Investigating the interaction of *Escherichia coli* stringent starvation protein A (EcSspA) with indanyloxyacetic acid 94 and ethacrynic acid.

by

Rethabile Mokoena

1076599

Dissertation

Submitted in fulfilment of the requirements for the degree

Master of Science

in

Molecular and Cell Biology

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

Supervisor: Professor H.W. Dirr

April 2020

Declaration:

I, **Rethabile Mokoena (1076599)**, am a student registered for the degree of **Master of Science (Dissertation) in Biochemistry** in the academic year **2019**.

I hereby declare the following:

- I am aware that plagiarism (the use of someone else's work without their permission and/or without acknowledging the original source) is wrong.
- I confirm that the research proposal submitted for assessment for the above degree is my own unaided work except where explicitly indicated otherwise and acknowledged.
- I have followed the required conventions in referencing the thoughts and ideas of others.
- I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.

Signature:  day of **April 2020**

Abstract:

Stringent starvation protein A (SspA) is a highly conserved protein found in Gram-negative bacteria, such as *Escherichia coli* (*E. coli*). SspA plays a critical role in the activation of defence systems in bacteria, such as acid tolerance, to aid in the survival of Gram-negative bacteria exposed to harmful conditions. *E. coli* SspA (EcSspA) is a bacterial homologue of human chloride intracellular channel 1 (hCLIC1) and both are members of the glutathione transferase (GST) superfamily. EcSspA and hCLIC1 exhibit similarities in structure and ion channel formation. Ion channel activity of membranous hCLIC1 and EcSspA is inhibited by indanyloxyacetic acid 94 (IAA-94) as well as the enzymatic activity of soluble hCLIC1, however, the effect that IAA-94 has on soluble EcSspA is unknown. Ethacrynic acid (EA), a structural homologue of IAA-94, is a known inhibitor of cytosolic GST enzymatic activity and its effect on EcSspA is also unknown. The aim of this study was to investigate the interaction of soluble EcSspA with IAA-94 and EA using 8-anilino-1-naphthalenesulfonic acid (ANS) displacement, isothermal titration calorimetry (ITC) and induced-fit molecular docking. EcSspA was purified to homogeneity using nickel ion immobilised metal affinity chromatography (Ni^{2+}-IMAC). Structural analysis of soluble EcSspA using far-UV circular dichroism, intrinsic tryptophan fluorescence spectroscopy as well as size exclusion high performance liquid chromatography, confirmed that soluble EcSspA has a predominantly α-helical backbone, with tryptophan residues that are partially solvent accessible and that it is a dimeric protein. Data obtained from ITC, ANS displacement and induced-fit molecular docking experiments, indicated that there was weak binding between EcSspA and both ligands at the dimer interface. The data further showed no interaction occurred at the EcSspA surface pocket. Therefore indicating that the dimer interface is a potential target site on soluble EcSspA. However, it remains unknown how the binding of IAA-94 and EA to soluble EcSspA would influence EcSspA function.

To my parents, family, friends and colleagues, ke leboha homenahane.

“Trust the process and embrace the journey” - *Unknown*

Acknowledgements:

I would like to thank my supervisor, Professor Dirr for his guidance, encouragement and for sharing his wealth of knowledge on the subject matter and on matters of life. I appreciate the opportunity he awarded me to be a part of the Protein Structure-Function Research Unit and to learn from all the members of the unit.

I would also like to thank my colleagues at the PSFRU, especially Roland Worth, Monare Thulo and Vinuchakkaravarthy Thangaraj, for all their help, support and encouragement.

I am grateful to the CSIR for the funding they awarded me to complete my master's degree.

agriculture,
forestry & fisheries

Department:
Agriculture, Forestry and Fisheries
REPUBLIC OF SOUTH AFRICA

Genetic Resources, Department of Agriculture, Forestry and Fisheries
Private Bag X973, Pretoria 0001

Enquiries: Bathobile Mahlangu • **Tel**: 012 319 6165 • **Fax:** 012 319 6298 • **E-mail**: BathobileM@daff.gov.za •**Ref**: 39.2/University of Witwatersrand -18/001

Prof H. W Dirr
University of Witwatersrand
School of Molecular and Cell Biology
East Campus
Gate House Room 431
Johannesburg
2050

+27 (0) 11 717 6352 **(Tel)**
+27 (0) 86 553 5708 **(Fax)**
Heinrich.dirr@wits.ac.za **(email)**

Dear Prof Dirr

RE: REGISTRATION OF FACILITY

With reference to the application to register a facility, submitted in terms of the Genetically Modified Organisms Act, 1997 (Act No. 15 of 1997). Registration number **39.2/University of Witwatersrand -18/103.**

The facility is hereby registered; please find attached your certificate which serves as proof of registration. Please familiarize yourself with the Standard Operating Procedure approved for Regulation 2(2) to determine whether your current activities or any future activities would require an additional contained use permit or not.

Please consult the website of the Department at www.daff.gov.za (Branches, Agricultural Production, Health & Food Safety/ Genetic Resources/ Biosafety) for the latest application forms and the SOP document referred to in the above paragraph.

If any of the provisions of the Genetically Modified Organisms Act, 1997 (Act No. 15 of 1997), including any condition of any permit issued in terms of the GMO Act, is not complied with at all times, you will be subject to prosecution in terms of Section 21 of the GMO Act, 1997.

Yours sincerely

DEPT. AGRICULTURE, FORESTRY & FISHERIES
DIRECTORATE GENETIC RESOURCES
2018 -01- 31
PRIVATE BAG X973 PRETORIA 0001
REGISTRAR - GMO ACT

Ms N L Mkhonza
Registrar: Genetically Modified Organisms Act, 1997 (Act No. 15 of 1997)

agriculture,
forestry & fisheries

Department:
Agriculture, Forestry and Fisheries
REPUBLIC OF SOUTH AFRICA

CERTIFICATE OF REGISTRATION

Registration Number: **39.2/UNIVERSITY OF WITWATERSRAND -18/103**

It is hereby certified that the facility situated at:

UNIVERSITY OF WITWATERSRAND
SCHOOL OF MOLECULAR AND CELL BIOLOGY
EAST CAMPUS
GATE HOUSE ROOM 431
JOHANNESBURG
2050

has been registered in the name(s) of: **PROF H. W DIRR**

on behalf of: **UNIVERSITY OF WITWATERSRAND**

in terms of Regulation 8 of the Genetically Modified Organisms Act, 1997 (Act No. 15 of 1997) in respect of the following-

type of facility: **LABORATORY**

containment level: **ONE (1)**

for period: **31 JANUARY 2018 – 31 JANUARY 2021**

Date issued: **31 JANUARY 2018**

DEPT. AGRICULTURE, FORESTRY & FISHERIES
DIRECTORATE GENETIC RESOURCES
2018 -01- 31
PRIVATE BAG X973 PRETORIA 0001
REGISTRAR - GMO ACT

Registrar for Genetically Modified Organisms

Date

UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG

Research Office

INSTITUTIONAL BIOSAFETY COMMITTEE
(R 14/16)

CLEARANCE CERTIFICATE **PROTOCOL NUMBER: 20180105Lab**

BRIEF DESCRIPTION OF APPLICATION:

Laboratory situated in room 411-434 URC Proten Structure-Fucntion Research Unit Gate House 4th Floor

APPLICANT: Prof Heini Dirr et al

SCHOOL/DEPARTMENT : Molecular and Cell Biology

DATE CONSIDERED: 29 January 2018

DECISION OF COMMITTEE: Approved unconditionally

These laboratories are considered to meet the standards of the Institutional Biosafety Committee (IBC), for BSL1 approval, as specified in the IBC's published Standard Operating procedures.
Any change to the type of biohazardous agents being handled should be reported to the IBC
DAFF- Reg no 39.2/University of the Witwatersrand-18/103 expiry date 31 January 2018 - 31 January 2021

1. This clearance certificate expires on 12 March 2023 and may be renewed on application.

DATE OF APPROVAL: 12 March 2018 **CHAIRPERSON:** Professor J Larkin

Radiation and Health Physics Unit
University of the Witwatersrand
Johannesburg, Republic of South Africa
Tel: +27 (0) 11 717 6631

DECLARATION OF APPLICANT:

To be completed in duplicate and **one copy** returned to the University of the Witwatersrand, Faculty of Health Sciences, Research Office, Office 301, Phillip Tobias Building, 29 Princess of Wales Terrace, Parktown,2193.

1. I have read, understood and accepted the approval conditions above
2. I note that the University Safety Officer, or his/her representative, may at any reasonable time inspect my laboratory or trial site to ensure compliance with current Health and Safety legislation. I undertake to offer my full co-operation in any such inspection.
3. I have read, understood and will comply with the *recommended standard operating procedures for the handling of biohazardous materials* posted at http://intranet.wits.ac.za/academic/uro/Pages/Institutional-Biosafety-Committee-(IBC).aspx

Signed: _____ **Date:** _____

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

Table of Contents

List of Figures i

List of Tables iii

List of Abbreviations iv

Chapter 1: Introduction 1

1.1 Survival of bacteria in varying environments 1
1.2 The role of SspA in *E. coli* 1
1.3 SspA orthologues 5
1.4 EcSspA structure 6
1.5 EcSspA surface pocket 9
1.6 EcSspA is a bacterial homologue of human CLIC1 12
1.7 IAA-94 and Ethacrynic acid (EA) potentially inhibit EcSspA function 13
1.8 Aim 16
1.9 Objectives 16

Chapter 2: Experimental Procedures 17

2.1 Materials 17
2.2 Methodology 17
2.2.1 pET-28a vector 17
2.2.2 Verification of cDNA sequence encoding EcSspA 17
2.2.3 Overexpression of EcSspA 18
2.2.4 Purification of EcSspA 19
2.2.5 Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE) 20
2.2.6 Protein concentration determination and purity assessment 21
2.2.7 Structural characterisation of EcSspA 22
2.2.8 Assessment of IAA-94 and EA interaction with EcSspA 25

Chapter 3: Results 34

3.1 Verification of cDNA sequence encoding EcSspA 34
3.2 Overexpression and purification of EcSspA 34
3.2.1 Overexpression of EcSspA 34
3.2.2 Purification of EcSspA 34
3.2.3 Verification of the purity of EcSspA 34
3.3 Protein concentration determination and purity assessment 39
3.4 Structural Characterisation of EcSspA 41

3.4.1 Far-UV CD spectropolarimetry 41
3.4.2 Near-UV CD spectropolarimetry 44
3.4.3 Intrinsic tryptophan fluorescence spectroscopy 44
3.4.4 SE-HPLC 44
3.5 Assessment of IAA-94 and EA interaction with EcSspA 48
3.5.1 ANS concentration determination and ANS extrinsic fluorescence 48
3.5.2 Displacement of ANS by IAA-94 and EA from EcSspA using fluorescence spectroscopy 51
3.5.3 ITC 53
3.5.4 Molecular docking of IAA-94 and EA to EcSspA 56
Chapter 4: Discussion 63
4.1 The overexpression and purification of EcSspA 63
4.2 The structural characterisation of EcSspA 64
4.3 Interaction of IAA-94 and EA with EcSspA 67
Conclusion: 71
Chapter 5: References 71

List of Figures:

Figure 1: Schematic of GDAR system in *E. coli* bacterium……**4**

Figure 2: Cartoon representation of GST B1-1 and YpSspA……**8**

Figure 3: Surface pocket of EcSspA modelled after YpSspA (PDB code: 1YY7) ……...**11**

Figure 4: Chemical structure of IAA-94 and EA represented in 2-D……...**15**

Figure 5: Schematic diagram of raw data obtained from an ITC experiment and accompanying isotherm……**28**

Figure 6: Simulated binding isotherm for various values of c……...**30**

Figure 7: Verification of the cDNA sequence encoding for EcSspA……...**35**

Figure 8: Growth curve of T7 *E. coli* cells……**36**

Figure 9: The purification elution profile of EcSspA……...**37**

Figure 10: Verification of the presence and purity of EcSspA……**38**

Figure 11: EcSspA concentration determination and purity assessment……...**40**

Figure 12: Far-UV CD and HT spectra of EcSspA……...**42**

Figure 13: Near-UV CD, absorbance and HT spectra of EcSspA……**45**

Figure 14: Tryptophan fluorescence emission spectrum of EcSspA……**46**

Figure 15: SE-HPLC elution profile of EcSspA and calibration curve……**47**

Figure 16: ANS concentration determination……...**49**

Figure 17: Fluorescence emission spectrum of ANS complexed with EcSspA……...**50**

Figure 18: Fluorescence emission spectrum of the displacement of ANS by IAA-94 and EA from EcSspA……**52**

Figure 19: ITC titrations of IAA-94 and EA into EcSspA……**54**

Figure 20: The homology model of EcSspA and the local quality estimate plot of the EcSspA model……**57**

Figure 21: The multiple sequence alignment of the EcSspA primary sequence against top two ranked template sequences and the superimposition of the template structure and the EcSspA model……**58**

Figure 22: The four highly ranked binding sites predicted by SiteMap……**60**

Figure 23: Docking of IAA-94 to EcSspA homology model……**61**

Figure 24: Docking of EA to EcSspA homology model……**62**

Figure 25: The location of tryptophan residues on the molecular surface of the EcSspA structure……**66**

List of Tables:

Table 1: The total α-helical, 3_{10}-helical and β-strand content of YpSspA and GSTB1-1 based on number of amino acid residues present in each secondary structural element using the Full Protein Feature View from RCSB PDB……**43**

Table 2: Thermodynamic parameters obtained from the titration of 2.1 mM IAA-94 and 2.48 mM EA into 100 µM EcSspA at 20˚C……**55**

List of Abbreviations:

2xYT	2 x Yeast and Tryptone
ANS	8-Anilino-1-naphthalenesulfonic acid ammonium salt
A.U.	Arbitrary units
BLAST	Basic Local Alignment Search Tool
CD	Circular Dichroism
CLIC	Chloride Intracellular Channel protein
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
EA	Ethacrynic acid
E. coli	*Escherichia coli*
EcSspA	*Escherichia coli* stringent starvation protein A
EDTA	Ethylenediaminetetraacetic acid
GABA	γ-aminobutyric acid
GSH	Glutathione
GST	Glutathione transferase
GST B1-1	beta-class GST subtype 1
hCLIC1	human Chloride Intracellular Channel protein 1
H-NS	Histone-like nucleoid structuring protein
IAA-94	Indanyloxyacetic acid-94
IMAC	Immobilised metal-ion affinity chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
kDa	kilo Dalton

NCBI	National Center for Biotechnology Information
OD_{600}	Optical density at 600 nm
PAGE	Polyacrylamide gel electrophoresis
PMSF	Phenylmethylsulfonyl fluoride
RNA	Ribonucleic acid
rpm	Revolutions per minute
SE-HPLC	Size Exclusion-High Performance Liquid Chromatography
SOC	Super Optimal broth with Catabolite repression
SDS	Sodium Dodecyl Sulfate
SspA	Stringent starvation protein A
TEMED	Tetramethylethylenediamine
Tris	tris(hydroxymethyl)aminomethane
UV-Vis	Ultraviolet-Visible
UV	Ultraviolet
YpSspA	*Yersinia pestis* stringent starvation protein A

The IUPAC one- and three-letter abbreviations for the 20 standard amino acids was used.

Chapter 1: Introduction

1.1 Survival of bacteria in varying environments

Bacteria are microorganisms that inhabit and thrive in varied environments, such as extreme temperatures (as high as 122˚C and as low as -20˚C), unusual chemical conditions or even anaerobic environments, all of which are due to their highly adaptable metabolism (Rampelotto, 2013). Bacteria have evolved mechanisms that allow them to survive and replicate in their preferred environments (Mekalanos, 1992). However, these environments can become hostile and unfavourable for bacteria and as a result, bacteria use their defence mechanisms to withstand changing environmental conditions (Murphy *et al*., 2006).

Bacteria possess genes that encode for proteins that are responsible for the optimal function of their defence mechanisms and ensure that the necessary proteins are expressed when challenged by the environment (Gao *et al*., 2011). For example, to combat environmental stress, some bacteria resort to minimising metabolic activity and lie dormant (Voet and Voet, 2011). However, bacteria like *Escherichia coli* (*E. coli*) for instance, express redox-sensitive transcriptional activator (SoxR) and hydrogen peroxide-inducible gene activator (OxyR) proteins to regulate oxidative stress (Hidalgo, 1998), as well as, *rpoS*-encoded sigma factor stationary phase responsive genes and RpoH to regulate heat-shock (Murphy *et al*., 2006). Additionally, stringent starvation protein A (SspA) is expressed to ensure acid tolerance in *E. coli* (Hansen *et al*., 2005b).

1.2 The role of SspA in *E. coli*

Enteric bacteria, like *E. coli* inhabit the gastrointestinal tract of mammalian hosts (Nguyen and Sperandio, 2012), and in order to successfully colonise this environment, enteric bacteria require a mechanism to tolerate the acidic conditions of the gut. This defence mechanism employs the use of SspA, a protein found to be highly conserved in Gram-negative bacteria, like *E. coli* (Hansen *et al*., 2005b). This acid tolerance mechanism ensures not only the survival of mutualistic enteric bacteria, but that of pathogenic enteric bacteria as well (Hansen *et al*., 2005b).

SspA is a RNA polymerase-associated protein, that is necessary in activating the transcription of bacteriophage P1 late lytic genes and does so, through protein-protein interactions with the RNA polymerase holo-enzyme (Williams *et al*., 1994), and thus, it has been suggested that SspA is involved in the regulation of gene expression (Hansen *et al*., 2005b). SspA expression is induced by a prolonged stationary phase, nutrient deprivation, as well as infection by λ

bacteriophage (Hansen *et al*., 2005b) and therefore, SspA has been implicated in stress responses.

When the expression of SspA is induced, it affects the histone-like nucleoid structuring (H-NS) protein by downregulating the level of H-NS protein present in *E. coli* (Hansen *et al*., 2005b). H-NS protein represses the expression of genes that are involved in numerous stress defence mechanisms like, oxidative stress, acid tolerance and multi-drug resistance, to name a few (Hansen *et al*., 2005b). For instance, it has been shown that H-NS protein represses *gadE*, a gene that is responsible for the expression of GadE, a protein that enables the expression of Gad proteins necessary for the adequate functioning of the glutamate decarboxylase acid resistance (GDAR) system (Hansen *et al*., 2005b).

The GDAR system is the primary mechanism for acid tolerance in *E. coli*. It is initiated in response to the entry of *E. coli* into the stationary phase as well as a decline in the environmental pH (Castanie-Cornet *et al*., 1999). When GadE is expressed, it binds to the transcriptional initiation sites of *gadA* and *gadBC*, allowing for the expression of GadA, GadB and GadC. GadA and GadB are glutamate decarboxylase isoforms that are biochemically identical but expressed separately in the GDAR system (Figure 1) (Capitani *et al*., 2003). GadC is a transmembrane protein that is co-transcribed with GadB and serves as a glutamate and γ-aminobutyrate (GABA) antiporter (Smith *et al*., 1992; Capitani *et al*., 2003). Both GadA and GadB catalyse the decarboxylation of intracellular glutamate and with the consumption of H^+ ions that leaked into the *E. coli* cytoplasm, neutral GABA is produced (Smith *et al*., 1992; Capitani *et al*., 2003) (Figure 1). In this way, GadA and GadB aid in maintaining the physiological pH of *E. coli* under acidic conditions (Castanie-Cornet *et al*., 1999). The expression of GadE is therefore vital in ensuring that GadA, GadB and GadC are expressed in the GDAR system to enable acid tolerance in *E. coli* (Hansen *et al*., 2005b).

As a result, the expression of GadE relies on the downregulation of H-NS protein by SspA to allow the GDAR system to be activated and thus indicates that SspA is a positive regulator of *gadE* by negatively regulating H-NS protein (Hansen *et al*., 2005b). Hansen *et al* (2005b) has further shown that *E. coli* with an in-frame deletion of *sspA*, had decreased expression of numerous genes including those responsible for acid tolerance (*gadA*, *gadB*, *gadC* and *gadE*) in comparison to wild-type *E. coli*. The expression of these acid tolerance genes was restored when a plasmid containing *sspA* was introduced to the *sspA* mutant. This further shows that

SspA is a positive regulator of the gad genes and emphasises that SspA plays a role in gene regulation (Hansen *et al*., 2005b).

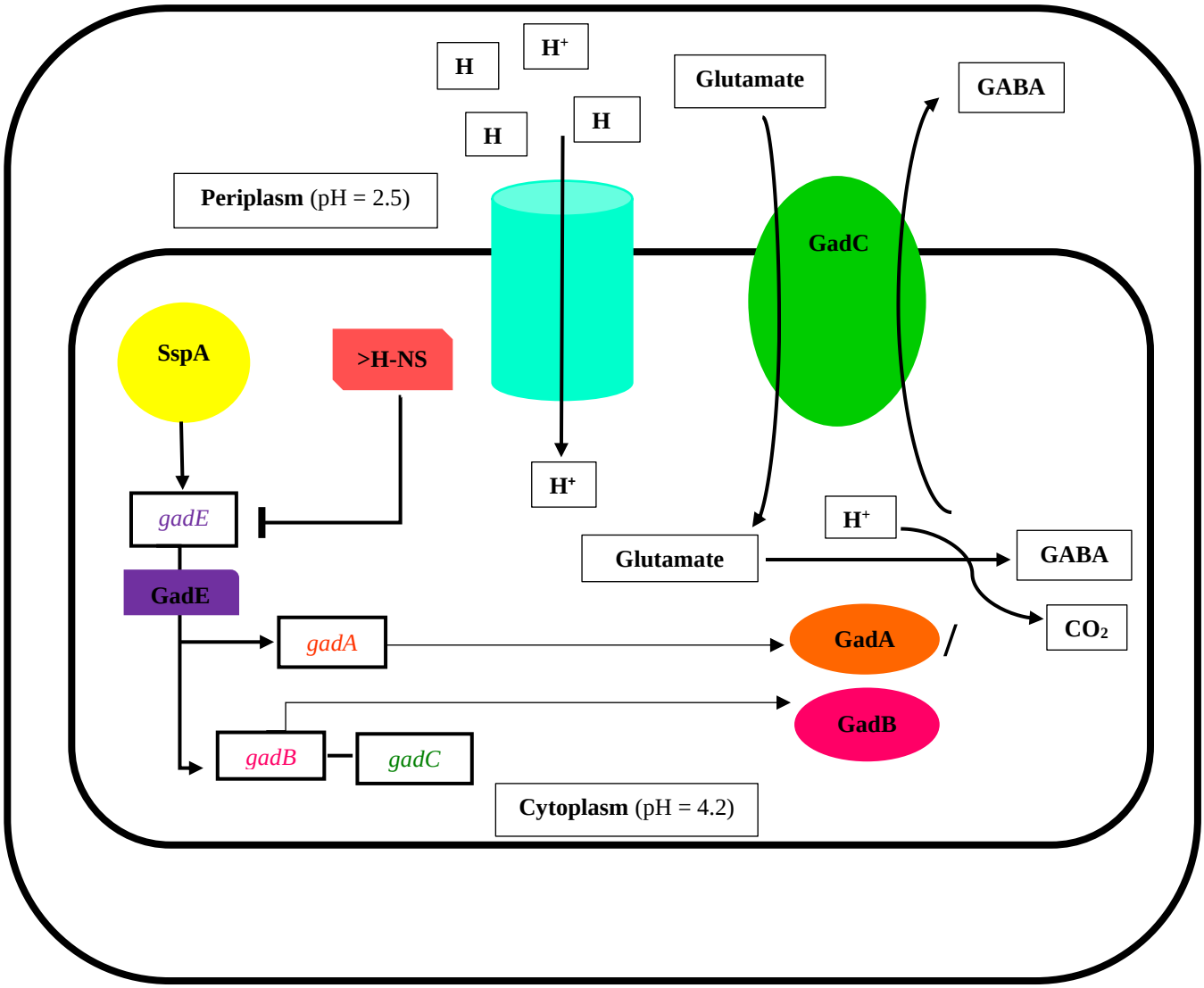


Figure 1: Schematic of GDAR system in *E. coli* bacterium. The GDAR system is activated by the expression of *gadE*, which relies on the presence of SspA to limit the number of H-NS protein in the *E. coli* cell. The successful expression of *gadE* will allow for *gadA* and *gadBC* to be expressed and enable GadA and GadB to convert intracellular glutamate into GABA by consuming H^+ ions that leaked into the cell. GABA is exported out of the bacterium cytoplasm using GadC. The GDAR system allows for the maintenance of the physiological pH of the *E. coli* cell in highly acidic environment.

When the accumulation of H-NS protein is limited by the presence of SspA in pathogenic *E. coli*, the genes responsible for acid tolerance (like *gadA*, *gadB*, *gadC* and *gadE*) and virulence become activated, allowing the survival of pathogenic *E. coli* in acidic environments (Hansen and Jin, 2012). This enables pathogenic *E. coli* to affect the health of their mammalian hosts. As a result, understanding the structure and function of SspA will aid in determining mechanisms to hinder SspA function and ensure the health of infected hosts.

1.3 SspA orthologues

Orthologues are genes present in different organisms that are related to one another due to common ancestry (Fitch, 2000). Orthologous proteins possess enough similarity in their primary structures to enable their tertiary structures to fold in a similar way (Andrykovitch *et al*., 2003; Hansen *et al*., 2005a). Therefore, in the event that the protein from the organism of interest is unable to form a crystal or the crystal formed is not satisfactory for structural determination, the use of an orthologous protein can be promising (Andrykovitch *et al*., 2003).

SspA orthologues from *Yersinia pestis*, *Vibrio cholerae* and *Pseudomonas aeruginosa* were crystallized, while crystals of *E. coli* SspA (EcSspA) were not suitable for structural determination (Andrykovitch *et al*., 2003). Only the *Y. pestis* SspA (YpSspA) crystal proved to be appropriate for predicting the structure of SspA (Andrykovitch *et al*., 2003; Hansen *et al*., 2005a). YpSspA was found to have an amino acid sequence that is 83% identical to that of EcSspA (Andrykovitch *et al*., 2003). Therefore, it can be proposed that YpSspA has a similar three-dimensional structure as EcSspA and as a result, YpSspA can be used to gain further insight into the function(s) of EcSspA (Andrykovitch *et al*., 2003). Additionally, a triple alanine mutation of EcSspA (P84A, H85A and P86A) was constructed and the SspA of *Y. pestis*, *V. cholerae* as well as *P. aeruginosa* were studied in this *E. coli sspA* mutant and showed that the SspA orthologues were able to provide acid resistance to the acid-sensitive *E. coli* (Hansen *et al*., 2005a). This indicated that they share the ability to support acid resistance in *E. coli* and emphasised that the SspA orthologues share a similar function in *E. coli* (Hansen *et al*., 2005a).

1.4 EcSspA structure

The EcSspA subunit is a 24.5 kDa protein (Gururaja Rao *et al*., 2017), that displays sequence and structural similarity with cytosolic members of the glutathione transferase (GST) family, such as beta-GSTs (Hansen *et al*., 2005a). Most cytosolic GSTs as well as EcSspA, exist and function as dimers (Hansen *et al*., 2005a; Atkinson and Babbitt, 2009), with each subunit containing two domains, a N-terminal and C-terminal domain. The N-terminal domain of both cytosolic GSTs and EcSspA, consists of a characteristic thioredoxin-fold (Atkinson and Babbitt, 2009; Gururaja Rao *et al*., 2017), further indicating the shared structural similarities between EcSspA and cytosolic GSTs.

Most GSTs can conjugate the thiol group of glutathione (GSH), a compound that binds to the G-site of the thioredoxin-fold, to toxic substrates and facilitate their metabolism, therefore illustrating the enzymatic capabilities of GSTs (Winayanuwattikun and Ketterman, 2004). For instance, 1-chloro-2,4-dinitrobenzene (CDNB) is a synthetic substrate that is often used to measure the enzymatic activity of most GSTs *in vitro* and can give an indication as to whether a GST-like protein can catalyse GSH and CDNB (Zablotowicz *et al*., 1995). Unlike GSTs, EcSspA is found to have no catalytic activity towards CDNB, nor can it bind GSH (Hansen *et al*., 2005a). The absence of GST-like enzymatic activity was assumed to be due to the replacement of the catalytic Cys^{10} residue in GST B1-1 by Phe^{21} and Tyr^{21} residues in YpSspA and EcSspA respectively, as shown in Figure 2. However, when the Tyr^{21} residue in EcSspA was substituted with a cysteine residue, EcSspA continued to lack catalytic activity towards CDNB (Hansen *et al*., 2005a). This highlights that although there are structural similarities between EcSspA and beta-GSTs, there are differences within their structure that make fulfilling the same function, like GST enzymatic activity improbable.

In comparison to GST B1-1, the proposed G-site in EcSspA has two vital residues altered at positions 21 and 58 that are required for protein-GSH interactions; where the catalytic residue Cys^{10} and Gln^{51} in GST B1-1 are replaced by Tyr^{21} and Ser^{58} in EcSspA, respectively (Hansen *et al*., 2005a). Phe^{21} and Tyr^{21} in YpSspA and EcSspA, respectively, are unable to form a hydrogen bond with the thiol group from GSH; furthermore, it was also proposed that the side chain of Phe^{21} and Tyr^{21} could cause some steric hinderance, therefore making it difficult to bind GSH or facilitate its conjunction to a xenobiotic substrate (Hansen *et al*., 2005a). The Thr^{58} and Ser^{58} in YpSspA and EcSspA, respectively, have side chains which are shorter than Gln^{51} in GST B1-1 (Figure 2) and as a result, alters the shape of the GSH binding site, with no

opportunity for a hydrogen bond to form between GSH and YpSspA and EcSspA (Hansen *et al*., 2005a).

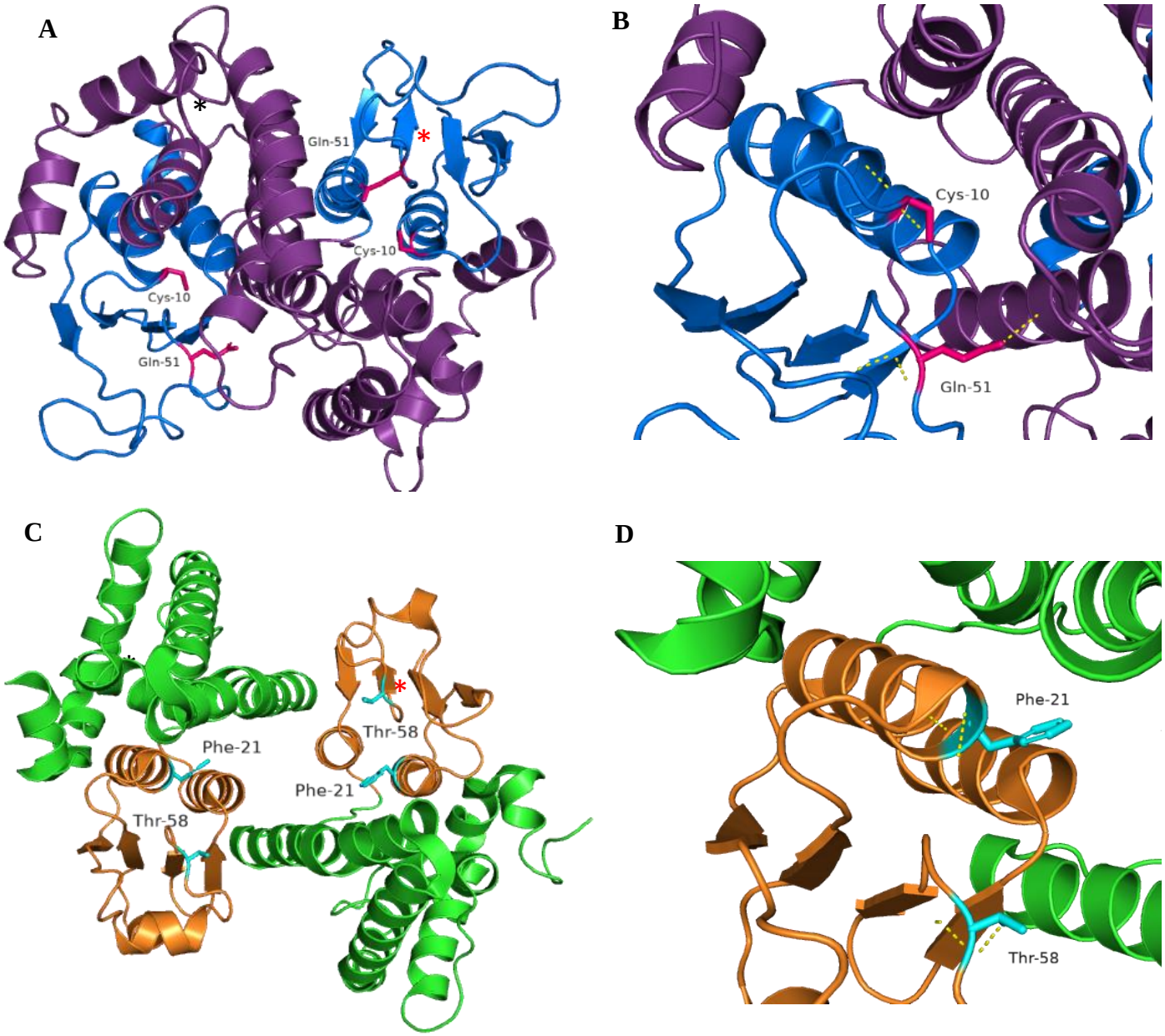


Figure 2: Cartoon representation of GST B1-1 and YpSspA. (A) Dimeric GST B1-1 is shown with the N and C-domains displayed in blue and purple respectively. Cys^{10} and Gln^{51} (in pink) are indicated in the N-domain, these residues are important in the binding of GSH in the G-site and the hydrogen bonds (yellow dashed lines) associated with Cys^{10} and Gln^{51} are shown in **B** (PDB code: 1A0F). **(C)** Dimeric YpSspA is shown with the N and C-domains illustrated in orange and green respectively. Phe^{21} and Thr^{58} (in blue) are indicated in the N-domain, in the proposed G-site and the hydrogen bond (yellow dashed lines) associated with Phe^{21} and Thr^{58} are shown in **D** (PDB code: 1YY7). The location of the G-site (*) and H-site (*) are indicated. Images were created using PyMol version 2.2.0.

1.5 EcSspA surface pocket

The crystal structure of YpSspA revealed a surface pocket at the dimer interface, which is formed by residues from both subunits (Hansen *et al*., 2005a). The residues that are involved in the formation of the surface pocket are located in the loop region between α-helices 3 and 4 and these residues, which are predominantly surface exposed, include: Arg82, Pro84, His85, Pro86, Leu88 and Tyr92 (Hansen *et al*., 2005a). Pro84, His85 and Pro86 are amongst 20 amino acid residues that were found to be strictly conserved in 50 SspA orthologues, including EcSspA, through multiple sequence alignment (Hansen *et al*., 2005a). A homology model of EcSspA was constructed using the structure of YpSspA to determine the role the surface pocket played in EcSspA function, as the function(s) of YpSspA is unknown (Hansen *et al*., 2005a) (Figure 3).

The surface pocket is mainly hydrophobic and has been suggested to be involved in forming protein-protein interactions between EcSspA and RNA polymerase (Gururaja Rao *et al*., 2017). The association of RNA polymerase with EcSspA would prove necessary for the transcription of phage P1 late lytic genes and genes associated with acid resistance (Hansen *et al*., 2005b). Furthermore, considering that His85 and Tyr92 can act as hydrogen bond donors, these residues could therefore strengthen the interaction between EcSspA and RNA polymerase (Hansen *et al*., 2005a).

The surface pocket has also shown to play a critical role in acid tolerance. Single alanine substitutions of the residues making up the surface pocket had a 4-8-fold decrease in acid tolerance of the mutants, with the exception of the P84A mutant, which exhibited acid tolerance capabilities like that of wild-type SspA (Hansen *et al*., 2005a). The substitution of either His85 or Tyr92 with alanine resulted in an 8-fold decrease in acid tolerance, where the substitution of both His85 and Tyr92 produced a 40-fold decrease in acid tolerance (Hansen *et al*., 2005a). The ring system produced by His85 and Tyr92 is therefore shown to be critical in ensuring that EcSspA fulfils its function in acid tolerance (Hansen *et al*., 2005a). It has been predicted that the proline residues in the surface pocket ensure that the ring system is well positioned to carry out the necessary functions of the surface pocket (Hansen *et al*., 2005a).

When a triple alanine substitution of P84A, H85A and P86A was conducted, EcSspA was unable to tolerate the acidic environment (Hansen *et al*., 2005a). This emphasises that the structure of the surface pocket is important for EcSspA to allow tolerance of a low pH environment. Other than being necessary for acid tolerance, the surface pocket has been

postulated to be a potential target site for inhibiting the function of EcSspA (Gururaja Rao *et al*., 2017; Hansen *et al*., 2005a).

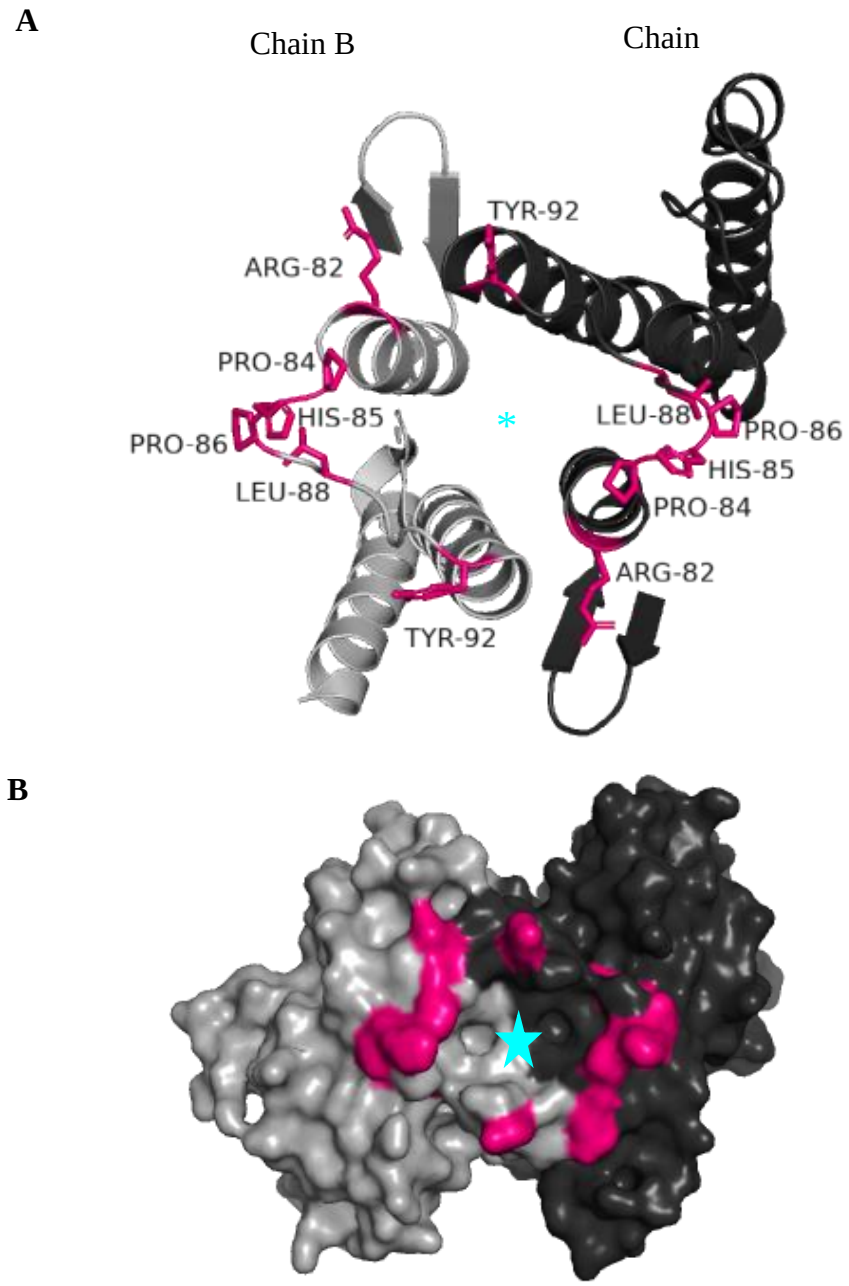


Figure 3: **Surface pocket of EcSspA modelled after YpSspA (PDB code: 1YY7).** (**A)** Cartoon representation of EcSspA homology model with residues in position 1-40 and 140-213 of Chain A (dark grey) and Chain B (light grey) removed for clearer illustration of the labelled residues. The labelled residues (Arg82, Pro84, His85, Pro86, Leu88 and Tyr92) shown in pink from each subunit make up the surface pocket (*). **(B**) Surface of EcSspA homology model with Chain A (dark grey) and Chain B (light grey) shown. Indicated in pink are the residues making up the surface pocket (blue star) from each subunit: Arg82, Pro84, His85, Pro86, Leu88 and Tyr92. Images created using PyMol version 2.2.0.

1.6 EcSspA is a bacterial homologue of human CLIC1

Human CLIC1 (hCLIC1) is a member of the chloride intracellular channel (CLIC) family of proteins. The CLIC proteins are classified as metamorphic proteins, as they can exist in both soluble and membranous forms (Valenzuela *et al*., 1997; Qian *et al*., 1999). The soluble form of the CLIC proteins are members of the GST superfamily (Dulhunty *et al*., 2001). The membrane form of the CLIC proteins, are located in plasma membranes as well as the membranes of intracellular organelles such as mitochondria and nuclei membranes, where they auto-insert into these membranes to form functional ion channels (Valenzuela *et al*., 1997; Berryman and Bretscher, 2000). CLIC proteins form selective ion channels and facilitate the transport of chloride ions in and out of cells; they are also associated with an array of physiological processes such as cell division and kidney function, to name a few (Valenzuela *et al*., 1997).

The soluble form of hCLIC1 is structurally homologous to soluble EcSspA (Gururaja Rao *et al*., 2017). Although hCLIC1 and EcSspA have a low sequence identity, they share structural similarities such as the distinctive thioredoxin-fold in the N-terminal domain, as well as a predominantly α-helical C-terminal domain (Harrop *et al*., 2001; Gururaja Rao *et al*., 2017). The ion channel activity of EcSspA was monitored by recording single ion channel currents of EcSspA bound to planar lipid bilayers and approximately 35 out of 50 recorded experiments revealed the presence of ion channel activity (Gururaja Rao *et al*., 2017). This indicated that soluble EcSspA is also metamorphic and forms functional ion channels similarly to hCLIC1 (Gururaja Rao *et al*., 2017). However, the mechanism in which both soluble hCLIC1 and EcSspA rearrange their structure to the membranous form and undergo insertion into membranes is still unknown (Tulk *et al*., 2000; Harrop *et al*., 2001).

Additionally, indanyloxyacetic acid 94 (IAA-94) was introduced to the EcSspA that had formed ion channels, to assess the effect it would have on the ion channel activity of EcSspA (Gururaja Rao *et al*., 2017). IAA-94 is a chloride channel blocker that is known to inhibit the ion channel activity of CLIC proteins (Tulk *et al*., 2000; Harrop *et al*., 2001) and in the presence of approximately 100 µM of IAA-94, the ion channel activity of EcSspA had decreased to a near inhibition state (Gururaja Rao *et al*., 2017). This suggests that the membrane form of EcSspA and hCLIC1 share a similarity in their function as membrane proteins.

Furthermore, it was shown that like the ion channel activity of hCLIC1, the ion channel activity of EcSspA is sensitive to pH and has increased ion channel activity in low pH environments

(Tulk *et al*., 2002; Gururaja Rao *et al*., 2017). The open probability of the EcSspA channel had increased significantly at a pH of 5.5, where changing the pH to 8.5, resulted in a significant decrease in the open probability of the EcSspA channel (Gururaja Rao *et al*., 2017). This further showed the similarities in the ion channel activity of hCLIC1 and EcSspA.

The structural similarities noted between soluble hCLIC1 and EcSspA, as well as the similarities noted in the ion channel activity of hCLIC1 and EcSspA show the relatedness between the proteins in both soluble and membrane form.

1.7 IAA-94 and Ethacrynic acid (EA) potentially inhibit EcSspA function

IAA-94 has been found to inhibit the ion channel activity of membrane bound hCLIC1 (Harrop *et al*., 2001; Tulk *et al*., 2002) as well as the enzymatic activity of soluble hCLIC1 in its monomeric form (Al Khamici *et al*., 2015). Soluble hCLIC1 possesses glutaredoxin-like glutathione-dependent oxidoreductase activity towards 2-hydroxyethyl disulfide (HEDS) and can catalyse known substrates of glutaredoxin-1, such as sodium selenite (Al Khamici *et al*., 2015). HEDS was shown to be reduced when hCLIC1 was coupled with GSH and glutathione reductase in the presence of NADPH (Al Khamici *et al*., 2015), therefore indicating that soluble hCLIC1 has glutathione-dependent oxidoreductase activity. Unlike soluble hCLIC1 and cytosolic GSTs, soluble EcSspA has shown no glutaredoxin-like or GST enzymatic activity, although soluble EcSspA has similar structural characteristics to soluble hCLIC1 and cytosolic GSTs (Hansen *et al*., 2005a).

EcSspA in pathogenic *E. coli* has shown to play a role in pathogenesis (Hansen *et al*., 2005a) and as a result, identifying potential inhibitors of EcSspA function would be beneficial in preventing disease caused by pathogenic enteric *E. coli*. Therefore, IAA-94 and EA, which are ligands that bind and inhibit the enzymatic activity of soluble hCLIC1 and cytosolic GSTs, respectively (Al Khamici *et al*., 2015; Rhodes and Twentyman, 1992), were used in this study to determine whether IAA-94 and EA would bind to soluble EcSspA and if so, where IAA-94 and EA are most likely to bind as well as the affinity that soluble EcSspA has for these ligands.

EA is a compound that inhibits the reabsorption of ions in the loop of Henle in the kidney and results in the excretion of the ions causing an increase in urine production (Allocati *et al*., 2018). This compound is a known enzyme inhibitor of GSTs and binds to the H-site of cytosolic GSTs, where in the absence of EA, the hydrophobic substrate would bind (Rhodes and Twentyman, 1992; Allocati *et al*., 2018). EA is a structural homologue of IAA-94 (Harrop *et al*., 2001; Tulk *et al*., 2002; Wang *et al*., 2017) and therefore, given the structural relatedness

of IAA-94 and EA (Figure 4), identifying the effect that IAA-94 and EA have on soluble EcSspA may prove to be beneficial in finding potential EcSspA inhibitors.

A **B**

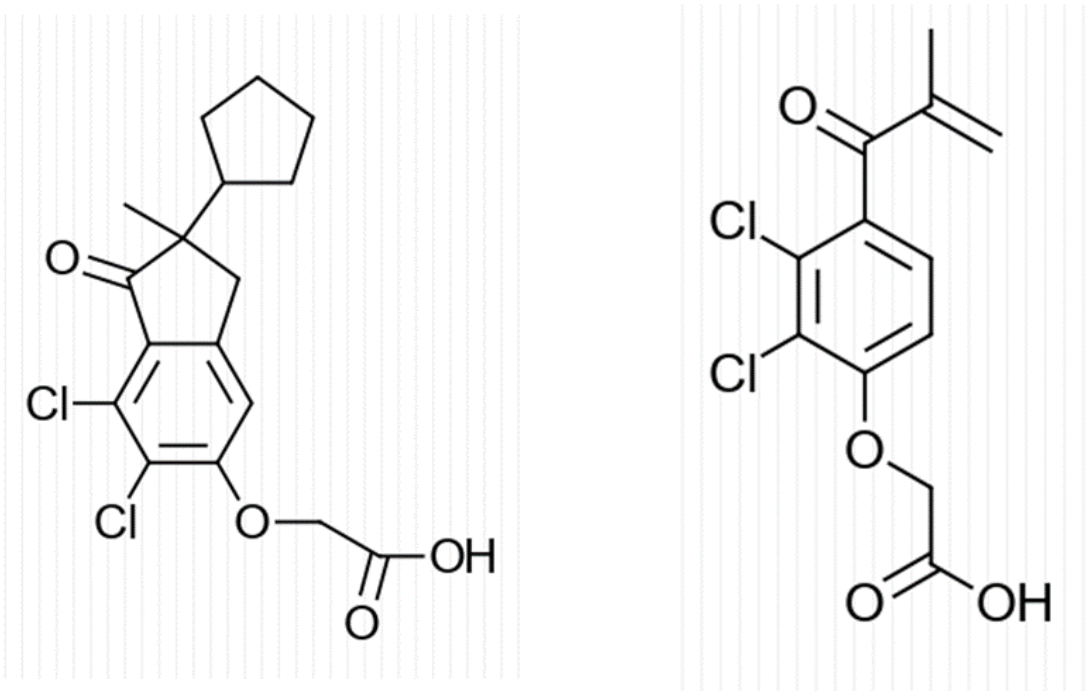


Figure 4: Chemical structure of IAA-94 and EA represented in 2-D. (A) 2-D chemical structure of IAA-94. **(B)** 2-D chemical structure of EA. Both **A** and **B** were created using ChemDoodles® (iChemLabs^{TM}, version 11.0, 2007-2020)

1.8 Aim

The aim of this project was to investigate the interaction of soluble EcSspA with IAA-94 and EA, respectively.

1.9 Objectives

1. To verify, through Inqaba Biotech^{TM}, the cDNA sequence encoding EcSspA, subcloned into the pET-28a vector.
2. To overexpress EcSspA in *E. coli* T7 cells and purify EcSspA using a nickel-immobilised metal affinity chromatography (Ni^{2+}-IMAC) column.
3. To verify the presence and purity of EcSspA using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).
4. To characterise the structural properties of EcSspA:
   - Secondary structure determination using far-UV circular dichroism (CD) spectropolarimetry.
   - Tertiary structure determination using intrinsic fluorescence spectroscopy and near-UV circular dichroism (CD) spectropolarimetry.
   - Quaternary structure determination using size exclusion-high performance liquid chromatography (SE-HPLC).
5. To characterise the ligand binding ability of EcSspA:
   - Displacement of ANS from EcSspA by IAA-94 and EA using extrinsic fluorescence spectroscopy.
   - Assess the interaction of IAA-94 and EA to EcSspA using isothermal titration calorimetry (ITC).
6. To computationally dock IAA-94 and EA to EcSspA using induced-fit molecular docking software from Schrödinger Maestro version 12.0.

Chapter 2: Experimental Procedures

2.1 Materials

The pET-28a vector encoding for EcSspA was purchased from GenScript (USA) and the T7 Express Competent *E. coli* cells were obtained from New England Biolabs (USA). The GeneJET® plasmid mini-prep kit used was purchased from ThermoScientific (USA) and the HisTrap Fast Flow (FF) Column was purchased from GE Healthcare (Sweden). Isopropyl β-D-1-thiogalactopyranoside (IPTG), kanamycin, chloramphenicol and the BLUeye Prestained Protein Ladder was purchased from Sigma-Aldrich (USA). The 8-anilino-1-naphthalenesulfonic acid ammonium salt (ANS), IAA-94 and EA were obtained from Sigma-Aldrich (USA). The gel filtration standards were purchased from Bio-Rad Laboratories (USA). All the other chemical reagents used in this study were of analytical grade.

2.2 Methodology

2.2.1 pET-28a vector

In addition to containing the cDNA sequence for EcSspA, the pET-28a vector (GenScript, USA) used possesses a kanamycin resistance gene that served as a selection marker for T7 *E. coli* cells (New England Biolabs, USA) that were transformed with the pET-28a vector. The vector also consists of a T7 promoter which allowed for the T7 RNA polymerase that was transcribed by the genome of the T7 *E. coli* cells, to transcribe and overexpress EcSspA (Tabor Stanley, 1990; Zalenskaya *et al*., 1990). Furthermore, a poly-histidine tag (his-tag) present at the N-terminus, was used to aid the purification of EcSspA. A thrombin cleavage site is also present in the vector to remove the his-tag, however, in this study the his-tag was not cleaved from EcSspA.

2.2.2 Verification of cDNA sequence encoding EcSspA

A 1 ml glycerol stock of T7 *E. coli* cells, that was previously transformed with the pET-28a vector was thawed on ice. The T7 *E. coli* cells used, were previously transformed using the heat-shock method, where 100 µl of chemically competent T7 *E. coli* cells were thawed and introduced to 6 µl of the thawed plasmid. Once mixed, the microcentrifuge tube containing the competent T7 *E. coli* cells and plasmid were immediately placed on ice for 15 minutes and then heat-shocked at 42˚C for 50 seconds and thereafter placed on ice for a further 2 minutes. The transformed cells were added to 1 ml of SOC media that was incubated at 37˚C and were then plated on Luria Broth (LB) agar plates (1.5% (w/v) agar, 1% (w/v) tryptone, 0.5% (w/v) yeast extract and 0.5% (w/v) sodium chloride supplemented with 30 µg/ml kanamycin and 35 µg/ml chloramphenicol) to determine whether the competent T7 *E. coli* cells were indeed transformed

with the plasmid. The colonies of the transformed cells were used to prepare the above-mentioned glycerol stock. Once thawed, the glycerol stock was used to inoculate 100 ml of 2xYT medium (1% (w/v) yeast extract, 1.6% (w/v) tryptone, 0.5% (w/v) sodium chloride) that was supplemented with 30 µg/ml kanamycin and 35 µg/ml chloramphenicol. Chloramphenicol was introduced to the medium to select for only T7 *E. coli* cells as they possess a chloramphenicol resistance gene (Kesik-Brodacka *et al*., 2012). The inoculated medium was well aerated and placed in a rotating incubator at 250 rpm, for 12-16 hours at 37˚C.

Subsequently, the cell culture was used to extract the plasmid, as per the instructions of the GeneJET™ plasmid mini-prep kit (ThermoScientific, USA). The extracted plasmid was stored at -20˚C and approximately 20 µl of the isolated plasmid was sent to Inqaba Biotech™ (Pretoria, South Africa) for Sanger sequencing, to verify that the cDNA sequence inserted in the plasmid encoded for EcSspA. The remaining cell culture was used to prepare glycerol stocks in a 1:1 ratio of glycerol (80% (v/v) glycerol) to cell culture.

The sequence received from Inqaba Biotech™ (Pretoria, South Africa) was extracted using Finch-TV (Version 1.4.0, Geospiza Inc.) and the sequence was analysed using NCBI BLASTx (https://blast.ncbi.nlm.nih.gov/Blast.cgi). NCBI BLASTx searched protein databases to find amino acid sequences that aligned with the nucleotide sequence sent by Inqaba Biotech™ (Pretoria, South Africa).

2.2.3 Overexpression of EcSspA

In preparation for overexpressing EcSspA, 7.5 L of fresh 2xYT medium was prepared and added to 15 Erlenmeyer flasks, each containing 500 ml of medium supplemented with 30 µg/ml kanamycin. One flask was inoculated with 750 µl of a thawed glycerol stock of transformed T7 *E. coli* cells and placed in an incubator for 12-16 hours at 37˚C, while rotating at 210 rpm to ensure the flask was well aerated during this period. Thereafter, 30 ml of the cell culture was transferred to each of the remaining 14 flasks. The 14 flasks were placed in a rotating incubator at 210 rpm, at 37˚C for 12-16 hours.

At 15-minute intervals, the optical density at 600 nm (OD_{600}) (mid-exponential growth phase) of the growing cell culture was recorded using a UV-Vis spectrophotometer and a growth curve was established by plotting a graph of the absorbance recorded at 600 nm as a function of time. When an OD_{600} of approximately 0.7 was reached, a final concentration of 0.2 mM of IPTG was added to each flask to induce the overexpression of EcSspA for 6-8 hours.

Subsequent to the overexpression of EcSspA, the cells were harvested through centrifugation at 5000 x g, for 10 minutes at 4˚C and a 1 ml sample of the supernatant was stored, and the remainder was discarded. The pellet was resuspended in 30 ml of resuspension buffer (20 mM Tris-HCl, 500 mM sodium chloride, 30 mM imidazole and 0.01% (w/v) sodium azide, pH 8) per litre of cell culture and 1 ml of 1 mM fresh benzamidine and phenylmethylsulfonyl fluoride (PMSF) were added to the resuspended pellet, to inhibit proteases that could be released during cell lysis, to ensure that no degradation of EcSspA would occur. The cells were placed in two 50 ml falcon tubes and stored at – 20˚C.

2.2.4 Purification of EcSspA

The cells (obtained in section 2.2.3) were placed in a beaker of water and were gently thawed over a benchtop stirrer. To aid in cell lysis, magnesium chloride (1 M), DNase (10 mg/ml) and lysozyme (10 mg/ml) (1 µl of each reagent was added per ml of cells) were added directly to the cells and left to incubate for 30 minutes. Thereafter, the cells were sonicated using the PRO 200 homogenizer (PRO Scientific Inc, USA) for 3 x 30s cycles, with the pulse at a medium setting and the cells were placed on ice between cycles. The cell lysate was further centrifuged at 23 000 x g for 10 minutes at 4˚C, the supernatant was collected and filtered using a 0.45 µm syringe filter. A sample of the supernatant and the pellet that remained after centrifugation was kept and stored at -20˚C and used to prepare samples for SDS-PAGE (see section 2.2.5).

Prior to applying the filtered supernatant onto the 5 ml Ni^{2+}-IMAC column (GE Healthcare, Sweden), the column was connected to the AKTA purification system and the equilibration buffer (20 mM Tris-HCl, 500 mM sodium chloride, 30 mM imidazole and 0.01% (w/v) sodium azide, pH 8) was pumped into the system at a flow rate of 4 ml/min to equilibrate it prior to EcSspA purification. When the AKTA purification system was equilibrated with 10 column volumes (this was noted by an A_{280} of 0 mAU), 50 ml of the filtered supernatant was loaded onto the column (this was noted by an increase in the A_{280}). The His-tagged EcSspA bound to the column and all the molecules that remained unbound were removed from the column using the equilibration buffer and collected as the flow-through. Elution of the bound His-tagged EcSspA was conducted when the A_{280} returned to 0 mAU by applying the elution buffer (20 mM Tris-HCl, 500 mM sodium chloride, 300 mM imidazole and 0.01% (w/v) sodium azide, pH 8) to the column. The eluted EcSspA was collected in three 5 ml fractions when the A_{280} started increasing.

The eluted EcSspA fractions were put in dialysis tubing with a 10 kDa molecular weight cut-off and placed in dialysis buffer (50 mM potassium phosphate, 100 mM sodium chloride, 1 mM EDTA, 0.01% (w/v) sodium azide, pH 7). Two buffer exchanges occurred in intervals of 3-4 hours as well as a last buffer exchange that occurred for 12-16 hours. Dialysis took place at 4˚C on a magnetic stirrer to aid the process of dialysis. Purified EcSspA was stored in the dialysis buffer at 4˚C and a litre of the buffer was filtered and kept for spectroscopic studies.

2.2.5 Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

To assess the presence and purity of the eluted EcSspA, a 12.5% (w/v) SDS-PAGE gel (Laemmli, 1970) was made. A 12.5% discontinuous SDS-PAGE gel consists of a 12.5% (w/v) separating gel (30% acrylamide/2.67% bis-acrylamide solution, 10% (w/v) SDS, 10% (w/v) ammonium persulfate, 100% (v/v) TEMED and 1.5 M Tris-HCl, pH 8.8) and a 4% stacking gel (30% acrylamide/2.67% bis-acrylamide solution, 10% (w/v) SDS, 10% (w/v) ammonium persulfate, 100% (v/v) TEMED and 0.5 M Tris-HCl, pH 6.8).

The samples that were analysed using SDS-PAGE were prepared as follows: 50 µl of each sample was added to 50 µl of reducing sample buffer (0.5 M Tris, pH 6.8, 20% (v/v) SDS, 25% (v/v) glycerol and 0.5% (w/v) bromophenol blue). Thereafter, β-mercaptoethanol was added to the samples in a 1:10 (v/v) ratio prior to heating the samples at 95˚C for ~ 5 minutes. Approximately 2 µl of the BLUeye Prestained Protein Ladder (GeneDireX Inc.) was loaded into the first well to serve as the molecular weight marker (MWM) and 5 µl of each sample was loaded into the other remaining wells. The gel was placed in the electrophoresis tank that was appropriately filled with 1x electrophoresis buffer (25 mM Tris-HCl, 192 mM glycine and 10% (w/v) SDS, pH 8.3) and electrophoretic separation of the samples occurred at 160 V for ~ 1 hour. All electrophoretic apparatuses used are from Bio-Rad Laboratories (USA).

Thereafter, the gel was removed from the electrophoretic tank and placed in a stain solution (0.25% (w/v) Coomassie Brilliant Blue R250, 45% (v/v) methanol and 10% (v/v) glacial acetic acid.) overnight. After that, the stain solution was removed and the de-stain solution (10% (v/v) glacial acetic acid, 40% (v/v) methanol and 50 % (v/v) distilled water) was poured over the gel until the gel was clear of the stain.

An SDS-PAGE calibration curve was also constructed by measuring the distance (cm) that the protein standards of known molecular mass migrated on the gel and compared it to the distance (cm) that the protein sample migrated on the gel relative to the dye front. The distance migrated by the protein standards was plotted against the log of the molecular mass of the protein

standards and the linear regression model from the fit of the data was used to determine the size of the EcSspA subunit. All plots were created using SigmaPlot version 12.0.

2.2.6 Protein concentration determination and purity assessment

The concentration of EcSspA was determined using absorbance spectroscopy. Five dilutions of the protein were prepared. The dilutions of EcSspA measured at A_{280} were plotted and a regression line was computed for the data. Obtaining the A_{280} and the molar extinction coefficient of EcSspA is important in calculating the concentration when using the Beer-Lambert Law:

$$\mathbf{A = \varepsilon . c . l} \quad (1)$$

where **A** represents the absorbance value at 280 nm, **ε** represents the molar extinction coefficient of EcSspA at 280 nm (M^{-1} cm^{-1}), **c** represents the concentration (M) of EcSspA and **l** represents the pathlength (1 cm) of the cuvette (Swinehart, 1962). To calculate the molar extinction coefficient the following equation was used:

$$\varepsilon_{280} = (nW \times \varepsilon_W) + (nY \times \varepsilon_Y) + (nC \times \varepsilon_C)$$

$$\varepsilon_{280} = (nW \times 5\,500\ M^{-1}\ cm^{-1}) + (nY \times 1\,495\ M^{-1}\ cm^{-1}) + (nC \times 125\ M^{-1}\ cm^{-1}) \quad (2)$$

where n represents the number of W (tryptophan), Y (tyrosine) or C (cystine) residues. The cysteine residues are considered negligible at A_{280} due to low contribution that the molar extinction coefficient has on the ε_{280} (Pace *et al*., 1995). The molar extinction coefficient of the EcSspA dimer was calculated to be 56 965 M^{-1} cm^{-1}.

An absorbance spectrum of EcSspA was also conducted to assess the purity of the protein. The absorbance spectrum was measured using the Jasco V-630 spectrophotometer over a wavelength range of 200-400 nm. The absorbance at 260 nm, 280 nm and 340 nm was closely observed. The A_{260}/A_{280} ratio can be used to determine the purity of the protein. An A_{260}/A_{280} ratio of 0.6 or lower indicates that the protein is free of nucleic acid contamination (Warburg and Christian, 1942) and monitoring whether a peak is observed at 340 nm, gave an indication of possible aggregation of the protein. The following parameters were entered into the Jasco V-630 spectrophotometer: data pitch was 1 nm, the scan speed was 200 nm/min and 3 accumulations were recorded. The data obtained was corrected for buffer (50 mM potassium phosphate, 100 mM sodium chloride, 1 mM EDTA, 0.01% (w/v) sodium azide, pH 7) contributions. All spectra and plots were plotted using SigmaPlot version 12.0.

2.2.7 Structural characterisation of EcSspA

2.2.7.1 Far-UV Circular Dichroism (CD) spectropolarimetry

Far-UV CD is a technique that was used to evaluate the secondary structural elements of EcSspA. CD makes use of circularly polarised light that is absorbed by the amide bonds present in the polypeptide backbone of proteins in the far-UV region (250-180 nm) (Kelly and Price, 2009). Due to the chiral nature of the amide bonds, they can absorb right and left-handed circularly polarised light unequally and the differential absorption in turn gives rise to the CD signal produced (Kelly and Price, 2009). Different secondary structural elements produce distinct spectra: (i) α-helices have troughs present at 208 nm and 222 nm and a crest at 195 nm, (ii) β-sheets have a trough present at 218 nm and a crest at 196 nm and (iii) random coils have a crest present at 212 nm and a trough at 195 nm (Beychok, 1966; Kelly and Price, 2009).

In addition, the CD spectropolarimeter generated a spectrum of high-tension voltage (HT) against the wavelength. The HT is an important parameter to monitor as it ensures that reliable CD data was obtained. A HT value of 700 V or greater, shows that an insufficient amount of light reached the detector and that CD data produced at those wavelengths are not reliable and therefore HT values below 700 V are more reliable to use (DiNitto and Kenney, 2012).

A Jasco J-1500 spectropolarimeter was used to measure the far-UV spectrum of EcSspA and the following parameters were put in place: data pitch was 1 nm, the scan speed was 100 nm/min, the bandwidth was 5 nm and the pathlength was 2 mm. A 2 mm quartz cuvette was used to measure the samples, and this occurred at 20˚C. Approximately 2.5 µM of EcSspA was used in the samples and the samples were prepared by directly diluting the sample with Milli-Q water as opposed to the buffer (50 mM potassium phosphate, 100 mM sodium chloride, 1 mM EDTA, 0.01% (w/v) sodium azide, pH 7), which created a lot of noise at a wavelength of 200 nm and below. An average of three samples was obtained and corrected for contributions made by the Milli-Q water. All spectra plotted using SigmaPlot version 12.0.

The secondary structural content of YpSspA was determined the using the Full Protein Feature View from RCSB PDB (PDB ID: 1YY7; Hansen *et al*., 2005) which showed the number of amino acid residues present in each secondary structural element.

2.2.7.2 Near-UV CD spectropolarimetry

Similar to far-UV CD, near-UV CD is a technique that makes use of the absorbance of right and left-handed circularly polarised light by chromophores. However, unlike far-UV CD, the chromophores that absorb light in the near-UV region (340-250 nm) are aromatic amino acid residues (Kelly and Price, 2009). The asymmetric environment of tryptophan, tyrosine and phenylalanine residues allows for their chiral nature, and therefore the differential absorption of the right and left-handed circularly polarised light by the aromatic amino acid residues will produce a characteristic CD signal (Kelly and Price, 2009). That is, phenylalanine will give account for the CD signal between 255-270 nm and tyrosine and tryptophan will give account for the CD signal observed between 275-282 nm and 290-305 nm, respectively (Beychok, 1966; Kelly and Price, 2009).

The Jasco J-1500 spectropolarimeter was also used to measure the near-UV spectrum of EcSspA as well as the corresponding HT spectrum. The same parameters as in section 2.2.7.1 were put in place, aside from the 5 mm quartz cuvette used to measure the samples and approximately 20 µM of EcSspA was used in each sample.

2.2.7.3 Intrinsic tryptophan fluorescence spectroscopy

Intrinsic tryptophan fluorescence spectroscopy is a sensitive application that is commonly used to analyse the local environment of the tryptophan residues in protein (Lakowicz, 2006). Fluorescence occurs when a fluorophore absorbs light at a specific wavelength and the electrons become excited, moving to a higher energy level. As the electrons move back to their ground state, light is emitted at a longer wavelength than it was excited at (Lakowicz, 2006). The aromatic amino acid residues are naturally occurring fluorophores in protein (Lakowicz, 2006). However, due to the poor fluorescence signal of phenylalanine and cysteine residues, their contribution to the intrinsic fluorescence of a protein are considered negligible (Lakowicz, 2006). Tyrosine and tryptophan residues are readily excitable at 280 nm, however, tryptophan residues can be selectively excited at 295 nm (Moller and Denicola, 2002; Lakowicz, 2006). Tryptophan residues are sensitive to the polarity of their local environment and the changes occurring in that environment can be monitored by observing the maximum emission wavelength of the tryptophan residues (Moller and Denicola, 2002; Lakowicz, 2006).

The Jasco FP-6300 spectrofluorometer was used to conduct the fluorescence measurements at 20˚C. The following parameters were put in place: data pitch was 1 nm, the excitation and emission band width were 5 nm, the scan speed was 200 nm/min, the excitation wavelength

was 295 nm and the emission spectra were recorded from 295-500 nm. A quartz cuvette with a pathlength of 10 mm was used and 3 accumulations of each sample were measured. Approximately, 2 µM of native EcSspA was used and an average of three samples was obtained, and the samples were corrected for buffer contributions (50 mM potassium phosphate, 100 mM sodium chloride, 1 mM EDTA, 0.01% (w/v) sodium azide, pH 7). All spectra were plotted using SigmaPlot version 12.0.

2.2.7.4 Size Exclusion-High Performance Liquid Chromatography (SE-HPLC)

SE-HPLC is a technique used to provide information on the oligomeric state of proteins by employing the principles of size exclusion chromatography (SEC). SE-HPLC separates molecules based on their hydrodynamic volumes, where larger molecules elute first and the smaller the molecule, the longer it is retained in the column prior to eluting (Tayyab *et al*., 1991). However, unlike SEC, the HPLC system allows for higher resolution and quick separation of molecules into their various components due to the high pressure associated with the system (Tayyab *et al*., 1991; Garcia-Lopera *et al*., 2005).

The Shimadzu UFLC LC-20A system, together with the TSK-Gel Super SW 2000 column, which has a fractionation range of 5-150 kDa and guard column were used to analyse the oligomeric state of EcSspA. Approximately, 20 µl of 164 µM purified EcSspA was injected into the SE-HPLC system and resolved over 30 minutes. The absorbance of the effluent was measured at 280 nm, at a flow rate of 0.2 ml/min. The mobile phase of the system was a buffer (50 mM potassium phosphate, 1 mM of EDTA, 0.01% (w/v) sodium azide and 500 mM sodium chloride, pH 7) that was used to equilibrate the HPLC system, as well as the column prior to injecting the EcSspA sample.

A set of gel filtration standards that consisted of: thyroglobulin (670 kDa), gamma globulin (158 kDa), ovalbumin (44 kDa), myoglobin (17 kDa) and vitamin B12 (1.5 kDa) (Bio-Rad Laboratories, USA) used to construct a calibration curve by plotting a curve of the log of the molecular mass of the gel filtration standards as a function of their corresponding retention times and a regression model was fitted to this data. The gel filtration standards (20 µl) were also injected into the system and resolved under identical conditions. The calibration curve was used to estimate the molecular mass of the EcSspA sample. The elution profiles and plots were plotted using SigmaPlot version 12.0.

2.2.8 Assessment of IAA-94 and EA interaction with EcSspA

2.2.8.1 ANS concentration determination and fluorescence

ANS is an amphiphilic dye that exhibits fluorescent capabilities in hydrophobic environments. ANS binds non-covalently to hydrophobic patches on the surface of protein and is sensitive to the polarity of the environment (Matulis and Lovrien, 1998). ANS has also been used considerably to characterise the ligand binding sites present on protein (Collini *et al*., 2000). Therefore, ANS was used to identify potential binding sites on EcSspA by using extrinsic ANS fluorescence.

Prior to conducting any fluorescence measurements of ANS with EcSspA, the concentration of ANS was first determined by means of absorbance spectroscopy. A dilution plot of ANS measured at an absorbance of 350 nm was constructed and a regression model was fitted to the data. Using the Beer-Lambert Law (see equation 1), the concentration of ANS was determined using the molar extinction coefficient of ANS, which is 4950 M^{-1} cm^{-1} (Matulis and Lovrien, 1998). The dilution series was measured using the Jasco V-630 spectrophotometer and the following parameters were in place: data pitch was 1 nm, the scan speed was 200 nm/min and 3 accumulations were recorded at 20˚C. The data was corrected for contributions made by the buffer (50 mM potassium phosphate, 100 mM sodium chloride, 1 mM EDTA, 0.01% (w/v) sodium azide, pH 7).

The fluorescence measurements were conducted using the Jasco FP-6300 spectrofluorometer, 200 µM of ANS was incubated in the dark, at room temperature in the appropriate buffer as well as in the presence of 2 µM of EcSspA for approximately 1 hour. The following parameters were used: data pitch was 1 nm, the scan speed was 200 nm/min, the excitation and emission band width were 5 nm, the excitation occurred at 390 nm and the emission spectrum ranged from 390-650 nm. Any contributions made by the buffer were subtracted from the data and the samples were measured in triplicates. All spectra and plots were plotted using SigmaPlot version 12.0.

2.2.8.2 The displacement of ANS by IAA-94 and EA from EcSspA

ANS displacement studies were conducted to determine whether IAA-94 and EA could displace ANS from EcSspA using extrinsic ANS fluorescence.

Samples of free ANS in buffer were prepared, as well as samples of ANS with 3% (v/v) DMSO, ANS with 2.1 mM IAA-94 and ANS with 2.48 mM EA. The control consisted of ANS with EcSspA only. Samples that consisted of ANS, EcSspA and 3% (v/v) DMSO were made, as

well as samples of ANS, EcSspA with 2.1 mM IAA-94 and ANS, EcSspA with 2.48 mM EA. Approximately, 200 µM of ANS and 2 µM of EcSspA was used for the ANS displacement studies.

Stock solutions of IAA-94 and EA were solubilised in 100% (v/v) DMSO and to ensure that EcSspA retains function and that minimal signal artefacts are present when conducting ITC titrations, 3% (v/v) DMSO was used in samples for ANS displacement and ITC (Velazquez-Campoy and Freire, 2006), to provide consistency to the samples used in the interaction studies. The samples were prepared in triplicate and incubated in the dark at room temperature prior to conducting fluorescence measurements.

The displacement of ANS was measured by monitoring the fluorescence intensity of ANS at the maximum emission wavelength. The same parameters as in section 2.2.8.1 were put in place, however, the emission spectra were measured over a wavelength range of 390-700 nm.

2.2.8.3 Isothermal titration calorimetry (ITC)

ITC is a technique used to study the binding of small molecules to larger macromolecules or the interaction of macromolecules in solution (Jelesarov and Bosshard, 1999). This technique measures the heat that is released or absorbed during the interaction between a ligand and a macromolecule, like a protein, in the sample cell with respect to a reference cell that contains only the buffer used in the titrations (Jelesarov and Bosshard, 1999). As the ligand is injected into the sample cell, the ligands bind to the macromolecules and peaks are produced as raw data (Figure 5A). The peaks represent the change in heat energy as binding occurs. A decrease in the peaks indicates the saturation of the macromolecule. These peaks are integrated by software and a Wiseman plot is produced and thermodynamic parameters can be derived from this plot (Figure 5B) (Wiseman *et al.*, 1989).

The isotherm produced thereafter can be used to directly determine the change in enthalpy (ΔH), the binding affinity constant (K_a) and the number of binding sites (N) (Wiseman *et al*., 1989) (Figure 5). The change in Gibbs free energy (ΔG) and the change in entropy (ΔS) can be indirectly determined using the following equations:

$$\Delta G = RT \ln (K_a) = \Delta H - T\Delta S \qquad (3)$$

where ΔG represents the change in Gibbs free energy (kJ/mol), the R represents the gas constant (8.314 J/mol/K), the T represents temperature in Kelvin and the K_a represents the binding

affinity constant (M or mol/l). The ΔH represents the change in enthalpy (kJ/mol) and the ΔS represents the change in entropy (kJ/mol/K).

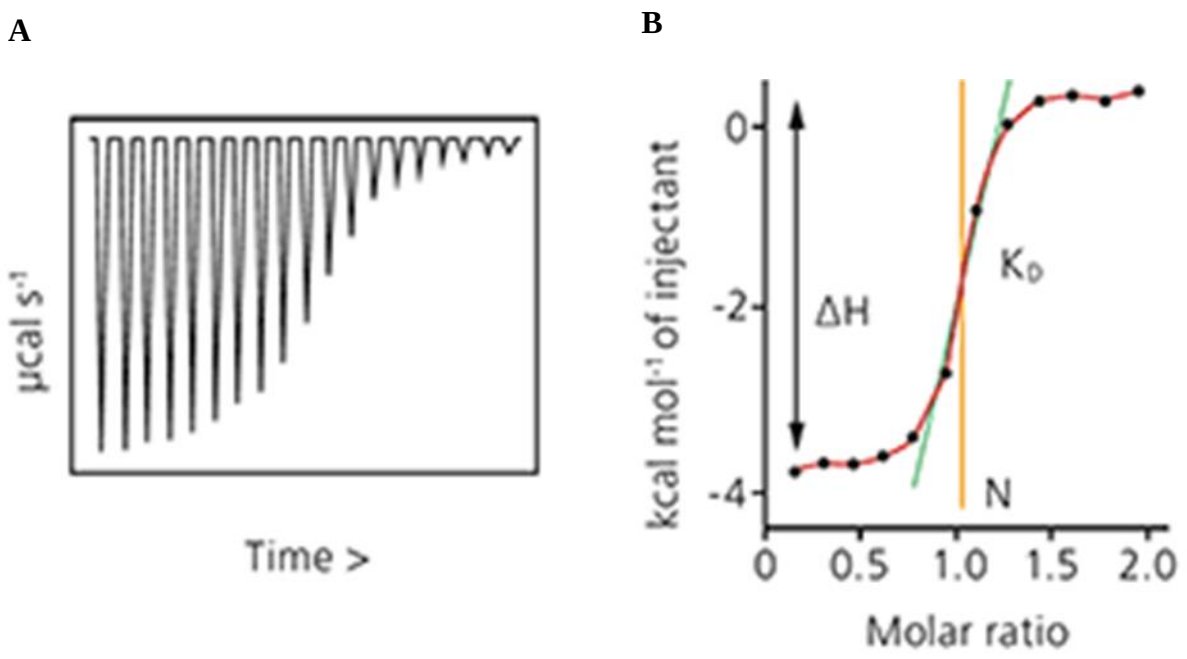


Figure 5: Schematic diagram of raw data obtained from an ITC experiment and accompanying isotherm. (A) Raw data shows downward peaks generated by the heat released during binding of ligand to the macromolecule until saturation was reached. (**B)** ITC isotherm generated from the integration of the peaks in **A**. The change in enthalpy (ΔH), the binding affinity (K_d) and the stoichiometry (N) can be directly determined from the isotherm. (**Source**: https://www.malvernpanalytical.com/en/products/technology/microcalorimetry/isothermal-titration-calorimetry)

When conducting ITC titrations, it is important to consider the c-value as reported by Wiseman *et al* (1989). The c-value is an important unitless parameter that ensures the reliability of binding isotherms and thus maximising the confidence in the binding parameters obtained from the titrations; the c-value is also considered a function of the binding constant (Ladbury and Chowdhry, 1996). The c-value is a product of the stoichiometry of the reaction (N) with the binding association constant (K_a) and the total macromolecular concentration in the sample cell of the calorimeter ($[M]_{total}$), $c = N \times K_a \times [M]_{total}$ (Wiseman *et al*., 1989). When the c-value ranges between $1 < c < 1000$, it reflects moderately tight binding reactions, with the c-value being considered optimal between $10 < c < 100$ (O'Brien *et al*., 2000). When the c-value is between 10 and 50, the shape of the binding isotherms is sigmoidal and therefore preferable for directly determining the ΔH, K_a and N (Wiseman *et al*., 1989). When the c-value is below 5, it represents weakly binding reactions and puts into question the reliability of directly determining the ΔH, K_a and N. If c-values are equal to or approach infinity, it results in very tight binding reactions (Figure 6) (Wiseman *et al*., 1989).

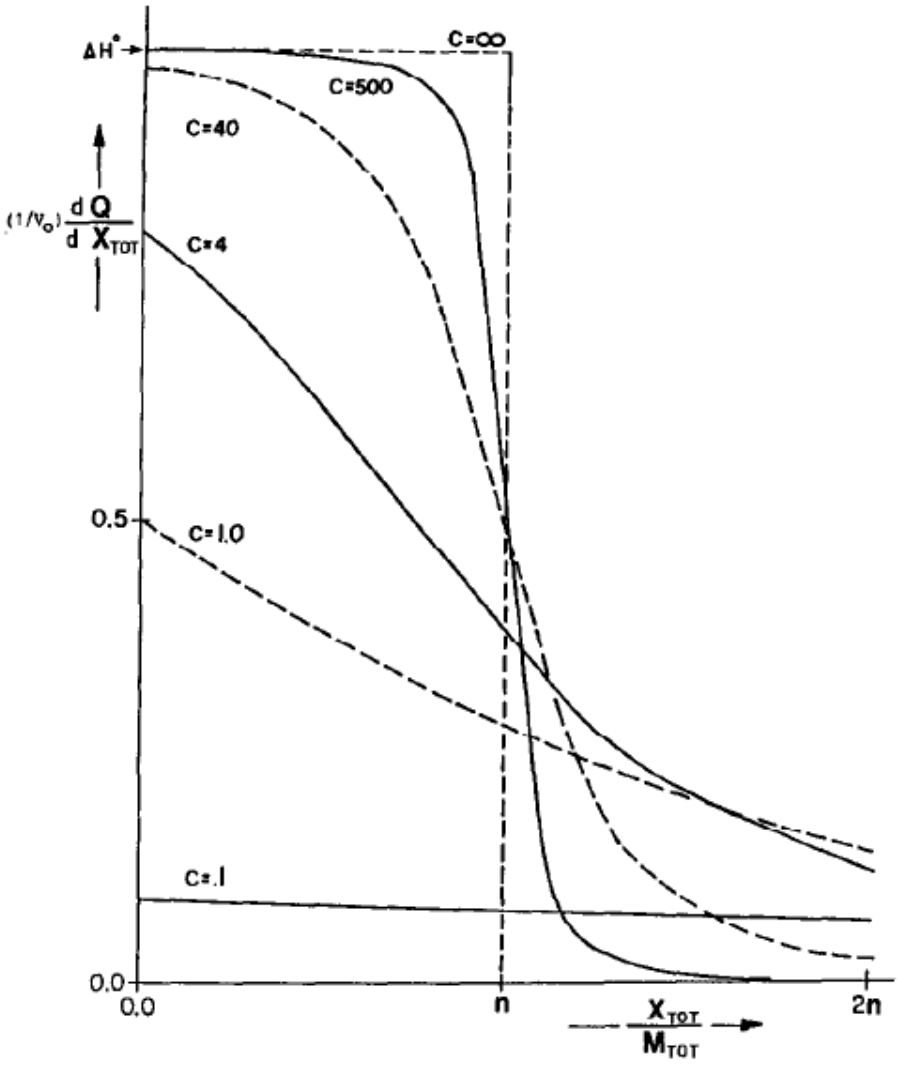


Figure 6: Simulated binding isotherm for various values of c. This binding isotherm indicates that low c-values result in shallow binding isotherms where binding parameters can be inaccurately determined and where high c-values result in steep isotherms where the binding constant cannot be determined accurately. Moderate c-values produce isotherms with a sigmoidal shape allowing for ΔH, K_d and N to be accurately and directly determined from the isotherm (Wiseman *et al.*, 1989).

ITC was used to determine the thermodynamic properties associated with the interaction between EcSspA and IAA-94 or EA. Vials of 50 mg of IAA-94 and EA were solubilized in 2 ml of 100% (v/v) DMSO, respectively, to yield a concentration of 25 mg/ml. EcSspA was dialysed (50 mM potassium phosphate, 1 mM EDTA and 0.01% (w/v) sodium azide, pH 7) and the dialysate was used to prepare the ligand solutions; 100 µM EcSspA with 3% (v/v) DMSO was prepared for the titrations. The stock concentration of IAA-94 was ~ 69.98 mM and that of EA was ~ 82.47 mM. To prepare for the titration, 30 µl of each ligand was added to 970 µl of the dialysate, to ensure no buffer mismatch occurred during the titrations. As a result, the concentration of IAA-94 and EA used in this study was 2.1 mM and 2.48 mM, respectively. All samples were degassed prior to titration.

All titration experiments were performed using a Nano-ITC calorimeter (TA instruments, USA). The Nano-ITC calorimeter was appropriately cleaned prior to all titration experiments. The sample cell contained 100 µM EcSspA with 3% (v/v) DMSO to a volume of 1 ml and the syringe contained IAA-94 or EA to a volume of 250 µl. Initial equilibration of the Nano-ITC occurred while the EcSspA was stirred at 250 rpm. Automated 5 µl injections with an initial injection of 2 µl of the ligand introduced into the sample cell, were incremental and occurred every 300s, with an initial and final baseline of 300s and ~ 40 injections were made per ITC run. All titration experiments occurred at 20˚C. Titrations of ligands into buffer alone (under identical conditions as the titrations of the ligands into protein) were also conducted to determine the heats of dilution for each ligand, as their addition causes a dilution effect which generates a heat signal. The heat signals generated due to the titration of EcSspA with ligand were subsequently corrected for heats of dilution.

NITPIC (version 1.2.7) (Keller *et al*., 2012; Scheuermann and Brautigam, 2015) and SEDPHAT (version 15.2b) are software used to correct data for the heats of dilution, obtain an unbiased baseline of titration data, to fit the data and obtain the thermodynamic parameters of the titrations. The data were fit using the A + B ↔ AB Hetero-association model implemented in SEDPHAT and represents a single molecule binding to a single site on a macromolecule. This interaction model was used to fit the data, as the data obtained from molecular docking studies (see section 2.2.8.4) revealed that the ligands used in this study could possibly bind to one site on the EcSspA structure. Data fit by SEDPHAT were exported to GUSSI (version 1.4.2) (Brautigam, C.A., 2015) to graphically represent the fitted data appropriately.

2.2.8.4 Molecular docking of IAA-94 and EA to EcSspA homology model

Homology Modelling

Homology modelling is a technique that is utilized to predict the three-dimensional structure of a protein, when only the amino acid sequence is available (Cavasotto and Phatak, 2009). Homology modelling uses the three-dimensional structure of a protein that has been solved experimentally, as a template, to determine a model structure of the target protein. The structure of EcSspA is yet to be solved and as a result, the modelling of EcSspA was necessary. For a template to be considered eligible for homology modelling, a sequence identity of >30% (Cavasotto and Phatak, 2009) is required. The sequence identity found between EcSspA and YpSspA (PDB: 1YY7) was found to be 83% (Andrykovitch *et al*., 2003) and as a result, YpSspA was used as a template. The amino acid sequence of EcSspA was put into SWISS-MODEL (Waterhouse et al., 2018a), which is an online homology modelling server and using YpSspA as a template, SWISS-MODEL was able to compute a model for EcSspA. This model was used in subsequent docking studies.

Prediction of potential binding sites

SiteMap is a tool incorporated in the Schrödinger Maestro version 12.0 Software (Halgren, 2007), that identifies potential binding sites on target proteins with a high degree of confidence. SiteMap identifies site points that are likely to contribute to strong protein-protein and/or protein-ligand binding. Various factors are calculated at each site point, such as the size of the site, the degree to which the site is exposed to solvent, the degree to which the site is enclosed by the protein of interest as well as the tightness that the site interacts with the ligand, the degree that the ligand in question can donate or accept hydrogen bonds and the hydrophobic and hydrophilic character of the site (Halgren, 2007). Using these factors, SiteMap can rank the potential binding sites. SiteMap can be applied to whole proteins and was used to determine potential binding sites on the EcSspA homology model (Halgren, 2007). The sites obtained on the EcSspA homology model by SiteMap were used to set up the docking studies.

Docking of ligands to EcSspA homology model structure

Prior to docking IAA-94 and EA to the EcSspA model, both ligands and the EcSspA homology model underwent preparation for docking using the ligand and protein preparation wizard provided by the Schrödinger Maestro version 12.0. Subjecting the structures of IAA-94 and EA as well as the EcSspA homology model to preparation by the preparation wizards ensured that the structure of the ligands and protein are correct, and that missing hydrogen atoms or

incomplete side chains were fixed. The OPLS 2005 force field was used to estimate the forces present between atoms within molecules and between molecules and therefore reproduce accurate structures of the protein and ligands (Friesner *et al*., 2004).

After the preparation of IAA-94, EA and the EcSspA structures, docking commenced. Docking is a computational tool that helps identify whether a ligand is likely to bind to a target protein. The Induced Fit Molecular Docking (IFMD) protocol was used to predict whether IAA-94 and EA would bind to the EcSspA homology model, respectively. IFMD makes use of Glide and Prime, which are methods used in the docking process, to thoroughly consider all possible binding modes and the conformational changes that occur at the receptor binding site (Friesner *et al*., 2004). IFMD docks the ligand in question to the protein structure using Glide, which generates a grid box that encompasses the whole protein, producing a range of ligand poses. Thereafter, a Prime structure prediction fits the ligand by re-orientating side chains that are nearby (Friesner *et al*., 2004). The ligand is then re-docked into its corresponding protein structure of lowest energy and the protein-ligand complexes are ranked according to the scoring function, GlideScore, which separates ligand poses based on whether they strongly bound to the receptor or not (Friesner *et al*., 2004).

Chapter 3: Results

3.1 Verification of cDNA sequence encoding EcSspA

Through Sanger DNA sequencing, Inqaba Biotech™ (Pretoria, South Africa) showed that the pET-28a vector contained the cDNA sequence that encoded for EcSspA. The nucleotide sequence received from Inqaba Biotech™ was retrieved (Figure 7A) and using NCBI BLASTx, the translated EcSspA nucleotide sequence yielded a 99% sequence identity to *E. coli* SspA (Figure 7B).

3.2 Overexpression and purification of EcSspA

3.2.1 Overexpression of EcSspA

Prior to inducing the overexpression of EcSspA with IPTG, a growth curve was established (Figure 8). The induction took place when the T7 *E. coli* cells reached their exponential growth phase, with an OD_{600} of approximately 0.7.

3.2.2 Purification of EcSspA

Subsequent to overexpression, the soluble fraction of the cell lysate was then applied onto a Ni^{2+}-IMAC column and the effluent was measured at an absorbance of 280 nm (A_{280}). The A_{280} of the effluent produced a large broad peak indicative of the flow-through, after the elution buffer was applied to the column a single peak was observed, which is indicative of the eluted EcSspA (see A-D in Figure 9). Three fractions of the eluted EcSspA were collected and eventually pooled together.

3.2.3 Verification of the purity of EcSspA

The MWM and pooled EcSspA fractions (**Pr**) were all resolved on a 12.5% glycine SDS-PAGE gel (Figure 10A). The Pr lane shows a thick, single band with no protein contaminants present, indicating that only EcSspA was present in the eluted fractions. The SDS-PAGE calibration curve (Figure 10B) was used to calculate the subunit size of EcSspA. Based on the calibration curve, the subunit size of EcSspA was calculated to be 25.8 kDa using the BLUeye Prestained Protein Ladder (GeneDireX, Inc.), which served as a MWM.

A

>SspA_T7_C04_07

TTCCCCCYTCCTAGAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATC
ATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTGTCGCTGCCAACAAACGTTC
GGTAATGACGCTGTTTTCCGGTCCTACTGACATCTATAGCCATCAGGTCCGCATTGTGCTGGCTGAG
AAAGGTGTAAGTTTCGAGATCGAACACGTGGAAAAGGACAATCCGCCTCAGGATCTGATTGACCTC
AACCCGAATCAGAGCGTTCCGACCCTGGTGGATCGTGAGCTGACCCTGTGGGAATCTCGCATCATTA
TGGAATATCTGGATGAGCGTTTCCCGCATCCGCCACTGATGCCTGTTTACCCGGTAGCTCGCGGTGA
AAGCCGTCTGTACATGCATCGCATCGAAAAAGACTGGTACACGCTGATGAACACCATCATCAACGGT
TCAGCTTCTGAAGCAGATGCCGCACGTAAGCAACTGCGCGAAGAACTGCTGGCGATTGCGCCGGTC
TTCGGTCAGAAGCCGTACTTCCTGAGCGATGAGTTCAGCCTGGTCGATTGCTATCTTGCTCCGCTGCT
GTGGCGTCTGCCGCAACTGGGCATCGAGTTCAGCGGCCCGGGTGCGAAAGAGCTGAAAGGCTATA
TGACCCGCGTCTTTGAGCGTGACTCTTTCCTTGCTTCTTTAACTGAAGCAGAACGTGAAATGCGTCTG
GGCCGGAGTTAAGCGGCCGCACTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAA

B

Score	Expect	Method	Identities	Positives	Gaps	Frame
436 bits(1121)	1e-151	Compositional matrix adjust.	211/212(99%)	212/212(100%)	0/212(0%)	+2

```
Query  110  MAVAANKRSVMTLFSGPTDIYSHQVRIVLAEKGVSFEIEHVEKDNPPQDLIDLNPNQSVP  289
            MAVAANKRSVMTLFSGPTDIYSHQVRIVLAEKGVSFEIEHVEKDNPPQDLIDLNPNQS+P
Sbjct  1    MAVAANKRSVMTLFSGPTDIYSHQVRIVLAEKGVSFEIEHVEKDNPPQDLIDLNPNQSIP  60

Query  290  TLVDRELTLWESRIIMEYLDERFPHPPLMPVYPVARGESRLYMHRIEKDWYTLMNTIING  469
            TLVDRELTLWESRIIMEYLDERFPHPPLMPVYPVARGESRLYMHRIEKDWYTLMNTIING
Sbjct  61   TLVDRELTLWESRIIMEYLDERFPHPPLMPVYPVARGESRLYMHRIEKDWYTLMNTIING  120

Query  470  SASEADAARKQLREELLAIAPVFGQKPYFLSDEFSLVDCYLAPLLWRLPQLGIEFSGPGA  649
            SASEADAARKQLREELLAIAPVFGQKPYFLSDEFSLVDCYLAPLLWRLPQLGIEFSGPGA
Sbjct  121  SASEADAARKQLREELLAIAPVFGQKPYFLSDEFSLVDCYLAPLLWRLPQLGIEFSGPGA  180

Query  650  KELKGYMTRVFERDSFLASLTEAEREMRLGRS  745
            KELKGYMTRVFERDSFLASLTEAEREMRLGRS
Sbjct  181  KELKGYMTRVFERDSFLASLTEAEREMRLGRS  212
```

Figure 7: Verification of the cDNA sequence encoding for EcSspA. (A) The cDNA sequence received from Inqaba Biotech™ (Pretoria, South Africa) showed that the cDNA sequence encoded for EcSspA. The sequence was viewed using FinchTV (Version 1.4.0, Geospiza Inc.) and saved as a text file. The start codon (ATG) of the cDNA sequence is indicated in blue, the open reading frame (ORF) is indicated in grey, the stop codon (TAA) is indicated in red, the codons representing the his-tag are shown in pink and the codons encoding the thrombin cleavage site is shown in green with a dotted black line where the cleavage is expected to take place. **(B)** BLASTx showed that the cDNA sequence received from Inqaba Biotech™ (Pretoria, South Africa) shared 99% sequence identity with EcSspA (Sequence ID: WP_097451043.1) and consists of 212 amino acid residues. The alignment further showed that there were 0% gaps between the query (orange box) and subject amino acid sequence (green box).

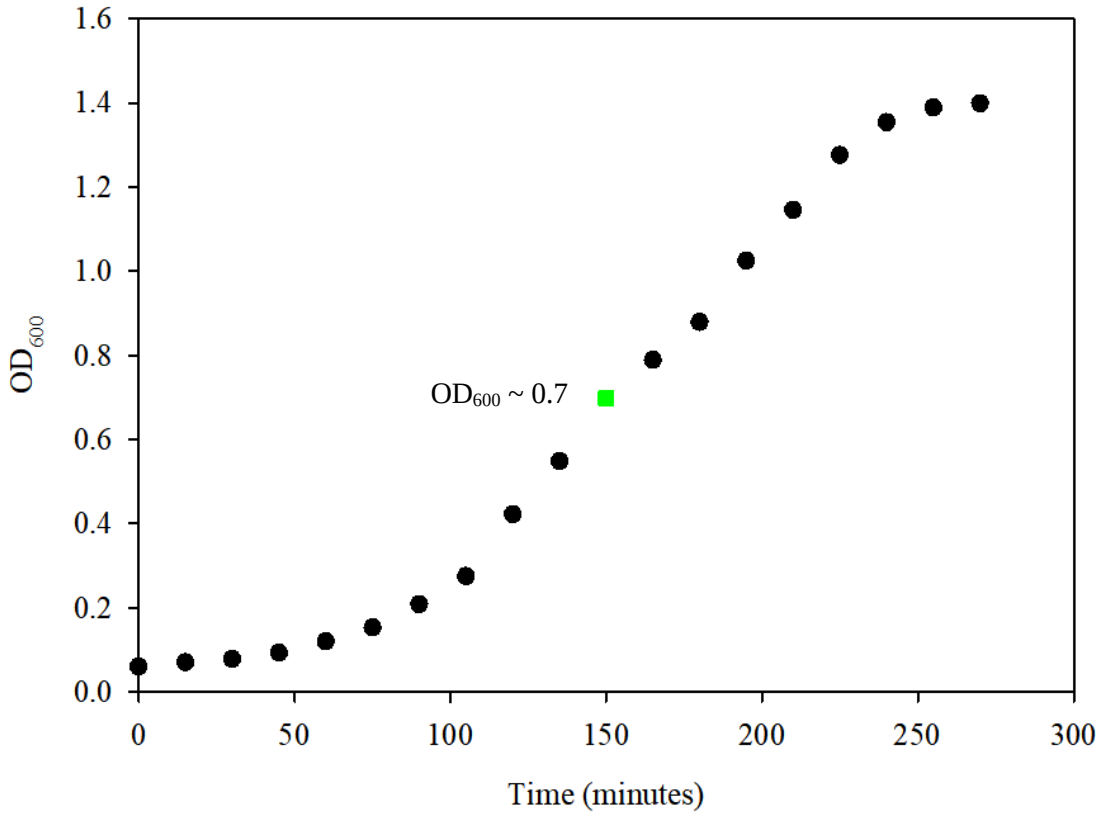


Figure 8: Growth curve of T7 *E. coli* cells. The growth curve shows that an OD_{600} of ~ 0.7 (green square) was reached after 150 minutes. To induce overexpression of EcSspA, a final concentration of 0.2 mM IPTG was added to the T7 *E. coli* cells upon reaching an OD_{600} of ~ 0.7. Overexpression of EcSspA occurred at 37˚C for 6-8 hours while rotating at 210 rpm.

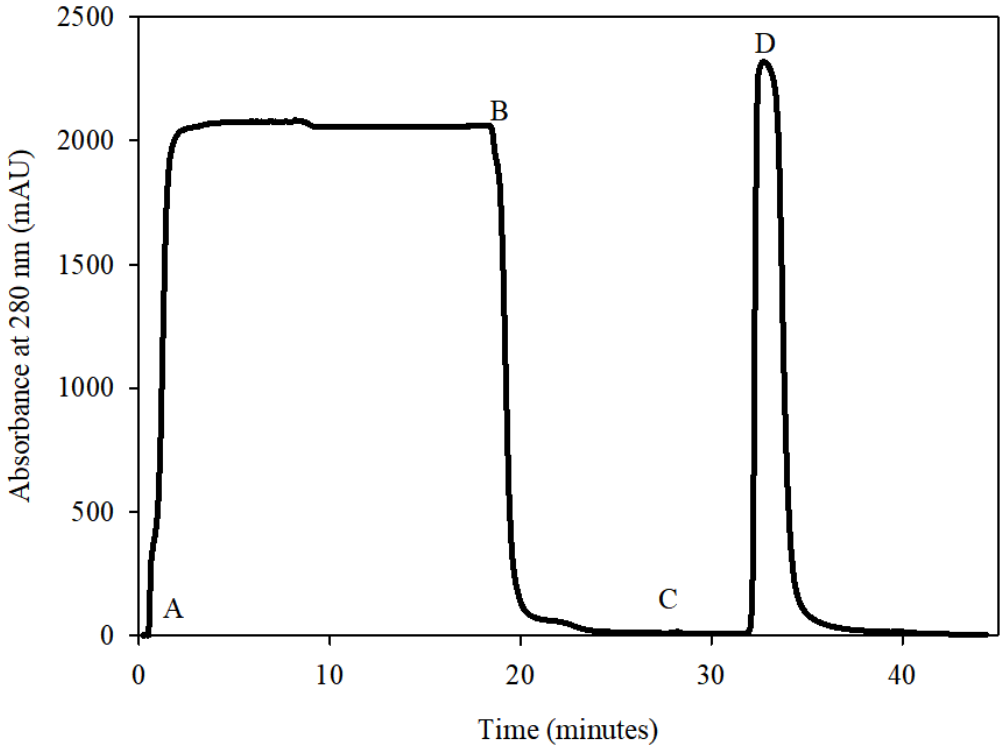


Figure 9: The purification elution profile of EcSspA. The elution profile of EcSspA shows the absorbance at 280 nm of the effluent, from the loading of the filtered supernatant to the elution of EcSspA. The large broad peak represents the loading of the supernatant (A) onto the Ni^{2+}-IMAC column, as well as introducing the equilibration buffer to wash off and remove any unbound molecules (B). The elution buffer, which consisted of 300 mM imidazole was applied to the column (C) and this resulted in the elution of EcSspA which is shown by the single symmetrical peak (D).

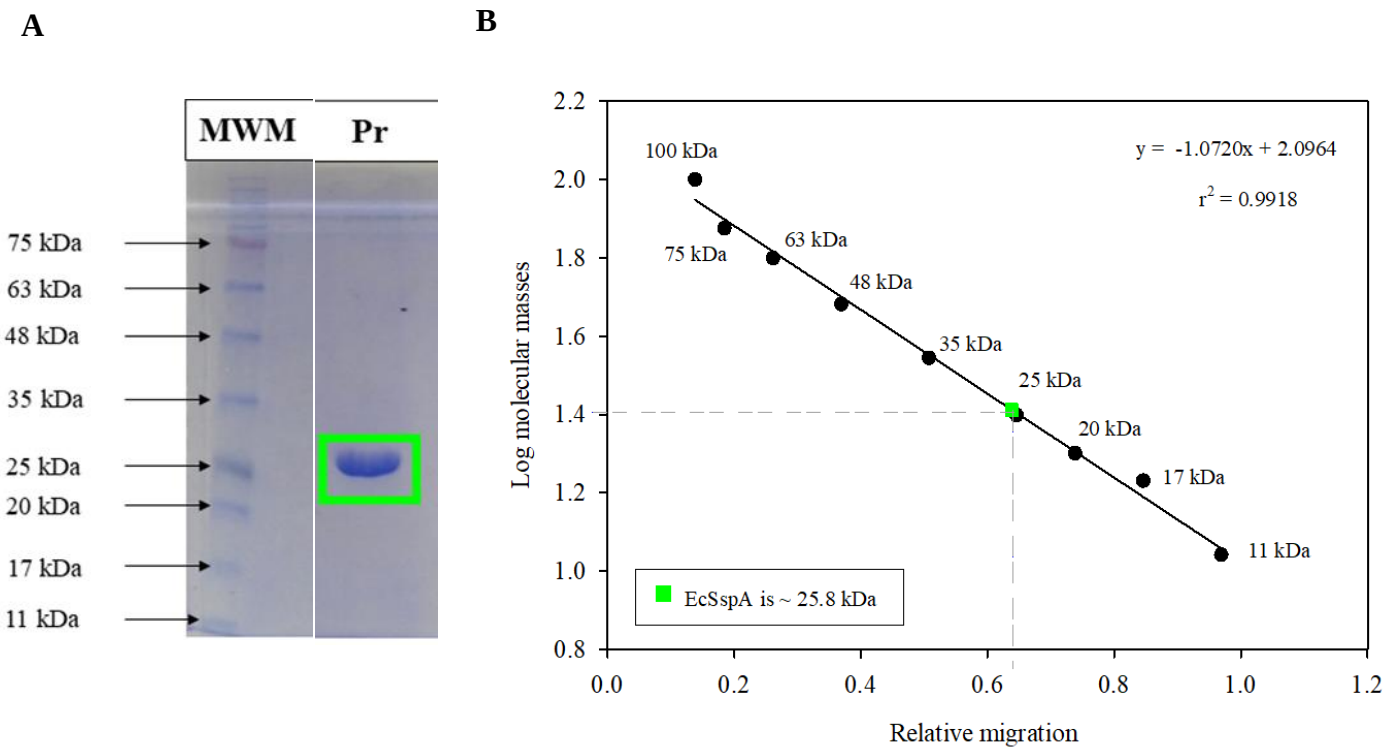
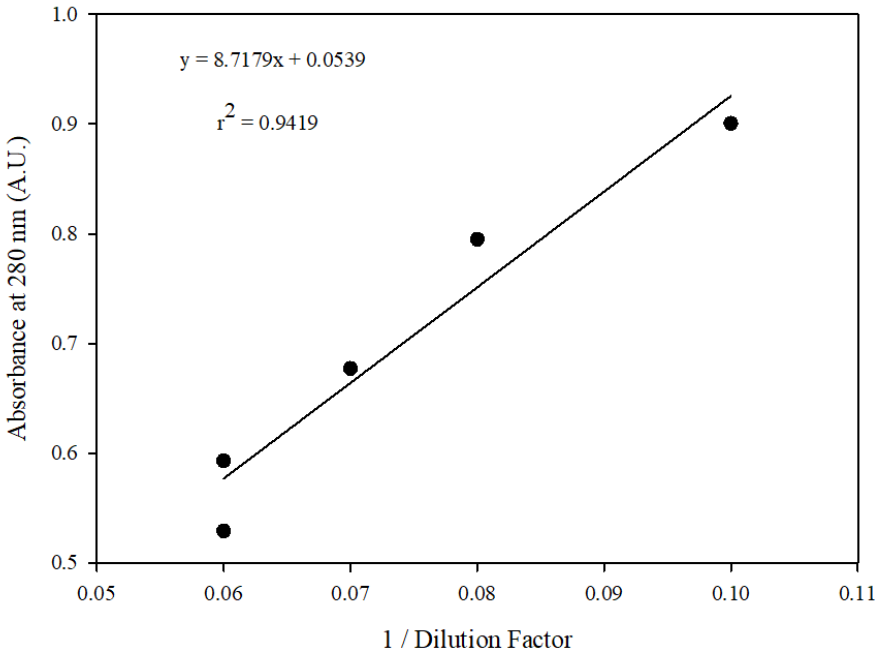


Figure 10: Verification of the presence and purity of EcSspA. (A) The 12.5% (w/v) SDS-PAGE gel was used to resolve samples of the effluent fractions collected during purification of EcSspA. The **Pr** lane shows a thick single band (green box) with no visible bands above or beneath it, indicating that EcSspA was pure. **(B)** The 12.5% (w/v) SDS-PAGE gel was also used to determine the subunit size of EcSspA using the SDS-PAGE calibration curve. Using the SDS-PAGE calibration curve with a linear regression model, the subunit size of EcSspA was determined to be 25.8 kDa, which is indicated on the curve by a green square.

3.3 Protein concentration determination and purity assessment

Subsequent to dialyzing purified EcSspA, its concentration was determined by constructing a dilution plot (Figure 11A), as well as using the Beer-Lambert Law (see equation 1) at an absorbance of 280 nm. The concentration was calculated to be ~164 µM. When using the measured A_{280} from the absorbance spectrum and calculating the concentration of EcSspA using the Beer-Lambert Law (see equation 1), the concentration of EcSspA is ~ 162 µM (not shown). This served as confirmation of the concentration calculated using the dilution plot. The A_{260}/A_{280} ratio calculated from the absorbance spectrum (Figure 11B) was found to be 0.488. The peaks present at 280 nm and 295 nm are indicative of tyrosine and tryptophan residues present in protein samples respectively. Whereas a peak present at 340 nm would show the presence of protein aggregates, however, the absorbance at 340 nm is ~ 0 A.U. indicating that no protein aggregates were present (Figure 11B).

A



B

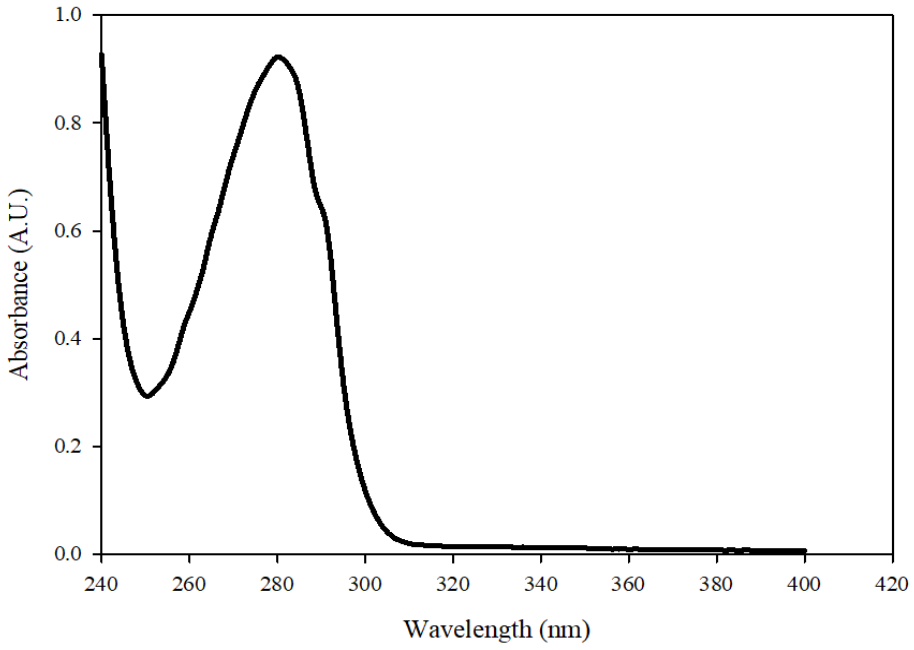


Figure 11: EcSspA concentration determination and purity assessment. (A) Dilution series plot of EcSspA measured at A_{280} and Beer-Lambert Law (see equation 1) used to determine the concentration. The gradient for the line of best fit represents the absorbance value for the protein sample and it was reported to be 8.7179 units and consequently the concentration of EcSspA was determined to be 164 µM. **(B)** The absorbance spectrum with a wavelength range of 240 – 400 nm of the EcSspA sample showed a A_{260}/A_{280} ratio of 0.487, indicating that EcSspA was free of nucleic acid contaminants.

3.4 Structural Characterisation of EcSspA

The structural elements of EcSspA were analysed using various techniques to ensure that the protein had maintained its structural integrity. The secondary structural elements were characterised using far-UV CD spectropolarimetry, the tertiary structure was characterised using intrinsic tryptophan fluorescence spectroscopy as well as near-UV CD spectropolarimetry. The quaternary structure of EcSspA was characterised using size exclusion-high performance liquid chromatography (SE-HPLC).

3.4.1 Far-UV CD spectropolarimetry

The secondary structure of EcSspA was characterised and analysed using far-UV CD spectropolarimetry. The far-UV CD spectrum of EcSspA (Figure 12A), displayed two characteristic troughs at 208 and 222 nm, as well as a characteristic peak at 195 nm, which indicated that the secondary structure of EcSspA is predominantly α-helical. This is supported by Table 1, that indicated that YpSspA has a total α-helical content of 52%. Along with the far-UV CD spectrum, a HT spectrum was constructed and showed that there was an increase in the voltage signal while the far-UV CD spectrum was being conducted but that the voltage remained below 700 V (Figure 12B).

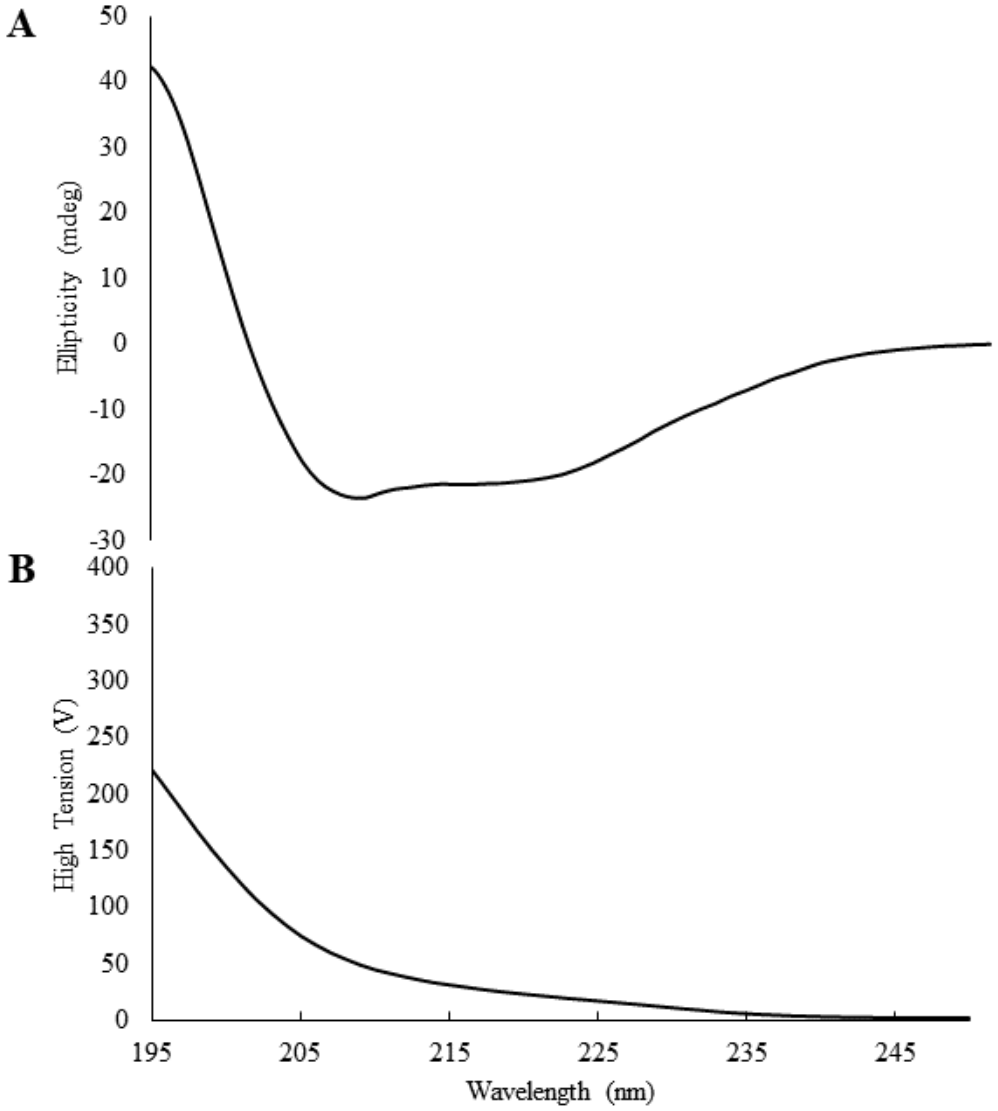


Figure 12: Far-UV CD and HT spectra of EcSspA. (A) The far-UV spectrum of 2.5 µM of EcSspA shows a minimum at 208 and 222 nm, as well as a maximum at 195 nm, which indicated that EcSspA is predominantly α-helical. **(B)** The HT signal increased as the wavelength decreased but remained below 700 V, indicating that the data shown in **A** is reliable. All spectra represent an average of three samples obtained and corrected for contributions made by the buffer used at 20˚C.

Table 1: The total α-helical, 3_{10}-helical and β-strand content of YpSspA and GSTB1-1 based on number of amino acid residues present in each secondary structural element using the Full Protein Feature View from RCSB PDB (PDB ID: 1A0F; Nishida *et al*., 1998) (PDB ID: 1YY7; Hansen *et al*., 2005a;).

	YpSspA	**GSTB1-1**
Total α-helical content	110	96
Total 3_{10}-helical content	10	9
Total β-strand content	18	18
Total number amino acid residues in protein	213	201
Percentage (%) of total α-helical content of the protein	51.6%	47.7%
Percentage (%) of total helical content of the protein	56.3%	52.2%

3.4.2 Near-UV CD spectropolarimetry

Near-UV CD spectropolarimetry is a technique used to provide insight into the tertiary structure of proteins that consist of tryptophan, tyrosine and/or phenylalanine residues. The near-UV CD spectrum (Figure 13A) indicated that phenylalanine, tyrosine and tryptophan residues are present in the structure of EcSspA by producing a signal between 255-270 nm, 275-282 nm and 290-305 nm respectively. This was confirmed by the absorbance spectrum (Figure 13B) which showed a spectrum in the wavelength regions that phenylalanine, tyrosine and tryptophan residues absorbed UV light, with peaks being present at 280 nm and 295 nm, which are indicative of the presence of tyrosine and tryptophan residues.

The HT spectrum (Figure 13C) that corresponds with the near-UV CD spectrum produced, showed that the HT remained below 700 V during the acquisition of the near-UV CD data, although an increase in the HT signal was observed.

3.4.3 Intrinsic tryptophan fluorescence spectroscopy

Intrinsic tryptophan fluorescence spectroscopy is a technique that investigates the local environment of tryptophan residues present in proteins and thus provides information on the tertiary structure of the protein (Lakowicz, 2006). Upon the selective excitement at 295 nm, the tryptophan fluorescence spectrum showed that there were two maximum emission wavelengths for the tryptophan residues within the structure of EcSspA, at 331 nm and 338 nm (Figure 14). This indicated that there was a blue shift in the emission wavelength of EcSspA tryptophan residues in comparison to maximum emission wavelength of free tryptophan in an aqueous environment, which is 350 nm (Lakowicz, 2006).

3.4.4 SE-HPLC

SE-HPLC was used to characterise the quaternary structure of EcSspA by providing information on the oligomeric state of EcSspA. SE-HPLC assess the hydrodynamic volume of samples to give an indication of the oligomeric state of samples (Tayyab *et al*., 1991).

The SE-HPLC elution profile of EcSspA showed two sharp symmetrical peaks representing the elution of EcSspA (Figure 15A). The elution profile of the gel filtration standards (Bio-Rad Laboratories, USA) showed sharp peaks which were annotated as peaks A-E. A calibration curve with a linear regression model (Figure 15B) was constructed to determine the molecular mass of EcSspA. Using the calibration curve, the molecular mass of EcSspA was calculated to be 46.5 kDa.

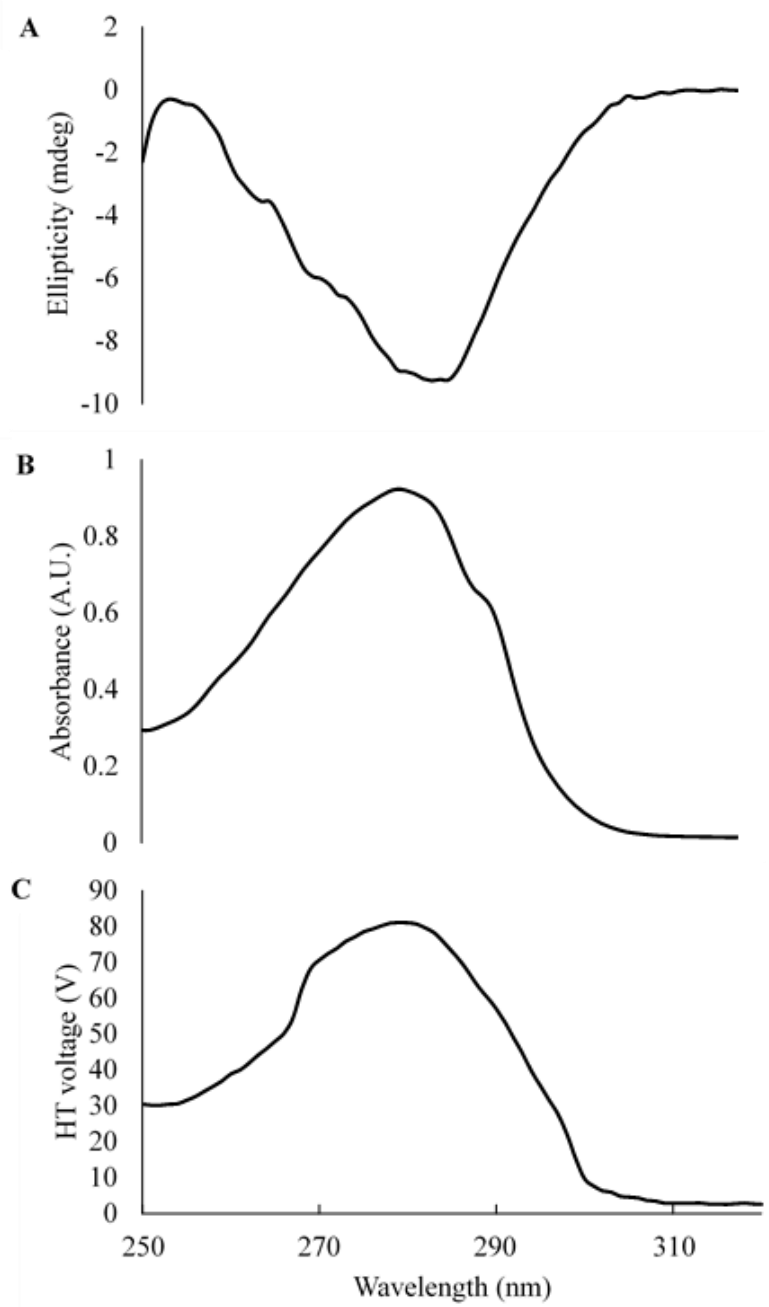


Figure 13: Near-UV CD, absorbance and HT spectra of EcSspA. (A) The near-UV spectrum of 20 µM of EcSspA showed the presence of phenylalanine (255 – 270 nm), tyrosine (275 – 282 nm) and tryptophan (290 – 305 nm) residues. **(B)** The presence of these aromatic residues is confirmed by the absorbance spectrum of EcSspA, that indicates that light is absorbed from 250 – 305 nm, corresponding to the signature wavelength ranges of phenylalanine, tyrosine and tryptophan residues as mentioned. **(C)** The HT spectrum showed an increase in the voltage, however the HT signal remained below 700 V. All spectra represent an average of three samples obtained at 20˚C and corrected for contributions made by Milli-Q water.

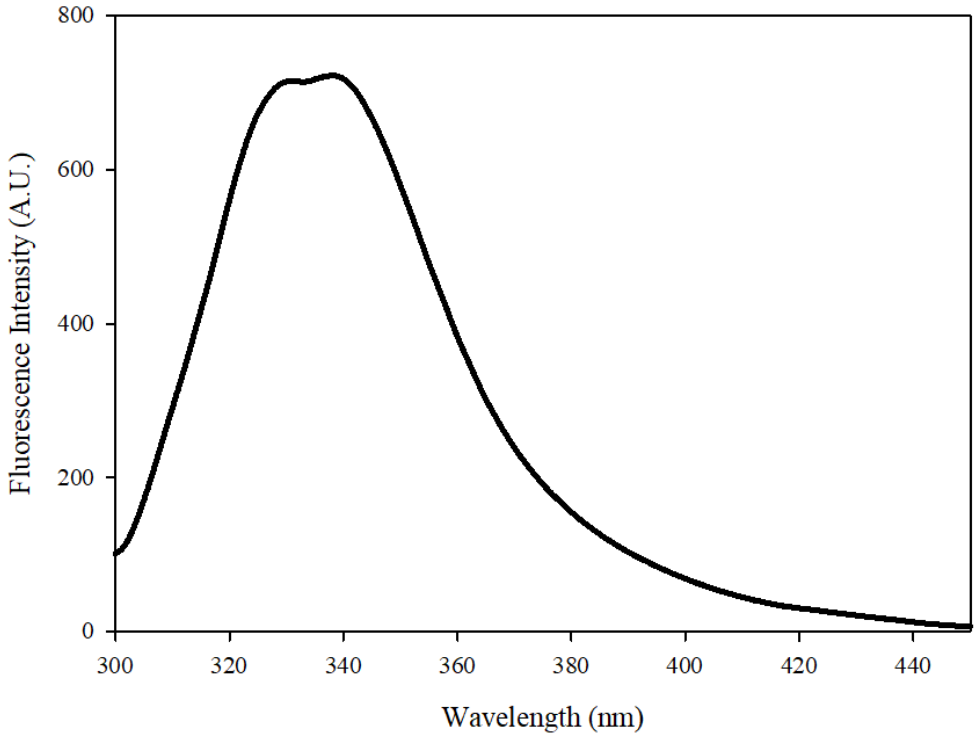
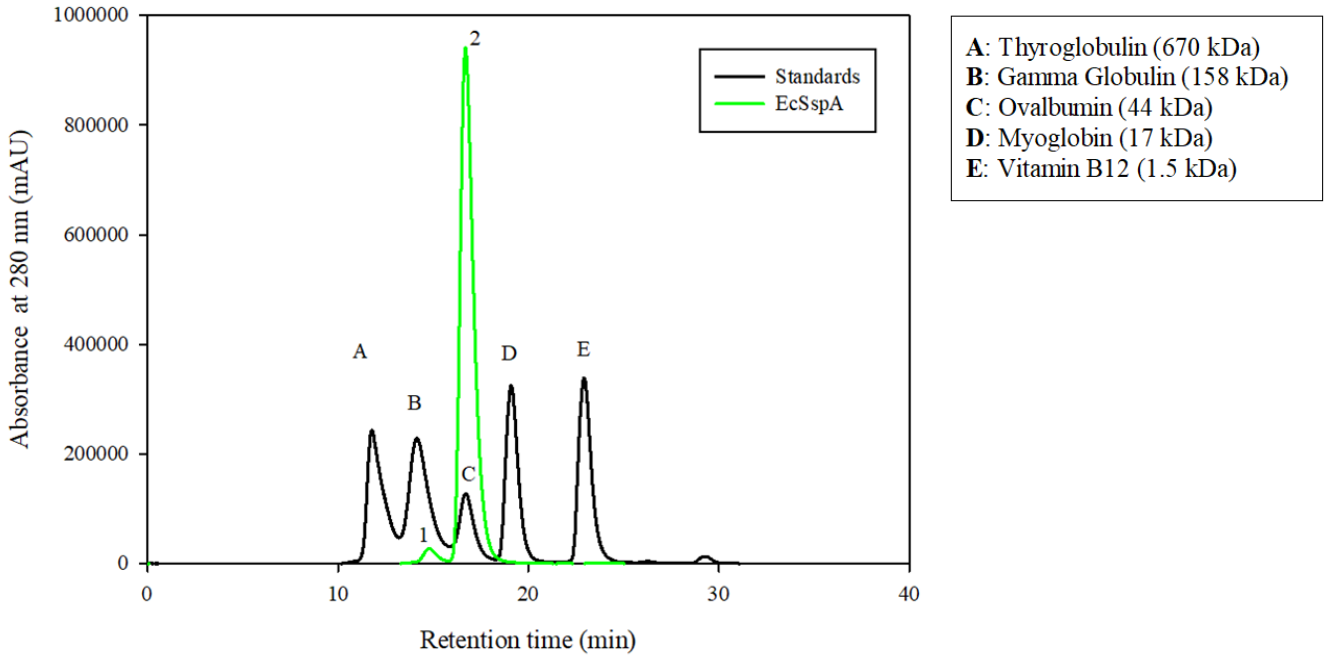


Figure 14: Tryptophan fluorescence emission spectrum of EcSspA. The fluorescence emission spectrum of 2 µM EcSspA showed that the tryptophan residues selectively excited at 295 nm, had a maximum emission wavelength of 338 nm. This emission spectrum represents an average of three samples obtained at 20˚C and corrected for contributions made by the buffer.

A



B

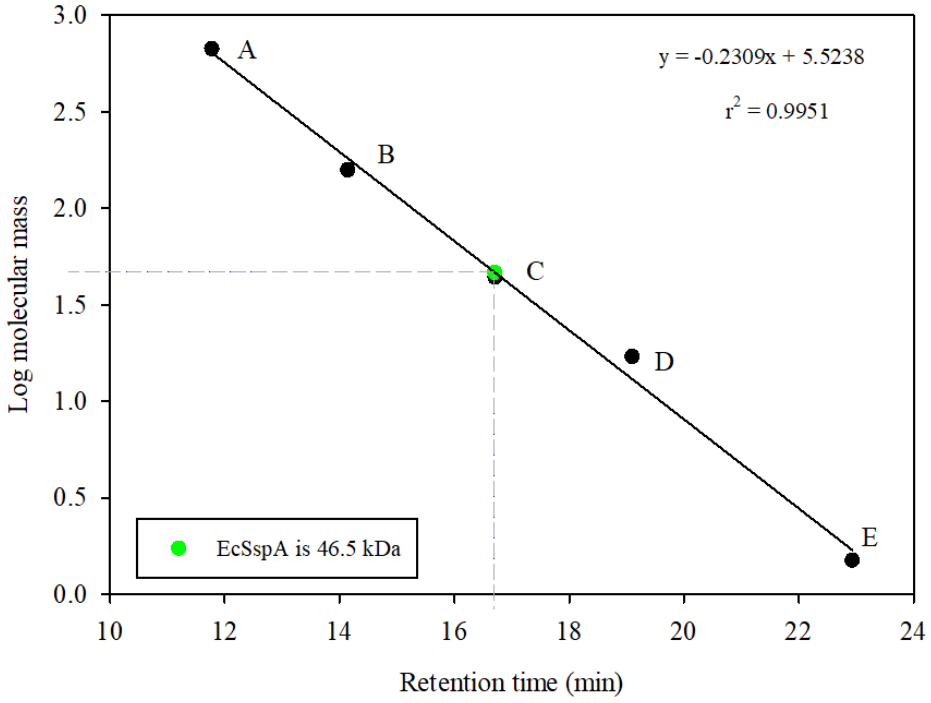


Figure 15: SE-HPLC elution profile of EcSspA and calibration curve. (A) SE-HPLC elution profile of 164 µM EcSspA indicated by the peak 1 and 2 as well as that of the gel filtration standards indicated as peaks A-E. All spectra represent an average of three samples obtained. **(B)** The SE-HPLC calibration curve was used to calculate the molecular mass of EcSspA under non-reducing conditions to be 46.5 kDa.

3.5 Assessment of IAA-94 and EA interaction with EcSspA

EcSspA plays a role in stationary phase induced stress responses, such as acid tolerance, in *E. coli* (Hansen *et al*., 2005b). These functional roles performed by EcSspA have been linked to the presence of a surface exposed pocket on the protein (see section 1.5). This surface exposed pocket has been postulated to be a potential protein binding site and therefore, it may serve as a target site to inhibit the function of EcSspA (Hansen *et al*., 2005a). As a result, being able to determine whether selected ligands can interact with EcSspA and where they could bind is important in order to assess if the surface exposed pocket is indeed a potential binding/target site. The surface pocket of EcSspA is predominantly hydrophobic (Hansen *et al*., 2005a) and it can be suggested that IAA-94 and EA could potentially bind to the surface pocket, considering their hydrophobic nature. As a result, conducting ANS displacement assays using IAA-94 and EA could potentially result in ANS being displaced from EcSspA by the ligands.

To investigate the interaction of IAA-94 and EA with EcSspA, the following techniques were used: the displacement of ANS dye by IAA-94 and EA from EcSspA, ITC as well as molecular docking of IAA-94 and EA to EcSspA.

3.5.1 ANS concentration determination and ANS extrinsic fluorescence

ANS is an amphiphilic, organic compound that can non-covalently bind to hydrophobic surfaces of protein and has been widely used to characterise binding sites of proteins, including GSTs (Ketley *et al*., 1975; Matulis and Lovrien, 1998). As a result, to determine whether ANS could bind to EcSspA, extrinsic ANS fluorescence was employed.

The concentration of ANS was determined by constructing a dilution plot (Figure 16) and using the Beer-Lambert Law (see equation 1) at an absorbance of 350 nm. The concentration was calculated to be approximately 77.5 mM (Figure 16). From the 77.5 mM stock solution of ANS, a final concentration of 200 µM was used for subsequent ANS fluorescence.

The fluorescence emission spectrum of free ANS (Figure 17) exhibited a maximum emission wavelength of 522 nm, whereas the fluorescence emission spectrum of ANS in the presence of EcSspA, exhibited a maximum emission wavelength of 482 nm; this was indicative of a blue shift in the maximum emission wavelength of ANS. Furthermore, ANS complexed with EcSspA had an increased fluorescence intensity in comparison to the fluorescence emission of free ANS, therefore showing that ANS did bind to the hydrophobic regions on the surface of EcSspA.

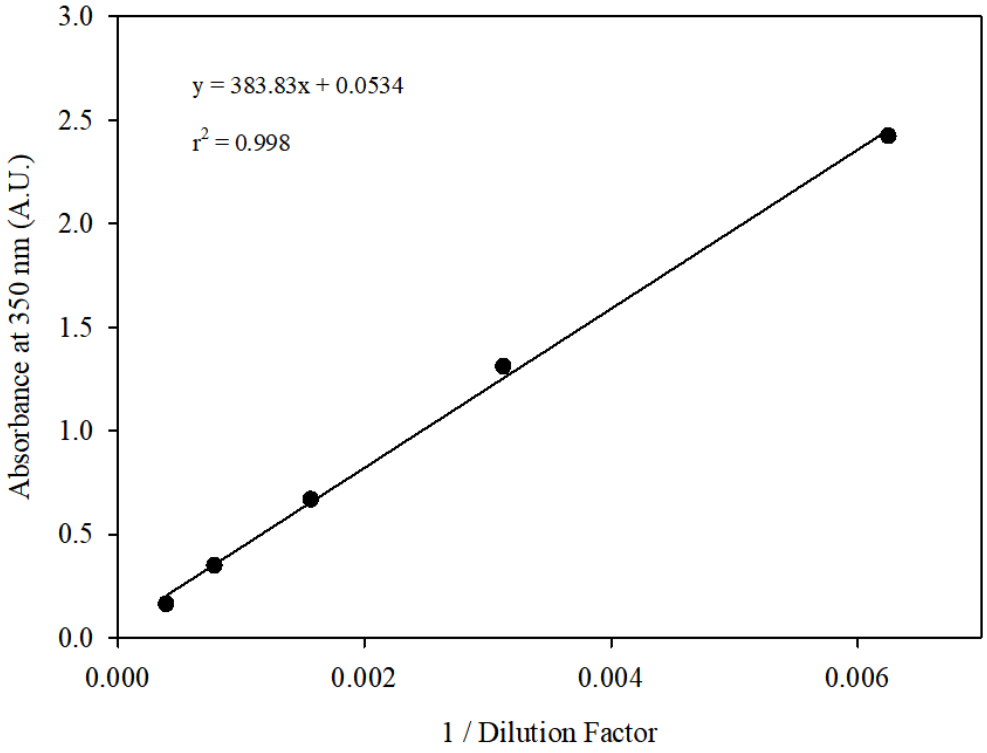


Figure 16: ANS concentration determination. Dilution plot of ANS measured at A_{350} and Beer-Lambert Law (see equation 1) used to determine the concentration of ANS. The gradient for the line of best fit is represented by the absorbance value and consequently the concentration of ANS was determined to be 77.5 mM.

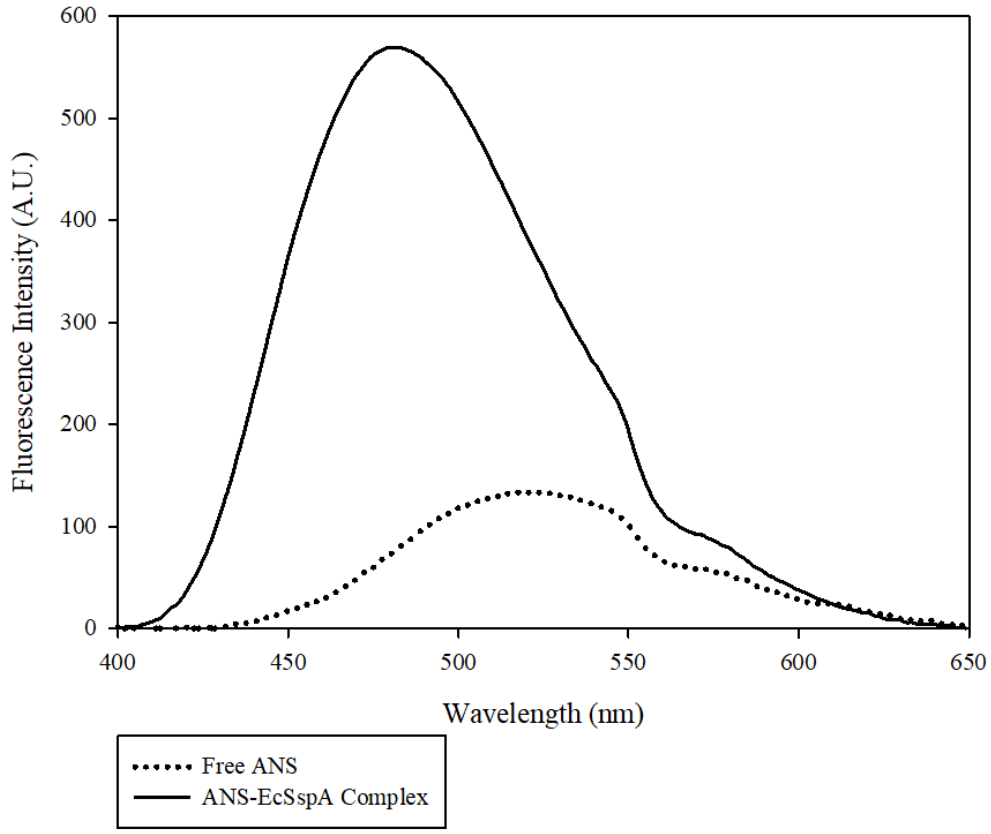


Figure 17: Fluorescence emission spectrum of ANS complexed with EcSspA. The spectrum indicated by the dotted line represents free ANS, which when excited at 390 nm, emitted maximally at a wavelength of 522 nm. The spectrum shown by the solid line represents ANS complexed with 2 µM EcSspA, which was also excited at 390 nm but emitted maximally at 482 nm, indicating a blue shift. The ANS-EcSspA spectrum also exhibited an increase in fluorescence intensity compared to the free ANS spectrum. All spectra represent an average of three samples obtained 20˚C and corrected for contributions made by the buffer used at. A final concentration of 200 µM of ANS was used.

3.5.2 Displacement of ANS by IAA-94 and EA from EcSspA using fluorescence spectroscopy

An ANS displacement assay was performed to observe the effect that the presence of IAA-94 and EA would have on the fluorescence emission of ANS bound to EcSspA and therefore, whether IAA-94 and EA could displace ANS bound to EcSspA.

When ANS was in the presence of DMSO (Figure 18), an increase in the fluorescence intensity was observed and the maximum emission wavelength experienced a blue shift in comparison to free ANS. Similar results were observed when ANS was in the presence of IAA-94 or EA only, as both ligands where solubilized in DMSO. Furthermore, ANS in the presence of DMSO and EcSspA yielded a greater fluorescence intensity, as well as a greater shift of the maximum emission wavelength to a shorter wavelength than what was observed when ANS was in the presence of DMSO only (Figure 18). This indicated that ANS did bind to the hydrophobic area(s) present on EcSspA in the presence of DMSO.

When IAA-94 was introduced to the EcSspA-ANS complex (Figure 18), there was a decrease in the fluorescence intensity of ANS. Similarly, when EA was added to the EcSspA-ANS complex, EA caused a decrease in the fluorescence intensity of ANS. However, EA did not decrease the fluorescence intensity of ANS to the extent in which IAA-94 did and as a result, this suggested that IAA-94 was able to displace more ANS molecules from EcSspA than what EA could. A final concentration of 2.1 mM IAA-94 displaced 60.9% of ANS molecules from EcSspA and that 2.48 mM EA displaced 15.3% of ANS bound to EcSspA.

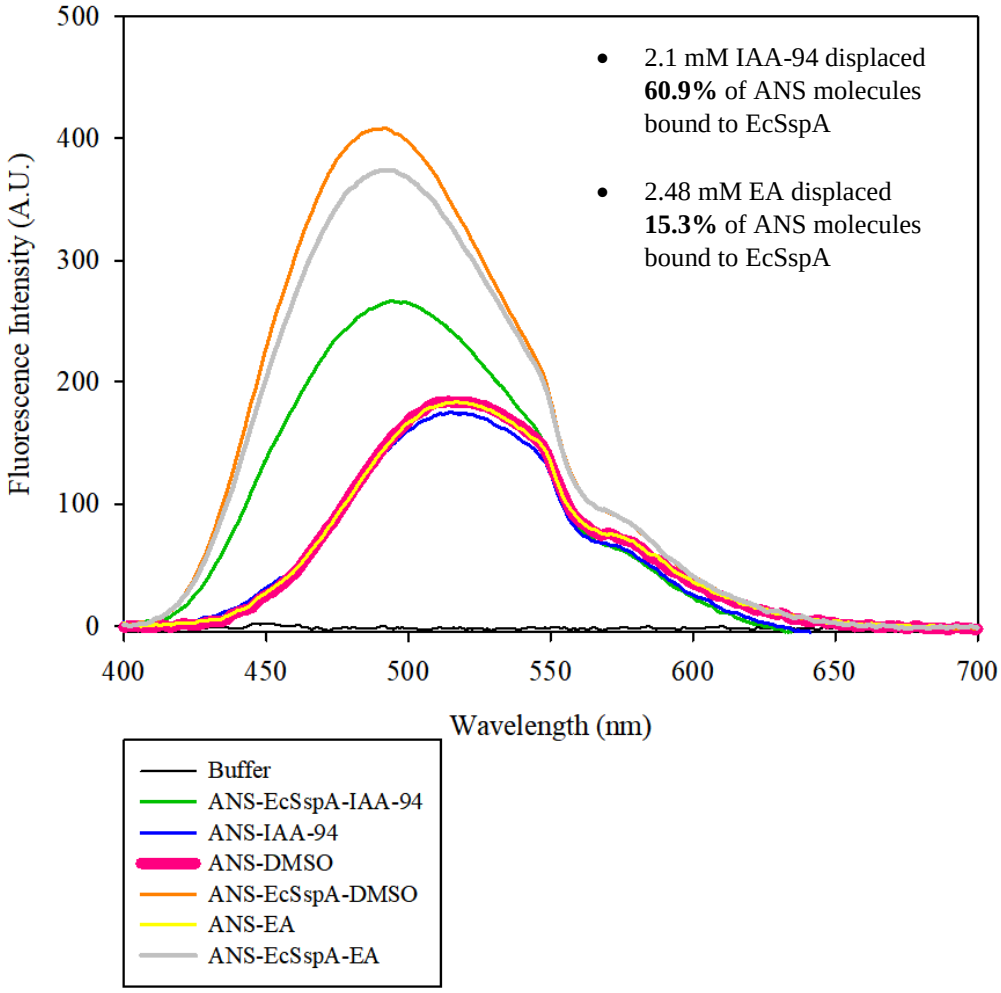


Figure 18: Fluorescence emission spectrum for the displacement of ANS by IAA-94 and EA from EcSspA. The ANS-DMSO spectrum (pink) served as a control for the ANS-IAA-94 spectrum (blue) and the ANS-EA spectrum (yellow). An increase in the fluorescence intensity of the ANS-DMSO spectrum was observed in comparison to the fluorescence intensity of ANS-IAA-94 spectrum, whereas the fluorescence intensity of the ANS-DMSO spectrum was similar to that of the ANS-EA spectrum. A decrease was observed in the fluorescence intensity of the ANS-EcSspA-IAA-94 spectrum (green) and the ANS-EcSspA-EA spectrum (grey), in comparison to the fluorescence intensity the ANS-EcSspA-DMSO spectrum (orange). It is shown that IAA-94 displaced 60.9% of ANS bound to EcSspA, whereas EA only displaced 15.3% and therefore showed that the fluorescence intensity of the ANS-EcSspA-IAA-94 spectrum had a greater decrease than that of the ANS-EcSspA-EA spectrum. All spectra represent an average of three samples obtained at 20˚C. A final concentration of 200 µM ANS, 2.1 mM IAA-94, 2.48 mM EA and 2 µM EcSspA were used in this study.

3.5.3 ITC

ITC was conducted to determine the thermodynamic parameters for the binding of IAA-94 and EA to EcSspA, respectively. The K_d, ΔH and N (stoichiomerty) are thermodynamic parameters that are directly obtained from the ITC isotherms and the ΔS, as well as ΔG values are obtained indirectly, through the use of equation 3 (Ladbury and Chowdhry, 1996).

The ITC titrations of IAA-94 into EcSspA (Figure 19A) showed that the K_d obtained from the interaction was 150 µM and the ΔH was -7.5 kcal/mol, the ΔS that was calculated as -8.1 cal/mol/K and the ΔG was calculated to be -5.1 kcal/mol. The ITC titrations of EA into EcSspA (Figure 19B) showed that the K_d obtained from the binding reaction was 531.9 µM and the ΔH was -18.2 kcal/mol, the ΔS was calculated as -47 cal/mol/K and the ΔG was calculated to be -4.4 kcal/mol. The ITC isotherms produced non-sigmoidal curves, with the error bars shown for each individual injection according to the fitting of the baseline by NITPIC. All ITC experiments were conducted at 20˚C (293 K) and the thermodynamic parameters obtained from the ITC titrations are further indicated in Table 2.

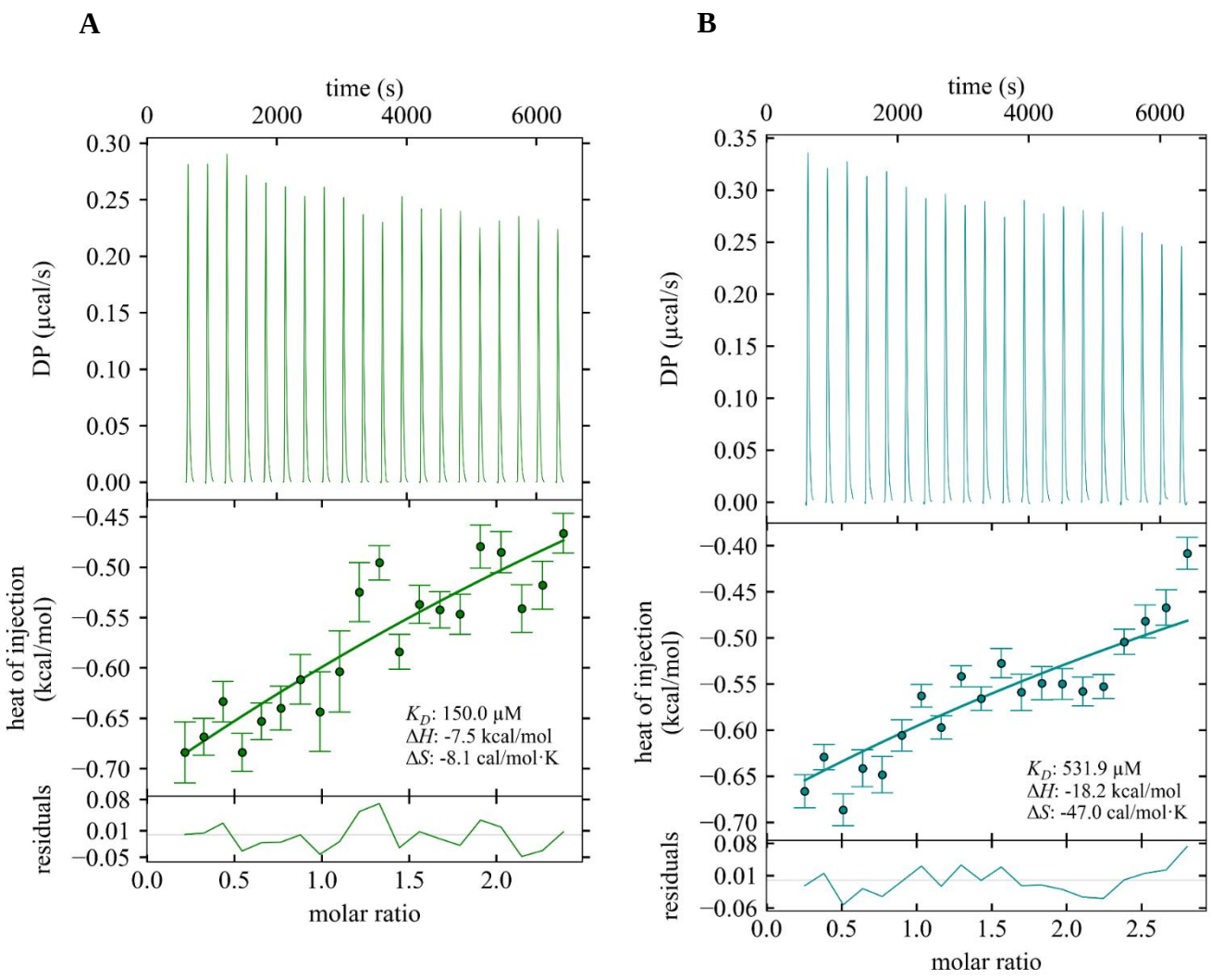


Figure 19: ITC titrations of IAA-94 and EA into EcSspA. (A) ITC graph indicating the isotherm and the integrated heats per injection, following the titration of 2.1 mM IAA-94. **(B)** ITC graph indicating the isotherm and the integrated heats per injection, following the titration of 2.48 mM EA. Both IAA-94 and EA were titrated into 100 µM of EcSspA at 20˚C. The heats of dilution were measured and subtracted from the baseline of the raw data prior to fitting by NITPIC (version 1.2.7) (Keller *et al.*, 2012; Scheuermann and Brautigam, 2015). The integrated heats were fitted using the A + B ↔ AB heteroassociation model implemented by SEDPHAT (version 15.2b). ITC graph created using GUSSI (version 1.4.2) (Brautigam, C.A., 2015). The error bars are shown for each injection in accordance to the fitting of the baseline by NITPIC.

Table 2: Thermodynamic parameters obtained from the titration of 2.1 mM IAA-94 and 2.48 mM EA into 100 µM EcSspA at 20˚C.

	IAA-94	**EA**
K_d (**µM**)	150	531
ΔG (**kcal/mol**)	-5.127	-4.429
ΔH (**kcal/mol**)	-7.5	-18.2
$T\Delta S$ (**kcal/mol**)	-2.37	-13.771

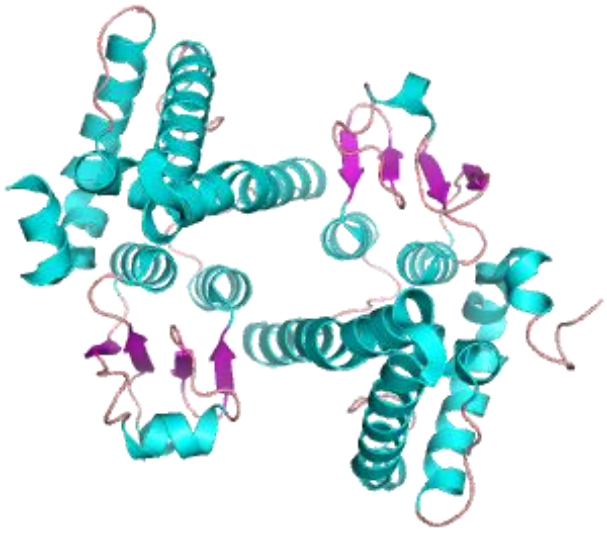
3.5.4 Molecular docking of IAA-94 and EA to EcSspA

Homology modelling

Through Swiss Model (Waterhouse *et al*., 2018), a homology model for EcSspA (Figure 20A) was made by using the crystal structure of YpSspA (PDB ID: 1YY7) as a template. The accuracy of the template and the homology model used was estimated using the global model quality estimation (GMQE) parameter (Waterhouse *et al*., 2018), which was calculated to be 0.94. The normalised qualitative model energy analysis (QMEAN) score, is also a parameter used to estimate the quality of the homology model in comparison to the QMEAN score of other experimentally determined structures of a similar size (Benkert *et al*., 2011). The normalised QMEAN score was calculated to be -0.30.

Swiss Model generated a local quality estimate plot, which shows the expected similarity that each residue in the homology model has to the native structure. The plot showed residues from chain A in red and those from chain B in blue. The plot indicated that most of the residues from the model scored above 0.6 (Waterhouse *et al*., 2018), emphasising that residues of the model are relatively similar to that of the native protein (Figure 20B). A multiple sequence alignment of the target sequence (EcSspA) against the top two ranked template sequences (YpSspA and RNAP) generated by Swiss Model indicated that both template sequences could be used to produce the EcSspA model and the sequence coverage of the templates was 99% for YpSspA and 100% for RNAP (not shown), which further emphasied the reliability of the model generated (Figure 21A). The 3-D superposition of the template structure (grey) and the EcSspA model (blue) was completed using PyMol version 2.2.0 (Figure 21B).

A



B

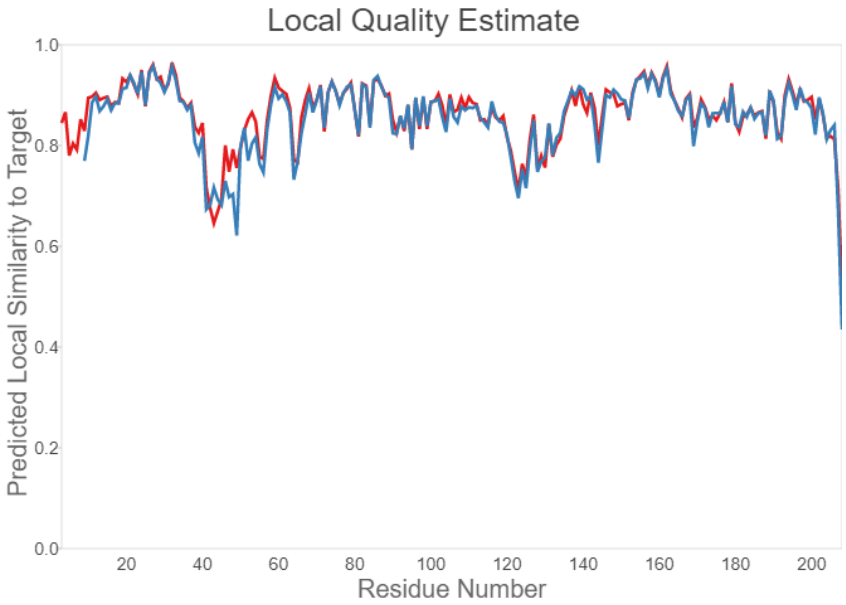
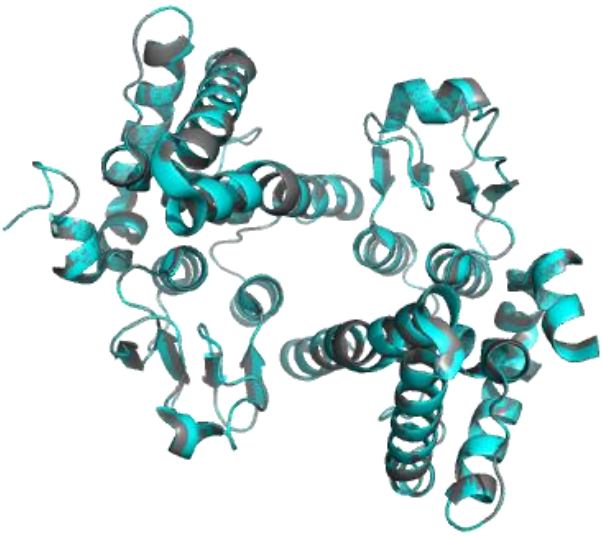


Figure 20: The homology model of EcSspA and the local quality estimate plot of the EcSspA model. (A) The homology model of EcSspA generated by Swiss Model with secondary structural elements highlighted: α-helices (blue), β-sheets (purple) and loops (pink). The structure of YpSspA was used as the template sequence (PDB: 1YY7). The image was created using PyMol version 2.2.0. **(B)** The Local Quality Estimate Plot generated by Swiss Model showing the expected similarity of each residue in the model to the native protein. The red and blue plots represent residues from chain A and B, respectively. The plots indicated that the model is of relatively good quality.

A



B

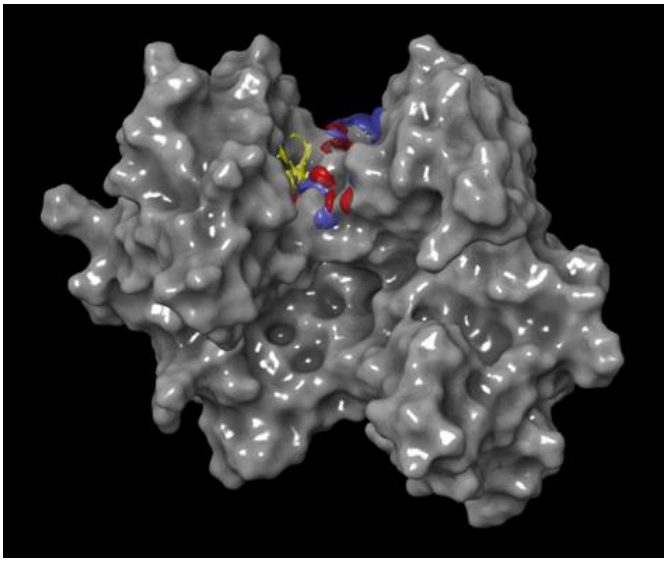
```
EcSspA      MAVAANKRSVMTLFSGPTDIYSHQVRIVLAEKGVSFEIEHVEKDNPPQDLIDLNPNQSVP   60
RNAP        MAVAANKRSVMTLFSGPTDIYSHQVRIVLAEKGVSFEIEHVEKDNPPQDLIDLNPNQSVP   60
YpSspA      MAVAANKRSVMTLFSGPTDIFSHQVRIVLAEKGVSVEIEQVEADNLPQDLIDLNPYRTVP   60
            ********************:*************.***:** ** ********* ::**

EcSspA      TLVDRELTLWESRIIMEYLDERFPHPPLMPVYPVARGESRLYMHRIEKDWYTLMNTIING   120
RNAP        TLVDRELTLWESRIIMEYLDERFPHPPLMPVYPVARGESRLYMHRIEKDWYTLMNTIING   120
YpSspA      TLVDRELTLYESRIIMEYLDERFPHPPLMPVYPVARGSSRLMMHRIEHDWYSLLYKIEQG   120
            *********:**************************.*** *****:***:*: .* :*

EcSspA      SASEADAARKQLREELLAIAPVFGQKPYFLSDEFSLVDCYLAPLLWRLPQLGIEFSGPGA   180
RNAP        SASEADAARKQLREELLAIAPVFGQKPYFLSDEFSLVDCYLAPLLWRLPQLGIEFSGPGA   180
YpSspA      NAQEAEAARKQLREELLSIAPVFNETPFFMSEEFSLVDCYLAPLLWRLPVLGIEFTGAGS   180
            .*.**:***********:*****.:.*:*:*:**************** *****:* *:

EcSspA      KELKGYMTRVFERDSFLASLTEAEREMRLGRS-   212
RNAP        KELKGYMTRVFERDSFLASLTEAEREMRLGRS-   212
YpSspA      KELKGYMTRVFERDAFLASLTEAEREMHLKTRS   213
            *************:************:*
```

Figure 21: The multiple sequence alignment of the EcSspA primary sequence against top two ranked template sequences and the superimposition of the template structure and the EcSspA model. (A) The multiple sequence alignment of the EcSspA amino acid sequence (UniProt accession number: P0ACA3) against that of YpSspA (UniProt accession number: A0A984KMD8) and an *E. coli* RNAP-promoter open complex (RPo) with SspA (UniProt accession number: H4XIP2) was generated using Clustal Omega (1.2.4) (Sievers *et al.*, 2011). The sequences were obtained from UniProtKB. The multiple sequence alignment produced had the amino acid residues colour coded: small and hydrophobic (red), acidic (blue), basic (pink), hydroxyl and sulfhydryl (green) residues. **(B)** A cartoon representation of the EcSspA model (blue) generated by Swiss-Model superimposed with the template structure (PDB: 1YY7, grey). The image was created using PyMol version 2.2.0.

Induced-fit molecular docking

Prior to molecular docking, Schrodinger SiteMap (Halgren, 2007) was used to determine the possible binding sites on the EcSspA homology model. Based on SiteMap, a total of 9 potential binding sites were predicted, however only the first 4 sites which yielded the highest SiteScore and largest volume are shown (Figure 22). Entry ID 43 was the best site as determined by SiteMap.

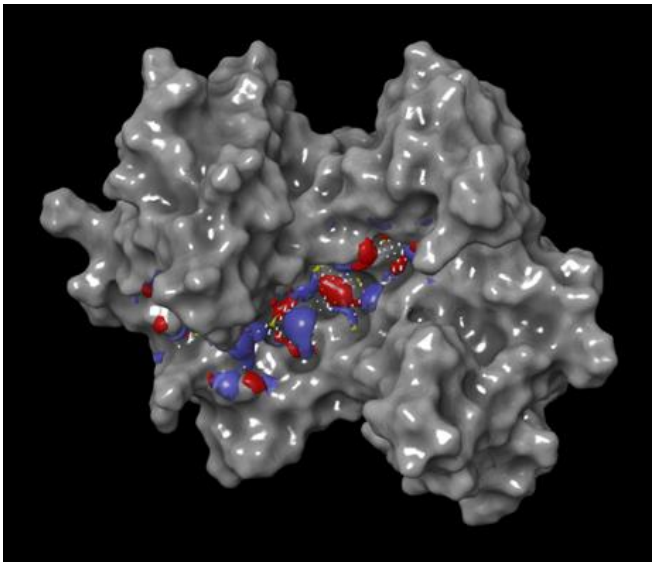
IFMD docked both IAA-94 (Figure 23A) and EA (Figure 24A) at entry ID 43, which is at the dimer interface of EcSspA. The two-dimensional representation of the interaction between IAA-94 and EcSspA (Figure 23B) showed that a chlorine atom formed a halogen bond with Lys108 and that a hydrogen bond was formed between the carboxyl group of IAA-94 and Tyr160. The two-dimensional representation of the interaction between EA and EcSspA (Figure 24B) showed that a halogen bond was formed between a chlorine atom and Arg73 from chain B of EcSspA and that hydrogen bonds by Tyr160 from chain A and B, Arg73 from chain A and Glu107 from chain B with EA.



A: Binding site 1 (Entry ID 42)

SiteScore: 1.005

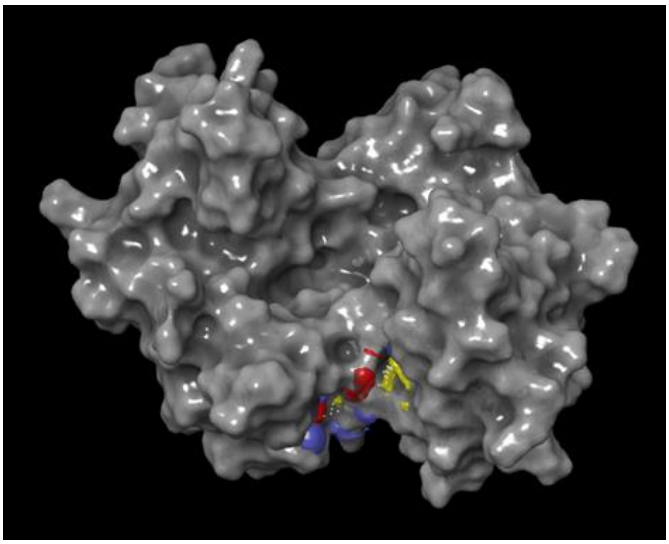
Volume: 341.285 units



B: Binding site 2 (Entry ID 43)

SiteScore: 1.002

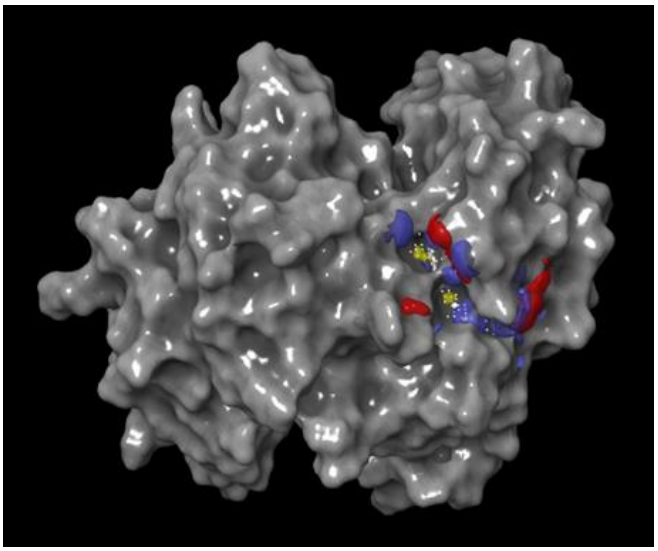
Volume: 553.945



C: Binding site 3 (Entry ID 44)

SiteScore: 0.827

Volume: 156.065 units



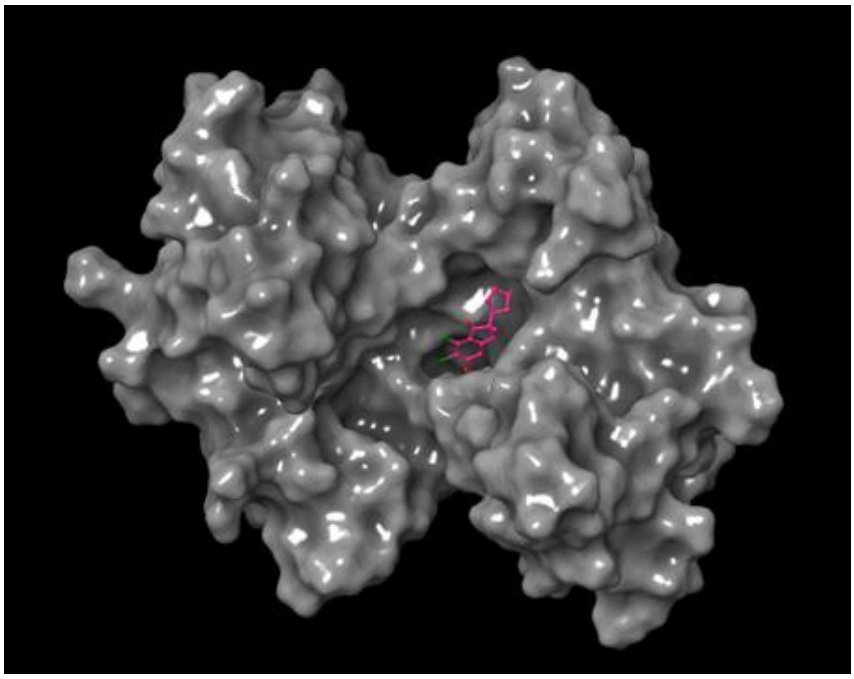
D: Binding site 4 (Entry ID 45)

SiteScore: 0.796

Volume: 159.495

Figure 22: The four highly ranked binding sites predicted by SiteMap. Entry ID 42, 43, 44 and 45 are the binding sites with the highest ranking SiteScores and the largest volumes of all the binding sites predicted by SiteMap. Entry ID 43 has the largest binding site which is found at the dimer interface of EcSspA.

A



B

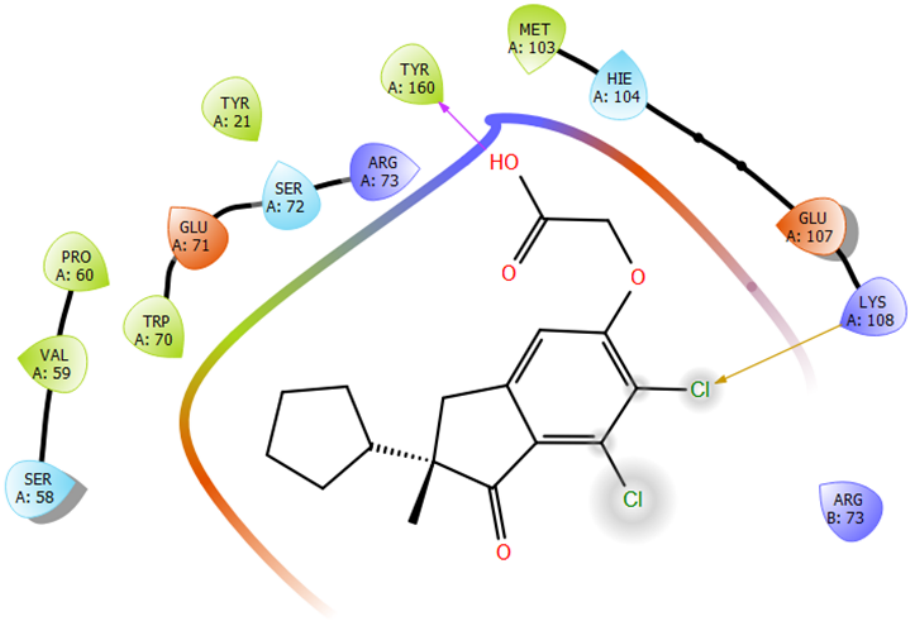
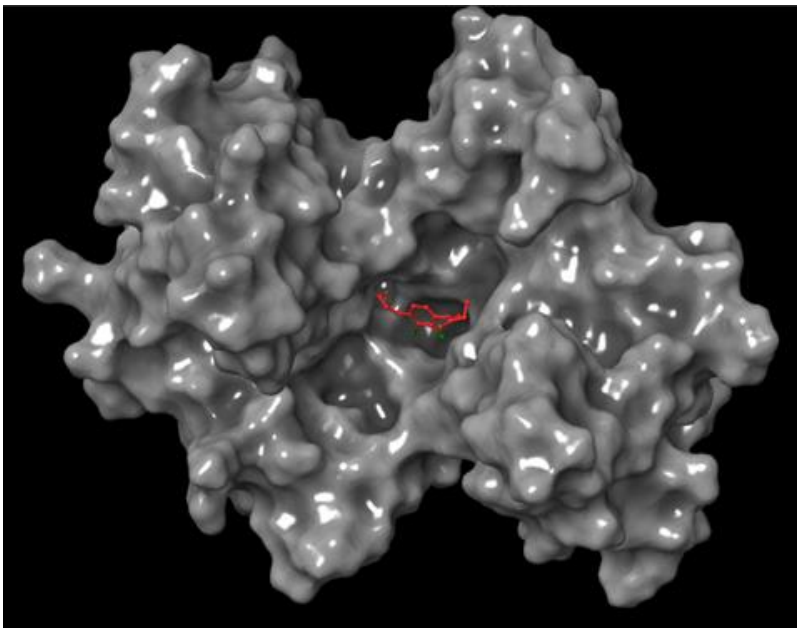


Figure 23: Docking of IAA-94 to EcSspA homology model. (A) The molecular surface of EcSspA (grey) shown with a ball-and-stick representation (pink) of IAA-94 docked to the dimer interface of EcSspA. **(B)** The 2-D image of the interaction of IAA-94 with the residues: Lys108 and Tyr160 forming a halogen bond and a hydrogen bond at the dimer interface respectively. The black line represents the polypeptide backbone, the pink arrow indicates a hydrogen bond and the gold arrow represents a halogen bond.

A



B

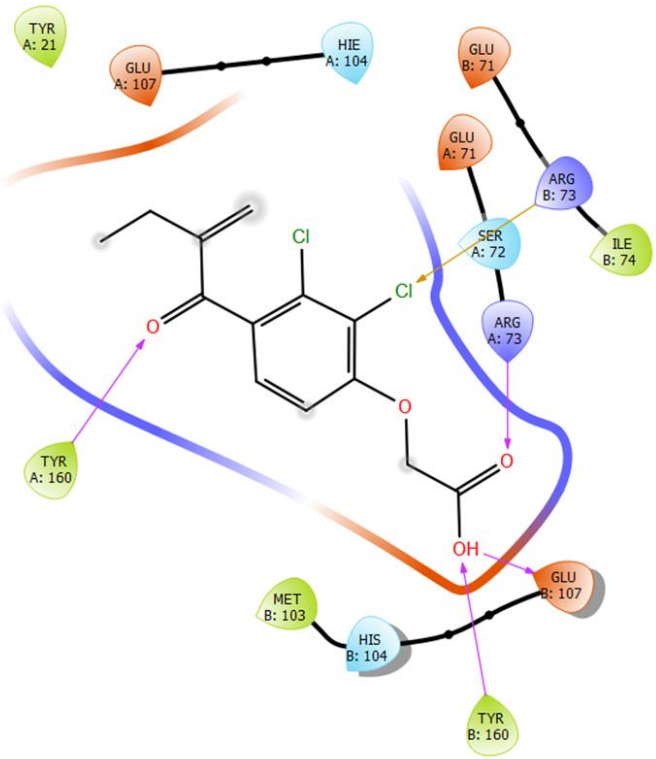


Figure 24: Docking of EA to EcSspA homology model. (A) The molecular surface of EcSspA (grey) shown with a ball-and-stick representation (red) of EA docked to the dimer interface of EcSspA. **(B)** The 2-D image of the interaction of Arg73 (chain B) forming a halogen bond and Tyr160, Arg73 and Glu107 forming hydrogen bonds at the dimer interface respectively. The black line represents the polypeptide backbone, the pink arrow indicates a hydrogen bond and the gold arrow represents a halogen bond.

Chapter 4: Discussion

This study aimed to investigate the interaction of IAA-94 and EA with soluble EcSspA. This aim was accomplished by first attaining EcSspA through inducing overexpression and purification. SDS-PAGE was then employed to verify the purity of EcSspA after purification. The structural characterisation of EcSspA followed and was achieved using UV-Visible spectroscopy, far and near-UV CD spectropolarimetry, intrinsic tryptophan fluorescence spectroscopy and SE-HPLC. Structural characterisation of EcSspA was important to ensure that the structural integrity of protein was intact, prior to assessing whether soluble EcSspA could bind IAA-94 and EA. To assess the interaction of IAA-94 and EA with EcSspA, ANS displacement assays were performed using ANS fluorescence spectroscopy as well as ITC and IFMD.

4.1 The overexpression and purification of EcSspA

The successful overexpression and purification of EcSspA, yielded his-tagged EcSspA from the soluble fraction obtained after cell harvesting. The single symmetrical peak shown in the purification elution profile indicated that EcSspA was eluted uniformly and the A_{260}/A_{280} ratio of the purified EcSspA sample indicated that EcSspA was free of nucleic acid contamination. Furthermore, the thick single band shown on the SDS-PAGE gel, with no additional bands present in lane Pr suggested that EcSspA was pure and free of protein contaminants of different molecular masses. Although the Sanger sequencing from Inqaba Biotech^{TM} indicated that the protein sample used in this study was EcSspA, to directly determine the true identity of the purified EcSspA sample, this study could have employed mass spectrometry techniques to further verify the EcSspA sample.

Additionally, under reducing conditions, the molecular mass of the recombinant EcSspA subunit is 25.8 kDa and the theoretical molecular mass of the EcSspA subunit is 24.3 kDa, as given by the ExPASy ProtParam Tool (Gasteiger *et al*., 2005). The theoretical molecular mass corresponds to the molecular mass of 24.5 kDa determined by Hansen *et al* (2005) (Hansen *et al*., 2005; Gururaja Rao *et al*., 2017). The his-tag present on purified EcSspA was possibly the cause of the difference in the theoretical and calculated molecular mass of the EcSspA subunit. His-tags have a size range of 1 - 2.5 kDa depending on the length of the his-tag used. In many instances, the relatively small size of the his-tag on a protein can be considered to have little effect on the three dimensional conformation of proteins (Kimple *et al*., 2013), although, his-tags have been shown to affect the dimerization of some proteins (Wu and Filutowicz, 1999). In the case of the EcSspA sample obtained, however, no aggregation occurred in downstream

experiments to show that the structure of EcSspA was compromised and the structural elements of the recombinant EcSspA indicated that the structural integrity was preserved. As a result, the his-tag attached to EcSspA remained uncleaved.

This showed that EcSspA was successfully purified and that the EcSspA samples used in downstream experiments were relatively free from any apparent contamination.

4.2 The structural characterisation of EcSspA

Secondary structure characterisation

The far-UV CD data indicated that EcSspA was predominantly α-helical, which was illustrated by the characteristic troughs at 208 nm and 222 nm as well as a peak at 195 nm (Kelly *et al*., 2005). The spectral data can be considered reliable, as the HT data, indicated that the voltage remained below 700 V, although an increase in the voltage was detected. Voltage readings below 700V suggest that sufficient light reached the detector of the CD spectrometer to ensure reliable CD data acquisition (DiNitto and Kenney, 2012).

Based on the high sequence identity between EcSspA and YpSspA, the predominantly α-helical secondary structure of EcSspA correlates with the secondary structure reported for YpSspA (Hansen *et al*., 2005a). The secondary structural content of YpSspA showed that α-helices make up approximately 52% of the YpSspA secondary structure (Table 1) using the Full Protein Feature View from RCSB PDB (PDB ID: 1YY7; Hansen *et al*., 2005).

Tertiary structure characterisation

Phenylalanine, tyrosine and tryptophan residues absorb UV light between 255-270 nm, 275-282 nm and 290-305 nm, respectively (Kelly and Price, 2009). With this being said, the near-UV CD data obtained showed that a spectrum was produced from 255-305 nm, suggesting that phenylalanine, tyrosine and tryptophan residues, are present in EcSspA. Similarly, the absorbance profile of EcSspA, revealed a spectrum from 255-305 nm, with the presence of peaks at 280 and 295 nm, indicating that phenylalanine and especially, tyrosine and tryptophan residues (Kelly *et al*., 2005; Lakowicz, 2006) are present in EcSspA. The absorbance profile of EcSspA therefore corresponds with the near-UV spectrum of EcSspA, as the absorbance of UV light occurs at the same wavelengths. The near-UV spectral data can be considered reliable as the HT data obtained, remained well below 700 V (DiNitto and Kenney, 2012) throughout the data acquisition in the near-UV region.

Additionally, the intrinsic tryptophan fluorescence spectrum of EcSspA showed that the tryptophan residues present in EcSspA are found to be buried within the protein structure. The maximum emission wavelength of free tryptophan in an aqueous environment is observed at 350 nm (Lakowicz, 2006), whereas the maximum emission wavelengths observed for tryptophan residues in EcSspA are 331 nm and 338 nm; this indicated that the maximum emission wavelength of the EcSspA tryptophan residues experienced a shift to a shorter wavelength, suggesting that they are located in a more hydrophobic environment (Lakowicz, 2006), such as the interior of the EcSspA structure compared to tryptophan residues found in an aqueous environment.

The location of the tryptophan residues on the molecular surface of EcSspA (Figure 25) was further evaluated and showed that Trp110 and Trp166 are more buried within the structure of EcSspA and thus are less solvent accessible than Trp70. Trp70 appears more clearly on the surface of EcSspA than Trp110 and Trp166, this suggests that Trp70 is partially exposed to the solvent environment. With this in consideration, the Trp70 residues could be responsible for the peak observed at 338 nm and the Trp110 and Trp166 could be the cause of the peak observed at 331 nm of the tryptophan fluorescence spectrum, due to the difference noted in their local environments, with Trp70 being partially exposed to the environment and Trp110 as well as Trp166 being less exposed to the solvent environment.

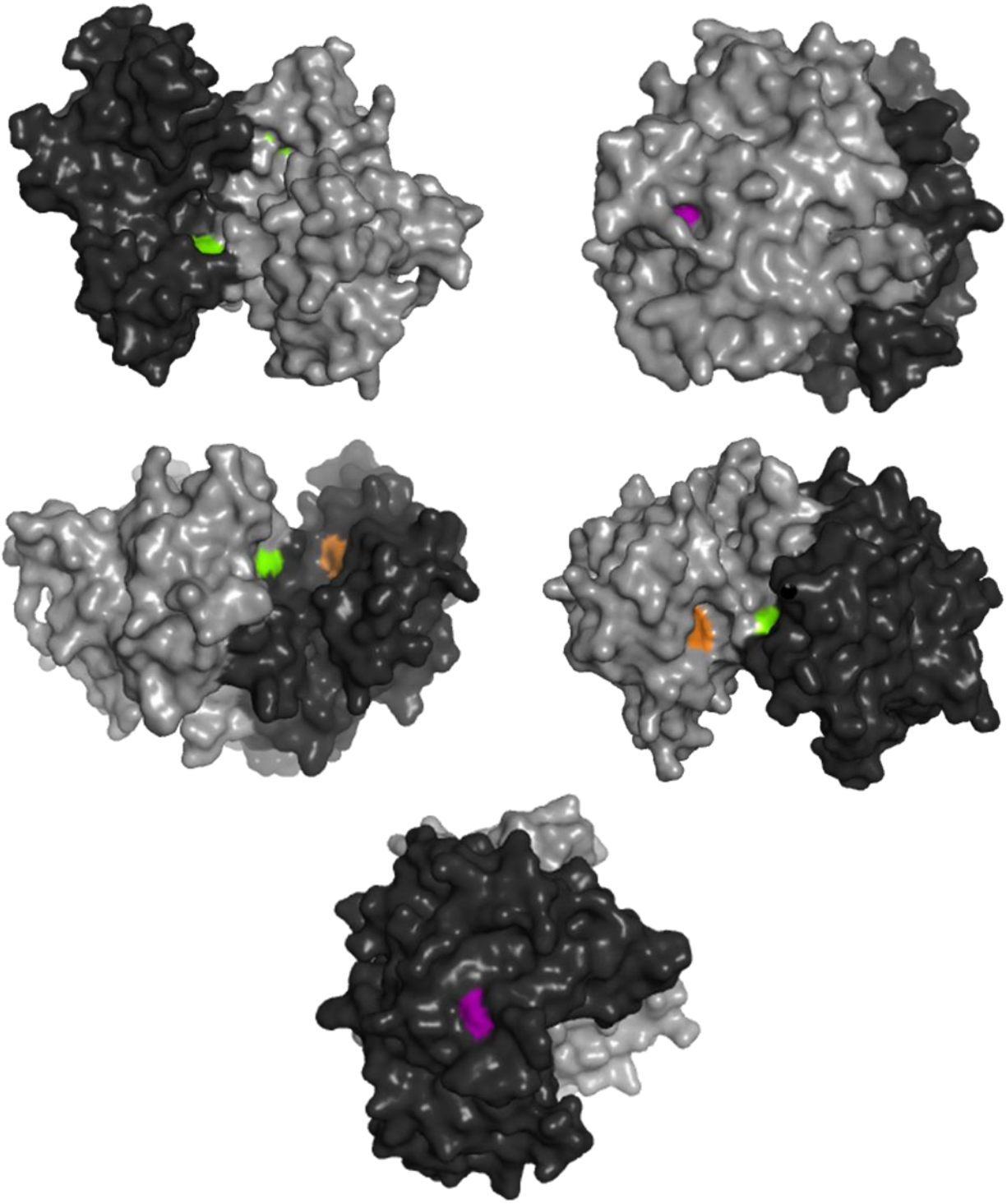


Figure 25: The location of tryptophan residues on the molecular surface of the EcSspA model. Molecular surface of EcSspA with chain A (dark grey) and chain B (light grey), showing Trp70 (in green), Trp110 (in orange) and Trp166 (in purple) residues from both chains. Trp70 is more visible on the surface of EcSspA, whereas Trp110 and Trp166 are less visible than Trp70 on the surface, indicating that they are less solvent exposed than Trp70. Images created with PyMol version 2.2.0.

Quaternary structure characterisation

The oligomeric state of EcSspA was characterised using SE-HPLC under non-reducing conditions. The chromatographic elution profile as well as the SE-HPLC calibration curve of EcSspA showed that the molecular mass of EcSspA was 46.5 kDa. Given that under reducing conditions, the molecular mass of EcSspA is 25.8 kDa (see section 4.1), this suggests that EcSspA under non-reducing conditions could be dimeric. Hansen *et al* (2005a) reported that EcSspA functions as a dimer and the theoretical molecular mass of EcSspA is given as 48.6 kDa by the ExPASy ProtParam Tool (Gasteiger *et al*., 2005), which suggests that the eluted EcSspA species is dimeric in nature. The discrepancy in the calculated and the theoretical molecular mass can be attributed to the globular shape of EcSspA, which allows EcSspA to take on a smaller tumble volume as it elutes, resulting in the elution of the protein sooner than expected (Tayyab *et al*., 1991; Bae *et al*., 2009); therefore making the molecular mass of the protein appear smaller than what it is.

It is important to note that the spectral data collected and shown in this section was unfortunately void of statistical analysis, which puts into question the precision of the data presented. Error bars could have been used to provide a clearer indication of the precision of the data, for experiments that were repeated independently (not just replicates of a single experiment) and thus improving the reliability of the data presented (Cumming *et al*., 2007).

4.3 Interaction of IAA-94 and EA with EcSspA

ANS has been widely used to characterise ligand binding sites on proteins, including cytosolic GSTs. Cytosolic GSTs, such as the mu class of GSTs, have been shown to bind ANS (Ketley *et al*., 1975; Kinsley *et al*., 2008). Similarly, through the use of extrinsic ANS fluorescence, soluble EcSspA was shown to bind ANS. The spectrum characteristically exhibited an increase in the fluorescence intensity and a blue shift of the maximum emission wavelength of the ANS-EcSspA complex, in comparison to the spectrum of free ANS (Qadeer *et al*., 2012). The binding of ANS to EcSspA indicated that the structure of EcSspA has hydrophobic sites on its surface, which could potentially serve as binding sites for other hydrophobic ligands.

With this knowledge, ANS displacement assays were performed to observe the effect that IAA-94 and EA would have on the fluorescence emission of the ANS-EcSspA complex. The presence of IAA-94 and EA resulted in the displacement of ANS molecules from EcSspA. Therefore, indicating that IAA-94 and EA interacted and bound to some of the hydrophobic regions of EcSspA, where ANS had previously bound. Furthermore, it was shown that 2.1 mM

of IAA-94 displaced more ANS molecules from EcSspA than the 2.48 mM of EA used; this suggested that EcSspA may have a greater affinity for IAA-94 than it does for EA.

It is important to note that IAA-94 and EA were solubilised in DMSO (see section 2.2.8.2) and as a result, 3% (v/v) DMSO was required in the samples used for ANS displacement as well as subsequent ITC titrations, to maintain the functional state of EcSspA. It has been shown that cytosolic GSTs retain functionality in DMSO concentrations of 1-3% (v/v) DMSO (Takahashi, 2017), which is also the suggested DMSO concentration range to ensure minimal signal artefacts when conducting ITC titrations in the presence of DMSO (Velazquez-Campoy and Freire, 2006).

Additionally, the ITC titrations that were performed in this study, show that EcSspA was interacting with IAA-94 and EA and that EcSspA may have a greater affinity for IAA-94 than EA. However, a closer look at the ITC isotherms revealed that the heats measured during the binding events were too low for the experiment. An appropriate peak height for the heats measured should at minimum be 0.5 µcal/s (Freyer and Lewis, 2008), whereas, this study yielded maximum peak heights of 0.3-0.35 µcal/s for the duration of the titrations. The peaks produced did not decrease during the course of the titrations, which indicated that both the reactions did not reach saturation. This influenced the shape of the isotherms and as a result, directly determining *N* accurately from the mid-point of the isotherm would not be possible (Wiseman *et al*., 1989; O'Brien *et al*., 2000). Therefore, the low heat signals produced for the binding of IAA-94 and EA to EcSspA begins to put into question the reliability of the thermodynamic information derived from the isotherms (Freyer and Lewis, 2008).

Due to the challenge of accurately deriving the *N* from the ITC isotherms, the IFMD of IAA-94 and EA to EcSspA was used to infer about the stoichiometry of the binding reactions. Both IAA-94 and EA were docked only at the dimer interface of the EcSspA model, which proposed that N=1 for the binding reactions. Taking into consideration the structural homology of IAA-94 and EA (Tulk *et al*., 2000; Harrop *et al*., 2001), this suggested that IAA-94 and EA could share a similar binding site on EcSspA and that the dimer interface of EcSspA is a potential binding site for ligands. Furthermore, docking at the dimer interface of EcSspA coincides with several GSTs that have shown to bind drugs at the dimer interface as well. For example, the *Schistosoma japonica* GST (SjGST) binds one molecule of praziquantel (an antischistosomal drug) at its dimer interface (McTigue *et al*., 1995) as well as a squid sigma class GST possessing non-catalytic ligand binding abilities at its dimer interface (Ji *et al*., 1996).

The interaction of IAA-94 and EA at the dimer interface revealed that hydrogen and halogen bonds are responsible for the binding occurring at the dimer interface. The interaction of IAA-94 and EcSspA shows that Lys108 and Tyr160 are involved in forming the halogen and hydrogen bonds at the dimer interface, respectively. The interaction of EA and EcSspA shows that Arg73 (from chain B) forms the halogen bond at the dimer interface and that Arg73 (from chain A), Tyr160 and Glu107 form hydrogen bonds at the dimer interface. These residues do not make up the structure of the EcSspA surface pocket, as the residues involved in forming the EcSspA surface pocket are: Arg82, Pro84, His85, Pro86, Leu88 and Tyr92 and none of these residues are shown to form intermolecular bonds with IAA-94 or EA. This suggests that although IAA-94 and EA bound at the dimer interface, they did not directly interact with the surface pocket. Therefore, indicating that the dimer interface could also serve as a target site for potential binding of ligands other than the proposed surface exposed pocket.

It is noteworthy to mention that the interaction between EcSspA and the ligands used in this study, would be better visualised on a 3-D model as opposed to a 2-D diagram. It would show the angles at which the interactions occurred and an illustration of the bond lengths would provide an indication of the strength of the bonds present in the interactions (Pauling, 1954). Additionally, a close-up view of the position of residues in the active site before and after docking of the ligands, would indicate how the side chains moved during the induced fit docking experiments. The images showing the molecular surface of the EcSspA model in section 3.5.4, would have benefitted from a white background as it would have provided better contrast of the EcSspA model depicted and the monomers of EcSspA could have been shown in separate colours to further highlight the dimer interface of the model. These changes would have enhanced the quality of the figures and the findings thereof.

Furthermore, although the GMQE and QMEAN scores of the EcSspA model show that the estimated accuracy and quality of the structure was relatively reliable based on the template used, the validity of the docking studies was not ensured and thus the reliability of the docking of IAA-94 and EA to the EcSspA model is questionable. In order to improve the reliability of the docking experiments, this study could have conducted the docking of IAA-94 and EA on the template structure as well as docking known ligands to the template and the EcSspA model to verify the reliability of the docking sites. This would then validate the docking of IAA-94 and EA at the dimer interface of the EcSspA model.

According to the thermodynamic parameters generated by SEDPHAT, the binding reaction of both IAA-94 and EA to EcSspA was exothermic and enthalpically driven. This indicates that the binding of these ligands to the dimer interface was energetically favourable (Du *et al*., 2016).

Although both IAA-94 and EA have shown to interact and favourably bind to EcSspA, the evaluation of the c-values indicated that IAA-94 and EA bind weakly to EcSspA. Using the K_d values provided by SEDPHAT, the c-values for the reaction of IAA-94 and EcSspA as well as EA and EcSspA were calculated and were both below 5, indicating that both reactions are weakly binding (Wiseman *et al*., 1989). As a result, determining the ΔH, K_d and N values directly from the binding isotherms proved to not be reliable, therefore compromising the validity of the ΔG and ΔS values obtained (Wiseman *et al*., 1989; O'Brien *et al*., 2000).

Rectifying the low c-value obtained for both protein-ligand interactions would improve the reliability of the thermodynamic parameters obtained. Although increasing the concentrations of EcSspA and the ligands may be an option, it can result in protein aggregation (Holdgate, 2001) and the appearance of signal artefacts due to the increased concentration of DMSO present in the binding reactions (Velazquez-Campoy and Freire, 2006). Other possible options would be to conducted the titration in reverse, therefore injecting EcSspA from the ITC syringe into the ligand solution in the sample cell (Holdgate, 2001) or to employ the displacement method, where EcSspA bound to the low affinity ligand is placed in the sample cell and a ligand of high affinity is titrated into the sample cell and displaces the low binding affinity ligand. The K_d of the strong binding ligand can then be used to determine the K_d of the weak binding ligand (Zhang and Zhang, 1998).

Although ITC is a sensitive technique, that can theoretically detect binding affinities from a nanomolar to a millimolar range (Freyer and Lewis, 2008), when the K_d or N are unknown at the beginning of the experiment (as in the case of this study), determining the accurate concentrations of the reactants can be a challenge and may require testing a range of concentrations (Freyer and Lewis, 2008). This can in turn produce poor heat signals and negatively affect subsequent derivations from the isotherm generated (Freyer and Lewis, 2008). If the reliability of the thermodynamic parameters cannot be improved by rectifying the c-values, it may indicate that ITC is not suitable to study the interaction/s at hand.

Overall, the findings of this study indicated that EcSspA interacted with both IAA-94 and EA and that the dimer interface of EcSspA could be a potential binding site for ligands, as long as the appropriate verification of experiments have been completed.

Conclusion:

This study intended to investigate whether IAA-94 as well as EA could interact with EcSspA and if so, where the binding would take place on the structure of EcSspA. The successful overexpression and purification of native, soluble EcSspA was followed by EcSspA concentration determination and structural characterisation. Structural characterisation confirmed that EcSspA is predominantly α-helical, with tryptophan residues that are less solvent exposed with the exception of Trp70, which is more solvent accessible than the other tryptophan residues within the EcSspA structure and that EcSspA is a dimeric protein that is approximately 46 kDa in size. The results of the structural characterisation were found to be in accordance with what has been reported in literature. The ANS displacement assays as well as the ITC data and IFMD studies revealed that IAA-94 and EA weakly bind to native, soluble EcSspA at the dimer interface but have no direct interaction with the EcSspA surface pocket. The ITC and IFMD data required necessary validation to provide reliability the above findings.

Chapter 5: References

Al Khamici, H., Brown, L.J., Hossain, K.R., Hudson, A.L., Sinclair-Burton, A.A., Ng, J.P.M., Daniel, E.L., Hare, J.E., Cornell, B.A., Curmi, P.M.G., Davey, M.W., Valenzuela, S.M. (2015). Members of the Chloride Intracellular Ion Channel Protein Family Demonstrate Glutaredoxin-Like Enzymatic Activity. PLoS ONE **10**, e115699.

Allocati, N., Masulli, M., Di Ilio, C., Federici, L. (2018). Glutathione transferases: substrates, inihibitors and pro-drugs in cancer and neurodegenerative diseases. Oncogenesis **7**.

Andrykovitch, M., Routzahn, K.M., Li, M., Gu, Y., Waugh, D.S., Ji, X. (2003). Characterization of four orthologs of stringent starvation protein A. Acta Crystallogr. D. **59**, 881–886.

Atkinson, H.J. and Babbitt, P.C. (2009). Glutathione Transferases Are Structural and Functional Outliers in the Thioredoxin Fold. Biochemistry. **48**, 11108–11116.

Bae, S.-H., Dyson, H.J., Wright, P.E. (2009). Prediction of the rotational tumbling time for proteins with disordered segments. J. Am. Chem. Soc. **131**, 6814–6821.

Benkert, P., Biasini, M., Schwede, T. (2011). Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics. **27**, 343–350.

Berryman, M. and Bretscher, A. (2000). Identification of a Novel Member of the Chloride Intracellular Channel Gene Family (CLIC5) That Associates with the Actin Cytoskeleton of Placental Microvilli. Mol. Biol. Cell. **11**, 1509–1521.

Beychok, S. (1966). Circular Dichroism of Biological Macromolecules. Science **154**, 1288–1299.

Brautigam, C.A. (2015). Calculations and publication-quality illustrations for analytical ultracentrifugation data. Methods Enzymol. **562**, 109-134

Capitani, G., De Biase, D., Aurizi, C., Gut, H., Bossa, F., Grütter, M.G. (2003). Crystal structure and functional analysis of Escherichia coli glutamate decarboxylase. EMBO J. **22**, 4027–4037.

Castanie-Cornet, M.-P., Penfound, T.A., Smith, D., Elliott, J.F., Foster, J.W. (1999). Control of Acid Resistance in Escherichia coli. J. Bacteriol. Res. **181**, 3525–3535.

Cavasotto, C.N. and Phatak, S.S. (2009). Homology modeling in drug discovery: current trends and applications. Drug Discov. Today. **14**, 676–683.

Collini, M., D'Alfonso, L., Baldini, G. (2000). New insight on beta-lactoglobulin binding sites by 1-anilinonaphthalene-8-sulfonate fluorescence decay. Protein Sci. **9**, 1968–1974.

Cumming, G., Fidler, F., Vaux, D. (2007). Error bars in experimental biology. J. Cell Biol. **177**, 7–11.

DiNitto, J.M. and Kenney, J.M. (2012). Noise characterization in circular dichroism spectroscopy. Appl. Spectrosc. **66**, 180–187.

Du, X., Li, Y., Xia, Y.L., Ai, S.M., Liang, J., Sang, P., Ji, X.L., Liu, S.Q. (2016). Insights into Protein–Ligand Interactions: Mechanisms, Models, and Methods. Int. J. Mol. Sci. **17**.

Dulhunty, A., Gage, P., Curtis, S., Chelvanayagam, G., Board, P. (2001). The Glutathione Transferase Structural Family Includes a Nuclear Chloride Channel and a Ryanodine Receptor Calcium Release Channel Modulator. J. Biol. Chem. **276**, 3319–3323.

Fitch, W.M. (2000). Homology: a personal view on some of the problems. Trends Genet. **16**, 227–231.

Freyer, M.W., Lewis, E.A. (2008). Isothermal titration calorimetry: experimental design, data analysis, and probing macromolecule/ligand binding and kinetic interactions. Methods Cell Biol. **84**, 79–113.

Friesner, R.A., Banks, J.L., Murphy, R.B., Halgren, T.A., Klicic, J.J., Mainz, D.T., Repasky, M.P., Knoll, E.H., Shelley, M., Perry, J.K., Shaw, D.E., Francis, P., Shenkin, P.S. (2004). Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. J. Med. Chem. **47**, 1739–1749.

Gao, H., Weitao, T., He, Q. (2011). Coping with the Environment: How Microbes Survive Environmental Challenges. Int. J. Microbiol. **2011**.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server, in: Walker, J.M. (Ed.), The Proteomics Protocols Handbook. Humana Press, Totowa, NJ. 571–607.

Gururaja Rao, S., Ponnalagu, D., Sukur, S., Singh, Harkewal, Sanghvi, S., Mei, Y., Jin, D.J., Singh, H. (2017). Identification and Characterization of a Bacterial Homolog of Chloride Intracellular Channel (CLIC) Protein. Sci. Rep. **7**.

Halgren, T.A. (2007). New Method for Fast and Accurate Binding-site Identification and Analysis. Chem. Biol. Drug Des. **69** (2): 146-148.

Hansen, A.M., Gu, Y., Li, M., Andrykovitch, M., Waugh, D.S., Jin, D.J., Ji, X. (2005a). Structural Basis for the Function of Stringent Starvation Protein A as a Transcription Factor. J. Biol. Chem. **280**, 17380–17391.

Hansen, A.M., Qiu, Y., Yeh, N., Blattner, F.R., Durfee, T., Jin, D.J. (2005b). SspA is required for acid resistance in stationary phase by downregulation of H-NS in Escherichia coli. Mol. Microbiol. **56**, 719–734.

Harrop, S.J., DeMaere, M.Z., Fairlie, W.D., Reztsova, T., Valenzuela, S.M., Mazzanti, M., Tonini, R., Qiu, M.R., Jankova, L., Warton, K., Bauskin, A.R., Wu, W.M., Pankhurst, S., Campbell, T.J., Breit, S.N., Curmi, P.M.G. (2001). Crystal Structure of a Soluble Form of the Intracellular Chloride Ion Channel CLIC1 (NCC27) at 1.4-Å Resolution. J. Biol. Chem. **276**, 44993–45000.

Hidalgo, E. (1998). The redox-regulated SoxR protein acts from a single DNA site as a repressor and an allosteric activator. EMBO J. **17**, 2629–2636.

Holdgate, G. (2001). Making cool drugs hot: isothermal titration calorimetry as a tool to study binding energetics. Biotechniques. **31**, 164–6, 168, 170.

Jelesarov, I., Bosshard, H.R. (1999). Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition. J. Mol. Recognit. **12**, 3–18.

Ji, X., Rosenvinge, E.C. von, Johnson, W.W., Armstrong, R.N., Gilliland, G.L. (1996). Location of a potential transport binding site in a sigma class glutathione transferase by x-ray crystallography. PNAS. **93**, 8208–8213.

Keller, S., Vargas, C., Zhao, H., Piszczek, G., Brautigam, C.A., Schuck, P. (2012). High-precision isothermal titration calorimetry with automated peak-shape analysis. Anal. Chem. **84**, 5066-5073

Kelly, S.M., Jess, T.J., Price, N.C. (2005). How to study proteins by circular dichroism. BBA-Proteins Proteom. **1751**, 119–139.

Kelly, S.M. and Price, N.C. (2009). Circular Dichroism: Studies of Proteins. Encyclopedia of Life Sciences. Chichester, UK. John Wiley & Sons, Ltd.

Kesik-Brodacka, M., Romanik, A., Mikiewicz-Sygula, D., Plucienniczak, G., Plucienniczak, A. (2012). A novel system for stable, high-level expression from the T7 promoter. Microb. Cell. Fact. **11**, 109.

Ketley, J.N., Habig, W.H., Jakoby, W.B. (1975). Binding of nonsubstrate ligands to the glutathione S-transferases. J. Biol. Chem. **250**, 8670–8673.

Kimple, M.E., Brill, A.L., Pasker, R.L. (2013). Overview of Affinity Tags for Protein Purification. Curr. Protoc. Protein. Sci. **73**, Unit-9.9.

Kinsley, N., Sayed, Y., Armstrong, R.N., Dirr, H.W. (2008). Characterization of the binding of 8-anilinonaphthalene sulphonate to rat class Mu GST M1-1. Biophys. Chem. **137**, 100–104.

Ladbury, J.E. and Chowdhry, B.Z. (1996). Sensing the heat: the application of isothermal titration calorimetry to thermodynamic studies of biomolecular interactions. Chem. Biol. **3**, 791–801.

Laemmli, U.K. (1970). Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. Nature **227**, 680.

Lakowicz, J.R. (2006). Principles of Fluorescence Spectroscopy, Third. ed. Springer Science & Business Media.

Matulis, D. and Lovrien, R. (1998). 1-Anilino-8-naphthalene sulfonate anion-protein binding depends primarily on ion pair formation. Biophys. J. **74**, 422–429.

McTigue, M.A., Williams, D.R., Tainer, J.A. (1995). Crystal Structures of a Schistosomal Drug and Vaccine Target: Glutathione S-Transferase fromSchistosoma japonicaand its Complex with the Leading Antischistomal Drug Praziquantel. J. Mol. Biol. **246**, 21–27.

Mekalanos, J.J. (1992). Environmental signals controlling expression of virulence determinants in bacteria. J. Bacteriol. **174**, 1–7.

Moller, M. and Denicola, A. (2002). Protein Tryptophan Accessibility Studied by Fluorescence Quenching. Biochem. Mol. Biol. Edu. **30**, 175–178.

Murphy, C., Carroll, C., Jordan, K.N. (2006). Environmental survival mechanisms of the foodborne pathogen Campylobacter jejuni. J. Appl. Microbiol. **100**, 623–632.

Nguyen, Y. and Sperandio, V. (2012). Enterohemorrhagic E. coli (EHEC) pathogenesis. Front. Cell. Infect. Mi. **2**, 1–7.

Nishida, M., Harada, S., Noguchi, S., Satow, Y., Inoue, H., Takahashi, K. (1998). Three-dimensional structure of Escherichia coli glutathione S-transferase complexed with glutathione sulfonate: catalytic roles of Cys10 and His10611. J. Mol. Biol. **281**, 135–147.

O'Brien, R., Ladbury, J.E., Chowdhry, B.Z. (2000). Isothermal titration calorimetry of biomolecules., in: Protein-Ligand Interactions: Hydrodynamics and Calorimetry. Chapter 10. Oxford, U.K. Oxford University Press.

Pace, C.N., Vajdos, F., Fee, L., Grimsley, G., Gray, T. (1995). How to measure and predict the molar absorption coefficient of a protein. Protein Sci. **4**, 2411–2423.

Pauling, L. (1954). The Dependence of Bond Energy on Bond Length. J. Phys. Chem. **58**, 662–666.

Qadeer, A., Rabbani, G., Zaidi, N., Ahmad, E., Khan, J.M., Khan, R.H. (2012). 1-Anilino-8-Naphthalene Sulfonate (ANS) Is Not a Desirable Probe for Determining the Molten Globule State of Chymopapain. PLOS ONE **7**, e50633.

Qian, Z., Okuhara, D., Abe, M.K., Rosner, M.R. (1999). Molecular Cloning and Characterization of a Mitogen-activated Protein Kinase-associated Intracellular Chloride Channel. J. Biol. Chem. **274**, 1621–1627.

Rampelotto, P.H. (2013). Extremophiles and Extreme Environments. Life (Basel). **3**, 482–485.

Rhodes, T. and Twentyman, P.R. (1992). A study of ethacrynic acid as a potential modifier of melphalan and cisplatin sensitivity in human lung cancer parental and drug-resistant cell lines. Br. J. Cancer. **65**, 684–690.

Scheuermann, T.H. and Brautigam, C.A. (2015). High-precision, automated integration of multiple isothermal titration calorimetric thermograms: New features of NITPIC. Methods. **76**, 87-98

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D., Higgins, D.G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. **7**, 539.

Smith, D.K., Kassam, T., Singh, B., Elliott, J.F. (1992). Escherichia coli has two homologous glutamate decarboxylase genes that map to distinct loci. J. Bacteriol. Res. **174**, 5820–5826.

Swinehart, D.F. (1962). The Beer-Lambert Law. J. Chem. Educ. **39**, 333.

Tabor, S. (1990). Expression Using the T7 RNA Polymerase/Promoter System. Curr. Protoc. Mol. Biol. **11**, 16.2.1-16.2.11.

Tayyab, S., Qamar, S., Islam, M. (1991). Size exclusion chromatography and size exclusion HPLC of proteins. Biochem. Educ. **19**, 149–152.

Tulk, B.M., Kapadia, S., Edwards, J.C. (2002). CLIC1 inserts from the aqueous phase into phospholipid membranes, where it functions as an anion channel. Am. J. Physiol. Cell Physiol. **282**, C1103–C1112.

Tulk, B.M., Schlesinger, P.H., Kapadia, S.A., Edwards, J.C. (2000). CLIC-1 Functions as a Chloride Channel when Expressed and Purified from Bacteria. J. Biol. Chem. **275** (35): 26989-26993

Valenzuela, S.M., Martin, D.K., Por, S.B., Robbins, J.M., Warton, K., Bootcov, M.R., Schofield, P.R., Campbell, T.J., Breit, S.N. (1997). Molecular Cloning and Expression of a Chloride Ion Channel of Cell Nuclei. J. Biol. Chem. **272**, 12575–12582.

Velazquez-Campoy, A., Freire, E. (2006). Isothermal titration calorimetry to determine association constants for high-affinity ligands. Nat. Protoc. **1**, 186–191.

Voet, D., Voet, J.G. (2011). Life. Biochemistry, Fourth. ed. Chapter 1, 3-6. New Jersey, U.S.A. John Wiley and Sons, Inc.

Wang, W., Wan, M., Liao, D., Peng, G., Xu, X., Yin, W., Guo, G., Jiang, F., Zhong, W., He, J. (2017). Identification of Potent Chloride Intracellular Channel Protein 1 Inhibitors from Traditional Chinese Medicine through Structure-Based Virtual Screening and Molecular Dynamics Analysis. Biomed Res. Int. **2017**.

Warburg, O. and Christian, W. (1942). Isolierung und Kristallisation des Garungs ferments Enolase. Biochem. Z. **310**, 384–421.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R., Schwede, T. (2018). SWISS-MODEL: Homology modelling of protein structures and complexes. Nucleic Acids Res. **46**, W296–W303.

Williams, M.D., Ouyang, T.X., Flickinger, M.C., (1994). Starvation-induced expression of SspA and SspB: the effects of a null mutation in sspA on Escherichia coli protein synthesis and survival during growth and prolonged starvation. Mol. Microbiol. **11**, 1029–1043.

Winayanuwattikun, P. and Ketterman, A.J. (2004). Catalytic and structural contributions for glutathione-binding residues in a Delta class glutathione S-transferase. Biochem J. **382**, 751–757.

Wiseman, T., Williston, S., Brandts, J.F., Lin, L. (1989). Rapid Measurement of Binding Constants and Heats of Binding Using a New Titration Calorimeter. Anal. Biochem. **179**, 131–137.

Wu, J., Filutowicz, M. (1999). Hexahistidine (His6)-tag dependent protein dimerization: a cautionary tale. Acta Biochim. Pol. **46**, 591–599.

Zablotowicz, R.M., Hoagland, R.E., Locke, M.A., Hickey, W.J. (1995). Glutathione-s-transferase activity and metabolism of glutathione conjugates by rhizosphere bacteria. Appl. Environ. Microbiol. **61**, 1054–1060.

Zalenskaya, K., Lee, J., Gujuluva, C.N., Shin, Y.K., Slutsky, M., Goldfarb, A. (1990). Recombinant RNA polymerase: inducible overexpression, purification and assembly of Escherichia coli rpo gene products. Gene. **89**, 7–12.

Zhang, Y.L. and Zhang, Z.Y. (1998). Low-affinity binding determined by titration calorimetry using a high-affinity coupling ligand: A thermodynamic study of ligand binding to protein tyrosine phosphatase 1B. Anal. Biochem. **261**, 139–148.