University of Florence
International Doctorate in Structural Biology
Cycle XVIII (01/2003 to 12/2005)
Magnetic Resonance Center
CERM

Expression and Structural Characterization of Bacterial Metalloproteins: Cytochrome c553 variants and Metal Translocating ATPases, PacS and ZiaA

Ph.D. thesis
Submitted by

MURUGENDRA VANAROTTI

Tutor

Prof.Alessandro Quattrone

Coordinator

Prof.Claudio Luchinat

Acknowledgements

I owe my deep regards to Prof. Ivano Bertini, Director, Magnetic Resonance Center (CERM) for giving me an opportunity to conduct my doctoral work over CERM and for giving me suggestions, help and hints.

I take this opportunity to sincerely thank Prof.Lucia Banci for their valuable suggestions.

To Simone Ciofi, I owe thanks beyond measure for his tremendous help and patience in completion of my thesis.

Many thanks to Luisa Poggi for helping me and for giving me the idea and data for my thesis.

I take this opportunity to sincerely thank Prof. Allessandro Quattrone and Prof.Antonio Rosato for their valuable suggestions.

There is a strong temptation to try to individually acknowledge everyone who has helped me to reach this point. Sadly, such a list would be much too long and it is inevitable that many people would be unintentionally omitted. I would thus like to thank everyone who, knowingly or otherwise, has provided support, encouragement and assistance along the way. However, there are others who have played a major role in the production of this thesis and who deserve a more personal note of gratitude.

Special thanks to all my friends and colleagues.

My deepest thanks go to my parents and my wife for their unconditional love and support.

LIST OF CONTENTS

RESULTS

LIST OF ABBREVIATIONS

NMR	Nuclear magnetic resonance spectroscopy
WT	Wild Type
GdmCl	Guanidinium chloride
E.coli	*Escherichia coli*
PCR	Polymerase Chain Reaction
DNA	Deoxynucleic acid
ATP	Adenosine triphosphate
CD	Circular Dichroism
PDB	Protein Data Bank
3D	Three-dimensional
HSQC	Heteronuclear single quantum coherence
NOESY	Nuclear Overhauser Effect Spectroscopy
NOE	Nuclear Overhauser Effect
cDNA	Complimentary 2'-deoxy-5'-ribonucleic acid
IPTG	Isopropyl-β-galactosidase
COSY	Correlated spectroscopy
RMSD	Root Mean Square Deviation
ssDNA	single-stranded DNA
Bpcytc	*Bacillus pateruii* cytochrome c
Met	Methionine
Tm	Melting Temperature
FPLC	Fast Protein Liquid Chromatography
DTT	Dithiothreitol
ADAM	A Disintegrin and Metalloprotease
ZNF9	Zinc Finger 9
SDS-PAGE	Sodium dodecyl sulphate- Polyacrylamide gel electrophoresis

INTRODUCTION

Seeing the three-dimensional structure of a biological molecule for the first time can be a revelation. For example, James Watson and Francis Crick's model of the DNA double helix offers immediate insight into how the molecule copies itself. The minute you see it, you can oversee how heredity works and one look gives you the idea. Since this achievement in 1953, the study of molecular structures has grown into a critically important science that is shaping much of biomedical research. When scientists can see the location of the individual atoms that make up a protein and examine just how the entire structure relates to the structures of other proteins, they can begin to understand the protein's function. They are no longer flying blind. The brightly colored images of newly solved structures now appear in leading science journals nearly every week, and the rate of discoveries in structural biology is accelerating rapidly. X-ray crystallography—the technique Max Perutz used in his painstaking, 22-year effort to find the structure of hemoglobin—has seen such dramatic improvements that structures can sometimes be solved within a few months. Nuclear magnetic resonance (NMR) spectroscopy, which previously could be used only with very small molecules, can now analyze macromolecules of 1 kDa to 50 kDa. Since it carries out this analysis while the molecules are in solution, it bypasses the lengthy and sometimes erratic process of making crystals of the proteins. At the same time, computer power is increasing geometrically, enabling scientists to tackle more difficult problems. This massive progress is changing the nature of structural biology. Whereas in the past researchers concentrated on solving the structures of whatever proteins yielded suitable crystals, now their main goal is to solve interesting biological problems. Increasingly, they define themselves by their subject matter: not as crystallographers or NMR spectroscopists, but as biologists who want to

find out how genes are turned on, for instance, or how drugs bind to enzymes. There has been an explosive increase in the number of structures that open up new areas of analysis and these structures often serve as springboards for other biological experiments. In the post genomic era, structural biology will become a central discipline for the explanation, linking and exploitation of biological data in the life sciences, in academic research as well as in applications in the biotechnological, agricultural and pharmaceutical industry. To bridge the widening gap between rapidly increasing information on genome sequences and limited knowledge on the function of gene products, a quantitative understanding of the three-dimensional structure of proteins, their folding and biogenesis, and their interactions with other molecules is required and will be the pivotal challenge in molecular biological sciences in the years ahead. Efficient experimental structure determination of soluble and membrane proteins and the biophysical study of the protein folding process will be a central component and provide the basis for the future development of basic research and practical applications such as drug design and delivery. The structure determination can be complemented, with NMR spectroscopy, by dynamic characterization. Local mobility of the protein could have a relevant correlation with the functional behavior of the proteins. Indeed, it is now widely recognized that fluctuations of macromolecules play a relevant role in determining their stability, their function and recognition with biological partners. NMR has given a great contribution to the study of local dynamics. In particular, analysis of ^{15}N spin relaxation has now become a powerful method to determine local mobility of the backbone. Local dynamics of side chains, despite seldom characterized, can also give very interesting information. Indeed, side chains have more degrees of freedom easily accessible than the backbone, and thus

residues on the surface of the protein can modulate the protein surface itself by changing, for example, the electrostatics of charged residues or by leaving internal regions more solvent exposed. This can affect the protein stability or the interaction with biological partners. In this work we combined protein expression and biosynthetic isotopic enrichment, which could be achieved by recently developed methods of biotechnology, with NMR spectroscopy.

Aims and overview of the research

The work carried out over three years of doctoral studies has been devoted to cloning, expression, purification and structural characterization of metalloproteins, and in particular cytochrome *c* variants, copper and zinc binding proteins. Metalloproteins contain a metal ion that confers peculiar features to the protein. The use of metals in living organisms is mainly due to their redox and acid/base properties and to the relatively easy feasibility of metal ion transfer in different locations or compartments of the cell, where they can be employed in various biological processes, such as electron transfer reactions, oxygen transport, and in a large variety of catalytic processes. This kind of proteins constitute a significant share of the total genome products (30-40% expected), but on the other hand they are not very much structurally studied; for instance, only the 0.5% of the structures deposited in the PDB are copper binding proteins (1). Therefore the structural characterization of this type of protein is an important aspect of structural biology.

The cytochromes *c* plays important roles in electron transfer in a large variety of eukaryotic and prokaryotic organisms. Despite the similarity of essential function of the molecule - passing of an electron by changing the oxidation state of the iron - there is significant divergence of sequence and structure of the protein as it appears in different organisms. It has been suggested that these differences are due to diverse adaptation to the metabolism in which the cytochromes function (2). Class I cytochrome c (covalent attachment of the c-type heme at two Cys residues, and His and Met ligation of the iron) can be broadly separated into two major families based on the size of the molecule. The structural and sequential similarity between eukaryotic cytochromes *c* and prokaryotic cytochromes *c2*, essentially from photosynthetic bacteria, groups these cytochromes to form a single category of large cytochromes (Ia).

In most of Gram-positive microorganisms cytochromes belonging to Ia subclass is usually membrane-bound proteins due to the lack of a true periplasmic space. The difficulties associated with the study of such water insoluble proteins may explain why knowledge of the properties and physiological role of c-type cytochromes in Gram-positive bacteria is very limited as compared to the information available for Gram-negative bacteria and eukaryotes [3].

Recently both solution (4) and crystal structure (5) of minimal 71-residue mono-heme cytochrome c: the soluble fragment of cytochrome c553 from *Bacillus pasteurii* (Bpcytc) was undertaken. There is practically no differences between the backbone conformations of reduced and oxidized states of the wild-type (WT) protein [6]. Also the dynamic properties of Bpcytc in the two physiologically relevant oxidation states are similar [6], with reduced Bpcytc being slightly more rigid, as commonly observed

(although to a larger extent) for c-type cytochromes [7]. WT oxidized Bpcytc features unchanged His/Met heme axial coordination at pH values up to 12, at variance with mitochondrial cytochromes c [8, 9]. Heme mis-ligation is induced at neutral pH by high concentration (>4 M) of guanidinium chloride (GdmCl), and occurs concomitantly with extensive unfolding of the polypeptide chain [8]. There are several elements of evidence, including the significant analogy of the organization of opening units across the protein structures [11], suggesting that the relatively uncommon stability of the Met loop in oxidized Bpcytc is mainly dependent on the local properties of the loop region itself, i.e. its primary sequence rich in Gly and Pro residues, its short size and the low number of long-range contacts with the remainder of the protein [11]. Since the binding of the axial Met represents a crucial and yet not fully understood aspect of the biochemistry of cytochromes c.

Trying to rationalize the increase of the stability of the Met loop region of Bpcytc, the first part of my PhD work is consisted of developing *Bacillus pasteruii* cytochrome c mutants at three locations within the Met loop. These appeared indeed to be residues potentially capable of affecting the loop stability (8). These sites have been probed through site-directed mutagenesis and the generated mutants (Q68K, P72A, P72G, I75A and I75V) have been investigated in the oxidized state through a combination of CD and NMR spectroscopy.

The second part of my work was focused on the study of copper and zinc binding proteins which are involved in the metal trafficking within bacterial organisms.

Copper is indeed a micronutrient that plays an essential role in biology, serving as a co-factor for several enzymes that include Cu, Zn-superoxide dismutase, cytochrome

oxidase, lysyl oxidase, and ceruloplasmin (11, 12). Dietary copper limitation studies in animals, as well as the existence of human genetic diseases of copper homeostasis such as Menkes and Wilson disease, underscore critical roles for proper copper absorption in the intestine and distribution to the organs and tissues to serve as an essential biochemical co-factor for enzymatic activities and other important biological processes (13-16). At the cellular level, copper is transported at the plasma membrane and distributed to cellular proteins and compartments for the incorporation of copper into copper-dependent proteins. Studies in yeast cells first identified genes encoding high affinity copper ion transport proteins in the plasma membrane. Either prior to or concomitant with high affinity uptake, Cu(II) is reduced to Cu(I) by one or more metalloreductases encoded by the *FRE1* through *FRE7* genes (17,18). Cu(I) is delivered across the plasma membrane by the high affinity transporter Ctr1 or Ctr3 in *S. cerevisiae* (19-22). After crossing the plasma membrane, copper is delivered to the secretory compartment, mitochondria, and cytosolic enzymes by the target-specific copper chaperone proteins, Atx1/Atox1, Cox17, and CCS, respectively (23-27). Similarly, bacteria utilize metallochaperone and metal transporters as heavy P_1-type of ATPases in copper and zinc trafficking and/or detoxification of process.

Cyanobacteria *Synechocystis* PCC 6803 contain internal thylakoid membranes where oxygen-evolving photosynthetic electron transport occurs. Respiratory electron transport occurs in both thylakoid and plasma membranes (28). Thylakoid membranes contain two protein complexes that include photosystems II and I. Within photosynthetically active cells, mobile soluble carriers shuttle electrons between these two complexes. Some cyanobacteria and green algae (29, 30) adapt to copper deficiency

by exploiting alternative carriers. In copper-sufficient *Synechocystis* PCC 6803, electrons transfer between the complexes via copper in plastocyanin (PetE) whereas under copper deficiency heme iron in cytochrome c_6 (PetJ) is used (31). Both PetE and PetJ are located "inside" the thylakoid lumen. One subset of proteins is imported into cyanobacterial (and plant chloroplast) thylakoids via the Sec system, whereas others are imported via a ΔpH-dependent pathway (32). The latter transports folded proteins, and its substrates tend to be proteins that require complex cofactors, thereby avoiding separate thylakoid import of the cofactors. Plastocyanin is imported using Sec indicating that a copper delivery system into this compartment is required when *Synechocystis* PCC 6803 switches from PetJ to PetE. Higher plant chloroplasts rely exclusively on plastocyanin for electron transport between the two photosystems (29), and therefore thylakoid copper import is predicted to be especially important in higher plants.

Ion transport proteins in subcellular compartments allow specialized biochemistry requiring specific ionic conditions. The integral membrane P-type ATPases are an important class of ion transport proteins that serve to maintain suitable ionic conditions. More than 100 P-type ATPases have been identified so far; they play critical roles in ion homeostasis in species as diverse as bacteria and humans, in a variety of membrane types and involve a number of different cations. Recently two such copper transporters in *Synechocystis* PCC 6803 have been described, CtaA and PacS. Both ATPases contains an N-terminal region with a conserved metal binding motif CXXC (34). They are required for efficient switching to the use of copper in plastocyanin rather than heme iron in cytochrome c_6 for photosynthetic electron transport (35). Disruption of *ctaA* gene reduced the total amount of copper cell^{-1}, whereas disruption of *PacS* gene

conferred copper sensitivity indicating that one ATPase contribute to copper import where the other one contributes to copper compartmentalization (Fig.1). This organism also contains two additional CPx-type ATPases, ZiaA and CoaT, which are known to be expressed in response to and required for growth in (elevated) zinc (36) and cobalt (37), respectively (Fig.1).Also these two ATPases possess a N-terminal tail containing potential metal binding motifs (34).

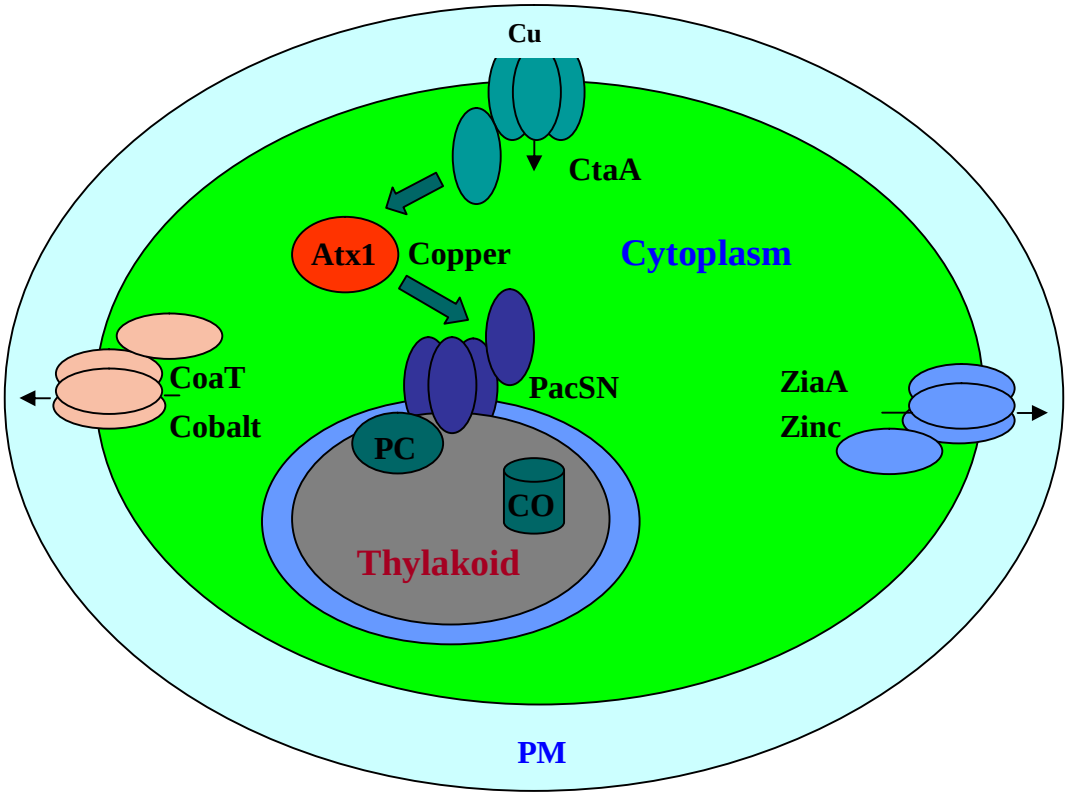


Figure 1. Copper transport and trafficking in *Synechocystis* PCC 6803.

Synechocystis PCC 6803 also contains soluble copper metallochaperone, Atx1 (ScAtx1) which is required for normal photosynthetic electron transfer via plastocyanin and for the

activity of a second thylakoid located copper protein, a caa_3-type cytochrome oxidase (33, 34). ScAtx1 directly interacts with soluble amino-terminal domains of P_1-type copper ATPases.

This simple bacterial organism represents an attractive model to investigate the metal selectivity and the action of ATX1 like metallochaperone in relation to understanding the mechanism of copper acquisition and release. In the second part of my research work I have therefore cloned, expressed and structurally characterized the N-terminal metal binding domain of PacS and ZiaA (here after $PacS_N$ and $ZiaA_N$, respectively) as well as performed protein-protein interaction studies to investigate the copper transfer between ScAtx1 and $PacS_N$.

References

1.Andreini, C., Bertini, I., & Rosato, A. (2004) *Bioinformatics.* 20(9) :1373- 80.

2. Pettigrew, G. W., and Moore, G. R. (1987) in *Cytochromes c: Biological Aspects*, Springer-Verlag, Berlin.

3. R. A. Scott, A. G. Mauk, Cytochrome c. A multidisciplinary approach, University Science Books, Sausalito, CA, 1996.

5. Banci L, Bertini I, Ciurli S, Dikiy A, Dittmer J, Rosato A, Thompsett AR.(2002) *Chembiochem.* 3(4):299-310.

6. Benini S, Gonzalez A, Rypniewski WR, Wilson KS, Van Beeumen JJ, Ciurli S. (2000) *Biochemistry*. 39(43):13115-26.

7. Bartalesi I, Bertini I, Rosato A (2003) *Biochemistry* 42:739–745.

8. Banci L, Bertini I, Luchinat C, Turano P (2000) In: Kadish KM, Smith KM, Guilard R

(eds) The porphyrin handbook.Academic Press, San Diego, pp 323–350.

9. Bartalesi I, Bertini I, Ghosh K, Rosato A, Turano P (2002) *J Mol Biol* 321:693–701.

10.Benini S, Borsari M, Ciurli S, Dikiy A, Lamborghini M (1998)*J Biol Inorg Chem* 3:371–382.

11. Bartalesi I, Rosato A, Zhang W (2003) *Biochemistry* 42:10923–10930.

12. Linder, M. C. (1991) *Biochemistry of Copper*, Plenum Press, New York 2. Peña, M. M. O., Lee, J., and Thiele, D. J. (1999) *J. Nutr.* **129**, 1251-1260.

13. Danks, D. M. (1988) *Annu. Rev. Nutr.* **8**, 235-257.

14. Prohaska, J. R. (2000) *Nutrition* **16**, 502-504.

15. DiDonato, M., and Sarkar, B. (1997) *Biochim. Biophys. Acta* **1360**, 3-16.

16. Schaefer, M., and Gitlin, J. D. (1999) *Am. J. Physiol.* **276**, G311-G314.

17. Hassett, R., and Kosman, D. J. (1995) *J. Biol. Chem.* **270**, 128-134.

18. Georgatsou, E., Mavrogiannis, L. A., Fragiadakis, G. S., and Alexandraki, D. (1997) *J. Biol. Chem.* **272**, 13786-13792.

19. Martins, L. J., Jensen, L. T., Simons, J. R., Keller, G. L., and Winge, D. R. (1998) *J. Biol. Chem.* **273**, 23716-23721.

20. Dancis, A., Yuan, D. S., Haile, D., Askwith, C., Eide, D., Moehle, C., Kaplan, J., and Klausner, R. D. (1994) *Cell* **76**, 393-402.

21. Knight, S. A. B., Labbé, S., Kwon, L. F., Kosman, D. J., and Thiele, D. J. (1996) *Genes Dev.* **10**, 1917-1929.

22. Labbé, S., Peña, M. M. O., Fernandes, A. R., and Thiele, D. J. (1999) *J. Biol. Chem.* **274**, 36252-36260.

23. Zhou, H., and Thiele, D. J. (2001) *J. Biol. Chem.* **276**, 20529-20535.

24. Valentine, J. S., and Gralla, E. B. (1997) *Science* **278**, 817-818.

25. Askwith, C., and Kaplan, J. (1998) *Trends Biochem. Sci.* **23**, 135-138.

26. Harrison, M. D., Jones, C. E., Solioz, M., and Dameron, C. T. (2000) *Trends Biochem. Sci.* **25**, 29-32.

27. Culotta, V. C., Lin, S. J., Schmidt, P., Klomp, L. W., Casareno, R. L., and Gitlin, J. (1999) *Adv. Exp. Med. Biol.* **448**, 247-254.

28. Peschek, G. A. (1981) *Biochim. Biophys. Acta* **635**, 470-475.

29. Kerfeld, C. A., and Krogmann, D. W. (1998) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **49**, 397-425

30. Moseley, J., Quinn, J., Eriksson, M., and Merchant, S. (2000) *EMBO J.* **19**, 2139- 2151

31. Zhang, L., McSpadden, B., Pakrasi, H. B., and Whitmarsh, J. (1992) *J. Biol. Chem.* **267**, 19054-19059.

32. Bogsch, E., Brink, S., and Robinson, C. (1997) *EMBO J.* **16**, 3851-3859.

33. Bogsch, E., Brink, S., and Robinson, C. (1997) *EMBO J.* **16,** 3851-3859.

34. Arnesano, F., Banci, L., Bertini, I., Cantini, F., Ciofi-Baffoni, S., Huffman, D. L., and O'Halloran, T. V. (2001) *J. Biol. Chem.* **276,** 41365–41376.

35. Tottey, S., Rich, P. R., Rondet, S. A. M., and Robinson, N. J. (2001) *J. Biol. Chem.* **276,** 19999-20004.

36. Thelwell, C., Robinson, N. J., and Turner-Cavet, J. S. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 10728-10733.

37. Rutherford, J. C., Cavet, J. S., and Robinson, N. J. (1999) *J. Biol. Chem.* **274**, 25827- 25832.

METHODS

A. Genome browsing for analysis of protein sequences

A number of bioinformatics web servers with databases and programs have been created before and after the advent of genome sequences with the aim of providing the scientific community with tools for searching gene banks, for the analysis of protein sequences, for the prediction of a variety of protein properties.

Some available databases and softwares used for genome browsing and sequence analysis are;

- **NCBI (www.ncbi.nlm.nih.gov/Entrez/) -** This web site integrates information from several databases (Swissprot, EMBL, all gene Bank, etc)
- **Pfam (http://pfam.wustl.edu)** - A collection of different protein family organized for different domain obtained from multiple alignment
- **BLAST (www.ncbi.nlm.nih.gov/BLAST/)**: Standard BLAST (Basic Local Alignment Search Tool) is a set of similarity search programs designed to explore all of the available sequence databases regardless of whether the query is protein or DNA. PHI-BLAST is designed to search for proteins that contain a pattern specified by the user, and simultaneously are similar to the query sequence.
- **PROSITE (www.expasy.org/prosite/)** SCANPROSITE allows to scan a protein sequence, provided by the user, for the occurrence of patterns and profiles stored in the PROSITE database, or to search in protein databases all sequences with a user entered pattern.
- **STRING (dag.embl-heidelberg.de)** STRING is a database of predicted functional associations among genes/proteins. Genes of similar function tend to be maintained in close neighborhood and tend to be present or absent together.

- **SMART:** SMART (a Simple Modular Architecture Research Tool) allows the identification and annotation of genetically mobile domains and the analysis of domain architectures.

These tools have been widely utilized during my PhD research work in order to select the protein constructs which I have studied.

B. Cloning, Expression and Purification of Recombinant Proteins

Structural genomics (1) or structural proteomics (2, 3) can be defined as the quest to obtain the three-dimensional structures of all proteins. However, the preparation of excellent protein sample for structural characterization remains a significant problem for this new field. The development of better and faster methods to clone, express and purify proteins is expected to generate new methods and reagents (clones, proteins, and purification procedures) that will benefit the general biological community as well as structural genomics projects.

In the following sections I will describe the main techniques related to the production of recombinant proteins.

1. Cloning strategy

To speed up protein production it's generally better to adopt a strategy of parallel cloning and expression. The gene of interest is cloned in parallel into a variety of expression vectors containing different tags and/or fusion partners, and into vectors for a variety of expression systems. This approach should not only gain us a lot of time but also result in a larger number of successfully expressed proteins.

The cloning strategy consists of the following steps:

- the gene of interest is amplified by PCR.
- the PCR product is cloned into a specific cloning or expression vector using one of the cloning methods described below.
- the sequence of a positive clone is checked by sequence analysis.

- the gene of interest is subcloned into a variety of expression vectors (for different expression systems) using a fixed set of restriction enzymes or specific recombination sites.

Cloning methods which are commonly used are;

1. Cloning using restriction enzymes
2. TA cloning and TOPO TA cloning
3. Recombination cloning systems
    - GATEWAY Cloning Technology (Invitrogen)
    - Creator (BD Clontech)
4. Dicistronic cloning

1.1. Cloning using restriction enzymes

Restriction enzymes (restriction endonucleases) are proteins that cut DNA at (or close to) specific recognition sites. Two types of restriction enzymes exist that differ in the way they cut the target DNA:

- **Blunt end cutters.** These enzymes cut both strand of the target DNA at the same spot creating blunt ends.
- **Sticky end cutters.** These enzymes cut both strand of the target DNA at different spots creating 3'- or 5'-overhangs of 1 to 4 nucleotides (so-called sticky ends).

To be able to clone a DNA insert into a cloning or expression vector, both have to be treated with two restriction enzymes that create compatible ends. At least one of the enzymes used should be a sticky end cutter to ensure that the insert is incorporated in the right orientation. It will save a lot of time when we could carry out the two digestions simultaneously (double digestion). Not all restriction enzymes work equally well in all commercially available buffers and, therefore, it is worthwhile to check (*e.g.* in the

reference appendix of the New England Biolabs catalog) which enzymes are compatible and in which buffer. To ensure efficient digestion the two recognition sites should be more than 10 base pairs apart. If one of the enzymes is a poor cutter or if the sites are separated 10 base pairs or less, the digestions should be performed sequentially. The first digest should be done with the enzyme that is the poorest cutter and the second enzyme added after digestion has been verified by running a sample of the reaction mix on an agarose gel.

1.1.1 Methylation of DNA

Methylation of DNA by the host strain can have a great effect on DNA cleavage and/or transformation. Most laboratory strains of *E. coli* contain 3 DNA methylases that methylate distinct nucleotides in specific DNA sequences:

- **Dam methylase :** methylates adenine residues in the sequence GATC
- **Dcm methylase :** methylates the internal cytocine residues in the sequence CCAGG and CCTGG
- **EcoK1 methylase:** methylates adenine residues in the sequence AAC(N_6)GTGC and GCAC (N_6) GTT

When the DNA is isolated from strains expressing one of these methylases, especially **Dam** or **Dcm** methylase because their sites are much more common than **EcoK1** sites, some or all sites for a restriction enzyme may be resistant to cleavage. This occurs because DNA is protected from cleavage when a particular base in the recognition site of the restriction enzyme is methylated. For example, plasmid DNA isolated from Dam+ *E. coli* is completely resistant to cleavage by Mbo I, which cleaves **GATC** sites. Not all restriction enzymes are sensitive to methylation. For example, BamHI still cleaves

GGATCC sites of plasmid DNA from Dam+ *E. coli*. Further, methylation of plasmid DNA can affect the frequency of transformation in special situations. For example, the transformation efficiency is reduced when Dam-modified plasmid DNA is introduced into Dam- strains or when Dam- or Dcm-modified DNA is introduced into certain other bacterial species. In some cases it will be necessary to use a Dam- and/or Dcm-negative *E. coli* strain, for instance **GM48** and **JM110**.

1.1.2 Vector preparation

- Digestion of vector DNA using (preferably) two restriction enzymes.
- Dephosphorylation of the ends using calf intestine or shrimp alkaline phosphatase. This reduces the background of non-recombinants due to self-ligation of the vector (especially when a single site was used for cloning).
- Purification of the digested vector by agarose electrophoresis to remove residual nicked and supercoiled vector DNA and the small piece of DNA that was cut out by the digestions. This usually reduces strongly the background of non-recombinants due to the very efficient transformation of undigested vector.

1.1.3 Insert preparation

- Digestion of insert DNA using (preferably) two restriction enzymes.
- **Purification of the digested insert:** Purification should be carried out by agarose gel electrophoresis when the insert is subcloned into a vector from a vector with the same selective marker or PCR amplified from a vector with the same selective marker. Otherwise, it can be purified using a commercial kit (such as Qiagen's PCR purification kit).

1.1.4 Ligation

The next step is the ligation of the insert into the linearized vector. This involves the formation of phosphodiester bonds between adjacent 5'-phosphate and 3'-hydroxyl residues, which can be catalyzed by two different ligases: *E. coli* DNA ligase and bacteriophage T4 DNA ligase. The latter is the preferred enzyme because it can also join blunt-ended DNA fragments.

The efficiency of the ligation reaction depends on:

- The **absolute DNA concentration**. The concentration should be high enough to ensure that intermolecular ligation is favored over self-ligation but not so high as to cause extensive formation of oligomeric molecules.
  For pET vectors, good results are obtained at a vector DNA concentration or approx. **1 nM** (*i.e.* 50-100 ng vector DNA per 20-ml ligation mix).
- The **ratio between vector and insert DNA**. The maximum yield of the right recombinants is usually obtained using a molar ratio of insert to vector DNA of approx. 2. If the concentration of insert DNA is substantially lower than that of the vector, the ligation efficiency becomes very low.
  In practice, we set-up ligation reactions with a molar ratio of insert to vector DNA from **2:1** to 6**:1**.
- The **cloning strategy**. Higher yields of the right recombinant are obtained when the vector and insert have been prepared using two restriction enzymes and the digested vector has been gel-purified before the ligation reaction (as shown in the figure on the right).

The ligation of blunt-ended fragments is less effective than that of sticky-ended ones. Blunt end ligation may be enhanced by:

- high concentrations of blunt-ended DNA fragments.
- a high concentration of ligase (10,000 NEB units/ml).
- a low concentration of ATP (0.1 mM).
- the addition of PEG 4000 [5% (w/v)]. (Ref)

1.2. TA and TOPO TA cloning

The **TA cloning** method takes advantage of the terminal transferase activity of some DNA polymerases such as *Taq* polymerase. This enzyme adds a single, 3'-A overhang to each end of the PCR product. This makes it possible to clone this PCR product directly into a linearized cloning vector with single, 3'-T overhangs. DNA polymerases with proofreading activity, such as *Pfu* polymerase, can not be used because they provide blunt-ended PCR products. TA cloning kits are available from different manufacturers. The TOPO TA cloning method combines the advantages of TA cloning with the ligation activity of topoisomerase I. This allows direct ligation of PCR products in just 5 minutes. TOPO TA cloning kits are available from *Invitrogen* (Fig. 2).

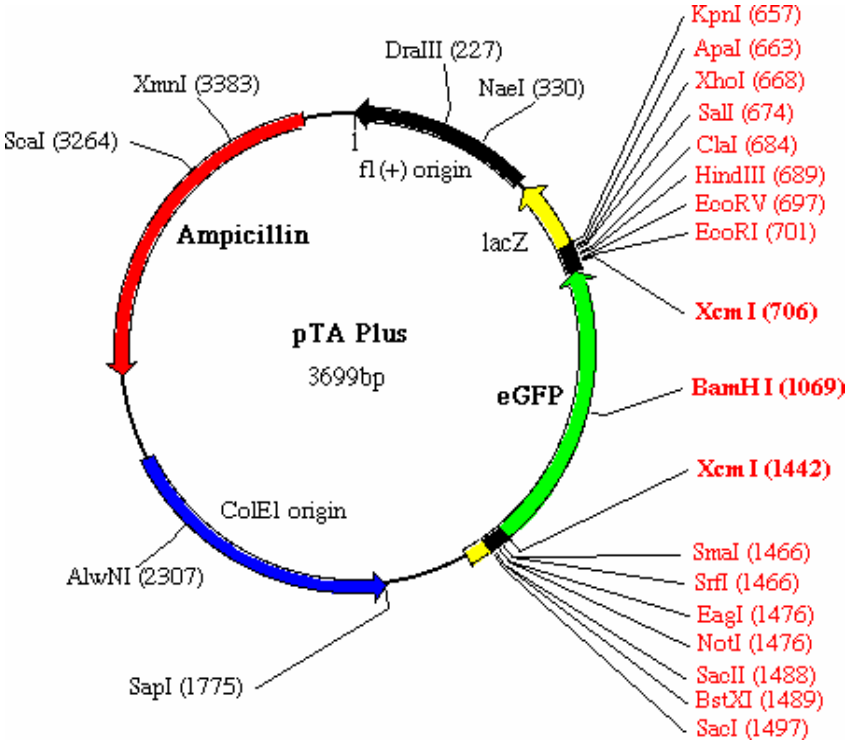


Figure 2. Map of pTA Plus

1.3. Recombination cloning systems

1.3.1 GATEWAY Cloning technology

The GATEWAY Cloning Technology is based on the site-specific recombination system used by phage l to integrate its DNA in the *E. coli* chromosome. Both organisms have specific recombination sites called *att*P in phage l site and *att*B in *E. coli*. The integration process (lysogeny) is catalyzed by 2 enzymes: the phage l encoded protein Int (Integrase) and the *E. coli* protein IHF (Integration Host Factor). Upon integration, the recombination between *att*B (25 nt) and *att*P (243 nt) sites generate *att*L (100 nt) and *att*R (168 nt) sites that flank the integrated phage l DNA (Fig.3).

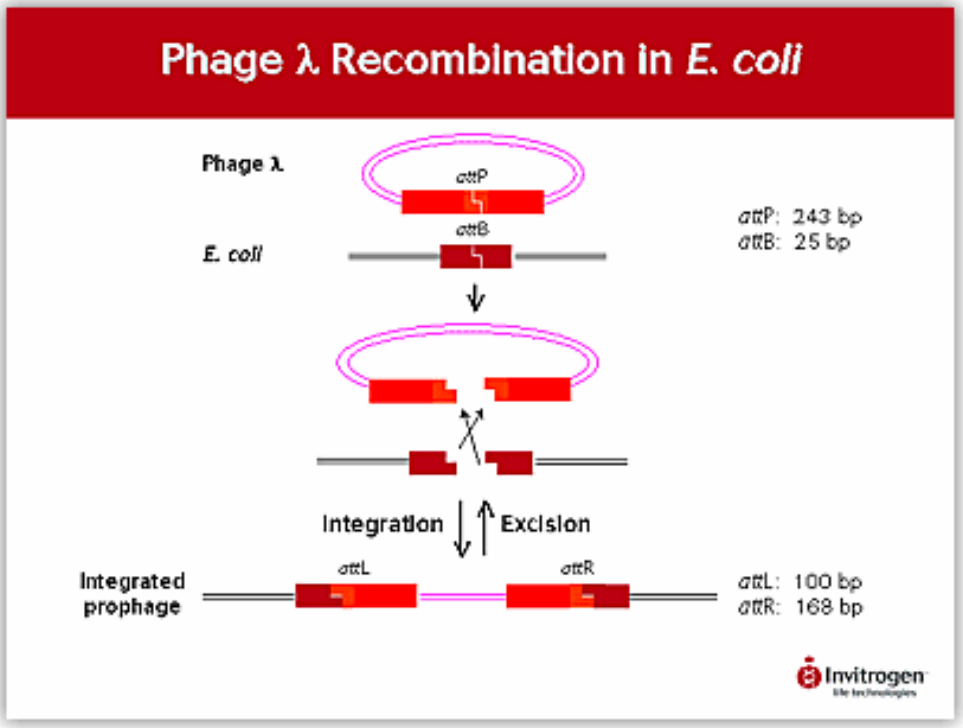


Figure 3. Phage l recombination into the *E. coli* chromosome

The process is reversible and the excision is again catalyzed Int and IHF in combination with the phage l protein Xis. The *att*L and *att*R sites surrounding the inserted phage DNA recombine site-specifically during the excision event to reform the *att*P site in phage l and the *att*B site in the *E. coli* chromosome. The GATEWAY reactions are *in vitro* versions of the integration and excision reactions. To make the reactions directional two slightly different and specific site were developed, *att*1 and *att*2 for each recombination site. These sites react very specifically with each other. For instance in the BP Reaction *att*B1 only reacts with *att*P1 resulting in *att*L1 and *att*R1, and *att*B2 only with *att*P2 giving *att*L2 and *att*R2. The reverse reaction (LR Reaction) shows the same specificity. The ultimate goal of the **GATEWAY reactions** is to make an expression clone. This is often a two step process:

Step 1 Cloning the gene of interest into an *Entry Vector* via the **BP Reaction**.

Step 2 Subcloning the gene of interest from the Entry Clone (Step 1) into a *Destination Vector* via the **LR Reaction** creating the *Expression Clone*.

In the LR Reaction of Step 2 (Fig.4) the gene of interest is cloned into an *Entry Vector* and flanked by the *att*L1 and *att*L2 recombination sites. The *Entry Vector* is transcriptionally silent and contains the gene for kanamycin resistance (Km^r). To produce the *Expression Clone* the gene has subcloned into a *Destination Vector* that contains all the sequence information necessary for expression, the gene for ampicillin resistance (Ap^r), and two recombination sites *(att*R1 and *att*R2) that flank a gene for negative selection, *ccd*B (the encoded protein is toxic for the standard *E. coli* strains). The two plasmids are mixed and the LR CLONASE Enzyme Mix is added. The reaction is directional and specific, so that *att*L1 only reacts with *att*R1 and *att*L2 with *att*R2. The recombination yields two constructs: the intended *Expression Clone* and a by-product (labelled in Figure 4 as *Donor Vector*). The produced expression clone is under two forms of selection: the antibiotic resistance and the negative selection by the toxic ccdB protein. As a result high levels of positive clones (typicaly more than 90 - 99%) are obtained after transformation to a standard cloning or expression strain like DH5a or BL21 (DE3). One of the main advantages of the GATEWAY Cloning Technology is that once we made an *Entry Clone* the gene of interest can be easily subcloned into a wide variety of *Destination Vectors* using the LR Reaction (Figure 5).

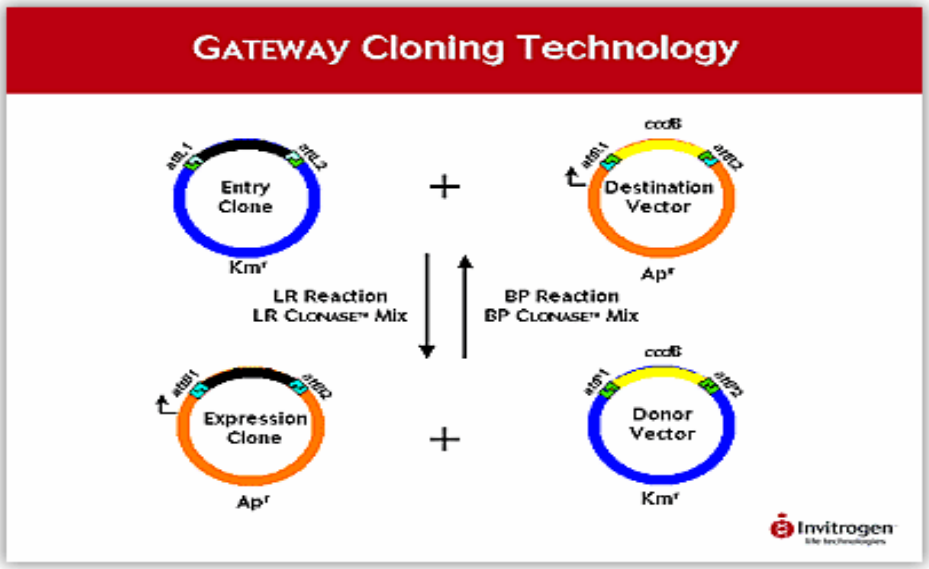


Figure 4. The GATEWAY reactions

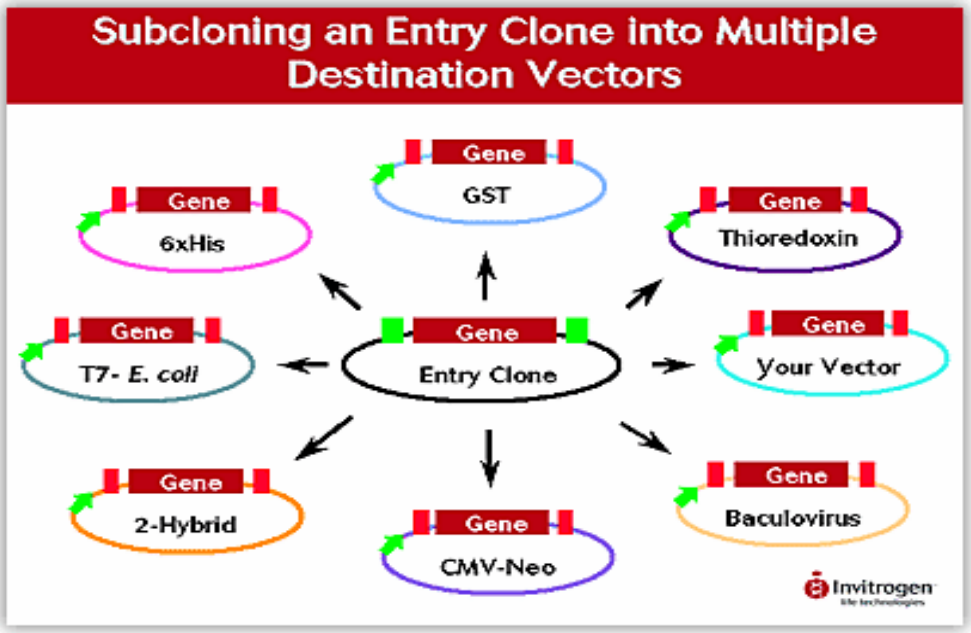


Figure 5. Subcloning an *Entry Clone* into multiple *Destination Vectors*

Usually the Destination Vectors do not contain a protease cleavage site. In order to have a protease cleavage site in our fusion protein to enable the removal of the fusion partner and/or tag, protease sequence was be incorporated into the appropriate primer and it has done in-frame with the gene of interest.

1.3.2 Creator (BD Clontech)

Creator is more efficient than conventional restriction enzyme-based cloning methods because it utilizes only one enzyme, **Cre Recombinase**, to shuttle a gene from a donor vector into an expression vector.

Cre Recombinase avoids problems associated with conventional cloning, such as:

- restriction enzyme site compatibility
- presence of cloning sites within the gene of interest
- the need to retain reading frame in translational fusions
- insert directionality
- restriction digestion and ligation efficiency
- gel purification
- the cost of maintaining a large restriction enzyme collection

Cre recombinase facilitates site-specific recombination at *loxP* sites. It recognizes and binds to inverted repeats that flank the spacer region where recombination occurs. The enzyme then uses a reactive tyrosine within its active site to cleave the DNA in the spacer region, creating a staggered cut with sticky ends. Cre then reattaches the 5' end of one *loxP* site to the 3' end of the other *loxP* at the site of the staggered cut, thus recombining the DNA from two different vectors. Multiple reactions between the *loxP* site in pDNR

and the two loxP sites in the acceptor vector occur simultaneously to transfer your gene and the chloramphenicol resistance gene into the acceptor vector.

1.4. Di- or multi-cistronic cloning

It is possible to co-express two or more proteins from the same vector.

There are two possible strategies:

- The proteins are expressed under the regulation of **different promoters**. In this case, the genes are cloned sequentially into a vector containing different multiple cloning sites, *e.g.* pFastBac DUAL (Invitrogen), a baculovirus expression vector.
- The proteins are cloned under the regulation of one promoter. In this case, the genes are cloned into a normal expression vector containing one promoter and one multiple cloning site (**di- or multi-cistronic cloning**). Each gene should have its own Shine-Dalgarno sequence and start codon.

For the dicistronic cloning pET vector are also available. The order in which the genes are cloned into the vector is important for the expression levels of the different proteins.

2. Transformation

Transformation is the process of getting the recombinant vector from a reaction mixture or vector solution into *E. coli* cells. To enable the cells to take up circular vector DNA they have to be made **competent**. The method for the preparation of competent cells depends on the transformation method used and transformation efficiency required.

The choice of the *E. coli* host strain depends on the goal of the transformation:

- The transformation of a ligation mix was done in a recA$^-$ cloning strain, such as DH5a, NovaBlue or XL1-Blue. Depending on the background of non-recombinants (from a ligation mix containing only digested vector) a number of transformants (3-12 colonies) should be picked and checked for the presence of the right insert by restriction analysis or colony PCR.
- The transformation of a vector for multiplication was also be done in a recA$^-$ strain, such as DH5a, NovaBlue or XL1-Blue.
- The transformation of a vector for protein expression was done in the appropriate expression host.

  pET vectors in BL21 (DE3)

  pBAD vectors in Top10, LMG194

3. Choice of expression system

Once we decided which protein or which domain(s) of a protein we would like to clone and express, we have to think also about which expression system we would like to use. At present there are many different expression systems available, but we limit ourselves to the following three systems since they are most suited for large-scale production of proteins:

3.1. Escherichia coli

The expression of proteins in ***E. coli*** is the easiest, quickest and cheapest method. There are many commercial and non-commercial expression vectors available with different N- and C-terminal tags and many different strains which are optimized for special applications. One of the major advantages of ***E.coli*** is production of isotope labeled protein for NMR is very easy in compare to other systems.

3.2. Yeast

Yeast is a eukaryotic organism and has some advantages and disadvantages over *E. coli.* One of the major advantages is that yeast cultures can be grown to very high densities, which makes them especially useful for the production of isotope labeled protein for NMR. The two most used yeast strains are *Saccharomyces cerevisiae* and the methylotrophic yeast *Pichia pastoris.*

3.3. Baculovirus infected insect cells

Insect cells are a higher eukaryotic system than yeast and are able to carry out more complex post-translational modifications than the other two systems. They also have the best machinery for the folding of mammalian proteins and, therefore, give you the best chance of obtaining soluble protein when you want to express a protein of mammalian origin. The disadvantages of insect cells are the higher costs and the longer duration before you get protein (usually 2 weeks).

4. Protein Purification

There are many purification methods available for the separation of macromolecules. To effectively resolve a crude mixture of substances, it may be necessary to use a combination of techniques. Some examples of liquid chromatographic techniques are;

4.1. Ion-Exchange Chromatography

Proteins are made up of twenty common amino acids. Some of these amino acids possess side groups ("R" groups) which are either positively or negatively charged. A comparison of the overall number of positive and negative charges will give a clue as to the nature of the protein. If the protein has more positive charges than negative charges, it is said to be a basic protein. If the negative charges are greater than the positive charges, the protein is acidic. When the protein contains a predominance of ionic charges, it can be bound to a support that carries the opposite charge. A basic protein, which is positively charged, will bind to a support which is negatively charged. An acidic protein, which is negatively charged, will bind to a positive support. The use of ion-exchange chromatography, then, allows molecules to be separated based upon their charge. Families of molecules (acidics, basics and neutrals) can be easily separated by this technique. This is perhaps the most frequently used chromatographic technique used for protein purification.

4.2. Hydrophobic Interaction Chromatography ("HIC")

Not all of the common amino acids found in proteins are charged molecules. There are some amino acids that contain hydrocarbon side-chains which are not charged and therefore cannot be purified by the same principles involved in ion-exchange chromatography. These hydrophobic ("water-hating") amino acids are usually buried away in the inside of the protein as it folds into its biologically active conformation. However, there is usually some distribution of these hydrophobic residues on the surface of the molecule. Since most of the hydrophobic groups are not on the surface, the use of HIC allows a much greater selectivity than is observed for ion-exchange chromatography. These hydrophobic amino acids can bind on a support which contains immobilized hydrophobic groups. It should be noted that these HIC supports work by a "clustering" effect; no covalent or ionic bonds are formed or shared when these molecules associate.

4.3. Gel-Filtration Chromatography

This technique separates proteins based on size and shape. The support for gel-filtration chromatography are beads which contain holes, called "pores," of given sizes. Larger molecules, which can't penetrate the pores, move around the beads and migrate through the spaces which separate the beads faster than the smaller molecules, which may penetrate the pores. This is the only chromatographic technique which does not involve binding of the protein to a support.

4.4. Affinity Chromatography

This is the most powerful technique available to the chromatographer. It is the only technique which can potentially allow a one-step purification of the target molecule. In order to work, a specific ligand (a molecule which recognizes the target protein) must be immobilized on a support in such a way that allows it to bind to the target molecule. A classic example of this would be the use of an immobilized protein to capture it's receptor (the reverse would also work). This technique has the potential to be used for the purification of any protein, provided that a specific ligand is available. Ligand availability and the cost of the specialized media are usually prohibitive at large-scale.

4.5. Other Methods

While the methods above are typically chosen for use in a purification process, there are in fact many others that can be used. Each of these methods or techniques takes advantage of a specific part of the protein being purified. The commonality is that all of the techniques employed are based on the protein's structure.

5. Isotope labeling of proteins for structural determination

For the isotope enrichment the protein is usually overexpressed in a bacterial system. Fully isotope labeled proteins can be produced by growing the bacteria on minimal medium containing ^{15}NH4Cl and ^{13}C-glucose as sole nitrogen and carbon sources.

6. Protein Characterization

Proteins are composed of complex polypeptide chains with unique 3-dimensional structures. These structures are stabilized by a combination of electrostatic and hydrophobic interactions, combined with a large degree of flexibility inside the structure of the molecule.

Popular biophysical protein characterization techniques are:

- Circular dichroism (CD).
- Nuclear magnetic resonance (NMR).

6.1. Circular Dichroism (CD)

Circular Dichroism (CD) is an excellent method for analyzing protein and nucleic acid secondary structure in solution. It can be used to follow the changes in folding as a function of temperature or denaturant, and is also useful for measuring protein-ligand and nucleic acid-ligand interactions.

Circular dichroism spectroscopy is particularly good for:

- determining whether a protein is folded, and if so characterizing its secondary structure and the structural family to which it belongs.
- comparing the structures of a protein obtained from different sources (*e.g.* species or expression systems) or comparing structures for different mutants of the same protein.

- studying the conformational stability of a protein under stress –thermal stability, pH stability, and stability to denaturants -- and how this stability is altered by buffer composition or addition of stabilizers.
- determining whether protein-protein interactions alter the conformation of protein.

6.2 Nuclear magnetic resonance (NMR)

In many cases it will be necessary to know the three-dimensional (3D) structure of a protein to understand its function, and two techniques play a dominant part: x-ray crystallography and nuclear magnetic resonance (NMR). These are the two main techniques that can provide structures of macromolecules at atomic resolution. Several features of solution-state NMR make it particularly suitable for structure-function analysis. Structural analysis by NMR does not require protein crystals. Most ($\approx$75%) of the NMR structures in the Protein Data Bank (PDB) do not have corresponding crystal structures, and many of these simply do not provide diffraction quality crystals. Moreover, NMR studies can be carried out in aqueous solution under conditions quite similar to the physiological conditions under which the protein normally functions. While most crystal structures are determined under physiologically relevant conditions, in many cases somewhat exotic solution conditions are required for crystallization. NMR measurements not only provide structural data but reach much further and can supply information on dynamics, conformational equilibrium, folding and intra- as well as intermolecular interactions (4-8). The power of NMR over other spectroscopic techniques

results from the fact that every NMR active nucleus gives rise to an individual signal (resonance line) in the spectrum that can be resolved by multi-dimensional NMR techniques. This becomes more difficult for larger molecular structures and puts a practical limit to the molecular size that can be studied in detail by NMR (9-11). The principles of a NMR structure determination can be summarized as follows: preparation of the protein solution, the NMR measurements, the assignment of NMR signals to individual atoms in the molecule, identification of conformational constraints (e.g. distances between hydrogen atoms), the calculation of the 3D structure on the basis of the experimental constraints. The first step to solve the three dimensional structure of biological macromolecules in solution by NMR is the preparation of the protein solution. An efficient structure determination by NMR requires a highly purified protein preparation. The macromolecule under study should be stable in the chosen conditions for many weeks. The pH, ionic strength, and temperature can often be adjusted to mimic physiological conditions. Any buffers, cosolvents and additives (e.g. detergent molecules) used should be preferentially hydrogen-free or deuterated. For the NMR measurements the protein is dissolved in 0.25 to 0.5 ml of aqueous buffer that contains about 5 % of D2O which is necessary for the stabilization of the NMR instrument during the measurement. The inherent low sensitivity of the technique requires protein concentrations of about 1 mM. Proteins with a molecular weight larger than 10 kDa must be isotope enriched in ^{15}N and ^{13}C for an efficient structure determination. The natural abundance of ^{15}N and $_{13}C$ is only 0.37% and 1.1%, respectively, whereas levels close to 100% are required for efficient NMR experiments with macromolecules. The sensitivity obtainable with these types of nuclei greatly varies even if the sample is fully isotope

labelled with 13C or 15N. The proton offers the best sensitivity and hence constitutes the preferred nucleus for detection of the NMR signal. The other nuclei are usually measured during evolution periods of multidimensional NMR experiments and their information is transferred to protons for detection. For unlabelled proteins smaller than 10 kDa the combination of the two 2D spectra, [^{1}H, ^{1}H]- COSY and [^{1}H, ^{1}H]-NOESY often allows the assignment of most proton NMR signals (12).The first experiment, the [^{1}H, ^{1}H]-COSY, detects through-bond interactions between protons and correlates protons, which are separated by up to three chemical bonds. With this experiment the protons within an amino acid can be correlated, however, neighboring amino acids in the polypeptide sequence cannot be connected. The set of correlated proton nuclei is referred to as spin system. However, as soon as an amino acid occurs more than once in a polypeptide chain a direct assignment to a specific sequence position is not possible through COSY experiment. For this purpose the second experiment, the [^{1}H, ^{1}H]-NOESY (14), is measured, where NOESY stands for NOE spectroscopy (13, 14). NOEs connect pairs of hydrogen atoms separated by less than 0.5 nm. In contrast to COSY-type experiments the nuclei involved in the NOE correlation can belong to amino acid residues that may be far apart along the protein sequence but close in space. For larger proteins extensive signal overlap prevents complete assignments of all 1H signals in proton spectra. This barrier can be overcome with 2D and 3D NMR techniques and uniformly ^{13}C and ^{15}N labelled proteins. With these methods, systems with molecular weights up to approximately 35 kDa can be studied. In [^{13}C, ^{15}N]-labelled proteins a sequential assignment strategy can be used which is based on through-bond correlations across the peptide-bond between sequential amino acids. After the assignment of all, or nearly all, resonances of a protein

has been performed, experiments for the extraction of structural parameters have to be analyzed. The most important parameter for NMR-based structure determination are H-H distances, which are derived from NOE intensities, and dihedral angles, which are obtained from 3J coupling constants. Programs available for the calculation of three-dimensional structures utilize, together with experimental constraints, information about the covalent structure of the protein - the amino acid sequence, bond lengths, bond angles, chiralities, and planar groups - as well as by steric repulsion between non-bonded atom pairs. It must be kept in mind that the experimental constraints do not uniquely describe one exact 3D structure because NMR-derived constraints typically describe a range of possible values and many distances cannot be determined. The structure calculation is thus repeated many times to determine an ensemble of structures consistent with the input data set. For this reason the NMR structures are usually represented by a bundle of structures. The quality of structures is assessed by the calculation of the root mean square deviation (RMSD) between the atoms of individual conformers in the bundle. At present not many structures of proteins above 30 kDa molecular weight have been solved by NMR. One can anticipate, however, that in the not too distant future many more NMR structures of larger proteins and protein complexes will become available by the widespread use of novel NMR experiments and creative isotope labeling schemes. In addition, development of higher magnetic fields and improved spectrometer hardware will result in gains in resolution and sensitivity and will further increase the upper molecular weight limit for structural studies by NMR.

7. Advent of recombinant DNA technology

Numerous methods have been developed or exploited to mutate DNA to make a large number of variants of proteins. Initially all approaches focused on the generation of random mutations in chromosomal DNA such as those induced by X-rays (29) and chemicals (15). While these methods of random mutagenesis provided a valuable tool for classical genetic studies, they were limited by their inability to target the mutation to a specific gene or genetic element. Techniques for randomly mutagenizing a genome required screening or selection from massive numbers of mutants to obtain the desired mutation (16, 17). The ability to manipulate DNA *in vitro*, through the use of plasmid vectors, became a driving force for newer technologies, which allowed precise changes in discrete, manageable segments of the genome with relatively little effort.

Recombinant DNA technology removes the limitations imposed by biological systems and allows a variety of *in vitro* techniques to be used to create these alterations and, in many cases, alterations beyond those that could be made by a biological system. Strictly speaking, the common procedure of sub-cloning could be considered a mutagenesis technique. We will limit our discussion to those techniques designed to alter DNA for the sake of studying the effect of that alteration on a regulatory element or gene product. Site-directed mutagenesis methods first benefited from recombinant DNA technology in the 1970s when isolated genes were exposed to conditions such as nucleotide analog incorporation or chemical agents to localize their mutagenic effects. During this time, the use of plasmid vectors for DNA replication greatly enhanced the study of mutations. Mutagenesis targeted to a defined region of DNA includes many techniques, some more popular than others. *In vitro* approaches to site-directed

mutagenesis can be grouped generally into three categories (18):

i) methods that restructure fragments of DNA, such as cassette mutagenesis (19);

ii) localized random mutagenesis; and

iii) oligonucleotide-directed mutagenesis.

Of these methods, oligonucleotide-directed mutagenesis is by far the most commonly used method.

7.1 Oligonucleotide-directed mutagenesis

All oligonucleotide-directed mutagenesis is based on the same concept an oligonucleotide encoding the desired mutation(s) is annealed to one strand of the DNA of interest and serves as a primer for initiation of DNA synthesis. In this manner, the mutagenic oligonucleotide is incorporated into the newly synthesized strand. Mutagenic oligonucleotides incorporate at least one base change but can be designed to generate multiple substitutions, insertions or deletions. In vitro site-directed mutagenesis is an invaluable technique for studying protein structure-function relationships and gene expression, and for carrying out vector modification. Several approaches to this technique have been published, but these methods generally require single-stranded DNA (ssDNA) as the template (20-23) and are labor intensive or technically difficult. Stratagene's Quick-Change site-directed mutagenesis kit allows site specific mutation in virtually any double-stranded plasmid, thus eliminating the need for sub cloning into M13-based bacteriophage vectors and for ssDNA rescue (24). In addition, the Quick-Change site-directed mutagenesis system requires no specialized vectors, unique restriction sites, or multiple transformations. This rapid four-step procedure generates mutants with greater

than 80% efficiency. The protocol is simple and uses either miniprep plasmid DNA or cesium-chloride-purified DNA. For long (~8 kb) or difficult targets, Stratagene offers the Quick-Change XL site directed mutagenesis kit (Catalog #200516).The Quick-Change site-directed mutagenesis kit is used to make point mutations, switch amino acids, and delete or insert single or multiple amino acids. The Quick-Change site-directed mutagenesis method is performed using *PfuTurbo* DNA polymerase and a temperature cycler. *PfuTurbo* DNA polymerase replicates both plasmid strands with high fidelity and without displacing the mutant oligonucleotide primers. The basic procedure utilizes a super coiled double-stranded DNA (dsDNA) vector with an insert of interest and two synthetic oligonucleotide primers containing the desired mutation (Fig.6). The oligonucleotide primers, each complementary to opposite strands of the vector, are extended during temperature cycling by *PfuTurbo* DNA polymerase. Incorporation of the oligonucleotide primers generates a mutated plasmid containing staggered nicks. Following temperature cycling, the product is treated with *Dpn* I. The *Dpn* I endonuclease (target sequence: 5´-Gm6ATC-3´) is specific for methylated and hemimethylated DNA and is used to digest the parental DNA template and to select for mutation-containing synthesized DNA (25). DNA isolated from almost all *E. coli* strains is dam methylated and therefore susceptible to *Dpn* I digestion. The nicked vector DNA containing the desired mutations is then transformed into XL1-Blue supercompetent cells. The small amount of starting DNA template required to perform this method, the high fidelity of the *PfuTurbo* DNA polymerase, and the low number of thermal cycles all contributes to the high mutation efficiency and decreased potential for generating random mutations during the reaction.

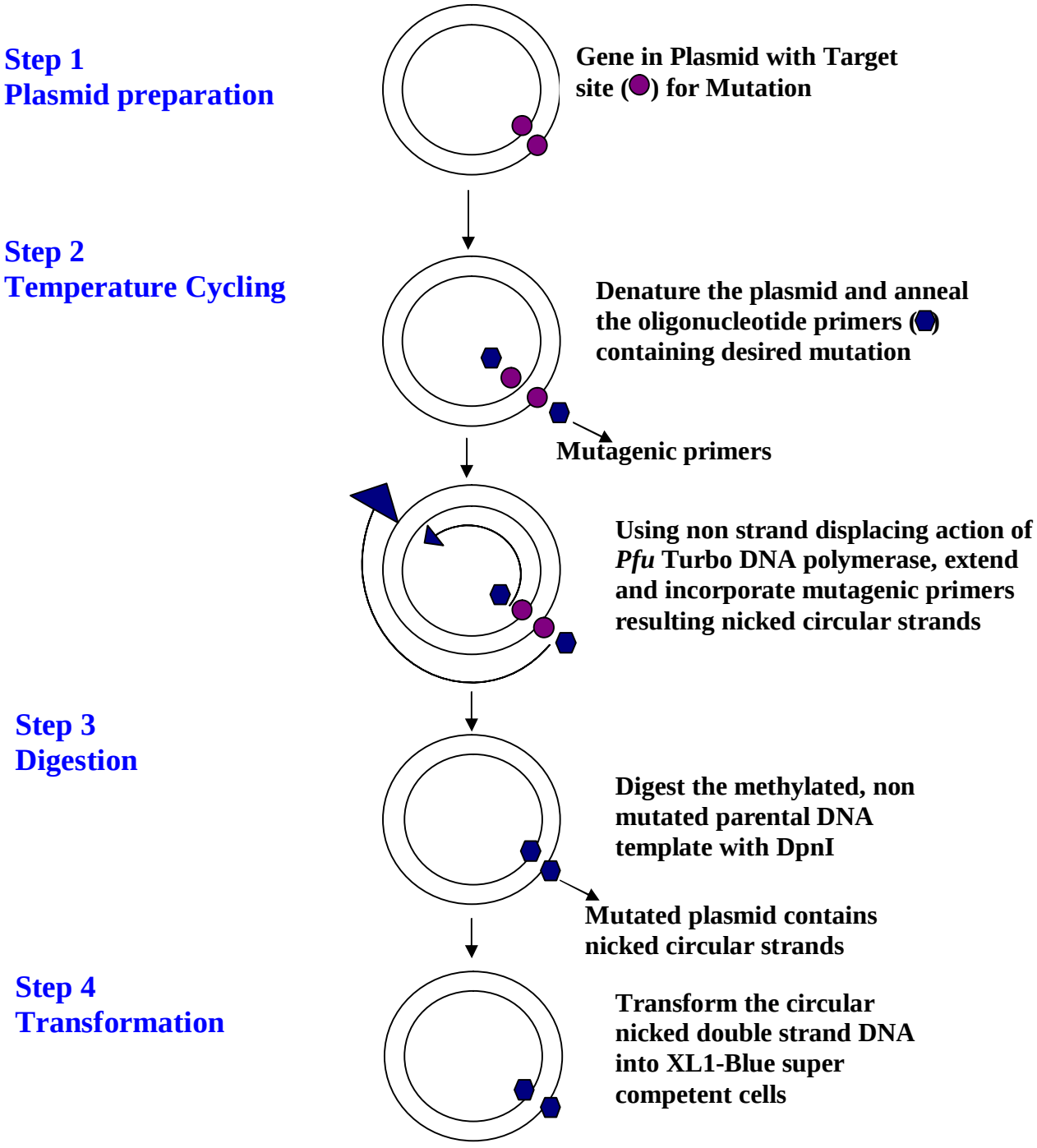


Figure 6.Overview of Quick change site directed mutagenesis method

8. References

1. Kim, S.H. 1998. *Nature Struct. Biol.* **5**, 643−645.

2. Christendat, D. *et al*.2000. *Nature Struct. Biol.* **7**, 903−909.

3. Christendat, D. *et al*. 2000 *Progr. Biophys. and Mol. Biol.* in the press.

4. Dyson, H.J. and Wright, P.E. (1996) *Annu.Rev.Phys.Chem. 47*,369-395.

5. Farrow, N.A., Zhang, O., Forman-Kay, J.D., and Kay, L.E. (1997) *Biochemistry 36*, 2390-2402.

6. Kay, L.E. (1998), *Nature Struct.Biol.* 513-517.

7. Palmer, A.G., III (1997), *Curr Opin Struct Biol 7*, 732-737.

8. Wuthrich, K., Billeter, M., Guntert, P., Luginbuhl, P., Riek, R., and Wider, G. (1996), *Faraday Discuss.* 245-253.

9. Clore, G.M. and Gronenborn, A.M. (1998), *Trends Biotechnol. 16*, 22-34.

10 Kay, L.E. and Gardner, K.H. (1997) *Curr.Opin.Struct.Biol.7*, 722-731.

11. Wuthrich, K. (1998), *Nat.Struct.Biol. 5 Suppl*, 492-495.

12. Wider, G., Macura, S., Kumar, A., Ernst, R.R., and Wüthrich, K. (1984), *J.Magn.Reson. 56*, 207-234.

13. Jeener, J., Meier, B.H., Bachmann, P., and Ernst, R.R. (1979), *J.Chem.Phys. 71*, 4546-4553.

14. Kumar, A., Ernst, R.R.,Wuthrich, K. (1980), *Biochem.Biophys.Res.Commun. 95*,1-6.

15. Muller, H.J. (1927) *Science* **66**, 84.

16. Auerbach, C. and Robson, J.M. (1947) *Proc. R. Soc. Edinburgh B* **62**, 279.

17. Hong, J.S. and Ames, B.N. (1971) *Proc. Natl. Acad. Sci. USA* **68**, 3158.

18. Botstein, D. and Shortle, D. (1985) *Science* **229**, 4719.

19. Lo, K.-M. et al. (1984) *Proc. Natl. Acad. Sci. USA* **, 81**, 2285.

20. Kunkel, T. A. (1985) *Proc Natl Acad Sci U S A* 82(2):488–92.

21. Vandeyar, M. A., Weiner, M. P., Hutton, C. J. and Batt, C. A. (1988) *Gene* 65(1):129–33.

22.Sugimoto, M., Esaki, N., Tanaka, H. and Soda, K. (1989) *Anal Biochem*179(2) :309–11.

23. Taylor, J. W., Ott, J. and Eckstein, F. (1985) *Nucleic Acids Res* 13(24):8765–85.

24. Papworth, C., Bauer, J. C., Braman, J. and Wright, D. A. (1996) *Strategies* 9(3):3–4.

25. Nelson, M. and McClelland, M. (1992) *Methods Enzymol* 216:279–303.

In the following sections I will describe in detail the methodologies used for the proteins I have studied;

A .*Bacillus pasteruii* Cytochrome c Mutants (Bpcytc)

B. Copper Transporting ATPases PacSN of Cyanobacteria *Synechocystis* PCC 6803

C. Zinc Transporting ATPases ZiaA of Cyanobacteria *Synechocystis* PCC 6803

D. Cellular Nucleic Acid Binding/ Zinc Finger 9 (CNBP/ZNF9)

E. Catalytic Domain of A Disintegrin and Metalloprotease 10 (ADAM10)

A. *Bacillus pasteruii* Cytochrome c Mutants (Bpcytc)

A.1. Experimental Methods

The plasmids used to produce protein samples were generated from the pAT1 plasmid encoding WT Bpcytc. To generate the mutants Q68K, P72A, P72G and I75A (Fig.10) the Bpcytc gene was mutated following the procedure of the Quick-Change site-directed mutagenesis kit.

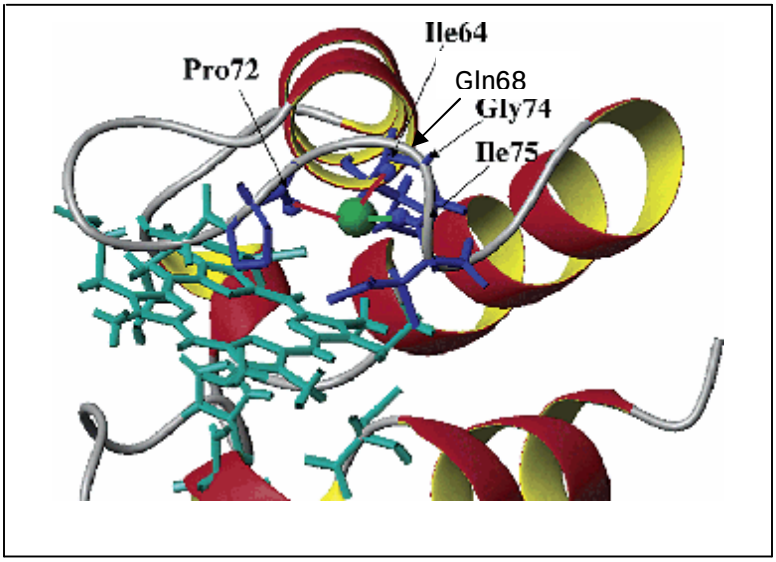


Figure 10. Residues selection from Met loop of B.pasteruii cytochrome c

A.1.1. Primer Design Guidelines

The mutagenic oligonucleotide primers used in this protocol have designed individually according to the desired mutation. The following considerations have taken care while designing the mutagenic primers:

♦ both of the mutagenic primers must contain the desired mutation and anneal to the same sequence on opposite strands of the plasmid.

♦ primers should be between 25 and 45 bases in length, with a melting temperature (*T*m) of ≥78°C. Primers longer than 45 bases may be used, but using longer primers increases the likelihood of secondary structure formation, which may affect the efficiency of the mutagenesis reaction. The following formula was used for estimating the *T*m of primers:

$$\mathbf{T_m = 81.5 + 0.41(\%GC) - 675/N - \%\ mismatch}$$

• *N* is the primer length in bases

• values for %GC and % mismatch are whole numbers

For calculating *T*m for the primers intended to introduce insertions or deletions following modified formula was used:

$$\mathbf{T_m = 81.5 + 0.41(\%GC) - 675/N}$$

where N does not include the bases which are being inserted or deleted.

♦ the desired mutation (deletion or insertion) should be in the middle of the primer with ~10–15 bases of correct sequence on both sides.

♦ the primers optimally should have a minimum GC content of 40% and should terminate in one or more C or G bases.

A.1.2 Mutant Strand Synthesis Reaction (Thermal Cycling)

For carrying the thermal cycle for mutagenesis plasmid DNA template was isolated from a dam^+ E. coli strain. The majority of the commonly used E. coli strains are dam^+. Plasmid DNA isolated from dam^- strains (e.g. JM110 and SCS110) is not suitable. To maximize temperature cycling performance thin-walled tubes were used, these ensure ideal contact with the temperature cycler's heat blocks.

The following protocol was optimized using thin walled tubes;

1. Two complimentary oligonucleotides containing the desired mutation were synthesized which are flanked by unmodified nucleotide sequence.

2. Control and Sample reaction (s) were prepared by following the instructions in stratgene's site directed mutagenesis kit.

3**.** Each cycle was done by using cycling parameters mentioned in Table 2.

Segment	Cycles	Temperature	Time
1	1	95 ^{0}C	30 seconds
2	12-18	95 ^{0}C	30 seconds
		55 ^{0}C	1 minute
		68 ^{0}C	1 min/kb plasmid length

4**.** Segment two of the cycling parameters were adjusted according to the type of mutation

Type of Mutation Desired	Number of Cycles
Point Mutation	12
Single amino acid change	16
Multiple amino acid deletions or additions	18

5. *Dpn* I Digestion of the Amplification Products

Dpn I restriction enzyme (10 U/µl) was added directly to each amplification and reaction mixture was spin down in microcentrifuge for 1 minute and immediately incubated each reaction at 37°C for 1 hour to digest the parental (i.e., the nonmutated) supercoiled dsDNA.

6. Transformation to XL1-Blue Supercompetent Cells

7. 5-6 colonies were picked from each plate and mutation was analyzed by sequencing results, positive clones were further used for protein expression.

A.1.3 Expression and Purification of *Bacillus pasteurii* cytochrome c mutants

For protein expression all the mutated clones were transformed in C41 competent cells which were co-transformed with pAT1 and pEC86, kindly provided by Dr. Thony Meyer. Plasmid pEC86 encodes the complete ccmA-H gene cluster from E. coli, which is needed for incorporation of the heme moiety in the expressed cytochrome, under the control of the Tet promoter, [29] and carries a marker for chloramphenicol. Cells harboring both plasmids were selected for their ability to grow on 2xYT plates containing ampicillin (100 μgmL^{-1}) and chloramphenicol (50 μgmL^{-1}). Rich media cultures were performed with the 2xYT culture medium supplemented with antibiotics. Cultures were usually incubated overnight with shaking at 37^0C, then induced with IPTG (IPTG; 0.5mM) and finally harvested after 24 h by centrifugation at 4^0C. For preparation of ^{15}N-enriched samples, cultures were grown on minimal medium consisting of M9 salts supplemented with MgSO4, trace metal, and vitamin solutions, and also containing δ-aminolevulinic acid (0.1 mM) and 2-mercaptoethanesulfonic acid (1 mM).The nitrogen

source was (^{15}NH4)2SO_4 (1.2 gL^{-1}) and the carbon source was glucose (4 gL^{-1}).The protein was extracted from E. coli cells by isolating the periplasmic contents, which were loaded directly onto a DE52 column equilibrated with 2 mM sodium ascorbate and 0.1M tris (hydroxymethyl) aminomethane (Tris)-HCl at pH 8.0 and eluted with a 0 ±300 mM NaCl gradient in the same buffer. Red fractions were concentrated, applied to a gel filtration column, and then eluted with 15 mM sodium phosphate buffer (pH 7.5). Samples for NMR spectroscopy were at a concentration of 1.0 ± 1.5 mM protein in the elution buffer of the gel filtration column. Immediately prior to NMR experiments on oxidized Bpcytc, O_2 was bubbled into the sample solution for 15-30 minutes. Reduction of Bpcytc was achieved by addition of an equimolar amount of sodium dithionite to the solution under anaerobic conditions.

A.1.4 Denaturation studies with GdmCl

Complete oxidation of the protein was achieved by washing samples with a five- to tenfold excess of $K_3Fe(CN)_6$. Protein samples in GdmCl were prepared by pipetting aliquots of stock GdmCl solutions into phosphate-buffered ([phosphate] =100 mM) protein solutions. Stock solutions of GdmCl were freshly prepared prior to each experiment. All unfolding measurements were carried out at pH 7.0. The pH of the final solutions was always checked after addition of the denaturant, and, when necessary, adjusted again to the desired value with concentrated NaOH or HCl. In all cases, the pH of the solutions was checked also after completion of the experiments. Samples for NMR spectroscopy were 0.5–1.5 mM in protein. Samples for CD studies were 5–20 μM in protein. Titrations with denaturant were followed on unlabelled samples through ^{1}H-1D

spectra at 700 MHz and CD spectra were acquired on a Jasco J-800 spectropolarimeter at room temperature.

A.1.5 NMR and CD Experiments

Detailed experimental procedure on NMR experiments and CD study were explained in the attached paper in results section (Chapter 3).

B. Copper Transporting ATPases PacS of Cyanobacteria *Synechocystis* PCC 6803

B.1 Experimental Methods

B.1.1 Cloning, Expression and Purification $PacS_N$

Amino terminal protein Sequence of $PacS_N$;

MAQTINLQLEGMRCAACASSIERAIAKVPGVQSCQVNFALEQAVVSYHGETTPQI LTDAVERAGYHARVLKQQVLSSQQTEDRKPVFSAKLVTGL **(NCBI** P73241)

Synechocystis PCC 6803 genomic DNA was used as template for PCR with primers 5'-GAACATATGGCCCAAACCATC-3' and 5'-GAAGAATTCTCATAACCCC GTTACCAATTTGGCCGA-3'. The amplified fragment of DNA containing codons 1-95 encoding the entire amino-terminal region of $PacS_N$ was ligated into the *NdeI/EcoRI* sites of pET29a to create $pETPacS_N$ (Fig.16).

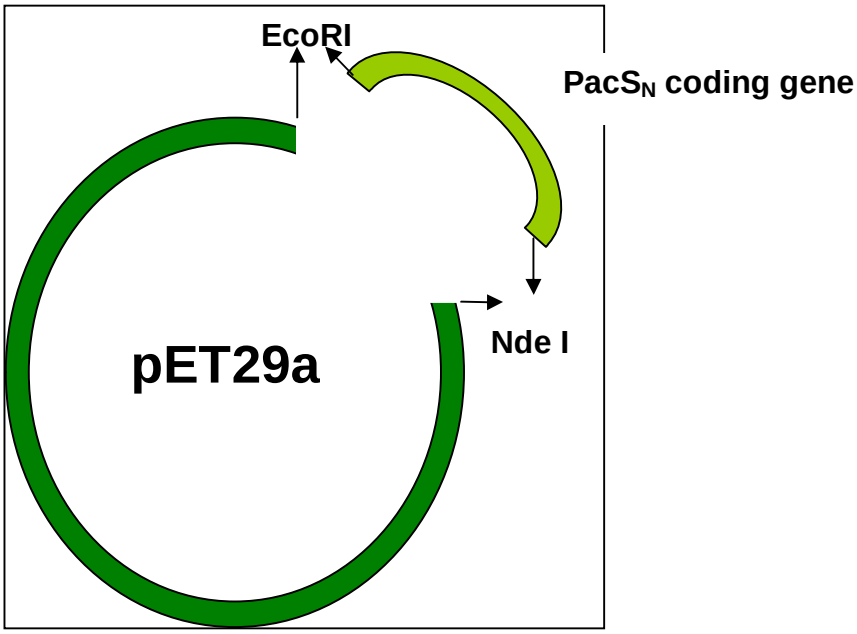


Figure 16.Cloning of amino terminal gene PacS$_N$ in pET29a

Recombinant protein was expressed as a soluble protein in *E. coli*. The entire expression tests were conducted in Minifor fermentor. To obtain unlabelled protein, BL21 cells harboring the pETPacS$_N$ plasmid were grown at 37°C, exposed to 1mM IPTG (isopropyl-1-thio-D-galactopyranoside) and 1 mM of copper sulfate in case of rich media. For expression of labeled PacS$_N$, M9 minimal medium (16) supplemented with [^{13}C] glucose and (^{15}NH4)2SO$_4$ as only sources of carbon and nitrogen was used. Cells were grown at 37°C to an A$_{600}$ nm of 0.6 prior to induction with 1 mM isopropyl-1-thio-D-galactopyranoside (IPTG), copper sulfate was added to the growth medium to a final concentration of 50 µM. Cells were harvested after 6 hrs in case of rich media (LB) and 7 hrs in case of minimal medium. Cells were then lysed by freeze-thawing and repeated sonication. The lysate was clarified by centrifugation at 100,000 × g for 40 min at 4 °C. The supernatant was filtered through a low protein-binding 0.45 µm filter and loaded

onto a 5ml HiTrap Q XL column (Amersham Pharmacia Biotech) which was pre-equilibrated with Tris 25 mM pH 9. Chromatography was done using a Pharmacia fast protein liquid chromatography (FPLC) unit. The protein was eluted by a 100 ml 50mM NaCl step. The fractions containing $PacS_N$ were concentrated and subjected to size exclusion chromatography on a HiLoad 16/60 Superdex 75 column (Amersham Pharmacia Biotech) equilibrated in 25 mM Tris pH 9. To prevent disulfide formation, which might occur because of the presence of two cysteines, the protein samples were kept in anaerobic conditions. In these conditions the buffer was changed by an ultra filtration device (Amicon) against sodium phosphate pH 7.0 with the addition of reducing agent dithiothreitol (DTT) at a final concentration of 4 mM. Protein expression and purification was monitored by sodium dodecyl phosphate-polyacrylamide (SDS) gel electrophoresis in 17% polyacrylamide gels stained with Coomasie brilliant blue R-250 against Perfect Protein marker (Novagen).

B.1.2 Metallation of $PacS_N$ with Cu (I)

$ApoPacS_N$ was diluted to 200 μM in 20 mL of Tris-MES (100mM, pH 8.0) in presence of 5 mM DTT in the glove box. 1.5 equivalents of [Cu (CH_3CN) $^+PCl^-$] was added to the protein solution and allowed to equilibrate for ~ 30 min. The unbound metal was washed away with three buffer exchanges without DTT (50:1 dilutions each time). Then protein was concentrated to 1 mM by using amicon ultrafiltration cell..

B.1.3 NMR Experiments and Structure Calculations of $apoPacS_N$

Detailed experimental procedures on NMR experiments, structural calculation,

mobility studies and protein–protein interaction were explained in the attached paper (Chapter 3).

C. Zinc Transporting ATPases ZiaA of Cyanobacteria *Synechocystis* PCC 6803

C.1 Experimental Methods

C.1.1 Cloning, expression and structural characterization of $ZiaA_N$

Protein Sequence of amino terminal $ZiaA_N$;

MTQSSPLKTQQMQVGGMDCTSCKLKIEGSLERLKGVAEASVTVATGRLTVTYD PKQVSEITIQERIAALGYTLAEPKSSVTLNGHKHPHSHREEGHSHSHGAGEFNLK QEL (NCBI: BAA17301)

Synechocystis PCC 6803 genomic DNA was used as template for PCR with primers 5'-GGATCCATGACCCAATCTTCACCGCTCAAAAC-3' with 5'-CTCGAGT AGTTCTTGTTTCAGATTAAATTC-3'. The amplified fragment of DNA containing codons (1-111 residue) encoding the entire amino-terminal region of ZiaA was ligated into the *Bam*HI/*Xho*I sites of pET29a to create pETZiaA (Fig.17).

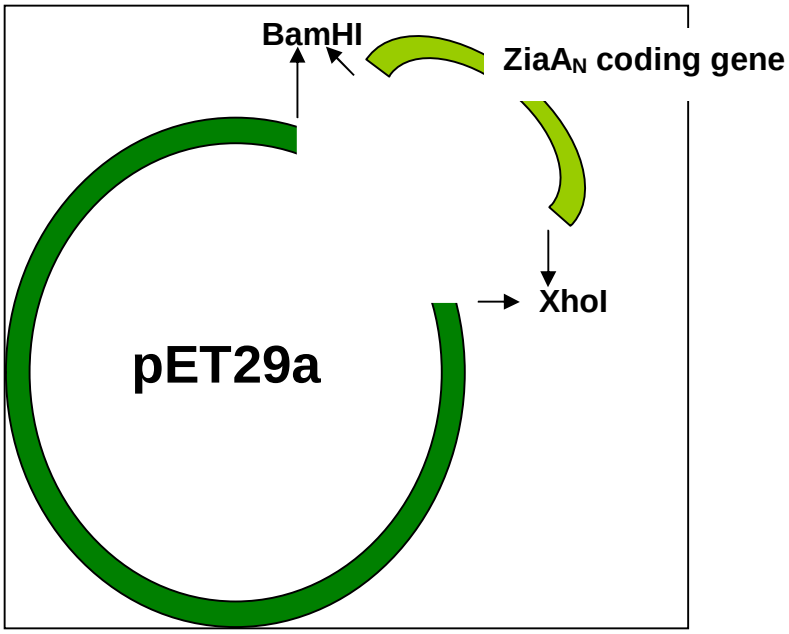


Figure 17. Cloning of amino terminal domain of ZiaA$_N$ in pET29a

Recombinant protein was expressed as a soluble protein in *E. coli*. The entire expression test was conducted in Minifor fermentor in order increase the protein yield. To obtain unlabelled protein, BL21 cells harboring the pETZiaA$_N$ plasmid were grown at 37°C, exposed to 0.5mM IPTG (isopropyl-1-thio-D-galactopyranoside) and 1 mM of zinc sulfate in case of rich media. For expression of labeled ZiaA$_N$, M9 minimal medium supplemented with [^{13}C] glucose and (^{15}NH4)2SO4 as only sources of carbon and nitrogen was used. Cells were grown at 37°C to an A$_{600}$ nm of 0.6 prior to incubation with 1 mM IPTG, zinc sulfate was added to the growth medium to a final concentration of 75 µM. Cells were harvested after 6 hrs in case of rich media and 7 hrs in case of minimal media. Cells were then lysed by freeze-thawing and repeated sonication. The lysate was clarified by centrifugation at 100,000 × g for 40 min at 4 °C. The supernatant was filtered through a low protein-binding 0.45 µm filter and loaded onto a 5ml Hi-Trap

Q XL column (Amersham Pharmacia Biotech) equilibrated with Tris 25 mM pH 9. Chromatography was done using a pharmacia fast protein liquid chromatography (FPLC) unit. The protein was eluted by a 100 ml 50mM NaCl step. The fractions containing $ZiaA_N$ were concentrated and subjected to size exclusion chromatography on a Hi-Load 16/60 Superdex 75 column (Amersham Pharmacia Biotech) equilibrated with 25 mM Tris pH 9. To prevent disulfide formation, which might occur because of the presence of two cysteines, the protein samples were kept in anaerobic conditions. In these conditions the buffer was changed by an ultra filtration device (Amicon) against sodium phosphate pH 7.0 with the addition of reducing agent dithiothreitol (DTT) at a final concentration of 4 mM. Protein expression and purification was monitored by sodium dodecyl phosphate-polyacrylamide (SDS) gel electrophoresis in 17% polyacrylamide gels stained with Coomasie brilliant blue R-250 against Perfect Protein marker (Novagen).

C.1.2 Preparation of Zn-ZiaA samples

$ZiaA_N$ was diluted to 200 μM in 10 mL HEPES (100mM, pH 7.0) buffer in presence of 5 mM DTT. One equivalent of $ZnSO_4$ was added to the protein solution and allowed to equilibrate for ~30 minutes. The unbound metal was washed away by passing the solution again in PD10 desalting column. Finally, the sample was concentrated to 500 μl and 10% D_2O was added.

C.1. 3 NMR Experiments and Structure Calculations of $ZiaA_N$

Detailed experimental procedures on NMR experiments, structural modeling and mobility studies were explained in the Chapter 3.

D. Zinc Finger 9 /Cellular Nucleic Acid Binding (CNBP/ZNF9)

D.1 Experimental Methods

D.1.1 Cloning and Expression of ZNF9 protein

Protein sequence of ZNF9

MSSNECFKCGRSGHWARECPTGGGRGRGMRSRGRGGFTSDRGFQFVSSSLPDICYRCGESGHLAK
DCDLQEDACYNCGRGGHIAKDCKEPKREREQCCYNCGKPGHLARDCDHADEQKCYSCGEFGHIQ
KDCTKVKCYRCGETGHVAINCSKTSEVNCYRCGESGHLARECTIEATA

The genes encoding human ZNF9 (1-177 residues) was PCR amplified from human brain cDNA library. The PCR amplificate was first cloned into an entry vector, pENTR/TEV/D-TOPO (Invitrogen) and further sub-cloned into an expression vector (pDEST-17), able to express N-terminal His-tagged protein (Fig.18). Bacterial clones were screened by either restriction digestion and/or PCR amplification by using a mixture of vector plasmid specific M13 forward primer and the gene specific reverse primer. Finally, positive bacterial clones were confirmed by DNA sequencing and one clone was chosen for expression. Bacterial cultures were induced with 0.2% L-arabinose at OD_{600} reached 0.6. Over-expression of the protein was seen after 6 hrs in rich medium.

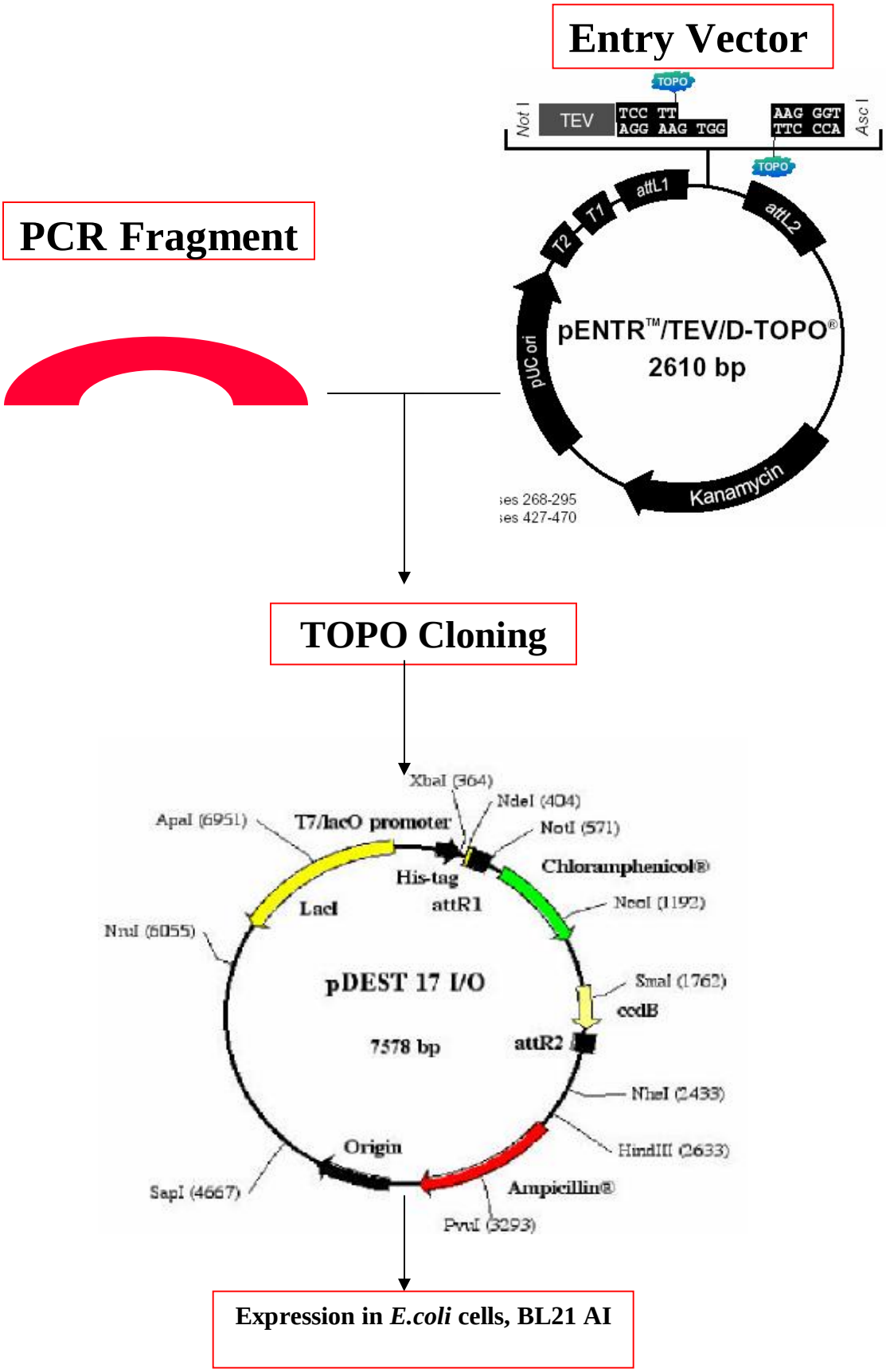


Figure 18. Schematic representation of the strategy of cloning of human ZNF9 protein in *E.coli*

D.1.2 Refolding of ZNF9 polypeptide

ZNF9 was expressed in inclusion bodies. Protein was purified under denaturing condition by using 8M urea. Since the protein was expressed with 6X Histidine tag, His-tag was exploited for purification and refolding. Hi-trap chelating column charged with Ni (SO_4) was used for purification and urea was removed by running a gradient. The final concentration of urea in the protein was 2 M.

Finally protein was refolded by using following additives;

1. Glycerol (10%)
2. Sarcosine(0.5M)
3. N-acetyl glucosamine (NAG) (1M)
4. Sulphobetaines (0.5M)
5. Tween (0.5M)
6. L-Arginin (1M)

E. Catalytic Domain of A Disintegrin and Metalloprotease 10 (ADAM10)

E.1 Experimental Methods

E.1.1 Cloning, Expression and Purification of ADAM10 protein

Protein sequence of ZNF9

QLYIQTDHLFFKYYGTREAVIAQISSHVKAIDTIYQTTDFSGIRNISFMVKRIRINTTADEKDPTNPFR
FPNIGVEKFLELNSEQNHDDYCLAYVFTDRDFDDGVLGLAWVGAPSGSSGGICEKSKLYSDGKKK
SLNTGIITVQNYGSHVPPKVSHITFAHEVGHNFGSPHDSGTECTPGESKNLGQKENGNYIMYARAT
SGDKLNNNKFSLCSIRNISQVLEKKRNNCFVESG

The genes encoding human ADAM10 (1-234 residues) was PCR amplified from human brain cDNA library. ADAM 10 was cloned and expressed by similar cloning and expression strategy to that of ZNF9, with slight modification in the induction time. The expression of ADAM 10 was maximized at 4 hours of post L-arabinose induction. ADAM 10 was purified by using Hi-trap chelating column charged with $Ni(SO_4)$.In addition size exclusion chromatography was used here. The protein was very pure at the end of second column. The preliminary protein folding was checked by 1D NMR. Further ^{15}N labeled sample was prepared by using ($^{15}NH4)2SO_4$ as only nitrogen source.^{1}H-^{15}N HSQC was performed to check the folding secondary folding of the protein.

RESULTS

First part of my research work during the period of my doctorate was expression, purification and structural characterization of *Bacillus pasteruii* cytochrome c mutants. Second part of my work was cloning, expression, purification and structural characterization of N-terminal copper and zinc binding domains of P1-type of ATPases PacS and ZiaA from *Synechocystis* PCC 6803.In addition to these proteins I have also worked with cloning, expression and purification of two more human zinc binding proteins, ZNF9 and ADAM10.

A. Cytochrome *c* Mutants of *Bacillus pasteruii*

A.1 Overview

Bacillus species are Gram-positive microorganisms lacking a true periplasm and are obliged to store electron-transfer proteins as membrane-bound forms. Three types of class I *c*-type cytochromes have been identified in Gram-positive bacteria: (i) cytochromes fused as integral domains of subunit II in membrane-bound aa_3-type terminal oxidases, (ii) cytochromes bound to the membrane through an N-terminal hydrophobic polypeptide, and (iii) cytochromes bound to the membrane via a diacyl-glyceryl-cysteine moiety (1). The difficulties associated with the study of such water-insoluble proteins may explain why only a few *c*-type cytochromes have been isolated from Gram-positive bacteria (2-13). As a consequence, their properties and functional roles are scarcely known, and no structural information is available, in contrast with cytochromes from Gram-negative bacteria and eukaryotes (14-15). The Gram-positive alkaliphilic soil bacterium *Bacillus pasteurii* produces large amounts of membrane-bound cytochromes of the *b*- and *c*-types, as well as a terminal oxidase of the aa_3-type (16-17). The soluble form of a *c*-type cytochrome (named cytochrome c_{553}) has been purified (12).

B. pasteurii cytochrome c_{553} is a small (9.6 kDa) acidic (p*I* = 3.3) protein (12), having a chain length of 92 amino acids and a single heme-binding sequence pattern (Cys-X-X-Cys-His) located in the first half of the polypeptide chain (18). These features indicate that the protein belongs to class I cytochromes (19). Its amino acid sequence suggests that in vivo cytochrome c_{553} is bound to the cytoplasm membrane through a diacyl-glyceryl-cysteine anchor located at the N-terminus tail and that the soluble form is obtained by cleavage of this tail during cell disruption (18). A paramagnetic NMR spectroscopic study on the oxidized Fe(III) form confirmed the presence of a *c*-type heme and additionally allowed the identification of a hexacoordinate low-spin Fe ion axially bound to His and Met residues (Figure 7) (20). The thermodynamic parameters for the one-electron reduction process suggested that the heme group is highly exposed to the solvent and that extrusion of water molecules from the protein hydration shell occurs upon reduction (20). The function of cytochrome c_{553} in *B. pasteurii* is unknown, but it probably plays a role in respiratory metabolism (18). For most members of this class the heme binding site is preceded by an N-terminal peptide stretch of 12 to 18 residues. In those class I proteins of which the 3D-structure has been determined, this region appears to form the so-called N-terminal helix. *B. pasteurii c553* therefore can be said to contain an N-terminal extension of a 15 residues compared to the N-terminal helix region of most class I cytochromes. The occurrence of a methionine residue as the sixth heme ligand in the second part of the polypeptide chain, 34 residues downstream the fifth heme ligand histidine of the heme binding site, suggests that *B. pasteurii c553* is a member of the cytochrome *c* subclass as occurs in the prototype sequence of *Pseudomonas aeruginosa* cytochrome *c*-551 [21,22].

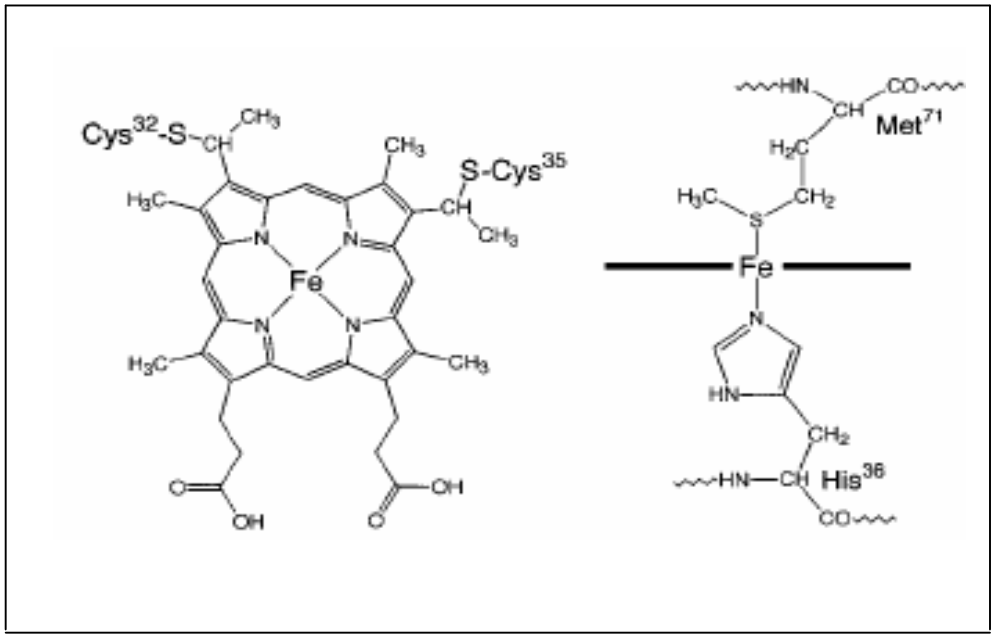


Figure 7. Scheme of the heme prosthetic group as detected in *B. pasteurii* cytochrome c_{553} by paramagnetic NMR spectroscopy. Top and side views are shown.

There is, however, very little sequence similarity with the latter cytochrome, except for a few residues. In Fig. 8 we can see the sequence of *B. pasteurii c553* with the sequences of all other *Bacillus* cytochromes *c* known so far. The comparison starts from Thr19, because this position corresponds to the N-terminal residue of most of the other *Bacillus* cytochromes. If we consider the chain length between the two heme axial ligand binding residues His and Met (positions 36 and 71 in *B. pasteurii c553*), there appear to be two groups of *Bacillus* cytochromes *c:* those with around 35 (subgroup S, sequences (7)–(13)) and those with around 50 residues between these sites (subgroup L, sequences (1)–(6)).

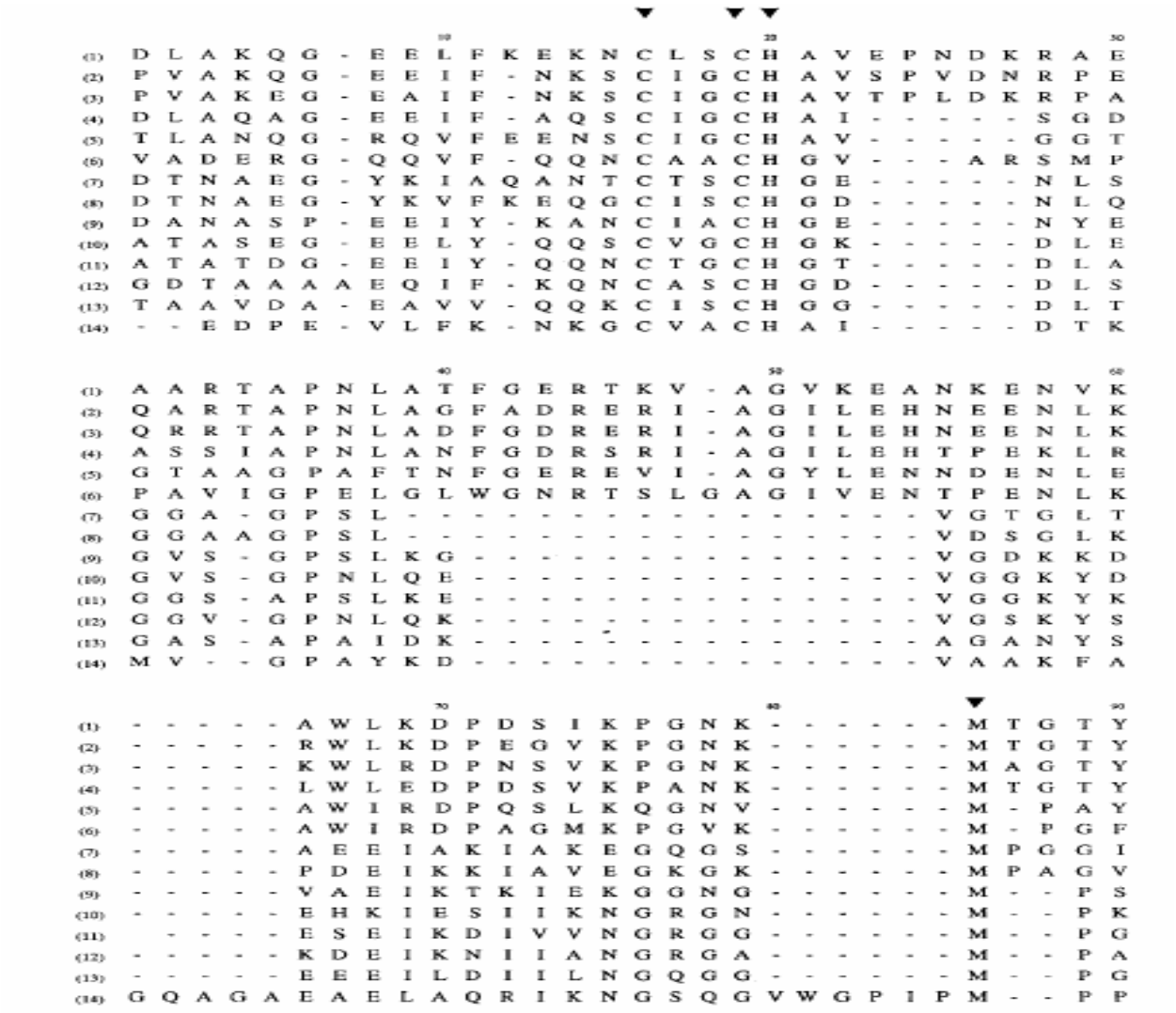


Figure 8. Amino acid sequence of the cytochromes c from several Bacillus strains starting the alignment with Thr_{19} of *B.pasteruii c553*.The heme binding residues are indicated by▼. Sequence (14) is from *Pseudomonas aeurigonasa* cytochrome c-551 for which also the 3D structure is known.

(1) ***Bacillus subtilis*** **CTAC**
(2) ***Bacillus stearothermophilus***
(3) ***Bacillus sp*** **PS3 CTAC**
(4) ***Bacillus sp*** **YN2000**
(5) ***Bacillus firmus*** **CTAC**
(6) ***Thermus thermophilus*** **CTAC**
(7) ***Bacillus stearothermophilus*** QCRC
(8) ***Bacillus subtilis*** **QCRC**
(9) ***Bacillus subtilis*** **c-550**
(10) ***Bacillus subtilis*** **c-551**
(11) ***Bacillus licheniformis*** **c-551**
(12) ***Bacillus subtilis*** **PS3 c-551**
(13) ***Bacillus pasteurii*** **c553**
(14) ***Pseudomonas aeurigonasa*** **c-551**

B. pasteurii cytochrome *c553* belongs to subgroup S, even without having to consider insertions or deletions compared to the other members of that group. A slightly decreased similarity is observed with *Bacillus sp.* PS3 cytochrome *c*-551 (45%) and *B. subtilis* cytochrome *c*-551 (41%), while much less similarity is detected with the *B. subtilis* cytochrome *c*-550 (23%).

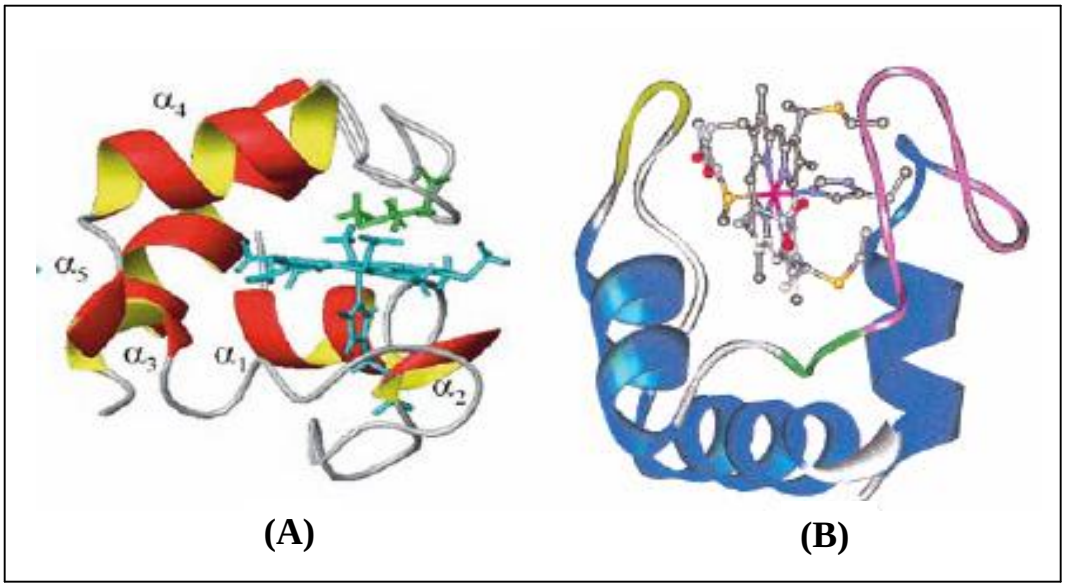


Figure 9. Solution (A) and Crystal structure (B) of *Bacillus pateruii* cytochrome c553.

The solution structure (Fig.9A) of oxidized *Bacillus pasteurii* cytochrome c553 was reported (24) and compared with the available crystal structure (Fig.9B) (23). The solution structure closely resembles the fold observed in the crystal structure. At variance with mitochondrial cytochrome c proteins, this protein does not experience pH-dependent coordination equilibria (24). Detailed characterization of the unfolding mechanisms and of the factors determining the thermodynamic stability of a cytochrome c553 was also

deeply investigated [24, 25, 26, 27, and 28]. Since the binding of the axial Met represents a crucial and yet not fully understood aspect of the biochemistry of cytochromes *c*, we have investigated the structural properties of the loop containing the iron axial ligand Met71 in oxidized Bpcyt*c* using site directed mutagenesis at different sequence positions, in order to obtain better insight into the determinants of the peculiar high stability of the axial coordination towards GdmCl and alkaline pH.

The aim of this research work was to undertake a detailed characterization of the unfolding mechanisms and of the factors determining the thermodynamic stability of a minimal 71-residue mono-heme cytochrome *c*. The other objective of this research work was to gain insight into the general factors and system-specific variations associated with the folding properties of the very widespread cytochrome c fold.

A.2 References

1. Sone, N., and Toh, H. (1994) *FEMS Microbiol. Lett. 122*, 203-210.
2. Jacobs, A. J., Kalra, V. K., Cavari, B., and Brodie, A. F. (1979) *Arch. Biochem. Biophys. 194*, 531-541.
3. Woolley, K. J. (1987) *Arch. Biochem. Biophys. 254*, 376-379.
4. Davidson, M. W., Gray, K. A., Knaff, D. B., and Krulwich, T. A. (1988) *Biochim. Biophys. Acta 933*, 470-477.
5. Sone, N., Kutoh, E., and Yanagita, Y. (1989) *Biochim. Biophys. Acta 977*, 329-334.
6. von Wachenfeldt, C., and Hederstedt, L. (1990) *J. Biol. Chem. 265*, 13939-13948.
7. Hreggvidsson, G. O. (1991) *Biochim. Biophys. Acta 1058*, 52-55.

8. Yumoto, I., Fukumori, Y., and Yamanaka, T. (1991) *J. Biochem. 110*, 267-273.
9. Saraiva, L. M., Denariaz, G., Liu, M.-Y., Payne, W. J., LeGall, J., and Moura, I. (1992) *Eur. J. Biochem. 204*, 1131-1139.
10. Fujiwara, Y., Oka, M., Hamamoto, T., and Sone, N. (1993) *Biochim. Biophys. Acta 1144*, 213-219.
11. Nitschke, W., Schoepp, B., Floss, B., Schricker, A., Rutherford, A. W., and Liebl, U. (1996) *Eur. J. Biochem. 242*, 695-702.
12. Benini, S., Ciurli, S., Rypnjewski, W., and Wilson, K. S. (1997) *Proteins: Struct. Funct. Genet. 28*, 580-585.
13. Albert, I., Rutherford, A. W., Grav, H., Kellermann, J., and Michel, H. (1998) *Biochemistry 37*, 9001-9008.
14. Moore, G. R., and Pettigrew, G. W. (1990) in *Cytochrome c: Evolutionary, Structural and Physicochemical Aspects*, Springer-Verlag, Berlin.
15. Scott, R. A., and Mauk, A. G. (1996) in *Cytochrome c. A Multidisciplinary Approach*, University Science Books, Sausalito, CA.
16. Haddock, B. A. and Cobley, J. G. (1976) *Biochem. Soc. Trans. 4*, 709-711.
17. Hoddinott, M. H., Reid, G. A., and Ingledew, W. J. (1978) *Biochem. Soc. Trans. 6*, 1295-1298.
18. Vandenberghe, I. H. M., Guisez, Y., Ciurli, S., Benini, S., and Van Beeumen, J. J. (1999) *Biochem. Biophys. Res. Commun. 264*, 380-387.
19. Pettigrew, G. W., and Moore, G. R. (1987) in *Cytochromes c: Biological Aspects*, Springer-Verlag, Berlin.

20. Benini, S., Borsari, M., Ciurli, S., Dikiy, A., and Lamborghini, M. (1998) *J. Biol. Inorg. Chem. 3*, 371-382.

21. Hicks, D. B., and Krulwich, T. A. (1995) *Biochim. Biophys. Acta 1229*, 303-314.

22. Ambler, R. P. (1963) *Biochem. J.* **89,** 349–378.

23. S. Benini, W. Rypniewski, K. S. Wilson, J. Van Beeumen, S. Ciurli, *Biochemistry* 2000, 39, 13115 -13126.

24. Banci L, Bertini I, Ciurli S, Dikiy A, Dittmer J, Rosato A,Sciara G, Thompsett A (2002) *ChemBioChem* 3:299–310

25. Bartalesi I, Bertini I, Ghosh K, Rosato A, Turano P (2002) *J Mol Biol* 321:693–701

26. Bartalesi I, Bertini I, Rosato A (2003) *Biochemistry* 42:739–745

27. Bertini I, Ghosh K, Rosato A, Vasos P (2003) *Biochemistry* 42:3457–3463

28. Bartalesi I, Rosato A, Zhang W (2003) *Biochemistry* 42:10923–10930

29. Arslan, H. Schulz, R. Zufferey, P. Kuenzler, L. Thoeny-Meyer, *Biochem. Biophys. Res. Commun.* 1998, 251, 744-747.

A.3 Cloning, expression and characterization of the mutants

The plasmid used to produce mutants was generated from the pAT1 plasmid encoding WT Bpcytc. Following are the mutants developed by using the Quick-Change Site-directed mutagenesis kit (Stratagene);

1. Q68K (Glutamine ⟶ Lysine)
2. P72A (Proline ⟶ Alanine)
3. P72G (Proline ⟶ Glycine)
4. I75A (Isoleucine ⟶ Alanine)

Two clones were selected from each mutant for screening by DNA sequencing. The positive clone from each mutant was selected for the protein expression. I75V mutant was selected for test expression and the culture was induced at $O.D_{600}$ reaching 1.2 with 0.5 mM IPTG at 37 ^{0}C. Samples were collected at different time intervals (0h, 4h, 6h, 8h, 16h and 24 h).The protein was well expressed after 24 hrs (Figure 19). The other three mutants were also expressed with the same procedure. All the mutants were purified by using anion exchange chromatography and size exclusion chromatography. In all the mutants the protein was pure at the end of final column more than 90% (Fig.20). Folding status of I75V was checked by 1D NMR (Fig. 21) using unlabelled protein. ^{1}H-^{15}N HSQC was done at 298 K to follow up the state of fold of I75V mutant (Fig.22) and compared with wild type ^{1}H-^{15}N HSQC.

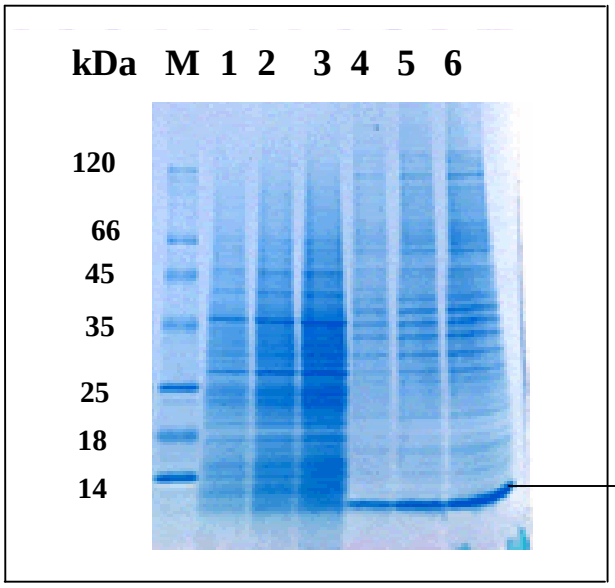


Figure 19.17% SDS PAGE profile of test expression for I75V mutant, (1) 0 hr (2) 4 hr (3) 6 hr (4) 8 hr (5) 16 hr (6) 24 hr. M-Protein Marker.

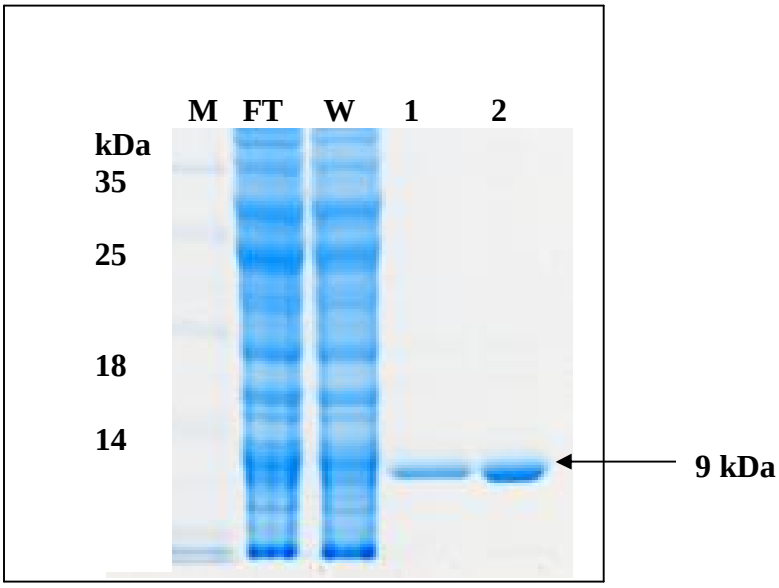


Figure 20.17% SDS PAGE showing the purity of protein after size exclusion. M-Marker, FT-Flow through, W-Wash, 1&2-I75V protein samples.

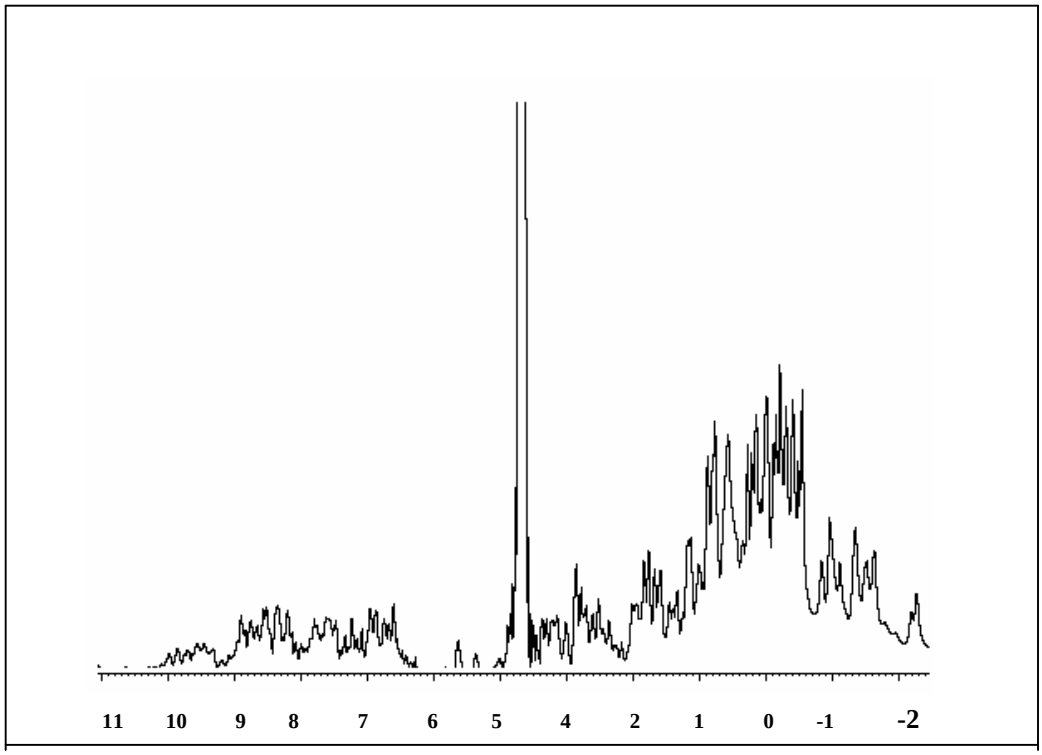


Figure 21.1D NMR spectra of I75V mutant at 298 K

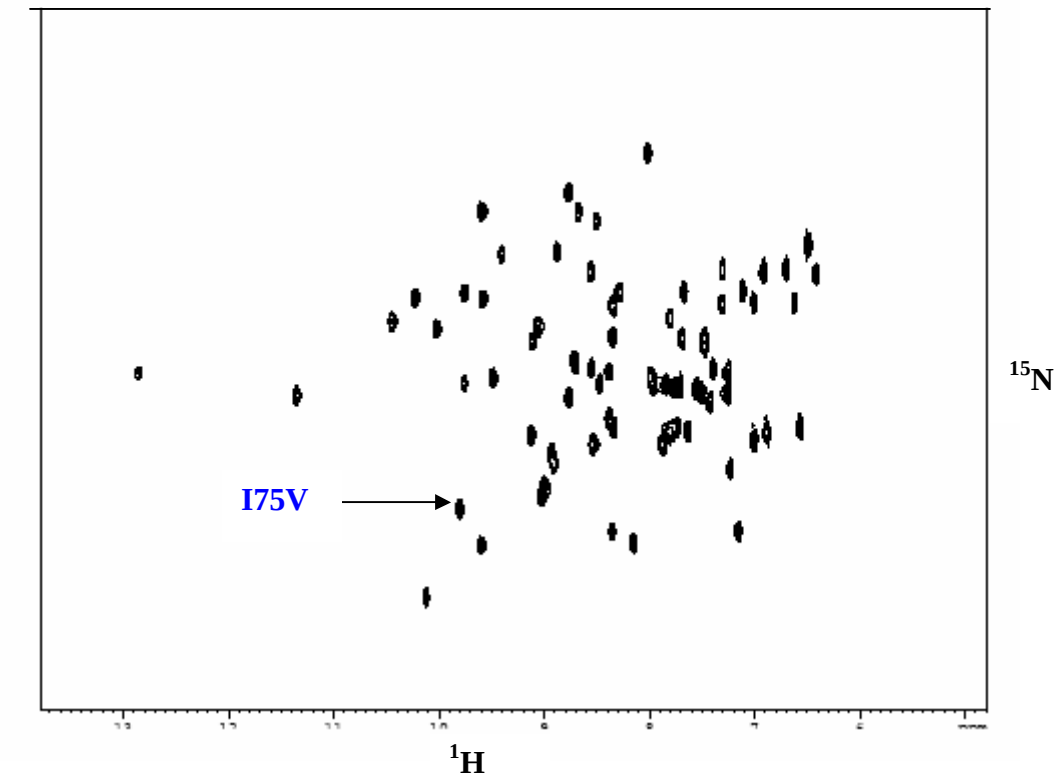


Figure 22.^{1}H-^{15}N HSQC of I75V mutant at 298 K

A.4 NMR and CD results

Further all four mutants were used to study the unfolding by using GdmCl. Stability of each mutant was analyzed by using CD and NMR techniques. All the results were described in the following attached article;

Protein stability and mutations in the axial methionine loop of a minimal cytochrome c

Ilaria Bartalesi Ivano Bertini Giulia Di Rocco Antonio Ranieri Antonio Rosato
Murugendra Vanarotti *Paul R. Vasos Maria Silvia Viezzoli*

J Biol Inorg Chem **(2004) 9: 600–608**

J Biol Inorg Chem (2004) 9: 600–608
DOI 10.1007/s00775-004-0558-8

ORIGINAL ARTICLE

Ilaria Bartalesi · Ivano Bertini · Giulia Di Rocco
Antonio Ranieri · Antonio Rosato
Murugendra Vanarotti · Paul R. Vasos
Maria Silvia Viezzoli

Protein stability and mutations in the axial methionine loop of a minimal cytochrome *c*

Received: 17 July 2003 / Accepted: 3 May 2004 / Published online: 3 June 2004
© SBIC 2004

Abstract The minimal mono-heme ferricytochrome *c* from *Bacillus pasteurii*, containing 71 amino acids, has been further investigated through mutagenesis of different positions in the loop containing the iron ligand Met71. These mutations have been designed to sample different aspects of the loop structure, in order to obtain insights into the determinants of the stability of the iron(III) environment. In particular, positions 68, 72 and 75 have been essayed. Gln68 has been mutated to Lys to provide a suitable alternate ligand that can displace Met71 under denaturing conditions. Pro72 has been mutated to Gly and Ala to modify the range of allowed backbone conformations. Ile75, which is in van der Waals contact with Met71 and partly shields a long-lived water molecule in a protein cavity, has been substituted by Val and Ala to affect the network of inter-residue interactions around the metal site. The different contributions of the above amino acids to protein parameters such as structure, redox potential and the overall stability against unfolding with guanidinium hydrochloride are analyzed. While the structure remains essentially the same, the stability decreases with mutations. The comparison with mitochondrial *c*-type cytochromes is instructive.

I. Bartalesi · I. Bertini (✉) · A. Rosato · M. Vanarotti
P. R. Vasos · M. S. Viezzoli
Magnetic Resonance Center (CERM),
University of Florence, Via L. Sacconi 6,
50019 Sesto Fiorentino, Italy
E-mail: bertini@cerm.unifi.it
Tel.: +39-055-4574272
Fax: +39-055-4574271

I. Bertini · A. Rosato · M. S. Viezzoli
Department of Chemistry, University of Florence,
Via della Lastruccia 3, 50019 Sesto Fiorentino,
Italy

G. Di Rocco · A. Ranieri
Department of Chemistry,
University of Modena and Reggio Emilia,
Via G. Campi 183, 41100 Modena,
Italy

Keywords Cytochrome *c* · NMR spectroscopy · Protein unfolding · Site-directed mutagenesis · Water-protein interaction

Abbreviations *Bp*cyt*c*: soluble fragment of cytochrome c_{553} from *Bacillus pasteurii* · GdmCl: guanidinium chloride · I75A: Ile75 to Ala mutant · I75V: Ile75 to Val mutant · P72A: Pro72 to Ala mutant · P72G: Pro72 to Gly mutant · Q68K: Gln75 to Lys mutant · WT: wild type

Introduction

We have recently undertaken a detailed characterization of the unfolding mechanisms and of the factors determining the thermodynamic stability of a minimal 71-residue mono-heme cytochrome *c*: the soluble fragment of cytochrome c_{553} from *Bacillus pasteurii* (*Bp*cyt*c* hereafter) [1, 2, 3, 4, 5]. The aim of this research is to obtain information that can be used, in comparison also with the intensely studied mitochondrial *c*-type cytochromes, to gain insight into the general factors and system-specific variations associated with the folding properties of the very widespread cytochrome *c* fold.

There are practically no differences between the backbone conformations of reduced and oxidized states of the wild-type (WT) protein [3]. Also the dynamic properties of *Bp*cyt*c* in the two physiologically relevant oxidation states are similar [3], with reduced *Bp*cyt*c* being slightly more rigid, as commonly observed (although to a larger extent) for *c*-type cytochromes [6]. Recently, we have identified within the protein frame of *Bp*cyt*c* a long-lived water molecule [4], which does not correspond to any of the buried waters bound to mitochondrial cytochromes *c*. This water molecule is not significantly affected by variations in the redox state of the heme iron [4], probably as a consequence of its being involved in three hydrogen bonds with the polypeptide chain (as suggested by inspection of the high-resolution crystal structure of oxidized *Bp*cyt*c* [7]).

WT oxidized *Bp*cyt*c* features unchanged His/Met heme axial coordination at pH values up to 12, at variance with mitochondrial cytochromes *c* [2, 8]. Heme mis-ligation is induced at neutral pH by high concentration (>4 M) of guanidinium chloride (GdmCl), and occurs concomitantly with extensive unfolding of the polypeptide chain [2]. There are several elements of evidence, including the significant analogy of the organization of opening units across the protein structures [5], suggesting that the relatively uncommon stability of the Met loop in oxidized *Bp*cyt*c* is mainly dependent on the local properties of the loop region itself, i.e. its primary sequence rich in Gly and Pro residues, its short size and the low number of long-range contacts with the remainder of the protein [2]. Since the binding of the axial Met represents a crucial and yet not fully understood aspect of the biochemistry of cytochromes *c*, we have investigated some of the key factors relevant to it.

In this work, three locations within the Met loop have been identified which appeared potentially capable of directly affecting the loop stability: Gln68, Pro72 and Ile75. These sites have been probed through site-directed mutagenesis; the mutants generated (Q68K, P72A, P72G, I75A, I75V) have been investigated in the oxidized state through a combination of UV-Vis, CD and NMR spectroscopy. In addition, the reduction thermodynamics for the WT species and some mutants have been determined through direct electrochemistry experiments and compared with those of other class I cytochromes *c*.

Materials and methods

The plasmids used to produce protein samples was generated from the pAT1 plasmid encoding WT *Bp*cyt*c*. To generate the mutants Q68K, P72A, P72G and I75A, the *Bp*cyt*c* gene was mutated following the procedure of the QuickChange site-directed mutagenesis kit (Stratagene). Protein production and purification was done as previously reported [1]. All mutants were studied in the oxidized state in 100 mM phosphate buffer at pH 7.0 unless otherwise stated.

The approach used for resonance assignment and structure calculation was the same as in previous works [1, 3]. 1D ^{1}H NOE measurements were performed with the methodology described earlier [9], and with parameters analogous to those used for WT oxidized *Bp*cyt*c* [1]. Structure calculations were run with the program CYANA [10], using CANDID [11] for automated NOESY assignment.

Protein dynamics and backbone amide proton exchange rates were characterized using data acquired at 400 MHz and/or 600 MHz. The sub-nanosecond and milli- to microsecond time scales are characterized based on the measured values of ^{15}N nuclear relaxation rates and ^{1}H–^{15}N heteronuclear NOEs. R_1 and R_2 values and exchange rates were obtained by fitting the cross peak intensities to a single exponential as a function of the relaxation delay. The errors on the fit were estimated by repeating the fit 500 times while varying the data randomly within the error range of the experiment. Heteronuclear NOE values were calculated as the ratio of peak volumes in spectra recorded with and without saturation. The error in NOE values was calculated by recording a second data set and repeating the analysis. Only relaxation data for which the error was lower than 20% were used in the fit. Overlapping signals have been discarded.

Protein–water interactions were probed through CLEANEX, TrROESY and ePHOGSY experiments, as described elsewhere [4].

Samples for protein unfolding studies were prepared as described previously [2]. All unfolding measurements were carried out at pH 7.0. Titrations with denaturant were followed on unlabelled samples through ^{1}H 1D spectra at 700 MHz, acquired using spectral windows between 70 and 200 ppm and recycle times between 0.5 and 1.0 s. UV-Vis spectra were acquired on a Varian Cary 3 UV-visible scanning spectrophotometer at room temperature. CD spectra were acquired on a Jasco J-800 spectropolarimeter at room temperature. Denaturation profiles were obtained by UV-Vis and CD spectroscopy at pH 7.0 as described previously [2], and analyzed through the well-known linear extrapolation method [12]:

$$\Delta G^{\mathrm{D}} = \Delta G^{\mathrm{H_2O}} - m[\mathrm{GdmCl}] \quad (1)$$

where ΔG^{D} is the free energy difference between the folded and unfolded forms (called respectively LS and LS_1 in [2]) at a specific concentration of GdmCl, defined by [GdmCl], ΔG^{H2O} is the free energy difference between the folded and unfolded forms in the absence of denaturant, and m is a constant that relates to the increase in the degree of exposure of the protein during the transition [13]. This equation holds in the assumption of only two states being present, each insensitive to changes in [GdmCl].

Cyclic voltammetry (CV) experiments were performed with a potentiostat/galvanostat PAR model 273A at different scan rates (v = 0.02–0.5 V/s) using a small-volume cell (V = 0.5 mL) under argon. A 2-mm diameter gold disc was used as the working electrode, while a calomel and a platinum sheet were used as the reference and counter electrodes, respectively. The electric contact between the SCE electrode and the sample solution was obtained with a Vycor set. Potentials were calibrated against the MV^{2+}/MV^{+} couple (MV = methylviologen). All reduction potentials reported in this paper are referred to the standard hydrogen electrode (SHE). The working electrode was polished with an alumina (particle size of about 0.015 μm) water slurry on cotton wool subsequent to immersion in absolute ethanol for 10 min, then the electrode was treated in a ultrasonic pool for 10 min. Finally, the electrode was subjected to 10 voltammetric cycles between +0.7 and −0.6 V (scan rate

0.1 V/s) to minimize residual adsorbed impurities. Modification of the electrode surface was performed by dipping the polished electrode into a 1 mM solution of 4-mercaptopyridine for 30 s, then rinsing it with nanopure water. A 100 mM sodium chloride solution (in the presence of 10 mM phosphate buffer) was used as base electrolyte. Protein solutions were freshly prepared before use and their concentration, varying over 0.1–0.3 mM, was checked spectrophotometrically. The temperature dependence of the reduction potential was determined with a ''nonisothermal'' cell [14], in which the reference electrode is kept at constant temperature (20 ± 0.1 °C), while the half-cell containing the working electrode and the Vycor junction to the reference electrode is under thermostatic control with a water bath. With this experimental configuration, the reaction entropy for reduction of the oxidized protein ($\Delta S^{\circ\prime}_{rc}$) is given by [14, 15]:

$$\Delta S^{\circ\prime}_{rc} = S^{\circ\prime}_{red} - S^{\circ\prime}_{ox} = nF\left(dE^{\circ\prime}/dT\right) \quad (2)$$

Thus, $\Delta S^{\circ\prime}_{rc}$was determined from the slope of the plot of $E^{\circ\prime}$ versus temperature, which turns out to be linear under the hypothesis that $\Delta S^{\circ\prime}_{rc}$ is constant over the limited temperature range investigated. With the same assumption, the enthalpy change ($\Delta H^{\circ\prime}_{rc}$) was obtained from the Gibbs–Helmholtz equation, namely from the slope of the $E^{\circ\prime}/T$ versus $1/T$ plot. The nonisothermal behavior of the cell was carefully checked by determining the $\Delta H^{\circ\prime}_{rc}$ and $\Delta S^{\circ\prime}_{rc}$ values of the ferricyanide/ferrocyanide couple [14, 15, 16]. For each species, the experiments were performed at least two times and the reduction potentials were found to be reproducible within ± 2 mV.

Results

NMR spectroscopy under native conditions

Figure 1 shows a comparison of the 1D ^{1}H spectra of WT *Bp*cyt*c* and the mutants generated in this work. There are two patterns of distribution of the paramagnetically shifted signals of the heme substituents and of the axial ligands: the first pattern is that of the WT, Q68K, P72A and P72G proteins, while the second pattern is characteristic of the two Ile75 mutants. The two patterns differ mostly for the chemical shifts of the methyl groups of the heme and for the signals of the axial Met71 (Fig. 1). Figure 2 shows combined chemical shift differences of the backbone amide moieties ($\Delta\delta$) [17] with respect to WT for P72G and I75A, derived from ^{1}H–^{15}N HSQC spectra. We show results on these mutants because they are representative of the different behavior with respect to GdmCl (described in the next section). Chemical shift data are very similar for the P72G and P72A mutants. It is apparent that the vast majority of the polypeptide chain is essentially unaffected by the various mutations considered, with $\Delta\delta$ values well within

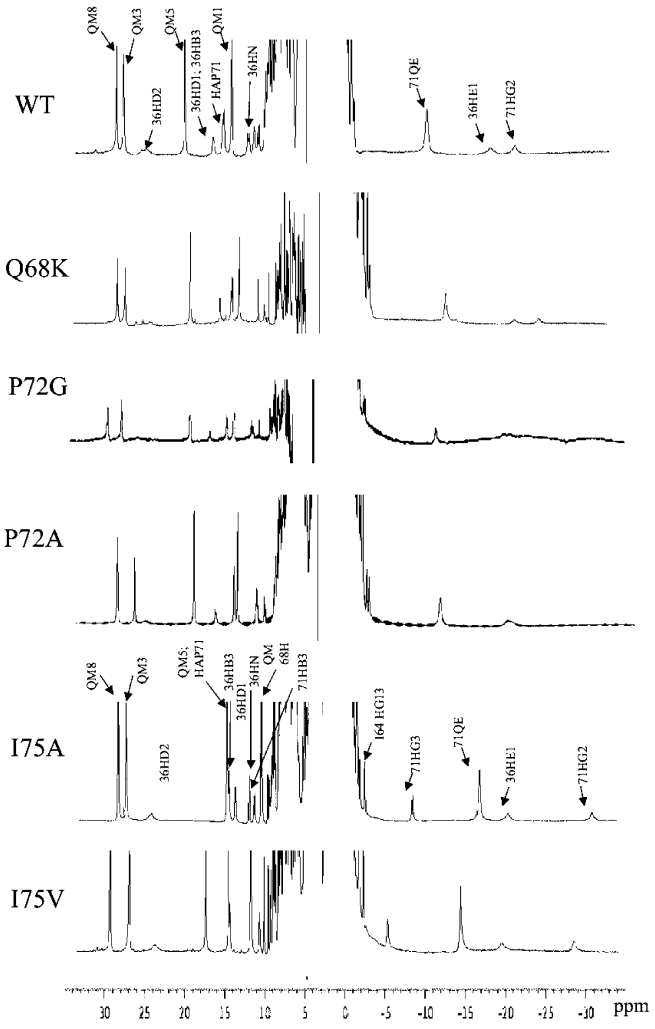


Fig. 1 1D NMR proton spectra of oxidized WT *Bp*cyt*c* and of various mutants studied in this work. Signals in the spectra of WT and I75A *Bp*cyt*c*, which are representative of the two different patterns observed, are labeled with the assignment of some hyperfine shifted signals. The labeled signals belong to the heme cofactor and to the axial ligands His36 and Met71. The spectra are recorded at 296 K. The samples were in 100 mM phosphate buffer, pH 7.0

0.15 ppm. The largest chemical shift variations are observed for the mutated amino acid and/or its neighbors, with other significant variations being confined within the Met loop. The only consistent exception to this general behavior is observed for the amide moiety residue His36 (one of the two iron ligands), which generally appears more sensitive to mutations than that of the other ligand, Met71. Overall, these data indicate that differences in the three-dimensional structure of WT and the mutants *Bp*cyt*c*, if any, should be very small.

Denaturation with GdmCl

Figure 3 shows the denaturation profiles of the various mutants, measured through CD spectroscopy. Best-fit

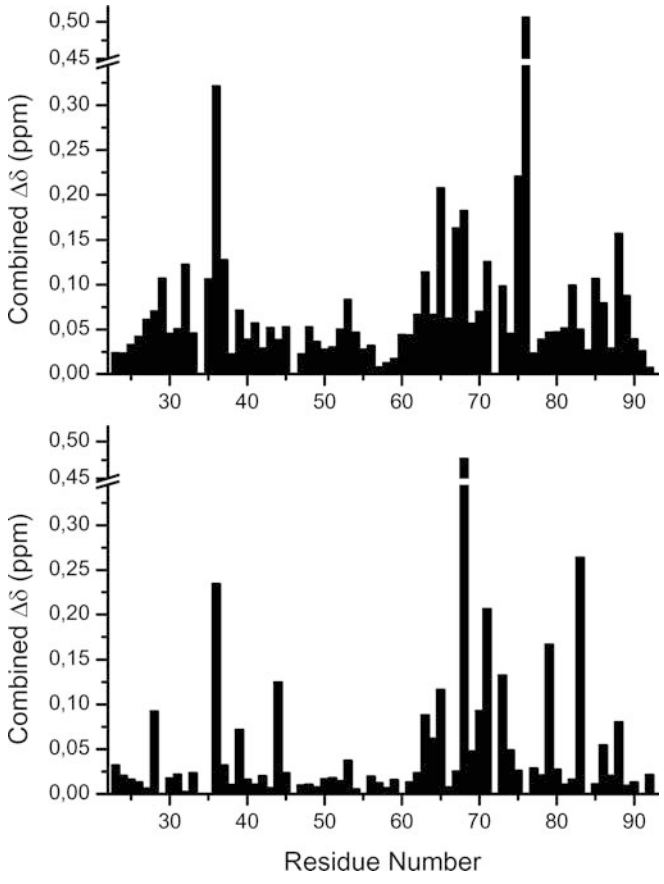


Fig. 2 Combined chemical shift variation with respect to WT *Bp*cyt*c* ($\Delta\delta$) [17] for backbone amide groups (*top*: I75A mutant; *bottom*: P72G mutant), derived from ^{1}H–^{15}N HSQC spectra

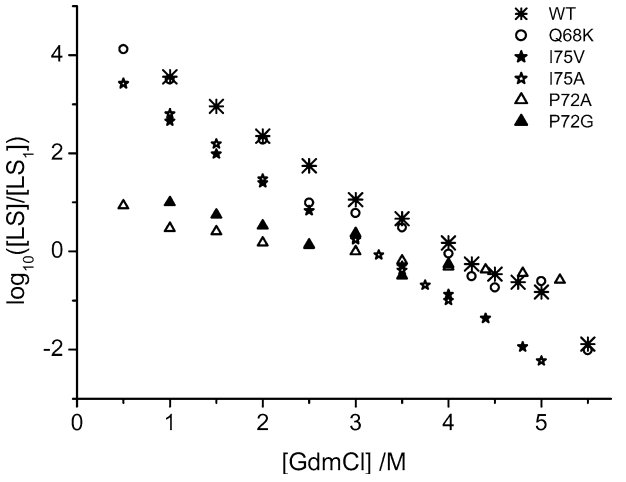


Fig. 3 Logarithm of the [LS]/[LS1] ratio estimated from CD spectra as a function of GdmCl concentration. *Asterisks*: WT; *open circles*: Q68K; *stars*: I75A (*open*) and I75V (*filled*); *triangles*: P72A (*open*) and P72G (*filled*). The samples were in 100 mM phosphate buffer, pH 7.0. The thermodynamic parameters resulting from the best fit of the data to a line are given in Table 1

Table 1 Thermodynamic parameters for denaturation of oxidized *Bp*cyt*c* induced by GdmCl. The parameters have been obtained from a best fit of the data in Fig. 3

	ΔG^{H2O} (kJ mol^{-1})	m (kJ mol^{-1} M^{-1})
WT	26.5 ± 0.68	−6.50 ± 0.17
Q68K	25.5 ± 0.90	−6.50 ± 0.38
I75A	22.6 ± 0.51	−6.99 ± 0.17
I75V	22.2 ± 0.23	−6.89 ± 0.10
P72A	4.90 ± 0.34	−1.60 ± 0.12
P72G	8.1 ± 1.2	−2.62 ± 0.45

thermodynamic parameters are given in Table 1. The results for the WT protein and the Q68K mutant are essentially the same, while the two I75 mutants are similar to one another but different from WT in that unfolding is more easily induced (i.e. lower ΔG^{H2O}). The two P72 mutants are strikingly more different with significantly lower ΔG^{H2O} and m. Analogous titrations by UV-Vis spectroscopy yield, within error, essentially the same thermodynamic parameters as the above-mentioned titrations by CD spectroscopy. CD and UV-Vis spectroscopy monitor different processes, as the first detects variations in the secondary structure of the protein, while the second is sensitive to variations in the coordination sphere of the heme iron. The good agreement between CD and UV-Vis data indicates that the loss of secondary structure and the variation of the heme coordination environment take place simultaneously, as already observed for the WT protein [2].

NMR titrations with GdmCl are in good agreement with the denaturation profiles obtained by CD and UV-vis spectroscopy. Figure 4 shows the NMR spectra at pH 7.0 and high (i.e. sufficient to induce > 50% denaturation) [GdmCl] for WT, I75A, P72G and P72A oxidized *Bp*cyt*c*. It appears that under denaturing conditions these four systems display the same pattern of

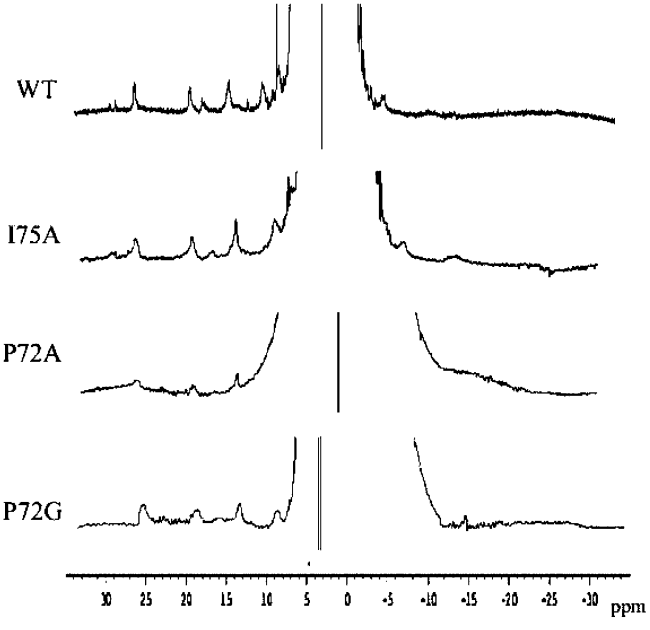


Fig. 4 Comparison of 1D NMR proton spectra at high [GdmCl]. The pH of all samples was 7.0

hyperfine shifted signals. Given the sensitivity of the hyperfine shift to the chemical environment of the iron(III) ion [18, 19, 20, 21, 22], this finding strongly suggests that the metal coordination under denaturing conditions is the same in the WT and mutant proteins.

Solution structure and solvation of the I75A mutant

NMR resonance assignments of the oxidized I75A mutant were performed with the same strategy used for both oxidized and reduced WT *Bp*cyt*c* [1, 3], yielding an assignment with a comparable degree of completeness (95% of all resonances of the I75A mutant have been assigned). The solution structure of I75A was obtained using 1010 meaningful upper distance limits, yielding an average target function of 0.6 ± 0.1 Å^2 within the family of the 20 best conformers. The structure is very similar to that of oxidized WT *Bp*cyt*c*, with an RMSD for the backbone atoms of only 0.7 Å (Fig. 5). In the present mutant, the loop containing residues 42–47 features extensive conformational mobility on the sub-microsecond time scale, which hinders observation of NOEs. Therefore, these residues have not been used for the superposition. All elements of secondary structure observed in the WT protein are also present in I75A (Fig. 5). Some differences are observed in the loop region containing the axial Met and the mutation.

Dipolar interactions between the redox-independent long-lived water molecule detected in WT *Bp*cyt*c* [4] and the amide protons of residues 74 and 75 are also detected in the P72A and I75A mutants. Consequently, it can be concluded that the water molecule is still present in the mutants (not shown). There are also no changes in the positions at which amide protons of the polypeptide chain are in fast exchange with the solvent, as detected by CLEANEX.

Dynamics of the I75A and P72A mutants

Figure 6 shows a per-residue comparison of the ^{15}N R_2/R_1 ratios for WT, P72A and I75A oxidized *Bp*cyt*c*. The

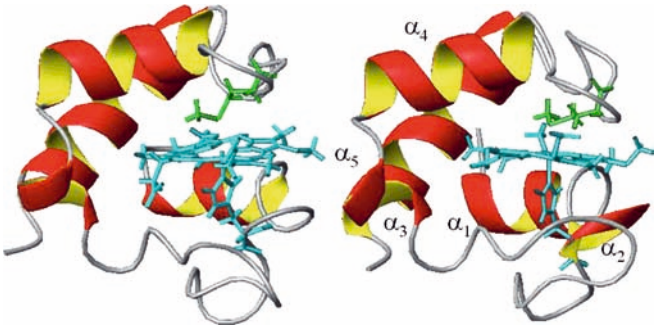


Fig. 5 Comparison of the solution structures of I75A (*left*) and WT (*right*) *Bp*cyt*c*, showing the main structural elements. The heme cofactor and the axial ligands are indicated: His36 and the heme in *cyan*, Met71 in *green*. This figure was prepared with the program MOLMOL [43]

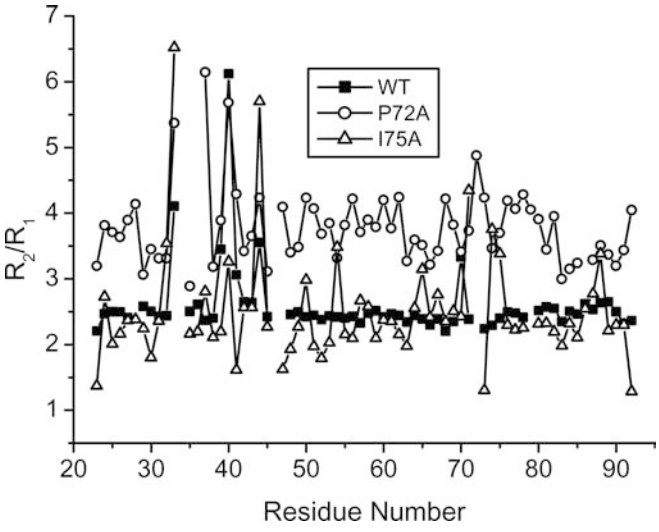


Fig. 6 Comparison of the R_2/R_1 ratios obtained from ^{15}N relaxation measurements for oxidized *Bp*cyt*c*. *Filled squares*: WT protein; *open triangles*: I75A mutant; *open circles*: P72A mutant

values for WT and I75A are comparable, and thus indicative of only small, local variations in the dynamics of the protein backbone. Such small variations can be highlighted by a detailed model-free [23] analysis of the data (not shown), and suggest enhanced mobility in the I75A mutant with respect to the WT protein in helix α_1 and the following loop, and in helix α_5. For the P72A mutant a strikingly different behavior is instead observed, with a generalized increase of the R_2/R_1 ratio over essentially the entire sequence. The present data for P72A are quite similar to the data obtained on WT *B*cyt*c* in the presence of 3.75 M GdmCl [2]. A possible indication from these measurements is thus that the P72A mutation makes largely unfolded conformations accessible also under native conditions, while the I75A mutation does not.

Backbone amide exchange rates can provide information on the processes of structural opening taking place on time scales of minutes and longer [24, 25]. The opening reactions are determined by local, segmental and global unfolding reactions [26, 27, 28]. By using a widely accepted model for amide proton exchange, it is possible to derive residue-specific equilibrium constants for the opening process [25]. Figure 7 shows the ratio between the exchange rates (R^{ex}) measured for the I75A and P72A mutants and the rates measured for the WT protein [5] (plotted on a logarithmic scale, so that the value is proportional to the $\Delta\Delta G$ of the opening reaction). For the I75A mutant, the largest systematic variations are observed for the C-terminal part of the terminal helix α5, which opens up more readily than for WT. Easier opening, but to a smaller extent than for helix α5, is also observed in I75A for residues 30–32, which are in between the first two helices. These patterns of variation are also observed for P72A, with a similar magnitude in the region 30–32, and with much larger magnitude in helix α5. In addition, the P72A mutant

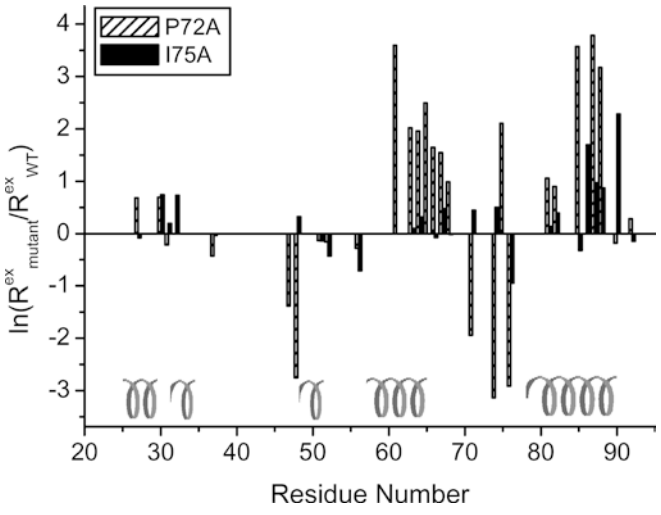


Fig. 7 Logarithmic plot of the ratio between backbone amide hydrogen exchange rates for mutant and WT oxidized *Bp*cyt*c*. *Hatched bars*: P72A; *filled bars*: I75A. The secondary structure is schematically represented

features strongly enhanced exchange (i.e. equilibrium shifted towards the open conformation) also in the region of helix α4 (residues 60–66).

Note that helix α5 exchanges through global opening [5], so the exchange reaction for the residues in this region is expected to become faster the lower the ΔG^{H2O} of the system, as it is indeed observed (compare Fig. 7 and Table 1). Helix α4 is an independent opening unit, with local folding/unfolding kinetics different from that of the neighboring residues and the C-terminal helix [5]. In mitochondrial cytochrome *c*, helix α4 has been shown to constitute an independent folding unit along the folding pathway [26]. The differential effect of mutations at position 72 or 75 on the amide exchange properties in this region could reflect a variation in the interaction between the different folding units. It is also relevant that the long-lived internal water molecule of *Bp*cyt*c* forms hydrogen bonds with both Pro72 and Ile64, and this may provide a pathway by which a mutation at position 72 can affect the 60–65 region.

Reduction thermodynamics

*Bp*cytc yields a one-electron, quasi-reversible voltammetric wave (CV) on the gold modified electrode, due to the Fe^{3+}/Fe^{2+} equilibrium of the heme iron [29, 30, 31, 32]. Peak separation in CV experiments varied from 70 to 120 mV within the range of scan rates investigated. Anodic and cathodic peak currents were found to be almost identical, and both were proportional to the protein concentration and $v^{1/2}$ (v = scan rate), indicating a diffusion-controlled electrochemical process. Given the quasi-reversibility of the electrochemical process, the symmetrical shape of the voltammograms and the almost negligible influence of the scan rate on the half-wave potential, the $E_{1/2}$ values can be confidently assumed to represent the $E^{\circ\prime}$ values. Analogous results were obtained for the mutants investigated. The temperature dependence of $E^{\circ\prime}$ for the WT protein and the P72A and I75 mutant proteins was measured from 10 to 40 °C (Fig. 8). However, at 40 °C the signal for the P72A mutant deteriorated so much as to make the data unreliable. The reduction entropies and enthalpies (the latter from the $E^{\circ\prime}/T$ vs. $1/T$) determined from these plots turn out to be the same within experimental error except for the P72A mutant (Table 2).

All *Bp*cyt*c* variants show negative reduction enthalpies and entropies, as is typical for class I cytochromes*c* [29, 30, 31, 32]. The negative enthalpy change is mainly the result of the selective stabilization of the reduced over the oxidized state due to ligand binding interactions (especially the axial methionine ligation), the hydrophobic heme environment and the limited accessibility of the heme to the solvent. The entropy loss on reduction is predominantly due to solvent reorganization effects within the hydration sphere of the molecule, together with the decrease in conformational (vibrational, torsional) degrees of freedom of the polypeptide chain indicated by a number of NMR solution structures [3, 6, 33, 34, 35].

The reduction potential of WT *Bp*cyt*c* (+68 mV) is less positive than that of mitochondrial and other bacterial cytochromes (namely cytochromes c_2 from *Rhodopseudomonas palustris*, *Rhodobacter sphaeroides* and *Rhodobacter capsulatus*), which feature $E^{\circ\prime}$ values of approximately +260 mV and + 350 mV, respectively [29, 30, 31, 32]. From Table 2 and the above-mentioned published data it is apparent that such a decrease in $E^{\circ\prime}$ is totally enthalpic in origin. This effect can be rather confidently ascribed to the much higher solvent accessibility of the heme in *Bp*cyt*c* as compared with the above species (the solvent exposition of the heme varies from 135 to 160 Å^2 [1, 7] against less than 50–60 Å^2 in mitochondrial cytochrome *c*), which favors the ferriheme (bearing a net charge of +1 against the null net

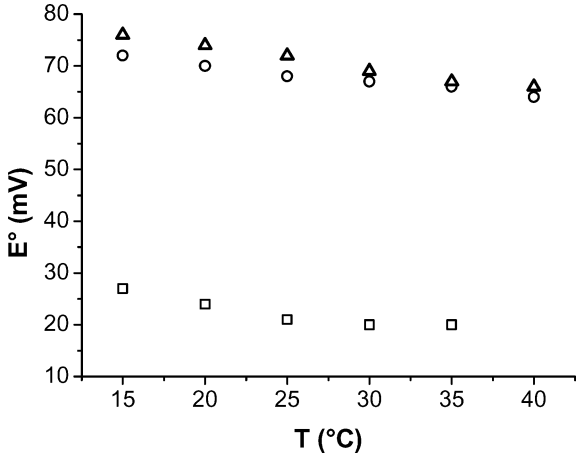


Fig. 8 Temperature dependence of the reduction potential of *Bp*cyt*c*. WT protein (*triangles*); P72A (*squares*); I75A mutant (*circles*). Base electrolyte, 100 mM NaCl, 10 mM phosphate buffer, pH 7

Table 2 Thermodynamic parameters for the reduction of the soluble fragment of cytochrome c_{553} from WT *B. pasteurii* and its I75A mutant[a]

Species	$E^{\circ\prime}$[b] (mV)	$\Delta S^{\circ\prime}_{rc}$ (J mol^{-1} K^{-1})	$\Delta H^{\circ\prime}_{rc}$ (kJ mol^{-1})	$T\Delta S^{\circ\prime}_{rc}/F$[b,c] (mV)	$-\Delta H^{\circ\prime}_{rc}/F$[b,c] (mV)
WT *Bp*cytc	+68	−28	−15	−85	+155
Q68K	+65	−26	−14	−78	+146
P72A	+21	−36	−13	−113	+134
I75A	+72	−30	−16	−93	+164

[a]All values obtained in 100 mM NaCl, 10 mM phosphate buffer. Average errors on $\Delta H^{\circ\prime}_{rc}$ and $\Delta S^{\circ\prime}_{rc}$ values are ±2 kJ mol^{-1}and ±6 J mol^{-1} K^{-1}, respectively

[b]The reduction potentials are measured at 25 °C and are referred to SHE

[c]$-\Delta H^{\circ\prime}_{rc}/F$ and $T\Delta S^{\circ\prime}_{rc}/F$ often do not sum up exactly to $E^{\circ\prime}$ since, because of the experimental error, the $\Delta H^{\circ\prime}_{rc}$ and $\Delta S^{\circ\prime}_{rc}$ values are rounded to the closest integer

charge of the ferrous form). The increased solvent exposure of the heme in the present species is consistent with the less negative $\Delta S^{\circ\prime}_{rc}$values featured by *Bp*cyt*c* as compared with other class I cytochromes *c* (in the range −40 to −60 J K^{-1} mol^{-1}). In fact, the greater the exposure of the heme to the solvent, the more effective the mechanism that leads to a decreased ordering of the water molecules in the surroundings of the heme following the one unit decrease in the positive heme charge upon reduction. Within the above context, the data of Table 2 for the P72A mutant can be interpreted as due to an enhanced solvent exposure in this system with respect to WT *Bp*cyt*c*.

Discussion

An interesting characteristics of *Bp*cyt*c* is that it does not present any pH-dependent variation in iron(III) ligation over the pH range 2–13 [2, 8]. In particular, it does not experience the so-called alkaline transition, observed for mitochondrial *c*-type cytochromes as well as cytochromes c_2, in which the native Met ligand is replaced by another endogenous iron(III) with apparent pK_a values in the range 8.5–9.5. In practice, an equilibrium is obtained involving different alkaline forms with different alternate ligands. Alkaline isomerization is facilitated by chemical denaturants and/or elevated temperatures, thereby suggesting a connection with the mechanism of folding/unfolding. This correlation has been directly demonstrated through kinetic studies based on native-state hydrogen exchange methods [36]. The latter study has indeed shown that the events leading to Met replacement in mitochondrial cytochrome *c* are the same as the early steps in the cooperative protein unfolding reaction (induced by denaturant) [36, 37]. This result is of particular interest also because mitochondrial cytochrome *c* and *Bp*cyt*c* share the same organization of folding units, thus suggesting that similar reaction mechanisms could be operative [5].

At first glance, the absence of an alkaline transition in *Bp*cyt*c* could be attributed to the absence of a suitable alternate ligand (typically a Lys side chain) in the proximity of the metal site. In this work, the Q68K mutant was produced, which introduces a Lys well within reach of the iron(III) ion, in a position corresponding to that of an alternate ligand in the alkaline form(s) of mitochondrial cytochrome *c*. Q68 is the N-terminal residue of the Met loop in *Bp*cyt*c*. However, the properties of the Q68K mutant are indistinguishable from that of WT *Bp*cyt*c* both at alkaline pH and in the presence of GdmCl, thereby excluding that the mere presence of an alternate ligand [even though in principle better suited than Met for binding an iron(III) ion] makes cytochrome *c* structural rearrangements readier.

The alkaline transition in mitochondrial cytochrome *c* starts with the unfolding of the so-called nested yellow loop [36] (residues 40–57 in mitochondrial cytochrome *c*). It has been proposed that this event allows a proton acceptor to enter into the protein core, leading to deprotonation of an internal group. This first step is needed for the subsequent replacement of the axial ligand. The above loop is absent altogether in the shorter *Bp*cyt*c* sequence. This should favor, at least kinetically, the alkaline transition. As this is not observed, it can be postulated that in *Bp*cyt*c*there are some specific factors preventing the structure change necessary to physically remove the native Met71 ligand from the iron(III) ion and bring in the alternate ligand. In mitochondrial cytochrome *c* this structure change has been identified with unfolding of the Met loop, both from equilibrium structural data [38] and from kinetic measurements [36].

In *Bp*cyt*c*, two possible determinants for extra stabilization of the Met loop (residues 68–76) could be its peculiar amino acid composition and/or the presence a long-lived water molecule, which plays a central role in a hydrogen-bond network connecting different regions of the *Bp*cyt*c* structure through the involvement of Ile64, Pro72 and Ile75 [4]. With respect to the first point, Pro72 is of particular interest, as it is strictly conserved in all *Bp*cyt*c* homologs [1], whereas in mitochondrial *c*-type cytochromes the corresponding residue is generally Ile, Val or Ala [39]. With respect to the long-lived water molecule, its close contacts with the hydrophobic side chain of Ile75 could play a significant role in determining the energetics of folding. Mutation of Pro72 could also in principle have an effect on this water molecule.

The I75V and I75A mutants show a reduced stability towards GdmCl with respect to the WT protein (Table 1), but direct measurements clearly show that the equilibrium position of the long-lived water is essentially unaffected. It is thus more reasonable to assume that this decreased stability is due to the removal of a number of long-range hydrophobic contacts with the axial ligand Met71 as well as with the rest of the polypeptide chain. This is also in qualitative agreement with the observed enhancement of protein mobility in helix α_1 and the following loop, and in helix α_5. Note that in the I75V mutant only the contact with Met71 should be removed; however, it cannot be ruled out that, because of the space available to the smaller Val side chain, also the other contacts of residue 75 are loosened. Indeed, the similar distribution of hyperfine-shifted signals observed for the two I75 mutants strongly suggests that the conformation of the axial Met71 should be the same in both systems. This is because the lone pair of the Sδ atom of the axial Met appears to be most important in determining the in-plane directions of the magnetic susceptibility axes for the iron(III) ion [40], and therefore even slight structural variations in the conformation of the Met side chain are expected to result in significant changes of the hyperfine shifts.

The most striking effect on protein stability occurs upon mutation of Pro72 (Fig. 3). Mutation of Pro72 does not affect significantly the 1D ^{1}H spectrum (Fig. 1), indicating that in the absence of denaturant the conformation at the iron(III) site is unchanged with respect of WT. As mentioned, the small variation of the chemical shifts of backbone amide groups excludes major rearrangements of the polypeptide chain conformation, as was also the case for the I75A mutant. CLEANEX experiments show that also the long-lived water molecule is unaffected by the P72A mutation. Prolines have been reported to contribute favorably to the stability of the native fold of proteins in some specific cases where they occur in rigid loops. This is most likely accomplished by reducing the entropic cost of the restriction of the conformational space available to the polypeptide chain that occurs upon protein folding [41, 42]. In the present system the Met71 loop must adopt a well-defined and relatively rigid conformation to allow proper metal ion coordination, and thus the presence of a Pro in the loop should be beneficial to stabilize the protein fold. In this context, it is also important that the conformation of Pro72 in *Bp*cyt*c* is within the allowed conformational space of Gly and Ala. At high [GdmCl], the iron(III) coordination appears to be the same in the Pro72 mutants as for WT. However, we do not expect a significant effect of Pro72 on metal coordination in the denatured state, also because replacement of Met71 by a Lys residue requires significant unfolding of the entire polypeptide chain. Denaturation of the two Pro72 mutants is characterized also by a somewhat reduced *m* value (Table 1). Backbone dynamics data show quite clearly that the P72A mutant in the native state can access largely unfolded conformations much more easily than the WT or I75A mutant. The ^{15}N relaxation data of Fig. 6 in fact indicate that the existing conformational exchange events are not localized but involve the entire polypeptide chain. Backbone amide exchange data (Fig. 7) show that exchange rates are enhanced in the two most slow-exchanging regions, i.e. helix α_4 and α_5, despite the fact that in the WT protein these two regions do not open coordinately for exchange [5]. However, this is in agreement with the idea that the P72A significantly samples largely unfolded conformations under native conditions, thereby strongly enhancing the contribution of global unfolding to the amide exchange reaction at all protein amide moieties. This picture is also in agreement with electrochemical measurements on P72A, which suggest that the heme is more solvent accessible than in WT *Bp*cyt*c*.

The low *m* value of the P72A mutant can thus be rationalized on the same basis used above to interpret the backbone dynamics data: if the protein samples extended conformations already under native conditions, which does not hold for WT *Bp*cyt*c*, the variation in solvent accessibility of the polypeptide chain upon complete denaturation will be somewhat lower for P72A than for WT *Bp*cyt*c*. In turn, this results in a somewhat lower *m*, as observed.

Conclusions

In summary, we have investigated the loop containing the iron axial ligand Met71 in oxidized *Bp*cyt*c* using site-directed mutagenesis of different positions, in order to obtain better insight into the determinants of the high stability of the axial coordination towards GdmCl and alkaline pH. It appears that one of the main determinants of this stability is the high energy barrier for the rearrangement of the Met loop conformation, as indicated by the fact that a mutant (Q68K) where a Lys has been introduced at the N-terminus of the Met loop is as stable as the WT protein. The Met loop rearrangement is a necessary step for replacement of the axial ligand [36, 38]. Met71 is in close contact with Ile75, which is important to tune the conformation of the side chain of the former residue. Mutation of Ile75 with less bulky amino acids results in a subtle rearrangement of the Met71 side chain and a small but significant destabilization of the protein fold. A key role is instead played by Pro72, which is conserved also in *Bp*cyt*c* homologs. Its mutation in fact dramatically affects the stability of *Bp*cyt*c* towards denaturants and makes largely unfolded conformations much more readily accessible even under native conditions.

The present results demonstrate that the amino acid composition of the axial Met loop and the corresponding network of inter-residue interactions are crucial for the stability of *c*-type cytochromes through a complex interplay of conformational propensities, hydrogen bonds (also with the solvent) and hydrophobic contacts.

Acknowledgements Financial support from MIUR is gratefully acknowledged. P.R.V. thanks the European Commission for financial support under contract number HPRN-CT-2000-00092.

References

1. Banci L, Bertini I, Ciurli S, Dikiy A, Dittmer J, Rosato A, Sciara G, Thompsett A (2002) ChemBioChem 3:299–310
2. Bartalesi I, Bertini I, Ghosh K, Rosato A, Turano P (2002) J Mol Biol 321:693–701
3. Bartalesi I, Bertini I, Rosato A (2003) Biochemistry 42:739–745
4. Bertini I, Ghosh K, Rosato A, Vasos P (2003) Biochemistry 42:3457–3463
5. Bartalesi I, Rosato A, Zhang W (2003) Biochemistry 42:10923–10930
6. Banci L, Bertini I, Luchinat C, Turano P (2000) In: Kadish KM, Smith KM, Guilard R (eds) The porphyrin handbook. Academic Press, San Diego, pp 323–350
7. Benini S, Rypniewski W, Wilson KS, Van Beeumen J, Ciurli S (2000) Biochemistry 39:13115–13126
8. Benini S, Borsari M, Ciurli S, Dikiy A, Lamborghini M (1998) J Biol Inorg Chem 3:371–382
9. Banci L, Bertini I, Luchinat C, Piccioli M, Scozzafava A, Turano P (1989) Inorg Chem 28:4650–4656
10. Güntert P, Mumenthaler C, Wüthrich K (1997) J Mol Biol 273:283–298
11. Herrmann T, Güntert P, Wüthrich K (2002) J Mol Biol 319:209–227
12. Pace CN (1986) Methods Enzymol 131:266–280
13. Myers JK, Pace CN, Scholtz JM (1995) Protein Sci 4:2138–2148
14. Yee EL, Cave RJ, Guyer KL, Tyma PD, Weaver MJ (1979) J Am Chem Soc 101:1131–1137
15. Taniguchi VT, Sailasuta-Scott N, Anson FC, Gray HB (1980) Pure Appl Chem 52:2275–2281
16. Koller KB, Hawkridge FM (1985) J Am Chem Soc 107:7412–7417
17. Grzesiek S, Bax A, Clore GM, Gronenborn AM, Hu J-S, Kaufman J, Palmer I, Stahl S, Wingfield P (1996) Nat Struct Biol 3:340–345
18. Turner DL, Williams RJP (1993) Eur J Biochem 211:555–562
19. Santos H, Turner DL (1992) Eur J Biochem 206:721–728
20. Bertini I, Luchinat C, Parigi G, Walker FA (1999) J Biol Inorg Chem 4:515–519
21. Banci L, Bertini I, Luchinat C, Pierattelli R, Shokhirev NV, Walker FA (1998) J Am Chem Soc 120:8472–8479
22. Bertini I, Luchinat C, Parigi G (2000) Eur J Inorg Chem 2473–2480
23. Lipari G, Szabo A (1982) J Am Chem Soc 104:4546–4559
24. Bai Y, Milne JS, Mayne L, Englander SW (1994) Proteins Struct Funct Genet 20:4–14
25. Huyghues-Despointes BM, Pace CN, Englander SW, Scholtz JM (2001) Methods Mol Biol 168:69–92
26. Bai YW, Sosnick TR, Mayne L, Englander SW (1995) Science 269:192–197
27. Milne JS, Mayne L, Roder H, Wand AJ, Englander SW (1998) Protein Sci 7:739–745
28. Maity H, Lim WK, Rumbley JN, Englander SW (2003) Protein Sci 12:153–160
29. Battistuzzi G, Borsari M, Sola M, Francia F (1997) Biochemistry 36:16247–16258
30. Battistuzzi G, Borsari M, Cowan JA, Eicken C, Loschi L, Sola M (1999) Biochemistry 38:5553–5562
31. Battistuzzi G, Borsari M, Sola M (2001) Eur J Inorg Chem 2989–3004
32. Battistuzzi G, Borsari M, Cowan JA, Ranieri GA, Sola M (2002) J Am Chem Soc 124:5315–5324
33. Baistrocchi P, Banci L, Bertini I, Turano P, Bren KL, Gray HB (1996) Biochemistry 35:13788–13796
34. Banci L, Bertini I, Bren KL, Gray HB, Sompornpisut P, Turano P (1997) Biochemistry 36:8992–9001
35. Banci L, Bertini I, Gray HB, Luchinat C, Reddig T, Rosato A, Turano P (1997) Biochemistry 36:9867–9877
36. Hoang L, Maity H, Krishna MMG, Lin Y, Englander SW (2003) J Mol Biol 331:37–43
37. Tezcan FA, Findley WM, Crane BR, Ross SA, Lyubovitsky JG, Gray HB, Winkler JR (2002) Proc Natl Acad Sci USA 99:8626–8630
38. Assfalg M, Bertini I, Dolfi A, Turano P, Mauk AG, Rosell FI, Gray HB (2003) J Am Chem Soc 125:2913–2922
39. Banci L, Bertini I, Rosato A, Varani G (1999) J Biol Inorg Chem 4:824–837
40. Bertini I, Faraone-Mennella J, Gray BH, Luchinat C, Parigi G, Winkler JR (2004) J Biol Inorg Chem 9:224–230
41. Watanabe K, Chishiro K, Kitamura K, Suzuki Y (1991) J Biol Chem 266:24287–24294
42. Querol E, Perez-Pons JA, Mozo-Villarias A (1996) Protein Eng 9:265–271
43. Koradi R, Billeter M, Wüthrich K (1996) J Mol Graphics 14:51–55

B. The N-terminal domain of the copper(I) transporting ATPase PacS from *Synechocystis PCC 6803*

B.1 Overview

Copper provides a challenge to biological systems. It is indeed essential to all organisms, but it's useful redox properties, become potentially toxic when copper is present in excessive amounts. The ability of Cu to cycle between a stable oxidised, Cu(II), and unstable reduced, Cu(I), states is used by cuproenzymes involved in redox reactions, e.g. Cu/Zn superoxide dismutase and cytochrome oxidase. However, the Cu (II) $\leftrightarrow$ Cu (I) transitions can in certain circumstances also result in the generation of reactive oxygen species, e.g. superoxide radical and hydroxyl radical, which, if not detoxified efficiently, can damage susceptible cellular components. Copper can also bind with high affinity to histidine, cysteine and methionine residues of proteins which can result in their inactivation. The need to provide Cu to essential enzymes without ensuing cellular toxicity has necessitated evolution of tightly regulated Cu homeostatic mechanisms some of which have been identified only recently. Maintenance of non-toxic but physiologically essential intracellular concentrations of Cu can be achieved through a Cu homeostasis system that involves regulated Cu uptake, vectorial intracellular transport, Cu sequestration/storage and regulated efflux. While these principles can be relatively well assimilated within a single cell system (bacteria and yeast), the regulation of Cu homeostasis in multicellular organisms or humans appears to be more complex, primarily because the efflux of Cu is not only essential for detoxification but is also vital for intercellular Cu transport. This complexity was recognized only recently, when novel mechanisms and pathways of Cu regulation were identified. Although the

characterization in more evoluted organisms is far from complete, the investigation of copper homeostasis in simple organisms is largely improving the comprehension of copper trafficking at molecular level.

B.1.1 Cyanobacteria command an evolutionary transition in metal availability

Cyanobacteria altered the solubility of metals on a global scale as a by-product of dioxygen-evolving photosynthesis. Anaerobic earth was 'polluted' with dioxygen some 2.7×10^9 years ago according to broadly accepted theory [1]. This closely correlates with the appearance of cyanobacteria and hence dioxygen-generating photosynthesis. *Synechocystis* PCC 6803 genome was among the first to be fully sequenced and has been followed by several other cyanobacteria: *Anabaena* PCC 7120, *Prochlorococcus marinus* MED4, *Synechococcus* WH 8102 and *Thermosynechococcus elongatus* BP-1. At least five more are in progress and nearing completion at the time of writing: *Gloeobacter violaceus* PCC 7421, *Nostoc punctiforme, Trichodesmium erythraem, Synechococcus* PCC 7002, *Synechococcus* PCC 6301 and *Synechococcus* PCC 7942 (Fig.11).

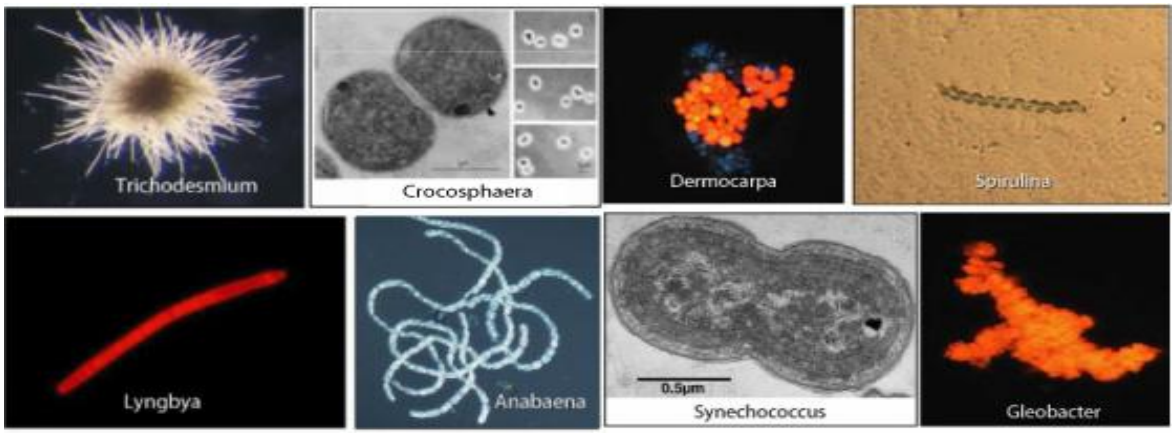


Figure 11. Various species of cyanobacteria

In cynobacteria the existence of thylakoids creates special metal-trafficking requirements peculiar to these bacteria. For example, four atoms of manganese are required to generate the holo-form of the water-splitting oxygen-evolving complex associated with the luminal face of PSII (Fig.12). Manganese acquisition in *Synechocystis* PCC 6803 involves MntABC, an ABC-type permease that mediates high-affinity transport under starvation conditions [2], a second high-affinity transporter for manganese-sufficient conditions [3] and a low-affinity transporter, the latter being indirectly observed from transport kinetics but the proteins involved are currently unknown. Manganese uptake is controlled via a two-component signal transduction pathway [4].

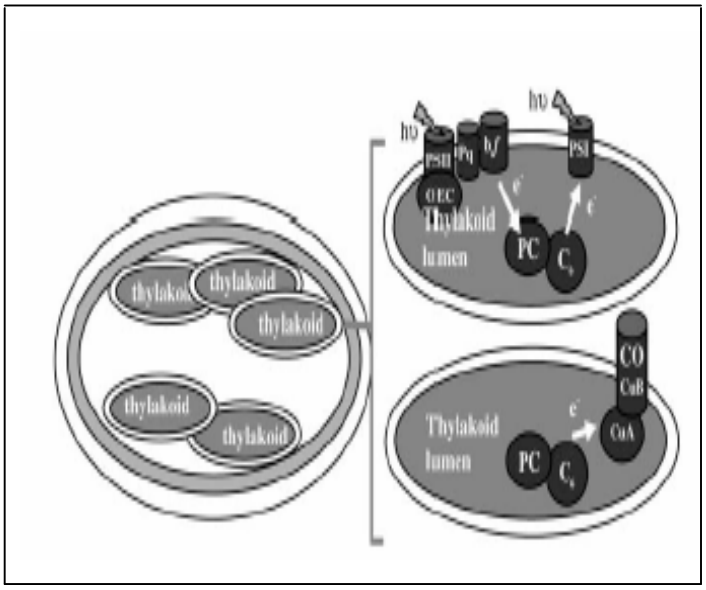


Figure 12. A simplistic representation of photosynthetic and respiratory electron transport in *Synechocystis* PCC 6803.

It is feasible that trafficking of the manganese cofactor is a late step in assembly of the photosynthetic electron transport machinery, which dictates the site of activity (thylakoidal versus plasma membrane). The thylakoids of *Synechocystis* PCC 6803 also contain two copper proteins, plastocyanin and a caa_3-type cytochrome oxidase, raising similar questions about how this metal cofactor is supplied to these enzymes, and the extent to which copper trafficking determines/controls their sites of activity.

B.1.2 Cyanobacterial Metallochaperone

Synechocystis PCC 6803 mutants analysis showed that Atx1 (ScAtx1) is required for normal photosynthetic electron transfer via plastocyanin and for the activity of a second thylakoid-located copper protein, a caa_3-type cytochrome oxidase (5, 6). In common with related polypeptides from other bacteria, yeast, and man (7-10), ScAtx1 directly interacts with soluble amino-terminal domains of P_1-type copper ATPases (114). However, unlike other copper metallochaperones, ScAtx1 from *Synechocystis* PCC 6803 associates with two such proteins, and there is a presumption that the vectors for copper transfer alternate in each of these two interactions. In *Synechococcus* PCC 7942, PacS is located in thylakoid membranes (11), whereas CtaA is thought to import copper at the plasma membrane (12). The phenotypes of Δ*ctaA* and Δ*PacS* mutants of *Synechocystis* PCC 6803 are consistent with both ATPases transporting copper in an inward direction into the cytosol and then into the thylakoid lumen (6). This provides an attractive system for studying the process of copper transfer between a copper metallochaperone and its partners.

Solution structure of ScAtx1 (Fig.13) shows that the overall folding of the copper-free and the copper-bound forms of ScAtx1 are almost identical (13). In Cu(I)ScAtx1 the copper(I) is coordinated by two cystein residues (in loop 1) and one His (C-terminus). Localized surface charge suggests how ScAtx1 may recognize soluble domains of copper ATPases but avoid a zinc ATPase (Fig.14) (13). The resulting specificities of these partnerships, rather than the inherent metal-binding preferences of each N-terminal domains may select which metals become available for transport. Metal-ligand arrangements of Cu(I)ScAtx1 coupled with observed backbone motions provide insight into how copper ions can pass from the plasma membrane to an intracellular compartment along a ligand gradient, without release into the cytosol (7) (Fig.15). Cys-12 and -15 are part of an α -helix in Cu (I) ScAtx1 but not in apoScAtx1. Characterization of Cu(I)ScAtx1 helped to point out that it forms a stable homodimer where the two subunits are linked through coordination bonds between the copper(I) ions and one bridge ligand (7).

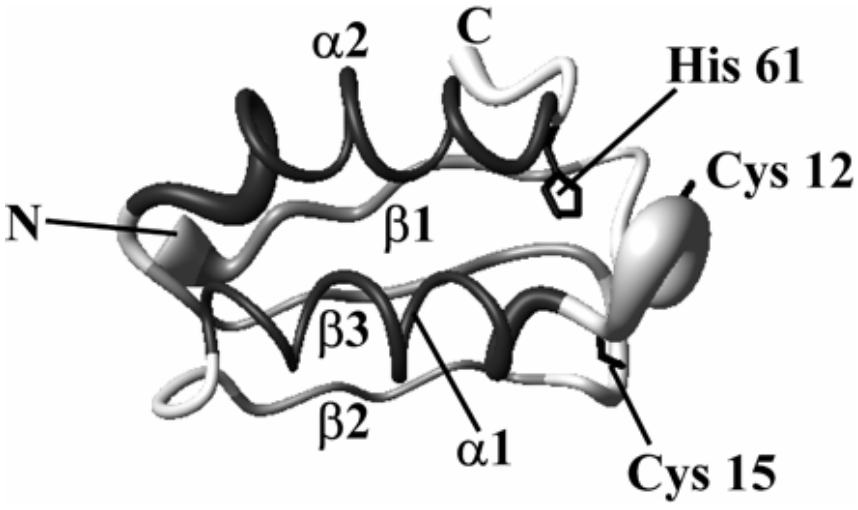


Figure 13. Solution structure of apoScAtx1; the side chains of His-61, Cys-12, and Cys-15 are shown.

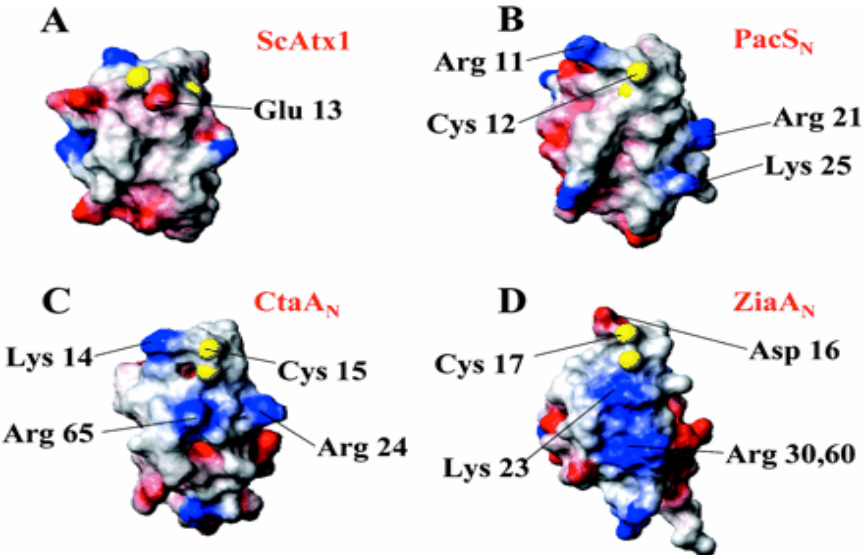


Figure 14. Electrostatic potential surface of ScAtx1 (*A*) and, of the structural models of PacSN$_N$ (*B*), CtaA$_N$ (*C*), and ZiaA$_N$ (*D*)

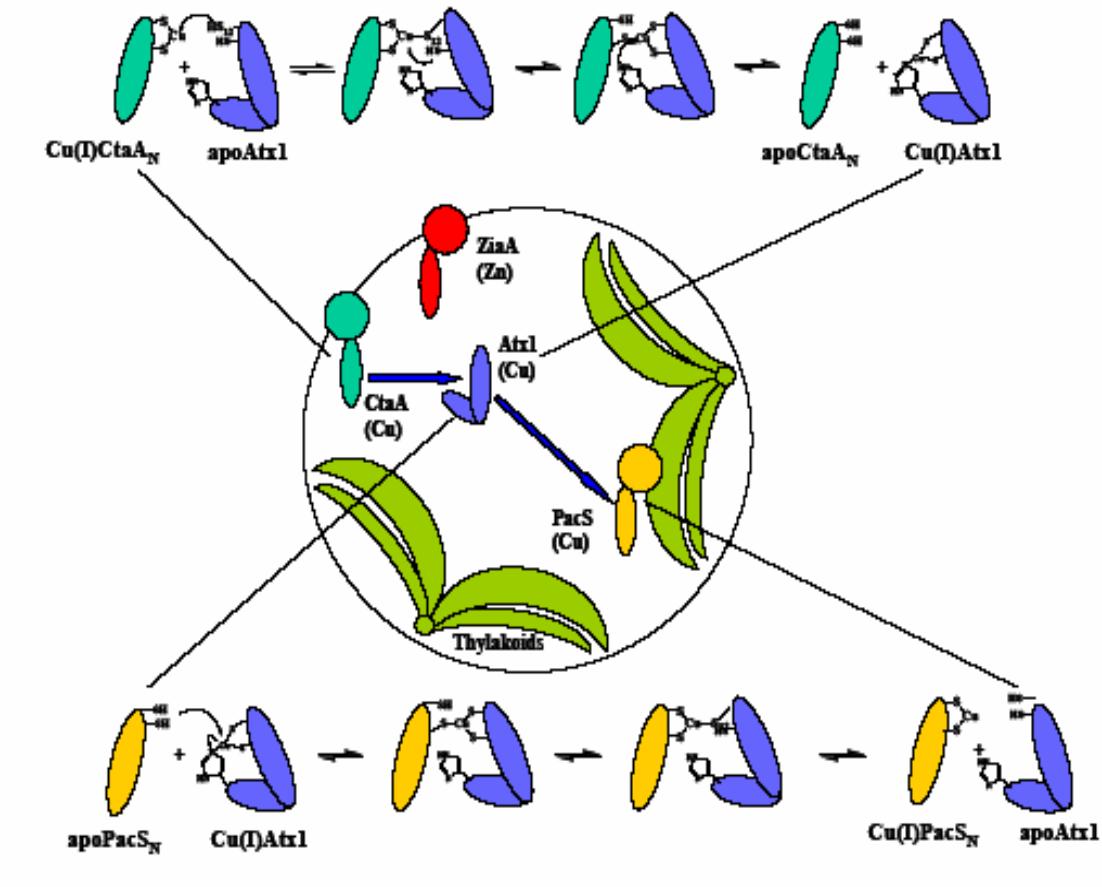


Figure 15. Proposed pathway for copper transfer from CtaA to PacS through ScAtx1.

The functional relevance of dimerization remains to be addressed, although the apo form is clearly monomeric, and it is speculated that dimerization may only occur at high concentrations *in vitro*. Structural data, showed that a third donor atom belonging to the N-terminus of P1-type of ATPases enters the $Cu(I)S_2$ coordination sphere of various eukaryotic and bacterial copper(I) chaperones in order to obtain an efficient transfer of the metal ion (17). This is consistent with digonal copper (I) tending to expand its coordination number. Most intriguing, a model of the soluble parts of $PacS_N$ does not display an evident extensive electrostatic complementary surface with ScAtx1 (Fig.14). It has been hypothesized that His-61 in ScAtx1 may influence the directionality of copper(I) transfer with respect to the two different P_1-type ATPases, PacS and CtaA (13). Greater structural flexibility of loop 1 suggests that one or other Cys residue of apoScAtx1 is more likely to first invade the copper-loaded site of $CtaA_N$, as proposed for Atx1 and Ccc2a from *S. cerevisiae* (14), rather than His-61. Subsequently, entrance of His-61 into the copper(I) site will nonetheless promote copper(I) acquisition by the copper metallochaperone with the concomitant displacement of CtaA. His-61 can render the copper(I) coordination environment more favorable in ScAtx1 than in CtaA because of formation of a coordination bond with $N^{\varepsilon 2}$ of His-61.

B.1.3 How is this process reversed upon interaction with $PacS_N$?

When Cu(I)ScAtx1 approaches apo$PacS_N$, it is speculated that His-61 is somehow displaced from the coordination sphere allowing one of the two cysteine residues of $PacS_N$ to encroach into a shared copper(I) site as proposed for Atx1/Ccc2a. Thus the favored order of copper (I) association with ScAtx1 ligands is 1) Cys-12, 2) Cys-15, and

finally 3) His-61 at acquisition from CtaA with the converse order of release at interaction with $PacS_N$. Localized negative charge from Glu-13, close to the metal-binding Cys residues of ScAtx1, could support interaction with complementary Arg and Lys in equivalent locations in $PacS_N$ and CtaA (Fig. 15). Lys/Arg residues are only adjacent to these Cys in metal-transporting P_1-type ATPases of *Synechococcus* PCC 7942, *Synechocystis* PCC 6803, and *Caulobacter crescentus* (16). So trafficking via sequential ligand exchange (Fig.15) will prevent copper from entering thermodynamic traps, adventitious high affinity copper (I) sites, while *en route* to thylakoids. Such sites are likely to include ones that should be occupied by other metals but that nonetheless have tighter affinity for copper (I).

In this work cloning, expression and structural characterization of N-terminal $PacS_N$ from *Synechocystis* PCC 6803 was performed. In addition, interaction studies between ScAtx1 metallochaperone and $PacS_N$ was also performed to understand the role of His-61 in the copper transfer mechanism at molecular level.

B.1.4 References

1. Frausto da Silva, J.J.R. and Williams, R.J.P. (2001) The Biological Chemistry of the Elements: The Inorganic Chemistry of Life, 2nd edn. Clarendon Press, Oxford.
2. Pakrasi, H.B. (2001) *Proc. Natl. Acad. Sci.USA* 98, 13443-13448.
3. Bartsevich, V.V. and Pakrasi, H.B. (1995) *EMBO J.* 14, 1845-1853.
4. Bartsevich, V.V. and Pakrasi, H.B. (1996) *J. Biol. Chem.* 271, 26057-26061.
5. Tottey, S., Rich, P. R., Rondet, S. A. M., and Robinson, N. J. (2001) *J. Biol.Chem.* **276,** 19999–20004.

6. Tottey, S., Rondet, S. A., Borrelly, G. P., Robinson, P. J., Rich, P. R., and Robinson, N. J. (2002) *J. Biol. Chem.* **277,** 5490–5497.

7. Banci, L., Bertini, I., Ciofi-Baffoni, S., Del Conte, R., and Gonnelli, L. (2003) *Biochemistry* **42,** 1939–1949.

8. Radford, D. S., Kihlken, M. A., Borrelly, G. P. M., Harwood, C. R.,Le Brun, N. E., and Cavet, J. S. (2003) *FEMS Microbiol. Lett.* **220,** 105–112.

9. Wernimont, A. K., Huffman, D. L., Lamb, A. L., O'Halloran, T. V., and Rosenzweig, A. C. (2000) *Nat. Struct. Biol.* **7,** 766–771.

10. Arnesano, F., Banci, L., Bertini, I., Cantini, F., Ciofi-Baffoni, S., Huffman, D. L., and O'Halloran, T. V. (2001) *J. Biol. Chem.* **276,** 41365–41376.

11. Kanamaru, K., Kashiwagi, S., and Mizuno, T. (1994) *Mol. Microbiol.* **13,**369–377.

12. Phung, L. T., Ajlani, G., and Haselkorn, R. (1994) *Proc. Natl. Acad. Sci.U. S. A.* **91,** 9651–9654.

13. Banci,L., Bertini,I.,Ciofi-Baffoni,S.,Cheng Su,X., Borrelly and Robinson,N.J (2004) *J. Biol. Chem.* **279,** 27502–27510.

14. Cavet, J. S., Borrelly, G. P., and Robinson, N. J. (2003) *FEMS Microbiol. Rev.* **27,** 165–181.

15. Pufahl, R., Singer, C. P., Peariso, K. L., Lin, S.-J., Schmidt, P. J., Fahrni, C. J., Cizewski Culotta, V., Penner-Hahn, J. E., and O'Halloran, T. V. (1997) *Science* **278,** 853–856.

16. Arnesano, F., Banci, L., Bertini, I., Ciofi-Baffoni, S., Molteni, E., Huffman, D. L., and O'Halloran, T. V. (2002) *Genome Res.* **12,** 255–271.

17. Tottey, S., Rich, P.R., Rondet, S.A. and Robinson, N.J. (2001) *J. Biol. Chem.* 276, 19999-20004.

B. 2 Cloning, Expression, Purification and Characterization of $PacS_N$

The gene encoding the N-terminal tail (1-95 amino acids) of $PacS_N$ was PCR amplified from *Synechocystis* PCC 6803 genomic DNA and digested with *EcoRI/NdeI* prior to ligation into pET29a vector (Novagen) to create $pETPacS_N$. Bacterial clones after ligation were screened by DNA sequencing and positive clone was transformed to BL21 (DE3) expression cells. Cells were grown in minimal medium and were induced at $O.D_{600}$ reaching 0.6 with 0.5 mM and 1 mM IPTG concentration. Protein was well expressed after 6 hrs (Fig.23) at 1 mM concentration of IPTG. The entire test expression was done in minifor fermentor by using minimal medium. Protein was purified under native condition by using anion exchange chromatography and size exclusion chromatography in 25 mM Tris buffer pH 9.0. Protein purity was checked through SDS-PAGE gel (Fig. 24)

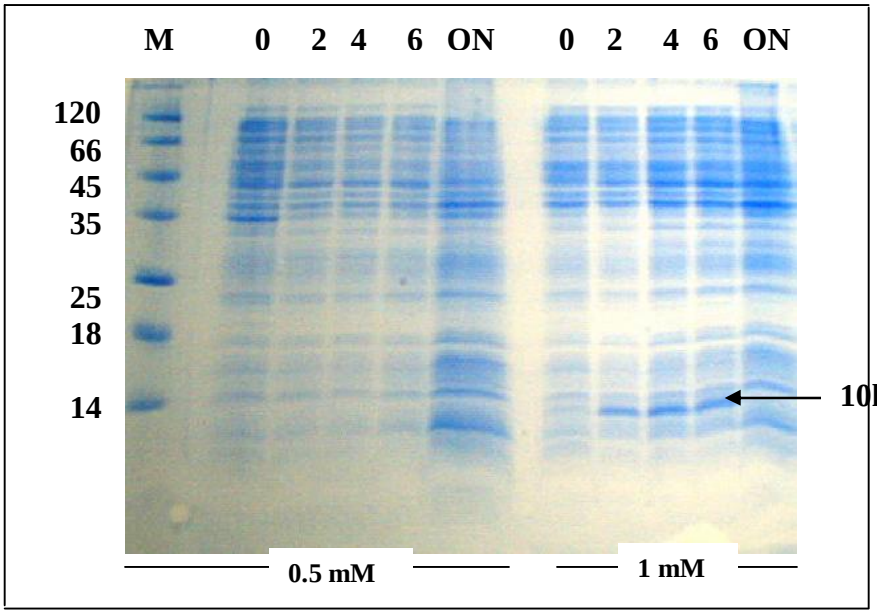


Figure 23.17% SDS-PAGE profile of over-expression of $PacS_N$ at different intervals of time in minimal medium (hrs), M-Marker.

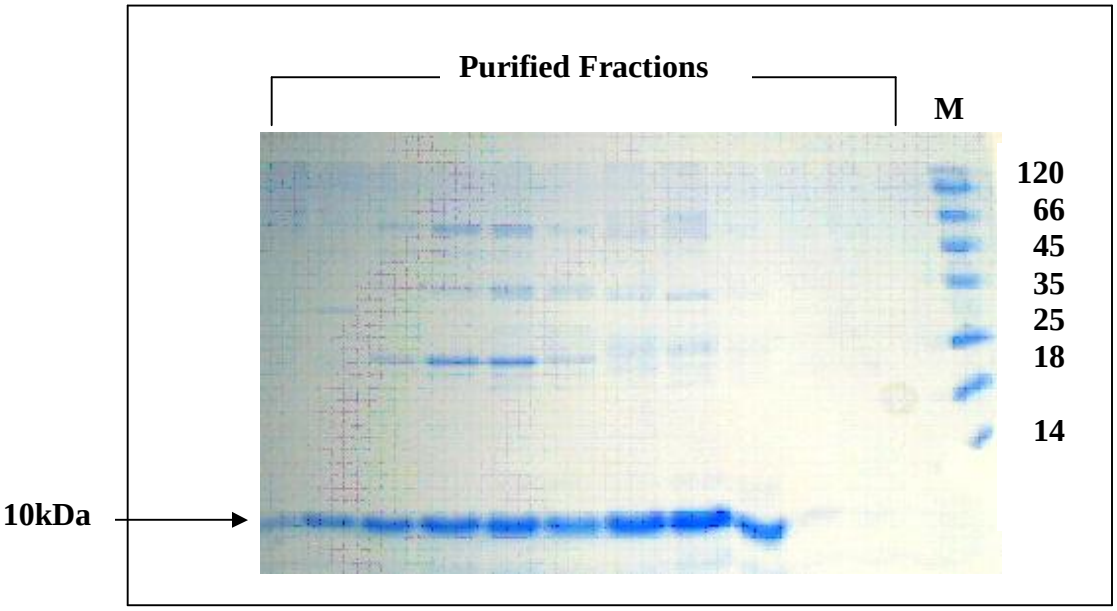


Figure 24.17% SDS-PAGE profile showing the purity of $PacS_N$ after size exclusion chromatography, M-Marker.

NMR samples were prepared by exchanging Tris with phosphate (Pi) buffer at pH 7.0 in the presence of 4M DTT in glove box. The final purified $PacS_N$ was concentrated and protein folding was checked by 1D NMR (Fig.25) and ^{1}H-^{15}N HSQC (Fig.26). Both 1D and 2D spectra shows that protein is well folded in the apo form. Double labeled ^{15}N and ^{13}C sample was prepared by using labeled nitrogen and carbon sources. Both single and double labeled samples were used for NMR experiments to derive the apo$PacS_N$ structure (Fig.27).

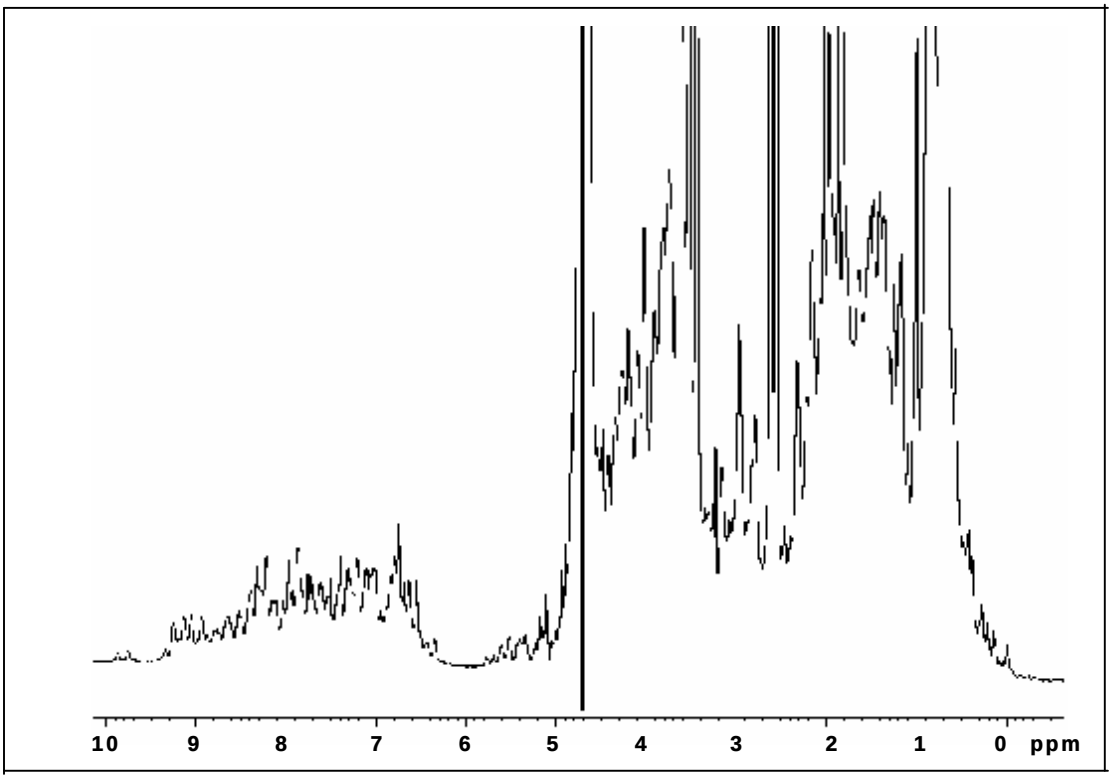


Figure 25.1D NMR spectra of apo$PacS_N$ at 298 K.

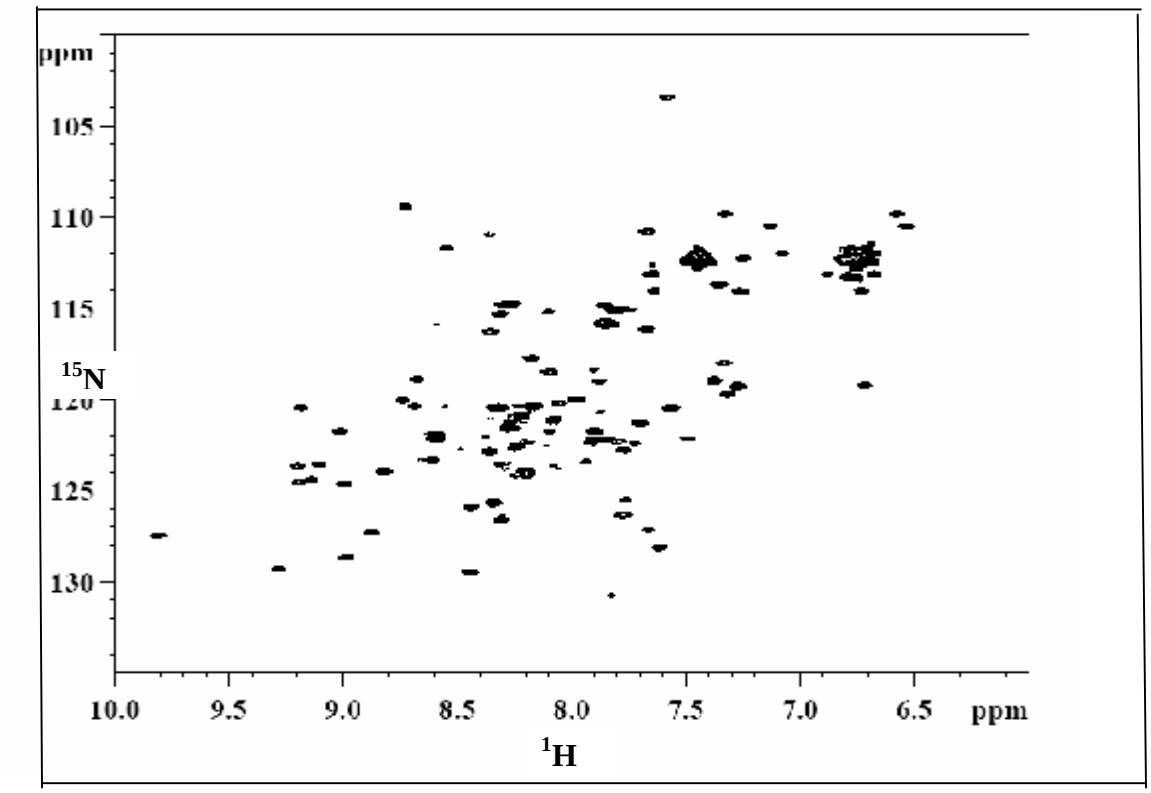


Figure 26.^{1}H-^{15}N HSQC of apoPacS$_N$ recorded at 800 MHz with 298 K.

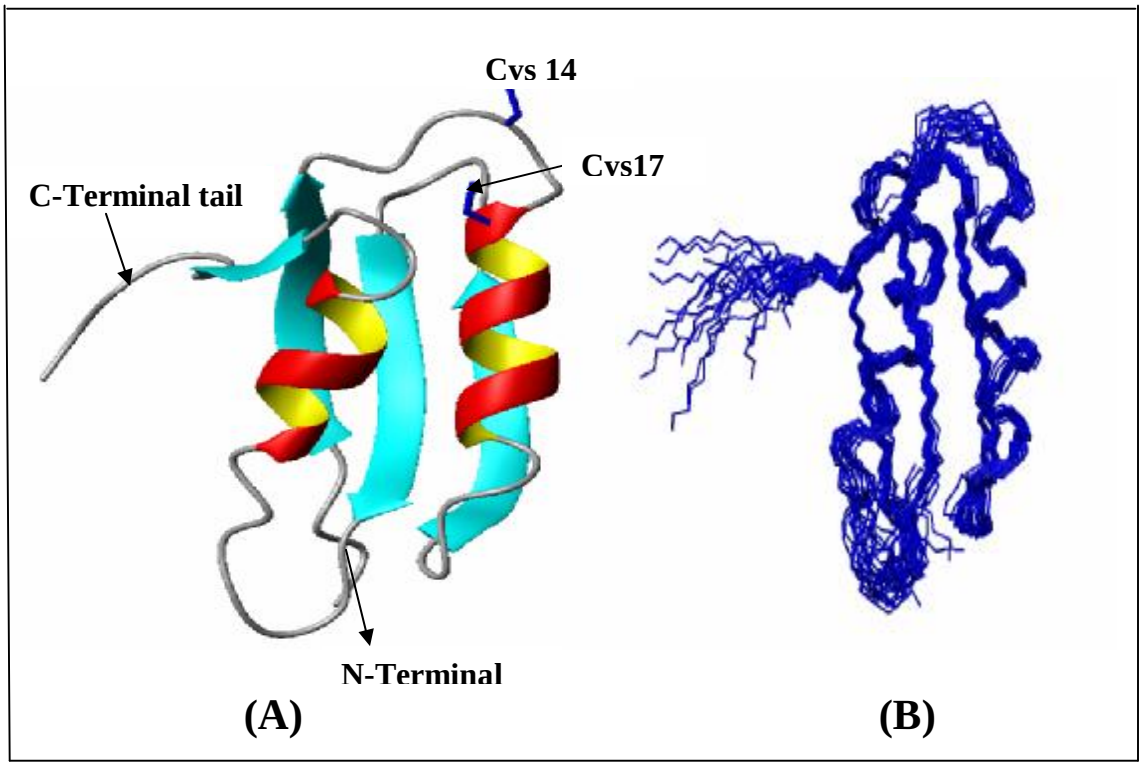


Figure 27. (A)Solution structure of apoPacS$_N$ showing Ferrodoxin-like fold βαββαβ

(B) Backbone structure of apoPacS$_N$.

B.2.1 NMR experiments, Structure calculation and Dynamics for apoPacS_N

Details on the apoPacS_N NMR structure calculation, mobility studies and interaction with copper metallochaperone ScAtx1 is explained in the following attached article:

His 61 mediates the copper transfer from the metallochaperone Atx1 to the amino-terminal region of the copper ATPase PacS

Lucia Banci, Ivano Bertini, Simone Ciofi-Baffoni, Xun-Cheng Su, ***Murugendra Vanarotti****, Nigel J. Robinson*

Submitted

His 61 mediates the copper transfer from the metallochaperone Atx1 to the amino-terminal region of the copper ATPase PacS

Lucia Banci, Ivano Bertini*, Simone Ciofi-Baffoni, Xun-Cheng Su, **Murugendra Vanarotti**, Nigel J. Robinson†

#Magnetic Resonance Center CERM and Department of Chemistry, University of Florence, Via Luigi Sacconi 6, 50019, Sesto Fiorentino, Florence, Italy

†Cell and Molecular Bioscience, Medical School, University of Newcastle, NE2 4HH, UK.

*To whom correspondence should be addressed: Prof. Ivano Bertini, CERM and Department of Chemistry, University of Florence, Via L. Sacconi 6, Sesto Fiorentino, ITALY 50019. Tel.: +39-055-4574272; Fax: +39-055-4574271; bertini@cerm.unifi.it.

RUNNING TITLE: Atx1/PacS interaction

Introduction

Most of the biosphere is either directly, or indirectly, reliant upon effective copper delivery to the thylakoids of primary producers. The conversion of light into useful chemical energy by cyanobacteria and plants involves the transfer of electrons within the thylakoid lumen between two membraneous photosystems. Electrons are commonly transferred via plastocyanin-bound copper. Plastocyanin is located within the thylakoid lumen and is imported as an unfolded protein (1), necessitating a separate copper supply in order to form the holoenzyme.

Analyses of mutants of the cyanobacterium *Synechocystis* PCC 6803 established that two copper transporting P_1-type ATPases, PacS and CtaA, plus a small soluble copper metallochaperone, Atx1 (herein referred to as ScAtx1) are required for normal photosynthetic electron transfer via plastocyanin and for the activity of a second thylakoid located copper protein, a caa_3-type cytochrome oxidase (2;3). In common with related polypeptides from other bacteria, yeast and man (4-7), ScAtx1 directly interacts with soluble, amino-terminal domains of P_1-type copper ATPases (3). However, unlike other copper metallochaperones, ScAtx1 from *Synechocystis* PCC 6803 associates with two such proteins, PacS and CtaA. In *Synechococcus* PCC 7942, PacS is located in thylakoid membranes (8), while CtaA is thought to import copper at the plasmamembrane (9), and the phenotypes of *DctaA* and *DpacS* mutants of *Synechocystis* PCC 6803 are consistent with both ATPases transporting copper in an inward direction into the cytosol and then into the thylakoid lumen (3). This provides an attractive system for studying the process of copper transfer between a copper metallochaperone and its partners. In addition to PacS and CtaA, *Synechocystis* PCC 6803 also contains a P_1-type ATPase, ZiaA, that transports zinc and not copper (10), but has an amino terminal domain with higher affinity for copper than for zinc (11). A sub-domain of ZiaA ($ZiaA_N$) is predicted to form a ferredoxin-like fold analogous to the amino-terminal regions of the two copper transporters, but ScAtx1 gave no detectable two-hybrid

interaction with this sub-domain of ZiaA (11) nor with the entire soluble amino-terminal region of ZiaA (3). Provided there is no freely available cytosolic copper in *Synechocystis* PCC 6803, as implied for *E. coli* (12), lack of ScAtx1-ZiaA$_N$ interaction will impose a kinetic barrier to discourage formation of otherwise thermodynamically favored copper-ZiaA$_N$ complexes (11) while copper trafficks to the thylakoid.

There is considerable interest in understanding the sub-molecular processes by which copper is handed between partner proteins in cellular trafficking pathways to avoid intracellular metal release. Recently, NMR derived solution structure of ScAtx1 (13) and EXAFS data revealed that loop 1, containing Cys-ligands, and loop 5, containing His 61, are involved in copper coordination, approaching themselves to form a symmetrical copper-bridge homodimer. In yeast, it has been proposed that the observed flexibility in loop 1 and 5 of apoAtx1 may provide a trigger for copper release and allow the chaperone to adapt to its two partners; downstream of CTR1 and upstream of Ccc2 (15). Loop 1 of copper-CopZ from *Bacillus subtilis* assumes the conformation of apo- rather than copper-CopZ, upon contact with the N-terminal domain of CopA (16). This implies copper release from the metallochaperone to the exporter and *DcopZ* mutants were subsequently shown to be copper sensitive (17), consistent with a role in export rather than import. It presumed that any analogous mechanism for ScAtx1 must be somehow adapted to encourage copper transfer to the metallochaperone from one P_1-type ATPase, as well as release to another, and His 61 has been suggested to play a significant role to drive this copper transfer mechanism. Indeed, conversion of His 61 to Arg altered two-hybrid interaction with PacS, but not with CtaA, implying that the His 61 can interact differently with the complementary surfaces of the two different copper-ATPases (18).

Here we determined the solution structure of the N-terminal region of PacS (PacS$_N$) in its apo form and studied the protein-protein interaction between apoPacS$_N$ and Cu(I)ScAtx1 in order to

investigate the role of His 61 in the copper transfer mechanism. We establish that apoPacS$_N$ is organized into a ferrodoxin-like fold comprising residues 1-72, while residues 74-95, constituting the linker between the ferrodoxin-like domain and the first transmembrane helix of PacS ATPase, are not structured and flexible. We also establish that apoPacS$_N$ forms a heterodimeric complex with Cu(I)ScAtx1, by breaking the homodimeric state of Cu(I)ScAtx1, and that the copper ion is largely associated in the protein complex with PacS$_N$ rather than with ScAtx1, being His 61 no more involved in copper coordination.

Experimental Procedures

PacS and ScAtx1 protein expression and purification

Synechocystis PCC 6803 genomic DNA was used as template for PCR with primers 5'-GAACAT ATGGCCCAAACCATC-3' and 5'-GAAGAATTCTCATAACCCCGTTACCAATTTGGCCGA-3'. The amplified fragment of DNA containing codons 1-95 encoding the entire amino-terminal region of PacS was ligated into the *NdeI/EcoRI* sites of pET29a to create pETPacS. Recombinant protein was expressed as a soluble protein in *E. coli*. The entire expression tests were conducted in Minifor fermentor. BL21 cells harboring the pETPacS plasmid were grown at 37°C, exposed to 1mM IPTG (isopropyl-1-thio-D-galactopyranoside) and 50 μM of copper sulfate. M9 minimal medium (19) supplemented with [^{13}C] glucose and (^{15}NH4)2SO$_4$ as only sources of carbon and nitrogen was used. Cells were grown at 37°C to an A_{600} nm of 0.6 prior to induction with 1 mM isopropyl-1-thio-D-galactopyranoside (IPTG), copper sulfate was added to the growth medium to a final concentration of 50 μM. Cells were then lysed by freeze-thawing and repeated sonication. The lysate was clarified by centrifugation at 100,000 × g for 40 min at 4 °C. The supernatant was filtered through a low protein-binding 0.45 μm filter and loaded onto a 5ml Hi-Trap Q XL column (Amersham Pharmacia Biotech) which was pre-equilibrated with Tris 25 mM pH 9. Chromatography was done using a Pharmacia fast

protein liquid chromatography (FPLC) unit. The protein was eluted by a 100 ml 50mM NaCl step. The fractions containing PacS were concentrated and subjected to size exclusion chromatography on a HiLoad 16/60 Superdex 75 column (Amersham Pharmacia Biotech) equilibrated in 25 mM Tris pH 9. To prevent disulfide formation, which might occur because of the presence of two cysteines, the protein samples were kept in anaerobic conditions. In these conditions the buffer was changed by an ultra filtration device (Amicon) against sodium phosphate pH 7.0 with the addition of reducing agent dithiothreitol (DTT) at a final concentration of 4 mM. Protein expression and purification was monitored by sodium dodecyl phosphate-polyacrylamide (SDS) gel electrophoresis in 17% polyacrylamide gels stained with Coomasie brilliant blue R-250 against Perfect Protein marker (Novagen). Isotopic labeling, protein expression, purification and metallation of ScAtx1 was performed as previously described (13).

NMR experiments and structure calculations

NMR spectra were performed on Avance 800, 700, 600 and 500 Bruker spectrometers operating at proton nominal frequencies of 800.13 MHz, 700.13 MHz, 600.13 MHz and 500.13 MHz, respectively. All the triple resonance (TXI 5-mm) probes used were equipped with Pulsed Field Gradients along the z-axis. The 800 and 500 MHz machines are equipped with a triple resonance cryoprobe.

The NMR experiments, reported on **Table 1** of the Supplementary Material, were recorded on $^{13}C/^{15}N$ and ^{15}N labelled and unlabelled samples. To identify the coordination mode of copper(I) binding histidine, a ^{1}H-^{15}N HSQC experiment was performed for measuring $^{2}J_{N\varepsilon H\delta}$, $^{2}J_{N\varepsilon H\varepsilon}$, $^{2}J_{N\delta H\varepsilon}$, and $^{3}J_{N\delta H\delta}$ coupling constants (20). In this experiment, the INEPT delay was set to 22 ms. All 3D and 2D

spectra were collected at 298 K, processed with the standard Bruker software (XWINNMR) and analysed with CARA program..

An automated CANDID approach combined with DYANA torsion angle dynamics algorithm (21) was used to assign the ambiguous NOE cross-peaks and to have a preliminary apoPacS_N structure. Structure calculations were then performed through iterative cycles of DYANA (22) followed by Restrained Energy Minimization with AMBER 5.0 (23) applied to each member of the family. The quality of the structures was evaluated using the programs PROCHECK-NMR (24). The figure was generated with the program MOLMOL (25).

The atomic coordinates of apoPacS_N have been deposited in the Protein Data Bank.

Relaxation Measurements and Analysis

Relaxation experiments were performed on Bruker Avance 600 MHz spectrometers at 298 K. ^{15}N R_1, R_2, and steady-state heteronuclear NOEs were measured with pulse sequences as described by Farrow ***et al*** (30). In all experiments the water signal was suppressed with 'water flipback' scheme (26). R_1 and R_2 experiments were acquired with 8 or 16 scans, while $\{^1H\}-^{15}N$ NOE spectra were acquired with 64 scans. Duplication of measurements was performed to estimate the experimental uncertainty. A recycle delay of 3 seconds was used for R_1 and R_2 relaxation experiments except for the NOE experiments in which the recycle delay was 5 seconds. A total of 2048 K (^{1}H) × 128 (^{15}N) data points were collected. All spectra were processed with the XWINNMR program (Bruker) and analyzed with CARA software. Relaxation rates R_1 and R_2 were determined by fitting the cross-peak intensities measured as a function of the delay within the pulse sequence, to a single-exponential decay. Errors on the rates were estimated through a Monte Carlo approach. The heteronuclear NOE values were obtained from the ratio of the peak intensity for ^{1}H-saturated and unsaturated spectra. The experimental

relaxation rates were used to map the spectral density function values, $J(\omega_H)$, $J(\omega_N)$, $J(0)$, following a procedure available in literature (27).

The overall rotational correlation time τ_m values were estimated from the R_2/R_1 ratio using the program *quadric_diffusion*, available from the Web site of A. G. Palmer, III (28). In this analysis, care was taken to remove from the input relaxation data those NHs having an exchange contribution to the R_2 value or exhibiting large-amplitude internal motions on a time scale longer than a few hundred picoseconds, identified from low NOE values, as inclusion of these data would bias the calculated tensor parameters (29,30). The input structure for *quadric diffusion* is the energy-minimized average solution structure of apoPacS_N.

NMR Titration of the Two Proteins

Titrations of Cu(I)^{15}NScAtx1 with unlabeled apoPacS_N and of apo^{15}NPacS_N with unlabeled Cu(I)ScAtx1 was performed with NMR spectroscopy, following the spectral changes in the ^{1}H-^{15}N HSQC spectra upon the addition of increasing amounts of the unlabeled partner. Aliquots were added in a Coy chamber under nitrogen atmosphere at 298 K. Two-dimensional TOCSY and NOESY, and three-dimensional NOESY-^{15}N HSQC experiments were recorded on ^{15}NScAtx1/PacS_N and ^{15}NPacS_N/ScAtx1 mixtures with protein concentration ratios of 1:0.5 and 1:1.

Results and Discussion

Solution structure and mobility of apoPacS$_N$

The ^{1}H-^{15}N HSQC spectrum of the apo form of PacS$_N$ (95 amino acids) shows a good spreading of signals, typical of proteins with a structural organization. However, several peaks are clustered in a spectral region typical of unfolded polypeptides (amide proton resonances clustered between 8 and 8.5 ppm). 78 out of the expected 91 ^{15}N backbone amide resonances were observed and assigned. The backbone NH resonances are missing for residues belonging to loop regions (Ala 15, Ala 16, Gly 49, Glu 50, Thr 51 and 52) and to the C-terminal tail (Ser 77, Gln 78, Gln 79, Val 86, Phe 87, Leu 91). The ^{1}H, ^{13}C and ^{15}N resonance assignments of apoPacS$_N$ are reported in **Table 1** of the Supplementary Material.

The solution structure of apoPacS$_N$ was obtained by using 1356 meaningful upper distance limits and 86 angle restraints for 42 ϕ and 44 ψ angles. After Restrained Energy Minimization with AMBER program on each of 20 lowest target function structures, obtained from DYANA calculations, the RMSD (root mean square deviation) for protein backbone and heavy atoms to the mean structure (for residues 4-70) is 0.63 Å (with a variability of 0.19 Å) and 1.08 Å (with a variability of 0.22 Å), respectively. The penalties for distance constraints and angle constraints are 0.15 ± 0.06 Å^2 and 0.02 ± 0.01 rad^2, respectively. The statistical analysis of the REM family of apoPacS$_N$ structures are reported in **Table 2** of the Supplementary Material. The structure displays the following secondary structure elements: 3-10 (β1), 17-26 (α1), 31-37 (β2), 41-48 (β3), 54-62 (α2), 66-68 (β4) in accordance with the $^{3}J_{HNH\alpha}$ coupling constants, the $d_{\alpha N}(i-1,i)/d_{N\alpha}(i,i)$ ratios and the NOEs patterns. In **Fig. 1**, the 20 conformer structure of apoPacS$_N$ is shown as a tube, whose radius is proportional to the backbone RMSD of each residue.

^{15}N R_1, R_2, and ^1H-^{15}N NOE values, which provide information on internal mobility, were measured at 600 MHz for all assigned backbone NH resonances. Such average values are 2.00 s^{-1} (with a variability of 0.17 s^{-1}), 8.44 s^{-1} (with a variability of 1.71 s^{-1}) and 0.75 s^{-1} (with a variability of 0.09 s^{-1}), respectively, calculated excluding the C-terminal tail (residues 71-95). The experimental relaxation data are reported in the **Fig. 2** of the Supplementary Material. The correlation time for molecular reorientation (τ_m), as estimated from the R_2/R_1 ratio, is 5.6 ± 0.4 ns, as expected for a protein of this size in a monomeric state (15). R_1, R_2 and ^1H-^{15}N NOE values were mostly homogeneous along the polypeptide sequence from residues 1 to 70, with the exception of a few residues mainly located in loop regions, while the C-terminal tail encompassing residues 71-95 shows very low or negative ^1H-^{15}N NOEs values. From the spectral density function analysis (**Fig. 3**), it appears that the apo$PacS_N$ protein can be considered as a rigid body with a C-terminal flexible region, which have $J(\omega_H)$ significantly higher than average, thus indicating for these residues the presence of local motions in ns-ps timescale, i.e. faster than the overall protein tumbling rate. In addition, considering that the NHs belonging to the C-terminal tail are all clustered in the spectral region typical of unfolded polypeptides, it can be concluded that the C-terminal tail of the protein is fluctuating in solution in a random coil conformation. A certain degree of local backbone flexibility on both ns-ps and ms-μs timescales (J(0) and $J(\omega_H)$ values higher than the average (**Fig.3**) is also observed for some residues located close to the metal binding motif, in loop 1,3 and at the beginning of helix α1.

Titration of Cu(I)^{15}NScAtx1 with unlabeled apo$PacS_N$ and of apo^{15}N$PacS_N$ with unlabeled apoScAtx1

The interaction between the copper chaperone ScAtx1 and the N-terminal metal binding domain $PacS_N$ has been investigated recording ^1H-^{15}N HSQC spectra on Cu(I)^{15}NScAtx1 and apo^{15}N$PacS_N$ samples titrated with the unlabelled partner apo$PacS_N$ and Cu(I)ScAtx1, respectively. In both titrations,

the ^{1}H and ^{15}N resonances showed the formation of a new species which increase in intensity upon increasing amounts of unlabeled proteins. In both titrations the ^{1}H-^{15}N HSQC spectra of the new species is not fully reproducing the ^{1}H-^{15}N HSQC spectra of the isolated apoScAtx1 and Cu(I)PacS_N proteins, respectively. Furthermore, in both titrations ^{1}H-^{15}N resonances of a few residues close to the metal binding site experience some broadening. Such spectral changes suggest that (i) ScAtx1 is interacting with PacS_N and (ii) the new species represents the complex ScAtx1/PacS_N, whose lifetime is long relative to the difference between $\delta_{complex}$ and δ_{free}, indicating that the proteins in the complex are in slow exchange with the free proteins in solution.

The spectral changes were followed up to a final protein ratio of 1:2. **Fig. 4** shows the ^{1}H-^{15}N HSQC of the isolated Cu(I)^{15}NScAtx1 protein or apo^{15}NScAtx1 (*blue contours*) overlaid onto that of the ScAtx1/PacS_N mixture containing equal concentrations of the two proteins (*red contours*). Cross-peaks of Cu(I)ScAtx1 residues 11-14, comprising the copper ligand Cys 12, broadened beyond detection. Comparing the spectra **A** with **B** in **Fig. 4**, it results that the ^{1}H and ^{15}N resonances of ScAtx1 in the complex are more similar to the apoScAtx1 form than to the Cu(I)ScAtx1. This data suggests that ScAtx1 in the complex is present in a conformation similar to the apo form. Accordingly, the ^{1}H and ^{15}N chemical shift differences between Cu(I)ScAtx1 or apoScAtx1 alone and the ScAtx1/PacS_N 1:1 mixture (**Fig. 4**) shows that major changes are present in the first chemical shift comparison, Cu(I)ScAtx1-ScAtx1/PacS_N 1:1 mixture. In particular, the NHs chemical shifts of loop 3 (residues 38-41), which is in close contact with the metal binding site (**Fig. 1**), are the same observed in the apoScAtx1 (**Fig. 4**), suggesting that ScAtx1 in the complex resembles to a conformation similar to apoScAtx1. In addition, in both cases the chemical shift changes are mainly localized in two regions of the protein corresponding to stretches 9-25 and 55-64 (**Fig. 4**). These stretches constitute loop 1, helices

α_1 and α_2, and C-terminus. The same regions are involved in the protein-protein interaction between the yeast and bacterial homologuos complexes Atx1/Ccc2a and CopZ/CopAb, respectively (31).

Fig. 4 shows the ^{1}H-^{15}N HSQC of the isolated apo^{15}NPacS$_N$ protein or Cu(I)^{15}NPacS$_N$ (*blue contours*) overlaid onto that of the ScAtx1/PacS$_N$ mixture containing equal concentrations of the two proteins (*red contours*). Cross-peaks of apoPacS$_N$ residues 13, 14, 17, 19-21 comprising the copper ligands Cys 15 and Cys 17, broadened beyond detection. Ala 18 is the only residue in the metal binding region whose NH is still detectable and its chemical shift, which is different with respect to apo and Cu(I)PacS$_N$, clearly suggests the formation of the protein-protein complex (**Fig.5**). Comparing spectra **A** with **B** in **Fig. 5**, it results that there is not drastic difference between the ^{1}H and ^{15}N resonances of PacS$_N$ in the complex and isolated apo or Cu(I)PacS$_N$ forms. Accordingly, the ^{1}H and ^{15}N chemical shift differences between Cu(I)PacS$_N$ or apoPacS$_N$ alone and the ScAtx1/PacS$_N$ 1:1 mixture shows that major changes are localized in two regions of the protein corresponding to stretches 13-24 and 61-64 (**Fig. 4**). These stretches constitute loop 1, helices α_1 and loop 5, which are on par with yeast Ccc2a/Atx1 complex (31).

EXAFS data indicate that one His (His-61 in loop 5) and two Cys (Cys-12 and -15 in loop 1) of ScAtx1 are involved in copper(I) coordination. ^{1}H-^{15}N HSQC experiments optimized for the detection of the $^{2}J_{NH}$ couplings in the imidazole ring, performed on apo and copper(I) forms to investigate the coordination status of His-61 at room temperature in solution, are consistent with the presence of a coordination bond between N$^{\varepsilon 2}$ of His-61 and copper(I) in a dimeric structural arrangement . In order to follow the coordination state of His 61 in the ScAtx1/PacS$_N$ protein interaction studies, $^{2}J_{NH}$ ^{1}H-^{15}N HSQC experiments were performed during the titration between Cu(I)^{15}NScAtx1 and apoPacS$_N$. From these data, it results that the typical pattern of His 61 bound to copper(I) disappears during the additions of apoPacS$_N$ and a broad peak in the $^{2}J_{NH}$ ^{1}H-^{15}N HSQC spectrum appears whose intensity increases

upon additions of apoPacS_N and its chemical shift is similar to that observed in the pattern of apoPacS_N (**Fig. 6**). These data suggest that His 61 is no more involved in copper coordination in the ScAtx1/PacS_N complex. It is likely that ^{1}H resonance of Hε1 of His 61 is not observed in the complex as a consequence of broadening beyond detection after the interaction with its protein partner.

Dynamic Characterization of the complex of Cu(I)ScAtx1 with apoPacS_N

The ^{15}N longitudinal (R_1) and transverse (R_2) relaxation rates are very sensitive to intermolecular association processes and the calculated correlation tumbling from the ratio of R_2 and R_1 could be a indicator to molecular aggregation in solution state. R_2 and R_1 experiments, performed on the Cu(I)^{15}NScAtx1/apoPacS_N mixture in the two steps of the titration (2:1 and 1:1 ScAtx1/PacS_N ratio) showed that, overall, the transverse relaxation rates are increased with respect to pure apoScAtx1 while it remains unchanged with respect to the pure Cu(I)ScAtx1 (**Table 3**). For Cu(I)-^{15}NScAtx1 in the mixture, τ_m was estimated using the R_2/R_1 ratio and found to be 7.1± 0.4 ns at 1:1 protein ratio. The resulted solution of Cu(I)ScAtx1 and PacS_N was checked with BCS, a strong cuprous ligand. BCS can take the copper out from the mixture of Cu(I)ScAtx1 and PacS_N, producing the apo proteins. In presence of 1mM BCS, the complex formed by 0.38 mM CuScAtx1 and 0.34mM PacS_N was turned into apo forms, resulting in a red color BCS-copper complex. The 2D ^{1}H-^{15}N HSQC spectrum is almost identical to apoScAtx1. Furthermore, the addition of the copper(I) chelator BCS determines a decrease of the correlation time τ_m becoming similar to that of a monomeric state of apoScAtx1 (**Table 3**). These data indicates that a protein complex is formed and the picture of adduct formation is consistent with that coming out from the titration.

R_2 and R_1 experiments, performed on the apo^{15}NPacS_N/Cu(i)ScAtx1 mixture in the final step of the titration showed that, overall, the transverse relaxation rates are increased with respect to pure apo^{15}NPacS_N and the correlation time τ_m of Cu(I)-^{15}NPacS_N in the mixture, as estimated using the R_2/R_1

ratio, was found to be 7.9 ± 0.9 ns at 1:1 protein ratio. The latter value increases with respect to the pure apo$PacS_N$ (5.6 ± 0.4 ns), indicating the formation of the protein adduct.

Conclusions

In conclusion, we have structurally characterized the N-terminal metal binding region of the copper(I) ATPase PacS in its apo state and characterized its interaction with the copper(I) metallochaperone ScAtx1. From our data it emerges a picture of the copper transfer between the two protein partners and how His 61 mediates the copper(I) transfer in the protein complex. Indeed, its removal from metal coordination sphere upon interaction with $PacS_N$ can be the determinant for triggering the copper transfer from the metallochaperone to the metal binding site of $PacS_N$.

Acknowledgments

This work was supported by the European Community (SPINE n°QLG2-CT-2002-00988, "Structural proteomics in Europe") and by MIUR COFIN 2003 "Il ruolo degli ioni metallici nei processi metabolici". NJR acknowledges financial support from BBSRC.

References

1. Bogsch, E., Brink, S., and Robinson, C. (1997) *EMBO J.* **16,** 3851-3859.

2. Tottey, S., Rich, P. R., Rondet, S. A. M., and Robinson, N. J. (2001) *J. Biol. Chem.* **276,** 19999-20004.

3. Tottey, S., Rondet, S. A., Borrelly, G. P., Robinson, P. J., Rich, P. R., and Robinson, N. J. (2002) *J. Biol. Chem.* **277,** 5490-5497.

4. Banci, L., Bertini, I., Ciofi-Baffoni, S., Del Conte, R., and Gonnelli, L. (2003) *Biochemistry* **42,** 1939-1949.

5. Radford, D. S., Kihlken, M. A., Borrelly, G. P. M., Harwood, C. R., Le Brun, N. E., and Cavet, J. S. (2003) *FEMS Microbiol. Lett.* **220,** 105-112.

6. Wernimont, A. K., Huffman, D. L., Lamb, A. L., O'Halloran, T. V., and Rosenzweig, A. C. (2000) *Nat. Struct. Biol.* **7,** 766-771.

7. Arnesano, F., Banci, L., Bertini, I., Cantini, F., Ciofi-Baffoni, S., Huffman, D. L., and O'Halloran, T. V. (2001) *J. Biol. Chem.* **276,** 41365-41376.

8. Kanamaru, K., Kashiwagi, S., and Mizuno, T. (1994) *Mol. Microbiol.* **13,** 369-377.

9. Phung, L. T., Ajlani, G., and Haselkorn, R. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 9651-9654.

10. Thelwell, C., Robinson, N. J., and Turner-Cavet, J. S. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 10728-10733.

11. Borrelly, G. P. M., Rondet, S. A. M., Tottey, S., and Robinson, N. J. (2004) *Mol. Microbiol.,* in press.

12. Changela, A., Chen, K., Xue, Y., Holshen, J., Outten, C. E., O'Halloran, T. V., and Mondragon, A. (2003) *Science* **301,** 1383-1387.

13. Banci L, Bertini I, Ciofi-Baffoni S, Su XC, Borrelly GP, Robinson NJ.(2004) J Biol Chem. 2004 25;279(26):27502-10.

14. Bogsch, E., Brink, S., and Robinson, C. (1997) *EMBO J.* **16,** 3851-3859.

15.Arnesano, F., Banci, L., Bertini, I., Cantini, F., Ciofi-Baffoni, S., Huffman, D. L., and O'Halloran, T. V. (2001) *J. Biol. Chem.* **276,** 41365-41376

16. Banci, L., Bertini, I., Ciofi-Baffoni, S., Del Conte, R., and Gonnelli, L. (2003) *Biochemistry* **42,** 1939-1949

17. Radford, D. S., Kihlken, M. A., Borrelly, G. P. M., Harwood, C. R., Le Brun, N. E., and Cavet, J. S. (2003) *FEMS Microbiol. Lett.* **220,** 105-112

18. Tottey S, Rondet SA, Borrelly GP, Robinson PJ, Rich PR, Robinson NJ.(2002). J Biol Chem. 2002 Feb 15; 277(7):5490-7

19. Sambamurti, K., Sevlever, D., Koothan, T., Refolo, L.M., Pinnix, I., Gandhi, S., Onstead, L., Younkin, L., Prada, C.M., Yager, D. *et al.* (1999). *J. Biol. Chem.*, **274**, 26810–26814.

20. Eijkelenboom, A. P., Van den Ent, F. M., Vos, A., Doreleijers, J. F., ard,K.,Tullius,T.D.,Plasterk, R. H., Kaptein, R., and Boelens, R. (1997) *Curr. Biol.* **7,** 739-746.

21. Herrmann, T., Güntert, P., and Wüthrich, K. (2002) *J. Mol. Biol.* **319,** 209-227.

22. Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) *J. Mol. Biol.* **273,** 283-298.

23. Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E., Ferguson, D. M., Seibel, G. L., Singh, U. C., Weiner, P. K., and Kollman, P. A. (1997) *AMBER, Version 5.0*, University of California, San Francisco.

24. Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR* **8,** 477-48.

25.Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graphics* **14,** 51-55

26. Grzesiek, S., and Bax, A. (1993) *J. Am. Chem. Soc.* **115,** 12593-12594.

27. Peng, J. W., and Wagner, G. (1992) *J. Magn. Reson.* **98,** 308-332.

28. Mandel, M. A., Akke, M., and Palmer, A. G., III (1995) *J. Mol. Biol.* **246,** 144-163.

29. Kay, L. E., Torchia, D. A., and Bax, A. (1989) *Biochemistry* **28,** 8972-8979.

30. Tjandra, N., Feller, S. E., Pastor, R. W., and Bax, A. (1995) *J. Am. Chem. Soc.* **117,** 12562-12566.

31. Banci L, Bertini I, Ciofi-Baffoni S, Huffman DL, O'Halloran TV. (2001)

*J Biol Chem.*16; 276(11):8415-26.

Supplementary

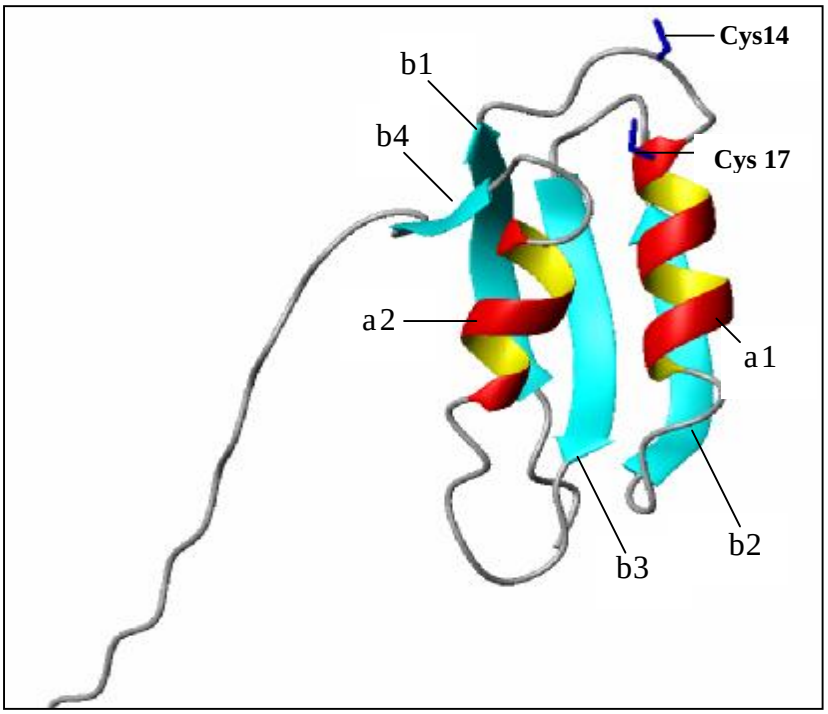


Fig. 1

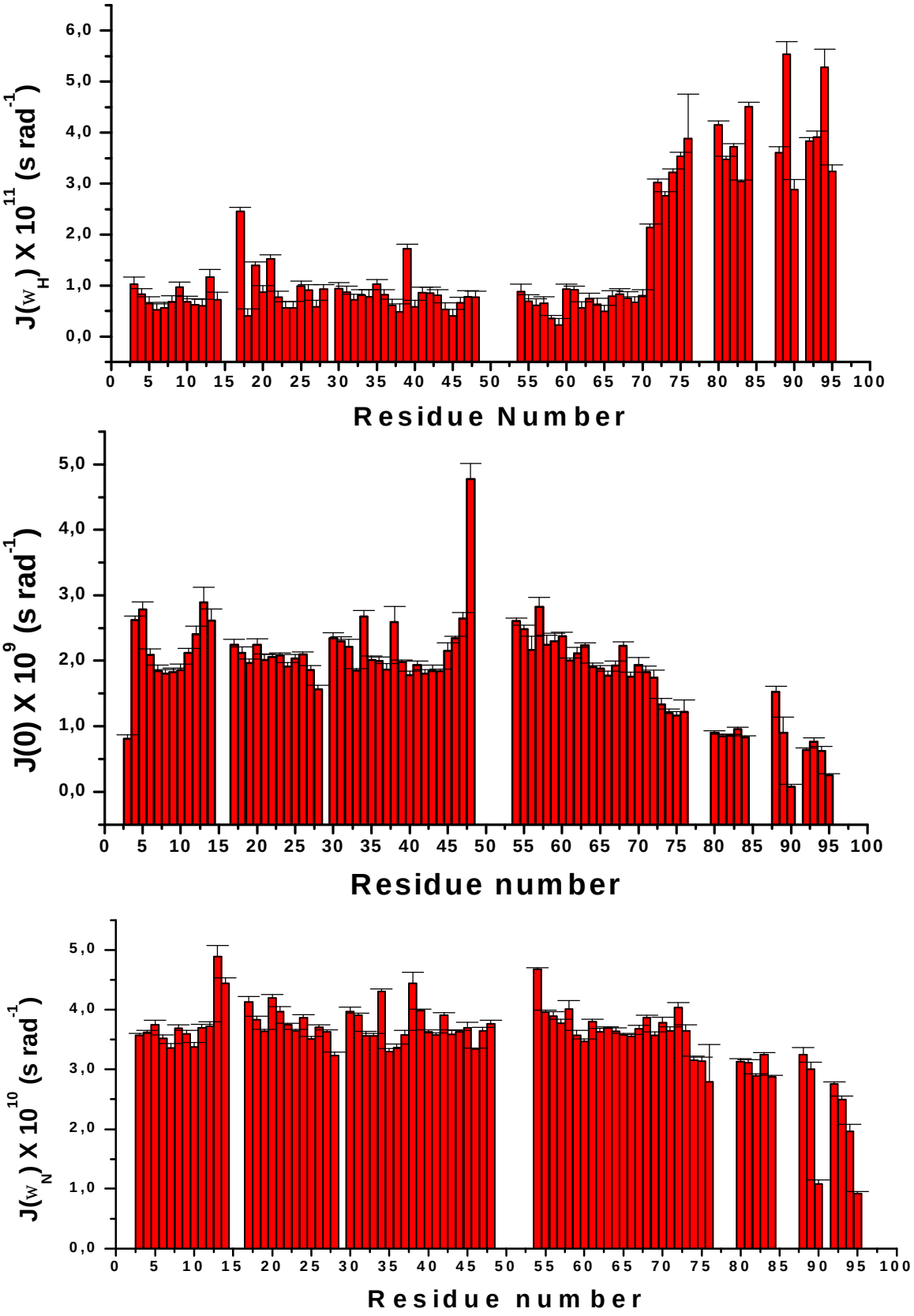

J(ω_H) X 10^11 (s rad^-1)
Residue Number
J(0) X 10^9 (s rad^-1)
Residue number
J(ω_N) X 10^10 (s rad^-1)
Residue number

Fig. 3

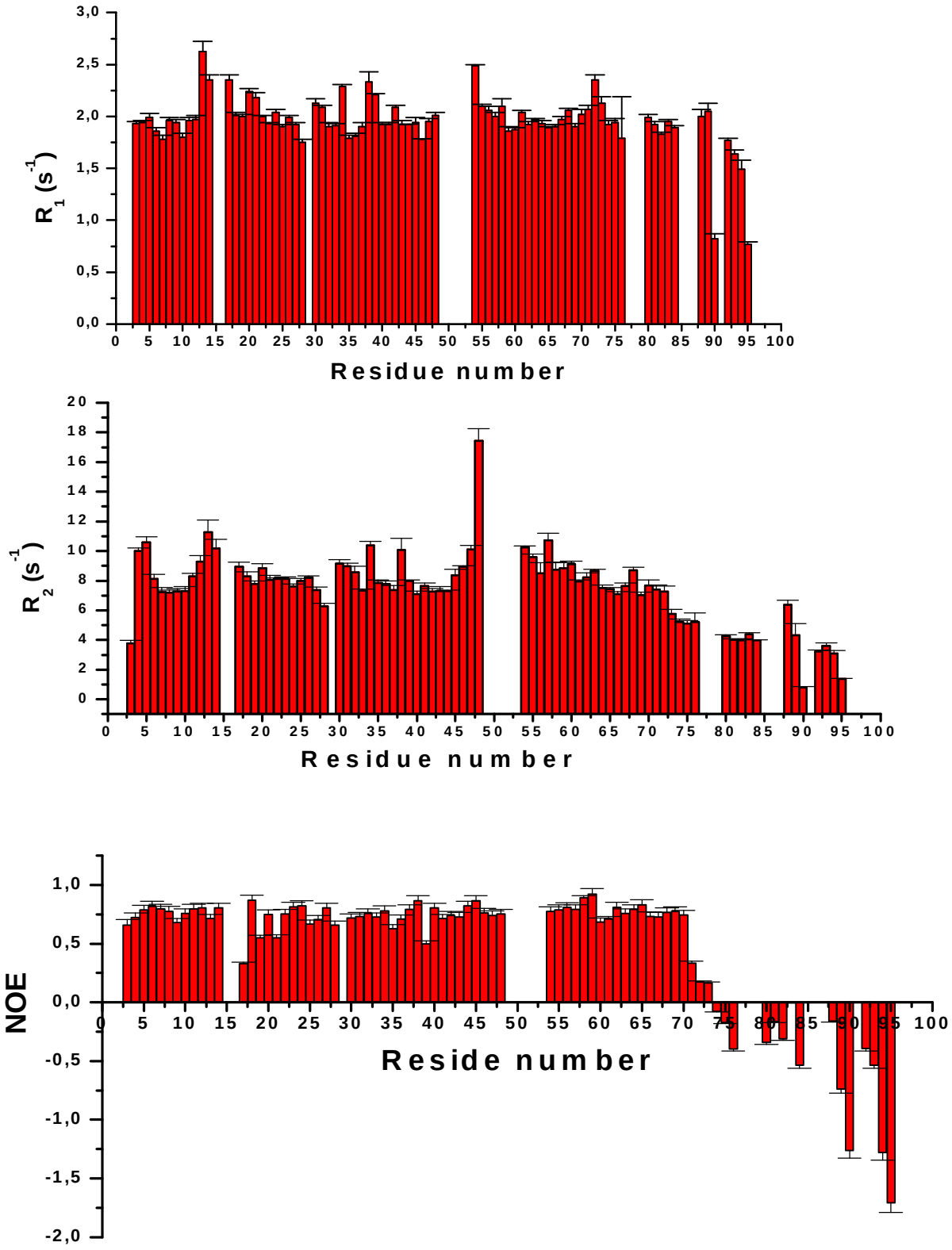


Fig. 2

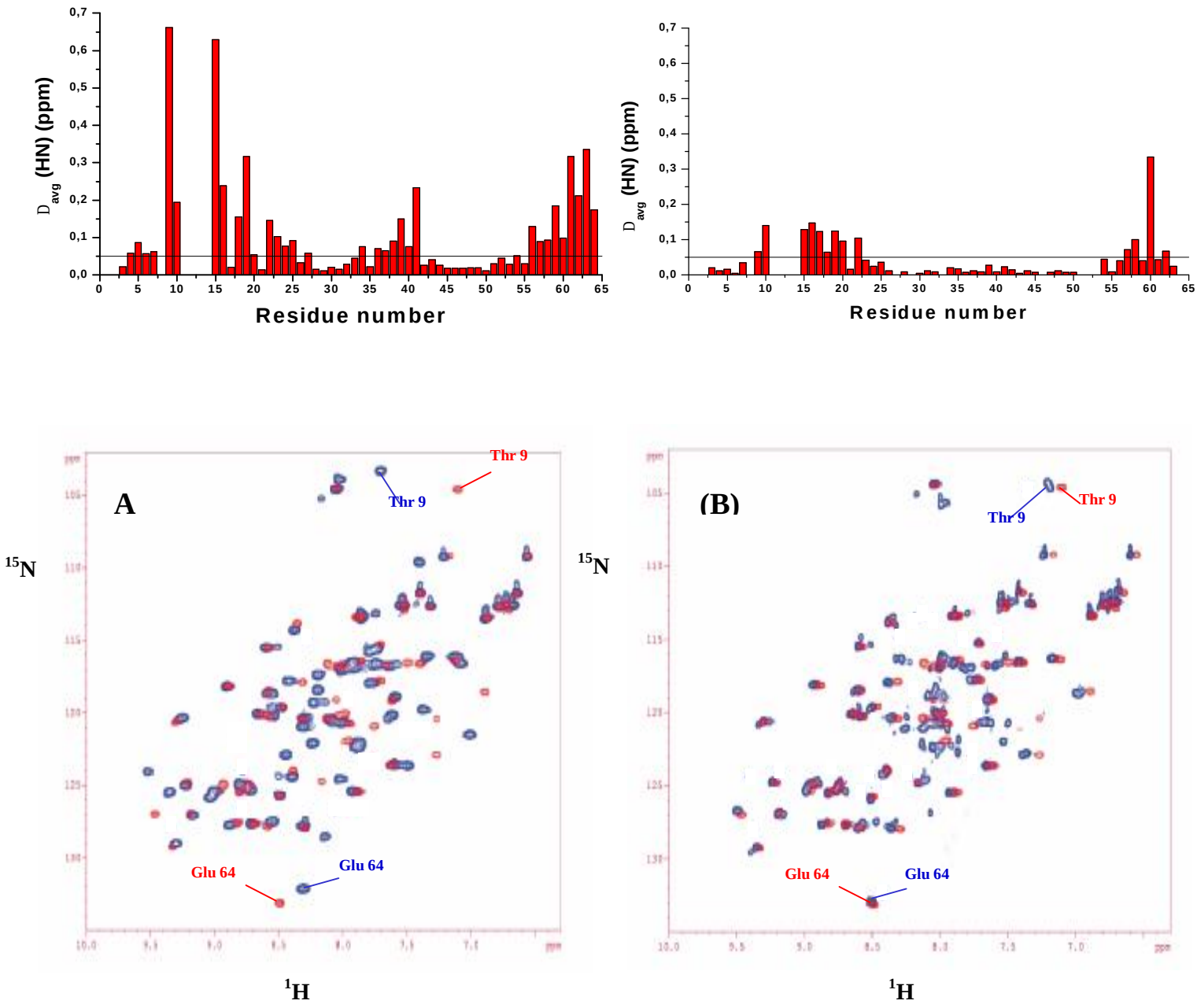
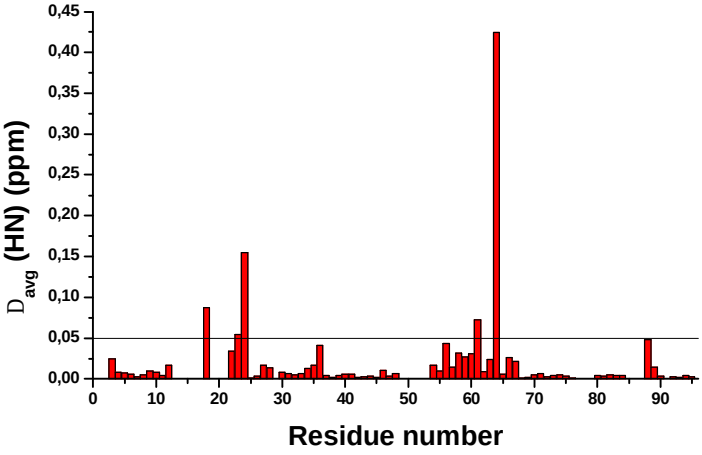
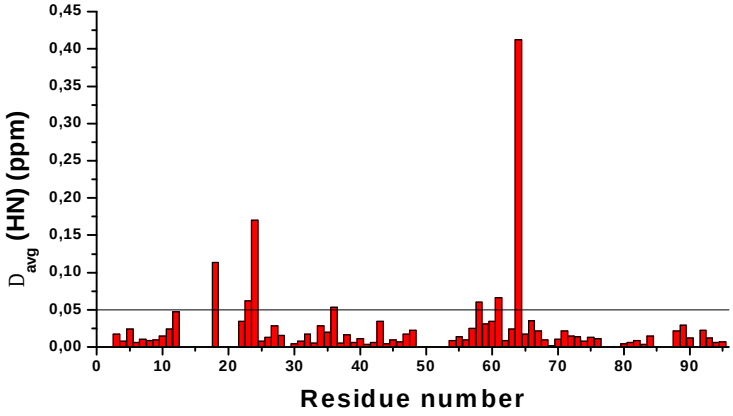
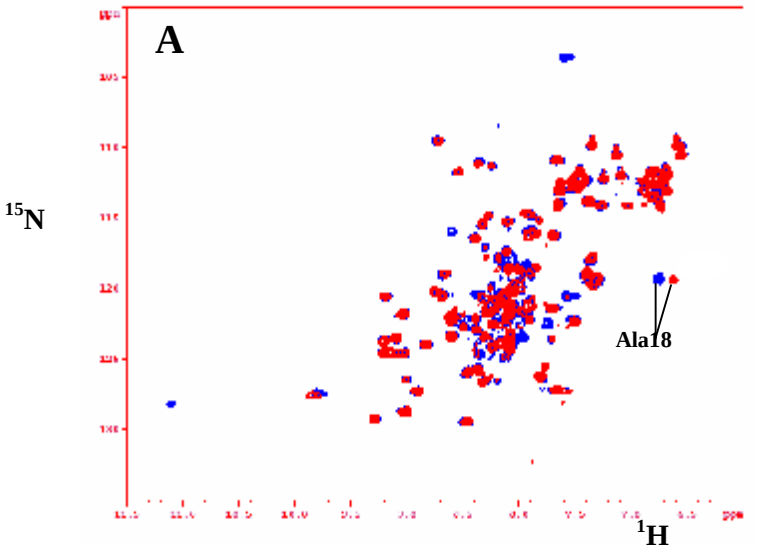


Fig. 4


D avg (HN) (ppm)
0,45
0,40
0,35
0,30
0,25
0,20
0,15
0,10
0,05
0,00
0 10 20 30 40 50 60 70 80 90
Residue number


D avg (HN) (ppm)
0,45
0,40
0,35
0,30
0,25
0,20
0,15
0,10
0,05
0,00
0 10 20 30 40 50 60 70 80 90
Residue number


A
15N
Ala18
1H

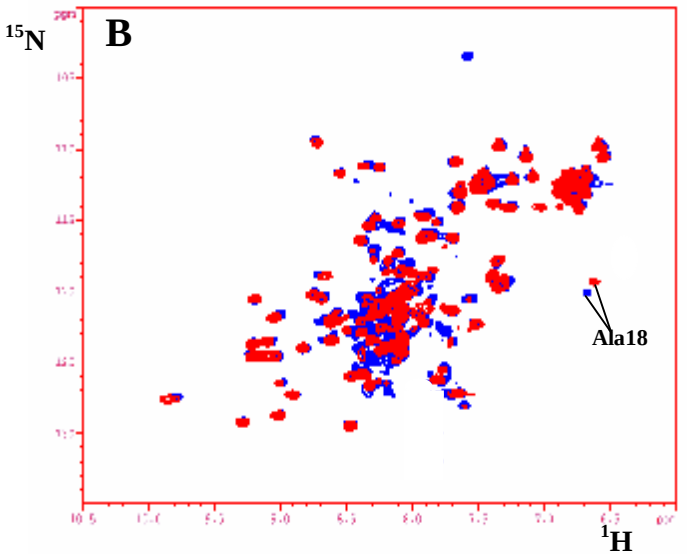

15N
B
Ala18
1H

Fig.5

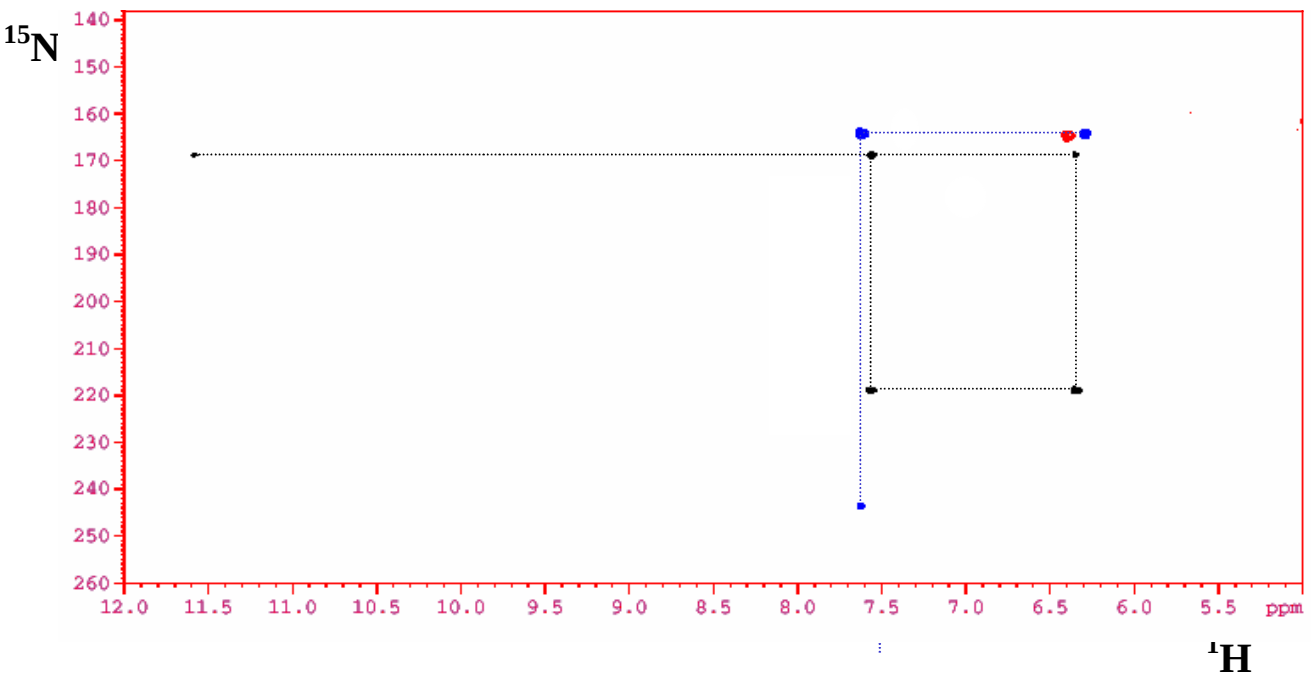

15N
140
150
160
170
180
190
200
210
220
230
240
250
260
12.0
11.5
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
ppm
1H

Fig.6

Table 2. Statistical quality analysis of the REM family and the mean structure of apoPacS_N from *Synechocystis* PCC 6803[a]

RSM violations per experimental distance constraint (Å)[b]	**REM (20 structures)**	**<REM>**
Intraresidue (204)	0.0148 ± 0.0021	0.0153
Sequential (299)	0.0080 ± 0.0017	0.0097
Medium range[c] (241)	0.0115 ± 0.0027	0.0151
Long range (273)	0.0134 ± 0.0019	0.0110
Total (1017)	0.0120 ± 0.0011	0.0125
Phi (44) (deg)	0	0
Average number of violations per structure		
Intraresidue	4.65 ± 1.65	2
Sequential	2.75 ± 0.88	5
Medium range[c]	4.90 ± 1.4	3
Long range	6.70 ± 1.62	4
Total	19.00 ± 3.00	14
Phi	0.2 ± 0.05	0
Average no. of NOE violations larger than 0.3 Å	0	0
Total NOE square deviations ($Å^2$)	0.15 ± 0.06	0.15
Average torsion deviations (rad^2)	0.02 ± 0.01	0
RMSD to the mean structure (3-64) (Å)	0.49 ± 0.12 Å (BB)	
	0.83 ± 0.14 (HA)	

Structural analysis[d]		
% of residues in most favourable regions	75.5	74.6
% of residues in allowed regions	21.4	20.3
% of residues in generously allowed regions	2.8	3.4
% of residues in disallowed regions	1.0	1.3
H-bond energy (kJ mol^{-1})	3.63 ± 0.12	3.63
Overall G-factor	-0.22 ± 0.02	-0.30

[a]REM means the energy minimized ensemble of 20 structures, <REM> is the energy minimized average structure of the ensemble.

[b]The number of experimental constraints for each structure is reported in the ensemble.

[c]Medium range distance constraints are those within residues (i,i+2), (i,i+3), (i,i+4) and (i,i+5).

[d]resulted from the Ramachandran plot analysis over the assigned residues. In the PROCHECK statistics, the average hydrogen-bond energy within 2.5-4.0 kJ mol^{-1} and overall G-factor over -0.5 is expected to be a good-quality structure.

Table 3. Average R_1 and R_2 relaxation rates (s^{-1}) for amide ^{15}N nuclei measured for apo and Cu(I)ScAtx1 from *Synechocystis* PCC 6803 and for Cu(I)ScAtx1 in the presence of apoPacS_N. Correlation time for molecular tumbling τ_m(ns) as estimated from the R_2/R_1 ratio is also reported.

	$R_1(s^{-1})$	$R_2(s^{-1})$	τ_m(ns)
ApoScAtx1(0.6mM)	2.46 ± 0.07	7.09 ± 0.19	4.3 ± 0.3
Cu(I)ScAtx1(0.8mM)	1.66 ± 0.14	10.98 ± 0.40	7.6 ± 0.7
[a]Cu(I)ScAtx1 in presence apoPacS_N at 2:1 ratio	2.53 ± 0.14	10.52 ± 0.44	6.8 ± 0.4
[a]Cu(I)ScAtx1 in presence of apoPacS_N at 1:1 ratio	2.42 ± 0.18	10.90 ± 0.22	7.1 ± 0.4
Cu(I)ScAtx1 in presence of apoPacS_N at 1:1 ratio + 1.0mM BCS	2.62 ± 0.20	7.36 ± 0.35	4.1 ± 0.4

[a] performed at 500MHz.

C. The N-terminal domain of the zinc transporting ATPase ZiaA from *Synechocystis PCC 6803*

C.1 Overview

Despite the essential role of zinc as a structural and catalytic cofactor in numerous metalloproteins, mechanisms of zinc homeostasis in bacteria are poorly understood (1,2). As for other essential metal ions, it is likely that both the uptake and efflux of zinc are tightly regulated in response to availability. Zinc uptake and efflux proteins have recently been identified in several bacteria, including *Synechocystis* PCC 6803 (6), and *Escherichia coli* (3, 4, 5). Expression of these transporters, where known, seems to be regulated by zinc-sensing metalloregulatory proteins (6, 7).

The *ziaA* gene in *Synechocystis* PCC 6803 encodes a polypeptide with sequence similarity to metal-transporting P_1-type ATPases [6]. Transcription of *ziaA* is induced by zinc, but not by other metal ions, and is mediated by the SmtB-related zinc-responsive repressor ZiaR [6]. Mutants deficient in *ziaA* have reduced tolerance to zinc and show reduced export of zinc to the periplasm. *ziaA*-mediated restoration of zinc tolerance can be used as a selectable marker supporting a role as a zinc exporter. *ziaA* mutants also show reduced tolerance to lead but transcription of *ziaA* is not induced by lead, even at inhibitory concentrations, indicating that lead detoxification is not the primary role for ZiaA.

The amino-terminal region of ZiaA contains a single GM*X*C*XX*C metal binding motif within a predicted ferrodoxin-like domain, followed by a region containing seven histidine residues arranged in H*X*H motifs. However, neither the amino-terminal region

$ZiaA_N$ nor $CoaT_N$ (a 38-amino acid region that does not contain an obvious metal-binding motif) formed detectable two-hybrid interactions with Atx1. The detected ScAtx1 interactions are specific to the amino-terminal regions of the copper transporters. There is no free zinc in the cytosol of *E. coli* (8), and it is anticipated that the same is true for *Synechocystis* PCC 6803 and for cobalt, raising questions about what molecules interact with and supply metals ions to $ZiaA_N$, and $CoaT_N$. It is unclear how widespread metallochaperones are and how significant their associations are in defining which inorganic elements are acquired by metalloproteins *in vivo*.

In order to address this question I have performed cloning, expression, purification and structural characterization of N-terminal domain of zinc transporting ATPases $ZiaA_N$ from *Synechocystis* PCC 6803.

C.1.2 References

1. Silver, S. 1996. Transport of inorganic cations, p. 1091-1102. *In* F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: Cellular and molecular biology, 2nd ed. American Society for Microbiology, Washington, D.C.
2. Silver, S., and M. Walderhaug. 1992. *Microbiol. Rev.* 56:195-228.
3. Beard, S. J., R. Hashim, J. Membrillo-Hernandez, M. N. Hughes, and R. K. Poole. (1997). *Mol. Microbiol*. 25:883-891.
4. Patzer, S. I., and K. Hantke.**(**1998) *Mol. Microbiol*. 28:1199-1210.

5. Rensing, C., B. Mitra, and B. P. Rosen. (1997). *Proc. Natl. Acad. Sci. USA* **94:**14326-14331

6. Thelwell, C., Robinson, N.J. and Turner-Cavet, J.S. (1998) *Proc. Natl. Acad. Sci. USA* 95, 10728-10733.

7. Patzer, S. I., and K. Hantke. (1998). *Mol. Microbiol. **28:***1199-1210

8. Outten, C. E., and O'Halloran, T. V. (2001) *Science* **292**, 2488-2492.

C. 2 Cloning, Expression, Purification and Characterization of $ZiaA_N$

Gene encoding the N-terminal tail (1-111 amino acids) of $ZiaA_N$ was PCR amplified from *Synechocystis* PCC 6803 genomic DNA and digested with *Bam*HI/*Xho*I site to create pET$ZiaA_N$. Bacterial clones after ligation were screened by DNA sequencing and positive clone was transformed to BL21 (DE3) expression cells. Cells were grown in minimal medium and were induced at $O.D_{600}$ reaching 0.6 with 1 mM IPTG. Expression test was done to scale up the expression time and was expressed very well after 6 hrs (Fig.28). The entire test expression was done in minifor fermentor by using minimal medium. Protein was purified under native condition by using anion exchange chromatography and size exclusion chromatography in 25 mM Tris buffer pH 9.0. Purity of the protein was checked by SDS-PAGE gel (Fig. 29).

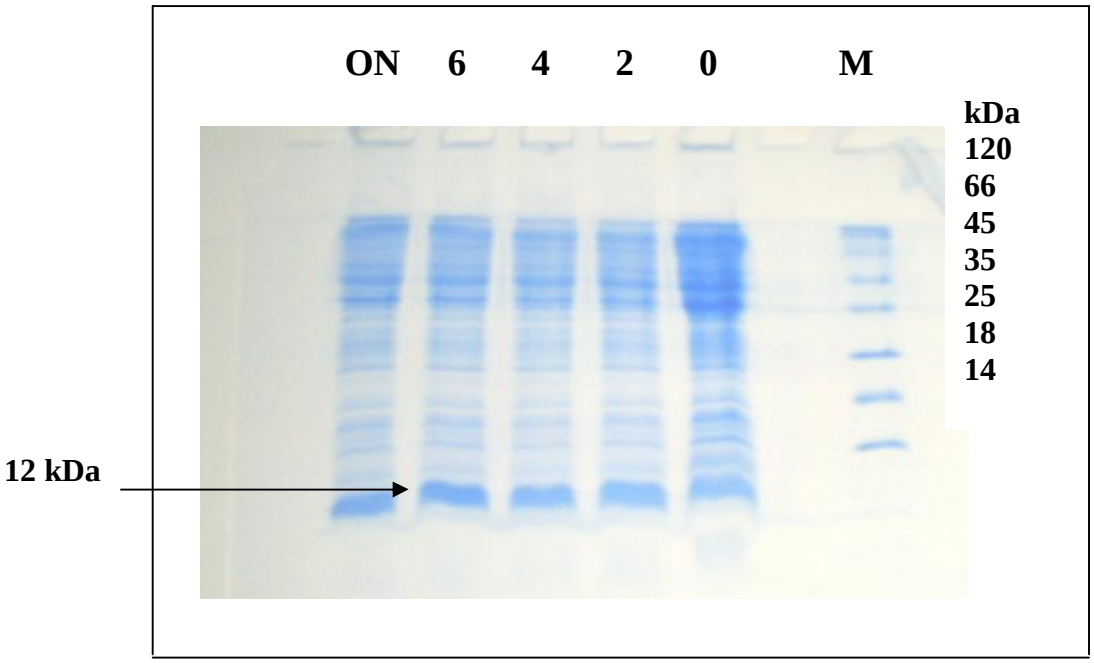


Figure 28. 17% SDS-PAGE profile of over-expression of $ZiaA_N$ at different time intervals (hrs), M-Protein Marker.

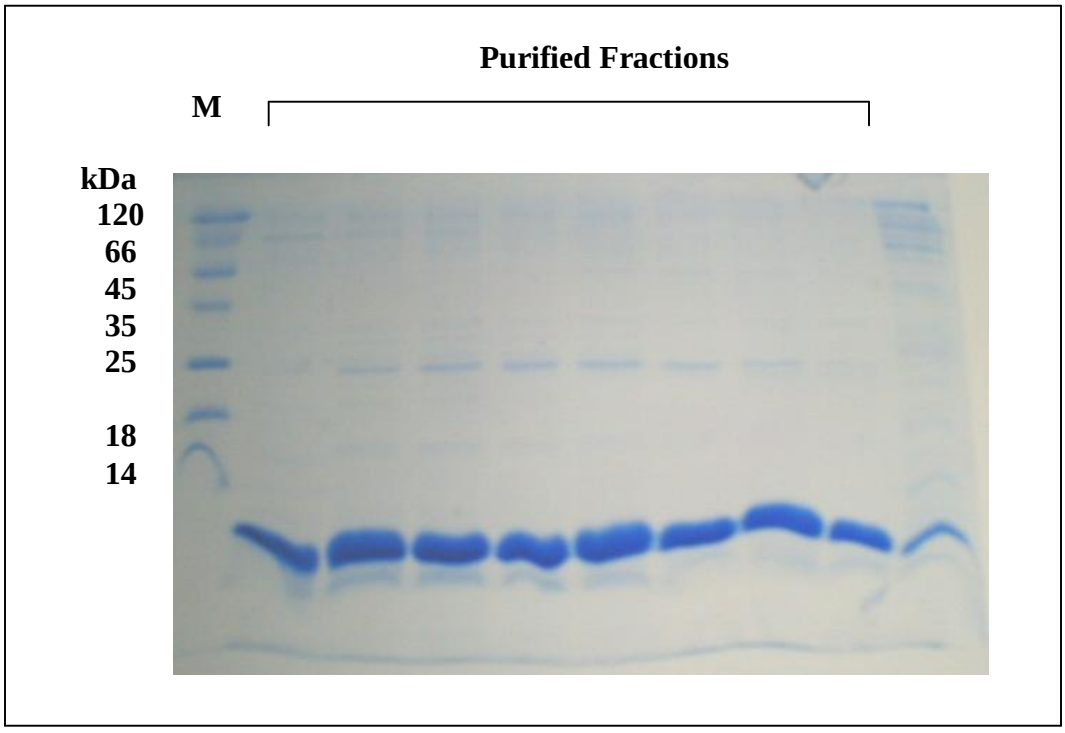


Figure 29.17% SDS-PAGE profile showing the purity of protein after size exclusion chromatography, M-Protein Marker.

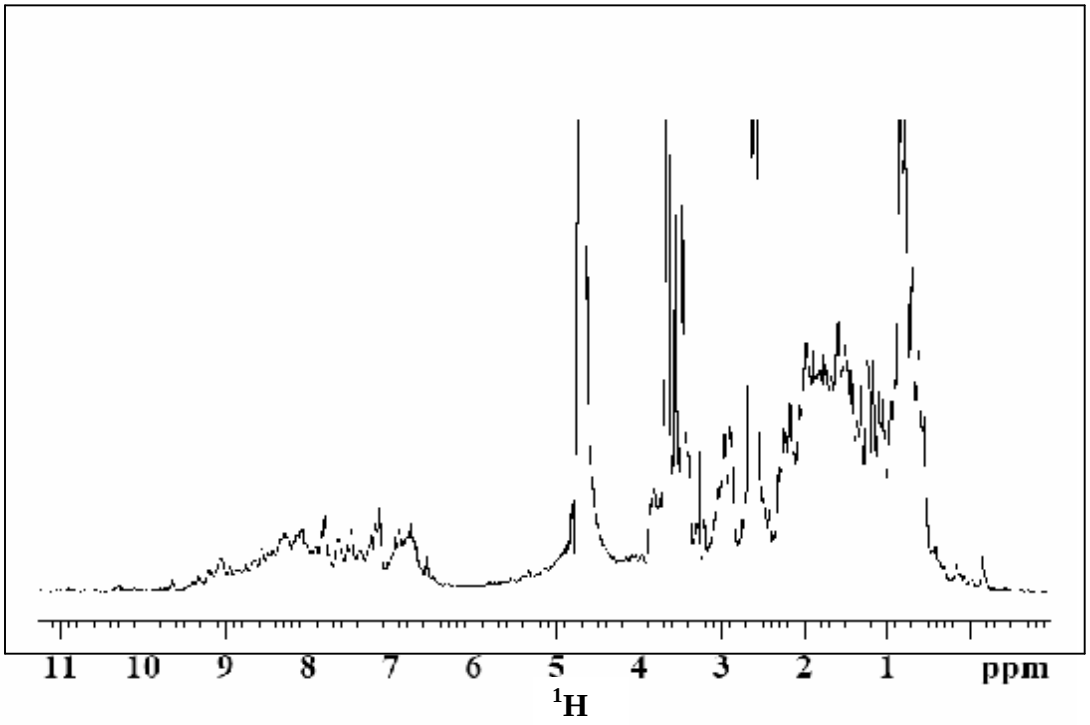


Figure 30. 1D NMR spectra of apoZiaA_N at 800 MHz, 303 K

Protein folding was checked by 1D NMR (Fig. 30) and ^{1}H-^{15}N HSQC (Fig. 31) at 303 K. Results from both spectra shows that protein is well folded. In both cases the protein was reduced with 4 mM DTT. For backbone and sidechain assignment ^{13}C and ^{15}N labeled sample was prepared by using labeled carbon and nitrogen source.

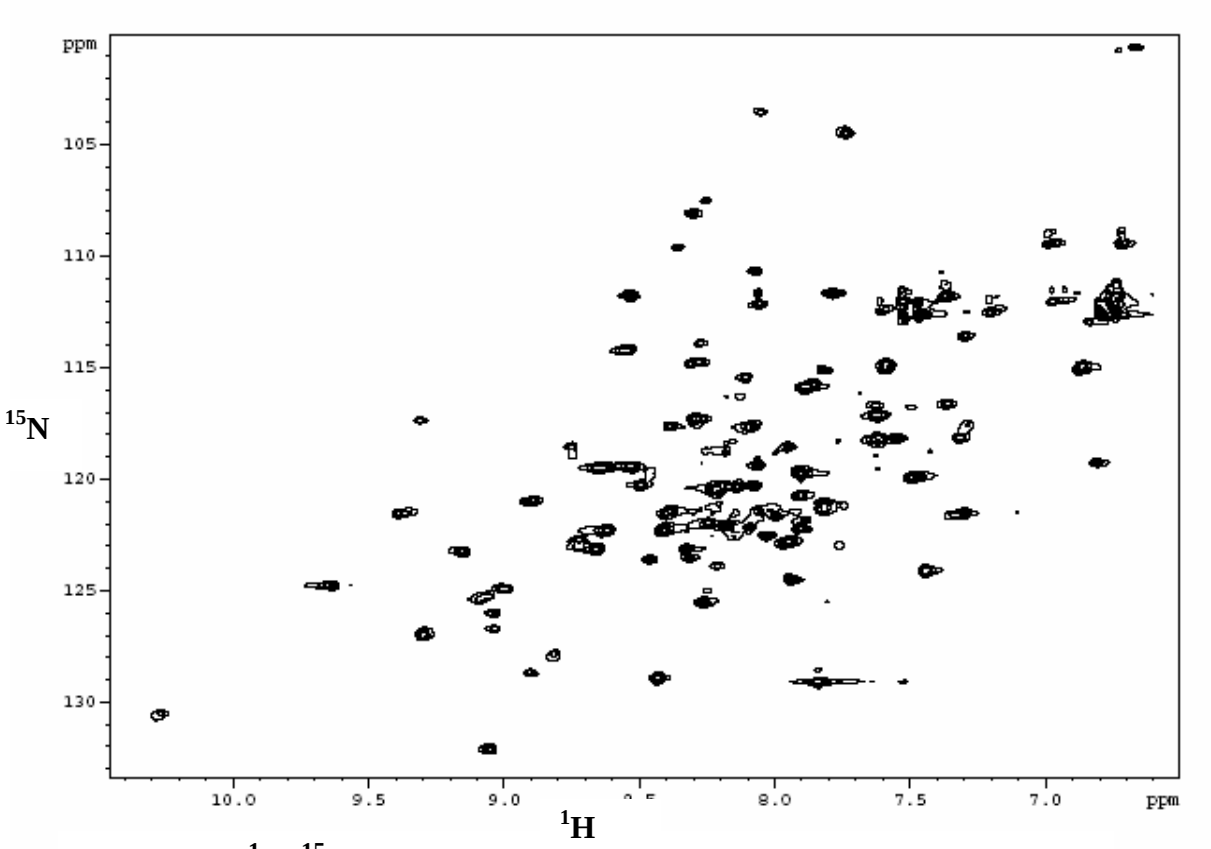


Figure 31. ^{1}H-^{15}N HSQC of apoZiaA_N at 800 MHz, 303 K

C.2.1 NMR experiments and Relaxation studies

Detailed results on NMR experiments and mobility study were mentioned in the following attached section:

Structural Characterization of N-terminal metal binding domain of Zinc transporting ATPases ZiaA from *Synechocystis* PCC 6803

*Lucia Banci, Ivano Bertini, Simone Ciofi-Baffoni, **Murugendra Vanarotti**, Luisa Poggi,*

Nigel J. Robinson

In Preparation

Introduction

Zinc is an essential nutrient: it forms the catalytic center in numerous enzymes and has an important structural role in a wide range of proteins. However, it can be toxic if its level and distribution is not carefully regulated, and its inappropriate binding may compromise cellular function. Therefore, organisms maintain cytoplasmatic zinc concentration at a nontoxic level that is sufficient for growth and homeostatic systems create the cellular environments in which the metal is acquired by metalloproteins {Frausto da Silva, 2001 8060 /id}.

Cyanobacteria have metal requirements often absent in other bacteria. It has been argued that cyanobacteria altered the solubility of metals on a global scale as a by-product of dioxygen-evolving photosynthesis. It is reasonably to suppose that the great autonomy provided by photosynthesis allowed early cyanobacteria to colonize vast vacant habitats of anaerobic earth some 2.7×10^9 years ago, life previously being restricted to niches with exploitable sources of chemical energy. In the absence of competition, rapid evolutionary divergence of ancestral cyanobacteria was likely to occur under minimal constraint {Cavet, 2003 8595 /id}. Moreover, at a time when cyanobacteria were evolving rapidly, the requirements for metal homeostasis were changing swiftly {Frausto da Silva, 2001 8060 /id}.

Metal-transporting P_1-type ATPases include numerous bacterial transporters commonly associated with resistance to excess Cd^{2+}, Zn^{2+}, Pd^{2+}, Ag^{2+}, Co^{2+} and copper. Features of this family of ATPases are an intramembranous metal binding site (the CPC motif) and a soluble amino-terminal metal-binding domain typically containing the motif CXXC.

The *ziaA* gene in *Synechocystis* PCC 6803 encodes a polypeptide with sequence similarity to metal-transporting P_1-type ATPases {Thelwell, 1998 8988 /id}. Transcription of *ziaA* is induced by zinc, but not by other metal ions, and is mediated by the zinc-responsive repressor ZiaR {Thelwell, 1998 8988 /id}.Mutants deficient in *ziaA* have reduced tolerance to zinc and show reduced export of zinc to the periplasm. *ZiaA*-mediated restoration of zinc tolerance can be used as a selectable marker supposing a role as a zinc exporter {Thelwell, 1998 8988 /id}. It has been shown that this P_1-type ATPases, ZiaA, transports zinc and not copper {Thelwell, 1998 8988 /id} and has an amino-terminal domain ($ZiaA_N$) with higher affinity for copper than for zinc (4).

High metal specificity is a common feature of all the four related P_1-type ATPases found in *Synechocystis* PCC 6803: CtaA, CoaT, PacS and ZiaA. A model for the action of each of these ATPases is pictured in **Figure 1**. At present, the metal ion transported and the direction of transport cannot be predicted from the sequence of the P_1-type ATPase {Tottey, 2001 7684 /id}.

Here we want to investigate the structural basis of the observed selective metal acquisition/exclusion by $ZiaA_N$.

Materials and Methods

Cloning, Production, and Purification of Recombinant $ZiaA_N$—

Synechocystis PCC 6803 genomic DNA was used as template for PCR with primers 5'-GGATCCATGACCCAATCTTCACCGCTCAAAAC-3' with 5'-CTCGAGTAGTTCT TGTTTCAGATTAAATTC-3' (the latter annealing to DNA 3' of the ZiaA stop codon). The amplified fragment of DNA containing codons 1-111 encoding the entire amino-terminal region of ZiaA **(Fig. 2)** was ligated into the *Nde*I/*Eco*RI sites of pET29a to

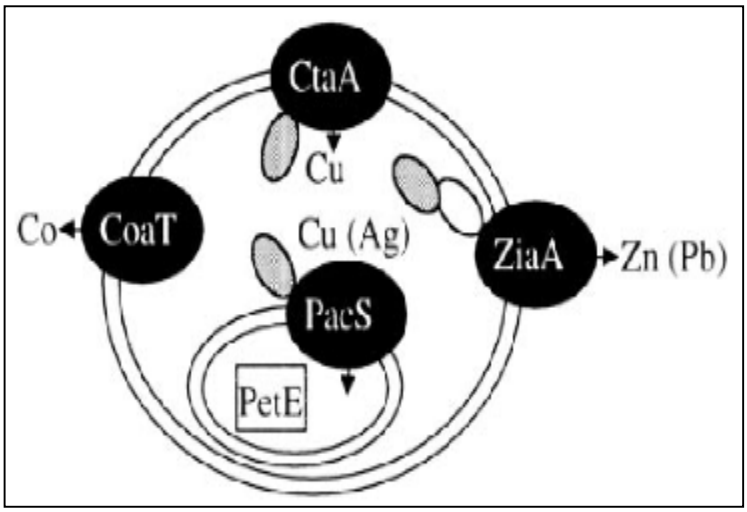


Figure 1. (A) A model for the action of each of the four related CPxt-type transporters in *Synechocystis* PCC 6803. *Arrows* indicate the direction of transport. Metals for which homeostasis is known to be altered following deletion of the respective transporter are shown. The model locates PacS within thylakoid membranes supplying copper for plastocyanin, PetE, in the thylakoid lumen. *Shaded ovals* represent a single amino-terminal metal-associated motif (C*XX*C), and a *nonshaded oval* (ZiaA) represents an additional H*X*H motif. (B) Sequence of the soluble amino-terminal region of ZiaA ($ZiaA_N$).

create pETZiaA. Recombinant protein was generated in *E. coli* (BL21) exposed to zinc (75 μM). Lysates were applied to Hi-Trap Q XL (Amersham) column and protein was eluted with one step 100 ml 50 mM NaCl. Fractions were collected and checked on 17% SDS-PAGE gel and low molecular weight protein fractions were applied on Sephadex G-75 (1.5 × 20 cm), and fractions were eluted in 50 mM potassium phosphate buffer, pH 7.0. A single prominent band of the anticipated size was detected by PAGE. An aliquot of protein was hydrolyzed and analyzed for amino acid composition (Alta Bioscience) to allow calibration of colorimetric estimation of $ZiaA_N$.

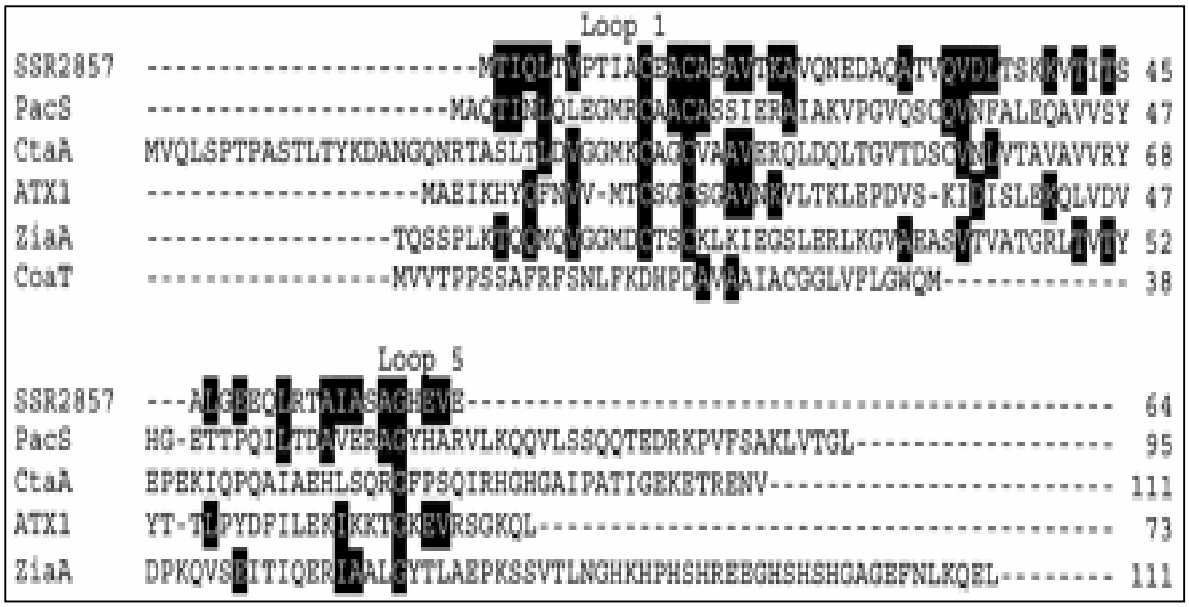


Figure 2. Alignment of predicted hydrophilic amino-terminal regions of the CPx-type transporters PacS, CtaA, ZiaA, and CoaT from *Synechocystis* PCC 6803.

NMR Experiments.

NMR spectra were acquired on Avance 800, 700, 600 and 500 Bruker spectrometers operating at proton nominal frequencies of 800.13, 700.13, 600.13 and 500.13 MHz, respectively. All the triple resonance (TXI 5-mm) probes used were equipped with pulsed field gradients along the *z* axis. The 800 and 500 MHz spectrometers were equipped with a triple resonance cryoprobe.

The NMR experiments were recorded on ^{15}N,^{13}C-labeled and ^{15}N-labeled apoZiaA$_N$ samples. The NMR experiments used for the backbone and the aliphatic side-chain assignment and for obtaining structural restraints are summarized in **Table 1**. The ^{1}H, ^{13}C and ^{15}N resonance assignments of apoZiaA$_N$ are reported in **Table 1.**

Experiments	Dimension of acquired data (Nucleus)			Spectral Width (ppm)			
	t1	t2	t3	F1	F2	F3	NS
[^{1}H-^{1}H]-NOESY	1024(H)	2048(H)		15	15		64
[^{1}H-^{1}H]-TOCSY	1024(H)	2048(H)		15	15		64
[^{1}H-^{15}N]-HSQC	512(^{15}N)	1024(H)		40	7		8
[^{1}H-^{13}C]-HSQC	256(^{13}C)	2048(H)		70	14		8
CBCA(CO)NH	124(^{13}C)	48(^{15}N)	1024(H)	88	40	16	8
CBCANH	124(^{13}C)	48(^{15}N)	1024(H)	88	40	16	8
HNCO	64(^{13}C)	48(^{15}N)	1024(H)	16	40	12	8
HN(CA)CO	64(^{13}C)	48(^{15}N)	1024(H)	16	40	12	16
^{13}C(H)CCH-TOCSY	272(^{13}C)	68(^{13}C)	1024(H)	88	88	12	16
^{15}N-edited-[^{1}H-^{1}H]-NOESY	272 (^{1}H)	48(^{15}N)	1024(H)	15	40	15	16
^{13}C-edited-[^{1}H-^{1}H]-NOESY	272 (^{1}H)	96(^{13}C)	1024(H)	13	86	13	16
HNHA	128 (^{1}H)	40(^{15}N)	1024(H)	15	40	15	16

Table 1. Acquisition parameters for NMR experiments performed on apoZiaA_N.

A structural model of the soluble domain of ZiaA, which shows a sequence identity with ZntA (32%), was calculated using the program MODELLER, version 4.0{Sali, 1993 4645 /id} and is shown in **Figure 3.** ^{15}N R_1, R_2 and steady-state heteronuclear NOEs were measured using the pulse sequences proposed by Farrow *et al.* {Farrow, 1994 7528 /id}. R_2 were measured using a refocusing delay of 450 µs. In all experiments the water resonance was suppressed with a "water flip-back" scheme {Grzesiek, 1993 3805 /id}.

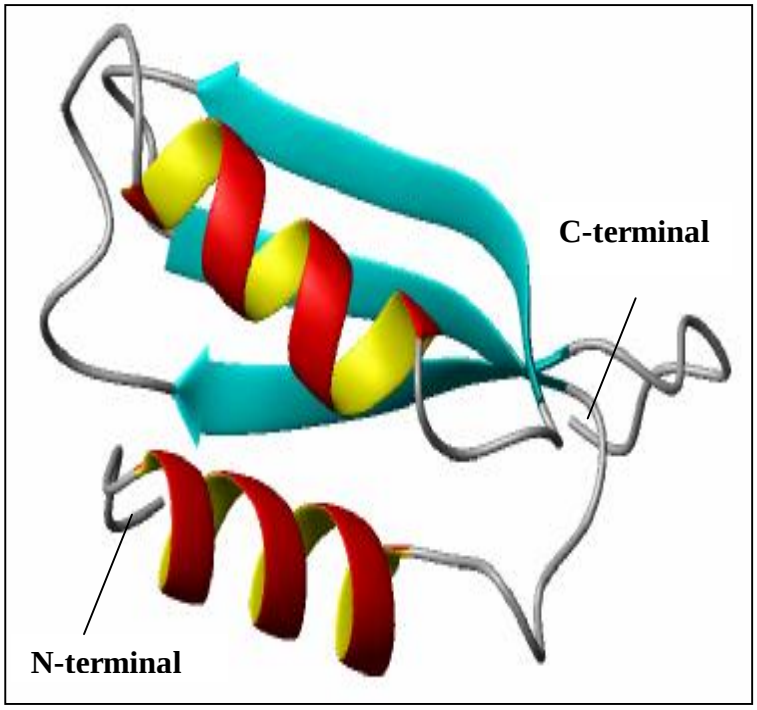


Figure 3. Structural model of the soluble domain of ZiaA (residues 1-73), which shows a sequence identity with ZntA (32%), calculated using the program MODELLER, version 4.0{Sali, 1993 4645 /id}

The experimental relaxation rates were used to map the spectral density functions, $J(\omega_H)$, $J(\omega_N)$ and $J(0)$ following the procedure reported in literature {Peng, 1992 4549 /id}.Fast amide proton exchange rates were measured using ^{15}N-(CLEANEX-PM)-FHSQC

{Hwang, 1998 4450 /id;Hwang, 1998 4448 /id} experiments with a mixing time of 100 ms

Calculation of Chemical Shift Index (CSI) for apoZiaA$_N$

Once the assignment of the backbone atoms had been completed, the chemical shift values of $^{1}H^{\alpha}$, $^{13}C^{\alpha}$, $^{13}C^{\beta}$ and ^{13}CO were used for the chemical shift index (CSI) analysis according to Wishart and Sykes [13]

Results and Discussion

Backbone Assignment and Dynamic Characterization of apoZiaA$_N$.

The ^{1}H-^{15}N HSQC spectrum of apoZiaA$_N$ show well dispersed resonances indicative of an essentially folded protein **(Figure 4)**. However, only 62% (69 out of 111) of the expected ^{15}N backbone amide resonances were possible to assign for apoZiaA$_N$.

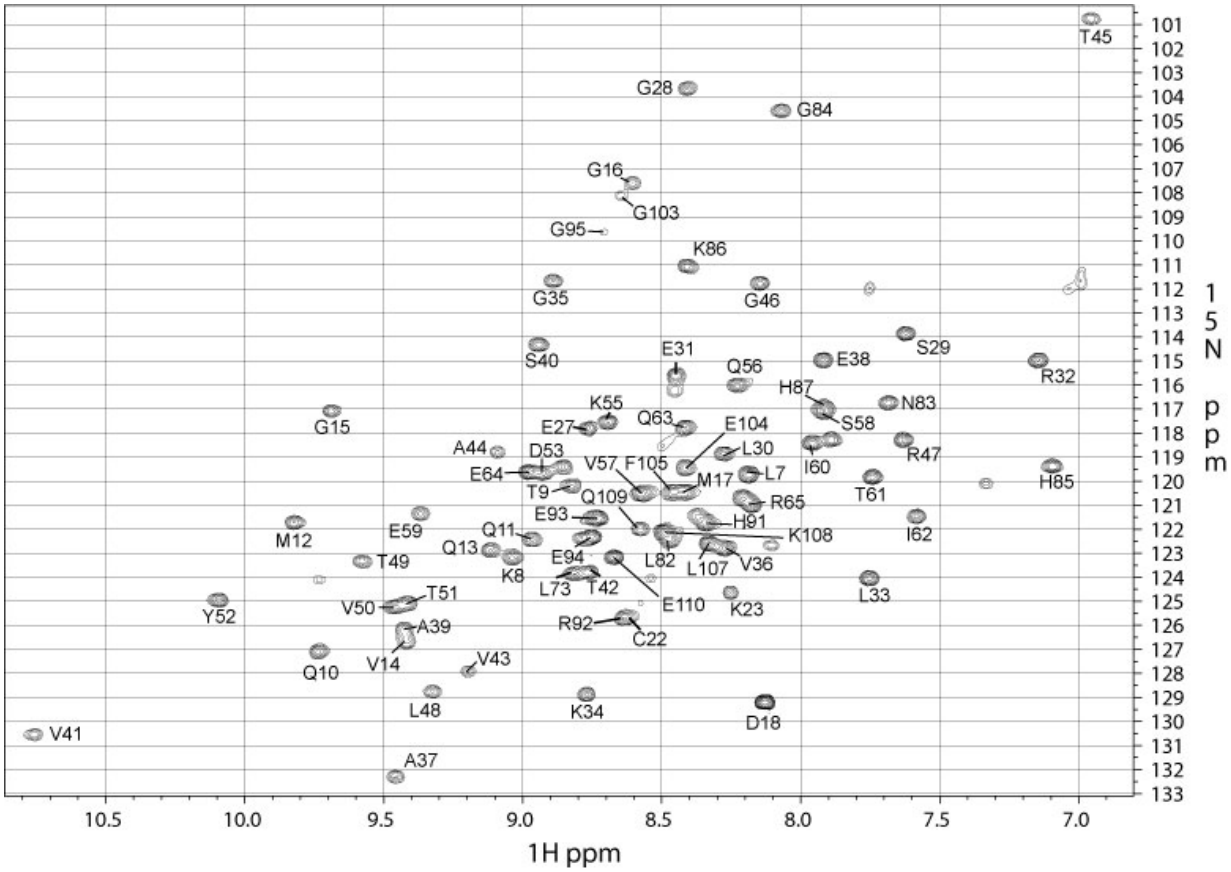


Figure 4. ^{1}H-^{15}N HSQC spectrum of apoZiaA$_N$ recorded at 700 MHz, 303 K, on a 1.0 mM sample in 20mM phosphate buffer pH 7 with 5 mM DTT.

^{15}N R_1, R_2 and ^{1}H-^{15}N NOE values can provide information on internal mobility as well as on the overall protein tumbling rate. From the analysis of 60 backbone HN signals, which are well resolved in the ^{1}H-^{15}N HSQC spectra of apoZiaA_N, average ^{15}N R_1, R_2 and ^{1}H-^{15}N NOE values of 2.00 ± 0.23, 8.22 ± 1.81 and 0.42 ± 0.68 s^{-1} are found, respectively, at 600 MHz. The experimental relaxation data are shown in **Figure 5** and look essentially homogeneous along the polypeptide chain, with the exception of residues located at the C and N termini. The correlation time for the molecule tumbling (τ_m) as estimated from the R_2/R_1 ratio is 5.69 ± 0.93 ns, as expected for a protein of this size in a monomeric state {Stokes, 1956 258 /id}{Einstein, 1956 261 /id}.

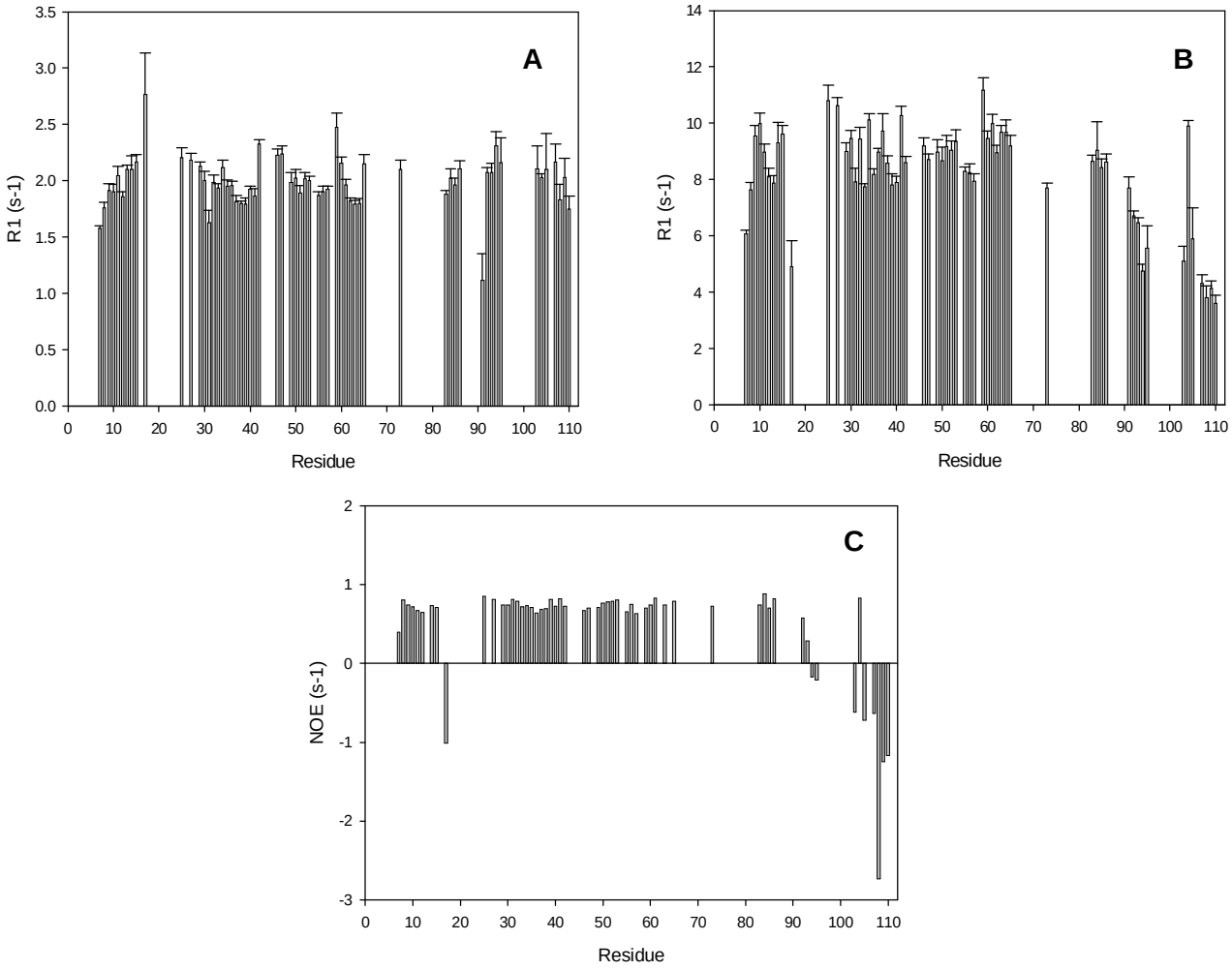


Figure 5. Experimental relaxation data R_1, R_2 and heteronuclear NOEs measured at 600 MHz for apoZiaA_N at 303 K on a 1.0 mM sample in 20mM phosphate buffer pH 7 with 5 mM DTT.

Conformational exchange phenomena that open parts of the protein backbone and expose it to bulk solvent are expected to give rise to relatively fast amide proton exchange kinetics. Residues affected by such amide proton exchange with solvent can be determined by the use of the ^{15}N-(CLEANEXPM)-FHSQC approach, as shown in **Figure 6**. A remarkable amount of assigned peaks clearly shows fast solvent exchange, as well as a few new peaks appearing only in the CLEANEX spectrum.

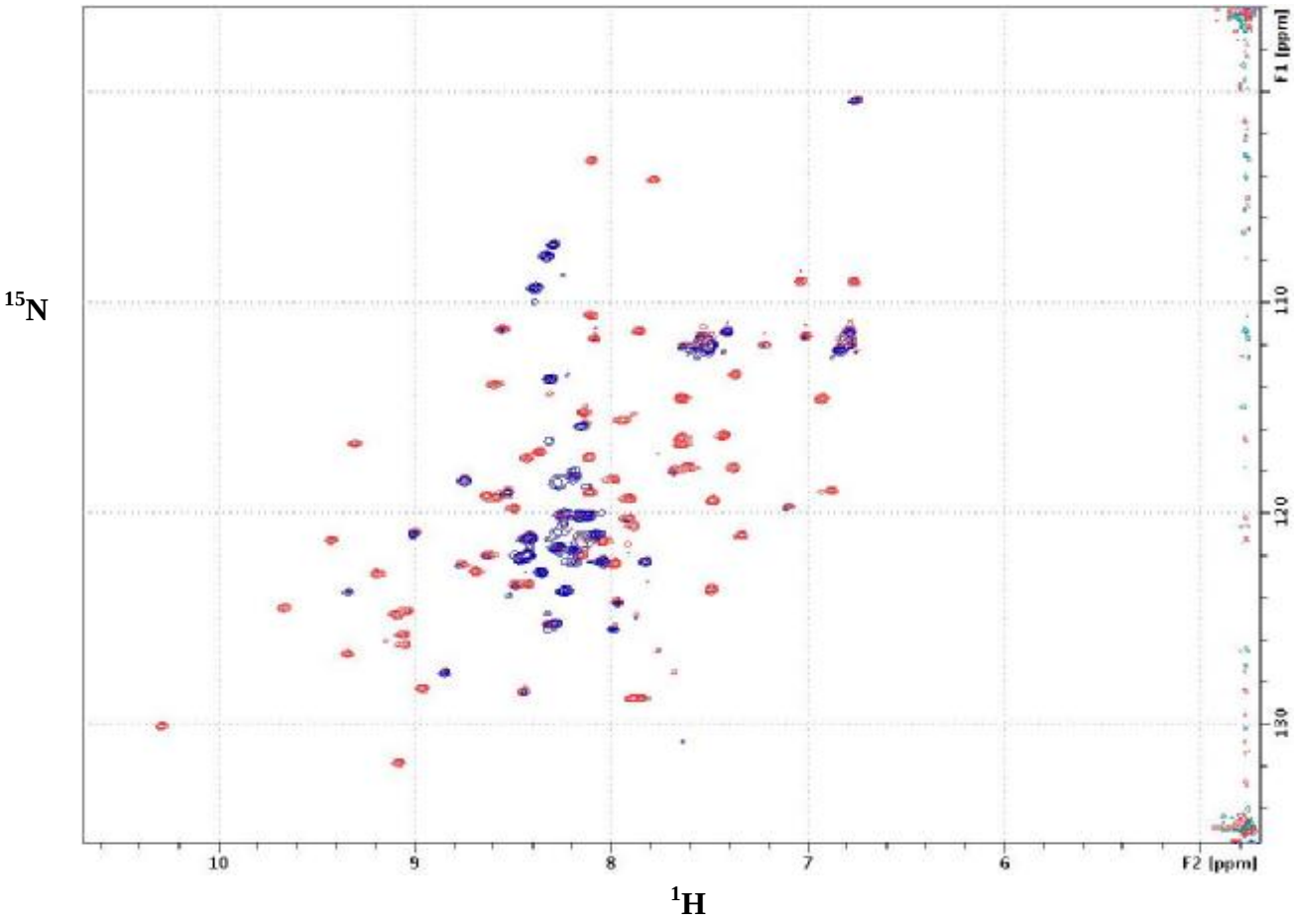


Figure 6. Superposition of a 1H-15N HSQC spectrum (red) and a ^{15}N-(CLEANEXPM)-FHSQC spectrum (blue) of apoZiaA$_N$ recorded at 700 MHz, 303 K, on a 1.0 mM sample in 20mM phosphate buffer pH 7 with 5 mM DTT.

CSI analysis has been performed (**Fig.7**) .It results that $ZiaA_N$ possesses βαββα secondary structural elements, in agreement with the derived structural model and as expected for N-terminal domains of P1-type of ATPases (4).CSI analysis shows that the last b strand is absent, at variance of the typical ferredoxin-like fold of homologous domains (4).

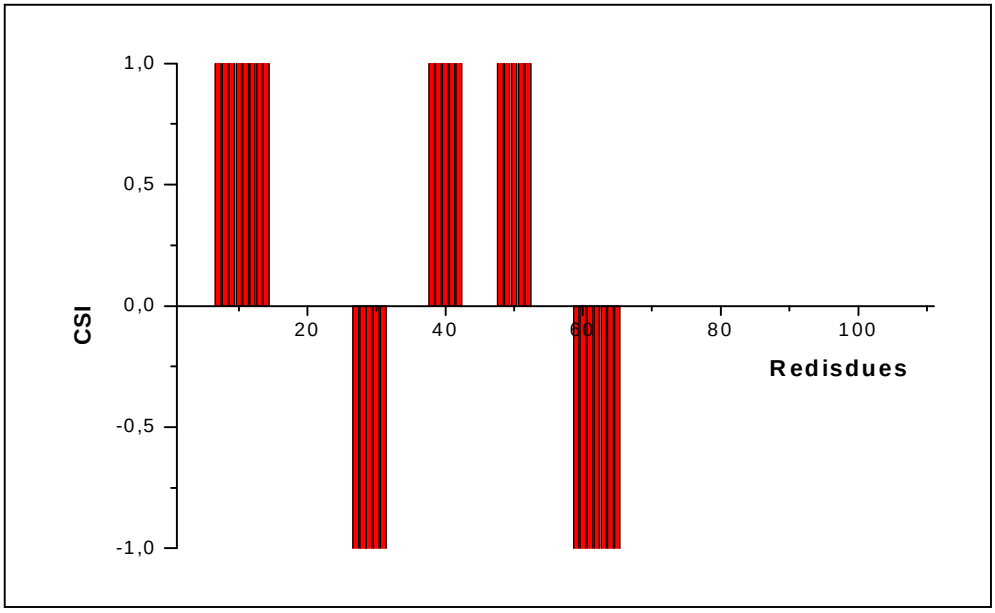


Figure 7. Consensus CSI result obtained from $^{1}H^{\alpha}$, $^{13}C^{\alpha}$, $^{13}C^{\beta}$ and ^{13}CO chemical shifts apo$ZiaA_N$, where an index of +1 or -1 indicates elements of β-strand or α-helical structure, respectively, while zero value predicts a random coil conformation.

References

1. Fraústo da Silva, J.J.R. and Williams, R.J.P. (2001) The Biological Chemistry of the Elements: The Inorganic Chemistry of Life, 2nd edn. Clarendon Press, Oxford.
2. Robinson, N.J., Rutherford, J.C., Pocock, M.R. and Cavet, J.S. (2000) Metal metabolism and toxicity: Repetitive DNA. In: The Ecology of Cyanobacteria (Whitton, B.A. and Potts, M., Eds.), pp. 443–463. Kluwer Academic, Dordrecht.
3. C. Thelwell, N.J. Robinson and J.S. Turner-Cavet, (1998) *Proc. Natl. Acad. Sci. USA* **95**. 10728–10733.
4. Tottey S, Rondet SA, Borrelly GP, Robinson PJ, Rich PR, Robinson NJ. (2002) *J Biol Chem*. 15; 277(7):5490-7.
5. Tottey, P.R. Rich, S.A. Rondet and N.J. Robinson,(2001). *J. Biol. Chem.* **276** 19999–20004.
6. Sali A, Blundell TL.(1993) *J Mol Biol*. 5; 234(3):779-815.
7. Grzesiek S, Vuister GW, Bax A.(1993) *J Biomol NMR* .l;3(4):487-93.
8. Farrow NA, Muhandiram R, Singer AU, Pascal SM, Forman-Kay JD, Kay LE. (1994) *Biochemistry,* 17; 33(19):5984-6003.
9. Stokes, G. (1956) *Trans. Cambridge Philos. Soc.* **9,** 5-10.
10. Einstein, A. (1956) *Investigations on the Theory of the Brownian Movement*, Dover, New York.
11. Peng, J. W., and Wagner, G. (1992) *J. Magn. Reson.* **98,** 308-332.
12. Hwang, T. L. & Shaka, A. J. (1998) *J. Magn. Reson.* **135,** 280–287.
13. Wishart, D.S. & Sykes, B.D. (1994) *J. Biomol. NMR* **4**, 171–180
14. Outten, C. E., and O'Halloran, T. V. (2001) *Science* **292**, 2488-2492.

D. Zinc Finger 9 /Cellular Nucleic Acid Binding (CNBP/ZNF9)

D.1 Overview

Myotonic dystrophy is autosomal-dominant multisystemic diseases (1) with peculiar phenotypic features affecting skeletal muscle, heart, eye and endocrine system. There are two kind of Myotonic dystrophy, namely type 1(DM1) and type 2 (DM2). DM1 is caused by an expanded CTG repeat in the 3′ UTR of the dystrophia myotonica protein kinase gene (DMPK) on chromosome 19q13.3 (2) and the protein involve is MBNL. Where as DM2 is caused by a dominantly transmitted CCTG repeat expansion in intron 1 of the zinc finger protein 9 (ZNF9) gene on chromosome 3q (3). In both cases the number of repeats depends on the severity of disease. In DM2 these repeat ranges in size from 5–37 repeats in the normal population, 50–1000 repeats in mild to classical symptom adult-onset patients with premature cataract formation, myotonia and muscle weakness, and >1000 repeats in severely affected congenital individuals (CDM) with neonatal respiratory distress, hypotonia and mental retardation. Huntington's disease is example for this class of disease which is caused by expansion of CAG repeat in a region of the gene that is translated into protein. The accumulation of these transcripts in numerous intranuclear foci has led to the search for proteins that interact with the repeats. One such protein is ZNF9, which co-localize with RNA foci in the cells (3). There is no structural information available either on MBNL or ZNF9.

This work includes cloning, expression, purification and characterization of ZNF9 protein.

D.1.1 References

1 Ranum, LP and J.W. Day, (2002). *Curr Neurol Neurosci Rep* 2, 465–470.

2. Christina L. Liquori, Kenneth Ricker, Melinda L. Moseley, Jennifer F. Jacobsen, Wolfram Kress, Susan L. Naylor, John W. Day, Laura P. W. Ranum (2001), *Science*, 293, 5531, 864-867.

3. Yoshihiro Kino, Daisuke Mori, Yoko Oma, Yuya Takeshita, Noboru Sasagawa and Shoichi Ishiura (2004), *Human Molecular Genetics*, 13, 5: 495-507.

D. 2 Cloning, Expression, Purification and Characterization of ZNF9

Protein was purified from under denaturing condition by utilizing N-terminal His-Tag of the protein. Impurities after first column were removed by size exclusion chromatography. Protein was enough pure after first column (Fig. 32).

Protein was refolded in the presence of several additives and additives were removed in several steps of slow dialysis. Primary protein folding checked by 1D NMR shows that the protein is partially folded (Fig. 33). ZNF9 was titrated with Zn^{2+} and 1D NMR spectra was collected during the titration steps. The addition of Zn^{2+} does not dramatically improve the fold of the protein (Fig.34), suggesting that the protein is not well structured in the absence of RNA repeats.

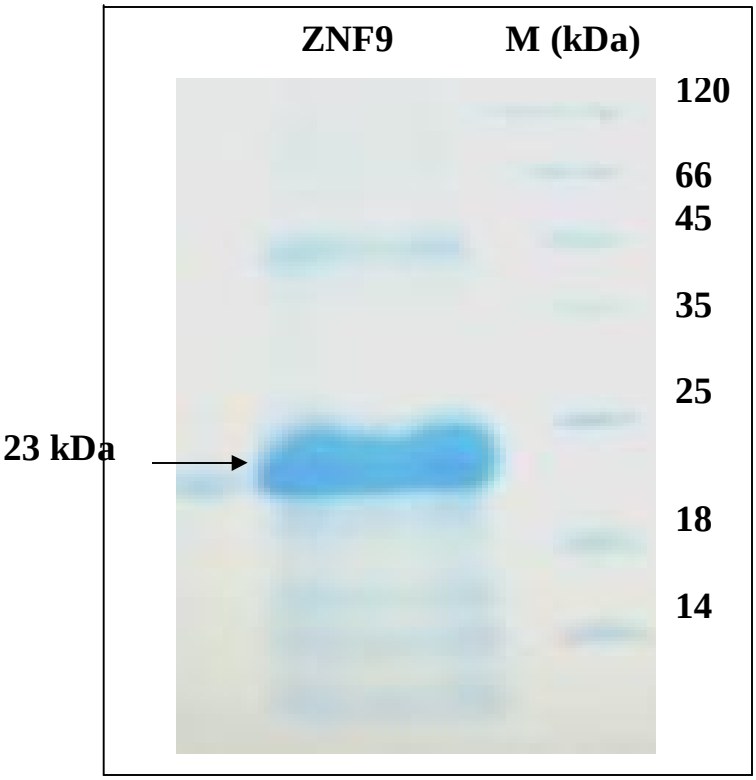


Figure 32. 17% SDS-PAGE of ZNF9 after size exclusion chromatography.

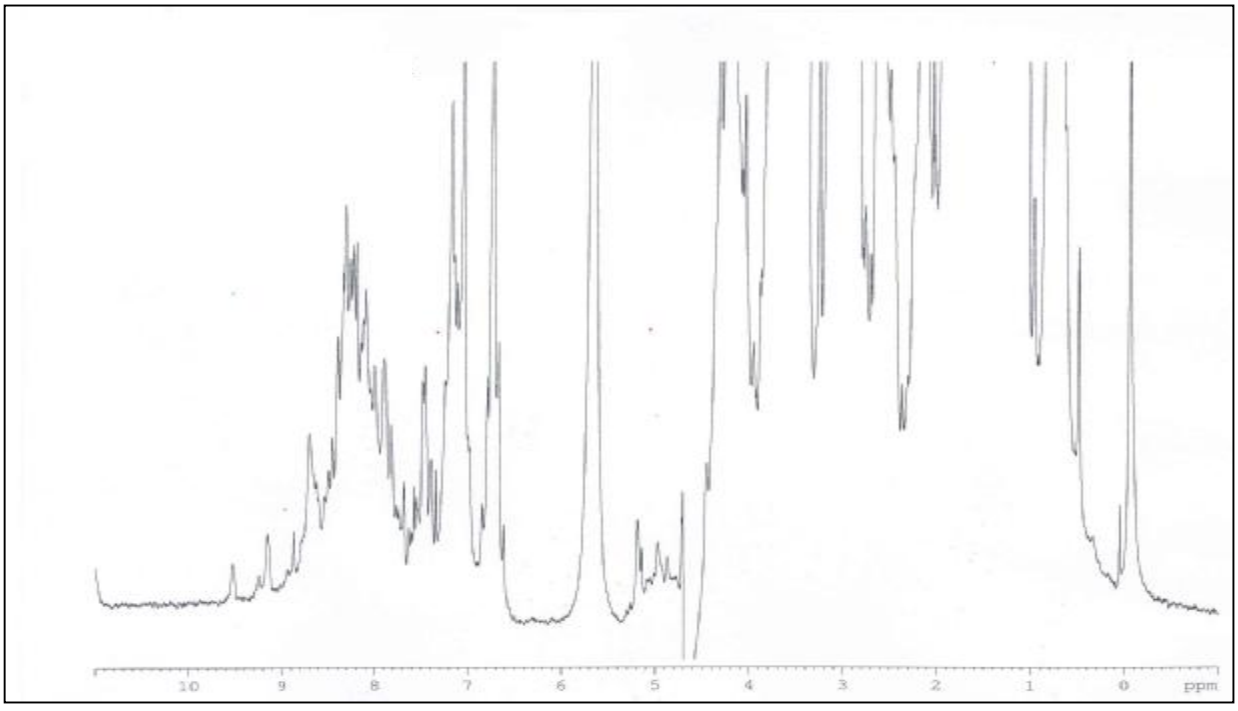


Figure 33. 1D NMR spectra of ZNF9 protein.

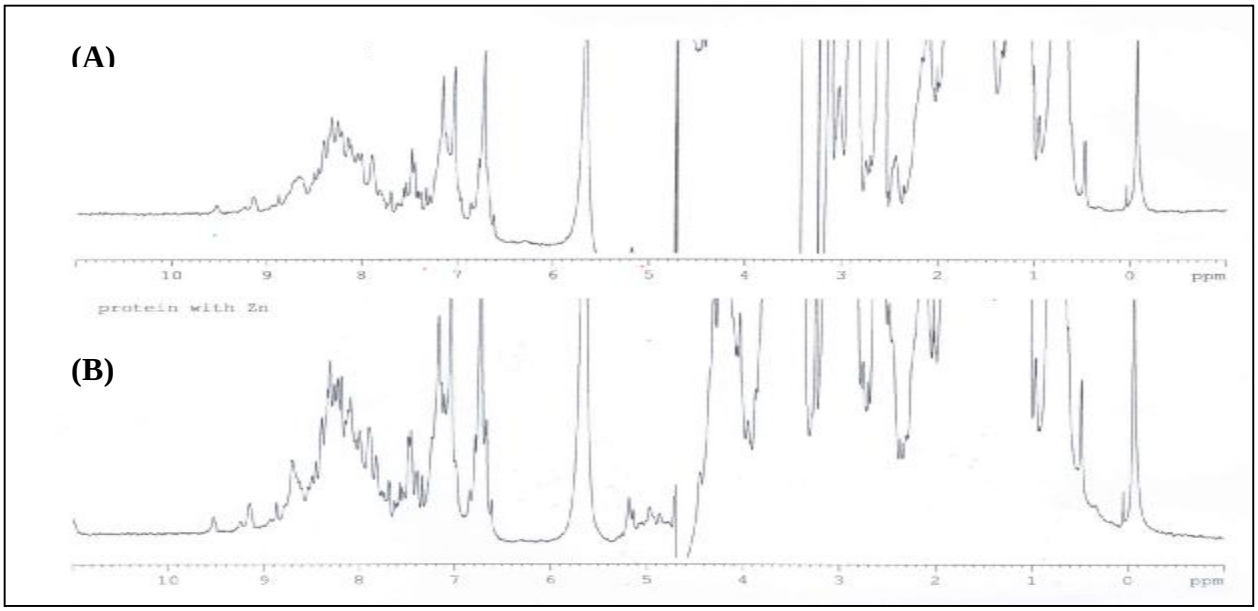


Figure 34. 1D NMR spectra of ZNF9 before (A) and after (B) adding zinc (1:7)

E. Catalytic Domain of A Disintegrin and Metalloprotease 10 (ADAM10)

E.1 Overview

The ADAM (A Disintegrin And Metalloprotease) family includes proteins containing disintegrin-like and metalloprotease-like domains (1). They are also referred to as MDC (Metalloprotease, Disintegrin, Cysteine-rich) proteins. The extracellular domains of a diverse range of cell proteins are cleaved and released from cells in a soluble form by a process known as ectodomain shedding. Membrane anchored proteins that undergo such cleavage include cytokines, cytokine receptors, growth factors, growth factor receptors, cell adhesion molecules (1), and proteins of unknown function, including the amyloid precursor protein (APP) (2). The ectodomain shedding of these proteins

exhibits three main characteristics. First, shedding is enhanced by agents, such as phorbol esters, that activate protein kinase C (3, 4). Second, shedding is sensitive to hydroxamic acid-based metalloproteinase inhibitors (5, 6). Third, shedding appears to occur at or near the cell surface (7). Proteolytic processing of the APP appears to be central to the etiology of Alzheimer's disease (AD). Recently it was reported that metalloprotease, ADAM10 is involved in this process. However, it has been suggested that inhibition metalloprotease (ADAM10) could have value as an alternative strategy in the treatment of Alzheimer's disease (8).

To address this challenge I have attempted to clone and express catalytic domain of ADMA10 protein.

E.1.1 References

1. Arribas, J., Coodly, L., Vollmer, P., Kishimotot, T. K., Rose-John, S., and Massagué, J. (1996) *J. Biol. Chem.* **271,** 11376–11382.
2. Arribas, J., and Massagué, J. (1995) *J. Cell Biol.* **128,** 433–441.
3. Ehlers, M. R. W., and Riordan, J. F. (1991) *Biochemistry* **30,** 10065–10074.
4. Massagué, J., and Pandiella, A. (1993) *Annu. Rev. Biochem.* **62,** 513–541.
5. Parvathy, S., Hussain, I., Karran, E. H., Turner, A. J., and Hooper, N. M. (1998) *Biochemistry* **37,** 1680–1685.
6. Amit, T., Hochberg, Z., Yogev-Falach, M., and Youdim, M. B. H. (2001) *J.Endocrinol.* **169,** 397–407.
7. Teixido, J., Wong, S. T., Lee, D. C., and Massagué, J. (1990) *J. Biol. Chem.* **265,** 6410–6415

8. Benjannet, S., Elagoz, A., Wickham, L., Mamarbachi, M., Munzer, J. S., Basak, A., Lazure, C., Cromlish, J. A., Sisodia, S., Checler, F., Chrétien, M., and Seidah, N. G. (2001) *J. Biol. Chem.* **276,** 10879–10887.

E. 2 Cloning, Expression, Purification and Characterization of ADAM10

Catalytic domain of ADAM10 was expressed in BL21 AI cells. Protein was expressed as soluble form. Purification was done by using His-tag and protein was enough pure after first column (Fig.35). His -tag was cleaved by using Factor Xa enzyme.

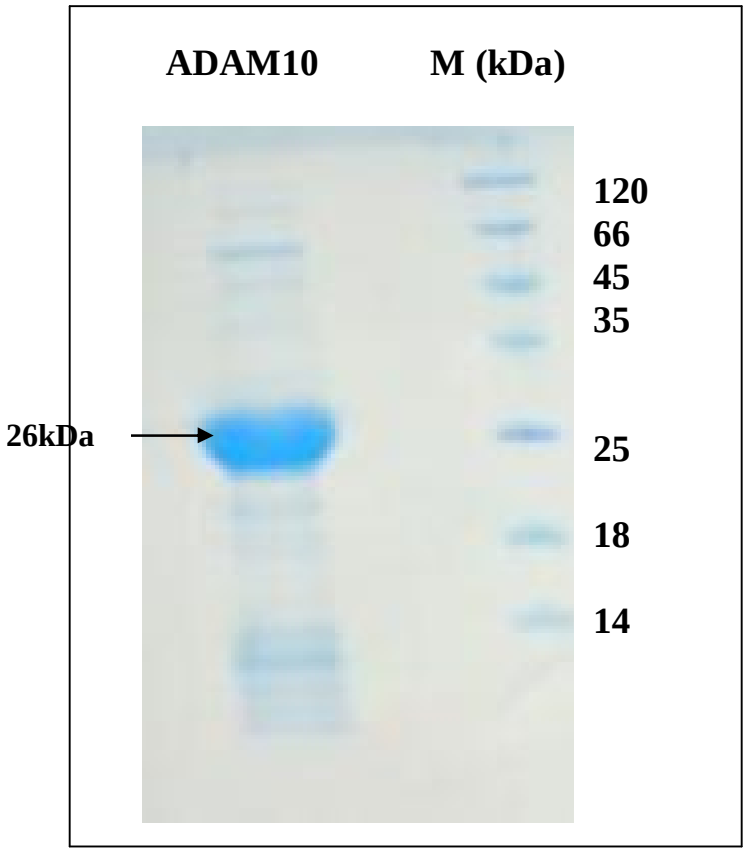


Figure 35. 17% SDS-PAGE of ADAM10 after Hi-Trap affinity column.

^{15}N labeled sample for NMR was prepared by using M9 medium. Folding of ADAM10 was checked by 1D NMR and 1H-15N HSQC spectra. Protein results well folded. ^{1}H-^{15}N NMR spectra showed (Fig. 36) indeed well dispersed number of peaks. However the protein is very unstable in these conditions; it completely precipitates after 6-8 hrs. This protein has also high auto-proteolytic activity and screening to find an efficient inhibitor is now under investigation.

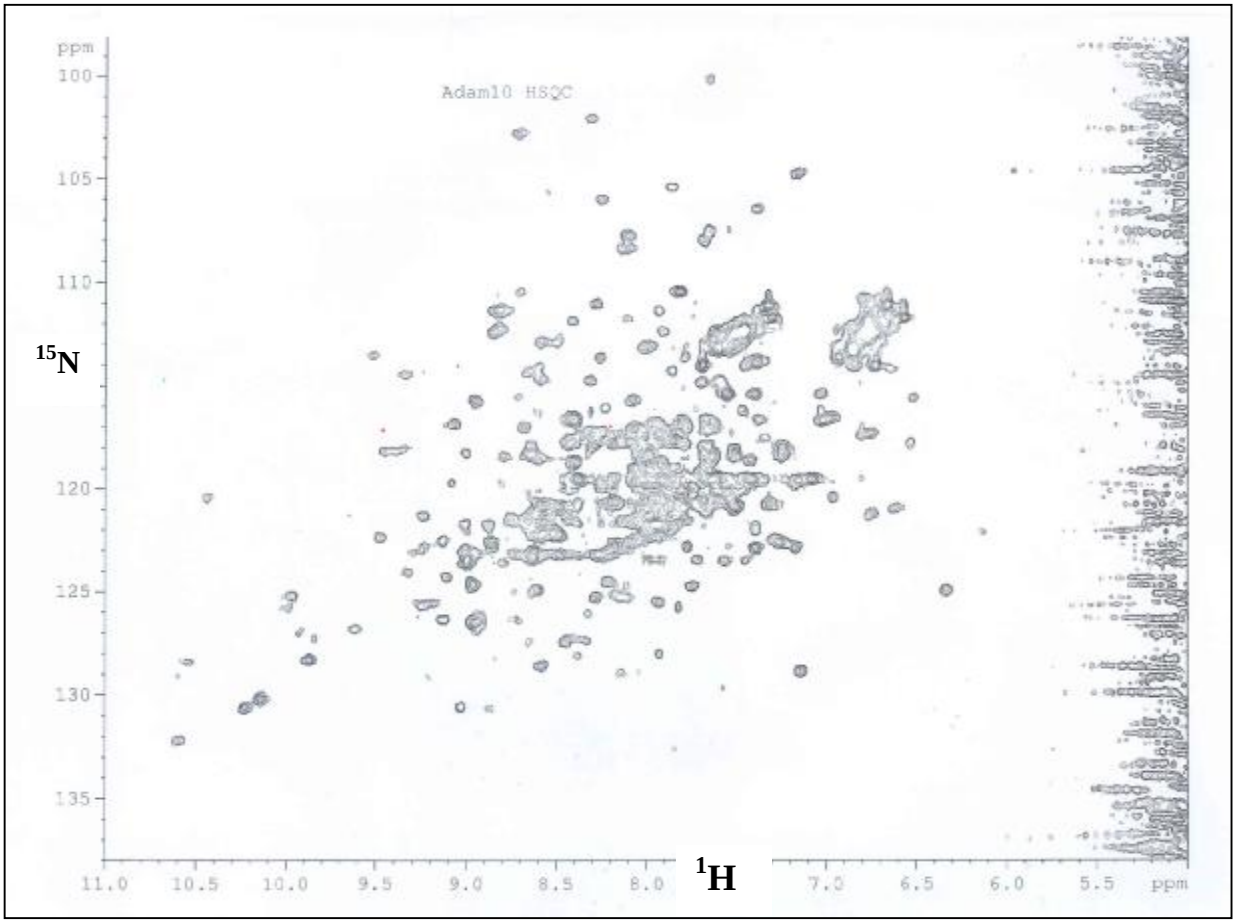


Figure 36.^{1}H-^{15}N HSQC of ADAM10 at 800 MHz with 298 K

GENERAL DISCUSSION AND PERSPECTIVES

S*tructural biology* provides a detailed analysis at atomic level of the structure of biological macromolecules, with the ultimate goal of investigating the relationship between structure and function. In the case of metalloproteins one of the most relevant aspects is to understand the role of the metal cofactor and to investigate the structural changes induced by the binding of the metal ion. ***Structural biology*** of metalloproteins can be utilized to explore the numerous modes of coordination of metals found in the protein matrix, for example as found in metal ion transporters. In fact, in this sense, a protein can be seen as a big ***complicated*** ligand chelating the metal and subtle conformational changes can affect the metal ion specificity. The protein matrix exerts more control than a small synthetic metal complex because there are also tertiary interactions, electrostatic and conformational features, which modulate metal uptake and release.

During the PhD course my attention was focused on to understand protein stability in the axial methionine loop of a minimal cytochrome c of ***Bacillus pasteurii*** and structural characterization metal binding soluble domains of P1-type of ATPAses PacS and ZiaA. In addition, I have also attempted to characterize two human zinc binding proteins ZNF9 (Zinc Finger 9) and ADAM10 (A Disintegrin and Metalloprotease).

In the first part of my PhD research work we have investigated the loop containing the iron axial ligand Met71 in oxidized Bpcytc using site directed mutagenesis at different sequence positions, in order to obtain better insight into the determinants of the high stability of the axial coordination towards GdmCl and alkaline pH. It appears that one of the main determinants of this stability is the high energy barrier for the rearrangement of the Met loop conformation. The Met loop rearrangement is a necessary

step for replacement of the axial ligand [1]. Met71 is in close contact with Ile75, which is important to tune the conformation of the side chain of the former residue. Mutation of Ile75 with less bulky amino acids results in a subtle rearrangement of the Met71 side chain and a small but significant destabilization of the protein fold. A key role is also played by Pro72, which is conserved also in Bpcytc homologs. Its mutation in fact dramatically affects the stability of Bpcytc towards denaturants and makes largely unfolded conformations much more readily accessible even under native conditions.

This work demonstrates that the amino acid composition of the axial Met loop and the corresponding network of inter-residue interactions are crucial for the stability of c-type cytochromes through a complex interplay of conformational propensities, hydrogen bonds (also with the solvent) and hydrophobic contacts.

In the second part of PhD research work I have studies copper and zinc proteins involved in metal trafficking within cyanobacteria *Synechocystis* PCC 6803.This work addressed the mechanism of copper transfer from the metallochaperone ScAtx1 to the $PacS_N$ domain, particularly investigating the role of His-61.Furthermore structural characterization of $ZiaA_N$ is the starting point for investigating the metal binding properties and the metal selectivity with respect to $PacS_N$ and ScAtx1.

The interaction studies of Cu(I)ScAtx1 with apo$PacS_N$ show indeed that a heterodimeric complex is formed and, if the spectrum of the complex is compared with those of the apo and copper bound forms of ScAtx1 protein, it may be concluded that ScAtx1 in the complex is largely present in the apo form. In addition, comparing the

mobility properties in the absence and presence of BCS, it can be concluded that the copper transport between Cu(I)ScAtx1 and $PacS_N$ is mainly copper affinity dependent. His 61 of ScAtx1 (that is involved in copper binding together with Cys12 and Cys15) is likely no more bound to copper(I) in the protein complex. These data on therefore in agreement with the following mechanism, when Cu(I)ScAtx1 approaches apo$PacS_N$, His 61 is somehow displaced from the coordination sphere allowing one of the two cysteine residues of $PacS_N$ to encroach into a shared copper(I) site as proposed for Atx1/Ccc2a (2). Thus the favored order of copper(I) dissociation of ScAtx1 ligands during the interaction with $PacS_N$ is (I) His 61 (II) Cys 15 finally (III) Cys 12.

Interaction of *Synechocystis* PCC 6803 Atx1 with $ZiaA_N$ domain will indeed exemplifies how the specificity of the contact surfaces between the metallochaperone and soluble metal binding domain, as opposed to the inherent metal-binding preferences of the latter, could dictate which metals are removed from the cytosol. This may be as important for preventing the 'wrong' metals from binding to a protein as facilitating delivery of the correct metals. For example, in the absence of copper metallochaperones (at least within the cytosol of copper-requiring organisms) copper might displace zinc from a sub-set of zinc proteins.

References

1. Bartalesi I, Bertini I, Ghosh K, Rosato A, Turano P (2002) *J Mol Biol.* 321(4): 693- 701.

2. Arnesano, F., Banci, L., Bertini, I., Cantini, F., Ciofi-Baffoni, S., Huffman, D. L., and O'Halloran, T. V. (2001) *J. Biol. Chem.* **276,** 41365–41376.