The role of surface amino acid residues in protein folding and stability

Yan Huang

Institute for Cell and Molecular Biosciences

Thesis submitted for the degree of Doctor of Philosophy

Newcastle University

May 2014

Declaration

I hereby declare that this thesis is my own work and effort and that it has not been submitted elsewhere for any award. Any contribution made to the project by others is explicitly acknowledged in this thesis.

Abstract

Colicin A is a plasmid encoded bacteriocin produced by *Escherichia coli* (*E. coli)* regarded as a potential antibiotic. Colicin A is unusual in that it requires acidic lipids for its activity and its pore-forming domain (ColA-P) forms an acidic molten globule state. Previously it was shown that surface aspartyl groups are essential for the stability of colicin A in solution. Learning the role of these surface amino acids in folding and stability of Colicin A pore-forming domain requires knowledge of the folded and unfolded states. To do this the aspartates were replaced by alanine, asparagine, glutamate and glutamine respectively. The folded and unfolded states of the protein were probed experimentally by near and far UV CD spectra and NMR ^{1}H-^{15}N HSQC-maps. The urea unfolding processes were studied and analysed by fluorescence emission spectra. This showed the unexpected result that this single domain protein unfolds in two distinct transitions, the first one being destabilized, and the second one being stabilized, by low pH. The thermal denaturation transitions of the protein were measured by differential scanning calorimetry (DSC) and CD spectroscopy at 209 nm and 295 nm. Unexpectedly the DSC data showed that the measured ΔH_{tr} was always half the $\Delta H_{vt'H}$ (van't Hoff) value and possible reasons for this are discussed. The data show that the aspartate to alanine replacements destabilize the protein significantly and the asparagine mutations do not. ColA-P D216, D274 and D372 are N-cap residues in the full length protein and this may explain why protonated Asp destabilizes the protein at low pH. The glutamate and glutamine mutants stabilize the protein more than expected for true N-capping residues. The biochemical function of the surface aspartates were studied comparing the NMR ^{1}H-^{15}N HSQC-maps of the proteins. The NMR experiments show that the aspartate to alanine mutation destabilize the protein in the same way as the acidic environment.

Contents

Acknowledgements

The first and most important person I would like to thank is my supervisor Prof. Jeremy Lakey. Without your incredible amounts of patience, guidance and training, I will never have this thesis.

Prof. Geoff Moore and Dr. Colin Macdonald, I would like to thank you two too. Both of you provide a lot of help on the knowledge of NMR. Without you two, I will not be able to finish the NMR chapter of my thesis.

Chris, thank you for all your help in the lab, and thank you to help me ordered all the stuffs from Blue Heron.

Helen, you are a truly helpful and organised person, and without your lab organisation, I would not be able to carry on my experiments.

Massive thanks to all the members of our lab, Ana, Nat, Daria and Hannah, I will never forget our happy dinners, you girls bring me a lot of happiness during these years.

Many thanks to the members from Geoff's team, without the massive help from Geoff's team, I will not be able to learn to use NMR spectroscopy and the complex analyse software.

Thanks to my Mum and Dad, when I feel lonely, you always spend a lot times to video chat with me and comfort me.

Abbreviations

Amp	Ampicillin
BSA	Bovine Serum Albumin
ColX-P	Colicin X (where X= A,B,N,E1 etc) pore forming domain
DNA	Deoxyribonucleic acid
E. coli	Escherichia coli
EDTA	Ethylenediaminetetraacetate
HSQC	Heteronuclear Single Quantum Coherence
IPTG	Isopropyl-Beta-d-thiogalactopyranoside
LB	Lysogeny broth
NMR	Nuclear magnetic resonance
OD	Optical Density
RNA	Ribonucleic acid
rRNA	Ribosomal ribonucleic acid
SDS	Sodium Dodecyl Sulphate

SDS-PAGE Sodium	Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
TEMED	N, N, N’, N’-tetramethylethylenediamine
tRNA	Transfer ribonucleic acid
WT	Wild type

Chapter One.

Introduction

1.1 Bacterial toxins

Bacterial toxins are ribosomally synthesized antimicrobial peptides produced by both Gram-positive and Gram-negative bacteria. At least 35% of protein toxins are membrane-damaging toxins (MDTs) (Alouf, 2001); they kill the target eukaryotic cells by damaging or disrupting the membrane, while others disturb cell function and cell signaling after entry into the cytosol (Offermanns and Rosenthal, 2008).

Generally, bacteria produce protein toxins for three reasons: the first one is to enter their target cells; the second one is to inhibit the immune system of the eukaryotic cells; the last one is to produce an appropriate host environment for their own development (Offermanns and Rosenthal, 2008). Bacteriocins have been divided into different groups in terms of molecular size, mode of actions, producer organism, immunity mechanisms etc., but a "universal" classification of bacteriocins are still not generally acceptable (Nes et al., 2007, Riley and Chavan, 2006). Currently, bacteriocins produced by Gram-positive bacteria have attracted most attention because most of them are safe for human (Nes et al., 2007), thus they are broadly used in food and feed fermented products (Nes et al., 2007). The colicins are the most extensively studied Gram-negative bacteria produced bacteriocins because they can be used as valuable tools in cell research; for example, they were used to examine the translocation process of protein through outer membrane and inner membrane in both directions, and the channel forming activity was used as a model system to study the protein insertion and pore formation process in membrane (Braun et al., 1994, Riley and Wertz, 2002). Halocin family is the only characterized Archaeocin produced by a

Archaea - halobacteria. Halocin is extremely stable which may make it the weapon of choice in the microbial world. (Riley and Wertz, 2002)

1.2 Colicins

Colicins are bacterial toxins produced by immune *Escherichia coli* (Gokce and Lakey, 2003). These toxic proteins are lethal for competitive strains of *E. coli* or closely related bacteria but not active against the producing bacteria (Nomura, 1967, Pugsley, 1984, Pugsley, 1985, Wolstenholme and O'Connor, 2008). Colicin gene clusters are encoded on specialized plasmids and consist of two closely linked genes: a colicin gene, which encodes the toxin and an immunity gene, which encodes a specific immunity protein to protect the producer cell by inactivating the colicins. A lysis gene may also be located on the colicin plasmid, which encodes a lysis protein to induce cell lysis and release the colicin into the medium (Gordon and O'Brien, 2006, Riley and Wertz, 2002). The production of colicins is induced under the SOS condition; the agents that trigger the SOS response range from physical agents such as UV light to chemical drugs such as DNA damaging mitomycin C (Cascales et al., 2007). The immunity is provided by a specific inhibitor protein which is able to prevent colicins from killing the producing bacteria and whose function depends upon the type of killing mechanism.

Colicins have been studied for nearly a century. The first colicin was identified by Gratia in 1925 as a heat-labile product present in cultures of *E. coli* V, which is toxic for *E. coli* φ (Gratia, 1925). The first entire colicin to be crystallized was colicin Ia (Wiener et al., 1997). The earlier structure of the C-terminal domain of colicin A (Parker et al., 1994) demonstrated the 10-helix bundle which is similar to other pore forming colicins.(see Figure.1.1)

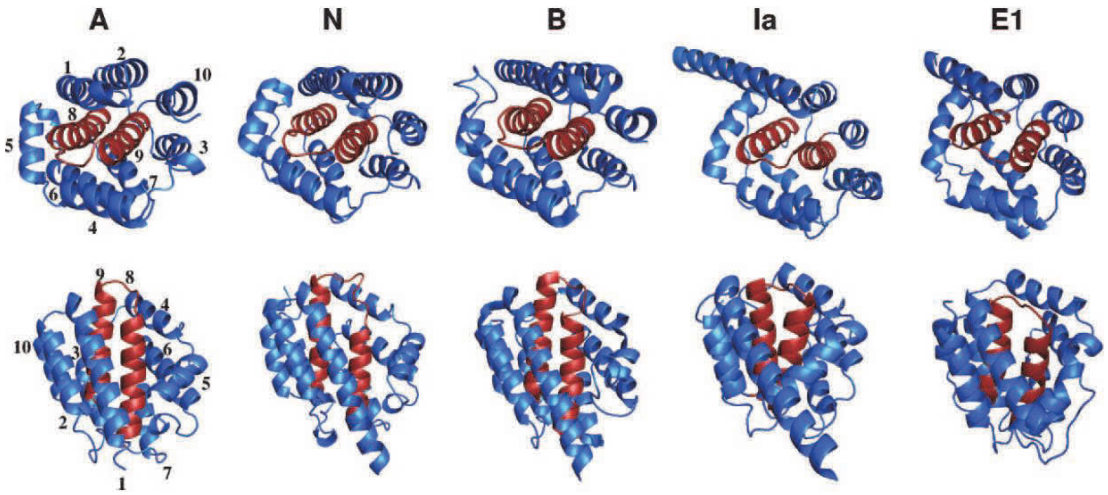


Figure. 1.1 Comparison of the structure of the pore forming domains of colicins A, N, B, Ia, and E1. Each molecule is shown in two orientations: one approximately parallel to the hydrophobic hairpin (upper) and the other perpendicular to it (lower). The central two hydrophobic helices (helix VIII and helix IX) are highlighted in red, whereas the other eight helices are shown in blue. Helix numbers are shown explicitly for colicin A. and the overall similarity in the folds of these molecules is apparent (Wiener et al., 1997).

The earliest studies on the mode of action of various colicins demonstrated that colicins kill the sensitive strains of *E. coli* by different effects. For example: type K colicins inhibit all macromolecular synthesis and many energy-dependent processes; colicin E2 breaks down DNA; colicin E3 stops protein synthesis (Nomura, 1963, Wendt, 1970, Konisky and Nomura, 1967). Later on, more killing actions of colicins have been revealed, such as: colicin E3 which hydrolyses rRNA; Colicins D and E5 digest tRNA (Cascales et al., 2007); Colicin A forms ion channels in the cytoplasmic membrane and allows potassium ion efflux (Parker et al., 1989, Parker et al., 1992, Parker et al., 1990); Colicin M was initially described as inhibiting peptidoglycan synthesis, but was later demonstrated to degrade lipid peptidoglycan intermediates (El Ghachi et al., 2006). Thus the modes of toxic activity can be summarised as follows: 1)

inhibition of macromolecular synthesis; 2) formation of voltage-gated channels on the membrane of the target cells; 3) digestion of nucleic acids.

Colicins have been purified and discovered to be proteins of high molecular mass ranging from 40 kDa to 80kDa. All the colicins are organized into three domains that correspond to one step of colicin action respectively: the central domain (receptor domain) is responsible for the receptor-binding activity; the N-terminal domain (translocation domain) is involved in translocation; the C-terminal domain possesses the killing function (Cavard and Lazdunski, 1981, Hardy et al., 1973, Vetter et al., 1998).

Colicin evolution

The investigation of colicin evolution has become a useful tool to understand the changes in the colicin-encoding genes, the toxic protein sequence and their killing modes (Riley and Chavan, 2006). With support of a large number of studies on colicins, it has become a model to study molecular evolution. Even though the colicin-encoded plasmids are various (Pugsley, 1987), the colicin molecules can be divided into many types and the killing modes of colicins are different, they share four features in common: (1) the colicin gene is encoded on plasmids (2) the colicin-operon is structured with a colicin activity gene, an immunity gene and a lysis gene (in group A colicins; see later) (3) toxic colicins are released from their host bacteria; and (4) the immunity protein protects the producers from external colicin (Hardy, 1975, Luria and Suit, 1987, Pugsley, 1984).

There are two main evolutionary lineages in colicin evolution: pore-formation and nuclease activity (Cascales et al., 2007). The former colicins kill cells by forming channels in the cytoplasmic membrane of the target cells, and kill the cells by disrupting the osmotic pressure balance. The latter ones kill the target cells by cleaving or inhibiting the synthesis of DNA or RNA. Most of the pore-forming

colicins share one or more regions with high levels of protein sequence similarity. On the other hand, the nuclease colicins share a recent common ancestry, their DNA sequences are very similar, ranging from 50% to 97% sequence identity, most divergence of nuclease colicins occurs in immunity region. (Cascales et al., 2007). (see Figure 1.2)

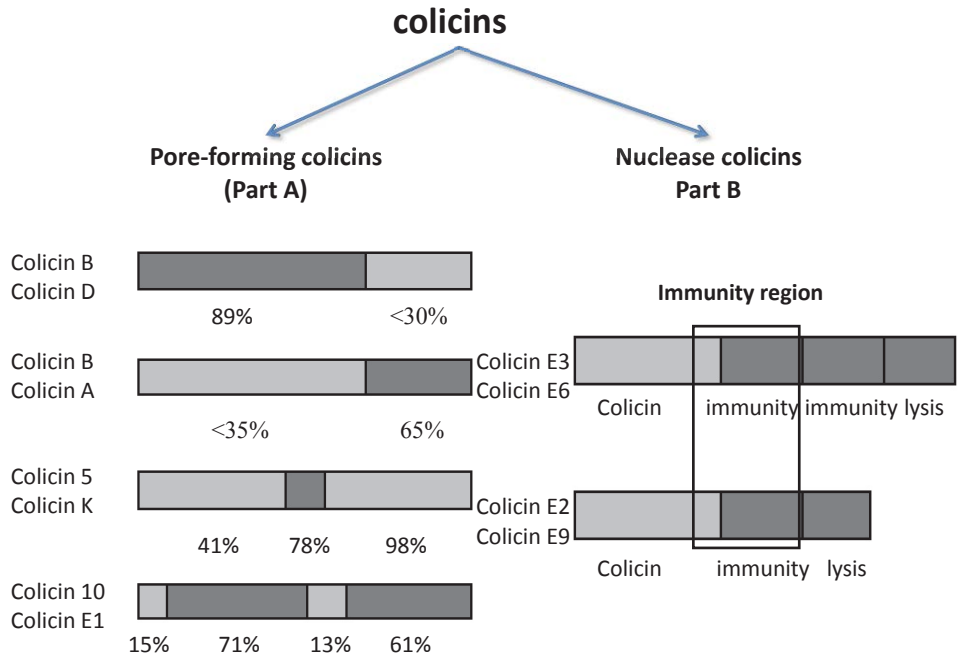


Figure. 1.2 Evolutionary lineages in colicin evolution. Part A indicating the pair wise comparison of pore-forming colicin protein sequences (Colicin B/D; Colicin B/A; Colicin 5/K and Colicin 10/E1). Values below each comparison show percent sequence identity for the highly homologous region (Dark grey) and divergent region (Light grey). Part B indicating the pair wise comparison of nuclease colicin Gene clusters (Colicin E3/E6 and Colicin E2/E9). Most divergence between nuclease colicins appear in the immunity region.

According to the result of protein sequence comparisons, five distinct groups of colicins exist: group 1 consists of colicin A, B, E1, Ia, Ib and N; group 2 consists of colicin B and D; group 3 consists of colicin DF13, E2, E3, E5, E6, E8 and E9; group

4 of consists of only colicin M and group 5 consists of only colicin V. Each member of a group either shared a recent common ancestor; or shared a remote common ancestor with either portions of sequence having been constrained or having retained a highly homologous amino acid sequence in one particular domain; this is particularly true of pore forming colicins (Riley, 1993).

The evolutionary relationships of 16 colicins are categorized into five groups and listed in table 1.1 (Riley, 1993)

Group NO.	Consists of colicins	Similarities in evolution	Divergence/difference in evolution
Group one	Colicin A/B/E1/Ia/Ib	1. They share a similar killing function. 2. They share a similar C-terminal domain.	1. They import into the sensitive cells in different way. 2. They have different immunity protein. 3. They have different N-terminal domain.
Group Two	Colicin B/D	1. They share the same receptor specificity 2. The N-terminal 400 amino acids are highly homologous.	1. Their killing functions are different. 2. They have different immunity protein.
Group Three	Colicin DF13/E2/E3/E5/E6/E8	1. They import into the sensitive cells in a similar mode. 2. The N-terminal 400 amino acids are highly homologous.	1. The C-terminal 100 amino acids are divergent.
Group four	Colicin M	1. Colicin M kills sensitive cells by inhibiting murein biosynthesis.	1. It shares no amino sequence similarity with any other colicins.
Group Five	Colicin V	1. Colicin Forms pores in the cytoplasmic membrane.	1. It shares no amino sequence similarity with any Group one colicins.

Colicin Immunity

Colicins are lethal to their related strains of *E. coli*; however none of them are active against their host bacteria. It is because a specific immunity protein which is also synthesized and released by the same *E. coli* strains to protect the cells from the killing activity of the colicin (Cascales et al., 2007). The immunity proteins block the killing activity via two different modes: (1) the immunity protein of pore forming colicins is anchored in the inner membrane, preventing the toxic protein opening its pore; (2) the immunity protein of a nuclease colicin is synthesized and located in the cytoplasm of the producing cells; here it can form a very high affinity complex with the incoming toxic protein and inhibit catalytic activity in the producing bacteria (Weaver et al., 1981). Thus nuclease toxins are secreted with a bound immunity protein which is released upon entry into the target cell.

Based on the length of the hydrophobic helix of the pore-forming domain of colicins, the corresponding immunity proteins were divided into two types: the type A immunity proteins inhibit the toxicity function of colicins A, B, N and U; the type E1 immunity proteins inhibit the lethality function of colicins E1, 5 K, 10, Ia and Ib. These two types of immunity proteins directly inhibited the specific functional channel in the inner membrane (Geli and Lazdunski, 1992, Zhang and Cramer, 1993). Type A immunity proteins can inactivate colicins before the formation of the pore channel; while it was believed that type E1 immunity proteins prevent the colicin activity in an open state or just before pore channels are open (Cascales et al., 2007).

Nuclease-specific immunity proteins are more varied in their structures. However, according to their inhibitory mechanism, two groups of immunity proteins can be distinguished: (1) one group inhibits the nuclease activity by targeting the enzyme active site (tRNase-specific Im proteins); (2) while another group inhibit the nuclease activity by blocking the access to the target enzymes (DNase- and rRNase-specific Im proteins (Capaldi et al., 2002, Gsponer et al., 2006). Colicin M is released bound to its

immunity protein and its cytotoxic activity is directly blocked by inhibiting the enzyme synthesis (Braun et al., 2002, Harkness and Braun, 1989)

Colicin release and expression

Colicins were first observed by Gratia in the culture medium of producing cells (Gratia, 1925). The mechanism of colicin secretion involves only one protein, the colicin lysis protein. After the synthesis of both colicins and the lysis protein, it allows colicins to be released from the producing cells (Cascales et al., 2007, Hakkaart et al., 1981, Jakes et al., 1974, Pugsley and Schwartz, 1983, Sabik et al., 1983, van der Wal et al., 1995). It not only differs from the protein releasing pathways in Gram-negative bacteria, it is also differs from the action of phage lysis proteins (Young, 1992).

Only group A colicins are released into the culture media, the group B colicins are not secreted. Group A colicin plasmids encode a lysis protein, which is functionally related to the release of the colicin, while the gene is missing from group B colicin plasmids. Only small amounts of colicins are generally present in the culture of colicinogenic cells in normal conditions. After treating the cells with SOS agents, UV light or mitomycin C, the amount of group A colicins released starts to increase rapidly. It reaches a maximum level after 120 to 150 min of induction and is about 1,000 times higher than the amount before induction. For the group A colicins, the synthesis is a toxic event for cells, so the amount of viable cells immediately starts to reduce after induction; for the group B colicin, their synthesis is not lethal for the producing cells (Cavard, 1997, Hardy et al., 1973, Jacob et al., 1952, Vetter et al., 1998).

1.3 Colicin classification

Colicins are encoded by colicinogenic plasmids-pCol, and two types of colicin plasmids were classified: type I, a group of small plasmids present in about 20 copies

by cell, which can be amplified and are mobilisable in the presence of conjugative plasmids and type II, large plasmids of about 40 kDa, which can be conjugative and promote the horizontal transfer of genetic material between donor and recipient cells by physical contact. The colicins were thus sorted into two groups, the group A colicins are mainly encoded by type I pCol while the group B colicins are encoded by type II pCol (Bouveret et al., 1997).

More recently, it has been found that colicins enter the target cells by using two different systems: Tol and Ton systems (see Figure 1.2) (Cascales et al., 2007). It allowed a new classification of colicins: group A colicins, such as colicins A, E1 to 9, k, L, N, S4 and Y, require the periplasmic Tol proteins to enter the cells,; while group B, such as colicins B, D, H, Ia, Ib, M, colicins need TonB, ExbB and ExbD as a receptor on the outer membrane of the cells, (Davies and Reeves, 1975a, Davies and Reeves, 1975b). As stated above, group A colicins are encoded by small plasmids and groups B colicins are encoded by large plasmids. Some colicins, such as Colicin 5 and 10, belong to group A colicins, but share similar protein sequences with group B colicins (Pilsl and Braun, 1995b, Pilsl and Braun, 1995a).

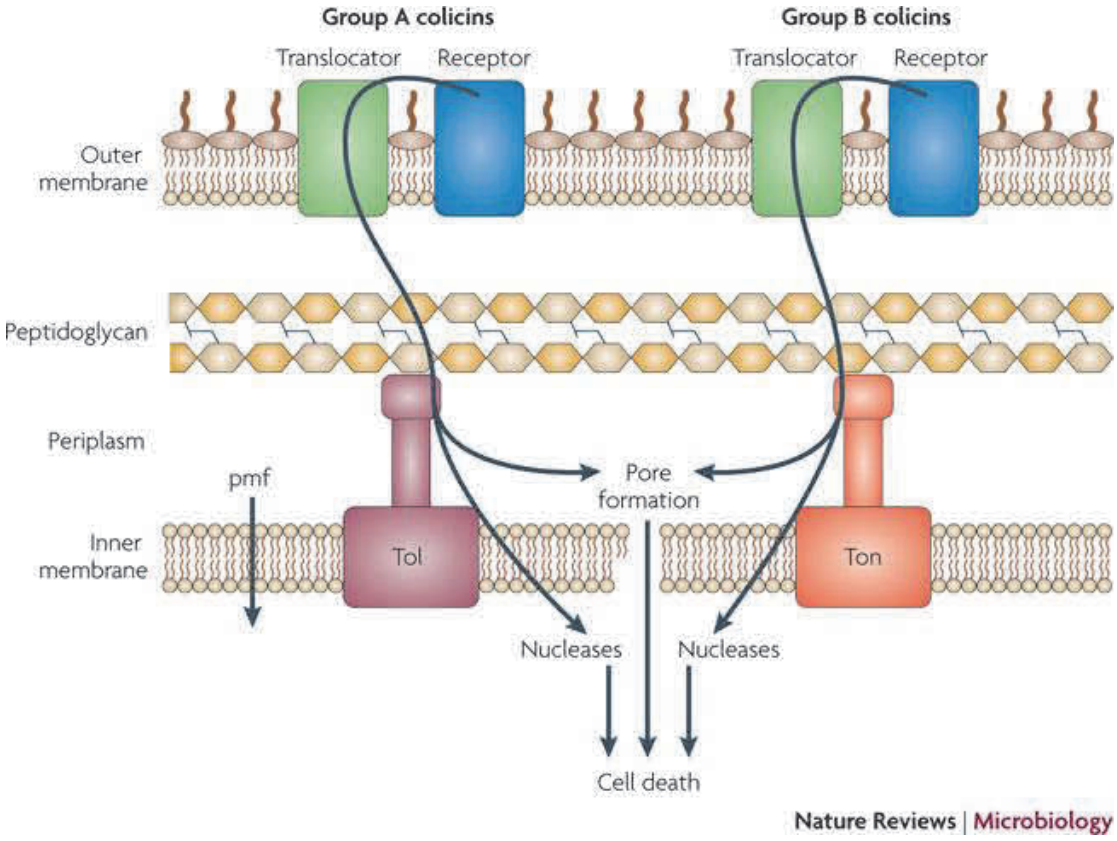


Figure. 1.2 Model of Tol and Ton systems. This figure shows that during colicin import, colicins firstly bind to their outer membrane receptor (blue) and then translocator proteins (green) help them to contact periplasmic and inner membrane proteins, finally Group A colicins translocate through the periplasm by using the Tol system, whereas group B colicins use the Ton system. (Kleanthous, 2010)

Group A colicins

Colicins target bacterial cells by binding to specific outer membrane proteins. Most of the group A colicins needs two receptors. The primary receptors, such as BtuB; the target of colicins A and E1 to E9, OmpA; the target of colicin U and Tsx; the target of colicin K, are cell surface proteins but usually not involved in the translocation process; the second receptor (usually OmpF or TolC) is required for the colicin to translocate across the outer membrane. Colicin N is the only group A colicin, which only uses a single receptor-OmpF for targeting and translocation.

Group A colicins were categorized by its receptors and second receptors, and listed in table 1.2

Group A colicins	Primary receptor	Secondary receptor
Colicin A	BtuB	OmpF, TolA/B/Q/R
Colicin E1	BtuB	TolC, TolA/Q
Colicin E2/E3/E4/E5/E6/E7/E8/E9	BtuB	OmpF, TolA/B/Q/R
Colicin N	OmpF	OmpF, TolA/Q/R
Colicin K	Tsx	OmpF, OmpA, TolA/B/Q/R
Colicin U	OmpA	OmpF, TolA/B/Q/R

Most of the group A colicin translocation processes follow the same steps. After recognizing and binding to the surface receptor, OmpF is employed to form a Receptor-OmpF-colicin complex with an unstructured segment of the colicins T-domain (Housden et al., 2005). Then the T-domain translocates into the cell through an OmpF pore to recruit a Tol protein (Loftus et al., 2006). Finally, it has been proposed that the colicin's killing domain (pore-forming domain or nuclease domain) which should be very close to an OmpF trimer now, translocates into the periplasm using another OmpF (Zakharov et al., 2006, Clifton et al., 2012a).

Colicin N cannot kill the target cells without outer membrane protein OmpF. After the receptor binding process, like all colicins, the entire colicin N unfolds during the translocation process. It is usually proposed that the colicin will pass through the membrane using the OmpF channel. However recent studies from our group show that the pore-forming domain could enter the cells at the OmpF-lipid interface (Clifton et al., 2012b).

Group A bacteria need the Tol system as periplasmic receptors. In most Gram-negative bacteria, this is a multi-protein complex composed of five proteins of the periplasm and inner membrane. Although, the Tol system has been studied by many different biological methods, the physiological role is still unclear (Lazzaroni et al., 2002, Lloubès R et al., 2001). Recent work in our group and elsewhere has demonstrated that although colicins A and N are group A colicins, and both of them exploit TolAIII to enter the target cells, they bind TolAIII at different sites.

Group B colicins

Most group B colicins require only one receptor to bind to and translocate, such as FepA-the only receptor for colicin B and D, FhuA- the only receptor for colicin M, Cir-the only receptor for colicin Ia and Ib. Colicin 5 and 10 are the only group B colicins which use two outer membrane proteins to bind and transport (Cascales et al., 2007).

Currently, only very limited knowledge about the translocation process of group B colicins has been revealed. Two steps are probably involved in the group B colicin translocation. First of all, colicin Ia binds to Cir and starts an interaction between Cir and Ton B, then a transient pore in Cir could be opened by pulling the plug domain out of the barrel. After the translocation domain enters the cell through the Cir pore and accesses the periplasm, an interaction between the colicin and the TonB from the inner membrane must be involved in the group B colicin translocation process (Cascales et al., 2007).

1.4 Colicin A

Colicin A (col A) is a plasmid-encoded pore-forming colicin which belongs to group A (Tol dependent). It kills *E. coli* by forming a voltage-gated channel in the cytoplasmic membrane of the competitive cells. Colicin A is imported from the

outside of the cell to the inner membrane through three steps: the first step, recognizing and binding onto the specific receptor protein at the external surface of sensitive cells (OmpF and BtuB are both required in this step); the second step, translocates to their specific target, the cell envelope with the aid of Tol A Tol B, Tol R, and Tol Q; the third step, the pore forming domain inserts into the inner membrane and forms a ion channel (Duche, 2002, Lazdunski et al., 1998, Benedetti et al., 1991).

Structure of colicin A

Colicin A is a 592 residue water soluble protein toxin organized into three functional domains (See Figure. 1.3):

(1) The central domain (receptor binding domain) which recognizes and binds the vitamin B12 receptor, BtuB at the cell surface.

(2) The N-terminal domain (translocation domain) which crosses the cell envelope and transports the pore-forming domain in to the periplasm.

(3) The C-terminal domain (pore-forming domain) consists of 10 alpha-helices. Its two hydrophobic helices (helix VIII and IX) are tightly surrounded by the remaining 8 amphipathic helices (helix I to helix VII and helix X). It has been shown that the hydrophobic core and the surrounding helices cross the membrane to form the voltage-dependent channel (Kienker et al., 1997).

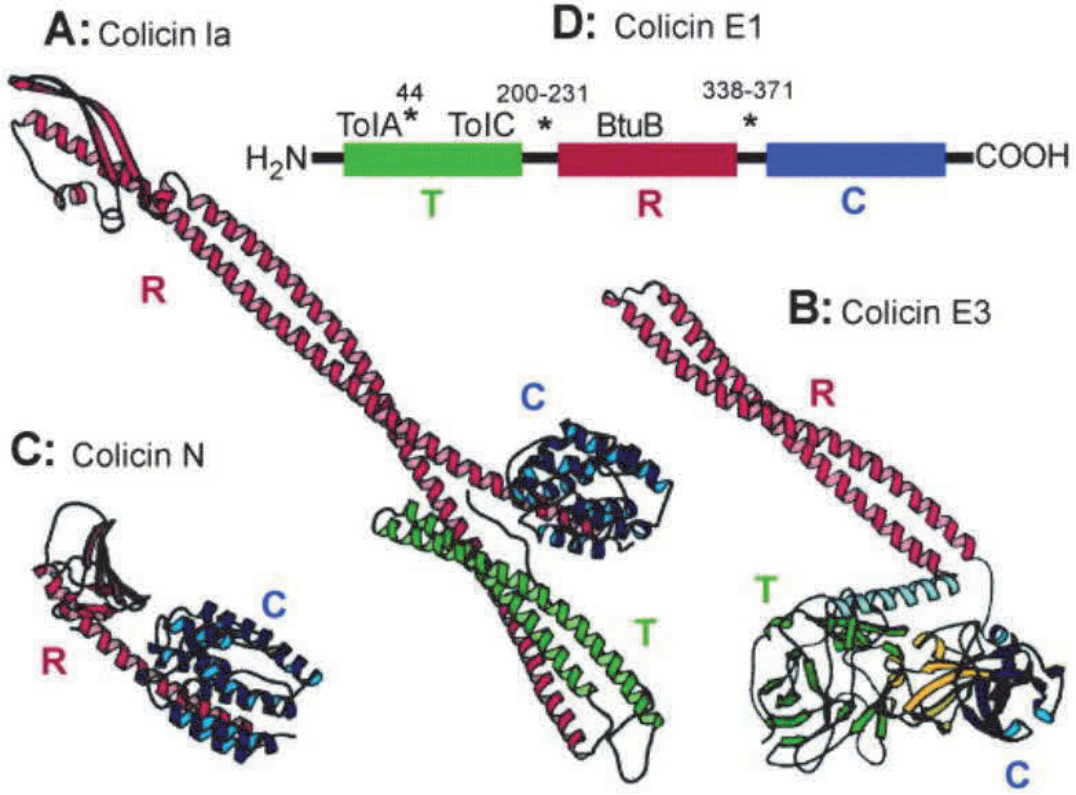


Figure. 1.3 Structures of Colicins Ia (A), E3 (B), N (C) and E1 (D). The regions correspond to Central domain (Red), translocation domain (Green) and C-terminal domain (Blue) (Zakharov and Cramer, 2002). Colicin A structure is unknown but falls between Ia and E3 in size.

Structures of the pore forming domains

Colicin A pore-forming domain is a 203 residue water soluble protein (Parker et al., 1989) which can form a molten globule state in an acidic environment (González-Mañas et al., 1992, van der Goot et al., 1993). The Colicin A pore-forming domain was the first colicin structure to be determined (Lakey and Slatin, 2001, Parker et al., 1989). Later on, the crystal structures of another four pore-forming proteins (colicins E1, B, Ia and N) have also been revealed (Elkins et al., 1994, Ghosh et al., 1994, Hilsenbeck et al., 2004, Vetter et al., 1998)

It has been proved that the molten globule is special structure of some colicins, such as colicin A and B, which occurs during their insertion and leads to the formation of a voltage-gated channel. However some pore-forming proteins, such as colicin E1 and N, do not form this state (Cascales et al., 2007). Thus, the *in vivo* acidic molten globule state of colicin A is a great biological interest for us especially as it does not occur in the closely related colicin N (Vetter et al., 1996). Thus, it is important to compare the two groups but this is made difficult by the fact that the pore-forming domains are at the C-terminus and thus their amino acid residues are numbered very differently. The different numbering systems of the amino acid residues (colicin A numbering system and colicin N numbering system) are shown in figure 1.4. which also shows the primary structures and secondary of the pore-forming-domain of colicin N and colicin E1.

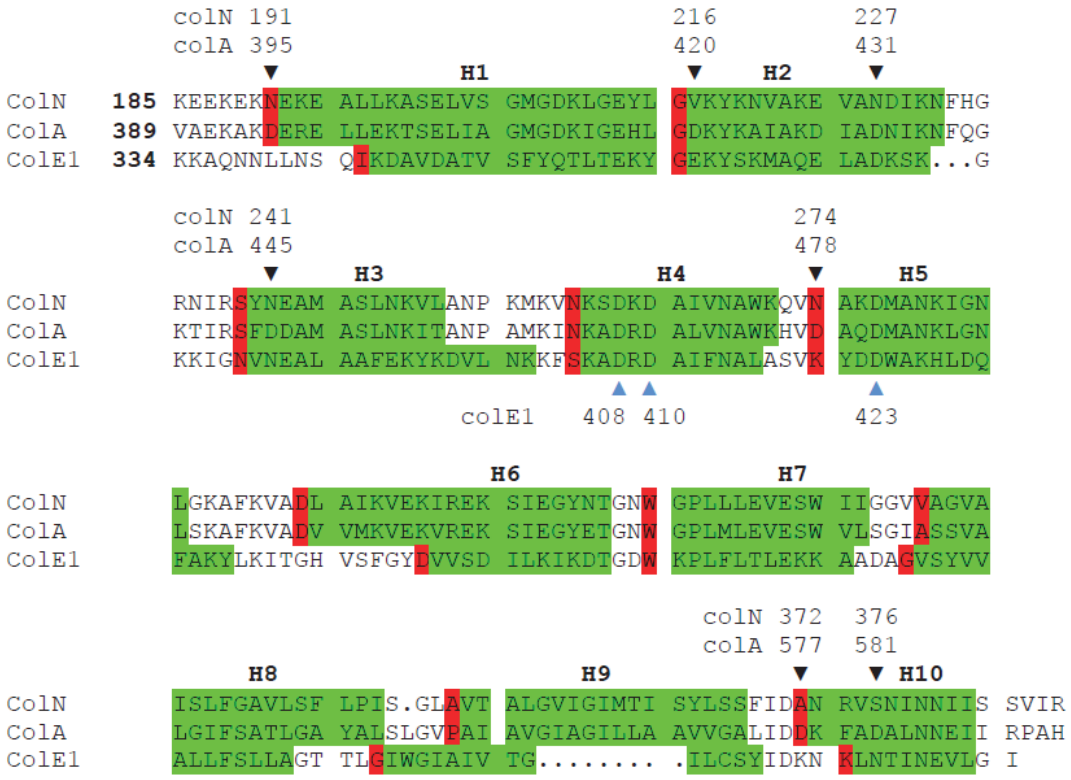


Figure 1.4 the amino acid residues of colicins N, A and E1 with their corresponding residue numbering system. The green boxes represent the ten alpha helices structures of the protein; the red boxes indicate the N capping point of each helix (Doig and Baldwin, 1995).

It is clear that the pore-forming domains of colicins A and N share more than 50% sequence identity, but they exhibit significantly different behaviour in the acidic environment. As in the observed channel forming activity, in solution the colicin A pore-forming (ColA-P) domain forms a stable intermediate state (molten globule state) below pH 3.0, but colicin N pore-forming domain remains its native structure below pH 1.0 (Fridd and Lakey, 2002)

Comparison of the Pore-forming domain of colicin A and other Pore-forming colicins.

According to the known crystal structures of colicin pore-forming domains, all of the pore-forming domain form pores in a similar way. The structure of the pore-forming domain of colicin A (Parker et al., 1989), Ia (Wiener et al., 1997), E1 (Elkins et al., 1994), N (Vetter et al., 1998), and B (Hilsenbeck et al., 2004) have now been solved, all of which are organized as 10 helices. Helices 8 and 9 are hydrophobic segments, which should insert into the membrane and interact with the hydrophobic core of the bilayer when forming the voltage-dependent channels; while the other helices are mostly amphipathic, which splay out onto the surface of the cytoplasmic membrane (Cascales et al., 2007).

Only colicins A and B form the acidic molten globule state below pH 3.0 (Braun and Maas, 1984, Evans et al., 1996, Ortega et al., 2001); the pore-forming domain of some colicins, such as colicin E1 shows limited pH sensitivity (Schendel and Cramer, 1994); even more, some pore-forming domains do not show pH sensitivity at all, such as colicin N (Evans et al., 1996). Although Colicins A and N are highly homologous, a few extra aspartate residues on the surface of ColA-P domain have been found to cause the significant differences between the pI and the pH stability of the pore-forming domains of colicins A and N (Fridd and Lakey, 2002).

1.5 Colicin A activity

Membrane channels: voltage-dependent channels

According to the chemical or physical modulator controlling gating activity, the ion channels that are formed by membrane-inserted proteins can be classified as: ligand-gated channels, voltage-gated channels, second messenger-gated channels, or mechanically-gated channels (Aidley and Stanfield, 1996)

By the end of the 1970s, two types of voltage-gated channels: Na/K channels and Ca channels were believed to be included in the voltage-gate channels (Aidley and Stanfield, 1996). Potassium efflux from the cytoplasm has been found in colicin A activity, which confirmed that colicin A acts by forming voltage-gated channels in the cytoplasmic membrane of *E. coli* (Bourdineaud et al., 1990). Although the colicin A toxin forms channels in pure lipid bilayers (Pattus et al., 1983a) it may also form a pore by combining with the K^+ channel proteins in *E. coli* and inhibiting the closure of the channel.

The mechanism of the open voltage-dependent channel is difficult to reveal due to its complex activity kinetics. The colicin pore forming system is not a simple two-state system and, besides, the closed and open states of colicin channels are not well defined (Lakey and Slatin, 2001).

The umbrella model and the penknife model

At present, two models, the umbrella model and the penknife model are the most popular hypothetical conformations of the closed channel state of colicin A. As described above, the colicin A C terminal domain consists of a hydrophobic hairpin that comprises two hydrophobic helices (H8 and H9) and eight amphipathic helices (H1-7 and H10). In the umbrella model, the hydrophobic hairpin is inserted into the bilayer; in the penknife model, the hydrophobic and the amphipathic helices are arranged nearly parallel to the surface of the membrane (Fig 1.2)

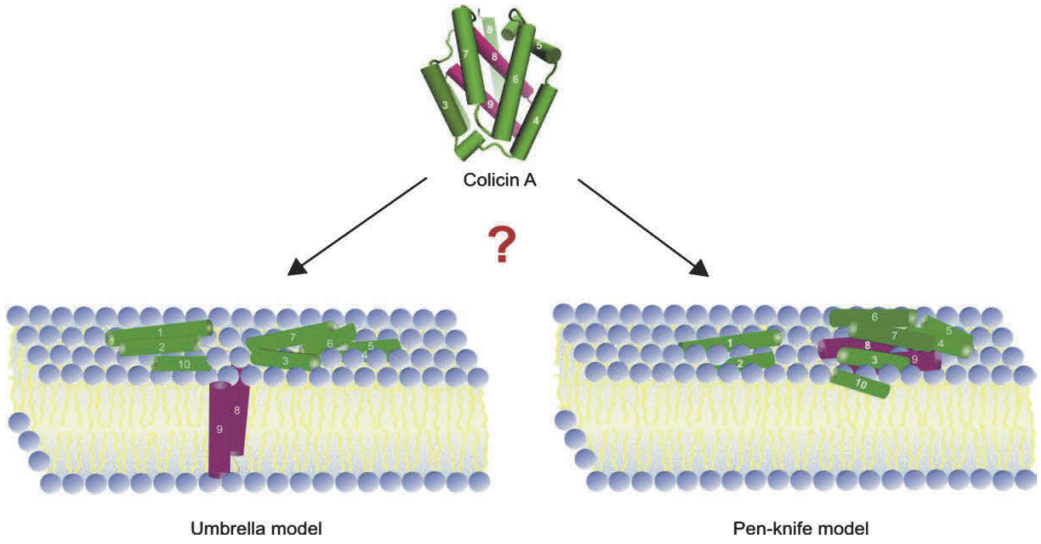


Fig 2. Umbrella and penknife models for the closed channel state of ColA. The ColA pore-forming domain is rendered in cartoon style. The amphipathic helices (H1–H7 and H10) are shown in green, and the hydrophobic hairpin (H8 and H9) is shown in magenta. Figure taken from (Padmavathi and Steinhoff, 2008).

For colicin A, the umbrella model is supported by fluorescence spectroscopy (Lakey et al., 1991b). In contrast the penknife model is supported by fluorescence resonance energy transfer experiments (Lakey et al., 1991a, Lakey et al., 1993), proteolysis, neutron scattering, disulfide bond engineering and electron paramagnetic resonance spectroscopy (Padmavathi and Steinhoff, 2008). Hence, understanding the conformation of the closed state of colicin A is still an essential point for studying the function of colicin A pore-forming domain.

The open channel of colicin A

Reasonable evidence has shown that the first three helices of colicin A are not involved in the formation of the pore and the following two helices are not essential in pore-forming process. It means only five to seven helices relate to the formation of the voltage-dependent channel (Baty et al., 1990, Collarini et al., 1987, Lindeberg et al., 2000, Nardi et al., 2001). Most or all of helices II to V were thought to be

translocated through the inner membrane of the cells in a later study (Cascales et al., 2007).

All of the colicins channels are highly selective for protons over other cations, however, colicin A, with a predicted 10-Å ion channel, has a proton flux that is so high that the proton selectivity is unexpected and far beyond expectation (Cascales et al., 2007). Furthermore the mechanism of the voltage-dependent channel is still unclear (Kienker et al., 2002).

1.6 Molten globule state of Colicin A Pore-forming domain

Introduction to the molten globule state

Generally, a protein should experience three stages in order to be folded into a precise three-dimensional arrangement before acquiring its full biological activity (Fridd and Lakey, 2002, Ptitsyn, 1995): secondary structure formation, intermediate association of secondary structure elements, and the final local adjustment. The molten globule was considered as useful model of the intermediate stage between the secondary structure formation and the final folded state (Evans et al., 1996).

The characteristics of a "molten globule state are briefly described as: (1) a native-like secondary structure, which is usually indicated by far UV CD spectra; (2) loss of tertiary structure, which is usually indicated by near UV CD spectra or ANS binding; (3) retention of some buried residues, which can be indicated by intrinsic fluorescence (only of the buried aromatic residues); (4) native-like compactness; and (5) absence of thermal transitions, which can be indicated by DSC and temperature dependent CD spectra (Ptitsyn, 1992, van der Goot et al., 1992b).

Molten globule state of colicin A pore-forming domain

With small, single domain proteins, the folding transitions are often regarded as simple two state systems, with measurable equilibrium constants between the folded and unfolded states. However, the molten globule state has sometimes been revealed as an intermediate step between the folded and unfolded state the unfolding process. An intermediate similar to the "molten globule state" (Wiener et al., 1997) has been found in the colicin A pore forming activity process (van der Goot et al., 1991).

Studies of the water soluble ColA-P protein in solution has shown that the protein folds to its native state at neutral pH and is partially unfolded at low pH, when it exhibits all of the characteristic of the protein "molten globule state" (van der Goot et al., 1991, van der Goot et al., 1992b). Further studies of the ColA-P interaction with artificial membrane vesicles showed that decreasing the pH of the membrane-solution interface can increase the insertion rate and the accumulation of intermediate proteins on the artificial membrane surface (van der Goot et al., 1991, van der Goot et al., 1992b). The fraction of protein in the molten globule state was proportional to the rate of membrane insertion. Hence, the molten globule state was considered a very important element in the colicin A pore-formation process.

1.7 The role of surface amino acid residues in protein stability and folding

Often, the molten globule state may be short-lived or absent, but in some cases it may be the most significant point for the study of a folding pathway. The previous study (Evans et al., 1996) demonstrated that the molten globule state is formed by the pore-forming domain of Colicin A (with a pI of 5.82) in acidic pH while the highly homologous colicin N (with a pI of 10.25) does not form a molten globule. The pore forming domain of Colicin B (with a pI of 5.48) is also able to form the molten globule state below pH 3, (Evans et al., 1996). Colicin N C-terminal domain has

seven fewer Asp/Glu residues than Colicin A or B, and thus it was considered that the absence of aspartate residues at key points was responsible for the inability of colicin N to form a molten globule state at low pH. Fridd and Lakey showed that inserting Asp residues into Colicin N Ala sites where colicin A has Asp had no effect upon the pore domains stability. However replacement of the corresponding individual Asp residues of Colicin A with Ala caused a significant destabilisation and even molten globule formation at neutral pH in some cases (Fridd and Lakey, 2002). Thus, the charge on only a few Asp residues is required for the stability of ColA. These residues are surface exposed and it was not clear how they stabilised the protein since surface residues are thought to contribute weakly to the stability of proteins and each Asp constitutes only 4% of the total charge.

Many pore-forming colicins such as colicin A (Baldwin et al., 2004, van der Goot et al., 1992b) and E1 (Musse and Merrill, 2003) become destabilised at low pH but only ColA has been shown to both form a distinct molten globule state and need an acidic environment to perform its lethal activity (see below) (van der Goot et al., 1993). Musse and Merrill (Musse and Merrill, 2003) demonstrated that charged acidic residues such as aspartate and glutamate can form ion pairs and hydrogen bonds in E1 helices 4 and 5, which could contribute to helix stabilization. Their studies showed that the substitution of these acidic residues enabled the mutants to bind to the membrane at higher pH values than the wild type. Thus, it was believed that those acidic residues were acting as a pH sensor in the toxic activity. However these residues are conserved in all colicins including colicin N which shows no pH sensitivity (Musse and Merrill, 2003). Thus, the very pH sensitive colicin A should possess a characteristic pH switch not shared by other colicins.

The role of negatively charged phospholipids in the colicin A pore forming action

Previous research (van der Goot et al., 1993, van der Goot et al., 1991) indicated that acidic lipids are involved in the protein translocation step *in vivo* and *in vitro.* The fluorescence emission spectrum of the Col-A pore-forming domain does not change after insertion into normal phospholipid vesicles but it does change after the addition of brominated dioleoylphosphatidylglycerol (Br-DOPG) vesicles which quench the tryptophan fluorescence. This enabled the role of different lipids to be measured (González-Mañas et al., 1992), and showed that negatively charged lipids and low pH are essential for the insertion step. Van der Goot *et al* (van der Goot et al., 1993) showed that colicin A, but not colicin N, was inhibited by removal of acidic lipids in the inner membranes of certain mutants of *E coli*. Thus, it appears that the negative surface charge is the essential for the formation of the molten globule state of colicin A pore-forming domain *in vivo* and *in vitro*. This is partly due to the reduced local pH at the membrane surface which amplifies the effect of the reduced pH of the media (van der Goot et al., 1991).

1.8. Project Aims

Pore-forming toxins present an important challenge to understand their translocation process. To understand the transition of water soluble Colicin A P-domain into a pore-forming protein, a study of the structural transition of Colicin A P-domain *in vitro* will be carried on in my studies. Since pH sensitivity is a real biological function of this protein (van der Goot et al., 1993) and a preliminary study (Fridd and Lakey, 2002) has shown that surface aspartate residues stabilize the folded state, I will investigate the link between these observation. This is rare because in most cases surface mutations have very slight effects on stability (Spector et al., 2000). The understanding of this will provide fundamental new information for the understanding

of protein stability. So, I am very interested in the further research of this phenomenon in my future studies. In order to study the pore forming process, and the mis-folding and mis-assembly processes of Colicin A P-domain, some mutants of Colicin A P-domain have been made (Fridd and Lakey, 2002)I shall make a larger set in order to fully understand their role in the stability and unfolding of this protein.

Chapter 2

Materials and Methods

2.1 Materials

All chemicals were either from Sigma-Aldrich (Poole, Dorset, UK) or from Melford Laboratories Ltd (UK) unless otherwise stated. Ampicillin, DNase I and RNase (Ribonuclease) were purchased from Sigma-Aldrich. For molecular biology, genes were synthesised and ligated into plasmid pET3c by Blue Heron Biotechnology. Ni-NTA resin was purchased from Qiagen. Wizard Plus SV Minipreps DNA purification system "Mini prep kit" was purchased from Promega. Centrifugation tubes were purchased from Beckman Coulter.

2.2 Microbiology

Bacterial strains The bacterial strains used throughout the project and their genotypes are listed in table 2.1

***E. coli* strain**	**Genotypes**	**Antibiotic resistance**	**Source**
***E. coli* BL21(DE3)**	F$^-$ *ompT hsdS*B(rB$^-$, mB$^-$) *gal dcm* **(DE3)**	Ampicillin	Life Technologies Ltd
***E. coli* BL21-AI**	F$^-$ *ompT hsdS*B(rB$^-$, mB$^-$) *gal dcm ara*B::*T7RNAP-tetA*	Ampicillin	Life Technologies Ltd
***Epicurian coli.* XL1-blue**	*recA1 endAq gyrA96 thi-1 hsd R17 supE44 relA1lac*[F' *proAB lacI*^{q} *Z ∆M15* Tn*10*(Tet^r)]	Tetracyline	Agilent Technologies UK limited

Escherichia coli (*E. coli*) BL21 (DE3) competent cells were used as an expression host. The genes for colicin A pore-forming domain wild type, D191A, D216A, D227A or D372A were ligated into pET8c vector expressed under the control of the bacteriophage T7 promoter from the host strain. This strain contains a copy of the T7 RNA polymerase under the control of *lac* UV5 promoter and *lac* I gene (Life Technologies). The colicins were transcribed and expressed following induction of the lac promoter and T7 RNA polymerase by the non-metabolisable lactose homolog isopropyl β-D-1-thiogalactopyranoside (IPTG).

Epicurian Coli XL1-Blue supercompetent cells were used to amplify plasmids for purification.

BL21-AI competent cells were used to express Hexa histidine tagged Colicin A pore-forming domain wild type and all the remaining mutants. This strain carries a gene of T7 polymerase under the control of *araBAD* promoter.

Chemically competent cells

Competent cells are used in transformations for cloning and protein expression because they can take up plasmid DNA. 5µl of frozen BL21(DE3) cells were streaked onto an LB agar plate with appropriate antibiotic and incubated at 37 ^{o}C for 16 to 20 hours. A single colony from the streak culture was used to inoculate 5 ml of LB media with the relevant antibiotic and grown at 37 ^{o}C with shaking for 16 to 20 hours. 1 ml of the culture grown from a single colony was used to inoculate 50 ml of LB media with the relevant antibiotic, and grown at 37 ^{o}C with shaking until the cells reached an optical density of 0.3 to 0.4 at 600nm. The cells were chilled on ice for 10 minutes, and then pelleted by centrifugation for 4 minutes at 3000 g at 4 ^{o}C. The supernatant was discarded and the pellet was resuspended in 50 ml frozen storage buffer (FSB) and chilled on ice for 20 minutes. The cells were re-pelleted by centrifugation for 4 minutes at 3000 g, 4 ^{o}C and resuspended in 8 ml 0 ^{o}C FSB then divided into 0.4 ml

aliquots and flash frozen on liquid nitrogen. Finally the chemically competent cells were stored at -80 ^{o}C.

Frozen storage buffer (FSB): 100mM KCl, 50mM CaCl2, 10% (w:v) glycerol and 10 mM potassium acetate pH 6.2, Dimethylsulphoxide (molecular biology grade) was added to a final concentration of 7% (v:v), autoclaved and stored at 4 ^{o}C.

Transformations

Under certain conditions, exogenous DNA can be taken up by pre-treated *E. coli* cells. If the DNA carries a gene that confers resistance to an antibiotic, cells transformed with the plasmids can be easily selected by growing them on an antibiotic-containing medium (Muga et al., 1993b). Plasmid DNA and frozen competent cells were thawed on ice. 5µl of plasmid DNA was added to 5µl of *E. coli* BL21-AI cells were in transformation buffer and incubated on ice for 30 minutes. The mixture was heat shocked at 42 ^{o}C for 45 seconds and then rested on ice for 5 minutes. 500 µl of LB media was added to the mixture and then incubated at 37 ^{o}C with shaking for 1 hour. 20 µl and 200 µl of cultures were then spread on two LB agar plates with the appropriate antibiotics. The plates were incubated at 37 ^{o}C for 16 to 18 hours.

E. coli (AI) cells transformation buffer: 100mM MOPS, 50mM $CaCl_2$ pH 6.5

Overnight bacterial cultures

One colony was picked out from a transformation plate and used to inoculate 50ml of LB media with the appropriate antibiotics. The cultures were incubated at 37 ^{o}C with shaking for 16-18 hours.

2.3 Molecular Biology

Techniques related to DNA purification and protein purification are described in the following sections.

Plasmids

The plasmid used throughout the project and their vectors are listed in table 2.2

Plasmid	Protein	Gene length bp	Vector
ColAWT	Colicin A-P domain W.T	1370	pET3a
ColAD241A	Colicin A-P domainD241A	1370	pET3a
ColAD274A	Colicin A-P domainD274A	1370	pET3a
ColAD376A	Colicin A-P domainD376A	1370	pET3a
ColAD191N	Colicin A-P domainD191N	1370	pET3a
ColAD216N	Colicin A-P domainD216N	1370	pET3a
ColAD227N	Colicin A-P domainD227N	1370	pET3a
ColAD241N	Colicin A-P domainD241N	1370	pET3a
ColAD274N	Colicin A-P domainD274N	1370	pET3a
ColAD372N	Colicin A-P domainD372N	1370	pET3a
ColAD376N	Colicin A-P domainD376N	1370	pET3a
ColAD216E	Colicin A-P domainD216E	1370	pET3a
ColAD216Q	Colicin A-P domainD216Q	1370	pET3a
ColAS239A	Colicin A-P domainS239A	1370	pET3a
ColAS239AD241A	Colicin A-P domainS239AD241A	1370	pET3a
ColAD274E	Colicin A-P domainD274E	1370	pET3a
ColAD274Q	Colicin A-P domainD274Q	1370	pET3a
ColAD372E	Colicin A-P domainD372E	1370	pET3a
ColAD372Q	Colicin A-P domainD372Q	1370	pET3a

The protein column describes the protein that is expressed with 0.2% L-arabinose. Plasmids ColAD191A, ColAD216A, ColAD227A and ColAD372A were ligated into pET8c vector by Susan Fridd and protein expression induced with ITPG.

The protein amino acid sequences and the respective DNA sequences of Colicin A pore forming domain that were used in this project are shown in Appendix 1.

Agarose gel electrophoresis

Agarose gel electrophoresis was used to determine the approximate size of plasmids, and the DNA sample was cut out of the gel for further analysis. When an electric field is applied the negatively charged DNA fragments move toward the anode at a rate dependent upon their size. The Relative Molecular Mass (M*r*) of a DNA fragment can be determined by comparison to a standard DNA marker of known M*r*. DNA sample in agarose gel can be cut out from the gel, returned to solution and recovered by heating the gel to 65 ^{o}C. Plasmids ColAD227A and ColNN191D were digested by BamHI restriction enzyme. The band was cut out of the gel and the recovered DNA fragments were sent for DNA sequencing (Vandergoot et al., 1993).

Digestion reaction: 1 µl BamHI, 1 µl T4 DNA ligase buffer, 1 µl BSA (10X), 2µl Plasmids and 5µl sterilized water incubated at 37 ^{o}C for 1 hour.

To make a 1% agarose gel: 1 g agarose and 100 ml TBS buffer were added into a small flask. The flask was placed in a microwave oven and heat at 100% power until the agarose was completely dissolved. 5 µl ethidium bromide was added to the flask when it cooled down to approximately 60 ^{o}C. Then the solution was poured into a mould without generating bubbles and a comb was inserted into the agarose gel. The gel was chilled for 30 minutes before electrophoresis.

1 litre 1XTBS buffer: 108 g Tris base, 55 g boric acid and 40 ml 0.5 M EDTA, pH8.0

Plasmid DNA purification

Small scale (from 5 ml culture) plasmid DNA extraction and purification was carried out using the Promega SV Wizard mini-prep Kit. Large scale (from 50 ml culture)

plasmid DNA extraction and purification was carried out using Qiagen Plasmid Midi-prep Kit.

2.4 Biochemical techniques

Sodium dodecylsulphate-polyacrylamide gel electrophoresis (SDS PAGE)

Sodium dodecylsulphate-polyacrylamide gel electrophoresis was used to monitor protein purification in the project. The pores in a SDS gel are much smaller than the pores in an agarose gel, so it is commonly used to separate proteins. Since the SDS gel separates protein according to size and shape, the proteins were usually boiled for 5 min in sample buffer to reduce disulphide bridges and denatured by SDS before electrophoresis in order to reduce the influence of the protein shape. The mass of protein can be determined by comparing its mobility with protein standard of known molecular weight that is run alongside on the same gel (Tuszynski and Kurzynski, 2003).

15% Sodium dodecylsulphate-polyacrylamide gel: 10 ml resolving gel was made from 3.38 ml acrylamide/bis acrylamide, 40% solution, 2.25 ml water and 3.38 ml Lower buffer and polymerized by adding 50 µl 10% ammonium persulphate (APS) and 20µl N,N,N',N'-tetramethylethylenediamine(TEMED). 4 ml stacking gel was made from 0.35 ml Acrylamide/bis acrylamide, 40% solution, mixed with 2.5 ml water and 1 ml Upper buffer and achieved polymerisation by adding 50µl APS and 10 µl TEMED.

Lower buffer: 182 g Trizma Base, 4 g SDS, pH 8.8

Upper buffer: 60.5 g Trizma Base, 4 g SDS, pH 8.0

Protein sample buffer: 15% (v/v) glycerol, 0.125 M Tris, 2% (w/v) SDS, 0.1% (w/v) bromophenol blue, 1% (w/v) β-mercaptoethanol, pH6.8

SDS PAGE stain: 0.05 % Coomassie Brilliant Blue, 10% (v/v) glacial acetic acid, 10% (v/v) propan-2-ol

Native polyacrylamide gel electrophoresis

Native polyacrylamide gel electrophoresis was used to detect the biological activity such as identification of the monomer and dimer structure of protein in non-denaturing conditions. The proteins were separated only according to their different electrophoretic mobilities and the friction resistance of the gel. The protein was pre-treated separately by EDTA (Ethylenediaminetetraacetic acid) and Ni2+. 5%EDTA can deactivate protein by destroying the dimer complex of the protein. 5% Ni2+ can catalyse the formation of dimer. Original (ColA-P WT without any treatment) protein and bovine serum albumin (BSA) were prepared as standards. All samples were mixed with sample buffer (without SDS) before the electrophoresis (Tuszynski and Kurzynski, 2003).

7X Native Gel Upper (Stacking) Buffer: 5.7g Tris base pH to 6.7 with H_3PO_4 water to 100ml

4X Native Gel Lower (Resolving) Buffer: 18.2 Tris base to pH 8.9 with HCl water to 100ml

50X Running Buffer: 7.5g Tris base 36g Glycine Water to 250 ml

Determining protein/DNA concentration by UV absorbance/Nano photometer (Nanodrop)

Ultraviolet-visible absorbance with a Nano photometer (Nanodrop-Labtech International Ltd) was used to measure the purified protein concentration during the project. There are three tyrosine residues and three tryptophan residues in Colicin A pore forming domain wild type and its mutant. Those aromatic residues exhibit a peak absorbance signal at a wavelength close to 280nm. However nucleic acids give an

absorbance as much as 10 times that of protein at this wavelength. Hence the low absorbance at 260 nm is an indication that there is no contamination by nucleic acids. The molar extinction coefficients of proteins can be calculated by multiplying the number of those aromatic residues with their respective molar extinction coefficients (Wu and Filutowicz, 1999, Vandergoot et al., 1993). The concentration of the protein can be calculated by using the Beer-Lambert law:

$A = \varepsilon_{280nm} cl$ (Vandergoot et al., 1993)

Where ε = the molar extinction coefficient for the absorber

c = molar concentration of protein solution

l = the pathlength through the solution (1 cm pathlength Hellma quartz cuvette was use in the project, so here l = 1cm)

The extinction coefficients and molecular weights of protein used in the project are listed in table 2.3

Protein	Mr	Extinction coefficient(ε_λ) $M^{-1}cm^{-1}$ at 280nm
Colicin A-P domain W.T	22918.3	20970
Colicin A-P domainD241A	22874.9	20970
Colicin A-P domainD274A	22874.9	20970
Colicin A-P domainD372A	22874.9	20970
Colicin A-P domainD376A	22874.9	20970
Colicin A-P domainD191N	22917.3	20970
Colicin A-P domainD216N	22917.3	20970
Colicin A-P domainD227N	22917.3	20970
Colicin A-P domainD241N	22917.3	20970
Colicin A-P domainD274N	22917.3	20970
Colicin A-P domainD372N	22917.3	20970
Colicin A-P domainD376N	22917.3	20970
Colicin A-P domainD216E	22921.3	20970
Colicin A-P domainD216Q	22922.3	20970
Colicin A-P domainS239A	22912.9	20970
Colicin A-P domainS239AD241A	22858.9	20970
Colicin A-P domainD274E	22921.3	20970
Colicin A-P domainD274Q	22922.3	20970
Colicin A-P domainD372E	22921.3	20970
Colicin A-P domainD372Q	22922.3	20970
Colicin A-P domain D191A	22874.9	20970
Colicin A-P domain D216A	22874.9	20970
Colicin A-P domain D227A	22874.9	20970
Colicin A-P domain D372A	22874.9	20970

All proteins are starting with an N-terminal His tag (MHHHHHHSS). Colicin A-P domain S239AD241A contains two amino acids change of residues Ser239 and Asp241. The original amino acid residues are substituted by alanine. All other mutants have one amino acid change.

To accurately measure the protein concentration, when the protein concentration was lower than 1 mg/ml, the Cary UV-Vis spectrophotometer was used to measure the protein concentration; when the protein concentration was higher than 1 mg/ml, the Nanodrop ND-100 photometer was used.

Size exclusion chromatography

Size exclusion (Gel Filtration) chromatography was used to separate proteins on the basis of their molecular size in the project. In gel filtration, distribution of particular compounds is dependent on its molecular size, which is represented by distribution coefficient (K_d). In this technique the large protein molecules will flow through the interstitial spaces between gel particles and will appear first in the eluate. The smaller molecules will pass through the pores of the gel particles and will appear in the eluate later. The distribution coefficient, K_d, for a given type of gel is an approximately linear function of the logarithm of protein molecular weight. If the molecule is large enough and completely excluded from the pores of the gel, $K_d = 0$ whereas, if the molecule is small it has complete access to the pores, $K_d = 1$. Elution volumes (Vs) of two analytes can be shown as below:

$Vs = (K'_d - K''_d)\ Vi$

Where Vi is the elution volume when $K_d = 1$.

K'_d and K''_d are the distribution coefficient of two analytic compounds.

The distribution coefficient K_ds of protein can be calculate as below:

$K_d=(V_R-V_0)/Vi$

The slope of the linear calibration curves can be worked out using the logarithm of the molecular weights of two given proteins and their K_ds as below:

$K=(\log'_{MW} -\log''_{MW})/(K''_d - K'_d)$

The distribution coefficient K'_d of an unknown protein can be worked out as below:

$K'_d=(V_R-V_0)/Vi$

A series of proteins PA heptamer, BSA, Aprotinin, Carbonic Anhydrase, β-amylase, ColN-full length, cPA83 and V-antigen were used as standards to achieve the linear section by plot of the logarithm of their molecular mass and their K_d s. The distribution coefficient of an unknown protein can be worked out from its elution volume (Switzer and Garrity, 1999).

Urea unfolding/guanidine unfolding

Urea and guanidine were used to unfold Colicin A pore forming domain and its mutations at different pH in the project. Although, the mechanism of the action of these agents is not fully understood, proteins can be unfolded gradually in increasing concentration of urea and guanidine solution by disrupting a protein's non-covalent bonds (Lemaster and Richards, 1982, Pace et al., 1989).

2.5 Protein purification

Colicin A-pore forming domain wild type and its mutants were expressed using *E. coli* competent cells and purified using the hexa-histidine tag on a Ni-NTA affinity resin in the project (Evans et al., 1996).

Large-scale cultures for protein expression

Large-scale cultures have been widely used for protein expression. All proteins (see table 2.3) with a His tag were expressed using *E. coli* BL21-AI cells or BL21 (DE3) cells. After transforming cells with the ColA plasmids, one single colony was picked from ampicillin LB agar plate and inoculated in 50 ml LB media (with 100µg/ml ampicillin). The cells were cultivated for 16 to 18 hours at 37 ^{o}C in an orbital incubator, shaking at 180 rpm. 4 X 500 ml LB (containing 100µg/ml ampicillin) in 2L shake flasks were inoculated with 5 ml culture and grown at 37 ^{o}C with shaking until the turbidity of cells reached an OD_{600} of 0.8-1.0. The expression of protein was induced by 0.2% arabinose or 1mM isopropyl ß–D- thiogalactopyranoside (ITPG). The cell culture was grown for a further 3 hours until the turbidity of cells reached an OD_{600} of 1.8-2.0, by this point the protein has been expressed. The cells were pelleted by centrifugation using Beckman Avanti series centrifuge JA-10 rotor at 10k × g for 10 minutes at 4 ^{o}C.

Large-scale cultures for ^{15}N labelled protein Expression

^{15}N labelled proteins were expressed using M9 media. Comparing with the routinely used large-scale LB medium, M9 media is relatively simple having less nutrients. Hence, the cells grow slower than in LB media. Before setting up a large-scale culture, the cells from a pre-culture were pelleted and re-suspended in M9 media. On average, 4 to 5 hours were needed until the turbidity of cells reached to OD_{600} of 0.8-1.0; another 4 hours were needed for protein expression.

M9 media preparation:

Stock salt solutions: 1M $CaCl_2$, 1M $MgSO_4$, 10X M9 salts (60g Na_2HPO_4, 30g KH_2PO_4, 5g NaCl) autoclave

Fresh prepared solutions: 1% thiamine, 0.01 M $FeCl_3$, 20% (v/v) D-glucose, 100 mg/ml ampicillin

In each 2 L shake flask: 450 ml autoclaved dH_2O, 50 ml M9 salts, 150 µl $CaCl_2$, 500 µl $MgSO_4$, 10 ml D-glucose, 5 ml MEM vitamins, 100 mg/ml ampicillin were added. (Filter sterilised using 0.22 µM filters)

Histidine tagged Colicin A pore-forming domain purification

Affinity chromatography using Ni-NTA agarose affinity resin in which a transition metal ion Ni^{2+} was used to bind the histidine tagged proteins selectively by reaction with the imidazole groups of histidine residues. The proteins were subsequently eluted and recovered using an excess of imidazole.

The first step in purification was the preparation of soluble protein from crude extracts. Firstly, the cell pellets were re-suspended in column loading buffer. DNase I, RNase, Lysozyme (from chicken egg white), and protease inhibitor cocktail were subsequently added to the cell suspension. The cells were then stored in an ice bucket and disrupted using a sonicator for 30 minutes with 80% power. After sonication, the unbroken cells and insoluble cytoplasmic organelles were pelleted using a Beckman Avanti series centrifuge JA-20 rotor at 18k × g for 60 minutes at 4 ^{o}C. All the soluble proteins and the histidine tagged protein were dissolved in supernatant.

In the second step, the histidine tagged protein was bound to the Ni-NTA affinity resin column and then eluted with imidazole. Firstly, the column was equilibrated with 25 ml column loading buffer. The supernatant was subsequently loaded onto the column and the flow through was collected for further analysis. The column was then washed with 2 X 5 ml column loading buffer, 3 X 5 ml column washing buffer and 8 X 3 ml column elution buffer in order. All the fractions were collected separately and labelled as L1, L2, W1, W2, W3, E1, E2, E3, E4, E5, E6, E7 and E8 respectively.

In the last step, samples of flow through and all fractions were analyzed using SDS gel electrophoresis. The fractions containing the protein of interest were pooled and dialysed to remove the imidazole.

Column loading buffer: 50 mM Phosphate buffer pH8.0, 300 mM NaCl

Column washing buffer: 50 mM Phosphate buffer pH 8.0, 300 mM NaCl, 25 mM imidazole

Column elution buffer: 50 mM Phosphate buffer pH8.0, 300 mM NaCl, 350 mM imidazole

2.6 Biophysical techniques

Refractometer

Refractometry is an optical spectroscopy method that measures the refractive index of solutions. The molarities of the urea and guanidine hydrochloride stock solutions were calculated from the difference in refractive index between the denaturant solutions and buffer ΔN. The equation for urea solution is M(urea) = $117.66(\Delta N)+29.753(\Delta N)^2+185.56(\Delta N)^3$ and the equation for guanidine solution is M (guanidine HCl) = $57.147(\Delta N)+38.68(\Delta N)^2-91.60(\Delta N)^3$. (Lakey et al., 1991b, Pace et al., 1989)

For example: if the refractive index of the urea stock solution is 1.412 and of the buffer is 1.336.

ΔN = 1.412-1.336=0.076

M(urea) = $117.66(0.076)+29.753(0.076)^2+185.56(0.076)^3$=9.1026 M

Circular dichroism spectroscopy

Circular dichroism is an optical spectroscopy method in molecular biophysics that gives information about the secondary structure of proteins by measuring the difference between left and right circularly polarised light absorbed by asymmetric molecules.

Electromagnetic waves are composed of coupled electric and magnetic fields, which oscillate along directions perpendicular to the propagation direction of the wave. Light as a high frequency electromagnetic wave can circularly polarise when the electric vector rotates along the direction of propagation with a constant length. The right circularly polarised light is formed when the vector rotates as a right handed-helix. The left circularly polarised light is formed when the vector rotates as a left-handed helix. The linearly or plane polarised light can be resolved into a left and a right circularly polarized light of equal magnitude. Ellipticity is formed when the left and the right circularly polarized light of unequal magnitude are superimposed. In CD it happens when a plane-polarised beam is shone through an optically active sample and the left and right circularly polarised light are absorbed differently (Wu and Filutowicz, 1999).

Ellipticity is defined as angle Θ, which is the tangent of the ratio of the minor to major axes of the ellipse traced by the tip of the electric vector of the transmitted beam.

The optically rotation is defined by the angle α, which is the major angle axis of ellipse and incident light.

The ellipticity Θ and the optical rotation angle α of elliptically polarised light are shown in figure 2.2.

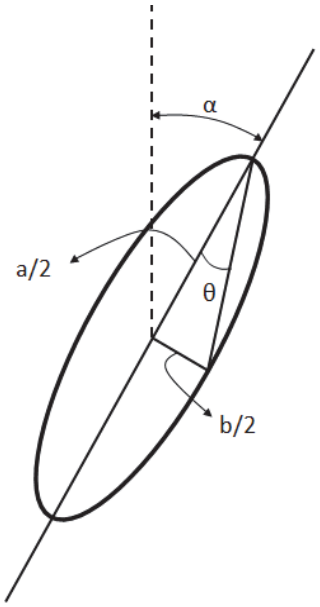


Figure 2.1 Ellipticity Θ and the optical rotation angle α of elliptically polarised light.

The ellipse is formed by the electric vector projection in elliptically polarised light.

Where Θ =ellipticity angle

α=the optical rotation angle

a/2 =semi-major axes

b/2 =semi-minor axis

According to Beer-Lambert law, the absorption of a certain polarised light A, can be expressed in terms of the decadic molar extinction coefficient ε:

$A= \varepsilon Cl$

CD is defined as the difference between the absorption of the two orientations of circularly polarised light for a chiral material such as a protein:

$\Delta A=A_L-A_R=\varepsilon_L Cl -\varepsilon_R Cl=\Delta\varepsilon Cl$

Where ΔA=the difference between the absorption of the two orientations of circular polarised light

Al=the absorption of the left circular polarised components of the transmitted light

Ar= the absorption of the right circular polarised components of the transmitted light

Δε=decadic molar CD

C=molar concentration of the sample

I=optical path length of cuvette

CD can also be defined as the molar ellipticity [Θ], it can be expressed in terms of ellipticity, Θ, and ΔA:

$\Theta = 32.98\ \Delta A$

$[\Theta] = 3298\ \Delta\varepsilon$

Ellipticity, Θ, is in now in degrees (Wu and Filutowicz, 1999).

The CD signal is strong in the far UV region (170 nm-250 nm), where the peptide bond absorbs; therefore it is the most appropriate technique to monitor the secondary structure of a protein in solution. According to analysis of the protein structure database, colicin A pore forming domain is an α-helical protein. The CD spectra for a classic α-helix show positive bands below 203 nm with a strong maximum at around 192nm resulting from ππ* transitions; the negative bands between 203 nm and 240 nm has minima at around 209 nm and 222 nm, caused by ππ* and nπ* transitions. The CD spectra of β proteins are various according to their spectra and β structure type, however the regular β protein has maximum from 190 nm to 200 nm and minimum between 210 nm and 225 nm. The most abundant secondary structures of protein are the α-helix and β-sheet.

The far UV CD spectra of α-helix and β-sheet are shown in figure 2.3.

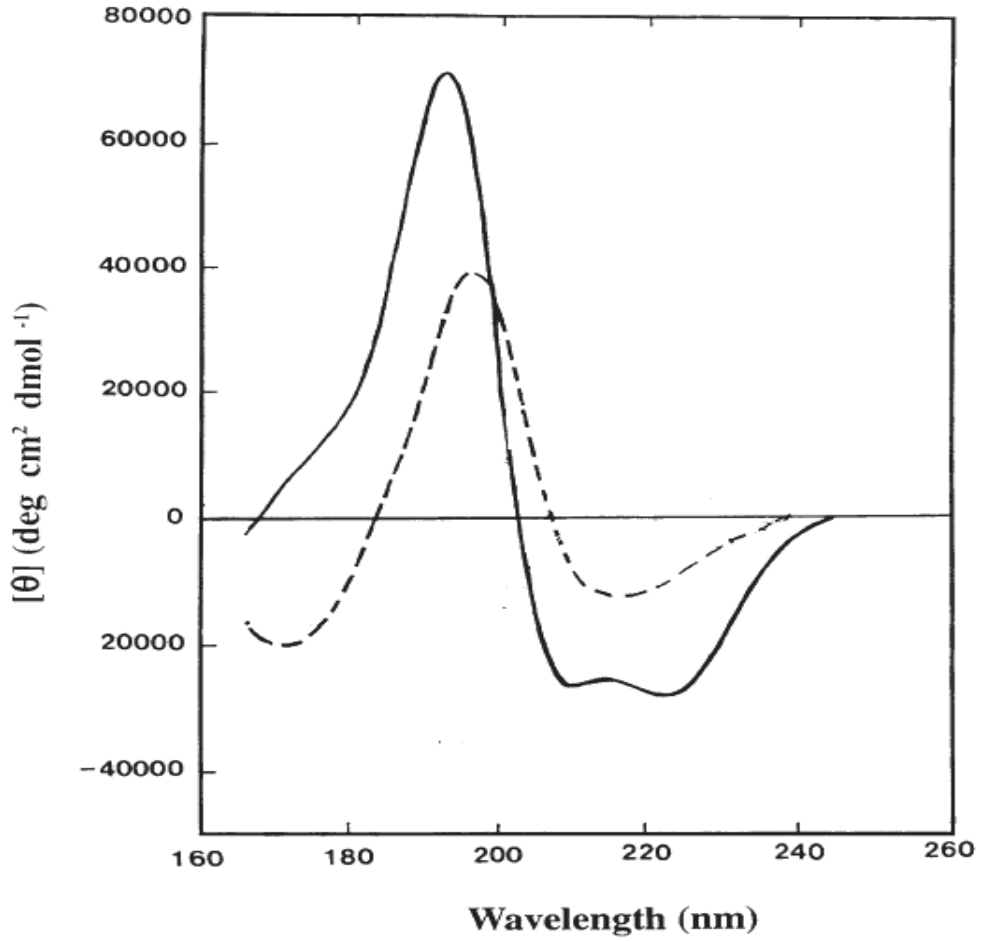


Figure 2.2 Far UV CD spectra of α-helix and β-sheet

Where the solid line indicates is the far UV CD spectra of α-helix, the dashed line is the far UV CD spectra of anti-parallel β-sheet (Kelly et al., 2005).

Because the peptide bond region between 209 nm and 222 nm is sensitive to the presence of α-helix, the folding state of α-helix protein such as colicin A pore forming domain and its mutants can be analysed by comparing the proportion of α-helix content. The proportion of helical structure can be expressed in terms of decadic molar CD, Δε, therefore a low Δε at 209 nm and 222 nm represents more α-helix content (Serdyuk et al., 2007).

The CD signal is also strong in the near UV region (250 nm-320 nm), where the aromatic amino acids absorb; therefore it is also appropriate to assess the tertiary structure of a protein. Colicin A pore forming domain contains three Tyr, three Trp and five Phe. The absorption spectrum close to 290 nm with fine structure between 290 nm and 305 nm corresponds to Trp residues. The bands between 275 nm and 282 nm probably represents the Tyr residues. A weaker and sharper peak close to 255 nm relates to Phe residues (Serdyuk et al., 2007).

Protein concentration, cuvette path length and the instrument parameter for near UV and Far UV CD spectra in this project were as below (table 2.4)

Parameter	**Far UV CD**	**Near UV CD**
Protein concentration (mg/ml)	0.5	2
Cuvette path length (cm)	0.02	0.5
Wavelength (nm)	190-250	250-320
Band width (nm)	2	2
Response time (sec)	4	4
Scanning rate (nm/min)	20	20
Data pitch (nm)	0.5	0.5
Accumulations	10	10
Temperature (^{o}C)	20	20

Protein concentration, cuvette path length and the instrument parameter for Near UV and Far UV thermal melting experiment in this project was settled as below (table 2.5)

Parameter	**Far UV CD**	**Near UV CD**
Protein concentration (mg/ml)	0.5	1
Cuvette path length (cm)	0.1	1
Wavelength (nm)	223	295
Band width (nm)	2	2
Response time (sec)	32	32
Scanning rate (^{o}C /min)	1	1
Data pitch (^{o}C)	0.2	0.2
Temperature (^{o}C)	20-90	20-90

Fluorescence spectroscopy

Fluorescence spectroscopy is a sensitive biophysical technique to assess the changes in the local environment of a chromophore, therefore it is an excellent probe to detect the conformation changes of protein, such as unfolding.

Fluorescence begins as a result of a molecule being excited from its ground electronic state to one of the vibrational states by absorbing a photon. Then the excited molecules lose vibrational energy until it reaches the lowest vibrational state of the excited electronic state. Because the molecule now can drop down from these various vibrational levels to the ground electronic state, the emitted photon can have a different energy. Holding the excitation light at a constant wavelength, the different frequencies of fluorescent light (emission wavelength) emitted by a sample can be measured and recorded.

The fluorescence comes from the aromatic amino acids within proteins. The environmental sensitivity of tryptophan fluorescence is much higher than the tyrosine and phenylalanine. The emission of tyrosine in water occurs at 303 nm and is insensitive to the local polarity. The emission of phenylalanine occurs at 282 nm, which is much shorter than tyrosine and tryptophan. The emission of tryptophan in water occurs at 350 nm and is highly sensitive to the microenvironment (Lakowicz, 2007). In a hydrophobic environment, the tryptophan emission maximum is usually shifted to smaller wavelengths. In an unfolded protein, when the tryptophan is exposed in aqueous environment, the fluorescence intensity is usually decreased and the emission is usually red-shifted. Thus, the intrinsic tryptophan fluorescence can be used as an important probe to monitor the unfolding state of proteins. (Creighton, 1997)

A Cary Eclipse fluorescence spectrophotometer from Varian was used to study the excited state properties of Colicin A pore-forming domain and its mutants in the

project. Three tryptophan residues (Trp-86, Trp-130, and Trp-140) are contained in the pore forming domain of colicin A. If the colicin A pore-forming domain is denatured by increasing concentration of Urea or Guanidine, a red-shifted emission spectrum will appear, which means the tryptophan residues is exposed in an aqueous environment and the protein is unfolded. The fluorescence emission curves are collected from 295 nm to 440 nm and the excitation wavelength is 280 nm. Fluorescence spectra were obtained using a 0.5 cm pathlength cuvette at a protein concentration of 0.03 μg/ml.

In order to understand the mechanism by which surface mutations destabilize the folding state of ColA-P, quantitative data on the effects of the mutants was obtained.

Fluorescence spectra of ColA-P Ala mutants, Asn mutants and the wild type were measured at pH 8.0 in urea solution concentration range from 0 M to 9.1 M and in guanidine solution concentration range between 0 M and 7.8 M. The barycentric means (BCM) of the emission wavelength were calculated using Microsoft Office Excel and the formula: BCM $\lambda = \frac{\sum F \times \lambda}{\sum \lambda}$ in the range of λ = 295-440. The BCM values for urea and guanidine unfolding experiments reveal the two transitions of the ColA-P unfolding process and by fitting a three-state model to the fluorescence data the free energy of folding in water ΔG (H_2O) can be calculated (Wilson and Walker, 2010). Here ΔG is the free energy difference between the folded state and unfolded state of ColA-P in the isothermal and isotonic system. ΔG_1 is the free energy between the folded state and molten globule state; ΔG_2 is the free energy between the molten globule state and folded state. When ColA-P folded = ColA-P molten globule, the $K_{eq1} = \frac{\text{ColA-P folded}}{\text{ColA-P molten globule}} = 1$; when ColA-P molten globule= ColA-P unfolded, the $K_{eq2} = \frac{\text{ColA-P molten globule}}{\text{ColA-P molten globule}} = 1$; The equilibrium constant K_{eq1} and K_{eq2} can be calculated for any point on the unfolding curve but far away from 1, the accuracy of measurement decreases. Thus, it is necessary to fit the data to a three state model to

determine ΔG (H_2O) in the absence of urea. The ΔG in the absence of urea can be calculated by fitting to and extrapolating the ΔG =ΔG1+ΔG2=-($RTInK_{eq1}$+ $RTInK_{eq2}$) data for a series of urea concentrations (Bouveret et al., 1998).

Differential scanning calorimetry

Differential scanning calorimetry is a biophysical technology that can be used to measure the heat capacity change of a soluble protein in solution for a designed temperature range at a constant pressure. A molar heat capacity of a protein as a function of temperature curve can be obtained in a DSC experiment. Purity of protein sample solution for calorimetry experiments should be less than 3% of contaminants and the absolute concentration should be measured in advance. (Serdyuk et al., 2007).

The molar heat capacity C_p of protein versus temperature curve is shown in figure 2.3.

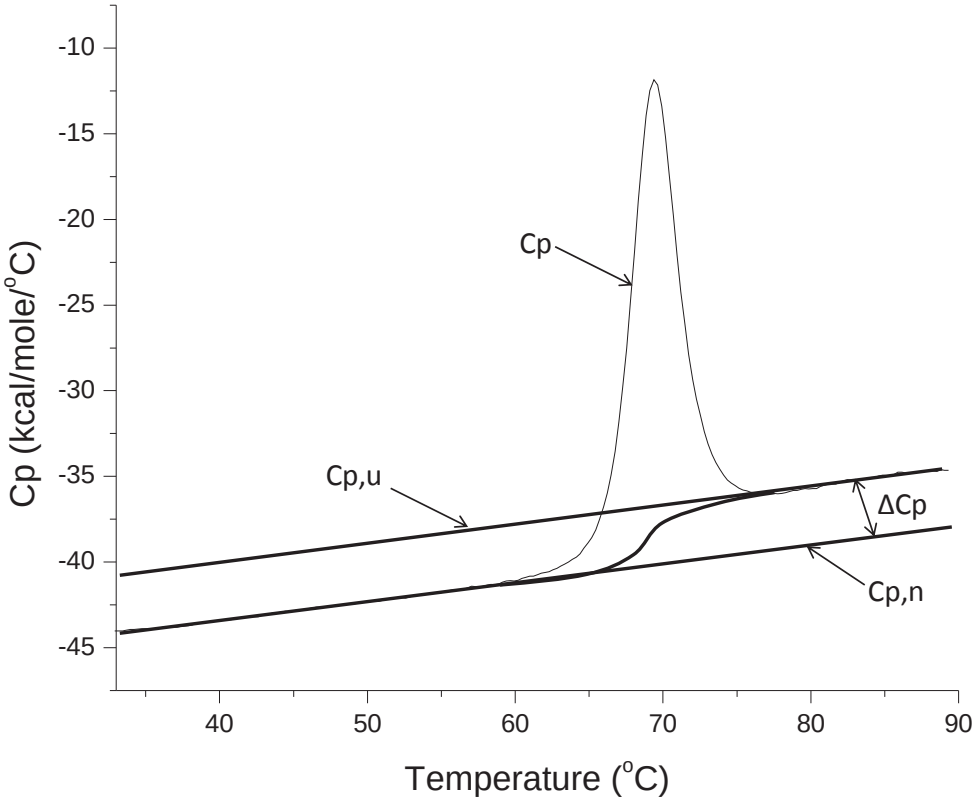


Figure 2.3 An idealised molar heat capacity C_p of a protein versus temperature

Where C_p =the molar heat capacity of the protein

$C_{p,u}$ =the molar heat capacity of the unfolded states

$C_{p,n}$ =the molar heat capacity of the native states

ΔC_p = the heat capacity change associated with the transition

The experimental calorimetric transition enthalpy $\Delta H_{tr,exp.}$ can be obtained from the DSC curves, which can be given by $\Delta H_{tr.exp.} = \int \Delta C_{p,tr.,exp}\,(T)\,dT$.

Where $\Delta C_{p,tr.,exp}$ (T)= the molar capacity in the transition peak above experimental baseline $C_p(T_{tr})$ H_{tr}

The heat capacity curve after baseline subtraction is shown in Figure 2.4.

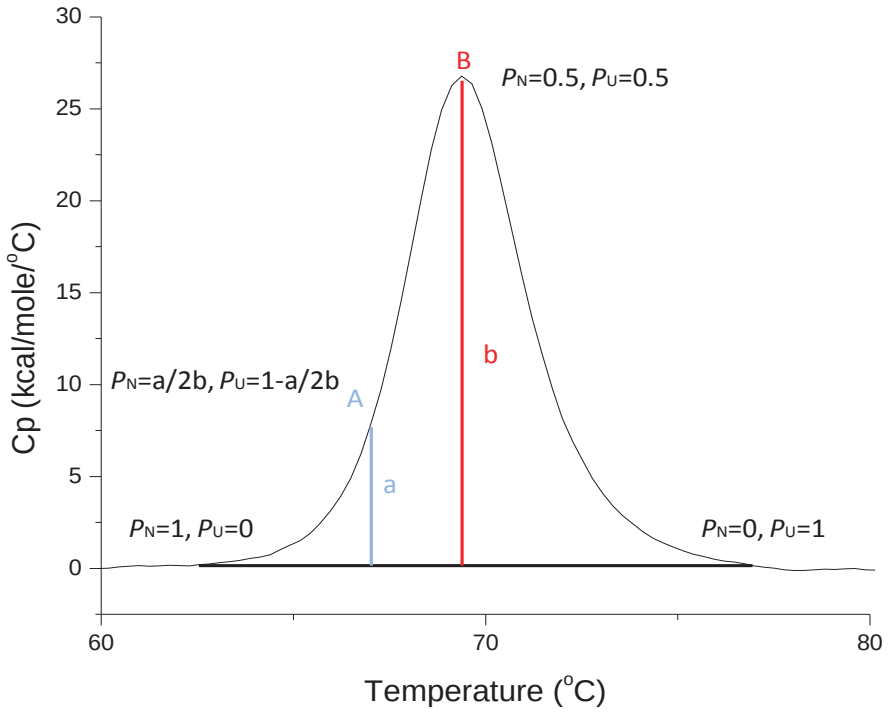


Figure 2.4 The heat capacity curve after baseline subtraction

Where P_N = the population fraction of native states

P_U = the population fraction of unfolded states

For a two state transition, the van't Hoff transition enthalpy can be obtained from the DSC curve, which can be given by $\Delta H_{v't H.} = 4RT^2 \frac{\Delta Cp(Ttr)}{\Delta Htr}$.

Where $\Delta C_p(T_{tr})$ = the heat capacity value at the peak point

ΔH_{tr} = the calorimetric enthalpy absorbed in the transition peak

The calorimetric enthalpy represents the actual thermodynamic energy change of one mole of protein in the reaction system. The van't Hoff enthalpy can reflect the thermodynamic energy change based on the transformation per mole in the system.

When the cooperative unit is the whole protein the van't Hoff enthalpy should be equal to the calorimetric enthalpy. If the cooperative unit is smaller than the protein molecule, as it would be for a dimer, the van't Hoff enthalpy is smaller than the calorimetric enthalpy. On the other hand, if the van't Hoff enthalpy is larger than the calorimetric enthalpy, the cooperative unit is larger than the protein molecule, suggesting the formation of intermediates.

The differential scanning calorimetry parameter in this project is given as below (table 2.6)

Parameter	**ColA P-domain samples**
Starting temperature	25 ^{o}C
Final temperature	90 ^{o}C
Scan rate	1 ^{o}C/min
Pre-Scan Thermostat	15 min
Post-Scan Thermostat	0 min
Filtering period	16 sec
Pre-Scan temperature	25 ^{o}C

Analytical Ultracentrifugation

Analytical ultracentrifugation (AUC) is a technique for accurate determination of purity, mass, shape and binding properties of macromolecules in solution (Serdyuk et al., 2007). The sedimentation coefficient distributions, C(s) were used to estimate protein molar mass and oligomeric state of the colicin A pore forming domain. Sedimentation velocity (SV) experiments were carried out in a Beckman Coulter (Palo Alto, CA, USA) ProteomeLab XL-I analytical ultracentrifuge using both absorbance at 280 nm and interference optics. All AUC runs were carried out at a

rotation speed of 48,000 rpm and an experimental temperature 4^{o}C using an 8-hole AnTi50 rotor and double-sector aluminium-epon centerpieces. The sample volume was 400 µl for the SV experiments, and the sample concentrations ranged between 0.1 and 0.76 mg/ml. Sedimentation velocity profiles were treated using the size-distribution c(s) model implemented in the program SEDFIT . In order to determine the mass of each species, the c(s) distribution was converted to c(M) distribution. Each peak on the distribution plot was integrated in order to obtain the weight-averaged values for sedimentation coefficient and molecular mass. The integrated values of sedimentation coefficient (s) obtained at experimental conditions were converted to the standard conditions ($s_{20,w}$) (which is the value of sedimentation coefficient in water at 20°C).

Nuclear magnetic resonance

Nuclear magnetic resonance is an analytical tool that is widely applied in the chemical, physical and biological fields to study the conformation and dynamics of molecules. This technique is based on the magnetic properties of atomic nuclei containing an odd number of nucleons. When placed in a magnetic field such nuclei act like tiny spinning spherical bodies and tend to line up in the direction of the field. (Cooper, 2004) These nuclei of particular importance to molecular bioscience (^{1}H, ^{13}C, ^{15}N and ^{31}P) have spin values of 1/2. According to quantum mechanics, in strong magnetic field, spin 1/2 nuclei give rise to only two states corresponding to either parallel or anti-parallel. The magnetic moment is related to a proportionality constant, λ, known as gyromagnetic ratio:

$$\mu = 1/2\hbar\lambda \ (\hbar = h/2\pi)$$

So that the energy spacing between them is given by:

$$\Delta E = 2\mu_z B = \hbar\lambda B$$

Where B is the strength of the magnified field, and μ_z is the Z component of the nuclear magnetic moment along the field.

The resonance frequency is related to the energy space:

$$Hv = \Delta E = \hbar\lambda B = h\lambda B/2\pi$$

So that the resonance frequency is:

$$V = \lambda B/2\pi$$

The resonance frequency of a protein depends on two factors: the nature of the atomic nuclei with an odd number of nucleons and the magnetic field. The magnetic field is made up from both the external applied field and other nuclei and chemical groups. Because the tiny spinning nuclei act like a small magnets they will shift the local magnetic field, NMR resonance frequencies are extremely sensitive to the chemical environment of particular nuclei. (Cooper, 2004)

2D ^{15}N-^{1}H heteronuclear single quantum (HSQC) correlates the chemical shift of proton with the chemical shift of the directly bonded nitrogen (Decatur, 2011). Therefore, the signal comes from the protein backbone and the chemical shifts correspond to the environment of each amino acid of the sample.

Sample preparation: ^{15}N labelled Colicin A pore forming domain solution was concentrated to 8-12 mg/ml, and placed in the bottom of the NMR sample tube, the minimum sample volume should be more than 500 µl. Thereafter, 5-10% (v/v) of D_2O was added and mixed with the sample. The NMR sample tube was sealed with a plastic cap and labelled with the sample name, sample concentration, solvent buffer, and the date.

NMR procedure:

I used a Bruker 800 MHz NMR spectrometer running Bruker Topspin 2.1 in the laboratory of Prof G. Moore School of Chemistry, University of East Anglia, Norwich UK. Sample injection: Before putting the new sample into the instrument, the sample chamber was set to 298 K; after releasing the gas from the instrument, the sample was inserted and the temperature allowed equilibrating to 298 K. The original parameter was set by Dr Olli Hecht and data was collected and treated as previously described (Hecht et al., 2009).

Chapter Three

The stability of aspartate to alanine mutations of the Colicin A pore forming domain.

The previously investigations indicated that, after binding to the receptor protein in the outer membrane of the sensitive cells but before inserting into the inner membrane, the colicin A pore forming domain forms an insertion-competent state (van der Goot et al., 1991). This conformational change could be induced *in vitro* by low pH. (Ortega et al., 2001) *In vitro*, Colicin A C-terminal domain (pore forming domain) forms a molten globule state below pH 3.5. (Fridd and Lakey, 2002, Evans et al., 1996) It is proposed that the insertion –competent state of colicin A pore forming domain is equivalent to a "molten globule state". (Fridd and Lakey, 2002) And low pH is a trigger to the "molten globule state" formation. *In vivo*, colicin A requires acidic phospholipids in the target inner membrane to kill *E.coli* cells. On the inner membrane surface, bulk pH is not necessary and local pH can be reduced by acidic groups, thus acidic lipids could be sufficient to induce insertion-competent state formation. (Fridd and Lakey, 2002) This membrane-associated conformation change is associated with the protein insertion and pore formation activity. Hence, a study of the stability and unfolding pathways of ColA-P is crucial for understanding its toxic activity and may provide information on the pH sensitivity of proteins in general. Comparing the sequences of colicin A and colicin N C-terminal domain, there are seven extra surface aspartyl residues in ColA-P which, when mutated to alanine, cause the formation of an equilibrium molten globule at acidic pH; (Fridd and Lakey, 2002). It is proposed that removal of the negative charges by protonation at low pH or by alanine mutation would destabilize the pH sensitive colicin A pore forming domain. (Fridd and Lakey, 2002)

In the previous study, it was found that sequential replacement of the surface aspartate of colicin A by alanine destabilized the protein at neutral pH. (Fridd and Lakey, 2002) Most pH dependent protein models shows that buried residues or those residues in ion-pairs or hydrogen bonds with anomalous pKas would stabilize or destabilize the protein in different pH ranges. (Yang and Honig, 1993) However, the surface aspartyl residues of colicin A pore forming domain do not clearly fit to the mechanism. Hence, we believed that further characterization of the role of the surface aspartyl residues by mutagenesis and biophysics would reveal a different pH dependence mechanism of the protein.

3.1 Aims

To investigate the role of the surface aspartyl residues in protein folding and stability, this chapter of the thesis studies the conformational changes of colicin A pore forming domain wild type and its aspartate to alanine mutations *in vitro* induced by temperature and pH

The colicin A pore forming domain was heated and the structural information are obtained from differential-scanning calorimetry (DSC) and circular Dichroism (CD).

The pH dependent “molten globule state” was dialysed in a range of low pH solutions and the structural changes induced by urea denaturation were measured by fluorescence and CD spectroscopy.

3.2 Results

Studying the stability of Aspartic acid to Alanine mutants of ColA-P domain by fluorescence spectroscopy

The intrinsic fluorescence of colicin A P-domain arises from tryptophan residues that are located on helix 4, helix 7 and in the loop between helix 6 and helix 7 and is thus highly sensitive to the polarity of local environment. It was used to study the urea induced unfolding of the protein and its mutants. The aspartyl residues on the surface of the protein which, when mutated to alanine, destabilize the colicin A pore forming domain are shown in Figure 3.1.

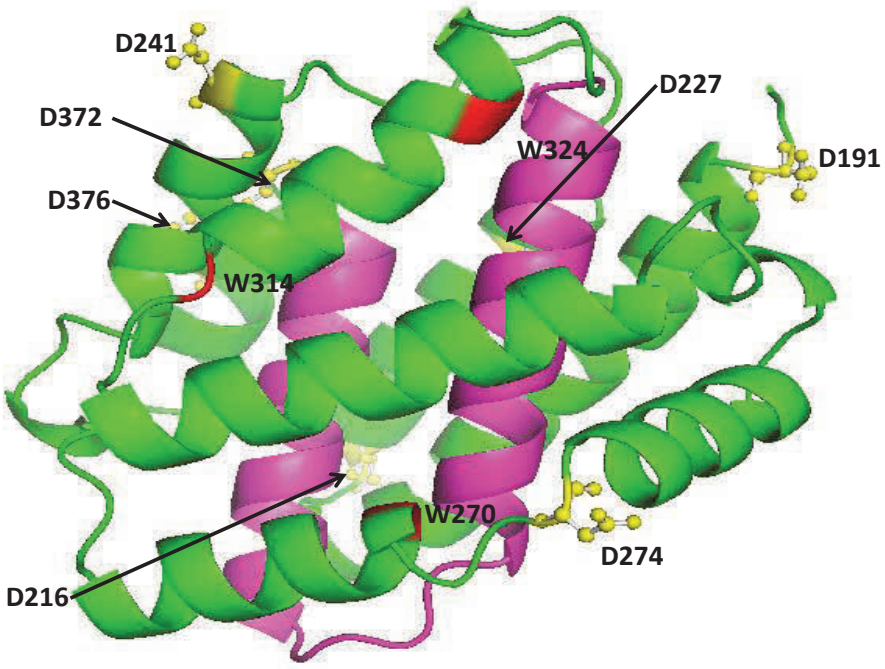


Figure 3.1: The crystal structure of ColA-P domain (PyMOL) highlighted with the critical surface aspartyl residues (Yellow sticks) and tryptophan residues (Red ribbon).

The tryptophan fluorescence of the Colicin A Pore forming domain in solution was measured between 295 nm and 440 nm, at an excitation wavelength of 280 nm. The tryptophan emission spectra of colicin A pore forming domain were analysed by calculating the barycentric mean wavelength. In fluorescence, the barycentric mean is the middle of an integrated emission curve and is very sensitive to small changes in emission wavelength that depends upon the local solvent environment around the intrinsic tryptophan residues (Cooper, 2004). Hence, the barycentric mean of the intrinsic protein fluorescence was used as a sensitive probe to assess the stability of the protein.

Powerful protein denaturants such as urea and guanidine hydrochloride were used to gradually unfold colicin A-pore forming domain by measuring samples at equilibrium in increasing concentrations.

To determine the stability of ColA-P domain and its alanine mutants, the intrinsic fluorescence spectra of ColA-P domain wild type, D216A, D227A, D274A and D372A were measured in a series of concentrations of urea at pH 7.0 in MOPS buffer. The barycentric mean of each emission spectrum was calculated and plotted to yield a urea unfolding curve. (Figure 3.2).

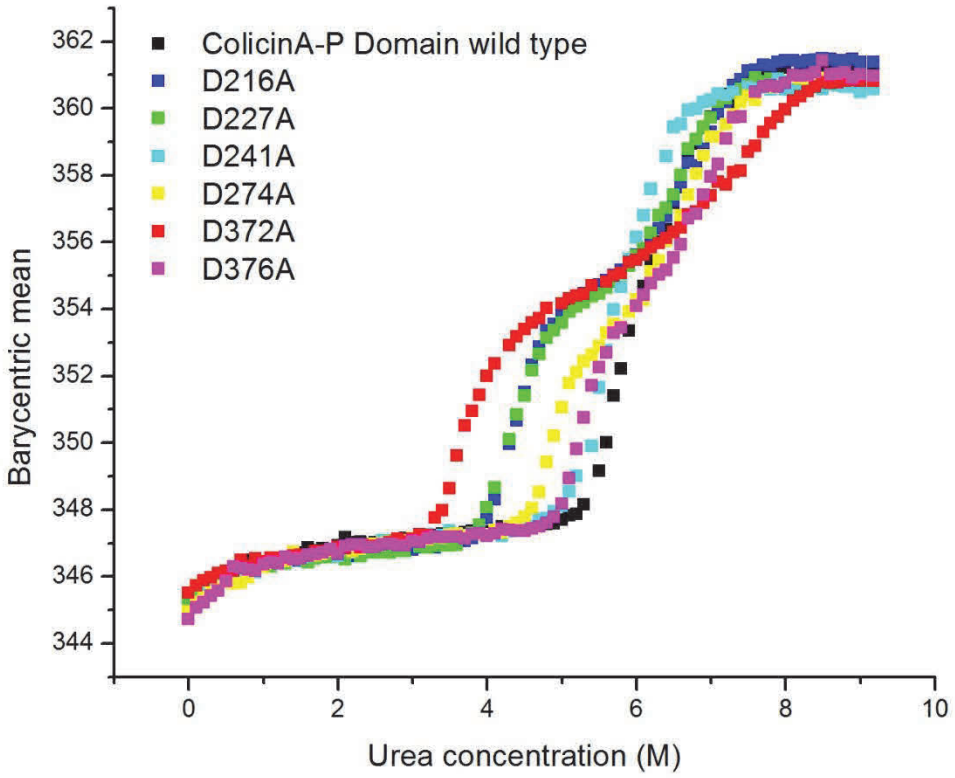


Figure 3.2: Urea unfolding transitions of the ColA-P wild type (Black curve), D216A (Blue curve), D227A (Green curve), D241A(Cyan), D274A (Yellow curve) , D372A (Red curve) and D376A (Magenta) at pH 7.0 in MOPS buffer.

Figure 3.2 shows the unfolding process of Colicin A wild type and aspartate to alanine mutant P-domain proteins triggered by urea. Two significant results are shown in this figure. The first one is that ColA-P wild type starts unfolding at 5.5M urea concentration, while the alanine mutations start to transform from their folded states to molten globule state in the range of 3.2-4.7M concentration of urea. These results demonstrate that the alanine mutations are less stable than its wild type protein. The second result is entirely unexpected: the ColA-P domain which is a single compact domain, was unfolded by urea in two stages. However, the urea unfolding curve of the wild type protein did not distinctly reveal its biphasic transition since the transitions

were close together. Thus guanidine HCl was used to separate the two transitions. ColA-P D372A and D227A, the most unstable and a medium stability alanine mutant respectively were chosen to compare with the guanidine HCl biphasic unfolding transition of ColA-P wild type. The intrinsic fluorescence of WT-ColA-P domain, D227A and D372A in a series of concentrations of guanidine HCl at pH 7.0 in MOPS buffer were measured and the barycentric mean of each emission spectrum plotted to yield a guanidine HCl induced unfolding curve. (Figure 3.3)

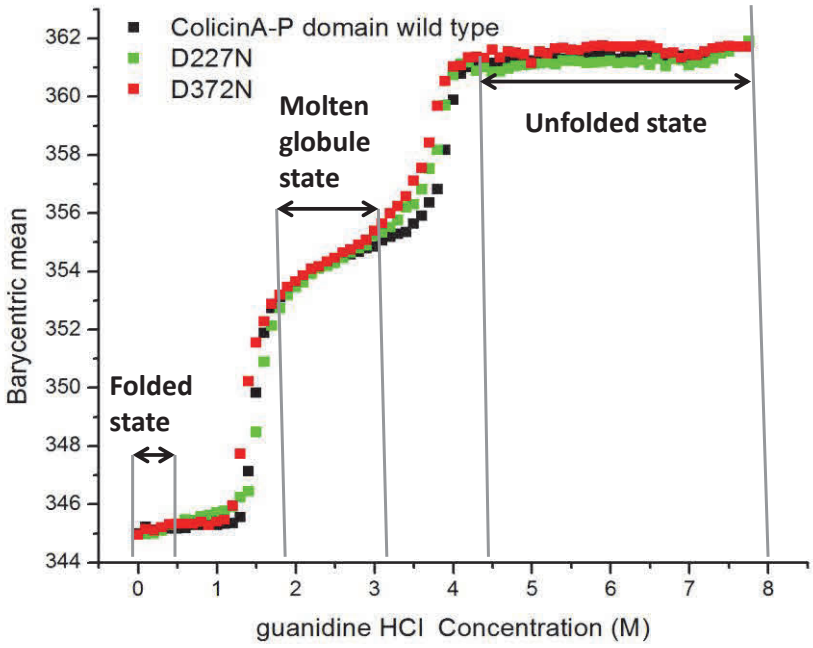


Figure 3.3: Biphasic guanidine hydrochloride unfolding transitions of the ColA-P wild type (Black curve), D227A (Red curve), D372A (Green curve) at pH7.0.

Figure 3.3 shows the biphasic unfolding process of Colicin A wild type and aspartate to alanine mutant P-domain proteins triggered by guanidine HCl. Two separated transitions of colicin A wild type and its alanine mutants can be observed in this Figure. Thus, three distinct states of ColA-P wild type and its mutants could well

defined - folded states, molten globule states and unfolded states. A red shift of the barycentric mean of the alanine mutants in fluorescence indicates that the tryptophan residues, which are located in the core of the protein, are more exposed to the solvent. Compared with the wild type protein, when the protein transformed from its native state to the molten globule state at a defined guanidine HCl concentration, the barycentric mean of the alanine mutants red shifts to a longer wavelength, demonstrating that the alanine mutants are easier to unfold to the molten globule state than the wild type protein. In other words, the ColA-P domain alanine mutants are less stable than the wild type. However, when unfolding from the molten globule state to the unfolded state, Col A-P D372A is similar to the wild type protein, showing that the single mutations don't always affect the stability of the second transition. These results indicate that the common effect of the surface mutations is only to enhance the possibility of the molten globule state (and therefore the insertion-competent state in vivo) formation. In conclusion: Figure 3.2 shows that the stability of the first unfolding transition of ColA-P aspartate to alanine mutants is: WT > D241A > D376A > D274A > D216A > D227A > D372A at pH 7.0 in urea solution. Figure 3.3 shows that the stability of the first unfolding phase of ColA-P aspartate to alanine mutants is: WT > D227A > D372A at pH 7.0 in guanidine solution. The second transition is not affected by all the mutations.

The Gibbs free energy is a way to define the thermodynamic properties of a system. The change in Gibbs free energy (ΔG) is used to define the energy absorbed or released from a system, for example during a reaction. For folded proteins, the ΔG between the unfolded and folded states should be negative and its magnitude can be used as a measure of their relative stability. Thus, the changes in Gibbs free energy (ΔG) of the ColA-P protein biphasic unfolding process were used to measure the relative stability of the folded state of ColA-P compared to its molten globule state and the stability of the molten globule state compared to its unfolded state. Thus we derive two (ΔG) values, ΔG_1 is the change in Gibbs free energy of the first transition

and ΔG_2 is the change in Gibbs free energy of the second transition. Each ΔG was calculated using a non-linear least squares fitting procedure in Microcal Origin that was kindly supplied by Dr David Hough, Bath University, UK. The Gibbs free energy (Wolstenholme and O'Connor, 2008) of ColA-P urea unfolding process were calculated and shown in table 3.1. The actual values obtained do need to be treated with caution since the fitting routine assumes a two state model which is not the case here.

$$N \leftrightarrow U$$

Which gives a simple equilibrium constant $K_{eq}=\frac{[U]_{eq}}{[N]_{eq}}$

This leads directly to a calculation of the free energy difference between native state (N) and unfolded state (U) for the system.

$$\Delta G_{N\rightarrow U} = -RTInK_{eq}$$

To define the free energy of unfolding we need to know the value in the absence of denaturant and this is obtained by extrapolation since the K_{eq} here is very small and immeasurable. The extrapolation can either be of a linear plot of ΔG verses denaturant concentration or the method used here which uses a non-linear curve fitting procedure. In either case the assumption of a two state equilibrium is important and is not easily proven here. Thus the values are only approximations which allow us to compare the mutants.

Table 3.1. The Gibbs free energy of biphasic unfolding for ColA-P WT and its alanine mutants detected by urea unfolding at pH 7.0

Mutation	$\Delta G1$ kCal.mol^{-1}	$\Delta G2$ kCal.mol^{-1}	$\Delta G1+\Delta G2$ kCal.mol^{-1}	First transition (Urea concentration M)	Second transition (Urea concentration M)
D216A	14.70	12.70	23.73	3.8-5.0	6.1-7.5
D227A	14.26	13.80	28.06	3.7-5.1	6.0-7.3
D241A	8.40	11.12	19.52	4.9-5.7	5.7-7.0
D274A	8.05	10.71	18.76	4.4-5.1	6.1-7.5
D372A	13.36	11.61	24.96	3.2-4.3	5.9-8.5
D376A	25.28	13.41	38.26	4.7-5.7	6.5-7.7
WT	21.83	10.45	32.28	5.2-6.0	6.1-7.5

The change of Gibbs free energy of the first transition ($\Delta G1$) represents the energy that is needed for the protein transformation from folded state to the molten globule state. $\Delta G1$ of D241A and D274A is distinctly lower than other mutants, which means D274A and D241A needs less energy to unfold to the intermediate stage.

The Gibbs free energies of ColA-P WT and mutants in guanidine HCl unfolding were calculated and are shown in table 3.2.

Table 3.2. The Gibbs free energy of ColA-P guanidine HCl biphasic unfolding process at pH 7.0

Mutation	$\Delta G1$ Kcal.mol^{-1}	$\Delta G2$ Kcal.mol^{-1}	$\Delta G1+\Delta G2$ Kcal.mol^{-1}	First transition (Gdn HCl concentration M)	Second transition (Gdn HCl concentration M)
D227A	8.85	22.47	31.32	0.8-1.4	3.1-3.9
D372A	6.55	17.32	23.87	0.6-1.2	3.5-4.3
WT	13.16	32.03	45.19	1.3-1.8	3.6-4.2

Guanidine HCl unfolding experiment separates the unfolding transition better than urea. The Ala mutations of ColA-P start the first unfolding transition at lower urea or guanidine HCl concentration than its wild type. The $\Delta G1$ of D372A $>\Delta G1$ of D227A $>\Delta G1$ of the wild type, which means D372A needs less energy to unfold to the

intermediate stage than D227A, and D227A needs less energy to unfold to the intermediate stage than the wild type protein.

Studying the pH dependent stability of Aspartate acid to Alanine mutants of ColA-P domain by fluorescence spectroscopy

As I mentioned before, ColA-P protein forms a molten globule state below pH 3.5 in vitro, and an acidic environment would trigger the formation of an insertion competent state in vivo. To investigate the relationship between the structure change and the pore forming activity of ColA-P protein, it would the most interesting to study the stability of the protein in an acidic environment. In order to determine the stability of ColA-P WT in acidic environments, the intrinsic fluorescence of ColA-P domain wild type in a series of concentrations of urea at pH 2.0, pH 2.5, pH 3.0, pH 3.5, pH 4.0 and pH 4.5 in sodium citrate buffer were measured. The barycentric mean of each emission spectrum was calculated and plotted as a urea unfolding curve. (Figure 3.4)

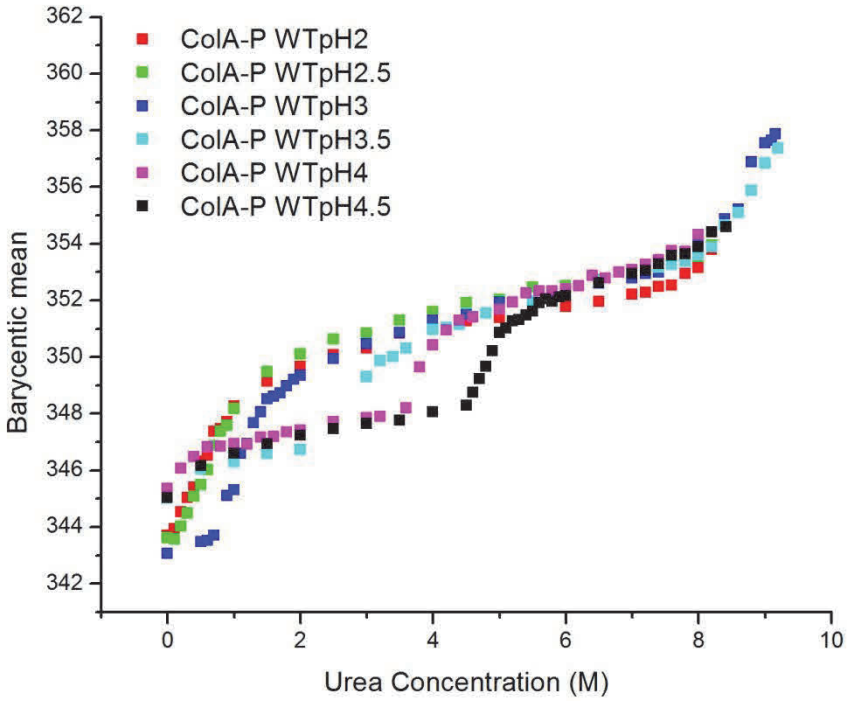


Figure 3.4: The urea unfolding curve of the wild type colicin A P-domain at pH 2.0 (Red curve), pH 2.5 (Green curve), pH 3.0 (Blue curve), pH 3.5 (Cyan Curve), pH 4.0 (Magenta Curve) and pH 4.5 (Black Curve) in sodium citrate buffer.

Figure 3.2 shows that alanine mutations at D216, D227 and D372 are less stable than the wild type protein and other mutations. Thus we will concentrate on the stability of these three mutants in this section. To compare the stability of ColA-P WT with D216A, D227A and D372A in low pH, the intrinsic fluorescence of ColA-P mutant domains in a series of concentrations of urea at pH 2.0, pH 2.5, pH 3.0, pH 3.5, pH 4.0 and pH 4.5 in sodium citrate buffer were measured. The barycentric mean of each emission spectrum was calculated and plotted to yield a urea unfolding curve. (Figures 3.5, 3.6 and 3.7)

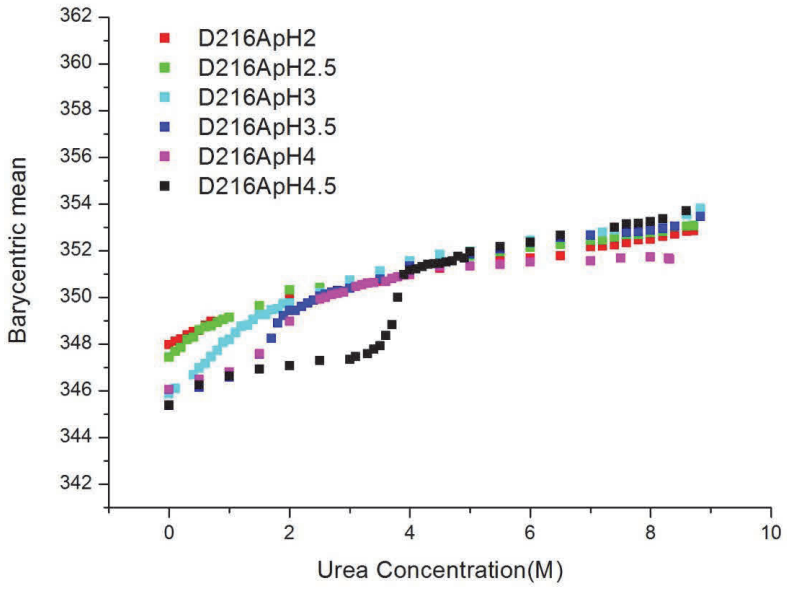


Figure 3.5: The urea unfolding curve of colicin A P-domain mutant D216A at pH 2.0 (Red curve), pH 2.5 (Green curve), pH 3.0 (Cyan curve), pH 3.5 (Blue Curve), pH 4.0 (Magenta Curve) and pH 4.5 (Black Curve) in sodium citrate buffer

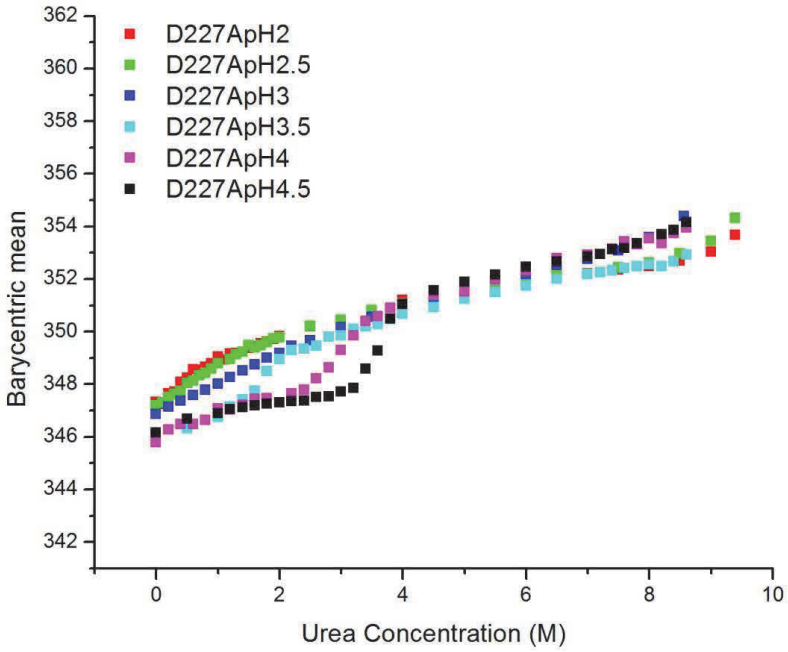


Figure 3.6: The urea unfolding curve of colicin A P-domain mutant D227A at pH 2.0 (Red curve), pH 2.5 (Green curve), pH 3.0 (Blue curve), pH 3.5 (Cyan Curve), pH 4.0 (Magenta Curve) and pH 4.5 (Black Curve) in sodium citrate buffer.

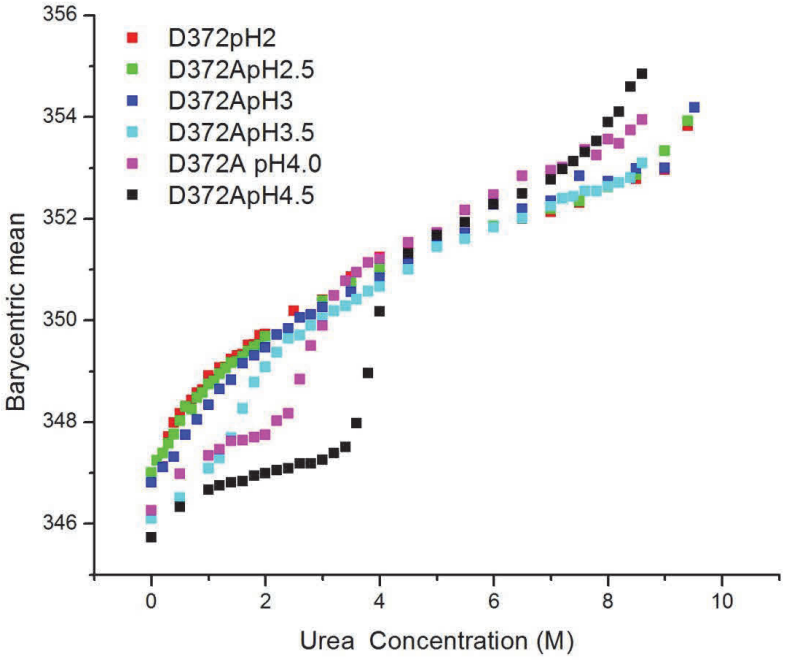


Figure 3.7: The urea unfolding curve of colicin A P-domain mutant D372A at pH 2.0 (Red curve), pH 2.5 (Green curve), pH 3.0 (Blue curve), pH 3.5 (Cyan Curve), pH 4.0 (Magenta Curve) and pH 4.5 (Black Curve) in sodium citrate buffer.

Figure 3.4 shows that obvious transitions appear on ColA-P WT's urea unfolding process from pH 3.0 to pH 4.5 in low urea concentration, and these transitions started from 0 M urea concentration at pH 2.0 and pH2.5. Figure 3.5 shows that obvious transitions only appear on ColA-P D216A's unfolding process between pH 3.5 and pH 4.5 in low urea concentration, the transition started from 0 M urea concentration at pH 3.0, and these transitions are abolished between pH 2.5 and pH 2.0. Figure 3.6 and Figure 3.7 show these transitions only appear on ColA-P D227A and D372A's unfolding process between pH 3.5 and pH 4.5 and they were abolished below pH 3.0. This means alanine mutations destabilized the proteins at low pH in various degrees, ColA-P D216A is slightly more stable than D227A and D372A, and it is less stable than ColA-P WT. The relationship between the unfolding transitions and stability of the protein were summarised and are shown in table 3.3.

Table 3.3 The unfolding transitions and the stability of ColA-P protein at pH 2.0, pH 2.5, pH 3.0, pH 3.5, pH 4.0 and pH 4.5.

Mutations	Transition at pH 2.0 (Urea concentration M)	Transition at H 2.5 (Urea concentration M)	Transition at pH 3.0 (Urea concentration M)	Transition at pH 3.5 (Urea concentration M)	Transition at pH 4.0 (Urea concentration M)	Transition at pH 4.5 (Urea concentration M)
D216A	abolished	abolished	0-1.6	1.0-2.2	1.5-2.6	3.3-4.0
D227A	abolished	abolished	abolished	1.0-2.2	2.4-3.6	3.2-4.0
D372A	abolished	abolished	abolished	1.2-2.4	2.0-3.2	3.4-4.5
WT	0-1.5	0-1.5	0.6-1.5	2.5-4.0	3.6-4.4	4.5-5.3

Table 3.3 shows that ColA-P D216A, D227A and D372A are less stable than the wild type protein in acidic environment. This result confirms that removal of the negative charges by protonation at low pH or by alanine mutation might destabilize the pH sensitive colicin A pore-forming domain in the same way.

Circular Dichroism studies of Colicin A pore-forming domain and its Aspartate to Alanine mutants

The protein CD spectrum in the Far UV is reported on a per residue basis and used here to detect helical content. The Near UV CD is generally reported using protein molar concentration and used to measure protein tertiary structure. (Luria and Suit, 1987)

To determine the secondary structure of ColA-P wild-type and its alanine mutants at different pH, the far UV CD spectra of Colicin A wild-type and D216A, D227A and D372A mutant P-domain proteins were measured and shown in Figure 3.8.

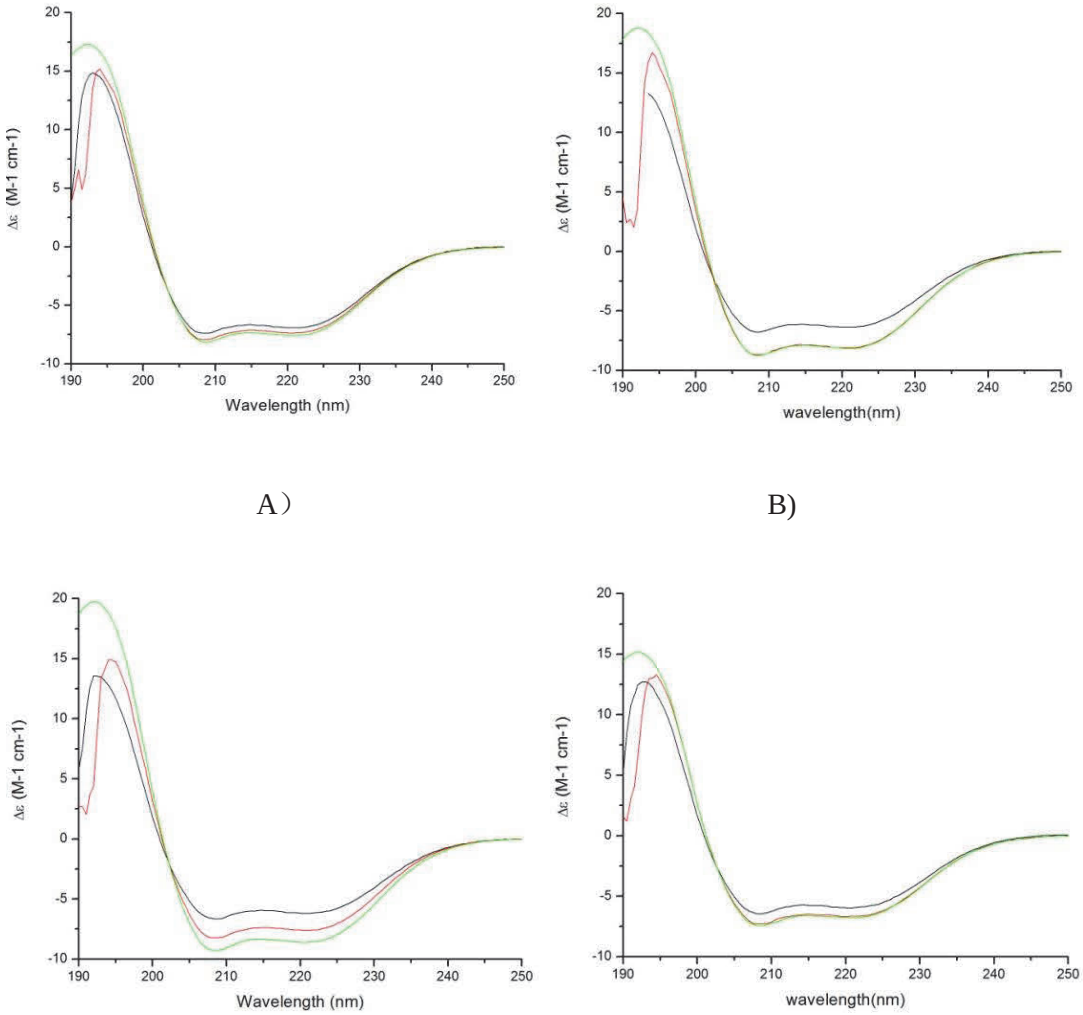


Figure 3.8 : A) Far UV CD spectra of the ColA-P WT at pH 3.0 (Black line), pH 5.0 (Red line), and pH 8.0 (Green line) B) Far UV CD spectra of ColA-P D216A at pH 3.0 (Black line), pH 5.0 (Red line), and pH 8.0 (Green line) C) Far UV CD spectra of ColA-P D227A at pH 3.0 (Black line), pH 5.0 (Red line), and pH 8.0 (Green line) D) Far UV CD spectra of ColA-P D372A at pH 3.0 (Black line), pH 5.0 (Red line), and pH 8.0 (Green line)

The Far UV CD spectra of ColA-P wild type and its mutant proteins at neutral pH show a signal indicating broadly α–helical structure. ColA-P D372A shows a slight

loss of secondary structure at pH 8.0. Each of the proteins show a slight loss of secondary structure at pH 5.0 and pH 3.0.

To determine the tertiary structure signals of ColA-P wild-type and its alanine mutants at different pH, the near UV CD spectra of Colicin A wild-type and D216A, D227A and D372A mutant P-domain proteins are shown in Figure 3.9.

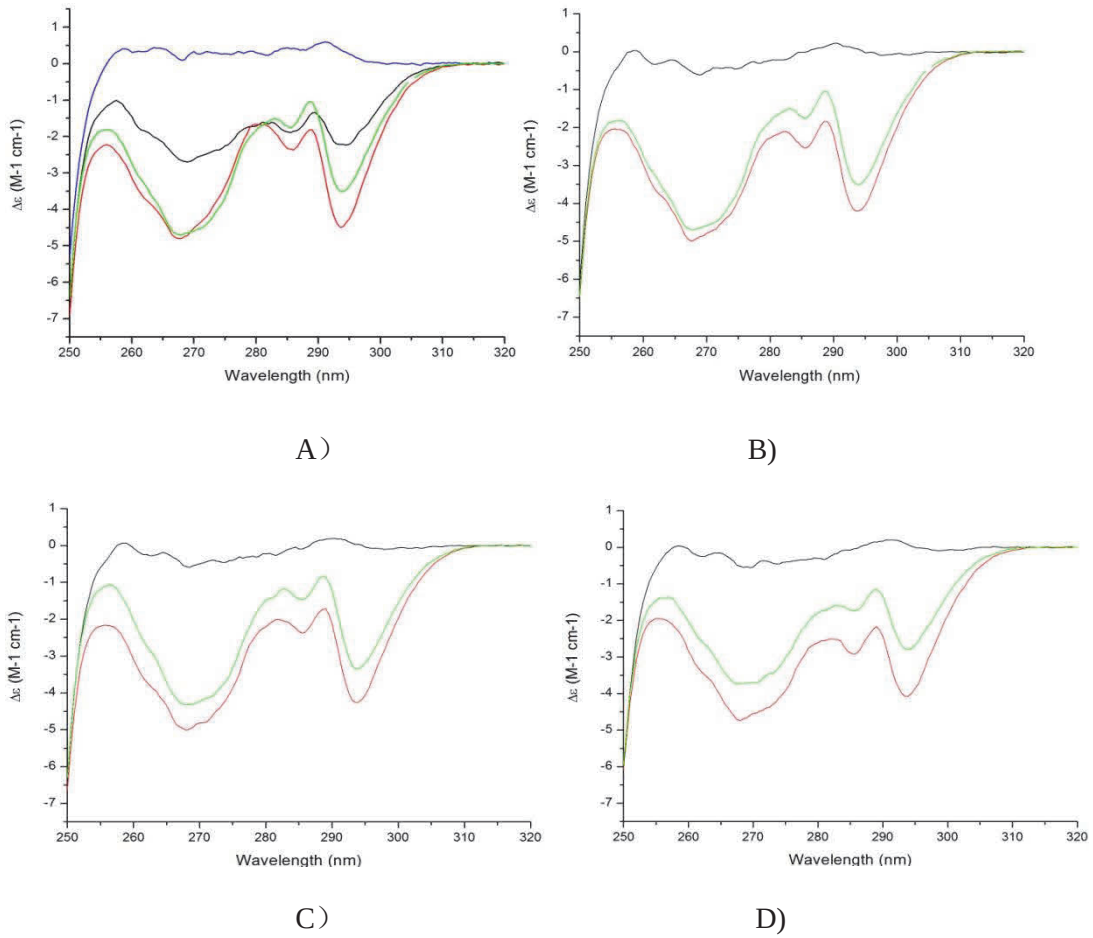


Figure 3.9: A) Near UV CD spectra of the ColA-P WT at pH 2.0 (Blue line), pH 3.0 (Black line), pH 5.0 (Red line), and pH 8.0 (Green line); B) Near UV CD spectra of ColA-P D216A at pH 3.0 (Black line), pH 5.0 (Red line), and pH 8.0 (Green line); C) Near UV CD spectra of ColA-P D227A at pH 3.0 (Black line), pH 5.0 (Red line), and pH 8.0 (Green line); D) Near UV CD spectra of ColA-P D372A at pH 3.0 (Black line), pH 5.0 (Red line), and pH 8.0 (Green line)

The near UV CD spectra of the mutant proteins show similar tertiary structure at pH 8.0 when compared with the wild-type, which is fully folded. At pH 5.0, both the wild type and the mutant proteins show differing degrees of loss of tertiary structure. At pH 3.0, the mutant proteins show complete loss of tertiary structure. The wild-type protein shows loss of tertiary structure at pH 2.0.

The stability of Ala mutants without urea and in 2 M urea and 6 M urea at pH 3.0, pH 5.0 and pH 8.0 studied by CD spectroscopy

A π-π^* transition of the α-helix has a negative CD signal at 208 nm. (Kelly et al., 2005) Colicin A-P is over 60% helix therefore the intensity of Far UV circular dichroism signals at 208 nm can be used to represent the secondary structure of ColA-P wild type and its mutants. Urea was used to compare the relative stabilities of the secondary structure of the mutants at each pH.

Colicin A pore-forming domain alanine mutants were initially dialyzed in 50 mM phosphate buffer with no urea, 2 M urea and 8 M urea at pH 8.0, pH 5.0 and pH 3.0 for more than 24 hours. The Far UV CD spectra and the intensity of Far UV CD signals at 208 nm of ColA-P proteins were measured and are shown in Figure 3.13.

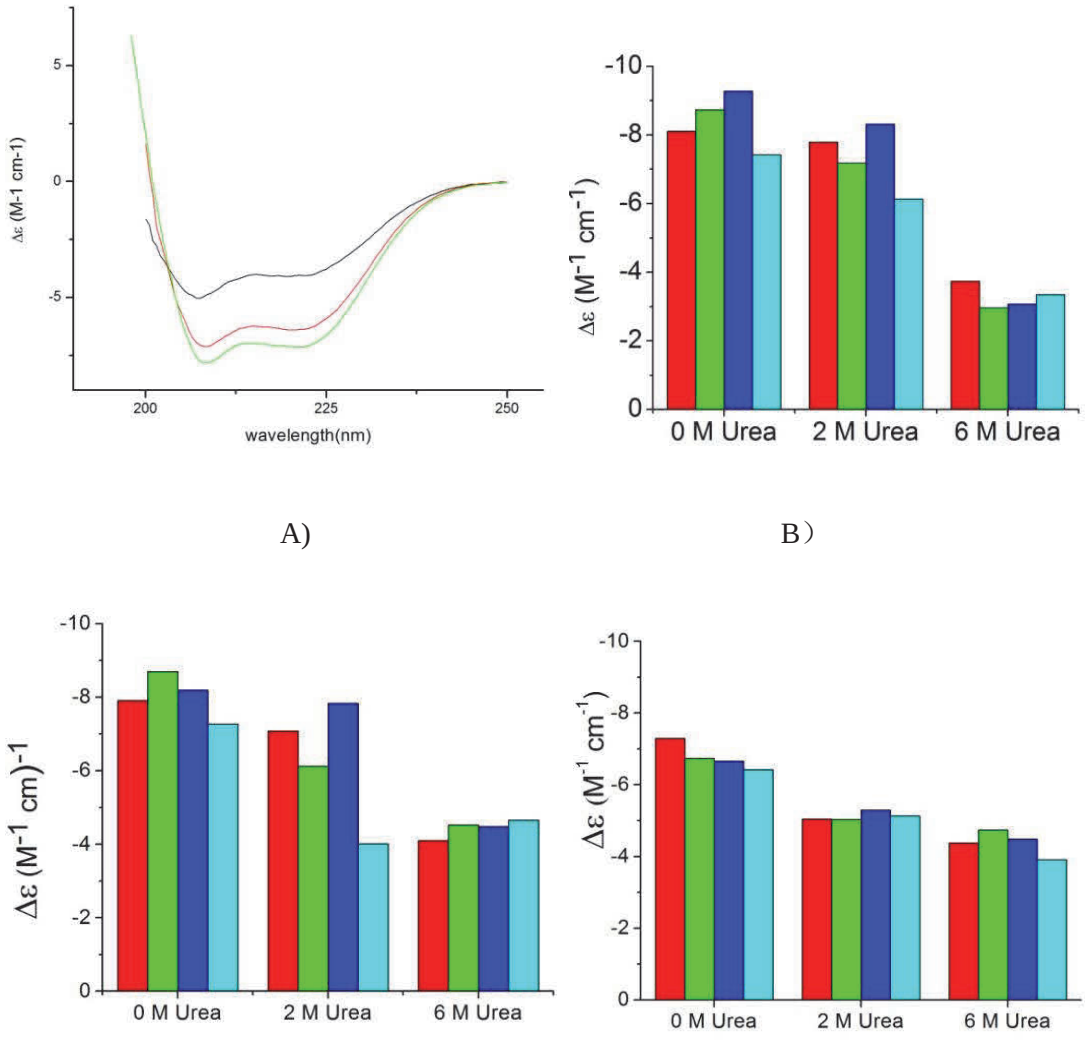


Figure 3.10: A) Far UV CD spectra of the ColA-P WT at pH 3.0 (Black line), ColA-P WT at pH 5.0 (Red line), ColA-P WT at pH 8.0 (Green line) in 2 M urea solution. B) The CD signal intensities of ColA-P WT (Red column), D216A (Green column), D227A (Blue column) and D372A (Cyan column) at 208 nm at pH 8.0. C) The CD signal intensities of ColA-P WT (Red column), D216A (Green column), D227A (Blue column) and D372A (Cyan column) at 208 nm pH 5.0. D) The CD signal intensities of ColA-P WT (Red column), D216A (Green column), D227A (Blue column) and D372A (Cyan column) at 208 nm at pH 3.0.

Figure 3.10 shows that the Far UV CD signal of ColA-P wild type is greater than the mutants at pH 3.0 in the absence of urea.(Figure 4.3 lower right) It reveals that ColA-P domain wild type retains its secondary structure in acid environments and its mutants are not as acid resistant as the wild type. The intensity of CD of ColA-P proteins at 208nm wavelength are most different at pH 5.0 in the intermediate concentration of urea (2M). Some mutants show higher intensity than ColA-P wild type. This confirms that ColA-P domain wild type retains its secondary structure in acidic environments and its mutants do not differ greatly from the wild type. The Far UV CD signals at 208 nm of D227A show the biggest difference between pH 3.0 and pH 8.0 and the mutants tend to be less stable at acidic pH.

The near UV CD spectrum around 290 nm arises largely from tryptophans and it gives a valuable fingerprint of the tertiary structure of the protein (Kelly et al., 2005). Thus the intensity of near UV CD signals at 295 nm was used to detect tertiary structure of ColA-P wild type and its mutants.

The Far UV CD spectra and the intensity of near UV CD signals at 295 nm of ColA-P proteins were measured and shown in Figure 3.11.

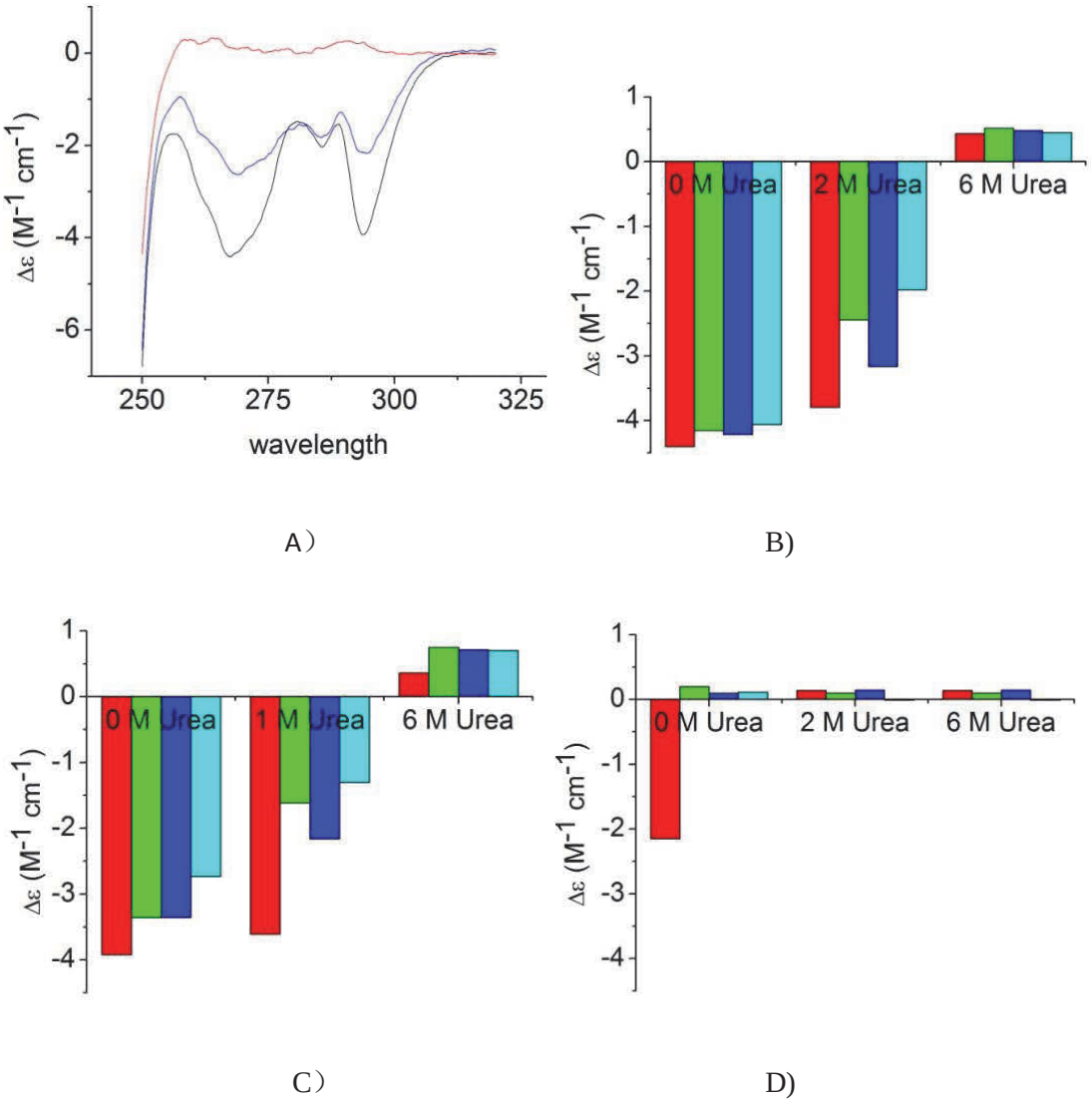


Figure 3.11: A) Near UV CD spectra of the ColA-P WT at pH 8.0 (Black line), ColA-P WT at pH 3.0 (Blue line) and ColA-P at pH 3.0 in 6M urea solution (Red line). B) The CD signal intensities of ColA-P WT (Red column), D216A (Green column), D227A (Blue column) and D372A (Cyan column) at 295 nm at pH 8.0. C) The CD signal intensities of ColA-P WT (Red column), D216A (Green column), D227A (Blue column) and D372A (Cyan column) at 295 nm pH 5.0. D) The CD signal intensities of ColA-P WT (Red column), D216A (Green column), D227A (Blue column) and D372A (Cyan column) at 295 nm at pH 3.0.

The Figure 3.11 shows that all of the mutants lose their near UV signal at pH 3.0. It is obvious that only the wild type of ColA-P retains most of its tertiary structure at pH 3.0 and this might be essential for the existence of the intermediate stage. The wild type loses tertiary structure at 2 M urea, thus the mutants show a significant effect near the point of formation of the molten globule.

Thermal transition temperature and the calorimetric transition enthalpy study by Differential Scanning Calorimetry

Since this single domain protein undergoes two unfolding transitions it would be interesting to investigate the relevant thermodynamic and structural information by differential-scanning calorimetry (DSC).

The melting temperature, calorimetric enthalpy and Van't Hoff enthalpy of colicin A-P domain were detected and analysed by DSC. The protein samples were studied in a temperature region of 25 to 90 ^{o}C at a scanning rate of 1 ^{o}C/min. The proteins were dissolved in 50 mM phosphate buffer to a concentration of 0.5 mg/ml at pH 8.0, pH 7.0 and pH 6.0; in 50 mM sodium citrate buffer to a concentration of 0.5 mg/ml at pH 5.0 and pH 3.0. Characteristic DSC scans for ColA-P at different pH are shown in Figure 3.12.

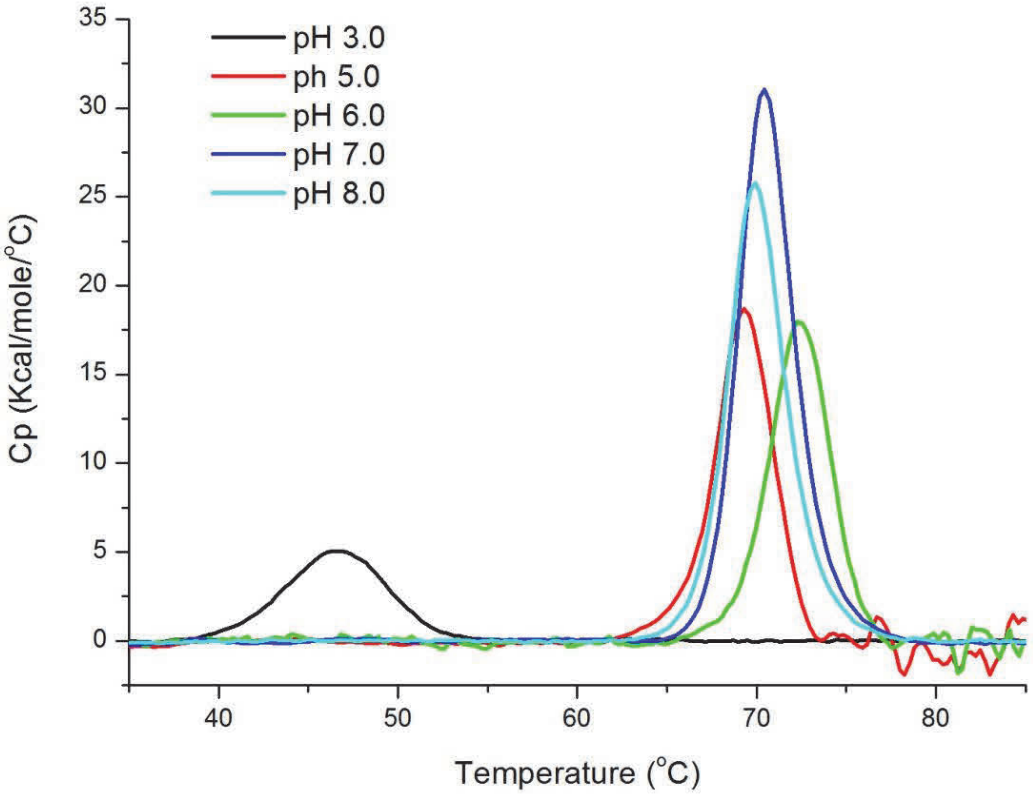


Figure 3.12: Characteristic heat-capacity curves for ColA-P WT at pH 8.0 (Cyan line), at pH 7.0 (Blue line), at pH 6.0 (Green line), at pH 5.0 (Red line), at pH 3.0 (Black line) with buffer baselines subtracted and normalized for protein concentrations.

Table 3.4. Calorimetric data of ColA-P WT in aqueous solution at 1℃ /min

Mutation	Thermal transition temperature Tm(℃) ±error	ΔH (kcal/mol) ±error	ΔHv (kcal/mol) ±error
WT pH3.0	46.57±0.02	36.90±0.08	140.0±0.05
WT pH5.0	69.26±0.03	75.54±0.06	228.7±0.05
WT pH6.0	72.38±0.04	77.19±0.35	222.8±0.93
WT pH7.0	70.55±0.01	124.3±0.10	231.1±0.40
WT pH8.0	70.01±0.01	107.8±0.08	220.7±0.02

The T_m of ColA-P WT at pH 6.0> T_m at pH 7.0 > T_m at pH 8.0 > T_m at pH 5.0 > T_m at pH 3.0 indicates that the thermal stability of the protein did not correlate to its pH. Figure 3.10 shows that the protein is more stable around pH 6.0 than pH 8.0 and pH 3.0 in solution. The calorimetric enthalpy (ΔH) of ColA-P WT at pH 7.0> T_m at pH 8.0 > T_m at pH 6.0 > T_m at pH 5.0 > T_m at pH 3.0 indicates that the energy required in

the thermal denaturation process of most of ColA-P proteins correlated to its pH, which means the protein need more energy to transform from its folded stage to its unfolded form at neutral pH than in low pH. The figures in the table 3.4 display that $\Delta H_V/\ \Delta H$=2.8 at pH 6.0 and reaches 3.8 at pH 3.0. Down to pH 5.0 this is due solely to a large drop in ΔH.

Thermal denaturation of ColA-P WT, D216A, D227A, D241A and D372A mutant P-domain proteins in solution were also studied by DSC. The protein samples were studied in a temperature range of 25 to 90 ^{o}C at a scanning rate of 1 ^{o}C/min. The proteins were dissolved in 50 mM phosphate buffer to a concentration of 0.5 mg/ml at pH 7.0. Characteristic DSC scans for colicin A wild-type and mutant P-domain proteins are shown in Figure 3.13.

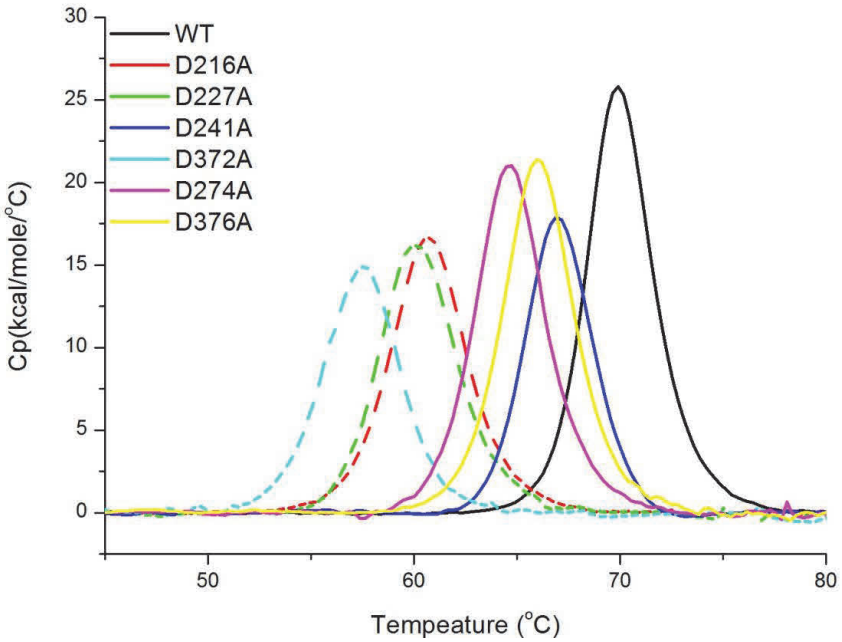


Figure 3.13 Characteristic heat-capacity curves for ColA-P WT (Black line), D216A (Red dash line), D227A (Green dash line), D241A (Blue line), D372A (Cyan dash line), D274A (Magenta line) and D376A (Yellow line) with buffer baselines subtracted and normalized for protein concentrations

Each of the ColA-P proteins showed only one DSC peak denaturation temperature within the temperature region of 25 to 90 ^{o}C.

ColA-P is a small single-domain protein, thus its temperature unfolding was expected firstly to be a two-state transition and by fitting the data the calorimetric transition enthalpy can be calculated. The denaturation temperature, calorimetric enthalpy and Van't Hoff enthalpy are shown in Table 3.3

Table 3.5. Thermodynamic Characterization of Colicin A wild type and Alanine mutant P-domain proteins

Mutation	Thermal transition temperature *Tm*(℃) ±error	*ΔH* (kcal/mol) ±error	*ΔHv* (kcal/mol) ±error
D216A	60.76±0.018	80.77±1.637	178.55±1.76
D227A	60.21±0.010	79.3±0.444	181±1.07
D241A	67.01±0.015	75.14±0.605	218.75±2.28
D274A	64.64±0.018	92.49±0.826	196.55±2.22
D376A	65.96±0.015	94.25±0.67	199.3±1.77
D372A	57.46±0.012	71.82±0.377	179.8±1.18
WT	70.01±0.0087	107.8±0.491	220.7±0.655

The T_m of ColA-P WT > T_m of ColA-P D241A > T_m of ColA-P D376A > T_m of ColA-P D274A > T_m of ColA-P D216A > T_m of ColA-P D227A > T_m of ColA-P D372A indicating that the thermal stability of the protein was destabilized by alanine mutations sequentially from D241 to D372. The calorimetric enthalpy (*ΔH*) of ColA-P WT >*ΔH* of ColA-P D376A >*ΔH* of ColA-P D274A >*ΔH* of ColA-P D216A >*ΔH* of ColA-P D227A >*ΔH* of ColA-P D241A >*ΔH* of ColA-P D372A indicates that the energy required in the thermal denaturation process of most of the ColA-P proteins correlated to its thermal stability, which means the more stable protein needs more energy to transform from its folded stage to its unfolded stage. ColA-P D241A is an exception.

Calorimetric enthalpy (ΔH) represents the actual heat energy change of one mole of protein of the system during the transition. (Cooper, 2004) The Van't Hoff enthalpies reflect the heat change associated with the transformation of one mole of cooperative unit of the system.; this is obtained from a fit to the peak shape.

If the values of ΔH and ΔHv are the same then the transition can be regarded as a two-state transition. (Serdyuk et al., 2007) If the cooperative unit is smaller than the macromolecule, as would be the case for a dimer, the ΔH is half of the ΔHv. In contrast, in the case of a non-2-state processes, ΔH is larger than ΔHv (Serdyuk et al., 2007).

The figures in table 3.3 show that surprisingly for ColA-P $\Delta H_V/ \Delta H=2$, and there are three initial possible explanations.

First, all the proteins, wild type and mutants are dimers in solution and unfold cooperatively with prior dissociation and without intermediates. This would be the simplest explanation.

Secondly, the protein concentration has been consistently over-estimated but this is unlikely to be the case so consistently for all samples.

Thirdly, irreversible exothermic effects during the transition and sharpening the transition peaks. Again, it is unlikely to be so consistent for wild type and all mutants by the value of 2 for the ratio.

Another potential explanation, is that colicin A-P domain thermally unfolds via the intermediate shown in the urea experiments, and the two steps are widely separated, with only one of the steps unfolding in the given temperature range. Temperature unfolding followed by CD spectroscopy represents two obvious separated transitions which match the biphasic unfolding process of urea and guanidine induced unfolding.

To examine if the protein is a dimer in solution we used Native PAGE, Size exclusion chromatography and AUC.

Native polyacrylamide gel electrophoresis

Native Page was used to examine if ColA-P is a dimer in solution since SDS-PAGE will usually break non-covalent interactions. 45 µl of colicin A-P WT with either 5 µl of EDTA or 5µl of $NiCl_2$ were run on the native PAGE respectively. 50 µl ColA-P WT proteins and BSA were run on native PAGE at 240V for 45 minutes as a contrast. The image of the dried Native PAGE was shown in Figure 3.14.

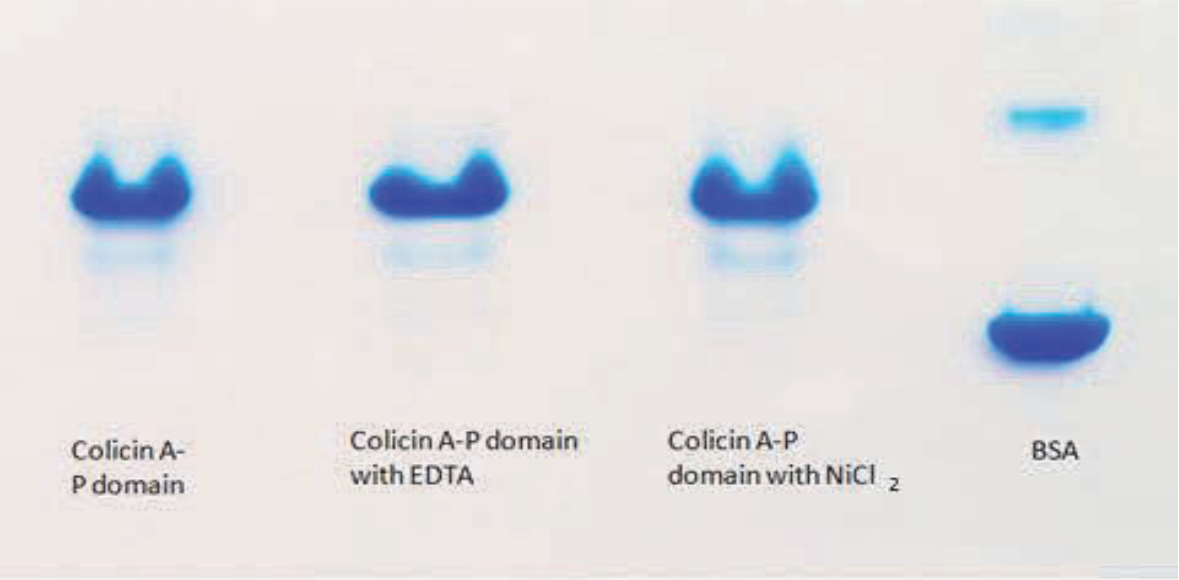


Figure 3.14 The Native gel of Colicin A-P domain

Ni^{2+} might cause dimers of His tagged proteins (Serdyuk et al., 2007), EDTA can remove the residual Ni^{+}, however the same migration of the samples on native gel showed that three samples had the same molecular mass and conformation. SDS-PAGE separates the molecules only by molecular mass; the native gel also separates the molecules by the cross-sectional area, thus the migration of the samples on native gel is also dependent on the shape of the overall structure. Therefore, native

gels cannot measure the molecular mass directly but this result was initial evidence that Colicin A-P domain is a single species in solution rather than a monomer- dimer equlibrium.

Size exclusion chromatography

Size exclusion chromatography was used to estimate the size of ColA-P in solution. The gel filtration column (Superose 12 10.300 GE Healthcare UK) was equilibrated with 50 mM sodium phosphate, 150mM NaCl at pH 7.4. A series of proteins; PA heptamer (Soliakov et al., 2010), BSA, Aprotinin, Carbonic Anhydrase and β-amylase, Colicin N, CaF1 (Chalton and Lakey, 2010) and V-antigen (Carr et al., 1999) were used as standards to achieve a linear plot of the $\log_{10}$ of their molecular mass and their K_d s. The distribution coefficient between stationary and mobile phases, and hence the molecular mass of an unknown protein can be worked out from its elution volume.

The calibration curves produced by plots of the logarithm of the molecular weights of standard proteins versus their K_d is shown in figure 3.15.

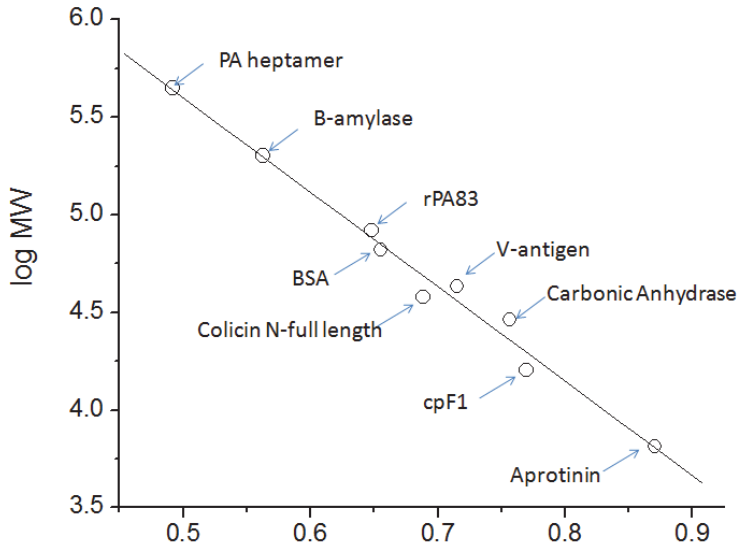


Figure 3.15: The calibration curves of the logarithm of the molecular weights

The calculated molecular weight of ColA-P D372A is 21182 Da

This corresponds to the expected molecular weight of ColA-P D372A which is 22920 Da, and provides further evidence that ColA-P is monomer in solution.

Analytical Ultracentrifugation

To complete the study of the oligomeric state of Colicin A P-domain a series of sedimentation distributions were obtained from a series of sedimentation velocity experiments with increasing concentration of Colicin A wild type P-domain protein (0.1 mg/ml, 0.134 mg/ml, 0.23 mg/ml, 0.327 mg/ml, 0.455 mg/ml, 0.678 mg/ml and 0.766 mg/ml). Before the start of the experiment, the protein was equilibrated at 4 ^{o}C for 2 hours. It was then run in high vacuum at 48,000 rpm in a Beckman Optima Ultracentrifuge, the sedimentation coefficient was calculated according to the rate of movement of the boundary per unit of centrifugal field.

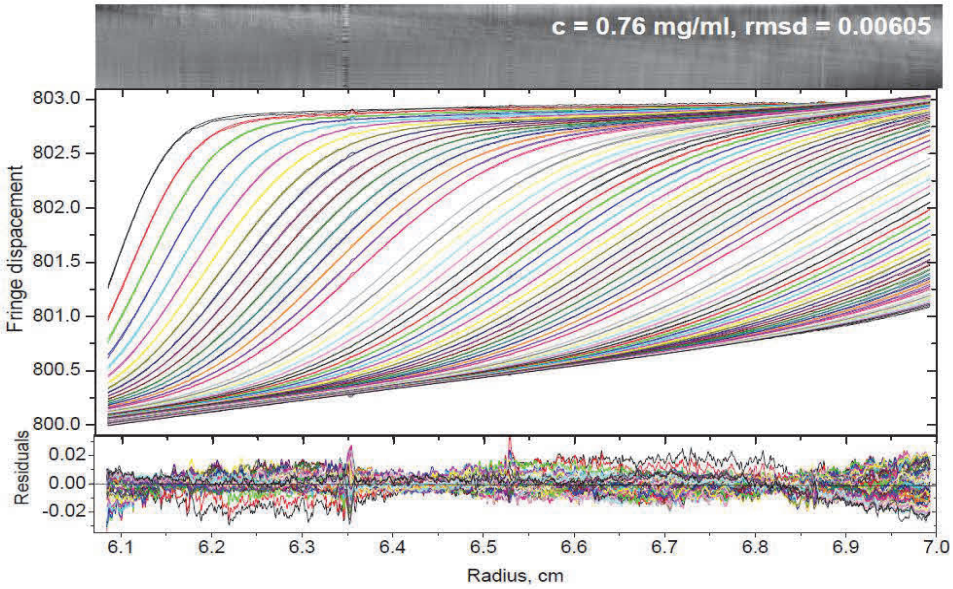
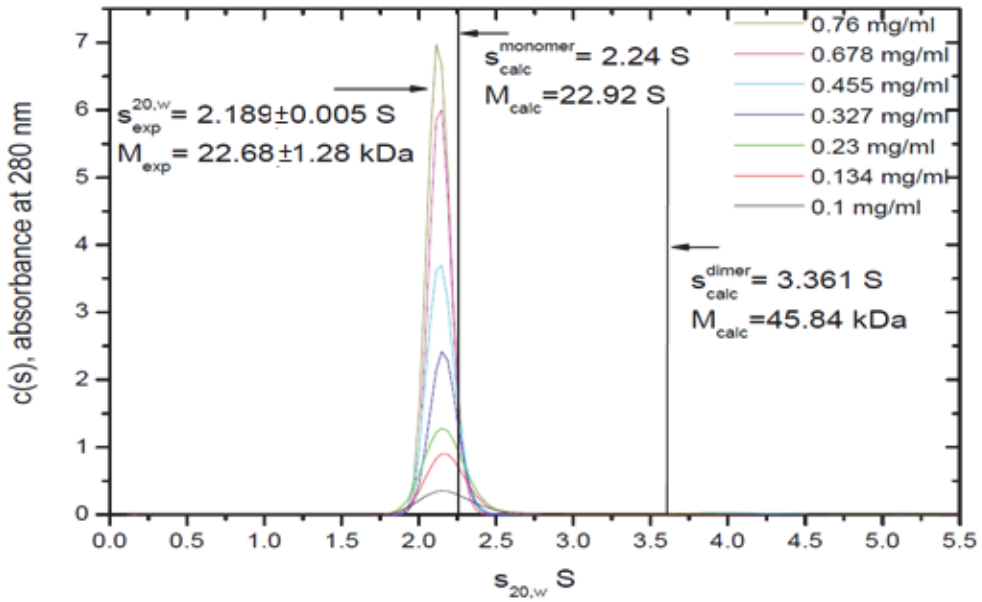


Figure 3.16 Sedimentation boundary modelled as if arising from a single sedimenting species.



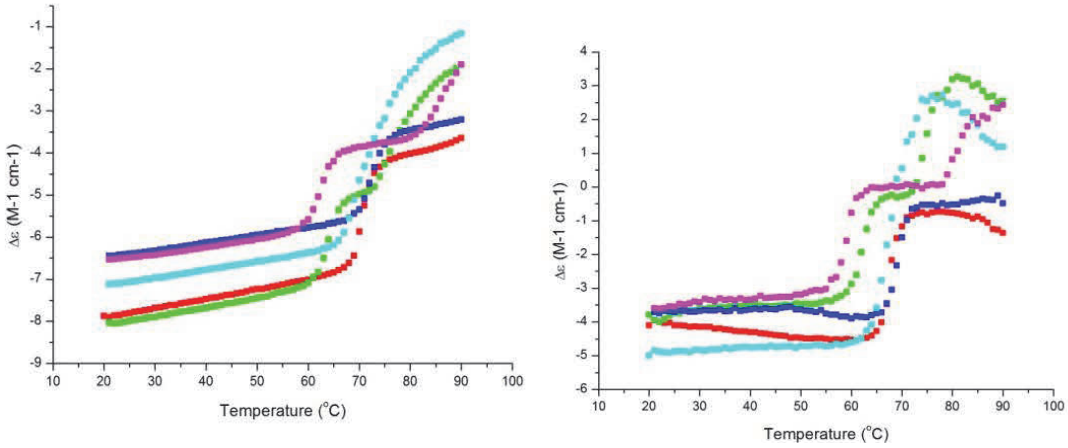
The Molar mass distribution c(M) of ColA-P WT monomer calculated by amino acid composition is 22.92 kDa, transformed into a Sedimentation coefficient distribution c(s), the peak of the monomer protein should appear at 2.24 S (S=Svedbergs). The

molar mass of ColA-P WT dimer calculated by amino acid composition is 45.84 kDa which corresponds to a c(s) distribution of 3.361 S. Thus, the peak of the dimer protein should appear at 3.361 S. The results obtained from a series of experiments with different concentrations of ColA-P WT (figure 3.11) show a range of peaks with the same c(s) distribution at 2.189 S. Transformed into a c (M) distribution, the peaks correspond to a species of 22.68 kDa, close to that of a monomer. This result suggests that Colicin A pore-forming is monomer and this protein purity is more than 95% of the protein present.

This result can be regarded as a final solid evidence that colicin A-P domain is monomer in solution.

Variable temperature measurement

The CD transition $n \rightarrow \pi^*$ is primarily responsible for the negative bands at 222nm, which can also be used as a detector of the characteristic of the α-helix (Hakkaart et al., 1981), whilst, as mentioned above, the CD spectrum in the near UV region gives the information about the tertiary structure of the protein and the 295 nm Tryptophan peak can be used to detect of the characteristics of the tertiary structure. (Sabik et al., 1983) The far and near UV CD spectra of folded histidine tagged colicin A P-domain wild type and its mutants show their characteristic minima near 222 nm and 295 nm respectively. Therefore, the CD intensities at 222 nm and 295 nm were acquired over a range of temperatures between 20 ^{o}C and 90 ^{o}C, using 0.1 cm and 1cm path length cuvettes respectively. The heating rate was 1 ^{o}C/min. Samples were pre-equilibrated at 20 ^{o}C for at least 20 minutes. The thermal denaturation transition curves of the ColA-P proteins are shown in Figure 3.17.



A) B)

Figure 3.17: A) Temperature-dependant CD spectra of ColA-P WT (Red curve), D227A (Green curve), D241A (Blue curve), D274A (Cyan curve) and D372A (Magenta curve) at wavelength of 222nm. B) Temperature-dependant CD spectra of ColA-P WT (Red curve), D227A (Green curve), D241A (Blue curve), D274A (Cyan curve) and D372A (Magenta curve) at wavelength of 295 nm.

The CD spectra of colicin A-P domain in the Far-UV region were only slightly reduced upon increasing the temperature between 20 ^{o}C and 65 ^{o}C, indicating that the tertiary structure of the wild type protein is largely stable within this temperature range. A single sharp transition occurs between 65 ^{o}C and 72 ^{o}C. The CD spectra of the wild type colicin A-P domain in the near-UV region were not affected too much by temperature from 20 ^{o}C to 70 ^{o}C. A single transition appears between 70 ^{o}C and 77 ^{o}C.

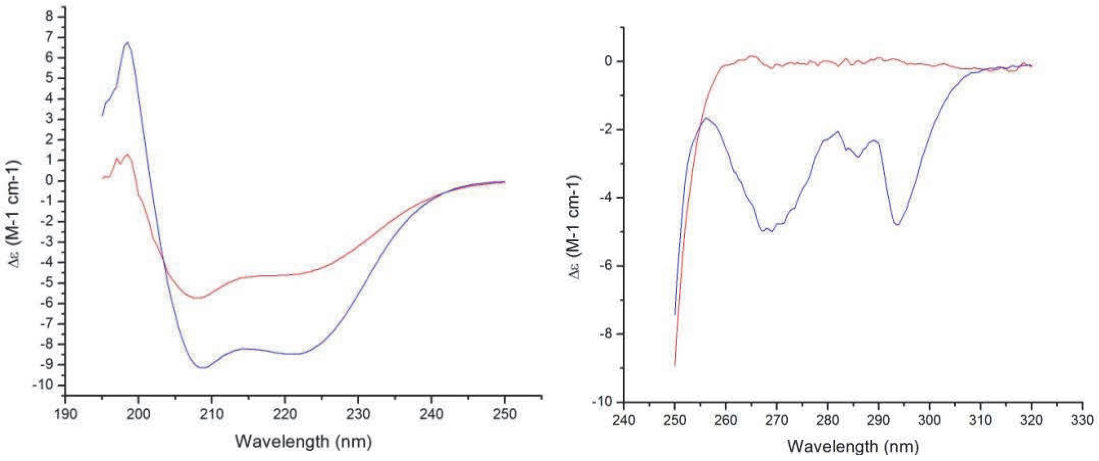
ColA-P D227A and ColA-P D274D mutants are stable in the temperature range between 20 ^{o}C and 59 ^{o}C. However, a biphasic thermal denaturation transition was observed by CD in the far-UV region, suggesting that the secondary structure of mutants were denatured by heating in two steps. The thermally unfolded intermediate of these mutants, as measured by both near and far UV signal exists over a much smaller region that the wild type which is still stable at 90 ^{o}C.

ColA-P D372A mutant also showed biphasic thermal transitions in both the far and near UV region. The first transition started from 55 ^{o}C in the far-UV region and from 59 ^{o}C in the near-UV region, suggesting that ColA-P D372A mutant is less stable than the other mutants. However, the second transition of ColA-P D372A started from 78 ^{o}C in the far-UV region and the near-UV region, suggesting that the intermediate stage of ColA-P D372A mutant is more stable than the other mutants but less stable than the wild type.

The temperature-dependant Near-UV CD spectra showed that the D227A mutant completed its first thermal denaturation process at 68 ^{o}C reaching the $\Delta\varepsilon$ of 0 (M^{-1} cm^{-1}) and the second transition was stopped at 75 ^{o}C by the $\Delta\varepsilon$ of 3 (M^{-1} cm^{-1}). Conversely the wild type completed its thermal transition process at 70 ^{o}C reaching the $\Delta\varepsilon$ of 0 (M cm). This suggests that the thermal denaturation process of colicin A-P domain wild type might also include two steps, but the second transition appears beyond the temperature range between 20 ^{o}C and 90 ^{o}C.

The Far UV spectra of both D227A and D274A indicated that, the mutated proteins keep some of their secondary structure until 70 ^{o}C where the $\Delta\varepsilon$ is around -+5 (M^{-1} cm^{-1}), and did not fully unfold until 90 ^{o}C. However, the $\Delta\varepsilon$ of wild type protein at the end of the single transition was around -4 (M cm). Therefore, the secondary structure of colicin A-P domain wild type appears to unfold by two steps, and the second step is widely separated from the first one.

Temperature-dependant CD spectra of the wild type colicin A P-domain at both near and far UV region have shown a single unfolding transition between 20 ^{o}C and 90 ^{o}C. However, it is unusual that the $\Delta\varepsilon$ at 295 nm in near UV region rise above 0 (M cm). Thus we measured the complete far and near UV CD spectra of the protein at 25 ^{o}C and 80 ^{o}C to examine the nature of the temperature-dependent CD spectra, which are shown in Figure 3.18.



B)

Figure 3.18: A) Far UV CD Spectra of ColA-P wild type at 25 ^{o}C (Blue line) and 80 ^{o}C (Red line). B) Near UV CD Spectra of ColA-P wild type at 25 ^{o}C (Blue line) and 80 ^{o}C (Red line).

The near and far UV spectra show that the protein retains part of its secondary structure and completely losses its tertiary structure at 80 ℃. However, the thermal denaturation experiment performed by CD spectroscopy showed that the protein still retained part of its tertiary structure at 80℃. The problem may be caused when insufficient equilibration time (one minute) is allowed at the end of a temperature programmed run.

3.3 Discussion

Biphasic unfolding of Colicin A pore-forming domain

The molten globule state is a major class of folding intermediate, which can be formed by acidic pH, high temperature, or intermediate concentrations of denaturation. (Ptitsyn et al., 1990) A stable molten globule state with a native-like content of secondary structure and flexible tertiary structure was found in ColA-P at low pH

(van der Goot et al., 1992b). Further studies indicated that the ColA-P forms a molten globule configuration at low pH and allows the protein to insert into the inner membrane much more efficiently *in vitro* (Lakey et al., 1994, Pattus et al., 1983b, van der Goot et al., 1991, Lakey et al., 1991b, Lakey and Slatin, 2001). Here the far UV CD spectra at pH 3.0, pH 5.0 and pH 8.0 confirm that the secondary structure of ColA-P wild type and alanine mutants were slightly unfolded in acidic environments (Figure 3.7). However, the near UV CD spectra at different pH demonstrated that the wild type retained tertiary structure at pH 3.0 whereas the mutants appeared fully unfolded. The wild type completely lost its tertiary structure below pH 3.0. This confirms that the compact single ColA-P domain can form a molten globule state under pH 3.0 and that the alanine substitutions, previously shown to destabilise at neutral pH affect this process. (Fridd and Lakey, 2002, Fink et al., 1994, Evans et al., 1996)

Thermal unfolding of ColA-P WT showed a single transition around 70 ^{o}C which leaves about 50% of the alpha helical signal remaining. The near UV tertiary signal is zero above this transition but may not represent a fully unfolded form (see later). The temperature dependant CD spectra of the ColA-P alanine mutants in both near and far UV regions represented two-step thermal transitions which lead to bigger decreases in signal between 20^{o}C and 90 ^{o}C (Figure 3.17). This is clearest for D227A and D372A but appear merged in D241A and D274A. A stable intermediate stage thus occurs between 66 ^{o}C to 80 ^{o}C in the far UV region for the mutants and remains folded even at 90^{o}C in the wild type. This means a stable intermediate stage with a part of the helix structure can be retained at a high temperature. The full Near UV CD spectrum of ColA-P D227A at 80 ^{o}C do not agree with the data. According to the Far-UV molar ellipticity differences between the low temperature and the intermediate temperature, we can obtain the similar conclusion that the wild type protein was not fully unfolded until 90 ^{o}C. Far and near UV CD Spectra of ColA-P wild type at 25 ^{o}C and 80^{o}C confirmed that ColA-P WT retained part of its secondary structure but completely lost

its tertiary structure at 80 ^{o}C (Figure 3.16). Importantly, there is no evidence in the DSC data of the biphasic thermal transition observed by CD in ColA-P alanine mutants. However, the pH dependent thermal dynamic transitions measured by DSC show a broader transition of ColA-P and lower calorimetric enthalpy at pH 3.0, (Figure 3.8) which might be due to the separation of the two transitions at low pH. This is in agreement with the more extensive study of ColA-P at low pH (Muga et al., 1993a), and shows that the alanine mutants destabilize the protein to various degrees. According to the temperature-dependent far UV CD spectra, the melting temperature value (T_1) of the initial transition of ColA-P D372A $<T_1$ of ColA-P D227A $<T_1$ of D274A $<T_1$ of D241 $<T_1$ of WT. (Figure 3.17) In conclusion, a biphasic transition occurs only in the mutant ColA-P domains thermal denaturation process, and ColA-P D372A destabilizes more than other alanine mutants. The intermediate stage of ColA-P WT is very stable at high temperature, and more stable than the intermediate stages of ColA-P alanine mutants are unstable at higher temperature.

The urea and guanidine unfolding experiments also revealed the surprising result that ColA-P is a compact single domain protein which unfolds through two steps. (Figure 3.2 and 3.3) Except the point mutation D372A and D376A which are located at helix 10, most of the alanine mutants only destabilize the initial transition and do not have a big effect on the later transition. (Figure 3.2) The formation of the molten globule state would reduce the energy to insert into the membrane of cells. (van der Goot et al., 1992a, Muga et al., 1993a) This has also been shown to correlate with the presence of acidic lipids *in vivo* (van der Goot et al., 1993).

Thus, the initial transition appears to be more biologically relevant than the later one. When we lower the pH, the mid points of the initial unfolding phases of ColA-P WT will move towards lower urea concentrations. (Figure 3.5, 3.6 and 3.7) The initial phases of ColA-P D227A and D372A were completely suppressed below pH 3.0 and that of ColA-P D216A below pH 2.5. (Figure 3.4) The barycentric means of the

fluorescence emission wavelengths of ColA-P D227A and D372A between pH 2.0 and pH 3.0 almost reach their intermediate stage in the absence of urea, indicating that the structures of ColA-P D227A and D372A are already in their molten globule state in the absence of urea below pH 3.0. The barycentric mean of ColA-P D216A at pH 2.0 and pH 2.5 without urea is also close to its intermediate stage in mild concentrations of urea, which means D216 also forms its intermediate stage below pH 2.5. (Figure 3.5) The barycentric mean of ColA-P proteins will reach 362 nm at neutral pH, when they are fully unfolded in urea or guanidine. (Figure 3.2 and Figure 3.3) However, in acidic environment, the barycentric mean of the fluorescence emission wavelength of ColA-P alanine mutants does not extend beyond 353 nm in 8 M urea. (Figure 3.5, 3.6 and 3.7) This means the proteins remain in the intermediate stage even at very high urea concentrations at low pH, thus this form seems to be stabilized by low pH and mutation Figure 3.4. This indicates that the second transition tends to shift toward higher urea concentration when low down the pH. Except for D376A, the estimated ΔG_1 of the alanine mutants is lower than the wild type, which confirms the hypothesis that it is easier for an alanine mutant to transform from its folded stage to the intermediate stage.

The Near and Far UV molar ellipticity of ColA-P WT and its mutants at 208 nm and 295 nm were used to describe the secondary structure and the tertiary structure contents of the protein respectively. Figure 3.13 and 3.14 show that ColA-P WT and D216A remain folded in 0 M urea at pH 3.0 and in 0 M and 2 M urea solution at pH 8.0. It forms the intermediate stage in 2 M and 6 M urea at pH3.0 and in 6 M urea solution at pH 8.0. ColA-P D227A and D372A retains its folded stage in 0 M and 2 M urea solution at pH 8.0. It forms the intermediate stage in 0 M, 2 M and 6 M urea at pH3.0 and 6M urea solution at pH 8.0. Figure 3.10 shows that ColA-P WT and its alanine mutants keep most of their secondary structure in 0 M and 2 M urea at pH 8.0 and lose half of their secondary structure in 6 M urea solution at pH 8.0. At pH 3.0, these proteins keep around 80% of their secondary structure in 0 M urea solution and

lose half of its helical content in 2 M and 6 M urea solution. This means ColA-P protein retains half of its secondary structure at its intermediate stage. (Figure 3.11) ColA-P WT and its alanine mutants keep more than 50% of their tertiary structure in 2 M urea solution at pH 8.0, but totally lose their tertiary structure in 6 M urea at pH 8.0. At pH 3.0, only the wild type protein retains around 50% of tertiary structure. This means, the proteins start to lose their tertiary structure during the initial transition, and have completely lost rigid structure in the intermediate stage. The tertiary structure transition occurs without urea at pH 3.0 for the mutants whereas the wild type needs 2 M urea. (Figure 3.11)

Aspartate to alanine mutants destabilize the Colicin A pore-forming domain

Alanine is the simplest amino acid with a chiral center. The side chain of alanine is nonpolar, which tends to stabilize protein structure by means of hydrophobic interactions (Nelson and Cox, 2008). However, the sequential replacement of surface aspartate of ColA-P domain by alanine revealed five sites where this change destabilized the protein at both neutral pH and low pH.

The urea and guanidine HCl barycentric mean unfolding curves showed that the first unfolding transition of the alanine mutants occurred at lower denaturant concentrations than wild type at neutral pH (Figure 3.1 and Figure 3.2). They also showed that the aspartate to alanine mutation destabilized the protein at low pH. ΔG_1 is the change in Gibbs free energy which can be used to detect the energy that is needed for the protein transformation from folded stage to unfolded state (Pace and Scholtz, 1997). The lower $\Delta G1$ of ColA-P alanine mutants compared to wild type means that the alanine mutants need less energy to unfold from fully folded stage to the molten globule state, which confirms that the alanine mutants are less stable than the wild type protein and ColA-P D372A is the most unstable aspartate to alanine mutant. (Tables 3.1 and 3.2) ΔG_2 is the change in Gibbs free energy which can be used to detect the energy that is needed for the protein transformation from its molten

globule state to the unfolded stage (Pace and Scholtz, 1997). Table 3.1 has shown that, for ColA-P proteins, ΔG_2 are similar as each other at neutral pH, which means that for all of the alanine mutants, similar energies were needed for the proteins unfold from its molten globule stage to its unfolded stage. It might give us the information that the later transition is not effected by aspartate to alanine mutations. Although the second transitions were abolished at low pH (Figure 3.4; 3.5; 3.6 and 3.7), the start of the second transitions were obviously delayed to high urea concentration conditions in acidic environments. It confirmed that, the second transition is related to the acidic environment rather than the point mutations; and the acidic environment helped the protein to keep its molten globule state *in vitro*.

The urea unfolding experiment revealed that the aspartate to alanine mutants destabilized Colicin A pore forming domain in acidic environment (Figure 3.4, Figure 3.5, Figure 3.6 and Figure 3.7). Beside, the barycentric mean wavelengths of ColA-P alanine mutants are greater than the wild type values in the absence of urea below pH 3.0, which means the conformation of ColA-P alanine mutants is already less compact in zero urea below pH 3.0 and the protein has been more destabilized than the wild type by low pH. However, the barycentric means of ColA-P WT between pH 2.0 and pH 3.0 are lower than those between pH 3.5 and pH 4.5 in the absence of urea, and it might be because the wild type protein is more compact in the absence of urea below pH 3.0. (Figure 3.4)

The temperature-dependant CD spectra also demonstrated that ColA-P domain is destabilised by the alanine mutations. The melting temperatures of the initial transitions of the alanine mutants are lower than that of the wild type, which correlate with the decrease in the initial transition stability observed by urea and guanidine HCl denaturation. (Figure 3.2 & 3.17) The wild type protein did not fully unfold even at 90 ^{o}C. It means both transitions of the alanine mutants are less heat resistant than the wild type with the effect that the upper transition becomes visible for the first time.

The DSC profile (Figure 3.7) also confirms that aspartate to alanine mutations destabilize the protein and ColA-P D372A is the most unstable mutant of all the alanine mutants. The T_m value ("melting temperature" = midpoint of the thermal transition) of ColA-P WT at pH 3.0 is much lower than that at neutral pH, which indicates the protein is destabilized in low pH environments.(Figure 3.8) (Muga et al., 1993a). The difference between calorimetric enthalpy ΔH and Van't Hoff enthalpy ΔH_v, which was not reported in the previous work, suggests that the thermal dynamic transition cannot be regarded as a two-state transition. The immediate explanation that it is due to the existence of a dimer state (Tables 3.3 and 3.4) (Igor N. Serdyuk et al., 2007) is not supported by the results of Native PAGE, Size Exclusion Chromatography and Analytical Ultracentrifugation which show solid evidence that ColA-P protein is monomer in solution. Thus we cannot find a clear explanation for the difference between ΔH and ΔH_v (Table 3.1). The thermal denaturation experiments are repeatable but not reversible since the reverse scans appeared to show aggregated proteins and are not shown in this thesis.

It is clear from the pH experiment that the ratio between ΔH and ΔH_v increases at low pH with a steep drop in ΔH and thus may be related to the two phases of unfolding. Since the structural change to the molten globule becomes less evident in DSC at low pH it may be that tertiary structure unfolding is broad and not picked up above the baseline.

Another potential explanation is that colicin A-P domain unfolding via intermediate has two steps that are widely separated with only one of the steps unfolding in the given temperature range. The CD data present two obvious separated transitions which match the biphasic unfolding process of urea and guanidine induced unfolding.

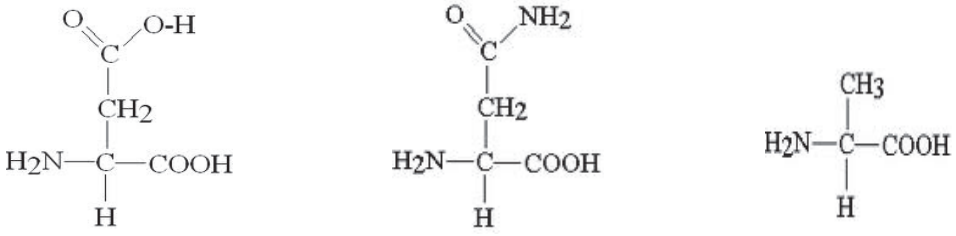
In conclusion, alanine mutants destabilized the colicin A pore forming domain in various ways. Mainly, it facilitates the formation of molten globule state and reduces the thermal stability of the protein. Some of the aspartate to alanine mutants, such as

ColA-P D216A, D227A and D372A destabilise the protein more than others, such as ColA-P D241A, D274A and D376A. So far the position of each mutant does not explain how it destabilises the protein since they are spread over the whole sequence. Unexpectedly the ColA-P domain, a small compact protein, unfolds in two steps in solution but the molten globule state is likely to be the insertion-competent state which leads to its pore forming activity.

Chapter Four

The effect of aspartate to asparagine mutations on the stability of the Colicin A P-domain

The preliminary study demonstrated that ColA-P was destabilized by Asp to Ala replacements at pH 7.0. (Fridd and Lakey, 2002) Work shown in Chapter three of this thesis revealed that these alanine mutants of ColA-P were destabilized at neutral pH, acidic environments between pH 2.0 and pH 4.5 and at high temperature. Two phases of ColA-P unfolding were defined by urea and guanidine HCl denaturation methods with the initial phase most affected by the mutations. Both phases were affected by low pH. From this data it is possible that the surface negative charges from the aspartates , which are lost at low pH and upon mutation, play an essential role in stabilizing the protein. In an Asp to Asn mutant the side chain of asparagine replaces the charged carboxylate with an amine hydrogen bond donor and retains the hydrogen bond accepting carbonyl group. It is thus a conservative mutation compared to alanine and, thus, mutation to asparagine has also been chosen to study the effect of ionic charge on the stability of the protein.



A) B) C)

Figure 4.1: Structures of A) Aspartic acid; B) Asparagine; C) Alanine

4.1 Aims

In this chapter, the surface charges due to Aspartic acid residues, previously shown to stabilise the ColA-P domain, were removed by mutation to asparagine. This retains the hydrogen bond accepting carbonyl group and is designed to probe the negative charge dependence of ColA-P domain stability. These mutants were studied using the same methods as in Chapter 3.

4.2 Results

Studying the stability of Aspartic acid to Asparagine mutants of the ColA-P domain by fluorescence spectroscopy

The intrinsic fluorescence of ColA-P asparagine mutants in a series of urea concentrations at pH 7.0 in MOPS buffer was measured. The barycentric mean of the emission wavelength has been calculated and plotted to yield a urea unfolding curve. (Figure 4.2)

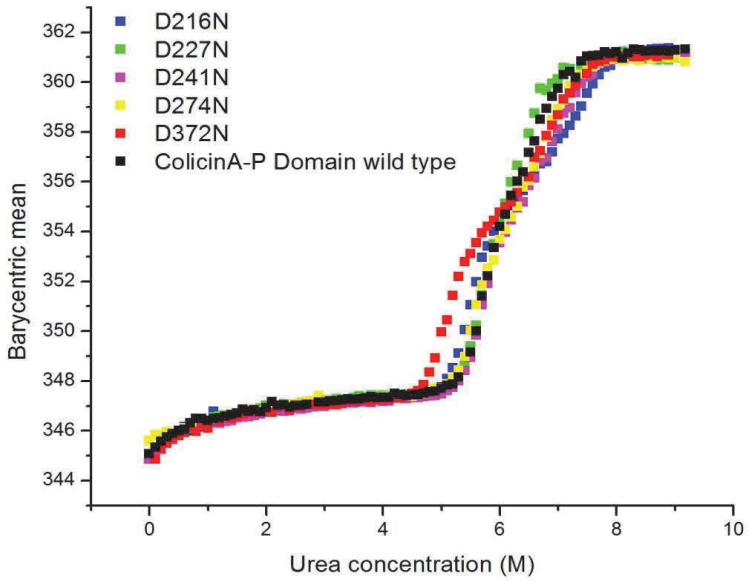


Figure 4.2: Biphasic urea unfolding transitions of the ColA-P wild type (Black curve), D216N (Blue Curve), D227N (Green curve), D241N (Magenta curve), D274N (Yellow curve) and D372N (Red curve) at pH7.0 in MOPS buffer.

Figure 4.2 shows the unfolding process of Colicin A P-domain wild type and aspartate to asparagine mutant proteins triggered by urea. Compared with the alanine mutants (figure 3.2), the first transitions of all the asparagine mutations resemble more the wild type protein. As with D372A, the first transition of D372N is clearly the least stable of the mutant series. On the other hand, compared with the wild type protein, the two transitions of D227N were even closer to each other making the urea unfolding curve of D227N is very close to a two state transition. Hence, D372N, D227N and the wild type ColA-P were selected for the guanidine hydrochloride unfolding experiment.

The intrinsic fluorescence of ColA-P domain wild type, D227A and D372A in a series of concentrations of guanidine hydrochloride at pH7.0 in MOPS buffer were measured. The barycentric mean of each emission spectrum was calculated and was plotted against guanidine hydrochloride concentration to yield the unfolding curve. (Figure 4.3)

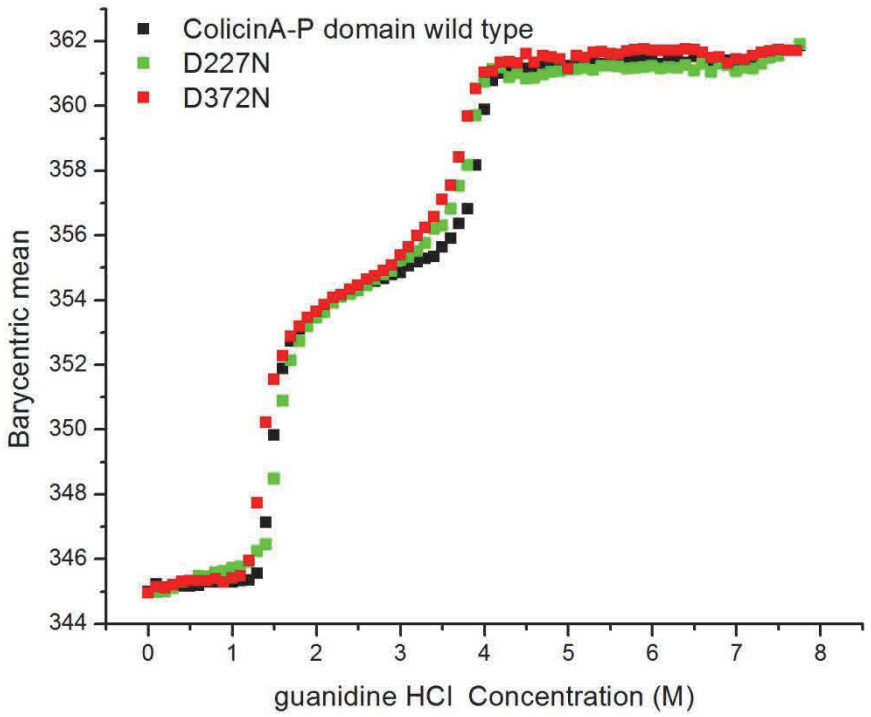


Figure 4.3: Biphasic guanidine hydrochloride unfolding transitions of the ColA-P wild type (Black curve), D227N (Green curve) and D372N (Red curve) at pH 7.0 in MOPS buffer.

Figure 4.3 shows the biphasic unfolding process of wild type and aspartate to asparagine mutant ColA-P domain proteins triggered by guanidine HCl. Two separate transitions of D372N and D227N can be observed in this figure. Thus, three distinct states of D227N not observable by urea denaturation were confirmed. The guanidine HCl unfolding curves of ColA WT, D372N and D227N are similar to each other, which means the stability of the mutated proteins are very similar to wild type. As the same for the urea unfolding curve (Figure 4.2), Figure 4.3 also shows that D372N is slightly less stable than the wild type protein. The second transitions of both D227N and D372N are more stable than the wild type protein.

Gibbs free energy values were calculated for the folded, intermediate and unfolded states of ColA-P asparagine mutants. Unfortunately, the double transitions of ColA-P asparagine mutants denatured by urea are too close to each other to calculate these values. Additionally, for mutants such as D227N, D241N and D274N, only a single transition is visible and the Gibbs free energy of the two transitions cannot be easily defined. However, the biphasic transitions of ColA-P D216N and D372N can be defined, and the Gibbs free energy of the two transitions of D216N and D372N were calculated shown in table 4.1.

Table 4.1. The Gibbs free energy of Biphasic unfolding for ColA-P asparagine mutants detected by urea unfolding at pH 7.0

Mutation	$\Delta G1$ Kcal.mol^{-1}	$\Delta G2$ Kcal.mol^{-1}	$\Delta G1+\Delta G2$ Kcal.mol^{-1}	First transition (Urea concentration M)	Second transition (Urea concentration M)
D216N	21.42	18.62	40.04	5.0-6.1	6.8-8.0
D372N	18.62	10.98	29.60	4.6-6.0	6.2-7.7
WT	21.83	10.45	32.38	5.2-6.0	6.1-7.5

Table 4.1 indicates that $\Delta G1$ of ColA-P D372N are distinctly lower than D216N and wild type, which means that less energy is needed to unfold to the molten globule state. On the other hand, $\Delta G2$ of ColA-P D216N is greater than ColA-P WT and other asparagine mutants, which means that D216N needs more energy to unfold from molten globule state to its

unfolded state. This result may be explained by the shallower gradient of this transition even though its mid-point is the highest.

In chapter three, I showed that ColA-P alanine mutations at the site point D227 and D372 were significantly destabilizing. D227N and D372N were thus selected as a comparison to those Ala mutations. The urea unfolding experiments shows that D372N is the most unstable mutants of all of the asparagine mutants and D227N more resembles the wild type. The Gibbs free energy of the two transitions for ColA-P D227N and D372N was calculated and shown in table 4.2.

Table 4.2. The Gibbs free energy of ColA-P asparagine mutants' biphasic unfolding process at pH 7.0 in guanidine HCl compared to the corresponding alanine mutants (Chapter Three).

Mutation	$\Delta G1$ Kcal.mol^{-1}	$\Delta G2$ Kcal.mol^{-1}	$\Delta G1+\Delta G2$ Kcal.mol^{-1}	First transition (Gdn HCl concentration M)	Second transition (Gdn HCl concentration M)
D227N	13.76	33.48	47.24	1.3-1.9	3.5-4.1
D227A	8.85	22.47	31.32	0.8-1.4	3.1-3.9
D372N	11.32	18.55	29.87	1.1-1.7	3.2-4.1
D372A	6.55	17.32	23.87	0.6-1.2	3.5-4.3
WT	13.16	32.03	45.19	1.3-1.8	3.6-4.2

Table 4.2 shows that The $\Delta G1$ values decrease in the order: wild type $\approx$ D227N $>$ D372N $>$ D227A $>$ D372A, (alanine mutant results from Fig 3.2 shown for comparison). This means D372N needs less energy to unfold to its molten globule state than D227N, just as D372A needs less energy to unfold to its molten globule state than D227A. It also indicates that the alanine mutants D227A and D372A need less energy to unfold to their molten globule states than asparagine mutants D227N and D372N. Thus asparagine mutants are more stable than alanine mutants; and very similar to the wild type protein.

Studying the stability of aspartate to asparagine mutants of ColA-P domain at low pH by fluorescence spectroscopy

In order to determine the stability of ColA-P aspartate to asparagine mutants in an acidic environment, the intrinsic fluorescence of ColA-P domain wild type, D216N, D227N and D372N were measured in a series of concentrations of urea at pH3.0 in sodium citrate buffer. In order to detect the higher transition of ColA-P WT and its asparagine mutants at pH 3.0, a very high concentration of urea solution has to be used. The barycentric mean of each emission spectrum was calculated and plotted to provide a urea unfolding curve. (Figure 4.4)

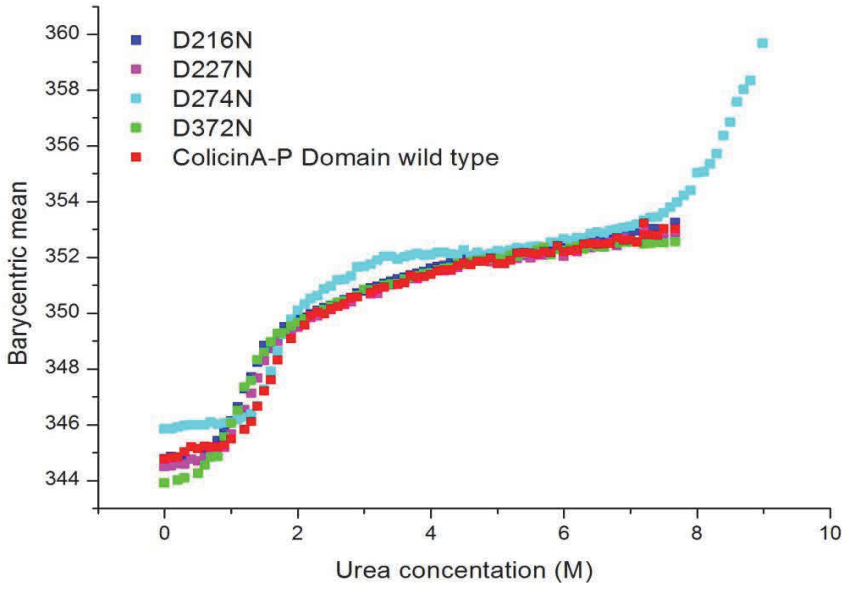


Figure 4.4: The urea unfolding curves of the ColA-P wild type (Red curve), D216N (Blue Curve), D227N (Magenta curve), D274N (Cyan curve) and D372N (Green curve) at pH3.0 in sodium citrate buffer.

Figure 4.4 shows that a clear transition appears at low urea concentration in the urea unfolding of ColA-P WT and its asparagine mutants at pH 3.0. Comparing with wild type, all of the asparagine mutants were just slightly destabilized at low pH. ColA-P D372N and D227A were found to be more unstable than D216N. The second transition appears on ColA-P D274N at very high urea concentrations at pH 3.0. It

confirms that the later transition shifts towards the higher urea concentration s for both WT and asparagine mutants when the pH is decreased.

The Gibbs free energy for the unfolding of each mutant in urea solution at pH 3.0 was calculated as above and shown in table 4.3.

Table 4.3. The Gibbs free energy of ColA-P asparagine mutants' biphasic unfolding process in urea at pH 3.0

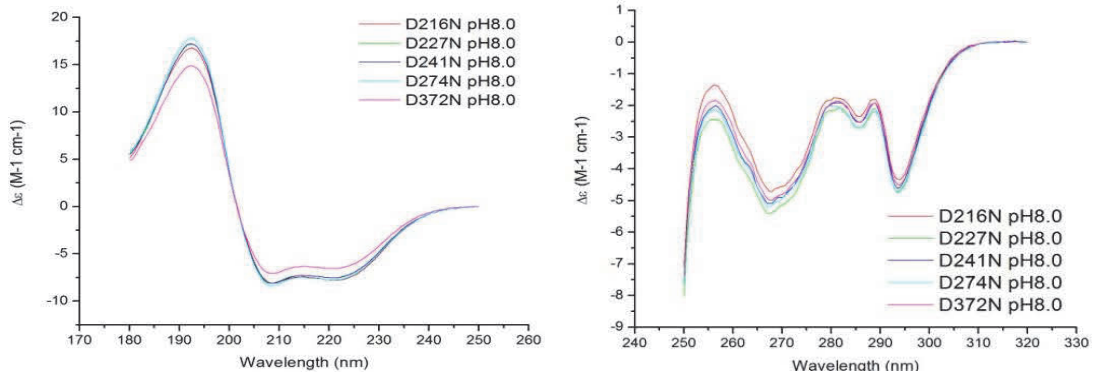
Mutation	$\Delta G1$ Kcal.mol^{-1}	$\Delta G2$ Kcal.mol^{-1}	$\Delta G1+\Delta G2$ Kcal.mol^{-1}	First transition (urea concentration M)	Second transition (urea concentration M)
D216N	3.33	nd	nd	0.5-1.8	nd
D227N	3.94	nd	nd	0.8-2.3	nd
D274N	3.65	6.61	10.26	1.2-2.9	nd
D372N	2.60	nd	nd	0.7-1.8	nd
WT	4.66	nd	nd	1.0-2.2	nd

nd (not determined as only a single transition is visible at this pH)

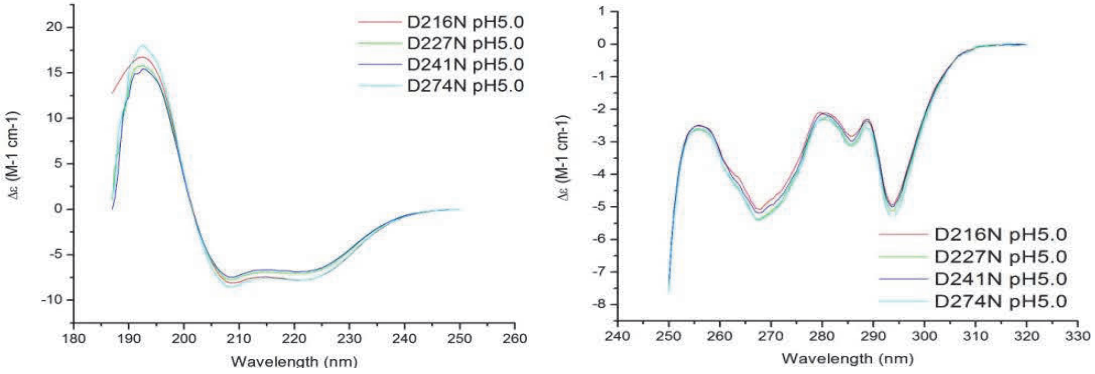
Table 4.3 shows that the $\Delta G1$ of the wild type$>\Delta G1$ of D227N $>\Delta G1$ of D274N $>$ $\Delta G1$ of D216N $>\Delta G1$ of D372N at pH3.0.

Circular Dichroism studies of ColA-P Aspartate to Asparagine mutants.

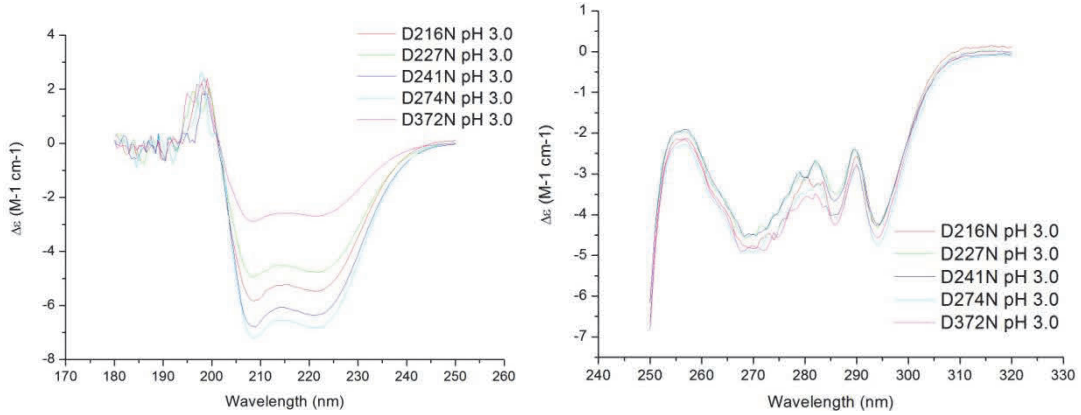
The urea and guanidine HCl unfolding experiments showed that ColA-P aspartate to asparagine mutation does not destabilize the protein as much as the alanine mutants. To further characterise this, the far and near UV CD spectra of the Colicin A-P aspartate to asparagine mutants were measured at pH 8.0, pH 5.0 and pH3.0 to determine their secondary and tertiary structure respectively.



A)



B)



C)

Figure 4.5 A)The Far and near UV CD spectra of ColA-P D216N (Red line), D227N (Green line), D241N (Blue line), D274N (Cyan line) and D372N (Magenta line) at pH 8.0; B) The Far and Near UV CD spectra of ColA-P D216N (Red line), D227N (Green line), D241N (Blue line) and D274N (Cyan line) at pH 5.0; C) The Far and Near UV CD spectra of ColA-P D216N (Red line), D227A (Green line), D241N (Blue line) and D274N (Cyan line) at pH 3.0.

The characteristic intensity maxima in both near and far UV CD spectra of ColA-P asparagine mutants were not significantly affected by the absence of a single negative charge on the surface of the protein at pH 8.0 and pH 5.0. ColA-P D372N has less helical content compared with other asparagine mutants and the WT. The Far UV CD spectra of asparagine mutants was significantly affected at pH 3.0. Figure 4.5 E) shows the helical contents of asparagine mutants as below: D372N < D227N < D216N <D241N < D274N. Surprisingly Figure 4.5 F) shows that they retain most of their tertiary structure signal at pH 3.0.

The thermal transition temperature and the calorimetric transition enthalpy

Thermal denaturation of ColA P-domain mutants D216N, D227N, D241N, D274N and D372N was studied by differential scanning calorimetry. The melting temperature, calorimetric enthalpy and Van't Hoff enthalpy were calculated and analysed. The protein samples were studied in a temperature region of 25 to 90 ^{o}C at a scanning rate of 1 ^{o}C/min. The proteins were dissolved in 50 mM phosphate buffer to a concentration of 0.5 mg/ml, and the pH was adjusted to pH 8.0. Characteristic DSC scans for ColA-P aspartate to asparagine mutant are shown in figure 4.5.

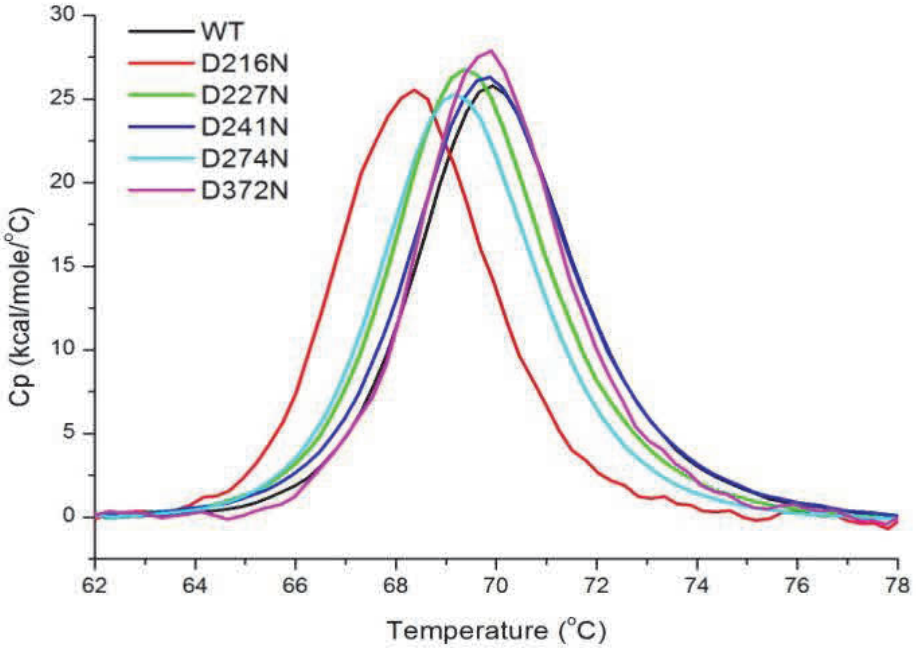


Figure 4.5 Characteristic heat-capacity curves for ColA-P WT (Black line), D216N (Red line), D227N (Green line), D241N (Blue line), D274N (Cyan line) and D372N (Magenta line) with buffer baselines subtracted and normalized for protein concentrations. Each of the ColA-P proteins showed only one heat capacity peak within the temperature region 25 to 90^{o}C.

Table 4.3 Calorimetric data of ColA-P mutants in aqueous solution at 1 ^{o}C /min

Mutation	Thermal transition temperature *Tm*(℃) ±error	*ΔH* (kcal/mol) ±error	*ΔHv* (kcal/mol) ±error
D216N	68.41±0.035	93.12±0.67	208.25±4.51
D227N	69.50±0.050	106.1±2.77	228.8±8.03
D241N	69.83±0.038	110.3±1.97	210.5±4.84
D274N	69.26±0.026	103±1.39	219.95±2.90
D372N	69.88±0.076	105.3±0.01	239.6±2.09
WT	70.01±0.087	107.8±0.49	220.7±0.66

The ColA-P Asparagine mutants were considered firstly to unfold with a two-state transition and by fitting the data to this model the calorimetric transition enthalpy can be calculated.

ColA-P WT and its asparagine mutants measured by DSC all show the similar melting temperatures. The temperature midpoint of the thermal transition (Tm) of ColA D216N is slightly lower than the other asparagine mutants. The Tm of D227N, D241N, D274N and D372N are almost the same as the wild type. The calorimetric enthalpy of ColA-P D216N is slightly lower than the other asparagine mutants. Col A-P D227N, D241N, D274N and D372N almost need the same amount of energy as its wild type to perform the thermal unfolding process.

4.3 Discussion

The stability of Asparagine mutants

A previous study had shown that surface aspartate residues are important in ColA-P stability (Fridd and Lakey, 2002). In chapter three we showed that the protein was destabilized by protonation of negatively charged side chains at low pH and by Ala mutation. Thus, destabilisation by alanine insertion was thought to be due to the loss of the negative charge. Asparagine is structurally similar to aspartate but contains an amino group instead of the hydroxyl group, which means an Asp to Asn mutation can also remove the negative charge. Therefore the replacement of Asp by Asn was expected to destabilise the ColA-P.

The urea and guanidine HCl unfolding experiments revealed that ColA-P asparagine mutants also unfolded through two closely spaced steps at neutral pH. (Figure4.2 and 4.3) At low pH, the initial unfolding transitions of ColA-P asparagine mutants move towards the low urea concentration direction and the later phases of the asparagine mutants move towards the high urea concentration (Figure 4.4). This is clear evidence to prove that the biphasic transition exists in the urea unfolding process at low pH. Furthermore this result also indicates that the second unfolding transition of the ColA-P proteins is delayed by low pH and therefore the intermediate molten globule

state is stabilized by acidic condition. Table 4.3 indicates that the free energy change for ColA D372N transformation from its folded state to its intermediate stage, and the wild type protein needs more energy to complete the unfolding process. In conclusion, although the stability of asparagine mutants is almost the same as its wild type, the mutation D372 slightly destabilizes the protein at low pH.

The Far and Near Circular Dichroism spectra show that ColA-P D372N slightly loses its secondary structure at neutral pH and the rest of the asparagine mutants retain their secondary and tertiary structure at pH 5.0 and pH 8.0 (Figure 4.5). At pH 3.0, all of the asparagine mutants lose their secondary structure to various degrees but retain their tertiary structure. According to the definition of the molten globule state, (van der Goot et al., 1992b) although the asparagine mutants such as D227N and D372N lose part of its secondary and tertiary structure, they are still in its native structure rather than in the molten globule state at pH 3.0. The far UV CD spectra at pH 3.0 also indicate that the mutations at D216, D227 and D372 destabilised the secondary structure protein at low pH. This is surprising considering the stability of ColN-P. It suggests that the aspartate to asparagine point mutations destabilise parts of the helical coils whilst having no effect on the stability of the hydrophobic core of the protein at low pH.

Thermal denaturation measured by differential scanning calorimetry shows that the stability of asparagine mutants is very similar to the WT. The chemical denaturation experiment shows that point mutations at D372 and D227 slightly destabilize the initial transition of the protein. (Figure 4.2 and 4.3) However, unlike the chemical stability, the thermal denaturation experiment shows that the mutation at D216 slightly destabilizes the protein. (Figure 4.4 and table 4.3) It might be because the mutation at D372 destabilizes the initial phase and stabilizes the second phase. It also could be due to the differences between the chemical denaturation and thermal denaturation process of the protein (Serdyuk et al., 2007).

In conclusion, this chapter shows us an unexpected result that aspartate to asparagine mutants do not have a big effect on the stability of the protein. The chemical denaturation experiment (urea and guanidine HCl unfolding experiment) shows that ColA-P D372N was the least stable mutant and D216N was least stable in thermal denaturation experiment (DSC).

Fridd has shown that introduction of a single aspartate residue at these at these critical sites into Colicin N does not destabilize the protein (Fridd and Lakey, 2002). The results in this chapter shows that the aspartate to asparagine mutations on colicin A P domain also does not destabilize the protein. Thus, the side chain charge may not be the crucial in the colicin A P domain stability. Chapters three and four demonstrated that although both aspartate to alanine mutation and aspartate to asparagine mutation remove the negative charge from the surface of the protein, asparagine mutants do not have a big effect on the stability of ColA-P protein. According to the N-capping hypotheses for alpha helix stabilisation (Doig and Baldwin, 1995), both aspartate and asparagine can play the role of a hydrogen bond acceptor on the N-terminal end of a helix to stabilize the helix structure. Amino acids such as alanine do not have a hydrogen bond acceptor on their side chains, which make them a poor N-capping residue in helix stability. This hypothesis may explain the reason that aspartate to alanine mutations destabilize the protein while aspartate to asparagine mutations do not. In Chapter Five I will develop and test the N-capping hypothesis for the pH sensitivity of ColA-P.

Chapter Five

Testing the hypotheses that N-capping residues promote Colicin A pore-forming domain unfolding at low pH

The colicin A pore forming domain consists of 10 α-helices. It has been recognized that some amino acids show more preferences to be in α-helix, while others show preferences for other secondary structures such as β-sheet or β-turn. (Chou and Fasman, 1974) It has also been realized that four terminal amide hydrogen groups (-NH) do not have the hydrogen bonding partners when at the end of N-terminals of α-helices. (Doig and Baldwin, 1995) The residues located on the N-terminal side of these unpaired residues (N1, N2, N3 and N4) are called helix N-cap residues. (Figure 5.1) In the previous chapters it was shown that protonation of helix N-capping residues could explain the destabilisation of the colicin A-pore forming domain at low pH.

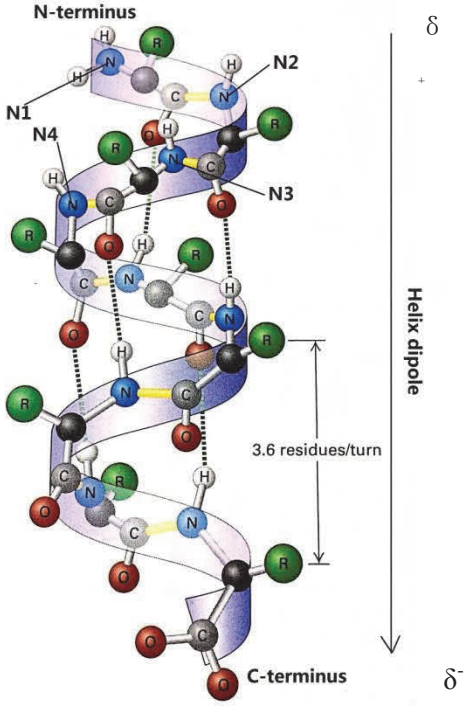


Figure 5.1: A ribbon diagram of an α helix structure. The polypeptide backbone is folded into an α-helix. Four hydrogen bonds between the carbonyl oxygen atom and the amide hydrogen atom are shown. Unpaired residues located at the N1-4 positions of α helix are shown. The electric dipole of a polypeptide is transmitted along an α helix with its positive pole at the N-terminus and negative pole at the C-terminus (Lodish et al., 2007).

A study on 215 alpha helices from 45 globular protein has shown that certain amino acid residues are favoured as the N-cap residues (Richardson and Richardson, 1988). In 1995, Doig and Baldwin have shown experimentally that, both aspartate and asparagines at the N-cap position stabilize the helix structure (Doig and Baldwin, 1995). As shown previously, some of the surface aspartates in ColA-P protein are located at the N-terminal side of the helix (Figure 5.1). Thus, N-cap hypothesis could explain why aspartate to alanine replacement destabilizes ColA-P while aspartate to asparagine replacement does not.

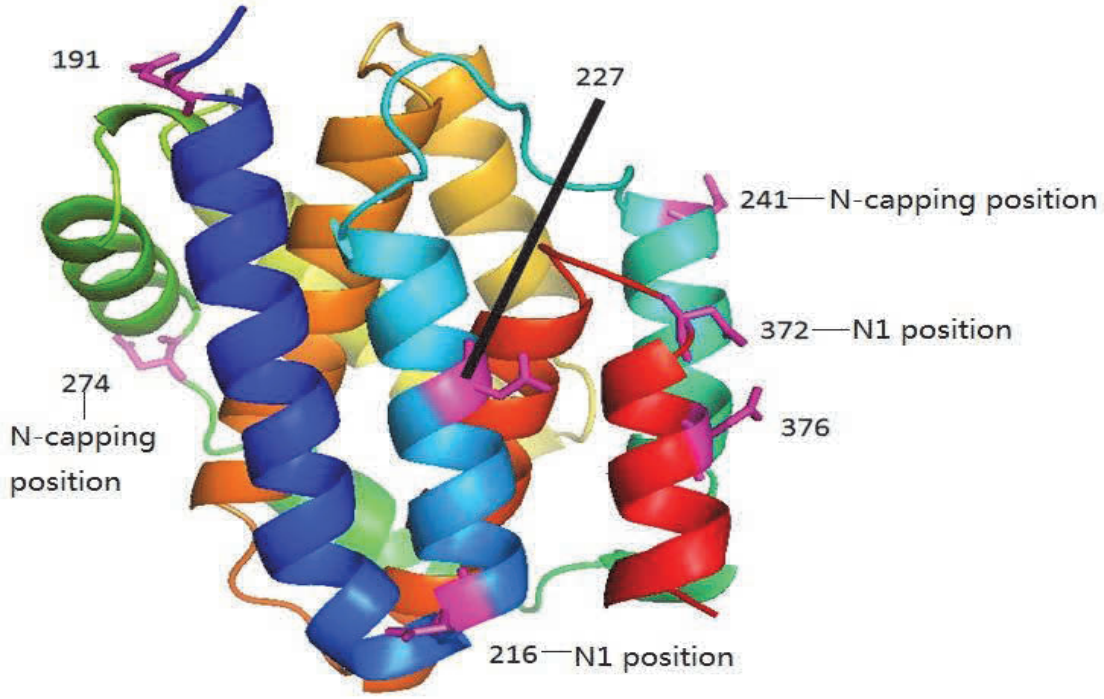


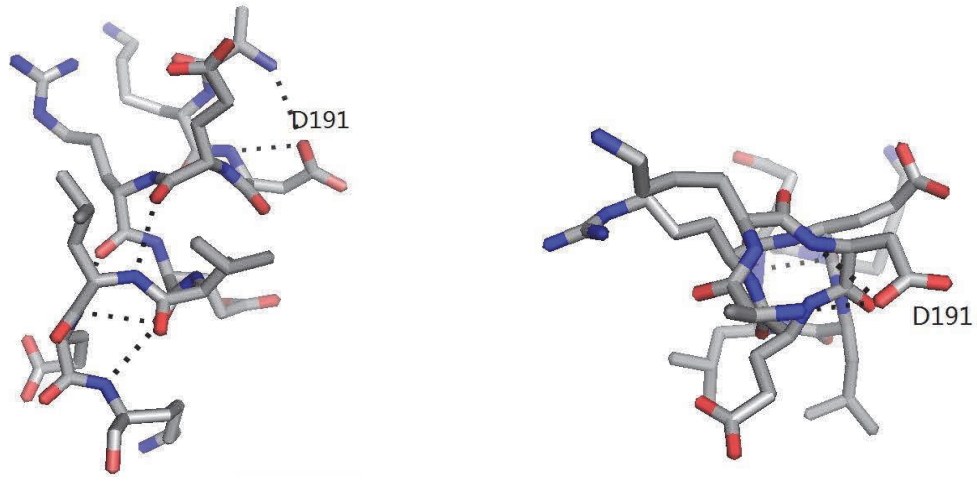
Figure 5.2: The position in a ribbon diagram of the seven surface aspartate residues of the pore-forming domain of colicin A not conserved in colicin N and their N capping positions

5.1 Aims

Although glutamate and glutamine are equally as good hydrogen bond acceptors as aspartate and asparagine, their side chains point away from the helix and the long distance between its methylene group (-CH_2-) and the backbone amide hydrogen groups (-NH-) makes it more flexible and unlikely to be a good hydrogen bond acceptor (Richardson and Richardson, 1988). This was confirmed in helical model peptides by Doig and Baldwin who showed that the free energy of stabilisation of glutamate and glutamine at the N cap position was -0.7 kcal.mol^{-1} and 2.5 kcal.mol^{-1} respectively compared to -1.6 kcal.mol^{-1} for aspartate, -1.7 kcal.mol^{-1} for asparagine and 0 kcal.mol^{-1} for alanine (Doig and Baldwin, 1995). They also investigated the preferences at the N1-N3 sites which are defined as the first, second and third residues in a helix and thus have unpaired backbone amide groups. The N-cap residue precedes N1 and unlike N1-N3 does not have alpha helical ψ angles (Gsponer et al., 2006, Capaldi et al., 2002, Lindeberg et al., 2000).

In order to examine if the N-capping hypothesis is fully suitable to explain the stability of the colicin A pore forming domain, aspartate to glutamate and aspartate to glutamine mutants were introduced and studied in my project.

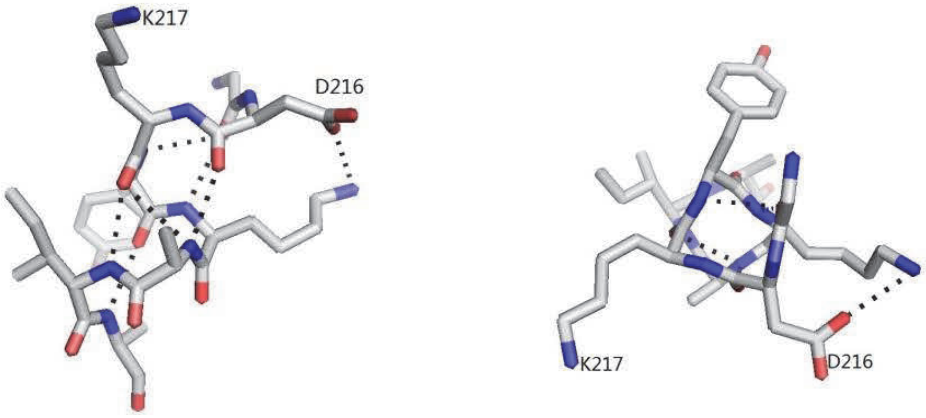
The site at D191 (Figure 5.2) is not located at the end of the N-terminal side of helix one in full length colicin A. However, it becomes one when helix one is cut in half to make the P-domain and faces towards the helix. Thus it is an artificial N-cap residue.



A B

Figure 5.3: Structure of the helix N-terminus near D191 and the hydrogen bonds within the region. A side view and B end on view of helix.

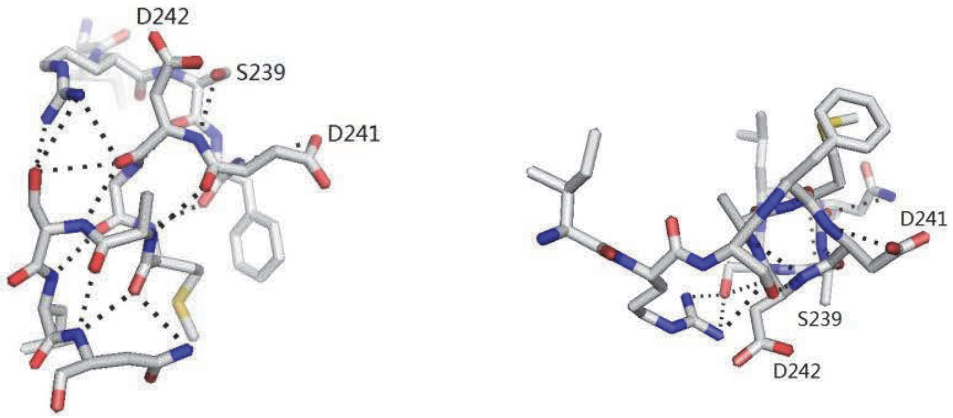
It appears that residue D216 (Figure 5.3) is located at the N1 position on the N-terminal side of helix two; however its side chain is pointing away from the unpaired amide hydrogens (-NH). Besides, there is a lysine at N2 position which should be disfavoured and it might be balancing D216.



A B

Figure 5.4: Structure of the N-terminal side near point site D216 and the hydrogen bonds within the region. A side view and B end on view of helix.

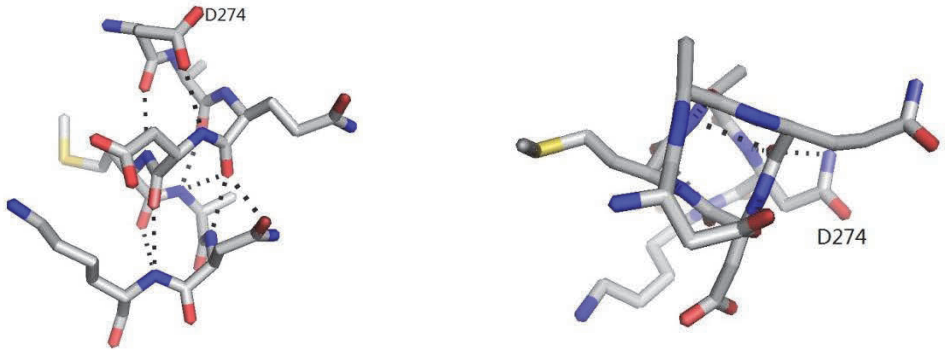
The residue D241 is located at the N-capping position on helix three; however it points away from the helix three and forms a ring at the top of the helix with residues S239 and D242. Actually, S239 points towards the helix and is closer to the N-terminus of the helix and might have a role in stabilizing the protein. (Figure 5.5)



A B

Figure 5.5: Structure of the N-terminal side near point site D241 and the hydrogen bonds within the region. A side view and B end on view of helix.

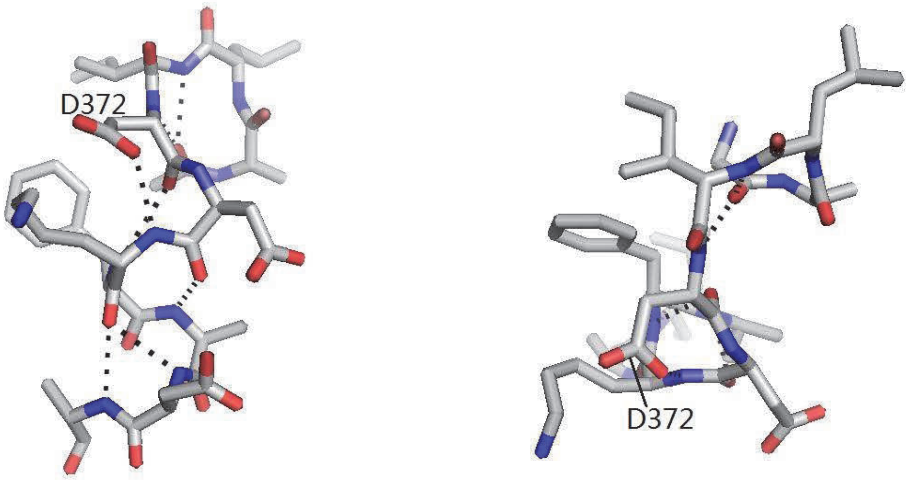
D274 is a perfect N-capping residue at the top of the N-terminal side of helix five. (Figure 5.6)



A B

Figure 5.6: Structure of the N-terminal side near point site D274 and the hydrogen bonds within the region. A side view and B end on view of helix.

Point site D372 is located at N1 position on N-terminal side of helix ten. (Figure 5.7)



A B

Figure 5.7: Structure of the N-terminal side near D372 and the hydrogen bonds within the region. A side view and B end on view of helix.

In light of the above information aspartate to glutamate and glutamine substitutions were made at positions 216, 241, 274 and 372. In order to study the effect of the ring located at the N terminal of the helix around 241, numerous substitutions including S239A and S239D-D241A were also prepared in the project.

5.2 Results

Urea unfolding curves of Colicin A pore-forming domain Aspartate acid to Glutamine and Glutamate mutants

To assess the N capping preferences of asparagine, alanine, glutamate and glutamine in helix two of colicin A-P domain, the intrinsic fluorescence of ColA-P D216A, D216N, D216E and D216Q in a series of concentrations of urea (MOPS pH 7.0) was measured. The barycentric mean of each emission spectrum was calculated and plotted to yield a urea unfolding curve. (Figure 5.8)

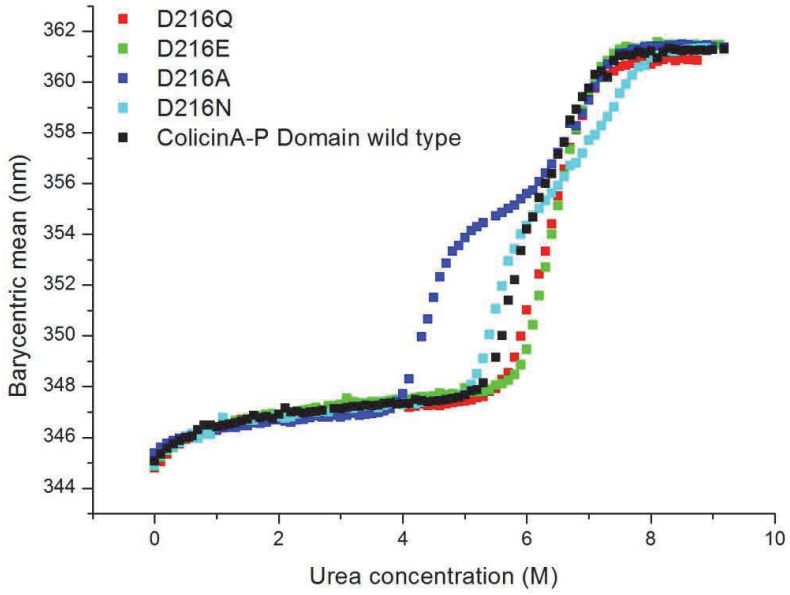


Figure 5.8: Biphasic urea unfolding transitions of the ColA-P D216A (Blue Curve), D216E (Green curve) D216N (Cyan curve), D216Q (Red curve) and wild type (Black curve) at pH 7.0 in MOPS buffer.

Figure 5.8 shows that the stability of the initial unfolding transition of the D216 site point mutations is: D216E = D216Q > D216N > D216A at pH 7.0. The asparagine mutation is the only one to stabilize the later phase.

To assess the N capping preferences of asparagine, alanine, glutamate and glutamine in helix five of colicin A-P domain, the intrinsic fluorescence of ColA-P D274A, D274N, D274E and D274Q in a series of concentrations of urea (MOPS pH 7.0) was measured. (Figure 5.9)

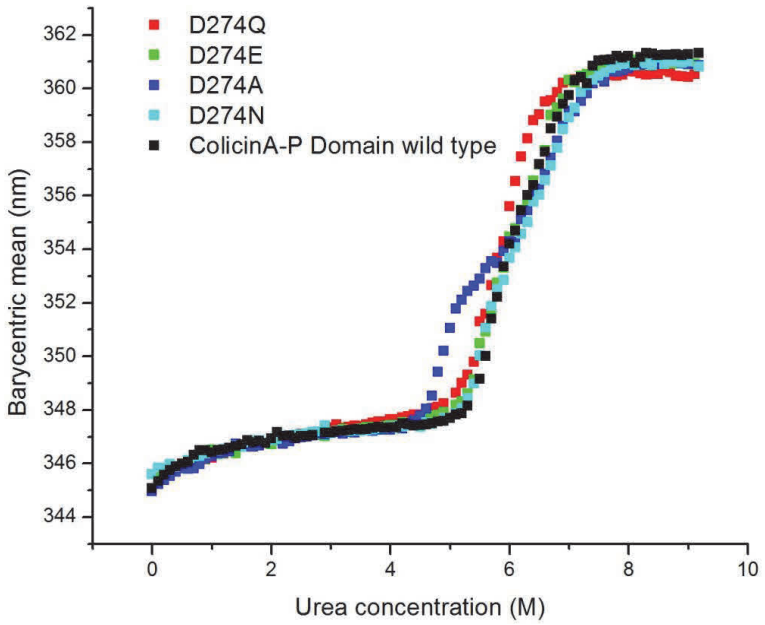


Figure 5.9 Biphasic urea unfolding transitions of the ColA-P D274A (Blue Curve), D274E (Green curve) D274N (Cyan curve), D274Q (Red curve) and wild type (Black curve) at pH 7.0 in MOPS buffer.

Figure 5.9 shows that the alanine mutation at D274 destabilized the initial transition; the glutamine mutation at this point site destabilized the later phase; and asparagine and alanine mutants slightly stabilize the later phase.

To assess the N capping preferences of asparagine, alanine, glutamate and glutamine in helix ten of colicin A-P domain, the intrinsic fluorescence of ColA-P D372A, D372N, D372E and D372Q in a series of concentrations of urea (MOPS pH 7.0) buffer was measured as above. (Figure 5.10)

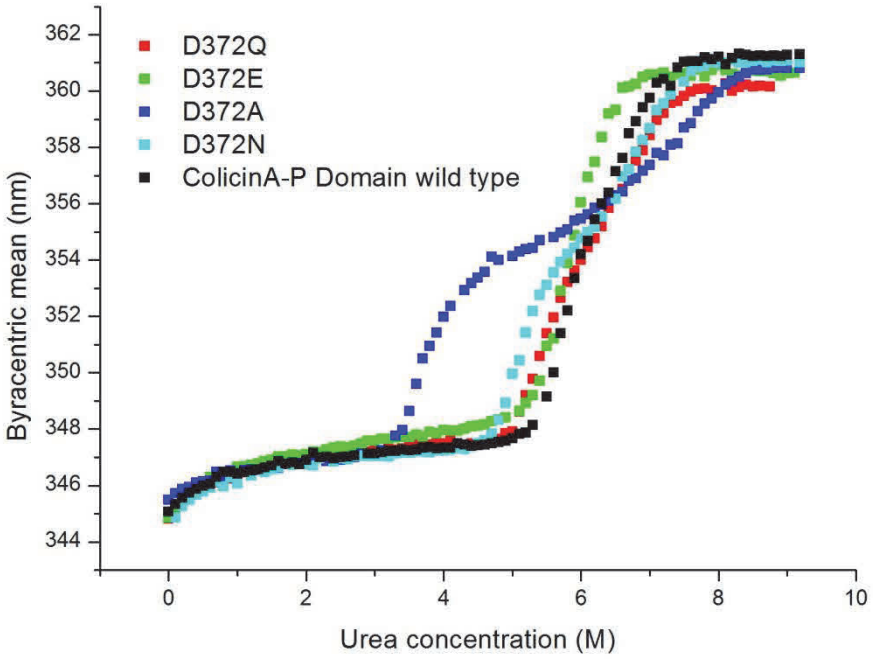


Figure 5.10: Biphasic urea unfolding transitions of the ColA-P D372A (Blue Curve), D372E (Green curve) D372N (Cyan curve), D372Q (Red curve) and wild type (Black curve) at pH7.0 in MOPS buffer.

Figure 5.10 shows that the stability of the initial unfolding transitions of the D372 site point mutations is: D372E > D372Q > D372N > D372A at pH 7.0. The glutamate mutant destabilizes the later phase and the alanine mutant destabilizes the early phase.

To assess the effect of the ring located at the N terminal in the helix ten of colicin A-P domain, the intrinsic fluorescence of ColA-P D241A, D241N, S239D and S239D-D241A in a series of concentrations of urea (MOPS pH 7.0) was measured. (Figure 5.11)

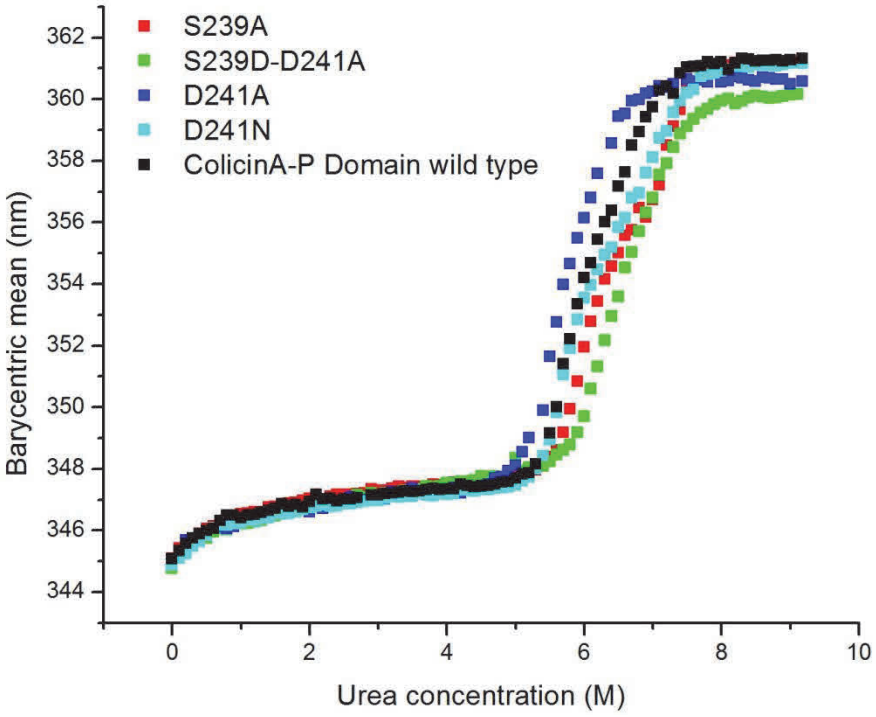


Figure 5.11: Biphasic urea unfolding transitions of the ColA-P D241A (Blue Curve), S239D-D241A (Green curve) D241N (Cyan curve), S239A (Red curve) and wild type (Black curve) at pH 7.0 in MOPS buffer.

Figure 5.11 shows that the stability of the initial unfolding transitions of point mutations and double mutation of ColA-P at N-capping position on helix three is: ColA-P S239D-D241A > S239A > D241N > D241A.

To assess the N capping preferences of asparagine, alanine, glutamate and glutamine in helix two of colicin A-P domain at low pH, the intrinsic fluorescence of ColA-P D216A, D216N, D216E and D216Q in a series of concentrations of urea at pH 3.0 (sodium citrate) was measured. (Figure 5.12)

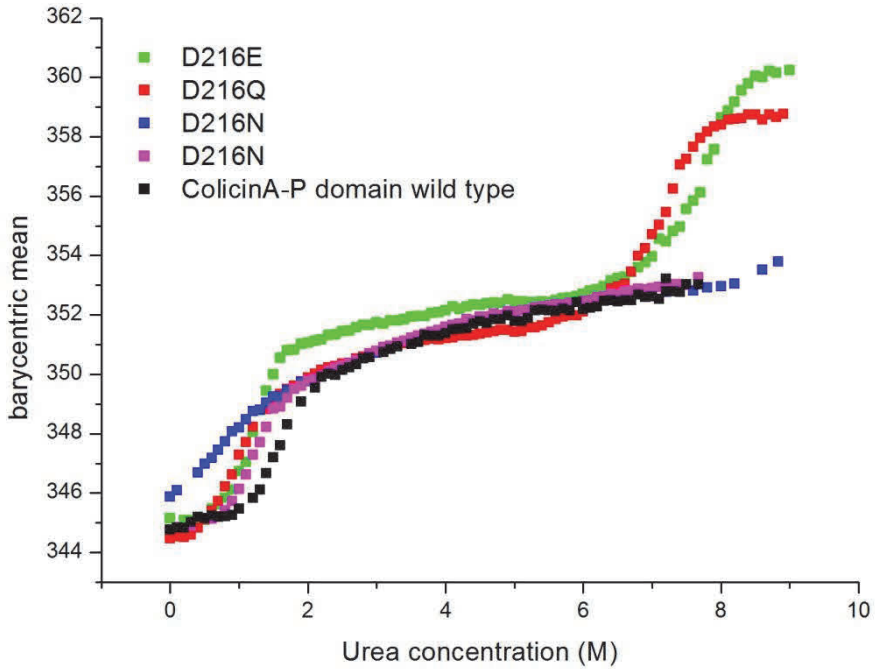


Figure 5.12: The urea unfolding curve of the ColA-P D216A (Blue Curve), D216E (Green curve) D216N (Magenta curve), D216Q (Red curve) and wild type (Black curve) at pH 3.0 in sodium citrate buffer.

Figure 5.12 shows that the stability of the initial unfolding transitions of the D216 site point mutations is: D216N > D216Q > D216E > D216A at pH 3.0, however the initial unfolding phase of D216N, D216Q and D216E is very similar. Most significantly the glutamate and glutamine mutants destabilize the later transition.

To assess the N capping preferences of asparagine, alanine, glutamate and glutamine in helix five of colicin A-P domain at low pH, the intrinsic fluorescence of ColA-P D274A, D274N, D274E and D274Q in a series of concentrations of urea (sodium citrate pH 3.0) was measured. (Figure 5.13)

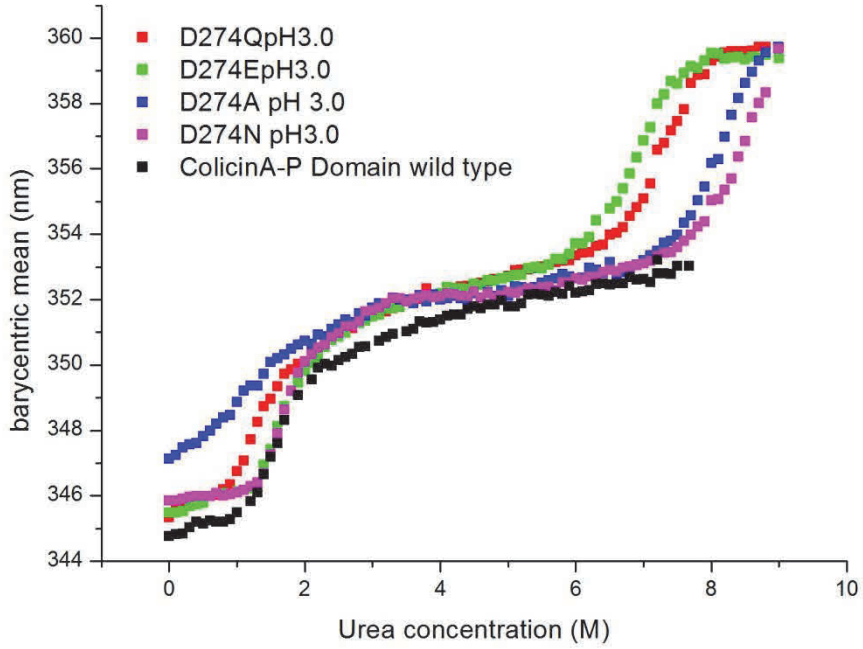


Figure 5.13: The urea unfolding curve of the ColA-P D274A (Blue Curve), D274E (Green curve) D274N (Magenta curve), D274Q (Red curve) and wild type (Black curve) at pH 3.0 in sodium citrate buffer.

Figure 5.13 shows that the stability of the initial unfolding transitions of the D274 point mutations is: D274N > D274E > D274Q > D274A at pH 3.0 and stability of the later unfolding transitions is: D274N > D274A > D274Q > D274E. Again the glutamate and glutamine mutations clearly destabilise the second transition.

To assess the N capping preferences of asparagine, alanine, glutamate and glutamine in helix ten of colicin A-P domain at low pH, the intrinsic fluorescence of ColA-P D372A, D372N, D372E and D372Q in a series of concentrations of urea (sodium citrate pH 3.0) was measured. (Figure 5.14)

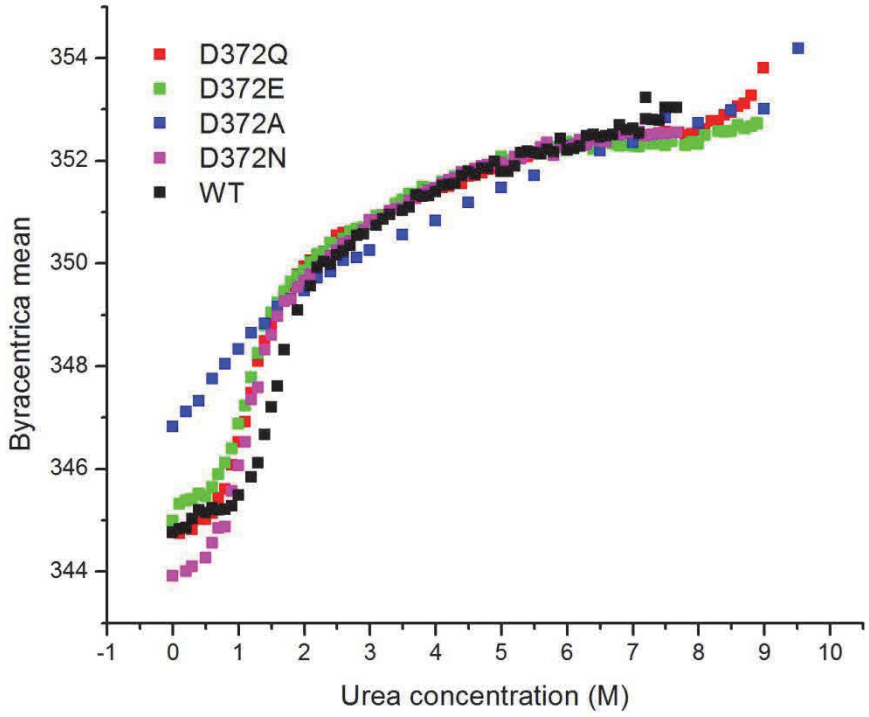


Figure 5.14: The urea unfolding curve of the ColA-P D372A (Blue Curve), D372E (Green curve) D372N (Magenta curve), D372Q (Red curve) and wild type (Black curve) at pH 3.0 in sodium citrate buffer.

Figure 5.14 shows that the stability of the initial unfolding transitions phases of the D372 site point mutations is: D372N > D372Q > D372E > D372A. Although the second transitions are almost suppressed at low pH, the start of the second transition can be observed in D372Q. The figure shows that the later phase of glutamate mutants is more stable than the glutamine mutant.

To assess the effect of the ring located at the N terminal in the helix ten of colicin A-P domain at low pH, the intrinsic fluorescence of ColA-P S239D and S239D-D241A in a series of concentrations of urea (sodium citrate pH 3.0) was measured. (Figure 5.15)

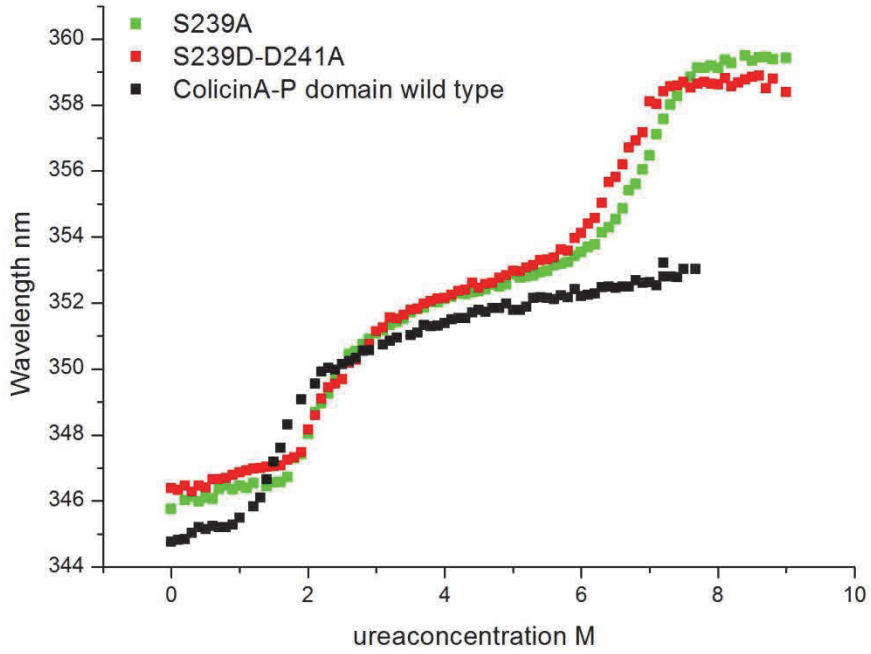


Figure 5.15: Biphasic urea unfolding transitions of the S239D-D241A (Green curve), S239A (Red curve) and wild type (Black curve) at pH 3.0 in sodium citrate buffer.

Figure 5.15 shows that the stability of the initial unfolding transitions of ColA-P S239A and ColA-P S239D D241A at pH 3.0 is similar to wild type and that the second transition is destabilised significantly.

The Gibbs free energy of folding of colicin A pore-forming domain glutamine and glutamate mutants

The Gibbs free energy for each mutation in urea solution at pH 7.0 and pH 3.0 were calculated using a non-linear least squares fitting procedure in Microcal Origin that was kindly supplied by Dr David Hough, Bath University, UK. (See methods) The Gibbs free energy of point mutation and double mutation of ColA-P at N-capping

position in the urea unfolding process were calculated and shown in tables 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7 and 5.8.

Table 5.1. The Gibbs free energy of Biphasic unfolding for point mutations at D216 of colicin A- pore forming domain determined by urea unfolding at pH 7.0

Mutation	$\Delta G1$ Kcal.mol^{-1}	$\Delta G2$ Kcal.mol^{-1}	$\Delta G1+\Delta G2$ Kcal.mol^{-1}	First transition (urea concentration M)	Second transition (urea concentration M)
D216Q	nd	nd	12.38	5.3-7.4	nd
D216E	nd	nd	14.79	5.7-7.5	nd
D216A	14.70	12.70	23.73	3.8-5.0	6.1-7.5
D216N	21.42	18.62	28.15	5.0-6.1	6.8-8.0
WT	21.83	10.45	32.38	5.2-6.0	6.1-7.5

nd = not determined, as the transitions are not separated at this pH

Table 5.1 shows that the $\Delta G1$ of the wild type $>\Delta G1$ of D216N $>\Delta G1$ of D216A at pH 7.0. For ColA-P D216E and D216Q, only a single transition is visible at pH 7.0.

Table 5.2. The Gibbs free energy of Biphasic unfolding for point mutation at D274 of colicin A- pore forming domain determined by urea unfolding at pH 7.0

Mutation	$\Delta G1$ Kcal.mol^{-1}	$\Delta G2$ Kcal.mol^{-1}	$\Delta G1+\Delta G2$ Kcal.mol^{-1}	First transition (Urea concentration M)	Second transition (Urea concentration M)
D274Q	15.14	12.04	27.28	4.9-5.8	5.9-6.9
D274E	15.81	16.22	32.03	5.1-6.0	6.1-7.2
D274A	8.05	10.71	18.76	4.4-5.1	6.1-7.5
D274N	25.47	12.83	38.30	5.1-5.8	5.9-7.5
WT	21.83	10.45	32.38	5.2-6.0	6.1-7.5

Table 5.2 shows that the $\Delta G1$ of ColA-P D274N $>$ $\Delta G1$ of the wild type $>\Delta G1$ of D274E $>$ $\Delta G1$ of D274Q $>\Delta G1$ of D274A at pH 7.0. However, $\Delta G1$ of ColA-P D274E and D274Q are very close to each other. It reveals that ColA-P D274E and D274Q needs more energy to unfold to the intermediate stage than the wild type and needs less energy to form its intermediate stage than D274N at low pH. The result reveals that ColA-P D274N is more stable than D274E, D274Q and its wild type.

Table 5.3. The Gibbs free energy of biphasic unfolding for site point mutation at D372216 of colicin A- pore forming domain detected by urea unfolding at pH 7.0

Mutation	$\Delta G1$ Kcal.mol^{-1}	$\Delta G2$ Kcal.mol^{-1}	$\Delta G1+\Delta G2$ Kcal.mol^{-1}	First transition (urea concentration M)	Second transition (urea concentration M)
D372Q	20.70	8.54	29.24	5.0-5.8	5.9-7.6
D372E	16.87	11.72	28.59	5.1-5.8	5.9-6.7
D372A	13.36	11.61	24.96	3.2-4.3	5.9-8.5
D372N	18.62	10.98	29.60	4.6-6.0	6.2-7.7
WT	21.83	10.45	32.38	5.2-6.0	6.1-7.5

Table 5.3 shows that the $\Delta G1$ of the wild type $>\Delta G1$ of D372Q $>\Delta G1$ of D372N $>\Delta G1$ of D372E $>\Delta G1$ of D372A at pH7.0.

Table 5.4. The Gibbs free energy of biphasic unfolding for point mutations at D372 of colicin A- pore forming domain determined by urea unfolding at pH 7.0

Mutation	$\Delta G1$ Kcal.mol^{-1}	$\Delta G2$ Kcal.mol^{-1}	$\Delta G1+\Delta G2$ Kcal.mol^{-1}	First transition (Gdn concentration M)	Second transition (Gdn concentration M)
D241A	8.4	11.12	19.52	4.9-5.7	5.7-7.0
S239D	21.38	9.10	30.48	5.4-6.5	6.8-7.8
S239D-D241A	17.61	10.31	27.92	5.7-6.6	5.7-7.9
D241N	22.28	14.29	36.27	5.2-6.0	6.1-7.7
WT	21.83	10.45	32.38	5.2-6.0	6.1-7.5

nd = not determined, as only a single transition is visible at this pH

Table 5.4 shows that the $\Delta G1$ of ColA-P D241N > $\Delta G1$ of the wild type > $\Delta G1$ of S239D > > $\Delta G1$ of S239DD241A > $\Delta G1$ of D241A at pH3.0. It means S239D-D241A needs less energy to unfold to the intermediate stage than D239D or D241N at low pH. The result reveals that the double mutant needs more energy to unfold to its intermediate stage than ColA-P D241A and needs less energy than S239D or D241N.

Table 5.5. The Gibbs free energy of Biphasic unfolding for point mutation at D216 of colicin A- pore forming domain determined by urea unfolding at pH 3.0

Mutation	*ΔG1* Kcal.mol^{-1}	*ΔG2* Kcal.mol^{-1}	*ΔG1+ ΔG2* Kcal.mol^{-1}	First transition (Gdn concentration M)	Second transition (Gdn concentration M)
D216Q	1.72	12.77	14.49	0.3-1.8	6.0-8.0
D216E	3.80	11.70	15.50	0.6-1.7	6.6-8.6
D216A	nd	nd	nd	nd	nd
D216N	3.33	nd	nd	0.5-1.8	nd
WT	4.66	nd	nd	1.0-2.2	nd

nd = not determined, two separate r transitions are not defined at this pH

D216A proceeds to the intermediate stage in the absence of urea, thus the *ΔG1* of D216A at pH 3.0 can be regarded as 0. It means that the *ΔG1* of the wild type > *ΔG1* of D216E > *ΔG1* of D216N > *ΔG1* of D216Q at pH 3.0 >*ΔG1*of D216A at pH 3.0.

Table 5.6. The Gibbs free energy of biphasic unfolding for site point mutation at D274 of colicin A- pore forming domain detected by urea unfolding at pH 3.0

Mutation	*ΔG1* Kcal.mol^{-1}	*ΔG2* Kcal.mol^{-1}	*ΔG1+ ΔG2* Kcal.mol^{-1}	First transition (urea concentration M)	Second transition (urea concentration M)
D274Q	3.16	14.17	17.33	0.8-2.4	6.4-8.2
D274E	4.64	12.85	17.49	1.2-2.7	5.9-7.9
D274A	nd	14.88	nd	nd	7.0-8.99
D274N	3.65	6.61	10.26	1.2-2.9	7.3-9.0
WT	4.66	nd	nd	1.0-2.2	nd

nd = not determined, two separate transitions are not defined at this pH

*The*ΔG2* of D274N is inaccurate because the second transition was not completed.

As D274A like D216A proceeds to the intermediate stage in the absence of urea, thus the *ΔG1* of D274A at pH 3.0 can be regarded as 0. It means that the *ΔG1* of the wild

type > *ΔG1* of D274E > *ΔG1* of D274N > *ΔG1* of D274Q at pH 3.0 >*ΔG1*of D216A at pH 3.0.

Table 5.7. The Gibbs free energy of Biphasic unfolding for point mutation at D372of colicin A- pore forming domain detected by urea unfolding at pH 3.0

Mutation	*ΔG1* Kcal.mol^{-1}	*ΔG2* Kcal.mol^{-1}	*ΔG1+ ΔG2* Kcal.mol^{-1}	First transition (urea concentration M)	Second transition (urea concentration M)
D372Q	2.27	nd	nd	0.4-2.0	nd
D372E	2.74	nd	nd	0.5-2.1	nd
D372A	nd	nd	nd	nd	nd
D372N	2.66	nd	nd	0.6-1.7	nd
WT	4.66	nd	nd	1.0-2.2	nd

nd = not determined, two separate transitions are not defined at this pH

Again, D372A forms to the intermediate stage without urea, thus the *ΔG1* of D274A at pH 3.0 can be regarded as 0. It means that the *ΔG1* of the wild type > *ΔG1* of D372E > *ΔG1* of D372N > *ΔG1* of D372Q at pH 3.0 >*ΔG1*of D372A at pH 3.0.

Table 5.8. The Gibbs free energy of Biphasic unfolding for point mutation at D372 of colicin A- pore forming domain determined by urea unfolding at pH 3.0

Mutation	*ΔG1* Kcal.mol^{-1}	*ΔG2* Kcal.mol^{-1}	*ΔG1+ ΔG2* Kcal.mol^{-1}	First transition (urea concentration M)	Second transition (urea concentration M)
D241A	--	--	--	--	--
S239D	4.52	12.90	17.42	1.8-3.3	5.8-7.3
S239DD241A	5.68	13.69	19.39	1.6-3.0	6.0-7.7
D241N	--	--	--	--	--
WT	4.66	nd	nd	1.0-2.2	nd

nd = not determined, two separate transitions are not defined at this pH--data not collected

Table 5.8 shows that the *ΔG1* of S239D-D241A > *ΔG1* of the wild type > *ΔG1* of S239A at pH 3.0.

Circular Dichroism studies of Colicin A pore-forming domain glutamine and glutamic acid mutants

According to the N-capping hypotheses, it was suggested that glutamine can actively stabilize non-helical structure (Chakrabartty et al., 1991, Richardson and Richardson, 1988), glutamate is also not a strongly N-capping preference amino acid, and however the negatively charged side chains make it have a higher N-capping preference than glutamine. (Doig and Baldwin, 1995)

In order to compare the stability of the secondary structure of ColA-P Ala, Asn, Glu and Gln mutants, the far UV CD spectra of the proteins was measured at pH 8.0 and pH 3.0. The molar CD at 222 nm was used to represent the helix content of ColA-P WT and its mutants (Serdyuk et al., 2007). The helix content of each mutant was measured at pH 8.0 where the protein is negatively charged (the pI of ColA-P domain is 5.4, so at pH 8.0 it is negatively charged), and at pH 3.0, where the protein is positively charged. The results are shown in table 5.10.

Name of the mutants	$[\Delta\varepsilon]_{222\ nm}$ at pH8.0	$[\Delta\varepsilon]_{222\ nm}$ at pH 3.0
ColA-P D216A	-8.1403	-6.35053
ColA-P D216N	-7.22308	-5.45104
ColA-P D216E	-8.23084	-7.79449
ColA-P D216Q	-8.04419	-6.84214
ColA-P D274A	-7.94277	-6.30529
ColA-P D274N	-7.69998	-6.80967
ColA-P D274E	-8.08672	-6.41927
ColA-P D274Q	-7.09382	-6. 78091
ColA-P D372A	-6.74827	-7.95262
ColA-P D372N	-6.51328	-6.6824
ColA-P D372E	-8.69675	-6.58196
ColA-P D372Q	-7.09382	-7.28921
ColA-P S239D	-7.87742	-6.50154
ColA-P S239DD241A	-7.35595	-6.88023
ColA-P WT	-7.53036	-6.89326

Table 5.10 shows that at pH 8.0 $[\Delta\varepsilon]_{222\ nm}$ of glutamate mutants > $[\Delta\varepsilon]_{222\ nm}$ of alanine mutants > $[\Delta\varepsilon]_{222\ nm}$ of glutamine mutants> $[\Delta\varepsilon]_{222\ nm}$ of asparagine mutants.

$[\Delta\varepsilon]_{222\ nm}$ values of ColA-P alanine mutants, asparagine mutants, glutamate mutants and glutamine mutants at 216, 274 and 372 positions at pH 3.0 are compared and shown in Figure 5.12

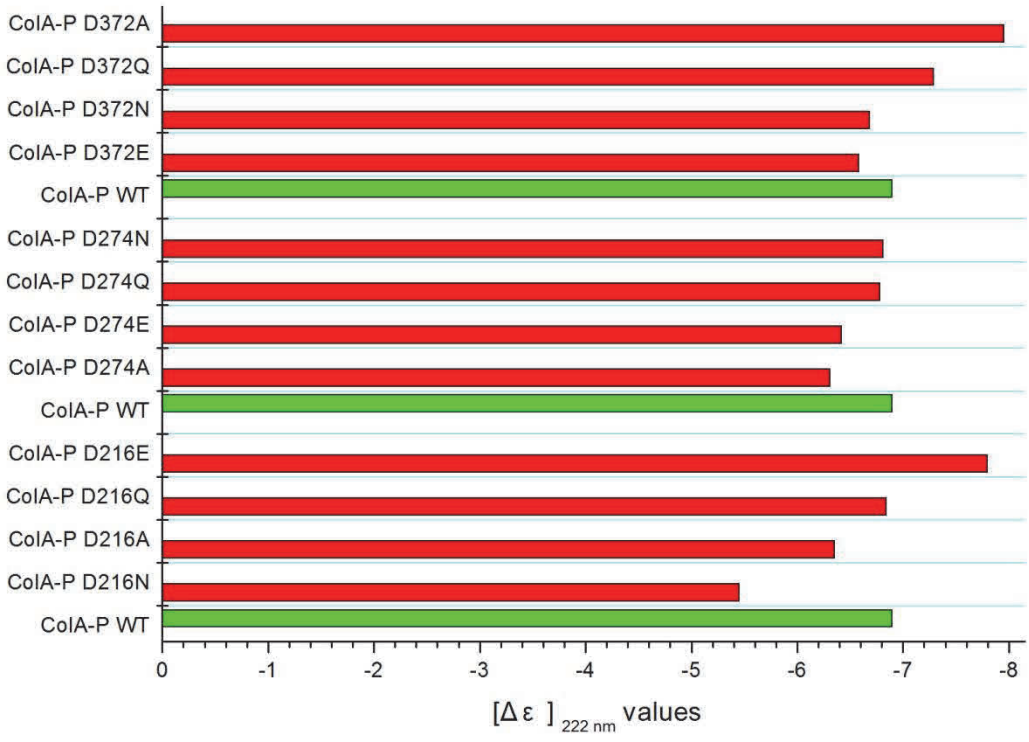
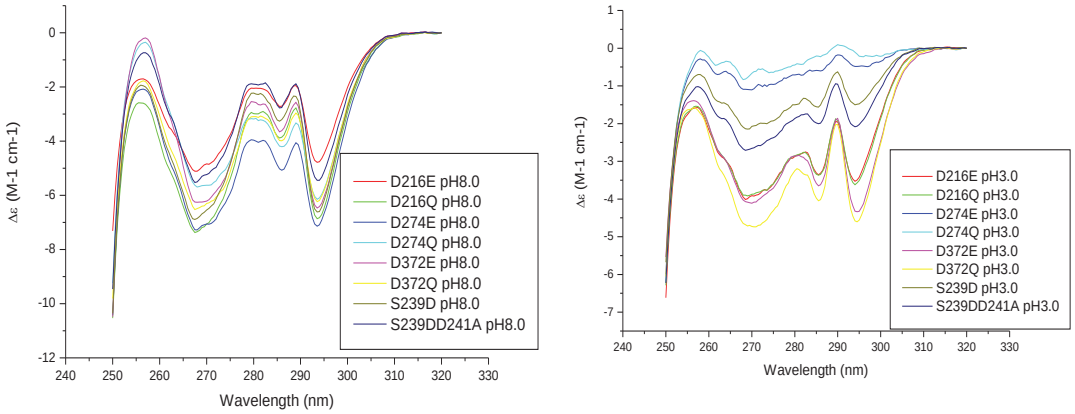


Figure 5.15: $[\Delta\varepsilon]_{222\ nm}$ values of ColA-P alanine mutants, asparagine mutants, glutamate mutants and glutamine mutants at 216, 274 and 372 positions at pH 3.0. At this wavelength the signal results almost entirely from the alpha helix content

The near UV CD spectra of ColA-P glutamate and glutamine mutants were measured at pH 8.0 and pH 3.0. (Figure 5.12)



A) B)

Figure 5.16: A) The Near UV CD spectra of ColA-P D216E (Red line), D216Q (Green line), D274E (Blue line), D274Q (Cyan line), D372E (Magenta line), D372Q (Yellow line), S239D (Dark yellow line) and S239DD241A (Navy line) at pH 8.0; B) The Near UV CD spectra of ColA-P D216E (Red line), D216Q (Green line), D274E (Blue line), D274Q (Cyan line), D372E (Magenta line), D372Q (Yellow line), S239D (Dark yellow line) and S239DD241A (Navy line) at pH 3.0

The Far UV CD spectra show that the glutamate mutants are less stable than the glutamine mutants at D216 and D372 position at pH 8.0. ColA-P D274E is more stable than D274Q at this pH. The Far UV CD spectra also show that the glutamine mutants are less stable than the glutamate mutants at D274 and D372 position at pH 3.0. ColA-P D216Q is more stable than D216E at low pH.

The thermal transition temperature and the calorimetric transition enthalpy

In order to compare the thermal stability between ColA-P protein and their point mutations with two different N capping residues, the protein samples were studied under the same conditions by DSC as before. Characteristic DSC scans for colicin A P-domain aspartate and asparagine mutants are shown in figure 5.16.

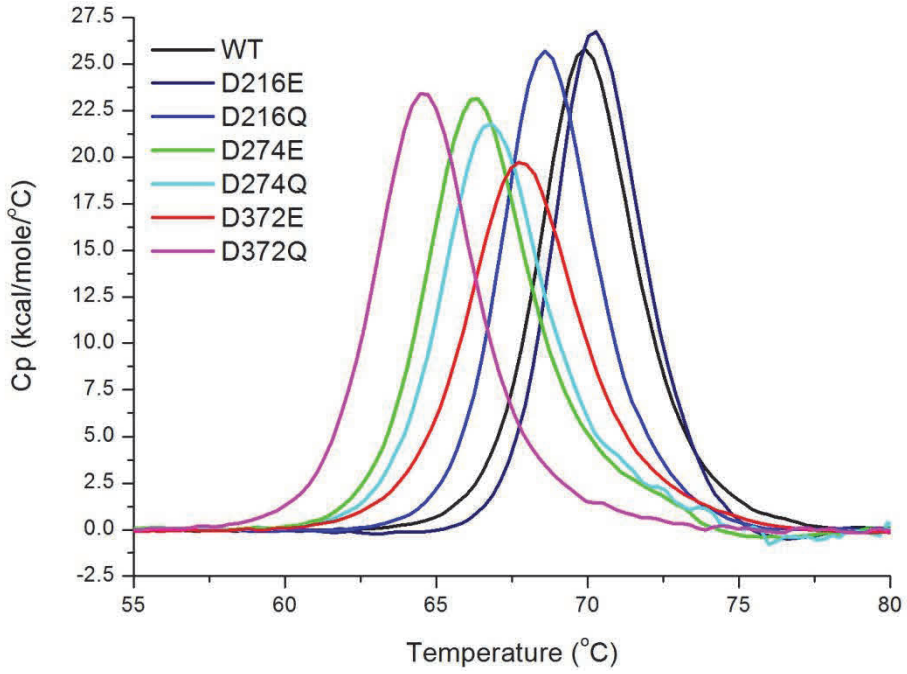


Figure 5.17: Characteristic heat-capacity curves for ColA-P WT (Black line), D216E (Navy line) D216Q (Blue line), D274E (Green line), D274Q (Cyan line), D372E (Red line), D372Q (Magenta line) with buffer baselines subtracted and normalized for protein concentrations

Each of the ColA-P glutamate and glutamine mutants showed only one denaturation peak within the temperature region of 25 to 90 ^{o}C.

Although it is clear from chapter three that the enthalpy does not follow a simple model, temperature unfolding was considered firstly as a two-state transition and by fitting the data the calorimetric transition enthalpy was calculated. The denaturation temperature, calorimetric enthalpy and Van't Hoff enthalpy are shown in table 5.11 with data represented from previous chapters for comparison.

Mutation	Thermal transition temperature Tm (oC) ±error	*ΔH* (kcal/mol) ±error	*ΔHv* (kcal/mol) ±error
D216N	68.41±0.035	93.12±0.67	208.25±4.51
D216E	70.35±0.012	102.0±0.668	246.4±2.06
D216A	60.76±0.018	80.77±1.637	178.55±1.76
D216Q	68.72±0.096	106.1±0.527	222.8±1.40
D274N	69.26±0.026	103.0±1.39	219.95±2.90
D274E	66.44±0.019	107.0±0.914	193.9±2.08
D274A	64.64±0.018	92.49±0.826	196.55±2.22
D274Q	66.94±0.017	100.9±0.810	194.4±1.96
D372N	69.88±0.076	105.3±0.013	239.6±2.09
D372E	67.89±0.012	97.79±0.505	183.4±1.19
D372A	57.46±0.012	71.82±0.337	179.8±1.18
D372Q	64.65±0.012	103.6±0.606	202.2±1.49
S239D	67.73±0.021	86.68±0.872	202.0±2.56
S239DD241A	68.50±0.028	106.2±1.64	234.6±4.59
WT	70.01±0.087	107.8±0.491	220.7±0.655

The values of Δ*H* and Δ*Hv* of ColA-P glutamate and glutamine mutants are different, which means the thermal denaturation transition is not a two-state transition (Serdyuk et al., 2007).

The T_m of ColA-P D216E > T_m of ColA-P D216Q > T_m of ColA-P D372E > T_m of ColA-P D274Q > T_m of ColA-P D274E > T_m of ColA-P D372Q indicating that the melting temperature of the glutamine mutants at positions D216 and D372 position is lower than glutamate mutants. ColA-P D274Q is slightly more stable than D274E. ColA-P D372E shows an obviously broader peak, which might due to the separation of the biphasic thermal transitions. However it still displays the overall pattern of $\Delta H = 0.5\,\Delta Hv$ (discussed in chapter three)

5.3 Discussion

The stability of Glutamine mutants and Glutamate mutants

The urea unfolding experiment shows that the initial unfolding transitions of glutamine and glutamate mutants at positions D216 and D372 are more stable than alanine mutants and asparagine mutants at neutral pH (Figure 5.8, and 5.10) They are more stable than the alanine mutants but less stable than asparagine mutants at pH 3.0. It means that, compared with the asparagine mutation, glutamine and glutamate mutations at D216 and D372 are more sensitive to pH.(Figure 5.12, 5.13 and 5.14) The initial unfolding transitions of glutamate and glutamine mutants at position D274 are more stable than alanine mutants and resemble the asparagine mutants at neutral pH. (Figure 5.9) The glutamate mutation is less stable than glutamine and asparagine mutants and more stable than alanine mutants at pH 3.0. It means that the glutamate mutation at D274 is more sensitive to the pH than glutamine and asparagine mutations.

The results in chapter three show that point mutations at N-capping position such as D372, D216 and D227 destabilize the protein more than the point mutations at D241 D274 and D376. In this chapter, it was shown that the first transition of the chemical denaturation of D372 and D216 aspartate to asparagine, aspartate to glutamate and aspartate to glutamine mutations started at 0.5 M urea, and the D274 aspartate to asparagine, aspartate to glutamate and aspartate to glutamine mutations started its initial transition at 1.2 M urea, which means the point mutation at D372 and D216 destabilize the protein more than the mutations at D274 at low pH. (Tables 5.6, 5.7 and 5.8)

The thermal denaturation experiment performed by DSC shows that except Cola-P D216E, the melting temperature of other glutamate and glutamine mutants is lower

than the wild type protein and asparagine mutants and higher than the alanine mutants.

In conclusion, the glutamate and glutamine mutations at N-capping position of the helix of colicin A pore-forming domain are less stable than its alanine mutants, but at low pH they are more stable than its asparagine mutations. Most of the glutamate and glutamine mutations reduce the Gibbs free energy for the protein to form to its molten globule state, which means the glutamate and glutamine mutations enhance the pH sensitivity of the protein (Table 5.1, 5.2 5.3, 5.4, 5.5, 5.6, 5.7 and 5.8).

N-capping hypothesis

According to the earlier studies of the N-capping theory, the N-capping preference residues such as aspartate and asparagine significantly stabilize a small isolated helical peptide (Doig and Baldwin, 1995). However, for a large protein such as colicin A pore forming domain the situation is much more complex than a short peptide. Thus my results may not fully agree with the N-capping theory.

The urea unfolding result in this chapter indicates the stability of the initial phase of Asp to Asn, Asp to Glu and Asp to Gln is very similar to each other. Although, Glu and Gln at N-capping position are supposed to destabilize the helical peptide (Doig and Baldwin, 1995), it seems that they do not significantly destabilize the colicin A pore forming domain, which means that a simple N-capping hypotheses may not be applicable in the full length proteins.

The $[\Delta\varepsilon]_{222\ nm}$ at pH8.0 and pH 3.0 (table 5.10) shows that the secondary structure of the glutamate mutants and glutamine mutants are almost as the same as the wild type protein, and confirm that a simple N-capping hypothesis may not be suitable to use in the full length proteins.

However, the thermal denaturation experiment performed by DSC indicates that the protein was destabilized by aspartate to glutamate and glutamine mutations and thus indicates that the geometry of glutamate and glutamine may render them less effective N-capping residues than aspartate and asparagine.

Earlier studies of the N-capping hypothesis suggested that alanine mutations should be more stable than the glutamine mutation and less stable than the glutamate mutation, whilst asparagine has the most N-capping preference (Doig and Baldwin, 1995). However, the definition of whether our mutations are at the N-Cap or the N1 position is sometimes uncertain. A later N-capping study suggests that alanine has a high N1 preference and asparagine at N-capping position destabilizes the protein. (Capaldi et al., 2002) My results shows that both alanine and asparagine mutation destabilize the initial unfolding transitions compared with its glutamate and glutamine mutants. (Figure 5.1, 5.2 and 5.3) Finally, the glutamate and glutamine mutations at N capping positions destabilize the later transitions. This observation cannot be easily explained by the N-Capping theory and needs further investigation.

pH related N capping stability.

It seems all the N-capping studies agree that the negatively charged aspartate (Asp-) and non charged aspartate (Asp^0) have significant different N-capping preference, which means when we lower the pH, the negative charged aspartate residue becomes uncharged, and this change could largely destabilize the protein. (Gsponer et al., 2006, Capaldi et al., 2002, Doig and Baldwin, 1995, Lindeberg et al., 2000)

The results in Chapter Three show that aspartate to alanine mutations significantly destabilize the protein. The results in Chapter four and this chapter show that aspartate to asparagine mutations at some N-capping position slightly destabilize the first transition in the urea unfolding process, whereas the aspartate to glutamate or glutamine substitutions actually stabilize the protein more than expected.

Comparing the primary structure of the pH sensitive colicin pore forming domains such as ColA-P, ColB-P and ColE1-P and non pH sensitive colicin such as ColN-P, reveals negatively charged amino acid residues such as aspartate and glutamate appear at the N-capping position on the pH sensitive colicin (such as positions D216, D241 and D372) whereas neutral amino acid residues are found at the N-capping position of the non pH sensitive colicin. Thus, the N-capping theory could still be adapted to explain the stability of pH dependent colicins.

pH sensitivity of the colicin A pore forming domain

Fridd et al examined the stability of mutations of the colicin N pore forming domain which introduced aspartate residues at equivalent sites where they occur in ColA-P. This was thus the inverse experiment to the one I have conducted here. They included valine to aspartate, alanine to aspartate and asparagine to aspartate point mutations. Using near UV CD spectroscopy the results showed small changes in tertiary structure and no change is pH sensitivity (Fridd and Lakey, 2002)

The result in my study indicates that asparagine, glutamate and glutamine mutants do not destabilize the protein as much as alanine mutants. The N-capping hypothesis therefore is a good generally applicable explanation of why those surface aspartate residues are essential to keep ColA-P in its native state at neutral pH and unfolding to the molten globule state at low pH. Thus further study of N-capping theory may still explain the role of other "surface" residues in protein stability and folding.

Table 5.2, 5.3, 5.6, and 5.7 show that ΔG_1 of glutamate and glutamine mutants dramatically decrease when reducing the pH, which means that ColA-P WT and its mutants are easy to transform from their folded stage to their molten globule state at low pH. At neutral pH, ΔG_1 of the wild type is 21.83 kcal, which means the protein cannot easily to form to the molten globule state at this pH. At pH 3.0, the ΔG_1 of ColA-P WT drops to 4.66 kcal mol $^{-1}$, which means that it is much easier for the

protein to unfold to its molten globule state in an acidic environment. In chapter three, I have discussed how the formation of the molten globule state would reduce the energy required to insert into the membrane of cells. (van der Goot et al., 1992b) The acidic environment would reduce the energy for the protein to unfold to its molten globule stage. Thus the low pH would be essential for colicin A pore forming domain to transport or insert to the *E. coli* inner membrane and this is confirmed by the need for acidic phospholipids to allow colicin A to function. It is obvious that mutations at positions D216, D274 and D372 would be preferential for unfolding the intermediate stage at pH 3.0. Generally, all the glutamate and glutamine mutants need less energy to unfold to their intermediate stage at both neutral pH and low pH and alanine mutants can form the molten globule state without denaturant at low pH. (Tables 5.1, 5.2, 5.3, 5.5, 5.6, 5.6)

The change of Gibbs free energy ($\Delta\Delta G$) measured here represents the energy difference between two states of protein. Low ΔG_1 means Colicin A pore forming domain needs less energy to transform from its native state to its molten globule state. Thus, the change of ΔG_1 between pH 7.0 and pH 3.0 can be used to detect how sensitive the protein can be affected by pH. The change of ΔG_1 is shown in table 5.10.

Table 5.10: The Gibbs free energy and the changes of Gibbs free energy of the initial phase for colicin A- pore forming domain and its mutants determined by urea unfolding at pH 7.0 and pH 3.0

Mutation	$\Delta G1_a$ at pH 7.0 Kcal.mol^{-1}	$\Delta G1_b$ at pH 3.0 Kcal.mol^{-1}	$\Delta G1_a$- $\Delta G1_b$ Kcal.mol^{-1}	First transition at pH 7.0 (Urea concentration M)	Second transition at pH 3.0 (Urea concentration M)
D216A	14.70	0	14.70	3.8-5.0	0-
D216N	21.42	3.33	18.09	5.0-6.1	0.5-1.8
D216E	nd	3.80	nd	5.7-	0.6-1.7
D216Q	nd	1.70	nd	5.3-	0.3-1.8
D274A	8.05	0	8.05	4.4-5.1	0-
D274N	25.47	3.65	21.82	5.1-5.8	1.2-2.9
D274E	15.81	4.64	11.17	5.1-6.0	0.8-2.4
D274Q	15.14	3.16	11.98	4.9-5.8	1.2-2.7
D372A	13.36	0	13.36	3.2-4.3	0-
D372N	18.62	2.66	15.96	4.6-6.0	0.6-1.7
D372E	16.87	2.74	14.13	5.1-5.8	0.5-2.1
D372Q	20.70	2.27	18.43	5.0-5.8	0.4-2.0
D241A	8.4	--	--	4.9-5.7	--
D241N	22.28	--	--	5.2-6.0	--
S239A	21.38	4.52	16.86	5.4-6.5	1.8-3.3
S239D-D241A	17.61	5.86	11.75	5.7-6.6	1.6-3.0
WT	21.83	4.66	18.23	5.2-6.0	1.0-2.2

nd = not determined, two separate transitions are not defined at this pH * --- data not collected When the mutant forms to the intermediate stage without urea, $\Delta G1$ was regarded as 0.

All the mutants at the N-capping position destabilize at low pH but almost all of the aspartate to alanine, aspartate to asparagine, aspartate to glutamine and aspartate to glutamate mutation at the N-capping position reduce the protein pH sensitivity, which

confirms that aspartate residues at N-capping position make the helical peptide sensitive to the pH and indicates that the aspartate at N-capping position is essential to maintain the pH sensitivity of proteins.

Colicin A and N pore forming domains show a high level of sequence homology indicating divergence from a common ancestor. Colicin A has been shown to be sensitive to the presence of acidic lipid in the inner membrane of *E. coli* and to form an acidic molten globule whilst colicin N does neither. They must therefore have adopted different ways to unfold and insert into the inner membrane, the evolution of which has involved the insertion of aspartates at specific sites in colicin A and their avoidance in colicin N. This raises one intriguing question, how does colicin N insert into the membrane?

Chapter Six

Studying the local structure changes in the Colicin A pore-forming domain by NMR

Colicin A pore-forming domain forms an ion channel in the inner membrane of bacterial cells. In all channel forming colicins, a hydrophobic, 32-48 residue segment exists near the C-terminus of the protein and it is through this "hydrophobic hairpin" that the channel is proposed to anchor to the membrane . The remaining amphiphilic helices would then insert into the membrane to form the ion channel driven by a trans-membrane potential (Merrill and Cramer, 1990, Slatin et al., 1994, van der Wal et al., 1995). For the colicin A channel forming domain, it is believed that only some of the helices (helix 6, 7, 8, 9 and 10) insert into the inner membrane and form the channel, while the rest of the helices (helix 1, 2, 3 and 4) incompletely translocate and play no role in the channel. Colicin A and colicin E1 are members of the channel-forming subfamily of colicins, and whilst both of their pore-forming activities are enhanced by an acidic environment, colicin A shows a much clearer transition (Lakey and Slatin, 2001). A precise pH-sensitive region within the soluble channel forming domain was identified in a study of the mechanism of colicin E1 low pH activation (Merrill et al., 1997), but this is entirely different to the mechanism of Colicin A activation described in the previous chapters. Thus, in order to assess if the colicin A pore-forming domain also has a specific pH-sensitive region, it is interesting to study the stability of each helix in colicin A pore-forming domain at low pH. The methods used so far do not provide information on which part of colicin A P-domain are unfolded at low pH or how the mutants influence the structure. Thus here I used NMR to define the structure changes more precisely.

6.1 Aims

In the last two decades, nuclear magnetic resonance (NMR) has become an important spectroscopic technique for determining the three dimensional structures of proteins in solution. NMR spectra can also characterize dynamic properties of proteins as well as allow their study under conditions directly relevant to their biological function. The two-dimensional heteronuclear single quantum coherence (HSQC) NMR experiment is a sensitive technique for the determination of heteronuclear connectivities e.g. within proteins, by correlating the ^{1}H and ^{15}N resonances of NH bonds via their one-bond coupling constant (Hecht et al., 2009). A protein of even a modest molecular weight has a very large number of chemically different protons, which makes their one-dimensional ^{1}H spectra difficult to fully analyse, but with the ^{1}H-^{15}N-HSQC spectrum only the protons coupled to ^{15}N are detected and this, together with the extra dimension, makes such spectra much easier to analyse. NMR assignments, and the backbone dynamics of colicin A pore-forming domain in solution monitored through ^{15}N relaxation data, were reported in 2009 (Ibanez de Opakua et al., 2010). The assignments were deposited in the BioMagRes databank and can be used as the basis for further studies of the channel forming protein. Since I have found that ColA-P unfolds in two steps, which might correlate with its channel forming activity, I decided to employ NMR to investigate structural features of the unfolding with the aim of identifying whether there are destabilized helices during the initial transformation. A combination of NMR spectra of wild-type and site-directed mutants of colicin A at neutral pH and at low pH will be used.

6.2 NMR spectroscopy

All spectra reported here were acquired using a Bruker Avance III 800 MHz spectrometer equipped with a triple resonance, pulsed field gradient probe, operating at a ^{1}H frequency of 800.23 MHz and ^{15}N frequency of 81.09 MHz, using pulse sequences incorporated into the Bruker Topspin 2.1 software. 2D ^{1}H-^{15}N- HSQC spectra were recorded at 298 K. The ^{1}H carrier frequency was positioned at the resonance of the water during the experiments. The ^{15}N carrier frequency was at 115 ppm. The spectrometer was housed in the School of Chemistry, University of East Anglia UK. Data were processed using NMRPipe [30] or Bruker TopSpin 2.1 software and analysed with CCPNMR (Collaborative Computing Project for NMR) analysis software or NMRView. Before Fourier transformation, a cosine-bell window function was applied to each dimension for apodization. The indirect dimensions were first linear-predicted to double the number of data points, and then zero-filled to round up the number of data points.

^{1}H-^{15}N-HSQC NMR spectra

On a ^{1}H-^{15}N-HSQC NMR spectrum, the X-axis indicates the chemical shifts of protons and the Y-axis indicates the chemical shifts of nitrogen, with cross-peaks giving the chemical shifts of spin-spin coupled proton and nitrogen nuclei. So this experiment provides a chemical shift map of the NH groups in the protein. The intensity of the cross-peak, reflected in the numbers of contours it has, and its line width are other parameters that can be readily extracted from the spectrum. The major class of NH groups detected in this experiment are the backbone amide groups, but some side chain NH groups are also seen, particularly the indole NH of tryptophan and the side chain amide groups of asparagine and glutamine. Histidine side chain NH

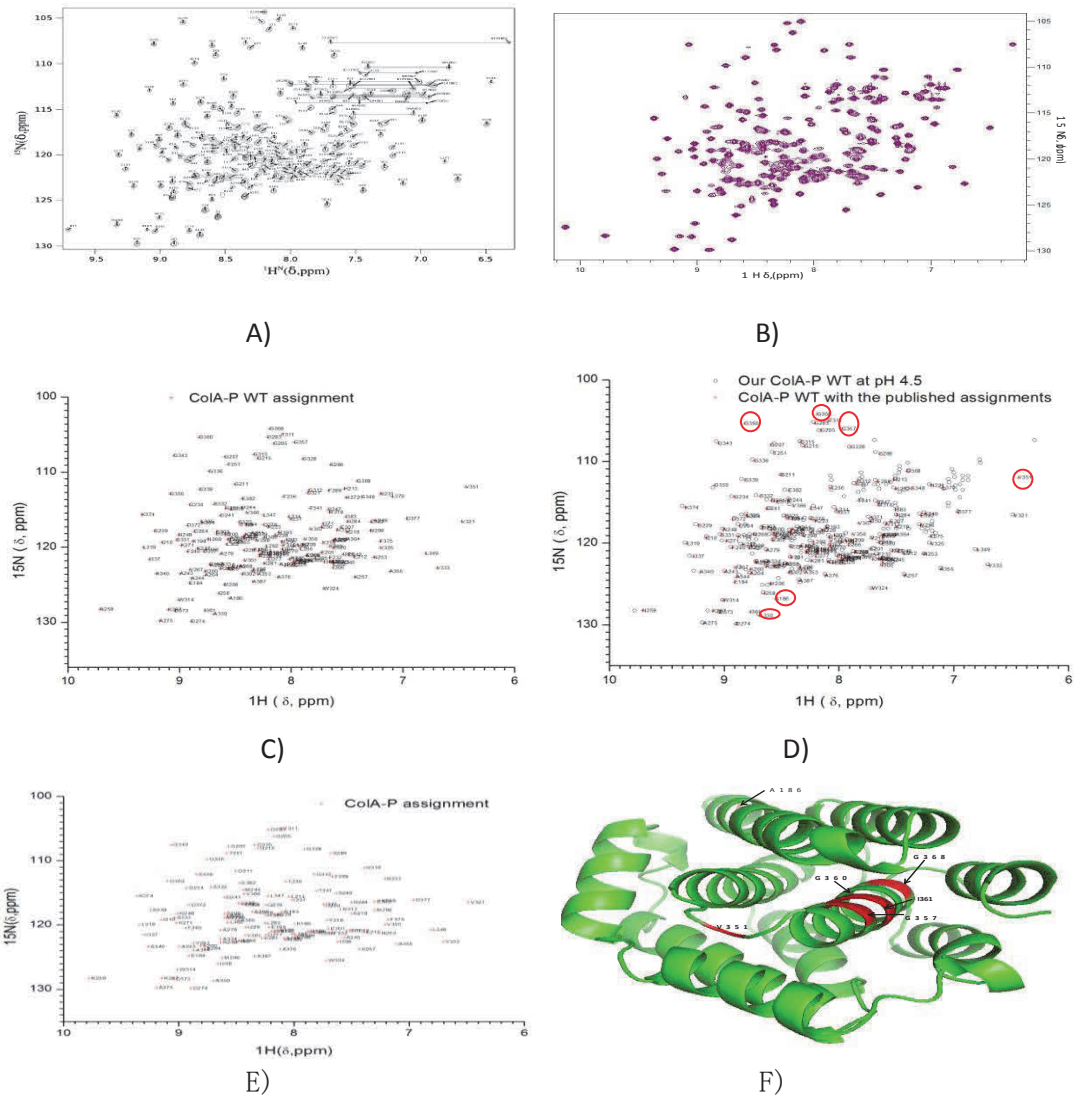
groups can be detected at much greater ^{15}N chemical shifts than the backbone NH region, and so are not seen in my spectra. Lysine and arginine NH signals are occasionally seen in the backbone NH region but usually they are not detected as the exchange of the proton with the solvent protons effectively bleaches their resonances. Uniformly ^{15}N-labeled ColA-P WT (with histidine tag) and site-directed mutants of it were expressed in *E. coli* strain BL21 (AI) in M9 minimal medium, purified as described in Chapter Two and dissolved in 90% H_2O/10% D_2O (v/v). Buffer conditions, temperature and urea concentrations, if any, for the NMR samples are given in the figure captions.

6.3 Results

NMR spectra of ColA-P WT at pH 4.5

NMR assignments of the ColA-P domain in solution at pH 4.5 have been reported by Ibañez de Opakua et al, 2010 (Ibanez de Opakua et al., 2010, Kienker et al., 2002) including those for the ^{1}H-^{15}N-HSQC spectrum. These were used to assign the signals in spectra of ColA-P in this study. The published ^{1}H-^{15}N-HSQC spectrum of ColA-P domain was collected at 298K and at pH 4.5, so I recorded a ^{1}H-^{15}N-HSQC spectrum of ColA-P WT under the same conditions except that I used citrate buffer as in the previous chapters and they used acetate buffer. The published spectrum is available in the BioMagRes data bank, through a listing of each peak's chemical shift and these were superimposed on my spectrum of ColA-P WT using OriginPro 7.0.

The spectra are reproduced in the following figures but larger scale copies are presented as an appendix to this chapter.



* The enlarged figures are appended on the following pages.

Figure 6.1. A) Published 800 MHz ^{1}H-^{15}N-HSQC spectrum of ColA-P WT at pH 4.5, 50 mM sodium citrate buffer, with residue specific assignments reproduced from (Ibanez de Opakua et al., 2010). B) My 800 MHz ^{1}H-^{15}N-HSQC spectrum of ColA-P WT at pH 4.5, 50 mM sodium citrate buffer C) Centre of cross-peaks of published ColA-P WT ^{1}H-^{15}N-HSQC spectra from the deposited chemical shifts (http://www.bmrb.wisc.edu) D) Overlaid cross-peak centres of the published ColA-P WT with assignments (Red cross) and my 1-H-15-N-HSQC spectra (Black circles) at pH 4.5 E) The cross-peaks of my ColA-P WT 1-H-15-N-HSQC spectra with residues

assignments indicated by colicin N number F) Cartoon of the structure of ColA-P domain (Green) highlighted with the residues missing in my spectrum (Red).

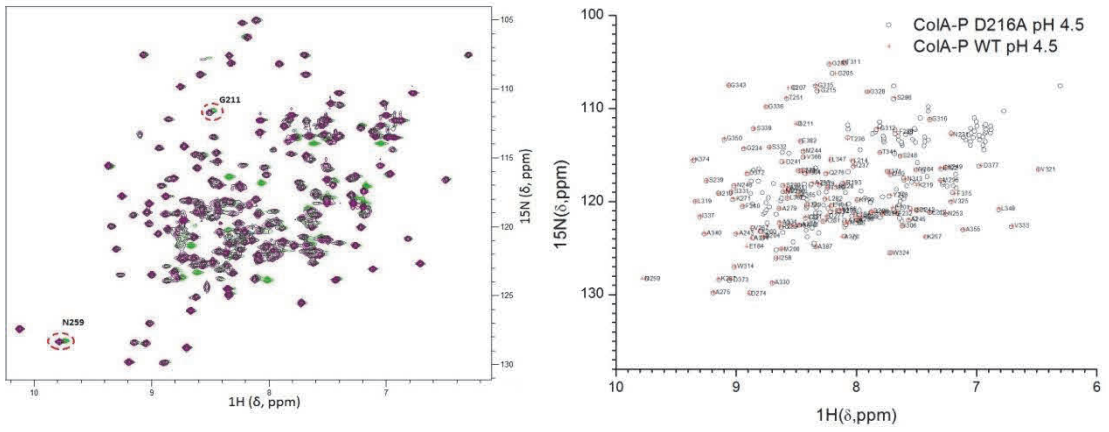
Figure 6.1 A) and B) show that the published spectrum of ColA-P domain and my spectrum are very similar to each other. However, there are a few differences between the two spectra which are probably caused by one or more of the following: i) we have a histidine tag on the N-terminus of my protein samples whilst they do not; ii) they have three extra amino acids on the N-terminal end of their protein sample; iii) their protein was dissolved in acetate buffer; but in order to retain the same buffer that we used in the urea unfolding experiments, we dissolved the protein in sodium citrate buffer.

The crowded central region of the spectrum is difficult to assign so the reliable assignments occur in the outer region. The Figure 6.1 B shows that six peaks in the non-crowded region of the spectra appear in the published data but were not observed in my spectrum at the same chemical shifts. These six residues are located in the hydrophobic core of the protein and far away from the N-terminal of the protein where the histidine tag is present (Figure 6.1 F). This difference may be a consequence of the differences in the protein construct and/or buffer difference affecting the hydrophobic hairpin. The line widths of resonances are extremely sensitive to dynamic processes affecting their nuclei and it is likely that whatever dynamics there are affecting the residues of the missing resonances have caused them to broaden beyond detection.

For the convenience of the further study, 123 peaks of the expected 206 main chain NH have been assigned in the uncrowded region based on the published data. I consider my assignments to be reliable because the assigned peaks are the resonances that exactly overlapped on the published ColA-P WT ^{1}H-^{15}N-HSQC spectra and my WT spectra. Assigned peaks in the crowded region were not included in my analysis. Fig 6.1 F) shows the X-ray structure of colA-P with residues in the uncrowded region

whose NH signal we have not observed coloured red. As can be seen, though we have an incomplete assignment, we have assigned signals throughout the structure. We are thus able to monitor conformational perturbations resulting from changes in solution conditions or mutations throughout the structure.

^{5}N-^{1}H-HSQC spectrum of D216A at pH 4.5



A) B)

* The enlarged figures are appended.

Figure 6.2. A) Overlaid 800 MHz 1-H-15-N-HSQC spectrum of ColA-P WT (Purple) and ColA-P D216A (Green) at pH 4.5 B) Overlaid peaks of ColA-P WT with assignments (red cross) and ColA-P D216A (Black circle) at pH 4.5

The overlaid spectra of ColA-P WT and D216A (Figure 6.2 A) show that only two peaks on the non-crowded region , residues G211 and N259 of ColA-P D216A, show a small chemical shift change (approximately 1 ppm on both the proton and nitrogen axes) compared with the wild type protein at pH 4.5(Figure 6.2 B). G207 is located in the first alpha helix of colicin A-P domain and N259 is located on the loop between helix III and helix IV of the protein.

Since G211 is close to D216 it is probably directly affected by the mutation, N259 is far away and located in a flexible inter-helix loop (Figure 6.3). The origin of its chemical shift difference is not clear but may be a consequence of the loop flexibility.

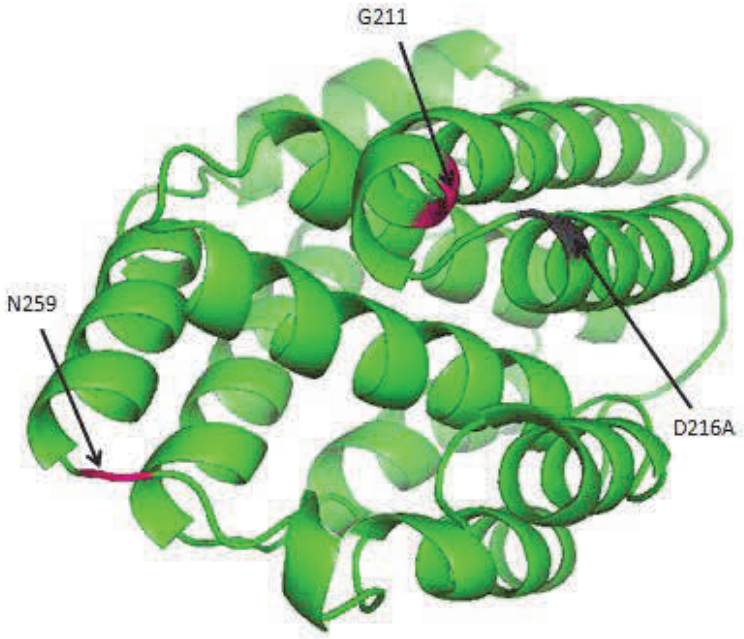
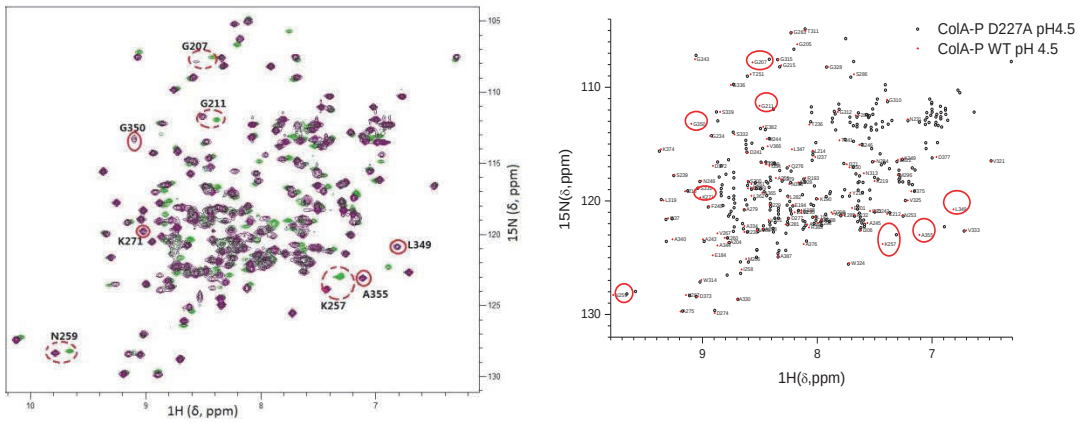


Figure 6.3. The crystal structure of ColA-P domain (PyMOL) highlighted with the residues with small chemical shifts (Magenta) in ColA-P D216A (the point mutation was highlight with dark grey) at pH 4.5

The chief conclusion from Figures 6.2 and 6.3, however, is that the structures of ColA-P D216A and the wild type protein at pH 4.5 are almost identical to each other.

^{15}N-^{1}H-HSQC spectrum of ColA-P D227A at pH 4.5



A）　　B）

* The enlarged figures are appended on the following pages.

Figure 6.4. A) Overlaid 800 MHz ^{1}H-^{15}N-HSQC spectrum of ColA-P WT (Purple) and ColA-P D227A(Green) at pH 4.5 B) Overlaid peaks of ColA-P WT with assignments (red cross) and ColA-P D227A (Black circle) at pH 4.5

There are four peaks (K271, L349, G350 and A355) in the non-crowded region of the WT spectrum not visible in the spectrum of ColA-P D227A (Figure 6.4 A) with another four peaks (G207, G211, K257 and N259) showing an obvious chemical shift perturbation (approximately 3 ppm on the proton axis and 3 ppm on the nitrogen axis). These residues are highlighted in the crystal structure of ColA-P wild type protein (Figure 6.5). L349, G350 and A355 are located in the hydrophobic core of the protein while G207, G211, K257, N259 and K271 are spread over the helices I, III and IV and the loop between helices III and IV. Thus the ColA-P D227A mutation cuases wide spread changes in helix II, helix III, helix IV and the hydrophobic core of the protein.

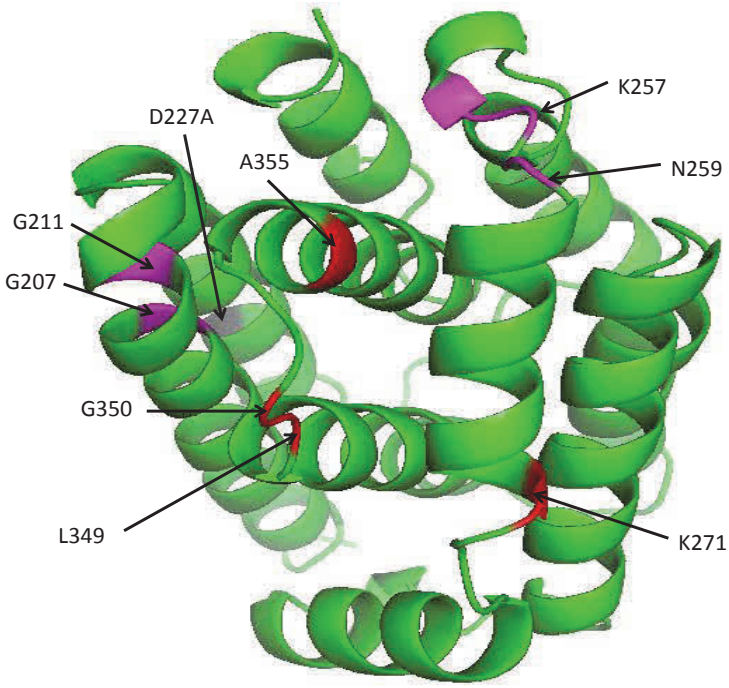
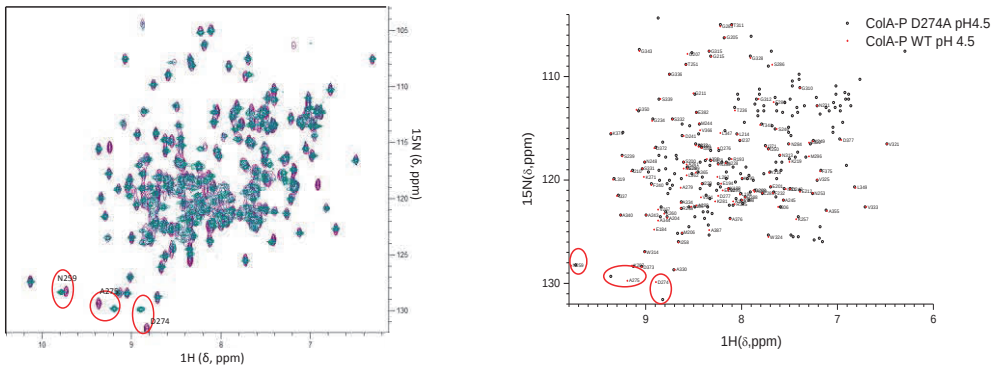


Figure 6.5. The crystal structure of ColA-P domain (PyMOL) showing the residues whose resonances are missing (Red) or have perturbed chemical shifts (Magenta) in ColA-P D227A (the point mutation is highlight with dark grey) at pH 4.5.

The point mutation D227A is located in the middle of the second helix. However, most of the missing peaks are located on the hydrophobic hairpin. Therefore, this point mutation perturbs the hydrophobic core of the protein.

Two important factors that can be learned from Figures 6.4 and 6.5. The first one is that D227A perturbs the protein more than D216A at pH 4.5. The second one is that the hydrophobic core is affected by the mutation, and helix I, helix III and helix IV of the protein are more flexible than the remaining helices at low pH. As with D216, however, the chief observation is that the structure of D227 ColA-P is similar to the WT structure.

$^{1}H^{15}N$-HSQC spectrum of ColA-P D274A at pH 4.5



A) B)

* The enlarged figures are appended.

Figure 6.6. A) Overlaid 800 MHz ^{1}H-^{15}N-HSQC spectrum of ColA-P WT (Purple) and ColA-P D274A (Dark Green) at pH 4.5 B) Overlaid peaks of ColA-P WT with assignments (Red cross) and ColA-P D274A (Black circle) at pH 4.5

Only three peaks (D274, A275 and N259) in the non-crowded region of the ColA-P D274A spectrum show a small chemical shift change (Figure 6.6 A). Figure 6.6 B) indicates that the peaks show a chemical shift perturbation of approximately 1 ppm on the proton axis and 20 ppm on nitrogen axis compared with the wild type protein at pH 4.5. Residues D274 and A275 are located on the N-terminal end of helix V of the protein and N259 is located on the loop between helix III and helix IV of the protein.

The residues whose peaks have shifted are shown in the crystal structure of ColA-P wild type protein. (Figure 6.7)

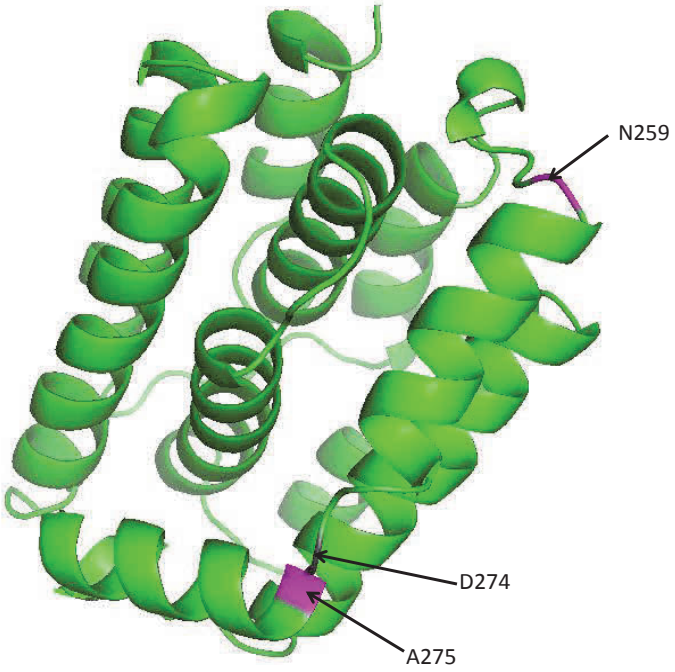
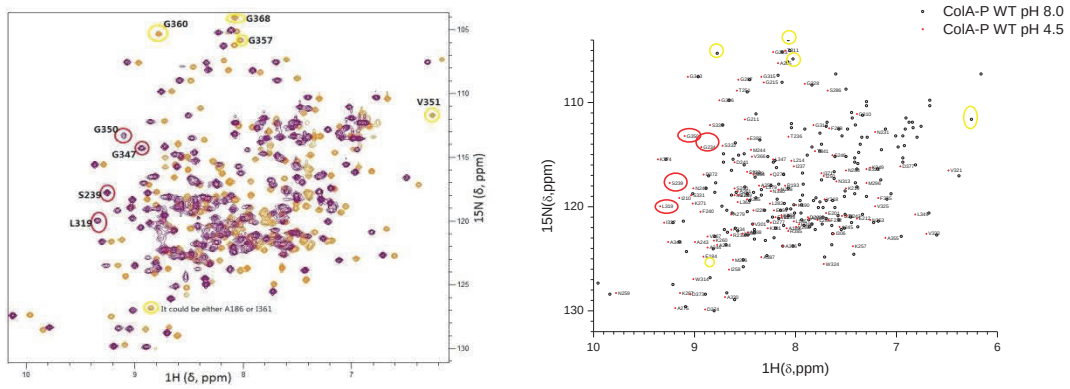


Figure 6.7. The crystal structure of ColA-P domain (PyMOL) highlighted with the residues with small chemical shift changes (Magenta) in ColA-P D274A (the point mutation is highlighted with dark grey) at pH 4.5

Figure 6.7 shows that the effects on D274 (the point mutation) and its neighbour A275 are caused locally by the mutation. However, residue N259 is far away and is evidence of longer distance destabilisation effects. This result is interesting since N259 was also affected by mutations at positions 216 and 227 (see above) thus this loop appears to be generally susceptible to destabilisation.

Overall, we can conclude that the structures and dynamics of ColA-P WT, D216A and D274A are very similar at pH 4.5 (Figures 6.2, 6.3, 6.4, 6.5, 6.6 and 6.7) whilst the mutation ColA-P D227A has more widespread effects.

^{1}H-^{15}N-HSQC spectrum of ColA-P WT at pH 8.0



A) B)

* The enlarged figures are appended.

Figure 6.8. A) Overlaid 800 MHz ^{1}H-^{15}N-HSQC spectrum of ColA-P WT at pH 8.0 (Orange) and pH 4.5 (Purple) B) Overlaid peaks of ColA-P WT with assignments (Red cross) and at pH 4.5 (Black circle)

Figure 6.8. A) shows that most of the peaks of ColA-P WT at pH 8.0 are shifted away from the corresponding peaks at pH 4.5. It means the structure of ColA-P WT at pH 4.5 and at pH 8.0 are different. This result confirms that pH plays an important role in colicin A-P domain stability even above pH 4.5.

Although, most of the peaks are shifted, they can still be observed in the spectrum. Only the peaks of residues S239, L319, G347 and G350 cannot be observed in the spectrum of the wild type protein at pH 8.0 (Figure 6.8. B). These are located on the loop between helix II and helix III, helix VII and the hydrophobic core of the protein. More interesting though, is that the resonances of V351, G357, G360 and G368, which are not visible in my spectrum of the wild type protein at pH 4.5, appear in the spectrum at pH 8.0. All of these four residues are located in the hydrophobic core of the protein as shown by the crystal structure of ColA-P wild type protein. (Figure 6.9)

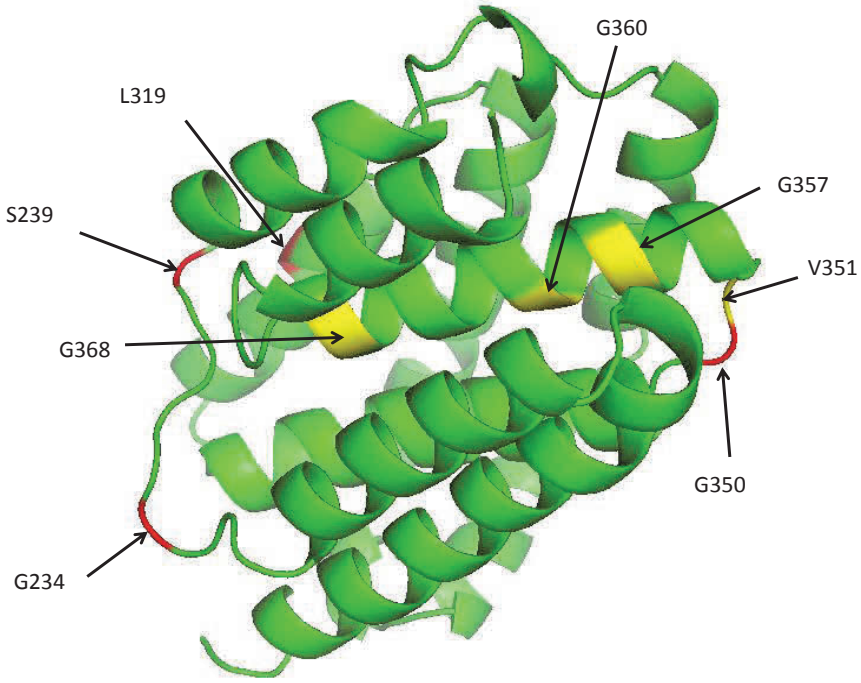
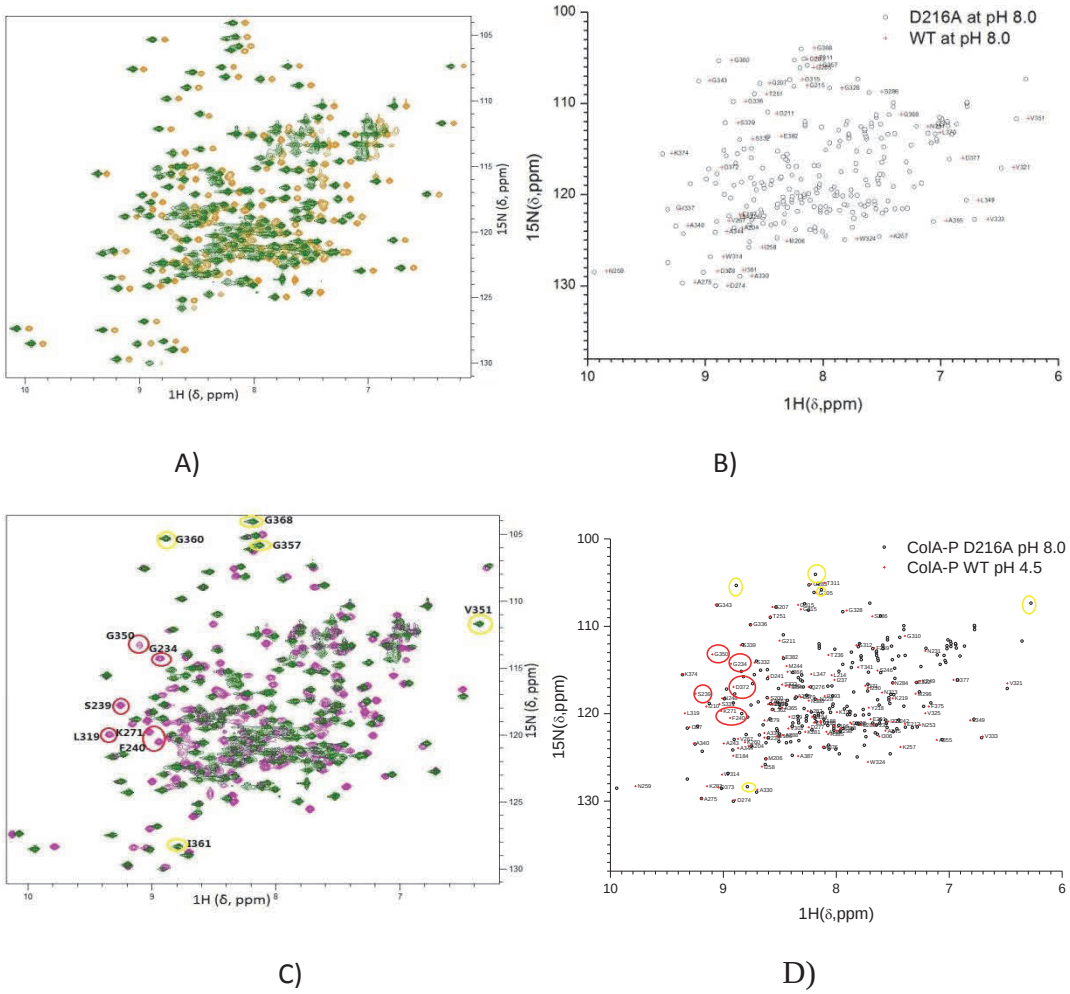


Figure 6.9. The crystal structure of ColA-P domain (PyMOL) with the residues whose resonances detected by Ibanez de Opakua *et al* and only observed in my spectra at pH 8.0 are coloured yellow. Residues with the additional missing peaks at pH 8.0 are coloured red.

Figure 6.9. shows that most of the ColA-P WT resonances affected by the change from pH 4.5 to pH 8.0 are from residues located in the hydrophobic core of the protein or the loops between the helices of the protein. It leads us to think that the hydrophobic hairpin and the loop regions are the most pH sensitive areas of the protein.

NMR spectra of ColA-P mutants at pH 8.0



* The enlarged figures are appended on the following pages.

Figure 6.10. A) Overlaid 800 MHz ^{1}H-^{15}N-HSQC spectrum of ColA-P WT(Orange) and ColA-P D216A (Green) at pH 8.0 B) Overlaid peaks of ColA-P WT with assignments (red cross) and ColA-P D216A (Black circle) at pH 8.0 C) Overlaid 800 MHz ^{1}H-^{15}N-HSQC spectrum of ColA-P WT (Purple) at pH 4.5 and ColA-P D216A (Green) at pH 8.0 D) Overlaid peaks of ColA-P WT at pH 4.5 with assignments (Red cross) and ColA-P D216A (Black circle) at pH 8.0

Figure 6.10 A) shows that most of the peaks of ColA-P D216A at pH 8.0 shift toward lower chemical shifts on the proton axis compared with the wild type protein at pH

8.0. Thus most of the peaks in the spectrum of ColA-P D216A at pH 8.0 and the wild type protein at pH 4.5 overlap each other. This means that the solution structures of ColA-P D216A at pH 8.0 and the wild type at pH 4.5 are similar to each other and that the destabilizations caused by the mutation and low pH are similar to each other. In other words, both the mutation and the lower pH (4.5) destabilize the protein to the same degree.

However, there are some differences between the spectra of D216A at pH 8.0 and the wild type protein at pH 4.5: the peaks of G234, S239, F240, K271, S319 and G350 which appear in the spectrum of the wild type protein at pH 4.5 cannot be observed in the spectrum of ColA-P D216A at pH 8.0. As above the peaks of V351, G357, G360 and G368 which cannot be observed in my spectrum of the wild type protein at pH 4.5 also reappear in the spectrum of ColA-P D216A at pH 8.0. (Figure 6.11)

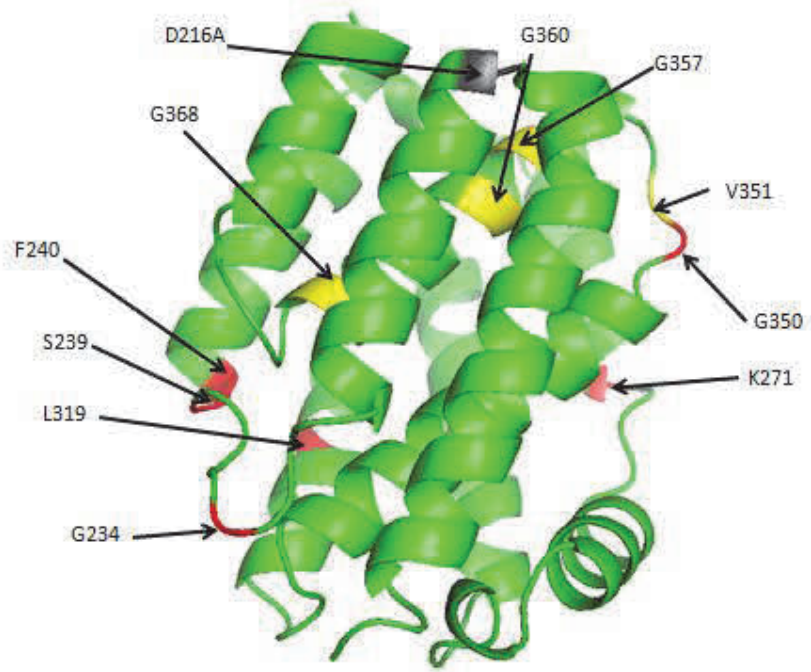
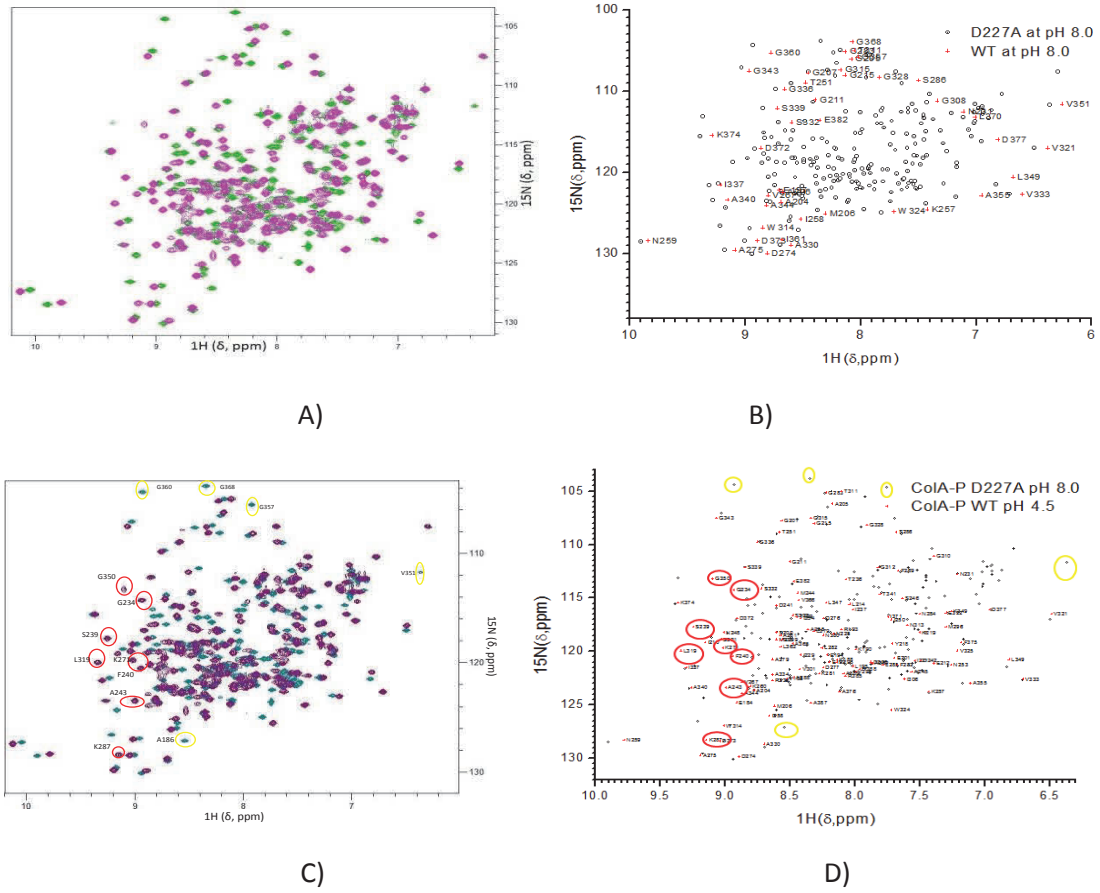


Figure 6.11. The crystal structure of ColA-P domain (PyMOL) highlighted with the residues that cannot be observed on my spectra of the wild type protein at pH 4.5 but

appear on the spectra of ColA-P D216A at pH 8.0 (Yellow). The missing peaks on the spectra compared to that of ColA-P WT at pH 8.0 are highlighted in Red colour.

Figure 6.11. shows that the resonances which either disappear or appear on the spectra of ColA-P D216A at pH 8.0 compared with the spectra of wild type protein at pH 4.5 are all located on the hydrophobic core of the protein or the loops between the helices of the protein or on tip of the helices. It confirms that the hydrophobic hairpin, the loops and the tip of the helices are more sensitive to this mutation than other regions of the protein.



* The enlarged figures are appended on the following pages.

Figure 6.12 A) Overlaid 800 MHz 1-H-15-N-HSQC spectrum of ColA-P WT (Green) and ColA-P D227A (Purple) at pH 8.0 B) Overlaid peaks of ColA-P WT with assignments (Red cross) and ColA-P D216A (Black circle) at pH 8.0 C) Overlaid 800 MHz ^{1}H-^{15}N-HSQC spectrum of ColA-P WT (Purple) at pH 4.5 and ColA-P D227A

(Dark Green) at pH 8.0 D) Overlaid peaks of ColA-P WT at pH 4.5 with assignments (Red cross) and ColA-P D216A (Black circle) at pH 8.0 .

Figure 6.12 A) shows that most of the peaks of ColA-P D227A at pH 8.0 also shift towards lower chemical shift values on the proton axis compared with the wild type at pH 8.0. However, figure 6.12 C) shows that most of the peaks from the spectra of ColA-P D227A at pH 8.0 and the wild type at pH 4.5 overlap each other. It means that the structures of ColA-P D227A at pH 8.0 and the wild type at pH 4.5 in solution are similar to each other. Thus, just like D216A, D227A destabilizes the protein at neutral pH and the spectrum at pH 8.0 resembles that of the wild type at pH 4.5.

Figure 6.12. C) and D) show peaks of residues G234, S239, F240, A243, K271, K287, S319 and G350 which appear on the spectra of the wild type protein at pH 4.5 cannot be observed on the spectra of ColA-P D216A at pH 8.0. The peaks of residues A186 V351, G357, G360 and G368 which cannot be observed on the spectra of the wild type protein at pH 4.5 appear on the spectra of ColA-P D227A at pH 8.0. These residues are located on the region of the hydrophobic core, helix III, C-terminal of helix IV, helix VII or the loop between helix II and III and the loop between helix X and XI. (The observation only included the peaks located at the non-crowded region of the spectra)

The peaks which have shifted in ColA-P D227A are highlighted and shown on the crystal structure of ColA-P wild type protein. (Figure 6.13)

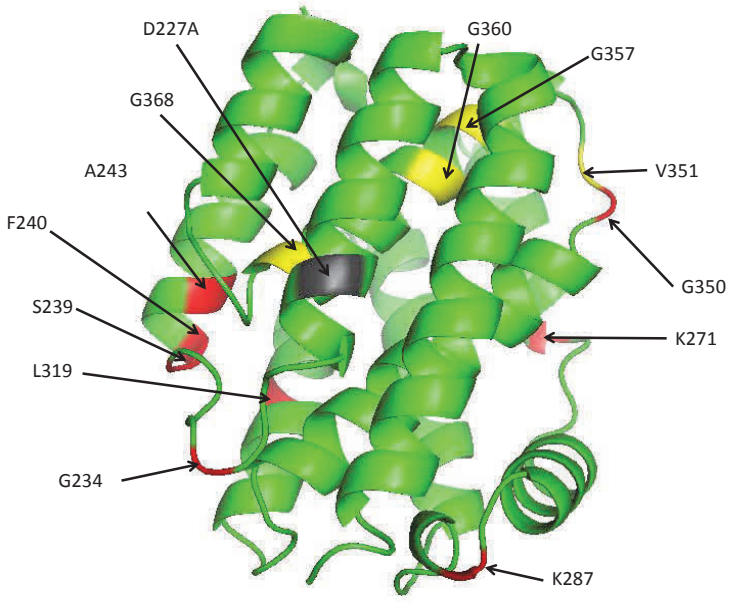
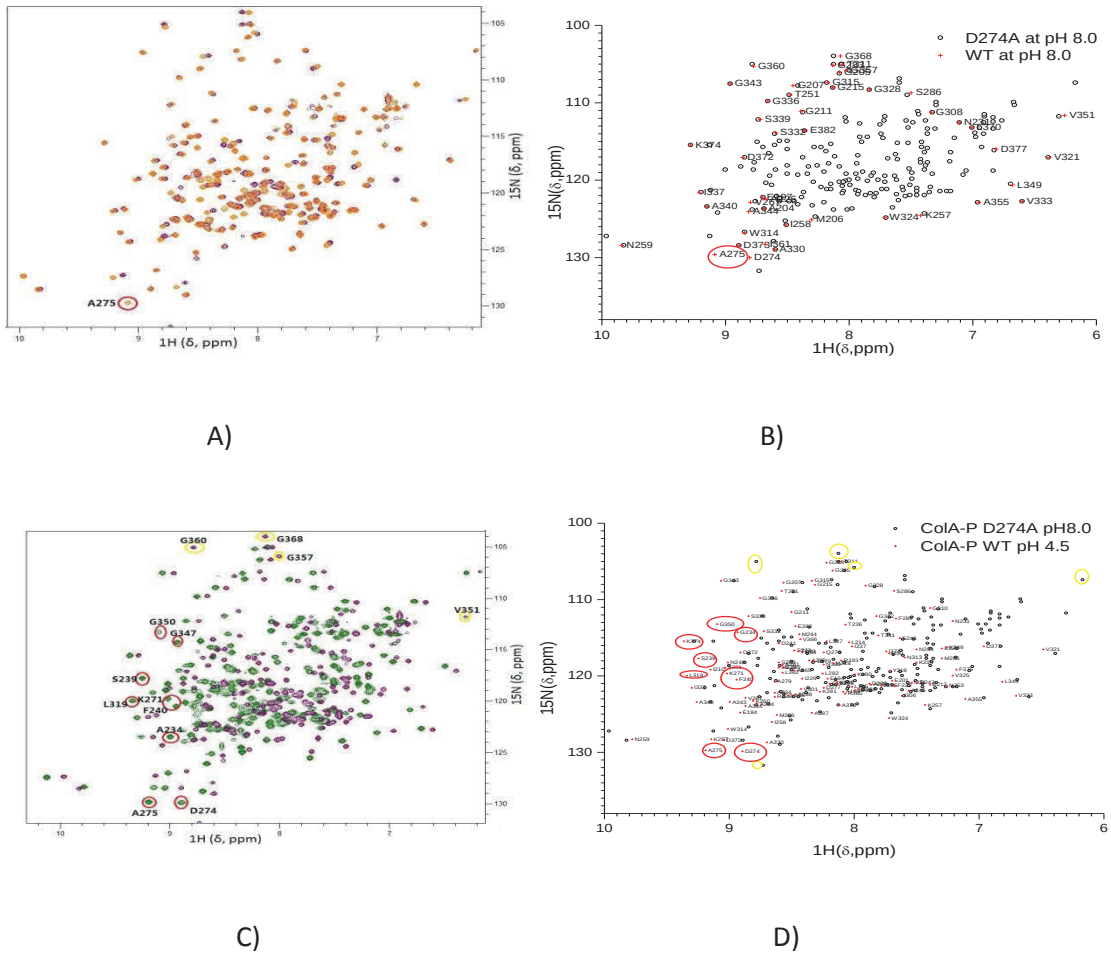


Figure 6.13. The crystal structure of ColA-P domain (PyMOL) highlighted with the residues that cannot be observed on my spectra of the wild type protein at pH 4.5 but which appear on the spectra of ColA-P D227A at pH 8.0 (Yellow). The missing peaks on the spectra of ColA-P WT at pH 8.0 are highlighted in red .

Figure 6.13. shows that most of the ColA-P D227A resonances which either disappear or appear on the spectrum of at pH 8.0 compared with the spectrum of wild type protein at pH 4.5 are also located on the hydrophobic core of the protein or the loops between the helices of the protein or on top of the helices. Comparing figure 6.11 and 6.13, we can see that helix III of ColA-P D227 could be more flexible than it of D216A at pH 8.0.



* The enlarged figures are appended on the following pages.

Figure 6.14 A) Overlaid 800 MHz ^{1}H-^{15}N-HSQC spectrum of ColA-P WT (Orange) and ColA-P D274A (Purple) at pH 8.0 B) Overlaid peaks of ColA-P WT with assignments (red cross) and ColA-P D216A (Black circle) at pH 8.0 C) Overlaid 800 MHz ^{1}H-^{15}N-HSQC spectrum of ColA-P WT at pH 4.5 (Green) and ColA-P D274A at pH 8.0 (Purple) D) Overlaid peaks of ColA-P WT with assignments (red cross) and ColA-P D274A (Black circle) at pH 4.5

Figure 6.14 A) shows that almost all of the peaks of ColA-P WT and D274A overlapped with each other at pH 8.0. Only one resonance-K287 disappears in the spectrum of ColA-P D274A at pH 8.0. Unlike ColA-P D216A and D227A, figure 6.14 C) shows that most of the peaks from the spectra of ColA-P D274A shift

compared with the wild type at pH 4.5. It means that the structures of ColA-P D274A and the wild type at pH 8.0 in solution are very similar to each other. From the previous chapters we have learnt that D274A destabilizes the protein less at neutral pH than ColA-P D216A and D227A which agrees with the NMR results.

The only peaks (in the non-crowded region of the spectrum) to show obvious chemical shifts differences at pH 8.0 compared to the wild type are from D274, A275 and G360. The first two are local to the mutation but G360 is located on the hydrophobic core of the protein (Figure 6.15)

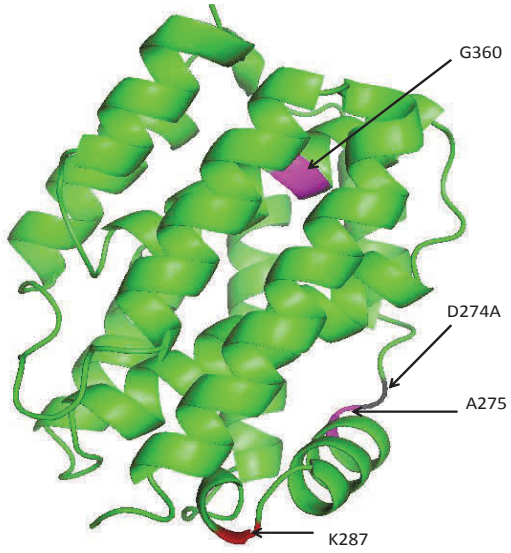
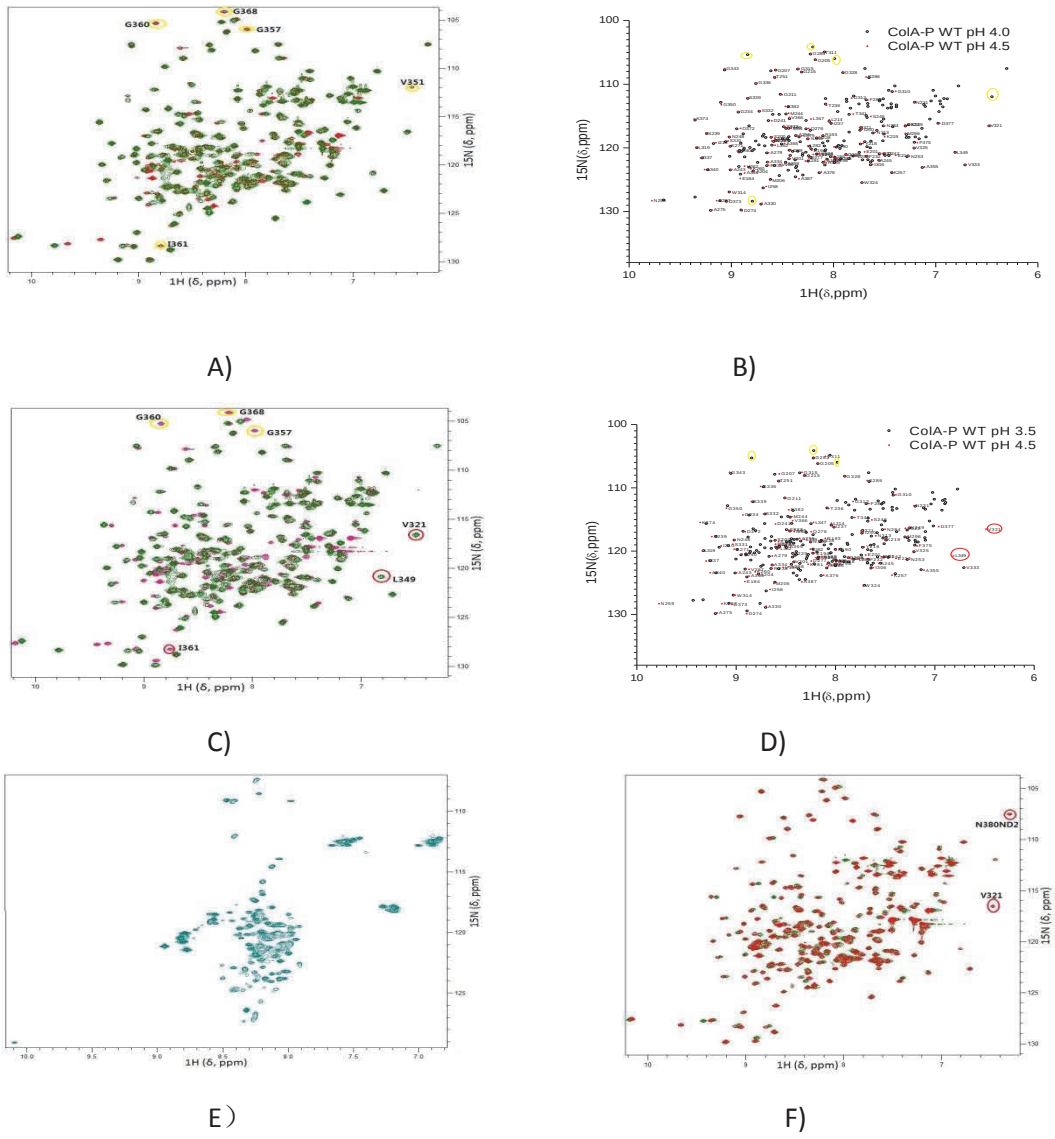


Figure 6.15. The crystal structure of ColA-P domain (PyMOL) highlight with the residues with chemical shifts(Magenta) the peaks missing on the spectra of ColA-P D274 at pH 8.0 (Red).

NMR spectra of ColA-P WT between pH 3.0 and pH 4.0

The previous chapters showed that colicin A pore-forming domain unfolded to a molten globule state at low pH. In order to assess the effect of pH on the structure of the protein, the ^{15}N labelled ColA-P WT was prepared as before and the ^{1}H-^{15}N-HSQC spectra collected at various pH values.



* The enlarged figures are appended on the following pages.

Figure 6.16. A) Overlaid 800 MHz 1-H-15-N-HSQC spectrum of ColA-P WT at pH 4.5 (Green) and pH 4.0 (Red) B) Overlaid peaks of ColA-P WT with assignments at pH 4.5 (red cross) and at pH 4.0 (Black Circle) C) Overlaid 800 MHz ^{1}H-^{15}N-HSQC spectrum of ColA-P WT at pH 4.5 (Green) and at pH 3.5 (Pink) D) Overlaid peaks of ColA-P WT with assignments (Red cross) at pH 4.5 and at pH 3.4 (Black circle) E) 800 MHz ^{1}H-^{15}N-HSQC spectrum of ColA-P WT at pH 3.0 F) Overlaid 800 MHz ^{1}H-^{15}N-HSQC spectrum of ColA-P WT at pH 4.0 (Green) and ColA-P WT at pH 3.5 (Green)

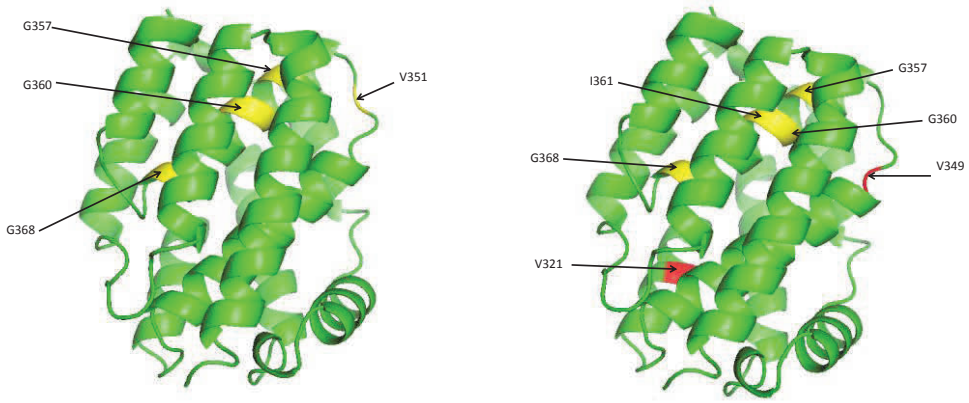
Figure 6.16 A) and C) show that the spectra of ColA-P WT at pH 3.5, pH 4.0 and pH 4.5 are very similar to each other, which means the structure of the wild type protein is little changed between pH 3.5 and 4.5.

The main differences between the spectra of ColA WT at pH 3.5, pH 4.0 and pH 4.5 is that the peaks of residues V351, G357, G360 and G368, located in the hydrophobic core of the protein, are absent from the spectrum of the protein at pH 4.5 but appear at pH 3.5 and 4.0.

In addition the peaks of residues V321 and V349 are missing from the spectrum at pH 3.5. V349 is located on the loop between the hydrophobic helices. V321 located on the interface between helix VII and the hydrophobic hairpin of the protein.

Figure 6.18 shows a spectrum of ColA-P WT at pH 3.0, but it is poorly dispersed making it impossible to obtain resonance assignments by inspection as was done above for the structured states of the wild type and mutant proteins. Although the CD spectra (Figure 3.14) show that the wild type protein loses only half of its tertiary structure at pH 3.0, the HSQC spectrum shows that the resonances of the amide groups of the asparagine and glutamine residues were accumulated into one ill-defined signal with poor dispersion on the proton axis. These characteristics are typical of a molten globule state of a protein in which variable dynamics can lean to sharp resonances and broad resonances co-existing even for residues in secondary structure elements (Dyson and Wright, 2004).

Figure 6.18 F) shows that the spectra of ColA-P WT at pH 3.5 and pH 4.0 are very similar apart from the peak of residue I361 appears only in the spectra at pH 3.5 The peaks which have shifted are highlighted and shown in the crystal structure of ColA-P wild type protein. (Figure 6.17)



A) B)

Figure 6.17. The crystal structure of ColA-P domain (PyMOL) highlighted (Yellow) with the residues that are not be observed on my spectrum of the wild type protein at pH 4.5 but which appear on the spectra of the wild type protein at pH 4.0 A) and pH 3.5 B). The missing peaks on the spectra of ColA-P WT at pH 8.0 are highlighted in Red colour.

Figure 6.17 A) and B) show that the hydrophobic core ColA-P of the protein is largely destabilized by the acidic environment. Although V321 is not located on the hydrophobic hairpin, it faces toward the hydrophobic part of the protein.

Since there is a large change in the NMR spectra of the wild type protein between pH 3.0 and pH 3.5 more data in this region should be collected in the future. Unfortunately, I would not be able to finish this study due to the time limitation of my PhD study.

NMR spectra of ColA-P D216A, D227 and D274A at pH 3.0

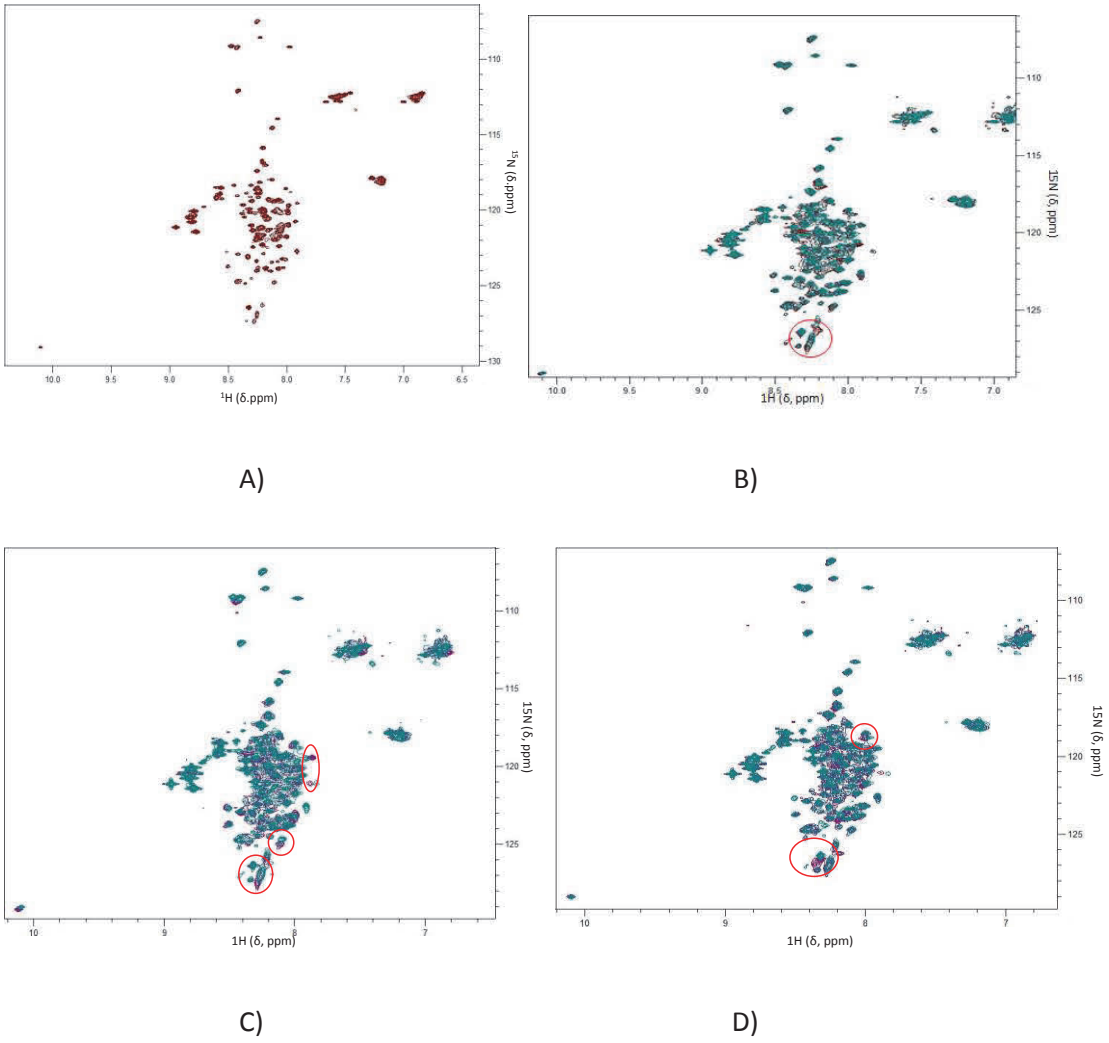


Figure 6.18. A) 800 MHz 1H-15N-HSQC spectrum of ColA-P WT at pH 3.0 (Spectra of the wild type protein at pH 3.0 from figure 6.16 shown for comparison) B) Overlaid 800 MHz 1H-15N-HSQC spectra of ColA-P WT (Dark Green) and D216A (Dark red) at pH 3.0 C) Overlaid 800 MHz 1H-15N-HSQC spectra of ColA-P WT (Dark Green) and D227A (Purple) at pH 3.0 D) Overlaid 800 MHz 1H-15N-HSQC spectra of ColA-P WT (Dark Green) and D274A (Purple) at pH 3.0

Figure 6.18 B) shows that the ^{1}H-^{15}N-HSQC spectra of ColA-P WT and D216A at pH 3.0 are very similar to each other. Only the area in the red circle shows some

differences. We are unable to tell which region of the protein correlated to the differences of the spectra; at least it shows that the conformational change of certain regions of the protein cause the differences of the spectra.

Figure 6.18 C) and D) show that the ^{1}H-^{15}N-HSQC spectra of ColA-P WT, D227A and D274A at pH 3.0 are very similar as each other. Again the areas in the red circles on figures 6.18 C and D show the obvious differences between the two spectra.

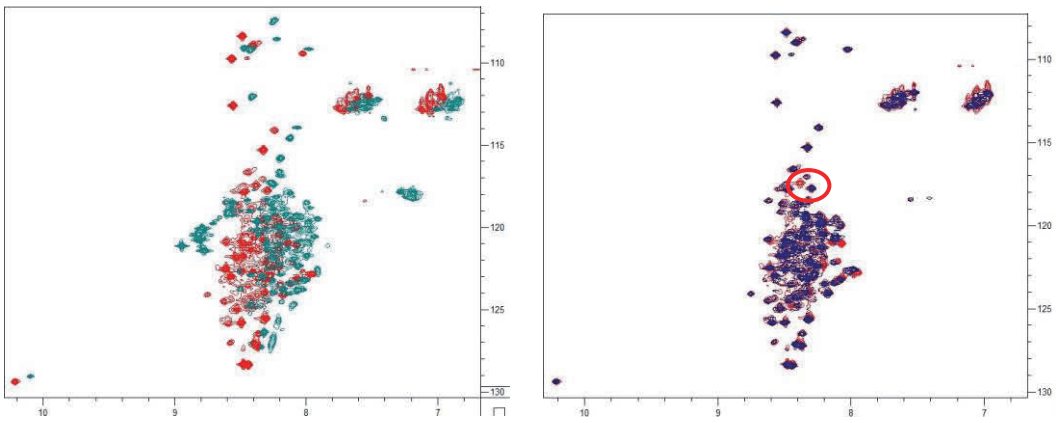
Figure 6.18 shows that ColA-P D227A at pH 3.0 is shows the most difference from ColA-P WT and D216A. According to the CD spectra, the wild type protein is more structured than both the alanine mutants at pH 3.0 but here the bigger difference between the NMR spectra of ColA-P WT and D227A means that this mutant is even less structured than D216A.The pH dependent urea unfolding experiment shows that D227A has already forms to its molten globule stage at pH 3.0, but ColA D216A forms to it molten globule stage between pH 2.5 and pH 3.0. It means ColA-P D216D should be more structured than D227A at pH 3.0 and the NMR spectra seem to agree with the earlier results.

Comparing with ColA-P D216A, the ^{1}H-^{15}N-HSQC spectrum of D274A is more similar to the wild type protein at pH 3.0. The area in the red circle shows slight differences between the spectra. The urea unfolding experiment shows that unlike ColA-P WT and D216A, ColA-P D274A does not unfold to its molten globule state at pH 3.0. The ^{1}H-^{15}N-HSQC spectra do not show many differences between the spectra of ColA-P WT and D274A at pH 3.0, which means ColA-P D274A is more structured than D216A.

NMR spectra of ColA-P WT and D216A in 5.6 M urea at pH 8.0

Both ColA-P wild type and D216A unfold in two steps; a stable intermediate stage (molten globule state) appears at pH 3.0 or in 5.6 M urea at neutral pH. In order to

learn more about the local conformation change of the intermediate stage of ColA-P WT and D216A the HSQC spectra of ^{15}N-labeled samples were recorded in 5.6 M Urea solution at pH 7.0 and analyzed using CCPN.



A) B)

Figure 6.19. A) Overlaid 800 MHz 1H-15N-HSQC spectra of ColA-P WT in 5.6 M urea at pH 7.0 (Dark cyan) and in 0 M urea at pH 3.0 (Red) B) Overlaid 800 MHz ^{1}H-^{15}N-HSQC spectra of ColA-P WT (Red) and D216A (Navy) in 5.6 M urea at pH 7.0

Figure 6.19 shows that the ^{1}H-^{15}N-HSQC spectrum of ColA-P WT in 5.6 M urea at pH 7.0 and in 0 M urea at pH 3.0 are different with each other. The spectrum of the wild type protein in 5.6 M urea at pH 7.0 is even less dispersed than it is at pH 3.0. In the previous study we have shown that ColA-P WT forms an intermediate state in 5.6 M urea at pH 7.0 (Figure 3.2). The CD spectra (Figure 3.9) also show that the wild type protein retains part of its tertiary structure at pH 3.0. Now these NMR spectra confirm the differences between the intermediate state (ColA-P WT in 5.6 M urea at pH 7.0) and the molten globule state (ColA-P at pH 3.0).

The 1H-15N-HSQC spectrum of ColA-P D216A is highly resembled to the spectrum of the wild type protein in 5.6 M urea at pH 7.0. Only one resonance appears on the

spectrum of wild type protein at pH 3.0 and disappears on the spectra of ColA-P D216A. We cannot get more information from the poorly dispersed spectra.

NMR spectra of ColA-P WT in 7.4 M urea at pH 8.0

The urea unfolding experiment shows that ColA-P WT is fully unfolded in 7.4 M urea at pH 8.0. In order to study the local conformational changes of ColA-P protein in the from its molten globule state to its fully unfolded state, the HSQC spectra of 15N-labeled ColA-P D216A were also recorded in 5.6 M urea at pH 8.0 in 50 mM phosphate buffer and analyzed using CCPN. (Figure 6.19)

Figure 6.20 shows that the spectra of ColA-P WT in 5.6 M urea and in 7.4 M urea are similar as each other. However, the number of cross-peaks observed increases in higher concentration of urea. The peaks that are observed in the HSQC spectra in increasing concentrations of urea correspond to peaks from the unfolded form of ColA-P WT. The increase in peak intensity corresponds to an increase in the population of unfolded protein with increasing urea concentration, it is clear that all of the expected peaks are observed in 7.4M urea (Greene et al., 2006).

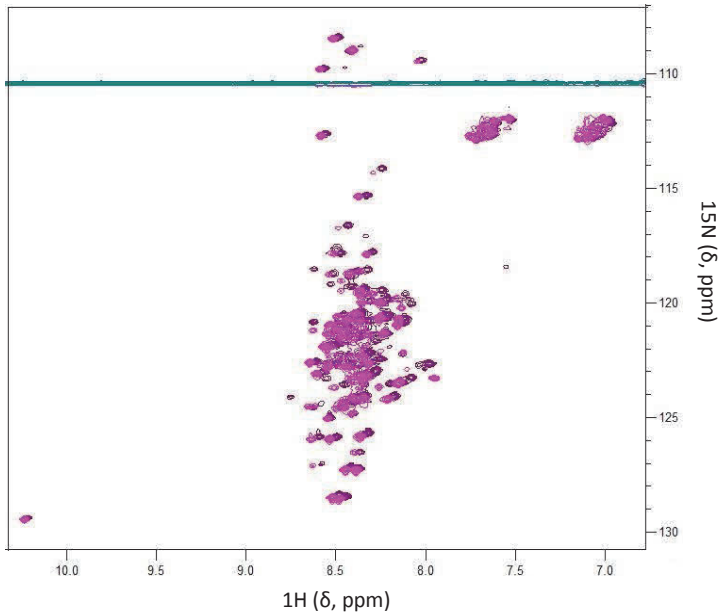


Figure 6.20. Overlaid ^{1}H-^{15}N-HSQC spectra of ColA-P WT in 5.6 M urea (Purple) and 7.4 M urea (Magenta) at pH 7.0

Detailed analysis of the chemical shifts.

Table 6.1: the chemical shift of the assigned residues

	Residues Residue type and number	Chemical shifts D216A at pH 4.5 VS WT at pH 4.5	Chemical shifts D274A at pH 4.5 VS WT at pH 4.5	Chemical shifts D216A at pH 8.0 VS WT at pH 4.5	Chemical shifts D274A at pH 8.0 VS WT at pH 4.5
loop	E184	*	*	*	*
	N206	+	*	+	--
	G207	-	---	+	+
I	I210	+	-	+	--
	G211	-	---	--	--
	E212	+	+	+	*
loop	G215	+	--	+	+
	N231	+	+	+	+
loop	G234	+	+	*	---
	S239	+	+	*	*
	F240	+	+	*	--
	D241	+	+	+	+
	A243	+	+	*	*
III	M244	+	+	*	*
	S246	+	+	+	+
	N248	+	+	+	+
	T251	+	+	+	+
	N253	+	+	+	+
	K257	+	---	*	+
	I258	+	*	+	+
	N259	-	---	*	*
	K260	+	--	*	+
IV	V267	+	*	*	*
	K271	+	*	--	*
loop	D272	+	+	+	--

Where * means the cross peak which exists in the NMR spectrum of the ColA-P WT at pH 4.5 is not visible in the corresponding spectrum in the column of the table; # means the cross peak which does not appear in the spectrum of the ColA-P at pH 4.5

appears in the corresponding spectrum on the column of the table; - means the cross peak with chemical shift; + means the cross peak without chemical shift.

	Residues Residue type and number	Chemical shifts D216A at pH 4.5 VS WT at pH 4.5	Chemical shifts D274A at pH 4.5 VS WT at pH 4.5	Chemical shifts D216A at pH 8.0 VS WT at pH 4.5	Chemical shifts D274A at pH 8.0 VS WT at pH 4.5
	A275	+	+	+	+
V	G283	+	+	-	-
	N284	+	+	+	+
loop	S286	+	-	-	-
	K287	+	--	*	*
VI	I306	+	+	-	-
	T311	+	-	-	--
loop	N313	-	--	*	-
	W314	+	+	-	-
	G315	+	+	-	-
	L319	+	-	*	*
VII	V321	+	+	--	--
	W324	+	+	*	*
	V325	+	+	--	--
loop	G328	+	+	-	-
	A330	+	+	-	-
	S331	+	+	--	---
	I332	+	-	-	-
	V333	+	+	+	+
	G336	+	--	+	-
VIII	I337	+	+	+	--
	S339	+	-	+	+
	A340	+	--	+	+
loop	L349	+	*	-	--
	G350	-	---	*	---

Where * means the cross peak which exists in the NMR spectrum of the ColA-P WT at pH 4.5 is not visible in the corresponding spectrum in the column of the table; # means the cross peak which does not appear in the spectrum of the ColA-P at pH 4.5

appears in the corresponding spectrum on the column of the table; - means the cross peak with chemical shift; + means the cross peak without chemical shift.

	Residues Residue type and number	Chemical shifts D216A at pH 4.5 VS WT at pH 4.5	Chemical shifts D274A at pH 4.5 VS WT at pH 4.5	Chemical shifts D216A at pH 8.0 VS WT at pH 4.5	Chemical shifts D274A at pH 8.0 VS WT at pH 4.5
	A355	+	---	-	--
IX	V366	+	--	-	*
loop	I371	+	-	+	+
	D372	-	--	--	+
	D373	+	-	--	*
X	K374	+	+	+	-
	D377	-	--	---	-
loop	A387	+	--	--	--

Where * means the cross peak which exists in the NMR spectrum of the ColA-P WT at pH 4.5 is not visible in the corresponding spectrum in the column of the table; # means the cross peak which does not appear in the spectrum of the ColA-P at pH 4.5 appears in the corresponding spectrum on the column of the table; - means the cross peak with chemical shift; + means the cross peak without chemical shift.

The changes in chemical shift is due to the conformation change of the protein. The missing residues could be caused by a process of slow conformation changes in the region (Kienker et al., 2002, Parker et al., 1989). The peaks which cannot be observed on my spectra of ColA-P WT at pH 4.5 but reappear on the spectra of the wild type protein and the mutants at pH 8.0 could be caused by the forming of structured conformation. In other words, a process of slow conformation changes could happen in the protein at pH 4.5, but it may not happen at pH 8.0. Only the cross peaks isolated from the non-crowded area of the spectra were selected and analyzed in table 6.1.

There are three important facts that can be concluded from table 6.1.

The first one is ColA-P D216A is more stable than D227A at both pH 4.5 and pH 8.0. The second one is that some helices, such as helix IV and IX, are more flexible than other helices; whereas some helices, such as helix V, VI and VIII, are more stable than the others. Helix IV of colicin A pore-forming domain is quite reminiscent of S4 helix of T-(translocation) domain of Ca^{2+} and Na^{+} channel, the positively charged S4 was proposed as the voltage sensor of the channel (Cramer et al., 1983). It might be the reason of the flexibility property of helix IV of ColA-P proteins Helix V contains an alignment of hydroxyl group-containing residues; Helix VIII contains an alignment of small uncharged side chain containing residues (Cramer et al., 1983). It might be the reason that those two alpha helices are more stable than the others. The most interesting point it that helix VIII and IX are located at the hydrophobic core of the protein, but their stability varies. Helix VIII is comparatively more stable than the other helices, but helix IX is very flexible. Thus, I suggest that a big conformation change could happen on the helix IV and IX in the ion channel forming process.

The third one is that the cross peaks which cannot be observed on the spectra of ColA-P proteins at pH 4.5 but appear on the spectra of ColA-P proteins at pH 8.0 all face towards to helix I; II and X. Helix I; II and X are not as stable as helix V and VI. Thus, I suggest that the acidic environment may change the stability of the hydrophobic core.

6.4 Discussion

According to the results that I have shown here, four valuable conclusions can be drawn. The first one is that my HSQC spectra of WT recorded at pH 4.5 are similar to the published data (ColA-P WT at pH 4.5). Although my protein samples and the samples that have been used for the published data are not exactly the same, 121 cross-peaks (60%) of my data are coincident with the published data. (195 residues were assigned in the published data) (Ibanez de Opakua et al., 2010)

The second one is that although HSQC spectra of ColA-P WT; D216A; D227A and D274A at pH 4.5 and pH 8.0 are similar to each other some resonances missing at pH 4.5 reappear at pH 8.0 (Figure 6.1, 6.2, 6.4 and 6.6). These correspond to a series of residues on the hydrophobic helix.

The third one is that at pH 8.0 the mutations D216A or D227A have similar effects to pH 4.5 on the structure of the ColA-P WT protein and are more destabilising than D274A. The spectra of ColA-P D274A and the wild type protein at pH 8.0 are similar to each other, which mean that ColA-P D274A does not destabilize the protein as D216A and D227A.

Finally, when a structured protein loses it tertiary and secondary structure, the local environment of protons become equal to each other, all the spectra of proton crowd together between 8 ppm and 9 ppm. It is clear that, the ColA-P protein partially retains its secondary structure at pH 3.0, and completely loses its secondary structure in 5.6 M urea and 7.4 M urea. This is important as the alpha helices are thought to be the main structural elements in the ion channel.

Colicin A pore forming domain is a compact and water-soluble protein, and the pathway of its transformation into a voltage-gated ion channel is not known (Cascales et al., 2007). The previous chapters reveal that some aspartate to alanine mutations and acidic environments reduce the protein stability, but the experiments results studied by CD, fluorescence and DSC cannot reveal the regional conformational changes of the protein. The summary of the chemical shifts of the assigned residues of the protein caused by either low pH or aspartate to alanine mutation at D216 and D227 can tell us some information about the stabilities of each helices of the protein. However, more NMR data will be needed to analyse the local conformationals changes of the molten globule stage of ColA-P proteins.

To conclude, NMR is a very sensitive technology to assess the regional structure changes of a protein, thus it is a very useful methodology in the study of protein

activity. ColA-P domain is a pH sensitive protein, it needs an acidic environment to perform it biological activity, and the NMR spectra result shown in this chapter gave us a general idea that the effect of some aspartate to alanine mutation and the acidic environment are similar as each other. To form the ion channel, some helices, such as helix IV and IX, need more conformational changes, and some helices, such as helix V, VI and VIII, almost need no conformational change to form the ion channel.

Chapter 6 Appendix . Enlarged Figures

The NMR spectra presented in Chapter 6 are all reproduced on the following pages at larger scale for clarity. The figure numbers are the same as those used in the main text.

Figure 6.1 A

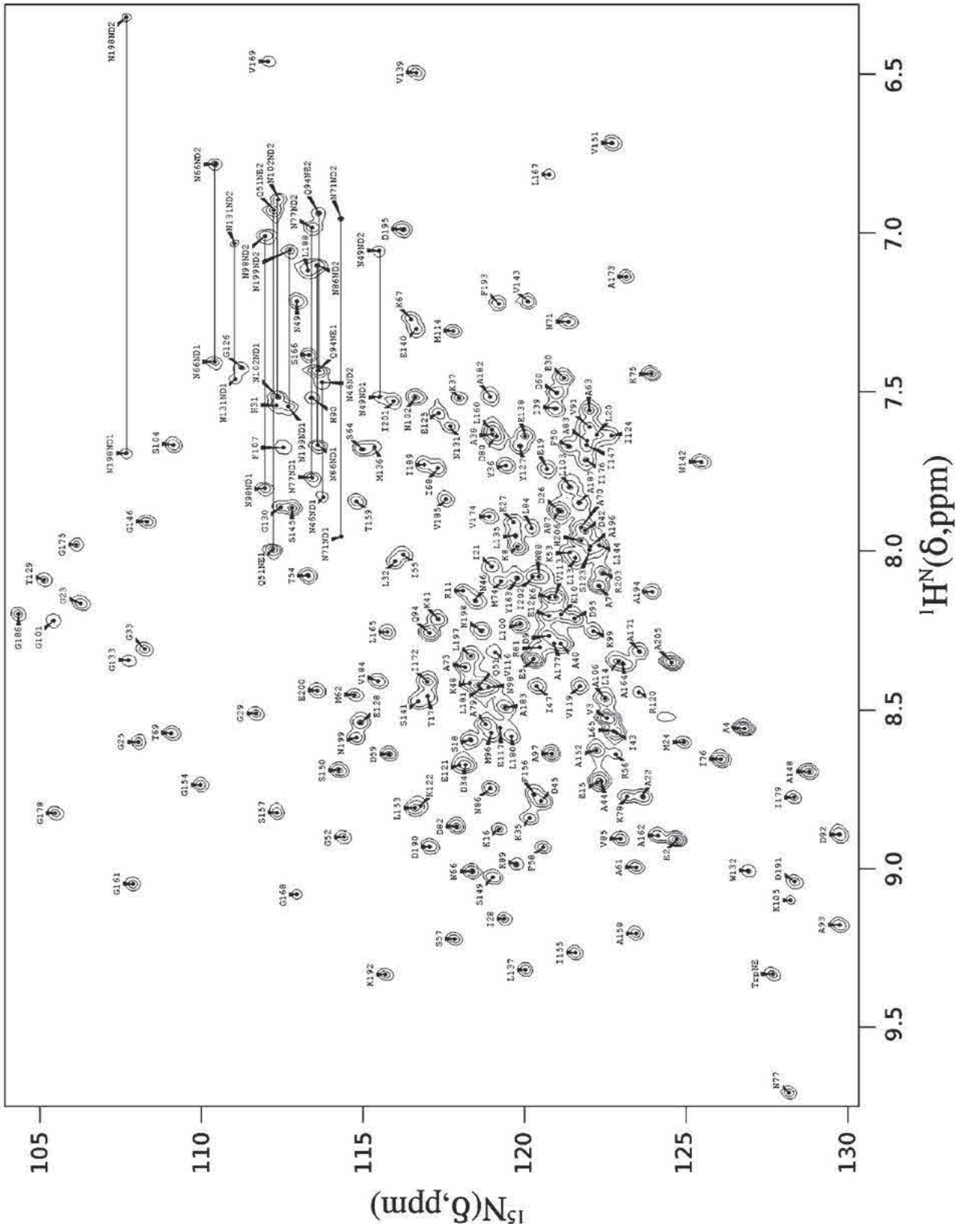


Figure 6.1B

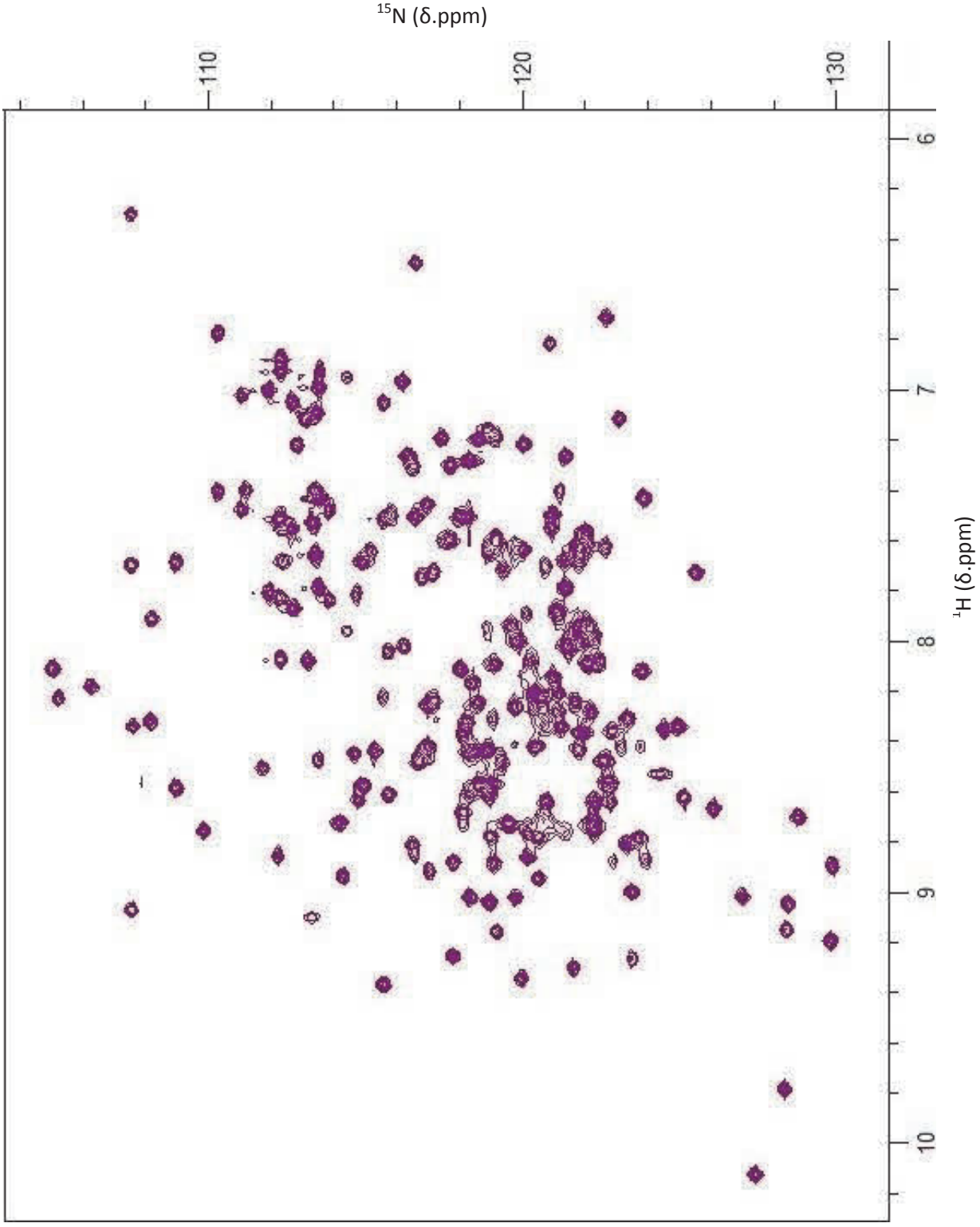


Figure 6.1 C

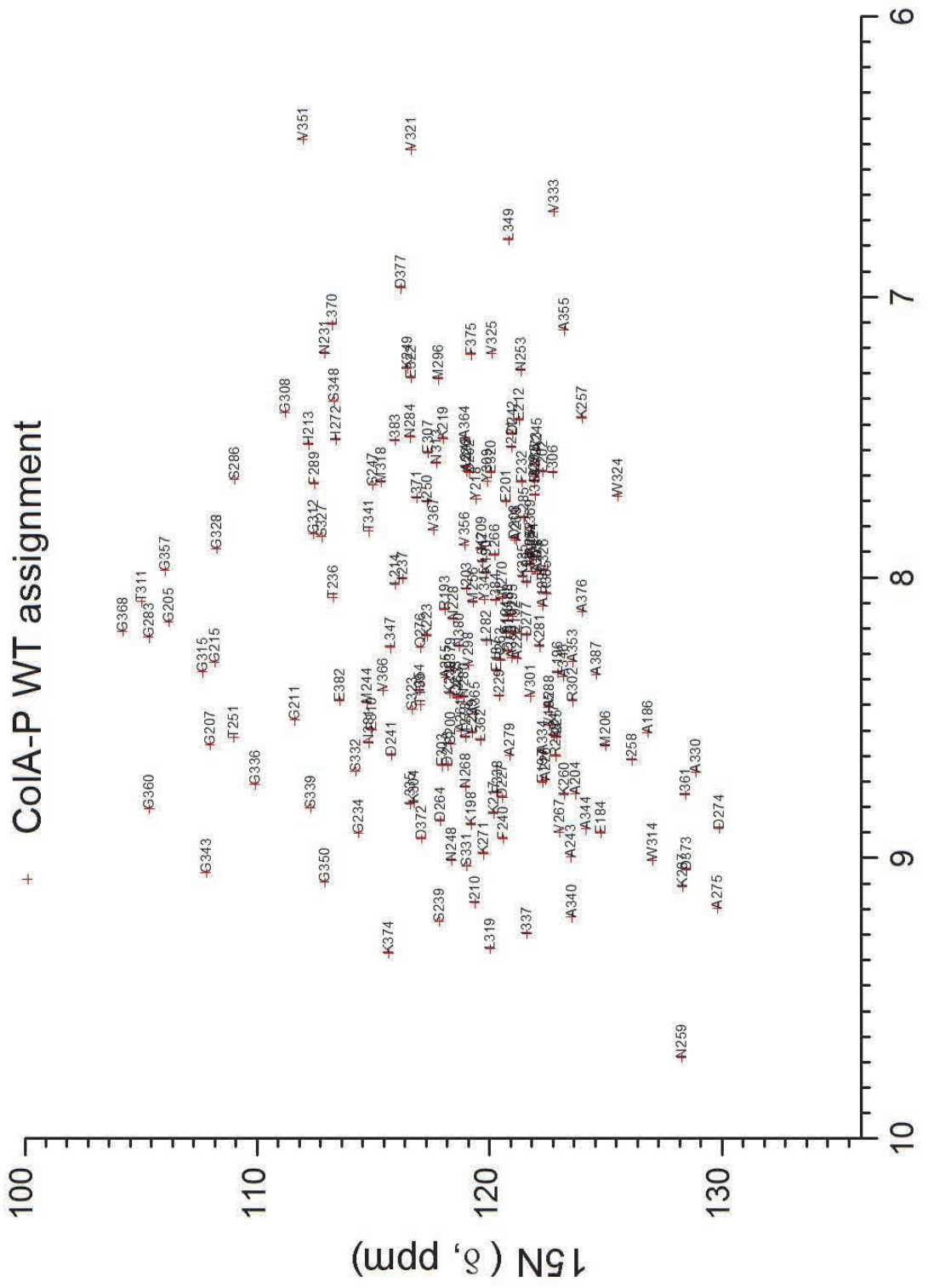


Figure 6.1 D

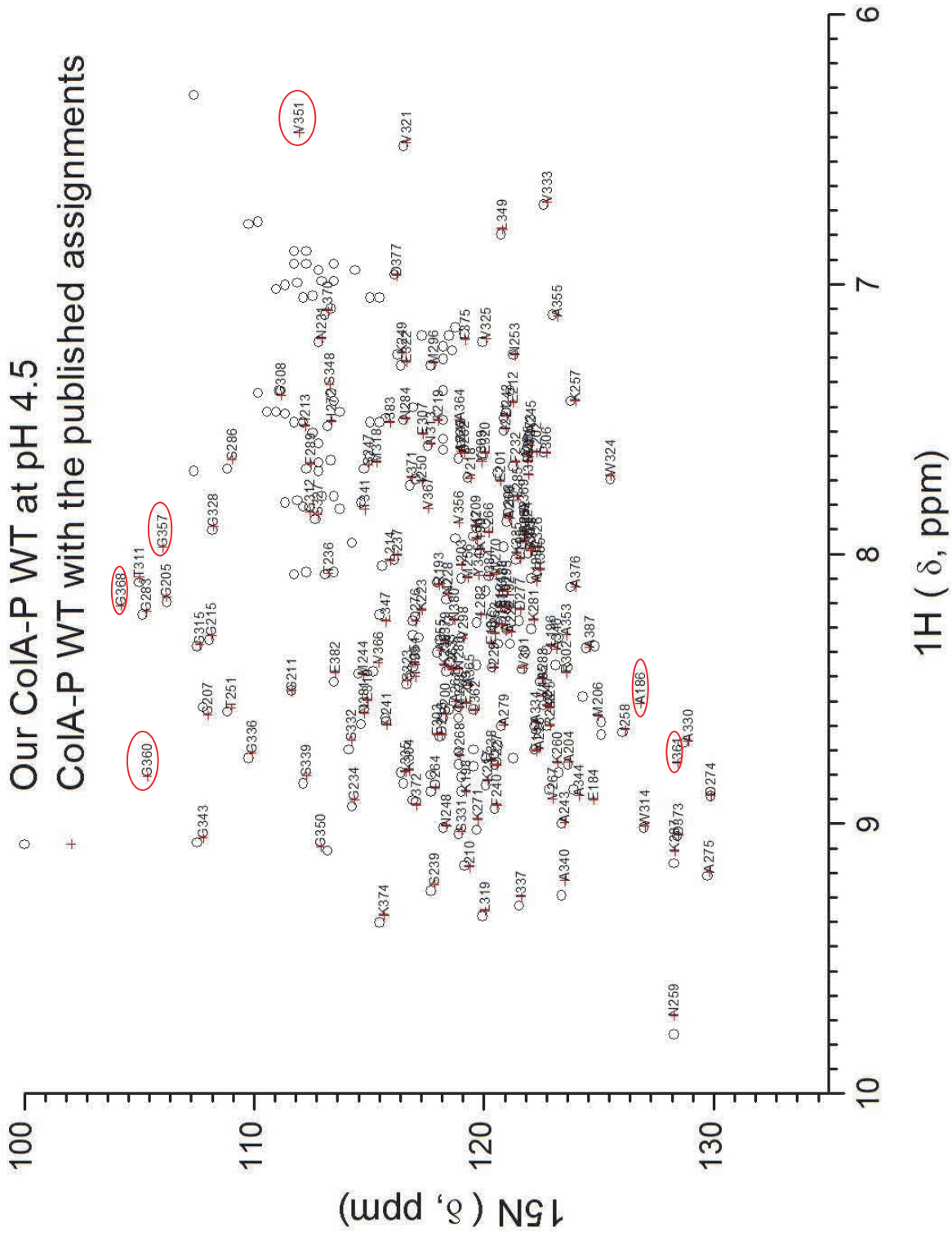

Our ColA-P WT at pH 4.5
ColA-P WT with the published assignments
15N (δ, ppm)
1H (δ, ppm)
100
110
120
130
6
7
8
9
10

Figure 6.1 E

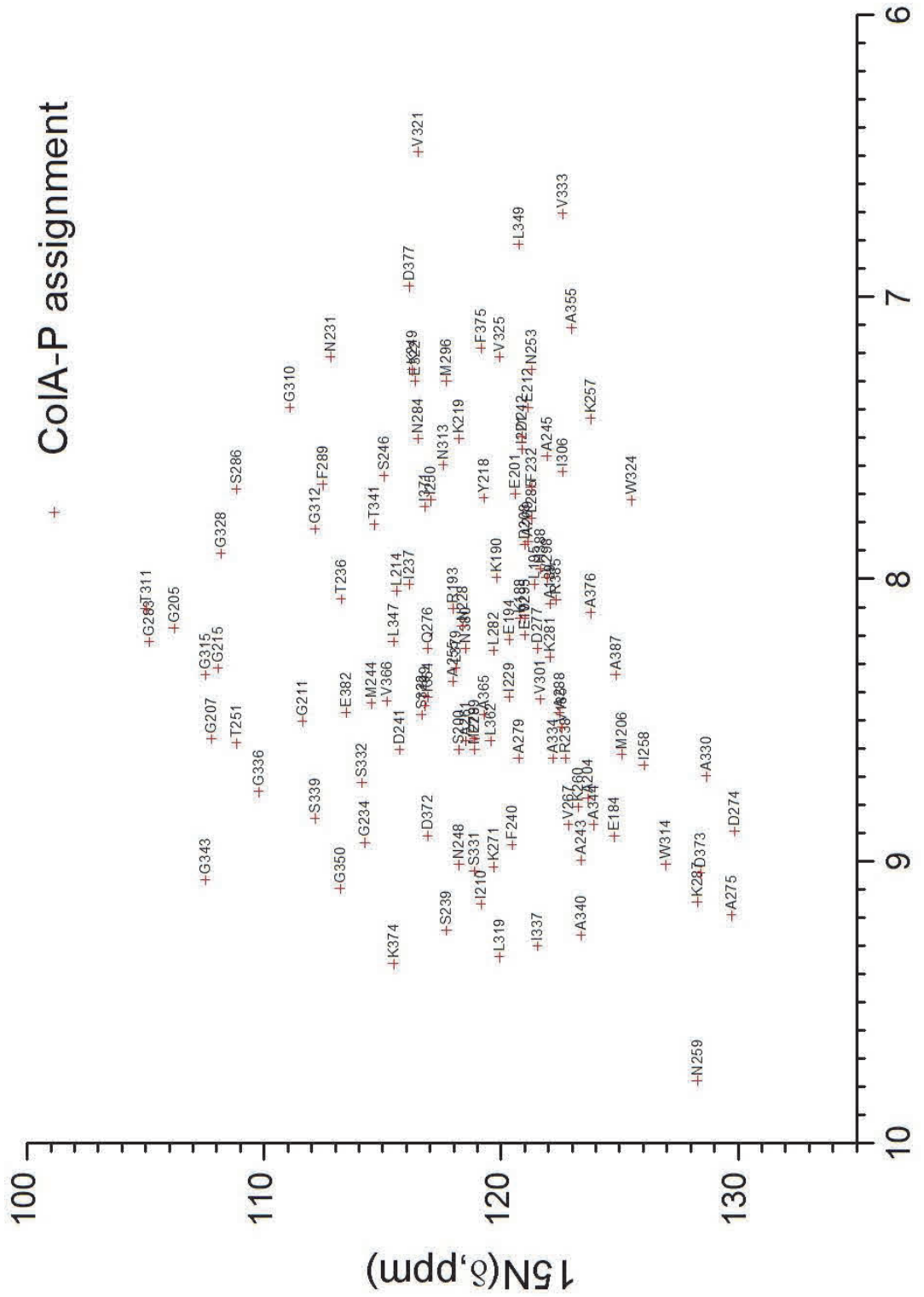


Figure 6.2 A

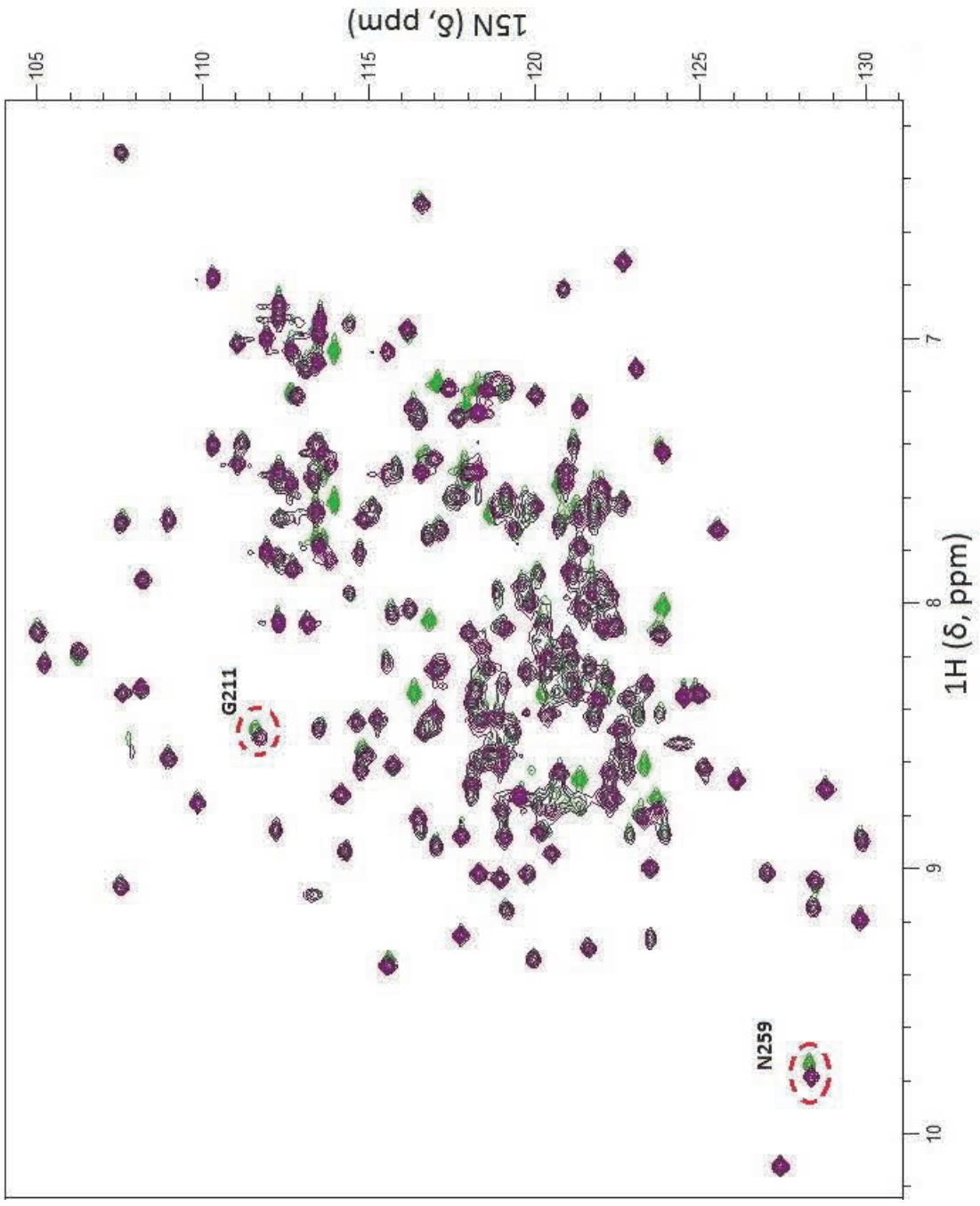


Figure 6.2 B

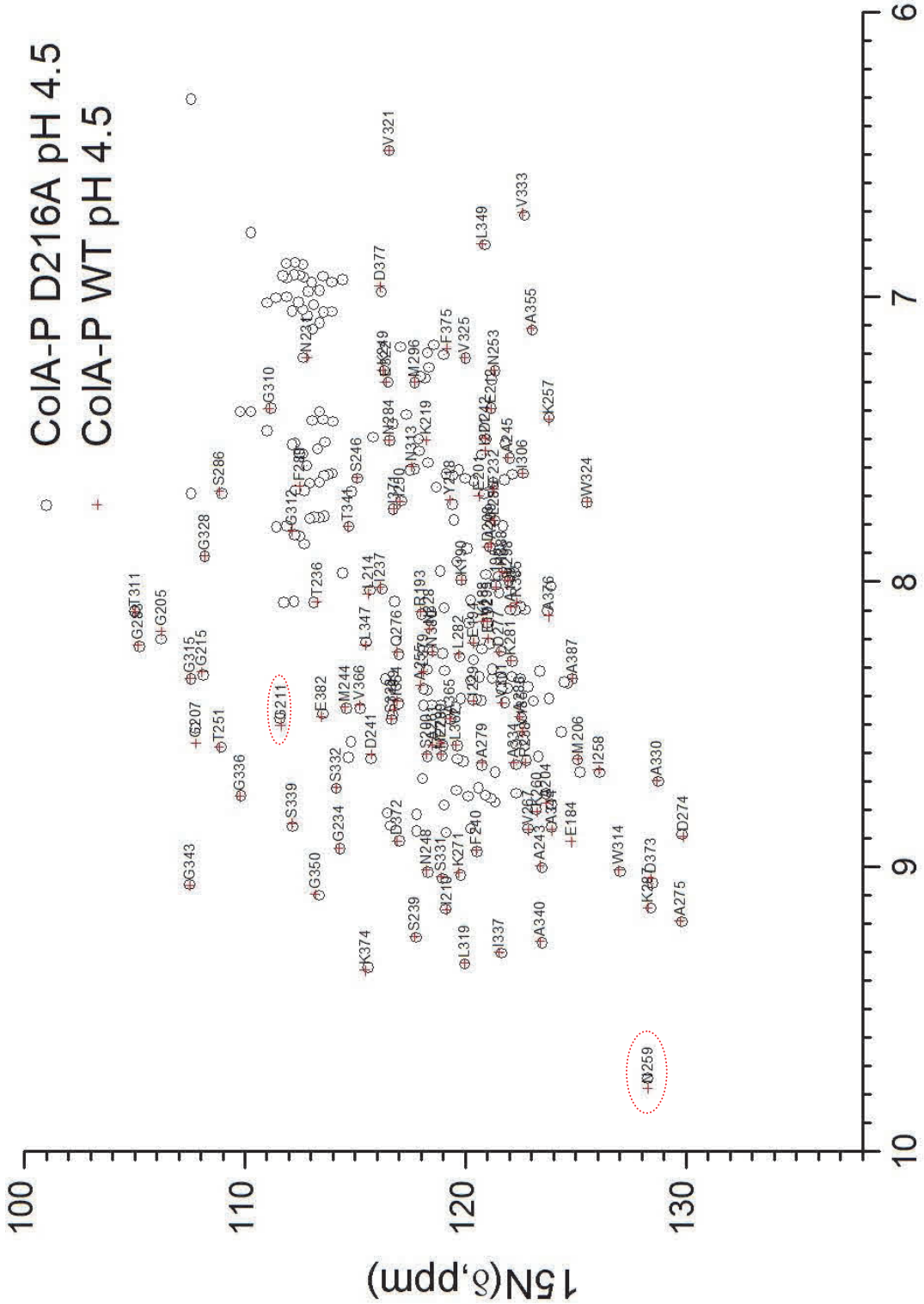

ColA-P D216A pH 4.5
ColA-P WT pH 4.5
15N(δ,ppm)
1H(δ,ppm)
100
110
120
130
6
7
8
9
10

Figure 6.4 A

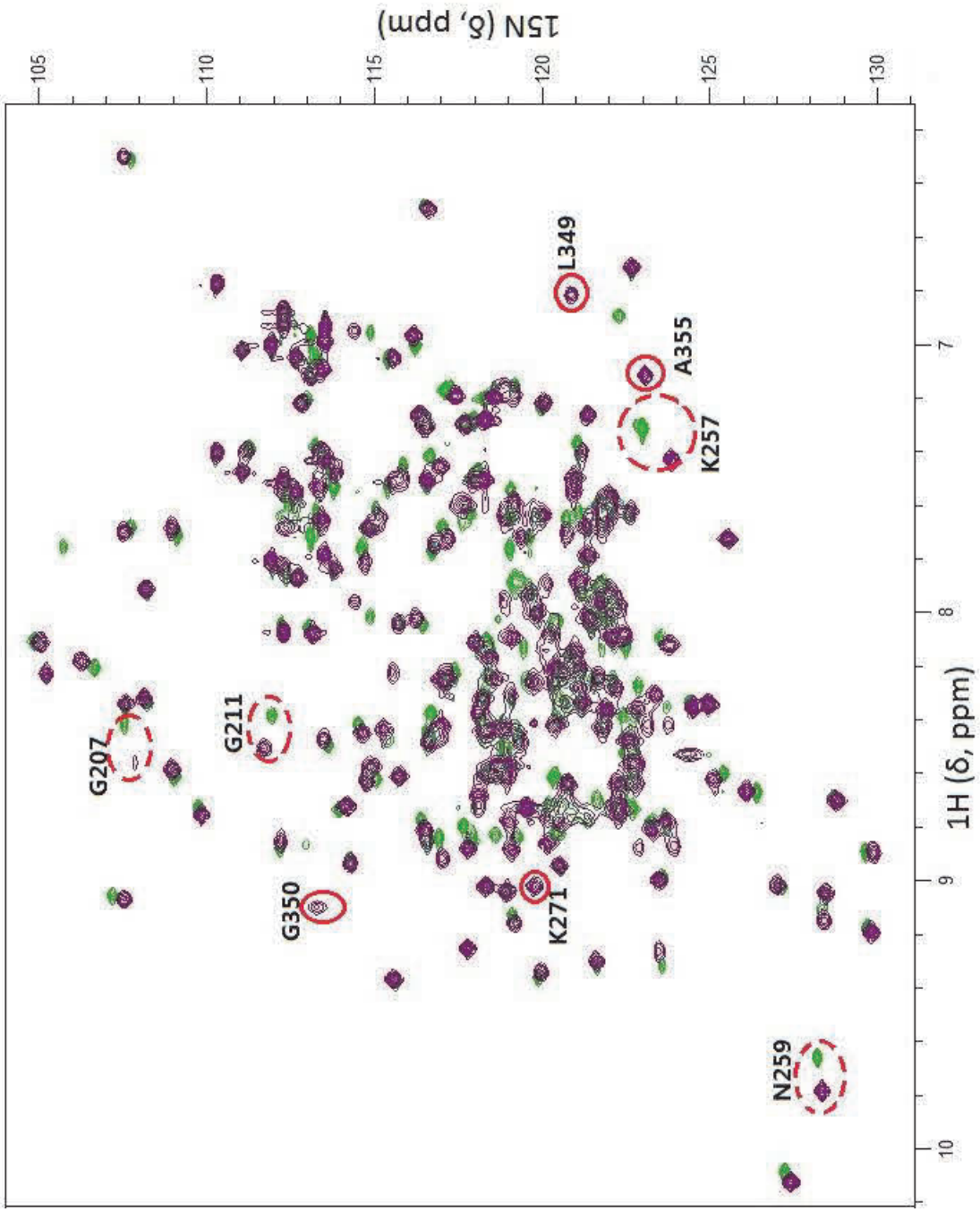

15N (δ, ppm)
105
110
115
120
125
130
1H (δ, ppm)
6
7
8
9
10
L349
A355
K257
G207
G211
G350
K271
N259

Figure 6.4 B

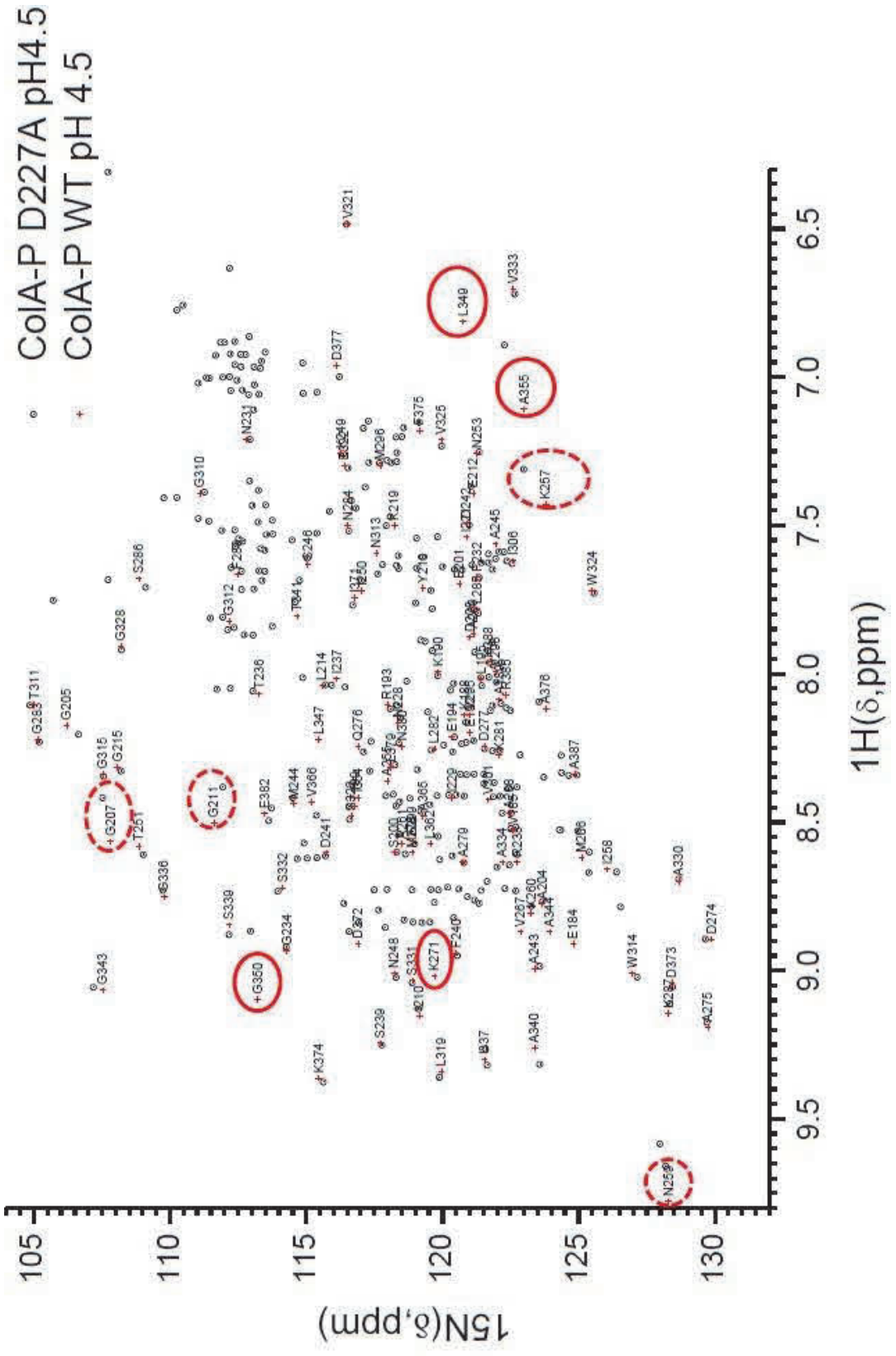

ColA-P D227A pH4.5
ColA-P WT pH 4.5
15N(δ,ppm)
1H(δ,ppm)
105
110
115
120
125
130
9.5
9.0
8.5
8.0
7.5
7.0
6.5
G350
G211
G207
K271
A355
L340
K257

Figure 6.6 A

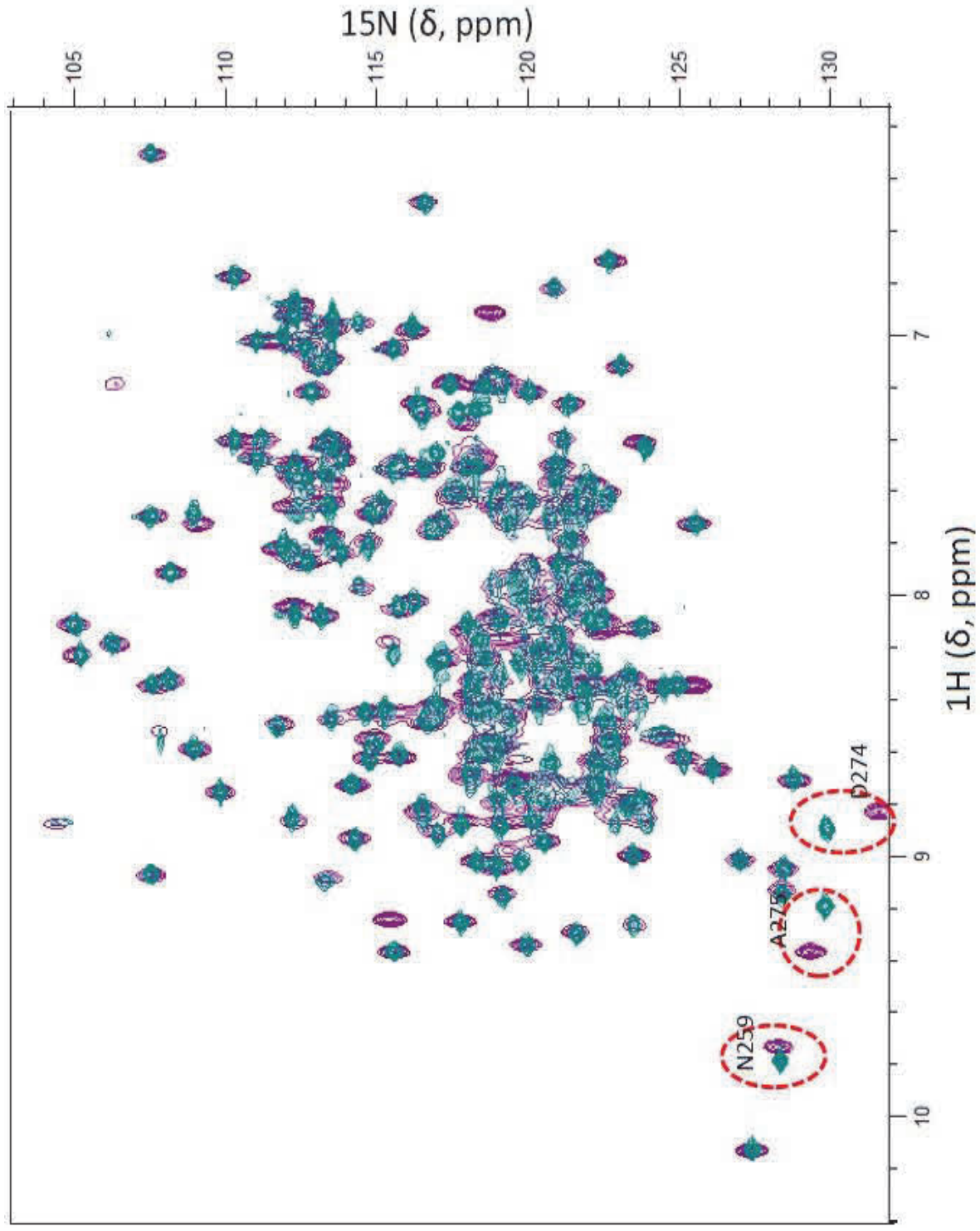


Figure 6.6 B

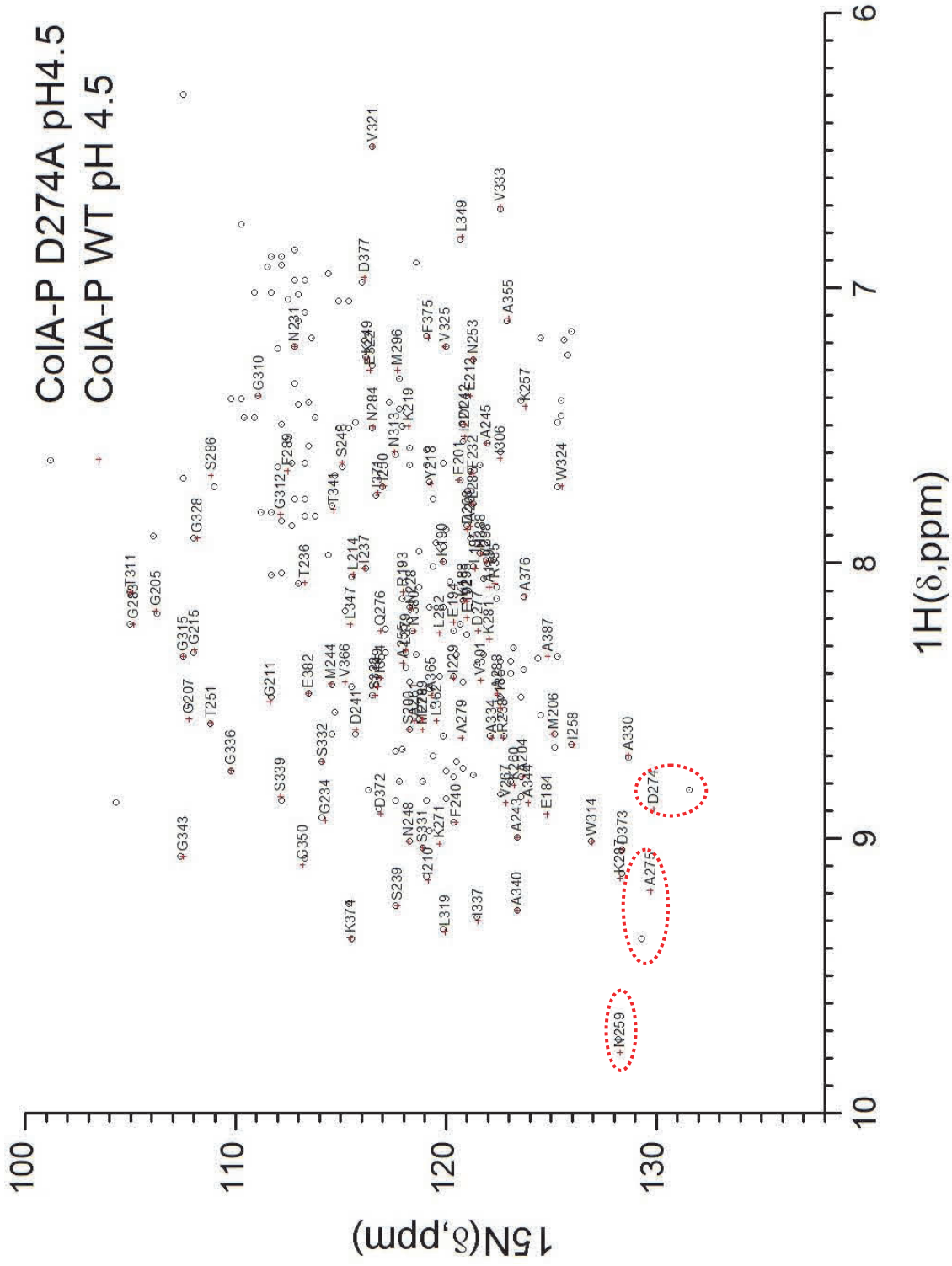

ColA-P D274A pH4.5
ColA-P WT pH 4.5
15N(δ,ppm)
1H(δ,ppm)
100
110
120
130
6
7
8
9
10

Figure 6.8 A

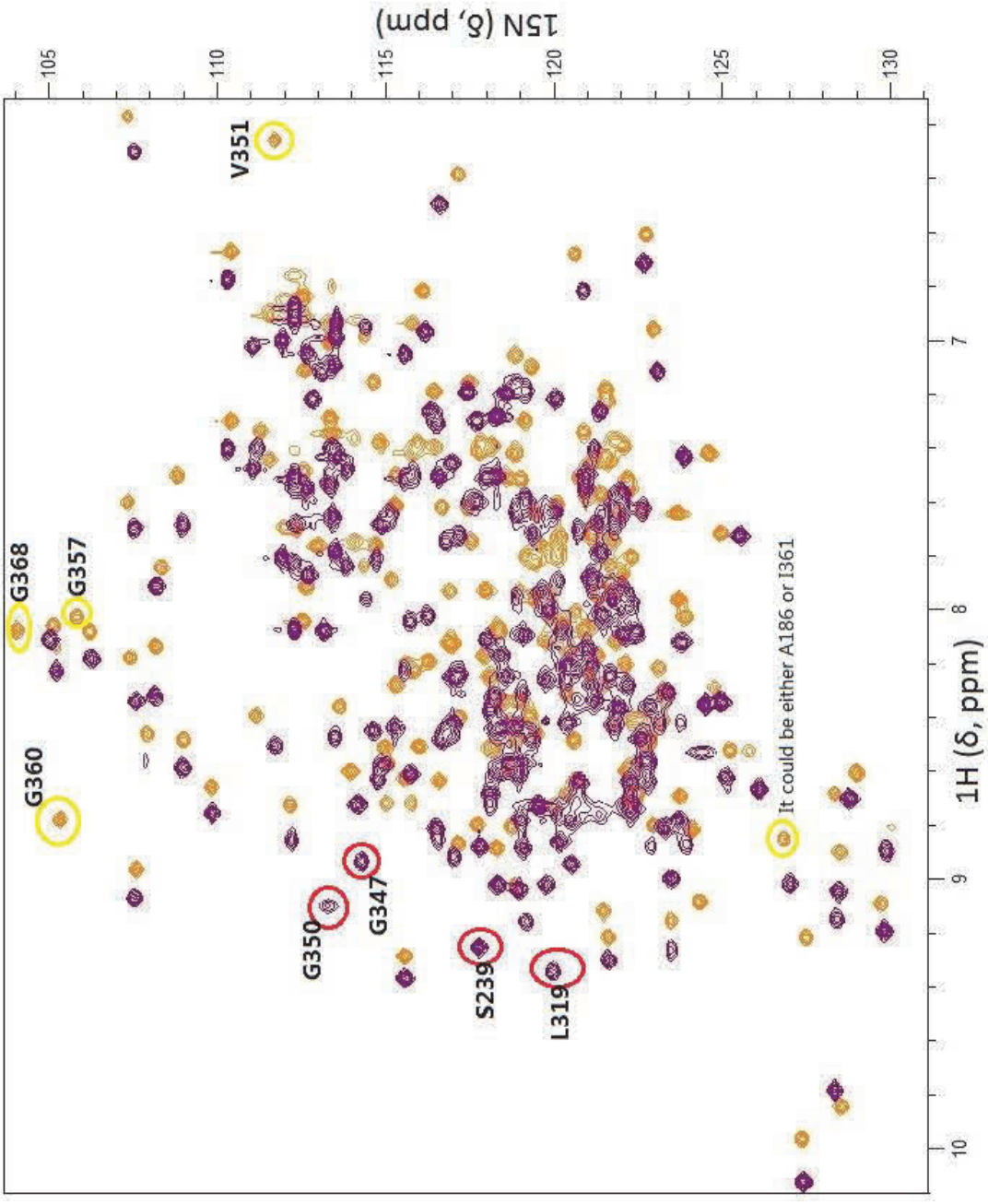

15N (δ, ppm)
105
110
115
120
125
130
1H (δ, ppm)
7
8
9
10
V351
G368
G357
G360
G350
G347
S239
L319
It could be either A186 or I361

Figure 6.8 B

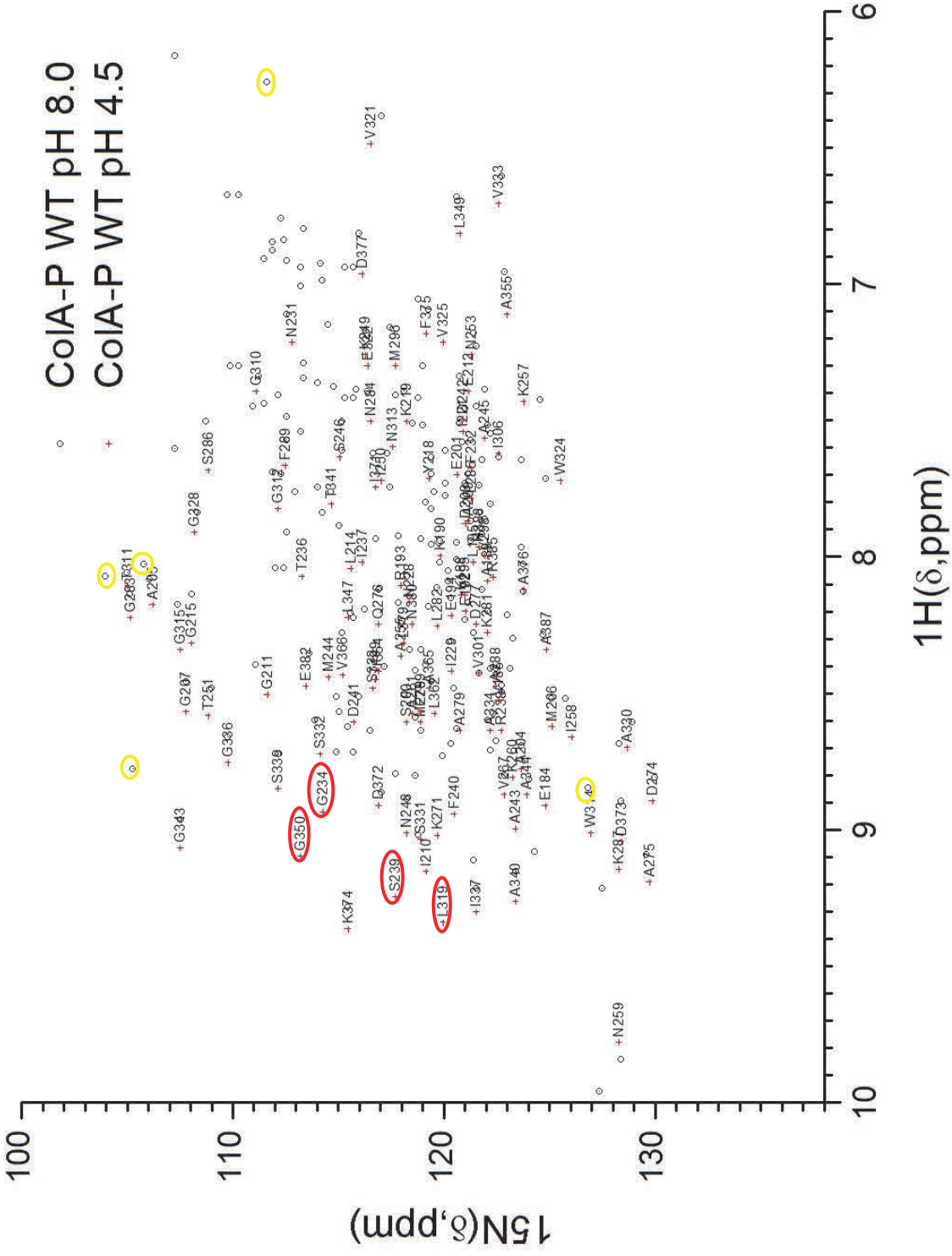

ColA-P WT pH 8.0
ColA-P WT pH 4.5
15N(δ,ppm)
1H(δ,ppm)
100
110
120
130
6
7
8
9
10

Figure 6.10 A

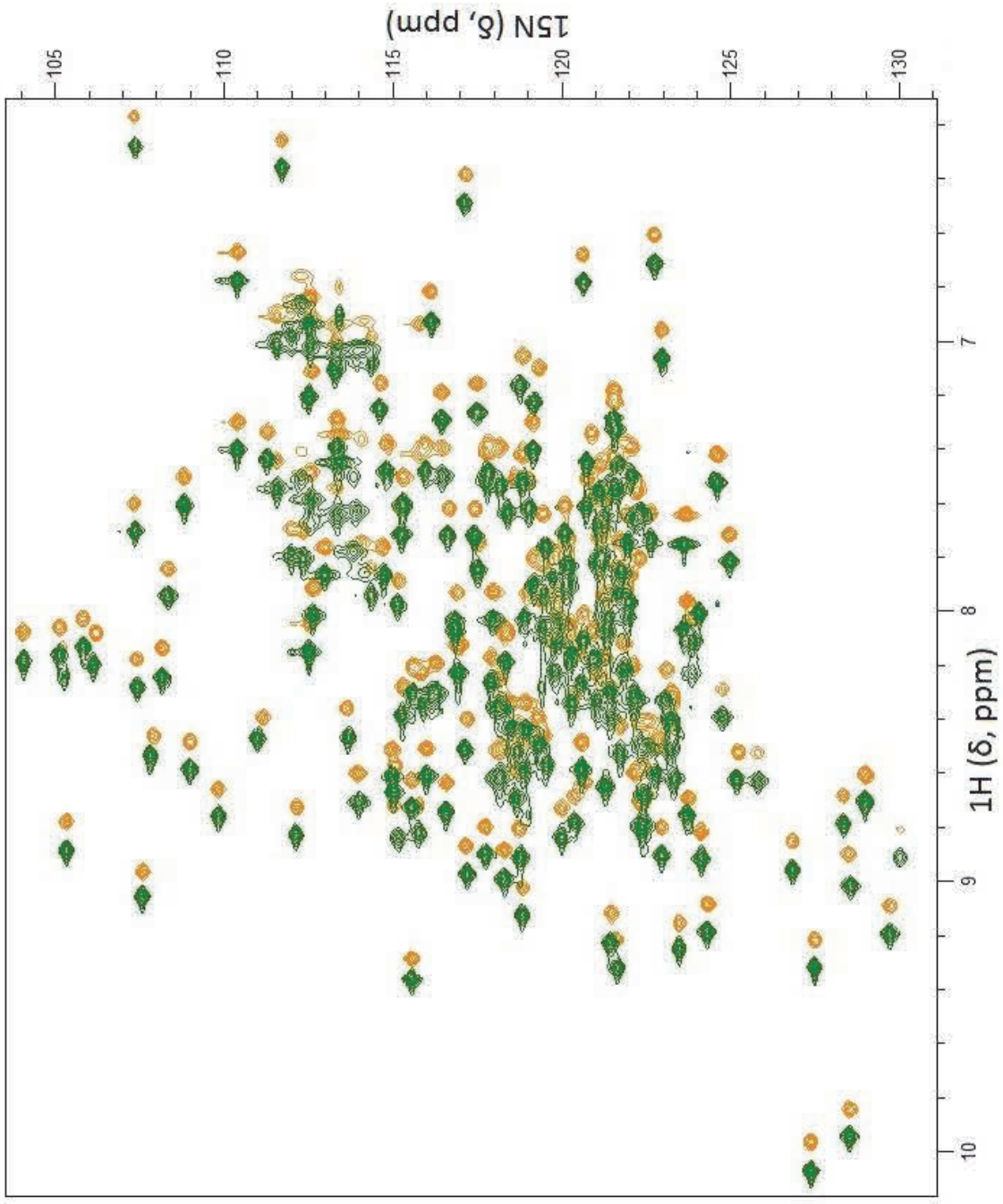


Figure 6.10 B

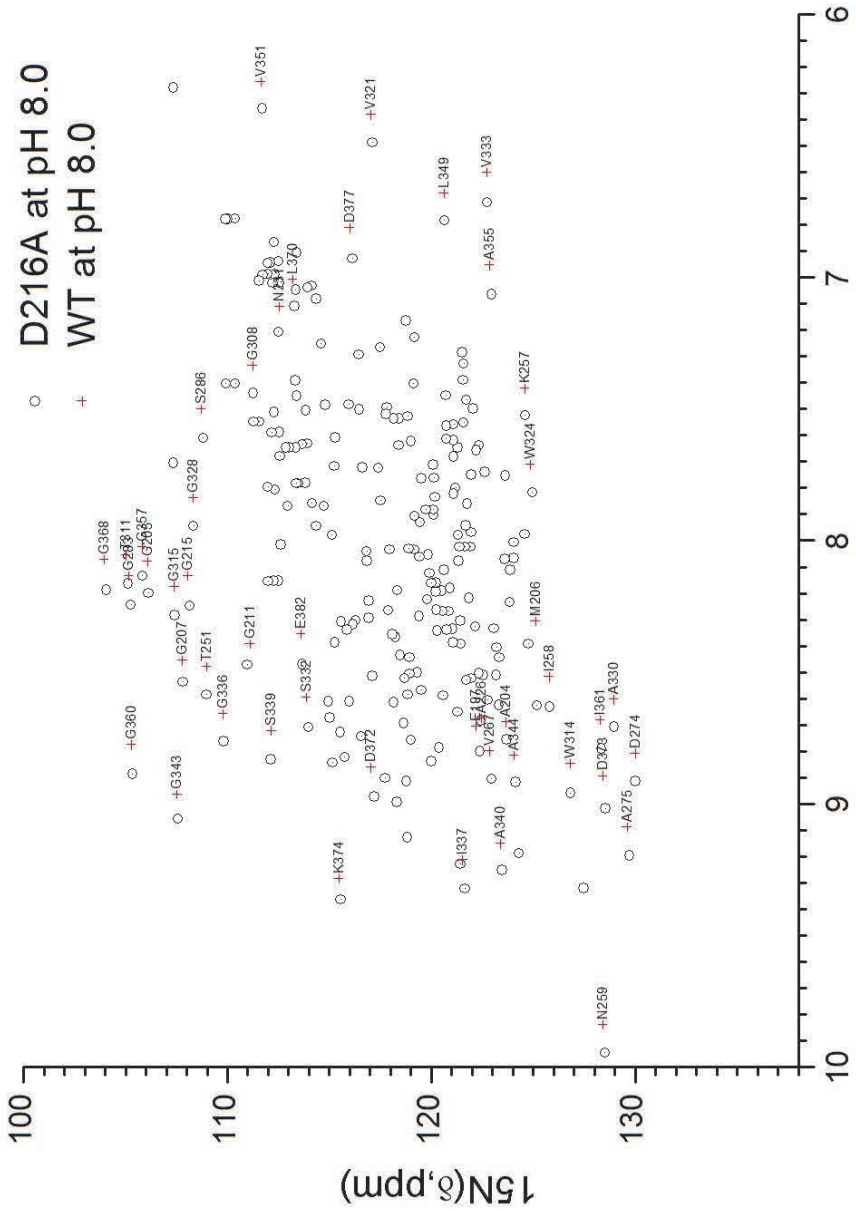

D216A at pH 8.0
WT at pH 8.0
15N(δ,ppm)
1H(δ,ppm)
100
110
120
130
6
7
8
9
10
+G368
+G360
+G343
+G207
+T251
+G336
+G211
+S339
+S332
+E382
+K374
+D372
+G315
+G215
+G328
+S286
+G308
+V351
+L370
+D377
+V321
+L349
+V333
+A355
+K257
+W324
+M206
+I258
+W314
+I361
+A330
+D274
+A275
+A340
+I337
+N259

Figure 6.10 C

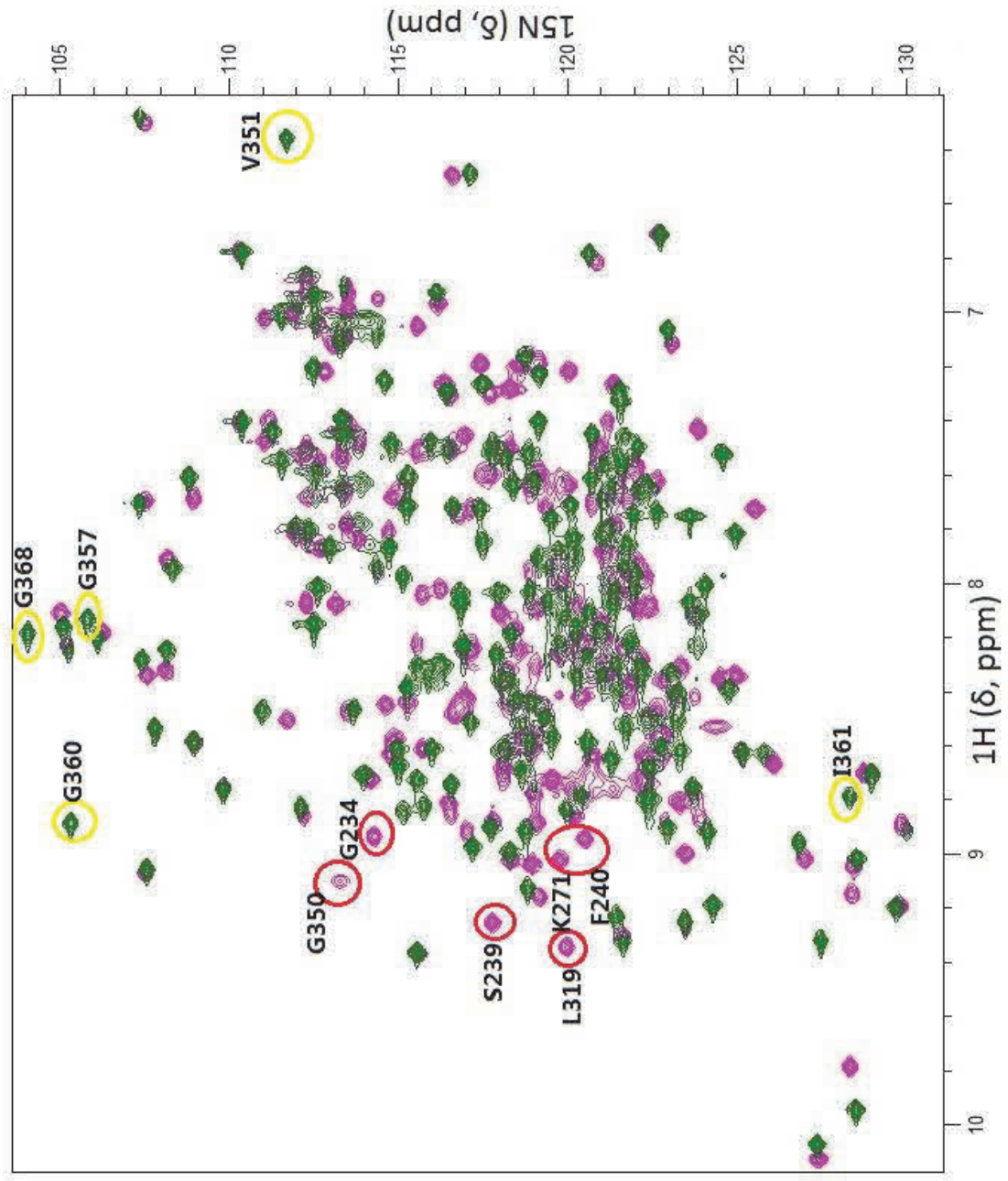

15N (δ, ppm)
105
110
115
120
125
130
1H (δ, ppm)
7
8
9
10
V351
G368
G357
G360
G234
G350
S239
L319
K271
F240
I361

Figure 6.10 D

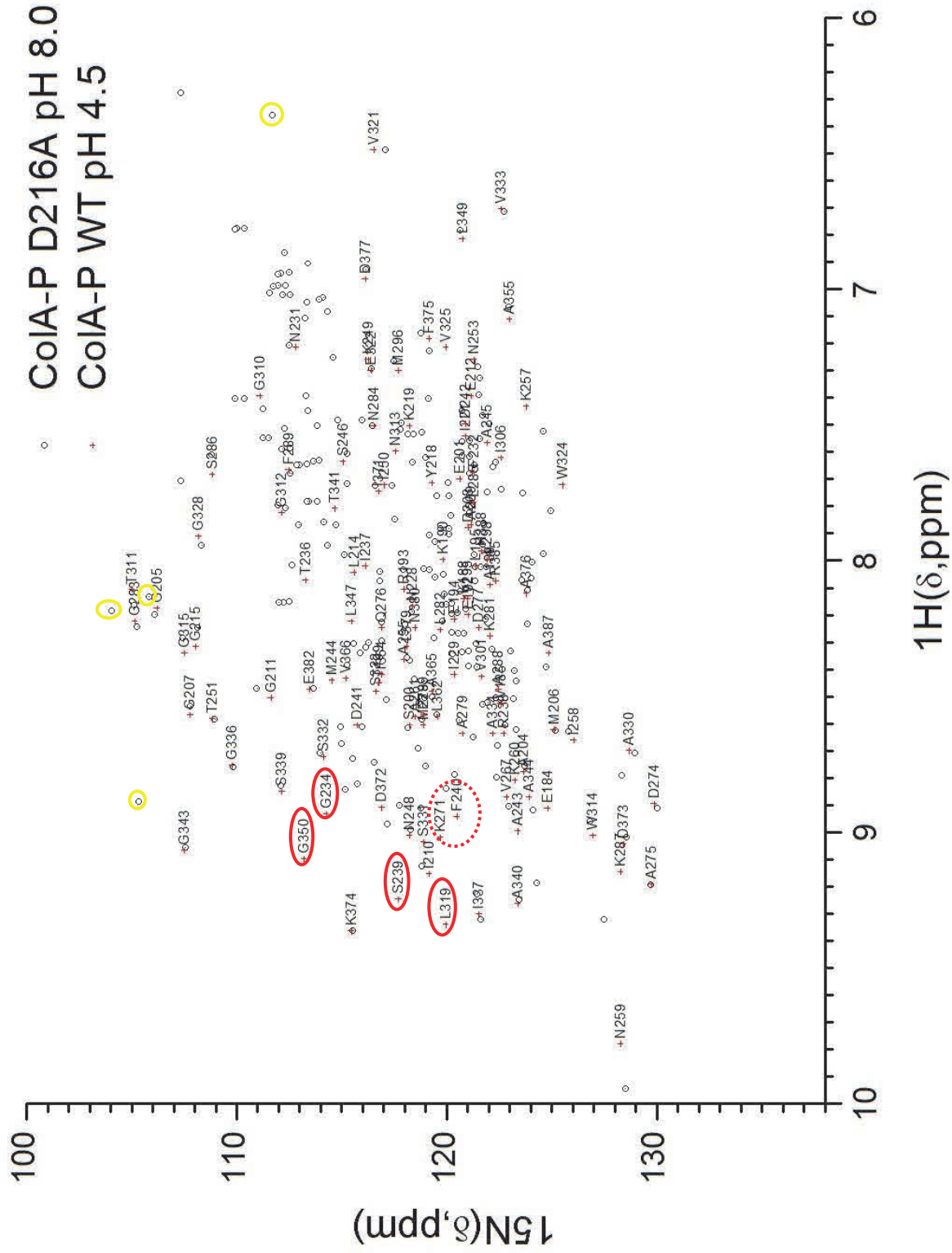

ColA-P D216A pH 8.0
ColA-P WT pH 4.5
15N(δ,ppm)
1H(δ,ppm)
100
110
120
130
6
7
8
9
10
G350
S239
L319
G234
F240

Figure 6.12 A

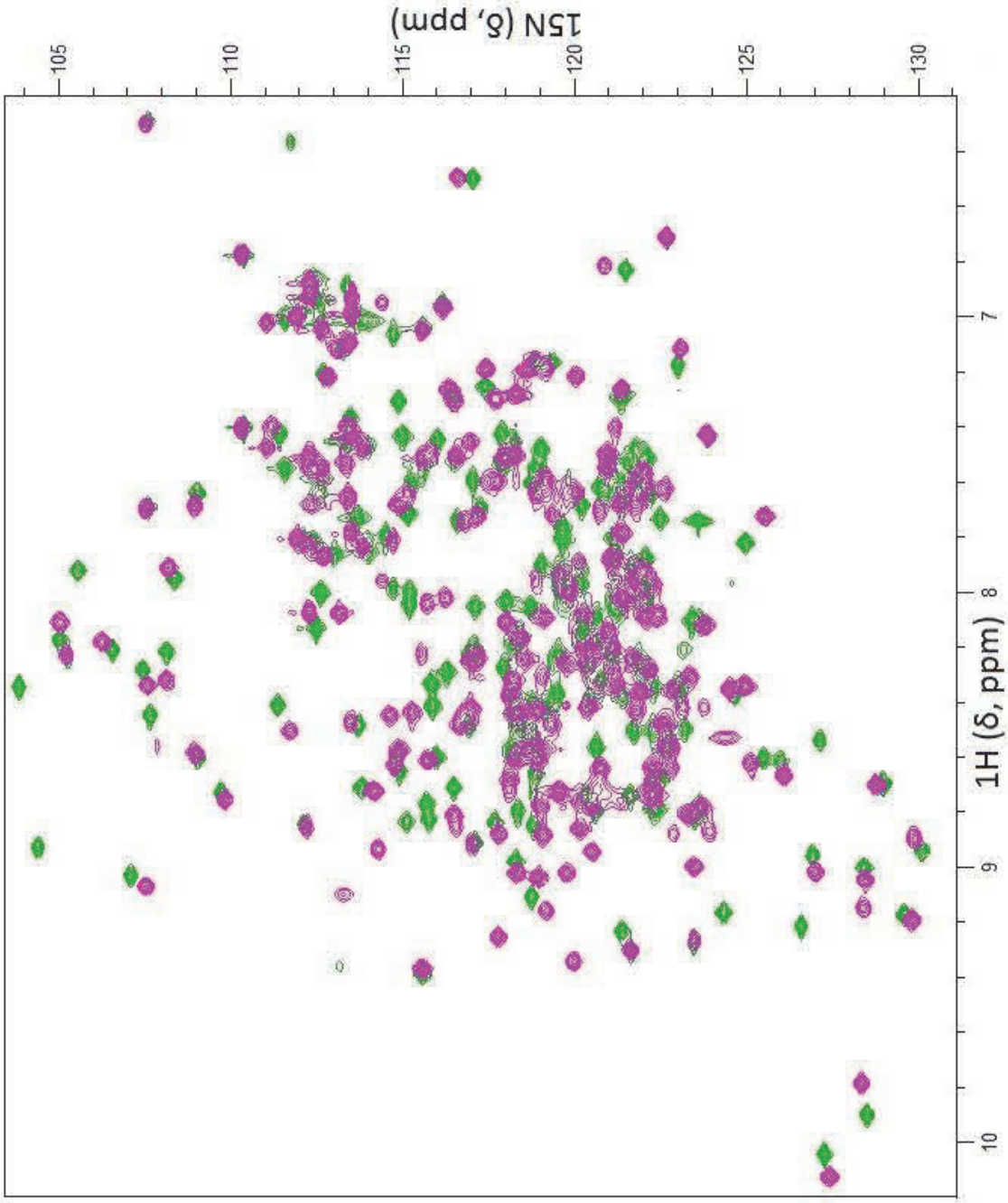


Figure 6.12 B

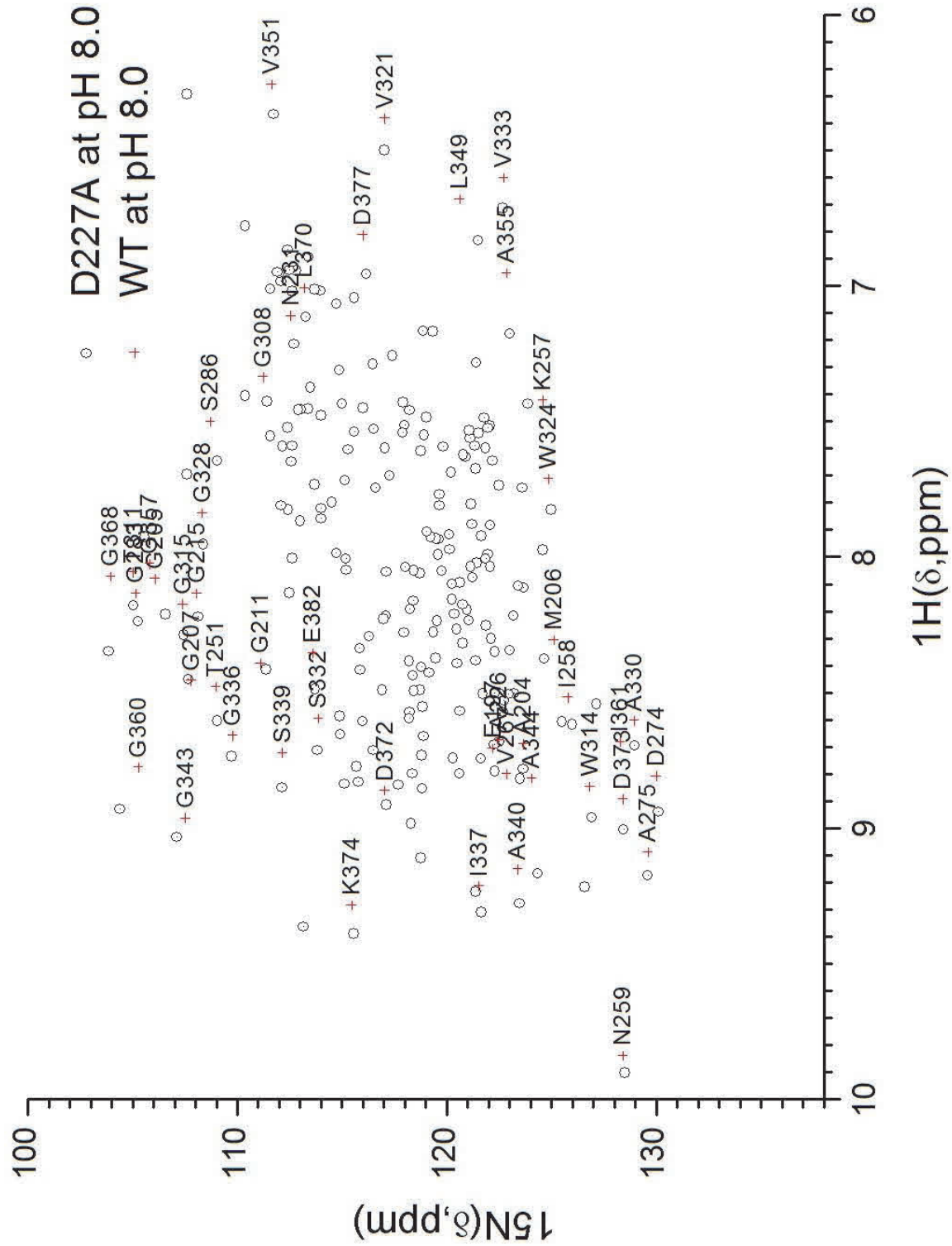

D227A at pH 8.0
WT at pH 8.0
15N(δ,ppm)
1H(δ,ppm)
100
110
120
130
6
7
8
9
10
G368
G360
G343
G207
G315
G215
G328
S286
T251
G336
G211
S339
S332
E382
G308
V351
N231
L370
D377
V321
L349
A355
V333
K257
W324
M206
I258
W314
D372
K374
I337
A340
A344
D373
I361
A330
A275
D274
N259

Figure 6.12 C

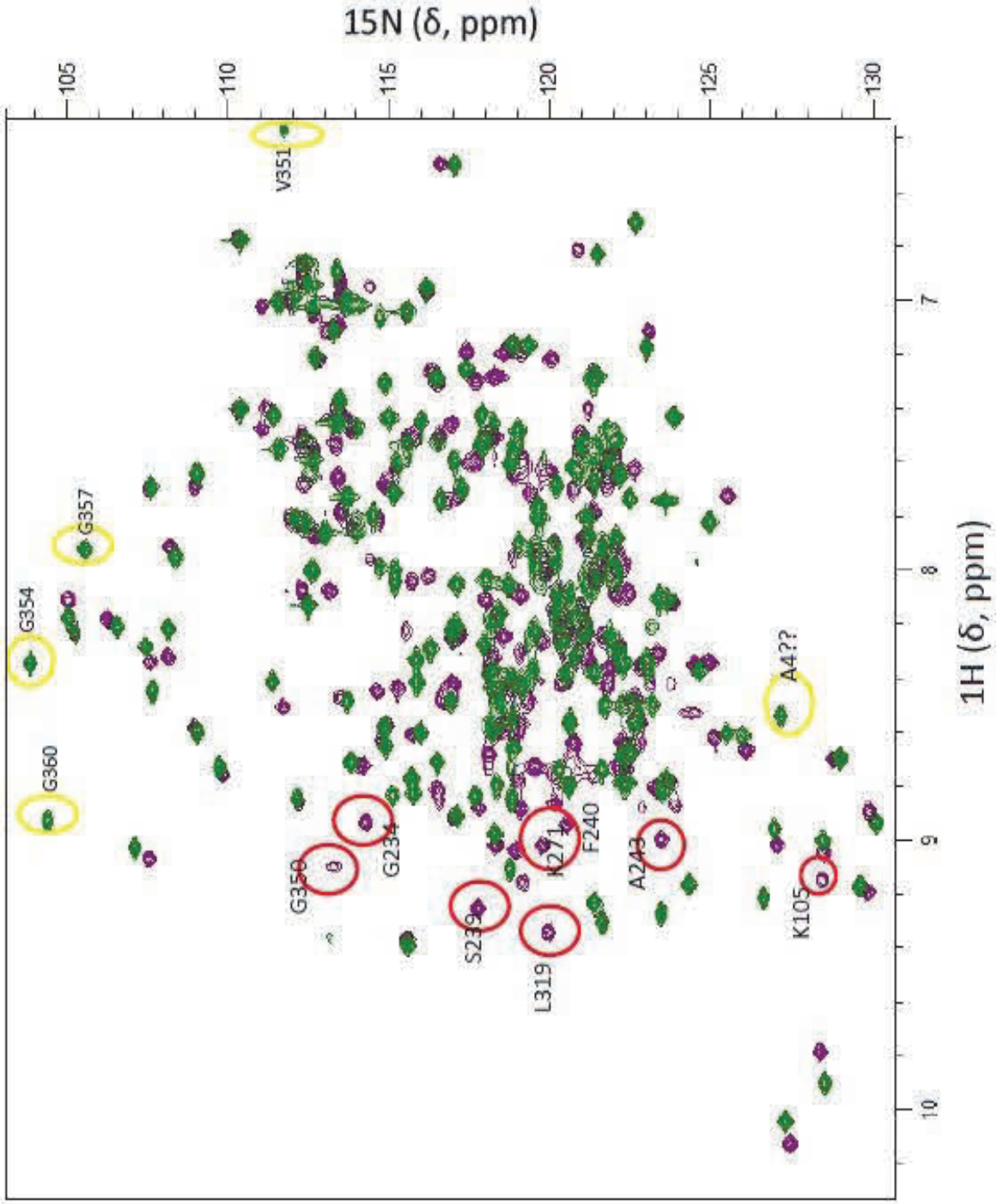

15N (δ, ppm)
105
110
115
120
125
130
1H (δ, ppm)
7
8
9
10
V351
G357
G354
G360
G350
G234
S239
L319
K271
F240
A243
A4??
K105

Figure 6.12 D

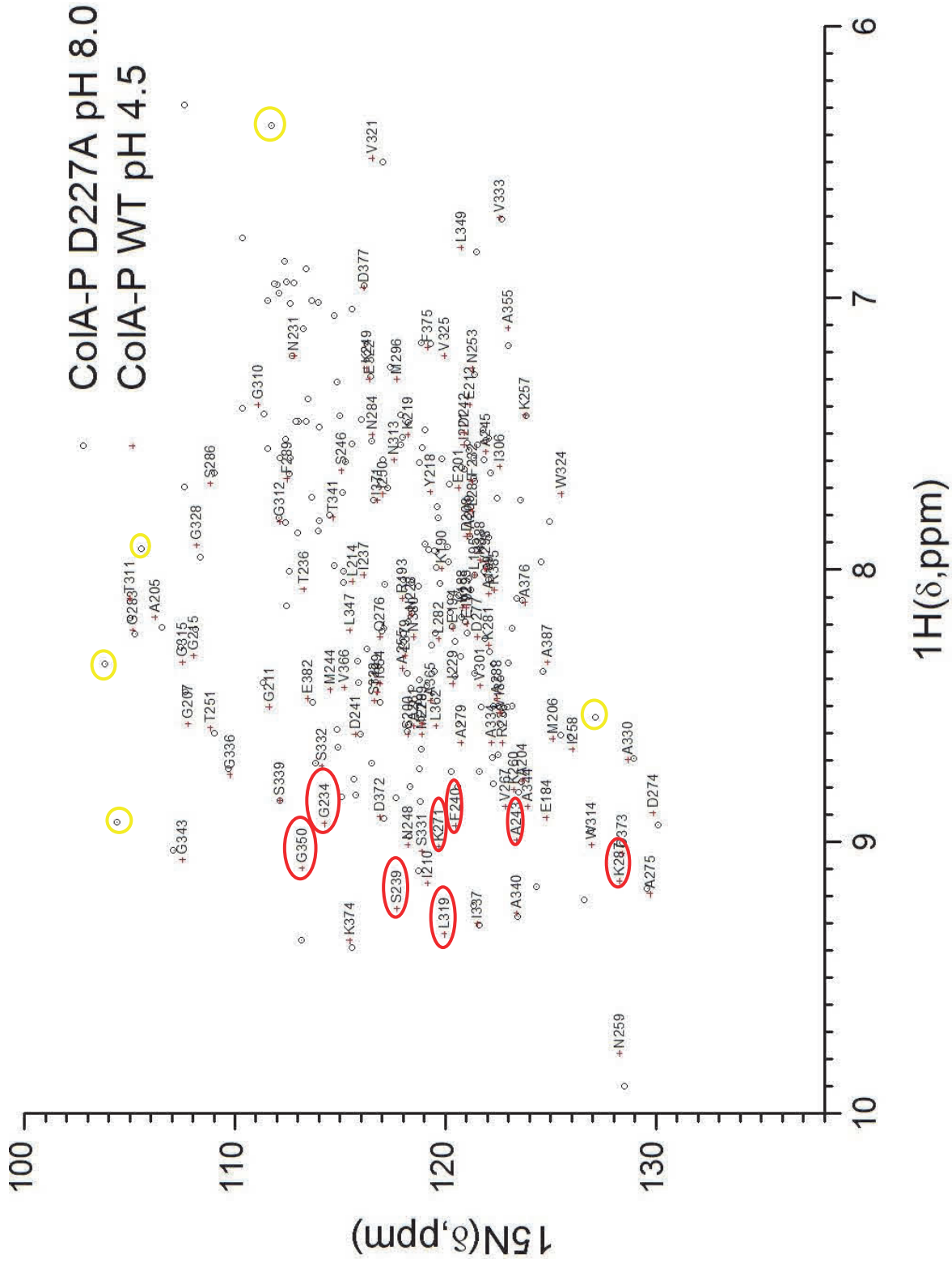

ColA-P D227A pH 8.0
ColA-P WT pH 4.5
15N(δ,ppm)
1H(δ,ppm)
100
110
120
130
6
7
8
9
10

Figure 6.14 A

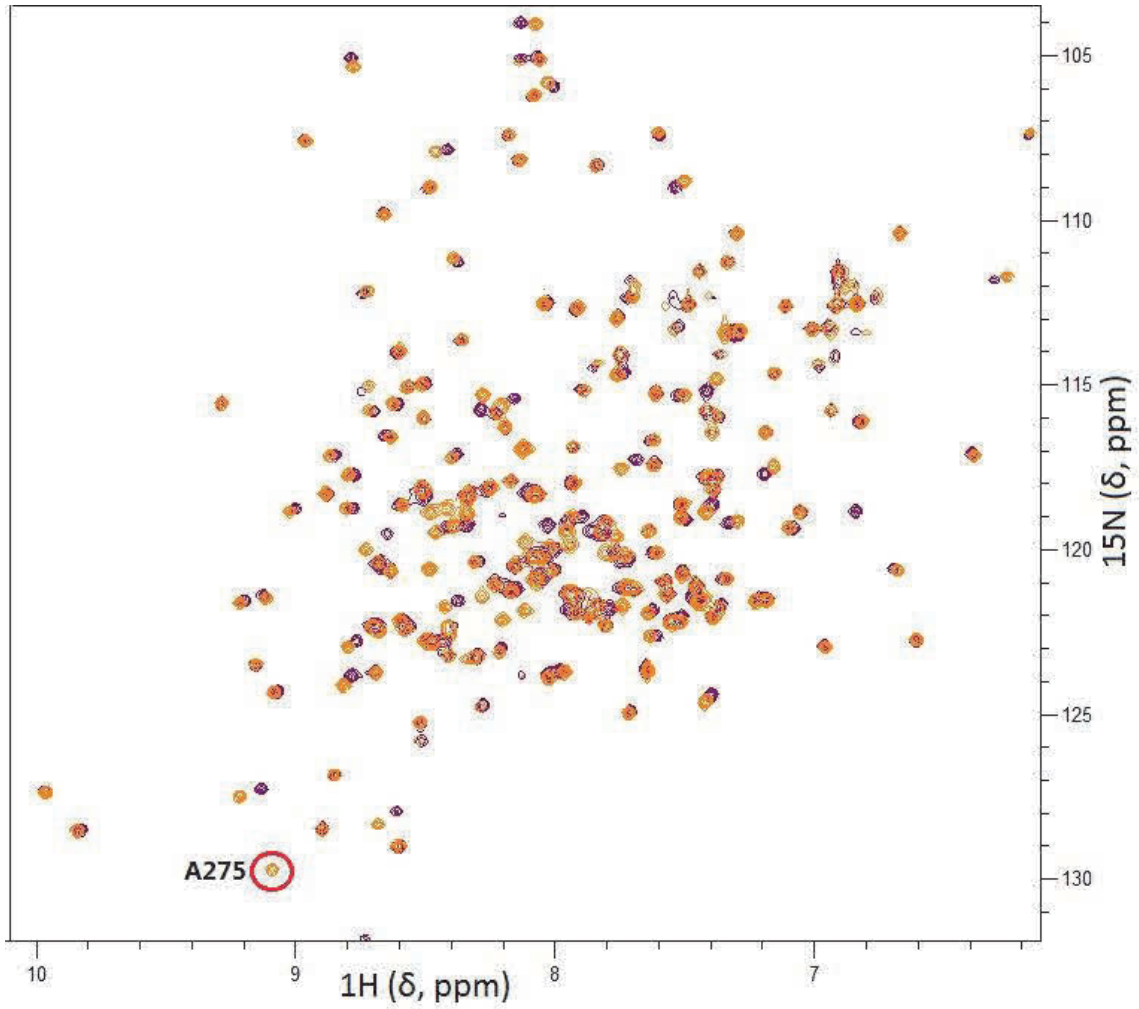


Figure 6.14 B

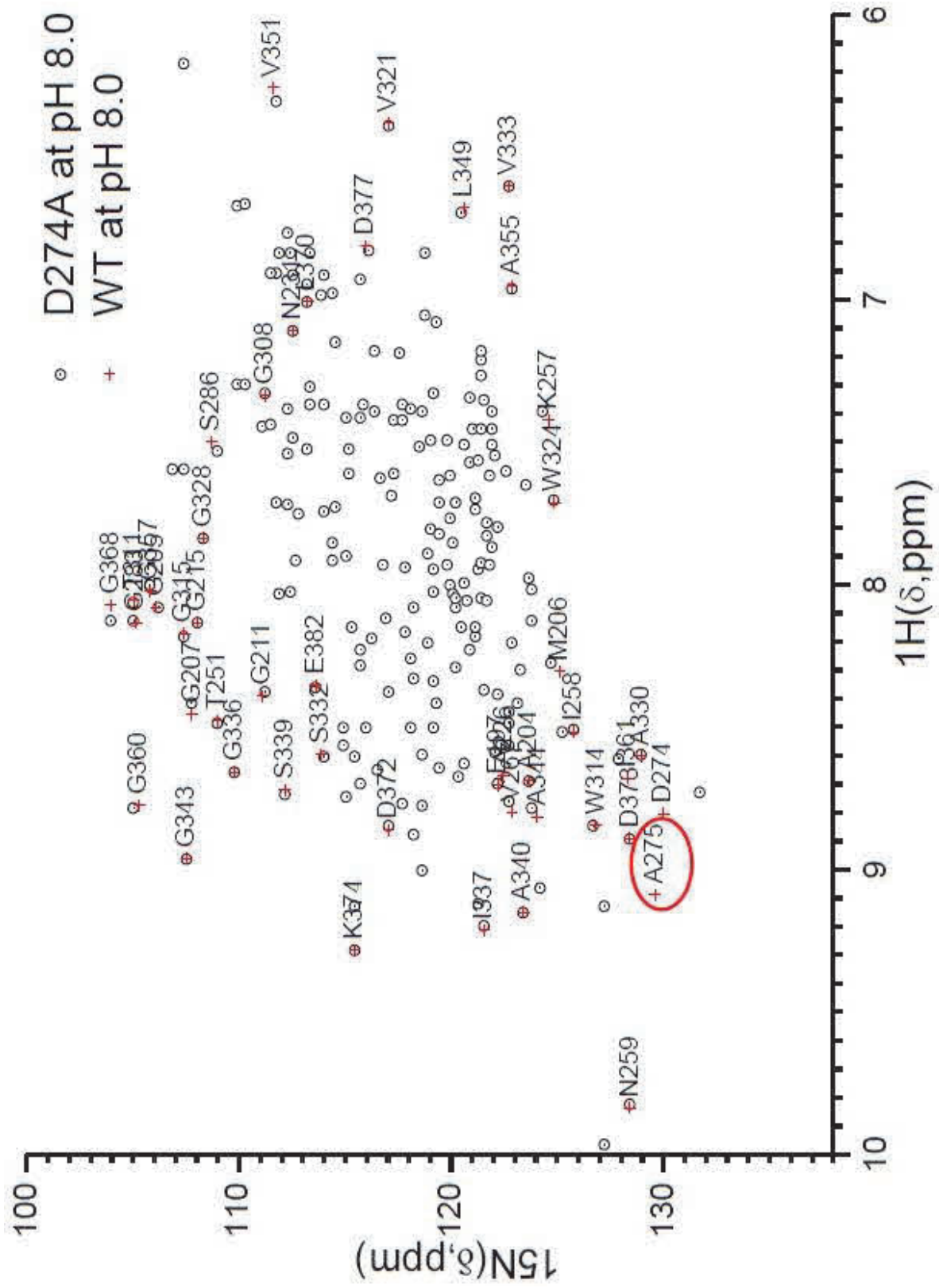

D274A at pH 8.0
WT at pH 8.0
15N(δ,ppm)
1H(δ,ppm)
100
110
120
130
6
7
8
9
10
G368
G360
G343
G207
T251
G336
G211
S339
S332
E382
D372
K374
I237
A340
A204
W314
N259
D378
A330
D274
A275
M206
I258
W324
K257
S286
G328
G308
V351
V321
D377
L349
A355
V333

Figure 6.14 C

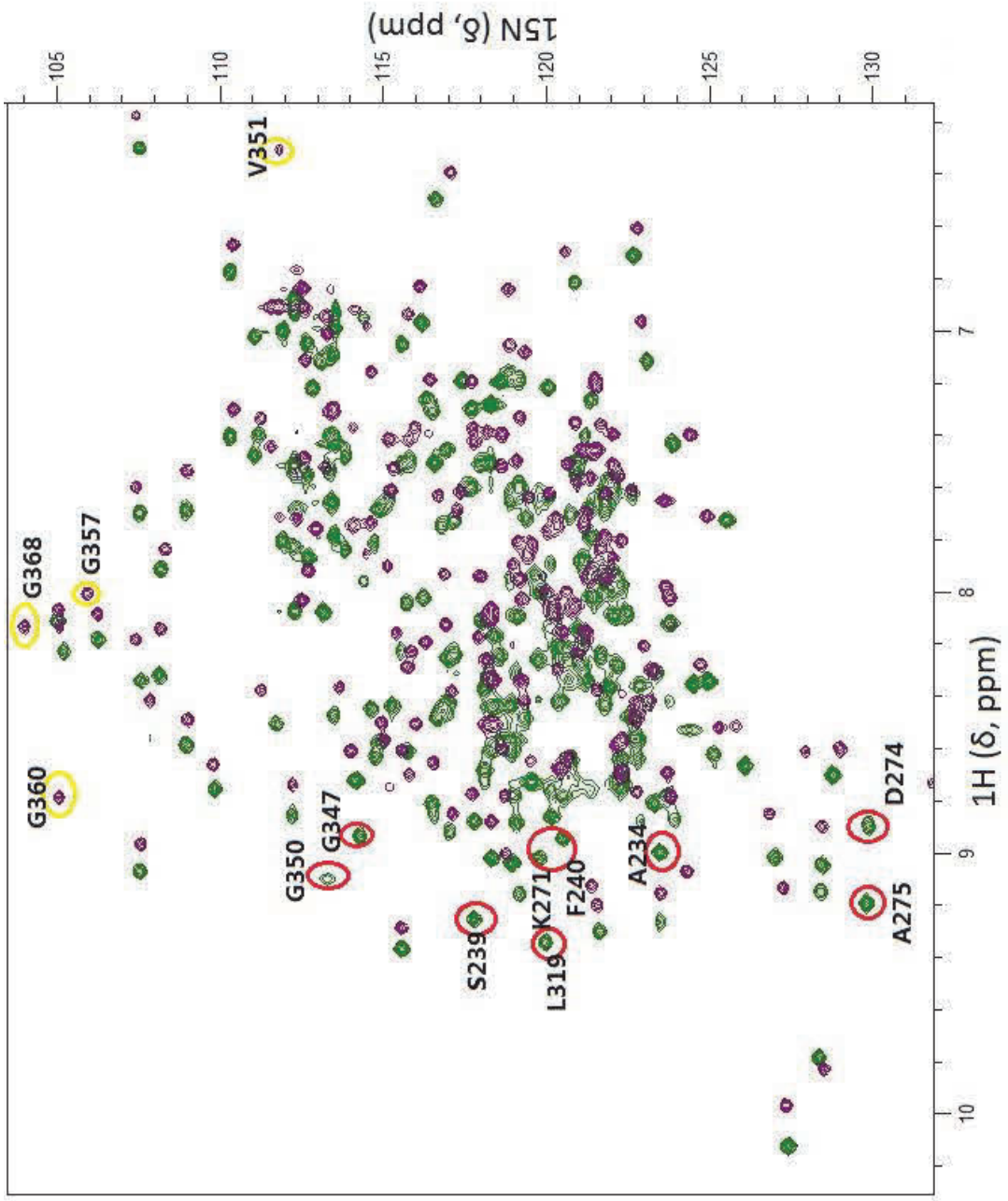

15N (δ, ppm)
105
110
115
120
125
130
1H (δ, ppm)
7
8
9
10
V351
G368
G357
G360
G350
G347
S239
L319
K271
F240
A234
A275
D274

Figure 6.14 D

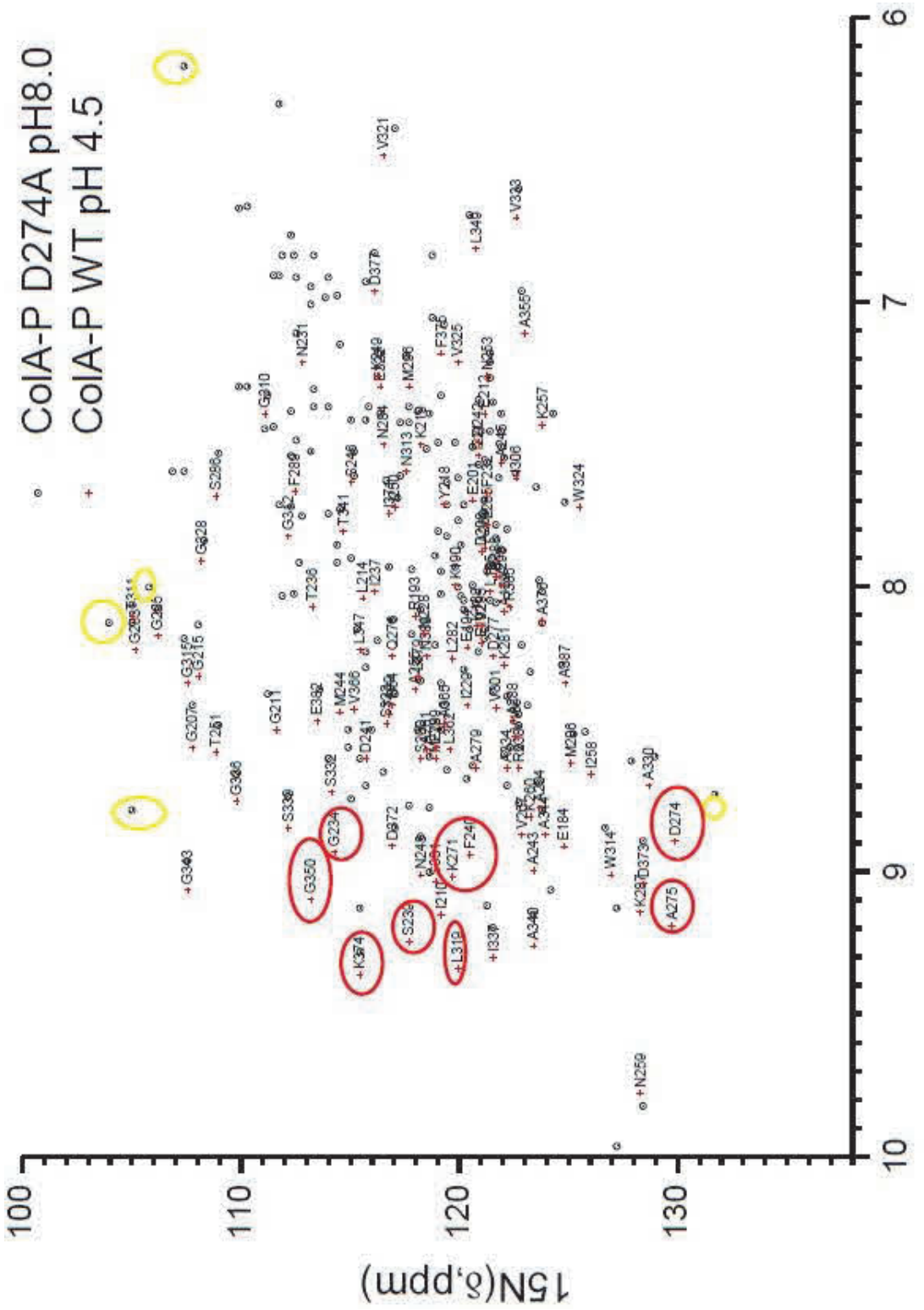

ColA-P D274A pH8.0
ColA-P WT pH 4.5
15N(δ,ppm)
1H(δ,ppm)
100
110
120
130
6
7
8
9
10
G350
G234
K271
F241
S236
L319
K374
D274
A275

Figure 6.16 A

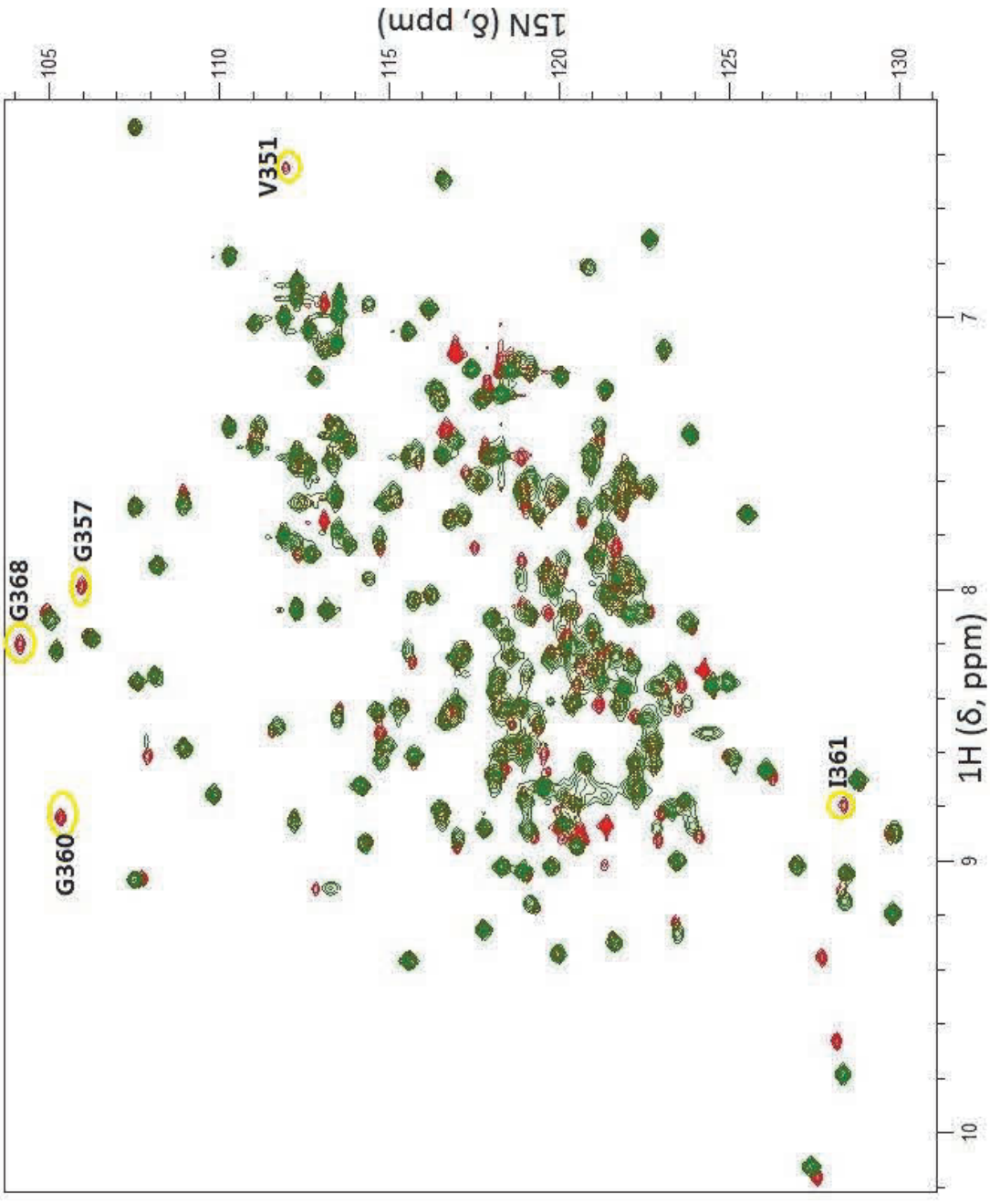

15N (δ, ppm)
105
110
115
120
125
130
1H (δ, ppm)
7
8
9
10
V351
G368
G357
G360
I361

Figure 6.16 B

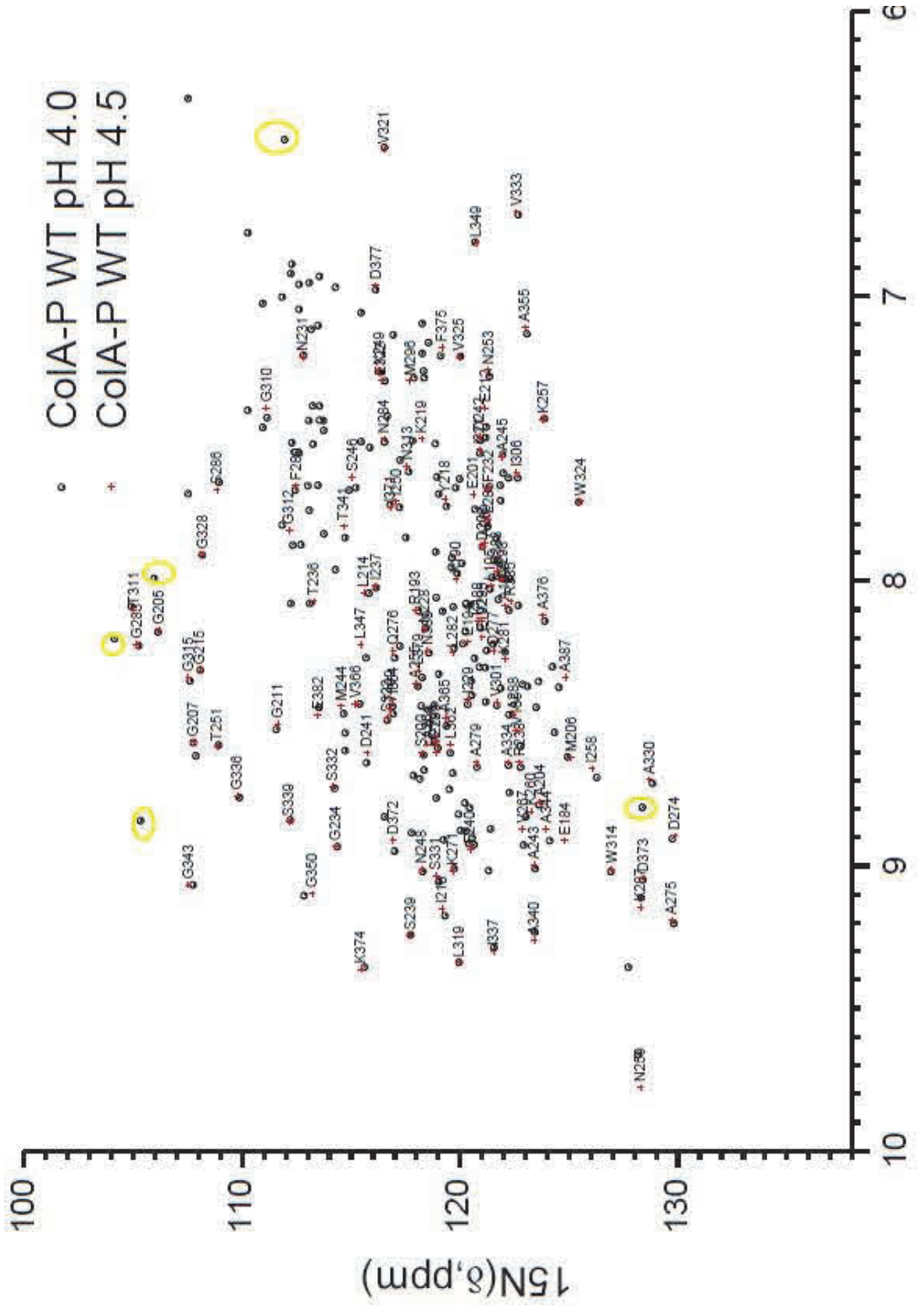

ColA-P WT pH 4.0
ColA-P WT pH 4.5
15N(δ,ppm)
1H(δ,ppm)
100
110
120
130
10
9
8
7
6

Figure 6.16 C

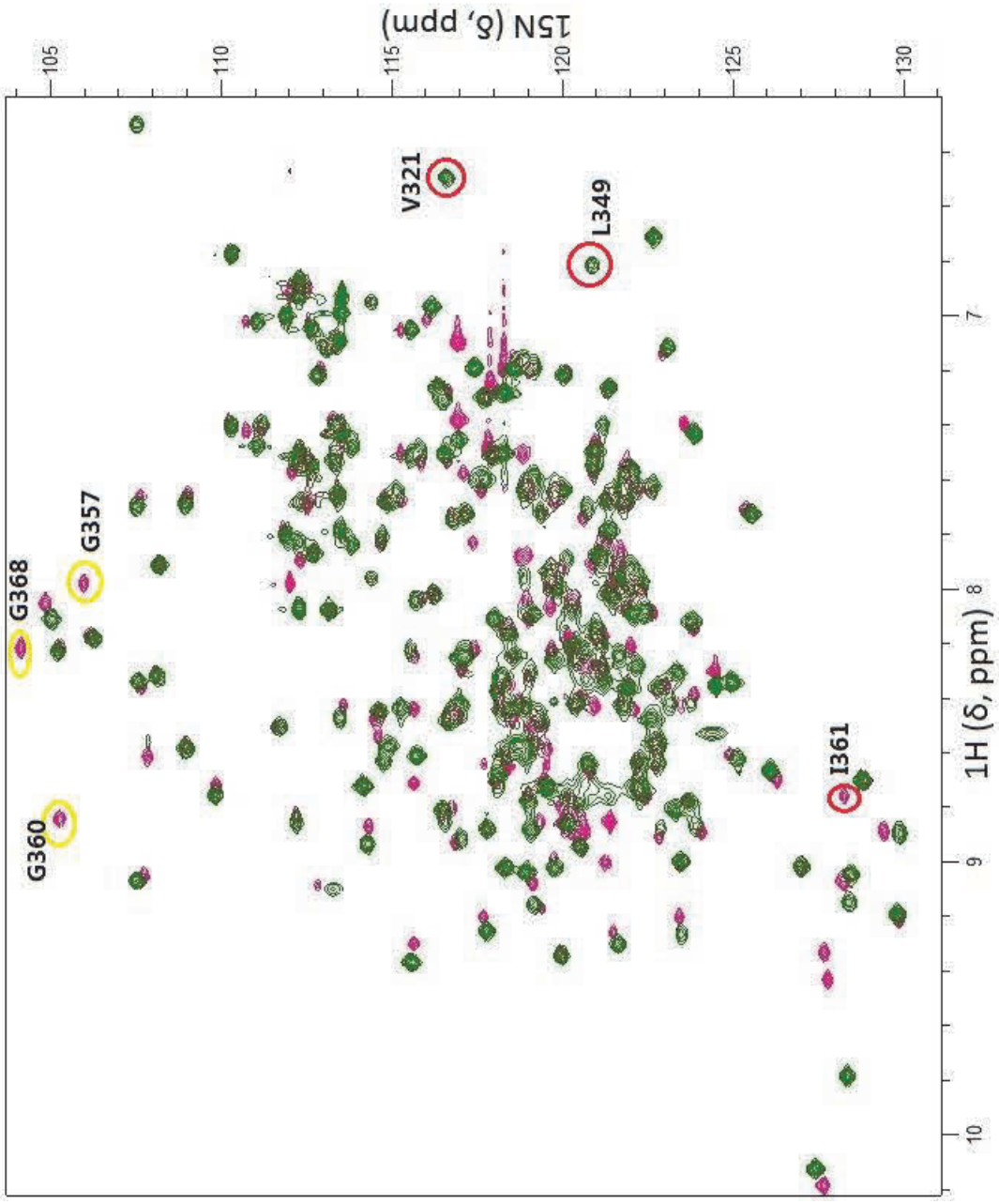

15N (δ, ppm)
105
110
115
120
125
130
1H (δ, ppm)
6
7
8
9
10
G368
G357
G360
V321
L349
I361

Figure 6.16 D

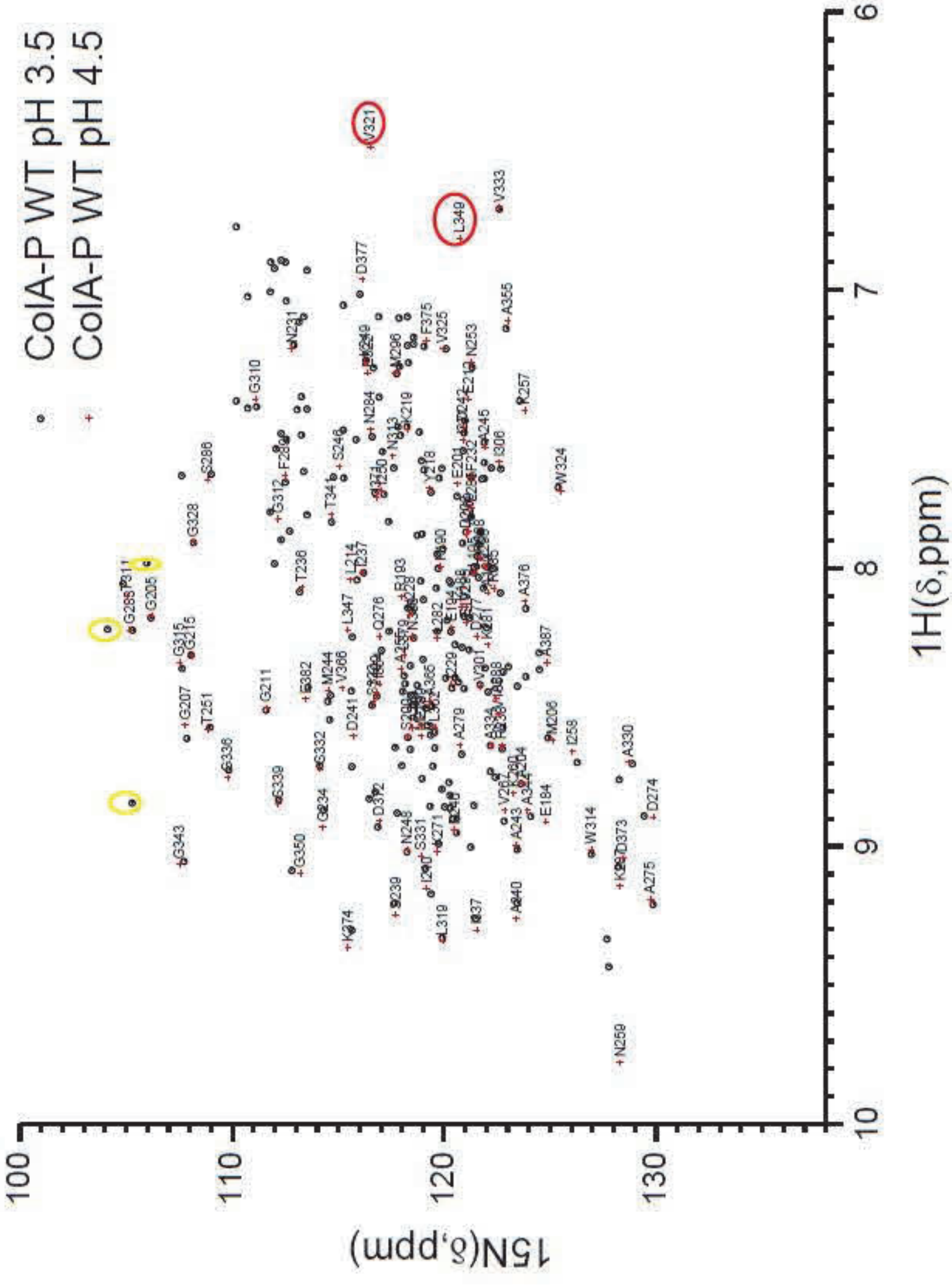

ColA-P WT pH 3.5
ColA-P WT pH 4.5
1H(δ,ppm)
15N(δ,ppm)
100
110
120
130
10
9
8
7
6

Figure 6.16 E

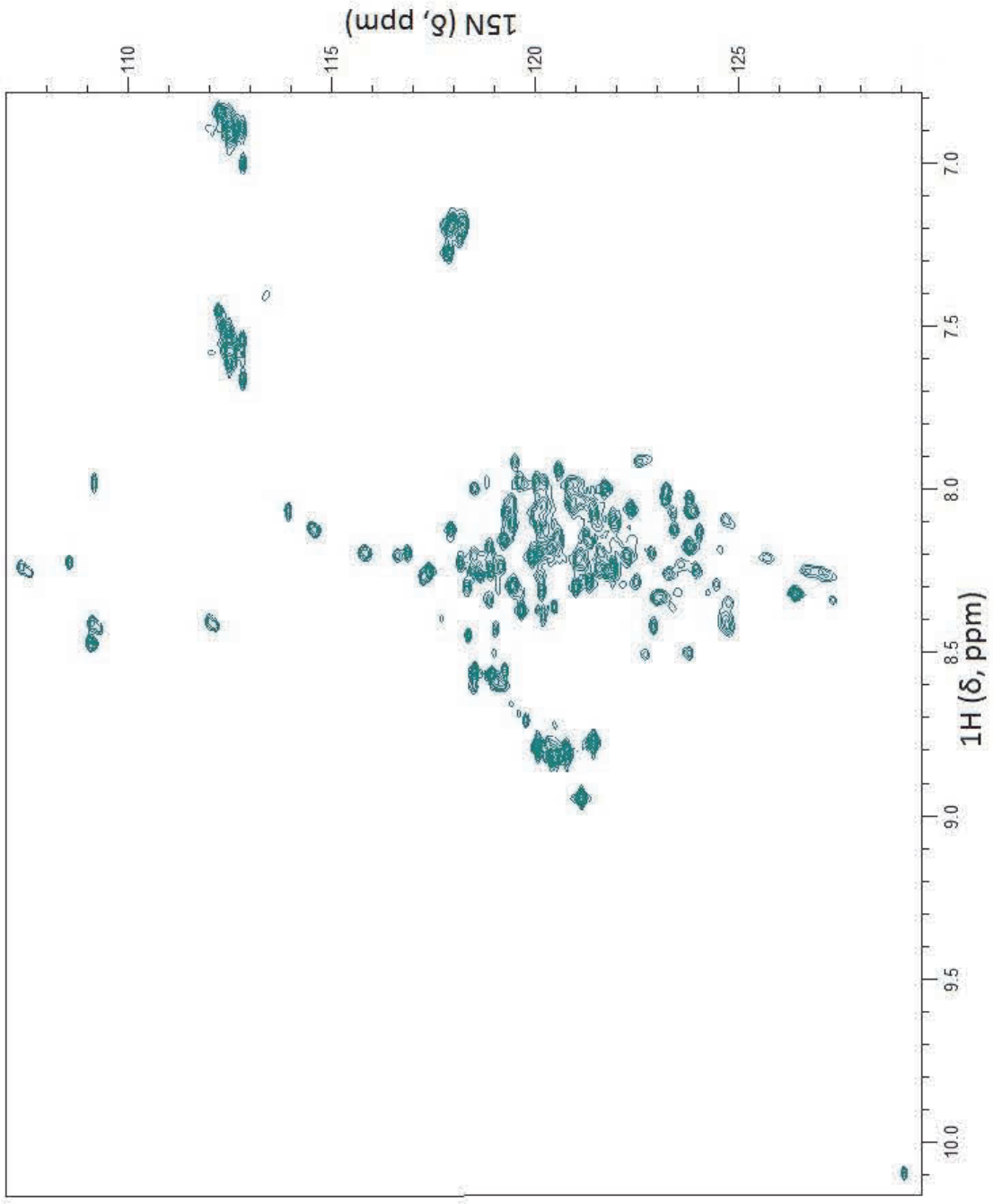
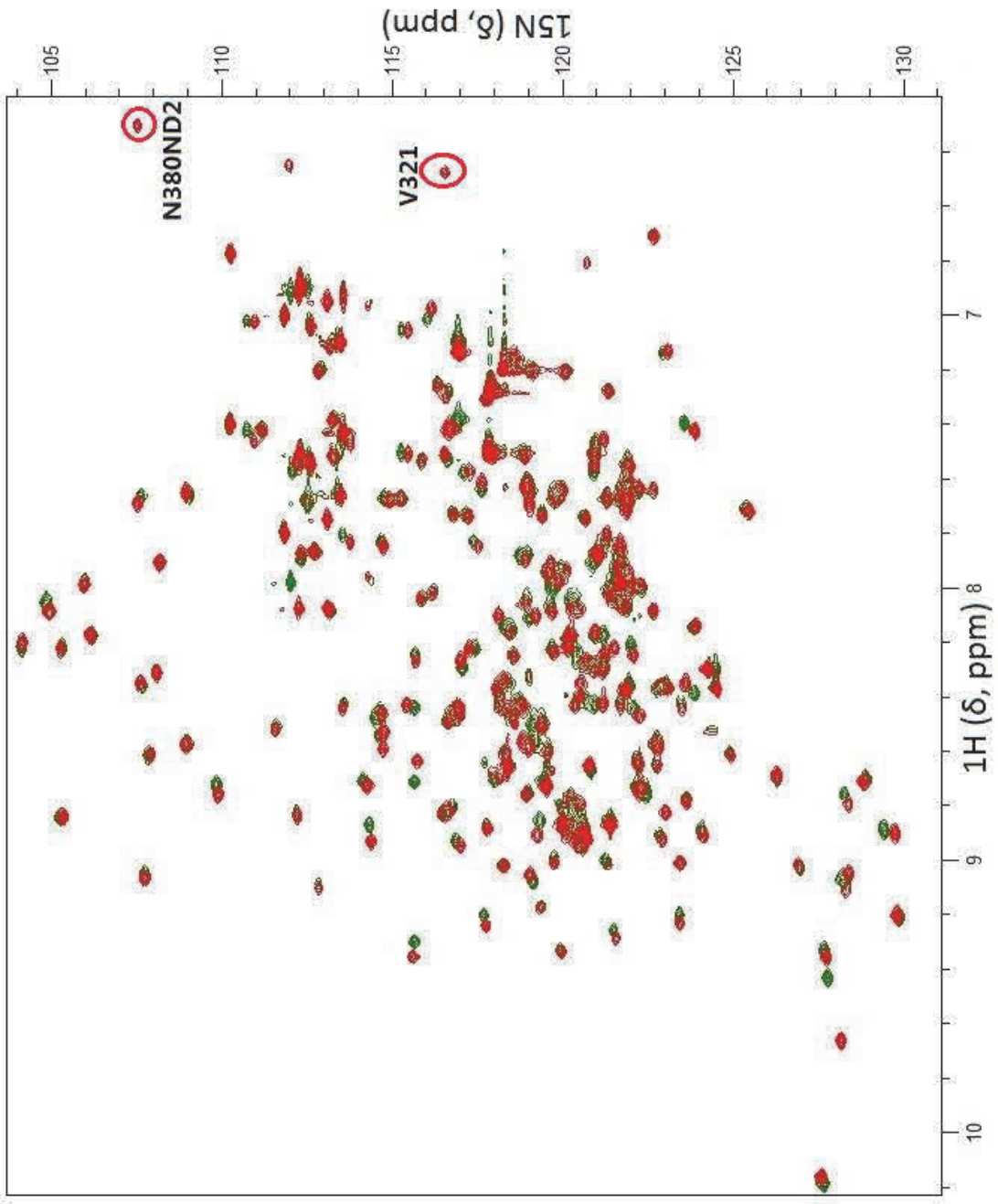


Figure 6.16 F



Chapter Seven

Conclusions and Future work

7.1. Conclusions:

Chapter Three confirmed previous studies that the negatively charged surface aspartate residues play an essential role in ColA-P protein folding and stability (Fridd and Lakey, 2002) and further showed that some aspartate to alanine mutants, such as ColA-P D216A, D227A and D372A destabilize the protein more than others, such as ColA-P D241A, D274A and D376A. The urea and guanidine HCl unfolding experiments followed by fluorescence further revealed that although the ColA-P domain is a small compact protein it unfolds in two steps in solution. The thermal denaturation experiments monitored by CD and DSC also show a two step transition process, which further confirms the existence of the intermediate state of the protein. The CD spectra indicated that at low pH the intermediate state of ColA-P protein lost part of its tertiary structure but retained its secondary structure, which confirmed that the intermediate state is equivalent to the molten globule state of the protein. This two stage unfolding leads to the creation of an intermediate state which has all the properties expected of the molten globule state which has been shown to be important *in vivo*. Low pH lowers the free energy barrier to the formation of the molten globule but increases the barrier to complete unfolding, in this way the intermediate is effectively stabilised. In showing that aspartate to asparagine mutants do not have a big effect on protein stability Chapter Four indicates that negative charge is not essential for protein stability at these sites. Therefore, unlike alanine, both aspartate and asparagine can play the role of a hydrogen bond acceptor and we also found that most of the negatively charged aspartate residues in ColA-P domain are located at the

N-terminal end of the helices. These observations can be explained by the N-capping hypothesis (Doig and Baldwin, 1995) which indicates that at the N terminus of an α-helix, a hydrogen bond acceptor, such as aspartate and asparagine can stabilize the helix structure. When aspartate is protonated at low pH it is less able to stabilise the helix and this provides an explanation for the effect of low pH on ColA-P stability. This suggested that most of the critical surface aspartate residues in ColA-P domain stabilize the protein as N-capping residues. The results in Chapter Five show that the helix structure are not destabilized by aspartate to glutamate and glutamine mutations as expected from the N-capping hypothesis (Doig and Baldwin, 1995), Thus, it seems that a simple N-capping hypothesis may not be applicable to full length proteins, such as ColA-P domain, and that individual stabilising side chains may be present in several N terminal sites (e.g. N1-3) of the helix. However, it is now clear that the acidic molten globule of colicin A is a result of the destabilisation of the N-termini of several alpha helices.

The NMR chapter confirms that the low pH effects influence the structure of the ColA-P domain but in addition it shows the subtle effects of the specific aspartate residues above pH 4.5. In particular when the WT ColA-P domain was measured at pH 8.0 we observed signals from residues in the hydrophobic core which were invisible at pH 4.5. When the mutants were investigated at pH 8.0 this transition did not occur, implying that they kept their acidic conformation even above pH 7.0. In addition, most of the mutations which affect signals from the hydrophobic core of the protein showing a long distance and potentially important role of these residues that is invisible to the other techniques used. In destabilising the hydrophobic core they help to convert the ColA-P domain from a water soluble to a membrane inserted protein.

7.2 Future work

The main unfinished work is in the NMR studies of the unfolding. The structure of ColA-P domain has changed a lot in acidic conditions or by aspartate to alanine mutations. In order to learn more about these conformation changes of ColA-P domain from ^{1}H-^{15}N-HSQC NMR spectra, it would ne necessary to have my own complete assignments based on ^{15}N-HSQC, ^{13}C-HSQC, and 3D HNCO, HN(CA)CO, HNCACB, HN(CO)CACB, HNCA, HN(CO)CA, HN(CA)HA, HN(COCA)HA and NNH-NOESY experiments. Especially, it would be important to have the complete assignments of ColA-P domain at pH 8.0 and pH 4.5, which can enable me to further localise the conformation change in the protein. With these assignments, we will be able to learn the relationship between the effect of acidic environment and the point mutations on the surface of the protein. One important question is whether the partially unfolded secondary structure of the intermediate maps onto a specific region of the protein or whether this reflects a dynamic equilibrium where 50% of the residues are unfolded at any one time.

One further area which has not been studied with these mutants is their effect of the toxin *in vivo* on living bacteria. Since we now have the ability to reduce stepwise the influence of pH on the ColA-P domain by insertion of asparagine mutants we can seek to understand why colicin A needs this behaviour. If we prevent its acidic unfolding will it stop working or become independent of acidic condition and the need for acidic lipids and thus resemble colicin N. Similarly we would now be able to predict the correct asparagine insertions to make colicin B insensitive to pH and thus compare this effect in a colicin which relies upon the Ton system for translocation rather than the Tol system used by Colicins A and N. These experiments will provide useful information about these diverse translocation systems and provide new information on how to devise new antimicrobial peptides.

References

Information on materials and methods is available on Science Online.

AIDLEY, D. J. & STANFIELD, P. R. 1996. *Ion channels - molecules in action.,* Cambridge, Cambridge University Press.

ALOUF, J. E. 2001. Pore-forming bacterial protein toxins: An overview. *Curr.Top. Microbiol. Immunol.,* 257**,** 1-14.

BALDWIN, M. R., LAKEY, J. H. & LAX, A. J. 2004. Identification and characterisation of the *Pasteurella multocida* toxin translocation domain. *Mol. Microbiol.,* 54**,** 239-250.

BATY, D., LAKEY, J., PATTUS, F. & LAZDUNSKI, C. 1990. A 136-amino-acid-residue COOH-terminal fragment of colicin A is endowed with ionophoric activity. *Eur. J. Biochem.,* 189**,** 409-13.

BENEDETTI, H., FRENETTE, M., BATY, D., KNIBIEHLER, M., PATTUS, F. & LAZDUNSKI, C. 1991. Indicidual domains of colicins confer specificity in colicin uptake, in pore properties and in immunity requirement. *J. Mol Biol.,* 217**,** 429-439.

BOURDINEAUD, J. P., BOULANGER, P., LAZDUNSKI, C. & LETELLIER, L. 1990. In vivo properties of colicin-A - channel activity is voltage dependent but translocation may be voltage independent. *Proc. Natl. Acad. Sci. USA.,* 87**,** 1037-1041.

BOUVERET, E., RIGAL, A., LAZDUNSKI, C. & BENEDETTI, H. 1997. The N-terminal domain of colicin E3 interacts with TolB which is involved in the colicin translocation step. *Mol. Microbiol.,* 23**,** 909-920.

BOUVERET, E., RIGAL, A., LAZDUNSKI, C. & BENEDETTI, H. 1998. Distinct regions of the colicin A translocation domain are involved in the interaction with TolA and TolB proteins upon import into *Escherichia coli*. *Mol. Microbiol.,* 27**,** 143-157.

BRAUN, V. & MAAS, E. 1984. Colicin-B Consists of a Single Polypeptide-Chain. *Fems Microbiol. Lett.,* 21**,** 93-97.

BRAUN, V., PATZER, S. I. & HANTKE, K. 2002. Ton-dependent colicins and microcins: modular design and evolution,. *Biochimie,* 84**,** 365-380.

BRAUN, V., PILSL, H. & GROSS, P. 1994. Colicins-structures, modes of action, transfer through membranes, and evolution. *Arch. Microbiol.,* 161**,** 199-206.

CAPALDI, A. P., KLEANTHOUS, C. & RADFORD, S. E. 2002. Im7 folding mechanism: misfolding on a path to the native state. *Nat. Struct. Biol.,* 9**,** 209-216.

CARR, S., MILLER, J., LEARY, S. E. C., BENNETT, A. M., HO, A. & WILLIAMSON, E. D. 1999. Expression of a recombinant form of the V antigen of Yersinia pestis, using three different expression systems. *Vaccine.,* 18**,** 153-159.

CASCALES, E., BUCHANAN, S. K., DUCHE, D., KLEANTHOUS, C., LLOUBES, R., POSTLE, K., RILEY, M., SLATIN, S. & CAVARD, D. 2007. Colicin biology. *Microbiol. Mol. Biol. R.,* 71**,** 158-229.

CAVARD, D. 1997. Role of the colicin A lysis protein in the expression of the colicin A operon. *Microbiol. UK.,* 143**,** 2295-2303.

CAVARD, D. & LAZDUNSKI, C. 1981. Involvement of BtuB and OmpF proteins in the binding and uptake of colicin A. *Fems Microbiol. Lett.,* 12**,** 311-316.

CHAKRABARTTY, A., SCHELLMAN, J. A. & BALDWIN, R. L. 1991. Large differences in the helix propensities of alanine and glycine. *Nature.,* 351**,** 586-8.

CHALTON, D. A. & LAKEY, J. H. 2010. Simple Detection of Protein Soft Structure Changes. *Anal. Chem.,* 82**,** 3073-3076.

CHOU, P. Y. & FASMAN, G. D. 1974. Conformation parameters for amino-asids in helical, beta-sheet, and random ciol regions calculated from proteins. *Biochemistry-US.,* 13**,** 211-222.

CLIFTON, L. A., JOHNSON, C. L., SOLOVYOVA, A. S., CALLOW, P., WEISS, K. L., RIDLEY, H., LE BRUN, A. P., KINANE, C. J., WEBSTER, J. R., HOLT, S. A. & LAKEY, J. H. 2012a. Low resolution structure and dynamics of a Colicin-Receptor complex determined by neutron scattering. *J. Biol. Chem.,* 287**,** 337-346.

CLIFTON, L. A., JOHNSON, C. L., SOLOVYOVA, A. S., CALLOW, P., WEISS, K. L., RIDLEY, H., LE BRUN, A. P., KINANE, C. J., WEBSTER, J. R. P., HOLT, S. A. & LAKEY, J. H. 2012b. Low Resolution Structure and Dynamics of a Colicin-Receptor Complex Determined by Neutron Scattering. *J. Biol. Chem.,* 287**,** 337-346.

COLLARINI, M., AMBLARD, G., LAZDUNSKI, C. & PATTUS, F. 1987. Gating processes of channels induced by colicin A, its C-terminal fragment and colicin E1 in planar lipid bilayers. *Eur. Biophys. J.,* 14**,** 147-53.

COOPER, A. (2004). *Biophysical. Chemistry.* Cambrisge The Royal Society of Chemistry.

CRAMER, W. A., DANKERT, J. R. & URATANI, Y. 1983. The membrane channel-forming bacteriocidal protein, colicin El. *Biochim. Biophys. Acta.,* 737**,** 173-93.

CREIGHTON, T. E. (ed.) 1997. *Protein structure*: Oxford University Press.

DAVIES, J. K. & REEVES, P. 1975a. Genetics of resistance to colicins in Escherichia coli K-12: cross-resistance among colicins of group A. *J. Bacteriol.,* 123**,** 102-17.

DAVIES, J. K. & REEVES, P. 1975b. Genetics of resistance to colicins in Escherichia coli K-12: cross-resistance among colicins of group B. *J. Bacteriol.,* 123**,** 96-101.

DECATUR, J. 2011. *'HSQC and HMBC for Topspin'*.

DOIG, A. J. & BALDWIN, R. L. 1995. N-and C-capping preferences for all 20 amino acids in a-helical peptides. *Protein. Sci.,* 4**,** 1325-1336.

DUCHE, D. 2002. The pore-forming domain of colicin A fused to a signal peptide: a tool for studying pore-formation and inhibition. *Biochimie.,* 84**,** 455-464.

DYSON, H. J. & WRIGHT, P. E. 2004. Unfolded proteins and protein folding studied by NMR. *Chem. Rev.,* 104**,** 3607-3622.

EL GHACHI, M., BOUHSS, A., BARRETEAU, H., TOUZE, T., AUGER, G., BLANOT, D. & MENGIN-LECREULX, D. 2006. Colicin M exerts its bacteriolytic effect via enzymatic degradation of undecaprenyl phosphate-linked peptidoglycan precursors. *J. Biol. Chem.,* 281**,** 22761-22772.

ELKINS, P. A., SONG, H. Y., CRAMER, W. A. & STAUFFACHER, C. V. 1994. Crystallization and characterization of colicin e1 channel-forming polypeptides. *Proteins.,* 19**,** 150-157.

EVANS, L. J. A., GOBLE, M. L., HALES, K. & LAKEY, J. H. 1996. Different sensitivities to acid denaturation within a family of proteins; Implications for acid unfolding and membrane translocation. *Biochemistry,* 35**,** 13180-13185.

FINK, A. L., CALCIANO, L. J., GOTO, Y., KUROTSU, T. & PALLEROS, D. R. 1994. Classification of acid denaturation of proteins - intermediates and unfolded states. *Biochemistry-US.,* 33**,** 12504-12511.

FRIDD, S. L. & LAKEY, J. H. 2002. Surface aspartate residues are essential for the stability of colicin A P-domain: A mechanism for the formation of an acidic molten-globule. *Biochemistry-US.,* 41**,** 1579-1586.

GELI, V. & LAZDUNSKI, C. 1992. An α-helical hydrophobic hairpin as a specific determinant in Protein-Protein interaction occurring in *Escherichia coli* colicin A and B immunity systems. *J. Bacteriol.,* 174**,** 6432-6437.

GHOSH, P., MEL, S. F. & STROUD, R. M. 1994. The domain-structure of the ion channel-forming protein colicin Ia. *Nat. Struct. Biol.,* 1**,** 597-604.

GOKCE, I. & LAKEY, J. H. 2003. Production of an E-coli toxin protein; Colicin A in E-coli using an inducible system. *Turk. J. Chem.,* 27**,** 323-331.

GONZALEZ-MANAS, J. M., LAKEY, J. H. & PATTUS, F. 1992. Brominated phospholipids as a tool for monitoring the membrane insertion of colicin A. *Biochemistry-US.,* 31**,** 7294-300.

GORDON, D. M. & O'BRIEN, C. L. 2006. Bacteriocin diversity and the frequency of multiple bacterlocin production in Escherichia coli. *Microbiology+.,* 152**,** 3239-3244.

GRATIA, A. 1925. Sur un remarquable example d'antagonisme entre deux souches de colibacille. *CR. Soc.Biol.,* 93**,** 1040-1041.

GREENE, L. H., WIJESINHA-BETTONI, R. & REDFIELD, C. 2006. Characterization of the molten globule of human serum retinol-binding protein using NMR Spectroscopy. *Biochemistry-US.,* 45**,** 9475-9484.

GSPONER, J., HOPEARUOHO, H., WHITTAKER, S. B. M., SPENCE, G. R., MOORE, G. R., PACI, E., RADFORD, S. E. & VENDRUSCOLO, M. 2006. Determination of an ensemble of structures representing the intermediate state of the bacterial immunity protein Im7. *Proc. Natl. Acad. Sci. USA.,* 103**,** 99-104.

HAKKAART, M. J., VELTKAMP, E. & NIJKAMP, H. 1981. Protein H encoded by plasmid Clo DF13 involved in lysis of the bacterial host. I. Localisation of the gene and identification and subcellular localisation of the gene H product. *Mol. Gen.,* 183(2).

HARDY, K. G. 1975. Colicinogeny and related phenomena. *Bacteriol. Rev.,* 39**,** 464-515.

HARDY, K. G., MEYNELL, G. G., DOWMAN, J. E. & SPRATT, B. G. 1973. Two major groups of colicin factors: their evolutionary significance. *Mol. Gen. Genet.,* 125**,** 217-30.

HARKNESS, R. E. & BRAUN, V. 1989. Colicin M inhibits peptidoglycan biosynthesis by interfering with lipid carrier recycling. *J. Biol. Chem.,* 264**,** 6177-82.

HECHT, O., RIDLEY, H., LAKEY, J. H. & MOORE, G. R. 2009. A Common Interaction for the Entry of Colicin N and Filamentous Phage into Escherichia coli. *J. Mol. Biol.,* 388**,** 880-893.

HILSENBECK, J. L., PARK, H., CHEN, G., YOUN, B., POSTLE, K. & KANG, C. 2004. Crystal structure of the cytotoxic bacterial protein colicin B at 2.5 Å resolution. *Mol. Microbiol.,* 51**,** 711-720.

HOUSDEN, N. G., LOFTUS, S. R., MOORE, G. R., JAMES, R. & KLEANTHOUS, C. 2005. Cell entry mechanism of enzymatic bacterial colicins: Porin recruitment and the thermodynamics of receptor binding. *Proc. Natl. Acad. Sci. USA.,* 102**,** 13849-13854.

IBANEZ DE OPAKUA, A., DIERCKS, T., VIGUERA, A. R. & BLANCO, F. J. 2010. NMR assignment and backbone dynamics of the pore-forming domain of colicin A. *Biomol. NMR. Assign.,* 4**,** 33-36.

SERDYUK, I.N., N.R. & ZACCAI, J. 2007. *Methods in Molecular Biophysics*, Cambridge University Press.

JACOB, F., SIMINOVITCH, L. & WOLLMAN, E. 1952. Biosynthesis of a colicin and its mode of action. *Ann. Inst. Pasteur.,* 83**,** 295-315.

JAKES, K., ZINDER, N. D. & BOON, T. 1974. Purification and properties of colicin E3 immunity protein. *J. Biol. Chem.,* 249**,** 438-44.

KELLY, S. M., JESS, T. J. & PRICE, N. C. 2005. How to study proteins by circular dichroism. *Biochim. Biophys. Acta.,* 1751**,** 119-139.

KIENKER, P., SLATIN, S. L. & FINKELSTEIN, A. 2002. Colicin channels have a shockingly high proton permeability. *Biophys. J.,* 82**,** 555A-555A.

KIENKER, P. J., QIU, X. Q., SLATIN, S. L., FINKELSTEIN, A. & JAKES, K. S. 1997. Transmembrane insertion of the colicin Ia hydrophobic hairpin. *J. Membrane. Biol.,* 157**,** 27-37.

KLEANTHOUS, C. 2010. Swimming against the tide: progress and challenges in our understanding of colicin translocation. *Nat. Rev. Microbiol.,* 8**,** 843-848.

KONISKY, J. & NOMURA, M. 1967. Interaction of colicins with bacterial cells. II. Specific alteration of Escherichia coli ribosomes induced by colicin E3 in vivo. *J. Mol. Biol.,* 26**,** 181-95.

LAKEY, J. H., BATY, D. & PATTUS, F. 1991a. Fluorescence energy transfer distance measurements using site-directed single cysteine mutants. The membrane insertion of colicin A. *J. Mol. Biol.,* 218**,** 639-53.

LAKEY, J. H., DUCH , D., GONZ LEZ-MA AS, J.-M., BATY, D. & PATTUS, F. 1993. Fluorescence energy transfer distance measurements: the hydrophobic helical hairpin of Colicin A in the membrane bound state. *J. Mol. Biol.,* 230**,** 1055-1067.

LAKEY, J. H., MASSOTTE, D., HEITZ, F., DASSEUX, J. L., FAUCON, J. F., PARKER, M. W. & PATTUS, F. 1991b. Membrane insertion of the pore-forming domain of colicin A. A spectroscopic study. *Eur. J. Biochem.,* 196**,** 599-607.

LAKEY, J. H., PARKER, M. W., GONZALEZMANAS, J. M., DUCHE, D., VRIEND, G., BATY, D. & PATTUS, F. 1994. Electrostatics and the Unfolding of colicins at the membrane surface. *In:* FREER, J., AITKEN, R., ALOUF, J. E., BOULNOIS, G., FALMAGNE, P.,

FEHRENBACH, F., MONTECUCCO, C., PIEMONT, Y., RAPPUOLI, R., WADSTROM, T. & WITHOLT, B. (eds.) *Bacterial Protein Toxins.* Heidelberg: Gustav Fischer.

LAKEY, J. H. & SLATIN, S. L. 2001. Pore-forming colicins and their relatives. *In:* VAN DER GOOT, F. G. (ed.) *Pore-Forming Toxins.* Heidelberg: Springer Verlag.

LAKOWICZ, J. R. 2007. *Principles of Fluorescence Spectroscopy*, Springer.

LAZDUNSKI, C. J., BOUVERET, E., RIGAL, A., JOURNET, L., LLOUBES, R. & BENEDETTI, H. 1998. Colicin import into Escherichia coli cells. *J. Bacteriol.,* 180**,** 4993-5002.

LAZZARONI, J. C., DUBUISSON, J. F. & VIANNEY, A. 2002. The Tol proteins of Escherichia coli and their involvement in the translocation of group A colicins. *Biochimie.,* 84**,** 391-7.

LEMASTER, D. M. & RICHARDS, F. M. 1982. PREPARATIVE-SCALE ISOLATION OF ISOTOPICALLY LABELED AMINO-ACIDS. *Anal. Biochem.,* 122**,** 238-247.

LINDEBERG, M., ZAKHAROV, S. D. & CRAMER, W. A. 2000. Unfolding pathway of the colicin E1 channel protein on a membrane surface. *J. Mol. Biol.,* 295**,** 679-692.

LLOUB S R, CASCALES E, WALBURGER A, BOUVERET E, LAZDUNSKI C, BERNADAC A & L., J. 2001. The Tol-Pal proteins of the Escherichia coli cell envelope: an energized system required for outer membrane integrity? *Res. Microbiol.,* 152**,** 523-529.

LODISH, H., BERK, A., KAISER, C. A., KRIEGER, M., SCOTT, M. P., BRETSCHER, A., PLOEGH, H. & MATSUDAIRA, P. 2007. *Molecular Cell Biology*, W. H. Freeman.

LOFTUS, S. R., WALKER, D., MATE, M. J., BONSOR, D. A., JAMES, R., MOORE, G. R. & KLEANTHOUS, C. 2006. Competitive recruitment of the periplasmic translocation portal TolB by a natively disordered domain of colicin E9. *Proc. Natl. Acad. Sci. USA.,* 103**,** 12353-12358.

LURIA, S. E. & SUIT, J. L. 1987. Colicins and col plasmids, in Escherichia coli and Aalonella typhomurium cellular and molecular biology *American Society for Microbiology, washington, D.C,* 2.

MERRILL, A. R. & CRAMER, W. A. 1990. Identification of a voltage-responsive segment of the potential-gated colicin E1 ion channel. *Biochemistry,* 29**,** 8529-34.

MERRILL, A. R., STEER, B. A., PRENTICE, G. A., WELLER, M. J. & SZABO, A. G. 1997. Identification of a chameleon-like pH-sensitive segment within the colicin E1 channel domain that may serve as the pH-activated trigger for membrane bilayer association. *Biochemistry-US.,* 36**,** 6874-6884.

MUGA, A., GONZAL Z-MA AS, J. M., LAKEY, J. H., PATTUS, F. & SUREWICZ, W. K. 1993a. pH-dependent stability and membrane interaction of the pore-forming domain of colicin-A. *J. Biol. Chem.,* 268**,** 1553-1557.

MUGA, A., GONZALEZMANAS, J. M., LAKEY, J. H., PATTUS, F. & SUREWICZ, W. K. 1993b. PH-DEPENDENT STABILITY AND MEMBRANE INTERACTION OF THE PORE-FORMING DOMAIN OF COLICIN-A. *J. Biol. Chem.,* 268**,** 1553-1557.

MUSSE, A. A. & MERRILL, A. R. 2003. The molecular basis for the pH-activation mechanism in the channel-forming bacterial colicin E1. *J. Biol. Chem.,* 278**,** 24491-24499.

NARDI, A., SLATIN, S. L., BATY, D. & DUCHE, D. 2001. The C-terminal half of the colicin A pore-forming domain is active in vivo and in vitro. *J. Mol. Biol.,* 307**,** 1293-1303.

NELSON, D. L. & COX, M. M. 2008. Principles of biochemistry. *PRINCIPLES OF BIOCHENMISTRY.* New York: W.H freeman and company.

NES, I. F., YOON, S.-S. & DIEP, D. B. 2007. Ribosomally synthesiszed antimicrobial peptides (Bacteriocins) in lactic acid bacteria: A review. *Food Sci. Biotechnol.,* 16**,** 675-690.

NOMURA, M. 1963. Mode of action of colicins. Cold spring harbor symp. Quant. *Biol.,* 28**,** 314-324.

NOMURA, M. 1967. Colicins and related bacteriocins. *Annu. Rev. Microbiol.,* 21**,** 257-84.

OFFERMANNS, S. & ROSENTHAL, W. 2008. *Encyclopedia of Molecular Pharmacology*, Springer.

ORTEGA, A., LAMBOTTE, S. & BECHINGER, B. 2001. Calorimetric investigations of the structural stability and interactions of colicin B domains in aqueous solution and in the presence of phospholipid bilayers. *J. Biol. Chem.,* 276**,** 13563-13572.

PACE, C. N. & SCHOLTZ, J. M. 1997. Protein structure:A practical approach *In:* HAMES, B. D. (ed.) *The practical approach series.* second edition ed.

PACE, C. N., SHIRLEY, B. A. & THOMPSON, J. A. 1989. Measuring the conformational stability of a protein. *In:* CREIGHTON, T. E. (ed.) *Protein structure; a practical approach.* Oxford: IRL Press.

PADMAVATHI, P. V. L. & STEINHOFF, H. J. 2008. Conformation of the closed channel state of colicin a in proteoliposomes: An umbrella model. *J. Mol. Biol.,* 378**,** 204-214.

PARKER, M. W., BUCKLEY, J. T., POSTMA, J., TUCKER, A. D., LEONARD, K., PATTUS, F. & TSERNOGLOU, D. 1994. Structure of the aeromonas toxin proaerolysin in its water-soluble and membrane-channel states. *Nature.,* 367**,** 292-295.

PARKER, M. W., PATTUS, F., TUCKER, A. D. & TSERNOGLOU, D. 1989. Structure of the membrane-pore-forming fragment of colicin A. *Nature.,* 337**,** 93-6.

PARKER, M. W., POSTMA, J. P., PATTUS, F., TUCKER, A. D. & TSERNOGLOU, D. 1992. Refined structure of the pore-forming domain of colicin A at 2.4 A resolution. *J. Mol. Biol.,* 224**,** 639-57.

PARKER, M. W., TUCKER, A. D., TSERNOGLOU, D. & PATTUS, F. 1990. Insights into membrane insertion based on studies of colicins. *Trends. Biochem. Sci.,* 15**,** 126-9.

PATTUS, F., CAVARD, D., VERGER, R., LAZDUNSKI, C., ROSENBUSCH, J. & SCHINDLER, H. 1983a. Formation of voltage dependent pores in planar bilayers by colicin A. *In:* G.SPACH (ed.) *Physical Chemistry of Transmembrane Ion Motions.* Amsterdam: Elsevier.

PATTUS, F., MARTINEZ, M. C., DARGENT, B., CAVARD, D., VERGER, R. & LAZDUNSKI, C. 1983b. Interaction of colicin A with phospholipid monolayers and liposomes. *Biochemistry-US.,* 22**,** 5698-5707.

PILSL, H. & BRAUN, V. 1995a. Novel colicin-10 - assignment of 4 domains to tonb-dependent and tolc-dependent uptake via the tsx receptor and to pore formation. *Mol. Microbiol.,* 16**,** 57-67.

PILSL, H. & BRAUN, V. 1995b. Strong function-related homology between the pore-forming colicin K and colicin 5. *J. Bacteriol.,* 177**,** 6973-6977.

PTITSYN, O. B. 1992. The Molten Globule State. *In:* CREIGHTON, T. E. (ed.) *Protein Folding.* New York: W.H.Freeman.

PTITSYN, O. B. 1995. Molten globule and protein folding. *Adv. Protein. Chem., Vol 47,* 47**,** 83-229.

PTITSYN, O. B., PAIN, R. H., SEMISOTNOV, G. V., ZEROVNIK, E. & RAZGULYAEV, O. I. 1990. Evidence for a molten globule state as a general intermediate in protein folding. *Febs. Lett.,* 262**,** 20-24.

PUGSLEY, A. P. 1984. The ins and outs of colicins. Part I: Production, and translocation across membranes. *Microbiol. Sci.,* 1**,** 168-75.

PUGSLEY, A. P. 1985. Escherichia coli K12 strains for use in the identification and characterization of colicins. *J. Gen. Microbiol*.,131, 367-76.

PUGSLEY, A. P. 1987. Nucleotide sequencing of the structural gene for colicin N reveals homology between the catalytic, C-terminal domains of colicins A and B. *Mol Microbiol,* 1**,** 317-25.

PUGSLEY, A. P. & SCHWARTZ, M. 1983. Expression of a gene in a 400-base-pair fragment of colicin plasmid ColE2-P9 is sufficient to cause host cell lysis. *J. Bacteriol.,* 156**,** 109-14.

RICHARDSON, J. S. & RICHARDSON, D. C. 1988. Amino-acid preferences for specific locations at the ends of alpha-helices. *Science.,* 240**,** 1648-1652.

RILEY, M. A. 1993. Molecular mechanisms of colicin evolution. *Mol. Biol. Evol.,* 10**,** 1380-1395.

RILEY, M. A. & CHAVAN, M. A. 2006. *Bacteriocins: Ecology and Evolution*, Springer.

RILEY, M. A. & WERTZ, J. E. 2002. Bacteriocins: Evolution, ecology, and application. *Annu. Rev. Microbiol.,* 56**,** 117-137.

SABIK, J. F., SUIT, J. L. & LURIA, S. E. 1983. cea-kil operon of the ColE1 plasmid. *J. Bacteriol.,* 153**,** 1479-1485.

SCHENDEL, S. L. & CRAMER, W. A. 1994. On the nature of the unfolded intermediate in the in-vitro transition of the colicin E1 channel domain from the aqueous to the membrane phase. *Protein. Sci.,* 3**,** 2272-2279.

SERDYUK, I. N., ZACCAI, N. R. & ZACCAI, J. 2007. *Methods in Molecular Biophysics: Structure, Dynamics, Function*, Cambridge University Press.

SLATIN, S. L., QIU, X. Q., JAKES, K. S. & FINKELSTEIN, A. 1994. Identification of a translocated protein segment in a voltage-dependent channel. *Nature.***,** 158-161.

SOLIAKOV, A., HARRIS, J. R., WATKINSON, A. & LAKEY, J. H. 2010. The structure of Yersinia pestis Caf1 polymer in free and adjuvant bound states. *Vaccine.,* 28**,** 5746-5754.

SPECTOR, S., WANG, M. H., CARP, S. A., ROBBLEE, J., HENDSCH, Z. S., FAIRMAN, R., TIDOR, B. & RALEIGH, D. P. 2000. Rational modification of protein stability by the mutation of charged surface residues. *Biochemistry-US.,* 39**,** 872-879.

SWITZER, R. L. & GARRITY, L. F. 1999. *Experimental Biochemistry*, W. H. Freeman.

TUSZYNSKI, J. A. & KURZYNSKI, M. 2003. *Introduction to Molecular Biophysics*, Taylor & Francis.

VAN DER GOOT, F. G., DIDAT, N., PATTUS, F., DOWHAN, W. & LETELLIER, L. 1993. Role of acidic lipids in the translocation and channel activity of colicins A and N in *Escherichia coli* cells. *Eur. J. Biochem.,* 213**,** 217-221.

VAN DER GOOT, F. G., GONZ LEZ-MA AS, J. M., LAKEY, J. H. & PATTUS, F. 1991. A 'molten-globule' membrane-insertion intermediate of the pore-forming domain of colicin A. *Nature.,* 354**,** 408-10.

VAN DER GOOT, F. G., LAKEY, J., PATTUS, F., KAY, C. M., SOROKINE, O., VANDORSSELAER, A. & BUCKLEY, J. T. 1992a. Spectroscopic study of the activation and oligomerization of the channel-forming toxin aerolysin - identification of the site of proteolytic activation. *Biochemistry-US.,* 31**,** 8566-8570.

VAN DER GOOT, F. G., LAKEY, J. H. & PATTUS, F. 1992b. The molten globule intermediate for protein insertion or translocation through membranes. *Trends. Cell. Biol.,* 2**,** 343-348.

VAN DER WAL, F. J., LUIRINK, J. & OUDEGA, B. 1995. Bacteriocin release proteins: Mode of action, structure, and biotechnological application. *Fems. Microbiol. Rev.,* 17**,** 381-399.

VANDERGOOT, F. G., DIDAT, N., PATTUS, F., DOWHAN, W. & LETELLIER, L. 1993. ROLE OF ACIDIC LIPIDS IN THE TRANSLOCATION AND CHANNEL ACTIVITY OF COLICIN-A AND COLICIN-N IN ESCHERICHIA-COLI-CELLS. *Eur. J. Biochem.,* 213**,** 217-221.

VETTER, I. R., PARKER, M. W., PATTUS, F. & TSERNOGLOU, D. 1996. Insights into membrane insertion based on studies of colicins. *In:* PARKER, M. W. (ed.) *Protein Toxin Structure.* Austin TX: R. G. Landes Company.

VETTER, I. R., PARKER, M. W., TUCKER, A. D., LAKEY, J. H., PATTUS, F. & TSERNOGLOU, D. 1998. Crystal structure of a colicin N fragment suggests a model for toxicity. *Structure.,* 6**,** 863-874.

WEAVER, C., REDBORG, A. H. & KONISKY, J. 1981. Plasmid-determined immunity of Esherichia coli K-12 to colicin Ia is mediated by a plasmid encoded membrane protein. *J. Bacteriol.,* 148**,** 817-828.

WENDT, L. 1970. Mechanism of colicin action-early events. *J. Bacteriol.,* 104**,** 1236-&.

WIENER, M., FREYMANN, D., GHOSH, P. & STROUD, R. M. 1997. Crystal structure of colicin Ia. *Nature.,* 385**,** 461-464.

WILSON, K. & WALKER, J. 2010. *Principles and Techniques of Biochemistry and Molecular Biology*, Cambridge University Press.

WOLSTENHOLME, G. E. W. & O'CONNOR, C. M. (eds.) 2008. *Drug Resistance in Micro-Organisms: Mechanisms of Development*: John Wiley & Sons.

WU, J. W. & FILUTOWICZ, M. 1999. Hexahistidine (His(6))-tag dependent protein dimerization: A cautionary tale. *Acta. Biochim. Pol.,* 46**,** 591-599.

YANG, A. S. & HONIG, B. 1993. On the pH-dependence of protein stability. *J. Mol. Biol.,* 231**,** 459-474.

YOUNG, R. 1992. bacteriophage lysis: mechanism and regulation. *Microbiol. Rev.,* 56**,** 430-481.

ZAKHAROV, S. D. & CRAMER, W. A. 2002. Colicin crystal structures: pathways and mechanisms for colicin insertion into membranes. *BBA-Biomembranes.,* 1565**,** 333-346.

ZAKHAROV, S. D., ZHALNINA, M. V., SHARMA, O. & CRAMER, W. A. 2006. The colicin E3 outer membrane translocon: Immunity protein release allows interaction of the cytotoxic domain with OmpF porin. *Biochemistry-US.,* 45**,** 10199-10207.

ZHANG, Y. L. & CRAMER, W. A. 1993. Intramembrane helix-helix interaction as the basis of inhibition of colicin e1 ion channel by its immunity protein. *Biophys. J.***,** 345.

Appendix

DNA sequence of wild type of colicin A pore forming domain that I used in the project:

Master gene sequence with NdeI (5') and BamHI (3') restriction sites

catatgcatcaccatcaccatcactcgagcgttgcggaaaaagccaaagatgagcgggagctgcttgaaaaaaccagtga
actgattgctggtatgggagataaaatcggcgagcatcttggagataaatataaggcgatagcgaaagatattgcggacaat
attaaaaattttcaggggaagaccatccgtagctttgatgatgcaatggcatcgctgaataaaatcacagccaacccagccat
gaaaattaataaggcggacagagatgctctggttaatgcctggaaacatgttgatgctcaggatatggcgaataaactgggt
aatctcagcaaggcttttaaagtcgccgacgtggtgatgaaggttgagaaggtccgggagaagagcattgaggggtatga
aaccgggaactgggggccgctgatgctggaggtggaatcctgggtgctcagtggtatagcttcctctgttgctctggggatt
ttttccgctacattaggagcgtatgccttatctcttggagttcctgctattgctgttggtatcgccggtattctactcgcagcagtt
gttggtgcgttaattgatgataagtttgcagatgctttgaataatgaaataatccgacctgcacattaagtttgattaaaaggca
gtgttactgccttttttttaataatttaaatgcagtgactggtgtgaatagcgtgatataggttgtaaagaaaattattgagtacag
accaatataaaataacaatagtgttgcgtcattggtggccattaatctgacaggccttccggctgttgttaactcaaaatttcgg
aagcagaatagataaacaattgctgcatagaccaatggacttaataaacaagatctaagtacagtgttcctgtcggtgttgttg

attaattttctgattttaaaggttaggatgtatagtattaaggcaaggaaaggcgctgttttacaataaatatccataacctttgtc
attaatgggttgtaggaggatgttattgacgggagattctcagtgcttgtagcaattttacgtaaaagtaaagcgtctggcgtttt
gtagtaaacaacaaaaatgcacaataatggtattaatccgattattataaataaatacaatgcgttattggcctttctgttgtccgt
atctattgagtgttcattcatcattattaacctatataattataataaaaggtatatatgaaaaaaattattatttgtgtcattttactgg
cgatcatgctgcttgccgcatgtcaagtcaacaatgtcagggatactggaggtggttctgtttcaccatcatctatcgttaccg
gagtgagttaaaggatcc

Yellow = Colicin A P-domain C-terminal His-tag (5’-3’) Red = Colicin A immunity protein (3’-5’) Green = Colicin A lysis protein (5’-3’)

Protein sequence of wild type of colicin A pore forming domain and the mutants that I used in the project.

Colicin A P-domain amino acid sequence (5’-3’)

MHHHHHHSSVAEKAKDERELLEKTSELIAGMGDKIGEHLGDKYKAIAKDIAD
NIKNFQGKTIRSFDDAMASLNKITANPAMKINKADRDALVNAWKHVDAQD
MANKLGNLSKAFKVADVVMKVEKVREKSIEGYETGNWGPLMLEVESWVLS
GIASSVALGIFSATLGAYALSLGVPAIAVGIAGILLAAVVGALIDDKFADALNN
EIIRPAH

colicin A immunity protein amino acid sequence (3’-5’)

MMNEHSIDTDNRKANNALYLFIIIGLIPLLCIFVVYYKTPDALLLRKIATSTENL
PSITSSYNPLMTKVMDIYCKTAPFLALILYILTFKIRKLINNTDRNTVLRSCLLS
PLVYAAIVYLFCFRNFELTTAGRPVRLMATNDATLLLFYIGLYSIIFFTTYITLF
TPVTAFKLLKKRQ

Colicin A lysis protein amino acid sequence: (5’-3’)

MKKIIICVILLAIMLLAACQVNNVRDTGGGSVSPSSIVTGVS

1: ColA-P WT

MHHHHHHSSVAEKAKDERELLEKTSELIAGMGDKIGEHLGDKYKAIAKDIAD
NIKNFQGKTIRSFDDAMASLNKITANPAMKINKADRDALVNAWKHVDAQD
MANKLGNLSKAFKVADVVMKVEKVREKSIEGYETGNWGPLMLEVESWVLS
GIASSVALGIFSATLGAYALSLGVPAIAVGIAGILLAAVVGALIDDKFADALNN
EIIRPAH

2: ColA-P D216A

MHHHHHHSSVAEKAKDERELLEKTSELIAGMGDKIGEHLGAKYKAIAKDIAD
NIKNFQGKTIRSFDDAMASLNKITANPAMKINKADRDALVNAWKHVDAQD
MANKLGNLSKAFKVADVVMKVEKVREKSIEGYETGNWGPLMLEVESWVLS
GIASSVALGIFSATLGAYALSLGVPAIAVGIAGILLAAVVGALIDDKFADALNN
EIIRPAH

3: ColA D227A

MHHHHHHSSVAEKAKDERELLEKTSELIAGMGDKIGEHLGDKYKAIAKDIAA
NIKNFQGKTIRSFDDAMASLNKITANPAMKINKADRDALVNAWKHVDAQD
MANKLGNLSKAFKVADVVMKVEKVREKSIEGYETGNWGPLMLEVESWVLS
GIASSVALGIFSATLGAYALSLGVPAIAVGIAGILLAAVVGALIDDKFADALNN
EIIRPAH

4: ColA D241A

MHHHHHHSSVAEKAKDERELLEKTSELIAGMGDKIGEHLGDKYKAIAKDIAD
NIKNFQGKTIRSFADAMASLNKITANPAMKINKADRDALVNAWKHVDAQD
MANKLGNLSKAFKVADVVMKVEKVREKSIEGYETGNWGPLMLEVESWVLS
GIASSVALGIFSATLGAYALSLGVPAIAVGIAGILLAAVVGALIDDKFADALNN
EIIRPAH

5: ColA D274A

MHHHHHHSSVAEKAKDERELLEKTSELIAGMGDKIGEHLGDKYKAIAKDIAD
NIKNFQGKTIRSFDDAMASLNKITANPAMKINKADRDALVNAWKHVAAQD
MANKLGNLSKAFKVADVVMKVEKVREKSIEGYETGNWGPLMLEVESWVLS
GIASSVALGIFSATLGAYALSLGVPAIAVGIAGILLAAVVGALIDDKFADALNN
EIIRPAH

6: ColA D372A

MHHHHHHSSVAEKAKDERELLEKTSELIAGMGDKIGEHLGDKYKAIAKDIAD
NIKNFQGKTIRSFDDAMASLNKITANPAMKINKADRDALVNAWKHVDAQD
MANKLGNLSKAFKVADVVMKVEKVREKSIEGYETGNWGPLMLEVESWVLS
GIASSVALGIFSATLGAYALSLGVPAIAVGIAGILLAAVVGALIDAKFADALNN
EIIRPAH

7: ColA D376A

MHHHHHHSSVAEKAKDERELLEKTSELIAGMGDKIGEHLGDKYKAIAKDIAD
NIKNFQGKTIRSFDDAMASLNKITANPAMKINKADRDALVNAWKHVDAQD
MANKLGNLSKAFKVADVVMKVEKVREKSIEGYETGNWGPLMLEVESWVLS
GIASSVALGIFSATLGAYALSLGVPAIAVGIAGILLAAVVGALIDDKFAAALNN
EIIRPAH 206

8: ColA D216N

MHHHHHHSSVAEKAKDERELLEKTSELIAGMGDKIGEHLGNKYKAIAKDIAD
NIKNFQGKTIRSFDDAMASLNKITANPAMKINKADRDALVNAWKHVDAQD
MANKLGNLSKAFKVADVVMKVEKVREKSIEGYETGNWGPLMLEVESWVLS
GIASSVALGIFSATLGAYALSLGVPAIAVGIAGILLAAVVGALIDDKFADALNN
EIIRPAH

9: ColA D227N

MHHHHHHSSVAEKAKDERELLEKTSELIAGMGDKIGEHLGDKYKAIAKDIAN
NIKNFQGKTIRSFDDAMASLNKITANPAMKINKADRDALVNAWKHVDAQD
MANKLGNLSKAFKVADVVMKVEKVREKSIEGYETGNWGPLMLEVESWVLS
GIASSVALGIFSATLGAYALSLGVPAIAVGIAGILLAAVVGALIDDKFADALNN
EIIRPAH

10: ColA D241N

MHHHHHHSSVAEKAKDERELLEKTSELIAGMGDKIGEHLGDKYKAIAKDIAD
NIKNFQGKTIRSFNDAMASLNKITANPAMKINKADRDALVNAWKHVDAQD
MANKLGNLSKAFKVADVVMKVEKVREKSIEGYETGNWGPLMLEVESWVLS
GIASSVALGIFSATLGAYALSLGVPAIAVGIAGILLAAVVGALIDDKFADALNN
EIIRPAH

11: ColA D274N

MHHHHHHSSVAEKAKDERELLEKTSELIAGMGDKIGEHLGDKYKAIAKDIAD
NIKNFQGKTIRSFDDAMASLNKITANPAMKINKADRDALVNAWKHVNAQD
MANKLGNLSKAFKVADVVMKVEKVREKSIEGYETGNWGPLMLEVESWVLS

GIASSVALGIFSATLGAYALSLGVPAIAVGIAGILLAAVVGALIDDKFADALNN
EIIRPAH

12: ColA D372N

MHHHHHHSSVAEKAKDERELLEKTSELIAGMGDKIGEHLGDKYKAIAKDIAD
NIKNFQGKTIRSFDDAMASLNKITANPAMKINKADRDALVNAWKHVDAQD
MANKLGNLSKAFKVADVVMKVEKVREKSIEGYETGNWGPLMLEVESWVLS
GIASSVALGIFSATLGAYALSLGVPAIAVGIAGILLAAVVGALIDNKFADALNN
EIIRPAH

13: ColA D216E

MHHHHHHSSVAEKAKDERELLEKTSELIAGMGDKIGEHLGEKYKAIAKDIAD
NIKNFQGKTIRSFDDAMASLNKITANPAMKINKADRDALVNAWKHVDAQD
MANKLGNLSKAFKVADVVMKVEKVREKSIEGYETGNWGPLMLEVESWVLS
GIASSVALGIFSATLGAYALSLGVPAIAVGIAGILLAAVVGALIDDKFADALNN
EIIRPAH

14: ColA D274E

MHHHHHHSSVAEKAKDERELLEKTSELIAGMGDKIGEHLGDKYKAIAKDIAD
NIKNFQGKTIRSFDDAMASLNKITANPAMKINKADRDALVNAWKHVEAQDM
ANKLGNLSKAFKVADVVMKVEKVREKSIEGYETGNWGPLMLEVESWVLSGI
ASSVALGIFSATLGAYALSLGVPAIAVGIAGILLAAVVGALIDDKFADALNNEI
IRPAH

15: ColA D372E

MHHHHHHSSVAEKAKDERELLEKTSELIAGMGDKIGEHLGDKYKAIAKDIAD
NIKNFQGKTIRSFDDAMASLNKITANPAMKINKADRDALVNAWKHVDAQD
MANKLGNLSKAFKVADVVMKVEKVREKSIEGYETGNWGPLMLEVESWVLS
GIASSVALGIFSATLGAYALSLGVPAIAVGIAGILLAAVVGALIDEKFADALNN
EIIRPAH

16: ColA D216Q

MHHHHHHSSVAEKAKDERELLEKTSELIAGMGDKIGEHLGQKYKAIAKDIAD

NIKNFQGKTIRSFDDAMASLNKITANPAMKINKADRDALVNAWKHVDAQD
MANKLGNLSKAFKVADVVMKVEKVREKSIEGYETGNWGPLMLEVESWVLS
GIASSVALGIFSATLGAYALSLGVPAIAVGIAGILLAAVVGALIDDKFADALNN
EIIRPAH

17: ColA D274Q

MHHHHHHSSVAEKAKDERELLEKTSELIAGMGDKIGEHLGDKYKAIAKDIAD
NIKNFQGKTIRSFDDAMASLNKITANPAMKINKADRDALVNAWKHVQAQD
MANKLGNLSKAFKVADVVMKVEKVREKSIEGYETGNWGPLMLEVESWVLS
GIASSVALGIFSATLGAYALSLGVPAIAVGIAGILLAAVVGALIDDKFADALNN
EIIRPAH

18: ColA D372Q

MHHHHHHSSVAEKAKDERELLEKTSELIAGMGDKIGEHLGDKYKAIAKDIAD
NIKNFQGKTIRSFDDAMASLNKITANPAMKINKADRDALVNAWKHVDAQD
MANKLGNLSKAFKVADVVMKVEKVREKSIEGYETGNWGPLMLEVESWVLS
GIASSVALGIFSATLGAYALSLGVPAIAVGIAGILLAAVVGALIDQKFADALNN
EIIRPAH

19: ColA D227A-A269D

MHHHHHHSSVAEKAKDERELLEKTSELIAGMGDKIGEHLGDKYKAIAKDIAD
NIKNFQGKTIRSFDDAMASLNKITANPAMKINKADRDALVNDWKHVDAQD
MANKLGNLSKAFKVADVVMKVEKVREKSIEGYETGNWGPLMLEVESWVLS
GIASSVALGIFSATLGAYALSLGVPAIAVGIAGILLAAVVGALIDDKFADALNN
EIIRPAH

20: ColA S239D-D241A

MHHHHHHSSVAEKAKDERELLEKTSELIAGMGDKIGEHLGDKYKAIAKDIAD
NIKNFQGKTIRDFANDAMASLNKITANPAMKINKADRDALVNAWKHVDAQ
DMANKLGNLSKAFKVADVVMKVEKVREKSIEGYETGNWGPLMLEVESWVL

SGIASSVALGIFSATLGAYALSLGVPAIAVGIAGILLAAVVGALIDDKFADALN
NEIIRPAH