Synthesis and evaluation of flexible pyrimethamine analogues as antifolates against drug-resistant malaria



Thesis by Matthew Maree

Submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Doctor of Philosophy

Supervisor: Professor Amanda Rousseau

Co-Supervisor: Dr Kennedy Ngwira

2025

Declaration

I declare that the work presented in this thesis was carried out exclusively by myself under the supervision of Professor Amanda Rousseau and Dr Kennedy Ngwira. Biological assessments were conducted by the BIOTEC research group in Thailand, the results of which are included in this thesis. This thesis is submitted for the Doctor of Philosophy Degree in Science at the University of the Witwatersrand, Johannesburg and has not been submitted before for any other degree or qualification.



Matthew Maree

March 2025

Abstract

In many parts of the world today, malaria still represents a major health crisis, with millions of cases and hundreds of thousands of deaths reported each year. The major malaria causing parasite is *Plasmodium falciparum*, which accounts for the majority of cases and deaths worldwide. This parasite has shown a remarkable ability to rapidly mutate and develop resistance against initially effective drug molecules. The continued threat of malaria today motivates the development of new molecules which can remain active despite point mutations that arise in the active site of the targeted enzymes.

Previous work done by A. Rousseau and colleagues showed that modifying existing class II antifolates, which target the bifunctional parasite enzyme dihydrofolate reductase – thymidylate synthase (DHFR-TS), by adding a flexible four-atom linker allowed for high levels of antiplasmodial activity to be maintained against drug-resistant forms of the parasite.

In the present work, molecular modelling techniques were used to optimize the structural elements of flexible pyrimethamine analogues for binding ability against mutant enzymes. The effects of various functional groups around a conserved structure were evaluated, and a library of compounds was selected for synthesis. Numerous flexible pyrimethamine analogues were successfully synthesized making use of partially established experimental procedures, and were submitted for various antiplasmodial assessments. The synthesized compounds showed excellent activity in single enzyme assays, inhibiting drug-resistant *Pf*DHFR enzymes at low nanomolar concentrations. Whole cell assays were also conducted, where a significant decrease in activity was observed, with the most active compounds inhibiting the parasite cells at low micromolar concentrations. These results suggested that while the compounds were effective binders of the target enzyme, they had some pharmacokinetic limitations which prevented them from effectively exhibiting their mode of action inside the cell.

A second-generation of analogues was then envisaged, taking inspiration from existing antifolate compounds which are known to have favourable pharmacokinetic properties. Methods for the synthesis of the devised second-generation analogues were developed as part of this work, and the binding ability of the compounds was validated with further molecular modelling studies. We are currently awaiting the results for the biological assessments of the second-generation analogues.

Acknowledgments

First and foremost, my deepest thanks go to my devoted supervisor, Professor Amanda Rousseau. Your diligent feedback and tireless support made the completion of this project possible. I feel incredibly fortunate to have worked with you over the past four years.

I am also deeply grateful to Dr Kennedy Ngwira for providing additional support and invaluable feedback during the writing process. Special thanks also to Dr Kamogelo Butsi, whose expertise was crucial in guiding the early stages of this project and my research.

Many aspects of this work would not have been possible without the biological assessments provided by the BIOTEC research group in Thailand. My heartfelt thanks to all its members for their hard work in delivering the biological results.

I also acknowledge the CSIR for hosting the CHPC, where all the molecular dynamics simulations for this project were conducted.

To the members of the organic lab at Wits, especially my lab mates in 318 — Bianca, Aneesa, and Khanya—thank you for your camaraderie and for keeping me sane during the trials and tribulations of organic synthesis.

To my close friends, Matthew Bracken and Alex Plakas, whom I met during honours in 2020 and continued this journey with, thank you for the countless lunches, coffee breaks, and opportunities to vent about failed reactions and endless columns.

I owe an immeasurable debt to my parents, who supported me throughout my academic journey and provided an environment that allowed me to focus on my work.

Finally, my deepest gratitude to Agi for your unwavering love and support during the most challenging times.

Table of contents

Contents

Declaration ii
Abstract iii
Acknowledgments iv
Table of contents v
Chapter 1 – Introduction 1
1.1 Malaria 1
1.1.1 Brief history and introduction 1
1.1.2 Summary of *Plasmodium falciparum* life cycle 2
1.1.3 History of malaria prevention and treatment 4
1.1.4 Social and economic burden of malaria 14
1.2 Details of folate metabolism and review of antifolate drugs 17
1.2.1 Overview of folate metabolism 17
1.2.2 Antifolates in the treatment of various diseases 21
1.2.3 Resistance to antifolates in the treatment of malaria. 28
1.2.4 Overcoming resistance to class II antifolates in the treatment of malaria 31
1.3 Molecular modelling in drug discovery 33
1.4 Project introduction and aims 34
1.4.1 Background of the current project 34
1.4.2 Aims 37
1.4.3 Proposed methodology 38
Chapter 2 – Molecular Modelling 42
2.1 Introduction 42
2.2 Aims and objectives 45
2.3 Validation of *Pf*DHFR-TS crystal structures and ligand binding mode 47
2.3.1 Assessment of protein crystal structures 47
2.3.2 Identification of ligand binding mode and self-docking experiments 51
2.3.3 Enrichment docking 57
2.4 Optimizing interactions in the hydrophobic pocket 59
2.5 Assessment of X-substituents around the terminal ring 69
2.6 Molecular dynamics simulations 80
2.7 *In silico* estimation of physiochemical parameters 85
2.8 Concluding remarks and recommendations for synthesis 86

2.9 Additional molecular modelling experiments 88
Chapter 3 – Results and discussion part A: Synthesis and biological evaluations of first-generation analogues 90
3.1 Synthesis of flexible pyrimethamine analogues 90
3.1.1 Alkylation of phenols with propargyl bromide 92
3.1.2 Esterification of aliphatic carboxylic acids 95
3.1.3 Formation of beta-ketonitriles 97
3.1.4 Microwave assisted formation of 2,4-diaminopyrimidines 100
3.1.5 Iodination of 2,4-diaminopyrimidines 105
3.1.6 Sonogashira coupling reaction 108
3.1.7 Hydrogenation of Sonogashira reaction products 115
3.2 Biological assessments of synthesized flexible pyrimethamine analogues 119
3.2.1 Single enzyme inhibition assays 123
3.2.2 Whole *P. falciparum* cell assays 133
3.2.3 Conclusions and path forwards 139
Chapter 4 – Results and discussion part B: Selection, synthesis, and modelling studies of second-generation analogues 142
4.1: Design of second-generation analogues 142
4.2: Proposed synthesis of second-generation analogues 146
4.2.1: Formation of propargyl substituted phenols and anilines 147
4.2.2 Sonogashira coupling reaction 162
4.2.3 Hydrogenation of Sonogashira reaction products 171
4.2.4 Removal of Boc protecting group 175
4.2.5 Hydrolysis of esters after Sonogashira and hydrogenation 176
4.2.6 Addition of glutamic acid 178
4.2.7 Deprotection of glutamic acid 196
4.3 Molecular modelling studies of second-generation analogues 198
4.3.1 Introduction 198
4.3.2 Induced fit docking 200
4.3.3 Molecular dynamics simulations 208
4.3.4 *In silico* estimation of physiochemical parameters 211
Chapter 5 – Conclusions and future work 213
Chapter 6 – Experimental methods and procedures 217
6.1 Synthesis of first-generation analogues 219
6.1.1 Alkylation of phenols with propargyl bromide 219
6.1.2 Synthesis of aliphatic Esters 226

6.1.3 Synthesis of Beta-ketonitriles (3-oxopropanenitriles) 227
6.1.4 Synthesis of 2,4–Diaminopyrimidines 229
6.1.5 Iodination of 2,4–Diaminopyrimidines 231
6.1.6 Sonogashira coupling reaction 233
6.1.7 Hydrogenation of alkynes 250
6.2 Synthesis of second-generation analogues 264
6.2.1 Esterification of carboxylic acid containing phenols and anilines 264
6.2.2 Alkylation of phenols for second-generation analogues 267
6.2.3 Boc protection of anilines 270
6.2.4 Alkylation of protected anilines 272
6.2.5 Deprotection of alkylated anilines 274
6.2.6 Methylation of alkylated anilines 276
6.2.7 Sonogashira coupling reaction for second-generation analogues 278
6.2.8 Hydrogenation of second-generation alkynes 284
6.2.9 Removal of Boc group after Sonogashira and reduction 290
6.2.10 Hydrolysis of esters after Sonogashira and reduction 291
6.2.11 Protection of glutamic acid 294
6.2.12 Addition of glutamic acid 295
6.2.13 Deprotection of glutamic acid 297
References 299
Appendices 305
Appendix A – Parameters for molecular modelling experiments 305
Appendix B – Predicted physiochemical properties 314
Appendix C – Full set of data obtained from biological assays 321
Appendix D – Predicted physiochemical properties for second-generation analogues 327
Appendix E – Characterization data for synthesized compounds 328

Chapter 1 – Introduction

1.1 Malaria

1.1.1 Brief history and introduction

Malaria is an infectious disease caused by single-celled parasites of the genus *Plasmodium*. Its effects have been known for thousands of years[1]. In ancient times it was thought to arise from the inhalation of contaminated air[2], however, in 1880, the French doctor Alphonse Laveran discovered that it was in fact caused by blood-borne microorganisms[1]. He was awarded the 1907 Nobel Prize in medicine for this discovery. By the end of the nineteenth century several other major breakthroughs had been made in the study of this parasite. Three distinct malaria causing species were uncovered by Camillo Golgi (*Plasmodium vivax*, *Plasmodium malariae*), and Marchiafava and Celli (*Plasmodium falciparum*) in 1886 and 1890, respectively[1]. The long-standing puzzle of malaria transmission was solved by Ronald Ross in 1897, where it was proven that the parasite was transmitted by *Anopheles* mosquitos[1]. For his work in studying the life cycle of the malaria parasite in the mosquito host, Ross received the 1902 Nobel Prize in medicine. The next fifty years after this ground-breaking discovery saw many insights gained into the details of the parasite life cycle. It was revealed that the parasite undergoes different stages of development in its two animal hosts – humans and mosquitos – and it was demonstrated that humans could be infected via transmission of sporozoites identified in infected mosquitos[3]. In 1948, sporozoites were identified for the first time in a human by studying liver samples from an infected individual. This aided the understanding of the full parasite life cycle in humans, from injection of sporozoites into the human host during a mosquito blood meal, subsequent proliferation in the liver and eventual infection of red blood cells by merozoites, to the uptake of differentiated gametocytes upon feeding by another mosquito[3]. In addition to *P. vivax*, *P. malariae*, and *P. falciparum*, two other *Plasmodium* species capable of causing malaria in humans were classified in the early twentieth century. *Plasmodium ovale* was named by John Stephens in 1922[4], and *Plasmodium knowlesi*, which

was first identified in simian hosts in 1932 by Campbell and Napier, was later shown to be able to infect humans via zoonosis[5]. Of the five *Plasmodium* species known to cause malaria in humans, *P. falciparum* accounts for the vast majority of cases and deaths worldwide[6].

1.1.2 Summary of *Plasmodium falciparum* life cycle

The life cycle of *P. falciparum* involves two different host organisms, namely humans and female *Anopheles* mosquitos, and is shown in brief in figure 1.1. Parasite sporozoite cells are first transferred to humans through the bites of infected mosquitos [Fig 1, A]. They migrate through the epithelium into the blood stream and within a few hours reach the liver where they infect hepatocyte cells[3,7] [Fig 1, B]. In the initial seven days after being infected, individuals exhibit no symptoms, and the disease is not transmissible. During this period, the parasite cells undergo rapid asexual proliferation, and become differentiated into multinucleate schizont cells. These rupture the infected liver cells, releasing numerous merozoites into the systemic circulation [Fig 1, C]. It is estimated that a single sporozoite can produce up to 10 000 further differentiated merozoite cells in the liver stage of the infection[7,8]. Once in the blood stream, merozoites quickly infect red blood cells and continue rapidly dividing [Fig 1, D]. The infected cells rupture, releasing the newly produced parasite cells into the blood stream. The cycle of infection continues and is known as the blood stage [Fig 1, E]. The red blood cell population in the body becomes severely diminished, which along with the effects of the human immune response, results in symptoms of the disease, namely fatigue, fever, and anaemia[3,9]. A small number of merozoites become differentiated into gametocytes during the blood stages and remain inside infected red blood cells [Fig 1, F]. When a mosquito takes a blood meal, infected cells are delivered to the mid gut of the insect [Fig 1, G]. The gametocytes differentiate into male and female gametes, which fuse to form a zygote [Fig 1, H]. This develops further into a motile ookinete, which exits the midgut, attaching itself to the outer gut membrane. There it develops into an oocyst, a multinucleate cell which produces thousands of new sporozoites over several days[8] [Fig 1, I]. Once released, the sporozoites travel to the salivary glands of the mosquito, where they can subsequently be injected along with the saliva into a new human host when the mosquito feeds[3,8] [Fig 1, J].

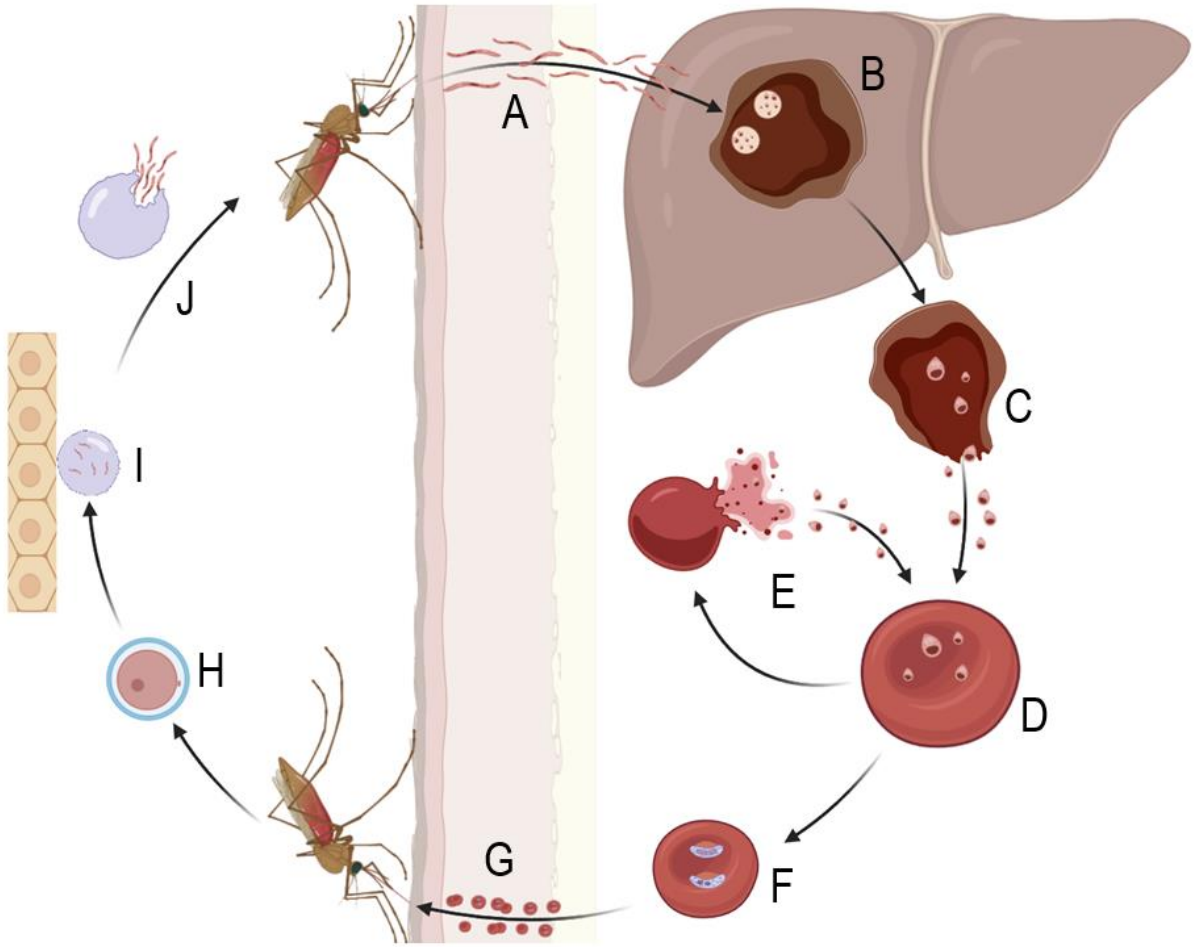


Figure 1.1: Summary of *P. falciparum* parasite life cycle in human and mosquito hosts. [Figure created with BioRender.com]

Symptoms of malaria typically show 10 - 15 days after the initial infection, and if not treated rapidly, can lead to severe illness and death. Children under five years of age are the most vulnerable, and account for two thirds of malaria deaths worldwide[6]. Malaria is treatable and preventable, however, over half a million people die and hundreds of millions more are infected each year.

1.1.3 History of malaria prevention and treatment

1.1.3.1 Small molecule drugs

There is a long history of treatments and preventative methods against malaria. Some progress was made in these efforts before the nature of the *Plasmodium* parasite and the disease itself was understood. The earliest known treatment for malaria comes from the bark of the *Cinchona* tree, which is indigenous to parts of South America[9,10]. The effect of the bark on fevers was long known by local populations, and in 1820 the active ingredient, quinine [Fig 1.2], was isolated by French chemists. An effective extraction procedure was developed, and it quickly became mass produced and used across the world[10]. It was the predominant malaria treatment until other, synthetic drugs were created in the 1920s[11]. It is still occasionally used in combination with other drugs for mild malaria cases today. Quinine combats malaria by preventing parasite cells from digesting haemoglobin during the blood stages of the infection, although the full mode of action remains unresolved[11].

Chloroquine, shown in figure 1.2, was the next major drug to be utilized. Developed in Germany in the 1930s as an alternative to quinine, this compound rapidly entered mainstream use and like quinine, also targets the blood stages of the *Plasmodial* infection[12]. Although an extremely effective antimalarial in the first few years of its use, parasites quickly developed resistance to chloroquine-based treatments[10]. The ability of the *Plasmodium* parasite to mutate so readily is one of the single most challenging factors in the eradication of malaria.

Time and time again, new promising antimalarial agents have been developed and have shown excellent activity in the initial years of use, before the parasite invariably develops key mutations to give it resistance against these new drugs. It has long been a challenge for chemists to create molecules which are able to target the parasite over many years and remain active despite any mutations in the active site of the targeted enzyme. By the second half of the twentieth century, chloroquine resistance was widespread[10], so new drug regimens were required.

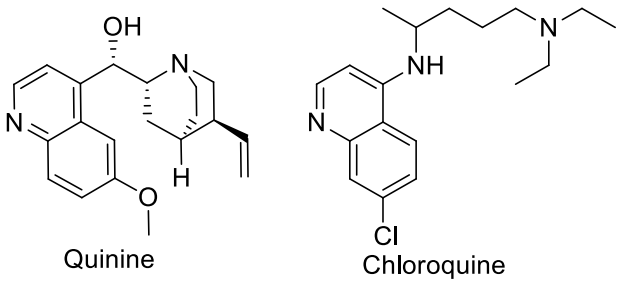


Figure 1.2: Structures of early antimalarials quinine and chloroquine

Two new compounds, proguanil and pyrimethamine, included in figure 1.3, were developed in 1945 and 1951, respectively, and quickly became widely used. Proguanil, which rearranges in the body into the active form of the drug cycloguanil, and pyrimethamine are both antifolates which target the bifunctional parasite enzyme dihydrofolate reductase - thymidylate synthase (DHFR-TS). This enzyme is a very attractive drug target as it is essential for folate metabolism in the parasite life cycle. In both the liver stage and the blood stages of the parasite infection in humans, the parasite undergoes large rounds of cellular proliferation, requiring a multitude of folate cofactors produced by the parasite's folate metabolism. By interrupting this process, the growth of the parasite can be halted[13]. Because they target both the liver stage and the blood stages, antifolate drugs can be used to prevent the symptoms of malaria from developing in the first place, eliminating the parasite in the liver before it enters the blood stream, or can be used to treat the symptoms of malaria after they manifest, halting the progress of the parasite in the blood stages[14].

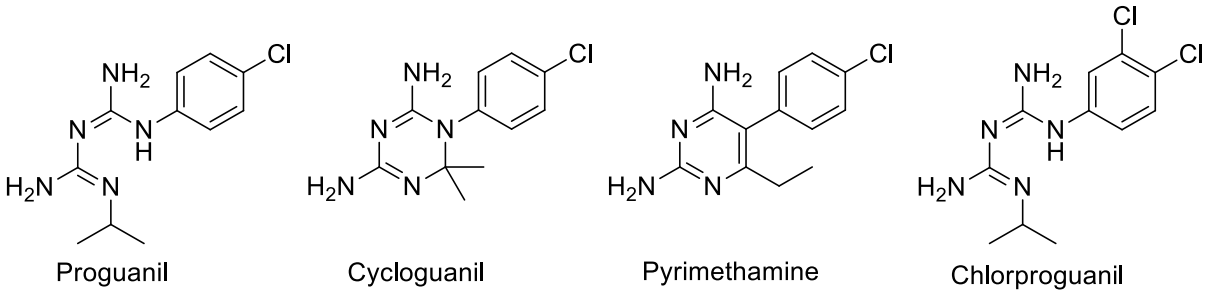


Figure 1.3: Structures of antimalarial class II antifolates cycloguanil and pyrimethamine, and the prodrugs proguanil and chlorproguanil

After the great initial success of these drugs, the issue of parasite resistance due to specific mutations however soon became apparent, and new combination treatment strategies were adopted[3]. Sulfones and sulfonamides, shown in figure 1.4, are another class of compounds which interrupt the folate metabolism in parasite cells. These molecules target the enzyme dihydropteroate synthase (DHPS) and are thus known as class I antifolates. Cycloguanil, pyrimethamine, and other compounds which target the DHFR-TS enzyme are called class II antifolates[15]. By combining both class I and class II antifolates into a single treatment regime, the advent of drug-resistance can be delayed, as mutations in two separate parasite enzymes are required for complete resistance[10]. Combination treatments became the norm, with pyrimethamine and sulfadoxine being sold under the brand name Fansidar, and chlorproguanil (an analogue of proguanil) [Fig 1.3] and dapsone marketed as Lapdap[16]. Despite early success, these treatments soon also lost their effectiveness against mutant parasite strains.

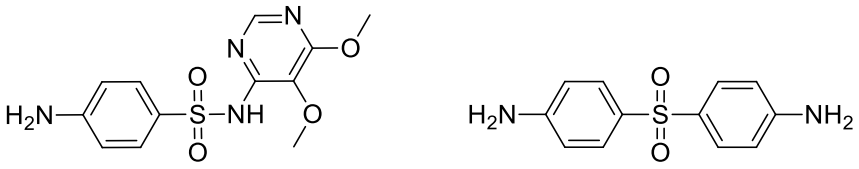
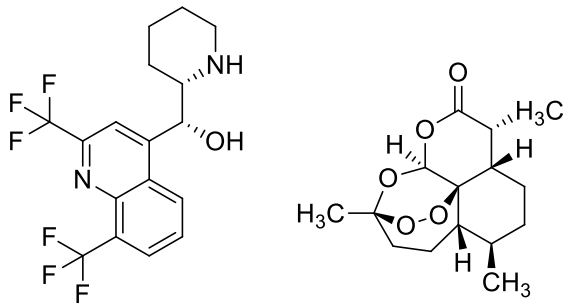


Figure 1.4: Common antimalarial class I antifolates sulfadoxine and dapsone

Mefloquine, included in figure 1.5, is a compound that was developed as a result of a large-scale screening campaign conducted in the 1960s and 1970s with the aim of finding new antimalarials[17]. Like quinine and chloroquine, this molecule contains a quinoline functional group, and similarly targets parasite development in the blood stages of the infection[12]. Mefloquine was first used in parts of Asia in the 1970s and become licenced for use in the United States in 1989. Today, parasite resistance to mefloquine is widespread[10].

In the modern era the most widely used antimalaria treatment is artemisinin combination therapy (ACT). The key component of this regimen is artemisinin, which is shown in figure 1.5. This natural product was first isolated from the *Artemisia annua* plant by Chinese researchers

in 1972[9]. The medicinal value of this plant has been long recognized by indigenous Chinese populations, and the isolation of the active ingredient led to much further research. Structurally, artemisinin differed drastically from other antimalarial compounds known at the time. It is comprised of a complex fused ring system, with the sesquiterpene lactone and endoperoxide rings proving essential for activity[10]. Artemisinin acts through a complicated multistep mechanism, involving the generation of free radicals from the peroxide bridge which disable key parasite proteins[18]. The exact steps involved in this process are still unclear. During the blood stages of a malaria infection, artemisinin inhibits both the asexual reproduction of parasite merozoites, and the formation of gametocytes. The disruption of gametogenesis prevents further transmission of the disease from infected individuals treated with artemisinin[19]. This compound has shown excellent activity against many mutant strains of *Plasmodium* and is used today in combination for the first line treatment of drug-resistant malaria infections[10]. In an effort to delay the advent of resistance against artemisinin, combination strategies are used. Despite this, partial artemisinin resistance was first observed in the early 2000s[20]. The development of resistance against artemisinin-based therapies poses a major threat to the ability to treat malaria, as no other drugs are poised to take their place as a first line defence against new infections[10].



Mefloquine Artemisinin

Figure 1.5: Structures of mefloquine and artemisinin antimalarials

Table 1.1 includes a brief summary of several small molecule drugs, highlighting details about their usage, mode of action, and information about ongoing clinical trials if relevant.

Table 1.1: Summary of small molecule antimalarial drugs

Compound	Target/Mechanism	Comments	Clinical Approval Status
Quinine	Inhibits haem detoxification	One of the oldest antimalarial agents; still used for severe and resistant cases	Approved – Historical use since the 17th century; now largely replaced by newer treatments
Chloroquine	Inhibits haem polymerization	Former first-line therapy; effectiveness severely limited by resistance	Approved – Introduced in the 1940s; limited current use due to widespread resistance
Cycloguanil	Class II antifolate; Inhibits dihydrofolate reductase (DHFR), blocking folate synthesis	Active metabolite of proguanil; sometimes included in combination therapies and as a model for research into new drugs	Research/Combination Use – Not used as a standalone agent
Pyrimethamine	Class II antifolate; Inhibits dihydrofolate reductase (DHFR), blocking folate synthesis	Commonly used in combination with sulfadoxine	Approved – Widely used in combination therapies despite resistance challenges
Sulfadoxine	Class I antifolate; Inhibits dihydropteroate synthase (DHPS), interfering with folate synthesis	Used in synergy with pyrimethamine; standalone efficiency reduced due to resistance	Approved – Employed in combination with noted resistance issues
Mefloquine		Effective against resistant strains; some negative	Approved – In clinical use since the 1980s

Compound	Target/Mechanism	Comments	Clinical Approval Status
	Inhibits haem detoxification, similar to quinine	side effects and reduced efficacy due to resistance	
Artemisinin	Generates reactive free radicals that damage parasite proteins upon activation	Cornerstone of modern malaria treatment; rapidly acting	Approved – Integral component of Artemisinin-Based Combination Therapies (ACTs)
Artesunate	Artemisinin derivative with improved solubility and bioavailability	Preferred for severe malaria due to rapid action	Approved – WHO-recommended for severe malaria treatment
Lumefantrine	Inhibits haem detoxification	Typically combined with other antimalarials for enhanced efficacy	Approved – Key component of ACTs since the early 2000s
WR99210	Class II antifolate; Potent inhibitor of dihydrofolate reductase (DHFR), blocking folate synthesis	Known for its high potency as an antifolate; primarily used as a research tool and currently undergoing early clinical trials	Under Early-Phase/Exploratory Clinical Evaluation

1.1.3.2 Vaccines

It has been a goal of scientists for many decades to produce a vaccine that is effective in preventing malaria, but this has proved extremely challenging. By the end of the 20th century, successful vaccines for many common diseases had already been developed and implemented across the globe. The earliest crude vaccine was the smallpox vaccine, created by Edward Jenner in 1796. Almost one hundred years later, Louis Pasteur introduced attenuated vaccines, and effective vaccines against rabies and cholera in humans were later developed. The 1920s saw the creation of vaccines for diphtheria and tetanus, and the first polio vaccine was

produced in the 1950s by Polish virologist Hilary Koprowski[21]. Many other key vaccines were also produced in the second half of the 20th century, including inoculations against *Haemophilus influenzae*, Hepatitis A, and Rotavirus, to name a few.

While many other diseases have all but been eradicated by the implementation of vaccines, the threat of malaria is still ever-present today. Only recently have vaccines for this disease become available, and they appear to show only moderate efficacy. This raises the question as to why a malaria vaccine has been out of reach for so long while huge progress has been made in the development of vaccines for other diseases. The answer to this is many-fold, with the single biggest factor being the greater biological complexity of the malaria parasite. While bacterial and viral infections are characterized by a limited number of antigens that can result in an immune response, parasite infections often involve several different stages of development in their life cycle, each with their own unique suite of associated proteins and other potential antigens[22]. This means that there is no single antigen which would be able to elicit a broad human immune response against all the stages of parasite development. In addition, the *Plasmodium* parasite has proven to be highly adept at evading the human immune system, and even in individuals who have been previously infected with malaria, a complete immunity is not developed[22]. The vast number of potential antigens displayed by the malaria parasite, and the inadequacies of the natural human immune system against this disease have made selecting antigens for a potential malaria vaccine very challenging[2,22].

So far two different broad types of vaccine have been investigated. Single target and combination vaccines. Single target vaccines contain a single parasite antigen against which an immune response is developed, while combination vaccines contain a cocktail of different antigens that can elicit a broader immune response. Combination vaccines are currently held to a higher standard than their single component counterparts as they may be able to generate an immune response against the parasite throughout the course of its life cycle. For a single target vaccine to be viable, it needs to be completely effective at eliminating all parasite cells before they can progress to the next stage of development. For example, a hypothetical single target vaccine which creates immunity against sporozoites will lose much of its effectiveness if even a single sporozoite is able to enter the liver and become differentiated into a hepatic schizont, which will go onto produce numerous merozoites and initiate the blood stages of the disease[22]. Similarly, for a single target vaccine which creates

immunity against the blood stages of the parasite, some merozoites may escape and become further differentiated into gametocytes, which would allow for the subsequent transmission of the disease to a new feeding mosquito. An additional challenge faced by single target vaccines is the high degree of protein polymorphism exhibited by the parasite, particularly in the blood stages of the infection[2]. This results in a best-case scenario of only partial immunity against this stage of the disease. For these reasons, it is unlikely that any single target vaccine will ever be highly effective. The development of this type of vaccine is still essential, however, as several different single target vaccines which are effective against specific stages of parasite development, or effective against specific polymorphic proteins, may be combined into a combination vaccine which is capable of generating robust immunity against each stage in the parasite's life cycle.

One last factor which cannot be overlooked in contributing to the difficulty in producing a malaria vaccine is the lack of financial incentives for large pharmaceutical companies. Malaria is a disease that largely affects impoverished tropical areas, and pharmaceutical companies are hesitant to spend millions of dollars in the research and development of a vaccine that will never generate large profits. The funding for malaria vaccine research has come mostly from humanitarian initiatives, such as the Bill and Melinda Gates Foundation. The rapid success in the development of the Covid 19 vaccine serves to demonstrate that if a collective, global effort is enacted for the development of a new vaccine, then progress can be made rapidly. Apart from the biological challenges in developing a malaria vaccine discussed earlier, this lack of broad international interest in vaccine development for a disease that largely infects impoverished communities is the largest factor influencing the slow progress in reaching this goal[22].

While several vaccine candidates are currently undergoing clinical trials, only two have been approved by the World Health Organization (WHO) for widespread use. The first of which is the RTS,S vaccine, developed by GlaxoSmithKline (GSK) pharmaceuticals[2,23]. Conceptualized in the 1970s, this vaccine underwent more than 40 years of development and testing before being recommended for the prevention of malaria in children by the WHO in 2021. Today, it is distributed mainly in sub-Saharan Africa, and as of the end of 2023, has been administered to more than two million children[6]. This vaccine makes use of the *P. falciparum* circumsporozoite protein (CSP), which is expressed on the surface of parasite cells during early

liver phase development[21]. This protein is a major antigen recognized by the immune system in the early stages of parasite infection and leads to a strong immune response. The vaccine is made up of a truncated portion of the CSP antigen, lacking the *N*-terminus, which is required for entry into host cells[21,23], fused with hepatitis B surface antigen (HBsAg), which acts as a carrier and immunogenic adjuvant[22]. In field trials, the RTS,S vaccine exhibited moderate efficacy of 30 – 50 %, and was shown to decrease the severity of malaria cases in children by lowering the parasite load which enters the blood stages[22]. Despite this, the immunity developed after administration of the RTS,S vaccine diminishes quickly, and 18 months after receiving the final dose no appreciable immunity can be observed[24].

Outlined in the malaria vaccine road map, the WHO has set a goal for the development of a vaccine which shows both lasting protective efficacy of at least 75 %, and transmission blocking ability by the year 2030[6]. While the RTS,S vaccine falls short of this target, it is still an important first step in realizing this goal and can nevertheless be used in the meantime as a tool for easing the burden of malaria until a superior vaccine candidate is developed.

The second malaria vaccine to date which has been approved for widespread use is the R21 vaccine. This vaccine was developed by scientists at the University of Oxford in collaboration with Novavax, a biotechnology company. The R21 vaccine has many similarities to the RTS,S vaccine. It also makes use of a truncated portion of the CSP to incite an immune response and is supported by hepatitis B derived adjuvants. The main difference between the two vaccines is the formulation of the virus-like-particles (VLPs) which are responsible for presenting the chosen antigen to the immune system. The RTS,S vaccine contains both fused CSP-HBsAg proteins as well as monomeric HBsAg, while the R21 vaccine is made up of only the recombinant CSP-HBsAg protein. This allows for the R21 VLPs to display a higher surface density of the CSP antigen, eliciting a stronger immune response[23]. The early-stage clinical trials of the R21 vaccine yielded very promising results. Efficacy of up to 75 % was observed in children 12 months after the initial three dose regimen, with a fourth dose extending the duration of immunity[6]. This vaccine was recommended by the WHO in October 2023, and is currently being distributed alongside the RTS,S vaccine in areas of high risk.

1.1.3.3 Vector control methods

Apart from medicines and vaccines, vector control remains one of the most promising strategies for combating malaria. Hypothetically, if every single *Anopheles* mosquito was eradicated today, then malaria would soon disappear completely without any medical interventions. This, however, poses an even greater challenge than the development of new antimalarial drugs or a universal vaccine. Despite this, vector control methods have been widely implemented, and continue to play a crucial role in limiting the number of cases.

The two main methods of mosquito vector control used today are the indoor residual spraying (IRS) of insecticides, and the use of insecticide treated nets (ITNs). The former involves coating the interior surfaces of dwellings with potent insecticides which are able to eliminate any mosquitos that come into contact with it. The latter combines the use of nets, which serve to create a physical barrier between humans and mosquitos, blocking any potential transmission, with insecticides that enhance this effect[25].

The most widely used insecticide for IRS to date is dichlorodiphenyltrichloroethane (DDT), the structure of which is included in figure 1.6. This compound was first synthesized in the late 1800s, but it was not until 1939 that its remarkable insecticidal properties were identified[10]. In low concentrations, this compound is able to disrupt the nervous system of arthropods by interfering with the normal functioning of sodium ion channels, leading rapidly to death[26]. The protection offered by DDT IRS lasts up to six months[27], and it has been a very effective method for reducing the incidence of malaria in highly affected areas[28,29]. This compound is however not without its drawbacks. After more than 60 years of widespread use, mosquito populations have developed resistance to DDT[27,29]. In addition, there are numerous concerns about its toxicity, with some evidence suggesting a link between DDT and the development of certain types of cancer in humans[30]. Additionally, DDT is classified as a persistent organic pollutant for its ability to remain intact in ecosystems long after its introduction[29]. For these reasons, the widespread use of DDT was banned in 1972 by the US environmental protection agency. Despite the ban, several countries received exemptions and continued the use of DDT as a vector control agent in the face of malaria outbreaks. In 2006 the WHO declared DDT acceptable for use in areas of high malaria risk and it is still often used today[25].

For use in ITNs, pyrethrums, shown in figure 1.6, are the insecticide of choice[25]. This class of compound was first extracted from the flowers of the *Tanacetum cinerariifolium* plant, the insecticidal properties of which have been known for thousands of years[4]. Numerous synthetic analogues have also been created, which are named pyrethroids [Fig 1.6]. These compounds exhibit a similar mode of action to DDT, involving the disruption of sodium ion channels inside the cell membranes of arthropod cells[2]. Although widely successful for a number of years, many cases of mosquito resistance to pyrethroid based insecticides have been reported. Due to this, various combination approaches are currently recommended for ITNs in areas of pyrethroid resistance[6].

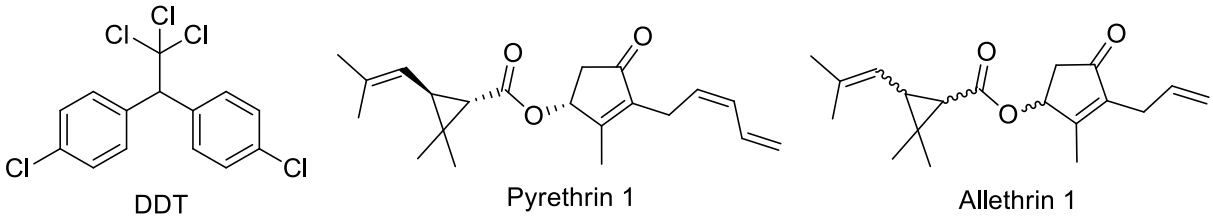


Figure 1.6: Structures of the insecticides DDT, the natural pyrethrin 1 and synthetic allethrin 1

1.1.4 Social and economic burden of malaria

Although malaria has been around for many thousands of years, this disease still represents a massive health crisis today. In 2022, the estimated number of malaria cases reached 249 million, with more than 600 000 deaths reported[6]. While malaria is endemic to more than 80 countries across the world, the greatest burden of this disease is felt in Africa, which accounted for more than 90 % of the total cases and deaths in 2022[6].

Due to continued, directed efforts at combating malaria over the last 20 years, substantial progress has been made. In the year 2000, an estimated number of 262 million infections, and 864 000 deaths were reported[6]. These figures have dropped each year reaching a minimum of 229 million cases and 576 000 deaths in 2019. However, due to the impacts of the global Covid-19 pandemic which started in early 2020, the number of cases and deaths saw a slight

increase. Humanity is still recovering from these setbacks, and the number of deaths and cases reported in 2022 are marginally above the pre-pandemic figures[6].

The global malaria mortality rate has also fallen consistently since 2000. This is defined by the number of deaths reported per 100 000 individuals at risk. Improvements in this statistic are made both by limiting the number of at-risk individuals, mainly through vector control methods, and also improvements in the treatment of malaria cases through medical interventions, lowering the proportion of deaths after infection. In 2000 the mortality rate was 29 deaths per 100 000 individuals at risk. This reached its lowest value of 14 in 2019. An increase in mortality rate was observed in 2021, reaching 15.2. Since then however some improvement has been made and the mortality rate reported in 2022 was 14.3, which is close to the pre-pandemic levels[6].

One key reason for the success achieved in the last 20 years is the increase in the coverage of vector control methods. These are relatively cheap, and highly effective at reducing transmission rates. In 2022, 282 million ITNs were distributed in malaria endemic countries. More than half of these were pyrethroid-PBO nets, which are treated with a combination of pyrethroid and piperonyl butoxide (PBO) insecticides. In many sub-Saharan African nations ITNs are the main form of vector control, and it is estimated that in 2022 around 70 % of households in sub-Saharan Africa had at least one ITN, a substantial increase from 5 % reported in 2000. This increase has however stagnated since 2015[6]. In 2022, 47 countries implemented IRS campaigns to prevent the spread of malaria, although the coverage of this is even less than that of ITNs, with an estimated 1.8 % of at-risk populations being protected by IRS[6].

Children under the age of five years and pregnant women are the most vulnerable to malaria, and so special attention needs to be paid to these demographics. Young children accounted for 76 % of the total number of deaths from malaria in 2022. Women who have been infected with malaria during pregnancy are more likely to die during childbirth, and children born under these conditions have a higher risk of child mortality and experiencing various developmental issues. In some heavily afflicted areas, up to 40 % of pregnant women had malaria infections in 2022. This is very alarming given the potential effects of malaria during pregnancy and makes combating malaria during pregnancy an issue of vital importance. It is

therefore recommended that pregnant women who may be exposed to malaria regularly take sulfadoxine-pyrimethamine formulations as a preventative measure[6].

In 2016, the WHO and partners introduced a global technical strategy (GTS) for the eventual eradication of malaria. This proposal outlined specific goals and provided recommendations for reducing the malaria mortality rate and lowering the number of malaria endemic countries by the year 2030. Compared to the baseline figures obtained in 2015, a decrease in the mortality rate of 40 % by 2020, 75 % by 2025, and 90 % by 2030 were targeted. Furthermore, it is hoped that by 2030, at least 35 new countries will be declared malaria free. As of the year 2022 these targets are not on track to be met.

As efforts to combat malaria continue to scale up, it is evident that an increase in funding is required. In 2022, an estimated $ 4.1 billion was invested in the fight against malaria. This represents an increase from previous years, although still far short of the $ 7.8 billion target defined by the GTS for 2022[6]. Most of this money is required for the manufacturing and distribution of pharmaceuticals and ITNs, as well as the rollout of the recently developed vaccines. Aside from this, it is also vital to allocate funds for the research and development of new strategies for fighting malaria. ACTs continue to be the pharmaceutical front line defence against malaria, but their long-term effectiveness isn't guaranteed. It is crucial that the drug development pipeline continues to produce new molecules to take the place of ACTs, which have been extensively used over the last two decades.

The WHO estimates that due to human intervention against malaria, namely the widespread distribution of antimalarial drugs and the implementation of vector control methods, a total of 2.1 billion cases and 11.7 million deaths have been prevented between 2000 and 2022[6]. This is a remarkable achievement, and it is important to consider this when looking at the vast number of cases and deaths that are still reported each year. Even though the threat of malaria is still ever-present in many parts of the world today, the situation could be significantly worse, and there is still continued hope that the complete eradication of this disease is a feasible undertaking.

1.2 Details of folate metabolism and review of antifolate drugs

1.2.1 Overview of folate metabolism

Folate metabolism is an ancient metabolic pathway that exists in some form in all organisms. Folate derivatives acquired from this process are essential cellular cofactors required for nucleic acid biosynthesis and DNA replication. Rapidly dividing cells such as cancer, bacteria, and parasite cells are heavily dependent on a large supply of folate cofactors for their continued proliferation[16]. The use of compounds which can inhibit the folate metabolism in these cells are highly effective at halting their progress, and as such antifolate drugs have been widely used in the treatment of various diseases.

During the *P. falciparum* life cycle in humans, there are several stages which involve large rounds of asexual cellular proliferation, namely the initial formation of merozoites during the liver stage and the continued cycle of infection – proliferation – lysis which repeats many times during the blood stages. The use of antifolates in the treatment of malaria disrupts these processes, and can serve as an effective prophylaxis, halting the development in the liver stage, as well as a symptomatic treatment, combating the continued proliferation of parasite cells during the blood stages.

Cells obtain folates in one of two ways, either through *de novo* synthesis or via a folate salvage pathway[31]. The former involves numerous steps and enzymes, and results in the formation of the key cofactor dihydrofolate (DHF), which is subsequently converted to tetrahydrofolate (THF) by the enzyme dihydrofolate reductase (DHFR). Animals lack serval key enzymes involved in *de novo* synthesis and must therefore obtain their folate cofactors through a different pathway. Typically, the required folate precursors are obtained through diet in the form of folic acid. This is converted by the animal's own DHFR enzymes to DHF, which is then further metabolised to THF by the same enzyme. As humans are unable to synthesize their own folates from scratch, folic acid, also called vitamin B9, is an essential nutrient[32]. *Plasmodium* cells have access to both metabolic pathways for producing their folate cofactors and can also scavenge folates from the infected host cells.

Folate derivatives are structurally conserved and are made up of three key components. The first is a pterin ring, which is attached to *para*-aminobenzoic acid (PABA), and finally a glutamic acid residue. These elements are highlighted on the structure of dihydrofolate in figure 1.7.

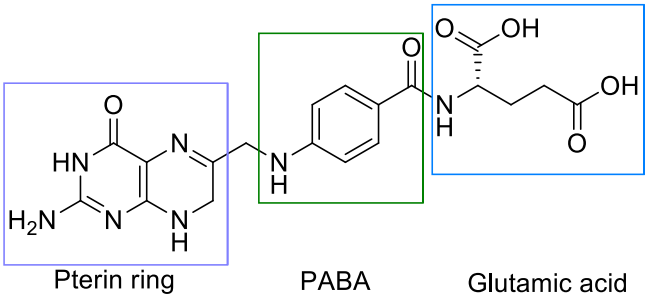
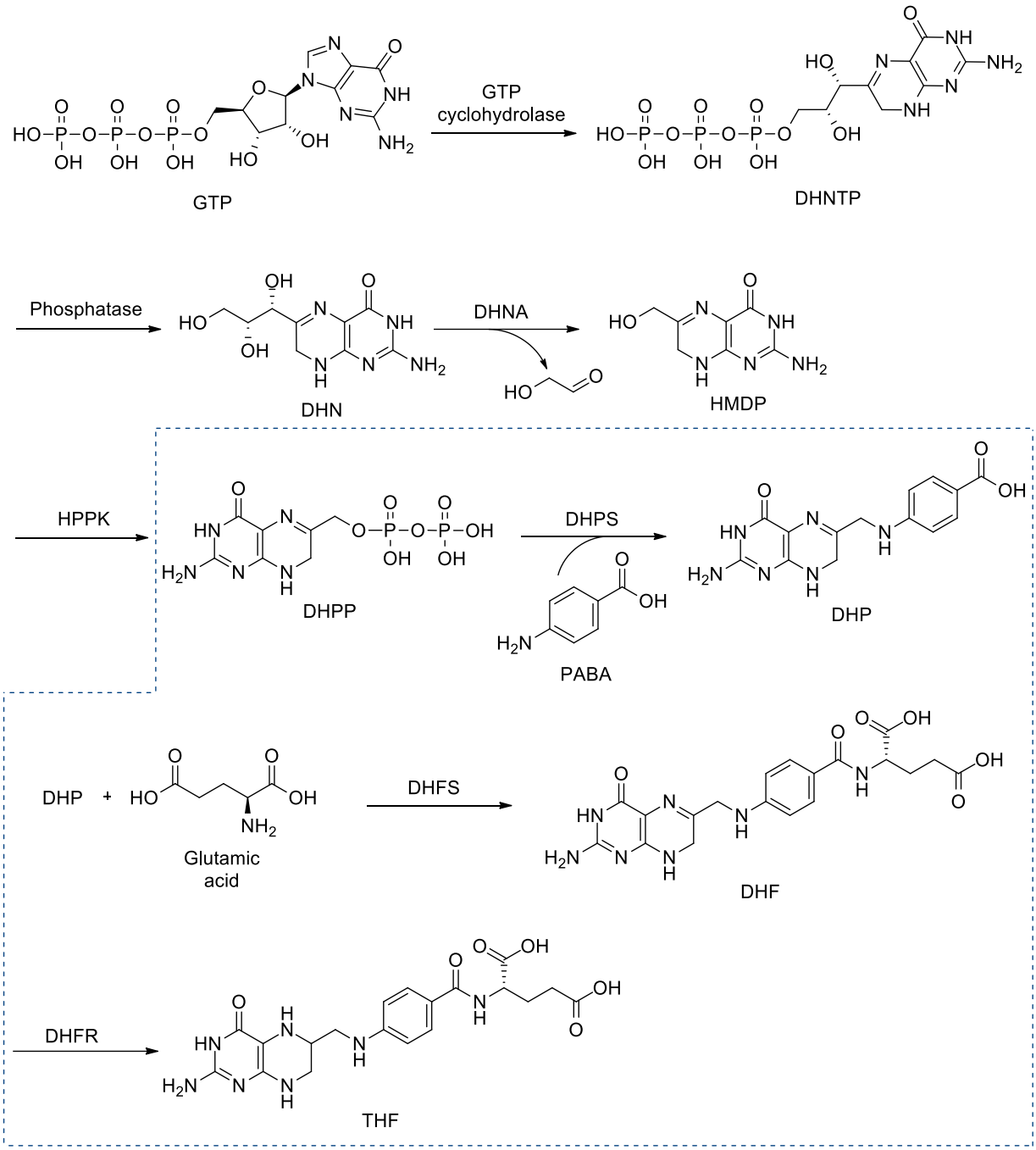


Figure 1.7: Key structural elements of DHF

An overview of *de novo* folate metabolism in *Plasmodium* cells is shown in scheme 1.1. The first step entails the rearrangement of gaunosine-5'-triphosphate (GTP) to 7,8-dihydroneopterin triphosphate (DHNTP) and is achieved by the enzyme GTP cyclohydrolase. This forms the core pterin ring system present in all folate derivatives. While the GTP cyclohydrolase enzyme is also found in animal cells, it does not play the same role in folate synthesis as several of the enzymes required for the later steps are absent[33]. The triphosphate tail of the DHNTP is then cleaved by phosphatase enzymes, resulting in the formation of 7,8-dihydroneopterin (DHN), which is subsequently converted to hydroxymethyldihydropterin (HMDP) by the enzyme dihydroneopterin aldolase (DHNA), with the liberation of glycolaldehyde. A bifunctional parasitic enzyme in *Plasmodium* cells catalyses the next two transformations. A diphosphate group is first added to the terminal hydroxy of HMDP by the hydroxymethyldihydropterin pyrophosphokinase (HPPK) domain. The resulting dihydropterin pyrophosphate (DHPP) then undergoes a condensation reaction with *p*-aminobenzoic acid (PABA), catalysed by the dihydropteroate synthase (DHPS) domain of the bifunctional enzyme, to form dihydropteroate (DHP)[33]. Finally, DHP is modified by the addition of a glutamic acid residue, facilitated by the enzyme dihydrofolate synthase (DHFS). The resulting product is dihydrofolate (DHF), which is reduced to tetrahydrofolate (THF) by dihydrofolate reductase (DHFR)[14].



Scheme 1.1: Overview of *de novo* folate metabolism in *Plasmodium* cells, steps also involved in the folate salvage route indicated by dashed blue outline

The folate salvage pathway is a convergent route that bypasses many of the early steps of THF synthesis found in *de novo* folate metabolism. While *de novo* synthesis begins with GTP and proceeds through several intermediates, the salvage pathway relies on later folate precursors that are typically scavenged from the host's bloodstream. The points at which the salvage pathway converge with the *de novo* synthesis route are included in scheme 1.1 above in the

dashed blue area. Class I antifolates inhibit the enzyme DHPS, effectively blocking the *de novo* folate synthesis pathway. However, this can be bypassed via the salvage pathway if later folate precursors, such as DHP or DHF, are scavenged from the host. In contrast, Class II antifolates target DHFR, inhibiting the conversion of DHF to THF, effectively blocking both the *de novo* and salvage routes of folate metabolism.

Inside the cell, the later folate derivatives usually exist in polyglutamated forms, containing up to ten additional glutamic acid residues. While monoglutamate forms are readily transported, the polyglutamate forms are not, and build up to high concentrations within the cell. The addition and removal of extra glutamate residues serve to regulate the concentration of folate cofactors, ensuring that cells requiring large amounts of folates have sufficient access to them in the polyglutamated form. In addition, the polyglutamated forms have enhanced abilities as enzyme substrates[34]. Two main enzymes facilitate this process, extra glutamic acid residues are sequentially added to the *C*-terminus of the glutamic acid present on the folate derivatives by the enzyme folylpolyglutamate synthase (FPGS). A second enzyme, gamma-glutamyl hydrolase, is able to cleave the extra glutamate residues resulting in the monoglutamate form of the folate cofactors for subsequent cellular transport[34]. These compounds are highly polar and require active transport mechanisms to cross cell membranes. In human cells there are three main pathways for facilitating the entry of the mono-glutamated folate cofactors into cells. These include the transmembrane proteins reduced folate carrier (RFC), the proton-coupled folate transporter (PCFT), and various folate receptors[34].

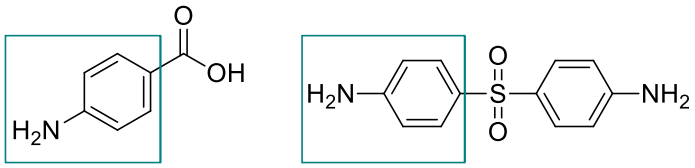
THF is the final product of both the *de novo* biosynthesis pathway, and the folate salvage route. This molecule is able to facilitate numerous key transformations required for cell survival. In particular it acts as an essential cofactor in one-carbon-transfer reactions[33]. Single carbon units are added to THF in various forms by enzymes, including formyl, methylene, and methyl C1 units[33]. These are then transferred to other substrates in several key biosynthetic pathways, including purine and thymidylate synthesis, which are required for DNA and RNA production, as well as the synthesis of various amino acids needed for protein formation. Once these transformations are complete, the C1 THF derivatives become oxidized back to DHF, which can subsequently be reduced to THF, continuing the metabolic cycle.

1.2.2 Antifolates in the treatment of various diseases

1.2.2.1 Treatment of malaria with antifolates

Many different compounds have been investigated as antifolates for their role in disrupting folate metabolism. While there are dozens of enzymes involved in this process, the most success has come from targeting two in particular. These are dihydropteroate synthase (DHPS) and dihydrofolate reductase (DHFR). Drugs which inhibit the former enzyme are known as class I antifolates, while drugs inhibiting the latter enzyme are class II antifolates. Antifolate drugs have shown to be very effective against *P. falciparum* strains but show significantly lower activity against other malaria causing *Plasmodium* species[13].

In the treatment of *P. falciparum* infections, both class I and class II antifolates have been widely used, commonly in combination to enhance the effects. The class I antifolates largely consist of sulfadrugs, such as dapsone and sulfadoxine [Fig 1.4]. These resemble one of the natural substrates for the DHPS enzyme, PABA. The structural similarities are presented in figure 1.8. Dapsone is the most potent known inhibitor of *P. falciparum* DHPS. It was originally synthesized in an attempt to create new dye molecules in the early 1900s and was shown to have a significant antimicrobial effect in the 1930s. Further testing also demonstrated the effects of this compound as an antimalarial, and today it is used to treat a number of diseases, including bacterial and fungal infections, as well as malaria[14]. While this drug is somewhat ineffective when administered in isolation, by combining it with class II antifolates a significant increase in the antimalarial effect is observed[14]. While the target enzyme of class I antifolates, DHPS, may be blocked by class I antifolates, the parasite can still obtain folate cofactors through the salvage pathway. In order for a complete reduction in folates available for parasite cells, the second pathway also needs to be targeted. This can be achieved by combining both class I and class II antifolates into a single treatment. While combination treatments have the advantage of delaying the development of parasite resistance, resistance still invariably emerges after prolonged use of the same active compounds.



PABA Dapsone

Figure 1.8: Conserved structural elements of PABA and the class I antifolate dapsone

Widely used class II antifolates in the treatment of malaria include compounds such as proguanil, chlorproguanil, and pyrimethamine [Fig 1.3]. These inhibit the parasite enzyme dihydrofolate reductase by mimicking the natural enzyme substrate DHF. Typically, class II antifolates are competitive reversible inhibitors, and work by outcompeting DHF at the active site of the DHFR enzyme[35]. Cycloguanil and its analogues contain a 2,4-diamino-dihydrotriazine core, while pyrimethamine and its analogues contain a 2,4-diaminopyrimidine core. These functional groups resemble the pterin moiety found in DHF, as shown in figure 1.9, and are essential for binding to the active site of the target enzyme. The binding mode of the class II antifolates to the active site of *P. falciparum* DHFR is highly conserved, and a minimum required pharmacophore for these compounds has been identified. This is shown in figure 1.10, and includes hydrogen bonds to residues Ile14, Ile164, Asp54, and an additional charge mediated interaction with Asp54. Some of these binding elements are not present for the natural enzyme substrate DHF, and so compounds which contain all of them can outcompete DHF in the active site of the target DHFR enzyme.

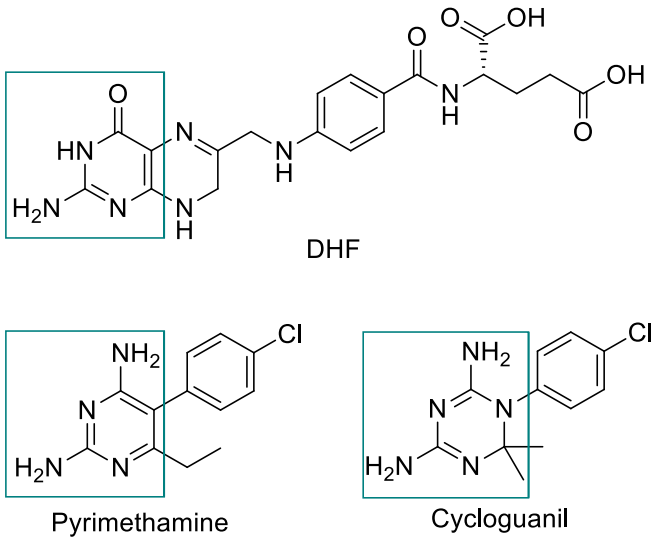


Figure 1.9: Conserved structural elements of DHF and class II antifolates

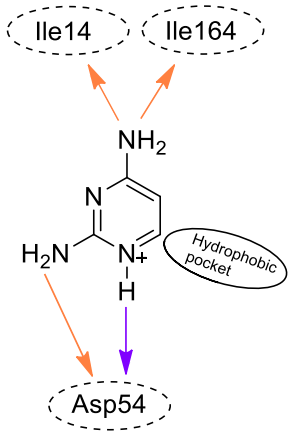
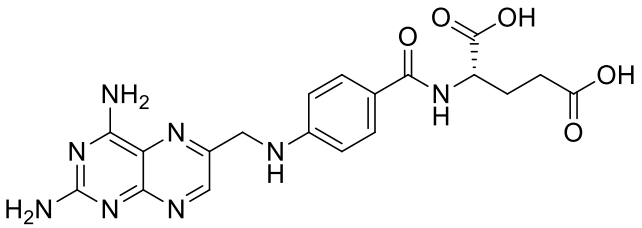


Figure 1.10: Binding mode of class II antifolates in the *Pf*DHFR enzyme

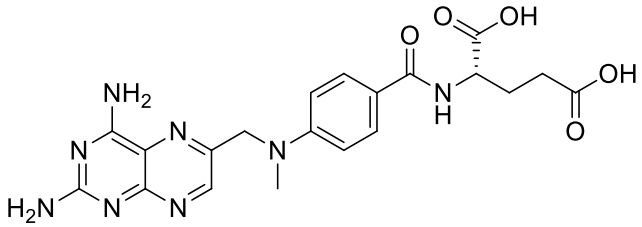
1.2.2.2 Treatment of cancer with antifolates

Cancerous cells are marked by rapid, uncontrolled proliferation. In this process, cells require large amounts of folate cofactors, and thus antifolate drugs which limit their supply are effective agents in the treatment of cancer. One of the first antifolate compounds to be used in this manner is aminopterin[35], shown in figure 1.11. This molecule closely resembles the natural substrate of the DHFR enzyme, DHF. Both contain a pteridine core moiety, but aminopterin has an amino substituent in the 4-position, while DHF has a carbonyl. This

enables additional interactions in the active site, with the primary amine of aminopterin able to both accept and donate hydrogen bonds to nearby residues, allowing it to outcompete the natural enzyme substrate. Aminopterin was first used to treat certain types of leukaemia in children. It showed great promise but was phased out because it had an undesirable toxicity profile. Methotrexate (MTX), included in figure 1.11, was its replacement for widespread use. Although this compound still displayed some toxicity towards healthy cells, it was a marked improvement over aminopterin.



Aminopterin

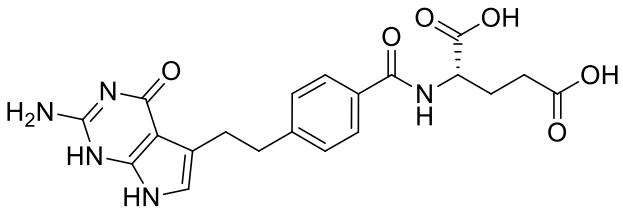


MTX

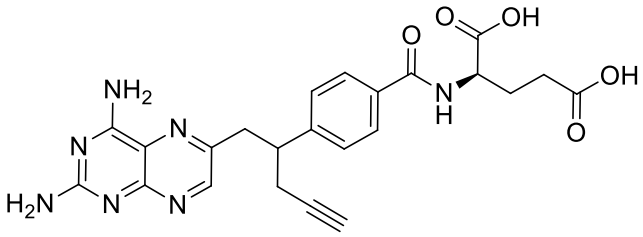
Figure 1.11: Structure of anticancer antifolates aminopterin and methotrexate (MTX)

The structures of aminopterin and methotrexate are very similar. The nitrogen present in the linker of aminopterin is a secondary amine, while in methotrexate it is methylated, becoming a tertiary amine. This small change improved the toxicity profile of the drug and allowed for a similar level of anti-cancer activity. MTX became the antifolate of choice in the treatment of certain cancers in the 1950s and is still in widespread use today[35]. While successful in practice, the detailed mechanism of how MTX is able to inhibit tumour cells was unknown for many years. This lack of a resolved mode of action led to challenges in developing other antifolate compounds in subsequent years. New anticancer antifolates only become available in the early 21st century. Pemetrexed and pralatrexate, shown in figure 1.12, are two such

compounds. Both are structurally related to MTX, with pemetrexed containing a modified guanine core and lacking the nitrogen atom in the linker. Pralatrexate also lacks the nitrogen atom in the linker, instead containing a carbon with a propargyl substituent. Both of these compounds are also effective anticancer antifolates and remain in use today.



Pemetrexed



Pralatrexate

Figure 1.12: Structures of recently developed anticancer antifolates pemetrexed and pralatrexate

One important aspect of treating cancer with antifolate drugs is the cellular transport of these compounds. Anticancer antifolates can be broken down into two general categories. The first are classical antifolates; these are defined as containing a glutamic acid residue[36]. Like the natural folates, these are able to become polyglutamated, which has significant effects on the distribution of these drugs throughout the body. Upon administration, they exist in the monoglutamated form. As they are extremely polar, and exist as anions under physiological conditions, various transmembrane transport proteins are required for entry into cells. Once inside the cell, they undergo polyglutamation, preventing them from exiting and resulting in bioaccumulation. This is a favourable process as it ensures that the drugs become concentrated inside tumour cells, where they can exhibit their desired mode of action.

Resistance of cancer cells to this type of antifolate is largely due to changes in the structure of the various transport proteins required for entry into the cancerous cells[34]. If mutations in these enzymes prevent the classical antifolates from entering the targeted cells, then they will have no means of exerting their desired action. To overcome this, another class of anticancer antifolates have been developed. These are named non-classical, or lipophilic antifolates[34]. Lacking the glutamic acid portion of their classical counterparts, they are not dependent on active cellular transport, and their increased lipophilicity allows them to enter cells though passive diffusion. One such compound is trimetrexate, shown in figure 1.13.

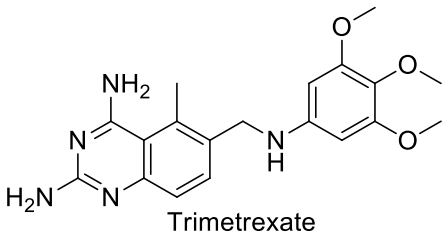


Figure 1.13: Structure of lipophilic antifolate trimetrexate

Trimetrexate is a structural analogue of MTX, containing a 5-methyl quinazoline ring instead of the pterin ring found in MTX. In addition, it lacks the glutamic acid substituent on its terminal ring, and instead contains a trimethoxy functionality. This compound has shown good efficacy against MTX resistant cancer strains as it can operate independently of mutated transport proteins which confer resistance to the classical antifolates. Despite this, some level of resistance can still develop against the non-classical antifolates, although in this case the resistance is due to mutations in the active site of the target enzyme itself[35].

1.2.2.3 Treatment of bacterial infections with antifolates

Folate metabolism is an essential pathway for all bacterial species. Compounds inhibiting this process have been widely used as antibiotics. The earliest known being prontosil [Fig 1.14], which was first synthesized as a dye molecule and later shown to have potent anti-microbial activity[37]. It is a prodrug for the active compound sulfanilamide [Fig 1.14]. Containing similar

functional groups to PABA, this molecule binds to the bacterial enzyme DHPS, blocking the *de novo* folate synthesis pathway in bacteria cells. Many other antibiotic compounds exist which target this enzyme, largely consisting of similar sulfur containing PABA mimics[38].

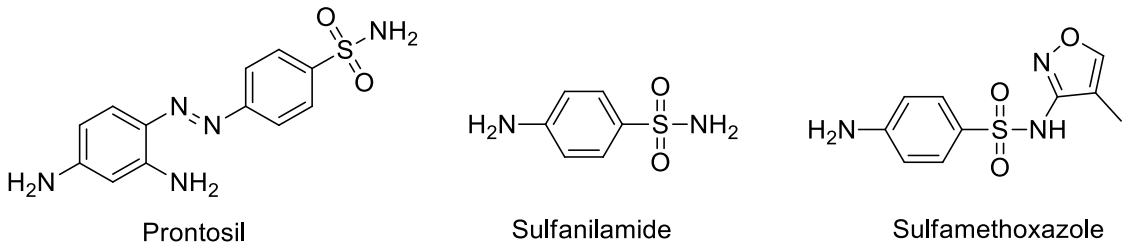


Figure 1.14: Structures of antibacterial antifolates sulfanilamide and sulfamethoxazole, and the prodrug of sulfanilamide, prontosil

In addition to DHPS, DHFR is the other key enzyme target for antibiotics inhibiting folate metabolism. Trimethoprim, shown in figure 1.15, was introduced in the late 1960s and has been used in combination with another DHPS inhibitor, sulfamethoxazole [Fig 1.14] very successfully for a number of years[37]. Unfortunately, bacterial resistance has developed against antifolate medications. Point mutations in the active site of the target enzymes render the currently used drugs unable to bind effectively[37].

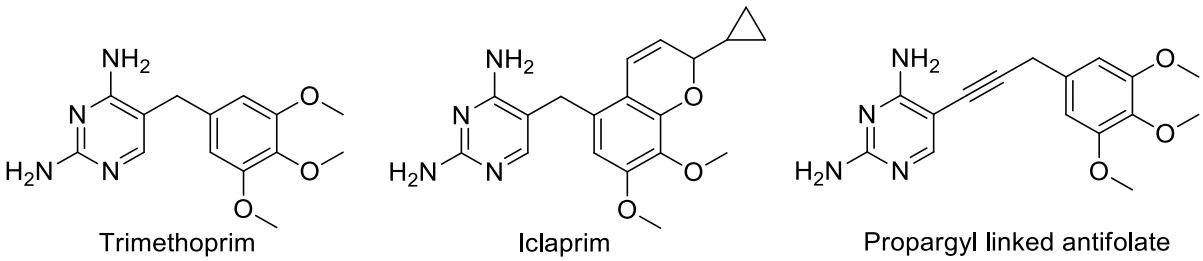


Figure 1.15: Structures of antibacterial antifolates targeting DHFR, trimethoprim, iclaprim, and a propargyl linked antifolate

Recent efforts are directed at developing new compounds which can overcome this resistance. Iclaprim, included in figure 1.15, is a potential drug candidate which has shown good levels of activity against trimethoprim resistant strains and progressed through various clinical trials

until some negative side effects were identified[37]. Other compounds in development include the propargyl-linked antifolates [Fig 1.15], which contain an unsaturated three-atom linker between its ring systems[39]. Various analogues of this nature have been produced and show some promise for development into new generation antibiotics.

1.2.3 Resistance to antifolates in the treatment of malaria.

Widely used in the treatment of malaria for many decades, both class I and class II antifolates have lost much of their activity due to the development of resistance in parasite populations. Point mutations in the active site of the target enzymes prevent drug molecules from binding, while maintaining high affinity for the enzymes' natural substrates.

A single A437G point mutation in the active site of the parasite DHPS enzyme leads to significant resistance against sulfadoxine and other related class I anti-folates[33]. This single amino acid substitution, alone or combined with other point mutations including A581G, S436A, and K540E, results in a many-fold loss of activity against sulfadoxine[13]. Many *P. falciparum* strains have been identified in various parts of the world containing one or more of the above listed drug-resistant mutations. It is thought that many of these mutations initially developed in parts of South-East Asia and South America but have become widespread today due to human migration patterns[33].

Although initially highly successful, class II antifolates have also been rendered ineffective due to mutations developing in the target enzyme DHFR. In *P. falciparum*, a single bifunctional DHFR-TS enzyme is encoded for, containing a DHFR domain and a thymidylate synthase (TS) domain. While the DHFR domain converts DHF to THF in the presence of NADPH, the TS domain is responsible for thymidine synthesis and makes use of a THF derivative as a cofactor in a reductive methylation reaction[31]. Mutations in the active site of the DHFR domain arise as a result of drug pressure after years of continued use of the same active compounds. Resistance to pyrimethamine was first observed in the 1950s, shortly after this compound was deployed for widespread use[13]. Initial strategies to combat this newfound resistance included the use of combination therapies, but soon after their implementation they too were found to be ineffective. *P. falciparum* resistance against pyrimethamine arises primarily through a

S108N point mutation in the active site of the parasite DHFR enzyme. This single amino acid substitution results in a 25-fold decrease in pyrimethamine activity when compared to the wild-type enzyme[13]. Subsequent mutations further decrease the activity of pyrimethamine, with quadruple mutant strains containing the above mutation in combination with C59R, N51I, and I164L substitutions causing a loss in activity of three orders of magnitude[14]. These mutations cause the shape of the active site of the parasite DHFR enzyme to change, adopting a new configuration for which pyrimethamine has a very low binding affinity[40]. While these changes also confer high levels of resistance to cycloguanil, distinct mutations are observed as a result of cycloguanil drug pressure. These include A16V and S108T substitutions, which are also able to drastically lower the effectiveness of pyrimethamine[14]. Cycloguanil and pyrimethamine are somewhat rigid molecules, both containing inflexible bi-aryl axes between their two rings. Because of this, they have a high energy barrier for rotation around this bond, meaning they are effectively locked into a single conformation. As such, they are unable to be accommodated in the active site of DHFR enzymes containing mutant residues which sterically hinder the binding pocket or alter its shape, preventing optimal alignment and interactions. Development of new class II antifolates is focused mainly on increasing the flexibility of active compounds, with the aim of mimicking the ability of the natural enzyme substrate, DHF, to fit optimally within the enzyme's active site, ensuring effective interactions and inhibition, even when mutations alter the shape or steric properties of the binding pocket.

Table 1.2 summarizes the key single amino acid substitutions discussed above and their roles in contributing to drug-resistance. It should be noted, however, that these mutations rarely occur in isolation. Instead, resistant *P. falciparum* strains contain multiple DHFR mutations that collectively enhance resistance. For instance, the quadruple mutant S108N/N51I/C59R/I164L is strongly associated with pyrimethamine resistance, while the A16V/S108T combination confers high levels of resistance to cycloguanil.

Table 1.2: Single amino acid mutations in drug-resistant *Pf*DHFR and their impact on inhibitor activity

Mutation:	**Associated drug-resistance:**	**Inhibitor activity:**
S108N	Pyrimethamine	Initial mutation conferring moderate resistance to pyrimethamine.
N51I	Pyrimethamine	When combined with S108N, increases resistance to pyrimethamine.
C59R	Pyrimethamine	Enhances resistance in combination with S108N and N51I mutations.
I164L	Pyrimethamine and cycloguanil	High-level resistance to pyrimethamine and cycloguanil, especially when combined with other mutations.
A16V	Cycloguanil	Confers resistance to cycloguanil but not pyrimethamine.
S108T	Cycloguanil	When combined with A16V, significantly increases resistance to cycloguanil.

While different amino acid substitutions can hinder the effectiveness of drug molecules, the mutational possibilities are limited by the fact that the natural substrate still needs to bind and undergo the required reaction[13]. This is fortunate as it means the enzyme cannot indefinitely mutate to avoid inhibition without compromising its essential function, thereby preserving the potential for the development of new effective class II antifolates.

There are numerous crystal structures of the *Pf*DHFR enzyme available in the Protein Data Bank (PDB), containing a variety of bound small molecule ligands and bearing several different mutations in the enzyme active site, which are responsible for the enzyme's resistance to inhibition. Understanding the structure of the enzyme is key for designing effective class II antifolates which can remain active in the presence of resistance conferring mutations.

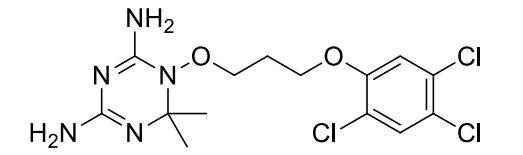
1.2.4 Overcoming resistance to class II antifolates in the treatment of malaria

Much insight has been gained into the mechanisms of parasite resistance thanks to the resolution of crystal structures of mutant enzymes. These detailed structures have revealed how specific mutations alter the active site and affect drug binding, providing valuable information for designing more effective inhibitors. As mentioned above, one attractive strategy for overcoming resistance is to introduce increased flexibility in drug molecules, thereby allowing them to arrange themselves favourably in the active site maintaining the same desired interactions but avoiding steric clashes from mutant residues[40]. In the development of new cycloguanil and pyrimethamine analogues, additional flexibility is imparted by adding a multi-atom linker between the two ring systems. In the case of cycloguanil and pyrimethamine, the two rings are directly joined, greatly limiting the rotational possibilities of these molecules due to the rigid bi-aryl axis. With the addition of a flexible linker, the rotational freedom between the rings is increased, allowing the molecules to adopt conformations that can better fit into the active site of mutant enzymes. This enhanced flexibility can improve binding affinity and effectiveness, potentially overcoming parasite resistance mechanisms[13].

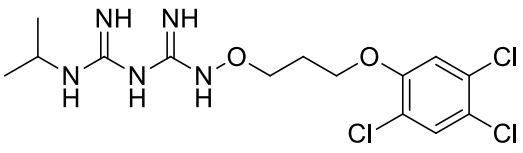
Many flexible analogues of cycloguanil and pyrimethamine have been developed in recent years and have shown improved binding ability against mutant DHFR parasite enzymes. One of the first compounds of this nature to be synthesized is WR99210, shown in figure 1.16. Like cycloguanil, this molecule contains a 2,4-diamino-dihydrotriazine core, with a dimethyl substituent in the 6-position. It however has a five-atom flexible linker in the 5-position connected to the halogenated terminal ring. Although highly potent against drug-resistant strains of *P. falciparum*, this compound has never been developed into a successful drug candidate for various pharmacokinetic and toxicity related factors[13]. Among these is low oral bioavailability. It is suggested that under the acidic conditions of the gastrointestinal tract, the triazine core of WR99210 becomes fully protonated, and as such, is unable to diffuse across cell membranes for absorption[14]. Cycloguanil is able to overcome this by being administered in pro-drug form. Rearrangement into the active compound only occurs in the liver, after it has already entered systemic circulation. PS-15 [Fig 1.16] is a pro-drug form of WR99210,

acting in the same manner as proguanil, and has shown greatly improved oral bioavailability[13]. Pyrimidines are slightly less basic than triazines and exist in an equilibrium between protonated and unprotonated species in the gastrointestinal tract. This allows for increased levels of absorption and may make them more suitable for the development of new drugs[41].

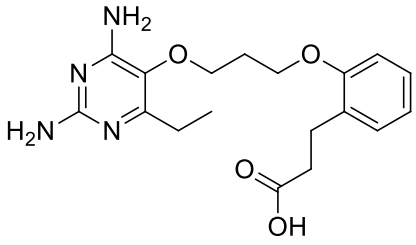
P218, shown in figure 1.16, is a flexible pyrimethamine analogue reported in 2012. It retains the conserved 2,4-diamino-6-ethyl-pyrimidine core of pyrimethamine[41]. Similarly to WR99210, it features a five-atom linker connected to a substituted benzene ring. However, instead of halogen substituents, P218 contains a 2-carboxyethyl group attached to its terminal ring. This allows for additional charge mediated interactions to asparagine 122 in the active site of the target enzyme, analogous to the binding of the natural substrate DHF[41]. This compound displayed excellent activity *in vitro*, inhibiting quadruple mutant drug-resistant *P. falciparum* strains in the low nanomolar range in whole cell assays. P218 is currently undergoing clinical trials for administration in humans. Early results show promising efficacy for this compound, including favourable toxicity and pharmacokinetic profiles[42].



WR99210



PS-15



P218

Figure 1.16: Structures of flexible class II antifolates, WR99210 and P218. PS-15, a pro-drug of WR99210, is also included

Both WR99210 and P218 were developed using rational drug design principles. Insights into the binding mode of the natural substrate and the mechanisms of resistance to existing drugs were obtained by studying the crystal structures of mutated enzymes. This information guided the design of effective inhibitors of drug-resistant *P. falciparum* DHFR. In the case of WR99210, a homology model was constructed, and it was predicted that increased flexibility between the rings of cycloguanil would enhance binding against mutant enzymes[13]. In the development of P218, various protein crystal structures containing synthesized molecules bound in the active site of the target enzyme were obtained. This allowed researchers to identify additional potential binding interactions for new drug molecules and led to the selection of the 2-carboxyethyl substituent to form charge mediated bonds to asparagine 122.

The development of new therapeutic compounds is often guided by Target Candidate Profiles (TCP) and Target Product Profiles (TPP), which outline the essential characteristics a drug must possess to be a viable therapeutic option. For new antimalarials, these criteria have been well defined by the Medicines for Malaria Venture (MMV), which states that novel compounds must exhibit potent activity against *P. falciparum*, favourable pharmacokinetics suitable for single dose or short course treatments, and low risk for the development of parasite resistance[42]. Several specific TCP goals have also been established, ranging from molecules that block parasite transmission, by targeting gametocytes, to compounds that act on the insect vector rather than the parasite itself. Class II antifolates align most strongly with TCP-1[42], which includes compounds capable of interrupting the asexual blood-stage of the parasite infection, either as monotherapies or in combination with other suitable antimalarials.

1.3 Molecular modelling in drug discovery

Molecular modelling is a powerful tool which can accelerate many of the most resource intensive and time-consuming stages in the drug development process. High-throughput screening can be used to identify hit compounds from large libraries of potential candidates, and further molecular docking experiments are useful in the early-stages of lead development to assess the specific effects of functional group changes. Molecular dynamics (MD) simulations offer insights into the dynamic nature of protein-ligand interactions by accounting

for the inherent flexibility of amino acid residues. Unlike static crystal structures or simple docking experiments, MD simulations capture the full range of conformational changes and interactions that occur in realistic biological systems. Furthermore, various physicochemical parameters of drug candidates can be evaluated and their pharmacokinetic characteristics predicted using ADMET (absorption, distribution, metabolism, excretion, toxicity) prediction tools.

Until recently, the identification of new hit compounds required screening hundreds or thousands of potential candidates in various *in vitro* assays. This is an immense undertaking, with no guarantee of success. Computational high-throughput screening makes it possible for this process to be done entirely *in silico* and allows for a greater range of structural diversity to be assessed as evaluated compounds do not actually have to be synthesized or isolated. In lead development, chemists often synthesize vast libraries of analogues, making slight functional group changes to elucidate the nature of substrate binding to the target enzyme and to generate a required minimum pharmacophore through trial and error. This process can be expedited using molecular modelling techniques, which allow for specific interactions to be predicted and optimized *in silico*, thereby reducing the need for extensive laboratory synthesis and *in vitro* testing. The prediction of physicochemical parameters also greatly assists in the drug development process. Optimization of pharmacokinetics by means of various ADMET properties can be done before any synthesis takes place, allowing for a more efficient and streamlined process, and increasing the likelihood of developing a successful drug candidate.

1.4 Project introduction and aims

1.4.1 Background of the current project

In 2016, A. Rousseau and colleagues reported a series of flexible cycloguanil analogues which showed excellent activity against drug-resistant *P. falciparum* strains in both single enzyme and whole cell assays[40]. Based on the rationale that a flexible linker would allow for improved

binding against mutant enzymes, a series of compounds was synthesized containing linkers of different lengths. The compounds featured a conserved 2,4-diamino-dihydrotriazine core, with a phenyl group in the 6-position, and the flexible linker attached to the 5-position of the dihydrotriazine ring, given by the general structure shown in figure 1.17.

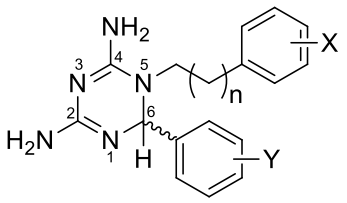
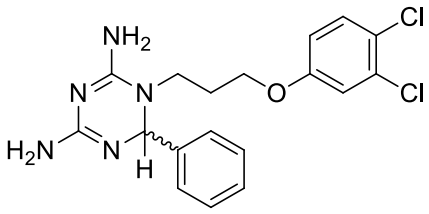


Figure 1.17: General structure of flexible cycloguanil analogues reported in 2016

Various linker lengths were assessed, ranging from two to six atoms. It was found that analogues with a four-atom linker performed the best, with the most potent compounds inhibiting drug-resistant *P. falciparum* in low nanomolar concentrations in whole cell assays. The effect of various substituents around the two phenyl rings were also investigated, and it was shown that while various groups were tolerated around the terminal ring, an unsubstituted phenyl ring in the 6-position of the dihydrotriazine core was optimal. The most active compound synthesized in this series is shown in figure 1.18.

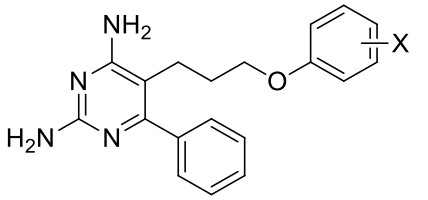


Pf FCR-3 IC_{50} = 39.8 nM

Figure 1.18: Structure of the most active compound of the series synthesized in 2016

Although these compounds displayed excellent anti-plasmodial activity, their synthesis was not well optimized, and they contained a chiral centre. Without the possibility of enantioselective synthesis in their preparation, racemic mixtures were evaluated for activity.

For these reasons a second-generation of compounds was prepared containing a pyrimidine core instead of the dihydrotriazine core[44]. The pyrimidine core is fully aromatised, and the carbon in the 6-position of the ring sp^2 hybridized. This eliminated the chiral centre and allowed for a more effective synthetic route to be employed. It was hoped that these new compounds would display comparable activity to their dihydrotriazine predecessors, this however was unfortunately not the case. The most active compounds in this series only inhibited drug-resistant *P. falciparum* in the low micromolar range in whole cell assays, representing a significant loss in activity. The general structure of these second-generation analogues is shown in figure 1.19.



Pf FCR-3 IC_{50} = 0.86 - 26.55 µM

Figure 1.19: General structure of second-generation flexible analogues

It was postulated that the observed loss in activity was due to the inclusion of a new rigid bond in the second-generation analogues. As the pyrimidine core is fully aromatized, a high energy barrier exists for the rotation of the phenyl ring at the 6-position of the pyrimidine core around the newly formed bi-aryl axis. This rigidity effectively locks molecules into a single conformation, preventing them from avoiding mutant residues in the active site of the target enzyme, drastically lowering their binding ability. While disappointing, these results informed the next stages of lead development.

The investigations conducted in the present project are a continuation of this work. The same core functionalities are conserved, specifically the 2,4-diaminopyrimidine core with a four-atom flexible linker in the 5-position, attached to a substituted terminal ring. Instead of a phenyl substituent at the 6-position, however, various linear or cyclic aliphatic hydrophobic groups are included. This approach preserves the desired hydrophobic interactions in the active site while allowing for free rotation of the new substituent. This ensures that the

molecules can remain highly flexible and are able to adopt favourable conformations for binding to mutant enzymes.

1.4.2 Aims

Based on the earlier findings that a four-atom linker is optimal, and that a pyrimidine core with an aliphatic hydrophobic substituent in the 6-position has some advantages over a dihydrotriazine core, compounds with the general structure given in figure 1.20 are investigated.

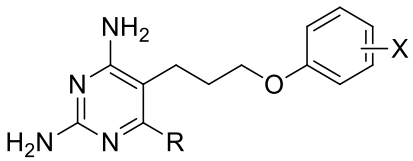


Figure 1.20: General structure of compounds investigated in this project

Molecular modelling tools are used to evaluate the effectiveness of different substituents on the provided general structure. To investigate the dimensions of the hydrophobic pocket within the active site, various cyclic and linear aliphatic hydrophobic groups of differing sizes are examined as R-substituents. Additionally, numerous analogues with common aromatic substituents at the X-position are assessed to potentially identify new interactions that could enhance binding affinity to mutant enzymes. ADMET prediction tools are also employed to ensure that the compounds exhibit favourable pharmacokinetic properties which are essential for the development of successful drug candidates. Furthermore, molecular dynamics (MD) simulations are performed on some of the more promising compounds identified in the molecular docking experiments to gain insight into the specific nature of the binding interactions to the target enzyme. Based on the results of the combined molecular modelling experiments, several candidates are selected for synthesis using partially established methods, which are optimized for the efficient synthesis of the selected compounds. The prepared compounds are then put forward for *in vitro* antimalarial assessments, including

both single enzyme *Pf*DHFR and whole cell assays against wild-type and drug-resistant strains of *P. falciparum*. Following these initial evaluations, structural changes may be implemented in the design of second-generation compounds for improved activity or pharmacokinetic properties.

1.4.3 Proposed methodology

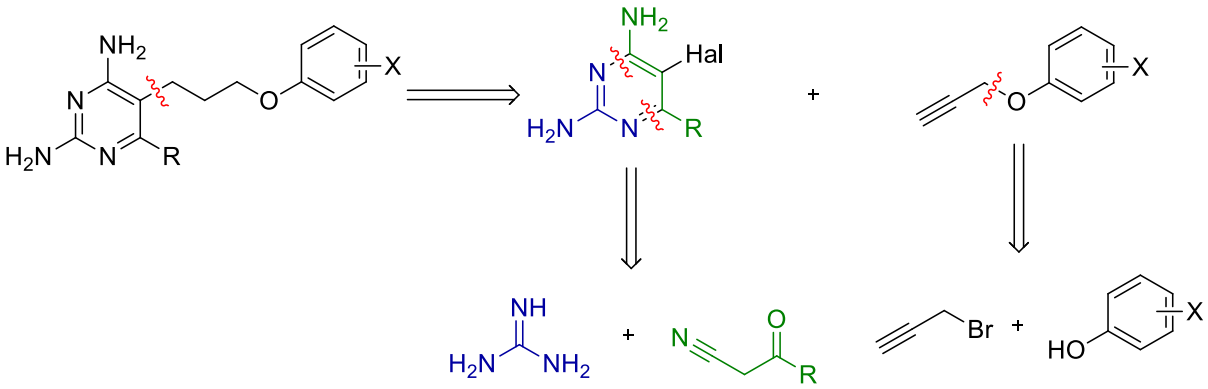
1.4.3.1 Molecular modelling

For the molecular modelling portion, various protocols and algorithms are implemented, all of which are contained within the Schrodinger molecular modelling suite of software. These methods are discussed in detail in chapter 2, but a brief outline is provided below.

Relevant protein crystal structures of drug-resistant *Pf*DHFR enzymes are imported from the protein data bank (PDB). They are prepared for subsequent modelling by refinement with epik, and energy minimization using the OPLS4 force field, as this is the most advanced force field available in the version of the modelling software used. Ligand design and library enumeration are carried out, generating numerous structures containing different R and X-substituents. Docking scores for these compounds are then calculated using the Glide algorithm. Additional Induced Fit Docking (IFD) procedures are also implemented, in order to account for protein flexibility in the docking experiments. Further molecular dynamics simulations are conducted on some of the top scoring compounds in order to accurately assess the binding mode of the compounds to the target enzyme. In addition to the methods mentioned above, QikProp is used for the estimation of various physiochemical properties in order to fully assess the suitability of the compounds of interest as potential drug candidates for oral administration.

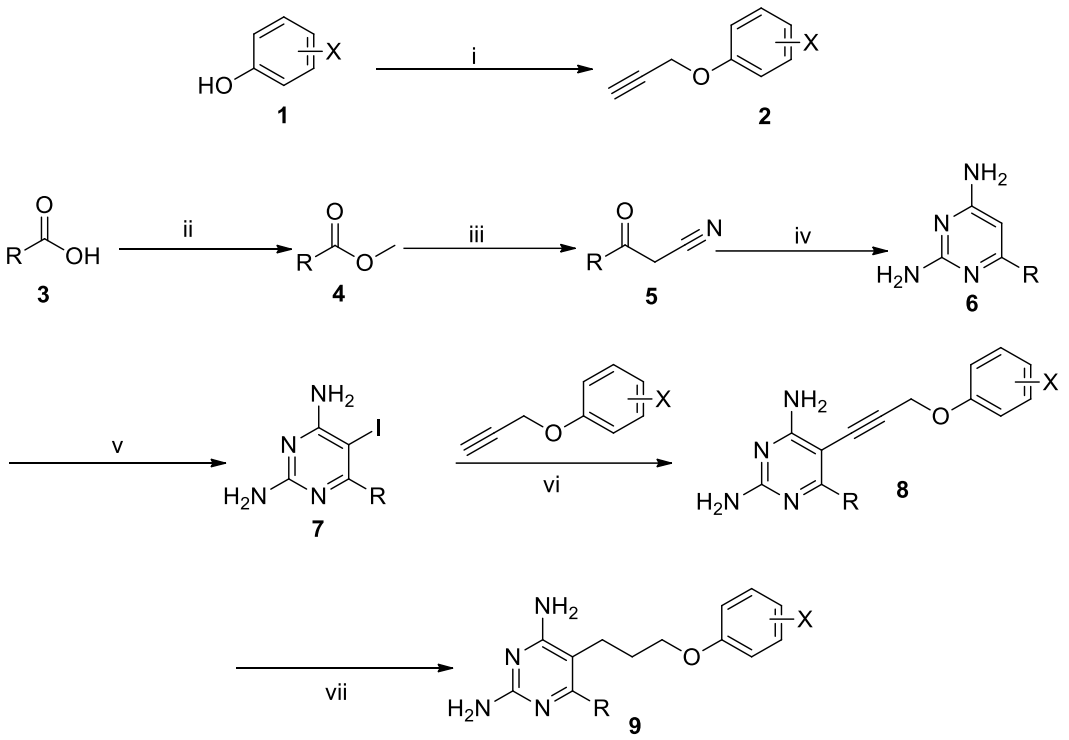
1.4.3.2 Synthesis

The target compounds are formed from the attachment of the flexible linker and terminal ring portions to the 5-position of the 2,4-diaminopyrimidine core. A disconnection approach to the synthesis is shown in scheme 1.2



Scheme 1.2: Retrosynthesis of flexible pyrimethamine analogues

The general synthetic method for the preparation of the final compounds is shown in scheme 1.3 and described in the following paragraph.



Scheme 1.3: Synthetic procedure for the production of flexible pyrimethamine analogues.

Reagents and conditions: i) Propargyl bromide, K_2CO_3, acetone, RT, 24 hrs; ii) H_2SO_4, MeOH, reflux, 24 hrs; iii) KOtBu, MeCN, IPA, 2-MeTHF, RT, 2 hrs; iv) Guanidine HCl, KOtBu, 2-MeTHF, MW, 150 W, 150 °C; v) I_2, HIO_3, H_2SO_4, CH_3COOH, H_2O, reflux, 24 hrs; vi) $Pd(PPh_3)_2Cl_2$, CuI, DIPA, DMF, reflux, 24 hrs; vii) Pd/C, H_2(g), MeOH, RT, 24 hrs.

Appropriately substituted phenols (**1**) are alkylated with propargyl bromide under basic conditions, yielding the four-atom linker and terminal ring portion of the desired compounds (**2**). The formation of the 2,4-diaminopyrimidine core occurs over several steps. Starting with commercially available carboxylic acids containing appropriate R-substitutions (**3**), an esterification reaction is performed with methanol under acidic conditions to obtain the corresponding methyl esters (**4**). The esters then undergo a base mediated nucleophilic substitution reaction with acetonitrile, forming the desired beta-keto nitrile products (**5**). Next, these are combined with guanidine hydrochloride under basic conditions in a microwave assisted ring closure reaction for the formation of the 2,4-diaminopyrimidine core (**6**). For the subsequent attachment of the linker and terminal ring portions in a cross-coupling reaction,

the 5-position of the pyrimidine core needs to be halogenated. This is achieved by a reaction with iodine under acidic conditions, forming the desired 5-iodo pyrimidines (**7**). The iodinated intermediates are then combined with the previously prepared propargyl substituted phenols in a palladium catalysed Sonogashira coupling reaction. This connects the 2,4-diaminopyrimidine core to the linker and terminal ring portions, affording the alkyne containing intermediates **8**. Finally, the triple bond is reduced in a hydrogenation reaction to yield the fully flexible final compounds (**9**) over a total of seven steps.

1.4.3.3 *In vitro* antimalaria testing

Synthesized compounds are evaluated for anti-plasmodial activity and human toxicity in single enzyme and whole cell assays. *Pf*DHFR and *Hs*DHFR enzymes are expressed, and inhibition constant (Ki) values reported. Compounds are further tested against TM4/8.2 (wild type *Pf*DHFR) and V1/S (QM *Pf*DHFR) strains cultured in human erythrocyte cells, and two mammalian derived cell lines for cytotoxicity assessments. The assay details, and discussion of the results are presented in chapter 3. All assays are conducted by the BIOTEC research group in Thailand. The general method followed for the *in vitro* experiments is provided in chapter 6.

Chapter 2 – Molecular Modelling

2.1 Introduction

This chapter deals with the extensive molecular modelling studies that were conducted to identify compounds for synthesis and testing against drug-resistant strains of *Plasmodium falciparum*. The Schrödinger molecular modelling suite was used, and all simulations were run in the Maestro user interface. Schrödinger release version 2023-2[45] was used for generating the molecular dynamics simulations, and version 2024-2[46] was used for everything else. The specific parameters used for each experiment are included in screenshots of the Maestro user interface contained in appendix A.

Protein crystal structures of mutant *Plasmodium falciparum* DHFR-TS enzymes were imported directly from the RCSB Protein Data Bank (PDB). All structures were obtained by X-ray crystallography and have various ligands bound in the dihydrofolate reductase active site, along with the cofactor NADPH. Protein structures were prepared for subsequent analysis using the built-in protein preparation workflow[47,48]. This is a vital first step in any molecular modelling work as protein structures obtained from the PDB are often incomplete. The protein preparation workflow fills in missing portions of the protein structure, including absent hydrogen atoms and highly disordered, solvent exposed loops which are not able to be resolved from crystallography. The hydrogen bonding network within the protein is optimized at a given pH, and the final protein structure is minimized using an energy force field. The OPLS4 force field[49] was used for this portion of the project.

The virtual screening workflow was used for method validation and to gain insight into the binding mode of the compounds assessed. This tool combines ligand preparation[50], receptor grid generation, and Glide docking[51–55] into a single user interface. Ligand preparation generates the possible ionization and tautomerization states of the compounds at a specified pH and uses Epik to calculate associated energy penalties for each state. Receptor grids are required for molecular docking and are generated by including the protein region surrounding the bound ligand, along with any specified docking constraints, into a three-dimensional structure file.

For Glide docking, three different algorithms are available: High Throughput Virtual Screening (HTVS), Standard Precision (SP), and Extra Precision (XP). In each method, numerous ligand conformations are generated and sampled for positive interactions against the defined receptor grid. The sampled poses are scored by calculating various energy terms related to protein-ligand contacts. These include hydrogen bonding, lipophilic interactions, electrostatic interactions, and van der Waals dispersion forces. These parameters are combined into the overall GlideScore, which is well suited for distinguishing between compounds with high binding ability and those with weak affinity for the active site.

The three Glide algorithms vary in the extent to which they sample ligand conformational possibilities, representing a trade-off between speed and accuracy. GlideHTVS is the fastest and performs only limited conformational sampling, while GlideSP employs an exhaustive sampling approach. GlideXP is the most resource-intensive algorithm, using an anchor-and-grow method that iteratively refines good initial poses to identify the most favourable ligand binding conformations. GlideXP also introduces additional energy parameters to the GlideScore, including a hydrophobic enclosure term that rewards ligands for their potential to displace water molecules from hydrophobic sites.

The ligand state penalties calculated during ligand preparation are incorporated into the GlideScore to produce the Docking Score. This ensures that scores for unfavourable ligand states, which might otherwise rank highly, are adjusted accordingly. GlideScore and Docking Score are both reported in units of kcal/mol, however, they are not directly comparable with experimentally obtained binding energies.

Further Induced Fit Docking (IFD)[56–58] experiments were conducted to assess the flexible nature of the protein active site, and to determine any potential new binding interactions. Traditional docking methods are limited by the fact they do not account for the inherent flexibility of amino acid residues, which can undergo conformational changes to enhance ligand binding in real biological systems.

The IFD protocol combines Glide docking and Prime refinement tools into a single procedure to sample both ligand and protein conformational states. Initial ligand poses are generated by performing a Glide docking experiment on the defined receptor grid. For each pose, Prime refines the amino acids in the active site, optimizing their conformations to best accommodate

the docked ligand. New receptor grids are then created from the refined protein structures. The ligands are redocked, and the final protein-ligand poses are scored.

Two sampling methods are available for use in IFD experiments: the Standard and Extended sampling protocols. Standard sampling generates up to 20 final protein-ligand poses for each ligand and allows for the selection of GlideSP or GlideXP in the final docking step. Extended sampling is much more rigorous, creating up to 80 final protein-ligand poses for each ligand using optimized docking parameters and additional iterations.

The final poses are ranked by a weighted IFDScore, combining energy terms from the Glide docking experiments and the total Prime energy, which is calculated for the active site of the ligand-protein complex.

In the Standard sampling method, the IFDScore is defined as follows:

$$IFDScore = GlideScore + 0.05\ X\ Prime_Energy$$

For Extended sampling, the IFDScore is calculated using:

$$IFDScore = 1.0\ X\ Prime_Energy + 9.057\ X\ GlideScore + 1.428\ X\ Glide_Ecoul$$

The calculated IFDScore, and the various terms which comprise this function, are given in units of kcal/mol. It is important to note that the ligand state penalties calculated during ligand preparation are not considered in the final IFDScore, as this parameter is derived from the GlideScore, not the adjusted Docking Score. Because the two sampling methods make use of different weightings and terms to calculate the overall IFDScores, the results obtained for the two protocols are not directly comparable.

Molecular dynamics (MD) simulations were run utilising Desmond[59,60]. These simulations made use of some of the best scoring poses generated from the IFD experiments and were submitted to the Centre for High Performance Computing (CHPC) cluster hosted by the CSIR.

MD simulations are used to study the dynamic nature of protein – ligand interactions and output many useful figures and parameters. Root mean square deviation (RMSD) plots are generated which show the relative changes in the positions of the ligand and protein molecules during the simulation, based on their initial conformations. MD simulations of tightly bound ligands will have RMSD plots that reach a steady state, indicating that an optimal binding mode is found and maintained throughout the simulation time. The overall RMSD values can be seen for the specific amino acid residues, and the constituent atoms of the ligand molecules. Protein – ligand contact diagrams are also produced, which show the interactions experienced between the ligand molecule and the active site residues. The strength of the interactions is indicated by the proportion of the total simulation time that they exist for.

Various key physicochemical parameters of the compounds investigated in this project were estimated with QikProp[61] to determine their suitability as drug molecules for synthesis.

Many of the figures presented in this section were created with PyMOL version 3.0.3[62]. This software allows for the generation of high quality three-dimensional images using ray tracing methods.

2.2 Aims and objectives

In this work, various *Pf*DHFR protein crystal structures containing the desired drug-resistant quadruple mutation N51I/C59R/S108N/I164L are evaluated, and the core binding modes of known inhibitors identified. The reliability of the docking methods used are assessed by performing self-docking and enrichment docking experiments.

Compounds with the general structure shown in figure 2.1 are investigated for binding ability against the target enzyme using various molecular modelling techniques.

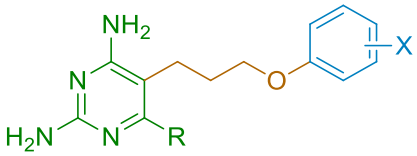


Figure 2.1: General structure of project compounds

The general structure of the compounds assessed is broken down into three parts. The first is the core 2,4-diaminopyrimidine ring, which is indicated in figure 2.1 in green. This portion of the molecule is essential for binding to the active site of the target enzyme and contains an R-substituent in the 6-position of the ring. The second group is the flexible four-atom linker, attached to the core 2,4-diaminopyrimidine ring at the 5-position. The flexibility imparted by this group allows molecules to bind despite drug-resistant mutations in the active site of the target enzyme. The final structural element is the terminal ring, indicated in blue in figure 2.1 and is attached to the end of the four-atom linker.

Different aliphatic hydrophobic R-groups are assessed as substituents in the 6-position of the 2,4-diaminopyrimidine ring. There is a hydrophobic pocket in this region of the active site[40], and flexible docking methods are used to probe its extent. The flexibility of the amino acid residues comprising the hydrophobic pocket are simulated to accurately account for any conformational changes that may allow for binding of bulky R-substituents.

Further flexible docking studies are conducted on compounds containing various common aromatic X-substituents around the terminal ring, with the aim of identifying any potential new interactions in this region of the active site. Cationic and anionic groups, as well as groups containing H-bond donors and acceptors are evaluated to maximize the possibility of interactions with the nearby amino acid residues.

Molecular dynamics simulations are run on notable analogues from the initial docking experiments to further verify the results obtained, and to provide insight into the dynamic nature of ligand – protein interactions.

Finally, various physiochemical parameters are estimated *in silico* to determine the compounds' suitability for subsequent synthesis.

Based on the combined results of the molecular docking experiments and the prediction of physicochemical properties, a series of analogues is selected for synthesis and evaluation against drug-resistant strains of *Pf*DHFR.

2.3 Validation of *Pf*DHFR-TS crystal structures and ligand binding mode

2.3.1 Assessment of protein crystal structures

Molecular modelling experiments rely on high quality protein crystal structures. To date, there are 34 crystal structures of *Plasmodium falciparum* DHFR available on the PDB. 19 of these structures contain the drug-resistant quadruple mutant with N51I/C59R/S108N/I164L substitutions, which is of relevance in this work. Filtering out structures with resolution greater than 2.5 Å leaves the following 11 structures listed in table 2.1 for potential analysis. Given the numerous available structures, a criterion was defined to include only those with a resolution of 2.5 Å or better, ensuring that only high-quality structures were carried forward for analysis.

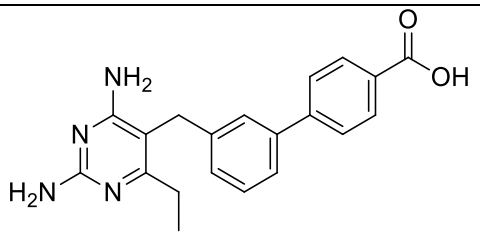
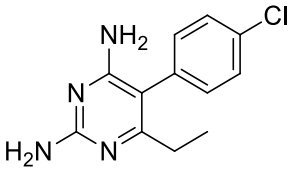
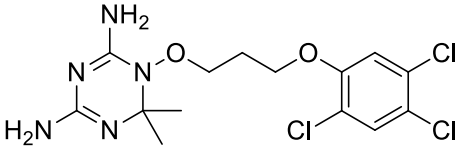
Table 2.1: High resolution PDB structures of N51I/C59R/S108N/I164L mutant *Pf*DHFR enzymes

PDB Code:	**Resolution (Å):**	**Bound ligand:**	**Year:**	**Reference:**
8JFC	2.30	B21591	2023	63
8JFD	2.30	B21594	2023	63
3QG2	2.30	Pyrimethamine	2011	64
6A2L	2.38	BT1	2019	65
6A2N	2.45	BT2	2019	65
1J3K	2.10	WR99210	2003	66
6KOT	2.15	B12128	2019	67

PDB Code:	Resolution (Å):	Bound ligand:	Year:	Reference:
6LHJ	2.40	C452	2020	68
6KP2	1.97	B10042	2019	67
6KP7	1.97	B12154	2019	67
6KPR	2.10	B12155	2019	67

Various ligands are included in the DHFR active sites of the above protein structures, along with the bound cofactor NADPH. The structures 1J3K, 3QG2, and 8JFC were selected for analysis based on the similarity of their bound ligands to the project compounds investigated. These ligands are highlighted in table 2.2.

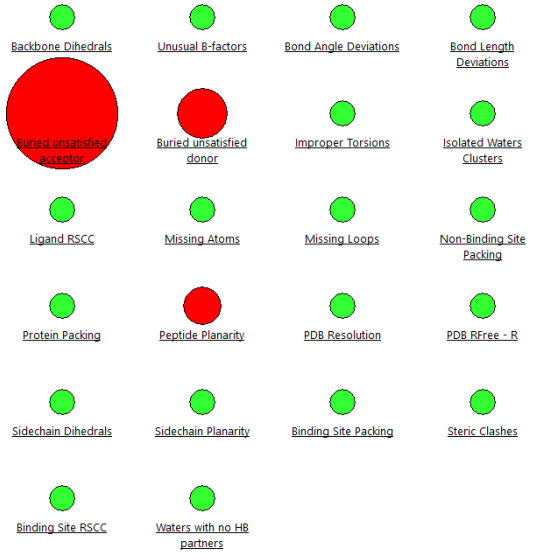
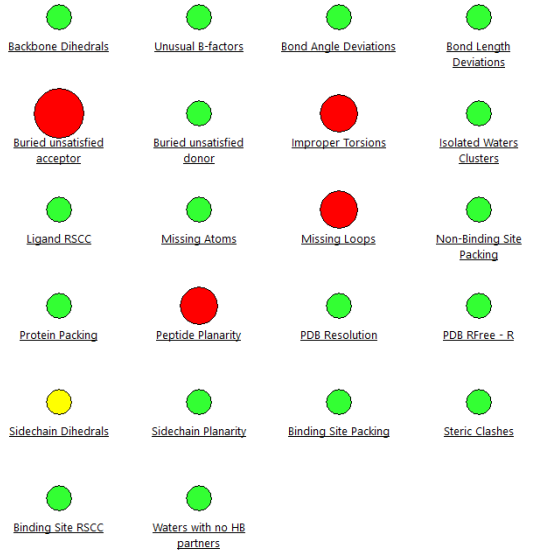
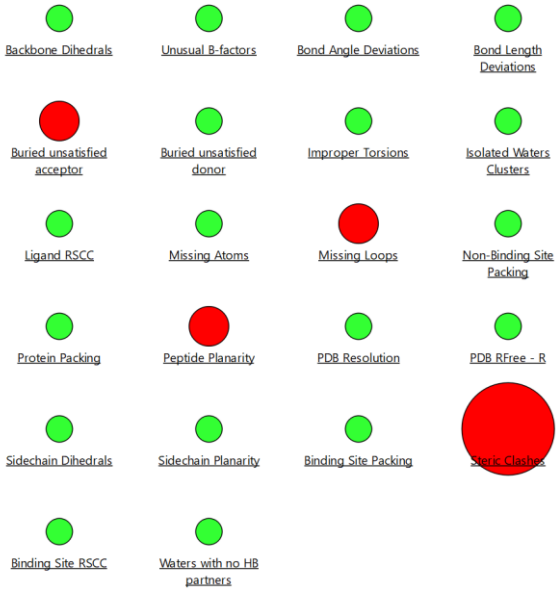
Table 2.2: Structures of bound ligands included in selected protein crystal structures

PDB Code:	Bound ligand:	Ligand structure:
8JFC	B21591	
3QG2	Pyrimethamine	
1J3K	WR99210	

The chosen protein structures were imported directly from the PDB into the protein preparation wizard. The default protein preparation settings were largely used (Appendix. A 1.1), with ionization states generated at a pH of 7.4 ± 2 and the option to fill in missing loops with Prime selected. All waters of crystallization were deleted, and the finalized protein structures were then minimized using the OPLS4 energy force field. Protein reliability reports

were generated for each of the prepared proteins. The ligand was chosen from one of the DHFR active sites, on chain A, and the area within 15 Å of the selected ligand was assessed (Appendix. A 1.2). The protein reliability report tool is useful for determining the quality of the prepared structures, and highlights deviations from expected values. Parameters falling within the normal range are indicated by small green circles, with an increasing radius and a colour shift to red indicative of greater deviations. The obtained reliability reports for the three prepared proteins are presented in table 2.3.

Table 2.3: Protein reliability reports for prepared structures 1J3K, 3QG2, and 8JFC

1J3K	3QG2	8JFC
		

Overall, this is a good result for each of the three protein structures, with most of the assessed parameters falling within the expected ranges. Structure 1J3K shows a deviation in the amount of buried unsatisfied H-bond acceptor atoms, however, these were identified as the oxygens present in the linker of bound ligand WR99210, and do not reflect the quality of the protein structure as a whole. Some steric clashes can be seen in structure 8JFC. These are due to a slight overlap in the position of the bound ligand B21591 and some nearby residues. This, however, will not affect any subsequent docking experiments which take ligand flexibility into account to minimize unfavourable steric interactions

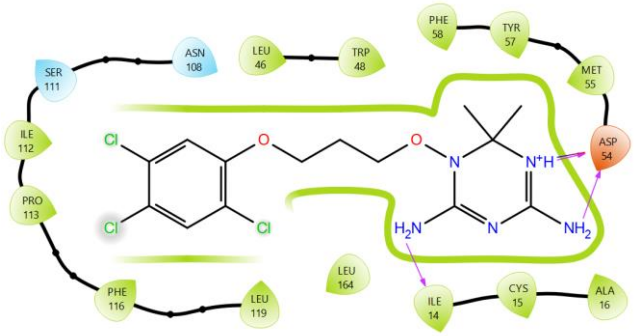
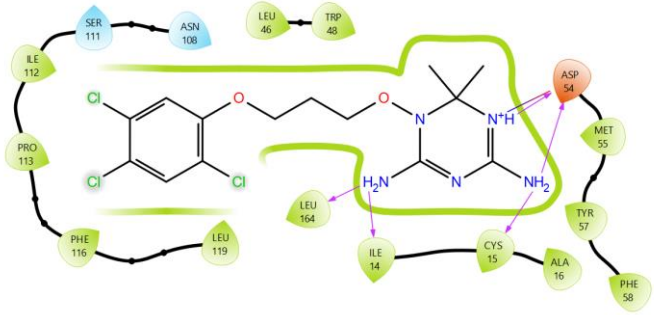
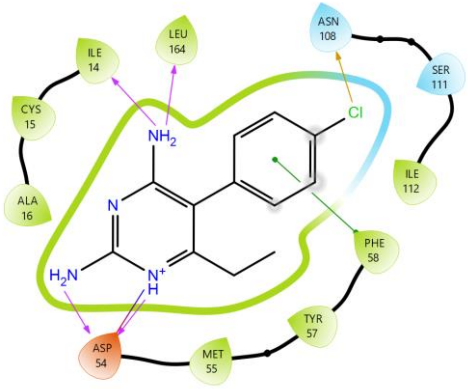
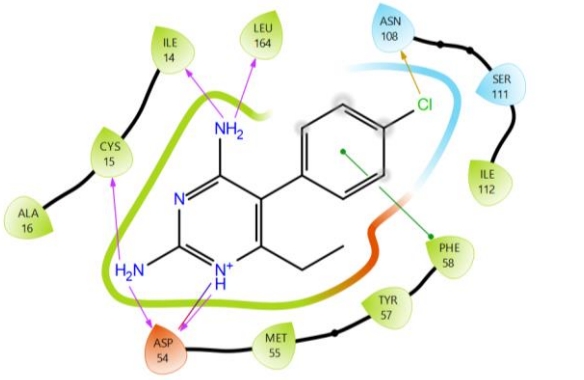
2.3.2 Identification of ligand binding mode and self-docking experiments

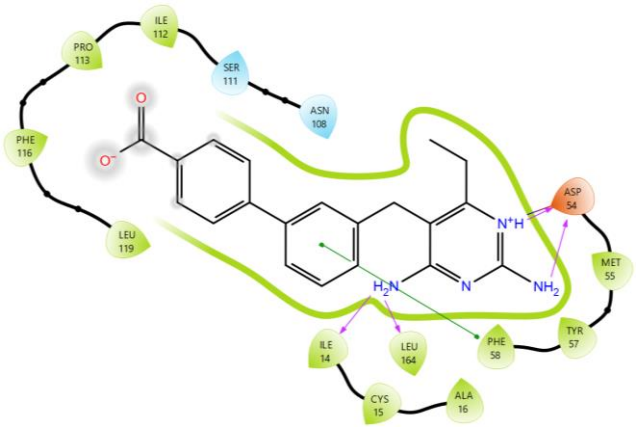
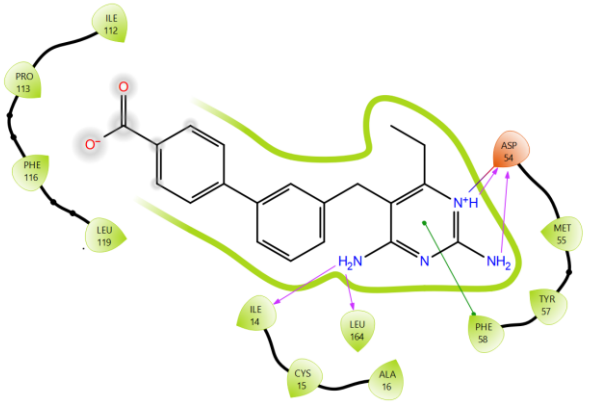
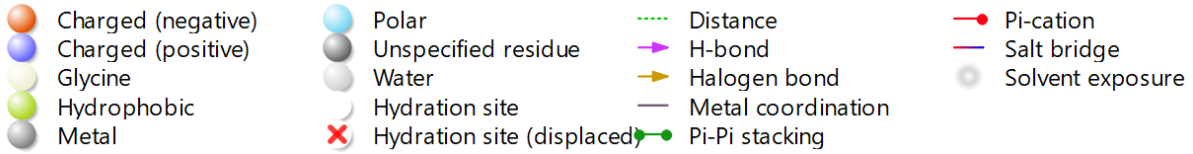
Examination of the three protein structures revealed that a conserved binding mode is adhered to for each of the bound ligands in the DHFR active site. The conserved binding mode includes H-bonds donated to both oxygens of the Asp54 carboxylate side chain, and additional H-bonds donated to the backbones of residues Ile14 and Leu164. Furthermore, the nitrogen in the 1-position of the core diamino ring is protonated for each ligand, facilitating an additional electrostatic interaction with the carboxylate anion of residue Asp54.

A self-docking experiment was conducted by extracting the ligands from the crystal structures and re-docking them into the active site of their respective proteins. This was done to ensure that the conserved binding interactions were adhered to, and that the ligand pose found in the crystal structure could be recreated by the docking methods used. To this end, the virtual screening workflow (VSW) was implemented (Appendix. A 1.3). The different possible ionization states of the ligands WR99210, pyrimethamine (Pyr), and B21591 were generated at pH 7.4 ± 2. Receptor grids were then created from the prepared protein structures 1J3K, 3QG2, and 8JFC, and the various ligand states docked into their respective active sites. Glide high-throughput virtual screening (HTVS) was used to produce initial poses, followed by docking with the more rigorous GlideXP algorithm. It was found that for the initial GlideHTVS docking, H-bond constraints to residue Asp54 were required to consistently obtain the desired binding poses.

The series of figures presented in table 2.4 highlights the specific interactions between the bound ligands and the residues in the protein active site. Interaction diagrams are shown for the ligand bound in the crystal structures, as well as the best poses generated from the self-docking experiments. 2D interaction diagrams are created by flattening out three-dimensional poses into a two-dimensional image. They are useful for showing protein – ligand interactions, but all spatial and conformational information is lost.

Table 2.4: 2D interaction diagrams of bound ligands in structures 1J3K, 3QG2, and 8JFC as well as the top poses generated in the self-docking experiments

Structure and ligand:	**Crystal structure:**	**Self-docked:**
1J3K WR99210		
3QG2 Pyr		

Structure and ligand:	Crystal structure:	Self-docked:
8JFC B21591		
Legend:		

The core binding mode is shown in the figures above, with the self-docked ligands adhering to each key interaction. The best scoring ionization state for each ligand matches those found in the crystal structure, with protonation of the 1-N of the diamino ring. The docked poses also show the potential for an additional H-bond to Cys15, and pi-pi stacking with Phe58. A detailed three-dimensional view of the ligand WR99210 self-docked in the active site of 1J3K is provided in figure 2.2.

Some interactions observed in the self-docking experiments were not present in the poses derived directly from the crystal structures. This is likely because docking experiments consider multiple ligand binding poses and select the single most optimal configuration. While these additional interactions may also occur in the actual ligand binding process, they may not always be captured in crystal structures due to the dynamic nature of ligand–protein interactions. Since crystal structures represent only a single static conformation, some valid binding interactions may not be evident.

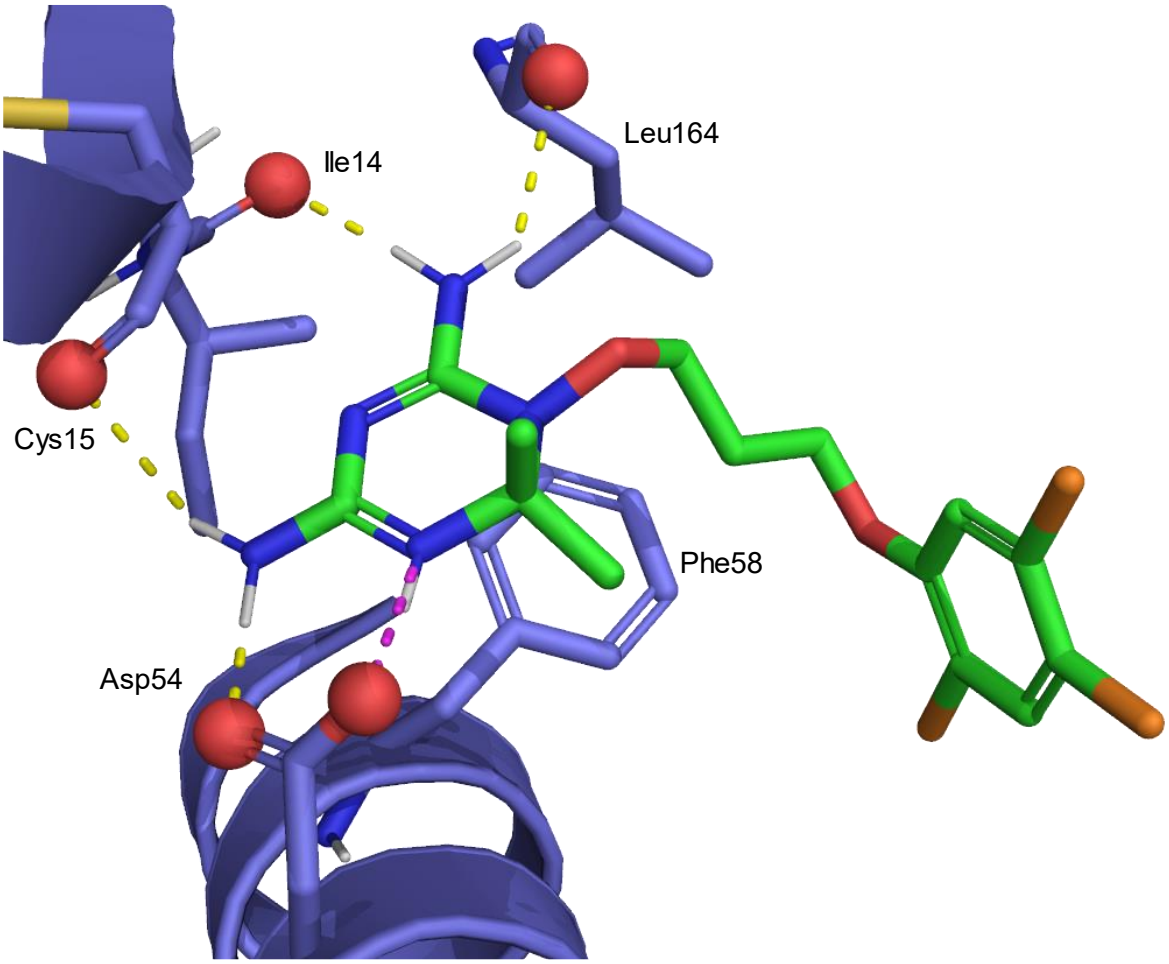


Figure 2.2: Docking pose generated from self-docking of ligand WR99210 into the active site of 1J3K

Figure 2.2 highlights the conserved binding mode. The nitrogen in the 1-position of the diamino ring is protonated, allowing for an ionic interaction with the adjacent Asp54

carboxylate side chain. The 2-amino group donates H-bonds to the side chain of Asp54, and to the backbone of Cys15. Additional H-bonds are seen between the 4-amino group and the backbone of residues Ile114 and Leu164. The phenyl side chain of residue Phe58 is also displayed, showing the potential for pi-pi stacking with either of the ligand's aromatic rings.

Figure 2.3 displays pyrimethamine bound in the active site of *Pf*DHFR, as found in the protein crystal structure 3QG2. Relevant amino acids are highlighted, including mutant residues and the nearby residues involved in binding interactions. Notably, the side chains of the mutant residues Leu164 and Asn108 are in close proximity to the phenyl ring of pyrimethamine, introducing steric hindrance that reduces its binding affinity. Additionally, mutations Ile51 and Arg59 alter the shape of the active site, further decreasing pyrimethamine's affinity for the quadruple mutant enzyme.

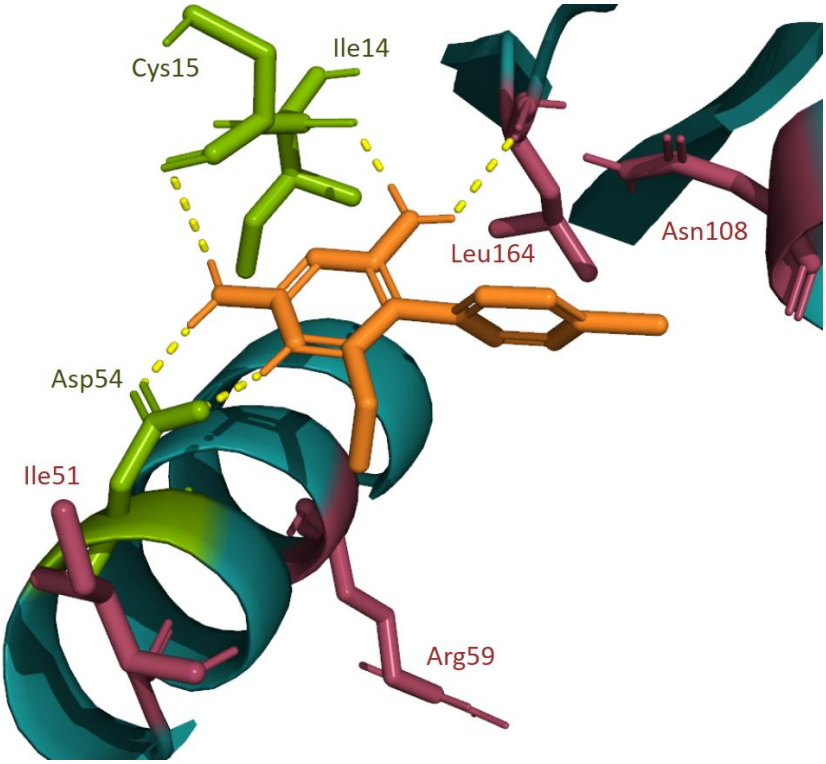


Figure 2.3: Active site of protein crystal structure 3QG2, showing bound pyrimethamine (orange) and relevant nearby amino acid residues. The mutant residues are coloured red.

Table 2.5: Structure of the bound ligands from the crystal structures (orange) superimposed with the poses generated from self-docking experiments (blue)

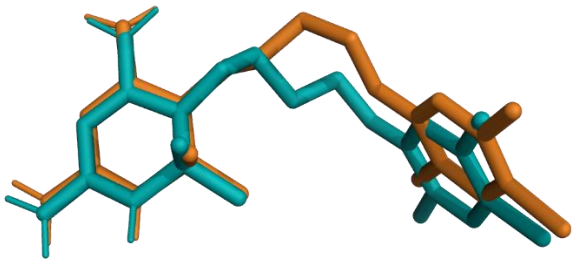
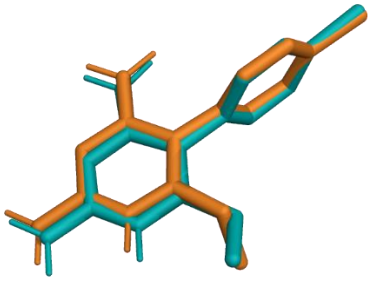
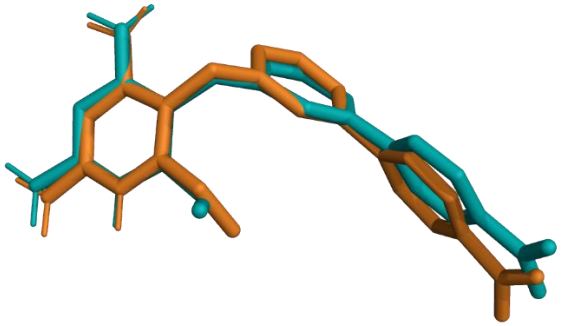
Protein structure: 1J3K **Ligand: WR99210**	
Protein structure: 3QG2 **Ligand: Pyrimethamine**	
Protein structure: 8JFC **Ligand: BT21591**	

Table 2.5 demonstrates the excellent alignment between the poses generated from the self-docking experiments (blue) and the poses extracted from the crystal structures (orange). In each case, the calculated RMSD values between the poses are less than 1 Å, indicating the known binding pose was faithfully recreated by the docking methods used.

2.3.3 Enrichment docking

Docking methods were further validated by performing enrichment docking. A few known binders were combined with a large library of decoy compounds and docked against the target receptor. Good docking methods will identify the known binders, and give them higher docking scores than the decoy molecules on average. In general this is done to reduce the possibility of false positives in docking experiments, where non-active compounds are erroneously given high docking scores.

The bound ligands in the three selected protein structures, along with DHF, P65, and a dihydrotriazine compound were chosen as the known binders and are included in figure 2.4. P65 is a pyrimidine analogue of WR99210 and has shown good activity in biological assays[65]. The dihydrotriazine compound was selected from a series of potent analogues prepared by A. Rousseau and colleagues in 2016[40], which form the foundation of the present work.

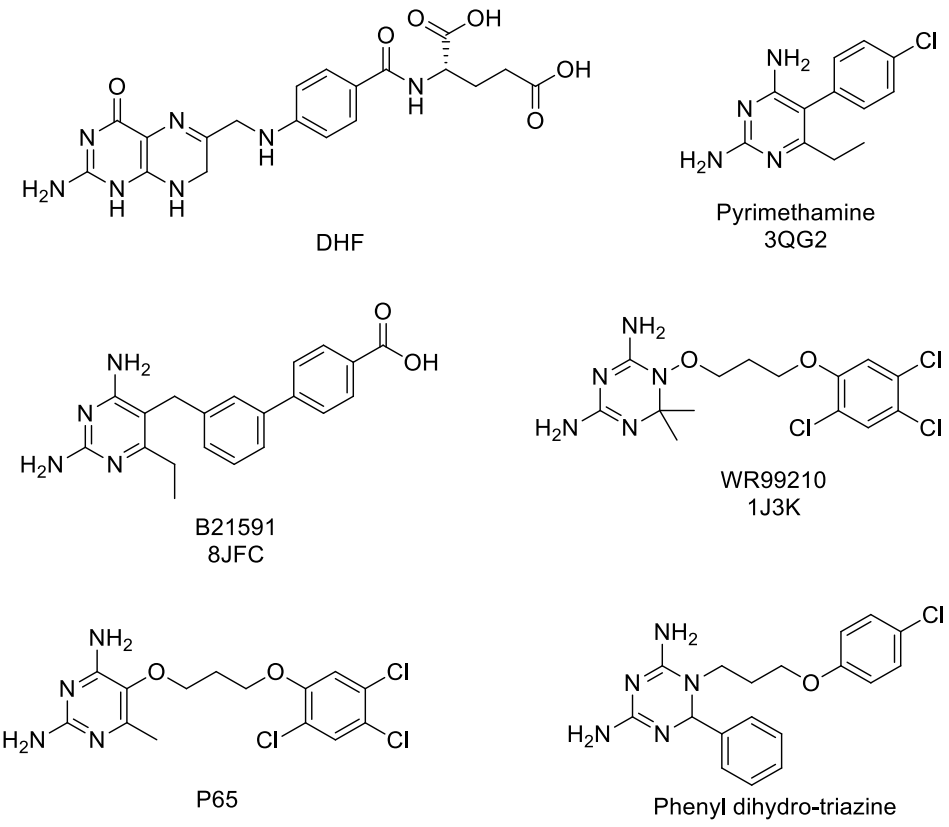
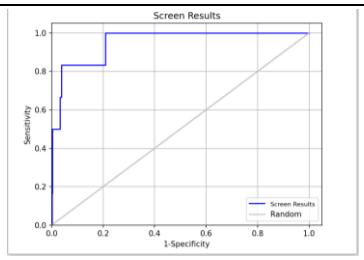
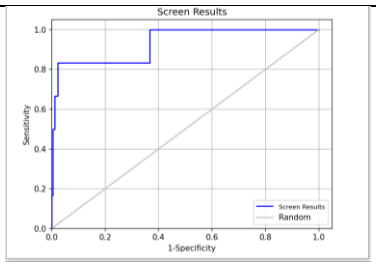
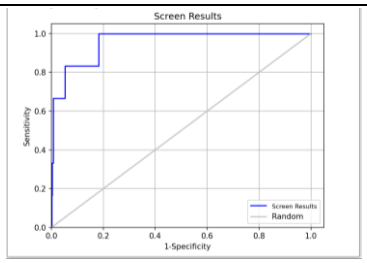


Figure 2.4: Structures of compounds selected as known binders for enrichment docking

The known binders were combined with a library of 1000 random molecules, containing various functional groups relevant for biological activity and an average molar mass of 360 g/mol. The decoy library was obtained from the Schrödinger database. The combined compounds were then docked agaisnt the active sites of the three protein structures using the VSW (Appendix. A 1.4). Steps include ligand preparation, receptor grid generation with H-bond constraints to Asp54 for GlideHTVS docking, and further docking with GlideXP.

The output for each receptor is an ordered list of docked ligands, ranked by their GlideXP score. The enrichment calculator tool (Appendix. A 1.5) was used to create receiver operating characteristic (ROC) plots based on the positions of the known binders in the ordered lists. The area under the ROC curve (AUC) represents a measure of how well the methods performed and has a maximum value of 1. The ROC plots and AUC values for the enrichment docking experiments conducted on each of the three crystal structures are shown in table 2.6.

Table 2.6: ROC plots and AUC values for enrichment docking of 1J3K, 3QG2, and 8JFC protein structures

1J3K	**3QG2**	**8JFC**
		
AUC: 0.95	AUC: 0.93	AUC: 0.96

In each case the known binders were highly ranked by the docking methods used, and AUC values of greater than 0.90 were obtained. This confirms the reliability of the docking methods to not only correctly reproduce the conserved binding mode, but also in successfully identifying the known binders out of a large library of decoy compounds.

Based on these results, any of the three crystal structures evaluated would be suitable for further analysis. The structure 1J3K, however, was selected ahead of the other two as the bound ligand, WR99210, is the most similar to the project compounds. In addition, the

structure 1J3K has been used in many past molecular modelling campaigns, allowing for greater continuity in this work.

2.4 Optimizing interactions in the hydrophobic pocket

There is a hydrophobic pocket in the active site adjacent to the 6-position of the core 2,4-diaminopyrimidine ring in the conserved binding mode[40]. Flexible docking methods were employed to simulate the flexibility of the residues comprising this hydrophobic pocket, with the aim of determining the size of hydrophobic groups which it can accommodate. A range of compounds containing various aliphatic hydrophobic groups in the 6-positon of the 2,4-diaminopyrimidine ring were assessed for binding ability using IFD. Ten different R-groups were selected, ranging in size from a methyl substituent to bulky cyclopentyl and cyclohexyl rings. The requirement for flexibility of the R-groups has been established in earlier work, and as such R-groups were selected which allow for rotation around the C-C bond in the 6-position of the 2,4-diaminopyrimidine ring. The general structure of the compounds assessed is shown in figure 2.5, while the different R-groups selected are indicated in table 2.7.

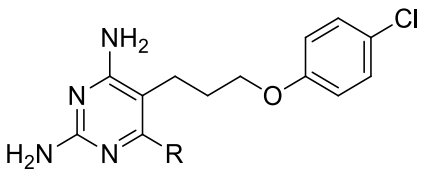
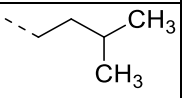
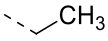
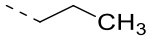
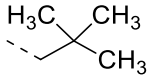
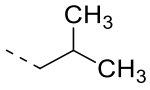
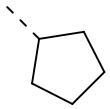
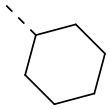


Figure 2.5: General structure of compounds assessed in section 2.4

Table 2.7: Structures of different R-groups assessed

R-group:	Abbreviation:	Structure:		R-group:	Abbreviation:	Structure:
Methyl	Meth			Isobutyl	iBu	
Ethyl	Eth			Cyclobutyl	Cbu	
Propyl	Prop			*tert*-Butyl	tBu	
Isopropyl	iPr			Cyclopentyl	Cpe	
Cyclopropyl	Cpr			Cyclohexyl	Che	

The R-groups were selected based on their relative sizes, with a 4-chloro substituent around the terminal ring included for each analogue so that any changes in binding scores between compounds were due only to the different R-groups. In addition to these project analogues, some known binders were also included in the IFD experiments, namely WR99210, DHF, and compound P65.

For the IFD method, the standard sampling protocol was selected (Appendix. A 1.6). The receptor grid was defined by selecting the bound ligand WR99210 in the active site of protein structure 1J3K, and the earlier identified H-bond constraints to residue Asp54 were added. The IFD workflow does not include ligand preparation, and so this was done separately. It has already been identified that the ligand state in which the 1-N of the core 2,4-diaminopyrimidine ring is protonated is the most active, and so only this state for each ligand was carried forward from the ligand preparation step to save computational resources. For the initial Glide docking and Prime refinement steps, the default settings were selected, with GlideXP chosen for the final re-docking step.

The output of the IFD job is numerous ligand-receptor poses for each ligand, representing the various conformational possibilities of both the ligand molecule and the amino acid residues

in the active site. Each pose is scored by an overall IFDScore, which has already been described.

The IFD results for the ten project analogues assessed, and the three selected known binders are presented in tables 2.8 and 2.9. Each parameter is expressed in units of kcal/mol, where more negative scores represent more favourable binding ability. Table 2.8 shows the compounds ordered by IFDScore, while table 2.9 orders the compounds by Docking Score.

Table 2.8: IFD results for different R-groups and known binders, ordered by IFDScore

Title	Docking Score (kcal/mol)	**IFDScore** (kcal/mol)	Prime Energy (kcal/mol)	XPGScore (kcal/mol)
4-Cl, **tBu**	-11.821	-2519.68	-50150.3	-12.167
P65	-10.888	-2519.64	-50158.8	-11.698
4-Cl, **Cbu**	-12.016	-2519.55	-50144.1	-12.340
4-Cl, **Prop**	-11.881	-2519.15	-50139.0	-12.195
4-Cl, **iPr**	-10.718	-2519.13	-50161.3	-11.064
4-Cl, **Cpe**	-11.093	-2519.05	-50153.1	-11.399
4-Cl, **iBut**	-10.835	-2518.87	-50154.5	-11.150
4-Cl, **Che**	-9.766	-2518.85	-50169.6	-10.368
4-Cl, **Cpr**	-10.596	-2518.70	-50155.2	-10.942
4-Cl, **Eth**	-10.557	-2518.52	-50152.8	-10.882
4-Cl, **Meth**	-11.686	-2517.72	-50114.2	-12.007
DHF	-9.968	-2514.48	-50046.5	-12.154
WR99210	-11.268	-2513.62	-50045.1	-11.364

Table 2.9: IFD results for different R-groups and known binders, ordered by Docking Score

Title	**Docking Score** (kcal/mol)	IFDScore (kcal/mol)	Prime Energy (kcal/mol)	XPGScore (kcal/mol)
4-Cl, **Cbu**	-12.016	-2519.55	-50144.1	-12.340
4-Cl, **Prop**	-11.881	-2519.15	-50139.0	-12.195
4-Cl, **tBu**	-11.821	-2519.68	-50150.3	-12.167
4-Cl, **Meth**	-11.686	-2517.72	-50114.2	-12.007
WR99210	-11.268	-2513.62	-50045.1	-11.364
4-Cl, **Cpe**	-11.093	-2519.05	-50153.1	-11.399
P65	-10.888	-2519.64	-50158.8	-11.698
4-Cl, **iBut**	-10.835	-2518.87	-50154.5	-11.150
4-Cl, **iPr**	-10.718	-2519.13	-50161.3	-11.064
4-Cl, **Cpr**	-10.596	-2518.70	-50155.2	-10.942
4-Cl, **Eth**	-10.557	-2518.52	-50152.8	-10.882
DHF	-9.968	-2514.48	-50046.5	-12.154
4-Cl, **Che**	-9.766	-2518.85	-50169.6	-10.368

Sorting the results by IFDScore and Docking Score reveals different trends. IFDScore considers the XPGScore produced in the final re-docking stage, along with the Prime Energy calculated for the final minimized ligand-receptor complex. Docking Score does not include the Prime Energy term, and instead adjusts the XPGScore by applying the ligand state penalties generated during ligand preparation. When considered together, these two scoring metrics allow for the assessment of the strength of the ligand binding interactions, along with the energy requirements for the protein structure to undergo the specific conformational changes and the energy penalties associated with the particular ligand states.

The IFDScores are very similar for the compounds assessed. The intermediate sized groups, consisting of cyclobutyl, propyl, and isopropyl substituents performed the best, while the small methyl and ethyl substituents yielded the worst IFDScores. This may give an indication of the optimal size of substituents for this position. It can be inferred that the small groups were unable to effectively fill the hydrophobic space in the active site, producing fewer

positive interactions. On the other hand, the larger cyclopentyl and cyclohexyl R-groups were able to maximize favourable hydrophobic interactions due to the flexibility of the nearby residues in accommodating these substituents. However, the energy requirements in adopting these conformations led to lower IFDScores. The intermediate sized groups struck the right balance between maximizing hydrophobic interactions, and requiring minimal energetically straining conformational changes, resulting in the best IFDScores for these compounds.

The Docking Scores show a slightly different trend. The intermediate sized R-groups still perform the best, but the relative positions of some of the other substituents are altered. This gives an indication of the energy penalties acquired by these compounds in attaining the desired protonation states for optimal binding.

With these results in hand, the various poses generated in the IFD experiment were further studied in order to gain additional insight into the binding mode of the compounds assessed. As done previously, the pose of the bound ligand in the protein structure 1J3K was compared to the best pose produced for this ligand from the docking experiments. This was done to ensure that the known binding mode was correctly reproduced by the docking methods used. Figure 2.6 shows the two structures superimposed, with the bound ligand conformation indicated in orange and the best pose generated for the ligand WR99210 from the IFD experiment in blue.

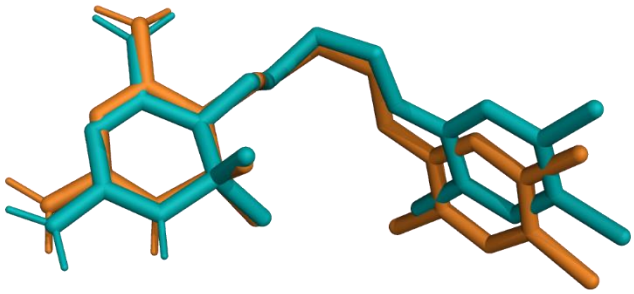


Figure 2.6: Superimposed structures of WR99210 as found in protein structure 1J3K (orange), and the best pose for this ligand obtained in the IFD experiment (blue)

This figure shows a significant overlap in the two poses examined, with a calculated RMSD value of 0.84 Å between the two structures. This result confirms that the known binding mode was reproduced, even when protein flexibility was simulated using IFD. Typically, RMSD values of less than 2 Å are considered acceptable in recreating known binding modes with docking procedures[69].

By combining multiple ligand poses generated from the IFD experiment, the extent of the ligand conformational possibilities in the active site can be seen. Figure 2.7 includes the top five IFD poses of the ligand WR99210.

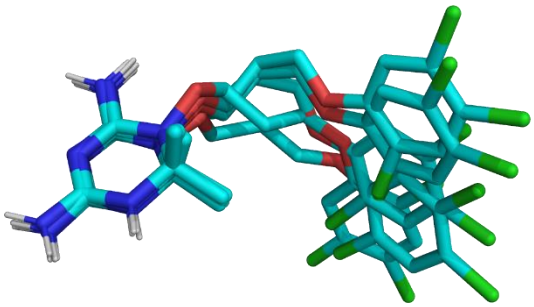


Figure 2.7: Superimposition of top 5 IFD poses of WR99210 docked into active site of 1J3K

The above figure shows that the position of the core ring was highly conserved between poses. This is due to the strong binding interactions to the nearby residues, including the forced docking constraint to Asp54, which kept this portion of the molecule tightly bound. The linker and terminal ring portions display greater variability in their conformations as there were no strong interactions holding them in place.

The top scoring IFD poses for each of the project compounds assessed are combined in figure 2.8. This figure further highlights the conserved binding of the core 2,4-diaminopyrimidine ring, with some variations shown in the positions of the linker and terminal rings. This confirms that all of the project compounds investigated adhered to the known binding mode despite differences in the added R-group.

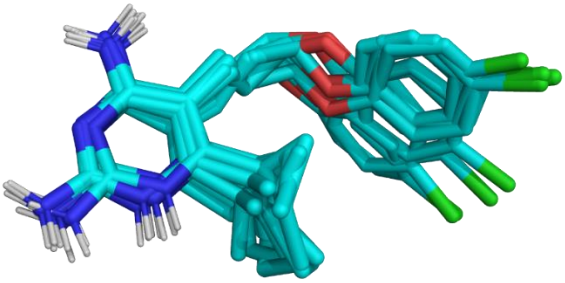


Figure 2.8: Superimposed IFD poses of project compounds with various R groups docked into the active site of 1J3K

The fact that all of the investigated project compounds adhered to the known binding mode, and that they all had very similar IFDScores, suggests that each R-group was able to be accommodated by the hydrophobic pocket. The amino acid residues which comprise the hydrophobic pocket were identified based on their proximity to the dimethyl group of bound ligand WR99210 in protein structure 1J3K. The residues within 5 Å of this group are Met55, Leu46, Trp48, Asp54, Phe58, and Ala16, which form the interior surface of the hydrophobic pocket. The extent of the flexibility of each of these residues in accommodating the various R-groups was determined by calculating RMSD values between the conformations generated in the IFD experiment and the original conformations as found in the crystal structure 1J3K. Root mean square deviation (RMSD) is a quantitative measure of the average distance between the corresponding atoms of two superimposed structures. This metric gives an indication of the similarity between two poses, with higher RMSD values indicating greater conformational changes. RMSD is calculated from two superimposed structures as follows, where N is the number of atom pairs examined and δ is the distance in Angstroms between the compared atoms.

$$RMSD = \frac{1}{N}\sum_{i=1}^{N}\delta_i^2$$

For this analysis, RMSD values were calculated for each of the six residues identified to make up the hydrophobic pocket. The conformations of the residues were extracted from the top IFD poses for each of the assessed project compounds and were superimposed with their original conformations as found in the crystal structure 1J3K. The Superimposition function in Maestro was used to calculate the different RMSD values, all atoms except for the non-polar hydrogens were considered, and the calculation was done without performing prior alignment.

Table 2.10: RMSD values (Å) calculated for residues comprising the hydrophobic pocket by comparing their conformations in the top scoring IFD poses and their original conformations as found in the protein structure 1J3K

R-group	**Met55**	**Phe58**	**Leu46**	**Asp54**	**Trp48**	**Ala16**	**Avg**
Cpe	0.3269	0.5856	1.518	0.9247	0.6099	0.1664	**0.6886**
Che	1.1288	0.6315	0.469	0.3937	0.4978	0.2205	**0.5569**
Prop	2.0929	0.3994	0.1231	0.2526	0.1738	0.1610	**0.5338**
Cbu	1.0313	0.5277	0.7388	0.2691	0.2087	0.0897	**0.4776**
Meth	1.6014	0.3919	0.0816	0.2270	0.0838	0.0901	**0.4126**
tBu	1.1818	0.4468	0.2554	0.2845	0.1598	0.1070	**0.4059**
iPr	0.7734	0.6390	0.3041	0.3253	0.1943	0.0889	**0.3875**
iBut	0.6435	0.5682	0.3820	0.3488	0.2437	0.1267	**0.3855**
Cpr	1.0509	0.4701	0.2099	0.2159	0.1889	0.0941	**0.3716**
Eth	1.0213	0.4056	0.1512	0.2838	0.1682	0.0873	**0.3529**
Avg	**1.0824**	**0.5095**	**0.4101**	**0.3518**	**0.2477**	**0.1181**	

Table 2.10 shows the results of the RMSD analysis for the amino acid residues that constitute the hydrophobic pocket. The table includes the specific RMSD values for the selected residues from the IFD experiments, calculated for each of the assessed project compounds. The average RMSD values are included for each residue and each docked ligand. These values provide an indication of the extent to which each ligand caused conformational changes in the hydrophobic pocket residues, and the extent to which each of these residues conformed to accommodate the various ligands. The table is ordered from top to bottom by decreasing average conformational changes induced by each ligand, and from left to right by decreasing average conformational changes displayed by each residue.

The results presented here fall within expectations, with the larger R-groups inducing the greatest conformational changes in the hydrophobic pocket residues, as the side chains rotated out of the way to accommodate the increased steric bulk. In contrast, the smaller R-groups induced lower conformational changes on average, since the original positions of the hydrophobic pocket residues were already orientated to accommodate groups of this size. Notably, a large RMSD value is observed for residue Met55 in the docking experiment of the methyl R-group. This occurred because the side chain of this residue rotated inwards towards the small methyl group in order to maximize hydrophobic interactions.

Looking at the RMSD values for each residue, it can be seen that Met55 exhibited the greatest conformational changes on average, followed by Phe58, Leu46, Asp54, and Trp48, with Ala16 experiencing minimal conformational changes in accommodating each ligand.

The various conformational possibilities of residue Met55 are displayed in figure 2.9, with the docked cyclopentyl ligand included and the original conformation of this residue extracted from protein structure 1J3K highlighted in orange. The specific positions of the Met55 side chain obtained from the docking experiments of the methyl and cyclohexyl R-groups are also included.

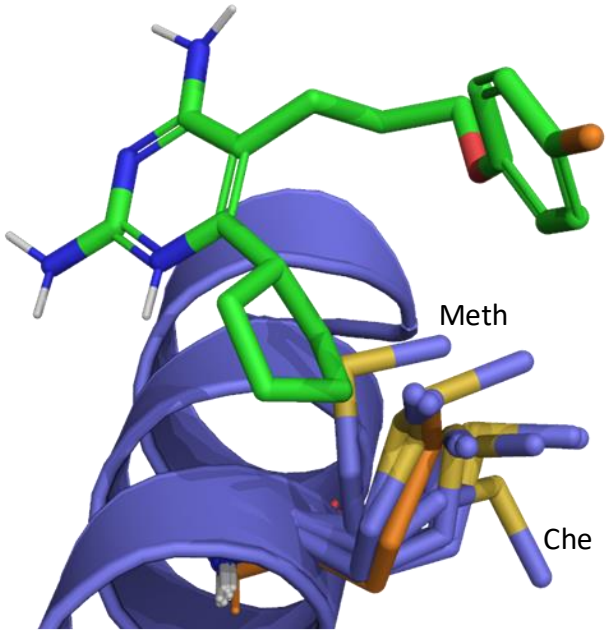


Figure 2.9: Conformational possibilities of residue Met55, with the original conformation highlighted in orange, and the docked cyclopentyl ligand included

The above figure shows the high variability in the position of residue Met55 in accommodating the incoming ligands. It can be seen that for the cyclohexyl R-group, the side chain of Met55 rotated out of the way to accommodate the increased steric bulk of this group, while for the methyl R-group, it moved inwards to maximize hydrophobic interactions.

While these results effectively showcase the conformational possibilities of the hydrophobic pocket residues in accommodating R-groups of different sizes, the analysis is limited by the fact that the IFD jobs output several protein-ligand poses for each docked ligand. Only the top scoring IFD pose for each ligand was examined; however, this is not necessarily the pose in which the greatest conformational changes occur. Nonetheless, the findings confirm that all of the R-groups assessed were able to fit within the hydrophobic pocket through conformational changes in the nearby residues, maximizing positive hydrophobic interactions and minimizing steric clashes.

The effects of the conformational changes on the overall shape of the hydrophobic pocket are displayed in figure 2.10.

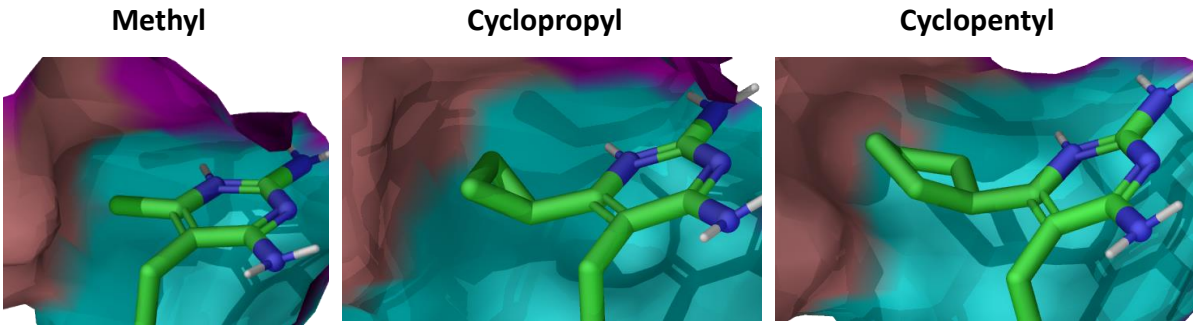


Figure 2.10: Project compounds docked into active site of 1J3K with IFD, the interior surface of hydrophobic pocket shown

The figure above includes the docked poses of the methyl, cyclopropyl, and cyclohexyl project compounds from the IFD experiment, with the interior surface of the hydrophobic pocket shown. It can be seen that the shape of the hydrophobic pocket conformed to fit the specific R-group present, maximizing the hydrophobic interactions and minimizing unfavourable steric clashes.

Overall, from the analysis of the various scoring metrics produced by the IFD experiment and the examination of the resulting poses, it can be concluded that all of the assessed R-groups were accommodated in the hydrophobic pocket while adhering to the conserved binding mode. The amino acids comprising this pocket have been identified, and the extent of their conformational changes to accommodate the various ligands has been quantified by RMSD calculations. Finally, images showing how the interior surface of the hydrophobic pocket conformed to fit the various R-groups have been presented.

2.5 Assessment of X-substituents around the terminal ring

Further IFD experiments were conducted on a new set of project analogues with the aim of identifying and optimizing interactions involving substituents around the terminal ring. From the earlier docking experiments, it has been concluded that while the core 2,4-diaminopyrimidine ring is tightly held in the active site of the target enzyme, the linker and terminal ring portions display much greater variability in their binding poses. This is because of the inherent flexibility of the four-atom linker combined with the fact that this portion of the active site is more open, allowing for varied ligand conformations to be adopted.

Figure 2.11 includes surface images of different PDB structures of *Pf*DHFR with the bound ligands displayed. Included are the wild-type structure 4DPD, with the natural enzyme substrate DHF bound, and the two quadruple mutant structures 3QG2 and 1J3K, with their respective ligands, pyrimethamine and WR99210 included.

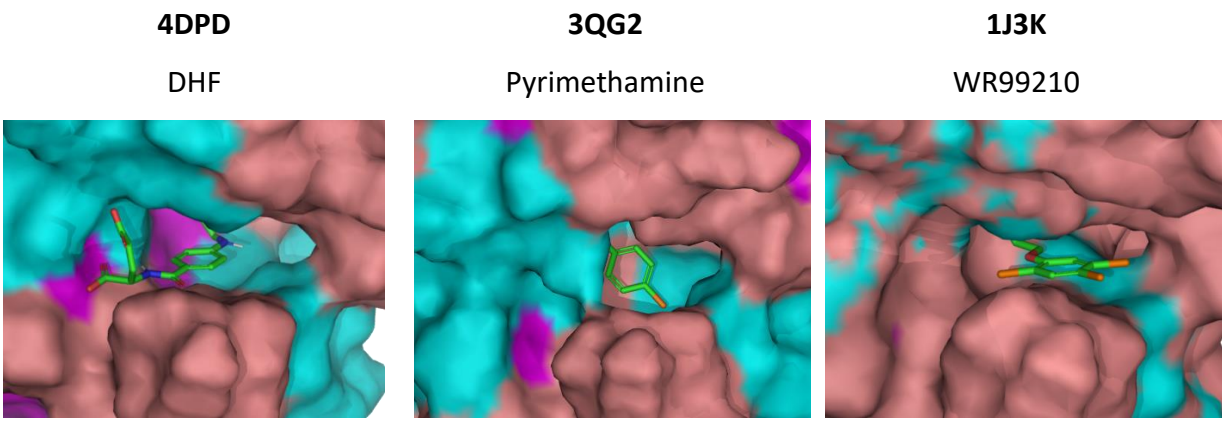


Figure 2.11: Surface images of *Pf*DHFR structures with bound ligands shown in the active site

The images presented above show how the ligand core functionality is buried inside the active site in each structure displayed. Examining structure 3QG2, it can be seen that the ligand pyrimethamine is entirely contained within the active site. The larger ligands DHF and WR99210, however, cause a conformational change in the active site, opening it up and positioning the terminal rings on the exterior surface of the protein structure. The amino acid residues on the exterior surface of proteins are typically highly polar, and as such have the potential to engage in polar interactions with the ligand molecules. Figure 2.12 shows the 2D-interaction diagram for DHF bound in the active site of wild-type protein structure 4DPD. From this figure some elements of the core binding mode can be observed, namely H-bonds to residues Asp54 and Cys15. Additional H-bonds and an electrostatic interaction to the charged residue Arg122 on the exterior surface of the protein structure are also evident.

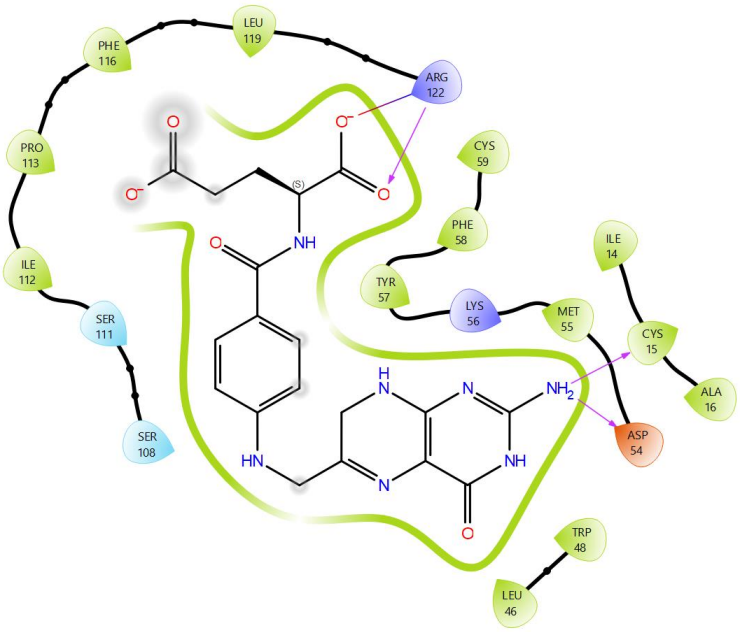
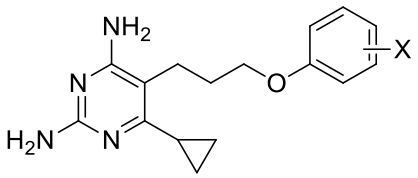


Figure 2.12: 2D interaction diagram of bound ligand DHF in active site of wild-type structure 4DPD

In order to identify potential new interactions in the surface exposed region of the active site, various common aromatic groups were investigated as substituents around the terminal ring. Because the residues exposed on the surface of proteins are often highly variable in their

conformations, protein flexibility was considered by using IFD to maximize the possibility of identifying new interactions. Functional groups capable of accepting and donating H-bonds were assessed, in addition to cationic and anionic moieties for potential electrostatic interactions to the nearby charged residues.

Compounds with the general structure presented in figure 2.13 were assessed, with the selected aromatic substitutions also included. A cyclopropyl R-groups was included in all structures for this analysis, so that any differences in the docking parameters produced were due to only the substitution of the terminal ring.



X = Cl, F, OMe, Me, OH, COOH, NH_2, NO_2

Figure 2.13: General structure of compounds assessed for identification and optimization of new interactions in exposed regions of the active site

A library of analogues based on the general structure shown above was enumerated, with the defined substituents added in the 2, 3, or 4-position of the terminal ring for a total of 24 compounds generated. The cyclopropyl group in the 6-position of the pyrimidine ring was selected as this group is very similar in size to the dimethyl group of WR99210, and it showed good binding affinity in the earlier run IFD experiments. The halogen substituents were included as they are very common aromatic substituents for drug molecules, and are present in many antifolate compounds, for example WR99210 and pyrimethamine. The 4-chloro group also serves as a basis for comparison with the earlier IFD experiments conducted on the various R-groups. The nitro and methyl groups alter the electronic nature of the terminal ring and also introduce steric constraints which may be beneficial to ligand binding. The carboxylic acid, hydroxyl, and amino groups have the capacity to both accept and donate H-bonds, with the methoxy group representing an additional H-bond acceptor. Finally, under physiological

conditions both the carboxylic acid and amino substituents can become charged, with the former losing a proton to form the negatively charged carboxylate ion, and the latter gaining a proton to form the positively charged conjugate acid of the amino group. The 24 ligand structures were put through ligand preparation to generate the possible ionization states. Because it has already been identified as the most active state, only ligands with protonation of the 1-N of the core 2,4-diaminopyrimidine ring were carried forward. For the carboxylic acid and amino substituted compounds, both the neutral and charged forms of these groups were included in the subsequent analysis, bringing the total number of ligands assessed to 30.

Docking of the generated ligands was done using IFD, with the same settings applied as for the previous IFD analysis (Appendix. A 1.6). The standard sampling protocol was selected, and the receptor grid was generated from the bound ligand WR99210 in protein structure 1J3K. Docking constraints were added to residue Asp54, and the GlideXP algorithm was selected for the final re-docking stage.

The IFD results are presented in tables 2.11 and 2.12. Each of the included energy parameters, namely XPGScore, Docking Score, IFDScore, and Prime Energy are listed separately and are ordered by the functional group assessed, and its position around the terminal ring. Conditional formatting has been applied to the cells, highlighting the top scoring compounds (most negative values) in green. All parameters are reported in units of kcal / mol.

Table 2.11: XPGScores and Docking Scores of project compounds from IFD ordered by X-group with conditional formatting of cells applied

XPGScore (kcal/mol)			
X-Group	2	3	4
Cl	-11.519	-10.816	-10.942
F	-10.879	-11.245	-10.542
OMe	-11.003	-10.361	-9.797
Me	-12.026	-11.027	-10.821
OH	-12.212	-10.273	-10.077
NH_2	-11.673	-10.651	-9.804
NH_3^+	-9.724	-10.091	-11.683
COOH	-10.527	-12.101	-11.779
COO^-	-14.058	-15.061	-15.237
NO_2	-10.487	-10.774	-9.921

Docking Score (kcal/mol)			
X-Group	2	3	4
Cl	-11.158	-10.457	-10.596
F	-10.52	-10.886	-10.184
OMe	-10.647	-10.004	-9.440
Me	-11.67	-10.671	-10.465
OH	-11.865	-9.927	-9.729
NH_2	-11.326	-10.304	-9.456
NH_3^+	-7.234	-5.174	-9.545
COOH	-9.324	-10.917	-10.596
COO^-	-13.711	-14.714	-14.889
NO_2	-10.124	-10.412	-9.557

Table 2.12: IFDScores and Prime Energy values of project compounds from IFD ordered by X-group with conditional formatting of cells applied

IFDScore (kcal/mol)			
X-Group	2	3	4
Cl	-2518.70	-2518.25	-2518.70
F	-2517.38	-2518.04	-2517.99
OMe	-2518.84	-2517.99	-2517.36
Me	-2518.05	-2519.04	-2517.74
OH	-2520.15	-2518.25	-2517.87
NH_2	-2519.19	-2518.51	-2516.94
NH_3^+	-2517.93	-2519.08	-2520.13
COOH	-2518.68	-2520.92	-2519.64
COO^-	-2521.48	-2524.10	-2523.97
NO_2	-2517.12	-2517.74	-2517.19

Prime Energy (kcal/mol)			
X-Group	2	3	4
Cl	-50143.7	-50148.6	-50155.2
F	-50130.1	-50136	-50148.9
OMe	-50156.7	-50152.7	-50151.3
Me	-50120.4	-50160.3	-50138.3
OH	-50158.8	-50159.5	-50155.8
NH_2	-50150.4	-50157.1	-50142.7
NH_3^+	-50164.2	-50179.7	-50168.9
COOH	-50163.1	-50176.3	-50157.3
COO^-	-50148.5	-50180.7	-50174.7
NO_2	-50132.6	-50139.4	-50145.4

Looking at the XPGScores and IFDScores, the ligands containing the carboxylate anion performed the best, with this group in the 3 or 4-positions of the terminal ring exhibiting the strongest binding interactions. This form of the ligands performed better than the neutral

carboxylic acid form, suggesting that the excellent binding affinities were due to electrostatic interactions. Aside from these groups, the other substituents performed similarly. The Docking Scores, which take the energy penalties associated with the specified ligand states into account, show that the carboxylate anion is favourable at physiological pH, while protonation of the amino group in the 2 or 3-positions of the terminal ring is not.

In order to gain a complete understanding of the various interactions between the docked ligands and the active site residues, the top poses for each of the assessed compounds were evaluated. In each case the required core binding mode was adhered to, with H-bonds formed to residues Asp54, Ile14, and Leu164 and an additional electrostatic interaction to Asp54.

Considering contacts involving the terminal ring substituents, several new conserved interactions were seen to some of the nearby residues. These residues and their potential interactions to the functional groups around the terminal ring are summarized in figure 2.14. This figure displays the residues' side chains in black and the backbone carbonyl groups in red, along with the identified pi – pi stacking, H-bonding, and electrostatic interactions to the docked ligands. The specific nature of these interactions is described in the text below.

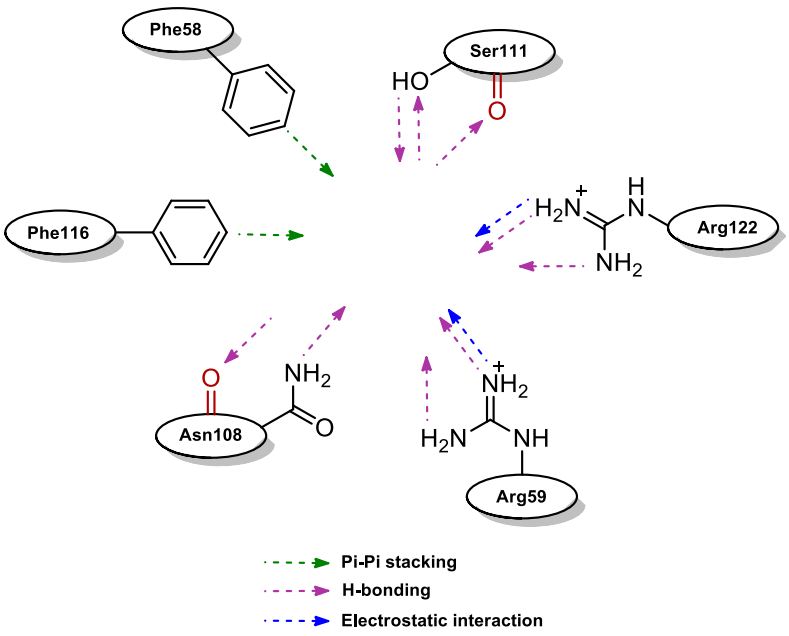


Figure 2.14: Pharmacophore displaying amino acid residues near to the terminal ring in the active site, with the potential binding interactions indicated

Strong interactions to residues Arg59 and Arg122 were seen in many of the IFD poses for the docked project molecules. These two residues are exposed on the exterior surface of the protein active site and possess positively charged guanidinium side chains. Ligands with appropriate substituents around the terminal ring accepted multiple H-bonds from the Arg side chains, and the ligands containing carboxylate anions engaged in electrostatic interactions. These residues also showed the ability to accept halogen bonds from the chloro and fluoro substituted analogues. Ligands were seen to further donate H-bonds to the backbone carbonyl and hydroxyl side chain of residue Ser111, while also accepting H-bonds from the hydroxyl side chain. Some poses showed ligands donating additional H-bonds to the backbone of Asn108 and the oxygen in the linker of ligand molecules could accept H-bonds from the primary amide side chain of this residue. Pi-pi stacking interactions between the terminal ring and the aromatic residues Phe58 and Phe116 were also observed.

Figure 2.15 shows the 2D interaction diagram of the 3-Me analogue, highlighting the pi – pi stacking interactions to the nearby aromatic residues and the H-bond formed between the linker oxygen and the side chain of residue Asn108.

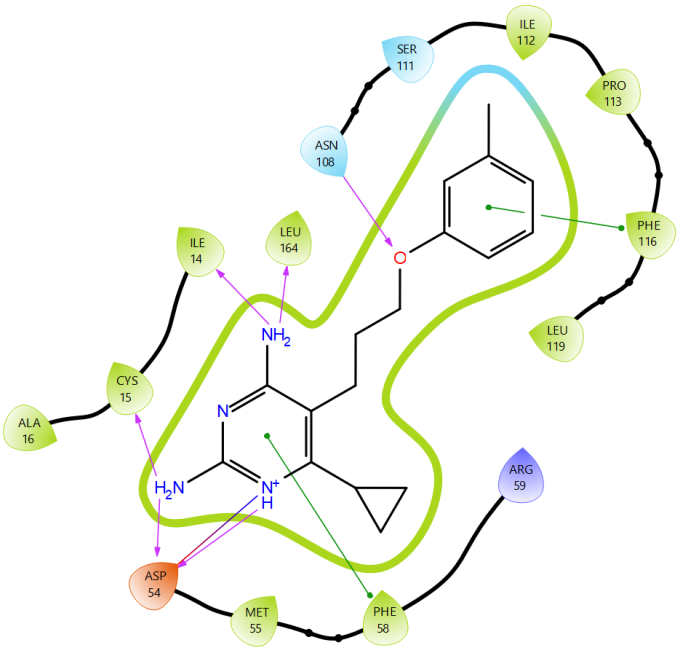


Figure 2.15: 2D interaction diagram of 3-Me analogue docked into the active site of 1J3K with IFD

It is important to note that not all of these interactions were seen in every pose. Different ligand binding modes favour some interactions over others and only by considering multiple good scoring poses for each ligand could the full range of potential interactions be seen.

Figures 2.16 and 2.17 show some of the project compounds docked into the active site of 1J3K. The surfaces were modified by removing some residues to get a better view into the active site. These figures served to highlight the different amino acid residues which can engage in interactions with the substituents around the terminal ring.

The amino acid residues in figures 2.16 and 2.17 are colour coded according to table 2.13, which also summarizes the relevant potential interactions to the ligand molecules.

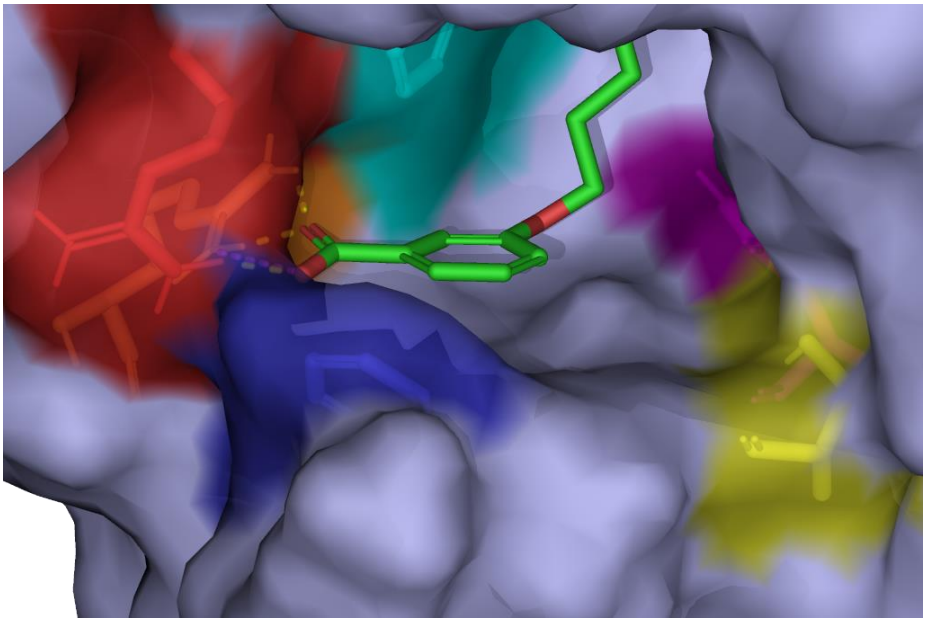


Figure 2.16: Modified surface diagram of 3-COO^- ligand docked into active site of 1J3K with IFD, highlighting interactions to the terminal ring. Relevant amino acid residues are colour coded and binding interactions to Arg59 and Arg122 shown

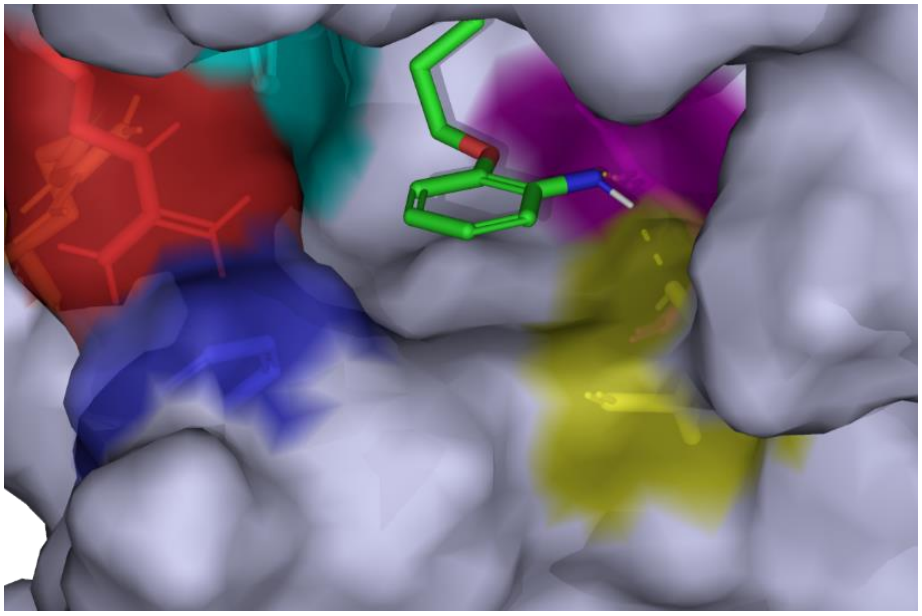


Figure 2.17: Modified surface diagram of 2-NH_2 ligand docked into active site of 1J3K with IFD, highlighting interactions to the terminal ring. Relevant amino acid residues are colour coded and binding interactions to Ser111 and Asn108 shown

Table 2.13: Colour key and potential interactions of amino acid residues to docked ligand molecules

Residue:	**Colour:**	**Possible interactions:**
Phe58	Light blue	Pi-pi stacking
Arg59	Red	H-bond donor Electrostatic bond
Asn108	Purple	H-bond donor H-bond acceptor
Ser111	Yellow	H-bond donor H-bond acceptor
Phe116	Dark blue	Pi-pi stacking
Arg122	Orange	H-bond donor Electrostatic bond

The surface images presented in figures 2.16 and 2.17 show the relative positions of the amino acids which interacted with the substituents around the terminal ring of the docked ligand molecules. It can be seen that the two arginine residues are very close together, on the left-hand side of the active site as shown in figures 2.16 and 2.17 while residues Ser111 and Asn108 are on the right-hand side. Both of these sites were accessible to the ligand molecules, and interactions were formed based on how the flexible linker orientated the substituents on the terminal ring. The ligands adopted poses which formed the most favourable interactions to either of these two groups of residues.

The corresponding 2D interaction diagrams for the ligand poses are included in figures 2.18 and 2.19.

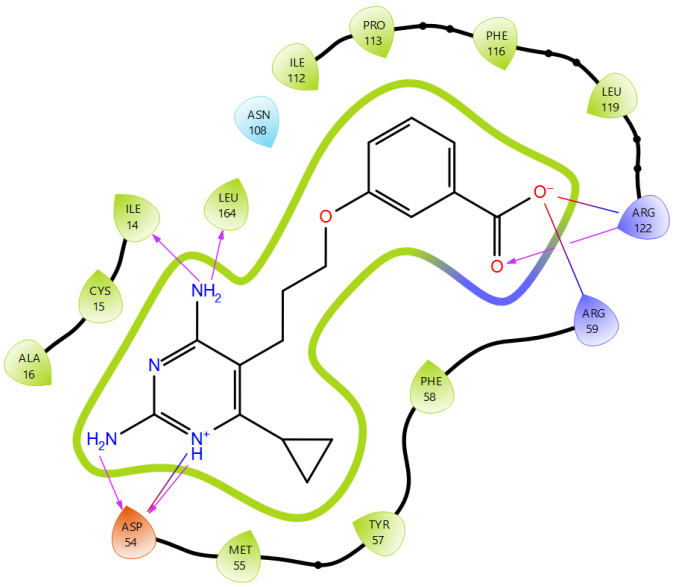


Figure 2.18: 2D interaction diagram of 3-COO^- ligand docked into the active site of 1J3K with IFD

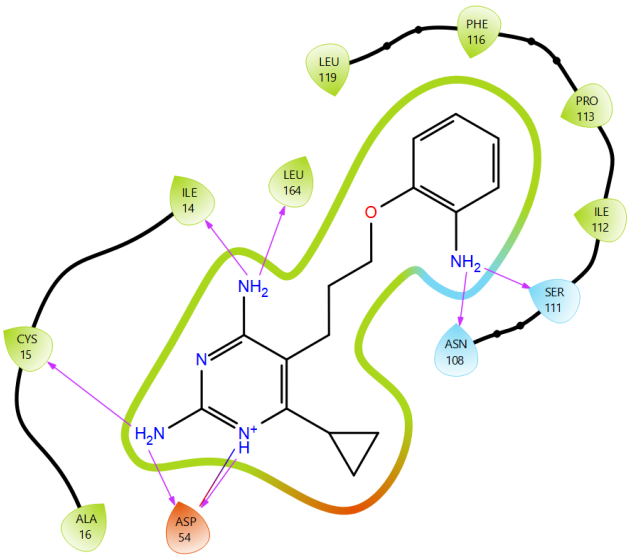


Figure 2.19: 2D interaction diagram of 2-NH_2 ligand docked into the active site of 1J3K with IFD

Figure 2.20 includes all of the top ligand poses from the IFD experiment. From this figure it can be seen that the binding mode of the core 2,4-diaminopyrimidine ring was highly conserved, and that the position of the terminal ring varied between the ligand poses based on which residues interactions were formed to with the terminal ring substituents.

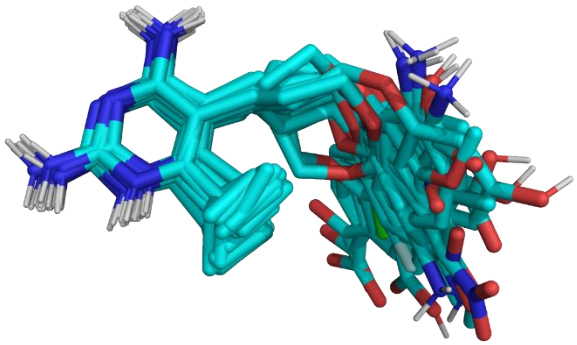


Figure 2.20: Superimposed IFD poses of ligands docked into the active site of 1J3K

In summary, a diverse library of ligands containing various functional groups substituted around the terminal ring were assessed using IFD. The compounds were evaluated based on

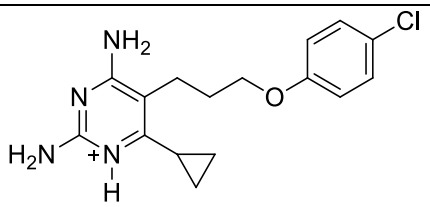
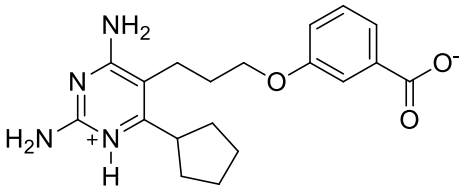
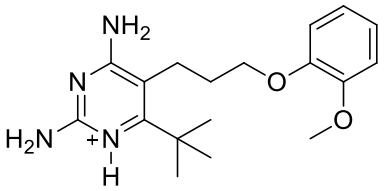
the different energy parameters produced by IFD, and several new strong binding interactions to residues near the terminal ring were identified. Ligands containing a carboxylate substitution performed the best, as this group enabled strong H-bonding and electrostatic interactions to residues Arg59 and Arg122. The other compounds assessed performed similarly in the docking experiments, showing that a wide range of functional groups which engage in different interactions to active site residues were all tolerated.

2.6 Molecular dynamics simulations

Molecular dynamics (MD) simulations were run on some project compounds, and the bound ligand of protein structure 1J3K. The simulations were initialized from the poses generated from the IFD experiments. Some new ligands were enumerated for the MD simulations, containing R and X-groups not combined together in the IFD experiments. For these ligands, additional IFD jobs were run to generate the starting poses for the MD simulations.

WR99210 was included in the MD simulations as a basis of comparison for the other ligands and to further validate the binding interactions established thus far. The project compounds assessed include a variety of R-groups and X-substituents in order to investigate the different interactions identified previously. R-groups include cyclopentyl, cyclopropyl, and *tert*-butyl. They were selected based on their IFD scores and offer a range of sizes. The 4-chloro, 3-carboxylate, and 2-methoxy groups on the terminal ring were included to assess the potential H-bonding, electrostatic, and pi-pi stacking interactions to the nearby amino acid residues. The project compounds chosen for MD analysis are indicated in table 2.14.

Table 2.14: Structures of project compounds subjected to MD analysis

R-group:	**X-group:**	**Structure:**
Cyclopropyl	4-Cl	
Cyclopentyl	3-COO^-	
tert-Butyl	2-OMe	

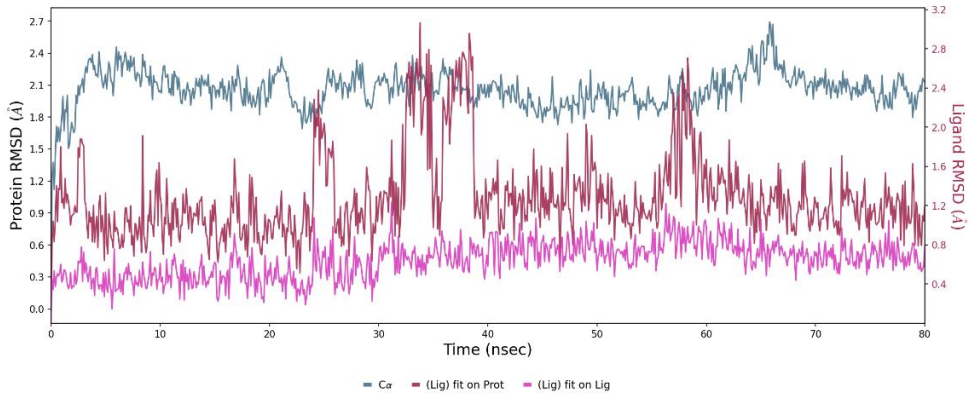
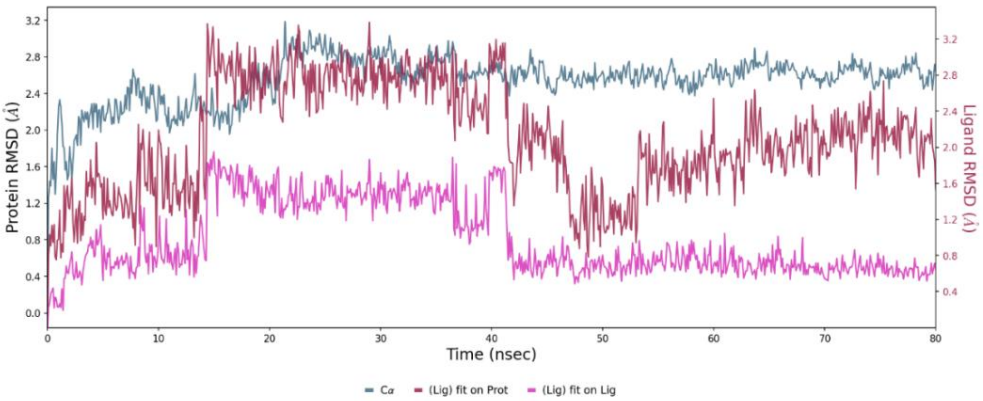
The system builder tool in Maestro (Appendix. A 1.7) was used to generate the input files for the subsequent MD simulations from the top poses generated with IFD. The SPC solvent model was selected and the OPLS_2005 force field[70] used. Prior to this, the other protein chains in structure 1J3K were deleted, leaving only chain A with the docked ligand and the bound NADHP cofactor in order to save computational resources. The MD simulations (Appendix. A 1.8) were run over 80ns with the NPT ensemble class selected.

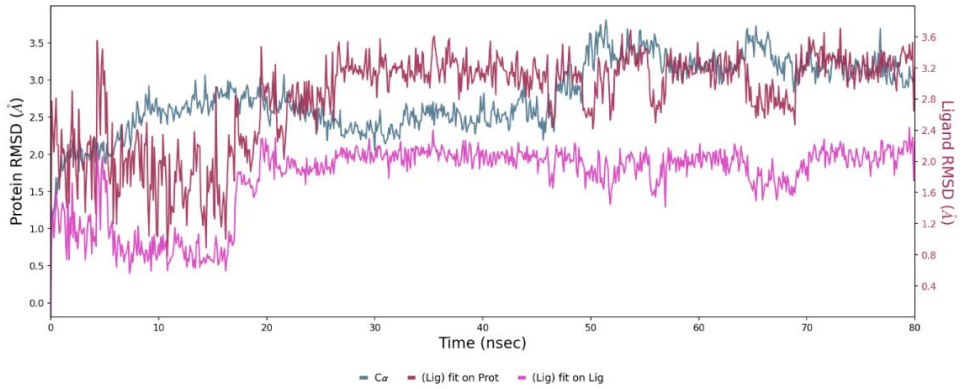
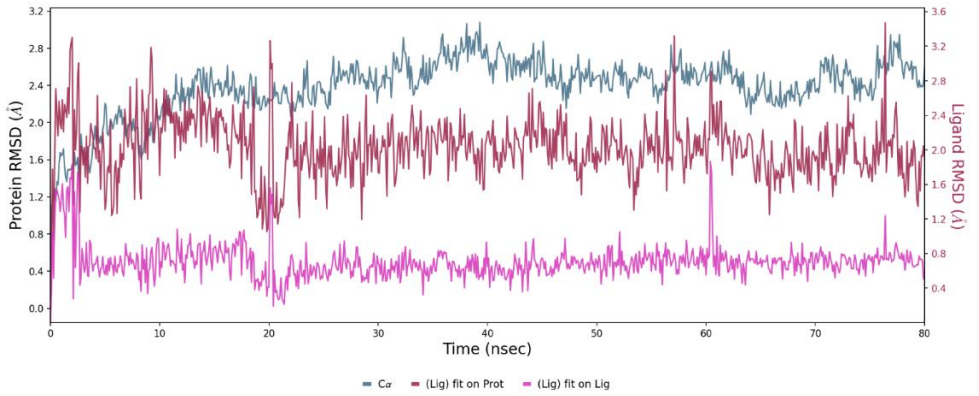
RMSD plots generated from the four MD simulations are presented in table 2.15. These figures show the variations in the protein and ligand positions over the course of the simulation time. Protein RMSD is shown in blue and is calculated based on the initial protein conformation. The ligand RMSD can be calculated based on the initial conformation of the protein, which is shown in red, or the initial conformation of the ligand, shown in pink.

For each of the simulations, the protein rapidly adopted a stable conformation, with steady RMSD values within 0.5 Å for most of the simulation time. The ligand RMSD plots show that the ligands occupied conserved binding modes during the simulation, with RMSD values remaining largely stable (within 1 Å) after an optimal conformation was found. These results confirm that the ligands were tightly bound in the active site, without any significant protein

structural changes occurring. In the case of the Cpr, 4-Cl ligand, some more significant variations can be seen in the ligand RMSD plots between 40 to 55 ns of the simulation. Close inspection of the ligand binding pose throughout the simulation revealed that this was due to the orientation of the terminal ring undergoing some minor conformational changes, however, after 55 ns it maintained a single stable binding pose for the remaining simulation time.

Table 2.15: RMSD plots for molecular dynamics simulations of WR99210 and three project compounds

Ligand:	**RMSD plot:**
WR99210	
Cpr, 4-Cl	

Cpe, 3-COO$^-$	
Tbu, 2-OMe	

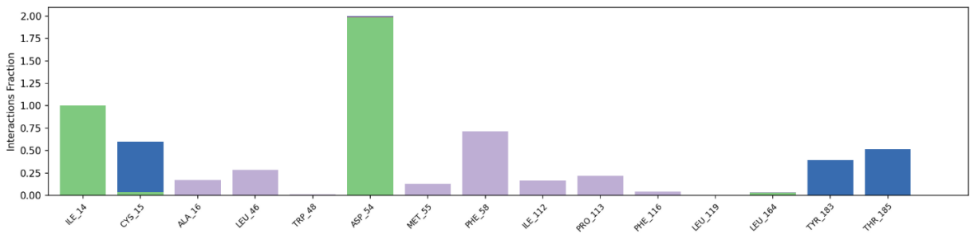
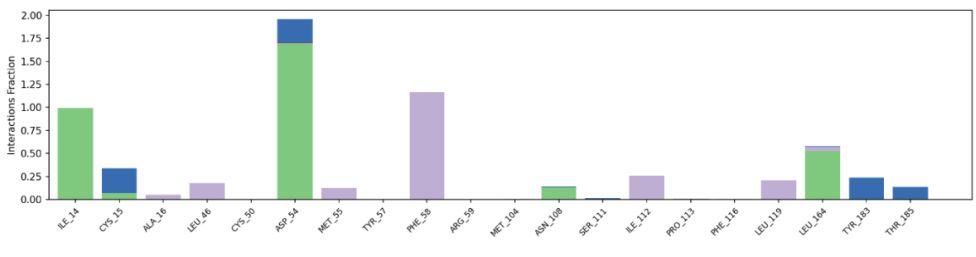
The interactions between the ligands and the amino acid residues in the active site occurring during the simulation are summarized by the protein – ligand contacts diagrams included in table 2.16. These diagrams display the residues that engaged in interactions with the ligand molecules. The type of interaction is shown by colour, and the proportion of simulation time the interactions existed for are indicated on the y-axis.

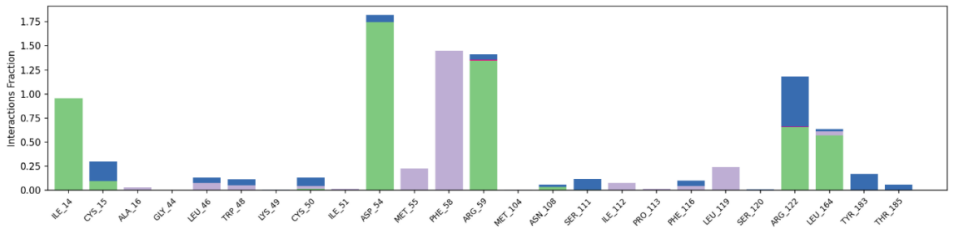
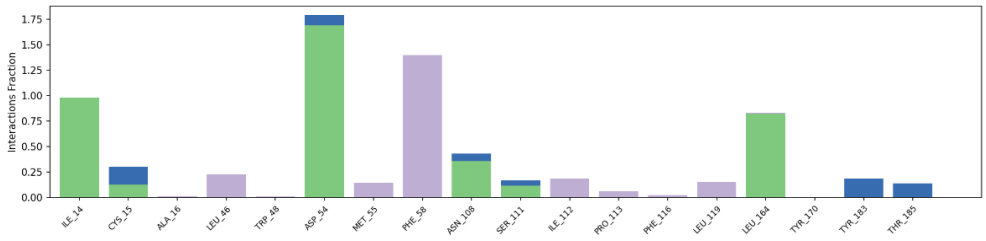
In each of the simulations, the core binding mode of the ligands was adhered to throughout. H-bonds to residues Ile14 and Asp54 can be seen for nearly the entire simulation time. The strong electrostatic interaction to Asp54 occurred alongside one of the H-bonds to this residue, so all of these interactions are included under H-bonding in the protein – ligand contacts diagrams. Significant H-bonding to Cys15 is also evident, largely through bridging water molecules. The three project compounds exhibited H-bonds to residue Leu164, but this interaction is not prevalent for WR99210. Phe58 shows a large amount of hydrophobic contacts with the ligands, which include pi-pi stacking, pi-cation, and other non-specific

lipophilic interactions. Other residues which displayed some hydrophobic contacts to the ligands include Met55, Leu46, and Ala16, which comprise the hydrophobic pocket along with Phe58. Both Arg59 and Arg122 engaged in H-bonds with the carboxylate side of the Cpe, 3-COO^- ligand for a large portion of the simulation time. The bonds formed to Arg122 were partially assisted by bridging water molecules. The potential for H-bonds to residues Asn108 and Ser111 is shown for the tBu, 2-OMe ligand, although these interactions did not occur for the majority of the simulation time.

The MD simulations conducted on WR99210 and the three project compounds confirmed the identified core binding interactions. In addition, the ability of R-groups of different sizes to be accommodated in the active site was shown, and hydrophobic interactions to the residues which make up the hydrophobic pocket in this portion of the active site were seen. All of the binding interactions involving the substituents around the terminal ring identified with IFD were adhered to in the MD simulations, showing the potential for increasing the affinity of drug molecules for the active site by adding specific groups in this position.

Table 2.16: Protein ligand contacts diagrams

Ligand:	**Contact diagram:**
WR99210	
Cpr, 4-Cl	

Cpe, 3-COO$^-$	
Tbu, 2-OMe	
Key	

2.7 *In silico* estimation of physiochemical parameters

The QikProp tool in Maestro was used to estimate various physiochemical parameters of the compounds generated in sections 2.4 and 2.5 (Appendix. A 1.9), to provide insights into their potential behaviour in biological systems (pharmacokinetics). The carboxylic acid and amino X-substituents from section 2.5 were assessed in their neutral forms.

Sixteen relevant metrics for each analogue are included in Appendix B. The calculated parameters are defined and discussed, and the analogues' drug-likeness is reviewed by comparison to expected values of known drug molecules. A summary of these results is given below, highlighting the key findings and any potential shortcomings of the compounds assessed.

Overall, the analogues generated in sections 2.4 and 2.5 displayed good predicted physiochemical properties. Many key parameters are presented in appendix B, including the molecular weight of the compounds, the number of H-bond donor and H-bond acceptor

groups for each molecule, the number of likely metabolites, and the extent of their predicted oral absorptivity. Additional metrics regarding the solubility and membrane permeability of the compounds are also included. The compounds' adherence to two general criteria for drug-likeness, namely Lipinski's rule of five and Jorgensen's rule of three, are further discussed.

No potential issues were identified for any of the compounds generated in section 2.4. These analogues all contain a 4-chloro substituent around the terminal ring and differ in the nature of the aliphatic hydrophobic R-group attached to the 6-position of the 2,4-diaminopyrimidine core. For this set of analogues, all of the evaluated metrics fell well within the standard ranges defined by known drug molecules. This suggests that any of the R-groups assessed would be suitable for further synthesis and biological evaluation.

The analogues enumerated in section 2.5 contain an array of different substituents around the terminal ring, all with a conserved cyclopropyl R-group in the 6-position of the 2,4-diaminopyrimidine core. The different groups assessed vary greatly in their polarity. Highly polar substituents include hydroxyl, amino, carboxyl, and nitro groups. While the less polar substituents include halogen, methyl, and methoxy groups. All of the less polar analogues performed well in this analysis, with each adhering to the standard range defined by known drugs in each of the assessed criteria. The more polar analogues, however, showed some potential shortcomings in their physiochemical profiles. As they are highly polar, they may experience difficulties in crossing cell membranes to enter the systemic circulation. This is highlighted by the poor results obtained in the predicted caco-2 cell permeability and oral absorptivity parameters.

The relevant physicochemical properties discussed above were also calculated for cycloguanil and pyrimethamine, the results for which are likewise included in the appendix. Both showed good predicted parameters, however, it should be noted that cycloguanil is administered in prodrug form, as proguanil, and undergoes rearrangement into the active compound after entering the systemic circulation.

2.8 Concluding remarks and recommendations for synthesis

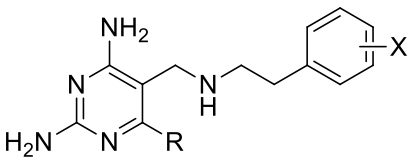
In this section an extensive molecular modelling campaign was conducted with the aim of identifying the conserved binding mode of *Pf*DHFR inhibitors and evaluating how some project compounds for potential synthesis and biological assessment performed under various computational analyses. Three different protein crystal structures of the quadruple mutant *Pf*DHFR enzyme bearing N51I/C59R/S108N/I164L substitutions were evaluated in section 2.3. From this analysis, the binding mode of the co-crystalized ligands was identified and the structure 1J3K selected for subsequent modelling work. Induced fit docking was used for the prediction of binding ability of project compounds with various substitutions in different positions of the general structure given in figure 2.1. It was found that all of the aliphatic hydrophobic substituents assessed in the 6-position of the core 2,4-diaminopyrimidine ring were tolerated for substrate binding. Groups of varying sizes were included, and the results presented in section 2.4 show how the amino acid residues in this region of the active site conformed to optimally accommodate each R-group. Section 2.5 describes the Induced Fit Docking results obtained from analogues bearing different substitutions around the terminal ring. Some new interactions to amino acid residues in this region of the active site were identified, in particular the potential for some groups to engage in H-bonding and electrostatic interactions to the charged residues Arg59 and Arg122, among others. Additional molecular dynamics simulations were conducted on some of the project compounds. These results are discussed in section 2.6 and serve to confirm the identified binding mode of the known inhibitors, and to further validate all of the other interactions identified in sections 2.4 and 2.5. Finally, various relevant physiochemical parameters for each analogue were predicted using QikProp and are discussed in section 2.7. Overall, the physiochemical predictions showed good results for the evaluated analogues, suggesting that they should exhibit suitable pharmacokinetic properties for oral administration. The only compounds for which any issues were identified were those containing highly polar substituents around the terminal ring. Analogues with hydroxyl, amino, carboxyl, and nitro groups may be too polar to effectively diffuse across cell membranes, potentially leading to low oral absorptivity.

Based on the results obtained in the IFD, MD, and QikProp experiments some recommendations can be made regarding the choice of compounds to carry forwards for synthesis and biological evaluation. While all of the R-groups assessed showed good docking

results and favourable physiochemical properties, the intermediate sized R-groups such as *tert*-butyl, cyclopropyl, and isopropyl performed the best. These groups should therefore be included for synthesis, along with some larger R-groups such as cyclopentyl or cyclohexyl to validate the claim that the amino acid residues in this region of the active site are flexible enough to accommodate the increased steric bulk. In terms of substituents around the terminal ring, the carboxylate group performed the best in the docking experiments. This group, however, along with some of the other highly polar substituents, showed poor predicted physiochemical properties. The less polar halogen and methoxy substituents are therefore recommended for synthesis as these groups still performed well in the docking experiments and showed the potential to engage in some of the new identified interactions while still exhibiting good predicted physicochemical properties.

2.9 Additional molecular modelling experiments

Some additional related analogues were also investigated for activity against drug-resistant strains of *P. falciparum* in our laboratories. The synthesis and biological results for these compounds is not included in the present work, however, some molecular modelling studies were conducted on this range of compounds, and were included in a publication[71]. The general structure of the compounds investigated in this publication is described in figure 2.21.



X = 2, 3, or 4 - Cl, F, or OMe
R = Phenyl, Cyclohexyl, Cyclopropyl

Figure 2.21: General structure of compounds investigated by K. Somandi et al, 2022[71]

The general structure of the compounds investigated by a co-worker in our laboratories is very similar to the compounds investigated in this project. The compounds differ only in the four-

atom linker; where the structures for the current work have an oxygen in the 4-position of the linker, the analogues investigated in this publication have an NH group in the 2-position of the linker. The same groups were investigated as X- and R-substituents around the general structure, included in figure 2.21. For the compounds synthesized for the publication, a phenyl R-group was also included.

Similar molecular modelling experiments as described in this chapter were conducted on the nitrogen containing compounds, including induced fit docking and molecular dynamics experiments. The results obtained for the nitrogen compounds were very similar to the results for the oxygen compounds. Both sets of analogues exhibited similar docking scores and adhered to the same conserved binding mode. The full discussion of the results obtained can be found in the publication: Molecular Docking Studies, Synthesis and Biological Evaluation of Substituted Pyrimidine-2,4-diamines as Inhibitors of Plasmodium falciparum Dihydrofolate Reductase. **https://doi.org/10.1002/cmdc.202200418**

Chapter 3 – Results and discussion part A: Synthesis and biological evaluations of first-generation analogues

3.1 Synthesis of flexible pyrimethamine analogues

In this section the synthesis and characterization of flexible pyrimethamine analogues is described. An existing method for the synthesis of these types of compounds had already been established in our laboratories and was adapted in this work to form a library of compounds with functional groups selected based on the results of the previously conducted molecular modelling campaign. The final compounds are comprised of three key structural elements, as described in chapter 2. These are indicated in figure 3.1, along with the R- and X-groups included in the synthesis.

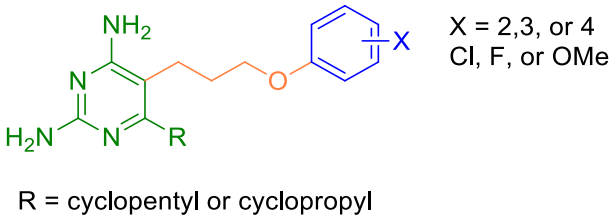
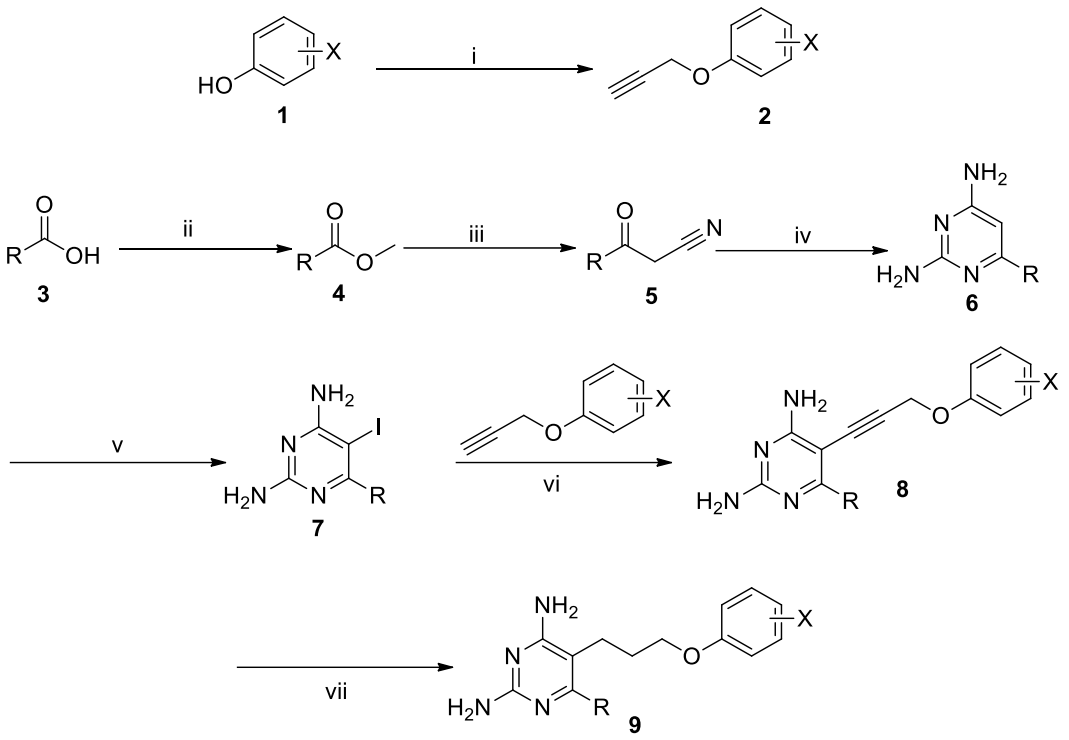


Figure 3.1: General structure of project compounds

The final fully flexible pyrimethamine analogues were assembled over several steps, starting with the linker and terminal ring portions, which are indicated in orange and blue, respectively, in figure 3.1. The core 2,4-diaminopyrimidine ring, coloured in green in figure 3.1, was constructed next, followed by the attachment of the linker and terminal ring portions to this group in a Sonogashira coupling reaction.

The overall reaction pathway is provided in scheme 3.1, whereby the final compounds were formed over a total of seven steps, starting from appropriately substituted phenols and carboxylic acids.



Scheme 3.1: Synthetic procedure for the production of flexible pyrimethamine analogues

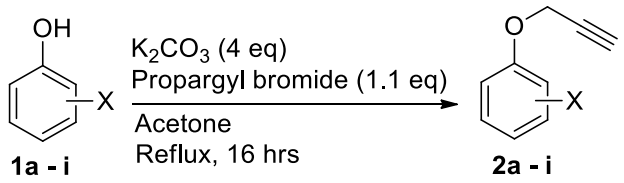
Reagents and conditions: i) Propargyl bromide, K_2CO_3, acetone, RT, 16 hrs; ii) H_2SO_4, MeOH, DCM, reflux, 16 hrs; iii) KOtBu, MeCN, IPA, 2-MeTHF, RT, 2 hrs; iv) Guanidine HCl, KOtBu, 2-MeTHF, MW, 150 W, 150 °C, 1 hr; v) I_2, HIO_3, H_2SO_4, CH_3COOH, reflux, 16 hrs; vi) $Pd(PPh_3)_2Cl_2$, CuI, DIPA, DMF, reflux, 16 hrs; vii) Pd/C, H_2(g), EtOH, RT, 24 hrs.

Appropriately substituted phenols (**1**) were alkylated with propargyl bromide under basic conditions, yielding the four-atom linker and terminal ring portion (**2**) of the desired compounds. The formation of the 2,4-diaminopyrimidine core occurred next over several steps. Starting with commercially available carboxylic acids containing appropriate R-substitutions (**3**), an esterification reaction was performed with methanol under acidic conditions to obtain the corresponding methyl esters (**4**). The esters then underwent a base mediated nucleophilic substitution reaction with acetonitrile, forming the desired beta-ketonitrile product (**5**). Next, this product was combined with guanidine hydrochloride under basic conditions in a microwave assisted ring closure reaction for the formation of the 2,4-diaminopyrimidine core (**6**). For the subsequent attachment of the linker and terminal ring

portions in a cross coupling reaction, the 5-position of the pyrimidine core required halogenation. This was achieved by a reaction with iodine under acidic conditions, forming the desired 5-iodo pyrimidine (**7**). The iodinated intermediates were then combined with the previously prepared propargyl substituted phenols (**2**) in a palladium-catalysed Sonogashira cross coupling reaction. This connected the 2,4-diaminopyrimidine core to the linker and terminal ring portions, affording the alkyne containing intermediate **8**. Finally, the triple bond was reduced in a hydrogenation reaction to yield the fully flexible final compounds (**9**) over a total of seven steps.

3.1.1 Alkylation of phenols with propargyl bromide

The first step in the overall synthesis of the flexible pyrimethamine analogues involved the formation of the terminal ring and linker portions of the final compounds. Based on the results of the previously conducted molecular modelling campaign, the functional groups selected as X-substituents around the terminal ring included chloro, fluoro, and methoxy substitutions in the 2, 3, and 4-positions of the ring. The general conditions used are described in scheme 3.2, and were adapted from well described literature methodology[72].



Scheme 3.2 Alkylation of phenols with propargyl bromide

To prepare the desired compounds, appropriately substituted phenols (**1**) were dissolved in acetone and treated with potassium carbonate and propargyl bromide and heated to reflux for 16 hours. The progress of the reaction was monitored by thin layer chromatography (TLC), where the products showed Rf values distinct to the starting materials. In each case, the desired propargyl substituted phenols (**2a – i**) were isolated as light-yellow oils in good yields ranging from 74.5 % to 99.5 % after purification by column chromatography. A total of nine

different substituted phenols were formed from readily available starting materials. The range of compounds formed, along with their corresponding yields, is presented in table 3.1.

Table 3.1: List of alkylated phenols (**2a – i**)

Number:	**X-substituent:**	**Yield (%):**
2a	4-Cl	83.0
2b	3-Cl	88.4
2c	2-Cl	93.3
2d	4-F	74.5
2e	3-F	82.8
2f	2-F	93.4
2g	4-OMe	99.5
2h	3-OMe	94.9
2i	2-OMe	90.3

Full characterization of the products was conducted with NMR and IR spectroscopy. The assignments of the NMR spectra were made in accordance with the numbering system indicated in figure 3.2.

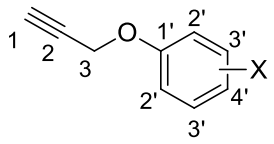


Figure 3.2: General numbering system for propargyl substituted phenols (**2**)

In the ^{1}H NMR spectra of the formed compounds, signals with a total integration of four are present in the aromatic region. The nature of these signals is dependent on the particular X-substituent, and the position around the ring. The assignments for these signals were made based on their integration and splitting patterns. In the case of the fluorine containing analogues, additional $J_{F\text{-}H}$ coupling is observed.

Peaks corresponding to the protons on carbons C1 and C3 of the added propargyl group appear at approximately 2.5 ppm and 4.7 ppm, respectively, in the ^{1}H NMR spectra. The C1 proton gives rise to a triplet signal with an integration of one, while the C3 protons produce a doublet signal which integrates for two. These two signals couple each other through the triple bond, resulting in a $^{4}J_{H\text{-}H}$ coupling constant of 2.4 Hz. However, this long-range coupling is not well defined in the case of formed products **2a** and **2i**, with the two signals appearing as either broad singlets or multiplets due to overlapping peaks. In the case of the methoxy containing analogues, an additional signal is evident in the ^{1}H NMR spectra. This signal appears at approximately 3.8 ppm and is a singlet with an integration of three.

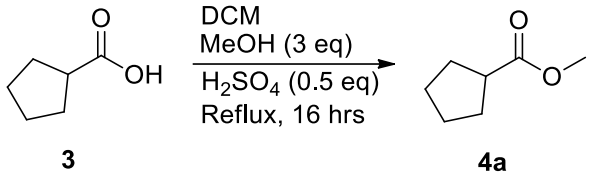
The ^{13}C NMR spectra show a number of signals in the aromatic region which were assigned based on their correlation to the already assigned ^{1}H NMR spectroscopic peaks in HSQC and HMBC spectra. The assignments were further assisted by the splitting in the ^{13}C NMR spectra of the fluorine containing compounds, where $^{1}J_{C\text{-}F}$ values of around 240 Hz, $^{2}J_{C\text{-}F}$ values of around 24 Hz, $^{3}J_{C\text{-}F}$ of around 8 Hz, and $^{4}J_{C\text{-}F}$ values of around 2.4 Hz are all observed. For each compound, three additional peaks are present in the ^{13}C NMR spectra which correspond to the added propargyl group. The signals for the C1, C2, and C3 carbons appear at approximately 76 ppm, 79 ppm, and 56 ppm, respectively. In each case, the assignments of these peaks were made by consulting the DEPT, HSQC, and HMBC spectra of the formed products. The analogues with methoxy groups have one additional peak in the ^{13}C NMR spectra, which is observed at approximately 55 ppm.

The terminal alkyne functional group present in the formed compounds results in two key vibrational modes which are clearly seen in the IR spectra. The C≡C-H stretching mode is observed as a sharp peak at around 3290 cm^{-1}, while the C≡C stretching mode is seen as a weak to medium intensity peak at approximately 2120 cm^{-1}. These two peaks are evident in the IR spectra of all the formed compounds, in addition to the other expected C-H, C=C, and C-O stretching vibrations.

3.1.2 Esterification of aliphatic carboxylic acids

The next portion of the synthesis was the formation of the 2,4-diaminopyrimidine ring, starting from commercially available carboxylic acids with appropriate R-substituents (**3**). Based on the results of the molecular modelling campaign, several different groups were suitable for inclusion in this position. The synthesis described here involves the use of cyclopentyl and cyclopropyl R-groups.

Aliphatic carboxylic acids, as well as methyl and ethyl esters of these, are readily available starting materials. In the case of this project, the ethyl ester of cyclopropyl carboxylic acid was on hand, while only the unprotected cyclopentyl carboxylic acid was available. Therefore, the initial esterification step was only required for the cyclopentyl carboxylic acid. The conditions used for this transformation are included in scheme 3.3.



Scheme 3.3: Esterification of aliphatic carboxylic acids

Standard esterification conditions were employed, with the desired product obtained by refluxing a mixture of cyclopentyl carboxylic acid (**3**), sulfuric acid, and methanol in dichloromethane for 16 hours. Unfortunately, neither the product nor the starting material could be seen on a TLC plate. Several different stains were tried but yielded no positive results. It was determined however that 16 hours was sufficient time for complete conversion of the starting material. The methyl ester of cyclopentyl carboxylic acid (**4a**) was isolated as a volatile clear oil in a yield of 92.6 % after workup of the reaction without the need for further purification.

The formed product was characterized by NMR and IR spectroscopy. The ^{1}H NMR and ^{13}C NMR spectra of the product show an additional peak each in comparison to the spectra of the starting material, indicating the successful esterification with methanol. In the ^{1}H NMR

spectrum, this peak appears as a singlet with an integration of three at 3.66 ppm. In the ^{13}C NMR spectrum, the corresponding signal is observed at 51.73 ppm. The peak for the carbonyl carbon is present in the ^{13}C NMR spectrum at 177.44 ppm.

The additional peaks to assign in the NMR spectra arise from the cyclopentyl group. This group is divided by a plane of symmetry and contains inequivalent axial and equatorial protons which are attached to the same carbon atoms. This is described in figure 3.3, which includes both a side on and overhead view of the formed methyl cyclopentane carboxylate, showing the mirror plane and positions of the axial and equatorial hydrogens, and defining the numbering system for the NMR spectroscopic assignments.

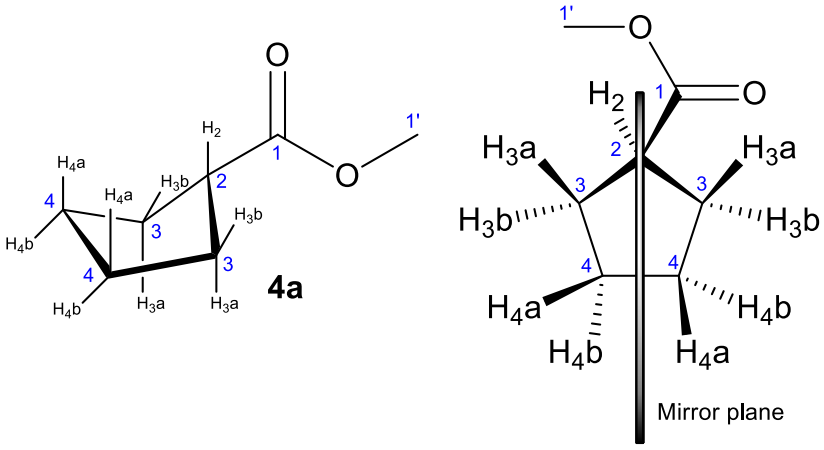


Figure 3.3: Atom numbering and three-dimensional configuration of cyclopentyl group of **4a**

The cyclopentyl ring adopts an envelope shape, where four of the carbon atoms lie in the same plane and the fifth extends slightly upwards. In this configuration there is a plane of symmetry which bisects the ring into two. This can be seen in the figure above, where carbon atoms on either side of the mirror plane are equivalent to each other.

There are three unique carbon environments for this group, resulting in three signals in the ^{13}C NMR spectrum. Signals at 43.82 ppm, 30.15 ppm, and 25.93 ppm were assigned to C2, C3, and C4, respectively. Five unique hydrogen environments are seen in the ^{1}H NMR spectrum, which correspond to the C2 hydrogen, and the axial and equatorial hydrogens on C3 and C4.

A peak at 2.72 ppm with an integration of one was assigned to the C2 hydrogen. Four other signals are observed slightly upfield of the C2 signal, each a multiplet integrating to two.

The coupling between the C2 and C3 hydrogens seen in the COSY spectrum allowed for definitive assignments to be made for the C3 and C4 ^{1}H NMR spectroscopic signals. The hydrogens on C3 and C4 have two peaks each in the ^{1}H NMR spectrum, due to the different environments experienced by the axial and equatorial protons.

Distinction between the axial and equatorial signals was tentatively made based on the chemical shifts of the peaks, where the axial protons are expected to be shifted slightly upfield from their equatorial counterparts. The two equatorial protons for C3 were assigned to the peak centred at 1.87 ppm, the equatorial C4 protons to the peak centred at 1.78 ppm, the axial C3 protons to the peak centred at 1.69 ppm, and the two axial C4 protons to the peak centred at 1.57 ppm.

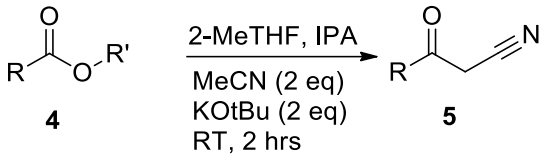
Analysis of the IR spectrum reveals the characteristic C=O stretching mode at 1732 cm^{-1}, and the absence of the broad O-H stretching peak, which would be observed in the starting material at around 3200 cm^{-1}.

3.1.3 Formation of beta-ketonitriles

The next step in the assembly of the 2,4-diaminopyrimidine core was the formation of beta-ketonitriles (**5**) from the appropriately R-substituted esters (**4**). In the case of the cyclopentyl containing compounds, the previously synthesized methyl ester was used, and for the cyclopropyl analogues, the ethyl ester which was already available in our laboratories was used.

A method for the formation of beta-ketonitriles from esters and acetonitrile, with the addition of potassium *tert*-butoxide, has previously been developed in our laboratories[73]. While this is a green and economical approach, the reported yields were relatively low (< 50 %). Some efforts were therefore undertaken to devise higher yielding reaction conditions. Several different bases and solvents, as well as the effects of temperature and reaction time were assessed for the formation of the beta-ketonitrile products from the ester starting materials and acetonitrile. It was found that regardless of the base and solvent used, heating of the

reaction led to a complex mixture of decomposed starting material. When performed at room temperature, traces of the product could be seen for reactions conducted with both sodium hydride and *n*-butyllithium in tetrahydrofuran (THF) or 2-methyltetrahydrofuran (2-MeTHF), along with mostly unreacted starting material. Increasing the equivalents of base, or the reaction time did not result in improvement in yield. After trying several different combinations of reaction conditions with no success, it was decided to proceed with the moderately yielding method described by Pienaar et al, 2019[73]. The conditions for this transformation are shown in scheme 3.4.



4a: R = cyclopentyl, R' = methyl
4b: R = cyclopropyl, R' = ethyl

5a: R = cyclopentyl, 55 % yield
5b: R = cyclopropyl, 55 % yield

Scheme 3.4: Formation of beta-ketonitriles

The ester starting materials (**4**) were dissolved in 2-MeTHF, along with a small amount of isopropyl alcohol. One equivalent each of potassium *tert*-butoxide and acetonitrile were then added, and after an hour, a further equivalent of these two reagents was added. The reaction was left to proceed for another hour before quenching with saturated aqueous ammonium chloride. While the ester starting materials were not able to be visualized by TLC, application of a potassium permanganate stain revealed the product as a vivid purple spot. This allowed for purification by column chromatography, which was used to obtain the products as yellow oils in moderate yields (55 %). The formed beta-ketonitrile products (**5**) have a distinct odour and are somewhat unstable. The yellow oils were seen to darken over time, and ^{1}H NMR spectra of old samples revealed some decomposition after a few weeks of standing at room temperature. The possible decomposition products were not identified, as the beta-ketonitriles were used in the subsequent reactions shortly after their formation.

NMR and IR spectroscopy were employed to confirm the successful formation of the desired products. The assignments of the NMR spectra were made in accordance with the general numbering system defined in figure 3.4.

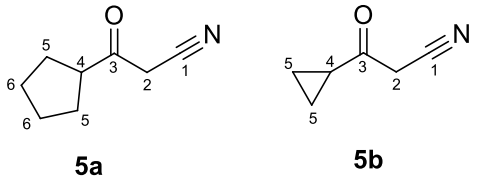


Figure 3.4: General numbering system of beta-ketonitriles (**5**)

In the ^{1}H and ^{13}C NMR spectra, the peak/s corresponding to the methyl or ethyl ester groups of the starting materials are absent, and new peaks resulting from the added acetonitrile moiety are evident. The ^{13}C NMR spectra show the characteristic carbonyl carbon peak at approximately 200 ppm. Two additional peaks are also observed in the ^{13}C NMR spectra at around 114 and 32 ppm, which were assigned to carbons C1 and C2, respectively, for each product. The ^{1}H NMR spectroscopic signal for the C2 protons is observed as a singlet which integrates for two at approximately 3.5 ppm. The nature of the peaks corresponding to the cyclopentyl group have already been discussed in some depth, and only show slight changes between the spectra of the starting material and products for this reaction.

The cyclopropyl group displays similar behaviour to the cyclopentyl group, where a fixed ring system with axial and equatorial protons is divided by a plane of symmetry. Two views of the formed cyclopropyl beta-ketonitrile are provided in figure 3.5. Here the plane of symmetry running through the cyclopropyl ring is indicated, as well as the axial and equatorial positions of the hydrogens on carbon atoms C5.

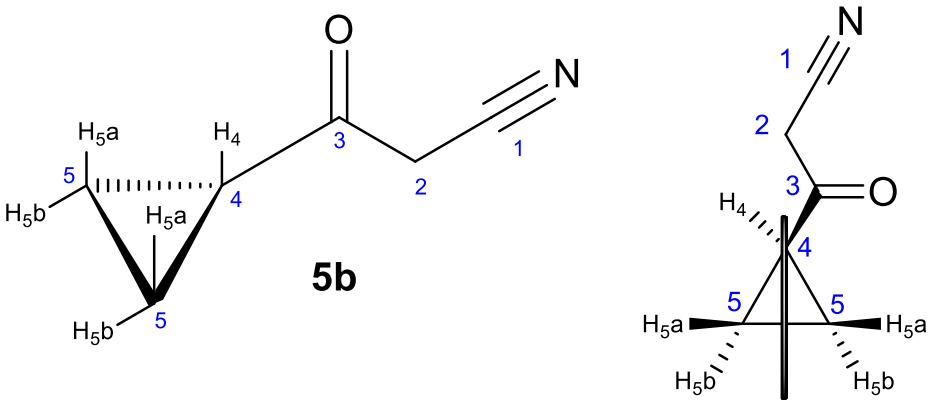


Figure 3.5: Three-dimensional configuration and atom numbering of cyclopropyl group of **5b**

In the ^{13}C NMR spectrum, two signals are observed for this group. The C4 carbon was assigned to the peak at 20.25 ppm and the C5 carbons assigned to the peak at 13.04 ppm. The ^{1}H NMR spectrum reveals three distinct peaks in addition to those assigned for the C2 protons. The proton of C4 is a quintet which integrates for one at 2.10 ppm, while the equatorial C5 protons produce a multiplet centered at 1.20 ppm integrating to two, and the C5 axial protons produce a multiplet centered at 1.08 ppm which also integrates for two.

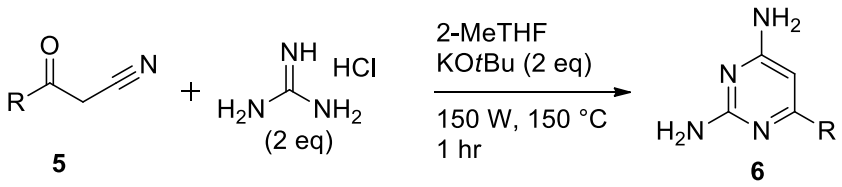
The characteristic C≡N stretching vibration is observed in the IR spectrum as a weak to medium intensity peak at around 2260 cm^{-1}, in addition to the other expected C=O and C-H stretching modes.

3.1.4 Microwave assisted formation of 2,4-diaminopyrimidines

The isolated beta-ketonitrile products (**5**) were subjected to a microwave assisted ring closure reaction with guanidine hydrochloride (HCl) in the presence of base to form the core 2,4-diaminopyrimidne ring (**6**).

Initially it was hoped to perform both the previous transformation and this one in a single one-pot two-step reaction. To realize this, the same base and solvent was assessed for the ring closure portion, starting from the isolated beta-ketonitrile intermediates (**5**). However, very low yields of the desired 2,4-diaminopyrimidine products (**6**) were afforded when the reaction was performed at room temperature. It was found that significant amounts of the product

were only formed when the reaction was heated to reflux. Following this, the one-pot two-step method was attempted by following the procedure described for the beta-ketonitrile synthesis, and then adding guanidine HCl and heating the mixture to reflux for several hours after the initial two-hour reaction period was complete. Although the formation of the beta-ketonitrile intermediate could be observed by TLC analysis, very low yields of the final 2,4-diaminopyrimidine products were isolated. Due to this, the one-pot two-step approach was abandoned in favour of performing the two reactions as distinct steps. Details of the reaction are provided in scheme 3.5.



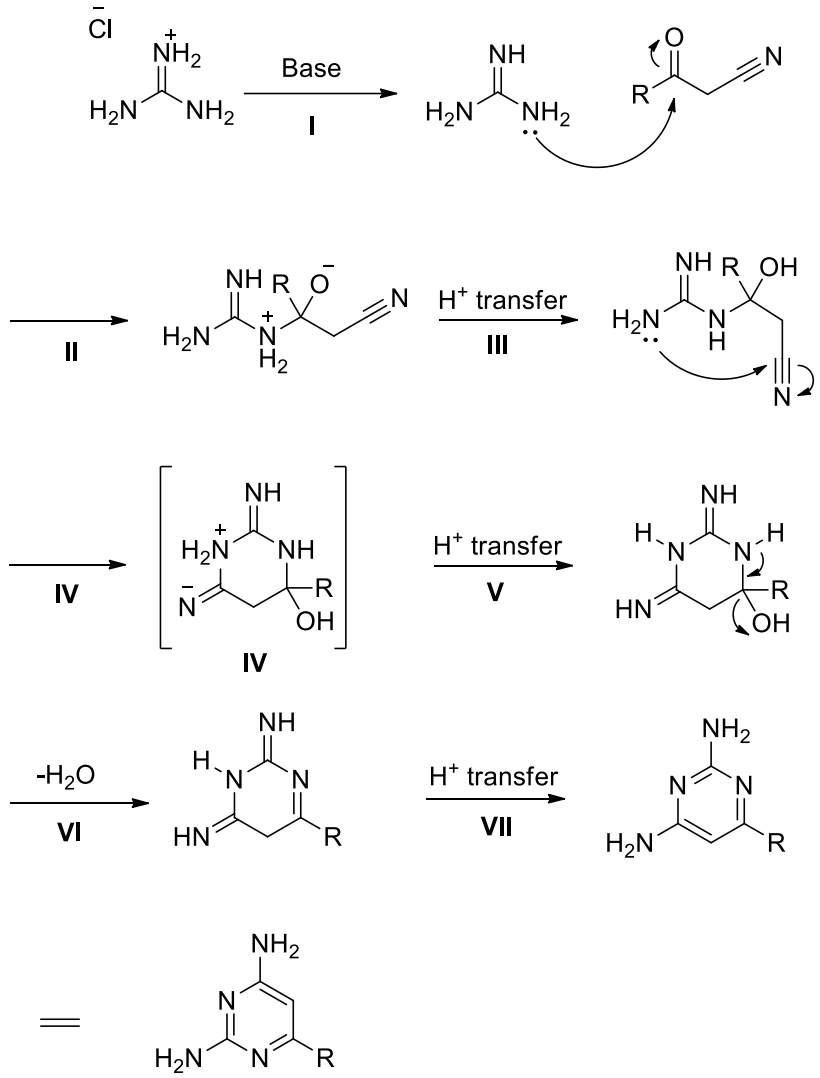
5a: R = cyclopentyl
5b: R = cyclopropyl

6a: R = cyclopentyl, 68 % yield
6b: R = cyclopropyl, 74 % yield

Scheme 3.5: Microwave assisted formation of 2,4-diaminopyrimidines

Yields obtained from normal heating of the reaction mixture were somewhat low, but it was found that heating with microwave irradiation improved the yields significantly and allowed for shortened reaction times. Several rounds of reaction condition optimization were conducted, and it was determined that two equivalents each of potassium *tert*-butoxide and guanidine HCl, along with a minimal amount of 2-MeTHF solvent produced the best results when heated to 150 °C at 150 W power for one hour in a microwave reactor. The formed 2,4-diaminopyrimidines (**6**) were found to be highly water soluble, and so in the initial stages of reaction optimization much of the product was lost into the aqueous phase when doing an aqueous work up. Great improvements were made in the yields when the aqueous work up was avoided altogether, and instead the crude reaction mixture was filtered through celite, and the filtrate adsorbed directly onto silica for subsequent purification by column chromatography. The TLC of the filtered reaction mixture showed many highly polar compounds formed; however the desired product was the most polar among them and could be eluted last from the column. After purification by column chromatography, the products

were isolated as yellow solids in acceptable yields of around 70 %. A proposed mechanism for the ring closure reaction is provided in scheme 3.6.



Scheme 3.6: Proposed mechanism for ring closure reaction in the formation of 2,4-diaminopyrimidines (**6**)

The guanidine HCl is neutralized by the base [**I**], and one of the nucleophilic nitrogen atoms then adds to the electrophilic carbonyl carbon of the beta-ketonitrile [**II**]. An H^+ transfer step occurs next [**III**], followed by the addition of one of the other nitrogens from the guanidine starting material to the electrophilic nitrile carbon of the beta-ketonitrile [**IV**], forming a six

membered ring. Aromatization of the ring occurs through the loss of water, and several proton transfer steps [**V**, **VI**, **VII**] to form the final 2,4-diaminopyrimidine ring.

The formed products were definitively characterized by IR and NMR spectroscopy, as well as mass spectrometry. Melting points were also recorded. The assignments for the NMR spectra were made with reference to the numbering system defined in figure 3.6. ^{1}H and ^{13}C NMR spectra of formed product **6a** are included in figures 3.7 and 3.8, respectively, with the assignments of the various spectroscopic signals indicated.

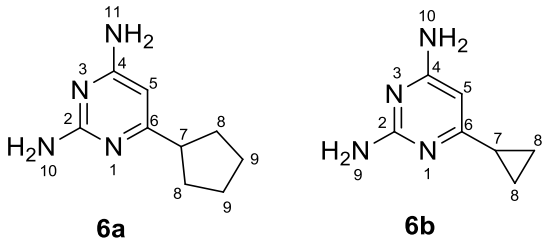


Figure 3.6: General numbering system for 2,4-diaminopyrimdines (**6**)

The NMR spectra show many clear differences from the beta-ketonitrile starting materials. In the ^{1}H NMR spectra, three signals are observed for the formed pyrimidine ring, in addition to the peaks corresponding to the different R-groups. The two primary amine groups produce a signal each, which show as broad singlets that integrate for two. These peaks show high variability in their chemical shifts. This is typical of amines and is likely due to differences in the concentrations of the prepared NMR samples. Distinction was made between these two groups based on HMBC spectra, where long-range coupling of the amine hydrogens through the nitrogen heteroatom to some of the carbons in the pyrimidine ring could be observed. The amine group attached to C4 of the pyrimidine ring typically appears between 6.0 and 7.0 ppm, while the peak for the amine group on C2 is found slightly upfield from the first peak, at around 5.5 to 6.5 ppm. The other ^{1}H NMR spectroscopic signal arising from the pyrimidine ring is a singlet with an integration of one and was assigned to the proton on C5.

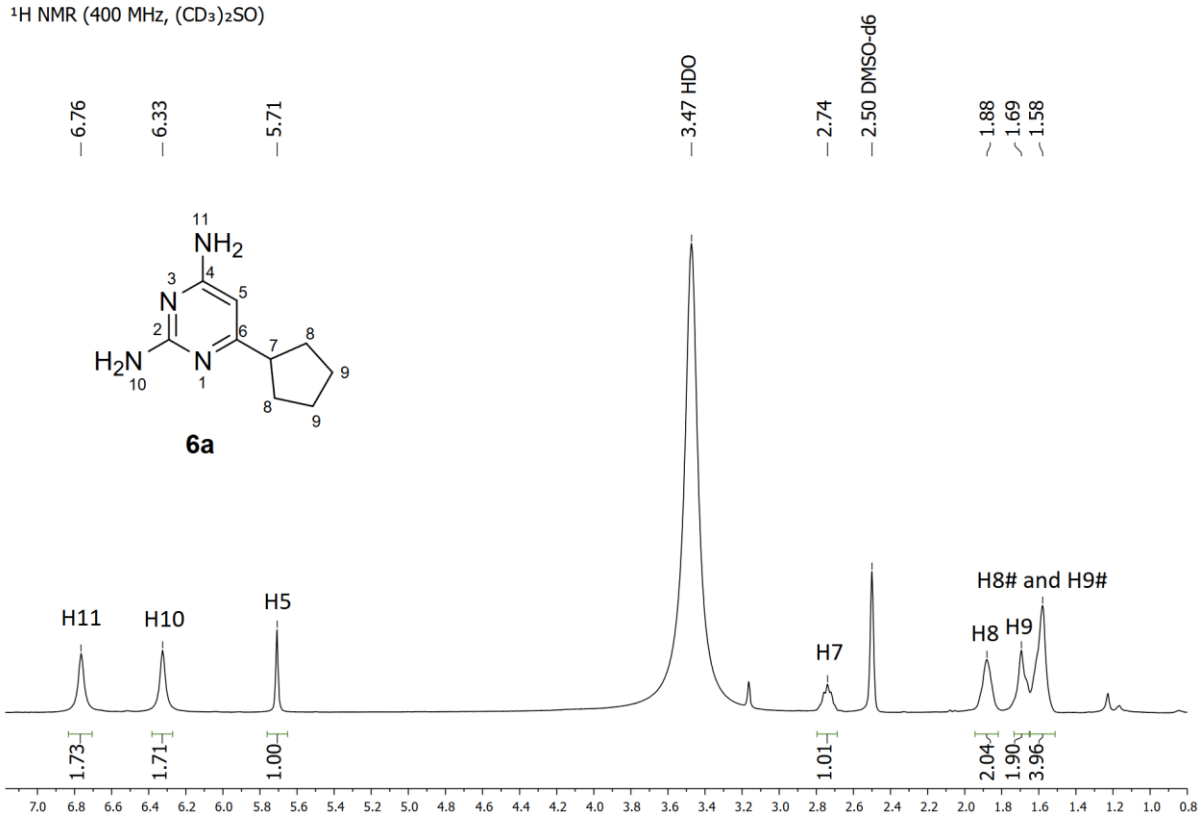


Figure 3.7: ^{1}H NMR spectrum of **6a** with spectroscopic signal assignments included

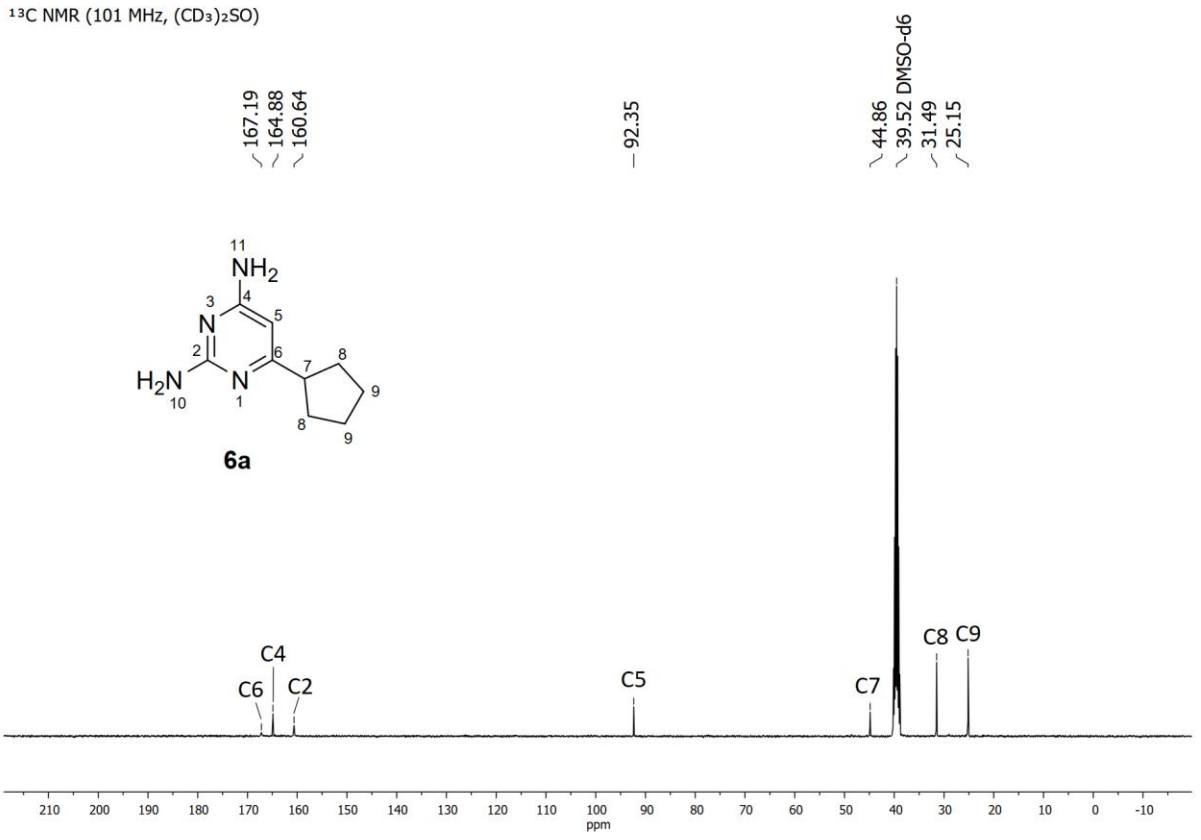


Figure 3.8: ^{13}C NMR spectrum of **6a** with spectroscopic signal assignments included

The ^{13}C NMR spectra show four peaks which correspond to the pyrimidine ring, in addition to those arising from the different R-groups. The assignments of these signals were made based on HMBC spectra which showed multiple bond coupling to the already assigned proton signals. The peak for the C5 carbon is found at approximately 91 ppm, while the peaks for carbons C6, C4, and C2 appear at around 167 ppm, 164 ppm, and 162 ppm, respectively.

The ^{13}C and ^{1}H NMR spectroscopic signals for the cyclopentyl and cyclopropyl R-groups are consistent with what has already been described for the starting materials.

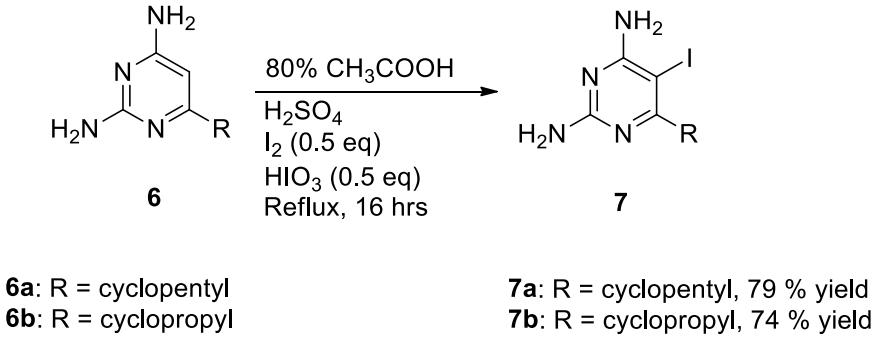
The IR spectra confirm the consumption of the starting materials as the key C≡N and C=O vibrations of the beta-ketonitriles are absent. Additional peaks corresponding to the NH_2 amine stretching modes, as well as the C=C and C=N stretching modes of the newly formed pyrimidine ring are apparent.

Mass spectrometry was run on the formed compounds, with prominent $[M+H]^+$ peaks identified in positive ion mode corresponding to the expected exact masses of the target compounds. For compound **6a**, the $[M+H]^+$ peak has an m/z value of 179.1287, and for **6b** 151.0980, which are in close agreement with the expected masses for these products of 179.1291 and 151.0978, respectively.

Melting points were recorded, revealing that the cyclopentyl analogue **6a** melts in the range of 124 – 126 °C, and the cyclopropyl compound **6b** at 138 – 140 °C.

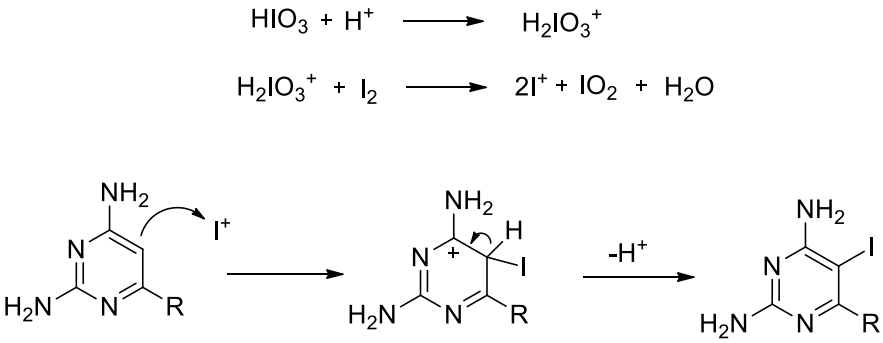
3.1.5 Iodination of 2,4-diaminopyrimidines

To facilitate the subsequent attachment of the formed propargyl substituted phenols (**2**) in a Sonogashira coupling reaction, the 5-position of the pyrimidine ring required iodination. Fortunately, there is only a single site around the ring which the iodine could be added to, allowing for high selectivity to be achieved. The conditions used are described in scheme 3.7



Scheme 3.7: Iodination of 2,4-diaminopyrimdines

The respective 2,4-diaminopyrimidine starting materials (**6**) were dissolved in 80 % acetic acid, and iodine and iodic acid were added. Next a small amount of sulfuric acid was added to ensure the reaction pH was sufficiently low for the acid catalysed reaction to occur. The reaction mixture was then heated to reflux under an inert atmosphere. The reaction progress was monitored by TLC, where the products showed a higher Rf value than the starting materials. Complete conversion was apparent after 16 hours, after which the formed products were isolated and purified with silica gel column chromatography. Decent yields of around 75 % were obtained of yellow polycrystalline material. The mechanism for this reaction is provided in scheme 3.8.



Scheme 3.8: Mechanism for acid catalysed aromatic iodination

The reaction is initiated by the formation of the electrophilic I^+ species. This occurs over two steps. First the iodic acid (HIO_3) is protonated by H_2SO_4 forming the $[H_2IO_3]^+$ intermediate. This is a strong oxidizing agent and oxidizes the molecular iodine forming two equivalents of I^+,

along with the reduced IO_2, which plays no further part in the reaction, and one equivalent of water. Once formed, the electrophilic I^+ adds to the ring in an electrophilic aromatic substitution reaction, forming the desired iodinated product.

Full characterization was achieved by the methods of NMR and IR spectroscopy, as well as mass spectrometry and melting point analysis. Assignment of the peaks in the NMR spectra were made with reference to the numbering system defined in figure 3.9.

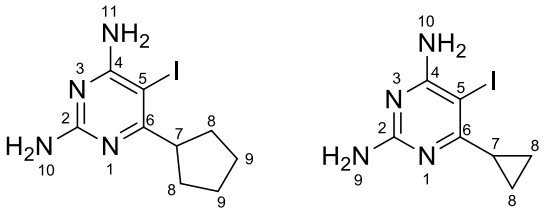


Figure 3.9: General numbering system for iodinated 2,4-dimaminopyrimdines (**7**)

Successful formation of the products is readily seen by comparison of the ^1H and ^{13}C NMR spectra with those of the starting materials. In the ^1H NMR spectra, the peak assigned to the C5 hydrogen is absent for each product, as this hydrogen was replaced by an iodine atom. All of the other peaks are very similar. In the ^{13}C NMR spectra, the position of the C5 carbon shifts drastically after the addition of the iodine. For the starting materials, this peak is at around 92 ppm. After addition of iodine this peak shifts upfield to close to 65 ppm. The chemical shifts for the other carbon atoms of the pyrimidine ring also shift slightly in position, but the order of the peaks remains the same as in the starting materials, as confirmed by HMBC spectra. The ^1H and ^{13}C NMR signals arising from the cyclopentyl and cyclopropyl groups are largely unchanged from the starting materials.

IR spectroscopy is less useful for assessing the success of the reaction, however, all of the expected peaks are seen in the spectra of the products, including N-H stretches, and C=C and C=N stretching vibrations.

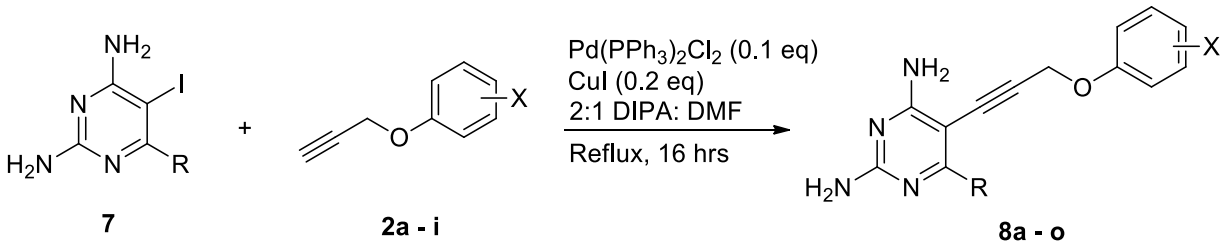
Additional characterization was done by HRMS, where $[M+H]^+$ peaks are observed for the formed products in positive ion mode. The successful addition of the iodine is confirmed by

noting the increase in the m/z value of the product $[M+H]^+$ peaks by 125.8977 (accurate to the last decimal place), which corresponds to the mass of iodine minus one hydrogen.

A small increase in the melting point of the iodinated pyrimidines is observed when compared to the starting materials, with the cyclopentyl compound **7a** melting between 143 and 146 °C, and the cyclopropyl compound **7b** melting between 145 and 146 °C.

3.1.6 Sonogashira coupling reaction

The Sonogashira coupling reaction represents a key step in the complete synthesis of the flexible pyrimethamine analogues. Here, the previously formed propargyl substituted phenols (**2a - i**) were attached to the iodinated 2,4-diaminopyrimidne rings (**7a** or **7b**). Extensive optimization of the reaction conditions was conducted in our laboratories before the advent of the present work. Different solvents, bases, and ligands around the Pd metal centre were assessed, as well as variations in the reaction temperature and duration. A robust general method was developed for the formation of the desired products and was implemented in this work. These conditions are indicated in scheme 3.9.



7a: R = cyclopenyl
7b: R = cyclopropyl

Scheme 3.9: Sonogashira coupling reaction

Dimethylformamide (DMF) and diisopropylethylamine (DIPA) in a 1:2 ratio was used as the solvent for the reaction. DIPA also performed a function as the base. This mixture was first degassed by bubbling argon or nitrogen gas through it, before the addition of the iodinated pyrimidine starting material (**7a** or **7b**), $Pd(PPh_3)_2Cl_2$ catalyst, and CuI cocatalyst. The mixture

was then heated to 80 °C to ensure the added solid materials were fully dissolved. The relevant propargyl substituted phenols (**2a – i**) were then added as a solution in degassed DMF to the reaction mixture. After 16 hours of heating at reflux (80 °C), complete conversion of the starting materials could be seen with TLC. After workup, the crude product was adsorbed onto silica for subsequent purification by column chromatography, affording the desired products as yellow solids in variable yields, ranging from around 35 % to 80 %, with most of the analogues giving yields close to 70 %.

This reaction was performed a number of times for the formation of 15 different analogues, which are indicated in table 3.2. The respective yields for each analogue are included, along with the identified $[M+H]^+$ peaks from HRMS experiments.

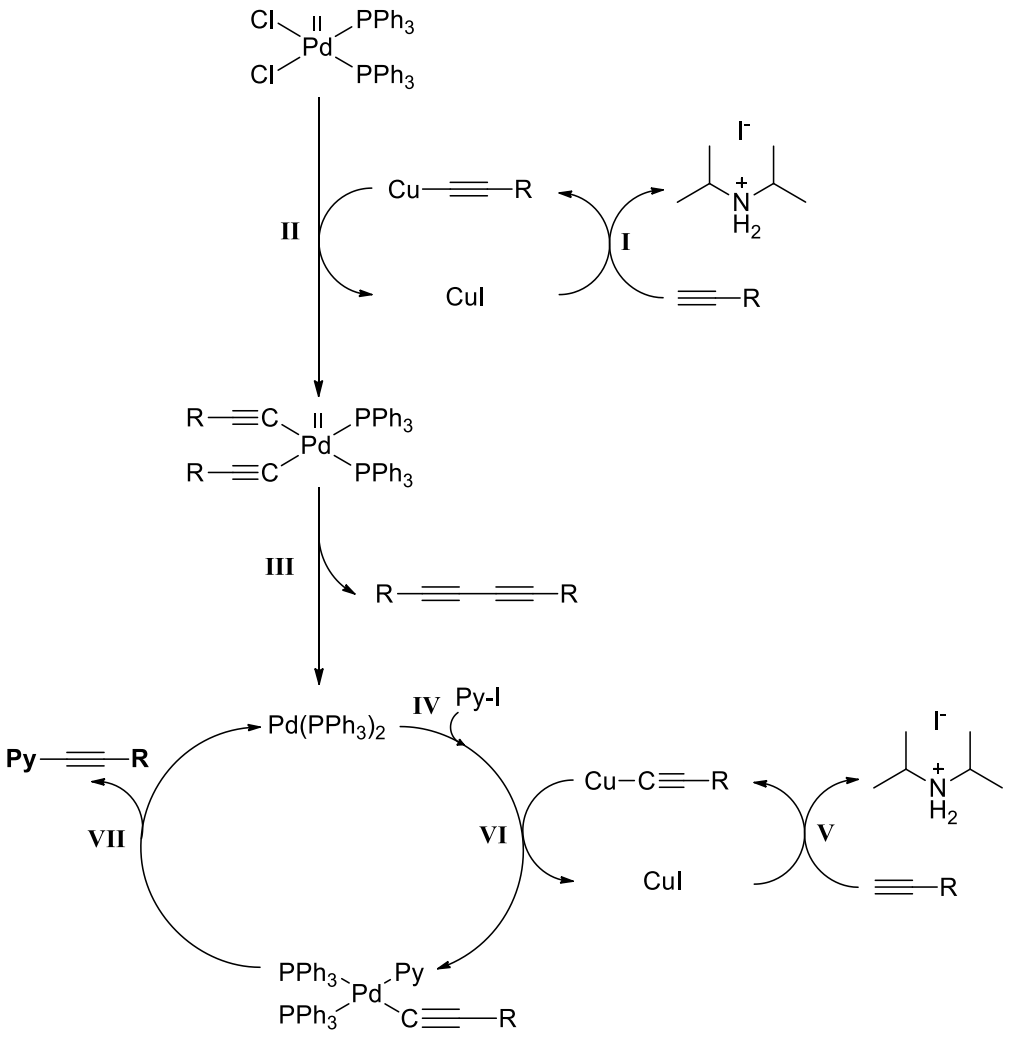
Table 3.2: List of Sonogashira reaction products (**8a** – **o**)

Number:	**R-group:**	**X-substituent:**	**Yield (%):**	**HRMS $[M+H]^+$ (m/z):**	**[M+H]+ expected (m/z):**
8a	Cpe	4-Cl	43.4	343.1291	343.1320
8b	Cpe	3-Cl	38.7	343.1297	343.1320
8c	Cpe	2-Cl	34.2	343.1292	343.1320
8d	Cpe	4-F	69.4	327.1620	327.1616
8e	Cpe	3-F	65.7	327.1641	327.1616
8f	Cpe	2-F	61.9	327.1639	327.1616
8g	Cpe	4-OMe	79.0	339.1836	339.1816
8h	Cpe	3-OMe	71.4	339.1840	339.1816
8i	Cpe	2-OMe	72.8	339.1848	339.1816
8j	Cpr	4-Cl	55.3	315.1011	315.1007
8k	Cpr	3-Cl	67.0	315.0996	315.1007
8l	Cpr	4-F	64.6	299.1317	299.1303
8m	Cpr	3-F	68.0	299.1315	299.1303
8n	Cpr	4-OMe	70.1	311.1519	311.1503
8o	Cpr	3-OMe	73.6	311.1517	311.1503

A general trend in the yields for the various analogues can be seen. The chlorine containing compounds were obtained in the lowest yields, followed by the fluorine analogues, while the methoxy compounds afforded the best results. In general, the moderate yields can be attributed to difficulties in the purification of the formed products, where triphenylphosphine oxide ($OPPh_3$) byproducts co-eluted from the column along with the desired compounds. Several methods were investigated in this work for the effective removal of the $OPPh_3$ impurity. Some success was achieved in recrystallizing the fractions obtained from the column with ethyl acetate / hexane solvent systems. Upon cooling, the pure product precipitated out of the mixture while the $OPPh_3$ impurity remained dissolved, allowing for removal of the byproduct by filtration. Another method investigated which showed good results was performing a second column chromatography purification step after the desired product and $OPPh_3$ impurity were isolated together from the initial column. By using basic alumina as the stationary phase, the $OPPh_3$ would elute first from the column when flushed with a moderately polar 50 % ethyl acetate / hexane solvent system. Under these conditions the desired product remained strongly adhered to the top of the column, and only eluted when the polarity of the mobile phase was increased up to 10 % MeOH / DCM. Neither of these two methods were completely optimal, however, and often had to be repeated multiple times to ensure complete removal of the $OPPh_3$ byproduct.

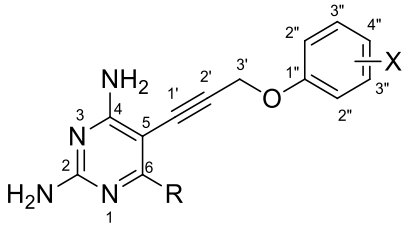
The overall catalytic cycle for this reaction is described in scheme 3.10. In this scheme, Py-I represents the iodinated 2,4-diaminopyrimidine starting materials (**7a** and **7b**), while R corresponds to the various substituted phenols employed. The first phase is the formation of the active Pd^0 catalyst species from the $Pd^{II}(PPh_3)_2Cl_2$ catalyst precursor. This is initiated in **step I** of the mechanism as described in scheme 3.10, where the terminal alkyne is deprotonated by the added base and coordinates to the copper cocatalyst. Transmetallation occurs next in **step II**, where two of the alkyne ligands are transferred to the Pd catalyst centre, which loses its existing chlorine ligands. **Step III** describes the subsequent reductive elimination, where a new bond is formed between the two terminal alkynes, generating a homocoupling byproduct and reducing the Pd metal centre to the 0-oxidation state. This occurs once for each equivalent of the catalyst until all of the Pd present in the reaction mixture is in the 0-oxidation state. For each equivalent of Pd catalyst added, one equivalent of the homocoupling byproduct is formed. These reaction byproducts were identified from the crude reaction mixtures but not

fully characterized. The result of this is the *in situ* generation of the active $Pd^0(PPh_3)_2$ catalyst species, which facilitates the catalytic cycle responsible for the formation of the desired product. The iodinated pyrimidines coordinate to the Pd metal centre in an oxidative insertion reaction in **step IV**, with the loss of iodine. Simultaneously, **step V** occurs, which entails the deprotonation of the terminal alkyne and coordination to the copper cocatalyst. **Step VI** is a transmetallation reaction, where the alkyne coordinated to the copper cocatalyst is transferred to the Pd metal centre. The complex then isomerizes so that the coordinated pyrimidine and alkyne are adjacent, followed by the reductive elimination in **step VII**. A new carbon – carbon bond is formed between the terminal alkyne and the pyrimidine ring, generating the desired product along with the oxidation of the Pd catalyst back to the active 0-oxidation state. This cycle continues until all of the starting materials have been consumed.



Scheme 3.10: Catalytic cycle for Sonogashira coupling reaction

Characterization of the formed products was done by NMR and IR spectroscopy, as well as mass spectrometry. Melting points for each analogue were also recorded. NMR assignments were made with reference to the numbering system defined in figure 3.10. ^{1}H and ^{13}C NMR spectra of formed product **8g** are included in figures 3.11 and 3.12, respectively, with the assignments of the various spectroscopic signals indicated.



8

Figure 3.10: General numbering system for Sonogashira products (**8**)

The ^{1}H NMR spectrum for each compound reveals peaks corresponding to both the 2,4-diaminopyrimidine core and the added linker and ring portions. Two broad singlets, each integrating to two, are observed in the ^{1}H NMR spectra and were assigned to the amine groups attached to C4 and C2 of the pyrimidine ring. The C4 amine peak appears at around 5.1 ppm when the spectra are run in $CDCl_3$, and at around 6.4 ppm when run in DMSO-*d6*. The C2 amine peak is slightly upfield from the C4 amine peak, at around 5.0 ppm in $CDCl_3$ and 6.1 ppm in DMSO. The peak for the hydrogens on the linker carbon C3' appears as a singlet integrating to two protons at around 5.0 ppm. In addition, peaks integrating to a total of four protons are seen in the aromatic region of the ^{1}H NMR spectra. The nature of these peaks is conserved between the propargyl substituted phenols (**2a – i**) used as starting material and the products formed in this reaction. The assignments were made based on the integration and splitting patterns of the peaks and are consistent with the assignments made for the starting materials.

The ^{13}C NMR spectrum of each product contains three characteristic peaks which were assigned to carbons C1', C2', and C3' of the linker, appearing at around 92 ppm, 81 ppm, and 57 ppm, respectively. The assignments were made based on HSQC and HMBC spectra. Four additional peaks are seen in the ^{13}C NMR spectra which correspond to the carbons present in

the pyrimidine ring of each product. HMBC spectra confirmed that the order of these peaks is preserved from the starting materials, with C6 at around 174 ppm, C4 around 164 ppm, C2 around 161 ppm, and C5 close to 88 ppm. The peaks in the ^{13}C NMR spectra of each product which arise from the carbons of the substituted phenyl ring are much the same as in the propargyl substituted phenol starting materials (**2a – i**). Likewise with the starting materials, the assignments for these peaks were made based on HSQC and HMBC spectra and were assisted by observing the expected $J_{C\text{-}F}$ coupling constants for the flouring containing analogues. The compounds bearing a methoxy substituent have an extra peak in both the ^{1}H and ^{13}C NMR spectra. The ^{1}H NMR spectroscopic signal appears as a singlet with an integration of three protons close to 3.8 ppm, while the corresponding ^{13}C NMR spectroscopic signal is observed at approximately 56 ppm.

The remaining peaks to assign in the ^{13}C NMR and ^{1}H NMR spectra arise from the cyclopentyl and cyclopropyl R-groups. These peaks show only minor differences from the starting materials, where they have already been described in some detail.

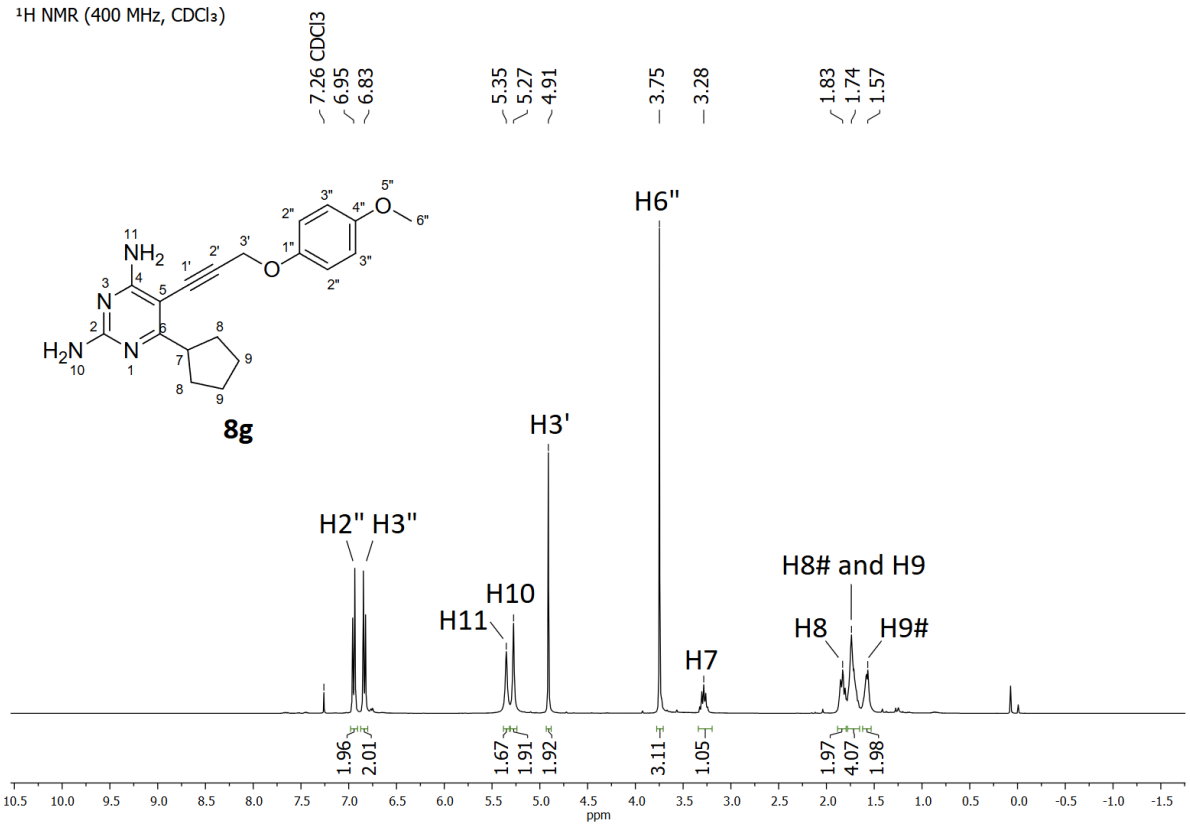


Figure 3.11: ^{1}H NMR spectrum of **8g** with spectroscopic signal assignments included

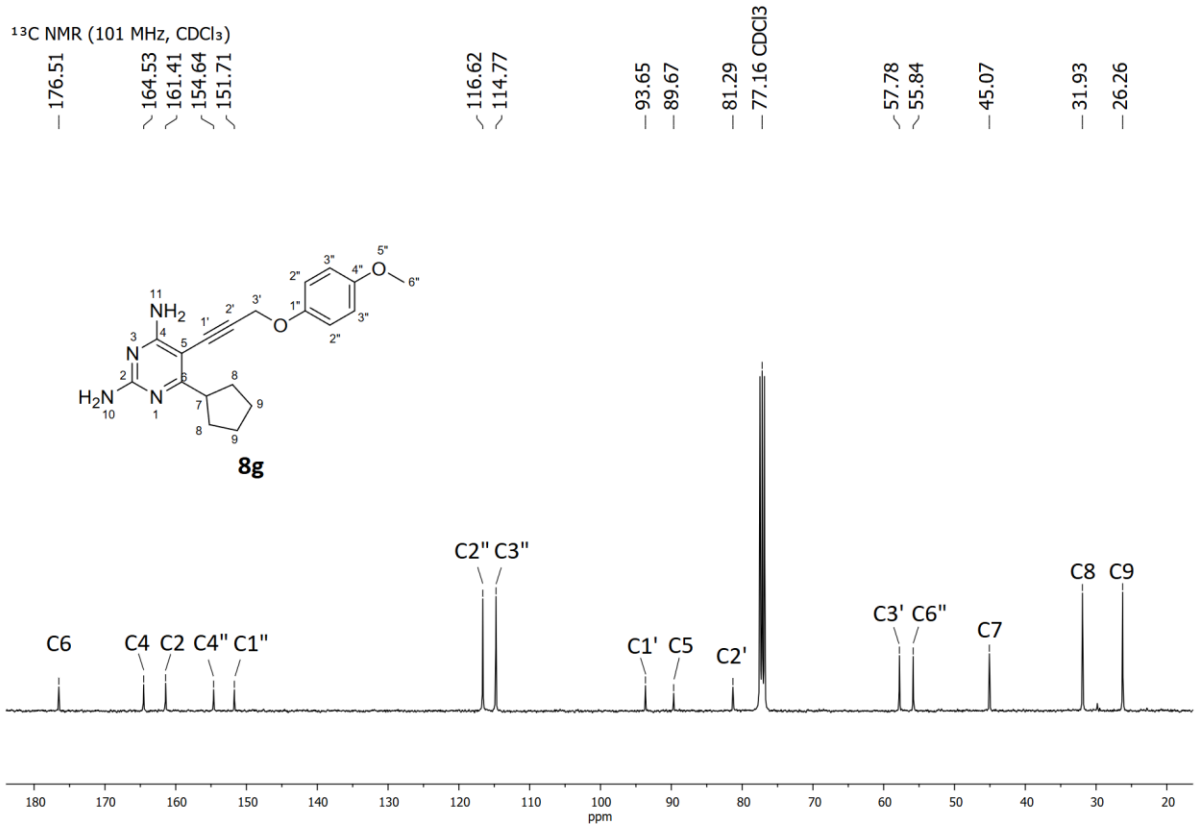


Figure 3.12: ^{13}C NMR spectrum of **8g** with spectroscopic signal assignments included

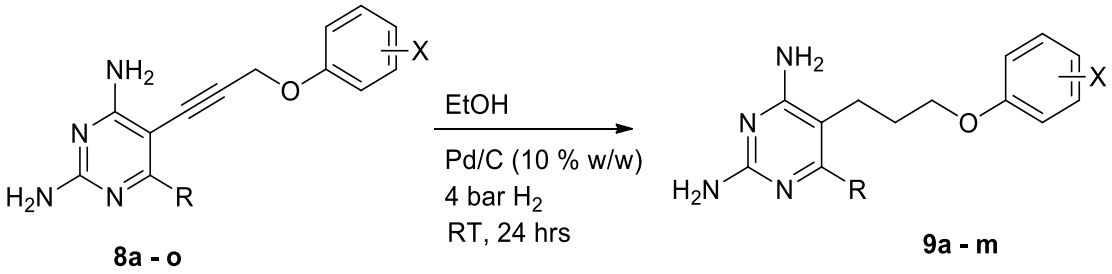
In the IR spectra, a weak intensity peak at around 2200 cm^{-1} reveals the presence of the C≡C bond. This is seen in addition to the other expected N-H, C=N, C=C, and C-O stretching modes.

All formed products were further characterized with mass spectrometry. When run in positive ion mode, prominent $[M+H]^+$ peaks were observed for each compound. The observed and expected values are included in table 3.2.

Melting points were recorded for each analogue and vary greatly depending on the nature of the X-substituent and its position around the ring.

3.1.7 Hydrogenation of Sonogashira reaction products

The final step in the synthesis of the flexible pyrimethamine analogues was the reduction of the triple bond present in the four-atom linker under hydrogenation conditions indicated in scheme 3.11.



Scheme 3.11: Hydrogenation of Sonogashira reaction products

The alkyne starting material (**8**) was dissolved in ethanol and placed in a high-pressure reaction vessel. Palladium on carbon was added and the heterogenous mixture stirred at room temperature under 4 bar hydrogen gas for 24 hours. In some cases, the reaction went to completion and the desired product was obtained in quantitative yields without further purification. More often than not, however, products were isolated after filtering as a mixture of partially reduced components, which contained traces of the alkyne starting material and semi-reduced alkene products along with the desired fully saturated alkanes. This was still sometimes evident even after subjecting the mixture to the hydrogenation conditions for a second time. Additional challenging purification steps were required in such cases. Recrystallization was found to be ineffective for the separation, and column chromatography was not ideal as the Rf values of the starting material, partially reduced product, and fully reduced product are nearly identical. Some success in the purification of the desired products was achieved by performing multiple sequential silica gel column chromatography steps, with a 10 % MeOH / DCM solvent system, although this extended purification process resulted in significantly diminished yields. For the hydrogenation reactions conducted on alkyne intermediates **8b** and **8k**, which are the 3-Cl analogues with cyclopentyl and cyclopropyl R-groups, respectively, sufficient purity of the corresponding fully reduced final products could

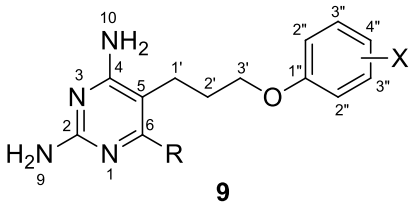
not be attained. After performing several sequential column chromatography purification steps for each of these compounds, traces of the starting materials and partially reduced products were still evident. With insufficient material on hand to continue to the purification process, it was decided to leave these two compounds out of the list of final products. For all of the other alkyne intermediates, the corresponding final products were successfully isolated as white polycrystalline solids, ranging in yield from 20 % to 100 %, based on the purification requirements. The 13 analogues produced this way are listed in table 3.3, including their respective yields and m/z values obtained from mass spectrometry.

Table 3.3: List of hydrogenation reaction products (**8a** – **o**)

Number:	**R-group:**	**X-substituent:**	**Yield (%):**	**HRMS $[M+H]^+$ (m/z):**	**[M+H]+ expected (m/z):**
9a	Cpe	4-Cl	66.7	347.1642	347.1633
9b	Cpe	2-Cl	23.4	347.1623	347.1633
9c	Cpe	4-F	44.3	331.1927	331.1929
9d	Cpe	3-F	50.4	331.1933	331.1929
9e	Cpe	2-F	35.6	331.1899	331.1929
9f	Cpe	4-OMe	75.0	343.2103	343.2129
9g	Cpe	3-OMe	55.2	343.2117	343.2129
9h	Cpe	2-OMe	62.2	343.2147	343.2129
9i	Cpr	4-Cl	85.5	319.1335	319.1320
9j	Cpr	4-F	51.3	303.1612	303.1616
9k	Cpr	3-F	63.0	303.1619	303.1616
9l	Cpr	4-OMe	100	315.1819	315.1816
9m	Cpr	3-OMe	96.6	315.1825	315.1816

The successful reduction of the triple bond was confirmed by the methods of NMR and IR spectroscopy, as well as mass spectrometry. Melting points were also determined for each compound. The assignments of the peaks in the NMR spectra were made with reference to the numbering system defined in figure 3.13. ^{1}H and ^{13}C NMR spectra of formed product **9i**

are included in figures 3.14 and 3.15, respectively, with the assignments of the various spectroscopic signals indicated.



9

Figure 3.13: General numbering system for hydrogenated products (**9**)

In the ^{1}H NMR spectra, the peak assigned to the hydrogens on C3' shows a slight shift upfield from the starting materials, appearing now at approximately 4.0 ppm. Additional splitting is also seen for this peak, which resolves as a triplet integrating to two protons. Two new peaks are evident in the ^{1}H NMR spectra of the formed compounds and were assigned to the hydrogens on carbons C1' and C2'. These two peaks are observed at approximately 2.7 ppm and 2.0 ppm, respectively, and integrate for two protons each. They show the expected multiplicity (triplet and quintet) and were assigned based on this and the COSY spectra.

The ^{13}C NMR spectroscopic signals for the C1' and C2' carbons shift drastically upfield after the reduction of the triple bond. These two peaks are found at around 22 ppm and 29 ppm, respectively, and were assigned with the aid of HSQC spectra. The C3' signal is also shifted slightly upfield, to a new position of around 67 ppm. The C5 carbon on the pyrimidine ring shifts slightly downfield from the starting materials to a new position of approximately 105 ppm. The remaining ^{1}H NMR and ^{13}C NMR spectroscopic signals show no notable changes from the starting materials.

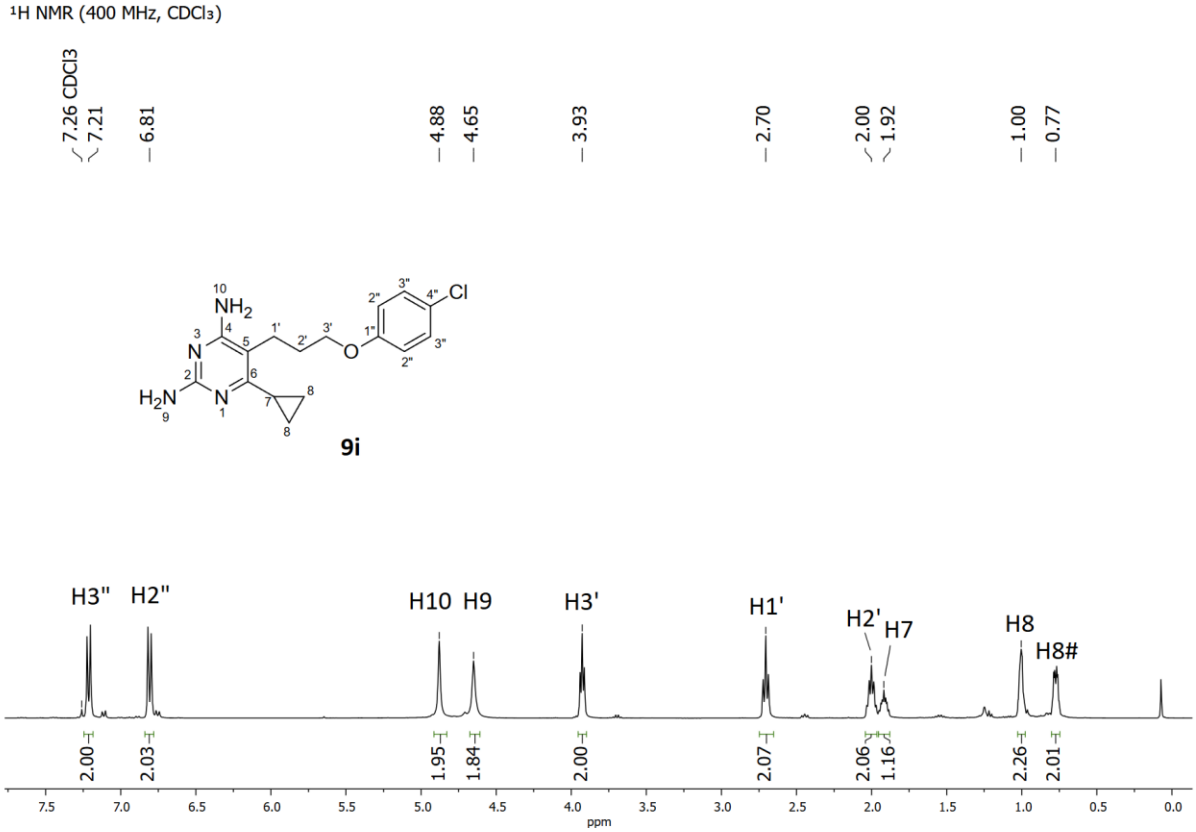


Figure 3.14: ^{1}H NMR spectrum of **9i** with spectroscopic signal assignments included

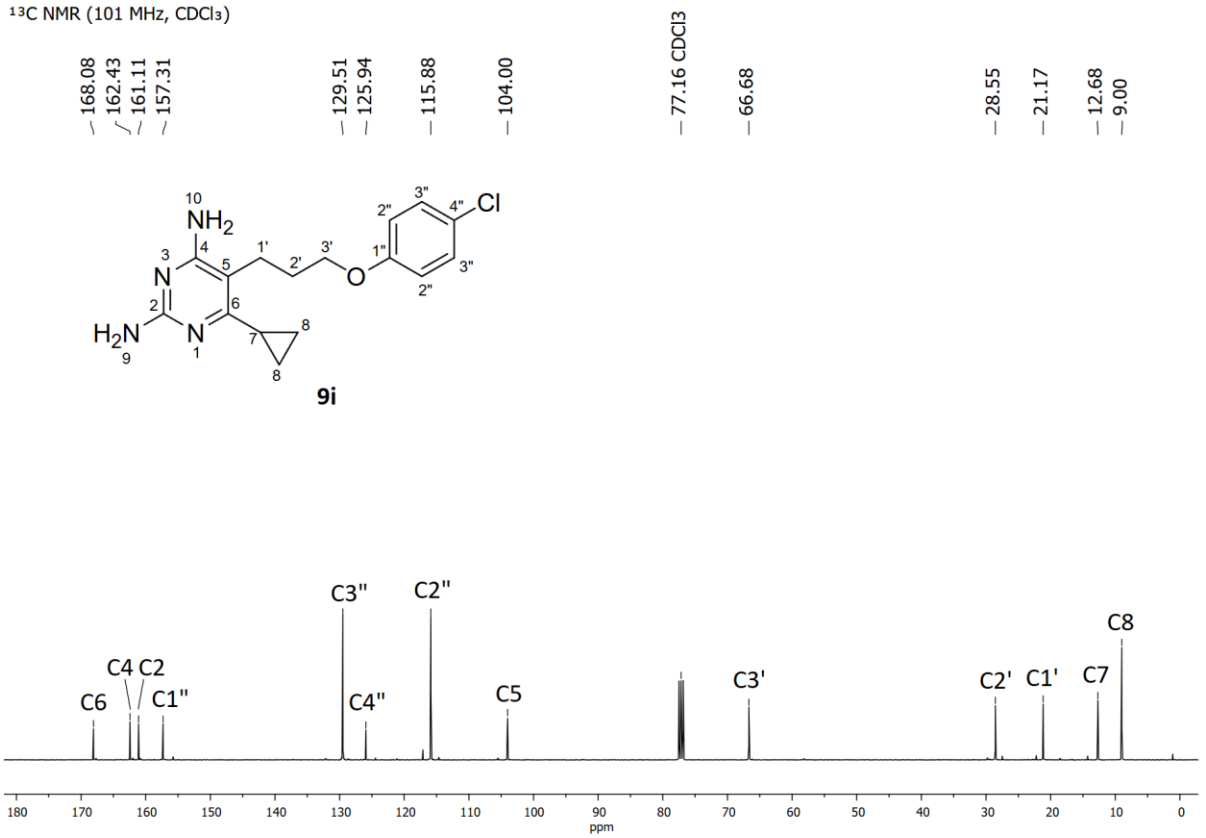


Figure 3.15: ^{13}C NMR spectrum of **9i** with spectroscopic signal assignments included

IR spectroscopy shows the absence of the key alkyne stretching peak, which is observed for the unsaturated starting materials at around 2200 cm^{-1}. Mass spectrometry reveals a prominent $[M+H]^+$ peak in positive ion mode. For each compound, the m/z values are greater than those of the starting materials by an amount equal to the mass of four protons, confirming the complete reduction of the triple bond. The observed and expected massed for each formed product are included in table 3.3. Melting points were also recorded for each analogue and differ slightly from those obtained for the starting materials.

3.2 Biological assessments of synthesized flexible pyrimethamine analogues

Activity data for a range of synthesized compounds against *P. falciparum* were obtained from collaborators at the BIOTEC research group in Thailand. Both single enzyme inhibition and whole cell assays were conducted. The enzyme inhibition assays targeted wild-type and quadruple mutant *Pf*DHFR enzymes, along with the analogous human *Hs*DHFR enzyme. Whole cell assays were performed against *P. falciparum* strains carrying either wild-type (TM4/8.2 strain) or quadruple mutant (V1/S strain) DHFR enzymes. Additional assays were conducted on mammalian Vero cells and cancerous KB cells to assess cytotoxicity. The experimental details of all the performed assays can be found in chapter 6.

The library of synthesized compounds submitted for biological evaluation was obtained through a collaborative effort in our laboratories. The general structure is provided in figure 3.16. Based on the results of the previously conducted molecular modelling experiments, cyclopentyl, cyclopropyl, *tert*-butyl, and isopropyl R-groups were included, while variations in the X-position consisted of chloro, fluoro, and methoxy substituents in the 2, 3, or 4-positions of the terminal ring.

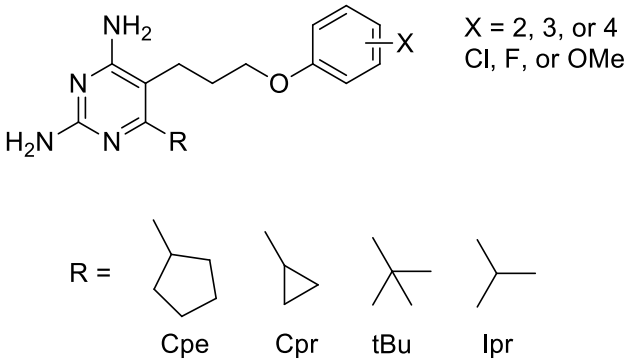


Figure 3.16: General scaffold and range of functional groups for the full set of investigated fully flexible compounds

The cyclopentyl and cyclopropyl analogues were synthesized as part of the present work, while the *tert*-butyl and isopropyl compounds were made by Dr Kamogelo Butsi as part of the broader research project. Discussion of the biological results obtained for these additional compounds is included in this section because they are part of the full range of analogues designed in the molecular modelling work, which was conducted as part of this study. The inclusion of these results allows for a more comprehensive analysis of structure-activity relationships and a better comparison to the predicted modelling outcomes.

The alkyne intermediates **8**, described in figure 3.17, involved in the synthesis of the final fully flexible analogues were also included in the initial biological assessments. While these compounds were not studied in the already conducted molecular modelling portion of this work, sufficient material was on hand to include in the testing, and variations in the degree of flexibility of the four-atom linker might provide insight into the role this group plays in binding to mutant enzymes. In addition, a series of related compounds, included in figure 3.13, with a propargyl linker have shown promise as antibacterial antifolates[39]. Similarities between these compounds and the alkyne intermediates further justified their inclusion in the preliminary biological studies.

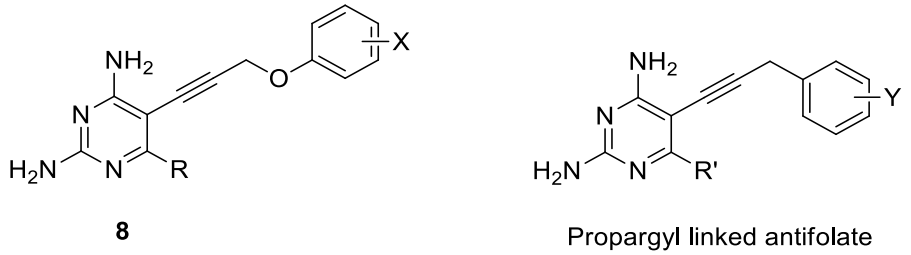


Figure 3.17: General structures of alkyne intermediates (**8**) and propargyl linked antifolates

A total of 22 final compounds (**9**) were assessed, as well as all of the 30 alkyne intermediates (**8**) formed. Due to challenges in the purification of the final products obtained in the last step of the synthesis, several fully reduced compounds were not suitable for inclusion in the biological studies. In the case of the *tert*-butyl analogues, the final reduction step was very low yielding, allowing for only two of the fully reduced *tert*-butyl analogues to be obtained in sufficient yield and purity for biological testing.

The full list of compounds included in the various assays is presented in table 3.4, along with the general structures of the alkane and alkyne analogues contained in figure 3.18. Compounds are assigned abbreviated names based on their structural elements. The analogues whose synthesis is not included as part of this work are indicated with an asterisk.

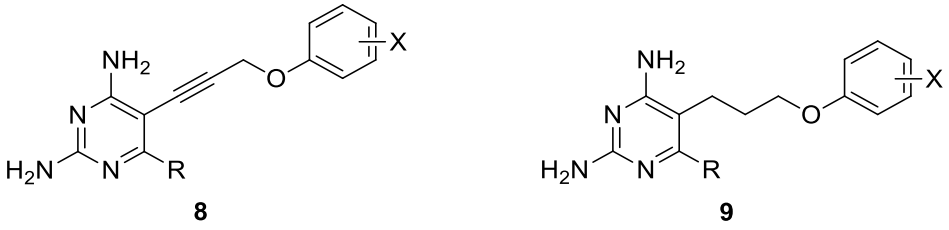


Figure 3.18: General structure of alkyne (**8**) and alkane (**9**) analogues for biological evaluation

Table 3.4: Full range of compounds evaluated for biological activity

Number:	R:	X:	Name:	Number:	R:	X:	Name:
8a	Cpe	4-Cl	Cpe-y-4Cl	9a	Cpe	4-Cl	Cpe-a-4Cl
8b	Cpe	3-Cl	Cpe-y-3Cl	9b	Cpe	2-Cl	Cpe-a-2Cl
8c	Cpe	2-Cl	Cpe-y-2Cl	9c	Cpe	4-F	Cpe-a-4F
8d	Cpe	4-F	Cpe-y-4F	9d	Cpe	3-F	Cpe-a-3F
8e	Cpe	3-F	Cpe-y-3F	9e	Cpe	2-F	Cpe-a-2F
8f	Cpe	2-F	Cpe-y-2F	9f	Cpe	4-OMe	Cpe-a-4OMe
8g	Cpe	4-OMe	Cpe-y-4OMe	9g	Cpe	3-OMe	Cpe-a-3OMe
8h	Cpe	3-OMe	Cpe-y-3OMe	9h	Cpe	2-OMe	Cpe-a-2OMe
8i	Cpe	2-OMe	Cpe-y-2OMe	9i	Cpr	4-Cl	Cpr-a-4Cl
8j	Cpr	4-Cl	Cpr-y-4Cl	9j	Cpr	4-F	Cpr-a-4F
8k	Cpr	3-Cl	Cpr-y-3Cl	9k	Cpr	3-F	Cpr-a-3F
8l	Cpr	4-F	Cpr-y-4F	9l	Cpr	4-OMe	Cpr-a-4OMe
8m	Cpr	3-F	Cpr-y-3F	9m	Cpr	3-OMe	Cpr-a-3OMe
8n	Cpr	4-OMe	Cpr-y-4OMe	9n*	tBu	4-Cl	tBu-a-4Cl
8o	Cpr	3-OMe	Cpr-y-3OMe	9o*	tBu	4-OMe	tBu-a-4OMe
8p*	tBu	4-Cl	tBu-y-4Cl	9p*	Ipr	4-Cl	Ipr-a-4Cl
8q*	tBu	3-Cl	tBu-y-3Cl	9q*	Ipr	2-Cl	Ipr-a-2Cl
8r*	tBu	4-F	tBu-y-4F	9r*	Ipr	4-F	Ipr-a-4F
8s*	tBu	3-F	tBu-y-3F	9s*	Ipr	3-F	Ipr-a-3F
8t*	tBu	2-F	tBu-y-2F	9t*	Ipr	4-OMe	Ipr-a-4OMe
8u*	tBu	4-OMe	tBu-y-4OMe	9u*	Ipr	3-OMe	Ipr-a-3OMe
8v*	tBu	3-OMe	tBu-y-3OMe	9v*	Ipr	2-OMe	Ipr-a-2OMe
8w*	Ipr	4-Cl	Ipr-y-4Cl				
8x*	Ipr	3-Cl	Ipr-y-3Cl				
8y*	Ipr	2-Cl	Ipr-y-2Cl				
8z*	Ipr	4-F	Ipr-y-4F				
8α*	Ipr	3-F	Ipr-y-3F				
8β*	Ipr	4-OMe	Ipr-y-4OMe				
8γ*	Ipr	3-OMe	Ipr-y-3OMe				
8δ*	Ipr	2-OMe	tBu-y-2OMe				

*Analogues prepared by Dr K Butsi as part of the broader project.

3.2.1 Single enzyme inhibition assays

Inhibition constants (Ki) were reported for the full range of compounds, along with the standards pyrimethamine (pyr) and cycloguanil (cyc) in wild-type *Pf*DHFR, quadruple mutant *Pf*DHFR, and *Hs*DHFR *in vitro* assays by monitoring NADPH depletion after adding various concentrations of the investigated compounds in the presence of the relevant enzyme, dihydrofolate, and the cofactor NADPH. The quadruple mutant *Pf*DHFR enzyme bears N51I/C59R/S108N/I164L substitutions which are responsible for the loss in activity of existing class II antifolates. A discussion of the results is provided below, along with the inhibition constants of a few select compounds and graphical representations of the full range of data. All of the compounds assessed inhibited both wild-type and quadruple mutant *Pf*DHFR enzymes in the low nanomolar range and showed good selectivity for parasite enzymes over the human *Hs*DHFR enzyme. The full range of data obtained in the single enzyme inhibition assays are contained in appendix C.

Table 3.5 includes the results for ten of the top performing compounds in the initial assays, compared with the standards pyrimethamine and cycloguanil. The compounds not synthesized as part of this work are indicated with an asterisk. Interestingly, both alkyne and alkane linkers are represented in the top few compounds, as well as all of the R-groups and X-substituents included for the different analogues.

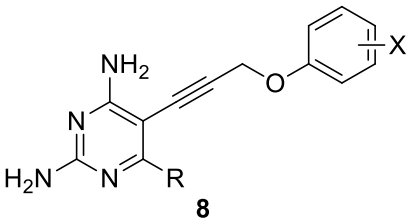
Table 3.5: Biological data for select compounds in single enzyme assays

Number:	**Name:**	**Ki (nM) *Pf*DHFR-WT:**	**Ki (nM) *Pf*DHFR-QM:**	**Ki (nM) *Hs*DHFR:**
8g	Cpe-y-4OMe	0.27 ± 0.02	1.45 ± 0.09	16 ± 0.2
8r*	tBu-y-4F	0.29 ± 0.02	1.28 ± 0.06	27 ± 0.4
8u*	tBu-y-4OMe	0.26 ± 0.02	0.85 ± 0.04	30 ± 1.6
8δ*	Ipr-y-3OMe	0.30 ± 0.01	1.40 ± 0.03	25 ± 0.6
9d	Cpe-a-3F	0.41 ± 0.01	1.07 ± 0.15	137 ± 7
9g	Cpe-a-3OMe	0.27 ± 0.02	0.82 ± 0.02	64 ± 1
9i	Cpr-a-4Cl	0.31 ± 0.01	13.0 ± 0.50	56 ± 4
9m	Cpr-a-3OMe	0.35 ± 0.02	7.10 ± 0.30	171 ± 5

Number:	Name:	Ki (nM) *Pf*DHFR-WT:	Ki (nM) *Pf*DHFR-QM:	Ki (nM) *Hs*DHFR:
9q*	Ipr-a-2Cl	0.36 ± 0.01	0.43 ± 0.02	170 ± 5
9v*	Ipr-a-2OMe	0.69 ± 0.02	0.58 ± 0.02	630 ± 32
	Cyc	0.90 ± 0.04	222 ± 18.9	235 ± 15
	Pyr	0.35 ± 0.01	233 ± 57.3	127 ± 15

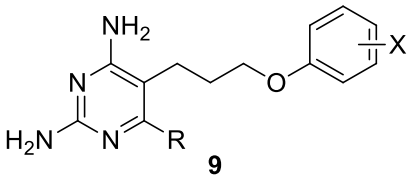
Error values are expressed as the standard deviation (SD), with the number of experiments (n) = 3.

The top performing molecule in the wild-type single enzyme inhibition assay was the tBu-y-4OMe analogue (**8u***), with a Ki value of 0.26 nM. Apart from a few exceptions, the evaluated compounds (**8a – δ** and **9a – v**) all inhibited the wild-type *Pf*DHFR enzyme in concentrations of less than 1 nM, as seen from the full range of results in appendix C. This is in the same order of activity as the standards cycloguanil (0.90 nM) and pyrimethamine (0.35 nM), indicating high levels of potency against this enzyme for the project compounds. A summary of the activities of the alkane and alkyne analogues with different R-groups against the wild-type enzyme is provided in figure 3.19.



8

R = Cpe: 0.32 - 1.80 nM
R = Cpr: 0.39 - 0.92 nM
R = tBu: 0.26 - 2.00 nM
R = Ipr: 0.28 - 0.60 nM



9

R = Cpe: 0.27 - 5.70 nM
R = Cpr: 0.31 - 0.52 nM
R = tBu: 0.36 - 0.99 nM
R = Ipr: 0.36 - 0.69 nM

Figure 3.19: Summary of inhibition constants (Ki) obtained against WT-*Pf*DHFR, grouped by linker type and R-group

Little distinction can be made between the Ki values for the alkynes and the corresponding alkane compounds, suggesting that in the absence of any mutations both types of linkers

could facilitate effective binding in the active site. On average, the isopropyl R-groups showed the lowest Ki values, followed by the *tert*-butyl, cyclopropyl, and finally cyclopentyl groups, although this trend is not universal. All of the different X-substituents included also performed well, displaying similar Ki values between the different functional groups and positions around the ring. Figure 3.20 displays the full range of data for all the compounds assessed as well as the standards cycloguanil and pyrimethamine in the wild-type single enzyme *Pf*DHFR inhibition assay.

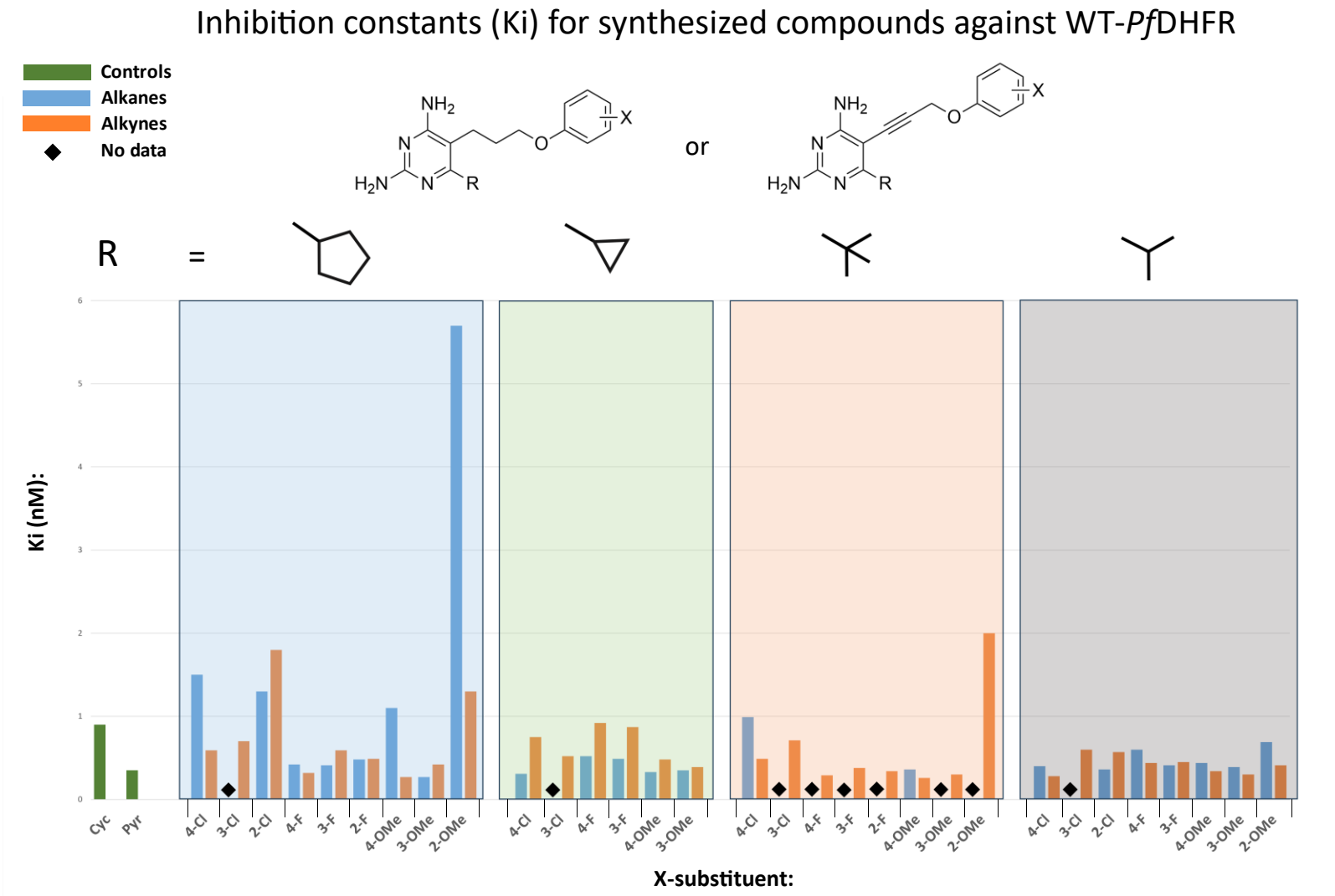
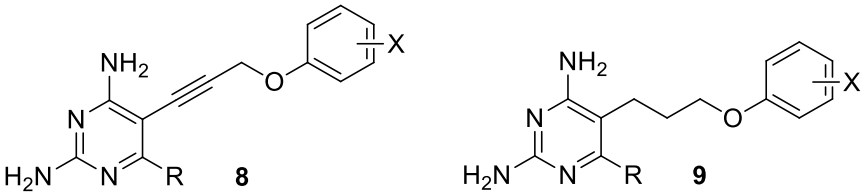


Figure 3.20: Inhibition constants (Ki) for synthesized compounds against WT *Pf*DHFR in single enzyme inhibition assays

In the quadruple mutant *Pf*DHFR single enzyme inhibition assay, all of the assessed compounds displayed activity in low nanomolar concentrations, albeit slightly higher than in the wild-type enzyme assay. This is not unexpected, as the quadruple mutant enzyme has mutations that alter the shape of its active site, reducing inhibitor binding affinity and resulting in higher activity compared to the wild-type enzyme. The compound with the lowest Ki in the quadruple mutant *Pf*DHFR assay was the Ipr-a-2Cl analogue (**9q***), which inhibited the drug-resistant enzyme at a concentration of 0.43 nM. A summary of the activities of the alkane and alkyne analogues with different R-groups against the quadruple mutant enzyme is provided in figure 3.21.



R = Cpe: 1.45 - 20.0 nM
R = Cpr: 5.70 - 20.0 nM
R = tBu: 0.78 - 38.0 nM
R = Ipr: 1.40 - 3.50 nM

R = Cpe: 0.82 - 34.0 nM
R = Cpr: 8.30 - 43.0 nM
R = tBu: 1.30 - 4.81 nM
R = Ipr: 0.43 - 1.10 nM

Figure 3.21: Summary of inhibition constants (Ki) obtained against QM-*Pf*DHFR, grouped by linker type and R-group

It can be seen from figure 3.22 that the standards cycloguanil and pyrimethamine lost much of their activity against the quadruple mutant enzyme, with inhibitory concentrations of greater than 200 nM. The fact that the project compounds remained highly active against this form of the enzyme was an excellent result as it showed that the flexible linker is able to facilitate effective binding in the active site despite the presence of the bulky mutant residues. Interestingly, both the alkyne and alkane linkers demonstrated high binding affinity for the mutant enzyme, with a few exceptions. This suggests that despite the reduced flexibility of the alkyne linkers compared to the alkanes, they were still able to adopt effective binding poses that facilitated the required interactions while avoiding steric clashes with bulky mutant residues. A similar trend in the activities of the analogues bearing different R-groups as in the

wild-type enzyme assays is seen, where the isopropyl compounds performed the best on average, followed by the *tert*-butyl, cyclopropyl, and cyclopentyl R-groups. As in the wild-type enzyme assays, no clear trend can be seen in the activities for analogues with differing X-substituents. This may suggest that these groups do not engage in any particular binding interactions or encounter significant steric clashes in the active site, so a range of functionalities around the terminal ring are tolerated for binding to the target enzyme. Figure 3.22 displays the full range of data for all the compounds assessed as well as the standards cycloguanil and pyrimethamine in the quadruple mutant single enzyme *Pf*DHFR inhibition assay.

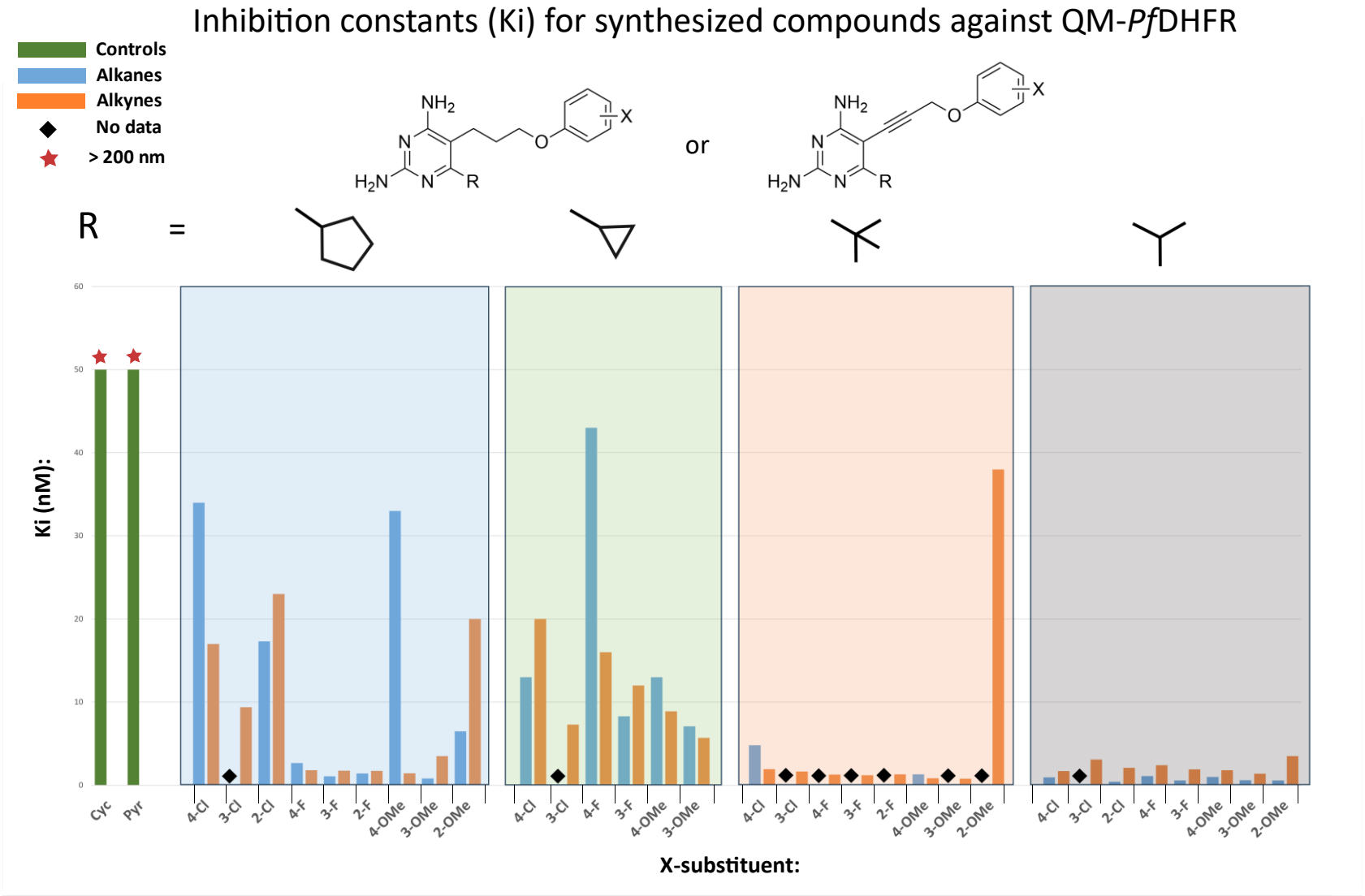


Figure 3.22: Inhibition constants (Ki) for synthesized compounds against QM *Pf*DHFR in single enzyme inhibition assays

The results obtained from the quadruple mutant single enzyme assays can be compared to the modelling results obtained in chapter 2. In the modelling experiments, a range of compounds with various aliphatic R-groups and a variety of X-substituents were evaluated for binding ability against the quadruple mutant *Pf*DHFR enzyme structure 1J3K.

Only the fully reduced final products were assessed in this way, however, due to the surprising activity displayed by the alkyne intermediates in the single enzyme assays, some additional molecular modelling experiments were conducted on this range of compounds to verify that they adhere to the same conserved binding mode as their alkane counterparts. An IFD docking experiment was conducted on the full range of alkyne compounds, using the same experimental settings as described in chapter 2. The results obtained from this confirmed that both alkane and alkyne compounds do indeed bind to the target enzyme in the same manner, with both types of analogues producing similar docking scores. This is illustrated in figure 3.23, which superimposes the top scoring docking poses of the alkane and alkyne Cpr-4Cl analogues generated from IFD experiments against the quadruple mutant *Pf*DHFR enzyme structure 1J3K.

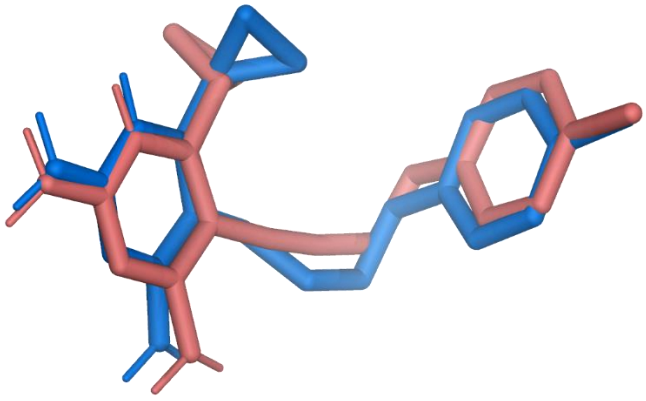


Figure 3.23: Superimposed docking poses of Cpr-4Cl alkyne (red) and Cpr-4Cl alkane (blue) analogues obtained from IFD experiments

Overall, the molecular modelling studies predicted that all of the R-groups which were selected for synthesis would allow for good binding to the target enzyme. The isopropyl and

tert-butyl R-groups performed the best in the docking experiments, and this result is reflected in the Ki values obtained experimentally in the QM *Pf*DHFR single enzyme inhibition assays. Numerous X-substituents were also evaluated, and the three groups selected for synthesis were done so based on their good docking scores as well as their favourable predicted physiochemical properties. In the modelling experiments, the chloro, fluoro, and methoxy groups all performed similarly, with little differences in the docking scores obtained for the substituents in the 2, 3, or 4-positions of the terminal ring. This was also validated in the QM *Pf*DHFR single enzyme assays, where all of these groups in the various positions around the terminal ring showed similar Ki values.

It can be concluded that the molecular modelling experiments therefore successfully predicted that the synthesized analogues would have good binding affinity against the target quadruple mutant *Pf*DHFR enzyme.

Single enzyme inhibition assays were also conducted for all the synthesized compounds against the human enzyme *Hs*DHFR in order to assess the analogues' specificity. The assessed compounds were seen to inhibit *Hs*DHFR in the nanomolar range, although in concentrations of one or two orders of magnitude higher than against the parasite enzyme. The controls cycloguanil and pyrimethamine inhibited the human enzyme in a similar concentration range to the assessed project analogues. In most cases the alkyne linker compounds showed significantly lower inhibition constants against *Hs*DHFR than their alkane counterparts. Despite the similar levels of activity for these two classes of compounds in the parasite enzyme assays, the improved specificity of the alkanes may make them more suitable drug candidates overall.

Figure 3.24 displays the full range of data for all the compounds assessed as well as the standards cycloguanil and pyrimethamine in the human *Hs*DHFR single enzyme inhibition assay.

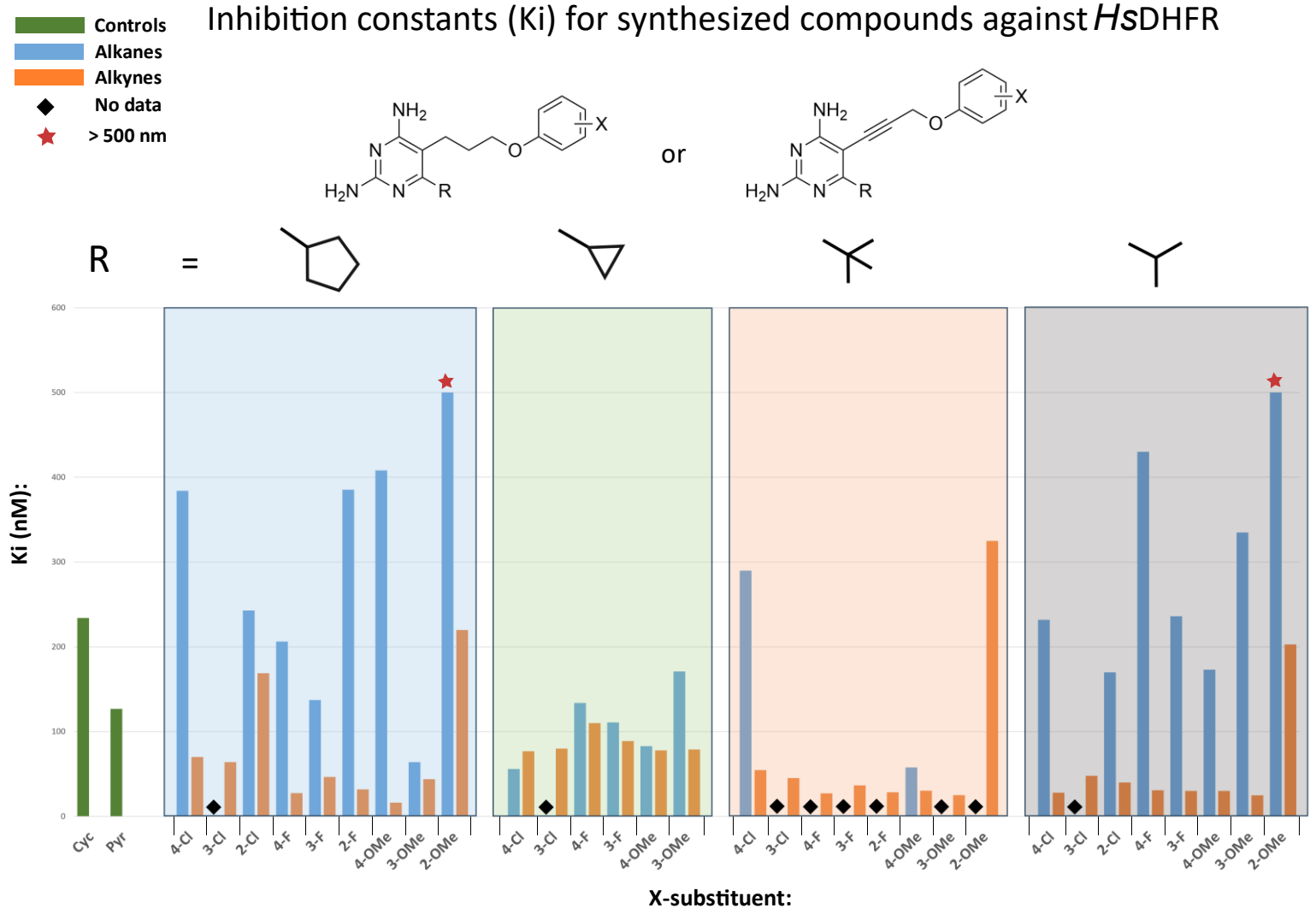


Figure 3.24: Inhibition constants (Ki) for synthesized compounds against *Hs*DHFR in single enzyme inhibition assays

3.2.2 Whole *P. falciparum* cell assays

Given the promising results obtained in the single enzyme assays, additional whole cell assays were conducted on the full range of synthesized compounds. Antiplasmodial activity was determined in ^{3}H hypoxanthine incorporation assays against two strains of *P. falciparum*: the drug-sensitive TM4/8.2 strain which carries the wild-type *Pf*DHFR enzyme; and the drug-resistant V1/S strain with N51I/C59R/S108N/I164L *Pf*DHFR mutations. Compounds were also screened for cytotoxicity against mammalian Vero cell cultures and possible anti-cancer activity against KB tumour cell lines.

The full set of results for the whole cell assays is contained in appendix C. IC_{50} values for some of the most active compounds is included in table 3.6, along with the results for the standards cycloguanil and pyrimethamine. The compounds not synthesized as part of this project are indicated with an asterisk. The selectivity index (SI) was also calculated for each compound, as the ratio of the activity against the Vero cell line to the quadruple mutant V1/S parasite strain, to provide an indication of the possible cytotoxicity of the project compounds.

Overall, a drastic decrease in activity was seen in the whole cell assays when compared to the results for the single enzyme assays. Inhibitory concentrations (IC_{50}) in the micromolar range were reported for most compounds, and many failed to inhibit the drug-resistant parasite cells at concentrations of 50 µM or less. However, the majority of compounds performed better than the standards cycloguanil and pyrimethamine (IC_{50} > 100 µM) in this assay.

Table 3.6: Inhibitory concentrations (IC_{50}) of select compounds in whole cell assays

Number:	**Name:**	***Pf*TM4/8.2 (µM):**	***Pf*V1/S (µM):**	**Vero (µM):**	**KB (µM):**	**SI:**
8l	Cpr-γ-4F	0.71 ± 0.06	4.0 ± 0.6	30 ± 10	> 50	7.5
8s*	tBu-γ-3F	1.1 ± 0.3	41 ± 7.6	12 ± 3.7	40.2 ± 8.4	0.98
8v*	tBu-γ-3OMe	0.28 ± 0.02	22 ± 3	0.7 ± 0.2	9.6 ± 3.6	0.44
8y*	Ipr-γ-3Cl	0.54 ± 0.18	45 ± 3	24 ± 8	> 50	0.53
8α*	Ipr-γ-4F	0.58 ± 0.08	50 ± 7	13 ± 3	> 50	0.26
8γ*	Ipr-γ-4OMe	0.27 ± 0.06	29 ± 7	10 ± 3	> 50	0.34

Number:	Name:	*Pf*TM4/8.2 (μM):	*Pf*V1/S (μM):	Vero (μM):	KB (μM):	SI:
9a	Cpe-a-4Cl	3.3 ± 0.7	6.0 ± 1.5	18 ± 2	22.0 ± 0.4	3.0
9b	Cpe-a-2Cl	3.1 ± 0.5	5.8 ± 1.0	14 ± 1	32.3 ± 0.9	2.4
9d	Cpe-a-3F	3.3 ± 0.3	6.5 ± 0.7	17 ± 2	21.9 ± 1.2	2.6
9g	Cpe-a-3OMe	2.6 ± 0.2	5.1 ± 0.1	15 ± 1	21.8 ± 0.6	2.9
	Cyc	0.037 ± 0.01	> 100	> 100	> 100	
	Pyr	0.086 ± 0.05	> 100	26 ± 2	60 ± 25	< 0.26

Error values are expressed as the standard deviation (SD), with the number of experiments (n) = 3.

Against the wild-type *Pf*TM4/8.2 strain, most compounds inhibited the parasite at concentrations of less than 5 μM, with all under 30 μM. The most active compound in the series was the Ipr-y-4OMe analogue (**8y***), with an IC_{50} value of 0.27 μM. Figure 3.25 displays the full range of data for all the compounds assessed as well as the standards cycloguanil and pyrimethamine in the wild-type *P. falciparum* TM4/8.2 whole cell assay. The drastic loss in activity of the synthesized compounds in the whole cell assay can be seen by comparison with the results obtained for the known standards cycloguanil and pyrimethamine. These two compounds maintained the high levels of activity seen in the single enzyme assay, with reported IC_{50} values of 0.037 μM and 0.086 μM, respectively, against the drug-sensitive *Pf*TM4/8.2 strain. Studying the data obtained for the synthesized compounds, it can be seen that the isopropyl R-group performed the best on average, with both alkyne and alkane linkers as well as the full range of X-substituents showing consistent activity of below 5 μM. The same is true for the analogues with cyclopropyl R-groups, although the IC_{50} values for these compounds were slightly higher than for the isopropyl analogues. The cyclopentyl compounds showed clear differences in activity between the alkyne and alkane containing analogues, with the alkanes inhibiting the wild-type parasite cells in lower concentrations on average than the corresponding alkynes.

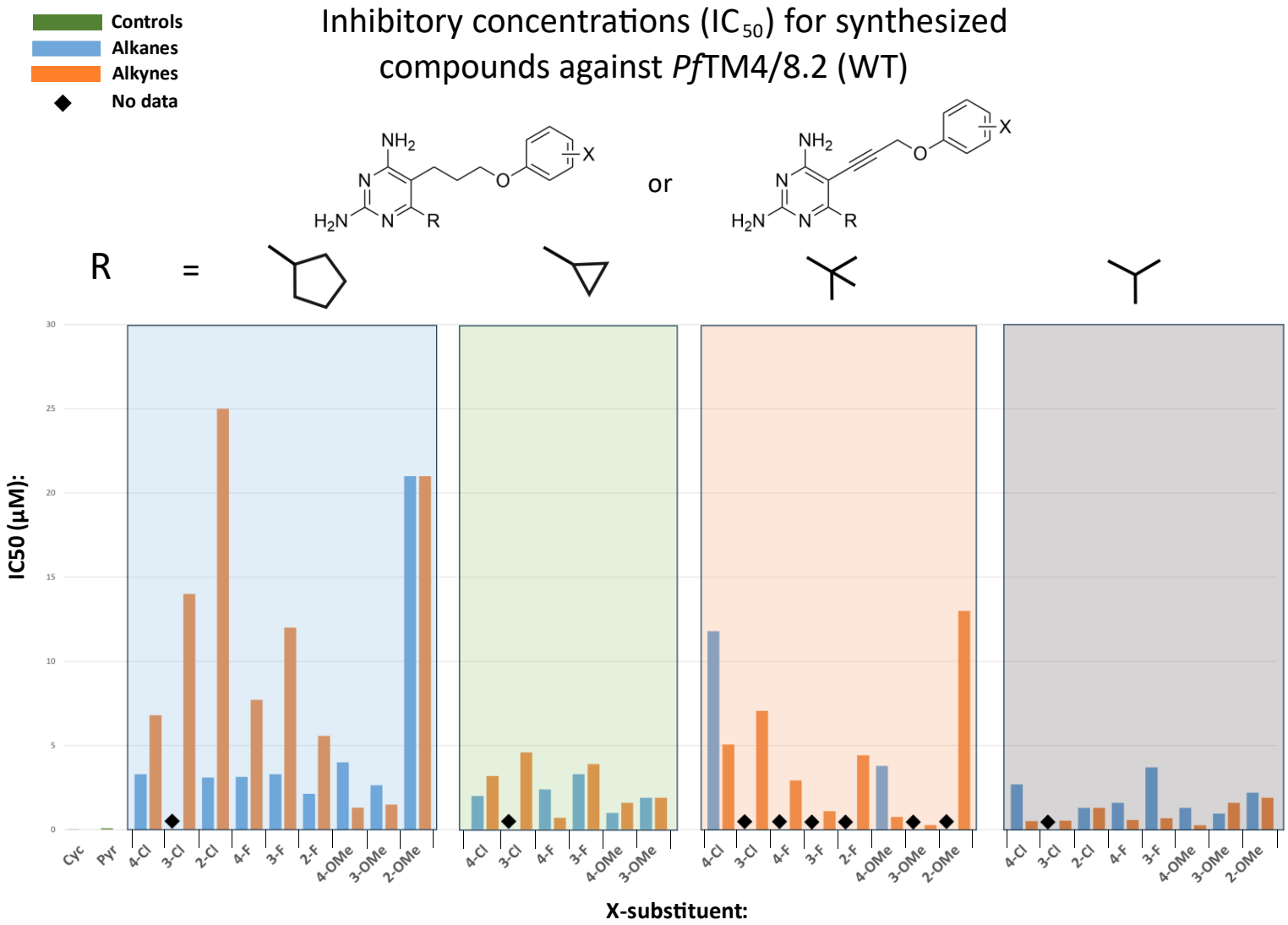


Figure 3.25: Inhibitory concentrations (IC_{50}) for synthesized compounds against wild-type *P. falciparum* strain TM4/8.2

In the whole cell assays against the drug-resistant V1/S *P. falciparum* strain, the top performing analogue was the Cpr-y-4F (**8l**) compound, with an IC_{50} value of 4.0 µM. Many of the analogues assessed in this assay failed to inhibit the growth of the parasite cells at concentrations of 50 µM. This is particularly true for the alkyne containing compounds, where only 18 of the 30 alkyne analogues showed IC_{50} values of 50 µM or less. Figure 3.26 displays the full range of data for all the compounds assessed as well as the standards cycloguanil and pyrimethamine in the quadruple *P. falciparum* V1/S whole cell assay. The alkanes displayed more consistent activity, with inhibitory concentrations of around 15 µM in general. The two standards cycloguanil and pyrimethamine were inactive in this assay, with IC_{50} values of greater than 100 µM. This was expected given that the mutant form of the *Pf*DHFR enzyme found in V1/S cell cultures is known to be highly resistant to inhibition by these compounds. All of the different X-substituents assessed performed similarly, with no clear trends observed in activity among these groups. The most active analogues in general were the cyclopentyl compounds with fully reduced alkane linkers (**9a – h**) . Out of the eight compounds of this description six displayed inhibitory concentrations of less than 10 µM, and all were less than 25 µM.

Additional whole cell assays were conducted against mammalian Vero cell lines, and cancerous KB cells to assess the potential cytotoxicity and anticancer activity of the synthesized compounds. The results for the Vero assays are displayed in figure 3.27, and the results for the KB assay are included in appendix C. The project analogues showed moderate cytotoxicity against Vero cells, in general inhibiting cellular growth in concentrations of around 30 µM. Although it is unlikely that these concentrations will be reached in systemic circulation *in vivo*. A similar trend in the activity of the alkynes vs the alkanes was seen as in the *Hs*DHFR single enzyme assays, where the alkynes proved to be more cytotoxic than their alkane counterparts. The majority of the assessed compounds failed to inhibit the growth of cancerous KB cells in concentrations of 50 µM or less. The most active compounds in these assays were the alkane containing cyclopropyl analogues, with IC_{50} values of around 30 µM.

The selectivity index (SI) for each compound is also included in table 3.6. These results display that the project compounds may have some cytotoxicity issues, with many inhibiting the Vero cells in lower concentrations than the quadruple mutant V1/S parasite strain.

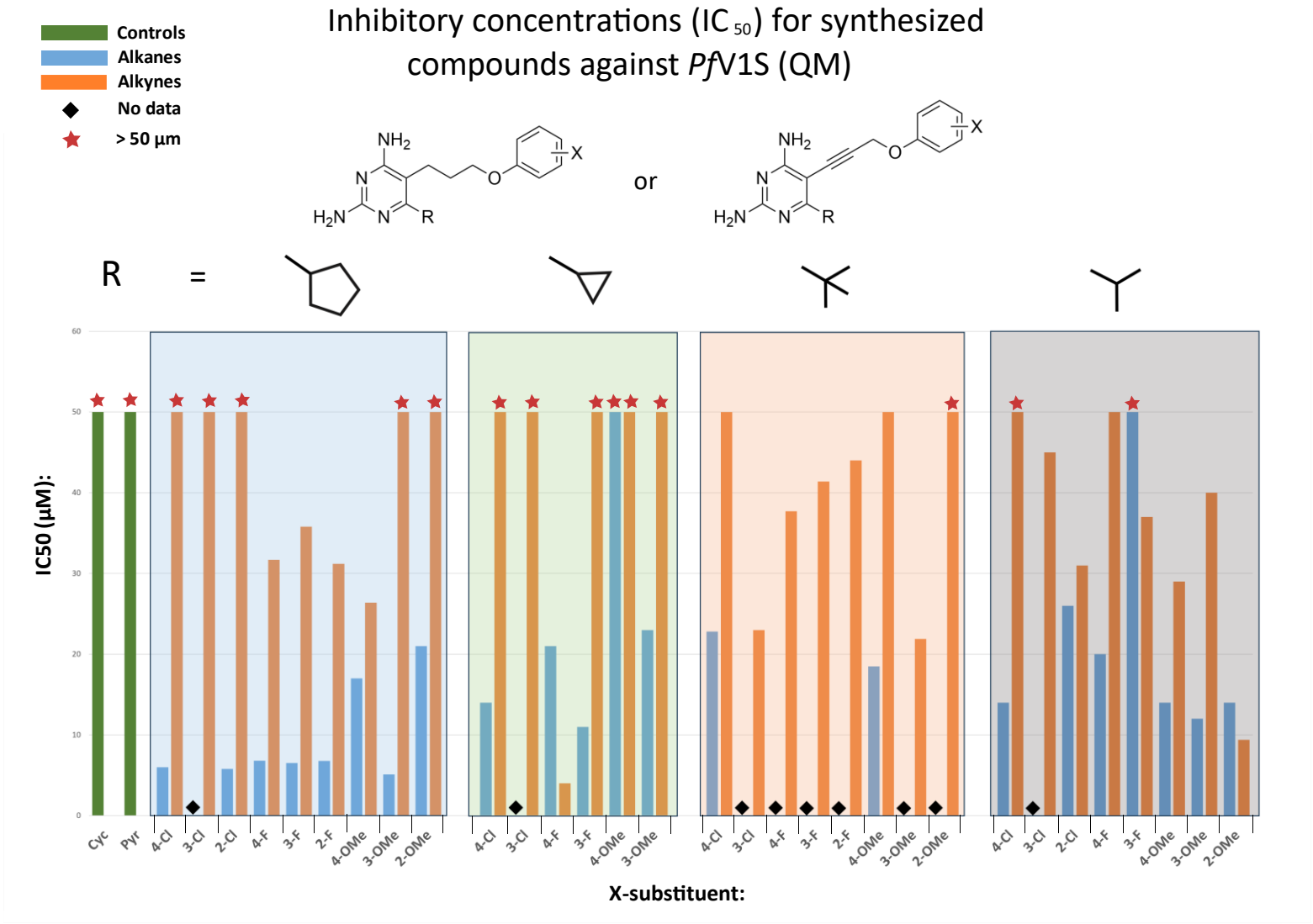


Figure 3.26: Inhibitory concentrations (IC_{50}) for synthesized compounds against quadruple mutant *P. falciparum* strain V1/S

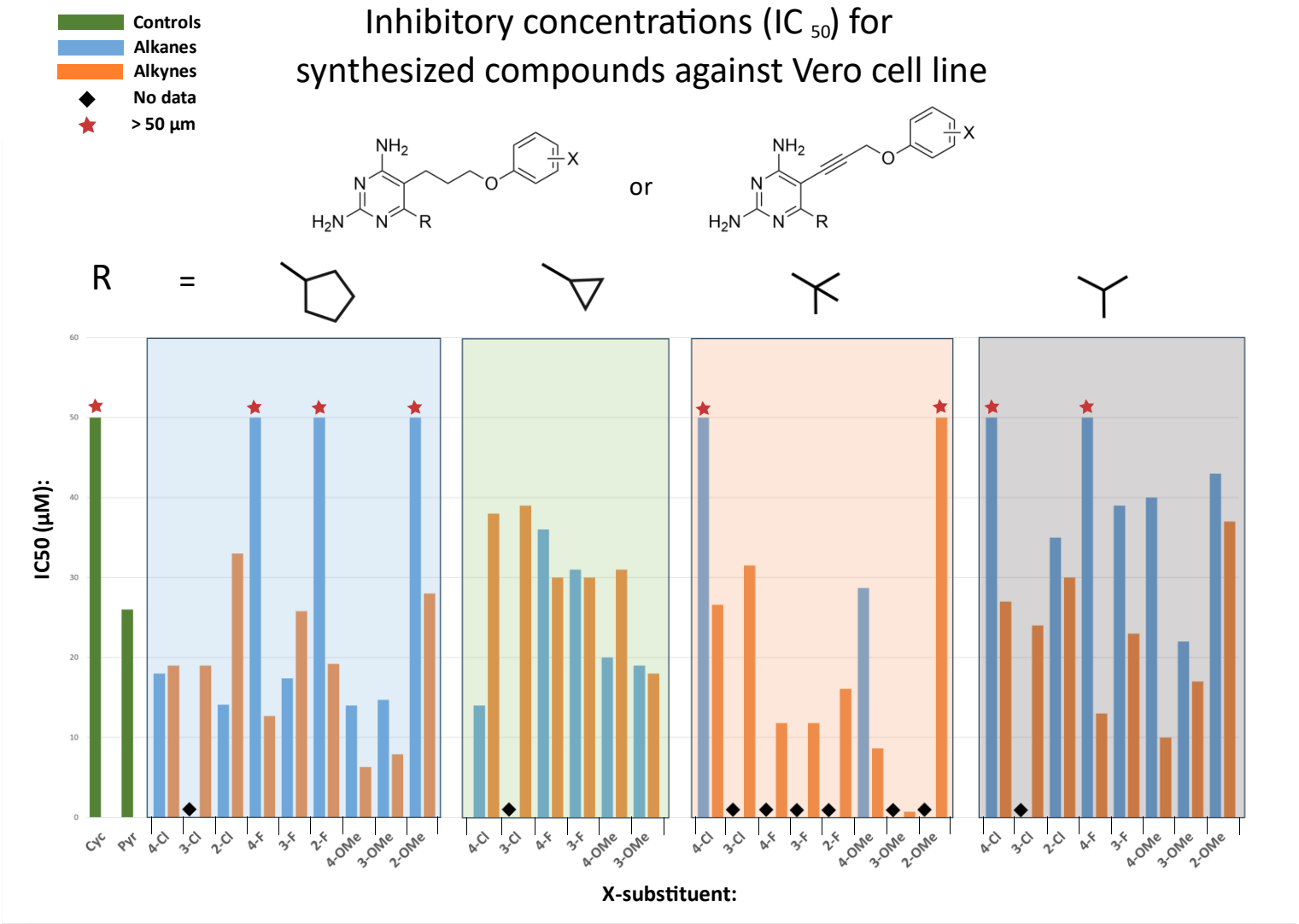


Figure 3.27: Inhibitory concentrations (IC_{50}) for synthesized compounds against mammalian Vero cell line

3.2.3 Conclusions and path forwards

Several different biological assays were conducted on the full range of compounds synthesized in this work, as well as an additional set of analogues made by Dr Kamogelo Butsi as part of the broader research project, with the aim of determining the compound's antiplasmodial activity and cytotoxicity. Single enzyme inhibition assays were performed on wild-type and quadruple mutant *Pf*DHFR enzymes, and excellent Ki values in the low nanomolar concentration range were reported for the synthesized compounds. In these assays the isopropyl analogues showed the lowest inhibition constants on average, although little distinction could be made between the activities of compounds containing alkyne and alkane linkers, and between compounds bearing various different substituents around the terminal ring. In the wild-type single enzyme assays, the assessed compounds showed activity in the same order as the known standards cycloguanil and pyrimethamine, although the synthesized analogues outperformed the standards in the drug-resistant quadruple mutant *Pf*DHFR assays. Additional single enzyme inhibition assays were conducted against human *Hs*DHFR to determine the potential cytotoxicity of the synthesized compounds. Good specificity was displayed by the project compounds, where each showed inhibition constants for the human enzyme greater than for either the wild-type or quadrupled mutant *Pf*DHFR enzymes. Notably the *Hs*DHFR Ki values for the compounds with alkyne linkers (**8a – δ**) were lower than the fully flexible alkane analogues (**9a – v**), indicating that the alkane series of compounds may be more suitable drug candidates than the alkynes due to their lower cytotoxicity.

Further whole cell assays were conducted on all of the project compounds based on the promising results obtained in the single enzyme inhibition assays. Inhibitory concentrations were measured for the compounds against wild-type TM4/8.2 and drug-resistant V1/S parasite strains, in addition to mammalian Vero cell lines and cancerous KB cells. Disappointingly, inhibitory concentrations in the micromolar range were reported for the synthesized compounds against the parasite cell cultures, representing a significant decrease in activity from the nanomolar concentrations obtained in the single enzyme assays. In the wild-type *P. falciparum* assays, the isopropyl and cyclopropyl compounds performed the best, and some improvements in the activity of the alkanes could be seen over the activity of the alkynes. In the drug-resistant *P. falciparum* assays, the compounds with the lowest IC_{50} values in general were those with fully flexible alkane linkers and containing cyclopentyl R-groups.

Seemingly only minimal differences in activity were seen between analogues with different X-substituents around the terminal ring, with no clear trends observed. The Vero mammalian whole cell assays showed moderate cytotoxicity for the compounds assessed, with the alkynes having slightly lower IC_{50} values than the alkanes. In the cancerous KB whole cell assay, most compounds showed negligible activity apart from the cyclopropyl alkanes, which had activity in the micromolar concentration range.

Overall, it can be seen that the compounds were highly potent binders of the free *Pf*DHFR enzymes, as seen in the single enzyme inhibition assays. In these assays, the nature of the R-group played a larger role in the binding affinity of the compounds than the type of linker or the X-substituents. Compounds displayed acceptable specificity for parasite enzymes over the human *Hs*DHFR equivalent. However, the synthesized compounds only inhibited parasite strains in whole cell assays in micromolar concentrations, representing around a thousand-fold loss in activity compared to the single enzyme assays. The isopropyl compounds showed the best activity, with little difference in the activity observed between X-substituents. The alkane compounds were slightly more active than the alkynes and showed better selectivity in the Vero mammalian whole cell assay.

There are many possible reasons for the observed decrease in activity from the single enzyme inhibition assays to the whole cell assays. In single enzyme assays, compounds are only required to bind strongly to the target enzyme, outcompeting the natural substrate, for good levels of activity to be measured. In the whole cell assays, a more realistic cellular environment is present where compounds need to display good pharmacokinetic properties in addition to high binding affinity for the target enzyme to inhibit the growth of the cells. The results obtained suggest that it is to do with some pharmacokinetic factors that the compounds were far less effective in the whole cell assays. Several different factors fall under the broad category of pharmacokinetics, including solubility, membrane permeability, metabolic stability, and enzyme specificity, among many others. It is likely that while the synthesized compounds are potent binders of the target enzyme, they were unable to reach their target in the cellular environment to effectively exert their desired effect.

Overall, the biological results obtained in this section for the synthesized compounds can be considered a partial success. The modelling results were validated by the excellent inhibition constants seen against the quadruple mutant *Pf*DHFR enzyme, although further work on the compounds was required to address the pharmacokinetic shortcomings which manifested in

the whole cell assays. These aspects of the project are considered in the following chapter and include potential structural modifications to the project compounds, based on functional groups present in existing drugs, which may enhance their pharmacokinetics.

Chapter 4 – Results and discussion part B: Selection, synthesis, and modelling studies of second-generation analogues

4.1: Design of second-generation analogues

While the first-generation flexible pyrimethamine analogues showed excellent activity in single enzyme *Pf*DHFR inhibition assays, they lost much of their activity against the parasite in the conducted whole cell assays. This suggests that although the compounds are effective binders of the target enzyme, they are unable to effectively reach their target in the cellular environment due to pharmacokinetic limitations. Extensive molecular modelling studies were carried out on the first-generation analogues, which correctly predicted that they would show excellent affinity for the target enzyme but failed to account for the lower activity observed in the whole cell assays. As part of initial modelling studies, numerous physiochemical properties of the target compounds were predicted, with each falling inside the expected ranges for known drug molecules. The fact that the experimentally obtained results for the whole cell assays indicate that there are indeed some pharmacokinetic shortcomings of the assessed compounds calls the reliability of the predicted physiochemical properties into question. For this reason, the design of a second-generation of flexible pyrimethamine analogues could not be based on predicted modelling outcomes. Furthermore, an *in vitro* assessment of the pharmacokinetic properties of the first-generation compounds was too costly to be considered for this project. Instead, several known drugs with similar structures to the compounds investigated in this work were considered, and functional groups which have shown to be effective at inhibiting their targets, as well as displaying favourable pharmacokinetics, were incorporated into the general structure of the project analogues. Several structurally related known antifolate drugs are included in figure 4.1.

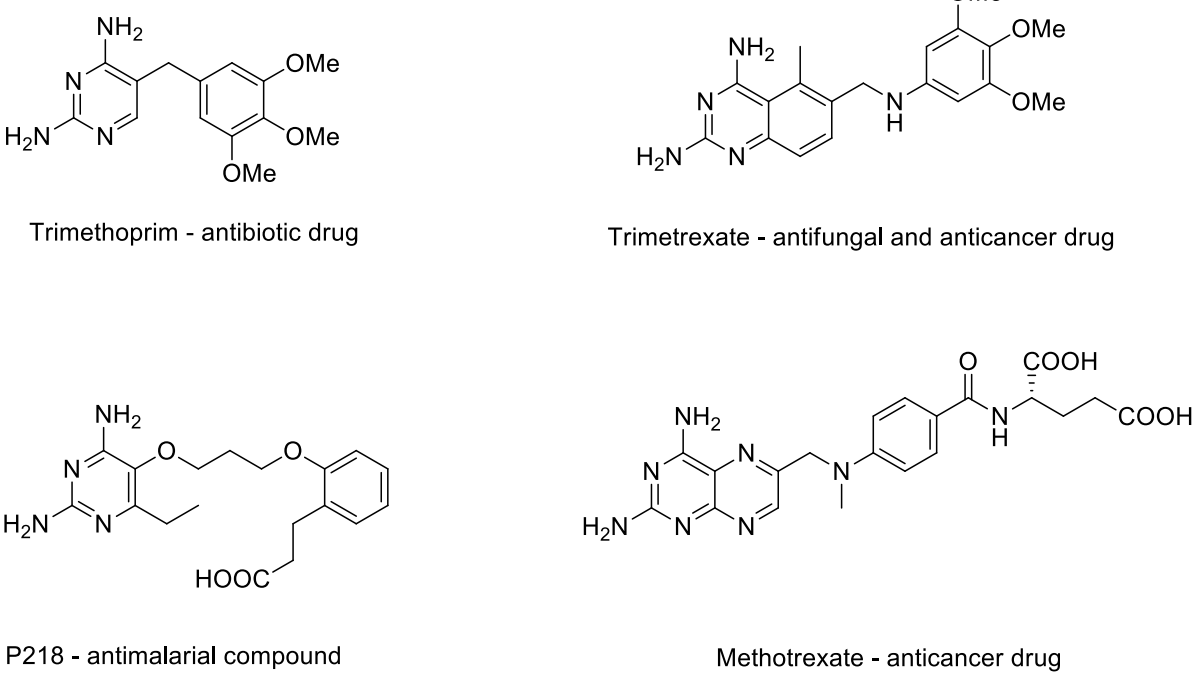


Figure 4.1: Structures of effective known antifolate compounds

Trimethoprim and trimetrexate, shown in figure 4.1, are antifolate compounds which have been used in the treatment of various diseases for a number of years. They both contain the conserved 2,4-diaminopyrimidine core, and have a 3,4,5-trimethoxy substituted aromatic ring. While the linker portion of these compounds differs from the structures investigated in this work, the trimethoxy group can be incorporated into the general scaffold to hopefully improve the compounds' pharmacokinetic properties. Although the first-generation project compounds included analogues with 2, 3, or 4-methoxy aromatic substituents, trimethoxy aromatic substitutions are a common moiety in medicinal chemistry and may improve the solubility and oral bioavailability of the investigated compounds[74]. Furthermore, trimetrexate has a nitrogen atom attached directly to the substituted aromatic ring. The first-generation analogues described in the previous section all contained an oxygen atom in the equivalent position. Substituting the oxygen in this position for nitrogen may lead to further improvements in the compounds' pharmacokinetics by modulating the overall pKa values of the target compounds.

With this in mind, two new project compounds were envisaged, the structures of which are included in figure 4.2. The first is very similar to the first-generation analogues, differing only

by the inclusion of the 3,4,5-trimethoxy substituent around the terminal aromatic ring. For the second compound, the 3,4,5-trimethoxy substituted aromatic ring is also included, in addition to having a nitrogen atom instead of oxygen in the 4-position of the flexible linker.

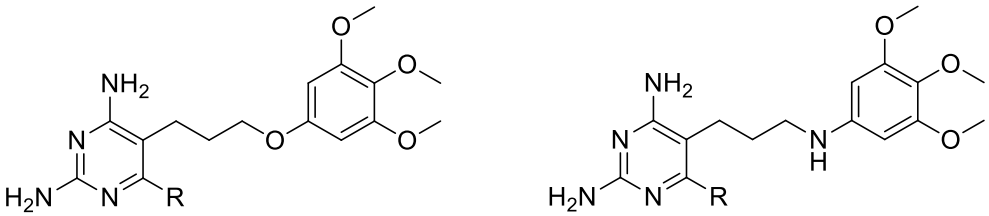


Figure 4.2: Proposed trimethoxy containing second-generation analogues

P218, included in figure 4.1, is a promising antimalarial candidate which has already been discussed in chapter 1. This compound contains the conserved 2,4-diaminopyrimidine core, a five-atom flexible linker, and a propionic acid substituent in the 2-position of the terminal ring. P218 is currently undergoing various early-stage drug trials and has shown excellent activity against the target enzyme and promising pharmacokinetics. This motivated the inclusion of a 2-propionic acid substituent in the general structure of the compounds investigated in this project for a new second-generation analogue. This group may allow for improved aqueous solubility and has been shown to facilitate additional binding interactions in the active site of the target enzyme. The proposed structure is included in figure 4.3.

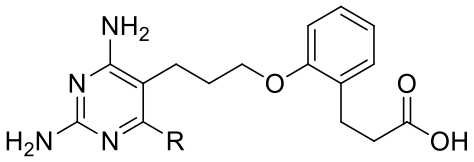


Figure 4.3: Proposed 2-propionic acid containing second-generation analogue

One last existing anti-folate compound considered is methotrexate, included in figure 4.1. This drug has been widely used in the treatment of cancer, and some other diseases, for many years. Importantly, this molecule contains a glutamic acid substituent which allows for it to

take advantage of active folate transport mechanisms to enter cells and reach the enzyme target. It is hoped that the inclusion of a glutamic acid substituent into the general scaffold of the project compounds will allow for them to likewise be able to access the active folate transport mechanisms for improved distribution to the targeted areas. In addition, it is expected that the inclusion of the glutamic acid group will enhance the compounds' aqueous solubility. Notably, methotrexate has a methylated nitrogen atom attached directly to the aromatic ring in the equivalent position of the oxygen present in the four-atom linker of the first-generation analogues. Based on this, two new structures were proposed as second-generation analogues. The first maintains the oxygen of the four-atom linker and includes the glutamic acid substituent on the terminal ring, while the second has a methylated nitrogen atom instead of oxygen in the four-atom linker, and also contains the glutamic acid substituent on the terminal ring. The two structures are shown in figure 4.4.

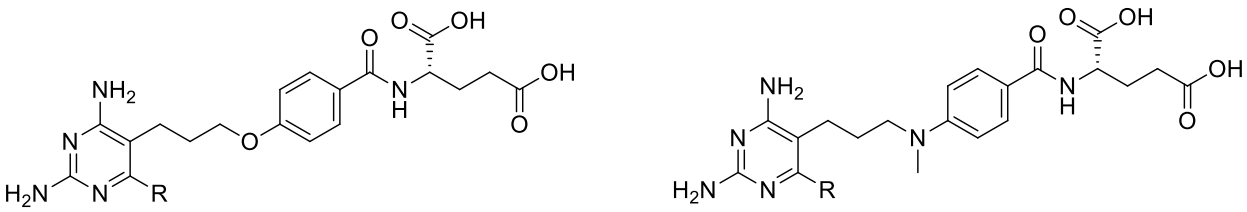


Figure 4.4: Proposed glutamic acid containing second-generation analogues

For all of the second-generation analogues, cyclopropyl was selected as the R-group in the 6-position of the 2,4-diaminopyrimidine ring. This group showed similar activity to the other assessed R-group in the whole cell assays. In addition, some other practical considerations included the fact that a large amount of the iodinated 2,4-diamino-6-cyclopropyl-pyrimdine (**7b**) intermediate was on hand after the synthesis of the first-generation analogues, and the distinct location of the NMR spectroscopic signals arising from this group would assist in the assignment of the NMR spectra of the second-generation analogues, particularly for the glutamic acid and propionic acid containing compounds which each have two additional CH_2 groups.

Overall, five new structures were proposed. It was hoped that by including functional groups present on drug molecules known to have favourable pharmacokinetic properties that some of these favourable properties would be imparted onto the project compounds. Some preliminary molecular modelling studies were conducted on the designed analogues to ensure that they were able to recreate the known binding mode against the target quadruple mutant *Pf*DHFR enzyme. The results of the preliminary studies showed that not only did all of the new compounds bind in the correct fashion, but that they also exhibited some new binding interactions that may enhance their antiplasmodial activity. A full discussion of the molecular modelling studies conducted on the second-generation analogues is included at the end of this chapter.

4.2: Proposed synthesis of second-generation analogues

In the synthesis of the proposed second-generation analogues, many of the same synthetic steps used in the preparation of the first-generation analogues could be employed. The second-generation analogues differ from the first-generation only in the substitution of the terminal ring, and for some a nitrogen atom is included in the linker instead of an oxygen. The previously prepared iodinated 2,4-diaminopyrimidines were carried over from the synthesis of the first-generation analogues and coupled with newly formed propargyl substituted anilines and phenols. An additional step was required for the formation of the glutamic acid containing analogues. The glutamic acid residue could be directly attached to a carboxylic acid group on the terminal ring in an amide coupling reaction to form the desired final compounds. The formation of the fully reduced carboxylic acid containing compound could be achieved using the same methods as employed in the synthesis of the first-generation analogues, starting from 4-carboxylic acid substituted phenol and aniline, respectively.

4.2.1: Formation of propargyl substituted phenols and anilines

Appropriately substituted phenols and anilines (**10a - e**), included in figure 4.5, were selected as the starting materials for the initial attachment of the propargyl group.

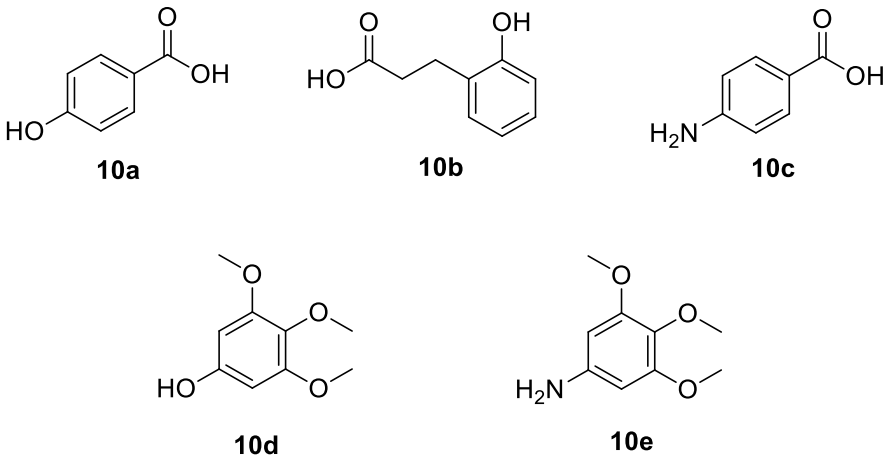
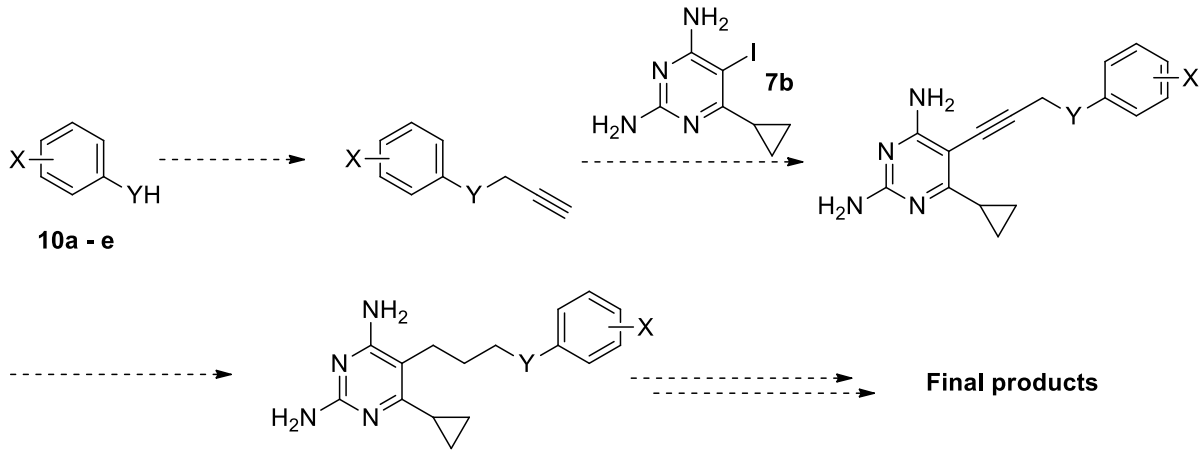


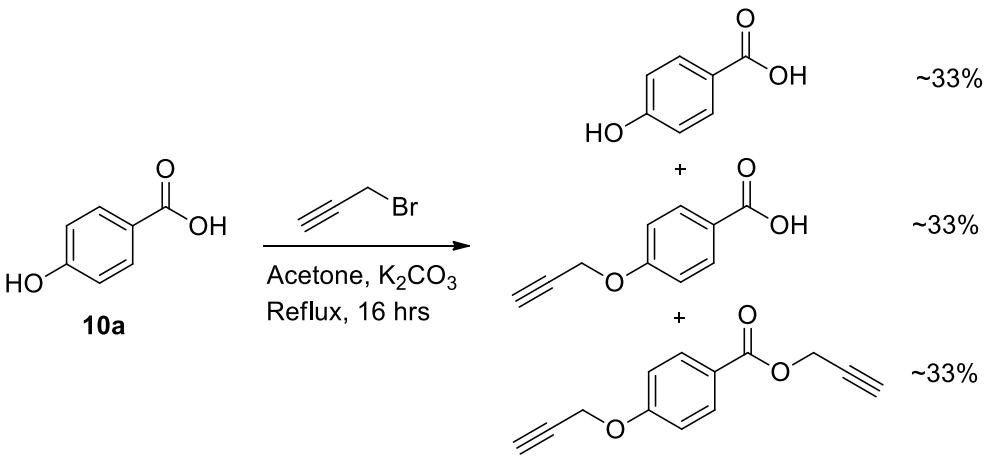
Figure 4.5: Starting materials **10a – e** in the complete synthesis of second-generation analogues

A brief review of the overall synthesis is provided in scheme 4.1, including the attachment of the propargyl group to the selected starting materials **10a – e**, the Sonogashira coupling reaction with the previously prepared iodinated 2,4-diamino-6-cyclopropyl-pyrimidine intermediate **7b**, and the subsequent hydrogenation of the triple bond. For some of the desired final compounds, additional steps were required after performing the hydrogenation reaction.



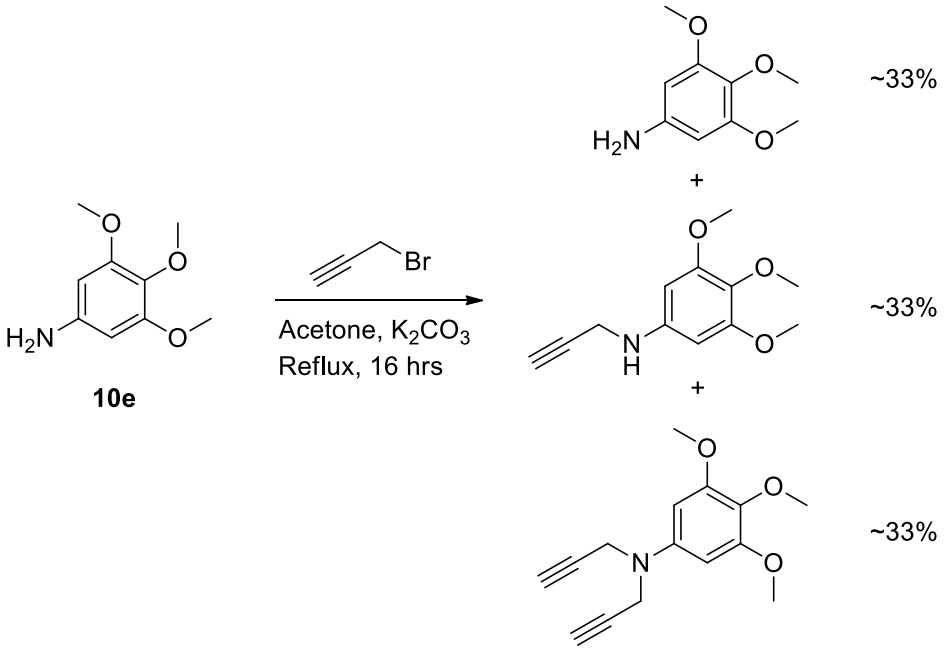
Scheme 4.1: Summary of overall synthesis for the second-generation analogues

Initial attempts to achieve the desired propargyl substitution were made using the same methods as described for this reaction in the synthesis of the first-generation analogues; namely, refluxing the starting materials in a mixture of acetone, propargyl bromide, and potassium carbonate. While these conditions afforded excellent results for the alkylation of starting material **10d**, it was found that the propargyl group would also add to the carboxylic acid oxygen of starting materials **10a**, **10b**, and **10c**, resulting in mixtures comprised of roughly equal proportions of the dialkylated product, the desired monoalkylated product, and unreacted starting material, as described with reference to compound **10a** in scheme 4.2.



Scheme 4.2: Results of alkylation of **10a** with propargyl bromide and K_2CO_3

For the aniline starting materials **10c** and **10e**, the propargyl group was also seen to add twice to the primary amine group, as described with reference to the starting material **10e** in scheme 4.3.

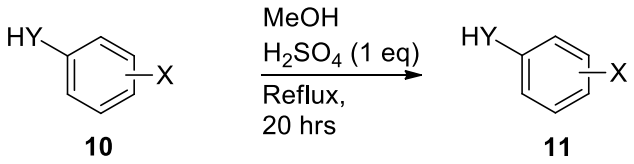


Scheme 4.3: Results of the alkylation of **10e** with propargyl bromide and K_2CO_3

For this reason, additional protection steps were required for starting materials **10a**, **10b**, **10c**, and **10e** before the propargyl group could be added. Protection of the carboxylic acid groups of **10a**, **10b**, and **10c** was achieved by esterification with methanol under acidic conditions, while the primary amines of **10c** and **10e** were protected by the addition of a *tert*-butoxycarbonyl (Boc) group. Starting material **10c** contains both a carboxylic acid group and a primary amine, so both protection steps were required for this compound.

4.2.1.1: Esterification of carboxylic acid containing phenols and anilines

The esterification of the carboxylic acid starting materials **10a**, **10b**, and **10c** was achieved through a reaction with methanol under acidic conditions, as described in scheme 4.4.



10

11

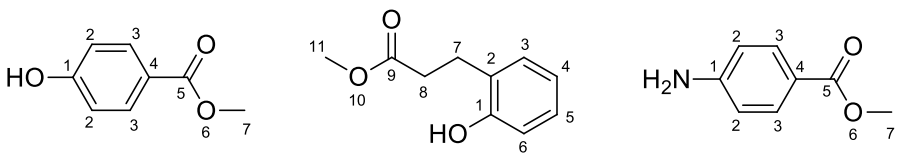
10a: Y = O, X = 4-COOH
10b: Y = O, X = 2-CH_2CH_2COOH
10c: Y = NH, X = 4-COOH

11a: Y = O, X = 4-COOMe, 97 % yield
11b: Y = O, X = 2-CH_2CH_2COOMe, 97 % yield
11c: Y = NH, X = 4-COOMe, 100 % yield

Scheme 4.4: Esterification of carboxylic acids

The respective starting materials **10a – c** were dissolved in methanol and a small amount of sulfuric acid was added. The mixture was heated to reflux, and the progress of the reaction monitored by TLC. After 20 hours, complete conversion was apparent. Compounds **11a** and **11c** were isolated after work up as white solids without the need for further purification. Compound **11b** was isolated as a yellow oil after purification with silica gel column chromatography. Excellent yields in excess of 95 % were achieved in each case.

Characterization of the formed compounds was conducted by NMR and IR spectroscopy. Compounds **11a** and **11c** are solids and melting points for these were recorded. The NMR spectroscopic assignments were made in accordance with the number system defined in figure 4.6.



11a **11b** **11c**

Figure 4.6: Numbering system for methyl carboxylate phenols and anilines (**11**)

In the ^{1}H NMR and ^{13}C NMR spectra of each product, an additional peak is evident in comparison to the spectra of the starting materials. This arises from the added OMe group and appears at approximately 3.7 ppm in the ^{1}H NMR spectra, and 52 ppm in the ^{13}C NMR spectra. For all the formed analogues, signals integrating to a total of four protons are observed in the aromatic region of the ^{1}H NMR spectra. Assignments of these peaks were made based on integration and splitting patterns, as well as the COSY spectra obtained for each product. Several ^{13}C NMR spectroscopic signals are observed in the aromatic region and were assigned with reference to HSQC and HMBC spectra which show single bond and multiple bond coupling, respectively, to the already assigned ^{1}H NMR spectroscopic signals. The carbonyl carbon peaks appear at around 166 ppm for compounds **11a** and **11c**, and at 176.12 ppm for **11b** in the ^{13}C NMR spectra. **11b** shows two additional peaks in both ^{1}H and ^{13}C NMR spectra, arising from the CH_2CH_2 group. The signals in the ^{1}H NMR spectrum resolve as multiplets with integration of two protons each. The assignments for these peaks were made based on the coupling seen in the COSY, HSQC, and HMBC spectra. The peak centered at 2.91 ppm is assigned to the hydrogens on C7, while the peak centered at 2.72 ppm is assigned to the hydrogens on C8. In the ^{13}C NMR spectra, the peaks corresponding to carbons C7 and C8 appear at 24.79 ppm and 35.08 ppm, respectively. The assignments were made from the coupling to the assigned proton signals in the HSQC spectrum.

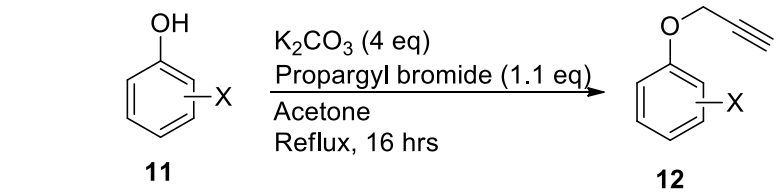
The IR spectra for each compound shows the expected peaks, including C=O, C=C aromatic, and O-H or N-H stretching modes.

11a and **11c** were isolated as white polycrystalline solids, while **11b** is a yellow oil. Metling points for **11a** and **11c** were recorded and are in line with values reported in literature[75,76]. **11a** melted in the range of 126 – 128 °C and **11c** in the range of 106 – 108 °C.

4.2.1.2 Alkylation of phenols for second-generation analogues

With intermediates **11a – c** in hand, the alkylation reaction was attempted again using the same conditions as described for this reaction in the preceding chapter. It has already been noted that these conditions worked well for the trimethoxy substituted phenol **10c**, included in figure 4.5. Good results were also obtained for esterified phenols **11a** and **11b**, forming the desired monoalkylated product in good yields. The Boc-protection step, however, was still

required for the aniline compounds **11c** and **10e** before the alkylation reaction could be performed, which is discussed later. The conditions used for the successful alkylation of **10c**, **11a**, and **11b** are described in scheme 4.5.



11

12

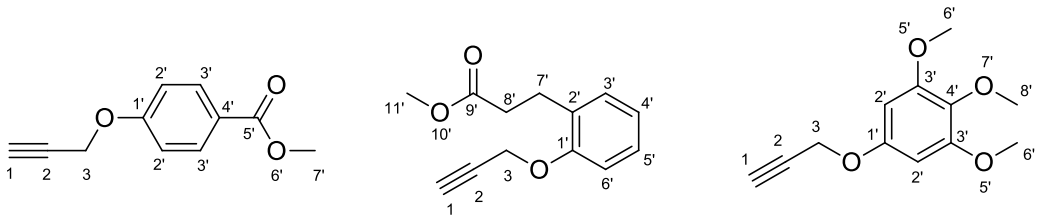
11a: X = 4-COOMe
11b: X = 2-CH_2CH_2COOMe
10c: X = 3,4,5-triOMe

12a: X = 4-COOMe, 83% yield
12b: X = 2-CH_2CH_2COOMe, 97% yield
12c: X = 3,4,5-triOMe, 90% yield

Scheme 4.5: Alkylation of phenols with propargyl bromide

The relevant starting materials were dissolved in acetone before the addition of potassium carbonate and propargyl bromide, and the mixture heated to reflux. The progress of the reaction was monitored by TLC, and after 16 hours no more starting material was observed. The formed products **12a – c** were purified with silica gel column chromatography and isolated in good yields of greater than 80 % in each case.

Successful formation of the propargyl substituted phenols **12a – c** was confirmed by NMR and IR spectroscopy. Melting points were also recorded where appropriate. The assignments of the NMR spectra were made with reference to the numbering system defined in figure 4.7.



12a **12b** **12c**

Figure 4.7: Numbering system for alkylated phenols **12a – c**

The aromatic region of the ^{1}H NMR spectra for compounds **12a** and **12b** show peaks with a total integration of four protons. The assignments for these signals was made based on the integration and splitting patterns, and the coupling observed in the COSY spectra. Intermediate **12c** contains only one signal in the aromatic region of the ^{1}H NMR spectrum, which is assigned to the hydrogens on carbon C2'. This peak is a singlet with an integration of two protons. The corresponding aromatic ^{13}C NMR spectroscopic signals were assigned based on the coupling seen in the HSQC and HMBC spectra, as well as the expected chemical shifts.

The NMR spectroscopic signals arising from the methyl ester groups present on compounds **12a** and **12b** are conserved from the starting materials, which have already been described. Intermediate **12c** shows two peaks in the ^{1}H NMR spectrum for the hydrogens on carbons C8', and the two equivalent hydrogen environments on carbons C6'. Both are singlets, with the C8' hydrogens appearing at 3.79 ppm and integrating to three protons, while the signal for the C6' hydrogens appears slightly downfield at 3.84 ppm and integrates for six protons.

Several new peaks appear in the NMR spectra of the formed compounds which confirms the successful addition of the propargyl group. The ^{1}H NMR spectrum of each product reveals peaks at around 2.5 ppm and 4.8 ppm which were assigned to the hydrogens on carbons C1 and C3, respectively. The signals for the C1 hydrogens are triplets integrating to a single proton, while the signals for the C3 hydrogens are doublets with an integration of two. The two peaks are seen to couple each other through the triple bond, with a $^{4}J_{H\text{-}H}$ coupling constant of 2.4 Hz. Three additional signals are seen in the ^{13}C NMR spectra of the products when compared to the spectra of the starting materials. Carbons C1, C2, and C3 were assigned to the signals at approximately 75 ppm, 78 ppm, and 56 ppm, respectively, based on the coupling seen in HSQC and HMBC spectra.

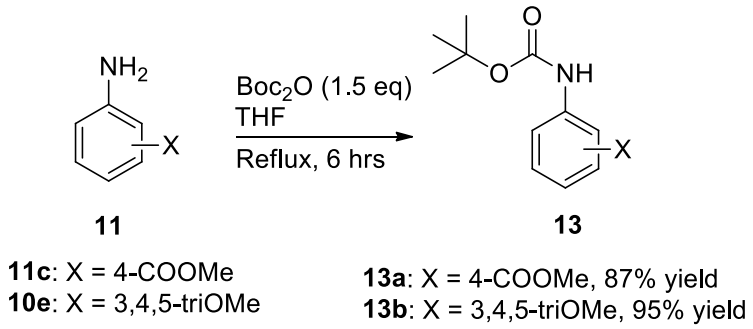
IR spectra were recorded for each compound, revealing the characteristic C≡C-H and C≡C stretching modes at around 3270 cm^{-1} and 2120 cm^{-1}, respectively.

Intermediate **12a** was isolated as a white solid, with a recorded melting point of 64 – 65 °C. This is in agreement with melting point values reported in literature[77]. Melting points were not recorded for the two other formed compounds, as **12b** was isolated as an oil, and **12c** as a semi-solid (m.p. < 30 °C).

4.2.1.3 Boc-protection of anilines

As mentioned above, a Boc-protection step was required for the aniline starting materials **10e** and **11c** to prevent the propargyl group from adding twice to the aromatic amine. Before the Boc-protection was performed, some efforts were made to selectively alkylate the aniline only once, including using lower reaction temperatures and shorter reaction times, as well as controlling the rate of addition of the propargyl bromide and using alternative bases for the reaction. However, no improvements in the roughly 33 % yield of the desired monoalkylated anilines were realized.

A literature search was conducted to uncover any possible methods for achieving the desired monoalkylation without resorting to the additional Boc-protection steps. Several methods making use of the same starting materials were found. However, for these procedures the reported yields were quite low (< 50 %), as both monoalkylated and dialkylated products were formed in the reaction, with the desired monoalkylated product isolated with silica gel column chromatography[78, 79]. It was decided to proceed with the additional Boc-protection steps in order to achieve higher yields overall for the desired monoalkylated products. Standard Boc-protection conditions were employed for the protection of anilines **10e** and **11c** and are described in scheme 4.6.



Scheme 4.6: Boc-protection of anilines

The aniline starting materials **11c** and **10e**, respectively, were dissolved in a small amount of THF, and di-*tert*-butyl dicarbonate was added before heating the mixture to reflux. It was found that satisfactory yields could be obtained for this transformation in the absence of the

base after six hours of heating. TLC was used to monitor the progress of the reaction and confirm complete conversion of the starting materials.

Characterization of the Boc-protected amines **13a** and **13b** was done by NMR and IR spectroscopy, as well as melting point analysis. No reports of compound **13b** were found in literature, so HRMS analysis was conducted for this molecule. The assignments for the NMR spectra were made in reference to the numbering system defined in the figure 4.8.

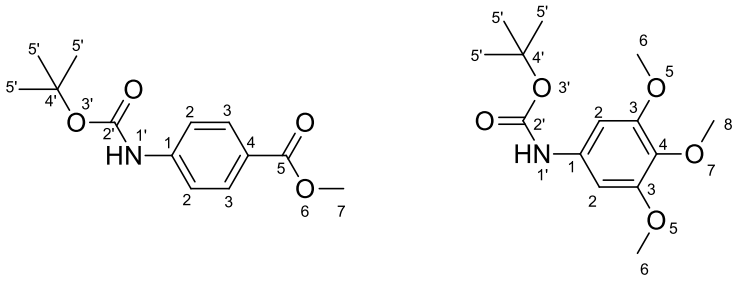


Figure 4.8: Numbering system for Boc-protected anilines **13a** and **13b**

The successful addition of the Boc group can be readily seen in the NMR spectra of the formed compounds. The ^{1}H NMR spectrum of each product reveals one additional signal when compared to the spectra of the starting materials. This peak corresponds to the hydrogens on the equivalent C5' carbon atoms and is a singlet integrating to nine protons at around 1.5 ppm. In addition, the integration of the peak for the amine protons decreases from two in the spectra of the starting materials to one in the spectra of the formed products and is seen to shift upfield by about 2.5 ppm.

Three new peaks are apparent in the ^{13}C NMR spectra of the products when compared to the spectra of the starting materials. Carbons C2', C4', and C5' were assigned to the signals appearing at approximately 152 ppm, 81 ppm, and 28 ppm, respectively. The assignments were made based on the coupling seen in the HSQC and HMBC spectra.

The IR spectra of the formed compounds include all of the expected vibrational modes, namely C=O, C=C aromatic, and C-O stretches.

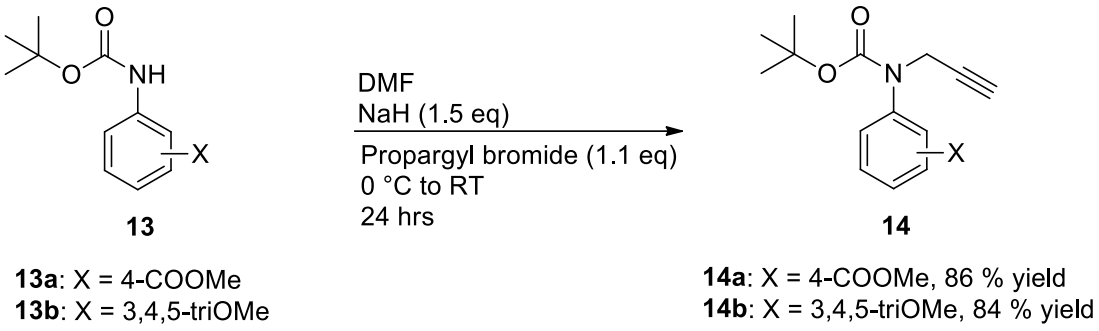
13a and **13b** were both isolated as solids, and their melting points were recorded. **13a** was found to melt between 158 and 160 °C, which is in agreement with the melting point values reported in literature for this compound[80]. The melting point for **13b** was determined to be between 150 and 151 °C.

HRMS data was collected for compound **13b**. The parent $[M+H]^+$ peak was observed at 284.1489 m/z when run in positive ion mode, which is in agreement with the expected mass of 284.1492 m/z.

4.2.1.4 Alkylation of protected anilines

After the successful Boc protection of anilines **11c** and **10e**, resulting in the formation of the corresponding protected intermediates **13a** and **13b**, the subsequent alkylation reaction could be performed. First attempts at this reaction were made using the standard conditions employed previously in this project, namely dissolving the starting materials in acetone before the addition of potassium carbonate and propargyl bromide and heating of the mixture to reflux for several hours. Reaction progress was monitored by TLC, but even after 24 hours of heating only a single spot corresponding to the starting materials was observed. The Boc-group reduces the nucleophilicity of the amine by withdrawing electron density, making it less reactive. As a result, a stronger base was required to drive the equilibrium toward complete deprotonation of the starting materials to ensure that the active nucleophile was present in sufficient quantity for the subsequent nucleophilic attack to reliably occur.

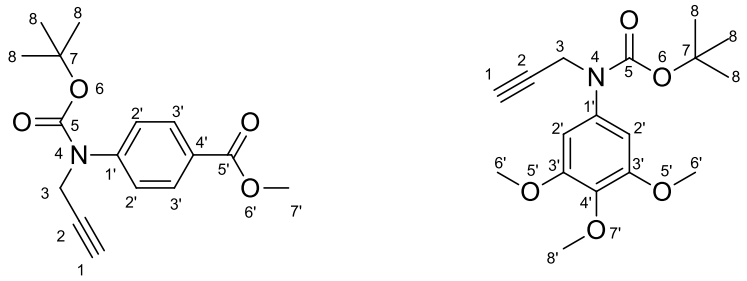
Several different combinations of solvents and bases were assessed for this transformation, with the best results obtained from mixtures of NaH, DMF, and propargyl bromide. Using these conditions, successful addition of the propargyl group onto the Boc protected intermediates **13a** and **13b** was achieved. Although there have been reports of hazards in the use of NaH with DMF, including the potential formation of explosive mixtures[81], risks were minimized by performing the reaction on at most one gram of starting material, using only 1.5 equivalents of NaH, and conducting the reaction from 0 °C to room temperature. While this allowed for safe lab-scale synthesis of the desired alkylated anilines, different conditions would need to be investigated if this step of the synthesis is to be scaled up. The optimized conditions are described in scheme 4.7.



Scheme 4.7: Alkylation of Boc-protected anilines

The relevant starting materials, **13a** and **13b**, were dissolved in DMF before cooling the mixture to 0 °C in an ice-bath. Sodium hydride was then added, and the mixture stirred at 0 °C for 20 minutes. After this period, propargyl bromide was added slowly and the mixture allowed to reach room temperature. The progress of the reaction was monitored by TLC. While the starting material and product have very similar Rf values, application of a ninhydrin stain allowed for differentiation between the secondary amine starting materials and the tertiary amine products. After 24 hours, complete consumption of the starting materials was apparent. The products **14a** and **14b** were isolated in good yields of around 85 % after purification by silica gel column chromatography.

Characterization of the formed compounds was done by the methods of NMR and IR spectroscopy. Melting points were recorded where appropriate. The assignments of the NMR spectra were made with reference to the numbering scheme defined in figure 4.9.



14a **14b**

Figure 4.9: Numbering system for alkylated, Boc protected anilines **14a** and **14b**

Comparison of the ^{1}H and ^{13}C NMR spectra of the starting materials and the products isolated for this reaction confirm the successful addition of the propargyl group. In the ^{1}H NMR spectra of products **14a** and **14b**, two new peaks are observed. The peak at approximately 2.3 ppm, which resolves as a triplet with an integration of one proton, was assigned to the hydrogen on carbon atom C1. The second additional peak appears at close to 4.3 ppm and is a doublet which integrates for two protons. This peak was assigned to the hydrogens on carbon atom C3. These two peaks were seen to couple each other through the triple bond, with a resulting $^{4}J_{H\text{-}H}$ coupling constant of 2.4 Hz. The assignments for these peaks were made based on their multiplicity and integration. In addition, the peak for the amine proton present in the ^{1}H NMR spectra of the starting materials is absent in the ^{1}H NMR spectra of the products.

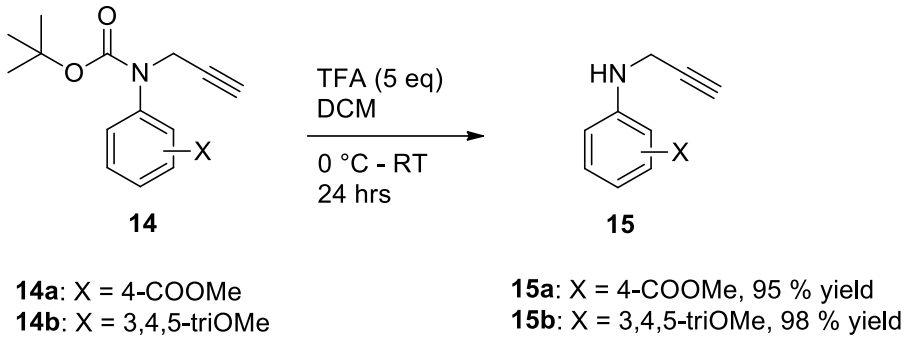
Three new peaks appear in the ^{13}C NMR spectra of each product. They are found at approximately 80 ppm, 72 ppm, and 40 ppm and were assigned to carbons C2, C1, and C3, respectively. The assignments were made based on the HSQC and HMBC spectra, where coupling to the already assigned ^{1}H NMR spectroscopic signals was evident. The remaining peaks in the NMR spectra of the formed products are conserved from the respective starting materials.

IR spectroscopy is also useful in confirming the success of the reaction, as two new characteristic stretching vibrations are expected for the added terminal alkyne group. For both formed products, the C≡C-H stretching mode is seen at around 3270 cm^{-1}, while the C≡C stretching mode is observed at close to 2120 cm^{-1}. The other expected C=O, C=C aromatic, and C-O stretching vibrations are also observed for each product.

Compound **14a** was isolated as on off-white solid, and the melting point recorded as 89 – 92 °C, which is in agreement with the reported literature value[82]. **14b** was isolated as a low-melting yellow solid, with a melting point of less than 40 °C.

4.2.1.5 Deprotection of alkylated anilines

The next step in the preparation of the alkylated anilines was the removal of the Boc protecting group before the subsequent Sonogashira coupling reaction. This was achieved using standard reaction conditions which are described in scheme 4.8.



Scheme 4.8: Deprotection of alkylated anilines

The respective Boc protected anilines **14a** and **14b** were dissolved in dichloromethane and the mixture cooled to 0 °C in an ice bath before trifluoracetic acid (TFA) was slowly added. The mixture was then allowed to reach room temperature, and the reaction progress was monitored by TLC. After 24 hours complete conversion of the starting materials was observed.

NMR spectroscopy confirmed the successful removal of the Boc group. Additional characterization was also done using IR spectroscopy and melting point analysis. The assignments of the peaks in the NMR spectra were made with reference to the numbering system defined in figure 4.10.

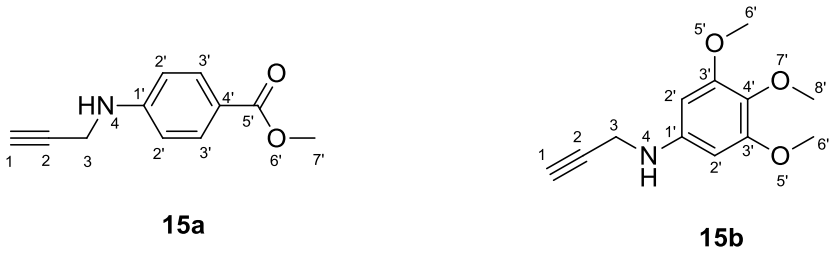


Figure 4.10: Numbering system for alkylated deprotected anilines **15a** and **15b**

The peaks in the NMR spectra of the starting materials corresponding to the Boc protecting group are absent in the NMR spectra of the products. An additional peak is seen in the ^{1}H NMR spectra of **15a** and **15b** when compared to the spectra of the relevant starting materials. This peak is a broad singlet with an integration of one and was assigned to the newly added

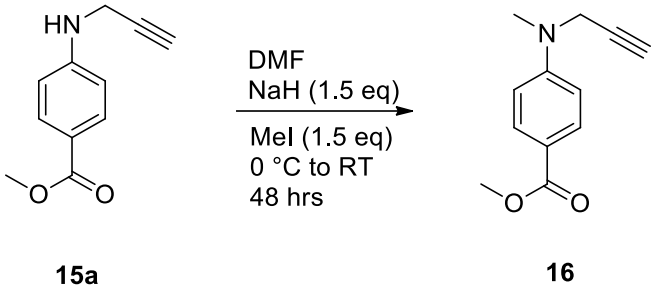
hydrogen on nitrogen atom N4. The other peaks seen in the ^{13}C and ^{1}H NMR spectra are consistent with what has already been described for the starting materials.

IR spectroscopy reveals the expected vibrational modes for the formed compounds, including C≡C-H, C≡C, C=C aromatic, C-O, and N-H stretches.

15a and **15b** were both isolated as solids, and their melting points were recorded. **15a** was found to melt in the range of 88 – 90 °C, and **15b** in the range of 69 – 72 °C. Both of these are in agreement with the melting point values reported in literature for these compounds[83].

4.2.1.6 Methylation of alkylated anilines

One last transformation was required for the deprotected intermediate **15a** before the subsequent Sonogashira coupling reaction could be conducted: namely methylation of the aniline nitrogen. Before reaching this point, some attempts were made to add a single methyl group to the unprotected aniline **11c**, with the hope of controlling the reaction specificity at this point to avoid the Boc protection step. However, regardless of the conditions assessed for this transformation, the same recurring problem of dialkylation was encountered. The addition of the methyl group was therefore conducted on the intermediate **15a**, after the series of steps including the initial Boc-protection, alkylation with propargyl bromide, and the removal of the Boc protecting group. Various different sets of conditions were assessed for this reaction, involving different solvents, bases and methylating agents. Methyl iodide (MeI) and dimethyl sulfate (DMS) were both assessed as methylating agents, with the former allowing for slightly improved yields. The best combination of solvent and base was found to be DMF and NaH, which facilitated the formation of the desired product in a yield of 75.5 %. These are similar conditions as used for the alkylation of the Boc protected anilines **13a** and **13b**, included in scheme 4.7. As mentioned, there are some potential hazards associated with mixtures of DMF and NaH, but as before, risks were minimized by performing the reaction on small scales with minimal equivalents of NaH and without heating above room temperature. The optimized reaction conditions employed for the methylation of intermediate **15a** are described in scheme 4.9.

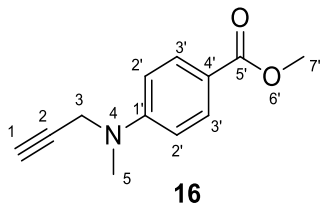


15a **16**

Scheme 4.9: Methylation reaction

15a was dissolved in DMF and the mixture cooled to 0 °C in an ice-bath. Sodium hydride was then added, and the mixture stirred at 0 °C for 20 minutes. Methyl iodide was introduced dropwise next, and the reaction mixture allowed to reach room temperature. Consumption of the starting materials was monitored by TLC, and after 48 hours complete conversion was observed. The methylated product **16** was isolated after purification by silica gel column chromatography as a pale-yellow solid in 75.5 % yield.

Successful formation of the desired product was confirmed with NMR spectroscopy and mass spectrometry. IR spectroscopy and melting point analysis were also conducted. The assignments for the NMR spectra of intermediate **16** were made based on the numbering system defined in figure 4.11.



16

Figure 4.11: Numbering system for methylated intermediate **16**

The added methyl group results in one additional signal seen in the ^{1}H NMR spectrum of the product **16** when compared to the spectrum of the starting material **15a**. This peak appears at 3.06 ppm and is a singlet with an integration of three protons. In addition, the peak corresponding to the secondary amine proton of **15a** is absent in the ^{1}H NMR spectrum of the

formed compound **16**. One additional peak is also observed in the ^{13}C NMR spectrum of the product isolated from this reaction. This peak appears at 38.48 ppm and was assigned to carbon C5 based on coupling to the corresponding ^{1}H NMR spectroscopic signal seen in the HSQC spectrum.

IR spectroscopy was conducted on the formed product, revealing all of the expected vibrations, including C≡C-H, C≡C, C-H, C=O, and C=C aromatic stretches.

Data obtained from positive ion mass spectrometry further confirmed the successful addition of the methyl group. An $[M+H]^+$ peak with an m/z value of 204.1005 was observed, which is in agreement with the expected mass of $C_{12}H_{14}NO_2^+$ of 204.1019 m/z.

Product **16** was isolated as a pale-yellow solid, and the melting point was recorded as 92 – 95 °C.

4.2.2 Sonogashira coupling reaction

After the successful formation of the various propargyl substituted phenols and anilines, included in figure 4.12, the key Sonogashira coupling reaction could be performed to facilitate the attachment of the previously prepared 2,4-diamino-5-iodo-6-cycloproyl-pyrimidine (**7b**), as referenced in scheme 4.1.

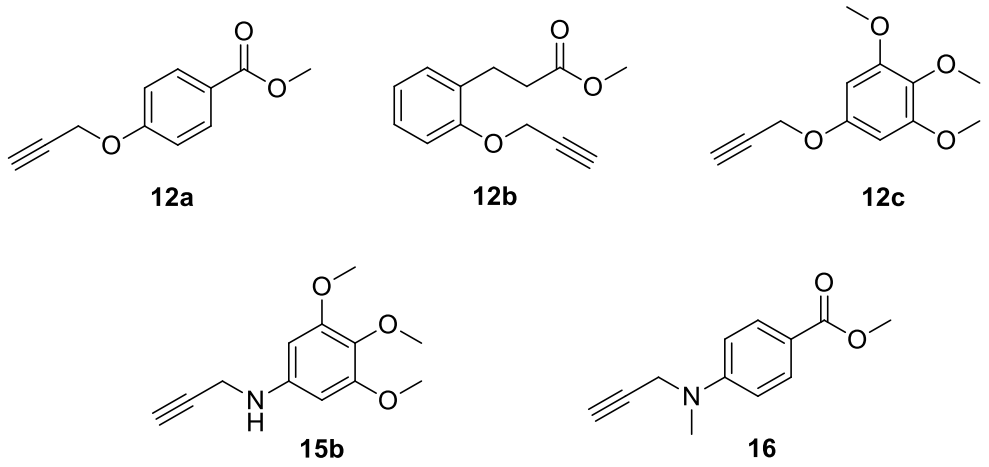
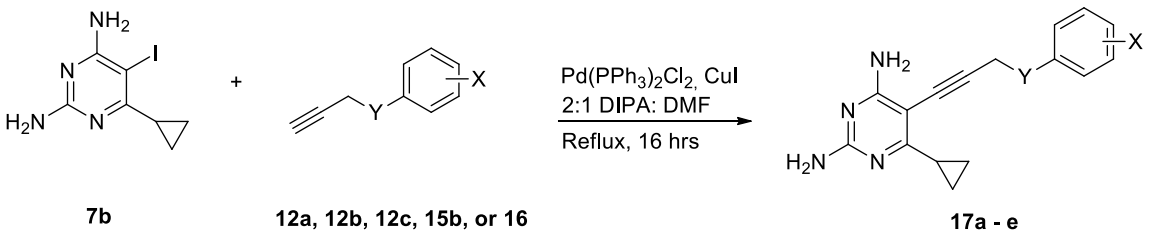


Figure 4.12: Propargyl substituted phenols and anilines for Sonogashira coupling reaction

The same robust conditions as described in the previous chapter for this transformation were used for the preparation of the second-generation analogues, and are included in scheme 4.10.

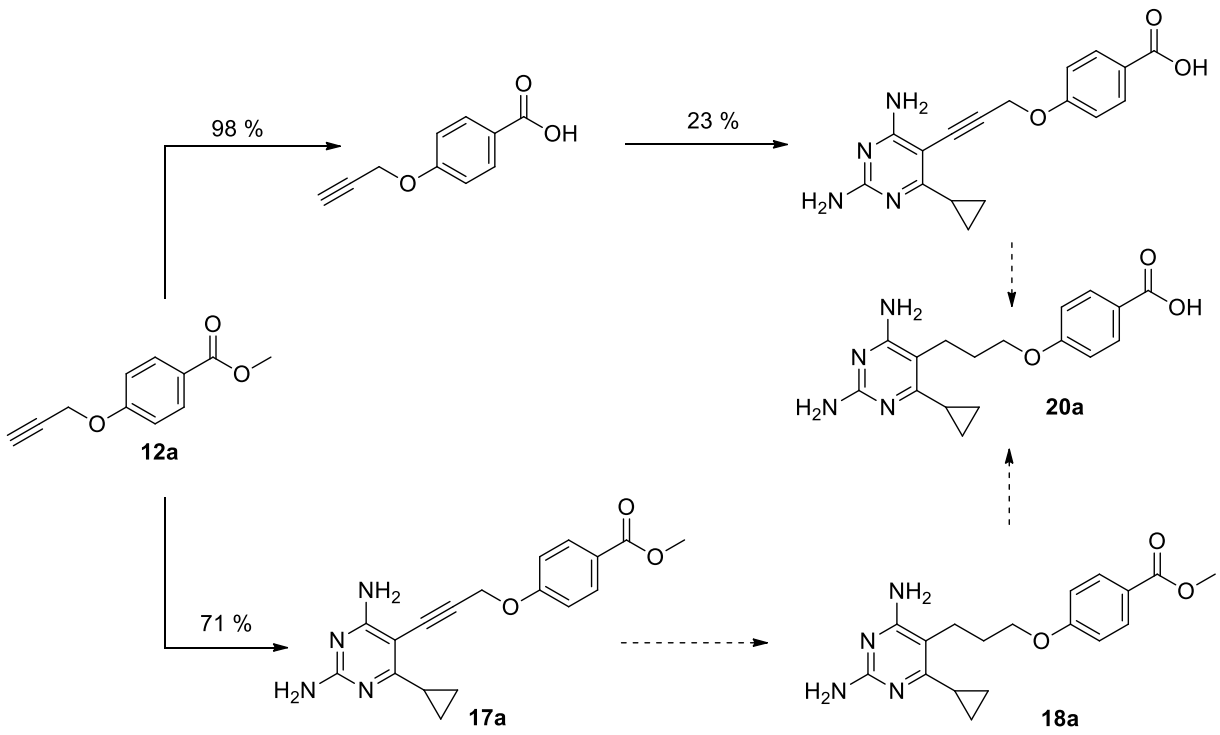


Scheme 4.10: Sonogashira coupling reaction

Dimethyl formamide (DMF) and diisopropyl amine (DIPA) in a 1:2 ratio was used as the solvent for the reaction. DIPA also performed a function as the base. This mixture was first degassed by bubbling argon or nitrogen gas through it, before the addition of the iodinated pyrimidine starting material (**7b**), $Pd(PPh_3)_2Cl_2$ catalyst, and copper iodide (CuI) cocatalyst. The mixture was then heated to 80 °C to ensure the added solid materials were fully dissolved. The relevant propargyl substituted phenols or anilines were then added as a solution in degassed DMF to the reaction mixture. After 16 hours of heating at reflux (80 °C), complete conversion of the starting materials could be seen with TLC. After workup, the crude product was adsorbed onto silica for subsequent purification by column chromatography, followed by an additional purification step either by recrystallization or a second column using basic alumina as the stationary phase for the removal of the $OPPh_3$ byproduct. Products **17a – 17e** were isolated as yellow or orange solids in yields ranging from 62 % to 77 % after purification.

Initially, for starting materials **12a**, **12b**, and **16**, it was thought to hydrolyze the methyl ester to the corresponding carboxylic acid before performing the Sonogashira coupling reaction. The required hydrolysis reactions were readily achieved by refluxing the starting materials in a mixture of methanol and water along with sodium hydroxide for several hours. However, lower than expected yields were obtained from the subsequent Sonogashira coupling reactions, and additional challenges were encountered in the removal of the $OPPh_3$ byproduct. When the same reaction conditions were used on the methyl ester substrates

(**12a**, **12b**, and **16**), good yields of the desired products were obtained (70 – 75 %) after purification by an initial silica gel column, followed by either recrystallization or an additional basic alumina column, for the removal of the $OPPh_3$ byproduct. The hydrolysis step was therefore moved to the end of the synthesis, after the Sonogashira coupling reaction and subsequent hydrogenation to obtain the final carboxylic acid containing analogues, as described with reference to the starting material **12a** in scheme 4.11.



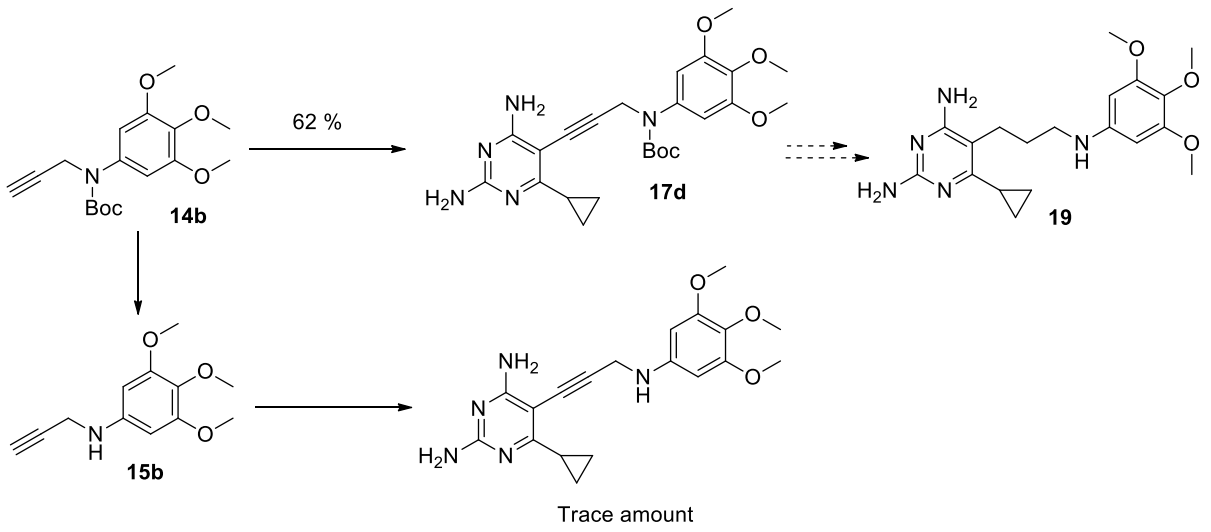
Scheme 4.11: Alternative routes for the formation of reduced carboxylic acid containing compound **20a**

Interestingly, when the reaction was conducted on starting material **15b**, only traces of the desired product were formed. The homocoupling byproduct of **15b** was identified from the crude reaction mixture, along with large amounts of de-iodinated 2,4-diaminopyrimidine starting material, which was the major product isolated. This suggests that the active catalyst species was formed successfully *in situ*, confirmed by the presence of the homocoupling byproduct and the fact that the reaction mixture turned black after the addition of the alkyne **15b**. However, either the transmetallation step which transfers the coordinated alkyne from

the cocatalyst to the active Pd catalyst, or the reductive elimination step which forms the desired carbon-carbon bond between the terminal alkyne and the 5-position of the 2,4-diaminopyrimidine ring did not occur satisfactorily, resulting in very low yields of the target compound. The overall catalytic cycle for the Sonogashira coupling reaction is included in scheme 3.13 in the previous chapter.

The reaction was attempted again under the same conditions except using the Boc-protected propargyl substituted aniline **14b**. The desired product was isolated after purification from this reaction in an acceptable yield of 61.7 %. Overall, this suggests that it is to do with the electronics of the alkyne group which prevented the reaction from working on the deprotected starting material **15b,** while allowing for the reaction to take place when the Boc-protected starting material **14b** was used. The Boc group is highly electron withdrawing, which lowers the electron density around the alkyne group. It has been reported that electron deficient alkynes tend to work better under Sonogashira coupling reaction conditions, as they have a greater ability to coordinate to the Pd catalyst center and undergo the required transformations[84].

After isolating the Boc-protected Sonogashira coupling product, the removal of the Boc-group was still required, and was performed after the hydrogenation reaction. This sequence of steps is described in scheme 4.12.



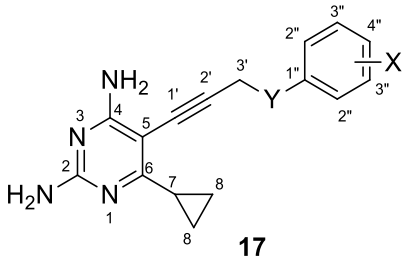
Scheme 4.12: Alternative routes for the formation of product **19**

The five different products obtained from the Sonogashira coupling reaction are included in table 4.1, which describes their structural elements and lists the respective yields and m/z values obtained from HRMS experiments.

Table 4.1: List of Sonogashira reaction products (**17a – e**)

Number:	**Linker (Y):**	**X-substituent:**	**Yield (%):**	**HRMS $[M+H]^+$ (m/z):**	**[M+H]+ expected (m/z):**
17a	O	4-COOMe	70.8	339.1439	339.1452
17b	O	2-CH_2CH_2COOMe	75.9	367.1755	367.1765
17c	O	3,4,5-triOMe	77.1	371.1693	371.1714
17d	NBoc	3,4,5-triOMe	61.7	470.2393	470.2398
17e	NMe	4-COOMe	70.6	352.1764	352.1768

The successful formation of the desired products was confirmed by IR and NMR spectroscopy, as well as mass spectrometry. All of the formed products **17a – 17e** were isolated as solids, and their melting points were recorded. The assignments made for the NMR spectra were based on the numbering system defined in figure 4.13. ^{1}H and ^{13}C NMR spectra of formed product **17e** are included in figures 4.14 and 4.15, respectively, with the assignments of the various spectroscopic signals indicated.



17

Figure 4.13: Numbering system for Sonogashira coupling products **17**

The NMR spectra of each formed compound contains peaks arising from both the 2,4-diamino-6-cyclopropyl pyrimidine ring as well as the added linker and terminal ring portions.

Four signals in the ^{1}H NMR spectra of each product were assigned to the 2,4, and 6-substituents around the pyrimidine ring. Two broad singlets with an integration of two each are observed between 5.0 and 6.4 ppm and were assigned to the protons on the amine groups attached to carbons C2 and C4. Three additional ^{1}H NMR spectroscopic signals arise from the cyclopropyl group and are observed at approximately 2.2 ppm, 1.0 ppm, and 0.9 ppm, respectively. The first peak listed resolves as a quintet with an integration of one and was assigned to the hydrogen on carbon atom C7. Due to the nature of the cyclopropyl group, two signals are expected for the axial and equatorial environments of the hydrogens on carbon atoms C8. The signal at close to 1.0 ppm is a multiplet with an integration of two and was assigned to the equatorial hydrogens of C8 based on the chemical shift. The remaining peak at around 0.9 ppm is also a multiplet with an integration of two protons and was assigned to the axial hydrogens of C8.

In the ^{13}C NMR spectra of the formed products **17a – 17e** four signals arise from the carbon atoms of the pyrimidine ring. These signals appear at around 172 ppm, 164 ppm, 162 ppm, and 87 ppm, and were assigned to carbons C6, C4, C2, and C5, respectively, based on the long-range coupling seen in the HMBC spectra of each product. The relative positions of these peaks is conserved from the starting material **7b**. Two peaks in the ^{13}C NMR spectra were assigned to the cyclopropyl group, the signal for carbon atom C7 appears at around 15 ppm, and the C8 signal appears at around 9 ppm. These two peaks were assigned based on the coupling to the already assigned ^{1}H NMR spectroscopic signals seen in the HSQC spectra.

The hydrogens attached to carbon atom C3’ of the linker appear as a singlet in the ^{1}H NMR spectrum of each formed compound. This peak is found at around 4.8 ppm and integrates for two protons. Three signals are seen in the ^{13}C NMR spectrum of each product corresponding to the carbon atoms C1’, C2’ and C3’ of the linker. The signals for the C1’ and C2’ carbons are conserved in position, appearing at close to 94 ppm and 80 ppm, respectively. The assignments were made based on coupling seen in the HMBC spectra. The signal for carbon atom C3’ varies in position depending on the nature of the heteroatom present in the 4-position of the linker. Formed products **17a**, **17b**, and **17c** have an oxygen atom in this position, so the peak for the C3’ carbon appears at around 58 ppm. **17d** has a Boc-protected nitrogen in the 4-position of the linker, which causes the chemical shift of the C3’ carbon to shift upfield to a new position of 40.54 ppm due to the electron withdrawing nature of the Boc-group.

Product **17e** has an N-Me group in the linker 4-position, resulting in the signal for the C3' carbon to be observed at 51.41 ppm. The assignment of the C3' carbon signals in the ^{13}C NMR spectrum of each product was made based on the coupling seen to the already assigned C3' hydrogens in the HSQC spectra.

Reaction products **17a**, **17b**, and **17e** display peaks with a total integration of four protons in the aromatic region of the ^{1}H NMR spectra. The assignments for these peaks were made based on their integration and splitting patterns and are consistent with the assignments made for the corresponding starting materials. Products **17c** and **17e** have only one signal in the aromatic region of the ^{1}H NMR spectra, which is a singlet with an integration of two protons and was assigned to the hydrogens on carbon atoms C2". The assignments of the ^{13}C NMR spectroscopic signals of the aromatic carbons for each product were made based on the coupling seen in the HSQC and HMBC spectra and are likewise consistent with the assignments made for the respective starting materials.

Reaction products **17a** and **17e** contain a methyl ester group attached to the 4-position of the terminal ring. The signal for the methyl hydrogens appears as a singlet with an integration of three protons at close to 3.8 ppm in the ^{1}H NMR spectra of these two products. **17e** has an additional methyl group attached to the nitrogen atom present in the four-atom linker. The corresponding ^{1}H NMR spectroscopic signal is seen at 3.05 ppm as a singlet integrating to three protons. Distinction between the two methyl signals for product **17e** was made based on the coupling observed in the HSQC spectra, where the ester methyl hydrogens were seen to couple to the corresponding adjacent carbon atom. The methyl ester group produces two signals in the ^{13}C NMR spectra of products **17a** and **17e**. The carbonyl carbon appears at close to 166 ppm for both products, while the methyl carbon is seen at 51.89 ppm for **17a** and 42.61 ppm for **17e**. The assignments for these peaks was made based on the coupling seen in the HSQC and HMBC spectra. The ^{13}C NMR spectroscopic signal for the methyl group attached to the nitrogen atom present in the linker of compound **17e** appears at 38.15 ppm and was assigned with assistance from the HSQC spectrum.

Products **17c** and **17d** have a 3,4,5-trimethoxy substituent around the terminal ring. This group is seen to produce two peaks in the ^{1}H NMR spectra of these products. Both peaks are singlets, appearing at close to 3.8 ppm. The assignments for these peaks were made based on their integration, with the methoxy group in the 4-position of the ring giving rise to a singlet

with an integration of three protons, and the methoxy groups on the equivalent carbon atoms in the 3-position of the ring forming a singlet with an integration of six protons. The corresponding signals in the ^{13}C NMR spectra of these products are seen at around 55 ppm and 60 ppm, respectively, and were assigned based on the coupling seen in the HSQC spectra.

17b has a methyl propionate group attached to the 2-position of the terminal ring. Three signals are seen in the ^{1}H NMR spectrum for this group. The methyl ester hydrogens give rise to a singlet with an integration of three protons at 3.65 ppm. The two other signals are observed at 2.97 ppm and 2.63 ppm, and both integrate for two protons. The assignments for these two peaks were made based on the long-range coupling seen in the HMBC spectra and are consistent with the assignments made for the starting material **12b.** In the ^{13}C NMR spectrum of product **17b**, four peaks are seen for the methyl propionate group. The carbonyl carbon was assigned to the peak at 173.87 ppm, while the methyl carbon was assigned to the peak at 51.67 ppm. The two remaining peaks appear at 34.27 ppm and 26.25 ppm and were assigned based on the coupling seen in the HSQC and HMBC spectra.

For the formed products **17a – 17e**, the only NMR spectroscopic signals not discussed above arise from the Boc-group present in compound **17d**. In the ^{1}H NMR spectrum of this compound, a singlet integrating to nine protons at 27.97 ppm was assigned to the hydrogens of the *tert*-butyl moiety of the Boc group. Three signals are observed in the ^{13}C NMR spectrum of **17d** which arise from the Boc group. The carbonyl carbon peak is seen at 152.44 ppm, while the quaternary *tert*-butyl carbon appears at 78.18 ppm. The final signal for the methyl carbons of the *tert*-butyl group is seen at 29.97 ppm. The positions of these peaks are conserved from the starting material **14b** and were assigned based on the HSQC and HMBC spectra.

IR spectroscopy of the Sonogashira reaction products reveals the key C≡C stretching vibration of the internal alkyne at around 2200 cm^{-1}. This is seen in addition to the other expected N-H, C=N, C=C, and C-O stretching modes.

High resolution mass spectrometry was also conducted on all of the formed compounds, revealing prominent $[M+H]^{+}$ peaks at the expected m/z values in each case. The m/z values obtained for each of the formed products are listed in table 4.1.

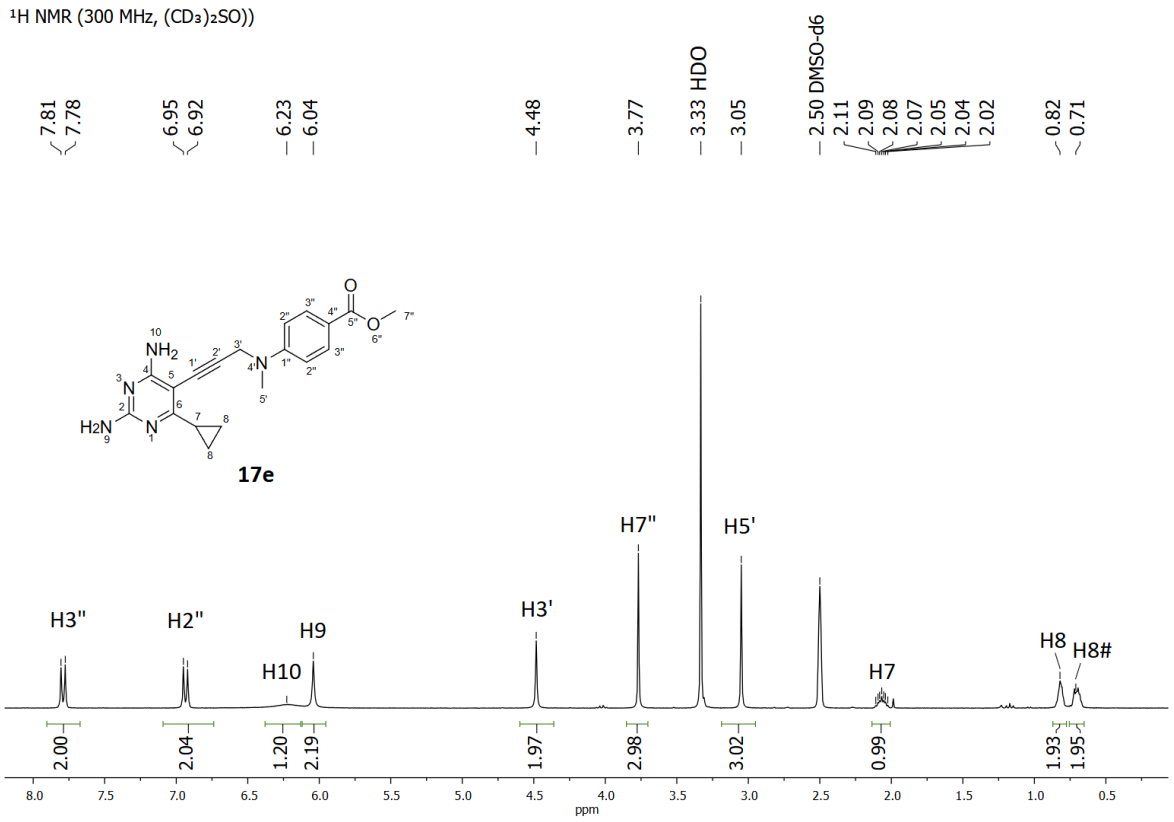


Figure 4.14: ^{1}H NMR spectrum of **17e** with spectroscopic signal assignments included

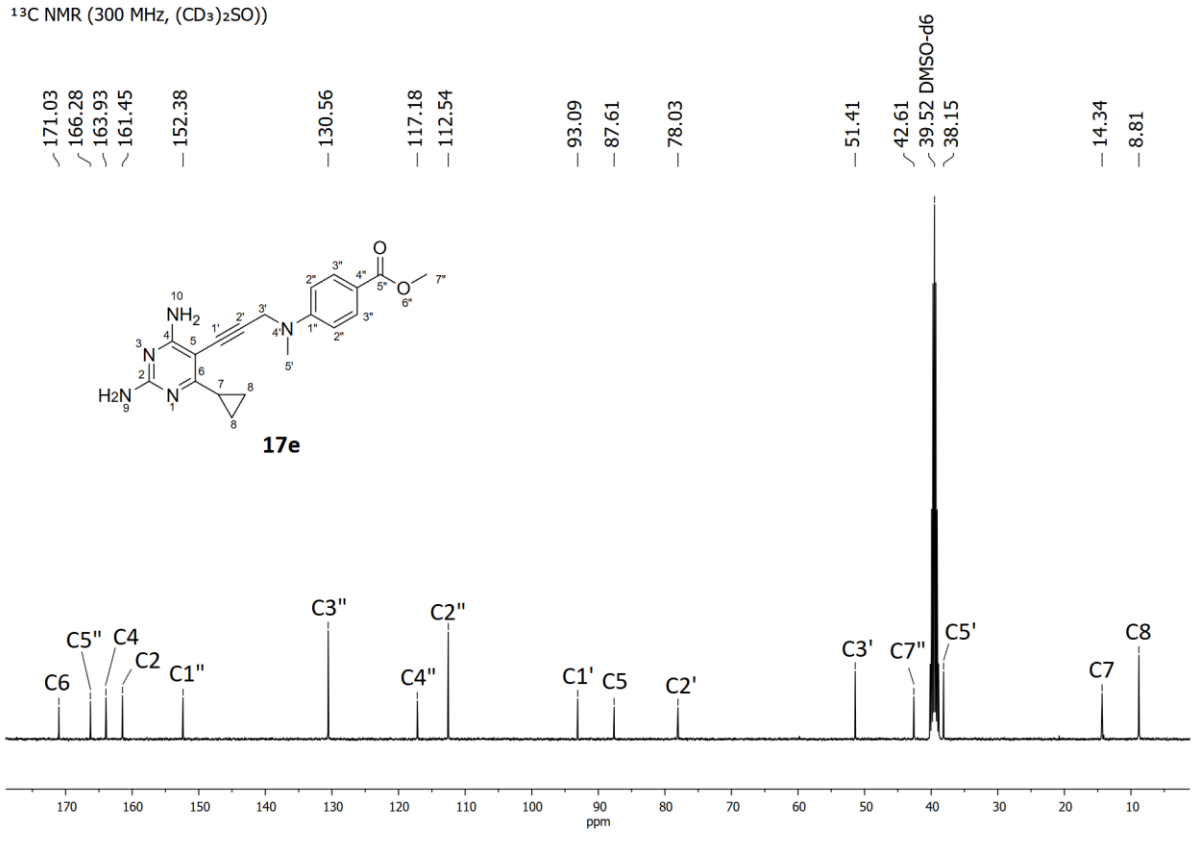
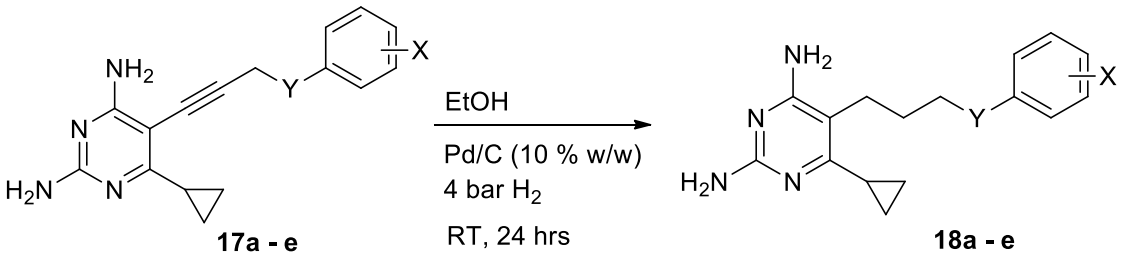


Figure 4.15: ^{13}C NMR spectrum of **17e** with spectroscopic signal assignments included

Melting points were recorded for each product, showing high variations between the different analogues due to the varied nature of the linker and X-substituents around the terminal aromatic ring.

4.2.3 Hydrogenation of Sonogashira reaction products

Hydrogenation reactions were successfully performed on all of the isolated Sonogashira coupling reaction products **17a – e**, facilitating the formation of the fully flexible compounds **18a – e**. The standard hydrogenation conditions used for this transformation were carried over from the synthesis of the first-generation analogues discussed in the previous chapter and are described in scheme 4.13.



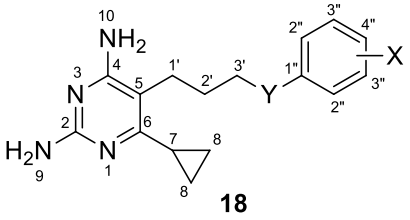
Scheme 4.13: Hydrogenation of Sonogashira reaction products

The respective alkyne starting materials (**17a – e**) isolated from the Sonogashira coupling reaction were dissolved in ethanol and transferred to a high-pressure reaction vessel before the addition of palladium on carbon. The mixture was stirred at room temperature under an atmosphere of hydrogen gas at a pressure of 4 bar. After this time the mixture was filtered through Celite, and purification was done if necessary by silica gel column chromatography. All of the desired products were isolated in good yields (85 – 95 %) as white crystalline solids. The five compounds formed in this way are listed in table 4.2, which categorises them by their structural elements and shows the reaction yields along with the m/z values obtained from HRMS.

Table 4.2: List of hydrogenation products **18a – e**

Number:	**Linker (Y):**	**X-substituent:**	**Yield (%):**	**HRMS $[M+H]^+$ (m/z):**	**[M+H]+ expected (m/z):**
18a	O	4-COOMe	94.5	343.1749	343.1765
18b	O	2-CH_2CH_2COOMe	86.2	371.2061	371.2078
18c	O	3,4,5-triOMe	85.1	375.2037	375.2027
18d	NBoc	3,4,5-triOMe	91.3	474.2708	474.2711
18e	NMe	4-COOMe	94.4	356.2075	356.2081

The successful reduction of the triple bond for each product was confirmed by the methods of NMR and IR spectroscopy, as well as mass spectrometry. Melting points were also determined. The assignments of the peaks in the NMR spectra were made with reference to the numbering system defined in figure 4.16. ^{1}H and ^{13}C NMR spectra of formed product **18d** are included in figures 4.17 and 4.18, respectively, with the assignments of the various spectroscopic signals indicated.



18

Figure 4.16: General numbering system for hydrogenated compounds **18**

In the ^{1}H NMR spectra of each formed compound, two additional peaks are apparent when compared to the spectra of the starting materials. These peaks arise from the hydrogens added across the triple bond in the hydrogenation reaction. The hydrogens of carbon atom C1' produce a triplet integrating to three protons at around 2.7 ppm, while the hydrogens on C2' resolve as a quintet with an integration of two protons at close to 2.0 ppm. The assignments for these two signals were made based on their multiplicity and the $^3J_{H\text{-}H}$ coupling seen in the COSY spectra. Upon the successful reduction of the triple bond, the signal

corresponding to the hydrogens on carbon atom C3' becomes split into a triplet and is shifted slightly upfield from its position in the ^{1}H NMR spectra of the starting materials.

The same number of signals are seen in the ^{13}C NMR spectra of the starting materials and the products of the hydrogenation reaction. However, the peaks corresponding to the carbon atoms on either side of the newly reduced triple bond show notable changes in their positions. The C5 carbon of the pyrimidine ring moves downfield in the ^{13}C NMR spectra of the hydrogenation products, to a new conserved position of around 105 ppm. The peaks for carbon atoms C1' and C2' move slightly upfield, with the C1' and C2' signals appearing at approximately 22 ppm and 30 ppm, respectively, in the ^{13}C NMR spectra of the formed products.

IR spectroscopy was conducted on each of the isolated products, and the successful reduction of the triple bond is confirmed by the absence of the key C≡C stretching vibration which is observed at around 2200 cm^{-1} in the spectra of the starting materials.

Mass spectrometry reveals prominent $[M+H]^{+}$ peaks for each compound. The observed m/z values are included in table 4.2 and differ from the starting materials by an amount equal to the mass of four protons, confirming the complete reduction of the triple bond in each case.

Melting points were recorded for each analogue and differ slightly from those reported for the respective starting materials.

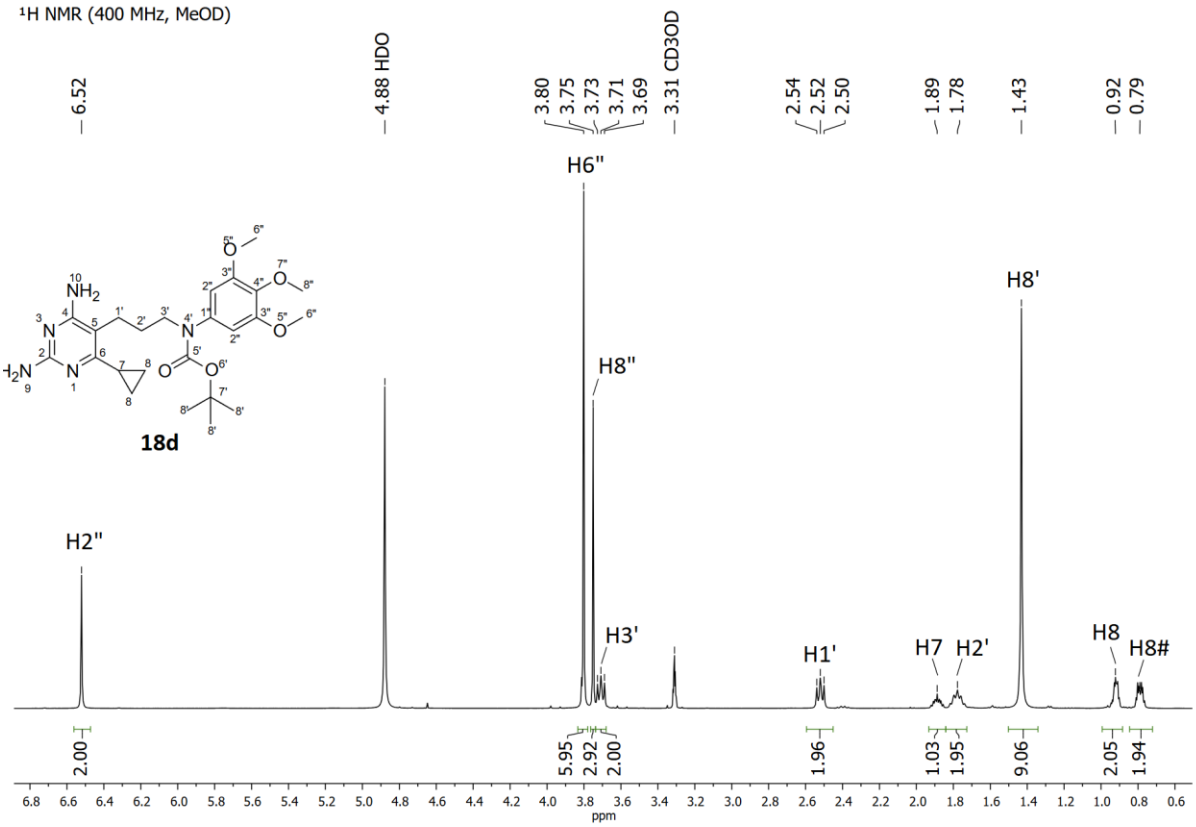


Figure 4.17: ^{1}H NMR spectrum of **18d** with spectroscopic signal assignments included

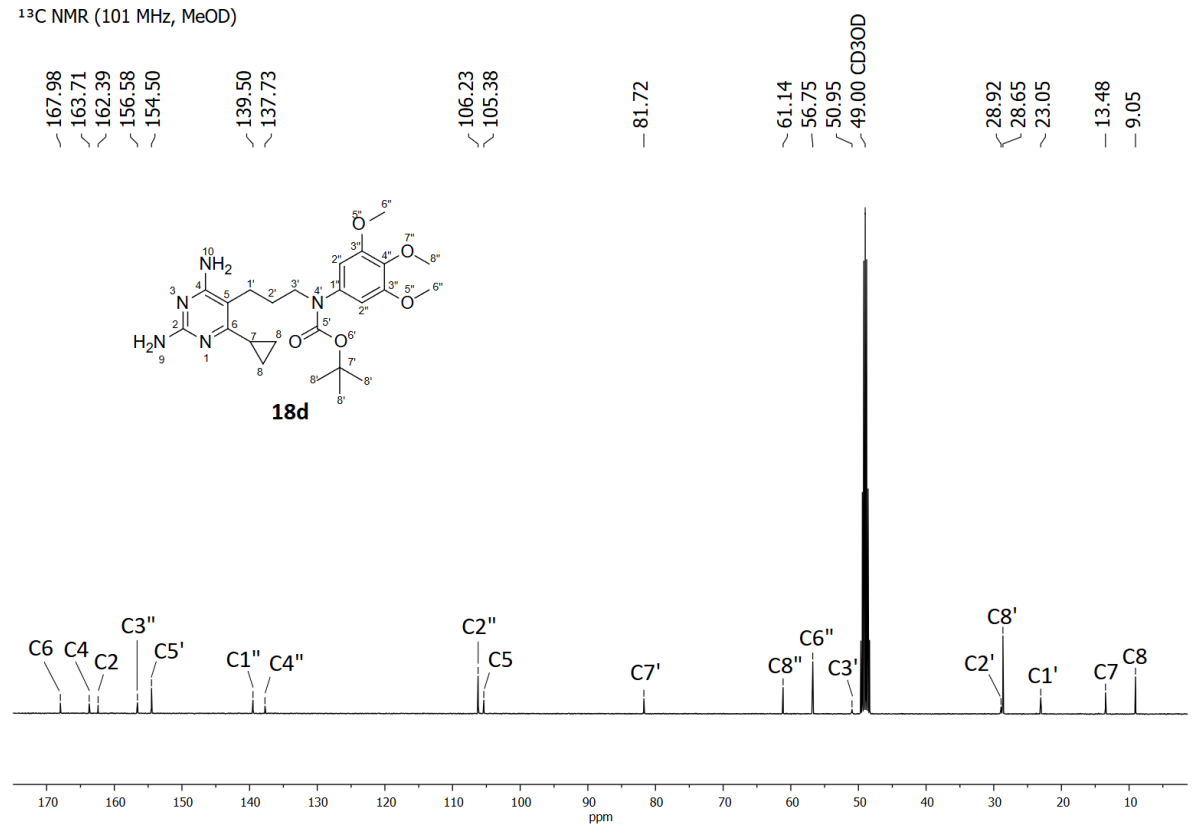
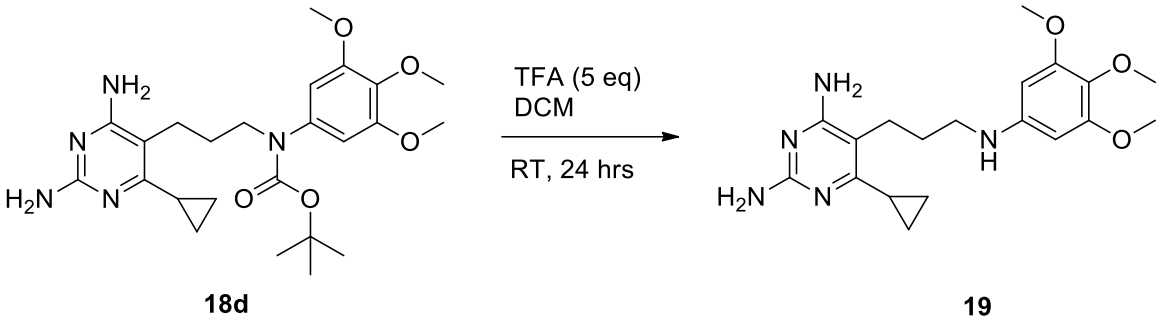


Figure 4.18: ^{13}C NMR spectrum of **18d** with spectroscopic signal assignments included

4.2.4 Removal of Boc protecting group

Intermediate **18d** isolated from the prior hydrogenation reaction still had the Boc-group present. To obtain the desired final product, an additional deprotection step was required. It was hoped that under the hydrogenation conditions employed that the Boc-group would be removed along with the reduction of the triple bond. This, however, did not occur, and so standard Boc deprotection conditions were used for the formation of the final product. The reaction conditions are described in scheme 4.14.



Scheme 4.14: Removal of Boc-group from **18d**

The Boc-protected starting material **18d** was dissolved in DCM and the mixture cooled to 0 °C in an ice bath. Trifluoroacetic acid (TFA) was then added slowly, and the mixture was left to stir at room temperature for several hours. The progress of the reaction was monitored by TLC, and after 24 hours no more starting material could be observed. Purification was achieved by silica gel column chromatography, to afford the desired product as an orange solid in a yield of 80.0 %.

Successful removal of the Boc-group was confirmed with NMR spectroscopy and mass spectrometry. IR and melting point analysis were also conducted on the formed compound. The assignments for the NMR spectra were made in reference to the numbering system defined in figure 4.19.

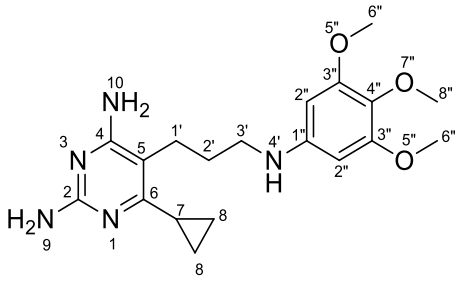


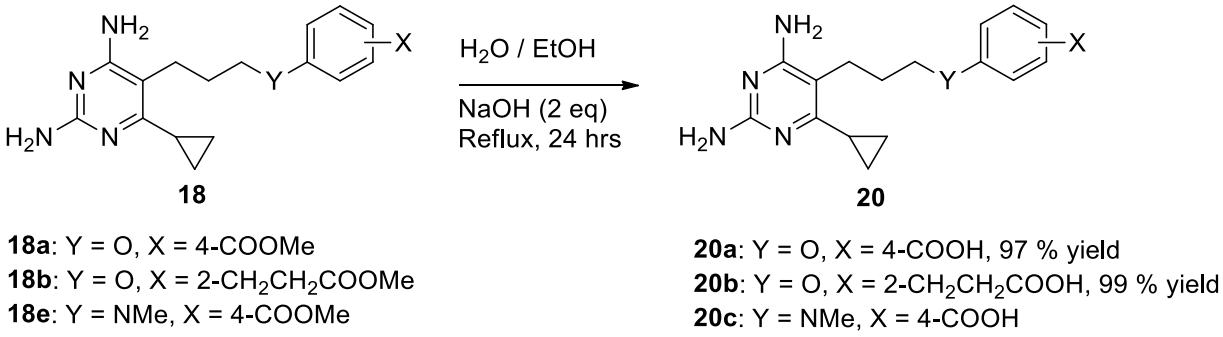
Figure 4.19: Numbering system for **19**

The peaks in the ^{1}H and ^{13}C NMR spectra of the starting material **18d** which correspond to the Boc-group are absent in the spectra of the formed product **19**. In addition, an extra signal is observed in the ^{1}H NMR spectrum of **19**, a broad singlet at 3.38 ppm which integrates for one proton. This peak was assigned to the newly added hydrogen atom on the secondary amine in the linker.

IR spectroscopy reveals the expected peaks, including N-H, C=N, C=C, and C-O stretching modes. HRMS of the formed compound shows a prominent $[M+H]^{+}$ peak at the expected mass in positive ion mode. The observed m/z value for **19** differs from the value obtained from the starting material **18d** by the mass of the Boc-group plus one hydrogen atom, confirming the successful removal of the Boc-group. The final product **19** was isolated as an orange solid, and the melting point recorded as 127 – 129 °C, which is slightly higher than the melting point of the Boc-protected intermediate **18d**.

4.2.5 Hydrolysis of esters after Sonogashira and hydrogenation

For intermediates **18a**, **18b**, and **18e**, the Sonogashira coupling reaction was conducted on the methyl ester starting materials. To proceed with the synthesis of the desired final products, a hydrolysis reaction was required to convert the methyl ester into a carboxylic acid. This was achieved by refluxing the starting materials in a mixture of ethanol and water, with added sodium hydroxide, as described in scheme 4.15.



Scheme 4.15: Hydrolysis of hydrogenation products **18**

The starting materials were dissolved in a 1:1 mixture of ethanol and water and sodium hydroxide was added. The reaction mixture was heated to reflux, and the consumption of the starting material was monitored by TLC. After 24 hours, complete conversion was apparent. Products **20a** and **20b** were isolated as white solids in near quantitative yields after recrystallization from the reaction mixture. The reaction was not performed on the starting material **18e**, due to time constraints encountered in the completion of the project.

The successful hydrolysis of intermediates **18a** and **18b**, forming compounds **20a** and **20b**, respectively, was confirmed by the methods of IR and NMR spectroscopy, as well as mass spectrometry. **20a** and **20b** were both isolated as white solids, and their melting points recorded. The assignments made for the NMR spectra were based on the numbering system defined in figure 4.20.

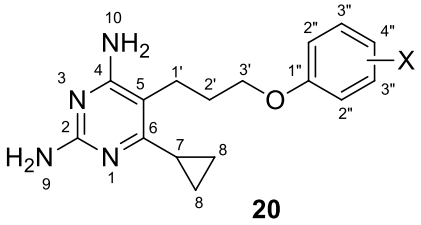


Figure 4.20: Numbering system for hydrolysis products **20a** and **20b**

The peaks corresponding to the methyl group in the ^{1}H and ^{13}C NMR spectra of the starting materials **18a** and **18b** are absent in the spectra of the formed products **20a** and **20b**. In

addition, broad O-H stretching vibrations are observed in the IR spectra of the products at around 3300 cm^{-1}, confirming the success of the hydrolysis reactions. High resolution mass spectrometry was conducted on the isolated products, revealing prominent $[M+H]^+$ peaks at the expected m/z values. The $[M+H]^+$ peak for **20a** was observed at 329.1586, while the $[M+H]^+$ peak for **20b** was observed at 357.1915, which are in close agreement with the expected values for these products of 329.1608 and 357.1921 m/z, respectively.

Melting points were also recorded for the products and differ from those recorded for the starting materials.

4.2.6 Addition of glutamic acid

One final transformation was required to form the desired glutamic acid substituted final product of the synthesis, namely the addition of the glutamic acid in an amide bond formation step. This could be done directly on the methyl ester compound **18a**, or on the carboxylic acid containing intermediate **20a**. Amide bond formation is a key process in medicinal chemistry as many existing drug molecules contain amide linkages. In general, amide bonds are usually formed between carboxylic acid starting materials and amines, in the presence of a suitable coupling reagent[85]. The coupling reagents are required as direct reactions between carboxylic acids and amines only readily occur at extremely high temperatures. They serve to activate the carboxylic acid starting materials by adding to it first, forming a new highly reactive intermediate to which the amine can subsequently add. Often a secondary coupling reagent is also included, which reacts with the initially formed activated intermediate, forming a new intermediate which is the reactive species for the addition of the amine. There are several different classes of amide coupling reagents, which have advantages for different types of substrates. Carbodiimides are among the most widely used coupling reagents and have shown good results for the formation of amides between various carboxylic acids and amines. Carbodiimides have the general structural formula of RN=C=NR, where the R groups differ between different carbodiimide coupling reagents. Two examples are included in figure 4.21, *N*,*N*'-dicyclohexylcarbodiimide (DCC) and 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC).



Figure 4.21: Structures of DCC and EDC carbodiimide coupling reagents

Other types of coupling reagents include various phosphonium and uronium salts, which form activated P-O and P-N bonds, respectively, for the subsequent addition of a secondary-coupling reagent or direct addition of the added amine. Reagents which facilitate the *in situ* formation of acid chloride intermediates have also been widely used for amide bond formation. These conditions work well on some substrates but are often too harsh for more sensitive starting materials.

There are numerous other amide coupling reagents which have been described to facilitate successful formation of amide bonds between carboxylic acids and amines, all work in a similar fashion as described above: by first reacting with the carboxylic acid starting material to form a new reactive intermediate which has a lower energy barrier for the subsequent addition of the amine[85].

One possible downside to the use of coupling reagents in amide bond formation is poor atom economy. Stoichiometric amounts of the coupling reagent, and sometimes other reagents, are required for the formation of one equivalent of the amide product. Apart from green chemistry concerns, this also makes the purification of the final products more challenging, as there are additional byproducts formed in the reaction mixture to remove.

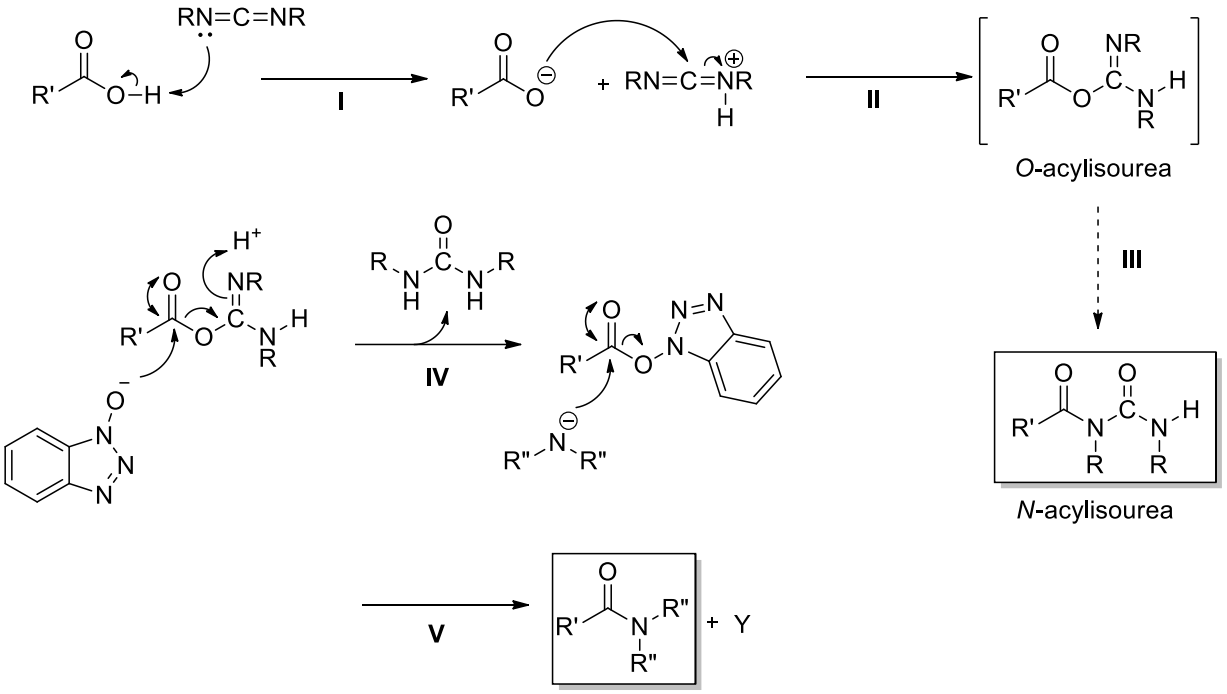
Less commonly, amide bonds are also sometimes formed from ester containing starting materials and amines. Due to the lower reactivity of esters compared to carboxylic acids, harsher reaction conditions are often required for the direct amidation of esters, often including the addition of Lewis acids to enhance the electrophilicity of the ester carbonyl carbon.

For the present synthesis, direct amidation of the methyl ester starting material **18a** was seen as the most attractive route as it avoided the additional hydrolysis step of this intermediate in the overall synthesis. Some patents were found which claim to form amide bonds between

the monosodium salt of glutamic acid (MSG) and folate-like starting materials containing a 4-methyl benzoate group, similar in structure to **18a**[86, 87]. The conditions described involve refluxing the starting materials in a mixture of *n*-butanol and water for several hours, to afford the desired amide product in good yields without the use of any additional reagents.

When this was attempted on **18a**, no reaction was observed and only starting material was recovered from the mixture. Some modifications were made to these conditions, including the use of different forms of the glutamic acid starting material, either MSG or the HCl salt, adding various amounts of HCl or K_2CO_3 (2 – 5 eq) to control the pH of the reaction mixture, and also varying the solvent system used. The patent methods[86, 87] make use of a mixture of 70 % *n*-butanol / water for this transformation, and different variations of this mixture were assessed, as well as the use of other single solvents systems such as acetonitrile and DMF. The reaction was also attempted with 1.5 equivalents of NaH, either in THF (reflux, 24 hours) or DMF (RT, 24 hours), but in each case either only unreacted starting material, or partially hydrolysed starting material, was recovered. It was thought that the addition of a Lewis acid could allow for the desired reaction to take place. To this end reactions were performed on the starting material **18a** and either MSG or the hydrochloride salt of glutamic acid (Glu HCl) in the presence of different Lewis acids and various solvents and bases. 0.1 – 3 Equivalents of $FeCl_3$, $AlCl_3$, or boron trifluoride diethyl etherate were assessed alongside ACN, THF, or DMF and 0.5 – 2 equivalents of K_2CO_3 or Cs_2CO_3 for this transformation but in each case no traces of the desired amide product could be identified. It was soon realized that the amide coupling reaction may afford better results if conducted on the carboxylic acid starting material **20a** with the use amide coupling reagents, and not directly on the methyl ester precursor **18a**.

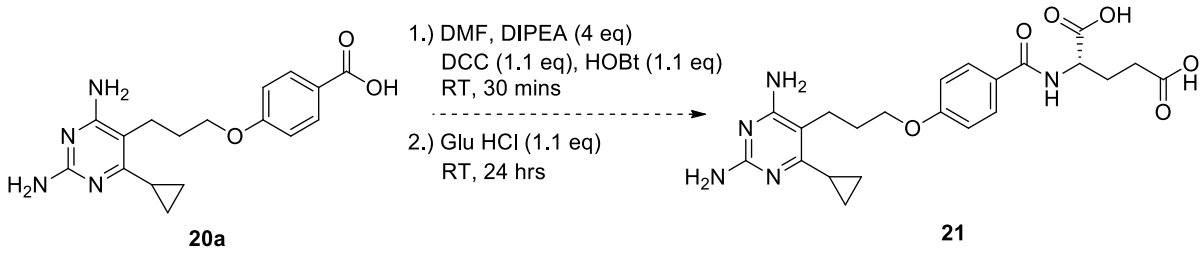
The coupling reaction was then attempted on the carboxylic acid starting material **20a** using *N*,*N*'-Dicyclohexylcarbodiimide (DCC) in the presence of hydroxybenzotriazole (HOBt) and *N*,*N*-Diisopropylethylamine (DIPEA) in DMF, before the subsequent addition of Glu HCl. HOBt acts as the secondary coupling reagent in this transformation, the overall mechanism of which is described in scheme 4.16. For DCC, R = cyclohexyl.



Scheme 4.16: Mechanism for amide bond formation with DCC and HOBt

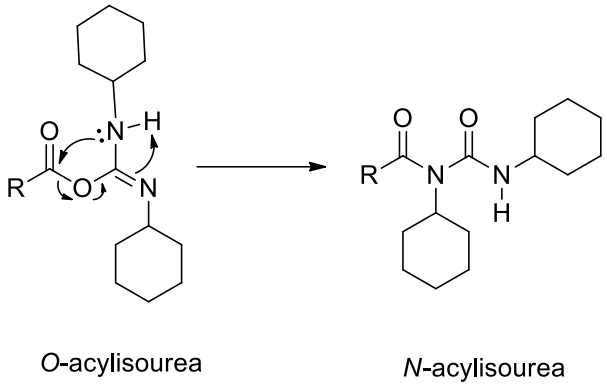
The reaction is initiated by the transfer of the acidic carboxylic acid proton to one of the basic carbodiimide nitrogens [**I**], followed by the nucleophilic attack of the newly formed carboxylate anion to the electrophilic central carbon of the coupling reagent [**II**], forming a new reactive *O*-acylisourea intermediate. For coupling reactions performed with carbodiimides, a secondary-coupling reagent is often required. (Scheme 4.16 includes HOBt as the secondary coupling reagent). The activated *O*-acylisourea intermediates formed *in situ* are highly unstable and show a tendency to isomerize into inactive *N*-acylisourea forms [**III**]. The secondary coupling reagent prevents this by adding first to the *O*-acylisourea before the unwanted isomerization can take place [**IV**], forming a new reactive intermediate to which the amine can add for the formation of the desired final product [**V**]. For the coupling of chiral amines, such as amino acids, the secondary coupling reagents can also limit the extent of epimerization of the chiral centre during the addition step. A suitable base is required for these types of transformations, typically mild nitrogen containing bases are used as they do not interfere with the formation of the activated intermediate but allow for the deprotonation of the added amine for subsequent nucleophilic attack.

The starting material **20a** was found to be soluble only in highly polar solvents such as DMF, DMSO, and alcohols. Fortunately, these types of transformations are often performed in DMF or DMSO. For the coupling reaction, it was initially thought to use the monosodium salt of glutamic acid (MSG), however, solubility issues were encountered with this compound in DMF and DMSO, so the hydrochloride salt of glutamic acid (Glu HCl), which showed sufficient solubility, was used instead. This required the addition of extra equivalents of base, to ensure the successful deprotonation of the amine group. The conditions used for the reaction of starting material **20a** are described in scheme 4.17.



Scheme 4.17: Reaction conditions for attempted amide coupling with DCC and HOBt

The carboxylic acid starting material **20a** was dissolved in (DMF) and DCC and HOBt were added followed by diisopropylethylamine (DIPEA) and the mixture stirred at room temperature. The consumption of the starting material was monitored by TLC, and after 30 minutes no starting material could be observed. Glu HCl was then added, and the mixture stirred at room temperature for a further 24 hours. After workup and purification by silica gel column chromatography, a single major product was isolated. High resolution mass spectrometry revealed the product to be the inactive *N*-acylisourea, which showed the correct m/z value of this product at 535.3388 (535.3391 expected) for the $[M+H]^+$ peak. As discussed above, this product forms from the isomerization of the active *O*-acylisourea intermediate *in situ* after the addition of DCC to the carboxylic acid. The mechanism for the isomerization is included in scheme 4.18.



O-acylisourea *N*-acylisourea

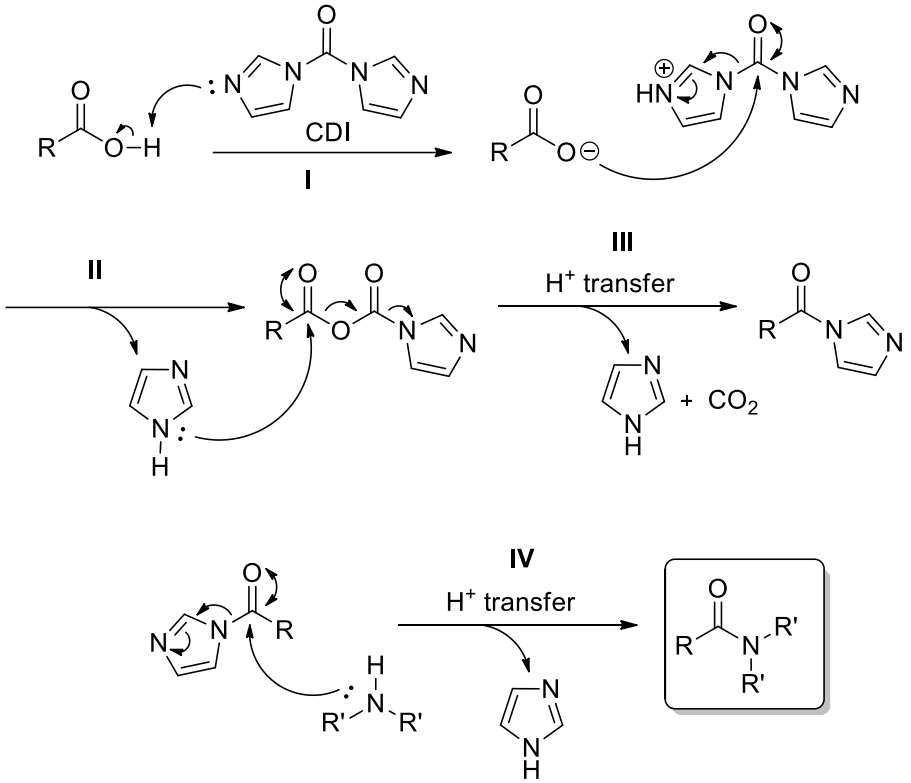
Scheme 4.18: Isomerization of *O*-acylisourea intermediate into inactive *N*-acylisourea form

The formation of the *N*-acylisourea product confirms that the first part of the reaction mechanism proceeded correctly, however, after the successful addition of DCC to the carboxylic acid, instead of subsequent addition by the secondary coupling reagent, HOBt, the *O*-acylisourea intermediate isomerized into the inactive form. There are many possible reasons as to why this unwanted isomerization occurred, including poor solubility, degradation of the HOBt, or insufficient reaction temperature, to list a few.

Modifications were made to the experimental methods described above to hopefully improve the outcome of the reaction. For all subsequent attempts, it was ensured that the HOBt was added to the reaction mixture before the addition of DCC, to ensure that as soon as the active *O*-acylisourea formed *in situ*, the secondary coupling reagent was available and could add immediately. A different source of HOBt was also used, as this reagent is prone to decomposition over time. Different combinations of bases and solvents were assessed, including performing the reaction in DMSO instead of DMF, and the use of TEA and DIPA as alternative bases to DIPEA. The number of equivalents of added base was also varied from two to six, however, in each case the major product isolated was the inactive *N*-acylisourea. The reaction was further attempted at 60 °C for 24 hours instead of room temperature, but in this case seemingly decomposed starting material was recovered. As a final attempt at the DCC coupling, DMAP was used as the secondary coupling reagent in the place of HOBt. Although for coupling reactions involving chiral amines, HOBt is preferred over DMAP as it leads to lower proportions of epimerization, it was thought to try DMAP to see if the reaction

would proceed past the formation of the active *O*-acylisourea intermediate, however, no improvements were realized under these conditions.

After many failed attempts at performing the coupling reaction with DCC, an alternative coupling reagent, carbonyldiimidazole (CDI), was employed next. Like all amide coupling reagents, CDI operates by forming a reactive intermediate with the carboxylic acid starting material to which the amide can subsequently be added to, forming the desired amide bond. A general mechanism for amide coupling with CDI is provided in scheme 4.19.

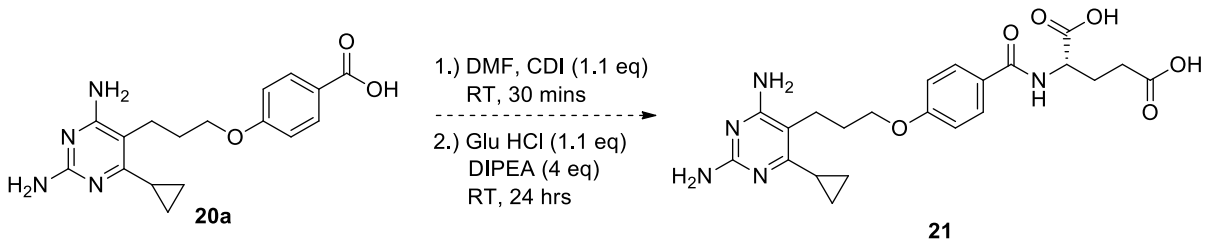


Scheme 4.19: General mechanism of amide bond formation with CDI

The reactive intermediate is formed over several steps. First the carboxylic acid starting material is deprotonated by one of the basic nitrogens of CDI [**I**], followed by the addition of the newly formed carboxylate anion to the carbonyl carbon of CDI, with the elimination of imidazole to form a new anhydride intermediate [**II**]. The free imidazole then adds back to the carbonyl carbon of the original carboxylic acid starting material, with the release of the other imidazole group and one equivalent of carbon dioxide [**III**] to form the reactive

carbonylimidazole. The amide bond is then formed between this intermediate and the added amine to form the desired final product [**IV**].

The initial conditions for the coupling of starting material **20a** and Glu HCl with CDI are described in scheme 4.20.



Scheme 4.20: Reaction conditions for attempted amide coupling with CDI

The carboxylic acid starting material **20a** was dissolved in DMF and CDI was added. The mixture was stirred at room temperature and the consumption of the starting material was monitored by TLC. After 30 minutes, no more starting material could be observed and so Glu HCl was added to the reaction mixture alongside DIPEA. The mixture was then left to stir at room temperature for a further 24 hours. After this time, the mixture was neutralized with aqueous HCl, and a white precipitate formed. The mixture was filtered, and the isolated precipitate subjected to analysis by NMR spectroscopy and mass spectrometry, which revealed it to be the unreacted starting material **20a**, recovered in near quantitative amounts. Slight variations in the reaction conditions described above were then attempted, including heating of the mixture to 60 °C and the use of different bases such as TEA and DIPA, and adjusting the equivalents of base added. The reaction was also repeated in DMSO, but in each case the major product isolated was the unreacted carboxylic acid starting material **20a**. The consumption of the starting material observed by TLC for each attempted reaction before the addition of the glutamic acid salt suggests that the initial addition of the CDI to the carboxylic acid was successful, however, it is unclear as to whether or not the carbonylimidazole intermediate formed, or if the anhydride intermediate hydrolyzed back to the starting material before the subsequent addition of the free imidazole, as described in scheme 4.16. Although less likely, it is also possible that the correct product did form during the reaction, but the acid

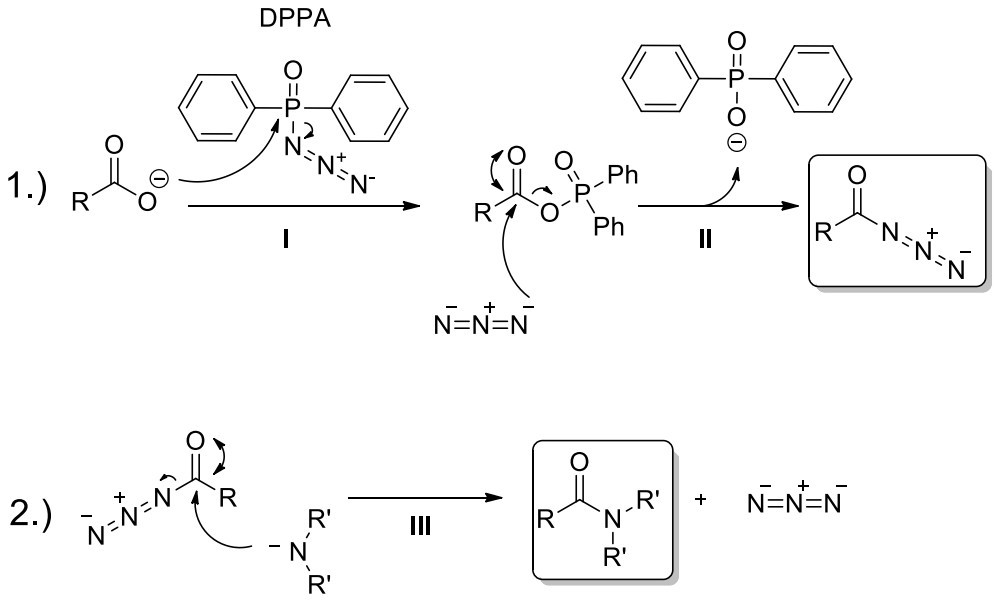
added for neutralization during workup hydrolyzed the amide bond. Alternative acid-free workup methods were assessed, but with no success. Overall, it was decided to abandon the CDI approach due to the poor results obtained.

Before any further attempts were made to achieve the desired transformation, a literature search was conducted to assess the various additional possibilities for this reaction. The search was limited to reactions performed on similar benzoic acid containing starting materials and glutamic acid. Several general methods for this transformation were uncovered, mostly involving various coupling reagents and carboxylic acid starting materials[88]. The only methods found which describe the formation of the amide product directly on ester containing starting materials were patents, which have already been discussed above[86, 87]. For all of the other methods found, the coupling was conducted on carboxylic acid starting materials, often requiring an extra hydrolysis step to form the carboxylic acid from ester precursors before the coupling reaction was performed. More commonly than not, the coupling reactions were performed on methyl or ethyl esters of glutamic acid. This is advantageous because it prevents any possibility of dimerization of the glutamic acid, where the coupling reagents add to one of the carboxylic acids of the glutamic acid starting material instead of the desired site on the other starting material. However this can also be controlled by only adding the glutamic acid after sufficient time has passed for all of the carboxylic acid starting material to be activated by the coupling reagent. In addition, the formed products of the coupling reaction with the ester protected glutamic acid may allow for easier purification due to their decreased polarity compared to the unprotected glutamic acid products. Despite the advantages, however, an extra hydrolysis step is required in these cases to form the final unprotected glutamic acid containing amide product.

Nair et al, 1989[89] and DeGraw et al, 1993[90] reported the synthesis of large libraries of folate-like compounds with glutamic acid substituents. The key amide coupling step was conducted using solid phase synthesis techniques. The carboxylic acid starting materials were adsorbed onto a solid support matrix, before the addition of isobutyl chloroformate coupling reagent and 4-methylmorpholine base, and the final addition of dimethyl glutamic acid. Good yields were reported for the transformation, but an extra hydrolysis step was required after the coupling reaction to deprotect the dimethyl glutamic acid of the coupled product.

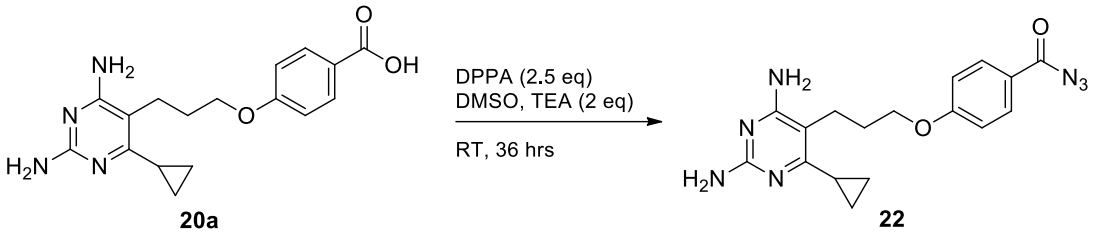
Other coupling reagents employed for similar transformations include diethyl cyanophosphonate (Borrel et al, 2001)[91] and phenyl *N*-phenylphosphoramidochloridate (Varney et al, 1997)[92], which were both described to successfully facilitate the desired coupling reaction on similar substrates, using the methyl or ethyl ester of glutamic acid and a suitable base.

A promising method was reported for the total synthesis of methotrexate and some analogues[93]. Methotrexate is an anticancer folate analogue which contains a glutamic acid substituent. The key amide bond formation was achieved over two distinct steps. The first step [**I** and **II**] involves the formation of a benzoyl azide intermediate which can be achieved in a number of ways, with the best results reported for a reaction between the carboxylic acid starting material and diphenylphosphoryl azide (DPPA) in the presence of triethylamine (TEA). Once isolated, the benzoyl azide was treated with glutamic acid and a suitable base for the formation of the desired amide bond [**III**]. A general reaction scheme for this two-step transformation is provided in scheme 4.21.



Scheme 4.21: General mechanism for amide bond formation through benzoyl azide intermediate

Excellent yields and straightforward purification procedures were reported for both steps, in addition the similarity of the substrates used to the compounds investigated in this project motivated the use of this particular method over some of the others discussed above. The conditions employed for the two-step transformation are included in schemes 4.22 and 4.23.

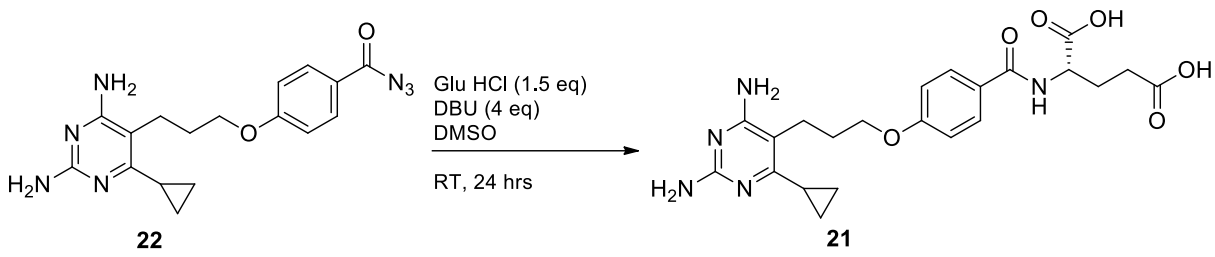


20a **22**

Scheme 4.22: Formation of benzoyl azide intermediate **22**

The starting material **20a** was dissolved in DMF before the addition of DPPA and TEA. The mixture was stirred at room temperature and the progress of the reaction monitored by TLC. After 48 hours, all of the starting material had been consumed. In the literature procedure, the benzoyl azide product was isolated from the reaction mixture by precipitation with THF and subsequent filtration or centrifugation. When this procedure was repeated for the reaction of **20a**, however, no precipitate was formed. The product was instead purified by silica gel column chromatography, allowing for the isolation of the benzoyl azide compound **22** in a yield of 78 %. The success of the reaction was confirmed by high resolution mass spectrometry, where the $[M+H]^+$ peak of the formed product was identified at 354.1671 m/z (354.1673 expected).

With the benzoyl azide intermediate **22** in hand, the next part of the two-step coupling procedure could be attempted, as described in scheme 4.23.



22 21

Scheme 4.23: Conditions for coupling of glutamic acid to benzoyl azide intermediate **22**

The benzoyl azide starting material **22** was dissolved in DMSO followed by the addition of Glu HCl and DBU. An excess of the base was required to first neutralize the HCl salt of glutamic acid, and to further ensure the deprotonation of the amine group for the subsequent addition to the benzoyl azide. After all the reagents were added, the mixture was left to stir at room temperature while monitoring the progress of the reaction with TLC. After 24 hours, no more starting material could be observed and a single new highly polar spot was seen at the base of the TLC plate. The mixture was then filtered through celite to remove any undissolved material, before slowly adding acetonitrile to the filtrate, causing the formed product to precipitate out of solution. The precipitated material was isolated by centrifugation and washed with portions of dilute aqueous HCl (1 %) and diethyl ether before drying under vacuum for several hours.

High resolution mass spectrometry was conducted on the isolated product, confirming the presence of the desired amide compound at an m/z value of 458.2000 (458.2034 expected), along with a large amount of left over DBU from the reaction mixture.

Because of the two carboxylic acid groups present on the glutamic acid adduct, it is likely that the product was isolated as a salt, with the two negative charges of the carboxylic acid groups neutralized by two equivalents of the conjugate acid of DBU formed during the reaction. Due to the zwitterionic nature of the formed product, it was not feasible to isolate it in a neutral form, and for the planned biological assessments, the inclusion of DBU as part of the formed salt was not desirable. Its presence could interfere with the assay results by affecting the cells or enzymes through its own activity, independently of the compound being tested.

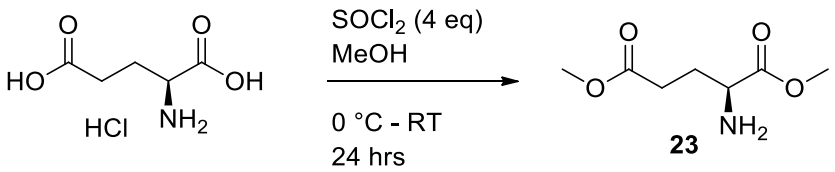
In order to replace the DBU cation with a more biologically inert ion, the formed product was recrystalized from a dilute aqueous solution of NaOH. However, the product isolated from this

procedure still contained significant amounts of DBU. It is possible that the desired ion exchange did not occur, or that the DBU precipitated out of the mixture alongside the sodium salt of the product. This procedure was repeated several times making use of different concentrations of aqueous NaOH and various recrystallization methods, however, the same result was obtained each time. Recrystallization was also attempted with aqueous KOH solutions, but the complete removal of the DBU impurity could not be realized.

Subsequent attempts were made to isolate the product in its conjugate acid form, namely through the neutralization of the carboxylic acid groups and the protonation of one of the basic nitrogens in the 2,4-diaminopyrimidine ring. To this effect recrystallization was attempted from dilute aqueous solutions of HCl, or by bubbling HCl gas through an aqueous solution of the product. However, in each case recovery of the product was very challenging, and the DBU impurity persisted.

While the successful formation of the desired amide compound **21** for the two-step coupling reaction with DPPA was an improvement over the other attempted methods, the complete purification of the final compound was not achieved. For this reason, it was decided to esterify the glutamic acid starting material first before performing the coupling reaction, to hopefully allow for easier purification of the formed amide. The protection of the glutamic acid carboxylic acid groups should allow for the desired product to be isolated in a neutral form under basic conditions, and for subsequent purification by recrystallization or column chromatography. Once the protected amide compound is isolated, the additional hydrolysis step to form the final compound would be conducted.

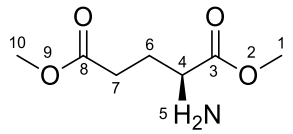
The esterification of glutamic acid was readily achieved by reaction of Glu HCl with $SOCl_2$ in methanol, as described in scheme 4.24.



Scheme 4.24: Esterification of glutamic acid

The glutamic acid HCl salt starting material was dissolved in methanol and the mixture cooled to 0 °C in an ice bath, before the slow addition of thionyl chloride ($SOCl_2$). The mixture was allowed to reach room temperature and stirred for a further 24 hours. TLC analysis showed that by this time all of the starting material had been consumed. The formed product **23** was purified by silica gel column chromatography and isolated as a yellow oil in a yield of 96.5 %.

To confirm the formation of the desired compound, NMR spectroscopy was conducted. The assignments made for the NMR spectra were based on the numbering system defined in figure 4.22.



23

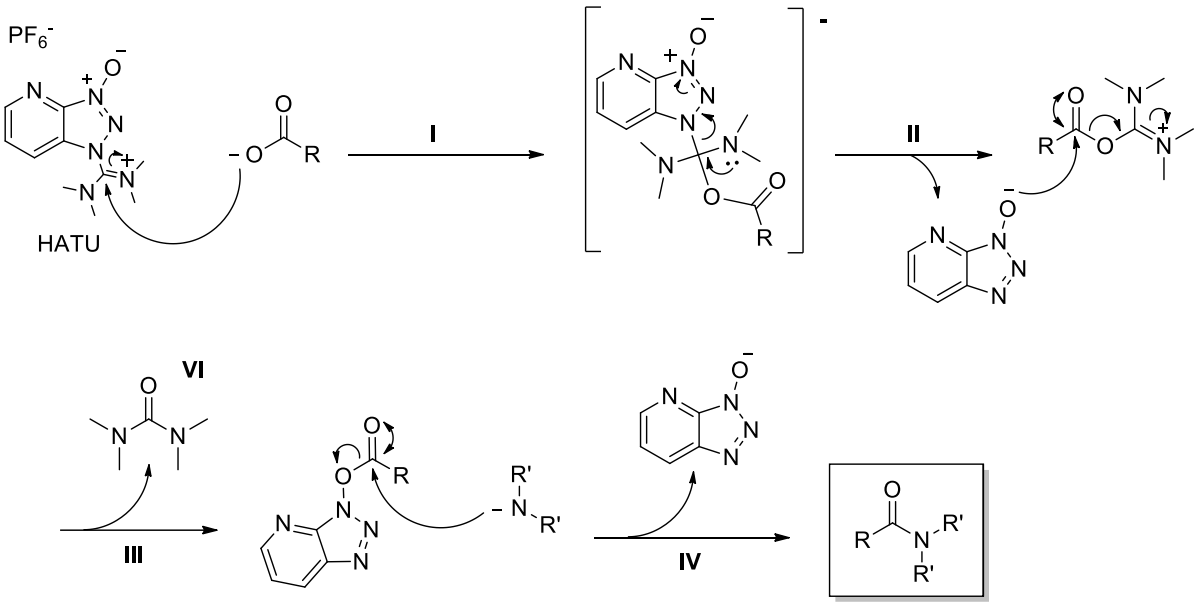
Figure 4.22: Numbering system for dimethyl protected glutamic acid

The ^{1}H and ^{13}C NMR spectra each revealed two peaks corresponding to the added methyl groups. In the ^{1}H NMR spectrum, these peaks appear at 3.72 ppm and 3.67 ppm and are both singlets with an integration of three protons. The peak at 3.72 ppm was assigned to the hydrogens on carbon C1, while the peak at 3.67 was assigned to the C10 hydrogens. The corresponding peaks in the ^{13}C NMR spectrum appear at 52.09 ppm and 51.68 ppm, respectively. The assignments for these signals were made based on coupling seen in the HSQC and HMBC spectra.

With the dimethyl glutamate in hand, the coupling reaction could be attempted once again. Most of the formed carboxylic acid starting material **20a** was lost in the previous attempts, so whilst efforts were made to produce more material, the coupling reaction was performed on a surrogate starting material to hopefully devise conditions which would also work on starting material **20a**.

Because of the extra hydrolysis step required for the deprotection of the glutamic acid after the amide coupling reaction, other single-step methods were assessed over the two-step DPPA mediated transformation to reduce the number of steps required overall.

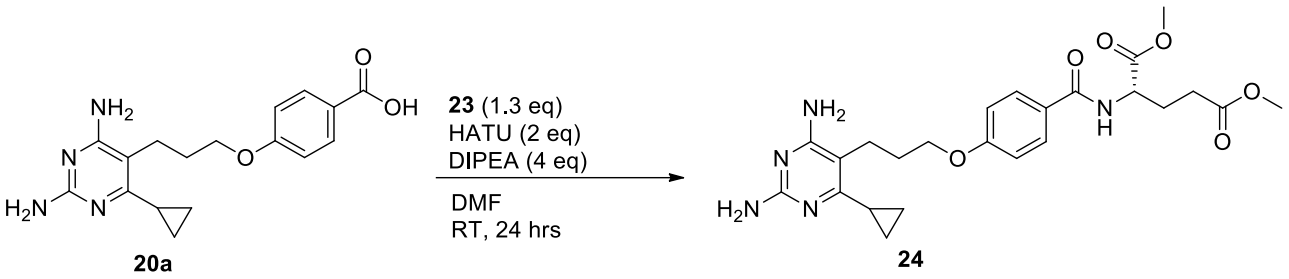
The first attempt made at coupling the dimethyl glutamate involved the use of hexafluorophosphate azabenzotriazole tetramethyl uronium (HATU) as the amide coupling reagent, and afforded excellent results when combined with 4-methoxy benzoic acid, which was selected as the surrogate starting material as it is expected to have a similar electronic environment around the aromatic ring as **20a**. Under standard reaction conditions, the amide product was formed and successfully purified with silica gel column chromatography in a yield of 96 %. The subsequent hydrolysis step also afforded high yields (98 %) and straightforward purification. A general mechanism for amide coupling with HATU is included in scheme 4.25.



Scheme 4.25: General mechanism for amide coupling with HATU

The reaction is initiated by the deprotonation of the carboxylic acid hydrogen by the added base. The resulting carboxylate anion then adds to the central carbon of the guanidine group of HATU [**I**]. The formed transition state rearranges with the loss of the hydroxy benzotriazole type group [**II**]. This then adds back to the carbonyl group replacing the tetramethyl uronium in the formation of a new reactive intermediate [**III**]. After this is formed, the amine can be added to the reaction mixture. The amine becomes deprotonated by the base and reacts with the activated intermediate to form the desired amide product [**IV**].

Due to the success of the coupling reaction with HATU, 4-methoxy benzoic acid, and dimethyl glutamate, the same conditions were applied to the starting material **20a**, as described in scheme 4.26.



Scheme 4.26: Conditions for amide coupling of **20a** and dimethyl glutamate (**23**) with HATU

The starting material **20a** was dissolved in DMF and DIPEA was added followed by HATU. The mixture was left to stir at room temperature while monitored with TLC. After 30 minutes all of the starting material had been consumed, and dimethyl glutamate (**23**) was added. The mixture was stirred at room temperature for 24 hours. After this time, all of the formed intermediate identified on the initial TLC plate had been consumed, and a single new spot was observed at a distinct Rf value to the starting materials. After workup, the product was purified by silica gel column chromatography and isolated as a white solid in a yield of 94.5 %.

The success of the reaction was confirmed by NMR and IR spectroscopy as well as mass spectrometry. The melting point was also recorded. The assignments made for the recorded NMR spectra were made with reference to the numbering system defined in figure 4.23. ^{1}H and ^{13}C NMR spectra of formed product **24** are included in figures 4.24 and 4.25, respectively, with the assignments of the various spectroscopic signals indicated. The provided in text ^{1}H NMR spectrum [Fig 4.24] is after D_2O exchange.

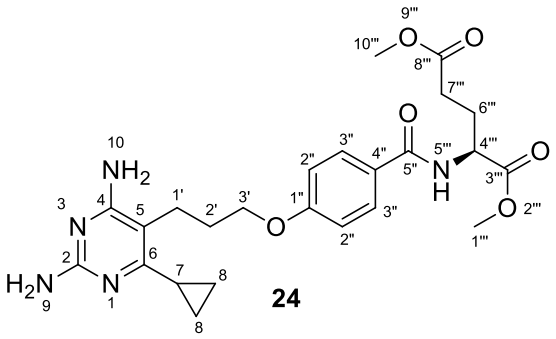


Figure 4.23: Numbering system for dimethyl glutamate adduct **24**

Comparison of the NMR spectra of the product formed in this reaction to the spectra of the starting material **20a** reveals several additional peaks which correspond to the added dimethyl glutamate group. The remaining peaks are highly conserved with the starting material **20a**, which have been discussed above. In the ^{1}H NMR spectrum, seven extra signals are observed for the seven unique hydrogen environments of the added dimethyl glutamate. The two protons attached to carbon atom C6”’ are diastereotopic, and produce a signal each. Both are multiplets which integrate for one proton, centred at 2.15 ppm and 2.31 ppm. The signal for the hydrogens on C7”’ is a multiplet centred at 2.47 ppm and has an integration of two protons. Two singlets integrating to three protons are seen in the ^{1}H NMR spectrum of **24** at 3.77 ppm and 3.64 ppm and were assigned to the hydrogens on carbon atoms C1”’ and C10”’, respectively. The signal for the hydrogen on the secondary amide nitrogen atom N5”’ is a doublet integrating to one proton and is observed at 6.91 ppm. The one remaining hydrogen environment to assign was that of the proton on carbon atom C4”’. Initially the ^{1}H NMR experiment was run in $CDCl_3$ and the signal for this proton was overlapped by one of the amine groups around the core 2,4-diaminopyrimidine ring. After D_2O exchange, the signal is clearly visible as a doublet of doublets at 4.79 ppm with an integration of one proton. The assignments for the ^{1}H NMR spectroscopic signals were made based on the coupling seen in COSY, HSQC, and HMBC spectra and are consistent with both starting materials **20a** and **23**.

Seven signals corresponding to the added dimethyl glutamic acid group appear in the ^{13}C NMR spectrum of compound **24**. The two carbonyl carbons C3”’ and C8”’ are seen at 172.74 ppm and 173.88 ppm, respectively, and were assigned based on the coupling seen in the HMBC spectrum. The methoxy carbons C1”’ and C10”’ show signals at 52.40 ppm and 52.05 ppm,

respectively, with the assignments for these two signals made based on the $^{1}J_{C\text{-}H}$ coupling seen in the HSQC spectrum. The three remaining extra signals arise from the carbons atoms C4”’, C6”’, and C7”’, and appear at 52.73 ppm, 28.46 ppm, and 30.40 ppm, respectively, in the ^{13}C NMR spectrum of compound **24**. The assignments for these peaks were made based on the coupling seen in the HSQC and HMBC spectra.

IR spectroscopy was also conducted on the isolated product, revealing all of the expected vibrational modes. Analysis by high resolution mass spectrometry showed the $[M+H]^+$ peak at an m/z value of 486.2342, which is in agreement with the expected exact mass of the $[M+H]^+$ ion at 486.2347. The melting point for **26** was recorded in the range of 87 – 88 °C.

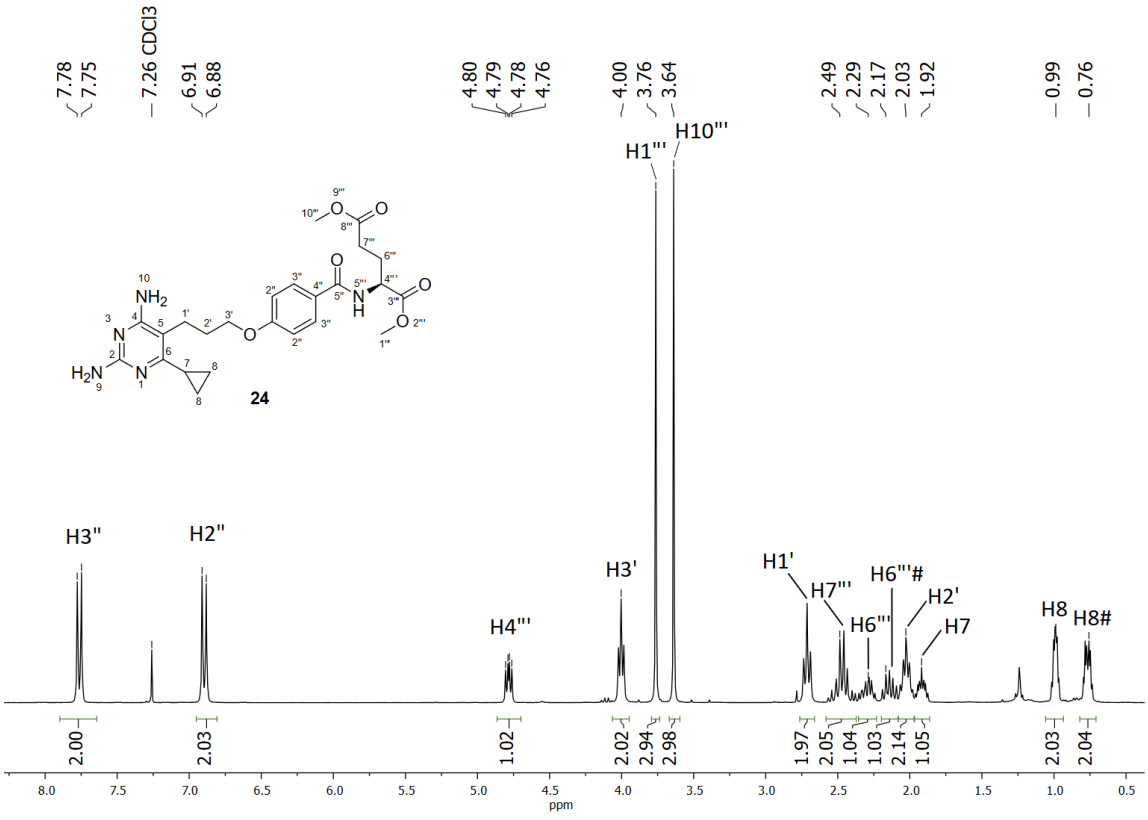


Figure 4.24: ^{1}H NMR spectrum of **24**, after D_2O exchange, with spectroscopic signal assignments included

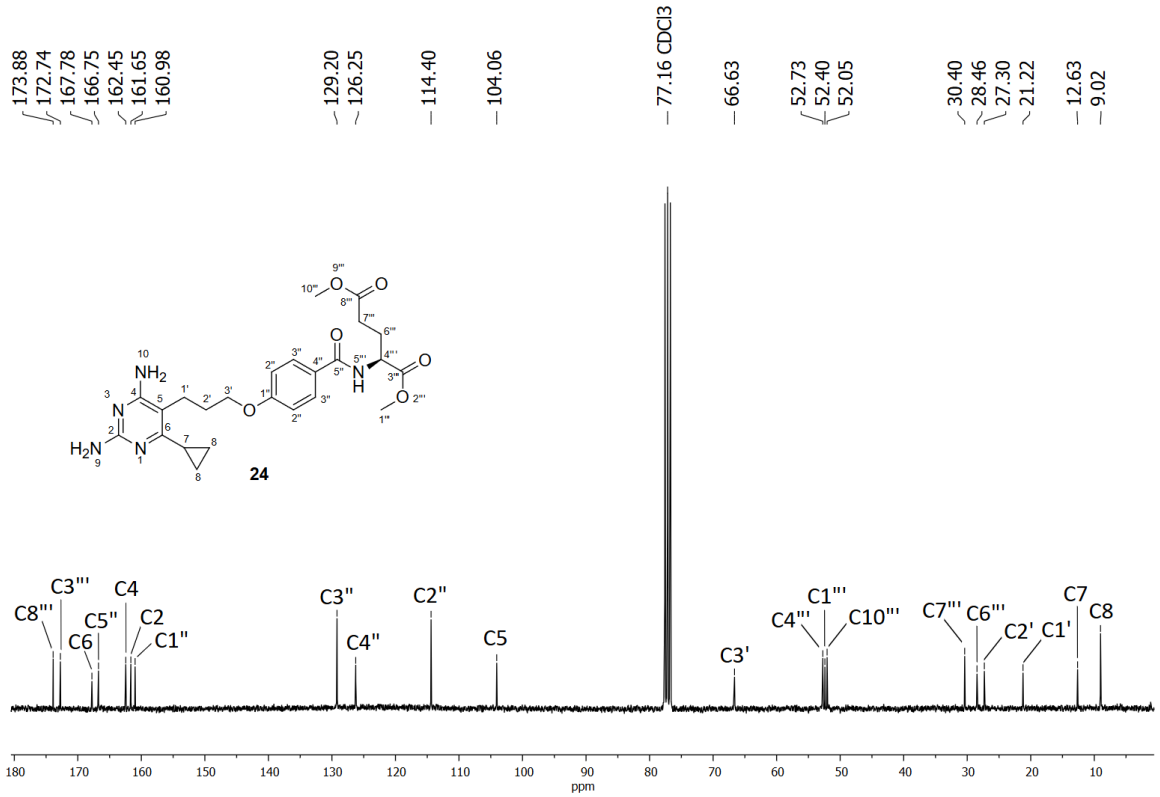
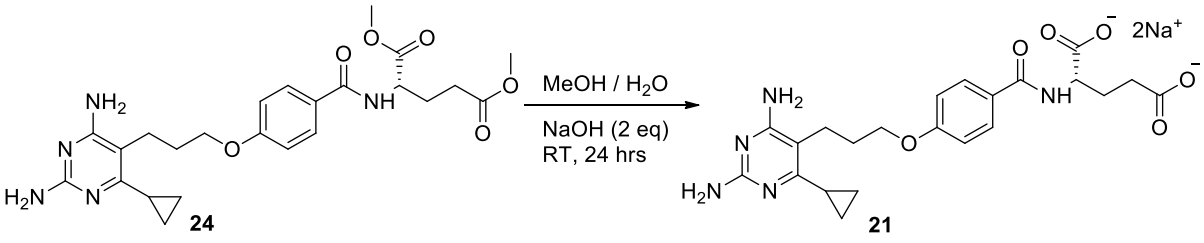


Figure 4.25: ^{13}C NMR spectrum of **24** with spectroscopic signal assignments included

4.2.7 Deprotection of glutamic acid

The hydrolysis reaction for the formation of the final compound **21** was performed next using the standard conditions described in scheme 4.27.



Scheme 4.27: Hydrolysis of dimethyl glutamate adduct **24**

The starting material **24** was dissolved in a 1:1 mixture of methanol and water before the addition of sodium hydroxide. The mixture was left to stir at room temperature whilst

monitoring the progress of the reaction with TLC. After five hours, all of the starting material had been consumed and a single new highly polar spot was observed at the base of the TLC plate. The solvent was then removed on a rotary evaporator, and the residue dissolved in a minimal amount of water. Cold acetone was added, and the product precipitated out of the solution and was subsequently filtered and dried under vacuum for several hours to afford a yield of 96 %.

The success of the reaction was confirmed by NMR and IR spectroscopy and mass spectrometry. The melting point was also recorded. The assignments for the NMR spectra were made with reference to the numbering system defined in figure 4.26.

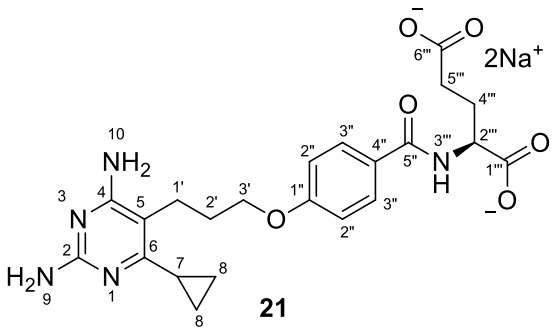


Figure 4.26: Numbering system for hydrolysis product **21**

The NMR spectra for the formed compound **21** differ from the spectra of the starting material **24** mainly by the loss of the peaks corresponding to the methyl ester groups. In both the ^{1}H and ^{13}C NMR spectra of **21**, these two peaks are absent which confirms the desired product was formed.

High resolution mass spectrometry revealed the $[M+H]^+$ at an m/z value of 458.2028, which is in line with the expected mass for this compound of 458.2034. An additional peak with low intensity was observed at an m/z value of 502.1667, which corresponds to the $[M-H+2Na]^+$ form of the compound (502.1673 expected), confirming the isolation of the disodium salt.

The melting point for compound **21** was recorded in the range of 273 – 275 °C, which is significantly higher than the dimethyl ester starting material **24** due to the additional strong intermolecular electrostatic interactions of the formed salt.

4.3 Molecular modelling studies of second-generation analogues

4.3.1 Introduction

Before beginning the synthesis of the second-generation analogues, a brief molecular modelling study was conducted to ensure that the proposed compounds would be able to recreate the desired binding mode, and form interactions in the active site similar to known inhibitors of the enzyme. Compounds were docked into the active site of quadruple mutant *Pf*DHFR enzyme structure 1J3K with GlideXP. Highly conserved docking poses were obtained, with each second-generation analogue seen to occupy nearly identical binding conformations to the first-generation compounds. Some additional interactions were also identified involving the various substituents around the terminal ring. These results were satisfactory and suggested that the new compounds may also be effective inhibitors of *Pf*DHFR, while also hopefully displaying improved pharmacokinetics based on the specific functional groups added.

Upon completion of the synthesis of the target compounds, a more rigorous modelling study was conducted with the aim of assessing the specific nature of any new binding interactions and comparing the overall binding ability of the second-generation analogues to the compounds studied previously in chapter 2. IFD experiments were conducted along with molecular dynamics simulations. Various physiochemical properties were also predicted for comparison with the predicted physiochemical properties of the first-generation analogues, although these results proved to be questionable.

In addition to the five targeted final compounds, some of the various intermediates involved in the synthesis were included in the present molecular modelling studies. Figure 4.27

contains the structures of the 12 analogues assessed *in silico*. The synthesis of compounds **20c**, **21b**, and **22b** was not completed; however, they were also included in the modelling studies as they may be synthesized and subjected to biological evaluations in the future.

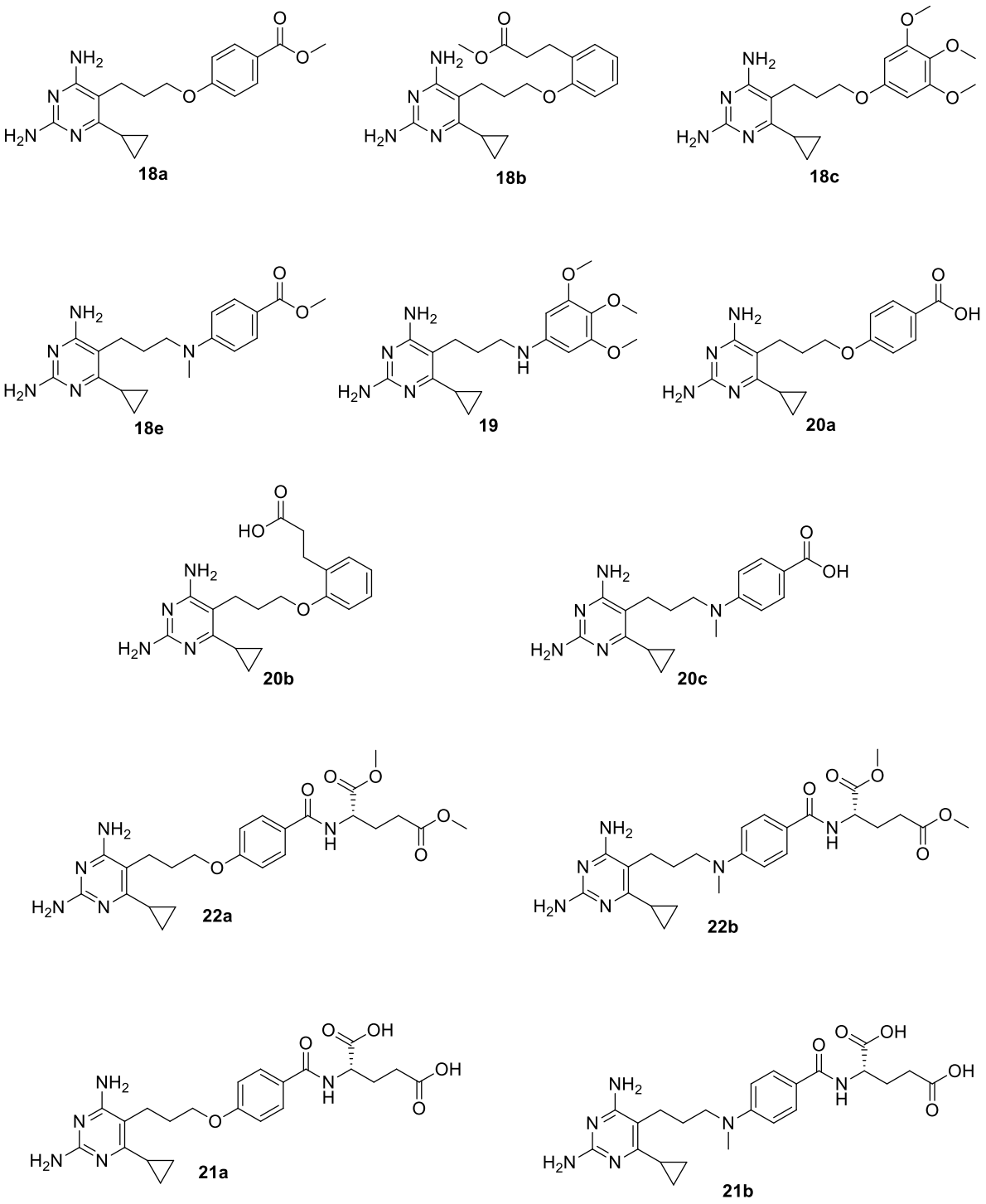


Figure 4.27: Structures of second-generation analogues assessed in molecule modelling studies

4.3.2 Induced fit docking

The prepared protein structure 1J3K, containing the quadruple mutant N51I/C59R/S108N/I164L *Pf*DHFR enzyme was carried over from the initial molecular modelling studies. Before running any docking experiments, the 12 project analogues were subjected to ligand preparation. This generated the possible ionization states at physiological pH and calculated the various state energy parameters which contribute to some of the docking scores discussed later. It has already been demonstrated that protonation of the nitrogen in the 1-position of the core 2,4-diaminopyrimidine ring leads to the most favourable binding interactions, namely an electrostatic interaction and strong hydrogen bonding to residue Asp54. Only this state of the ligands was used in the subsequent docking procedures. It was seen that all of the carboxylic acid groups contained by some analogues were deprotonated in the physiological pH range of 7.4 ± 2. This means that these compounds are zwitterionic, as the basic 1-N of the pyrimidine ring is protonated at this pH.

The prepared ligands were then docked into the active site of enzyme structure 1J3K with IFD. The same experimental parameters described for the IFD experiments carried out in chapter 2 were used here, including the addition of an H-bond docking constraint to residue Asp54. The results for the second-generation analogues are presented in table 4.3, along with the docking scores obtained for some of the first-generation analogues and the co-crystallized ligand of 1J3K, WR99210. For each compound, four scoring parameters are included: Docking Score, IFDScore, Prime energy, and XPGScore. The relevance of each parameter and how they are derived has already been discussed in chapter 2. All of the project analogues assessed have cyclopropyl R-groups.

Table 4.3: IFD results for second-generation analogues, ordered by IFDScore

Name	Number	Docking score (kcal/mol)	**IFDScore** (kcal/mol)	Prime Energy (kcal/mol)	XPGScore (kcal/mol)
O-2propOOH	**20b**	-17.054	-2527.10	-50194.5	-17.375
O-Glu	**21a**	-14.883	-2526.48	-50225.4	-15.211
O-GluOMe	**22a**	-14.111	-2526.00	-50231.4	-14.434
N(Me)-Glu	**21b**	-13.503	-2525.46	-50232.4	-13.844
N(Me)-GluOMe	**22b**	-13.159	-2524.11	-50212.3	-13.494
O-4COOH	**20a**	-14.889	-2523.97	-50174.7	-15.237
N(Me)-4COOH	**20c**	-14.020	-2522.73	-50167.5	-14.354
O-propOOMe	**18b**	-11.318	-2521.36	-50194.4	-11.640
N-triOMe	**19**	-11.824	-2520.72	-50171.2	-12.158
N(Me)-4OOMe	**18e**	-12.251	-2520.62	-50160.6	-12.585
O-4COOMe	**18a**	-11.146	-2519.42	-50159.1	-11.467
O-4Cl*	**9i**	-10.942	-2518.70	-50155.2	-10.942
O-triOMe	**18c**	-9.696	-2517.65	-50152.7	-10.018
O-4OMe*	**9l**	-9.440	-2517.37	-50151.3	-9.797
WR99210		-11.268	-2513.62	-50045.1	-11.364

The compounds are sorted in table 4.3 by IFDScore, with more negative values indicating stronger predicted binding ability. All of the second-generation analogues, including the additional intermediates involved in the synthesis of the final compounds, showed excellent scores for all of the different docking parameters. In general, the carboxylic acid containing compounds **20a – c** and **21a – b** performed better than their methyl ester counterparts **18a**, **18b**, **18e**, and **22a – b**, which in turn showed improved results over the trimethoxy analogues **19** and **18c**. Studying the docking poses produced for each analogue reveals that the carboxylate containing compounds were able to form strong charge mediated hydrogen bonds to the nearby residues Arg59 and Arg122. While H-bonds were still formed between these

residues and the methyl ester analogues, the interactions were not as strong as for the carboxylate compounds as they were not enhanced by the additional electrostatic attraction. Compound **20b**, containing a propionic acid group in the 2-position of the terminal ring, shows the best IFDScore and Docking Score. In general, the second-generation analogues showed improved binding ability over most of the compounds assessed in the initial docking experiments performed in chapter 2. The 2D interaction diagram of **20b** docked into the active site of 1J3k obtained from the IFD experiment is included in figure 4.28, along with the 2D interaction diagram of the Cpr, 4-Cl analogue (**9i**) assessed in chapter 2.

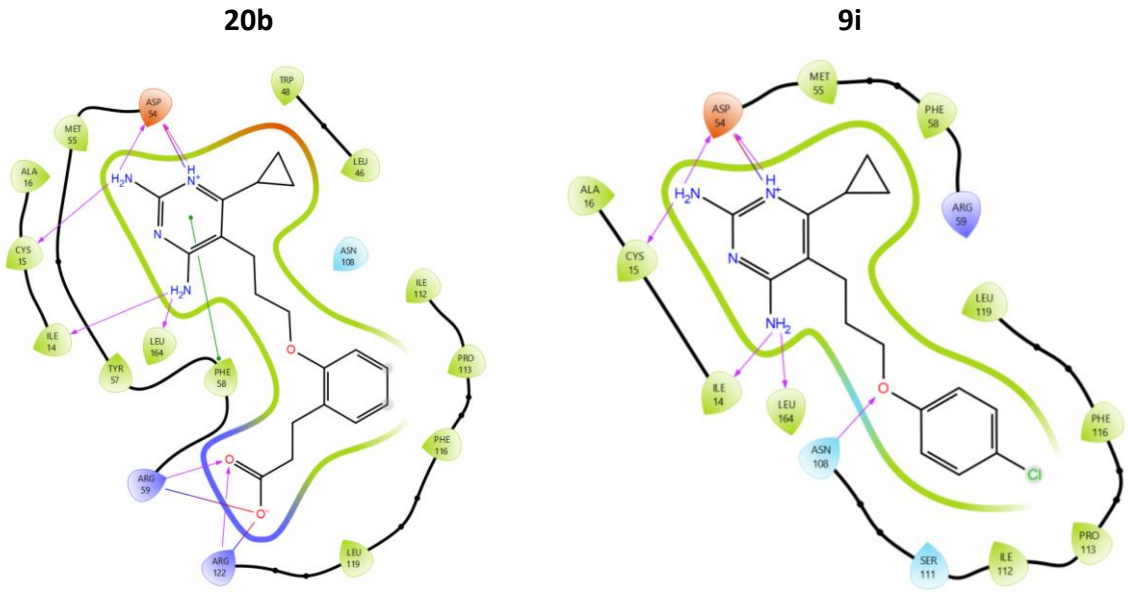


Figure 4.28: 2D interaction diagrams of **20b** and **9i** docked into the active site of 1J3K with IFD

From figure 4.28 it can be seen that all of the conserved interactions involving the core 2,4-diaminopyrimidime ring were adhered to. Including an electrostatic interaction and multiple H-bonds to Asp54, as well as H-bonds to residues Cys15, Leu164, and Ile14. Of note are the additional interactions between the carboxylate group and charged residues Arg59 and Arg122.

Figure 4.29 shows the three-dimensional conformation of the compound **20b** and some nearby residues obtained from the IFD experiment. The carboxylate group is optimally positioned to engage in strong interactions to both adjacent arginines.

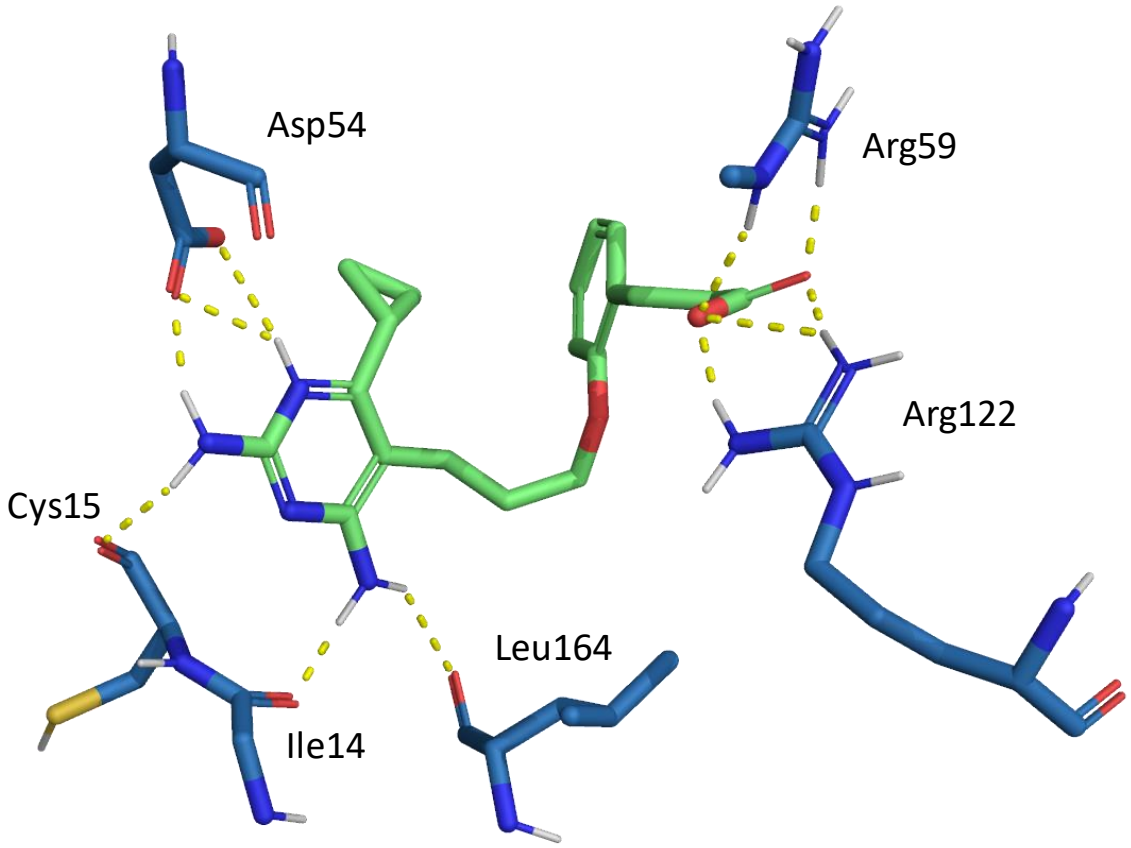


Figure 4.29: Docking pose of **20b** from IFD experiment with the positions of key amino acid residues and specific interactions shown

Arginine residues Arg59 and Arg122 have been identified as key targets for binding interactions in the enzyme active site. Other work has shown the potential for exploiting this, in particular the success of the drug candidate P218[41]. This molecule was envisaged as part of a rational drug design campaign, where the 2-propionic acid group was included to facilitate the specific interactions to residues Arg59 and Arg122. The excellent potential for this particular group is highlighted by the docking scores obtained in the present analysis, where the top performing compound, **20b**, also includes a 2-propionic acid group around its terminal ring. Furthermore, similar interactions were also identified for the various carboxylate containing first-generation analogues assessed in chapter 2.

The docking pose of **20b** can also be compared to the configuration of P218, as found in the

crystal structure 4DP3. This structure contains the quadruple mutant *Pf*DHFR enzyme with P218 bound in the active site. The comparison of the binding poses is shown in figure 4.30, with P218 coloured in orange and **20b** coloured in blue. An alignment score can be calculated between the two poses, however, as they are not the same molecule, the maximum common substructure method is used instead of a straightforward RMSD calculation. This aligns the shared portions of the two distinct compounds, and outputs and RMS value. In the case of P218 and **20b**, the shared portions are the core 2,4-diaminopyrimidine ring, and the 3-phenylpropionic acid moiety. An RMS value of 1.93 Å was calculated between the two structures, indicating that they occupy very similar configurations in the active site. Note how the positions of the carboxylic acid groups are very closely aligned, despite the differences in the lengths of the linkers, as they both engage in ionic interactions to residues Arg59 and Arg122.

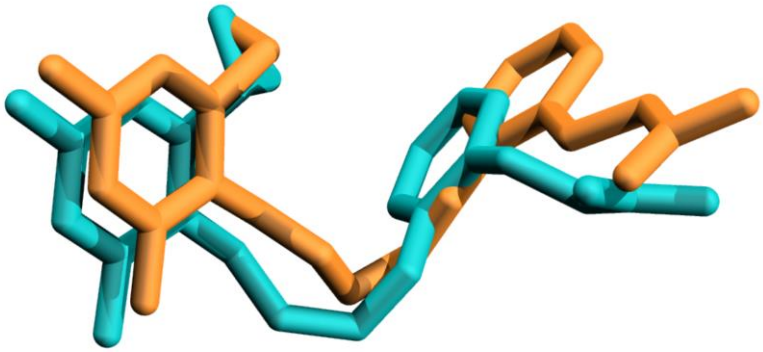


Figure 4.30: Comparison of binding poses of P218 (orange), from protein crystal structure 4DP3, and **20b** (blue), outputted from the IFD experiment

Arginine residues have a basic propyl guanidine side chain, which is positively charged at physiological pH. The propyl linker between the alpha carbon of the amino acid residue and the guanidine group allows for a high degree of flexibility. This was seen in the conducted IFD experiments, as residues Arg59 and Arg122 exhibited significant conformational changes to position the positive charge carried by the guanidine group nearby the carboxylate groups of the docked ligands. A general pharmacophore of arginine is included in figure 4.31, displaying

the various interactions that the residues Arg59 and Arg122 were observed to facilitate with the docked carboxylate containing ligands.

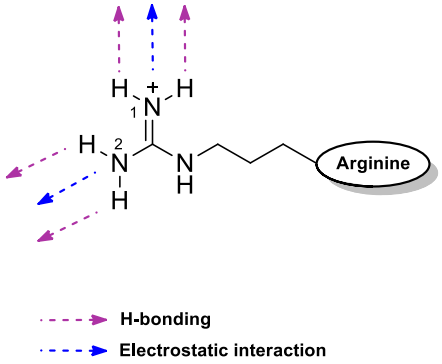


Figure 4.31: Possible binding interactions involving arginine

Each arginine residue can donate multiple H-bonds, and due to resonance around the guanidine group, either of the nitrogen atoms labelled as “1” or “2” in figure 4.31 can carry the positive charge which facilitates the additional electrostatic interaction.

To further study the identified interactions to residues Arg59 and Arg122, the top poses for each carboxylic acid containing ligand obtained from the IFD experiment were superimposed. This revealed that the position of the carboxylate groups was highly conserved between the binding poses of the different ligands. Figure 4.32 superimposes the top scoring binding poses generated for ligands **20a**, **20b**, and **21a** along with the average positions of residues Arg59, Arg122 and Asp54. The conserved charge mediated H-bonds are also included.

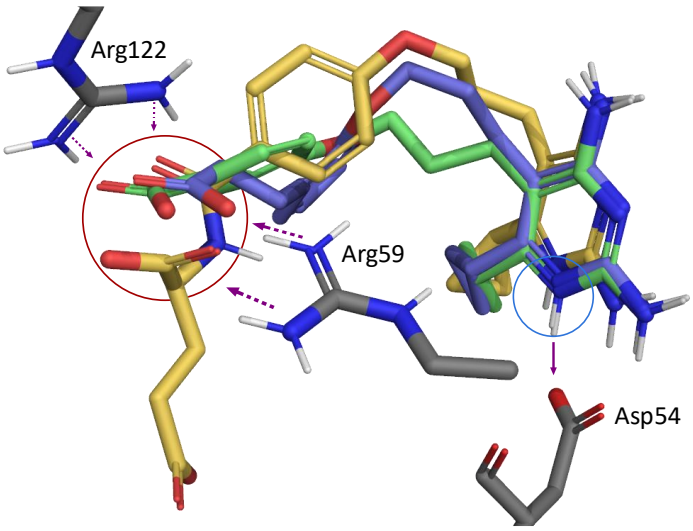


Figure 4.32: Superimposed binding poses of carboxylate containing analogues **20a** (green), **20b** (blue), and **21a** (yellow)

The different analogues orientated themselves in order to satisfy two key binding interactions. The first is the well described charged mediated H-bond from the 1-N protonated nitrogen to Asp54. The second interaction is between one or both residues Arg59 and Asp122 and the carboxylate group, which has a highly conserved position between the different analogues. These two interactions together contributed to the very high IFDScores and Docking Scores obtained from the relevant compounds.

The methyl ester analogues **18a**, **18b**, **18e**, **22a**, and **22b** were also seen to accept H-bonds from the two arginine residues, although did not engage in the additional electrostatic component of this interaction. The trimethoxy compounds **18c** and **19** had no interactions with the arginine residues, but still bound strongly due to the conserved interactions between the groups around the 2,4-diaminopyrimidine core and the adjacent residues. Compound **19**, which has a secondary amine in the 4-position of the linker, shows an additional H-bond to Asn108, which gave it higher docking scores than the oxygen containing equivalent **18c**.

2D Interaction diagrams for some of the ester and trimethoxy analogues are included in figures 4.33 – 4.36.

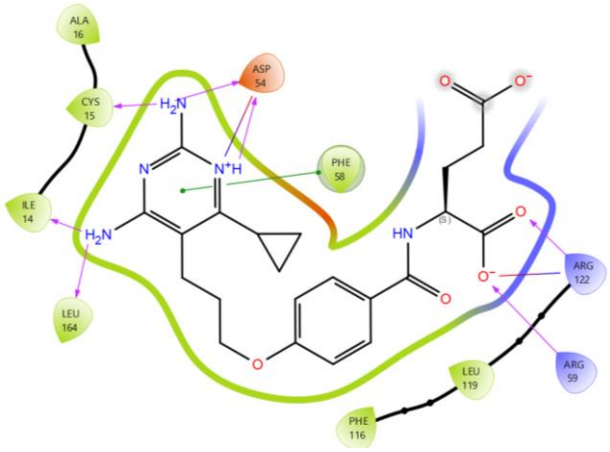


Figure 4.33: 2D interaction diagram of **21a** docked into the active site of 1J3K with IFD

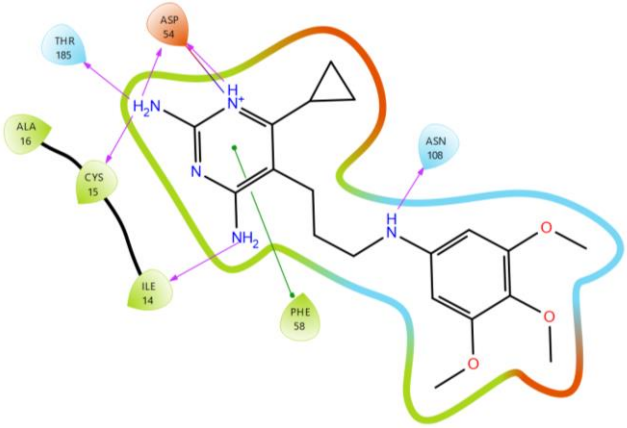


Figure 4.34: 2D interaction diagram of **19** docked into the active site of 1J3K with IFD

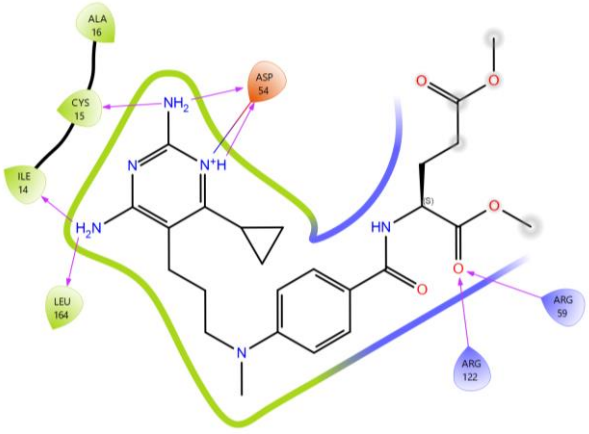


Figure 4.35: 2D interaction diagram of **22b** docked into the active site of 1J3K with IFD

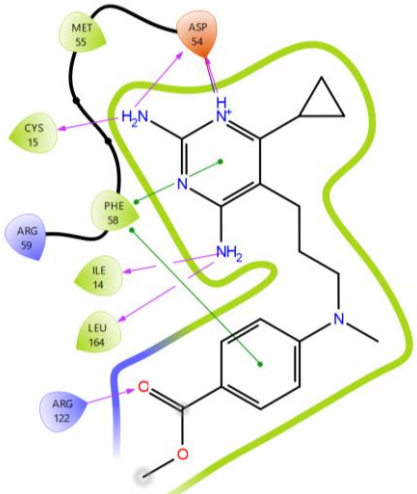


Figure 4.36: 2D interaction diagram of **18e** docked into the active site of 1J3K with IFD

All of the 12 project analogues showed excellent predicted activity when assessed for binding ability against the quadruple mutant *Pf*DHFR enzyme structure 1J3K with IFD. Each compound produced scoring parameters similar to or better than the best performing first-generation analogues assessed in chapter 2.

It must be noted however that the modelling methods used are best suited for identifying hit compounds and predicting specific binding interactions but are less well suited for differentiating between multiple compounds all with favourable binding scores. Furthermore, the design of the second-generation analogues was based on improving the compounds' pharmacokinetics, and not with the aim of improving the binding ability to the target enzyme. Nevertheless the excellent docking scores obtained suggest that the compounds should effectively inhibit the target enzyme in single enzyme assays, and further should hopefully show improved results over the first-generation analogues in the whole cell assays due to the specific functional groups added, which were selected from a range of known drug molecules for improved pharmacokinetics.

4.3.3 Molecular dynamics simulations

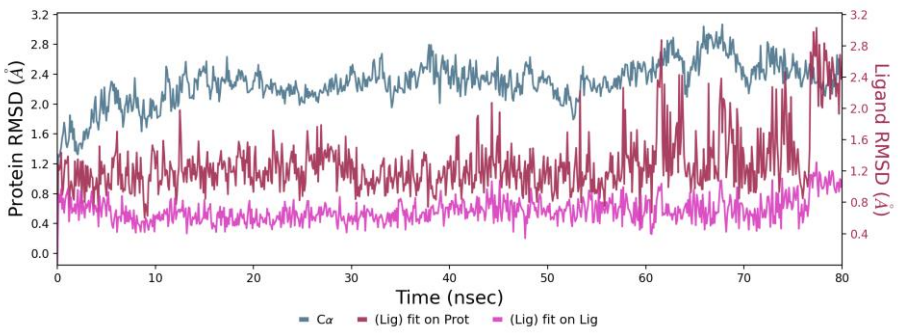
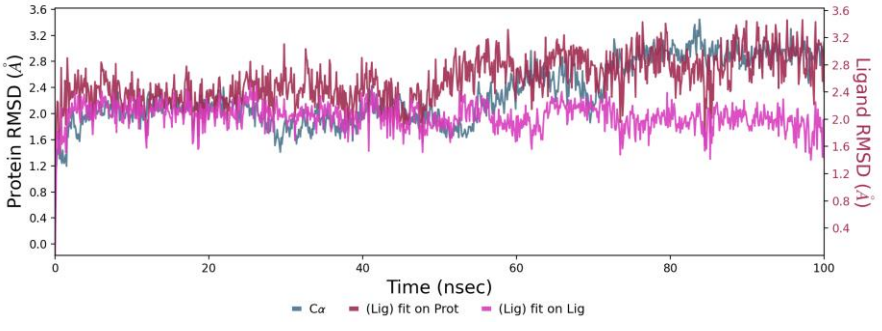
MD experiments were conducted on compounds **19**, **21a**, and **22a** to verify the binding interactions discussed above.

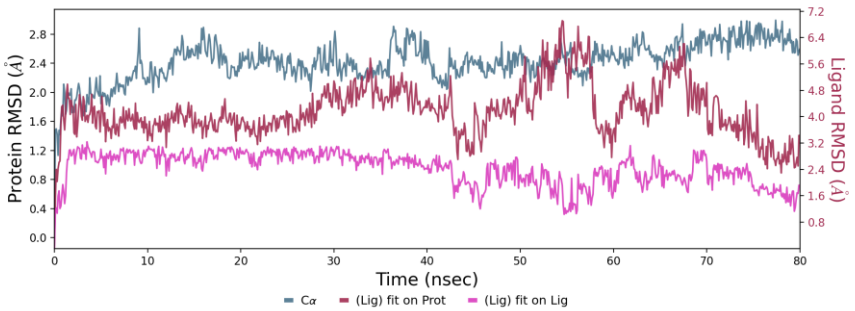
The same methods as described in chapter 2 were used for these simulations, with the initial poses likewise obtained from IFD and the simulations run over 80 ns.

The RMSD plots generated from the MD simulations of the three project analogues are included in table 4.4. The RMSD plot for compound **19** shows a steady state for both the ligand molecule and the protein residues throughout the simulation time. From the start of the simulation up to about 60 ns, the RMSD for the ligand based on the initial protein conformation is seen to be held steady, with minor deviations within 0.5 Å observed. After 60 ns, some larger conformational changes are observed but are still within 1.5 Å of the steady state reached earlier in the simulation. The RMSD plots for compound **21a** also maintain a steady state throughout the simulation time, and show minor variations in the RMSD values of the ligand and amino acid residues within 1 Å. Compound **22a** exhibits greater variation in

the ligand RMSD plots. Initially, it maintains a stable binding configuration up to approximately 45 ns of simulation time. Beyond this point, more significant deviations occur. Analysis of the ligand's conformation throughout the simulation revealed that these deviations primarily arise from fluctuations in the protected glutamic acid portion of the molecule, which shifts near the exterior of the protein active site. Unlike compound **21a**, whose unprotected carboxylic acid groups form strong ionic interactions with Arg59 and Arg122, **22a** lacks these interactions, resulting in weaker retention in the outer regions of the active site. Nevertheless, the core regions of the molecule remain strongly bound in the more buried areas of the active site, experiencing only very minor positional fluctuations.

Table 4.4: RMSD plots for compounds **19**, **21a**, and **22a** generated from MD simulations

N-triOMe (**19**)	
O-Glu (**21a**)	

O-GluOMe (**22a**)	

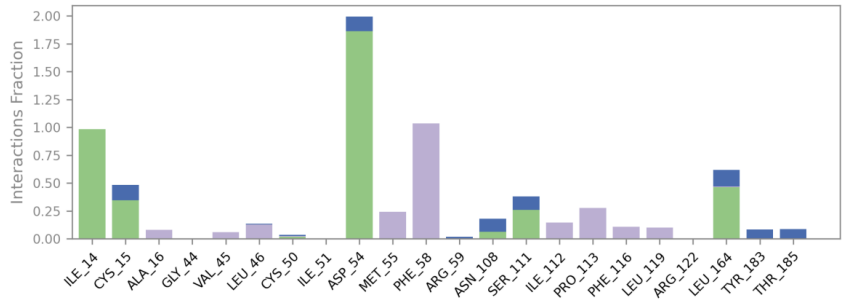
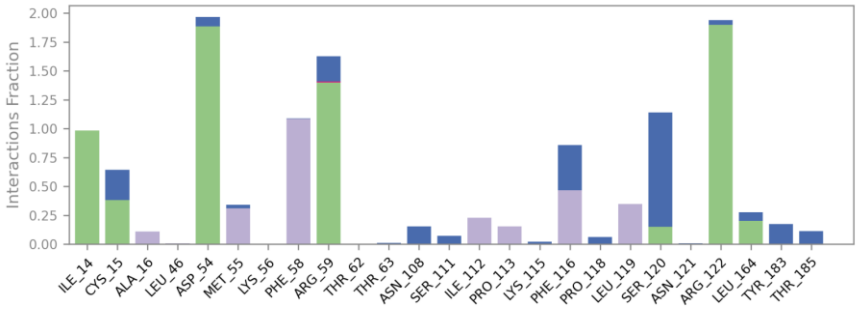
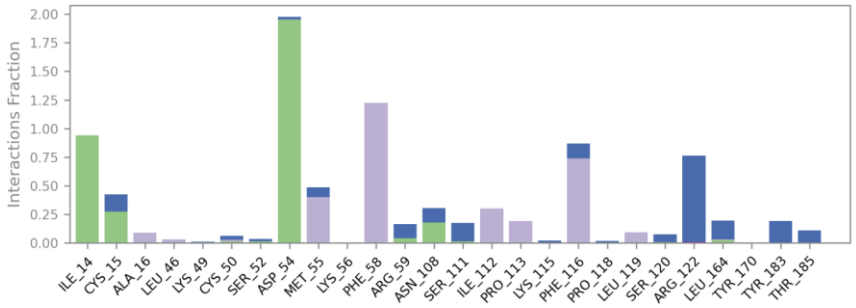
The protein-ligand contacts diagrams generated from the MD simulations for the three assessed compounds are included table 4.5.

These figures display the specific amino acid residues which engaged in interactions with the simulated ligands, and the proportion of the simulation time that these interactions existed for. For each analogue, all of the conserved interactions involving the 2,4-diaminopyrimidine cores were adhered to for the majority of the simulation time, including strong H-bonds to residues Asp54 and Ile14, and additional more transient H-bonding interactions to residues Cys15 and Leu164. Significant Pi-Pi stacking interactions were also observed to the aromatic residues Phe58 and Phe116, with the former residue showing multiple contacts to the ligand molecules throughout the simulation time.

As expected from the results of the IFD experiments, compound **21a**, containing an unprotected glutamic acid substituent, engaged in strong charge mediated H-bonds to residues Arg59 and Arg122 throughout the simulation time. H-bonding interactions were also seen to these residues from the methyl ester analogue **22a**, however, they were not assisted by the electrostatic attraction and occurred to a lesser extent over the course of the simulation.

Overall, compounds **19** and **21a** remained tightly bound in the enzyme active site over the course of the simulation time, while compound **22a** experienced some more significant deviations in its ligand RMSD plots, as discussed above. All three compounds engaged in numerous strong binding interactions with the nearby residues, further suggesting that the second-generation analogues will be effective inhibitors of the target enzyme and validating the interactions discussed from the IFD experiments.

Table 4.5: Protein-ligand contacts diagrams for compounds **19**, **21a**, and **22a**

N-triOMe (**19**)	
O-Glu (**21**)	
O-GluOMe (**22a**)	
Key:	

4.3.4 *In silico* estimation of physiochemical parameters

QikProp was used for the estimation of various relevant physiochemical parameters of the 12 assessed second-generation analogues. Some selected properties are included for each of the compounds in appendix D. The trimethoxy containing analogues **18c** and **19** showed good predicted oral absorptivity, in addition to favourable scores for all the other estimated properties. The other compounds however were flagged for poor oral absorptivity, in particular the highly polar carboxylic acid containing compounds were predicted to be insufficiently lipophilic to effectively cross cell membranes. The methyl ester analogues showed improvements in the related parameters over the carboxylic acid compounds, but still fall slightly outside of the desirable ranges. The QikProp predictions do not take active transport mechanisms into account however, which should hopefully play a role in the distribution of the glutamic acid containing compounds via the various folate transport pathways. Overall, the poor predicted parameters are not too concerning as the excellent predicted physiochemical properties of the first-generation analogues did not result in good activity in the conducted whole cell assays. The strong calculated binding ability of the second-generation analogues and the specific selection of functional groups from existing drug molecules which have known favourable pharmacokinetics motivated their inclusion in single enzyme inhibition and whole cell antiplasmodial assays, which have not yet been conducted at the completion of the present work.

Chapter 5 – Conclusions and future work

A series of flexible pyrimethamine analogues was designed based on prior work done by A. Rousseau and colleagues, and were further optimized for binding ability to the target quadruple mutant drug-resistant *Plasmodium falciparum* DHFR enzyme with molecular modelling techniques. Induced fit docking (IFD) and molecular dynamics (MD) experiments confirmed the compounds' high affinity for the enzyme active site and served to validate several conserved binding interactions. Several analogues were selected for synthesis based on the results of the modelling studies, and were successfully formed over multiple steps making use of partially established experimental procedures and readily available starting materials. The range of functional groups included in the synthesized compounds are summarized in figure 5.1.

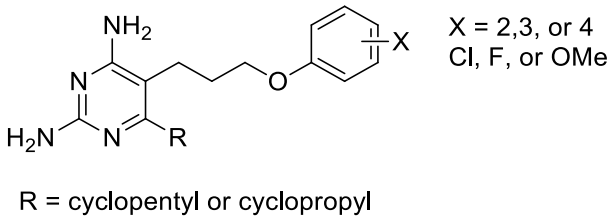
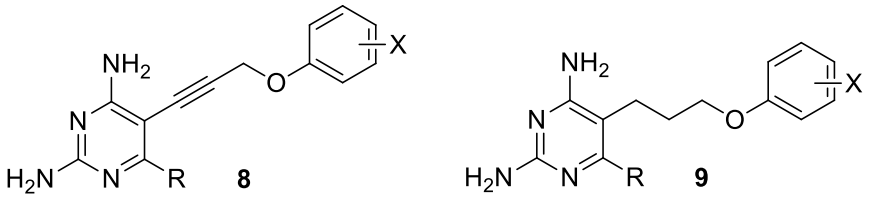


Figure 5.1: General structure of synthesized first-generation compounds

The synthesized compounds were subjected to biological studies in single enzyme inhibition and whole cell assays, which were conducted by collaborators at the BIOTEC research group in Thailand. A total of 28 first-generation analogues synthesized as part of this project were evaluated for biological activity, including 15 analogues with alkyne functional groups present in the four-atom linker (**8**) and a further 13 with fully reduced flexible four-atom linkers (**9**). The compounds displayed excellent activity in the single enzyme assays, inhibiting the drug-resistant *Pf*DHFR enzyme in low nanomolar concentrations. The inhibition constants (Ki) of the synthesized compounds against the quadruple mutant enzyme are summarized in figure 5.2.

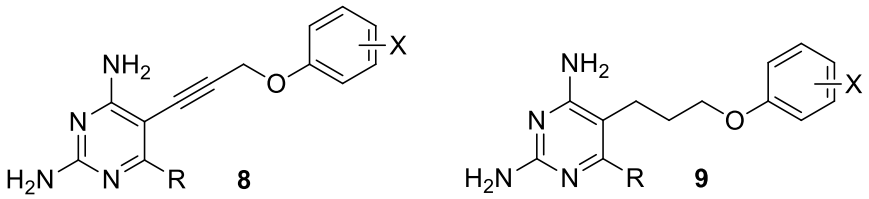


R = Cpe: 1.45 - 20.0 nM
R = Cpr: 5.70 - 20.0 nM

R = Cpe: 0.82 - 34.0 nM
R = Cpr: 8.30 - 43.0 nM

Figure 5.2: Inhibition constants (Ki) of synthesized compounds against QM-*Pf*DHFR enzymes

Both alkyne (**8**) and alkane (**9**) compounds performed similarly in the single enzyme assays, with the most potent compounds inhibiting the drug-resistant quadruple mutant *Pf*DHFR enzyme in concentrations of less than 1 nM. These results were extremely promising, and further whole cell assays were conducted on the synthesized compounds. Disappointingly, a drastic drop in activity was recorded against both wild-type and drug-resistant *P. falciparum* strains in the whole cell assays, with the most active compounds inhibiting the growth of parasite cells in the low micromolar concentration range. The results of the whole cell assays conducted against the drug-resistant quadruple mutant *P. falciparum* V1/S strain are summarized in figure 5.3.



R = Cpe: 26 - >50 µM
R = Cpr: 4.0 - >50 µM

R = Cpe: 5.1 - 21 µM
R = Cpr: 11 - >50 µM

Figure 5.3: Inhibitory concentrations (IC_{50}) for synthesized compounds against quadruple mutant *P. falciparum* strain V1/S

In the whole cell assays, the fully flexible alkane analogues (**9**) showed a marked improvement on average over the alkyne (**8**) compounds, where the most active compounds displayed IC_{50}

values of less than 10 μM, representing a many-fold decrease in activity when compared to the results obtained from the single enzyme inhibition assays.

This suggested that while the synthesized compounds were potent binders of the target enzyme, some pharmacokinetic factors limited their efficacy in the cellular environment of the whole cell assays. A second-generation of project compounds was therefore devised by making modifications to the initial general structure [Fig 5.1], namely by incorporating functional groups found in similar existing drug molecules with known favourable pharmacokinetics to hopefully improve the pharmacokinetic properties of the project compounds. The synthesis of a second-generation of flexible pyrimethamine analogues was successfully conducted, involving the development of new experimental procedures for the formation of the desired final products. Further molecular modelling studies were also conducted on the new range of compounds to validate their potential binding abilities before the relevant biological assessments were conducted. The results of the biological assays for the second-generation analogues have not yet been finalised.

Further work on the project compounds can be done once the results for the biological evaluations of the second-generation analogues are available. Depending on the activity of the synthesized compounds in both the single enzyme and whole cell assays some further modifications may be required to realize the potential of the project compounds as effective class II antifolates against drug-resistant malaria. In addition, the synthesis of some of the devised second-generation analogues was not completed during the present work, and so the synthesis of these extra compounds can also be included as future work.

Additional molecular modelling experiments can also be performed on the human DHFR enzyme to optimize the selectivity of the compounds. This may help identify structural modifications that enhance binding affinity for *Pf*DHFR while minimizing interactions with human DHFR, reducing potential off-target effects. Computational studies, such as induced fit docking or molecular dynamics simulations, can further refine compound design to improve selectivity and efficacy.

Based on the *in vitro* biological assessments of the second-generation analogues, the results of which are still pending, further steps may be taken to advance their development into viable drug candidates. This includes evaluating the potential for resistance development, identifying

key areas of the parasite life cycle that the compounds are the most effective against, conducting comprehensive analyses of their physiochemical properties and pharmacokinetic profiles, and identifying potential partner compounds for combination therapies, among others. Additionally, it may be valuable to screen the synthesized compounds against other *Plasmodium* species to assess their ability as broad spectrum antimalarials. Further investigation into their activity against other Apicomplexan parasites could also provide insights into their broader therapeutic applicability.

Chapter 6 – Experimental methods and procedures

Starting materials were purchased from Sigma-Aldrich (Steinheim, Germany), Merck KGaA (South Africa), and MK Chemicals (South Africa) were of reagent grade and were used without any further purification unless specified. Solvents used for chromatography (ethyl acetate and hexane) were distilled before use.

Reaction monitoring was done with Merck aluminum foil F_{254} plates coated with silica gel 60, and visualized under UV-light (254 nm). For purification with silica gel column chromatography, silica gel 60 (Machery-Nagel, particle size 0.063 – 0.200 mm) was used.

NMR spectroscopic data was collected using either Bruker Avance 300 MHz, Avance III 400 MHz, or Avance III 500 MHz spectrometers, using appropriate deuterated solvents (CDCl3, DMSO-*d*6, acetone-*d*6, MeOD, or D_2O). Infrared spectroscopy was conducted using a Bruker Tensor-27 Fourier Transform spectrometer, sample transmittance was measured in cm^{-1}. Uncorrected melting points were determined using a Stuart SMP10 melting point apparatus. The melting points of samples with reported values were correlated with those available in literature.

High resolution mass spectrometry data were obtained from a Bruker Compact Q-TOF mass spectrometer, with sample diluted to a concentration of 10 µg / ml in HPLC grade methanol. HRMS experiments were run in positive ion mode of the identification of $[M+H]^+$ peaks. Mass spectrometry data was only recorded and reported for novel compounds.

Microwave assisted reactions were conducted using a CEM Discover microwave.

Biological assays were performed by the BIOTEC research group, making use of the following general methods, as described in K. Somandi et al, 2022[71].

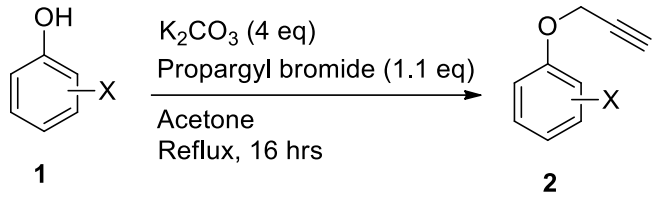
Enzyme preparation and inhibition assays: *Pf*DHFR and *Hs*DHFR enzymes were expressed under T7 promotor of pET17b plasmids in *E. Coli* BL21(DE3)[94] and purified from the crude bacterial lysate of IPTG (isopropyl-β-D-thiogalactopyranoside) induced cells by metho-trexate-agarose affinity chromatography. Enzyme inhibition assay was determined in 96-well plates at 25 °C using Multiskan GO microplate spectrophotometer (Thermo Scientific) and the Ki-values were calculated as previously described[95, 96]. Briefly, the DHFR inhibition assay contains 100 µM each of dihydrofolate (natural substrate) and reduced nicotinamide adenine dinucleotide

phosphate (NADPH, the cofactor), various concentrations of an inhibitor, and PfDHFR. The reactions were monitored at 340 nm for depletion of NADPH by a kinetic plate reader (Labsystems, Finland). The enzyme inhibition constant or Ki values of each inhibitor against the wild-type and mutant enzymes were calculated by non-linear curve fitting using KaleidaGraph program.

Parasite culture, antimalarial and cytotoxicity testing *in vitro*: Two *P. falciparum* strains, TM4/8.2 (wild type PfDHFR, CQ sensitive) and V1/S (QM PfDHFR, CQ resistant) were cultured in human erythrocytes at 37 °C under 3 % CO_2 in RPMI 1640 culture media containing 25 mM HEPES, 0.2 % $NaHCO_3$, 40 µg/mL gentamicin and 8 % human serum. In vitro antimalarial activity was determined by ^{3}H-hypoxanthine incorporation method[96, 97]. Cytotoxicity of each compound was determined against two mammalian derived cell lines, Vero (African green monkey kidney epithelial cells) and KB (human epithelial carcinoma cells) by quantitative staining of cellular proteins using sulforhodamine B protein staining[96, 98].

6.1 Synthesis of first-generation analogues

6.1.1 Alkylation of phenols with propargyl bromide



1

2

1a: X = 4-Cl
1b: X = 3-Cl
1c: X = 2-Cl
1d: X = 4-F
1e: X = 3-F
1f: X = 2-F
1g: X = 4-OMe
1h: X = 3-OMe
1i: X = 2-OMe

2a: X = 4-Cl
2b: X = 3-Cl
2c: X = 2-Cl
2d: X = 4-F
2e: X = 3-F
2f: X = 2-F
2g: X = 4-OMe
2h: X = 3-OMe
2i: X = 2-OMe

Scheme 6.1: Alkylation of phenols with propargyl bromide

The respective phenol starting materials **1a – i** (1.00 g, 1 eq) were dissolved in acetone (20.0 ml) and the flask flushed with N_2. After this K_2CO_3 (4 eq) was added, and the mixture was stirred for ten minutes. Propargyl bromide (1.1 eq) was then added slowly, and the mixture was heated to reflux (55 °C) under N_2 overnight (16 hrs). The mixture was allowed to cool and was filtered through celite to remove the unreacted base. The solvent was evaporated under reduced pressure and the residue dissolved in EtOAc (100 ml). This was washed with saturated aqueous NH_4Cl (100 ml) and brine (100 ml) before being dried with $MgSO_4$ and filtered again through celite. The solvent was removed on a rotary evaporator and the product was purified by silica gel column chromatography (10 % EtOAc / Hex). The reaction information is summarized in scheme 6.1.

6.1.1.1 1-(Prop-2-yn-1-yloxy)-4-chlorobenzene (**2a**)

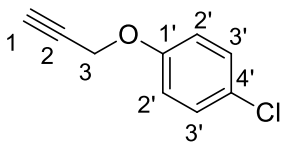


Figure 6.1: 1-(Prop-2-yn-1-yloxy)-4-chlorobenzene (**2a**)

A mixture of 4-chlorophenol **1a** (1.00 g, 7.78 mmol, 1 eq), potassium carbonate (4.30 g, 31.1 mmol, 4 eq), and propargyl bromide (1.02 g, 8.58 mmol, 1.1 eq) in acetone (20.0 ml) formed **2a** [Fig 6.1] as a light-yellow oil after purification (1.08 g, 83.0%).

^{1}H NMR (Chloroform-*d*, 400 MHz): δ = 7.25 (H-3', d, *J* = 7.7 Hz, 2H), 6.90 (H-2', d, *J* = 7.7 Hz, 2H), 4.66 (H-3, s, 2H), 2.54 – 2.51 (H-1, m, 1H); **^{13}C NMR** (Chloroform-*d*, 101 MHz): δ = 156.21 (C-1'), 129.50 (C-3'), 126.68 (C-4'), 116.39 (C-2'), 78.28 (C-2), 75.97 (C-1), 56.18 (C-3); **IR** (vmax / cm^{-1}): 3295 (C≡C-H stretch), 2922 (C-H stretch), 2871 (C-H stretch), 2133 (C≡C stretch), 1487 (C=C stretch), 1216 (C-O stretch); **r.f.** 0.70 (10 % EtOAc / Hex).

6.1.1.2 1-(prop-2-yn-1-yloxy)-3-chlorobenzene (**2b**)

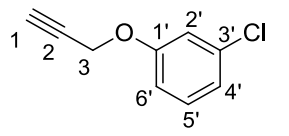


Figure 6.2: 1-(Prop-2-yn-1-yloxy)-3-chlorobenzene (**2b**)

A mixture of 3-chlorophenol **1b** (1.00 g, 7.78 mmol, 1 eq), potassium carbonate (4.30 g, 31.1 mmol, 4 eq), and propargyl bromide (1.02 g, 8.58 mmol, 1.1 eq) in acetone (20.0 ml) formed **2b** [Fig 6.2] as a light-yellow oil after purification (1.15 g, 88.4%).

^{1}H NMR (Chloroform-*d*, 400 MHz): δ = 7.22 (H-5', t, *J* = 8.4 Hz, 1H), 7.00 – 6.97 (H-2' and 4', m, 2H), 6.87 (H-6', ddd, *J* = 8.4, 2.4, 1.0 Hz, 1H), 4.69 (H-3, d, *J* = 2.4 Hz, 2H), 2.54 (H-1, t, *J* = 2.4

Hz, 1H); **^{13}C NMR** (Chloroform-*d*, 101 MHz): δ = 158.37 (C-1'), 135.05 (C-3'), 130.40 (C-5'), 121.95 (C-4'), 115.59 (C-2'), 113.50 (C-6'), 78.13 (C-2), 76.11 (C-1), 56.14 (C-3); **IR** (vmax / cm^{-1}): 3295 (C≡C-H stretch), 2921 (C-H stretch), 2872 (C-H stretch), 2133 (C≡C stretch), 1487 (C=C stretch), 1215 (C-O stretch); **r.f.** 0.61 (10 % EtOAc / Hex).

6.1.1.3 1-(Prop-2-yn-1-yloxy)-2-chlorobenzene (**2c**)

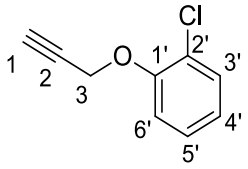


Figure 6.3: 1-(Prop-2-yn-1-yloxy)-2-chlorobenzene (**2c**)

A mixture of 2-chlorophenol **1c** (1.00 g, 7.78 mmol, 1 eq), potassium carbonate (4.30 g, 31.1 mmol, 4 eq), and propargyl bromide (1.02 g, 8.58 mmol, 1.1 eq) in acetone (20.0 ml) formed **2c** [Fig 6.3] as a light-yellow oil after purification (1.21 g, 93.3%).

^{1}H NMR (Chloroform-*d*, 300 MHz): δ = 7.38 (H-3', dd, *J* = 7.9, 1.6 Hz, 1H), 7.27 – 7.20 (H-5', m, 1H), 7.09 (H-6', dd, *J* = 8.3, 1.4 Hz, 1H), 6.95 (H-4', td, *J* = 7.6, 1.4 Hz, 1H), 4.79 (H-3, d, *J* = 2.4 Hz, 2H), 2.54 (H-1, t, *J* = 2.4 Hz, 1H); **^{13}C NMR** (Chloroform-*d*, 75 MHz): δ = 153.27 (C-1'), 130.65 (C-3'), 127.72 (C-5'), 123.49 (C-2'), 122.58 (C-4'), 114.59 (C-6'), 78.16 (C-2), 76.28 (C-1), 56.96 (C-3); **IR** (vmax / cm^{-1}): 3293 (C≡C-H stretch), 2924 (C-H stretch), 2871 (C-H stretch), 2124 (C≡C stretch), 1480 (C=C stretch), 1228 (C-O stretch); **r.f.** 0.65 (10 % EtOAc / Hex).

6.1.1.3 1-(Prop-2-yn-1-yloxy)-4-fluorobenzene (**2d**)

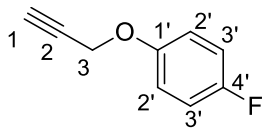


Figure 6.4: 1-(Prop-2-yn-1-yloxy)-4-fluorobenzene (**2d**)

A mixture of 4-fluorophenol **1d** (1.00 g, 8.92 mmol, 1 eq), potassium carbonate (4.93 g, 35.7 mmol, 4 eq), and propargyl bromide (1.17 g, 9.81 mmol, 1.1 eq) in acetone (20.0 ml) formed **2d** [Fig 6.4] as a light-yellow oil after purification (0.998g, 74.5%).

^{1}H NMR (Chloroform-*d*, 400 MHz): δ = 7.05 – 6.88 (H-2’ and H-3’, m, 4H), 4.66 (H-3, d, *J* = 2.4 Hz, 2H), 2.52 (H-1, t, *J* = 2.4 Hz, 1H); **^{13}C NMR** (Chloroform-*d*, 101 MHz): δ = 157.93 (C-4’, d, J_{C-F} = 239.2 Hz), 153.78 (C-1’, d, J_{C-F} = 2.3 Hz), 116.33 (C-2’, d, J_{C-F} = 8.0 Hz), 116.03 (C-3’, d, J_{C-F} = 23.3 Hz), 78.58 (C-2), 75.76 (C-1), 56.64 (C-3); **IR** (νmax / cm^{-1}): 3297 (C≡C-H stretch), 2922 (C-H stretch), 2867 (C-H stretch), 2124 (C≡C stretch), 1501 (C=C stretch), 1194 (C-O stretch); **r.f.** 0.68 (10 % EtOAc / Hex).

6.1.1.5 1-(Prop-2-yn-1-yloxy)-3-fluorobenzene (**2e**)

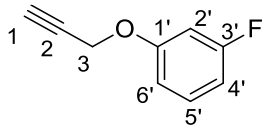


Figure 6.5: 1-(Prop-2-yn-1-yloxy)-3-fluorobenzene (**2e**)

A mixture of 3-fluorophenol **1e** (1.00 g, 8.92 mmol, 1 eq), potassium carbonate (4.93 g, 35.7 mmol, 4 eq), and propargyl bromide (1.17 g, 9.81 mmol, 1.1 eq) in acetone (20.0 ml) formed **2e** [Fig 6.5] as a light-yellow oil after purification (1.11 g, 82.8%).

^{1}H NMR (Chloroform-*d*, 300 MHz): δ = 7.31 – 7.19 (H-5', m, 1H), 6.80 – 6.66 (H-2' and H-4' and H-6', m, 3H), 4.68 (H-3, d, *J* = 2.4 Hz, 2H), 2.54 (H-1, t, *J* = 2.4 Hz, 1H); **^{13}C NMR** (Chloroform-*d*, 75 MHz): δ = 163.67 (C-3', d, J_{C-F} = 245.6 Hz), 158.98 (C-1', d, J_{C-F} = 10.9 Hz), 130.41 (C-5', d, J_{C-F} = 9.9 Hz), 110.76 (C-6', d, J_{C-F} = 3.0 Hz), 108.58 (C-4', d, J_{C-F} = 21.3 Hz), 102.97 (C-2', d, J_{C-F} = 25.1 Hz), 78.19 (C-2), 76.03 (C-1), 56.18 (C-3); **IR** (νmax / cm^{-1}): 3298 (C≡C-H stretch), 2924 (C-H stretch), 2872 (C-H stretch), 2125 (C≡C stretch), 1609 (C=C stretch), 1593 (C=C stretch), 1263 (C-O stretch); **r.f.** 0.64 (10 % EtOAc / Hex).

6.1.1.6 1-(Prop-2-yn-1-yloxy)-2-fluorobenzene (**2f**)

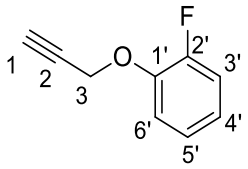


Figure 6.6: 1-(Prop-2-yn-1-yloxy)-2-fluorobenzene (**2f**)

A mixture of 2-fluorophenol **1f** (1.00 g, 8.92 mmol, 1 eq), potassium carbonate (4.93 g, 35.7 mmol, 4 eq), and propargyl bromide (1.17 g, 9.81 mmol, 1.1 eq) in acetone (20.0 ml) formed **2f** [Fig 6.6] as a light-yellow oil after purification (1.25 g, 93.4%).

^{1}H NMR (Chloroform-*d*, 400 MHz): δ = 7.16 – 7.04 (H-3' and H-4' and H-5', m, 3H), 7.00 – 6.93 (H-6', m, 1H), 4.77 (H-3, d, *J* = 2.4 Hz, 2H), 2.54 (H-1, t, *J* = 2.4 Hz, 1H); **^{13}C NMR** (Chloroform-*d*, 101 MHz): δ = 153.18 (C-2', d, J_{C-F} = 246.1 Hz), 145.60 (C-1', d, J_{C-F} = 10.6 Hz), 124.34 (C-4', d, J_{C-F} = 4.0 Hz), 122.51 (C-6', d, J_{C-F} = 7.0 Hz), 116.59 (C-3', d, J_{C-F} = 18.3 Hz), 116.36 (C-5'), 78.24 (C-2), 76.24 (C-1), 57.39 (C-3); **IR** (νmax / cm^{-1}): 3295 (C≡C-H stretch), 2927 (C-H stretch), 2872 (C-H stretch), 2124 (C≡C stretch), 1501 (C=C stretch), 1256 (C-O stretch); **r.f.** 0.67 (10 % EtOAc / Hex).

6.1.1.7 1-(Prop-2-yn-1-yloxy)-4-methoxybenzene (**2g**)

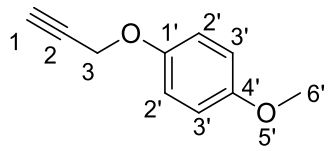


Figure 6.7: 1-(Prop-2-yn-1-yloxy)-4-methoxybenzene (**2g**)

A mixture of 4-methoxyphenol **1g** (1.00 g, 8.06 mmol, 1 eq), potassium carbonate (4.45 g, 32.2 mmol, 4 eq), and propargyl bromide (1.06 g, 8.87 mmol, 1.1 eq) in acetone (20.0 ml) formed **2g** [Fig 6.7] as a light-yellow oil after purification (1.30 g, 99.5%).

^{1}H NMR (Chloroform-*d*, 400 MHz): δ = 6.96 – 6.90 (H-2', m, 2H), 6.88 – 6.82 (H-3', m, 2H), 4.64 (H-3, d, *J* = 2.4 Hz, 2H), 3.77 (H-6', s, 3H), 2.51 (H-1, t, *J* = 2.4 Hz, 1H); **^{13}C NMR** (Chloroform-*d*, 101 MHz): δ = 154.63 (C-4'), 151.82 (C-1'), 116.29 (C-2'), 114.75 (C-3'), 79.04 (C-2), 75.43 (C-1), 56.76 (C-1), 55.83 (C-6'); **IR** (vmax / cm^{-1}): 3286 (C≡C-H stretch), 2835 (C-H stretch), 2120 (C≡C stretch), 1503 (C=C stretch), 1204 (C-O stretch); **r.f.** 0.59 (10 % EtOAc / Hex).

6.1.1.8 1-(Prop-2-yn-1-yloxy)-3-methoxybenzene (**2h**)

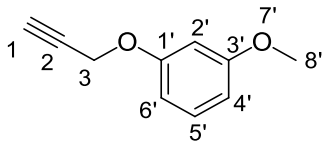


Figure 6.8: 1-(Prop-2-yn-1-yloxy)-3-methoxybenzene (**2h**)

A mixture of 3-methoxyphenol **1h** (1.00 g, 8.06 mmol, 1 eq), potassium carbonate (4.45 g, 32.2 mmol, 4 eq), and propargyl bromide (1.06 g, 8.87 mmol, 1.1 eq) in acetone (20.0 ml) formed **2h** [Fig 6.8] as a light-yellow oil after purification (1.24 g, 94.9%).

^{1}H NMR (Chloroform-*d*, 400 MHz): δ = 7.23 – 7.17 (H-5', m, 1H), 6.61 – 6.51 (H-2' and H-4' and H-6', m, 3H), 4.68 (H-3, d, *J* = 2.4 Hz, 2H), 3.79 (H-8', s, 3H), 2.53 (H-1, t, *J* = 2.4 Hz, 1H); **^{13}C NMR** (Chloroform-*d*, 101 MHz): δ = 160.95 (C-3'), 158.93 (C-1'), 130.06 (C-5'), 107.34 (C-6'), 107.01 (C-4'), 101.61 (C-2'), 78.67 (C-2), 75.64 (C-1), 55.97 (C-3), 55.44 (C-8'); **IR** (νmax / cm^{-1}): 3286 (C≡C-H stretch), 2940 (C-H stretch), 2837 (C-H stretch), 2121 (C≡C stretch), 1590 (C=C stretch), 1145 (C-O stretch); **r.f.** 0.72 (20 % EtOAc / Hex).

6.1.1.9 1-(Prop-2-yn-1-yloxy)-2-methoxybenzene (**2i**)

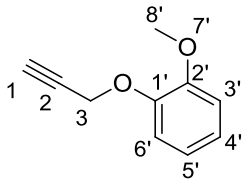
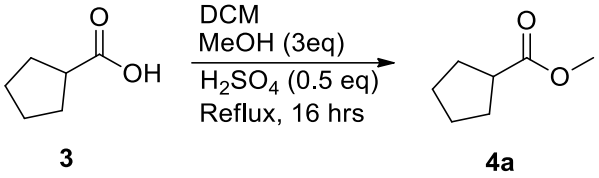


Figure 6.9: 1-(Prop-2-yn-1-yloxy)-2-methoxybenzene (**2i**)

A mixture of 2-methoxyphenol **1i** (1.00 g, 8.06 mmol, 1 eq), potassium carbonate (4.45 g, 32.2 mmol, 4 eq), and propargyl bromide (1.06 g, 8.87 mmol, 1.1 eq) in acetone (20.0 ml) formed **2i** [Fig 6.9] as a light-yellow oil after purification (1.18 g, 90.3%).

^{1}H NMR (Chloroform-*d*, 400 MHz): δ = 7.05 (H-6', d, *J* = 7.7 Hz, 1H), 7.03 – 6.95 (H-3', m, 1H), 6.92 (H-4' and H-5', t, *J* = 7.5 Hz, 2H), 4.77 (H-3, s, 2H), 3.88 (H-8', s, 3H), 2.52 – 2.49 (H-1, m, 1H); **^{13}C NMR** (Chloroform-*d*, 101 MHz): δ = 149.81 (C-2'), 146.86 (C-1'), 122.38 (C-5'), 120.76 (C-4'), 114.59 (C-6'), 111.92 (C-3'), 78.75 (C-2), 75.82 (C-1), 56.83 (C-3), 55.93 (C-8'); **IR** (νmax / cm^{-1}): 3285 (C≡C-H stretch), 2942 (C-H stretch), 2837 (C-H stretch), 2120 (C≡C stretch), 1501 (C=C stretch), 1247 (C-O stretch); **r.f.** 0.79 (30 % EtOAc / Hex).

6.1.2 Synthesis of aliphatic Esters



Scheme 6.2: Synthesis of aliphatic Esters

Cyclopentane carboxylic acid **3** (10.0 g, 87.6 mmol 1 eq) was added to a round bottom flask containing DCM (20.0 ml). MeOH (8.43 g, 263 mmol, 3 eq) and concentrated H_2SO_4 (4.30 g, 43.8 mmol, 0.5 eq) were then added, and the mixture heated at reflux overnight (16 hours). After this, the reaction was quenched slowly with saturated aqueous Na_2CO_3 (200 ml), and the organic components of the mixture were extracted with DCM (200 ml). The organic layer was then washed with brine (200 ml) and dried using $MgSO_4$. The mixture was filtered through celite, and the solvent was removed on a rotary evaporator at room temperature. The product **4a** [Fig 6.10] was isolated without purification as a volatile clear oil in good yield (10.4 g, 92.6 %). Unable to visualize these compounds using TLC. The reaction information is summarized in scheme 6.2.

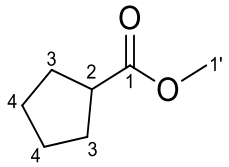
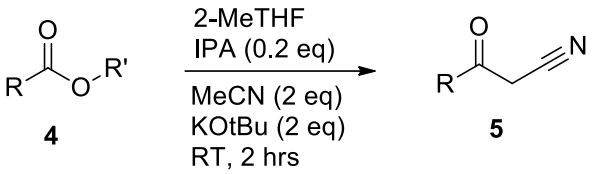


Figure 6.10: Methyl cyclopentane carboxylate (**4a**)

^{1}H NMR (400 MHz, Chloroform-*d*): δ = 3.66 (H-1', s, 3H), 2.72 (H-2, q, *J* = 7.8 Hz, 1H), 1.93 – 1.83 (H-3, m, 2H), 1.83 – 1.73 (H-4, m, 2H), 1.73 – 1.64 (H-3#, m, 2H), 1.63 – 1.50 (H-4#, m, 2H); **^{13}C NMR** (101 MHz, Chloroform-*d*): δ = 177.44 (C-1), 51.73 (C-1'), 43.82 (C-2), 30.15 (C-3), 25.93 (C-4); **IR** (ν_{max} / cm^{-1}): 2953 (C-H stretch), 2872 (C-H stretch), 1732 (C=O stretch), 1157 (C-O stretch).

6.1.3 Synthesis of Beta-ketonitriles (3-oxopropanenitriles)



4a: R = cyclopentyl, R' = methyl
4b: R = cyclopropyl, R' = ethyl

5a: R = cyclopentyl, 54.9% yield
5b: R = cyclopropyl, 54.8% yield

Scheme 6.3: Synthesis of Beta-ketonitriles (3-oxopropanenitriles)

Either the respective ethyl or methyl ester starting material (**4a** or **4b**) (1 eq) was added to a round bottom flask containing 2-MeTHF under N_2. IPA (0.2 eq), MeCN (1 eq), and KOtBu (1 eq) were then added sequentially. The resulting mixture was left to stir under N_2 for one hour. After this, additional portions of MeCN (1 eq) and KOtBu (1 eq) were added, and the mixture stirred for a further hour. The reaction was then quenched by adding saturated aqueous NH_4Cl (200 ml) and the product was extracted into DCM (200 ml). The organic layer was washed with brine (200 ml) before being dried using $MgSO_4$ and filtered through celite. The solvent was removed on a rotary evaporator and the product was purified by silica gel column chromatography (30 % EtOAc / Hex). $KMnO_4$ stain required for beta-ketonitrile product to be seen on a TLC plate. The reaction information is summarized in scheme 6.3.

6.1.3.1 3-Cyclopentyl-3-oxopropanenitrile (**5a**)

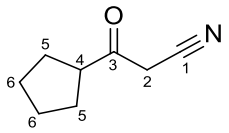


Figure 6.11: 3-Cyclopentyl-3-oxopropanenitrile (**5a**)

Reacting **4a** (10.0 g, 78.1 mmol, 1 eq) with acetonitrile (6.40 g, 156 mmol, 2 eq), potassium *tert*-butoxide (17.5 g, 156 mmol, 2 eq), and isopropyl alcohol (0.937 g, 15.6 mmol, 0.2 eq) in 2-methyl tetrahydrofuran (100 ml) yielded **5a** [Fig 6.11] as a yellow oil after purification (5.88 g, 54.9%).

^{1}H NMR (400 MHz, Chloroform-*d*): δ = 3.51 (H-2, s, 2H), 3.06 (H-4, q, *J* = 7.7 Hz, 1H), 1.96 – 1.84 (H-5, m, 2H), 1.84 – 1.73 (H-6, m, 2H), 1.73 – 1.56 (H-5# and H-6#, m, 4H); **^{13}C NMR** (101 MHz, Chloroform-*d*): δ = 200.02 (C-3), 114.13 (C-1), 50.92 (C-4), 31.30 (C-2), 28.94 (C-5), 26.04 (C-6); **IR** (ν_{max} / cm^{-1}): 2955 (C-H stretch), 2872 (C-H stretch), 2257 (C≡N stretch), 1720 (C-O stretch); **HRMS(ESI)**: *m/z* calculated $C_8H_{12}NO^+$: 138.0913 $[M + H]^+$; found: 138.0919 (4.34 ppm error); **r.f.** 0.36 (40 % EtOAc / Hex).

6.1.3.2 3-Cyclopropyl-3-oxopropanenitrile (**5b**)

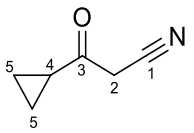
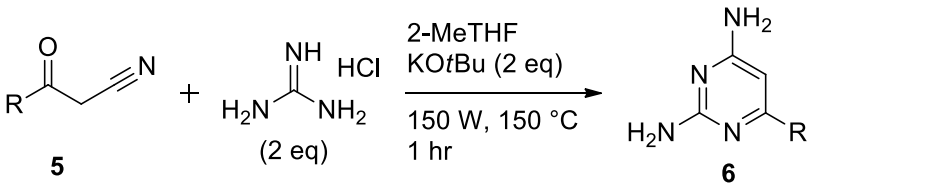


Figure 6.12: 3-Cyclopropyl-3-oxopropanenitrile (**5b**)

Reacting **4b** (5.00 g, 43.8 mmol, 1 eq) with acetonitrile (3.60 g, 87.6 mmol, 2 eq), potassium *tert*-butoxide (9.83 g, 87.6 mmol, 2 eq), and isopropyl alcohol (0.526 g, 8.76 mmol, 0.2 eq) in 2-methyl tetrahydrofuran (50.0 ml) yielded **5b** [Fig 6.12] as a yellow oil after purification (2.62 g, 54.8 %).

^{1}H NMR (400 MHz, Chloroform-*d*): δ = 3.61 (H-2, s, 2H), 2.10 (H-4, tt, *J* = 7.7, 4.5 Hz, 1H), 1.24 – 1.16 (H-5, m, 2H), 1.13 – 1.04 (H-5#, m, 2H); **^{13}C NMR** (101 MHz, Chloroform-*d*): δ = 197.74 (C-3), 114.03 (C-1), 32.78 (C-2), 20.25 (C-4), 13.04 (C-5); **IR** (vmax / cm^{-1}): 3014 (C-H stretch), 2951 (C-H stretch), 2919 (C-H stretch), 2260 (C≡N stretch), 1707 (C=O stretch); **HRMS(ESI)**: *m/z* calculated $C_6H_8NO^+$: 110.0600 $[M + H]^+$; found: 110.0591 (8.18 ppm error); **r.f.** 0.22 (30 % EtOAc / Hex).

6.1.4 Synthesis of 2,4–Diaminopyrimidines



5a: R = cyclopentyl
5b: R = cyclopropyl

6a: R = cyclopentyl, 67.6% yield
6b: R = cyclopropyl, 74.1% yield

Scheme 6.4: Synthesis of 2,4–Diaminopyrimidines

To form the desired 2,4 – diaminopyrimidine products (**6a – b**), the beta-ketonitrile starting material **5a** or **5b** (1.00 g, 1 eq) was added along with guanidine HCl (2 eq) and KOtBu (2 eq) to a sealed microwave vessel. A small amount of 2-MeTHF (4.00 ml) was then added to form a paste. The mixture was heated to 150 °C at 150 W in a microwave reactor for one hour. After this the mixture was allowed to cool, and the crude product was dissolved in hot MeOH (100 ml) and filtered to remove the unreacted base and starting material. The filtrate was then concentrated under vacuum and adsorbed onto silica before purification using column chromatography (10 % MeOH / DCM). The reaction information is summarized in scheme 6.4.

6.1.4.1 6-Cyclopentylpyrimidine-2,4-diamine (**6a**)

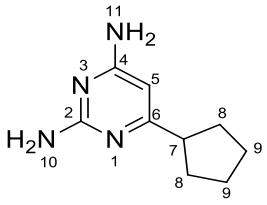


Figure 6.13: 6-Cyclopentylpyrimidine-2,4-diamine (**6a**)

Microwave assisted reaction of **5a** (1.00 g, 7.23 mmol, 1 eq), guanidine hydrochloride (1.39 g, 14.6 mmol, 2 eq), and potassium *tert*-butoxide (1.64 g, 14.6 mmol, 2 eq) in 2-methyl tetrahydrofuran (4 ml) yielded **6a** [Fig 6.13] as a yellow powder after purification (0.871 g, 67.6 %).

m.p. 124 – 126 °C; **^{1}H NMR** (400 MHz, DMSO-d_6): δ = 6.76 (H-11, brs, 2H), 6.33 (H-10, brs, 2H), 5.71 (H-5, s, 1H), 2.74 (H-7, q, J = 7.5 Hz, 1H), 1.94 – 1.82 (H-8, m, 2H), 1.73 – 1.65 (H-9, m, 2H), 1.65 – 1.51 (H-8# and H-9# m, 4H); **^{13}C NMR** (101 MHz, DMSO-d_6): δ = 167.19 (C-6), 164.88 (C-4), 160.64 (C-2), 92.35 (C-5), 44.86 (C-7), 31.49 (C-8), 25.15 (C-9); **IR** (ν_{max} / cm^{-1}): 3321 (N-H stretch), 3157 (N-H stretch), 2949 (C-H stretch), 2868 (C-H stretch), 1633 (C=N stretch), 1545 (C=C stretch); **HRMS(ESI)**: *m/z* calculated $C_9H_{15}N_4^+$: 179.1291 [M + H]$^+$; found: 179.1287 (2.23 ppm error); **r.f.** 0.30 (10 % MeOH / DCM).

6.1.4.2 6-Cyclopropylpyrimidine-2,4-diamine (**6b**)

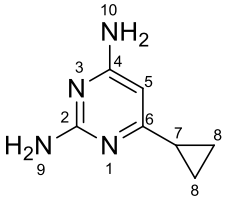
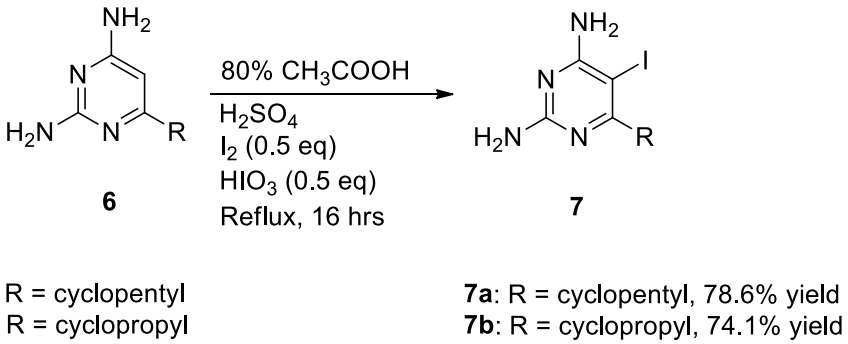


Figure 6.14: 6-Cyclopropylpyrimidine-2,4-diamine (**6b**)

Microwave assisted reaction of **5b** (1.00 g, 9.16 mmol, 1 eq), guanidine hydrochloride (1.75 g, 18.3 mmol, 2 eq), and potassium *tert*-butoxide (2.05 g, 18.3 mmol, 2 eq) in 2-methyl tetrahydrofuran (4 ml) yielded **6b** [Fig 6.14] as a yellow powder after purification (1.02 g, 74.1 %).

m.p. 138 – 140 °C; **^{1}H NMR** (300 MHz, DMSO-d_6): δ = 6.04 (H-10, brs, 2H), 5.65 (H-9, brs, 2H), 5.61 (H-5, s, 1H), 1.67 – 1.57 (H-7, m, 1H), 0.83 – 0.77 (H-8, m, 2H), 0.76 – 0.70 (H-8#, m, 2H); **^{13}C NMR** (75 MHz, DMSO-d_6): δ = 168.82 (C-6), 164.06 (C-4), 163.28 (C-2), 91.79 (C-5), 16.13 (C-7), 7.87 (C-8); **IR** (ν_{max} / cm^{-1}): 3488 (N-H stretch), 3350 (N-H stretch), 3174 (C-H stretch), 3080 (C-H stretch), 1630 (C=N stretch), 1541 (C=C stretch); **HRMS(ESI)**: *m/z* calculated $C_7H_{11}N_4^+$: 151.0978 [M + H]$^+$; found: 151.0980 (1.32 ppm error); **r.f.** 0.21 (10 % MeOH / DCM).

6.1.5 Iodination of 2,4–Diaminopyrimidines

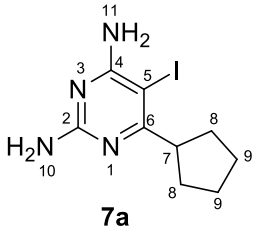


Scheme 6.5: Iodination of 2,4–Diamino–iodopyrimidines

The 2,4-Diaminopyrimidine starting material **6a** or **6b** (0.500 g, 1 eq) was added to a 3-neck round bottom flask along with I_2 (0.5 eq) and HIO_3 (0.5 eq). 80.0 % acetic acid (20.0 ml) was then added followed by a small amount of concentrated H_2SO_4 (0.200 ml). The reaction mixture was refluxed at 100 °C under an N_2 atmosphere overnight (16 hours). After this the mixture was allowed to cool and diluted with H_2O (30.0 ml) before adjusting the pH to 9 using a 10M aqueous NaOH solution. The product was extracted into EtOAc (200 ml), and the organic layer was washed with saturated aqueous $Na_2S_2O_3$ (200 ml) to quench the unreacted

I_2, followed by an additional wash with brine (200 ml). The organic layer was then dried with $MgSO_4$ and filtered through celite before being concentrated under vacuum. The product was purified by column chromatography (60 % EtOAc / hex). The reaction information is summarized in scheme 6.5.

6.1.5.1 6-Cyclopentyl-5-iodopyrimidine-2,4-diamine (**7a**)



7a

Figure 6.15: 6-Cyclopentyl-5-iodopyrimidine-2,4-diamine (**7a**)

A mixture of **6a** (500 mg, 2.81 mmol, 1 eq), iodine (359 mg, 1.41 mmol, 0.5 eq), iodic acid (248 mg, 1.41 mmol, 0.5 eq), 80.0 % acetic acid (10.0 ml), and concentrated sulfuric acid (0.200 ml) yielded **7a** [Fig 6.15] as a yellow crystalline solid after purification (672 mg, 78.6%).

m.p. 143 – 146 °C; **^{1}H NMR** (400 MHz, DMSO-d_6): δ = 6.25 (H-11, brs, 2H), 5.94 (H-10, brs, 2H), 3.29 – 3.21 (H-7, m, 1H), 1.86 – 1.78 (H-8, m, 2H), 1.75 – 1.63 (H-8# and H-9, m, 4H), 1.62 – 1.54 (H-9#, m, 2H); **^{13}C NMR** (101 MHz, DMSO-d_6): δ = 172.26 (C-6), 162.92 (C-4), 162.59 (C-2), 65.08 (C-5), 48.55 (C-7), 31.27 (C-8), 25.57 (C-9); **IR** (ν_{max} / cm^{-1}): 3462 (N-H stretch), 3403 (N-H stretch), 3314 (N-H stretch), 3141 (N-H stretch), 2954 (C-H stretch), 2856 (C-H stretch), 1620 (C=N stretch), 1544 (C=C stretch); **HRMS(ESI)**: *m/z* calculated $C_9H_{15}N_4^+$: 305.0258 [M + H]$^+$; found: 305.0285 (8.85 ppm error); **r.f.** 0.52 (70 % EtOAc / Hex).

6.1.5.2 6-Cyclopropyl-5-iodopyrimidine-2,4-diamine (**7b**)

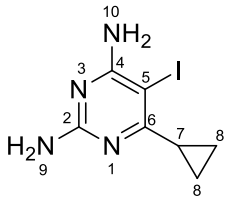
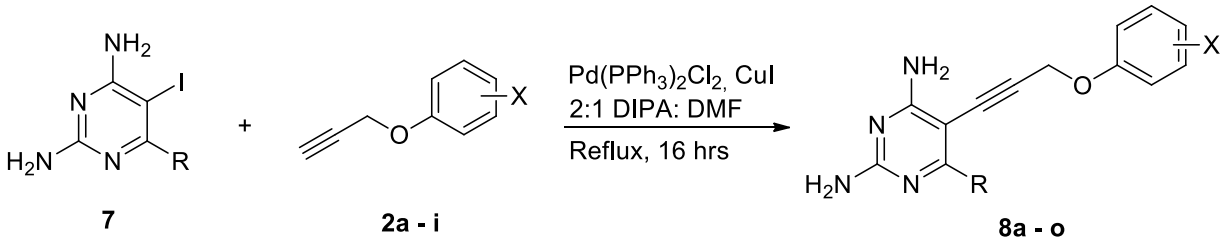


Figure 6.16: 6-Cyclopropyl-5-iodopyrimidine-2,4-diamine (**7b**)

A mixture of **6b** (500 mg, 3.33 mmol, 1 eq), iodine (421 mg, 1.66 mmol, 0.5 eq), iodic acid (292 mg, 1.66 mmol, 0.5 eq), 80.0 % acetic acid (10.0 ml), and concentrated sulfuric acid (0.200 ml) yielded **7b** [Fig 6.16] as a yellow crystalline solid after purification (681 mg, 74.1%).

m.p. 145 – 146 °C; **^{1}H NMR** (400 MHz, Chloroform-*d*): δ = 5.22 (H-10, brs, 2H), 4.70 (H-9, brs, 2H), 2.25 (H-7, tt, *J* = 8.1, 4.7 Hz, 1H), 1.05 – 0.97 (H-8, m, 2H), 0.95 – 0.86 (H-8#, m, 2H); **^{13}C NMR** (101 MHz, Chloroform-*d*): δ = 171.53 (C-6), 162.45 (C-4), 162.37 (C-2), 66.33 (C-5), 19.94 (C-7), 10.19 (C-8); **IR** (ν_{max} / cm^{-1}): 3402 (N-H stretch), 3307 (N-H stretch), 3173 (C-H stretch), 1620 (C=N stretch), 1545 (C=N stretch); **HRMS(ESI)**: *m/z* calculated $C_7H_{10}IN_4^+$: 276.9945 [M + H]$^+$; found: 276.9942 (1.08 ppm error); **r.f.** 0.47 (70 % EtOAc / Hex).

6.1.6 Sonogashira coupling reaction



7a: R = cyclopenyl
7b: R = cyclopropyl

Scheme 6.6: Sonogashira coupling reaction

DMF (5.00 ml) was added to a 50.0 ml 3-neck round bottom flask equipped with a condenser. The solvent was degassed under vacuum and Ar was bubbled through for several minutes to remove any dissolved oxygen. DIPA (10.0 ml) was then added, and Ar bubbled through the solution for several more minutes. After this the 2,4-diamino-5-iodopyrimidine starting material **7a** or **7b** (200 – 300 mg, 1 eq) was added together with $Pd(PPh_3)_2Cl_2$ (0.1 eq) and CuI (0.2 eq). The mixture was then slowly heated to reflux (80 °C) while stirring until all the solid material was dissolved. The relevant previously prepared alkyne (**2a – i**) (2 eq) was then added, and the resulting solution heated overnight (16 hours) under Ar. After this the mixture was allowed to cool and was quenched with saturated aqueous NH_4Cl (200 ml). The product was extracted into EtOAc (200ml), and the organic layer was washed with several portions of brine (3 x 200 ml) and dried using $MgSO_4$. The mixture was then filtered through celite and concentrated on a rotary evaporator. The crude product was adsorbed onto silica and purified by column chromatography (70 % EtOAc / hex). The reaction information is summarized in scheme 6.6, and the full list of products formed from this reaction is displayed in table 6.1.

Table 6.1: List of Sonogashira coupling reaction products (8)

Pyrimidine:	**Alkyne:**	**Product:**
7a, R = Cpe	**2a**, X = 4-Cl	**8a**
	2b, X = 3-Cl	**8b**
	2c, X = 2-Cl	**8c**
	2d, X = 4-F	**8d**
	2e, X = 3-F	**8e**
	2f, X = 2-F	**8f**
	2g, X = 4-OMe	**8g**
	2h, X = 3-OMe	**8h**
	2i, X = 2-OMe	**8i**
7b, R = Cpr	**2a**, X = 4-Cl	**8j**
	2b, X = 3-Cl	**8k**
	2d, X = 4-F	**8l**
	2e, X = 3-F	**8m**

Pyrimidine:	Alkyne:	Product:
	2g, X = 4-OMe	**8n**
	2h, X = 3-OMe	**8o**

6.1.6.1 6-Cyclopentyl-5-[3-(4-chlorophenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8a**)

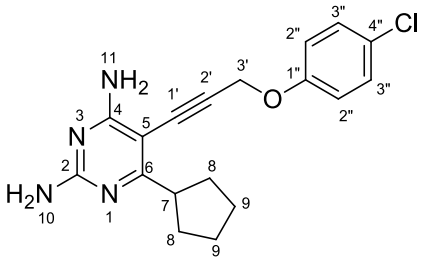


Figure 6.17: 6-Cyclopentyl-5-[3-(4-chlorophenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8a**)

A mixture of **7a** (200 mg, 0.658 mmol, 1 eq), bis(triphenylphosphine)palladium(II) dichloride (46.2 mg, 0.0658 mmol, 0.1 eq), copper iodide (25.1 mg, 0.132 mmol, 0.2 eq), **2a** (220 mg, 1.32 mmol, 2 eq), dimethylformamide (5.00 ml), and diisopropylamine (10.0 ml) yielded **13a** [Fig 6.17] as a yellow polycrystalline material after purification (98.0 mg, 43.4 %)

m.p. 65 – 68 °C; **^{1}H NMR** (400 MHz, Chloroform-*d*): δ = 7.26 (H-3”, d, *J* = 8.8 Hz, 2H), 6.95 (H-2”, d, *J* = 8.8 Hz, 2H), 5.10 (H-11, brs, 2H), 4.96 (H-3’, s, 2H), 4.94 (H-10, brs, 2H), 3.27 (H-7, q, *J* = 8.1 Hz, 1H), 1.87 – 1.78 (H-8, m, 2H), 1.80 – 1.69 (H-8# and H-9, m, 4H), 1.63 – 1.55 (H-9#, m, 2H); **^{13}C NMR** (101 MHz, Chloroform-*d*): δ = 176.58 (C-6), 164.52 (C-4), 161.47 (C-2), 156.20 (C-1”), 129.46 (C-3”), 126.61 (C-4”), 116.61 (C-2”), 92.87 (C-1’), 89.30 (C-5), 81.82 (C-2’), 57.20 (C-3’), 45.09 (C-7), 31.88 (C-8), 26.21 (C-9); **IR** (vmax / cm^{-1}): 3590 (N-H stretch), 3487 (N-H stretch), 3356 (N-H stretch), 3148 (N-H stretch), 2947 (C-H stretch), 2869 (C-H stretch), 2216 (C≡C stretch), 1662 (C=N stretch), 1618 (C=C stretch), 1539 (C=C stretch); **HRMS(ESI)**: *m/z*

calculated for $C_{18}H_{20}ClN_4O^+$: 343.1320 [M + H]$^+$; found: 343.1291 (8.45 ppm error); **r.f.** 0.23 (70 % EtOAc / Hex).

6.1.6.2 6-Cyclopentyl-5-[3-(3-chlorophenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8b**)

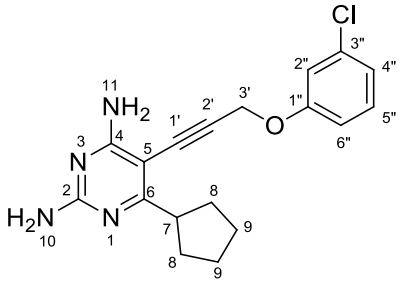


Figure 6.18: 6-Cyclopentyl-5-[3-(3-chlorophenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8b**)

A mixture of **7a** (200 mg, 0.658 mmol, 1 eq), bis(triphenylphosphine)palladium(II) dichloride (46.2 mg, 0.0658 mmol, 0.1 eq), copper iodide (25.1 mg, 0.132 mmol, 0.2 eq), **2b** (220 mg, 1.32 mmol, 2 eq), dimethylformamide (5.00 ml), and diisopropylamine (10.0 ml) yielded **8b** [Fig 6.18] as a yellow polycrystalline material after purification (87.4 mg, 38.7 %).

m.p. 95 – 98 °C; **^{1}H NMR** (400 MHz, Chloroform-*d*): δ = 7.22 (H-5”, t, *J* = 8.1 Hz, 1H), 7.03 (H-2”, s, 1H), 6.98 (H-4”, d, *J* = 8.0 Hz, 1H), 6.90 (H-6”, d, *J* = 8.3 Hz, 1H), 5.13 (H-11, brs, 2H), 4.97 (H-3’, s, 2H), 4.95 (H-10, brs, 2H), 3.31 (H-7, q, *J* = 8.3 Hz, 1H), 1.91 – 1.81 (H-8, m, 2H), 1.81 – 1.69 (H-8# and H-9, m, 4H), 1.66 – 1.55 (H-9#, m, 2H); **^{13}C NMR** (126 MHz, Chloroform-*d*): δ = 176.69 (C-6), 164.43 (C-4), 161.36 (C-2), 158.23 (C-1”), 134.92 (C-3”), 130.27 (C-5”), 121.70 (C-4”), 115.47 (C-2”), 113.75 (C-6”), 92.63 (C-1’), 89.23 (C-5), 81.90 (C-2’), 57.00 (C-3’), 44.95 (C-7), 31.81 (C-8), 26.15 (C-9); **IR** (vmax / cm^{-1}): 3527 (N-H stretch), 3453 (N-H stretch), 3312 (N-H stretch), 3155 (N-H stretch), 2925 (C-H stretch), 2866 (C-H stretch), 2216 (C≡C stretch), 1643 (C=N stretch), 1591 (C=C stretch), 1548 (C=C stretch); **HRMS(ESI)**: *m/z* calculated for $C_{18}H_{20}ClN_4O^+$: 343.1320 [M + H]$^+$; found: 343.1297 (6.70 ppm error); **r.f.** 0.26 (70 % EtOAc / Hex).

6.1.6.3 6-Cyclopentyl-5-[3-(2-chlorophenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8c**)

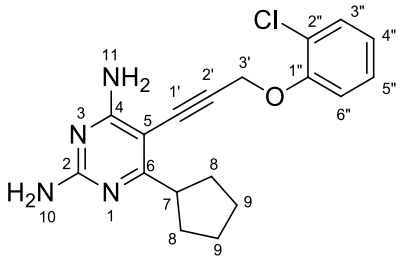


Figure 6.19: 6-Cyclopentyl-5-[3-(2-chlorophenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8c**)

A mixture of **7a** (200 mg, 0.658 mmol, 1 eq), bis(triphenylphosphine)palladium(II) dichloride (46.2 mg, 0.0658 mmol, 0.1 eq), copper iodide (25.1 mg, 0.132 mmol, 0.2 eq), **2c** (220 mg, 1.32 mmol, 2 eq), dimethylformamide (5.00 ml), and diisopropylamine (10.0 ml) yielded **8c** [Fig 6.19] as a yellow polycrystalline material after purification (77.2 mg, 34.2 %).

m.p. 57 – 59 °C; **^{1}H NMR** (400 MHz, Chloroform-*d*): δ = 7.39 (H-3”, d, *J* = 7.9 Hz, 1H), 7.23 (H-5”, t, *J* = 7.8 Hz, 1H), 7.13 (H-6”, d, *J* = 8.2 Hz, 1H), 6.95 (H-4”, t, *J* = 7.6 Hz, 1H), 5.10 (H-11, brs, 2H), 5.08 (H-3’, s, 2H), 4.94 (H-10, brs, 2H), 3.28 (H-7, q, *J* = 7.7 Hz, 1H), 1.86 – 1.80 (H-8, m, 2H), 1.79 – 1.71 (H-8#, H-9, m, 4H), 1.63 – 1.55 (H-9#, m, 2H); **^{13}C NMR** (101 MHz, Chloroform-*d*): δ = 176.65 (C-6), 164.52 (C-4), 161.45 (C-2), 153.24 (C-1”), 130.75 (C-3”), 127.57 (C-5”), 123.73 (C-2”), 122.55 (C-4”), 114.80 (C-6”), 92.80 (C-1’), 89.42 (C-5), 82.13 (C-2’), 58.02 (C-3’), 45.11 (C-7), 31.90 (C-8), 26.23 (C-9); **IR** (vmax / cm^{-1}): 3528 (N-H stretch), 3450 (N-H stretch), 3310 (N-H stretch), 3154 (N-H stretch), 2925 (C-H stretch), 2865 (C-H stretch), 2212 (C≡C stretch), 1645 (C=N stretch), 1591 (C=C stretch), 1548 (C=C stretch); **HRMS(ESI)**: *m/z* calculated for $C_{18}H_{20}ClN_4O^+$: 343.1320 $[M + H]^+$; found: 343.1292 (8.16 ppm error); **r.f.** 0.25 (70 % EtOAc / Hex).

6.1.6.4 6-Cyclopentyl-5-[3-(4-fluorophenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8d**)

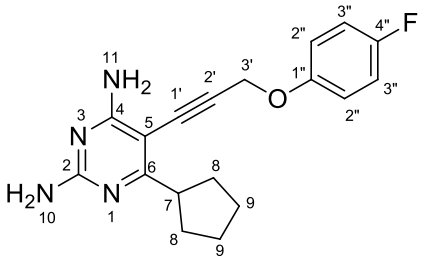


Figure 6.20: 6-Cyclopentyl-5-[3-(4-fluorophenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8d**)

A mixture of **7a** (200 mg, 0.658 mmol, 1 eq), bis(triphenylphosphine)palladium(II) dichloride (46.2 mg, 0.0658 mmol, 0.1 eq), copper iodide (25.1 mg, 0.132 mmol, 0.2 eq), **2d** (198 mg, 1.32 mmol, 2 eq), dimethylformamide (5.00 ml), and diisopropylamine (10.0 ml) yielded **8d** [Fig 6.20] as a yellow polycrystalline material after purification (149 mg, 69.4 %).

m.p. 59 – 62 °C; **^{1}H NMR** (400 MHz, Chloroform-*d*): δ = 7.03-6.92 (H-2" and H-3", m, 4H), 5.27 (H-11, brs, 2H), 5.18 (H-10, brs, 2H), 4.94 (H-3', s, 2H), 3.27 (H-7, q, *J* = 8.3 Hz, 1H), 1.87-1.79 (H-8, m, 2H), 1.78-1.66 (H-8# and H-9, m, 4H), 1.64-1.52 (H-9#, m, 2H); **^{13}C NMR** (101 MHz, Chloroform-*d*): δ = 176.62 (C-6), 164.53 (C-4), 161.44 (C-2), 157.88 (C-4", d, $J_{C\text{-}F}$ = 239.5 Hz), 153.74 (C-1"), 116.57 (C-2", d, $J_{C\text{-}F}$ = 7.9 Hz), 116.02 (C-3", d, $J_{C\text{-}F}$ = 23.2 Hz), 93.17 (C-1'), 89.48 (C-5), 81.62 (C-2'), 57.65 (C-3'), 45.09 (C-7), 31.91 (C-8), 26.24 (C-9); **IR** (ν_{max} / cm^{-1}): 3522 (N-H stretch), 3442 (N-H stretch), 3307 (N-H stretch), 3161 (N-H stretch), 2954 (C-H stretch), 2856 (C-H stretch), 2222 (C≡C stretch), 1640 (C=N stretch), 1606 (C=C stretch), 1545 (C=C stretch); **HRMS(ESI)**: *m/z* calculated for $C_{18}H_{20}FN_4O^+$: 327.1616 $[M + H]^+$; found: 327.1620 (1.22 ppm error); **r.f.** 0.29 (70 % EtOAc / Hex).

6.1.6.5 6-Cyclopentyl-5-[3-(3-fluorophenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8e**)

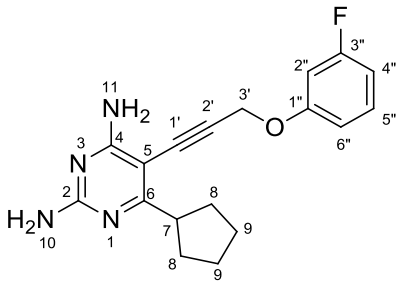


Figure 6.21: 6-Cyclopentyl-5-[3-(3-fluorophenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8e**)

A mixture of **7a** (200 mg, 0.658 mmol, 1 eq), bis(triphenylphosphine)palladium(II) dichloride (46.2 mg, 0.0658 mmol, 0.1 eq), copper iodide (25.1 mg, 0.132 mmol, 0.2 eq), **2e** (198 mg, 1.32 mmol, 2 eq), dimethylformamide (5.00 ml), and diisopropylamine (10.0 ml) yielded **8e** [Fig 6.21] as a yellow polycrystalline material after purification (141 mg, 65.7 %).

m.p. 56 – 59 °C; **^{1}H NMR** (500 MHz, Chloroform-*d*): δ = 7.25 (H-5”, m, 1H), 6.80 (H-6”, dd, *J* = 8.3, 2.0 Hz, 1H), 6.75 (H-2”, dt, *J* = 10.8, 2.2 Hz, 1H), 6.71 (H-4”, td, *J* = 8.3, 2.1 Hz, 1H), 5.15 (H-11, brs, 2H), 4.98 (H-10, brs, 2H), 4.97 (H-3’, s, 2H), 3.31 (H-7, q, *J* = 7.9, 1H), 1.90-1.81 (H-8, m, 2H), 1.81-1.68 (H-8# and H-9, m, 4H), 1.64-1.49 (H-9#, m, 2H); **^{13}C NMR** (126 MHz, Chloroform-*d*): δ = 176.77 (C-6), 164.54 (C-4), 163.69 (C-3”, d, $J_{C\text{-}F}$ = 245.78 Hz), 161.47 (C-2), 158.98 (C-1”, d, $J_{C\text{-}F}$ = 10.7 Hz), 130.39 (C-5”, d, $J_{C\text{-}F}$ = 9.9 Hz), 110.98 (C-6”, d, $J_{C\text{-}F}$ = 2.9 Hz), 108.47 (C-4”, d, $J_{C\text{-}F}$ = 21.3 Hz), 103.09 (C-2”, d, $J_{C\text{-}F}$ = 25.0 Hz), 92.82 (C-1’), 89.39 (C-5), 81.91 (C-2’), 57.17 (C-3’), 45.11 (C-7), 31.92 (C-8), 26.23 (C-9); **IR** (vmax / cm^{-1}): 3593 (N-H stretch), 3490 (N-H stretch), 3365 (N-H stretch), 3151 (N-H stretch), 2948 (C-H stretch), 2872 (C-H stretch), 2216 (C≡C stretch), 1664 (C=N stretch), 1612 (C=C stretch), 1541 (C=C stretch); **HRMS(ESI)**: *m/z* calculated for $C_{18}H_{20}FN_4O^+$: 327.1616 $[M + H]^+$; found: 327.1641 (7.64 ppm error); **r.f.** 0.33 (70 % EtOAc / Hex).

6.1.6.6 6-Cyclopentyl-5-[3-(2-fluorophenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8f**)

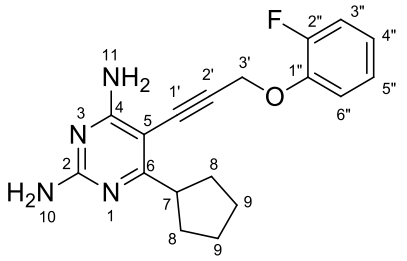


Figure 6.22: 6-Cyclopentyl-5-[3-(2-fluorophenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8f**)

A mixture of **7a** (200 mg, 0.658 mmol, 1 eq), bis(triphenylphosphine)palladium(II) dichloride (46.2 mg, 0.0658 mmol, 0.1 eq), copper iodide (25.1 mg, 0.132 mmol, 0.2 eq), **2f** (198 mg, 1.32 mmol, 2 eq), dimethylformamide (5.00 ml), and diisopropylamine (10.0 ml) **8f** [Fig 6.22] as a yellow polycrystalline material after purification (133 mg, 61.9 %).

m.p. 72 – 74 °C; **^{1}H NMR** (400 MHz, Methanol-d_4): δ = 7.27 (H-5”, t, J = 7.8 Hz, 1H), 7.17-7.07 (H-6” and H-4”, m, 2H), 6.97 (H-3”, m, 1H), 5.10 (H-3’, s, 2H), 3.26 (H-7, q, J = 8.6 Hz, 1H), 1.83-1.72 (H-8 and H-9, m, 4H), 1.71-1.65 (H-8#, m, 2H), 1.61-1.54 (H-9#, m, 2H); **^{13}C NMR** (101 MHz, Methanol-d_4): δ = 176.61 (C-6), 166.03 (C-4), 163.01 (C-2), 154.54 (C-2”, d, $J_{C\text{-}F}$ = 244.6 Hz), 146.84 (C-1”, d, $J_{C\text{-}F}$ = 10.8 Hz), 125.37 (C-4”, d, $J_{C\text{-}F}$ = 2.1 Hz), 123.20 (C-6”, d, $J_{C\text{-}F}$ = 6.7 Hz), 117.70 (C-5”), 117.21 (C-3”, d, $J_{C\text{-}F}$ = 18.5 Hz), 94.10 (C-1’), 89.73 (C-5) , 82.15 (C-2’), 58.81 (C-3’), 46.05 (C-7), 32.62 (C-8), 27.08 (C-9); **IR** (vmax / cm^{-1}): 3591 (N-H stretch), 3488 (N-H stretch), 3358 (N-H stretch), 3150 (N-H stretch), 2947 (C-H stretch), 2868 (C-H stretch), 2216 (C≡C stretch), 1666 (C=N stretch), 1617 (C=C stretch), 1539 (C=C stretch); **HRMS(ESI)**: *m/z* calculated for $C_{18}H_{20}FN_4O^+$: 327.1616 $[M + H]^+$; found: 327.1639 (7.03 ppm error); **r.f.** 0.34 (70 % EtOAc / Hex).

6.1.6.7 6-Cyclopentyl-5-[3-(4-methoxyphenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8g**)

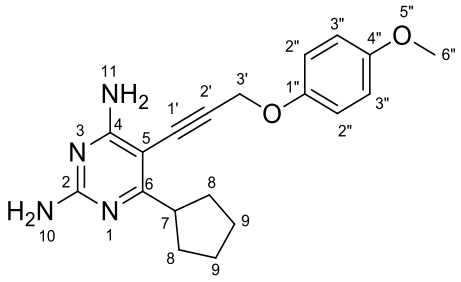


Figure 6.23: 6-Cyclopentyl-5-[3-(4-methoxyphenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8g**)

A mixture of **7a** (200 mg, 0.658 mmol, 1 eq), bis(triphenylphosphine)palladium(II) dichloride (46.2 mg, 0.0658 mmol, 0.1 eq), copper iodide (25.1 mg, 0.132 mmol, 0.2 eq), **2g** (214 mg, 1.32 mmol, 2 eq), dimethylformamide (5.00 ml), and diisopropylamine (10.0 ml) yielded **8g** [Fig 6.23] as a yellow polycrystalline material after purification (176 mg, 79.0 %).

m.p. 57 – 59 °C; **^{1}H NMR** (400 MHz, Chloroform-*d*): δ = 6.95 (H-2", d, *J* = 9.1 Hz, 2H), 6.84 (H-3", d, *J* = 9.1 Hz, 2H), 5.35 (H-11, brs, 2H), 5.27 (H-10, brs, 2H), 4.91 (H-3', s, 2H), 3.75 (H-6", s, 3H), 3.28 (H-7, q, *J* = 8.2 Hz, 1H), 1.90-1.78 (H-8, m, 2H), 1.79-1.65 (H-8# and H-9, m, 4H), 1.67-1.51 (H-9#, m, 2H); **^{13}C NMR** (101 MHz, Chloroform-*d*): δ = 176.51 (C-6), 164.53 (C-4), 161.41 (C-2), 154.64 (C-4"), 151.71 (C-1"), 116.62 (C-2"), 114.77 (C-3"), 93.65 (C-1'), 89.67 (C-5), 81.29 (C-2'), 57.78 (C-3'), 55.84 (C-6"), 45.07 (C-7), 31.93 (C-8), 26.26 (C-9); **IR** (ν_{max} / cm^{-1}): 3590 (N-H stretch), 3453 (N-H stretch), 3326 (N-H stretch), 3177 (N-H stretch), 2957 (C-H stretch), 2876 (C-H stretch), 2215 (C≡C stretch), 1653 (C=N stretch), 1619 (C=C stretch), 1541 (C=C stretch), 1200 (C-O stretch); **HRMS(ESI)**: *m/z* calculated for $C_{19}H_{23}N_4O_2^+$: 339.1816 $[M + H]^+$; found: 339.1836 (5.90 ppm error); **r.f.** 0.25 (70 % EtOAc / Hex).

6.1.6.8 6-Cyclopentyl-5-[3-(3-methoxyphenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8h**)

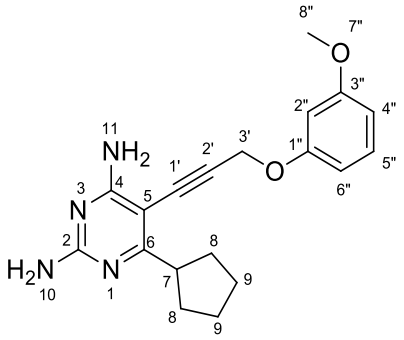


Figure 6.24: 6-Cyclopentyl-5-[3-(3-methoxyphenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8h**)

A mixture of **7a** (200 mg, 0.658 mmol, 1 eq), bis(triphenylphosphine)palladium(II) dichloride (46.2 mg, 0.0658 mmol, 0.1 eq), copper iodide (25.1 mg, 0.132 mmol, 0.2 eq), **2h** (214 mg, 1.32 mmol, 2 eq), dimethylformamide (5.00 ml), and diisopropylamine (10.0 ml) yielded **8h** [Fig 6.24] as a yellow polycrystalline material after purification (159 mg, 71.4 %).

m.p. 148 – 150 °C; **^{1}H NMR** (400 MHz, Chloroform-*d*): δ = 7.21 (H-5”, t, *J* = 8.2 Hz, 1H), 6.62 (H-6”, ddd, *J* = 8.2, 2.4, 0.8 Hz, 1H), 6.58 (H-2”, t, *J* = 2.3 Hz, 1H), 6.55 (H-4”, dd, *J* = 8.2, 0.8 Hz, 1H), 5.13 (H-11, brs, 2H), 4.96 (H-3’, s, 2H) 4.95 (H-10, brs, 2H), 3.79 (H-8”, s, 3H), 3.32 (H-7, q, *J* = 8.3 Hz, 1H), 1.91-1.81 (H-8, m, 2H), 1.80-1.68 (H-8# and H-9, m, 4H), 1.65-1.55 (H-9#, m, 2H); **^{13}C NMR** (101 MHz, Chloroform-*d*): δ = 176.54 (C-6), 164.54 (C-4), 161.38 (C-2), 160.99 (C-3”), 158.85 (C-1”), 130.03 (C-5”), 107.24 (C-6”), 107.13 (C-4”), 101.91 (C-2”), 93.31 (C-1’), 89.61 (C-5), 81.45 (C-2’), 56.95 (C-3’), 55.43 (C-8”), 45.06 (C-7), 31.93 (C-8), 26.23 (C-9); **IR** (ν_{max} / cm^{-1}): 3487 (N-H stretch), 3392 (N-H stretch), 3312 (N-H stretch), 3162 (N-H stretch), 2942 (C-H stretch), 2866 (C-H stretch), 2197 (C≡C stretch), 1643 (C=N stretch), 1613 (C=C stretch), 1541 (C=C stretch) , 1149 (C-O stretch); **HRMS(ESI)**: *m/z* calculated for $C_{19}H_{23}N_4O_2^+$: 339.1816 [M + H]$^+$; found: 339.1840 (7.08 ppm error); **r.f.** 0.24 (70 % EtOAc / Hex).

6.1.6.9 6-Cyclopentyl-5-[3-(2-methoxyphenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8i**)

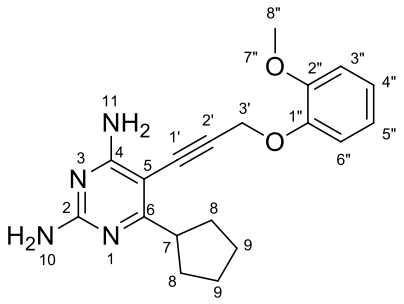


Figure 6.25: 6-Cyclopentyl-5-[3-(2-methoxyphenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8i**)

A mixture of **7a** (200 mg, 0.658 mmol, 1 eq), bis(triphenylphosphine)palladium(II) dichloride (46.2 mg, 0.0658 mmol, 0.1 eq), copper iodide (25.1 mg, 0.132 mmol, 0.2 eq), **2i** (214 mg, 1.32 mmol, 2 eq), dimethylformamide (5.00 ml), and diisopropylamine (10.0 ml) yielded **8i** [Fig 6.25] as a yellow polycrystalline material after purification (162 mg, 72.8 %).

m.p. 72 – 75 °C; **^{1}H NMR** (400 MHz, Chloroform-*d*): δ = 7.12-7.03 (H-4”, m, 1H), 6.99-6.93 (H-6”, m, 1H), 6.92-6.85 (H-5” and H-3”, m, 2H), 5.27 (H-11, brs, 2H), 5.19 (H-10, brs, 2H), 5.05 (H -3’, s, 2H), 3.86 (H-8”, s, 3H), 3.28 (H-7, q, *J* = 8.3 Hz, 1H), 1.87-1.77 (H-8, m, 2H), 1.78-1.65 (H-8# and H-9, m, 4H), 1.62-1.50 (H-9#, m, 2H); **^{13}C NMR** (101 MHz, Chloroform-*d*): δ = 176.53 (C-6), 164.53 (C-4), 161.40 (C-2), 150.11 (C-2”), 146.82 (C-1”), 122.42 (C-6”), 120.71 (C-3”), 114.92 (C-4”), 112.12 (C-5”), 93.45 (C-1’), 89.69 (C-5), 81.64 (C-2’), 57.90 (C-3’), 55.99 (C-8”), 45.06 (C-7), 31.92 (C-8), 26.24 (C-9); **IR** (ν_{max} / cm^{-1}): 3488 (N-H stretch), 3439 (N-H stretch), 3360 (N-H stretch), 3147 (N-H stretch), 2949 (C-H stretch), 2870 (C-H stretch), 2211 (C≡C stretch), 1661 (C=N stretch), 1618 (C=C stretch), 1541 (C=C stretch), 1246 (C-O stretch); **HRMS(ESI)**: *m/z* calculated for $C_{19}H_{23}N_4O_2^+$: 339.1816 $[M + H]^+$; found: 339.1848 (9.43 ppm error); **r.f.** 0.22 (70 % EtOAc / Hex).

6.1.6.10 6-Cyclopropyl-5-[3-(4-chlorophenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8j**)

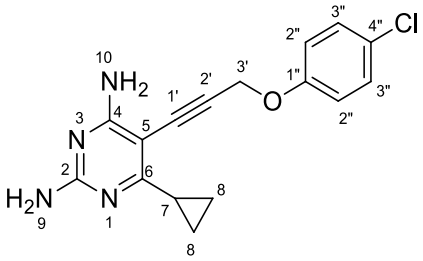


Figure 6.26: 6-Cyclopropyl-5-[3-(4-chlorophenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8j**)

A mixture of **7b** (300 mg, 1.09 mmol, 1 eq), bis(triphenylphosphine)palladium(II) dichloride (76.5 mg, 0.109 mmol, 0.1 eq), copper iodide (41.5 mg, 0.218 mmol, 0.2 eq), **2a** (363 mg, 2.18 mmol, 2 eq), dimethylformamide (5.00 ml), and diisopropylamine (10.0 ml) yielded **8j** [Fig 6.26] as a yellow polycrystalline material after purification (190 mg, 55.3 %).

m.p. 166 – 168 °C; **^{1}H NMR** (400 MHz, DMSO-d_6): δ = 7.34 (H-2”, d, *J* = 8.5 Hz, 2H), 7.10 (H-3”, d, *J* = 8.5 Hz, 2H), 6.42 (H-10, brs, 2H), 6.16 (H-9, brs, 2H), 5.06 (H-3’, s, 2H), 2.18 – 2.05 (H-7, m, 1H), 0.90 – 0.82 (H-8, m, 2H), 0.82 – 0.74 (H-8#, m, 2H); **^{13}C NMR** (101 MHz, DMSO-d_6): δ = 171.71 (C-6), 164.16 (C-4), 161.70 (C-2), 156.14 (C-1”), 129.09 (C-3”), 124.71 (C-4”), 116.93 (C-2”), 92.51 (C-1’), 86.89 (C-5), 82.04 (C-2’), 57.01 (C-3’), 14.37 (C-7), 8.94 (C-8); **IR** (vmax / cm^{-1}): 3477 (N-H stretch), 3372 (N-H stretch), 3326 (N-H stretch), 3171 (N-H stretch), 3012 (C-H stretch), 2924 (C-H stretch), 2217 (C≡C stretch), 1600 (C=C stretch), 1546 (C=C stretch); **HRMS(ESI)**: *m/z* calculated for $C_{16}H_{16}ClN_4O^+$: 315.1007 $[M + H]^+$; found: 315.1011 (1.27 ppm error); **r.f.** 0.41 (70 % EtOAc / Hex).

6.1.6.11 6-Cyclopropyl-5-[3-(3-chlorophenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8k**)

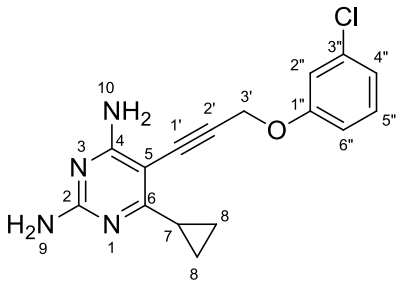


Figure 6.27: 6-Cyclopropyl-5-[3-(3-chlorophenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8k**)

A mixture of **7b** (300 mg, 1.09 mmol, 1 eq), bis(triphenylphosphine)palladium(II) dichloride (76.5 mg, 0.109 mmol, 0.1 eq), copper iodide (41.5 mg, 0.218 mmol, 0.2 eq), **2b** (363 mg, 2.18 mmol, 2 eq), dimethylformamide (5.00 ml), and diisopropylamine (10.0 ml) yielded **8k** [Fig 6.27] as a yellow polycrystalline material after purification (230 mg, 67.0 %).

m.p. 122 – 125 °C; **^{1}H NMR** (400 MHz, DMSO-d6): δ = 7.32 (H-5”, t, *J* = 8.1 Hz, 1H), 7.19 (H-2”, s, 1H), 7.07 – 6.97 (H-4” and H-6”, m, 2H), 6.43 (H-10, brs, 1H), 6.16 (H-9, brs, 2H), 5.10 (H-3’, s, 2H), 2.20 – 2.09 (H-7, m, 1H), 0.90 – 0.84 (H-8, m, 2H), 0.83 – 0.75 (H-8#, m, 2H); **^{13}C NMR** (101 MHz, DMSO-d6): δ = 171.79 (C-6), 164.21 (C-4), 161.74 (C-2), 158.20 (C-1”), 133.63 (C-3”), 130.77 (C-5”), 120.90 (C-4”), 115.01 (C-2”), 114.46 (C-6”), 92.47 (C-1’), 86.88 (C-5), 82.28 (C-2’), 57.02 (C-3’), 14.41 (C-7), 8.95 (C-8); **IR** (vmax / cm^{-1}): 3547 (N-H stretch), 3479 (N-H stretch), 3353 (N-H stretch), 3099 (C-H stretch), 3016 (C-H stretch), 2921 (C-H stretch), 2222 (C≡C stretch), 1657 (C=N stretch), 1545 (C=C stretch); **HRMS(ESI)**: *m/z* calculated for $C_{16}H_{16}ClN_4O^+$: 315.1007 $[M + H]^+$; found: 315.0996 (3.49 ppm error); **r.f.** 0.45 (70 % EtOAc / Hex).

6.1.6.12 6-Cyclopropyl-5-[3-(4-flourophenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8l**)

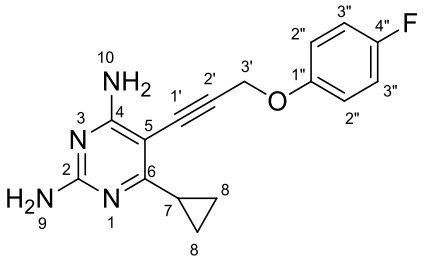


Figure 6.28: 6-Cyclopropyl-5-[3-(4-flourophenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8l**)

A mixture of **7b** (300 mg, 1.09 mmol, 1 eq), bis(triphenylphosphine)palladium(II) dichloride (76.5 mg, 0.109 mmol, 0.1 eq), copper iodide (41.5 mg, 0.218 mmol, 0.2 eq), **2d** (327 mg, 2.18 mmol, 2 eq), dimethylformamide (5.00 ml), and diisopropylamine (10.0 ml) **8l** [Fig 5.68] as a yellow polycrystalline material after purification (210 mg, 64.6 %).

m.p. 115 – 118 °C; **^{1}H NMR** (300 MHz, DMSO-d_6): δ = 7.17 – 7.04 (H-2" and H-3", m, 4H), 6.39 (H-10, brs, 2H), 6.14 (H-9, brs, 2H), 5.04 (H-3', s, 2H), 2.13 (H-7, tt, J = 8.1, 4.8 Hz, 1H), 0.90 – 0.82 (H-8, m, 2H), 0.84 – 0.74 (H-8#, m, 2H); **^{13}C NMR** (75 MHz, DMSO-d_6): δ = 171.73 (C-6), 164.19 (C-4), 161.73 (C-2), 156.78 (C-4", d, $J_{C\text{-}F}$ = 236.3 Hz), 153.61 (C-1", d, $J_{C\text{-}F}$ = 1.9 Hz), 116.48 (C-2", d, $J_{C\text{-}F}$ = 8.0 Hz), 115.74 (C-3", d, $J_{C\text{-}F}$ = 23.0 Hz), 92.78 (C-1'), 86.99 (C-5), 81.85 (C-2'), 57.20 (C-3'), 14.43 (C-7), 9.01 (C-8); **IR** (vmax / cm^{-1}): 3505 (N-H stretch), 3399 (N-H stretch), 3314 (N-H stretch), 3169 (N-H stretch), 3004 (C-H stretch), 2919 (C-H stretch), 2210 (C≡C stretch), 1660 (C=N stretch), 1538 (C=C stretch); **HRMS(ESI)**: *m/z* calculated for $C_{16}H_{16}FN_4O^+$: 299.1303 $[M + H]^+$; found: 299.1317 (4.68 ppm error); **r.f.** 0.38 (70 % EtOAc / Hex).

6.1.6.13 6-Cyclopropyl-5-[3-(3-flourophenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8m**)

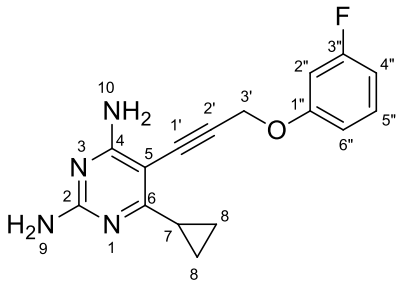


Figure 6.29: 6-Cyclopropyl-5-[3-(3-flourophenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8m**)

A mixture of **7b** (300 mg, 1.09 mmol, 1 eq), bis(triphenylphosphine)palladium(II) dichloride (76.5 mg, 0.109 mmol, 0.1 eq), copper iodide (41.5 mg, 0.218 mmol, 0.2 eq), **2e** (327 mg, 2.18 mmol, 2 eq), dimethylformamide (5.00 ml), and diisopropylamine (10.0 ml) yielded **8m** [Fig 6.29] as a yellow polycrystalline material after purification (221 mg, 68.0 %).

m.p. 97 – 101 °C; **^{1}H NMR** (400 MHz, DMSO-d_6): δ = 7.37 – 7.28 (H-5”, m, 1H), 6.99 (H-6”, d, J = 11.4 Hz, 1H), 6.90 (H-2”, d, J = 8.3 Hz, 1H), 6.78 (H-4”, t, J = 8.3 Hz, 1H), 6.43 (H-10, brs, 2H), 6.16 (H-9, brs, 2H), 5.08 (H-3’, s, 2H), 2.20 – 2.11 (H-7, m, 1H), 0.90 – 0.84 (H-8, m, 2H), 0.82 – 0.75 (H-8#, m, 2H); **^{13}C NMR** (101 MHz, DMSO-d_6): δ = 171.79 (C-6), 164.26 (C-4), 162.95 (C-3”, d, $J_{C\text{-}F}$ = 242.5 Hz), 161.77 (C-2), 158.80 (C-1”, d, $J_{C\text{-}F}$ = 11.1 Hz), 130.61 (C-5”, d, $J_{C\text{-}F}$ = 10.1 Hz), 111.56 (C-6”d, $J_{C\text{-}F}$ = 2.5 Hz), 107.56 (C-4”, d, $J_{C\text{-}F}$ = 21.1 Hz), 102.61 (C-2”, d, $J_{C\text{-}F}$ = 25.1 Hz), 92.51 (C-1’), 86.89 (C-5), 82.17 (C-2’), 57.09 (C-3’), 14.45 (C-7), 8.98 (C-8); **IR** (vmax / cm^{-1}): 3467 (N-H stretch), 3396 (N-H stretch), 3306 (N-H stretch), 3154 (N-H stretch), 3001 (C-H stretch), 2920 (C-H stretch), 2201 (C≡C stretch), 1664 (C=N stretch), 1545 (C=C stretch); **HRMS(ESI)**: *m/z* calculated for $C_{16}H_{16}FN_4O^+$: 299.1303 $[M + H]^+$; found: 299.1315 (4.01ppm error); **r.f.** 0.38 (70 % EtOAc / Hex).

6.1.6.14 6-Cyclopropyl-5-[3-(4-methoxyphenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8n**)

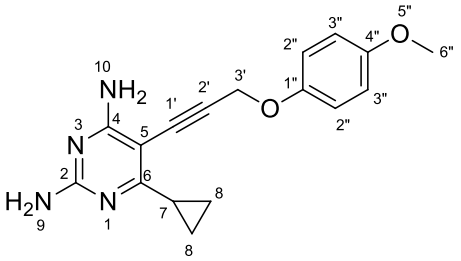


Figure 6.30: 6-Cyclopropyl-5-[3-(4-methoxyphenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8n**)

A mixture of **7b** (300 mg, 1.09 mmol, 1 eq), bis(triphenylphosphine)palladium(II) dichloride (76.5 mg, 0.109 mmol, 0.1 eq), copper iodide (41.5 mg, 0.218 mmol, 0.2 eq), **2g** (354 mg, 2.18 mmol, 2 eq), dimethylformamide (5.00 ml), and diisopropylamine (10.0 ml) yielded **8n** [Fig 6.30] as a yellow polycrystalline material after purification (237 mg, 70.1 %).

m.p. 148 – 150 °C; **^{1}H NMR** (400 MHz, DMSO-d_6): δ = 6.99 (H-2”, d, *J* = 8.4 Hz, 2H), 6.85 (H-3”, d, *J* = 8.4 Hz, 2H), 6.37 (H-10, brs, 2H), 6.13 (H-9, brs, 2H), 4.98 (H-3’, s, 2H), 3.68 (H-6”, s, 3H), 2.29 – 1.94 (H-7, m, 1H), 0.92 – 0.81 (H-8, m, 2H), 0.83 – 0.72 (H-8#, m, 2H); **^{13}C NMR** (101 MHz, DMSO-d_6): δ = 171.75 (C-6), 164.18 (C-4), 161.74 (C-2), 153.78 (C-4”), 151.30 (C-1”), 116.23 (C-2”), 114.53 (C-3”), 93.25 (C-1’), 87.20 (C-5), 82.04 (C-2’), 57.01 (C-3’), 55.43 (C-6”), 14.49 (C-7), 9.10 (C-8); **IR** (vmax / cm^{-1}): 3465 (N-H stretch), 3416 (N-H stretch), 3341 (N-H stretch), 3173 (N-H stretch), 3010 (C-H stretch), 2934 (C-H stretch), 2209 (C≡C stretch), 1650 (C=N stretch), 1534 (C=C stretch), 1504 (C=C stretch), 1223 (C-O stretch); **HRMS(ESI)**: *m/z* calculated for $C_{17}H_{19}N_4O_2^+$: 311.1503 $[M + H]^+$; found: 311.1519 (5.14 ppm error); **r.f.** 0.42 (70 % EtOAc / Hex).

6.1.6.15 6-Cyclopropyl-5-[3-(3-methoxyphenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8o**)

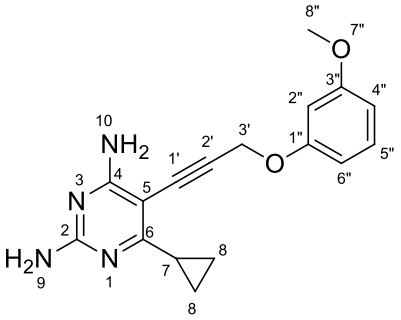
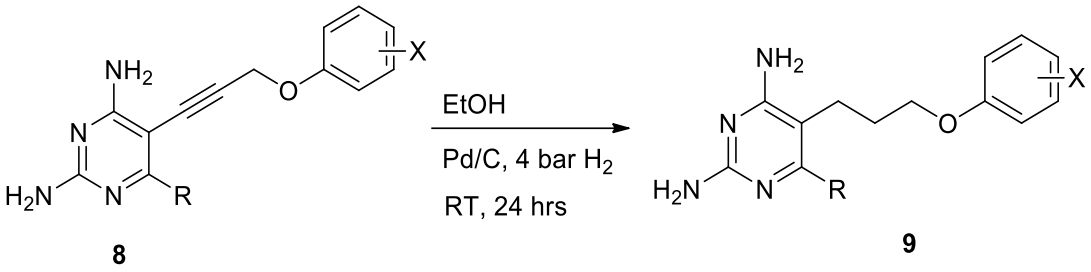


Figure 6.31: 6-Cyclopropyl-5-[3-(3-methoxyphenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**8o**)

A mixture of **7b** (300 mg, 1.09 mmol, 1 eq), bis(triphenylphosphine)palladium(II) dichloride (76.5 mg, 0.109 mmol, 0.1 eq), copper iodide (41.5 mg, 0.218 mmol, 0.2 eq), **2h** (354 mg, 2.18 mmol, 2 eq), dimethylformamide (5.00 ml), and diisopropylamine (10.0 ml) yielded **8o** [Fig 6.31] as a yellow polycrystalline material after purification (249 mg, 73.6 %).

m.p. 115 – 118 °C; **^{1}H NMR** (300 MHz, DMSO-d_6): δ = 7.19 (H-5”, t, *J* = 8.5 Hz, 1H), 6.69 – 6.58 (H-2” and H-6”, m, 2H), 6.53 (H-4”, dd, *J* = 8.2, 1.4 Hz, 1H), 6.38 (H-10, brs, 2H), 6.13 (H-9, brs, 2H), 5.04 (H-3’, s, 2H), 3.72 (H-8”, s, 3H), 2.17 (H-7, tt, *J* = 8.1, 4.4 Hz, 1H), 0.94 – 0.84 (H-8, m, 2H), 0.83 – 0.70 (H-8#, m, 2H); **^{13}C NMR** (75 MHz, DMSO-d_6): δ = 171.67 (C-6), 164.19 (C-4), 161.72 (C-2), 160.45 (C-3”), 158.55 (C-1”), 129.85 (C-5”), 107.38 (C-6”), 106.54 (C-4”), 101.30 (C-2”), 92.94 (C-1’), 87.03 (C-5), 81.75 (C-2’), 56.66 (C-3’), 55.11 (C-8”), 14.44 (C-7), 8.97 (C-8); **IR** (vmax / cm^{-1}): 3510 (N-H stretch), 3469 (N-H stretch), 3329 (N-H stretch), 3127 (N-H stretch), 3020 (C-H stretch), 2962 (C-H stretch), 2221 (C≡C stretch), 1653 (C=N stretch), 1543 (C=C stretch), 1146 (C-O stretch); **HRMS(ESI)**: *m/z* calculated for $C_{17}H_{19}N_4O_2^+$: 311.1503 $[M + H]^+$; found: 311.1517 (4.50 ppm error); **r.f.** 0.37 (70 % EtOAc / Hex).

6.1.7 Hydrogenation of alkynes



Scheme 6.7: Hydrogenation of alkynes

The relevant unsaturated starting material (**8a – o**) (50 - 300 mg, 1 eq) was dissolved in EtOH (5.00 – 20.0 ml) and added to a high-pressure reaction vessel followed by the addition of Pd/C (10 % m/m). The mixture was then stirred under H_2 at 4 bar of pressure for 24 hours. After this the mixture was filtered through celite and the solvent evaporated under reduced pressure. Purification was achieved using silica gel column chromatography (10% MeOH / DCM) if required. The reaction information is summarized in scheme 6.7, and the full list of products formed from this reaction is displayed in table 6.2.

Table 6.2: Full list of products from hydrogenation reaction

Product	**R**	**X**
9a	Cpe	4-Cl
9b	Cpe	2-Cl
9c	Cpe	4-F
9d	Cpe	3-F
9e	Cpe	2-F
9f	Cpe	4-OMe
9g	Cpe	3-OMe
9h	Cpe	2-OMe
9i	Cpr	4-Cl

Product	R	X
9j	Cpr	4-F
9k	Cpr	3-F
9l	Cpr	4-OMe
9m	Cpr	2-OMe

6.1.7.1 6-Cyclopentyl-5-[3-(4-chlorophenoxy)propyl] pyrimidine-2,4-diamine (**9a**)

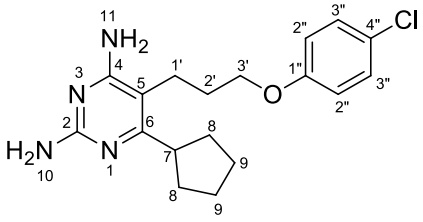


Figure 6.32: 6-Cyclopentyl-5-[3-(4-chlorophenoxy)propyl] pyrimidine-2,4-diamine (**9a**)

Hydrogenation of **8a** (50.0 mg, 0.146 mmol) with palladium on carbon (5.00 mg, 10 % m/m) in ethanol (5.00 ml) afforded **9a** [Fig 6.32] as a white crystalline solid (33.3 mg, 66.7 %).

m.p. 137 – 140 °C; **^{1}H NMR** (400 MHz, Methanol-d_4): δ = 7.24 (H-3", d, *J* = 8.8 Hz, 2H), 6.90 (H-2", d, *J* = 8.7 Hz, 2H), 3.98 (H-3', t, *J* = 5.7 Hz, 2H), 3.17 (H-7, q, *J* = 8.5 Hz, 1H), 2.62 (H-1', t, *J* = 7.4 Hz, 2H), 1.91 (H-2', q, J = 6.3 Hz, 2H), 1.85 – 1.62 (H-8 and H-9 and H-8#, m, 6H), 1.62 – 1.47 (H-9#, m, 2H); **^{13}C NMR** (101 MHz, Methanol-d_4) δ 170.52 (C-6), 164.48 (C-4), 162.53 (C-2), 158.98 (C-1"), 130.34 (C-3"), 126.55 (C-4"), 116.99 (C-2"), 105.22 (C-5), 67.92 (C-3'), 43.15 (C-7), 33.52 (C-8), 29.65 (C-2'), 26.97 (C-9), 22.07 (C-1'); **IR** (vmax / cm^{-1}): 3457 (N-H stretch), 3326 (N-H stretch), 3150 (N-H stretch), 2956 (C-H stretch), 2868 (C-H stretch), 1622 (C=N stretch), 1560 (C=C stretch); **HRMS(ESI)**: *m/z* calculated for $C_{18}H_{24}ClN_4O^+$: 347.1633 $[M + H]^+$; found: 347.1642 (2.59 pm error); **r.f.** 0.18 (70 % EtOAc / Hex).

6.1.7.2 6-Cyclopentyl-5-[3-(2-chlorophenoxy)propyl] pyrimidine-2,4-diamine (**9b**)

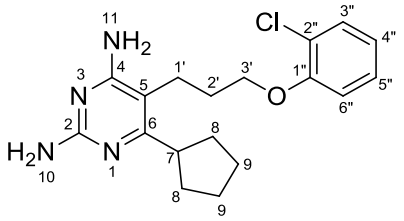


Figure 6.33: 6-Cyclopentyl-5-[3-(2-chlorophenoxy)propyl] pyrimidine-2,4-diamine (**9b**)

Hydrogenation of **8c** (50.0 mg, 0.146 mmol) with palladium on carbon (5.00 mg, 10 % m/m) in ethanol (5.00 ml) **9b** [Fig 6.33] as a white crystalline solid (12.0 mg, 23.4 %).

m.p. 55 – 58 °C; **^{1}H NMR** (400 MHz, Methanol-d_4): δ = 7.37 (H-3”, d, *J* = 7.8 Hz, 1H), 7.25 (H-5”, t, *J* = 7.8 Hz, 1H), 7.04 (H-6”, d, *J* = 8.2 Hz, 1H), 6.91 (H-4”, t, *J* = 7.6 Hz, 1H), 4.07 (H-3’, t, *J* = 5.4 Hz, 2H), 3.31 – 3.20 (H-7, m, 1H), 2.70 (H-1’, t, *J* = 7.5 Hz, 2H), 1.95 (H-2’, q, *J* = 8.4 Hz, 2H), 1.79 – 1.66 (H-8 and H-9 and H-8#, m, 6H), 1.61 – 1.52 (H-9#, m, 2H); **^{13}C NMR** (101 MHz, Methanol-d_4): δ = 169.18 (C-6), 163.10 (C-4), 161.14 (C-2), 154.20 (C-1”), 129.74 (C-3”), 127.71 (C-5”), 122.24 (C-2”), 121.00 (C-4”), 112.98 (C-6”), 103.83 (C-5), 66.93 (C-3’), 41.65 (C-7), 32.11 (C-8), 28.25 (C-2’), 25.56 (C-9), 20.69 (C-1’); **IR** (vmax / cm^{-1}): 3487 (NH_2 stretch), 3350 (NH_2 stretch), 3129 (NH_2 stretch), 2955 (C-H stretch), 2864 (C-H stretch), 1632 (C=N stretch), 1599 (C=C stretch), 1562 (C=C stretch); **HRMS(ESI)**: *m/z* calculated for $C_{18}H_{24}ClN_4O^+$: 347.1633 $[M + H]^+$; found: 347.1623 (2.88 pm error); **r.f.** 0.21 (70 % EtOAc / Hex).

6.1.7.3 6-Cyclopentyl-5-[3-(4-fluorophenoxy)propyl] pyrimidine-2,4-diamine (**9c**)

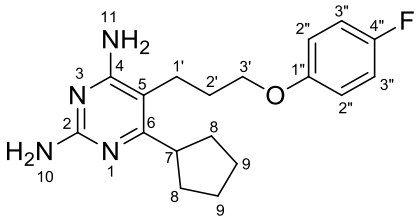


Figure 6.34: 6-Cyclopentyl-5-[3-(4-fluorophenoxy)propyl] pyrimidine-2,4-diamine (**9c**)

Hydrogenation of **8d** (100 mg, 0.306 mmol) with palladium on carbon (10.0 mg, 10 % m/m) in ethanol (10.0 ml) afforded **8c** [Fig 6.34] as a white crystalline solid (44.8 mg, 44.3 %).

m.p. 56 – 58 °C; **^{1}H NMR** (400 MHz, Methanol-*d*$_4$): δ = 6.99 (H-3", t, *J* = 8.7 Hz, 2H), 6.90 (H-2", dd, *J* = 9.2, 4.3 Hz, 2H), 3.95 (H-3', t, *J* = 5.8 Hz, 2H), 3.19 (H-7, q, *J* = 8.6 Hz, 1H), 2.62 (H-1', t, *J* = 7.4 Hz, 2H), 1.90 (H-2', q, *J* = 6.2 Hz, 2H), 1.84 – 1.64 (H-8 and H-9 and H-8#, m, 6H), 1.61 – 1.49 (H-9#, m, 2H); **^{13}C NMR** (101 MHz, Methanol-*d*$_4$): δ = 169.93 (C-6), 164.57 (C-4), 162.28 (C-2), 158.63 (C-4", d, $J_{C\text{-}F}$ = 236.7 Hz), 156.44 (C-1"), 116.71 (C-3", d, $J_{C\text{-}F}$ = 17.3 Hz), 116.56 (C-2"), 105.33 (C-5), 68.08 (C-3'), 43.05 (C-7), 33.49 (C-8), 29.71 (C-2'), 26.96 (C-9), 22.05 (C-1'); **IR** (vmax / cm^{-1}): 3478 (N-H stretch), 3383 (N-H stretch), 3154 (N-H stretch), 2953 (C-H stretch), 2870 (C-H stretch), 1684 (C=N stretch), 1619 (C=C stretch), 1561 (C=C stretch); **HRMS(ESI)**: *m/z* calculated for $C_{18}H_{24}FN_4O^+$: 331.1929 $[M + H]^+$; found: 331.1927 (0.604 pm error); **r.f.** 0.24 (70 % EtOAc / Hex).

6.1.7.4 6-Cyclopentyl-5-[3-(3-fluorophenoxy)propyl] pyrimidine-2,4-diamine (**9d**)

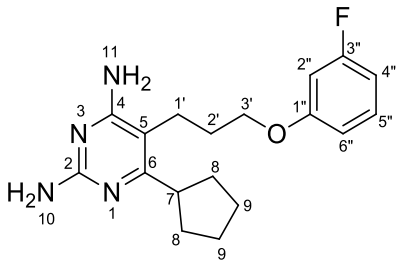


Figure 6.35: 6-Cyclopentyl-5-[3-(3-fluorophenoxy)propyl] pyrimidine-2,4-diamine (**9d**)

Hydrogenation of **8e** (100 mg, 0.306 mmol) with palladium on carbon (10.0 mg, 10 % m/m) in ethanol (10.0 ml) afforded **9d** [Fig 6.35] as a white crystalline solid (51.0 mg, 50.4 %).

m.p. 188 – 190 °C; **^{1}H NMR** (400 MHz, Methanol-*d*$_4$): δ = 7.26 (H-5”, q, *J* = 7.9 Hz, 1H), 6.75 (H-6”, d, *J* = 8.4 Hz, 1H), 6.72 – 6.61 (H-2” and H-4”, m, 2H), 4.00 (H-3’, t, *J* = 5.7 Hz, 2H), 3.22 (H-7, q, *J* = 8.9 Hz, 1H), 2.64 (H-1’, t, *J* = 7.4 Hz, 2H), 1.93 (H-2’, q, *J* = 6.4 Hz, 2H), 1.85 – 1.73 (H-8 and H-9, m, 4H), 1.72 – 1.62 (H-8#, m, 2H), 1.61 – 1.53 (H-9#, m, 2H).**^{13}C NMR** (Chloroform-*d*, 101 MHz): δ = 170.43 (C-6), 163.77 (C-3”, d, $J_{C\text{-}F}$ = 245.5 Hz), 162.84 (C-4), 160.78 (C-2), 159.99 (C-1”, d, $J_{C\text{-}F}$ = 10.8 Hz), 130.49 (C-5”, d, $J_{C\text{-}F}$ = 10.1 Hz), 110.38 (C-6”, d, $J_{C\text{-}F}$ = 2.8 Hz), 108.01 (C-4”, d, $J_{C\text{-}F}$ = 21.3 Hz), 104.12 (C-5), 102.33 (C-2”, d, $J_{C\text{-}F}$ = 24.9 Hz), 66.73 (C-3’), 42.19 (C-7), 32.79 (C-8), 28.90 (C-2’), 26.18 (C-9), 21.40 (C-1’); **IR** (vmax / cm^{-1}): 3478 (N-H stretch), 3384 (N-H stretch), 3163 (N-H stretch), 2954 (C-H stretch), 2872 (C-H stretch), 1682 (C=N stretch), 1615 (C=C stretch), 1562 (C=C stretch); **HRMS(ESI)**: *m/z* calculated for $C_{18}H_{24}FN_4O^+$: 331.1929 [M + H]$^+$; found: 331.1933 (1.21 pm error); **r.f.** 0.27 (70 % EtOAc / Hex).

6.1.7.5 6-Cyclopentyl-5-[3-(2-fluorophenoxy)propyl] pyrimidine-2,4-diamine (**9e**)

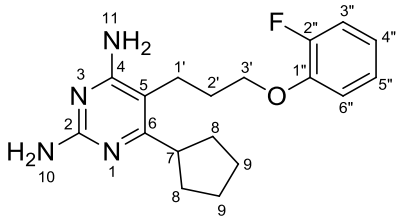


Figure 6.36: 6-Cyclopentyl-5-[3-(2-fluorophenoxy)propyl] pyrimidine-2,4-diamine (**9e**)

Hydrogenation of **8f** (50.0 mg, 0.153 mmol) with palladium on carbon (5.00 mg, 10 % m/m) in ethanol (5.00 ml) afforded **9e** [Fig 6.36] as a white crystalline solid (18.0 mg, 35.6 %).

m.p. 56 – 58 °C; **^{1}H NMR** (400 MHz, Methanol-*d*$_4$): δ = 7.14 – 7.03 (H-4" and H-5" and H-6", m, 3H), 6.96 – 6.86 (H-3", m, 1H), 4.07 (H-3', t, *J* = 5.7 Hz, 2H), 3.24 (H-7, q, *J* = 8.8, 1H), 2.66 (H-1', t, *J* = 7.5 Hz, 2H), 1.94 (H-2', q, *J* = 6.1 Hz, 2H), 1.83 – 1.74 (H-8 and H-9, m, 4H), 1.72 – 1.64 (H-8#, m, 2H), 1.62 – 1.55 (H-9#, m, 2H); **^{13}C NMR** (101 MHz, Methanol-*d*$_4$): δ = 168.59 (C-6), 164.78 (C-4), 161.70 (C-2), 154.01 (C-2", d, $J_{C\text{-}F}$ = 243.8 Hz), 148.15 (C-1", d, $J_{C\text{-}F}$ = 10.5 Hz), 125.63 (C-4", d, $J_{C\text{-}F}$ = 3.7 Hz), 122.15 (C-6", d, $J_{C\text{-}F}$ = 6.5 Hz), 116.93 (C-3", d, $J_{C\text{-}F}$ = 18.2 Hz), 115.85 (C-5"), 105.48 (C-5), 68.63 (C-3'), 42.80 (C-7), 33.41 (C-8), 29.62 (C-2'), 26.90 (C-9), 21.94 (C-1'); **IR** (vmax / cm^{-1}): 3498 (N-H stretch), 3395 (N-H stretch), 3223 (N-H stretch), 2953 (C-H stretch), 2871 (C-H stretch), 1658 (C=N stretch), 1615 (C=C stretch), 1563 (C=C stretch); **HRMS(ESI)**: *m/z* calculated for $C_{18}H_{24}FN_4O^+$: 331.1929 $[M + H]^+$; found: 331.1899 (9.06 pm error); **r.f.** 0.28 (70 % EtOAc / Hex).

6.1.7.6 6-Cyclopentyl-5-[3-(4-methoxyphenoxy)propyl] pyrimidine-2,4-diamine (**9f**)

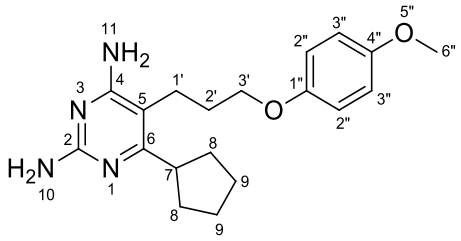


Figure 6.37: 6-Cyclopentyl-5-[3-(4-methoxyphenoxy)propyl] pyrimidine-2,4-diamine (**9f**)

Hydrogenation of **8g** (50.0 mg, 0.148 mmol) with palladium on carbon (5.00 mg, 10 % m/m) in ethanol (5.00 ml) afforded **9f** [Fig 6.37] as a white crystalline solid (38.0 mg, 75.0 %).

m.p. 56 – 58 °C; **^{1}H NMR** (400 MHz, Methanol-*d*$_4$): δ = 6.88 – 6.81 (H-2” and H-3”, m, 4H), 3.95 (H-3’, t, *J* = 5.8 Hz, 2H), 3.74 (H-6”, s, 3H), 3.20 (H-7, q, *J* = 8.5 Hz, 1H), 2.62 (H-1’, t, *J* = 7.4 Hz, 2H), 1.89 (H-2’, q, *J* = 6.0 Hz, 2H), 1.82 – 1.73 (H-8 and H-9, m, 4H), 1.72 – 1.63 (H-8#, m, 2H), 1.61 – 1.53 (H-9#, m, 2H); **^{13}C NMR** (101 MHz, Methanol-*d*$_4$): δ = 169.15 (C-6), 164.74 (C-4), 161.89 (C-2), 155.44 (C-4”), 154.23 (C-1”), 116.44 (C-2”), 115.74 (C-3”), 105.60 (C-5), 67.99 (C-3’), 56.13 (C-6”), 42.91 (C-7), 33.44 (C-8), 29.78 (C-2’), 26.95 (C-9), 22.06 (C-1’); **IR** (νmax / cm^{-1}): 3483 (N-H stretch), 3333 (N-H stretch), 3187 (N-H stretch), 2946 (C-H stretch), 2868 (C-H stretch), 1622 (C=N stretch), 1562 (C=C stretch), 1226 (C-O stretch); **HRMS(ESI)**: *m/z* calculated for $C_{19}H_{27}N_4O_2^+$: 343.2129 $[M + H]^+$; found: 343.2103 (7.58 pm error); **r.f.** 0.20 (70 % EtOAc / Hex).

6.1.7.7 6-Cyclopentyl-5-[3-(3-methoxyphenoxy)propyl] pyrimidine-2,4-diamine (**9g**)

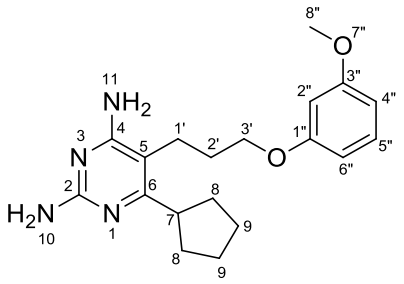


Figure 6.38: 6-Cyclopentyl-5-[3-(3-methoxyphenoxy)propyl] pyrimidine-2,4-diamine (**9g**)

Hydrogenation of **8h** (50.0 mg, 0.148 mmol) with palladium on carbon (5.00 mg, 10 % m/m) in ethanol (5.00 ml) afforded **9g** [Fig 6.38] as a white crystalline solid (28.0 mg, 55.2 %).

m.p. 59 – 63 °C; **^{1}H NMR** (400 MHz, Methanol-*d*$_4$): δ = 7.15 (H-5”, t, *J* = 8.1 Hz, 1H), 6.55 – 6.44 (H-2” and H-4” and H-6”, m, 3H), 3.97 (H-3’, t, *J* = 5.8 Hz, 2H), 3.76 (H-8”, s, 3H), 3.20 (H-7, q, *J* = 8.7 Hz, 1H), 2.63 (H-1’, t, *J* = 7.4 Hz, 5H), 1.90 (H-2’, q, *J* = 6.3 Hz, 2H), 1.81 – 1.71 (H-8 and H-9, m, 4H), 1.71 – 1.62 (H-8#, m, 2H), 1.61 – 1.51 (H-9#, m, 2H); **^{13}C NMR** (101 MHz, Methanol-*d*$_4$): δ = 169.07 (C-6), 164.71 (C-4), 162.42 (C-3”), 161.87 (C-2), 161.40 (C-1”), 130.96 (C-5), 107.74 (C-6”), 107.30 (C-4”), 105.50 (C-5), 101.93 (C-2”), 67.39 (C-3’), 55.67 (C-8”), 42.91 (C-7), 33.44 (C-8), 29.62 (C-2’), 26.92 (C-9), 22.02 (C-1’). **IR** (νmax / cm^{-1}): 3397 (N-H stretch), 3221 (N-H stretch), 2949 (C-H stretch), 2870 (C-H stretch), 1683 (C=N stretch), 1616 (C=C stretch), 1561 (C=C stretch), 1156 (C-O stretch); **HRMS(ESI)**: *m/z* calculated for $C_{19}H_{27}N_4O_2^+$: 343.2129 $[M + H]^+$; found: 343.2117 (3.50 pm error); **r.f.** 0.22 (70 % EtOAc / Hex).

6.1.7.8 6-Cyclopentyl-5-[3-(2-methoxyphenoxy)propyl] pyrimidine-2,4-diamine (**9h**)

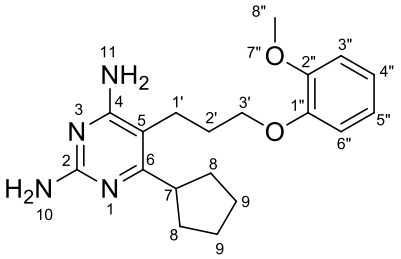


Figure 6.39: 6-Cyclopentyl-5-[3-(2-methoxyphenoxy)propyl] pyrimidine-2,4-diamine (**9h**)

Hydrogenation of **8i** (50.0 mg, 0.148 mmol) with palladium on carbon (5.00 mg, 10 % m/m) in ethanol (5.00 ml) afforded **9h** [Fig 6.39] as a white crystalline solid (31.5 mg, 62.2 %).

m.p. 124 – 127 °C; **^{1}H NMR** (400 MHz, Methanol-d4:) δ = 7.00 – 6.84 (H-3" and H-4" and H-5" and H-6", m, 4H), 4.03 (H-3', t, *J* = 5.5 Hz, 2H), 3.85 (H-8", s, 3H), 3.25 (H-7, q, *J* = 8.7 Hz, 1H), 2.66 (H-1', t, *J* = 7.4 Hz, 2H), 1.96 – 1.88 (H-2', m, 2H), 1.84 – 1.76 (H-8 and H-9, m, 4H), 1.72 – 1.64 (H-8#, m, 2H), 1.63 – 1.56 (H-9#, m, 2H).**^{13}C NMR** (101 MHz, Methanol-d_4): δ = 167.14 (C-6), 165.09 (C-4), 160.98 (C-2) 150.79 (C-2"), 149.51 (C-1"), 122.34 (C-4"), 122.15 (C-5"), 114.28 (C-6"), 113.17 (C-3"), 105.93 (C-5), 68.28 (C-3'), 56.40 (C-8"), 42.54 (C-7), 33.36 (C-8), 29.74 (C-2'), 26.94 (C-9), 22.02 (C-1'); **IR** (vmax / cm^{-1}): 3489 (N-H stretch), 3453 (N-H stretch), 3128 (N-H stretch), 2946 (C-H stretch), 2868 (C-H stretch), 1667 (C=N stretch), 1621 (C=C stretch), 1563 (C=C stretch), 1224 (C-O stretch); **HRMS(ESI)**: *m/z* calculated for $C_{19}H_{27}N_4O_2^+$: 343.2129 $[M + H]^+$; found: 343.2147 (5.25 pm error); **r.f.** 0.17 (70 % EtOAc / Hex).

6.1.7.9 6-Cyclopropyl-5-[3-(4-chlorophenoxy)propyl] pyrimidine-2,4-diamine (**9i**)

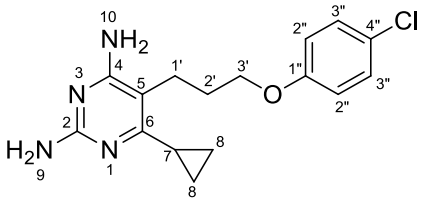


Figure 6.40: 6-Cyclopropyl-5-[3-(4-chlorophenoxy)propyl]pyrimidine-2,4-diamine (**9i**)

Hydrogenation of **8j** (300 mg, 0.953 mmol) with palladium on carbon (30.0 mg, 10 % m/m) in ethanol (20.0 ml) afforded **9i** [Fig 6.40] as a white crystalline solid (250 mg, 85.5 %).

m.p. 137 – 140 °C; **^{1}H NMR** (400 MHz, Chloroform-*d*): δ = 7.21 (H-3”, d, *J* = 9.0 Hz, 2H), 6.81 (H-2”, d, *J* = 8.9 Hz, 2H), 4.88 (H-10, brs, 2H), 4.65 (H-9, brs, 2H), 3.93 (H-3’, t, *J* = 5.6 Hz, 2H), 2.70 (H-1’, t, *J* = 7.1 Hz, 2H), 2.00 (H-2’, q, *J* = 6.6 Hz, 2H), 1.97 – 1.86 (H-7, m, 1H), 1.09 – 0.94 (H-8, m, 2H), 0.87 – 0.70 (H-8#, m, 2H); **^{13}C NMR** (101 MHz, Chloroform-*d*): δ = 168.08 (C-6), 162.43 (C-4), 161.11 (C-2), 157.31 (C-1”), 129.51 (C-3”), 125.94 (C-4”), 115.88 (C-2”), 104.00 (C-5), 66.68 (C-3’), 28.55 (C-2’), 21.17 (C-1’), 12.68 (C-7), 9.00 (C-8); **IR** (ν_{max} / cm^{-1}): 3478 (N-H stretch), 3383 (N-H stretch), 3175 (N-H stretch), 3009 (C-H stretch), 2933 (C-H stretch), 1609 (C=C stretch), 1560 (C=C stretch); **HRMS(ESI)**: *m/z* calculated for $C_{16}H_{20}ClN_4O^+$: 319.1320 [M + H]$^+$; found: 319.1335 (4.70 pm error); **r.f.** 0.35 (70 % EtOAc / Hex).

6.1.7.10 6-Cyclopropyl-5-[3-(4-fluorophenoxy)propyl] pyrimidine-2,4-diamine (**9j**)

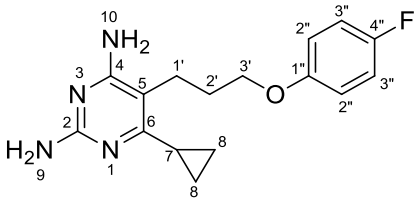


Figure 6.41: 6-Cyclopropyl-5-[3-(4-fluorophenoxy)propyl] pyrimidine-2,4-diamine (**9j**)

Hydrogenation of **8l** (50.0 mg, 0.168 mmol) with palladium on carbon (5.00 mg, 10 % m/m) in ethanol (5.00 ml) afforded **9j** [Fig 6.41] as a white crystalline solid (25.0 mg, 51.3 %).

m.p. 202 – 204 °C; **^{1}H NMR** (300 MHz, Chloroform-*d*): δ = 6.97 (H-3”, t, *J* = 8.6 Hz, 2H), 6.83 (H-2”, dd, *J* = 9.0, 4.3 Hz, 2H), 4.82 (H-10, brs, 2H), 4.55 (H-9, brs, 2H), 3.93 (H-3’, t, *J* = 5.5 Hz, 2H), 2.72 (H-1’, t, *J* = 7.0 Hz, 2H), 2.01 (H-2’, q, *J* = 6.5 Hz, 2H), 1.95 – 1.88 (H-7, m, 1H), 1.07 – 0.96 (H-8, m, 2H), 0.86 – 0.67 (H-8#, m, 2H); **^{13}C NMR** (101 MHz, Chloroform-*d*): δ = 168.06 (C-6), 162.53 (C-4), 161.14 (C-2), 157.46 (C-4”, d, $J_{C\text{-}F}$ = 238.6 Hz), 154.78 (C-1”, d, $J_{C\text{-}F}$ = 1.9 Hz), 116.04 (C-3”, d, $J_{C\text{-}F}$ = 23.1 Hz), 115.56 (C-2”, d, $J_{C\text{-}F}$ = 8.0 Hz), 104.09 (C-5), 66.87 (C-3’), 28.65 (C-2’), 21.16 (C-1’), 12.65 (C-7), 9.02 (C-8); **IR** (νmax / cm^{-1}): 3462 (N-H stretch), 3317 (N-H stretch), 3180 (N-H stretch), 3015 (C-H stretch), 2922 (C-H stretch), 1607 (C=C stretch), 1559 (C=C stretch), 1504 (C=C stretch); **HRMS(ESI)**: *m/z* calculated for $C_{16}H_{20}FN_4O^+$: 303.1616 $[M + H]^+$; found: 303.1612 (1.32 ppm error); **r.f.** 0.33 (70 % EtOAc / Hex).

6.1.7.11 6-Cyclopropyl-5-[3-(3-fluorophenoxy)propyl] pyrimidine-2,4-diamine (**9k**)

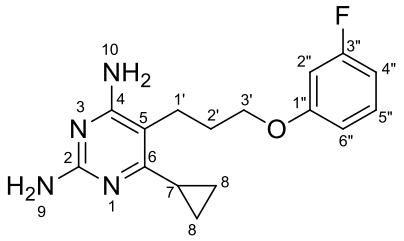


Figure 6.42: 6-Cyclopropyl-5-[3-(3-fluorophenoxy)propyl] pyrimidine-2,4-diamine (**9k**)

Hydrogenation of **8m** (100 mg, 0.335 mmol) with palladium on carbon (10.0 mg, 10 % m/m) in ethanol (10.0 ml) afforded **9k** [Fig 6.42] as a white crystalline solid (63.0 mg, 63.0 %).

m.p. 152 – 155 °C; **^{1}H NMR** (400 MHz, Acetone-d_6): δ = 7.25 (H-5”, q, J = 7.6 Hz, 1H), 6.95 – 6.43 (H-2”, H-4”, H-6”, m, 3H), 5.71 (H-10, brs, 2H), 5.22 (H-9, brs, 2H), 4.01 (H-3’, t, J = 5.8 Hz, 2H), 2.72 (H-1’, t, J = 7.2 Hz, 2H), 2.02 – 1.93 (H-2’ and H-7, m, 3H), 0.94 – 0.79 (H-8, m, 2H), 0.71 – 0.48 (H-8#, m, 2H).**^{13}C NMR** (101 MHz, Acetone-d_6): δ = 167.36 (C-6), 164.45 (H-3”, d, $J_{C\text{-}F}$ = 243.0 Hz), 163.56 (C-4), 162.44 (C-2), 161.52 (C-1”, d, $J_{C\text{-}F}$ = 11.0 Hz), 131.32 (C-5”, d, $J_{C\text{-}F}$ = 10.2 Hz), 111.42 (C-6”, d, $J_{C\text{-}F}$ = 2.7 Hz), 107.77 (C-4”, d, $J_{C\text{-}F}$ = 21.4 Hz), 103.87 (C-5), 102.69 (C-2”, d, $J_{C\text{-}F}$ = 24.9 Hz), 67.91 (C-3’), 29.29 (C-2’), 21.60 (C-1’), 12.74 (C-7), 8.91 (C-8); **IR** (vmax / cm^{-1}): 3493 (N-H stretch), 3393 (N-H stretch), 3188 (N-H stretch), 3011 (C-H stretch), 2934 (C-H stretch), 1612 (C=N stretch), 1559 (C=C stretch); **HRMS(ESI)**: *m/z* calculated for $C_{16}H_{20}FN_4O^+$: 303.1616 $[M + H]^+$; found: 303.1619 (0.990 pm error); **r.f.** 0.33 (70 % EtOAc / Hex).

6.1.7.12 6-Cyclopropyl-5-[3-(4-methoxyphenoxy)propyl] pyrimidine-2,4-diamine (**9l**)

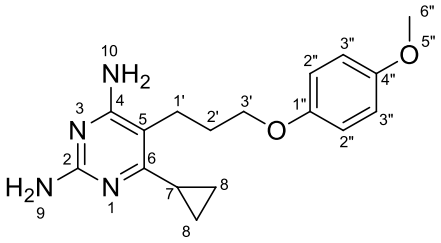


Figure 6.43: 6-Cyclopropyl-5-[3-(4-methoxyphenoxy)propyl] pyrimidine-2,4-diamine (**9l**)

Hydrogenation of **8n** (200 mg, 0.644 mmol) with palladium on carbon (20.0 mg, 10 % m/m) in ethanol (20.0 ml) afforded **9l** [Fig 6.43] as a white crystalline solid (196 mg, 100 %).

m.p. 209 – 211 °C; **^{1}H NMR** (400 MHz, Chloroform-*d*): δ = 6.81 (H-2" and H-3", s, 4H), 5.11 (H-10, brs, 2H), 4.81 (H-9, brs, 2H), 3.93 – 3.84 (H-3', m, 2H), 3.73 (H-6", s, 3H), 2.69 (H-1', t, *J* = 6.6 Hz, 2H), 2.05 – 1.82 (H-7 and H-2', m, 3H), 1.04 – 0.92 (H-8, m, 2H), 0.84 – 0.71 (H-8#, m, 2H); **^{13}C NMR** (101 MHz, Chloroform-*d*): δ = 167.76 (C-6), 162.56 (C-4), 161.08 (C-2), 153.95 (C-4"), 152.64 (C-1"), 115.43 (C-2"), 114.70 (C-3"), 104.03 (C-5), 66.70 (C-3'), 55.69 (C-6"), 28.65 (C-2'), 21.02 (C-1'), 12.58 (C-7), 8.86 (C-8); **IR** (vmax / cm^{-1}): 3483 (N-H stretch), 3381 (N-H stretch), 3187 (N-H stretch), 3005 (C-H stretch), 2934 (C-H stretch), 1612 (C=N stretch), 1562 (C=C stretch), 1505 (C=C stretch), 1148 (C-O stretch); **HRMS(ESI)**: *m/z* calculated for $C_{17}H_{23}N_4O_2^+$: 315.1816 $[M + H]^+$; Found: 315.1819 (0.952 ppm error); **r.f.** 0.38 (70 % EtOAc / Hex).

6.1.7.13 6-Cyclopropyl-5-[3-(3-methoxyphenoxy)propyl] pyrimidine-2,4-diamine (**9m**)

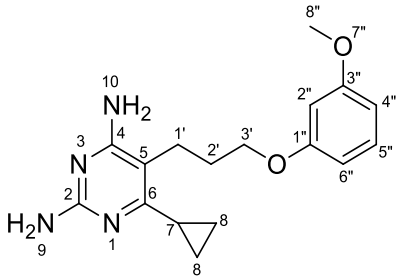


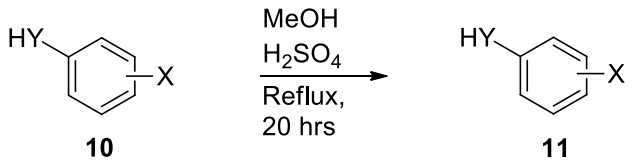
Figure 6.44: 6-Cyclopropyl-5-[3-(3-methoxyphenoxy)propyl] pyrimidine-2,4-diamine (**9m**)

Hydrogenation of **8o** (100 mg, 0.332 mmol) with palladium on carbon (10.0 mg, 10 % m/m) in ethanol (10.0 ml) afforded **9m** [Fig 6.44] as a white crystalline solid (97.0 mg, 96.6 %).

m.p. 132 – 135 °C; **^{1}H NMR** (400 MHz, Chloroform-*d*): δ = 7.17 (H-5”, t, *J* = 8.1 Hz, 1H), 6.53 – 6.47 (H-4” and H-6”, m, 2H), 6.46 (H-2”, s, 1H), 4.95 (H-10, brs, 2H), 4.66 (H-9, brs, 2H), 3.95 (H-3’, t, *J* = 5.1 Hz, 2H), 3.78 (H-8”, s, 3H), 2.72 (H-1’, t, *J* = 6.7 Hz, 2H), 2.05 – 1.97 (H-2’,m, 2H), 1.97 – 1.89 (H-7, m, 1H), 1.05 – 0.99 (H-8, m, 2H), 0.84 – 0.76 (H-8#, m, 2H); **^{13}C NMR** (101 MHz, Chloroform-*d*): δ = 167.91 (C-6), 162.53 (C-4), 161.01 (C-2), 160.96 (C-3”), 159.87 (C-1”), 130.11 (C-5”), 106.87 (C-6”), 106.53 (C-4”), 104.11 (C-5), 101.09 (C-2”), 66.21 (C-3’), 55.37 (C-8”), 28.62 (C-2’), 21.13 (C-1’), 12.65 (C-7), 8.99 (C-8); **IR** (νmax / cm^{-1}): 3492 (N-H stretch), 3371 (N-H stretch), 3162 (N-H stretch), 2999 (C-H stretch), 2924 (C-H stretch), 1664 (C=N stretch), 1559 (C=C stretch); **HRMS(ESI)**: *m/z* calculated for $C_{17}H_{23}N_4O_2^+$: 315.1816 $[M + H]^+$; found: 315.1825 (2.86 ppm error); **r.f.** 0.35 (70 % EtOAc / Hex).

6.2 Synthesis of second-generation analogues

6.2.1 Esterification of carboxylic acid containing phenols and anilines



10 **11**

10a: Y = O, X = 4-COOH
10b: Y = O, X = 2-CH_2CH_2COOH
10c: Y = NH, X = 4-COOH

11a: Y = O, X = 4-COOMe, 97.1% yield
11b: Y = O, X = 2-CH_2CH_2COOMe, 96.8% yield
11c: Y = NH, X = 4-COOMe, 99.7% yield

Scheme 6.8: Esterification of carboxylic acid containing phenols and anilines

The relevant carboxylic acid containing phenol (**10a – c**) (1 eq) was dissolved in MeOH and concentrated H_2SO_4 (1 eq) was added. The mixture was then heated to reflux under Ar overnight. Reactions were monitored by TLC, and after 20 hours no more starting material was observed. The mixture was allowed to cool and concentrated on a rotary evaporator before diluting with H_2O (50.0 ml) and adjusting the pH to 7 using saturated aqueous Na_2CO_3. The product was extracted in EtOAc (100 ml) and washed with brine (100 ml). The organic layer was dried using $MgSO_4$. The mixture was filtered through celite, and solvent removed under reduced pressure. The product was purified using silica gel column chromatography if required (10 – 30 % EtOAc / Hex). The details of the reaction conditions are summarized in scheme 6.8.

6.2.1.1 Methyl 4-hydroxybenzoate (**11a**)

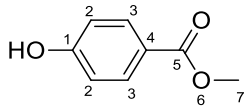


Figure 6.45: Methyl 4-hydroxybenzoate (**11a**)

A mixture of 4-hydroxybenzoic acid **10a** (10.0 g, 72.4 mmol, 1 eq) and concentrated sulfuric acid (3.00 ml) in methanol (100 ml) formed **11a** [Fig 6.45] as a pure white solid without purification (10.7 g, 97.1 %).

m.p. 126 – 128 °C; **^{1}H NMR** (DMSO-d_6, 400 MHz): δ = 10.36 (OH, s, 1H), 7.83 – 7.79 (H-3, m, 2H), 6.86 – 6.82 (H-2, m, 2H), 3.78 (H-7, s, 3H); **^{13}C NMR** (DMSO-d_6, 101 MHz): δ = 166.09 (C-5), 161.98 (C-1), 131.44 (C-3), 120.28 (C-4), 115.37 (C-2), 51.65 (C-7); **IR** (ν_{max} / cm^{-1}): 3287 (O-H stretch), 2963 (C-H stretch), 1676 (C=O stretch), 1606 (C=C stretch), 1587 (C=C stretch), 1273 (C-O stretch); **r.f.** 0.26 (20 % EtOAc / Hex).

6.2.1.2 Methyl 3-(2-Hydroxyphenyl)propanoate (**11b**)

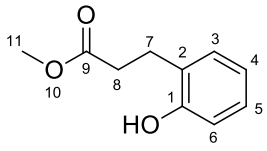


Figure 6.46: Methyl 3-(2-Hydroxyphenyl)propanoate (**11b**)

A mixture of 3-(2-hydroxyphenyl)propanoic acid **10b** (1.00 g, 6.02 mmol, 1 eq) and concentrated sulfuric acid (0.200 ml) in methanol (20.0 ml) formed **11b** [Fig 6.46] as a yellow oil after purification (1.05 g, 96.8 %).

^{1}H NMR (Chloroform-*d*, 400 MHz): δ = 7.14 – 7.07 (H-3 and H-5, m, 2H), 6.90 – 6.84 (H-4 and H-6, m, 2H), 3.69 (H-11, s, 3H), 2.94 – 2.88 (H-7, m, 2H), 2.75 – 2.69 (H-8, m, 2H); **^{13}C NMR**

(Chloroform-*d*, 101 MHz): δ = 176.12 (C-9), 154.28 (C-1), 130.67 (C-3), 128.14 (C-5), 127.31 (C-2), 120.98 (C-4), 117.22 (C-6), 52.38 (C-11), 35.08 (C-8), 24.79 (C-7); **IR** (ν_{max} / cm^{-1}): 3386 (O-H stretch), 2953 (C-H stretch), 1707 (C=O stretch), 1594 (C=C stretch), 1228 (C-O stretch); **r.f.** 0.56 (30 % EtOAc / Hex).

6.2.1.3 Methyl 4-aminobenzoate (**11c**)

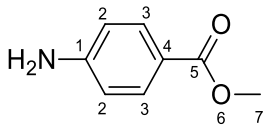
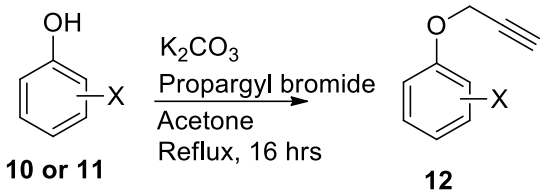


Figure 6.47: Methyl 4-aminobenzoate (**11c**)

A mixture of 4-animobenzoic acid **10c** (5.00 g, 36.5 mmol, 1 eq) and concentrated sulfuric acid (2.00 ml) in methanol (50.0 ml) formed **11c** [Fig 6.47] as a pure white solid without purification (5.50 g, 99.7 %).

m.p. 106 – 108 °C; **^{1}H NMR** (Chloroform-*d*, 300 MHz): δ = 7.87 – 7.82 (H-3, m, 2H), 6.66 – 6.59 (H-2, m, 2H), 4.09 (NH_2, s, 2H), 3.85 (H-7, s, 3H); **^{13}C NMR** (Chloroform-*d*, 75 MHz): δ = 167.29 (C-5), 150.98 (C-1), 131.70 (C-3), 119.78 (C-4), 113.89 (C-2), 51.71 (C-7); **IR** (ν_{max} / cm^{-1}): 3410 (N-H stretch), 3335 (N-H stretch), 1682 (C=O stretch), 1635 (C=C stretch), 1596 (C=C stretch), 1284 (C-O stretch); **r.f.** 0.44 (30 % EtOAc / Hex).

6.2.2 Alkylation of phenols for second-generation analogues



11a: X = 4-COOMe
11b: X = 2-CH_2CH_2COOMe
10d: X = 3,4,5-OMe

12a: X = 4-COOMe, 83.1% yield
12b: X = 2-CH_2CH_2COOMe, 96.6% yield
12c: X = 3,4,5-OMe, 89.5% yield

Scheme 6.9: Alkylation of phenols for second-generation analogues

The relevant phenol starting material (**11a, 11b, or 10d**) (1.00 g, 1 eq) was dissolved in acetone (20.0 ml) and the flask flushed with N_2. After this K_2CO_3 (4 eq) was added, and the mixture was stirred for 10 minutes. Propargyl bromide (1.1 eq) was then added slowly, and the mixture was heated to reflux (55 °C) under N_2 overnight (16 hrs). The mixture was allowed to cool and was filtered through celite to remove the unreacted base. The solvent was evaporated under reduced pressure and the residue dissolved in EtOAc (100 ml). This was washed with saturated aqueous NH_4Cl (100 ml) and brine (100 ml) before being dried with $MgSO_4$ and filtered again through celite. The solvent was removed on a rotary evaporator and the product was purified by silica gel column chromatography (10 % EtOAc / Hex). The details of the reaction conditions are summarized in scheme 6.9.

6.2.2.1 1-(Prop-2-yn-1-yloxy)-4-methylbenzoate (**12a**)

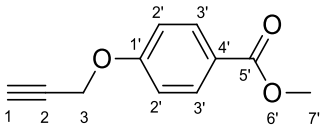


Figure 6.48: 1-(Prop-2-yn-1-yloxy)-4-methylbenzoate (**12a**)

A mixture of methyl-4-hydroxybenzoate **11a** (1.00 g, 6.57 mmol, 1 eq), potassium carbonate (3.63 g, 26.3 mmol, 4 eq), and propargyl bromide (0.860 g, 7.23 mmol, 1.1 eq) in acetone (20.0 ml) formed **12a** [Fig 6.48] as a white solid after purification (1.04 g, 83.1%).

m.p. 64 – 65 °C; **^{1}H NMR** (Chloroform-*d*, 400 MHz): δ = 8.03 – 7.98 (H-3', d, *J* = 9.0 Hz, 2H), 7.02 – 6.97 (H-2', d, *J* = 9.0 Hz, 2H), 4.75 (H-3, d, *J* = 2.4 Hz, 2H), 3.88 (H-7', s, 3H), 2.55 (H-1, t, *J* = 2.4 Hz, 1H); **^{13}C NMR** (Chloroform-*d*, 101 MHz): δ = 166.84 (C-5'), 161.29 (C-1'), 131.69 (C-3'), 123.60 (C-4'), 114.62 (C-2'), 77.95 (C-2), 76.21 (C-1), 55.96 (C-3), 52.07 (C-7'); **IR** (ν_{max} / cm^{-1}): 3242 (C≡C-H stretch), 2960 (C-H stretch), 2865 (C-H stretch), 2129 (C≡C stretch), 1703 (C=O stretch), 1606 (C=C stretch), 1287 (C-O stretch); **r.f.** 0.73 (30 % EtOAc / Hex).

6.2.2.2 [1-(Prop-2-yn-1-yloxy)phenyl]-2-methylpropanoate (**12b**)

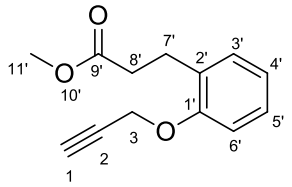


Figure 6.49: [1-(Prop-2-yn-1-yloxy)phenyl]-2-methylpropanoate (12b)

A mixture of methyl 3-(2-hydroxyphenyl)propanoate **11b** (1.00 g, 5.55 mmol, 1 eq), potassium carbonate (3.07 g, 22.2 mmol, 4 eq), and propargyl bromide (0.727 g, 6.11 mmol, 1.1 eq) in acetone (20.0 ml) formed **12b** [Fig 6.49] as a light-yellow oil after purification (1.17 g, 96.6%).

^{1}H NMR (Chloroform-*d*, 400 MHz): δ = 7.24 – 7.15 (H-3' and H-5', m, 2H), 6.99 – 6.90 (H-4 and H-6, m, 2H), 4.72 (H-3, d, *J* = 2.4 Hz, 2H), 3.67 (H-11', s, 3H), 2.98 – 2.93 (H-7, m, 2H), 2.66 – 2.60 (H-8, m, 2H), 2.49 (H-1, t, *J* = 2.4 Hz, 1H); **^{13}C NMR** (Chloroform-*d*, 101 MHz): δ = 173.89 (C-9'), 155.60 (C-1'), 130.35 (C-3'), 129.60 (C-2'), 127.63 (C-5'), 121.58 (C-4'), 111.95 (C-6'), 78.89 (C-2), 75.49 (C-1), 55.91 (C-3), 51.67 (C-11'), 34.20 (C-8'), 26.16 (C-7'); **IR** (ν_{max} / cm^{-1}): 3287 (C≡C-H stretch), 2951 (C-H stretch), 2122 (C≡C stretch), 1730 (C-O stretch), 1492 (C=C stretch), 1221 (C-O stretch); **HRMS(ESI)**: *m/z* calculated $C_{13}H_{15}O_3^+$: 219.1016 $[M + H]^+$; found: 219.1013 (1.37 ppm error); **r.f.** 0.50 (20 % EtOAc / Hex).

6.2.2.3 1-(Prop-2-yn-1-yloxy)-3,4,5-trimethoxybenzene (**12c**)

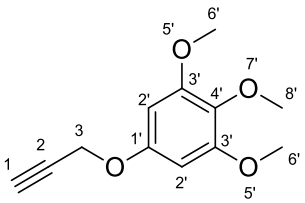
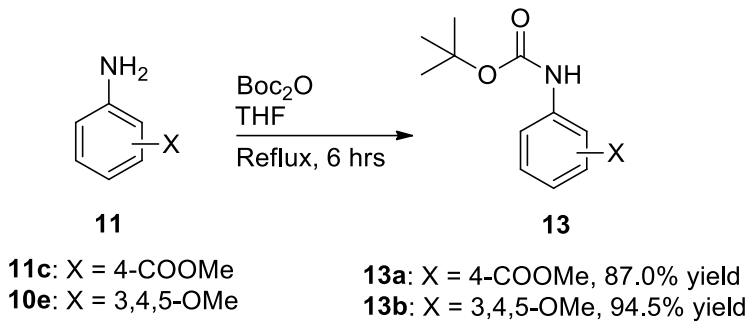


Figure 6.50: 1-(Prop-2-yn-1-yloxy)-3,4,5-trimethoxybenzene (**12c**)

A mixture of 3,4,5-trimethoxyphenol **10c** (1.00 g, 5.43 mmol, 1 eq), potassium carbonate (3.00 g, 21.7 mmol, 4 eq), and propargyl bromide (0.710 g, 5.97 mmol, 1.1 eq) in acetone (20.0 ml) formed **12c** [Fig 6.50] as a white low-melting solid after purification (1.08 g, 89.5%).

^{1}H NMR (Chloroform-*d*, 400 MHz): δ = 6.24 (H-2', s, 2H), 4.67 (H-3, d, *J* = 2.4 Hz, 2H), 3.84 (H-6', s, 6H), 3.79 (H-8', s, 3H), 2.55 (H-1, t, *J* = 2.4 Hz, 1H); **^{13}C NMR** (Chloroform-*d*, 101 MHz): δ = 154.30 (C-1'), 153.83 (C-3'), 133.08 (C-4'), 92.98 (C-2'), 78.71 (C-2), 75.77 (C-1), 61.16 (C-6'), 56.44 (C-3), 56.25 (C-8'); **IR** (ν_{max} / cm^{-1}): 3266 (C≡C-H stretch), 2943 (C-H stretch), 2840 (C-H stretch), 2128 (C≡C stretch), 1593 (C=C stretch), 1502 (C=C stretch), 1225 (C-O stretch), 1123 (C-O stretch); **HRMS(ESI)**: *m/z* calculated $C_{12}H_{15}O_4^+$: 223.0965 $[M + H]^+$; found: 223.0959 (2.69 ppm error); **r.f.** 0.44 (30 % EtOAc / Hex).

6.2.3 Boc protection of anilines



Scheme 6.10: Boc protection of anilines

The appropriate amine (**11c** or **10e**) (0.500 g, 1 eq) was dissolved in a minimal amount of THF (4.00 ml) before Boc_2O (1.5 eq) was added. The mixture was heated to reflux under Ar. The reaction was monitored by TLC. After six hours no more starting material was observed, and the mixture was allowed to cool. The solvent was removed on a rotary evaporator and purified immediately using silica gel chromatography (10 – 30 % EtOAc / Hex). The details of the reaction conditions are summarized in scheme 6.10.

6.2.3.1 Methyl 4-[(*tert*-butoxycarbonyl)amino]benzoate (**13a**)

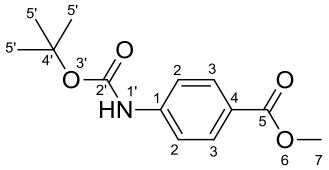


Figure 6.51: Methyl 4-[(*tert*-butoxycarbonyl)amino]benzoate (**13a**)

Reacting **11c** (0.500 g, 3.31 mmol, 1 eq) with di-*tert*-butyl dicarbonate (1.08 g, 4.97 mmol, 1.5 eq) in tetrahydrofuran (4.00 ml) afforded **13a** [Fig 6.51] as a white solid after purification (0.660 g, 87.0 %).

m.p. 158 – 160 °C; **^{1}H NMR** (Chloroform-*d*, 400 MHz): δ = 7.97 (H-3, d, *J* = 8.8 Hz, 2H), 7.43 (H-2, d, *J* = 8.8 Hz, 2H), 6.66 (H-1', brs, 1H), 3.89 (H-7, s, 3H), 1.52 (H-5', s, 9H); **^{13}C NMR** (Chloroform-*d*, 101 MHz): δ = 166.87 (C-5), 152.31 (C-2'), 142.81 (C-1), 131.04 (C-3), 124.55 (C-4), 117.48 (C-2), 81.39 (C-4'), 52.08 (C-7), 28.42 (C-5'); **IR** (ν_{max} / cm^{-1}): 3330 (N-H stretch), 2983 (C-H stretch), 1708 (C=O stretch), 1596 (C=C stretch), 1528 (C=C stretch), 1159 (C-O stretch); **r.f.** 0.53 (20 % EtOAc / Hex).

6.2.3.2 *tert*-Butyl-(3,4,5-trimethoxyphenyl)carbamate (**13b**)

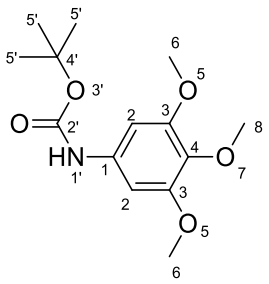
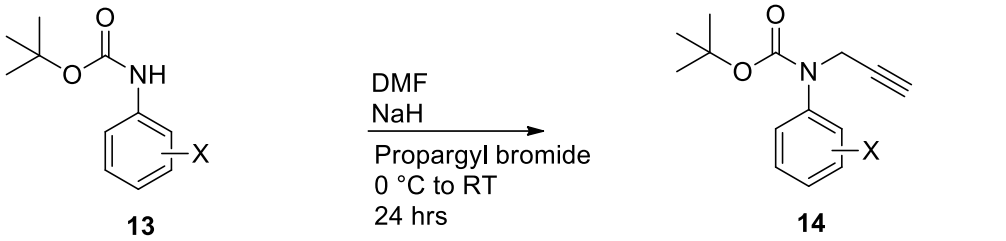


Figure 6.52: *tert*-Butyl-(3,4,5-trimethoxyphenyl)carbamate (**13b**)

Reacting **10e** (0.500 g, 2.73 mmol, 1 eq) with di-*tert*-butyl dicarbonate (0.895 g, 4.10 mmol, 1.5 eq) in tetrahydrofuran (4.00 ml) afforded **13b** [Fig 6.51] as a white solid after purification (0.728 g, 94.5 %).

m.p. 150 – 151 °C; **^{1}H NMR** (400 MHz, Chloroform-*d*) δ 6.65 (H-2, s, 2H), 6.46 (NH, brs, 1H), 3.83 (H-6, s, 6H), 3.79 (H-8, s, 3H), 1.51 (H-5', s, 9H); **^{13}C NMR** (101 MHz, Chloroform-*d*) δ 153.53 (C-2), 152.87 (C-2'), 134.69 (C-1), 133.84 (C-4), 96.22 (C-3), 80.65 (C-4'), 61.11 (C-8), 56.19 (C-6), 28.48 (C-5); **IR** (ν_{max} / cm^{-1}): 3304 (N-H stretch), 2980 (C-H stretch), 2939 (C-H stretch), 1715 (C=O stretch), 1610 (C=C stretch), 1227 (C-O stretch), 1158 (C-O stretch), 1130 (C-O stretch); **HRMS(ESI)**: *m/z* calculated for $C_{14}H_{22}NO_5^+$: 284.1492 $[M + H]^+$; found: 284.1489 (1.06 ppm error); **r.f.** 0.48 (30 % EtOAc / Hex).

6.2.4 Alkylation of protected anilines



13

13a: X = 4-COOMe
13b: X = 3,4,5-OMe

14

14a: X = 4-COOMe, 85.6% yield
14b: X = 3,4,5-OMe, 83.7% yield

Scheme 6.11: Alkylation of protected anilines

The selected starting material (**13a** or **13b**) (1.00 g, 1 eq) was dissolved in DMF (20 ml) and flushed with Ar. The solution was cooled to 0 °C in an ice bath and NaH (1.5 eq) was added. After stirring for 10 minutes, propargyl bromide (1.1 eq) was added. The mixture was then stirred at room temperature under Ar. The reaction was monitored by TLC. The Rf of the starting material and the product overlap, however, ninhydrin stain turns the product yellow and the starting material brown when developed. After 24 hours no more starting material could be observed. The reaction was slowly quenched with saturated aqueous NH_4Cl (50.0 ml), and the product extracted in EtOAc (100 ml). The organic layer was washed several times with brine (3 x 100 ml) and dried using $MgSO_4$. This mixture was filtered through celite, and the solvent removed on a rotary evaporator. Purification was achieved by silica gel column chromatography (10 – 20 % EtOAc / Hex). The details of the reaction conditions are summarized in scheme 6.11.

6.2.4.1 Methyl 4-[(*tert*-butoxycarbonyl)(prop-2-yn-1-yl)amino]benzoate (**14a**)

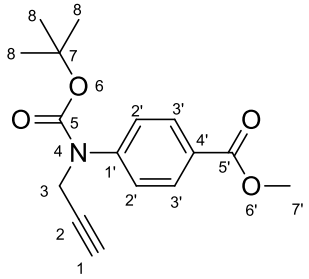


Figure 6.53: Methyl 4-[(*tert*-butoxycarbonyl)(prop-2-yn-1-yl)amino]benzoate (**14a**)

Alkylation of **13a** (1.00 g, 3.98 mmol, 1 eq) with propargyl bromide (0.521 g, 4.38 mmol, 1.1 eq) using sodium hydride (0.143 g, 5.97 mmol, 1.5 eq) in dimethylformamide (20.0 ml) afforded **14a** [Fig 6.53] as an off-white solid after purification (0.986 g, 85.6 %).

m.p. 89 – 92 °C; **^{1}H NMR** (Chloroform-*d*, 400 MHz): δ = 8.02 – 7.98 (H-3', m, 2H), 7.43 – 7.39 (H-2' m, 2H), 4.39 (H-3, d, *J* = 2.5 Hz, 2H), 3.89 (H-7', s, 3H), 2.26 (H-1, t, *J* = 2.4 Hz, 1H), 1.46 (H-8, s, 9H); **^{13}C NMR** (Chloroform-*d*, 101 MHz): δ = 166.62 (C-5'), 153.45 (C-5), 146.36 (C-1'), 130.22 (C-3'), 127.37 (C-4'), 125.22 (C-2'), 81.89 (C-7), 79.66 (C-2), 72.25 (C-1), 52.18 (C-7'), 39.50 (C-3), 28.30 (C-8); **IR** (ν_{max} / cm^{-1}): 3272 (C≡C-H stretch), 2981 (C-H stretch), 2119 (C≡C stretch), 1710 (C=O stretch), 1604 (C=C stretch); **r.f.** 0.56 (20 % EtOAc / Hex).

6.2.4.2 *tert*-Butyl prop-2-yn-1-yl(3,4,5-trimethoxyphenyl)carbamate (**14b**)

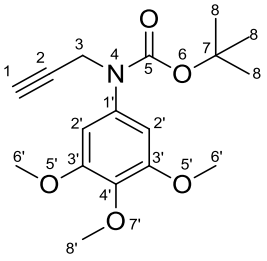
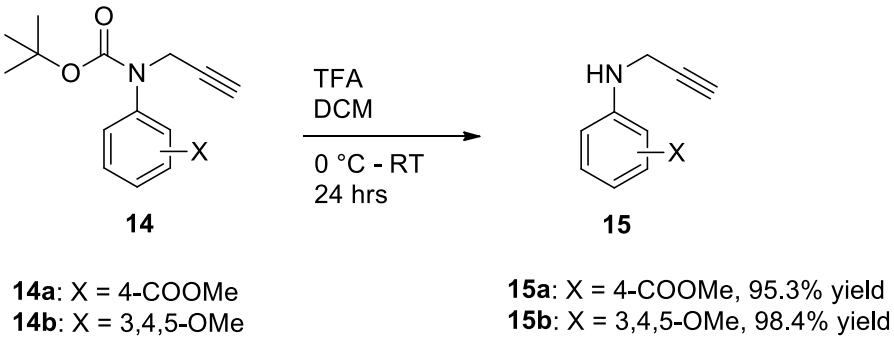


Figure 6.54: *tert*-Butyl prop-2-yn-1-yl(3,4,5-trimethoxyphenyl)carbamate (**14b**)

Alkylation of **13a** (1.00 g, 3.53 mmol, 1 eq) with propargyl bromide (0.521 g, 3.88 mmol, 1.1 eq) using sodium hydride (0.127 g, 5.30 mmol, 1.5 eq) in dimethylformamide (20.0 ml) afforded **14b** [Fig 6.54] as a low-melting yellow solid (m.p. < 30 °C) after purification (0.950 g, 83.7 %).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 6.58 (H-2', s, 2H), 4.32 (H-3, d, *J* = 2.1 Hz, 2H), 3.84 (H-6', s, 6H), 3.84 (H-8', s, 3H), 2.28 (H-1, t, *J* = 2.4 Hz, 1H), 1.48 (H-8, s, 9H); **^{13}C NMR** (101 MHz, Chloroform-*d*) δ 154.22 (C-3'), 153.16 (C-5), 138.12 (C-1'), 136.74 (C-4'), 104.30 (C-2'), 81.29 (C-7), 80.40 (C-2), 71.82 (C-1), 61.02 (C-8'), 56.26 (C-6'), 40.23 (C-3) 28.49 (C-8); **IR (ν_{max} / cm^{-1})**: 3286 (C≡C-H stretch), 2975 (C-H stretch), 2938 (C-H stretch), 1605 (C=O stretch), 1592 (C=C stretch), 1505 (C=C stretch), 1229 (C-O stretch), 1163 (C-O stretch), 1122 (C-O stretch); **r.f.** 0.42 (70 % EtOAc / Hex); **r.f.** 0.50 (20 % EtOAc / Hex).

6.2.5 Deprotection of alkylated anilines



Scheme 6.12: Deprotection of alkylated anilines

The BOC protected aniline (**14a** or **14b**) (1.00 g, 1 eq) was dissolved in DCM (20 ml) and cooled to 0 °C in an ice bath while under Ar. TFA (5 eq) was then added slowly over 10 minutes, and the mixture stirred at room temperature for 24 hours. The reaction was monitored by TLC. After all the starting material was consumed, the reaction was quenched with brine (50.0 ml), and the organic layer isolated. This was washed sequentially with saturated aqueous NH_4Cl (50.0 ml) and brine (50.0 ml). The organic layer was then dried with $MgSO_4$ and filtered through celite. The solvent was removed using a rotary evaporator, and the product purified

using silica gel column chromatography (20 % EtOAc / Hex). The details of the reaction conditions are summarized in scheme 6.12.

6.2.5.1 Methyl 4-[(*tert*-butoxycarbonyl)amino]benzoate (**15a**)

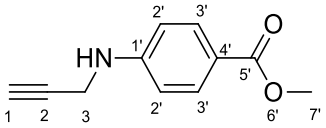


Figure 6.55: Methyl 4-[(*tert*-butoxycarbonyl)amino]benzoate (**15a**)

Deprotection of **14a** (1.00 g, 3.46 mmol, 1 eq) was achieved by a reaction with trifluoroacetic acid (1.97 g, 17.3 mmol, 5 eq) in dichloromethane (20.0 ml) to afford **15a** [Fig 6.55] as a yellow solid after purification (0.624 g, 95.3 %).

m.p. 88 – 90 °C**; ^{1}H NMR** (Chloroform-*d*, 400 MHz): δ = 7.91 (H-3', d, *J* = 8.7 Hz, 2H), 6.65 (H-2', d, *J* = 8.8 Hz, 2H), 4.34 (NH, brs, 1H), 3.99 (H-3, dd, *J* = 6.0, 2.4 Hz, 2H), 3.86 (H-7's, 3H), 2.25 (H-1, t, *J* = 2.5 Hz, 1H); **^{13}C NMR** (Chloroform-*d*, 101 MHz): δ = 167.32 (C-5'), 150.69 (C-1'), 131.61 (C-3'), 119.89 (C-4'), 112.33 (C-2'), 80.05 (C-2), 71.94 (C-3), 51.79 (C-7'), 33.25 (C-3); **IR** (ν_{max} / cm^{-1}): 3352 (N-H stretch), 3290 (C≡C-H stretch), 2949 (C-H stretch), 2885 (C-H stretch), 2119 (C≡C stretch), 1688 (C=O stretch), 1599 (C-O stretch); **r.f.** 0.42 (20 % EtOAc / Hex).

6.2.5.2 3,4,5-Trimethoxy-*N*-(prop-2-yn-1-yl)aniline (**15b**)

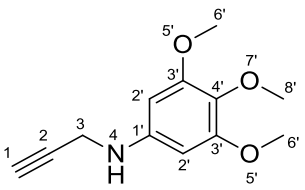
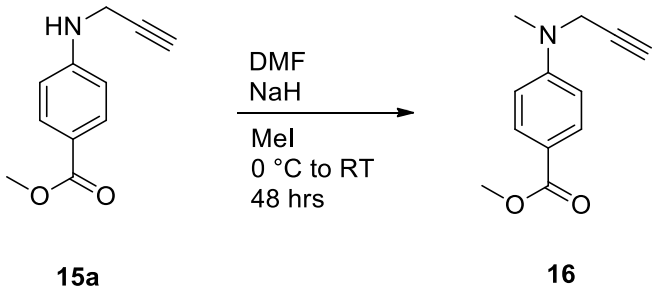


Figure 6.56: 3,4,5-Trimethoxy-*N*-(prop-2-yn-1-yl)aniline (**15b**)

Deprotection of **14b** (1.00 g, 3.11 mmol, 1 eq) was achieved by a reaction with trifluoroacetic acid (1.78 g, 15.6 mmol, 5 eq) in dichloromethane (20.0 ml) to afford **15b** [Fig 6.56] as an orange solid after purification (0.677 g, 98.4 %).

m.p. 69 – 72 °C; **^{1}H NMR** (300 MHz, Chloroform-*d*): δ = 5.93 (H-2', s, 2H), 3.91 (H-3, d, *J* = 2.1 Hz, 2H), 3.82 (H-6', s, 6H), 3.76 (H-8', s, 3H), 2.24 (H-1, t, *J* = 2.1 Hz, 1H); **^{13}C NMR** (75 MHz, Chloroform-*d*): δ = 154.03 (C-3'), 143.68 (C-1'), 131.03 (C-4'), 91.34 (C-2'), 81.05 (C-2), 71.56 (C-1), 61.15 (C-8'), 56.08 (C-6'), 34.15 (C-3); **IR** (ν_{max} / cm^{-1}): 3309 (N-H stretch), 3269 (C≡C-H stretch), 2929 (C-H stretch), 2112 (C≡C stretch), 1559 (C=C stretch), 1506 (C-O stretch), 1451 (C-O stretch); **r.f.** 0.33 (30 % EtOAc / Hex).

6.2.6 Methylation of alkylated anilines



Scheme 6.13: Methylation of alkylated anilines

The starting material **15a** (0.500 g, 2.64 mmol, 1 eq) was dissolved in DMF (15.0 ml), and the flask flushed with Ar. The mixture was then cooled to 0 °C in an ice bath, and NaH (0.0950 g, 3.96 mmol, 1.5 eq) was added. After stirring for 20 minutes, MeI (0.562 g, 3.96 mmol, 1.5 eq) was added slowly, and the reaction mixture stirred under Ar for 48 hours. The reaction was monitored by TLC, and after all starting material had been consumed the reaction was quenched with saturated aqueous NH_4Cl (50.0 ml) and the product was extracted into EtOAc (50.0 ml). The organic layer was washed with several portions of brine (3 x 50.0 ml), before being dried with $MgSO_4$ and filtered through celite. The solvent was removed on a rotary evaporator and the product purified with silica gel column chromatography (10 % EtOAc / Hex) to afford product **16** [Fig 6.57] as a pale yellow solid (0.463 g, 75.5 %). The details of the reaction conditions are summarized in scheme 6.13.

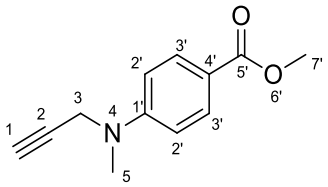
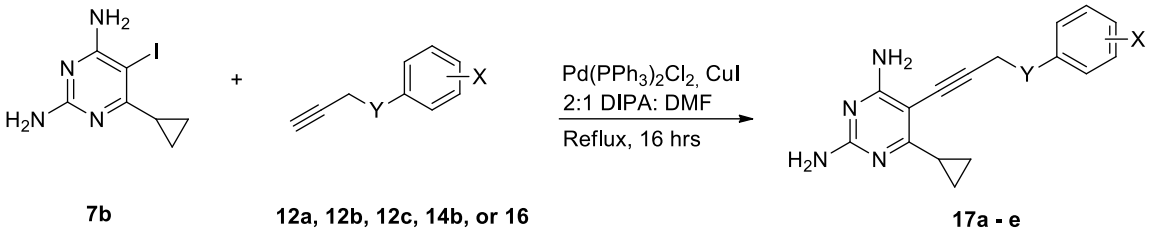


Figure 6.57: Methyl 4-[methyl(prop-2-yn-1-yl)amino]benzoate (**16**)

m.p. 92 – 95 °C; **^{1}H NMR** (Chloroform-*d*, 400 MHz): δ = 7.96 – 7.92 (H-2', m, 2H), 6.80 – 6.75 (H-3', m, 2H), 4.11 (H-3, d, 2H, *J*=2.4 Hz), 3.86 (H-7', s, 3H), 3.06 (H-5, s, 3H), 2.21 (H-1, t, 1H, *J*=2.4 Hz); **^{13}C NMR** (Chloroform-*d*, 101 MHz): δ = 167.40 (C-5'), 152.15 (C-1'), 131.38 (C-2'), 118.92 (C-4'), 112.30 (C-3'), 78.74 (C-2), 72.38 (C-1), 51.74 (C-7'), 41.96 (C-3), 38.48 (C-5); **IR** (ν_{max} / cm^{-1}): 3252 (C≡C-H stretch), 2992 (C-H stretch), 2853 (C-H stretch), 2109 (C≡C stretch), 1689 (C=O stretch), 1608 (C=C stretch); **HRMS(ESI)**: *m/z* calculated for $C_{12}H_{14}NO_2^+$: 204.1019 $[M + H]^+$; found: 204.1005 (6.86 ppm error); **r.f.** 0.22 (10 % EtOAc / Hex).

6.2.7 Sonogashira coupling reaction for second-generation analogues



7b **12a, 12b, 12c, 14b, or 16** **17a - e**

Scheme 6.14: Sonogashira coupling reaction for second-generation analogues

DMF (5.00 ml) was added to a 50.0 ml 3-neck round bottom flask equipped with a condenser. The solvent was degassed under vacuum and Ar was bubbled through for several minutes to remove any dissolved oxygen. DIPA (10.0 ml) was then added, and Ar bubbled through the solution for several more minutes. After this the 2,4-diamino-5-iodopyrimidine starting material **7b** (200 – 300 mg, 1 eq) was added together with Pd(PPh3)2Cl2 (0.1 eq) and CuI (0.2 eq). The mixture was then slowly heated to reflux (80 °C) while stirring until all the solid material was dissolved. The relevant previously prepared alkyne (**12a – c, 14b, or 16**) (2 eq) was then added, and the resulting solution heated overnight (16 hours) under Ar. After this the mixture was allowed to cool and was quenched with saturated aqueous NH4Cl (200 ml). The product was extracted into EtOAc (200ml), and the organic layer was washed with several portions of brine (3 x 200 ml) and dried using MgSO4. The mixture was then filtered through celite and concentrated on a rotary evaporator. The crude product was adsorbed onto silica and purified by column chromatography (70 % EtOAc / hex). The reaction information is summarized in scheme 6.14, and the full list of products formed from this reaction is displayed in table 6.3.

Table 6.3: Full list of Sonogashira reaction products in the synthesis of second-generation analogues

Pyrimidine:	Alkyne:	Product:
7b, R = Cpr	**12a**, X = 4-COOMe, Y = O	**17a**
	12b, X = 2-CH_2CH_2COOMe, Y = O	**17b**
	12c, X = 3,4,5-OMe, Y = O	**17c**
	14b, X = 3,4,5-OMe, Y = NBoc	**17d**
	16, X = 4-COOMe, Y = NMe	**17e**

6.2.7.1 Methyl 4-{[3-(2,4-diamino-6-cyclopropylpyrimidin-5-yl) prop-2-yn-1-yl]oxy} benzoate (**17a**)

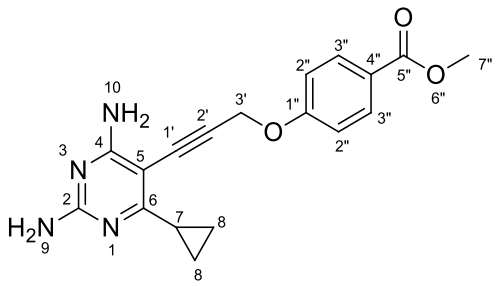


Figure 6.58: Methyl 4-{[3-(2,4-diamino-6-cyclopropylpyrimidin-5-yl)prop-2-yn-1-yl]oxy} benzoate (**17a**)

A mixture of **7b** (200 mg, 0.725 mmol, 1 eq), bis(triphenylphosphine)palladium(II) dichloride (50.9 mg, 0.0725 mmol, 0.1 eq), copper iodide (27.6 mg, 0.145 mmol, 0.2 eq), **12a** (276 mg, 1.45 mmol, 2 eq), dimethylformamide (5.00 ml), and diisopropylamine (10.0 ml) yielded **17a** [Fig 6.58] as a yellow polycrystalline material after purification (175 mg, 70.8 %).

m.p. 134 – 136 °C; **^{1}H NMR** (DMSO-d_6, 300 MHz): δ = 7.92 (H-3”, d, *J* = 8.8 Hz, 2H), 7.19 (H-2”, d, *J* = 8.8 Hz, 2H), 6.41 (H-10, brs, 2H), 6.14 (H-9, brs, 2H), 5.16 (H-3’, s, 2H), 3.81 (H-7”, s, 3H), 2.15 (H-7, tt, *J* = 8.1, 4.7 Hz, 1H), 0.90 – 0.83 (H-8, m, 2H), 0.82 – 0.74 (H-8#, m, 2H); **^{13}C NMR** (DMSO-d_6, 75 MHz): δ = 171.88 (C-6), 165.89 (C-5”), 164.26 (C-4), 161.77 (C-2), 161.24 (C-1”),

131.08 (C-3”), 122.29 (C-4”), 115.17 (C-2”), 92.35 (C-1’), 86.84 (C-5), 82.36 (C-2’), 57.09 (C-3’), 51.89 (C-7”), 14.46 (C-7), 9.02 (C-8); **IR** (νmax / cm^{-1}): 3416 (N-H stretch), 3335 (N-H stretch), 3203 (N-H stretch), 2221 (C≡C stretch), 1737 (C=O stretch), 1687 (C=N stretch), 1618 (C=C stretch), 1539 (C=C stretch), 1240 (C-O stretch); **HRMS(ESI)**: *m/z* calculated for $C_{18}H_{19}N_4O_3^+$: 339.1452 $[M + H]^+$; found: 339.1439 (3.83 ppm error); **r.f.** 0.23 (70 % EtOAc / Hex).

6.2.7.2 Methyl 3-{2-[(3-{2,4-diamino-6-cyclopropylpyrimidin-5-yl} prop-2-yn-1-yl)oxy]phenyl} propanoate (**17b**)

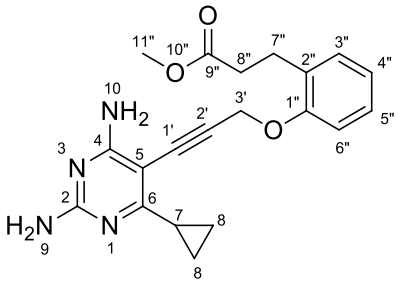


Figure 6.59: Methyl 3-{2-[(3-{2,4-diamino-6-cyclopropylpyrimidin-5-yl} prop-2-yn-1-yl)oxy]phenyl} propanoate (**17b**)

A mixture of **7b** (300 mg, 1.09 mmol, 1 eq), bis(triphenylphosphine)palladium(II) dichloride (76.5 mg, 0.109 mmol, 0.1 eq), copper iodide (41.5 mg, 0.218 mmol, 0.2 eq), **12b** (476 mg, 2.18 mmol, 2 eq), dimethylformamide (5.00 ml), and diisopropylamine (10.0 ml) yielded **17b** [Fig 6.59] as a yellow polycrystalline material after purification (303 mg, 75.9 %).

m.p. 288 – 291 °C; **^{1}H NMR** (Chloroform-*d*, 400 MHz): δ = 7.21 – 7.15 (H-3” and H-5”, m, 2H), 7.02 (H-6”, d, *J* = 8.1 Hz, 1H), 6.91 (H-4”, t, *J* = 7.0 Hz, 1H), 5.10 (H-10, brs, 2H), 5.02 (H-9, brs, 2H), 4.85 (H-3’, s, 2H), 3.65 (H-11”, s, 3H), 2.97 (H-7”, t, *J* = 7.8 Hz, 2H), 2.66 – 2.60 (H-8”, m, 2H), 2.26 – 2.18 (H-7, m, 1H), 1.05 – 1.00 (H-8, m, 2H), 0.90 – 0.84 (H-8#, m, 2H); **^{13}C NMR** (Chloroform-*d*, 101 MHz): δ = 173.87 (C-9”), 173.59 (C-6), 164.10 (C-4), 161.36 (C-2), 155.54 (C-1”), 130.46 (C-3”), 129.95 (C-2”), 127.48 (C-5”), 121.54 (C-4”), 112.37 (C-6”), 93.09 (C-2’), 89.56 (C-5), 81.42 (C-1’), 57.01 (C-3’), 51.67 (C-11”), 34.27 (C-8”), 26.25 (C-7”), 14.99 (C-7), 9.86 (C-8); **IR** (νmax / cm^{-1}): 3373 (N-H stretch), 3103 (C-H stretch), 2220 (C≡C stretch), 1666

(C=O stretch), 1598 (C=N stretch), 1544 (C=C stretch), 1388 (C-O stretch), 1209 (C-O stretch); **HRMS(ESI)**: *m/z* calculated for $C_{19}H_{27}N_4O_4^+$: 367.1765 $[M + H]^+$; found: 367.1755 (2.72 ppm error); **r.f.** 0.42 (70 % EtOAc / Hex).

6.2.7.3 6-Cyclopropyl-5-[3-(3,4,5-trimethoxyphenoxy) prop-1-yn-1-yl] pyrimidine-2,4-diamine (**17c**)

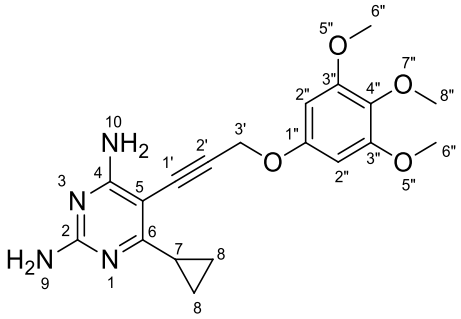


Figure 6.60: 6-Cyclopropyl-5-[3-(3,4,5-trimethoxyphenoxy)prop-1-yn-1-yl] pyrimidine-2,4-diamine (**17c**)

A mixture of **7b** (300 mg, 1.09 mmol, 1 eq), bis(triphenylphosphine)palladium(II) dichloride (76.5 mg, 0.109 mmol, 0.1 eq), copper iodide (41.5 mg, 0.218 mmol, 0.2 eq), **12c** (484 mg, 2.18 mmol, 2 eq), dimethylformamide (5.00 ml), and diisopropylamine (10.0 ml) yielded **17c** [Fig 6.60] as a yellow polycrystalline material after purification (310 mg, 77.1 %).

m.p. 122 – 124 °C; **^{1}H NMR** (Chloroform-*d*, 500 MHz): δ = 6.29 (H-2", s, 2H), 5.07 (H-10, brs, 2H), 4.96 (H-9, brs, 2H), 4.78 (H-3', s, 2H), 3.83 (H-6", s, 6H), 3.78 (H-8", s, 3H), 2.26 (H-7, tt, 1H, *J* = 8.3, 4.7 Hz), 1.08 – 1.01 (H-8, m, 2H), 0.91 – 0.84 (H-8#, m, 2H); **^{13}C NMR** (Chloroform-*d*, 126 MHz): δ = 173.72 (C-6), 164.16 (C-4), 161.41 (C-2), 154.29 (C-1"), 153.83 (C-3"), 133.08 (C-4"), 93.38 (C-2"), 92.95 (C-1'), 89.48 (C-5), 81.75 (C-2'), 61.15 (C-8"), 57.62 (C-3'), 56.25 (C-6"), 15.03 (C-7), 9.96 (C-8); **IR** (vmax / cm^{-1}): 3472 (N-H stretch), 3336 (N-H stretch), 3195 (N-H stretch), 2961 (C-H stretch), 2972 (C-H stretch), 2205 (C≡C stretch), 1594 (C=N stretch), 1503 (C=C stretch), 1118 (C-O stretch), 1024 (C-O stretch); **HRMS(ESI)**: *m/z* calculated for

$C_{19}H_{23}N_4O_4^+$: 371.1714 [M + H]$^+$; found: 371.1693 (5.66 ppm error); **r.f.** 0.34 (70 % EtOAc / Hex).

6.2.7.4 Methyl 4-[(*tert*-butoxycarbonyl)[3-(2,4-diamino-6-cyclopropylpyrimidin-5-yl)prop-2-yn-1-yl]amino] benzoate (**17d**)

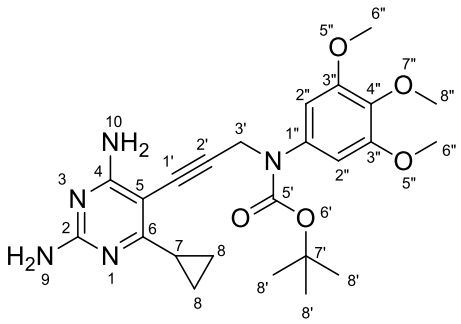


Figure 6.61: Methyl 4-[(*tert*-butoxycarbonyl)[3-(2,4-diamino-6-cyclopropylpyrimidin-5-yl)prop-2-yn-1-yl]amino] benzoate (**17d**)

A mixture of **7b** (200 mg, 0.725 mmol, 1 eq), bis(triphenylphosphine)palladium(II) dichloride (50.9 mg, 0.0725 mmol, 0.1 eq), copper iodide (27.6 mg, 0.145 mmol, 0.2 eq), **14b** (466 mg, 1.45 mmol, 2 eq), dimethylformamide (5.00 ml), and diisopropylamine (10.0 ml) yielded **17d** [Fig 6.61] as an off-white polycrystalline material after purification (210 mg, 61.7 %).

m.p. 174 – 176 °C; **^{1}H NMR** (400 MHz, DMSO-d_6): δ = 6.64 (H-2”, s, 2H), 6.27 (H-10, brs, 2H), 6.06 (H-9, brs, 2H), 4.69 (H-3’, s, 2H), 3.73 (H-6”, s, 6H), 3.63 (H-8”, s, 3H), 2.05 – 1.93 (H-7, m, 1H), 1.40 (H-8’, s, 9H), 0.88 – 0.80 (H-8, m, 2H), 0.77 – 0.67 (H-8#, m, 2H); **^{13}C NMR** (101 MHz, DMSO-d_6): δ = 170.95 (C-6), 163.93 (C-4), 161.52 (C-2), 153.33 (C-3”), 152.44 (C-5’), 137.47 (C-1”), 135.56 (C-4”), 104.25 (C-2”), 93.81 (C-2’), 87.61 (C-5), 80.23 (C-1’), 78.18 (C-7’), 60.05 (C-8”), 55.92 (C-6”), 40.54 (C-3’), 27.97 (C-8’), 14.23 (C-7), 8.76 (C-8); **IR** (νmax / cm^{-1}): 3475 (N-H stretch), 3356 (N-H stretch), 3156 (N-H stretch), 2961 (C-H stretch), 2835 (C-H stretch), 2193 (C≡C stretch), 1687 (C=O stretch), 1610 (C=N stretch), 1596 (C=C stretch), 1540 (C=C stretch), 1232 (C-O stretch); **HRMS(ESI)**: *m/z* calculated for $C_{24}H_{32}N_5O_5^+$: 470.2398 [M + H]$^+$; found: 470.2393 (1.06 ppm error); **r.f.** 0.34 (70 % EtOAc / Hex).

6.2.7.5 Methyl 4-{[3-(2,4-diamino-6-cyclopropylpyrimidin-5-yl) prop-2-yn-1-yl](methyl)amino} benzoate (**17e**)

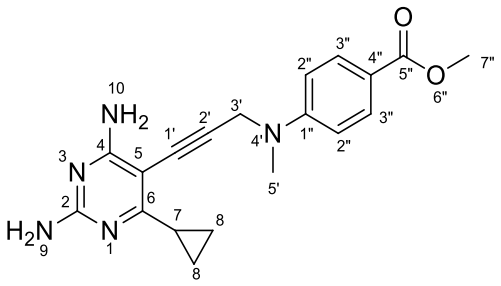
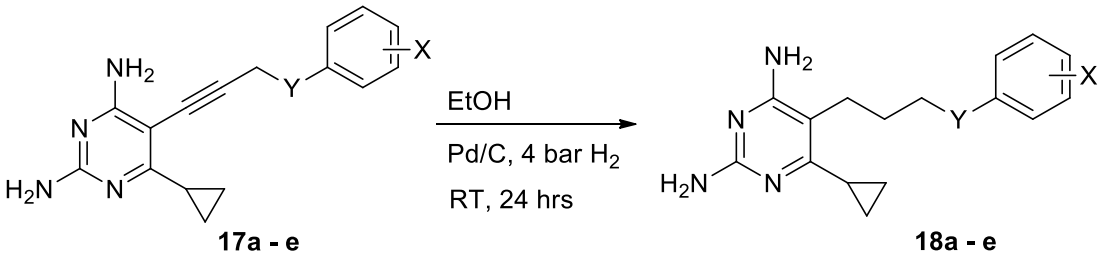


Figure 6.62: Methyl 4-{[3-(2,4-diamino-6-cyclopropylpyrimidin-5-yl) prop-2-yn-1-yl](methyl)amino} benzoate (**17e**)

A mixture of **7b** (200 mg, 0.725 mmol, 1 eq), bis(triphenylphosphine)palladium(II) dichloride (50.9 mg, 0.0725 mmol, 0.1 eq), copper iodide (27.6 mg, 0.145 mmol, 0.2 eq), **16** (315 mg, 1.45 mmol, 2 eq), dimethylformamide (5.00 ml), and diisopropylamine (10.0 ml) yielded **17e** [Fig 6.62] as an orange polycrystalline material after purification (180 mg, 70.6 %).

m.p. 219 – 221 °C; **^{1}H NMR** (300 MHz, DMSO-d_6): δ = 7.79 (H-3”, d, *J* = 9.0 Hz, 2H), 6.94 (H-2”, d, *J* = 9.0 Hz, 2H), 6.23 (H-10, brs, 2H), 6.04 (H-9, brs, 2H), 4.48 (H-3’. s, 2H), 3.77 (H-7”, s, 3H), 3.05 (H-5’, s, 3H), 2.07 (H-7, tt, *J* = 8.3, 4.8 Hz, 1H), 0.82 (H-8, m, 2H), 0.71 (H-8#, m, 2H); **^{13}C NMR** (101 MHz, DMSO-d6) δ 171.03 (C-6), 166.28 (C-5”), 163.93 (C-4), 161.45 (C-2), 152.38 (C-1”), 130.56 (C-3”), 117.18 (C-4”), 112.54 (C-2”), 93.09 (C-1’), 87.61 (C-5), 78.03 (C-2’), 51.41 (C-3’), 42.61 (C-7”), 38.15 (C-5’), 14.34 (C-7), 8.81 (C-8); **IR** (vmax / cm^{-1}): 3437 (N-H stretch), 3404 (N-H stretch), 3303 (N-H stretch), 2208 (C≡C stretch), 1691 (C=O stretch), 1598 (C=N stretch), 1544 (C=C stretch), 1435 (C=C stretch), 1286 (C-O stretch); **HRMS(ESI)**: *m/z* calculated for $C_{19}H_{22}N_5O_2^+$: 352.1768 [M + H]$^+$; found: 352.1764 (1.14 ppm error); **r.f.** 0.46 (70 % EtOAc / Hex).

6.2.8 Hydrogenation of second-generation alkynes



17a - e **18a - e**

Scheme 6.15: Hydrogenation of second-generation alkynes

The unsaturated starting material (**17a – e**) (100 – 150 mg, 1 eq) was dissolved in EtOH (~ 5.00 ml) and added to a high-pressure reaction vessel followed by the addition of Pd/C (10 % m/m). The mixture was then stirred under H_2 at 4 bar of pressure for 24 hours. After this the mixture was filtered through celite and the solvent evaporated under reduced pressure. Purification was achieved using silica gel column chromatography (10% MeOH / DCM) if required. The details of the reaction procedure are summarized in scheme 6.15, and the full list of products formed from this reaction is displayed in table 6.4.

Table 6.4: Full list of products from hydrogenation reaction in the synthesis of second-generation analogues

Product:	**Y:**	**X:**
18a	O	4-COOMe
18b	O	2-CH_2CH_2COOMe
18c	O	3,4,5-OMe
18d	NBoc	3,4,5-OMe
18e	NMe	4-COOMe

6.2.8.1 Methyl 4-[3-(2,4-diamino-6-cyclopropylpyrimidin-5-yl)propoxy]benzoate (**18a**)

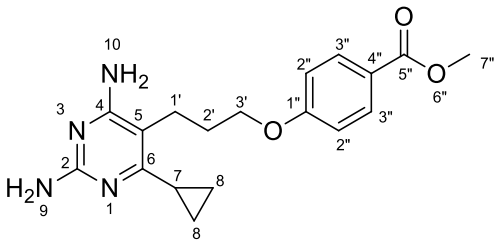


Figure 6.63: Methyl 4-[3-(2,4-diamino-6-cyclopropylpyrimidin-5-yl)propoxy]benzoate (**18a**)

Hydrogenation of **17a** (100 mg, 0.292 mmol) with palladium on carbon (10.0 mg, 10 % m/m) in ethanol (10.0 ml) afforded **18o** [Fig 6.63] as a white crystalline solid (94.5 mg, 94.5 %).

m.p. 94 – 96 °C; **^{1}H NMR** (Chloroform-*d*, 300 MHz): δ = 8.03 – 7.95 (H-3”, m, 2H), 6.95 – 6.87 (H-2”, m, 2H), 4.66 (H-10, brs, 2H), 4.44 (H-9, brs, 2H), 4.04 (H-3’, t, *J* = 5.6 Hz, 2H), 3.89 (H-7”, s, 3H), 2.75 (H-1’, t, *J* = 7.1 Hz, 2H), 2.06 (H-2’, q, *J* = 6.8 Hz, 2H), 1.93 (H-7, tt, *J* = 8.0, 4.7 Hz, 1H), 1.04 – 0.99 (H-8, m, 2H), 0.82 – 0.74 (H-8#, m, 2H); **^{13}C NMR** (Chloroform-*d*, 101 MHz): δ = 168.22 (C-6), 166.93 (C-5”), 162.51 (C-4), 162.31 (C-2), 161.03 (C-1”), 131.83 (C-3”), 123.05 (C-4”), 114.24 (C-2”), 104.00 (C-5), 66.63 (C-3’), 52.05 (C-7”), 28.48(C-2’), 21.24 (C-1’), 12.69 (C-7), 9.11 (C-8); **IR** (vmax / cm^{-1}): 3440 (N-H stretch), 3384 (N-H stretch), 3176 (N-H stretch), 2946 (C-H stretch), 1698 (C=O stretch), 1606 (C=C stretch), 1563 (C=C stretch), 1250 (C-O stretch); **HRMS(ESI)**: *m/z* calculated for $C_{18}H_{23}N_4O_3^+$: 343.1765 $[M + H]^+$; found: 343.1749 (4.66 ppm error); **r.f.** 0.18 (70 % EtOAc / Hex).

6.2.8.2 Methyl-3-{2-[3-(2,4-diamino-6-cyclopropylpyrimidin-5-yl) propoxy] phenyl} propanoate benzoate (**18b**)

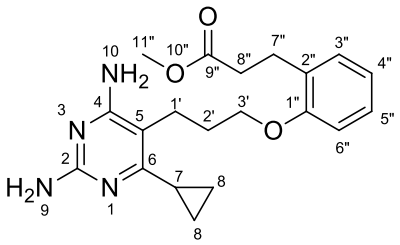


Figure 6.64: Methyl-3-{2-[3-(2,4-diamino-6-cyclopropylpyrimidin-5-yl) propoxy] phenyl} propanoate benzoate (**18b**)

Hydrogenation of **17b** (100 mg, 0.273 mmol) with palladium on carbon (10.0 mg, 10 % m/m) in ethanol (10.0 ml) afforded **18b** [Fig 6.64] as an off-white crystalline solid (87.2 mg, 86.2 %).

m.p. 98 – 101 °C; **^{1}H NMR** (Methanol-d_4, 400 MHz): δ = 7.21 – 7.15 (H-5”, m, 1H), 7.15 – 7.11 (H-3”, m, 1H), 6.92 (H-6”, d, *J* = 8.1 Hz, 1H), 6.86 (H-4”, t, *J* = 7.4 Hz, 1H), 4.09 (H-3’, t, *J* = 5.6 Hz, 2H), 3.64 (H-11”, s, 3H), 2.90 (H-7”, t, *J* = 7.7 Hz, 2H), 2.87 – 2.77 (H-8”, m, 2H), 2.59 (H-1’, t, *J* = 7.7 Hz, 2H), 2.17 – 2.09 (H-7, m, 1H), 2.08 – 1.99 (H-2’, m, 2H), 1.07 – 1.00 (H-8, m, 2H), 0.98 – 0.92 (H-8#, m, 2H); **^{13}C NMR** (Methanol-d_4, 101 MHz): δ = 175.35 (C-9”), 166.73 (C-6), 164.06 (C-4), 161.65 (C-2), 158.02 (C-1”), 130.81 (C-3”), 129.72 (C-2”), 128.79 (C-5”), 121.50 (C-4”), 112.26 (C-6”), 105.64 (C-5), 67.81 (C-3’), 52.04 (C-11”), 35.12 (C-8”), 29.59 (C-2’), 27.17 (C-7”), 22.38 (C-1’), 13.26 (C-7), 9.02 (C-8); **IR** (νmax / cm^{-1}): 3396 (N-H stretch), 3345 (N-H stretch), 2924 (C-H stretch), 1734 (C=O stretch), 1624 (C=N stretch), 1565 (C=C stretch), 1243 (C-O stretch); **HRMS(ESI)**: *m/z* calculated for $C_{19}H_{27}N_4O_4^+$: 371.2078 $[M + H]^+$; found: 371.2061 (4.58 ppm error); **r.f.** 0.38 (70 % EtOAc / Hex).

6.2.8.3 6-Cyclopropyl-5-[3-(3,4,5-trimethoxyphenoxy) propyl] pyrimidine-2,4-diamine (**18c**)

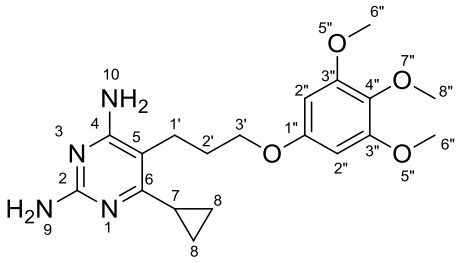


Figure 6.65: 6-Cyclopropyl-5-[3-(3,4,5-trimethoxyphenoxy) propyl] pyrimidine-2,4-diamine (**18c**)

Hydrogenation of **17c** (100 mg, 0.270 mmol) with palladium on carbon (10.0 mg, 10 % m/m) in ethanol (10.0 ml) afforded **18c** [Fig 6.65] as an off-white crystalline solid (86.0 mg, 85.1 %).

m.p. 143 – 145 °C; **^{1}H NMR** (Chloroform-*d*, 400 MHz): δ = 6.13 (H-2", s, 2H), 4.87 (H-10, brs, 2H), 4.61 (H-9, brs, 2H), 3.93 (H-3', t, *J* = 5.6 Hz, 2H), 3.82 (H-6", s, 6H), 3.77 (H-8", s, 3H), 2.72 (H-1', t, *J* = 7.0 Hz, 2H), 2.05 – 1.97 (H-2', m, 2H), 1.96 – 1.89 (H-7, m, 1H), 1.04 – 0.99 (H-8, m, 2H), 0.84 – 0.76 (H-8#, m, 2H); **^{13}C NMR** (Chloroform-*d*, 101 MHz): δ = 167.97 (C-6), 162.56 (C-4), 161.05 (C-2), 155.32 (C-1"), 153.86 (C-3"), 132.59 (C-4"), 104.07 (C-5), 92.35 (C-2"), 66.65 (C-3'), 61.13 (C-8"), 56.19 (C-6"), 28.67 (C-2'), 21.17 (C-1'), 12.65 (C-7), 9.02 (C-8); **IR** (vmax / cm^{-1}): 3502 (N-H stretch), 3451 (N-H stretch), 3384 (N-H stretch), 3351 (N-H stretch), 3006 (C-H stretch), 2936 (C-H stretch), 1607 (C=N stretch), 1560 (C=N stretch), 1446 (C=C stretch), 1424 (C=C stretch), 1225 (C-O stretch), 1121 (C-O stretch); **HRMS(ESI)**: *m/z* calculated for $C_{19}H_{27}N_4O_4^+$: 375.2027 $[M + H]^+$; found: 375.2037 (2.67 ppm error); **r.f.** 0.30 (70 % EtOAc / Hex).

6.2.8.4 *tert*-Butyl (3-(2,4-diamino-6-cyclopropylpyrimidin-5-yl) propyl) (3,4,5-trimethoxyphenyl) carbamate (**18d**)

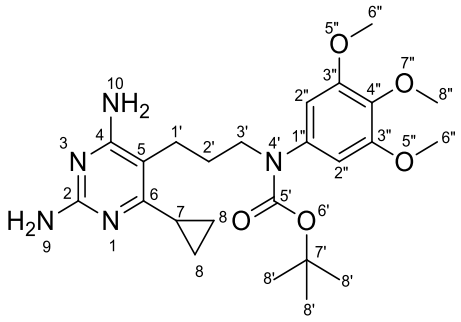


Figure 6.66: *tert*-Butyl (3-(2,4-diamino-6-cyclopropylpyrimidin-5-yl) propyl) (3,4,5-trimethoxyphenyl) carbamate (**18d**)

Hydrogenation of **17d** (150 mg, 0.319 mmol) with palladium on carbon (15.0 mg, 10 % m/m) in ethanol (15.0 ml) afforded **18d** [Fig 6.66] as a white crystalline solid (138 mg, 91.3 %).

m.p. 100 – 102 °C**; ^{1}H NMR** (400 MHz, Methanol-d_4): δ = 6.52 (H-2”, s, 1H), 3.80 (H-6”, s, 6H), 3.75 (H-8”, s, 3H), 3.73 – 3.68 (H-3’, m, 2H), 2.55 – 2.49 (H-1’, m, 2H), 1.89 (H-7, tt, *J* = 8.1, 4.8 Hz, 1H), 1.78 (H-2’, *J* = 7.6 Hz, 2H), 1.43 (H-8’, s, 9H), 0.96 – 0.88 (H-8, m, 2H), 0.83 – 0.76 (H-8#, m, 2H); **^{13}C NMR** (101 MHz, Methanol-d_4): δ = 167.98 (C-6), 163.71 (C-4), 162.39 (C-2), 156.58 (C-3”), 154.50 (C-5’) , 139.50 (C-1”), 137.73 (C-4”), 106.23 (C-2”), 105.38 (C-5), 81.72 (C-7’), 61.14 (C-8”), 56.75 (C-6”), 50.95 (C-3’), 28.92 (C-2’), 28.65 (C-8’), 23.05 (C-1’), 13.48 (C-7), 9.05 (C-8); **IR** (vmax / cm^{-1}): 3478 (N-H stretch), 3366 (N-H stretch), 3195 (N-H stretch), 2969 (C-H stretch), 2934 (C-H stretch), 1686 (C=O stretch), 1592 (C=N stretch), 1563 (C=C stretch), 1229 (C-O stretch), 1122 (C-O stretch); **HRMS(ESI)**: *m/z* calculated for $C_{24}H_{36}N_5O_5^+$: 474.2711 $[M + H]^+$; found: 474.2708 (0.63 ppm error); **r.f.** 0.28 (70 % EtOAc / Hex).

6.2.8.5 Methyl-4-{[3-(2,4-diamino-6-cyclopropylpyrimidin-5-yl) propyl] (methyl)amino} benzoate (**18e**)

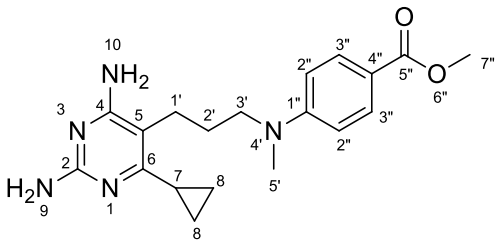
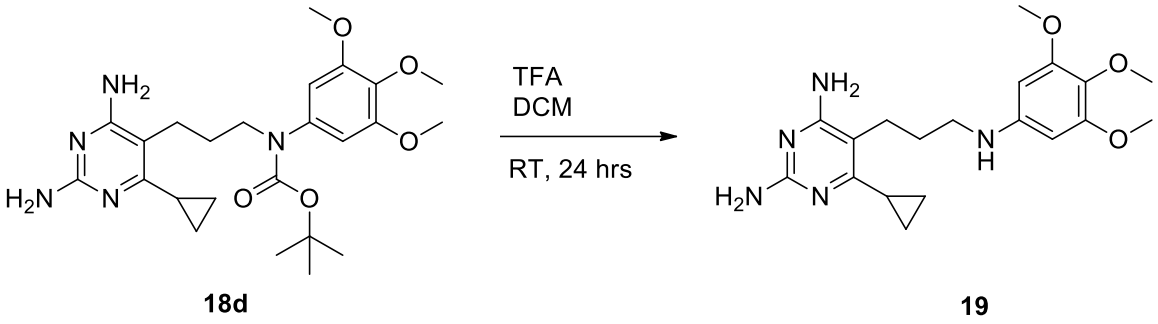


Figure 6.67: Methyl-4-{[3-(2,4-diamino-6-cyclopropylpyrimidin-5-yl)propyl](methyl)amino} benzoate (**18e**)

Hydrogenation of **17e** (150 mg, 0.427 mmol) with palladium on carbon (15.0 mg, 10 % m/m) in ethanol (15.0 ml) afforded **18e** [Fig 6.67] as a white solid (143 mg, 94.4 %).

m.p. 144 – 145 °C; **^{1}H NMR** (400 MHz, Methanol-d_4): δ = 7.81 (H-3", d, *J* = 9.0 Hz, 2H), 6.69 (H-2", d, *J* = 9.1 Hz, 2H), 3.82 (H-7", s, 3H), 3.57 – 3.41 (H-3', m, 2H), 3.04 (H-5', s, 3H), 2.64 – 2.47 (H-1', m, 2H), 1.92 (H-7, tt, *J* = 8.2, 4.8 Hz, 1H), 1.81 (H-2', q, *J* = 7.8 Hz, 2H), 0.95 – 0.89 (H-8, m, 2H), 0.82 – 0.75 (H-8#, m, 2H); **^{13}C NMR** (101 MHz, DMSO-d6): δ = 166.38 (C-6), 165.05 (C-5"), 162.25 (C-4), 161.12 (C-2), 152.26 (C-1"), 130.81 (C-3"), 115.28 (C-4"), 110.63 (C-2"), 102.88 (C-5), 51.26 (C-3'), 51.01 (C-7"), 37.90 (C-5'), 25.79 (C-2'), 21.55 (C-1'), 12.07 (C-7), 8.12 (C-8); **IR** (vmax / cm^{-1}): 3497 (N-H stretch), 3454 (N-H stretch), 3393 (N-H stretch), 3393 (N-H stretch), 3128 (N-H stretch), 2947 (C-H stretch), 2877 (C-H stretch), 1680 (C=O stretch), 1600 (C=N stretch), 1564 (C=C stretch), 1284 (C-O stretch); **HRMS(ESI)**: *m/z* calculated for $C_{19}H_{26}N_5O_2^+$: 356.2081 $[M + H]^+$; found: 356.2075 (1.68 ppm error); **r.f.** 0.39 (70 % EtOAc / Hex).

6.2.9 Removal of Boc group after Sonogashira and reduction



Scheme 6.16: Removal of Boc group after Sonogashira and reduction

Starting material **18d** (100 mg, 0.211 mmol, 1 eq) was dissolved in DCM (5.00 ml) and cooled to 0 °C in an ice bath while under Ar. TFA (121 mg, 1.06 mmol, 5 eq) was then added slowly over 10 minutes, and the mixture stirred at room temperature for 24 hours. The reaction was monitored by TLC. After all the starting material was consumed, the reaction was quenched with brine (20.0 ml), and the organic layer isolated. This was washed sequentially with saturated aqueous NH_4Cl (20.0 ml) and brine (20.0 ml). The organic layer was then dried with $MgSO_4$ and filtered through celite. The solvent was removed using a rotary evaporator, and the product purified using silica gel column chromatography (20 % EtOAc / Hex) to afford product **19** [Fig 6.68] as an orange solid (63.0 mg, 80.0 %). The details of the reaction conditions are summarized in scheme 6.16.

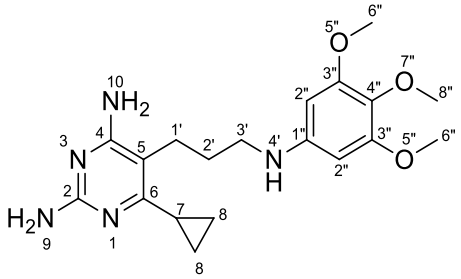
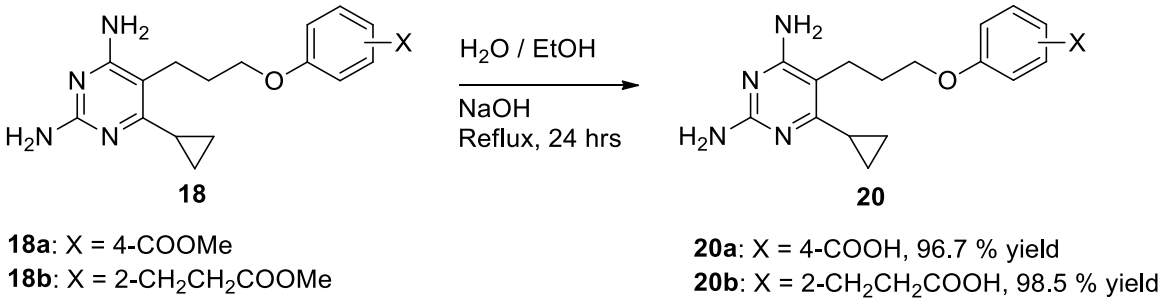


Figure 6.68: 6-Cyclopropyl-5-{3-[(3,4,5-trimethoxyphenyl)amino]propyl}pyrimidine-2,4-diamine (**19**)

m.p. 127 – 129 °C; **^{1}H NMR** (300 MHz, Chloroform-*d*): δ = 5.87 (H-2", s, 2H), 4.96 (H-10, brs, 2H), 4.56 (H-9, brs, 2H), 3.80 (H-6", s, 6H), 3.75 (H-8", s, 3H), 3.38 (H-4', brs, 1H), 3.10 (H-3', t, *J* = 6.3 Hz, 2H), 2.67 (H-1', t, *J* = 6.9 Hz, 2H), 1.97 – 1.83 (H-7 and H-2', m, 3H), 1.10 – 0.97 (H-8, m, 2H), 0.89 – 0.76 (H-8#, m, 2H); **^{13}C NMR** (75 MHz, Chloroform-*d*): δ = 167.84, 162.86, 161.03, 154.05, 145.09, 130.79, 104.42, 91.27, 61.17, 56.07, 43.35, 29.11, 22.11, 12.67, 9.00; **IR** (vmax / cm^{-1}): 3425 (N-H stretch), 3323 (N-H stretch), 3222 (N-H stretch), 2931 (C-H stretch), 2848 (C-H stretch), 1610 (C=N stretch), 1562 (C=C stretch); **HRMS(ESI)**: *m/z* calculated for $C_{19}H_{28}N_5O_3^+$: 374.2187 $[M + H]^+$; found: 374.2181 (1.60 ppm error); **r.f.** 0.25 (10 % EtOH / EtOAc).

6.2.10 Hydrolysis of esters after Sonogashira and reduction



Scheme 6.17: Hydrolysis of esters after Sonogashira and reduction

The carboxylic ester starting material (**18a** or **18b**) (100 mg, 1 eq) was dissolved in EtOH (10.0 ml). A separate solution containing NaOH (2 eq) and deionized water (10.0 ml) was prepared and added. The mixture was heated to reflux for 24 hours under argon. The reaction mixture was allowed to cool, and the pH adjusted to 8 using 10 % HCl. The EtOH from the reaction mixture was removed on a rotary evaporator and the product crashed out as a white solid. This was filtered off and washed with 3 portions of deionized water (5.00 ml). The details of the reaction conditions are summarized in scheme 6.17.

6.2.10.1 4-[3-(2,4-Diamino-6-cyclopropylpyrimidin-5-yl)propoxy] benzoic acid (**20a**)

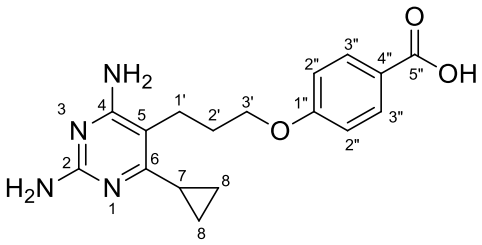


Figure 6.69: 4-[3-(2,4-Diamino-6-cyclopropylpyrimidin-5-yl)propoxy] benzoic acid (**20a**)

Hydrolysis of **18a** (100 mg, 0.292 mmol, 1 eq) with sodium hydroxide (23.4 mg, 0.584 mmol, 2 eq) in a 1:1 water – ethanol solvent mixture (20.0 ml) provided **20a** [Fig 6.69] as a white crystalline solid after filtering (92.8 mg, 96.7 %).

m.p. 303 – 306 °C; **^{1}H NMR** (Methanol-d_4, 500 MHz): δ = 7.97 (H-3", d, *J* = 8.8 Hz, 2H), 6.98 (H-2", d, *J* = 8.8 Hz, 2H), 4.14 (H-3', t, *J* = 5.9 Hz, 2H), 2.83 (H-1', t, *J* = 7.5 Hz, 2H), 2.18 – 2.11 (H-7, m, 1H), 2.05 (H-2', q, *J* = 6.4 Hz, 2H), 1.10 – 1.04 (H-8, m, 2H), 1.02 – 0.97 (H-8#, m, 2H); **^{13}C NMR** (DMSO-d_6, 101 MHz): δ = 167.17 (C-6), 165.09 (C-5"), 162.33 (C-4), 162.21 (C-2), 161.06 (C-1"), 131.34 (C-3"), 123.10 (C-4"), 114.23 (C-2"), 102.51 (C-5), 67.00 (C-3'), 28.24 (C-2'), 20.37 (C-1'), 11.99 (C-7), 8.18 (C-8); **IR** (vmax / cm^{-1}): 3443 (N-H stretch), 3299 (O-H stretch), 2940 (C-H stretch), 1678 (C=O stretch), 1643 (C=N stretch), 1627 (C=C stretch), 1251 (C-O stretch); **HRMS(ESI)**: *m/z* calculated for $C_{17}H_{20}N_4O_3^+$: 329.1608 $[M + H]^+$; found: 329.1586 (6.68 ppm error); **r.f.** 0.12 (10 % MeOH / DCM).

6.2.10.2 3-{2-[3-(2,4-Diamino-6-cyclopropylpyrimidin-5-yl) propoxy]phenyl} propanoic acid (**20b**)

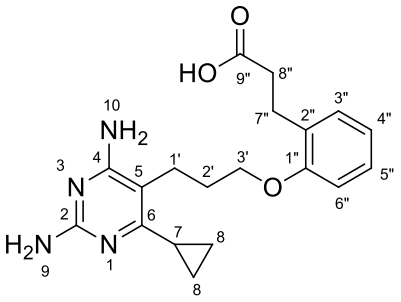


Figure 6.70: 3-{2-[3-(2,4-Diamino-6-cyclopropylpyrimidin-5-yl) propoxy]phenyl} propanoic acid (**20b**)

Hydrolysis of **18b** (100 mg, 0.270 mmol, 1 eq) with sodium hydroxide (21.6 mg, 0.540 mmol, 2 eq) in a 1:1 water – ethanol solvent mixture (20.0 ml) provided **20b** [Fig 6.70] as an off-white solid after filtering (94.8 mg, 98.5 %).

m.p. 135 – 136 °C; **^{1}H NMR** (DMSO-d_6, 400 MHz): δ = 7.20 – 7.10 (H-3” and H-5”, m, 2H), 6.91 (H-6”, d, *J* = 8.1 Hz, 1H), 6.83 (H-4”, t, *J* = 7.4 Hz, 1H), 5.98 (H-10, brs, 2H), 5.42 (H-9, brs, 2H), 4.01 (H-3’, t, *J* = 6.0 Hz, 2H), 2.78 (H-7”, t, *J* = 7.8 Hz, 2H), 2.63 (H-8”, t, *J* = 7.7 Hz, 2H), 2.48 – 2.43 (H-1’, m, 2H), 1.94 (H-7, tt, *J* = 8.4, 4.7 Hz, 1H), 1.85 (H-2’, q, *J* = 6.3 Hz, 2H), 0.83 – 0.77 (H-8, m, 2H), 0.66 – 0.61 (H-8#, m, 2H); **^{13}C NMR** (DMSO-d_6, 101 MHz): δ = 174.43 (C-9”), 165.75 (C-6), 162.73 (C-4), 161.67 (C-2), 156.85 (C-1”), 129.98 (C-3”), 128.97 (C-2”), 127.93 (C-5”), 120.52 (C-4”), 111.60 (C-6”), 103.13 (C-5), 67.11 (C-3’), 34.18 (C-8”); **IR** (νmax / cm^{-1}): 3326 (O-H stretch), 3171 (N-H stretch), 3119 (N-H stretch), 2927 (C-H stretch), 1651 (C=O stretch), 1626 (C=N stretch), 1535 (C=C stretch), 1234 (C-O stretch); **HRMS(ESI)**: *m/z* calculated for $C_{19}H_{25}N_4O_3^+$: 357.1921 [M + H]$^+$; found: 357.1915 (1.68 ppm error); **r.f.** 0.16 (10 % MeOH / DCM).

6.2.11 Protection of glutamic acid



Scheme 6.18: Protection of glutamic acid

L-glutamic acid HCl (500 mg, 2.72 mmol, 1 eq) was dissolved in MeOH (15 ml) and cooled to 0 °C in an ice bath under argon. $SOCl_2$ (1293 mg, 10.9 mmol, 4 eq) was then added slowly, and the mixture stirred for 24 hours at room temperature. Reaction progress was monitored by TLC, and after all starting material had been consumed the reaction mixture was concentrated on a rotary evaporator at room temperature. The residue was dissolved in EtOAc (100 ml) and washed with saturated aqueous $NaHCO_3$ (100 ml) and brine (100 ml). The organic layer was isolated and dried using MgSO4 before filtering through celite and removing the solvent on a rotary evaporator at room temperature. Purification was achieved by silica column chromatography (100 % EtOAc) to afford the product **23** [Fig 6.71] as a pale-yellow oil in high yield (460 mg, 96.5 %). A summary of the described reaction conditions is included in scheme 6.18.

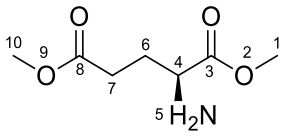
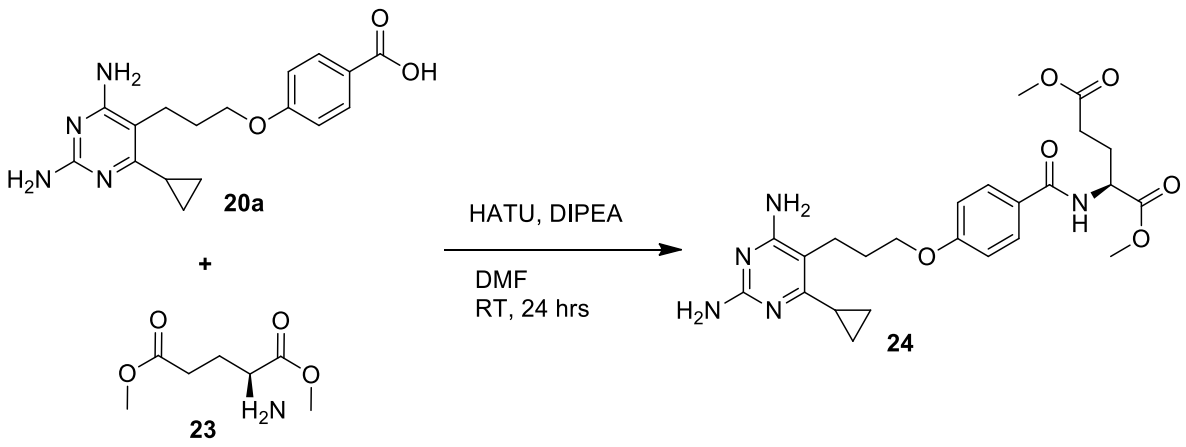


Figure 6.71: (*S*)-dimethyl 2-aminopentanedioate (**23**)

^{1}H NMR (300 MHz, Chloroform-*d*): δ = 3.72 (H-1, s, 3H), 3.67 (H-10, s, 3H), 3.48 (H-4, dd, *J* = 8.3, 5.3 Hz, 1H), 2.47 (H-7, t, *J* = 7.4 Hz, 2H), 2.16 – 2.00 (H-6, m, 1H), 1.84 (H-6#, dq, *J* = 14.7, 7.7 Hz, 1H), 1.52 (H-5, brs, 2H); **^{13}C NMR** (101 MHz, Chloroform-*d*): δ = 176.04 (C-8), 173.57 (C-3), 53.69 (C-4), 52.09 (C-1), 51.68 (C-10), 30.35 (C-7), 29.67 (C-6); **IR** (ν_{max} / cm^{-1}); 3253 (N-

H stretch), 2923 (C-H stretch), 2853 (C-H stretch), 1644 (C=O stretch), 1205 (C-O stretch); **r.f.** 0.34 (100 % EtOAc).

6.2.12 Addition of glutamic acid



Scheme 6.19: Addition of glutamic acid

The carboxylic acid starting material **20a** (100 mg, 0.305 mmol, 1 eq) was dissolved in DMF (5 ml). To this was added DIPEA (158 mg, 1.22 mmol, 4 eq), followed by HATU (232 mg, 0.610 mmol, 2 eq), and the mixture stirred at room temperature under an inert atmosphere for 30 minutes. After this time, dimethyl glutamic acid **23** (69.5 mg, 0.397 mmol, 1.3 eq) was added and the reaction stirred at room temperature for 24 hours. Reaction progress was monitored by TLC, using Hanessian stain which turns the amide product blue upon heating. Once all the starting material had been consumed, the reaction mixture was quenched with brine (100 ml) and the product extracted into EtOAc (50 ml). The organic layer was washed with additional portions of brine (3 x 50 ml) and dried with $MgSO_4$. The solvent was removed on a rotary evaporator and the product was purified with silica gel column chromatography (10 % EtOH / EtOAc) to afford product **24** [Fig 6.72] as a white solid in high yield (140 mg, 94.5 %). A summary of the reaction conditions described is included in scheme 6.19.

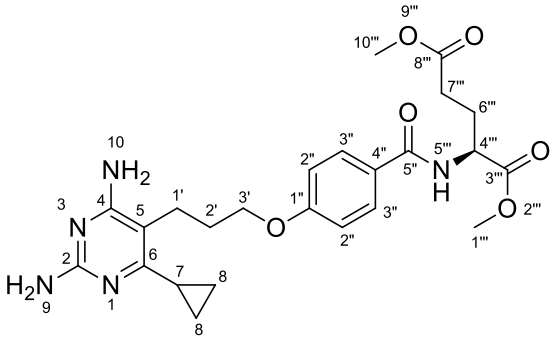
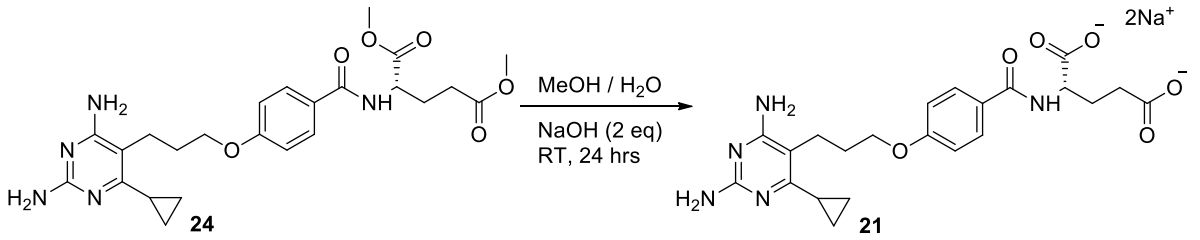


Figure 6.72: (*S*)-Dimethyl 2-{4-[3-(2,4-diamino-6-cyclopropylpyrimidin-5-yl)propoxy] benzamido}pentanedioate (**24**)

m.p. 87 – 88 °C; **^{1}H NMR** (300 MHz, Chloroform-*d*): δ = 7.77 (H-3”, d, *J* = 8.7 Hz, 2H), 7.01 (H-5’”, d, *J* = 7.5 Hz, 1H), 6.90 (H-2”, d, *J* = 8.8 Hz, 2H), 4.86 – 4.73 (H-10, brs, 2H), 4.79 (H-4”’, dd, 1H), 4.56 (H-9, brs, 2H), 4.01 (H-3’, t, *J* = 5.6 Hz, 2H), 3.77 (H-1”’, s, 3H), 3.64 (H-10”’, s, 3H), 2.72 (H-1’, t, *J* = 7.1 Hz, 2H), 2.55 – 2.39 (H-7”’, m, 2H), 2.37 – 2.24 (H-6”’, m, 1H), 2.21 – 2.08 (H-6”’#, m, 1H), 2.03 (H-2’, q, *J* = 6.0 Hz, 1H), 1.92 (H-7, tt, *J* = 8.2, 4.8 Hz, 1H), 1.05 – 0.94 (H-8, m, 2H), 0.83 – 0.72 (H-8#, m, 2H); **^{13}C NMR** (75 MHz, Chloroform-*d*): δ = 173.88 (C-8”’), 172.74 (C-3”’), 167.78 (C-6), 166.75 (C-5”), 162.45 (C-4), 161.65 (C-2), 160.98 (C-1”), 129.20 (C-3”), 126.25 (C-4”), 114.40 (C-2”), 104.06 (C-5), 66.63 (C-3’), 52.73 (C-4”’), 52.40 (C-1”’), 52.05 (C-10”’), 30.40 (C-7”’), 28.46 (C-6”’), 27.30 (C-2’), 21.22 (C-1’), 12.63 (C-7), 9.02 (C-8); **IR** (vmax / cm^{-1}): 3476 (N-H stretch), 3364 (N-H stretch), 3201 (N-H stretch), 2949 (C-H stretch), 1731 (C=O stretch), 1606 (C=N stretch), 1562 (C=C stretch), 1249 (C-O stretch), 1173 (C-O stretch); **HRMS(ESI)**: *m/z* calculated $C_{24}H_{32}N_5O_6^+$: 486.2347 $[M + H]^+$; found: 486.2342 (1.03 ppm error); **r.f.** 0.24 (10 % EtOH / EtOAc).

6.2.13 Deprotection of glutamic acid



Scheme 6.20: Deprotection of glutamic acid

The dimethyl glutamate starting material **24** (100 mg, 0.206 mmol, 1 eq) was dissolved in a mixture of MeOH (3 ml) and H_2O (6 ml) and NaOH (16.5 mg, 0.412 mmol, 2 eq) was added. The mixture was left to stir under an inert atmosphere at room temperature for 24 hours. After this period, TLC analysis revealed complete consumption of the starting material. The solvent was removed on a rotary evaporator, and the residue recrystallized from a mixture of water and acetone, to afford the desired product **21** [Fig 6.73] as a white solid in good yield (89.0 mg, 94.5 %). The reaction conditions are provided in scheme 6.20.

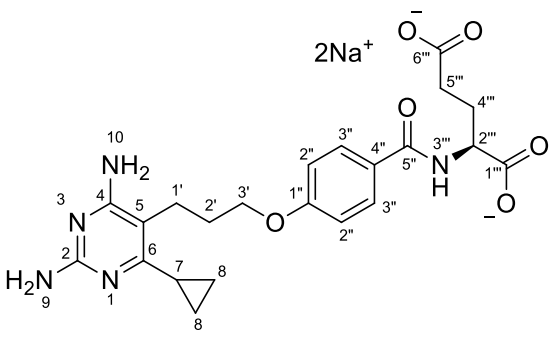


Figure 6.73: (S)-2-{4-[3-(2,4-diamino-6-cyclopropylpyrimidin-5-yl)propoxy]benzamido} pentanedioic acid (**21**)

m.p. 273 – 275 °C; **^{1}H NMR** (400 MHz, Deuterium Oxide): δ = 7.74 (H-3”, d, *J* = 8.8 Hz, 2H), 6.94 (H-2”, d, *J* = 8.9 Hz, 2H), 4.28 (H-2”’, dd, *J* = 8.9, 4.5 Hz, 1H), 4.07 (H-3’, t, *J* = 5.7 Hz, 2H), 2.67 (H-1’, t, *J* = 7.1 Hz, 1H), 2.28 (H-5”’, m, 2H), 2.12 (H4”’, m, 1H), 1.96 (H2’ and H7 and H4”’#, m,

4H), 0.73 (H-8, s, 2H), 0.71 (H-8#, s, 2H). **^{13}C NMR** (101 MHz, Deuterium Oxide): δ = 182.32 (C-6"'), 179.03 (C-1"'), 169.57 (C-6), 167.63 (C-5'), 162.47 (C-4), 161.08 (C-2), 160.75 (C-1"), 129.16 (C-3"), 125.94 (C-4"), 114.50 (C-2"), 105.36 (C-5), 67.32 (C-3'), 55.97 (H-2"'), 34.29 (C-5"'), 28.43 (C-4"'), 26.86 (C-2'), 20.66 (C-1'), 12.57 (C-7), 8.18 (C-8); **HRMS(ESI)**: *m/z* calculated $C_{22}H_{28}N_5O_6^+$: 458.2034 $[M + H]^+$; found: 458.2028 (1.31 ppm error) and $C_{22}H_{26}N_5Na_2O_6^+$: 502.1673 $[M + 2Na - H]^+$; found: 502.1667 (1.19 ppm error); **r.f.** 0.09 (10 % EtOH / EtOAc).

References

(**1**) K. Rajakumar and M. Weisse, *South Med J*, 1999, **92**, 567–571.

(**2**) V. Lorenz and P. Karanis, *Asian Pac J Trop Biomed*, 2011, **1**, 74–78.

(**3**) N. D. Salinas, W. K. Tang and N. H. Tolia, *J Mol Biol*, 2019, **431**, 4259–4280.

(**4**) R. P. Man, *Russell P. Man's Mastery of Malaria*, Oxford Univ Press, Oxford, UK, 1955.

(**5**) N. M. Anstey, M. J. Grigg, G. S. Rajahram, D. J. Cooper, T. William, S. Kho and B. E. Barber, *Adv Parasitol*, 2021, **113**, pp. 1–43.

(**6**) WHO world malaria report. World Health Organization, 2023. https://www.who.int/publications/i/item/9789240086173 Accessed June 2024.

(**7**) A. S. P. Yang, D. Dutta, K. Kretzschmar, D. Hendriks, J. Puschhof, H. Hu, K. E. Boonekamp, Y. van Waardenburg, S. M. Chuva de Sousa Lopes, G.-J. van Gemert, J. H. W. de Wilt, T. Bousema, H. Clevers and R. W. Sauerwein, *Nat Commun*, 2023, **14**, 4631.

(**8**) R. Ménard, *Microbes Infect*, 2000, **2**, 633–642.

(**9**) *Saving Lives, Buying Time: Economics of Malaria Drugs in an Age of Resistance*, ed. K.J. Arrow, C. Panosian, H. Gelband, National Academies Press, Washington (DC), 2004

(**10**) M. Dagen, in *Antimalarial Agents*, Elsevier, 2020, pp. 1–48.

(**11**) J. Achan, A. O. Talisuna, A. Erhart, A. Yeka, J. K. Tibenderana, F. N. Baliraine, P. J. Rosenthal and U. D'Alessandro, *Malar J*, 2011, **10**, 144.

(**12**) M. Foley and L. Tilley, *Int J Parasitol*, 1997, **27**, 231–240.

(**13**) A. Gregson and C. V. Plowe, *Pharmacol Rev*, 2005, **57**, 117–145.

(**14**) A. Nzila, *Journal of Antimicrobial Chemotherapy*, 2006, **57**, 1043–1054.

(**15**) C.H. Sibley, J.E. Hyde, P.F. Sims, C.V. Plowe, J.G. Kublin, E.K. Mberu, A.F. Cowman, P.A. Winstanley, W.M. Watkins and A.M. Nzila, *Trends Parasitol*, 2001, **17**, 582-588.

(**16**) A. Nzila, S. A. Ward, K. Marsh, P. F. G. Sims and J. E. Hyde, *Trends Parasitol*, 2005, **21**, 292–298

(**17**) L. Foquet, C. Schafer, N. K. Minkah, D. G. W. Alanine, E. L. Flannery, R. W. J. Steel, B. K. Sack, N. Camargo, M. Fishbaugher, W. Betz, T. Nguyen, Z. P. Billman, E. M. Wilson, J. Bial, S. C. Murphy, S. J. Draper, S. A. Mikolajczak and S. H. I. Kappe, *Front Immunol*, 2018.

(**18**) S. R. Meshnick, *Med Trop (Mars)*, 1998, **58**, 13–7.

(**19**) S. Krishna, A. Uhlemann and R. Haynes, *Drug Resistance Updates*, 2004, **7**, 233–244.

(**20**) C. Rasmussen, P. Alonso and P. Ringwald, *Expert Rev Anti Infect Ther*, 2022, **20**, 353–372.

(**21**) A. Saleh, S. Qamar, A. Tekin, R. Singh and R. Kashyap, *Cureus*, 2021, **13**.

(**22**) A. A. El-Moamly and M. A. El-Sweify, *Trop Med Health*, 2023, **51**, 29.

(**23**) D. I. Stanisic and M. F. Good, *BioDrugs*, 2023, **37**, 737–756.

(**24**) P. Jagannathan and A. Kakuru, *Nat Commun*, 2022, **13**, 2678.

(**25**) K. Raghavendra, T. K. Barik, B. P. N. Reddy, P. Sharma and A. P. Dash, *Parasitol Res*, 2011, **108**, 757–779.

(**26**) K. S. Silver, Y. Du, Y. Nomura, E. E. Oliveira, V. L. Salgado, B. S. Zhorov and K. Dong, *Adv Insect Physiol*, 2014, **46** pp. 389–433.

(**27**) S. S. Sahu, S. Thankachy, S. Dash, K. Nallan, S. Swaminathan, G. Kasinathan and J. Purushothaman, *Malar J*, 2020, **19**, 19.

(**28**) G. Benelli and J. C. Beier, *Acta Trop*, 2017, **174**, 91–96.

(**29**) T. E. Nkya, U. Fillinger, O. P. Sangoro, R. Marubu, E. Chanda and C. M. Mutero, *Malar J*, 2022, **21**, 279.

(**30**) K. A. McGlynn, C. C. Abnet, M. Zhang, X.-D. Sun, J.-H. Fan, T. R. O'Brien, W.-Q. Wei, B. A. Ortiz-Conde, S. M. Dawsey, J.-P. Weber, P. R. Taylor, H. Katki, S. D. Mark and Y.-L. Qiao, *JNCI*, 2006, **98**, 1005–1010.

(**31**) L. B. Bailey and J. F. Gregory, *J Nutr*, 1999, **129**, 779–782.

(**32**) B. Shane, *Food Nutr Bull*, 2008, **29**, S5–S16.

(**33**) I. B. Müller and J. E. Hyde, *Mol Biochem Parasitol*, 2013, **188**, 63–77.

(**34**) N. Gonen and Y. G. Assaraf, *Drug Resist Updat*, 2012, **15**, 183–210.

(**35**) M. Visentin, R. Zhao and I. D. Goldman, *Hematol Oncol Clin North Am*, 2012, **26**, 629–648.

(**36**) J. McGuire, *Curr Pharm Des*, 2003, **9**, 2593–2613.

(**37**) A. Estrada, D. L. Wright and A. C. Anderson, *Cold Spring Harb Perspect Med*, 2016, **6**, a028324.

(**38**) I. S. Kovalev, G. V. Zyryanov, S. Santra, A. Majee, M. V. Varaksin and V. N. Charushin, *Molecules*, 2022, **27**, 6229.

(**39**) W. Zhou, E. W. Scocchera, D. L. Wright and A. C. Anderson, *Medchemcomm*, 2013, **4**, 908.

(**40**) A. C. U. Lourens, D. Gravestock, R. L. van Zyl, H. C. Hoppe, N. Kolesnikova, S. Taweechai, Y. Yuthavong, S. Kamchonwongpaisan and A. L. Rousseau, *Org Biomol Chem*, 2016, **14**, 7899–7911.

(**41**) Y. Yuthavong, B. Tarnchompoo, T. Vilaivan, P. Chitnumsub, S. Kamchonwongpaisan, S. A. Charman, D. N. McLennan, K. L. White, L. Vivas, E. Bongard, C. Thongphanchang, S. Taweechai, J. Vanichtanankul, R. Rattanajak, U. Arwon, P. Fantauzzi, J. Yuvaniyama, W. N. Charman and D. Matthews, *Proc Natl Acad Sci U.S.A*, 2012, **109**, 16823–16828.

(**42**) M. F. Chughlay, M. El Gaaloul, C. Donini, B. Campo, P.-J. Berghmans, A. Lucardie, M. W. Marx, M. H. Cherkaoui-Rbati, G. Langdon, I. Angulo-Barturen, S. Viera, A. Rosanas-Urgell, J.-P. Van Geertruyden and S. Chalon, *Am J Trop Med Hyg*, 2021, **104**, 1348–1358.

(**43**) J. N. Burrows, S. Duparc, W. E. Gutteridge, R. H. van Huijsduijnen, W. Kaszubska, F. Macintyre, S. Mazzuri, J. J. Möhrle, T. N. C. Wells, *Malar J*, 2017, **16**, 26.

(**44**) T. D. Seanego, H. F. Klein, N. C. J. van Vuuren, R. L. van Zyl and A. L. Rousseau, *Arkivoc*, 2020, **2020**, 1–16.

(**45**) **Schrödinger Release 2023-2**: Maestro, Schrödinger, LLC, New York, NY, 2024

(**46**) **Schrödinger Release 2024-2**: Maestro, Schrödinger, LLC, New York, NY, 2024

(**47**) G. Madhavi Sastry, M. Adzhigirey, T. Day, R. Annabhimoju and W. Sherman, *J Comput Aided Mol Des*, 2013, **27**, 221–234.

(**48**) **Schrödinger Release 2024-2**: Protein Preparation Wizard; Epik, Schrödinger, LLC, New York, NY, 2024; Impact, Schrödinger, LLC, New York, NY; Prime, Schrödinger, LLC, New York, NY, 2024.

(**49**) C. Lu, C. Wu, D. Ghoreishi, W. Chen, L. Wang, W. Damm, G. A. Ross, M. K. Dahlgren, E. Russell, C. D. Von Bargen, R. Abel, R. A. Friesner, E. D. Harder, *J Chem Theory Comput*, 2021, **17** (7), 4291–4300.

(**50**) **Schrödinger Release 2024-2**: LigPrep, Schrödinger, LLC, New York, NY, 2024.

(**51**) R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis and P. S. Shenkin, *J Med Chem*, 2004, **47**, 1739–1749.

(**52**) T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard and J. L. Banks, *J Med Chem*, 2004, **47**, 1750–1759.

(**53**) R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin and D. T. Mainz, *J Med Chem*, 2006, **49**, 6177–6196.

(**54**) **Schrödinger Release 2024-2**: Glide, Schrödinger, LLC, New York, NY, 2024.

(**55**) Y. Yang, K. Yao, M. P. Repasky, K. Leswing, R. Abel, B. K. Shoichet and S. V. Jerome, *J Chem Theory Comput*, 2021, **17**, 7106–7119.

(**56**) W. Sherman, T. Day, M. P. Jacobson, R. A. Friesner and R. Farid, *J Med Chem*, 2006, **49**, 534–553.

(**57**) W. Sherman, H. S. Beard and R. Farid, *Chem Biol Drug Des*, 2006, **67**, 83–84.

(**58**) **Schrödinger Release 2024-2**: Induced Fit Docking protocol; Glide, Schrödinger, LLC, New York, NY, 2024; Prime, Schrödinger, LLC, New York, NY, 2024.

(**59**) **Schrödinger Release 2024-2**: Desmond Molecular Dynamics System, D. E. Shaw Research, New York, NY, 2024. Maestro-Desmond Interoperability Tools, Schrödinger, New York, NY, 2024.

(**60**) K. J. Bowers, E. Chow, H. Xu, R. O. Dror, M. P. Eastwood, B. A. Gregersen, J. L. Klepeis, I. Kolossvary, M. A. Moraes, F. D. Sacerdoti, J. K. Salmon, Y. Shan and D. E. Shaw, in *Proceedings of the ACM/IEEE Conference on Supercomputing (SC06)*, Tampa, Florida, 2006.

(**61**) **Schrödinger Release 2024-2**: QikProp, Schrödinger, LLC, New York, NY, 2024.

(**62**) The PyMOL Molecular Graphics System, Version 3.0.3 Schrödinger, LLC

(**63**) M. Hoarau, P. Sermmai, T. Varatthan, R. Thiabma, T. Jantra, R. Rattanajak, D. Vitsupakorn, J. Vanichtanankul, S. Saepua, Y. Yuthavong, C. Thongpanchang and S. Kamchonwongpaisan, *RSC Med Chem*, 2023, **14**, 1755–1766.

(**64**) J. Vanichtanankul, S. Taweechai, J. Yuvaniyama, T. Vilaivan, P. Chitnumsub, S. Kamchonwongpaisan and Y. Yuthavong, *ACS Chem Biol*, 2011, **6**, 905–911.

(**65**) B. Tarnchompoo, P. Chitnumsub, A. Jaruwat, P. J. Shaw, J. Vanichtanankul, S. Poen, R. Rattanajak, C. Wongsombat, A. Tonsomboon, S. Decharuangsilp, T. Anukunwithaya, U. Arwon, S. Kamchonwongpaisan and Y. Yuthavong, *ACS Med Chem Lett*, 2018, **9**, 1235–1240.

(**66**) J. Yuvaniyama, P. Chitnumsub, S. Kamchonwongpaisan, J. Vanichtanankul, W. Sirawaraporn, P. Taylor, M. D. Walkinshaw and Y. Yuthavong, *Nat Struct Mol Biol*, 2003, **10**, 357–365.

(**67**) S. Saepua, K. Sadorn, J. Vanichtanankul, T. Anukunwithaya, R. Rattanajak, D. Vitsupakorn, S. Kamchonwongpaisan, Y. Yuthavong and C. Thongpanchang, *Bioorg Med Chem*, 2019, **27**, 115158.

(**68**) S. Kamchonwongpaisan, N. Charoensetakul, C. Srisuwannaket, S. Taweechai, R. Rattanajak, J. Vanichtanankul, D. Vitsupakorn, U. Arwon, C. Thongpanchang, B. Tarnchompoo, T. Vilaivan and Y. Yuthavong, *Eur J Med Chem*, 2020, **195**, 112263.

(**69**) J. A. Erickson, M. Jalaie, D. H. Robertson, R. A. Lewis, M. Vieth, *J Med Chem*, 2004, **47** (1), 45–55.

(**70**) J. L. Banks, H. S. Beard, Y. Cao, A. E. Cho, W. Damm, R. Farid, A. K. Felts, T. A. Halgren, D. T. Mainz, J. R. Maple, R. Murphy, D. M. Philipp, M. P. Repasky, L. Y. Zhang, B. J. Berne, R. A. Friesner, E. Gallicchio, *J Comput Chem*, 2005, **26** (16), 1752–1780.

(**71**) K. Somandi, T. Seanego, T. Dlamini, M. Maree, C.B. de Koning, J. Vanichtanankul, R. Rattanajak, T. Saeyang, Y. Yuthavong, S. Kamchonwongpaisan, A.L. Rousseau, *ChemMedChem*, 2022, **17** (22).

(**72**) T. Batool, N. Rasool, Y. Gull, M. Noreen, F.H. Nasim, A. Yaqoob, M. Zubair, U. A. Rana, S.U.D. Khan, M. Zia-Ul-Haq, *PLoS ONE*, 2014, **9** (12).

(**73**) D.P. Pienaar, K.R. Butsi, A.L. Rousseau, D. Brady, D.A. Green, *Beilstein J Org Chem*, 2019, **15**, 2930–2935.

(**74**) M. A. Langarizadeh, A. Ameri, H. Forootanfar, *Med Chem Res*, 2023, **32**, 1943–1960.

(**75**) D. Serbezeanu, I.D. Carja, A. Nicolescu, M. Aflori, T. Vlad-Bubulac, M. Butnaru, R.D. Damian, S. Dunca, S. Shova, *J Mol Struct*, 2019, **1196**, 637–646.

(**76**) P. Kumar, S. Verma, K. Rathi, D. Chandra, V.P. Verma, J.L. Jat, *EurJOC*, 2022, **27**.

(**77**) T. Ishikawa, K. Nagai, N. Ohkubo, *Heterocycles*, 1994, **39** (1).

(**78**) D. Zhu, Z. Wu, B. Luo, Y. Du, P. Liu, Y. Chen, Y. Hu, P. Huang, S. Wen, *Org Lett*, 2018, **20** (16), 4815–4818

(**79**) Y. Ogiwara, H. Suzuki, H. Hori, N. Sakai, *Tetrahedron Lett*, 2017, **58** (1), 63–66.

(**80**) A.P. Ingale, S.V. Shinde, N.M. Thorat, *Synth Commun*, 2021, **51** (16), 2528–2543.

(**81**) Q. Yang, M. Sheng, J. J. Henkelis, S. Tu, E. Wiensch, *Org Process Res Dev*, 2019, **23** (10), 2210–2217.

(**82**) U. Streit, F. Birbaum, A. Quattropani, C.G. Bochet, *J Org Chem*, 2013, **78**, 6890–6910.

(**83**) D.J. McNamara, E.M. Berman, D.W. Fry, L.M. Werbel, *J Med Chem*, 1990, **33**, 2045–2051.

(**84**) R. Chinchilla, C. Najera, *Chem Rev,* 2007, **107**, 874–922.

(**85**) E. Valeur, M. Bradley, *Chem Soc Rev*, 2009, **38**, 606–631.

(**86**) W. Zhongteng Pharmaceutical Technology Co, Ltd, CN112010856A, 2020.

(**87**) H. Zhang, CN110105362A, 2019.

(**88**) F. Zhao, X.D. Wang, L.N. Erber, M. Luo, A.Z. Guo, S.S. Yang, J. Gu, B.J. Turman, Y.R. Gao, D.F. Li, Z.Q. Cui, Z.P. Zhang, L.J. Bi, A.D. Baughn, X.E. Zhang, J.Y. Deng, *Antimicrob Agents Chemother*, 2014, **58** (3), 1479–1487.

(**89**) M.G. Nair, B.R. Murthy, S.D. Patil, R.L. Kisliuk, J. Thorndike, Y. Gaumont, R. Ferone, D.S. Duch, M.P Edelstein, *J Med Chem*, 1989, **32**, 1277–1283.

(**90**) J.I. DeGraw, W.T. Colwell, J.R. Piper, F.M. Sirotnak, *J Med Chem*, 1993, **36**, 2228–2231.

(**91**) J.I. Borrell, J. Teixido, J.L. Matallana, B. Martinez-Teipel, C. Colominas, M. Costa, M. Balcells, E. Schuler, M.J. Castillo, *J Med Chem*, 2001, **44**, 2366–2369.

(**92**) M.D. Varney, C.L. Plamer, W.H. Romines, T. Boritzki, S.A. Margosiak, R. Almassy, C.A. Janson, C. Bartlett, E.J. Howland, R. Ferre, *J Med Chem*, 1997, **40**, 2502–2524.

(**93**) J. Luo, M.D. Smith, D.A. Lantrip, S. Wang, P.L. Fuchs, *J Med Chem*, 1997, **119**, 10004–10013.

(**94**) S. Kamchonwongpaisan, J. Vanichtanankul, B. Tarnchompoo, J. Yuvaniyama, S. Taweechai, Y. Yuthavong, *Anal Chem*, 2005, **77**, 1222–1227.

(**95**) J. Vanichtanankul, S. Taweechai, C. Uttamapinant, P. Chitnumsub, T. Vilaivan, Y. Yuthavong, S. Kamchonwongpaisan, *Antimicrob Agents Chemother*, 2012, **56**, 3928–3935.

(**96**) S. Kamchonwongpaisan, N. Charoensetakul, C. Srisuwannaket, S. Taweechai, R. Ratanachak, U. Arwon, J. Vanichtanankul, D. Vitsupakorn, C. Thongpanchang, B. Tarnchompoo, T. Vilaivan, Y. Yuthavong, *Eur J Med Chem*, 2020, **195**, 112263.

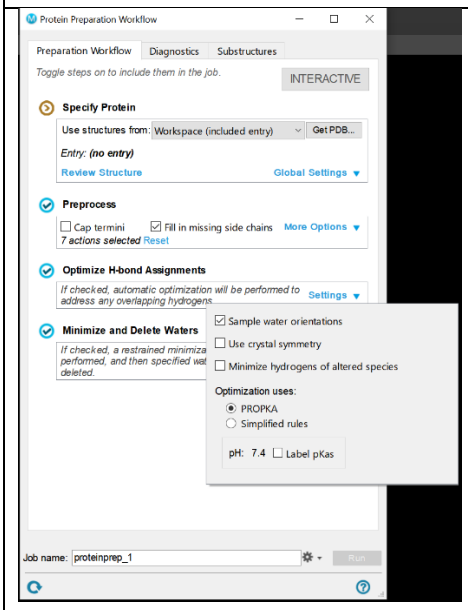
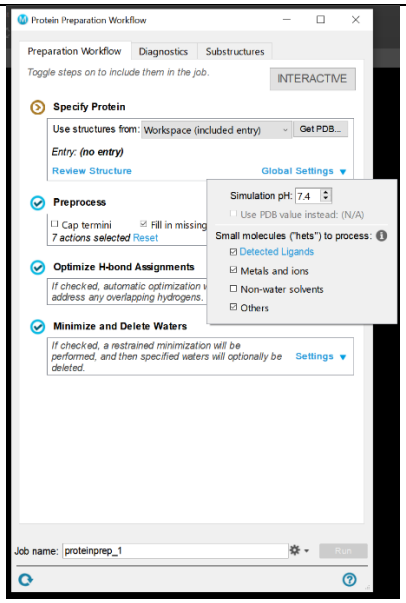
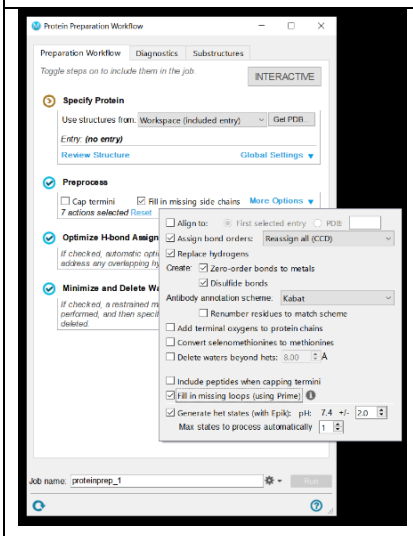
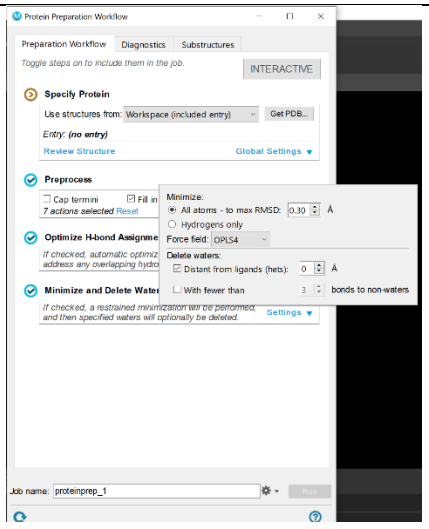
(**97**) R. E. Desjardins, C. J. Canfield, J. D. Haynes, J. D. Chulay, *Antimicrob Agents Chemother*, 1979, **16**, 710–718.

(**98**) V. Vichai, K. Kirtikara, *Nat Protoc*, 2006, **1**, 1112–1116.

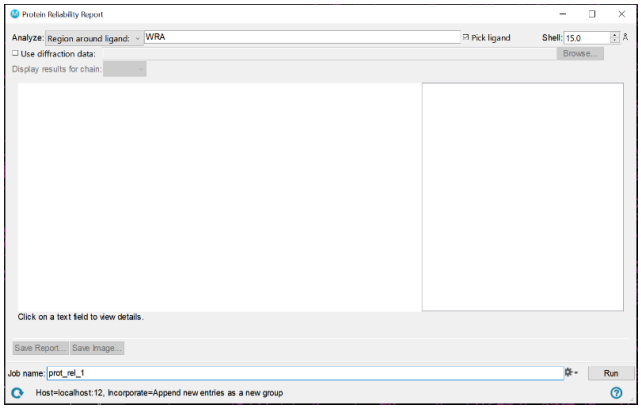
Appendices

Appendix A – Parameters for molecular modelling experiments

1.1 Protein preparation workflow:

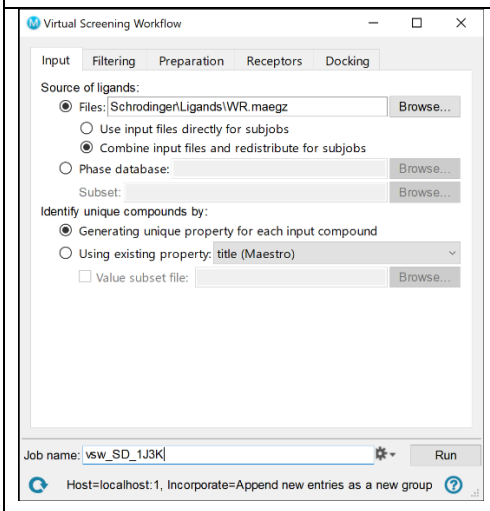
1.1.1 Global settings	1.1.2 Preprocess settings
	
1.1.3 H-bond optimization	**1.1.4 Minimization**
	

1.2 Protein reliability report

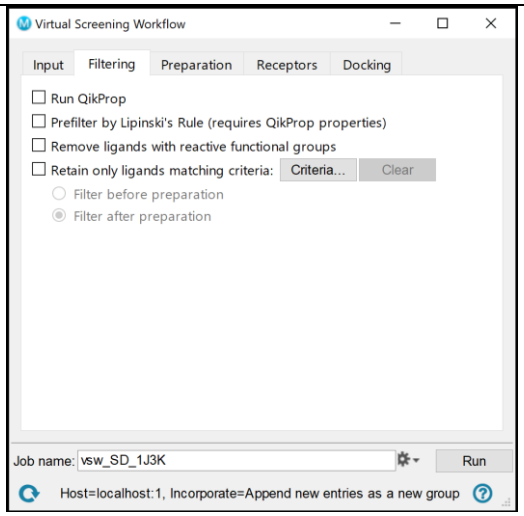


1.3 Virtual Screening Workflow for self-docking

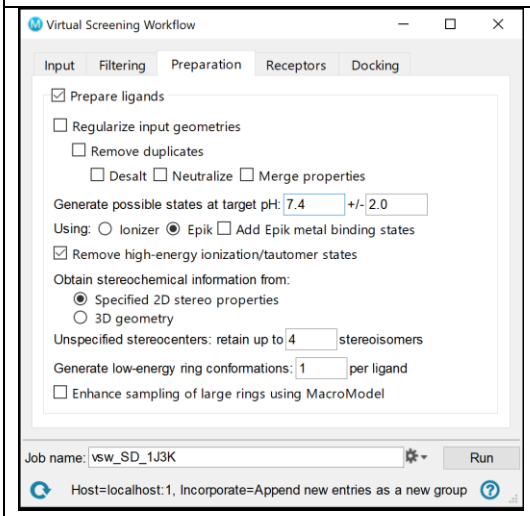
1.3.1 Input settings



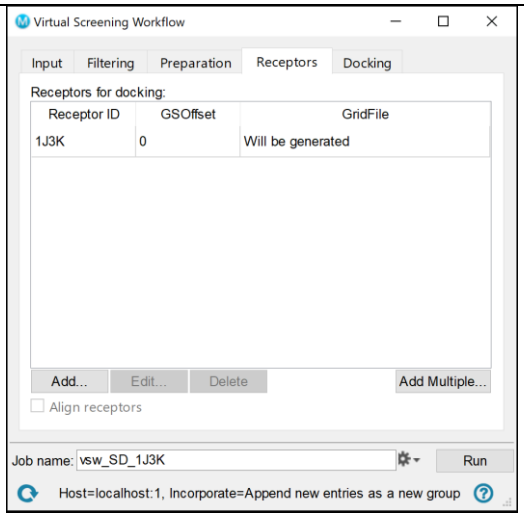
1.3.2 Filtering



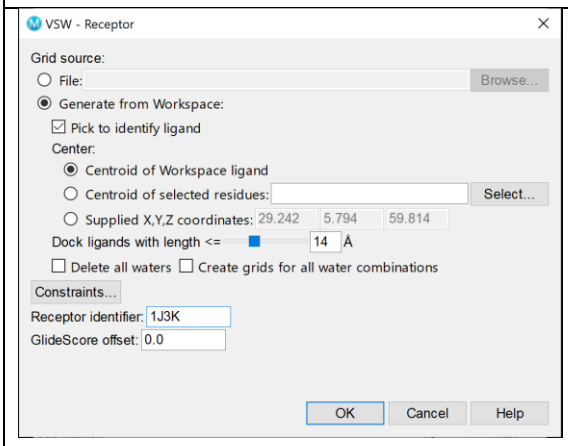
1.3.3 Preparation



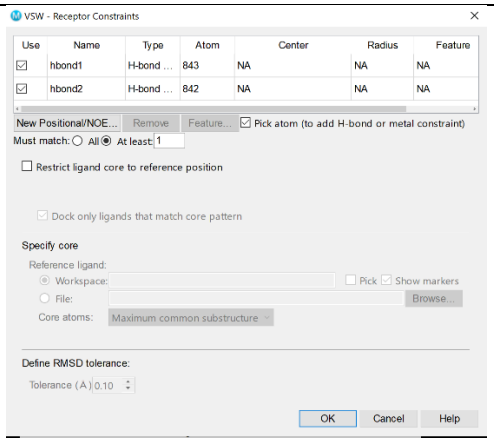
1.3.4 Receptors



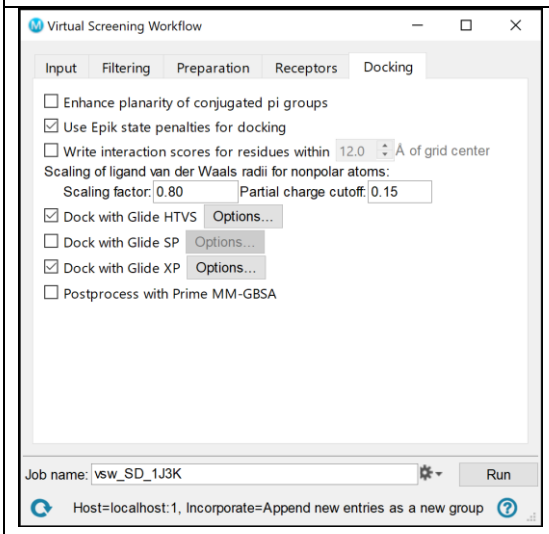
1.3.5 Receptor settings



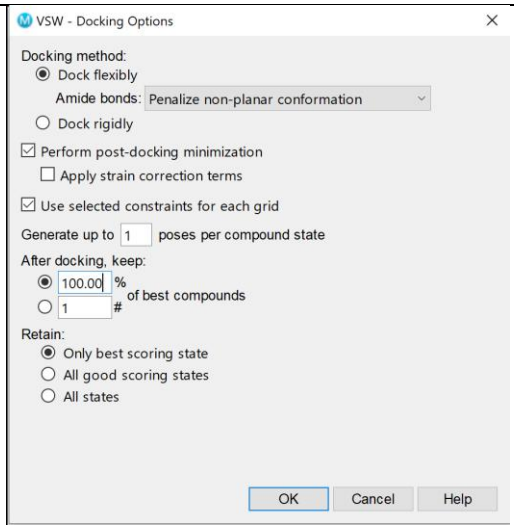
1.3.6 Constraints



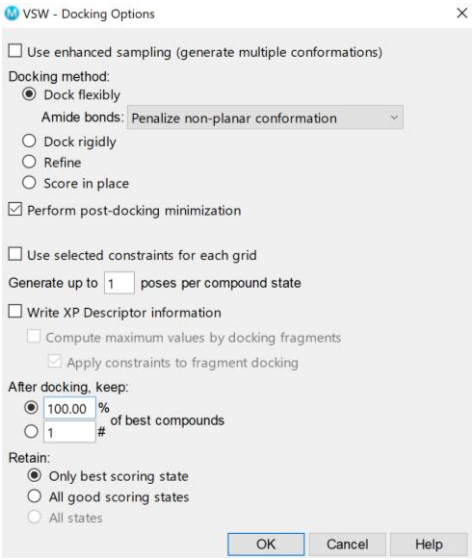
1.3.7 Docking



1.3.8 HTVS docking settings

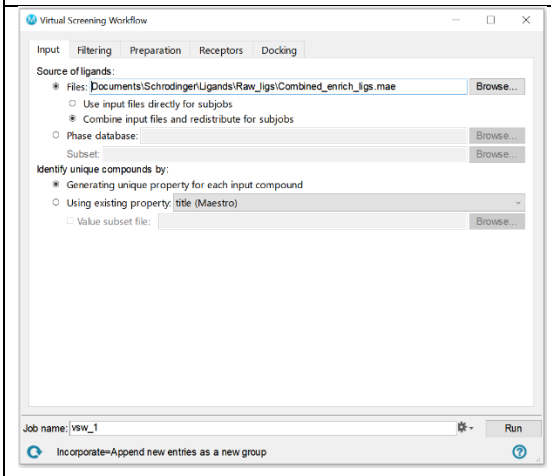


1.3.9 GlideXP docking settings

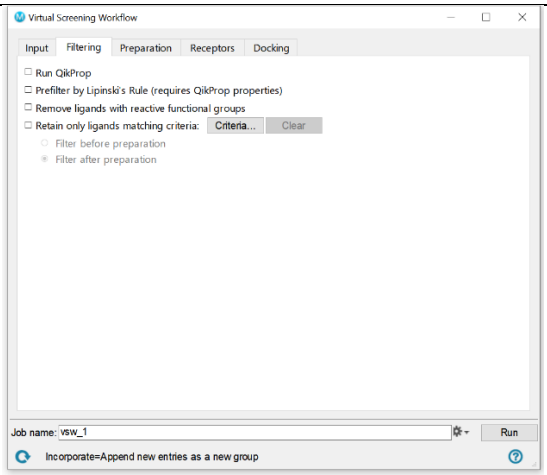


1.4 Virtual screening workflow for enrichment docking

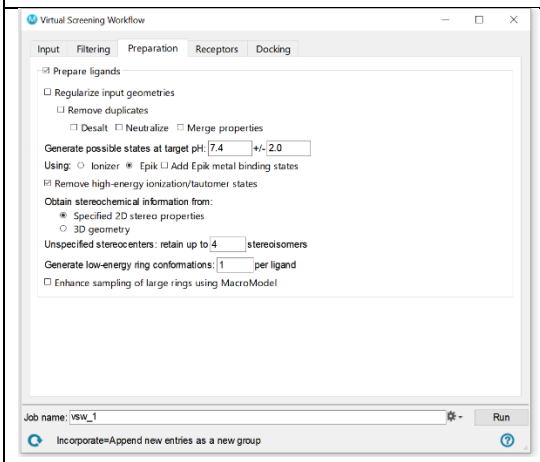
1.4.1 Input settings



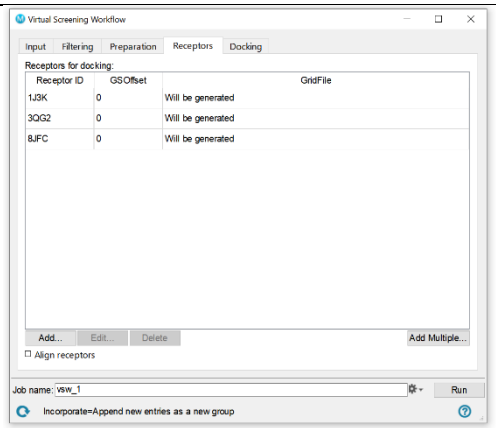
1.4.2 Filtering settings



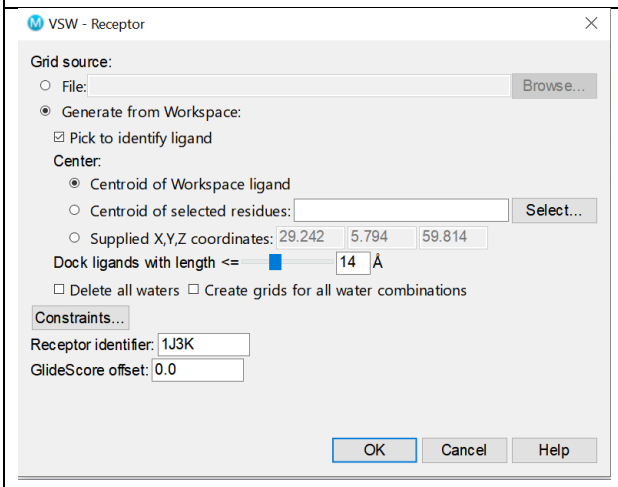
1.4.3 Ligand preparation



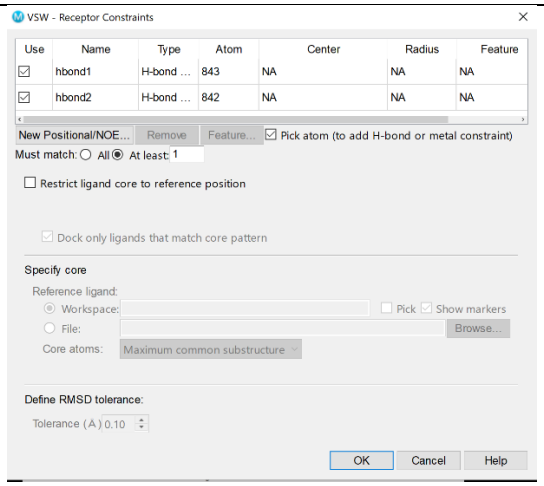
1.4.4 Receptors



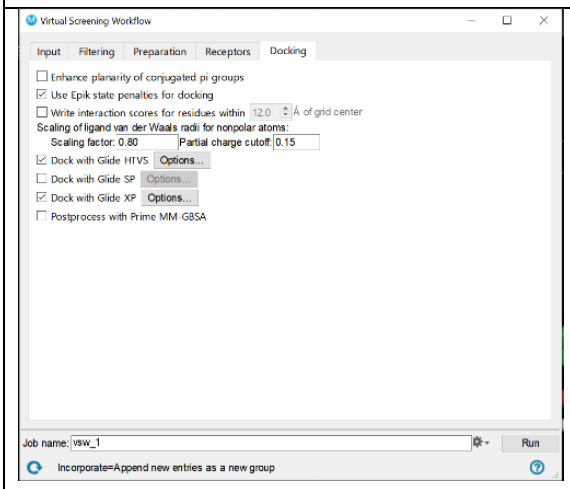
1.4.5 Receptor settings



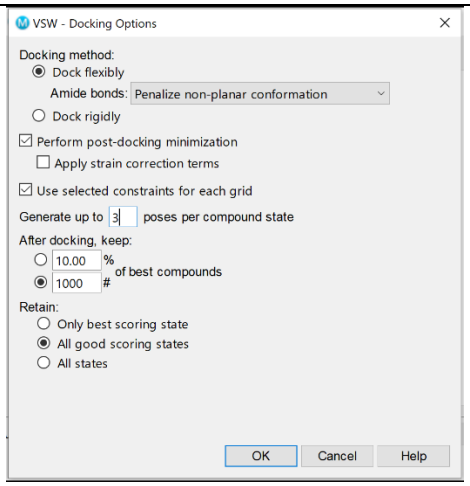
1.4.6 Constraints



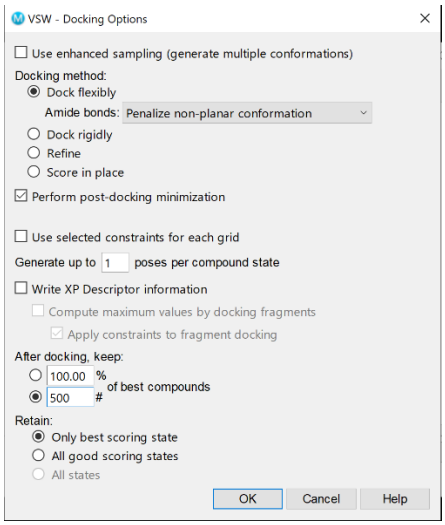
1.4.7 Docking



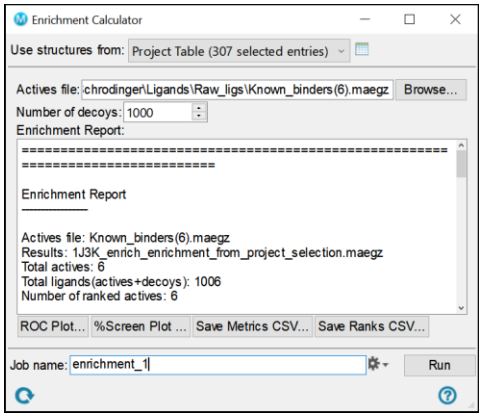
1.4.8 HTVS settings



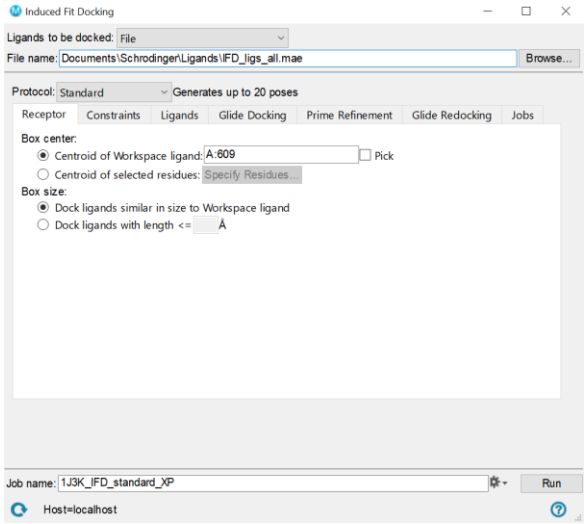
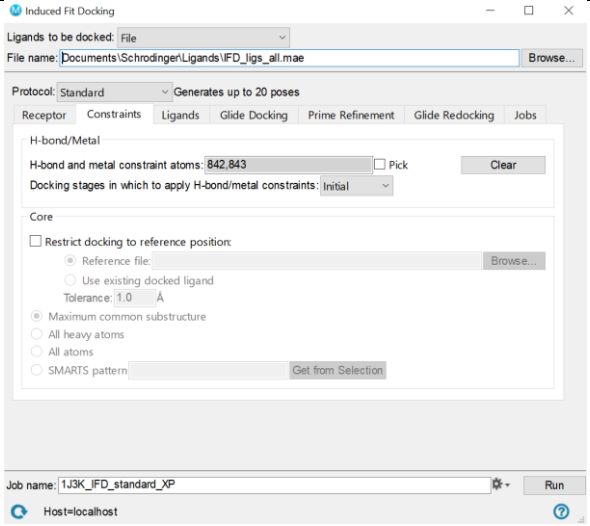
1.4.9 GlideXP settings



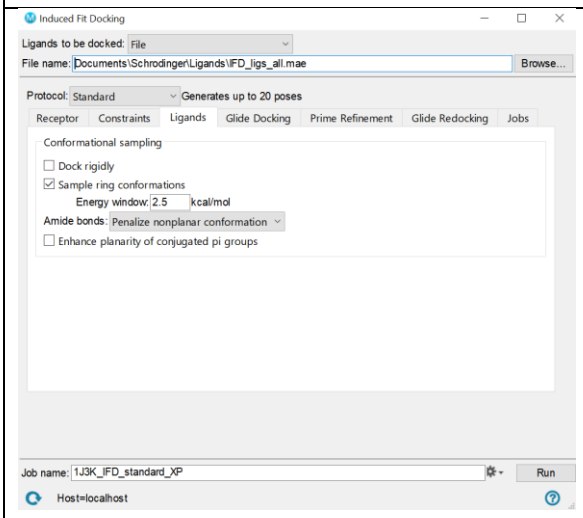
1.5 Enrichment report



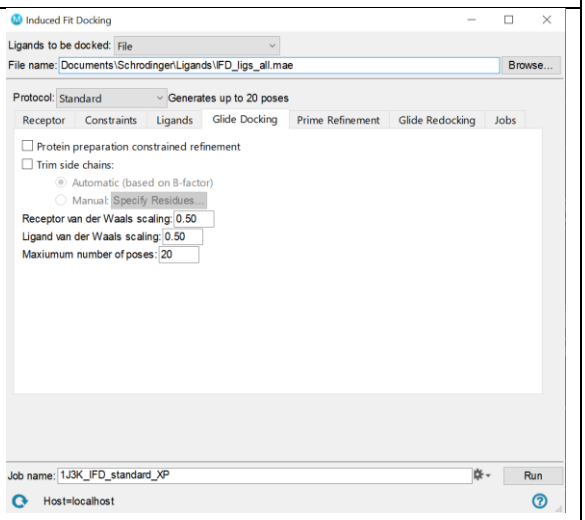
1.6 Induced fit docking

1.6.1 IFD Receptor	1.6.2 IFD Constraints
	

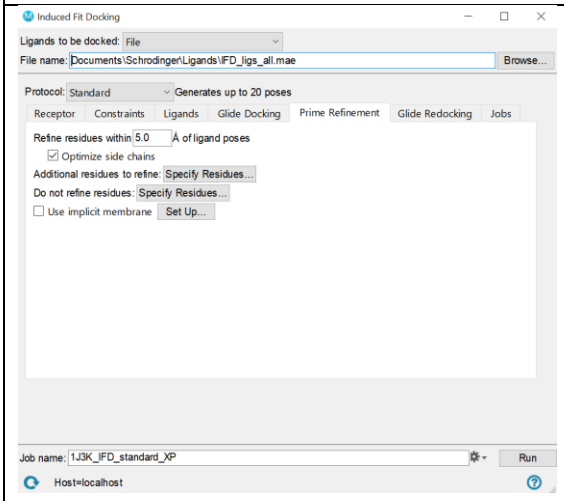
1.6.3 IFD Ligands



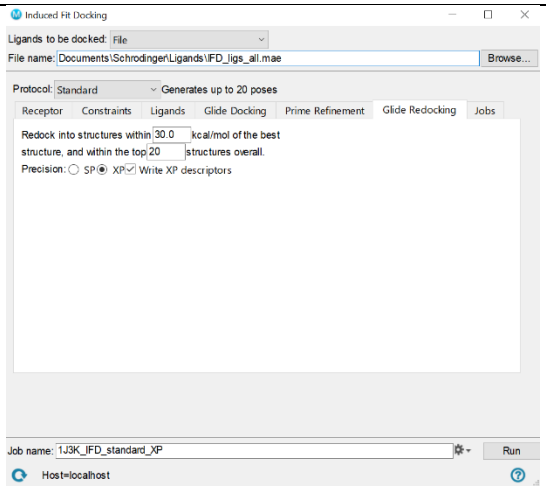
1.6.4 IFD Glide docking



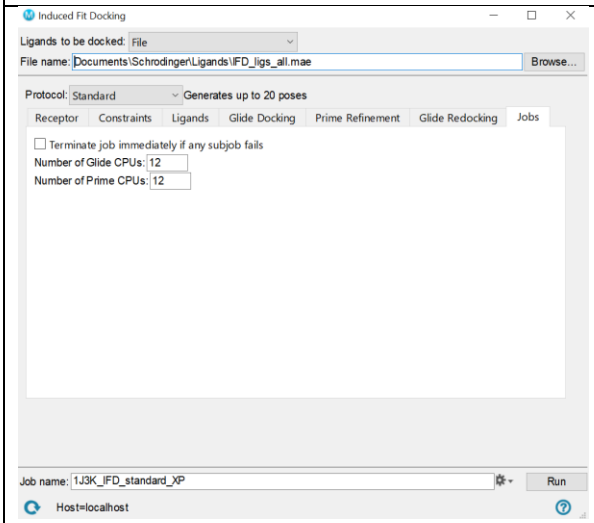
1.6.5 IFD Prime refinement



1.6.6 IFD Glide redocking

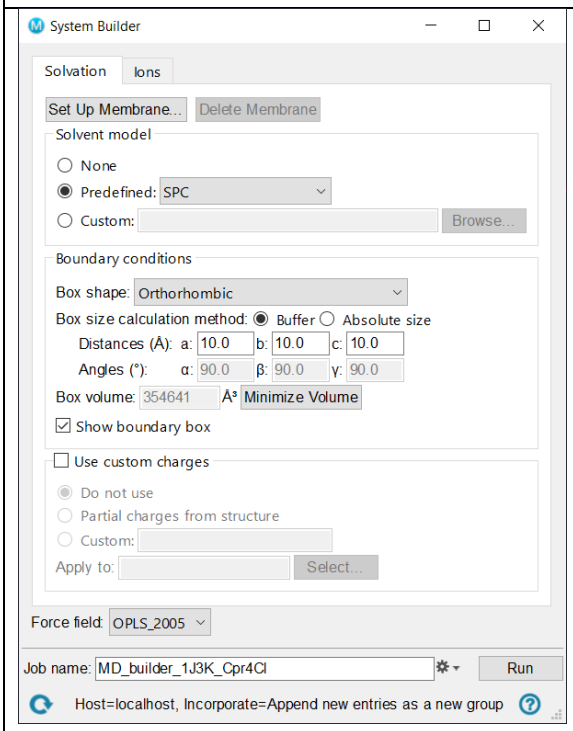


1.6.7 IFD jobs

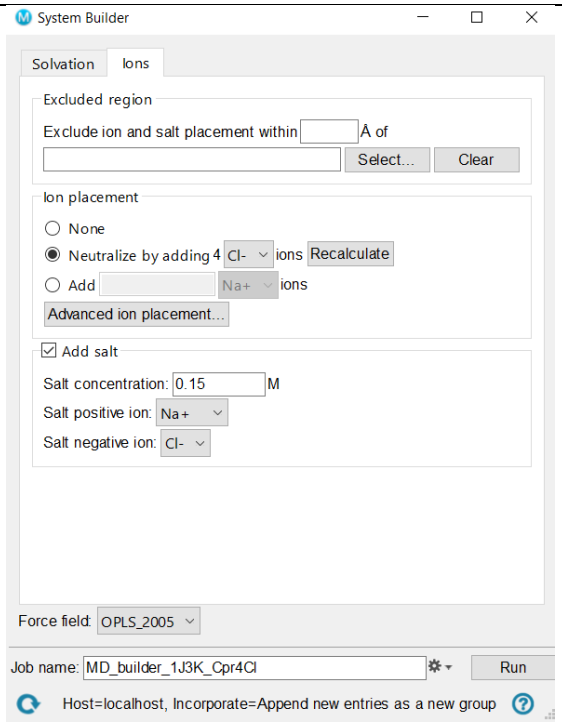


1.7 MD System builder

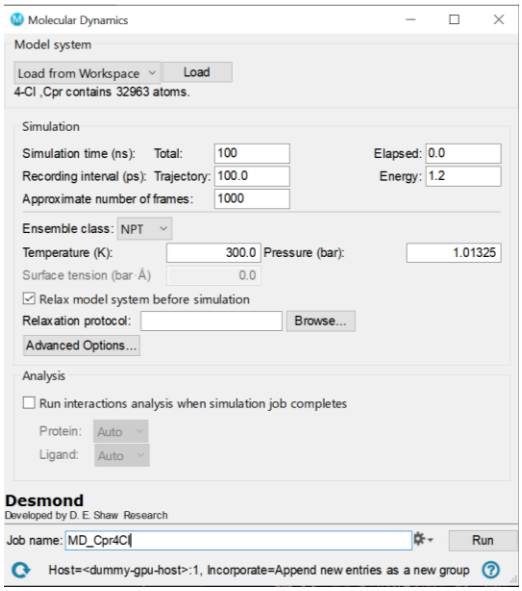
1.7.1 MD system builder solvation



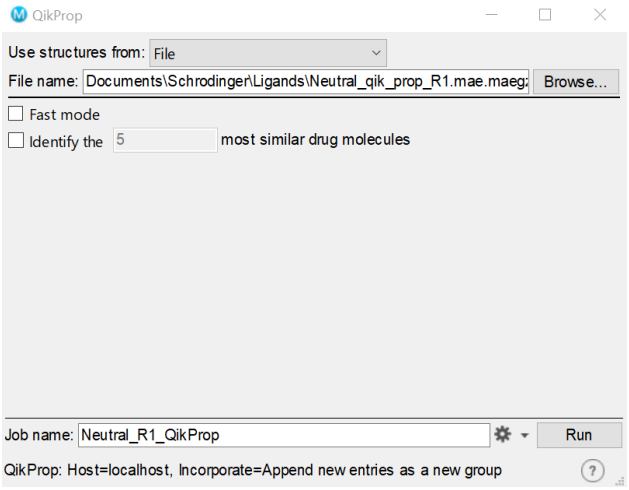
1.7.2 MD system builder ions



1.8 Molecular dynamics



1.9 QikProp


QikProp
Use structures from: File
File name: Documents\Schrodinger\Ligands\Neutral_qik_prop_R1.mae.maeg
Browse...
Fast mode
Identify the 5 most similar drug molecules
Job name: Neutral_R1_QikProp
Run
QikProp: Host=localhost, Incorporate=Append new entries as a new group

Appendix B – Predicted physiochemical properties

[a] Number of stars: Compounds are flagged with stars if any of their calculated parameters fall outside of a 95 % similarity range of a library of experimentally obtained results from known drug molecules. None of the analogues assessed were flagged with any stars, suggesting at a glance that all of the compounds are drug-like.

[b] Molecular weight: Typically, small molecule drugs have a molecular weight within a range of 130 – 725 amu. Compounds with masses outside of this range are flagged with stars. The analogues assessed have masses roughly between 300 and 350 amu, with each falling inside the defined range.

[c] Predicted dipole moment: The dipole moment is calculated by the vector sum of all polarized bonds in a molecule. This is dependent on the given conformation of the molecule. For this calculation, the lowest free energy conformation is used. The expected dipole moment for drug like molecules is between 1 and 12.5. All of the analogues fall within this range.

[d] Number of hydrogen bond donor groups: Multiple configurations for each molecule are assessed for groups which can donate H-bonds to water molecules in an aqueous solution. The value presented is an average over multiple configurations, allowing for non-integer numbers. The defined range for known drug molecules is 0 – 6 H-bond donors. The analogous assessed here have a number of H-bonds donor groups ranging from 4 to 5.5. The 2,4-diaminopyrimine core present in each analogue accounts for 4 of these, with 2 from each amine group. Some substituents included around the terminal ring introduce up to 2 additional H-bond donating groups, depending on the molecule configuration. All of the analogues assessed have fewer than 6 H-bond donating groups on average, falling within the specified range for known drug molecules.

[e] Number of hydrogen bond acceptor groups: Multiple configurations for each molecule are assessed for groups which can accept H-bonds from water molecules in an aqueous solution. The value presented is an average over multiple configurations, allowing for non-integer numbers. The defined range for known drug molecules is 2 – 20 H-bond donors. All of the analogues assessed fall within this range.

[f] Estimated octanol/water partition coefficient: This is a measure of the lipophilicity of the drug molecules. The expected range is between -2.0 and 6.5. Compounds which are too lipophilic may have low aqueous solubility, while non-lipophilic compounds will have poor membrane permeability. All of the analogues assessed have estimated octanol/water partition coefficients which fall within the desired range.

[g] Estimated water solubility: Known drugs have water solubility values falling between -6.5 and -0.5. The analogues assessed all lie in this range.

[h] Caco-2 cell permeability in nm/s: This is a model for transport across the gut-blood barrier, which is essential for orally administered drugs. This parameter only considers the diffusion across cell membranes and does not take active transport mechanisms into account. Values lower than 25 are considered poor, and values over 500 are considered great. Most of the analogous assessed have good permeability values of around 400 nm/s, however, the analogous with highly polar functional groups around the terminal ring show poor permeability of around 100 nm/s or less.

[i] Predicted blood-brain barrier permeability: Some drugs which act on receptors in the brain need to pass through the BBB, in the case of the antifolates described here crossing the blood brain barrier is not required, as they target cells primarily in the liver, and the systemic circulation. Compounds in this series have BBB permeability falling between -1.1 and -2.4, which is within the range for known drugs which do not act on the nervous system of -3.0 to -1.2.

[j] Number of likely metabolites: This includes many common metabolic processes, if this number if very high, then it is likely that the compounds will be metabolised quickly and not have chance to exhibit their desired effect. This series has #metab ranging from 2 to 4, with the acceptable value being less than 8.

[k] Human serum albumin binding coefficient: A high HSA binding coefficient is detrimental to the distribution of the drug, as compounds may become strongly bound to this enzyme in the systemic circulation and therefore unable to reach their target and exhibit the desired effect. Values ranging between -1.5 and 1.5 are considered acceptable. All of the analogues assessed lie within this range.

[l] Human oral absorptivity: Calculating by combing many relevant parameters and given integer values from 1 to 3, with 1 being poor, 2 moderate, and 3 excellent. The analogues assessed have values of 2 or 3.

[m] Percentage oral absorptivity: This value is used in the calculation of HOA and provides a percentage scale for oral absorptivity. For this parameters values less than 25 % are considered poor and values greater than 80 % are considered good. Most of the analogues assessed have values close to or above 90 %, while the compounds bearing highly polar functional groups around the terminal ring have values between 55 % and 75 %.

[n] Polar surface area: Calculated from the Van der Waals surface area of polar nitrogens and oxygens, and carbonyl carbons. Known drugs have values for this parameter ranging between 7 and 200 $Å^2$. The analogues evaluated here have polar surface areas between 80 and 130 $Å^2$.

[o] Number of violations of Lipinski's rule of 5: Lipinski's rule of 5 is made up of four criteria, the limits of which are multiples of 5. For compounds to be considered drug-like according to this definition they should have no more than 5 H-bond donors, no more than 10 H-bond acceptors, a molecular weight of 500 amu or less, and they must have octanol/water partition coefficients of less than 5. Most of the analogues assessed show no violations in Lipinski's rule of 5. The compounds containing amino substituents around the terminal ring, however, have a conformational average of 5.5 H-bond donating groups, which violates the first criteria.

[p] Number of violations of Jorgensen's rule of 3: Similar to Lipinski's rule of 5 in that these rules give an indications of a compounds drug-likeness based on its adherence to certain parameters. For Jorgensen's rule of 3, the three criteria are as follows: LogS must be greater than -5.7, caco-2 cell permeability must be greater than 22 nm/s, and the number of primary metabolites must not exceed 6. Some of the analogues assessed with highly polar functional groups around the terminal ring do not satisfy the final criteria. Aside from these compounds, however, every analogue obeys each of the three criteria of Jorgensen's rule of 3.

Table A.1: Selected physiochemical parameters estimated for analogues generated in section 2.4

Title:	**#stars [a]**	**MW [b]**	**Dipole [c]**	**DnrHB [d]**	**AccptHB [e]**	**QPlogPo/W [f]**	**QPlogS [g]**	**QPPCaco [h]**	**QPlogBB [i]**	**#metab [j]**	**QPlogKhsa [k]**	**HOA [l]**	**% HOA [m]**	**PSA [n]**	**Ro5 [o]**	**Ro3 [p]**
4-Cl, **iBut**	0	334.89	2.38	4.00	3.75	3.35	-4.59	439.29	-1.25	3	0.15	3	93.85	80.09	0	0
4-Cl, **Prop**	0	320.82	2.32	4.00	3.75	3.06	-4.34	389.76	-1.31	3	0.05	3	91.26	82.66	0	0
4-Cl, **Eth**	0	306.79	2.33	4.00	3.75	2.72	-4.03	389.76	-1.22	3	-0.05	3	89.23	82.66	0	0
4-Cl, **Meth**	0	292.77	2.24	4.00	3.75	2.38	-3.77	351.97	-1.18	3	-0.13	3	86.46	82.97	0	0
4-Cl, **Cpr**	0	318.81	2.48	4.00	3.75	2.96	-4.49	425.54	-1.14	2	0.08	3	91.34	81.07	0	0
4-Cl, **Cbu**	0	332.83	2.41	4.00	3.75	3.27	-4.87	466.44	-1.12	3	0.19	3	93.84	80.38	0	0
4-Cl, **Cpe**	0	346.86	2.42	4.00	3.75	3.54	-5.22	466.60	-1.13	3	0.30	3	95.42	80.39	0	0
4-Cl, **Che**	0	360.89	2.38	4.00	3.75	3.83	-5.55	475.74	-1.13	3	0.42	3	100.00	80.48	0	0
4-Cl, **iPr**	0	320.82	2.39	4.00	3.75	3.08	-4.40	475.87	-1.14	3	0.07	3	92.88	80.28	0	0
4-Cl, **Tbu**	0	334.85	2.14	4.00	3.75	3.43	-4.90	478.84	-1.17	2	0.21	3	95.02	79.61	0	0

Table A.2: Selected physiochemical parameters estimated for analogues generated in section 2.5

Title:	#stars [a]	MW [b]	Dipole [c]	DnrHB [d]	AccptHB [e]	QPlogPo/W [f]	QPlogS [g]	QPPCaco [h]	QPlogBB [i]	#metab [j]	QPlogKHsa [k]	HOA [l]	% HOA [m]	PSA [n]	Ro5 [o]	Ro3 [p]
Cpr, **4-Cl**	0	318.81	2.36	4.00	3.75	2.97	-4.44	428.31	-1.12	2	0.08	3	91.41	81.73	0	0
Cpr, **3-Cl**	0	318.81	3.67	4.00	3.75	2.97	-4.45	428.31	-1.12	2	0.08	3	91.44	81.75	0	0
Cpr, **2-Cl**	0	318.81	4.02	4.00	3.75	3.00	-4.48	428.31	-1.13	2	0.09	3	91.54	81.03	0	0
Cpr, **4-F**	0	302.35	2.37	4.00	3.75	2.71	-4.08	428.31	-1.16	2	0.01	3	89.94	81.73	0	0
Cpr, **3-F**	0	302.35	1.14	4.00	3.75	2.72	-4.09	428.31	-1.17	2	0.01	3	89.96	81.75	0	0
Cpr, **2-F**	0	302.35	4.10	4.00	3.75	2.76	-4.21	428.31	-1.18	2	0.02	3	90.18	81.54	0	0
Cpr, **4-OMe**	0	314.39	3.12	4.00	4.50	2.57	-3.96	428.31	-1.35	3	0.01	3	89.11	90.21	0	0
Cpr, **3-OMe**	0	314.39	2.29	4.00	4.50	2.57	-3.88	428.31	-1.33	3	0.02	3	89.12	90.21	0	0
Cpr, **2-OMe**	0	314.39	2.74	4.00	4.50	2.57	-3.97	428.31	-1.36	3	0.01	3	89.11	89.88	0	0
Cpr, **4-Me**	0	298.39	2.44	4.00	3.75	2.77	-4.26	428.31	-1.31	3	0.12	3	90.28	81.74	0	0

Title:	#stars [a]	MW [b]	Dipole [c]	DnrHB [d]	AccptHB [e]	QPlogPo/W [f]	QPlogS [g]	QPPCaco [h]	QPlogBB [i]	#metab [j]	QPlogKHsa [k]	HOA [l]	% HOA [m]	PSA [n]	Ro5 [o]	Ro3 [p]
Cpr, **3-Me**	0	298.39	2.60	4.00	3.75	2.77	-4.26	428.31	-1.31	3	0.12	3	90.28	81.74	0	0
Cpr, **2-Me**	0	298.39	1.85	4.00	3.75	2.81	-4.27	428.31	-1.30	3	0.13	3	90.51	80.06	0	0
Cpr, **4-OH**	0	300.36	1.33	5.00	4.50	1.72	-3.42	129.24	-1.89	3	-0.21	3	74.81	104.40	0	0
Cpr, **3-OH**	0	300.36	2.40	5.00	4.50	1.72	-3.42	129.24	-1.89	3	-0.21	3	74.81	104.40	0	0
Cpr, **2-OH**	0	300.36	2.97	5.00	4.50	1.74	-3.47	129.97	-1.90	3	-0.21	3	74.98	103.68	0	0
Cpr, **4-NH_2**	0	299.38	3.70	5.50	4.75	1.52	-3.34	108.51	-1.97	3	-0.27	2	59.32	108.64	1	0
Cpr, **3-NH_2**	0	299.38	4.08	5.50	4.75	1.52	-3.34	108.51	-1.97	4	-0.27	2	59.32	108.64	1	0
Cpr, **2-NH_2**	0	299.38	1.13	5.50	4.75	1.55	-3.39	110.74	-1.98	4	-0.27	2	59.65	107.57	1	0
Cpr, **4-COOH**	0	328.37	2.00	5.00	5.75	1.78	-3.54	11.10	-2.37	2	-0.38	2	56.08	129.91	0	1
Cpr, **3-COOH**	0	328.37	2.17	5.00	5.75	1.80	-3.70	10.81	-2.44	2	-0.37	2	55.98	131.58	0	1
Cpr, **2-COOH**	0	328.37	4.20	5.00	5.75	1.87	-3.67	12.38	-2.36	2	-0.36	2	57.47	128.56	0	1

Title:	#stars [a]	MW [b]	Dipole [c]	DnrHB [d]	AccptHB [e]	QPlogPo/W [f]	QPlogS [g]	QPPCaco [h]	QPlogBB [i]	#metab [j]	QPlogKHsa [k]	HOA [l]	% HOA [m]	PSA [n]	Ro5 [o]	Ro3 [p]
Cpr, **4-NO_2**	0	329.36	6.38	4.00	4.75	1.83	-4.00	52.09	-2.34	3	-0.02	3	68.38	126.08	0	0
Cpr, **3-NO_2**	0	329.36	4.52	4.00	4.75	1.83	-4.00	52.09	-2.34	3	-0.02	3	68.38	126.08	0	0
Cpr, **2-NO_2**	0	329.36	7.82	4.00	4.75	1.90	-4.10	51.89	-2.37	3	0.00	3	68.76	126.21	0	0
Cyc	1	251.72	4.07	4.00	3.50	1.62	-3.09	443.88	-0.59	0	-0.21	3	83.82	73.497	0	0
Pyr	0	248.71	2.58	4.00	3.00	1.82	-3.02	421.20	-0.78	1	-0.24	3	84.57	73.467	0	0

Appendix C – Full set of data obtained from biological assays

Table A.3: Inhibition constants (Ki) for synthesized compounds in single enzyme assays

Number:	**Name:**	**Ki (nM) *Pf*DHFR-WT:**	**Ki (nM) *Pf*DHFR-QM:**	**Ki (nM) HsDHFR:**
8a	Cpe-y-4Cl	0.59 ± 0.01	17.0 ± 0.90	70 ± 1
8b	Cpe-y-3Cl	0.70 ± 0.02	9.40 ± 0.20	64 ± 3
8c	Cpe-y-2Cl	1.80 ± 0.07	23.0 ± 0.70	169 ± 3
8d	Cpe-y-4F	0.32 ± 0.02	1.81 ± 0.09	27 ± 0.4
8e	Cpe-y-3F	0.59 ± 0.04	1.76 ± 0.04	47 ± 0.8
8f	Cpe-y-2F	0.49 ± 0.02	1.73 ± 0.02	32 ± 1.9
8g	Cpe-y-4OMe	0.27 ± 0.02	1.45 ± 0.09	16 ± 0.2
8h	Cpe-y-3OMe	0.42 ± 0.02	3.50 ± 0.20	44 ± 2
8i	Cpe-y-2OMe	1.30 ± 0.08	20.0 ± 0.60	220 ± 5
8j	Cpr-y-4Cl	0.75 ± 0.03	20.0 ± 0.80	77 ± 2
8k	Cpr-y-3Cl	0.52 ± 0.01	7.30 ± 0.20	80 ± 2
8l	Cpr-y-4F	0.92 ± 0.06	16.0 ± 0.60	110 ± 3
8m	Cpr-y-3F	0.87 ± 0.04	12.0 ± 0.40	89 ± 3
8n	Cpr-y-4OMe	0.48 ± 0.01	8.90 ± 0.20	78 ± 4
8o	Cpr-y-3OMe	0.39 ± 0.02	5.70 ± 0.30	79 ± 1
8p*	tBu-y-4Cl	0.49 ± 0.06	1.95 ± 0.04	55 ± 3.6
8q*	tBu-y-3Cl	0.71 ± 0.05	1.65 ± 0.07	45 ± 1.5
8r*	tBu-y-4F	0.29 ± 0.02	1.28 ± 0.06	27 ± 0.4
8s*	tBu-y-3F	0.38 ± 0.01	1.21 ± 0.05	37 ± 1.6
8t*	tBu-y-2F	0.34 ± 0.02	1.32 ± 0.05	28 ± 1
8u*	tBu-y-4OMe	0.26 ± 0.02	0.85 ± 0.04	30 ± 1.6
8v*	tBu-y-3OMe	0.30 ± 0.01	0.78 ± 0.02	25 ± 1

Number:	Name:	Ki (nM) *Pf*DHFR-WT:	Ki (nM) *Pf*DHFR-QM:	Ki (nM) HsDHFR:
8w*	tBu-γ-2OMe	2.00 ± 0.06	38.0 ± 1.00	325 ± 24
8x*	Ipr-γ-4Cl	0.28 ± 0.01	1.70 ± 0.05	28 ± 0.4
8y*	Ipr-γ-3Cl	0.60 ± 0.03	3.10 ± 0.20	48 ± 1
8z*	Ipr-γ-2Cl	0.57 ± 0.03	2.10 ± 0.07	40 ± 1
8α*	Ipr-γ-4F	0.44 ± 0.01	2.40 ± 0.09	31 ± 0.4
8β*	Ipr-γ-3F	0.45 ± 0.04	1.90 ± 0.07	30 ± 0.8
8γ*	Ipr-γ-4OMe	0.34 ± 0.01	1.80 ± 0.09	30 ± 1
8δ*	Ipr-γ-3OMe	0.30 ± 0.01	1.40 ± 0.03	25 ± 0.6
8ε*	Ipr-γ-2OMe	0.41 ± 0.01	3.50 ± 0.01	203 ± 3
9a	Cpe-a-4Cl	1.50 ± 0.04	34.0 ± 0.20	384 ± 16
9b	Cpe-a-2Cl	1.30 ± 0.02	17.3 ± 0.50	243 ± 7
9c	Cpe-a-4F	0.42 ± 0.01	2.66 ± 0.19	206 ± 8.7
9d	Cpe-a-3F	0.41 ± 0.01	1.07 ± 0.15	137 ± 7.1
9e	Cpe-a-2F	0.48 ± 0.01	1.41 ± 0.06	385 ± 6
9f	Cpe-a-4OMe	1.10 ± 0.07	33.0 ± 0.70	408 ± 6
9g	Cpe-a-3OMe	0.27 ± 0.02	0.82 ± 0.02	64 ± 1.0
9h	Cpe-a-2OMe	5.70 ± 0.40	6.50 ± 0.70	>100 000
9i	Cpr-a-4Cl	0.31 ± 0.01	13.0 ± 0.50	56 ± 4
9j	Cpr-a-4F	0.52 ± 0.02	43.0 ± 2.00	134 ± 5
9k	Cpr-a-3F	0.49 ± 0.02	8.30 ± 0.40	111 ± 2
9l	Cpr-a-4OMe	0.33 ± 0.03	13.0 ± 0.70	83 ± 5
9m	Cpr-a-3OMe	0.35 ± 0.02	7.10 ± 0.30	171 ± 5
9n*	tBu-a-4Cl	0.99 ± 0.06	4.81 ± 0.18	290 ± 5.6
9o*	tBu-a-4OMe	0.36 ± 0.05	1.30 ± 0.06	57.8 ± 0.9

Number:	Name:	Ki (nM) *Pf*DHFR-WT:	Ki (nM) *Pf*DHFR-QM:	Ki (nM) HsDHFR:
9p*	Ipr-a-4Cl	0.40 ± 0.03	0.95 ± 0.04	232 ± 6
9q*	Ipr-a-2Cl	0.36 ± 0.01	0.43 ± 0.02	170 ± 5
9r*	Ipr-a-4F	0.60 ± 0.02	1.10 ± 0.07	430 ± 34
9s*	Ipr-a-3F	0.41 ± 0.02	0.58 ± 0.02	236 ± 1
9t*	Ipr-a-4OMe	0.44 ± 0.03	1.00 ± 0.02	173 ± 2
9u*	Ipr-a-3OMe	0.39 ± 0.03	0.61 ± 0.06	335 ± 9
9v*	Ipr-a-2OMe	0.69 ± 0.02	0.58 ± 0.02	630 ± 32
	Cyc	0.90 ± 0.04	222 ± 18.9	235 ± 15
	Pyr	0.35 ± 0.01	232 ± 57.3	127 ± 15

Error values are expressed as the standard deviation (SD), with the number of experiments (n) = 3.

Table A.4: Inhibitory concentrations (IC_{50}) for synthesized compounds in whole cell assays

Number:	Name:	*Pf*TM4/8.2 (µM):	*Pf*V1/S (µM):	Vero (µM):	KB (µM):	SI:
8a	Cpe-y-4Cl	6.8 ± 2.7	> 50	19 ± 3.0	> 50	< 0.38
8b	Cpe-y-3Cl	14 ± 3	> 50	19 ± 5.0	> 50	< 0.38
8c	Cpe-y-2Cl	25 ± 4	> 50	33 ± 4.0	> 50	< 0.66
8d	Cpe-y-4F	7.7 ± 2.0	32 ± 0.5	13 ± 1.9	34.7 ± 2.3	0.41
8e	Cpe-y-3F	12 ± 2.9	36 ± 4.3	26 ± 4.4	55.2 ± 9.8	0.72
8f	Cpe-y-2F	5.6 ± 2	31 ± 2.2	19 ± 3.6	65.0 ± 5.4	0.61
8g	Cpe-y-4OMe	1.3 ± 0.45	26 ± 0.5	6.3 ± 1.1	31.1 ± 7.9	0.24
8h	Cpe-y-3OMe	1.5 ± 0.9	> 50	7.9 ± 2.9	> 50	< 0.16
8i	Cpe-y-2OMe	21 ± 8	> 50	28 ± 2.0	> 50	< 0.56

Number:	Name:	*Pf*TM4/8.2 (μM):	*Pf*V1/S (μM):	Vero (μM):	KB (μM):	SI:
8j	Cpr-γ-4Cl	3.2 ± 0.7	> 50	38 ± 6.0	> 50	< 0.76
8k	Cpr-γ-3Cl	4.6 ± 0.4	> 50	39 ± 6.0	> 50	< 0.78
8l	Cpr-γ-4F	0.71 ± 0.06	4.0 ± 0.6	30 ± 10	> 50	7.5
8m	Cpr-γ-3F	3.9 ± 1.3	> 50	30 ± 4.0	> 50	< 0.60
8n	Cpr-γ-4OMe	1.6 ± 0.5	> 50	31 ± 3.0	> 50	< 0.62
8o	Cpr-γ-3OMe	1.9 ± 0.3	> 50	18 ± 5.0	> 50	< 0.36
8p*	tBu-γ-4Cl	5.1 ± 0.64	50 ± 14	27 ± 17	44.2 ± 9.2	0.54
8q*	tBu-γ-3Cl	7.1 ± 0.73	23 ± 1	32 ± 9.9	> 50	1.4
8r*	tBu-γ-4F	2.9 ± 0.16	38 ± 7	12 ± 4	38.8 ± 6.3	0.32
8s*	tBu-γ-3F	1.1 ± 0.3	41 ± 7.6	12 ± 3.7	40.2 ± 8.4	0.29
8t*	tBu-γ-2F	4.4 ± 0.34	44 ± 9.1	16 ± 1	52 ± 18	0.36
8u*	tBu-γ-4OMe	0.77 ± 0.06	50 ± 7.7	8.6 ± 3.8	34 ± 14	0.17
8v*	tBu-γ-3OMe	0.28 ± 0.02	22 ± 2.8	0.7 ± 0.2	9.6 ± 3.6	0.032
8w*	tBu-γ-2OMe	13 ± 4	> 50	> 50	> 50	
8x*	Ipr-γ-4Cl	0.52 ± 0.13	> 50	27 ± 6.0	> 50	< 0.54
8y*	Ipr-γ-3Cl	0.54 ± 0.18	45 ± 3	24 ± 8	> 50	0.53
8z*	Ipr-γ-2Cl	1.3 ± 0.6	31 ± 2	30 ± 7	> 50	0.97
8α*	Ipr-γ-4F	0.58 ± 0.08	50 ± 7	13 ± 3	> 50	0.26
8β*	Ipr-γ-3F	0.69 ± 0.15	37 ± 5	23 ± 7	16 ± 5	0.62
8γ*	Ipr-γ-4OMe	0.27 ± 0.06	29 ± 7	10 ± 3	> 50	0.34
8δ*	Ipr-γ-3OMe	1.6 ± 1.3	40 ± 7	17 ± 6	17 ± 4	0.43
8ε*	Ipr-γ-2OMe	13 ± 4	> 50	> 50	> 50	
9a	Cpe-a-4Cl	3.3 ± 0.7	6.0 ± 1.5	18 ± 2	22.0 ± 0.4	3.0

Number:	Name:	*Pf*TM4/8.2 (μM):	*Pf*V1/S (μM):	Vero (μM):	KB (μM):	SI:
9b	Cpe-a-2Cl	3.1 ± 0.5	5.8 ± 1.0	14 ± 0.7	32.3 ± 0.9	2.4
9c	Cpe-a-4F	3.1 ± 0.33	6.8 ± 0.6	61 ± 6.7	46 ± 15	9.0
9d	Cpe-a-3F	3.3 ± 0.28	6.5 ± 0.7	17 ± 2.0	21.9 ± 1.2	2.6
9e	Cpe-a-2F	2.1 ± 1.1	6.8 ± 0.4	> 50	29.7 ± 11	> 7.4
9f	Cpe-a-4OMe	4.0 ± 0.5	17 ± 2	14 ± 7	23.0 ± 0.5	0.82
9g	Cpe-a-3OMe	2.6 ± 0.17	5.1 ± 0.1	15 ± 0.4	21.8 ± 0.6	2.9
9h	Cpe-a-2OMe	21 ± 1	21 ± 4.0	> 50	> 50	> 2.4
9i	Cpr-a-4Cl	2.0 ± 0.5	14 ± 4	14 ± 4	> 50	1
9j	Cpr-a-4F	2.4 ± 0.9	21 ± 8	36 ± 2	> 50	1.7
9k	Cpr-a-3F	3.3 ± 1	11 ± 2	31 ± 8	> 50	2.8
9l	Cpr-a-4OMe	1.0 ± 0.4	> 50	20 ± 14	> 50	< 0.40
9m	Cpr-a-3OMe	1.9 ± 0.2	23 ± 6	19 ± 9	> 50	0.83
9n*	tBu-a-4Cl	12 ± 1.6	23 ± 1.5	> 50	> 50	> 2.2
9o*	tBu-a-4OMe	1.3 ± 0.7	14 ± 4	40 ± 3	> 50	2.9
9p*	Ipr-a-4Cl	2.7 ± 0.8	14 ± 4	> 50	> 50	> 3.6
9q*	Ipr-a-2Cl	1.3 ± 0.5	26 ± 4	35 ± 4	23 ± 2	1.3
9r*	Ipr-a-4F	1.6 ± 0.4	20 ± 4	> 50	31 ± 14	> 2.5
9s*	Ipr-a-3F	3.7 ± 0.5	> 50	39 ± 5	35 ± 2	< 0.78
9t*	Ipr-a-4OMe	1.3 ± 0.7	14 ± 4	40 ± 3	> 50	2.9
9u*	Ipr-a-3OMe	0.96 ± 0.22	12 ± 5	22 ± 6	> 50	1.8
9v*	Ipr-a-2OMe	2.2 ± 0.8	14 ± 4	43 ± 4	> 50	3.1
	Cyc	0.037 ± 0.01	> 100	> 100	> 100	
	Pyr	0.086 ± 0.05	> 100	26 ± 2	60 ± 25	< 0.26

Error values are expressed as the standard deviation (SD), with the number of experiments (n) = 3.

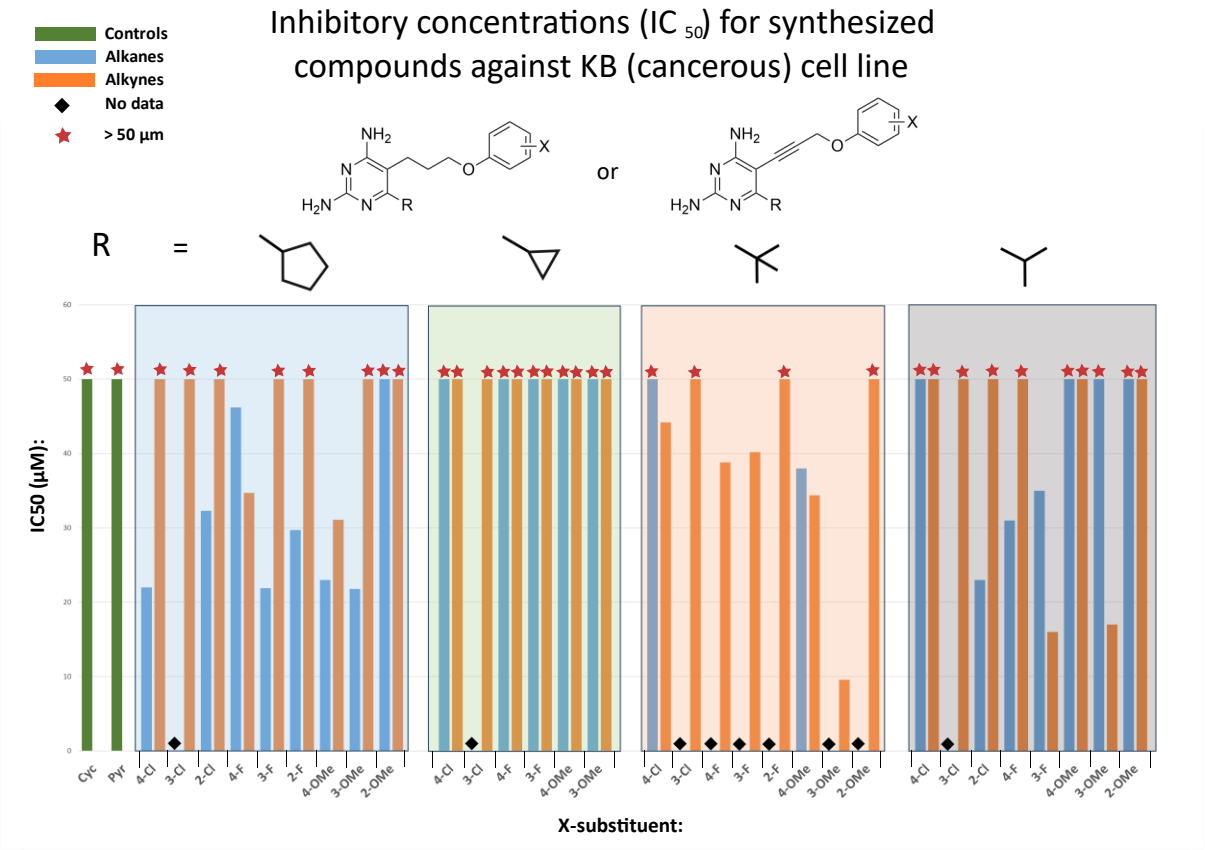


Figure A.1: IC_{50} values for synthesized compounds against KB cancerous cell line

Appendix D – Predicted physiochemical properties for second-generation analogues

Table A.5: Selected physiochemical parameters estimated for analogues generated in section 4.3

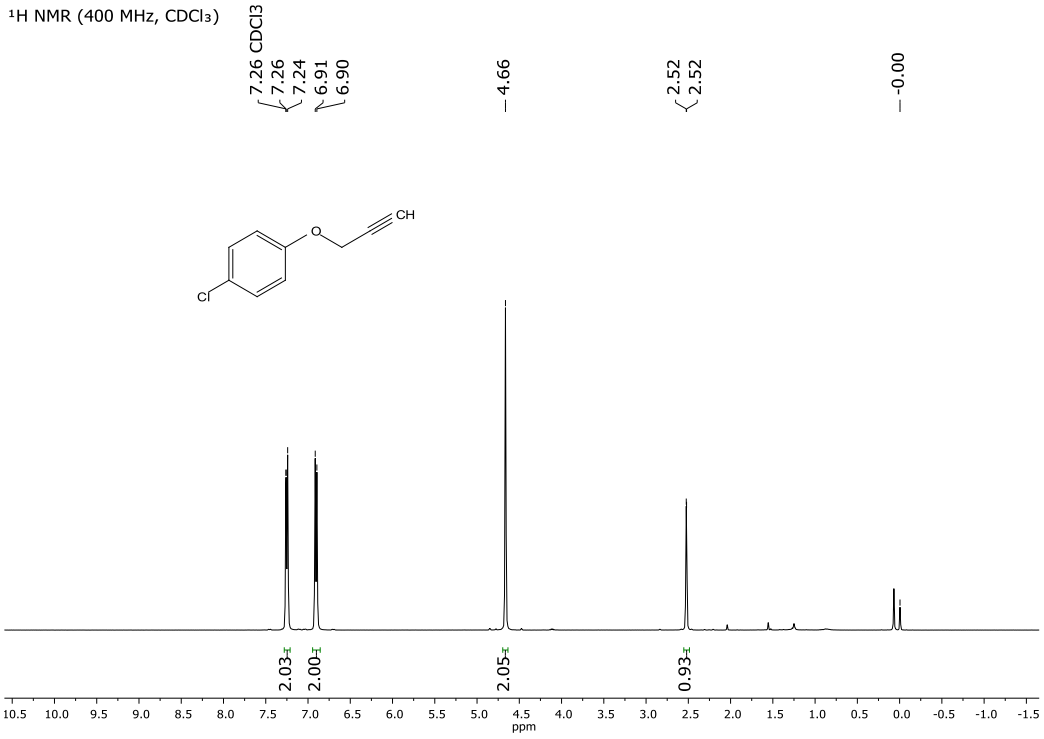
Title:	Num:	#stars [a]	MW [b]	Dipole [c]	DnrHB [d]	AccptHB [e]	QPlogPo/W [f]	QPlogS [g]	QPPCaco [h]	QPlogBB [i]	#metab [j]	QPlogKhsa [k]	HOA [l]	% HOA [m]	PSA [n]	Ro5 [o]	Ro3 [p]
O-4COOMe	**18a**	0	342.40	4.54	4.00	5.75	2.25	-4.63	131.46	-1.95	2	0.04	3	78.01	115.5	0	0
O-2propOOMe	**18b**	1	370.45	0.26	4.00	5.75	2.93	-4.62	167.50	-2.02	4	0.19	3	83.88	111.9	0	0
O-triOMe	**18c**	0	374.44	3.19	4.00	6.00	2.84	-4.59	417.59	-1.50	5	0.10	3	90.50	104.5	0	0
N(Me)-4COOMe	**18e**	0	355.44	4.66	4.00	6.00	2.58	-4.55	132.48	-2.00	2	0.18	3	80.05	110.9	0	0
N-triOMe	**19**	0	373.45	5.23	5.00	6.25	2.59	-4.43	323.02	-1.69	6	0.01	3	87.04	110.5	0	0
O-4COOH	**20a**	0	328.37	2.00	5.00	5.75	1.78	-3.54	11.10	-2.37	2	-0.38	2	56.08	129.9	0	1
O-2propOOH	**20b**	0	356.42	3.68	5.00	5.75	2.61	-3.68	19.78	-2.21	4	-0.21	2	65.41	129.7	0	1
N(Me)-4COOH	**20c**	0	341.41	4.38	5.00	6.00	2.09	-3.20	10.46	-2.43	2	-0.23	2	57.43	125.9	0	1
O-Glu	**21a**	4	457.49	1.02	6.25	9.50	1.92	-4.85	0.16	-4.62	4	-0.56	1	0.00	212.4	2	1
N(Me)-Glu	**21a**	4	470.53	9.48	6.25	9.75	2.27	-5.34	0.29	-4.14	4	-0.47	1	4.61	200.2	2	1
O-GluOMe	**22a**	2	485.54	4.73	4.25	9.50	2.81	-7.05	25.85	-3.59	4	0.21	1	55.71	181.8	1	1
N(Me)-GluOMe	**22b**	2	498.58	7.30	4.25	9.75	3.10	-6.20	27.22	-3.47	4	0.34	1	57.82	175.7	1	1

Appendix E – Characterization data for synthesized compounds

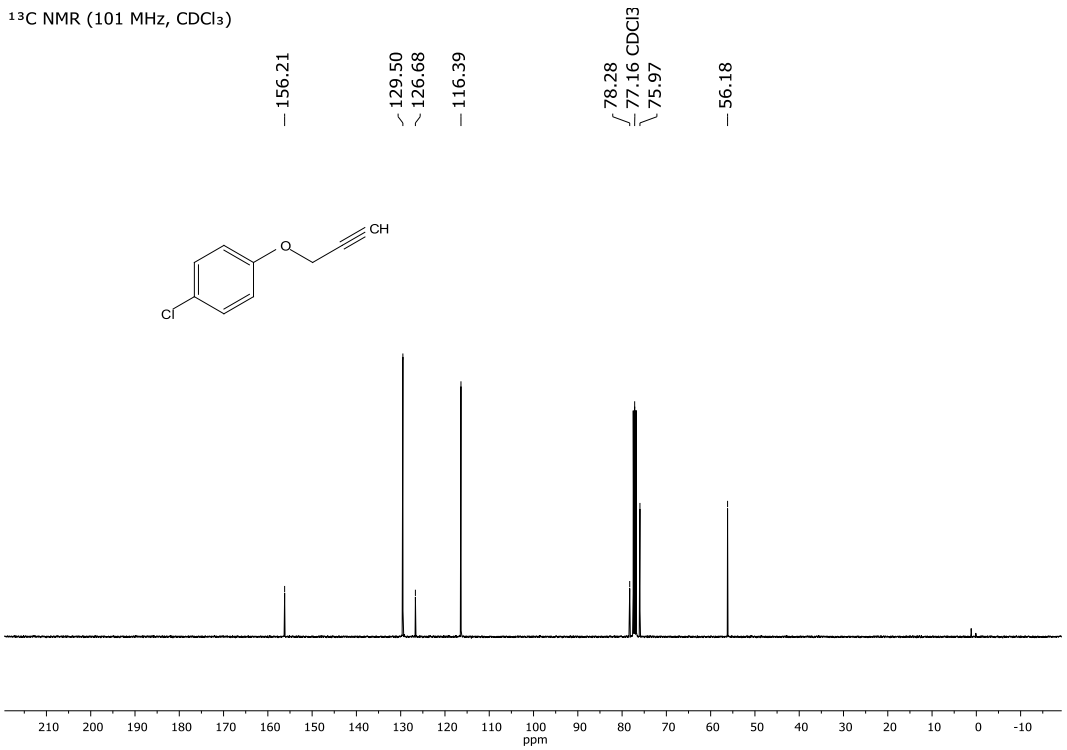
Alkylation of phenols with propargyl bromide (**2**)

1-(Prop-2-yn-1-yloxy)-4-chlorobenzene (**2a**)

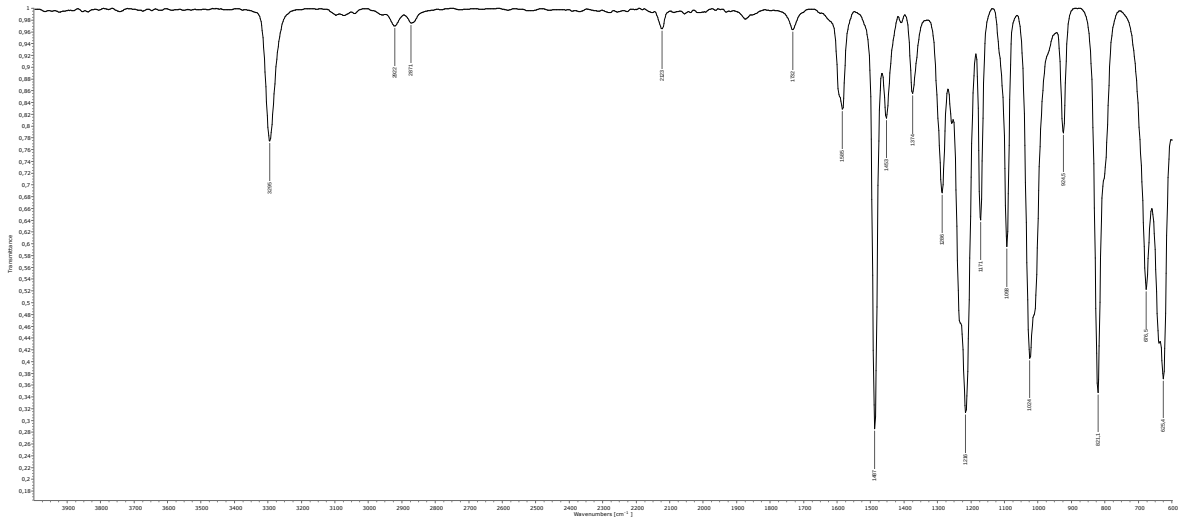
^{1}H NMR:



^{13}C NMR:



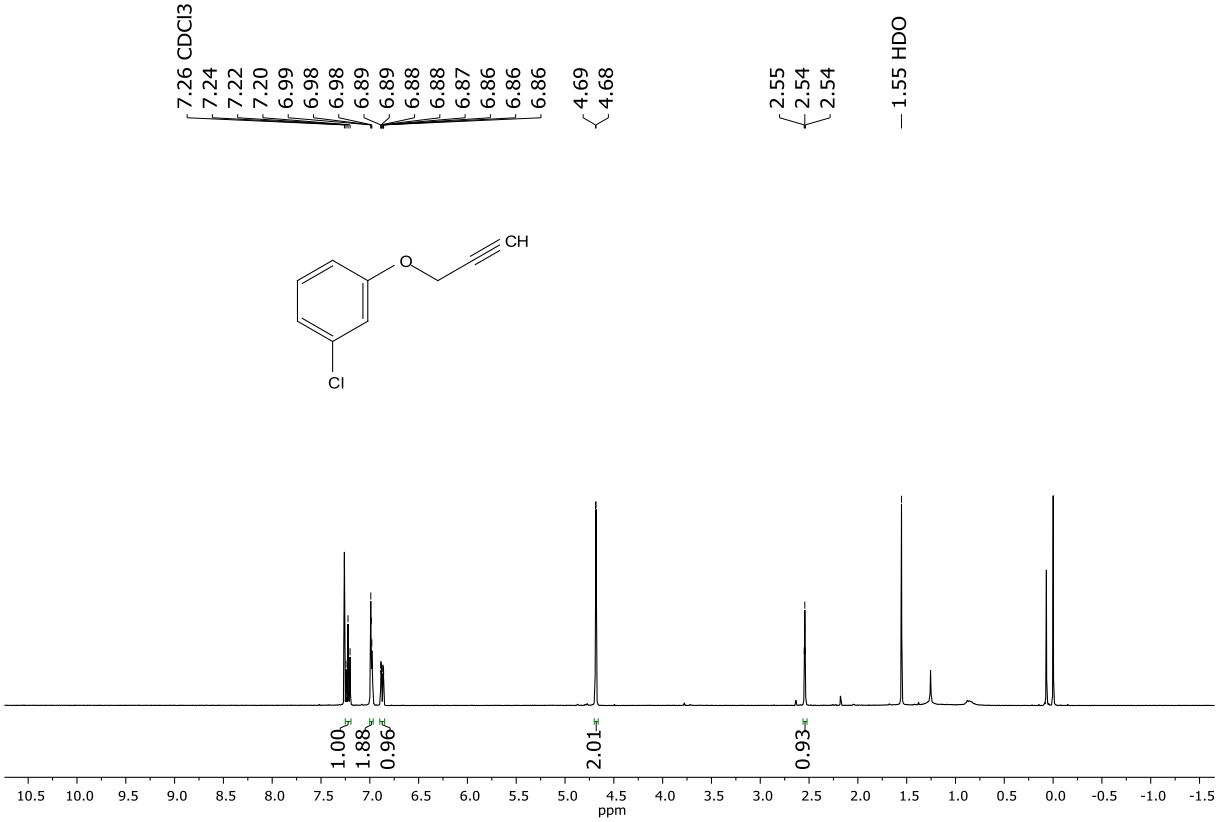
FTIR:


Transmittance
Wavenumbers (cm⁻¹)
3296
2922
2871
2323
1532
1385
1453
1374
1289
1171
1438
1487
1409
1024
1118
930.5
876.4
691.5

1-(prop-2-yn-1-yloxy)-3-chlorobenzene (**2b**)

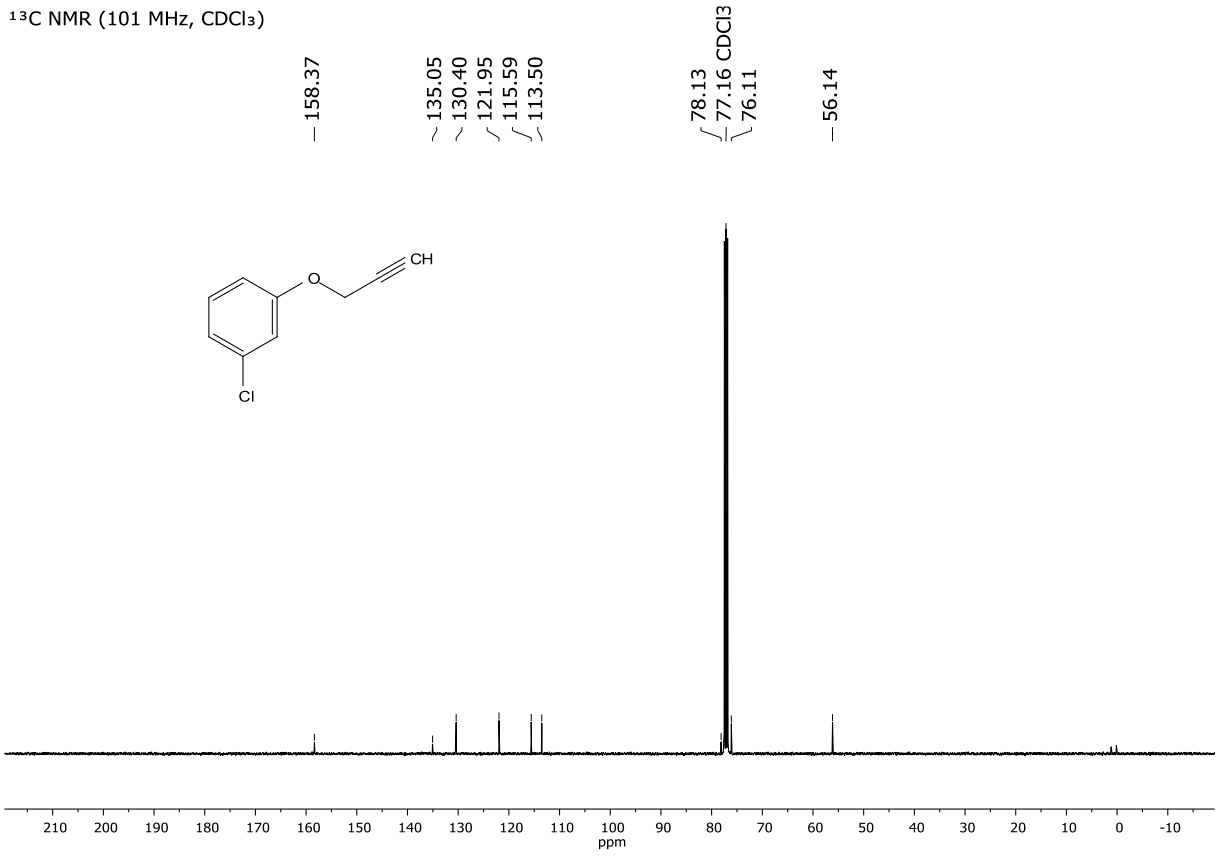
^{1}H NMR:

^{1}H NMR (400 MHz, CDCl$_3$)

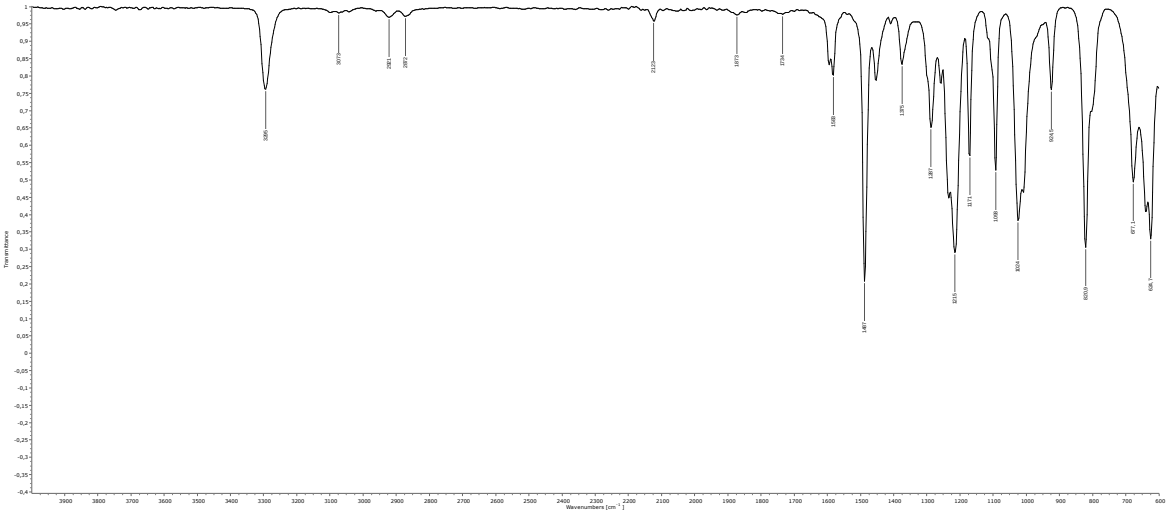


^{13}C NMR:

^{13}C NMR (101 MHz, CDCl$_3$)



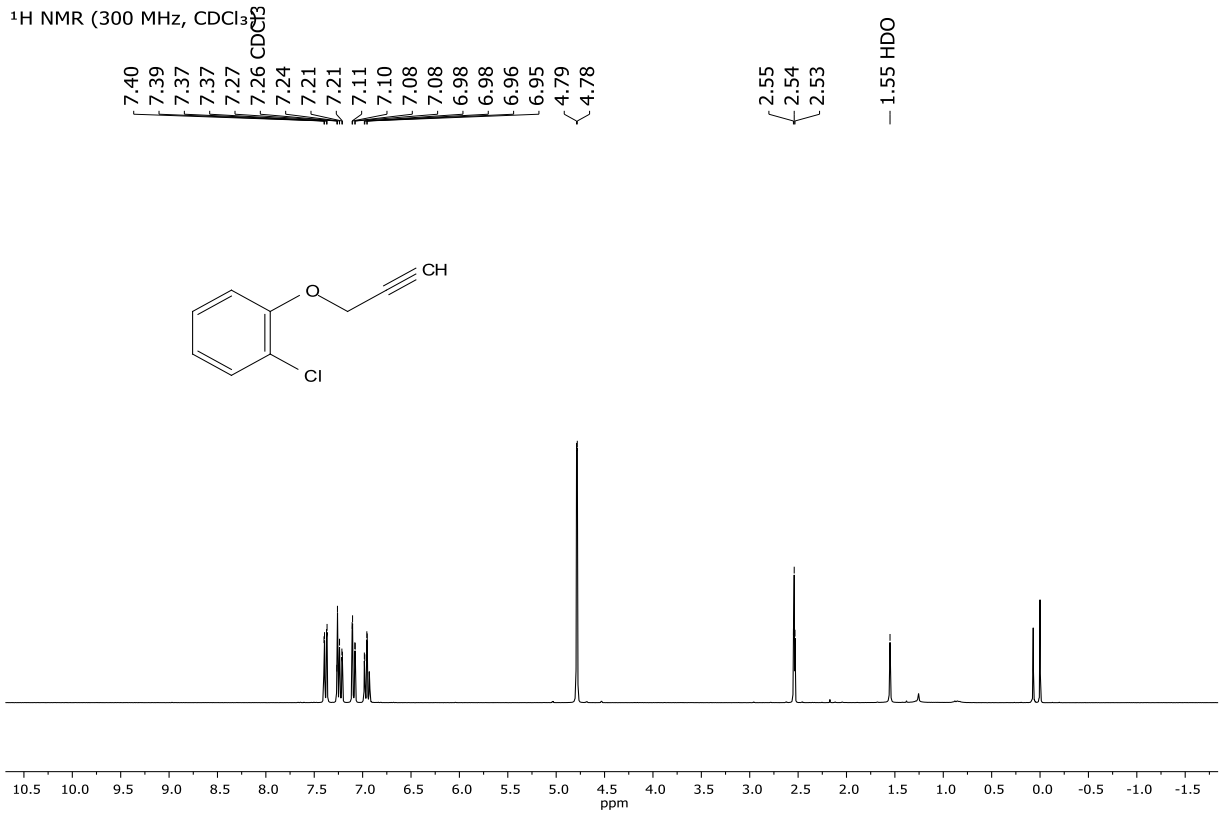
FTIR:


Transmittance
Wavenumbers [cm⁻¹]
3295
3073
2921
2872
2115
1973
1734
1588
1467
1375
1307
1210
1171
1098
1020
914
853
747
696

1-(Prop-2-yn-1-yloxy)-2-chlorobenzene (**2c**)

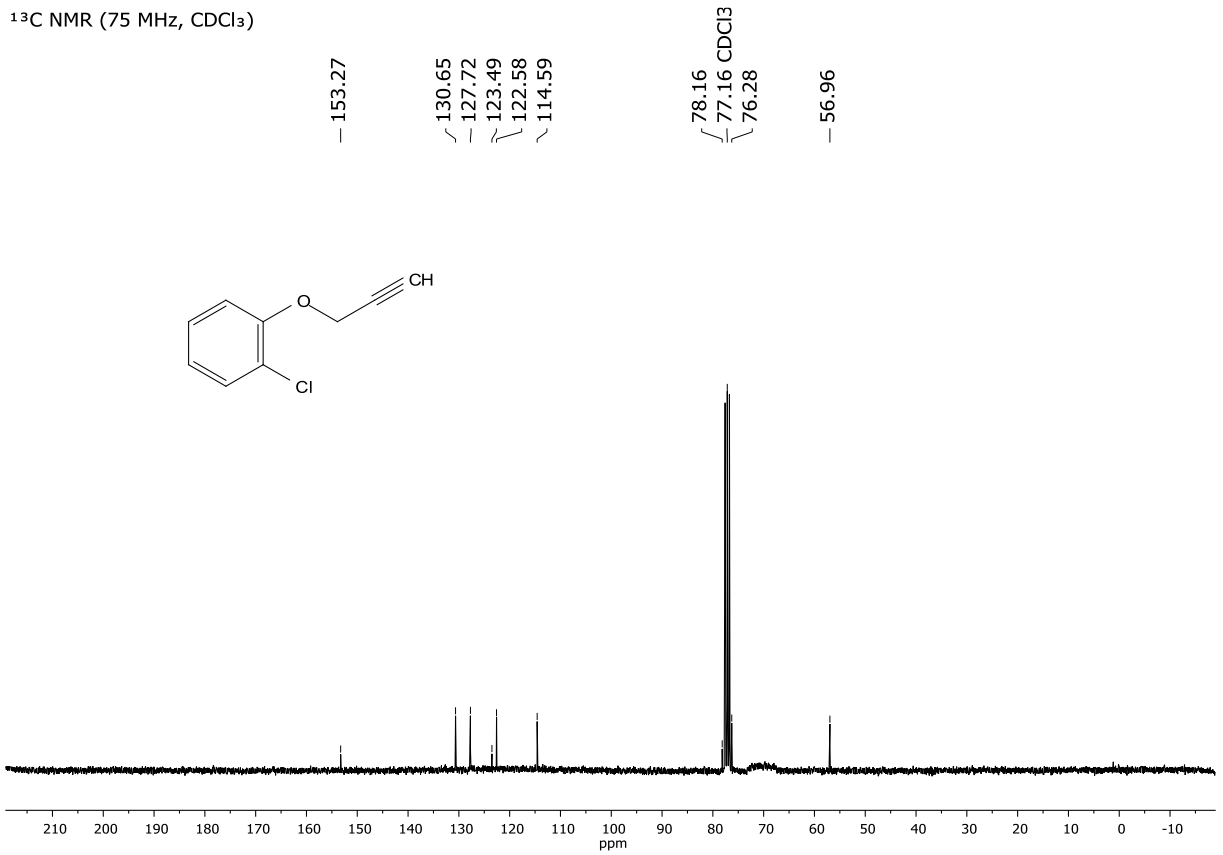
^{1}H NMR:

^{1}H NMR (300 MHz, $CDCl_3$)

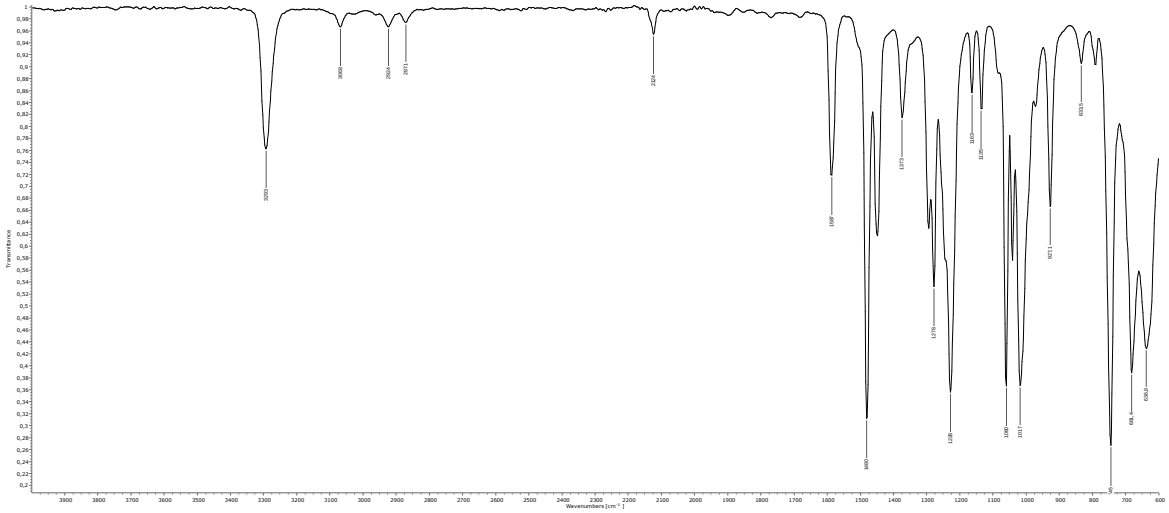


^{13}C NMR:

^{13}C NMR (75 MHz, $CDCl_3$)



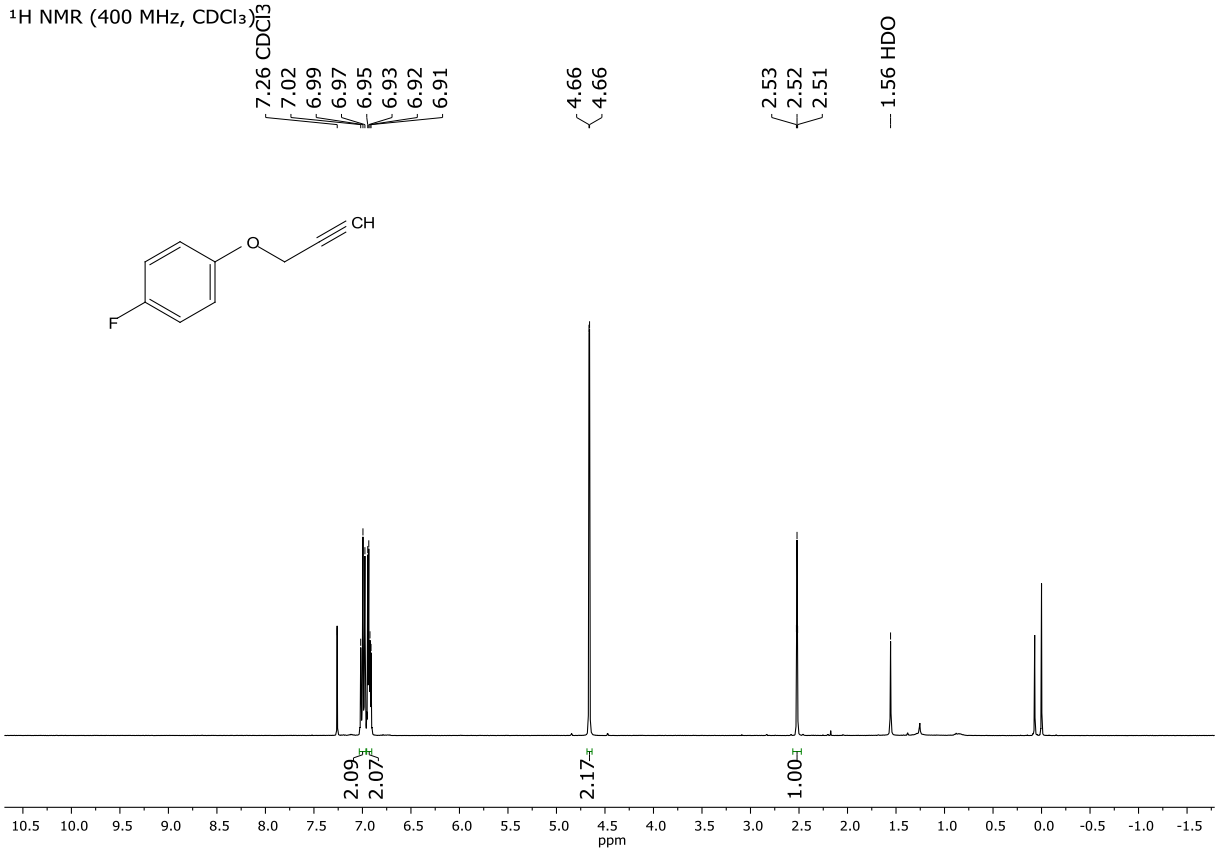
FTIR:


Transmittance
Wavenumbers [cm⁻¹]
3293
2924
1598
1370
1339
1211
1017
915
630.8

1-(Prop-2-yn-1-yloxy)-4-fluorobenzene (**2d**)

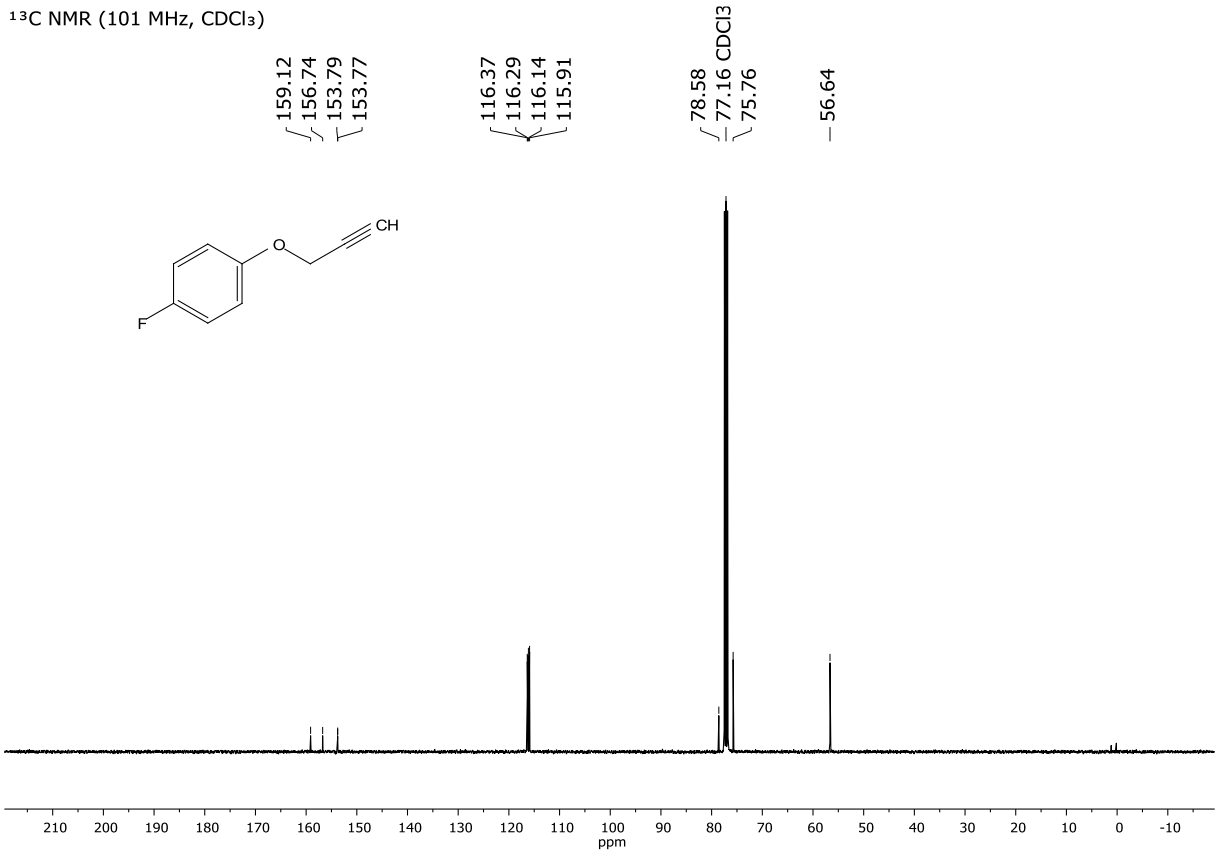
^{1}H NMR:

^{1}H NMR (400 MHz, $CDCl_3$)

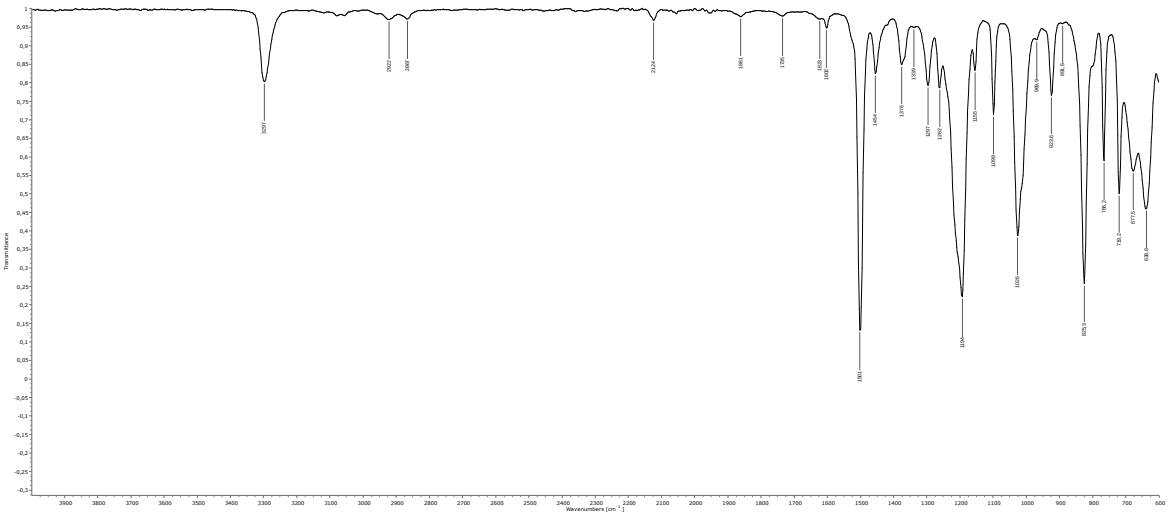


^{13}C NMR:

^{13}C NMR (101 MHz, $CDCl_3$)



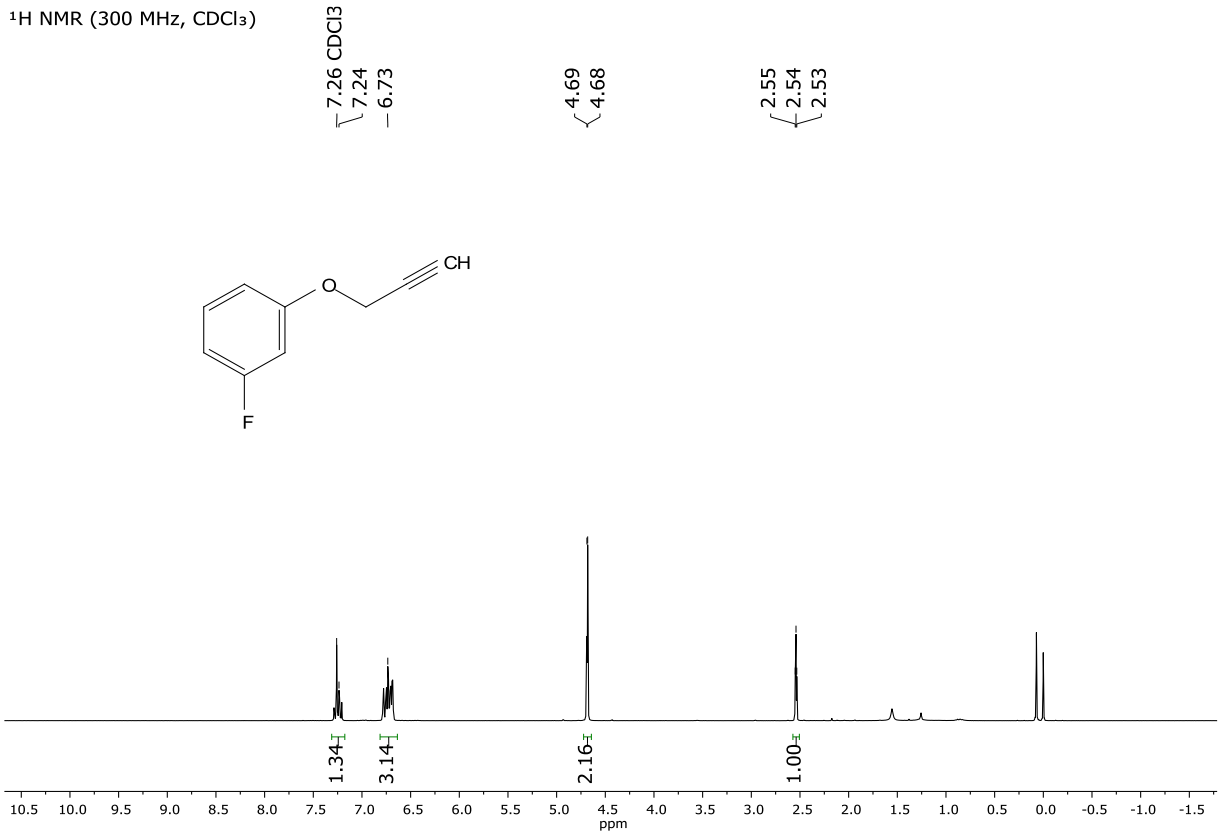
FTIR:


Transmittance
Wavenumbers [cm⁻¹]
3307
2922
2866
2124
1961
1730
1638
1604
1491
1454
1376
1339
1297
1245
1130
1110
1069
1020
984
933.4
898
825.3
786.2
730.2
677.4
638.9

1-(Prop-2-yn-1-yloxy)-3-fluorobenzene (**2e**)

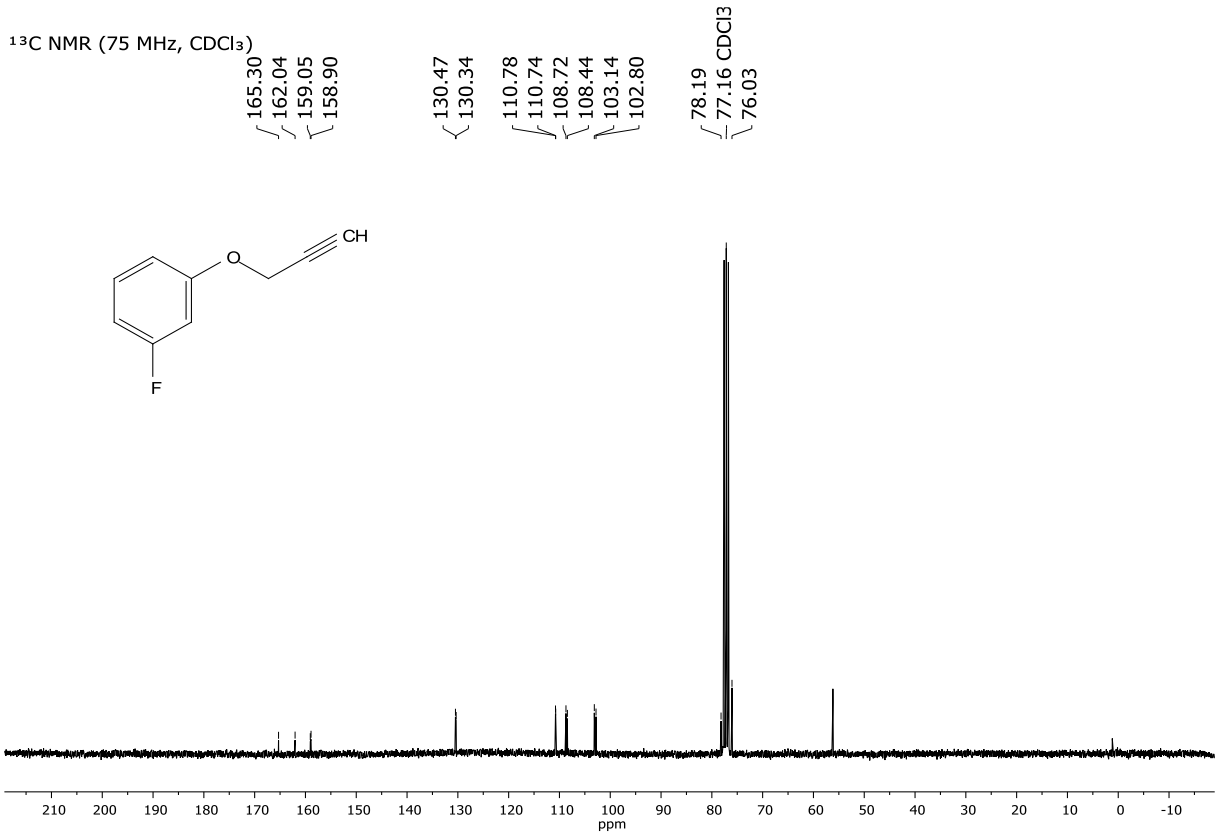
^{1}H NMR:

^{1}H NMR (300 MHz, $CDCl_3$)

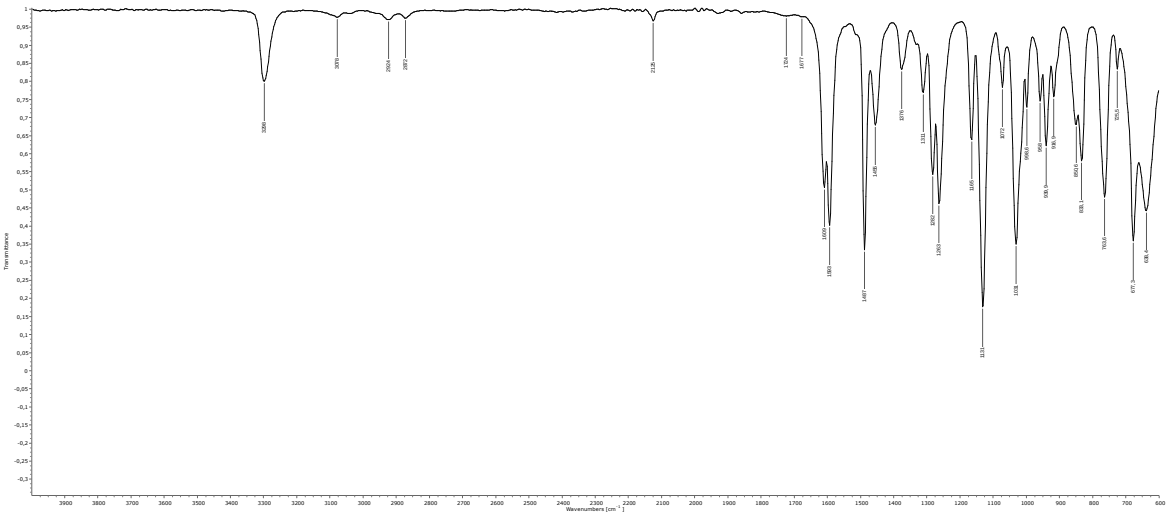


^{13}C NMR:

^{13}C NMR (75 MHz, $CDCl_3$)



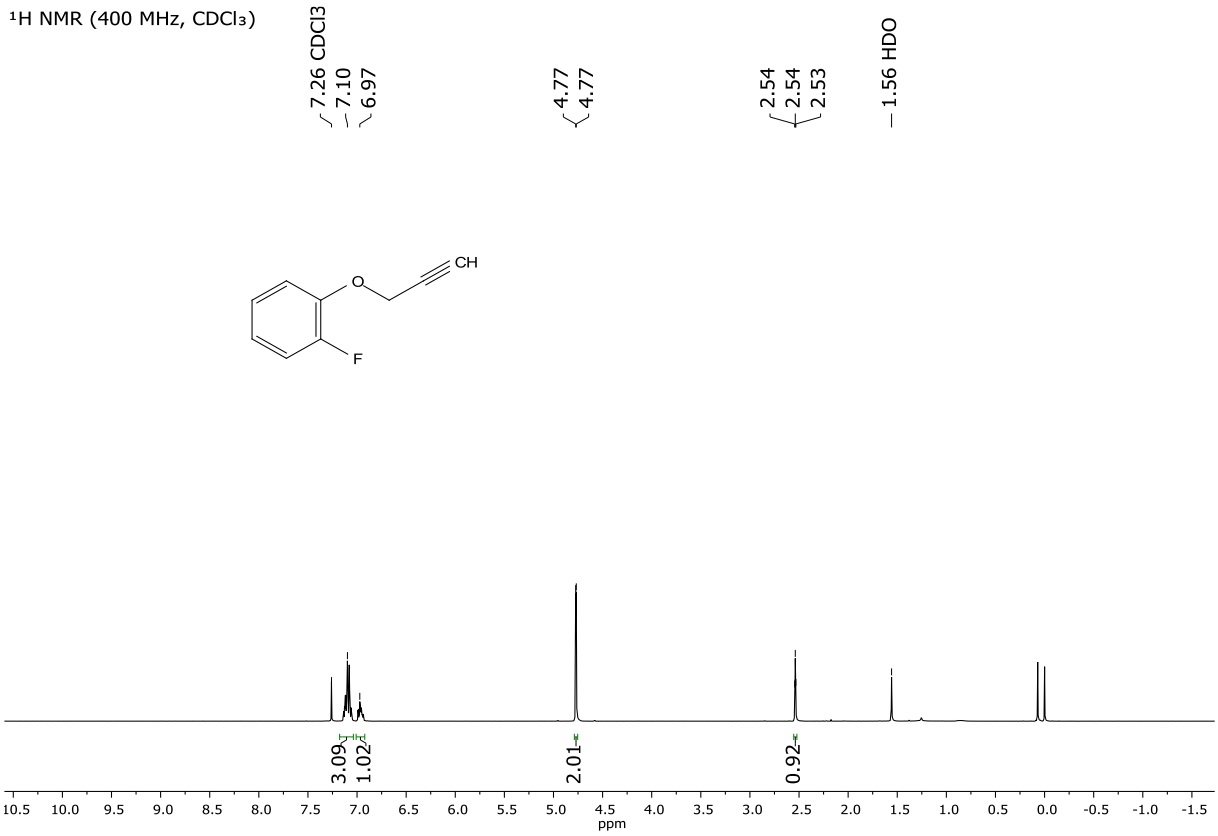
FTIR:


Transmittance
Wavenumbers [cm⁻¹]
3309
3095
2924
2872
2128
1721
1477
1600
1560
1497
1405
1371
1331
1292
1255
1101
1018
972
9984
958
936.5
936.1
891.6
781.9
712.2
677.3
636.4

1-(Prop-2-yn-1-yloxy)-2-fluorobenzene (**2f**)

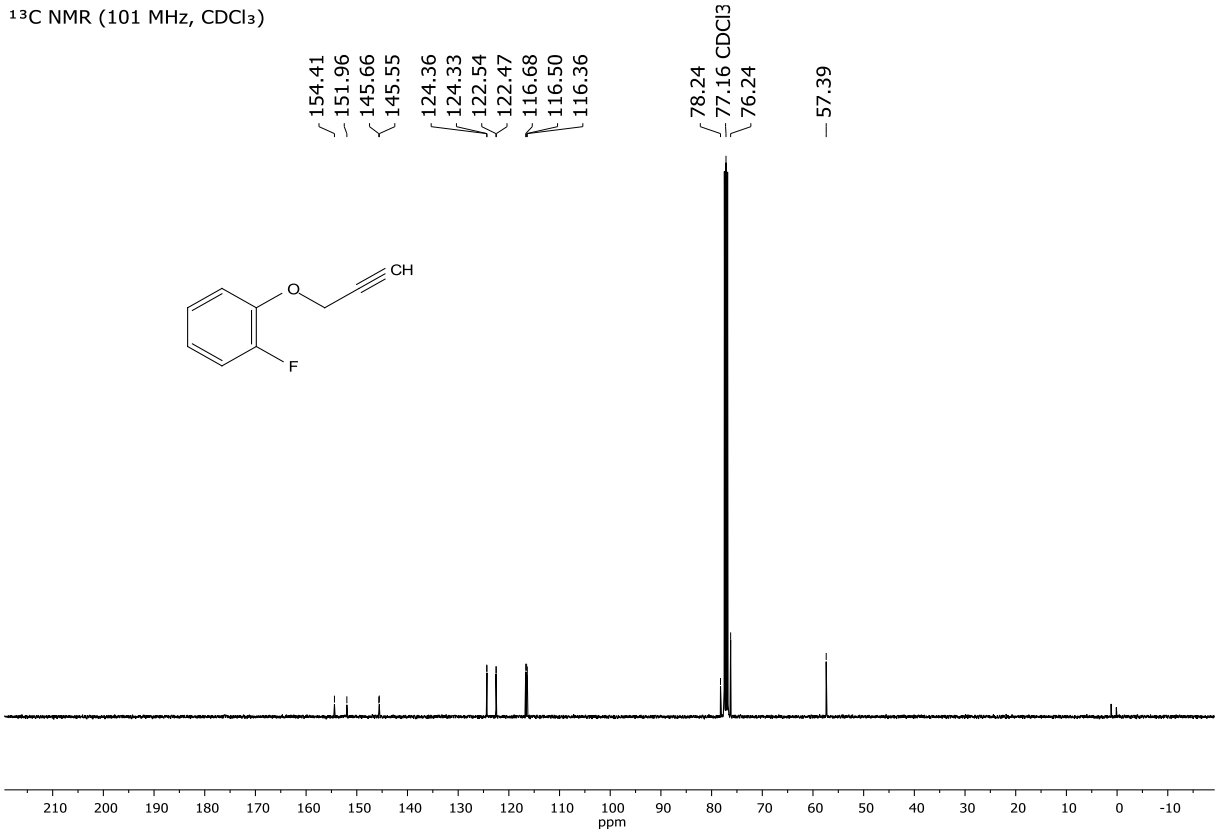
^{1}H NMR:

^{1}H NMR (400 MHz, $CDCl_3$)

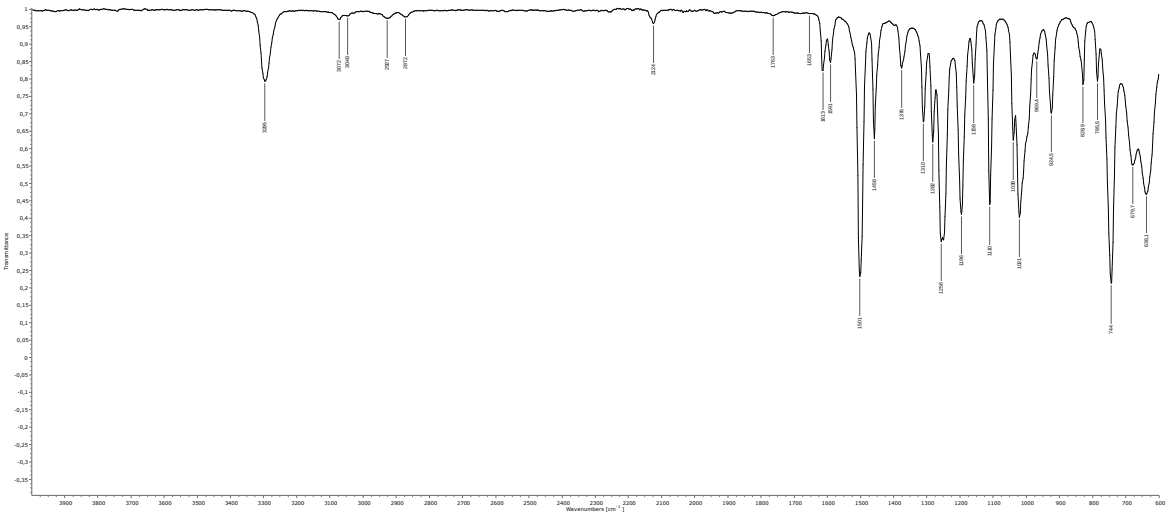


^{13}C NMR:

^{13}C NMR (101 MHz, $CDCl_3$)



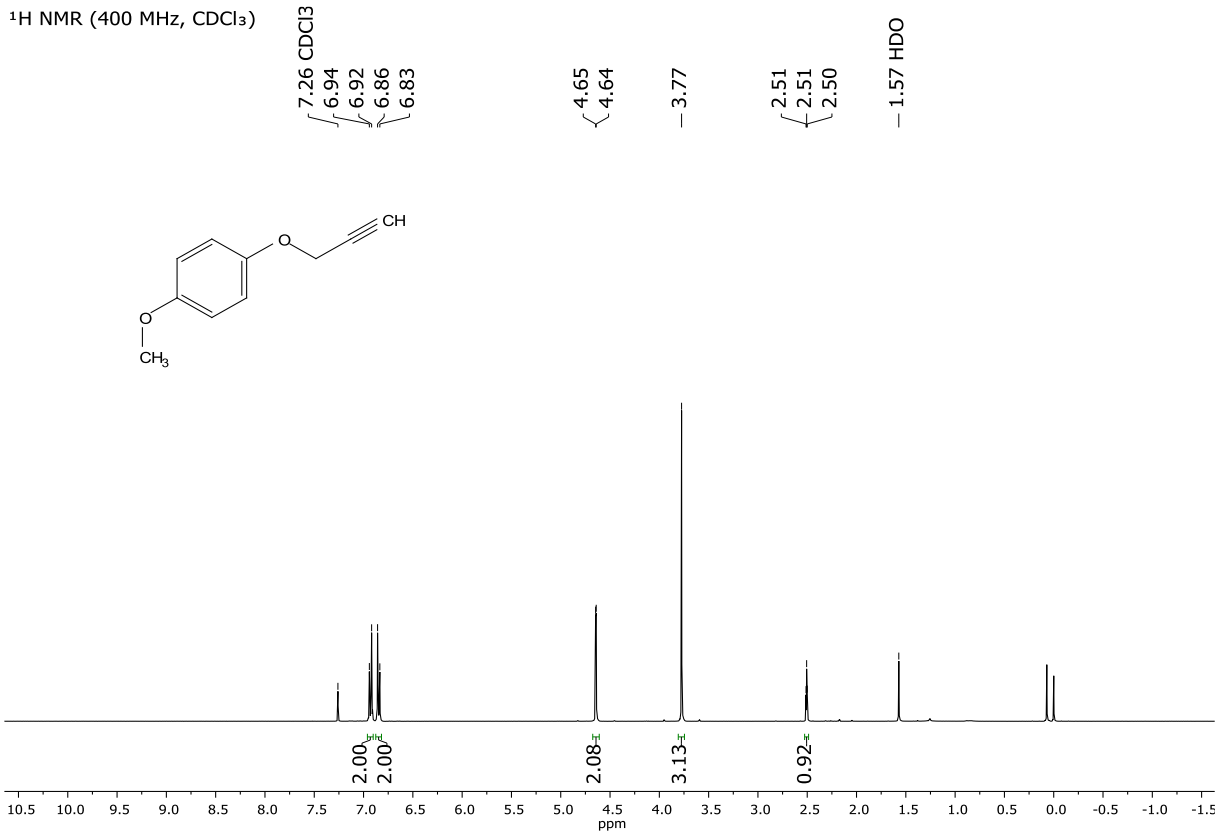
FTIR:


Transmittance
Wavenumbers [cm⁻¹]
3306
3072
3026
2967
2872
2324
1703
1651
1613
1593
1491
1376
1310
1298
1256
1196
1180
1117
1051
1018
1008
939
919
839
819
780
744
670.7
638.1

1-(Prop-2-yn-1-yloxy)-4-methoxybenzene (**2g**)

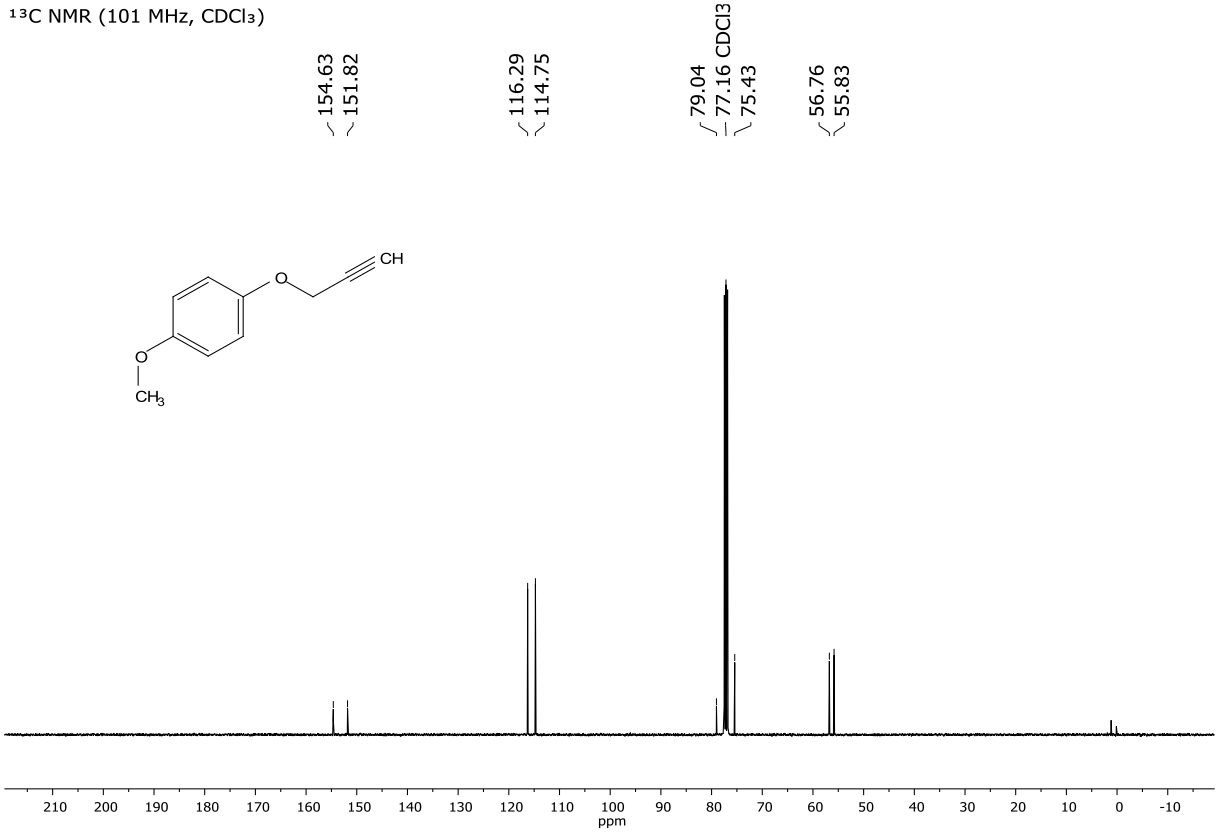
^{1}H NMR:

^{1}H NMR (400 MHz, $CDCl_3$)

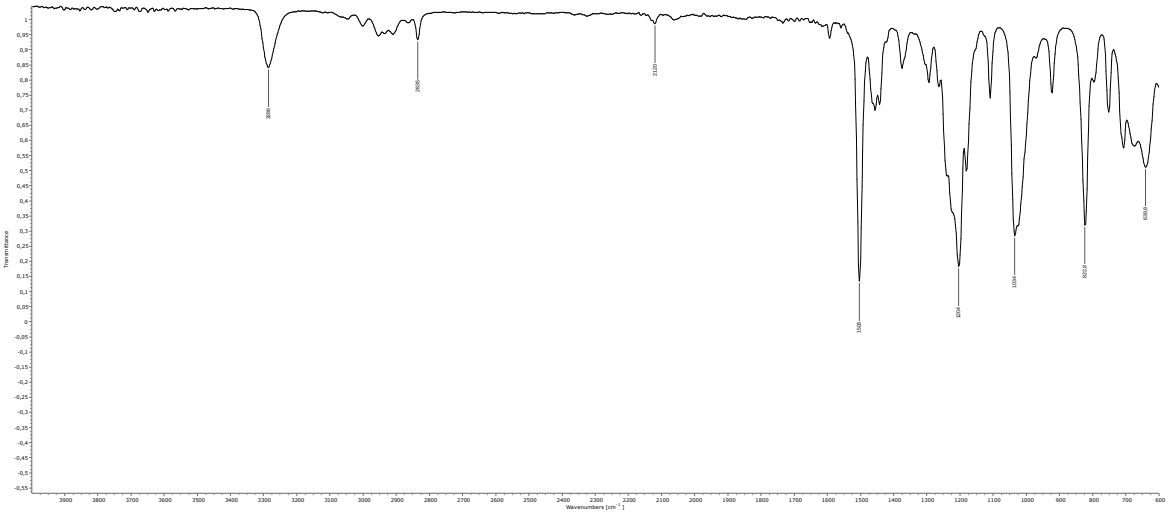


^{13}C NMR:

^{13}C NMR (101 MHz, $CDCl_3$)

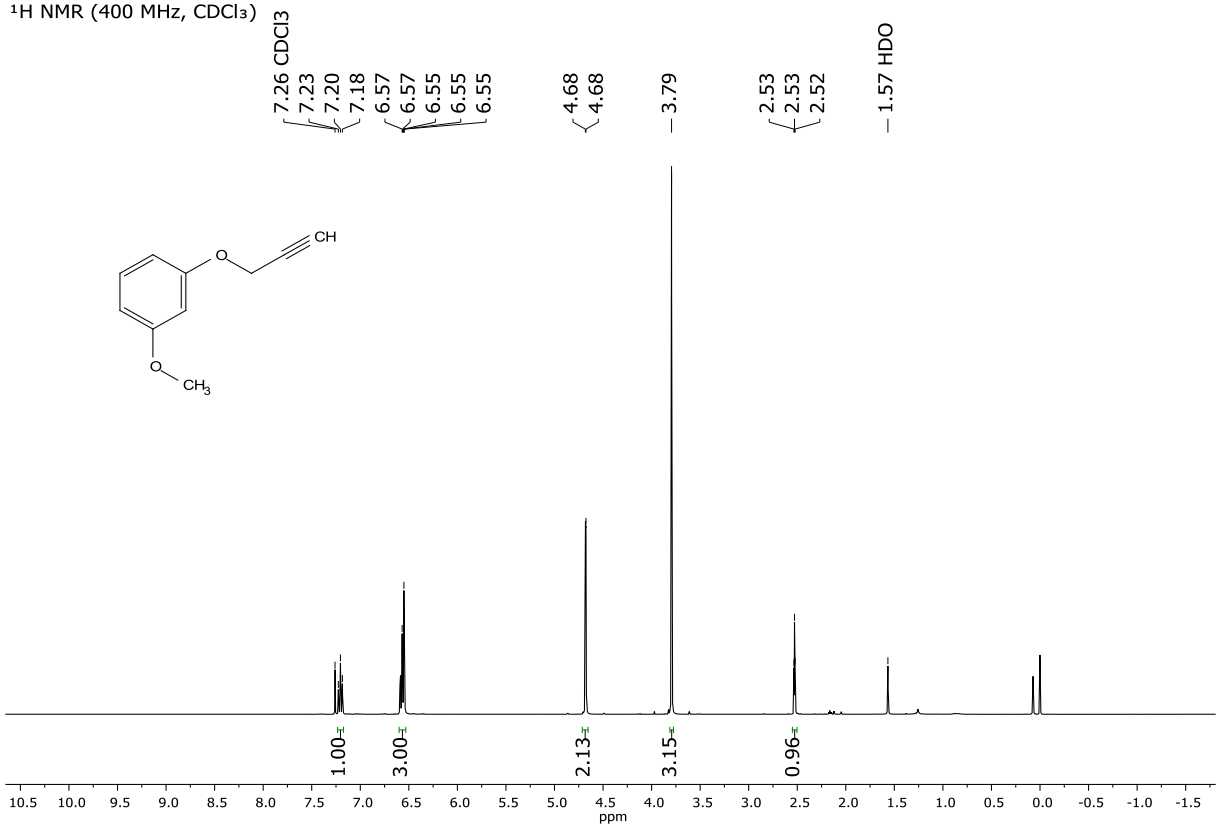


FTIR:

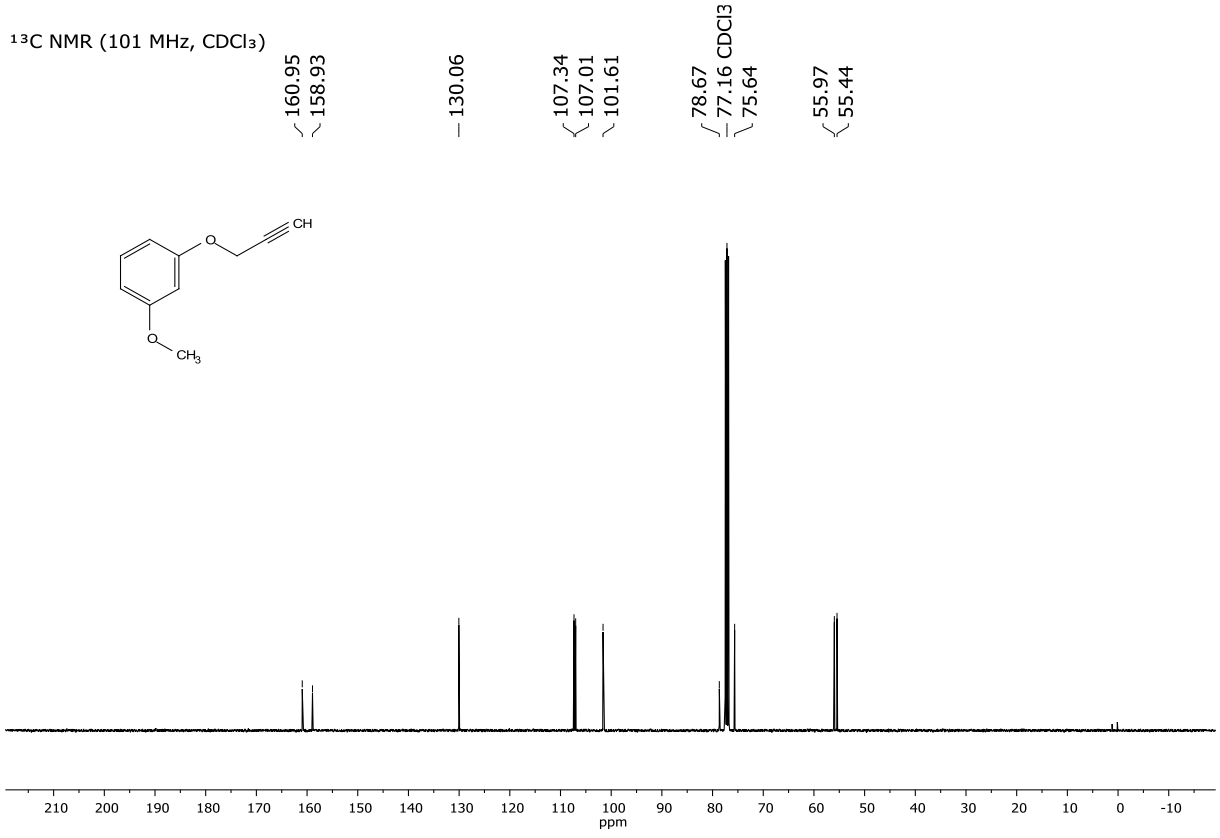

Transmittance
Wavenumbers [cm⁻¹]
3300
2935
2120
1500
1040
1031
820.0
630.0

1-(Prop-2-yn-1-yloxy)-3-methoxybenzene (**2h**)

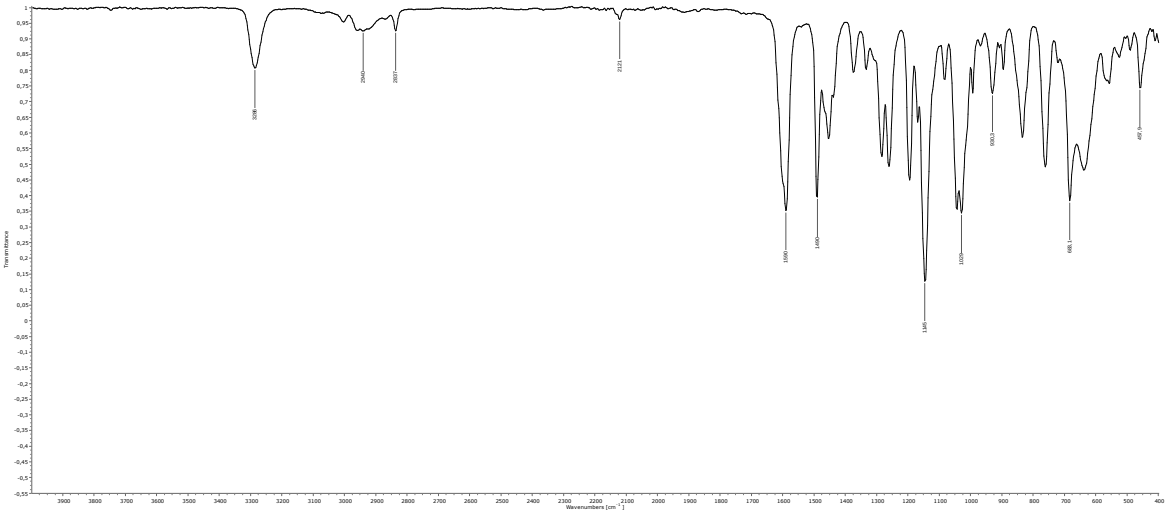
^{1}H NMR:



^{13}C NMR:



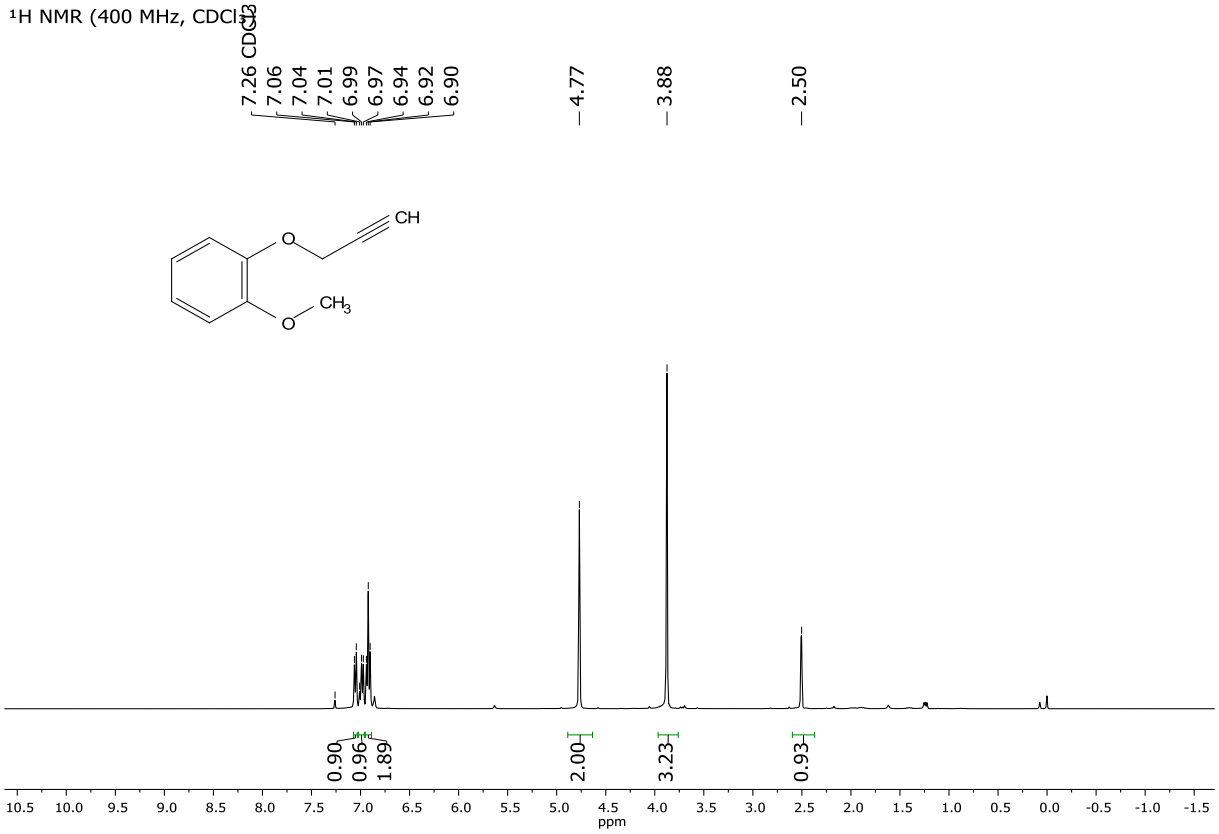
FTIR:


Transmittance
Wavenumbers [cm⁻¹]
3308
2940
2837
2121
1597
1450
1145
1029
699
489

1-(Prop-2-yn-1-yloxy)-2-methoxybenzene (**2i**)

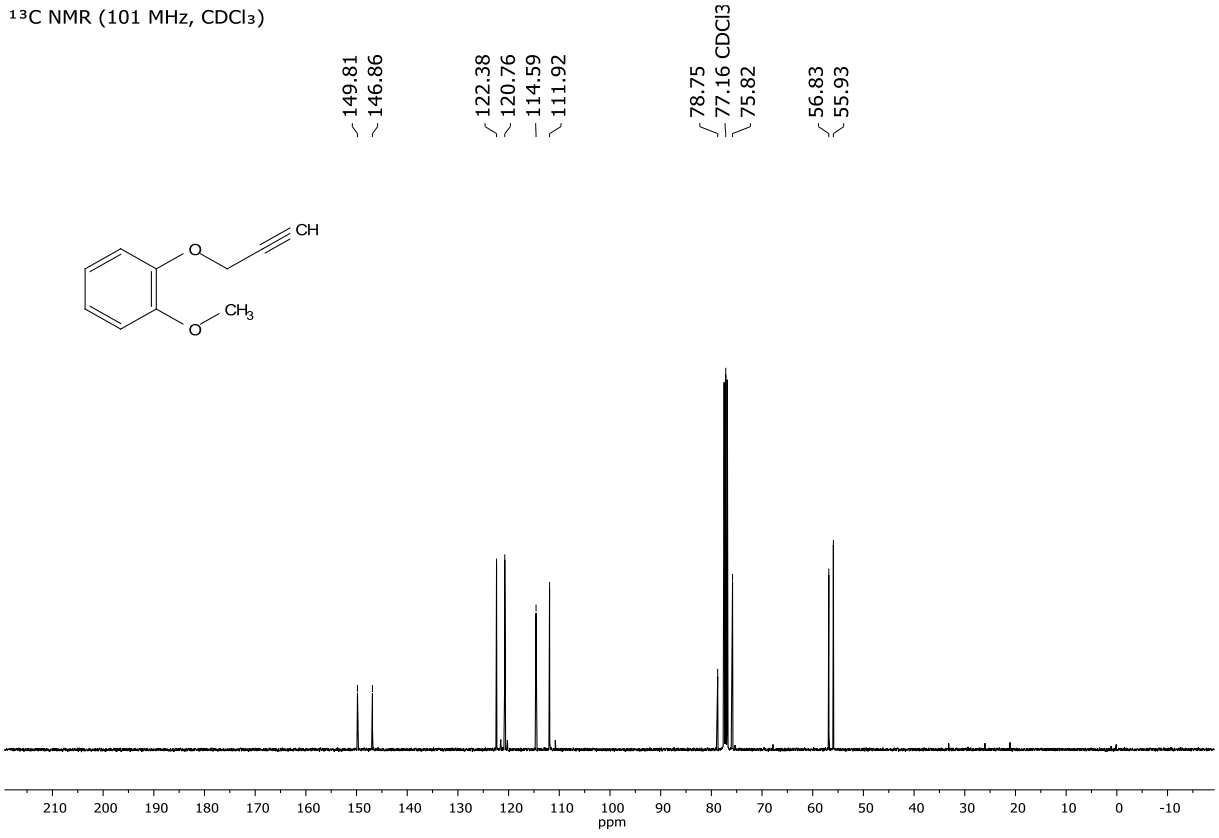
^{1}H NMR:

^{1}H NMR (400 MHz, $CDCl_3$)

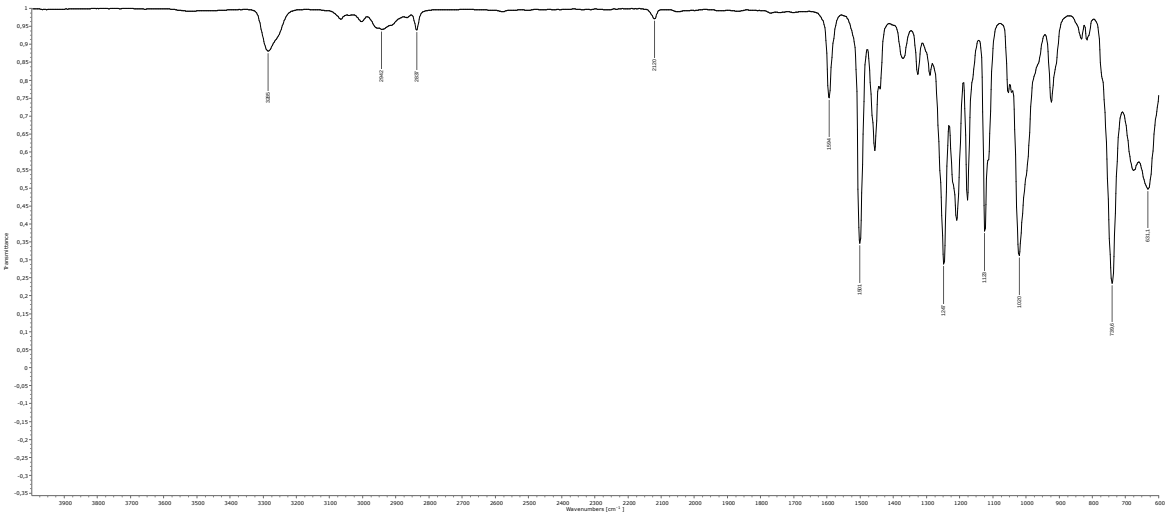


^{13}C NMR:

^{13}C NMR (101 MHz, $CDCl_3$)



FTIR:

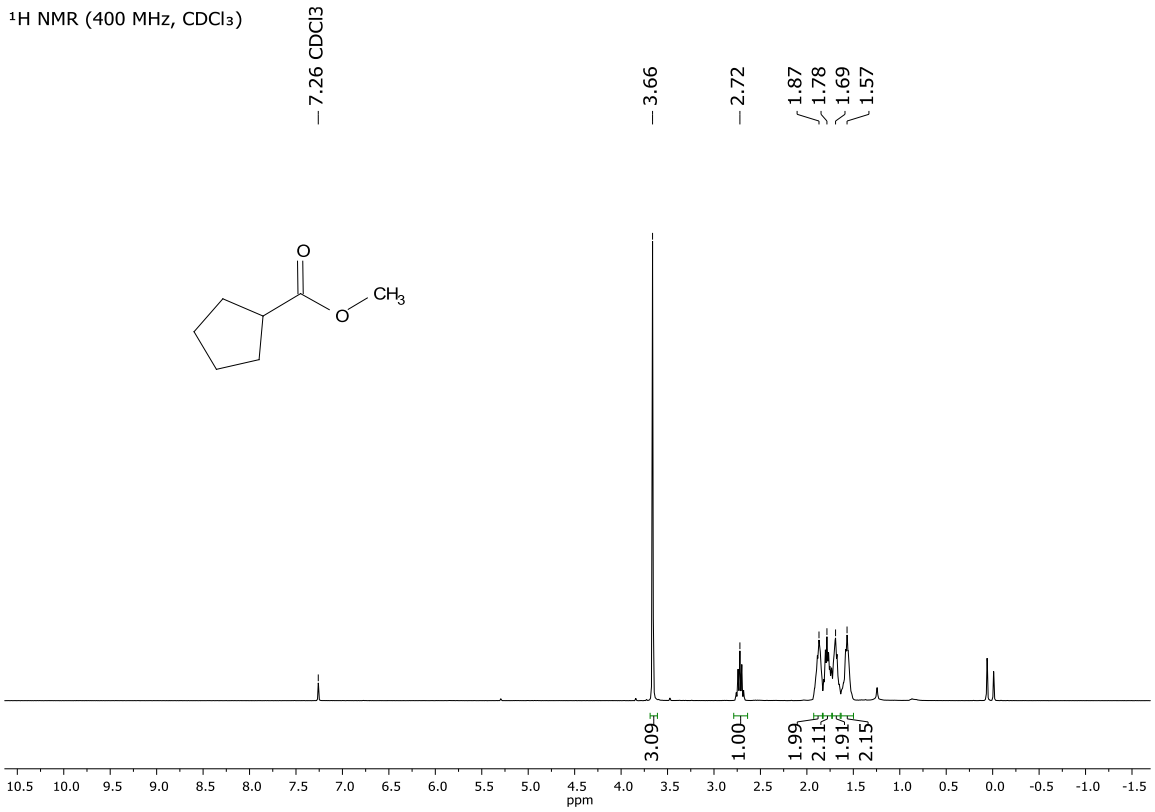

Transmittance
Wavenumbers [cm⁻¹]
3285
2942
2858
2120
1604
1381
1209
1118
1020
799.0
631.1

Synthesis of aliphatic Esters (**4**)

Methyl cyclopentane carboxylate (**4a**)

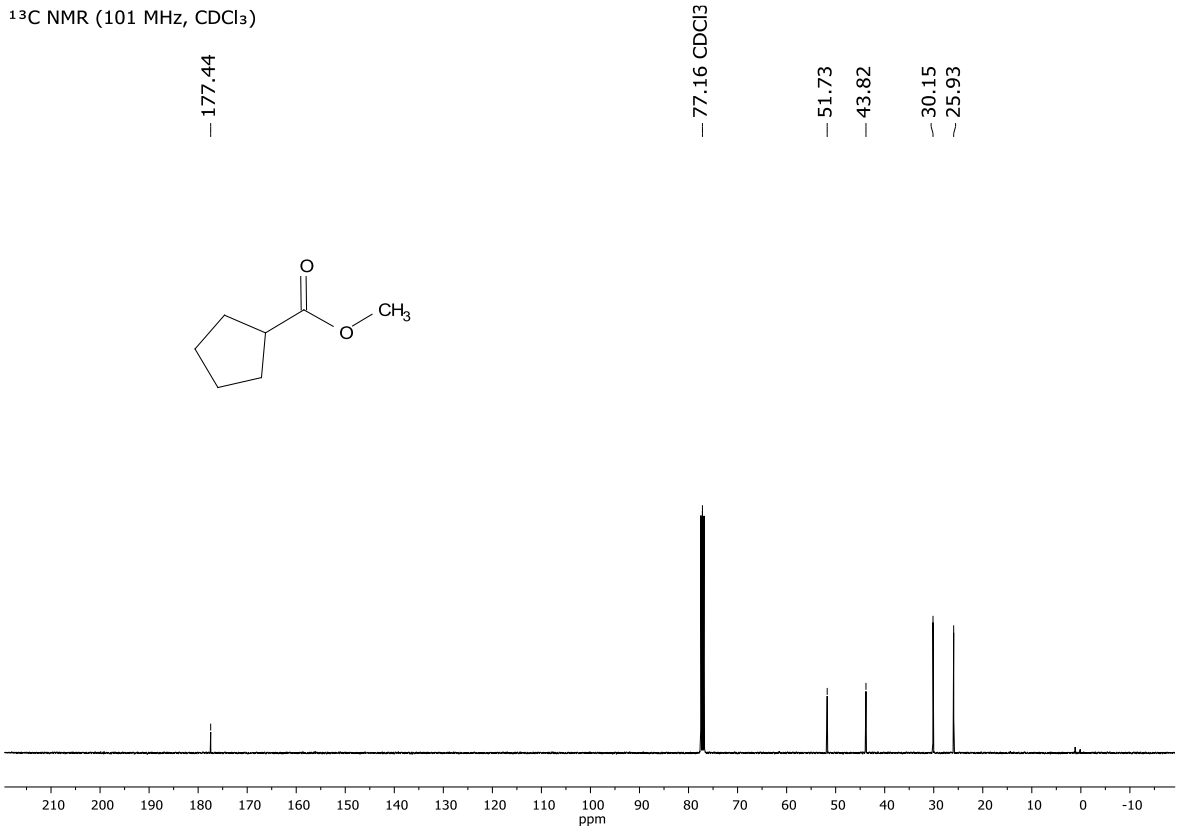
^{1}H NMR:

^{1}H NMR (400 MHz, $CDCl_3$)

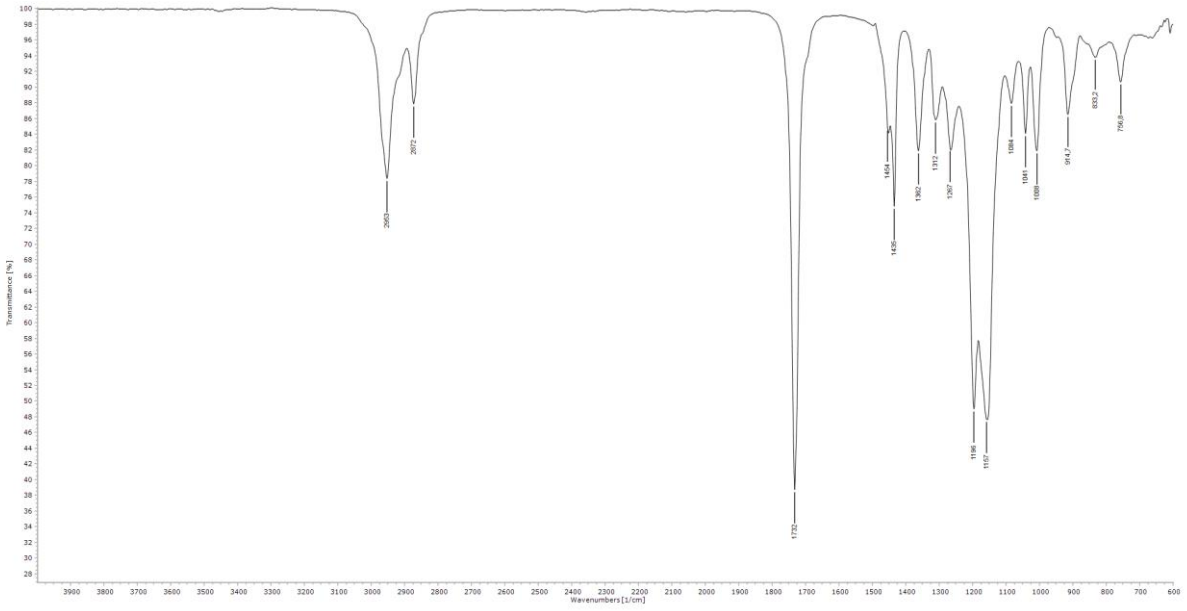


^{13}C NMR:

^{13}C NMR (101 MHz, $CDCl_3$)



FTIR:

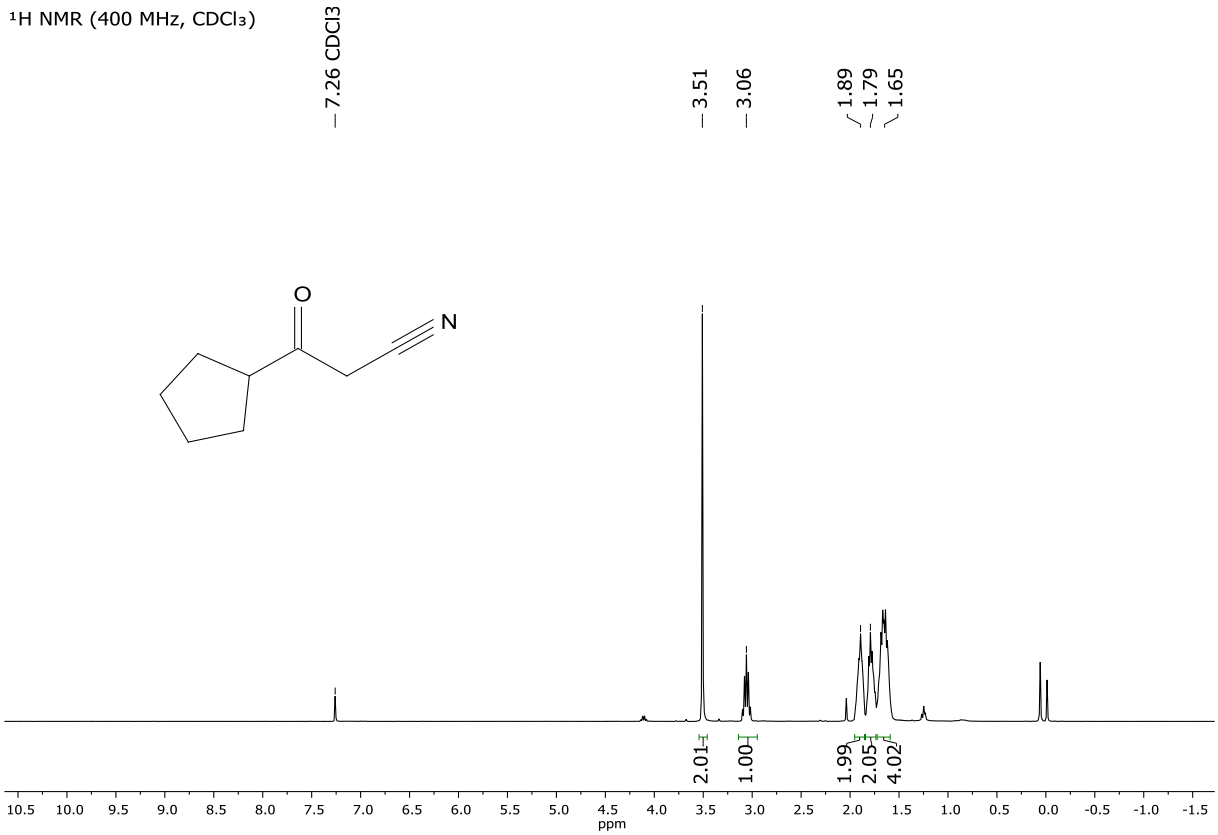

Transmittance [%]
Wavenumbers [1/cm]
2953
2872
1732
1454
1435
1382
1312
1267
1196
1107
1084
1041
1008
914.7
833.2
756.8

Synthesis of Beta-ketonitriles (3-oxopropanenitriles) (**5**)

3-Cyclopentyl-3-oxopropanenitrile (**5a**)

^{1}H NMR:

^{1}H NMR (400 MHz, $CDCl_3$)

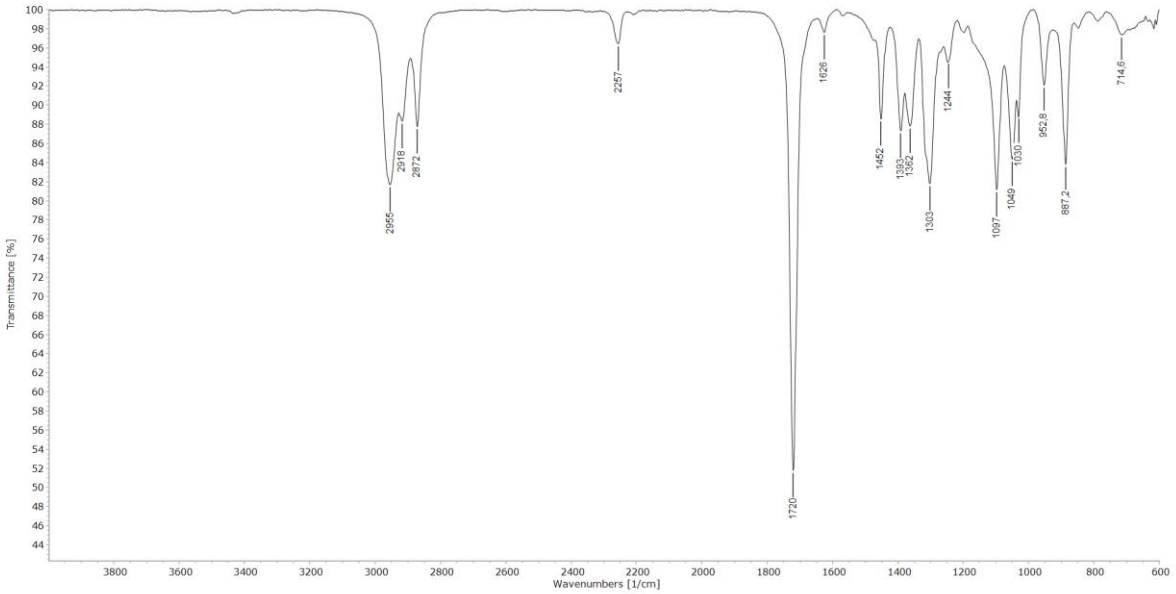


^{13}C NMR

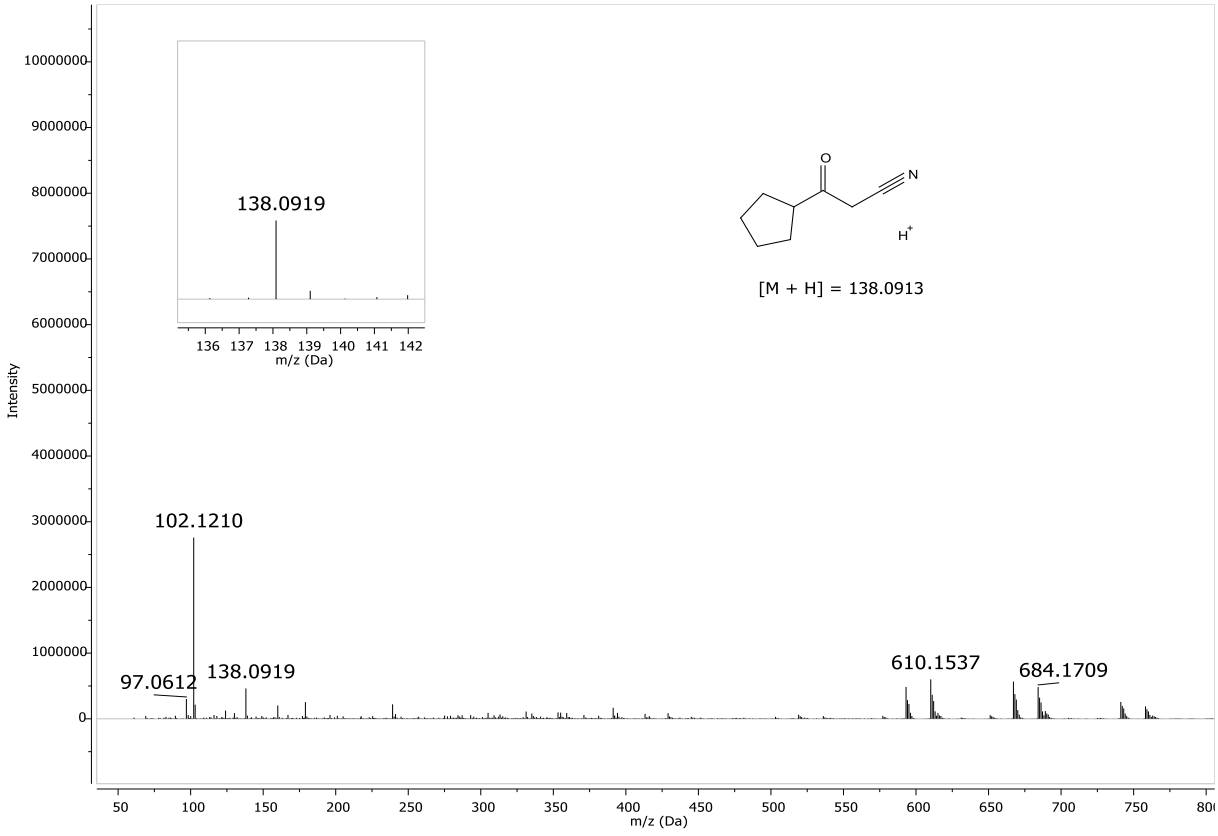
^{13}C NMR (101 MHz, $CDCl_3$)



FTIR:


Transmittance [%]
Wavenumbers [1/cm]
2955
2918
2872
2257
1720
1626
1452
1393
1362
1303
1244
1097
1049
1030
952.8
887.2
714.6

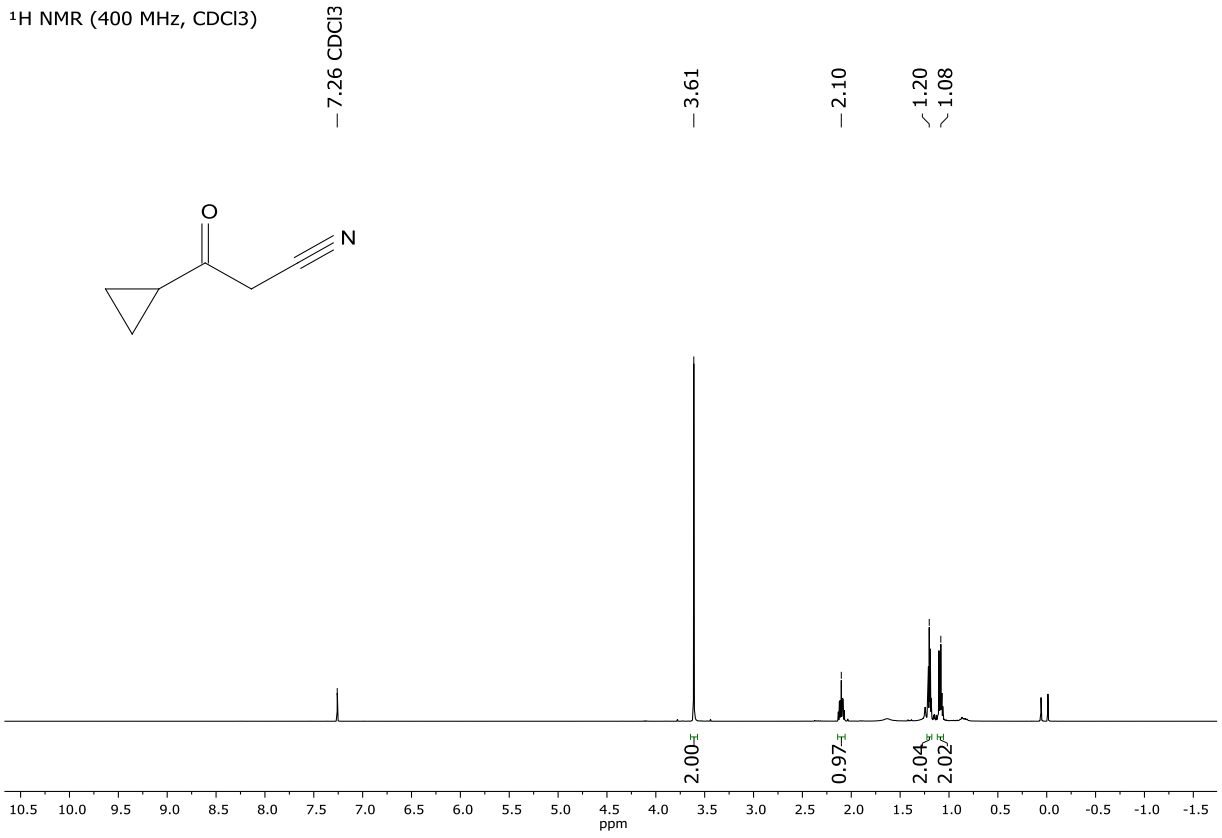
HRMS:


Intensity
m/z (Da)
138.0919
102.1210
97.0612
610.1537
684.1709
[M + H] = 138.0913

3-Cyclopropyl-3-oxopropanenitrile (**5b**)

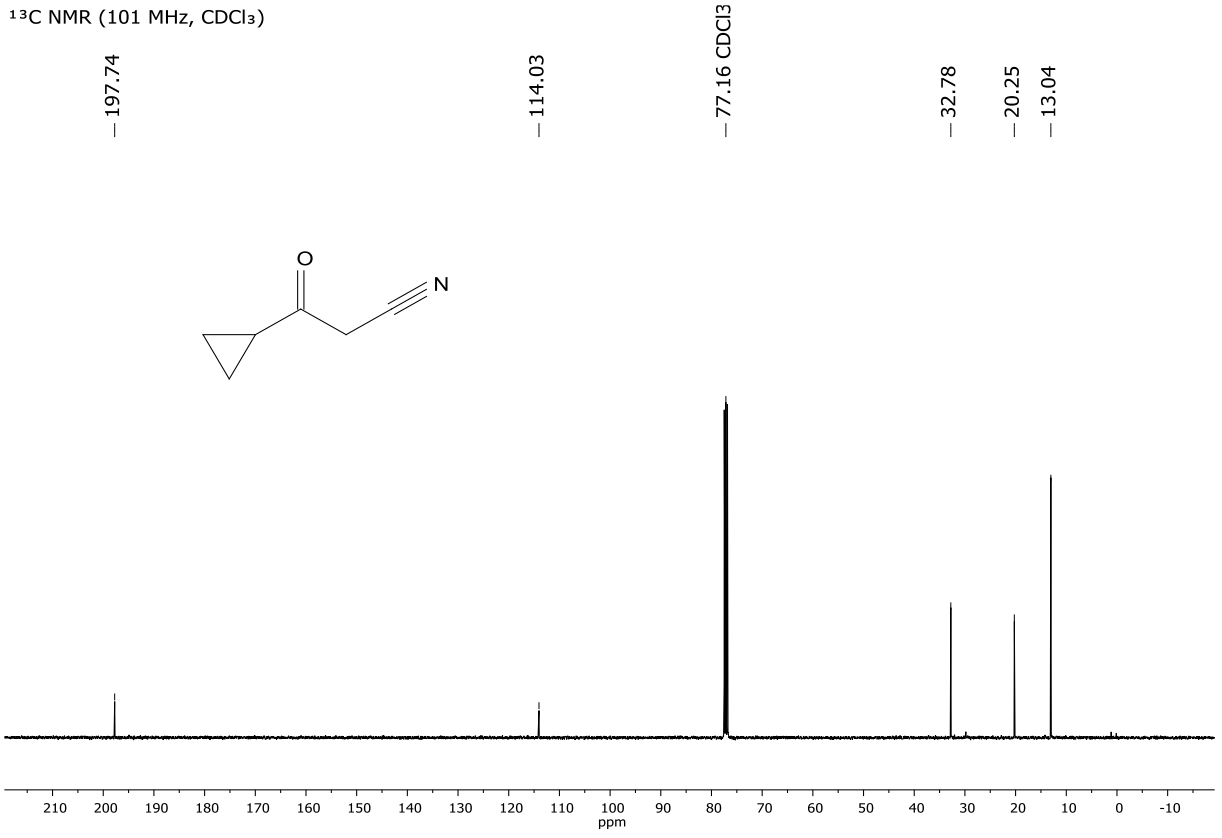
^{1}H NMR:

^{1}H NMR (400 MHz, CDCl3)

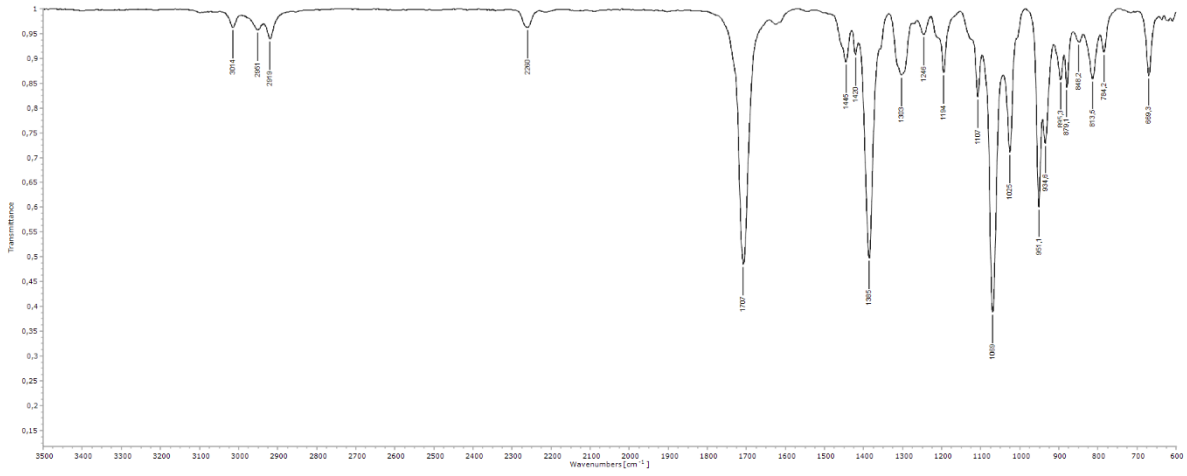


^{13}C NMR:

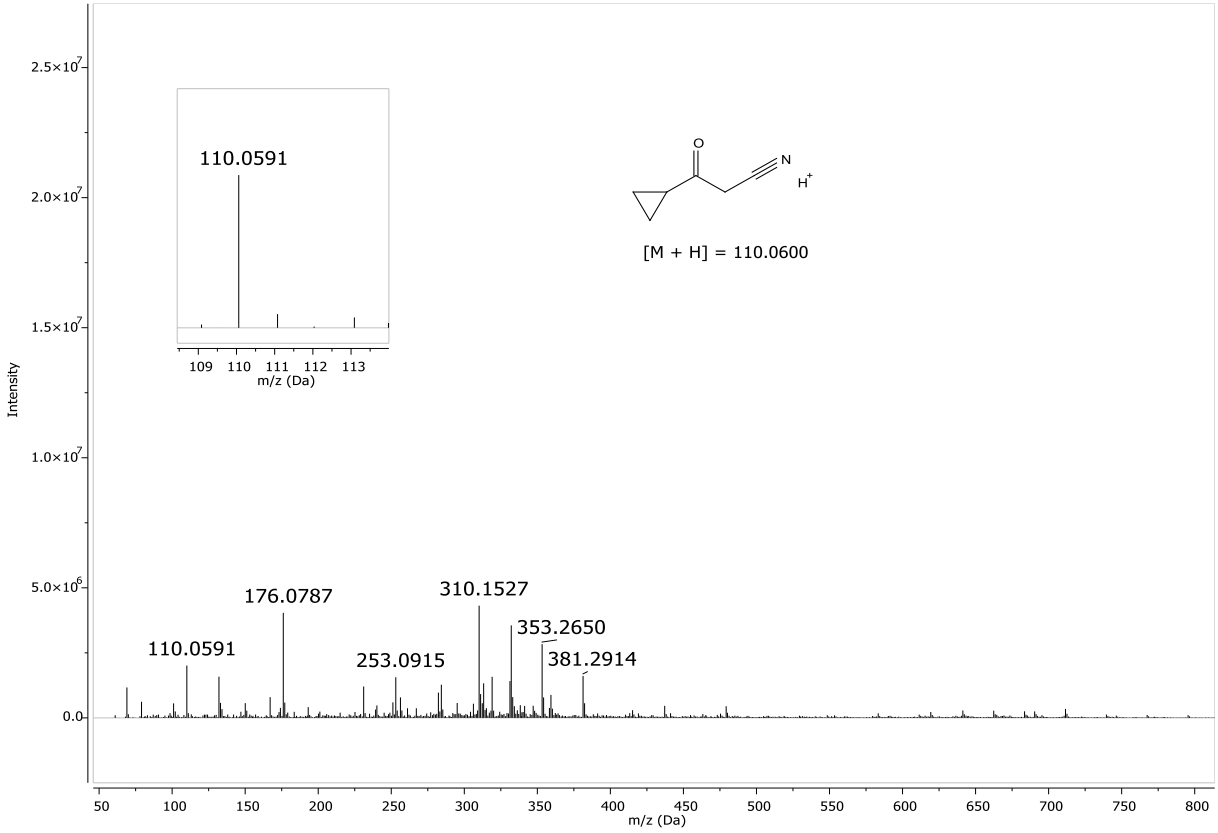
^{13}C NMR (101 MHz, $CDCl_3$)



FTIR:


Transmittance
Wavenumbers [cm⁻¹]

HRMS:

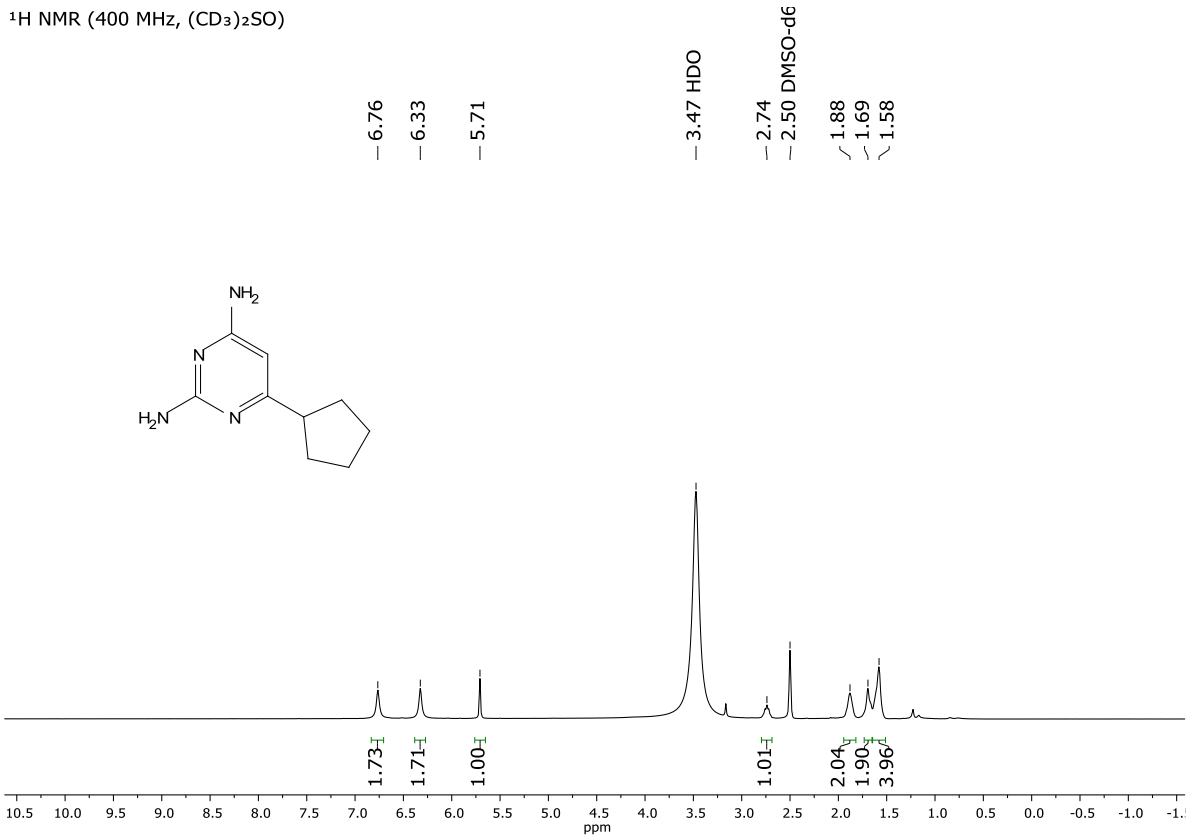

Intensity
m/z (Da)
110.0591
176.0787
310.1527
353.2650
381.2914
253.0915
110.0591
[M + H] = 110.0600

Synthesis of 2,4–Diaminopyrimidines (**6**)

6-Cyclopentylpyrimidine-2,4-diamine (**6a**)

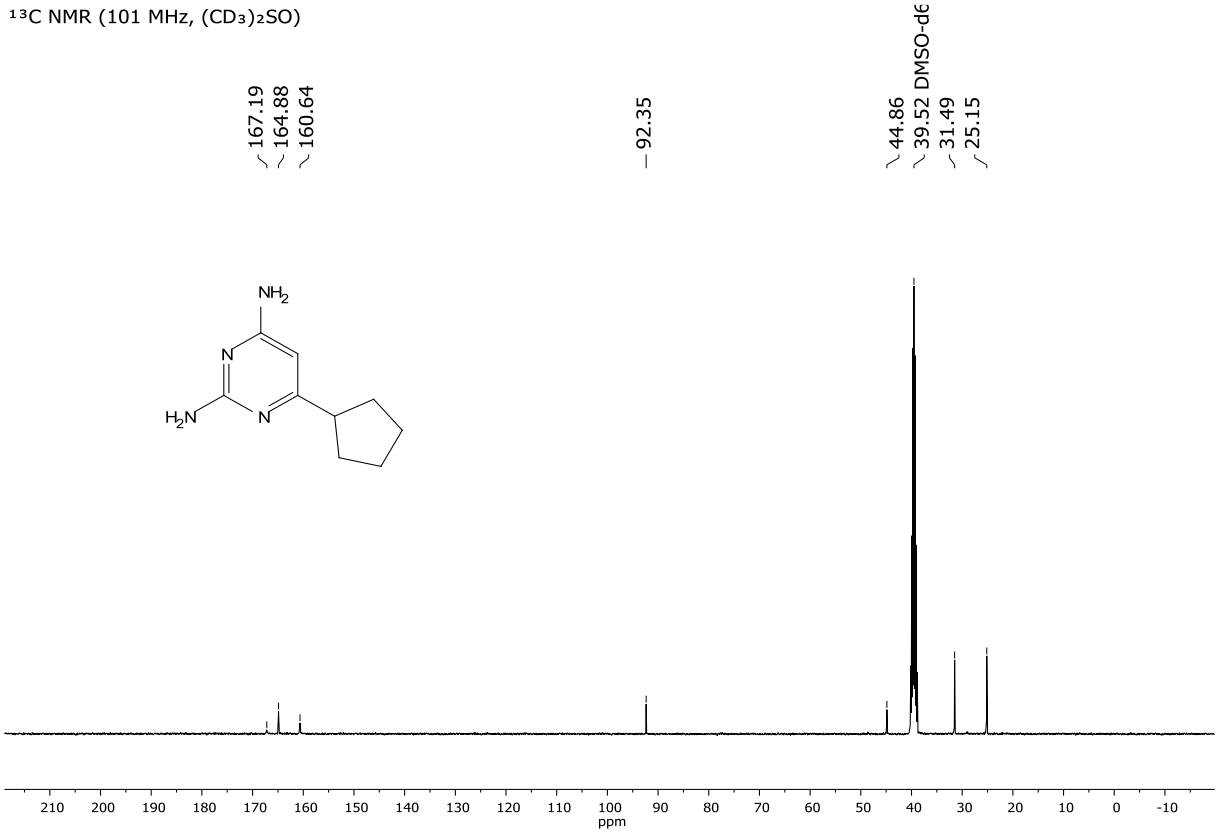
^{1}H NMR:

^{1}H NMR (400 MHz, $(CD_3)_2SO$)

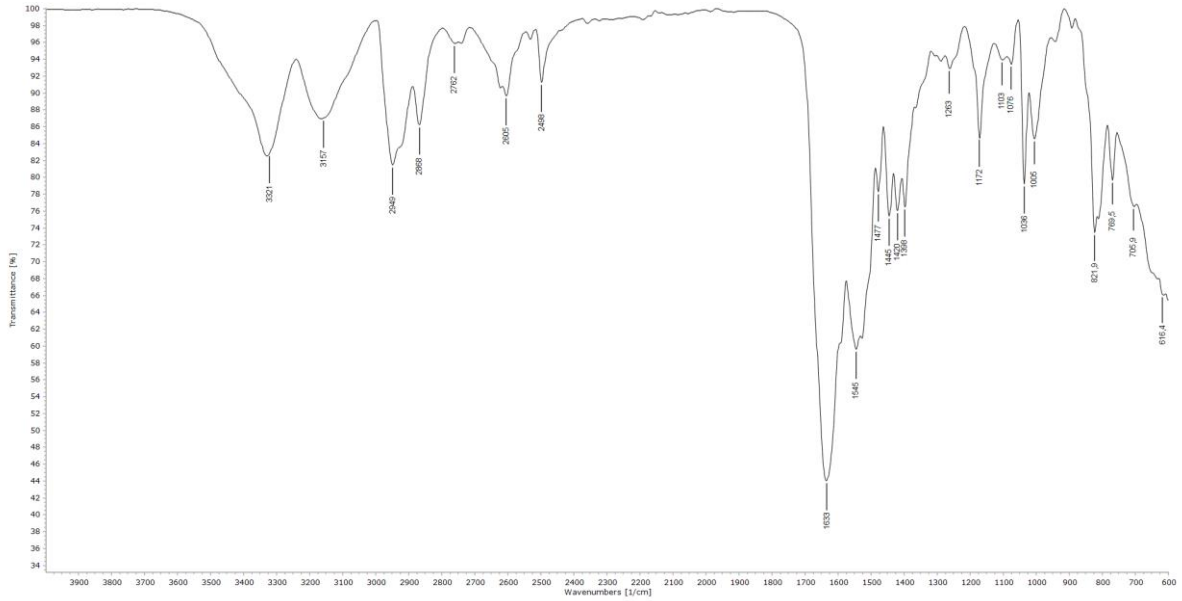


^{13}C NMR

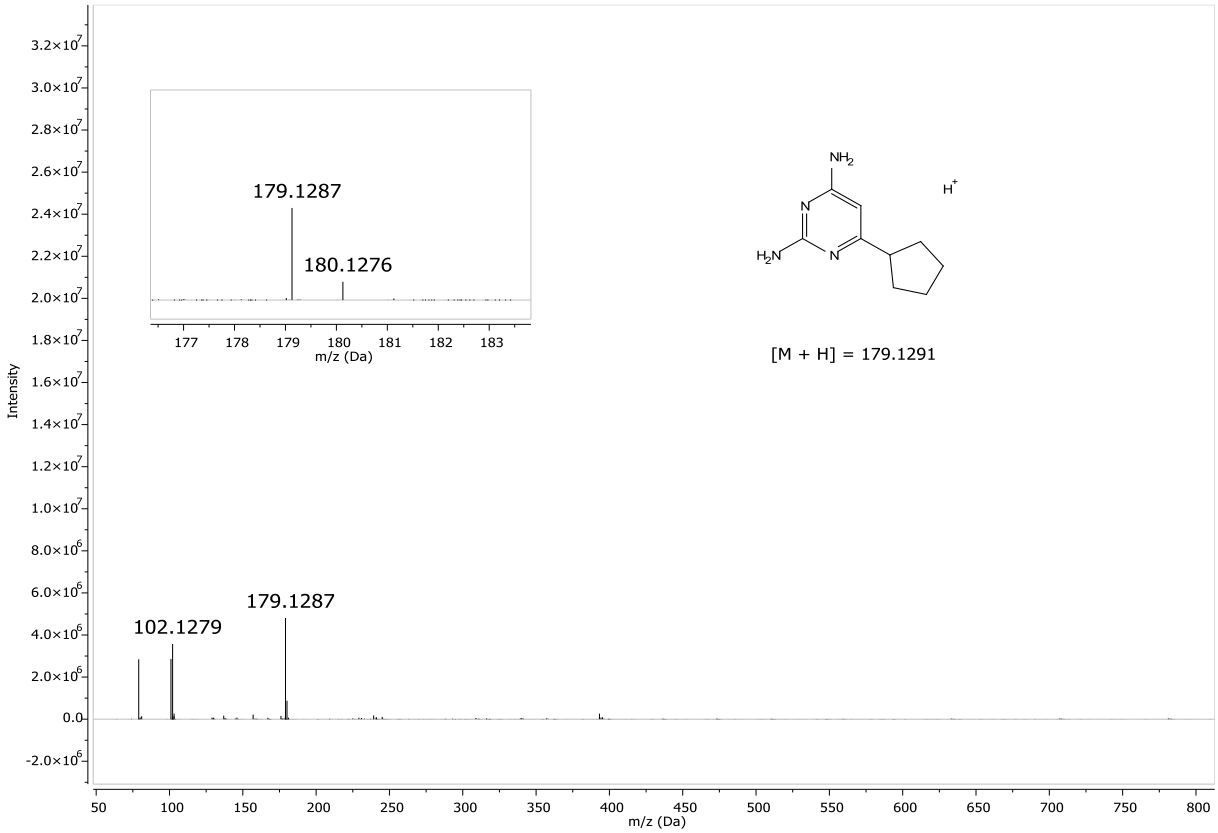
^{13}C NMR (101 MHz, $(CD_3)_2SO$)



FTIR:


Transmittance [%]
Wavenumbers [1/cm]
3321
3157
2949
2868
2702
2606
2468
1633
1546
1477
1445
1430
1366
1263
1172
1103
1076
1036
1005
821.9
769.5
705.9
616.4

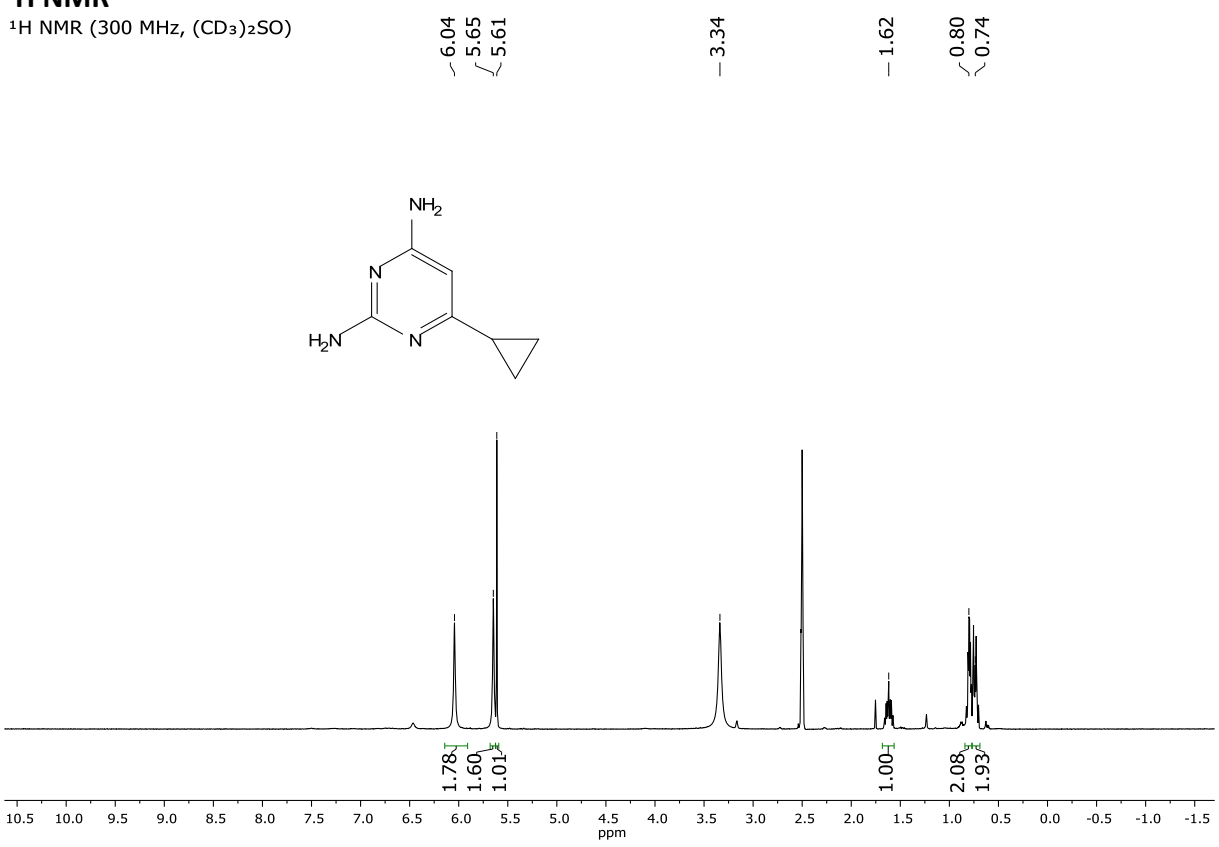
HRMS:


Intensity
m/z (Da)
179.1287
180.1276
102.1279
179.1287
NH2
H2N
N
N
H+
[M + H] = 179.1291

6-Cyclopropylpyrimidine-2,4-diamine (**6b**)

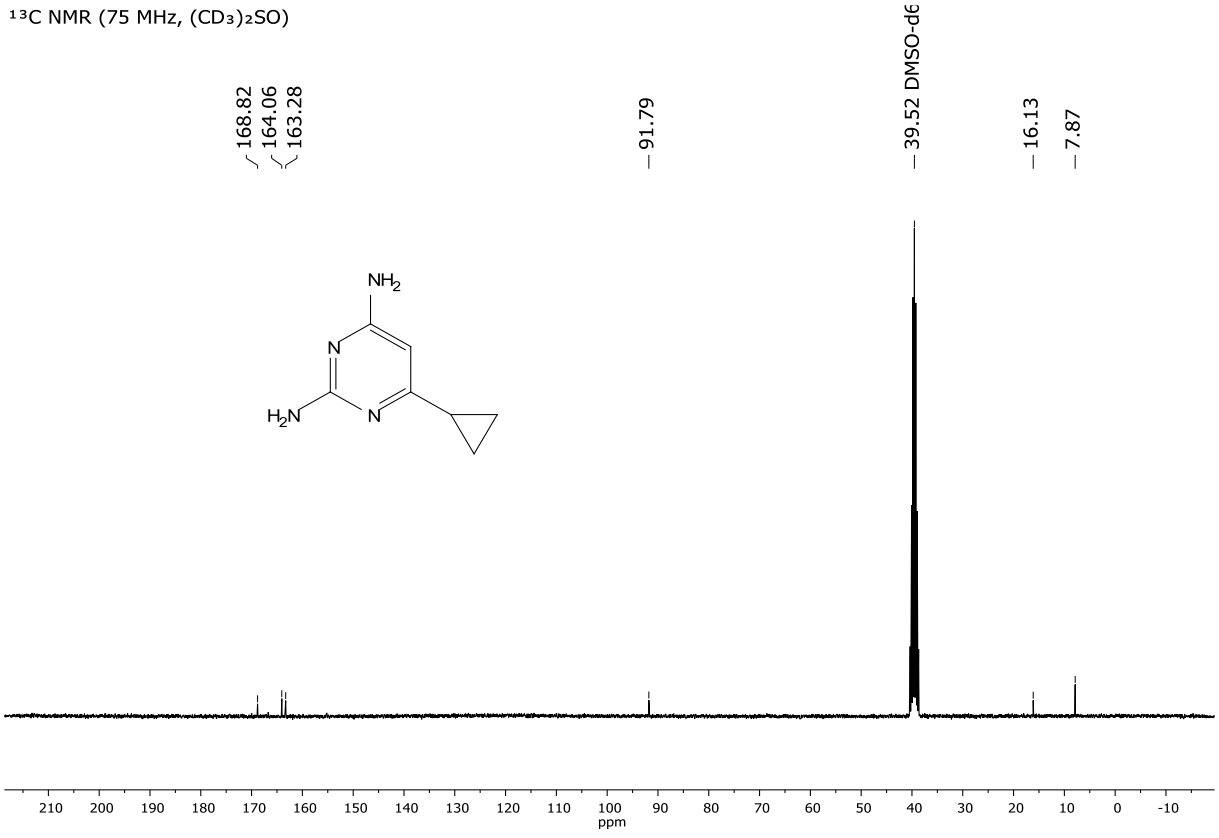
^{1}H NMR

^{1}H NMR (300 MHz, $(CD_3)_2SO$)

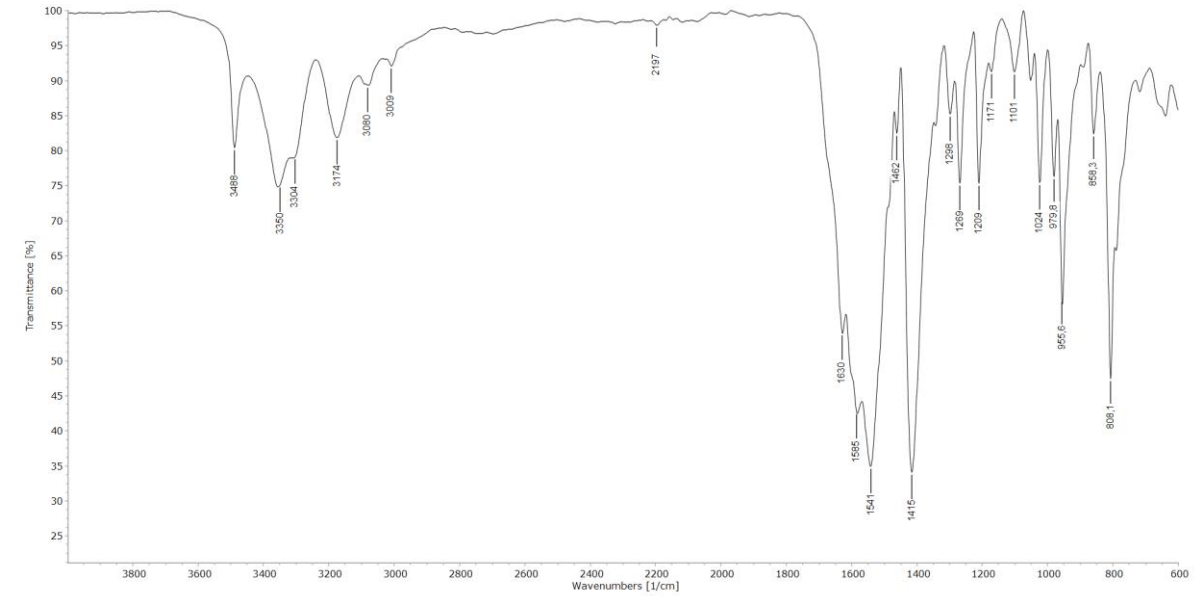


^{13}C NMR:

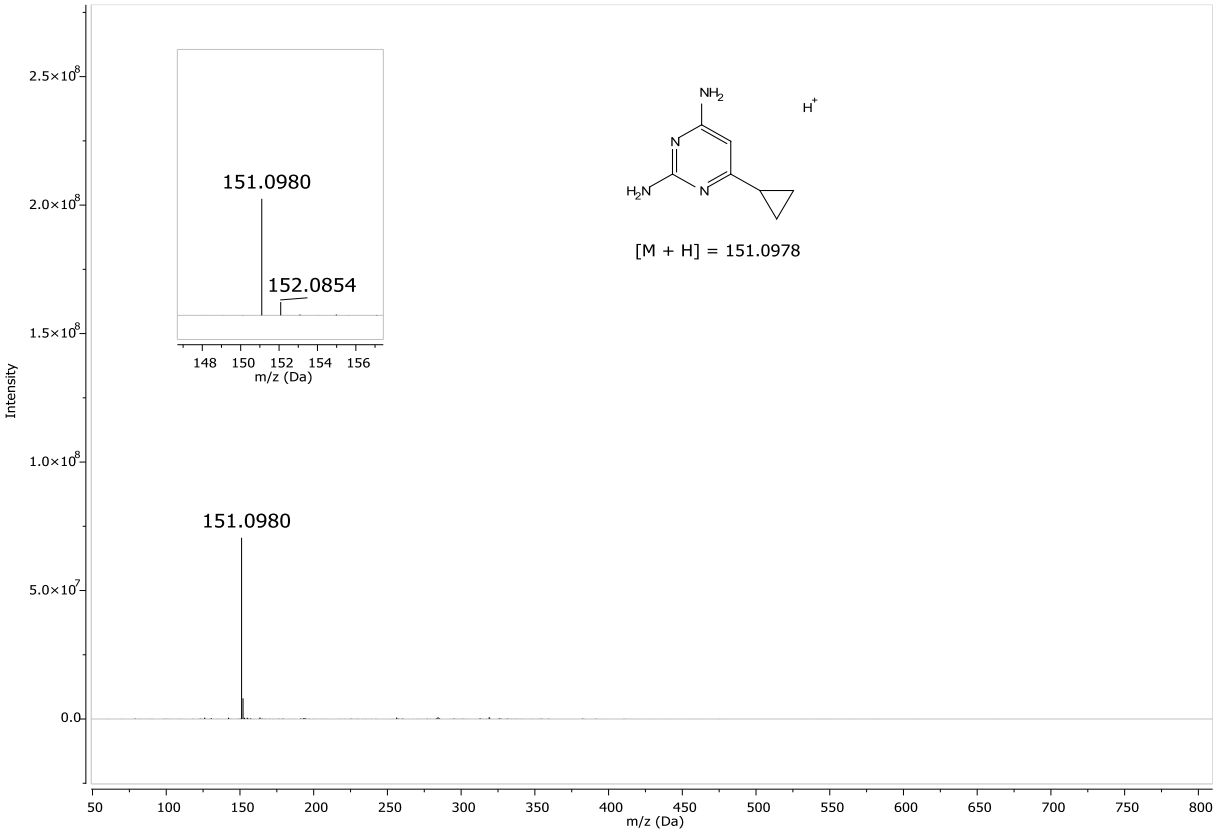
^{13}C NMR (75 MHz, $(CD_3)_2SO$)



FTIR:


Transmittance [%]
Wavenumbers [1/cm]
3488
3350
3304
3174
3080
3009
2197
1630
1585
1541
1462
1415
1298
1269
1209
1171
1101
1024
979.8
955.6
859.3
808.1

HRMS:

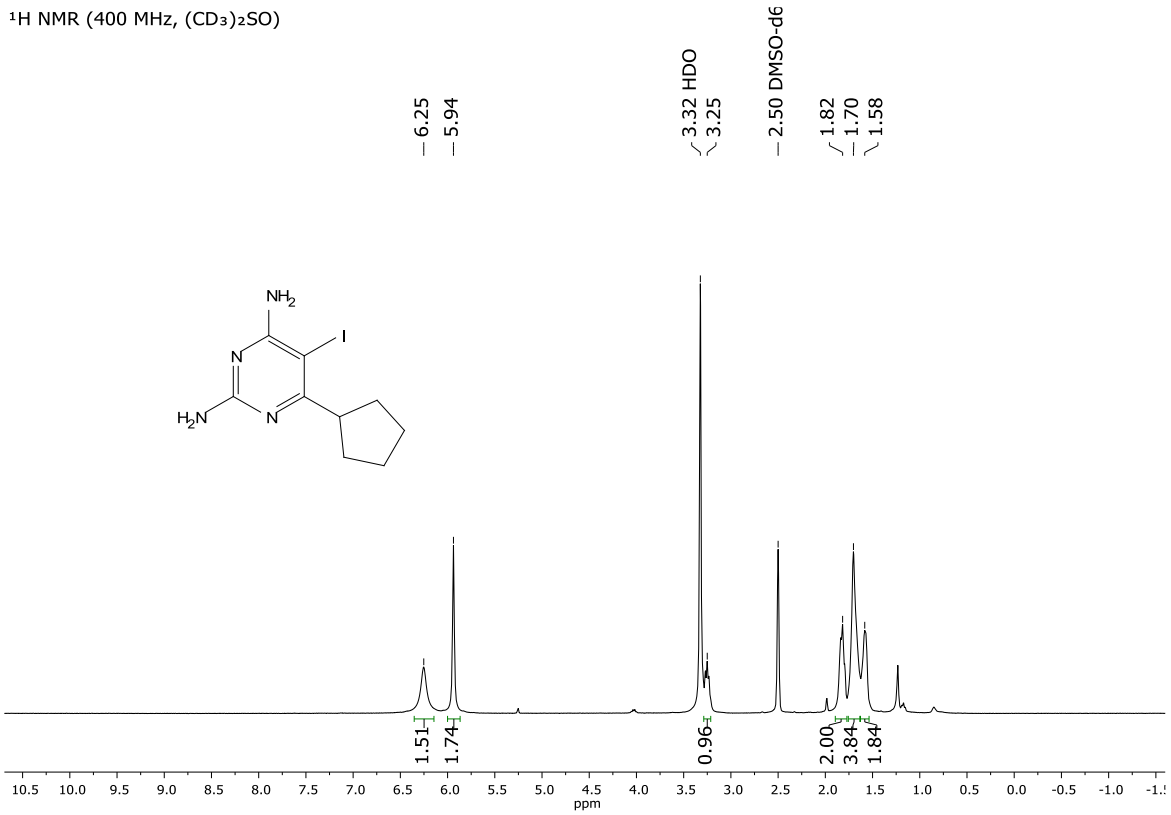

151.0980
152.0854
148 150 152 154 156
m/z (Da)
NH2
H+
N
H2N
N
[M + H] = 151.0978
Intensity
151.0980
50 100 150 200 250 300 350 400 450 500 550 600 650 700 750 800
m/z (Da)

Iodination of 2,4–Diamino–iodopyrimidines (**7**)

6-Cyclopentyl-5-iodopyrimidine-2,4-diamine (**7a**)

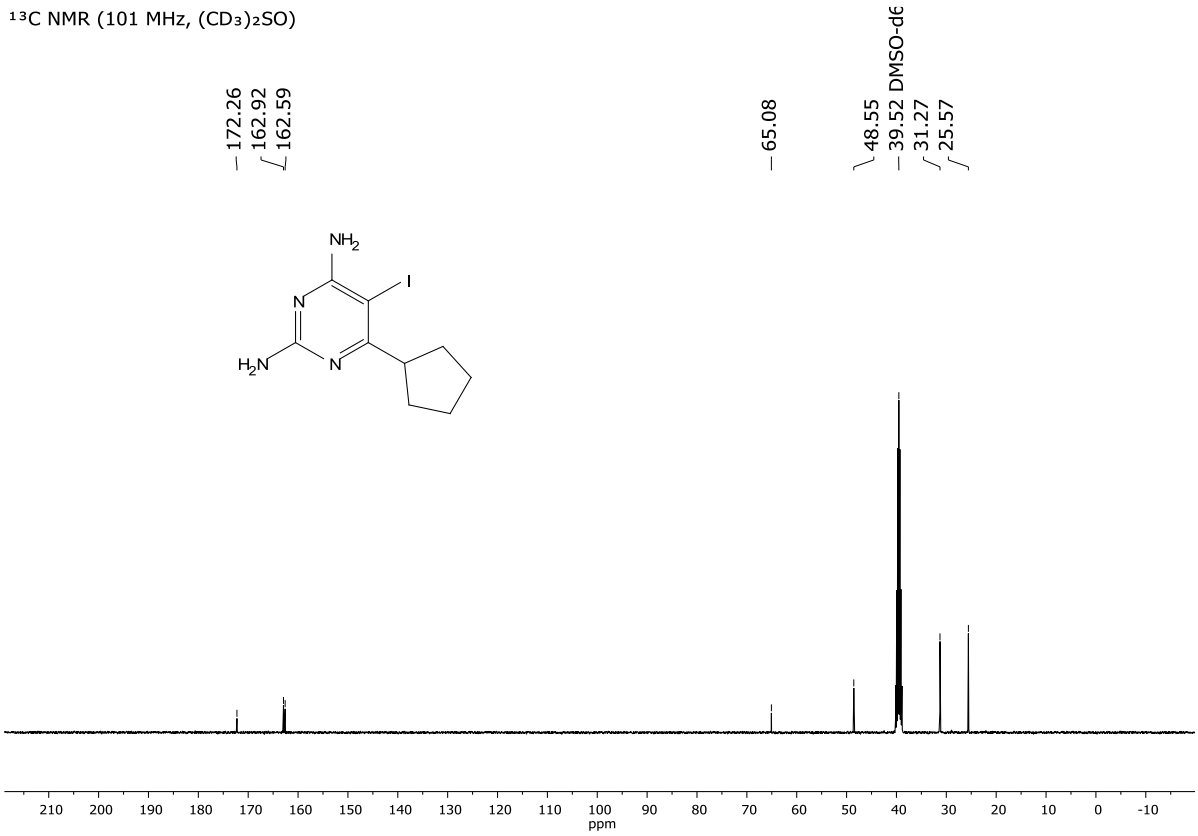
^{1}H NMR:

^{1}H NMR (400 MHz, $(CD_3)_2SO$)

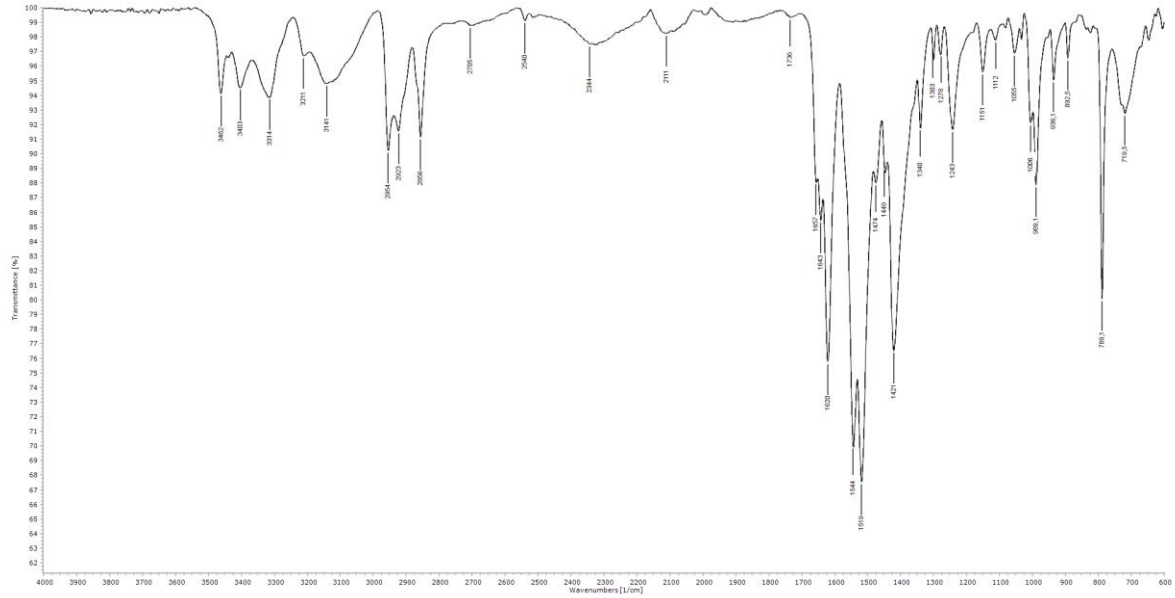


^{13}C NMR:

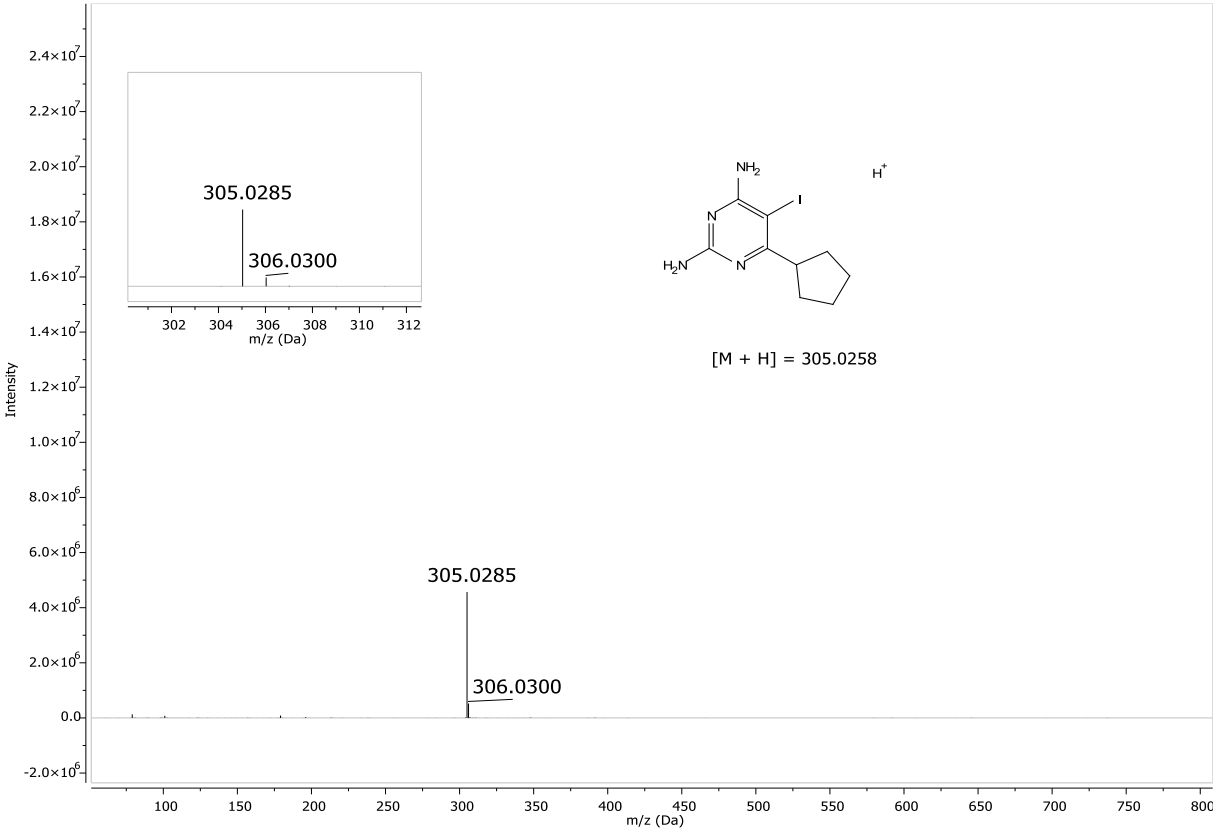
^{13}C NMR (101 MHz, $(CD_3)_2SO$)



FTIR:


Transmittance [%]
Wavenumbers [1/cm]

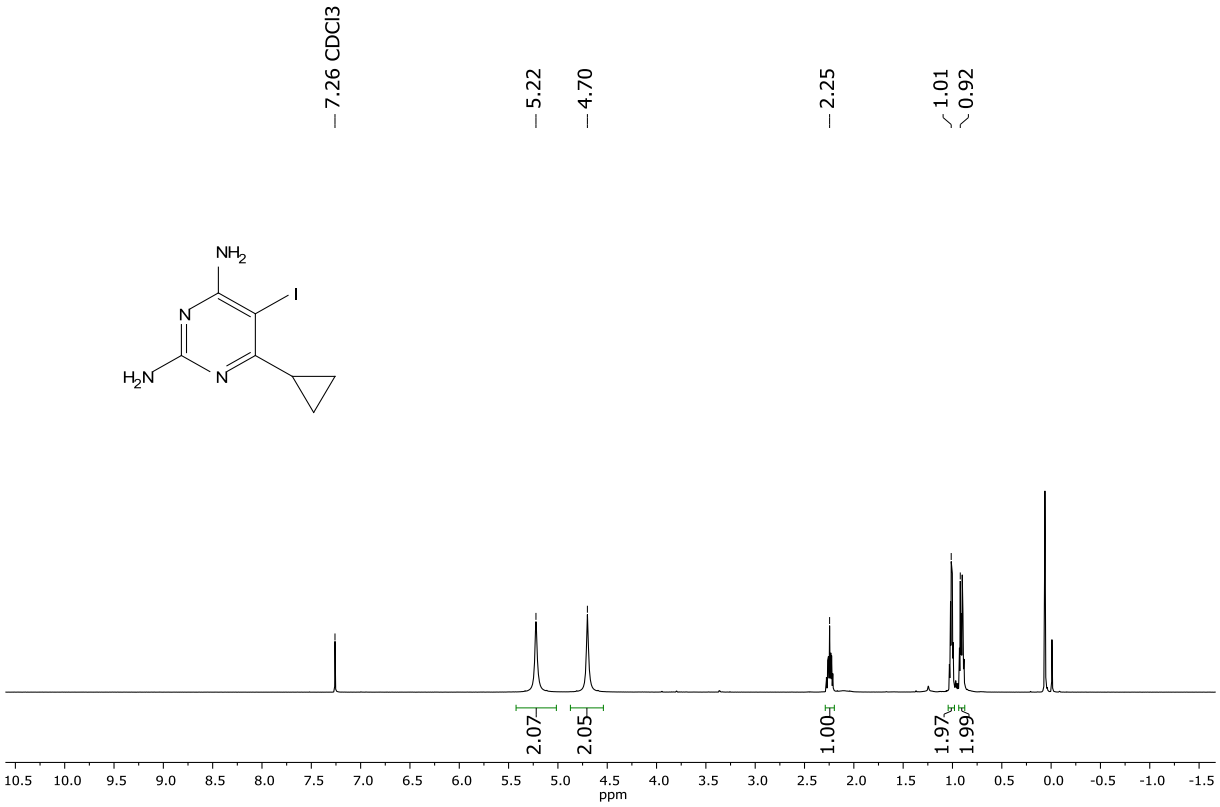
HRMS:


305.0285
306.0300
m/z (Da)
[M + H] = 305.0258
Intensity
305.0285
306.0300
m/z (Da)

6-Cyclopropyl-5-iodopyrimidine-2,4-diamine (**7b**)

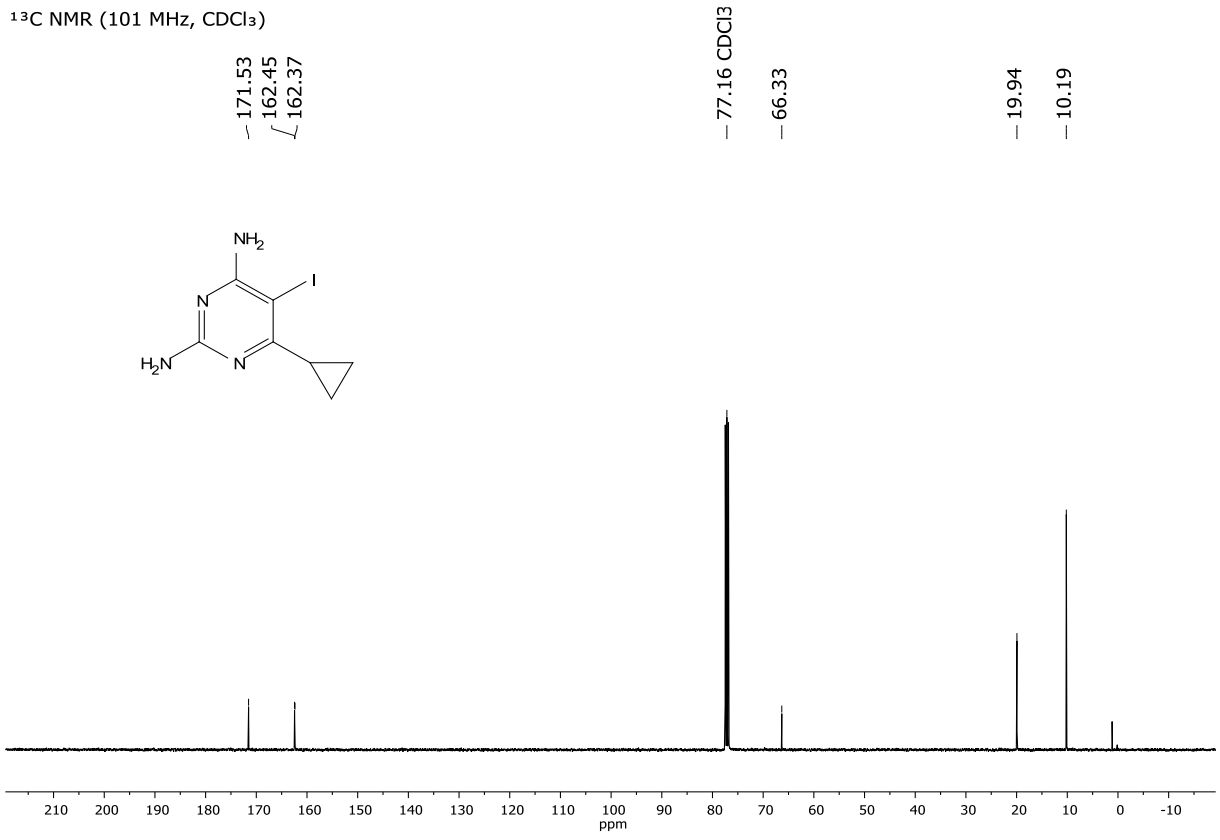
^{1}H NMR:

^{1}H NMR (400 MHz, $CDCl_3$)

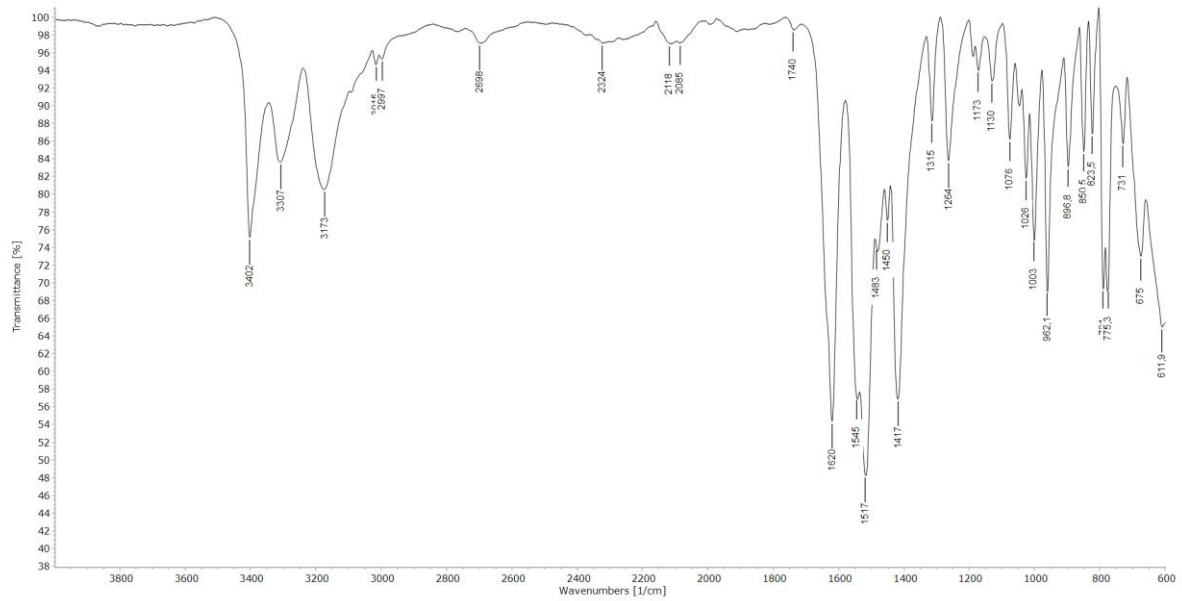


^{13}C NMR:

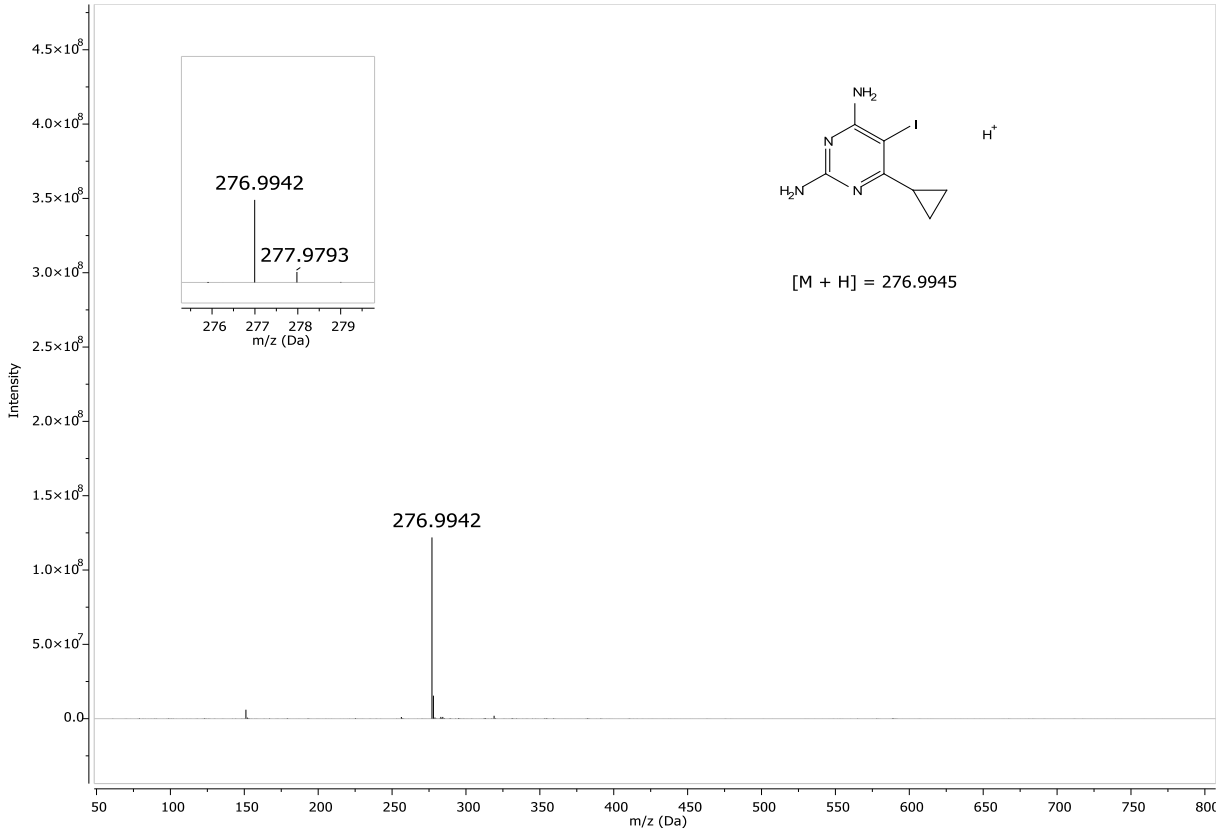
^{13}C NMR (101 MHz, $CDCl_3$)



FTIR:


Transmittance [%]
Wavenumbers [1/cm]
3402
3307
3173
2967
2698
2324
2118
2085
1740
1620
1545
1517
1483
1450
1417
1315
1264
1173
1130
1076
1026
1003
962.1
896.8
850.5
823.5
775.4
753.3
731
675
611.9

HRMS:

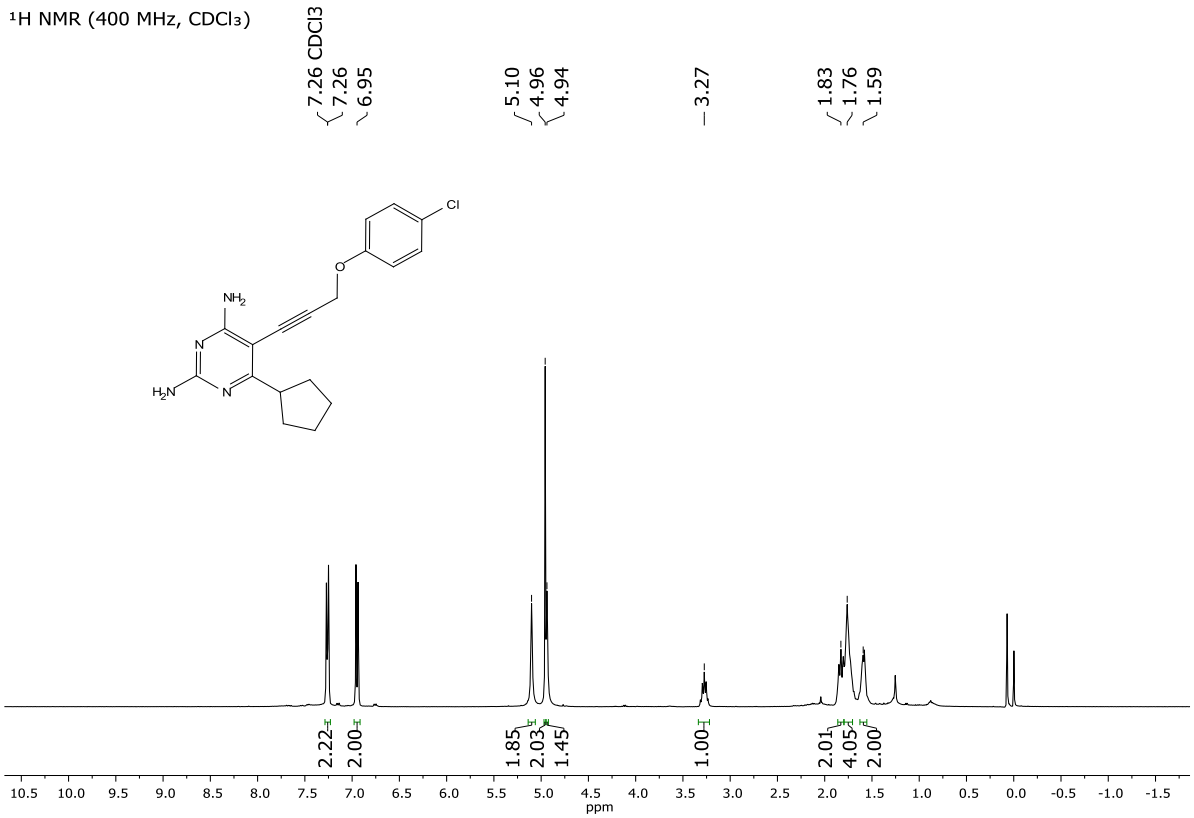

276.9942
277.9793
m/z (Da)
[M + H] = 276.9945
Intensity
276.9942
m/z (Da)

Sonogashira coupling reaction (**8**)

6-Cyclopentyl-5-[3-(4-chlorophenoxy)prop-1-yn-1-yl]pyrimidine-2,4-diamine (**8a**)

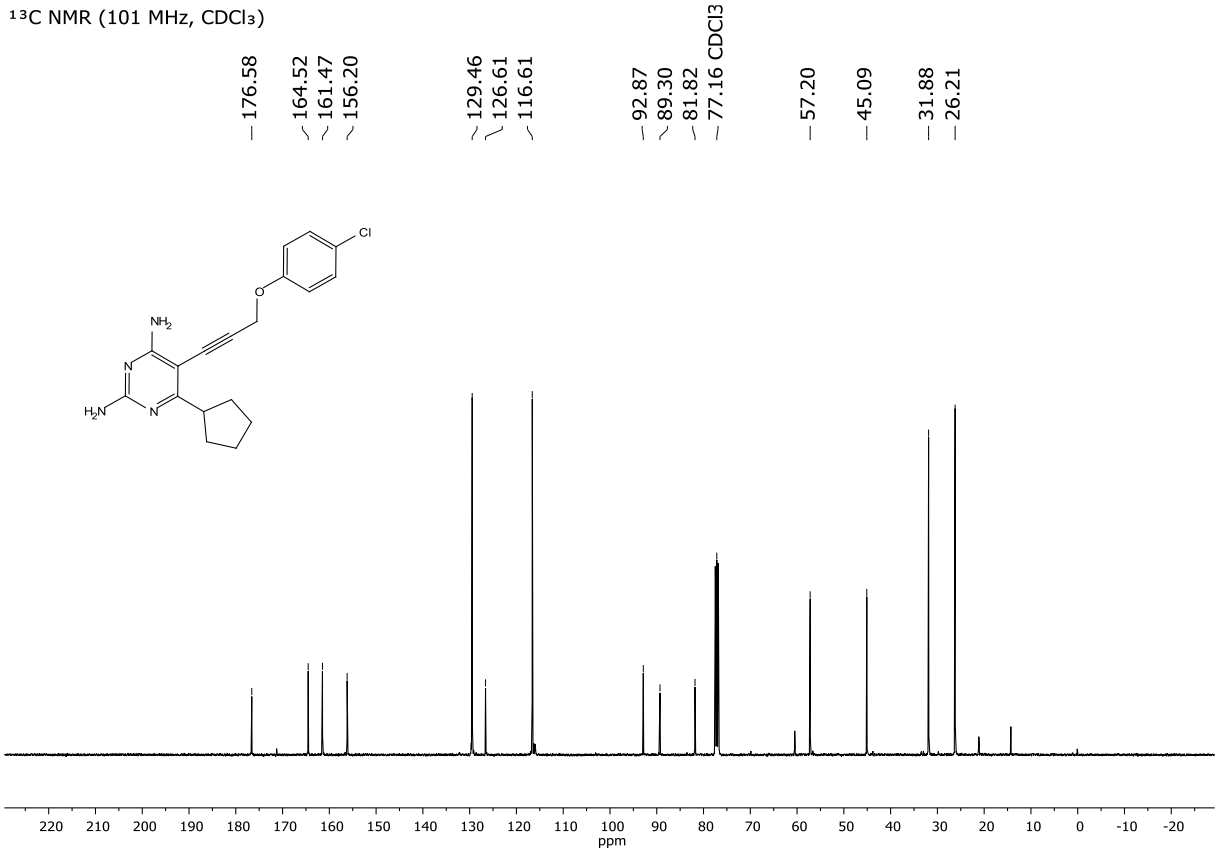
^{1}H NMR:

^{1}H NMR (400 MHz, CDCl$_3$)

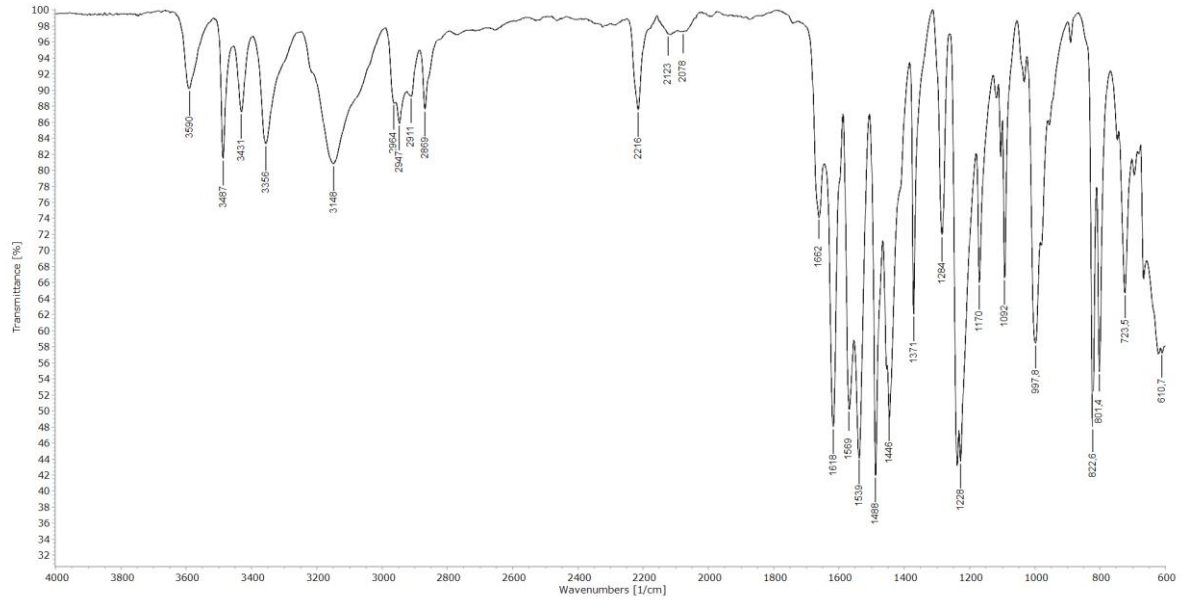


^{13}C NMR:

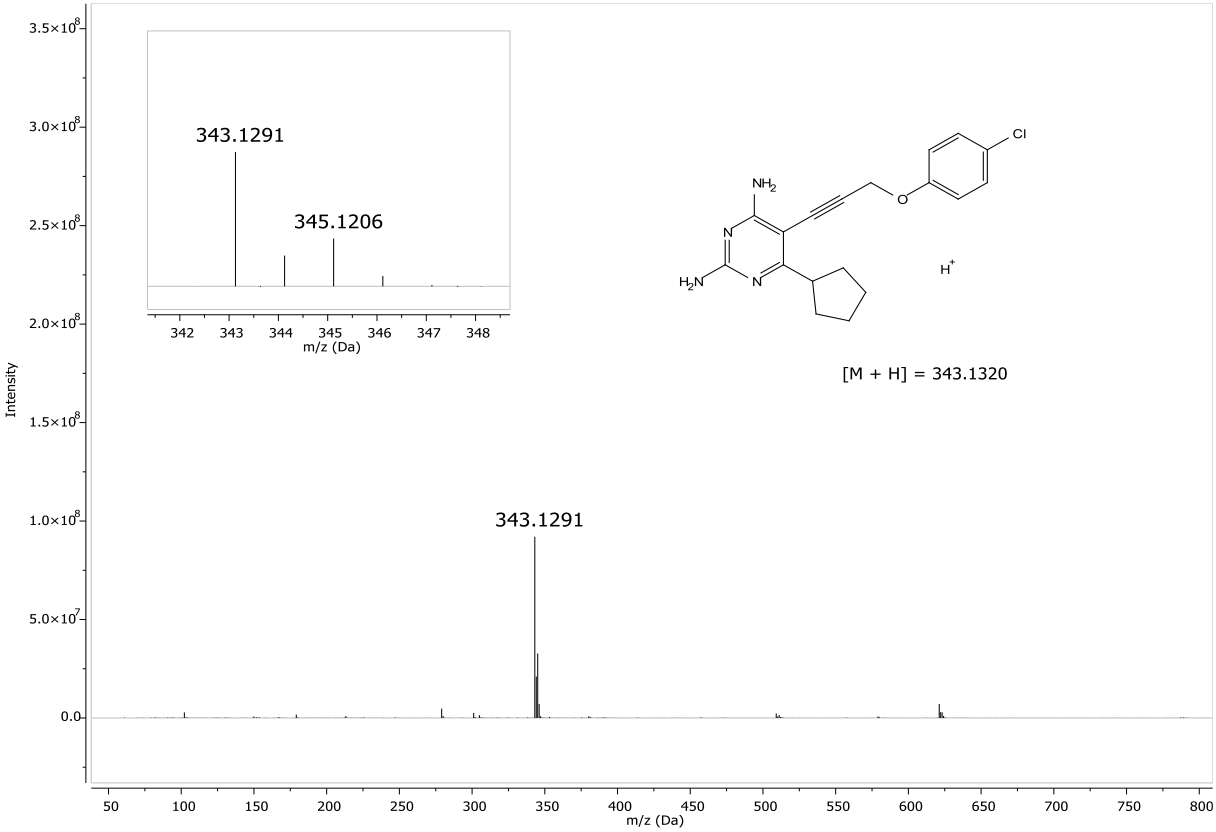
^{13}C NMR (101 MHz, CDCl$_3$)



FTIR:


Transmittance [%]
Wavenumbers [1/cm]
3590
3487
3431
3356
3148
2964
2947
2911
2869
2216
2123
2078
1662
1618
1569
1539
1488
1446
1371
1284
1228
1170
1092
997.8
861.4
822.6
723.5
610.7

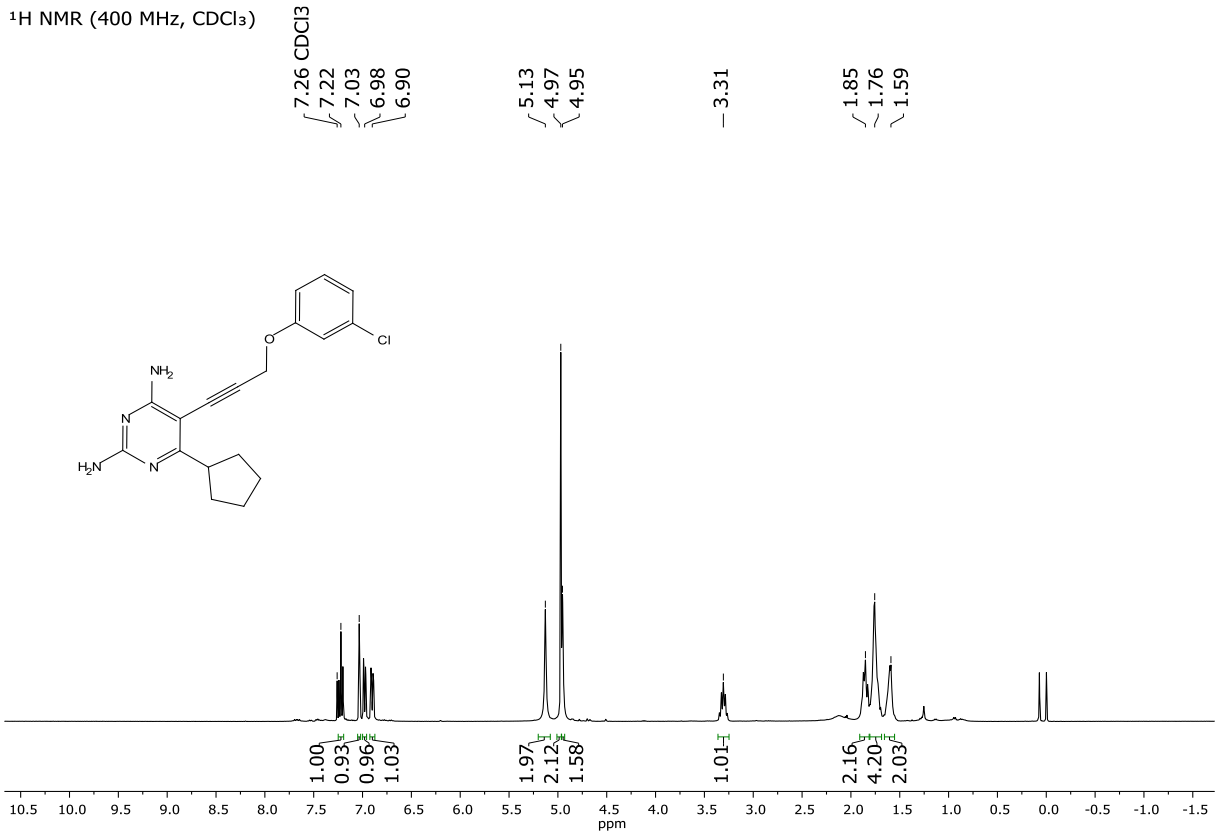
HRMS:


343.1291
345.1206
m/z (Da)
Cl
NH2
O
N
H2N
H+
[M + H] = 343.1320
Intensity
343.1291
m/z (Da)

6-Cyclopentyl-5-[3-(3-chlorophenoxy)prop-1-yn-1-yl]pyrimidine-2,4-diamine (**8b**)

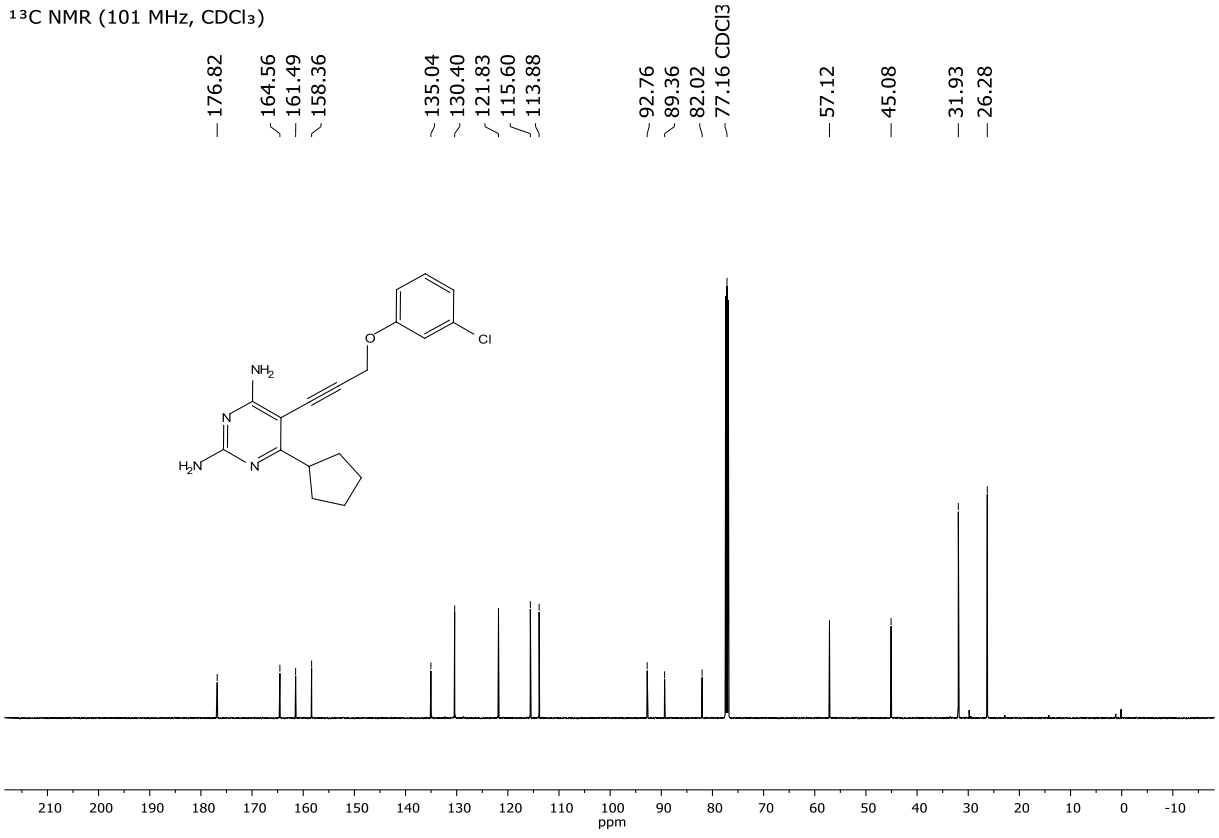
^{1}H NMR:

^{1}H NMR (400 MHz, CDCl$_3$)

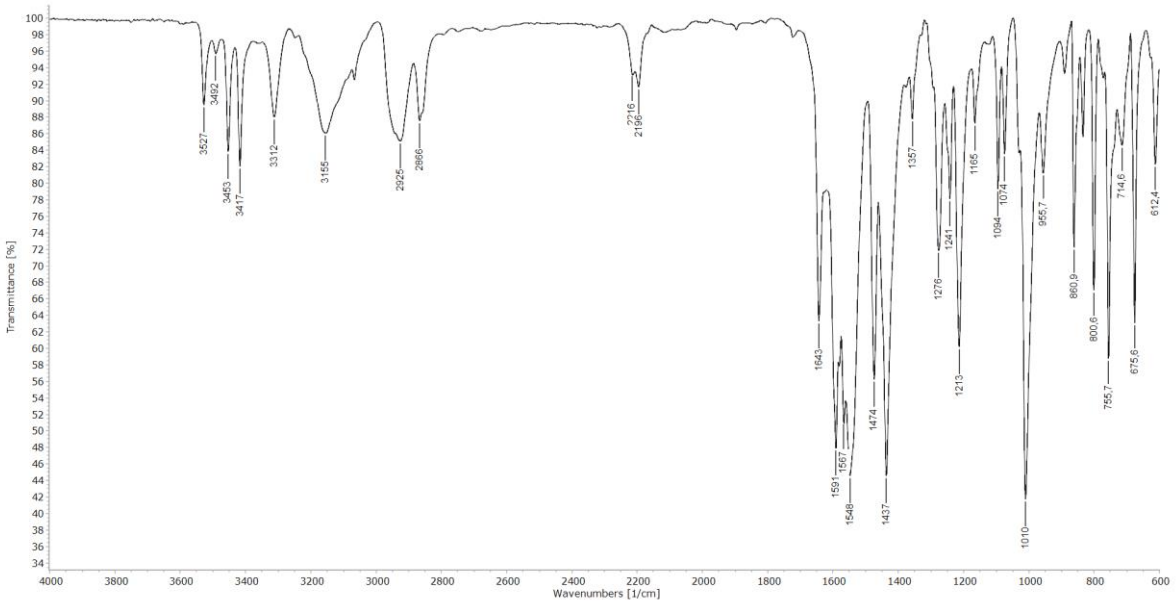


^{13}C NMR:

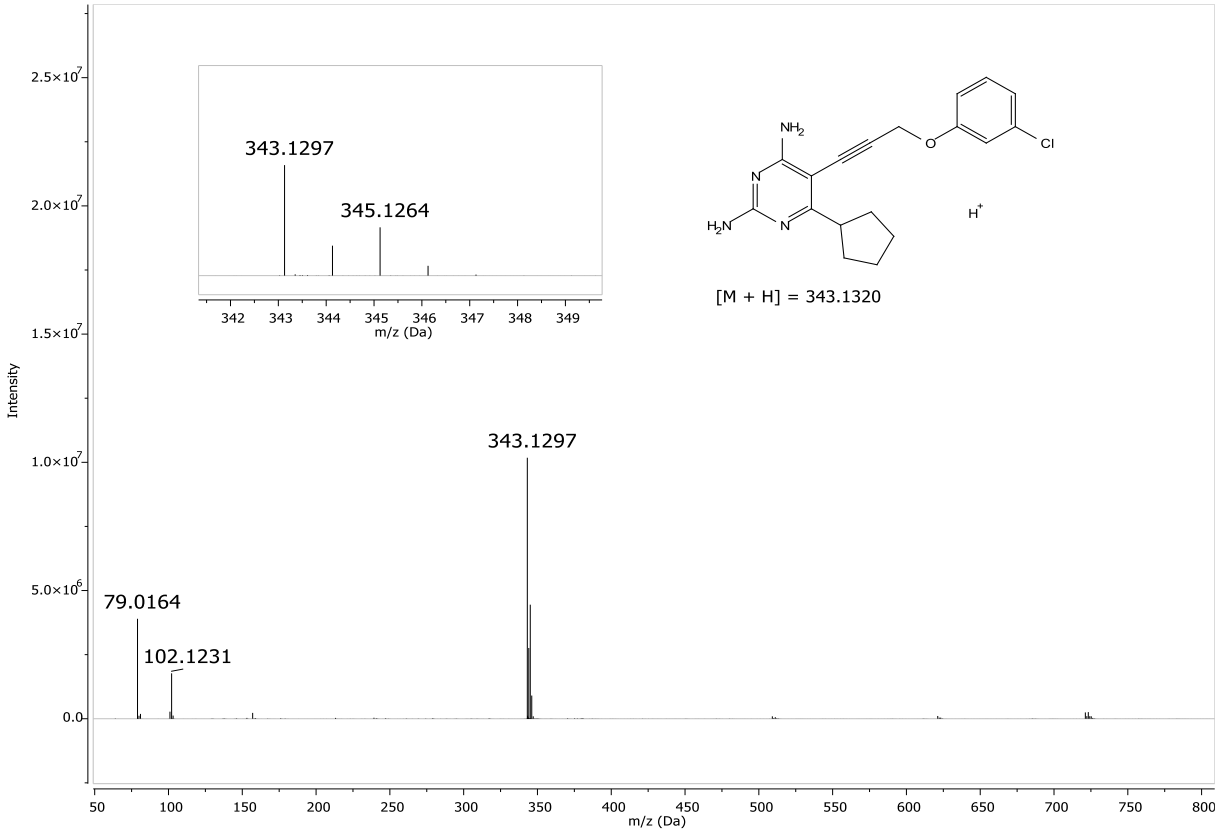
^{13}C NMR (101 MHz, CDCl$_3$)



FTIR:

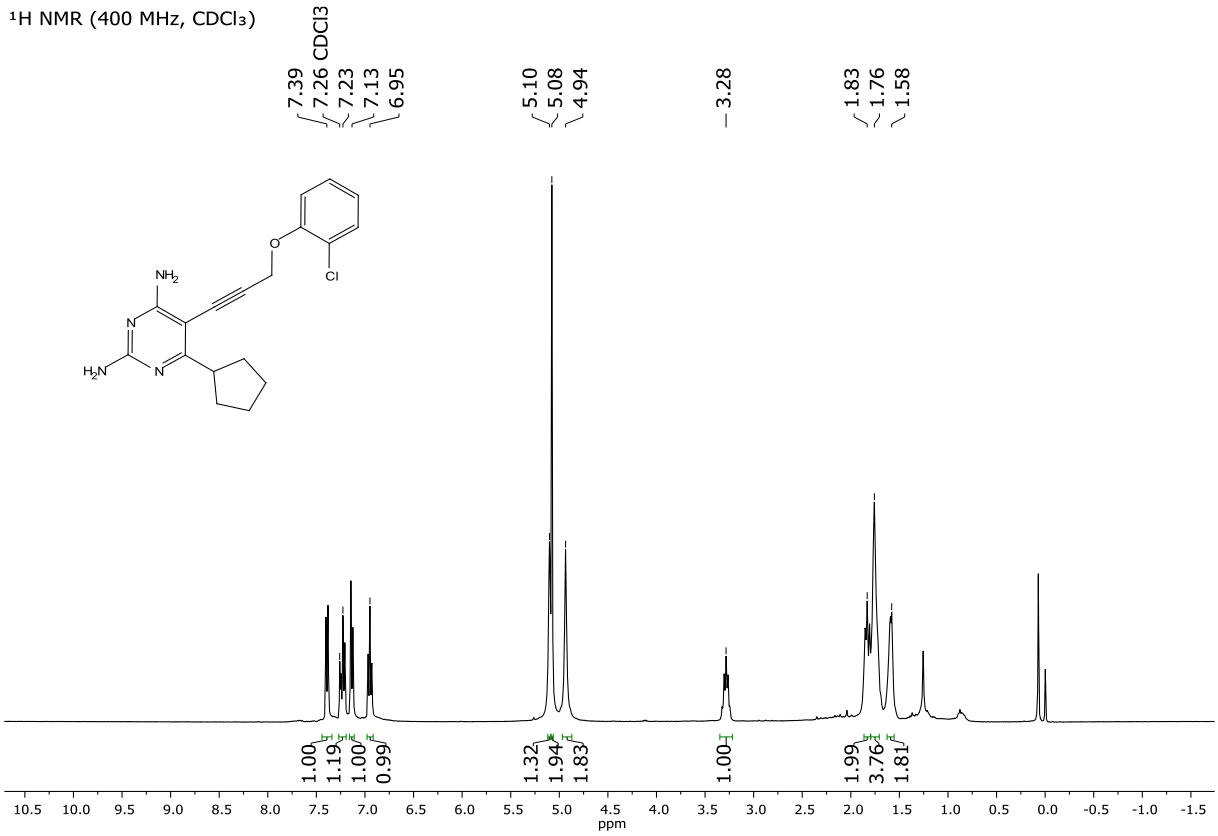

Transmittance [%]
Wavenumbers [1/cm]
3527
3492
3453
3417
3312
3155
2925
2866
2216
2196
1643
1591
1567
1548
1474
1437
1357
1276
1241
1213
1165
1094
1074
1010
955.7
860.9
800.6
755.7
675.6
714.6
612.4

HRMS:

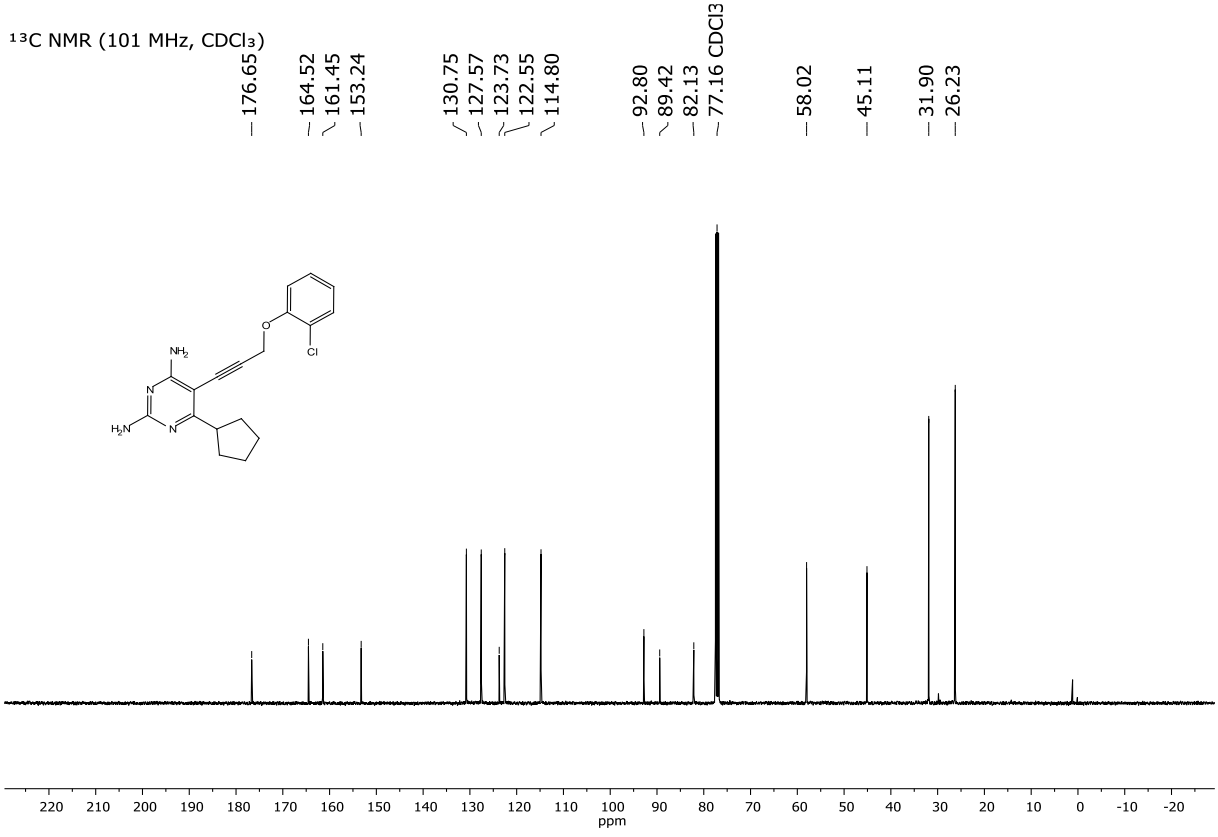

343.1297
345.1264
m/z (Da)
[M + H] = 343.1320
Intensity
343.1297
79.0164
102.1231
m/z (Da)

6-Cyclopentyl-5-[3-(2-chlorophenoxy)prop-1-yn-1-yl]pyrimidine-2,4-diamine (**8c**)

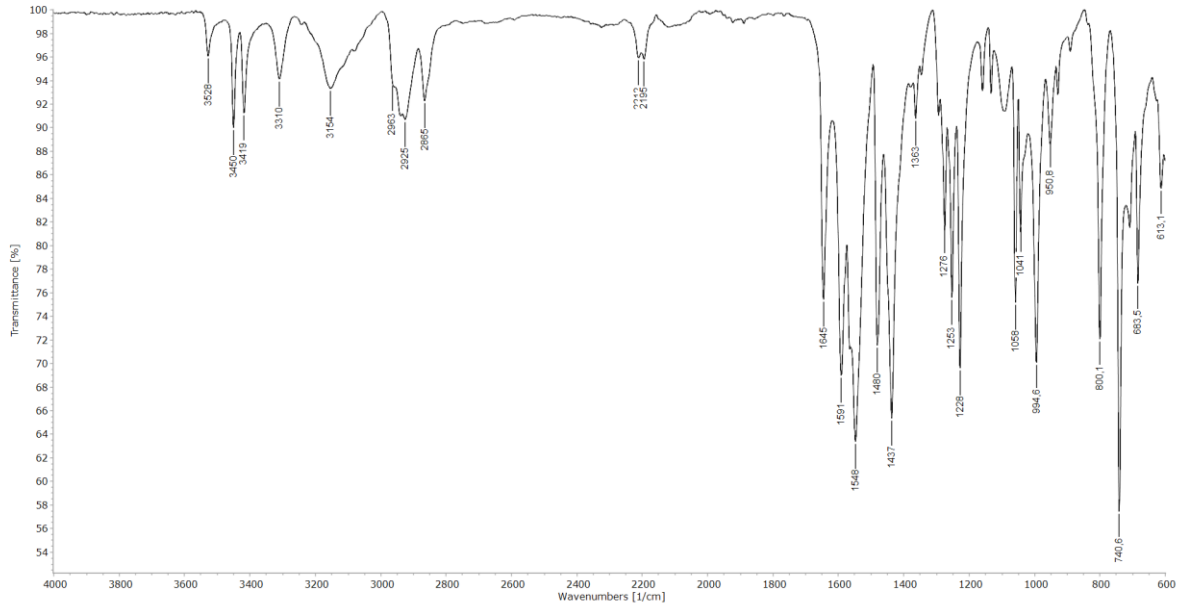
^{1}H NMR:



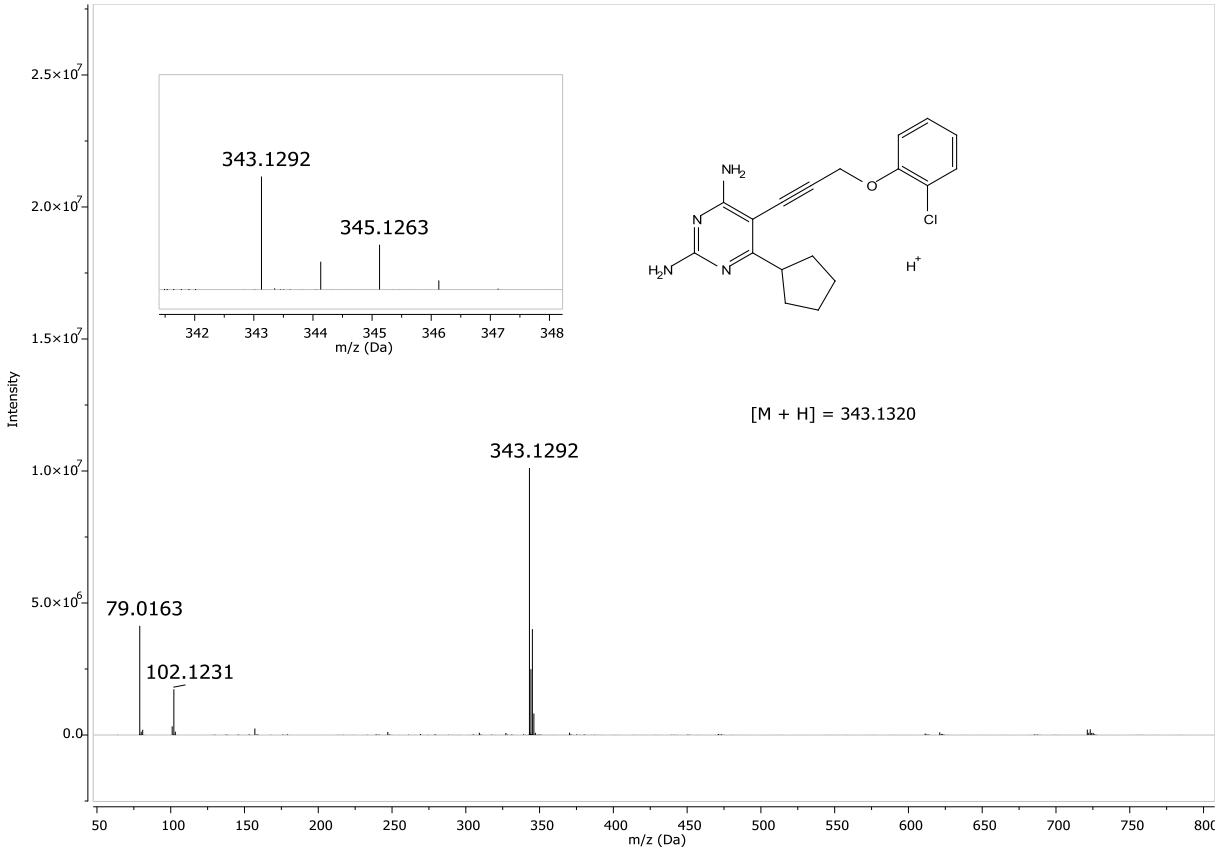
^{13}C NMR:



FTIR:


Transmittance [%]
Wavenumbers [1/cm]
3528
3450
3419
3310
3154
2963
2925
2865
2212
2195
1645
1591
1548
1480
1437
1363
1276
1253
1228
1058
1041
994,6
950,8
800,1
740,6
683,5
613,1

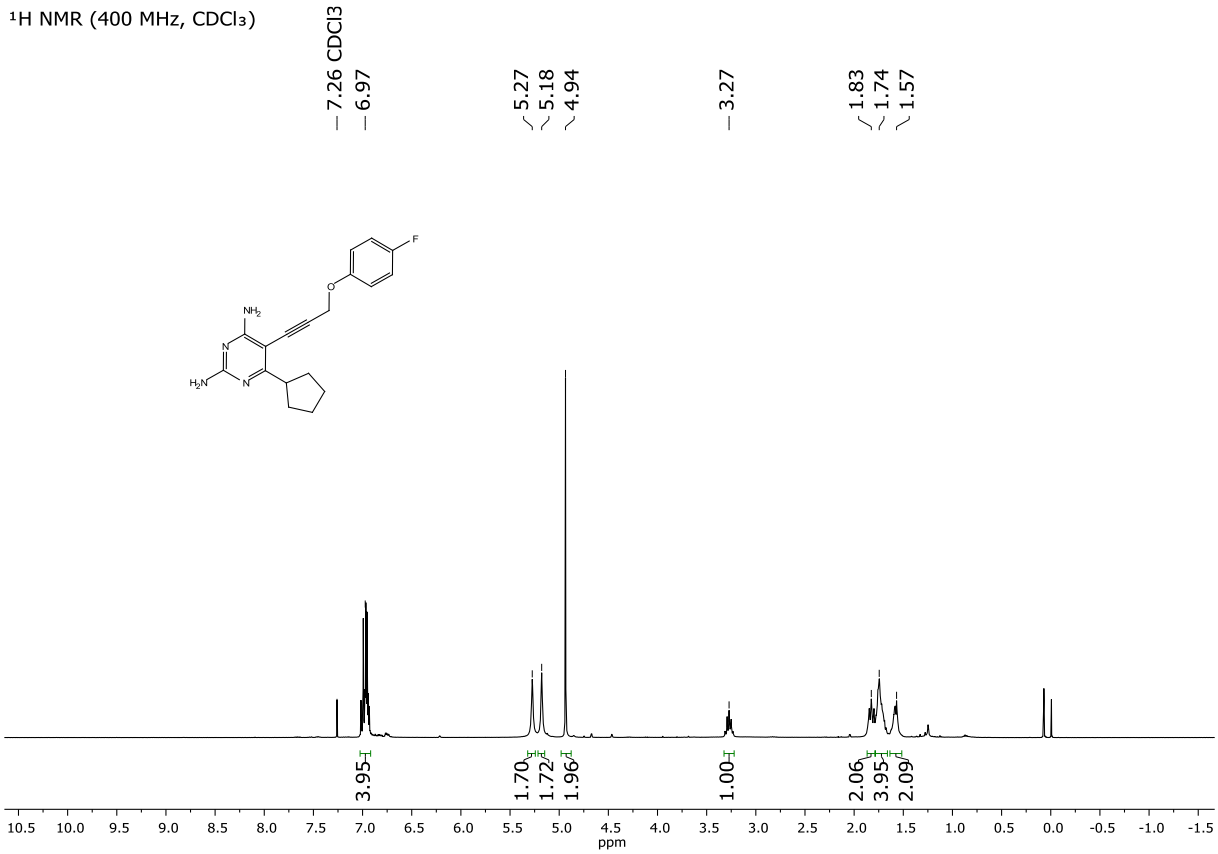
HRMS:


343.1292
345.1263
m/z (Da)
NH2
N
N
H2N
O
Cl
H+
[M + H] = 343.1320
Intensity
343.1292
79.0163
102.1231
m/z (Da)

6-Cyclopentyl-5-[3-(4-fluorophenoxy)prop-1-yn-1-yl]pyrimidine-2,4-diamine (**8d**)

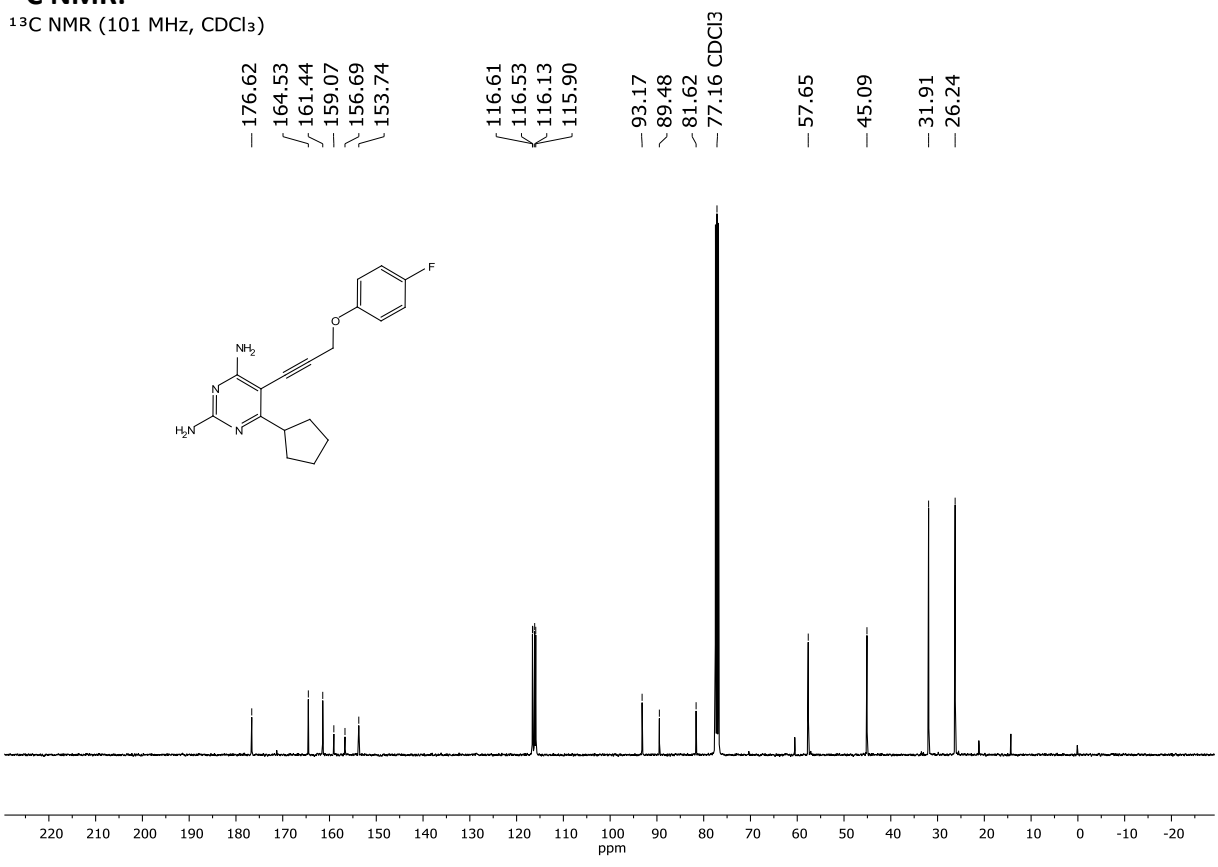
^{1}H NMR:

^{1}H NMR (400 MHz, CDCl$_3$)

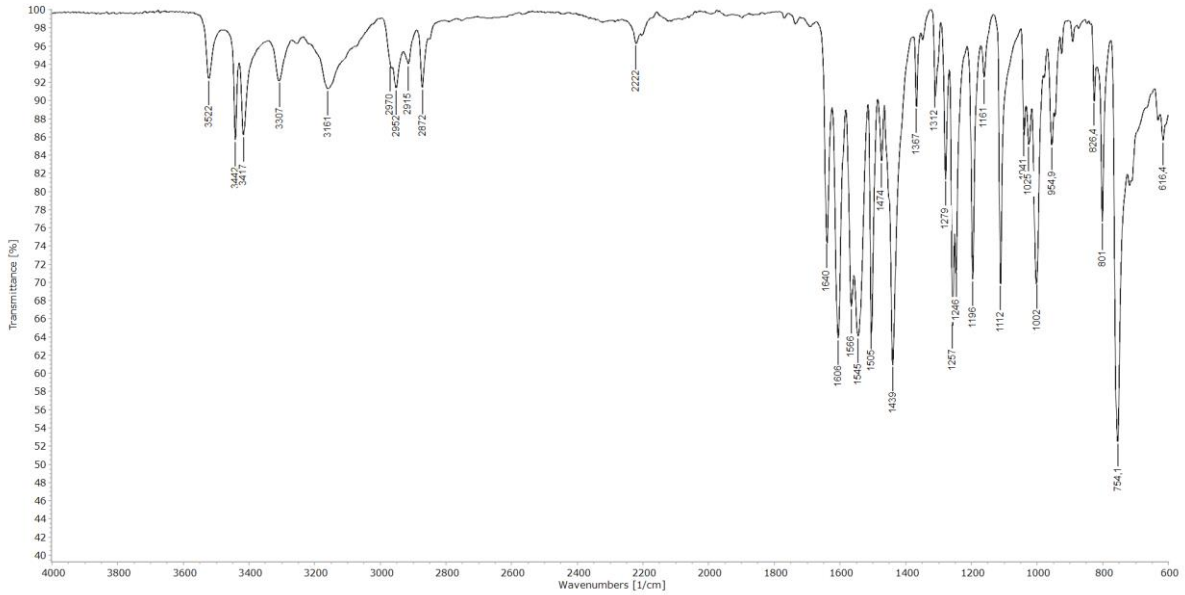


^{13}C NMR:

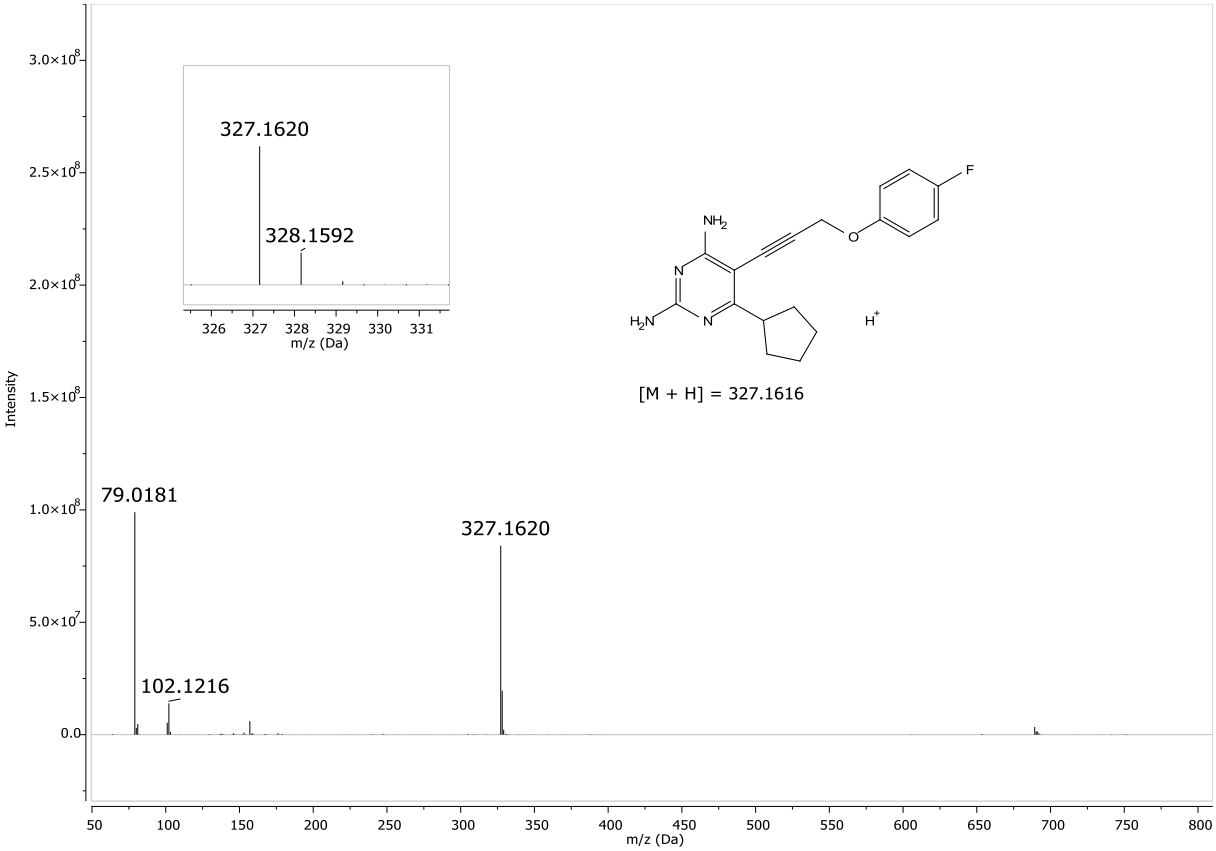
^{13}C NMR (101 MHz, CDCl$_3$)



FTIR:


Transmittance [%]
Wavenumbers [1/cm]
3522
3442
3417
3307
3161
2970
2952
2915
2872
2222
1640
1606
1568
1545
1505
1474
1439
1367
1312
1279
1257
1246
1196
1161
1112
1041
1025
1002
954.9
826.4
801
754.1
616.4

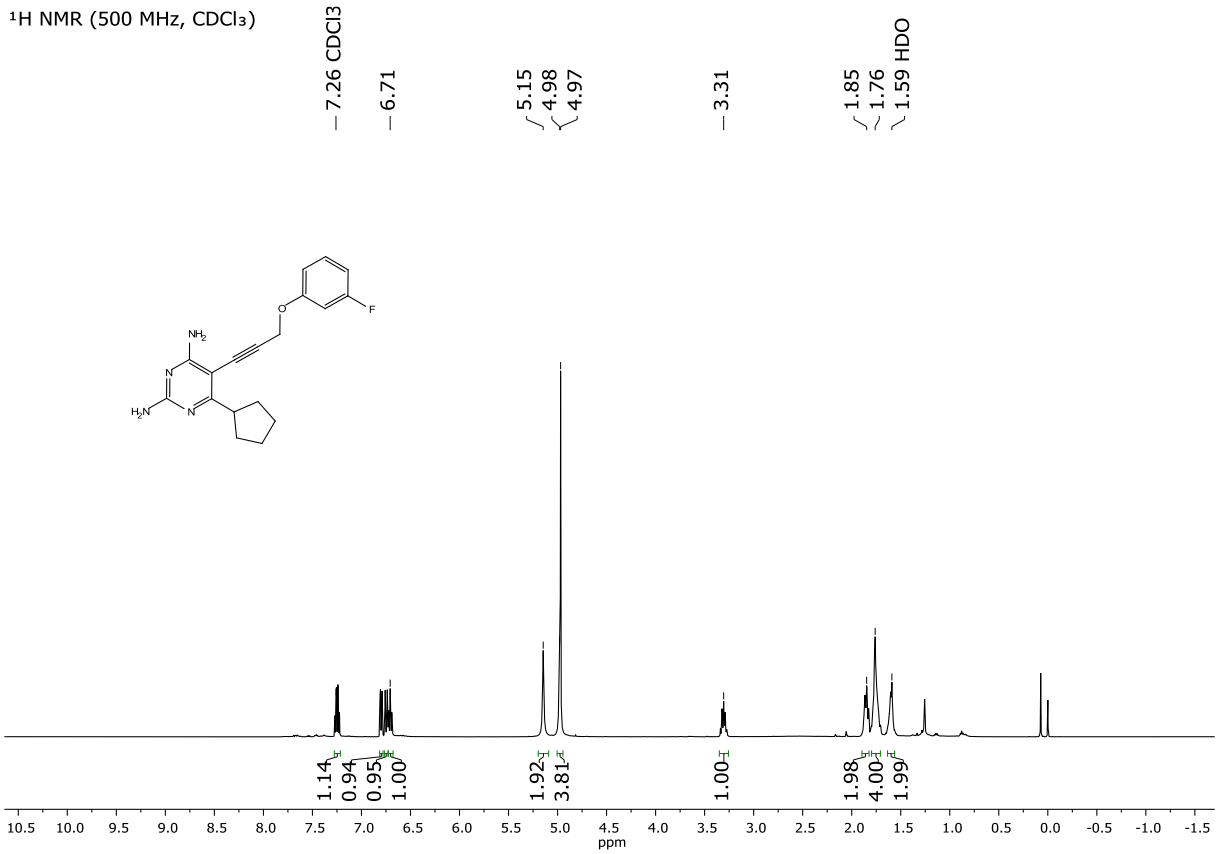
HRMS:


Intensity
m/z (Da)
327.1620
328.1592
79.0181
102.1216
[M + H] = 327.1616
F
O
N
NH2
H2N
H+

6-Cyclopentyl-5-[3-(3-fluorophenoxy)prop-1-yn-1-yl]pyrimidine-2,4-diamine (**8e**)

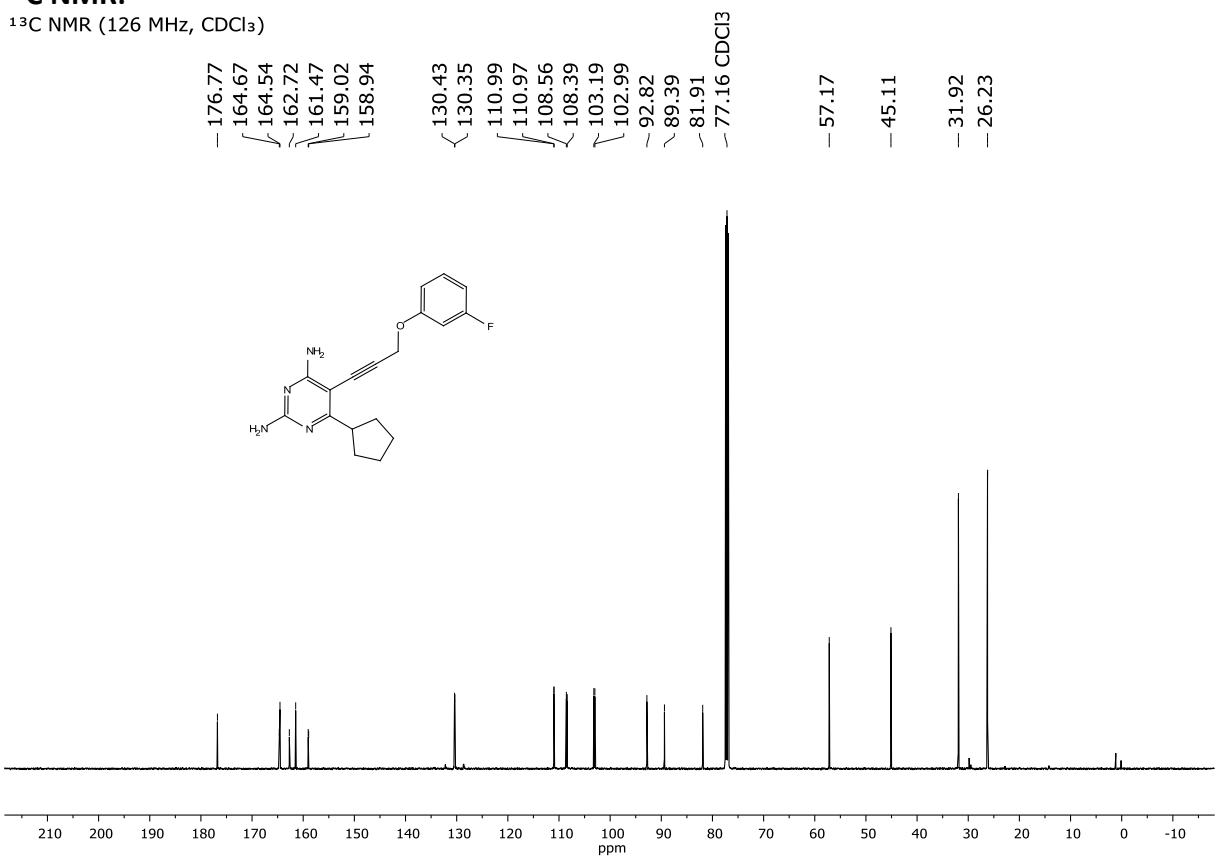
^{1}H NMR:

^{1}H NMR (500 MHz, CDCl$_3$)

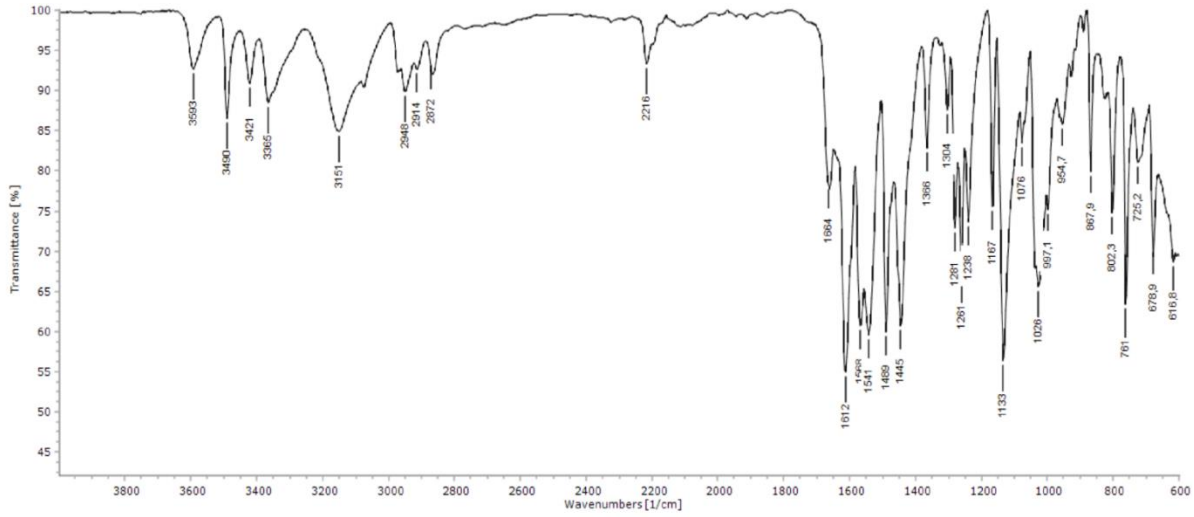


^{13}C NMR:

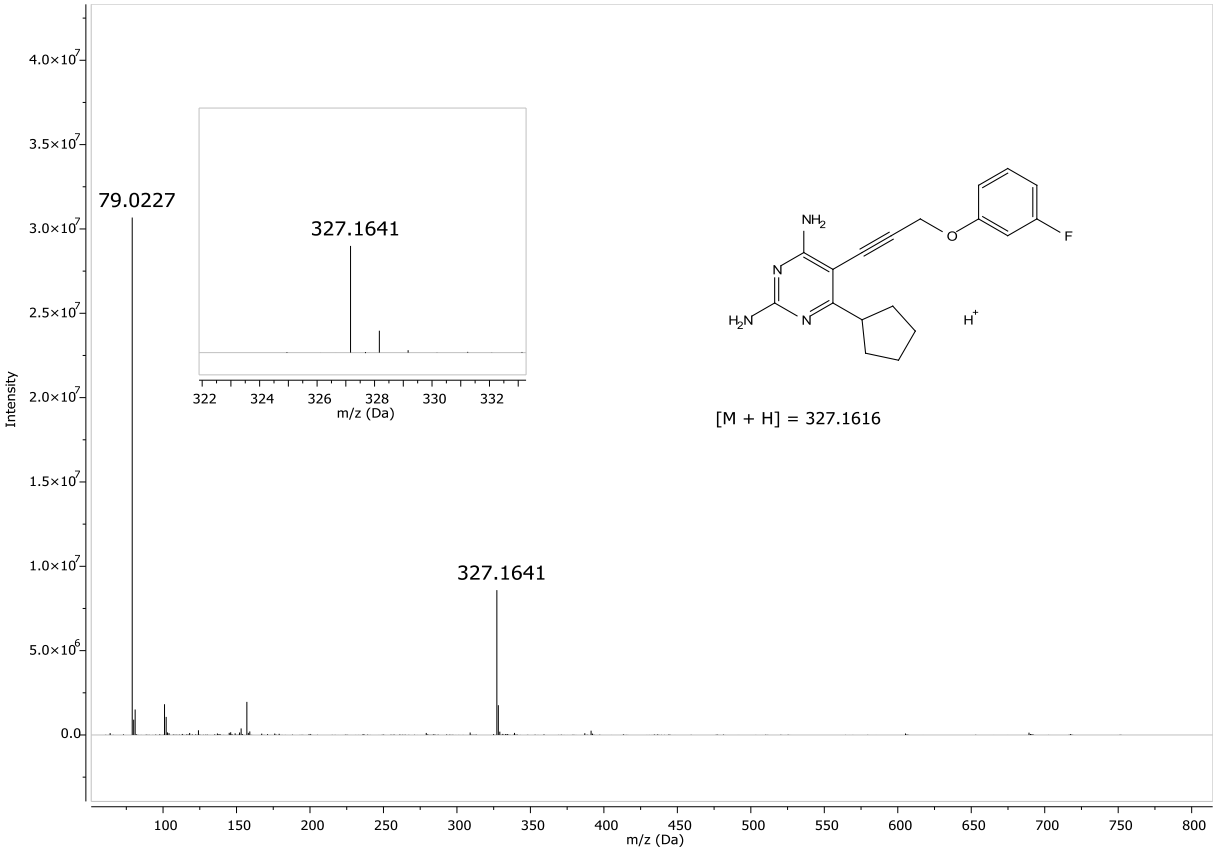
^{13}C NMR (126 MHz, CDCl$_3$)



FTIR:

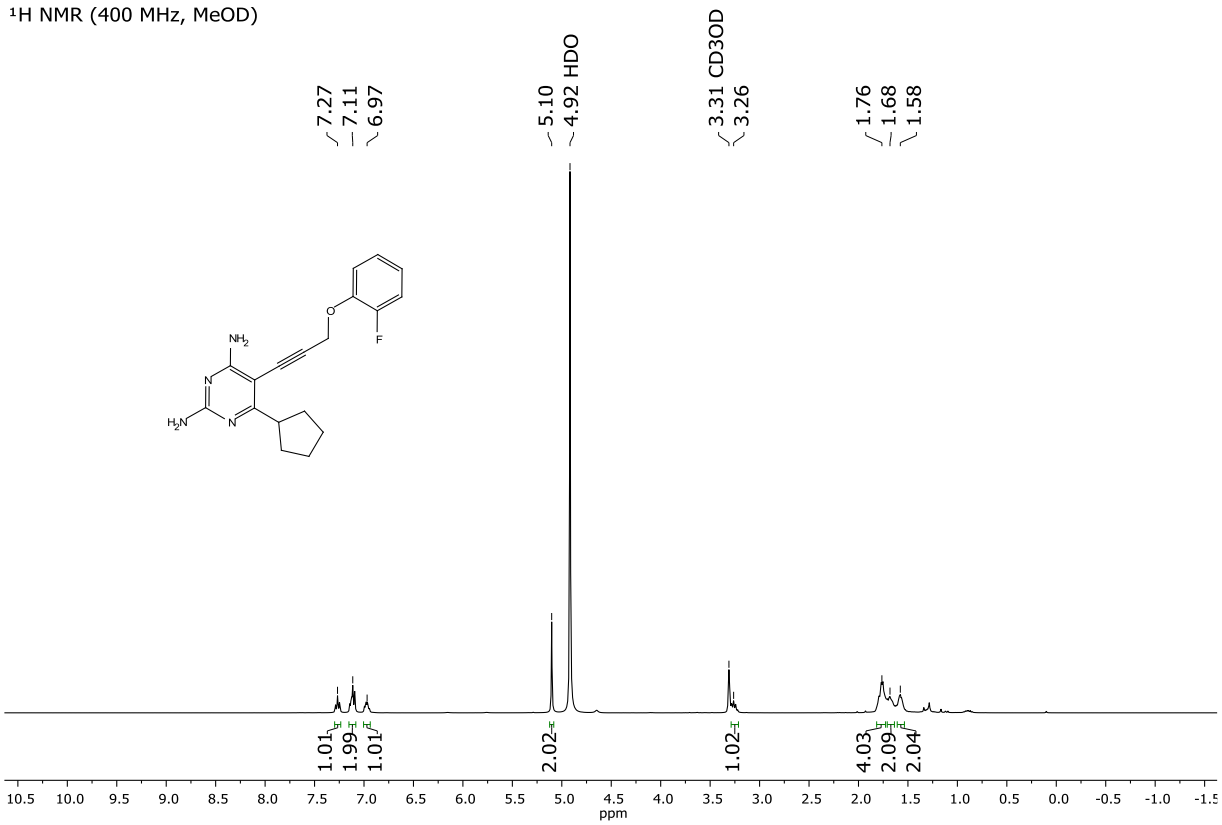

Transmittance [%]
Wavenumbers [1/cm]
3693
3490
3421
3365
3151
2948
2914
2872
2216
1984
1612
1541
1489
1445
1366
1304
1261
1238
1167
1133
1076
1026
997,1
954,7
867,9
802,3
761
725,2
678,9
616,8

HRMS:

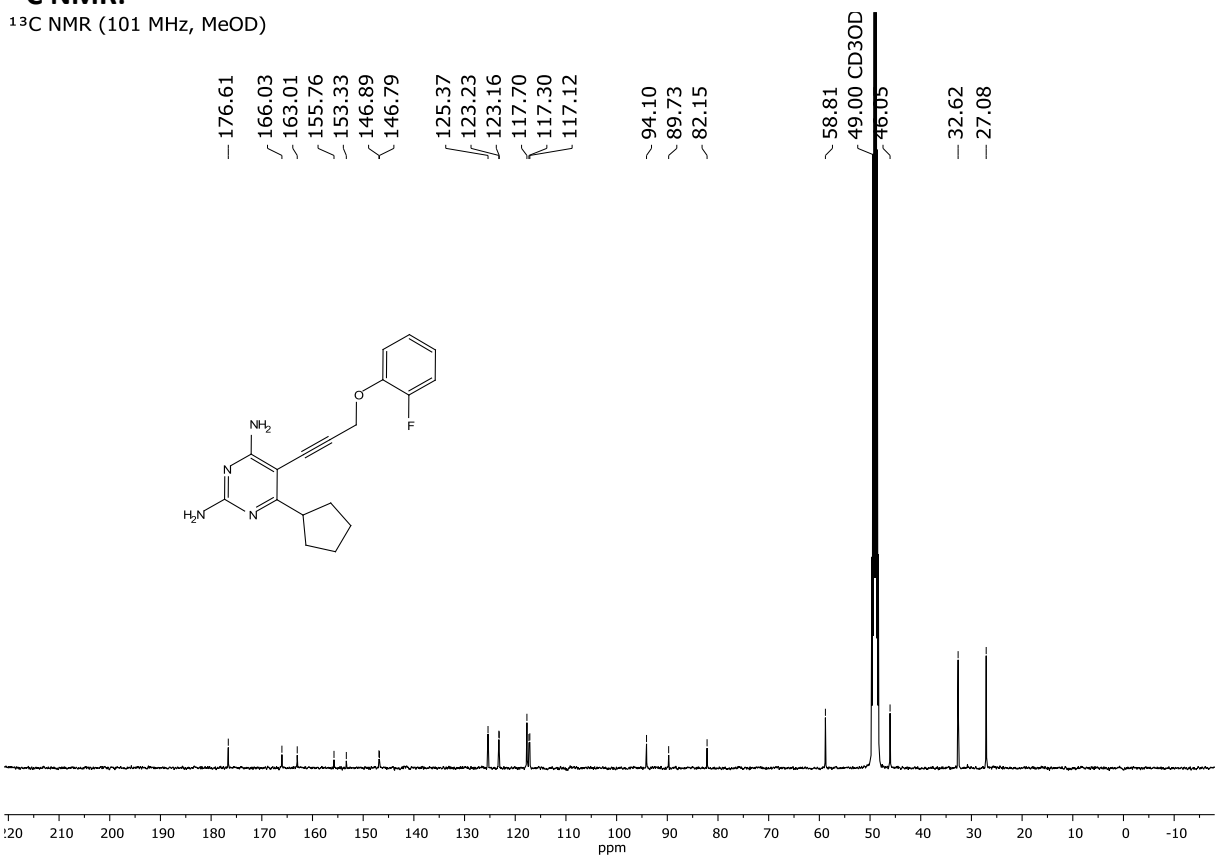

79.0227
327.1641
327.1641
Intensity
m/z (Da)
[M + H] = 327.1616
NH2
H2N
N
O
F
H+

6-Cyclopentyl-5-[3-(2-fluorophenoxy)prop-1-yn-1-yl]pyrimidine-2,4-diamine (**8f**)

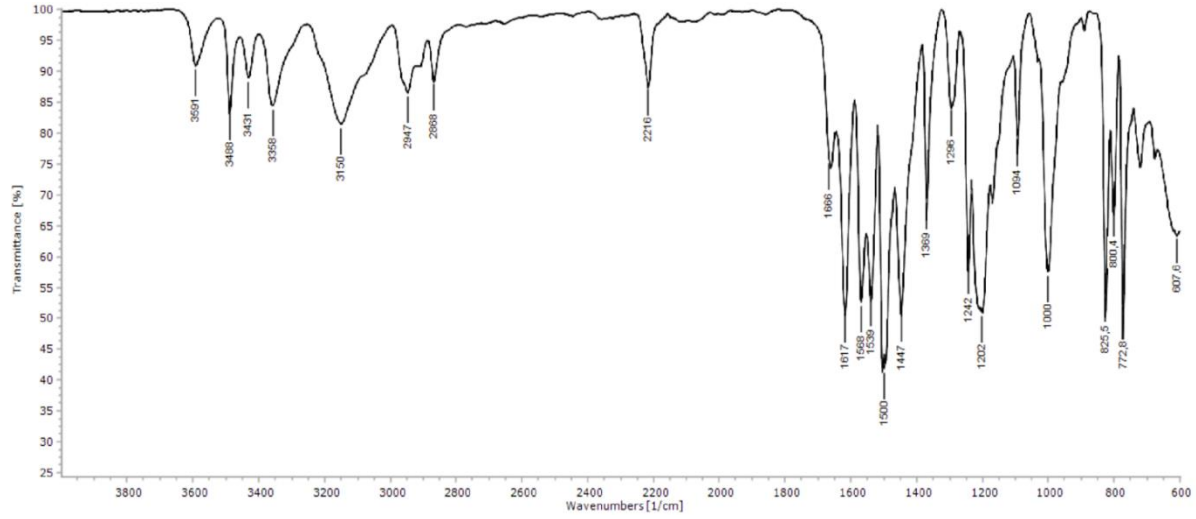
^{1}H NMR:



^{13}C NMR:



FTIR:


Transmittance [%]
Wavenumbers [1/cm]
3591
3488
3431
3358
3150
2947
2868
2210
1668
1617
1588
1539
1500
1447
1369
1296
1242
1202
1094
1000
825,5
800,4
772,8
607,6

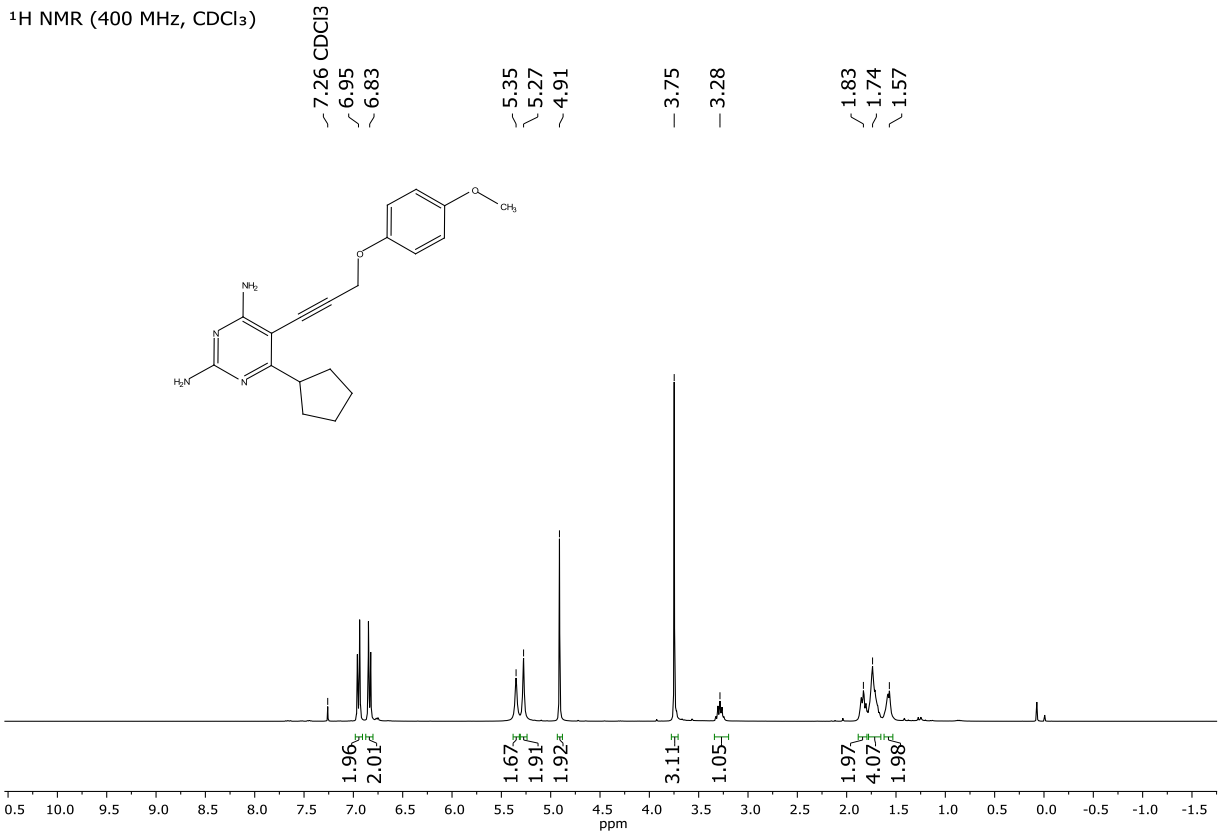
HRMS:


327.1639
328.1654
m/z (Da)
[M + H] = 327.1616
Intensity
79.0227
327.1639
m/z (Da)

6-Cyclopentyl-5-[3-(4-methoxyphenoxy)prop-1-yn-1-yl]pyrimidine-2,4-diamine (**8g**)

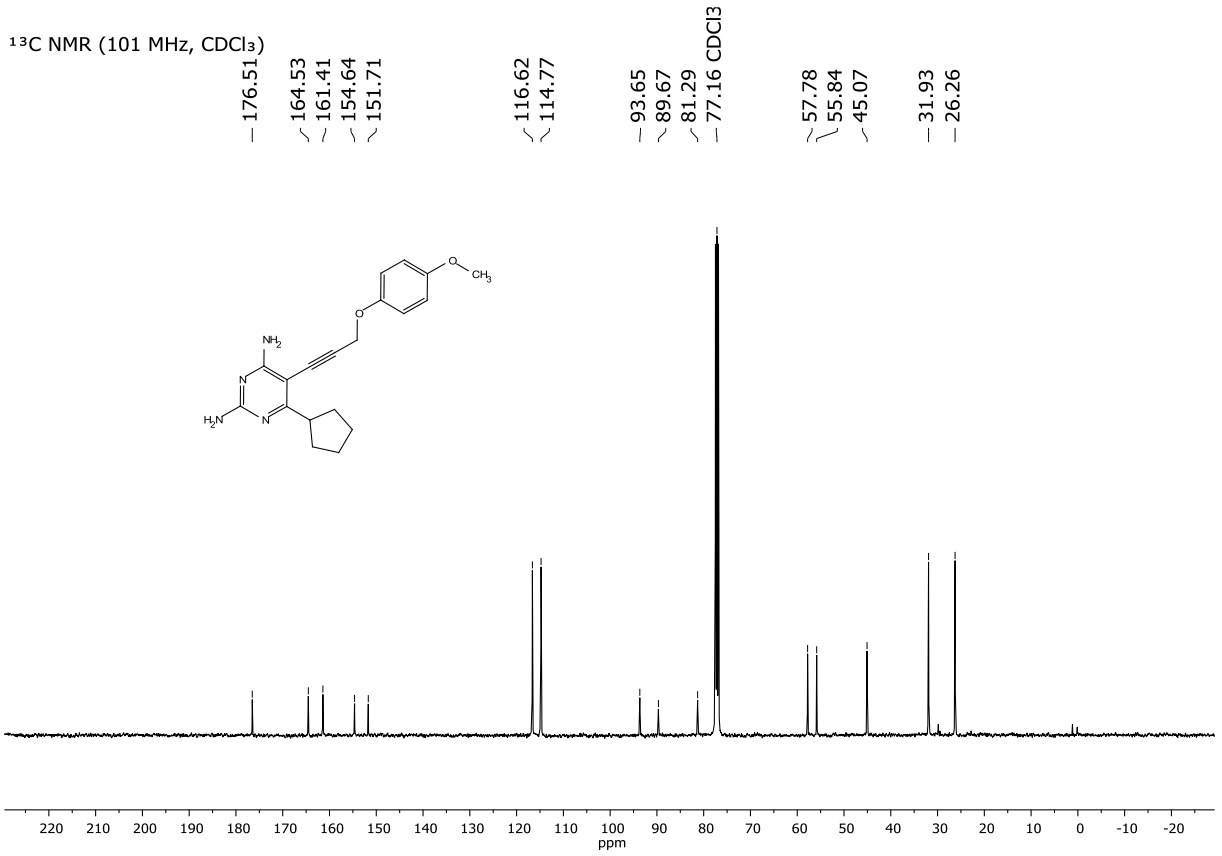
^{1}H NMR:

^{1}H NMR (400 MHz, $CDCl_3$)

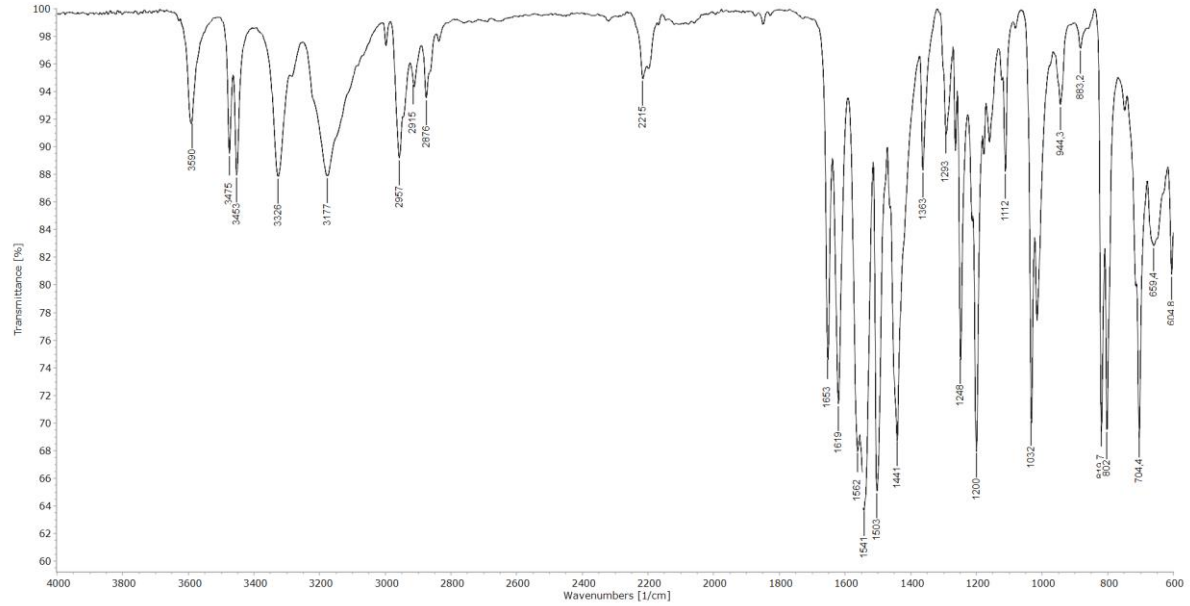


^{13}C NMR:

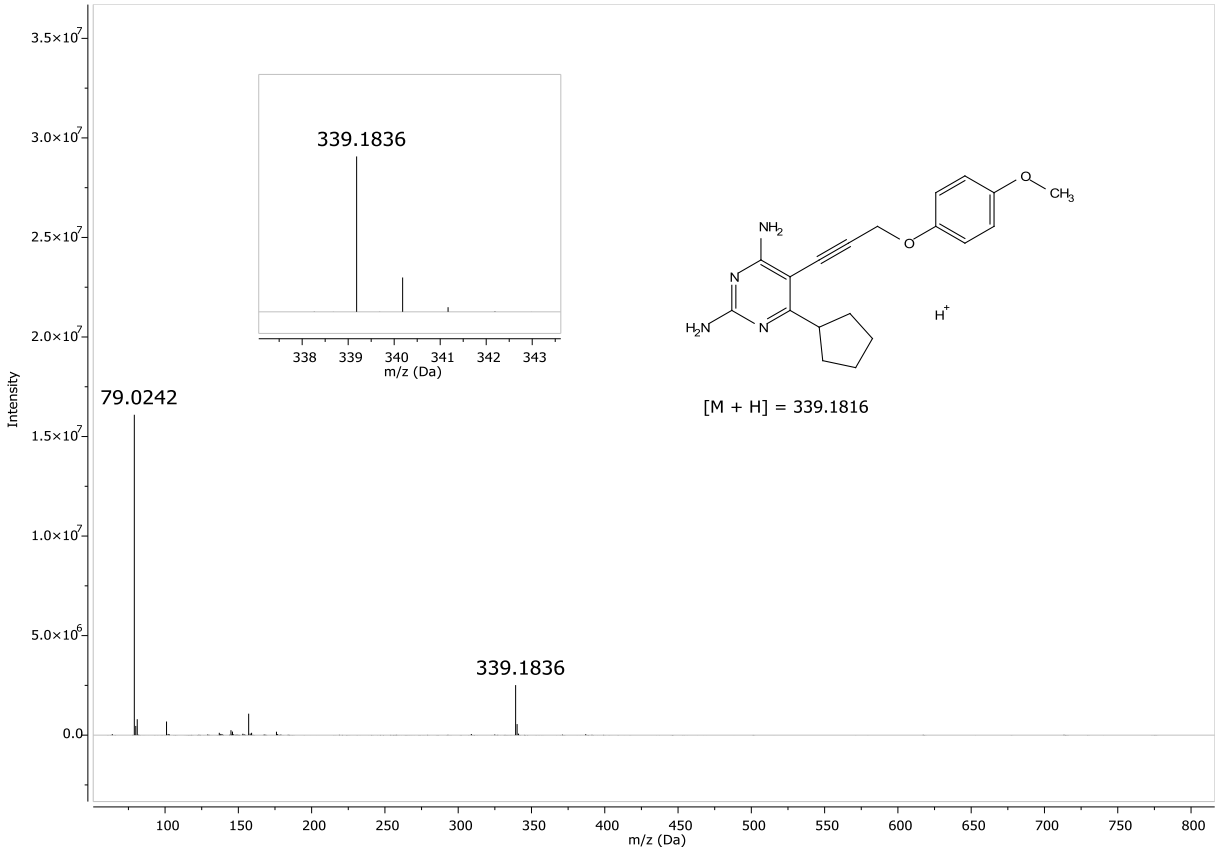
^{13}C NMR (101 MHz, $CDCl_3$)



FTIR:


Transmittance [%]
Wavenumbers [1/cm]
3590
3475
3450
3326
3177
2957
2915
2876
2215
1653
1619
1562
1541
1503
1441
1363
1293
1248
1200
1112
1032
944.3
883.2
810.7
802
704.4
659.4
604.8

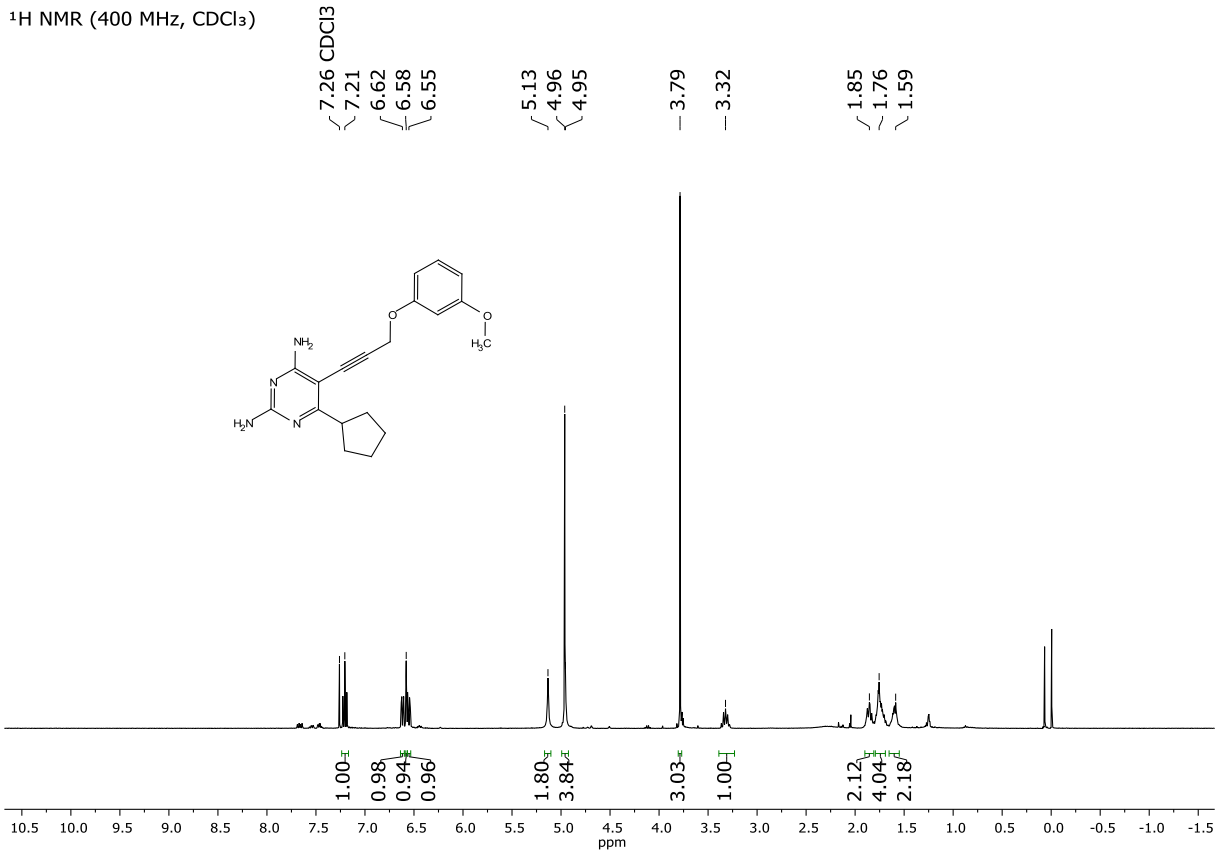
HRMS:


339.1836
m/z (Da)
79.0242
339.1836
Intensity
m/z (Da)
[M + H] = 339.1816

6-Cyclopentyl-5-[3-(3-methoxyphenoxy)prop-1-yn-1-yl]pyrimidine-2,4-diamine (**8h**)

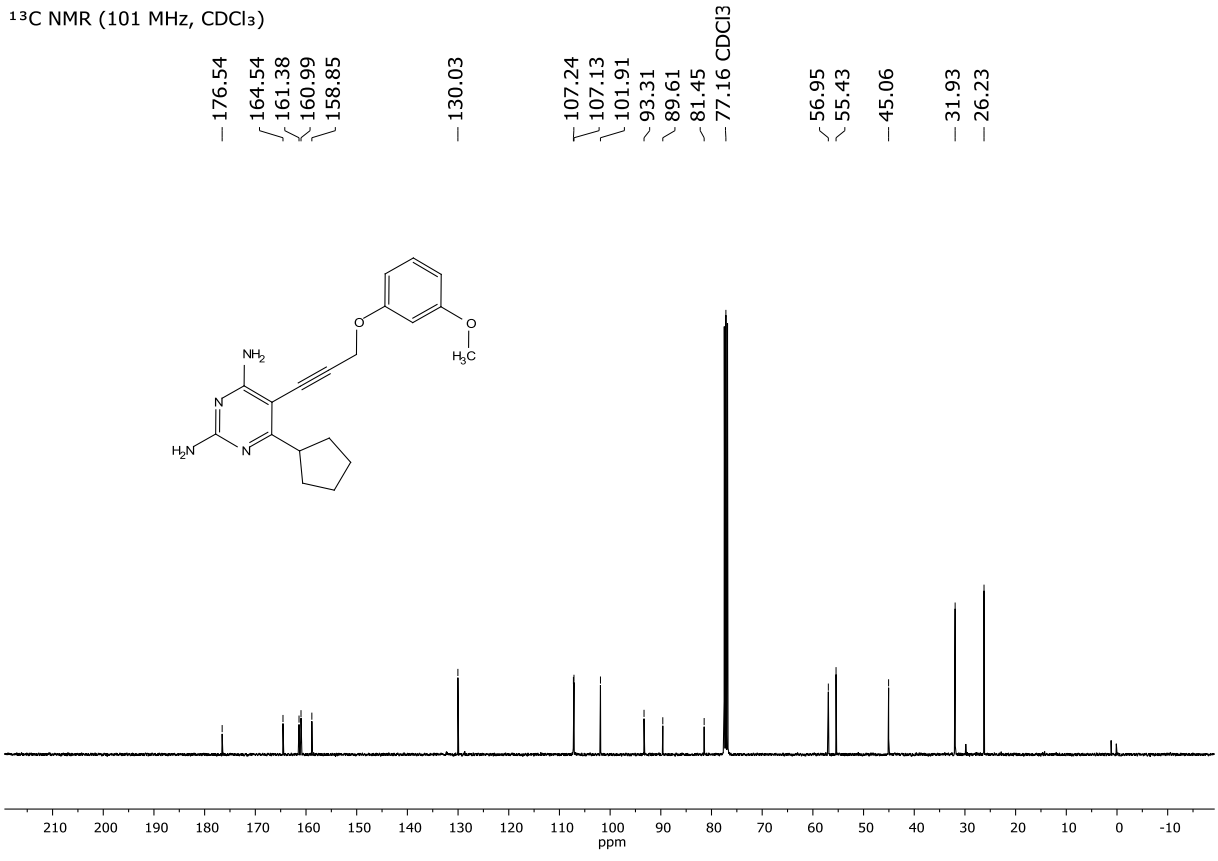
^{1}H NMR:

^{1}H NMR (400 MHz, CDCl$_3$)

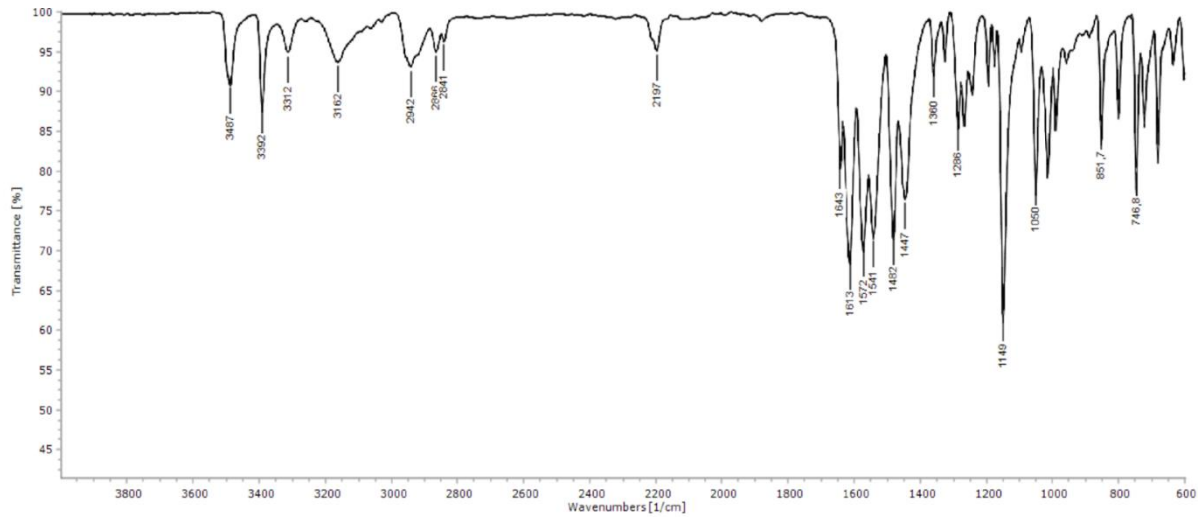


^{13}C NMR:

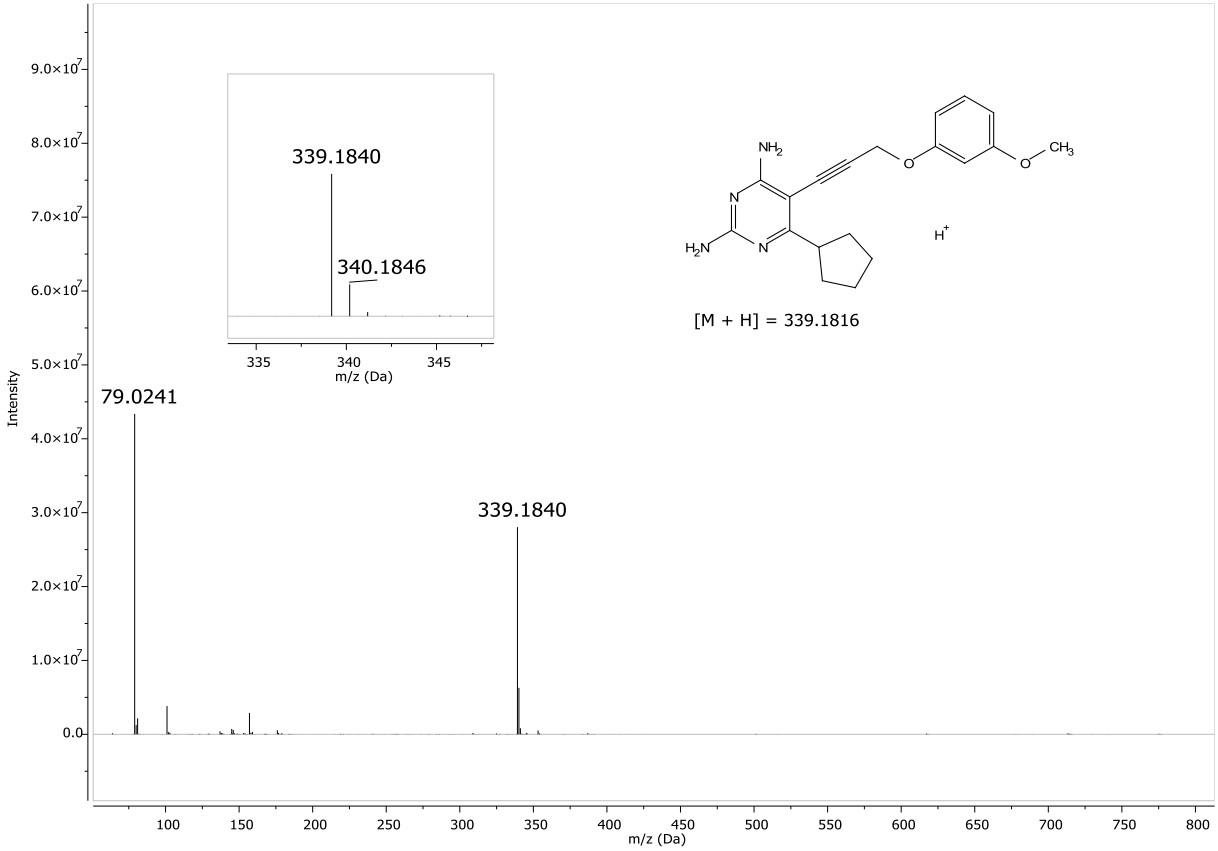
^{13}C NMR (101 MHz, CDCl$_3$)



FTIR:


Transmittance [%]
Wavenumbers [1/cm]
3487
3392
3312
3162
2942
2866
2841
2197
1643
1613
1572
1541
1492
1447
1360
1298
1149
1050
851,7
746,8

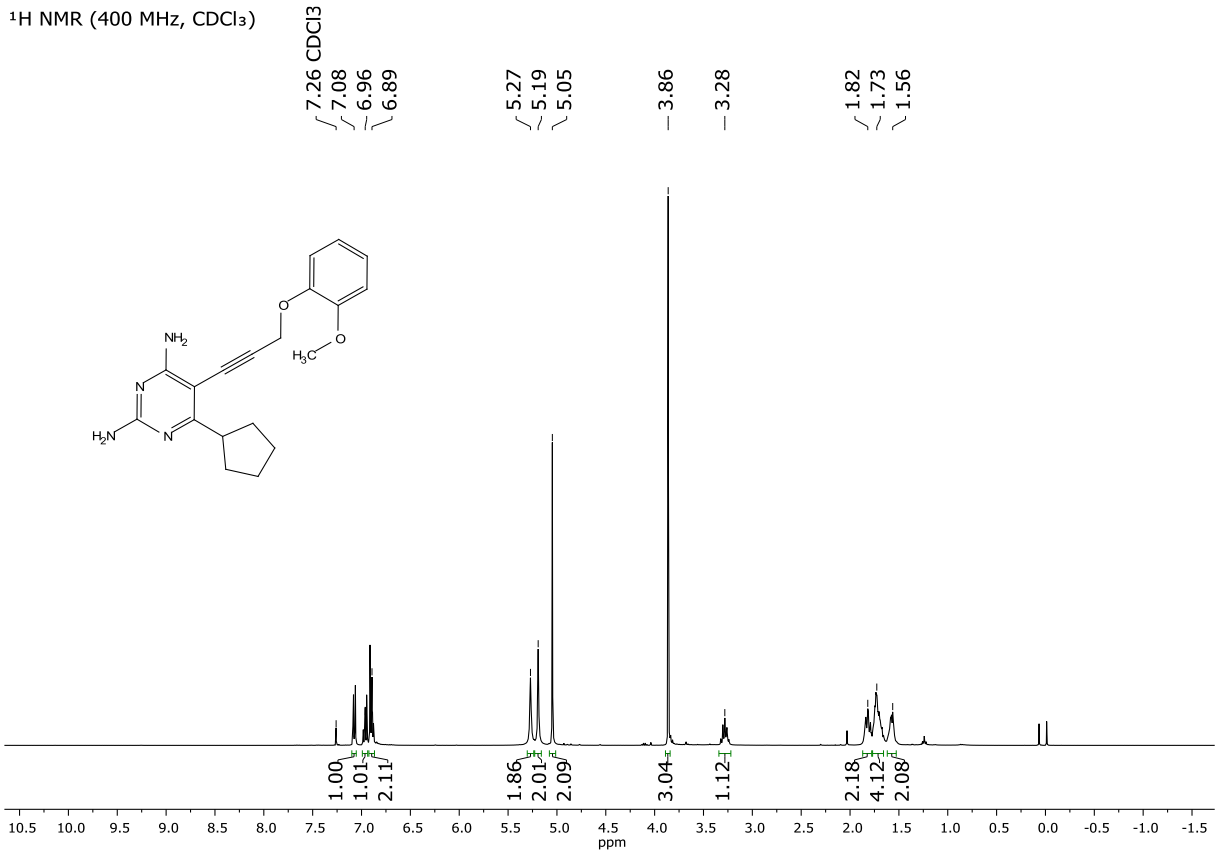
HRMS:


339.1840
340.1846
335
340
345
m/z (Da)
NH2
N
H2N
O
O
CH3
H+
[M + H] = 339.1816
Intensity
79.0241
339.1840
m/z (Da)

6-Cyclopentyl-5-[3-(2-methoxyphenoxy)prop-1-yn-1-yl]pyrimidine-2,4-diamine (**8i**)

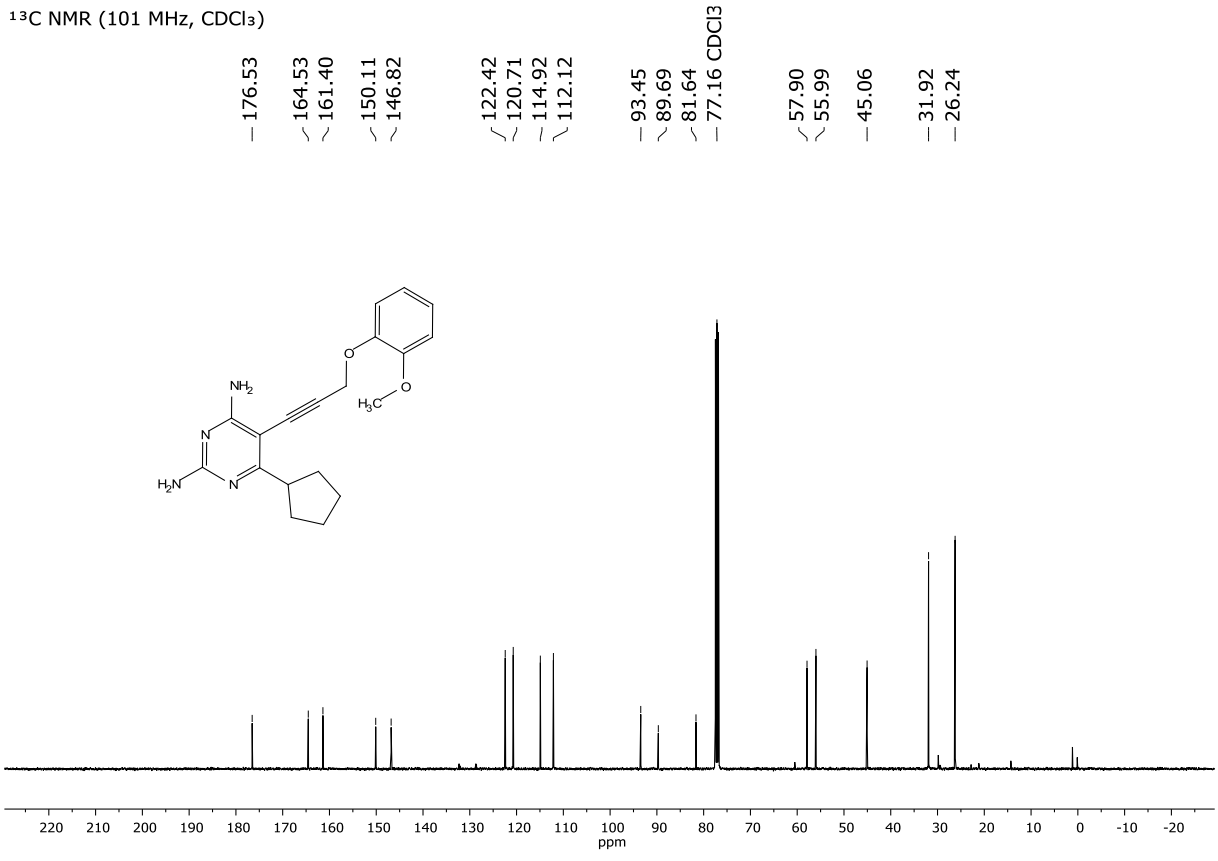
^{1}H NMR:

^{1}H NMR (400 MHz, CDCl$_3$)

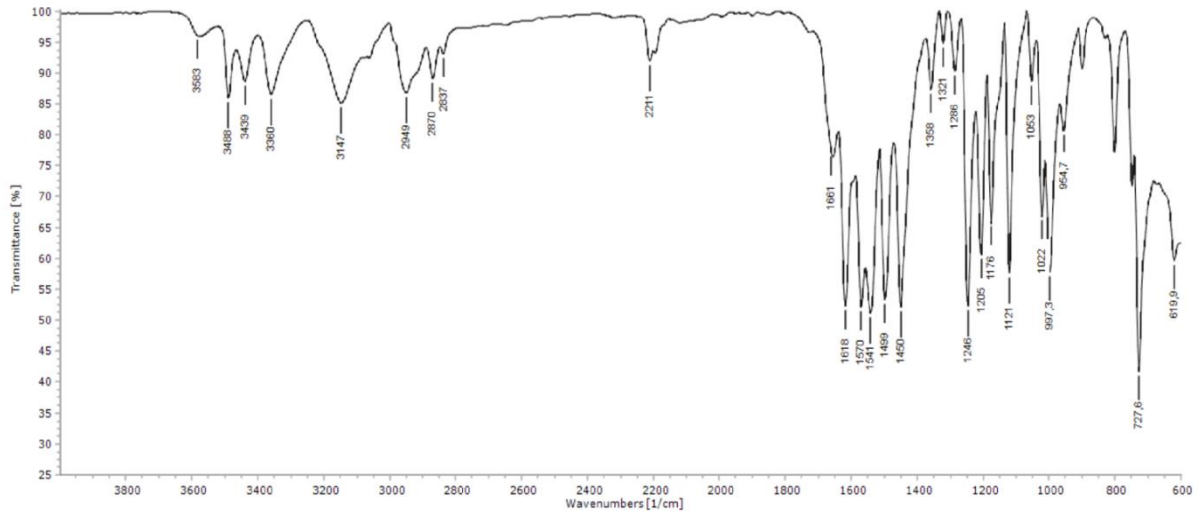


^{13}C NMR:

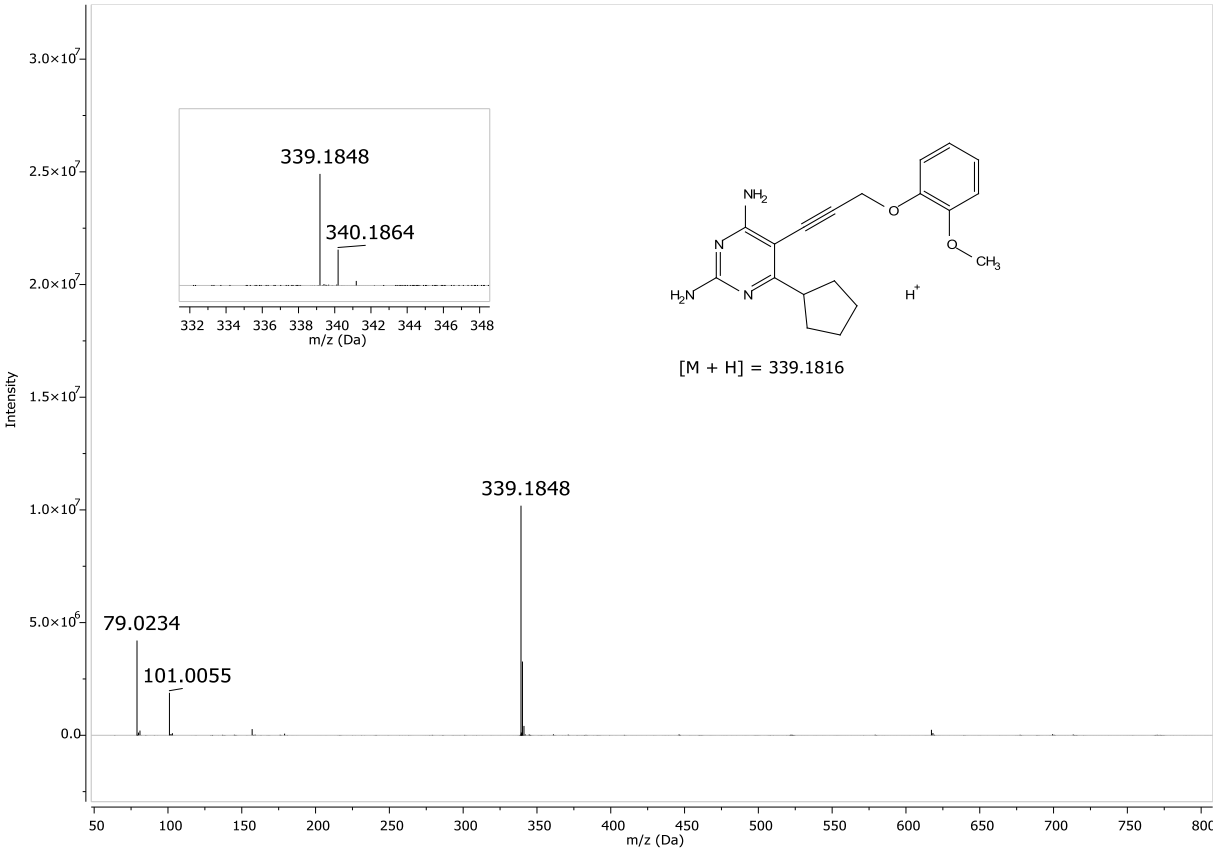
^{13}C NMR (101 MHz, CDCl$_3$)



FTIR:


Transmittance [%]
Wavenumbers [1/cm]
3563
3488
3439
3360
3147
2949
2870
2837
2211
1661
1618
1570
1541
1499
1450
1398
1321
1286
1246
1205
1176
1121
1053
1022
997,3
964,7
727,6
619,9

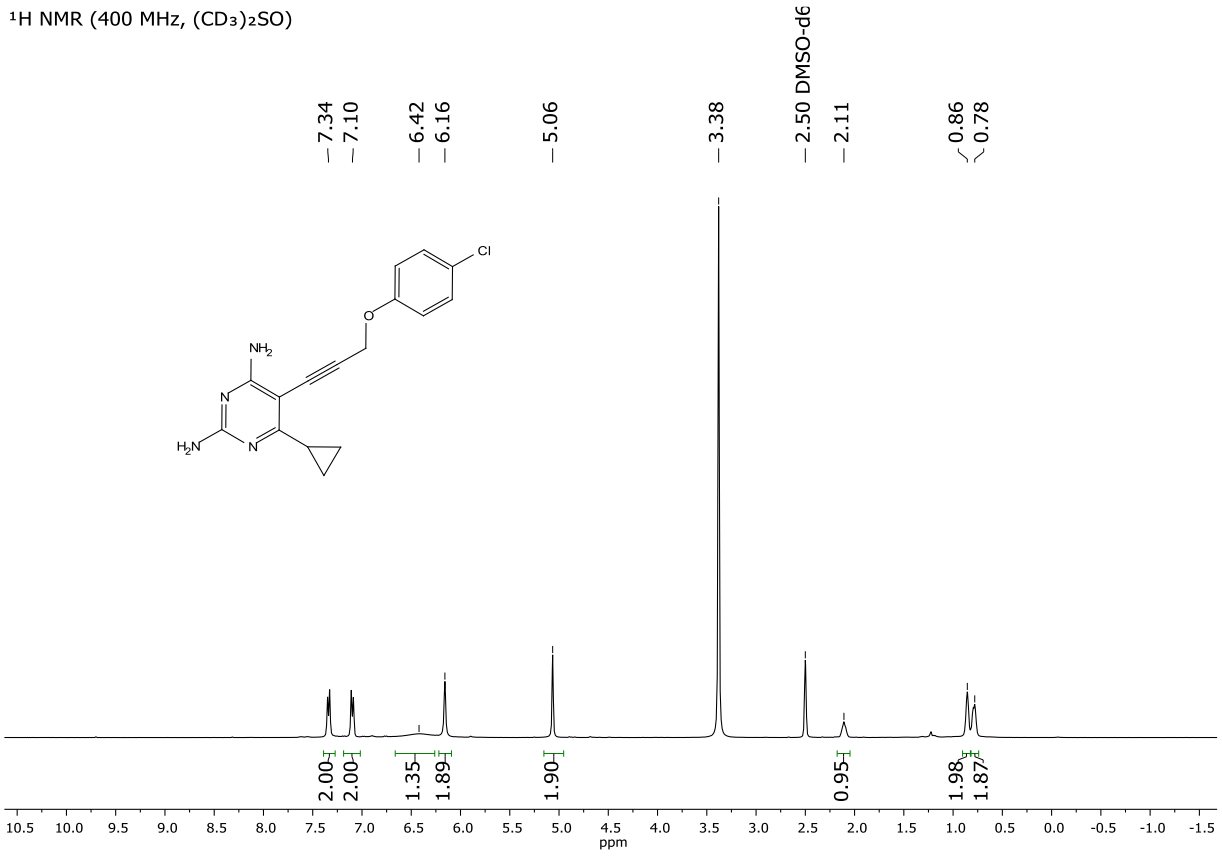
HRMS:


Intensity
339.1848
340.1864
m/z (Da)
[M + H] = 339.1816
339.1848
79.0234
101.0055
m/z (Da)

6-Cyclopropyl-5-[3-(4-chlorophenoxy)prop-1-yn-1-yl]pyrimidine-2,4-diamine (**8j**)

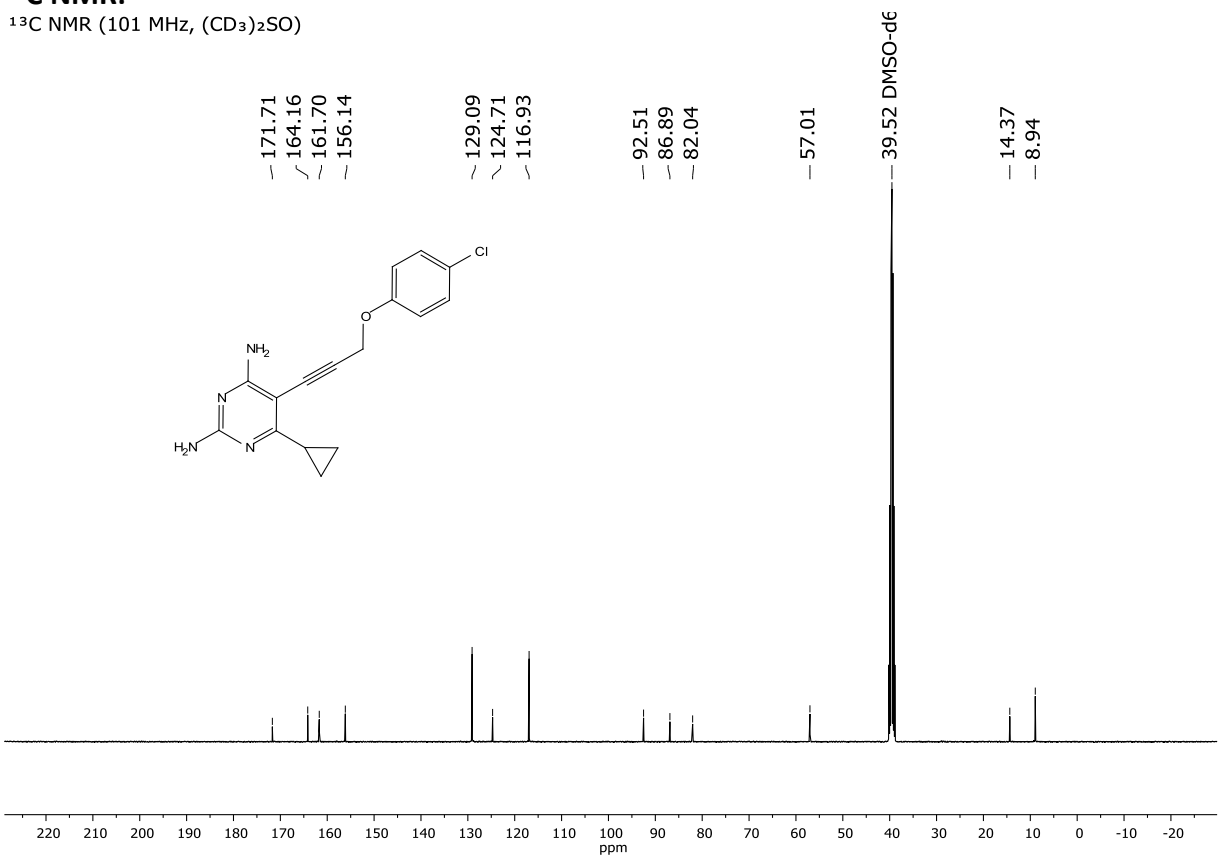
^{1}H NMR:

^{1}H NMR (400 MHz, $(CD_3)_2SO$)

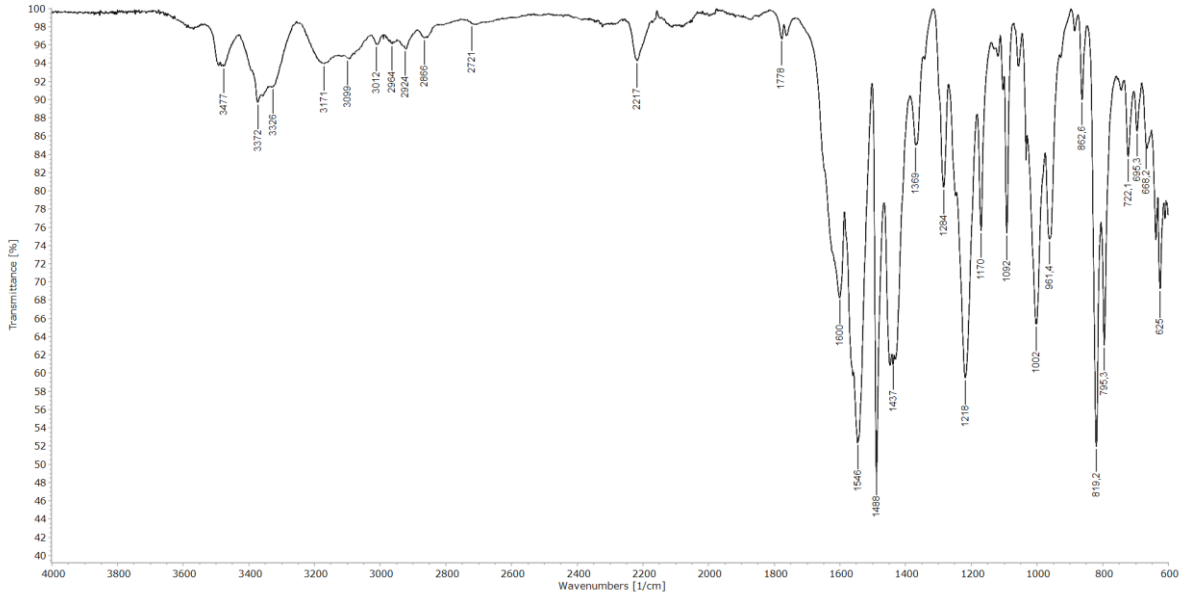


^{13}C NMR:

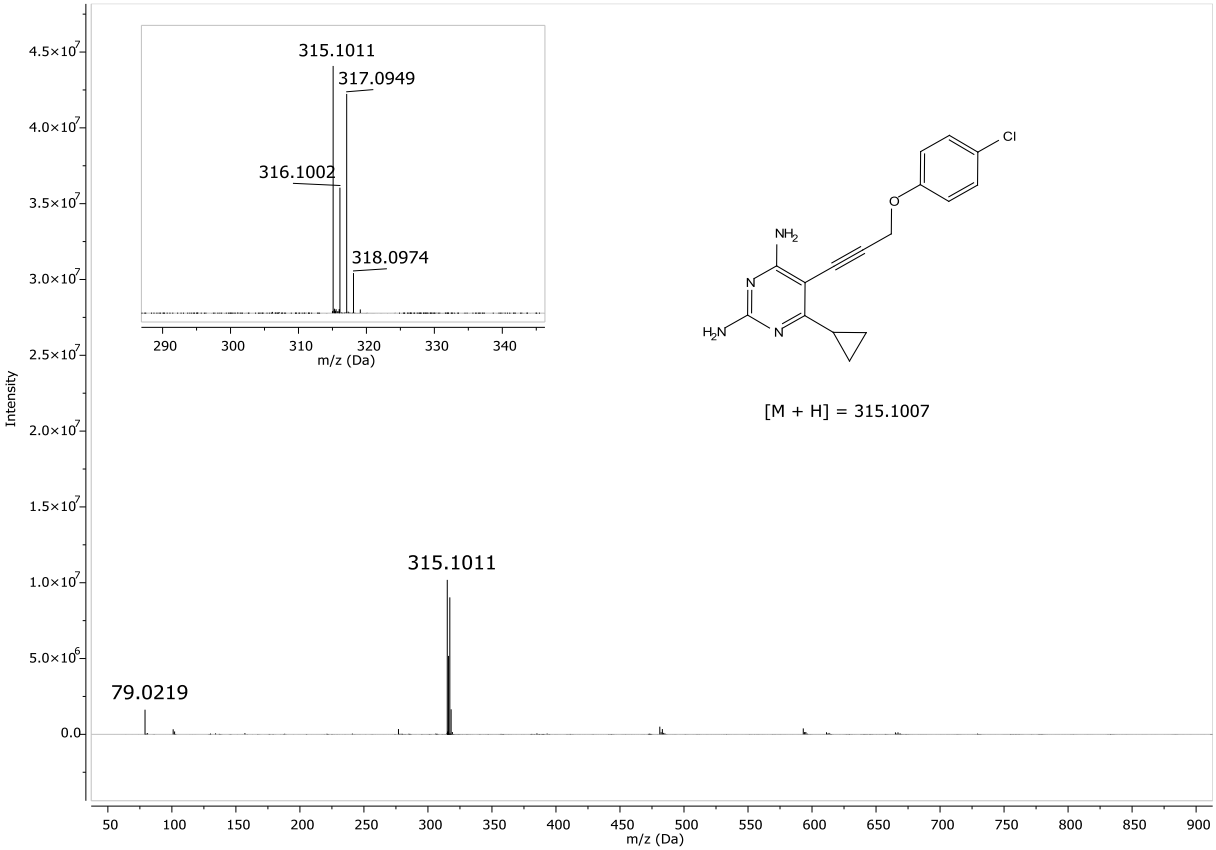
^{13}C NMR (101 MHz, $(CD_3)_2SO$)



FTIR:

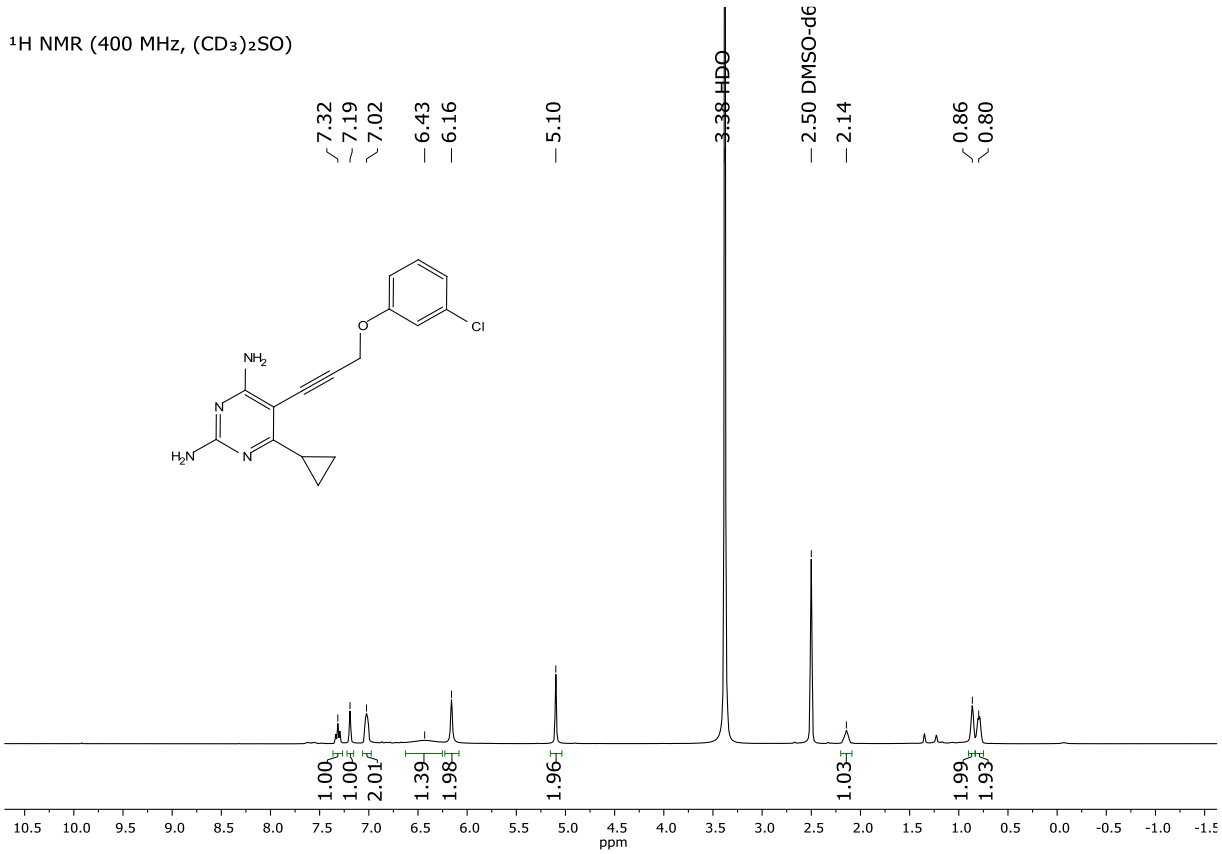

Transmittance [%]
Wavenumbers [1/cm]
3477
3372
3326
3171
3099
3012
2964
2924
2866
2721
2217
1778
1600
1546
1488
1437
1369
1284
1218
1170
1092
1002
961.4
862.6
819.2
795.3
722.1
695.3
665.2
625

HRMS:

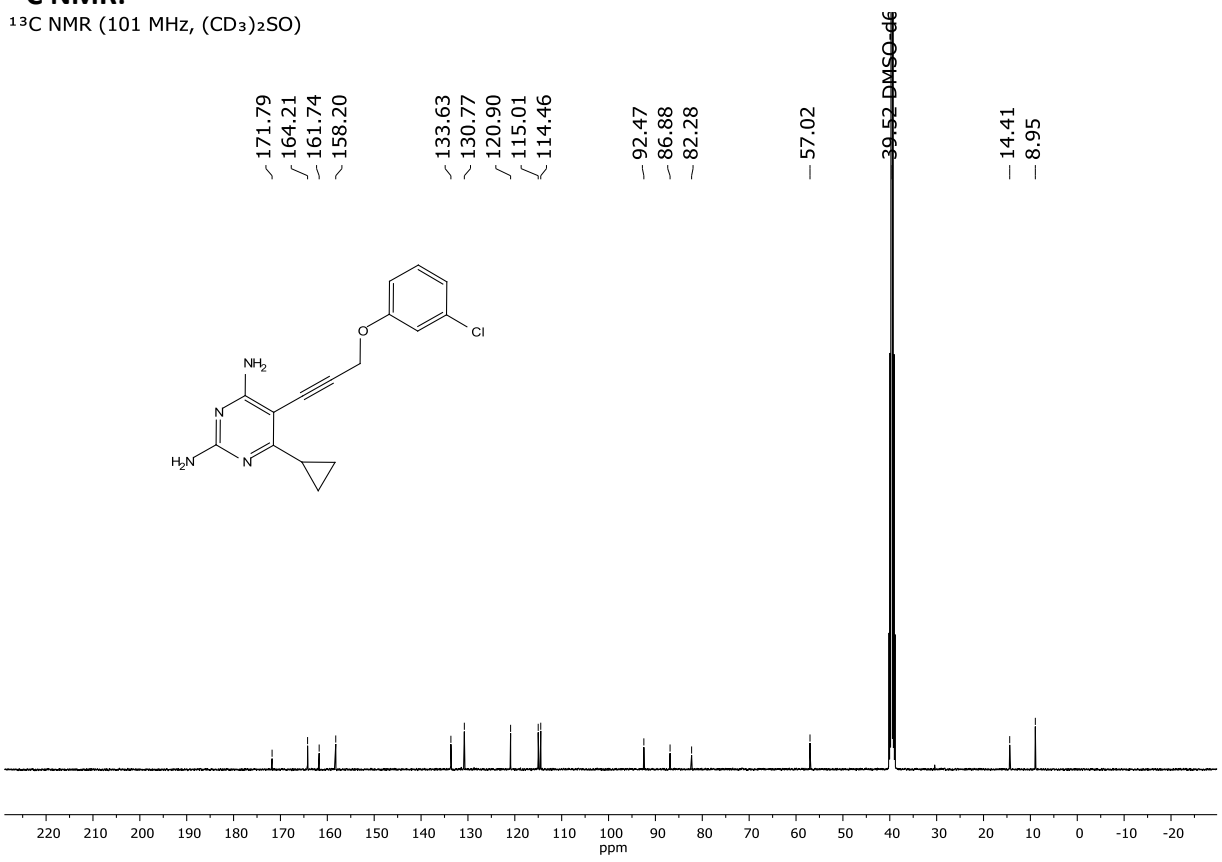

Intensity
m/z (Da)
315.1011
317.0949
316.1002
318.0974
79.0219
[M + H] = 315.1007
Cl
O
NH2
N
H2N

6-Cyclopropyl-5-[3-(3-chlorophenoxy)prop-1-yn-1-yl]pyrimidine-2,4-diamine (**8k**)

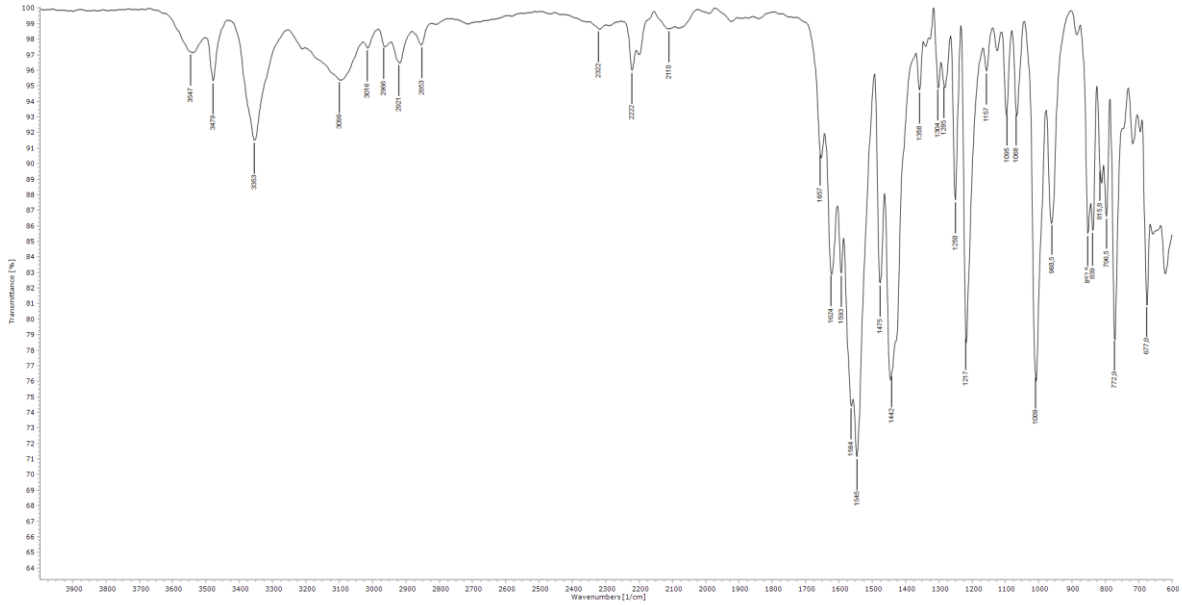
^{1}H NMR:



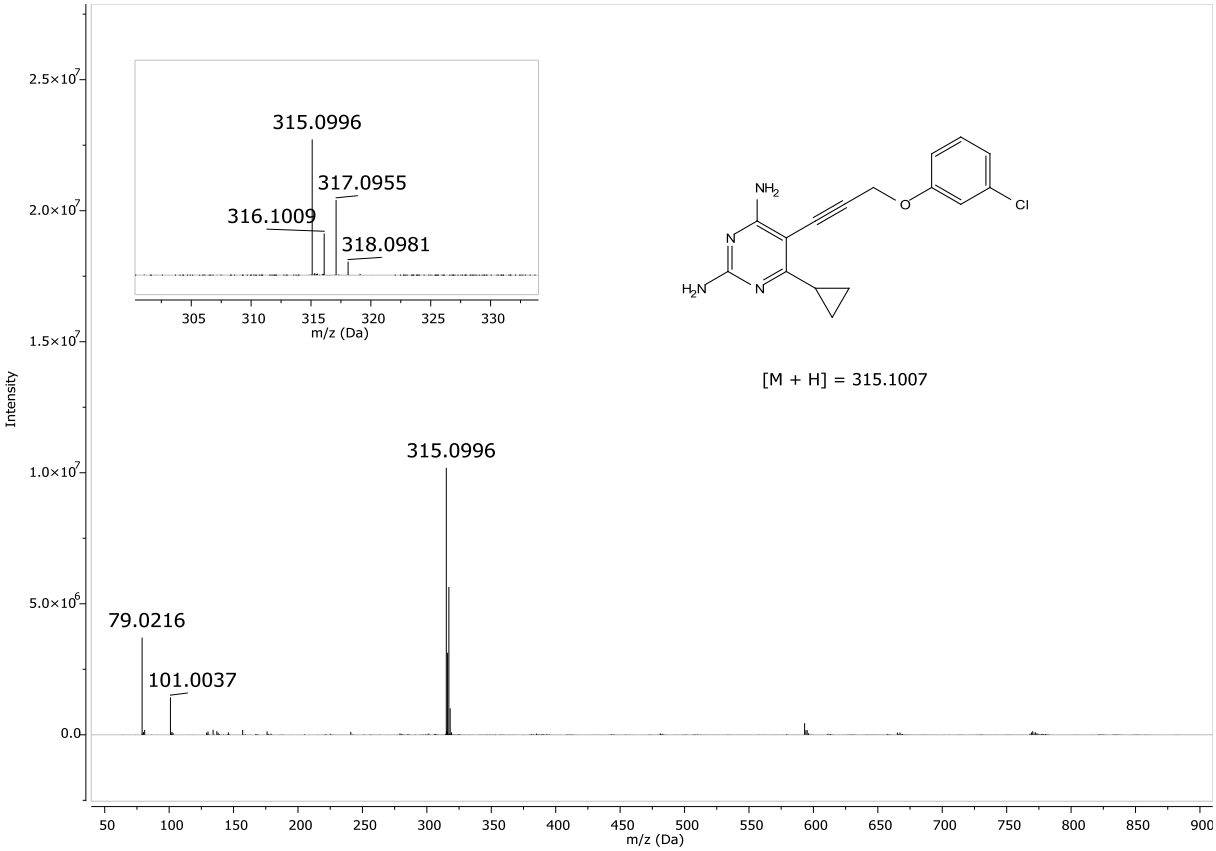
^{13}C NMR:



FTIR:



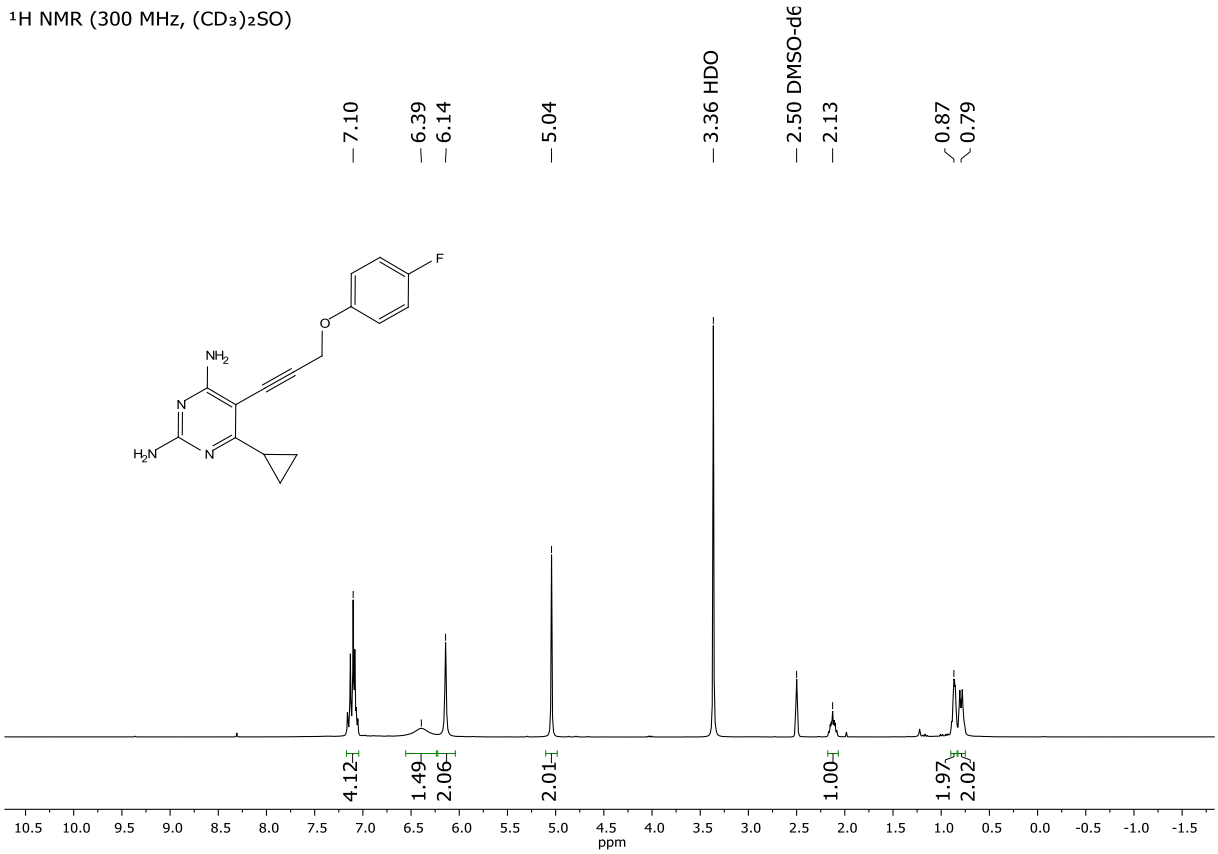
HRMS:


315.0996
317.0955
316.1009
318.0981
305
310
315
320
325
330
m/z (Da)
NH2
N
O
Cl
H2N
N
[M + H] = 315.1007
Intensity
2.5×10^7
2.0×10^7
1.5×10^7
1.0×10^7
5.0×10^6
0.0
315.0996
79.0216
101.0037
50
100
150
200
250
300
350
400
450
500
550
600
650
700
750
800
850
900
m/z (Da)

6-Cyclopropyl-5-[3-(4-flourophenoxy)prop-1-yn-1-yl]pyrimidine-2,4-diamine (**8l**)

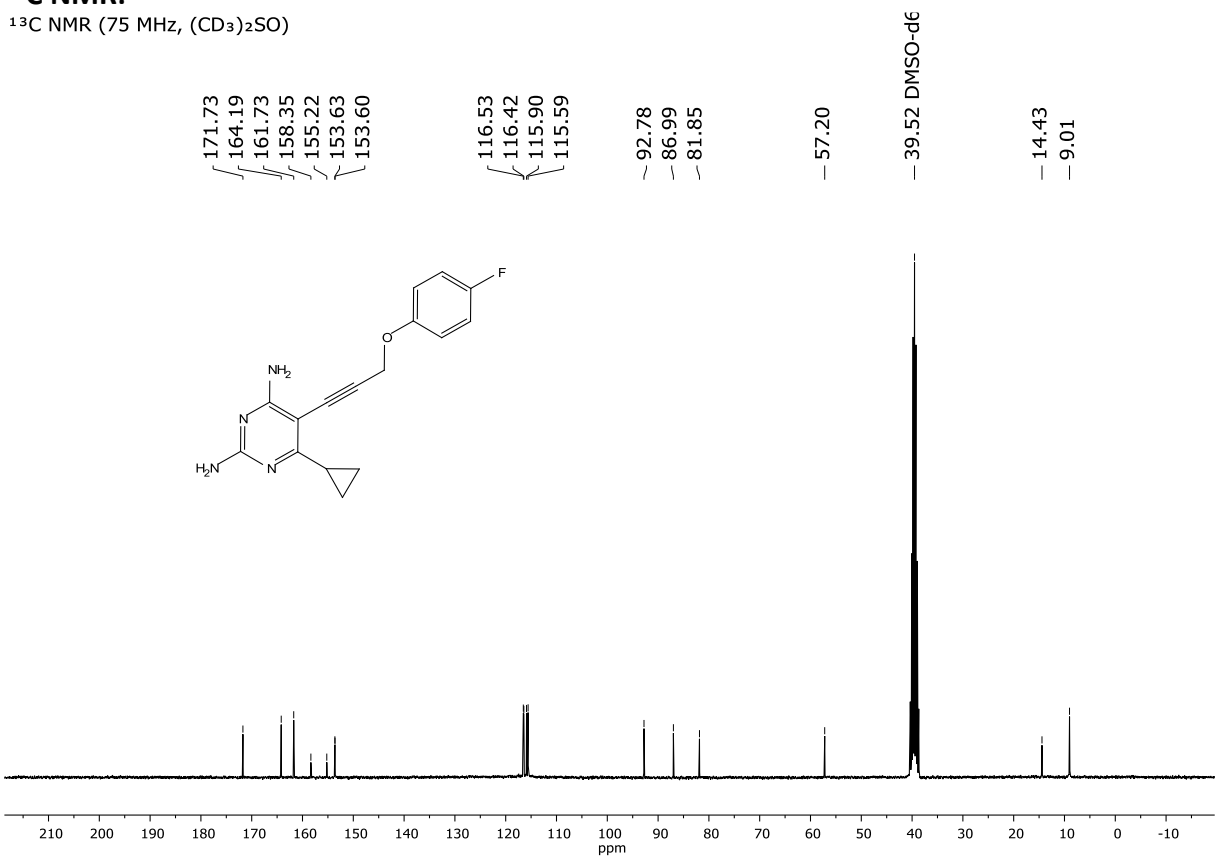
^{1}H NMR:

^{1}H NMR (300 MHz, $(CD_3)_2SO$)

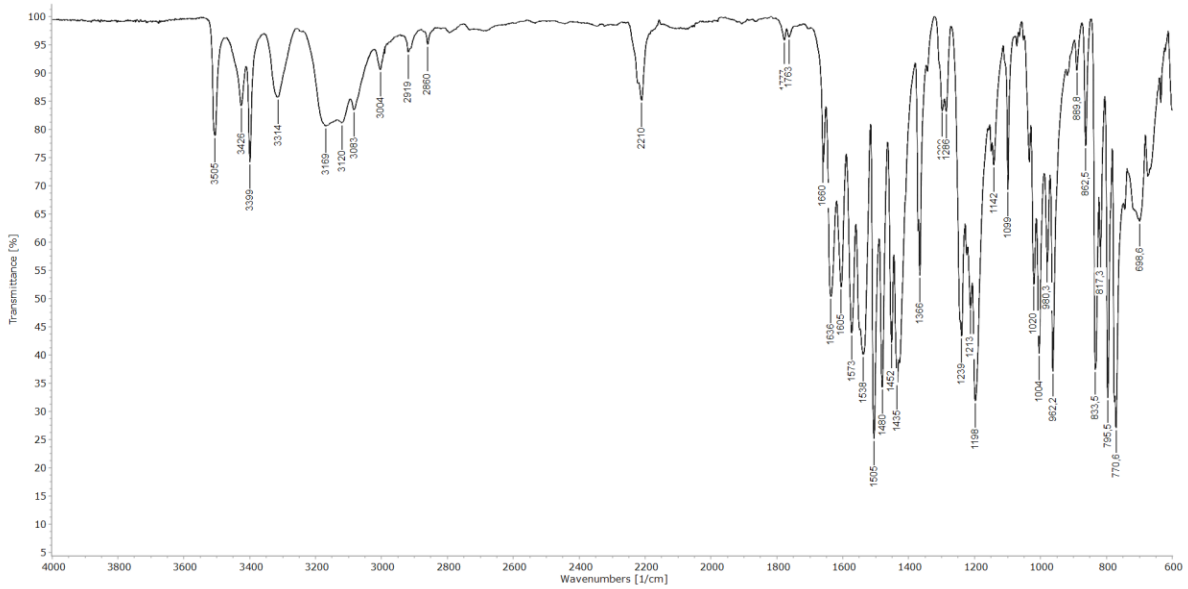


^{13}C NMR:

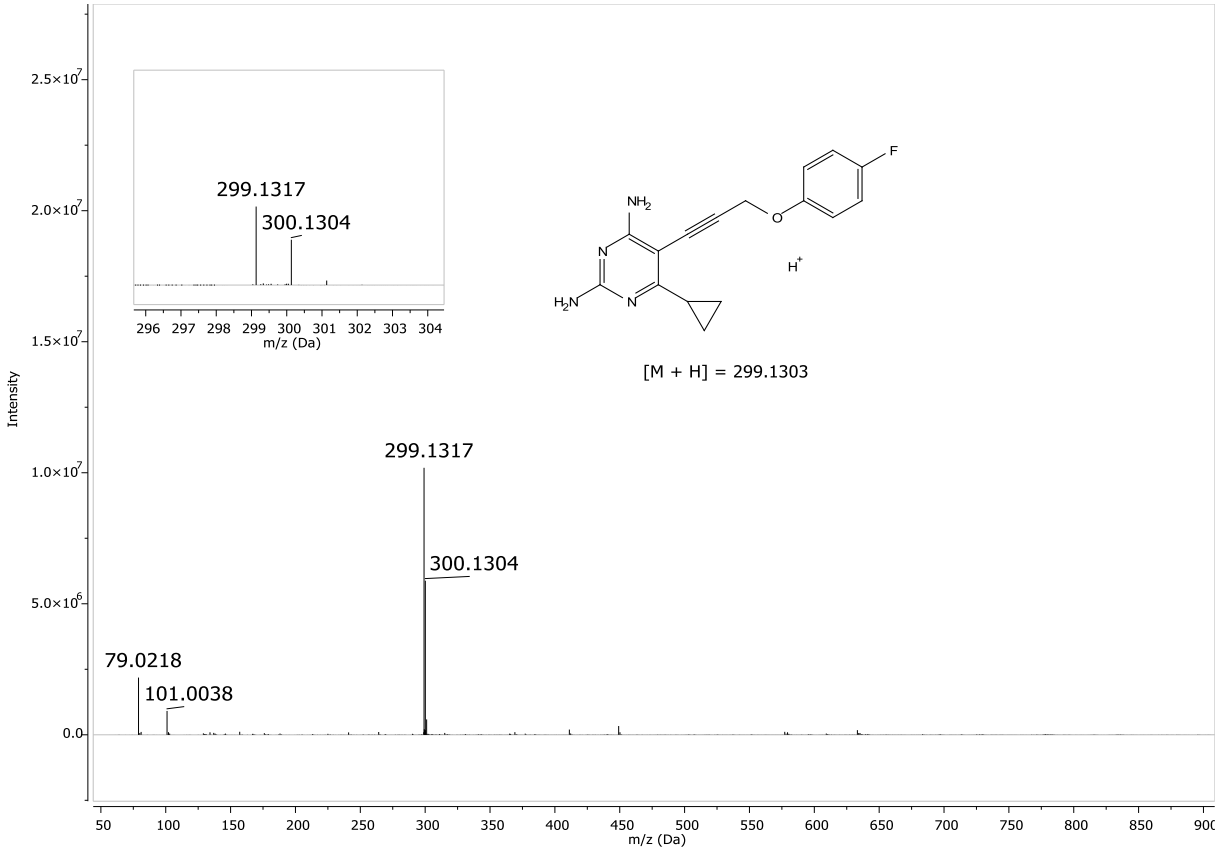
^{13}C NMR (75 MHz, $(CD_3)_2SO$)



FTIR:

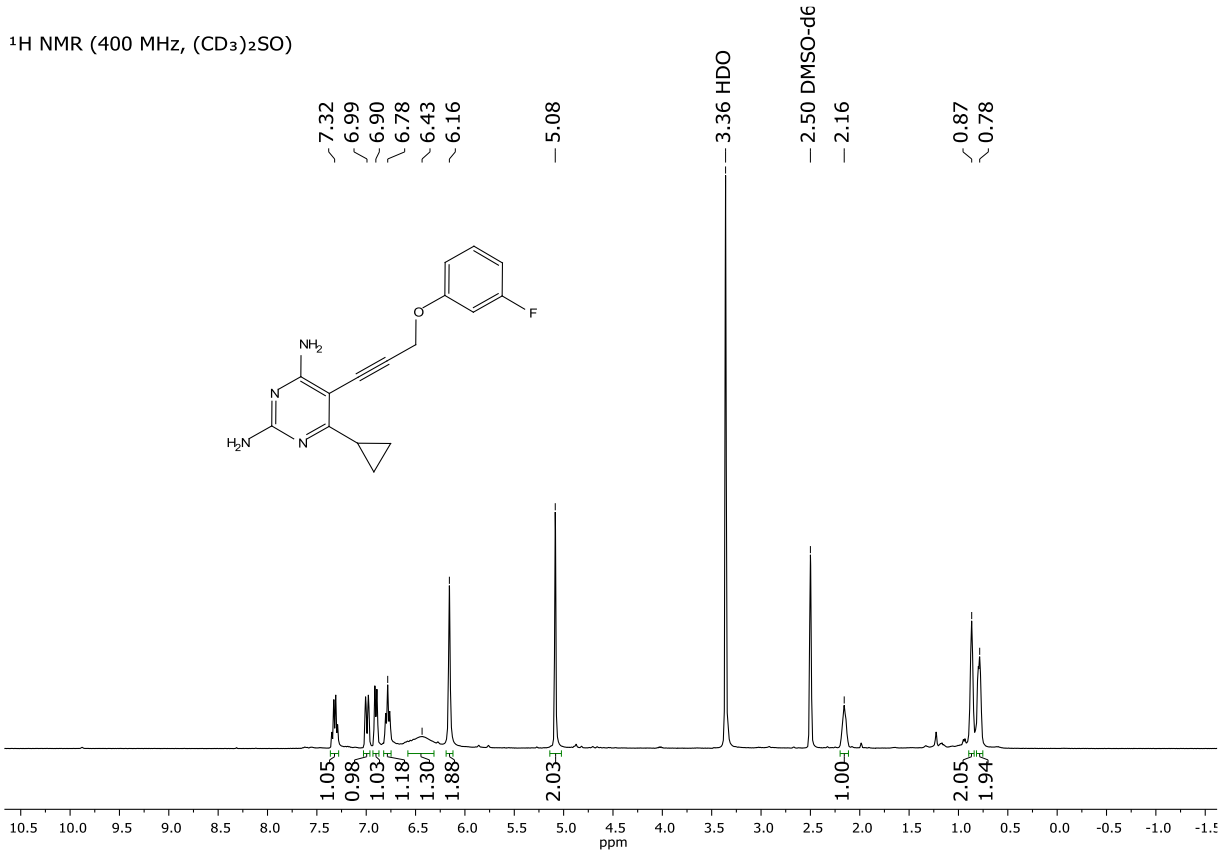

Transmittance [%]
Wavenumbers [1/cm]
3505
3426
3399
3314
3169
3120
3083
3004
2919
2860
2210
1777
1763
1660
1636
1605
1573
1538
1505
1480
1452
1435
1366
1286
1239
1213
1198
1142
1099
1020
1004
962.2
889.8
862.5
833.5
817.3
796.5
770.8
698.6

HRMS:

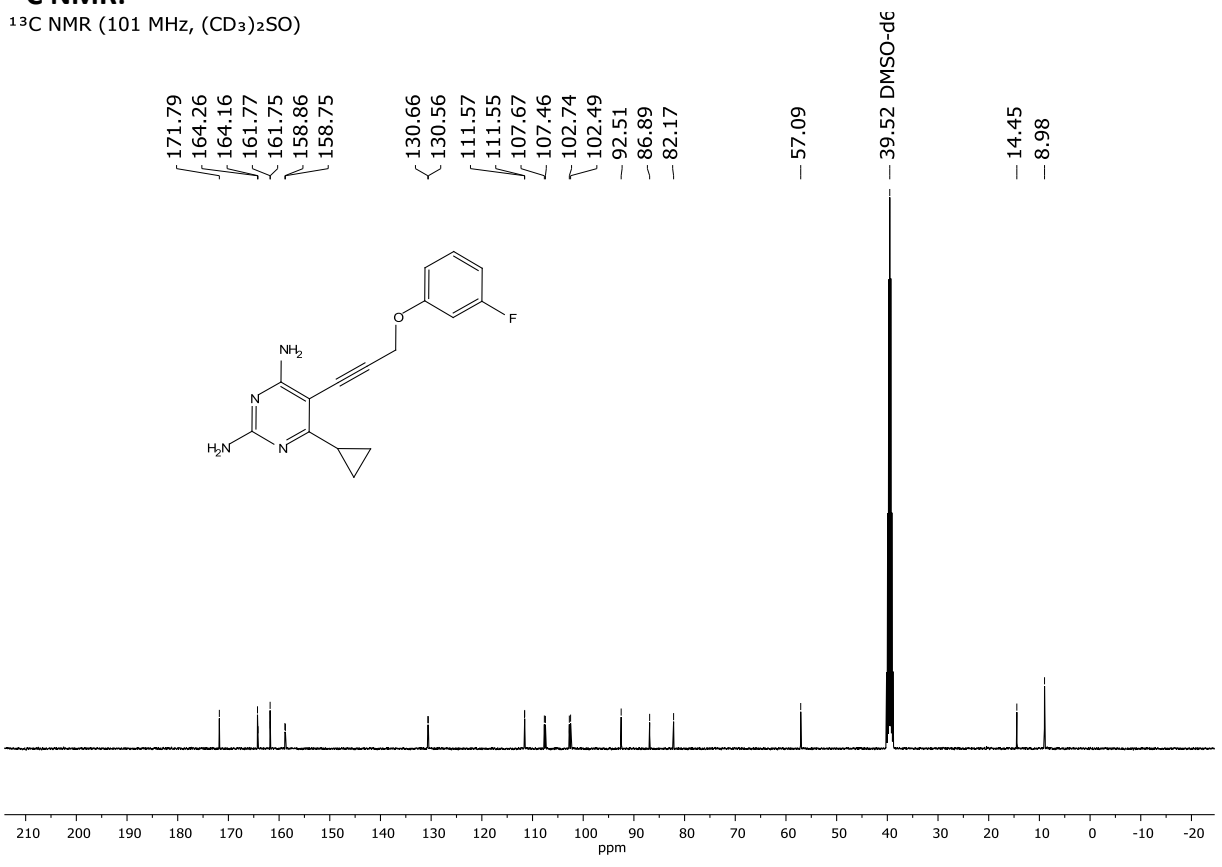

299.1317
300.1304
296 297 298 299 300 301 302 303 304
m/z (Da)
NH2
N
H2N
N
O
F
H+
[M + H] = 299.1303
Intensity
2.5×10^7
2.0×10^7
1.5×10^7
1.0×10^7
5.0×10^6
0.0
299.1317
300.1304
79.0218
101.0038
50 100 150 200 250 300 350 400 450 500 550 600 650 700 750 800 850 900
m/z (Da)

6-Cyclopropyl-5-[3-(3-flourophenoxy)prop-1-yn-1-yl]pyrimidine-2,4-diamine (**8m**)

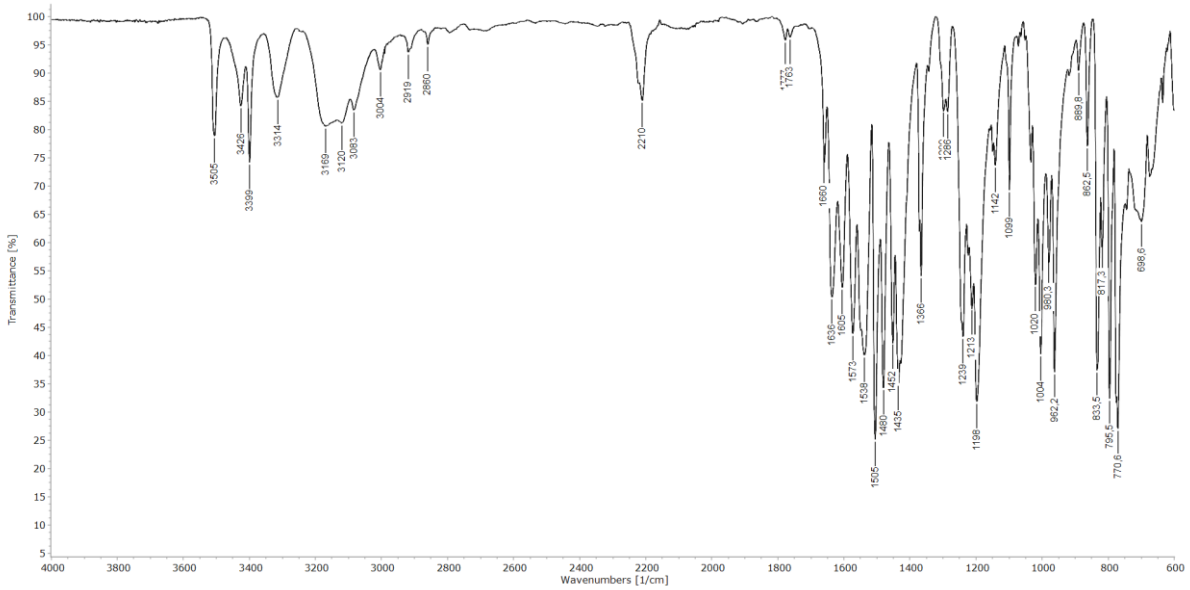
^{1}H NMR:



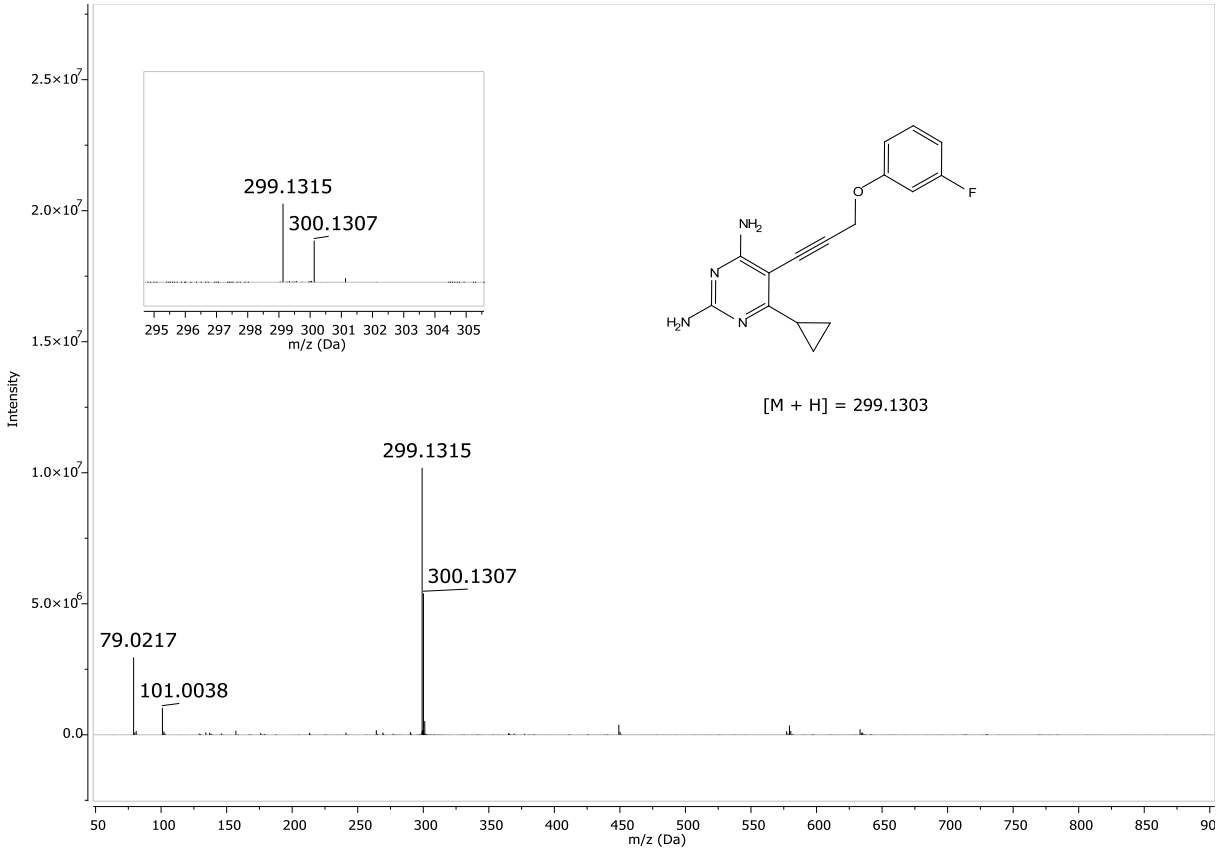
^{13}C NMR:



FTIR:

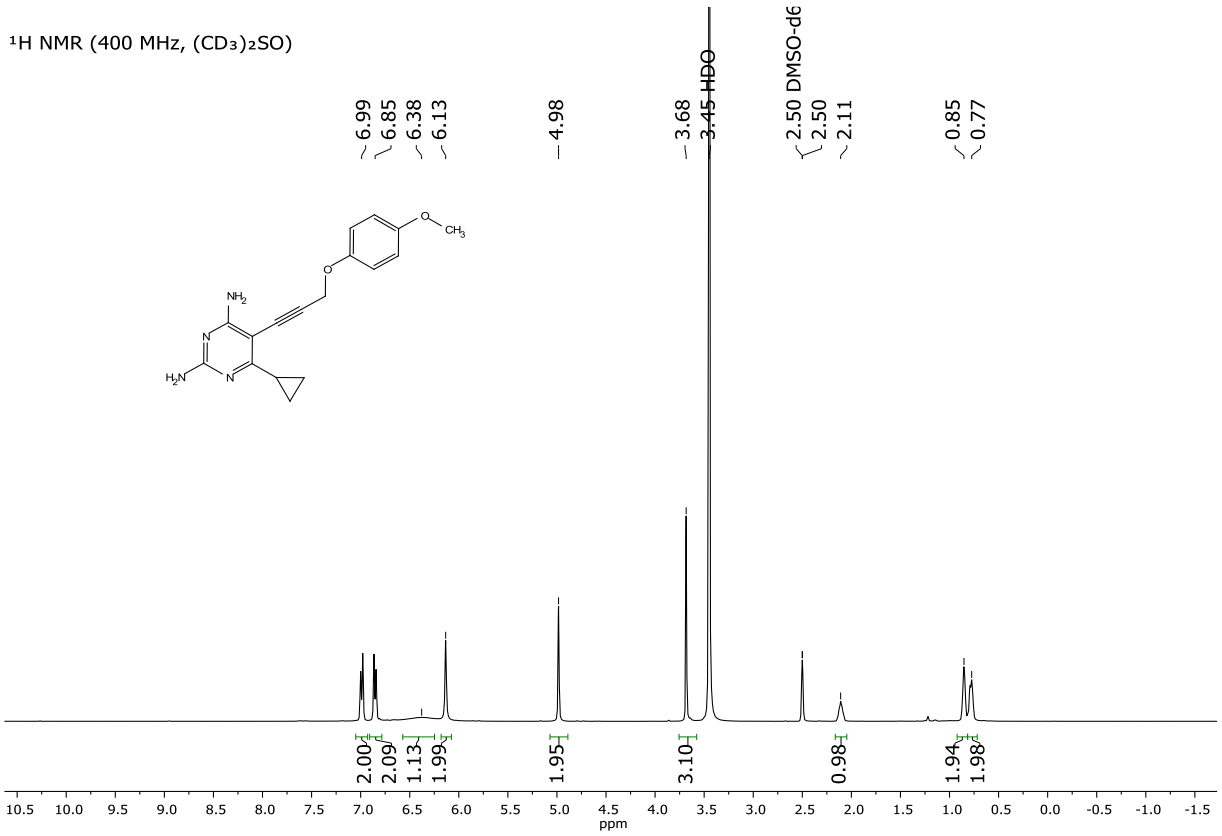

Transmittance [%]
Wavenumbers [1/cm]
3505
3426
3399
3314
3189
3120
3083
3004
2919
2860
2210
1777
1763
1660
1636
1605
1573
1538
1505
1480
1452
1435
1366
1299
1289
1239
1213
1198
1142
1099
1020
980.3
962.2
1004
889.8
862.5
833.5
817.3
795.5
770.8
698.6

HRMS:

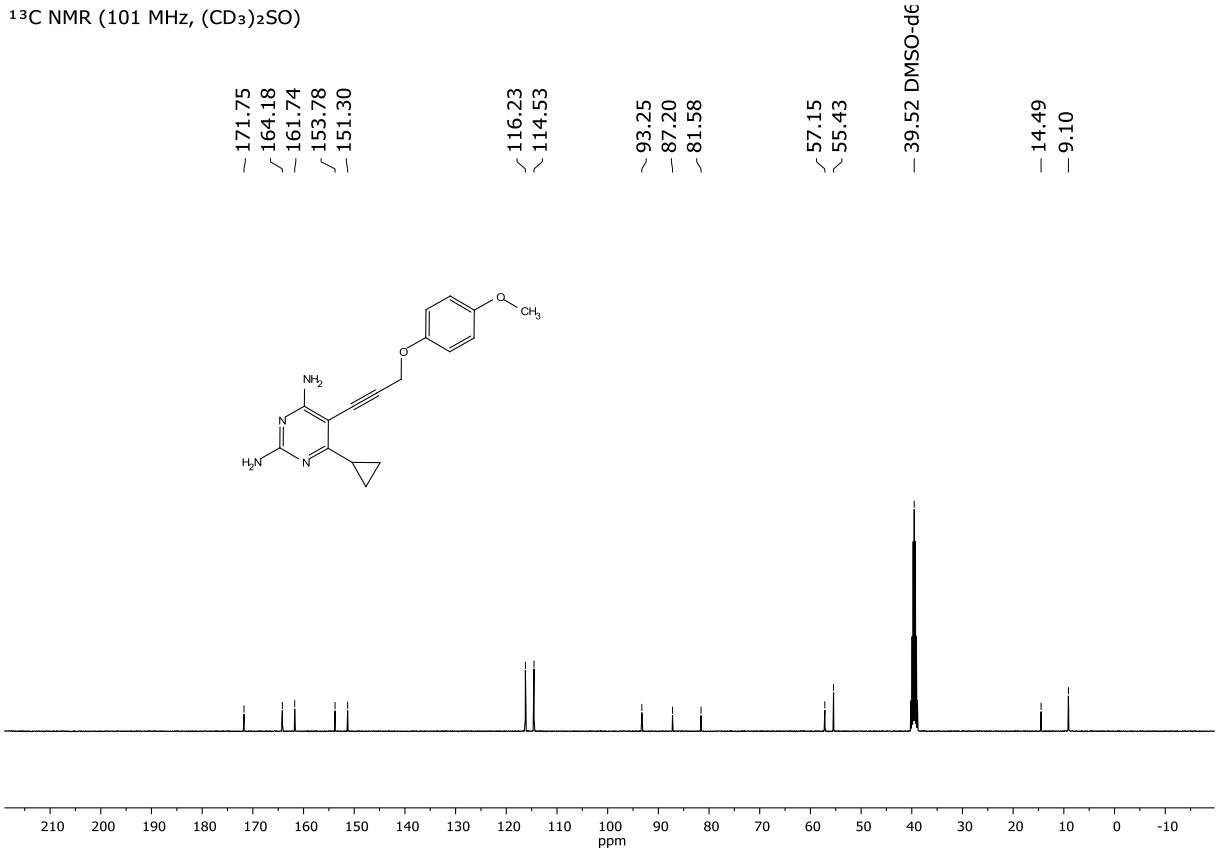

299.1315
300.1307
m/z (Da)
[M + H] = 299.1303
Intensity
299.1315
300.1307
79.0217
101.0038
m/z (Da)

6-Cyclopropyl-5-[3-(4-methoxyphenoxy)prop-1-yn-1-yl]pyrimidine-2,4-diamine (**8n**)

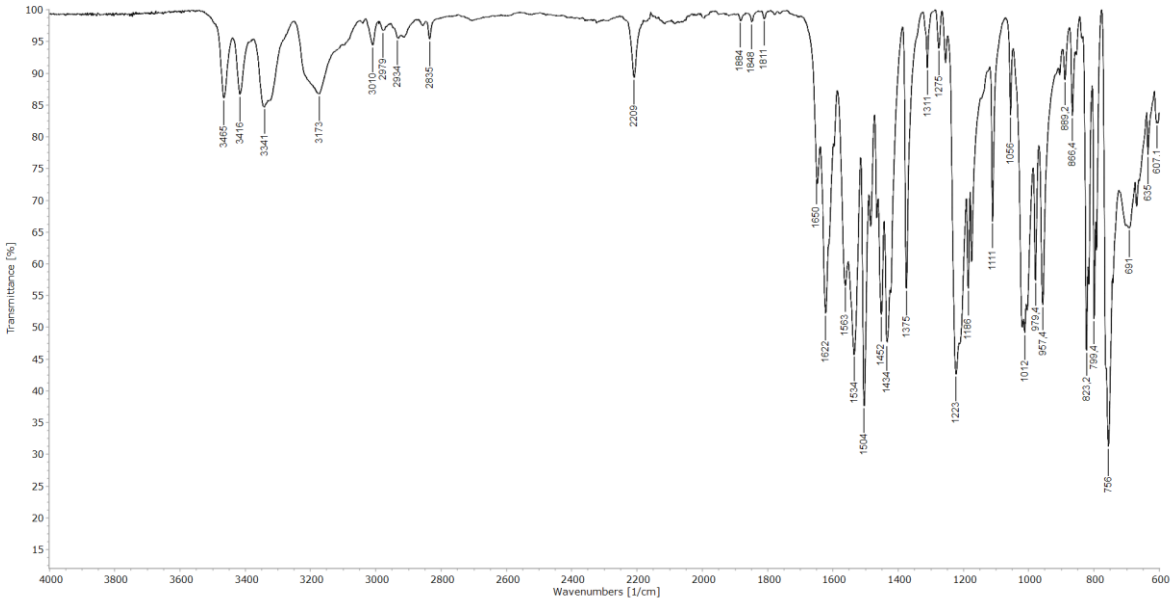
^{1}H NMR:



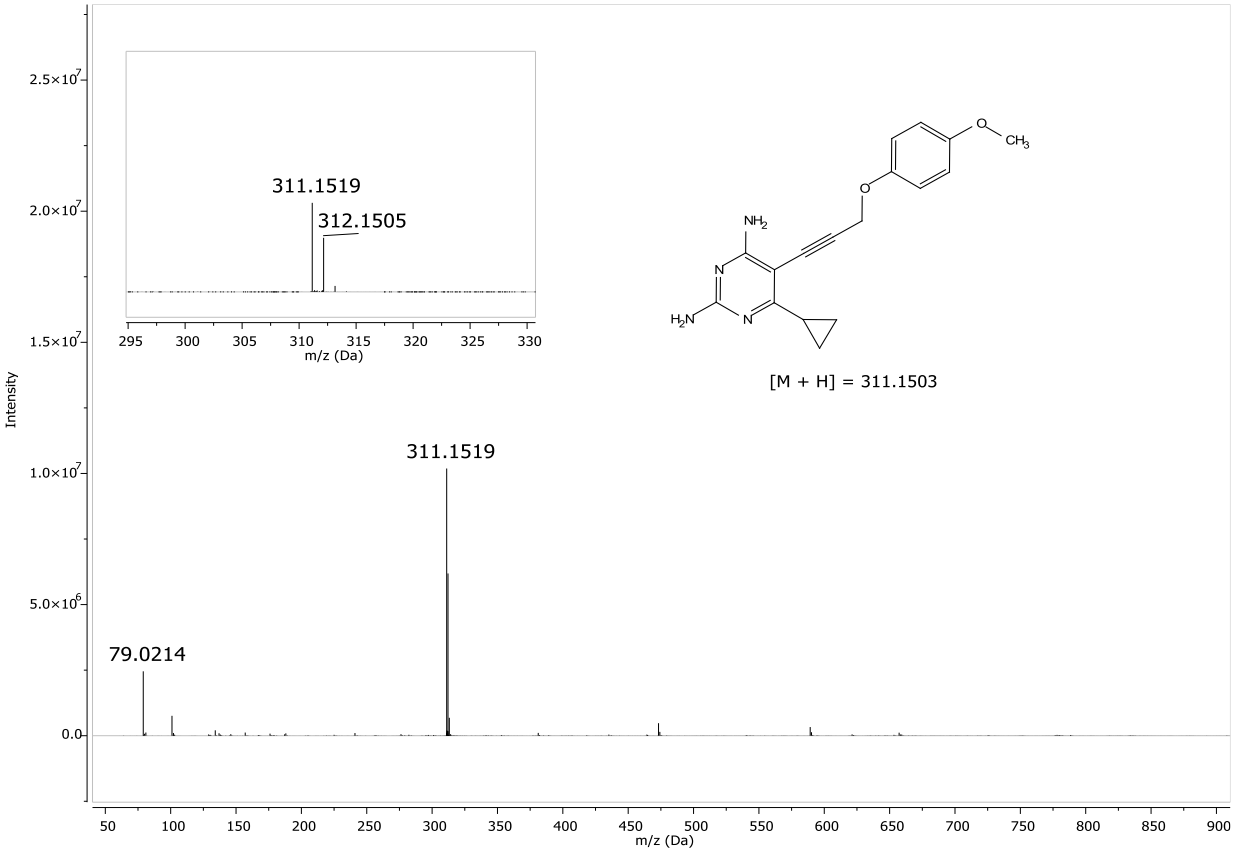
^{13}C NMR:



FTIR:


Transmittance [%]
Wavenumbers [1/cm]
3465
3416
3341
3173
3010
2979
2934
2835
2209
1884
1848
1811
1650
1622
1563
1534
1504
1452
1434
1375
1311
1275
1223
1186
1111
1056
1012
979.4
957.4
889.2
866.4
823.2
799.4
756
691
635
607.1

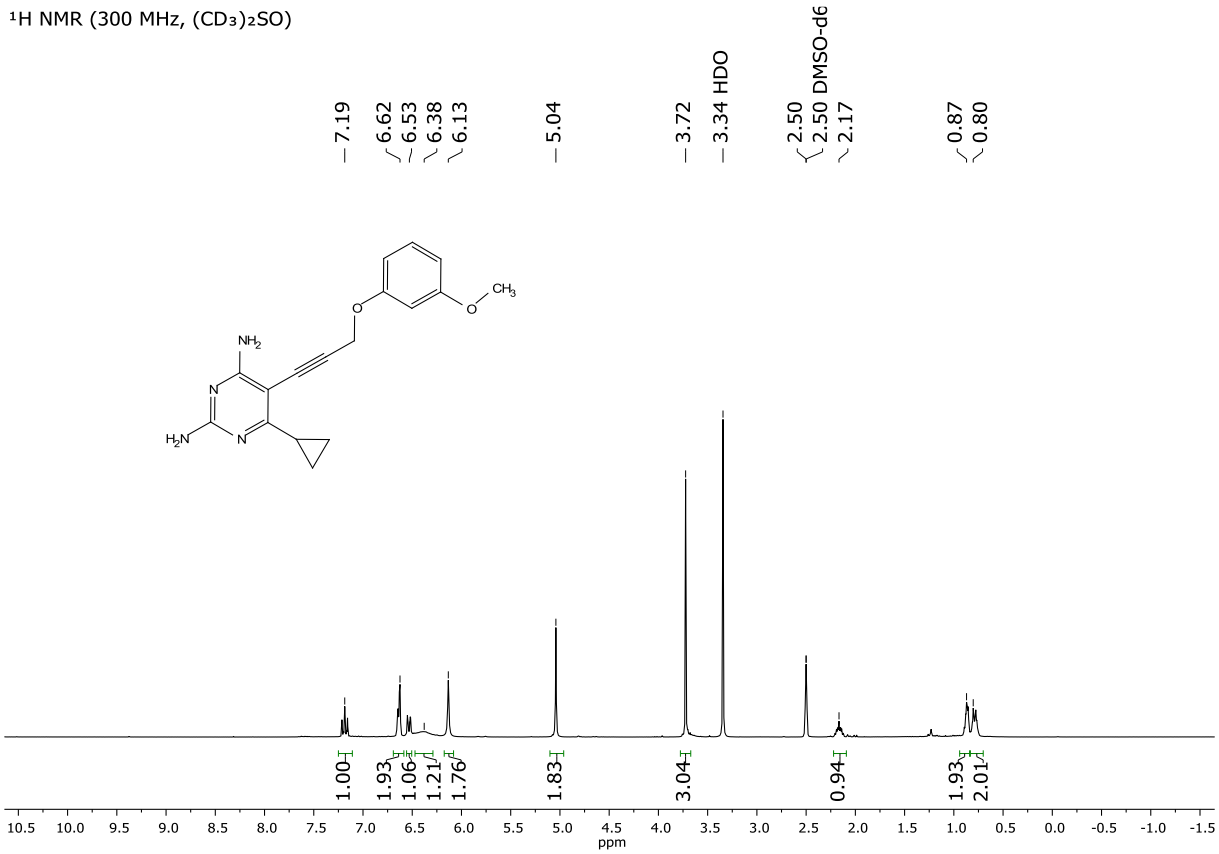
HRMS:


311.1519
312.1505
m/z (Da)
[M + H] = 311.1503
Intensity
311.1519
79.0214
m/z (Da)

6-Cyclopropyl-5-[3-(3-methoxyphenoxy)prop-1-yn-1-yl]pyrimidine-2,4-diamine (**8o**)

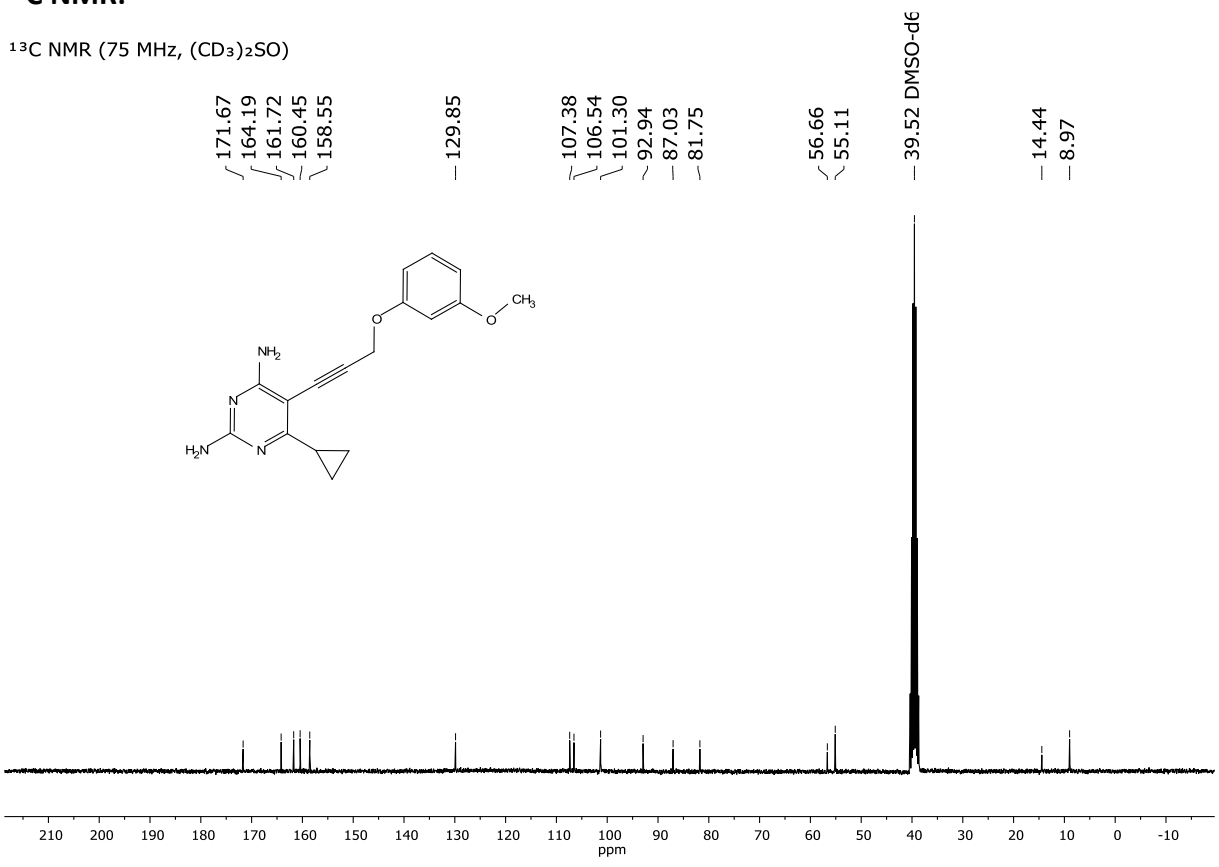
^{1}H NMR:

^{1}H NMR (300 MHz, $(CD_3)_2SO$)

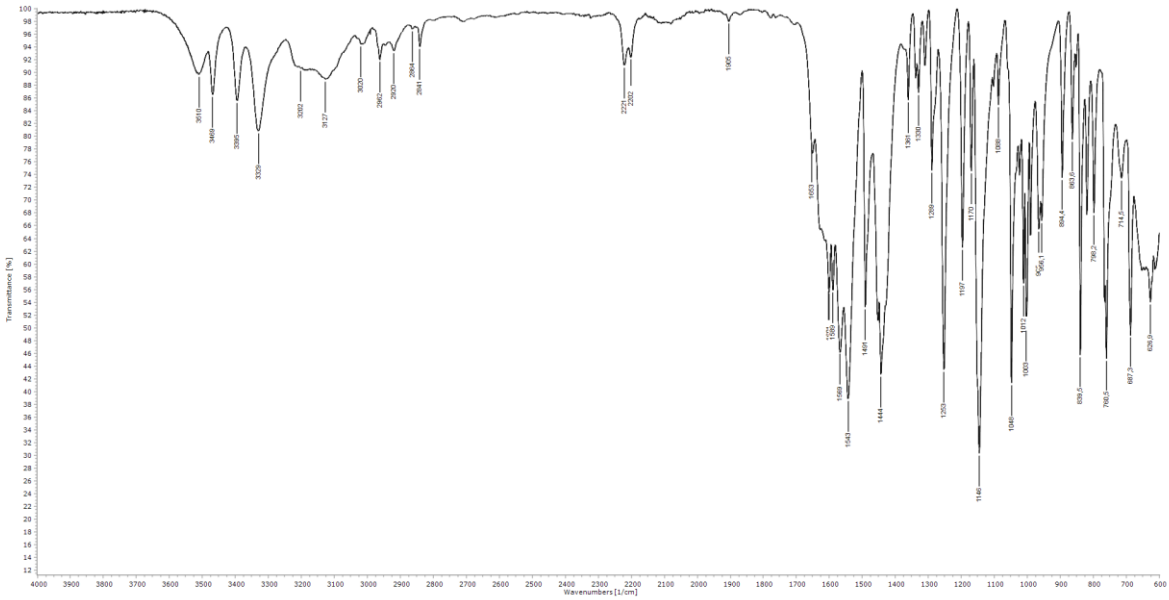


^{13}C NMR:

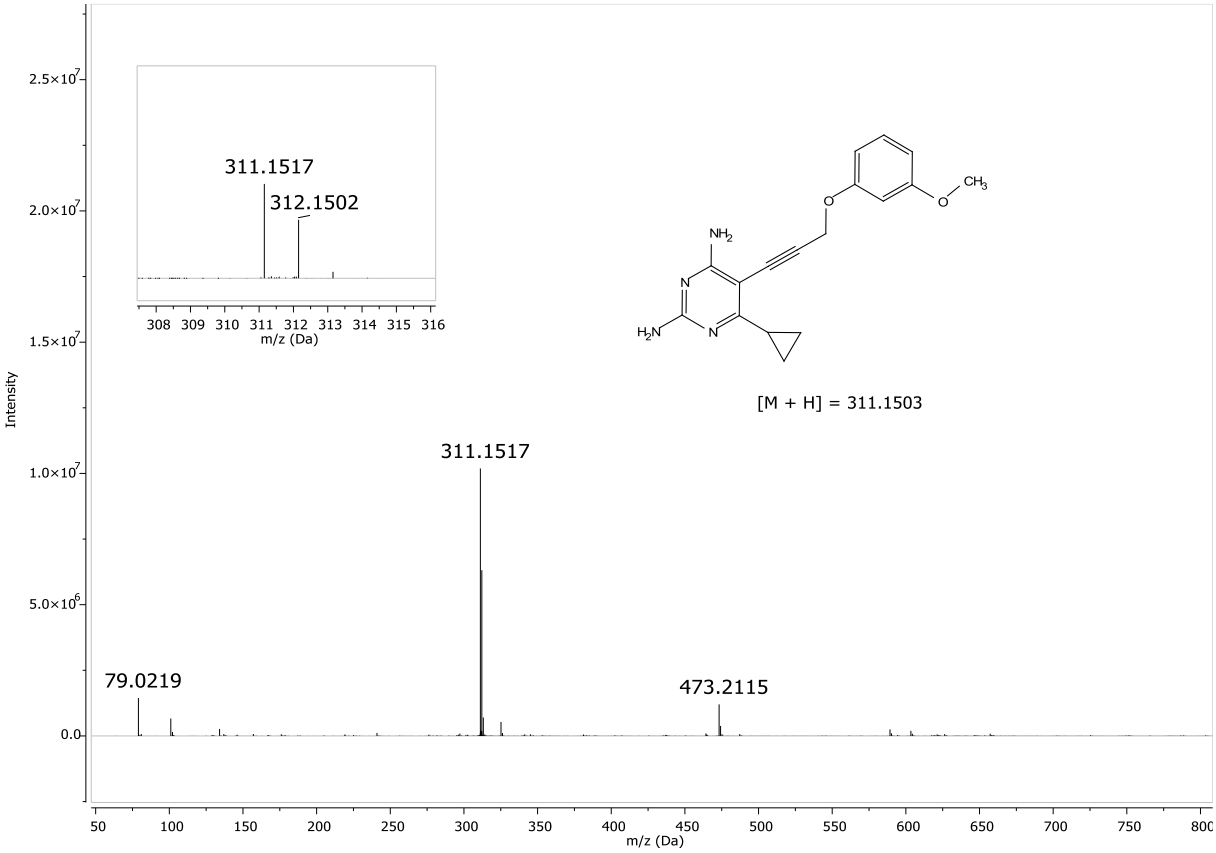
^{13}C NMR (75 MHz, $(CD_3)_2SO$)



FTIR:



HRMS:

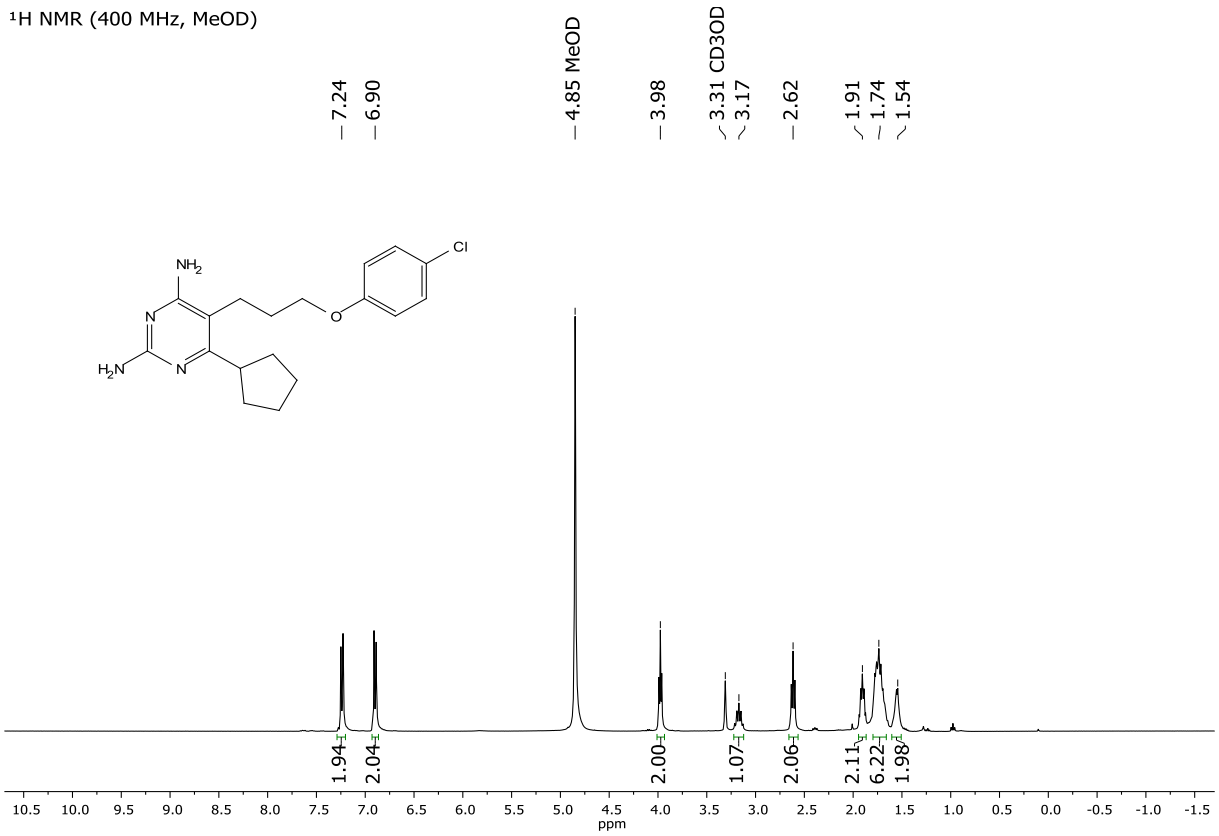

311.1517
312.1502
308 309 310 311 312 313 314 315 316
m/z (Da)
NH2
N
H2N
N
O
O
CH3
[M + H] = 311.1503
Intensity
311.1517
79.0219
473.2115
50 100 150 200 250 300 350 400 450 500 550 600 650 700 750 800
m/z (Da)

Hydrogenation of alkynes (**9**)

6-Cyclopentyl-5-[3-(4-chlorophenoxy)propyl]pyrimidine-2,4-diamine (**9a**)

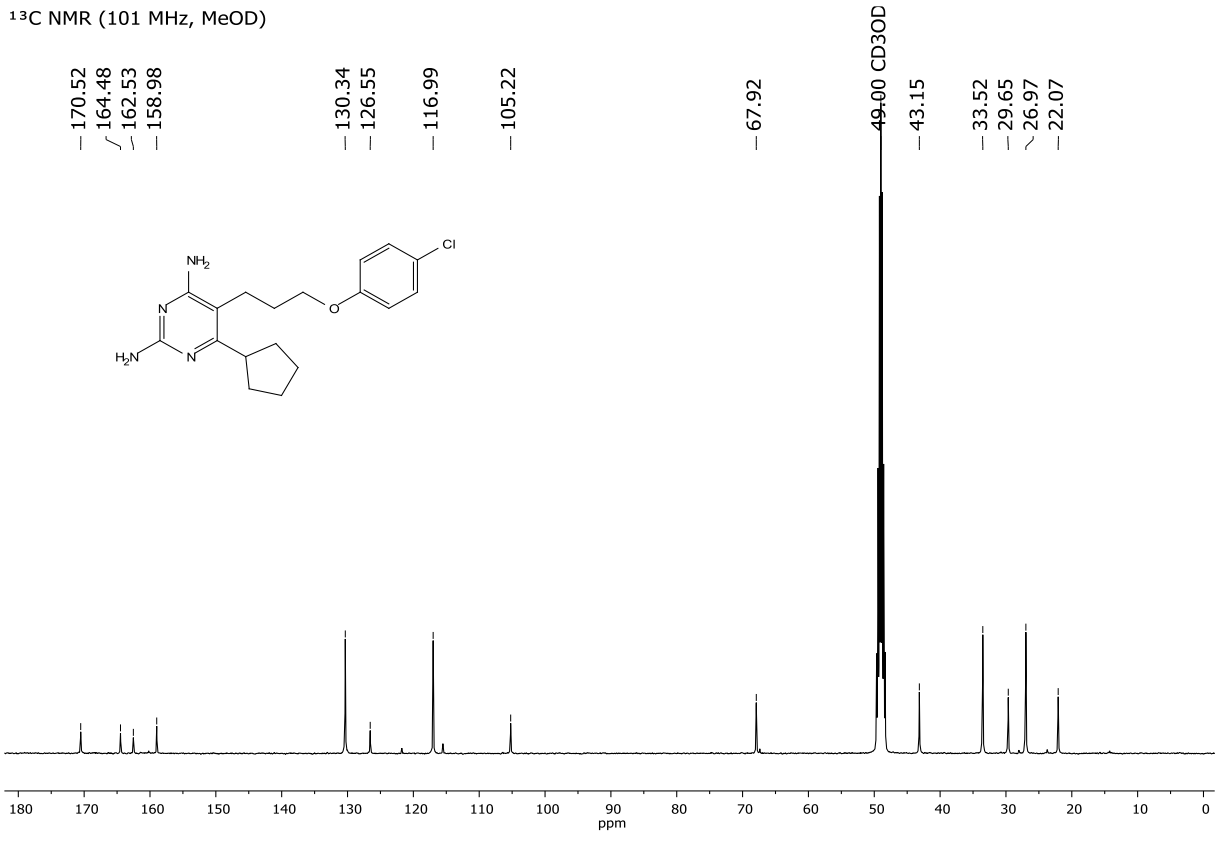
^{1}H NMR:

^{1}H NMR (400 MHz, MeOD)

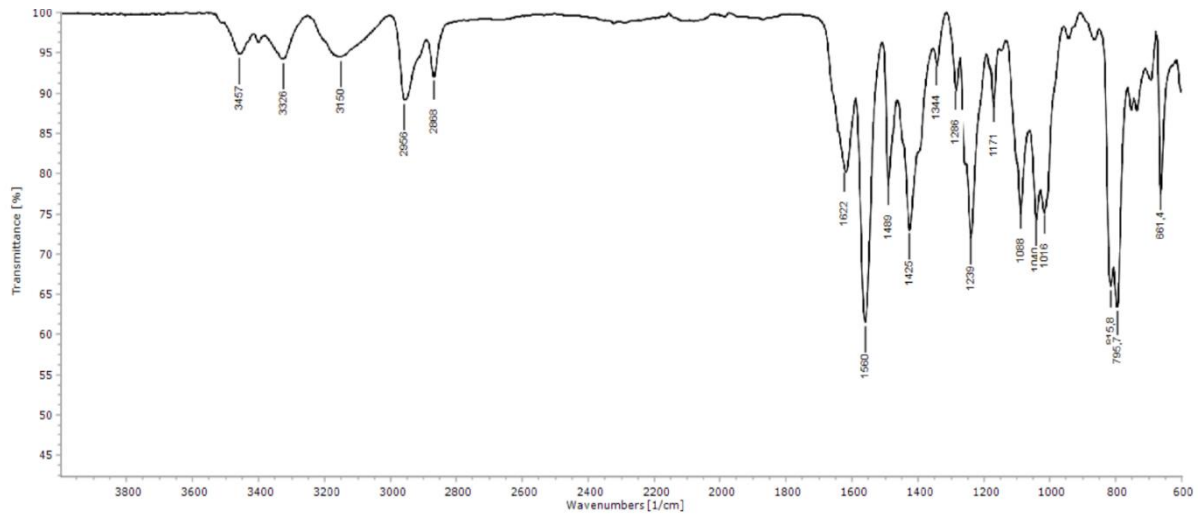


^{13}C NMR:

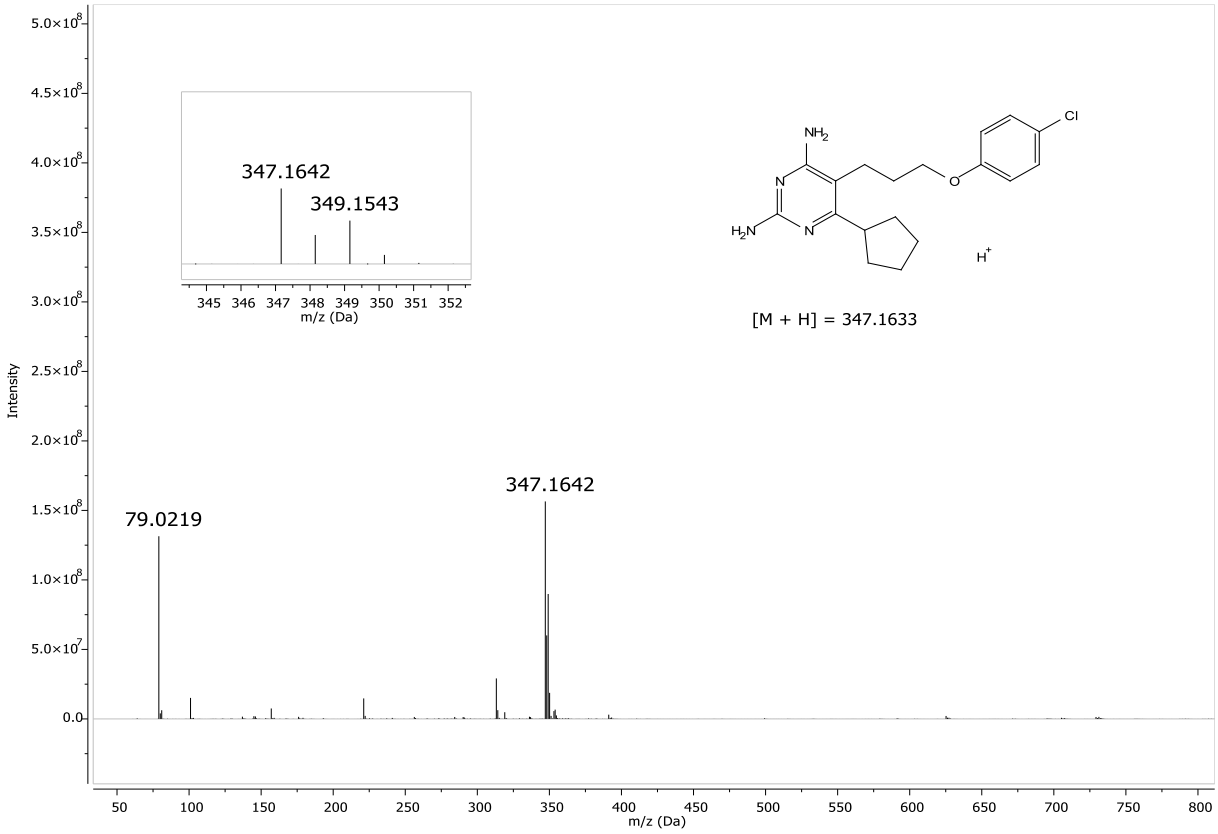
^{13}C NMR (101 MHz, MeOD)



FTIR:


Transmittance [%]
Wavenumbers [1/cm]
3457
3326
3150
2959
2868
1622
1580
1489
1425
1344
1288
1239
1171
1068
1049
1016
915,8
796,7
861,4

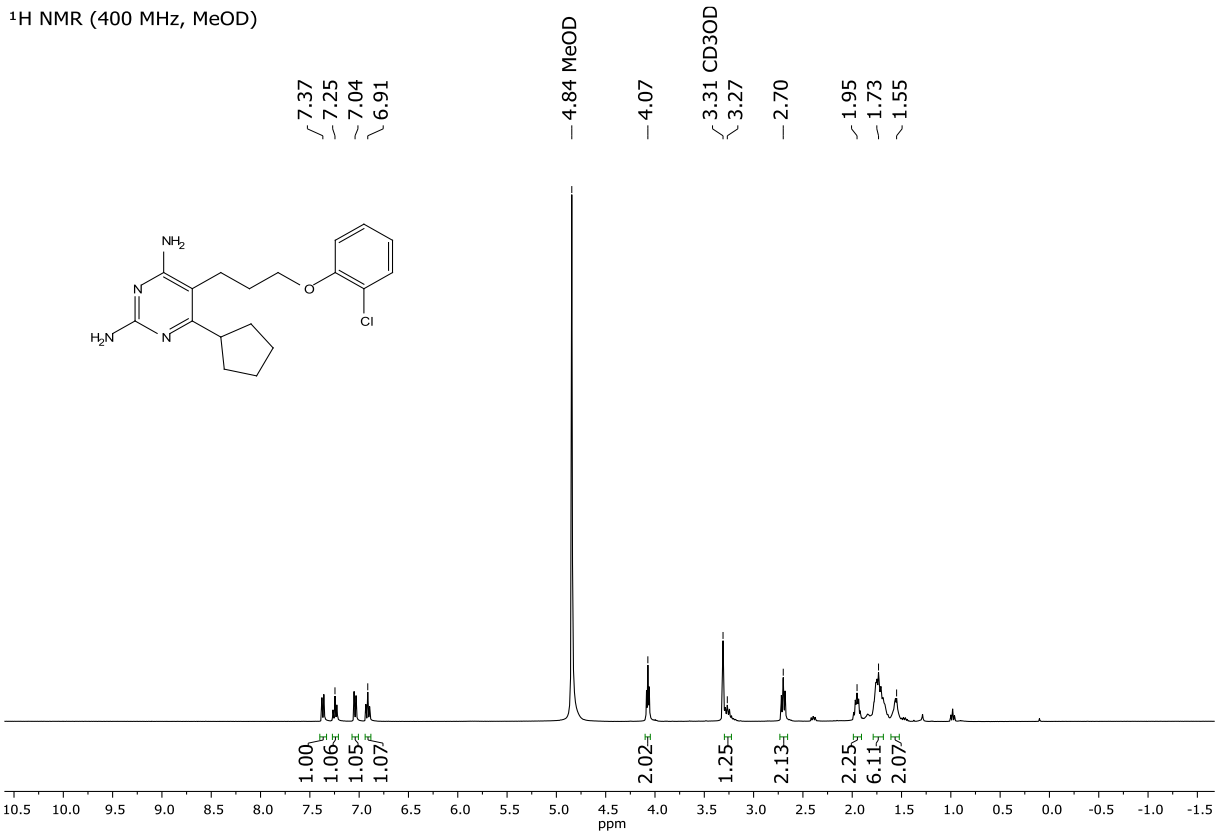
HRMS:


347.1642
349.1543
m/z (Da)
[M + H] = 347.1633
Intensity
79.0219
347.1642
m/z (Da)

6-Cyclopentyl-5-[3-(2-chlorophenoxy)propyl]pyrimidine-2,4-diamine (**9b**)

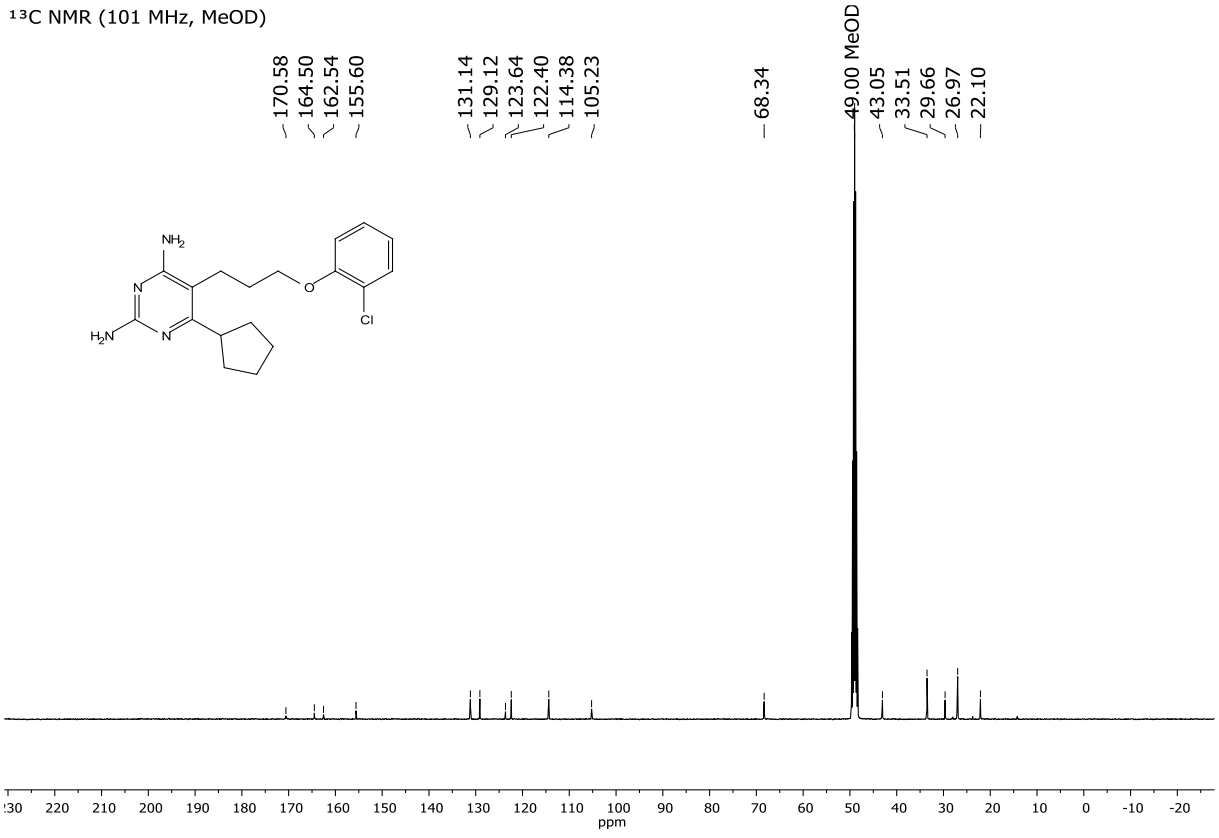
^{1}H NMR:

^{1}H NMR (400 MHz, MeOD)

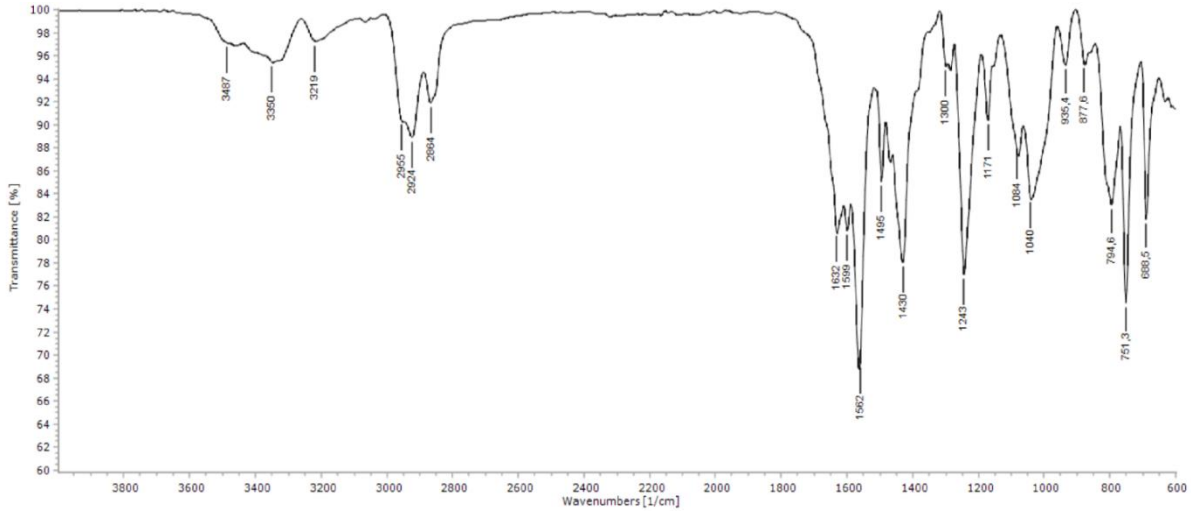


^{13}C NMR:

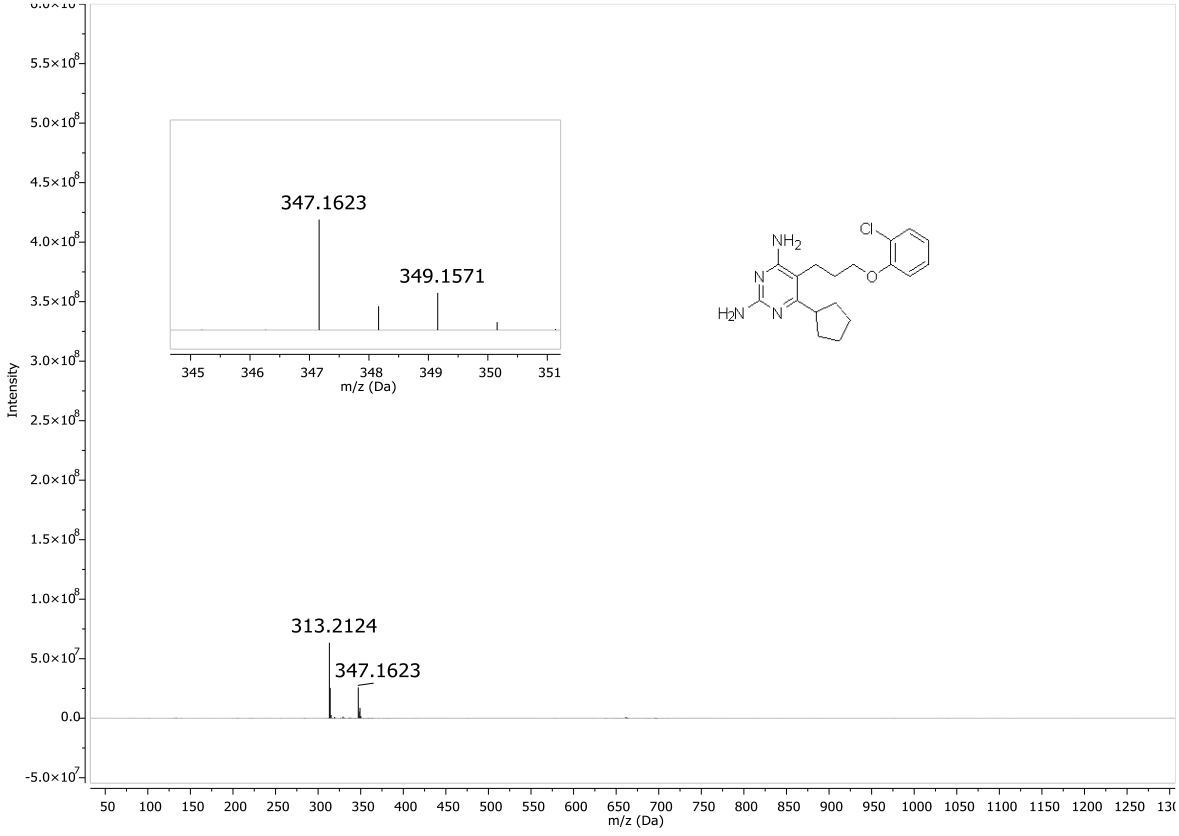
^{13}C NMR (101 MHz, MeOD)



FTIR:


Transmittance [%]
Wavenumbers [1/cm]
3487
3350
3219
2955
2924
2864
1632
1599
1562
1496
1430
1300
1243
1171
1084
1040
935,4
877,6
794,6
751,3
688,5

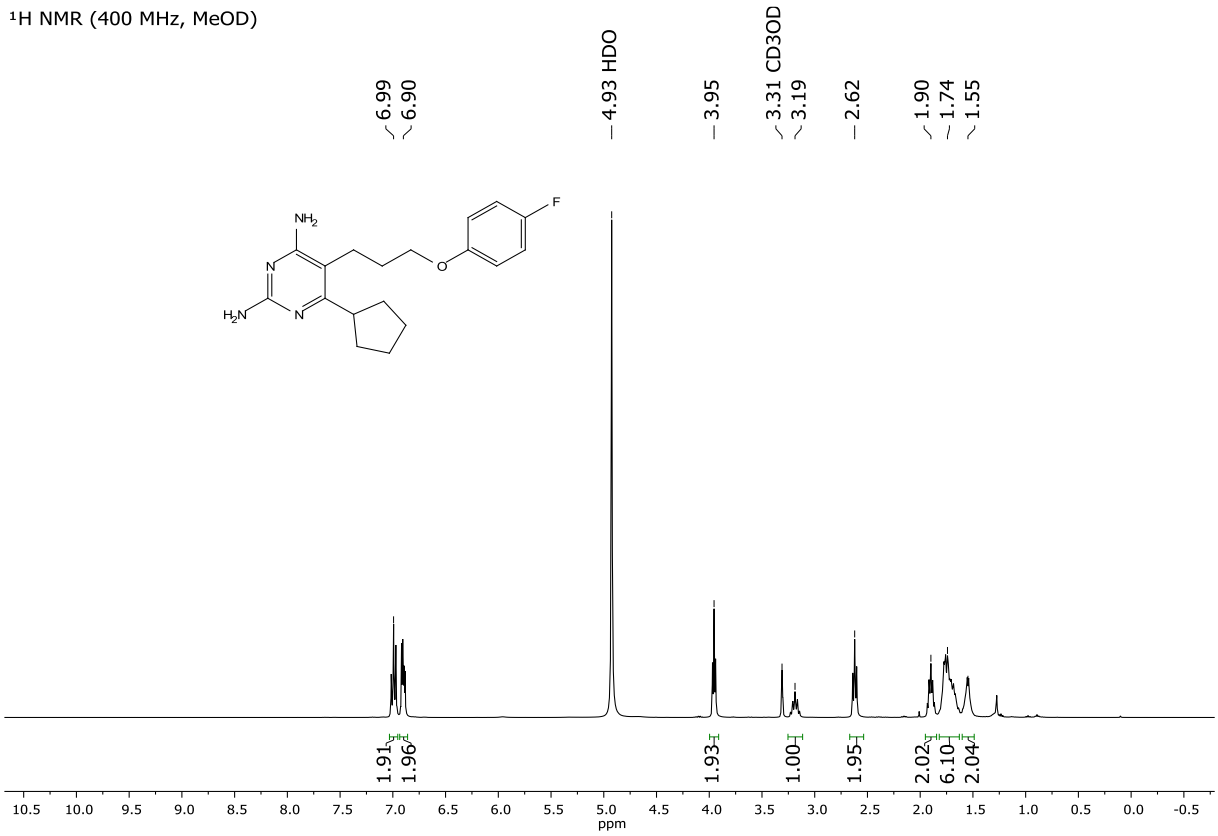
HRMS:


347.1623
349.1571
m/z (Da)
Intensity
313.2124
347.1623
m/z (Da)
NH2
Cl
O
N
H2N

6-Cyclopentyl-5-[3-(4-fluorophenoxy)propyl]pyrimidine-2,4-diamine (**9c**)

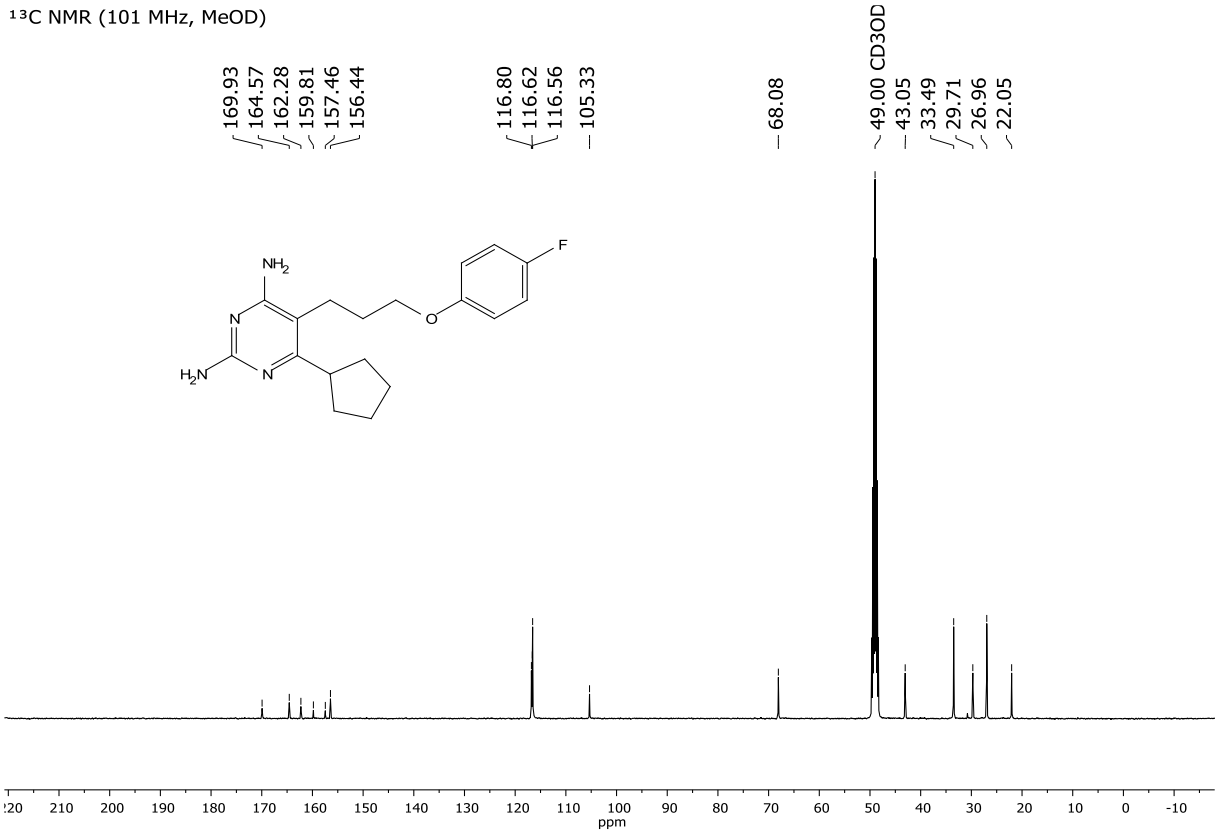
^{1}H NMR:

^{1}H NMR (400 MHz, MeOD)

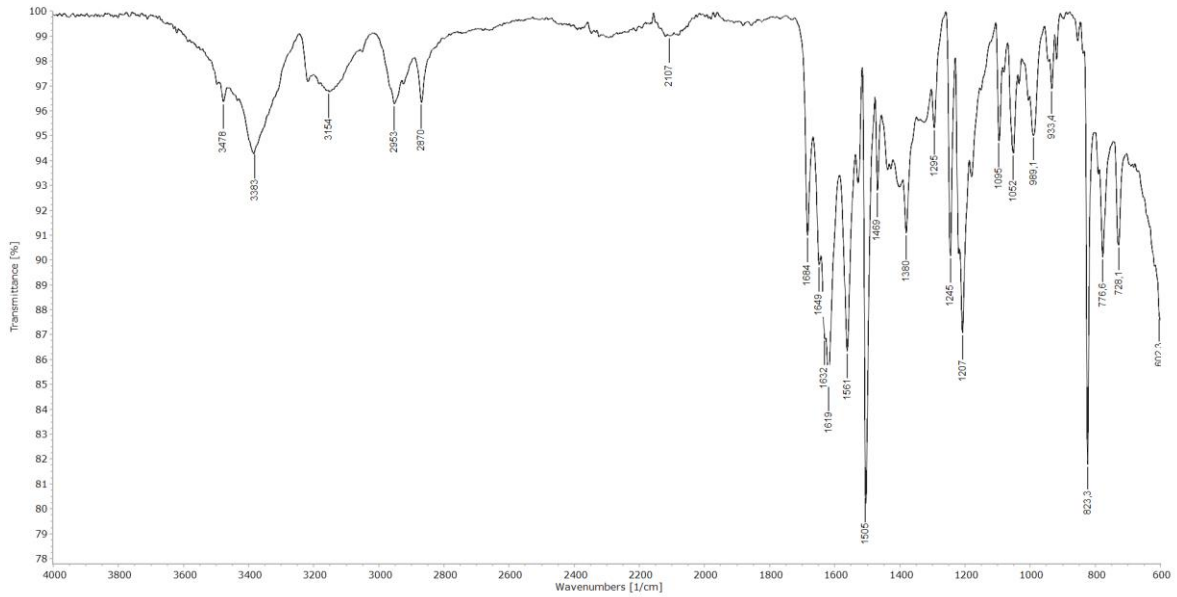


^{13}C NMR:

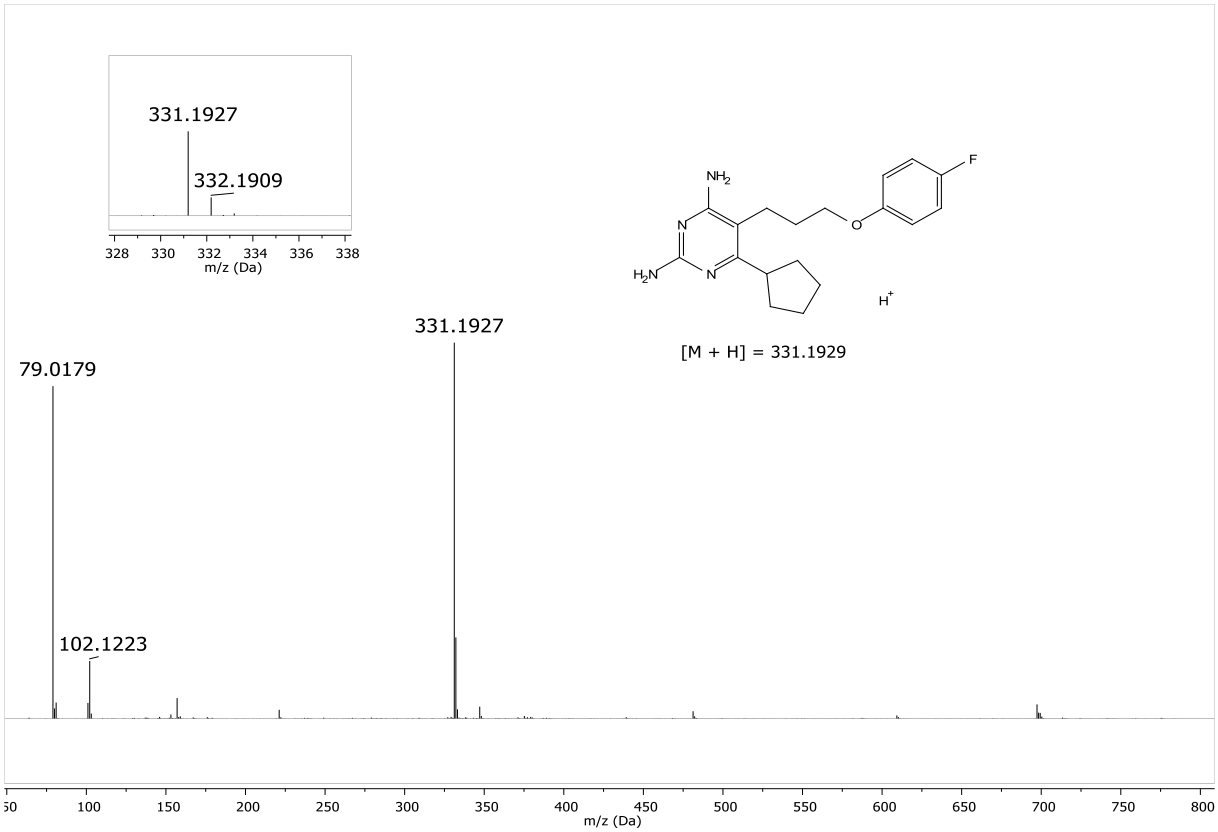
^{13}C NMR (101 MHz, MeOD)



FTIR:

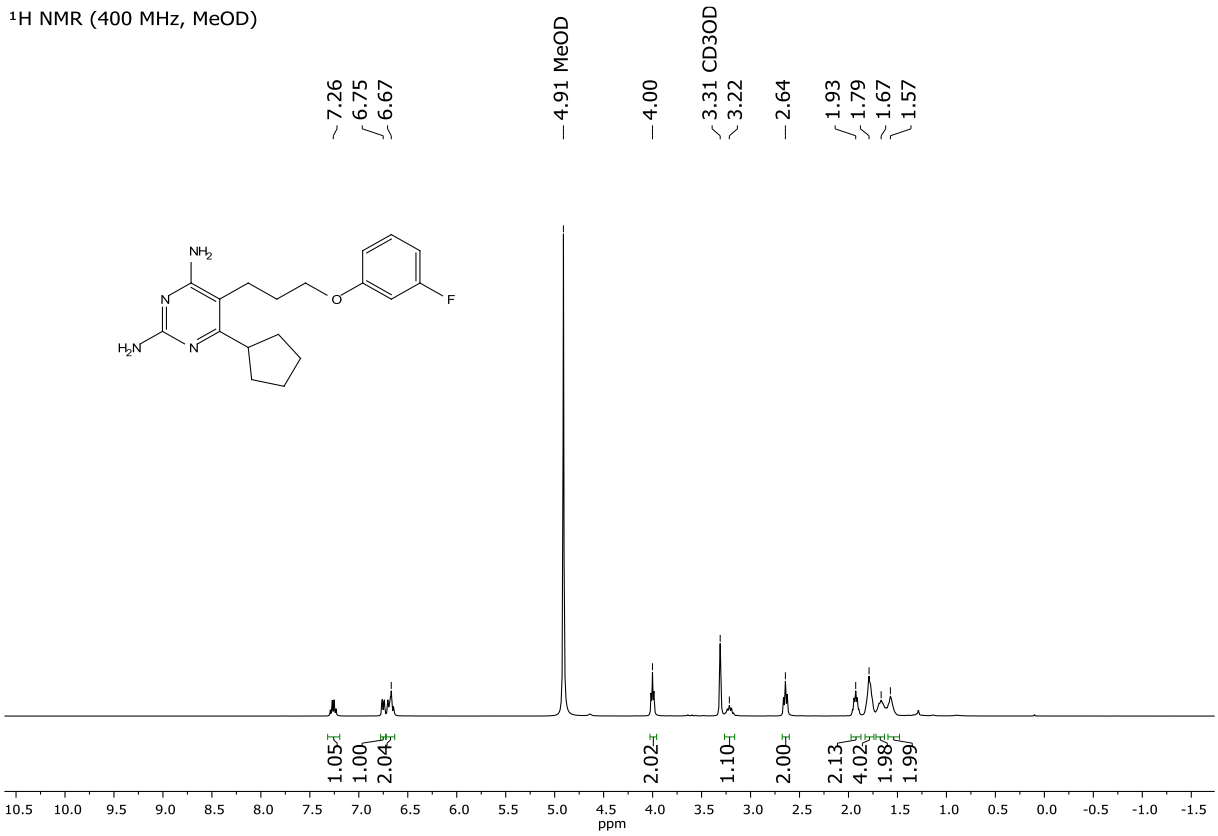

Transmittance [%]
Wavenumbers [1/cm]
3478
3383
3154
2963
2870
2107
1684
1649
1632
1619
1561
1505
1469
1380
1295
1245
1207
1095
1052
989.1
933.4
823.3
776.6
728.1

HRMS:

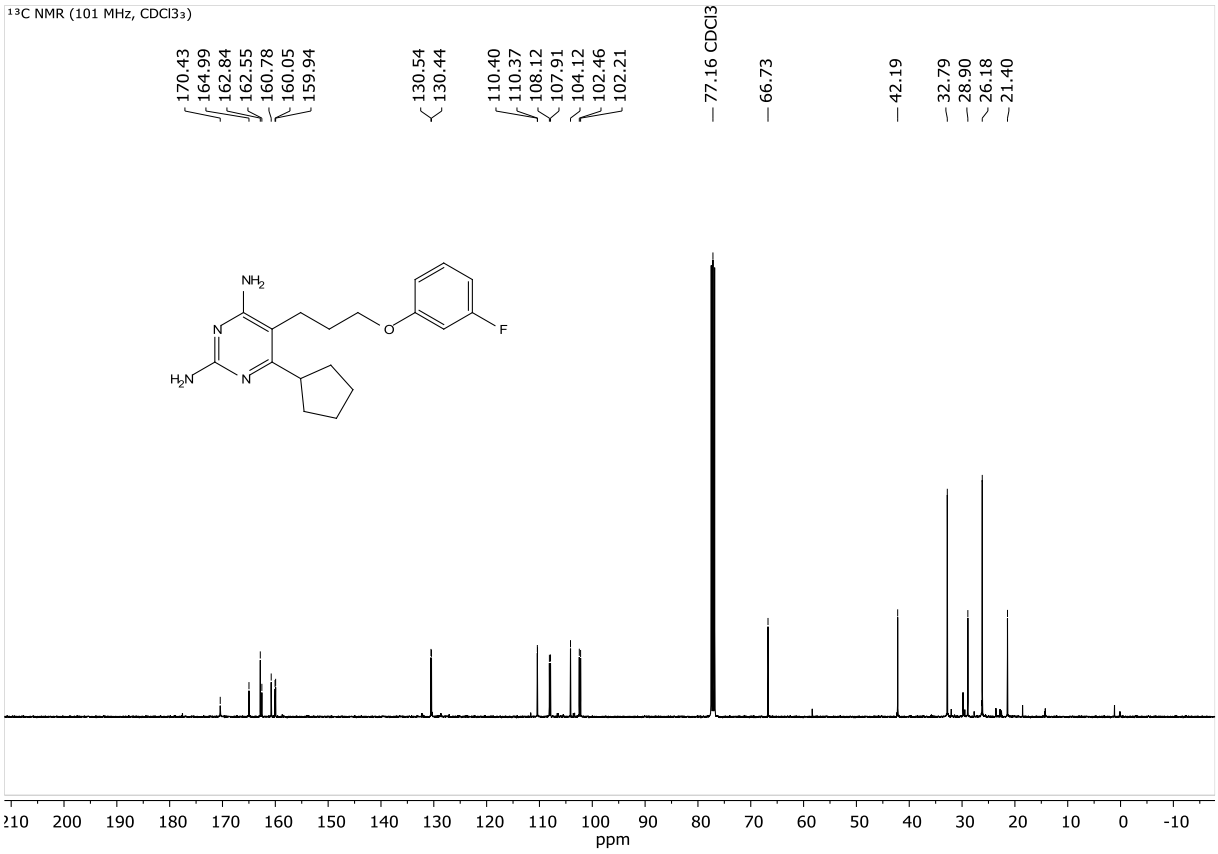

331.1927
332.1909
m/z (Da)
NH2
F
O
N
H2N
H+
[M + H] = 331.1929
331.1927
79.0179
102.1223
m/z (Da)

6-Cyclopentyl-5-[3-(3-fluorophenoxy)propyl]pyrimidine-2,4-diamine (**9d**)

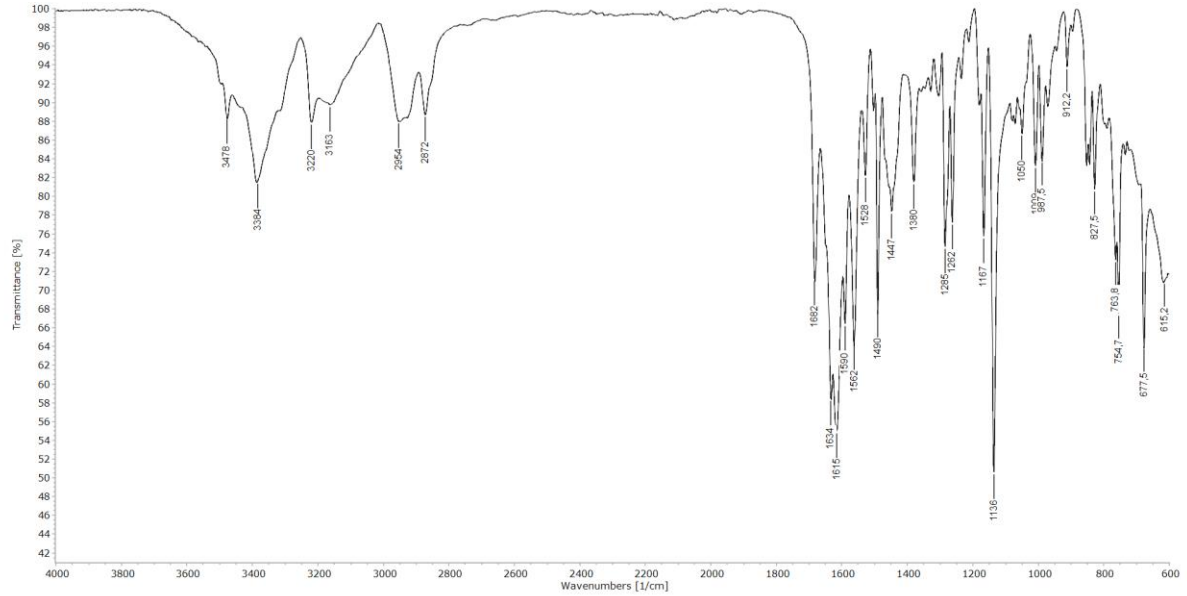
^{1}H NMR:



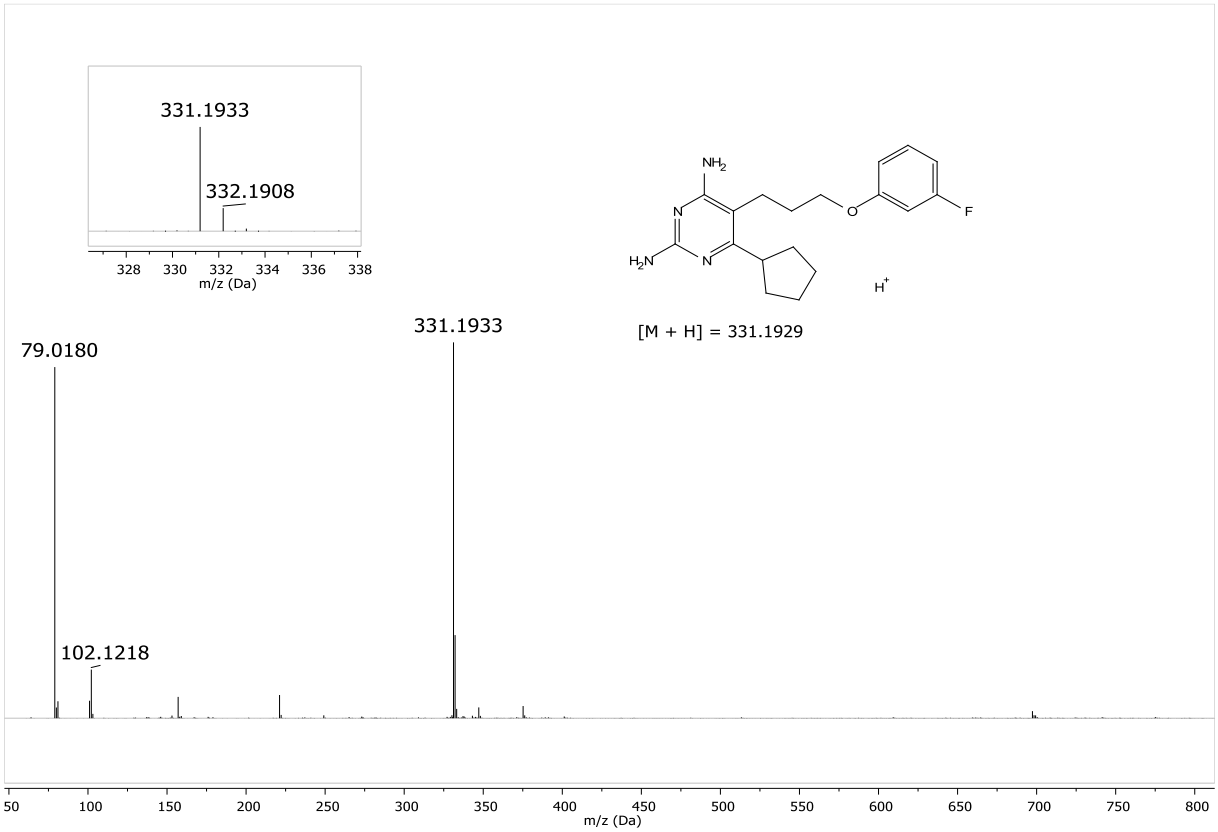
^{13}C NMR:



FTIR:

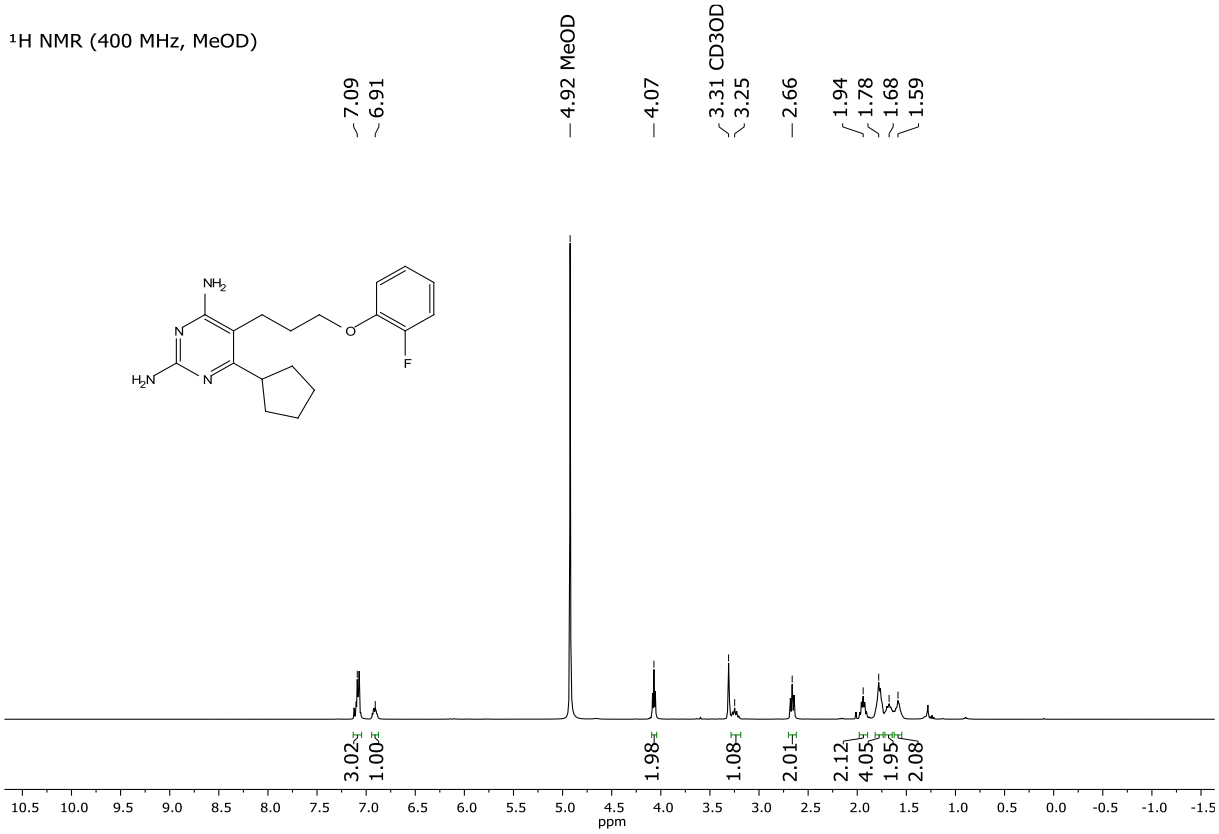

Transmittance [%]
Wavenumbers [1/cm]
3478
3384
3220
3163
2954
2872
1682
1634
1615
1590
1562
1528
1490
1447
1380
1285
1262
1167
1136
1050
1000
997.5
912.2
827.5
763.8
754.7
677.5
615.2

HRMS:

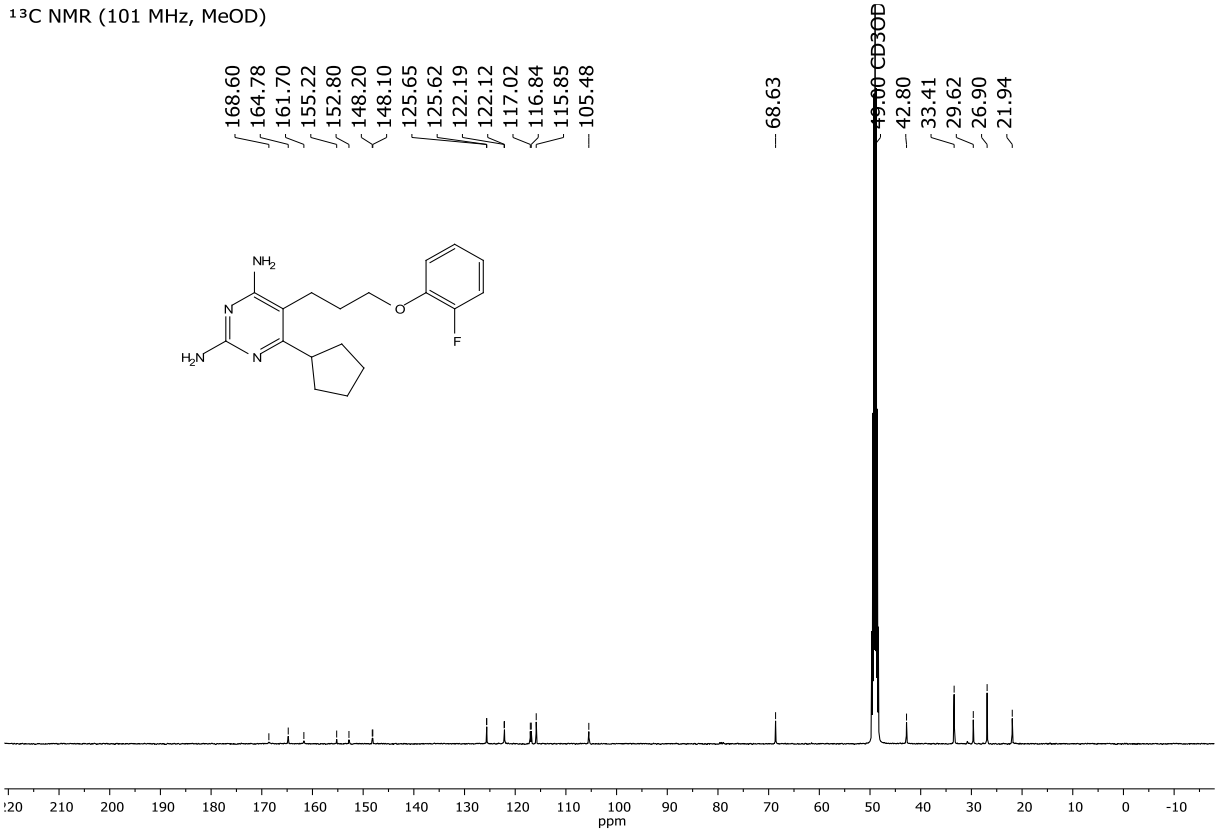

331.1933
332.1908
328 330 332 334 336 338
m/z (Da)
NH2
N
O
F
H2N
N
H+
[M + H] = 331.1929
79.0180
331.1933
102.1218
50 100 150 200 250 300 350 400 450 500 550 600 650 700 750 800
m/z (Da)

6-Cyclopentyl-5-[3-(2-fluorophenoxy)propyl]pyrimidine-2,4-diamine (**9e**)

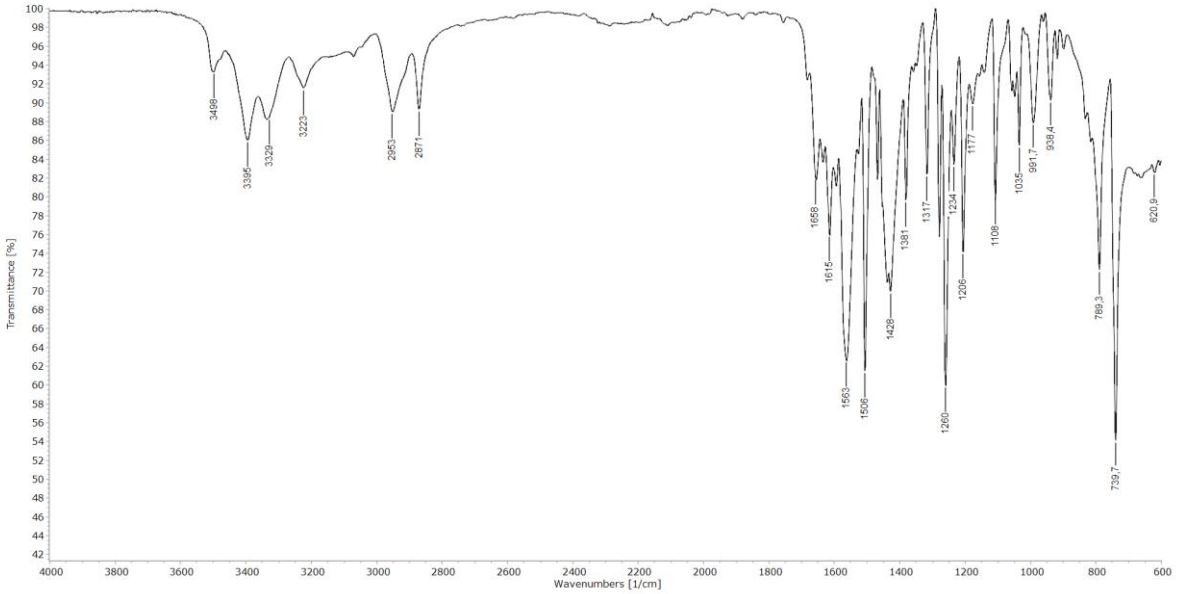
^{1}H NMR:



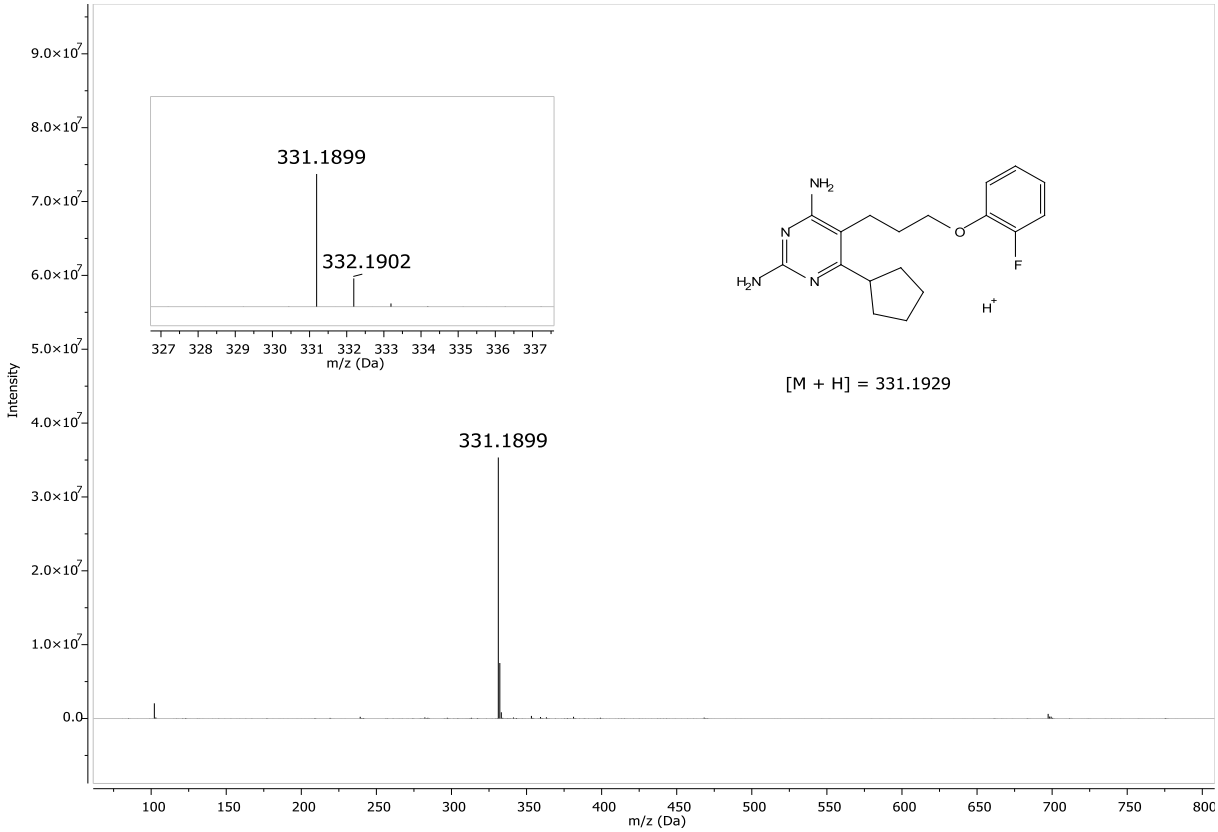
^{13}C NMR:



FTIR:


Transmittance [%]
Wavenumbers [1/cm]
3498
3395
3329
3223
2953
2871
1658
1615
1563
1506
1428
1381
1317
1260
1234
1206
1177
1108
1035
991.7
938.4
789.3
739.7
620.9

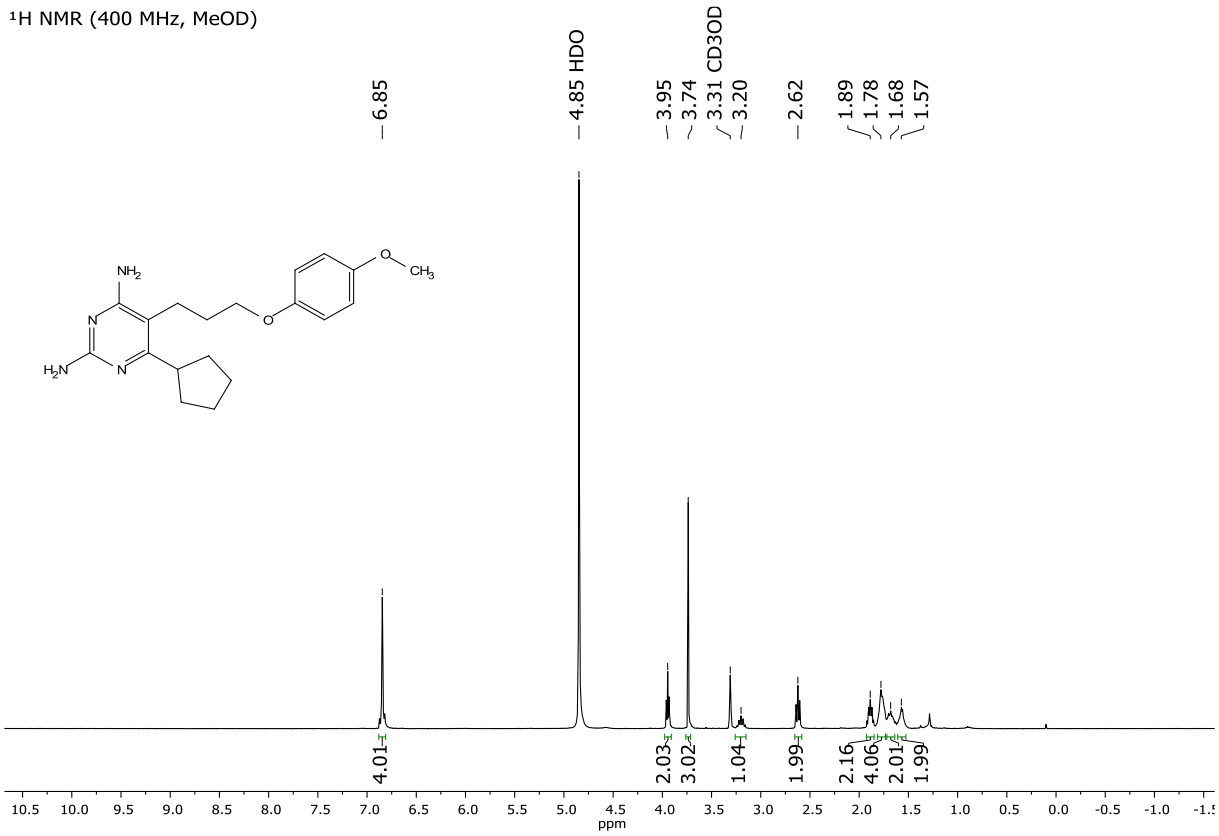
HRMS:


331.1899
332.1902
m/z (Da)
NH2
N
O
F
H2N
N
H+
[M + H] = 331.1929
Intensity
331.1899
m/z (Da)

6-Cyclopentyl-5-[3-(4-methoxyphenoxy)propyl]pyrimidine-2,4-diamine (**9f**)

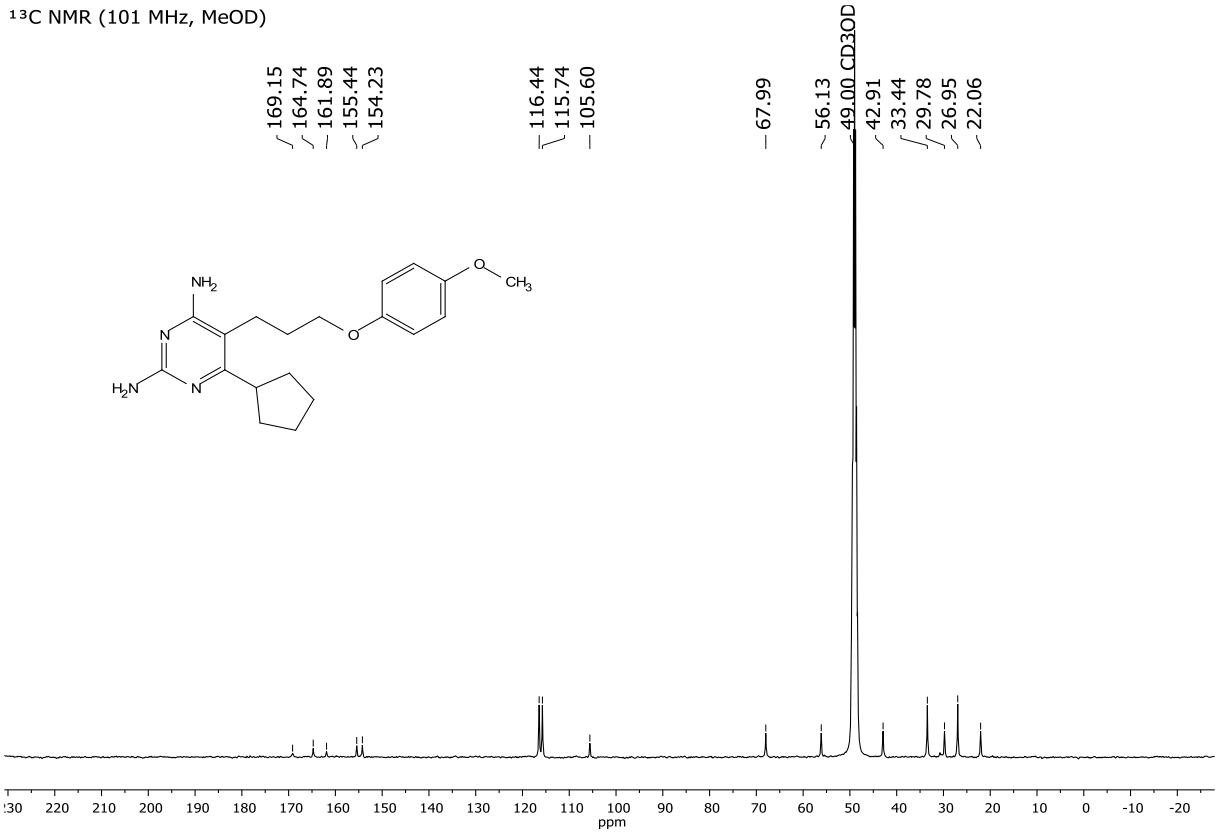
^{1}H NMR:

^{1}H NMR (400 MHz, MeOD)

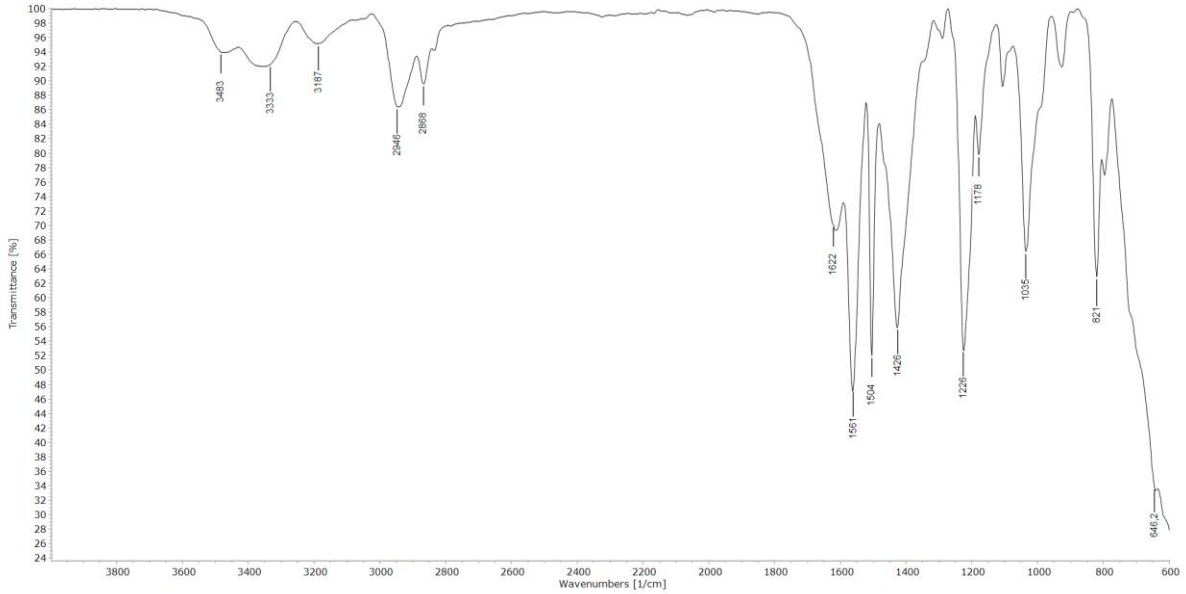


^{13}C NMR:

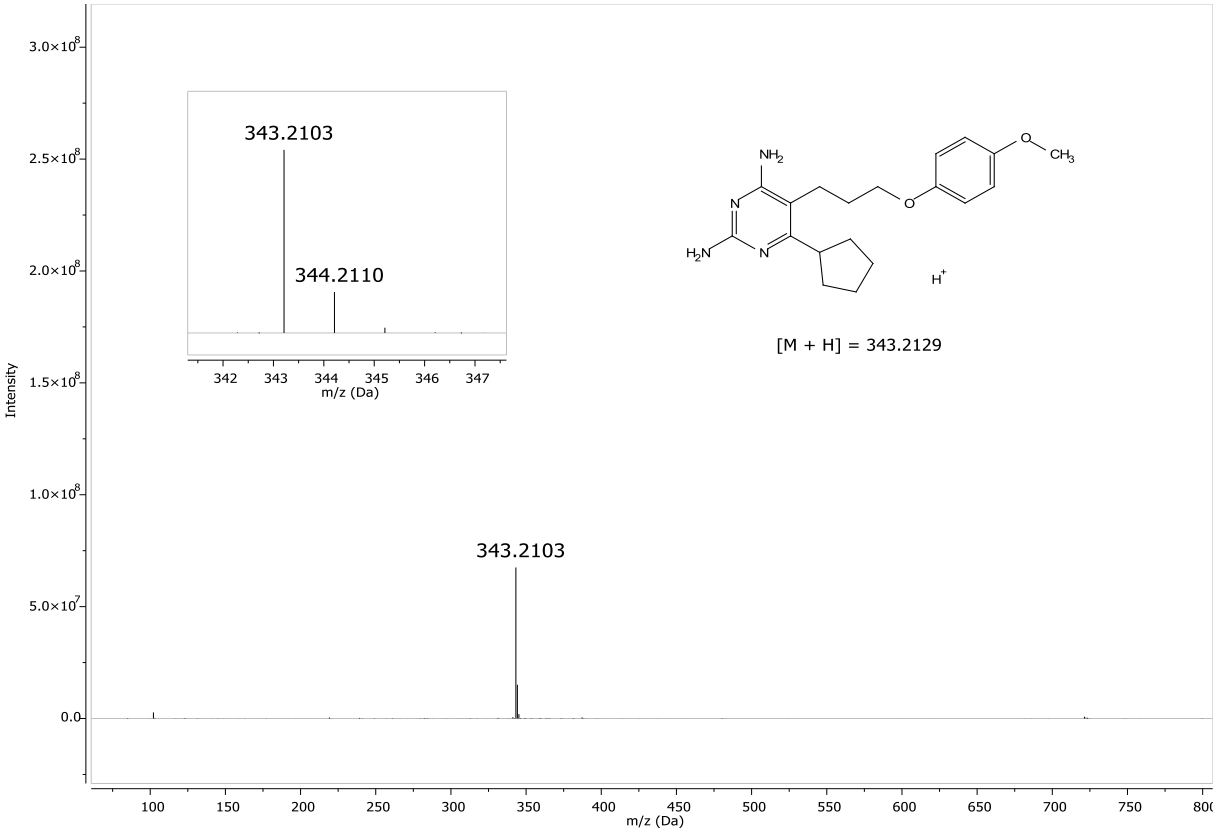
^{13}C NMR (101 MHz, MeOD)



FTIR:


Transmittance [%]
Wavenumbers [1/cm]
3483
3333
3187
2946
2868
1622
1561
1504
1426
1226
1178
1035
821
646.2

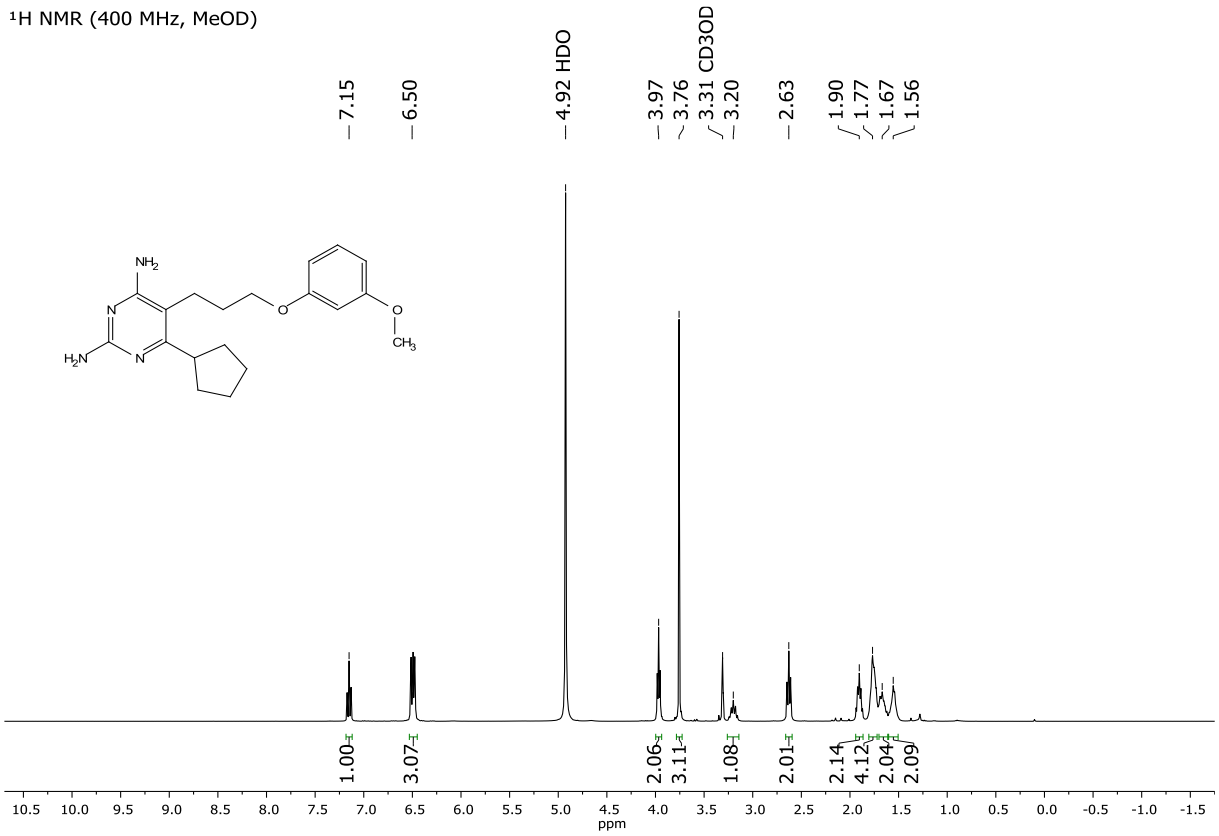
HRMS:


343.2103
344.2110
m/z (Da)
[M + H] = 343.2129
Intensity
343.2103
m/z (Da)

6-Cyclopentyl-5-[3-(3-methoxyphenoxy)propyl]pyrimidine-2,4-diamine (**9g**)

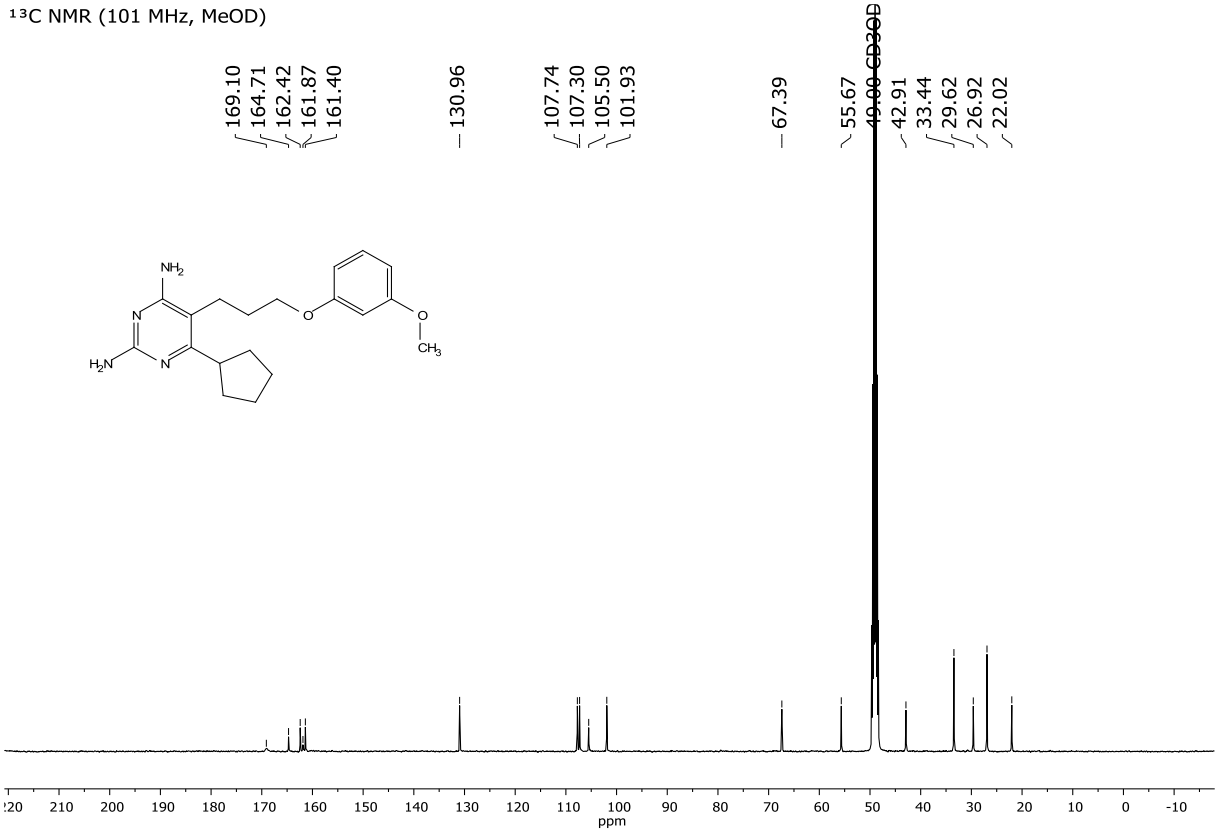
^{1}H NMR:

^{1}H NMR (400 MHz, MeOD)

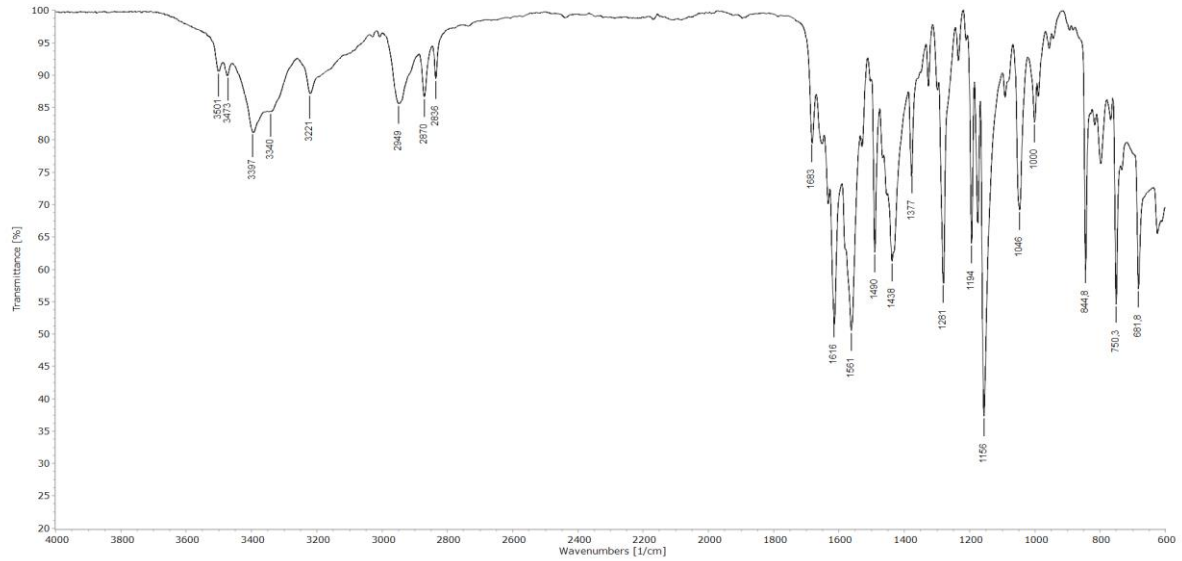


^{13}C NMR:

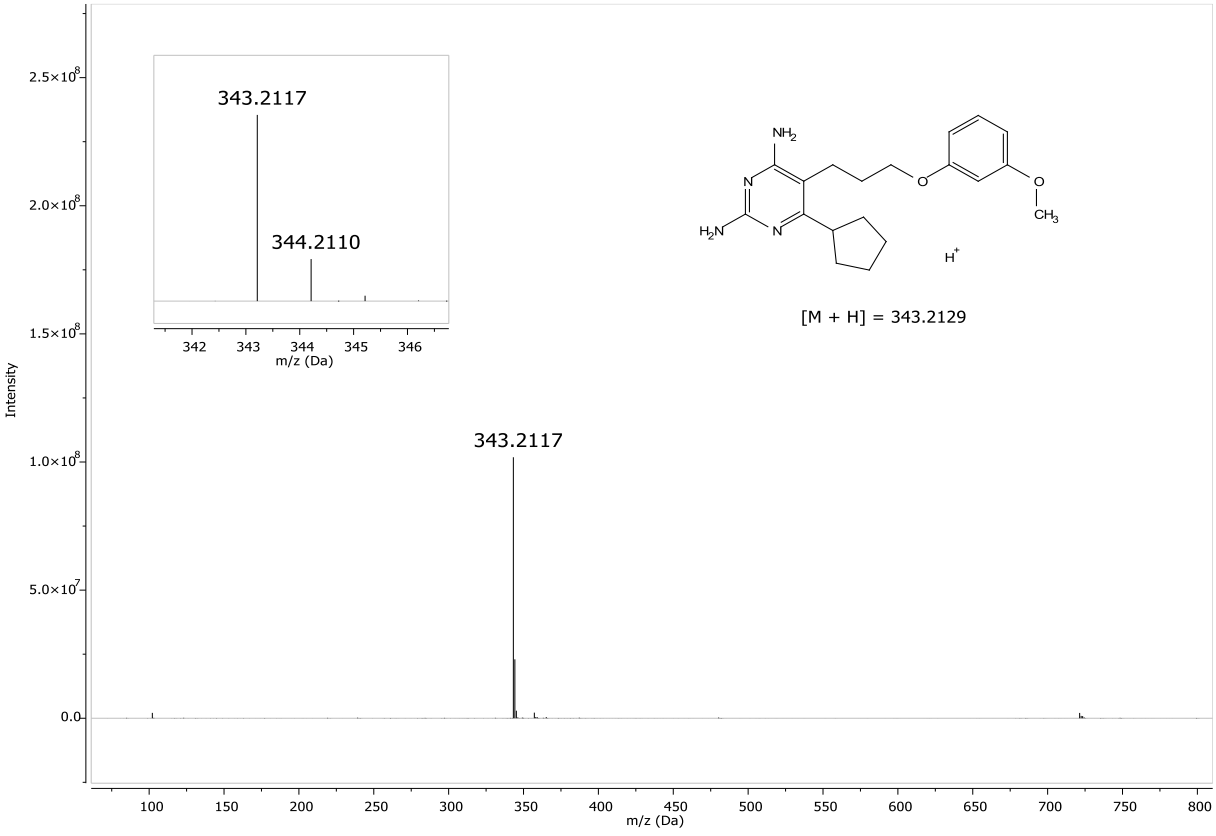
^{13}C NMR (101 MHz, MeOD)



FTIR:


Transmittance [%]
Wavenumbers [1/cm]
3501
3473
3397
3340
3221
2949
2870
2836
1683
1616
1561
1490
1438
1377
1281
1194
1156
1046
1000
844.8
750.3
681.8

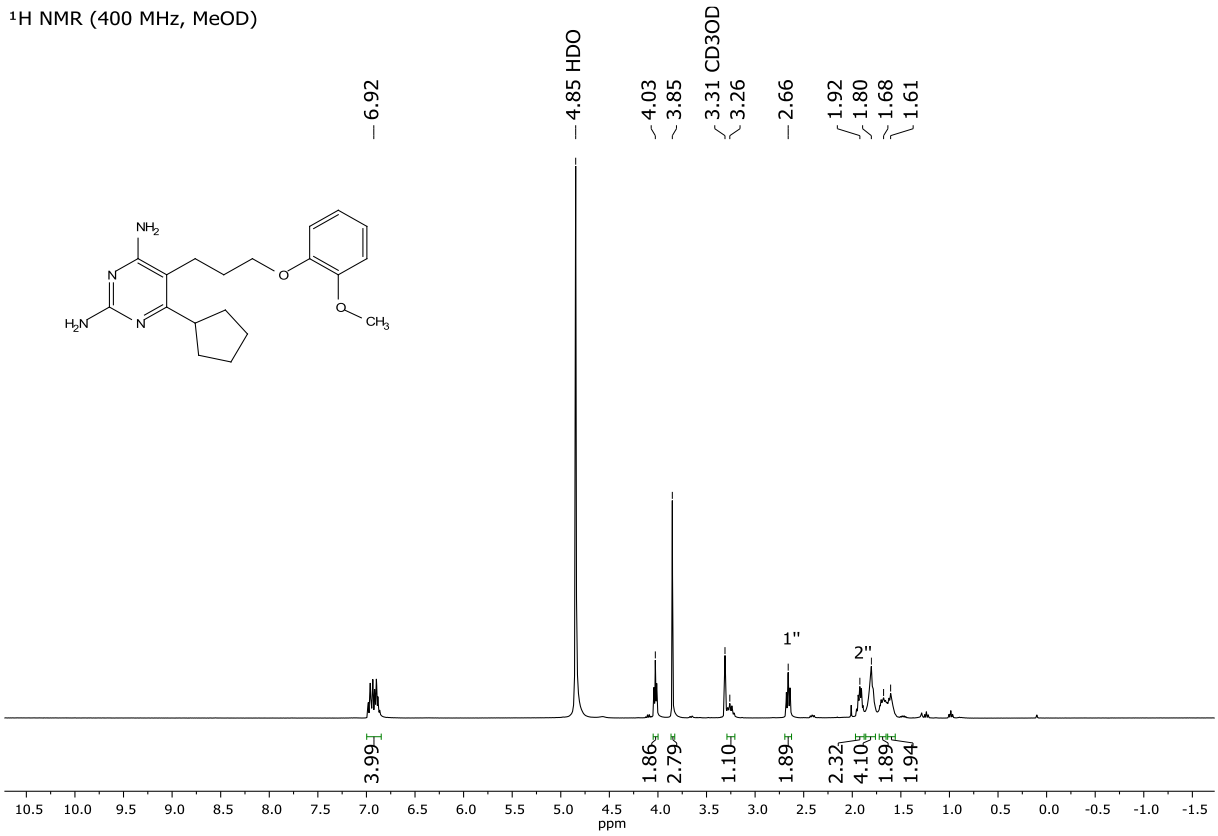
HRMS:


343.2117
344.2110
m/z (Da)
[M + H] = 343.2129
Intensity
343.2117
m/z (Da)

6-Cyclopentyl-5-[3-(2-methoxyphenoxy)propyl]pyrimidine-2,4-diamine (**9h**)

^{1}H NMR:

^{1}H NMR (400 MHz, MeOD)

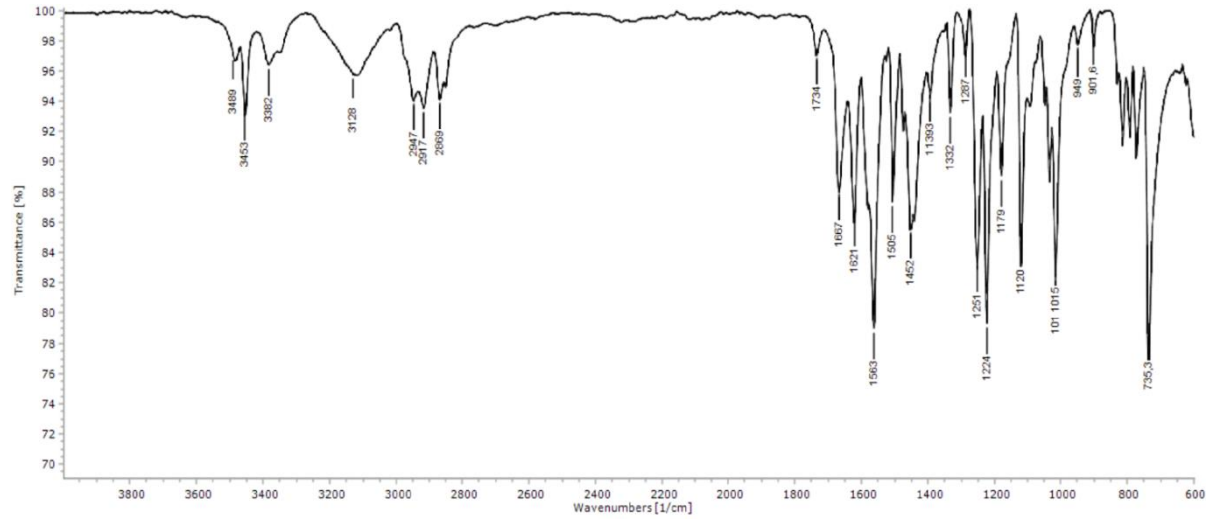


^{13}C NMR:

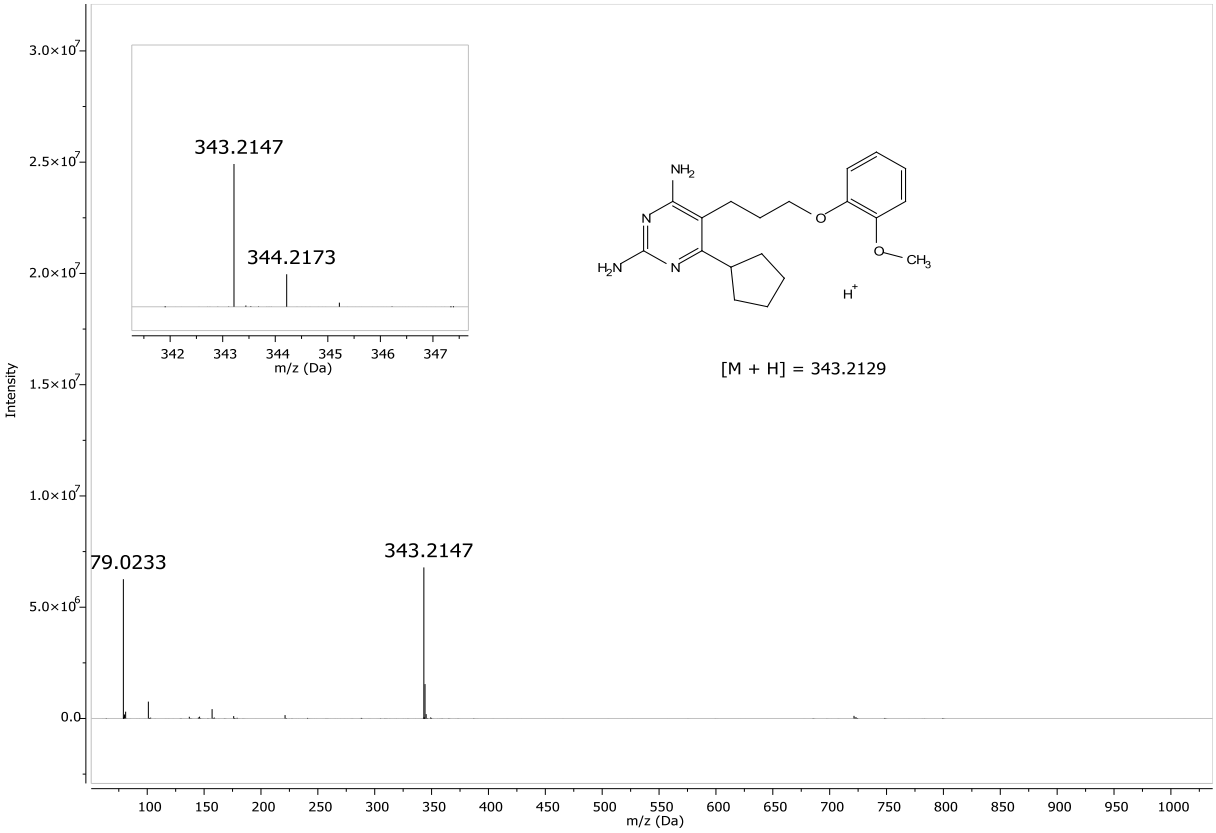
^{13}C NMR (101 MHz, MeOD)



FTIR:


Transmittance [%]
Wavenumbers [1/cm]
3489
3453
3382
3128
2947
2917
2669
1734
1667
1621
1563
1505
1452
1393
1332
1287
1251
1224
1179
1120
1015
949
901,6
735,3

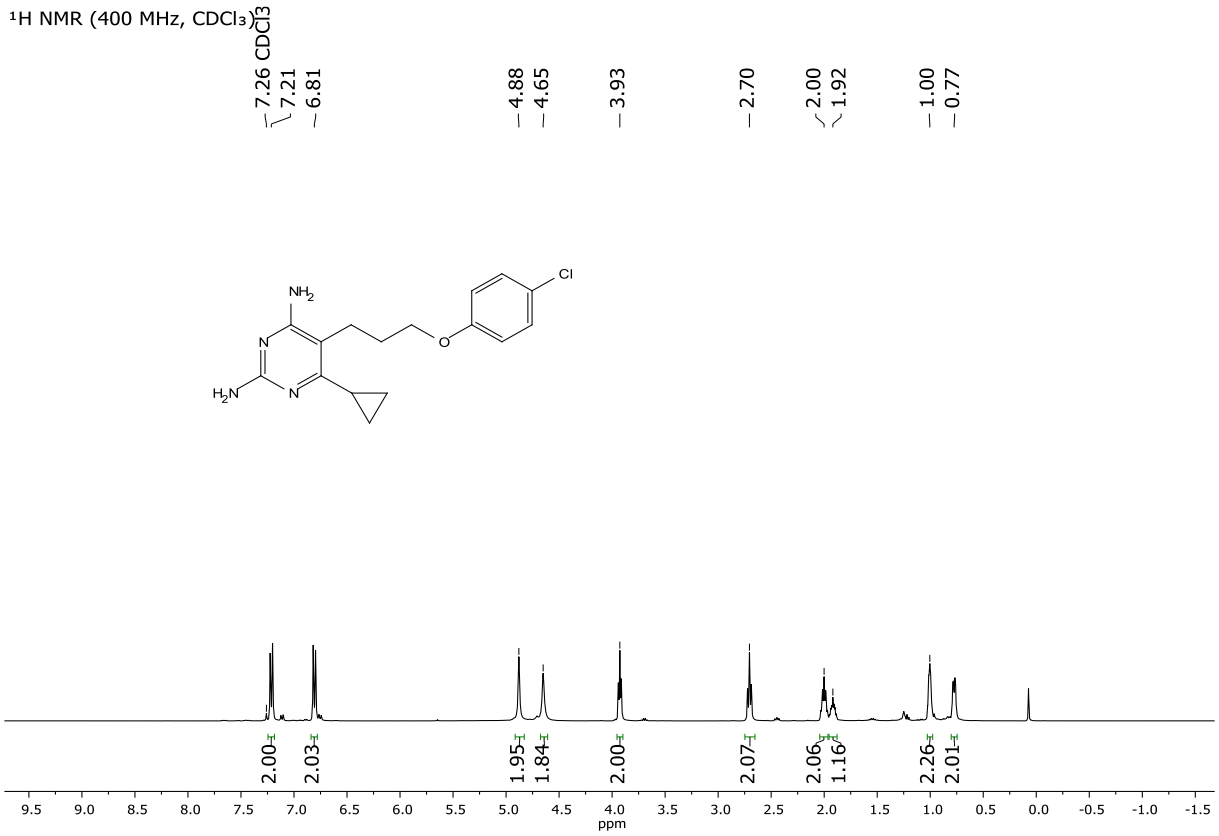
HRMS:


343.2147
344.2173
m/z (Da)
NH2
H2N
N
O
O
CH3
H+
[M + H] = 343.2129
Intensity
79.0233
343.2147
m/z (Da)

6-Cyclopropyl-5-[3-(4-chlorophenoxy)propyl]pyrimidine-2,4-diamine (**9i**)

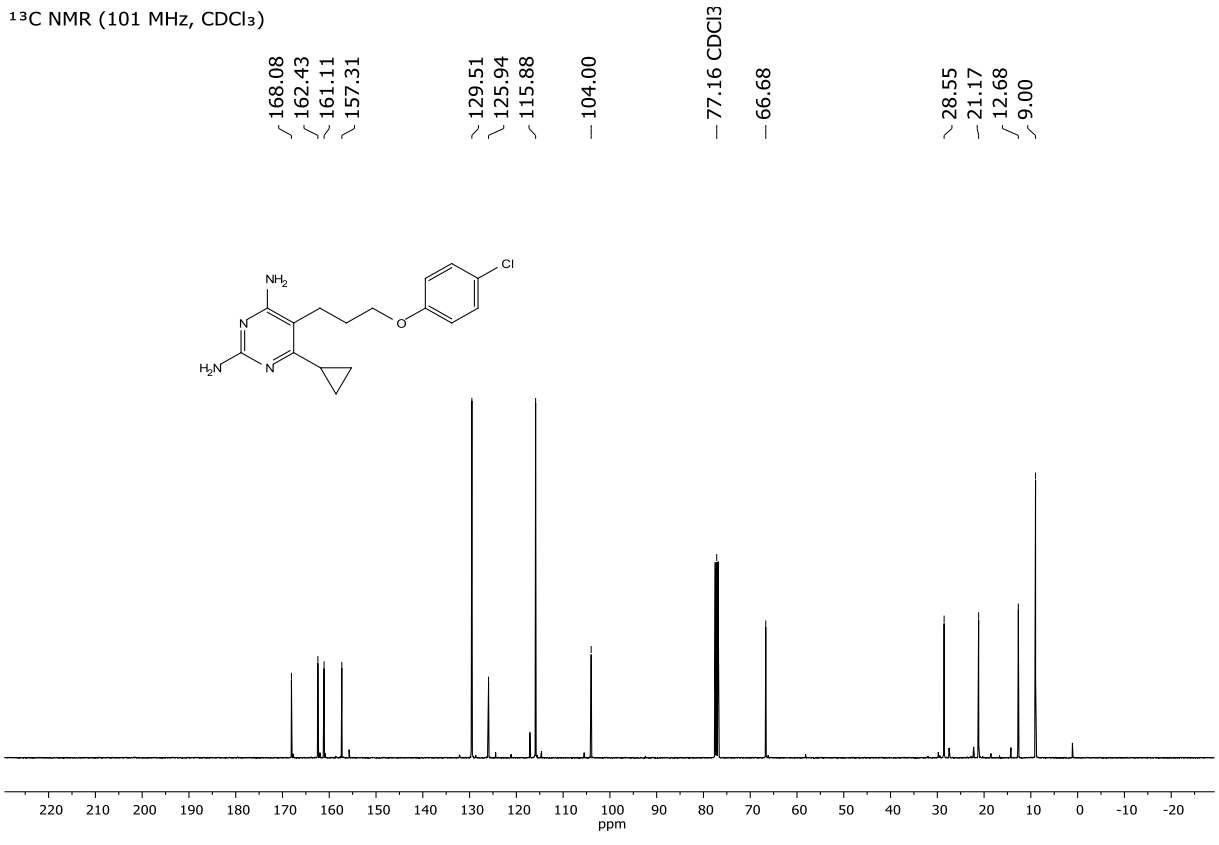
^{1}H NMR:

^{1}H NMR (400 MHz, $CDCl_3$)

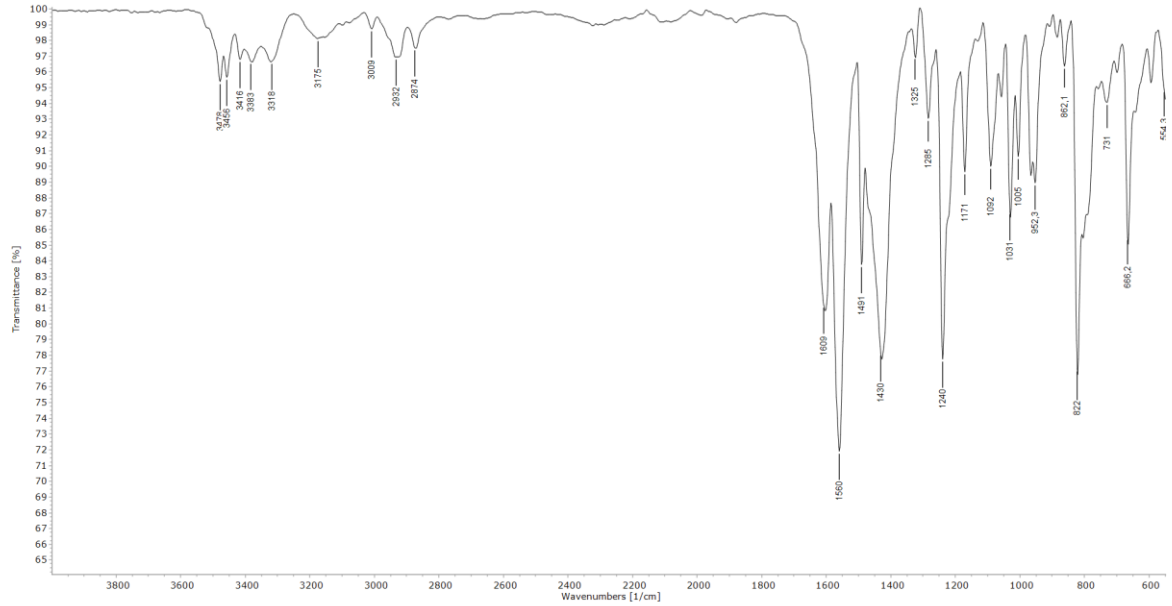


^{13}C NMR:

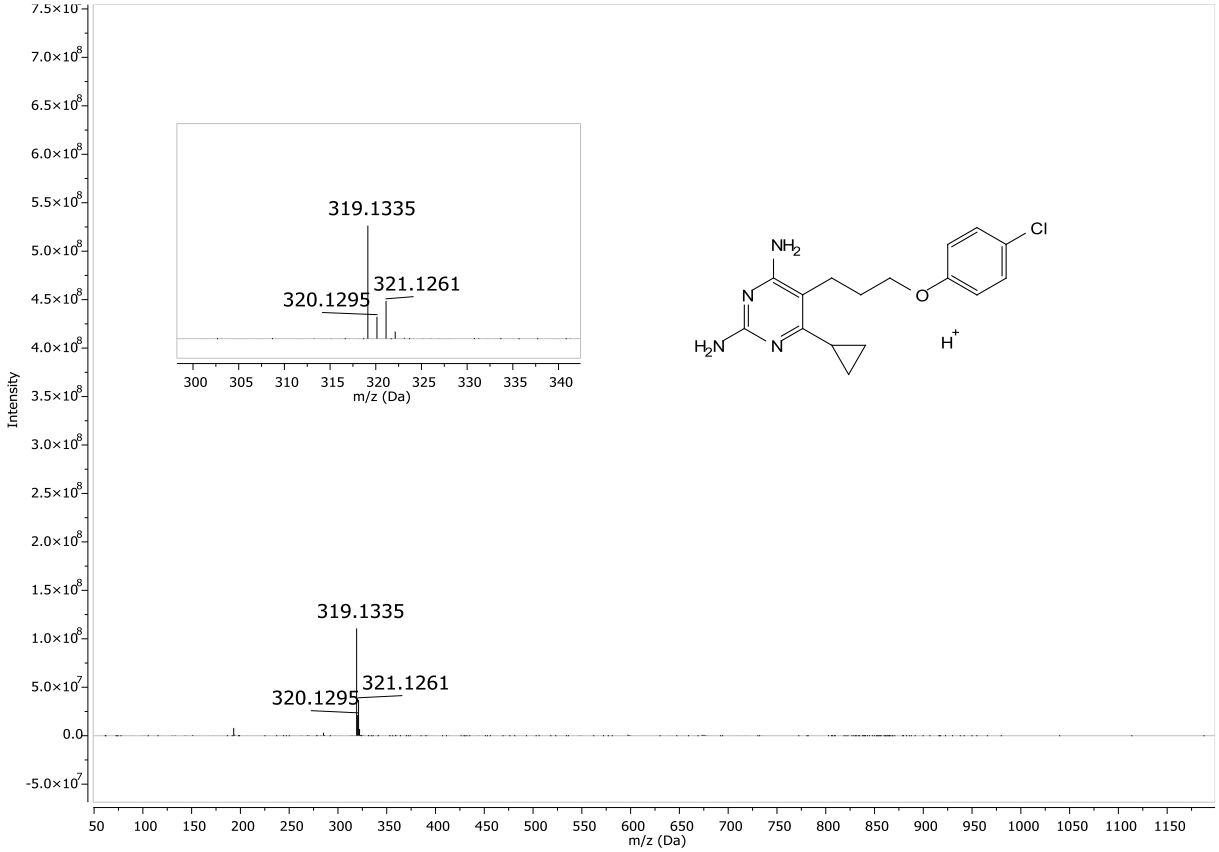
^{13}C NMR (101 MHz, $CDCl_3$)



FTIR:


Transmittance [%]
Wavenumbers [1/cm]
3478
3456
3416
3383
3318
3178
3009
2932
2874
1609
1560
1491
1430
1325
1285
1240
1171
1092
1031
1005
952,3
862,1
822
731
666,2

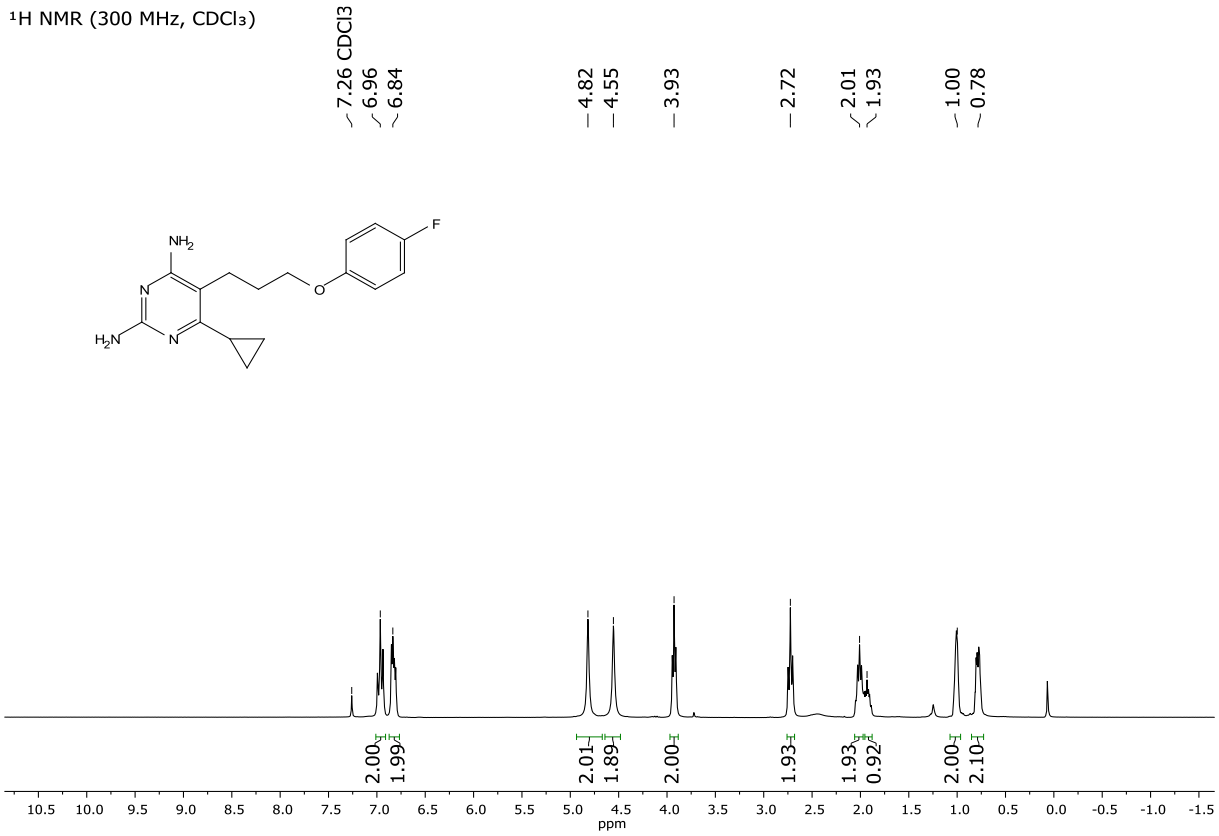
HRMS:


Intensity
m/z (Da)
319.1335
320.1295
321.1261
Cl
NH2
N
O
H2N
H+

6-Cyclopropyl-5-[3-(4-fluorophenoxy)propyl]pyrimidine-2,4-diamine (**9j**)

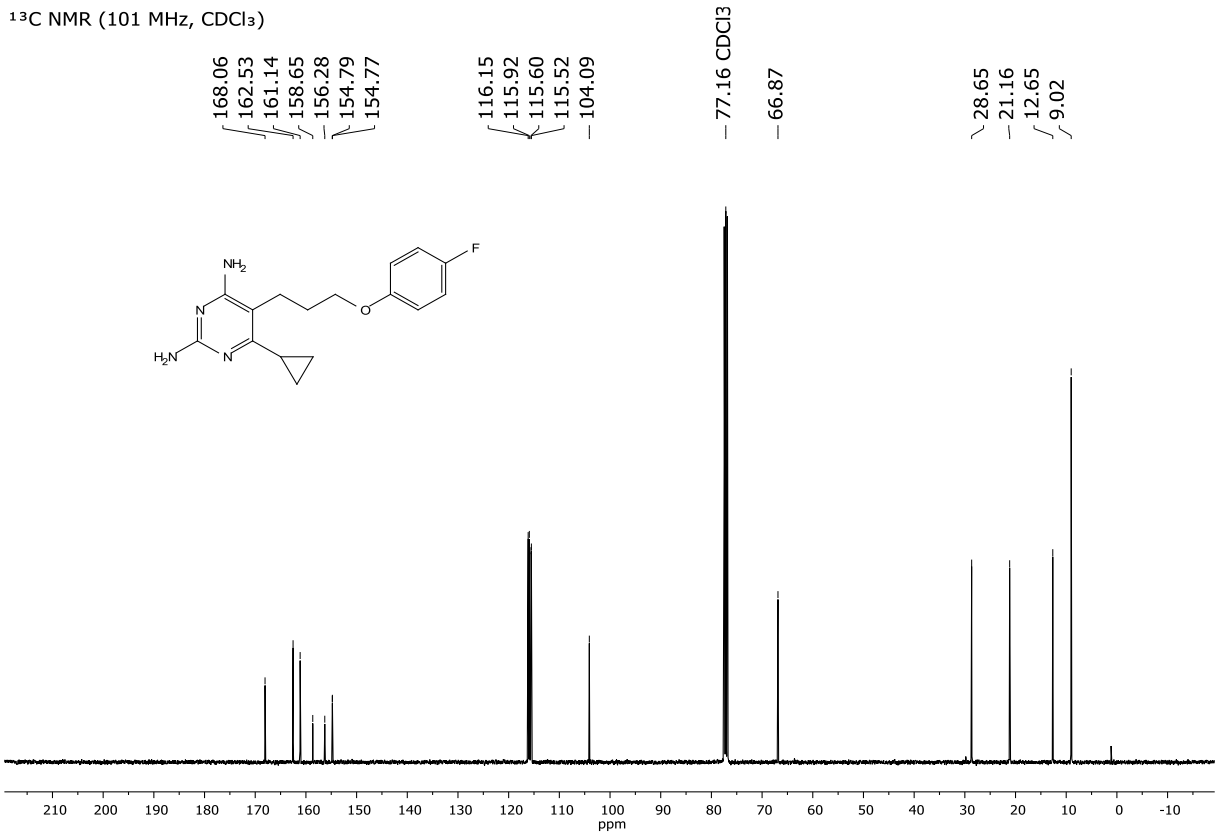
^{1}H NMR:

^{1}H NMR (300 MHz, $CDCl_3$)

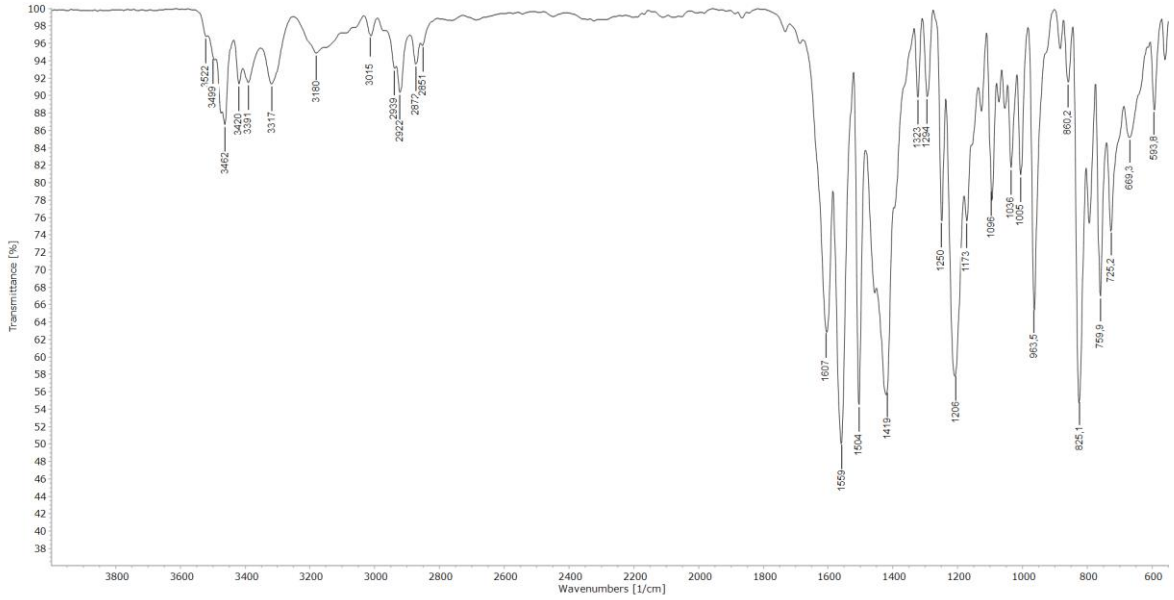


^{13}C NMR:

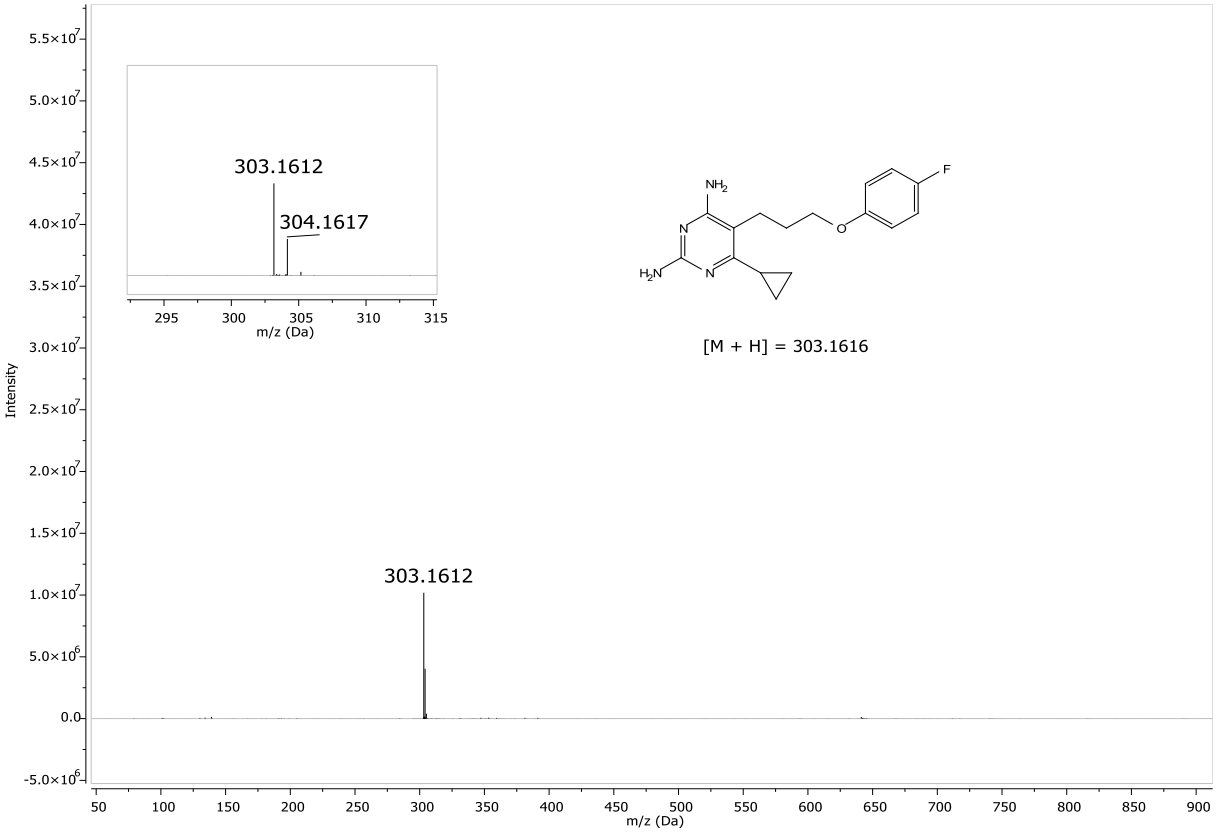
^{13}C NMR (101 MHz, $CDCl_3$)



FTIR:

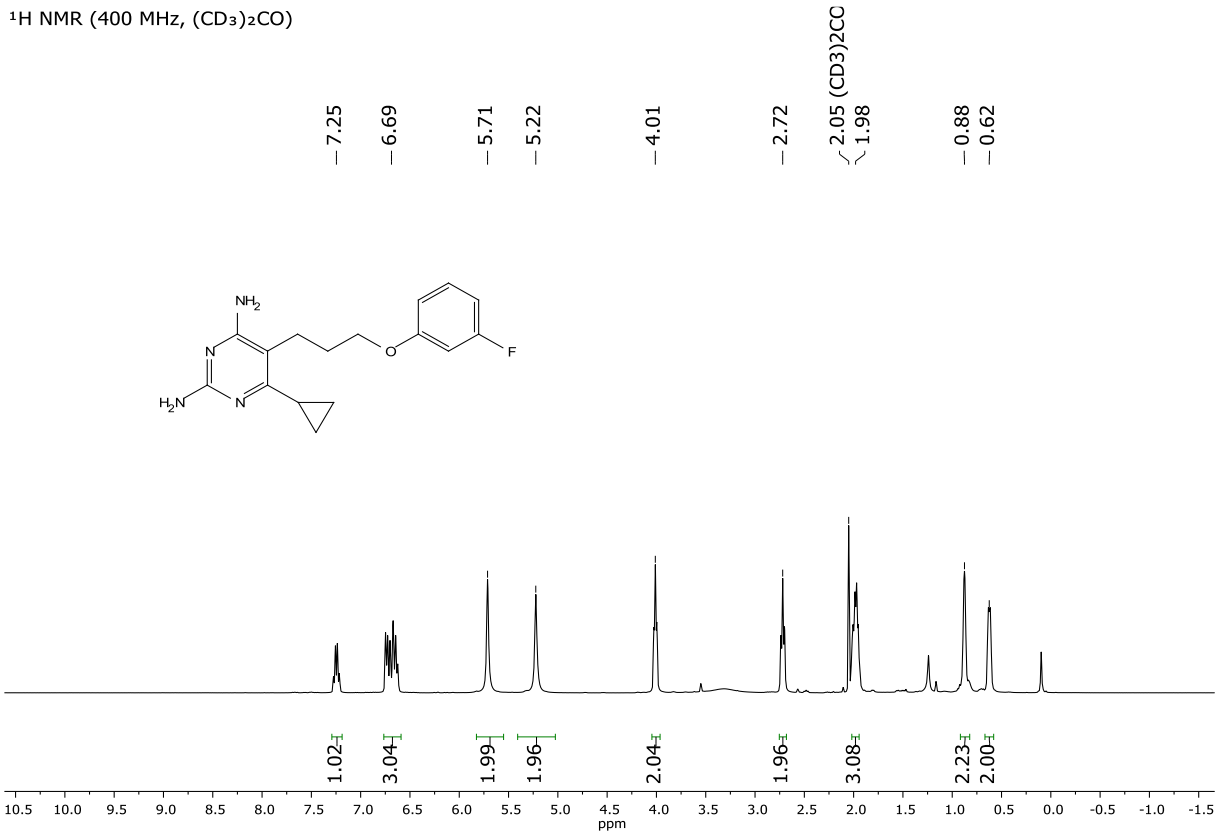

Transmittance [%]
Wavenumbers [1/cm]
3462
3499
3420
3381
3317
3180
3015
2922
2872
2851
1607
1559
1504
1419
1323
1294
1250
1206
1173
1096
1036
1005
963.6
860.2
825.1
759.9
725.2
669.3
593.8

HRMS:

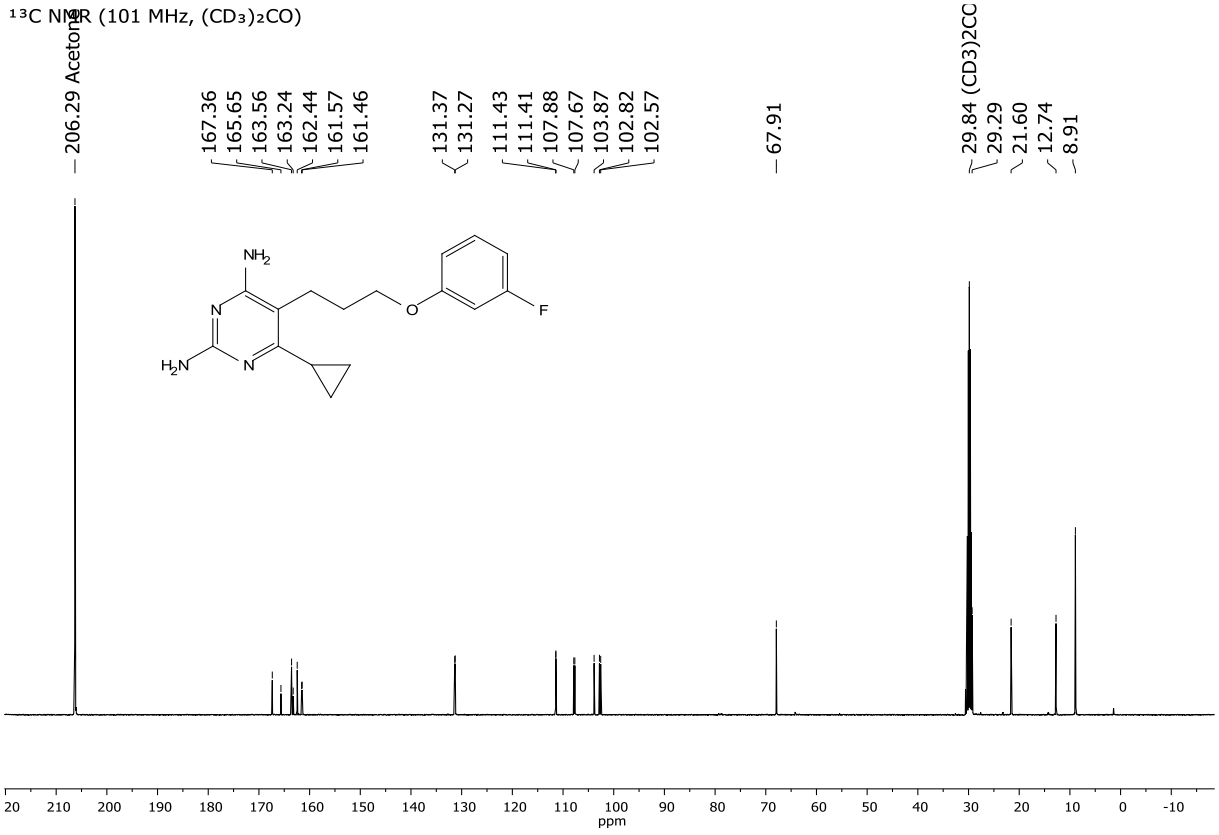

303.1612
304.1617
m/z (Da)
[M + H] = 303.1616
Intensity
303.1612
m/z (Da)

6-Cyclopropyl-5-[3-(3-fluorophenoxy)propyl]pyrimidine-2,4-diamine (**9k**)

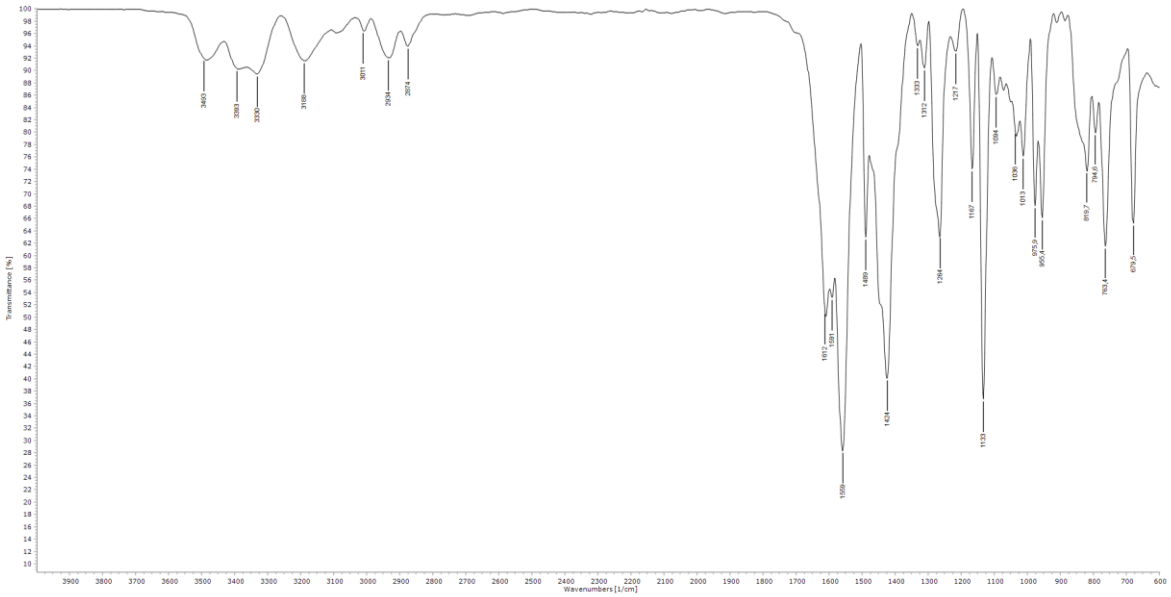
^{1}H NMR:



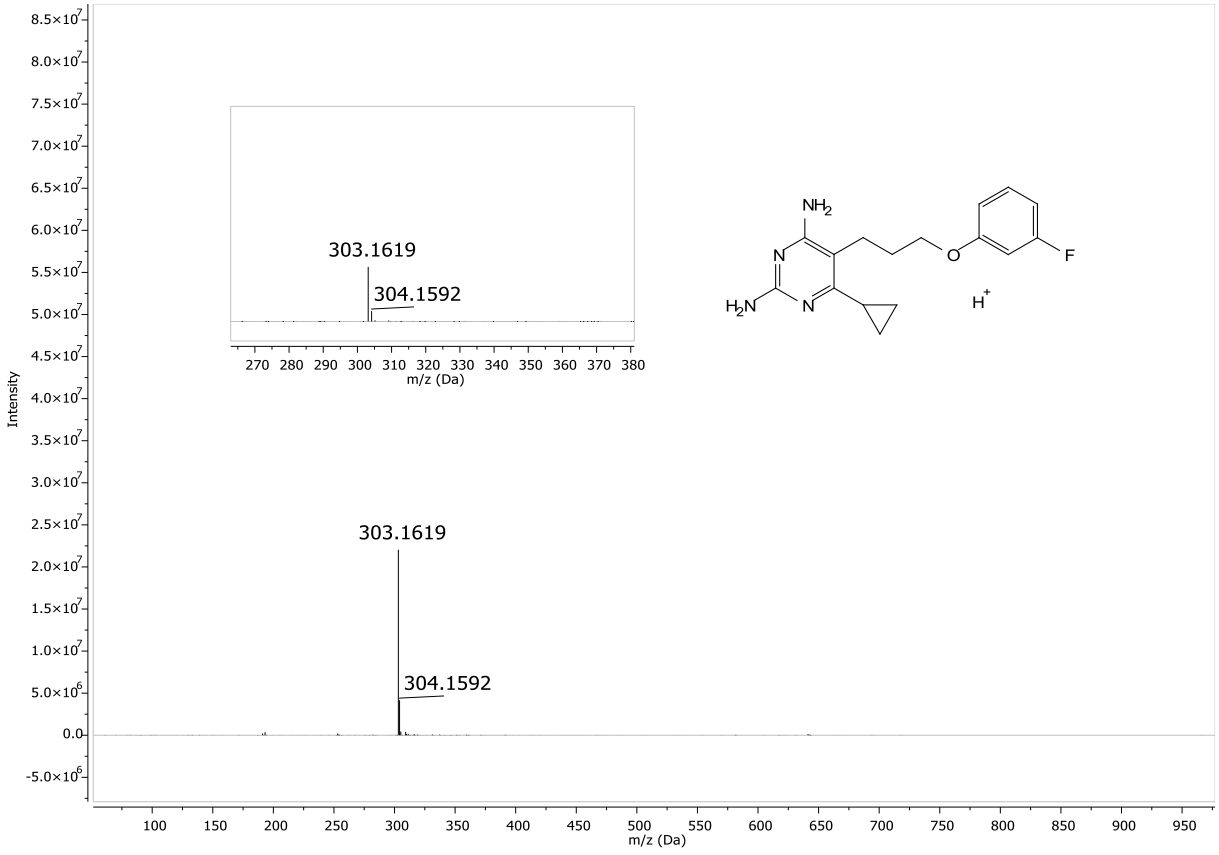
^{13}C NMR:



FTIR:


Transmittance [%]
Wavenumbers [1/cm]

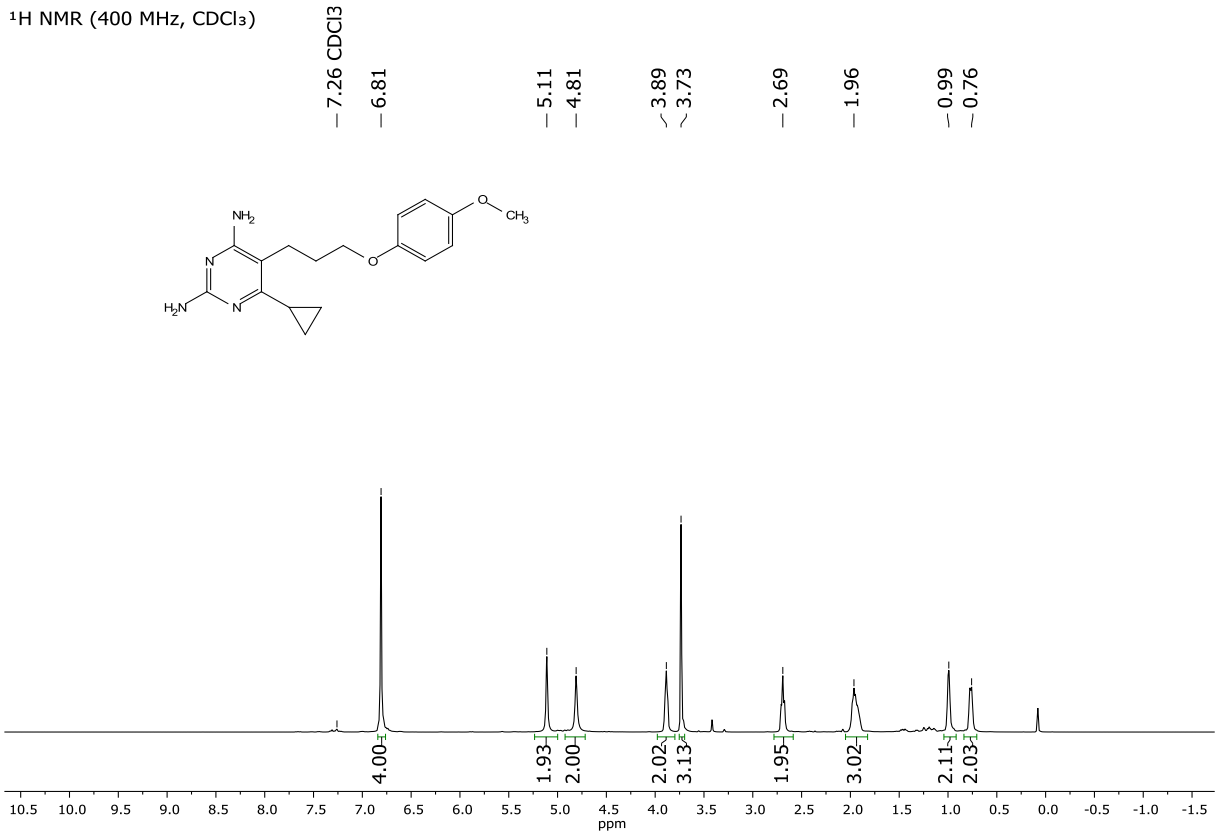
HRMS:


303.1619
304.1592
m/z (Da)
NH2
N
O
F
H2N
N
H+
Intensity
303.1619
304.1592
m/z (Da)

6-Cyclopropyl-5-[3-(4-methoxyphenoxy)propyl]pyrimidine-2,4-diamine (**9l**)

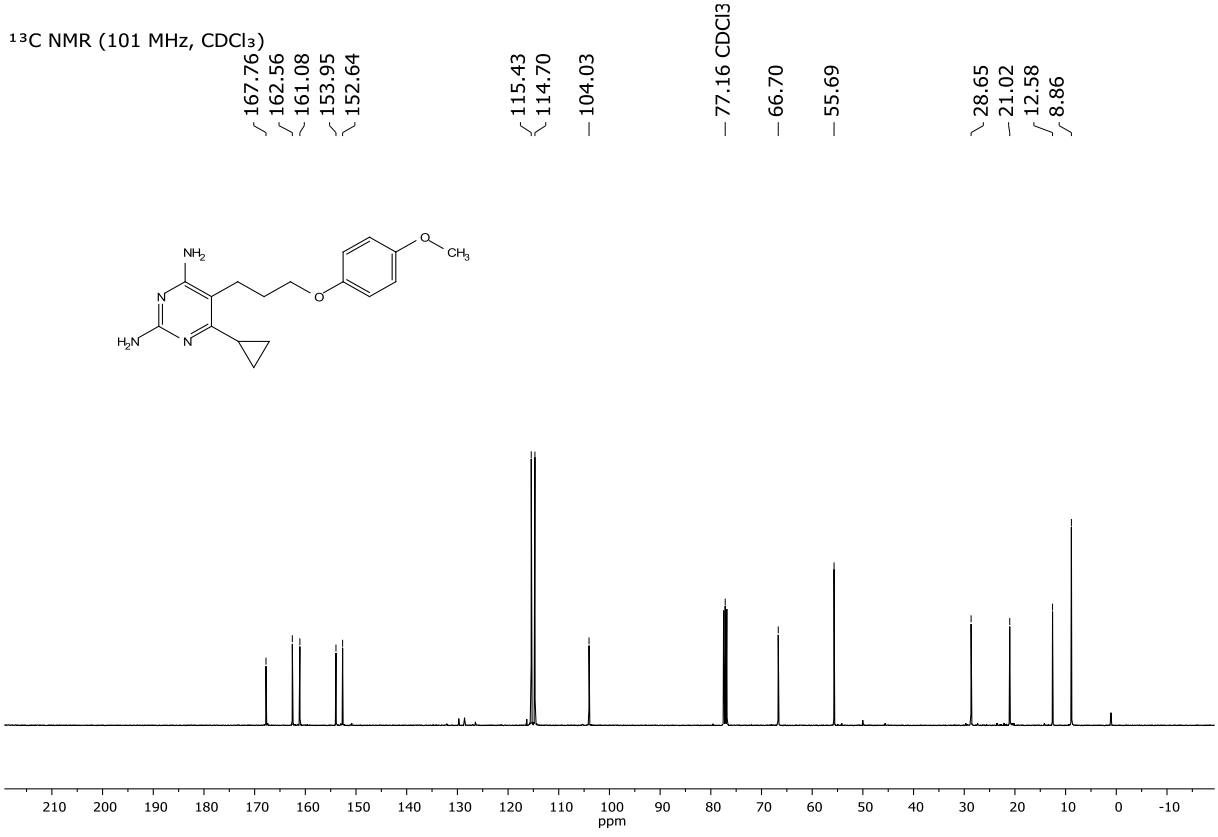
^{1}H NMR:

^{1}H NMR (400 MHz, $CDCl_3$)

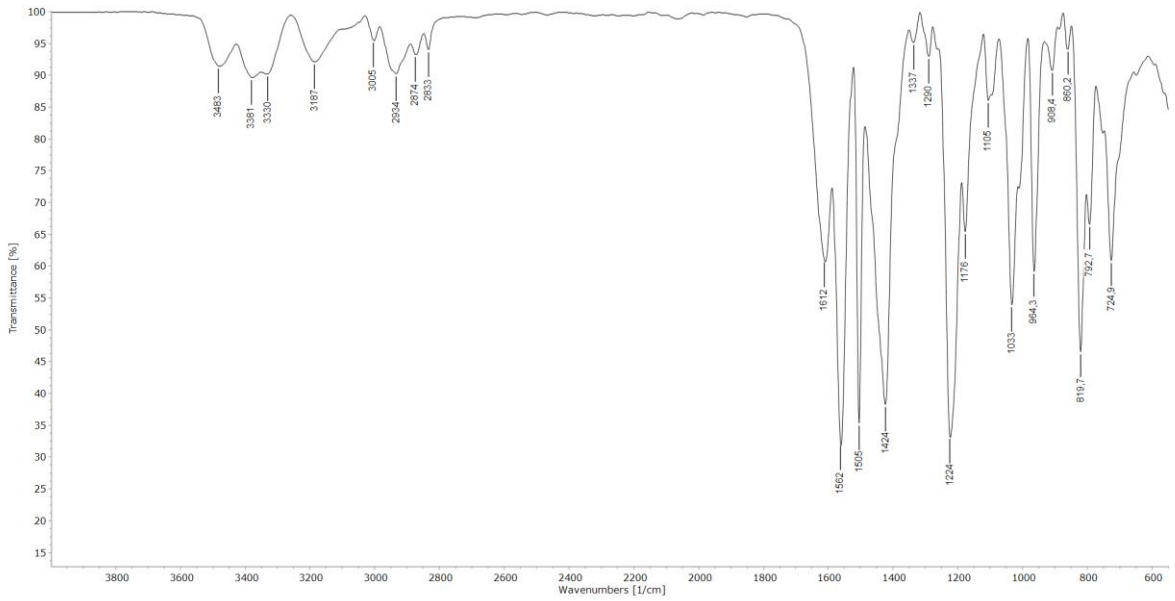


^{13}C NMR:

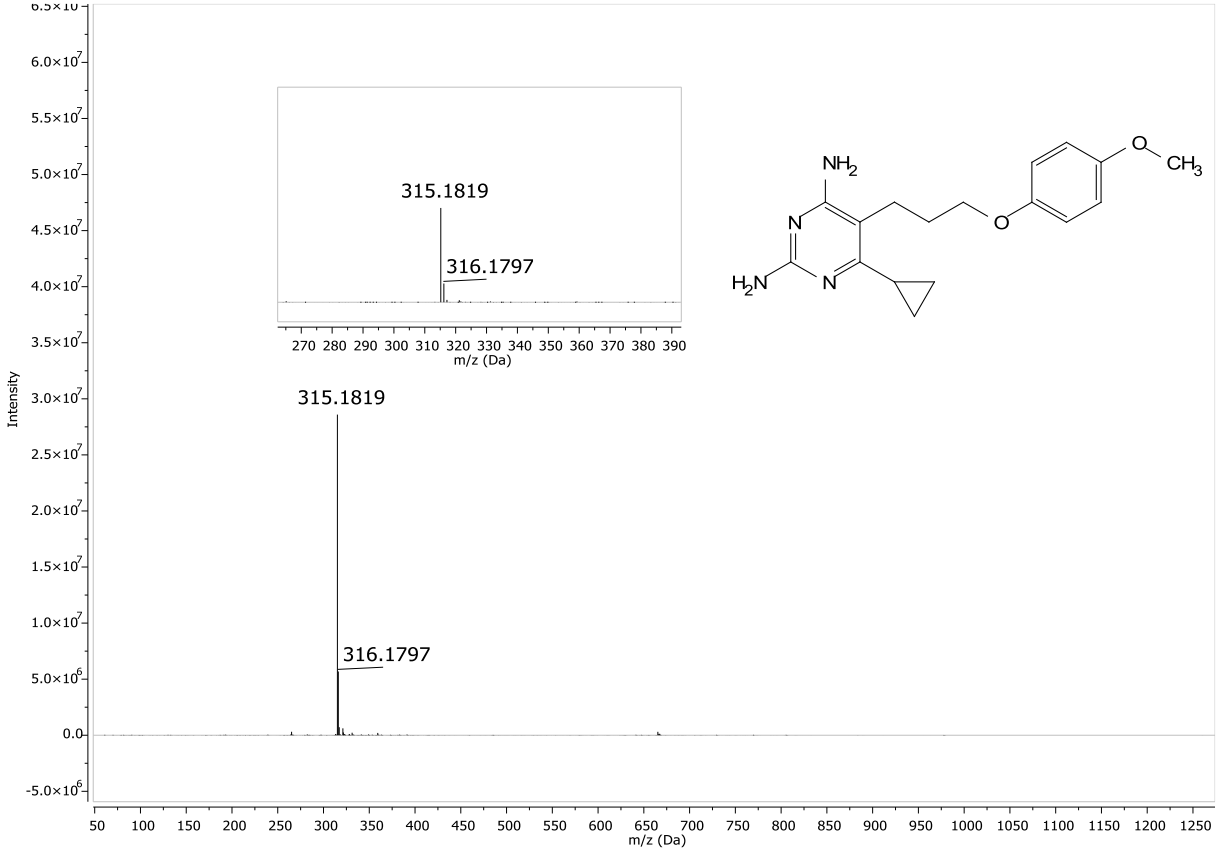
^{13}C NMR (101 MHz, $CDCl_3$)



FTIR:


Transmittance [%]
Wavenumbers [1/cm]
3483
3381
3330
3187
3005
2934
2874
2833
1612
1562
1505
1424
1337
1290
1224
1176
1105
1033
964.3
908.4
860.2
819.7
792.7
724.9

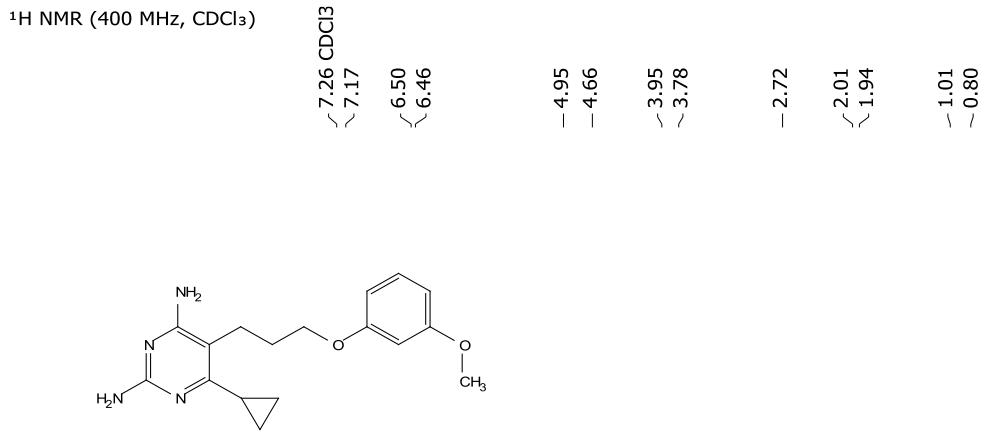
HRMS:

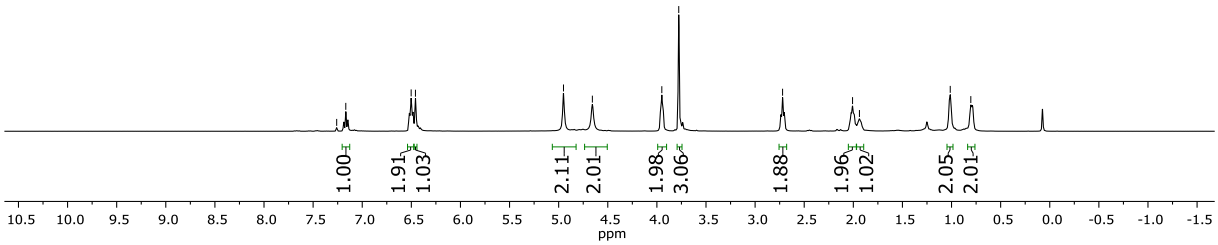

315.1819
316.1797
m/z (Da)
Intensity
315.1819
316.1797
m/z (Da)
NH2
N
H2N
N
O
O
CH3

6-Cyclopropyl-5-[3-(3-methoxyphenoxy)propyl]pyrimidine-2,4-diamine (**9m**)

^{1}H NMR:

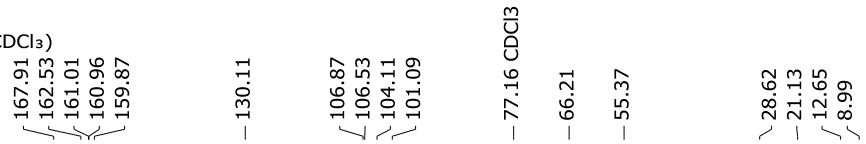
^{1}H NMR (400 MHz, $CDCl_3$)

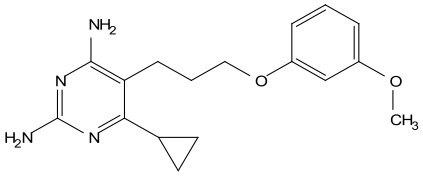


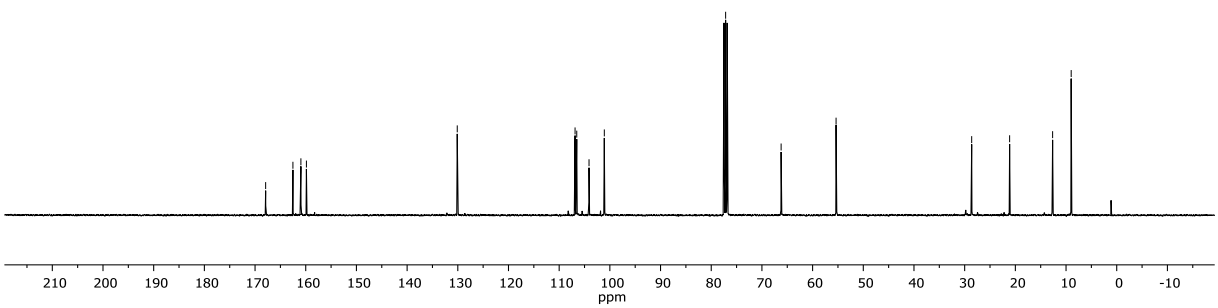


^{13}C NMR:

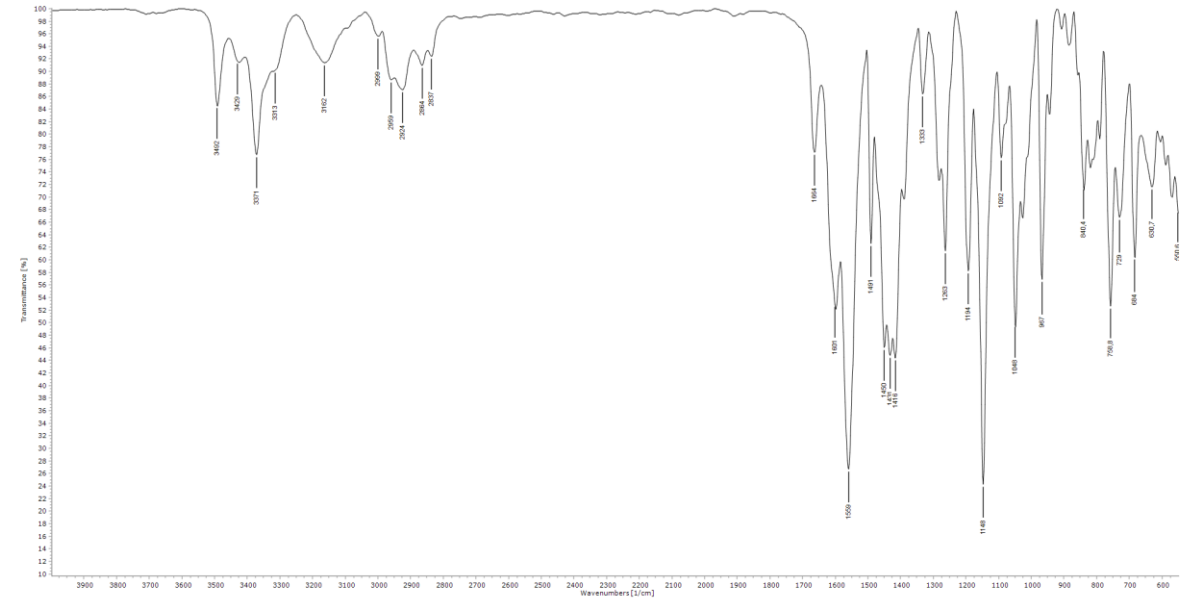
^{13}C NMR (101 MHz, $CDCl_3$)



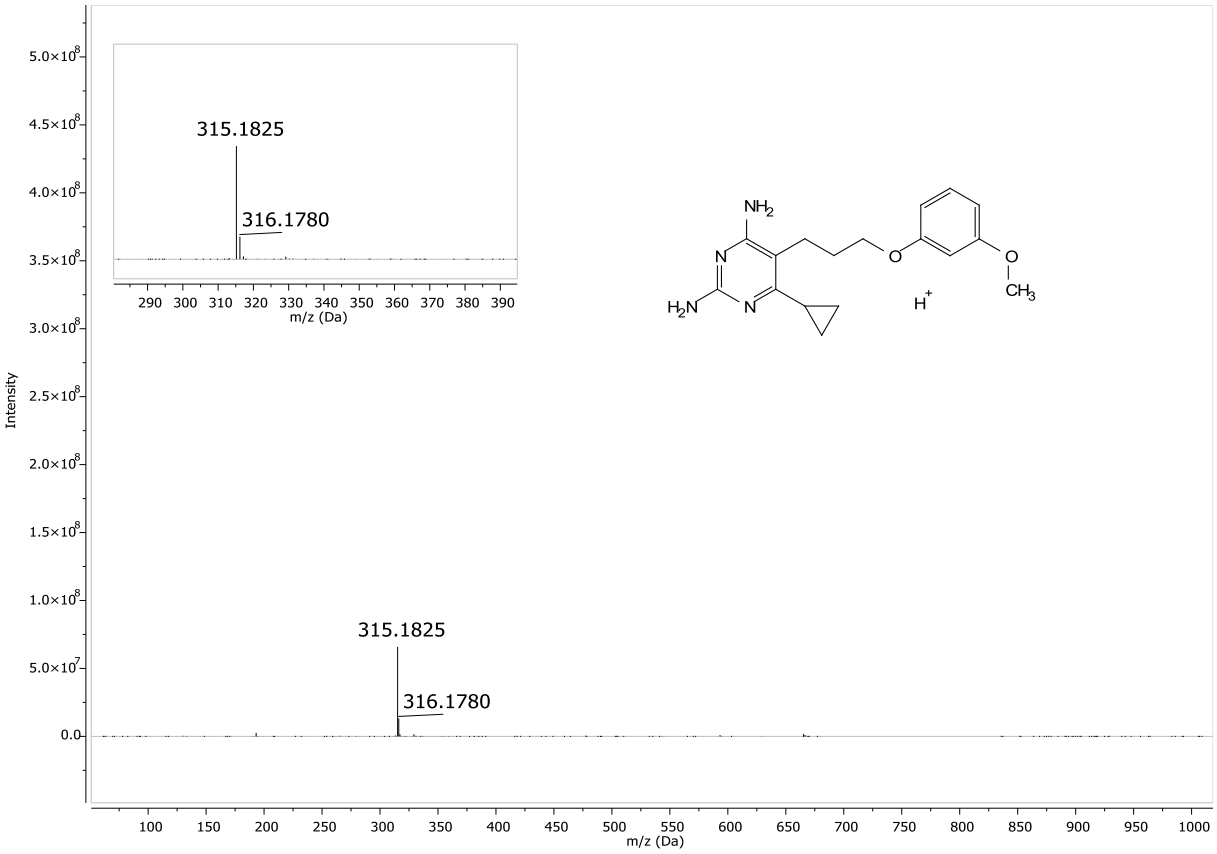




FTIR:


Transmittance [%]
Wavenumbers [1/cm]

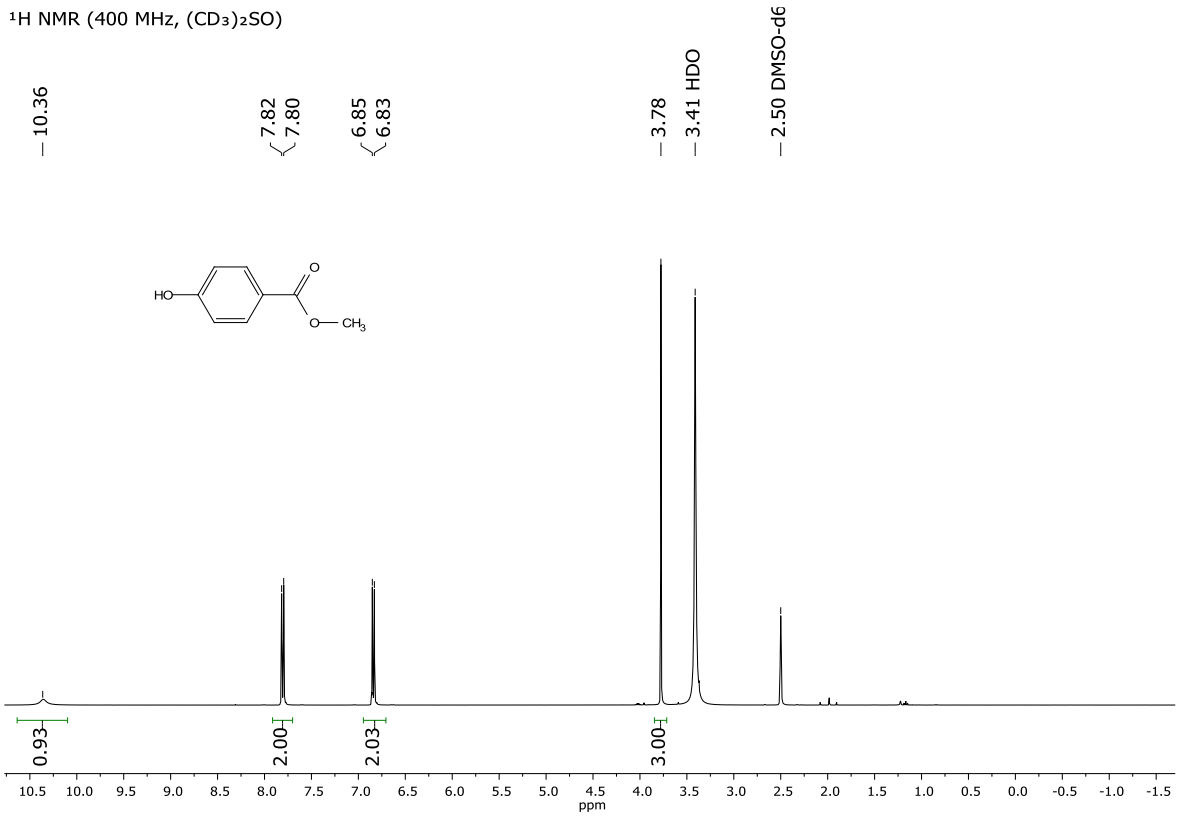
HRMS:


315.1825
316.1780
m/z (Da)
NH2
N
O
O
CH3
H2N
N
H+
Intensity
315.1825
316.1780
m/z (Da)

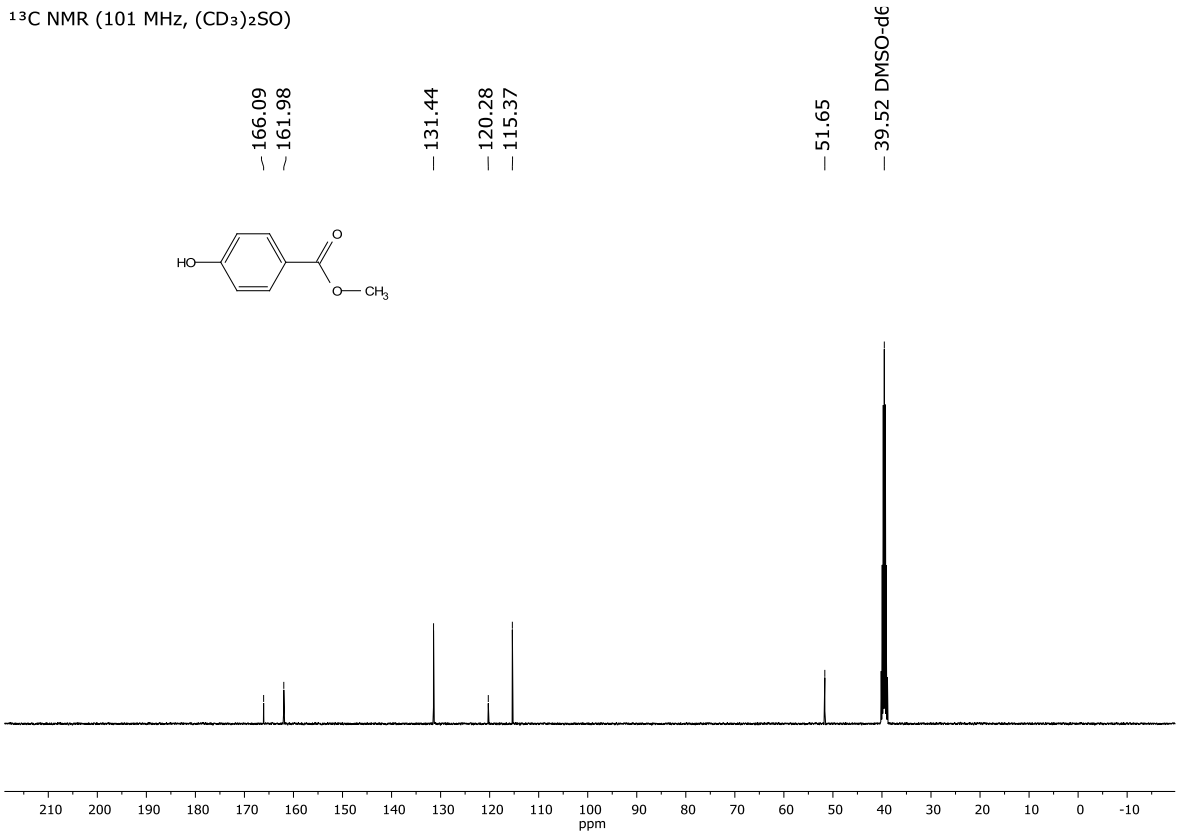
Esterification of carboxylic acid containing phenols and anilines (**11**)

Methyl 4-hydroxybenzoate (**11a**)

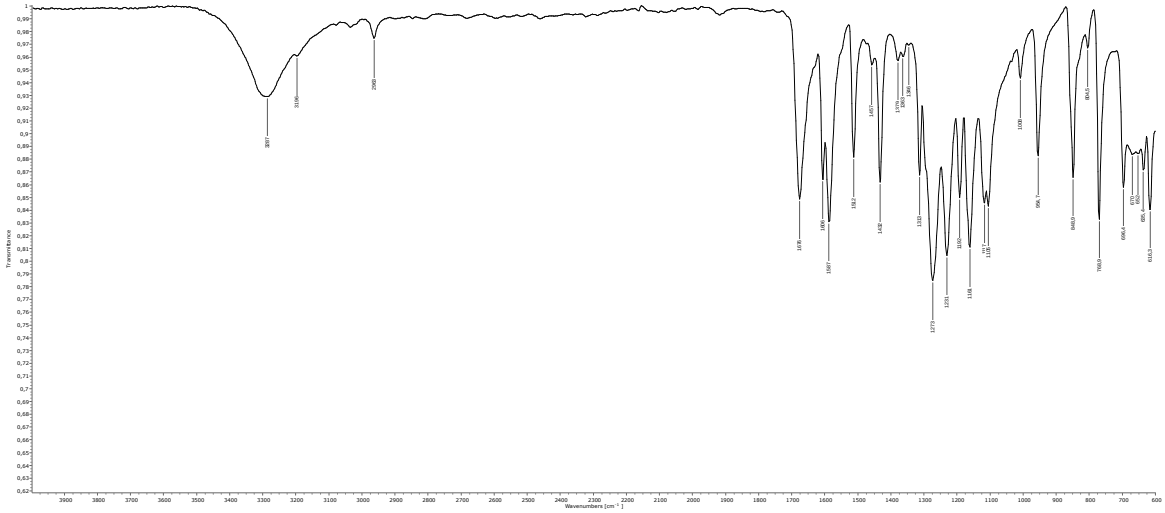
^{1}H NMR:



^{13}C NMR:



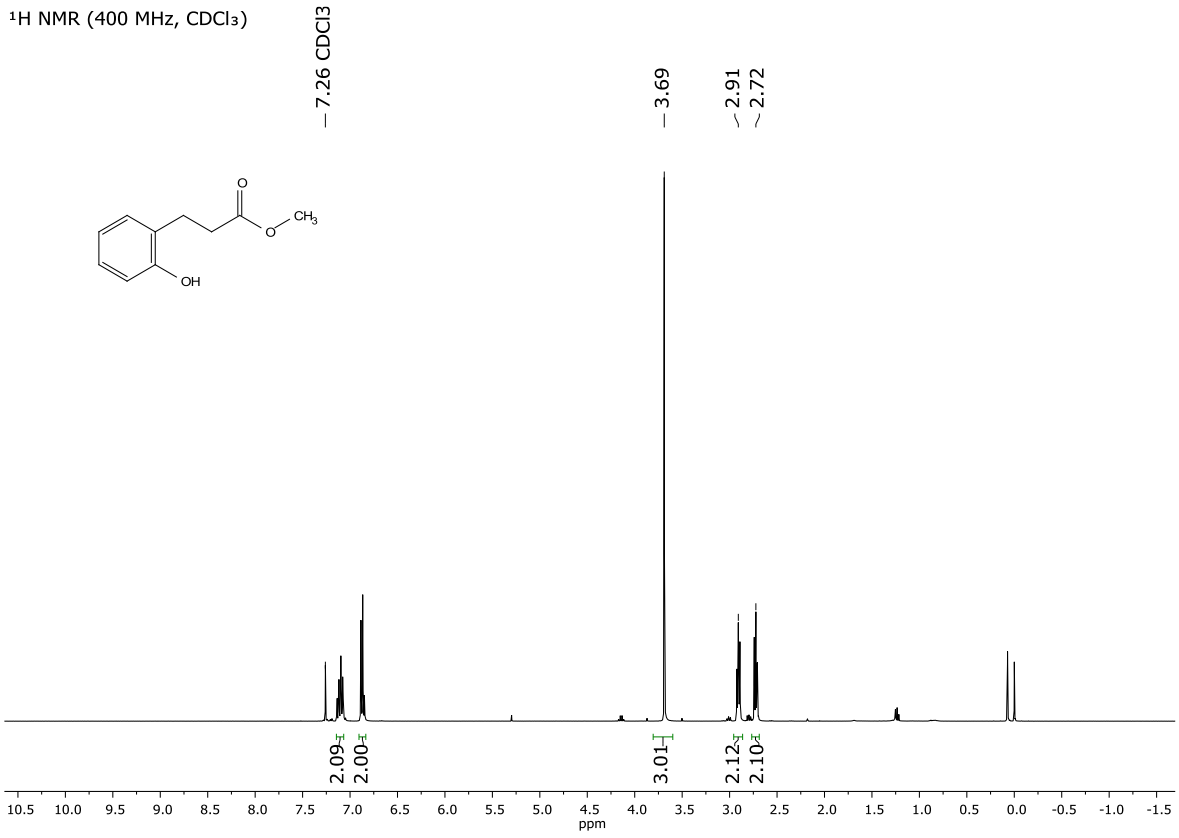
FTIR:



Methyl 3-(2-Hydroxyphenyl)propanoate (**11b**)

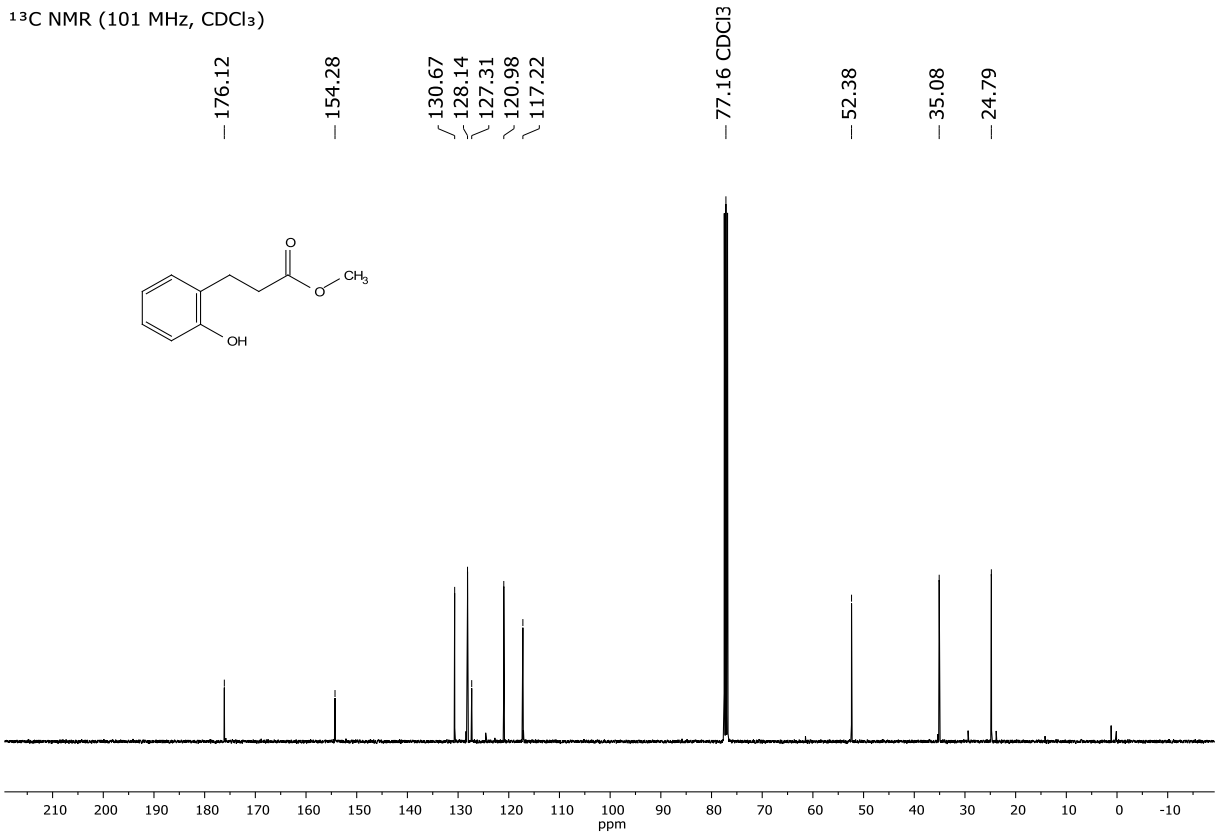
^{1}H NMR:

^{1}H NMR (400 MHz, $CDCl_3$)

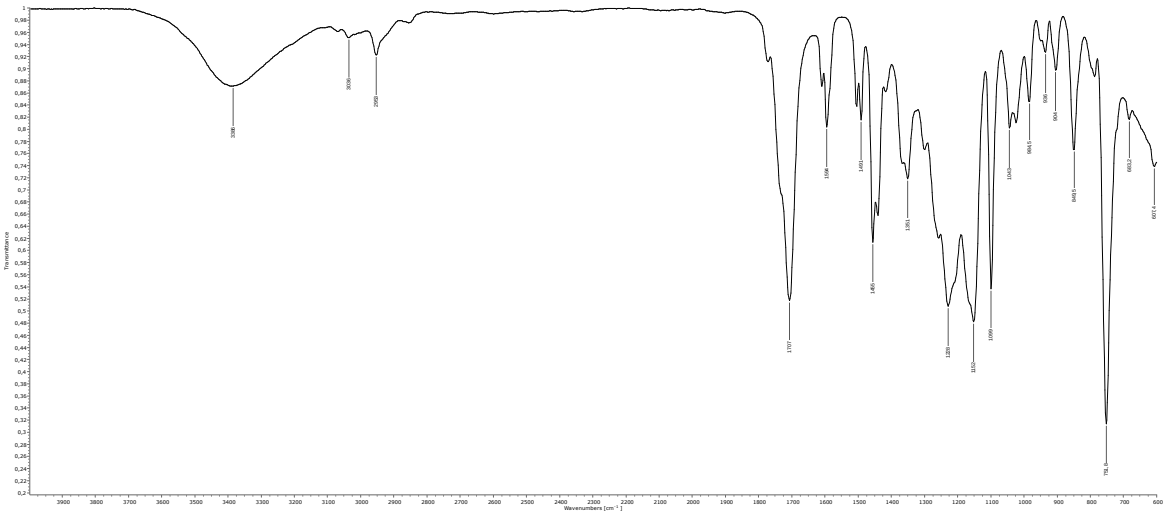


^{13}C NMR:

^{13}C NMR (101 MHz, $CDCl_3$)



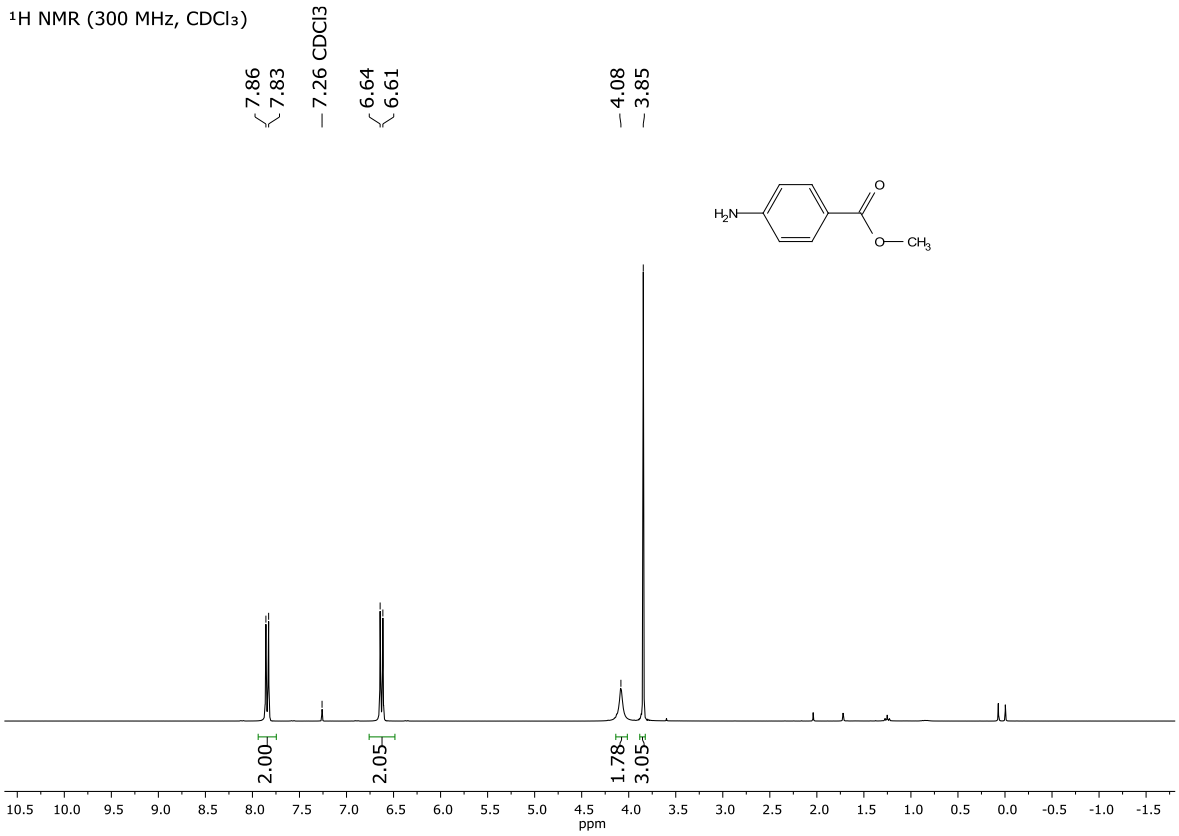
FTIR:


Transmittance
Wavenumbers [cm⁻¹]
3388
3035
2959
1707
1598
1493
1460
1351
1289
1101
1099
1043
994.1
935
906
849.9
781.9
693.2
617.4

Methyl 4-aminobenzoate (**11c**)

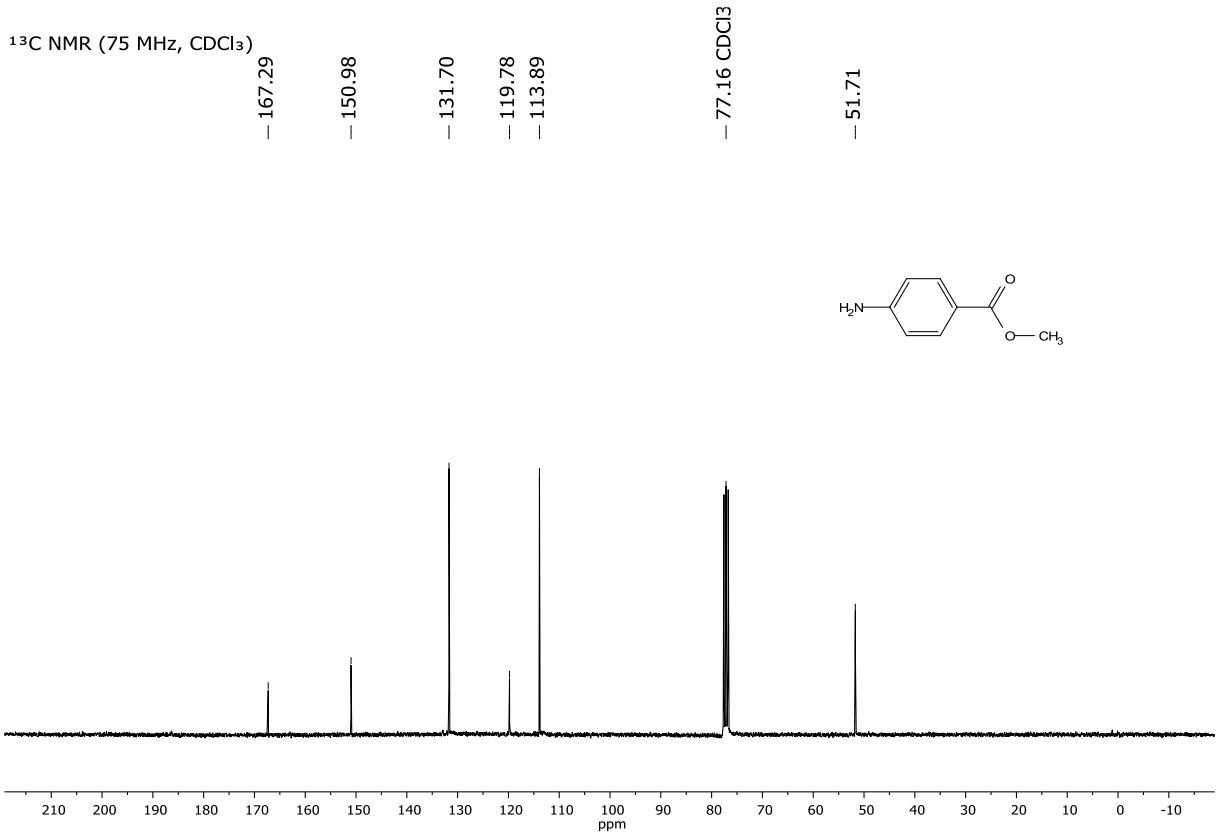
^{1}H NMR:

^{1}H NMR (300 MHz, $CDCl_3$)

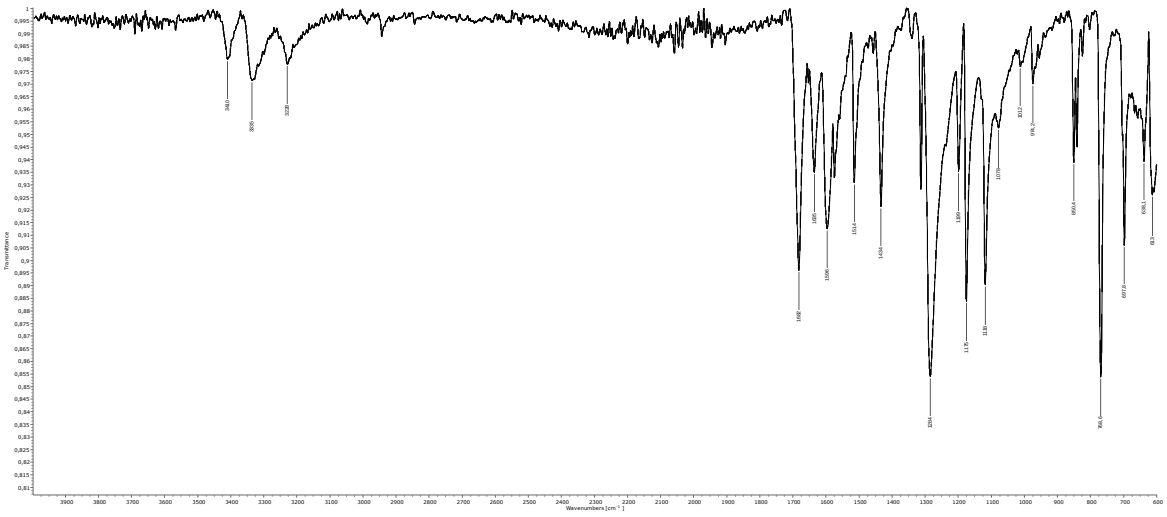


^{13}C NMR:

^{13}C NMR (75 MHz, $CDCl_3$)



FTIR:



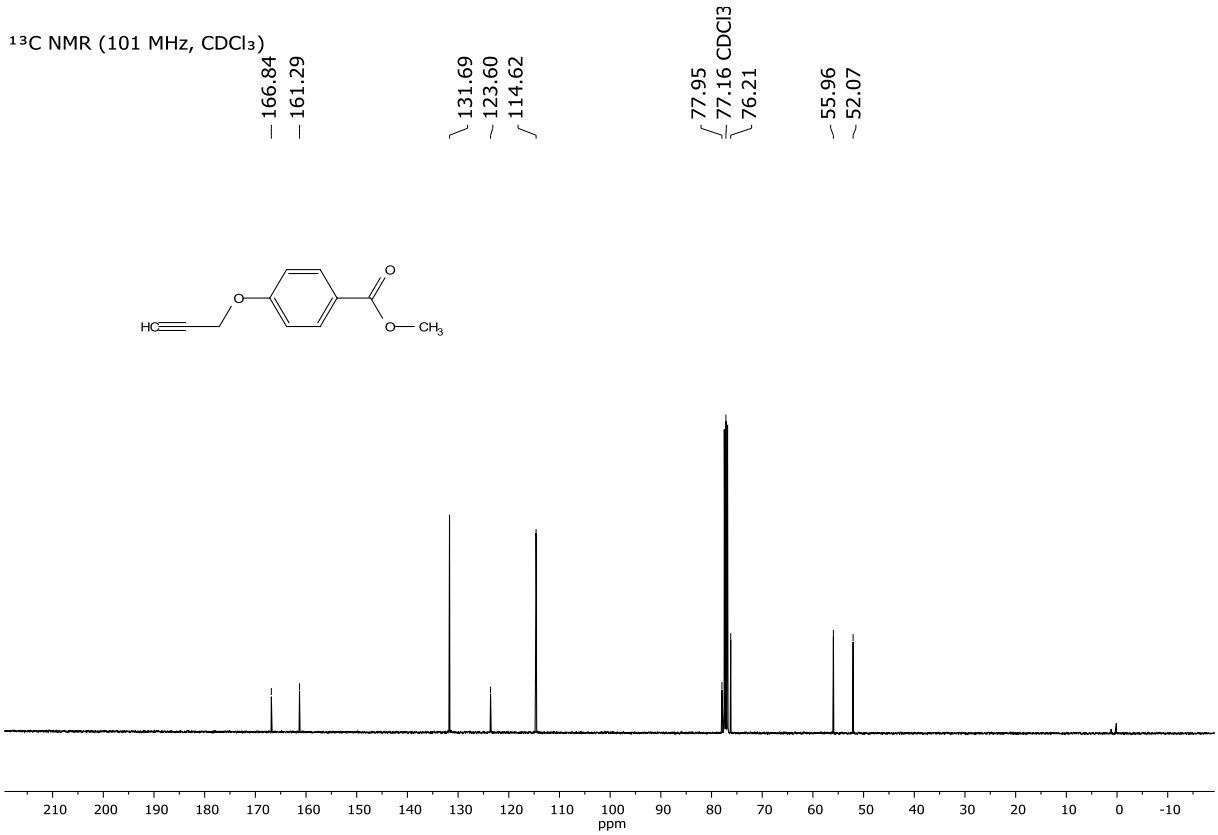
Alkylation of phenols for second-generation analogues(**12**)

1-(Prop-2-yn-1-yloxy)-4-methylbenzoate (**12a**)

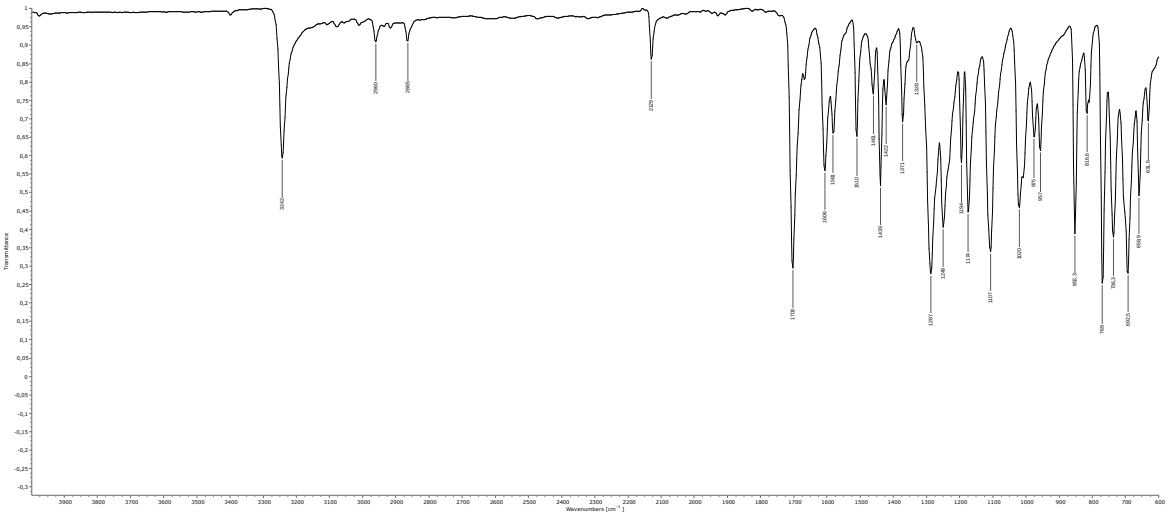
^{1}H NMR:



^{13}C NMR:



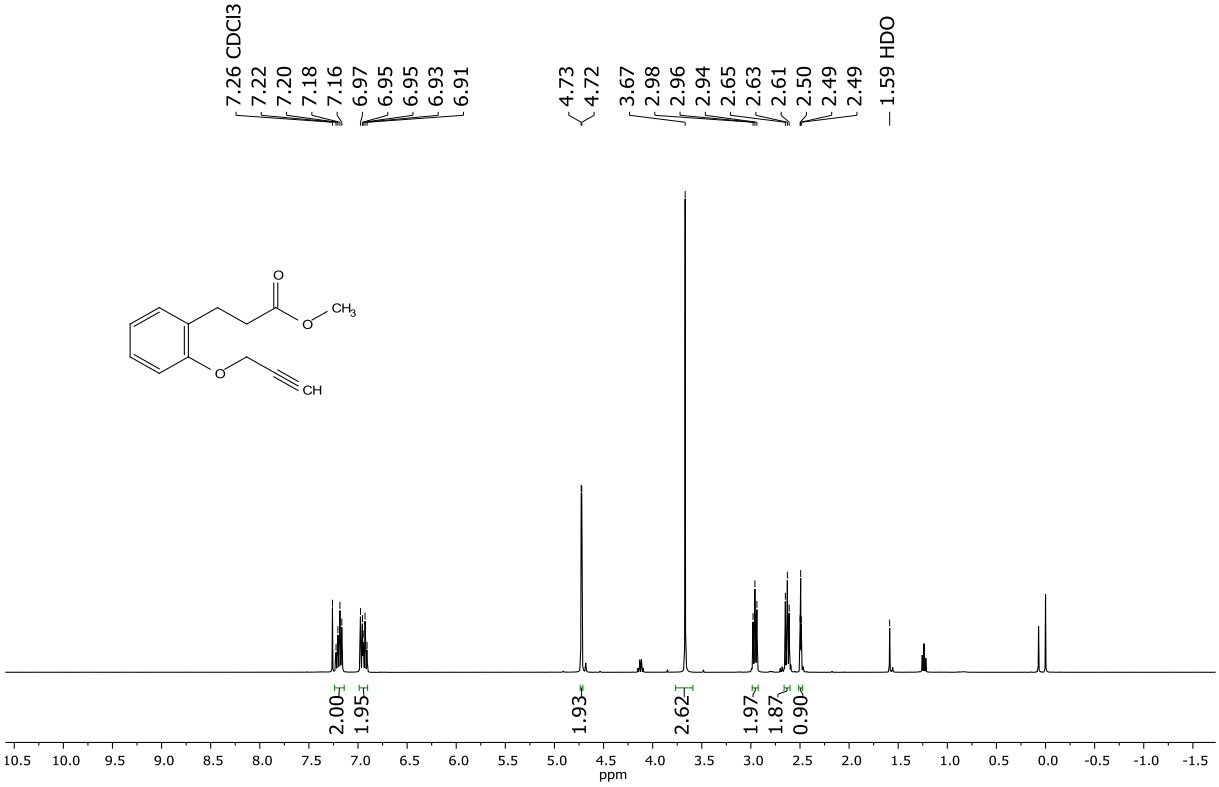
FTIR:


Transmittance
Wavenumbers [cm⁻¹]
3420
2960
2895
2325
1720
1650
1608
1510
1456
1422
1371
1320
1246
1178
1107
1030
997
970
890
816
760
706.1
669.9
625.4

[1-(Prop-2-yn-1-yloxy)phenyl]-2-methylpropanoate (**12b**)

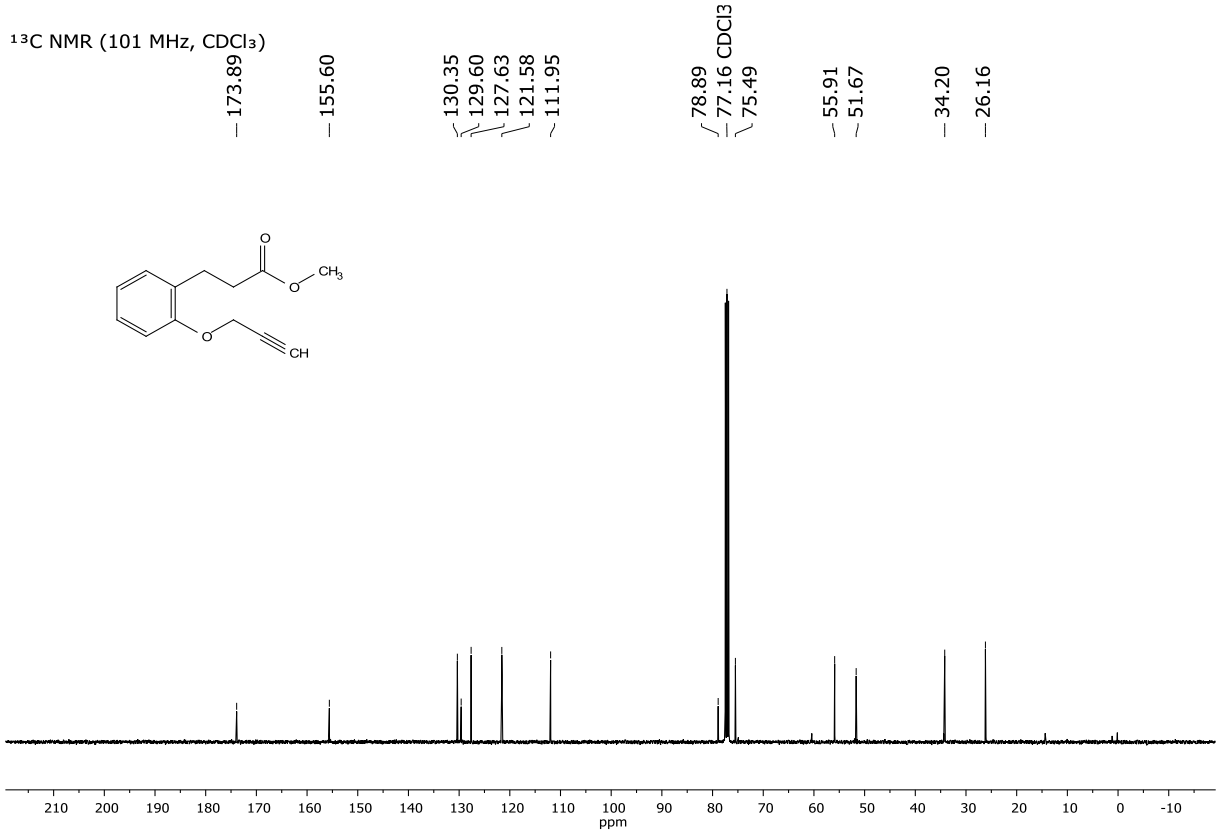
^{1}H NMR:

^{1}H NMR (400 MHz, $CDCl_3$)

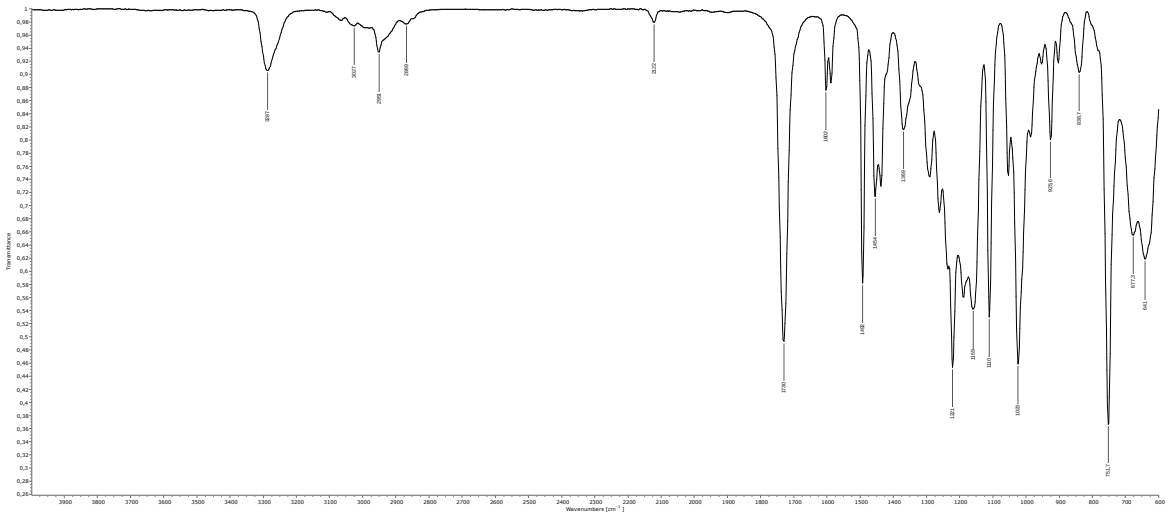


^{13}C NMR:

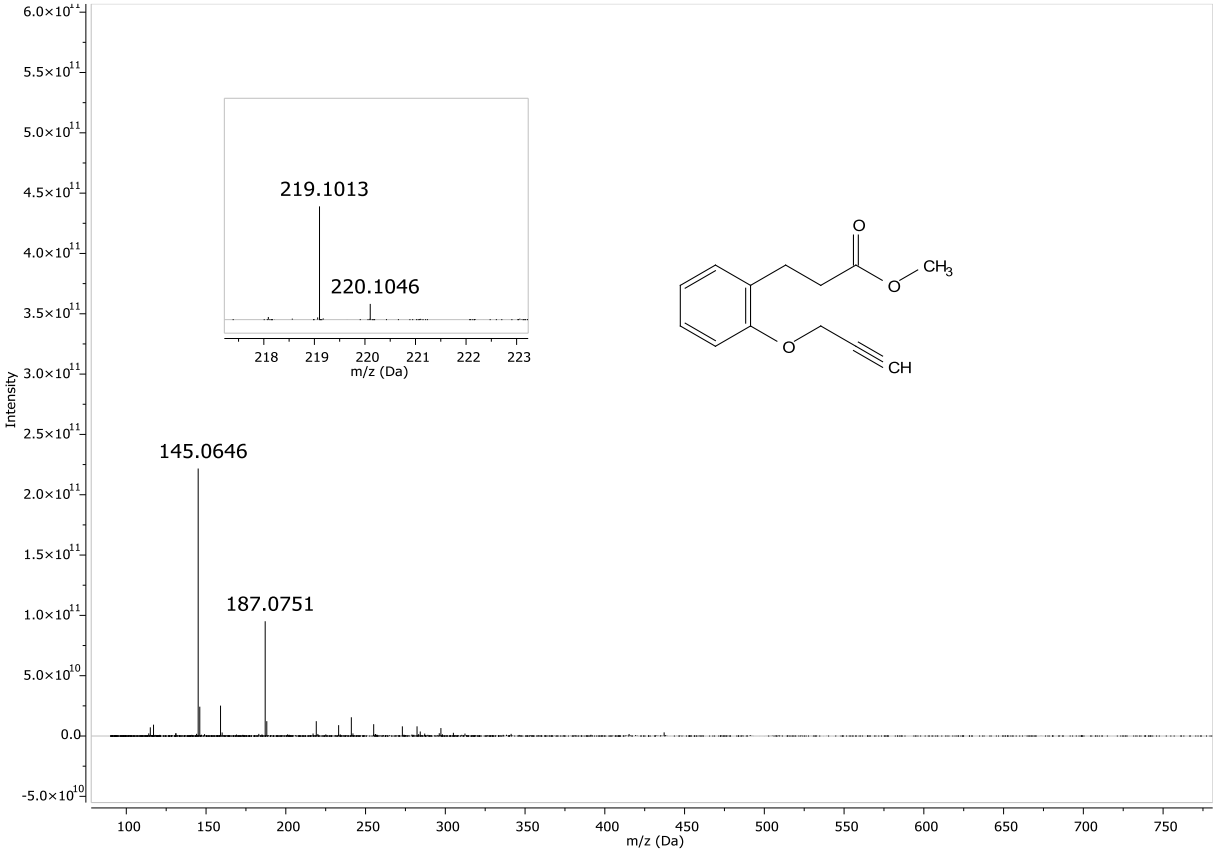
^{13}C NMR (101 MHz, $CDCl_3$)



FTIR:



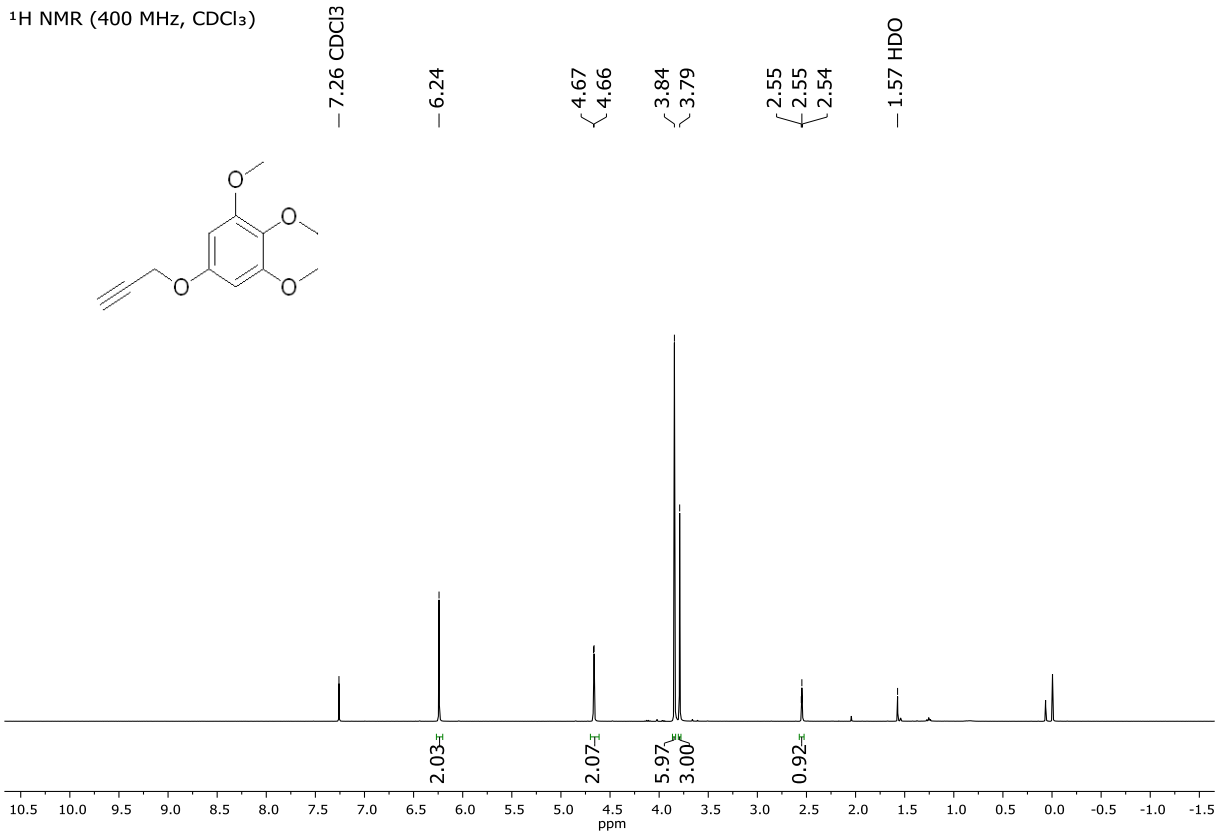
HRMS:


219.1013
220.1046
218 219 220 221 222 223
m/z (Da)
O
O
CH3
O
CH
145.0646
187.0751
Intensity
100 150 200 250 300 350 400 450 500 550 600 650 700 750
m/z (Da)

1-(Prop-2-yn-1-yloxy)-3,4,5-trimethoxybenzene (**12c**)

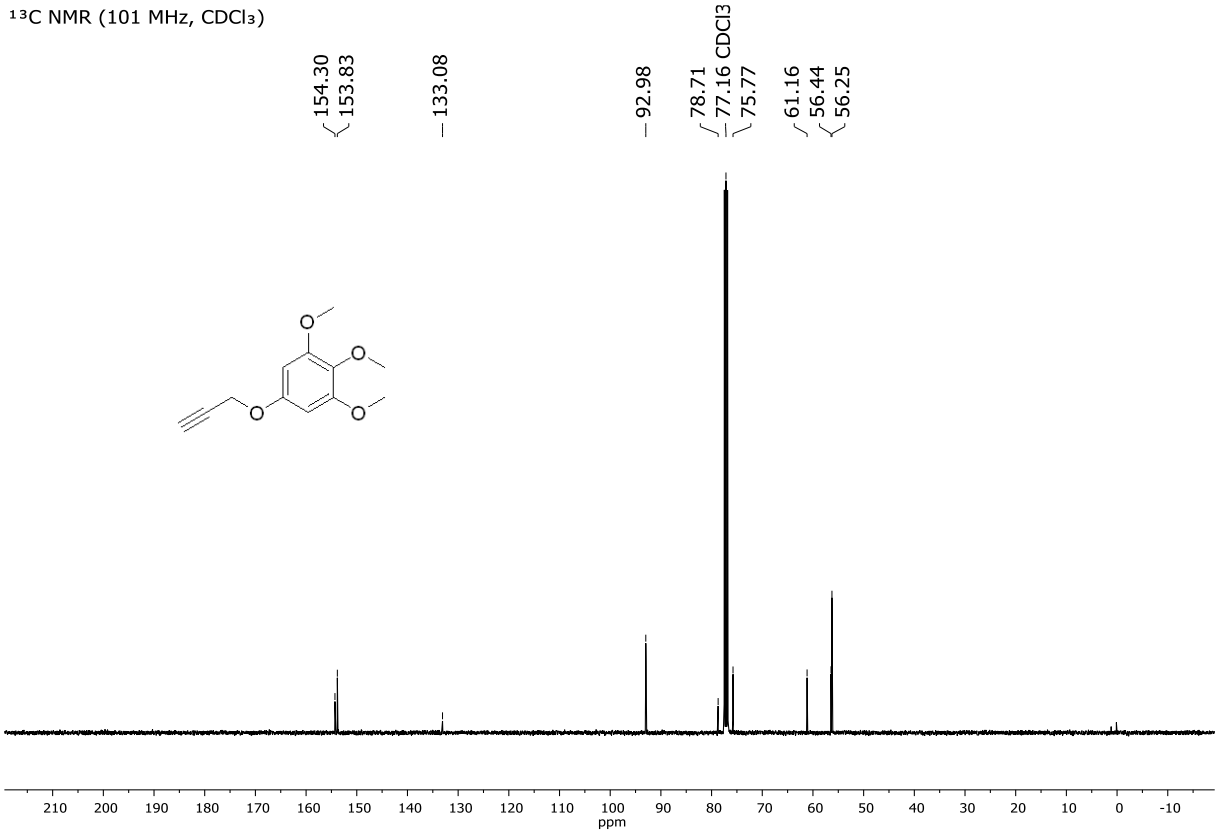
^{1}H NMR:

^{1}H NMR (400 MHz, $CDCl_3$)

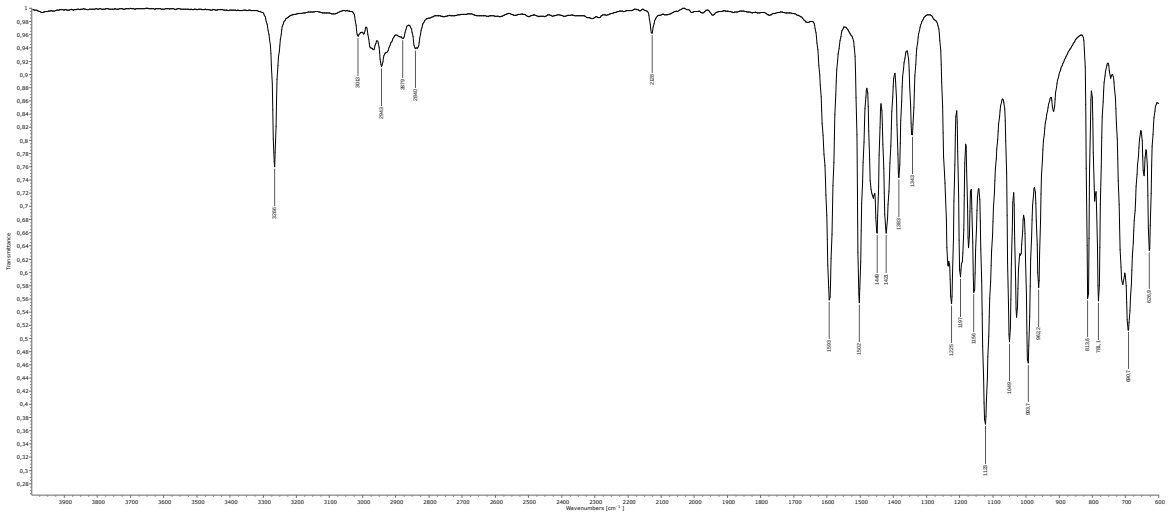


^{13}C NMR:

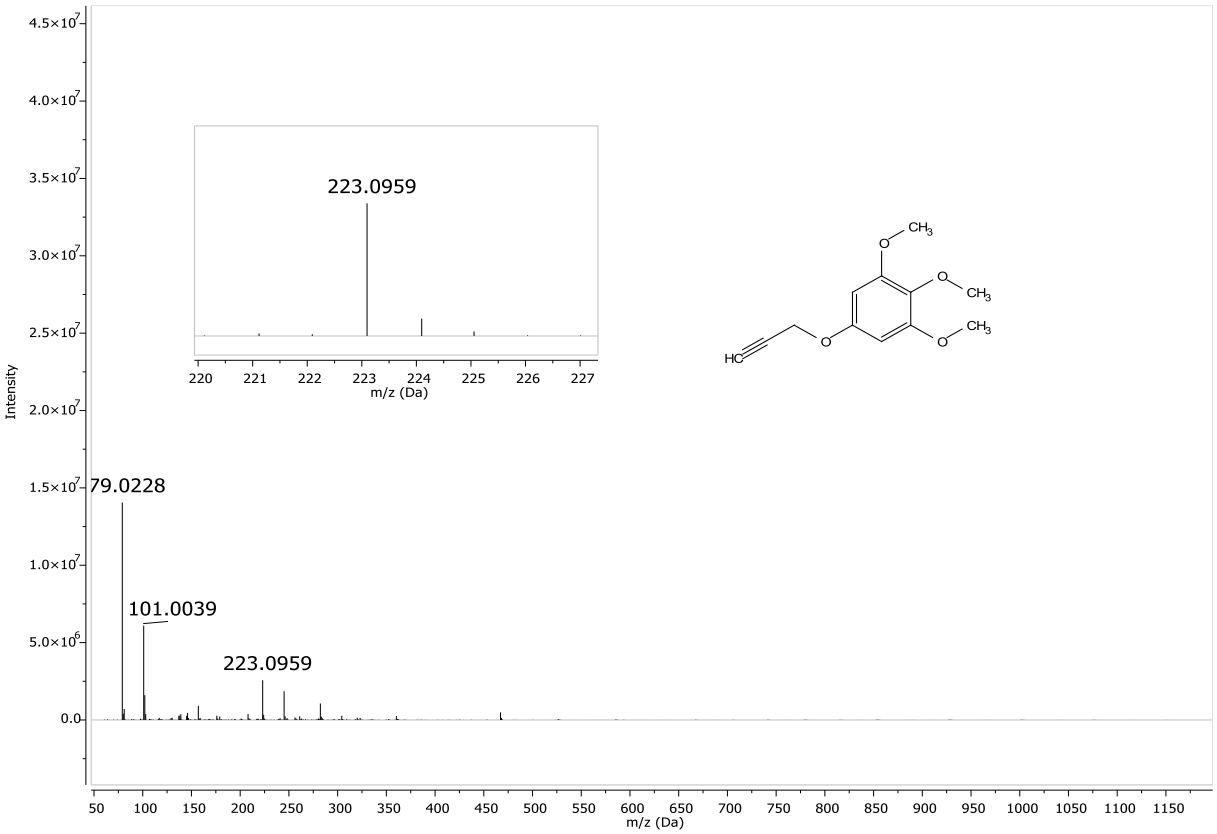
^{13}C NMR (101 MHz, $CDCl_3$)



FTIR:



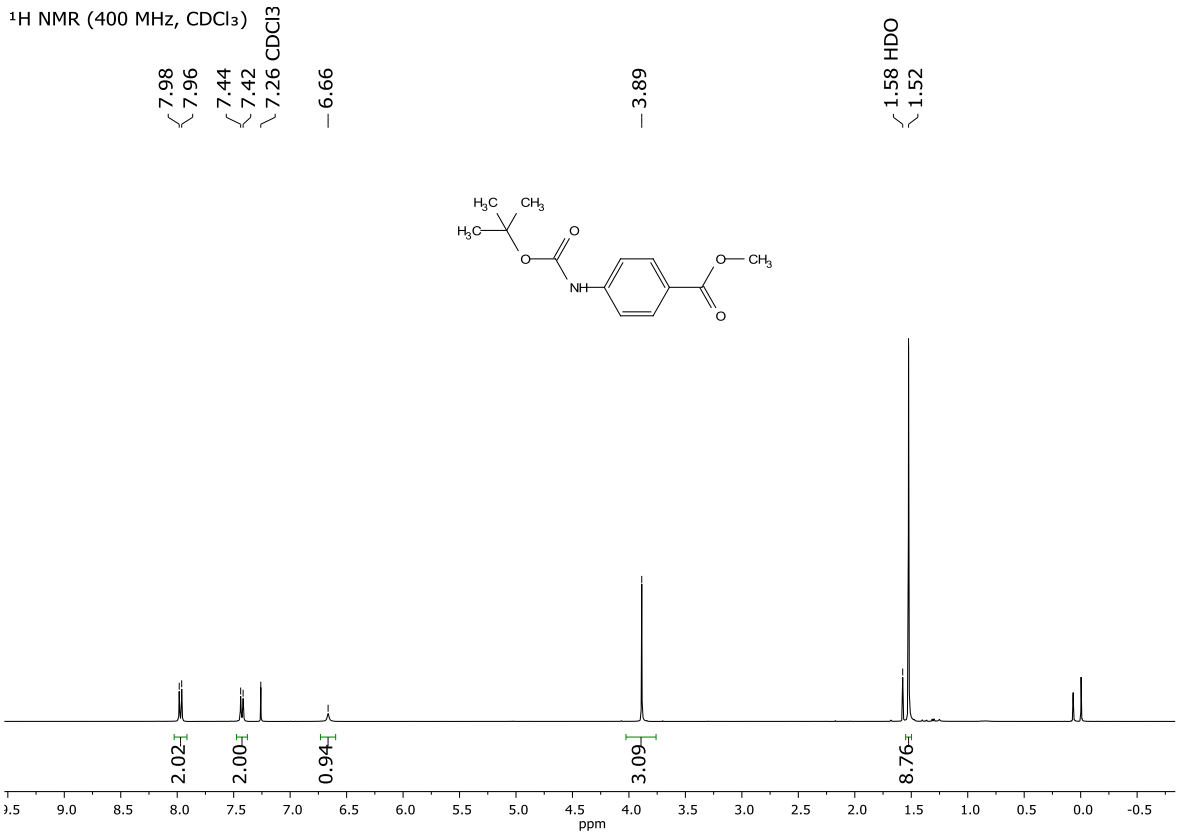
HRMS:


223.0959
m/z (Da)
CH3
O
CH3
CH3
HC
Intensity
79.0228
101.0039
223.0959
m/z (Da)

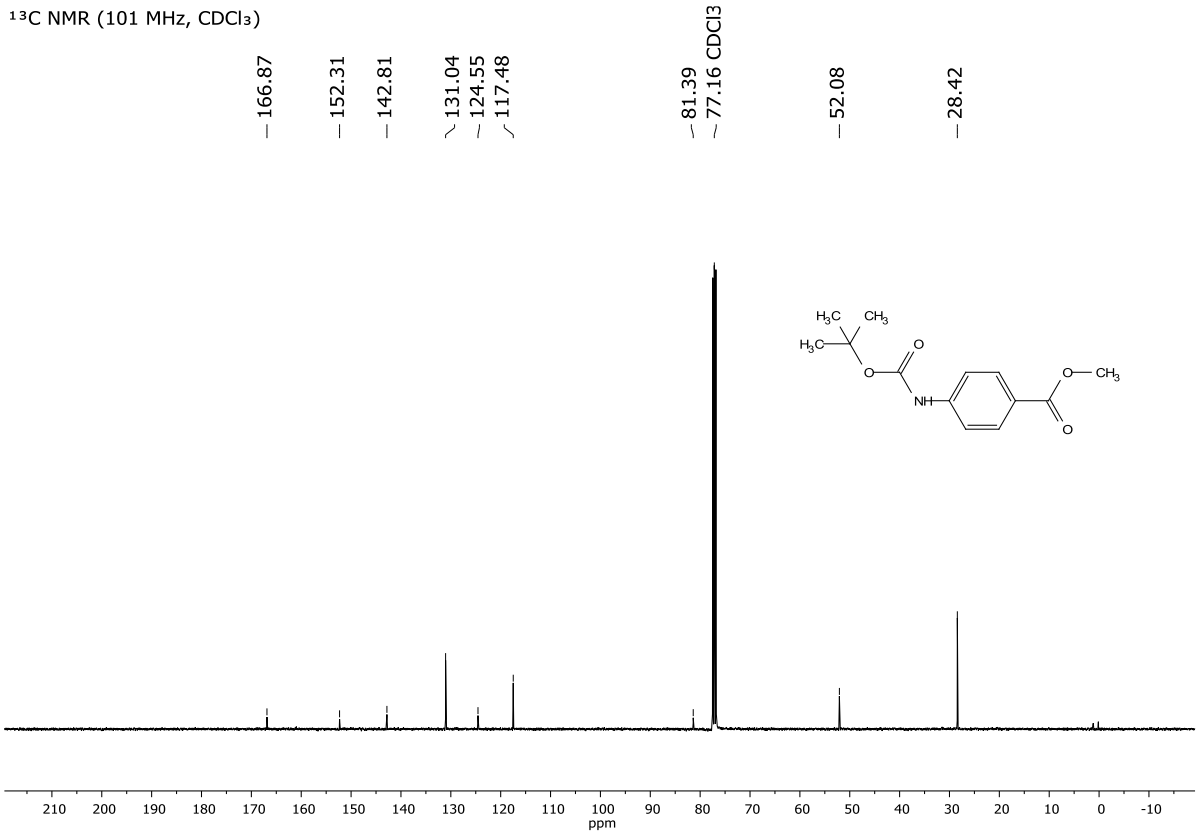
Boc protection of anilines (**13**)

Methyl 4-[(*tert*-butoxycarbonyl)amino]benzoate (**13a**)

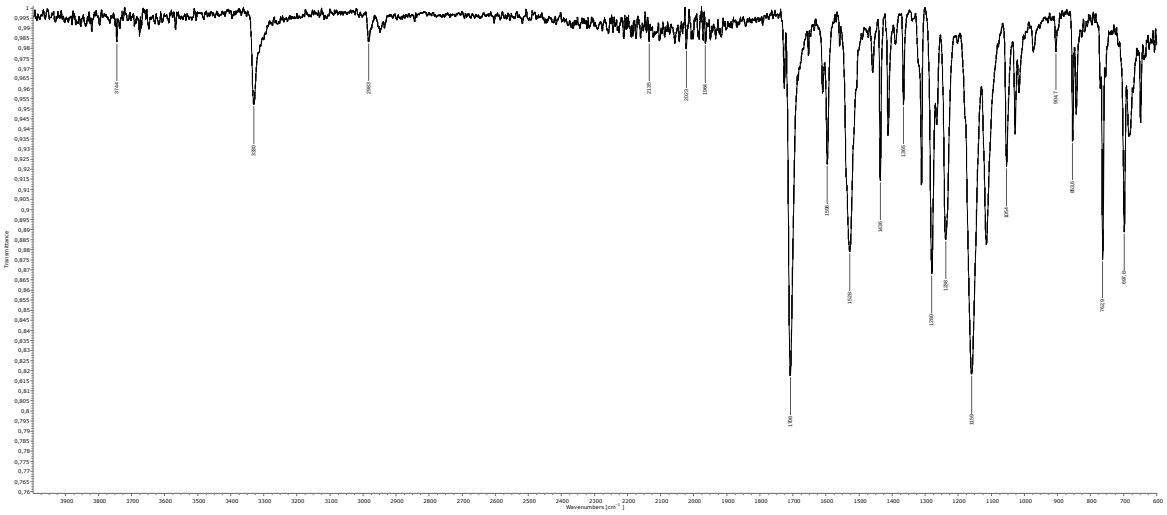
^{1}H NMR:



^{13}C NMR:



FTIR:


Transmittance
Wavenumbers [cm⁻¹]
3744
3385
2980
2135
2027
1966
1636
1508
1450
1365
1265
1169
1028
904.7
818.8
752.9
697.8

tert-Butyl-(3,4,5-trimethoxyphenyl)carbamate (**13b**)

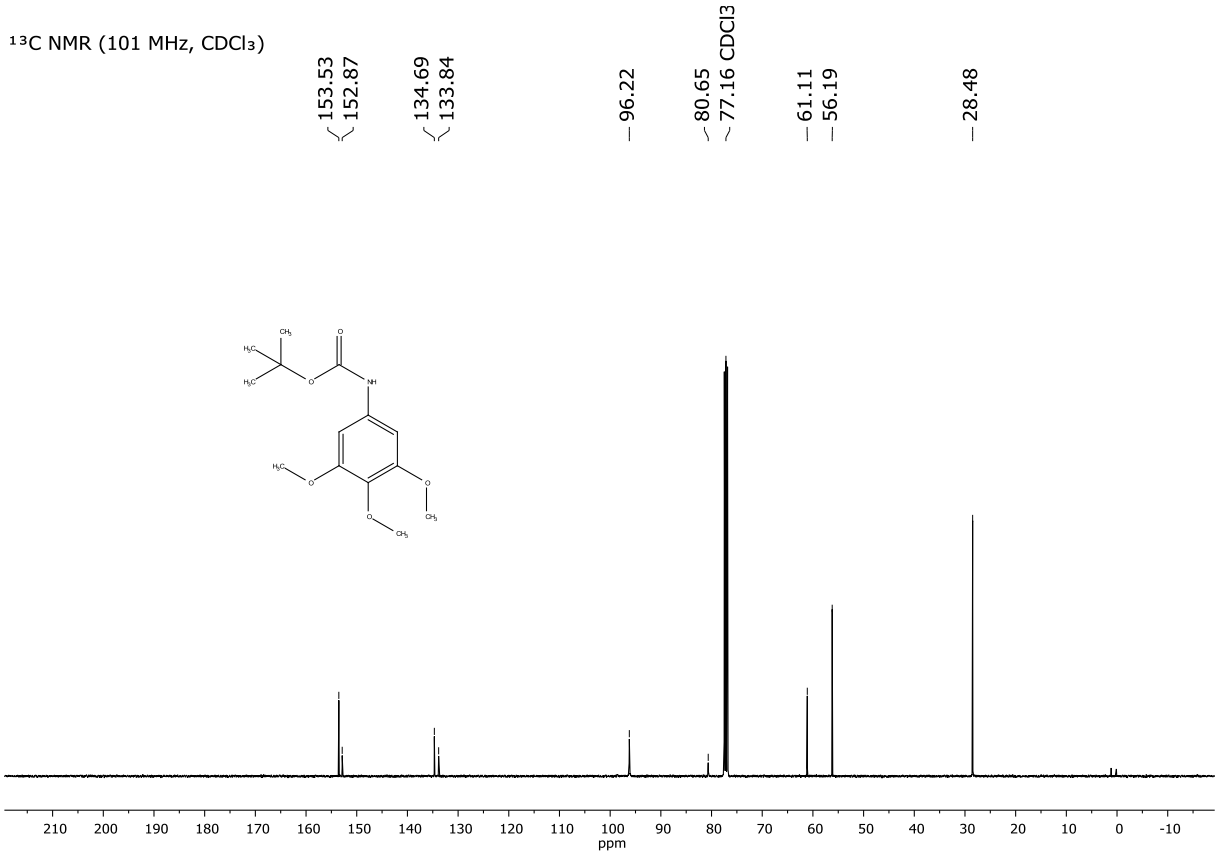
^{1}H NMR:

^{1}H NMR (400 MHz, $CDCl_3$)

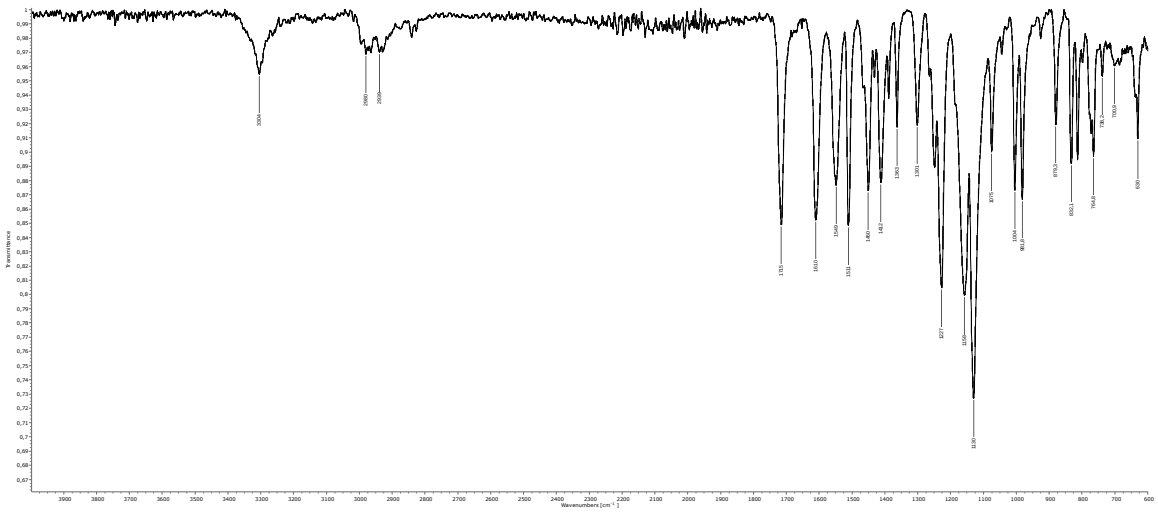


^{13}C NMR:

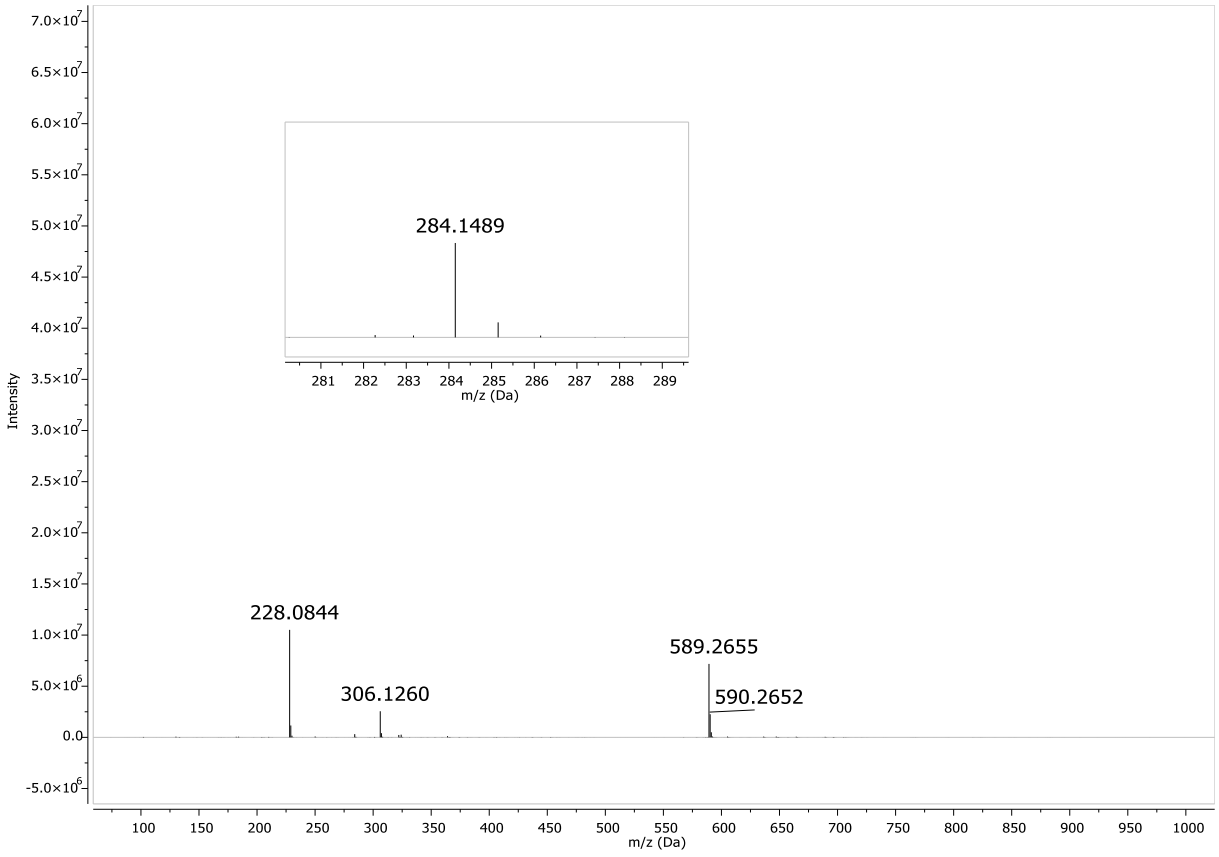
^{13}C NMR (101 MHz, $CDCl_3$)



FTIR:



HRMS:


284.1489
281 282 283 284 285 286 287 288 289
m/z (Da)
Intensity
228.0844
306.1260
589.2655
590.2652
100 150 200 250 300 350 400 450 500 550 600 650 700 750 800 850 900 950 1000
m/z (Da)

Alkylation of protected anilines (**14**)

Methyl 4-[(*tert*-butoxycarbonyl)(prop-2-yn-1-yl)amino]benzoate (**14a**)

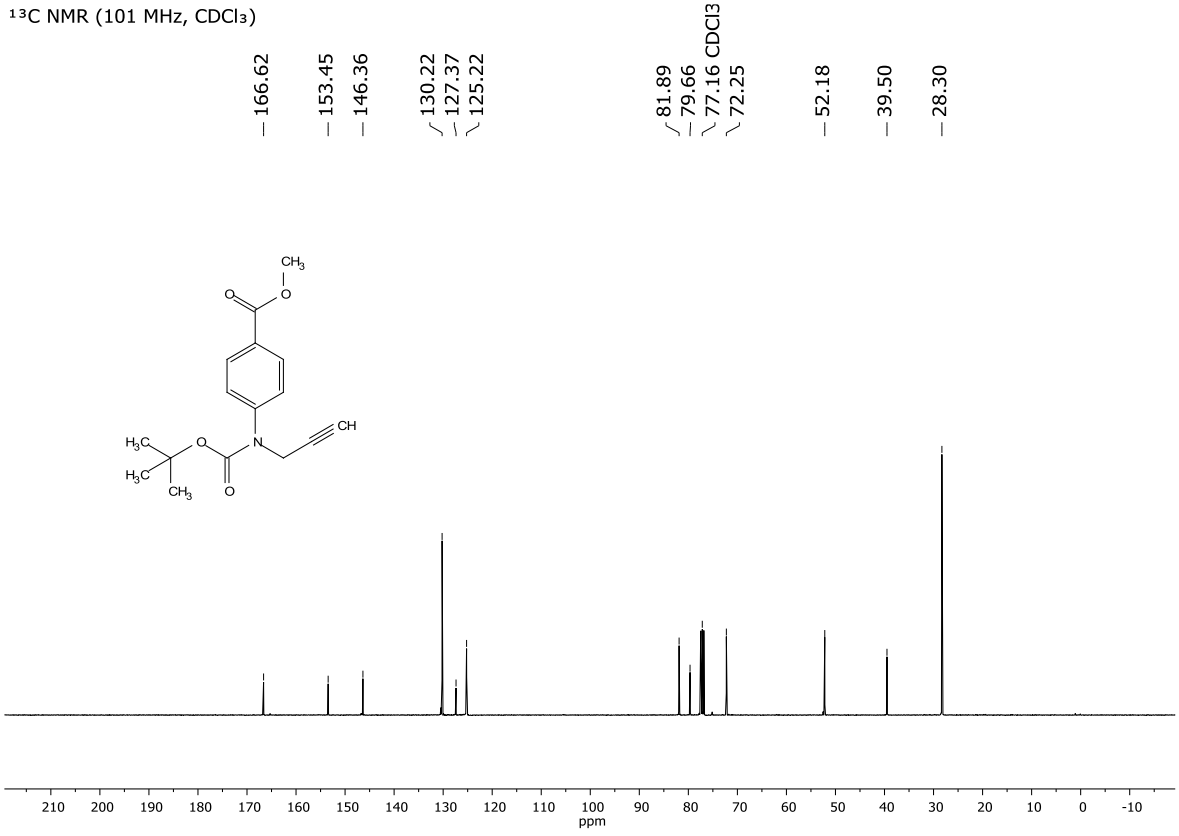
^{1}H NMR:

^{1}H NMR (400 MHz, $CDCl_3$)

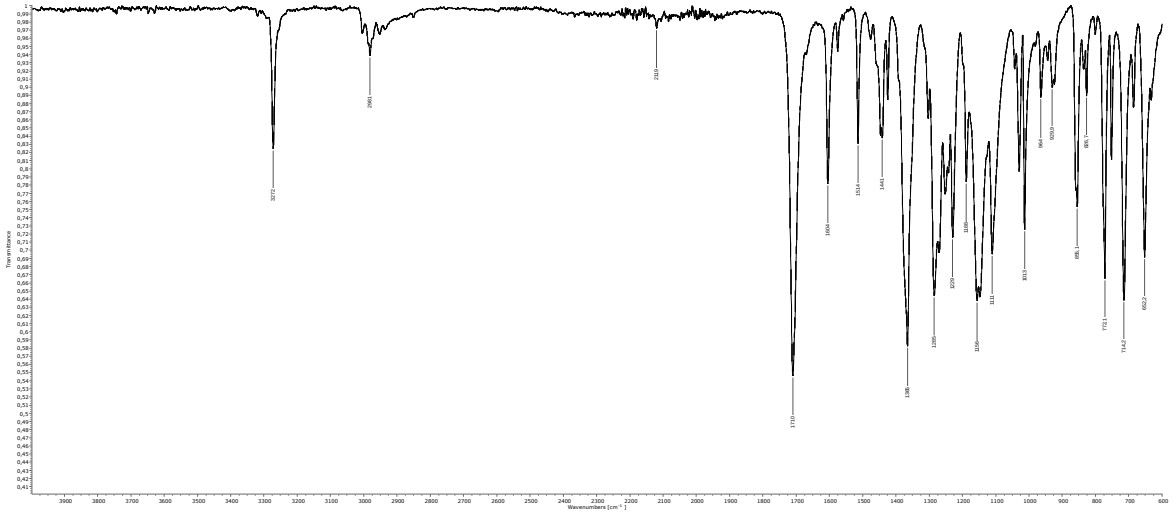


^{13}C NMR:

^{13}C NMR (101 MHz, $CDCl_3$)



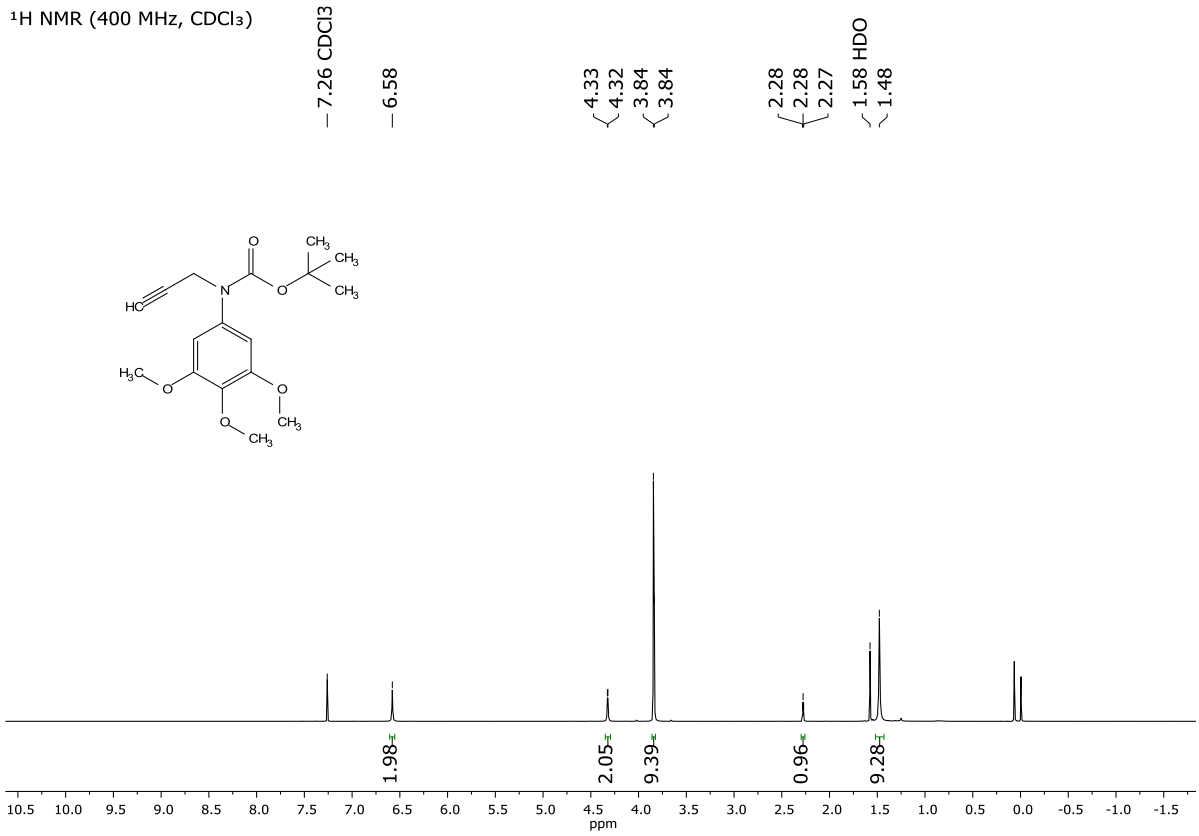
FTIR:


Transmittance
Wavenumbers [cm⁻¹]
3372
2981
2103
1710
1604
1514
1441
1385
1329
1305
1181
1111
1010
964
920
889
749
721
712
652

tert-Butyl prop-2-yn-1-yl(3,4,5-trimethoxyphenyl)carbamate (**14b**)

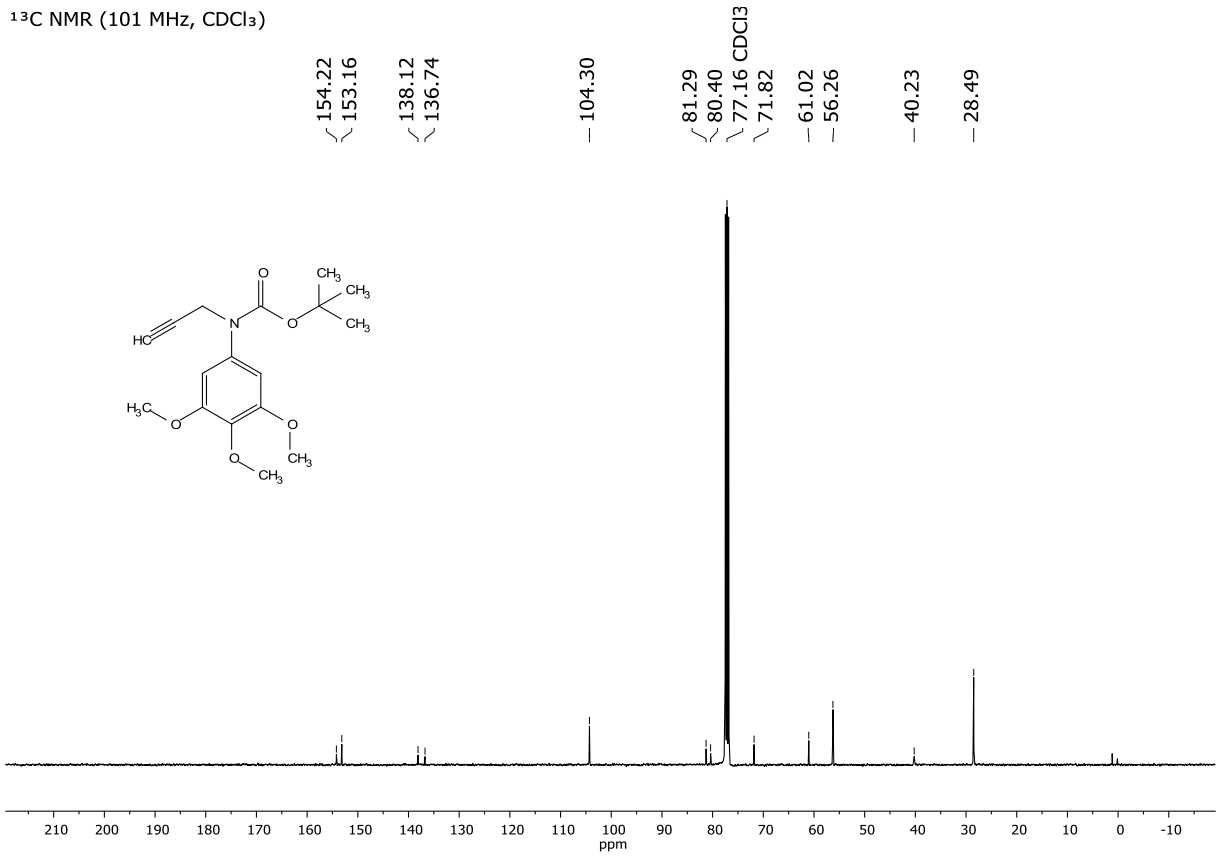
^{1}H NMR:

^{1}H NMR (400 MHz, $CDCl_3$)

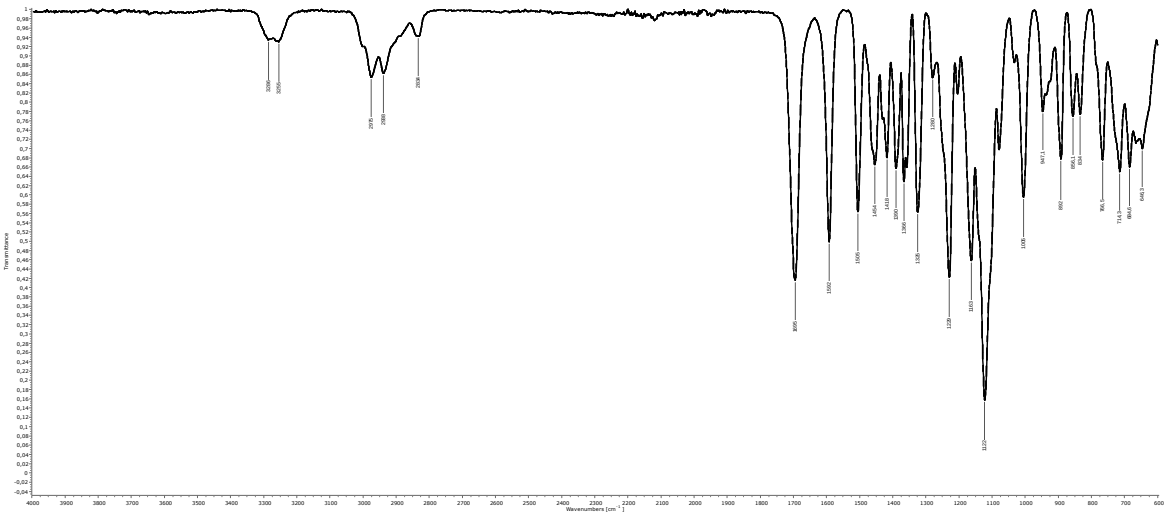


^{13}C NMR:

^{13}C NMR (101 MHz, $CDCl_3$)



FTIR:


Transmittance
Wavenumbers [cm⁻¹]
3295
3060
2975
2880
2539
1695
1550
1500
1450
1410
1384
1366
1335
1329
1222
1103
1021
947.1
892
856.1
824
766.5
714.5
680.0
649.3

Deprotection of alkylated anilines (**15**)

Methyl 4-[(*tert*-butoxycarbonyl)amino]benzoate (**15a**)

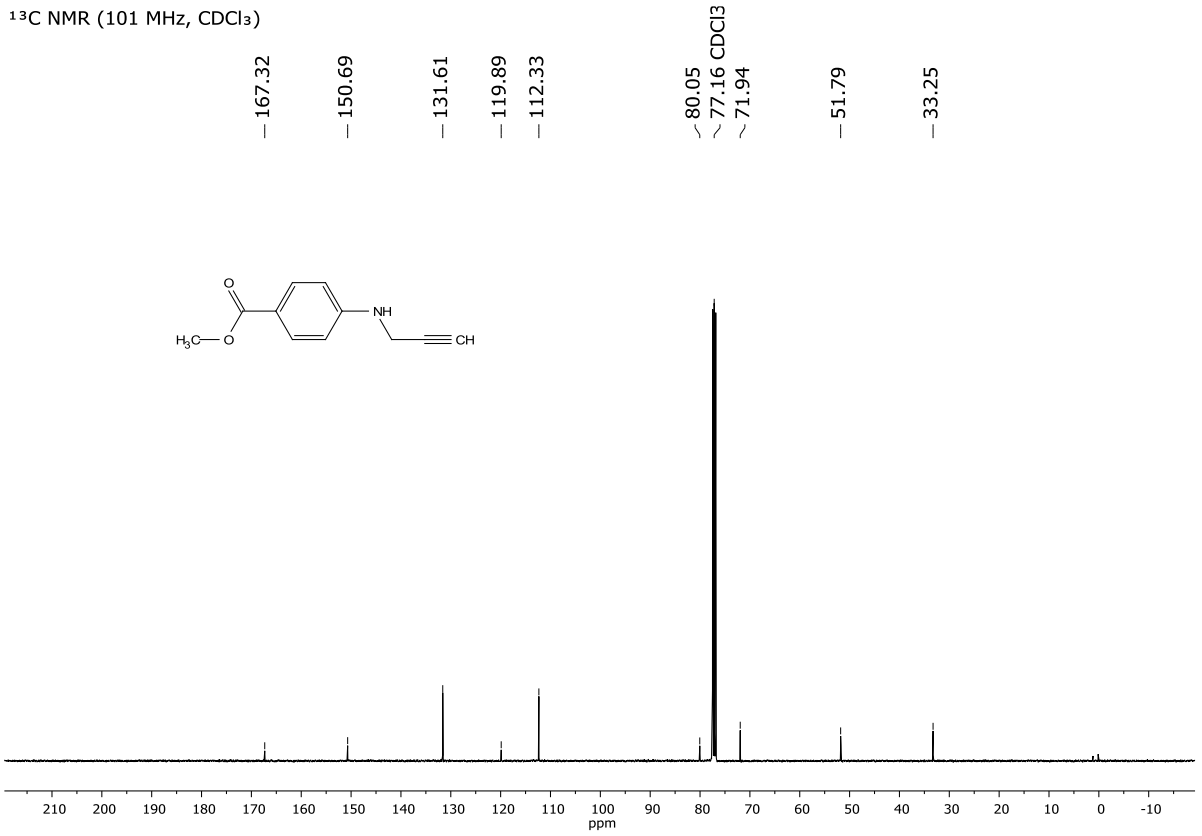
^{1}H NMR:

^{1}H NMR (400 MHz, CDCl$_3$)

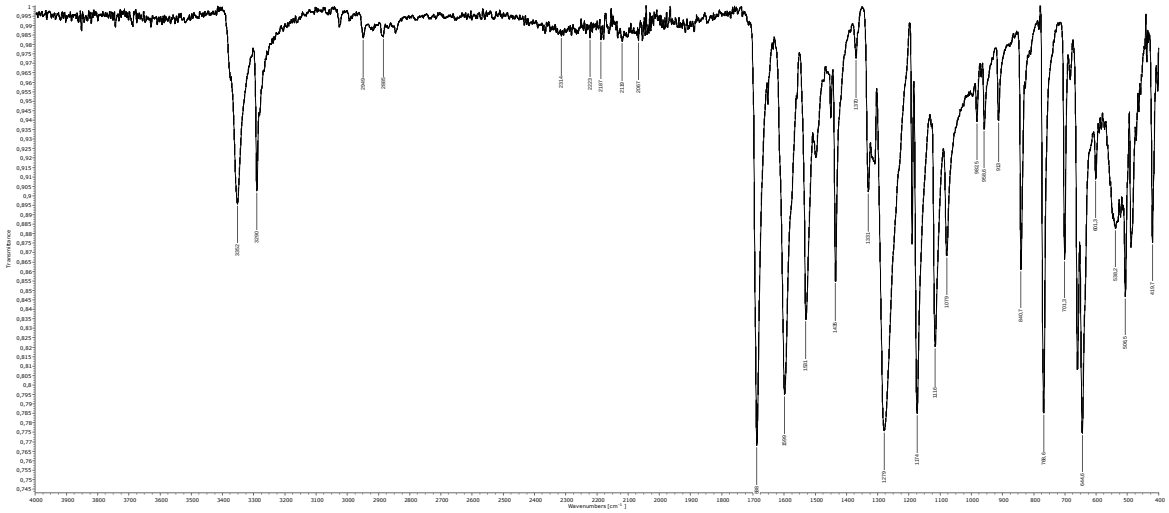


^{13}C NMR:

^{13}C NMR (101 MHz, CDCl$_3$)



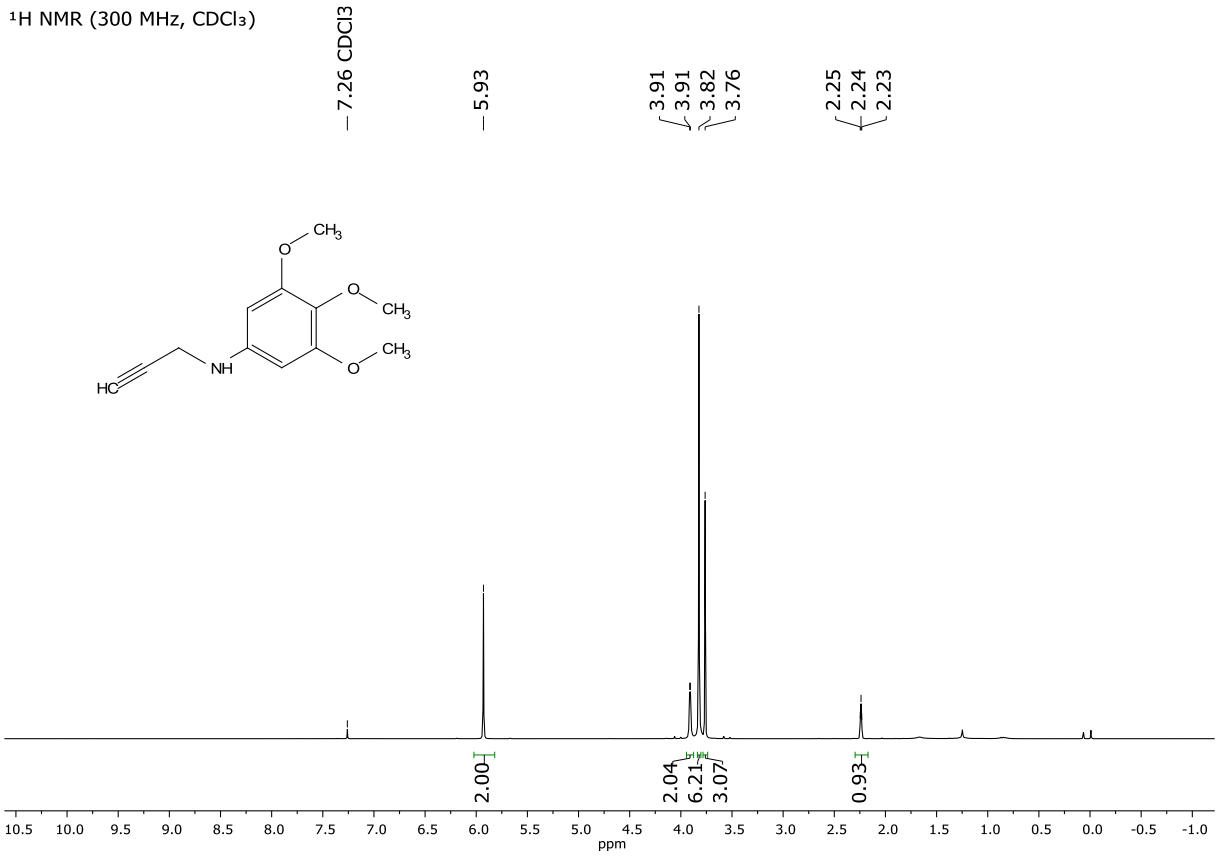
FTIR:



3,4,5-Trimethoxy-*N*-(prop-2-yn-1-yl)aniline (**15b**)

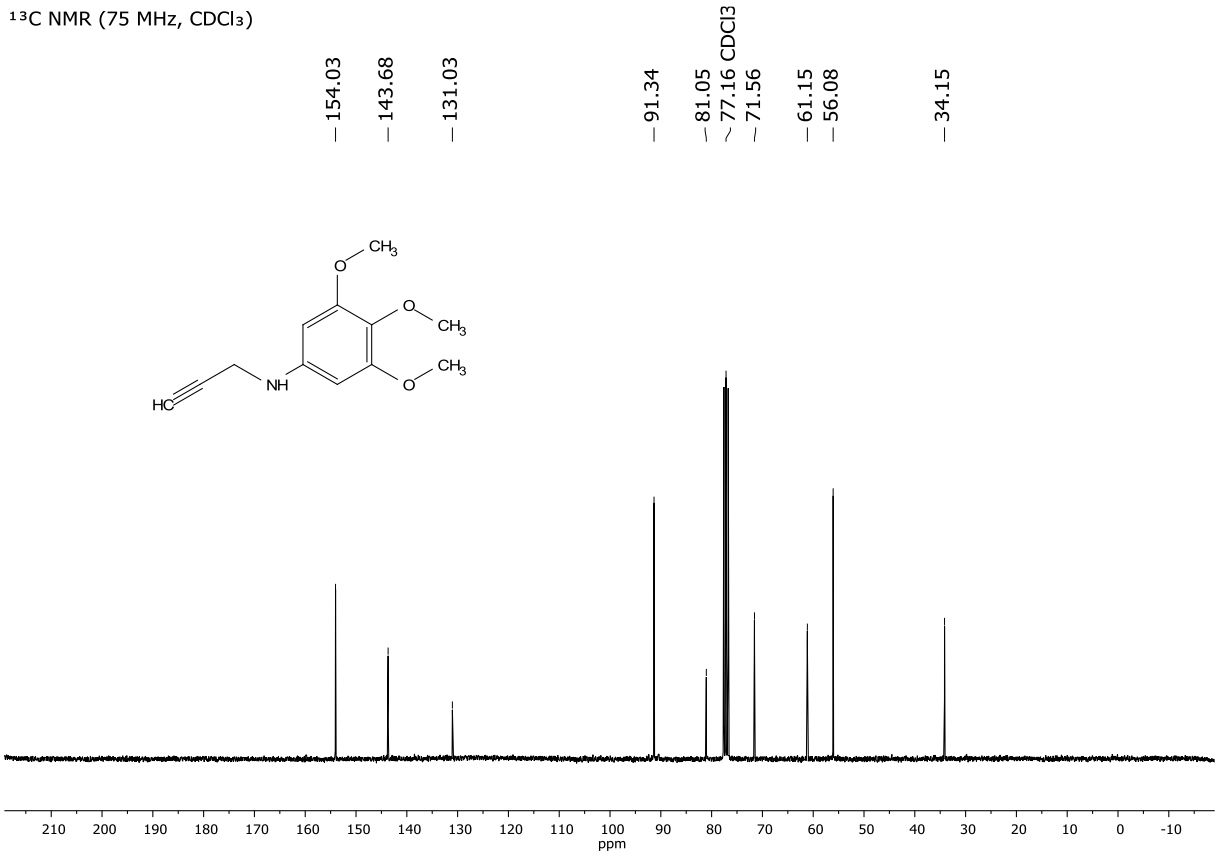
^{1}H NMR:

^{1}H NMR (300 MHz, $CDCl_3$)

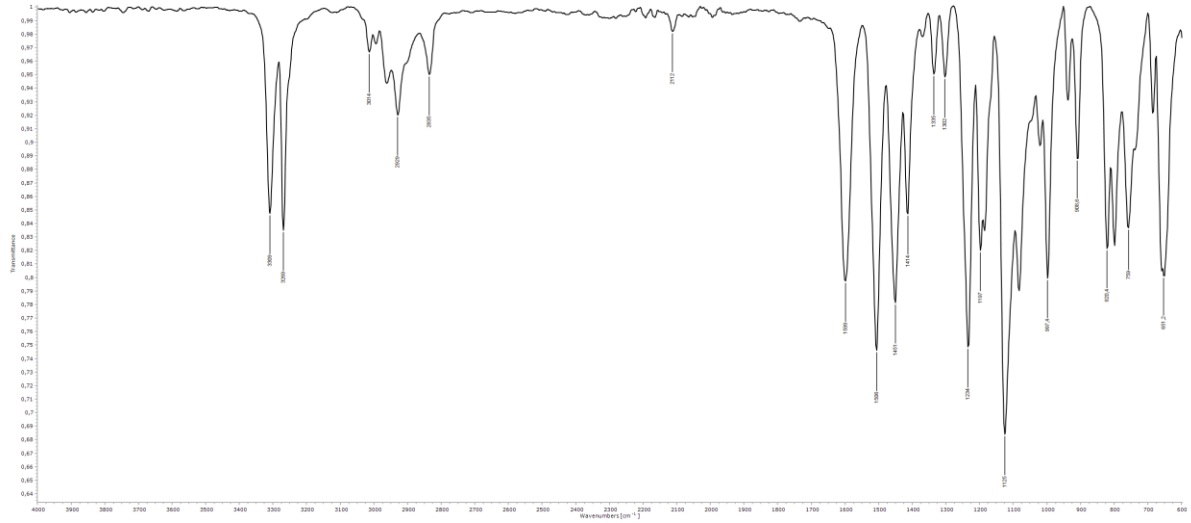


^{13}C NMR:

^{13}C NMR (75 MHz, $CDCl_3$)



FTIR:

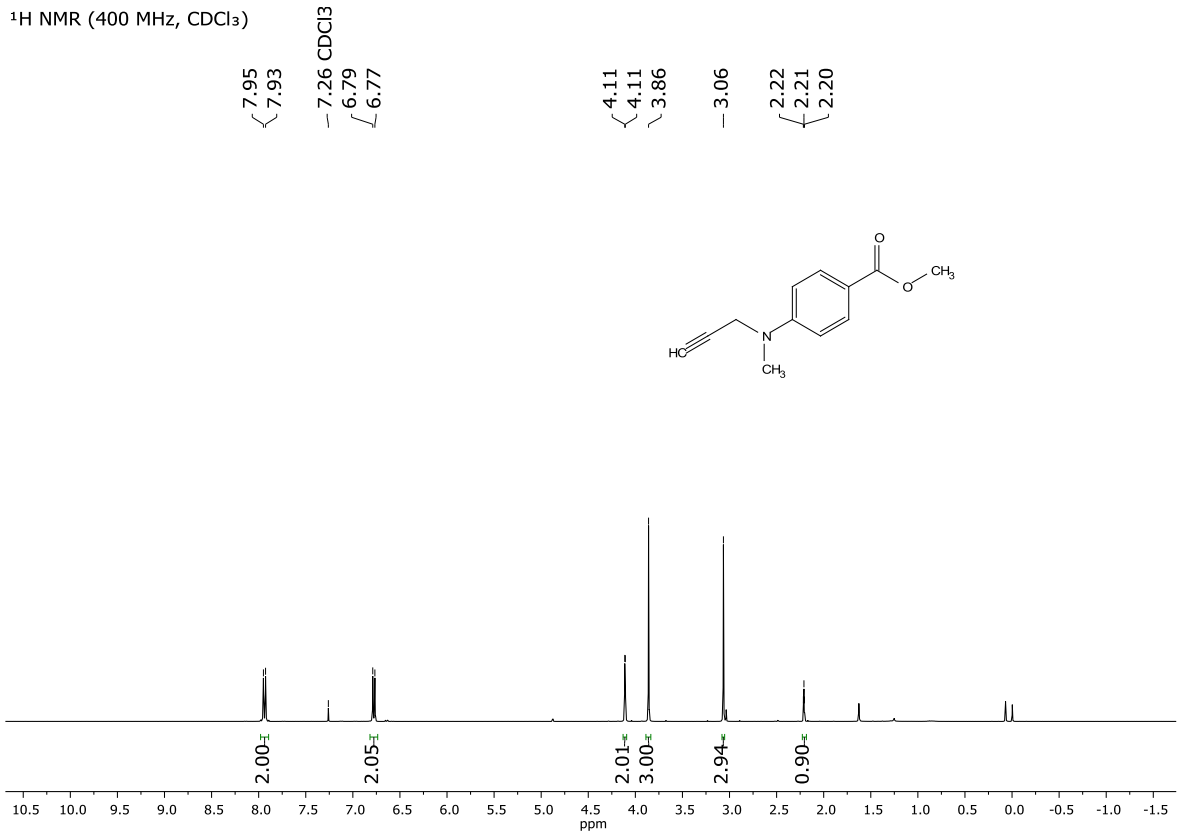


Methylation of alkylated anilines (**16**){1}

Methyl 4-[methyl(prop-2-yn-1-yl)amino]benzoate (**16**)

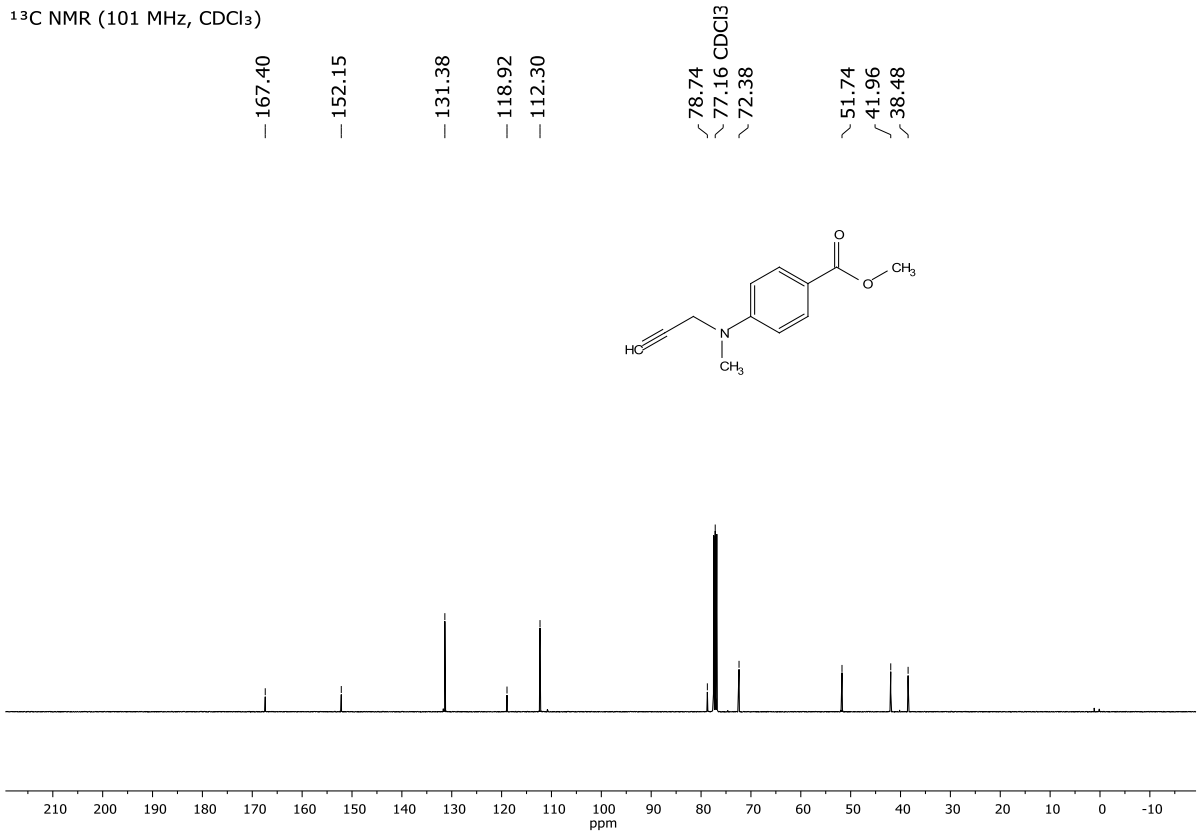
^{1}H NMR:

^{1}H NMR (400 MHz, $CDCl_3$)

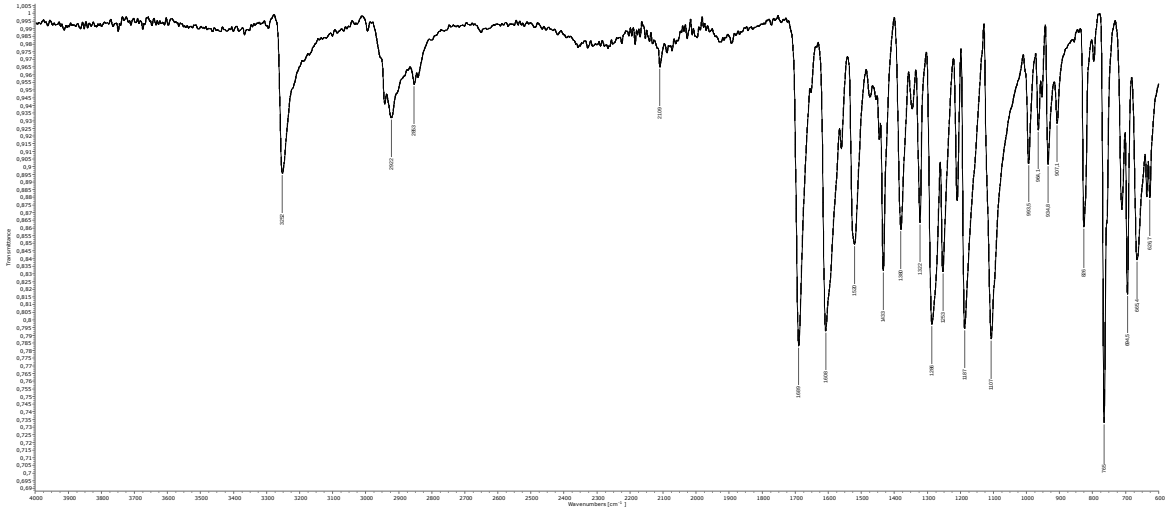


^{13}C NMR:

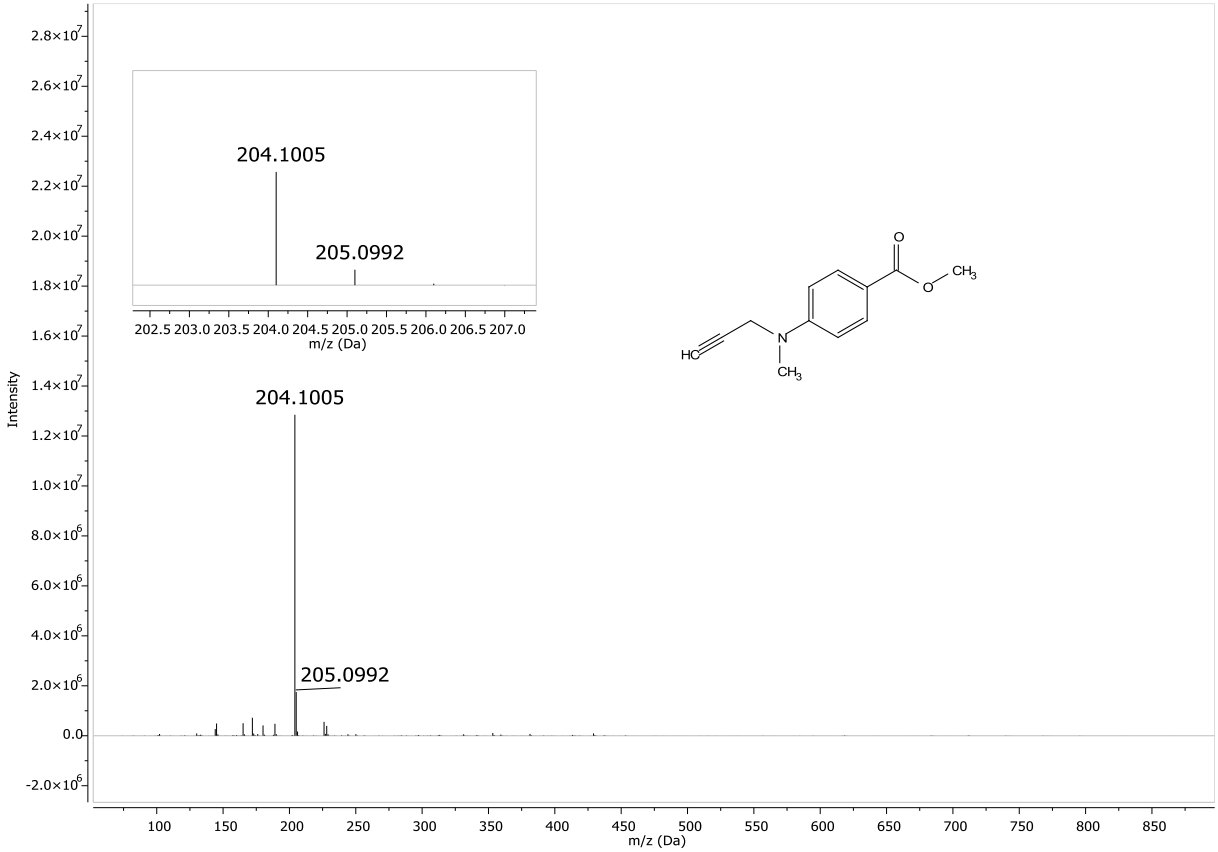
^{13}C NMR (101 MHz, $CDCl_3$)



FTIR:



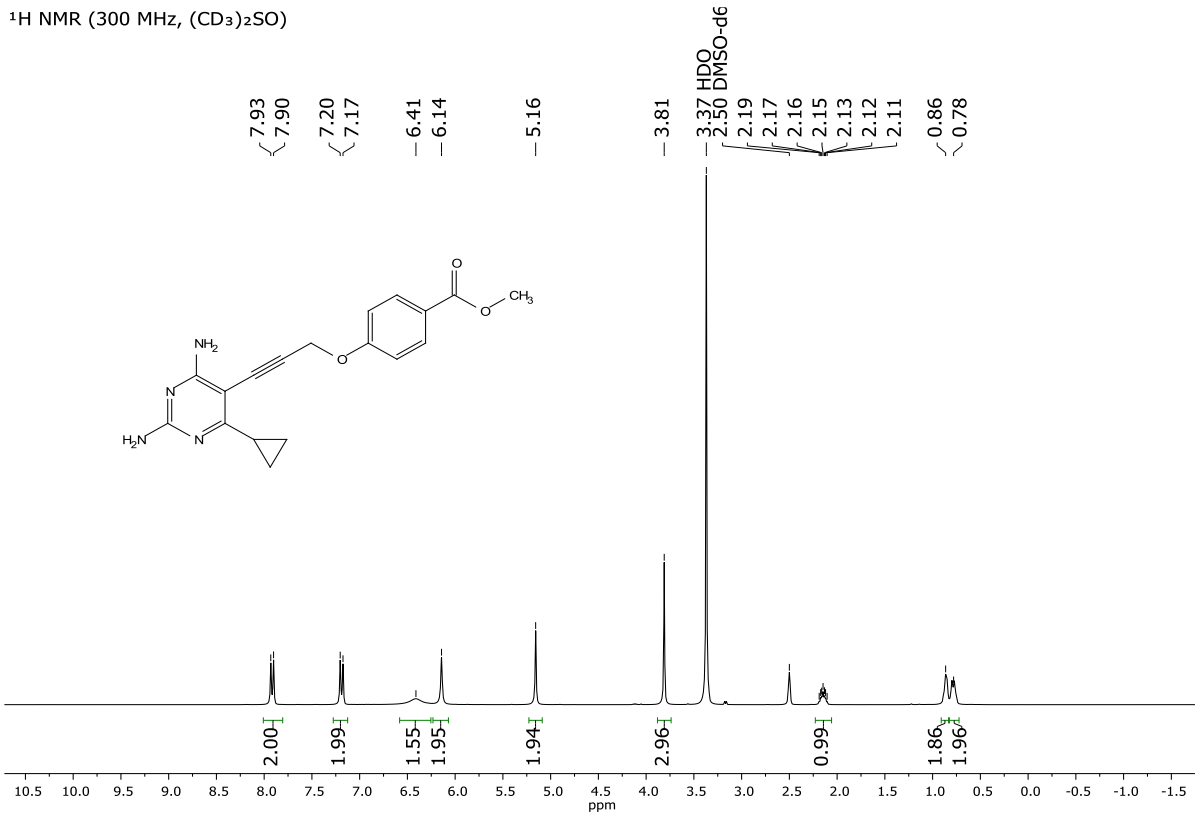
HRMS:


204.1005
205.0992
202.5 203.0 203.5 204.0 204.5 205.0 205.5 206.0 206.5 207.0
m/z (Da)
O
O
CH3
HC
N
CH3
Intensity
204.1005
205.0992
100 150 200 250 300 350 400 450 500 550 600 650 700 750 800 850
m/z (Da)

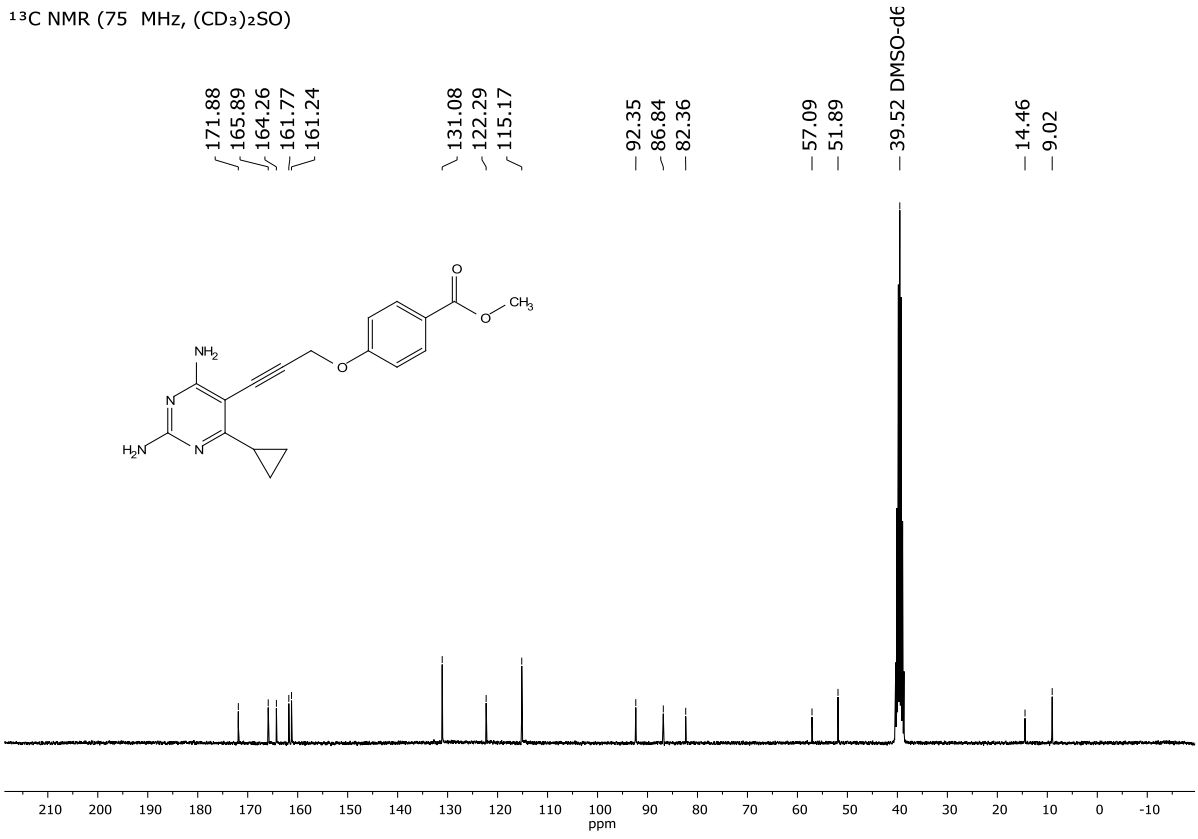
Sonogashira coupling reaction for second-generation analogues (**17**)

Methyl 4-{[3-(2,4-diamino-6-cyclopropylpyrimidin-5-yl)prop-2-yn-1-yl]oxy}benzoate (**17a**)

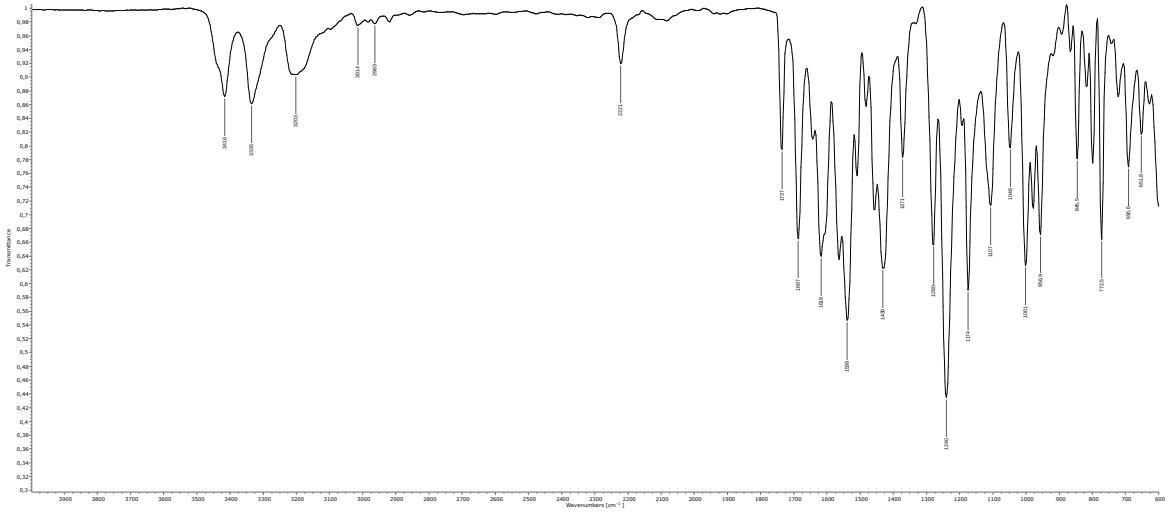
^{1}H NMR:



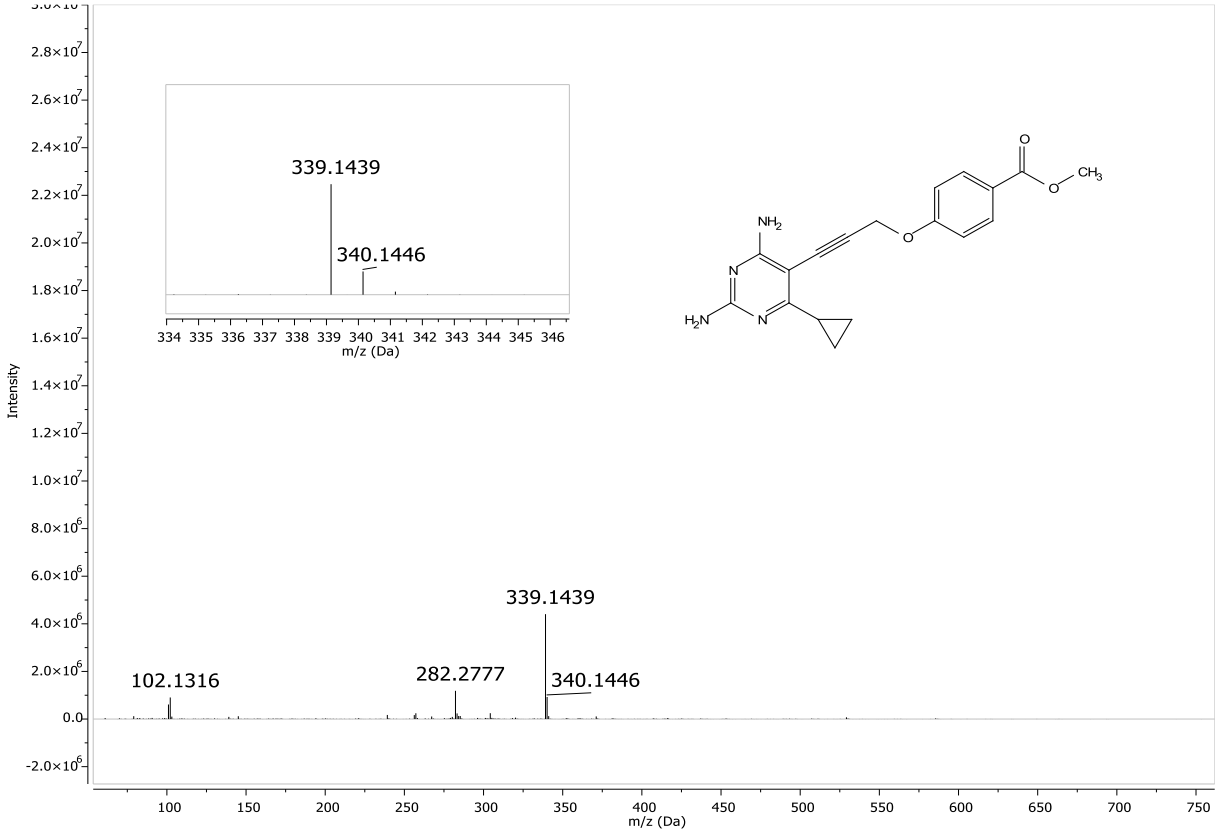
^{13}C NMR:



FTIR:



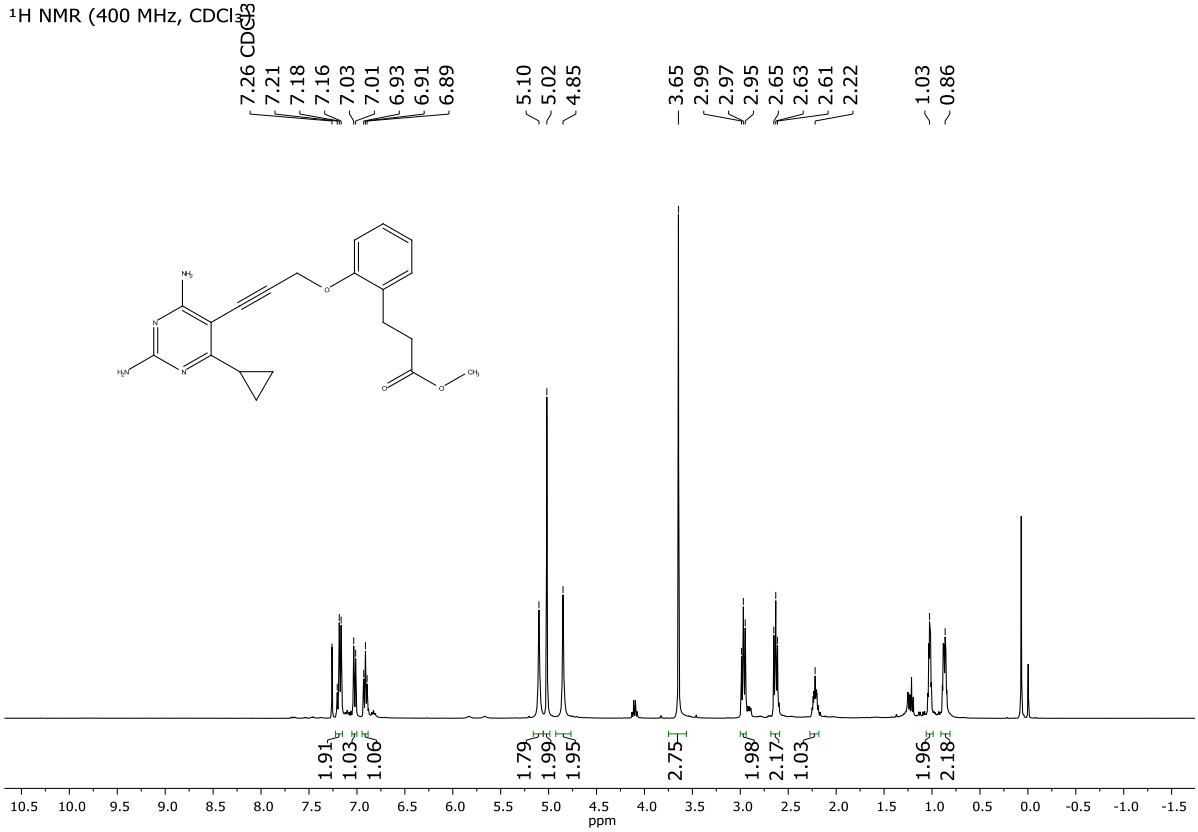
HRMS:


339.1439
340.1446
334 335 336 337 338 339 340 341 342 343 344 345 346
m/z (Da)
NH2
N
O
O
CH3
H2N
N
Intensity
339.1439
102.1316
282.2777
340.1446
100 150 200 250 300 350 400 450 500 550 600 650 700 750
m/z (Da)

Methyl 3-{2-[(3-{2,4-diamino-6-cyclopropylpyrimidin-5-yl} prop-2-yn-1-yl)oxy]phenyl} propanoate (**17b**)

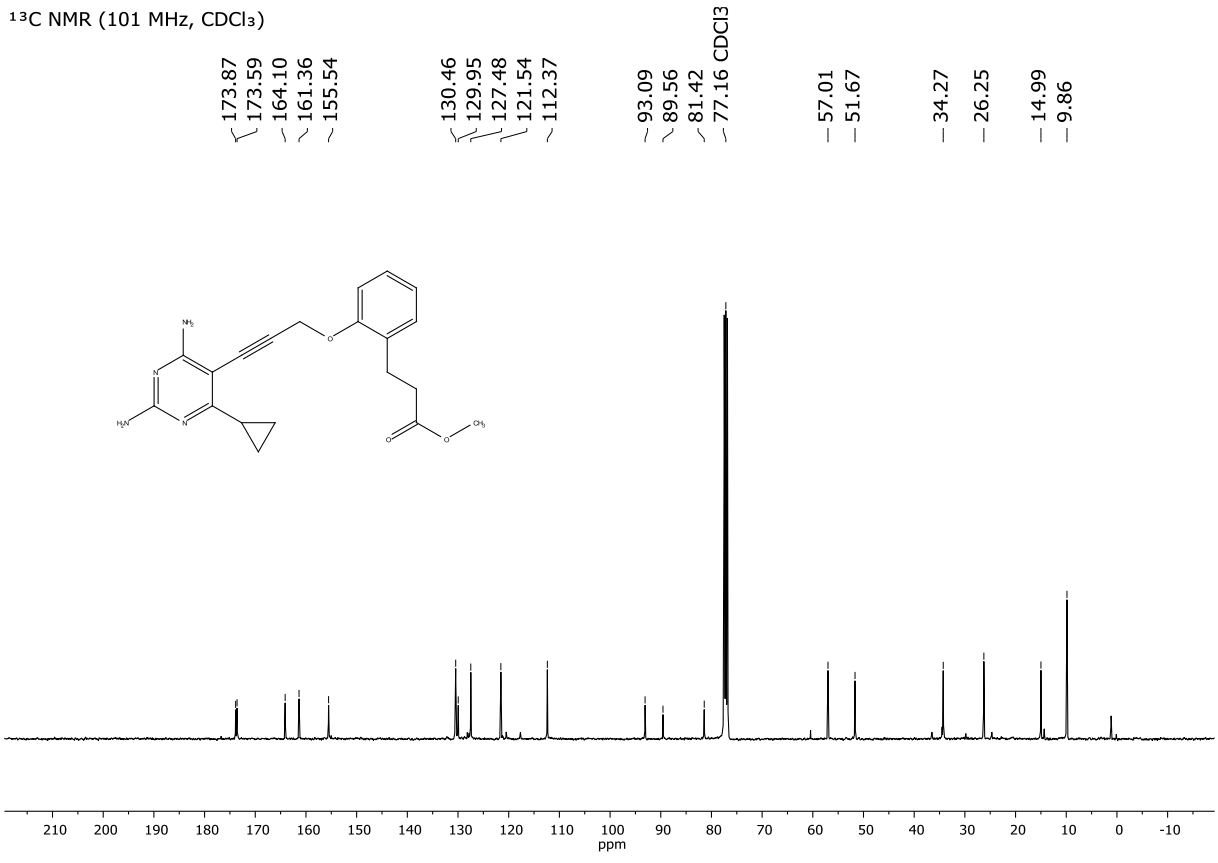
^{1}H NMR:

^{1}H NMR (400 MHz, CDCl$_3$)

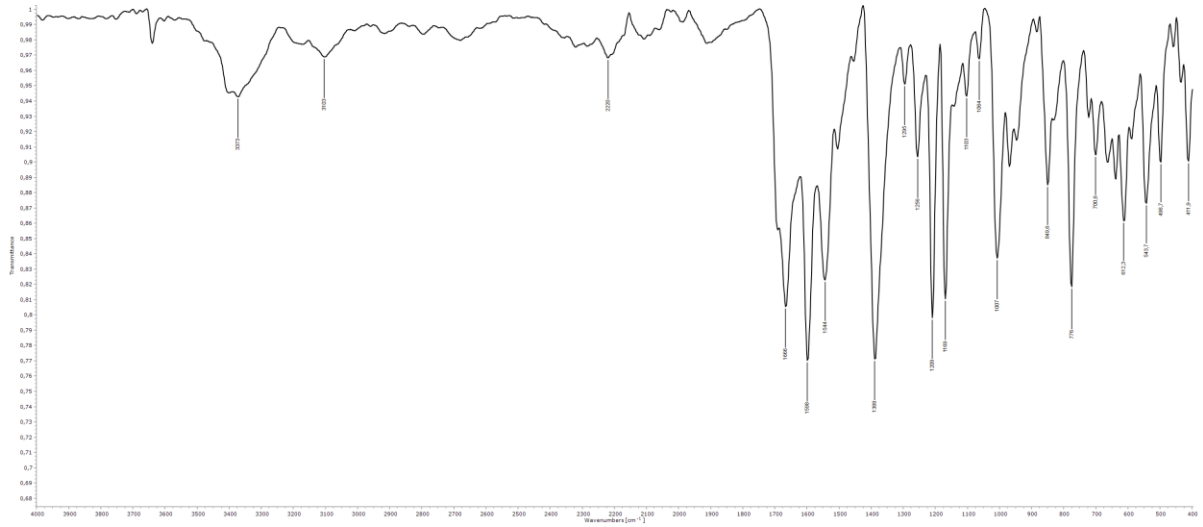


^{13}C NMR:

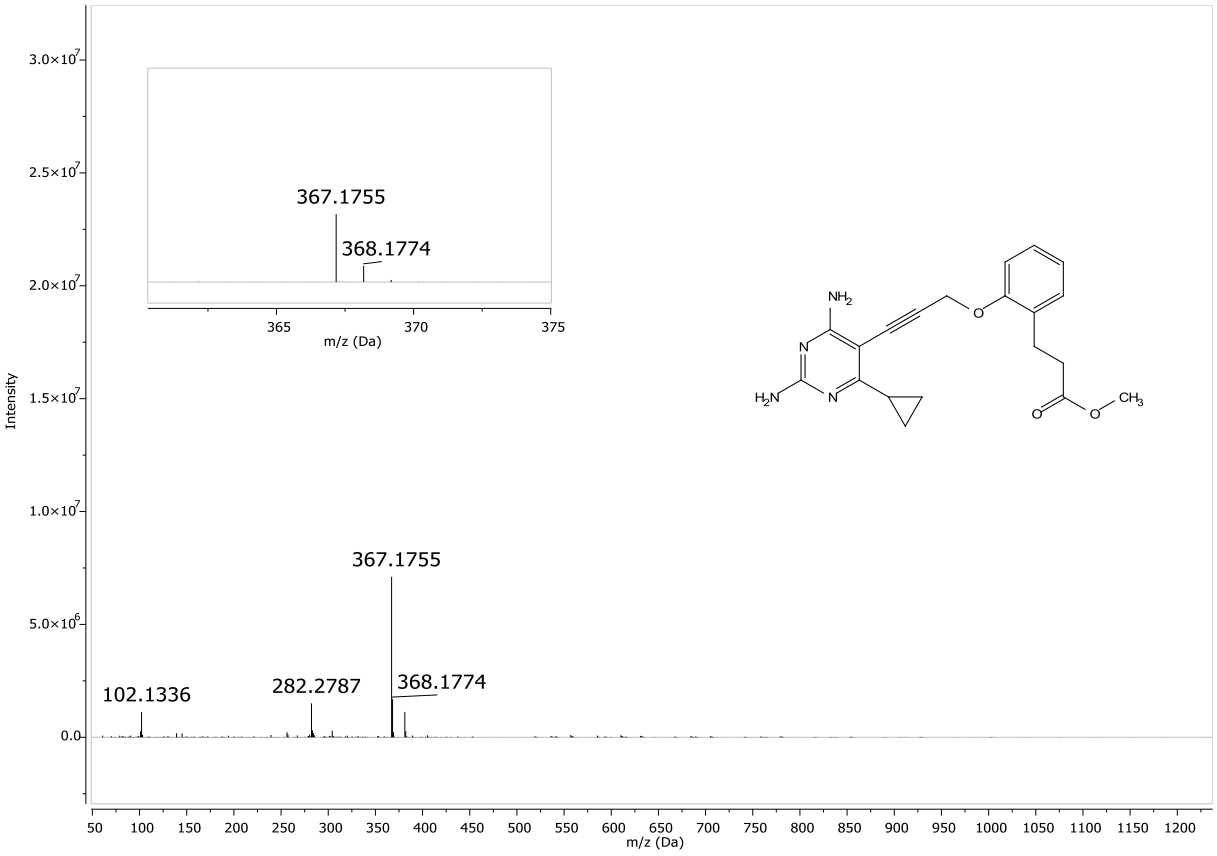
^{13}C NMR (101 MHz, CDCl$_3$)



FTIR:



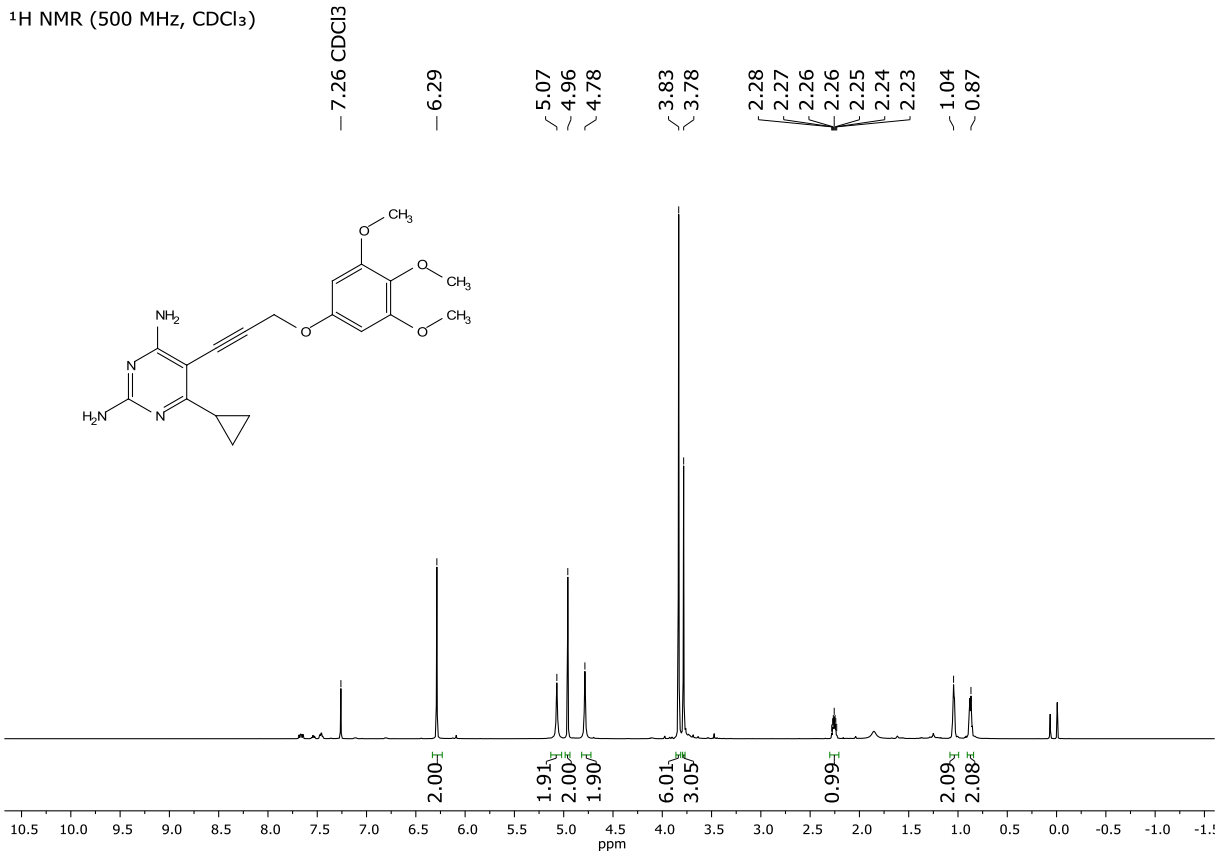
HRMS:


367.1755
368.1774
365
370
375
m/z (Da)
NH2
N
O
H2N
N
O
O
CH3
Intensity
3.0×10^7
2.5×10^7
2.0×10^7
1.5×10^7
1.0×10^7
5.0×10^6
0.0
367.1755
102.1336
282.2787
368.1774
50
100
150
200
250
300
350
400
450
500
550
600
650
700
750
800
850
900
950
1000
1050
1100
1150
1200
m/z (Da)

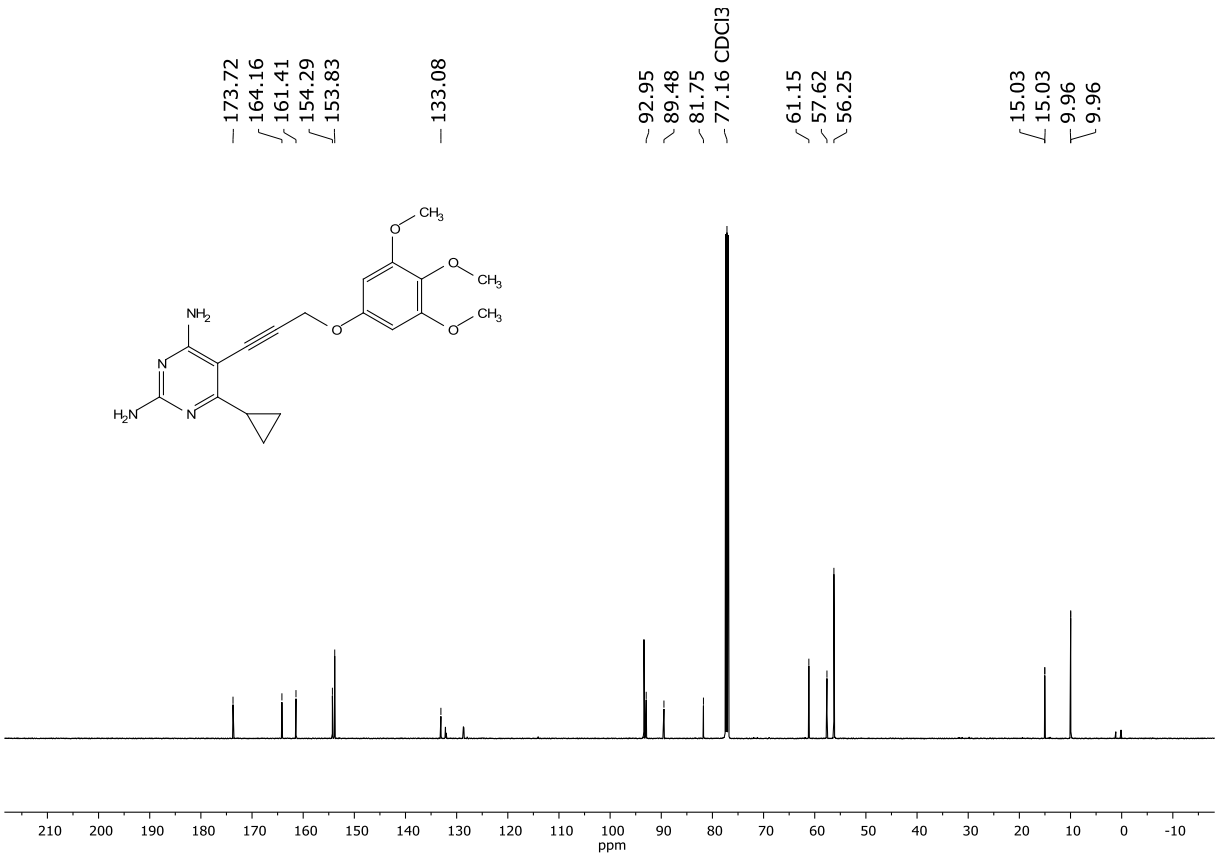
6-Cyclopropyl-5-[3-(3,4,5-trimethoxyphenoxy)prop-1-yn-1-yl]pyrimidine-2,4-diamine (**17c**)

^{1}H NMR:

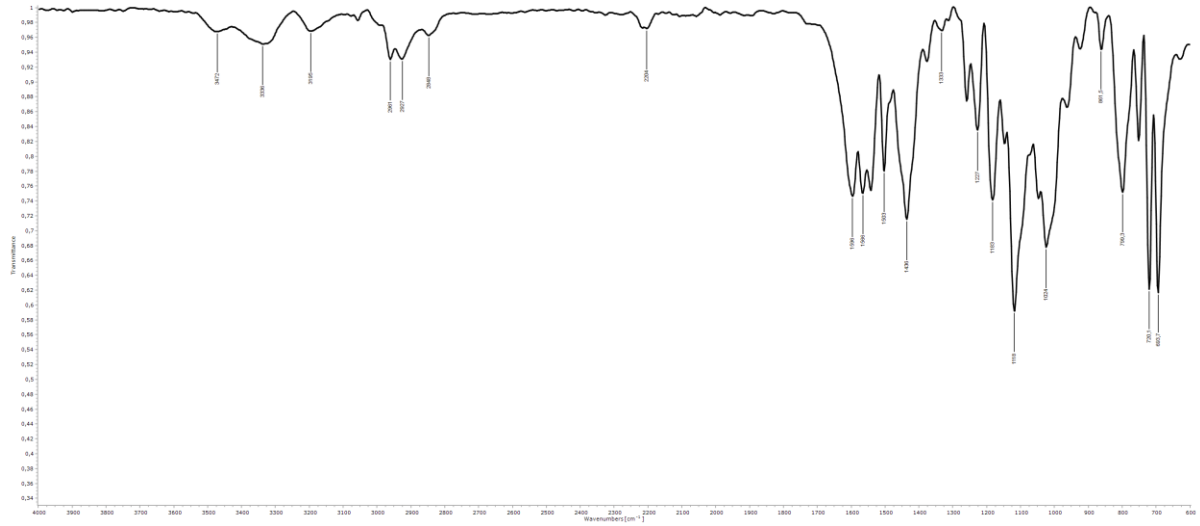
^{1}H NMR (500 MHz, $CDCl_3$)



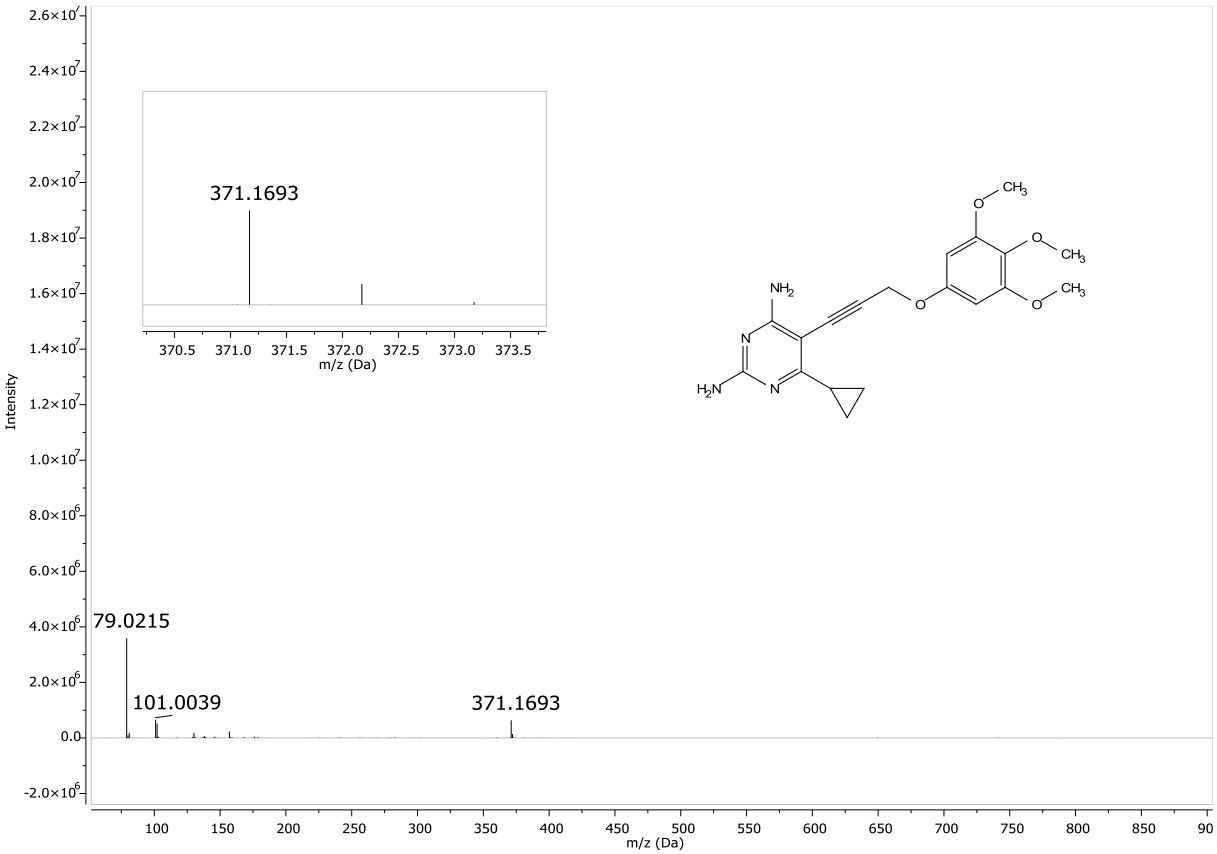
^{13}C NMR:



FTIR:

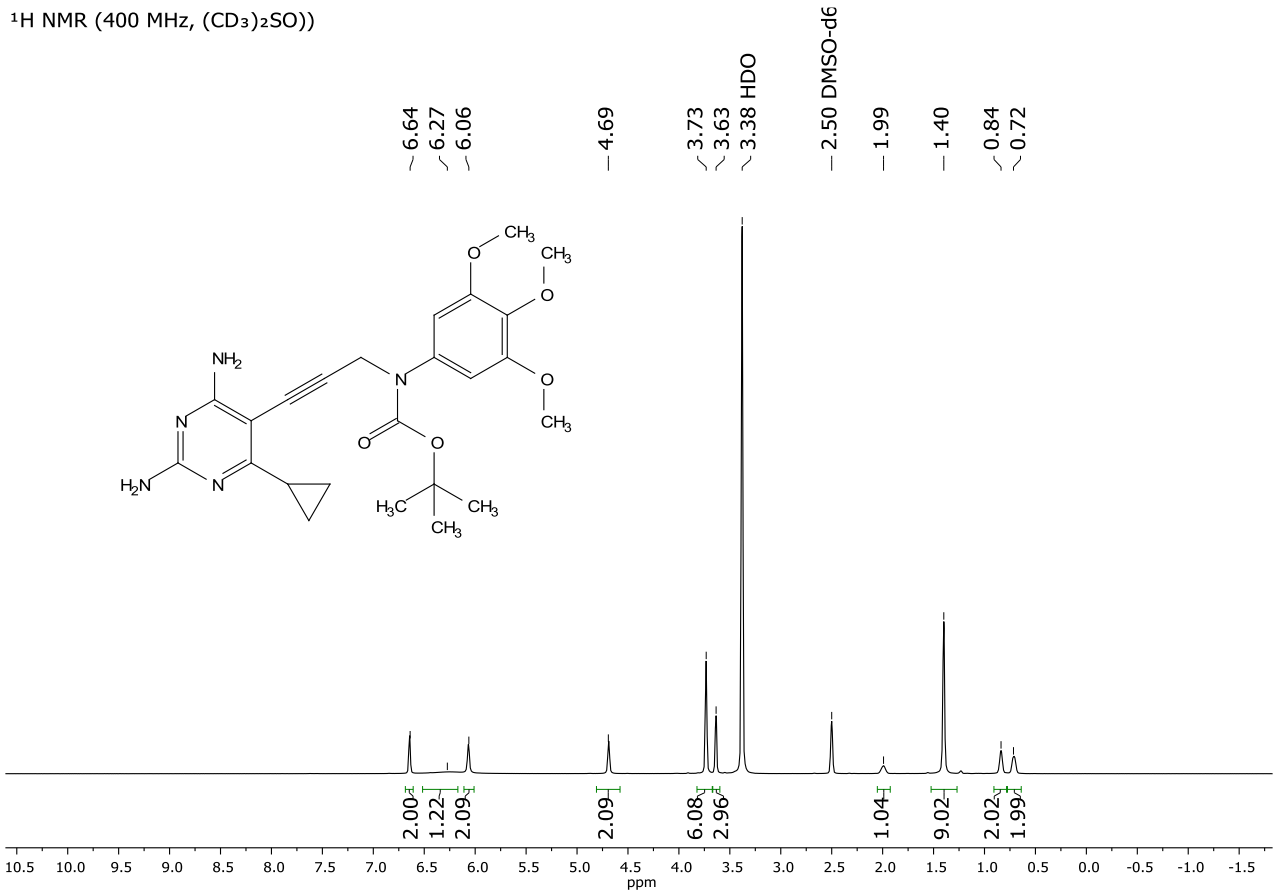


HRMS:


371.1693
370.5 371.0 371.5 372.0 372.5 373.0 373.5
m/z (Da)
Intensity
79.0215
101.0039
371.1693
m/z (Da)

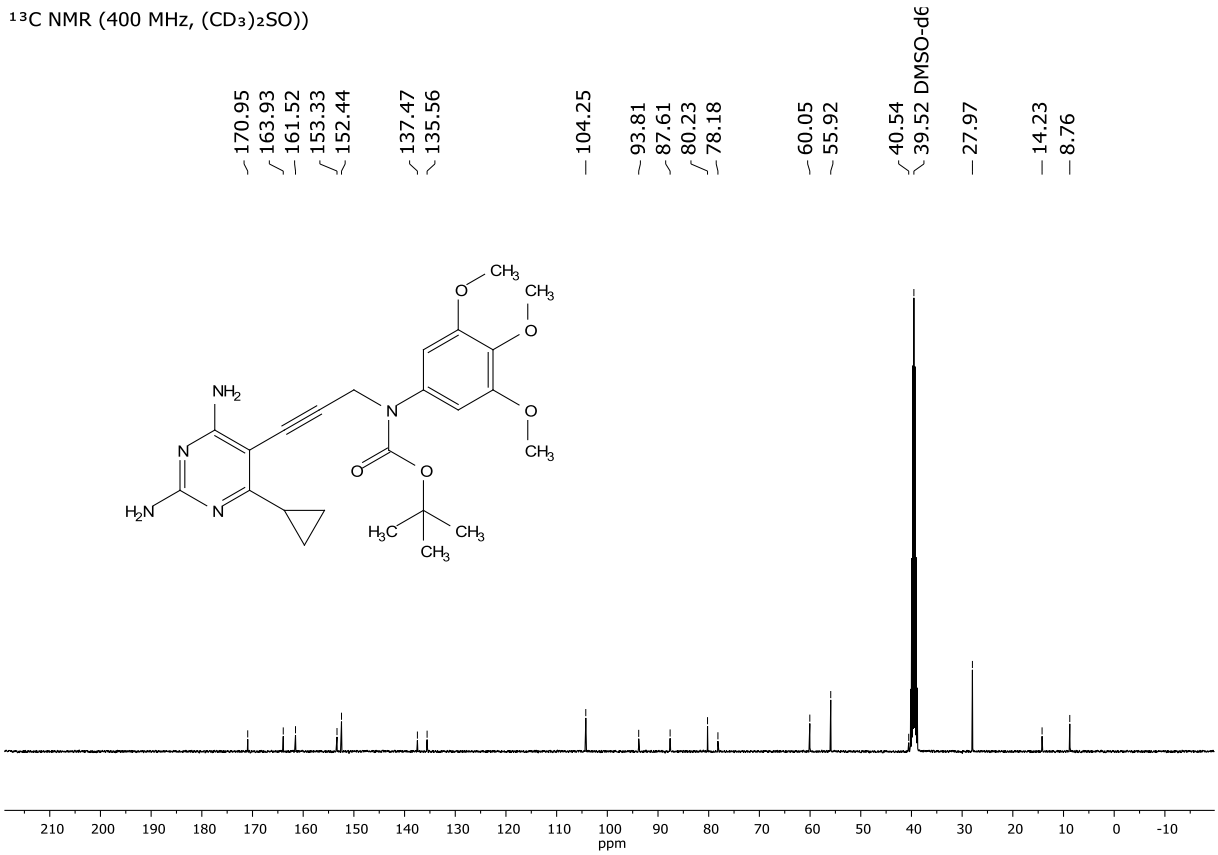
Methyl 4-[(*tert*-butoxycarbonyl)[3-(2,4-diamino-6-cyclopropylpyrimidin-5-yl)prop-2-yn-1-yl]amino]benzoate (**17d**)

^{1}H NMR:

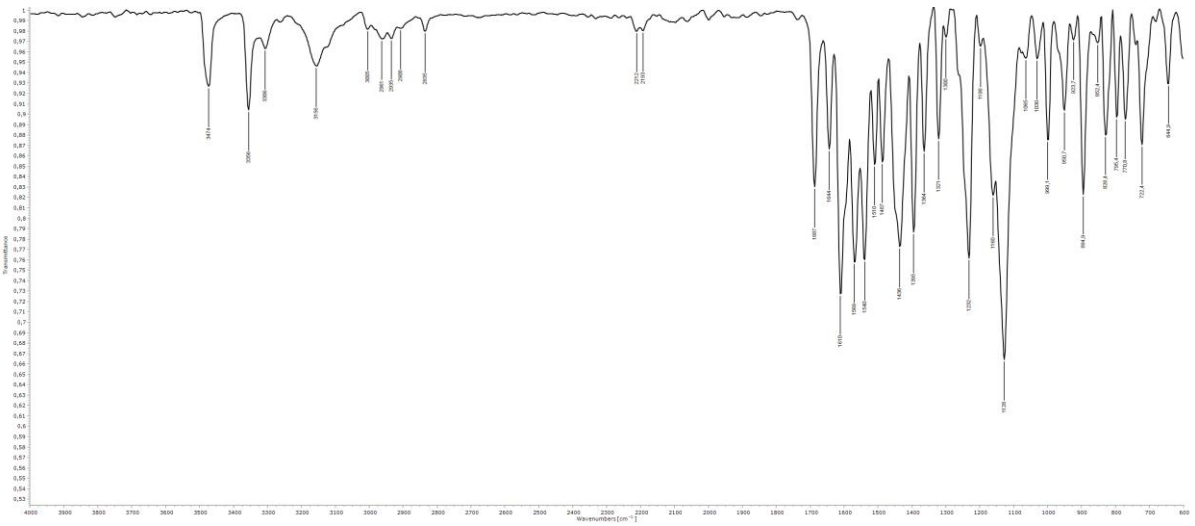


^{13}C NMR:

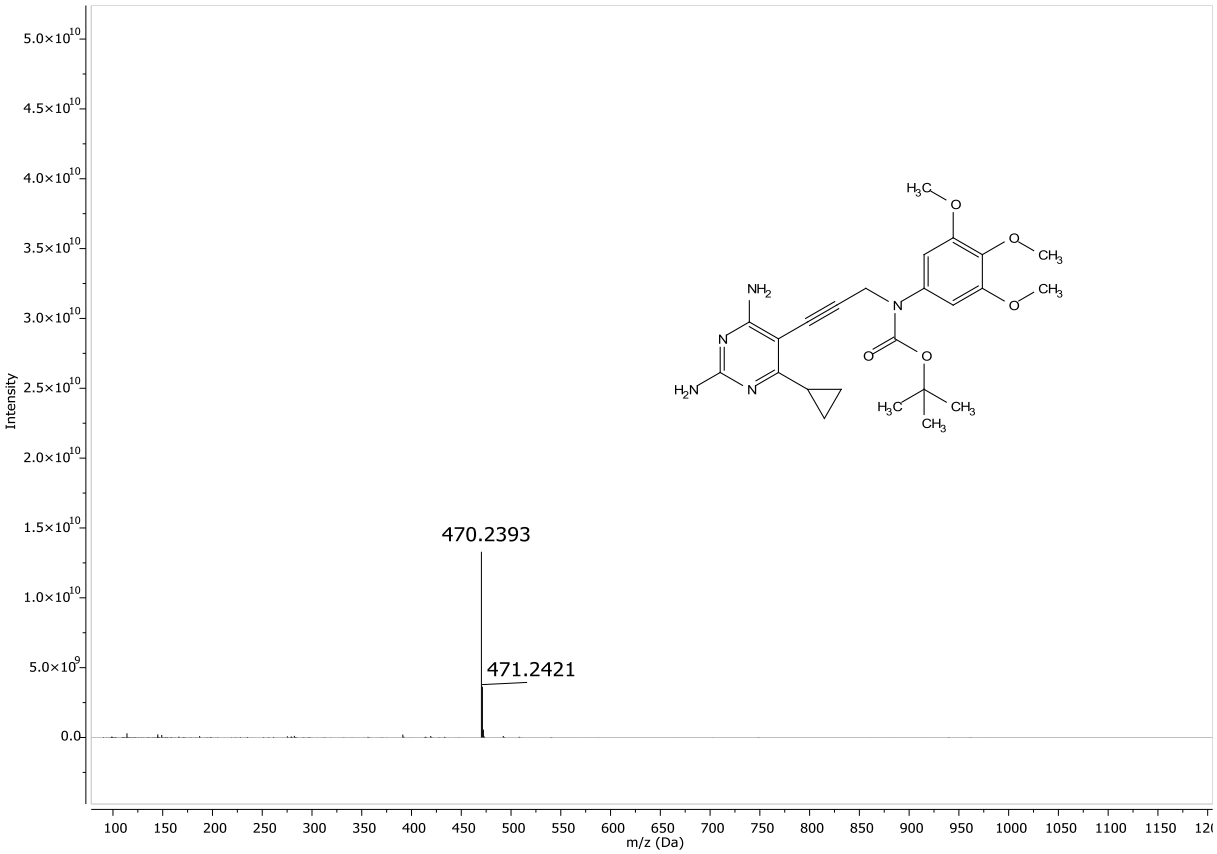
^{13}C NMR (400 MHz, $(CD_3)_2SO$))



FTIR:



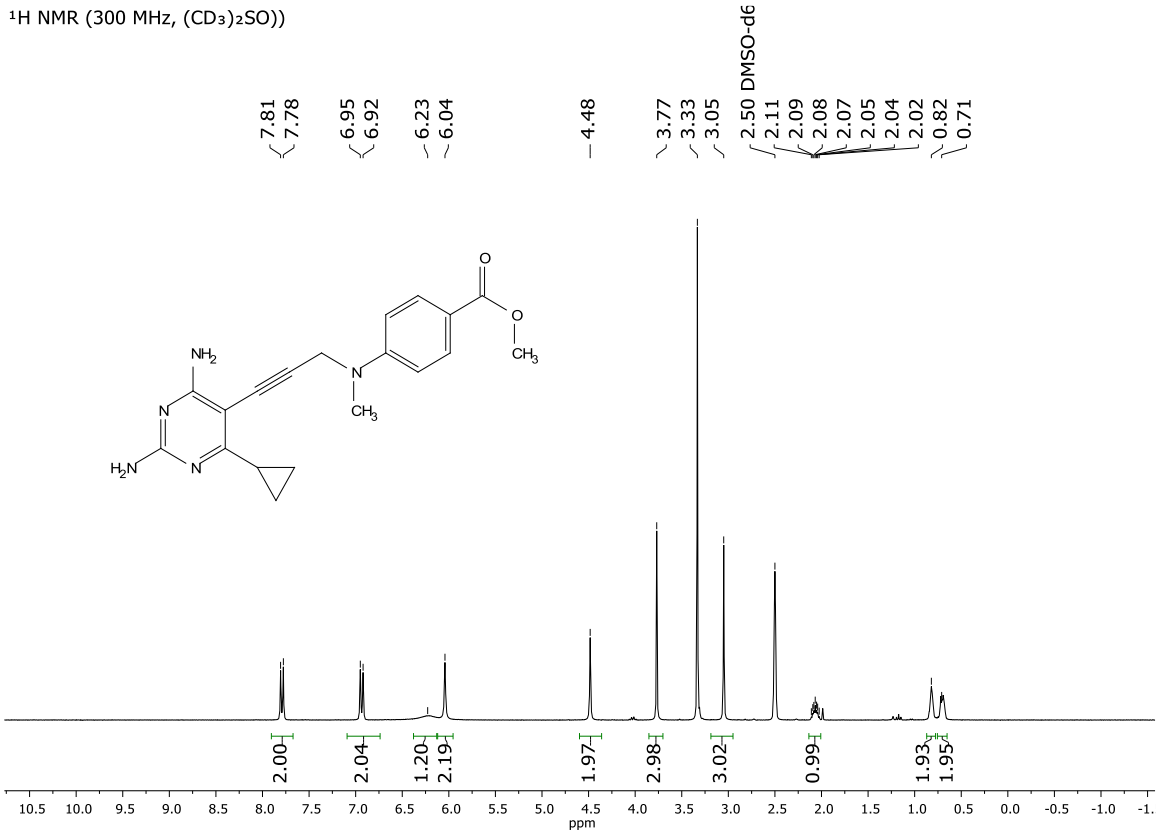
HRMS:



Methyl 4-{[3-(2,4-diamino-6-cyclopropylpyrimidin-5-yl)prop-2-yn-1-yl](methyl)amino} benzoate (**17e**)

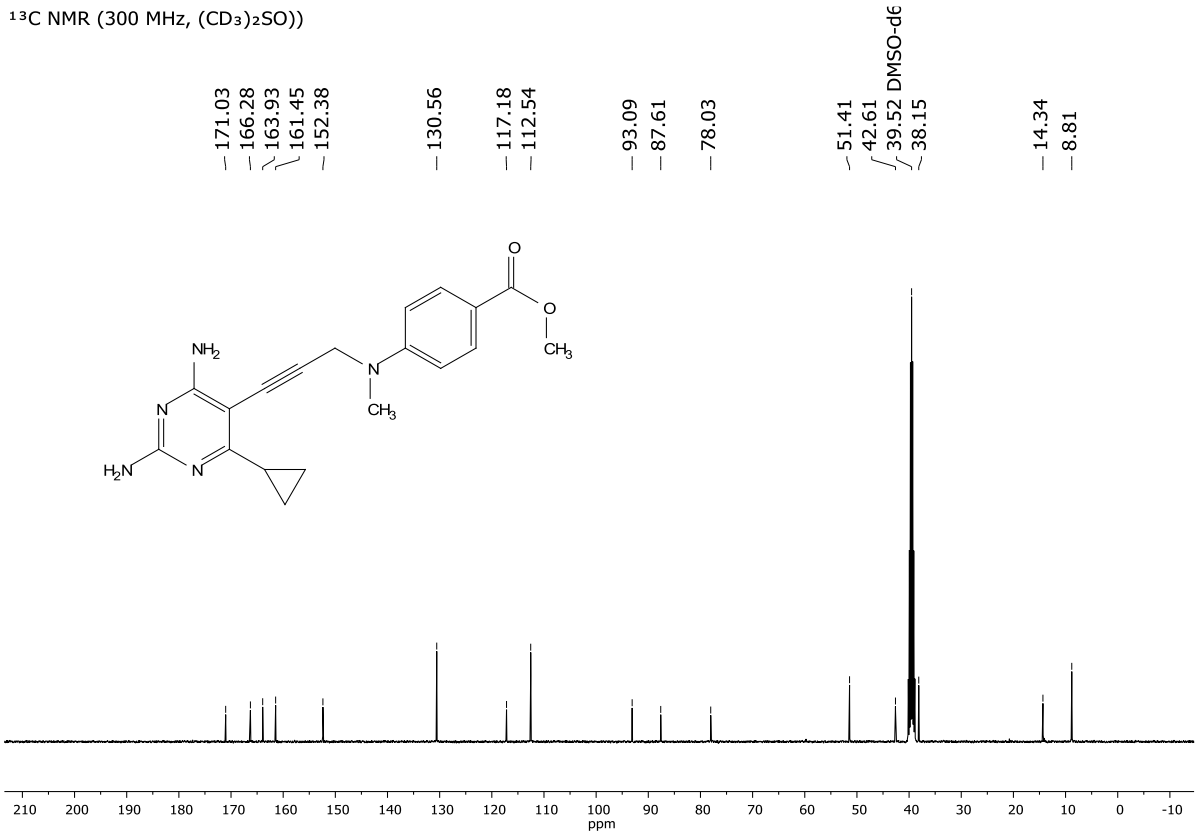
^{1}H NMR:

^{1}H NMR (300 MHz, $(CD_3)_2SO$))

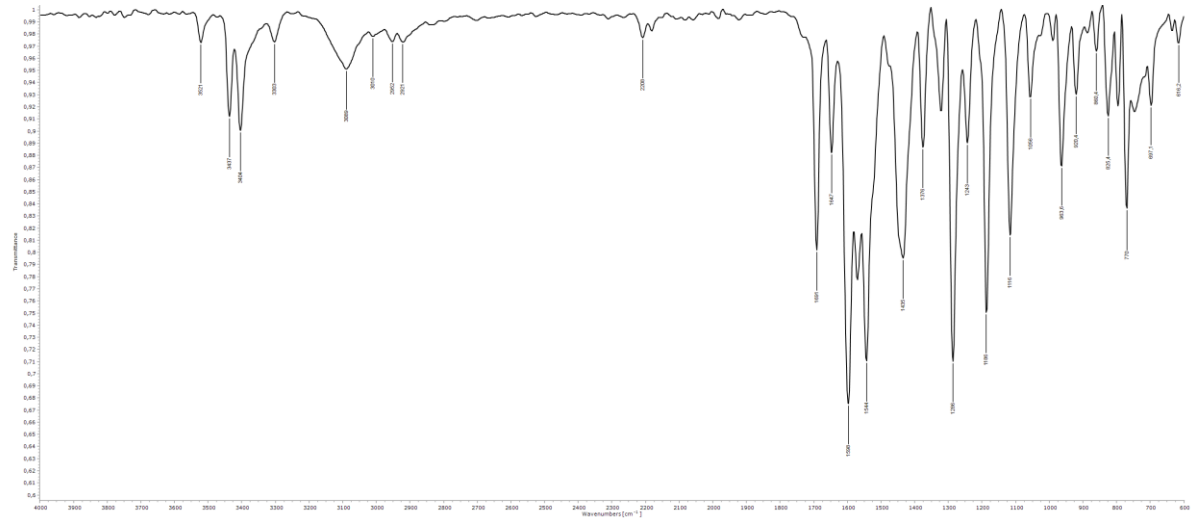


^{13}C NMR:

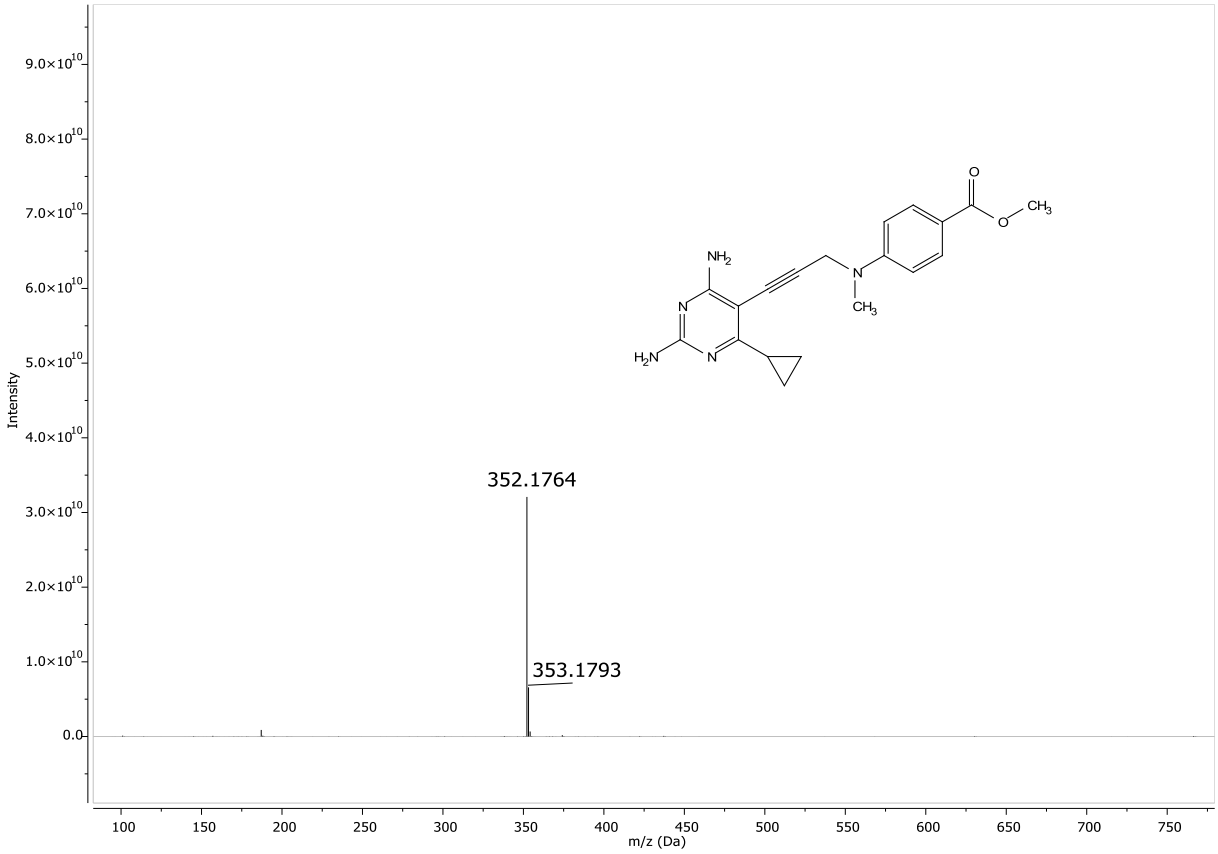
^{13}C NMR (300 MHz, $(CD_3)_2SO$))



FTIR:



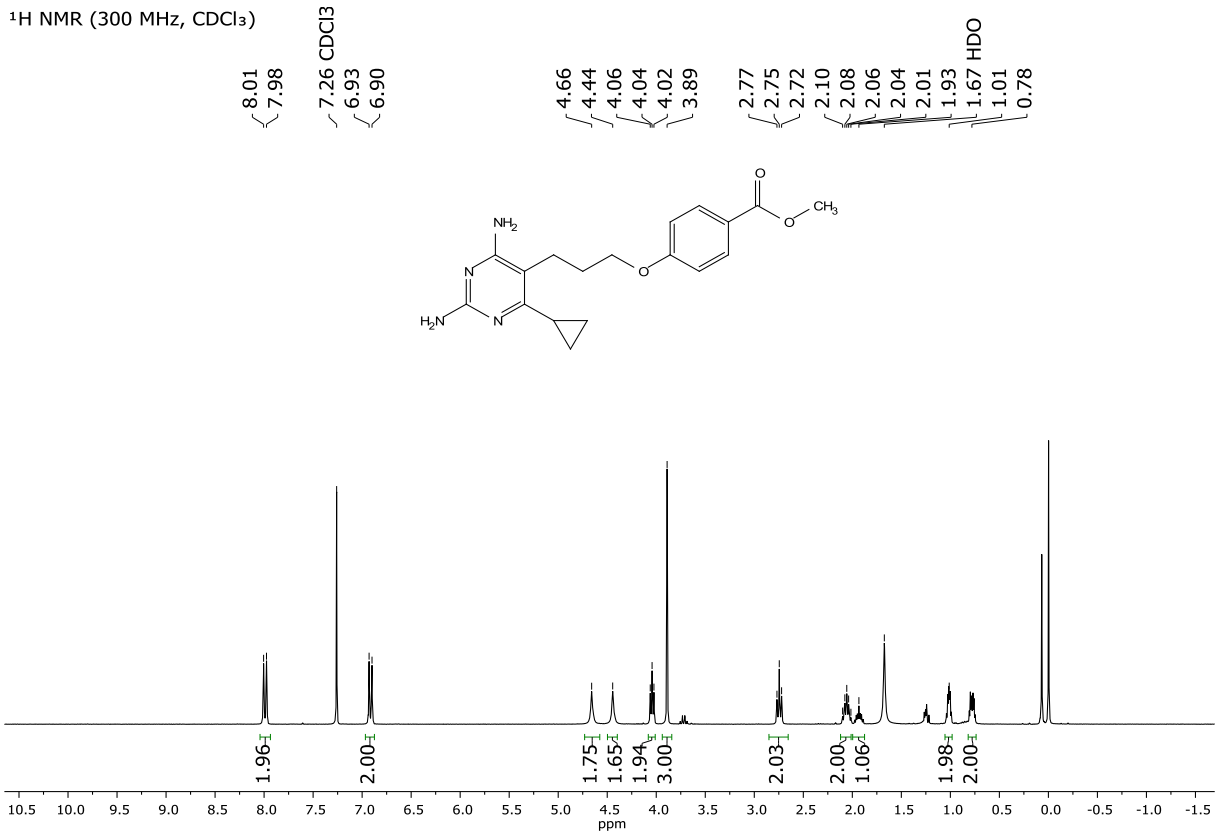
HRMS:



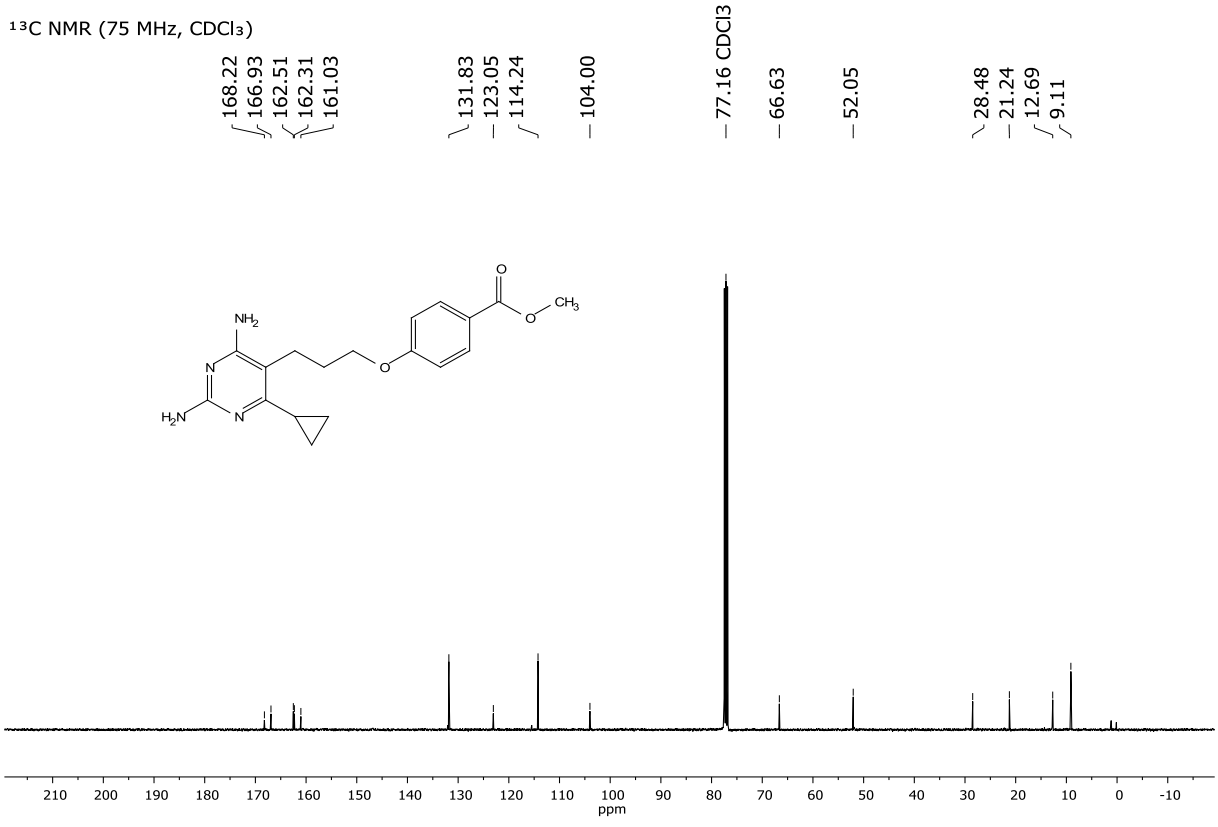
Hydrogenation of second-generation alkynes (**18**)

Methyl 4-[3-(2,4-diamino-6-cyclopropylpyrimidin-5-yl)propoxy]benzoate (**18a**)

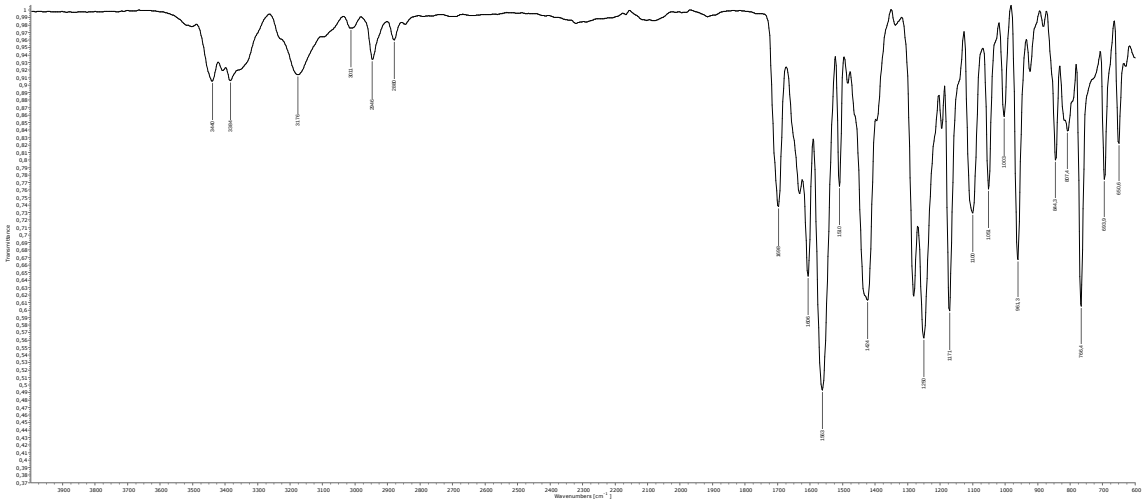
^{1}H NMR:



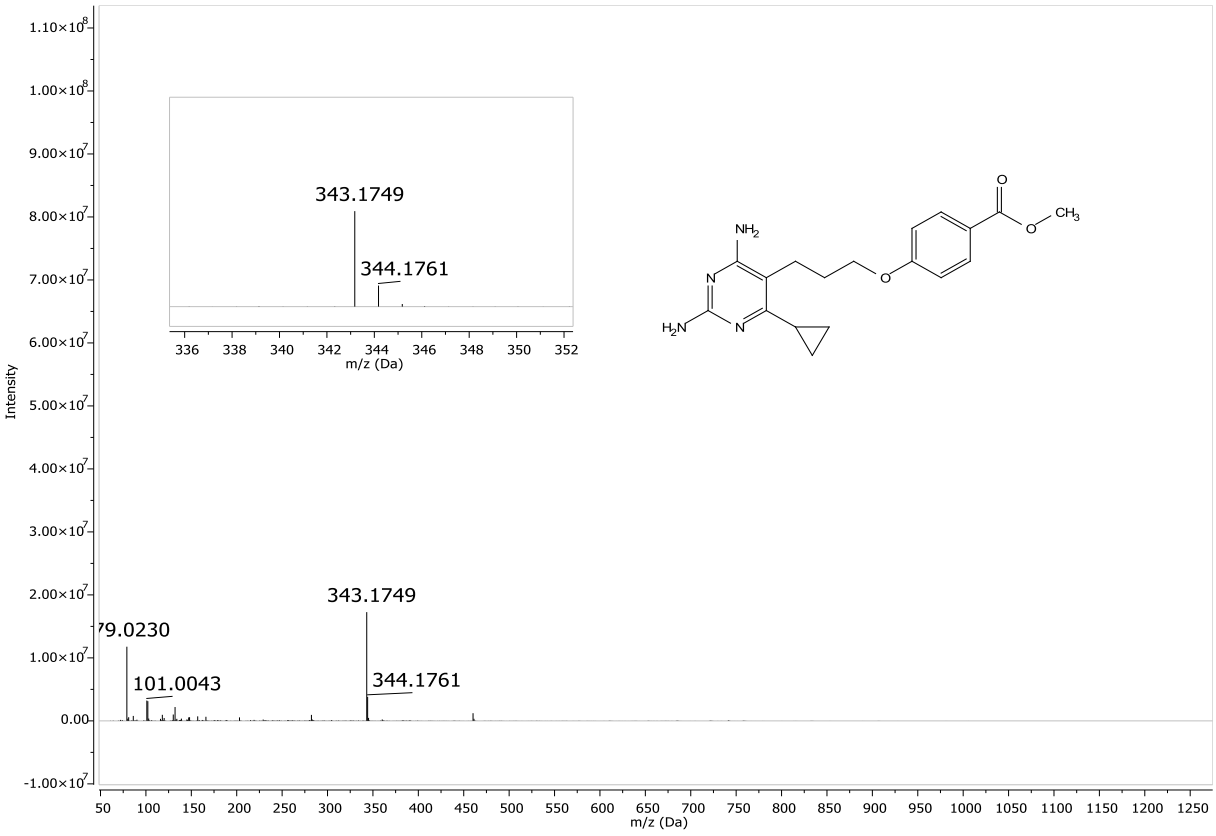
^{13}C NMR:



FTIR:

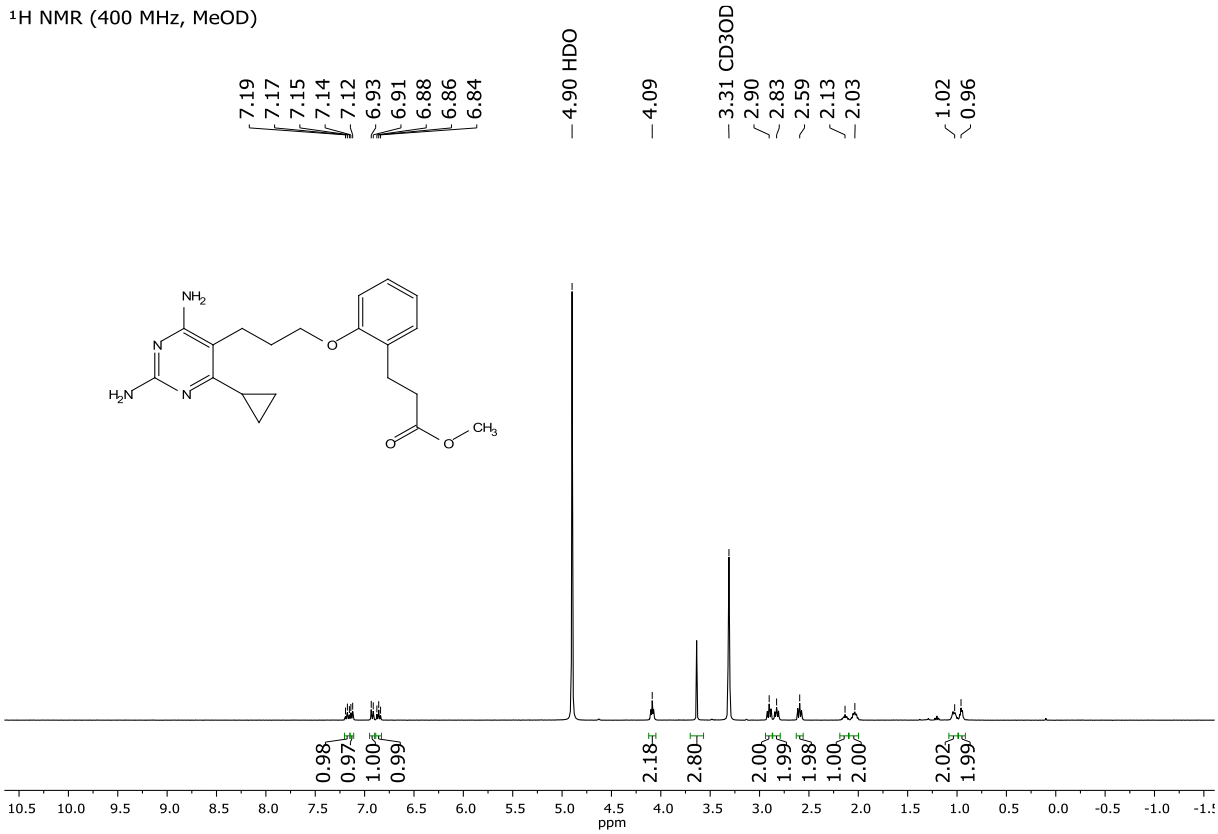


HRMS:

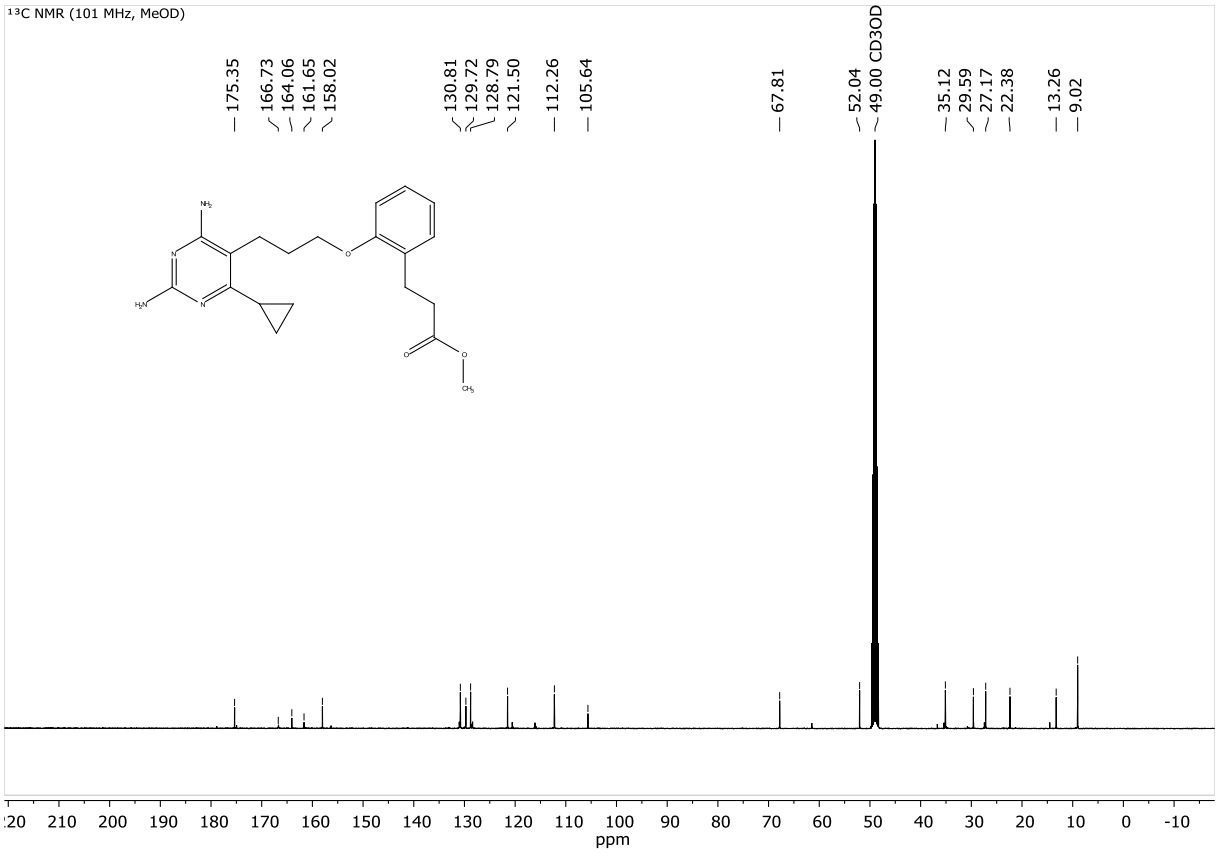

343.1749
344.1761
m/z (Da)
NH2
N
H2N
N
O
O
O
CH3
Intensity
343.1749
79.0230
101.0043
344.1761
m/z (Da)

Methyl-3-{2-[3-(2,4-diamino-6-cyclopropylpyrimidin-5-yl)propoxy]phenyl} propanoatebenzoate (**18b**)

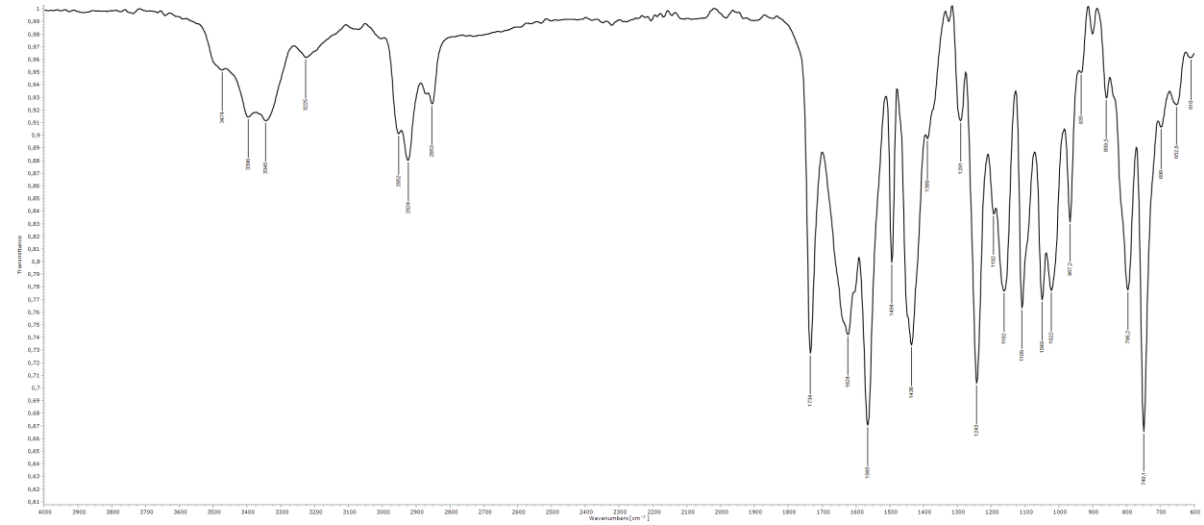
^{1}H NMR:



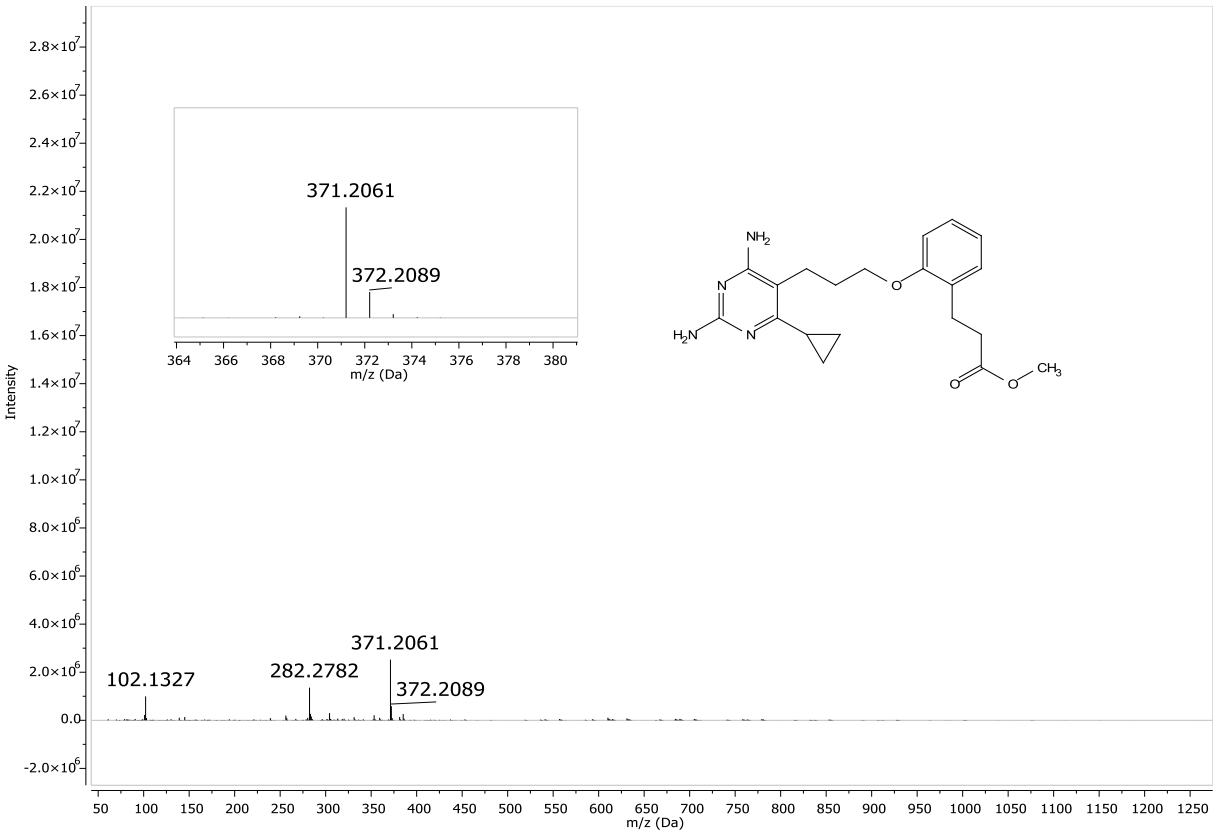
^{13}C NMR:



FTIR:



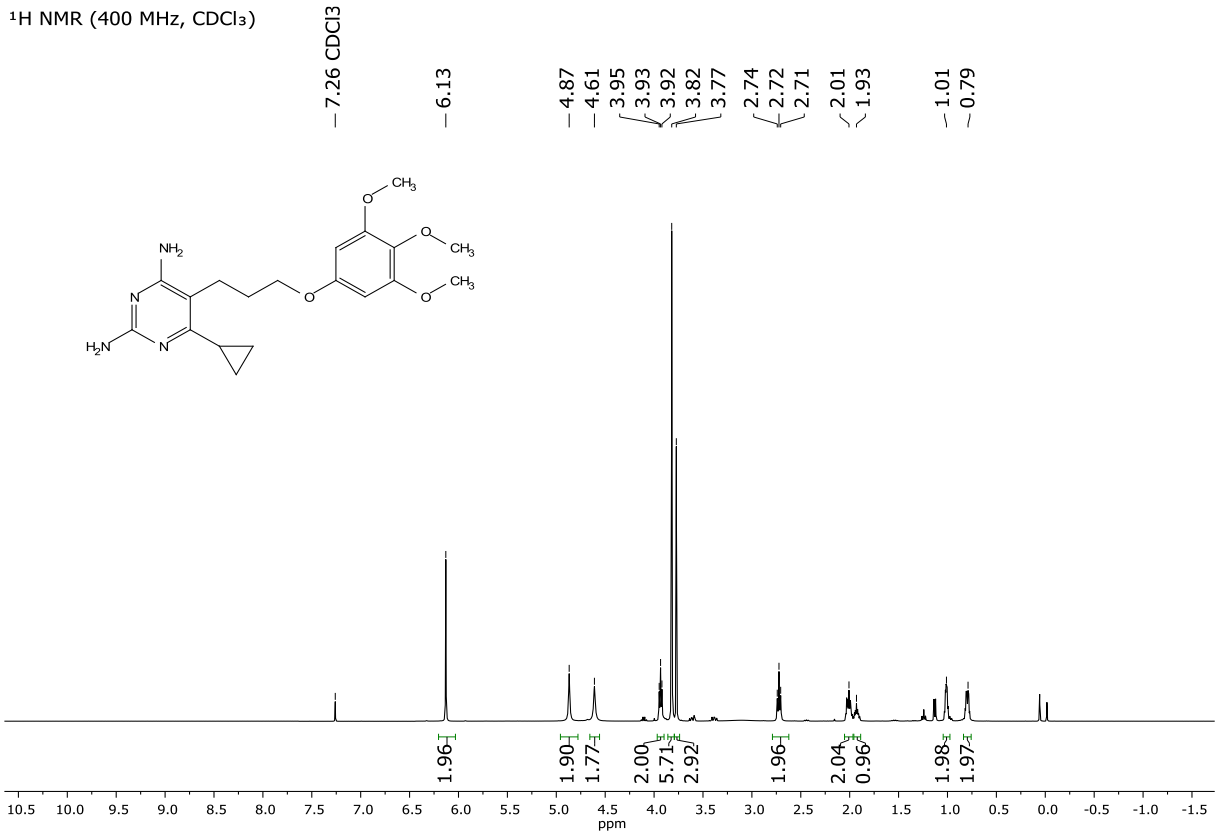
HRMS:


371.2061
372.2089
364 366 368 370 372 374 376 378 380
m/z (Da)
NH2
N
O
H2N
N
O
O
CH3
Intensity
2.8×10^7
2.6×10^7
2.4×10^7
2.2×10^7
2.0×10^7
1.8×10^7
1.6×10^7
1.4×10^7
1.2×10^7
1.0×10^7
8.0×10^6
6.0×10^6
4.0×10^6
2.0×10^6
0.0
-2.0×10^6
102.1327
282.2782
371.2061
372.2089
50 100 150 200 250 300 350 400 450 500 550 600 650 700 750 800 850 900 950 1000 1050 1100 1150 1200 1250
m/z (Da)

6-Cyclopropyl-5-[3-(3,4,5-trimethoxyphenoxy)propyl]pyrimidine-2,4-diamine (**18c**)

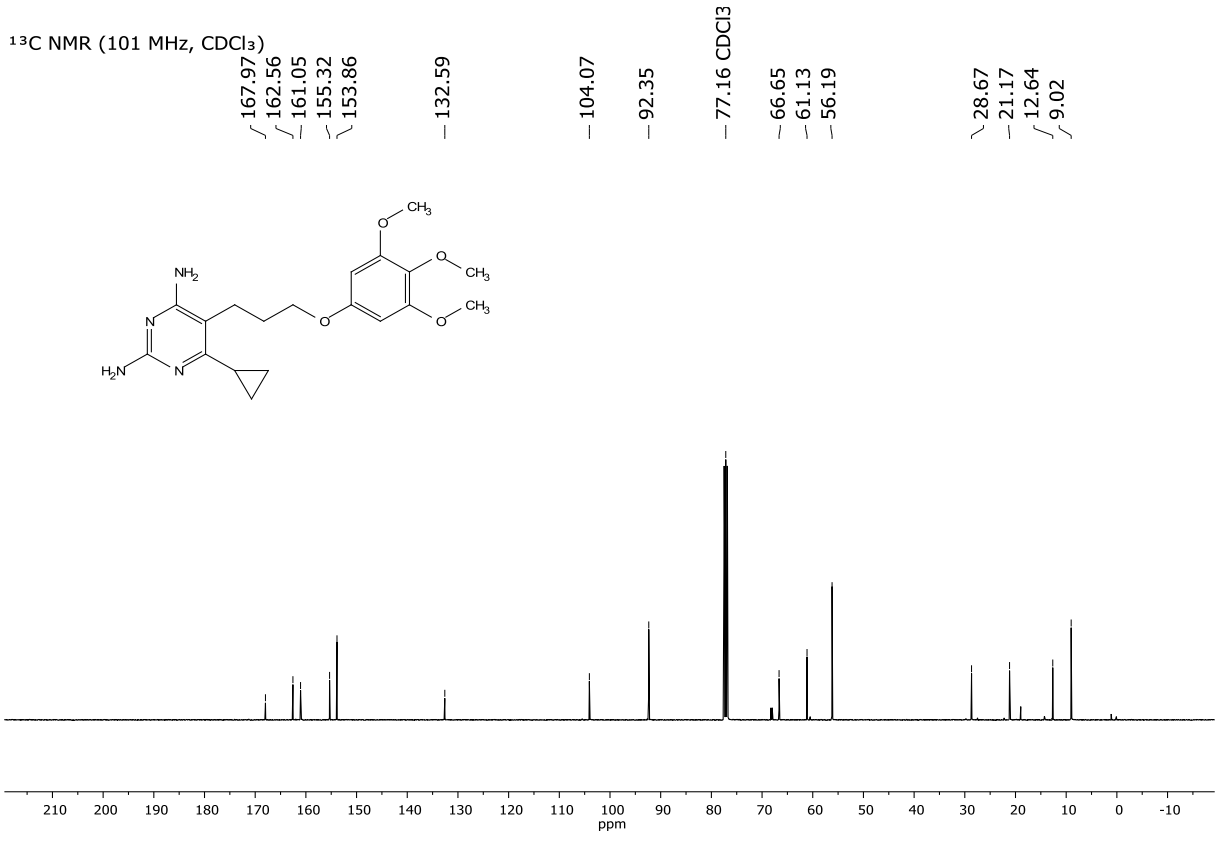
^{1}H NMR:

^{1}H NMR (400 MHz, $CDCl_3$)

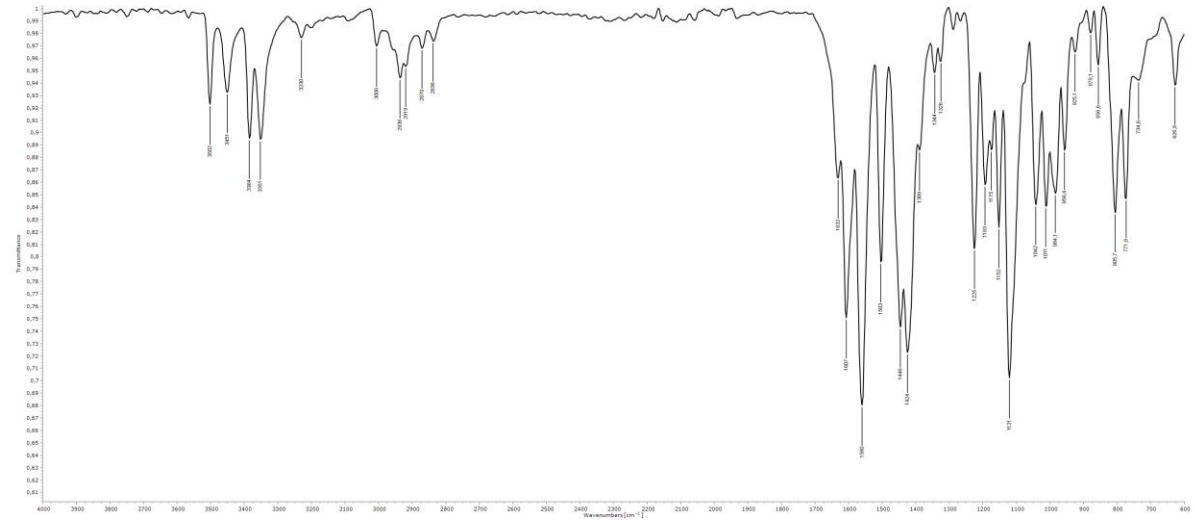


^{13}C NMR:

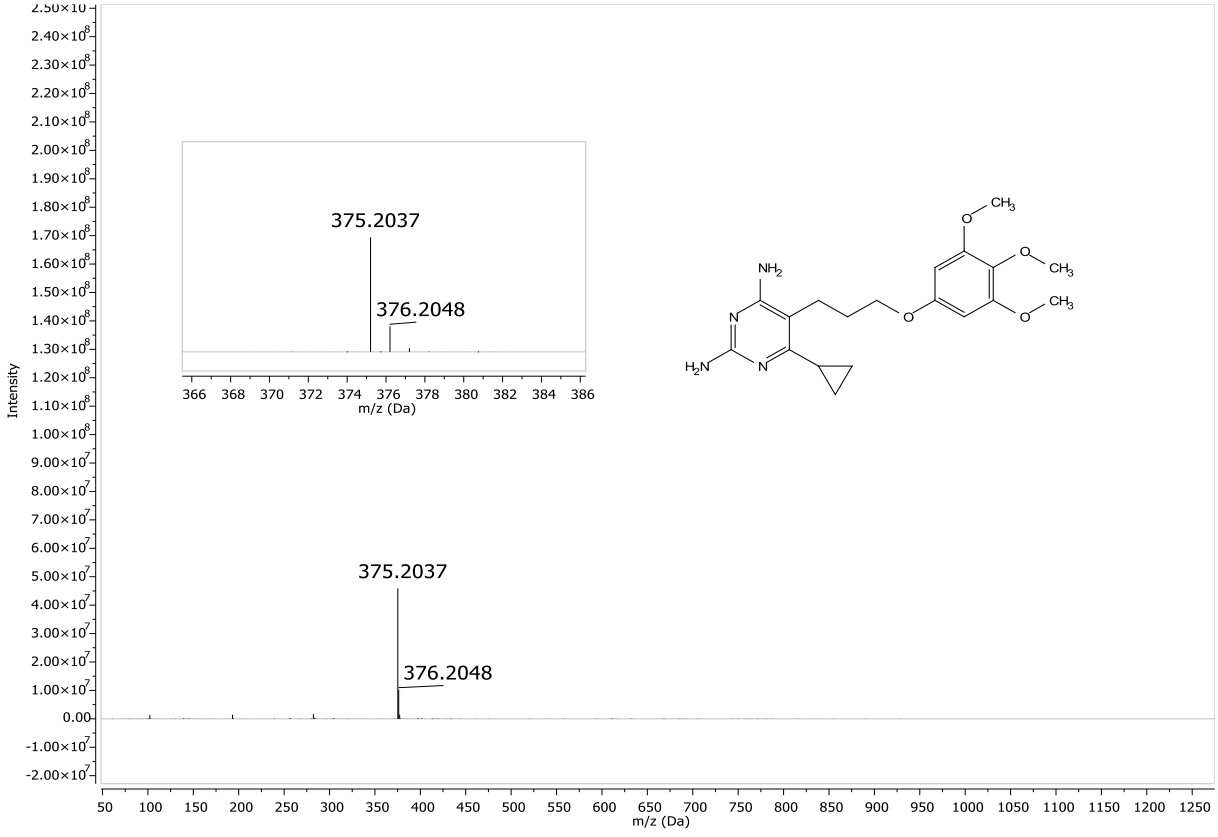
^{13}C NMR (101 MHz, $CDCl_3$)



FTIR:



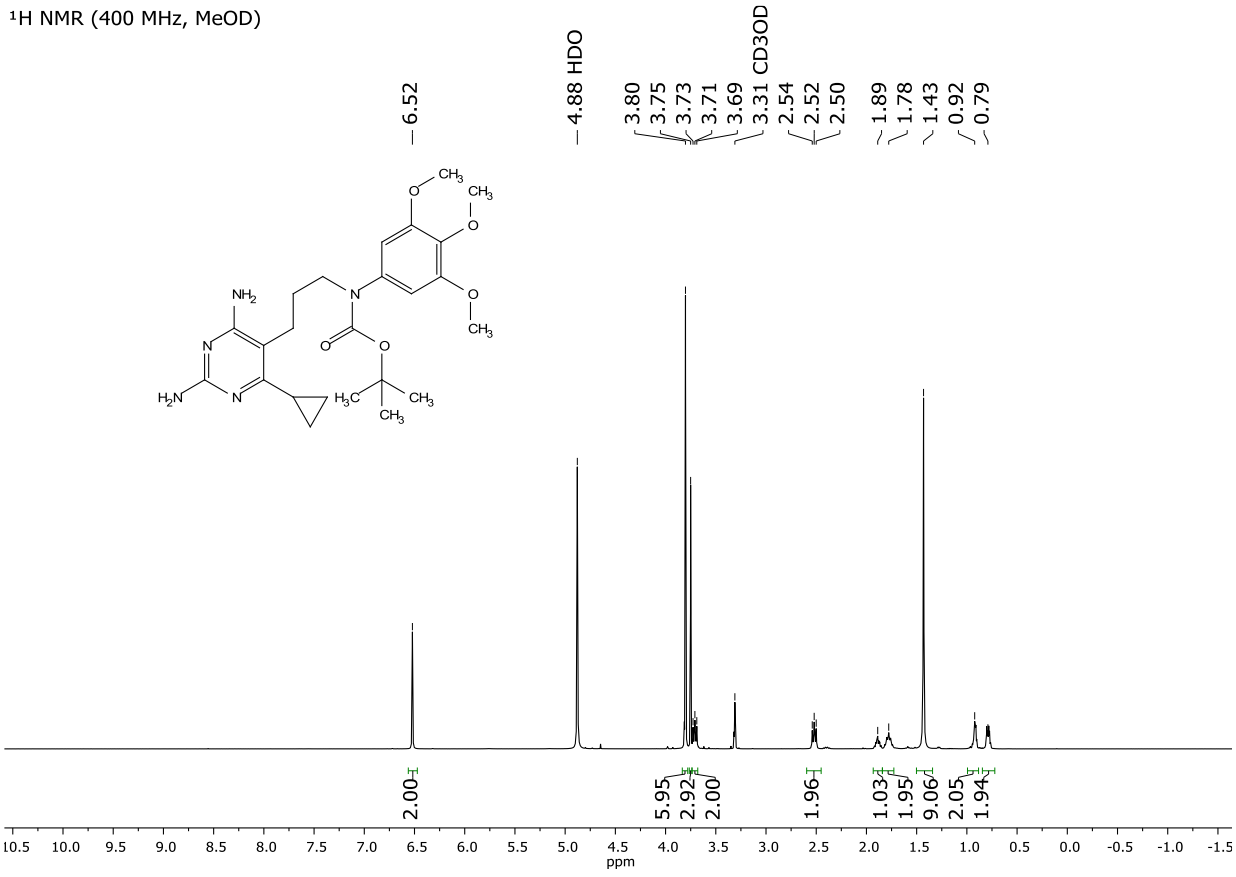
HRMS:



tert-Butyl (3-(2,4-diamino-6-cyclopropylpyrimidin-5-yl)propyl)(3,4,5-trimethoxyphenyl) carbamate (**18d**)

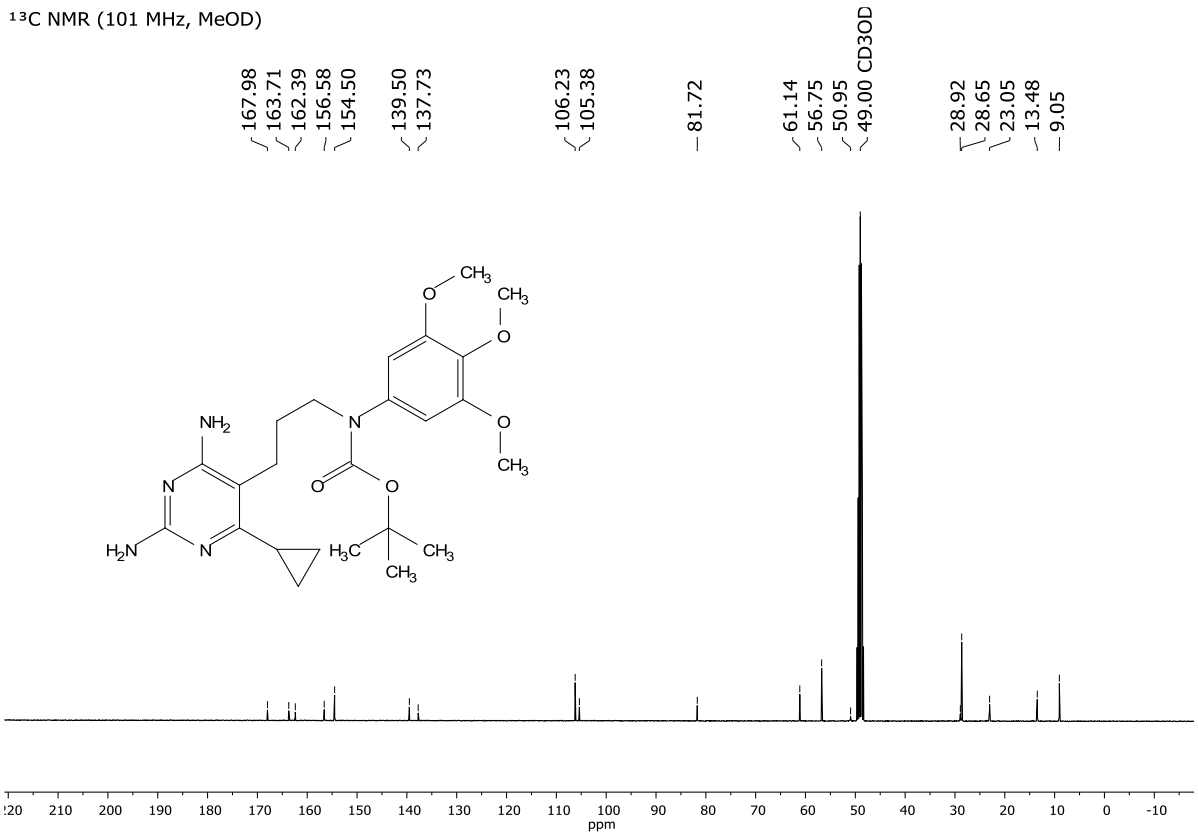
^{1}H NMR:

^{1}H NMR (400 MHz, MeOD)

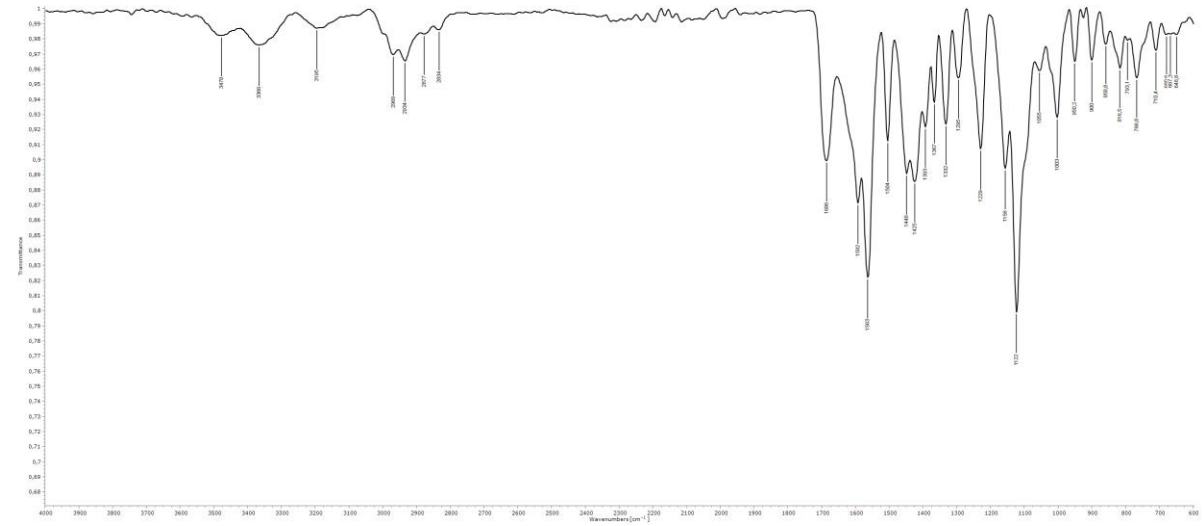


^{13}C NMR:

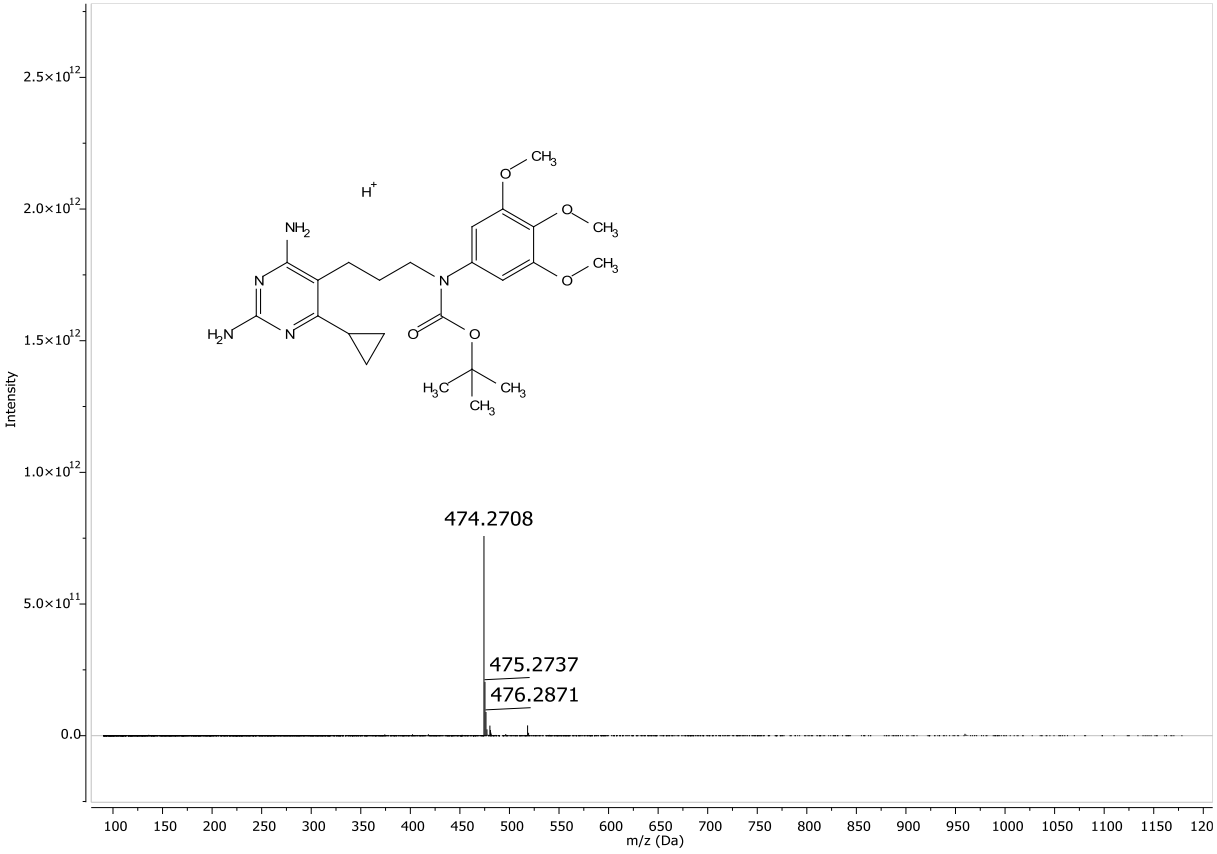
^{13}C NMR (101 MHz, MeOD)



FTIR:



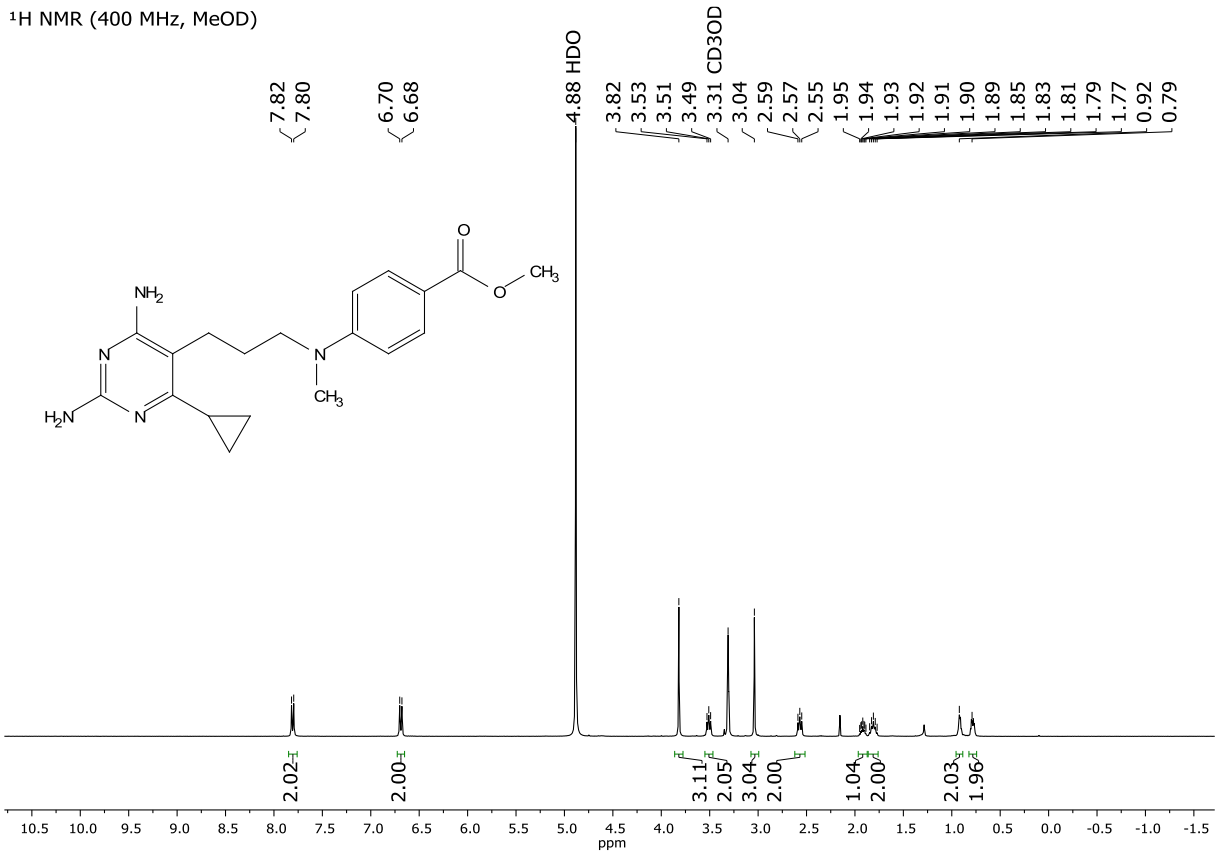
HRMS:



Methyl 4-{[3-(2,4-diamino-6-cyclopropylpyrimidin-5-yl)propyl](methyl)amino}benzoate (**18e**)

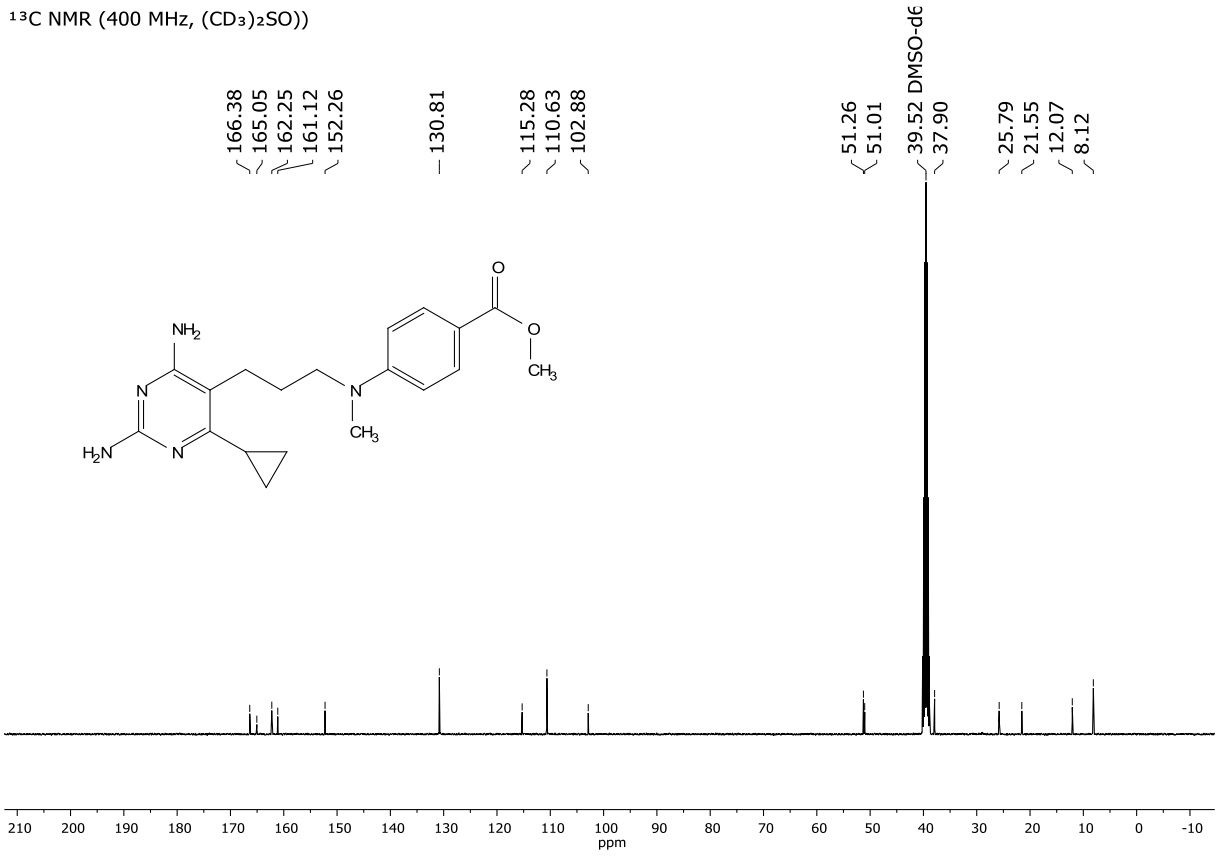
^{1}H NMR:

^{1}H NMR (400 MHz, MeOD)

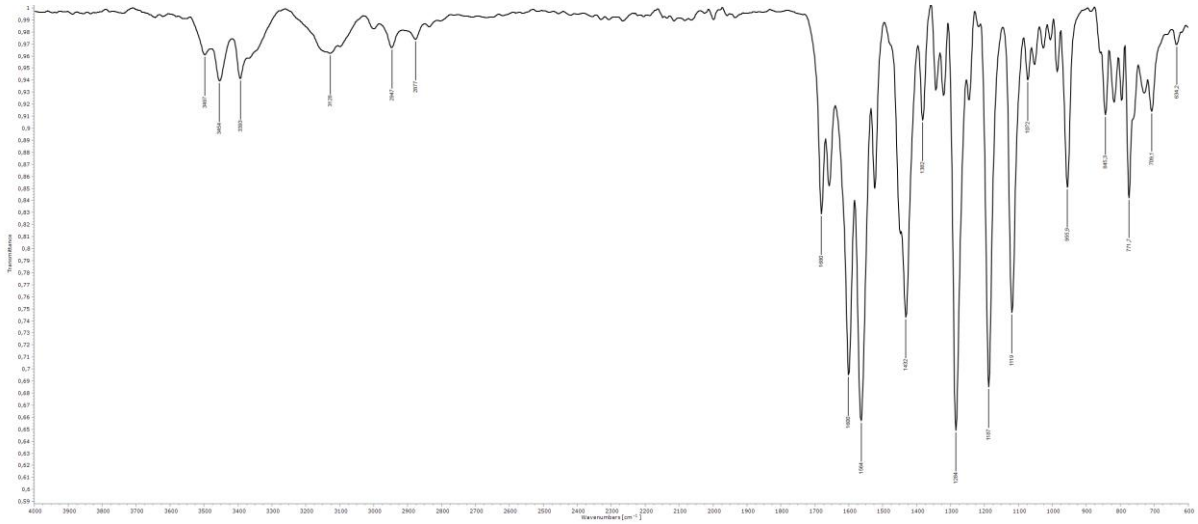


^{13}C NMR:

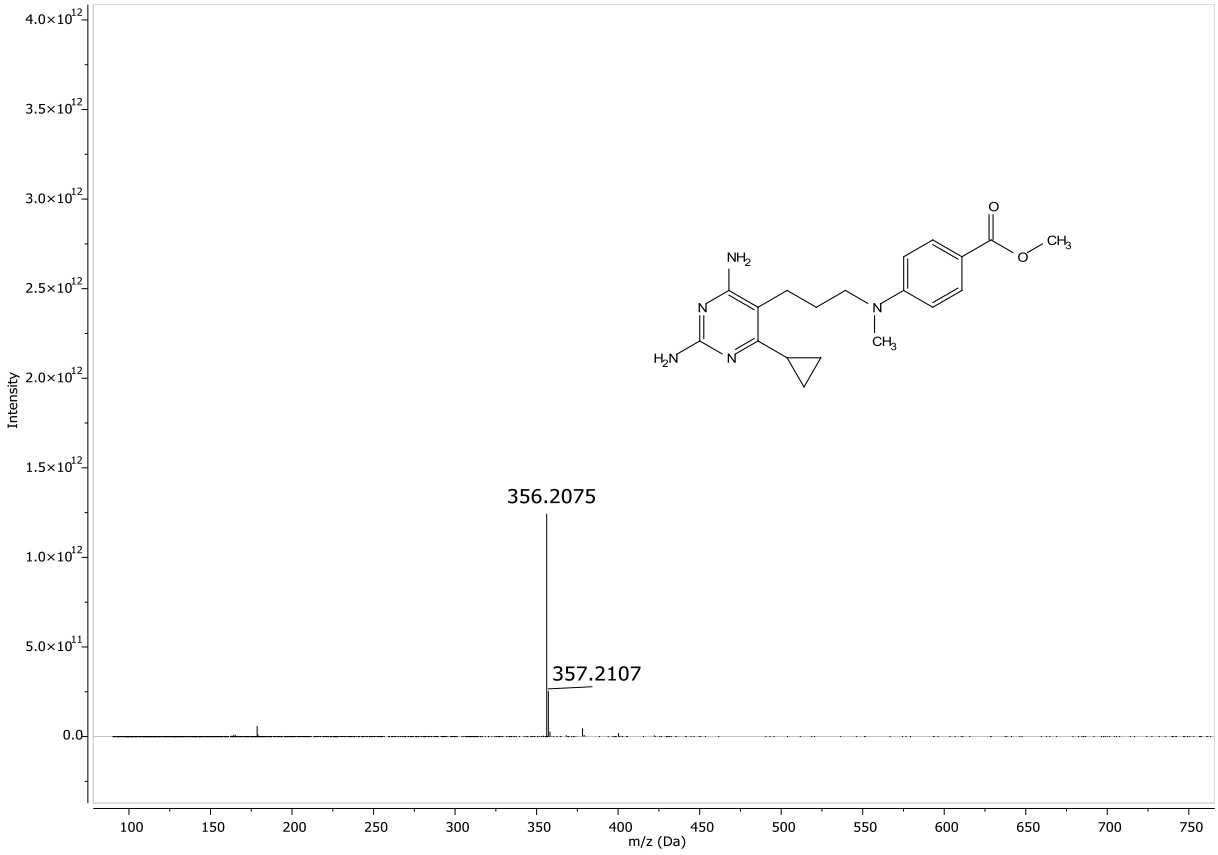
^{13}C NMR (400 MHz, $(CD_3)_2SO$))



FTIR:



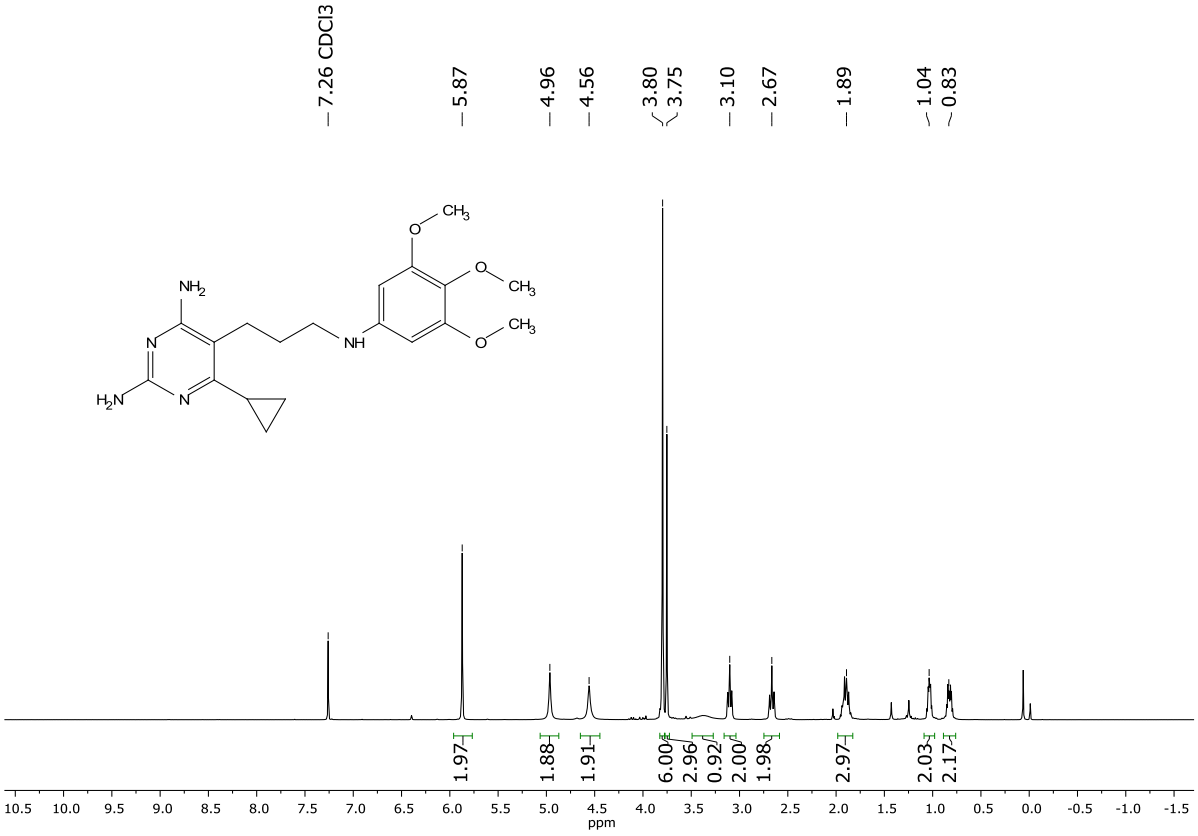
HRMS:


Intensity
4.0×10^12
3.5×10^12
3.0×10^12
2.5×10^12
2.0×10^12
1.5×10^12
1.0×10^12
5.0×10^11
0.0
356.2075
357.2107
100
150
200
250
300
350
400
450
500
550
600
650
700
750
m/z (Da)
NH2
H2N
N
N
N
CH3
O
O
CH3

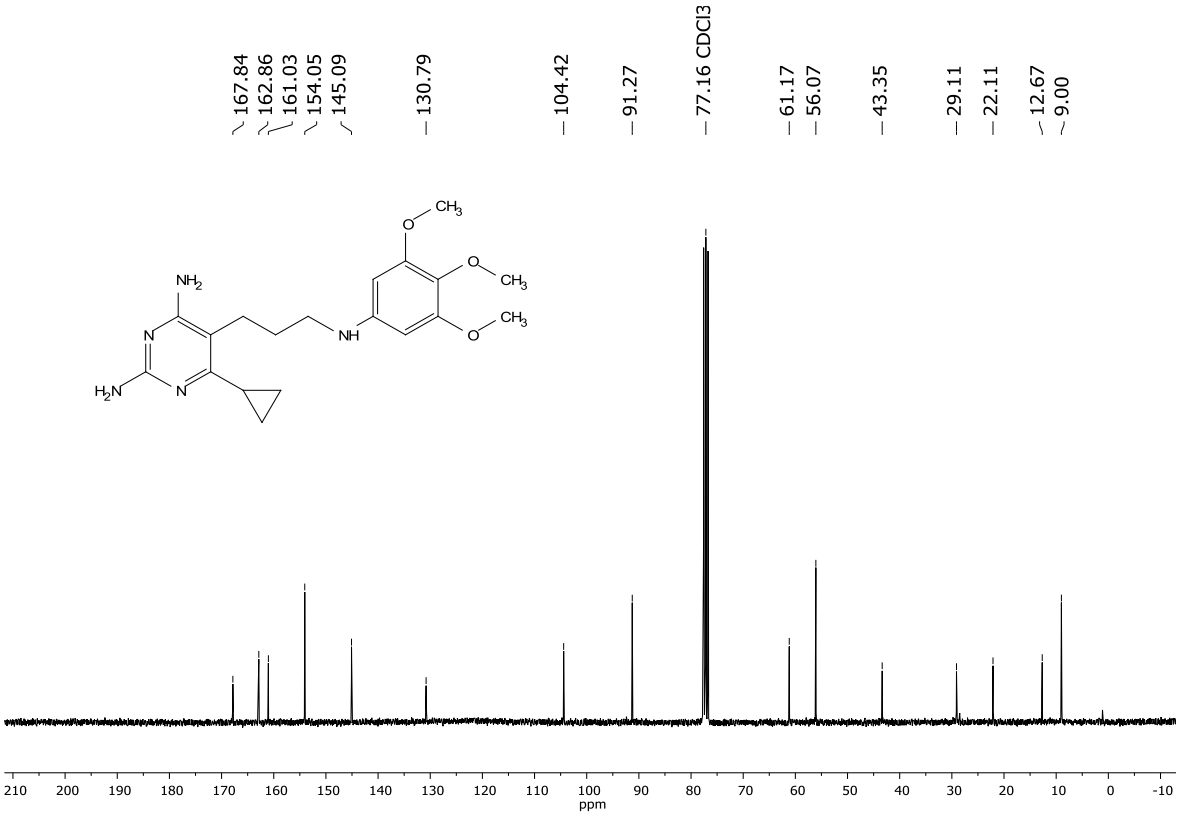
Removal of Boc group after Sonogashira and reduction (**19**)

6-Cyclopropyl-5-{3-[(3,4,5-trimethoxyphenyl)amino]propyl}pyrimidine-2,4-diamine (**19**)

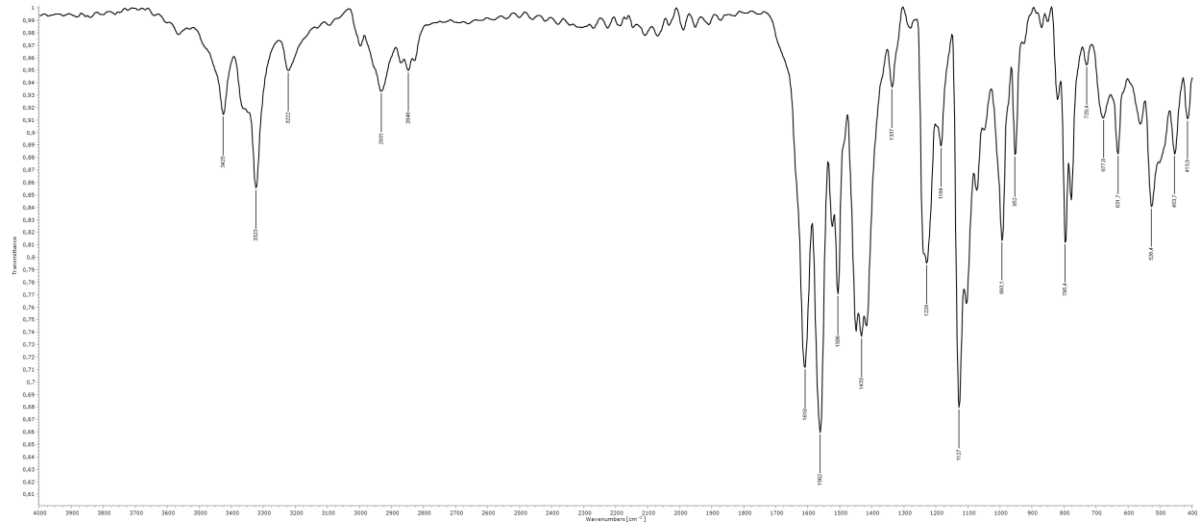
^{1}H NMR:



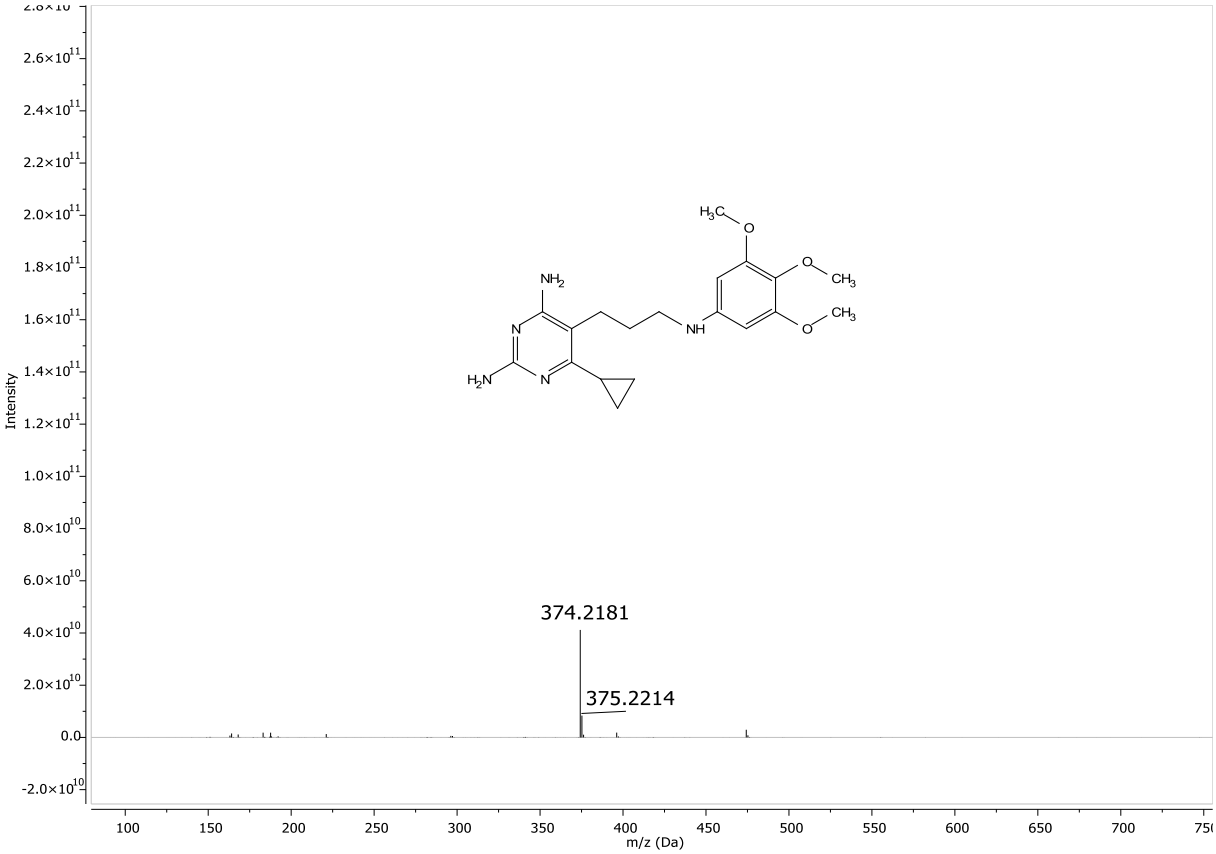
^{13}C NMR:



FTIR:



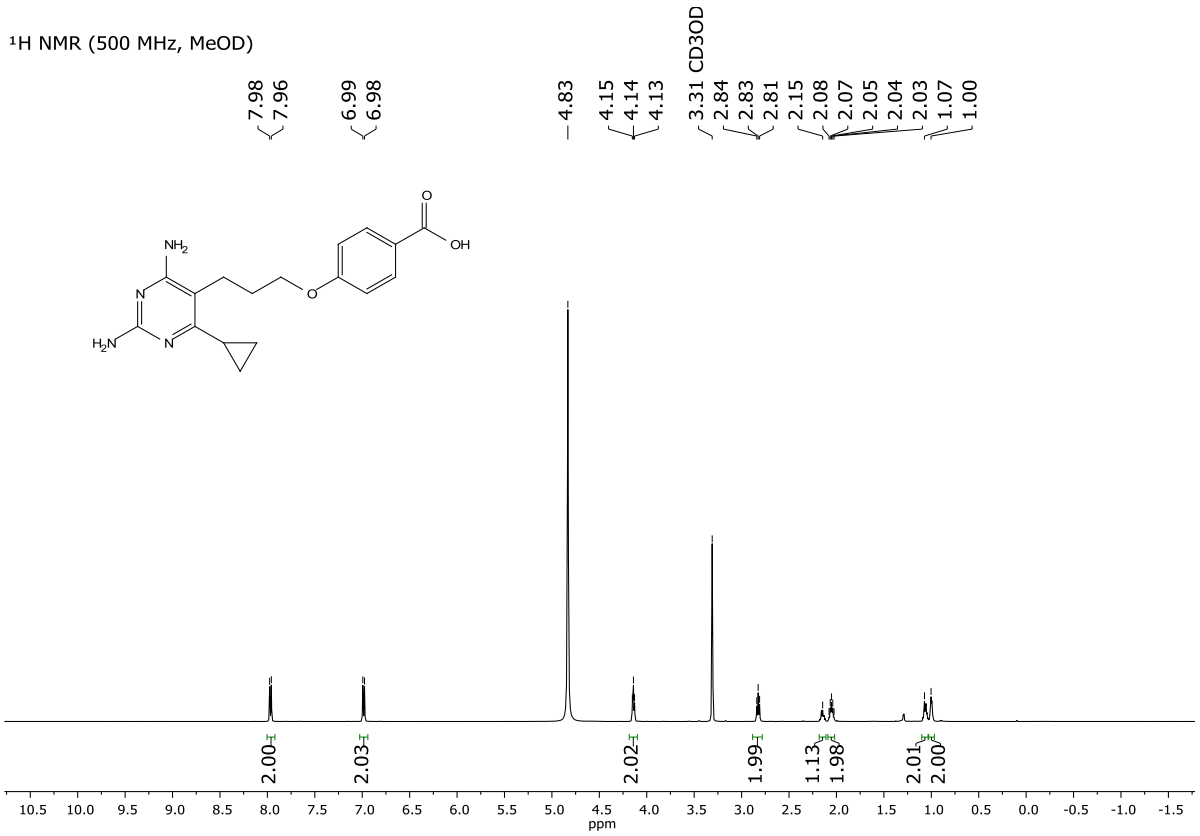
HRMS:



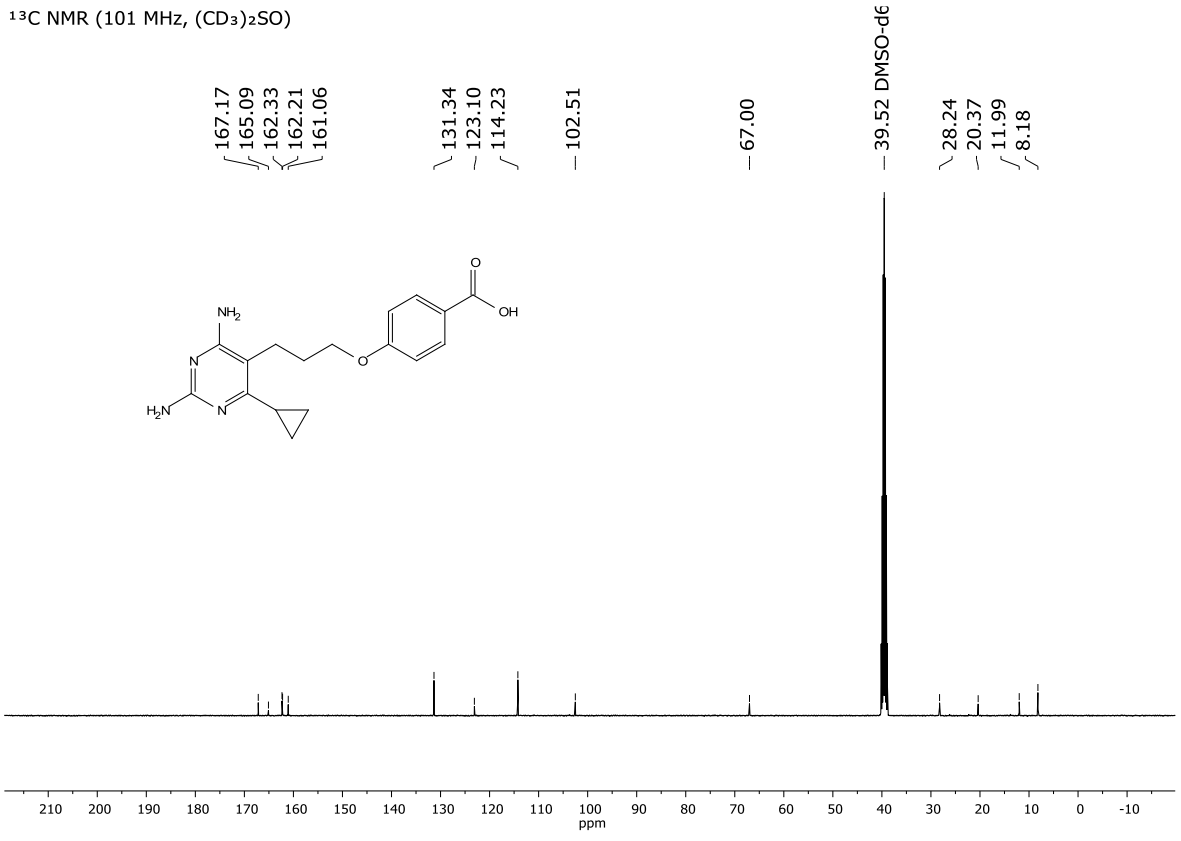
Hydrolysis of esters after Sonogashira and reduction (**20**)

4-[3-(2,4-Diamino-6-cyclopropylpyrimidin-5-yl)propoxy]benzoic acid (**20a**)

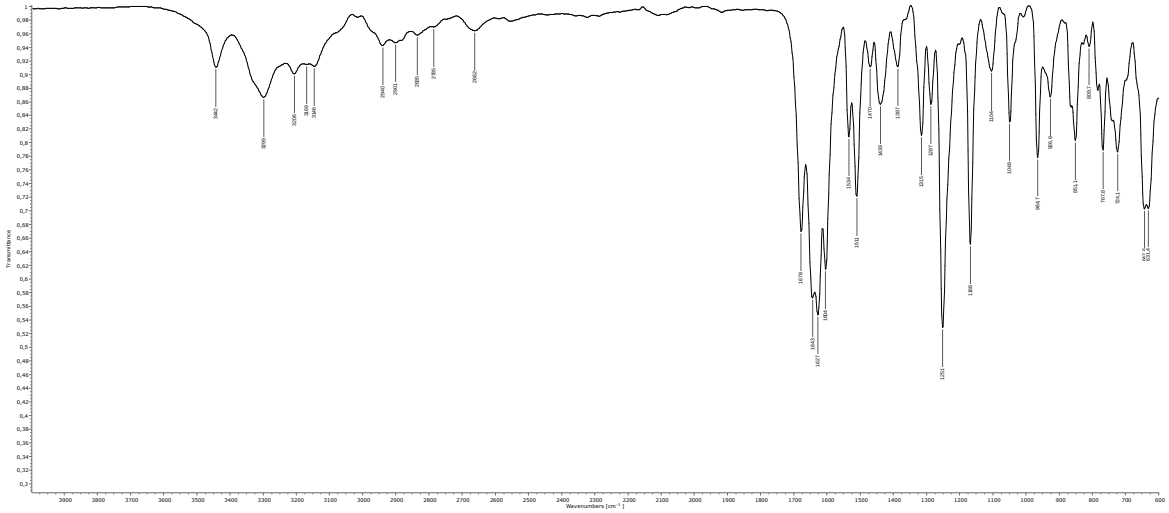
^{1}H NMR:



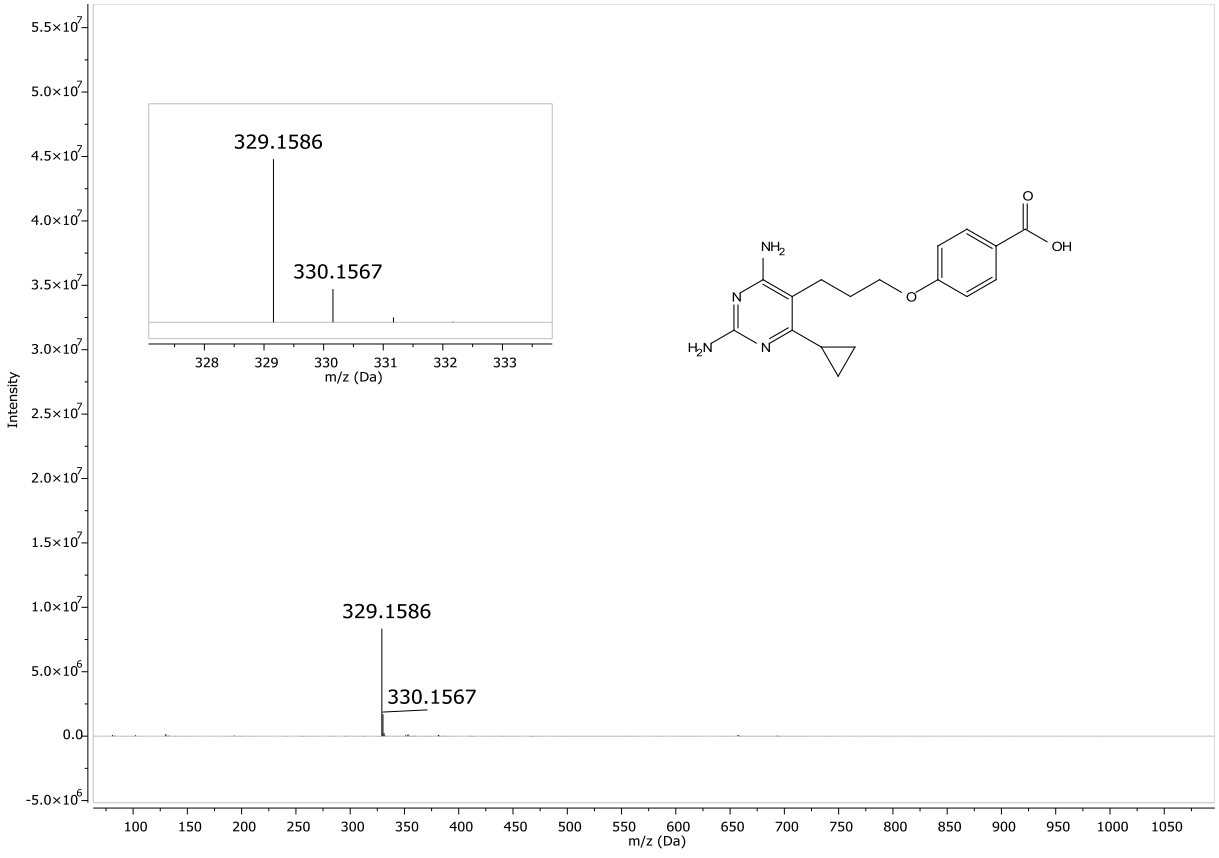
^{13}C NMR:



FTIR:



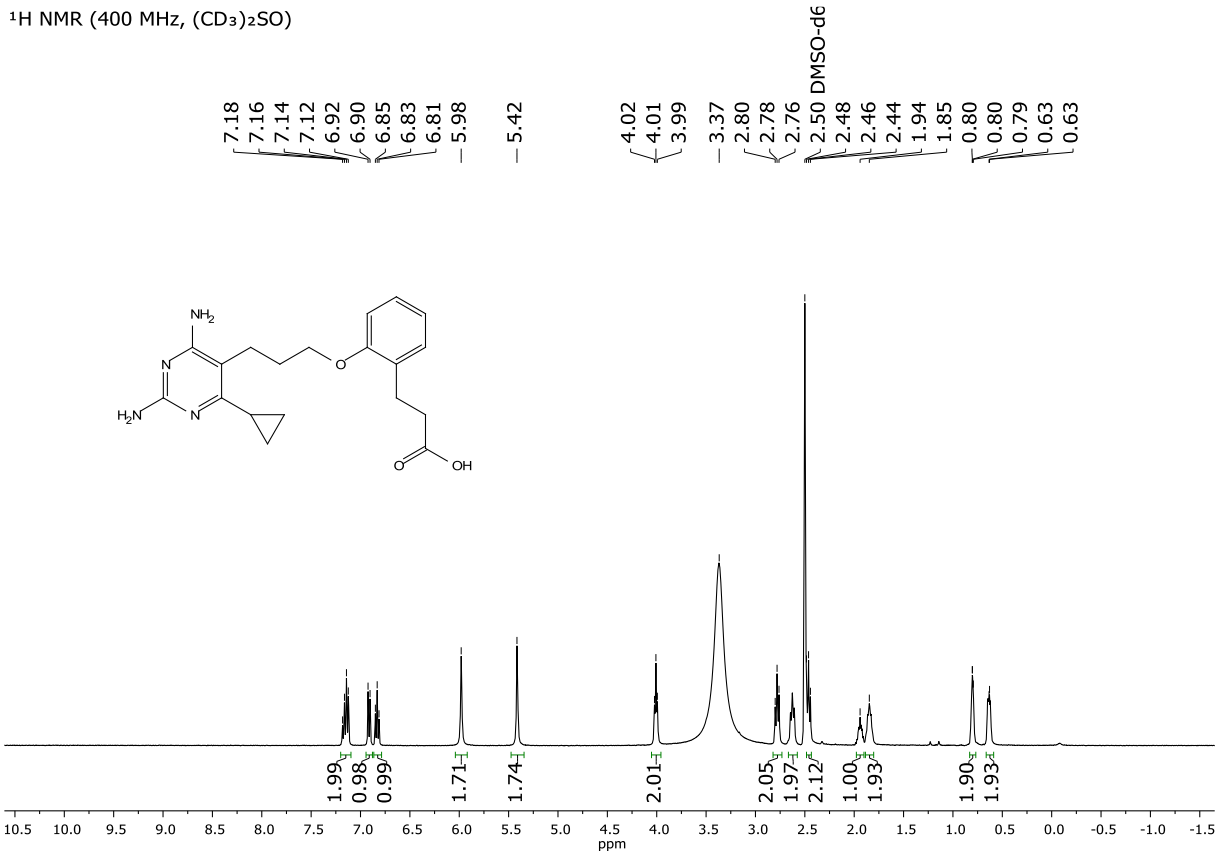
HRMS:


329.1586
330.1567
328
329
330
331
332
333
m/z (Da)
NH2
N
H2N
N
O
O
OH
Intensity
5.5×10^7
5.0×10^7
4.5×10^7
4.0×10^7
3.5×10^7
3.0×10^7
2.5×10^7
2.0×10^7
1.5×10^7
1.0×10^7
5.0×10^6
0.0
-5.0×10^6
329.1586
330.1567
100
150
200
250
300
350
400
450
500
550
600
650
700
750
800
850
900
950
1000
1050
m/z (Da)

3-{2-[3-(2,4-Diamino-6-cyclopropylpyrimidin-5-yl)propoxy]phenyl}propanoic acid (**20b**)

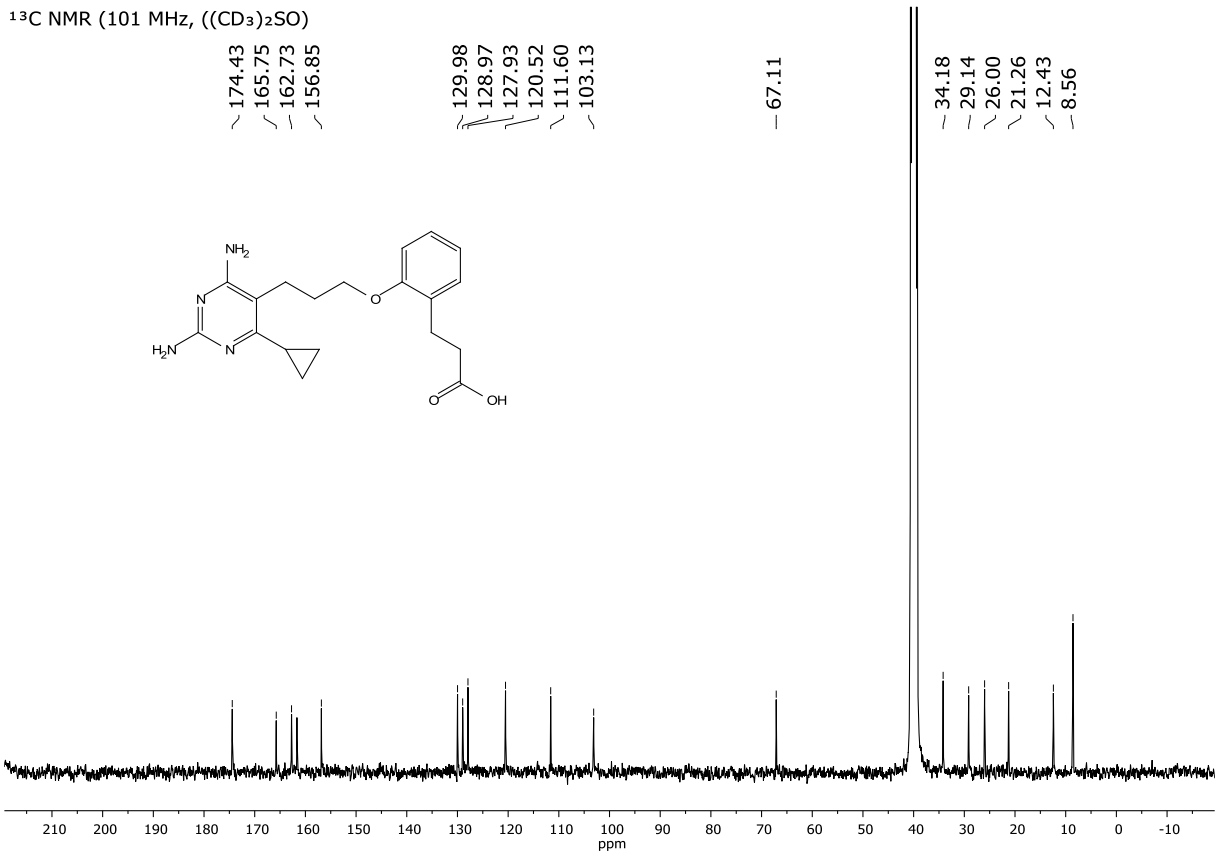
^{1}H NMR:

^{1}H NMR (400 MHz, $(CD_3)_2SO$)

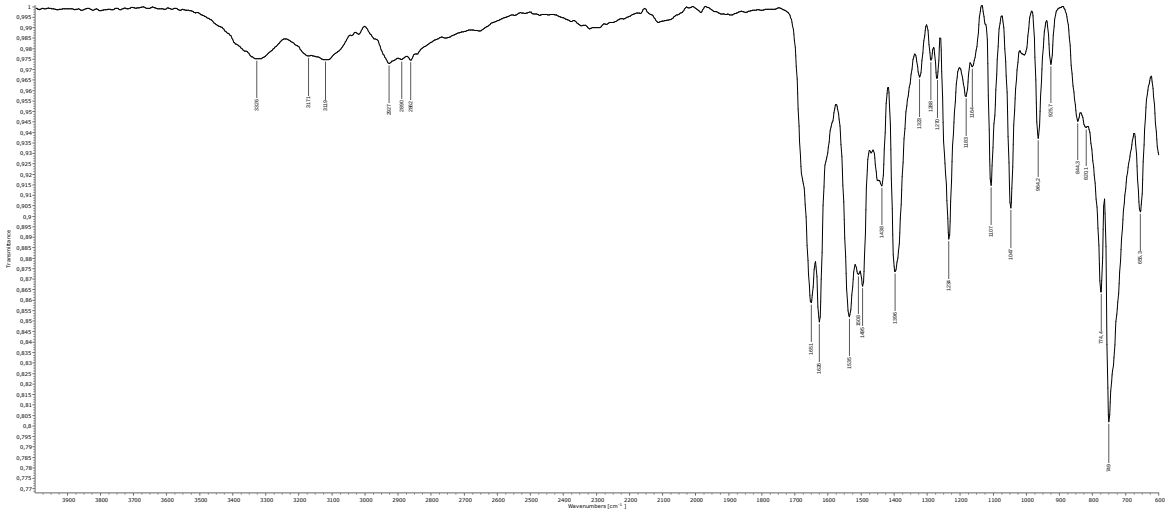


^{13}C NMR:

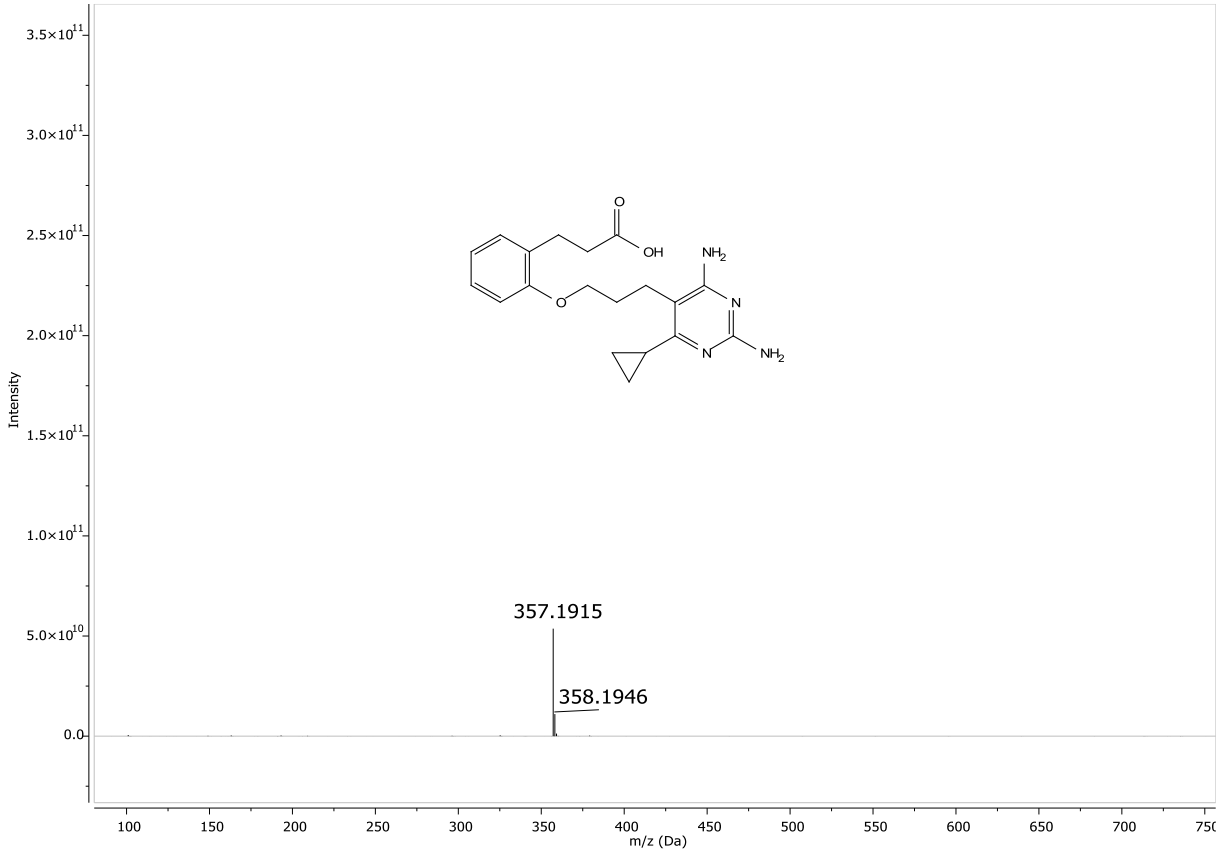
^{13}C NMR (101 MHz, $((CD_3)_2SO$)



FTIR:



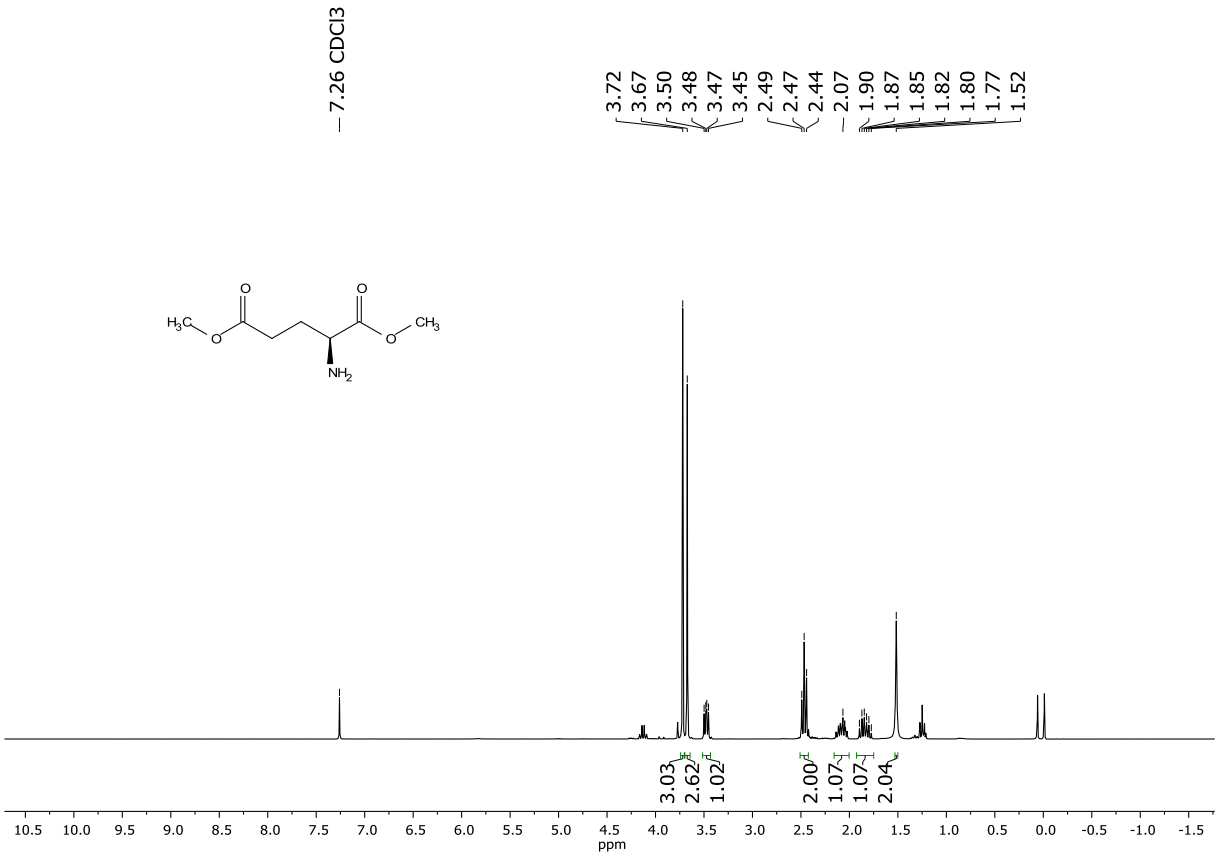
HRMS:


3.5×10^11
3.0×10^11
2.5×10^11
2.0×10^11
1.5×10^11
1.0×10^11
5.0×10^10
0.0
Intensity
357.1915
358.1946
100
150
200
250
300
350
400
450
500
550
600
650
700
m/z (Da)
O
OH
NH2
N
O
N
NH2

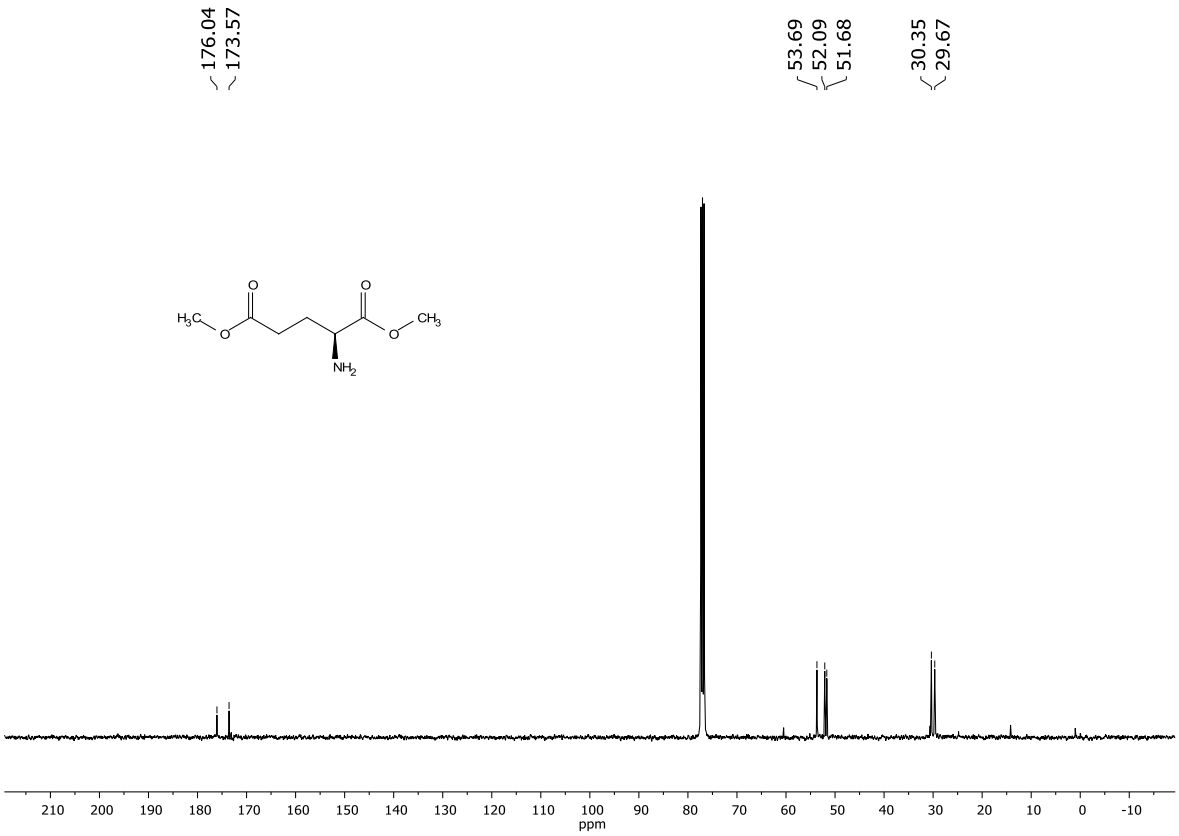
Protection of glutamic acid (**23**)

(S)-dimethyl 2-aminopentanedioate (**23**)

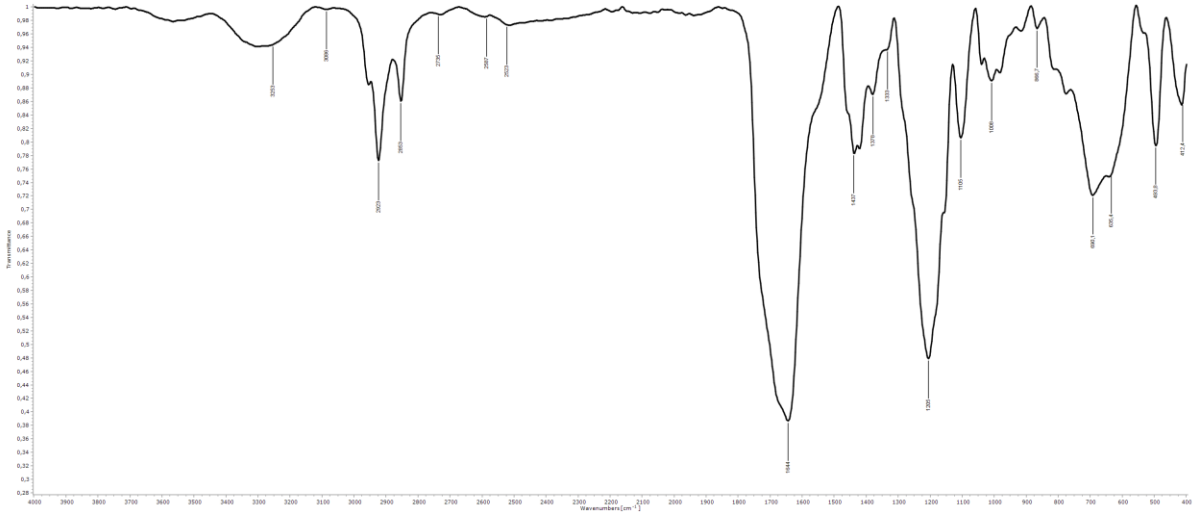
^{1}H NMR:



^{13}C NMR:


176.04
173.57
53.69
52.09
51.68
30.35
29.67
O
O
H3C
O
O
CH3
NH2
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
ppm

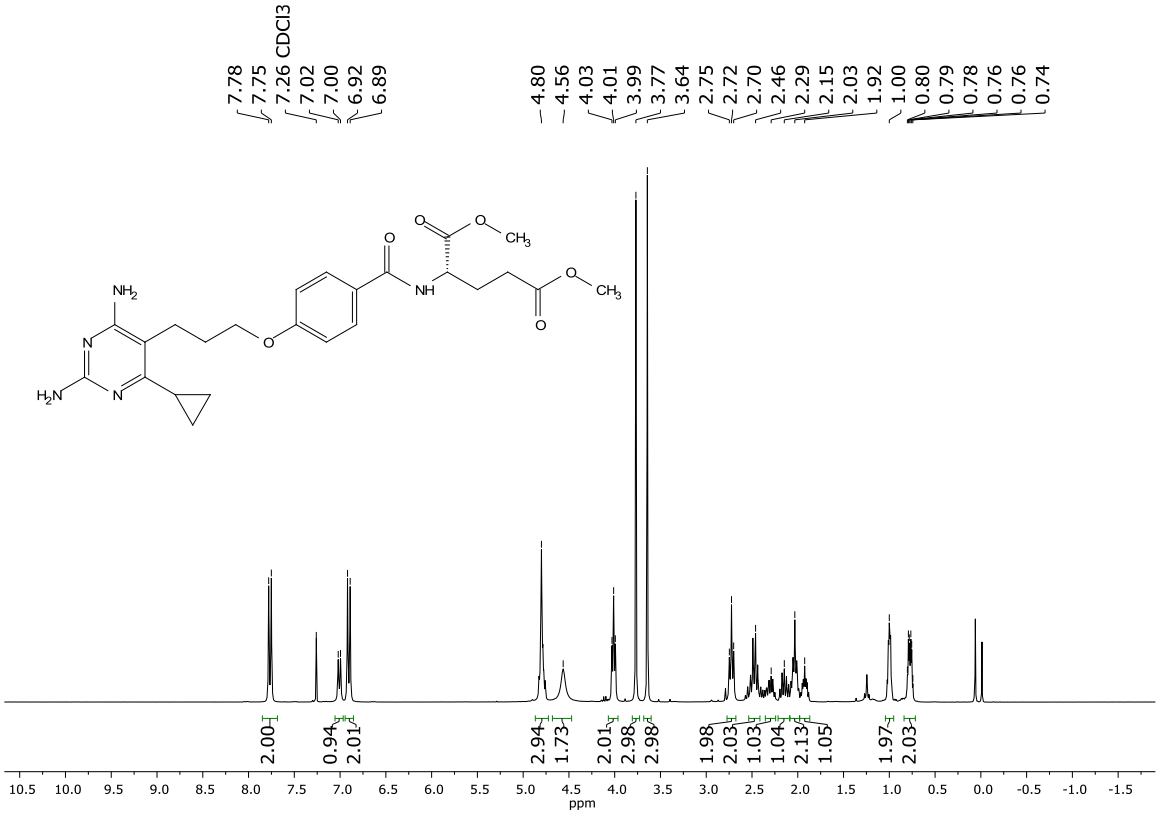
FTIR:



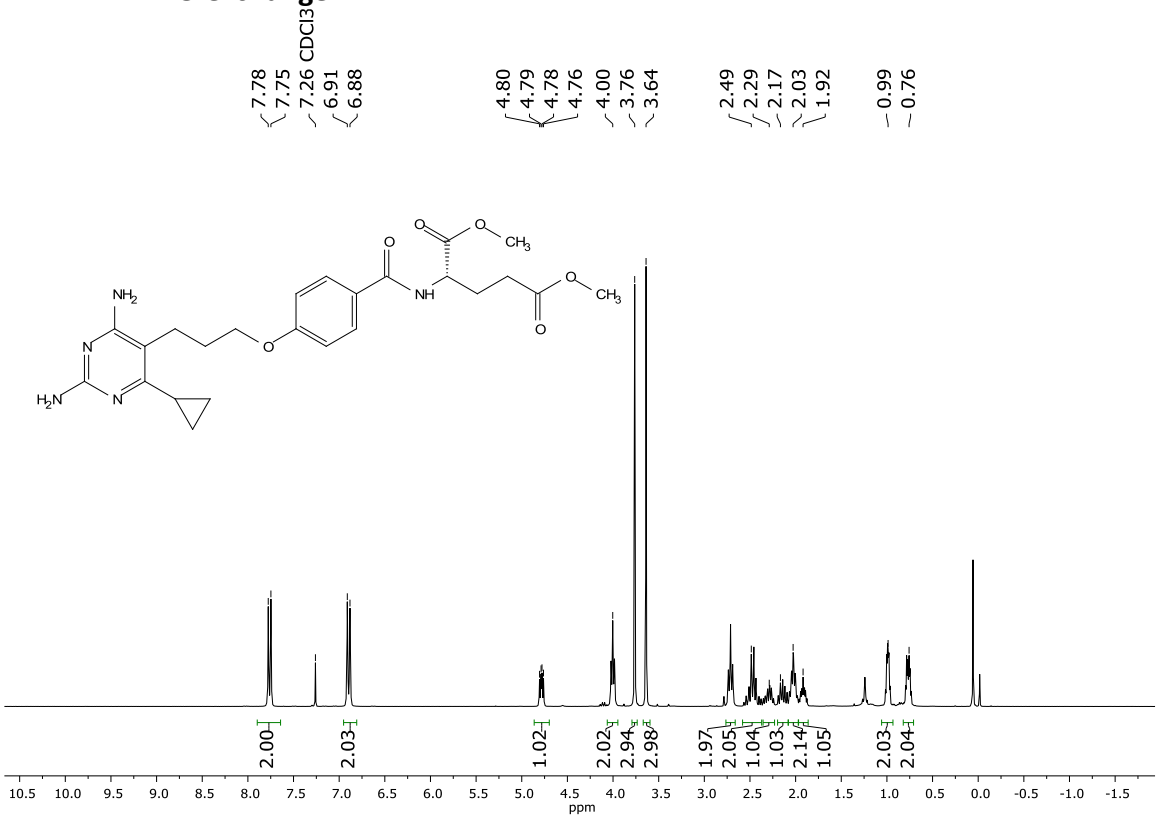
Addition of glutamic acid (**24**)

(S)-Dimethyl 2-{4-[3-(2,4-diamino-6-cyclopropylpyrimidin-5-yl)propoxy]benzamido} pentanedioate (**24**)

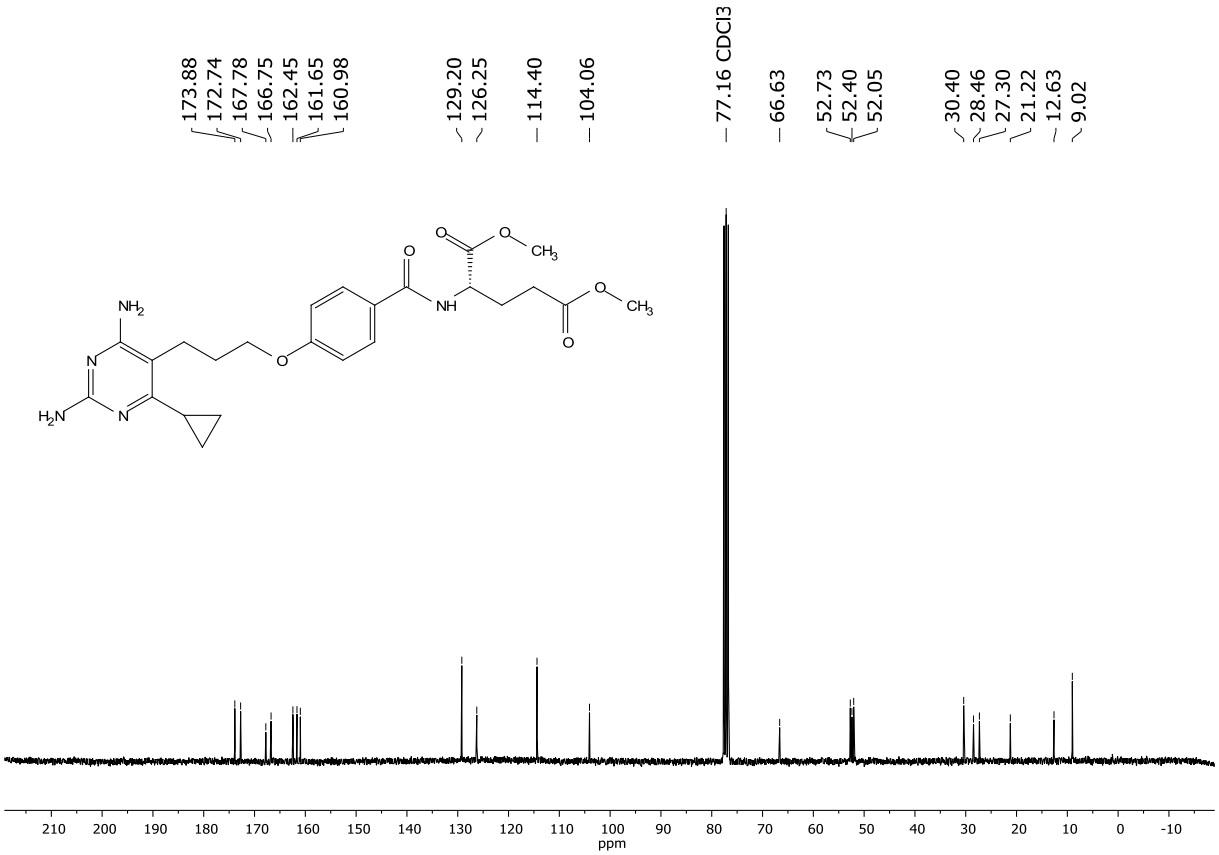
^{1}H NMR:



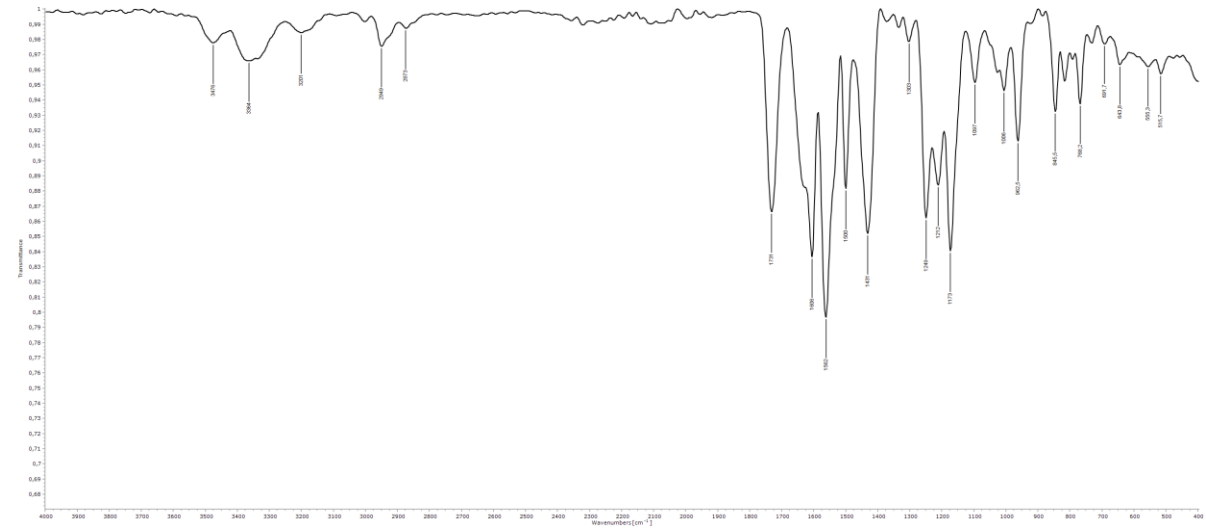
^{1}H NMR – D2O exchange:



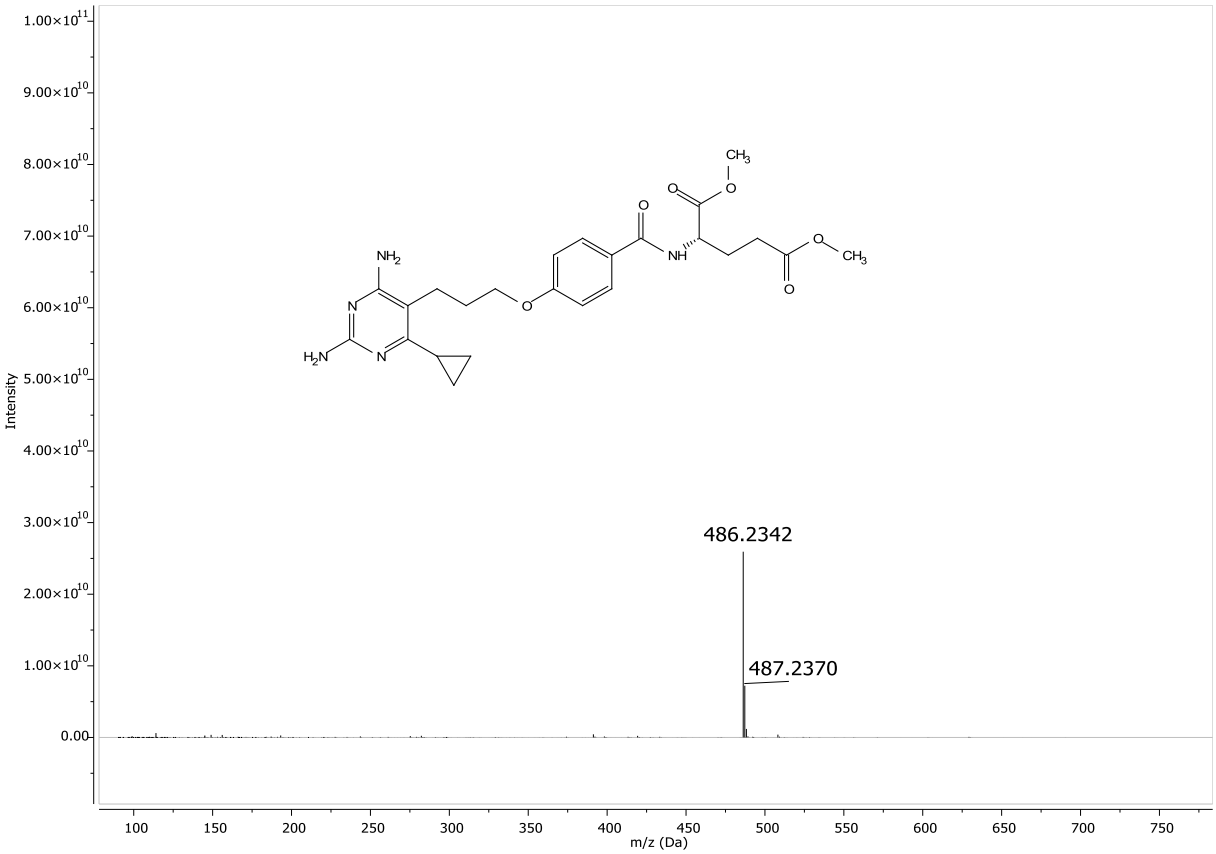
^{13}C NMR:



FTIR:



HRMS:



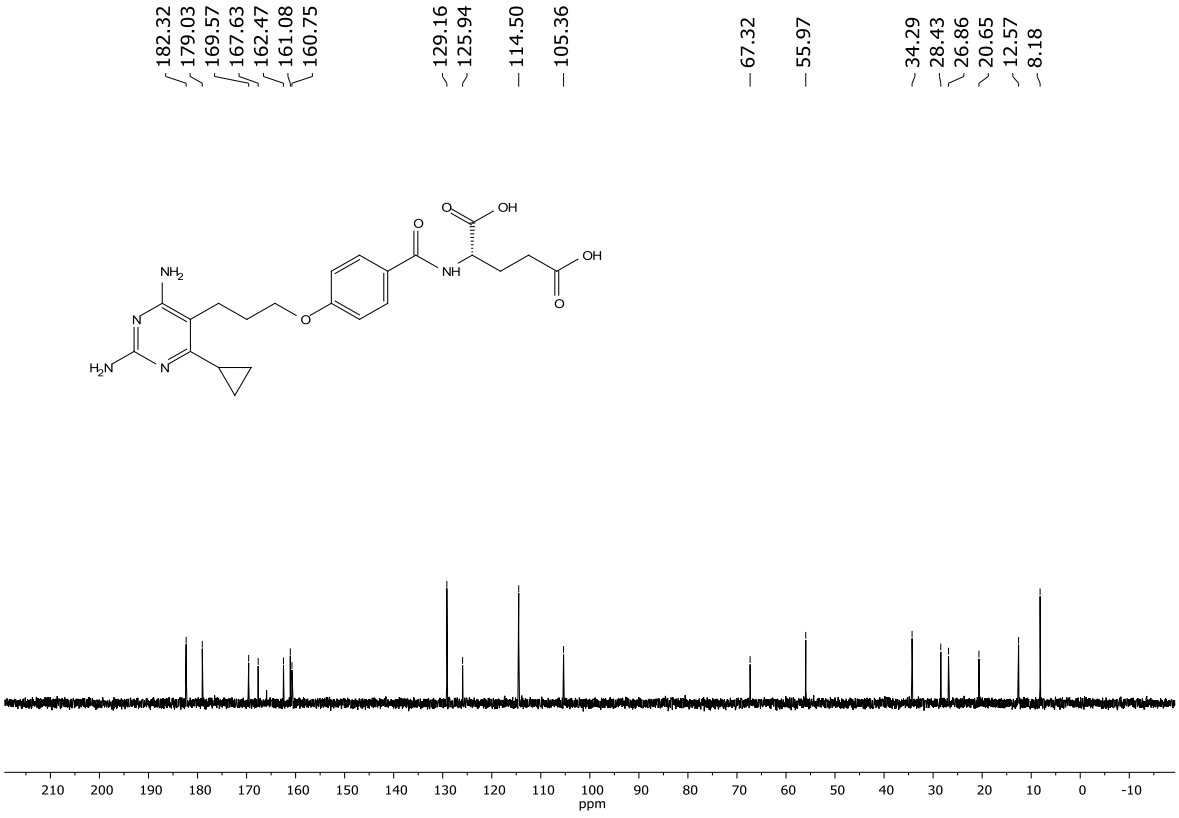
Deprotection of glutamic acid (**21**)

(*S*)-2-{4-[3-(2,4-diamino-6-cyclopropylpyrimidin-5-yl)propoxy]benzamido} pentanedioic acid (**25**)

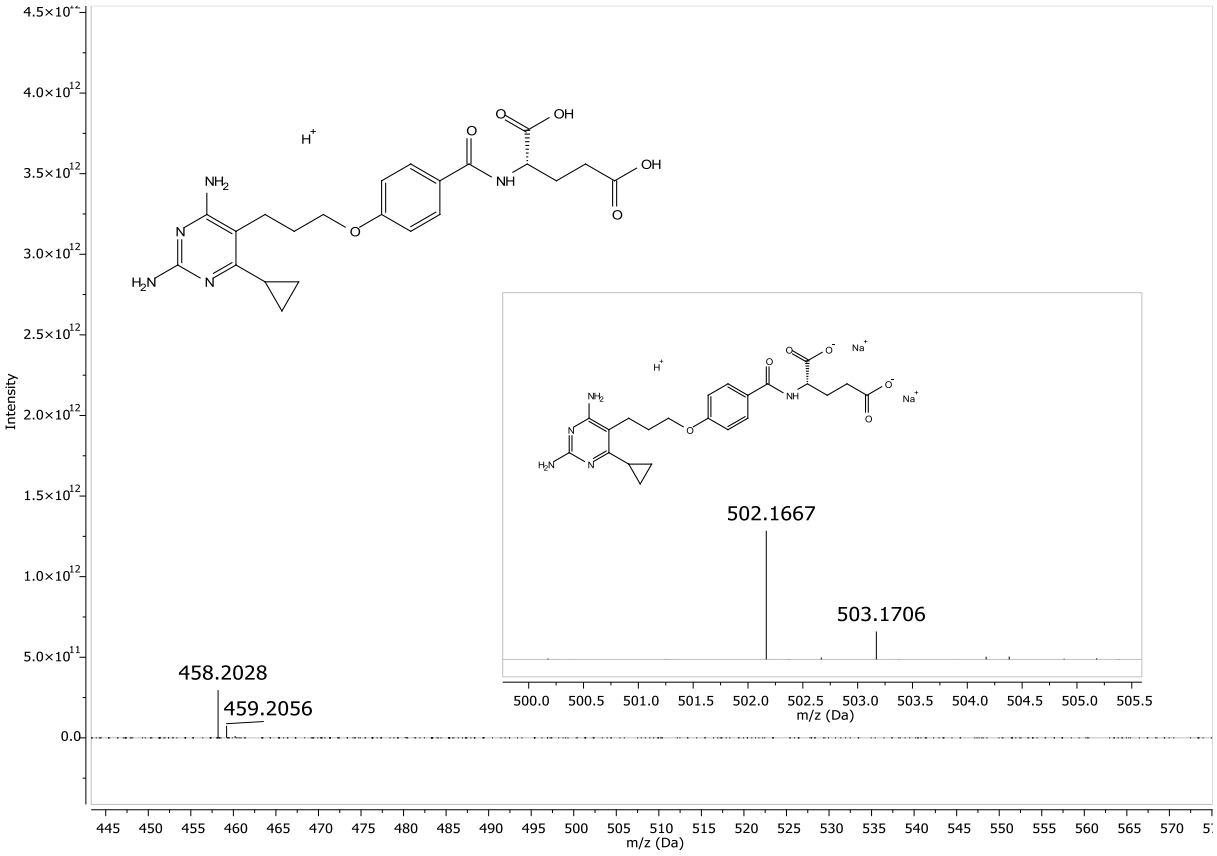
^{1}H NMR:



^{13}C NMR:

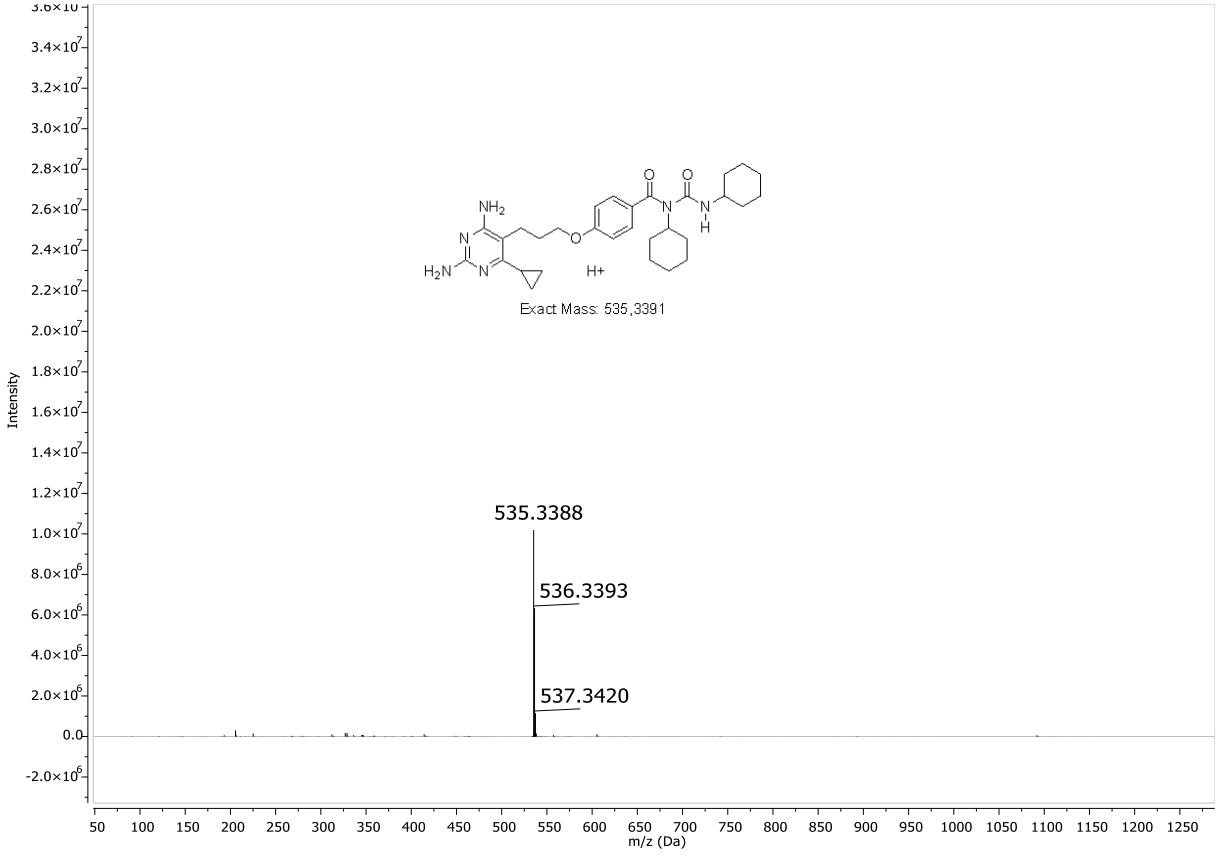


HRMS:

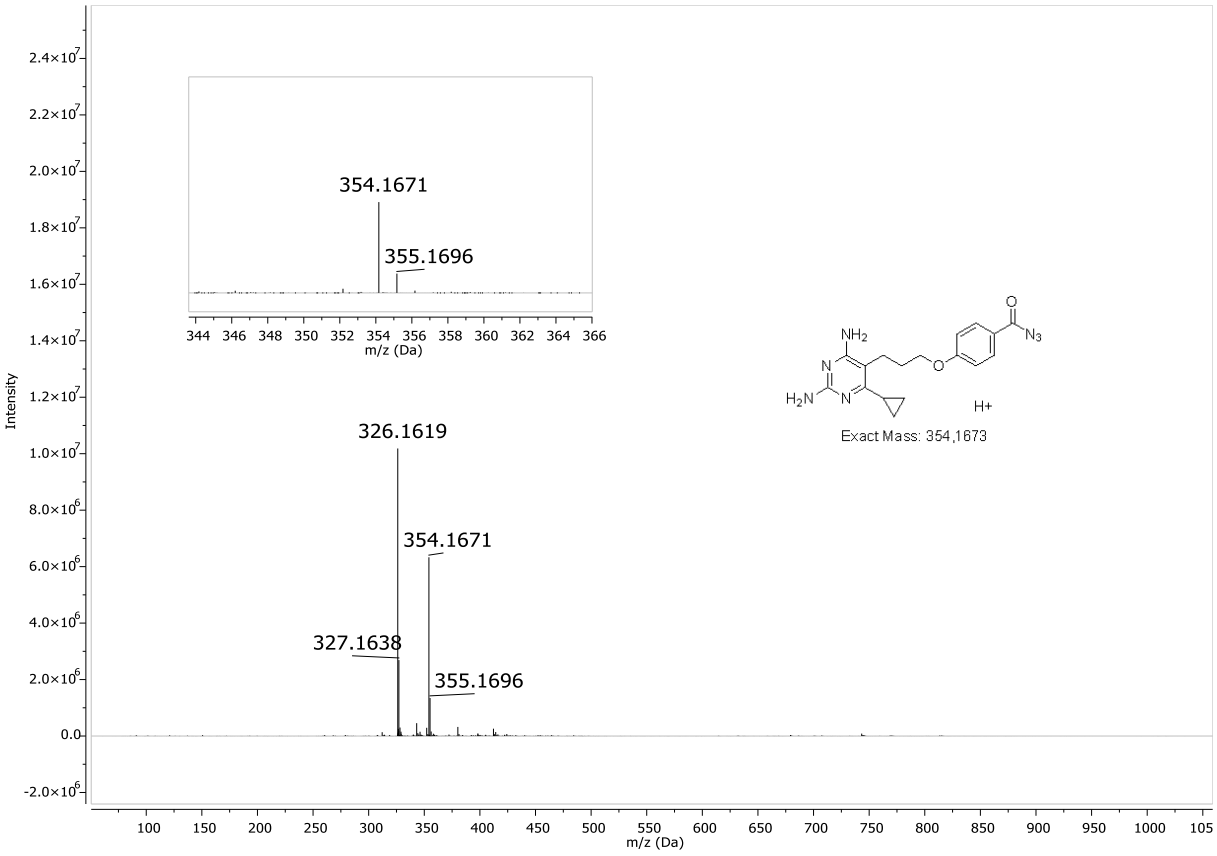


Additional mass spectra

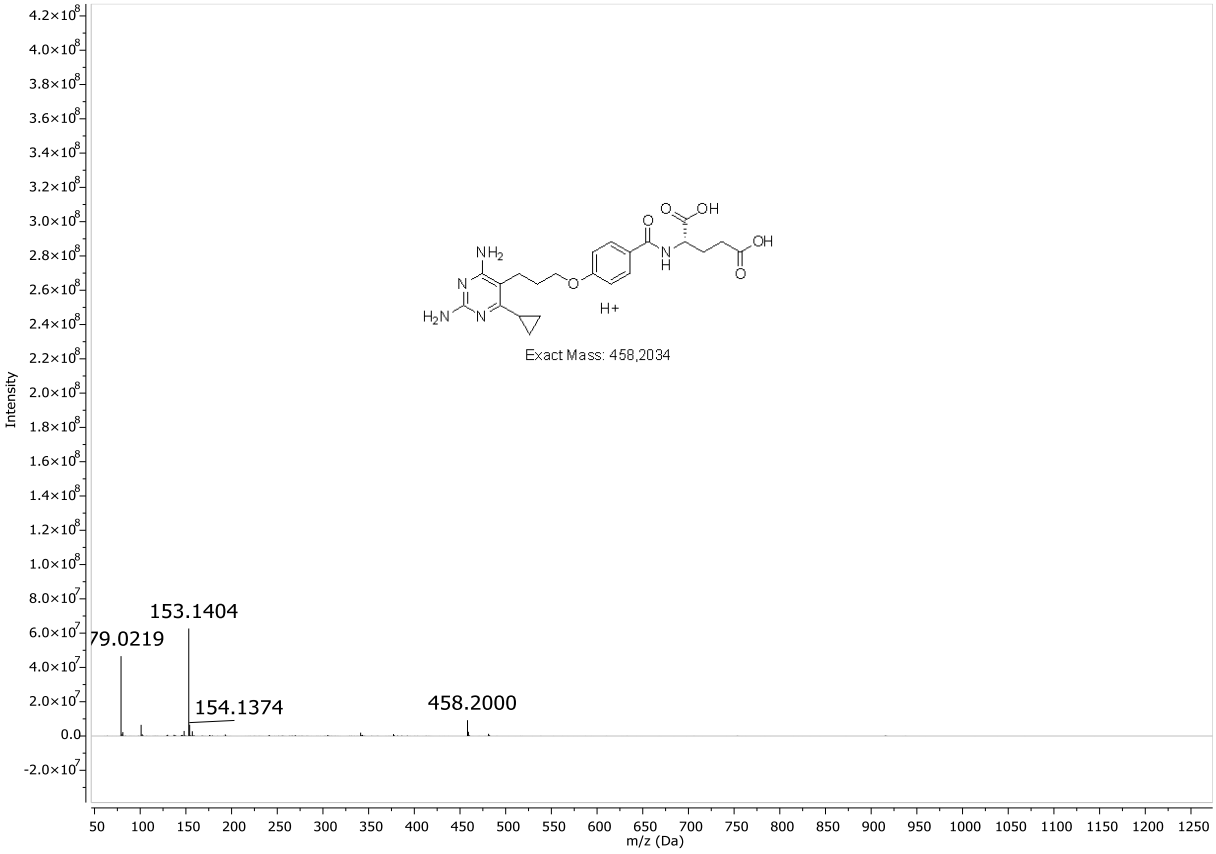
N-acylisourea bioproduct from of amide coupling of **21** with DCC.



Benzoyl azide intermediate **22** from reaction of **21** with DPPA.



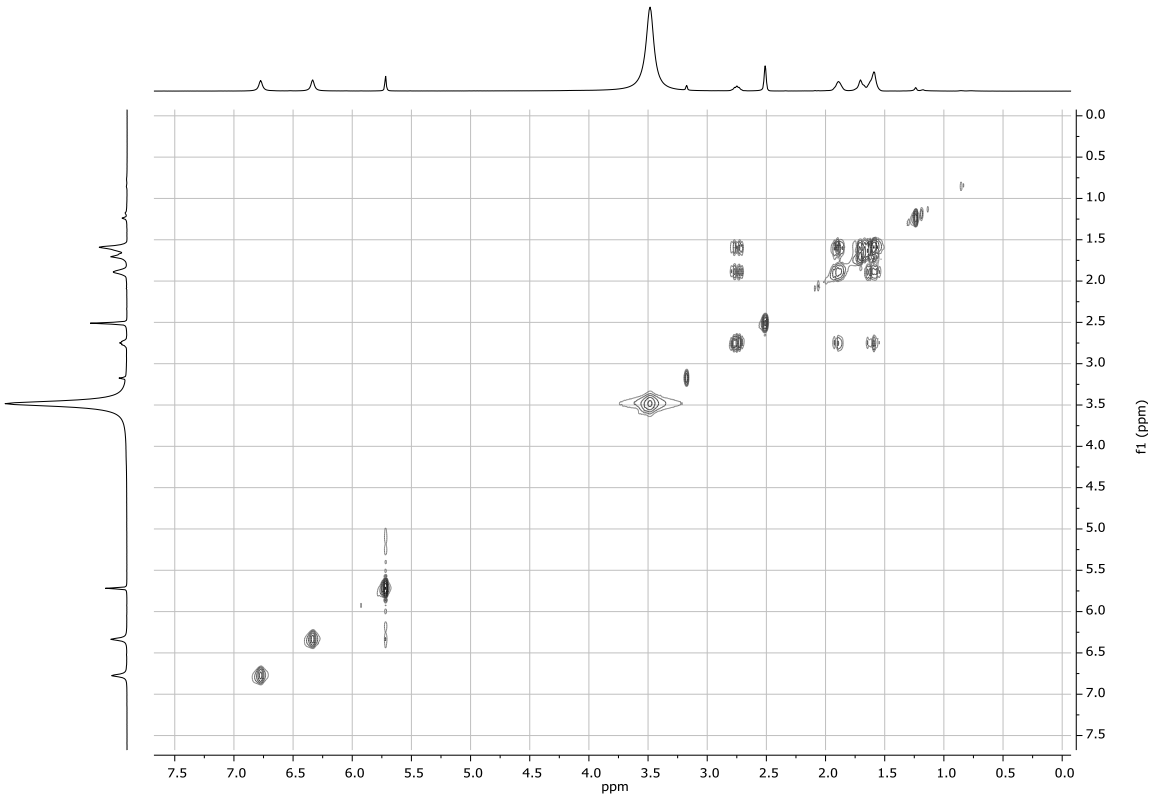
Amide product from reaction of benzoyl azide intermediate **22** with glutamic acid, showing persistent DBU impurity. $[M+H]^+$ for DBU expected at m/z value of 153.1386.



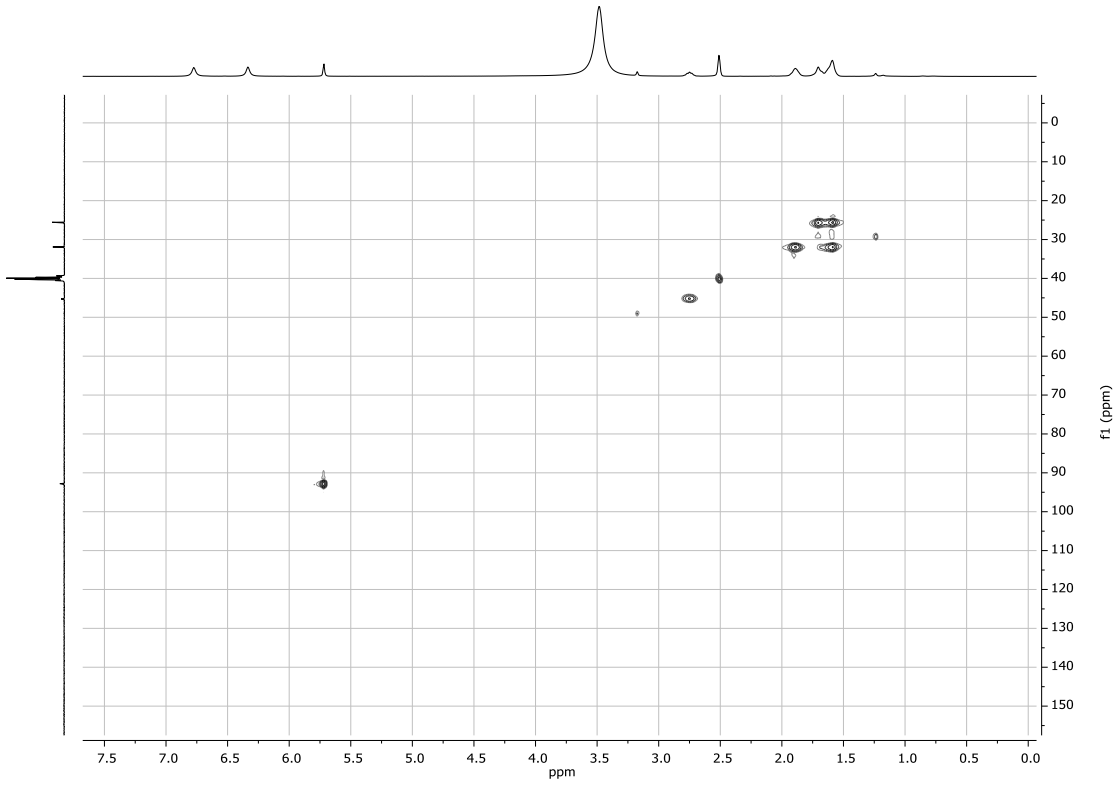
2D NMR spectra of representative compounds

6a

COSY:

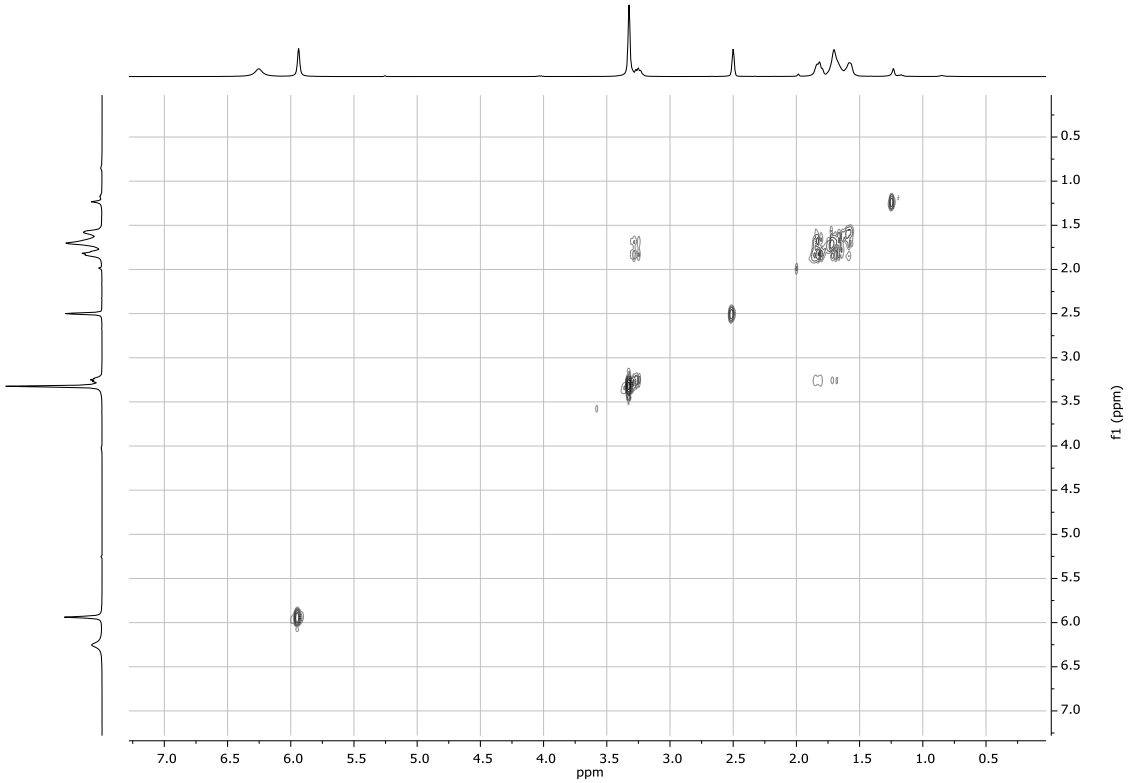


HSQC:

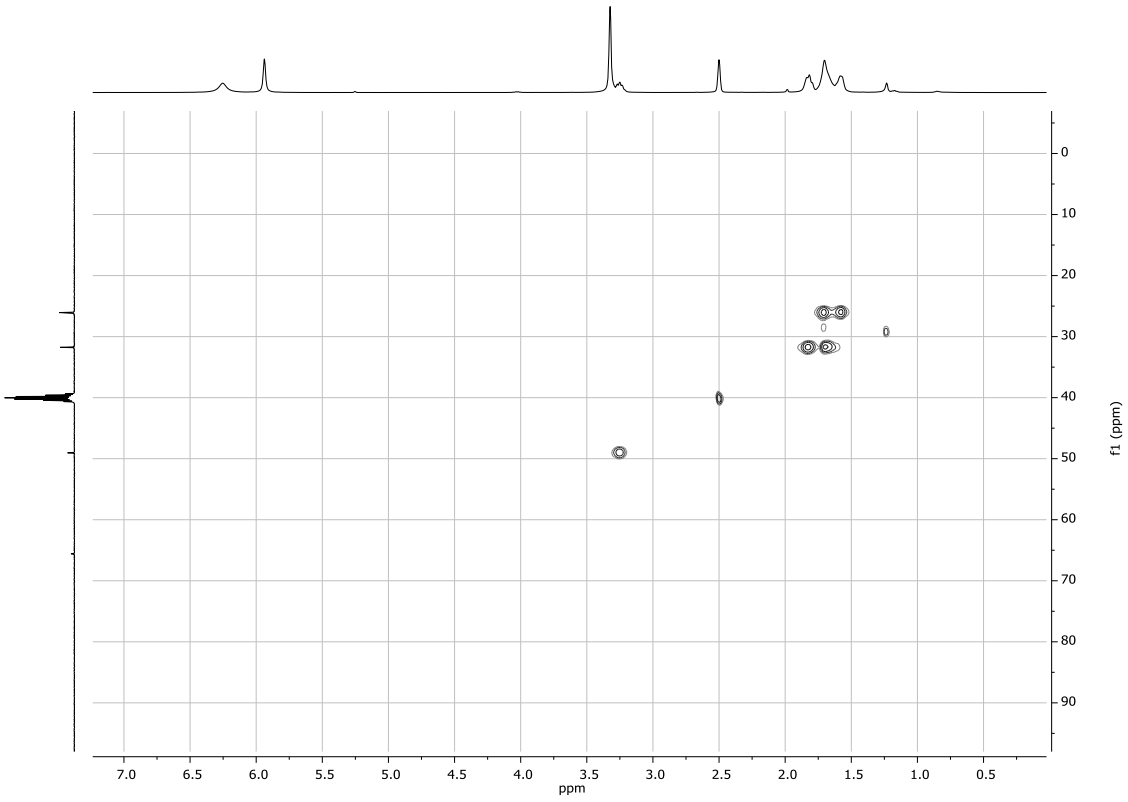


7a

COSY:

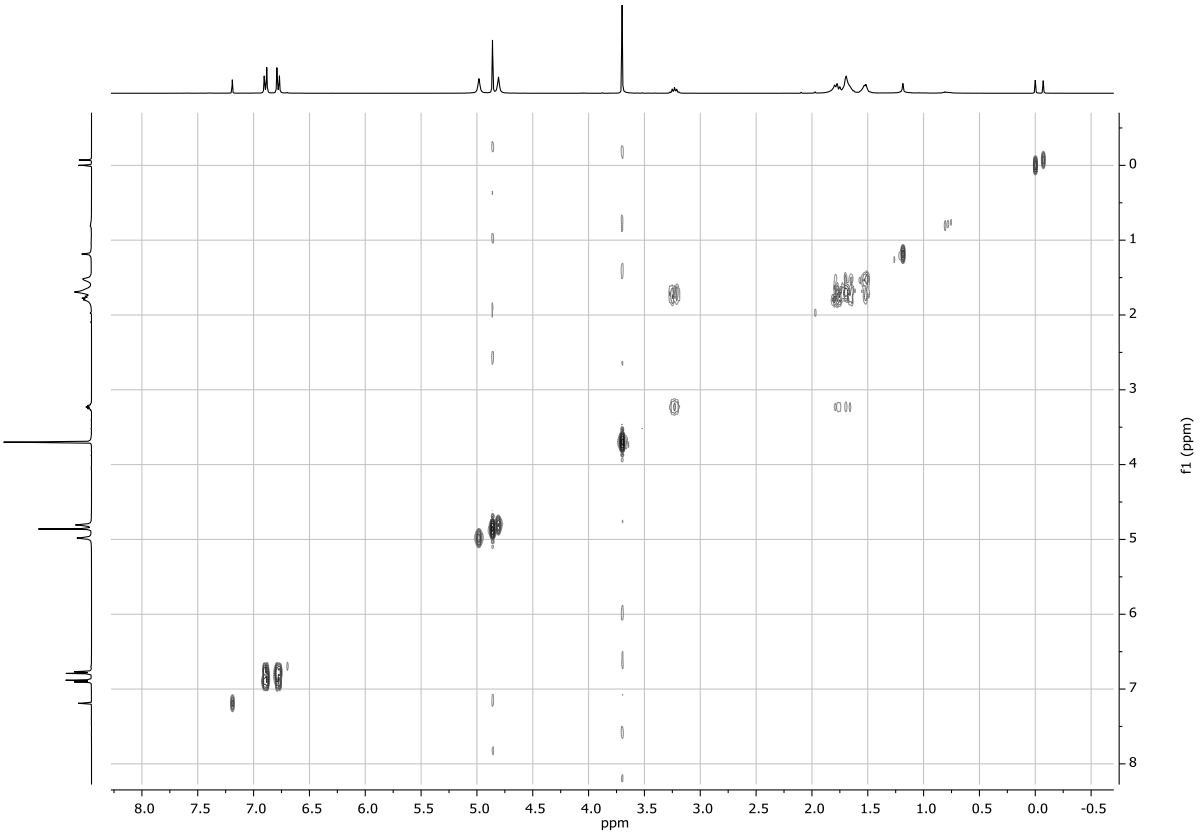

0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
f1 (ppm)
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
ppm

HSQC:

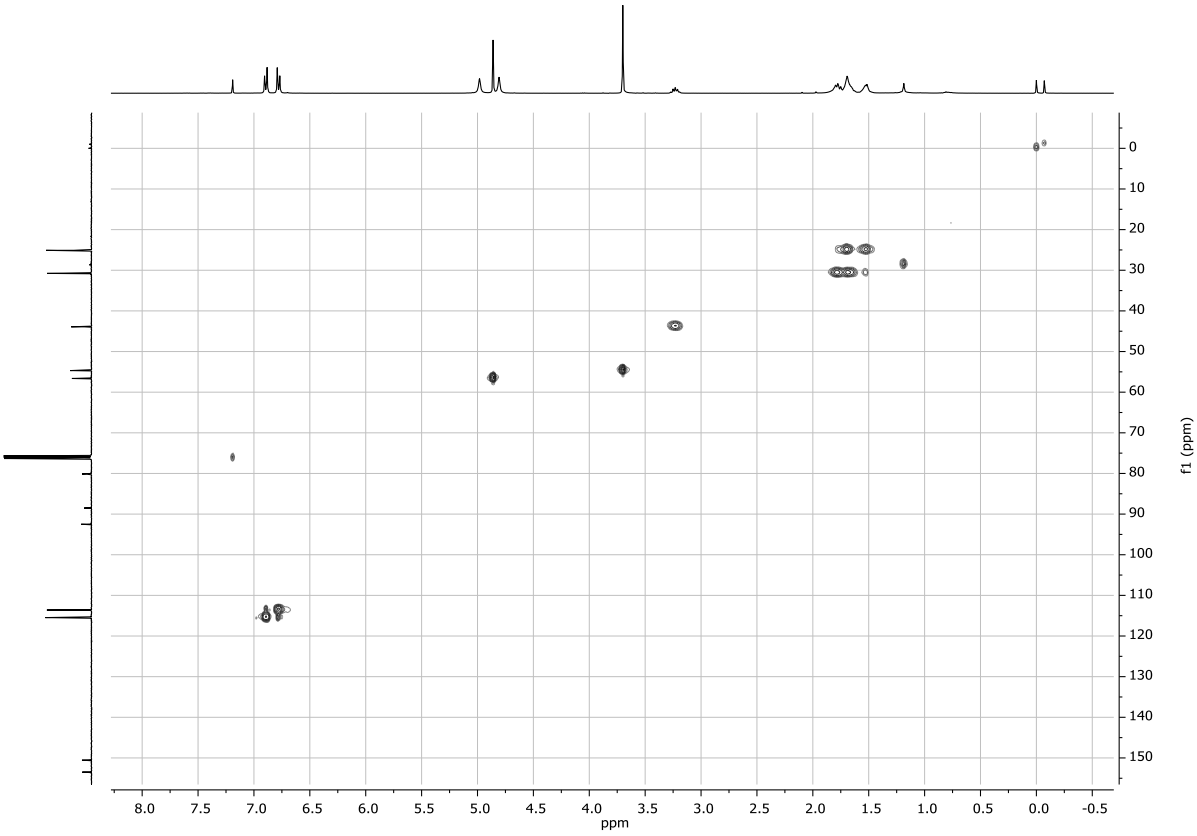

0
10
20
30
40
50
60
70
80
90
f1 (ppm)
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
ppm

8g

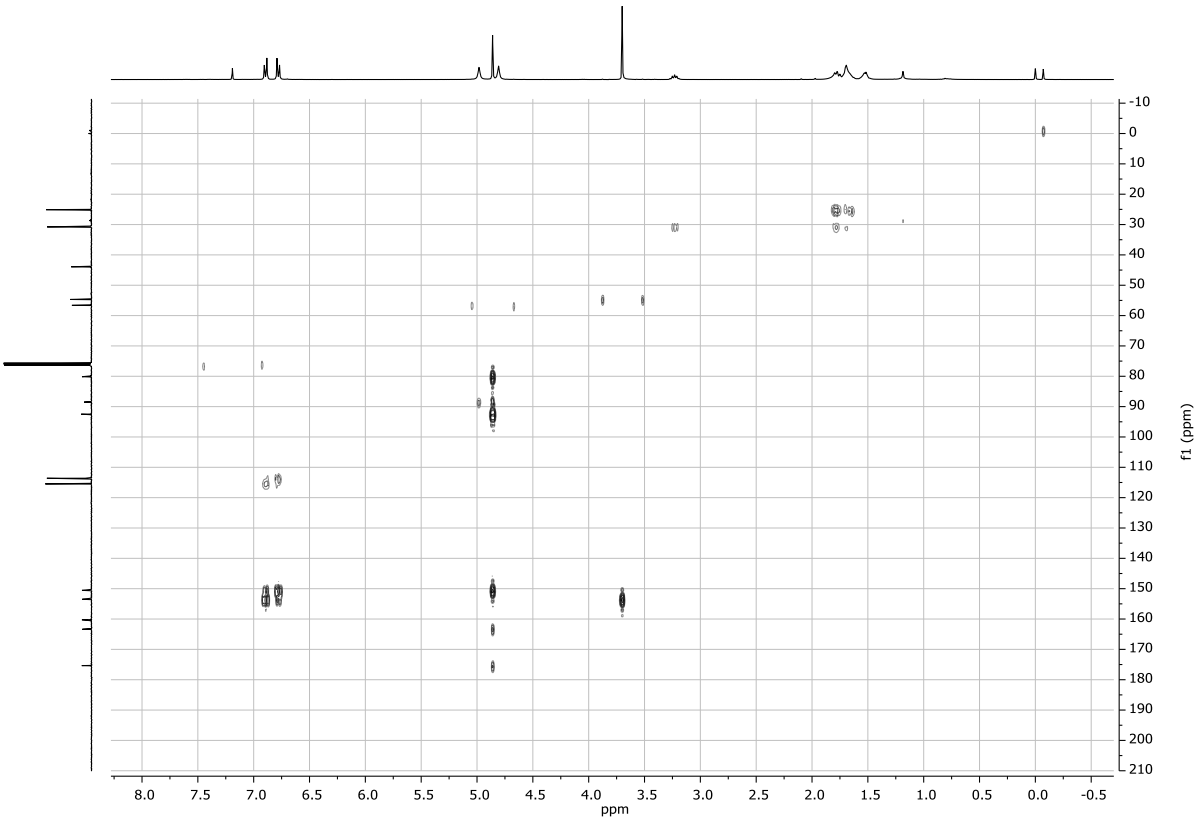
COSY:


8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
ppm
0
1
2
3
4
5
6
7
8
f1 (ppm)

HSQC:

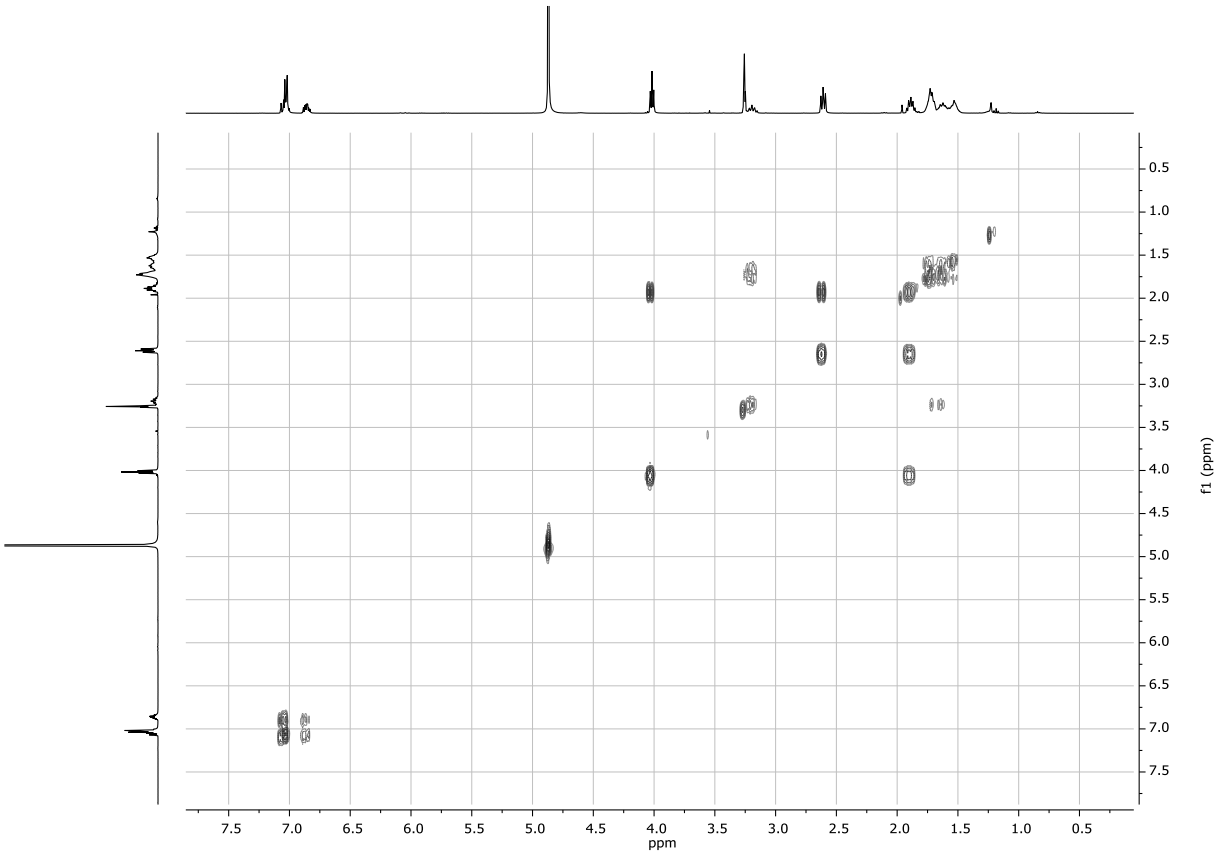


HMBC:

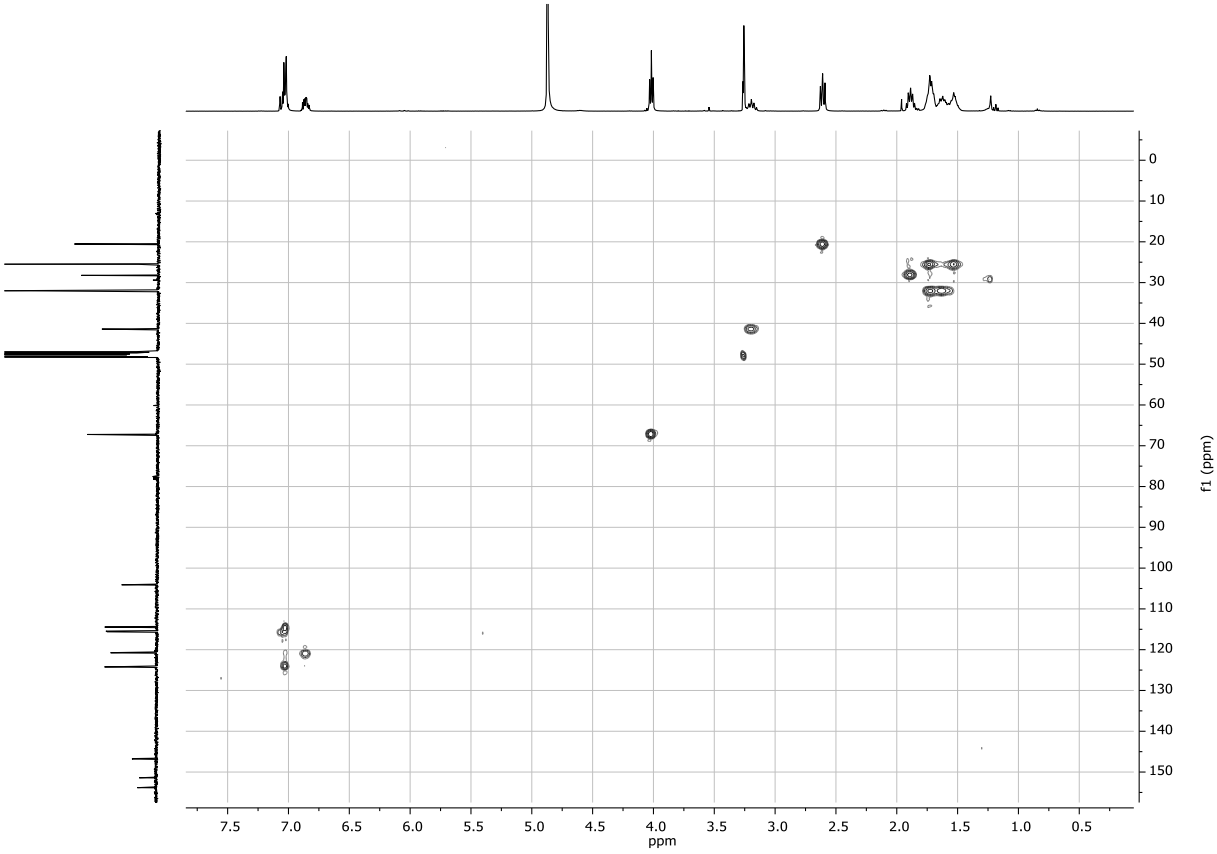


9e

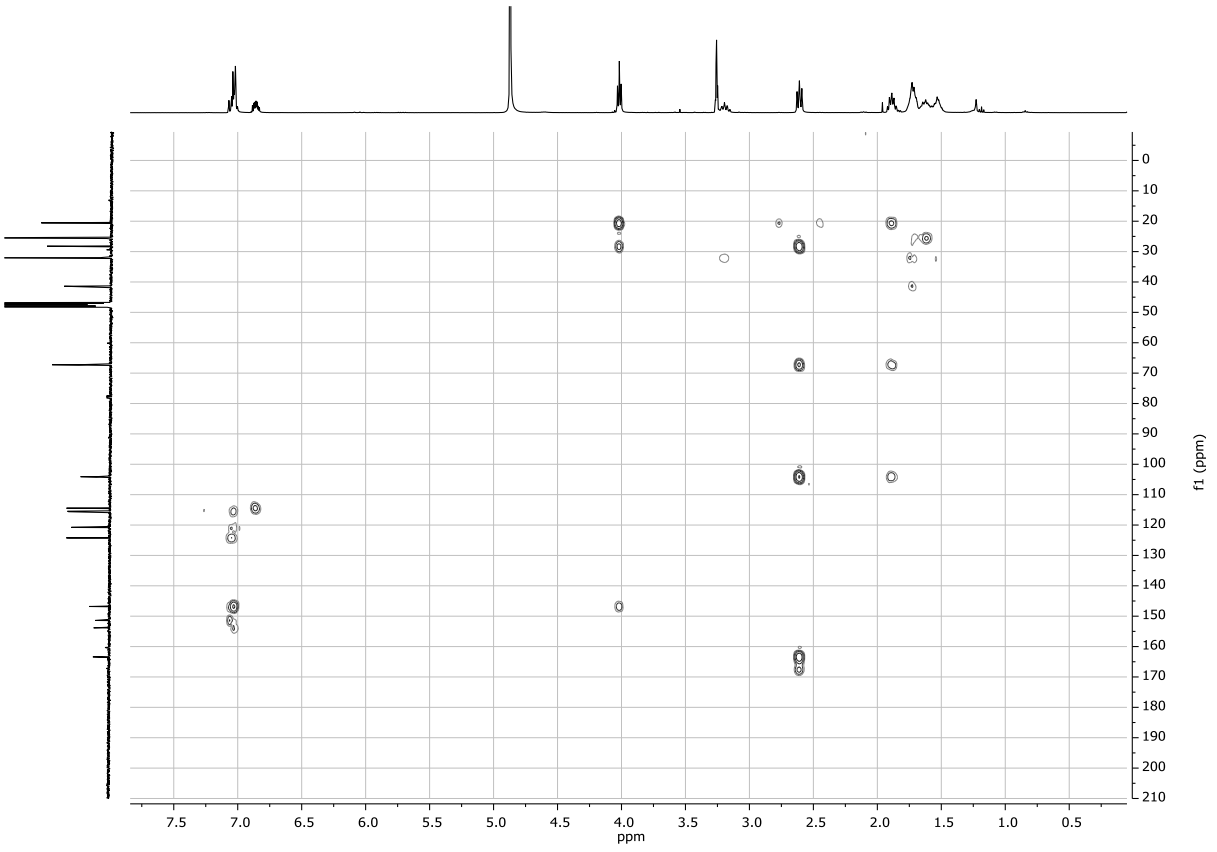
COSY:


7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
ppm
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
f1 (ppm)

HSQC:

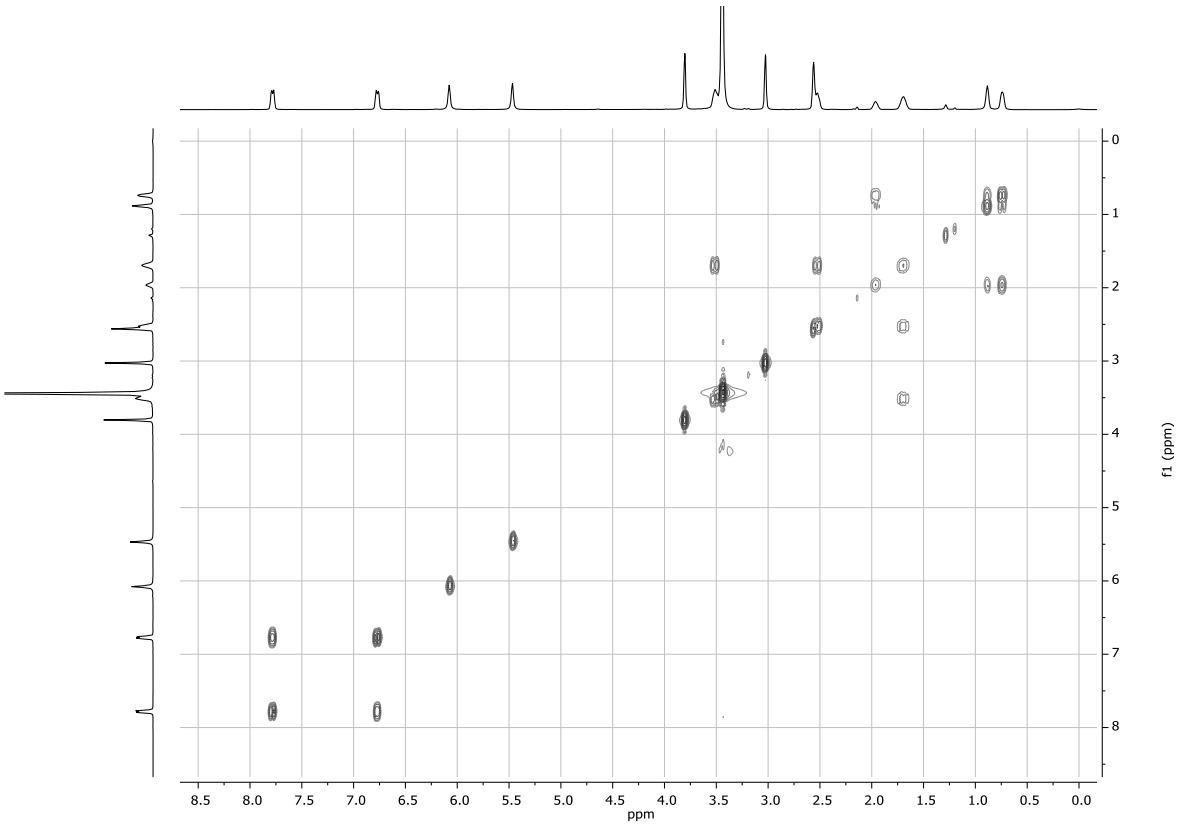

0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
f1 (ppm)
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
ppm

HMBC:


0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210
f1 (ppm)
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
ppm

18e

COSY:



HSQC:

