

https://dspace.org/

# Elucidating the Structure-Function Relationships of Enterococcus faecium Nicotinate-Nucleotide Adenylyltransferase through X-Ray Crystallography, Computational Modelling and Binding Studies

**Jeje, Olamide Adetomi**

**2024**

University of the Witwatersrand, Johannesburg

https://hdl.handle.net/10539/42156

Jeje, Olamide Adetomi . (2024). Elucidating the Structure-Function Relationships of Enterococcus faecium Nicotinate-Nucleotide Adenylyltransferase through X-Ray Crystallography, Computational Modelling and Binding Studies [PhD thesis, University of the Witwatersrand, Johannesburg]. WireDSpace.

Downloaded from a DSpace repository. https://dspace.org/




# Elucidating the Structure-Function Relationships of *Enterococcus faecium* Nicotinate-Nucleotide Adenylyltransferase through X-Ray Crystallography, Computational Modelling and Binding Studies

by

**Olamide Adetomi Jeje**

**(2219632)**

**A Thesis**

Submitted in fulfilment of the requirements for the degree of

**Doctor of Philosophy**

in

**Molecular and Cell Biology**

in the Faculty of Science, University of the Witwatersrand, Johannesburg, South Africa

**Supervisor**: Prof. Ikechukwu A. Achilonu

**Co-supervisor**: Dr Ramesh Pandian

June 2024

# DECLARATION

I declare that this dissertation is my own unaided work. It is being submitted for the Degree of Doctor of Philosophy at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.



Olamide A. Jeje

___5th___day of ___June___2024.

# ABSTRACT

Nicotinate nucleotide adenylyltransferase (NNAT) is a vital enzyme at the heart of NAD biosynthesis, catalysing a crucial reaction that leads to the formation of pyridine dinucleotides. $NAD^+$ is an essential coenzyme in numerous metabolic processes, DNA repair, and cellular signalling. Given its pivotal role, NNAT has emerged as a compelling drug target, particularly for its potential to disrupt the survival mechanisms of bacterial pathogens. By inhibiting NNAT, it is possible to undermine the metabolic integrity of these pathogens, making NNAT a promising focal point in the fight against bacterial infections and antibiotic resistance. However, understanding the structure-function relationship of *Enterococcus faecium* NNAT (EfNNAT) has remained elusive. Hence, this study aimed to address this gap by characterising EfNNAT and validating its potential as a druggable target. EfNNAT was overexpressed and purified using the *Escherichia coli* system and IMAC purification technique. Subsequently, biophysical characterisation was performed, followed by the determination of the three-dimensional structure in both apo and liganded forms using X-ray crystallography. High-throughput virtual screening, along with SP and XP docking, was conducted using a library of synthesizable flavonoids. Molecular dynamic simulation and fluorescence studies were employed to establish and validate the binding of identified inhibitors to EfNNAT. Successful expression and purification of EfNNAT yielded approximately 101 mg per 7.8 g of wet *E. coli* cells, with a purity exceeding 98%. High-resolution crystal structures of EfNNAT in native, adenine-bound, and NMN-bound forms were determined at 1.90 Å, 1.82 Å, and 1.84 Å, respectively. These structures provided insights into EfNNAT's substrate preference and revealed a potential allosteric site at the dimer interface of the NMN-bound structure. Virtual screening identified quercetin 3-O-beta-D-glucose-7-O-beta-D-gentiobioside as the only potential inhibitor from the flavonoid library used. A 500 ns atomistic molecular dynamics simulation showed the compound interacted through hydrogen bonding and water bridges, albeit unstable within the

receptor. ANS and mant-ATP fluorescence spectroscopy confirmed quercetin binding, while thermal shift assay revealed minimal impact of the inhibitor on the protein stability and structure. This study establishes a pipeline from expression and purification to structure solution and potential inhibitor identification for EfNNAT, validating its druggability. The mechanistic insights offer a foundation for advancing drug discovery efforts targeting EfNNAT and other bacterial NNAT enzymes.

# RESEARCH OUTPUTS

## 1. Published research work

**Jeje O.**, Pandian R., Sayed Y., and Achilonu I. (2023). Obtaining high yield recombinant *Enterococcus faecium* nicotinate nucleotide adenylyltransferase for X-ray crystallography and biophysical studies.

*International Journal of Biological Macromolecules*, 250, 126066.

Impact Factor 8.2

https://doi.org/10.1016/j.ijbiomac.2023.126066

**Contribution:**

Olamide Jeje: Methodology, formal analysis, investigation, visualisation, data curation, writing - original draft, review and editing.

Ramesh Pandian: Methodology, visualisation, supervision, data curation, formal analysis, review and editing.

Yasien Sayed: Review and editing, funding acquisition.

Ikechukwu Achilonu: Conceptualisation, methodology, visualisation, resources, review and editing, supervision, project administration, funding acquisition.

*Appendix A: Pdf screenshot of the publication*

## 2. Review paper

**Jeje O.**, Otun S., Aloke C., and Achilonu I. (2024). Exploring $NAD^+$ metabolism and NNAT: Insights from structure, function, and computational modelling.

*Biochimie* 220 (2024) 84-98.

Impact Factor 3.9

https://doi.org/10.1016/j.biochi.2024.01.002

**Contribution:**

Olamide Jeje: Writing - original draft, review and editing

Sarah Otun: Contributed to writing and editing

Chinyere Aloke: Review and editing

Ikechukwu Achilonu: Supervisior

*Appendix B: Pdf screenshot of the publication*

3. **Manuscript research work (Submitted)**

**Jeje O.** and Achilonu I. (2024). Exploring flavonoids as potential inhibitors for *Enterococcus faecium* nicotinate nucleotide adenylyltransferase: An integration of *in silico* with empirical studies.

*Journal of Molecular Structure*.

Impact Factor 3.8

**Contribution:**

Olamide Jeje: Methodology, formal analysis, investigation, visualisation, data curation, writing - original draft, review and editing.

Ikechukwu Achilonu: Conceptualisation, methodology, visualisation, resources, review and editing, supervision, project administration, funding acquisition.

*Appendix C: Pdf screenshot of the submitted work*

4. **Conference Proceedings**:

1. **Poster Presentation:** Obtaining high yield recombinant *Enterococcus faecium* nicotinate nucleotide adenylyltransferase for X-ray crystallography (MBRT 2023 Research Symposium, **Dec 2023**)
2. **Oral Presentation:** Obtaining high yield recombinant *Enterococcus faecium* nicotinate nucleotide adenylyltransferase for X-ray crystallography (African Physical Society Conference, **Nov 2023**)

To God Almighty, the giver of life, health, strength, inspiration, and wisdom.

# ACKNOWLEDGEMENT

Throughout this research, I have received tremendous support and assistance.

Firstly, I express my deepest gratitude to my supervisor, Professor Ikechuckwu Achilonu, for his extensive personal and professional guidance throughout this research. His invaluable expertise and constructive criticism were a massive motivation and catalyst to the successful completion of this thesis. I can only say a big thank you for your patience, understanding, encouragement, mentorship, availability and, finally, for believing in me. Your influence has left an indelible mark on my academic and professional life. Thank you for being an exceptional mentor.

My heartfelt appreciation goes to my co-supervisor, Dr Ramesh Pandian, for his assistance and expertise during some difficult phases of my research. I am truly fortunate to have had the opportunity to learn from you and to benefit from your profound knowledge and experience. Thank you for investing your time and effort in my development.

I also extend my unreserved acknowledgement to everyone in the Protein Structure-Function Research Unit (PSFRU), particularly Professor Yasien Sayed, the unit head and Dr Sylvia Fanucchi, for creating a conducive environment for learning and research. Special thanks to Olalekan Onisuru and Reabetswe Maake for their invaluable support and assistance within and outside the lab.

Special thanks are due to the MRC research grant for the financial support provided and to the Centre for High Performance Computing (CHPC) and the Council for Scientific and Industrial Research (CSIR) for access to the Schrödinger molecular modelling software (Maestro), which was a significant part of my project.

Finally, I express my sincere appreciation to my family; your love and support are inspiring. A special thank you to my wonderful parents for your prayers and

encouragement. I could not have completed this dissertation without the support of the most exceptional man in my life, my husband, who supported my career goals and provided financial and emotional support to pursue these goals. Thank you for believing in me. To my amazing children, Folayemi and Motolani, thank you for your patience and understanding. You provide unending inspiration.

Most importantly, I express my eternal gratitude to God almighty for the successful journey towards a PhD.

- **Table of Contents**

**DECLARATION** ..... ii

**ABSTRACT** ..... iii

**RESEARCH OUTPUTS** ..... v

**ACKNOWLEDGEMENT** ..... viii

**LIST OF FIGURES** ..... xv

**LIST OF TABLES** ..... xviii

**LIST OF ABBREVIATIONS** ..... xix

**Chapter 1: INTRODUCTION** ..... **1**

**1.1 Problem Statement** ..... **1**

**1.2 Rationale** ..... **3**

**1.3 Aims and Objectives** ..... **4**

**1.4 Thesis Structure** ..... **5**

**Chapter 2: LITERATURE REVIEW** ..... **8**

**2.1 *Enterococcus faecium* Infections and Antibiotic Resistance** ..... **8**

**2.1.1 *Enterococcus faecium*** ..... **8**

**2.1.2 Clinical significance of *Enterococcus faecium* infections** ..... **8**

**2.1.3 *Enterococcus faecium* prevalence** ..... **9**

**2.1.4 Mechanisms of antibiotic resistance in *Enterococcus faecium*** ..... **10**

**2.1.5 Current treatments and their mechanism of action against *E. faecium* bacterial infections** ..... **11**

**2.1.6 Impact of antibiotic resistance on treatment outcomes** ..... **13**

**2.2 The Structure and Function of Nicotinate-Nucleotide Adenylyltransferase (NNAT)** ..... **14**

**2.2.1 NNAT** ..... **14**

**2.2.2 Role of NNAT in bacteria $NAD^+$ biosynthesis.....15**

**2.2.3 NNAT role in human $NAD^+$ synthesis .....17**

**2.2.4 Importance of NNAT in bacterial physiology .....17**

**2.2.5 Structure of NNAT .....18**

**2.2.6 Comparison between bacterial and human NMNAT.....19**

**2.2.7 Overview of EfNNAT .....23**

**2.2.8 Studies suggesting inhibition of NNAT .....23**

**2.3 Protein Structure Determination .....24**

**2.3.1 X-ray crystallography .....24**

**2.3.2 Nuclear magnetic resonance (NMR) spectroscopy.....25**

**2.3.3 Cryo-electron microscopy (cryo-EM).....26**

**2.3.4 Small-angle X-ray scattering (SAXS).....26**

**2.3.5 Mass spectrometry (MS) .....27**

**2.4 X-Ray Crystallography Studies.....27**

**2.4.1 Methodology of X-ray crystallography studies.....28**

**2.4.2 Protein crystallisation and setting-up .....29**

**2.4.3 X-ray data collection and analysis .....30**

**2.4.4 Phasing.....30**

**2.4.5 Model building and structure refinement .....32**

**2.4.6 Structure validation and deposition.....35**

**2.4.7 X-ray crystallography studies of NNAT .....36**

**2.5 Computational Modelling .....36**

**2.5.1 Homology modelling .....37**

**2.5.2 Molecular docking.....38**

**2.5.3 Molecular dynamic (MD) simulation .....38**

**Chapter 3: Obtaining High Yield Recombinant *Enterococcus faecium* Nicotinate Nucleotide Adenylyltransferase for X-Ray Crystallography and Biophysical Studies** ..... **40**

**3.1 Highlights** ..... **41**

**3.2 Abstract** ..... **41**

**3.3 Graphical Abstract** ..... **42**

**3.4 Introduction** ..... **43**

**3.5. Materials and Methods** ..... **45**

**3.5.1 EfNNAT vector construct and overexpression** ..... **45**

**3.5.2 Recombinant EfNNAT purification** ..... **47**

**3.5.3 Dual enzyme-coupled assay** ..... **47**

**3.5.4 EfNNAT secondary structure analysis** ..... **48**

**3.5.5 Spectroscopy studies with fluorescent dyes** ..... **48**

**3.5.6 Thermal shift assays** ..... **49**

**3.5.7 EfNNAT quaternary structure determination** ..... **50**

**3.3.8 Crystallisation of EfNNAT** ..... **50**

**3.5.9 EfNNAT crystal data collection and structure solution** ..... **51**

**3.6 Results** ..... **52**

**3.6.1 Expression and purification of recombinant EfNNAT** ..... **52**

**3.6.2 Enzyme activity assay** ..... **53**

**3.6.3 Secondary structure analysis** ..... **54**

**3.6.4 Tertiary structure analysis** ..... **56**

**3.6.5 Structural stability study** ..... **57**

**3.6.6 Quaternary structure analysis** ..... **59**

**3.6.7 Three-dimensional structures of apo and adenine-bound EfNNAT** ..... **59**

**3.7 Discussion** ..... **66**

**Chapter 4: Co-Crystallisation of *Enterococcus faecium* NNAT with NMN and CTP: Crystal Structure and Preliminary X-Ray Analysis.** ..... **74**

**4.1 Introduction** ..... **75**

**4.2 Materials and methods** ..... **76**

**4.2.1 EfNNAT expression and purification** ..... **76**

**4.2.2 Crystallisation** ..... **76**

**4.2.3 Data collection and processing** ..... **76**

**4.2.4 EfNNAT-NMN structure solution** ..... **77**

**4.3 Results** ..... **77**

**4.3.1 EfNNAT co-crystallisation and crystallographic analysis** ..... **77**

**4.3.2 EfNNAT-CTP structure determination and refinement** ..... **79**

**4.3.3 Overview of EfNNAT-NMN complex structure** ..... **79**

**4.3.4 Structural flexibility upon NMN binding** ..... **81**

**4.3.5 Ligand interaction analysis** ..... **82**

**4.3.6 Dimerisation interface analysis** ..... **83**

**4.4 Discussion** ..... **86**

**Chapter 5: Exploring Flavonoids as Potential Inhibitors for *Enterococcus faecium* Nicotinate Nucleotide Adenylyltransferase: An Integration of *In Silico* and Empirical Studies** ..... **90**

**5.1 Abstract** ..... **91**

**5.2 Graphical Abstract** ..... **92**

**5.3 Introduction** ..... **93**

**5.4 Materials and Methods** ..... **96**

**5.4.1 Materials** ..... **96**

**5.4.2 Ligand library generation and preparation ..... 97**

**5.4.3 Protein preparation and refinement ..... 97**

**5.4.4 Cavity analysis ..... 98**

**5.4.5 Virtual screening and molecular docking ..... 99**

**5.4.6 Molecular dynamic simulation and data analysis ..... 100**

**5.4.7 Physiochemical properties of identified compounds ..... 100**

**5.4.8 8-Anilino-1-naphthalenesulfonic acid (ANS) fluorescence spectroscopy ..... 101**

**5.4.9 2'/3'-O-(N-Methyl-anthraniloyl)-adenosine-5'-triphosphate (mant-ATP) fluorescence spectroscopy ..... 102**

**5.4.10 SYPRO Orange thermal shift assay ..... 102**

**5.5 Results ..... 103**

**5.5.1 Energy minimised Vs MD simulated model ..... 103**

**5.5.2 Active site analysis of the EfNNAT modelled structures ..... 103**

**5.5.3 Virtual screening and compound identification ..... 106**

**5.5.4 MDS analysis ..... 110**

**5.5.5 Physiochemical properties and druglikeness of identified compounds 121**

**5.5.6 Empirical validation of identified compounds ..... 122**

**5.6 Discussion ..... 126**

**Chapter 6: CONCLUSION ..... 132**

**References ..... 134**

**Appendices ..... 148**

## LIST OF FIGURES

Figure 1.1 Map showing the global projection of 10 million deaths by 2050 attributed to antimicrobial resistance. ..... 2

Figure 1.2. Pictorial representation of workflow and methodology ..... 5

Figure 2.1. The reaction catalysed by nicotinate/nicotinamide mononucleotide adenylyltransferase. ..... 14

Figure 2.2. Schematic representation of the de novo and salvage synthesis of $NAD^+$ in bacteria and humans. ..... 16

Figure 2.3. A schematic representation of linked metabolic pathways involving the redox activity of NAD+/NADH to generate energy. ..... 19

Figure 2.4. The sequence alignment of NNAT from human and bacteria species showing variation in the amino acid sequences. ..... 20

Figure 2.5. Ribbon representation showing the difference between *E. coli* NNAT and human NMNAT protomer structure. ..... 21

Figure 2.6. Predicted structure of EfNNAT generated via SwissModel. ..... 23

Figure 2.7. Overview of X-ray crystallography methods utilised for protein structure determination. ..... 28

Figure 2.8. Outline of the different stages involved in the MD simulation of a protein. ..... 39

Figure 3.1. A pictorial representation of the vector construct used for expressing the recombinant EfNNAT. ..... 46

Figure 3.2. Expression and purification profile of recombinant *E. faecium* NNAT ..... 53

Figure 3.3 The activity profile of recombinant EfNNAT at different pH ..... 54

Figure 3.4. Far-UV circular dichroism spectra of EfNNAT ..... 55

Figure 3.5. Fluorescence emission spectra of (A) ANS interaction with EfNNAT apo (red solid), EfNNAT denatured (red dotted), and EfNNAT-ATP complex (blue).. ... 57

Figure 3.6. The fluorescence thermal unfolding profile of recombinant EfNNAT apo (red) and in the presence of ATP (blue). ..... 58

Figure 3.7. Size Exclusion-HPLC analysis of EfNNAT. ..... 59

Figure 3.8. Images of some high-quality diffractive crystals of EfNNAT obtained under different crystallisation screen conditions ..... 61

Figure 3.9. Overview of EfNNAT structures and comparison with gram-positive and negative bacterial NNAT ..... 63

Figure 3.10. Amino acid sequence alignment of NNAT from *E. faecium*, *B. anthracis*, *B. subtilis*, *M. tuberculosis*, *S. aureus*, *E. coli* and *P. aeruginosa* ..... 64

Figure 3.11. Stereo image of EfNNAT structural flexibility and the *Fo-Fc* electron density map showing the active site interaction ..... 65

Figure 3.12. Sequence alignment of seven *Enterococcus spp* ..... 66

Figure 4.1. The Crystal structure of *E. faecium* NNAT bound to NMN ..... 80

Figure 4.2 Structural alignments of EfNNAT apo, adenine-bound and NMN-bound ..... 82

Figure 4.3. The NMN interaction plot shows its stabilisation within the dimer interface ..... 85

Figure 4. 4 Comparative analysis of the dimerisation interfaces of EfNNAT. ..... 86

Figure 5.1. Outline of the workflow employed to identify potential inhibitors of *E. faecium* NNAT ..... 96

Figure 5.2. Structural comparison of EfNNAT energy-minimised Vs MD simulated model ..... 105

Figure 5. 3. Surface and pictorial representation of EfNNAT minimised model (left) and the MD simulated (right) cavities. ..... 106

Figure 5.4. C$\alpha$-RMSD and RMSF of EfNNAT apo and in complex with inhibitors .112

Figure 5.5. EfNNAT radius of gyration showing the protein's compactness in its apo and inhibitor complex forms. ..... 113

Figure 5.6. Plots depicting quercetin (red) and apigenin (blue) dynamism within the EfNNAT receptor over 500 ns simulation time. ..... 115

Figure 5.7. Graphical representation of the number of intra-hydrogen bond formation .....116

Figure 5.8. 3D and 2D cluster analysis of EfNNAT in complex .....118

Figure 5.9. Side-chain interaction bar chart of EfNNAT .....119

Figure 5.10. A detailed 2D atomic interaction plot of EfNNAT-quercetin and EfNNAT-apigenin complexes .....120

Figure 5.11. The bioavailability radar of (A) quercetin and (B) apigenin showing how their physiochemical properties fit into the druglikeness descriptors.....122

Figure 5.12. ANS and mant-ATP Fluorescence emission spectra showing quercetin and apigenin binding to EfNNAT.....123

Figure 5.13 The fluorescence thermal denaturation spectra for the binding of EfNNAT .....125

## LIST OF TABLES

Table 3.1. Summary of EfNNAT expression and purification yield .....52

Table 3.2. EfNNAT crystal data collection and refinement statistics .....62

Table 4.1. EfNNAT Co-crystallisation data collection. .....78

Table 4.2. Summary of EfNNAT-NMN refinement statistics.....79

Table 4.3. Dimerisation interface analysis of EfNNAT apo, adenine-bound and NMN-bound .....84

Table 5.1. Details of EfNNAT minimised and MD simulated model cavity analysis .....105

Table 5.2. Five topmost ligands identified post-HTVS for EfNNAT minimised and MDS models.....108

Table 5.3. Comparative analysis of the five topmost ligands selected by the EfNNAT energy-minimised and MD-simulated model after HTVS docking. .....109

# LIST OF ABBREVIATIONS

| | |
|---|---|
| ANS | 8-anilino-1-naphthalenesulfonic acid |
| ATP | Adenosine triphosphate |
| CID | Compound ID |
| CTP | Cytidine triphosphate |
| EC | Enzyme class |
| EfNNAT | *Enterococcus faecium* nicotinate nucleotide adenylyltransferase |
| hNMNAT | Human NMNAT |
| HTVS | High throughput virtual screening |
| IFD | Induced-fit docking |
| IMAC | Immobilised Metal Affinity Chromatography |
| IPTG | Isopropyl β-D-1-thiogalactopyranoside |
| kDa | Kilodalton |
| mant-ATP | 2'/3'-O-(N-Methyl-anthraniloyl)-adenosine-5'-triphosphate |
| MDS | Molecular dynamic simulation |
| MM | Molecular mechanics |
| MMGBSA | Molecular mechanics-generalised Born surface area |
| MSA | Molecular surface area |
| NaAD | Nicotinate adenine dinucleotide |
| $NAD^+$ | Nicotinamide adenine dinucleotide |
| NadD | Nicotinate Mononucleotide Adenylyltransferase gene |
| NADS | NAD synthetase |
| NadE | NAD synthetase gene |

| | |
|---|---|
| NaMN | Nicotinate mononucleotide |
| NaMNAT | Nicotinate mononucleotide adenylyltransferase |
| NMN | Nicotinamide mononucleotide |
| NMNAT | Nicotinamide mononucleotide adenylyltransferase |
| NNAT | Nicotinate nucleotide adenylyltransferase |
| NPT | Number of molecules/atoms in assembly Pressure and Temperature |
| OD | Optical density |
| OPLS | Optimised Potentials for Liquid Simulations |
| PDB | Protein Data Bank |
| QM | Quantum mechanics |
| Rog | Radius of gyration |
| RMSD | Root means square deviation |
| RMSF | Root means square fluctuation |
| SASA | Solvent accessible surface area |
| SDS-PAGE | Sodium dodecyl sulphate polyacrylamide gel electrophoresis |
| SE- HPLC | Size exclusion high-pressure liquid chromatography |
| SP | Standard precision |
| TIP3 | Transferable intermolecular potential with three points |
| $T_m$ | Melting temperature |
| XP | Extended precision |

The IUPAC-IUBMB one and three letters code for amino acids are used.

# Chapter 1
# INTRODUCTION

## 1.1 Problem Statement

Antibiotic resistance has become a global health crisis, threatening our ability to treat bacterial infections effectively. Despite the widespread use of antibiotics, bacterial pathogens are evolving mechanisms to resist the effects of these drugs, rendering them ineffective. Hence, patient care is becoming more expensive due to the replacement of first-line antibiotics with costlier medications, and the prolonged duration of illness and treatment not only increases healthcare costs but also imposes a more significant economic burden on patients and societies [1, 2].

The ESKAPE pathogens (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter species*) represent a group of multidrug-resistant bacteria that pose a severe threat to public health [1, 3]. These pathogens are notorious for their ability to "escape" the effects of antibiotics, leading to difficult-to-treat infections that are associated with high morbidity and mortality rates. Estimates have shown that if actions are not taken regarding the increasing drug resistance arising [2], especially from this group of multidrug-resistant pathogens, by 2050, antibiotic resistance could result in deaths of over 10 million lives across the globe annually [4-6]. This alarming figure highlights the significant threat drug-resistant pathogens pose if current trends continue unchecked. Figure 1.1 shows the global distribution of the projected death rate due to antimicrobial drug resistance [7].

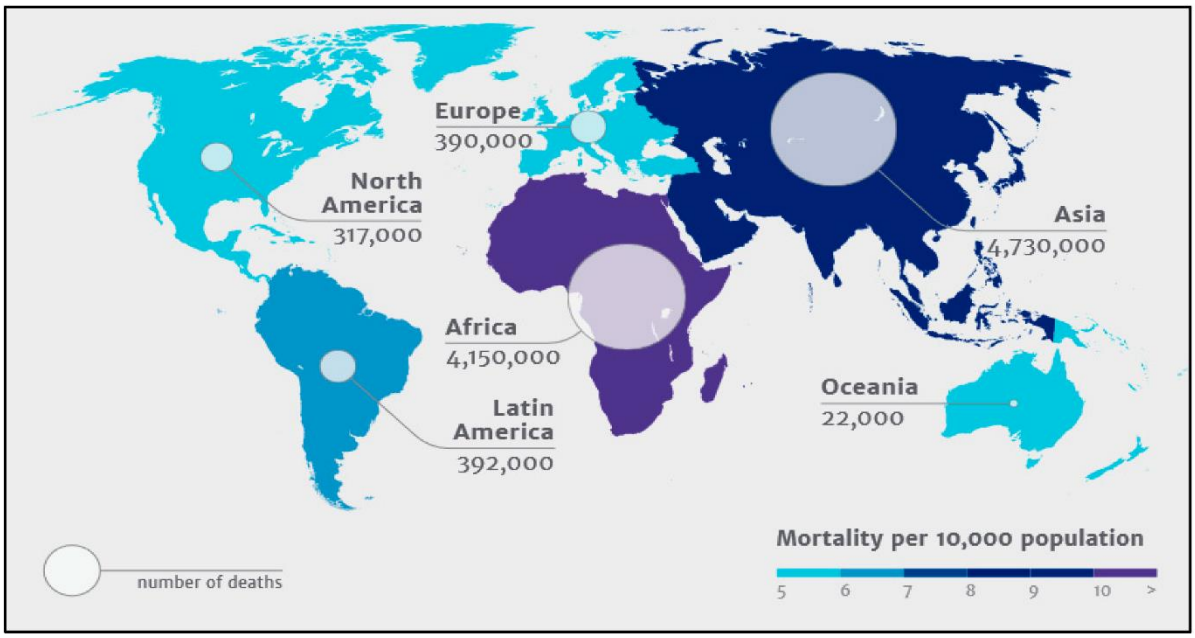



**Figure 1.1 Map showing the global projection of 10 million deaths by 2050 attributed to antimicrobial resistance.**
Source: Ananda *et al.*, 2022 (The Review on Antimicrobial Resistance – Tackling drug-resistant infections globally, The Wellcome Trust, The UK Department of Health, London, 2014) [7].

*Enterococcus faecium*, a Gram-positive bacterium, emerged as a significant clinical concern due to its increasing prevalence and notorious resistance to antibiotic treatments [8-11]. The pathogen substantially contributes to hospital-acquired infections, such as urinary tract infections, bloodstream infections, surgical site infections and endocarditis, a severe and often life-threatening infection of the heart valves [12, 13]. The ability of *E. faecium* to persist in the healthcare environment and survive on medical surfaces contributes to its role as a nosocomial superbug, facilitating the spread of antibiotic resistance [11, 14-16]. A critical challenge associated with *E. faecium* is its remarkable resistance to multiple antibiotics, including commonly used antibiotics such as vancomycin, ampicillin, gentamicin, ciprofloxacin, linezolid, and daptomycin [12, 17], which currently limits treatment options and underscores the urgency for alternative therapeutic strategies.

## 1.2 Rationale

Enzymes involved in essential biosynthetic pathways are often excellent candidates for therapeutic intervention because of target precision [18, 19], therefore minimising the risk of developing microbial resistance. Nicotinate nucleotide adenylyltransferase (NNAT), a pivotal enzyme in the nicotinamide adenine dinucleotide ($NAD^+$) biosynthesis pathway of *E. faecium*, emerges as a potential target for drug development, given its critical role in bacterial survival [20-23]. The biosynthesis of $NAD^+$ is a vital pathway for *E. faecium*, as it supplies the cell with a critical cofactor required for DNA repair, maintenance of cellular redox balance and the execution of numerous enzymatic reactions crucial for the energy metabolism of the bacterium [24-26]. Essentially, NNAT converts nicotinic acid mononucleotide (NaMN) into nicotinic acid adenine dinucleotide (NaAD), a precursor in the synthesis of $NAD^+$. It also facilitates the direct formation of $NAD^+$ from nicotinamide mononucleotide (NMN) but at a much slower rate [27].

Given the indispensable role of NNAT in the biosynthesis of $NAD^+$, inhibiting NNAT can lead to a deficiency in $NAD^+$ production, thus disrupting vital cellular processes and ultimately compromising bacterial survival [28, 29]. Additionally, studies have shown that inhibiting the NNAT enzyme has a potent bactericidal effect, as demonstrated by a reduction in viable cells over 1,000 times compared to the uninduced reference [29]. The fact that bacteria do not have an alternative source to $NAD^+$ underscores the central role of NNAT in the $NAD^+$ biosynthesis pathway and its significance for bacterial survival. Moreover, the human homologue enzyme exhibits sequence and structural variances, making it an attractive focal point for designing antimicrobial drugs to combat infections caused by this resilient pathogen.

This study focused on understanding the structure-function of the NNAT enzyme from *E. faecium* and validating it as a druggable target. Firstly, the protein was overexpressed, purified, and characterised biophysically and structurally. Then the three-dimensional structure of the protein was determined, and virtual screening was

performed to assess the druggability of the protein and identify potential inhibitors that could serve as therapeutic interventions.

## 1.3 Aims and Objectives

This study aimed to elucidate the structural and functional aspects of *E. faecium* NNAT using empirical and computational tools.

The secondary aims, which were structured as a publishable manuscript, and their objectives included:

1. To obtain highly pure recombinant EfNNAT for structure-function studies. This was achieved through the following objectives.
   1.1 Protein expression, purification, and quality assessment
   1.2 Activity assay by dual enzyme-coupled assay
   1.3 Secondary structure content determination by Far-UV Circular Dichroism
   1.4 Tertiary structure analysis via fluorescence spectroscopy and thermal shift assay
   1.5 Quaternary structure analysis via Size exclusion-high performance liquid chromatography
2. To obtain a high-resolution crystal structure of apo EfNNAT and liganded-EfNNAT
   2.1 Crystallisation of EfNNAT-apo and liganded
   2.2 X-ray diffraction and data collection
   2.3 Structure solution
   2.4 Structure validation and analysis
3. To use the high-resolution crystal structure for computational modelling studies towards the identification of potential inhibitors
   3.1 Ligand and protein preparation
   3.2 Virtual screening and molecular docking (HTVS, SP and XP)

3.3 Molecular Dynamic Simulation (MDS)

3.4 Physiochemical properties and druglikeness of identified compounds

4. To empirically validate the binding capacities of the *in-silico* derived inhibitors

4.1 Binding studies using fluorescence spectroscopy

4.2 Stability studies using thermal shift assay

The workflow employed is detailed in Figure 1.2.

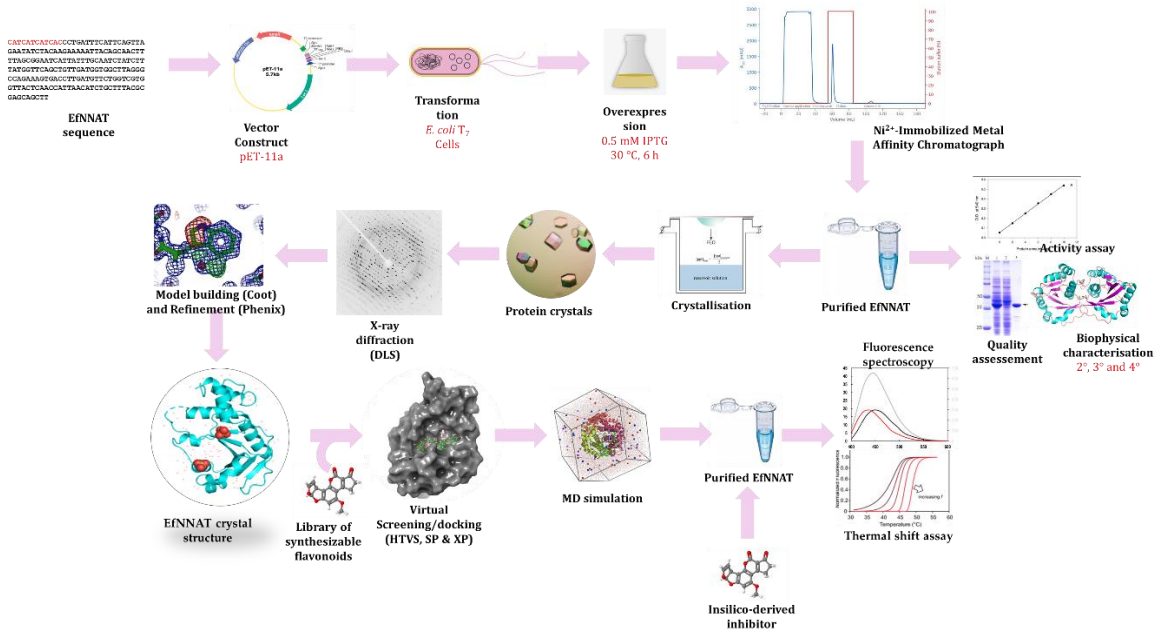



**Figure 1.2. Pictorial representation of workflow and methodology**

## 1.4 Thesis Structure

The thesis is structured as a publishable manuscript, and it is divided into six chapters.

**Chapter 1**

**Introduction:** This chapter introduces the research work, stating the problem statement, rationale, aims and objectives, and thesis outline.

**Chapter 2**

**Literature review:** This chapter offers a comprehensive review of *Enterococcus faecium* infections and antibiotic resistance, highlighting the challenges posed by this pathogen. It also delves into the physiological role of NNAT in bacteria, the structure and function of NNAT, focusing on *E. faecium* NNAT, and provides an overview of X-ray crystallography studies, including principles, methodology, and applications. Additionally, the chapter explores the role of computational modelling in this context, discussing techniques such as homology modelling, molecular docking, and molecular dynamic simulations. These computational methods are shown to be valuable in predicting and understanding biological processes at the atomic level, complementing experimental data and enhancing our knowledge of EfNNAT and its potential as a drug target.

**Chapter 3**

**Obtaining High-yield Recombinant *Enterococcus faecium* Nicotinate Nucleotide Adenylyltransferase for X-Ray Crystallography and Biophysical Studies:** This chapter presents a novel study on the expression, purification, biophysical characterisation, and X-ray structure determination of EfNNAT (both apo and adenine-bound forms). The development of an efficient protocol for expressing and purifying EfNNAT with high yield is described here, followed by detailed biophysical characterisation using various methods. The three-dimensional structure of EfNNAT is determined at high-resolution, providing valuable insights into its structural features. The chapter is presented in the required format for publication and represents the final version of the manuscript.

**Chapter 4**

**Co-Crystallisation of *Enterococcus faecium* NNAT with NMN and CTP: Crystal Structure and Preliminary X-Ray Analysis:** This chapter presents a detailed exploration of the co-crystallisation, X-ray diffraction, structure solution, and analysis

of EfNNAT in complex with NMN. Additionally, it includes an analysis of the X-ray diffraction data obtained from the co-crystals of EfNNAT with CTP. The co-crystallisation process was aimed at elucidating the structural details of the EfNNAT-NMN complex, providing insights into the binding interactions between EfNNAT and NMN.

**Chapter 5**

**Exploring Flavonoids as Potential Inhibitors for *Enterococcus faecium* Nicotinate Nucleotide Adenylyltransferase: An Integration of *In Silico* with Empirical studies:** This chapter capitalises on the high-resolution crystal structure of EfNNAT to conduct high-throughput virtual screening using a library of synthesizable flavonoids, aiming to identify potential inhibitors. Subsequently, a 500 ns molecular dynamic simulation of the selected ligands was carried out to analyse the structural stability and interaction of the complexes. Empirical studies were then employed to validate the binding of the inhibitors to EfNNAT. The findings of this chapter are presented as a publishable manuscript.

**Chapter 6**

This concluding chapter summarises and contextualises the key findings from the research presented in Chapters 3-5, highlighting their significance and contributions to the field of study.

# Chapter 2
# LITERATURE REVIEW

## 2.1 *Enterococcus faecium* Infections and Antibiotic Resistance

### 2.1.1 *Enterococcus faecium*

Until recent decades, Gram-positive *E. faecium* was classified as a low virulent pathogen [30] living as microflora in humans. [31]. However, this commensal role has evolved over the years to be pathogenic in humans due to excessive and improper administration of antibiotics, resulting in resistant strains. Furthermore, colonisation, extra-chromosome and mobile genetic elements, extracellular proteins biofilm formation, components cell wall, and adherence factors have also been identified as contributing factors to the ability of *E. faecium* to survive and cause infection in a host [32, 33]. Even though *E. faecium* are not naturally virulent as some other bacterial pathogens, the ability to acquire and exchange antibiotic resistance determinants, in addition to its innate resistance to many antibiotics selectively, provides *E. faecium* an added advantage in environments with heavy antibiotic usage such as hospitals, thereby allowing them to out-grow other species that would ordinarily keep them in check.

### 2.1.2 Clinical significance of *Enterococcus faecium* infections

Several nosocomial infections have been associated with *E. faecium*, including urinary tract infections, bloodstream infections, and endocarditis, with surgical wound infection accounting for its highest count [30, 34]. *E. faecium* is the second most prominent nosocomial-causing pathogen in ICU patients after *Staphylococcus aureus.* It is ranked as the third major cause of urinary tract infections and bloodstream infections [34-36]. The antimicrobial properties demonstrated by *E. faecium* are

characterised by the possession of plasmid-encoded resistance gene that allows them to transfer genetic information within themselves and between different species via horizontal gene transfer [30]. This characteristic, therefore, makes *Enterococci* infections persistent and increasingly challenging to treat [8-11].

Common infections caused by *E. faecium* include urinary tract infections, bloodstream infections, surgical site infections, and infective endocarditis [12, 13]. These infections can be particularly problematic in immunocompromised patients, elderly individuals, and those with underlying medical conditions. Furthermore, *E. faecium* infections are associated with increased morbidity, mortality, and healthcare costs, highlighting the clinical significance of this pathogen.

### 2.1.3 *Enterococcus faecium* prevalence

*E. faecium* is a globally prevalent pathogen, particularly in hospital settings, with increasing rates of antibiotic resistance posing significant challenges to public health. Its prevalence has increased dramatically over the past few decades, particularly due to its ability to acquire resistance to multiple antibiotics, including vancomycin, leading to the emergence of vancomycin-resistant *Enterococcus* (VRE). Records of high rates of VRE have been recorded in North America, southern and eastern Europe [37, 38], Asia [39-41], Australia [13] and New Zealand[42], with some reports in Africa [43-45].

Most *E. faecium* infections occur in hospital settings, particularly in intensive care units, where patients are more vulnerable due to invasive procedures and weakened immune systems. The widespread and often inappropriate use of antibiotics has also driven the selection of resistant *E. faecium* strains. This includes the use of vancomycin and other antibiotics in both human medicines. *E. faecium* is highly transmissible within healthcare settings and can spread via direct contact with contaminated surfaces, medical equipment, and the hands of healthcare workers. The ability of *E. faecium* to form biofilms on medical devices and surfaces has also

enhanced its survival and persistence in hospital environments, thus contributing to its prevalence [16].

The rise of VRE and other multidrug-resistant strains of *E. faecium* complicates treatment options and increases healthcare costs. It necessitates using last-resort antibiotics, which may have more severe side effects and lower efficacy. Effective infection control measures, including hand hygiene, environmental cleaning, and antimicrobial stewardship, are crucial to limiting the spread of *E. faecium* in healthcare settings [46]. Continuous surveillance and monitoring of *E. faecium* prevalence and resistance patterns are essential to inform treatment guidelines and public health strategies.

### 2.1.4 Mechanisms of antibiotic resistance in *Enterococcus faecium*

One of the major challenges in treating *E. faecium* infections is the development of antibiotic resistance. *E. faecium* is intrinsically resistant to many antibiotics, including cephalosporins, and has acquired resistance to others through horizontal gene transfer. The acquisition of resistance genes, such as those encoding vancomycin resistance (vanA and vanB), has led to the emergence of vancomycin-resistant *Enterococcus* (VRE) [13, 47], associated with increased morbidity, mortality, and healthcare costs.

The antibiotic resistance mechanisms in *E. faecium* are diverse and include altered cell wall synthesis, drug efflux pumps, and enzymatic inactivation of antibiotics [17, 48]. The presence of multiple resistance mechanisms in *E. faecium* strains contributes to the difficulty in treating infections and highlights the need for new treatment strategies.

### 2.1.5 Current treatments and their mechanism of action against *E. faecium* bacterial infections

Current treatments for *E. faecium* infections include various antibiotics targeting various bacterial processes such as cell wall synthesis, protein synthesis, and membrane integrity. The effectiveness of these treatments is often compromised by the bacteria's ability to develop resistance through various mechanisms, making treating multi-drug resistant *E. faecium* a significant challenge. Here are some current treatments and their mechanisms:

***Vancomycin***

Vancomycin inhibits cell wall synthesis by binding to the D-Ala-D-Ala terminus of peptidoglycan precursors, preventing their incorporation into the cell wall. The emergence of Vancomycin-resistant *E. faecium* (VRE) limits its use [49].

***Linezolid***

Linezolid is an oxazolidinone that inhibits bacterial protein synthesis by binding to the 50S ribosomal subunit, preventing the formation of the initiation complex for translation. Resistance can develop through mutations in the 23S rRNA or acquisition of the cfr gene, which methylates the ribosome, reducing linezolid binding [50].

***Daptomycin***

Daptomycin is a lipopeptide that inserts into the bacterial cell membrane, causing rapid depolarisation and cell death due to loss of membrane potential. Resistance is rare but can occur through changes in membrane charge and composition [51].

***Tigecycline***

Tigecycline is a glycylcycline that binds to the 30S ribosomal subunit, blocking protein synthesis by preventing the incorporation of amino-acyl tRNA. Resistance is less common but can occur through efflux pumps and ribosomal protection proteins [52].

*Quinupristin/Dalfopristin (Synercid)*

This is a combination of two streptogramins. Quinupristin binds to the 50S ribosomal subunit, preventing elongation, while dalfopristin binds nearby, causing conformational changes that enhance quinupristin binding. Resistance can develop through enzymatic modification of the drugs or ribosomal target site mutations [53].

*Aminoglycosides (in combination with a cell wall agent)*

Aminoglycosides like gentamicin inhibit protein synthesis by binding to the 30S ribosomal subunit, causing misreading of mRNA. They are often combined with β-lactams or vancomycin to enhance their uptake into bacterial cells. Resistance mechanisms include aminoglycoside-modifying enzymes, 16S rRNA methylation, and efflux pumps [54].

*Nitrofurantoin*

Nitrofurantoin is reduced by bacterial flavoproteins to reactive intermediates that inactivate or alter bacterial ribosomal proteins and other macromolecules. It is mainly used for urinary tract infections (UTIs) due to its high urinary concentration [55].

*Chloramphenicol*

Chloramphenicol inhibits protein synthesis by binding to the 50S ribosomal subunit and inhibiting peptidyl transferase. Generally reserved for severe infections when other antibiotics are ineffective or contraindicated due to their potential for serious side effects [56].

*Fosfomycin*

Fosfomycin inhibits cell wall synthesis by inactivating the enzyme MurA, which catalyses the first step in peptidoglycan biosynthesis. It is often used for UTIs, including those caused by VRE [57, 58].

***New and Investigational Drugs***

*Oritavancin* is a lipoglycopeptide that disrupts cell wall synthesis and membrane integrity [59, 60].

*Telavancin* is a derivative of vancomycin that inhibits cell wall synthesis and disrupts cell membrane potential [60, 61].

*Ceftaroline* is a cephalosporin that acts against MRSA and some VRE strains by binding to altered PBPs [62].

### 2.1.6 Impact of antibiotic resistance on treatment outcomes

The rise of antibiotic resistance in *E. faecium* infections has profound implications for public health and patients. Infections caused by antibiotic-resistant strains of *E. faecium* are associated with an increased risk of treatment failure, longer hospital stays, higher mortality rates, and higher healthcare costs [63, 64]. Furthermore, the spread of antibiotic-resistant *E. faecium* strains within healthcare settings contributes to the transmission of resistance genes to other bacteria, further exacerbating the problem of antibiotic resistance. The emergence of VRE has led to outbreaks in healthcare facilities worldwide. It underscores the urgent need for effective strategies to prevent and control the spread of antibiotic-resistant *E. faecium*. The limited treatment options for VRE infections further exacerbate the problem, as alternative antibiotics are often less effective and more toxic [63]. Addressing the problem of antibiotic resistance in *E. faecium* is crucial to ensuring effective treatment and reducing the burden of *E. faecium* infections in healthcare settings.

## 2.2 The Structure and Function of Nicotinate-Nucleotide Adenylyltransferase (NNAT)

### 2.2.1 NNAT

NNAT is an enzyme that plays a critical role in the biosynthesis of $NAD^+$, a coenzyme involved in various metabolic pathways, including metabolism, DNA repair, and redox reactions [65]. The primary function of NNAT is to catalyse the conversion of NaMN to NaAD in the $NAD^+$ biosynthesis pathway. It can sometimes catalyse the direct conversion of NMN to $NAD^+$ through the salvage pathway but with a lower affinity [27]. This reaction involves transferring an adenylyl group from ATP to the carboxyl group of NAMN/NMN, forming NaAD/$NAD^+$ and pyrophosphate (Figure 2.1). $NAD^+$ is a critical cofactor in various metabolic pathways, including glycolysis, the tricarboxylic acid cycle, and oxidative phosphorylation. By catalysing the formation of NaAD, NNAT plays a central role in maintaining cellular $NAD^+$ levels and regulating cellular metabolism.

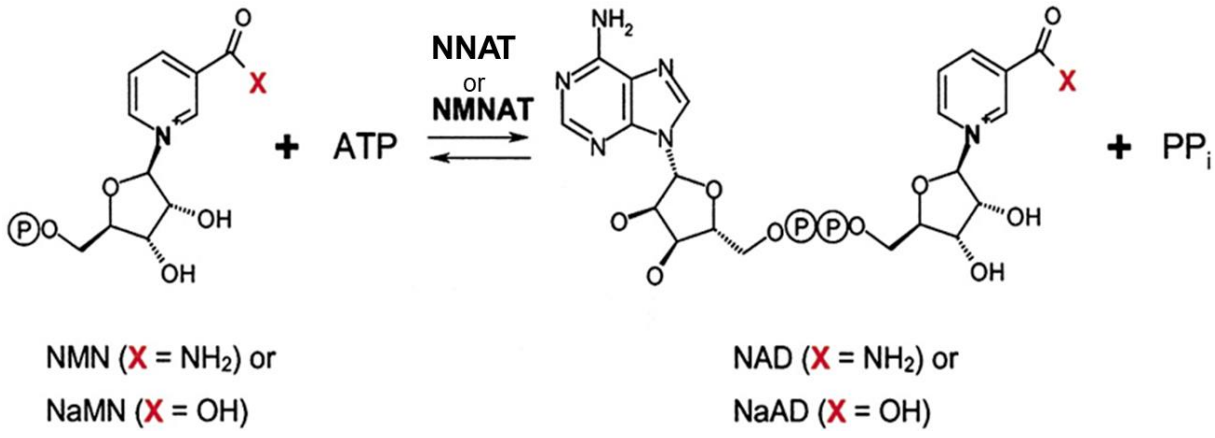



**Figure 2.1. The reaction catalysed by nicotinate/nicotinamide mononucleotide adenylyltransferase.** The diagram depicts the dual substrate catalysed reaction of NNAT/NMNAT enzyme resulting in $NAD^+$ or NaAD synthesis. Source: Zhang *et al.*, 2002 [66].

### 2.2.2 Role of NNAT in bacteria $NAD^+$ biosynthesis

In bacteria, the biosynthesis of $NAD^+$ is accomplished by two biochemical pathways, the de novo synthesis and the pyridine ring salvage pathway (Figure 2.2). The *de novo* synthesis of $NAD^+$ uses amino acids as precursors. Most bacteria synthesise $NAD^+$ anaerobically using aspartate and dihydroxyacetone phosphate as a precursor [67] to synthesise quinolinic acid in a two-step enzymatic reaction. In contrast, a few others synthesise $NAD^+$ aerobically from tryptophan via a 5-step enzymatic reaction to form quinolinic acid. The quinolinic acid formed from either tryptophan or aspartate precursor is phosphoribosylated and decarboxylated by the enzyme quinolinic acid phosphoribosyltransferase (encoded as NadC), giving rise to nicotinate mononucleotide (NaMN). The NaMN formed is further adenylated by the enzyme NNAT (NadD) using ATP as a cofactor to form nicotinate adenine dinucleotide (NaAD). The resulting NaAD is then amidated to $NAD^+$ in the presence of ammonia and ATP in a reaction catalysed by $NAD^+$ synthetase (NadE) [68, 69].

Breakdown products from $NAD^+$ turnover containing pyridine bases like nicotinate (Na) and nicotinamide (Nm) are recycled through the pyridine ring salvage pathway to resynthesise $NAD^+$. The most common salvage pathway in bacteria is the two enzymatic steps (PncA-PncB salvage route) that involve the deamination of nicotinamide to nicotinate by nicotinamidase (PncA) and the linking of nicotinate to 5-phosphoribosyl-1-pyrophosphate (PRPP) by the enzyme nicotinate phosphoribosyltransferase (PncB) to form NaMN. The NaMN formed converges at the *de novo* pathway, where it is adenylated and deamidated by NNAT and NAD synthase to form NaAD and $NAD^+$, respectively. Alternatively, nicotinamide is salvaged through NMN to synthesise $NAD^+$ by the enzyme nicotinamide phosphoribosyltransferase and NNAT, respectively (Figure 2.2). A few bacteria like *Haemophilus influenzae* which are $NAD^+$ auxotrophs, use different salvage pathways involving nicotinamide intermediate (ribosyl nicotinamide) to synthesise $NAD^+$. This is achieved through the PnuC-NadR nicotinamide riboside salvage

pathway involving a two-step conversion of ribosyl nicotinamide to $NAD^+$ via phosphorylation by the enzyme nicotinamide riboside kinase (NmR kinase) and adenylation by NNAT (NadR) respectively [70].

In summary, NaMN generated from the de novo pathway, as well as NaMN and NMN products generated from the salvage pathway, converge at the reaction catalysed by the enzyme NNAT to form their respective dinucleotides (NaAD and $NAD^+$) [71]. Thus, the essentiality of NNAT in the $NAD^+$ synthesis is established.

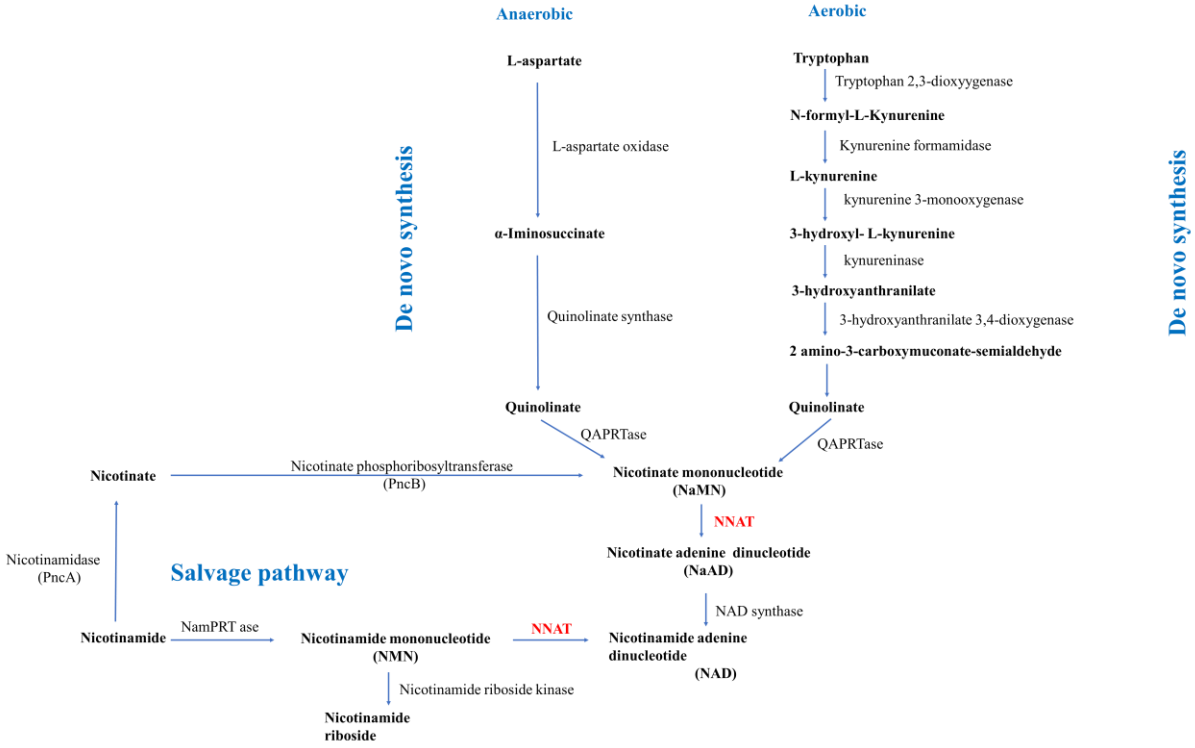



**Figure 2.2. Schematic representation of the de novo and salvage synthesis of $NAD^+$ in bacteria and humans.**

L-aspartate is the precursor amino acid utilised by bacteria anaerobically to synthesise $NAD^+$, while a few other bacteria and humans utilise tryptophan aerobically. NaMN generated from the *de novo* pathways employ NNAT to form $NaAD^+$, which is further deamidated to form $NAD^+$. Nicotinic acid, nicotinamide, and their corresponding ribosides resulting from $NAD^+$ breakdown and exogenous sources are salvaged to resynthesise $NAD^+$ via NaMN and NMN linked to the reaction catalysed by NNAT to form NaAD and $NAD^+$, respectively [72]. QAPRTase (quinolinic acid phosphoribosyltransferase), NADS (NAD synthetase), NRK (nicotinamide riboside kinase), NamPRT (nicotinamide phosphoribosyltransferase), and NAPRTase (nicotinate phosphoribosyltransferase).

### 2.2.3 NNAT role in human $NAD^+$ synthesis

NAD+ is also synthesised in humans via the de novo and pyridine salvage pathways. As the de novo aerobic $NAD^+$ synthesises in bacteria (Figure 2.2), tryptophan is the precursor amino acid. The sequential steps leading to the formation of $NAD^+$ are also the same for humans.

The $NAD^+$ salvage pathway in humans serves as the chief source of $NAD^+$. Nicotinate and nicotinamide generated from exogenous sources and $NAD^+$-utilising enzyme turnover are coupled directly to PRPP by NAPRTase and NamPRT to form NaMN and NMN, respectively. NMNAT further adenylates the generated mononucleotides to form NaAD and $NAD^+$. The NaAD formed is ultimately converted to $NAD^+$ by NAD synthase. Exogenous nicotinamide ribosides are also salvaged to synthesise $NAD^+$ via phosphorylation and adenylation reactions catalysed by NRK and NMNAT, respectively [73].

Interestingly, the successive biosynthetic steps to the formation of $NAD^+$ from quinolinic acid are well sustained in bacteria and humans. These reactions are catalysed by analogous enzyme sets, although with some notable disparities [74]. Amongst the enzymes that drive the downstream conversion of quinolinate to $NAD^+$, NNAT represents a potential antibacterial target because of its lowest sequence identity and more distant structure to its human homology [75]

### 2.2.4 Importance of NNAT in bacterial physiology

NNAT plays a pivotal role in the $NAD^+$ biosynthesis pathway in bacteria, supporting essential cellular processes such as energy metabolism, DNA repair, and stress responses. $NAD^+$ is indispensable for redox reactions in bacterial metabolism, including glycolysis, the tricarboxylic acid (TCA) cycle, and oxidative phosphorylation. It plays a key role in generating ATP, which is crucial for bacterial growth and survival (Figure 2.3). It is also involved in the activity of enzymes responsible for DNA repair and replication [76]. Adequate levels of $NAD^+$ are

necessary for maintaining genomic integrity and proper cell division in bacteria. Furthermore, $NAD^+$ is a substrate for sirtuins and other $NAD^+$-dependent enzymes that regulate bacterial responses to environmental stressors [77], such as oxidative stress and nutrient deprivation. These responses are vital for bacterial adaptation and survival under hostile conditions.

One of the many implications of NNAT activity and $NAD^+$ levels is pathogenicity. Many pathogenic bacteria rely on efficient $NAD^+$ biosynthesis for virulence. NNAT activity supports the energy demands and stress responses needed for infection and colonisation of host tissues. Furthermore, $NAD^+$ regulates biofilm formation, a key factor in bacterial persistence and antibiotic resistance. NNAT activity can affect the ability of bacteria to form and maintain biofilms. Moreover, $NAD^+$ is critical for the survival of bacteria under antibiotic pressure. NNAT and $NAD^+$ biosynthesis pathways can influence the effectiveness of certain antibiotics, and targeting these pathways may enhance antibiotic susceptibility.

NNAT represents a promising target for the development of new antibacterial agents. Inhibitors of NNAT can disrupt $NAD^+$ biosynthesis, leading to impaired bacterial metabolism and viability. Also, combining NNAT inhibitors with existing antibiotics can potentially enhance the antibacterial effects by weakening bacterial defences and making them more susceptible to treatment.

### 2.2.5 Structure of NNAT

NNAT are globular proteins containing about 160-400 amino acids, with subunit mass ranging between 20 and 50 kDa. Most characterised NNATs exist as homo-oligomers of two to six identical subunits. They are classified as the nucleotidyltransferase $\alpha/\beta$ phosphodiesterase superfamily, identified by a characteristic $\alpha/\beta$-fold, with a central $\beta$-sheet surrounded by $\alpha$-helices. Their subunit comprises a larger domain identified as the N-terminal and a small domain identified

as the C-terminal. The structure of NNAT is highly conserved across species, indicating its importance in cellular function [27].

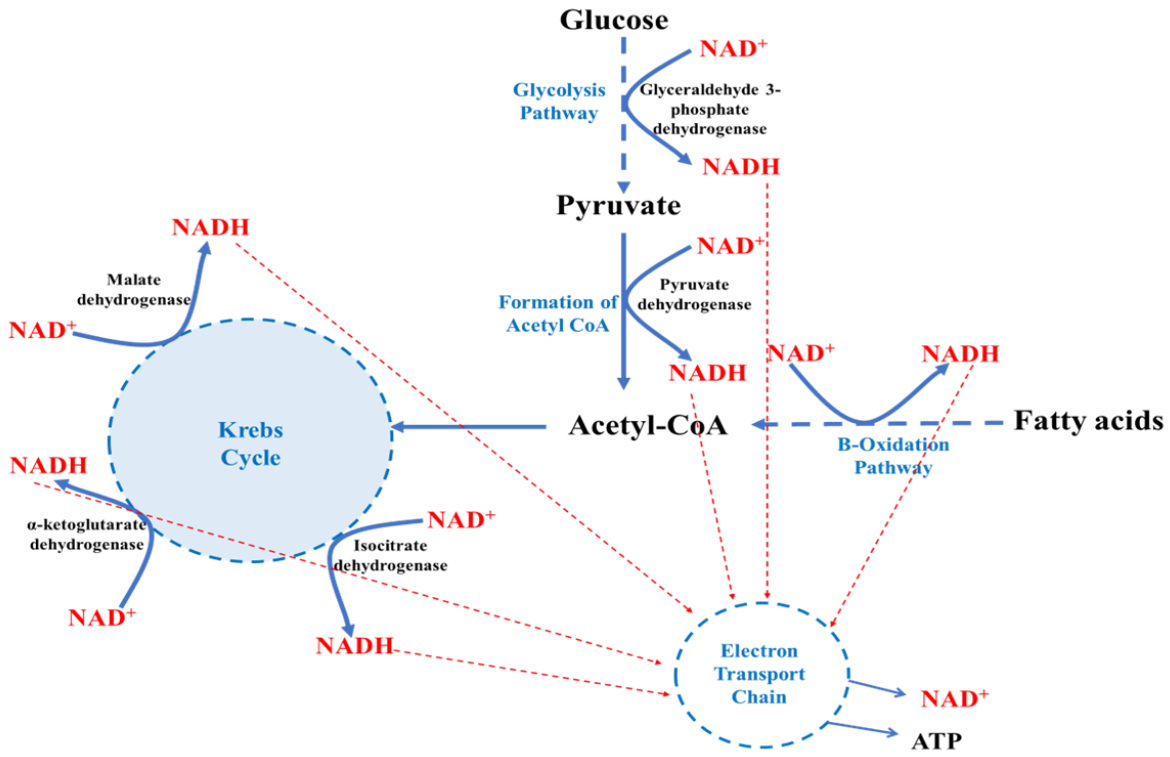



**Figure 2.3. A schematic representation of linked metabolic pathways involving the redox activity of $NAD^+$/NADH to generate energy.**

## 2.2.6 Comparison between bacterial and human NMNAT

Bacteria and human NNAT/NMNAT are affiliated with the same nucleotidyltransferase superfamily, yet they share less than 20% sequence identity and no more than 40% sequence similarity (Figure 2.4). Several biochemical studies related to NNAT disclose that both bacteria and human NMNAT are two unique subfamilies, exhibiting differences in sequence conservation patterns, mononucleotide substrate specificity, oligomeric states, biochemical and enzymatic properties (Figure 2.5) [71, 78].

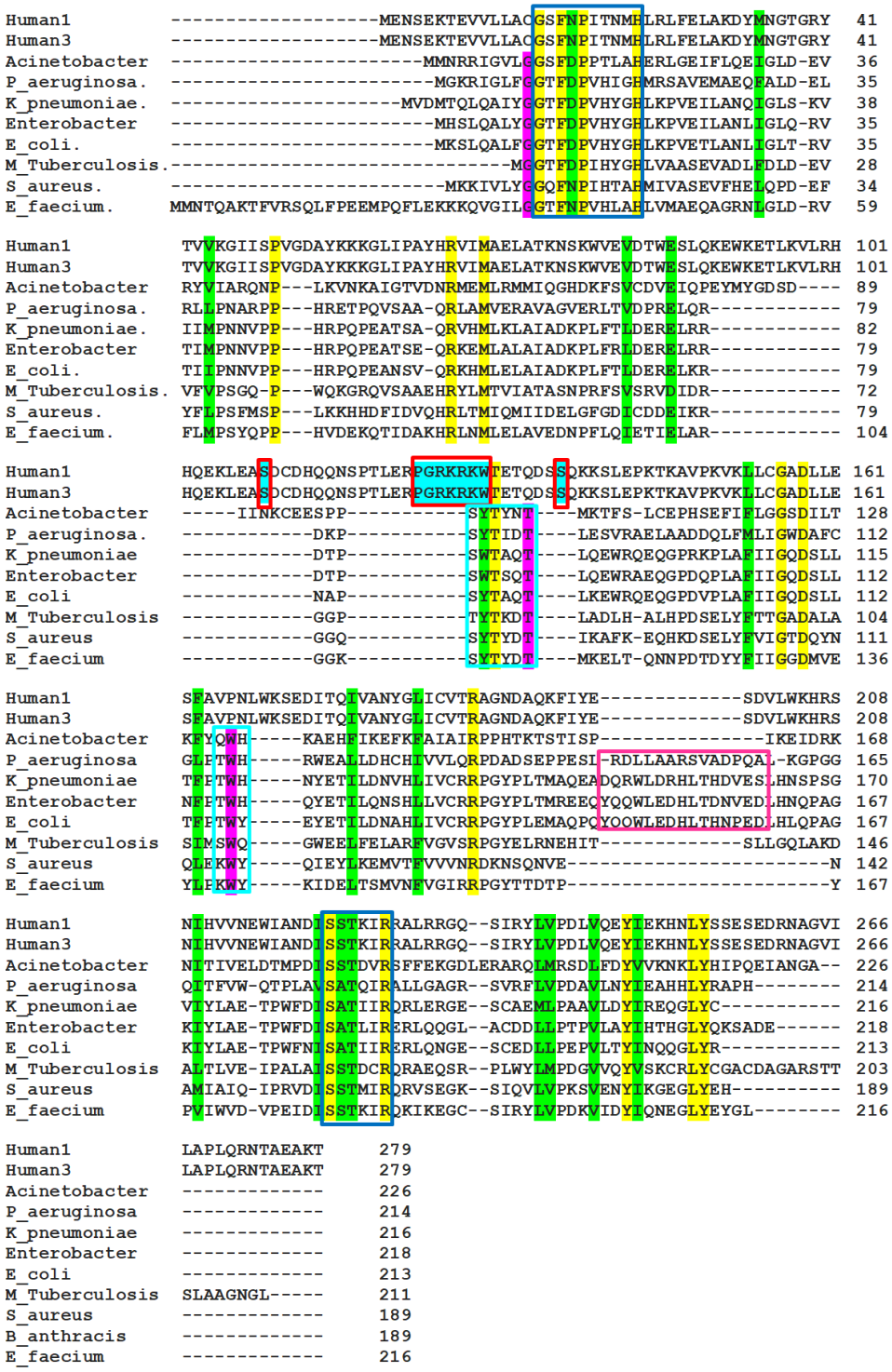



**Figure 2.4. The sequence alignment of NNAT from human and bacteria species showing variation in the amino acid sequences.**

The yellow highlight indicates positions containing conserved residues across all species, purple highlights conservation across bacteria species, and partially conserved or mutated residues are indicated in green. The absence of residues in the designated space indicates non-conservation. The region in the red box indicates the nuclear localisation signal (NLS) sequence and the two phosphorylation site serine residues conserved to the human NMNAT. The ATP binding motifs are shown in a blue box, the cyan box indicates the NaMN binding motif common to the bacteria species, and the additional $\alpha\beta\alpha$ unit insertion common to the Gram-negative bacteria NNAT is shown in pink box. The sequence alignment presented in this context was produced utilising the CLUSTAL Omega.

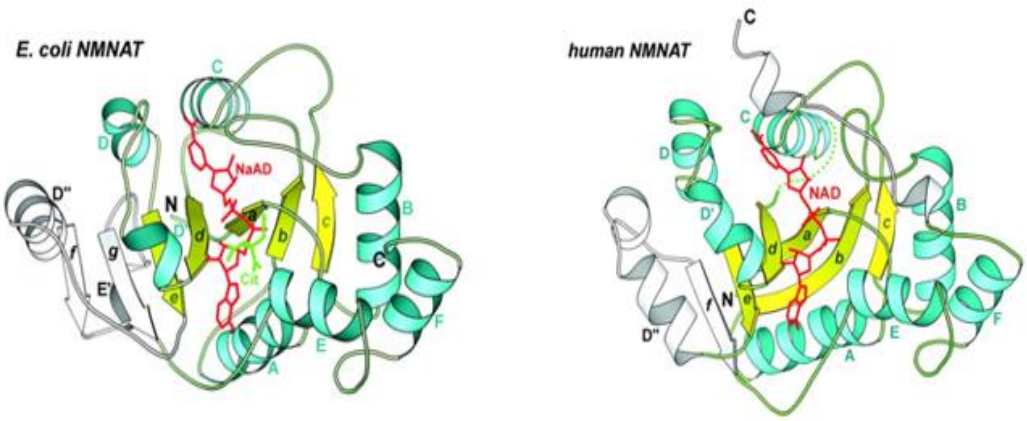



**Figure 2.5. Ribbon representation showing the difference between *E. coli* NNAT and human NMNAT protomer structure.**
The bound ligands are represented by red sticks for NaAD or $NAD^+$, while the $SO_4^{2-}$ and citrate are represented with green sticks. The secondary structures are labelled grey for the random coils, yellow for strands, and cyan for helices. The dotted line indicates the disordered hsNMNAT region absent in the bacteria NNAT. RMSD = 15.755 Å (492 to 492 atoms) [75].

### *Mononucleotide substrate specificity*

In contrast to bacteria NNAT (EC 2.7.1.18), which prefers NaMN over NMN [79], the human NMNAT (hNMNAT EC 2.7.1.1) exhibits equal dual substrates recognition towards both mononucleotides hence, can flexibly participate in both *pathways* leading to the synthesis of $NAD^+$. The diverse substrate specificities exhibited by NNAT/NMNAT are mediated by the carboxyl group and carboxamide group of the nicotinate and nicotinamide substrates. In bacteria *E. coli* NNAT (EcNNAT), the positive dipole of $\alpha$-helix C and the main chain amide group of Tyr85, Ala86, Leu117 create an anion-binding pocket, which is responsible for recognising the negative charge on NaMN carboxylate group (Figure 2.4). The surrounding hydrophobic residues further enhance the electrophilic nature of this pocket, thus increasing its ability to discern between a carboxamide (NMN) and a carboxylate (NaMN) and its reason for a higher affinity towards NaMN than NMN. Conversely, the hNMNAT can recognise both NaMN and NMN readily with no need for any substantial spatial arrangement changes because of the existence of a negatively charged Asp residue (Asp173) and several conserved water molecules in the hNMNAT pyridine-binding

site. This allows the enzyme to take up substrates of different electrostatic properties by changing the refined charge distributions of the binding site, thus making hNMNAT less rigid in stabilising different NMN/NaMN metabolic changes for $NAD^+$ synthesis [66, 75].

***Sequence conservation and secondary structure***

As revealed from the sequence alignment (Figure 2.4) and the crystallographic analysis on recombinant hNMNAT (Figure 2.5), the distinguished complex structures possess a disordered region encompassing a nuclear localisation signal (NLS) sequence PGRKRKW and two phosphorylation site serine residues (Ser109 and Ser136) which interacts with the nuclear-transporting proteins. However, the sequence complementary to this region is absent in bacterial NNAT [75, 80]. Furthermore, the monomer structure of hNMNAT exhibits a dinucleotide-binding fold consisting of a six-stranded central parallel β-sheet fringed by several helices, which are quite different from bacteria *Ec*NNAT that has an added αβαβ (**D′′fE′g**) section implanted at the β sheet C-terminal end summing up to seven-stranded central β-sheet instead (Figure 2.5) [66]. Besides, the hNMNAT C-terminal extension has a short helix (residues 267–274) situated near the pyridine-binding site, which possibly impacts substrate recognition [75], which is also absent in the bacterial NNAT [66].

***Oligomeric states***

As revealed from previous studies, human NMNAT exists in three different isoforms based on cellular activity and tissue distribution. These isoforms exhibit different functional units. Compared to the monomeric and dimeric structures typical of the bacterial NadD NNAT, the hNMNAT-1 exhibits a homo-hexameric structure, hNMNAT-2 a monomeric, and hNMNAT-3 a tetrameric structure. The variances between bacteria and human NNNAT/NMNAT active site, subunit interfaces, and limited sequence identity make NNAT a suitable target as it suggests that highly selective inhibitors may be found or designed [75, 81].

### 2.2.7 Overview of EfNNAT

The EfNNAT comprises 214 amino acids, with a predicted molecular weight of ~26 kDa and a theoretical pI of 5.79. Its molar extinction coefficient measured in water at 280 nm is 30370 $M^{-1}$ $cm^{-1}$ (ProtParam ExPASY). Based on the predicted homology model generated (Figure 2.6) using *Bacillus anthracis* (*Ba*NNAT- PDB ID:2QTN), which shares ~46% sequence identity (Swissmodel.expasy.org), EfNNAT is a homo-dimeric protein. The presence of a Rossman fold characterises its secondary structure, which is predominately $\alpha$-helical. The overall structure of EfNNAT shows a composition of larger and smaller domains referred to as N-terminal and C-terminal, respectively. The highly conserved "H/TXGH" signature motif sequence required for recognition and binding of ATP is identified as " HLAH " and located close to the N-terminus between residues 37 and 40.

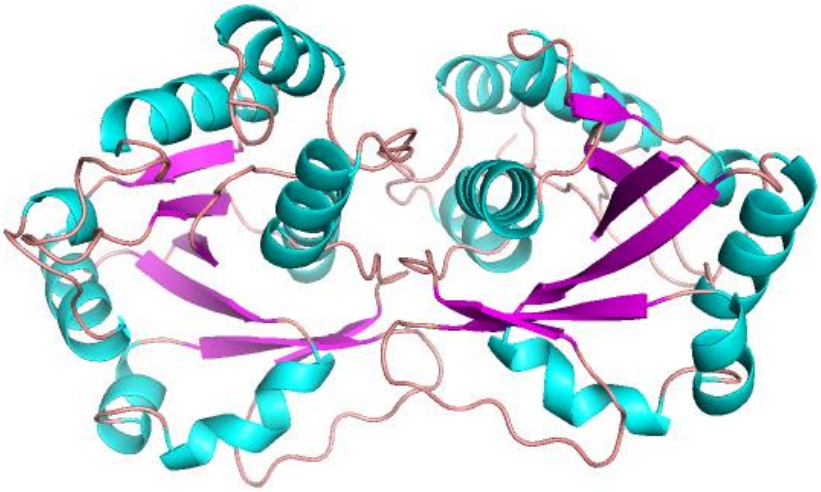


**Figure 2.6. Predicted structure of EfNNAT generated via SwissModel.**
EfNNAT was built using the 2qtn.1.A model through SwissModel. The $\alpha$-helices are represented in green, the β-strands in purple, and the unordered coils in brown.

### 2.2.8 Studies suggesting inhibition of NNAT

It is hypothesised that by suppressing the activity of NNAT in the cell, the production of $NAD^+$ will be disrupted, resulting in cell death. Rodionova *et. al* (2014)

confirmed this in *Mycobacterium smegmatis* model system using a protein knockdown approach to induce the deprivation of target enzyme NNAT, which consequently displayed a robust bactericidal effect with changes in the levels of NaMN and eventual depletion of NAD(H) pool. The depleting effect was more pronounced in vital pathways such as the Kreb cycle, cholesterol catabolism, fatty acid, and aromatic compound synthesis that depend so much on NAD(P) redox reaction. The metabolic activities were suppressed, consequently affecting the production of ATP energy [29]. The expectation is that inhibitors that would be designed against the NNAT enzyme must be able to deplete the $NAD^+$ pool and result in the entire collapse of the central metabolic system, consequently inhibiting bacterial growth [28, 82].

## 2.3 Protein Structure Determination

Protein structure determination involves several methods, each with its own strengths and limitations, and it is chosen based on the specific requirements of the study. The main techniques used for determining the structures of proteins are X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, cryo-electron microscopy (cryo-EM), and small-angle X-ray scattering (SAXS).

X-ray crystallography and cryo-EM are preferred for high-resolution structures of large proteins and complexes, while NMR is suitable for studying proteins in solution, providing insights into their dynamics. SAXS and mass spectrometry offer complementary information, which is particularly useful for understanding proteins' overall shape and interactions. Combining these methods often provides a complete understanding of protein structure and function.

### 2.3.1 X-ray crystallography

X-ray crystallography is the most widely used method for determining the atomic structure of proteins. It is a powerful technique for determining the three-

dimensional structure of biological macromolecules, providing valuable insights into their structure-function and mechanism of action. X-ray crystallography is suitable for high-resolution structures of large proteins and protein complexes, and it provides detailed information about the atomic arrangement of the protein. It is used extensively in drug design and understanding enzyme mechanisms. One of its limitations is that it requires high-quality crystals, which can be challenging to obtain for some proteins. Also, it provides static snapshots of proteins, which may not capture the dynamic aspects of protein function [83, 84].

### 2.3.2 Nuclear magnetic resonance (NMR) spectroscopy

Nuclear Magnetic Resonance (NMR) spectroscopy is a powerful analytical technique used to determine molecules' structure, dynamics, and interactions, particularly in solution. It measures the magnetic properties of atomic nuclei to determine the structure of proteins in solution. NMR spectroscopy is based on the magnetic properties of atomic nuclei. Certain isotopes, such as ^1H, ^13C, ^15N, and ^31P, have nuclear spins, which makes them NMR-active. When these nuclei are placed in a strong external magnetic field, they align with or against the field, creating distinct energy levels [85].

NMR is widely used in chemistry, biochemistry, and structural biology to elucidate the structure of small molecules, proteins, nucleic acids, and other macromolecules. It provides information about the distances and angles between atoms, which can be used to build a protein structure model. NMR is suitable for studying proteins in their native, soluble state and provides information on protein dynamics, folding, and interactions. This technique is useful for studying smaller proteins, typically up to 40 kDa, due to the complexity and time required for larger proteins. It also requires high concentrations of protein samples, which can sometimes be difficult to obtain [85].

### 2.3.3 Cryo-electron microscopy (cryo-EM)

Cryo-Electron Microscopy (cryo-EM) is a technique used in structural biology to visualise the architecture of macromolecules at near-atomic resolution without the need for crystallisation. Unlike traditional electron microscopy, cryo-EM involves the rapid freezing of samples to preserve their native state without the need for chemical fixation or staining. The frozen-hydrated samples are then examined with a transmission electron microscope (TEM). The electrons interact with the sample, creating images that contain information about the structure [86, 87].

This method has revolutionised the field of structural biology, enabling the study of complex and dynamic biomolecular structures that are difficult to analyse using other techniques. The method is ideal for studying large protein complexes, membrane proteins, and viruses. Furthermore, it provides insights into the structure and function of proteins in a near-native state, and it is suitable for heterogeneous samples and dynamic protein complexes [86, 87]. The limitation of the technique is that it requires expensive and sophisticated equipment, and the data processing and image reconstruction can be computationally intensive.

### 2.3.4 Small-angle X-ray scattering (SAXS)

Small-Angle X-ray Scattering (SAXS) is a versatile and powerful technique used to study the structure of materials at the nanoscale. It measures the scattering of X-rays as they pass through a protein solution, providing information about the overall shape and size of the protein, conformational changes of proteins in solution and the internal structure of particles in solution or solid-state systems. SAXS gives low-resolution information about a protein's structure in solution and is particularly valuable in the field of structural biology for analysing macromolecules, such as proteins, nucleic acids, and their complexes, in a near-native environment [88].

Its ability to analyse samples in a near-native state and in real-time makes it an essential tool for studying dynamic processes and complex systems. Despite its limitations in resolution and detailed atomic data interpretation, SAXS remains an indispensable method in both biological and material sciences, complementing other high-resolution structural techniques like X-ray crystallography and NMR. It is suitable for studying protein folding, dynamics, and interactions [88].

### 2.3.5 Mass spectrometry (MS)

Mass spectrometry (MS) is an analytical technique used to measure the mass-to-charge ratio (m/z) of ions. It is widely used to determine the molecular weight, identify the composition of a sample, quantify the amount of specific components within a sample, and elucidate the structure of molecules. It can be used to infer structural information and study protein modifications and interactions.

MS is a powerful tool in various fields, including chemistry, biochemistry, and molecular biology, for studying protein complexes, post-translational modifications, and protein-ligand interactions. It can provide insights into the stoichiometry and architecture of protein complexes but does not provide detailed structural information on its own. It is often used in conjunction with other techniques for comprehensive structural analysis. Despite its complexity and cost, the breadth of applications and the depth of information provided by MS continue to drive its widespread use and development [89].

## 2.4 X-Ray Crystallography Studies

The X-ray crystallography method was utilised in this study to determine the structure of EfNNAT in its apo and complexed form. This was done in order to obtain a detailed atomic arrangement of the protein, which is necessary to provide valuable insights into the structure, function, and mechanism of the protein.

### 2.4.1 Methodology of X-ray crystallography studies

X-ray crystallography relies on the diffraction of X-rays by the electron density of crystals to determine the arrangement of atoms in a molecule. When X-rays interact with a crystal, they are diffracted at specific angles, producing a diffraction pattern that can be used to determine the positions of atoms within the crystal lattice. Collecting and analysing these diffraction patterns can reconstruct the molecule's three-dimensional electron density map, revealing its atomic details [90, 91].

In X-ray crystallography studies, the first step is obtaining high-quality protein crystals. This typically involves purifying the protein to homogeneity and crystallising it under controlled conditions. Once suitable crystals are obtained, X-ray diffraction data is collected using a specialised X-ray diffractometer. The diffraction data is then processed and analysed using computational methods to determine the electron density map of the protein. Finally, the atomic model of the protein is built into the electron density map and refined to fit the experimental data. Figure 2.7 provides an overview of the x-ray crystallography steps.

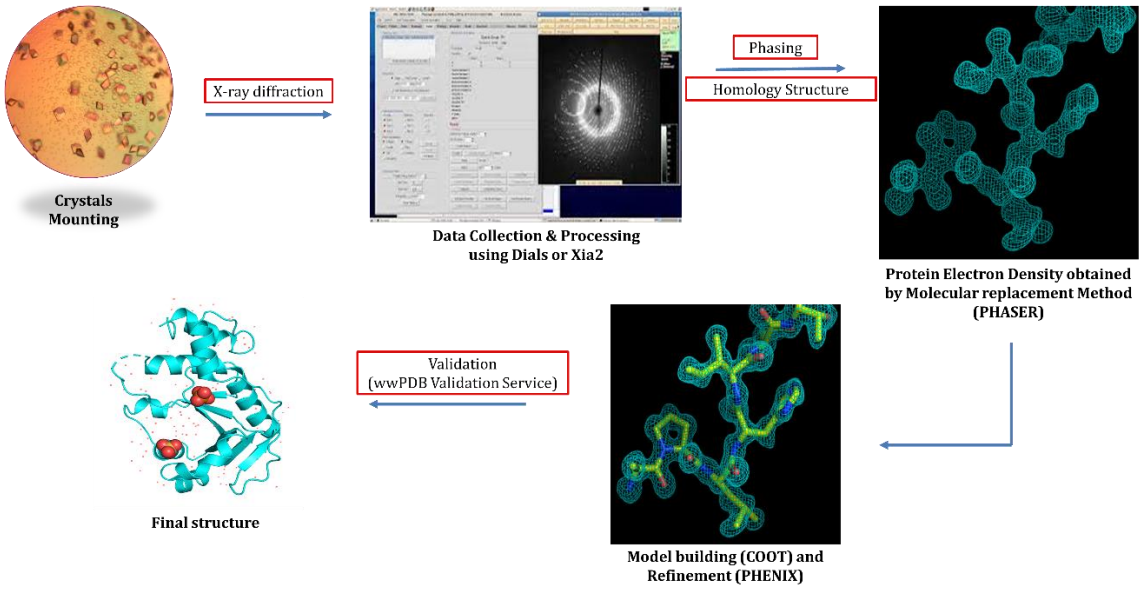



**Figure 2.7. Overview of X-ray crystallography methods utilised for protein structure determination.**

### 2.4.2 Protein crystallisation and setting-up

Protein crystallisation is a fundamental step in determining the three-dimensional structure of a protein via X-ray crystallography. The process involves producing a highly ordered, crystalline form of the protein, which can then diffract X-rays to provide data for structural analysis. Protein crystallisation is a complex, iterative process requiring careful optimisation of numerous variables, including protein concentration, precipitant type and concentration (ammonium sulfate, polyethylene glycol and polymers), pH, temperature, buffer and additives [84, 92].

Higher protein concentrations can promote crystallisation but may also lead to aggregation. Temperature and protein pH affect solubility and stability; hence needs to be screened over a wide range. The presence of addictives like small molecules or ions can stabilise the protein, thus promoting crystallisation. Because proteins are unique and crystallise under different conditions, crystal trials are often necessary. This is performed using various crystallisation screen buffers made up of different screening conditions [84, 93]. The ultimate goal is to obtain high-quality crystals that can sufficiently diffract X-rays to determine the protein's three-dimensional structure, providing insights into its function and potential for drug targeting. Different techniques are used for setting up protein crystallisation;

***Vapor Diffusion***

***Hanging Drop***: A small drop of protein solution is mixed with a reservoir solution on a cover slip, which is then inverted over the reservoir. Water evaporates from the drop to the reservoir, increasing the concentration of protein and precipitant, leading to crystallisation.

***Sitting Drop***: A drop containing protein and precipitant is placed in a well or on a pedestal above a reservoir solution. Similar to a hanging drop, water evaporates from the drop to the reservoir, increasing the concentration of protein and precipitant leading to crystallisation.

***Microbatch Under Oil***

Protein and precipitant solutions are mixed and covered with oil to prevent evaporation. This allows for slow equilibration and, finally, crystallisation.

***Dialysis***

The protein solution is placed inside a dialysis membrane and immersed in a precipitant solution. Slow diffusion of precipitant into the protein solution leads to crystallisation [94].

***Free Interface Diffusion***

Protein solution and precipitant solution are layered in a capillary or small tube. As the solutions diffuse into each other, a concentration gradient forms, leading to crystallisation at the interface.

### 2.4.3 X-ray data collection and analysis

In X-ray crystallography, the crystal of interest is placed in a beam of X-rays, and the resulting diffraction pattern is recorded. The diffraction pattern is a series of spots on a detector, each corresponding to a set of planes in the crystal lattice that diffracted the X-rays. The crystal is rotated to obtain a complete diffraction pattern, and data is collected at multiple angles. The resulting dataset contains information about the intensity and direction of scattered X-rays at different angles. This dataset are processed and analysed using computational tools, including Xia [95], Dials [96], autoPROC [97] to determine the spatial arrangement of atoms within the crystal. The structure can be refined by comparing the experimental diffraction data with theoretical models and an atomic-resolution model of the crystal generated.

### 2.4.4 Phasing

Phasing in X-ray crystallography is a crucial step in determining the three-dimensional structure of a molecule, typically a protein or nucleic acid, from its X-ray diffraction pattern. The primary challenge in crystallography is solving the phase

problem, which arises because while the intensities of the diffracted X-rays can be measured directly, the phases of the diffracted beams cannot be. Without phase information, it is impossible to directly construct an electron density map, which is necessary to model the molecule's structure [98]. The different Methods of Phasing include;

***Molecular Replacement (MR)***

This method uses a known structure of a similar molecule as a template to determine the phases. The known structure, often referred to as a search model, is positioned in the unit cell of the unknown structure to generate an initial phase estimate. Software tools such as PHASER of CCP4i suite [99], MOLREP in the CCP4 suite [100] and MrBUMP [101]are some of the widely used tools for performing molecular replacement.

***Single Isomorphous Replacement (SIR) and Multiple Isomorphous Replacement (MIR)***

This uses heavy atom derivatives of the crystal to introduce additional scattering that can be used to calculate phase angles. The Crystals are soaked in solutions containing heavy atoms, and the differences in diffraction patterns (with and without heavy atoms) are used to estimate phases. Some of the software tools used for SIR and MIR include SHARP and SHELX [102]. SHARP is often used for heavy atom site refinement and phase calculation, while the SHELX are used for locating heavy atoms and refining their positions [102].

***Single-wavelength Anomalous Dispersion (SAD) and Multi-wavelength Anomalous Dispersion (MAD)***

Here, the method derives phase information from the anomalous scattering properties of certain atoms, such as selenium, bromine, or sulfur. In SAD, data is collected at one wavelength, while in MAD, data is collected at multiple wavelengths to maximise the anomalous signal. PHENIX Autosol [103, 104] and SOLVE/RESOLVE

[87] are some of the software tools for SAD and MAD phasing and density modification.

***Direct Methods***

This uses mathematical algorithms to directly solve the phase problem, which is typically effective for small molecules solution and refinement rather than macromolecules. The method relies on statistical relationships between the intensities of the reflections. Examples of software tools used for direct method phasing are SHELXD [102, 105] and Olex2 [106].

## 2.4.5 Model building and structure refinement

Model building and refinement are critical steps in determining the three-dimensional structure of a protein using X-ray crystallography. These involve constructing an atomic model of the protein and iteratively improving its fit to the experimental data. After obtaining the initial phases and electron density maps, an initial model of the molecule is built based on the electron density map using software like Coot (Crystallographic Object-Oriented Toolkit) [107]. The protein sequence is aligned with the density map, and the polypeptide chain is accurately traced by assigning sequence positions to density, correcting any register shifts, and ensuring continuity of the main chain. Coot allows researchers to manually fit amino acid residues into the electron density and perform real-space refinement, density fit, and rotamer fitting. However, other automated programs automatically place amino acids into the density map, providing an initial model. These include programs like ARP/wARP [108], PHENIX AutoBuild [104], or Buccaneer [109]. The iterative steps of model building and refinement include;

***Refinement cycles***

This process involves adjusting the positions of the atoms in the model to minimise the difference between the calculated and observed diffraction patterns. Software

programs like REFMAC, PHENIX, and CNS (Crystallography and NMR System) are used to perform the various refinement cycles, which include rigid body refinement, positional refinement, B-factor refinement, anisotropic refinement, and TLS (translation, libration, screw) refinement.

***Validation and correction***

The refined model is validated, and necessary corrections are made to improve accuracy using tools like MolProbity, which evaluates model quality by checking geometry and steric clashes and fitting them to the density. The R-free and R-work are also monitored as they are indicators of refinement quality [110]. The R-free is calculated from a subset of reflections not used in refinement. If the refinement process advances as expected, the free R-factor will decrease accordingly. However, if the model contains significant errors, the free R-factor will remain unchanged.

***Addressing refinement issues***

Refinement issues such as model bias is reduced using composite omit maps, twinning issues are also identified and corrected, and adjustments relating to the model building are corrected on the difference density maps (Fo-Fc maps).

***Software Tools for Model Building and Refinement***

1. *Coot (Crystallographic Object-Oriented Toolkit)*

   It is an interactive program for model building, validation, and visualisation of macromolecular structures. It displays the electron density maps (*2Fo-Fc*, *Fo-Fc*, composite omit maps) and molecular models to be refined. Coot allows manual model building, including adding and adjusting residues, building peptide chains, fitting rotamers, fitting ligands, and modifying density. Real-space refinement can also be performed on Coot, where parts of the model are refined directly against the electron density. Coot can be a validation tool as it contains

features including Ramachandran plots, clash detection, and real-space correlation coefficients [107].

2. *PHENIX (Python-based Hierarchical Environment for Integrated Xtallography)*

   This is a comprehensive software suite for automated and manual structure determination, including phasing, model building, refinement, and validation analysis to ensure high-quality model. It is an interactive tool that provides a user-friendly graphical interface for accessing various tools and functions within the PHENIX suite, including automated model building and refinement, validation using MolProbity integration, phasing using PHASER and AutoSol, electron density map interpretation and modification, ligand and small molecule handling and fitting [103, 104, 111].

3. *REFMAC (Refinement of Macromolecular Structures)*

   Allows automated refinement of macromolecular structures using maximum likelihood methods. It is often used in conjunction with Coot for iterative refinement and manual adjustments. It provides detailed output on model statistics and quality. Some of its features include TLS refinement for modelling anisotropic displacement, positional, B-factor, anisotropic refinement, and validation tools [112, 113].

4. *CCP4 (Collaborative Computational Project No. 4)*

   CCP4 is a comprehensive suite of programs for all aspects of macromolecular crystallography, including data processing, model building, and refinement. It provides a comprehensive environment for crystallographic workflows. Its features include tools for phasing, model building, refinement, and validation. CCP4 utilises programs like Buccaneer for automated model building and integration with REFMAC and other refinement tools for refinement [101, 112].

5. *ARP/wARP*

   Is an automated model building and refinement, particularly effective at initial stages. Its features include automated tracing of protein chains, solvent modelling, and iterative model building [108].

6. *Chimera*

   Chimera is a tool for structural comparison, surface analysis, and distance measurements. It provides advanced visualisation of molecular structures and electron density maps. Interactive modelling, including manual adjustments, analysis of protein-ligand interactions, and presentation of structural data, can be done using this tool. Chimera provides a range of tools for exploring and interpreting structural models.

7. *MolProbity*

   MolProbity is a comprehensive validation tool for model geometry, including all-atom contacts, Ramachandran analysis, and rotamer evaluations. It suggests corrections for identified issues during refinement, such as flipping incorrect side chain orientations to improve model [114]. Overall, it ensures high-quality models by identifying and correcting common errors. It is often used for the final validation step before submitting structures to databases like the Protein Data Bank (PDB).

### 2.4.6 Structure validation and deposition

Once model building and refinement are complete, the final refined structure can be validated for stereochemical clashes using tools like MolProbity or web servers (wwPDB Validation Service) (https://validate-rcsb-2.wwpdb.org/)[115]. Structures can then be analysed to understand the arrangement of atoms in the crystal and their

interactions. Upon validation and satisfaction, the structure is deposited into the PDBe [116]

### 2.4.7 X-ray crystallography studies of NNAT

X-ray crystallography studies have contributed significantly to understanding NNAT across the prokaryotes and eukaryotes species, elucidating the structural basis of its catalytic activity and substrate specificity [75, 117-121]. X-ray crystallography studies have elucidated the structure of NNAT at atomic resolution, revealing a homo-oligomeric enzyme composed of identical subunits. Its overall fold is shown with each subunit adopting a characteristic $\alpha/\beta$-fold, with a central β-sheet surrounded by $\alpha$-helices [66, 117, 122]. The active site of NNAT, where catalysis occurs, has been identified and characterised, providing a detailed understanding of the catalytic mechanism of NNAT. Furthermore, X-ray crystallography studies have been used to identify different sequence motifs and key residues involved in substrate recognition and binding of ATP and NaMN/NMN [121, 123]. Additionally, the binding interactions between NNAT and its substrates ATP, NaMN, NMN, and NaAD in different species have been elucidated, shedding light on the substrate specificity and product binding of the enzyme.

Continued X-ray crystallography studies of NNAT hold promise for uncovering new insights into enzyme inhibition with allosteric modulators and small-molecule inhibitors, which have implications for developing novel therapeutics targeting NNAT for treating infections and diseases.

## 2.5 Computational Modelling

Computational modelling involves the simulation and study of complex systems using computers. It is a theoretical aspect of science that combines mathematics, physics, and computer science, which is demonstrated to be an

extremely useful complement to experimental investigations. Over the years, different sophisticated applications and techniques have been developed to simulate near reality and explore the molecular recognition and the dynamics of biomolecular events between a ligand and a receptor site [124, 125]. These techniques employ a range of physical methodologies, including classical molecular mechanics (MM) and quantum mechanics (QM). Because of the predictive power of computational biochemistry in providing insight and information about certain data that are sometimes inaccessible from experiments, the integration of computational modelling with empirical studies has been embraced to augment experimental data and provide further elucidation of biological processes at the atomic level, which can be extrapolated for rational drug design [126].

Some common computational approaches employed in evaluating and analysing the molecular interactions and ligand bioactivity include homology modelling, molecular docking, and molecular dynamic simulation.

### 2.5.1 Homology modelling

Due to the limited availability of experimental three-dimensional crystal structures of proteins, computational modelling has become another route for predicting or deriving unavailable protein structures [127, 128]. This is possible through homology modelling, which utilises the amino acid sequence of the target protein and an experimentally determined crystal structure of a related homologous protein as a model to construct a 3D structure of the protein. A library of experimentally determined protein structures is probed for a specific protein structure to be built to select appropriate templates. This template selection depends on the similarities in the structure and sequence alignment as the query sequence. Hence, a homology model depends on the sequence alignment quality and the template structure in use [129]. Homology modelling is applicable for purposes like protein-ligand interaction and drug design predictions, which require atomic-resolution data.

### 2.5.2 Molecular docking

Molecular docking studies the fitting together of two or more molecular structures. It involves the *in-silico* process of simulating ligands' binding, including substrates, drugs, or inhibitors, into the active site of macromolecules such as protein or nucleic acid to predict how the small molecules will interact with the macromolecule *in vivo*. The primary aim of molecular docking is to find the correct positions of ligands in the protein binding pocket and to envisage the affinity between the ligand and the protein. These are achieved through two basic docking processes: sampling and scoring. Sampling generates all probable positions, orientations, and ligand conformations within the protein sites. The scoring function aids in accurately determining the most energetically advantageous binding pose and classifies it in ranking order. Though the available docking programs operate slightly differently, these two basic docking processes are the essential composite of a molecular docking program [130, 131].

### 2.5.3 Molecular dynamic (MD) simulation

MD simulation is a computational method utilised to mimic complex molecular systems at an atomic level and compute the motions of atoms present in the actual environment of individual molecules. It analyses the physical movements and fluctuations of atoms and molecules over a set timeframe while the dynamic evolution of the system is sighted. MD simulation enables the sampling of the conformational space of protein molecules and the determination of the kinetic and thermodynamic properties, thus providing a connection between structure and dynamics and an understanding of the structural and mechanistic properties of the protein molecule.[132] The MD simulations of proteins require many sequential steps, as outlined in Figure 2.8. The processes include energy minimisation, solvation, system heating, equilibration, and production phase. The simulation of proteins is conducted to replicate experimental settings necessitating the consideration of several physical

factors such as pressure and temperature. Protein simulations often use the isothermal-isobaric (NPT) or the canonical ensemble (NVT).

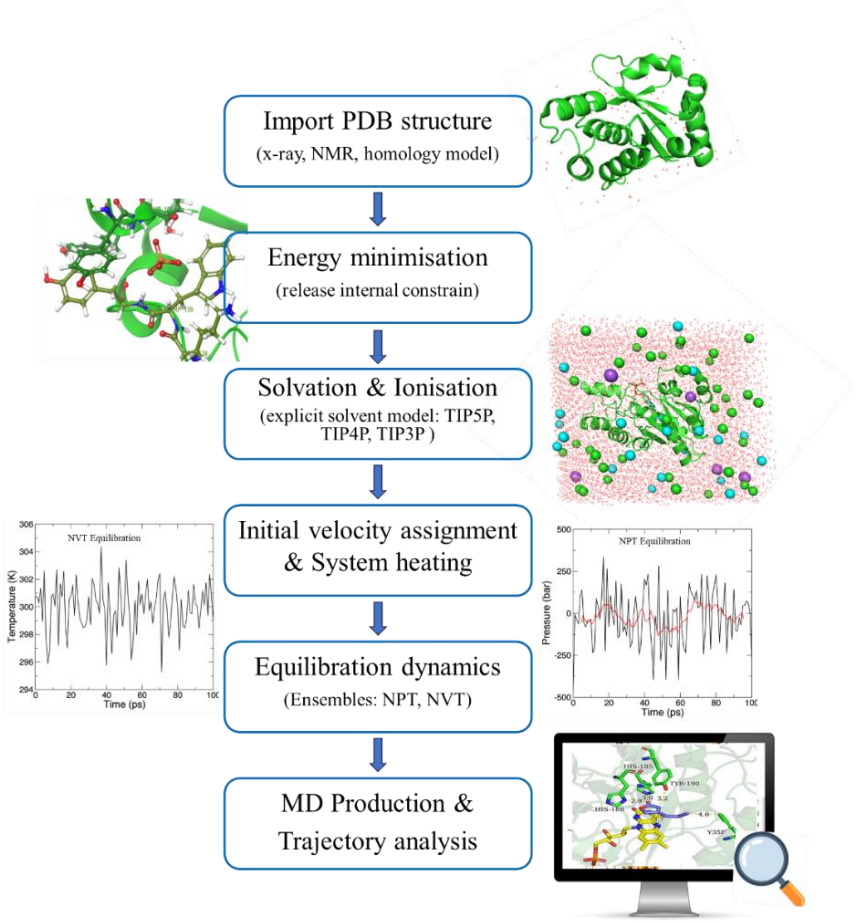



**Figure 2.8. Outline of the different stages involved in the MD simulation of a protein.**
NMR: Nuclear magnetic resonance; MD: Molecular Dynamics; NVT: Number of molecules/atoms in assemby, Volume, Temperature; NPT: Number of molecules/atoms in assembly, Pressure, Temperature; TIP3P/4P/5P: Transferable Intermolecular Potential 3/4/5-site water models [132].

# Chapter 3

# Obtaining High Yield Recombinant *Enterococcus faecium* Nicotinate Nucleotide Adenylyltransferase for X-Ray Crystallography and Biophysical Studies

Olamide Jeje, Ramesh Pandian, Yasien Sayed and Ikechukwu Achilonu*

Protein Structure-Function Research Unit, School of Molecular and Cell Biology, Faculty of Science, University of the Witwatersrand, Johannesburg 2050, South Africa.

Olamide.Jeje1@students.wits.ac.za, Ramesh.Pandian@wits.ac.za, Yasien.Sayed@wits.ac.za

*Correspondence: Ikechukwu.Achilonu@wits.ac.za; Tel.: +27-11-717-6349

## 3.1 Highlights

- Recombinant expression of the EfNNAT yielded approximately 101 mg of EfNNAT per litre of culture.
- More than 98 % homogeneously pure EfNNAT is obtainable using a single-step IMAC purification method.
- EfNNAT is highly stable and easily crystallised under several conditions.
- Two high-resolution crystal structures of EfNNAT were determined at 1.90 Å and 1.82 Å.

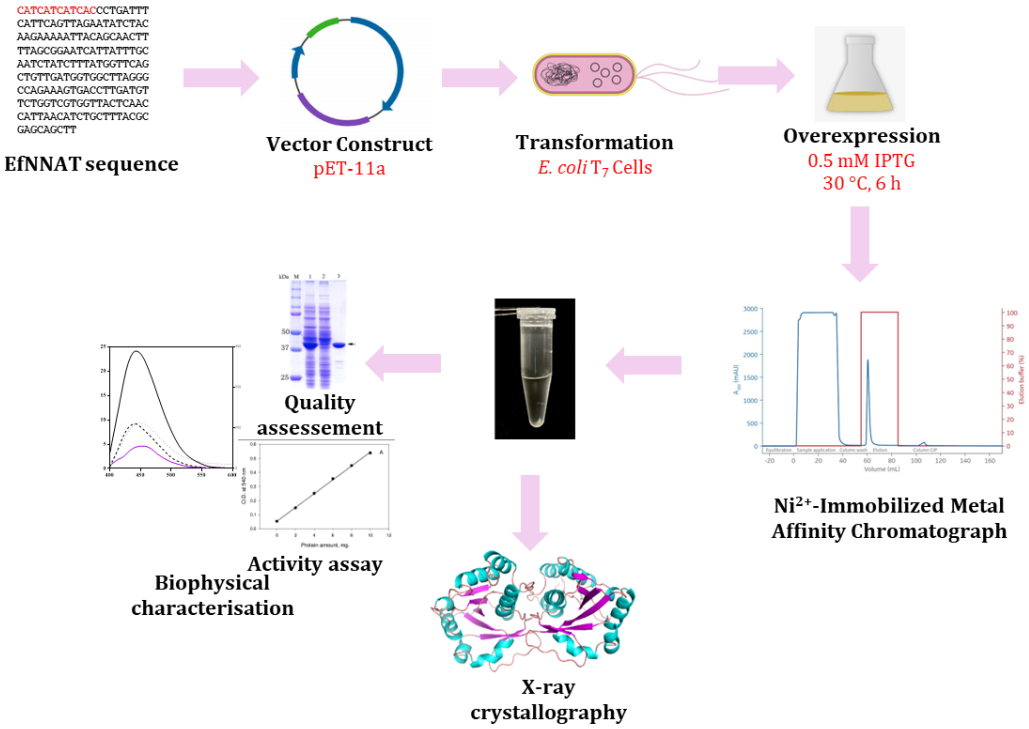
## 3.2 Abstract

Nicotinate nucleotide adenylyltransferase (NNAT) has been a significant research focus on druggable targets, given its indispensability in the biosynthesis of $NAD^+$, which is crucial to the survival of bacterial pathogens. However, no information is available on the structure-function of *Enterococcus faecium* NNAT (EfNNAT). This study established the expression and purification protocol for obtaining a high-yield recombinant EfNNAT using the *E. coli* expression system and a single-step IMAC purification method. Approximately 101 mg of EfNNAT was obtained per 7.8 g of wet *E. coli* cells, estimated to be >98 % pure. We further characterised the biophysical structure and determined the three-dimensional structure of the EfNNAT. Biophysical studies revealed a dimeric protein with a higher $\alpha$-helical composition. The highly stable protein crystallised in multiple conditions, yielding high-quality crystals diffracting between 1.78-2.8 Å. Two high-resolution crystal structures of EfNNAT in its native and adenine-bound forms were determined at 1.90 Å and 1.82 Å, respectively. The X-ray structures of the EfNNAT revealed the presence of phosphate and sulfate ions occupying and interacting with conserved amino acid residues within the putative substrate binding site, hence providing insight into the probable substrate preference of EfNNAT and, consequently, why

EfNNAT may not prefer β-nicotinamide mononucleotide as a substrate. With the accessibility to high-resolution structures of EfNNAT, further structural evaluation and drug-based screening can be achieved. Hence, we anticipate this study will provide the basis for discovering structure-based inhibitors against this enzyme.

**Keywords:** Nicotinate Nucleotide Adenylyltransferase (NNAT); Nicotinamide mononucleotide Adenylyltransferase (NMNAT); *Enterococcus faecium*; Expression; Purification; X-ray Crystallography.

## 3.3 Graphical Abstract



## 3.4 Introduction

The gram-positive *E. faecium* is a bacterial pathogen notorious for resisting most known antibiotics. It has been identified alongside other virulent and multidrug-resistant bacterial pathogens nick-named ESKAPE (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* spp.) as the major causative of nosocomial infections (Hospital-acquired infection) which threatens health care systems across the globe [1, 3]. The ESKAPE pathogens have been listed by the World Health Organisation (WHO) among the bacteria which urgently need new antibiotic intervention [133]. Given the substantial antibiotic resistance of *E. faecium* that has continued to increase over the years and the declining number of current antibiotics effective against this pathogen, *E. faecium* has been listed in the high-priority list of pathogens urgently needing new antibiotics [133]. Yet, only a few new antibiotics have been developed since 2010 [134].

The biosynthesis of $NAD^+$ and the enzymes involved in its downstream formation has been a foreground for drug targets in recent times [28, 29, 118, 135-140] because of the impact and essentiality of $NAD^+$ to cell survival. The vitalness of $NAD^+$ and its derivatives revolves around them regulating key metabolic reactions and pathways. Since these metabolic pathways are linked, disruption in the availability of $NAD^+$ would most likely result in system malfunctioning and possible collapse of the entire central metabolic system, ultimately resulting in cell death. The availability of $NAD^+$ in bacteria solely depends on $NAD^+$ biosynthesis via the *de novo* or pyridine salvage pathway. These two pathways meet at the reaction step catalysed by the enzyme NNAT. Since most bacteria cannot take up $NAD^+$ from their environment and there is no other known route leading to pyridine dinucleotide formation, NNAT activity is clearly indispensable in $NAD^+$ synthesis and, hence, its exploration as a druggable target.

NNAT is a dual substrate enzyme responsible for the adenylation of NaMN and NMN using ATP. It can catalyse the transfer of adenyl group unto NaMN and

NMN, leading to the formation of their respective dinucleotides, NaAD and $NAD^+$, but at a different degree of specificity [136]. Depending on the enzyme-substrate preference, it is either termed nicotinate nucleotide adenylyltransferase (NNAT) or nicotinamide mononucleotide adenylyltransferase (NMNAT). NNAT belongs to the superfamily of nucleotidyltransferase $\alpha/\beta$ phosphodiesterase (EC 2.7.7.18), characterised by a highly conserved "H/TXXH" signature motif and the presence of a Rossman fold (alternating β-strands and $\alpha$-helices folding to form an extended β-sheets surrounded by $\alpha$-helices) [141]. It is widely conserved in almost all species however, evolutionary divergence exists, reflecting their unique biophysical, structural, and catalytic properties. The structural variation exhibited between the human and bacterial NNAT and its conservation amongst bacterial pathogens makes it an attractive drug target for the design of antimicrobials [20, 75].

Despite the wide embrace of NNAT study in different species [66, 75, 119, 121, 142-147] until now, no information is available on the biophysical structure of EfNNAT. To provide this missing information while validating EfNNAT as a potential druggable target, the *nadD* gene encoding EfNNAT, a 216 amino acid, was explored. In this study, we reported the overexpression and purification of the recombinant protein because availability of highly soluble and purified EfNNAT is essential for the downstream characterisation of the enzyme and the basis towards uncovering the structure-function of the enzyme. Furthermore, the biophysical features of the enzyme were characterised, and the X-ray structure determined in its apo and adenine-bound forms to gain insight into the overall architecture of the enzyme and probable mechanisms of biologic activities at the atomic and molecular level necessary to facilitate drug development.

## 3.5. Materials and Methods

### 3.5.1 EfNNAT vector construct and overexpression

The EfNNAT expression vector was constructed using the EfNNAT gene open reading frame (UniProtKB-A0A133MWI0 (A0A133MWI0_ENTFC)), obtained from the EMBL database (ID: KXA08386). The sequence was codon-optimised to incorporate a NdeI restriction endonuclease site (5'CAT ATG 3') and a hexahistidine tag nucleotide sequence (CATCATCATCATCAC) at the 5'-end of the EfNNAT ORF sequence, and a BamHI restriction endonuclease site (5'GGA TCC 3') at the 3'-end. The codon-optimised sequence was synthesised and cloned into a pET-derived expression vector (pET-11a) via the NdeI and BamHI restriction endonuclease sites by GenScript Corporation (NJ, USA). The vector construct (pET-11a-$(His)_6$-EfNNAT) resulted in a recombinant EfNNAT protein with an uncleavable hexahistidine tag at the N-terminus (Figure 3.1), which was used for purification via IMAC. This amounted to a recombinant protein composed of 223 amino acid residues with an estimated molecular weight of ~25.8 kDa (ProtParam ExPASY) [148].

To overexpress the recombinant EfNNAT, *E. coli* T7 cells (*E. coli* T7-Express ™), which contains the chloramphenicol resistance plasmid pLysS were transformed with the resulting vector construct (pET-11a-$(His)_6$-EfNNAT), and grown on an LB-agar plate supplemented with 50 μg/mL ampicillin. A single colony of the successful transformants was inoculated into a 20 mL 2 × Yeast-Tryptone (2 × YT) media starter culture containing 50 μg/mL ampicillin and grown for 12-16 h in a rotating incubator (37 °C, 250 rpm). The starter culture was used to inoculate a freshly prepared 2 × YT media augmented with 50 μg/mL ampicillin and incubated with shaking (37 °C, 250 rpm) to an optical density ($OD_{600}$) of ~0.6. Overexpression of EfNNAT was induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) followed by incubation and shaking (30 °C, 200 rpm) for 6 h. The expressing cells were harvested by centrifugation (4 °C, 5000 × *g*) for 10 min, and the cell pellet resuspended in a 24 mL lysis buffer (Phosphate-buffered saline (PBS), pH 7.2, 25 mM imidazole, 0.01% (w/v)

$NaN_3$) per litre of culture. Resuspended cells were frozen at −20 °C for 20 h to facilitate cell lysis. Frozen cells were thawed and lysed through sonication on ice. The soluble fraction was extracted from the lysate by centrifugation (4 °C, 18,000 × *g*, 15 min). The overexpression of the recombinant protein was analysed on a 12.5% (w/v) discontinuous SDS-PAGE.

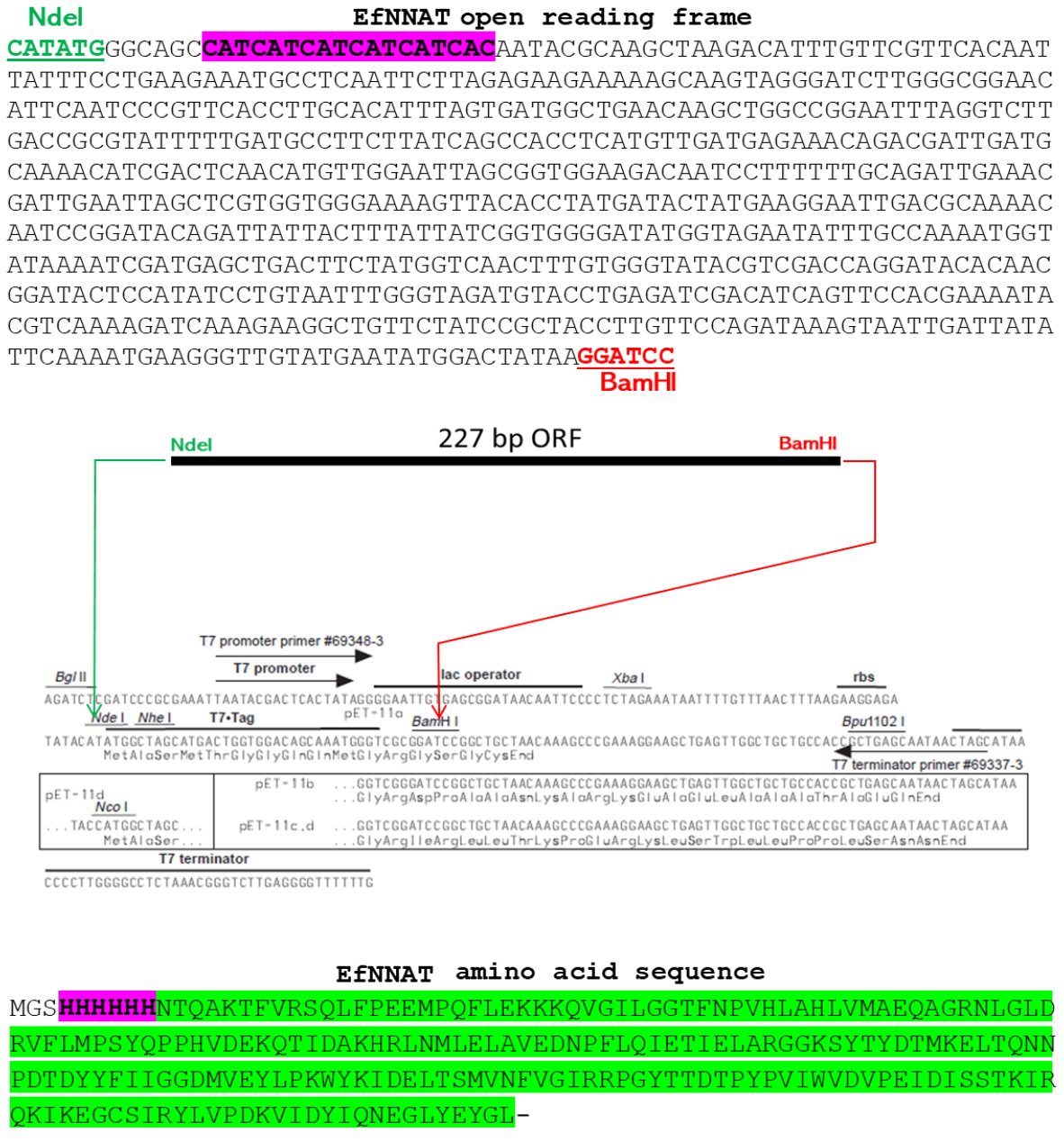



**Figure 3.1. A pictorial representation of the vector construct used for expressing the recombinant EfNNAT.**

Hexahistidine was factored into the N-terminus of the gene sequence, as shown in purple. The resulting sequence was inserted into pET-11a at the NdeI (green) and BamHI (red) sites.

### 3.5.2 Recombinant EfNNAT purification

The overexpressed EfNNAT protein was purified using nickel ion-immobilised metal affinity chromatography ($Ni^{2+}$-IMAC), performed at 20 °C with low pressure and a gravity pump (Model EP-1 Econo™, Bio-Rad). The soluble fraction containing the hexahistidine tag EfNNAT was passed through a $Ni^{2+}$-Sepharose resin column pre-equilibrated with 30-column volumes of lysis buffer. Upon collection of the unbound flow through, the resin was equilibrated and then washed with 0.1% (v/v) Tween-20 to remove non-specifically bound proteins from the column. Another round of equilibration followed this to remove any residual Tween-20. The bound EfNNAT was eluted with 350 mM imidazole. The quality and quantity of the eluate (purified hexahistidine tagged EfNNAT) were determined using SDS-PAGE and a UV-Vis spectrophotometer (Jasco V-630). All downstream experiments were performed with pure hexahistidine-tagged (His-tagged) EfNNAT.

### 3.5.3 Dual enzyme-coupled assay

A coupled enzyme assay involving EfNNAT and alcohol dehydrogenase (ADH) was used to investigate whether the expressed enzyme was active [149]. The product formed by EfNNAT ($NAD^+$) acts as a substrate for ADH to form NADH, which is then measured spectrophotometrically at 340 nm. Given that NNATs are dual-substrate enzymes capable of utilising both NMN and NaMN, this assay explored NMN alongside ATP as the initial substrate to form $NAD^+$. Reason being that with NaMN, another enzyme, NAD synthase, would be required first to convert the product NaAD to $NAD^+$.

The EfNNAT activity assay was carried out between pH 5.5–10.5 in different buffers, including 0.1 M sodium acetate, 0.1 M sodium phosphate, 0.1 M Tris-HCl, and 0.1 M sodium glycine, to determine its optimal activity [150]. A standard 1 mL reaction mixture containing 1 μM of recombinant EfNNAT, 5 mM $MgCl_2$, 1 mM ATP, 1% (v/v) ethanol, 10 U/mL ADH, and 1 mM DTT was used. The reaction was initiated

at 20 °C by adding 0.5 mM NMN. The absorbance of NADH produced from the reduction of $NAD^+$, which EfNNAT synthesised, was monitored spectrophotometrically using a UV-Vis spectrophotometer (Jasco V-630) at 340 nm for 60 sec. Linear progress curves were recorded, and the slope of the absorbance against time generated for each assay was plotted against the pH. The negative control for the assay was carried out in the absence of EfNNAT.

### 3.5.4 EfNNAT secondary structure analysis

The secondary structural composition of the recombinant EfNNAT and the impact of ATP ligand binding to its secondary structure were estimated by far-UV circular dichroism (CD) spectroscopy. The apo (5 μM EfNNAT), ATP bound (5 μM EfNNAT, 0.2 mM ATP, 1 mM $MgCl_2$), and denatured (5 μM EfNNAT, 8 M urea) forms of EfNNAT were analysed in dd$H_2O$ at ~20 °C with a CD spectrometer (Jasco J-1500). The experiments were performed in triplicate using 2 mm high transparency CD quartz cuvettes, 1 nm data pitch, 2 nm bandwidth, and 200 nm/min scanning speed. Each sample reading is an average of 5 accumulations. EfNNAT spectra data were accessible within 185 and 260 nm wavelength range. A plot of mean residue ellipticity [θ] against wavelength was generated by first converting the machine unit in millidegrees to mdeg $cm^2$ $dmol^{-1}$ $residue^{-1}$. *Sigma Plot* was used to process the resulting CD spectra, while GraphPad Prism 8 was used to generate the plot.

### 3.5.5 Spectroscopy studies with fluorescent dyes

The EfNNAT tertiary structure conformation, interactions and binding properties with the ATP ligand were investigated using extrinsic fluorescent dyes (8-Anilino-1-naphthalenesulfonic acid (ANS) and mant-ATP). The experiments were carried out on the EfNNAT native, denatured and in complex with ATP. A 1 mL reaction mixture of 50 μM ANS and 5 μM EfNNAT was used for the ANS binding

studies. To analyse the denatured form of EfNNAT, 8 M urea was augmented to the reaction mixture to denature the protein. The EfNNAT-ATP complex sample was prepared by supplementing the reaction mixture with 1 mM ATP and 2 mM $MgCl_2$. All samples were excited at 395 nm, and the ANS fluorescence emission was collected over a 400-650 nm wavelength range.

Similarly, the mant-ATP studies were performed using a 1 mL reaction mixture comprising 10 μM mant-ATP and 2.5 μM EfNNAT. The mixture was supplemented with 8 M urea to denature the protein, and 1 mM ATP and 2 mM $MgCl_2$ were used for the EfNNAT-ATP complex sample. Mant-ATP excitation was at 355 nm, and fluorescence emission was collected over a 400-600 nm wavelength range. For both the ANS and mant-ATP fluorescence assays, triplicates of samples were obtained over 5 accumulation readings. These were averaged and corrected for buffer absorption and free dyes. All assays were performed in 10 mM $NaH_2PO_4$, pH 7.5, 0.01% (w/v) $NaN_3$, and 1 mM DTT, at 20 °C using a quartz cuvette (10 mm path length) in a spectrofluorometer (Jasco FP-6300, 1 nm data pitch, 5 nm excitation and emission bandwidth, 200 nm/min scanning speed). *Sigma Plot* was used to process the resulting spectra, and GraphPad Prism 8 was used to generate the plots.

### 3.5.6 Thermal shift assays

The thermal denaturation and stability of recombinant EfNNAT were monitored with SYPRO Orange dye. The study was conducted with and without ATP using a Real-Time PCR Detector (CFX96 Touch, Bio-Rad). A 25 μL reaction volume comprising 25 mM Tris-HCl, pH 7.5, 0.01% (w/v) $NaN_3$, 20 μg of protein, and 10 × SYPRO Orange dye was used for the assay. In order to assay for EfNNAT-ATP interaction, 0.1 mM ATP and 2 mM $MgCl_2$ were added to the reaction mixture. All assays were carried out in a PCR microplate (96-well, Bio-Rad) in quadruplicate. A sealing Film (Microseal 'B' PCR, Bio-Rad) was used to seal the plate, vortexed for 5 sec, and then centrifuged for 30 sec before running in a Real-Time PCR Detection System

(20 - 80 °C in increments of 0.5 °C for 10 sec). The readings for the relative fluorescence unit (RFU) and the derivative of the fluorescence signal with respect to the temperature (-d(RFU)/dT) were accessible via CFX Maestro software and exported to Excel for further processing. The fraction unfolded (RFU), and the (- d(RFU)/dT) readings were plotted individually against temperature using GraphPad Prism 8.

### 3.5.7 EfNNAT quaternary structure determination

The overall native state and size of the recombinant EfNNAT in solution were determined by size exclusion high-performance liquid chromatography (SE-HPLC). A TSKgel SuperSW2000 30 cm × 4.6 mm Tosoh Bioscience column was used to measure EfNNAT retention time. Column equilibration was performed with 10 mM $NaH_2PO_4$, pH 7.5, 750 mM NaCl, and 0.01% (w/v) $NaN_3$ at 20 °C and calibrated (0.2 mL/min flow rate, 40 min) using 20 µL of 18 mg/mL standard protein markers (Bio-Rad gel filtration standard), ranging between 1.35 kDa and 670 kDa. The protein fractions were monitored at an absorbance of 280 nm. A 20 µL of EfNNAT (~6.5 mg/mL) was passed onto the column and analysed in triplicate under the same conditions as the standard markers. EfNNAT molecular weight was calculated by comparing its retention time with the standards through a calibration plot of log molecular weight versus the retention time of the standards. Based on the apparent molecular weight obtained, the oligomeric state of EfNNAT was deduced.

### 3.3.8 Crystallisation of EfNNAT

In order to obtain crystals with good diffraction quality, an optimal crystal growth condition is needed to determine the three-dimensional structure of EfNNAT. The Index crystallisation screen kit (HR2-144, Hampton Research) was used to identify the crystallisation conditions for EfNNAT using the sitting drop vapor diffusion method. Each crystallisation well was prepared using a 2 µL of 15 mg/mL EfNNAT

prepared in 50 mM HEPES, 150 mM NaCl, pH 7.2, 2 mM DTT, and 0.01% (w/v) $NaN_3$, and 2 μL of the reservoir solution. Each well was equilibrated against 250 μL of their respective reservoir solutions and incubated at 20 °C. Relatively big crystals of EfNNAT were obtained within 48-72 h under various crystallisation set-up conditions. Crystals were harvested and preserved with either Parabar 10312 or mother liquor containing 20% glycerol and immediately flash-frozen in liquid nitrogen. Samples were then sent to the Diamond Light Source Ltd (DLS Ltd) in the UK for data collection.

### 3.5.9 EfNNAT crystal data collection and structure solution

The X-ray diffraction intensities of EfNNAT apo and adenine-bound crystals were collected at the DLS Synchrotron with the I04 beamline at DLS Ltd, UK, equipped with Eiger2 XE 16 M detector. Data were collected under a liquid nitrogen stream at 100 K using a 0.5° oscillation angle coverage per frame. The scaling and reduction of the apo and adenine-bound forms of EfNNAT data sets were performed using XIA2 [95] and DIALS [96], respectively. The structure solution of both structures was executed by the molecular replacement method using MrBUMP of CCP4i suite [101, 151, 152].

A homology structure generated by AlphaFold [128] embedded in the Diamond computational pipeline [153] was used as the initial search model for the EfNNAT structures reported in this paper. Before its usage, this structure was improved using SCULPTOR [34] in PHASER [35]. Several rounds of refinement and model building were performed, along with the addition of solvent molecules using PHENIX [103] and COOT [107], which resulted in a complete model. The stereochemical quality for the final two structures was validated through an online-based server, wwPDB Validation Service (https://validate-rcsb-2.wwpdb.org/) [115]. The atomic coordinates and structure factors for the EfNNAT apo and adenine

complexes were deposited in the Protein Data Bank with accession codes 8AIH and 8AII, respectively.

## 3.6 Results

### 3.6.1 Expression and purification of recombinant EfNNAT

The use of a vector construct pET-11a-$(His)_6$EfNNAT in *E. coli* T7 cells, with the optimal expression conditions of 0.5 mM IPTG, 30 °C, 230 rpm shaking, and 6 h incubation, was successful in overexpressing the recombinant EfNNAT protein. Approximately 7.8 g of wet cell pellet per litre of EfNNAT expression culture was obtained. The analysed crude lysate showed a high yield of soluble His-tagged EfNNAT (~25.8 kDa) expressed in the supernatant (Figure 3.2, lane 3). The single-step $Ni^{2+}$-IMAC purification proved effective in isolating the His-tagged protein as well as in achieving pure EfNNAT (Figure 3.2, lane 7). Furthermore, the qualitative assessment of the eluate via ultraviolet-visible absorption spectrum monitored between 240 and 350 nm confirmed the absence of nucleic acid contamination and aggregations. EfNNAT concentration was estimated to be ~5.05 mg/mL using the Bradford assay, which amounted to a final purification yield of ~101 mg EfNNAT per litre of *E. coli* T7 expression culture (**Table 3.1**). The obtained His-tagged EfNNAT is highly stable even with a concentration as high as 50 mg/mL.

**Table 3.1. Summary of EfNNAT expression and purification yield**

| Steps | Av. Protein Concentration (μg/ml) | Vol (ml) | Total Protein (μg) | Total Protein (mg) | Amount of Protein Lost (mg) |
|---|---|---|---|---|---|
| Crude Lysate | 15578.33 | 33 | 514085.00 | 514.085 | 0 |
| Supernatant | 7174.58 | 28 | 200888.00 | 200.888 | 313.197 |
| IMAC | 5049.58 | 20 | 100991.70 | 100.992 | 413.093 |
| Buffer Exchange | 4547.14 | 20 | 90942.90 | 90.943 | 423.142 |

The assay was performed from a 7.8 g wet cell obtained from a 1 L culture. The protein concentration was determined by Bradford assay, using bovine serum albumin as standard. Each average protein concentration represents the absorbance reading of 5 accumulations performed in triplicate. A loss of total protein was observed with each step down the table.

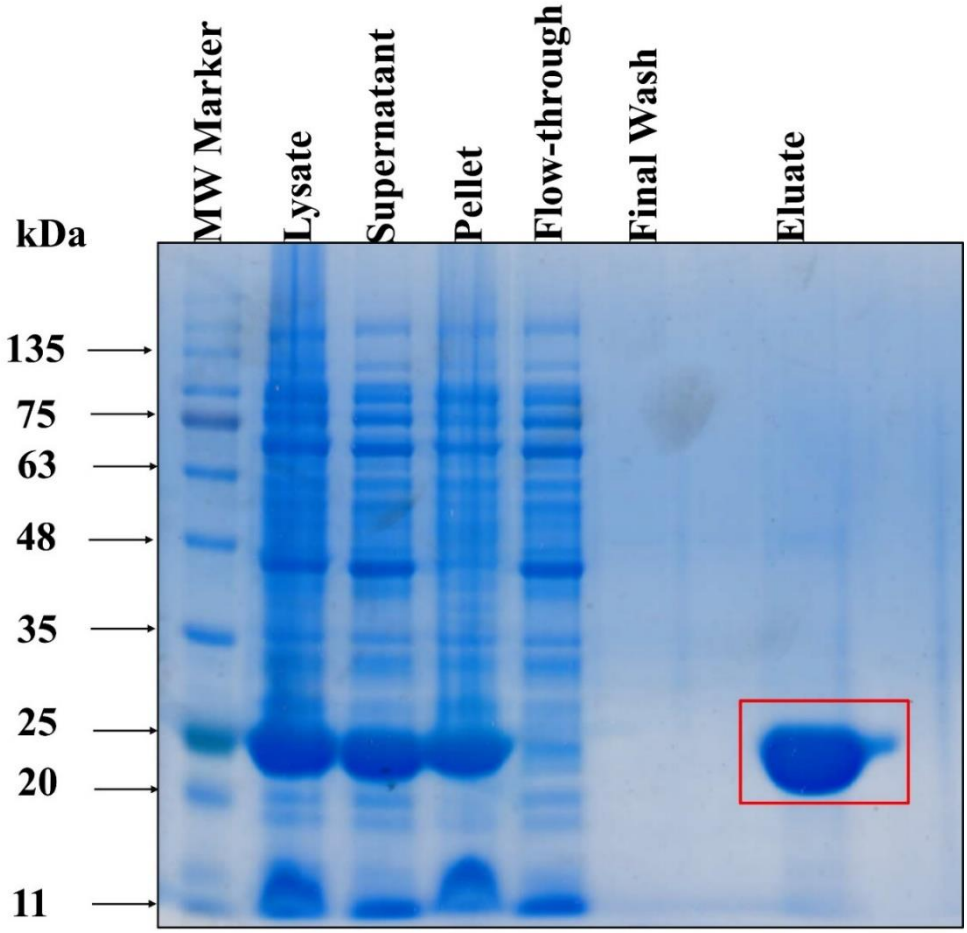



**Figure 3.2. Expression and purification profile of recombinant *E. faecium* NNAT**
Analysis of the crude lysate shows that the protein is overexpressed, evident by a prominent band corresponding to the theoretical molecular weight deduced for EfNNAT sequences (~25.8 kDa). The soluble fraction was separated by centrifugation and purified by $Ni^{2+}$-affinity chromatography using 350 mM im idazole in PBS, pH 7.2, for elution. The purified protein is indicated in the box. Discontinuous SDS-PAGE was used to analyse all fractions and visualised with Coomassie brilliant blue staining.

## 3.6.2 Enzyme activity assay

With the continuous coupled enzyme assay, the activity of the recombinant EfNNAT was examined at varying pH (section 2.3). Results of the investigation of the recombinant enzyme for NNAT activity showed that EfNNAT was only active at pH 8.0 (Figure 3.3). However, this activity was not easily reproducible; hence, the specific activity of the enzyme could not be determined. Nevertheless, the biophysical

structure of the enzyme was examined, which established that the protein is in its native conformation.

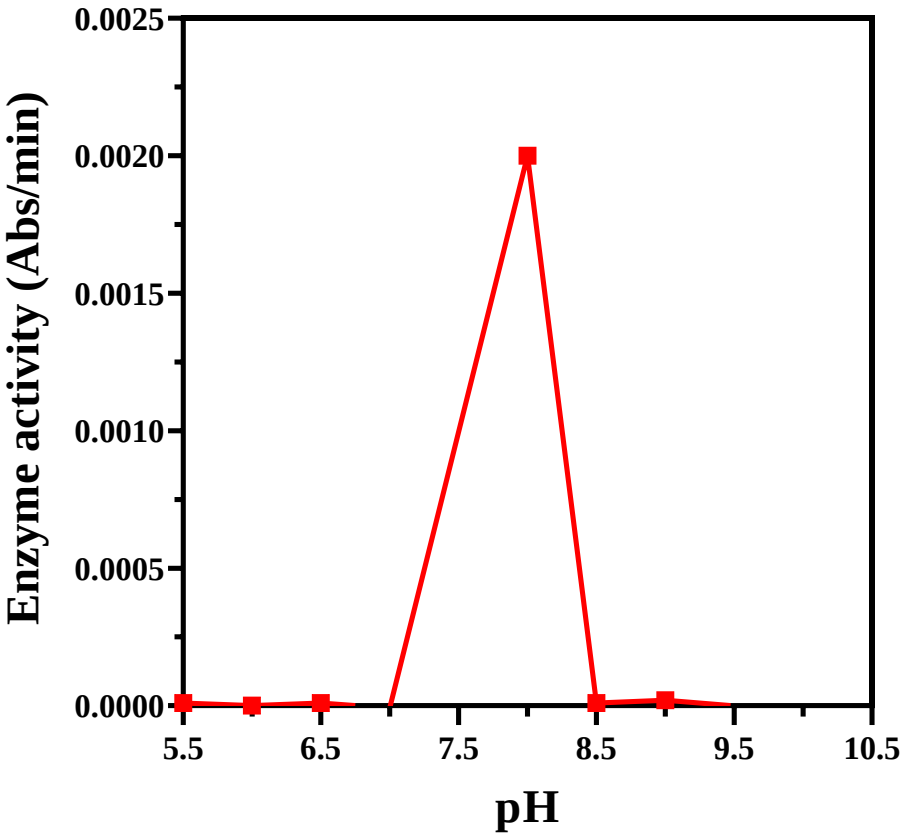



**Figure 3.3 The activity profile of recombinant EfNNAT at different pH**
Analysis was carried out in 0.1 M sodium acetate buffer (pH 5.5), 0.1 M sodium phosphate buffer (pH 6.0–8.0), 0.1 M Tris-HCl buffer (pH 8.5–9.0), and 0.1 M sodium glycine buffer (pH 9.0–10.5) using 1 μM of recombinant EfNNAT, 1 mm ATP and 0.5 mM NMN. EfNNAT demonstrated a narrow pH optimum and a pseudo activity with NMN at pH 8.0, which was not easily reproducible.

### 3.6.3 Secondary structure analysis

Based on the ability of protein peptide bonds to polarise light, the far-UV CD properties of EfNNAT were used to estimate its secondary structural composition. Analysis of the CD spectrum revealed a distinct peak maximum at 190±0.5 nm and minima at 220±0.5 nm and 208±0.5 nm (Figure 3.4), signifying a predominantly $\alpha$-helical secondary structure composition. Upon ATP binding, a decrease in the 190 nm peak maximum and an upward shift at 208 nm peak minimum was observed, suggesting conformational changes to the EfNNAT secondary structure. Upon

denaturation of the protein with 8 M urea, EfNNAT exhibited loss of both peaks implying that the protein existed in its native conformation.

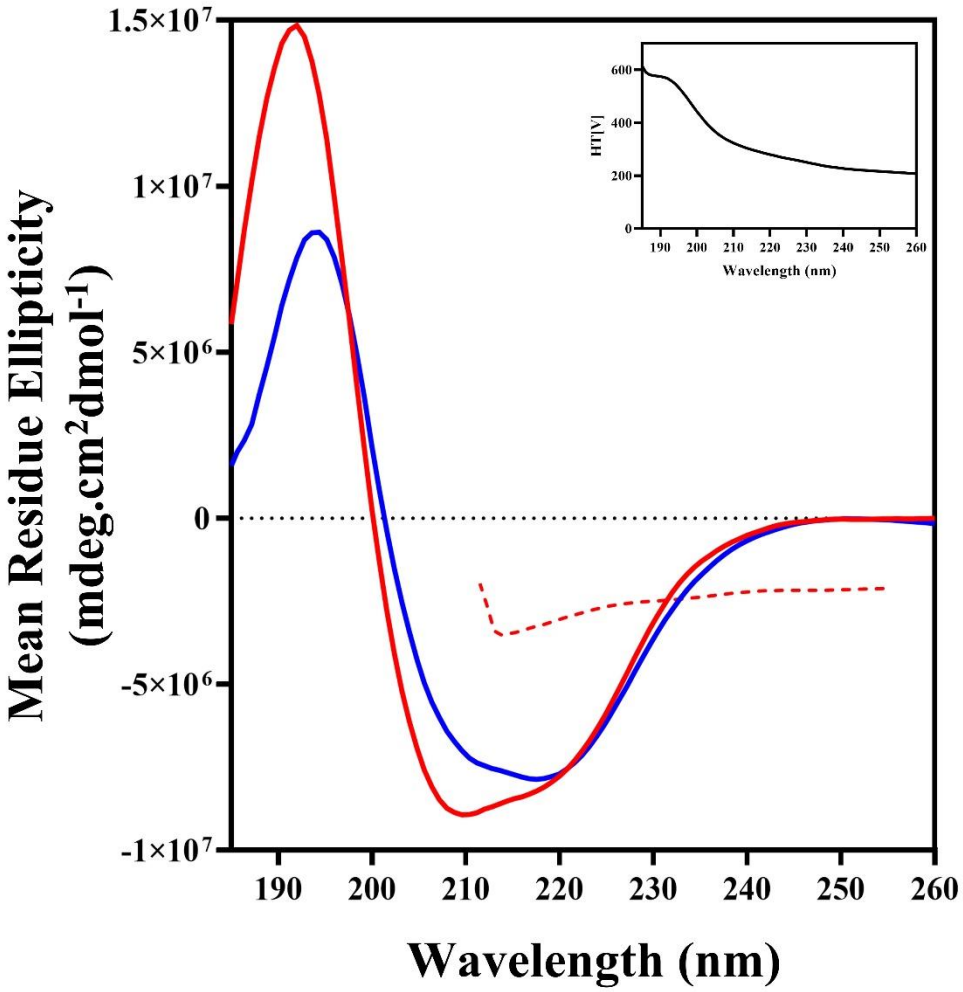



**Figure 3.4. Far-UV circular dichroism spectra of EfNNAT**
in its native (bold red line), ATP bound (blue line), and denatured (dotted red line) forms. 5µM EfNNAT protein was analysed in dd$H_2O$ at 20 °C using a 2 mm quartz cuvette. The EfNNAT-ATP complex form was performed in the presence of 0.2 mM ATP and 1 mM $MgCl_2$. A prominent peak of positive ellipticity at 190±0.5 nm and negative ellipticity at 220±0.5 nm and 208±0.5 nm is observed for the native and bound EfNNAT, which is lost when denatured with 8 M urea. Each spectrum represents an average of five accumulations performed in triplicate. Inset: HT[V] plot indicating the CD signal reliability.

### 3.6.4 Tertiary structure analysis

Given the spectral attributes of ANS (an anionic dye) and mant-ATP (a fluorescent analogue of ATP), the tertiary structure of EfNNAT was studied using fluorescence spectroscopy. The changes arising in the fluorescent and emission properties of the dyes while bound to EfNNAT were used to study the enzyme's binding site and hydrophobicity, as well as its possible structural conformation and binding properties to ATP. The non-covalent binding of ANS to a protein hydrophobic pocket or positively charged amino acid residues and mant-ATP to a nucleotide binding site often results in high fluorescence intensity followed by a blue-shift in the emission wavelength [154, 155]. The data obtained from the ANS binding studies (Figure 3.5) shows a significant increase in ANS quantum yield upon binding to EfNNAT. This increased fluorescence intensity was accompanied by a 7 nm blue-shift emission wavelength, indicating that ANS occupied a hydrophobic pocket in EfNNAT. In the presence of ATP, the ANS fluorescence intensity decreased, resulting in a lower quantum yield than the native EfNNAT. This signifies ANS' exposure to solvent due to displacement by ATP and the possibility of ANS occupying the ATP binding site on EfNNAT. Additionally, ATP binding resulted in a 5nm blue-shift indicating conformational changes in the tertiary structure of EfNNAT.

Further characterisation of the ATP binding site and the proclivity of EfNNAT to bind ATP was probed using a fluorescent analogue of ATP (Figure 3.5). Similar to the ANS binding studies, the interaction of mant-ATP with EfNNAT yielded a high mant-ATP fluorescence intensity complemented by a blue-shift from 450 nm to 442 nm, indicating the binding of mant-ATP to a nucleotide-binding site. However, the presence of ATP in the sample displaced mant-ATP as observed by the quenching of mant-ATP fluorescence coupled with a red-shift in emission wavelength from 442 nm to 454 nm, suggesting changes in the hydrophobicity of the binding site.

Compared to the native EfNNAT, a remarkable decrease in the fluorescence intensity of ANS and mant-ATP was observed for denatured EfNNAT, implying that

the fluorescent dyes are exposed due to limited access to binding pockets on the protein arising from its unfolding. Also, a 2 nm red-shift was observed in the ANS binding studies; however, no significant change was observed in the emission wavelength for mant-ATP, suggesting that the EfNNAT binding site is either surface-exposed or located in a more polar region.

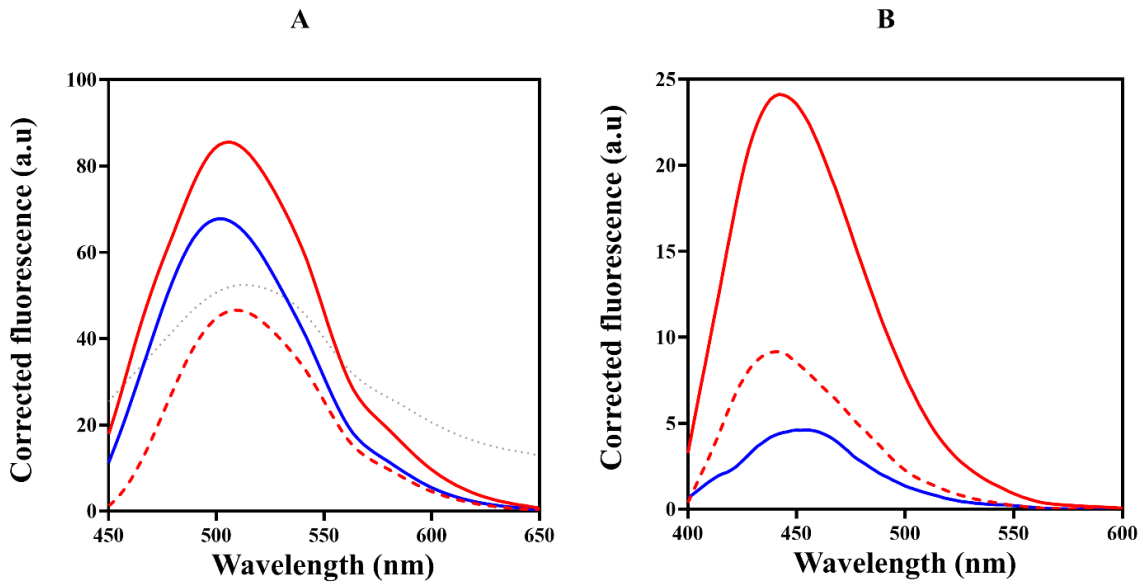



**Figure 3.5. Fluorescence emission spectra** of (A) ANS interaction with EfNNAT apo (red solid), EfNNAT denatured (red dotted), and EfNNAT-ATP complex (blue). Analysis was performed using 50 μM ANS and 2.5 μM protein. ANS excitation was at 395 nm, and emission spectra collected between 400 and 650 nm. Decreased ANS fluorescence and a blue-shift are observed in the presence of ATP. (B) Mant-ATP interaction with EfNNAT apo (red solid), EfNNAT denatured (red dotted) and EfNNAT-ATP complex (blue). Analysis was performed using 10 μM mant-ATP and 2.5 μM protein. Mant-ATP was excited at 355 nm, and emission spectra collected between 400 and 600 nm. The binding of ATP to EfNNAT resulted in the fluorescence quenching of mant-ATP. All analyses were performed in 10 mM $NaH_2PO_4$, pH 7.5, 0.01% (w/v) $NaN_3$, and 1 mM DTT at 20 °C. Each spectrum represents an average of 5 accumulations performed in triplicate.

### 3.6.5 Structural stability study

Insight into the structural stability and the impact of ATP on EfNNAT stability was accessed via a fluorescence thermal shift assay performed with SYPRO Orange dye. The binding of SYPRO Orange to proteins is such that as the temperature increases, the protein unfolds, allowing the dye to access more hydrophobic pockets,

consequently increasing its fluorescence. The significant increase in fluorescence intensity corresponds with increasing temperature, resulting in a sigmoidal curve that signifies the unfolding of the protein. The thermal shift assay results monitored between 20 °C and 80 °C showed EfNNAT apo to exhibit a $T_m$ of 43 °C (Figure 3.6). Upon the binding of ATP to EfNNAT, the unfolding curve and melting peak shifted to the right, with an increasing $T_m$ from 43 °C to 53.5 °C, indicating increasing protein stability.

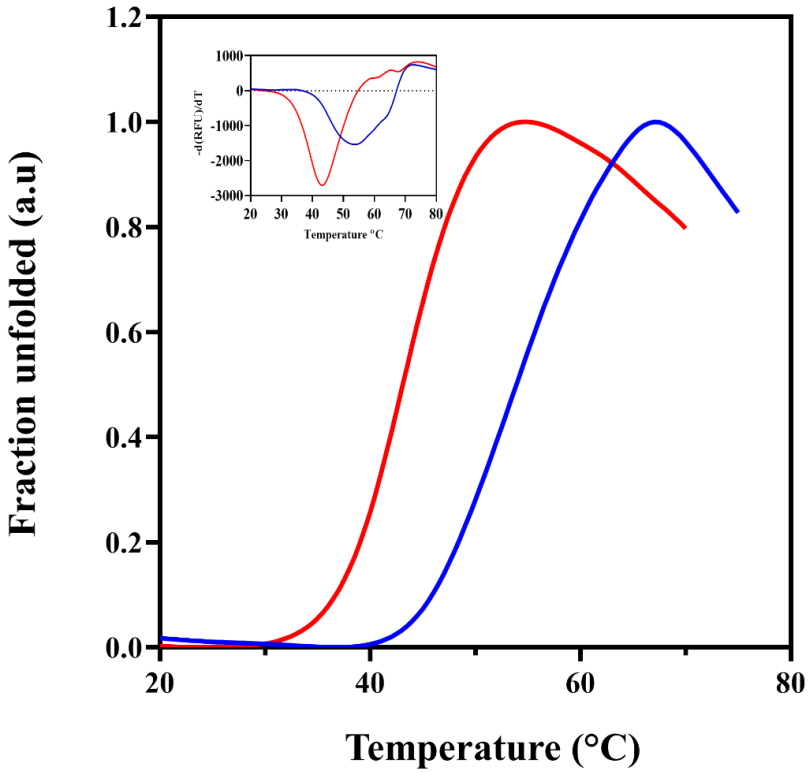



**Figure 3.6. The fluorescence thermal unfolding profile of recombinant EfNNAT apo (red) and in the presence of ATP (blue).**
Assay was carried out using 20 µg of protein and 5 µL of 50 × SYPRO Orange dye supplemented with 0.1 mM ATP and 2 mM $MgCl_2$ and monitored between 20 °C and 80 °C in 25 mM Tris-HCl, pH 7.5. As the temperature increases, fluorescence emission also increases, giving rise to a sigmoidal curve that shows the transition from folded to unfolded protein. Inset: melting peak of EfNNAT apo (red) and EfNNAT-ATP complex (blue). All assays were performed in quadruplicate.

### 3.6.6 Quaternary structure analysis

Using SE-HPLC, the apparent molecular weight and the overall native state of EfNNAT in solution were estimated (Figure 3.7). A single peak with a retention time of about 16.95 min was detected at 280 nm on the SE-HPLC chromatogram. This peak was estimated to be ~40 kDa on the calibration curve, approximating it to a dimer.

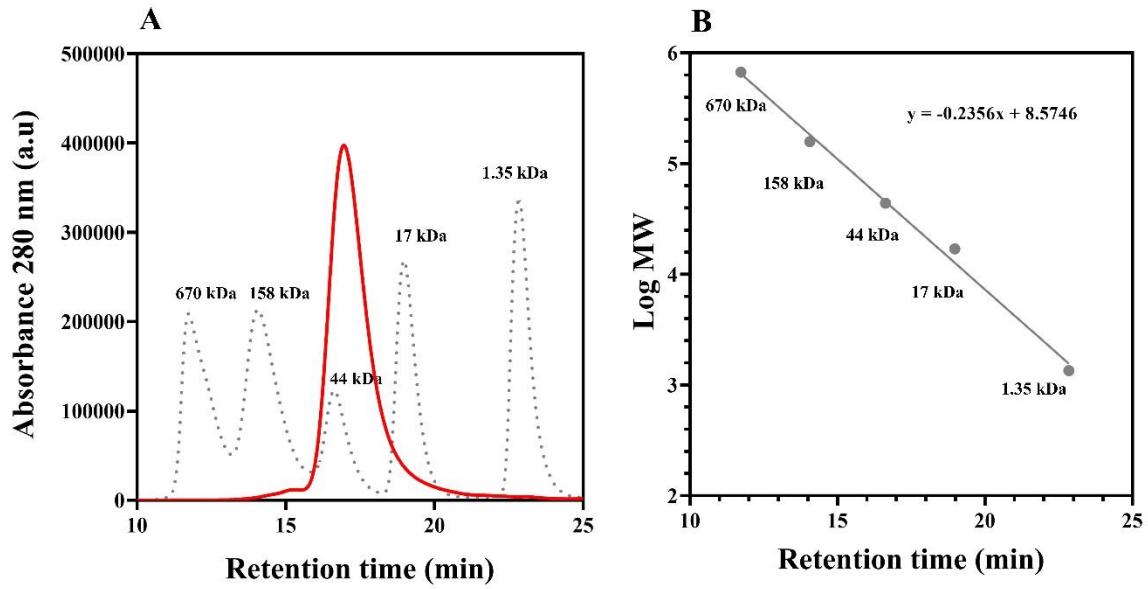



**Figure 3.7. Size Exclusion-HPLC analysis of EfNNAT.**
**(A)** Chromatogram of standard proteins and recombinant EfNNAT resolved at 20 °C in 10 mM $NaH_2PO_4$, pH 7.5, 750 mM NaCl, and 0.01% (w/v) $NaN_3$. The elution profile of the standard proteins is depicted in grey dotted lines, and EfNNAT (~ 6.5 mg/mL) in red. The standard proteins used and their corresponding molecular weight include Vitamin B 12 (1.35 kDa), Myoglobin (17 kDa), Ovalbumin (44 kDa), γ-globulin (158 kDa), and Thyroglobulin (670 kDa). **(B)** Constructed molecular weight calibration curve used to estimate the quaternary size of EfNNAT as ~40 kDa. Each chromatogram represents an average of three replicates.

### 3.6.7 Three-dimensional structures of apo and adenine-bound EfNNAT

The screening and setting up of EfNNAT crystal by sitting drop vapour diffusion using the Index crystallisation screen kit (HR2-144, Hampton Research) yielded many diffractive crystals under several crystallisation conditions (Figure 3.8). The presence of the N-terminal His-tag did not impact the crystal formation or the diffraction quality. The crystals were bipyramidal in morphology and belonged to the

I222 space group. Their diffraction quality ranged between 1.78 – 2.8 Å. The EfNNAT apo crystal used for this study was obtained from the crystallisation condition containing 0.1 M Tris pH 8.5, 2.0 M ammonium sulfate (Index Reagent #6). The crystal diffracted to a resolution of 1.9 Å, and its unit cell was determined as a = 63.650 Å, b = 64.540 Å, c = 108.540 Å. Only one crystallographic molecule of EfNNAT was observed in the asymmetric unit containing 184 amino acid residues, 105 water molecules, one sulfate ion ($SO_4^{2-}$), and one dihydrogenphosphate ion ($H_2PO_4^-$) (Figure 3.9).

The adenine-bound EfNNAT crystals were obtained under the crystallisation condition 0.2 M lithium sulfate monohydrate, 0.1 M Bis-Tris pH 5.5, 25% w/v polyethylene glycol 3350 (Index Reagent #74). The crystal diffracted to a resolution of 1.82 Å, with unit cell *a* = 62.511 Å, *b* = 65.644 Å, *c* = 108.373 Å. The X-ray structure also revealed one crystallographic molecule of EfNNAT in the asymmetric unit containing 181 amino acid residues, 90 water molecules, one adenine molecule, one sulfate ion ($SO_4^{2-}$), and one magnesium ion ($Mg^{2+}$) (Figure 3.9). The crystal was initially set up for EfNNAT apo, and adenine was not present in the crystallisation solution; however, a clear electron density for adenine was observed in the structure (Figure 3.11). Detailed crystallographic and refinement statistics of the EfNNAT structures and their PDB code is in Table 3.2.

The N-terminal His-tag (MGS-HHHHHH) along with the first 22/23 residues, were disordered in the EfNNAT apo and adenine-bound structures, respectively. Residues 69-73, 157, 213, and 214 in the EfNNAT apo and residues 65-72, 213, and 214 in the adenine-bound structures were also disordered. These regions, as indicated by the predicted secondary structure of the protein sequence executed via an online-based algorithm, PSSpred [156], are loop regions which are highly flexible (Figure 3.10) and are common to the *Enterococcus spp* (Figure 3.12). Superimposition of the EfNNAT apo structure on the adenine-bound (Figure 3.11) showed that dihydrogenphosphate and adenine molecule binds similar positions in the EfNNAT structures and are stabilised mainly through hydrogen bonds and non-bonded interactions with residues Gly32, Thr33, Phe34, His38, His41, Ser178 and 179. The

sulfates occupy a similar position in the two structures, and the interactions are almost the same, with hydrogen bonds being formed with the amide backbone of Thr108 and Tyr140 as well as non-bonded contact with Tyr107, Thr108, Lys138 and Trp139.

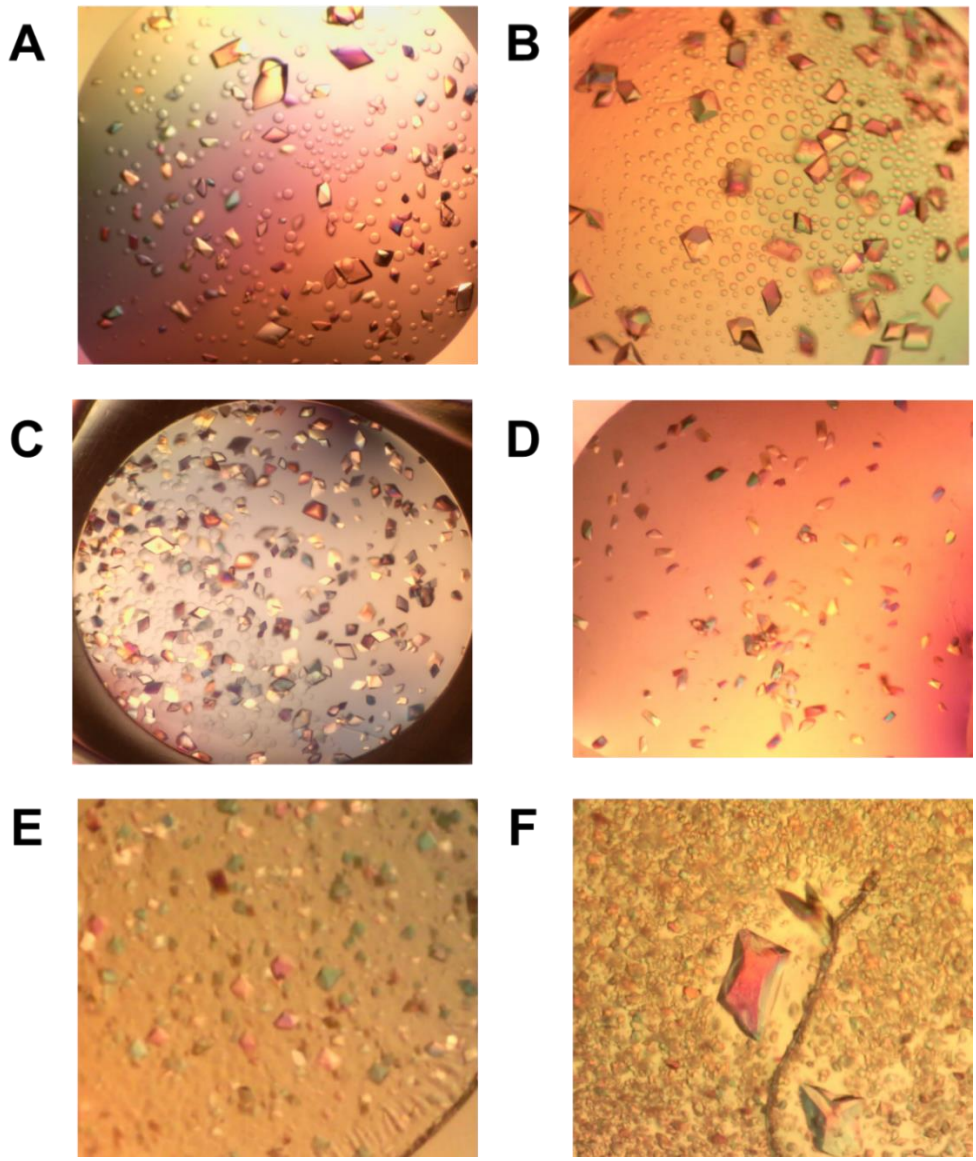



**Figure 3.8. Images of some high-quality diffractive crystals of EfNNAT obtained under different crystallisation screen conditions**

Images of some high-quality diffractive crystals of EfNNAT obtained under different crystallisation screen conditions using the Index crystallisation screen kit (HR2-144, Hampton Research) at 20 °C. **(A)** Index Reagent #6 **(B)** Index Reagent #74 **(C)** Index Reagent #35 **(D)** Index Reagent #17 **(E)** Index Reagent #32 **(F)** Index Reagent #88. The crystals were obtained by sitting drop vapour diffusion method using an initial protein concentration of 15 mg/mL prepared in 50 mM HEPES, 150 mM NaCl, pH 7.2, 2 mM DTT and 0.01% (w/v) $NaN_3$. The final protein concentration yielding these crystals amounted to 7.5 mg/mL.

**Table 3.2. EfNNAT crystal data collection and refinement statistics**

| | Apo | Adenine-bound |
|---|---|---|
| ***Data Collection*** | | |
| Space group | I222 | I222 |
| Unit-cell parameters | | |
| a, b, c (Å) | 63.65, 64.54, 108.54 | 62.51, 65.64, 108.37 |
| Resolution range (Å) | 45.32- 1.9 (1.93 - 1.9) | 45.27 - 1.82 (1.85 - 1.82) |
| Completeness (%) | 99.9 (99.8) | 98.7 (97.7) |
| No. of unique reflections | 17988 (874) | 19984 (979) |
| $R_{merge}$ | 0.22 (6.1) | 0.12 (3.4) |
| I/sigma(I) | 5.9 (0.6) | 10.4 (0.4) |
| Molecules per asymmetric unit | 1 | 1 |
| ***Refinement Statistics*** | | |
| Final overall $R_{factor}$ (%) | 23.8 | 20.7 |
| $R_{work}/R_{free}$(%) | 23.6/27.9 | 20.6/23.3 |
| Overall Mean B-value (Å$^2$) | 54.59 | 56.93 |
| No. of residues in asymmetric unit | 184 | 181 |
| No. of water molecules | 105 | 90 |
| No. of adenine molecules | 0 | 1 |
| No. of bound ions: | | |
| Sulfate ion ($SO_4^{2-}$) | 1 | 1 |
| Dihydrogenphosphate ion ($H_2PO_4^-$) | 1 | 0 |
| Magnesium ion ($Mg^{2+}$) | 0 | 1 |
| RMSD from ideal geometry: | | |
| Bond lengths (Å) | 0.014 | 0.007 |
| Bond angles (degrees) | 1.353 | 0.821 |
| Ramachandran statistics | | |
| Favored (%) | 96 | 94 |
| Allowed (%) | 4 | 6 |
| Outliers (%) | 0 | 0 |
| PDB ID | 8AIH | 8AII |

Values in parentheses represent the highest-resolution shell.

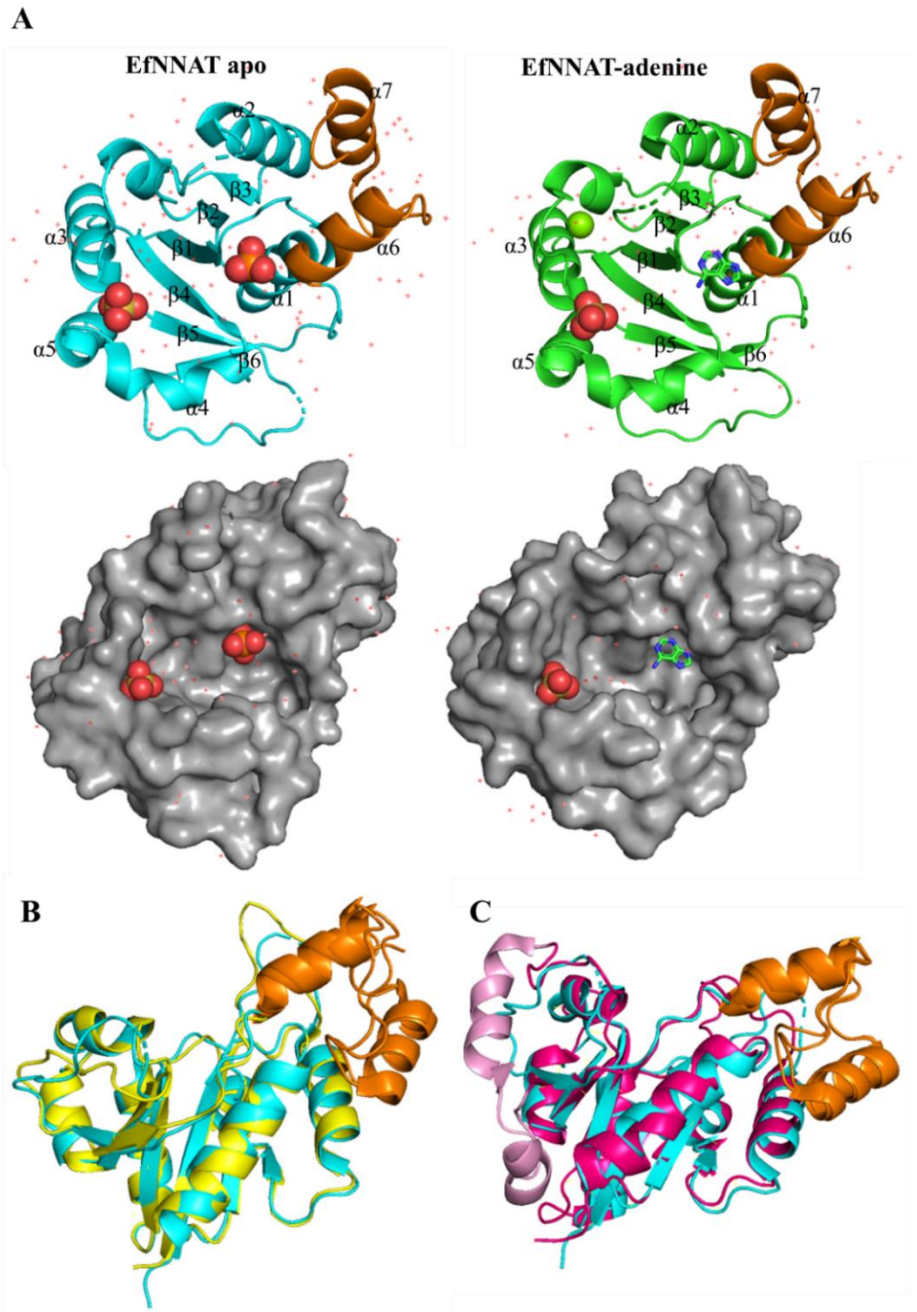



**Figure 3.9. Overview of EfNNAT structures and comparison with gram-positive and negative bacterial NNAT**

**(A)** Cartoon representation of EfNNAT apo (cyan) and EfNNAT-adenine complex (green) showing the secondary structure elements, the large N-terminal domain in cyan/green containing the Rossman fold, and the helical C-terminal domain in brown. The respective surface representations showing the putative binding site is directly below the cartoon structures. Bound to the apo structure are $SO_4^{2-}$ ions in red and yellow and $H_2PO_4^-$ in red and orange spheres. The adenine complex contains adenine molecules shown in green and blue sticks, $SO_4^{2-}$ in red and yellow spheres and a magnesium ion in green sphere. The red dots represent water molecules. **(B)** Superimposition of EfNNAT apo (cyan) with gram-positive *B. anthracis* NNAT 2qtm (yellow). The overall architecture and topology fold are similar with a rmsd of 0.779 Å for 164 atoms. **(C)** Superimposition of EfNNAT apo (cyan) with gram-negative *E. coli* NNAT 1k4k (pink). An additional αβα unit insertion shown in light pink is observed in *E. coli*, which is absent in EfNNAT. RMSD = 1.074 Å for 154 to 154 atoms.

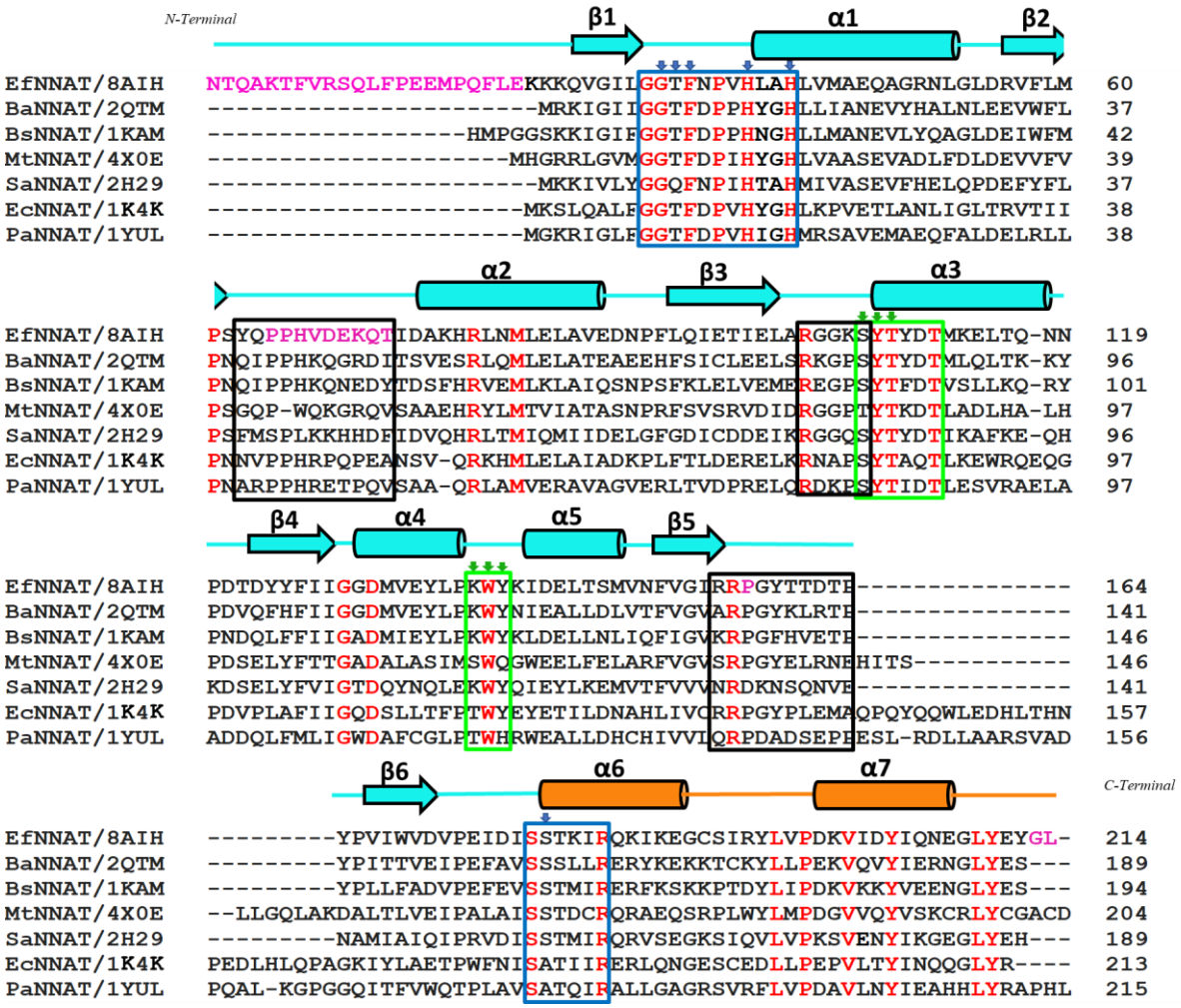



**Figure 3.10. Amino acid sequence alignment of NNAT from *E. faecium*, *B. anthracis*, *B. subtilis*, *M. tuberculosis*, *S. aureus*, *E. coli* and *P. aeruginosa***

The secondary structure elements of EfNNAT based on sequence prediction is shown on top of the sequence with the cylinder and arrows representing $\alpha$-helices and $\beta$-strands, respectively. The larger N-terminal domain is indicated in cyan, and the C-terminal domain in brown. Conserved residues are coloured red, and the missing residues from the EfNNAT crystal structures are coloured magenta. The ATP binding motifs are shown in a blue box, the green box indicates the NMN/NaMN binding motif, while the black boxes indicate loop regions with conformational flexibility in EfNNAT, which are also reported to undergo significant conformational change upon substrate/product binding in *E. coli*, *B. anthracis* and *B. subtilis*. Residues making bonded and non-bonded contact with adenine and dihydrogen phosphate are shown with a blue arrow, while those interacting with $SO_4^{2-}$ are shown with a green arrow. The sequence identifier code used is the pdb code for the protein structures. Sequence alignment was performed with Clustal Omega, and the EfNNAT predicted secondary structure element (SSE) was generated via PSSpred. The predicted SSE exhibits more than 95% similarity with the EfNNAT crystal structure.

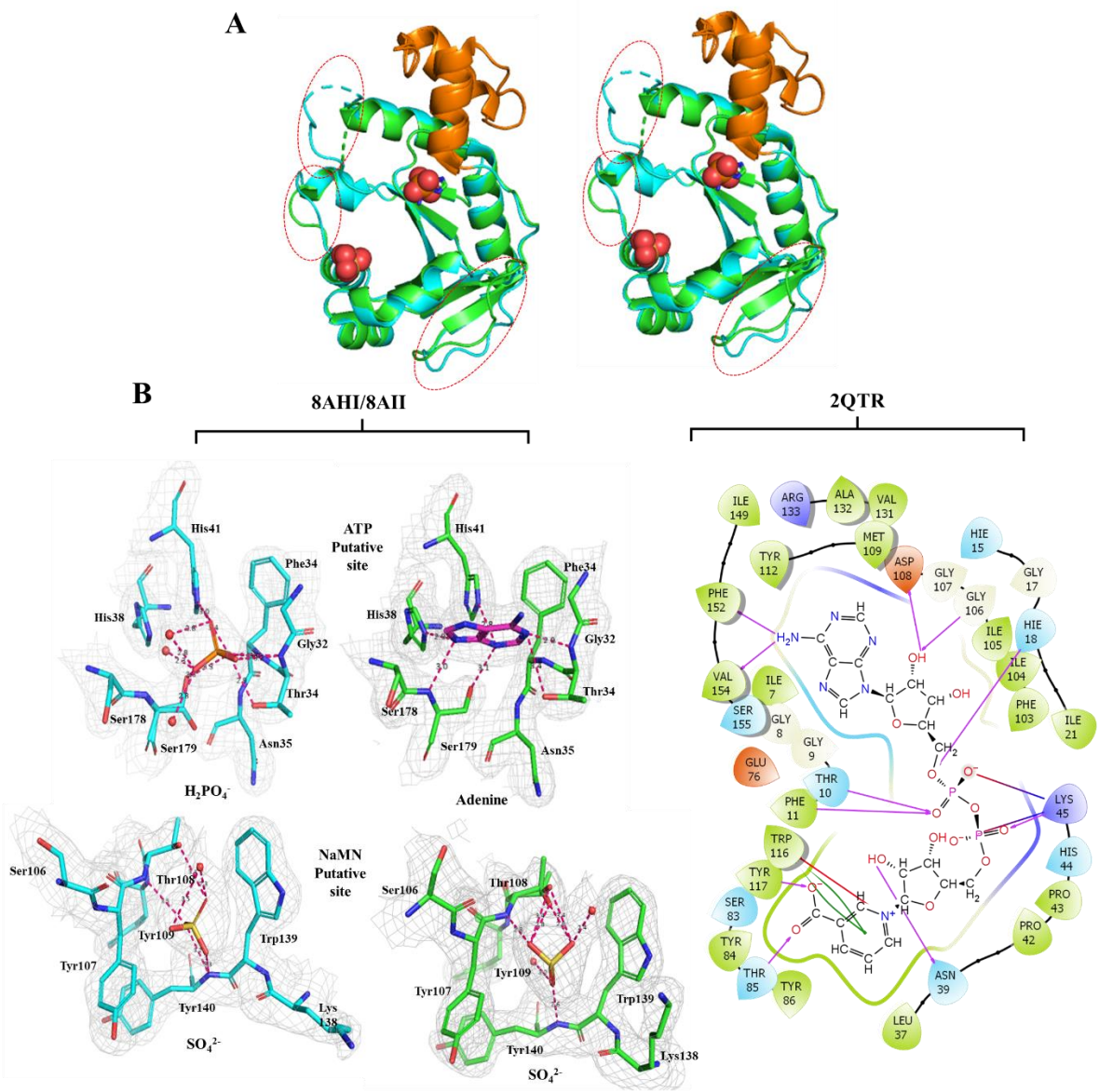



**Figure 3.11. Stereo image of EfNNAT structural flexibility and the Fo-Fc electron density map showing the active site interaction**

**(A)** Superimposition of EfNNAT apo (cyan) and EfNNAT-adenine complex (green) structures showing the bound $H_2PO_4^-$ in red and orange spheres overlaying adenine in green and blue sticks and the two $SO_4^{2-}$ ions in red and yellow spheres overlaying each other. Regions with structural deviations are depicted with dotted circles. **(B)** Comparison of EfNNAT active site interaction (8AHI/8AII) with *B. anthracis* NNAT NaAD complex (2QTR). Sticks in cyan represent the active site interaction of EfNNAT apo (8AHI), while the sticks in green represent the active site interaction of the EfNNAT-adenine complex (8AII). $H_2PO_4^-$ is shown in red and orange sticks, adenine in magenta and blue sticks, and $SO_4^{2-}$ in red and yellow sticks. The hydrogen bond interactions are depicted in pink broken lines. *B. anthracis* NNAT NaAD complex interaction is shown in 2D where residues in blue are polar, green are non-polar, orange is negatively charged, and violet positively charged. The hydrogen bond is represented as purple lines, salt bridge interactions as blue-red lines, π-π∗ interaction as green lines, and grey shades indicate exposure to solvent. The images were captured using a Maestro 2D interaction diagram.

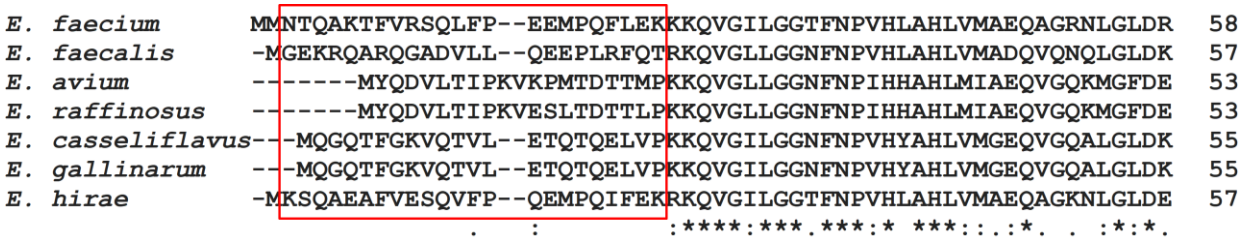



**Figure 3.12. Sequence alignment of seven *Enterococcus* spp**
Sequence alignment of seven *Enterococcus spp*. NNAT's with clinical relevance showing the first 22 amino acid residues uncommon to the other NadD bacterial NNAT. Conserved residues are indicated in asterisk (*****), partially conserved residues with strong similarity in colon (**:**), and those with weak similarity in period (**.**).

## 3.7 Discussion

For the first time, we overexpressed and purified to homogeneity, highly soluble recombinant EfNNAT by cloning an N-terminal hexahistidine fusion EfNNAT gene sequence in a pET-11a vector and growing the vector construct in *E. coli* T7 strain at 30 °C, for 6 h, followed by a one-step $Ni^{2+}$-IMAC purification. The success of the purification system employed and the extent of the protein purity is evident from the series of fractions obtained during the purification process (Figure 3.2). The absence of target protein in the flow-through and wash fractions, as indicated in lanes 5 and 6, suggests the successful trapping of the N-terminal hexahistidine tagged proteins in the $Ni^{2+}$ chelator resin column via selective coordinate covalent binding between $Ni^{2+}$ and the imidazole ring of the polyhistidine tag. Furthermore, the appearance of one band with a molecular mass corresponding to ~25.8 kDa (estimated sequence molecular weight) in the elution fraction (Figure 3.2, lane 7) indicates the successful isolation of the target protein. Compared to the preceding fractions analysed (lanes 2-6), the quality of the elution fractions (lane 7) showed more than 98% pure protein, free of nucleic acid contamination and aggregation. The expression of ~101 mg of pure EfNNAT per 7.8 g of wet *E. coli* cells per litre culture suggests that large-scale production of pure recombinant EfNNAT can be achieved using this protocol.

Insight from the functional studies revealed that recombinant *EfNNAT might be active* and possess NNAT activity, *as evident by an* increase in absorbance at 340 nm, pH 8.0 (Figure 3.3). Compared to most characterised NNAT enzymes, *which demonstrate* broad pH optima between 6.0 and 9.0 [120, 150, 157-159], EfNNAT only displayed significant activity at pH 8, which was not easily reproducible. This we presumed to be due to EfNNAT having a poor affinity for the NMN substrate, given that previous studies have revealed that the NadD bacterial NNAT exhibited a substrate preference for NaMN over NMN [160, 161], as demonstrated in *E. coli* NNAT [66], *S. aureus* NNAT *[119, 162]*, *B. anthracis* NNAT [144] and *B. subtilis* NNAT [145]. Moreover, Osterman and Begley identified that most gram-positive bacterial pathogens utilise the PncA-PncB route to salvage $NAD^+$ synthesis [21] rather than the concomitant usage of the NMN salvage route. Hence, the NNAT from gram-positive may display a reduced preference for NMN, as established by a study in our lab, in which NNAT from a gram-negative bacteria *K. pneumonia* demonstrated significant activity with NMN compared to NNAT from gram-positive *E. faecium* [150, 163].

The presence of an uncleavable histidine tag is unlikely to have affected the enzyme's activity, considering that another NNAT from *K. pneumonia,* which was designed in our laboratory with a slightly longer tag (histidine tag and thrombin cleavage site), showed activity with NMN [150, 163]. Moreover, different studies have reported histidine-tagged NNAT of varying species active and comparable to their untagged native form [144]. However, this does not rule out the possibility of the histidine tag interfering with the activity, as studies have also shown the presence of polyhistidine tag in the binding site of some recombinant proteins, interfering or acting as an inhibitor [164]. Again, the thought that the signature motif (H/TXXH) involved in ATP binding is located close to the N-terminal (His 38, Leu 39, Ala 40, His41) of the protein sequence and given the high flexibility of the polyhistidine region along with the first 22 amino acids (Figure 3.10), the structural conformation of the signature motif may have been impacted. However, as verified from the X-ray structures, these assumptions are incorrect, considering that no polyhistidine tag was

found in the binding site. Rather, ligands that could mimic the natural substrates were found in the binding sites (Figure 3.11). Furthermore, a comparison of the H/TXXH signature motif feature of the predicted secondary structure (Figure 3.10) with the three-dimensional structure of the protein showed that the conserved helical structure is sustained.

The sequence alignment of EfNNAT with other homologous bacterial NNAT (Figure 3.10) showed that the first 20-23 amino acids are not recorded for the other members of the NNAT NadD family, which is surprising. Hence, the sequence alignment of seven *Enterococcus spp*. NNAT with clinical relevance was carried out (Figure 3.12), which showed that these extended residues are not uncommon to the *Enterococcus spp*. Worth noting is that this region is not conserved, and the prediction of the secondary structure of the sequences using PSSpred software [156] further revealed this region of the protein as unordered in the *Enterococcus spp*., which justifies the poor visibility and lack of electron density observed for EfNNAT structure at this region. Given that this unordered region of the protein is not conserved, coupled with the fact that it is located at the beginning of the N-terminal of the sequence, it is reasonable to conclude that its absence from the crystal structure of EfNNAT should have no significant impact on the overall structure and function. Also, the disordered residues observed between 65-73 in the EfNNAT structures are not unusual, considering it is a loop region as indicated by the predicted SSE (Figure 3.10), which displays over 95% similarity with the EfNNAT crystal structure. Furthermore, this corresponding region in some NadD bacteria NNAT, including *B. subtilis* NNAT [145] and *M. tuberculosis* NNAT [146] apo structures, have also been reported disordered, which, however, becomes ordered in their complex form as seen in *B. subtilis* NNAT NaAD-complex structure.

Consistent with the far-UV CD studies, which show that EfNNAT possesses a secondary structure composition that indicates a dominant $\alpha$-helices (Figure 3.4), the 3D structure of EfNNAT comprises a larger N-terminal domain distinguished by alternating $\alpha$-helices and $\beta$-strands and a helical C-terminal domain (Figure 3.9 (A)).

The large N-terminal domain consisting of residues 23-177 makes up $\alpha$1-5 and β1-6, which reveals a dinucleotide-binding fold known as the Rossman fold [165]. This super-secondary structure is a well-established feature typical of the nucleotidyltransferase $\alpha$/β phosphodiesterase family and commonly described in both the prokaryotes and eucaryotes NNAT/NMNAT, but exhibit distinct topologies [66, 75, 119, 121, 145, 147, 166]. The parallel β-strands in the core of the EfNNAT N-terminal domain adopt a topological arrangement in the order of 3-2-1-4-5-6**.** This domain is linked to the C-terminal domain, which consists of residues 178-212 that form the $\alpha$-helices 6 and 7. Helix 6 is highly conserved and contains the ATP binding motif SxTxxR (Figure 3.10). The overall architecture and topology fold of EfNNAT is similar to the gram-positive bacterial NNATs like the 2H29 *S. aureus* [119] with an rmsd of 1.103 Å for 155 atoms and 1KAM *B. subtilis* [145] with an rmsd of 0.773 Å for 143 atoms. However, it closely resembles 2QTM *B. anthracis* [144] with an rmsd of 0.779 Å for 164 atoms (Figure 3.9 (B,C)).

In the EfNNAT crystal lattice, the molecule crystallised in the I222 space group with one molecule in the asymmetric unit. The body-centred tetragonal structure of the crystal suggests that another monomer could be generated based on crystallographic symmetry, resulting in a dimer as an oligomeric/functional state of the EfNNAT. Data from the SE-HPLC reveals the native molecular mass to be ~40 kDa (Figure 3.7), which exceeds the theoretical monomeric mass of ~25.8 kDa and the mass obtained from the SDS-PAGE, thus proposing that the enzyme is biologically active as a dimer in solution. Furthermore, the large buried surface area exhibited by the protein structure suggests that the dimerisation of the subunits would likely give rise to a "closed structure" with an intrinsic symmetry [167]. Thus, resulting in a molecular mass somewhat lower than the expected dimeric mass as observed from the SE-HPLC. Given these, we propose the possibility of recombinant EfNNAT functioning as a dimer; hence, agreeing with the consistent distinguishable features observed for NNAT from gram-positive bacteria like the *B. subtilis* [145], *S. aureus* [119], *B. anthracis* [117] and *M. tuberculosis* [146]. The higher proportions of $\alpha$-helices and β-strand than

unordered coils exhibited by EfNNAT, coupled with its dimeric nature, are perhaps factors contributing to the relatively high stability displayed by the enzyme (Figure 3.6) even at a concentration above 50 mg/mL and its ability to crystallise easily.

The co-crystallisation of sulfate ions, dihydrogenphosphate ion, and adenine with EfNNAT further aided the assessment of the binding site in the absence of the substrates or product (Figure 3.11). The presence of sulfate and dihydrogenphosphate ions in the crystal structure is no surprise, given that the crystallisation solution contained 0.2 M lithium sulfate monohydrate/2.0 M ammonium sulfate and the purification buffer contained phosphate. Though adenine was not in the crystallisation solution, the apparent electron density as seen in the map, is unmistakable and possibly acts as a crosslinker that helps to stabilise the protein to crystallise (Figure 3.11). Its presence in the crystallisation condition most likely resulted from the bacterial source during expression.

Inspection and comparative analysis of the residues involved in stabilising the anions and the adenine molecule in the binding sites of the EfNNAT crystal structures with the NaAD-bound *B. anthracis* NNAT site revealed the presence of conserved residues and motifs implicated in substrate and product binding (Figure 3.10 and 3.12**)**. This suggests that the sites occupied by these ligands are EfNNAT's putative binding sites. The surface representation of the enzyme, as indicated in Figure 3.9 A, shows that the EfNNAT binding site is a large cleft characteristic of most reported NNAT binding sites. This large cleft appeared exposed and solvent accessible, which possibly justifies the minimal or insignificant shift observed in the ANS and mant-ATP emission wavelength upon unfolding the protein in contrast to its native form.

The sulfate and dihydrogenphosphate/adenine bind to different sides of the cleft (Figure 3.9). The site occupied by both adenine and $H_2PO_4^-$ is presumed to be the EfNNAT putative ATP binding site, given the productive positioning of the highly conserved residues Gly32, Thr33, Phe34, His38, and His41 of the GXFXPX(H/T)XXH signature motifs and residues Ser178 and Ser179 of the SXTXXR motif which is required for ATP recognition and binding. An extensive hydrogen bonding network

is observed between the residues and the bound $H_2PO_4^-$/adenine (Figure 3.11 (B)). Residues Phe34 make hydrogen bond contact with oxygen2 atom of $H_2PO_4^-$, the nitrogen2 (NE2) atom of His41 forms hydrogen bonds with oxygen4 and the amide backbone of Thr33 hydrogen bonds with oxygen1 and oxygen2. Adenine is bound by the hydrogen-bonding network of the NE2 atom and side chain of His38 and the amide backbone and hydroxyl group of Ser179. These corresponding residues and interactions are also conserved and reported in the ATP moiety region of the *B. anthracis* NNAT NaAD complex, as shown in Figure 3.11 (B) and in the ATP complex structure of *P. aeruginosa* NNAT [166]. An additional hydrogen bond with three water molecules is observed in the dihydrogenphosphate interaction plot, which has also been reported in some bacterial NNAT to interact with the AMP-phosphate [144]. This interaction elucidates the changes observed in the hydrophobicity of EfNNAT binding site in the presence of ATP as indicated by a red-shift in the mant-ATP fluorescence emission wavelength, suggesting a solvent-accessible pocket.

Conversely, the stabilisation and interaction of the sulfate ions with the highly conserved residues Thr108 and Trp139, which are associated with pyridine ring stabilisation, and the partially-conserved Tyr140 suggest the possibility that the site occupied is the putative NMN/NaMN binding site. Visual inspection and comparative analysis of the interactions stabilising the sulfate ions in EfNNAT with the NaMN complexed structure of *B. anthracis* NNAT showed that the sulfates mimic the carboxylate group of NaMN. An extensive hydrogen-bond network is observed between the amide backbone of Tyr140 and Thr108 and the hydroxyl group of Thr108 with the oxygen atoms of the sulfates. Because the NaMN moiety of the NaAD occupies the same orientation as in the *B. anthracis* NNAT NaMN-complexed structure, a similar interaction is observed between the carboxylate group of the bound NaAD and the corresponding residues Tyr117 and Thr85 (Figure 3.11 (B)) [144]. The proximity and interactions of the conserved residues with the adenine molecule, phosphate and sulfate ions suggest their conformation is in a productive position. Hence, the poor activity demonstrated by the enzyme in solution is unlikely due to

poor structural conformity but rather substrate preference arising from the interactions with the carboxylate group of NaMN, which is absent in NMN. However, activity study using NaMN as a substrate and co-crystallisation of EfNNAT with either NMN or NaMN would help to validate this.

Structural comparison of the EfNNAT apo with the adenine-bound shows that both structures are almost similar, with an overall RMSD of 0.262 Å for 149 to 149 atoms. However, conformational flexibility is observed between the two structures at the loop regions of residues 63-74, which contains the partially conserved motif PPHX reported to function as an arm to recognise and bind NaMN [66, 119, 145, 166] or bind NaAD [144]. Loop regions of residues 102-106 close to the NaMN binding site and residues 155-165 located at the surface of the molecule (Figure 3.10 and 3.12 (A)). At least one or more of these corresponding loop regions in other characterised NadD bacterial NNAT have been shown to undergo conformational changes and become ordered upon substrate or product binding [66, 119, 145].

The conformational flexibility shown by the EfNNAT crystal structures is also reflected in solution as observed from the far-UV CD (Figure 3.4) and ANS fluorescence spectroscopy studies (Figure 3.5), where ATP binding induces changes to the secondary and tertiary configuration of the protein. Considering that EfNNAT exhibits a binding pocket that appears to be solvent exposed, we presume that the accommodation of substrate/product by EfNNAT may require reconfiguration of the side chains of some residues lining the binding domain, such that a transient hydrophobic subpocket is formed or a more closed active site as reported in *E. coli* [66], *B. anthracis [144]*, and *B. subtilis* [145] NNATs upon NaAD binding. This conformational flexibility displayed by EfNNAT may not only be necessary for substrate binding but perhaps play a role in bringing the substrates into close proximity to facilitate the nucleophilic attack necessary for product formation and product release.

Besides the secondary and tertiary conformational changes, the structural stability of EfNNAT is also impacted upon ligand binding, as evident from the thermal

shift assay where EfNNAT $T_m$ increases from 43 °C to 53.5 °C upon ATP binding (Figure 3.6). The conformational change induced by ATP binding possibly rigidifies the enzyme, thus enhancing its stability. Furthermore, the replacement of the glycine residue of the "TH/XGH" motif with alanine may have adopted a more rigid structure for EfNNAT, thus contributing to its high stability both in its bound and unbound forms and perhaps the ease of its crystallisation.

With the increasingly rising levels of antibiotic resistance across the globe, treating patients with nosocomial infections is becoming more and more difficult and impossible. While combating the multidrug-resistant *E. faecium* calls for novel targets, the availability of highly pure recombinant EfNNAT and accessibility to high-resolution structures of EfNNAT would provide further structural evaluation and drug-based screening, leading to the discovery of structure-based inhibitors against this enzyme.

Additionally, seeing that EfNNAT is easily expressed at a relatively large amount in *E. coli*, with a high yield of over 98% pure protein obtained via a one-step purification system, coupled with its ability to easily crystallise in multiple conditions both as apo and in complex with nucleotides, EfNNAT could serve as a teaching tool in protein crystallography. Unlike the standard lysozyme, which is only commercially available and can only be crystallised in the apo form under one condition- here is a protein whose expression and purification can be controlled, likewise co-crystallised with ligands including ATP, CTP, and NMN (manuscript in preparation). Because EfNNAT crystals are easily obtainable within 48-72 h, at 20°C requiring no special condition, and give a high diffraction quality, it could be used as a benchmark or standard and an alternative to lysozyme especially considering that it provides an added advantage of co-crystallisation. Furthermore, EfNNAT could serve as a research tool in drug discovery to study molecular mimicry of nucleotides and the orientation and conformation of these drugs when bound to a protein.

# Chapter 4

# Co-Crystallisation of *Enterococcus faecium* NNAT with NMN and CTP: Crystal Structure and Preliminary X-Ray Analysis.

Olamide Jeje, Ramesh Pandian, and Ikechukwu Achilonu*

Protein Structure-Function Research Unit, School of Molecular and Cell Biology, Faculty of Science, University of the Witwatersrand, Johannesburg 2050, South Africa.

Olamide.Jeje1@students.wits.ac.za, Ramesh.Pandian@wits.ac.za

*Correspondence: Ikechukwu.Achilonu@wits.ac.za; Tel.: +27-11-717-6349

## 4.1 Introduction

Bacterial infections remain a major global health concern, exacerbated by the rise of antibiotic resistance. *E. faecium* is a bacterial pathogen responsible for a range of infections, particularly in healthcare settings [12, 13]. Its increasing resistance to most current antibiotics justifies the urgent need for drug development. NNAT is a key enzyme in the biosynthesis of $NAD^+$ [27], an essential cofactor involved in numerous cellular processes [20-23]. Given its central role in bacterial metabolism and potential as a therapeutic target, NNAT has garnered interest as a target for the development of novel antimicrobial agents. Understanding the structure and function of NNAT in *E. faecium* is crucial for developing novel therapeutic strategies to combat this pathogen.

X-ray crystallography is a powerful technique used to determine the three-dimensional structure of biological macromolecules, including proteins, at atomic resolution, providing valuable insights into their function and mechanism of action. The co-crystallisation of a protein with its ligands or substrates offers a powerful approach to elucidate the structure-function relationships of the enzyme and bound ligand. This chapter presents the crystal structure of EfNNAT bound with NMN and the preliminary X-ray analysis for the dataset obtained for the CTP co-crystal. By elucidating the atomic details of NNAT in complex with NMN, we showed the impact of NMN binding to EfNNAT dimer interface and the possibility of EfNNAT exhibiting an allosteric site, thus laying the foundation for structure-guided drug design targeting this enzyme.

## 4.2 Materials and methods

### 4.2.1 EfNNAT expression and purification

The expression and purification of EfNNAT protein were as previously described (Chapter 3, sections 2.1 and 2.2). Briefly, the gene encoding EfNNAT (UniProtKB-A0A133MWI0 (A0A133MWI0_ENTFC)) was cloned into a pET-11a vector. *E. coli* T7 strains were transformed with the vector, and overexpression was induced by adding 0.5 mM IPTG. The strains were incubated for 6 hours at 30°C, shaking at 200 rpm. The cells were lysed by sonication, and the soluble fraction containing the EfNNAT was purified using $Ni^{2+}$-IMAC.

### 4.2.2 Crystallisation

Co-crystallisation trials were performed using the sitting drop vapour diffusion method. The various crystallisation conditions that yielded EfNNAT apo crystals (Chapter 3, Figure 3.8 - index reagents #6, #17, #32, #35, #74 and #88) were utilised for the trial setup. Samples were prepared by mixing 2 μL of the crystallisation buffer with 2 μL of the protein solution (15 mg/mL EfNNAT prepared in 50 mM HEPES, 150 mM NaCl, pH 7.2, 0.02 % (w/v) NaN3 and 1 mM DTT, supplemented with 2 mM NMN/$NAD^+$ or 5 mM CTP and 5mM $Mgcl_2$). Then, the mixture was equilibrated with 100 μL of their respective crystallisation buffer, incubating at 20 °C. Crystals were observed within 72 hours, then harvested and cryo-protected with either paraben or 20% glycerol mixed with the respective crystallisation buffer. Crystals were flash-frozen with liquid nitrogen and transported to DLS Ltd in the UK for data collection.

### 4.2.3 Data collection and processing

The crystals were subjected to X-ray diffraction at 100 K under a stream of liquid nitrogen, with each frame covering an oscillation angle of 0.10°. The EfNNAT-CTP crystal diffraction dataset was collected with the I24 beamline at DLS

Synchrotron, UK, and then processed and scaled using autoPROC [97]. The EfNNAT-NMN co-crystal diffraction dataset was collected with the I04 beamline at DLS and processed using DIALS [96].

### 4.2.4 EfNNAT-NMN structure solution

The structure of EfNNAT co-crystallised with NMN was determined by molecular replacement using the Phaser-MR (simple one-component interface) program in Phenix [103]. The EfNNAT apo structure (PDB ID: 8AIH) was used as the search model. Further refinement and several sessions of model building and solvation were performed using PHENIX [103]and COOT [107]. The resulting model was then validated for stereochemical quality via wwPDB Validation Service (https://validate-rcsb-2.wwpdb.org/).

## 4.3 Results

### 4.3.1 EfNNAT co-crystallisation and crystallographic analysis

The co-crystal trials were initially set up for EfNNAT-NMN, EfNNAT-NAD and EfNNAT-ATP complex. The EfNNAT-NMN co-crystals grew in the Hampton crystal screens index reagent #17 consisting of 1.26 M sodium phosphate monobasic monohydrate, 0.14M potassium phosphate dibasic pH 5.6. The Monoclinic crystal belonged to the $C12_1$ space group with cell sizes a = 94.45, b = 107.34 and c = 62.34 Å. The crystal diffracted to a resolution of 1.87 Å, and the dataset was measured to a completeness of 97.55%. Two molecules of the protein were observed in the crystallographic asymmetric Unit. Table 4.1 summarises the data-collection statistics.

The EfNNAT-NAD co-crystals were obtained from the crystal screen index reagents #17 and #32. Crystals from both conditions grew in the $C12_1$ space group, diffracting at a resolution between 3.18 Å and 3.5 Å. However, given the poor dataset obtained, further processing and refinement could not be achieved.

On the other hand, an attempt to get ATP to co-crystallise with EfNNAT was unsuccessful, as no ligand was detected in the analysed crystals. Hence, mant-ATP and CTP (a cytidine base nucleoside instead of adenine) were used as alternatives. EfNNAT-Mant-ATP co-crystals were obtained with the index reagents #26, #28, #35 and #88. However, all the diffraction datasets obtained were of low resolution, ranging between 2.65 and 3.75 Å, with completeness below 90%. Hence, the dataset could not be used.

The Ef-NNAT-CTP co-crystals grew in several crystallisation conditions, including #26, #28, #35, #80, and #88. The best EfNNAT-CTP crystals grew in 0.2 M ammonium acetate, 0.1 M HEPES pH 7.5, 25% w/v polyethylene glycol 3,350 (index reagent #80). Crystallographic analysis of the crystal revealed a tetragonal crystal system belonging to the P41212 space group with unit cell dimensions a = 51.89, b = 51.89, and c = 130.88 Å. Only one protein molecule was observed in the unit cell, which diffracted at 1.85 Å resolution. The dataset collected was 100% complete, with Rmerge and Rmeas of 15.2 and 15.7, respectively, and an I/sigma(I) of 10.1. Table 4.1 provides details of the EfNNAT-CTP dataset.

**Table 4.1. EfNNAT Co-crystallisation data collection.**

| | **EfNNAT-NMN** | **EfNNAT-CTP** |
|---|---|---|
| ***Data Collection*** | | |
| Space group | $C12_1$ | P41212 |
| Unit-cell parameters | | |
| A, B, C (Å) | 94.45, 107.34, 62.34 | 51.89, 51.89, 130.88 |
| $\alpha$, $\beta$, $\gamma$ (Å) | 90.00, 128.96, 90.00 | 90.00, 90.00, 90.00 |
| Resolution range (Å) | 48.45- 1.84 (1.87 - 1.84) | 47.63 - 1.86 (1.89 - 1.86) |
| Completeness (%) | 98.28 (97.55) | 100.00 (100.00) |
| Observations | 292,906 (14963) | 260,481 (10565) |
| No. of unique reflections | 41,063 (2032) | 15341 (736) |
| $R_{merge}$ | 0.098 (4.732) | 0.152 (3.318) |
| $R_{meas}$ | 0.106 (4.882) | 0.157 (3.435) |
| I/sigma(I) | 11.4 (0.3) | 10.1 (1.0) |
| Molecules per asymmetric unit | 2 | 1 |

The values in brackets indicate the highest-resolution shell.

### 4.3.2 EfNNAT-CTP structure determination and refinement

The EfNNAT-CTP structure solution was attempted by the molecular replacement method using Mr BUMP of CCP4i suite [101, 151, 152]**.** The MR BUMP structure template was used as the search model for the EfNNAT-CTP structure because the EfNNAT apo structure had not been deposited in the PDB at the time of data collection. Attempts to refine and build the model in PHENIX [103] and COOT [107] were unsuccessful, as no significant reduction was observed in the R-factor. After several interactive rounds of refinement and model building, the model could not be refined below an R-work and R-free value of 31 and 35, respectively.

### 4.3.3 Overview of EfNNAT-NMN complex structure

The crystal structure of EfNNAT-bound NMN was determined to have a resolution of 1.84 Å with an R-factor of 21.77%. Table 4.2 details the refined structure statistics. The crystal contains two molecules in the asymmetric unit, forming a homodimer of chains A and B. The first 20 amino acid residues, hexahistidine-tag, and 64-73, are disordered in chain A. Similarly, in chain B, the first 21 amino acid residues, the hexahistidine-tag and residues 66-72 are disordered. The two monomers associate such that they are anti-parallel to each other (Figure 4.1). Each monomer has a large N-terminal and a smaller C-terminal domain. A series of alternating beta strands and alpha helices define the N-terminal comprising the nucleotide-binding Rossmann fold. A predominant helical structure characterises its C-terminal. Its architecture resembles the EfNNAT apo and adenine-bound structures (Figure 4.2A and B). Also present in the co-crystal structure are one NMN, two sodium ions ($Na^+$), three PEGs, and 188 water molecules (Figure 4.1).

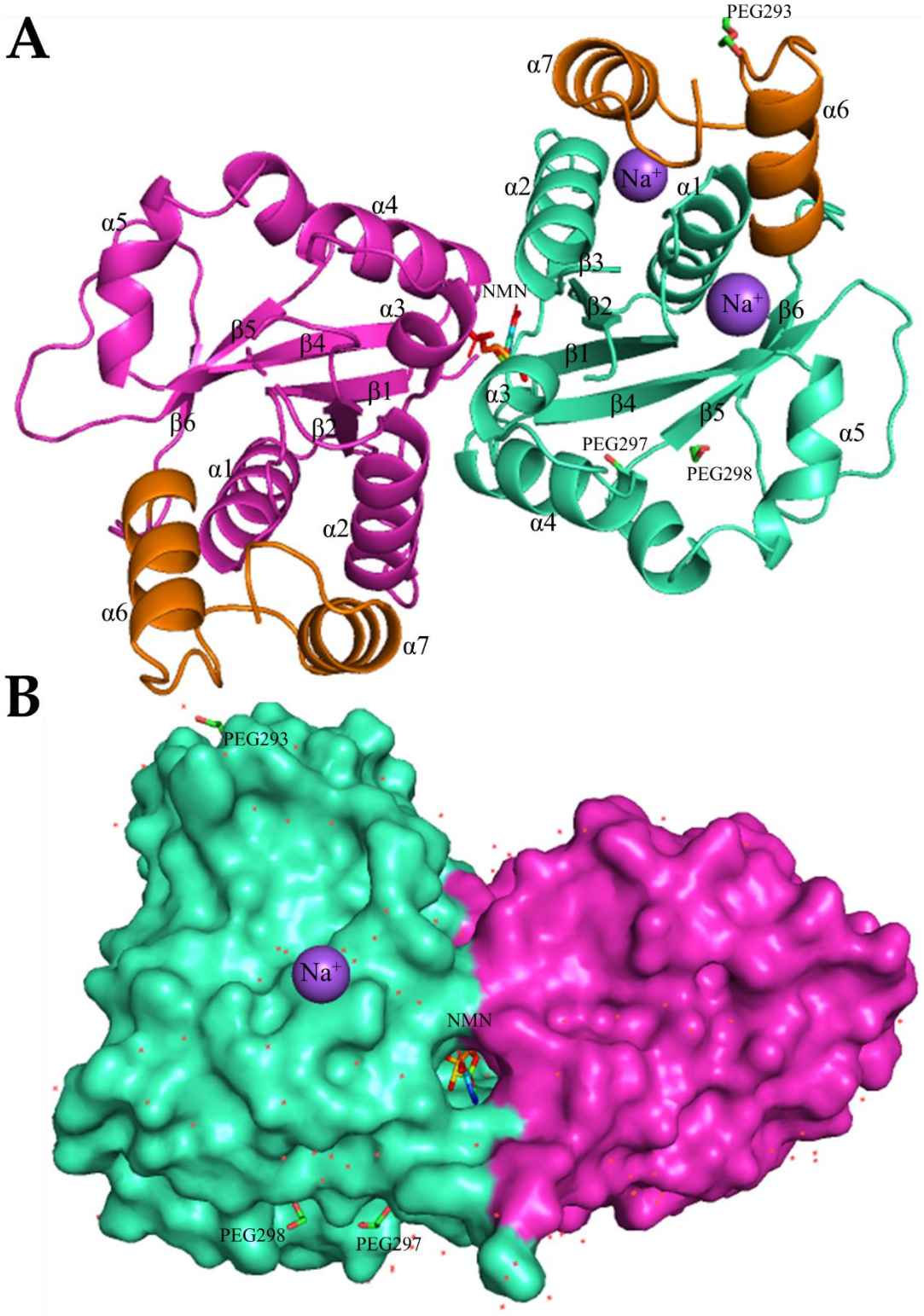



**Figure 4.1. The Crystal structure of *E. faecium* NNAT bound to NMN**

(A) Cartoon representation showing the monomeric subunits separated into the N-terminal domain (magenta/mint green) and C-terminal (brown). Their corresponding secondary structural elements are annotated reflecting the Rossmann fold in the structure and the structural similarity with the apo (Chapter 3). At the dimer interface is the bound NMN shown in rainbow sticks, PEGs in green and red sticks, and $Na^+$ in the purple sphere. (B) Surface representation colour code into monomeric subunits. Bound NMN is shown within the dimer pocket and water molecules in red dots.

**Table 4.2. Summary of EfNNAT-NMN refinement statistics**

| | EfNNAT-NMN |
|---|---|
| ***Refinement Statistics*** | |
| Final overall $R_{factor}$ (%) | 21.77 |
| $R_{work}$(%) | 21.50 |
| $R_{free}$(%) | 26.60 |
| Overall Mean B-value (Å$^2$) | 52.95 |
| No. of water molecules | 188 |
| No. of NMN molecules | 1 |
| Polyethylene glycol (PEG) | 3 |
| Sodium ion ($Na^+$) | 2 |
| RMSD from ideal geometry: | |
| Bond lengths (Å) | 0.009 |
| Bond angles (degrees) | 0.968 |
| Ramachandran statistics | |
| Favoured (%) | 97.49 |
| Allowed (%) | 2.51 |
| Outliers (%) | 0.00 |

Free R value Test set size (%) = 4.95

## 4.3.4 Structural flexibility upon NMN binding

Superimposition of the EFNNAT-NMN structure with the symmetry generated dimer of the EfNNAT apo and the adenine-bound showed a rmsd of 0.519 and 0.480, respectively (Figure 4.2A and B). The superimposition of the EfNNAT apo and the adenine-bound showed a rmsd of 0.346. A comparison of these rmsd showed that more structural deviation was observed upon NMN binding. These deviations are observed in the flexible regions 102-107 and 155-165, as previously described in the apo structures, as well as regions 136-139 and 171-176 (Figure 4.2C), implying that the binding of NMN did not stabilise this flexible region but may have induced some conformation changes.

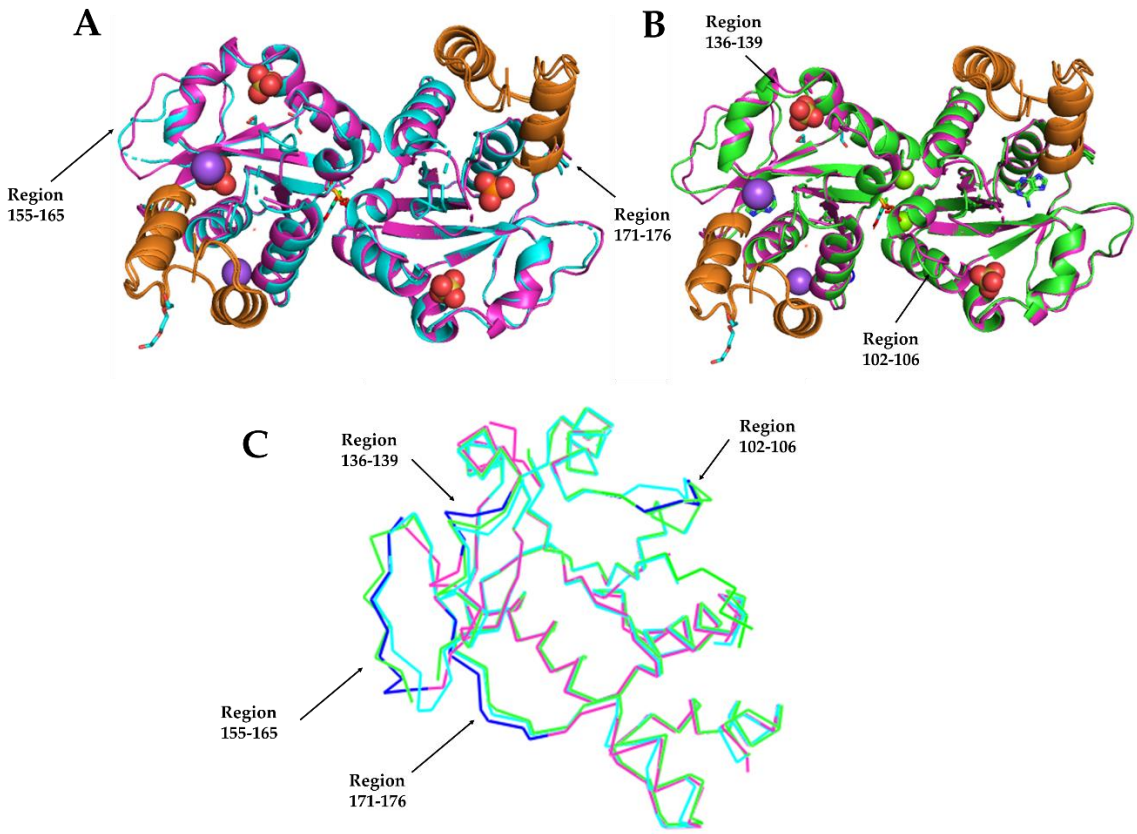



**Figure 4.2 Structural alignments of EfNNAT apo, adenine-bound and NMN-bound**
(A) Structural alignment of the apo (cyan) to NMN-bound (magenta) with rmsd 0.519 (B) Structural alignment of EfNNAT adenine-bound (green) to NMN-bound (magenta) with rmsd 0.480. The red and orange spheres represent the ions occupying the putative binding site, the purple sphere represents the $Mg2^{+}$ and $Na^{+}$, the green and blue sticks depict bound adenine, and the rainbow sticks represent NMN. (C) Alignment of the monomers of all three structures represented in ribbons. Apo (cyan), adenine-bound (green), and NMN-bound (magenta). The blue portion of the ribbon indicates regions of structural flexibility.

### 4.3.5 Ligand interaction analysis

The NMN molecule is bound between the dimer interface stabilised by five hydrogen bond contacts and 66 non-bonded contacts between chains A and B (Figure 4.1). The hydrogen bond contacts were formed between the oxygen atoms O18, O19 and O21 of NMN and residues Glu96, Tyr124 and Arg56 of chains A, respectively. Also, Asn119 of chain B formed hydrogen bond contacts with the O11 of NMN ribose sugar (Figure 4.3B). A hydrogen bond contact was also observed between NMN's

phosphate group oxygen (O7) and one water molecule. In total, four water molecules were seen networking between the sidechains and NMN.

Bound to chain A are three PEG molecules identified as Peg293, Peg297 and Peg298. The Peg293 molecule is stabilised by two hydrogen bond contacts with residues Cys190 and Ile192; Peg 297 is stabilised by six hydrogen bond contacts with Asp123, Val149, Asn150 and two water molecules, while Peg298 is stabilised by one hydrogen bond contact with Phe151. The two $Na^+$ interacted through non-bonded contacts with residues Glu46, Arg50 and Arg183. One of the $Na^+$ is positioned close to the ATP binding site. However, the sites occupied by NMN and the PEGs differ from those of the putative active site.

### 4.3.6 Dimerisation interface analysis

The total surface area of the EfNNAT-NMN dimerisation interface sums to 1769 $Å^2$ with chain A composed of 888 $Å^2$ and chain B 881 $Å^2$. Analysis of the interface showed that the monomers were stabilised majorly through salt bridges, hydrogen bonds and non-bonded contacts. Eight hydrogen bond contacts were formed between reciprocal residues Gln117-Arg56, Arg102-Lys77, Asn118-Gln94 and Lys25-Lys23. Additionally, two salt bridges were formed by Asp121-Lys25 and Glu84-Arg102 of chains A and B. In total, 96 interactions were recorded at the EfNNAT-NMN dimer interface. Details of the residue interactions and interface statistics are provided in Figure 4.4 and Table 4.3.

A comparison of the EfNNAT-NMN interface with the apo and adenine-bound symmetry generated by the dimer interface showed similar residue interaction across all three structures. The residues interact through the anti-parallel association of the two monomers, suggesting a pseudo-2-fold symmetry operation, thus resembling the dimerisations reported in Gram-positive bacterial NNAT [117, 119] (Figure 4.1A). The two hydrogen bonds and two salt bridges stabilising the apo structure exist in the NMN-bound dimer interface (Figure 4.4). Similarly, the two salt bridges stabilising

the adenine-bound structure are replicated in the NMN-bound structure. The highest interface residues, as well as a number of interactions, are observed with the NMN-bound structure, suggesting that the presence of NMN stabilises the monomers and possibly contributes to the dimerisation of the monomers, hence the larger buried surface area compared to the other structures.

**Table 4.3. Dimerisation interface analysis of EfNNAT apo, adenine-bound and NMN-bound**

| | Apo | Adenine-bound | NMN-bound |
|---|---|---|---|
| **No. of interface residues (Chain A+B)** | 26 | 26 | 38 |
| **Interface area (Å²) (Chain A+B)** | 1566 | 1538 | 1769 |
| **No. of hydrogen bonds** | 2 | 4 | 8 |
| **No. of salt bridges** | 2 | 2 | 2 |
| **No. of disulphide bonds** | 0 | 0 | 0 |
| **No. of non-bonded contacts** | 67 | 51 | 86 |
| **Total no. of interactions** | 71 | 57 | 96 |

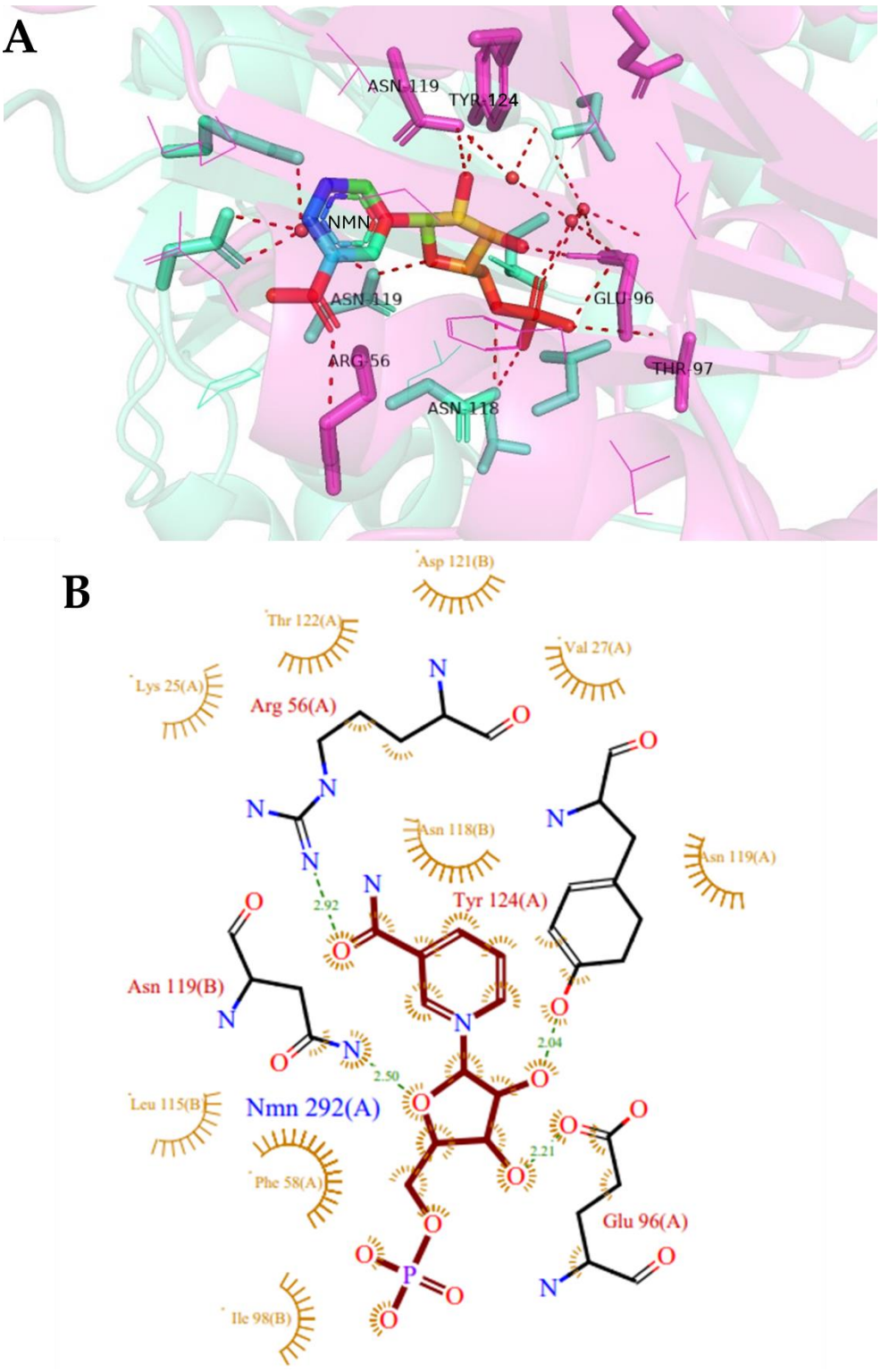



**Figure 4.3. The NMN interaction plot shows its stabilisation within the dimer interface**
(A)Cartoon representation showing NMN in the dimer interface facilitating residue interaction between both monomers represented in magenta and mint green. NMN is depicted as a rainbow stick (B) 2D interaction plot of NMN. Hydrogen bond contacts and their distance are indicated in green dashes and numbers. Non-bonded contacts are shown in gold. NMN is represented in magenta sticks, and the side chains involved in hydrogen bond contacts are in black sticks. The images were generated through PyMOL and PDBSum [168].

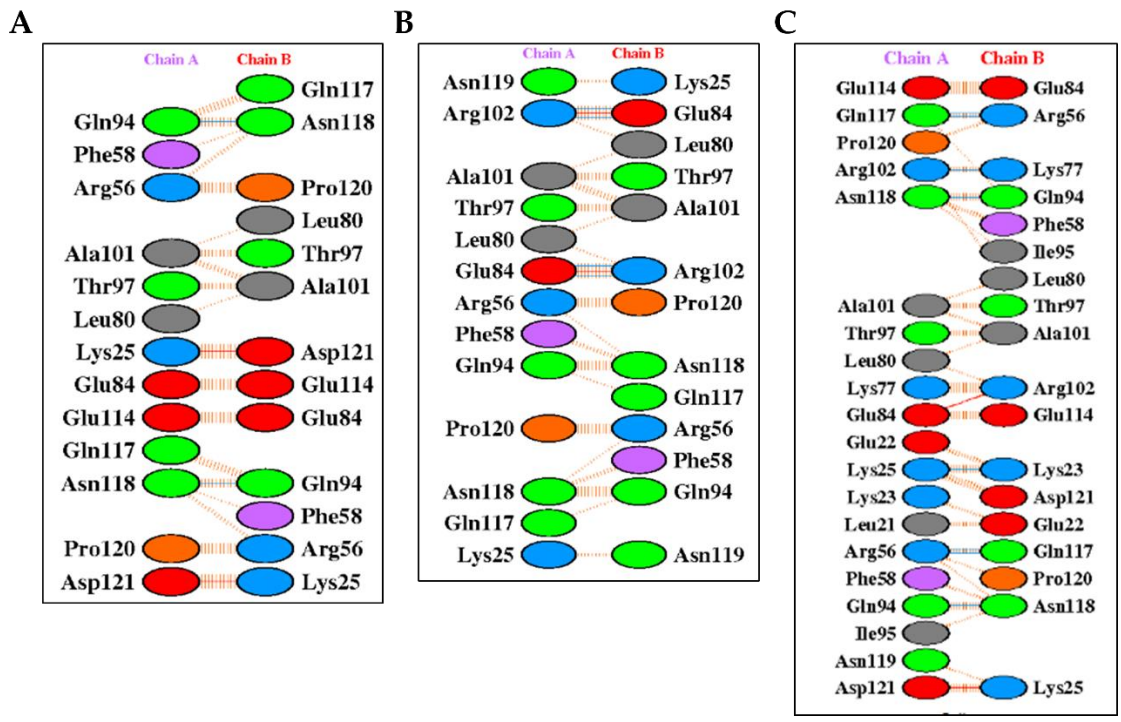



**Figure 4.4 Comparative analysis of the dimerisation interfaces of EfNNAT.**
(A) EfNNAT apo (B) EfNNAT-adenine and (C) EfNNAT NMN-bound. The respective lines represent; salt bridges, hydrogen bonds, and non-bonded contacts. The amount of red/blue lines between two residues indicates the number of potential salt bridges/hydrogen bonds that could form between them, while for non-bonded contacts, the thickness of the striped line correlates with the number of atomic contacts made. Blue, red, green, grey, and purple Residues depict positive, negative, neutral, aliphatic and aromatic residues, respectively. Proline/glycine and cysteine are shown in gold and yellow, respectively. The image was generated through PDBSum[168].

## 4.4 Discussion

To understand the binding and interaction mechanism of EfNNAT substrates and products, co-crystallisation of EfNNAT with ATP, NMN, CTP and $NAD^+$ was attempted. The co-crystallisation of EfNNAT-$NAD^+$ diffracted at a resolution between 3.18 Å and 3.5 Å, and because of the poor dataset obtained, further processing could not be achieved. Given the difficulty of co-crystallising ATP with the protein, CTP was used as an alternative. This study reported the X-ray diffraction data of EfNNAT-CTP crystal and the X-ray structure of EfNNAT in complex with NMN.

The attempt to solve the structure of EfNNAT-CTP was conducted using the molecular replacement (MR) method facilitated by Mr BUMP [101] within the CCP4i suite. Given that the EfNNAT apo structure had not been deposited in the PDB at the time of data collection, the MR BUMP structure template was used as the search model for the EfNNAT-CTP structure. Molecular replacement is a technique used to solve the phase problem in X-ray crystallography by using a known homologous structure as a starting model [108]. The process involves aligning the search model to the crystallographic data of the unknown structure to find the best fit. In this case, Mr BUMP was employed to identify and use a suitable template for EfNNAT-CTP. Despite this approach, subsequent refinement and model-building efforts in PHENIX [104] and COOT [107] were unsuccessful. The refinement did not yield significant reductions in the R-factors, with R-work and R-free values remaining at 31 and 35, respectively, even after multiple interactive rounds.

The R-factor (or R-value) is a critical metric in crystallography that measures the discrepancy between the observed structure factor amplitudes (*Fo*) and those calculated from the model (*Fc*). A lower R-factor indicates a better fit of the model to the experimental data [110]. Typically, well-refined structures exhibit R-factors below 20–25%, depending on the resolution and quality of the data. Additionally, the free R-factor (R-free), which is calculated using a subset of the data not included in the refinement process, should decrease alongside R-work if the refinement is proceeding correctly [169].

In the case of EfNNAT-CTP, the persistent high R-factors suggested potential issues. Despite several cycles of refinement and model building, the values did not improve significantly. This stagnation indicates that the model might contain significant errors or issues with the data itself. Possible reasons for this could include data quality issues [170]. Data collection, scaling, or reduction errors can lead to inaccurate diffraction data, hindering successful refinement. For instance, poor crystal quality, radiation damage, or issues during data processing can all contribute to high R-factors.

Another possible reason could be model inaccuracy. If the search model used in molecular replacement is not sufficiently homologous to the target structure, it can lead to incorrect initial phases and poor model fitting. The MR BUMP template may not have provided an accurate starting point for EfNNAT-CTP. Furthermore, statistical errors during data scaling or reduction, such as incorrect space group assignment or improper handling of anisotropy, can affect the quality of the data and the refinement process [169, 170].

The unsuccessful attempt to refine the EfNNAT-CTP structure underscores the importance of high-quality data and accurate initial models in crystallographic studies. The persistent high R-factors suggest that further investigation into the data quality and model accuracy is necessary. Potential steps to address these issues could include collecting new diffraction data, reassessing the space group assignment, and exploring alternative molecular-replace search models.

Contrary to previous reports that have shown that all NNAT ligands, including substrates and products (ATP, NaMN, NMN, NaAD and $NAD^+$), bind to the active site [66, 117, 166], here a novel structure of EfNNAT with NMN-bound to the dimer interface is presented. Although we anticipated NMN binding at the NaMN active site, analysis of the residues interacting with NMN confirms otherwise. Moreover, the superimposition of the NMN-bound structure with the EfNNAT apo and adenine showed that the putative NaMN binding site is empty and void of any ligand. Though our activity assay showed that EfNNAT has a low preference for NMN [123], we speculate that NMN would still have some affinity for the EfNNAT binding site. However, considering that only one NMN molecule was found bound to the protein at the dimer interface and no other molecule, including PEG or ions, is seen bound to the active sites, brings about the reasoning if the binding of NMN probably induces some form of conformational changes which could impact the active site.

Analysis of the three superimposed structures showed some structural flexibility at certain loop regions containing the ATP and NaMN binding motif and near the binding site. Previous reports have shown that substrate or product binding

induces conformational changes to these flexible regions to accommodate substrate binding or product release [66, 121, 144, 145]. However, given that NMN is not bound to the active site, the deviations observed at these regions could induce conformational changes to the active site. Perhaps this suggests that the site occupied by NMN may be a probable allosteric or regulatory site that could be used in drug discovery. However, further studies using computational tools such as high throughput virtual screening of small molecules that could bind to the site, molecular docking of ligands along with the substrate and molecular dynamic simulations would be required to validate this site.

In addition to the probable conformational flexibility induced by NMN binding to the protein, its position at the dimer interface influences certain changes in the local environment of the interface. More hydrogen bond contacts and networking through water molecules were observed between the residues of one monomer and the other (Figure 4.3 A), thus increasing the overall interactions and stability of the monomers and, consequently, promoting dimerisation. Furthermore, given that more residues are involved in stabilising the monomers, an increase in the proportion of buried surface area is expected. Though EfNNAT is predicted as a dimer and its large buried surface area suggests that the dimer might be biologically relevant, the presence of NMN possibly facilitating the dimerisation cannot be ruled out.

In conclusion, the X-ray structure of EfNNAT bound to NMN has been successfully determined and analysed. The structural analysis reveals the potential presence of an allosteric site within the EfNNAT protein dimer. Identifying such a site could significantly expand the scope of drug discovery efforts targeting EfNNAT and other bacterial NNAT enzymes.

# Chapter 5

# Exploring Flavonoids as Potential Inhibitors for *Enterococcus faecium* Nicotinate Nucleotide Adenylyltransferase: An Integration of *In Silico* and Empirical Studies

Olamide Jeje and Ikechukwu Achilonu*

Protein Structure-Function Research Unit, School of Molecular and Cell Biology, Faculty of Science, University of the Witwatersrand, Johannesburg 2050, South Africa.

Olamide.Jeje1@students.wits.ac.za

*Correspondence: Ikechukwu.Achilonu@wits.ac.za; Tel.: +27-11-717-6349

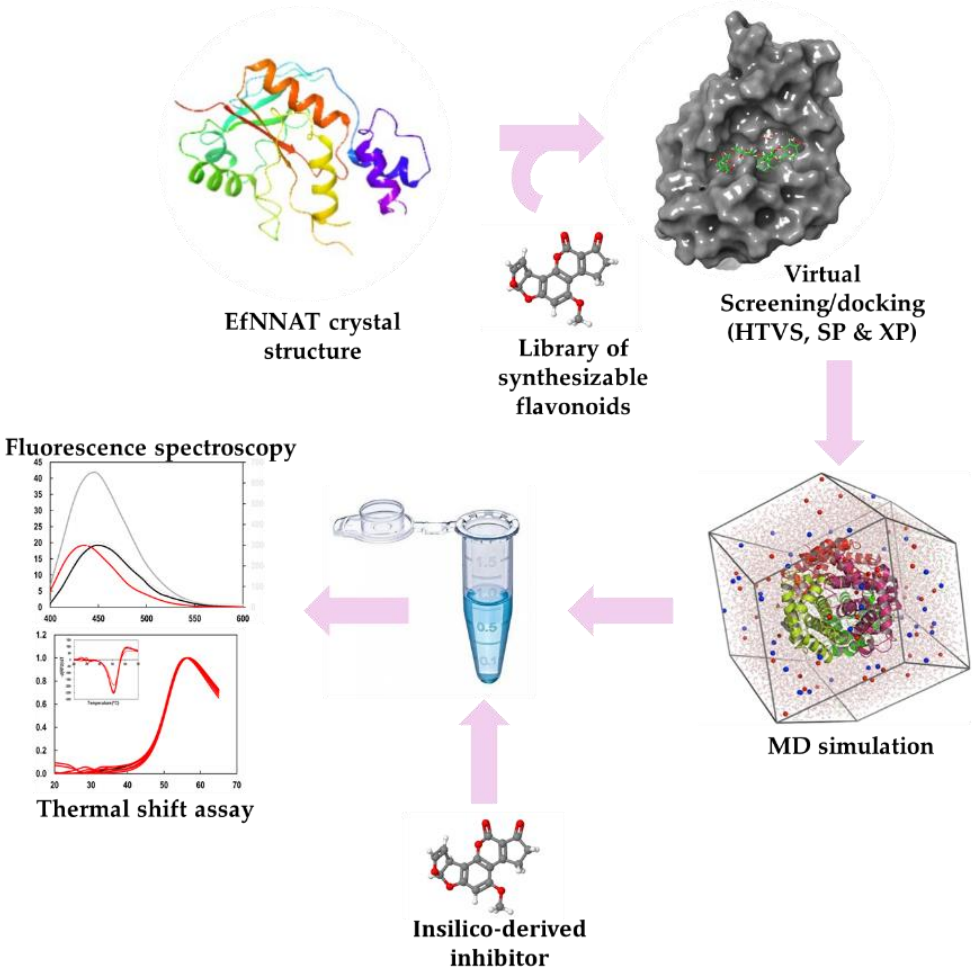
## 5.1 Abstract

Nicotinate nucleotide adenylyltransferase (NNAT) emerges as a promising target for drug development, given its pivotal role in the nicotinate nucleotide biosynthesis pathway crucial for bacterial survival. This study used computational modelling and fluorescence methods to screen a library of 472 synthesisable flavonoid compounds from MedChemExpress (MCE) as potential inhibitors against *Enterococcus faecium* NNAT (EfNNAT). We employed Maestro's Glide HTVS, SP, and XP docking protocols to identify compounds with high binding affinity. The screening utilised two models of the EfNNAT crystal structure: energy-minimised and 100 ns MD simulated model, hypothesising structural and conformational changes influencing ligand selection. Our findings revealed variations between both models, impacting the top five ligand selections post-HTVS screening. However, quercetin 3-O-beta-D-glucose-7-O-beta-D-gentiobioside emerged as the top inhibitor selected by both models after more rigorous sampling with SP and XP. A 500 ns atomistic MD simulation of the hit complex revealed that the compound interacts mainly through hydrogen bonding and water bridges. Validation of the *in-silico* studies via ANS and mant-ATP fluorescence spectroscopy confirmed quercetin's binding to EfNNAT by displacing the fluorescent probes from the protein binding site. Furthermore, thermal shift assay demonstrated minimal impact on protein stability and structure upon quercetin binding. Insights gleaned from this study offer a mechanistic basis that could be leveraged to develop new leads. Expansion of the compound library holds promise for identifying inhibitors with enhanced affinity for EfNNAT.

**Keywords:** Nicotinate nucleotide adenylyltransferase (NNAT), nicotinamide mononucleotide adenylyltransferase (NMNAT), *Enterococcus faecium*, ESKAPE pathogens, potential inhibitors, synthesisable flavonoids, high throughput virtual

screening, computational modelling, fluorescence spectroscopy, and thermal shift assay.

## 5.2 Graphical Abstract



## 5.3 Introduction

The global rise in antibiotic resistance and the emergence of new resistance mechanisms pose a critical and escalating threat to global public health, as the effectiveness of existing antibiotics and other antimicrobial agents is diminishing rapidly [1]. This alarming trend not only jeopardises the treatment of common infections but also undermines medical advancements and heightens the risk of untreatable diseases, leading to increased morbidity, mortality, and healthcare costs [2]. Estimates have shown that if actions are not taken regarding the increasing drug resistance arising especially from an identified group of multidrug-resistant pathogens called ESKAPE (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter species*), by 2050, antibiotic resistance could lead to a future where common infections become untreatable, thus highlighting the urgent need for action [4-6].

*Enterococcus faecium* is a bacterium that has emerged as a significant healthcare-associated pathogen, causing a range of challenging infections [12, 13]. One of the primary concerns with *E. faecium* is its increasing resistance to multiple antibiotics, including vancomycin [12, 17], which is often considered a last resort for treating serious infections. This rise in antibiotic resistance has led to limited treatment options and increased mortality rates associated with *E. faecium* infections. Furthermore, *E. faecium* is adept at acquiring and disseminating antibiotic-resistance genes, making it a formidable opponent in healthcare [11, 14-16]. The extensive antibiotic resistance, versatility, and the implications *E. faecium* poses for healthcare, patient outcomes, and public health underscore the critical need for comprehensive strategies to combat its spread and impact. The urgency of tackling *E. faecium* and other ESKAPE pathogens highlights the importance of research into new antibiotics and alternative treatments.

The biosynthesis of $NAD^+$ is a vital pathway for *E. faecium*, as it supplies the cell with a key cofactor required for DNA repair, maintenance of cellular redox balance and the execution of numerous enzymatic reactions crucial for the energy metabolism

of the bacterium [24-26]. NNAT plays a crucial role in the biosynthesis of $NAD^+$. Given its significance in bacterial survival and virulence, NNAT has emerged as a promising target for developing novel antimicrobial agents [20-23]. NNAT explicitly uses ATP to convert NaMN into NaAD, a precursor in the synthesis of $NAD^+$. It also facilitates the direct formation of $NAD^+$ from NMN but at a much slower rate [27]. Given the indispensable role of NNAT in the biosynthesis of $NAD^+$, inhibiting NNAT can lead to a deficiency in $NAD^+$ production, thus disrupting vital cellular processes and ultimately compromising bacterial survival [28, 29].

For several reasons, exploring synthesisable flavonoids as potential agents holds significant promise in drug discovery. Flavonoids, naturally occurring compounds in plants, exhibit various pharmacological properties, such as antibacterial, anti-inflammatory, and antioxidant effects [171-174]. They have shown potential as inhibitors in various biological pathways [175-178], making them attractive candidates for drug development. Their ability to interact with specific molecular targets, such as enzymes or receptors [179-181], provides a basis for designing compounds that can modulate key physiological processes. The structural versatility of flavonoids enables the tailoring of molecules to optimise their pharmacokinetic and pharmacodynamic profiles. Moreover, the unrestricted synthesizability of flavonoids offers a distinct advantage as it allows for the evaluation and validation of potential inhibitors empirically. Thus, targeting NNAT with synthesisable flavonoids facilitates the creation of diverse chemical libraries for systematic screening and accelerated lead compound identification and makes them empirically accessible for validation and characterisation.

Using computational tools in vaccine and drug design has revolutionised and significantly enhanced the ability to develop new therapeutics rapidly and efficiently [182, 183]. The ability of *in silico* techniques such as molecular docking and molecular dynamics simulations to provide a high degree of precision and accuracy in predicting molecular interactions at an atomic level contributes to identifying, optimising, and validating potential drug candidates and vaccines [183-185]. Furthermore,

computational techniques can predict potential off-target interactions and toxicities, which are critical for ensuring the safety of new therapeutics. While *in silico* techniques offer numerous advantages, they are most powerful when integrated with experimental methods. The validation of computational predictions through laboratory experiments and clinical trials ensures their accuracy and efficacy [186, 187]. This complementary relationship between *in silico* and *in vitro/in vivo* approaches creates a comprehensive vaccine and drug design framework, where computational insights guide experimental research and experimental data refine computational models.

In this work, we leverage computational techniques to search for potential inhibitors of EfNNAT from the MCE database of synthesisable flavonoids. High-throughput virtual screening (HTVS), followed by standard and extended precision (SP and XP) dockings, were performed using two EfNNAT model structures (an energy-minimised and a 100 ns simulated structure) to identify the most attractive compound. Only one hit resulted from this procedure. It was then subjected to a 500 ns time scale atomistic MD simulation and the topmost hit from the HTVS screening to analyse their interaction mechanism with EfNNAT. We further integrated fluorescence spectroscopy and thermal shift assay to experimentally validate the identified potential inhibitors' binding properties and impact on the EfNNAT structure. Figure 5.1 summarises the workflow employed in this study. This is the first investigation into the possible use of synthesisable flavonoids as potential agents against bacterial NNAT enzymes. This study provides the basis for further optimisation that could contribute to designing potential antibacterial drugs.

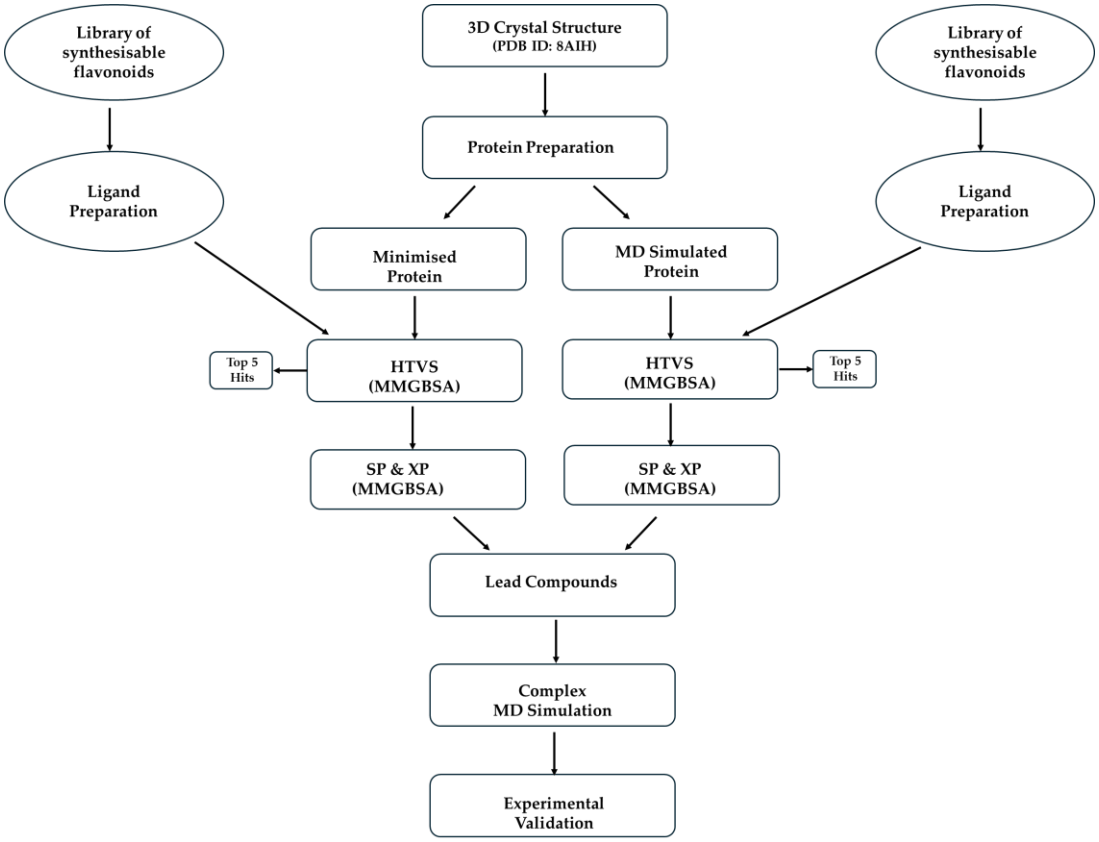



**Figure 5.1. Outline of the workflow employed to identify potential inhibitors of *E. faecium* NNAT**

## 5.4 Materials and Methods

### 5.4.1 Materials

Ligand and protein preparations, cavity analysis, receptor grid generation, virtual screening and docking studies were carried out with Maestro (Maestro 13.0, Schrödinger2021-4). The Molecular dynamic simulations and the post-dynamic data analysis were conducted using Desmond (Desmond-v7.4, Schrodinger2023-2) (Council for Scientific and Industrial Research, Cape Town, South Africa). Experimental validation was performed using a spectrofluorometer (Jasco FP- 6300) and a Real-Time PCR Detector (Bio-Rad CFX96 Touch). Quercetin-3-O-beta-D-glucose-7-O-beta-D-gentiobioside) and (apigenin 7-O-(2G-rhamnosyl) gentiobioside were purchased from MedChemExpress (MCE) through the Biotechnology Hub Africa (Pretoria, South Africa). SYPRO Orange dye, mant-ATP, 8-anilino-1-

naphthalene sulfonic acid (ANS), NMN, ATP, and $NAD^+$ were obtained from Sigma Merck (Johannesburg, South Africa). Other biochemicals and reagents utilised were of analytical grade.

### 5.4.2 Ligand library generation and preparation

We use a library of flavonoid compounds from the MCE database (Cat. No.: HY-L068), which can be obtained on their website (www.medchemexpress.com). The library had a collection of 472 flavonoids. The three-dimensional structures of these flavonoids were retrieved from PubChem and saved as a structure data file (SDF). Also added to the library are the EfNNAT natural substrates NMN (CID 14180) and NaMN (CID 121991) obtained from PubChem (https://pubchem.ncbi.nlm.nih.gov/)[188]. The 474 compounds retrieved were prepared in the LigPrep module of Maestro [189] by desalting and generating probable tautomeric forms and ionisation states at pH 7.0 ± 2.0 using the Epik ionisation tool [190, 191]. The compounds' stereoisomer chemistry was kept as stipulated in the library. The ligand molecules were then subjected to the LigPrep minimisation OPLS2005 (Optimised Potentials for Liquid Simulations) force field in Maestro to generate a stable three-dimensional structure for HTVS [192, 193]. An average output of up to thirty-two stereoisomers per ligand was generated, resulting in 1427 ligand structures.

### 5.4.3 Protein preparation and refinement

The coordinates of the recently deposited X-ray crystal structure of EfNNAT resolved at 1.90 Å (PDB ID: 8AIH) [123] were fetched from the Protein Data Bank (PDB, www.rcsb.org) [116] and used for this study. The protein structure was prepared in two forms: (i) energy-minimised and (ii) MD simulated. The energy-minimised structure was prepared using the Protein Preparation Wizard and executed

in Maestro [189]. The EfNNAT x-ray structure was pre-processed by removing the crystallographic water molecules and other hetero-atoms to obtain an apo structure. Hydrogen atoms that were absent were added, bond orders were determined, zero bond order and disulfide bonds were created, missing side chains and loops were also added, and the charges and protonation states corresponding to ionisable residues were assigned at pH 7.0 ± 2.0. Then, the structure was optimised to create a hydrogen bond network. Lastly, restrained minimisation was performed on the processed protein structure using the OPLS2005 force field with a heavy-atom RMSD threshold of 0.3 Å to relax the strained bond angles and steric hindrances [192].

To prepare the simulated structure of EfNNAT, the energy-minimised structure was subjected to MD simulation using Desmond [194]. Before MD simulation, the minimised structure was solvated and ionised using the System Builder module available in Desmond. This was executed by placing the structure in an orthorhombic boundary box at a 10 Å buffer distance from the protein molecule. The box was solvated using the TIP3P (transferable intermolecular potential with three points) explicit water model, and the system was neutralised by adding an appropriate number of counterions. Furthermore, 0.15 M NaCl was added to provide an iso-osmotic salt environment. The system was energy minimised with an OPLS2005 force field followed by a 100 ns MD simulation under an NPT ensemble at a constant pressure and temperature of 1.013 bars and 300 K, respectively, with relaxation enabled by default before the simulation. The system's energy was captured with each trajectory frame at an interval of 100 ps. The trajectory clustering tool in Desmond was then used to generate the average RMSD of the simulated structure.

### 5.4.4 Cavity analysis

Cavity analysis was performed on the minimised and simulated EfNNAT models to determine possible conformational changes that could impact ligand selection. This was achieved using the SiteMap module in the Maestro suite [195]. The

default setting on the SiteMap module was applied to identify the highest-ranked potential receptor binding sites. This setting utilises at least 15 site points per reported site, allows up to 5 site-point groupings, applies a restrictive definition of hydrophobicity, and uses a standard grid and a 4 Å crop site map from the nearest site point. Sites with a site score above 0.7 were selected and further evaluated.

### 5.4.5 Virtual screening and molecular docking

The receptor grid file for the two model structures was created using the receptor grid generation panel of the Glide module in Maestro. The centroid of the prepared protein residues was selected to generate a grid box. A 0.25 partial atomic cutoff with a 1.00 Van der Waals radius scaling factor and OPLS2005 force field were employed for generating the receptor grid. The library of prepared flavonoid compounds was docked to the receptors using the virtual screening workflow of the Glide module in Maestro [196, 197]. This was followed by a more rigorous and extensive screening using the standard precision (SP) and extended precision (XP) mode of Glide Maestro [198], which utilises a more complex sampling and scoring function to penalise compounds exhibiting weak binding affinity for the protein cavity [198]. The prime molecular mechanics-generalised Born surface area (MMGBSA) module was used to post-process the ligand's docked poses for rescoring to obtain their binding energies [199]. The binding free energy (MMGBSA $\Delta G_{bind}$) was calculated according to the equation below:

$$\Delta G_{bind} = \Delta G_{complex} - \Delta G_{protein} - \Delta G_{ligand}$$

$\Delta G_{complex}$, $\Delta G_{protein}$, and $\Delta G_{Ligand}$ represent the free energies of the optimised protein-ligand complex, the free protein receptor, and the free ligand.

### 5.4.6 Molecular dynamic simulation and data analysis

The EfNNAT-ligand complex of identified compounds was subjected to MD simulation to analyse the complex's structural stability and interaction. Similar to the previously described MD simulation protocol during protein preparation, the identified EfNNAT-inhibitor complexes were assessed by simulation using the system builder module in a TIP3P explicit solvated orthorhombic box. An interval of 10 Å was maintained between the simulation box boundary and the protein surface. The system was neutralised by adding appropriate $Na^+$ and $Cl^-$ counterions, and the physiological condition was maintained by adding 0.15 M NaCl. The MD simulations were conducted for 500 ns in the NPT ensemble using the default relaxation before simulation protocol in Desmond while maintaining a 300 K temperature and 1.013 bar pressure. The OPLS2005 force field was employed to compute the system's energy, which was captured over 1000 frames at intervals of 500 ps for each trajectory.

Analysis of the MD trajectories of the complexes was executed using the Schrodinger Maestro 13.0 suite to gain insight into the binding stability of the protein and ligands. The root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), protein-ligand contacts, solvent accessible surface area (SASA), molecular surface area (MSA), and intramolecular hydrogen bonds (intra-HB) were analysed using the simulation interaction diagram module. The Simulation Event Analysis Module was used to obtain the radius of gyration (Rog) of the systems, while the Desmond trajectory clustering tool was employed to generate the 2D and 3D protein-ligand clustering.

### 5.4.7 Physiochemical properties of identified compounds

The physiochemical properties and drug-likeness of the identified compounds were determined using SwissADME (http://www.swissadme.ch/), a free online-based tool [200, 201]. This was necessary to understand the stability of the compounds within the protein receptor while also validating it as a promising drug candidate. The

canonical SMILES utilised to predict the compounds were obtained from the PubChem database (https://pubchem.ncbi.nlm.nih.gov/) [188].

### 5.4.8 8-Anilino-1-naphthalenesulfonic acid (ANS) fluorescence spectroscopy

The flavonoid compounds were purchased in powdered form, dissolved in dimethyl sulfoxide (DMSO) and kept at 4 °C. The ANS binding studies were performed in the absence and presence of the identified flavonoid compounds and the natural substrates ATP, NMN, and $NAD^+$. The experiments were analysed in 10 mM $NaH_2PO_4$, pH 7.5, (w/v) $NaN_3$, and 1 mM DTT using 1 mL reaction volume. To empirically validate the binding of the identified flavonoid compounds to EfNNAT, 50 μM of the respective compounds was incubated with 10 μM of the protein at 37 °C for 10 min, then 50 μM ANS was added to the sample and incubated for another 10 min at 37 °C. As a basis for comparison and validation of the flavonoid compounds binding to EfNNAT, 1 mM ATP, NMN, and $NAD^+$ were respectively prepared and incubated with 10 μM of EfNNAT protein for 10 min at 37 °C, followed by the addition of 50 μM ANS and incubation for another 10 min at 37 °C. The fluorescence reading was taken at 37 °C, using a 10 mm path length quartz cuvette in a Jasco FP-6300 spectrofluorometer. Samples were excited at 380 nm, and emission spectra were recorded at 400–650 nm wavelength. Other measurement parameters include 1 nm data pitch, 5 nm excitation and emission bandwidth, and 200 nm/min scanning speed. Triplicates of three emission spectra were collected per experiment, averaged and corrected for free ANS and buffer absorption. The obtained spectra were analysed with SigmaPlot.

### 5.4.9 2'/3'-O-(N-Methyl-anthraniloyl)-adenosine-5'-triphosphate (mant-ATP) fluorescence spectroscopy

Mant-ATP was used to probe the EfNNAT ATP binding site for interaction and binding of the flavonoid compounds to the site. This was analysed using 2.5 μM dialysed EfNNAT and 10 μM mant-ATP in 10 mM $NaH_2PO_4$, pH 7.5, (w/v) $NaN_3$, and 1 mM DTT. Similar to the protocol described in the ANS binding study, 2.5 μM of EfNNAT was incubated with 10 μM of the respective compound at 37 °C for 10 min, then 10 μM mant-ATP was added and incubated for another 10 min at 37 °C. The protein was also incubated with 1 mM ATP and 2 mM $MgCl_2$ under the same condition as a basis for comparison. Fluorescence readings were taken at 37 °C with samples excitation at 355 nm and emission wavelength recorded at 400–650 nm. A 2.5 nm excitation and 5 nm emission bandwidth were used for spectra collection. All other measurement parameters are as detailed in the ANS study. Experiments were performed in triplicates and corrected for free mant-ATP and buffer absorption. Data processing was done using SigmaPlot.

### 5.4.10 SYPRO Orange thermal shift assay

A thermal stability assay was used to analyse the effect of the identified flavonoid on the structural stability of EfNNAT, consequently confirming the ability of the compounds to bind and stabilise the protein. This was achieved by using SYPRO Orange dye to monitor the thermal unfolding of the protein in the absence and presence of the flavonoid compounds. The assay was performed using a 25 μL reaction volume comprising 5 μL of 50 × SYPRO Orange dye, 10 μL of 50 μM EfNNAT protein, and 10 μL of buffer (10 mM $NaH_2PO_4$, pH 7.5, 1 mM DTT, 0.01 % (w/v) $NaN_3$, and 2 mM $Mgcl_2$) supplemented with varying concentration of the flavonoid compounds. The concentrations used include 0, 10, 25, 50, 100, 250 and 500 μM. The protein was incubated with the respective compound for 10 mins at 37 °C, and then

SYPRO Orange dye was added. All experiments were performed in quadruplicate using a Bio-Rad 96-well PCR microplate. The samples were vortexed and centrifuged for about 20 sec and then run in a BioRad CFX96 Touch Real-Time PCR Detector between 20 °C and 80 °C with 0.5 °C increments. The readings were accessed using the CFX Maestro software and processed in Excel. All samples were corrected for buffer absorption.

## 5.5 Results

### 5.5.1 Energy minimised Vs MD simulated model

Before screening the flavonoid library, we postulated that there might be structural and conformational variation with the protein's active site when simulated versus the energy-minimised model, which may influence ligand selection during virtual screening. Hence, we subjected the energy-minimised EfNNAT to a 100 ns simulation equilibration in a water model containing counter ions using the OPLS2005 force field. Superimposition of the EfNNAT minimised model with the 100 ns simulated model showed structural and conformational changes (Figure 5.2). Apparent structural differences were observed with transformation in the structural element of the protein from coil to beta-strand and helix, as indicated in Figures 5.2B and C. The simulation of the structure resulted in the formation of more order elements, indicating changes in the secondary structure content as well as the tertiary structure of the protein, which could influence ligand selection.

### 5.5.2 Active site analysis of the EfNNAT modelled structures

The cavities of the EfNNAT energy-minimised, and the MD simulated models were also analysed using Maestro's SiteMap module. The sites were selected for evaluation based on their site score. The site score measures the likelihood that a given site is an actual binding site. It is calculated based on many factors, including the size

and shape of the site, the number of hydrogen bonds and other interactions that can be formed between the site and a ligand, and the site's conservation across different species. A site score above 0.7 is generally considered a good indication that the site is an actual binding site and more likely to bind ligands with high affinity and specificity [195].

The topmost site score for the energy-minimised model was 1.011, with a volume of about 517.930 Å$^3$. The MD simulated model's topmost site score was 1.031 with 195.853 Å$^3$ volume. Its site_2 had a larger volume of 253.134 Å$^3$ with a site score of 0.961. Table 5.1 summarises the reports of the cavity analysis for the two models. Evaluation of the site_1 and _2 cavities of the simulated model (195.853 and 253.134 Å$^3$) in comparison with the site_2 cavity (517.930 Å$^3$) of the EfNNAT minimised model revealed overlapping of the sites. The site_2 of the minimised model comprises the site_1 and _2 residues of the simulated model, thus suggesting that the one big site of the initial minimised structure has been remodelled into two defined sites upon 100 ns simulation of the protein (Figure 5.3) in an explicit TIP3P water model. Perhaps it is a reminder that the protein is a dual substrate-binding protein, with each substrate having its binding site.

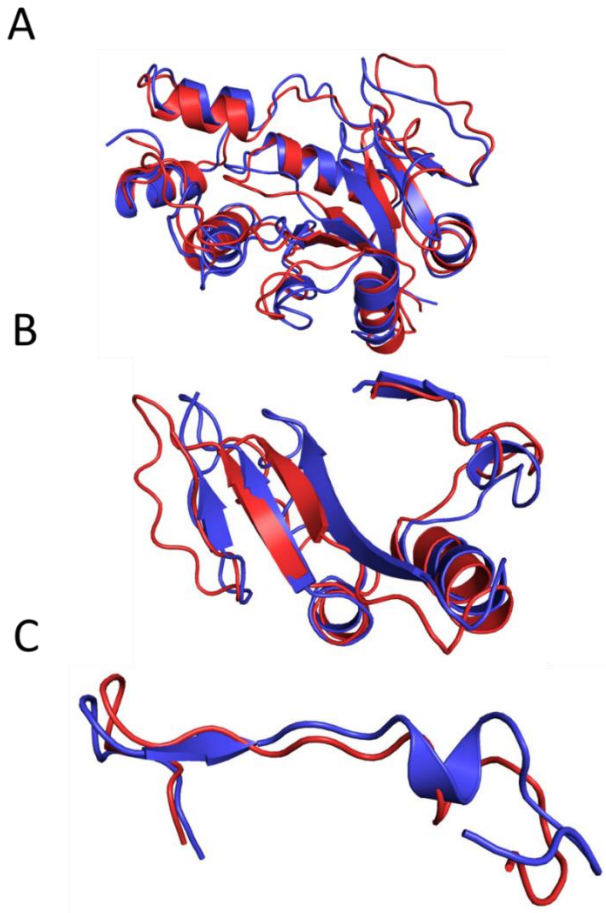



**Figure 5.2. Structural comparison of EfNNAT energy-minimised Vs MD simulated model**
(A) Superimposition of EfNNAT minimised (red) with the MD simulated model (blue) (B) Regions between the two models with structural deviation (C) Detailed image depicting the structural transformation of the minimised model (red) coil into beta-strand and helix (blue) upon 100 ns simulation.

**Table 5.1. Details of EfNNAT minimised and MD simulated model cavity analysis**

| | **Minimised** | | | **MDS** | | |
|---|---|---|---|---|---|---|
| **Site** | **Site score** | **Volume (Å3)** | **Residues** | **Site score** | **Volume (Å3)** | **Residues** |
| 1 | 0.979 | 254.506 | 33,35,63-67,71-76, 79,100,179,180,183,184,187,210-212 | 1.031* | 195.853* | 30,33,35,38,41,44,45,129132,154, 156-159,175-181,183 |
| 2 | 1.011* | 517.930* | 30-34,38,40,41,43,44,60,62,63,99, 105-109, 126-132,138-140,145, 154-156,172-183 | 0.961* | 253.134* | 29-31,60,62,63,99,102-109,111,112,115,126-129,132,135,139,140 |
| 3 | 0.625 | 67.228 | 130,133,134,137,155,156,159,160, 161,163-165,169 | 0.635 | 103.586 | 39,42,43,46,47,85-90,197-199 |
| 4 | 0.637 | 59.682 | 25,27,29,56,58,96,115,118,119,121, 122,124 | 0.532 | 63.455 | 116,120,122-125,148-150 |
| 5 | 0.544 | 58.310 | 186,189,190-193,201,204,205,211 | | | |
| **Total** | | **772.436** | | | **448.987** | |

*Indicates the sites considered for structural evaluation

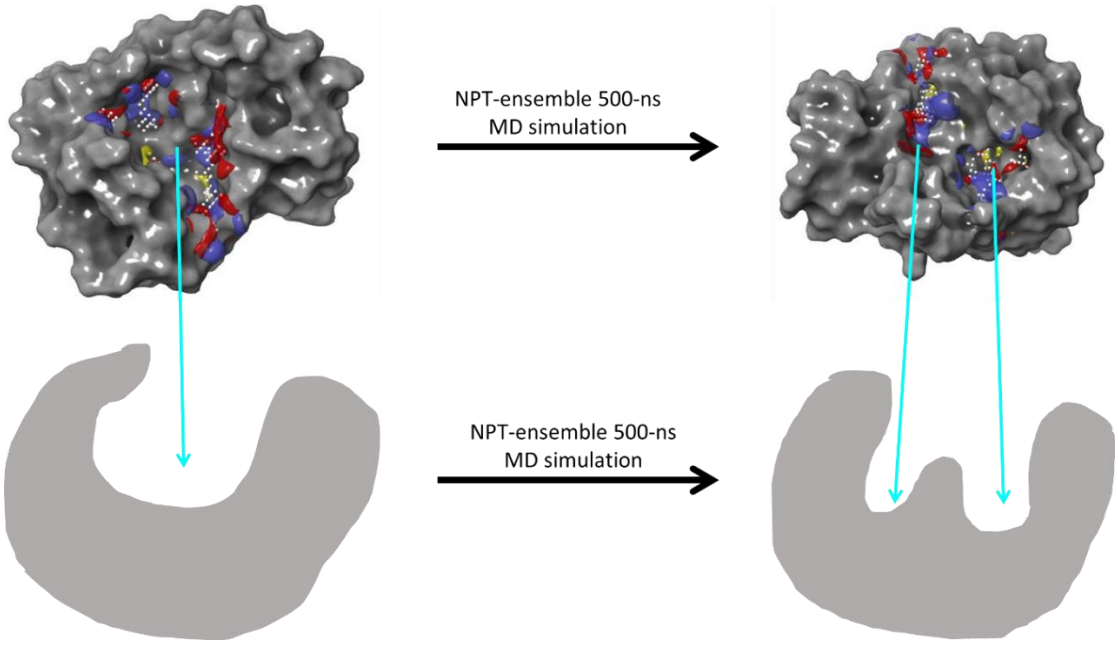



**Figure 5.3. Surface and pictorial representation of EfNNAT minimised model (left) and the MD simulated (right) cavities.**
White spheres are site points used by the algorithm for the binding site prediction; red, blue, and yellow patches represent hydrogen bond acceptors, hydrogen bond donors and hydrophobic interaction sites, respectively.

### 5.5.3 Virtual screening and compound identification

Considering that the minimised and the simulated models exhibit some conformational differences, both models were used to screen the prepared ligand library to see if the models would identify different or similar compounds. The five topmost compounds selected by the energy-minimised model after HTVS docking and MMGBSA prime were apigenin-7-O-(2G-rhamnosyl)gentiobioside, 6'''-feruloylspinosin, quercetin 3-O-beta-D-glucose-7-O-beta-D-gentiobioside, pinocembrin, and naringin dihydrochalcone. The MD simulated model identified apigenin 7-O-(2G-rhamnosyl)gentiobioside, pinocembrin, bavachalcone, quercetin 3-O-beta-D-glucose-7-O-beta-D-gentiobioside and sappanone A as its five topmost ligands. Tables 5.2 and 5.3 provide details of the five topmost ligands identified,

including their structures, molecular weight, docking scores, MMGBSA binding energies, docking poses and binding interactions.

Following the implemented HTVS protocol, a more rigorous and extensive sampling was performed using the SP and XP protocol to screen down and identify the ligand with the most binding affinity for the protein. Interestingly, only one compound was selected from the entire library by both models, with the compound identified as quercetin 3-O-beta-D-glucose-7-O-beta-D-gentiobioside showing an MMGBSA $\Delta G_{bind}$ of −59.64 and −35.97 kcal/mol for the minimised and MD simulated models respectively.

**Table 5.2. Five topmost ligands identified post-HTVS for EfNNAT minimised and MDS models**

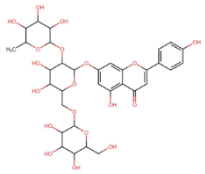
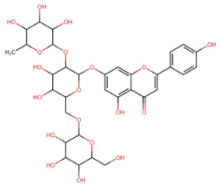
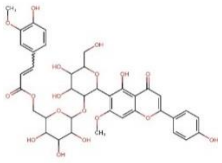
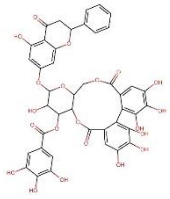
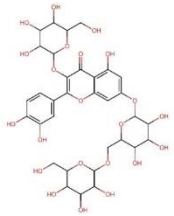
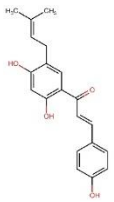
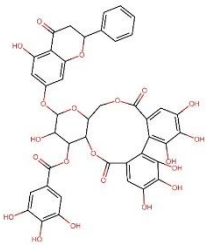
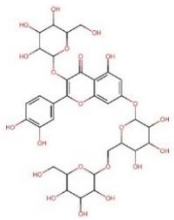
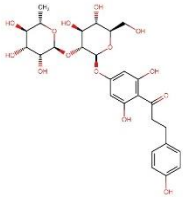
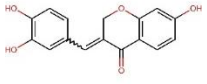
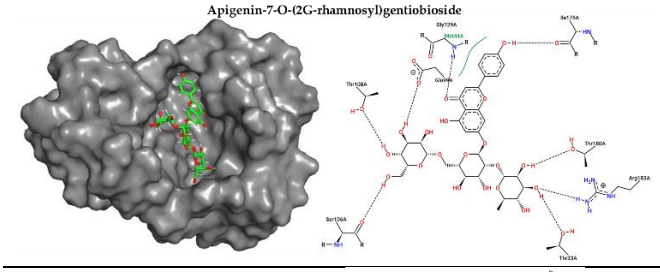
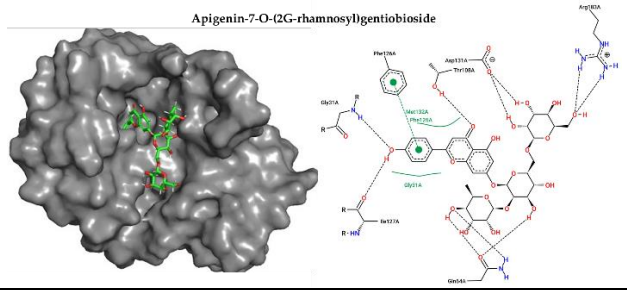
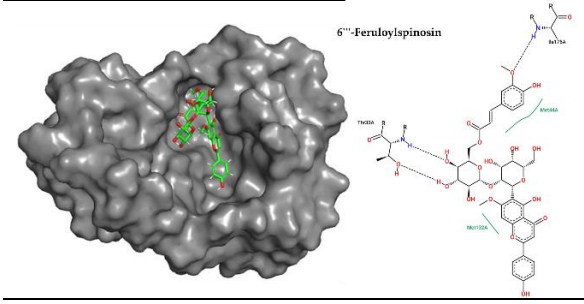
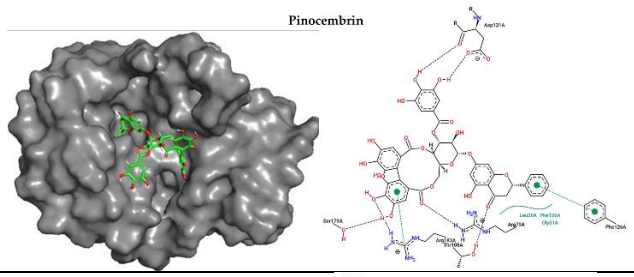
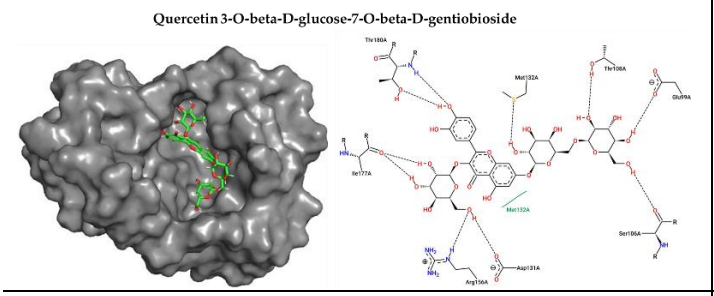
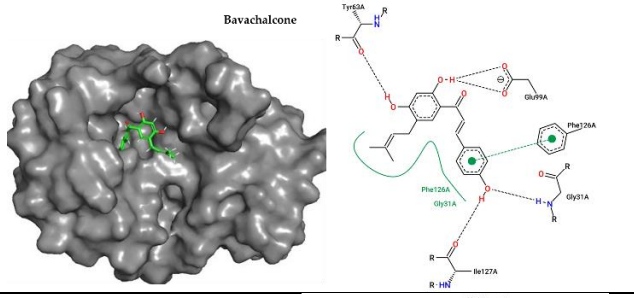
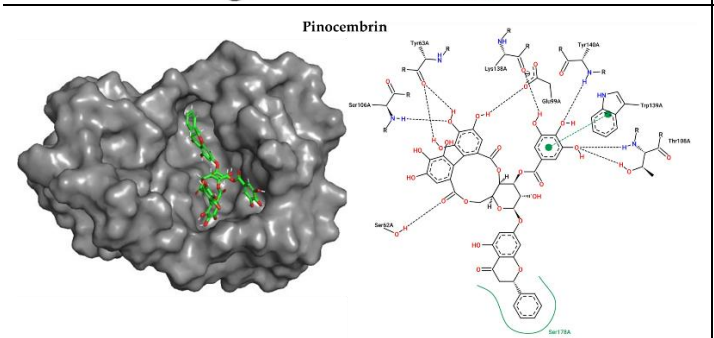
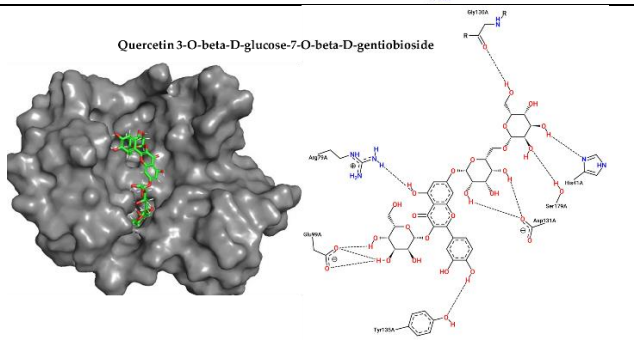
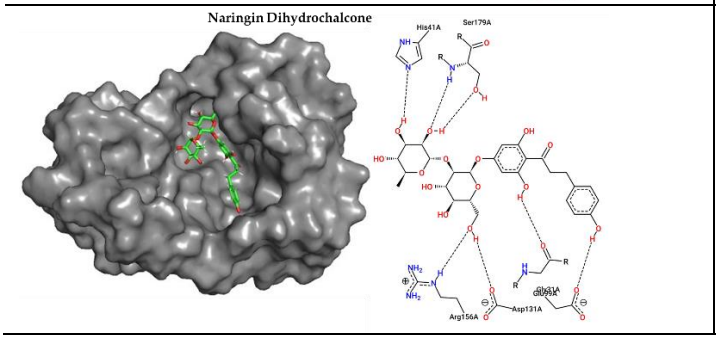
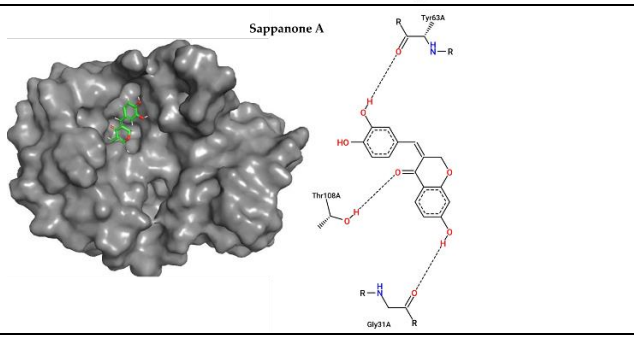
| Minimised | | | | MDS | | | |
|---|---|---|---|---|---|---|---|
| Compound Name and Structure | Mol wgt (g/mol) | Docking Score (kcal/mol) | MMGBSA $\Delta G_{bind}$ (kcal/mol) | Compound Name and Structure | Mol wgt (g/mol) | Docking Score (kcal/mol) | MMGBSA $\Delta G_{bind}$ (kcal/mol) |
|  Apigenin-7-O-(2G-rhamnosyl)gentiobioside | 740.67 | −8.321 | −74.24 |  Apigenin-7-O-(2G-rhamnosyl) gentiobioside | 740.67 | −7.312 | −60.84 |
|  6'''-Feruloylspinosin | 784.72 | −7.907 | −72.16 |  Pinocembrin | 872.70 | −7.017 | −55.55 |
|  Quercetin 3-O-beta-D-glucose-7-O-beta-D-gentiobioside | 788.67 | −7.421 | −60.36 |  Bavachalcone | 324.38 | −7.009 | −36.14 |
|  Pinocembrin | 872.70 | −7.300 | −71.17 |  Quercetin 3-O-beta-D-glucose-7-O-beta-D-gentiobioside | 788.67 | −6.973 | −39.38 |
|  Naringin Dihydrochalcone | 582.56 | −7.158 | −67.82 |  Sappanone A | 284.27 | −6.066 | −32.62 |

**Table 5.3. Comparative analysis of the five topmost ligands selected by the EfNNAT energy-minimised and MD-simulated model after HTVS docking.**

The 3D and 2D docking poses and binding interaction were generated via PyMOL and Protein Plus server [202].

| Minimised | MDS |
|---|---|
| Apigenin-7-O-(2G-rhamnosyl)gentiobioside  | Apigenin-7-O-(2G-rhamnosyl)gentiobioside  |
| 6‴-Feruloylspinosin  | Pinocembrin  |
| Quercetin 3-O-beta-D-glucose-7-O-beta-D-gentiobioside  | Bavachalcone  |
| Pinocembrin  | Quercetin 3-O-beta-D-glucose-7-O-beta-D-gentiobioside  |
| Naringin Dihydrochalcone  | Sappanone A  |

### 5.5.4 MDS analysis

The hit compound (quercetin 3-O-beta-D-glucose-7-O-beta-D-gentiobioside), along with the first topmost compound (apigenin 7-O-(2G-rhamnosyl)gentiobioside) identified by HTVS which is a common hit to both the energy-minimised and MD simulated models were subjected to MD simulation in complex to the EfNNAT. The MD simulations were performed using the EfNNAT simulated model, given that the models identified the same ligands. Insight into the stability and interaction of the compounds with the EfNNAT was gained upon 500 ns MD simulation. Due to the two compounds' long names, they will subsequently be referred to as quercetin and apigenin.

***Stability and fluctuation analysis***

The C$\alpha$-RMSD and C$\alpha$-RMSF generated via the Simulation Interaction Diagram algorithm were used to measure the structural stability and fluctuation of each EfNNAT-inhibitor complex. The C$\alpha$-RMSD assesses the dynamic movement observed in all C$\alpha$ atoms within the protein backbone, gauging the typical deviation of these atoms from their original arrangement over the trajectory. This offers an understanding of a protein's general structural steadiness throughout the simulation period. Variation in the C$\alpha$-RMSD of the protein with respect to the EfNNAT-quercetin complex and EfNNAT-apigenin complex was observed, as indicated in Figure 5.4A. The protein appeared stable without the inhibitors; however, the average C$\alpha$-RMSD (2.67 Å) was high. Compared to the apo form of the protein, the EfNNAT-quercetin complex was slightly unstable, but within 1-3 Å fluctuation, its average C$\alpha$-RMSD (2.04 Å) was lower. Similarly, the EfNNAT-apigenin complex exhibited a lower average C$\alpha$-RMSD of 2.34 Å compared to the unbound EfNNAT. Overall, the C$\alpha$-RMSD are comparable amongst the three systems, thus indicating no major deviation in the stability of the protein. However, quercetin appears to stabilise the protein backbone slightly.

Furthermore, the RMSF was used to characterise the regional changes in the C$\alpha$ of the amino acid residues by evaluating the extent of displacement of a particular atom or group of atoms in relation to its structure. The higher the RMSF fluctuation, the more flexible the residues are during complex formation. Figure 5.4B shows the RMSF results of the EfNNAT apo and the inhibitor complexes, providing insights into the regions with high fluctuations during the simulation. The amino acid residues within regions 60-73 and 98-110 showed the most fluctuation in the EfNNAT apo. The first region is identified as the loop region containing the partially conserved PPHX motif, which serves as an arm in the recognition and binding of NaMN [66, 119, 145, 166] or NaAD [144]. The second region is comprised of a loop region and the NaMN/NMN binding motif [123]. The EfNNAT-apigenin complex showed a slight decrease at region 60-73 and an increase in fluctuation at 98-110 in relation to the EfNNAT apo. The EfNNAT-quercetin complex, on the other hand, decreased the fluctuation of residues at regions 98-110 and 130-142, which contains partially conserved residues known to stabilise the NaMN pyridine ring. This suggests that quercetin has some influence on the protein stability.

***EfNNAT structural compactness***

The Rog of the protein and the inhibitor complex forms were measured in order to gain insight into the global compactness of the protein during MD simulation. The lower the Rog of a protein, the more compact it is and vice versa when the Rog is higher [203]. Figure 5.5 shows the Rog of the EfNNAT apo structure compared to the EfNNAT-quercetin and EfNNAT-apigenin complex. A few spikes in the Rog of the EfNNAT apo and the EfNNAT-apigenin complex were observed at the beginning of the MD simulation, stabilising within 100 ns of the simulation time. Also, towards the end of the simulation, a negligible shift in the compactness of the protein was observed between the EfNNAT-inhibitor complexes and the EfNNAT apo. Overall, the average Rog values of 17.126, 17.223, and 17.273 Å for the EfNNAT apo, EfNNAT-quercetin

and EfNNAT-apigenin, respectively, are comparable with no significant deviation in the global compactness of the protein.

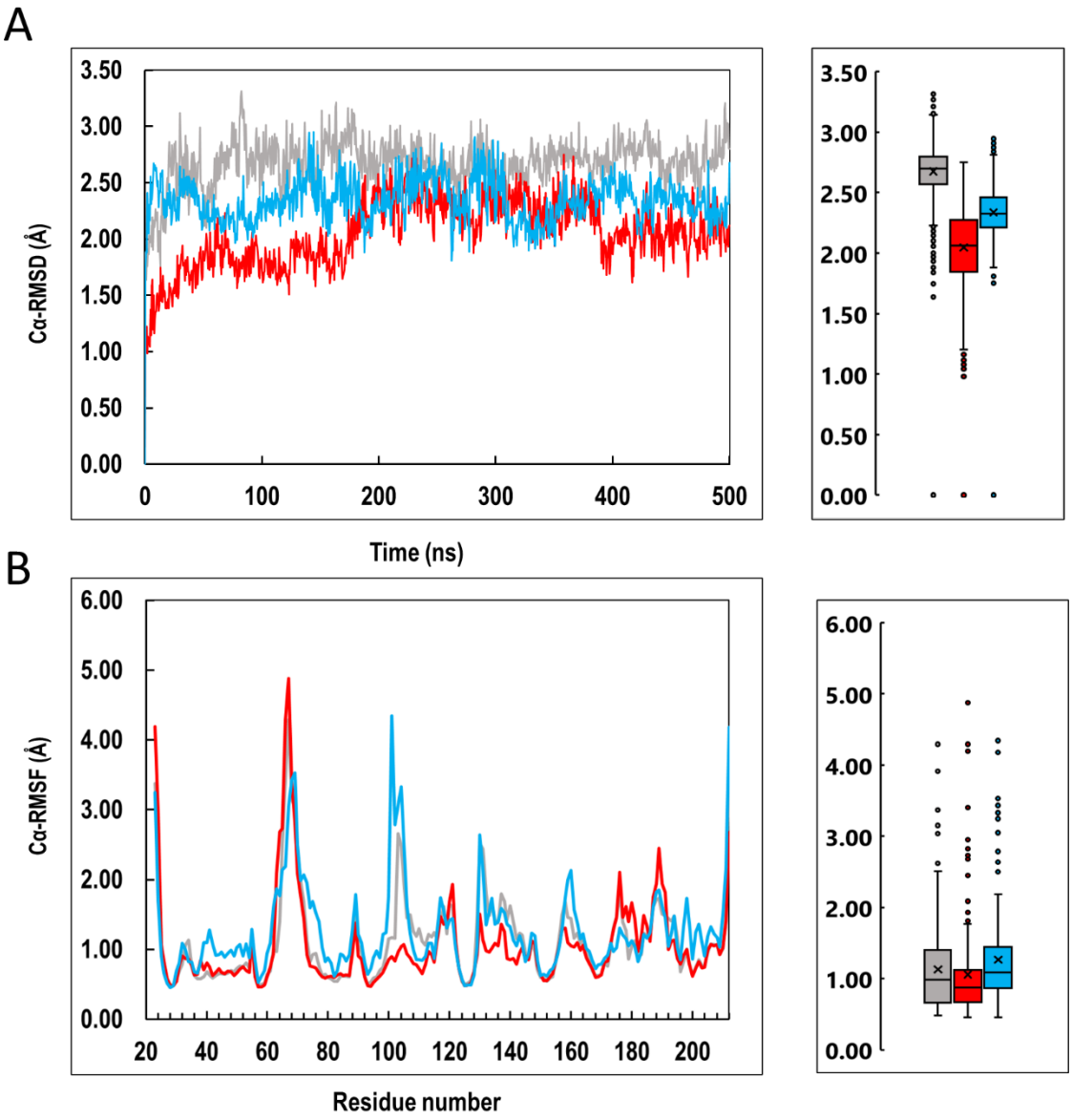



**Figure 5.4. Cα-RMSD and RMSF of EfNNAT apo and in complex with inhibitors**
(A) EfNNAT apo (grey), EfNNAT-quercetin (red) and EfNNAT-apigenin (blue) Cα-RMSD backbone measured with respect to 500 ns simulation time. The Cα atoms are comparable, but quercetin exhibited the least Cα-RMSD value. (B) Cα-RMSF indicating EfNNAT amino acid residue fluctuation in the absence and presence of quercetin and apigenin. More residue restriction is observed with quercetin than with the EfNNAT apo or the apigenin complex. To the right are the respective box and whisker plots showing the data distribution. The narrower the box, the more stable the trajectory and vice versa.

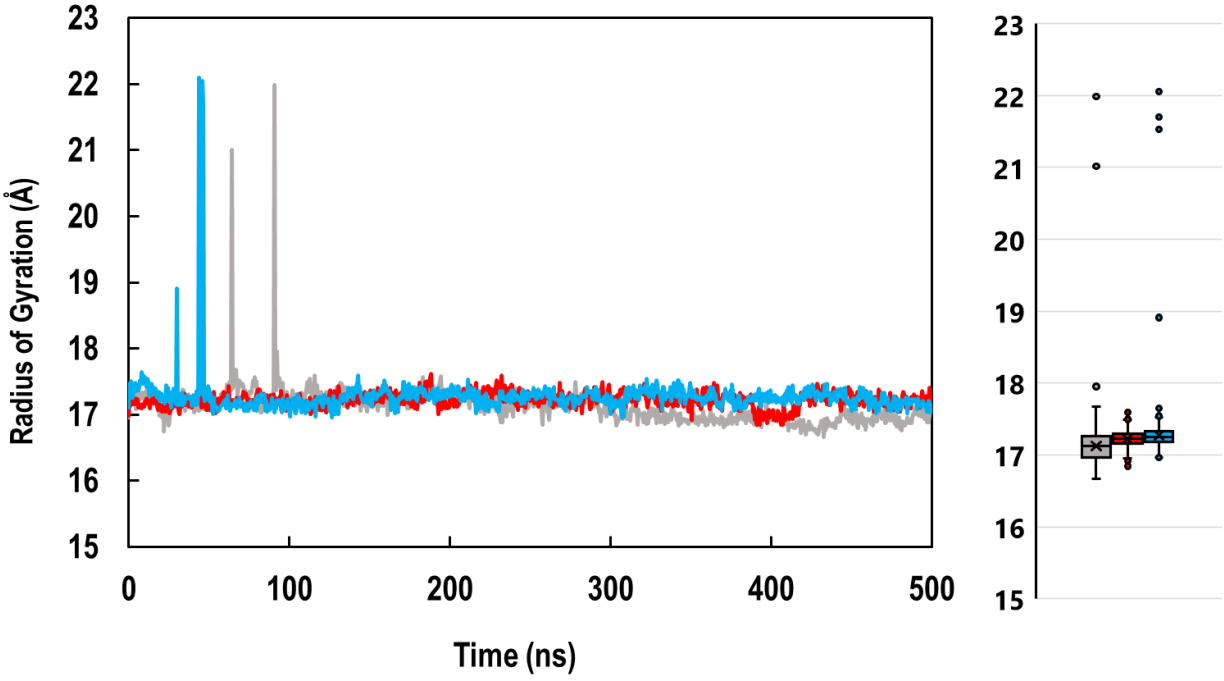



**Figure 5.5. EfNNAT radius of gyration showing the protein's compactness in its apo and inhibitor complex forms.**
The Rog of EfNNAT apo (grey), EfNNAT-quercetin (red), and EfNNAT-apigenin (blue) are comparable, as indicated by the box and whisker plot.

### *Conformational dynamics of identified compounds within EfNNAT receptor*

Insight into the conformational dynamics of quercetin and apigenin within the protein receptor was provided by analysing the ligand's RMSD, MSA, SASA, and intra-HB formation. The ligand RMSD with respect to the protein was used to measure the internal changes of the ligand atoms within the binding site. The ligand-RMSD of the EfNNAT-inhibitor complexes, as indicated in Figure 5.6A, showed that quercetin was stable in the binding site for the first 120 ns, after which an increasing RMSD fluctuation from about 2.5-6.5 Å was observed which stabilised to a large extent for the remaining of the simulation time. On the other hand, apigenin was increasingly fluctuating, with drops in the RMSD at around 160 ns and 250 ns. Then, an increasing RMSD fluctuation from 4.3-12.0 Å was observed, which dropped to about 7.5 Å while converging at 365 ns and stabilised till the end of the simulation. The increasing RMSD

fluctuations observed for both inhibitors suggest they are highly dynamic within the protein receptor.

The MSA and SASA were used to measure the area available for non-bonded interactions and water molecules to access the ligand within the receptor. This was necessary to provide information on how the surface area of the inhibitors evolves within the EfNNAT receptor. A stable trajectory indicates that the molecule maintains one conformation within the receptor. However, fluctuations are indications of conformational changes. Figures 5.6B and C show the MSA and the SASA of quercetin and apigenin within the EfNNAT receptor. Both inhibitors were unstable, as indicated by the constant fluctuation in their MSA and SASA trajectories, suggesting they are highly dynamic within the protein receptor.

The intra-HB contact of the two molecules was used to probe their dynamism further, considering that the higher the tendency of a molecule to form intra-HB contact, the more dynamic the molecule will be. Figure 5.7 shows the number of intra-HB formed by quercetin and apigenin per time throughout the 500 ns simulation. Both inhibitors formed many single, double, and triple hydrogen bonds, with quercetin forming as high as four intra-HB at certain times. Also, note worth is that throughout the 500 ns simulation time, at least one intra-HB is formed by the inhibitors, signifying how dynamic they are. Overall, quercetin formed more intra-HB contacts than apigenin.

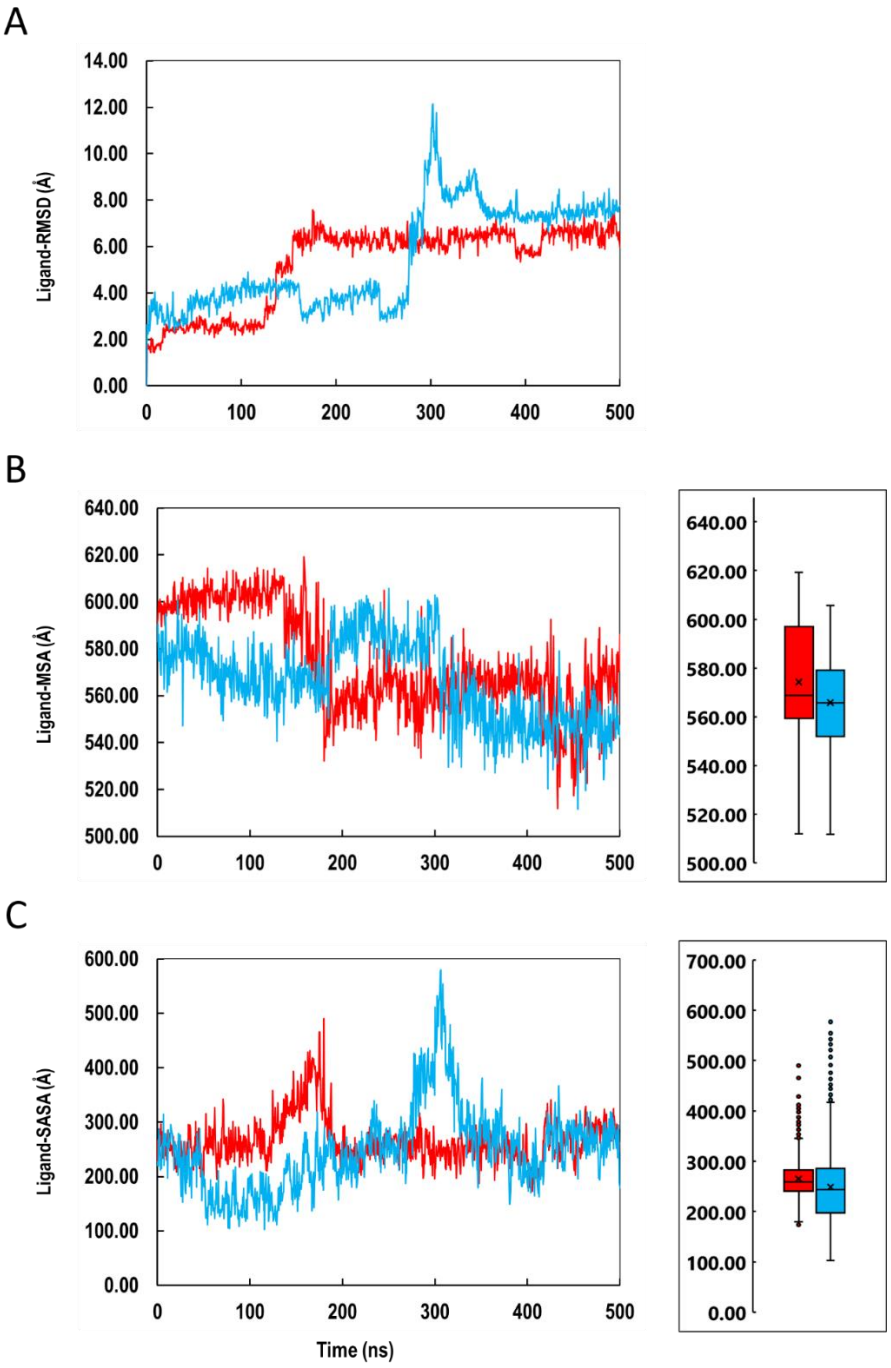



**Figure 5.6. Plots depicting quercetin (red) and apigenin (blue) dynamism within the EfNNAT receptor over 500 ns simulation time**.

(A) The ligand-RMSD (B) Molecular surface area (MSA) and (C) Solvent accessible surface area (SASA) of quercetin (red) and apigenin (blue) with respect to EfNNAT receptor. A wide box and whisker plots indicate an unstable trajectory. Overall, both ligands are unstable and highly dynamic within the protein receptor, as shown by the constant fluctuation in the trajectories.

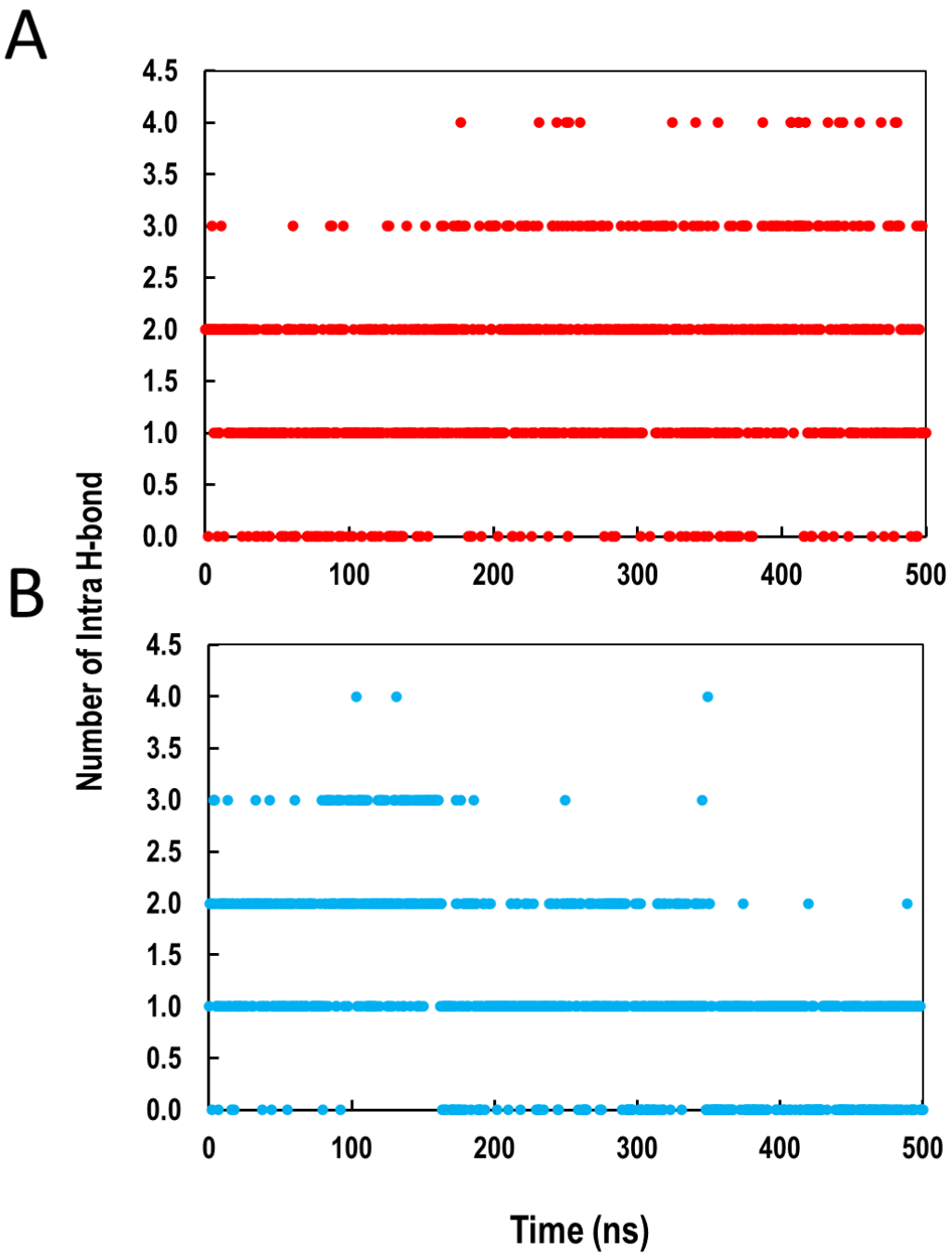



**Figure 5.7. Graphical representation of the number of intra-hydrogen bond formation between the atoms of (A) quercetin and (B) apigenin along the 500 ns trajectory time.** At least one intra-HB contact is formed per time throughout the simulation period.

***Interaction analysis of identified compound with EfNNAT***

The clustering analysis of the EfNNAT-inhibitor complexes showed that quercetin and apigenin interacted with EfNNAT through the active site by binding in such an extended conformation to span the two cavities (Figure 5.8). Quercetin formed direct hydrogen bond contact with Gln64, Glu99, and Asp131 residues while attracting more water solvent. The binding of quercetin was largely driven by water molecules, as indicated by its hydrogen bond interactions with 16 water molecules. Water bridges were formed with Gln64, Gly104, and Thr108. On the other hand, apigenin interacted with the protein, forming direct hydrogen bond contact with Ly32, Met60, Ser62, Glu99, and Ile175 and two water molecules. A water bridge with residue Arg156 was also observed.

The interaction of quercetin and apigenin with EfNNAT side chains is largely driven by hydrogen bonds and water bridges, as shown in Figure 5.9. Hydrophobic contacts and a few ionic interactions were also observed in the stabilisation of these compounds. The highly conserved residue Asp131 made the most interaction fraction of all the side chain residues for both compounds, as shown in the stacked bar charts (Figure 5.9). Detailed atomic interaction of the compounds with EfNNAT that occurs for over 30% of the 500 ns trajectory time further showed that two hydroxyl groups of quercetin interacted with the Asp131 for 55% and 36% while apigenin interacted with the residue for 35% of the simulation time (Figure 5.10). Gly103, Ser106, Thr108, and Tyr140 of the NMN binding motif also interacted with quercetin for 45%, 62%, 32%, and 47% of the simulation time, respectively, through hydrogen bonding and water bridge. The highest percentage interaction made by quercetin was with Glu99 for 74% of the simulation time. Note worth is the intra-HB contact made between a hydroxyl and carboxylic group of the compounds, which occurred for 74% of the simulation time in apigenin and two of such contacts in quercetin for about 48% and 49% of the simulation time.

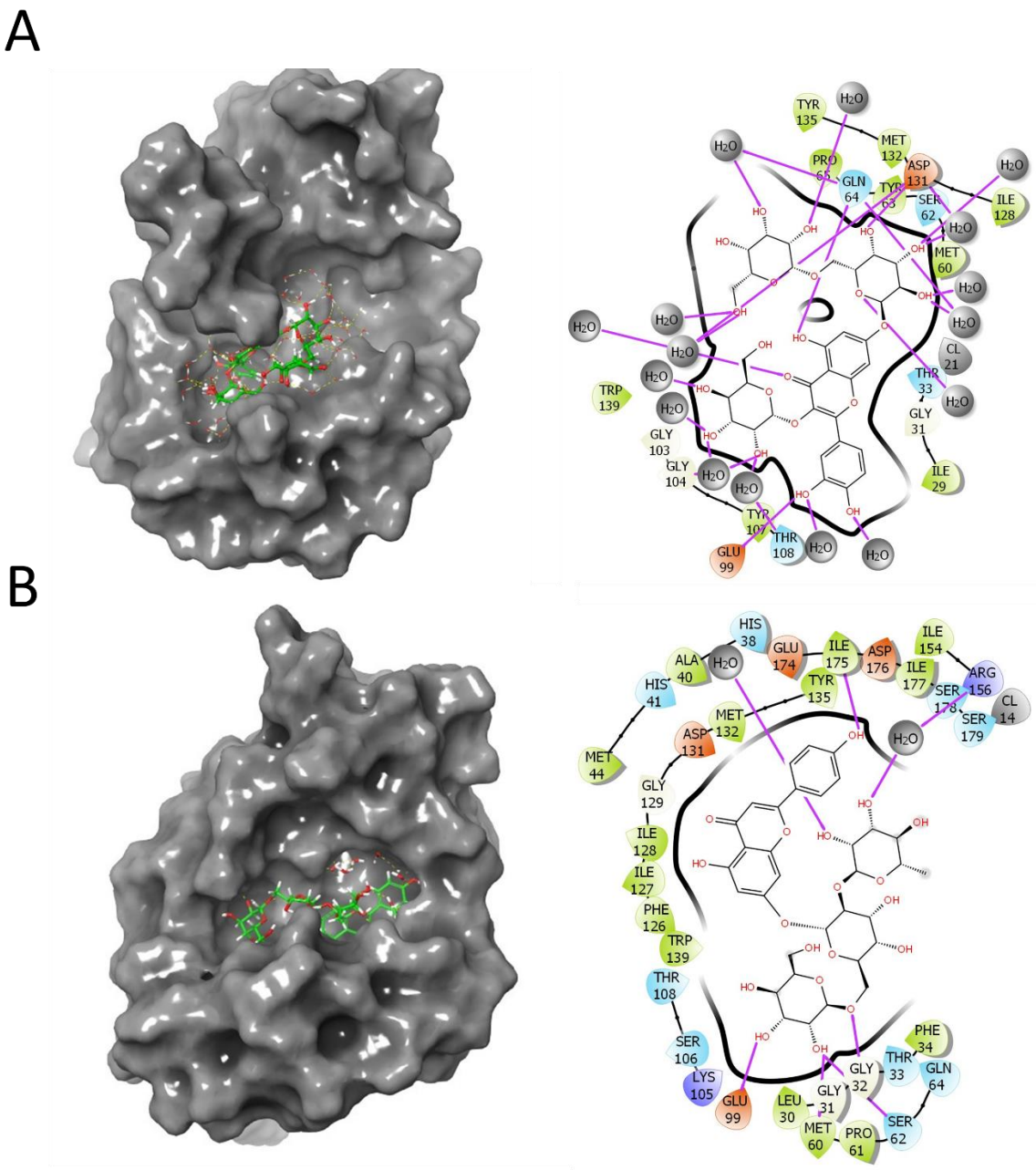



**Figure 5.8. 3D and 2D cluster analysis of EfNNAT in complex with (A) quercetin and (B) apigenin. EfNNAT in s**urface representation is complex with the respective ligands shown in green sticks. Hydrogen bond networking between the complex is shown in yellow dashed lines. Side chains in green, blue, white, violet, and orange depict non-polar, polar, glycine, positively and negatively charged residues. The purple lines indicate hydrogen bonds, while the grey shades around the ligand atoms and bonds show exposure to solvent. The screenshot illustrates the trajectory frame of the most dormant RMSD.

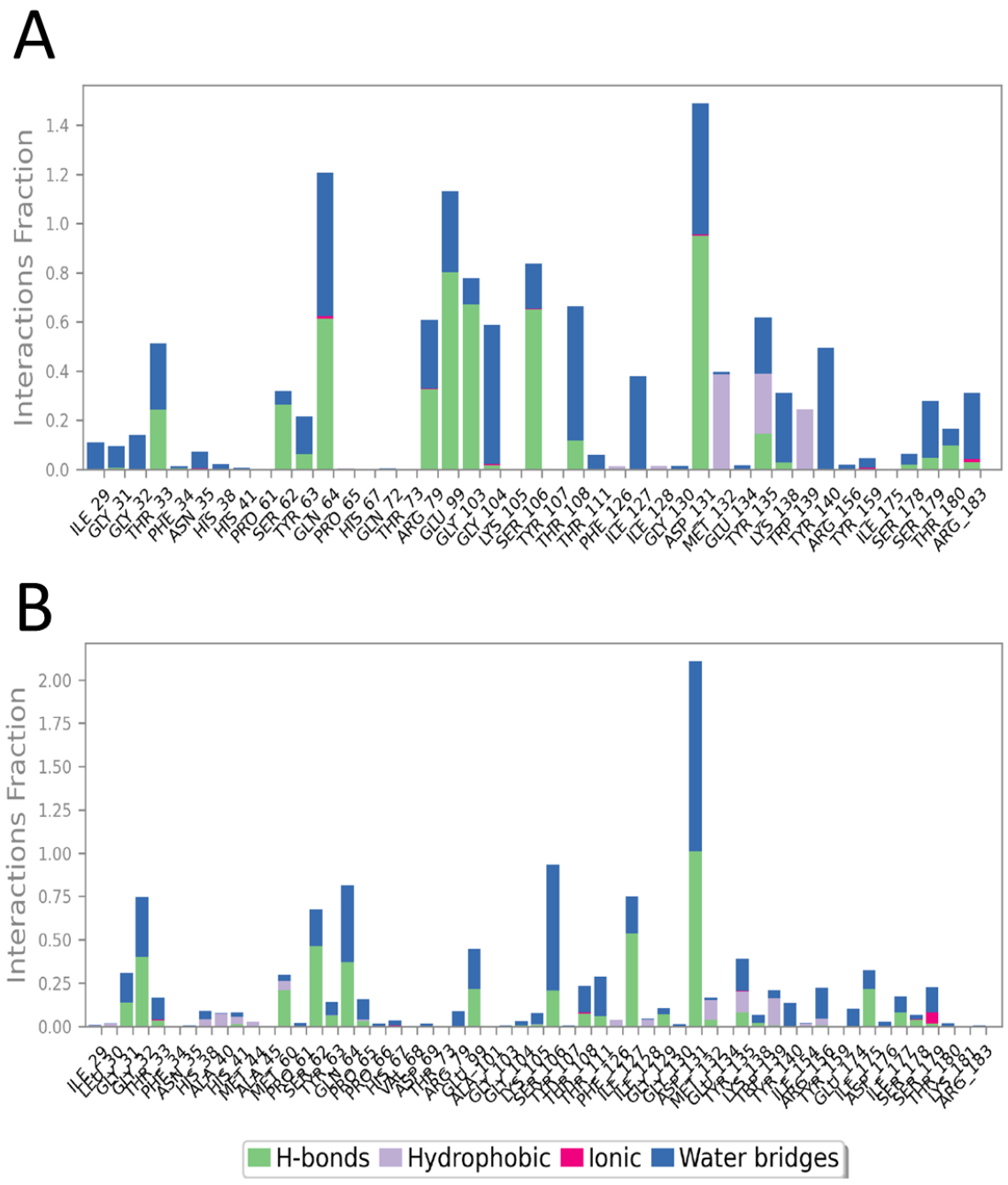



**Figure 5.9. Side-chain interaction bar chart of EfNNAT with (A) quercetin and (B) apigenin for 500 ns simulation time.** Hydrogen bonds and water bridges primarily influence the interactions. The stacked bar charts are adjusted to a standard scale throughout the trajectory.

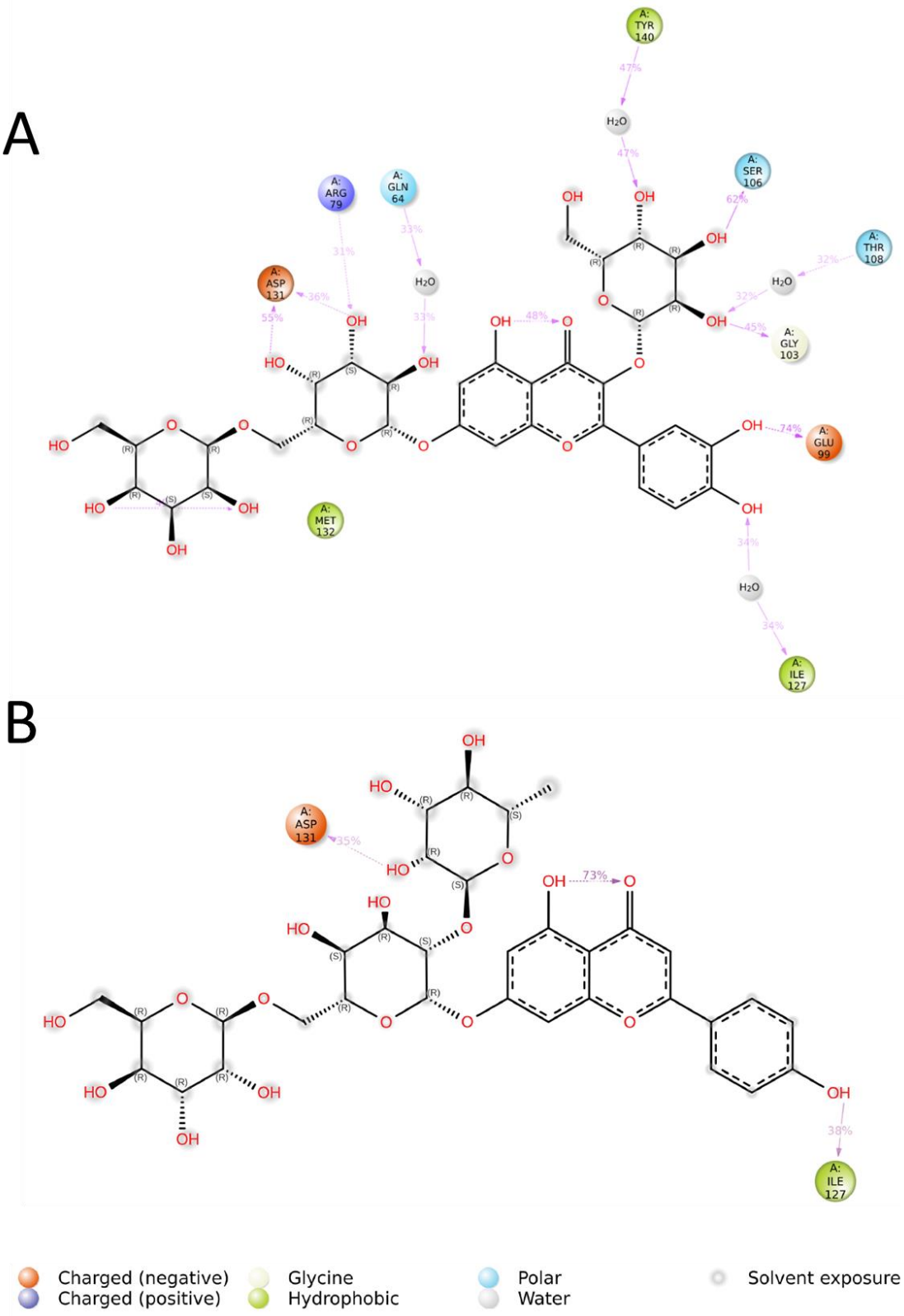



**Figure 5.10. A detailed 2D atomic interaction plot of EfNNAT-quercetin and EfNNAT-apigenin complexes**

(A) Shows specific interactions between quercetin atoms and EfNNAT residues. (B) apigenin atomic interaction with EfNNAT residues. The trajectory frame depicts interactions over 30% of the 500 ns simulation period.

### 5.5.5 Physiochemical properties and druglikeness of identified compounds

The physicochemical properties of quercetin and apigenin analysis showed that they are big molecules with a molecular weight of 788.66 g/mol and 740.66 g/mol, respectively. Quercetin possesses 22 hydrogen bond acceptors, 14 hydrogen bond donors, and ten rotatable bonds, while apigenin possesses 19 and 11 hydrogen bond acceptors and donor and nine rotatable bonds, making them highly flexible, hence, unstable and dynamic within the protein receptor. These properties elucidate the compounds' highly dynamic nature within the EfNNAT receptor.

Other physiochemical properties of the compounds include total polar surface area (TPSA) of 368.81 $Å^2$ and 308.12 $Å^2$, lipophilicity (XLOGP3) −2.64 and −1.43, solubility (log *S*) 2.14 and 0.35, and saturation (fraction Csp3) 0.55 for quercetin and apigenin respectively. The bioavailability radar in Figure 5.11 shows how these physiochemical properties fit into the druglikeness descriptors [200, 201] [204, 205]. Quercetin is predicted as too polar, big and flexible, while apigenin is predicted as too polar and big.

Furthermore, the pharmacokinetic properties of the compounds showed that they are both non-inhibitors of cytochrome P450 1A2, 2C19, 2C9, 2D6 and 3A4, eliminating potential toxicity of the compounds [201, 206, 207]. The Pan Assay Intefering Structures (PAINS) of quercetin indicate one alert score showing its structure contains catechol_A. However, apigenin did not display any form of alert [208].

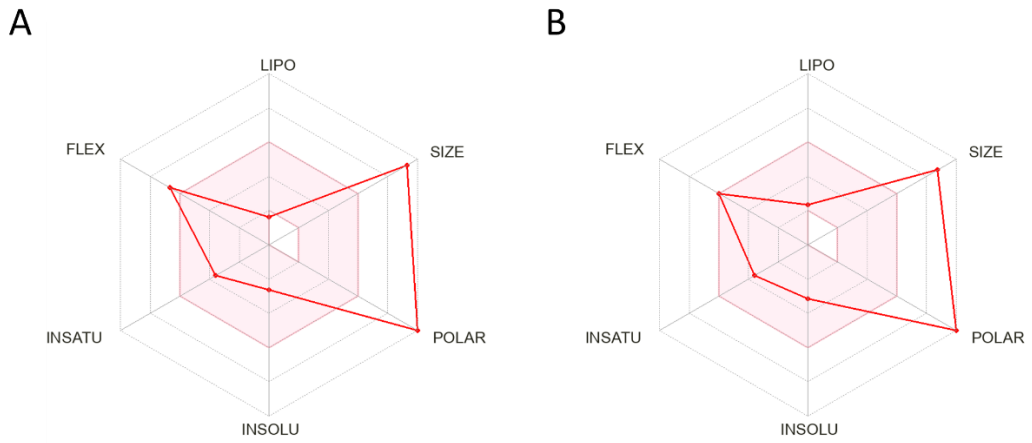



**Figure 5.11. The bioavailability radar of (A) quercetin and (B) apigenin showing how their physiochemical properties fit into the druglikeness descriptors**
The pink region indicates acceptable range, while the red points indicate where the analysed data falls on the radar. The physiochemical properties considered include solubility, polarity, saturation, lipophilicity, size and flexibility.

## 5.5.6 Empirical validation of identified compounds

Following the virtual screening of the flavonoid library to identify compounds with the best binding affinity for EfNNAT and the MD simulation to establish these compounds' interaction and stability with EfNNAT, empirically validating the computational studies was important.

Fluorescence spectroscopy was used to validate the binding of the two flavonoid compounds, quercetin and apigenin, to EfNNAT. The extrinsic fluorescent probes ANS and mant-ATP were used for this study. ANS is an anionic dye used to detect hydrophobic pockets in proteins. An increased fluorescence intensity and a blue shift usually accompany its binding and interaction with a protein cavity. Changes in the spectral properties of the dye are used to provide insight into the binding attributes of the protein [154]. Figure 5.12A shows the results of the EfNNAT-ANS studies in the presence and absence of the identified compounds.

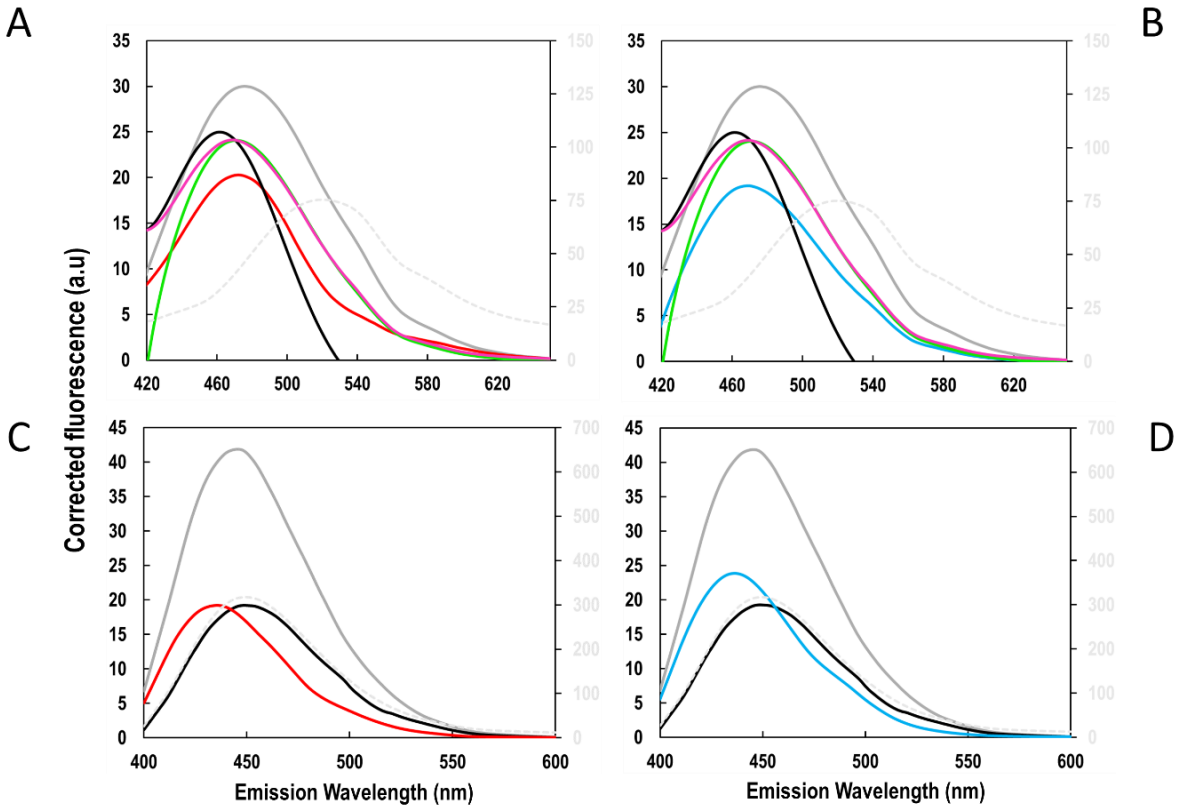



**Figure 5.12. ANS and mant-ATP Fluorescence emission spectra showing quercetin and apigenin binding to EfNNAT**

ANS binding studies for **(A)** quercetin (red) and **(B)** apigenin (blue) were conducted using EfNNAT apo (grey) and EfNNAT complex with its natural ligands ATP (black), NMN (green) and $NAD^+$ (purple) as a control. A decrease in ANS fluorescence emission relative to EfNNAT apo validates the binding of the compounds. Mant-ATP binding studies for **(C)** quercetin (red) and **(D)** apigenin (blue) were conducted using EfNNAT apo (grey) and EfNNAT-ATP complex (black) as a control. A decrease in mant-ATP fluorescence emission relative to EfNNAT apo validates the binding of the compounds. All assays were performed in 10 mM $NaH_2PO_4$, pH 7.5 at 37 °C.

The binding of ANS to the EfNNAT apo was established as indicated by a high fluorescent intensity and a blue shift in the emission wavelength compared to the free ANS in the buffer [155]. Upon incubation with quercetin and apigenin, respectively, a decrease was observed in the fluorescent intensity of ANS, implying its displacement and the binding of the compounds to EfNNAT. About 2 and 3 nm blueshift was observed in the ANS fluorescence emission wavelength in the presence of quercetin and apigenin, respectively, compared to the apo structure, suggesting minimal impact

on the structural conformation of the protein. The EfNNAT natural ligands were used as a control to confirm the binding of the identified compounds. As shown in Figure 5.12A, the binding of ATP, NMN and $NAD^+$ to EfNNAT yielded a lower ANS fluorescence compared to the EfNNAT apo, with ATP resulting in about 10 nm blueshift in the fluorescence emission wavelength and NMN and $NAD^+$ 3 nm, respectively. Compared to the natural ligands, quercetin and apigenin showed more displacement of the ANS by much lower fluorescence intensity, thus establishing the binding of the compounds to EfNNAT.

Also, mant-ATP, a fluorescent analogue of ATP, was used to probe further the compounds' propensity to bind EfNNAT. Similar to ANS spectral properties, the binding of mant-ATP to a nucleotide binding site increases the fluorescent intensity of mant-ATP, followed by a blue shift in the maximum emission wavelength. This attribute was observed with the EfNNAT apo in Figure 5.12B. Upon incubating the protein with quercetin and apigenin, respectively, mant-ATP fluorescence intensity decreased, indicating the displacement of the fluorescent probe from the binding site. The displacement observed is comparable with the protein when bound to ATP, thus validating quercetin and apigenin's ability to bind EfNNAT.

The impact of the potential inhibitors on the structural stability of EfNNAT was then validated using the thermal shift assay. The protein's melting temperature was monitored at a varying concentration of the compounds between 20 °C and 80 °C with SYPRO Orange dye. The ability of SYPRO Orange dye to access more hydrophobic pockets as the protein unfolds with increasing temperature yields increasing fluorescence emission, consequently generating a sigmoidal curve that shows the unfolding transition of the protein. A shift in this melting curve either to the left or right of the reference curve indicates decreasing or increasing stability, respectively. The assay result in Figure 5.13 showed no significant shift in the melting curve of the protein in the presence of quercetin and apigenin. Also, no significant change was observed in the protein's melting temperature except at 500 μM quercetin and apigenin concentration, where the $T_m$ of EfNNAT decreased by 0.5 °C. This suggests

that the binding of quercetin and apigenin has negligible significance on the overall structural stability of EfNNAT, and the compounds might not necessarily impact the structural stability of the stud. However, further interaction studies and activity assays such as isothermal titration calorimetry (ITC) and Surface plasmon resonance (SPR) would be required to validate this and confirm their inhibitory levels on EfNNAT.

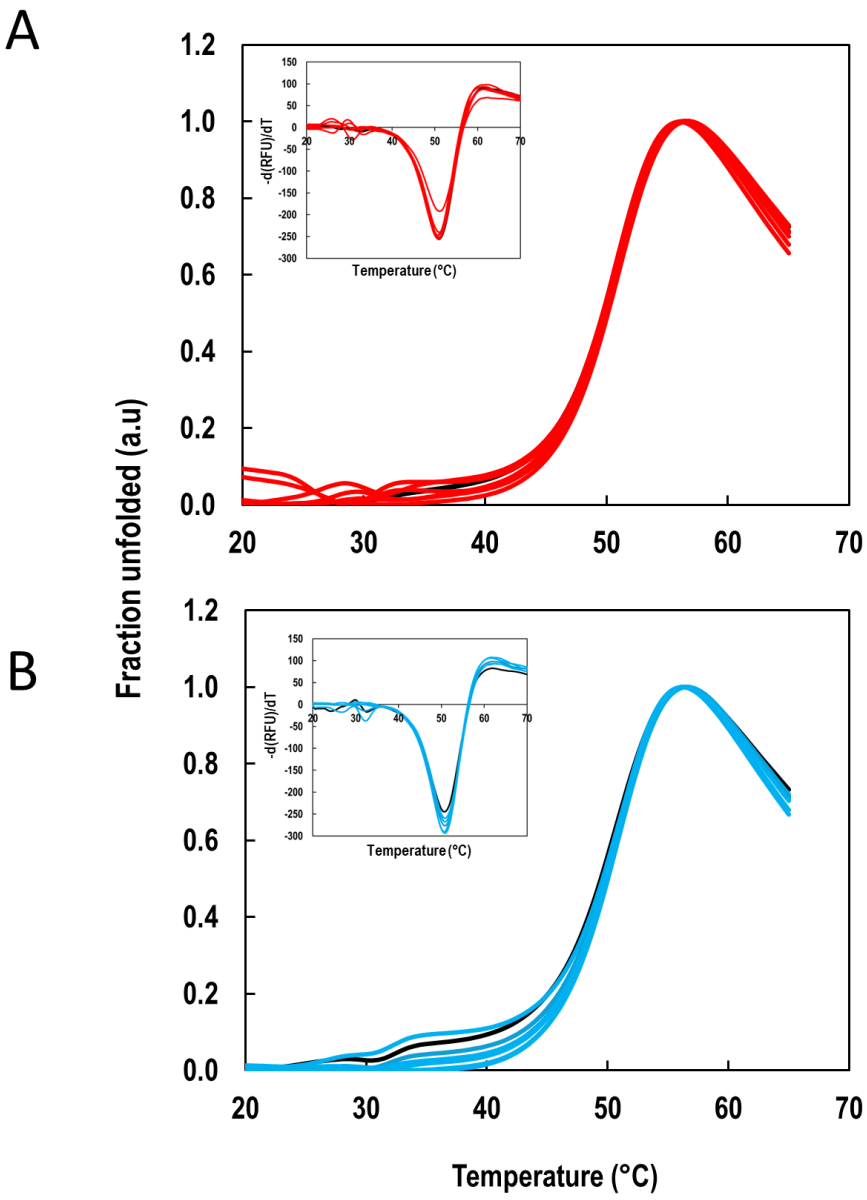



**Figure 5.13 The fluorescence thermal denaturation spectra for the binding of EfNNAT**
The unfolding transition and melting curves (inset) were obtained at varying concentrations of the compounds between 0-500 μM. The binding of the compounds showed no significant shift in the stability or $T_m$ of EfNNAT. The black spectrum indicates the EfNNAT apo.

## 5.6 Discussion

The availability of high-resolution crystal structures of EfNNAT [123] facilitated the virtual screening of structure-based inhibitors for the enzyme. In this study, we performed a virtual screening using a library of flavonoid compounds from MCE, given that they can be easily synthesised for empirical studies.

Prior to the flavonoid library screening, it was important to determine what structure to use for the HTVS, considering that the protein crystal structure used was crystallised in the apo form, which may vary with the active conformation. If this postulate is true, which conformation (Apo or active) should be used as the model for HTVS? We also postulated that by subjecting the structure to MD simulation, given that proteins are dynamic, the protein may adopt a more active conformation, implying variation in the ligand selection during HTVS. Our results affirm our postulate that MDS may induce structural and conformational perturbations as changes were observed in the secondary structural elements of the protein (Figure 5.2) and the active site conformations (Figure 5.3).

Contrary to the one large cavity exhibited by the EfNNAT minimised model, two distinct cavities were observed for the MD simulated model (Figure 5.3). Residue analysis of the two cavities (Table 5.1) showed that site_1 contains His38 and His41 of the H/TXXH signature motif and Ser178, Ser179, Thr180, Lys181, and Arg183 of the SXTXXR motif, which are required for ATP recognition and binding [66, 123, 166]. Site_2 contains Ser106, Tyr107, Thr108, Tyr109, and Thr111 of the XYTXXT motif and Trp139 and Tyr140 of the XYX motif involved in NMN/NaMN binding and pyridine ring stabilisation [123, 144]. Thus, proposing site_1 is an ATP binding site and site_2 is a NaMN/NMN binding site. These two sites were also observed in the minimised model but as one big cavity (Table 5.1 and Figure 5.3). The one large cavity observed for the minimised structure appears to be the conformation the protein would adopt upon product formation, given previous reports from other bacterial NNAT where NaAD binding to the NNAT occurs in an extended conformation. This structural

distinction suggests that the MD simulated model would probably be the more physiologically acceptable form of the structure, considering that the protein is a dual substrate binding protein with probable ATP and NaMN/NMN binding pockets.

Nonetheless, both models were utilised for ligand screening with the presumption that the models may select different ligands. Tables 5.2 and 5.3 confirm variation in the five topmost ligand selections after the initial HTVS. However, after more rigorous sampling of the ligands, both models (minimised vs MDS) selected quercetin 3-O-beta-D-glucose-7-O-beta-D-gentiobioside as the only suitable ligand with good affinity for the receptor. Quercetin binding spans both pockets of the simulated model, akin to NaAD/$NAD^+$, a putative product of the reaction catalysed by the enzyme [66, 144]. Contrary to the minimised model, whose selected ligands were all observed to span the large cavity, two of the selected ligands (Bavachalcone and Sappanone A) by the simulated model were small molecules fitting one of its two cavities (Table 5.3). This underscores that the simulated model can bind either a molecule that spans both binding sites, like the product is expected to bind, or a molecule occupying either of the substrate binding sites. We anticipated smaller molecules to bind the MD-simulated model than bulky molecules; however, a library of smaller molecular weights of applying the Lipinski Rule of five during the HTVS may increase the possibility of identifying ligands that will be more active towards the dual cavities of the MD-simulated models.

The interaction and stability of quercetin within the EfNNAT receptor were analysed using the simulated model. Analysis was performed to compare apigenin, the topmost compound identified after the initial HTVS. This was to provide a better understanding of the selection and binding of quercetin to EfNNAT, particularly because apigenin demonstrated a high binding free energy coupled with the fact that the two modelled structures commonly selected it during the initial HTVS.

As established through MD simulation and validated empirically through fluorescence spectroscopy studies (Figure 5.12), quercetin and apigenin bind EfNNAT. The binding of quercetin slightly stabilises the protein backbone as construed from

the Cα RMSD and RMSF of the protein, with restrictions observed particularly in the movement of flexible regions 98-110 and 130-142 close to the NaMN binding site (Figure 5.4)[123]. This was also evident from the detailed atomic interaction of quercetin with EfNNAT, where more than 30% of the 500 ns simulation time (figure 5.10) showed that the protein stabilisation occurred through residues Glu64, Arg79, Glu99, Gly103, Ser106, Thr108, Ile127, Asp131, and Tyr140. These residues are identified as NaMN binding motifs and are involved in stabilising the NaMN substrate[123, 144]. However, most of these interactions were hydrogen bonding and water bridges. They did not exceed 75% of the simulation time, nor were there slightly tighter interactions like ionic bonds involved, thus suggesting no much tight binder that could hold onto this molecule for as long as possible.

On the other hand, Apigenin showed minimal impact on stabilising the protein backbone, with more residue fluctuations observed than quercetin binding (Figure 5.4). Its percentage interaction plot (Figure 5.10) showed that there were no many interactions strong enough to hold onto the molecule throughout the 500 ns simulation time, hence acting as a weaker binder. Moreover, the few residues involved in the binding interacted mainly through hydrogen bonding (Figure 5.8) rather than ionic interaction, such as salt bridges, which would have contributed to slightly tighter interactions.

The inability of both compounds to proficiently establish a long-lasting stable interaction with the protein residues could be attributed to the highly dynamic nature of the ligands within the protein receptor as revealed from the ligand's RMSD, MSA and SASA. This concurs with the physicochemical properties of quercetin and apigenin, which show they are big molecules with high hydrogen bond potentials and rotatable bonds, making them highly dynamic and unstable [201]. Apigenin exhibited over 7 Å deviation in its RMSD, implying its instability within the receptor; hence, it could not form stable and prolonged contacts with the protein residues. Quercetin, though dynamic, could still establish some contact as its deviation was within 4 Å. Furthermore, molecules predisposed to forming intra-HB rarely make available

functional groups to interact with the amino acid side chains. This is evident from the detailed atomic interaction of quercetin and apigenin with EfNNAT (Figure 5.10). Rather than their functional groups forming contacts with the protein residues, intra-HB contacts were observed between the hydroxyl and carboxylic group of the ligands, with apigenin establishing intra-HB contacts as high as 73% of the simulation time.

At large, the dynamism of the inhibitors within the protein receptor elucidates why the protein's overall structure and global compactness were not impacted by the ligand, as shown from the Rog (Figure 5.5). This is also validated empirically through a negligible blue shift in the ANS emission wavelength and the melting temperature ($T_m$) and thermal unfolding of the protein (Figure 5.13) despite the increasing concentration of the compounds. We anticipated that the two ligands would enhance thermal stability, suggesting a relatively strong binding. Surprisingly, our results indicate that while both ligands bind to the enzyme, as validated through the fluorescence spectroscopy study (Figure 12), they fail to induce thermal stability. One possible explanation is that these ligands lack sufficient thermal stability to remain bound during heat denaturation. Previous studies have demonstrated that flavonoids, like the ones examined here, are prone to denaturation upon heating [209-212]. Additionally, molecular dynamics (MD) simulations revealed that both ligands experienced up to four intra-hydrogen bonds (intraHB) during the simulation. This suggests that the flavonoids may undergo intra-hydrogen bond stabilisation upon heating, potentially hindering hydrogen bond interactions with sidechains in the binding sites. Consequently, this could have impacted the overall $\Delta G_{bind}$ of the EfNNAT:flavonoid interaction.

Although both compounds may not bind as tightly as anticipated, quercetin binds better than apigenin, validating its selection after the SP and XP docking and ligand sampling. Its binding interface attracted more water molecules that could contribute favourably to free energy binding (Figure 5.8). Water molecules have been shown to play a significant role in ligand recognition and protein-ligand stabilisation [213, 214]. Their occupancy within a protein-ligand interface can facilitate ligand

binding by contributing favourably to free energy [213, 215]. The fact that most of the residue quercetin interacted with are identified as NaMN stabilisers suggests that the ligand could exhibit some inhibitory effect against EfNNAT. Hence, the need to perform further empirical work, including *in vivo* and *in vitro* studies such as antibacterial activity, enzyme activity assays, binding interactions and kinetics with ITC and SPR, would be essential to validate this computational inference and to determine the potency and inhibitory levels of the identified compounds against EfNNAT.

However, with the lack of adherence to Lipinski rule of five (MW>500, NorO>10, NHorOH>5) [216] and poor bioavailability predicators [201, 204, 205], the identified compound do not fit into the druglikeness descriptor. Moreover, the PAIN alert score of catechol_A substructure in quercetin indicates that a false positive result may be obtained regarding quercetin interaction with EfNNAT, given the potential of the substructure to interact non-specifically with a wide range of biological targets [208]. Catechol_A, like other PAINS, can form nonspecific interactions, chelate metal ions, or undergo redox reactions that can affect assay readouts, thereby complicating the identification of true active compounds [217]. Its presence may compromise the true activity of quercetin, thus reducing its efficacy as a drug candidate. However, thorough validation can be conducted using biochemical, biophysical, and cellular assays to ensure that the observed activity is specific and relevant to the targeted EfNNAT. Furthermore, the chemical structure of quercetin may be modified to remove PAINS features while retaining activity against the target [208].

In conclusion, considering flavonoids are indispensable in medicinal applications, a library of synthesiable flavonoids from MCE was explored as a drug discovery tool against EfNNAT. This study established a pipeline between identifying potential inhibitors and validating the biological activity of the compound towards EfNNAT protein, thus making EfNNAT a druggable target. Although the flavonoid identified in this study might not be the best drug candidate, it is because of the nature of the library used for this study, considering we wanted synthesisable ligands and

ligands that have no restriction regarding patents. The mechanistic insights provided a foundation for advancing drug discovery efforts targeting EfNNAT and other bacterial NNAT enzymes. With a more extensive library, inhibitors with better binding affinity and druglikeness can be identified.

# Chapter 6
# CONCLUSION

Bacterial infections remain a major global health concern, exacerbated by the rise of antibiotic resistance. NNAT has garnered attention as a promising druggable target due to its pivotal role in $NAD^+$ biosynthesis, which is critical for the survival of bacterial pathogens. Inhibiting NNAT could disrupt bacterial metabolism and render bacteria more susceptible to existing antibiotics. In this study, the structural and functional aspects of *E. faecium* NNAT were expounded using empirical and computational tools with the aim of validating EfNNAT as a druggable target.

The successful overexpression and purification of EfNNAT to homogeneity represent a significant milestone in this research, laying the groundwork for future breakthroughs. Biophysical characterisation and determination of the three-dimensional structure in both apo and liganded forms provided valuable insights into the structural and mechanistic attributes of EfNNAT. The apo, adenine-bound and NMN-bound structures were determined at a high-resolution, two of which have been deposited into the PDB as the first-ever structures of the *E. faecium* NNAT. The results obtained across the two studies were comparable.

Analysis of the apo and adenine-bound structures revealed crucial information about the EfNNAT binding site and its potential interactions with ATP, NaMN, and NMN, shedding light on its substrate specificity. Co-crystallisation attempts with ATP, NMN, and $NAD^+$, while successful for NMN, did not yield structures for ATP and $NAD^+$, leading to the use of an ATP analogue, CTP, for co-crystallisation. Although the structure of EfNNAT-CTP could not be refined with the current dataset, new diffraction data can be collected using the established procedure for obtaining high-quality EfNNAT-CTP crystals, allowing for the determination of the three-dimensional structure.

Insights from the NMN-bound structure hinted at the presence of another binding site in EfNNAT, distinct from the active site. Analysis of this site suggests the possibility of it being an allosteric site. Further computational studies, including molecular docking of small molecules into the site in the presence and absence of substrate and molecular dynamic simulations, are warranted to validate this site and explore its potential significance.

With the availability of high-resolution structures of EfNNAT, further structural evaluation and drug-based screening were performed in order to identify potential inhibitors. This was achieved using synthesisable flavonoids for empirical validation, given their unrestricted patency. Quercetin 3-O-beta-D-glucose-7-O-beta-D-gentiobioside was the only compound identified as a potential EfNNAT inhibitor and was validated empirically to bind the EfNNAT. Although the compound did not come as a good drug candidate, the results confirmed EfNNAT as a druggable target, highlighting the need for a more extensive inhibitor library to identify compounds with higher affinity.

Overall, this research has significantly contributed to bridging the gap in understanding the structure-function relationship of EfNNAT and its potential as a druggable target. The successful expression, purification, and structural elucidation of EfNNAT, along with the identification of potential inhibitors, underscore the druggability of this protein and pave the way for future drug discovery efforts targeting EfNNAT and other bacterial NNAT enzymes.

# References


1. Oliveira, D.M.P.D., et al., *Antimicrobial resistance in ESKAPE pathogens.* Clinical Microbiology Reviews, 2020. **33**(3): p. 10.1128/cmr.00181-19.

2. Murray, C.J.L., et al., *Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis.* The Lancet, 2022. **399**(10325): p. 629-655.

3. Rice, L.B., *Federal funding for the study of antimicrobial resistance in nosocomial pathogens: No ESKAPE.* The Journal of Infectious Diseases, 2008. **197**(8): p. 1079-1081.

4. Strathdee, S.A., S.C. Davies, and J.R. Marcelin, *Confronting antimicrobial resistance beyond the COVID-19 pandemic and the 2020 US election.* The Lancet, 2020. **396**(10257): p. 1050-1053.

5. O'Neill, J., *Tackling drug-resistant infections globally: final report and recommendations*, 2016.

6. Jim, O.N., *Antimicrobial resistance: tackling a crisis for the health and wealth of nations*, in Review Antimicrobial Resististance, 2014.

7. Ananda, T., et al., *Nosocomial infections and role of nanotechnology.* Bioengineering, 2022. **9**(2): p. 51.

8. Ramos, S., et al., *Enterococci, from harmless bacteria to a pathogen.* Microorganisms, 2020. **8**(8).

9. Guzman Prieto, A.M., et al., *Global emergence and dissemination of Enterococci as nosocomial pathogens: Attack of the clones?* Frontiers in Microbiology, 2016. **7**.

10. López-Luis, B.A., et al., *Risk factors and outcomes associated with vancomycin-resistant Enterococcus faecium and ampicillin-resistant Enterococcus faecalis bacteraemia: A 10-year study in a tertiary-care centre in Mexico City.* Journal of Global Antimicrobial Resistance, 2021. **24**: p. 198-204.

11. Lee, T., et al., *Molecular characterization and evolution of the first outbreak of vancomycin-resistant Enterococcus faecium in Western Australia.* International Journal of Antimicrobial Agents, 2019. **53**(6): p. 814-819.

12. Fiore, E., D. Van Tyne, and M.S. Gilmore, *Pathogenicity of Enterococci.* Microbiology Spectrum, 2019. **7**(4).

13. O'Toole, R.F., et al., *Vancomycin-resistant Enterococcus faecium and the emergence of new sequence types associated with hospital infection.* Research in Microbiology, 2023. **174**(4): p. 104046.

14. Abebe, T., *Biofilm formation by Enterococcus faecalis and Enterococcus faecium: Review.* International Journal of Research Studies in Biosciences, 2019: p. 2349-0365.

15. Da Silva Fernandes, M., et al., *Formation of multi-species biofilms by Enterococcus faecium, Enterococcus faecalis, and Bacillus cereus isolated from ricotta processing and effectiveness of chemical sanitation procedures.* International Dairy Journal, 2017. **72**: p. 23-28.

16. Șchiopu, P., et al., *An overview of the factors involved in biofilm production by the Enterococcus genus.* International Journal of Molecular Sciences, 2023. **24**(14).

17. Miller, W.R., J.M. Munita, and C.A. Arias, *Mechanisms of antibiotic resistance in enterococci.* Expert Review of Anti-infective Therapy, 2014. **12**(10): p. 1221-36.

18. Murima, P., J.D. McKinney, and K. Pethe, *Targeting bacterial central metabolism for drug development.* Chemistry and biology, 2014. **21**(11): p. 1423-1432.

19. Balducci, E., N. Braidy, and M. Migaud, *Editorial: NAD+ metabolism as a novel target against infection–Volume II.* Frontiers in Molecular Biosciences, 2022. **9**.

20. Gerdes, S.Y., et al., *From genetic footprinting to antimicrobial drug targets: Examples in cofactor biosynthetic pathways.* Journal of bacteriology, 2002. **184**(16): p. 4555.

21. Osterman, A.L. and T.P. Begley, *A subsystems-based approach to the identification of drug targets in bacterial pathogens*, in *Systems Biological Approaches in Infectious Diseases*, H.I. Boshoff and C.E. Barry, Editors. 2007, Birkhäuser Basel: Basel. p. 131-170.

22. Sorci, L., et al., *Genomics-driven reconstruction of Acinetobacter NAD metabolism: Insights for antibacterial target selection.* The Journal of biological chemistry, 2010. **285**: p. 39490-9.

23. Bi, J., H. Wang, and J. Xie, *Comparative genomics of NAD(P) biosynthesis and novel antibiotic drug targets.* Journal of Cellular Physiology, 2011. **226**(2): p. 331-340.

24. Chen, Y., et al., *From bacteria to biomedicine: Developing therapies exploiting NAD+ metabolism.* Bioorganic Chemistry, 2024. **142**: p. 106974.

25. Sultani, G., et al., *NAD(+) : A key metabolic regulator with great therapeutic potential.* Journal of Neuroendocrinol, 2017. **29**(10).

26. Xiao, W., et al., *NAD(H) and NADP(H) Redox couples and cellular energy metabolism.* Antioxidants and Redox Signaling, 2017. **28**(3): p. 251-272.

27. Lau, C., M. Niere, and M. Ziegler, *The NMN/NaMN adenylyltransferase (NMNAT) protein family.* Frontiers in Bioscience, 2009. **14**: p. 410-31.

28. Sorci, L., et al., *Targeting NAD biosynthesis in bacterial pathogens: Structure-based development of inhibitors of nicotinate mononucleotide adenylyltransferase NadD.* Chemistry and biology, 2009. **16**(8): p. 849-861.

29. Rodionova, I.A., et al., *Metabolic and bactericidal effects of targeted suppression of NadD and NadE enzymes in mycobacteria.* mBio, 2014. **5**(1): p. e00747-13.

30. Oprea, S.F., et al., *Molecular and clinical epidemiology of vancomycin-resistant Enterococcus faecalis.* Journal of Antimicrobial Chemotherapy, 2004. **53**(4): p. 626-630.

31. Robert C. Moellering, Jr., *Emergence of Enterococcus as a significant pathogen.* Clinical Infectious Diseases, 1992. **14**(6): p. 1173-1176.

32. Gao, W., B.P. Howden, and T.P. Stinear, *Evolution of virulence in Enterococcus faecium, a hospital-adapted opportunistic pathogen.* Current Opinion in Microbiology, 2018. **41**: p. 76-82.

33. Treitman, A.N., et al., *Emerging incidence of Enterococcus faecium among hospital Isolates (1993 to 2002).* Journal of Clinical Microbiology, 2005. **43**(1): p. 462.

34. Daniel, D.S., et al., *Public health risks of multiple-drug-resistant Enterococcus spp. in Southeast Asia.* Applied and environmental microbiology, 2015. **81**(18): p. 6090-6097.

35. Edwardson, S. and C. Cairns, *Nosocomial infections in the ICU.* Anaesthesia and Intensive Care Medicine, 2019. **20**(1): p. 14-18.

36. Elward, A.M., et al., *Nosocomial primary bloodstream infections in intensive care unit patients in a nonteaching community medical center: A 21-Month Prospective Study.* Clinical Infectious Diseases, 2001. **33**(8): p. 1329-1335.

37. Zhou, X., et al., *Enterococcus faecium: from microbiological insights to practical recommendations for infection control and diagnostics.* Antimicrobial Resistance and Infection Control, 2020. **9**(1): p. 130.

38. Cimen, C., et al., *Vancomycin-resistant enterococci (VRE) in hospital settings across European borders: a scoping review comparing the epidemiology in the Netherlands and Germany.* Antimicrobial Resistance and Infection Control, 2023. **12**(1): p. 78.

39. Shrestha, S., et al., *Prevalence of vancomycin-resistant enterococci in Asia—A systematic review and meta-analysis.* Journal of Clinical Pharmacy and Therapeutics, 2021. **46**(5): p. 1226-1237.

40. Chua, W.C. and A.R. Zaidah, *First reported cases of linezolid-resistant vancomycin-resistant enterococci in South-East Asia: A report of three cases and literature review.* Proceedings of Singapore Healthcare, 2021. **30**(4): p. 309-312.

41. Lin, P.-Y., et al., *Epidemiological profiles and pathogenicity of Vancomycin-resistant Enterococcus faecium clinical isolates in Taiwan.* PeerJ, 2023. **11**: p. e14859.

42. Cairns, K.A., et al., *Therapeutics for Vancomycin-resistant Enterococcal bloodstream infections.* Clinical Microbiology Reviews, 2023. **36**(2): p. e00059-22.

43. Kibwana, U.O., et al., *Antimicrobial resistance profile of Enterococcus species and molecular characterization of Vancomycin resistant Enterococcus faecium from the fecal samples of newly diagnosed adult HIV patients in Dar es Salaam, Tanzania.* Frontiers in Tropical Diseases, 2024. **5**.

44. Wada, Y., et al., *Vancomycin-resistant Enterococci (VRE) in Nigeria: The first systematic review and meta-analysis.* Antibiotics, 2020. **9**(9): p. 565.

45. Melese, A., C. Genet, and T. Andualem, *Prevalence of Vancomycin resistant enterococci (VRE) in Ethiopia: a systematic review and meta-analysis.* BMC Infectious Diseases, 2020. **20**(1): p. 124.

46. Browning, S., J.S. Davis, and B.G. Mitchell, *Have gloves and gowns had their day? An Australian and New Zealand practice and attitudes survey about contact precautions for MRSA and VRE colonisation.* Infection, Disease and Health, 2023. **28**(3): p. 221-225.

47. Lee, T., et al., *Antimicrobial-resistant CC17 Enterococcus faecium: The past, the present and the future.* Journal of Global Antimicrobial Resistance, 2019. **16**: p. 36-47.

48. Santajit, S. and N. Indrawattana, *Mechanisms of antimicrobial resistance in ESKAPE pathogens.* BioMed research international, 2016. **2016**: p. 2475067-2475067.

49. Barna, J.C. and D.H. Williams, *The structure and mode of action of glycopeptide antibiotics of the vancomycin group.* Annual Review of Microbiology, 1984. **38**: p. 339-57.

50. Brickner, S.J., et al., *Synthesis and antibacterial activity of U-100592 and U-100766, two oxazolidinone antibacterial agents for the potential treatment of multidrug-resistant gram-positive bacterial infections.* Journal of Medicinal Chemistry, 1996. **39**(3): p. 673-9.

51. Schriever, C.A., et al., *Daptomycin: A novel cyclic lipopeptide antimicrobial.* American Journal of Health-System Pharmacy, 2005. **62**(11): p. 1145-1158.

52. Petersen, P.J., et al., *In vitro and in vivo antibacterial activities of a novel glycylcycline, the 9-t-butylglycylamido derivative of minocycline (GAR-936).* Antimicrobial Agents and Chemotherapy, 1999. **43**(4): p. 738-744.

53. Jamjian, C., M.S. Barrett, and R.N. Jones, *Antimicrobial characteristics of quinupristin/dalfopristin (Synercid® at 30:70 ratio) compared to alternative ratios for in vitro testing.* Diagnostic Microbiology and Infectious Disease, 1997. **27**(4): p. 129-138.

54. Eustice, D.C. and J.M. Wilhelm, *Mechanisms of action of aminoglycoside antibiotics in eucaryotic protein synthesis.* Antimicrobial Agents and Chemotherapy, 1984. **26**(1): p. 53-60.

55. Osborne, N.G., *Macrodantin.* Journal of Gynecologic Surgery, 1992. **8**(3): p. 195-196.

56. Zelkowitz, L., G. Arimura, and A. Yunis, *Chloramphenicol and protein synthesis in mammalian cells.* The Journal of Laboratory and Clinical Medicine, 1968. **71**(4): p. 596-609.

57. Hendlin, D., et al., *Phosphonomycin, a new antibiotic produced by strains of streptomyces.* Science, 1969. **166**(3901): p. 122-123.

58. Christensen, B., et al., *Phosphonomycin: structure and synthesis.* Science, 1969. **166**(3901): p. 123-125.

59. Draghi, D., et al. *Anti-enterococcal activity profile of oritavancin, a potent lipoglycopeptide under development for use against Gram-positive infections*. in 47th Interscience Conference on Antimicrobial Agents and Chemotherapy. 2007.

60. Guskey, M.T. and B.T. Tsuji, *A Comparative Review of the Lipoglycopeptides: Oritavancin, Dalbavancin, and Telavancin.* Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy, 2010. **30**(1): p. 80-94.

61. Leadbetter, M.R., et al., *Hydrophobic vancomycin derivatives with improved ADME properties discovery of telavancin (TD-6424).* The Journal of Antibiotics, 2004. **57**(5): p. 326-336.

62. Zhanel, G.G., et al., *Ceftaroline: a novel broad-spectrum cephalosporin with activity against meticillin-resistant Staphylococcus aureus.* Drugs, 2009. **69**: p. 809-831.

63. Uda, A., et al., *Risk factors for the acquisition of enterococcus faecium infection and mortality in patients with Enterococcal bacteremia: A 5-Year Retrospective Analysis in a Tertiary Care University Hospital.* Antibiotics, 2021. **10**(1): p. 64.

64. Hemapanpairoa, J., et al., *Does vancomycin resistance increase mortality? Clinical outcomes and predictive factors for mortality in patients with Enterococcus faecium infections.* Antibiotics, 2021. **10**(2): p. 105.

65. Xie, N., et al., *NAD(+) metabolism: pathophysiologic mechanisms and therapeutic potential.* Signal Transduct Target Ther, 2020. **5**(1): p. 227.

66. Zhang, H., et al., *Crystal structures of E. coli nicotinate mononucleotide adenylyltransferase and its complex with deamido-nad.* Structure, 2002. **10**(1): p. 69-79.

67. Liu, G., et al., *Nucleoside salvage pathway for NAD biosynthesis in Salmonella typhimurium.* Journal of Bacteriology, 1982. **152**(3): p. 1111.

68. Kurnasov, O., et al., *NAD Biosynthesis: Identification of the tryptophan to quinolinate pathway in bacteria.* Chemistry and Biology, 2003. **10**(12): p. 1195-1204.

69. Begley, T.P., et al., *The biosynthesis of nicotinamide adenine dinucleotides in bacteria*, in *Vitamins and Hormones*. 2001, Academic Press. p. 103-119.

70. Kurnasov, O.V., et al., *Ribosylnicotinamide kinase domain of NadR protein: Identification and implications in NAD biosynthesis.* Journal of Bacteriology, 2002. **184**(24): p. 6906.

71. Magni, G., et al., *Enzymology of NAD+ synthesis.* Advances in Enzymology and Related Areas of Molecular Biology, 1999. **73**: p. 135-82.

72. Jayaram, H., P. Kusumanchi, and J.A. Yalowitz, *NMNAT expression and its relation to NAD metabolism.* Current Medicinal Chemistry, 2011. **18**: p. 1962-72.

73. Bieganowski, P. and C. Brenner, *Discoveries of nicotinamide riboside as a nutrient and conserved NRKgenes establish a preiss-handler independent route to nad in fungi and humans.* Cell, 2004. **117**(4): p. 495-502.

74. Katoh, A. and T. Hashimoto, *Molecular biology of pyridine nucleotide and nicotine biosynthesis*. Frontiers in Bioscience, 2004. Vol. 9. 1577-86.

75. Zhou, T., et al., *Structure of human nicotinamide/nicotinic acid mononucleotide adenylyltransferase: Basis for the dual substrate specificity and activation of the oncolytic agent tiazofurin.* Journal of Biological Chemistry, 2002. **277**(15): p. 13148-13154.

76. Wilkinson, A., J. Day, and R. Bowater, *Bacterial DNA ligases.* Molecular Microbiology, 2001. **40**(6): p. 1241-1248.

77. Imai, S.-i., et al., *Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase.* Nature, 2000. **403**(6771): p. 795-800.

78. Raffaelli, N., et al., The *Escherichia coli NadR regulator is endowed with nicotinamide mononucleotide adenylyltransferase activity.* Journal of Bacteriology, 1999. **181**(17): p. 5509.

79. Gazzaniga, F., et al., *Microbial NAD metabolism: lessons from comparative genomics.* Microbiology and Molecular Biology Reviews, 2009. **73**(3): p. 529-541.

80. Schweiger, M., et al., *Characterization of recombinant human nicotinamide mononucleotide adenylyl transferase (NMNAT), a nuclear enzyme essential for NAD synthesis.* FEBS Letters, 2001. **492**(1-2): p. 95-100.

81. Raffaelli, N., et al., *Identification of a novel human nicotinamide mononucleotide adenylyltransferase.* Biochemical and Biophysical Research Communications, 2002. **297**(4): p. 835-840.

82. Sorci, L., et al., *Quinolinate salvage and insights for targeting NAD biosynthesis in group A Streptococci.* Journal of Bacteriology, 2013. **195**(4): p. 726.

83. Ladd, M.F.C., R.A. Palmer, and R.A. Palmer, *Structure determination by X-ray crystallography*. Springer, 1977. Vol. 233.

84. Timofeev, V. and V. Samygina, *Protein Crystallography: Achievements and challenges.* Crystals, 2023. **13**(1): p. 71.

85. Khalil, A. and M. Kashif, *Nuclear magnetic resonance spectroscopy for quantitative analysis: a review for its application in the chemical, pharmaceutical and medicinal domains.* Critical Reviews in Analytical Chemistry, 2023. **53**(5): p. 997-1011.

86. Nogales, E. and J. Mahamid, *Bridging structural and cell biology with cryo-electron microscopy.* Nature, 2024. **628**(8006): p. 47-56.

87. Cheng, A. and Y. Yu, *Recent advances in data collection for Cryo-EM methods.* Current Opinion in Structural Biology, 2024. **86**: p. 102795.

88. Tants, J.-N. and A. Schlundt, *Advances, applications, and perspectives in small-angle x-ray scattering of RNA.* ChemBioChem, 2023. **24**(17): p. e202300110.

89. Khalikova, M., et al., *What is the role of current mass spectrometry in pharmaceutical analysis?* Mass Spectrometry Reviews, 2024. **43**(3): p. 560-609.

90. Maveyraud, L. and L. Mourey, *Protein x-ray crystallography and drug discovery.* Molecules, 2020. **25**(5): p. 1030.

91. Ameh, E.S., *A review of basic crystallography and x-ray diffraction applications.* The International Journal of Advanced Manufacturing Technology, 2019. **105**(7): p. 3289-3302.

92. Au - Sandy, J., et al., *Crystallization and in situ room temperature data collection using the crystallization facility at harwell and beamline VMXi, diamond light source.* Journal of Visualised Experiments, 2024. (205): p. e65964.

93. Tarver, C.L. and M.L. Pusey, *Effects of ionic liquids as additives on protein crystallization*, in *advanced methods in structural biology*. Springer, 2023. p. 187-197.

94. Kwan, T.O., et al., *High-throughput protein crystallization via microdialysis.* Journal of Visualised Experiments, 2023. (193): p. e64744.

95. Winter, G., *xia2: an expert system for macromolecular crystallography data reduction.* Journal of Applied Crystallography, 2010. **43**(1): p. 186-190.

96. Winter, G., et al., *DIALS: implementation and evaluation of a new integration package.* Acta Crystallographica Section D, 2018. **74**(2): p. 85-97.

97. Vonrhein, C., et al., *Data processing and analysis with the autoPROC toolbox.* A Acta Crystallographica Section D: Biological Crystallography, 2011. **67**(Pt 4): p. 293-302.

98. Hendrickson, W., *Facing the phase problem.* Journal of the International Union of Crystallography, 2023. **10**(5): p. 521-543.

99. McCoy, A.J., et al., *Phaser crystallographic software.* Journal of Applied Crystallography, 2007. **40**(4): p. 658-674.

100. Vagin, A. and A. Teplyakov, *Molecular replacement with MOLREP.* Acta Crystallographica Section D: Biological Crystallography, 2010. **66**(1): p. 22-25.

101. Keegan, R.M., et al., *Recent developments in MrBUMP: better search-model preparation, graphical interaction with search models, and solution improvement and assessment.* Acta Crystallographica Section D, 2018. **74**(3): p. 167-182.

102. Usón, I. and G.M. Sheldrick, *An introduction to experimental phasing of macromolecules illustrated by SHELX; new autotracing features.* Acta Crystallographica Section D: Structural Biology, 2018. **74**(2): p. 106-116.

103. Adams, P.D., et al., *PHENIX: a comprehensive Python-based system for macromolecular structure solution.* Acta Crystallographica Section D: Biological Crystallography, 2010. **66**(Pt 2): p. 213-21.

104. Adams, P.D., et al., *The Phenix software for automated determination of macromolecular structures.* Methods, 2011. **55**(1): p. 94-106.

105. Sheldrick, G.M., *Crystal structure refinement with SHELXL.* Acta Crystallographica Section C: Structural Chemistry, 2015. **71**(1): p. 3-8.

106. Bourhis, L.J., et al., *The anatomy of a comprehensive constrained, restrained refinement program for the modern computing environment - Olex2 dissected.* Acta Crystallographica Section A, 2015. **71**(1): p. 59-75.

107. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics.* Acta Crystallographica Section D, 2004. **60**(12 Part 1): p. 2126-2132.

108. Agirre, J., et al., *The CCP4 suite: integrative software for macromolecular crystallography.* Acta Crystallographica Section D: Structural Biology, 2023. **79**(6): p. 449-461.

109. Cowtan, K., *The Buccaneer software for automated model building. 1. Tracing protein chains.* Acta Crystallographica Section D: Biological Crystallography, 2006. **62**(9): p. 1002-1011.

110. Holton, J.M., et al., *The R-factor gap in macromolecular crystallography: an untapped potential for insights on accurate structures.* The FEBS Journal, 2014. **281**(18): p. 4046-4060.

111. Liebschner, D., et al., *Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix.* Acta Crystallographica Section D, 2019. **75**(10): p. 861-877.

112. Murshudov, G.N., et al., *REFMAC5 for the refinement of macromolecular crystal structures.* Acta Crystallographica Section D, 2011. **67**(4): p. 355-367.

113. Murshudov, G.N., A.A. Vagin, and E.J. Dodson, *Refinement of macromolecular structures by the maximum-likelihood method.* Acta Crystallographica Section D, 1997. **53**(3): p. 240-255.

114. Williams, C.J., et al., *MolProbity: More and better reference data for improved all-atom structure validation.* Protein Science, 2018. **27**(1): p. 293-315.

115. Berman, H., K. Henrick, and H. Nakamura, *Announcing the worldwide protein data bank.* Nature Structural and Molecular Biology, 2003. **10**(12): p. 980-980.

116. Berman, H.M., et al., *The protein data bank.* Nucleic Acids Research, 2000. **28**(1): p. 235-242.

117. Lu, S., et al., *Structure of nicotinic acid mononucleotide adenylyltransferase from Bacillus anthracis.* Acta crystallographica. Section F, Structural biology and crystallization communications, 2008. **64**(Pt 10): p. 893-898.

118. Bathke, J., et al., *Structural and functional characterization of Plasmodium falciparum nicotinic acid mononucleotide adenylyltransferase.* Journal of Molecular Biology, 2016. **428**(24, Part B): p. 4946-4961.

119. Han, S., et al., *Crystal structure of nicotinic acid mononucleotide adenylyltransferase from Staphyloccocus aureus: Structural basis for NaAD interaction in functional dimer.* Journal of Molecular Biology, 2006. **360**(4): p. 814-825.

120. Emanuelli, M., et al., *Identification and characterization of a second NMN adenylyltransferase gene in Saccharomyces cerevisiae.* Protein Expression and Purification, 2003. **27**(2): p. 357-364.

121. Saridakis, V., et al., *Insights into ligand binding and catalysis of a central step in NAD+ synthesis: structures of Methanobacterium thermoautotrophicum NMN adenylyltransferase complexes.* Journal of Biological Chemistry, 2001. **276**(10): p. 7225-32.

122. Wang, E., et al., *Structure and functional properties of the Bacillus subtilis transcriptional repressor Rex.* Molecular microbiology, 2008. **69**(2): p. 466-478.

123. Jeje, O., et al., *Obtaining high yield recombinant Enterococcus faecium nicotinate nucleotide adenylyltransferase for X-ray crystallography and biophysical studies.* International Journal of Biological Macromolecules, 2023. **250**: p. 126066.

124. Bras, N., et al., *Computational Biochemistry*. 2015.

125. Sivakumar, K.C., et al., *Prospects of multitarget drug designing strategies by linking molecular docking and molecular dynamics to explore the protein-ligand recognition process.* Drug Development Research, 2020. **81**(6): p. 685-699.

126. Rehman, A.U., et al., *Computational approaches for the design of modulators targeting protein-protein interactions.* Expert Opinion on Drug Discovery, 2023. **18**(3): p. 315-333.

127. Richard, E., et al., *Protein complex prediction with AlphaFold-multimer.* BioRxiv, 2022: p. 2021.10.04.463034.

128. Jumper, J., et al., *Highly accurate protein structure prediction with AlphaFold.* Nature, 2021. **596**(7873): p. 583-589.

129. Zheng, W., et al., *Protein structure prediction using deep learning distance and hydrogen-bonding restraints in CASP14.* Proteins, 2021. **89**(12): p. 1734-1751.

130. Thuluz Meza, M., J.-P. Claudia, and C.Z. Rossana, *Past, present, and future of molecular docking*, in *drug discovery and development*. IntechOpen, 2020. p. Ch. 2.

131. Boittier, E.D., et al., *Assessing molecular docking tools to guide targeted drug discovery of CD38 inhibitors.* International Journal of Molecular Sciences, 2020. **21**(15): p. 5183.

132. Sachin Patodia, A.B.a.D.C., *Molecular dynamics simulation of proteins: A brief overview.* Journal of Physical Chemistry and Biophysics, 2014. **Volume 4**( Issue 6 ): p. 4.

133. Tacconelli, E., et al., *Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis.* The Lancet Infectious Diseases, 2018. **18**(3): p. 318-327.

134. Mulani, M.S., et al., *Emerging strategies to combat ESKAPE pathogens in the era of antimicrobial resistance: A review.* Frontiers in Microbiology, 2019. **10**.

135. Rodionov, D.A., et al., *Transcriptional regulation of NAD metabolism in bacteria: Genomic reconstruction of NiaR (YrxA) regulon.* Nucleic Acids Research, 2008. **36**(6): p. 2032-2046.

136. Lau, C., M. Niere, and M. Ziegler, *The NMN/NaMN adenylyltransferase (NMNAT) protein family.* Frontiers in bioscience, 2009. **14**: p. 410-31.

137. Magni, G., et al., *NAD(P) biosynthesis enzymes as potential targets for selective drug design.* Current Medicinal Chemistry, 2009. **16**(11): p. 1372-90.

138. Pankiewicz, W.K., et al., *Nicotinamide adenine dinucleotide based therapeutics.* Current Medicinal Chemistry, 2008. **15**(7): p. 650-670.

139. Khan, J.A., et al., *Nicotinamide adenine dinucleotide metabolism as an attractive target for drug discovery.* Expert Opinion on Therapeutic Targets, 2007. **11**(5): p. 695-705.

140. Liang, L., et al., *Effects of overexpression of NAPRTase, NAMNAT, and NAD synthetase in the NAD(H) biosynthetic pathways on the NAD(H) pool, NADH/NAD+ ratio, and succinic acid production with different carbon sources by metabolically engineered Escherichia coli.* Biochemical Engineering Journal, 2013. **81**: p. 90-96.

141. Izard, T. and A. Geerlof, *The crystal structure of a novel bacterial adenylyltransferase reveals half of sites reactivity.* The EMBO Journal, 1999. **18**(8): p. 2021-2030.

142. O'Hara, J.K., et al., *Targeting NAD+ metabolism in the human malaria parasite Plasmodium falciparum.* PLOS ONE, 2014. **9**(4): p. e94061.

143. Zhang, X., et al., *Structural characterization of a human cytosolic NMN/NaMN adenylyltransferase and implication in human NAD biosynthesis.* Journal of Biological Chemistry, 2003. **278**(15): p. 13503-11.

144. Sershon, V.C., B.D. Santarsiero, and A.D. Mesecar, *Kinetic and X-ray structural evidence for negative cooperativity in substrate binding to nicotinate mononucleotide adenylyltransferase (NMAT) from Bacillus anthracis.* Journal of Molecular Biology, 2009. **385**(3): p. 867-888.

145. Olland, A.M., et al., *Identification, characterization, and crystal structure of Bacillus subtilis nicotinic acid mononucleotide adenylyltransferase.* The Journal of Biological Chemistry, 2002. **277**(5): p. 3698-3707.

146. Rodionova, I.A., et al., *Mycobacterial Nicotinate Mononucleotide Adenylyltransferase: Structure, mechanism, and implications for drug discovery*. Journal of Biological Chemistry, 2015. **290**(12): p. 7693-7706.

147. D'Angelo, I., et al., *Structure of nicotinamide mononucleotide adenylyltransferase: a key enzyme in NAD+ biosynthesis.* Structure, 2000. **8**(9). p. 993-1004.

148. Gasteiger, E., et al., *Protein identification and analysis tools on the ExPASy Server*, in *the proteomics protocols handbook*. Humana Press, 2005. p. 571-607.

149. Balducci, E., et al., *Assay methods for nicotinamide mononucleotide adenylyltransferase of wide applicability.* Analytical Biochemistry, 1995. **228**(1): p. 64-68.

150. Jeje, O., et al., *Effect of divalent metal ion on the structure, stability and function of Klebsiella pneumoniae nicotinate-nucleotide adenylyltransferase: empirical and computational studies.* International Journal of Molecular Sciences, 2022. **23**(1): p. 116.

151. Keegan, R.M. and M.D. Winn, *Automated search-model discovery and preparation for structure solution by molecular replacement.* Acta Crystallographica Section D, 2007. **63**(4): p. 447-457.

152. Winn, M.D., et al., *Overview of the CCP4 suite and current developments.* Acta Crystallographica Section D, 2011. **67**(4): p. 235-242.

153. J. Gildea, R., et al., *Embedding AI in the Protein crystallography workflow.* Synchrotron Radiation News, 2022. **35**(4): p. 51-54.

154. Niesen, F.H., H. Berglund, and M. Vedadi, *The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability.* Nature protocols, 2007. **2**(9): p. 2212-2221.

155. Gasymov, O.K. and B.J. Glasgow, *ANS fluorescence: potential to augment the identification of the external binding sites of proteins.* Biochimica et Biophysica acta, 2007. **1774**(3): p. 403-411.

156. Yan, R., et al., *A comparative assessment and analysis of 20 representative sequence alignment methods for protein structure prediction.* Scientific Reports, 2013. **3**: p. 2619.

157. Emanuelli, M., et al., *Molecular cloning, chromosomal localization, tissue mRNA levels, bacterial expression, and enzymatic properties of human NMN adenylyltransferase*. Journal of Biological Chemistry, 2001. **276**(1): p. 406-12.

158. Raffaelli, N., et al., *Characterization of nicotinamide mononucleotide adenylyltransferase from thermophilic archaea.* Journal of Bacteriology, 1997. **179**(24): p. 7718.

159. Raffaelli, N., et al., *Synechocystis sp. slr0787 protein is a novel bifunctional enzyme endowed with both nicotinamide mononucleotide adenylyltransferase and 'Nudix' hydrolase activities.* FEBS Letters, 1999. **444**(2-3): p. 222-226.

160. Hughes, K.T., et al., *An indispensable gene for NAD biosynthesis in Salmonella typhimurium.* Journal of Bacteriology, 1983. **155**(1): p. 213-221.

161. Mehl, R.A., C. Kinsland, and T.P. Begley, I*dentification of the Escherichia coli nicotinic acid mononucleotide adenylyltransferase gene.* Journal of Bacteriology, 2000. **182**(15): p. 4372.

162. *Enzymology of Nad+ Synthesis*, in *Advances*. Enzymology and Related Areas of Molecular Biology. p. 135-182.

163. Daya, T., et al., *Expression, Purification, and biophysical characterization of Klebsiella pneumoniae nicotinate nucleotide adenylyltransferase.* The Protein Journal, 2022. **41**(1): p. 141-156.

164. Majorek, K.A., et al., *Double trouble-buffer selection and His-tag presence may be responsible for nonreproducibility of biomedical experiments.* Protein Science, 2014. **23**(10): p. 1359-68.

165. Rossmann, M., et al., *The Enzymes.* Academic Press, 1975. *Vol. 11*.

166. Yoon, H.-J., et al., *Crystal structure of nicotinic acid mononucleotide adenylyltransferase from Pseudomonas aeruginosa in its apo and substrate-complexed forms reveals a fully open conformation.* Journal of Molecular Biology, 2005. **351**(2): p. 258-265.

167. Mei, G., et al., *The importance of being dimeric.* The FEBS Journal, 2005. **272**(1): p. 16-27.

168. Laskowski, R.A., et al., *PDBsum: Structural summaries of PDB entries.* Protein Sci, 2018. **27**(1): p. 129-134.

169. Karplus, P.A. and K. Diederichs, *Assessing and maximizing data quality in macromolecular crystallography.* Current Opinion in Structural Biology, 2015. **34**: p. 60-8.

170. Weiss, M., *Global indicators of X-ray data quality.* Journal of Applied Crystallography, 2001. **34**(2): p. 130-135.

171. Tungmunnithum, D., et al., *Flavonoids and other phenolic compounds from medicinal plants for pharmaceutical and medical aspects: An overview.* Medicines, 2018. **5**(3): p. 93.

172. Maleki, S.J., J.F. Crespo, and B. Cabanillas, *Anti-inflammatory effects of flavonoids.* Food Chemistry, 2019. **299**: p. 125124.

173. Mutha, R.E., A.U. Tatiya, and S.J. Surana, *Flavonoids as natural phenolic compounds and their role in therapeutics: an overview.* Future Journal of Pharmaceutical Sciences, 2021. **7**(1): p. 25.

174. Roy, A., et al., *Flavonoids a bioactive compound from medicinal plants and its therapeutic applications.* BioMed Research International, 2022. **2022**: p. 5445291.

175. Chang, C.-H., et al., *Resveratrol-induced autophagy and apoptosis in cisplatin-resistant human oral cancer CAR cells: A key role of AMPK and Akt/mTOR signaling.* International Journal of Oncology, 2017. **50**(3): p. 873-882.

176. Gao, F., et al., *Resveratrol suppresses human hepatocellular carcinoma via targeting HGF-c-Met signaling pathway.* Oncology Reports, 2017. **37**(2): p. 1203-1211.

177. Sun, Y., et al., *Curcumin inhibits the proliferation and invasion of MG-63 cells through inactivation of the p-JAK2/p-STAT3 pathway.* Onco Targets Therapy, 2019. **12**: p. 2011-2021.

178. Peng, H.L., et al., *Fisetin inhibits the generation of inflammatory mediators in interleukin-1β–induced human lung epithelial cells by suppressing the NF-κB and ERK1/2 pathways.* International Immunopharmacology, 2018. **60**: p. 202-210.

179. Kumar, R., et al., *Exploring the binding mechanism of flavonoid quercetin to phospholipase A2: fluorescence spectroscopy and computational approach.* European Journal of Experimental Biology, 2017. **7**: p. 33.

180. Guo, Y.Q., et al., *Prenylated flavonoids as potent phosphodiesterase-4 inhibitors from Morus alba: Isolation, modification, and structure-activity relationship study.* European Journal of Medicinal Chemistry, 2018. **144**: p. 758-766.

181. González Mosquera, D.M., et al., *Flavonoids from Boldoa purpurascens inhibit proinflammatory cytokines (TNF-α and IL-6) and the expression of COX-2.* Phytotherapy Research, 2018. **32**(9): p. 1750-1754.

182. Khan, M.T., et al., *Immunoinformatics and molecular modeling approach to design universal multi-epitope vaccine for SARS-CoV-2.* Informatics in Medicine Unlocked, 2021. **24**: p. 100578.

183. Valli, A. and I. Achilonu, *Molecular dynamics-derived pharmacophores of Schistosoma glutathione transferase in complex with bromosulfophthalein: Screening and analysis of potential inhibitors.* Journal of Molecular Graphics and Modelling, 2023. **122**: p. 108457.

184. Parihar, A., et al., *Phytochemicals-based targeting RdRp and main protease of SARS-CoV-2 using docking and steered molecular dynamic simulation: A promising therapeutic approach for Tackling COVID-19.* Computers in Biology and Medicine, 2022. **145**: p. 105468.

185. Parihar, A., et al., *Plant-based bioactive molecules for targeting of endoribonuclease using steered molecular dynamic simulation approach: a highly conserved therapeutic target against variants of SARS-CoV-2.* Molecular Simulation, 2023. **49**(12): p. 1267-1279.

186. Parihar, A., R. Shrivastava, and A. Dube, *Interaction of Cp6-his and Cp6 with bovine serum albumin and liver microsomes: Spectroscopic and molecular docking studies.* Journal of Photochemistry and Photobiology, 2021. **5**: p. 100013.

187. Clyde, A., et al., *High-throughput virtual screening and validation of a SARS-CoV-2 main protease noncovalent inhibitor.* Journal of Chemical Information and Modeling, 2022. **62**(1): p. 116-128.

188. Kim, S., et al., *PubChem in 2021: new data content and improved web interfaces.* Nucleic Acids Research, 2020. **49**(D1): p. D1388-D1395.

189. Madhavi Sastry, G., et al., *Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments.* Journal of Computer-Aided Molecular Design, 2013. **27**(3): p. 221-234.

190. Shelley, J.C., et al., *Epik: a software program for pKaprediction and protonation state generation for drug-like molecules.* Journal of Computer-Aided Molecular Design, 2007. **21**(12): p. 681-691.

191. Onufriev, A.V. and E. Alexov, *Protonation and pK changes in protein-ligand binding.* Q Rev Biophys, 2013. **46**(2): p. 181-209.

192. Roos, K., et al., *OPLS3e: Extending force field coverage for drug-like small molecules.* Journal of Chemical Theory and Computation, 2019. **15**(3): p. 1863-1874.

193. Harder, E., et al., *OPLS3: A force field providing broad coverage of drug-like small molecules and proteins.* Journal of Chemical Theory and Computation, 2016. **12**(1): p. 281-296.

194. Klepeis, J.L., et al., *Long-timescale molecular dynamics simulations of protein structure and function.* Current Opinion in Structural Biology, 2009. **19**(2): p. 120-127.

195. Halgren, T., *New method for fast and accurate binding-site identification and analysis.* Chemical Biology and Drug Design, 2007. **69**(2): p. 146-148.

196. Halgren, T.A., et al., *Glide: A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening.* Journal of Medicinal Chemistry, 2004. **47**(7): p. 1750-9.

197. Friesner, R.A., et al., *Glide: A New approach for rapid, accurate docking and scoring. 1. method and assessment of docking accuracy.* Journal of Medicinal Chemistry, 2004. **47**(7): p. 1739-1749.

198. Friesner, R.A., et al., *Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein−ligand complexes.* Journal of Medicinal Chemistry, 2006. **49**(21): p. 6177-6196.

199. Joseph, M.H. and A. Georgios, *MM-GB(PB)SA calculations of protein-ligand binding free energies*, in *molecular dynamics*. IntechOpen, 2012. p. Ch. 9.

200. Daina, A., O. Michielin, and V. Zoete, *SwissTargetPrediction: updated data and new features for efficient prediction of protein targets of small molecules.* Nucleic Acids Research, 2019. **47**(W1): p. W357-W364.

201. Daina, A., O. Michielin, and V. Zoete, *SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules.* Scientific Reports, 2017. **7**(1): p. 42717.

202. Stierand, K., P.C. Maaß, and M. Rarey, *Molecular complexes at a glance: automated generation of two-dimensional complex diagrams.* Bioinformatics, 2006. **22**(14): p. 1710-1716.

203. Lobanov, M.Y., N.S. Bogatyreva, and O.V. Galzitskaya, *Radius of gyration as an indicator of protein structure compactness.* Molecular Biology, 2008. **42**(4): p. 623-628.

204. Ritchie, T.J., P. Ertl, and R. Lewis, *The graphical representation of ADME-related molecule properties for medicinal chemists.* Drug Discovery Today, 2011. **16**(1): p. 65-72.

205. Lovering, F., J. Bikker, and C. Humblet, *Escape from flatland: increasing saturation as an approach to improving clinical success.* Journal of Medicinal Chemistry, 2009. **52**(21): p. 6752-6.

206. Huang, S.-M., et al., *New era in drug interaction evaluation: us food and drug administration update on cyp enzymes, transporters, and the guidance process.* The Journal of Clinical Pharmacology, 2008. **48**(6): p. 662-670.

207. Di, L., *The role of drug metabolizing enzymes in clearance.* Expert Opinion on Drug Metabolism and Toxicology, 2014. **10**(3): p. 379-93.

208. Baell, J.B. and G.A. Holloway, *New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays.* Journal of Medicinal Chemistry, 2010. **53**(7): p. 2719-40.

209. Liu, W.-N. and X.-H. Zhao, *Changes of the stability and bioactivity of quercetin and myricetin in BGC-823 cells in response to heat treatment and Fe2+/Cu2+ addition.* Journal of Food Measurement and Characterization, 2019. **13**(4): p. 3285-3297.

210. Gao, Y., et al., *Impact of thermal processing on dietary flavonoids.* Current Opinion in Food Science, 2022. **48**: p. 100915.

211. Ioannou, I., L. Chekir, and M. Ghoul, *Effect of the processing temperature on the degradation of food flavonoids: Kinetic and calorimetric studies on model solutions.* Journal of Food Engineering and Technology, 2019. **8**(2): p. 91-102.

212. Chaaban, H., et al., *Effect of heat processing on thermal stability and antioxidant activity of six flavonoids.* Journal of Food Processing and Preservation, 2017. **41**(5): p. e13203.

213. Ladbury, J.E., *Just add water! The effect of water on the specificity of protein-ligand binding sites and its potential application to drug design.* Chemistry and Biology, 1996. **3**(12): p. 973-980.

214. Maurer, M. and C. Oostenbrink, *Water in protein hydration and ligand recognition.* Journal of Molecular Recognition, 2019. **32**(12): p. e2810.

215. Connelly, P.R., et al., *Enthalpy of hydrogen bond formation in a protein-ligand binding reaction.* Proceedings of the National Academy of Sciences of the United States of America, 1994. **91**(5): p. 1964-1968.

216. Lipinski, C.A., et al., *Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings1PII of original article: S0169-409X(96)00423-1. The article was originally published in Advanced Drug Delivery Reviews 23 (1997) 3–25.1.* Advanced Drug Delivery Reviews, 2001. **46**(1): p. 3-26.

217. Baell, J.B., *Feeling Nature's PAINS: Natural products, natural product drugs, and pan assay interference compounds (PAINS).* Journal of Natural Products, 2016. **79**(3): p. 616-628.

# Appendices

## APPENDIX A

International Journal of Biological Macromolecules 250 (2023) 126066

Contents lists available at ScienceDirect

International Journal of Biological Macromolecules

journal homepage: www.elsevier.com/locate/ijbiomac

# Obtaining high yield recombinant *Enterococcus faecium* nicotinate nucleotide adenylyltransferase for X-ray crystallography and biophysical studies

Olamide Jeje, Ramesh Pandian, Yasien Sayed, Ikechukwu Achilonu *

*Protein Structure-Function Research Unit, School of Molecular and Cell Biology, Faculty of Science, University of the Witwatersrand, Johannesburg 2050, South Africa*

ARTICLE INFO

*Keywords:*
Nicotinate nucleotide adenylyltransferase (NNAT)
Nicotinamide mononucleotide adenylyltransferase (NMNAT)
*Enterococcus faecium*
Expression
Purification
X-ray crystallography

ABSTRACT

Nicotinate nucleotide adenylyltransferase (NNAT) has been a significant research focus on druggable targets, given its indispensability in the biosynthesis of $NAD^+$, which is crucial to the survival of bacterial pathogens. However, no information is available on the structure-function of *Enterococcus faecium* NNAT (EfNNAT). This study established the expression and purification protocol for obtaining a high-yield recombinant EfNNAT using the *E. coli* expression system and a single-step IMAC purification method. Approximately 101 mg of EfNNAT was obtained per 7.8 g of wet *E. coli* cells, estimated to be over 98 % pure. We further characterized the biophysical structure and determined the three-dimensional structure of the EfNNAT. Biophysical studies revealed a dimeric protein with a higher α-helical composition. The highly stable protein crystalizes in multiple conditions, yielding high-quality crystals diffracting between 1.78 and 2.80 Å. Two high-resolution crystal structures of EfNNAT in its native and adenine-bound forms were determined at 1.90 Å and 1.82 Å, respectively. The X-ray structures of the EfNNAT revealed the presence of phosphate and sulfate ions occupying and interacting with conserved amino acid residues within the putative substrate binding site, hence providing insight into the probable substrate preference of EfNNAT and, consequently, why EfNNAT may not prefer β-nicotinamide mononucleotide as a substrate. With the accessibility to high-resolution structures of EfNNAT, further structural evaluation and drug-based screening can be achieved. Hence, we anticipate that this study will provide the basis for the discovery of structure-based inhibitors against this enzyme.

### 1. Introduction

The gram-positive *Enterococcus faecium* is a bacterial pathogen notorious for resisting most known antibiotics. It has been identified alongside other virulent and multidrug-resistant bacterial pathogens nick-named ESKAPE (*Enterococcus faecium*, *Staphylococcus aureus*, *Kleb-*

urgent need of new antibiotics [3]. Yet, only a few new antibiotics have been developed since 2010 [4].

The biosynthesis of nicotinamide adenine dinucleotide ($NAD^+$) and the enzymes involved in its downstream formation has been a fore-ground for drug target in recent times [5–13] because of the impact and essentiality of $NAD^+$ to cell survival. The vitalness of $NAD^+$ and its de-

**APPENDIX B**

Biochimie 220 (2024) 84–98

Contents lists available at ScienceDirect

Biochimie

journal homepage: www.elsevier.com/locate/biochi

# Exploring NAD$^+$ metabolism and NNAT: Insights from structure, function, and computational modeling

Olamide Jeje [a], Sarah Otun [a, *], Chinyere Aloke [a, b], Ikechukwu Achilonu [a]

[a] Protein Structure-Function and Research Unit, School of Molecular and Cell Biology, Faculty of Science, University of the Witwatersrand, Braamfontein, Johannesburg, 2050, South Africa

[b] Department of Medical Biochemistry, Alex Ekwueme Federal University Ndufu-Alike, Ebonyi State, Nigeria

ARTICLE INFO

*Article history:*
Received 29 October 2023
Received in revised form
18 December 2023
Accepted 2 January 2024
Available online 3 January 2024

Handling Editor: Prof. J.M. Gimble

*Keywords:*
NAD$^+$ metabolism
Nicotinamide/nicotinate nucleotide adenylyltransferase
Structural insights
Computational modelling
Enzyme function
Therapeutic implications

ABSTRACT

Nicotinamide Adenine Dinucleotide (NAD+), a coenzyme, is ubiquitously distributed and serves crucial functions in diverse biological processes, encompassing redox reactions, energy metabolism, and cellular signalling. This review article explores the intricate realm of NAD + metabolism, with a particular emphasis on the complex relationship between its structure, function, and the pivotal enzyme, Nicotinate Nucleotide Adenylyltransferase (NNAT), also known as nicotinate mononucleotide adenylyltransferase (NaMNAT), in the process of its biosynthesis. Our findings indicate that NAD + biosynthesis in humans and bacteria occurs via the same de novo synthesis route and the pyridine ring salvage pathway. Maintaining NAD homeostasis in bacteria is imperative, as most bacterial species cannot get NAD+ from their surroundings. However, due to lower sequence identity and structurally distant relationship of bacteria, including *E. faecium* and *K. pneumonia*, to its human counterpart, inhibiting NNAT, an indispensable enzyme implicated in NAD + biosynthesis, is a viable alternative in curtailing infections orchestrated by *E. faecium* and *K. pneumonia*. By merging empirical and computational discoveries and connecting the intricate NAD + metabolism network with NNAT's crucial role, it becomes clear that the synergistic effect of these insights may lead to a more profound understanding of the coenzyme's function and its potential applications in the fields of therapeutics and biotechnology.

© 2024 Published by Elsevier B.V.

**Contents**

Journal of Molecular Structure

# Exploring Flavonoids as Potential Inhibitors for Enterococcus faecium Nicotinate Nucleotide Adenylyltransferase: An Integration of In Silico with Empirical Studies

--Manuscript Draft--

| | |
|---|---|
| **Manuscript Number:** | MOLSTRUC-D-24-02722 |
| **Article Type:** | Research Paper |
| **Keywords:** | Nicotinate nucleotide adenylyltransferase (NNAT); nicotinamide mononucleotide adenylyltransferase (NMNAT); Enterococcus faecium; ESKAPE pathogens; flavonoids; MD Simulation |
| **Corresponding Author:** | Ikechukwu Achilonu, Ph.D<br>University of the Witwatersrand Johannesburg<br>Johannesburg, Gauteng SOUTH AFRICA |
| **First Author:** | Olamide Jeje, MSc |
| **Order of Authors:** | Olamide Jeje, MSc |
| | Ikechukwu Achilonu, Ph.D |
| **Abstract:** | Nicotinate nucleotide adenylyltransferase (NNAT) emerges as a promising target for drug development, given its pivotal role in the nicotinate nucleotide biosynthesis pathway crucial for bacterial survival. This study screened a library of 472 compounds from MedChemExpress as potential inhibitors against Enterococcus faecium NNAT (EfNNAT) using computational modelling and fluorescence methods. We employed Maestro's Glide HTVS, SP, and XP docking protocols to identify compounds with high binding affinity. The screening utilized two models of the EfNNAT crystal structure: energy-minimized and 500 ns MD-simulated model, hypothesizing conformational change that may influence ligand selectivity in HTVS study. Our findings revealed variations between the minimized versus MD-simulated models, impacting ligand selectivity by each model post-HTVS. However, quercetin 3-O-beta-D-glucose-7-O-beta-D-gentiobioside (QGG) appeared to be the top inhibitor selected by both models after more rigorous docking using Maestro-implemented SP and XP. A 500 ns all atom MD simulation of the EfNNAT:ligand complex revealed that the compound interacts mainly through hydrogen bonding and water bridges. Validation of the in-silico studies via molecular spectroscopy confirmed QGG's binding to EfNNAT by displacing the fluorescent probes from the protein binding site. However, thermal shift assay demonstrated minimal impact on protein stability and structure upon QGG binding. Insights gleaned from this study offer a mechanistic basis that could be leveraged to develop new leads. Expansion of the compound library holds promise for identifying inhibitors with enhanced affinity for EfNNAT. |
| **Suggested Reviewers:** | Pritish Kumar Varadwaj, PhD<br>Researcher, Department of Bioinformatics, Indian Institute of Information Technology-Allahabad, CC2-4203, Jhalwa Campus, Deoghat, Allahabad-211012, Uttar Pradesh, India<br>pritish@iiita.ac.in<br>The Reviewer recently published an article titled " Molecular docking and dynamics simulation study of flavonoids as BET bromodomain inhibitors", in which flavonoids were used as potential inhibitors. This closely aligns with our study. |
| | Ruchi Malik, PhD<br>Researcher<br>ruchimalik1976@curaj.ac.in<br>The Reviewer recently published an article titled " Discovery of Benzoxazepines as a New Class of PIM1 Kinase Inhibitors Through Structure Based Virtual Screening, Biochemical Evaluation and Cytotoxicity Studies", in which flavonoids were used as potential inhibitors. This closely aligns with our study. |
| | Arpana Parihar, PhD<br>Researcher<br>arpana_parihar@yahoo.com<br>The Reviewer recently published an article titled " Plant-based bioactive molecules for targeting of endoribonuclease using steered molecular dynamic simulation approach: a highly conserved therapeutic target against variants of SARS-CoV-2", in which flavonoids were used as potential inhibitors. This closely aligns with our study. |

*Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation*