

THE UNIVERSITY OF CALGARY

Crystallographic Studies of Calcium-Binding Proteins: *Aeromonas salmonicida* Surface
Array Protein and Calmodulin

by

Robert J. Skene

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF BIOLOGICAL SCIENCES

CALGARY, ALBERTA

NOVEMBER, 2000

© Robert J. Skene 2000



National Library
of Canada

Bibliothèque nationale
du Canada

Acquisitions and
Bibliographic Services

Acquisitions et
services bibliographiques

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file *Votre référence*

Our file *Notre référence*

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-64978-4

Canada

ABSTRACT

Aeromonas salmonicida is a gram negative bacterial fish pathogen whose surface layer (S-layer) is a crucial virulence factor in the lethal disease furunculosis. A-protein monomers form a tetragonal lattice on the cell surface, and contain distinct core and linker domains. A high resolution structure of A-protein was attempted using X-ray crystallography. Modifications made to the purification protocol and sample conditions reduced isoelectric form and oligomeric heterogeneity. The resulting A-protein crystals displayed X-ray diffraction to a resolution of 3.5 Å. Unfortunately, these crystals were unstable at room temperature and a complete data set could not be collected.

Calmodulin (CaM) is a ubiquitous Ca^{2+} regulatory protein that binds and activates over 80 different target enzymes. CaM is known to bind the majority of its target proteins via two hydrophobic surfaces that are rich in methionine (Met) residues. X-ray crystal structures were determined for two calcium-saturated CaM variants (CT-CaM to 2.5 Å and Eth-CaM to 2.35 Å). The results indicate that the length and flexibility of the Met side chains are important in providing key hydrophobic contacts in the binding and activation of target proteins.

ACKNOWLEDGMENTS

I would like to express my sincerest gratitude to my supervisor Dr. Barry Phipps, for his insight, guidance, support and friendship during my time in his laboratory. I would also like to express my thanks to Dr. Hans Vogel for his support and friendship over the past years. Thanks to Dr. Tao Yuan who expressed and purified the calmodulin proteins used in this study. In addition, I will be forever indebted to the entire 'Vogel crew' for all the valuable assistance, support and friendship they have provided during my time here. I would like to thank Dr. Steve Litster, Robert Nicholson M.D., Dr. Les Tari and Doug Dougan (cab driver) who not only provided valuable assistance but also made coming to the lab every day a pleasure.

Finally, I would like to express my deepest gratitude to my partner Hilary Devrome, and to my parents and family for the all love and support they have given me over the years.

To Beans

TABLE OF CONTENTS

	Page
Approval Page -----	ii
Abstract -----	iii
Acknowledgements -----	iv
Dedication -----	v
Table of Contents -----	vi
List of Tables -----	viii
List of Figures -----	ix
List of Abbreviations -----	xii
1. INTRODUCTION: S-LAYERS -----	1
1.1 S-layer containing Cell Envelopes	1
1.2 S-layer Proteins	4
1.3 S-layer Ultrastructure	5
1.4 S-layer Biogenesis and Interactions	9
1.5 S-layer Functions	11
1.6 Biotechnological Applications of S-layers	14
1.7 <i>Aeromonas salmonicida</i>	15
1.8 Objective of thesis	22
2. INTRODUCTION: CALMODULIN -----	23
2.1 Biologically important calcium ions	23
2.2 Calmodulin in signal transduction	26
2.3 Calmodulin structure	28
2.4 Calcium binding and conformational change	32
2.5 Peptide binding	33
2.6 Target recognition	35
2.7 Objective of thesis	36
3. MATERIALS AND METHODS -----	38
3.1 Chemicals	38
3.2 Bacterial strains	39
3.3 Culture growth conditions	39
3.4 Cell fractionation	40
3.5 A-protein purification	41
3.6 Calmodulin source and purification	42
3.7 Protein concentration	42
3.8 Protein and LPS electrophoresis	43
3.9 Isoelectric focusing and 2D gels	44
3.10 Dynamic light scattering	45
3.11 Protein crystallization	45
3.12 Crystal mounting	46
3.13 Data collection	46
3.14 Data processing and refinement	47

4. RESULTS: <i>AEROMONAS SALMONICIDA</i> SURFACE	
ARRAY PROTEIN -----	49
4.1 Crystallization of A-protein	49
4.1.1 Purification	49
4.1.2 Crystallization	50
4.1.3 Diffraction studies	54
4.2 Crystallization with reduced isoelectric form heterogeneity	56
4.2.1 Purification	56
4.2.2 Crystallization	59
4.2.3 Diffraction studies	59
4.3 Assessment of oligomeric state	61
4.3.1 Dynamic light scattering	61
4.4 Crystallization with reduced oligomeric heterogeneity	67
4.4.1 Crystallization	67
4.4.2 Diffraction studies	69
5. RESULTS: CALMODULIN -----	73
5.1 Calmodulin sample	73
5.2 Crystallization	73
5.2.1 CT-CaM	74
5.2.2 Eth-CaM	74
5.2.3 Dfm-CaM	77
5.2.4 Nle-CaM	78
5.3 Detwinning crystals	79
5.4 Diffraction data	80
5.4.1 CT-CaM	81
5.4.2 Eth-CaM	81
5.4.3 Dfm-CaM	85
5.5 Data processing	85
5.6 Structure determination and refinement	88
5.6.1 CT-CaM	89
5.6.2 Eth-CaM	94
5.6.3 Dfm-CaM	97
5.7 Structural analysis	97
5.7.1 General structural comparison	97
5.7.2 Examination of Leu mutations in CT-CaM	103
5.7.3 Examination of Eth mutations in Eth-CaM	107
6. DISCUSSION: <i>AEROMONAS SALMONICIDA</i> SURFACE	
ARRAY PROTEIN -----	115
7. DISCUSSION: CALMODULIN -----	124
REFERENCES -----	139
APPENDIX A: Copyright permission -----	152

LIST OF TABLES

Table		Page
1	Some CaM-regulated enzymes and proteins grouped according to their function.	29
2	Crystallization trial using 5 mg/ml A-protein at room temperature.	53
3	Optimized crystallization condition used for the CaM proteins.	75
4	Data processing statistics for CT-CaM, Eth-CaM and Dfm-CaM.	87
5	Progress of CT-CaM and Eth-CaM structure determination as measured by the R-value and free R-value obtained at various stages of refinement.	91

LIST OF FIGURES

Figure		Page
1	The major classes of prokaryotic cell envelopes containing S-layers.	3
2	Illustration of lattice types found for S-layers.	6
3	Three-dimensional models of the protein mass distribution of the S-layer of <i>Bacillus coagulans</i> and <i>Bacillus stearothermophilus</i> .	8
4	A computer generated model from the 3-D reconstruction of a single layer of wild type <i>A. salmonicida</i> .	18
5	Schematic diagram depicting most of the major proteins involved in regulation of cytoplasmic Ca ²⁺ levels.	25
6	Schematic of two methods through which Ca ²⁺ -CaM is proposed to activate target proteins.	27
7	Ribbon diagrams of the three dimensional structures of CaM, apo-CaM and skMLCK-CaM.	31
8	Purification of A-protein from <i>A. salmonicida</i> A450 using DOC/6 M guanidine-HCl and 2 M guanidine-HCl.	51
9	Crystal of A-protein grown in 15% PEG 8000, 0.1 M HEPES pH 6.9 at 21°C utilizing the hanging drop technique.	55
10	Two-dimensional PAGE of A-protein samples.	57
11	Crystal of A-protein with reduced IEF microheterogeneity.	58
12	X-ray diffraction pattern of A-protein with reflections recorded to a resolution limit of 5 Å.	60
13	Particle size distribution analysis of the A-protein sample conditions.	63
14	Particle size distribution analysis of A-protein samples of increasing concentration.	64

15	Crystal of A-protein grown in the presence of 4 mM EGTA and 1% Tween-20.	68
16	Diffraction pattern of an A-protein crystal grown in the presence of 1% Tween with reflections recorded to a resolution limit of 3.5 Å.	71
17	Crystals of CaM displaying rod-like (needle) morphology.	76
18	Diffraction pattern of CT-CaM crystal with reflections recorded to a resolution limit of 2.5 Å.	82
19	Diffraction pattern of Eth-CaM crystal with reflections recorded to a resolution limit of 2.35 Å.	83
20	Diffraction pattern of Dfm-CaM crystal with reflections recorded to a resolution limit of 1.8 Å.	84
21	$2 F_o - F_c $ electron density map of the 2.5 Å structure of CT-CaM contoured around the mutated Leu residues of the C-terminal hydrophobic cleft.	93
22	$2 F_o - F_c $ electron density map of the 2.35 Å structure of Eth-CaM contoured around the biosynthetically incorporated Eth residues of the N-terminal hydrophobic cleft.	95
23	$2 F_o - F_c $ electron density map of the 2.35 Å structure of Eth-CaM contoured around the biosynthetically incorporated Eth residues of the C-terminal hydrophobic cleft.	96
24	Ramachandran plots of CT-CaM and Eth-CaM.	98
25	Overlay of the alpha-carbon backbones of CT-CaM and wild type mammalian CaM.	100
26	Overlay of the alpha-carbon backbones of Eth-CaM and wild type mammalian CaM.	101
27	Ribbon diagrams of CT-CaM and Eth-CaM.	102
28	Overlay of the alpha-carbon backbones from the C-terminal hydrophobic cleft of CT-CaM and wild type mammalian CaM.	104

29	Overlay of the mutated Leu residues in the C-terminal hydrophobic cleft of CT-CaM with the Met residues of wild type mammalian CaM.	106
30	Overlay of the alpha-carbon backbones from the N-terminal hydrophobic cleft of Eth-CaM and wild type mammalian CaM.	108
31	Overlay of the biosynthetically incorporated Eth residues in the N-terminal hydrophobic cleft of Eth-CaM with the Met residues of wild type mammalian CaM.	109
32	Overlay of the alpha-carbon backbones from the C-terminal hydrophobic cleft of Eth-CaM and wild type mammalian CaM.	111
33	Overlay of the biosynthetically incorporated Eth residues in the C-terminal hydrophobic cleft of Eth-CaM with the Met residues of wild type mammalian CaM.	113
34	Overlay of the relevant residues in the C-terminal hydrophobic clefts of CT-CaM, wild type-CaM and smMLCK-CaM.	131
35	Overlay of the relevant residues in the N-terminal hydrophobic cleft of Eth-CaM, wild type CaM and smMLCK-CaM.	134
36	Overlay of the relevant residues in the C-terminal hydrophobic cleft of Eth-CaM, wild type CaM and smMLCK-CaM.	136

LIST OF ABBREVIATIONS

Buffer A	20 mM HEPES pH 7.7, 1 mM EDTA
Buffer B	20 mM HEPES pH 7.7, 4 mM CaCl ₂ , 1 mM EDTA, 0.1 mM PMSF
Buffer C	20 mM Tris-Cl pH 7.7, 4 mM CaCl ₂ , 1 mM EDTA, 0.1 mM PMSF
Buffer D	20 mM Tris-Cl pH 7.7, 1 mM EDTA, 0.1 mM PMSF
Buffer E	20 mM Tris-Cl pH 7.7, 4 mM CaCl ₂ , 1 mM EDTA, and 50 mM NaCl
CaM	Calmodulin
CT-CaM	Calmodulin with four Met to Leu mutations in the C-terminal domain
CNS	Crystallography and NMR system
Dfm	Difluoromethionine
Dfm-CaM	Calmodulin with biosynthetically incorporated Dfm in place of Met
DOC	Deoxycholate
EDTA	Ethylene diamine tetraacetic acid
EGTA	Ethylene glycol bis-2-aminoethylether-N,N,N',N'- tetraacetic acid
Eth	Ethionine
Eth-CaM	Calmodulin with biosynthetically incorporated Eth in place of Met
HEPES	N-2-Hydroxyethyl piperazine-N'-2-ethanesulfonic acid
Leu	Leucine
LPS	Lipopolysaccharide
Met	Methionine
MLCK	Myosin light chain kinase
MPD	2-methyl-2,4-pentanediol

Nle	Norleucine
Nle-CaM	Calmodulin with biosynthetically incorporated Nle in place of Met
NSLS	National Synchrotron Light Source
PBS	Phosphate buffered saline pH 7.2 (0.075 M HPO₄²⁻/H₂PO₄⁻)
PEG	Polyethylene glycol
PMSF	Phenylmethylsulfonyl fluoride
R-value	$R = \sum F_{obs} - F_{calc} / \sum F_{obs}$
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SLS	Sodium lauryl sarcosine
TCA	Trichloroacetic acid

1. INTRODUCTION: S-LAYERS

1.1 S-layer Containing Cell Envelopes

Throughout the course of evolution, prokaryotic organisms have developed a number of unique cell envelope structures. As unicellular life forms, these envelope structures function as an important interface between the cell and its environment. In general, the cell envelope must maintain the chemical composition inside the cell while selectively allowing nutrients and waste products to pass in and out. In addition to this, the cell envelope must be strong enough to withstand the turgor pressure of the protoplast and is often directly involved in determining and maintaining cell shape (1). Although great diversity exists in the structure and sophistication of these bacterial cell envelopes, they can be classified into three main types, the gram-negative archaeal cell envelope, the gram-positive cell envelope in archaea and eubacteria and the more complex gram-negative eubacterial cell envelopes.

In addition to the basic cell envelope, prokaryotes often exhibit two-dimensional arrays of proteinaceous subunits on the cell surface (2). These crystalline protein arrays cover the entire cell surface and have been termed the surface layer or S-layer. The first “periodic macromolecular monolayer” as the outer component of the cell envelope was reported in 1953 by Houwink (3). Electron micrographs indicated a hexagonally patterned layer covering the cell surface of a *Spirillum* species. Today, S-layers have been observed in hundreds of prokaryotic species representing many different taxonomical groups and have been detected on the surface of some eukaryotic algae (4).

In archaea, S-layers are an almost universal feature of the cell envelope (5).

Gram-negative archaea cell envelopes generally have an S-layer as the only component external to the cytoplasmic membrane where gram-positive archaea have an additional layer of pseudomurein adjacent to the cytoplasmic membrane and an S-layer as the outermost element. The membrane lipids found in archaea are comprised of isoprenyl glycerol ether derivatives which are distinct from the ester-linked fatty acyl glycerol derivatives of eubacterial and eukaryotic membranes (6). In gram-positive eubacteria, the S-layer, if present, is located adjacent to the rigid sacculus covering the cytoplasmic membrane which is composed of peptidoglycan of variable thickness. The peptidoglycan layer is chemically and structurally related to the pseudomurein found in archaea. Gram-negative eubacteria have the most complex prokaryotic cell envelope. The cytoplasmic membrane is surrounded by a thin layer of peptidoglycan plus an additional membrane external to this called the outer membrane. The space between the two membranes is the periplasm and contains a number of important soluble proteins involved in uptake of nutrients across the cytoplasmic membrane (7). The outer membrane is an asymmetric lipid bilayer in that the inner leaflet contains phospholipid and the outer leaflet is composed of a significant amount of lipopolysaccharide (LPS). LPS is composed of lipid A which is a disaccharide of phosphorylated glucosamine residues substituted by saturated fatty acyl chains through ester and amide linkages (8). The lipid A moiety is generally covalently linked to a core oligosaccharide which is in turn linked to an O antigen (or O polysaccharide). The core oligosaccharide and O antigen extend out from the membrane surface while the lipid A portion forms the external leaflet of the outer membrane (9). S-layers, if present, are found external to the outer membrane and associated with certain outer membrane components. Figure 1 shows a schematic

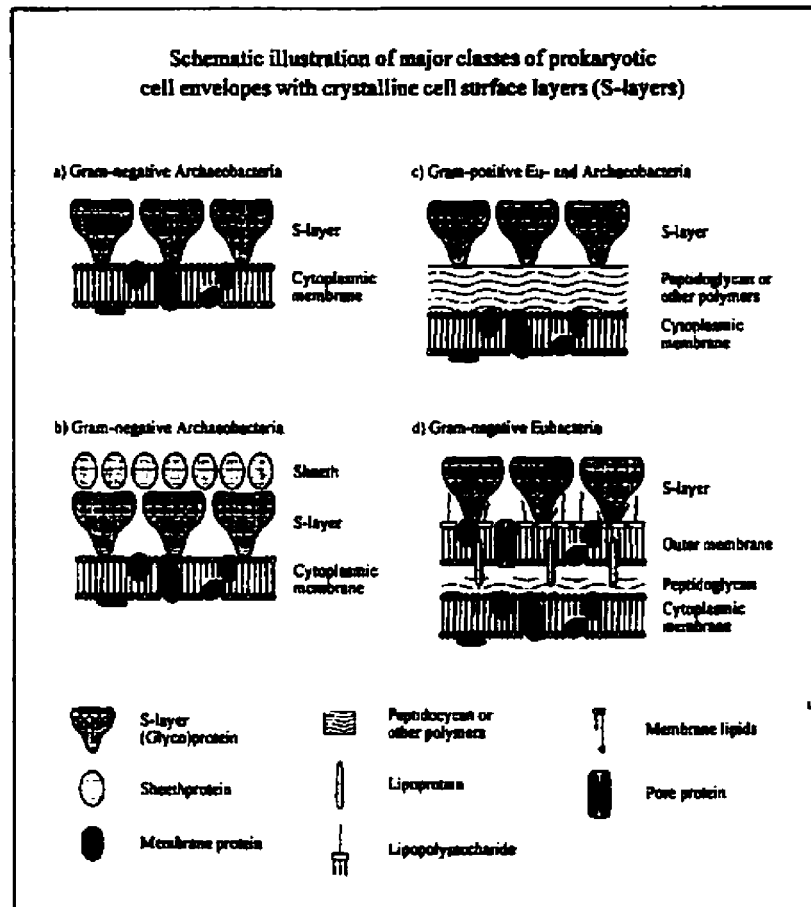


Figure #1 Illustration of the major classes of prokaryotic cell envelopes containing S-layers. (a) Cell envelope structure of Gram-negative archaea with S-layer as the only cell wall component external to the cytoplasmic membrane. (b) Gram-negative archaeal cell envelope with an additional sheath composed of regularly arranged subunits. (c) Gram-positive eubacteria and archaea cell envelope. (d) Gram-negative eubacteria with the S-layer closely associated with the outer membrane. (Reproduced with permission: From Slejtr UB, Messner P, Pum D, Sara M. 1996. Crystalline Bacterial Cell Surface Proteins. pp. 5-23. © R.G. Landes Company)

diagram of the major classes of prokaryotic cell envelopes containing S-layers (1).

1.2 S-layer Proteins

S-layers generally consist of a single protein species with molecular weights ranging from 40 to 220 kDa. In some cases, S-layers are composed of two or more subunit types of differing molecular weights as seen in *Clostridium difficile* (10). The majority of archaeal S-layer proteins undergo post-translational glycosylation; however, a few non-glycosylated S-layers have been identified among the methanogens. The eubacterial S-layer proteins are generally non-glycosylated with some exceptions found among the *Bacillaceae* family (11). In archaea, N-linked short heterosaccharides are the predominate species of glycan whereas in eubacteria, O-linked long-chain glycans are most prevalent (12). Little information is available regarding the biosynthesis of glycoproteins in archaea and eubacteria.

Analysis of the amino acid composition of S-layer proteins presents some common characteristics. S-layer proteins have a disproportionate amount of acidic amino acids. This results in isoelectric points (pI) ranging from 4 to 6 (13). A notable exception is found with some lactobacilli which have pI values around 9 (14). These proteins also possess large quantities of hydrophobic amino acids (40 to 60%) and include few sulfur-containing amino acids. The deficiency in sulfur-containing residues is a characteristic often seen in secreted proteins and proteins exposed at the cell surface such as flagellin (15). Despite the general similarities in amino acid composition, little sequence homology is observed between S-layer proteins (16). Studies indicate the majority of the sequence homology is located at the internal and C-terminal regions with very little present in N-

terminal regions. In addition to this, a multiple sequence alignment of a number of S-layer proteins reveals no conserved amino acid residues (1). This indicates bacterial S-layer proteins are non-conserved and are of limited taxonomical value (17).

Secondary structure predictions of a few S-layer proteins have been done using data from circular dichroism and infrared spectroscopy. This has revealed a general picture of the secondary structure of S-layer proteins. Approximately 30 to 40% of the structure is β -sheet with a variable amount of α -helix ranging up to 20%. The remaining structure is composed of aperiodic folding and β -turns which vary extensively among the different S-layer proteins (18). S-layer proteins when present in intact arrays are highly resistant to proteolytic attack but are susceptible, to varying degrees, in the monomeric form. This implies that the surface exposed regions have evolved to be inert to proteases present in the environment (19, 20).

1.3 S-layer Ultrastructure

Electron microscopy is commonly used in identifying and studying S-layers. One of the most effective techniques used to identify S-layers is freeze-etching of intact cells. These preparations indicate that S-layers form a crystalline array covering the entire surface of the archaeal or eubacterial cell (21). This coverage is complete during all stages of cell growth and division with no apparent gaps (22). The S-layer exhibits a regular two-dimensional crystalline lattice on the cylindrical portion of rod-shaped cells, while at cell poles they consist of randomly oriented patches of arrays or a regular lattice interrupted by lattice faults (21, 22, 23).

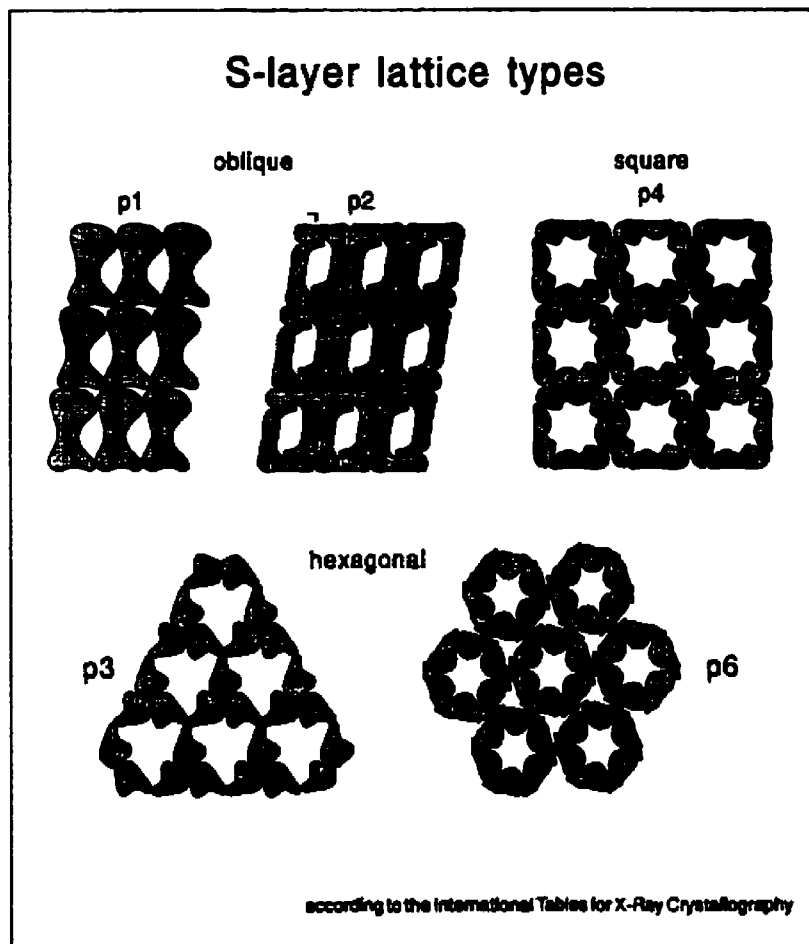


Figure #2 Illustration of lattice types found for S-layers. The lattice unit cells are composed of mono-, di-, tri-, tetra-, or hexamers. (Reproduced with permission: From Sleytr UB, Messner P, Pum D, Sara M. 1996. Crystalline Bacterial Cell Surface Proteins. pp. 5-23. © R.G. Landes Company).

Information about the mass distribution of S-layers has been obtained using high-resolution electron microscopy (electron crystallography) of negatively stained preparations or frozen hydrated specimens. Two-dimensional averaging of individual projection images and three-dimensional reconstruction from tilt series of electron micrographs have given structural information to a resolution of 2.0 to 0.6 nm (24). These indicate S-layers form crystalline lattices exhibiting oblique (p1, p2), square or tetragonal (p4), or hexagonal (p3, p6) symmetry (Figure #2). Archaeal S-layers most commonly exhibit hexagonal symmetry where oblique, square and tetragonal are generally found in eubacteria (25). These morphological units usually have center-to-center spacings ranging from 3 to 30 nm. In thin sections, the S-layer is clearly distinct from cell wall layers and regularly arranged integral membrane proteins (2). A single S-layer protein lattice can range from 5 to 25 nm in thickness; in some cases two distinct S-layers are present, one on top of the other (1).

The three-dimensional models obtained from electron microscopy reveal other common features of S-layers (Figure #3). S-layers have a smooth outer surface with a more corrugated inner surface and have been shown to exhibit a large degree of porosity (26). The presence of two or more distinct pores of differing size and shape has been observed in a number of S-layers. The pore sizes generally range from 2 to 8 nm with the porosity of the protein meshwork between 30 and 70% (27). The pores frequently appear funnel-shaped and occur in the space between protein molecules, not as a channel within the protein itself (28). Although clear boundaries between protein monomers are difficult to determine, these S-layer models suggest probable monomer shapes. S-layer proteins appear to adopt an elongated multiple domain structure with the presence of a

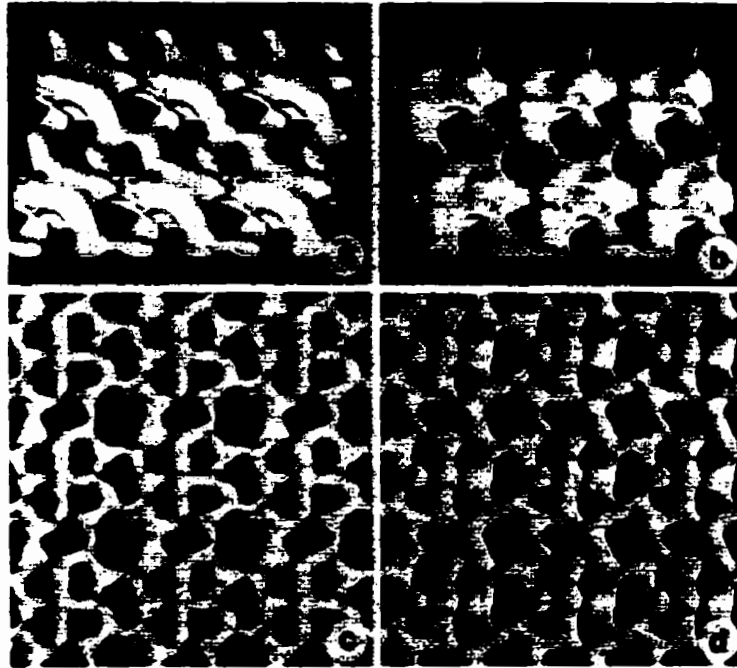


Figure #3 Three-dimensional models of the protein mass distribution of the S-layer of *Bacillus coagulans* E38-66 in (a) and (b), and of *Bacillus stearothermophilus* NRS 2004/3a in (c) and (d). Panels (a) and (c) show the outer face and panels (b) and (d) show the inner face of the S-layer. (Reproduced with permission: From Sleytr UB, Messner P, Pum D, Sara M. 1996. Crystalline Bacterial Cell Surface Proteins. pp. 5-23. © R.G. Landes Company)

heavy domain, a light domain and a small linker domain (29). The heavy domains group together forming large core structures while the light domains extend out and connect with each other in the formation of the S-layer (29, 30, 31, 32).

Studies comparing individual strains of a number of eubacterial species indicate extensive diversity in lattice types, lattice constants and chemistry of S-layers. Despite this, the large core structures exhibit significant similarities. They are located at the inner surface of the layer and enclose a pore frequently about 2.5 nm in size (2, 33). However, the light domains which are exposed on the outer surface display a structure that is highly variable even in closely related species (33, 34). Despite the apparent similarities, S-layers are non-conservative structures and generally represent only a strain-specific characteristic (25, 35). It has been suggested that the common features shared by S-layer proteins and the S-layers as a whole result from convergent evolution in order to satisfy structural and functional requirements (1).

1.4 S-layer Biogenesis and Interactions

The formation of protein subunits into two-dimensional S-layers occurs despite the obvious differences in cell envelope components of gram-positive and negative archaea and eubacteria. The interaction of S-layers with each other and the underlying cell envelope generally involves a combination of ionic bonds, hydrogen bonds and hydrophobic interactions (36, 37). Often the presence of a divalent cation is required to maintain the structural integrity of the S-layer. Chaotropic agents, pH changes, metal chelating agents and detergents have all been used to disrupt and solubilize S-layers. Although little is known about the specific interactions of S-layer proteins with the

underlying cell envelope, they are believed to interact with secondary cell wall polymers of glucose and glucosamine attached to the peptidoglycan layer of gram-positive bacteria and with LPS components in gram-negative bacteria (1).

When disrupting agents used in isolation are removed, the protein subunits demonstrate the ability to self-assemble two-dimensional arrays either in suspension or on a suitable surface (38). These recrystallizations indicate the assembly of the S-layer is an entropy-driven process in which the information necessary to form the arrays is present within the monomer subunits (39, 40). The array formation is thought to be initiated by nucleation into oligomeric precursors which then assemble into the larger lattice (41). The proper insertion and orientation of the S-layer seems to be dependent on the difference in net charge and hydrophobicity of the inner and outer surfaces of the S-layer and is independent of the supporting matrix or cell envelope (42).

The S-layer on an averaged sized prokaryote has been estimated to contain upwards of 5×10^5 monomers, making these proteins one of the most abundant cellular proteins. During cell growth, this would require the production, translocation and incorporation of at least 500 copies per second (38). Since S-layer genes are present on the chromosomes as single copies, they must have very efficient promoters (43) in addition to expedient transport of subunits to the cell surface. In most cases however, the rate of protein synthesis is strictly controlled with very little S-layer protein lost in the growth media. Export of these proteins to the surface of the cell envelope requires the passage through intermediate layers of peptidoglycan and/or outer membrane. In gram-negative eubacteria, transport across the outer membrane occurs either through fusion regions with the cytoplasmic membrane or as an independent step across the outer

membrane requiring an S-layer specific transport mechanism (44, 45). In gram-positive eubacteria, insertion of new subunits is seen to occur in defined growth areas on the cylindrical part of the cell while insertion of new subunits in gram-negative eubacteria appears to occur at random sites (46, 47).

1.5 S-layer Functions

In hundreds of eubacterial and almost all archaeal species, S-layers comprise a major component of the cell envelope structure. S-layer proteins can constitute up to 15% of all cellular protein and as a result, an extensive amount of cellular resources are spent in their synthesis (1). Its location at the cell surface and the large investment of cellular resources in its production, suggests an important biological function. Although much is known about the structure, assembly and chemistry of S-layers, in many cases the specific biological functions are still speculative. It is believed that since S-layers have been detected on a wide variety of bacterial species, they must fulfill a broad range of functions in the hostile and competitive environments of these cells (1). Some general functions attributed to S-layers include determination and maintenance of cell shape, cell adhesion and surface recognition, molecular sieving, trapping of certain molecules and ions, protection against hostile agents and, in specific cases, they may play key roles in pathogenesis (48).

It has been suggested that S-layer like structures could have provided a barrier and supporting function required by self-reproducing systems during the early period of biological evolution (38, 40). This function appears to be preserved in some archaeal cell envelopes which have S-layers as the only cell wall component external to the plasma

membrane. In this case, the rigid nature of the S-layer is thought to determine cell shape and provide protection against osmotic rupture (49). Examples of this are found in *Thermoproteus* and *Pyrobaculum* species in which the hexagonally ordered S-layer lattices are thought to be covalently linked providing a strong rigid structure even as cell free isolated layers (50, 51). *Halobacterium* species are shown to lose their rod shape resulting from disruption of the S-layer (52). However, this function is not thought to be significant in gram-positive and gram-negative eubacteria since strains devoid of S-layers demonstrate no morphological differences (21).

The location of S-layers at the cell surface may enable them to act as a promoter of cell adhesion and surface recognition. S-layers have demonstrated the ability to provide a surface with a variety of characteristics. A strongly hydrophobic surface, as seen with *Aeromonas salmonicida*, is shown to enhance autoagglutination and association with phagocytic monocytes (53). In contrast, the S-layer of *Campylobacter fetus* is hydrophilic and prevents uptake by phagocytes (54). Several *Bacillaceae* have demonstrated net neutral S-layers which act to mask the negative charge of the underlying cell envelope (35). Therefore, the S-layer provides cells with surface charge and/or hydrophobic character enabling interaction with particles and materials of different chemical properties. This enables the cell to resist or adhere to nonbiological substrates, host cells and each other (55).

S-layers provide cells a barrier able to act as molecular sieve and a protective coat. Permeability studies performed on a number of S-layers indicated a molecular weight cut off in the range of 30 000 to 40 000 Da (56). Some S-layers have even demonstrated the ability to modulate pore size suggesting the potential to regulate uptake and release of

macromolecules in response to environmental conditions (57). It is thought that S-layers in archaea and gram-positive eubacteria provide a compartment analogous to the periplasmic space. With a distinct, but much larger molecular weight cut off, the S-layer would be able to retain some important secreted macromolecules (58). A few organisms contain S-layers with the ability to protect the cell through exclusion of harmful lytic enzymes (59). In this case, a pore diameter of 2 to 3 nm corresponding to a molecular weight cut off of 3500 to 11 000 would exclude most proteases, phospholipases and lysozyme. S-layers on a number of gram-negative eubacteria (*Aquaspirillum serpens*, *A. salmonicida*, and *C. fetus*) were found to mask outer membrane attachment sites for the bacterial parasite *Bdellovibrio bacteriovorus* (60).

A variety of specific animal and human bacterial pathogens contain S-layers that have proven to be important virulence factors (21, 25). *C. fetus* causes infertility in sheep and cattle and is responsible for extraintestinal infections in humans. The S-layer resists phagocytosis and repels binding of complement proteins (61). Disruption or removal of S-layer proteins leaves these cells susceptible thereby preventing bacterial infection. The S-layer of *Clostridium difficile*, a pathogen causing human pseudomembrane colitis, was also found to participate in virulence. The layer enables adhesion to human embryonic intestinal cells, and to adult colon cells (62). Other pathogenic bacteria such as *A. salmonicida*, *Rickettsia* species, *Lactobacilli* species and *Bacillus anthracis* all contain S-layers implicated in pathogenesis. Many other human, animal and plant pathogens have been shown to contain S-layers which may be relevant virulence factors (21, 25).

1.6 Biotechnological Applications of S-layers

The ability of S-layer proteins to spontaneously assemble two-dimensional arrays with defined molecular architecture can be exploited for many biotechnological applications. S-layers can be used as ultrafiltration membranes (UFM). Conventional membranes are produced using amorphous polymers such as cellulose acetate and polysulfone. These polymers display pore sizes that vary widely, often by as much as one order of magnitude (63). S-layers provide ideal systems for the production of isoporous UFM's due to their uniform pore size and morphology. S-layer UFM have been produced by depositing the self assembly products on commercial microfiltration membranes resulting in a sharp molecular weight cut off of 30 to 40 000 (64). After cross linking, the surface of the membrane can be chemically modified in order to customize its rejection/adsorption properties.

Since the chemical properties of each unit in the S-layer lattice are identical, they provide a unique structure for defined binding of functional molecules (56). Electron microscopy shows that immobilized molecules are frequently regularly arranged in a monolayer reflecting the periodicity of the supporting lattice (65). The immobilization of functional molecules has applications as bioanalytical mono- and multi-enzyme sensors, enzyme and affinity membranes, immunoassays, dipsticks and affinity microparticles (64, 66). The S-layer matrix provides the additional advantage of extremely low nonspecific adsorption thereby reducing background reactions.

The presence of S-layers on the cell envelopes of a number of pathogenic organisms suggests they may be an important target for vaccine development (67). S-

layers have also been investigated as possible immobilization matrices for weakly immunogenic antigens and haptens. The S-layer conjugate vaccines greatly enhance the medicinal properties of these agents. S-layer vaccines do not cause observable side effects and depending on the S-layer, the antigenic conjugates elicit either cellular or humoral immune responses (64). S-layers have also been used to stabilize liposomes used in drug delivery. Liposomes are readily removed from circulation by the liver and spleen minutes after injection. Therefore, it is advantageous to coat the liposome with a crystalline protein matrix that can be chemically modified to suit a therapeutic use (64, 68).

S-layers also have great potential as supporting structures for functional lipid membranes. Investigation of structure, function and applications of lipid membranes are limited by the short life and poor mechanical stability. S-layers supported Langmuir-Blodgett lipid films demonstrate the ability to maintain structural integrity after heavy use and handling (64). The additional ability to incorporate functional molecules into stabilized lipid layers exposes many potential areas of development. There are potential applications in sensor technology, diagnostics, in electronic or optical devices and in many other areas of biotechnology, biomimetic and nanotechnology research (41, 64, 68).

1.7 *Aeromonas salmonicida*

Aeromonas salmonicida is a gram-negative eubacterial fish pathogen, that is non-motile, facultatively aerobic and rod-shaped in appearance. *A. salmonicida* is an important pathogen, causing both the systemic and ulcerative forms of the disease furunculosis in salmonid fish (69). The presence of a surface layer adjacent to the gram-

negative cell envelope of *A. salmonicida* was first discovered in 1978 by Udey and Fryer. This layer was described as the "additional layer" or A-layer (70). The A-layer (or S-layer) has shown to be the major virulence factor responsible for pathogenicity in fish. Isogenic mutants unable to produce the A-layer display greater than 10^5 fold reduction in virulence (71). Typical strains cause a systemic disease which proceeds rapidly from infection to colonization of body organs, septicemia, tissue necrosis and death. Atypical strains are responsible for the ulcerative form of the disease in both salmonid and non-salmonid fish (72). The persistence of infection even in immunized fish is a major problem in the commercial salmon industry.

Examination of individual cells by thin section electron microscopy shows the A-layer immediately peripheral to the outer membrane and completely surrounding the cell (70, 73). The A-layer is composed of a single non-glycosylated 50 kDa protein species called A-protein. It has been demonstrated that A-protein is exposed at the cell surface conferring hydrophobic character to the A-layer (53). The LPS found in *A. salmonicida* generally contain a nine sugar branched core oligosaccharide linked to lipid A and a complete O-antigen containing a linear tetrasaccharide repeat joined to the core oligosaccharide (74). Some O-antigens project through the surface while most of the O-antigens and the core oligosaccharides are situated beneath the A-layer (74, 75).

The A-layer is associated with the outer membrane of *A. salmonicida* through noncovalent forces. Specific interactions of A-protein monomers with the O-antigens of the LPS act to anchor the A-layer to the cell surface (43, 76). The presence of the divalent cation Ca^{2+} is essential for the proper assembly and attachment of the A-layer. In vitro studies indicate Ca^{2+} acts to enhance attachment of the A-layer to LPS O-antigen and is

required for the correct arrangement of A-protein into the interlocking tetragonal array (77). This indicates Ca^{2+} may form cationic bridges between A-protein monomers and/or induce a conformational change conducive for array formation (77).

Preparations of negatively stained *A. salmonicida* cells reveal large sheets of displaced A-layer. Electron crystallography indicates this layer exists as a tetragonal surface array with lattice spacing of 11 nm. A computer enhanced three-dimensional reconstruction of the A-layer from a tilt series of electron micrographs was reported by Dooley *et al* (79) in 1989 (Figure 4). This model confirms that A-protein monomers form a tetragonal array of alternating heavy and light domains in the A-layer (79). The heavy domains come together forming the common central core architecture seen in other S-layers. This central core structure forms a major fourfold axis of symmetry, containing the majority of the A-layer protein mass, and encloses a funnel shaped pore (79). The light domains act as linkers that extend out from the core structure. These connect to form a secondary fourfold axis of symmetry containing significantly less protein mass. Although the secondary tetragon does not contain a depression as seen in the core structure, oval shaped pores lie between the connecting arms (79). The thickness of the A-layer is estimated to lie between 5 and 8 nm with the linker arms elevated with respect to the plane of the core structures (79).

Biochemical evidence for the domain structure of A-protein has been obtained from protease digestion studies. Treatment of purified A-protein with trypsin produces an N-terminal fragment of 35.5 kDa and a 16.7 kDa C-terminal fragment. The N-terminal fragment is totally refractile to further digestion and contains the majority of the non-surface exposed residues (80). This likely corresponds to the major mass domain that

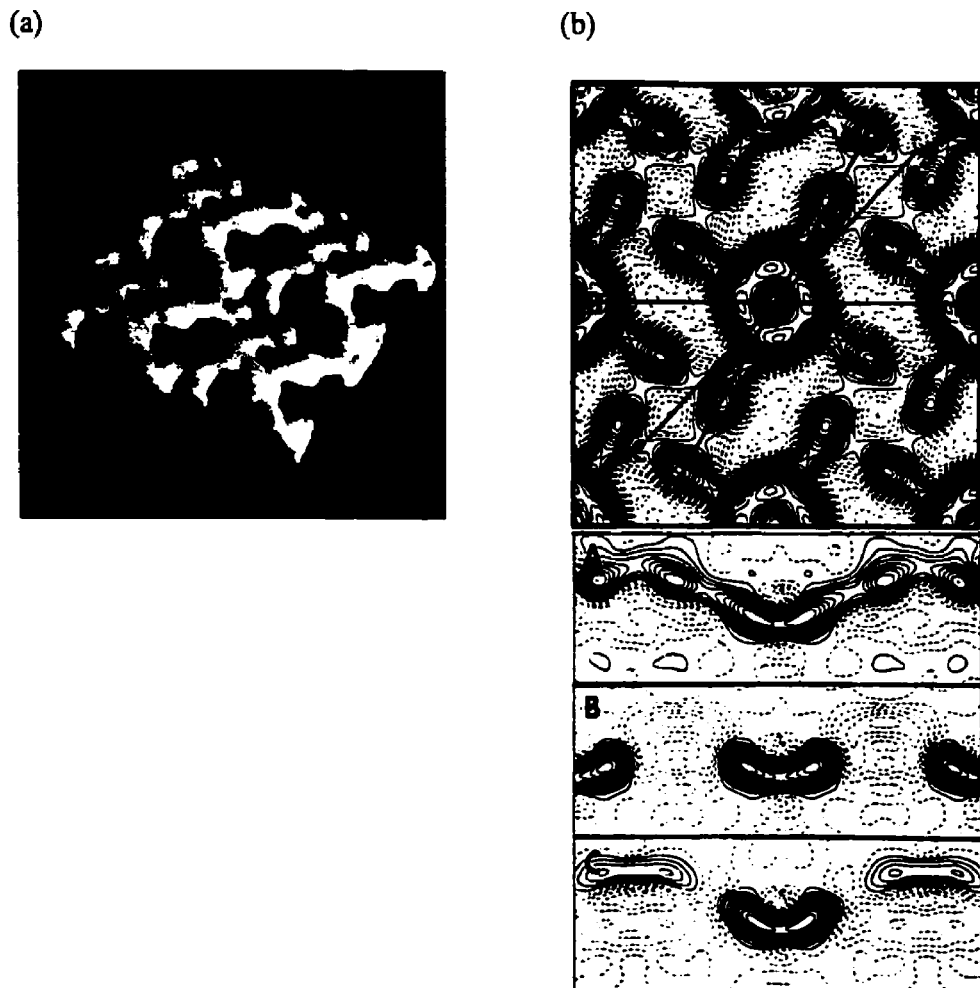


Figure #4 Panel (a) shows a computer generated model from the 3-D reconstruction of a single layer of wild type *A. salmonicida*. Panel (b) shows vertical sections through the 3-D reconstruction . The upper panel indicated the position of the vertical sections. (Reproduced with permission: From Dooley JSH *et al.* 1989. *J. Bacteriol.* 171:190-197. © American Society for Microbiology)

forms the core structure located towards the cell surface. The C-terminal fragment has intermediate resistance to further digestion and contains the majority of the surface exposed residues (80). This is consistent with the proposed minor domain located at the outer surface of the A-layer.

A-protein can be readily extracted from the cell surface using chaotropic agents or acid. These act to disrupt the non-covalent interactions within the A-layer and with the LPS O-antigen. Purified A-protein readily forms multimers and exhibits multiple isoelectric forms corresponding to single charge differences both in solution and on the cell surface (81). The amino acid sequence of A-protein reveals a similar composition to other S-layer proteins. A-protein contains more acidic residues than basic resulting in measured pI values distributed around 5.7 for the different isoforms (81). Like most S-layer proteins, A-protein contains a significant number of hydrophobic residues. However, A-protein has an unusually high hydrophobic amino acid content, one of the highest reported for an S-layer protein (82).

The gene encoding the 481 amino acid A-protein has been cloned and sequenced (83). Expression of the virulence array protein gene (*vapA*) was obtained at low levels in *Escherichia coli*. The gene is present as a single copy and the transcripts were found to be monocistronic. Two promoters (P1, P2) have been identified. The P1 promoter is located 181 nucleotides upstream of the structural gene and the P2 promoter is about 62 nucleotides upstream. It has been determined that *vapA* transcripts directed by the P1 promoter are the major species present in the natural host (83). The P1 promoter is RNA polymerase dependent, with a characteristic -35 consensus sequence, and a related -10 motif (83). Structural predictions of the 181 nucleotide leader segment of the *vapA*

transcript indicate the presence of two stem-loop structures (83). These stem-loop structures are believed to increase the stability of *vapA* transcripts which exhibit a half-life of 22 min compared to an average of 3 min for most prokaryotic transcripts. This stability is believed to contribute to the high levels of A-protein production during cell growth (83). Regulation of *vapA* expression has been observed in the presence of AbcA (84), which is a bifunctional protein containing an N-terminal ATP-binding domain required for synthesis of the O-antigen found in LPS and a C-terminal leucine zipper domain needed for increased *vapA* expression (85). This suggests that synthesis of A-protein and long chain LPS may be co-regulated. Transcription of *vapA* is also regulated by the ambient temperature, as growth of *A. salmonicida* at temperatures exceeding 25°C inhibits the production of A-protein (78). The *vapA* transcript is found to be most stable at 15°C which is expected in that *A. salmonicida* is normally found in fish living at temperatures below 20°C (83).

A-protein containing a 21 amino acid signal sequence is expressed in the cytoplasm and is readily transported across the cytoplasmic membrane. Transposon insertion mutants accumulate A-protein in the periplasm while retaining the ability to export other proteins (43). This indicates that transport of A-protein across the cytoplasmic membrane and outer membrane occurs as distinct steps. The product of the *apsE* gene in *A. salmonicida* exhibits homology with a number of proteins involved in bacterial protein secretion and has a conserved ATP-binding domain (45). Mutation of the *apsE* gene results only in the impaired secretion of A-protein. This indicates the gene product of *apsE* is involved in a specific secretion pathway that transports A-protein across the outer membrane independent of other extracellular proteins (45). Once the

mature A-protein is exported across the outer membrane it is able to spontaneously assemble or insert into the A-layer as seen with other S-layer proteins.

The intact A-layer of *A. salmonicida* is able to fulfill a number of important biological functions for the organism. A key function of the A-layer is as a protective barrier. It is primarily this protective function that makes the A-layer such an important virulence factor. The layer is able to protect the cell from host immune and proteolytic attack. The A-layer shields the cell from complement mediated killing by preventing complement proteins and immunoglobulins from binding to the outer membrane (82). The presence of an intact array also confers resistance to protease degradation thus enabling *A. salmonicida* to survive ingestion by host macrophages (86). The strongly hydrophobic character of the A-layer promotes association and internalization with host macrophages (53). It thus acts as an invasin utilizing macrophages as a means of spreading the infection throughout the host (86).

The A-layer is also able to bind the immunoglobulins IgG and IgM independent of their antigenic specificity (76). Also, the A-layer can bind the extracellular matrix proteins fibronectin, laminin and collagen (87). The binding of IgG and IgM is believed to function as a means of immunoglobulin depletion and/or a masking mechanism to avoid phagocytosis (69). The ability of A-layer to bind extracellular matrix proteins also likely plays a role in pathogenesis. *A. salmonicida* ability to bind these proteins enables the bacterium to persist within the host for long periods until stress precipitates the onset of an acute infection (88). The A-layer and A-protein itself are able to bind porphyrins, hemin and other heme analogues (89). This is proposed to function as a means of heme and iron uptake for *A. salmonicida*, in addition to existing siderophore iron sequestering

mechanisms.

1.8 Objectives of Thesis

As a result of their localization on the cell surface, S-layers are believed to play key roles in bacterial cell growth and survival. An understanding of the important protective and invasion properties, its binding and receptor functions and its ability to spontaneously assemble two-dimensional arrays will greatly enhance our knowledge of the *A. salmonicida* S-layer. To understand the molecular basis of these properties, a high resolution three-dimensional structure would be extremely beneficial. To achieve this, X-ray crystallography must be used to determine the protein structure in atomic detail. This would represent the first X-ray crystal structure determination of an S-layer protein. Therefore, this presents a unique opportunity to provide a significant contribution to our understanding of S-layers, and particularly their role in bacterial pathogenesis. The primary goal of this project is to obtain three-dimensional crystals of the A-layer constituent protein that can be used in structure determination. To achieve this goal, (i) highly purified and homogeneous A-protein must be obtained. (ii) Crystallization conditions must be determined and optimized to yield crystals suitable for X-ray diffraction studies. (iii) Diffraction data can be collected on crystals that diffract to the 3 Å range or better and used for a structure determination by multiple isomorphous replacement or multiwavelength anomalous dispersion.

2. INTRODUCTION: CALMODULIN

2.1 Biologically important calcium ions

In living organisms, certain metal ions are required for the performance of many important biological functions. Many of these metal ions, such as zinc and iron, are required in small quantities where others are needed in relatively high abundance (90). The monovalent cations of sodium (Na^+) and potassium (K^+) typically have an asymmetric distribution in living cells. ATP-driven cation pumps maintain a Na^+ concentration 10 to 20 times higher outside the cell whereas the reverse is true for K^+ . The energy provided by the resulting concentration gradients enable cells to regulate osmotic balance, and drive transport of sugars and amino acids (91). The divalent cation magnesium (Mg^{2+}) also plays an essential biological role. Mg^{2+} ions are present in mM concentrations in cells, are required in the active sites of a number of enzymes and are required to neutralize the negative charge of nucleotide phosphate groups (92).

One of the most important divalent cations necessary in biological systems is calcium (Ca^{2+}). In higher organisms, calcium is commonly found as an essential biomineral in bones and teeth. However, at a cellular level, Ca^{2+} acts as a ubiquitous intracellular messenger able to regulate many biological processes. The first direct evidence of Ca^{2+} ability to act as an intracellular mediator came in 1947 when Heilbrunn and Wiercenski observed that intracellularly injected Ca^{2+} induces the contraction of skeletal muscle (93). Since then it has become clear that Ca^{2+} is an important secondary messenger involved in a wide variety of cellular responses.

The amount of free Ca^{2+} in the cytosol of living cells is kept extremely low (10^{-7}

M) in comparison to the extracellular fluid (10^{-3} M) (94). The cell maintains this large concentration gradient through a number of mechanisms as depicted in Figure 5. Ca^{2+} -ATPases located in the plasma membrane utilize energy from ATP hydrolysis to pump Ca^{2+} out of the cell. In muscle and nerve cells, the removal of Ca^{2+} can also be coupled to the influx of Na^+ through Na^+ - Ca^{2+} exchange proteins (95). In addition, cytosolic Ca^{2+} can be sequestered in specialized intracellular compartments such as the sarcoplasmic and endoplasmic reticulum and even temporarily in mitochondria. These contain Ca^{2+} -ATPase pumps enabling uptake of Ca^{2+} from the cytosol, which is then stored loosely bound to Ca^{2+} -binding proteins such as calsequestrin (96, 97). This sets up a large concentration gradient enabling Ca^{2+} to flow across the plasma membrane and out of the intracellular compartments in response to certain stimuli. The resulting transient increase of cytosolic Ca^{2+} is then used as a secondary messenger to elicit further responses.

There are two pathways by which Ca^{2+} is able to enter the cytosol in response to an extracellular signal. The first method is used primarily in nerve cells. Depolarization of the plasma membrane by an ensuing action potential opens voltage gated Ca^{2+} channels. The resulting influx of Ca^{2+} into the nerve terminal initiates the secretion of neurotransmitter (98). The second method is a ubiquitous pathway used by almost all eukaryotic cells. In this case, extracellular signaling molecules bind to receptors at the cell surface inducing production of another intracellular messenger such as inositol triphosphate. This facilitates release of stored Ca^{2+} from the intracellular compartments resulting in increased cytoplasmic Ca^{2+} (99, 100).

Despite the large concentration gradient driving Ca^{2+} into the cytosol, the level of

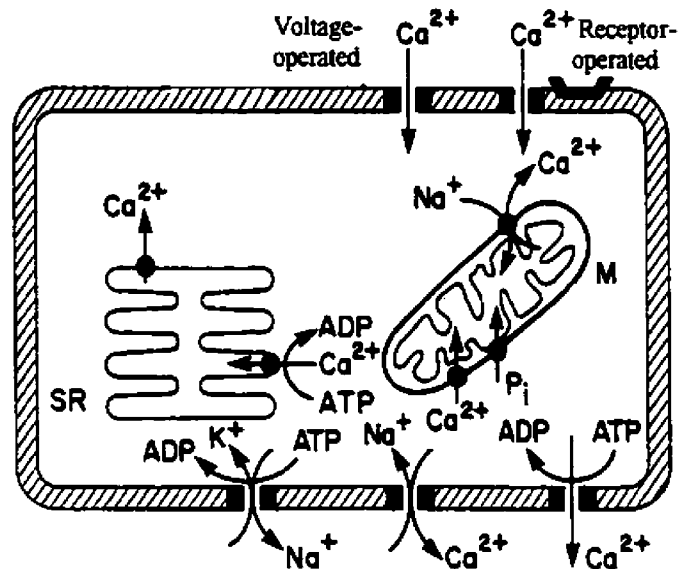


Figure #5 Schematic diagram depicting most of the major proteins involved in regulation of cytoplasmic Ca^{2+} levels. M, mitochondria; SR, sarcoplasmic or endoplasmic reticulum. (Modified from Hiraoki and Vogel. J. 1987. J. Cardiovasc. Pharmacol. 10:14-31.).

free Ca^{2+} in the activated cell is only increased by a factor of 10. Once the intracellular Ca^{2+} concentration reaches 10^{-6} M, feedback inhibition closes the Ca^{2+} channels and activates Ca^{2+} -ATPases thereby maintaining cytosolic Ca^{2+} concentration at that level (101). As a result of this, any molecule serving as a direct target for Ca^{2+} -dependent regulation must have a dissociation constant for Ca^{2+} in the micromolar range, and given the concentration of free Mg^{2+} in the cytosol, must have a selectivity for Ca^{2+} over Mg^{2+} of at least 1000-fold (102).

2.2 Calmodulin in signal transduction

The first protein found to meet the criteria mentioned above was troponin C which plays an important role in skeletal muscle contraction. Later, a closely related protein known as calmodulin (CaM) was discovered as a Ca^{2+} -dependant activator of cyclic nucleotide phosphodiesterase (103). CaM demonstrates a binding affinity in a range suitable to respond to intracellular Ca^{2+} fluctuations and the ability to rapidly react to these changes. As a result, CaM is well suited thermodynamically and kinetically to act in the Ca^{2+} signal transduction pathway (104).

The increase in cytosolic Ca^{2+} results in allosteric activation of the Ca^{2+} -CaM complex. Ca^{2+} -CaM has no enzymatic activity of its own, but elicits a response through binding other proteins. Two methods through which Ca^{2+} -CaM is able to activate target proteins have been proposed. In the first method depicted in Figure 6a, activation of CaM by a Ca^{2+} signal enables the C-terminal domain to bind the target protein. This in turn promotes binding of the N-terminal domain of CaM to the target protein. The

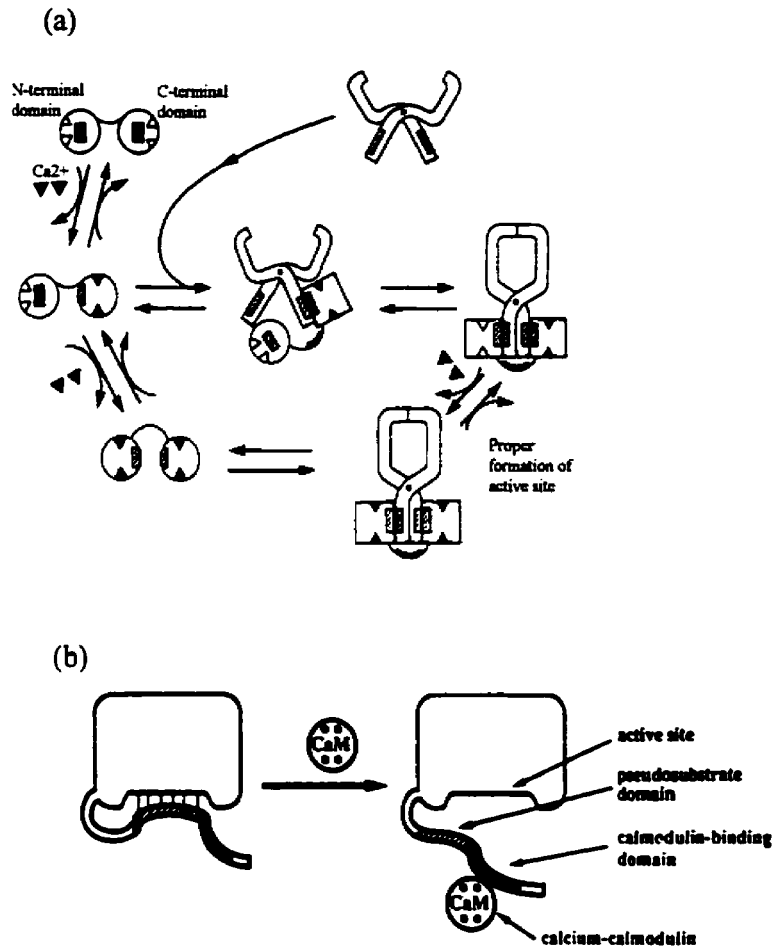


Figure #6 Panel (a) depicts the proposed sequence of events in which calcium ions binding to the C-terminal domain of CaM enables this domain to bind target protein. This is proposed to induce a conformational change facilitating interaction of target with the N-terminal domain of CaM. Panel (b) depicts the role of the proposed CaM-binding and pseudosubstrate domain in inhibition and release of target protein activity. (Panels (a) and (b) are separate figures modified from Vogel H. 1994. *Biochem. Cell Biol.* 72:357-375.).

formation of a complex with CaM then induces a conformational change in the target protein resulting in the formation of a complete active site (92). Support for this method was seen through activation studies conducted using mixtures of CaM fragments (105). In the second method shown in Figure 6b, it is suggested that target proteins contain an inhibitory pseudosubstrate domain in addition to a CaM binding domain. In this case, binding of activated CaM enables the release of the inhibitory domain from a suitably preformed active site (92). In a number of cases a pseudosubstrate domain can be removed through limited proteolysis resulting in a continuously active protein independent of CaM (106, 107). It is now clear that many enzymes activated by CaM use the second mode of action.

CaM is present in all eukaryotic cells and is intimately involved in many important signaling pathways. CaM is able to bind and activate a wide range of target enzymes and proteins in response to a Ca^{2+} signal (108). Some of these include contraction of smooth muscle through interaction with myosin light chain kinase, activation of enzymes in cyclic nucleotide metabolism, and activation of many protein kinases, plasma membrane Ca^{2+} -ATPases and ion channels (92, 109). CaM is also involved in nitric oxide synthesis, in nuclear RNA processing and is believed to be involved in learning processes in the brain (94, 110). In total, CaM is capable of activating upwards of 80 distinct enzymes and proteins some of which are listed in Table 1.

2.3 Calmodulin structure

CaM is an acidic protein whose 148 amino acid sequence is highly conserved in

Table #1 Some CaM-regulated enzymes and proteins grouped according to their function. (Modified from Vogel H. 1994. *Biochem. Cell Biol.* 72:357-375.).

Function	CaM-regulated enzyme or protein
Muscle contraction	MLCK (smooth and skeletal) Caldesmon Calponin
Cell messengers	Plasma membrane Ca ²⁺ -ATPase Cyclic nucleotide phosphodiesterase Adenylate cyclase Nitric oxide synthases Calcium and other ion channels Phospholamban Inositol triphosphate kinase
Protein phosphorylation and dephosphorylation	CaM kinases Phosphorylase kinase Calcineurin Growth factor receptor tyrosine kinases
Gene expression and cell proliferation	CaM kinases and calcineurin Basic helix-loop-helix transcription factors CaM-dependent endonuclease RNA helicase
Others	NAD kinase Phosphofructokinase Heat shock proteins Synapsin Human immunodeficiency virus transmembrane glycoprotein

all eukaryotes. Early NMR and CD data indicated CaM was a two-domain structure composed largely of α -helical regions (111). This was confirmed in 1985 by the X-ray crystal structure of Ca^{2+} -CaM reported by Babu *et al* (112). This revealed CaM to be a dumbbell shaped molecule with N-terminal and C-terminal globular domains linked together by a long central helix (Figure 7b). In this structure, the two globular domains are well defined and structurally homologous. However, residues found in the central helix display high temperature factors. This is indicative of structural heterogeneity or flexibility in this region. The likely structure of the central helix was determined from NMR studies of Ca^{2+} -CaM. The solution structure determined by Ikura *et al* in 1991 (113) and NMR dynamic studies (114) clearly indicate the central helix to be flexible in solution. As a result of this, the two globular domains are able to adjust positions with respect to each other.

Since CaM acts as a secondary messenger in response to a Ca^{2+} signal, the presence of bound Ca^{2+} is expected. The crystal structure of Ca^{2+} -CaM clearly shows four bound Ca^{2+} ions, two in each globular domain (112). These Ca^{2+} ions are bound in helix-loop-helix (EF-hand) binding motifs that are present in the globular domains. This motif consists of two perpendicularly placed α -helices and an interhelical loop that act to form a single Ca^{2+} binding site (115). Each of the highly conserved binding sites provides seven oxygen ligands. This forms a pseudo-bipyramidal coordination sphere with Ca^{2+} (116). The EF-hand helices show extensive contacts with each other producing the compact globular domains (104, 116).

The crystal structure of Ca^{2+} -CaM also reveals an additional feature believed to play a key role in CaM ability to function as an intracellular Ca^{2+} mediator. Specifically,

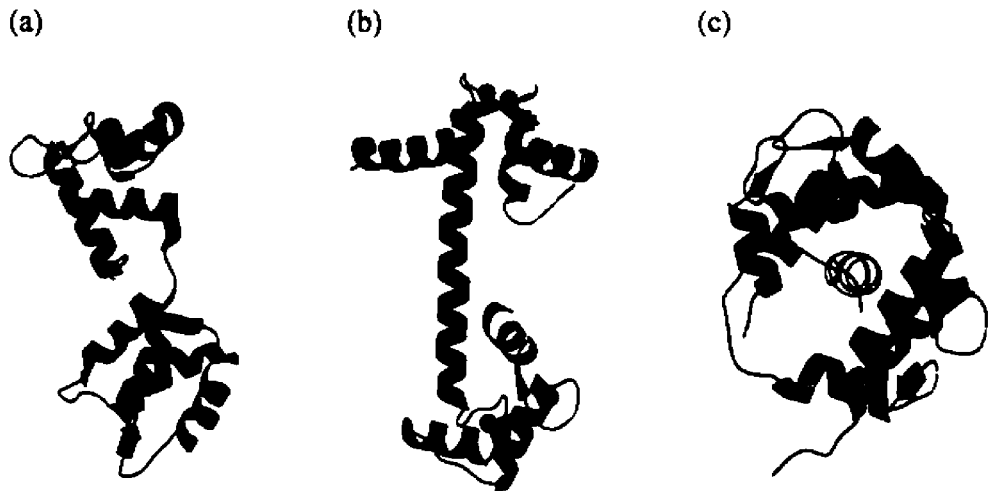


Figure #7 Ribbon diagrams of the three dimensional structures of CaM. (a) Structure of Ca²⁺-free CaM in blue (Zhang *et al.* 1995; 1DMO). (b) Structure of Ca²⁺-bound CaM in red (Chattopadhyaya *et al.* 1992; 1CLL). (c) The skMLCK-CaM complex structure in green (Ikura *et al.* 1992; 2BBM). The figure was prepared using SETOR.

the presence of two hydrophobic surface patches, one in each of CaM's globular domains. In the Ca^{2+} -CaM structure, these hydrophobic patches form a crescent shaped cavity consisting of the hydrophobic residues Met, Leu, Ile, Phe, and Val (117, 118). A striking feature of these hydrophobic surface patches is the large methionine (Met) content. Met residues are located at the surface of the hydrophobic cavity and contribute extensively to the exposed hydrophobic surface area (119). It is through these hydrophobic surface patches that CaM interacts with the majority of its target proteins.

2.4 Calcium binding and conformational change

The Ca^{2+} dependant functions of CaM result from a conformational change in the protein. This conformational change is triggered through binding of Ca^{2+} in the four EF-hand motifs. Ca^{2+} binding is believed to occur in sequential fashion, initially occupying the C-terminal binding sites and then the N-terminal sites (92). The C-terminal domain cooperatively binds two Ca^{2+} with a K_d of 10^{-6} , where the N-terminal domain cooperatively binds Ca^{2+} with 10 fold lower affinity (104, 120, 121). The positive cooperativity of the four binding sites, in the presence of target protein, results in a very steep Ca^{2+} binding curve. This property enables CaM to function effectively as a trigger mechanism over the narrow range that the Ca^{2+} concentration fluctuates in the cell (92).

The structure of CaM in the absence of Ca^{2+} has been solved using NMR by Kuboniwa *et al* and Zhang *et al* in 1995 (Figure 7a) (122). The conformational differences seen in the crystal structure of Ca^{2+} -CaM and the solution structure of apo-CaM indicate how CaM's interaction with target protein is regulated in a Ca^{2+} -dependent manner. Although apo-CaM shows the same secondary structure, there are some

important conformational differences. The position of the two helices constituting each EF-hand binding motif are altered in the presence and absence of Ca^{2+} . In apo-CaM, these helices adopt an antiparallel or closed conformation, which upon Ca^{2+} binding, is changed to a perpendicular or open conformation (104). The concerted movement of these helices in CaM is likely responsible for the high cooperativity of Ca^{2+} binding (115). Another important conformational difference is seen in the hydrophobic surface patches. In the Ca^{2+} -CaM crystal structure, each globular domain contains a large, solvent exposed hydrophobic surface patch. These hydrophobic surfaces have proven to be largely responsible for binding CaM to its targets (123, 124). However, in the absence of Ca^{2+} , the solution structure of CaM shows an important difference in the hydrophobic surface patch. In apo-CaM, the hydrophobic cavity is smaller with less surface exposure (104, 122). Therefore, CaM is only able to interact with its target proteins after Ca^{2+} triggers a conformational change transforming the protein into its active form.

2.5 Peptide binding

Since CaM is able to bind and activate a wide range of target proteins, it is reasonable to expect these target proteins to share a homologous amino acid recognition sequence. However, the CaM binding domains of various targets show almost no amino acid sequence similarity (125). Despite this, they do share some important common features. CaM binding domains are generally a sequence of 20 amino acid residues found at the C-terminal end or in flexible loop regions of the target protein. They are composed of basic and hydrophobic residues and have a tendency to form amphiphilic α -helices (104, 119, 125). In nearly all cases, target proteins form a 1:1 complex with CaM

although examples of multiple CaM attachment are found (109, 126). In order to study the interaction of CaM with its binding domains, synthetic peptides have been used. These peptides contain the same amino acid sequence as the target protein binding domains and are able to bind CaM with high affinity (125).

High-resolution structures of Ca^{2+} -CaM complexed with the binding domain peptide of myosin light chain kinase (MLCK) have been determined by NMR and X-ray crystallography (123, 124). The structure, seen in Figure 7c, has been described as two hands, representing the CaM globular domains, grabbing a rope or peptide (123). These structures indicate that peptide, which is unstructured in solution, adopts an α -helical conformation when bound to CaM. The hydrophobic residues of the peptide interact with the hydrophobic patches in CaM, and the positively charged residues form salt bridges with acidic residues in CaM (104). In the presence of peptide, the N- and C-terminal globular domains of CaM retain a similar conformation as seen in Ca^{2+} -CaM. There is however, a major difference in the central linker region. The central linker unwinds further enabling the two globular domains of Ca^{2+} -CaM to wrap around and bind the helical peptide (92,104). The Ca^{2+} -CaM-MLCK peptide complex adopts a more globular appearance, engulfing up to 80% of the bound peptide (92, 123). The MLCK peptide binds CaM in an antiparallel orientation. The peptide contains two bulky hydrophobic residues separated by 12 amino acids which act to anchor the N-terminal end of the peptide to the C-terminal globular domain of CaM and vice versa (104). CaM binds the peptide exclusively through amino acid side chains in which the majority of contacts are hydrophobic van der Waals interactions (92).

The structures of Ca^{2+} -CaM complexed with the target peptide of CaM Kinase II

(CaMKII), and CaM Kinase Kinase (CaMKK) both indicate a similar overall complex. In the crystal structure of the Ca^{2+} -CaM-CaMKII complex, the peptide is bound in a similar fashion, however the two globular domains of CaM are positioned differently (127). The CaMKII peptide is shorter than the MLCK peptide with the two key hydrophobic groups separated by only 8 residues. Therefore, the central helix of CaM must unwind to a greater extent in order to accommodate this peptide. In the solution structure of the Ca^{2+} -CaM-CaMKK complex some unique features are observed (160). This peptide binds CaM in the opposite orientation with the two key hydrophobic residues separated by 14 residues. In addition, a portion of the α -helical peptide forms a hair pin like loop in the C-terminal domain which is then involved in binding to CaM. Despite this unique target recognition motif, only small differences are observed in the N- and C-terminal domains of CaM as the flexible central linker region is able to properly position these domains. Further NMR and crystallographic data indicate the plasticity of the central linker enables CaM to reposition the globular domains such that hydrophobic surfaces are optimally positioned (125, 128, 129).

2.6 Target recognition

The ability of CaM to interact with over 80 different target proteins that share no binding domain sequence homology is a unique feature of this protein. Interactions between proteins are usually very specific, and the ability to bind multiple targets is generally a result of extensive sequence homology in these binding domains. CaM is seen to interact with its target protein exclusively through amino acid side chains (104, 123, 124). Most protein interactions involve backbone and side chain functional groups which

provides a more rigid and specific binding motif (92). The absence of any side chain - backbone or backbone - backbone contacts is uncommon, and is believed to contribute to the lack of sequence specificity for CaM binding domains (92, 104).

The presence of a large number of Met residues in the hydrophobic surfaces of CaM is also believed to play a key role in the multitarget recognition of CaM. A similar Met rich hydrophobic surface is believed to enable the 54-kDa subunit of the signal recognition particle to recognize multiple targets in a sequence independent manner (104, 130). The average abundance of Met in protein is 1.5%, however, in CaM the Met content is 6.1% (131). CaM contains a total of nine Met residues with the majority found in the two hydrophobic surface patches. Four Met residues are found in each globular domain with one present in the central linker region. The four Met residues in each lobe are located at the entrance to the hydrophobic surface cavity with the base of the pocket composed of aromatic and branched aliphatic residues (104). In total, Met contributes 46% of the total exposed hydrophobic surface area of the protein, which suggests these residues play an important role in binding target protein (112, 119). This is supported by the available structures of CaM-peptide complexes, which indicate Met residues are intimately involved in binding target protein (123, 124).

2.7 Objective of thesis

The ability of CaM to bind and activate a wide range of target proteins in a sequence independent manner is a key functional property of this protein. Met residues present in the hydrophobic surfaces of CaM are believed to play an important role in binding target protein. A quadruple mutant (CT-CaM) with all four C-terminal Met

residues replaced with Leu has been produced and used in binding studies with various target proteins (117, 133). In addition, a number of variant CaM's (Eth-CaM, Nle-CaM, and Dfm-CaM) have been produced with biosynthetically incorporated unnatural amino acids in place of the nine Met residues in CaM. These residues are ethionine (Eth), which has an ethyl group attached to the sulfur rather than the methyl group in Met, norleucine (Nle) which has a methylene group substituted for the sulfur atom in the side chain and difluoromethionine (Dfm) in which the terminal methyl group is difluorinated. These have been subsequently used to study interaction with target protein (131). The objective of this thesis is to obtain the X-ray crystal structures of CT-CaM, Eth-CaM, Nle-CaM, and Dfm-CaM. This will complement previous studies done with these proteins by confirming if no large scale structural changes have occurred and providing insight at the atomic level into the alterations made. From these structures valuable information can be obtained as to the importance of Met and in particular its sulfur group in the binding of target protein with CaM. In order to achieve this, purified mutant proteins will be crystallized. (i) The optimal crystallization conditions for each of these proteins must be determined to obtain crystals suitable for X-ray diffraction studies. (ii) A native data set that diffracts to better than 3 Å must be collected and processed. (iii) With these data sets, electron density maps will be calculated using molecular replacement and the known amino acid sequence fit to give a preliminary structure. (iv) Finally, the models will be refined and used to analyze the hydrophobic surface in atomic detail.

3. MATERIALS AND METHODS

3.1 Chemicals

The following reagents were used for growth of bacterial cultures. Yeast extract, tryptone, peptone and bacto-agar were obtained from DIFCO laboratories. Highly pure sodium chloride and glucose were obtained from BDH. The Na/K minimal media consisted of potassium phosphate (dibasic), sodium phosphate (dibasic) supplied by Sigma, potassium phosphate (monobasic) from Fisher and sodium dihydrogen orthophosphate and magnesium sulfate obtained from BDH.

Reagents and resins used in electrophoresis and chromatography were as follows. Sephadex G-25 Medium and DEAE Sepharose Fast Flow were obtained from Pharmacia Biotech. Electrophoresis grade acrylamide, TEMED, flavin mononucleotide and Bio-Lyte ampholytes (3/10 and 5/7) were purchased from BioRad. N,N'-methylenebisacrylamide, and ammonium persulfate were from Bethesda Research Laboratories. Ultra pure urea and glycine were supplied by ICN, sodium dodecyl sulfate (omni pure) and Coomassie R-250 were obtained from EM Science and Aldrich respectively.

The following chemicals and reagents were used in purification and crystallization trials. Highly pure HEPES, MES, and sodium acetate buffers were purchased from Sigma as were high quality sodium lauryl sarcosine, guanidine-HCl, calcium chloride and PMSF. Ultra pure Tris was obtained from ICN, EDTA was from Boehringer Mannheim, EGTA was supplied by EM science and the detergent Tween-20 was obtained from BioRad. Crystal Screen I, Crystal Screen II, ammonium sulfate Grid Screen, polyethylene glycol 6000 Grid Screen, and Detergent Screen I were all supplied by Hampton Research.

Crystallization grade polyethylene glycol 1000, 3550, 4000, 6000, and 8000 were also purchased from Hampton Research.

3.2 Bacterial strains

Two *A. salmonicida* wild type typical strains, A450 and A451 were used in this study. The strains were supplied by W.W. Kay (University of Victoria) and were originally isolated by C. Michel as strains TG72/78 and TG51/79, respectively. Both strains were isolated from furunculosis infections of trout in France.

3.3 Culture growth conditions

The *A. salmonicida* strains were routinely grown using Luria-Bertani (LB) broth (134) supplemented with 25 mM D-glucose. However, in place of NaCl, a modified Davis minimal medium or “NaK salts” (30 mM K_2HPO_4 , 30 mM Na_2HPO_4 , 16.5 mM KH_2PO_4 , 16.5 mM NaH_2PO_4 , 7.5 mM $(NH_4)_2SO_4$, 0.4 mM $MgSO_4$, pH 7.0) was added to supply buffering capacity and mild hypertonicity to the medium. This medium will be referred to as “LB/NaK salts”. For solid media, LB agar plates were used (15 g/l of agar) and incubated at 20-23°C.

A 100 ml culture was inoculated with a single colony of *A. salmonicida* from LB agar plates and grown in a shaker to an absorbance of 1.0 as measured at 650 nm in “LB/NaK salts” at 20-23°C. From this, a 10 ml inoculum was used to initiate growth in four 2800 ml flasks containing 1000 ml of LB/NaK salts. The broth cultures were grown at 20 to 23°C in a Lab-Line Orbit Environ-Shaker at 275 rpm to an A_{650} between 0.6 and 1.0, which corresponded to mid exponential phase. Liquid cultures were harvested by

centrifugation at 3000 x g for 10 minutes at a temperature of 4°C. The cell pellet from 4 liters of medium was washed and centrifuged 3x with PBS at pH 7.2 at 4°C yielding a wet mass ranging from 22 to 28 grams. If required, the cell pellet was frozen at -20°C for short term storage.

3.4 Cell fractionation

Freshly washed or frozen cell pellets were resuspended in PBS pH 7.2 at 0 - 4°C. Whole cells were ruptured by at least two passes through an Aminco French Pressure Cell. A pressure of at least 16 000 psi was required to achieve adequate cell disruption. Unbroken cells were then removed by centrifugation at 3000 x g for 10 minutes at 4°C. The resulting supernatant was then centrifuged for 45 minutes at 30 000 x g and 4°C to obtain a pellet containing whole membranes. The whole membrane pellet contains the cytoplasmic membrane and the outer membrane with the A-layer still closely associated. The pellet was resuspended in 20 mM HEPES pH 7.7, 1 mM EDTA (buffer A) containing 1% sodium lauryl sarcosine (SLS) and incubated at 22°C for 30 minutes according to the method of Filip *et al* (135). This results in the solubilization of cytoplasmic membrane components. The outer membrane and associated A-layer was isolated by centrifugation at 30 000 x g for 45 minutes at 4°C. The extraction with SLS was repeated a second time and the outer membrane pellet was washed 2 x with buffer A. The serine protease inhibitor PMSF was added at all stages to a final concentration of 0.1 mM.

3.5 A-protein purification

The standard purification protocol for A-protein as reported in Phipps et al (81) and a modified procedure were used. In the modified procedure, solubilization of the outer membrane pellet with deoxycholate (DOC) was omitted and 2 M guanidine-HCl was used in the final extraction instead of 6 M. Cells of *A. salmonicida* (A450 and A451) were grown, harvested and outer membrane pellets prepared as described above. The resulting pellet was suspended in 35 ml of 2 M guanidine-HCl in either buffer A or 20 mM HEPES pH 7.7, 4 mM CaCl₂, 1 mM EDTA, 0.1 mM PMSF (buffer B). This suspension was incubated with stirring for 1 hour at 0 - 4°C and then centrifuged at 135 000 x g for 3 hours at 4°C. The supernatant, which contained the A-protein, was then passed through a 0.2 µm filter.

All subsequent chromatography was done using a GradiFrac low pressure chromatography system with a P50 pump, UV-1 monitor, and XK-16 or XK-26 columns supplied by Pharmacia Biotech. The filtered supernatant was desalted at 4°C using a Sephadex G25 column (bed volume 200 ml) equilibrated with 20 mM Tris-Cl pH 7.7, 4 mM CaCl₂, 1 mM EDTA, 0.1 mM PMSF (buffer C) or, in some cases, with the same buffer but omitting CaCl₂ (buffer D). A-protein was found to elute in the void volume fractions. The pooled fractions were then applied to a DEAE-Sepharose column (bed volume 8 ml) equilibrated in either buffer C or D at 4°C. The column was washed with 3 bed volumes of buffer and eluted with a linear gradient of 0-500 mM NaCl in either buffer D or C. A-protein eluted between 220 and 250 mM NaCl. The sample was then dialyzed into either buffer D or C containing 50 mM NaCl.

3.6 Calmodulin source and purification

Expression and purification of CaM mutants was performed by Tao Yuan and H. Vogel (University of Calgary) using a synthetic bovine CaM gene expressed in *Escherichia coli*. The molecular cloning and site-directed mutagenesis used in the production of the quadruple mutant CT-CaM was reported by Zhang *et al* 1994 (117). CT-CaM was expressed in MM294 *E. coli* cells and purified using Phenyl-Sepharose hydrophobic chromatography as described in Zhang and Vogel 1993 (136). The unnatural amino acids ethionine, norleucine and difluoromethionine were incorporated into CaM and purified following the procedure as described in Zhang and Vogel, 1994 (137). A methionine auxotrophic strain of *E. coli* was used and grown to a desired cell density in a rich medium. Cells were then transferred to a defined medium in which Met was replaced by an amino acid analog. Protein synthesis was induced by addition of isopropylthio- β -galactoside (IPTG) resulting in an incorporation rate of amino acid analog of up to 90% (131). Purified samples of CT-CaM, Eth-CaM, Nle-CaM, and Dfm-CaM were provided in lyophilized form.

3.7 Protein concentration

A number of methods were employed to concentrate A-protein preparations. In some cases, Amicon stirred ultrafiltration cells of 10ml and 250 ml were employed, using Diaflo YM30 membranes. A positive pressure of 55 psi was supplied through a pressurized N₂ gas cylinder. Protein samples were also concentrated using a collodion apparatus fitted with 10 000 M_r cut-off collodion bags supplied by Schleicher and Schuell. Negative pressure was obtained using water aspirator vacuum. Protein was also

concentrated using dialysis against PEG 20 000 in either buffer C or D at 4°C. In some cases, dilute samples were concentrated prior to PAGE by precipitation with acetone. In this case, five volumes of acetone at -20°C were added to the sample and let stand for 20 minutes. The precipitate was centrifuged at 15 000 x g for 10 minutes, and resuspended in buffer. The concentration of purified A-protein was determined directly by reading the absorbance at $A_{280\text{nm}}$ using the established extinction coefficient ($0.606 \text{ mg}^{-1}\text{mlcm}^{-1}$). In the case of the CaM proteins, a known quantity of lyophilized sample was resuspended in buffer and final concentration confirmed by using the extinction coefficient ($\epsilon_{276} = 1.8 \text{ mg}^{-1}\text{mlcm}^{-1}$).

3.8 Protein and LPS electrophoresis

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) of protein and LPS samples were carried out using the Mini-PROTEAN II electrophoresis cell from BioRad. The miniature slab gels were prepared according to the method of Laemmli (138). For A-protein and LPS, the acrylamide concentration in the separating gel was 12% and 5% in the stacking gel. CaM samples were run using a 15% acrylamide separating gel. Samples were incubated in SDS sample buffer (0.0625 M Tris-Cl pH 6.8, 2% SDS, 5% 2-mercaptoethanol, 10% glycerol, 0.05% bromophenol blue) at 95°C for 5 minutes. Prior to this, protein was removed from LPS samples by incubation with proteinase K at 60°C for 1 hour. Electrophoresis was carried out at a constant voltage of 100 V through the stacking gel and 200 V through the separating gel. Protein gels were stained for 1 hour using 0.1% Coomassie blue R-250 in a fixative consisting of 40% methanol and 10% acetic acid. LPS was visualized using a silver stain procedure

according to Tsai and Frasch (139). Low range molecular weight standards were obtained from BioRad consisting of phosphorylase B (M_r 92 500), BSA (M_r 66 200), ovalbulmin (M_r 45 000), carbonic anhydrase (M_r 31 000), soybean trypsin inhibitor (M_r 21 000) and lysozyme (M_r 14 000).

3.9 Isoelectric focusing and two-dimensional gels

Isoelectric focusing was performed using the Model 111 Flatbed Mini IEF Cell from BioRad. Denaturing gels contained 8 M urea, 2% NP-40 in addition to 5% acrylamide, 5% glycerol and 20% carrier ampholyte. The appropriate ampholytes were selected to produce a linear pH gradient within the range of either 3 to 10 or 4 to 6. Sample buffer contained 8 M urea, 3% NP-40, 3% ampholytes, 7.5% 2-mercaptoethanol, and 7.5% glycerol was mixed 2:1 with protein samples. Focusing was carried out at a constant voltage increased in steps from 100 V to 200 V in 15 minute intervals and finally to 450 V for 60 minutes. Gels were fixed in 12.5% TCA, 30% methanol for 30 minutes and stained in 27% isopropanol, 10% acetic acid, 0.04% Coomassie blue R-250 with 0.5% CuSO_4 to remove background staining of ampholytes.

Two-dimensional polyacrylamide gel electrophoresis was performed using the Mini-PROTEAN II 2-D cell and Mini-PROTEAN II electrophoresis cell from BioRad. The first dimension IEF tube gels contain 4% acrylamide, 9.2 M urea, 2.0% Triton X-100, 1.6% 5/7 and 0.4% 3/10 ampholytes. The first dimension sample buffer consists of 9.5 M urea, 2.0% Triton X-100, 5% 2-mercaptoethanol and ampholytes as described above was mixed 1:1 with sample. Tube gels were run using a constant voltage of 500 V for 10 minutes followed by 750 V for 3.5 hours. The second dimension was run and developed

as described for SDS-PAGE, in which the tube gel is incubated in sample buffer and placed length wise along the slab gel.

3.10 Dynamic light scattering

Insight on how protein samples behaved in solution was obtained through dynamic light scattering using a Protein Solutions DynaPro Light Scattering instrument. Samples to be tested were prepared by centrifugation at 15 000 x g followed by filtering through either a 0.1 or 0.02 micron filter. Test solutions were injected into the sample cell and placed in the MicroSampler which was maintained at a temperature of 20°C. The Dynamics version 4.0 software was used to record and interpret the data obtained. Once the count rates were stable to within approximately 15% of the average intensity, measurements were taken. At least 10-20 measurements were required to provide adequate data for each experiment.

3.11 Protein crystallization

The two methods of protein crystallization used in this study were vapor diffusion and microdialysis. Crystallization experiments were primarily conducted using vapor diffusion via the hanging drop technique. In this case, Linbro plates were used with siliconized glass cover slips sealed to the reservoir with vacuum grease. Initial crystallization trials were conducted using Hampton Research sparse matrix and grid screens. For this purpose, reservoirs of 500 μ l were used with hanging drops containing 2 μ l protein and 2 μ l reservoir solution. Once initial conditions were found, reservoir volumes of 1000 μ l and hanging drops of 4 to 12 μ l were routinely used. In some cases,

microdialysis using dialysis buttons sealed with a semi-permeable membrane were used. In this case, dialysis buttons containing 5 to 50 μ l of protein sample were placed in a 1000 μ l reservoir solution and sealed in a Linbro plate with glass cover slips and vacuum grease. Crystallization trials were conducted at three temperatures. A-protein trials were primarily conducted at room temperature, however temperatures of 18°C and 4°C were also used. Crystallization trials of CaM were done primarily at 4°C although some room temperature trials were also carried out. All macroseeding was done using hanging drop setups in which pre-grown crystals are washed and introduced individually into a pre-equilibrated protein solution.

3.12 Crystal mounting

Protein crystals obtained as described above were mounted either for room temperature or cryo data collection. For room temperature data collection, crystals were mounted in capillary tubes supplied by Charles-Supper Co. Glass and quartz capillary tubes with internal diameters ranging from 0.1 to 0.7 mm were used and sealed with either capillary wax or epoxy. Capillaries were secured to a goniometer head using plastic pins or mounting clay. For cryocrystallography, crystals were mounted using the crystal cap system and magnetic cap assembly supplied by Hampton Research. The 20 micron nylon cryoloops ranging from 0.1 to 0.7 mm in loop diameter were used and crystals frozen by either submerging directly in liquid nitrogen or by placing in the cold stream (see below).

3.13 Data collection

The X-ray diffraction data was obtained either in house at the University of

Alberta, the University of Calgary, or at the Brookhaven National Laboratory. National Synchrotron Light Source (NSLS). Data collected in at the University of Alberta utilized a Nonius Dip 2030 image plate system with a Rigaku RU-200 rotating anode generator producing $\text{CuK}\alpha$ X-rays. Data collected in Calgary utilized a Mar 345 image plate detector with a Rigaku RU-H3RHB rotating anode generator producing $\text{CuK}\alpha$ X-rays at 50 kV and 100 mA (wavelength = 1.54 Å). Osmic Max-Flux confocal multilayer optics were used to focus and produce a monochromatic X-ray beam. An Oxford Cryosystems 600 Series Cryostream Cooler was used to maintain the crystal at a temperature of 100°K when cryocrystallography data collection was required. Individual diffraction images were collected by rotating the crystal in the X-ray beam through an oscillation angle (rotation about the phi axis) of either 1 or 1.5° with an exposure time ranging from 5 to 15 minutes. A complete data set consists of images collected over a phi angle range of 90 to 180°. The crystal to detector distance was generally set at 180 mm however a distance of 120 mm was also used. Data obtained at NSLS was collected using a Brandies B4 CCD detector and an Oxford Cryosystems 600 series Cryostream Cooler at the X12C beam line. The crystals were oscillated through 1° over 2 to 5 minute intervals with a crystal to detector distance of 200 to 300 mm.

3.14 Data processing and refinement

The X-ray diffraction data obtained was processed using the HKL software package distributed by HKL Research Inc (140). The HKL suite consists of three programs intended for the analysis of diffraction data. XDISPLAY was used for visualization of diffraction patterns, DENZO was used for data reduction and integration,

and SCALEPACK was used for merging and scaling of the obtained intensities.

Molecular replacement and subsequent crystallographic refinements were conducted using the program CNS (141). Once an initial model was obtained, a series of rigid body refinements, simulated annealing, B-factor refinements, and conjugate gradient minimizations were performed. Electron density maps ($2|F_o|-|F_c|$ and $|F_o|-|F_c|$) were calculated and examined on a Silicon Graphics R10000 Indigo 2 workstation using the molecular graphics program O (142). Progress of the refinement was monitored using both R and R_{free} -values (5% of data in test set).

4. RESULTS: *AEROMONAS SALMONICIDA* SURFACE ARRAY PROTEIN

4.1 Crystallization of A-protein

4.1.1 Purification

The success of crystallization experiments can often be correlated to the purity of the protein sample used. This is based on the fact that contaminants may compete for sites on the growing crystal and in doing so, generate lattice errors resulting in internal disorder and poor diffraction. The surface array protein of *Aeromonas salmonicida*, also known as A-protein, can be obtained in large quantities following the purification protocol published by Phipps *et al*, in 1983 (81). However, samples purified by this procedure have demonstrated a degree of LPS and porin contamination. For the purposes of crystallization, the purity of the A-protein obtained from this procedure had to be improved.

The following alterations were made to the established purification protocol. The first modification was to eliminate the 2% deoxycholate extraction performed on the outer membrane pellet. This extraction was thought to disrupt the integrity of the outer membrane. The resulting destabilized outer membrane releases unwanted porin and LPS during the subsequent guanidine-HCl extraction. The second modification made was the use of 2 M guanidine-HCl at 0-4°C to release the A-protein from the outer membrane pellet. Extraction of A-protein with 6 M guanidine-HCl at 22°C results in the liberation of both porin and LPS into the preparation. These two modifications taken together result in a more stable outer membrane pellet enabling the selective removal of A-protein. The lower concentration of guanidine-HCl was also preferable as it was less likely to

cause denaturation of the A-protein. The improved purity of the preparation can be seen in Figure 8. This indicates the modifications made virtually eliminate the presence of porin contamination during the guanidine-HCl extraction. In addition to this, the presence of LPS was virtually eliminated as judged by SDS-PAGE and silver staining (data not shown).

4.1.2 Crystallization

To date, the production of diffraction-quality crystals has not been reported for any surface layer proteins. Therefore, the crystallization of A-protein required an extensive search to determine possible conditions. The search for a unique condition that facilitates crystallization of a new protein can be a laborious task. Generally, a variety of sparse matrix crystallization screens and grid screens are used to obtain preliminary conditions. The sparse matrix allows an efficient screen of a broad range of precipitants such as salts, polymers, and organic solvents at a variety of pH's. Grid Screens utilize a single precipitant thereby screening a variety of concentrations at unique pH's, and are commonly used in conjunction with the sparse matrix screens. Results from the initial screens yield information on the solubility of the protein, and ideally a starting crystallization condition. If crystals are obtained in the initial screen, the condition is then optimized to produce crystals suitable for X-ray diffraction studies. In the event that no clear starting conditions are found, the solubility information provided by the initial trials is used to direct the strategy of subsequent screens.

For A-protein purified in the presence of excess Ca^{2+} , a sample was concentrated to 10 mg/ml in 20 mM Tris-Cl pH 7.7, 4 mM CaCl_2 , 1 mM EDTA, and 50 mM NaCl

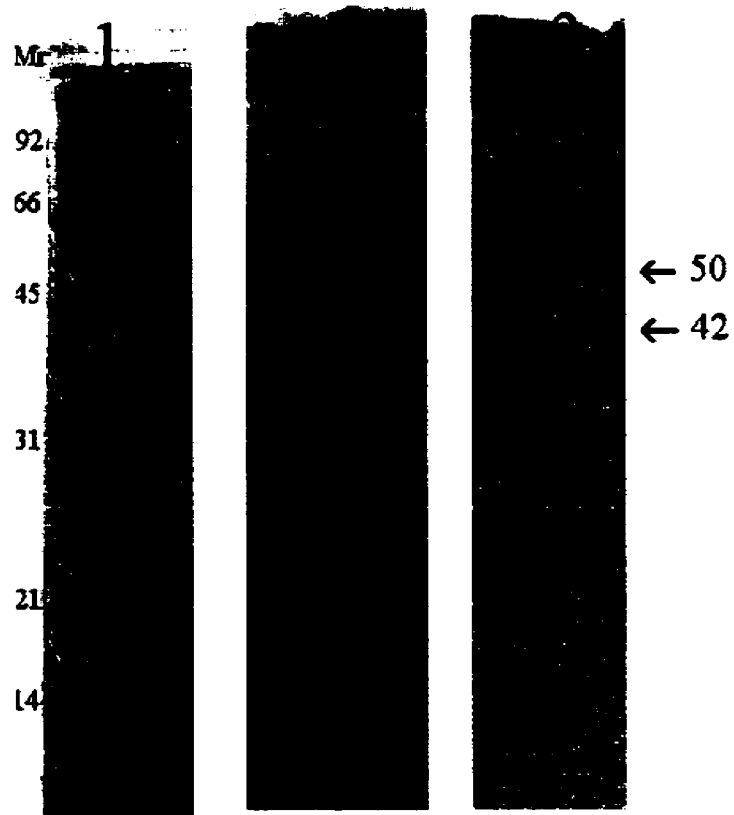


Figure #8 Purification of A-protein from *A. salmonicida* A450 using DOC/6 M guanidine-HCl and 2 M guanidine-HCl. Lane 1, molecular weight standards; lane 2, sample after DOC/6 M guanidine-HCl extraction displays porin contamination; lane 3, sample after 2 M guanidine-HCl extraction is virtually free of porin. Arrows indicate the positions of A-protein (M_r 50 000) and porin (M_r 42 000).

(buffer E). Initial crystallization trials were performed using the Crystal Screen I sparse matrix kit from Hampton Research which contains 50 unique reagents. This initial screen was conducted at both room temperature and 4°C. Observation of the hanging drops over a period of several days revealed the formation of amorphous precipitate in nearly all 50 conditions of Crystal Screen I. A-protein precipitated rapidly upon mixing with reservoir solutions and the appearance of the drops were unaltered beyond the second day.

Finding a condition able to induce a state of supersaturation in the protein sample is the goal of any crystallization experiment. Supersaturation is the driving force in the formation of a crystal nucleus and subsequent crystal growth. The poor solubility demonstrated by A-protein in nearly all ranges of precipitant concentration, pH, and temperature indicates supersaturation is being exceeded. This suggested the concentration of the protein sample needed to be reduced in subsequent trials. Therefore, A-protein at a concentration of 5 mg/ml in buffer E was used in subsequent crystallization trials. The sparse matrix Crystal Screens I and II plus the PEG and ammonium sulfate Grid Screens were set up as before at room temperature. Table 2 indicates the notable results obtained from these initial screens.

Information obtained from Crystal Screen I and PEG Grid Screen suggested PEG was the most successful precipitant to use with A-protein. The crystallization conditions were optimized, starting with initial conditions obtained from these screens. This involved systematically altering a single variable of the crystallization condition and monitoring subsequent crystal formation. Optimization led to the following conditions for growth of A-protein crystals.

Table 2. Crystallization trials using 5 mg/ml A-protein at room temperature

screen used and tube #	crystallization condition	observation of drop	follow up
CS I #9	30% PEG 4000, 0.1 M Na Citrate pH 5.6, 0.2 M Ammonium Acetate	Granular precipitate appears after 2 days	Further optimization gave no improvement
CS I #18	20% PEG 8000, 0.1 M Na Cacodylate pH 6.5, 0.2 M Magnesium Acetate	Microcrystalline material appears after 4 days	Optimization led to microcrystal formation
CS I #22	30 % PEG 4000, 0.1 M Tris-HCl pH 8.5, 0.2 M Na Acetate	Granular precipitate appears after 2 days	No improvement
CS I #40	20% PEG 4000, 20% iso-propanol, 0.1 M Na Citrate pH 5.6	Microcrystalline material appears after 5 days	Optimization led to microcrystal formation
CS I #41	20% PEG 4000, 10% iso-propanol, 0.1 M HEPES pH 7.5	Birefringent precipitate appears after 3 days	Optimization lead to microcrystal formation
CS II #26	30% PEG MME 5000, 0.1 M MES pH 6.5, 0.2 M Ammonium Sulfate	Granular precipitate appears after 1 day	No improvement
CS II #37	10% PEG 8000, 8% Ethylene Glycol, 0.1 M HEPES pH 7.5	Microcrystalline material appears after 4 days	Optimization lead to microcrystal formation
CS II #38	20% PEG 10000, 0.1 M HEPES pH 7.5	Granular precipitate appears after 1 day	No improvement
PGS #B3	10% PEG 6000, 0.1 M MES pH 6.0	Microcrystals grow from light amorphous precipitate in 7 days	Optimization lead to single crystals suitable for X-ray diffraction studies
PGS #B4	10 % PEG 6000 0.1 M HEPES pH 7.0	Microcrystalline material appears amongst light amorphous precipitate in 5 days	Optimized as with PGS #B3
PGS #C3	20% PEG 6000, 0.1 M MES pH 6.0	Microcrystals grow from heavy amorphous precipitate in 7 days	Optimized as with PGS #B3
PGS #C4	20% PEG 6000, 0.1 M HEPES pH 7.0	Microcrystalline material appears amongst heavy amorphous precipitate	Optimized as with PGS #B3

Note: CS = Crystal Screen
PGS = PEG grid screen

- 2 mg/ml A-protein in buffer E
- 15% PEG 8000
- 0.1 M HEPES pH 6.9
- 21°C

Crystallizations were set up using the hanging drop technique with 1000 μ l of reservoir solution and a drop containing 6 μ l protein solution and 3 μ l reservoir solution. The contents of the hanging drop were not actively mixed together during the setup procedure in order to promote a longer equilibration time. In addition, reducing the volume of reservoir solution added to the drop resulted in longer equilibration time. This caused the protein to approach the point of supersaturation more gently and reduced the amount of amorphous precipitate formed. Observation of the hanging drops indicated the formation of small quantities of amorphous precipitate within 2 days, followed by crystal growth in 14 to 28 days. These appeared with two distinct crystal morphologies, having smooth faces and sharp edges, forming either rectangular or diamond shaped crystals. The crystal shown in Figure 9 grew to 0.3 x 0.2 x 0.2 mm and was of suitable size to proceed with X-ray diffraction studies.

4.1.3 Diffraction studies

Crystals of A-protein were mounted at room temperature in a quartz capillary tube and used for X-ray diffraction studies. Crystals were exposed to X-rays for 15 minutes using an oscillation angle of 1°. The resulting diffraction was extremely poor. Very weak reflections were observed around the beam stop that correspond to a resolution limit of approximately 10 Å. This data could not be indexed and therefore no information

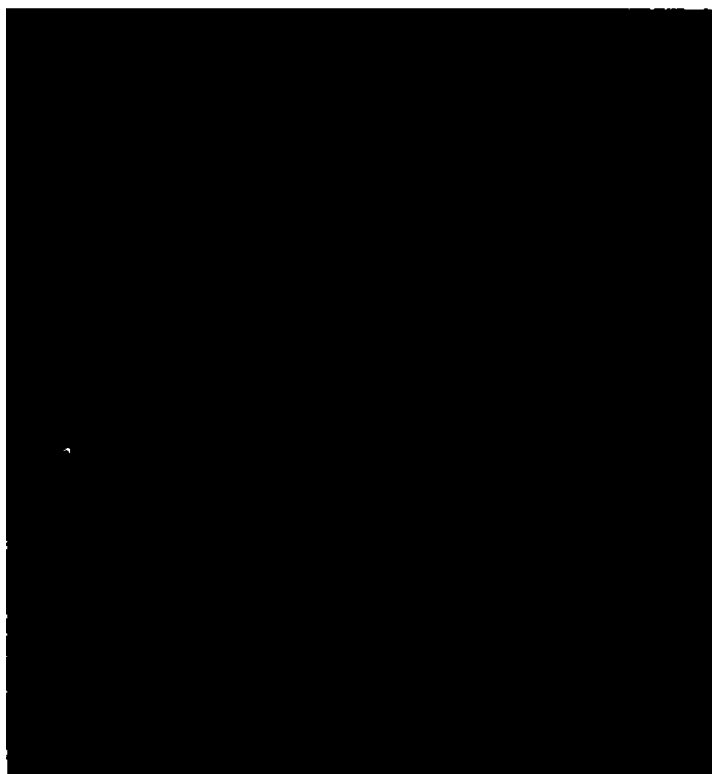


Figure #9 Crystal of A-protein grown in 15% PEG 8000, 0.1 M HEPES pH 6.9 at 21°C utilizing the hanging drop technique. Crystal dimensions are 0.3 x 0.2 x 0.2 mm with rectangular morphology. X-ray diffraction from these crystals was poor, likely resulting from microheterogeneity in the protein sample.

on the unit cell parameters could be obtained.

4.2 Crystallization with reduced isoelectric form heterogeneity

4.2.1 Purification

Failure of the initial A-protein crystals to diffract to a useful resolution range indicated the presence of significant internal disorder in the crystals. The protein sample used was essentially free of contaminants, however, disorder could result from heterogeneity among the A-protein molecules. Heterogeneity in a pure sample of macromolecules is referred to as microheterogeneity. From previous studies, A-protein was known to exhibit isoelectric form microheterogeneity. A-protein was shown to possess a number of isoforms with closely spaced pI's ranging from 4.8 to 5.3. The isoforms were spaced 0.05 to 1.0 pH units apart and probably represented single charge differences (143). In order to improve the diffraction of A-protein crystals, attempts were made to reduce the level of isoform heterogeneity in the sample.

Initially, the purification protocol was modified in an attempt to achieve isoelectric purity. The A450 strain was grown at a lower temperature and harvested earlier in the exponential growth phase. Protein obtained from cells grown at 20°C and harvested at an A_{650} of 0.6 gave an improved isoform distribution. In Figure 10, 2D gels indicate the improvement in sample microheterogeneity. Originally, purified A-protein contained four major isoforms, however, the modifications made reduced this to two major species. Although all isoforms were still represented in the A-protein sample, the overall quality of the protein preparation was improved. Further attempts to isolate a single isoform using chromatofocusing and preparative IEF in solution (BioRad Rotoform

(a)



(b)



Figure #10 Two-dimensional PAGE of A-protein samples. The first dimension tube gels separate protein according to isoelectric forms over a pH range of 4 to 9 (from left to right). The second dimension is a standard SDS-PAGE. Panel (a) shows the original A-protein sample. This sample contained four major and one minor isoelectric species of A-protein. Panel (b) shows A-protein obtained utilizing the modified purification procedure. This sample contained two major isoelectric forms with significant reduction in the other three isoforms.

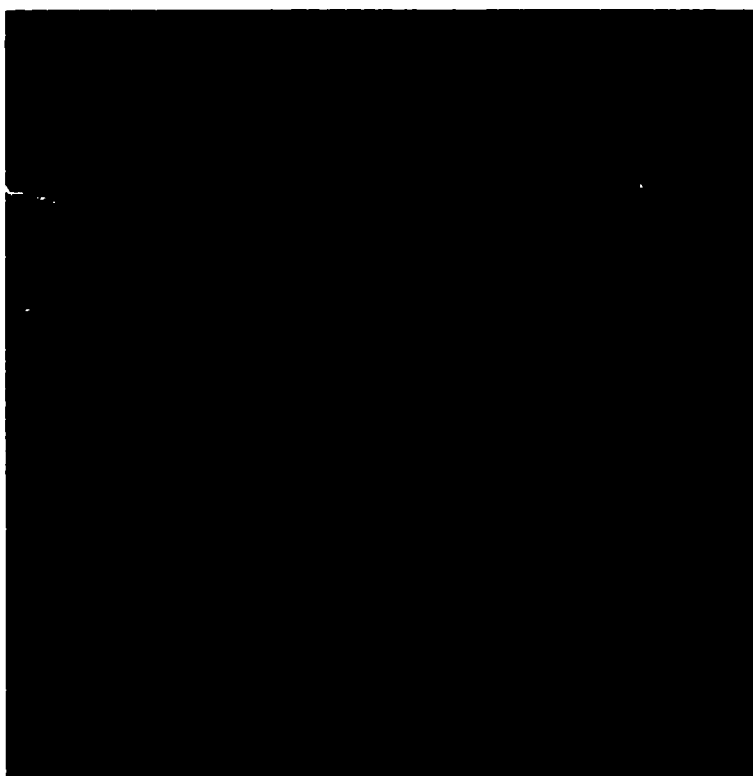


Figure #11 Crystal of A-protein with reduced IEF microheterogeneity grown from the previously reported condition utilizing the hanging drop technique. Appearance is similar to previous A-protein crystals with dimensions of 0.3 x 0.2 x 0.15 mm with diamond like morphology. X-ray diffraction from these crystals were significantly improved.

system) were unsuccessful in separating the closely spaced isoelectric forms of A-protein.

4.2.2 Crystallization

Having obtained A-protein with reduced microheterogeneity, crystallization experiments were initiated. A-protein samples of 2 to 5 mg/ml in buffer E were prepared and screened as before. These results indicated that the previously optimized conditions still produced the best A-protein crystals. Crystals obtained had the same morphology and grew to approximately the same size as those previously obtained. The crystal, seen in Figure 11, grew to 0.3 x 0.2 x 0.15 mm in 14 to 28 days and was used in X-ray diffraction studies.

4.2.3 Diffraction studies

Data was collected at room temperature using the in house MAR 345 image plate detector. The crystals were exposed to the X-ray beam for 15 minutes while oscillating through a phi angle of 1°. An example of the resulting diffraction is seen in Figure 12. These crystals demonstrated significantly improved diffraction. Reflections were observed to a resolution of 5 Å, and the data was indexed. Although X-ray diffraction was limited, information as to the unit cell was obtained. A-protein appeared to crystallize in a primitive monoclinic space group (P2 or P2₁) with the unit cell dimensions $a= 319.7$ $b= 175.3$ $c= 350.0$ Å and $\beta= 93.5^\circ$. This unit cell is very large and likely not accurate. The volume of this unit cell would require approximately 46 to 125 A-protein molecules in the asymmetric unit (assuming 80-30% of cell volume is occupied

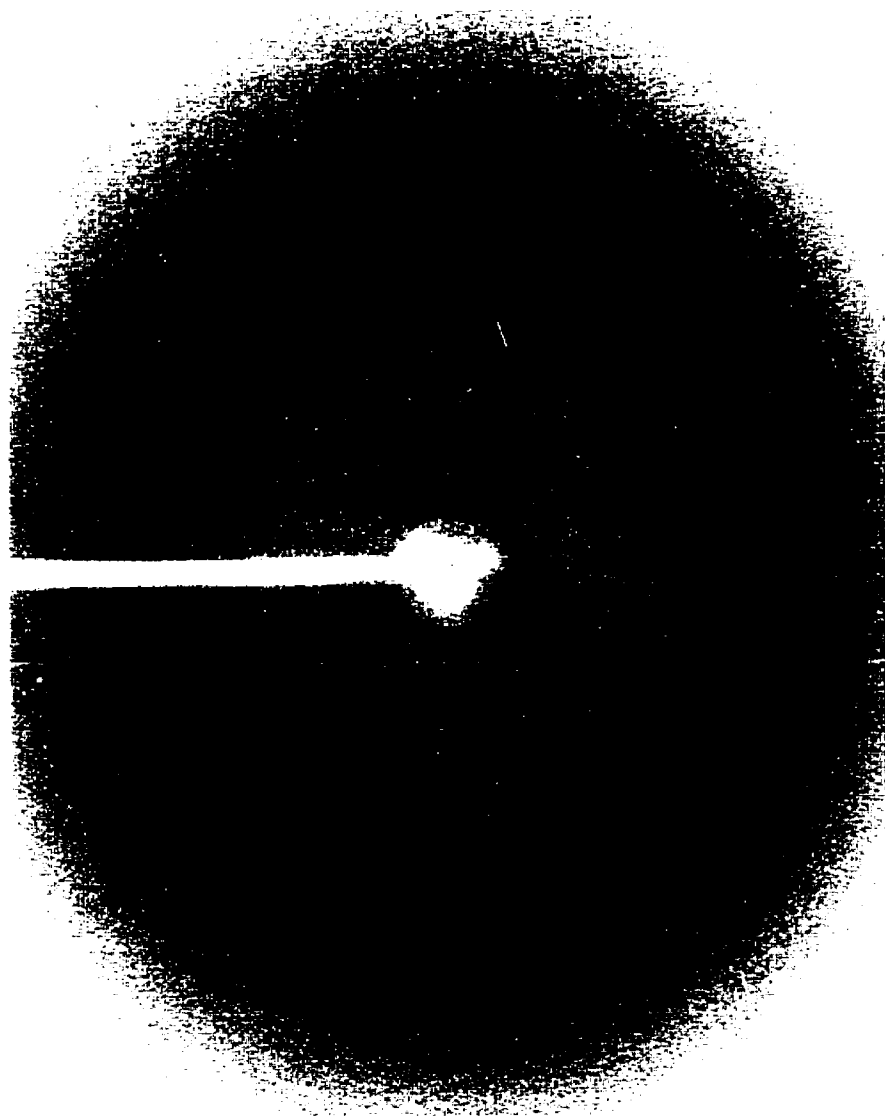


Figure # 12 X-ray diffraction pattern of A-protein. Reflections were recorded to a resolution limit of 5 Å. Data was collected at room temperature using a MAR 345 image plate with a crystal to detector distance of 180 mm and 15 min exposure. The preliminary space group determination was P2 or P2₁ with unit cell dimensions $a=319.7$ $b=175.3$ $c=350.0$ Å and $\beta=93.5^\circ$.

by solvent). The small number of observed reflections and weakness of the data may have resulted in incorrect indexing and space group assignment. Although improving the isoelectric microheterogeneity of the A-protein sample significantly enhanced X-ray diffraction, there still remained internal disorder in the crystals. The resolution limit of 5 Å was too low to permit a structure determination.

4.3 Assessment of oligomeric state

4.3.1 Dynamic light scattering

Many surface layer proteins have demonstrated the ability to spontaneously self assemble into two-dimensional arrays. As a result, solutions of purified surface layer proteins may contain a number of protein complexes of different sizes. The presence of various oligomerization states is another type of microheterogeneity found in protein samples that can prevent or lead to crystals with significant disorder. A-protein has been shown to spontaneously reform into A-layer sheets under the appropriate conditions. Therefore, in view of the limited resolution of A-protein crystals, an assessment of the multimeric state of the A-protein sample was undertaken by dynamic light scattering.

A-protein was purified in the presence of 4 mM CaCl₂, and analyzed at a concentration of 2 mg/ml using a standard dynamic light scattering protocol. Stable scattered light intensity counts could not be achieved indicating extreme heterogeneity and/or fluctuation in the multimeric state of the protein. Therefore, this sample was unable to yield useful data and no hydrodynamic radius could be determined.

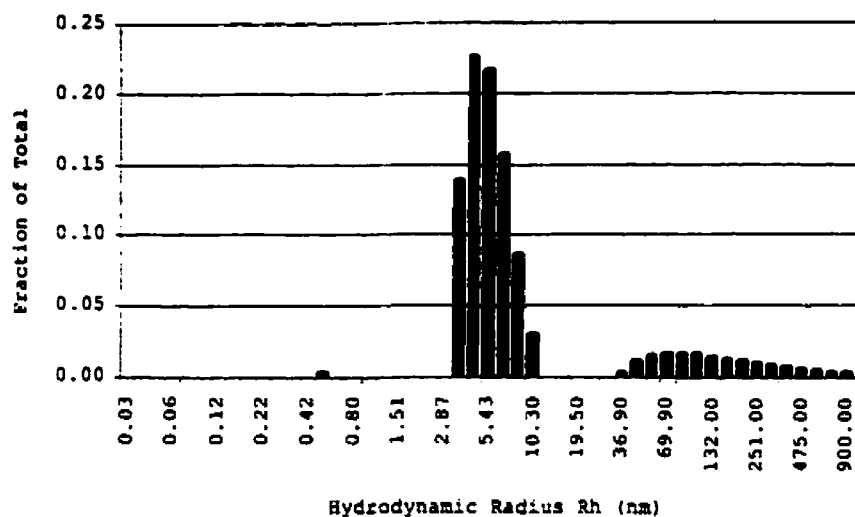
The presence of divalent cations is often required for surface layers to maintain their structural integrity. For A-protein, the presence of Ca²⁺ is believed to enhance

protein-protein interactions. Therefore, samples of A-protein were obtained in the absence of Ca^{2+} and analyzed for multimeric heterogeneity. A-protein originally purified in the presence Ca^{2+} was dialyzed against 20 mM Tris-Cl pH 7.7, 4 mM EGTA and 50 mM NaCl to obtain a Ca^{2+} free preparation. A 2 mg/ml sample was obtained and analyzed by dynamic light scattering. In this case, the scattered light intensity counts were stable and an autocorrelation function could be determined. Using the Dynals particle size distribution analysis, a monodisperse peak with a corresponding hydrodynamic radius of 4.91 nm was obtained (Figure 13). However, this analysis also revealed the presence of a large polydisperse molecular weight aggregate in the Ca^{2+} free sample. Increased concentrations of EGTA had no further effect on the radius obtained for the monodisperse peak or the large molecular weight aggregate.

In an attempt to eliminate the large molecular weight aggregate from the protein sample, non-ionic detergents were used. The effect of including either Tween-20 or NP-40 was investigated. Samples of 2 mg/ml A-protein in the presence of 2.0% Tween-20 gave stable scattered light intensity counts enabling calculation of a hydrodynamic radius of 4.91 nm. This result was similar to the previous sample containing EGTA, however, the presence of detergent prevented the formation of large molecular weight aggregate (Figure 13).

The determination of the molecular weight of the monodisperse A-protein species was attempted using a Zimm plot analysis. The linear equation plotted in this analysis is dependent on internal interference or shape of the molecule. However, this value is near unity as the wavelength of laser light is fifty times larger than the size of the observed

(a)



(b)

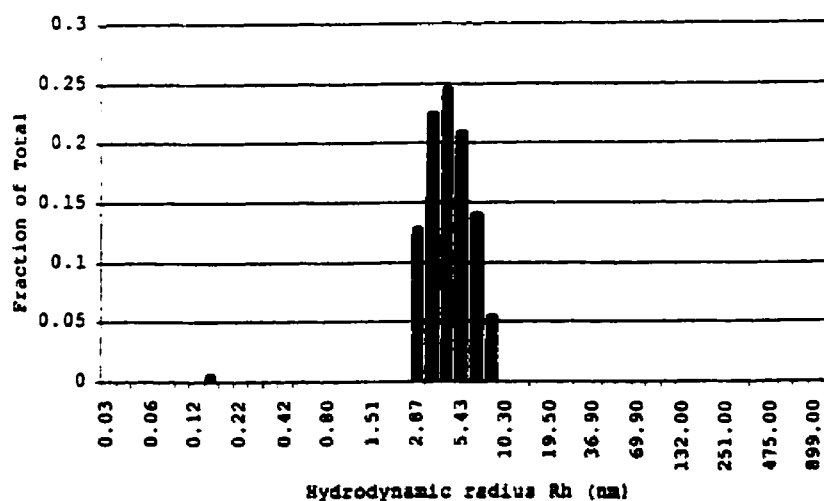
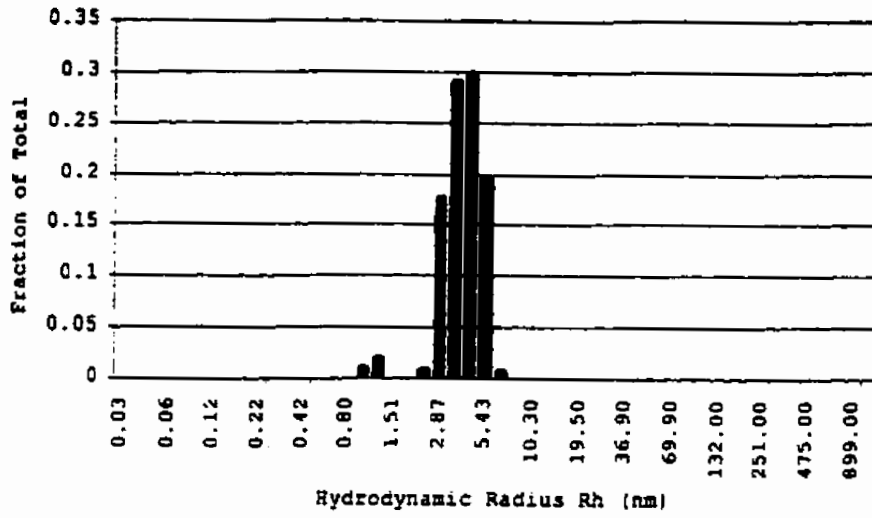


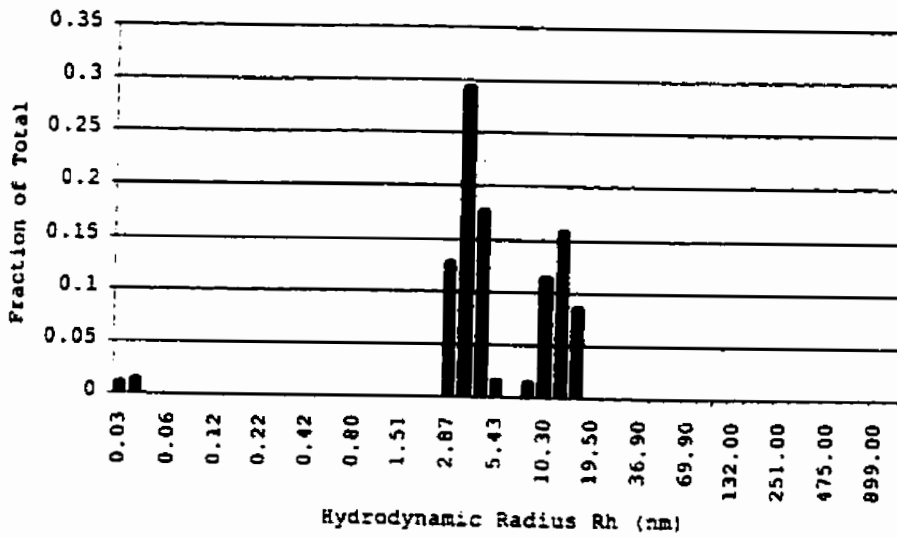
Figure #13 Particle size distribution analysis of A-protein sample conditions. Panel (a) represents 2.0 mg/ml A-protein purified in the presence of Ca^{2+} with subsequent addition of 4 mM EGTA. Contains a peak at mode (position) 4.91 nm plus a large molecular weight aggregate. Panel (b) represents 2.0 mg/ml A-protein in the presence of 2.0% v/v Tween-20. Contains a peak at mode (position) 4.91 nm with the absence of large molecular weight aggregate. (Figure prepared by Ivan Chebib).

Figure #14 Particle size distribution analysis of A-protein samples of increasing concentration. Panel (a) contains 0.5 mg/ml A-protein-EGTA, mode (position) = 4.91 nm, mean = 4.92 nm. Panel (b) contains 2.0 mg/ml A-protein-EGTA, peak1: mode (position) = 3.97 nm, mean = 4.21 nm, peak2: mode (position) = 14.0 nm, mean = 14.2 nm. Panel (c) contains 9.0 mg/ml A-protein-EGTA, subpeak1: mode (position) = 3.21 nm, mean = 3.49 nm, subpeak2: mode (position) = 11.5 nm, mean = 12.4 nm, all in the presence of 2.0% Tween-20. (Figure prepared by Ivan Chebib)

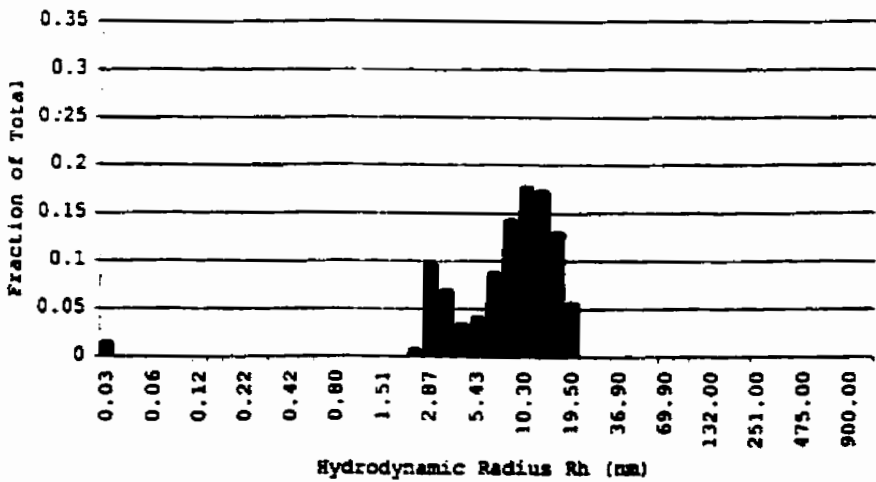
(a)



(b)



(c)



particle. Therefore, this essentially is a shape independent method for the determination of sample molecular weight. A-protein in the presence of 4 mM EGTA and 2.0% Tween-20 was prepared at concentrations of 0.25 to 10 mg/ml. However, the resulting Zimm plot was not linear and an accurate molecular weight could not be determined.

Investigation of the Dynals particle size distribution analysis obtained for the varying A-protein concentrations revealed the hydrodynamic radius obtained was not constant. As the concentration of A-protein increased, a second species emerged. This possessed a larger hydrodynamic radius and therefore, was a larger molecular weight species (Figure 14).

Since the Zimm plot analysis was unsuccessful in obtaining an accurate molecular weight of the monodisperse A-protein species, an approximate value was determined. The molecular weight of a protein species can be estimated from the hydrodynamic radius obtained. This is achieved through the application of data obtained from a model system with a defined frictional coefficient. In this case, two protein models were used. Since A-protein monomers have demonstrated the presence of two lobes connected with a small linker, a calmodulin model was used to approximate the molecular weight. For the monodisperse peak with a hydrodynamic radius of 4.91 nm a molecular weight of 50 kDa was calculated. However, when a globular protein model was used, a molecular weight of 200 kDa was determined. Since A-protein is approximately 50 kDa in size, the primary species in solution is likely either a monomer or a tetramer.

4.4 Crystallization with reduced oligomeric heterogeneity

4.4.1 Crystallization

The presence of a monodisperse protein species in solution greatly enhances the probability a crystallization experiment will be successful. Therefore, the information obtained from dynamic light scattering was used to determine sample conditions with a greater likelihood of leading to well-ordered crystals of A-protein. The removal of Ca^{2+} was shown to reduce the oligomeric heterogeneity and provide a monodispers protein sample, while the inclusion of the non-ionic detergent Tween-20 reduced the formation of large molecular aggregates. New crystallization trials were performed incorporating these modified sample conditions.

A-protein purified in the presence of Ca^{2+} was dialyzed against 20 mM Tris-Cl pH 7.7, 4 mM EGTA, and 50 mM NaCl. Protein samples of 2 mg/ml were prepared with the addition of 1.0% to 2.0% Tween-20 and left over night at 4°C to equilibrate (A-protein-EGTA-Tween). Crystallization experiments were conducted using the hanging drop technique with the previously optimized crystallization conditions. These trials resulted in the formation of mild precipitation in the hanging drops but no protein crystals. Therefore, with the new sample conditions, it was necessary to screen for new crystallization conditions.

The A-protein-EGTA-Tween sample was screened using Crystal Screen I and II at 4°C and room temperature. The best lead was obtained from Crystal Screen I tube 37 in which a new crystal form of A-protein was obtained. Using this condition, small crystals were obtained directly from the initial crystal screen. Optimization was



Figure #15 Crystal of A-protein grown in the presence of 4 mM EGTA and 1% Tween-20 using 5% PEG 4000 (w/v), 0.1 M sodium acetate pH 4.3 at 18°C. Crystal dimensions are 0.4 x 0.3 x 0.2 mm with a cubic morphology. These crystals diffracted to 3.5 Å in the C2 space group.

performed through systematic variations of the PEG concentration, the type of PEG used, pH and temperature. As a result of this, two separate crystallization conditions were optimized yielding crystals large enough to proceed with X-ray diffraction studies.

In both cases, the hanging drop technique was used with 1000 μ l reservoirs and drops containing 6 μ l of protein solution mixed with 6 μ l of reservoir solution. The first optimized crystallization condition consisted of the following:

- 10% PEG 1000 (w/v), 0.1 M sodium acetate pH 4.5, at 21°C

This resulted in A-protein crystals that grew to 0.4 x 0.1 x 0.1 mm in size over a two day period. These crystals appeared to possess tooth like cavities at either end and often grew irregular in shape. The second optimized crystallization condition differs primarily in the type and concentration of PEG as follows:

- 5% PEG 4000 (w/v), 0.1 M sodium acetate pH 4.3, at 18°C

The larger PEG polymer used at a lower concentration and temperature seemed to improve crystal growth. The A-protein crystals obtained using this condition grew larger in size at 0.4 x 0.3 x 0.2 mm over a period of seven days. These crystals generally appeared well formed with smooth faces and sharp edges as seen in Figure 15. Crystals grown using both conditions were subsequently tested for X-ray diffraction.

4.4.2 Diffraction studies

Crystals of A-protein obtained using both crystallization conditions were mounted and subjected to diffraction analysis at room temperature. The crystals grown in 10% PEG 1000 diffracted poorly and no useful data could be collected. However, the crystals grown in 5% PEG 4000 diffracted significantly better. Reflections were recorded out to a

resolution range of 3.5 Å (Figure 16). These diffraction patterns indicated a significant improvement in the internal order of the A-protein crystals through the reduction in oligomeric heterogeneity. The new crystals of A-protein formed in a C-centered monoclinic space group (C2) with the unit cell dimensions $a= 150.8$ $b= 63.3$ $c= 66.6$ and $\beta= 114.9^\circ$. The dimensions of this unit cell are likely more realistic. Calculation of the unit cell volume (assuming 80-30% of the cell volume occupied by solvent) indicated the asymmetric unit would contain 1 to 4 A-protein molecules. Since these crystals diffracted to a useful resolution range, a complete data collection was attempted. The data collection was set up to obtain a 180° sweep of diffraction data at 1° intervals, with an exposure time of 10 minutes per image. However, the A-protein-EGTA-Tween crystals were highly sensitive to X-ray induced damage and the diffraction quality failed rapidly after only a few images.

In protein crystallography, the decay of crystals in the X-ray beam is a common problem. X-ray induced radiation damage can be significantly reduced through the use of cryocrystallographic methods. This entails freezing the crystal in a cryogen and maintaining it at 100°K during the data collection. For A-protein crystals, the first step was to determine a suitable cryoprotectant to add before freezing the crystals. The cryoprotectant acts to prevent the formation of ice crystals that would damage the fragile protein crystal. Various concentrations of glycerol, MPD, and PEG were screened as possible candidates to use. The condition offering the best cryoprotection without any visible crystal disruption was 10% glycerol, 15% PEG 4000 in 0.1 M sodium acetate at pH 4.3. Crystals were briefly soaked in cryoprotectant before mounting in 20 micron

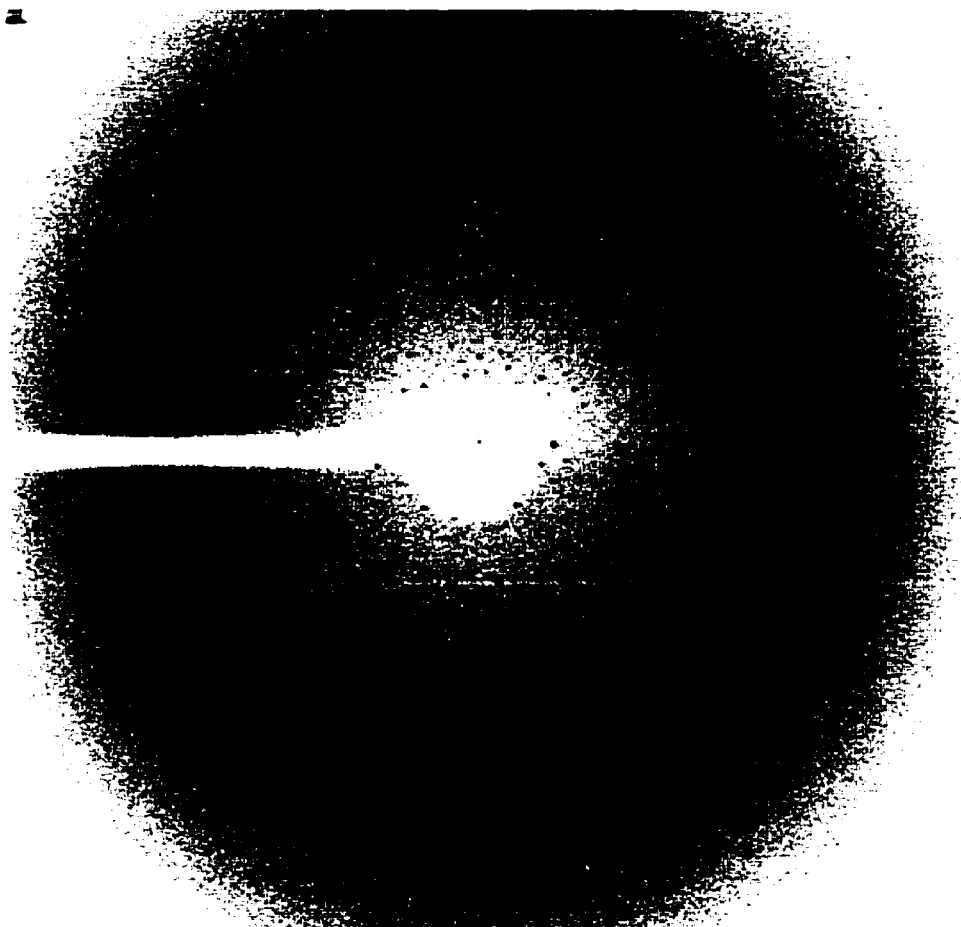


Figure #16 Diffraction pattern of an A-protein crystal grown in the presence of 1% Tween. Data was collected at room temperature using a MAR 345 image plate with a crystal to detector distance of 180 mm and 10 minute exposure time. Reflections were recorded to a resolution of 3.5 Å enabling a preliminary space group determination of C centered monoclinic with unit cell dimensions $a= 150.8$ $b= 63.3$ $c= 66.6$ and $\beta= 114.9^\circ$.

cryoloops and freezing in the liquid nitrogen cold stream.

These crystals were stored and taken to beam line X12C at the NSLS at the Brookhaven National Laboratory. Using a synchrotron X-ray source, data was collected using 1 to 2 minute exposure times over an oscillation angle of 1° . The resulting diffraction was strong out to 3.4 \AA but was cut off sharply at this point (data not shown). Unfortunately, the diffraction data could not be indexed and therefore no useful structural information could be obtained. Similar problems indexing frozen A-protein crystals was observed in house indicating these crystals do not tolerate the freezing process well.

5. RESULTS: CALMODULIN

5.1 Calmodulin sample

Lyophilized samples of CT-CaM, Eth-CaM, Dfm-CaM and Nle-CaM were resuspended to the desired protein concentration in 20 mM Tris-Cl pH 7.0, 5 mM CaCl₂ and maintained at 0 to 4°C at all times. The solubility of Eth-CaM and Dfm-CaM was excellent in this solution however, CT-CaM and Nle-CaM required substantial time before complete resuspension was obtained. The quality of the protein preparations were assessed by SDS-PAGE. The purity of the CaM protein samples appeared greater than 95% (data not shown) and crystallization trials were initiated.

5.2 Crystallization

The crystallization conditions used in the structure determination of wild type mammalian CaM (rat testis) were reported by Babu *et al* in 1985 (112). Since then, similar conditions have been used to crystallize CaM proteins from a variety of sources. In general, CaM proteins have responded best to solutions containing 2-methyl-2,4-pentanediol (MPD) buffered in sodium acetate in a pH range of 4.5 to 5.6 (112, 118, 145, 146, 147). The concentration of MPD required, temperature of crystallization, protein concentration used and the presence of other additives all vary to a certain extent in the reported CaM crystallizations. Since these conditions have proven successful, they served as a starting point in the attempt to crystallize the four mutated or biosynthetically altered CaM proteins.

5.2.1 CT-CaM

Crystallization experiments were set up as described by Babu *et al* (112) using 12 mg/ml CT-CaM solutions at room temperature. Under these conditions, the trials produced extensive amorphous precipitate. Subsequent reduction in the protein concentration used, in addition to variation of the MPD concentration and pH, failed to induce crystal formation.

The crystallization of recombinant CaM as reported by Chattopadhyaya *et al* in 1992 (118) required the inclusion of ethanol for crystal formation. Therefore, a number of alcohols were included in the CT-CaM screens in an attempt to induce crystallization. The presence of ethanol resulted in the formation of crystalline precipitate in trials set up at room temperature and in the formation of microcrystals when crystallizations were performed at 4°C. This initial lead was optimized to the crystallization condition listed in Table 3. Using hanging drops containing 6 µl of protein at a concentration of 8 mg/ml mixed with 6 µl of reservoir solution, small crystals were obtained in 4 to 7 days. The presence of significant quantities of amorphous precipitate often hampered the growth of CT-CaM crystals. However, the crystals obtained were used to macroseed pre-equilibrated hanging drops in order to induce formation of large CT-CaM crystals. Long rod-like crystals were grown from the smaller seeds in about 4 to 6 weeks as seen in Figure 17. Crystals of CT-CaM grew to approximately 1.6 x 0.2 x 0.1 mm in size.

5.2.2 Eth-CaM

Crystallization trials conducted using Eth-CaM followed a similar procedure as described for CT-CaM. Initial trials were performed using the crystallization conditions

Table3. Optimized crystallization condition used for the CaM proteins

Protein	Sample Conditions	Optimized Crystallization Conditions
CT-CaM	8.0 mg/ml protein in 20 mM Tris-Cl pH 7.0, 5 mM CaCl ₂	40% (v/v) MPD, 15% (v/v) ethanol, 5 mM CaCl ₂ in 0.1 M sodium acetate pH 4.5 at 4°C
Eth-CaM	10.0 mg/ml protein in 20 mM Tris-Cl pH 7.0, 5 mM CaCl ₂	50% (v/v) MPD, 15% (v/v) ethanol, 5 mM CaCl ₂ in 0.1 M sodium acetate pH 4.5 at 4°C
Dfm-CaM	10.0 mg/ml protein in 20 mM Tris-Cl pH 7.0, 5 mM CaCl ₂	45% (v/v) MPD, 10% (v/v) ethanol, 5 mM CaCl ₂ in 0.1 M sodium acetate pH 4.6 at 4°C
Nle-CaM	5.0 mg/ml protein in 20 mM Tris-Cl pH 7.0, 5 mM CaCl ₂	48% (v/v) MPD, 15% (v/v) ethanol, 5 mM CaCl ₂ in 0.1 M sodium acetate pH 4.2 at 4°C 9.0 mM 1-s-Octyl-β-D-thioglucoiside

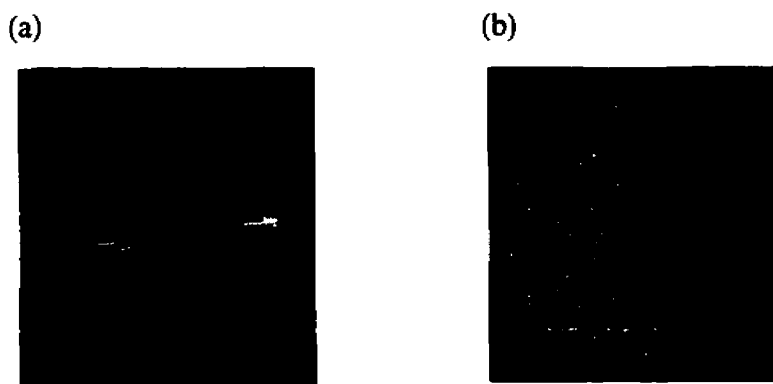


Figure #17 Crystals of CaM displaying rod-like (needle) morphology. Panel (a) contains twinned CT-CaM crystals 1.6 x 0.2 x 0.1 mm in dimension. Panel (b) contains twinned Eth-CaM crystals of similar dimensions. Crystals of Dfm-CaM are identical in appearance to these and are not shown.

reported by Babu *et al.* Eth-CaM displayed greater solubility in these conditions than CT-CaM, producing only limited amounts of amorphous precipitate. Despite the improved solubility, crystallization could not be induced. Therefore, the effects of including various concentrations of ethanol and of reducing the incubation temperature were explored. As with CT-CaM, crystals of Eth-CaM could be produced in the presence of ethanol when crystallizations were conducted at 4°C. The optimized crystallization conditions for Eth-CaM are given in Table 3. This condition differs from CT-CaM in that a larger protein concentration of 10 mg/ml is used, and a greater amount of the primary precipitant MPD (50%) was required for crystal formation. This is likely a result of Eth-CaM's greater solubility. As before, hanging drops containing 6 µl of protein mixed with 6 µl reservoir solution were used. Small crystals grew readily in about 3 to 4 days which were further used to macroseed pre-equilibrated hanging drops. Again, long rod-like crystals grew from the smaller seeds in about 4 to 6 weeks appearing similar to the crystals in Figure 17. The Eth-CaM crystals grew to approximately 2.0 x 0.2 x 0.1 mm in size.

5.2.3 Dfm-CaM

The initial crystallization trials conducted with Dfm-CaM gave results similar to those obtained with Eth-CaM. The protein demonstrated good solubility when used at high concentration and produced little amorphous precipitate. As with the other CaM proteins, the presence of ethanol and incubation at 4°C were required for nucleation and crystal formation. Following the optimization procedure, the final crystallization condition for Dfm-CaM was determined and is listed in Table 3. Crystallization of Dfm-CaM required a larger protein (10 mg/ml) and MPD (45%) concentrations than seen with

CT-CaM, however, the best crystal growth was obtained with a reduced ethanol content of 10%. As for CT-CaM and Eth-CaM, large hanging drops produced small crystals that were used to seed growth of larger crystals. The Dfm-CaM crystals grew to approximately 2.0 x 0.2 x 0.1 mm in size similar to the crystals in Figure 17.

5.2.4 Nle-CaM

Nle-CaM demonstrated poor solubility in both the crystallization conditions reported by Babu *et al* and those used to crystallize the other CaM mutants. Extensive amorphous precipitate was observed in all crystallization experiments. However, in the trials containing MPD/ethanol at 4°C small amounts of crystalline precipitate and clusters of crystal matter were observed amongst the amorphous precipitate. An extensive screening of various precipitate concentrations, protein concentrations, and additives led to little improvement. The abundance of amorphous precipitate in Nle-CaM crystallization trials suggested hydrophobic protein-protein interactions occur, resulting in non-specific aggregation. Non-specific aggregation can prevent or interfere with crystallization of the protein sample. To alleviate this problem, a number of detergents were included in the crystallization trials. From a screen of twenty-four detergents, the presence of 9 mM 1-s-Octyl- β -D-thioglycoside gave significant improvement. The condition listed in Table 3 resulted in the formation of very small rod like crystals in 2 to 3 weeks. These tiny crystals were subsequently used to macroseed pre-equilibrated hanging drops. However, the seeding experiments with Nle-CaM were unsuccessful as no crystals grew large enough to attempt X-ray diffraction studies.

5.3 Detwinning Crystals

During the growth of protein crystals, a form of intergrowth can sometimes occur. The formation of a twin results from either intergrown individuals joined symmetrically about an axis or a plane, or from the growth of individual crystals on top of each other (parallel growth) in such a way that all corresponding faces and edges are parallel (148). Examination of crystals using a microscope equipped with a crossed polarizer can often detect the presence of a twin. As the crystal is rotated on the stage, the polarized light is often extinguished at different points by the individual crystal lattices, revealing the presence of twins. Failing this means of detection, analysis of the subsequent diffraction pattern would reveal the presence of a twinned crystal.

Before the CaM crystals were mounted and tested for X-ray diffraction, they were examined for the presence of a twin. Using a polarized light source, all the CT-CaM, Eth-CaM and Dfm-CaM crystals obtained displayed signs of twinning. In all cases, the CaM crystals presented as parallel growth twins with at least two independent crystals joined along the long axis of the crystal. The two independent crystals could be clearly differentiated since their coloration was different under polarized light. Confirmation of the twinned state of these CaM crystals was obtained from analysis of X-ray diffraction images. These demonstrated characteristic spot doubling and overlaps in addition to the inability to determine a constant unit cell as the crystals are rotated in the X-ray beam.

The initial attempt to eliminate the presence of the parallel growth twins in the CaM crystals involved chemical modification of the mother liquor. Including certain additives during crystal growth has proven successful in preventing twinning in a number of cases (148, 149). Using concentrations ranging from 0.5 to 10 % (v/v), the effect of

adding glycerol, dioxane, spermine, acetone, various alcohols, etc. to the reservoir solution was determined. Unfortunately, no additive tested was successful in preventing the growth of parallel twins.

Since modifying the crystallization conditions failed to solve the problem, attempts were made to physically separate the adjacent twins. Protein crystals are extremely fragile and are often readily damaged by physical manipulation. Therefore, using a micro-needle tool (Hampton Research), a steady delicate touch was required to attempt twin separation. Using the polarized light to distinguish between the twins, the thinner twin was selected for removal through a careful process of picking away the crystal matter. The selected twin was completely destroyed during this procedure and often this manipulation resulted in irrevocable damage to both crystals. Many attempts were made to obtain segments of untwined CaM crystals until the procedure was eventually successful. For CT-CaM, a fragment of $0.8 \times 0.1 \times 0.05$ mm was obtained which proved to be free of the parallel growth twin. Subsequent X-ray diffraction studies confirmed the removal of the twined crystal. For Eth-CaM and Dfm-CaM, fragments of $1.0 \times 0.1 \times 0.05$ mm and $0.5 \times 0.1 \times 0.05$ mm were obtained, respectively. These also proved to be free of the parallel growth twins and were used in X-ray diffraction studies.

5.4 Diffraction Data

The detwinned crystals of CT-CaM and Eth-CaM were mounted in quartz capillary tubes with the length of the needle oriented along the length of the tube and X-ray diffraction data was collected at room temperature.

5.4.1 CT-CaM

For CT-CaM, an initial diffraction pattern, as seen in Figure 18, was obtained and indexed to obtain preliminary unit cell and space group information. Using this information, a data collection strategy was determined that would capture the required amount of data to yield a complete data set. CT-CaM crystallized in the triclinic space group P1, which possesses no rotational crystallographic symmetry. In the absence of rotational symmetry, all data collected in a 180° sweep around the phi axis of the goniometer (crystal mount) is unique. Therefore, it was necessary to rotate the crystal through at least 180° to collect a complete data set. Data was collected using 1.5° oscillations over a range of 0 to 240°. The 1.5° oscillation range was used since these crystals tended to decay significantly after extended exposure to X-rays. This enabled collection of the required range of data in a shorter time period before significant crystal decay. The detector was set at a distance of 180 mm from the crystal and 10 minute exposure times were used per 1.5° of data collected. The initial diffraction was strong and a complete data set was obtained containing 0 to 180° of data. However, the CT-CaM crystal demonstrated significant decay in the images collected from 180 to 240° and this redundant data could not be included.

5.4.2 Eth-CaM

The data collection procedure for Eth-CaM was essentially identical to that of CT-CaM. The initial diffraction pattern (Figure 19) revealed a triclinic space group and the same data collection parameters were used. In this case, the Eth-CaM crystal showed

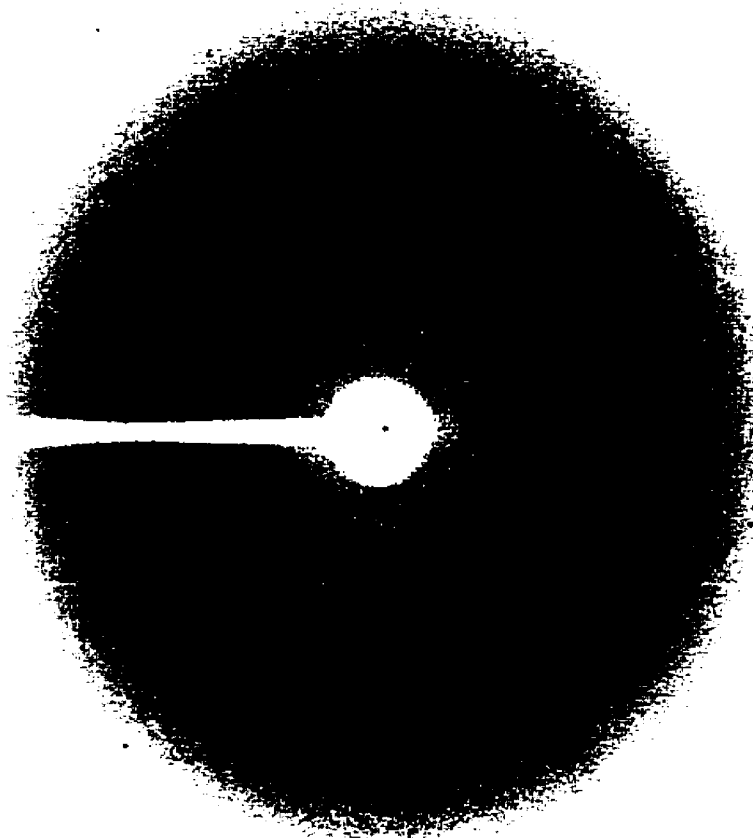


Figure #18 Diffraction pattern of CT-CaM crystal. Data collected at room temperature using a MAR 345 image plate, crystal to detector distance of 180 mm and 10 minute exposure time. Reflections were recorded to 2.5 Å giving the space group P1 with unit cell dimensions $a= 30.01$ $b= 53.89$ $c= 25.02$ Å and $\alpha= 93.47^\circ$ $\beta= 96.76^\circ$ $\gamma= 88.63^\circ$.

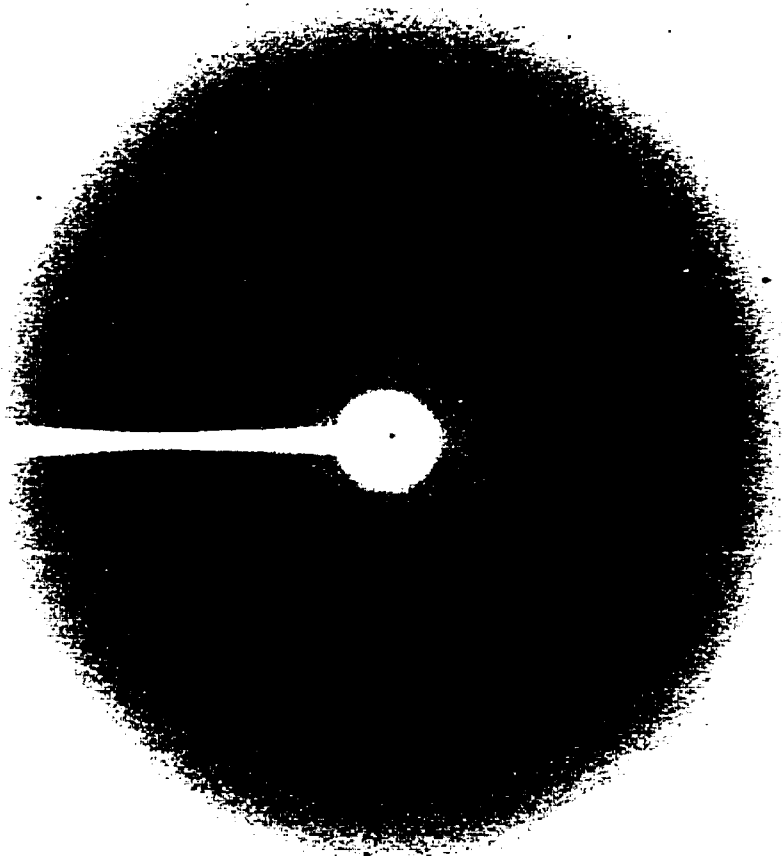


Figure #19 Diffraction pattern of Eth-CaM crystal. Data collected at room temperature using a MAR 345 image plate, crystal to detector distance of 180 mm and 10 minute exposure time. Reflections were recorded to 2.35 Å giving the space group P1 with unit cell dimensions $a=29.95$ $b=53.86$ $c=25.15$ Å and $\alpha=94.19^\circ$ $\beta=96.61^\circ$ $\gamma=88.73^\circ$.

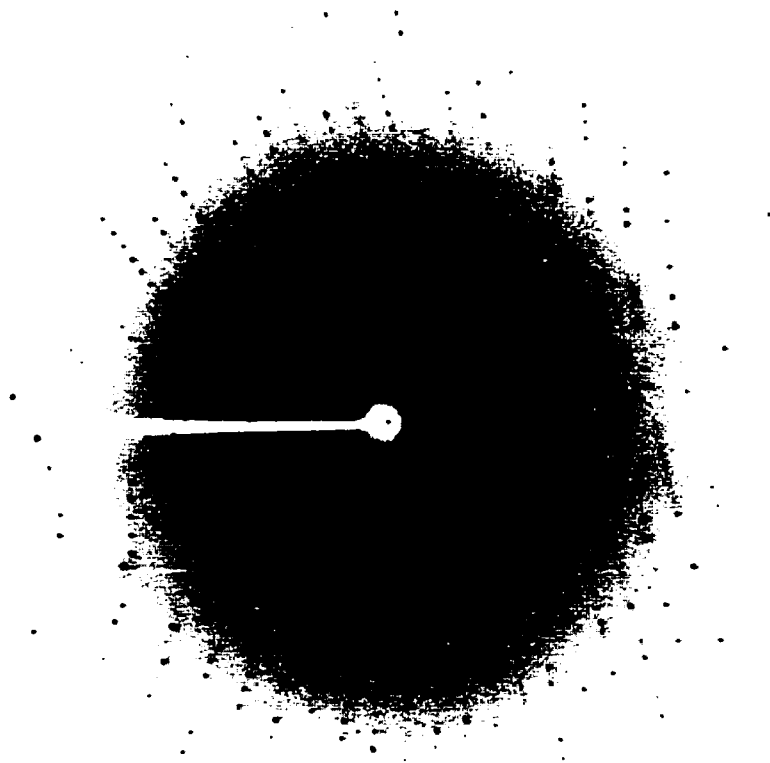


Figure #20 Diffraction pattern of Dfm-CaM crystal. Data collected at 100°K using a MAR 345 image plate, and Oxford Cryosystems Cryostream. A crystal to detector distance of 120 mm was used with a 10 minute exposure time. Reflections were recorded to 1.8 Å giving the space group P1 with unit cell dimensions $a=58.75$ $b=54.08$ $c=24.42$ Å and $\alpha=94.19^\circ$ $\beta=96.61^\circ$ $\gamma=88.73^\circ$.

only minor decay in the 180 to 240° range and this redundant data was included in subsequent processing.

5.4.3 Dfm-CaM

For Dfm-CaM, data was collected at 100°K using cryo-mounting procedures. The initial diffraction pattern obtained for cryo-mounted Dfm-CaM is seen in Figure 20. The Dfm-CaM crystals could be frozen directly from the mother liquor by scooping a crystal out with a cryoloop and freezing directly in the cold stream. No additional cryoprotectant was required. As with CT-CaM and Eth-CaM, the space group obtained was triclinic requiring 180° of data for a complete data set. The Dfm-CaM crystal diffracted significantly better than the previous CaM proteins requiring a reduction in the crystal to detector distance. By moving the detector closer to the crystal, reflections diffracting to a higher resolution (larger scattering angle) can be recorded at the edge of the image plate detector. Therefore, data was collected at a distance of 120 mm. Since the crystal was frozen at 100°K, there was little problem of crystal decay and 1° oscillations were used with 10 minute exposure times. Data was collected from 0 to 240° with no apparent decay in reflection intensity.

5.5 Data Processing

Once a complete data set is collected, analysis and reduction of the diffraction data is required. For single crystal diffraction data this consists of seven major steps (140).

1. Visualization and preliminary analysis of the original, unprocessed diffraction patterns
2. Indexing of each diffraction pattern

3. Refinement of the crystal and detector parameters
4. Integration of the diffraction maxima
5. Finding the relative scale factors between measurements
6. Precise refinement of crystal parameters using the whole data set
7. Merging and statistical analysis of the measurements related by space group symmetry

To accomplish this, the HKL software suite consisting of the programs Xdisplay, Denzo, and Scalepack, was used.

For each data set (CT-CaM, Eth-CaM and Dfm-CaM), the first four steps listed above were completed using the programs Xdisplay and Denzo. In this process, a single oscillation image was auto-indexed yielding a list of possible lattice types and unit cell dimensions. With this information the precise crystal and detector orientation parameters for the single image were determined. The crystal orientation parameters describe the orientation of the reciprocal lattice with respect to the spindle, beam and vertical axes of the image plate, while the detector and X-ray parameters include the wavelength, the crystal to detector distance and the precise coordinates of the direct beam (140). Next, all images in the data set were processed in series using the determined parameters. The output from this process was subsequently scaled, further refined and merged together using Scalepack, completing the final three step mentioned above. The quality of the data was assessed by the values obtained for χ^2 and R_{merg} statistics. The data processing statistics for CT-CaM, Eth-CaM, and Dfm-CaM are summarized in Table 4.

Table 4. Data processing statistics for CT-CaM, Eth-CaM and Dfm-CaM.

	CT-CaM	Eth-CaM	Dfm-CaM
Space group	P1	P1	P1
Unit cell (a, b, c in Å) (α, β, γ in °)	$a=30.01$ $\alpha=93.47$ $b=53.89$ $\beta=96.76$ $c=25.02$ $\gamma=88.63$	$a=29.95$ $\alpha=93.95$ $b=53.86$ $\beta=96.46$ $c=25.15$ $\gamma=88.57$	$a=58.75$ $\alpha=94.19$ $b=54.08$ $\beta=96.61$ $c=24.42$ $\gamma=88.73$
Unit cell volume (Å ³)	40102.6	40210.2	76846
Unique reflections	4878	5724	17516
Resolution (Å)	50 – 2.5	50 – 2.35	50 – 1.8
Completeness	89%	89.5%	85.2%
χ^2 value	1.07	1.07	1.05
R_{merg}	0.045	0.055	0.030

For CT-CaM, and Eth-CaM, the final merged data sets showed good completeness to 2.5 and 2.35 Å respectively, with χ^2 and R_{merg} values indicative of data with good internal consistency. The space group and final refined unit cells are the same as those observed previously for wild type CaM (112, 118, 145, 146, 147). For Dfm-CaM, the final merged data set also displays an acceptable level of completeness, but to a significantly higher resolution of 1.8 Å. The values obtained for the χ^2 and R_{merg} from this data set also indicate good quality data. The interesting feature of this data is found in the final refined unit cell. As seen in Table 4, these P1 CaM crystals contain a doubled a axis compared to previously reported CaM unit cells.

5.6 Structure Determination and Refinement

Having obtained processed data sets of CT-CaM, Eth-CaM, and Dfm-CaM, the next step is to produce an image of the molecule from the crystallographic data. To accomplish this, a computer is used to simulate the action of a lens by computing the electron density within the unit cell from the amplitude and phase of each diffracted X-ray beam or reflection (150). The relationship between an object and its diffraction data can be described using a Fourier transform. This enables conversion of a Fourier-series description of the diffraction spots or reflections, called a structure-factor equation, to a Fourier-series description of the electron density, called the electron density equation (150). Each structure factor consists of an amplitude and phase describing the diffracted X-ray wave giving rise to a particular reflection. The amplitudes are obtained from the list of indexed intensities, however, X-ray diffraction patterns fail to capture the phase of the reflection, so this vital piece of information is missing. Therefore, computation of the

electron density is impossible from the data available in native diffraction patterns. To solve this problem, molecular replacement was used to obtain the phase information for CT-CaM, Eth-CaM, and Dfm-CaM. Molecular replacement uses a known protein structure as a phasing model to calculate initial phases. This is done by placing the model structure in the unit cell and finding the position and orientation that yields phases most like those of the new protein (151). With starting phases, a structure determination of the new CaM proteins could be pursued.

5.6.1 CT-CaM

In the attempt to solve the structure of CT-CaM by molecular replacement, the structure of mammalian CaM was used as the initial phasing model. The atomic coordinates of the 1.7 Å model of mammalian CaM (1CLL) solved by Chattopadhyaya (118) were obtained from the Protein Data Bank. This model contained 1,130 atoms of the polypeptide chain from residues 4 to 147 of CaM. The four calcium ions and all the water molecules were removed from the initial model. In addition, the Met residues at positions 109, 124, 144, and 145 were replaced with alanine (Ala) residues. Since in CT-CaM, these positions have been mutated to Leu residues, the inclusion of side chain coordinates could have biased the initial electron density maps.

The first step in determining a molecular replacement solution was the rotation search to find the correct rotational orientation of the phasing model in the unit cell of CT-CaM. The rotation function using data in the 15 – 4 Å range gave a clear solution at Euler angles close to zero. Therefore, the phasing model and CT-CaM were essentially isomorphous. This was expected, as the only difference should be in the C-terminal

hydrophobic clefts. Since CaM crystallized in the triclinic (P1) space group there is only one translational position available in the unit cell. Therefore, with the phasing model in the proper rotational orientation, a rigid body refinement was done yielding the initial solution for CT-CaM. The initial R-value, defined as a measurement of overall agreement between the observed amplitudes and those calculated from the model, was 0.336. For proteins, the placement of a phasing model yielding R-values in the range of 0.3 – 0.4 often provides adequate initial estimates of phases (150). A $2|F_o|-|F_c|$ electron density map was calculated revealing interpretable electron density depicting the unit cell contents of CT-CaM.

Once the molecular replacement solution was found for CT-CaM, an iterative process of phase improvement and extension was conducted. The success of refinement and quality of the subsequent model was assessed by monitoring both the R-value, as mentioned above, and the free R-value as a cross-validation. In both cases, a reduction in the resulting value indicates improvement of phase accuracy during the refinement step. The refinement of the CT-CaM model was initiated by a rigid body refinement in which each of the C-terminal and N-terminal domain positions were allowed to be optimized independently. Simulated annealing using all the available data from 50 to 2.5 Å was subsequently done. This included initially “heating” the system to 5000 K and slowly “cooling” in increments of 25 K. This allows the model to move as if at high temperature, lifting it out of local energy minima, and then cool to find its preferred conformation (150). This gave an R-value and free R-value of 0.292 and 0.404 respectively. The progress of the R-values and free R-values during refinement are listed in Table 5. At this point, the initial model had the same B-factor, or temperature factor,

Table 5. Progress of CT-CaM and Eth-CaM structure determination as measured by the R-value and free R-value obtained at various stages of refinement.

	CT-CaM		Eth-CaM	
	R-value	free R-value	R-value	free R-value
Molecular Replacement solution	0.336	N/A	0.328	N/A
Simulated annealing of initial model	0.292	0.404	0.309	0.397
Group B-factor refinement	0.276	0.375	0.298	0.378
Group B-factor with Ca ²⁺	0.254	0.356	0.279	0.358
Simulated annealing of rebuilt model	0.235	0.352	0.230	0.315
Individual B-factor refinement	0.232	0.333	0.226	0.310
Final refined model	0.207	0.305	0.217	0.305

for all atoms. A group B-factor refinement was done which refined B-factors for the main-chain atoms and the side-chain atom separately. This resulted in a significant reduction in both R and free R-values.

Having obtained a good initial model that was amenable to subsequent refinement, the next step was to build in the missing parts. First, the locations of the four calcium ions were determined through analysis of $2|F_o|-|F_c|$ and $|F_o|-|F_c|$ electron density maps. Density for Ca^{2+} could clearly be seen in each of the four EF-hand binding motifs and these were built into the structure. This was followed by another group B-factor refinement and conjugate gradient minimization. Next, the mutated Leu residues (109, 124, 144, and 145) were built into the model along with a manual fitting of residues which did not properly occupy the available electron density. At this point, the residue at position 4 of CT-CaM was removed from the structure as a result of poor electron density. A subsequent cycle of simulated annealing, conducted as previously mentioned, resulted in a reduction in the R-value to 0.235 and the free R-value to 0.352. At this point, all the atoms in the model were subjected to an individual B-factor refinement followed by $2|F_o|-|F_c|$ and $|F_o|-|F_c|$ map calculations. These maps were used to incorporate solvent molecules into the structure which formed hydrogen bonds of less than 3.2 Å in length with the protein. The final model consisted of residues 5 –147, four calcium ions and 43 water molecules. The final R-value was 0.207 and the final free R-value was 0.305. The bond distances and bond angles had a rms deviation from ideal values of 0.0064 and 1.04 respectively. The position of the four C-terminal Met to Leu mutations are seen in Figure 21. This shows the Leu side chains clearly occupying the available electron density.

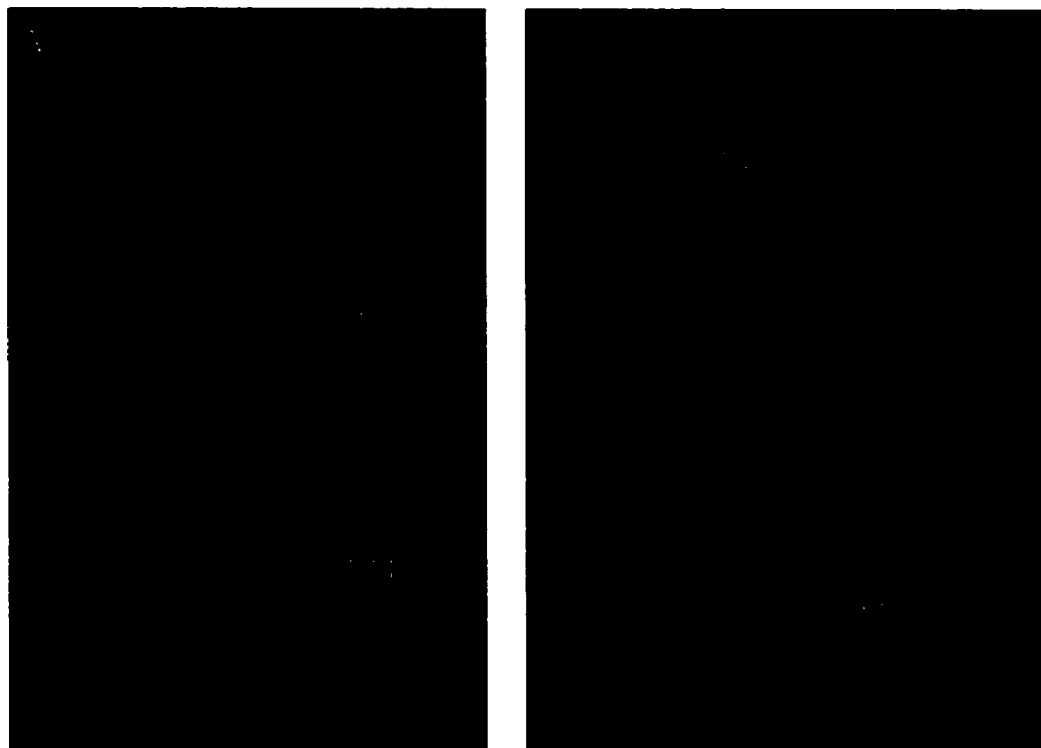


Figure #21 $2|F_o|-|F_c|$ electron density map of the 2.5 Å structure of CT-CaM contoured around the mutated Leu residues of the C-terminal hydrophobic cleft. Panel A contains Leu 109 and 124. Panel B contains Leu 144 and 145.

5.6.2 Eth-CaM

For the structure determination of Eth-CaM, the 1.7 Å model of mammalian CaM was again used as the initial phasing model in molecular replacement. In this case, the nine Met residues at positions 36, 51, 71, 72, 76, 109, 124, 144, and 145 were truncated to Ala and all calcium and water molecules were removed. The molecular replacement and subsequent refinement were conducted as described for CT-CaM. The rotation function gave a clear solution with Euler angles essentially zero indicating the model and Eth-CaM are also isomorphous. Following rigid body refinement the initial R-value was 0.328 and interpretable electron density was obtained.

A rigid body refinement allowing independent domain movement was done followed by simulated annealing using all the available data from 50 to 2.35 Å. After a subsequent group B-factor refinement, the resulting R-value and free R-value were 0.298 and 0.378 respectively. The progress of the R-values and free R-values during refinement of Eth-CaM are also listed in Table 5. At this point, the four missing calcium ions were placed in the model as determined by the electron density obtained from $2|F_o| - |F_c|$ and $|F_o| - |F_c|$ maps. Once another round of group B-factor and energy minimization were conducted, the Eth residues were built into the model. Eth residues could be placed at positions 36, 51, 71, and 72 of the N-terminal domain and at positions 109, 124, 144, and 145 of the C-terminal domain. However, the Eth residue at position 76, located in the central linker region, could not be placed since the electron density for this side chain was poor. In addition, a number of other residues were manually fit into available density before subsequent rounds of simulated annealing and individual B-factor refinement were conducted. At this stage of the refinement, the R-value and the free R-



Figure #22 $2|F_o|-|F_c|$ electron density map of the 2.35 Å structure of Eth-CaM contoured around the biosynthetically incorporated Eth residues of the N-terminal hydrophobic cleft. Panel A contains Eth 36 and 51. Panel B contains Eth 71 and 72.

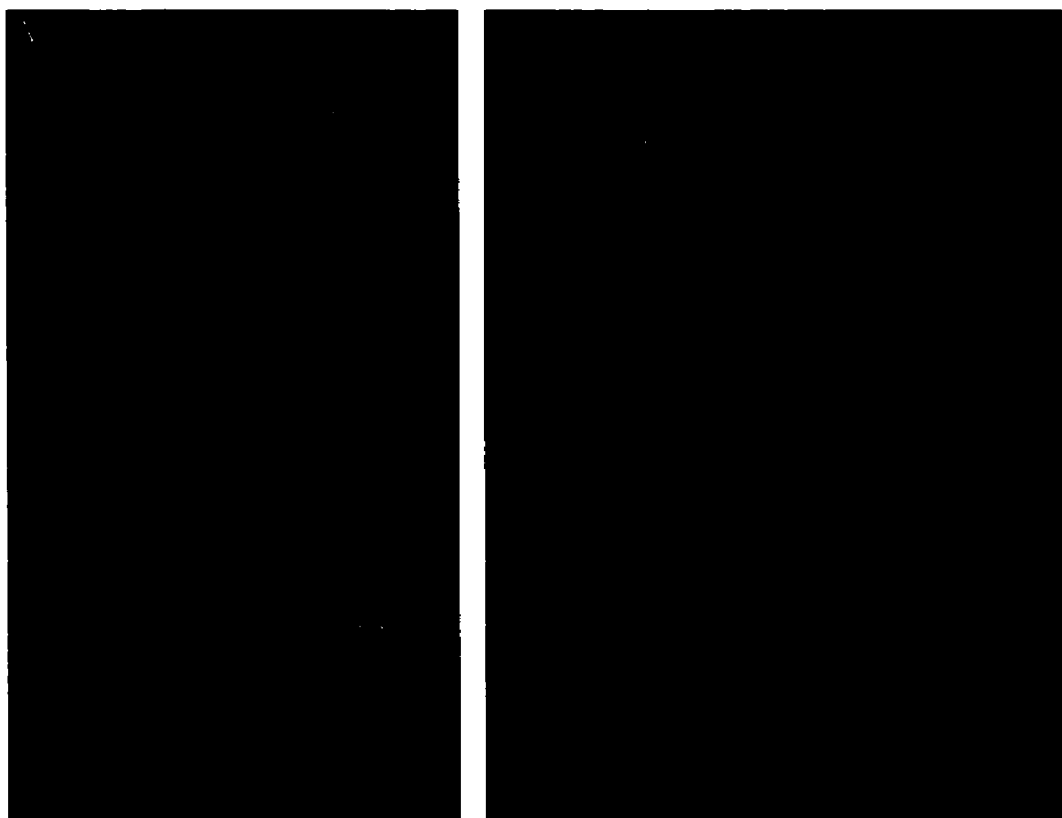


Figure #23 $2|F_o|-|F_c|$ electron density map of the 2.35 Å structure of Eth-CaM contoured around the biosynthetically incorporated Eth residues of the C-terminal hydrophobic cleft. Panel A contains Eth 109 and 124. Panel B contains Eth 144 and 145.

value show a good progression to acceptable values at 0.226 and 0.310 respectively. As with CT-CaM the residue at position 4 of Eth-CaM was removed due to poor electron density. Solvent molecules were incorporated and subsequent refinements led to a final model of Eth-CaM consisting of residues 5 – 147, four calcium ions and 30 water molecules. The final R-value was 0.217 and the final free R-value was 0.305. The bond distances and bond angles had a rms deviation from ideal values of 0.0059 and 1.08 respectively. The position of the four N-terminal and four C-terminal Eth residues are shown in Figure 22 and Figure 23 respectively. Clear well defined electron density was observed for the eight Eth side chains (with Eth 76 omitted).

5.6.3 Dfm-CaM

Although the data obtained from the crystal of Dfm-CaM processed with no apparent problem, attempts to solve this structure have been unsuccessful. The molecular replacement solutions obtained were not clear and yielded maps with significant main chain breaks and R-values greater than 0.5. Therefore, at present the structure of Dfm-CaM remains unsolved.

5.7 Structural Analysis

5.7.1 General structural comparison

The structures of two mammalian CaM proteins, CT-CaM and Eth-CaM, have been determined to 2.5 Å and 2.35 Å resolution. CT-CaM contains site-specific mutations of the four C-terminal Met residues to Leu. Eth-CaM contains the biosynthetically incorporated unnatural amino acid Eth in place of all nine Met residues.

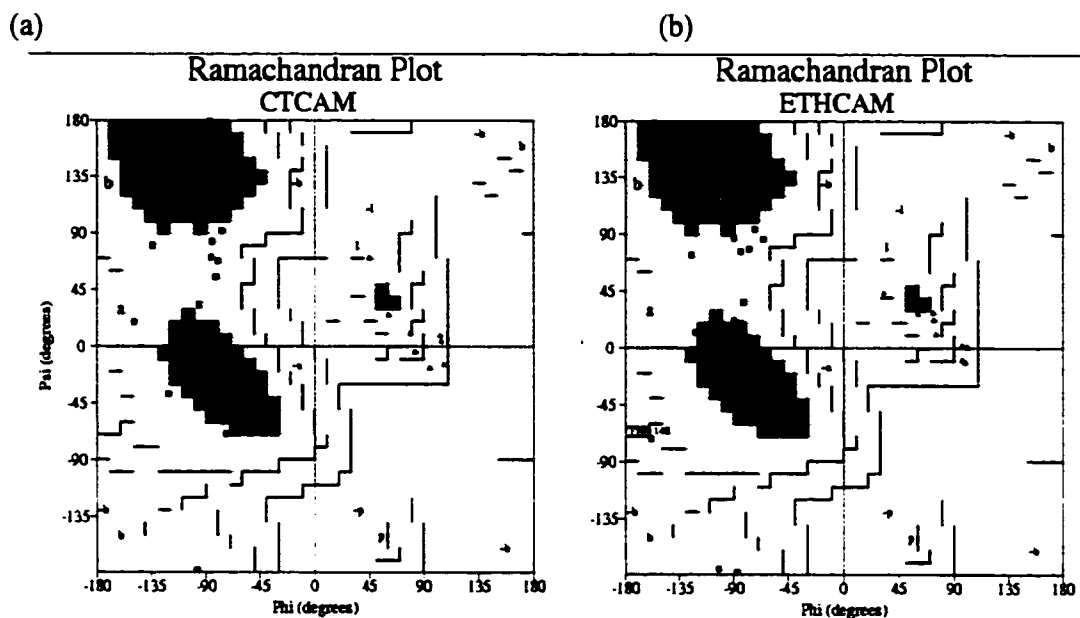


Figure #24 Ramachandran plots of CT-CaM and Eth-CaM. (a) For CT-CaM 91.3% of non-gly residues are in the most favored regions, 8.7% are in allowed regions, and 0% are in disallowed regions. (b) For Eth-CaM 90.6% of non-gly residues are in the most favored regions, 8.6% are in allowed regions, 0.8% are in generously allowed regions, and 0% are in disallowed regions. Non-gly residues are represented by squares and gly residues are represented by triangles).

The overall structures obtained for CT-CaM and Eth-CaM appear very similar to previously reported mammalian CaM structures (112, 118, 145, 146, 147). Both structures demonstrate a dumb-bell shaped protein containing two globular domains separated by a long central helix and are approximately $65 \times 30 \times 30 \text{ \AA}$ in size. A Ramachandran plot of CT-CaM, and Eth-CaM (Figure 24) shows the satisfactory location of all residues into allowed regions of conformational space.

As with wild type mammalian CaM, these structures display a large alpha-helical content. There are seven helical regions in CT-CaM and Eth-CaM which correspond to residues 7-19, 32-39, 45-55, 68- 92, 105-112, 118-128, and 141-146. Comparison of the 1.7 \AA model of wild type mammalian CaM with CT-CaM and Eth-CaM reveals a rms deviation in the alpha-carbon backbone of 0.324 \AA and 0.314 \AA respectively. Figure 25 and 26 show an overlay of the alpha-carbon backbones of CT-CaM and Eth-CaM with wild type CaM demonstrating the good alignment. The secondary and tertiary structure of CT-CaM and Eth-CaM are essentially identical to that of wild type CaM as seen in the ribbon diagrams in figure 27. These structures also contain the four EF hand calcium-binding motifs corresponding to residues 20-31, 56-67, 93-104, and 129-140. The bound calcium ions are hepta-coordinated by five side chain oxygens, one carbonyl oxygen and one water molecule. The coordination distances of the four calcium ions in each of CT-CaM and Eth-CaM are similar other mammalian CaM structures. The similarities in the structures indicate the mutation of the four C-terminal Met residues to Leu in CT-CaM, and the incorporation of the unnatural amino acid Eth for all nine Met residues in Eth-CaM, have no large scale structural consequences. This enables both CT-CaM and Eth-CaM to fold into structures very similar to those of wild type CaM.

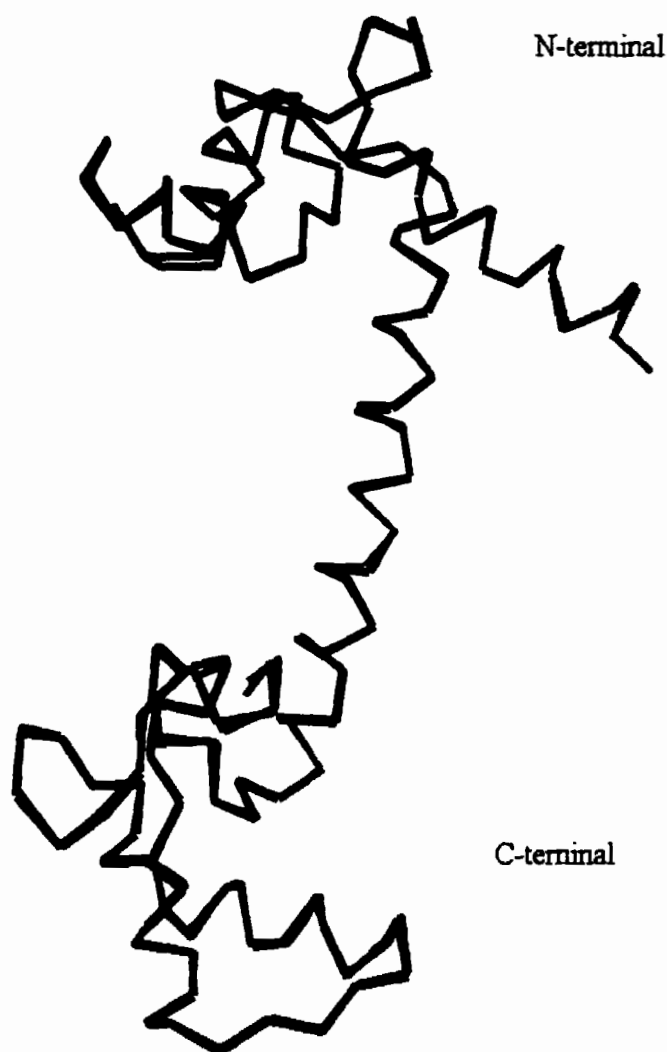


Figure #25 Overlay of the alpha-carbon backbones of CT-CaM and wild type mammalian CaM. The CT-CaM alpha-carbon backbone is depicted in red and the wild type CaM backbone is depicted in yellow. These two structures display an rms deviation of only 0.324 Å.

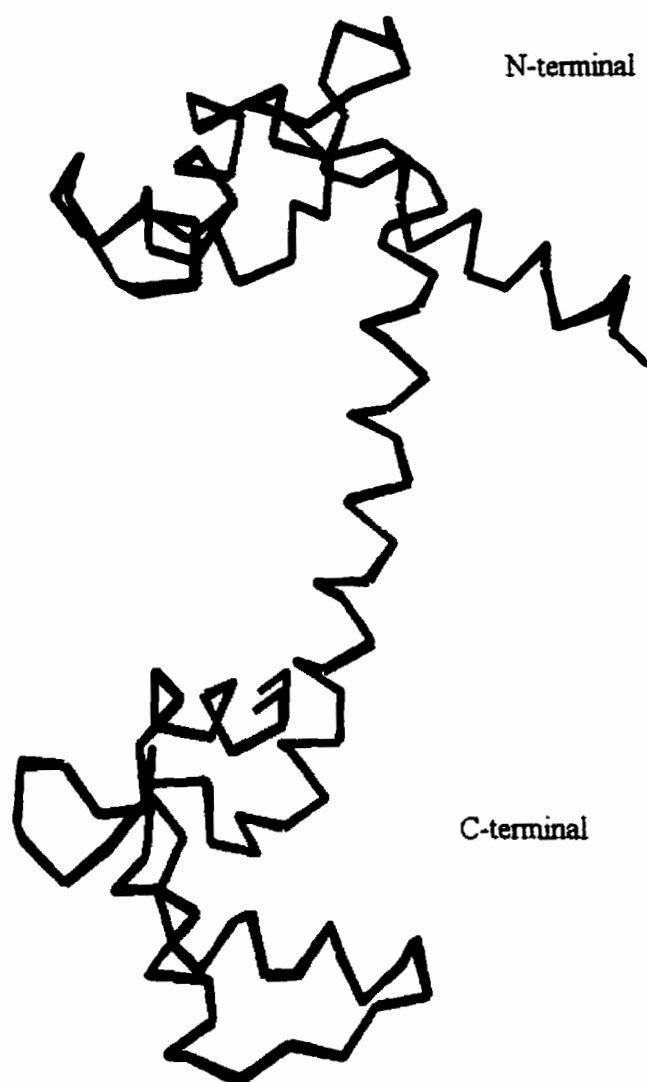


Figure #26 Overlay of the alpha-carbon backbones of Eth-CaM and wild type mammalian CaM. The Eth-CaM alpha-carbon backbone is depicted in blue and the wild type CaM backbone is depicted in yellow. These two structures display an rms deviation of only 0.314 Å.

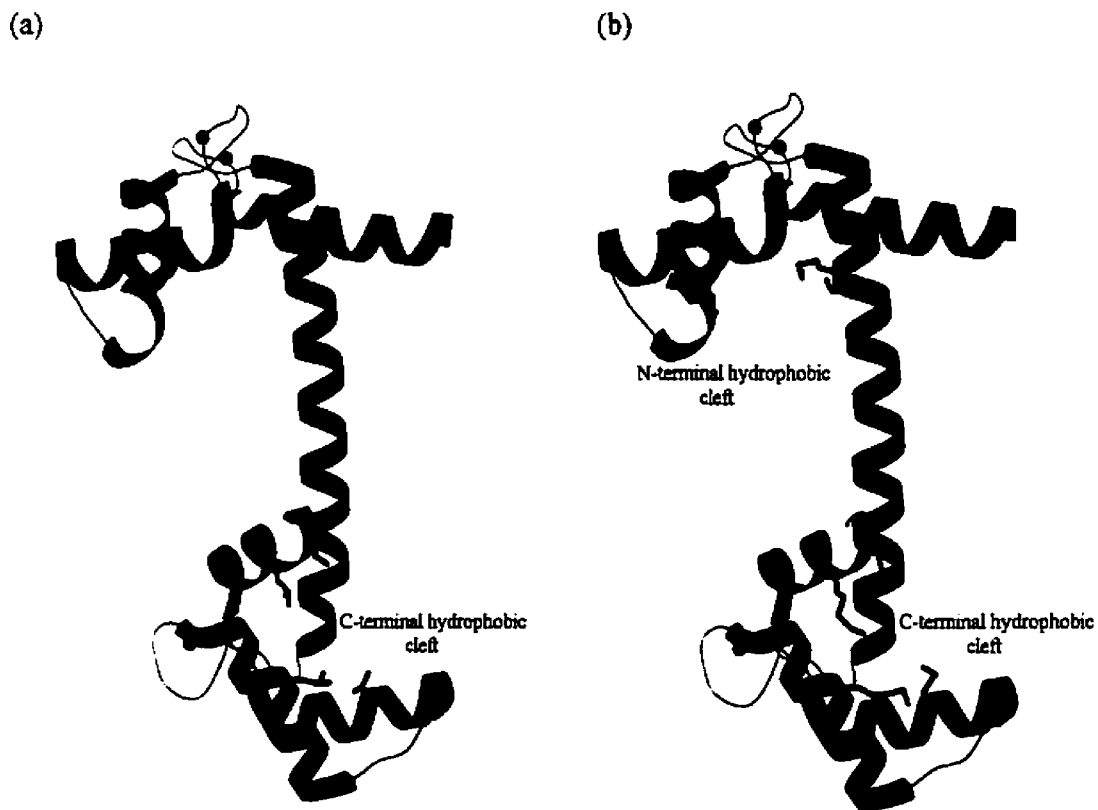


Figure #27 Ribbon diagrams of CT-CaM and Eth-CaM. Panel (a) shows CT-CaM in red with the mutated C-terminal Leu residues in green. Panel (b) shows Eth-CaM in blue with the N-terminal and C-terminal biosynthetically incorporated Eth residues in green.

5.7.2 Examination of Leu mutations in CT-CaM

The mutations of Met 109, 124, 144 and 145 to Leu in CT-CaM resulted in only minor changes to the overall structure of the C-terminal hydrophobic pocket. Comparison of the alpha-carbon backbone of CT-CaM and wild type mammalian CaM over residues 109 to 145 indicates a rms deviation of 0.288 Å. An overlay of residues involved in the formation of the C-terminal hydrophobic clefts of CT-CaM and wild type CaM is shown in Figure 28. This indicates the residues in the C-terminal hydrophobic pocket are in good alignment. The four Leu mutations are seen to occupy approximately the same space in the hydrophobic pocket as wild type CaM with some interesting differences. Figure 29a displays a closer look at the differences between Met and Leu 109 of wild type and CT-CaM. In wild type CaM, Met 109 extends approximately 4.45 Å into the hydrophobic cleft where in CT-CaM, the Leu residue in the same position only extends 3.88 Å across the cleft. In addition, Leu 109 in CT-CaM sits deeper in the hydrophobic cleft as the CE of Met is positioned approximately 2.4 Å above the CD1 of Leu. In Figure 29b, the positions of Met and Leu 124 can be seen. Met 124 of wild type CaM protrudes approximately 4.29 Å into the cleft where as Leu 124 of CT-CaM extends about 3.91 Å. The position of the sulfur atom in Met and the CD1 in Leu are similar, with the CE of Met extending further into the cleft. For the most part, these two residues occupy the same space in the hydrophobic cleft with Met only slightly extended. The same can generally be said for the residues at position 144 (Figure 29c). However at this position, the Leu residue of CT-CaM unexpectedly protrudes further into the hydrophobic cleft than the Met of wild type CaM at 3.89 and 3.11 Å respectively. This is seen as Met 144 bends along the side of the hydrophobic cleft instead of across. The CE of Met is

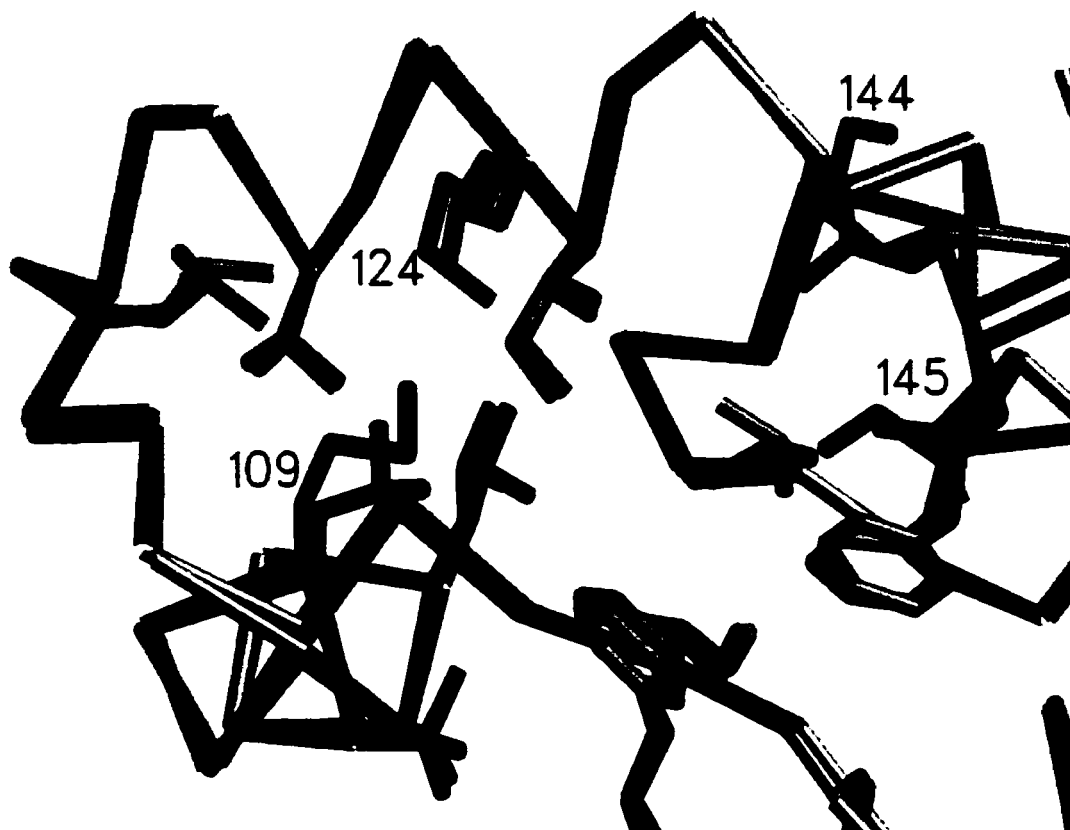


Figure #28 Overlay of the alpha-carbon backbones from the C-terminal hydrophobic cleft of CT-CaM and wild type mammalian CaM. The side chains involved in the formation of the hydrophobic clefts of CT-CaM and wild type CaM are displayed. The Met residues of wild type CaM are shown in green, the mutated Leu residues of CT-CaM are shown in red. Phe 92, Ile 100, Leu 105, Val 108, Leu 116, Val 121, Ile 125, Ala 128, Val 136, and Phe141 from both structures are shown in yellow.

positioned over 2 Å to the side of the closest CD atom of Leu. Finally, Figure 29d displays the location of residue 145 in wild type and CT-CaM. The Leu in CT-CaM is positioned such that it protrudes into the hydrophobic cleft only 3.12 Å where the Met of wild type is seen to extend out to 4.1 Å. The residues occupy similar space in the cleft with the SD of Met and CD2 of Leu positioned together. The only major difference is the extension of the Met CE 1 Å further into the hydrophobic cleft. The positions of these residues underscore the conformational flexibility of Met side chains. While the Leu side chains in CT-CaM are all positioned in a similar fashion, protruding from 3.12 to 3.91 Å into the hydrophobic cleft, the Met residues in wild type CaM display a variety of conformations that extend from 3.1 to 4.45 Å across and around the cleft.

Despite the mutation of four residues in CT-CaM, examination of the hydrophobic cleft revealed little change in the positions of neighboring side chains. Therefore, any effect resulting from the alterations in CT-CaM can be solely attributed to the replacement of Met. As a result of these mutations, the surface of the C-terminal hydrophobic cleft of CT-CaM possesses a wider opening. The distance across the cleft between the opposing side chains of residues 109 and 145 is approximately 1 Å greater in CT-CaM and the distance between 124 and 144 are comparable in CT-CaM and wild type CaM. In addition to this, the distance between residues 109 and 124 along the side of the hydrophobic cleft in CT-CaM is approximately 1.5 Å less than in wild type. This indicates that target proteins entering the C-terminal hydrophobic pocket in CT-CaM encounter a binding site with an opening approximately 1 Å wider across and 1.5 Å narrower on one side.

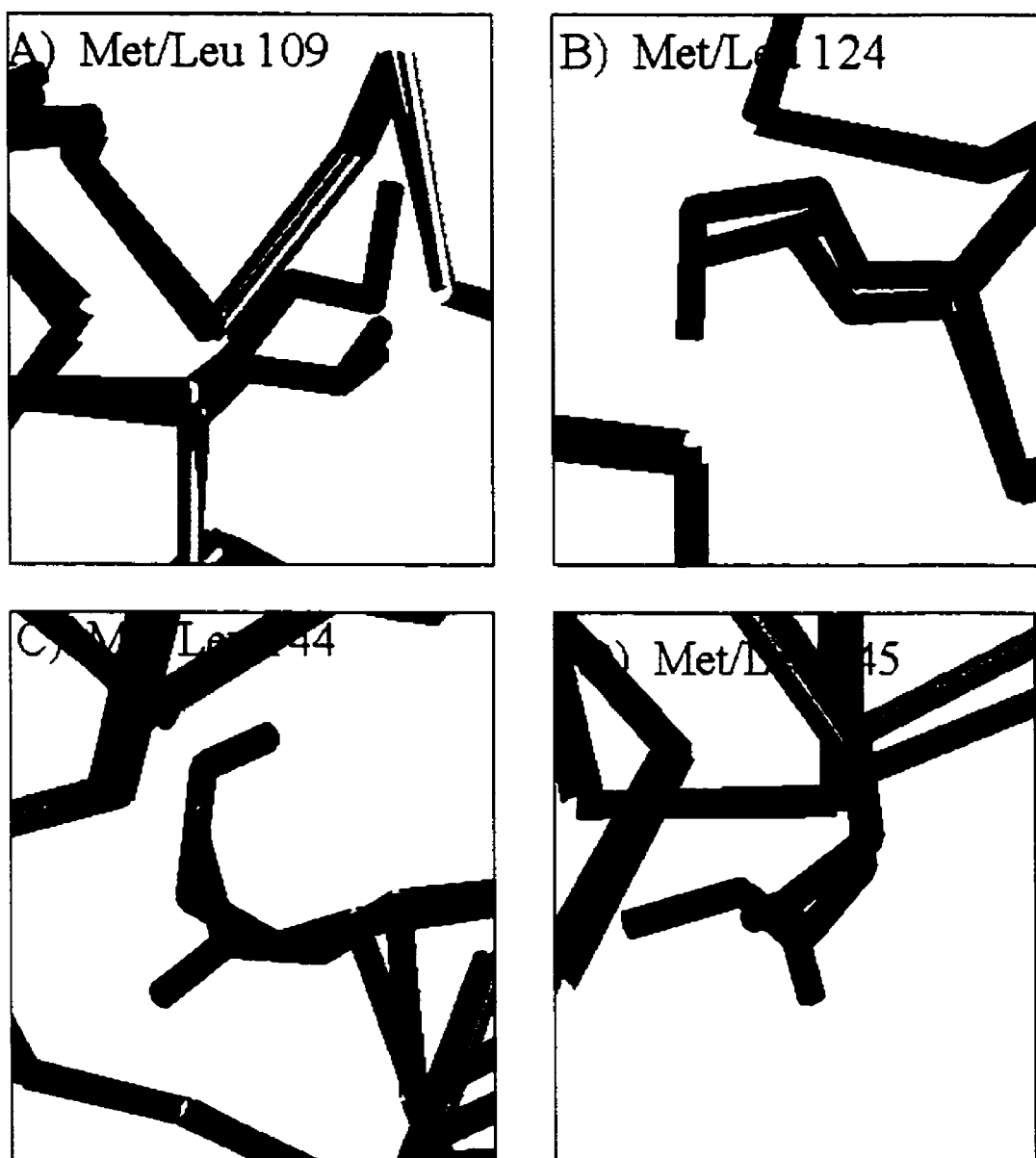


Figure #29 Overlay of the mutated Leu residues in the C-terminal hydrophobic cleft of CT-CaM with the Met residues of wild type mammalian CaM. The position of the mutated Leu side chains are shown in red and the position of the wild type Met side chains are shown in green. Panel A) contains the Met/Leu 109 residues. Panel B) contains the Met/Leu 124 residues. Panel C) contains the Met/Leu 144 residues. Panel D) contains the Met/Leu 145 residues.

5.7.3 Examination of ethionine substitutions in Eth-CaM

The incorporation of the unnatural amino acid ethionine at positions 36, 51, 71, 72, 76, 109, 124, 144 and 145 in Eth-CaM also resulted in only minor changes in the hydrophobic pockets. Eth 36, 51, 71 and 72 occupy the N-terminal hydrophobic pocket, Eth 109, 124, 144 and 145 occupy the C-terminal hydrophobic pocket and Eth 76 is found in the central linker region. Comparison of the alpha-carbon backbone of Eth-CaM and wild type mammalian CaM in the N-terminal hydrophobic pocket (residues 36 to 72) indicated a rms deviation of 0.248 Å. An overlay of residues involved in the N-terminal hydrophobic pockets of Eth-CaM and wild type CaM shows good side chain alignment (Figure 30). The Eth side chains are seen to occupy approximately the same space in the N-terminal hydrophobic cleft as the Met side chains of wild type CaM.

A closer look at the Eth side chains in the N-terminal hydrophobic cleft is shown in Figure 31. Examination of the position of Eth 36 in Eth-CaM reveals a placement virtually identical to that seen for Met 36 of wild type CaM (Figure 31a). The side chains are positioned essentially the same up to the CE atom. At this point, the extra methyl group of Eth extends out. It can be seen that Eth 36 of Eth-CaM protrudes 5.45 Å into the hydrophobic cleft where Met 36 in wild type CaM extends out 4.36 Å. The positions of the sulfur groups are very similar with a difference of only 0.24 Å. Therefore, the two side chains essentially occupy identical space in the hydrophobic cleft. In Figure 31b, the positions of Eth 51 in Eth-CaM and Met 51 in wild type CaM are compared. Again, these two side chains are located in similar positions up to the CE atoms. The Eth side chain bends away from the cleft and as a result protrudes only 5.16 Å across the opening. A similar distance is seen for Met 51 in wild type CaM at 5.26 Å. The two sulfur groups

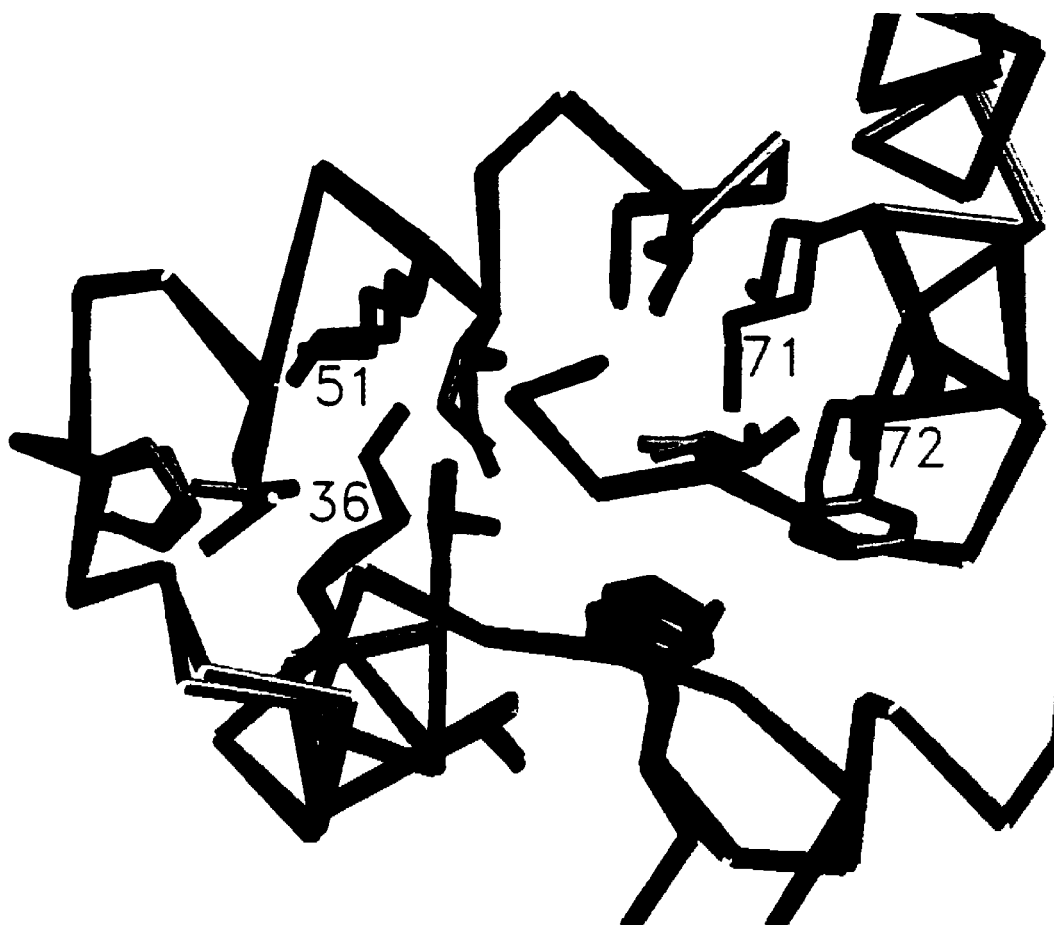


Figure #30 Overlay of the alpha-carbon backbones from the N-terminal hydrophobic cleft of Eth-CaM and wild type mammalian CaM. The side chains involved in the formation of the hydrophobic clefts of Eth-CaM and wild type CaM are displayed. The Met residues of wild type CaM are shown in green, the biosynthetically incorporated Eth residues of Eth-CaM are shown in blue. Phe 19, Ile 27, Leu 32, Val 35, Pro 43, Leu 48, Ile 52, Val 55, Ile 63, Phe 68 from both structures are shown in yellow.

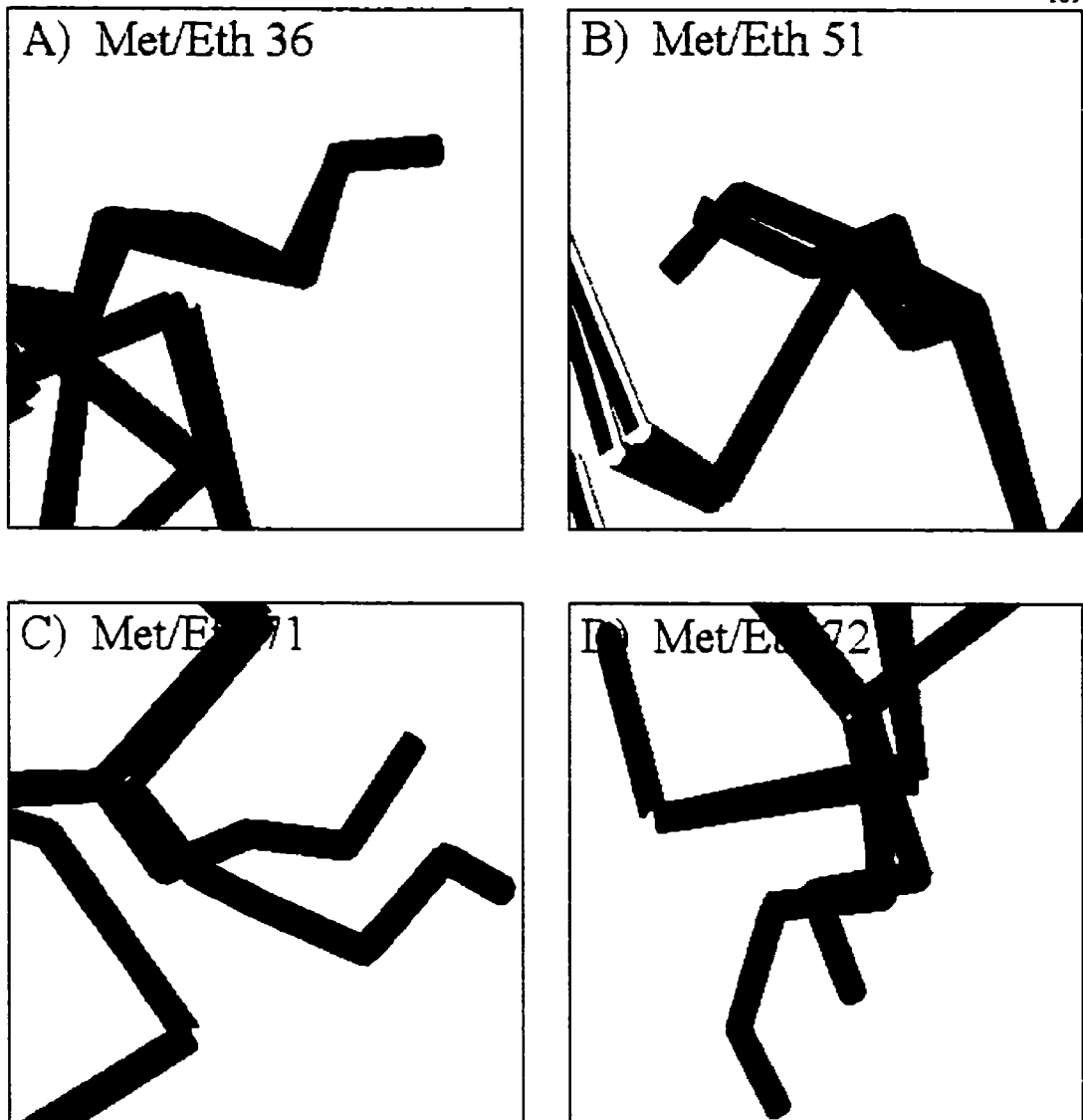


Figure #31 Overlay of the biosynthetically incorporated Eth residues in the N-terminal hydrophobic cleft of Eth-CaM with the Met residues of wild type mammalian CaM. The positions of the Eth side chains are shown in blue and the positions of the wild type Met side chains are shown in green. Panel A) contains the Met/Eth 36 residues. Panel B) contains the Met/Eth 51 residues. Panel C) contains the Met/Eth 71 residues. Panel D) contains the Met/Eth72 residues.

are located only 0.7 Å apart as the Eth side chain is positioned slightly higher at the surface of the hydrophobic cleft. The biggest difference in the N-terminal hydrophobic cleft is seen at residue 71. Figure 31c displays the positions of the Eth and Met side chains. Although they have a similar shape, the Eth side chain sits deeper in the hydrophobic pocket. The Eth side chain also projects further across the surface than Met at 5.84 Å versus 4.59 Å. The sulfur atom of Eth is positioned 1.4 Å below and further into the cleft than the Met of wild type CaM. The final residue of interest in the N-terminal hydrophobic cleft is Eth 72. The position of this residue and Met 72 of wild type CaM is seen in Figure 31d. These side chains follow each other closely up to the sulfur atoms which are located only 0.69 Å apart. The Eth side chain projects 5.79 Å into the hydrophobic cleft while the Met only extends 4.61 Å across. Eth 72 in Eth-CaM sits slightly lower in the pocket than Met but the CE and CZ atoms bend away from the cleft in a similar fashion to the CE of wild type Met.

The incorporation of Eth residues in the N-terminal hydrophobic cleft of Eth-CaM resulted in no significant alterations in the position of neighboring residues. In Eth-CaM, the opening of the hydrophobic cleft is narrowed by approximately 1 Å between residues 36 and 72 and between 51 and 71. In addition, there is less space between these residues at the side of the cleft. Target proteins entering the N-terminal hydrophobic cleft encounter a binding site with a smaller opening, and with the exception of Eth 71 the sulfur positions are essentially unchanged.

Comparison of the alpha-carbon backbone of the C-terminal hydrophobic pocket in Eth-CaM and wild type CaM (residues 109 to 145) reveals a rms deviation of 0.196 Å. Although the backbones match very closely, an overlay of the C-terminal hydrophobic



Figure #32 Overlay of the alpha-carbon backbones from the C-terminal hydrophobic cleft of Eth-CaM and wild type mammalian CaM. The side chains involved in the formation of the hydrophobic clefts of Eth-CaM and wild type CaM are displayed. The Met residues of wild type CaM are shown in green, the biosynthetically incorporated Eth residues of Eth-CaM are shown in blue. Phe 92, Ile 100, Leu 105, Val 108, Leu 116, Val 121, Ile 125, Ala 128, Val 136, and Phe141 from both structures are shown in yellow.

clefts indicate some differences in side chain alignment. (Figure 32). In general, the Eth side chains occupy approximately the same space as the Met side chains with the notable exception of Eth 144. A closer look at residue 109 is displayed in Figure 33a. The position of the Eth side chain closely follows that of Met in wild type CaM up to the CE atom. After this, the CZ of Eth projects into the hydrophobic cleft. Eth 109 protrudes approximately 5.6 Å across the cleft where the Met of wild type CaM extends 4.45 Å. The position of the sulfur atoms at residue 109 are well conserved with a difference of only 0.7 Å. Therefore, this side chain is positioned similar to the Met 109 of wild type CaM. In Figure 33b, the side chain of Eth124 appears to closely match the position of Met in wild type CaM up to the sulfur atom. The sulfur atoms are positioned only 0.19 Å apart in the hydrophobic cleft. However, after this point the CE and CZ atoms of Eth are rotated in the opposite direction of the CE atom of Met. Eth 124 projects approximately 5.15 Å into the hydrophobic cleft compared to 4.29 Å for Met in wild type CaM. The rotation of the CE and CZ atoms of Eth 124 may result from the position of Eth 144. This residue seen in Figure 33c, is extended across the hydrophobic cleft likely inducing a rotation of the CE atom of Eth 124. The location of Eth144 is the most notable variation in the structure of Eth-CaM. This side chain protrudes 6.2 Å directly into the hydrophobic cleft compared to only 3.11 Å by Met, since its side chain is rotated along the side of the cleft. This results in a difference in sulfur atom position of approximately 4.9 Å from wild type CaM. This residue represents the most significant deviation in Eth side chain position when compared to Met of wild type CaM. The final residue in the C-terminal cleft of Eth-CaM to mention is Eth 145. In Figure 33d, the position of this side chain appears similar to that of Met 145. Both display a curvature limiting the extent to which the residues

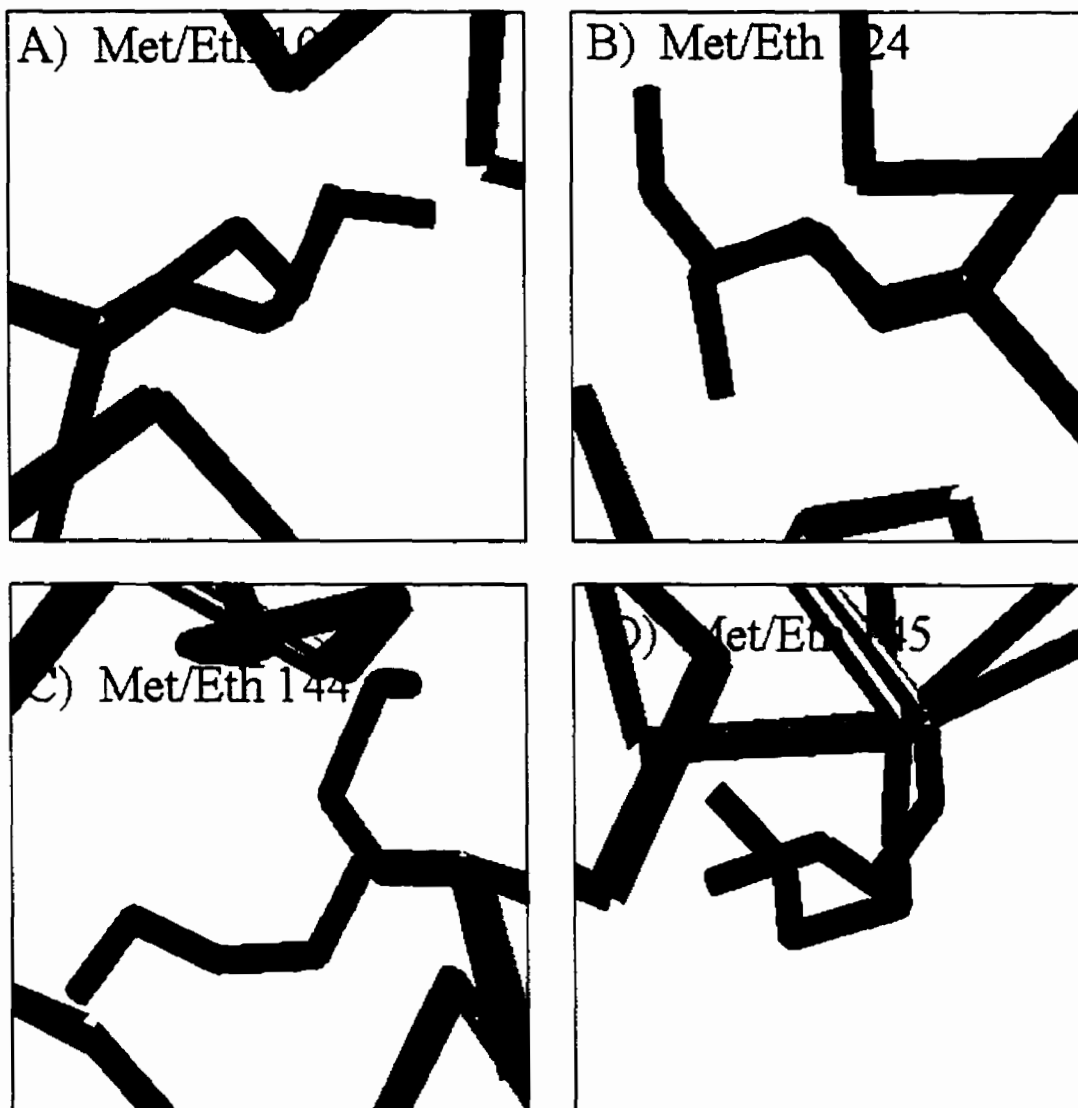


Figure #33 Overlay of the biosynthetically incorporated Eth residues in the C-terminal hydrophobic cleft of Eth-CaM with the Met residues of wild type mammalian CaM. The positions of the Eth side chains are shown in blue and the positions of the wild type Met side chains are shown in green. Panel A) contains the Met/Eth 109 residues. Panel B) contains the Met/Eth 124 residues. Panel C) contains Met/Eth 144 residues. Panel D) contains the Met/Eth 145 residues.

protrude into the hydrophobic cleft. Eth 145 extends 3.9 Å into the cleft compared to 4.1 Å for Met of wild type CaM. Despite a difference of 1.46 Å in the placement of the sulfur atoms, the side chains generally occupy the same space in the C-terminal hydrophobic pocket.

As in the N-terminal domain, the incorporation of Eth result in no significant changes in the surrounding residues of the C-terminal hydrophobic cleft. The distance across the cleft between the opposing residues 109 and 145 and between 124 and 144 is reduced in Eth-CaM by roughly 1.5 Å in both cases. However, when the distance across the cleft between the side chains of Eth 109 and 144 is compared to wild type CaM, there is a reduction of 4.5 Å in the opening of the hydrophobic pocket. In addition, the distances at the sides of the cleft between Eth 109 and 124 and between 144 and 145 are approximately 4.1 Å compared to about 5.1 Å in wild type CaM. Therefore, target protein entering the C-terminal hydrophobic cleft encounters a binding site with an opening approximately 4.5 Å smaller across and 1 Å narrower on the sides. Despite this difference, the sulfur group positions, with the exception of Eth 144, are similar to those in wild type CaM.

6. DISCUSSION: *AEROMONAS SALMONICIDA* SURFACE ARRAY PROTEIN

In attempting to solve the structure of a new macromolecule by X-ray crystallography, the formation of the protein crystal is often the most difficult step. Although it can be expected that most pure soluble proteins will crystallize under some set of conditions, the correct combination of reagents is often hard to find. In many cases, there is little apparent connection between the degree to which proper protocols are observed and the degree to which satisfactory results are obtained (148). Crystallization generally consists of the three seemingly simple steps of nucleation, growth, and cessation of growth. However, when dealing with biological macromolecules, there are significantly more parameters affecting the success of crystallization than with small molecules. Certain parameters can be manipulated with relative ease, such as concentration of the macromolecule and precipitants, the pH, temperature, volume and geometry of the samples and set-ups. However, the parameters dependent on the particular physiochemical properties of the macromolecule are generally difficult to control. These include its stability in aqueous media, conformational flexibility, and hydrophobicity versus hydrophilicity. All of these factors contribute to the success or failure of a crystallization experiment.

The goal of this project was to produce crystals of the surface layer protein of *A. salmonicida* suitable for a structure determination by X-ray crystallography. To accomplish this, large quantities of highly pure A-protein sample were required. Prior to any crystallization experiment, the purity of the protein sample is a major concern. Although some macromolecules can crystallize from an impure solution, it is generally preferred to use samples of 'crystallography-grade' purity. For A-protein preparations,

the established purification protocol produces a large quantity of highly pure protein.

However, there was a concern as to effect small quantities of LPS and porin contamination would have on subsequent crystallization trials. Since a protein crystal contains a large number of molecules (about 10^{20} for a cubic crystal of 0.1 mm on edge), a few p.p.m. of a contaminant corresponds to a large number of molecules that can interfere with crystal growth (148). The purity in the protein sample generally promotes the production of crystals with good mosaicity, X-ray stability, and quality diffraction.

The modifications made to the purification protocol of A-protein served to maintain the integrity of the outer membrane. This proved to virtually eliminate the presence of LPS and porin in the final sample. Since A-protein could be obtained to approximately 97 to 98 % purity, the sample was deemed to be of acceptable quality to begin crystallization trials. As mentioned in Results, an extensive search for crystallization conditions was initiated using a variety of sparse matrix screens, and grid screens. From this work, a successful lead was obtained and followed through to the production of A-protein crystals. However, the initial success in producing A-protein crystals was not translated into successful diffraction. Despite the fact that these crystals appeared well formed and were of sufficient size, they ultimately failed to diffract. Poorly diffracting crystals have an excessive degree of internal disorder which often results from impurities or heterogeneity in the protein sample.

This result underscores the importance of sample quality in protein crystallography. The concept of purity takes on a peculiar meaning in biological crystallogensis (152). Not only must the macromolecules be isolated from other unwanted molecules, the sample should ideally contain a homogeneous population of

macromolecules. The production of good monocrystals is generally enhanced when using solutions containing well defined entities with identical conformations and physiochemical properties (148). Microheterogeneity in a protein sample can result from a number of sources. These include variations in primary and/or tertiary structure resulting from proteolytic degradation, genetic variation, or improper folding or partial unfolding. In addition, the presence of different charge and aggregation states, oligomerization, flexible domains, and variations in post-translational modifications all result in heterogeneity in the protein sample (148).

For purified solutions of A-protein, two forms of microheterogeneity may have contributed to the poor diffraction of the crystals. Solutions of A-protein have been shown to exhibit isoelectric form variation and multimeric complex formation. The presence of a number of isoelectric forms of A-protein has been reported with pI's ranging from 4.8 to 5.3 (153). The existence of this microheterogeneity was confirmed by isoelectric focusing and 2D gels. These charge differences were previously shown not to result from variable levels of phosphorylation or glycosylation, nor do they result from the presence of multimers of differing size (82). These variations were also not an artefact of the purification protocol as heterogeneous charge species were present on the cell surface (153). Similar isoelectric form heterogeneity has been observed for other S-layer proteins, e.g. those of *Azotobacter* and *Clostridium* species (154). These variations were speculated to result from some enzymatic activity such as an amidohydrolase converting glutamine and asparagine into acidic residues or spontaneous deamidation during the mid log growth phase of these cells (155, 156). The pilin of *Neisseria gonorrhoeae* also displays isoelectric heterogeneity, and was originally determined to be non-glycosylated

by chemical methods (157). However, more sensitive techniques accompanied by the recent crystal structure, demonstrated the presence of O-linked disaccharide (158, 159). Therefore, this raises the possibility of undetected glycosylation of the A-protein that may contribute to sample microheterogeneity.

Reduction of isoelectric form heterogeneity was shown in the case of the Fab fragment of anti HIV gp-41 monoclonal antibody to improve crystal quality and the subsequent diffraction obtained (152). With this in mind, attempts were made to improve the isoform content of the A-protein samples. This involved modifications to the purification protocol with the intention of minimizing exposure to deamidating agents. By growing cells at a lower temperature, harvesting earlier during the growth phase and speeding up the purification steps, significant improvements were obtained. As mentioned in the results, this produced a protein sample containing fewer isoelectric forms. Unfortunately, subsequent attempts to isolate a single isoelectric form species of A-protein by chromatographic means were unsuccessful as they were too closely spaced to be resolved. With the improved sample, A-protein crystals were obtained under identical conditions as used previously. The crystal growth, shape and appearance were all similar to the previous A-protein crystals, however, the resulting X-ray diffraction was improved. Although these crystals still failed to diffract to a high enough resolution to be useful in structure determination, the reduction in isoelectric form heterogeneity appeared to significantly improve crystal quality. The mixture of A-protein molecules containing various net charges likely disrupts some crystal contacts due to their different attractive and repulsive character. Reduction of this heterogeneity surely contributed to the improved diffraction and therefore internal order of the A-protein crystals. It is therefore

likely that further reduction in isoelectric form heterogeneity would result in further improvements in crystal quality and diffraction.

The other source of microheterogeneity in the A-protein sample was the formation of high molecular weight oligomers or aggregates. Previous studies using sedimentation equilibrium and molecular sieving HPLC indicated samples of purified A-protein form complexes ranging in molecular weights to an excess of 500 000 (153). Whether the formation of A-protein multimers results from reconstituted segments of two-dimensional array or simply the formation of nonregular aggregates is not clear. A-protein has significant hydrophobic character that could lead to aggregate formation. In fact, an amorphous precipitate of A-protein is often observed during dialysis and concentration of purified protein. However, recent analysis using dynamic light scattering suggests that a combination of specific oligomerization and nonregular aggregation is occurring.

Dynamic light scattering was used to analyze the A-protein sample used in the previous crystallization trials. This provided insight as to the multimeric state of the protein sample used in the production of the initial A-protein crystals. The detection of multimeric heterogeneity by dynamic light scattering was not unexpected, however, the extent of this heterogeneity was surprising. The inability to obtain stable light intensity counts in the presence of Ca^{2+} was a clear indication of the extreme diversity in the A-protein sample. The presence of small sheets of variable size (flexibility of sheets may lead to an extremely variable hydrodynamic radius) and/or larger nonspecific aggregates may be responsible for the highly heterogeneous light scattering. From this analysis, it would seem very unlikely that crystals of A-protein could be produced from this sample. The fact that crystals were obtained speaks to the unpredictability of crystal growth. In

addition, the extent of multimeric heterogeneity in the protein sample would contribute significantly to the internal disorder of the A-protein crystals and thereby restrict the resolution limit of the diffraction obtained.

Calcium has shown to be involved in the formation and maintenance of the two-dimensional surface arrays found on *A. salmonicida* (77). The presence of Ca^{2+} ions was speculated to enhance array formation by providing a cationic bridge between monomers and/or by inducing a conformational change that facilitates binding between A-protein subunits (75). Support for a Ca^{2+} induced conformational change was obtained through the use of circular dichroism (Vikaryous, Skene, Phipps, unpublished results). Samples of A-protein were prepared with and without Ca^{2+} and the resulting secondary structure determined. This indicated that 10% of the α -helix content is converted to β -sheet in the presence of Ca^{2+} . This extra β -sheet content may be involved in the formation of S-layer arrays. With this in mind, Ca^{2+} free samples of A-protein were obtained and subjected to dynamic light scattering analysis. This was done in the hope of improving the multimeric heterogeneity through the reduction of reconstituted array segments.

The results of the dynamic light scattering analysis of the Ca^{2+} free samples of A-protein indicated improved sample dispersion. In this case, stable measurements were obtained and a hydrodynamic radius determined for the predominant species. An approximate molecular weight of the monodisperse species was obtained by applying a model system to the data obtained. When an extended bilobal model (calmodulin) was used the resulting molecular weight corresponded exactly to that of the A-protein monomer. The same determination made using a globular model suggested the species was an A-protein tetramer. Both results are reasonable given the propensity of A-protein

to form tetramers on the cell surface. However, regardless of which state is the correct one, the presence of a monodisperse species in the Ca^{2+} -free preparations is the key result in attempting crystallization. The removal of Ca^{2+} from the sample clearly limits the ability of A-protein to form multimers in solution, supporting a role for Ca^{2+} in A-protein intermolecular association in the formation of the surface array. From this information, it was reasonable to expect better success from crystallizations conducted in the absence of Ca^{2+} .

Despite the significantly reduced multimeric heterogeneity, the Ca^{2+} -free samples of A-protein still contained some large molecular weight aggregate. The existence of this aggregate was likely a result of hydrophobic interactions between A-protein molecules. This was likely detrimental to the formation of quality A-protein crystals and needed to be eliminated. A-protein is known to possess a significant hydrophobic character. This can present a problem in protein crystallization, as in the case of integral membrane proteins. The use of detergents in crystallization of membrane proteins is required to stabilize the hydrophobic portions of the protein. With A-protein, this was accomplished through the use of the nonionic detergent Tween-20. Subsequent dynamic light scattering of Ca^{2+} -free A-protein in the presence of detergent revealed a monodisperse sample devoid of large molecular weight aggregate. The nonionic detergent is likely acting to associate with hydrophobic regions preventing protein aggregation. At this point, a sample of A-protein free of multimeric heterogeneity could be produced and used for crystallization trials.

In an attempt to determine a model-independent molecular weight of the monodisperse species of A-protein in solution, a Zimm plot analysis was performed over a

range of protein concentrations. Nonlinearity of the resulting plot made calculation of the molecular weight impossible. However, the nonlinearity was due to an interesting feature, namely a dynamic structure change appearing as the concentration of A-protein increased. The particle size distribution analysis indicated as the samples became more concentrated, a second species of larger hydrodynamic radius was produced. This resulted in samples containing two distinct forms of A-protein. The emergence of the second A-protein species appears to occur around a concentration of 2 mg/ml and at the higher protein concentrations tested, the bidispersity remained in the sample. Crystallization using the new sample conditions have only been successful at low A-protein concentrations. In fact, a protein concentration of 2 mg/ml was generally the highest used (note: concentration in hanging drops are diluted two fold). This suggested the bidispersity in A-protein samples of higher concentration may act to inhibit crystal formation. The size of the second species potentially corresponds to an octamer. Therefore, further studies at high concentration may lead to the formation of a single high molecular weight species (monodisperse) suitable for crystallization.

A-protein in the presence of 4 mM EGTA and 1% Tween crystallized under unique conditions distinct from those observed for A-protein in the presence of Ca^{2+} . The resulting diffraction was extremely encouraging with reflections recorded out to 3.5 Å. This was a significant improvement since this level of diffraction could provide enough information to attempt a structure determination. Unfortunately, these crystals were extremely unstable during data collection at room temperature and reliable diffraction has not yet been obtained under cryo-conditions. The A-protein crystals suffered some kind of detrimental rearrangement upon addition of cryoprotectant or freezing of the crystals.

It may be possible to preserve the crystal organization through the use of different cryoprotectants, introducing cryoprotectants in a different manner (eg. gradual incrementing of cryoprotectant concentration), or using a different low temperature protocol. Alternatively, crystals at room temperature could potentially be stabilized by cross-linking (eg. with gluteraldehyde).

In the future, work with A-protein would benefit greatly from the establishment of a successful protein expression system. The gene has been cloned but expression is detrimental to *E. coli* (83). Use of an alternative expression system (eg. one allowing secretion of A-protein) may get around this toxicity. A good expression system would eliminate the need to expose A-protein to the harsh guanidine-HCl extraction. Partial denaturation of A-protein may result in some protein failing to refold correctly thereby introducing additional heterogeneity into the sample. In addition, the ability to genetically modify proteins has in many cases proven to be advantageous in obtaining suitable protein crystals. This may enable the production of a single isoelectric form species of A-protein through the mutation surface-exposed of glutamine and asparagine to their associated acidic residues. Also, this would facilitate the production of A-protein fragments that could be used in crystallization. Presently, A-protein fragments are produced through enzymatic degradation which often leaves imprecise cleavage sites introducing sample heterogeneity. Finally, since the structure of A-protein is unique, this presents a phasing problem when using X-ray crystallography. The ability to produce A-protein with selenomethionine incorporated would provide a means of solving the X-ray data by MAD phasing.

7. DISCUSSION: CALMODULIN

The crystallization and structure determination of mammalian (rat testis) CaM in the presence of calcium was first reported in 1985 by Babu *et al* (112). Since then, a number of CaM proteins from a variety of sources have been successfully crystallized and subsequent structure determinations made. Not surprisingly, the reported crystallization conditions for Ca²⁺-CaM share some common elements. In particular, crystallizations were conducted utilizing solutions containing MPD as the primary precipitant in buffers ranging in pH from 5.0 to 5.8 (112, 118, 145, 146, 147). However, despite these common elements, some variations in the crystallization conditions exist. The concentration of MPD in the reservoir ranged from as low as 17% up to around 55%. In some cases, successful crystallization could only be obtained in the presence of 10 to 15% ethanol. Different crystallization temperatures have also been reported from 21°, 18°, and 4° C. In most cases, once crystals were obtained, the use of macroseeding techniques were required to produce crystals of suitable size for X-ray diffraction studies. In addition, some variations in crystal morphology have been reported in the literature. The CaM crystals reported by Babu *et al* (116) appeared as large plate-like prisms, however, the crystals obtained by Chattopadhyaya *et al* (118) formed needle-like rods.

Crystallization experiments of the four CaM proteins CT-CaM, Eth-CaM, Nle-CaM, and Dfm-CaM, began with the exploration of the previously reported CaM crystallization conditions. This served as an excellent starting point, provided the alterations made in the four CaM proteins did not dramatically affect the proteins ability to crystallize under these conditions. A review of the CaM-CaM interactions within reported crystals revealed a number of key crystal contacts (118). The majority of these

intermolecular contacts were provided by water-mediated interactions between adjacent protein molecules. The largest number of direct protein-protein contacts were between the (0, 0, 1) and the (0, 1, 1) translational neighbors in the crystal. These interactions primarily involved residues 37 to 42 and 114 to 120 of adjacent CaM molecules. None of the nine Met residues in wild type CaM demonstrate any significant intermolecular involvement and provide no important crystal contacts. Therefore, it was reasonable to expect CT-CaM, Eth-CaM, Nle-CaM, and Dfm-CaM to respond favorably to the previously reported crystallization conditions.

During crystallization trials, the CaM proteins clearly preferred conditions closer to those reported by Chattopadhyaya *et al* (118). This entailed including 10 to 15 % ethanol in the reservoir solutions, utilizing a more acidic pH, and conducting the crystallization experiments at 4°C. In addition, the resulting crystal morphology closely resembled that reported by Chattopadhyaya since crystals appear as rod-like needles.

However, while Eth-CaM and Dfm-CaM crystallize readily, CT-CaM and Nle-CaM presented more difficulty. Both CT-CaM and Nle-CaM produced significant amounts of amorphous precipitate during crystallization trials. This was speculated to result from non-specific hydrophobic interactions between the CaM proteins in solution. An explanation for this behavior is suggested by considering the nature of the modifications made to these proteins. In CT-CaM, the four C-terminal Met residues are replaced by the more hydrophobic Leu residues, and in Nle-CaM, all nine Met residues are replaced with Nle. In both cases, the highly polarizable sulfur atoms of Met are removed from their respective hydrophobic patches. In wild type CaM, the sulfur atoms in Met side chains enable interaction with the polar environment. In addition, the sulfur

atoms of Met are believed to stabilize the solvent exposed hydrophobic cleft through partial electrostatic interactions with Phe (92). Therefore, since the stabilizing sulfur atoms are absent from both the N-terminal and C-terminal hydrophobic cleft of Nle-CaM, these surfaces may prefer to interact hydrophobically resulting in amorphous precipitate in the presence of MPD. This could explain the failure of Nle-CaM to form useful crystals under conditions used successfully with other CaM proteins. In addition, this explains the observed reduction in amorphous precipitate and microneedle formation in the presence of the detergent 1-s-Octyl- β -D-thioglucoiside under the normal CaM crystallization conditions. For CT-CaM, the removal of sulfur from the C-terminal hydrophobic cleft may act to partially destabilize this region of the protein. Although crystallizations of CT-CaM were ultimately successful, this could explain the reduced solubility of the protein and the relative difficulty in crystal formation as compared to wild type CaM. The crystallizations of Eth-CaM and Dfm-CaM progressed with little difficulty which may in part be due to the presence of sulfur in the hydrophobic clefts of these proteins. Likely, this enables Eth-CaM and Dfm-CaM to behave similarly to wild type CaM during crystallization.

Macroseeding was necessary to produce crystals large enough for X-ray diffraction studies, as reported previously for other CaM crystals. However, none of the reported CaM crystallizations refer to any difficulty with twinned crystals. The presence of parallel growth twins in the crystallizations of CT-CaM, Eth-CaM, and Dfm-CaM presented considerable difficulty in obtaining useful diffraction data. The reason these crystals grew twinned is not clear. High quality reagents were used during crystallizations, so it is unlikely that they contained impurities that might have induced

twinned growth. The presence of an undetected impurity or microheterogeneity in the protein samples could have resulted in twinned crystals. Whatever the reason, no modification to the setup procedure, or chemistry in the reservoir was able to eliminate twinned growth. Fortunately, the CaM crystals were amenable to delicate manipulation and single crystals could be obtained by physically removing the parallel growth twin.

Data collection of CT-CaM and Eth-CaM revealed unit cell dimensions very close to those reported by Chattopadhyaya *et al.* Although not identical, the cell dimensions are within 0.5 % of each other, and no angle deviates more than 0.8° . As with other CaM crystals, these proteins crystallized in the P1 triclinic space group. Although complete data sets were collected for CT-CaM and Eth-CaM, the crystals demonstrated considerable decay in the X-ray beam at room temperature. Therefore, in further CaM data collections, the use of cryo-crystallography is recommended. This was successfully employed in the data collection performed with Dfm-CaM crystals. The MPD in the mother liquor provides a suitable cryoprotectant and data can be collected at 100° K with no crystal decay.

Data collected from the Dfm-CaM crystals revealed an interesting variation in the unit cell dimensions. The cell parameters are virtually identical to previous CaM crystals with the exception of the a axis. For Dfm-CaM, the length of the a axis was essentially double what is normally expected (58.75 \AA vs 29.95 \AA). This doubled axis was not an artefact of the freezing process since data collected at room temperature gave the same unit cell dimensions. As a result, the unit cell of Dfm-CaM is essentially doubled in volume. This led to the assumption that two molecules of Dfm-CaM were contained within the unit cell. Therefore, it is speculated that one of the two molecules of Dfm-CaM

is in the same orientation as normal CaM crystals, while the other has a similar but different enough orientation, or a small translational shift which breaks the crystal symmetry. A report of this unique crystal form of wild type CaM (*Gallus gallus*) was made in 1996 by Bernhard Rupp (144). Dr. Rupp (personal communication) supported the above explanation, and informed us of the development of a molecular replacement and automated rebuilding protocol with which he recently solved the structure of the unusual crystal form. Hopefully the Dfm-CaM data will be amenable to this protocol and a structure determination can be made in the near future.

The purpose of these studies was to confirm the overall structure of the CaM proteins and to analyze the altered hydrophobic surfaces. This information would ideally shed light on the involvement of Met residues in binding target polypeptides. Met residues have some unique features believed to be important in CaM's interaction with target proteins. The first property of interest is the flexibility of the Met side chain. Met has an unbranched aliphatic side chain that contains a sulfur group. The presence of two S-C bonds in the unbranched side chain confers a large degree of conformational flexibility to Met (132). Since the S-C bonds are somewhat longer than a C-C bond, the energy barrier between the different rotamers relatively is low, and rotation around these bonds is essentially unhindered (92, 132). Secondly, the presence of the sulfur group provides a highly polarizable surface. The electron rich sulfur atom is able to adjust its electron distribution enabling Met to interact with polar and nonpolar environments (92). This is believed to be important in stabilizing the solvent exposed hydrophobic surface of Ca²⁺-CaM. This feature of Met enables a favorable interaction with water in the polar solvent while retaining the ability to interact with hydrophobic side chains of the target

binding domains (92, 104). In addition, the polarizable sulfur group is thought to strengthen the London dispersion forces and stabilize hydrophobic interactions. This is believed to greatly enhance the stickiness of the hydrophobic binding surface and the ability of CaM to interact with target peptide (92, 132). Therefore, in CaM the Met residues provide a malleable surface enabling the hydrophobic patch to adjust and bind to a number of target proteins in a sequence-independent manner.

In CT-CaM, the four C-terminal Met residues were simultaneously mutated to Leu. This amino acid was used as a replacement for Met since it is a hydrophobic residue with a preference to form α -helical structures, and as a result should not perturb the secondary structure (117). Therefore, the C-terminal hydrophobic cleft of CT-CaM contains a surface that is significantly less flexible and less polarizable than wild type CaM. Previous studies investigating functional effects resulting from these mutations in CT-CaM were performed by Zhang and Vogel in 1994 (117). In this case, CT-CaM demonstrated a significantly lower binding affinity (5x) for phosphodiesterase (PDE) but is able to fully activate the enzyme when provided with higher concentrations of CaM (117). In addition to this, CT-CaM displays a lower binding affinity for MLCK and calcineurin (133).

The crystal structure of CT-CaM was determined to 2.5 Å in order to provide a structural explanation for the functional data. The first question to be answered was whether the mutations made to CT-CaM resulted in any large scale structural changes that would adversely affect binding affinity. Since CT-CaM crystallized under similar conditions to those used for wild type CaM, and with similar unit cell parameters, it was likely the overall protein structure would be quite similar. This was confirmed when the

structure of CT-CaM was compared to that of wild type CaM. CT-CaM contained all seven helical regions, and had four calcium ions bound in EF-hand binding motifs. The overall similarity was clearly demonstrated through comparison of the alpha-carbon backbones of CT-CaM and wild type CaM seen in Figure 25, which had an rms deviation of only 0.324 Å. This supports previous NMR studies indicating these C-terminal mutations were well tolerated in CT-CaM (117). Therefore, the overall structure of CT-CaM is essentially unchanged from wild type CaM and should have no detrimental affect on binding target proteins.

Since no large scale structural changes account for the altered binding affinity of CT-CaM, the Leu mutations must have a more direct affect. The structure of the C-terminal hydrophobic pocket as seen in Figure 28 demonstrates no significant change in the alpha-carbon backbone. In addition, the position of residues neighboring the Leu mutations were well conserved. Therefore, the only significant differences in the C-terminal hydrophobic cleft are the mutated Leu residues. Although the Leu side chains occupy similar space in the hydrophobic pocket as the wild type Met side chains, some notable differences were observed. These resulted in a wider opening to the hydrophobic cleft between Leu 109 and 145 and a narrowing between Leu 109 and 124 at the side of the cleft. In order to see what effect the change in dimensions of the C-terminal hydrophobic cleft might have on interactions with substrate peptides, the C-terminal domain of CT-CaM was superimposed on that of the smMLCK-CaM complex. Figure 34 displays an overlay of the C-terminal Met side chains of wild type Ca²⁺-CaM, and smMLCK-CaM, plus the C-terminal Leu side chains of CT-CaM. From this, it is clear that the Met side chains adopt different orientations when bound to the smMLCK peptide

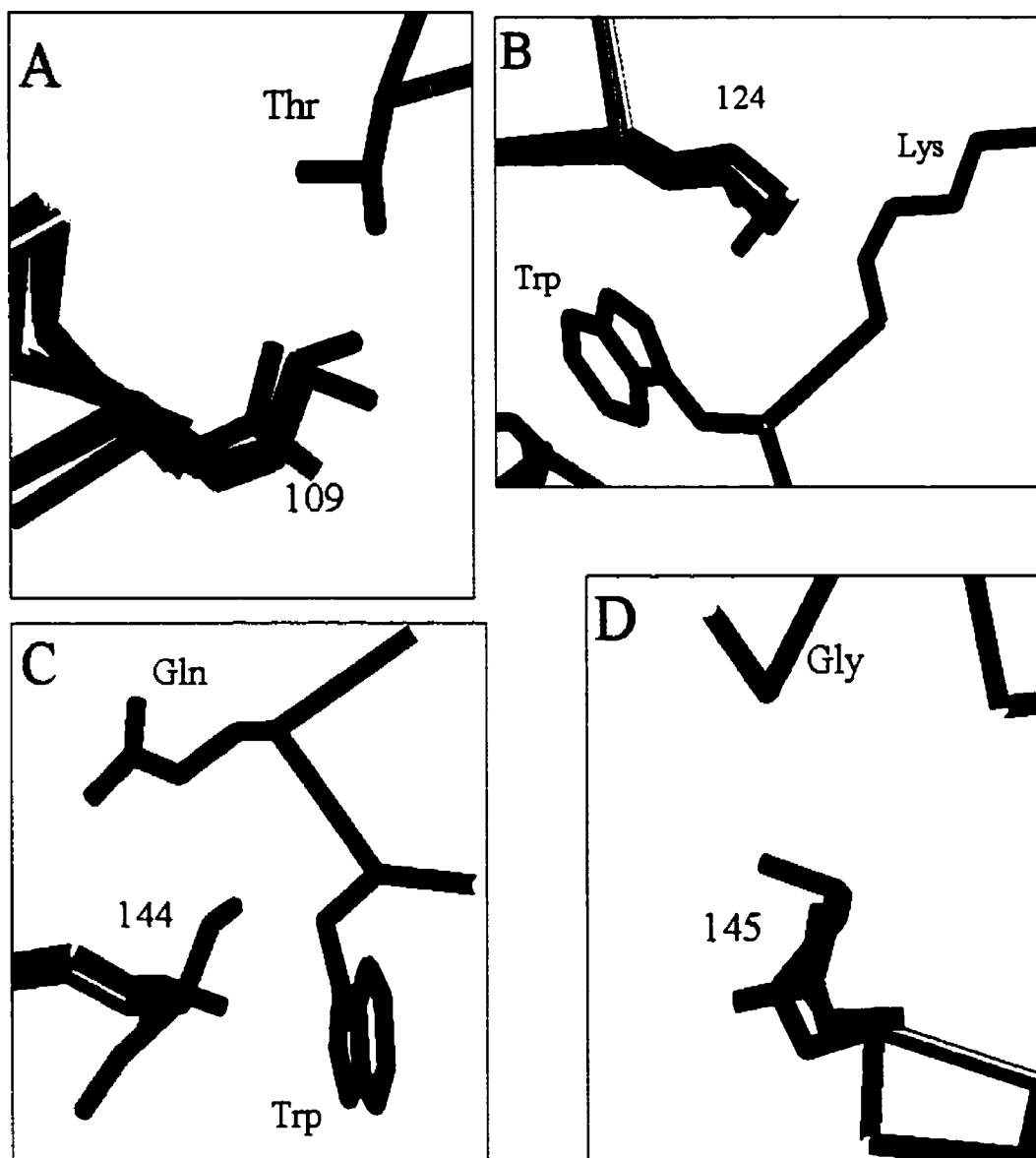


Figure #34 Overlay of the relevant residues in the C-terminal hydrophobic clefts of CT-CaM, wild type-CaM and smMLCK-CaM. CT-CaM is in red, wild type CaM is in green, smMLCK complexed CaM is in yellow, and the MLCK peptide is in purple. Panel A shows the position of Met/Leu 109 with the Thr residue of MLCK. Panel B shows the position of Met/Leu 124 with the Lys and Trp of MLCK. Panel C shows the position of Met/Leu 144 with the Gln and Trp of MLCK. Panel D shows the position of Met/Leu 145 with the Gly residue of MLCK.

than in the wild type structure. In order to make the necessary peptide contacts, residues 124 and 145 display considerable rotation of their SD atoms while the side chain of Met 144 is rotated nearly 180° at the CG atom. Comparison of these positions with the Leu side chains of CT-CaM provides some interesting points. The Leu residues at positions 109, 124 and 145 in CT-CaM appear close enough to provide hydrophobic contacts for the MLCK peptide. However, small differences at these positions do exist which likely add to the reduced binding affinity. A significant difference is seen at Leu 144. In this comparison, Leu 144 would be unable to adapt its position, as seen with Met 144, to maximize contacts along the plane of the Trp rings and sits too far away to make contacts with the Gln of MLCK.

From the crystal structure of CT-CaM, it is clear the results obtained in the functional experiments by Zhang and Vogel (117) are solely due to the presence of Leu in the C-terminal hydrophobic cleft, and not any other structural rearrangement. The reduced binding affinity of CT-CaM is likely due to a combination of reduced side chain length and flexibility that may abolish some important peptide contacts as seen with Leu 144. In addition, since polarizable sulfur atoms are believed to strengthen London dispersion forces, the hydrophobic contacts made by Leu 109, 124, 144 and 145 are likely weaker than in wild type CaM. All these factors would result in a reduced binding affinity in CT-CaM, which is in agreement with Zhang and Vogel (117).

The other CaM protein used to probe the importance of Met side chains was Eth-CaM. Incorporation of the unnatural amino acid Eth was shown to be 87% with the substitutions occurring randomly among the nine Met residues in CaM (131). The Eth residue retains many of the important qualities of Met, in particular the flexible side chain

and the presence of a polarizable sulfur atom. However, the side chain is extended in length with a more bulky ethyl group attached to the sulfur atom. The biological affect of these substitutions was previously investigated by assaying Eth-CaM ability to activate MLCK and calcineurin (131). In this study, the maximal activation level (P_{max}) and the dissociation constant K_D were determined for Eth-CaM in comparison to wild type CaM. When MLCK was assayed, Eth-CaM was shown to produce 90 % of the maximal activation with little change in the K_D value. In the case of calcineurin, Eth-CaM could only achieve 59 % of the maximal activation with a K_D larger than found with wild type CaM (131).

The crystal structure of Eth-CaM was determined to 2.35 Å in order to provide a structural explanation for the functional data. As with CT-CaM, the crystallization experiments suggest the overall structure of CaM was retained in Eth-CaM. Confirmation of this was seen by comparison of the Eth-CaM and wild type CaM structures. The overall similarity in the alpha-carbon backbones of Eth-CaM and wild type CaM is seen in Figure 26, and had an rms deviation of only 0.314 Å. This supports previous FTIR and CD spectroscopy which indicated no substantial secondary structure changes (131). As a result, the overall structure of Eth-CaM should not alter the proteins ability to interact with target proteins.

In order to explain the functional studies, a closer examination of both the N- and C-terminal hydrophobic pockets was required. The N-terminal hydrophobic cleft was structurally similar to that in wild type CaM with little deviation in the alpha-carbon backbone. In addition, the positions of residues neighboring Eth in the cleft were relatively unchanged (Figure 30). Although the dimensions of the cleft were slightly

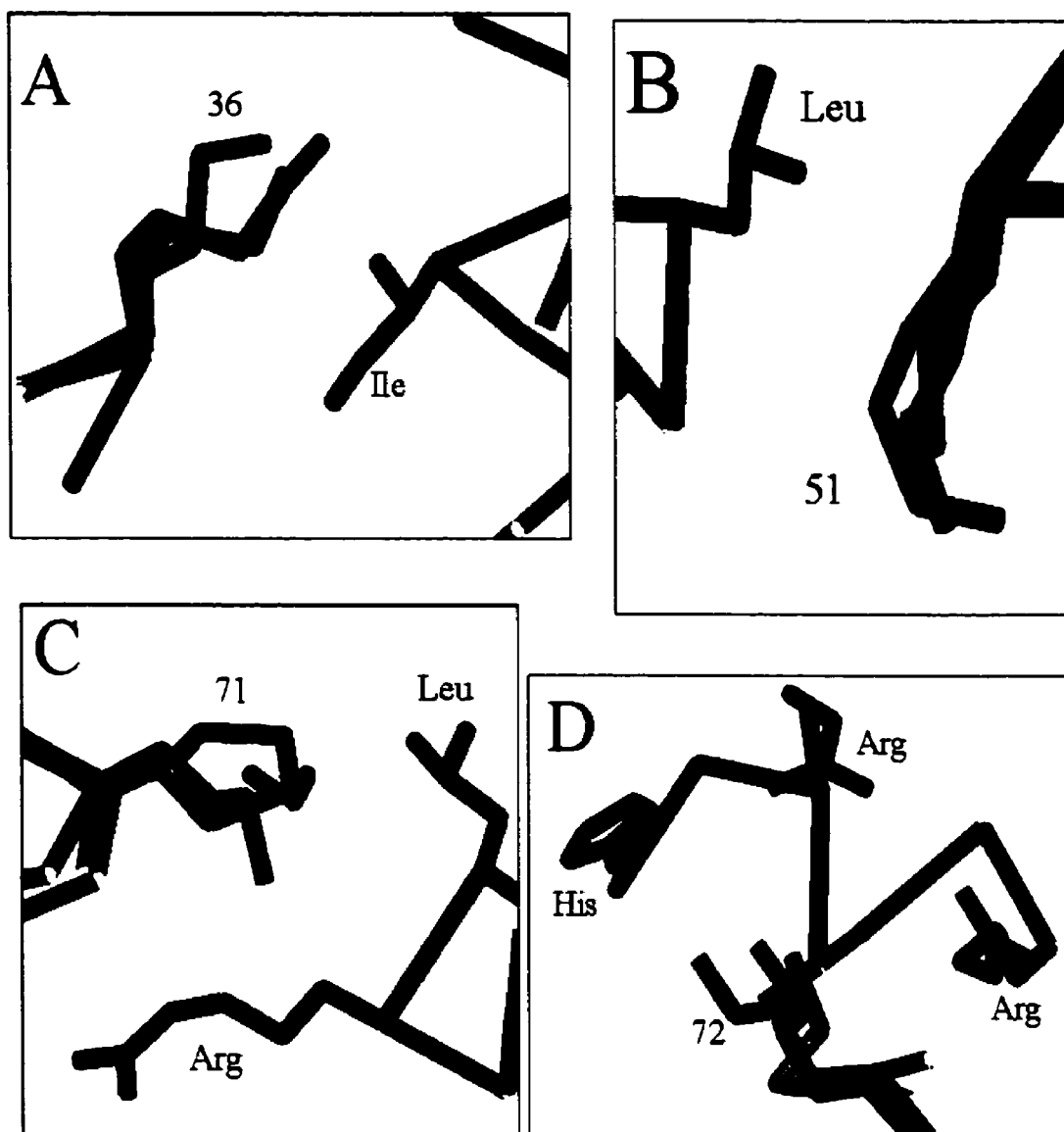


Figure #35 Overlay of the relevant residues in the N-terminal hydrophobic cleft of Eth-CaM, wild type CaM and smMLCK-CaM. Eth-CaM is in blue, wild type CaM is in green, smMLCK complexed CaM is in yellow, and MLCK peptide is in pink. Panel A shows the position of Met/Eth 36 with the Ile of MLCK. Panel B shows the position of Met/Eth 51 with the Leu of MLCK. Panel C shows the position of Met/Eth 71 with Leu and Arg of MLCK. Panel D shows the position of Met/Eth 72 with His, Arg and Arg of MLCK.

more constrained, the Eth side chains occupied similar positions to Met with the sulfur positions (with the exception of Eth 71) well conserved. Therefore, MLCK and calcineurin likely interact with an N-terminal hydrophobic pocket that is similar to that of wild type CaM. An idea of how the Eth side chains interact with the binding domain of MLCK can be obtained by superimposing the N-terminal domain of Eth-CaM with smMLCK-CaM. In Figure 35, the extended length of the Eth side chains would appear to produce no significant steric hindrance in binding of MLCK. The Eth 36 and 51 are positioned to provide hydrophobic contacts with Ile and Leu of MLCK respectively. Eth 71 and 72 appear to be slightly out of position to provide optimal contacts with the Leu, Arg and His, Ala, Arg, Arg of MLCK. However, due to free rotation around the two C-S bonds, Eth is likely able to adapt its position to accommodate these interactions.

In the C-terminal hydrophobic pocket the incorporation of Eth residues appears to make a greater difference. As in the N-terminal domain, the alpha-carbon backbone of the C-terminal hydrophobic cleft of Eth-CaM is virtually identical to wild type CaM, and the neighboring residues exhibit little change in position (Figure 32). However, the position of Eth 144 deviated considerably from the location of Met in wild type-CaM. The position of this residue dramatically narrows the hydrophobic cleft opening and would likely interfere with binding of target protein. In Figure 36, the C-terminal hydrophobic cleft of Eth-CaM is superimposed with smMLCK-CaM. This revealed the Eth side chains 109, 124 and 145 would likely be able to form the required contacts with MLCK, requiring only minor structural rearrangements (Eth 124 requiring more than the others). However, the position of Eth 144 would produce significant steric interference with MLCK and likely calcineurin in the cleft.

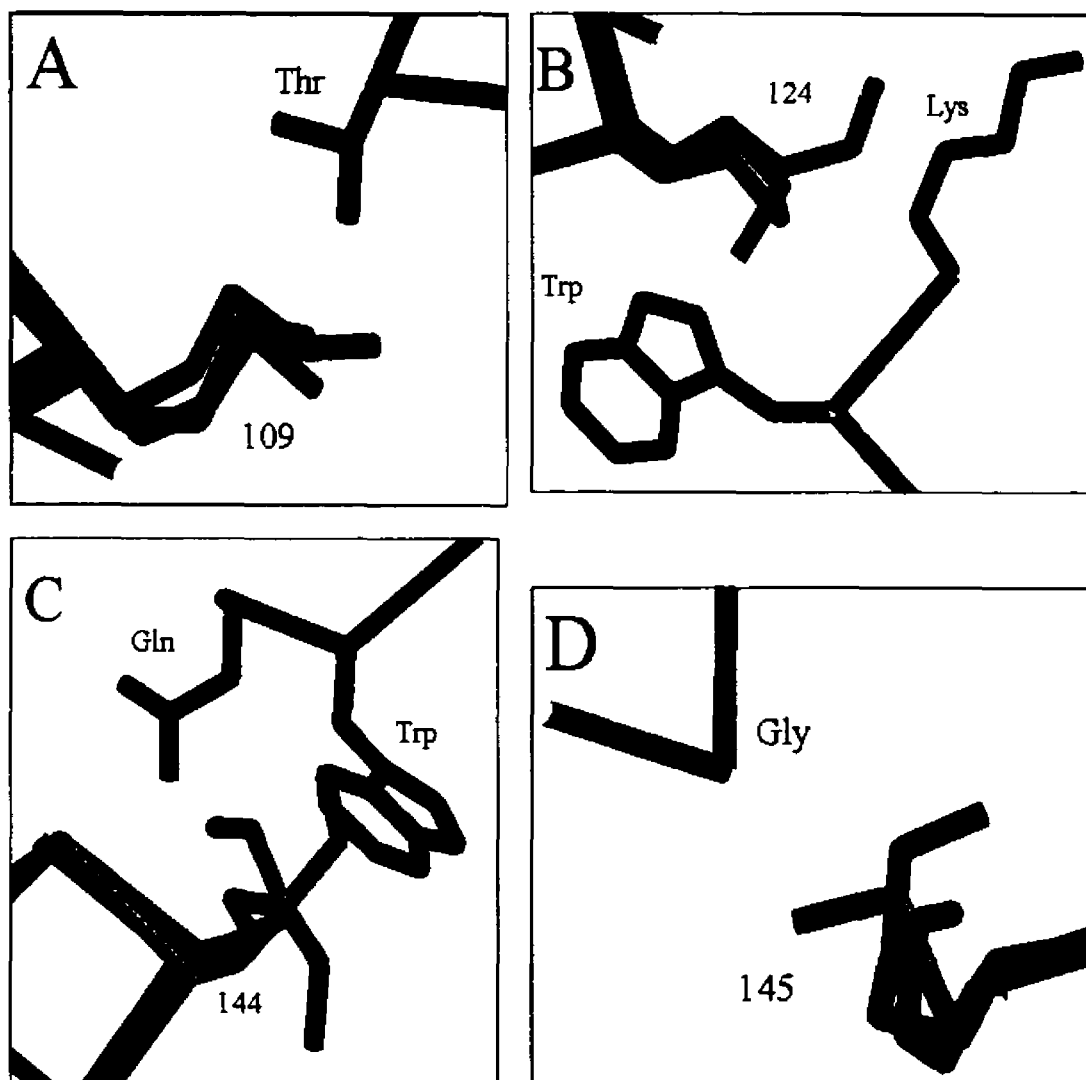


Figure #36 Overlay of the relevant residues in the C-terminal hydrophobic cleft of Eth-CaM, wild type CaM and smMLCK-CaM. Eth-CaM is in blue, wild type CaM is in green, smMLCK complexed CaM is in yellow, and MLCK peptide is in pink. Panel A shows the position of Met/Eth 109 with the Thr of MLCK. Panel B shows the position of Met/Eth 124 with the Lys and Trp of MLCK. Panel C shows the position of Met/Eth 144 with Gln and Trp of MLCK. Panel D shows the position of Met/Eth 145 with Gly of MLCK.

Since the ability of Eth-CaM to activate MLCK and calcineurin is considerably different, the Eth side chains must affect their binding differently. The fact that the N-terminal hydrophobic pocket is affected less by these substitutions suggests it is able to bind target protein in a similar fashion as wild type. However, the C-terminal hydrophobic cleft appears to play the pivotal role. Since Eth-CaM is able to achieve 90% activation of MLCK, this implies the position of Eth 144 is able to adapt and successfully bind this protein, where the same may not be true for calcineurin. The presence of the large hydrophobic Trp residue near Eth 144 of Eth-CaM may induce this side chain to shift position in order to maximize favorable hydrophobic contacts. In doing so, this would relieve the apparent steric hindrance with MLCK. Since no structure of a calcineurin peptide-CaM complex has been reported, the likely interactions with Eth-CaM are unclear. However, it can be speculated that the calcineurin binding domain would bind the N-terminal cleft of Eth-CaM in a similar fashion as wild type (since only minor structural differences exist in this cleft) where the C-terminal hydrophobic cleft may not be able to completely accommodate this protein as a result of interference from Eth 144. Therefore, some key hydrophobic contacts may not be formed reducing activation of calcineurin. The fact that Eth-CaM demonstrates little change in K_D for MLCK and only a small increase (1.6 x) in K_D with calcineurin is reasonable since Eth is likely able to provide enough hydrophobic contacts to bind MLCK and to some extent calcineurin. In addition, the presence of the sulfur atoms may strengthen these interactions and secure bound MLCK and calcineurin.

In summary, the results obtained from the crystal structures of CT-CaM and Eth-CaM indicate the presence of the polarizable sulfur atom in Met acts to increase the

affinity of CaM for its target proteins. Also, the length and flexibility of the Met side chain is a significant factor in providing key hydrophobic contacts in the binding and activation of target proteins. These results support previous investigations into the role of Met residues in the hydrophobic clefts of CaM (117, 131, 133). In the future, the structure of Dfm-CaM to 1.8 Å will be solved which may provide further insight as to the role of Met residues in CaM. In addition, crystallization of Nle-CaM should be pursued to determine conditions that may turn existing microneedles into crystals useful for a structure determination. Finally, crystallization of these CaM proteins in a peptide complex would provide the ultimate answers as to how the altered residues interact with the target protein.

REFERENCES

1. Sleytr, U.B., P. Messner, D. Pum, M. Sara. 1996. Occurrence, location, ultrastructure and morphogenesis of S-layers, p.5-23. *In* Crystalline Bacterial Cell Surface Proteins. Academic Press, R.G. Landes Company, Texas.
2. Sleytr, U.B., P. Messner. 1978. Regular arrays of macromolecules on bacterial cell walls: structure, chemistry, assembly and function. *Int. Rev. Cytol.* **53**:1-64.
3. Houwink, A.L. 1953. A macromolecular monolayer in the cell wall of *Spirillum* spec. *Biochim. Biophys. Acta.* **10**:360-366.
4. Roberts, K., C. Grief, G.J. Hills P.J. Shaw. 1985. Cell wall glycoproteins: structure and function. *J. Cell. Sci.* **2**:105-27.
5. Sleytr, U.B., P. Messner, M. Sara, D. Pum. 1986. Crystalline envelope layers in archaeobacteria. *System. Appl. Microbiol.* **7**:310-13.
6. De Rosa, M., A. Gambacorta, A. Gliozzi. 1986. Structure, biosynthesis, and physicochemical properties of archaeobacterial lipids. *Microbiol. Rev.* **50**:70-80.
7. Ames, G.F. 1986. Bacterial periplasmic transport systems: structure, mechanism, and evolution. *Annu. Rev. Biochem.* **55**:397-425.
8. Luderitz, O., M.A. Freudenberg, C. Galanos, V. Lehmann, E. Rietschel, D. Shaw. 1982. Lipopolysaccharides of gram-negative bacteria, p. 79-151. *In* S. Razin and S. Rottem (ed.), Current topics in membranes and transport, vol. 17. Membrane lipids of prokaryotes. Academic Press, New York.
9. Lugtenberg, B., L. van Alphen. 1983. Molecular architecture and functioning of the outer membrane of *Escherichia coli* and other gram-negative bacteria. *Biochim. Biophys. Acta* **737**:51-115.
10. Hagiya, H., T. Oka, H. Tsuji, K. Takumi. 1992. The S-layer composed of two different protein subunits from *Clostridium difficile* GAI 1152: a simple purification method and characterization. *J. Gen. Appl. Microbiol.* **38**:63-74.
11. Messner, P. 1996. Chemical composition and biosynthesis of S-layers, p. 35-64. *In* Sleytr, U.B., P. Messner, D. Pum, M. Sara (ed.), Crystalline Bacterial Cell Surface Proteins. Academic Press, R. G. Landes Company, Texas.
12. Messner, P., U.B. Sleytr. 1991. Bacterial surface layer glycoproteins. *Glycobiology.* **1**:545-51.

13. Smit, J., D.A. Grano, R.M. Glaeser, N. Agabain. 1981. Periodic surface array in *Caulobacter crescentus*: fine structure and chemical analysis. *J. Bacteriol.* **146**:1135-1150
14. Boot H.J., C. Kolen, J.M. van Noort, P.H. Pouwels. 1993. S-layer protein of *Lactobacillus acidophilus* ATCC 4356: purification, expression in *Escherichia coli*, and nucleotide sequence of the corresponding gene. *J. Bacteriol.* **175**:6089-96.
15. Glenn, A.R. 1979. Production of extracellular proteins by bacteria. *Annu. Rev. Microbiol.* **30**:41-62.
16. Dooley, J.S.G., W.D. McCubbin, C.M. Kay, T.J. Trust. 1988. Isolation and biochemical characterization of the S-layer protein from a pathogenic *Aeromonas hydrophila* strain. *J. Bacteriol.* **170**:2631-38.
17. Sleytr, U.B., M. Sara. 1997. Bacterial and archaeal S-layer proteins: structure-function relationships and their biotechnological applications. *TIBTECH.* **15**:20-26.
18. Kuen, B., W. Lubitz. 1996. Analysis of S-layer proteins and genes. p. 77-102. *In* Sleytr, U.B., P. Messner, D. Pum, M. Sara (ed.), *Crystalline Bacterial Cell Surface Proteins*. Academic Press, R.G. Landes Company, Texas.
19. Inatomi, K., M. Ohba, T. Oshima. 1983. Chemical properties of proteinaceous cell wall from an acidothermophile, *Sulfolobus acidocaldarius*. *Chem. Lett.* **8**:1191-1194.
20. Koval, S.F., R.G.E. Murray. 1984. The isolation of surface array proteins from bacteria. *Can. J. Biochem. Cell Biol.* **62**:1181-1189.
21. Sleytr, U.B., P. Messner. 1983. Crystalline surface layers on bacteria. *Annu. Rev. Microbiol.* **37**:311-339.
22. Pum, D., P. Messner, U.B. Sleytr. 1991. Role of the S-layer in morphogenesis and cell division of the archaeobacterium *Methanococcus sinense*. *J. Bacteriol.* **173**:6865-73.
23. Sleytr, U.B., A.M. Glauert. 1975. Analysis of regular arrays of subunits on bacterial surfaces: evidence for a dynamic process of assembly. *J. Ultrastruct. Res.* **50**:103-116.
24. Baumeister W., I. Wildhaber, B.M. Phipps. 1989. Principles of organization in eubacterial and archaeobacterial surface proteins. *Can. J. Microbiol.* **35**:215-27.
25. Messner, P., U.B. Sleytr. 1992. Crystalline bacterial cell surface layers. p. 213-75, *In* Rose, A.H. (ed.), *Advances in Microbial Physiology* vol 33, London: Academic Press.

26. Beveridge, T.J. 1994. Bacterial S-layers. *Curr. Opin. Struct. Biol.* 4:204-12.
27. Sleytr, U.M., M. Sara. 1986. Ultrafiltration membranes with uniform pores form crystalline bacterial cell envelope layers. *Appl. Microbiol. Biotechnol.* 25:83-90.
28. Saxton, W.O., W. Baumeister. 1986. Principles of organization in S-layers. *J. Mol. Biol.* 187:251-253.
29. Wildhaber, I., U. Santarius, W. Baumeister. 1987. Three-dimensional structure of the surface protein of *Desulfurococcus modilis*. *J. Bacteriol.* 169:5563-5568.
30. Bingle, W.H., P.W. Whippey, J.L. Doran, R.G. Murray, W.J. Page. 1987. Structure of the *Azobacter vinelandii* surface layer. *J. Bacteriol.* 169:802-810.
31. Deatherage, J.F., K.A. Taylor, L.A. Amos. 1983. Three-dimensional arrangement of the cell wall protein of *Sulfolobus acidocaldarius*. *J. Mol. Biol.* 167:823-852.
32. Karlsson, B., T. Vaara, K. Lounatmaa, H. Gyllenberg. 1983. Three-dimensional structure of the regularly constructed surface layer from *Synechocystis* sp. Strain CLII. *J. Bacteriol.* 156:1338-1343.
33. Baumeister, W., H. Engelhardt. 1987. Three-dimensional structure of bacterial surface layers, p. 109-154. *In* J.R. Harris, R.W. Horne (eds.), *Electron microscopy of proteins*, vol. 6. Membranous structures. Academic Press, New York.
34. Lepault, J., N. Martin. K. Leonard. 1986. Three-dimensional structure of the T-layer of *Bacillus sphaericus* P-1. *J. Bacteriol.* 168:303-308.
35. Sara, M., K. Moser-Their, U. Kainz, U.B. Sleytr. 1990. Characterization of S-layers from mesophilic bacillaceae and studies on their protective role toward muramidases. *Arch. Microbiol.* 153:209-214.
36. Beveridge, T.J. 1981. Ultrastructure, chemistry, and function of the bacterial wall. *Int. Rev. Cytol.* 72:229-317.
37. Koval, S.F., R.G.E. Murray. 1985. Effect of calcium on the in vivo assembly of the surface protein of *Aquaspirillum serpens* VHA. *Can. J. Microbiol.* 31:261-267.
38. Sleytr, U.B., P. Messner. 1989. Self-assemblies of crystalline bacterial cell surface layers. p. 13-31. *In* Plattner, H. (ed.), *Electron Microscopy of Subcellular Dynamics*. CRC Press, Boca Raton.
39. Sleytr, U.B., A.M. Glauert. 1975. Analysis of regular arrays of subunits on bacterial surfaces; evidence for a dynamic process of assembly. *J. Ultrastruct. Res.* 50:103-116.

40. Sleytr, U.B., R. Plohberger. 1980. The dynamic process of assembly of two-dimensional arrays of macromolecules on bacterial cell walls. P. 36-47. *In* Baumeister, W., W. Vogell. (eds.), *Electron Microscopy at Molecular Dimensions*. Berlin: Springer.
41. Pum, D., M. Weinhandl, C. Hodl, U.B. Sleytr. 1993. Large-scale recrystallization of the S-layer of *Bacillus coagulans* E38-66 at the air/water interface and on lipid films. *J. Bacteriol.* **175**:2762-2766.
42. Sleytr, U.B. 1981. Morphopoietic and functional aspects of regular protein membranes present on prokaryotic cell walls. p. 3-26. *In* Kiermayer, O. (ed.), *Cytomorphogenesis in Plants, Cell Biology Monographs*, vol 8. Wien: Springer.
43. Belland, R.J., T.J. Trust. 1985. Synthesis, export, and assembly of *Aeromonas salmonicida* A-layer analyzed by transposon mutagenesis. *J. Bacteriol.* **163**:877-881.
44. Smit, J. 1987. Protein surface layers of bacteria. p. 343-376. *In* Inouye, M. (ed.), *Bacterial Outer Membranes as Model Systems*. New York: Wiley.
45. Noonan, B., T.J. Trust. 1995. Molecular analysis of an A-protein secretion mutant of *Aeromonas salmonicida* reveals a surface layer-specific protein secretion pathway. *J. Mol. Biol.* **248**:316-327.
46. Gruber, K., U.B. Sleytr. 1988. Localization insertion of new S-layer during cell growth of *Bacillus stearothermophilus* strains. *Arch. Microbiol.* **149**:485-491.
47. Smit, J., N. Agabain. 1982. Cell surface patterning and morphogenesis: biogenesis of a periodic surface array during *Caulobacter* development. *J. Cell Biol.* **95**:41-49.
48. Sleytr, U.B., P. Messner. 1988. Crystalline surface layers in procaryotes. *J. Bacteriol.* **170**:2891-2897.
49. Sara, M., E.M. Egelseer. 1996. Functional aspects of S-layers. p. 103-131. *In* Sleytr, U.B., P. Messner, D. Pum, M. Sara (eds.), *Crystalline Bacterial Cell Surface Proteins*. Academic Press, R.G. Landes Company, Texas.
50. Messner, P., D. Pum, M. Sara, K.O. Stetter, U.B. Sleytr. 1986. Ultrastructure of the cell envelope of the archaebacteria *Thermoproteus tenax* and *Thermoproteus neutrophilus*. *J. Bacteriol.* **166**:1046-1054.
51. Phipps, B.M., R. Huber, W. Baumeister. 1991. The cell envelope of the hyperthermophilic archaebacterium *Pyrobaculum organothrophum* consists of two regularly arrayed protein layers: three-dimensional structure of the outer layer. *Mol. Microbiol.* **5**:253-265.

52. Sumper, M. 1993. S-layer glycoproteins from moderately and extremely halophilic archaeobacteria. p. 109-117. *In* Beveridge, T.J., S.F. Koval (eds.), *Advances in Bacterial Paracrystalline Surface Layers*. Plenum Press, New York.
53. Trust, T.J., W.W. Kay, E.E. Ishiguro. 1983. Cell surface hydrophobicity and macrophage association of *Aeromonas salmonicida*. *Curr. Microbiol.* 9:15-18.
54. McCoy, E.C., D. Doyle, K. Burda, L.B. Corbeil, A.J. Winter. 1975. Superficial antigens of *Campylobacter fetus*: characterization of an antiphagocytic component. *Infect. Immun.* 11:517-525.
55. Sara, M., Kalsner, U.B. Sleytr. 1988. Surface properties from the S-layer of *Clostridium Thermosaccharolyticum* D120-70 and *Clostridium Thermohydrosulfuricum* L111-69. *Arch. Microbiol.* 149:527-533.
56. Weigert, S., M. Sara. 1995. Surface modifications of an ultrafiltration membrane with crystalline structure and studies on interactions with selected protein molecules. *J. Membrane Sci.* 106:147-159.
57. Stewart, M. T.J. Beveridge, T.J. Trust. 1986. Two patterns in the *Aeromonas salmonicida* A-layer may reflect a structural transformation that alters permeability. *J. Bacteriol.* 166:120-127.
58. Breitwieser, A., K. Gruber, U.B. Sleytr. 1992. Evidence for an S-layer protein pool in the peptidoglycan of *Bacillus stearothermophilus*. *J. Bacteriol.* 174:8008-8015.
59. Engelhardt, H., O.W. Saxton, W. Baumeister. 1986. Three-dimensional structure of the tetragonal surface layer of *Sporosarcina urea*. *J. Bacteriol.* 168:309-317.
60. Koval, S.F., S.H. Hynes. 1991. Effect of paracrystalline protein surface layers on predation by *Bdellovibrio bacteriovorus*. *J. Bacteriol.* 173:2244-2249.
61. Blaser, M.J. 1993. Role of S-layer proteins of *Campylobacter fetus* in serum-resistance and antigenic variation: a model of bacterial pathogenesis. *Am. J. Med. Sci.* 306:325-329.
62. Takeoka, A., K. Takumi, T. Kawata. 1991. Purification and characterization of S-layer proteins from *Clostridium difficile* GAI 0714. *J. Gen. Microbiol.* 173:261-267.
63. Dietz, P., P.K. Hansma, O. Inacker, H.D. Lehmann, H.K. Hermann. 1992. Surface pore structure of micro- and ultrafiltration membranes imaged with the atomic force microscope. *J. Membrane Sci.* 65:101-111.
64. Sara, M., S. Kupcu, U.B. Sleytr. 1996. Biotechnological applications of S-layers. p. 133-159. *In* Sleytr, U.B., P. Messner, D. Pum, M. Sara (ed.), *Crystalline Bacterial*

Cell Surface Proteins. Academic Press, R. G. Landes Company, Texas.

65. Sara, M., S. Kupcu, C. Weiner, et al. 1993. Crystalline protein layers as isoporous molecular sieves and immobilization matrices. p. 71-86 *In* Sleytr, U.B., P. Messner, D. Pum, M. Sara, (eds.), *Immobilized Macromolecules: Application Potentials* p. 71-86. Springer, London.
66. Pum, D., M. Sara, U.B. Sleytr. 1993. Two-dimensional (glyco)protein crystals as patterning elements and immobilization matrices for the development of biosensors. p. 142-160. *In* Sleytr, U.B., P. Messner, D. Pum, M. Sara (eds.), *Immobilized Macromolecules. Application Potentials*. Springer, London.
67. Evenberg, D., P. De Graaff, B. Lugtenberg, W.B. Van Muiswinkel. 1988. Vaccine-induced protective immunity against *Aeromonas salmonicida* tested in experimental carp erythrodermatitis. *J. Fish Disease* 11:337-350.
68. Kupcu, S., M. Sara, U.B. Sleytr. 1995. Liposomes coated with crystalline bacterial cell surface proteins (S-layers) as immobilization structures for macromolecules. *Biochim. Biophys. Acta* 1235:263-269.
69. Kay, W.W., T.J. Trust. 1991. Form and functions of the regular surface array (S-layer) of *Aeromonas salmonicida*. *Experientia* 47:412-414.
70. Udey, L.R., J.L. Fryer. 1978. Immunization of fish with bacterins of *Aeromonas salmonicida*. *Mar. Fish Rev.* 40:12-17.
71. Chu, S., S. Cavaignac, J. Feutrier, B.M. Phipps, M. Kostryznska, W.W. Kay, T.J. Trust et al. 1991. Structure of the tetragonal surface virulence array protein and gene of *Aeromonas salmonicida*. *J. Biol. Chem.* 266:15258-15265.
72. McCarthy, D.H., R.J. Roberts. 1980. Furunculosis of fish – the present state of our knowledge. *Adv. Aquat. Microbiol.* 2:293-340.
73. Trust, T.J., P.S. Howard, J.B. Chamberlain, E.E. Ishiguro, J.T. Buckley. 1980. Additional surface protein in autoaggregating strains of atypical *Aeromonas salmonicida*. *FEMS Microbiol. Lett.* 9:35-38.
74. Chart, H., D.H. Shaw, E.E. Ishiguro T.J. Trust. 1984. Structural and immunological homogeneity of *Aeromonas salmonicida* lipopolysaccharide. *J. Bact.* 158:16-22.
75. Evenberg, D., R. Versluis, B. Lugtenberg. 1985. Biochemical and immunological characterization of the cell surface of the fish pathogenic bacterium *Aeromonas salmonicida*. *Biochim. Biophys. Acta* 815:233-244.
76. Phipps, B.M., W.W. Kay. 1988. Immunoglobulin binding by the regular surface array of *Aeromonas salmonicida*. *J. Biol. Chem.* 263:9298-9303.

77. Garduno, R.A., B.M. Phipps, W.W. Kay. 1995. Physical and functional S-layer reconstitution in *Aeromonas salmonicida*. *J. Bacteriol.* **177**:2684-2694.
78. Ishiguro, E.E., W.W. Kay, T. Ainsworth, J.B. Chamberlain, R.A. Austen, J.T. Buckley, T.J. Trust. 1981. Loss of virulence during culture of *Aeromonas salmonicida* at high temperature. *J. Bacteriol.* **148**:333-340.
79. Dooley, J.S.G., H. Engelhardt, W. Baumeister, W.W. Kay, T.J. Trust. 1989. Three-dimensional structure of an open form of the surface layer from fish pathogen *Aeromonas salmonicida*. *J. Bacteriol.* **171**:190-197.
80. Doig, P., W.D. McCubbin, C.M. Kay, T.J. Trust. 1993. Distribution of surface-exposed and non-accessible amino acid sequences among the two major structural domains of the S-layer protein of *Aeromonas salmonicida*. *J. Mol. Biol.* **233**:753-765.
81. Phipps, B.M., T.J. Trust, E.E. Ishiguro, W.W. Kay. 1983. Purification and characterization of the cell surface virulence A protein from *Aeromonas salmonicida*. *Biochemistry.* **22**:2934-2945.
82. Munn, C.B., E.E. Ishiguro, W.W. Kay, T.J. Trust. 1982. Role of surface components in serum resistance of virulent *Aeromonas salmonicida*. *Infect. Immun.* **36**:1069-1075.
83. Chu, S., C.E. Gustafson, J. Feutrier, S. Cavaignac, T.J. Trust. 1993. Transcriptional analysis of the *Aeromonas salmonicida* S-layer protein gene vapA. *J. Bacteriol.* **175**:7968-7975.
84. Chu, S., T.J. Trust. 1993. An *Aeromonas salmonicida* gene which influences A-protein expression in *Escherichia coli* encodes a protein containing an ATP-binding cassette and maps beside the surface array protein gene. *J. Bacteriol.* **175**:3105-3114.
85. Chu, S., B. Noonan, S. Cavaignac, T.J. Trust. 1995. Endogenous mutagenesis by an insertion sequence element identifies *Aeromonas salmonicida* AbcA as an ATP-binding cassette transport protein required for biogenesis of smooth lipopolysaccharide. *PNAS* **92**:5754-5758.
86. Kay, W.W., E.E. Ishiguro, B.M. Phipps, et al. 1986. Properties, organization and role in virulence of the surface protein array of *Aeromonas salmonicida*, p. 217-227. *In* Vivares, C.P., J.R. Bonami, E. Jaspers (eds.), *Pathology in marine aquaculture*, special publication no. 9 European Aquaculture Society, Bredene, Belgium.
87. Doig, P., L. Emody, T.J. Trust. 1992. Binding of laminin and fibronectin by the trypsin-resistant major structural domain of the crystalline virulence surface array protein of *Aeromonas salmonicida*. *J. Biol. Chem.* **267**:43-49.
88. Trust, T.J. 1986. Pathogenesis of infectious disease of fish. *Annu. Rev. Microbiol.* **40**:479-502.

89. Kay, W.W., B.M. Phipps, E.E. Ishiguro, T.J. Trust. 1985. Porphyrin binding by the surface array virulence A-protein from *Aeromonas salmonicida* strains. *J. Bacteriol.* **164**:1332-1336.
90. Frausto da Silva, J.J.R., R.J.P. Williams. 1991. The biological chemistry of the elements. p. 1-561, Clarendon Press, Oxford.
91. Rossier, B.C., K. Geering, J.P. Kraehenbuhl. 1987. Regulation of the sodium pump: how and why. *Trends Biochem. Sci.* **12**:483-487.
92. Vogel, H.J. 1994. Calmodulin: a versatile calcium mediator protein. *Biochem. Cell Biol.* **72**:357-375
93. Heilbrunn, L.V., F.J. Weircenski. 1947. The action of various cations on muscle protoplasm. *J. Cell Comp. Physiol.* **29**:15-32.
94. Carafoli, E. 1987. Intracellular calcium homeostasis, *Annu. Rev. Biochem.* **56**:395-433.
95. Carafoli, E., J.T. Penniston. 1985. The calcium signal, *Sci. Am.* **253**:70-78.
96. Michalak, M., R.E. Milner, K. Burns, M. Opas. 1992. Calreticulin. *Biochem. J.* **285**:681-692.
97. Yano, K., A. Zarain-Herzberg. 1994. Sarcoplasmic reticulum calsequestrins: structural and functional properties. *Mol. Cell. Biochem.* **135**:61-70.
98. Augustine, G.J., M.P. Charlton, S.J. Smith. 1987. Calcium action in synaptic transmitter release. *Annu. Rev. Neurosci.* **10**:633-693.
99. Berridge, M.J. 1993. Inositol triphosphate and calcium signaling. *Nature* **361**:315-325.
100. Galione, A. 1993. Cyclic ADP-ribose: a new way to control calcium. *Science.* **259**:325-326.
101. Pietrobon, D., F. DiVirgilio, T. Pozzan. 1990. Structural and functional aspects of calcium homeostasis in eukaryotic cells. *Eur. J. Biochem.* **193**:599-622.
102. Gerday, C., R. Gilles, L. Bolis. 1988. Calcium and Calcium Binding Proteins. Berlin, Springer-Verlag.
103. Teo, T.S., J.H. Wang. 1973. Mechanism of activation of a cyclic adenosine 3'5'-phosphodiesterase from bovine heart by calcium ions. *J. Biol. Chem.* **248**:5950-5955.
104. Zhang, M., T. Yuan. 1998. Molecular mechanisms of calmodulin's functional versatility. *Biochem. Cell Biol.* **76**:313-323.

105. Persechini, A., K. McMillan, P. Leakey. 1994. Activation of myosin light chain kinase and nitric oxide synthase activities by calmodulin fragments. *J. Biol. Chem.* **269**:16-154.
106. Tallant, E.A., W.Y. Cheung. 1984. Activation of bovine brain calmodulin-dependent protein phosphatase by limited trypsinization. *Biochemistry* **23**:973-979.
107. Kincaid, R.L., I.E. Stith-Coleman, M. Vaughan. 1985. Proteolytic activation of calmodulin-dependent cyclic nucleotide phosphodiesterase. *J. Biol. Chem.* **260**:9009-9015.
108. Klee, C.B., T.C. Vanaman. 1982. Calmodulin. *Adv. Protein Chem.* **35**:213-321.
109. Means, A.R., M.F.A. Bagchi, K.P. Lu, C.D. Rasmussen. 1991. Regulatory functions of calmodulin. *Pharmacol. Ther.* **50**:255-270.
110. Silva, A.J., S. Paylor, J.M. Wehner, S. Tonegawa. 1992. Impaired spatial learning in α -calcium-calmodulin kinase II mutant mice. *Science* **257**:206-211.
111. Forsen, S., H.J. Vogel, T. Drakenberg. 1986. Biophysical studies of calmodulin. p. 113-157. *In* Cheung, E. (ed.), *Calcium and Cell Function*. vol. 6, Academic Press. New York.
112. Babu, Y.S., J.S. Sack, T.J. Greenbrough, C.E. Bugg, A.R. Means, W.J. Cook. 1985. Three-dimensional structure of calmodulin. *Nature* **315**:37-40.
113. Ikura, M., S. Spera, G. Barbato, L.E. Kay, M. Krinks, A. Bax. 1991. Secondary structure and side-chain resonance assignments of calmodulin by heteronuclear multidimensional NMR spectroscopy. *Biochemistry* **30**:9216-9228.
114. Barbato, G., M. Ikura, L.E. Kay, R.W. Pastor, A. Bax. 1992. Backbone dynamics of calmodulin studied by ^{15}N relaxation using inverse detected NMR spectroscopy: the central helix is flexible. *Biochemistry* **31**:5269-5278.
115. Ikura, M. 1996. Calcium binding and conformational response in EF-hand proteins. *TIBS* **21**:14-17.
116. Babu, Y.S., C.E. Bugg, W.J. Cook. 1988. Structure of calmodulin refined at 2.2 Å resolution. *J. Mol. Biol.* **204**:191-204.
117. Zhang, M., L. Ming, J.H. Wang, H.J. Vogel. 1994. The effect of Met-Leu mutations on calmodulin's ability to activate cyclic nucleotide phosphodiesterase. *J. Biol. Chem.* **22**:15546-15552.
118. Chattopadhyaya, R., W.E. Meador, A.R. Means, F.A. Quijcho. 1992. Calmodulin structure refined at 1.7 Å resolution. *J. Mol. Biol.* **228**:1177-1192.

119. O'Neil, K.T., W.F. DeGrado. 1990. How calmodulin binds its targets: sequence independent recognition of amphiphilic α -helices. *TIBS* **15**:59-64.
120. Anderson, A., S. Forsen, E. Thulin, H.J. Vogel. 1983. Cadmium-113 NMR studies of proteolytic fragments of calmodulin, assignment of strong and weak cation binding sites. *Biochemistry* **22**:2309-2313.
121. Ikura, M., N. Hasegawa, S. Aimoto, M. Yazawa, K. Yagi, K. Hitachi. 1989. Cadmium-113 NMR evidence for cooperative interactions between amino and carboxyl terminal domains of calmodulin. *Biocem. Biophys. Res. Commun.* **161**:1233-1238.
122. Kuboniwa, H., N. Tjandra, S. Grzesiek, H. Ren, C.B. Klee, A.B. Bax. 1995. Solution structure of calcium-free calmodulin. *Nat. Struct. Biol.* **2**:768-776.
123. Ikura, M., G.M. Clore, A.M. Gronenborn, G. Zhu, C.B. Klee, A.Bax. 1992. Solution structure of a calmodulin-target peptide complex by multidimensional NMR. *Science* **256**:632-638.
124. Meador, W.E., A.R. Means, F.A. Quijano. 1992. Target enzyme recognition by calmodulin: 2.4 Å structure of a calmodulin-peptide complex. *Science* **257**:1251-1255.
125. Vogel, H.J., M. Zhang. 1995. Protein engineering and NMR studies of calmodulin. *Mol. Cell. Biochem.* **149/150**:3-15.
126. Gould, R., A.J. Bains. 1994. Evidence that two non-overlapping high-affinity calmodulin-binding sites are present in the head region of synapsin I. *Eur. J. Biochem.* **224**:229-240.
127. Meador, W.E., A.R. Means, F.A. Quijano. 1993. Modulation of calmodulin plasticity in molecular recognition on the basis of X-ray structures. *Science* **262**:1718-1721.
128. Zhang, M., H.J. Vogel. 1994. The calmodulin-binding domain of caldesmon binds to calmodulin in an α -helical conformation. *Biochemistry* **33**:1163-1171.
129. Meador, W.E., S.E. George, A.R. Means, F.A. Quijano. 1995. X-ray analysis reveals conformational adaptation of the linker in functional calmodulin mutants. *Nat. Struct. Biol.* **2**:943-945.
130. Bernstein, H.D., M.A. Poritz, K. Strub, P.J. Hoben, J. Brenner, P. Walter. 1989. Model for signal sequence recognition from amino-acid sequence of 54K subunit of signal recognition particle. *Nature* **340**:482-486.
131. Yuan, T., H.J. Vogel. 1999. Substitution of the methionine residues of calmodulin with the unnatural amino acid analogs ethionine and norleucine: biochemical and

- spectroscopic studies. *Prot. Sci.* **8**:113-121.
132. Gellman, S.H. 1991. On the role of methionine residues in the sequence independent recognition of nonpolar protein surfaces. *Biochemistry* **30**:6633-6636.
 133. Edwards, R.A., M.P. Walsh, C. Sutherland, H.J. Vogel. 1998. Activation of calcineurin and smooth muscle myosin light chain kinase by Met to Leu mutants of calmodulin. *Biochem. J.* **331**:149-152.
 134. Maniatis, T., E.F. Fritsch, J. Sambrook. 1982. *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor N.Y.
 135. Filip, C., G. Fletcher, J.L. Wulff, C.F. Earhart. 1973. Solubilization of the cytoplasmic membrane of *Escherichia coli* by the ionic detergent sodium-lauryl sarcosinate. *J. Bacteriol.* **115**:717-722.
 136. Zhang M., H.J. Vogel. 1993. Determination of the side chain pK_a values of the lysine residues in calmodulin. *J. Biol. Chem.* **268**:22420-22428.
 137. Zhang M., H.J. Vogel. 1994. Two-dimensional NMR studies of selenomethionine calmodulin. *J. Mol. Biol.* **239**:545-554.
 138. Laemmli, U.K. 1970. Cleavage of the structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**:680-685.
 139. Tsai, C.M., C.E. Frasch. 1982. A sensitive silver stain for detecting lipopolysaccharide in polyacrylamide gels. *Anal. Biochem.* **119**:115-119.
 140. Otwinowski, Z., W. Minor. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Met. Enzymology* **267**: Macromolecular crystallography, part A. p.307-326, Carter, Jr. & R.M. Sweet, Eds., Academic Press.
 141. Brunger, A.T. 1998. Crystallography and NMR System: A new software system for macromolecular structure determination. *Acta. Cryst.* **D52**:905-921.
 142. Kjeldgaard, M., T.A. Jones. 1993. O version 5.9. Department of Molecular Biology, BMC, Uppsala University, Sweden; Department of Chemistry, Aarhus University, Denmark.
 143. O'Farrell, P.H. 1975. High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* **250**:4007-4021.
 144. Rupp, B., D.R. Marshak, S. Parkin. 1996. Crystallization and preliminary X-ray analysis of two new crystal forms of calmodulin. *Acta Cryst.* **D52**:411-413.
 145. Taylor, D.A., J.S. Sack, J.F. Maune, K. Beckingham, F.A. Quijcho. 1991. Structure of a recombinant calmodulin from *Drosophila melanogaster* refined at 2.2-Å

- resolution. *J. Biol. Chem.* **266**:21375-21380.
146. Rao, S.T., S.Wu, K.A. Satyshur, K.Y. Ling, C. Kung, M. Sundaralingam. 1993. Structure of *Paramecium tetraurelia* calmodulin at 1.8 Å resolution. *Prot. Science* **2**:436-447.
 147. Ban, C., B. Ramakrishnan. 1994. Structure of the recombinant *Paramecium tetraurelia* calmodulin at 1.68 Å resolution. *Acta Cryst.* **D50**:50-63.
 148. Mikol, V., R. Giege. 1992. Physical chemistry of protein crystallization, p. 73-97. *In* Ducruix, A. and R. Giege eds. *Crystallization of Nucleic Acids and Proteins: A Practical Approach*. Oxford University press.
 149. Hampton Research. 1996-1999. Crystallization Tips. *In* *Crystallization Research Tools*. Volume 6-9.
 150. Rhodes, G. 1993. *Crystallography Made Crystal Clear*. p.1-197. Academic Press.
 151. Drenth, J. 1994. *Principles of Protein X-ray Crystallography*. Springer Verlag.
 152. Giege, R., A.C. Dock, D. Kern. 1986. The role of purification in the crystallization of proteins and nucleic acids. *J. Crystal Growth*. **76**:554.
 153. Phipps, B.M. 1988. Purification, Biochemical Characterization and Molecular Cloning of the Regular Surface Layer A Protein of *Aeromonas salmonicida*. Dissertation, University of Victoria.
 154. Sleytr, U.B., K.J.I. Thorne. 1976. Chemical characterization of the regularly arranged surface layers of *Clostridium thomosaccharolyticum* and *Clostridium thermohydrosulfuricum*. *J. Bacteriol.* **126**:377-383.
 155. Wriston, J.C., T.O. Yellin. 1973. L-asparaginase: a review. *Adv. Enzymol.* **39**:185-248.
 156. Geiger, T., S. Clarke. 1987. Deamidation, isomerization, and racemization at asparaginal and aspartyl residues in peptides: succinimide-linked reactions that contribute to protein degradation. *J. Biol. Chem.* **262**:785-794.
 157. Lambden, P.R. 1982. Biochemical comparison of pili from variants of *Neisseria gonorrhoeae*. *J. Gen. Microbiol.* **128**:2105-2111.
 158. Virji, M., J. R. Saunders, G. Sims, K. Makenpeace, D. Maskell, D.J.P. Ferguson. 1993. Pilus-facilitated adherence of *Neisseria meningitides* to human epithelial and endothelial cells: modulation of adherence phenotype occurs concurrently with changes in primary amino acid sequence and glycosylation status of pilin. *Mol. Microbiol.* **10**:1013-1028.

159. H. Parge, K.T. Forest, M.J. Hickey, D. Christensen, E. Getzoff, J. Tainer. 1995. Structure of the fibre-forming protein pilin at 2.6 Å resolution. *Nature* **378**:32-38.
160. Osawa, M., H. Tokumitsu, M.B. Swindells, H. Kurihara, M. Orita, T. Shibanuma, T. Furuya, M. Ikura. 1999. A novel target recognition revealed by calmodulin in complex with Ca²⁺-calmodulin-dependent kinase kinase. *Nat. Struct. Biol.* **6**:819-824.

APPENDIX A: Copyright permission



UNIVERSITY OF
CALGARY

153

FACULTY OF SCIENCE

Department of Biological Sciences
Telephone: (403) 220-5261
Fax: (403) 289-9311

December 6, 2000

R.G. Landes Company,
909 Pine Street, Georgetown
Texas, USA 78626
Fax (512) 863-0081

Dear Sir or Madam:

I am a graduate student at the University of Calgary in the Department of Biological Sciences. I am writing to request permission to reproduce the following figures from the book:

"Crystalline Bacterial Cell Surface Proteins" (1996) U.B. Sleytr, F. Messner, D. Purn,
and M. Sara.

- Figure 2.1 on page 7
- Figure 2.3 on page 10
- Figure 2.4 on page 11

These figures would be included in a literature review of bacterial surface layer proteins as a part of my Masters Thesis at the University of Calgary.

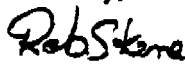
I can be contacted by email: rjskene@ucalgary.ca
phone: (403) 220-8560

The fax number for the Department of Biological Sciences at the U of C is:
Fax: (403) 289-9311

I appreciate your consideration

12/13/2000

Yours truly,


Robert J. Skene

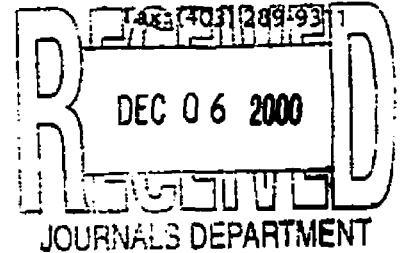
Permission is granted. Please give appropriate acknowledgement to authors, publication and publisher.


Sara Johnson



FACULTY OF SCIENCE

Department of Biological Sciences
Telephone: (403) 220-5261



December 6, 2000

American Society for Microbiology
1752 N Street, N.W., Washington, DC, USA 20036-2804
Fax (202) 942-9355

Journals Department

Dear Sir or Madam:

I am a graduate student at the University of Calgary in the Department of Biological Sciences. I am writing to request permission to reproduce the following figures from the "Journal of Bacteriology":

1. Dooley J.S.G., H. Engelhardt, W. Baumeister, W.W. Kay, and T.J. Trust. (1989). Three-Dimensional Structure of an Open Form of the Surface Layer from the Fish Pathogen *Aeromonas salmonicida*. *J. Bacteriology*, volume 171: 190-197.

- Figure 6 on page 194
- Figure 7 on page 195

These figures would be included in a literature review of bacterial surface layer proteins as a part of my Masters Thesis at the University of Calgary.

I can be contacted by email: rjskene@ucalgary.ca
phone: (403) 220-8560

The fax number for the Department of Biological Sciences at the U of C is:
Fax: (403) 289-9311

I appreciate your consideration

Yours truly,

Robert J. Skene

PERMISSION GRANTED CONTINGENT ON AUTHOR PERMISSION
AND APPROPRIATE CREDIT
American Society for Microbiology
Journals Department

Date 12-7-00