

Probing the Structure of Calbindin D_{28k}
Using Constraints from
Intramolecular Cross-links and Molecular Modeling

David Yong-Hoi Yeung

A Thesis in
The Department
of
Chemistry and Biochemistry

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Science at
Concordia University
Montreal, Quebec, Canada

August 2004

© David Yong-Hoi Yeung, 2004



Library and
Archives Canada

Bibliothèque et
Archives Canada

Published Heritage
Branch

Direction du
Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 0-612-94671-1

Our file Notre référence

ISBN: 0-612-94671-1

The author has granted a non-exclusive license allowing the Library and Archives 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 et Archives 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.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

Probing the Structure of Calbindin D_{28k} Using Constraints from Intramolecular Cross-links and Molecular Modeling

David Yong-Hoi Yeung

Calbindin D_{28k} is a calcium binding protein of unknown structure. It is believed to be composed of six EF-hand subdomains (helix-loop-helix motifs), which is supported in part by NMR analysis of a recombinant fragment containing the first and second EF-hands. A recent study demonstrated the structural characterization of fibroblast growth factor using sequence alignment and distance constraints based on several identified intramolecular chemical cross-links. A similar approach is applied in this thesis to structurally characterize human calbindin D_{28k}. Reactions of the protein with a bis(sulfosuccinimidyl) suberate, a lysine specific cross-linker, yields primarily protein modified with a single cross-linker. Identified cross-linked peptides in the MALDI mass fingerprints of tryptic digests of modified calbindin D_{28k} impose distance constraints between pairs of lysine residues. Applying these constraints, hypothetical structures of the protein were constructed. Of particular interest is the orientation of the EF-hand subdomains with respect to each other. Of the 103 configurations considered, only 14 respected the distance constraints. One of these 14 structures was selected to demonstrate the use of molecular mechanics for structure refinement.

Cross-linking of myoglobin, a protein of known structure, was first performed to evaluate the experimental method. Because the calcium loaded form of calbindin D_{28k} was studied, the calcium coordination spheres of four of its EF-hands were investigated using molecular modeling. Additionally, the expression and purification of recombinant human calbindin D_{28k} is described.

Acknowledgements

I would first like thank my research supervisors, Prof. Ann M. English and Prof. Gilles H. Peslherbe for their guidance and for allowing me the opportunity to undertake this project. I would like to acknowledge my committee members, Prof. Heidi Muchall and Prof. Joanne Turnbull for their valuable insight.

I would also like to thank an extensive list of fellow lab members for the many fruitful discussions and for making my M.Sc. experience a positive one. Tao-Nhan Nguyen, Denise Koch, Sean Hughes, Jeff Hasnain, Yin Wei, Lei Zhang, Robert Mawhinney, Sacha Zlatkova, Sichuan Xu, Etienne Paradis, Mohammed Fares, Celine Toubin, Pascal Larregaray, Mouna Sbata, Paul Loncke, Cyril Puel, Elena Ivanova, Qadir Timerghazin, Grygoriy Dolgonos, Svetlana Popenova, Georgia Kremmydiotis, Jiang Heng, Jacqueline Montalibet, Biao Shen, Mengwei Ye, Steve Spetsieris, Ernesto Moran, Adamo Petosa, Jean-Philippe Lambert, Kathryn Potter, Limei Tao, Andrea Romeo, John Wright, Meileen Joynt, Pascal Turcotte, Vira Patel, Robert Papp, Eaquab Ali, Mihai Ciortea, Farida Mohammed, Lekha Sleno, Dominic Cuerrier and Jesse Kao, thanks.

Special thanks go out to my parents Chak Kwan and Chi Kit for their unwavering support. I'd also like to thank my brother and sister, Rita and Hunt as well as my parents in-law, Siowa and Thérèse.

Finally, I would like to thank Mireille, my wife, for her encouragement and patience. You're so great!

Table of contents

List of Figures	vii
List of Tables	ix
List of Abbreviations	x
1.0. Introduction	1
1.1. Calbindin D _{28k}	1
1.2. Function of calbindin D _{28k}	1
1.3. Structural characteristics of calbindin D _{28k}	3
1.3.1. Calcium-binding stoichiometry	7
1.4. Mass spectrometry in the structural characterization of proteins	8
1.4.1. Chemical cross-linking as a tool for throughspace distance information	9
1.5. Purpose of research	11
1.6. Outline of thesis and contributions of colleagues	12
2.0. Calbindin D _{28k} expression system	14
2.1. Introduction	14
2.2. Material and methods	15
2.2.1. JM105-pTrc99A-His-calb system	15
2.2.2. Protein purification	19
2.3. Results	22
2.3.1. JM105-pTrc99A-His-calb expression system	22
2.3.2. Expression and purification of recombinant human calbindin D _{28k}	24
2.3.3. Analysis of thiol reactivity in His-calb vs. untagged calbindin D _{28k}	27
2.4. Discussion	30
3.0. Reactions of myoglobin and calbindin D _{28k} with the homocross-linker bis(sulfosuccinimidyl) suberate	33
3.1. Introduction	33
3.2. Materials and methods	36
3.2.1. Materials	36
3.2.2. BS3 Cross-linking reactions	36
3.2.3. ESI-MS of intact myoglobin and calbindin D _{28k} from BS3 cross-linking reactions	37
3.2.4. SDS-PAGE analysis of cross-linked myoglobin and cross-linked calbindin D _{28k} solutions	38
3.2.5. Digestion of the myoglobin and calbindin D _{28k} products from the BS3 cross-linking reaction	38
3.2.6. MALDI-MS analysis of the cross-linked peptides	39
3.3. Results	40

3.3.1. Cross-linking of myoglobin	40
3.3.2. Cross-linking of calcium-loaded calbindin D _{28k}	44
3.3.3. Identification of cross-linked sites by MALDI-ToF-MS	45
3.4. Discussion	48
3.4.1. Myoglobin and calbindin D _{28k} cross-linked products	48
3.4.2. Cross-linked peptides in myoglobin	55
3.4.3. Factors influencing the cross-linking of myoglobin	57
3.4.4. Cross-linked peptides in calbindin D _{28k}	57
3.4.5. Stability of the BS3 cross-linked peptides	58
3.4.6. Abundance of the cross-linked peptides	59
4.0. Molecular modeling	62
4.1. Introduction	62
4.1.1. Building the calbindin D _{28k} models	62
4.1.2. Calcium-binding in calbindin D _{28k}	64
4.2. Methods	64
4.2.1. Molecular modeling software	64
4.2.2. Modeling calcium-binding in calbindin D _{9k}	66
4.2.3. Building the EF-hand pairs of calbindin D _{28k}	67
4.3. Results	69
4.3.1. Calcium-binding in calbindin D _{9k}	69
4.3.2. Modeled EF-hand pairs of calbindin D _{28k}	70
4.3.3. Modeled full-length calbindin D _{28k} protein	71
4.3.4. Refinement of a full-length calbindin D _{28k} structure	75
4.4. Discussion	77
4.4.1. Screening the calbindin D _{28k} structures	77
4.4.2. Refinement of the proposed structure of human calbindin D _{28k}	78
4.4.3. Modeling calcium coordination sites	81
5.0. Summary and suggested further studies	85
5.1. Summary	85
5.2. Future study	87
Appendix A	89
Appendix B	91
Appendix C	93
References	105

List of Figures

Figure 1.1	A ribbon diagram of an EF-hand motif	4
Figure 1.2	Primary structure and secondary structural features of human calbindin D _{28k}	5
Figure 1.3	Consensus sequence for the calcium-binding loop of EF-hand subdomains	5
Figure 2.1	Plasmid maps of pTrc99A and pGYMX-His-calb	16
Figure 2.2	Restriction analysis of pTrc99A and pGYMX-His-calb	22
Figure 2.3	Restriction analysis of the plasmids isolated from JM105 cells transformed with pTrc99A-His-calb	23
Figure 2.4	10% SDS-PAGE analysis of steps involved in His-calb purification on the Ni-NTA column	25
Figure 2.5	MS analysis of purified His-calb	26
Figure 2.6	SDS-PAGE analysis of His-tag cleavage in His-calb	27
Figure 2.7	MS analysis of purified recombinant calbindin D _{28k}	28
Figure 2.8	Deconvoluted ESI-MS spectra of (a) His-calb and (b) calbindin D _{28k}	29
Figure 3.1	Stick diagram of bis(sulfosuccinimidyl) suberate (BS3)	34
Figure 3.2	12% SDS-PAGE analysis of myoglobin from the BS3 cross-linking reactions	41
Figure 3.3	(a) ESI-MS spectrum of myoglobin reacted with 3-fold molar excess of BS3	43
Figure 3.4	Deconvoluted ESI-MS spectrum of myoglobin reacted with 5-fold molar excess of BS3	44
Figure 3.5	15% SDS-PAGE analysis of calbindin D _{28k} from the BS3 cross-linking reactions	45
Figure 3.6	ESI-MS spectrum of calbindin D _{28k} reacted with 5-fold molar excess of BS3	46

Figure 3.7	MALDI mass fingerprint of tryptic digests of (a) unmodified myoglobin and (b) myoglobin reacted with a 3-fold excess of BS3	50
Figure 3.8	MALDI mass fingerprint of the in-gel tryptic digest of (a) unmodified myoglobin and (b) BS3-cross-linked (Mb-BS3)	51
Figure 3.9	MALDI mass fingerprint of a GluC digest of (a) unmodified calbindin D _{28k} and (b) calbindin D _{28k} reacted with 5-molar excess of BS3	52
Figure 3.10.	MALDI mass fingerprint of a trypsin and GluC double digest of (a) unmodified calbindin D _{28k} and (b) calbindin D _{28k} reacted with 5-molar excess of BS3	53
Figure 3.11.	Ribbon diagram of the NMR derived structure of myoglobin showing the six observed BS3 cross-links	56
Figure 3.12.	Stick diagram of disuccinimidyladipate (DSA)	60
Figure 4.1.	Residues used in modeling the individual EF-hands of calbindin D _{28k}	68
Figure 4.2.	Predicted calcium coordination in EF2 of calbindin D _{9k}	71
Figure 4.3.	Ribbon diagram of predicted structures of the three EF-hand pairs of calbindin D _{28k}	71
Figure 4.4.	Representation of EF-hands	72
Figure 4.5.	The 14 EF-hand configurations that are consistent with the experimentally derived distance restraints	73
Figure 4.6.	Energy fluctuation of calbindin D _{28k} during the 1000 ps MD simulation	75
Figure 4.7.	Refined structure of calbindin D _{28k}	76
Figure 4.8.	Linear arrangement of EF-hand pairs of calbindin D _{28k} .	78
Figure 4.9.	Ribbon diagram of refined structure of calbindin D _{28k} showing the positions of the cysteine residues	80

List of Tables

Table 1.1.	Intermolecular interactions of fragments representing EF-hands of calbindin D _{28k}	7
Table 1.2.	Top 20 sequence homology models for FGF	11
Table 2.1.	Recombinant calbindin D _{28k} expression systems	32
Table 3.1.	Possible products following modification of protein by BS3	42
Table 3.2.	Modified peptides detected in the MALDI mass fingerprint of horse heart myoglobin from a BS3 reaction	49
Table 3.3.	Modified peptides detected in the MALDI mass fingerprint of calbindin D _{28k} from a BS3 reaction	54
Table 3.4.	Relative abundance of ions arising from modified proteins in the ESI mass spectra	55
Table 4.1.	Comparison between predicted and X-ray derived calcium-oxygen distances in EF2 of calbindin D _{9k}	70
Table 4.2.	C _α separations of lysine residues of calbindin D _{28k} that were observed to be cross-linked	76
Table 4.3.	Predicted coordination sphere and calcium-oxygen distances (<i>r</i>) in the modeled calcium-binding loops of calbindin D _{28k}	77
Table 4.4.	Distances (Å) separating the sulfur atoms of the cysteine residues in the modeled structure of calbindin D _{28k}	80
Table 4.5.	Predicted calcium coordination spheres of calbindin D _{28k} vs. consensus ligands	83

List of Abbreviations

ACN	Acetonitrile
amp	Ampecilin
APS	Ammonium persulfate
ASAP	Automated Spectrum Assignment Program
bp	Base pairs
BS3	Bis(sulfosuccinimidyl) suberate
C _α	Alpha-carbon
CD	Circular dichroism
Da	Dalton
DSA	Disuccinimidyladipate
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
ESI	Electrospray ionization
FA	Formic acid
HCCA	4-hydroxy- α -cyanocinnamic-acid
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC	High pressure liquid chromatography
IPTG	Isopropyl- β -D-thiogalactopyranoside
LB	Luria-Bertani
m/z	Mass-to-charge
MAC	Metal affinity chromatography
MALDI-ToF	Matrix-assisted laser desorption ionization time-of-flight
Mb	Myoglobin
MD	Molecular dynamics
MKK1	MAP kinase kinase-1
MS	Mass spectrometry
NEM	N-ethylmaleimide
NEM	N-ethyl morphaline
Ni-NTA	Nickel-nitrilotriacetic acid
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
NOG	N-octylglucopyranosid
N _ζ	Zeta-nitrogen
PDB ID	Protein Data Bank Identification
ppm	Parts per million
RMSD	Root-mean-square deviation
SDS	Sodium dodecylsulfate
SDS-PAGE	Sodium dodecylsulfate polyacrylamide gel electrophoresis
TFA	Trifluoroacetic acid
Tris	Tris(hydroxymethyl)methylamine

1.0. Introduction

1.1. Calbindin D_{28k}

Calbindin D_{28k}, as the name implies, is a calcium-binding protein. Early studies of the protein showed that calbindin D_{28k} in chick duodenum exhibited vitamin dependant expression (1, 2). The “D” in the name indicates this property. However, the vitamin D dependence for protein expression does not hold true for all organisms (3). The “D” is nevertheless still employed in naming the protein. A recent review of calcium-binding proteins indicates that the distribution of calbindin D_{28k} in organisms varies between species (3, 4). In humans, it is found abundantly in neuronal cells (3), particularly in the human brain where its concentration ranges from 0.5 to 1.5 % (5). The mass of human brain calbindin D_{28k} was initially estimated to be 28 kDa as determined by SDS-PAGE analysis. Human calbindin D_{28k} was later sequenced (6, 7) and its theoretical average mass is calculated to be 29981 Da.

1.2. Function of calbindin D_{28k}

The precise physiological function of calbindin D_{28k} is not known. It is a water soluble cytosolic protein (3) found in high concentrations in specific regions of the brain (5). One of the proposed roles of calbindin D_{28k} is that it acts as a calcium buffer thereby preventing cell damage or excitotoxicity caused by elevated levels of calcium. Consistent with this proposal is that neurons affected by neurodegenerative diseases such as Huntington’s, Alzheimer’s and Parkinson’s diseases show low levels of calbindin D_{28k} expression (8). As an example, a mutation in the presenilin gene in neuronal cells

destabilizes calcium homeostasis. The cells become sensitive to apoptosis, which leads in some cases to Alzheimer's disease. However, upon overexpression of calbindin D_{28k} in these cells, apoptosis is averted (9). In this case, calbindin D_{28k} presumably acts as a buffer by regulating the calcium levels. Other known proteins that act as calcium buffers include calbindin D_{9k} and parvalbumin (3).

Calbindin D_{28k} has also been proposed to act as a calcium sensor. A characteristic of a calcium sensor is that the protein undergoes a conformational change upon binding calcium, which promotes interactions with a target molecule. The conformational change typically alters the surface accessibility of hydrophobic patches, thus affecting the sensor's ability to interact with target molecules (10). Calmodulin is a well documented example of a calcium sensor (10, 11). In its apo form, the protein is dumbbell shaped with two subunits connected through a long α -helix (11). In the calcium-loaded form, hydrophobic patches become more exposed allowing for interactions with ligands (11). In this state, calmodulin adopts a globular conformation (11).

Calbindin D_{28k} has also been shown to undergo structural changes upon binding or release of calcium (12-14). However, unlike most calcium sensors, the protein appears to have exposed hydrophobic patches in both the apo and calcium-loaded states (12). If calbindin D_{28k} does act as a sensor, its role will be determined by identifying molecules with which it interacts. Linse *et al.* identified myo-inositol monophosphatase as a potential target (15). Interaction with calbindin D_{28k} increases the activity of myo-inositol phosphatase, which is involved in the signaling pathway of inositol-1,4,5-triphosphate, a molecule responsible for opening calcium channels. It has also been shown *in vitro* that copper, zinc-superoxide dismutase catalyzes the S-nitrosation of

calbindin D_{28k} by S-nitrosoglutathione (16). This suggests that calbindin D_{28k} may function as a nitric oxide buffer or transporter.

1.3. Structural characteristics of calbindin D_{28k}

Efforts have been made to characterize at high resolution the structure of calbindin D_{28k} using conventional NMR spectroscopic (14, 17) and X-ray crystallographic methods (Prof. Michael Murphy, UBC, personal communication). X-ray crystallography is generally capable of providing protein structures with lower than 2 Å resolution. However, crystallographic analysis is sometimes limited to the quality (size and uniformity) of the crystals formed by the protein. Indeed, crystallization of calbindin D_{28k} has yielded crystals that were not of crystallographic quality (Prof. Michael Murphy, personal communication). NMR spectroscopy has also been employed to study the structure of calbindin D_{28k} (14, 17). The NMR peak assignments for a mutant form of calbindin D_{28k} have recently been published (17) but the structure has not yet been reported. The mutations were designed to circumvent deamination and disulfide bridge formation, which result in linebroadening and spectral heterogeneity. All five cysteines were mutated to serines and Q182 and N203 were mutated to glutamate and aspartate residues, respectively. Because of the large size of the protein (~30 kDa), ¹³C and ¹⁵N labeling was employed.

Although the tertiary structure of calbindin D_{28k} remains elusive, a number of its structural features have been described. The primary structure has been solved by gene (6) and protein (7) sequencing. Both show that the human protein is composed of 261

residues. It exists as a monomer under reducing conditions, but can form intra- and intermolecular disulfide bridges *in vitro* (12).

Through sequence homology modeling, the protein is thought to be composed of six helix-loop-helix motifs (6), termed EF-hands by Kretsinger (18). This motif is commonly observed in calcium-binding proteins. The EF-hand is an independent fold, but does not exist as a separate domain. Moreover, its structure is dependent on interactions with a second EF-hand partner (19). The EF-hand is therefore referred to as a subdomain (19). The EF-hand is so named because it was first observed that the fifth and sixth (or the E and the F) helices of parvalbumin formed this particular subdomain. The term “hand” stems from the observation that, if the thumb and the index finger were stretched while keeping the remaining fingers curled, the positions of the index and the thumb resemble the arrangement of the two helices in the subdomain. The calcium-binding site is a loop region between the two helices. The structure of an EF-hand of parvalbumin (PDB ID: 1b8r) is shown in Figure 1.1. The helices and the calcium-binding loop each span approximately 12 residues in length. The EF-hands are connected

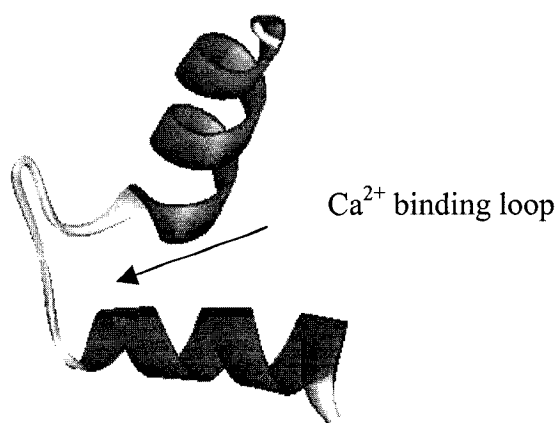


Figure 1.1. Ribbon diagram of an EF-hand motif. The figure is taken from the NMR derived structure of parvalbumin (PDB ID: 1b8r). Note that the E and the F helices are close to perpendicular.

by linker loops. The primary structure and the proposed secondary structural elements of calbindin D_{28k} are shown in Figure 1.2.

Linker loop	α -helix (E)	Ca ²⁺ binding-loop	α -helix (F)	Linker loop
MAESHLQSSLIT	ASQFFEIWLHF	DADGSGYLEGKE	LQNLIQEL	QQARK
KAGLELSP	EMKTFVDQY	GQRDDGKIGIVE	LAHVLPTEE	NFLLFR
CQQLKSCE	EFMKTWRKY	DTDHSGFIETEE	LKNFLKDLL	EKANKTVD
DTKLAE	YTDLMLKLF	DSNNDGKLELTE	MARLLPVQE	NFLLKFQ
GIKMCGK	EFNKAFELY	DQDNGYIDENE	LDALLKDL	CEKNKQDLDI
NN	ITTYKKNIM	ALSDGGKLYRTD	LALILCAGD	N

Figure 1.2. Primary structure and secondary structural features of human calbindin D_{28k}. This Figure is adapted from Reference (13).

Within the calcium-binding loop, oxygen atoms of the sidechains and the peptide carbonyl ligate the calcium ion (11). The consensus sequence for the calcium-binding site is shown in Figure 1.3.

1 (X)	2	3 (Y)	4	5 (Z)	6	7	8	9 (-X)	10	11	12 (-Z)
D	-	D/N	-	D/N	G	-	I/L/M	-	-	-	E

Figure 1.3. Consensus sequence for the calcium-binding loop of EF-hand subdomains. The dashes (-) represent weakly conserved residues. X, Y, Z, -X and -Z (E12 donates 2 oxygen ligands) refer to the positions of the 6 oxygens provided by the protein to the calcium coordination sphere (3, 20). The conserved residues listed above are present in more than 80% of EF-hand sequences. This figure is adapted from Reference .

In addition to the homology modeling, experimental data suggest that the protein is composed of EF-hands. The structure of a truncated mutant representing the first and second EF-hands of calbindin D_{28k} was determined by NMR spectroscopy (12). The derived structure indicates that the mutant protein is indeed composed of two EF-hands as predicted. Moreover, the structure shows the presence of a anti-parallel β -sheet region spanning 3 residues in the 12 residue calcium-binding loop, a feature commonly observed in calcium-binding proteins (11).

In an attempt to gain insight into the arrangement of the six EF-hands in calbindin D_{28k}, fragments representing the EF-hands have been constructed (19, 21, 22). Intermolecular interactions between these peptides were probed using spectroscopic and chromatographic methods (19, 21, 22). From these experiments calbindin D_{28k} was shown to be a single domain globular protein. For example, mixtures containing less than 6 EF-hands were examined by far-UV CD spectroscopy. The “missing” EF-hands were independently measured and mathematically summed. The assumption is that if separate domains exist, the sum of spectra of the mixtures should be similar to the spectrum of the native protein. The sums of spectra of mixtures of 2 to 5 peptides were compared to the native protein. The spectral summations could not reproduce the spectrum of the native protein. The only case where the CD spectrum of the native protein was reproduced was if all 6 EF-hands were present together in solution. Calbindin D_{28k} must therefore exist as a single domain globular protein. This is in contrast to the dumbbell-like arrangement of calcium-free calmodulin (11).

Using a chromatographic approach to detect interacting partners, the (negatively charged) constructs were mixed in a variety of combinations and injected into an anion exchange column. An increase in the peptide’s elution volume when in the presence of other fragments indicates that there is intermolecular interaction. This is because interaction between peptides results in a complex with greater negative charge than the isolated peptides. Peptides observed to interact are summarized in Table 1.1. The data for EF2, EF3 and EF4 indicate that these EF-hands are able to interact with several others. However, these are not necessarily representative of natural interacting partners.

Table 1.1. Intermolecular interactions of fragments representing EF-hands of calbindin D_{28k}^a

EF-hand ^b	Mixtures with increased retention time ^c
EF1	123456
EF2	123456; 23456; 126; 234; 12356; 12456; 23; 25; 12345; 2345; 235
EF3	123456; 345; 23456; 234; 12356; 12345; 2345
EF4	123456; 23456; 345; 234; 12456; 456; 13456; 12345; 2345

^a These data were taken from Reference (22).

^b Fragment being probed for interaction.

^c Numbers represent the peptides corresponding to the EF-hands of calbindin D_{28k}; e.g., 1 represents EF1, 2 represents EF2, etc.

1.3.1. Calcium-binding stoichiometry

Although calbindin D_{28k} is composed of 6 EF-hands, not all subdomains bind calcium. The calcium-binding stoichiometry as well as the identification of the EF-hands that bind calcium have been investigated. Calbindin D_{28k} was titrated with calcium and analyzed by ESI-MS (14, 23). This approach revealed 4 calcium ions bound to calbindin D_{28k}. The six synthetic EF-hand fragments described above were also probed for calcium-binding ability using agarose gel, NMR and circular dichroism (CD). It is assumed that if a fragment is capable of binding calcium, the corresponding EF-hands in the full-length protein should also bind calcium. Changes in CD and NMR signals and differences in migration rates across the agarose gels were observed in 5 of 6 fragments (the peptide representing EF2 did not bind calcium) (24). Heteronuclear NMR analysis of full-length rat calbindin D_{28k} has focused on the amide proton and nitrogen of the glycine residues (14). These signals are particularly sensitive to the presence of calcium. Changes in chemical shifts of approximately 2.1 and 15 parts per million (ppm) for ¹H

and ^{15}N , respectively of glycine residues in EF1, EF3, EF4, EF5 suggests that calcium binds to these subdomains. EF2 and EF6 did not exhibit such changes in chemical shifts. Recent NMR data on mutant human calbindin D_{28k} also show changes in the chemical shifts of glycine residues in the same four subdomains upon calcium binding (17). The experimental findings are consistent with the observation that EF2 and EF6 deviate from the calcium-binding loop consensus sequence (Figures 1.2 and 1.3). EF2 differs at the X and Y positions, and EF6 differs at the X, Y and Z positions.

1.4. Mass spectrometry in the structural characterization of proteins

Mass spectrometry (MS) is known as a powerful tool for accurate mass measurements of atoms and molecules. In the field of biochemistry and proteomics, mass spectrometry is now routinely used for the identification of a protein from cells or for the characterization of covalent protein modifications. MS has also gained recognition for its ability to provide insight into protein tertiary structures. A variety of approaches have been employed such as limited proteolysis where a protein is digested at its most accessible or surface-exposed cleavage sites. These sites are then identified by MS analysis. Berggård *et al.* employed this method to characterize calcium-loaded and calcium-free forms of calbindin D_{28k} (13). The most accessible trypsin cut sites were K59, K72, K98, K235, K236 and K245 for the calcium-loaded form and K59, K235 and R169 for the calcium-free form. Interestingly, the trypsin susceptible sites in the calcium-loaded protein are in EF2 and EF6 with the exception of K93, which is in the linker region between EF2 and EF3. EF2 and EF6 are also susceptible to proteolytic attack in the calcium-free form of the protein as well as EF6, which contains R169.

Thus, the subdomains that do not bind calcium appear to exhibit greater conformational flexibility in calbindin D_{28k} than the calcium-binding subdomains.

Resing *et al.* investigated the structural differences between wild-type MAP kinase kinase-1 (MKK1) and an activated mutant MKK1 (25). Briefly, the two forms of the protein were incubated in D₂O for a brief period of time allowing only the surface-exposed backbone hydrogens to exchange. The proteins were digested and injected into an HPLC-MS system. The peptides that exhibited increases in mass reflect areas of the protein with structural differences (25). Differences were observed in 6 peptides. Two peptides in the activated mutant protein showed greater H/D exchange rates than in wild type, and it was proposed that the activation of MKK1 is related to the structural flexibility of this region.

Other approaches take advantage of the nature of electrospray ionization (ESI) (14). Konnerman *et al.* used ESI to monitor the unfolding of lysozyme, cytochrome *c* and ubiquitin. The three proteins were unfolded under acidic conditions. When injected into the ESI source, the unfolded proteins had higher charge states than the native proteins, presumably because unfolding exposes basic amino acids that are buried in the native state (26). Also, due to its larger size, an unfolded protein can accommodate greater charge than a folded protein. Thus, ESI-MS can be used to monitor changes in surface accessibility of basic residues.

1.4.1. Chemical cross-linking as a tool for throughspace distance information

Intermolecular chemical cross-linking has been employed to study protein-protein interactions (27-30). Identification of the cross-links by MS can afford structural

information on proteins. A recent study revealed the possibility of protein structural characterization by combining sequence alignment, intramolecular chemical cross-linking and MALDI-ToF (matrix-assisted laser desorption ionization time-of-flight) MS (31). The model protein investigated in the published study was fibroblast growth factor (FGF), a protein with a known structure and a member of the β -trefoil family. Briefly, FGF was reacted with a primary amino specific bifunctional chemical cross-linker, bis(sulfosuccinimidyl) suberate (BS3), and digested with a proteolytic enzyme. The cross-linked peptides were identified by mass spectral analysis. Given that the bifunctional chemical cross-linker can only span a limited distance dictated by its spacer, the maximum distance separating the reacted lysine residues is obtained, and can be used as a distance constraint.

Concurrently, sequence alignment was employed using a database of proteins of known structure to generate a number of possible protein folds that FGF may adopt. Tertiary structures of FGF were proposed based on the fold identified by the alignment results. Apart from FGF-2 (itself) in Table 1.2, IL-1 β , the most homologous protein to FGF-2 with a β -trefoil fold, was ranked fifth out of the top twenty folds considered. The experimentally derived distance restraints were compared to the proposed models generated from the sequence alignment. Violations of the distance restraint implied that a proposed fold was incorrect. Through a process of elimination, the likely fold of FGF was identified as a β -trefoil fold. The tertiary structure of the protein was constructed based on the backbone of IL-1 β and following molecular dynamics simulation, the proposed structure was within 4 Å root-mean-squared-deviation (RMSD) of the crystal structure (31).

Since this study, a number of experimental refinements on the approach have been employed. These range from isotopically labeling the cross-linker to facilitate the identification of modified peptides (32), to automation by direct infusion of a modified protein into a mass spectrometer and tandem MS to pinpoint the specific modification (33).

Table 1.2. Top 20 sequence homology models for FGF

Name	Fold family	% Sequence identity	Threading rank ^a	Number of violations
FGF-2	β -Trefoil	98.6	1	0
IL-1 β	β -Trefoil	12.7	5	0
Gastrotropin	Lipocalin	7.1	8	1
Hisactophilin	β -Trefoil	8.6	12	2
Guanylate kinase	P-loop	12.4	9	4
NTP pyrophosphohydrolase	NTP pyrophosphohydrolase	9.3	6	3
Glutathione peroxidase	Thioredoxin	11.1	14	5
Retinol-binding protein	Lipocalin	9.1	18	3
Nucleoside diphosphokinase	Ferridoxin-like	8.8	20	2
Cytochrome <i>c</i> ₄	Cytochrome <i>c</i>	12.6	11	5
Aspartate carbamoyltransferase	Ferridoxin-like	9.8	13	4
D-UTPase	β -Clip	7.8	2	7
Disulfide bond formation protein	Thioredoxin	8.4	15	8
ASV integrase	Ribonuclease H-like	7.8	19	5
Endoglucanase C	Galactose binding	11.6	4	6
TATA-box-binding protein	TATA-box-binding protein-like	10.3	7	8
Phospholipase A2	Phospholipase A2	9.5	16	7
PRD paired domain	3-helix bundle	12.7	17	8

^a Threading rank is based on similarity between FGF and the proteins in the database. Note: the fourth and tenth threading ranks are not included in this table because cross-linked lysines were located in gaps or unresolved positions of the model protein.

1.5. Purpose of research

The purpose of the research reported in this thesis is two-fold. The first is to gain insight into the arrangement or the relative orientations of the EF-hand subdomains

within human calbindin D_{28k}. The second is to further evaluate the use of distance constraints for protein structure elucidation. The cross-linking approach is used to obtain throughspace distances between lysine residues of human calbindin D_{28k} in its calcium-loaded form. These distances are used to gain insight into the subdomain organization of the EF-hands of calbindin D_{28k}. Three-dimensional structures of the protein are proposed based on the experimental distance restraints and the modeled subdomains. These structures are then refined using molecular dynamics simulations. To date, the intramolecular chemical cross-linking approach has been applied only to model systems of known structure. Herein, the intramolecular cross-linking approach is applied to a protein for which there exist a number of structural studies but no known high-resolution three-dimensional structure. Currently, no structure is available for proteins composed of more than four EF-hands.

1.6. Outline of thesis and contributions of colleagues

This thesis is divided into three parts. The protein is not commercially available, so the development of an expression system and the purification of calbindin D_{28k} will first be described (Chapter 2). This work was performed in collaboration with Dr. Limei Tao (Concordia University). The second part describes chemical cross-linking and MS analysis of calbindin D_{28k} in order to obtain distance constraints (Chapter 3). The chemical cross-linking reaction and MS analysis was first applied to horse myoglobin as a model system. In the last part of the thesis, structures of calbindin D_{28k} based on the distance constraints are proposed (Chapter 4). This chapter will also deal with calcium

coordination sphere simulations as well as the refinement of a proposed calbindin D_{28k} structure.

2.0. Calbindin D_{28k} expression system

2.1. Introduction

Human calbindin D_{28k} is not commercially available and isolation of the protein from mammalian brain is inefficient (approximately 50 µg per 100 g of rat brain tissue) (34). Hence, to produce the recombinant protein an expression system had to be developed. To express recombinant protein, a host cell, typically yeast or bacterial, is transformed by the insertion of the gene coding for the protein of interest from any organism.

Two expression systems have been reported for recombinant human calbindin D_{28k} (35, 36). In one of these systems, the calbindin D_{28k} gene was cloned into a pET3a vector, which was then used to transform a BL21 *Escherichia coli* (*E. Coli*) cell line. The purification protocol involved a heat precipitation step to remove unwanted proteins followed by two rounds of ion-exchange chromatography. This system produced 20-50 mg of pure protein per litre culture.

Murphy and coworkers from the University of British Columbia (UBC) cloned the calbindin D_{28k} gene into a pGYMX vector (36). This vector, which is derived from the pGYM vector (2), contains a gene coding for a hexa-histidine tag to facilitate purification by metal affinity chromatography (MAC). Following ligation of the cDNA for calbindin D_{28k} into pGYMX the expressed fusion-protein was comprised of calbindin D_{28k} and the N-terminal tag, **MATSH₆IEGRAS**. This construct will henceforth be referred to as His-calb. The N-terminal tag includes the translation-initiating methionine residue, **M**. Following the **M** residue is a series of six consecutive histidines, which is

commonly referred to as a His-tag. This tag interacts strongly with immobilized metal ions such as Ni^{2+} via its imidazole sidechains, making MAC an effective purification method. **IEGR** is a recognition sequence for the protease factor Xa allowing for removal of the His-tag following MAC purification. The bases encoding residues **AS** correspond to the *NheI* restriction site, which was used in the engineering of pGYMX. These are two N-terminal residues of recombinant protein although they are not part of the native protein. When BL21 cells are transformed with pGYMX, the expression system is unstable since the plasmid appears to be expelled from the bacteria (36). The calbindin $\text{D}_{28\text{k}}$ cDNA along with the His-tag were subcloned into a second plasmid, pTrc99A. This plasmid was then used to transform JM-105, the host cell for pTrc99A (37). Using the latter system, expression and purification of the fusion protein (15 mg per litre culture, personal communication) was achieved.

Unfortunately, pTrc99A-His-calb is no longer available at UBC. Here the transformed JM05-pTrc99A-His-calb expression system was recreated starting from the pGYMX-His-calb plasmid, which was obtained from UBC.

2.2. Material and methods

2.2.1. JM105-pTrc99A-His-calb system

Vector digestion

The pGYMX-His-calb plasmid was provided by Prof. Michael Murphy (UBC). The plasmid was used to transform a $\text{DH5}\alpha$ *E. coli* cell line by electroporation. The cells were streaked on a LB plate containing 100 $\mu\text{g/mL}$ ampicillin (this amount of ampicillin was used throughout) (38). A single colony was picked from the plate and used to

prepare a starter culture in 3 mL of LB/amp medium (38). This culture was incubated at 37°C for 8 h with shaking at 250 rpm. One mL of this culture was used to inoculate 250 mL of LB/amp medium. The cells were grown at 37°C for 8 h with shaking at 300 rpm. The plasmid was then isolated using a Qiagen QIAquick PCR purification kit as per the manufacturer's instructions (39 1999). Quantitation of DNA was performed assuming A_{260} 50 µg/mL/cm for double stranded DNA (40). The pTrc99A vector was provided by Prof. Susan Aitken (Carleton University). The relevant features of the pTrc99A and pGYMX plasmids are shown in Figure 2.1.

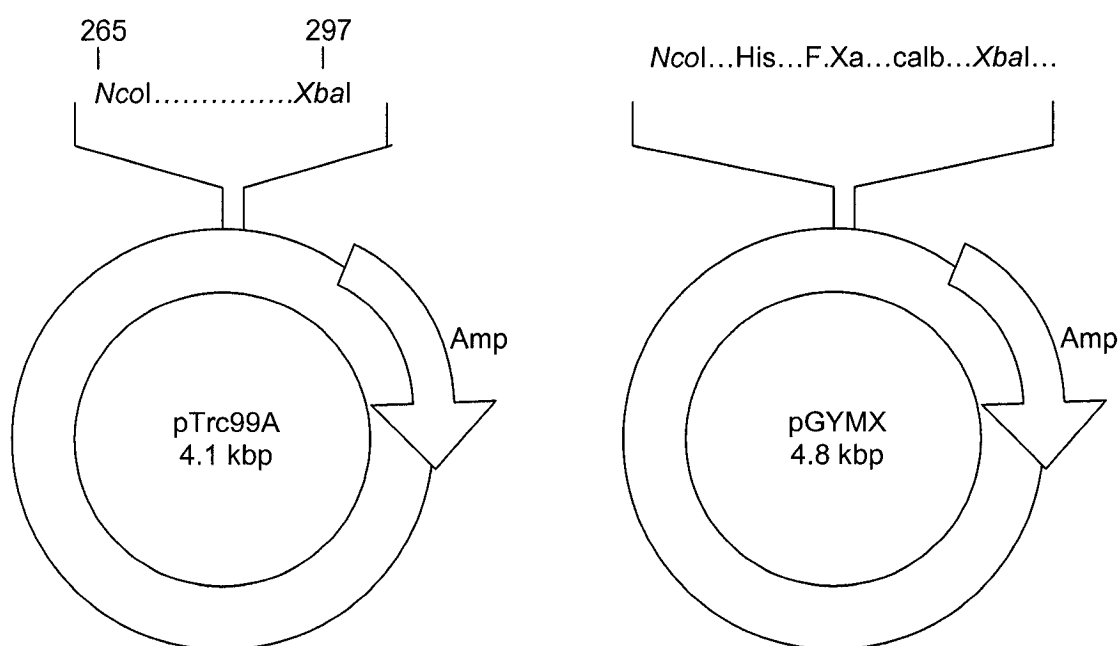


Figure 2.1. Plasmid maps of pTrc99A and pGYMX-His-calb. Data from References (2, 36, 37).

Both pTrc99A and pGYMX-His-calb contain *NcoI* and *XbaI* recognition sites. These two restriction enzymes were used to create the sticky ends for the ligation of pTrc99A vector with the insert. pTrc99A (30 µg) was digested with 1 µL of *NcoI* (10 units/µL, MBI Fermentas) and 1 µL of *XbaI* (10 units/µL, MBI Fermentas) in a total

volume of 100 μL Y^+ / Tango buffer (MBI Fermentas). pGYMX-His-calb (45 μg) was digested with 1 μL of *Nco*I and 1 μL of *Xba*I in a total volume of 115 μL Y^+ / Tango buffer (MBI Fermentas). The solutions were incubated at 37°C for 16 h. The restriction enzymes were thermally deactivated by heating the solutions at 65°C for 20 min. To verify that pGYMX-His-calb and pTrc99A were digested and to isolate the His-calb gene, 1 μL of the digest products was analyzed by gel electrophoresis.

Analysis of DNA by agarose gel electrophoresis

The procedure used was adapted from Reference (40). Agarose gel (1%) was prepared by mixing 0.5 g of agarose in 50 mL 1X Tris Borate EDTA (TBE) buffer (Appendix A). The solution was heated (but not boiled) using a microwave until the agarose dissolved, and was allowed to cool to approximately 60°C before being poured into a casting tray. A comb was immediately inserted into the agar solution. The comb was removed when the agarose cooled to room temperature and polymerized. Samples previously mixed with Loading Dye solution (MBI Fermentas, SM0321) were loaded in the wells and 100 V was applied through the gel for approximately 1 h. The gel was stained with 10 μL of ethidium bromide in 100 mL of 1X TBE buffer for 15 min.

Purification of the calbindin D_{28k} gene and the pTrc99A vector

Isolation of the calbindin D_{28k} insert from the pGYMX-His-calb digest was achieved by excising the gel band containing the insert. The DNA was extracted from the gel using a QIAquick gel extraction kit as per the manufacturer's instruction (39 1999). Digested pTrc99A was purified using a Qiagen QIAquick PCR purification kit

and to prevent religation, the vector was dephosphorylated by alkaline dephosphatase (2 μ L, 1 unit/ μ L) in 5 μ L of the supplied dephosphorylation buffer on incubation at 37°C for 1 h. Dephosphorylated pTrc99A was then purified with the QIAquick PCR kit.

Ligation of the calbindin D_{28k} cDNA into the pTrc99A vector

Dephosphorylated pTrc99A (60 ng) and 60 ng of the purified calbindin insert (1:5 vector to insert molar ratio) were ligated by T₄ DNA ligase (1 μ L, 5 units/ μ L) in 20 μ L of the supplied buffer (Roche, 1 635 379) on incubation for 10 min at 22°C. As recommended by Roche, inactivation of the ligase was not performed.

Transformation of JM105 by pTrc99A-His-calb

JM105 (a gift from Prof. Paul Joyce, Concordia University), the recommended host cell for pTrc99A (37 1998) was made chemically competent by the Inoue method (40). Heat shock was performed to transform JM105 with the pTrc99A-His-calb vector. JM105 was first chilled on ice for 30 min. One μ L of ligation product was added to 100 μ L of JM105 and immersed in a 42°C bath for 30 s. The sample was placed on ice for 2 min, 1 mL SOC medium (38) added and the solution was shaken at 37°C for 45 min. The cells were grown on a LB/amp plate for 16 h at 37 °C. Because the gene conferring ampicillin resistance is present in pTrc99A, ampicillin selection does not distinguish between pTrc99A and pTrc99A-His-calb ligation products. To confirm the presence of the insert, twelve colonies were picked from the plate and each was used to inoculate 3 mL of LB/amp media. Plasmids from the bacterial culture were isolated using the

QIAquick gel kit and digested with the *Nco*I and *Xba*I restriction enzymes as outlined previously. Finally, the products were analyzed by agarose gel electrophoresis.

2.2.2. Protein purification

Isolation of the His-calb fusion protein

JM105 cells transformed with pTrc99A-His-calb were screened on a LB/amp plate. A colony was used to inoculate 50 mL of LB/amp. The cells were grown at 30°C for 14 h. Ten mL of the culture was added to 1 L of LB/Amp medium and grown to an OD₆₀₀ of 0.5. Protein expression was induced with 1 mL of 1 M isopropyl-β-D-thiogalactopyranoside (IPTG, Bioshop IPT001) over a period of 14 – 16 h with shaking at 30°C. The culture was centrifuged at 6000 g at 4°C for 20 min and the supernatant discarded. The cell pellet (5 g) was collected and resuspended in 25 mL of BugBuster (Novagen) cell lysis solution. Benzonase Nuclease (25 μL, Novagen, 70746-3) and half a tablet of protease inhibitor cocktail (Roche) was added to the suspension. The nuclease reduces the viscosity of the solution, which facilitates manipulation and the protease inhibitor cocktail prevents digestion of the His-calb by endogenous proteases. The suspended cells in the mixture were lysed using a Dounce homogenizer. The mixture was incubated at 22°C for 20 min with shaking at 10 rpm on a Thermolyne Vari Mix shaker platform, and centrifuged at 16000 g at 4°C for 20 min. The supernatant was collected, mixed with 0.8 mL of Ni-NTA resin (bed volume) in a 2 x 20 cm fritted column. Five mM imidazole pH 8.0 buffer was added in order to decrease unspecific protein binding to the beads (note, all imidazole buffers mentioned herein also contained 20 mM sodium phosphate, 0.5 M NaCl, 5 % glycerol and 10 mM β-mercaptoethanol).

The His-calb was immobilized on the Ni-NTA beads by rotating the mixture for 1 h at 4 °C.

The Ni-NTA column was washed with 100 mL of 5 mM imidazole buffer followed by the following step gradients: 10 mL of 25 mM imidazole pH 8.0; 10 mL of 50 mM imidazole, pH 8.0; 4 mL of 300 mM imidazole, pH 6.0 and 6 mL of 500 mM imidazole pH 6.0. The His-calb was observed to elute in 300 mM imidazole. β -mercaptoethanol (10 mM) was added in all imidazole buffers to prevent disulfide bridge formation. The 300 mM imidazole eluate (4 mL) was added to a 12 000 – 14 000 molecular weight cutoff dialysis bag (Spectra/Por, 132 676) and dialyzed against 1 L of 20 mM TrisHCl, pH 7.4 with 0.1 mM CaCl_2 over 24 h with 3 rounds of buffer exchange. Identification of the His-calb protein was confirmed by ESI-MS.

Removal of the His-tag from recombinant calbindin D_{28k}

The factor Xa recognition site on the fusion protein is located between the His-tag and the N-terminal of calbindin D_{28k}. Thus, the His-tag can be removed by digesting the fusion protein with factor Xa (Novagen, 69037-3). The digestion reaction was carried out in 20 mM TrisHCl pH 7.4, 100 mM NaCl at 20°C for 16 h at a factor Xa:calbindin ratio of 1:250 (w:w).

Following His-tag removal, the calbindin D_{28k} sample was further purified on the Ni-NTA column. The sample was mixed with 1 mL of the Ni-NTA beads by rotating the mixture for 1 h at 4°C. Because the recombinant protein no longer bound to the column, it was collected in the eluate. The eluted calbindin D_{28k} was dialysed three times against 20 mM TrisHCl, pH7.4, 1 mM DTT and 0.1 mM CaCl_2 over a period of 24 h and stored

at -80°C . A 500 mM imidazole wash was used to recover species that remained bound to the column. The eluates were analysed by sodium dodecylsulfate-polyacrylamide gel (10%) electrophoresis (SDS-PAGE).

SDS-PAGE analysis

The procedure was adapted from Reference (40). Samples were prepared by mixing 1 part sample with 1 part sample buffer (v:v) (Appendix A) in 1.5 mL centrifuge tubes. The tubes were heated at 95°C for 5 min, and allowed to cool to room temperature before loading the samples on gel.

The resolving gel was prepared by mixing the resolving gel components (Appendix A). Ammonium persulfate (APS) promotes the polymerization and was the last component to be added. The resolving gel solution was poured between two glass plates separated by spacers until the solution level was approximately 2 cm from the top of the plates. A 2-mm layer of water saturated N-butanol was poured over the resolving gel solution in order to straighten the level of the resolving gel. The gel was allowed to polymerize for approximately 30 min and the upper liquid layer that contained N-butanol, water and unpolymerized resolving gel solution was removed. Stacking gel solution (Appendix A) was used to fill the remaining glass plate assembly. A comb was immediately inserted and the solution was allowed to polymerize. The glass assembly was placed into a SDS-gel tray, which was then filled with 1X glycine buffer (Appendix A). The comb was removed and the samples were loaded into the wells. A current of 30 mA was applied across the gel until the loading buffer dye reached 1 cm from the bottom of the gel. The gel was removed from the glass plate assembly, washed with H_2O for 15

min on a shaker platform, stained with Gelcode Blue stain reagent (Pierce, 24592) for 1 h and destained with H₂O.

2.3. Results

2.3.1. JM105-pTrc99A-His-calb expression system

Digestion of the plasmids was visualized by agarose gel electrophoresis (Figure 2.2). The digestion of pTrc99A (4.1 kbp) by *Nco*I and *Xba*I results in a loss of a 32 bp fragment. The small fragment is not observed because its fluorescence intensity was too weak and/or because it ran off the gel. The digestion product, therefore, appears as a single band on the gel (Lane 2). Digestion of pGYMX-His-calb results in the appearance of two fragments (Lane 3). The smaller band corresponds to the calbindin gene, which has a size of 846 bp.

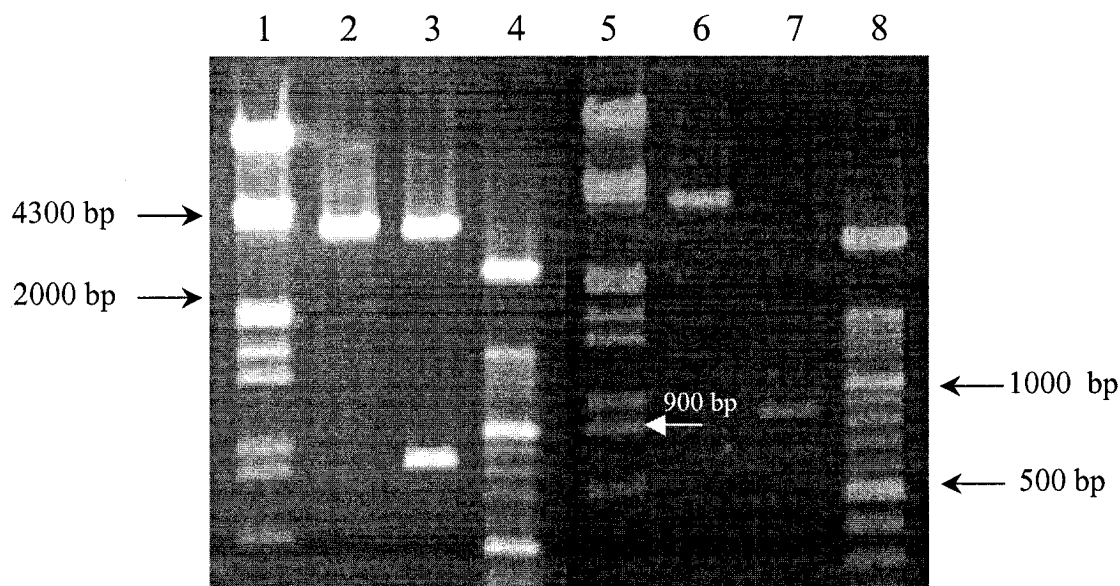


Figure 2.2. Restriction analysis of pTrc99A and pGYMX-His-calb. (2) *Xba*I and *Nco*I digested pTrc99A. (3) *Xba*I and *Nco*I digested pGYMX-His-calb. (6) Purified, dephosphorylated pTrc99A. (7) Purified His-calb insert. Lanes 1 and 5 contain DNA Molecular Weight Marker III (Roche) while lanes 4 and 8 DNA Molecular Weight Marker IV (Roche).

The dephosphorylation of the pTrc99A plasmid was performed to reduce the likelihood of pTrc99A self-ligation. Following dephosphorylation and purification, the plasmid was run on an agarose gel (Figure 2.2, Lane 6) and quantitated based on the intensity of the 1000 bp marker (Figure 2.2, Lane 8) which contains 113 ng of DNA. Purification of the His-calb gene insert was achieved by excising the gel band at 846 bp in Lane 3 and extracting the His-calb DNA. The His-calb insert was also quantitated on the agarose gel (Lane 7) by comparison with the 900 bp marker (Figure 2.2, Lane 5) which contains 19 ng of DNA.

The dephosphorylated pTrc99A vector and the His-calb insert were ligated and the resulting vector was used to transform JM105. The transformed cells were streaked on LB/agar plate containing ampicillin. Bacterial colonies were observed indicating that the transformation was successful. Twelve colonies were picked from the plate following overnight growth (Section 2.2.1), and their plasmids were isolated and doubly digested with the *NcoI* and *XbaI*. Analysis of the fragments indicates that all colonies selected

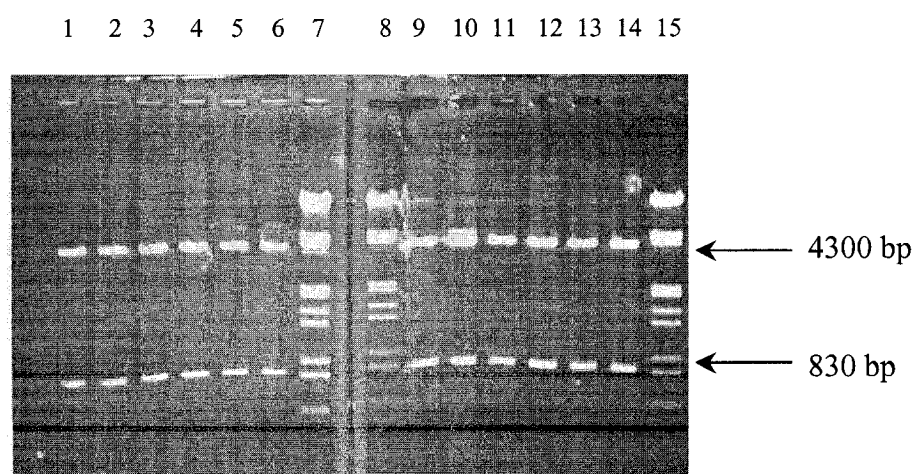


Figure 2.3. Restriction analysis of the plasmids isolated from JM105 cells transformed with pTrc99A-His-calb. Lanes 1 - 6 and 9 - 14 contain *NcoI* and *XbaI* digested pTrc99A-His-calb. Lanes 7, 8 and 15 contain the DNA Molecular Weight Marker III (Roche).

contained the pTrc99A-His-calb plasmid (Figure 2.3). Clearly, the 5-fold molar excess of His-calb insert used favored the desired ligation product.

2.3.2. Expression and purification of recombinant human calbindin D_{28k}

Approximately 4.2 g of cell paste was recovered from a 1 L culture of transformed JM105 cells centrifuged at 6000 g for 20 min at 4°C. The cells were mixed with BugBuster protein extraction reagent, lysed using a homogenizer and centrifuged at 16000 g for 20 min at 4°C. The supernatant was added to a 2 x 20 cm Ni-NTA column and eluted with a step-gradient of imidazole. The induction of calbindin D_{28k} is demonstrated by comparing Lane 1 with Lanes 2 and 3 (Figure 2.4). The sample in Lane 1 was not induced by IPTG and shows only a weak band corresponding to His-calb. The presence of His-calb in the supernatant (Lane 4) reveals that the recombinant tagged protein is water-soluble.

Purification of the His-calb fusion protein involved a Ni-NTA column. Impurities with no Ni-NTA affinity were allowed to flow through (Lane 5). Impurities with some affinity for column were eluted using imidazole step gradients, since free imidazole competes with the histidine side chain for Ni-NTA (Lanes 6 to 8). His-calb is observed to elute at 300 mM imidazole along with some impurities (Lanes 10 and 11). At 500 mM imidazole very little protein is eluted (Lane 12) indicating that 300 mM is sufficient to release most protein from the column.

The 300 mM imidazole eluate (Lanes 10 and 11) was analyzed by ESI-MS (Figure 2.5). The deconvoluted mass spectrum shows an intense peak at 31721Da. This is consistent with the theoretical mass of His-calb minus the initiating methionine

residue, thus confirming the identity of the wide band at approximately 30 kDa in Lanes 10 and 11 of Figure 2.4 as His-calb.

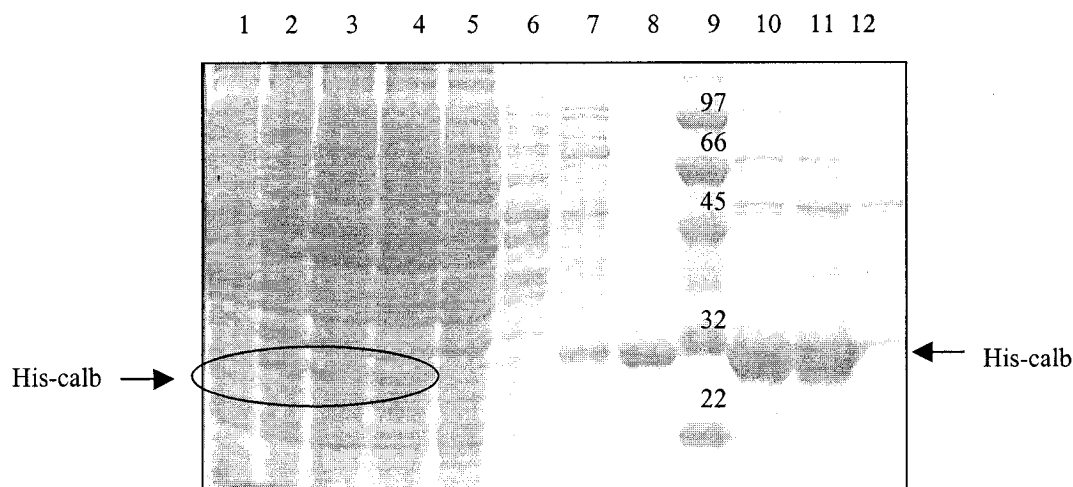


Figure 2.4. 10% SDS-PAGE analysis of steps involved in His-calb purification on the Ni-NTA column. Lane 1, whole cell lysate (no IPTG added); Lanes 2 and 3, whole cell lysate (with IPTG); Lane 4, supernatant of IPTG induced cell lysate; Lane 5, flow through from the Ni-NTA column; Lanes 6, 7, 8 and 12 contain 5 mM, 25 mM, 50 mM and 500 mM imidazole eluates, respectively. Lanes 10 and 11 both contain the 300 mM imidazole eluates. Lane 9 contains the protein molecular weight markers. The imidazole elution buffers contain 20 mM sodium phosphate, 0.5 M NaCl, 5 % glycerol and 10 mM β -mercaptoethanol. β -Mercaptoethanol was also present in to the sample buffer.

The His-tag was removed from the protein using factor Xa. The digest was passed through the Ni-NTA column and His-tag-free calbindin D_{28k} was not retained on the column (Figure 2.6, Lanes 2 and 3). To verify that factor Xa efficiently cleaved the tag, 500 mM imidazole was run through the column. SDS-PAGE analysis of the eluate shows a band corresponding to His-calb (Lane 4), indicating that not all the tag was cleaved by factor Xa. Lane 3 contained calbindin D_{28k} under non-reducing conditions. There is no evidence of intermolecular disulfide bridge. The eluate from Lanes 2 and 3 were collected and analyzed by ESI-MS, which confirmed that the His-tag was removed since the mass of the recombinant protein decreased by 1539 Da (Figure 2.7).

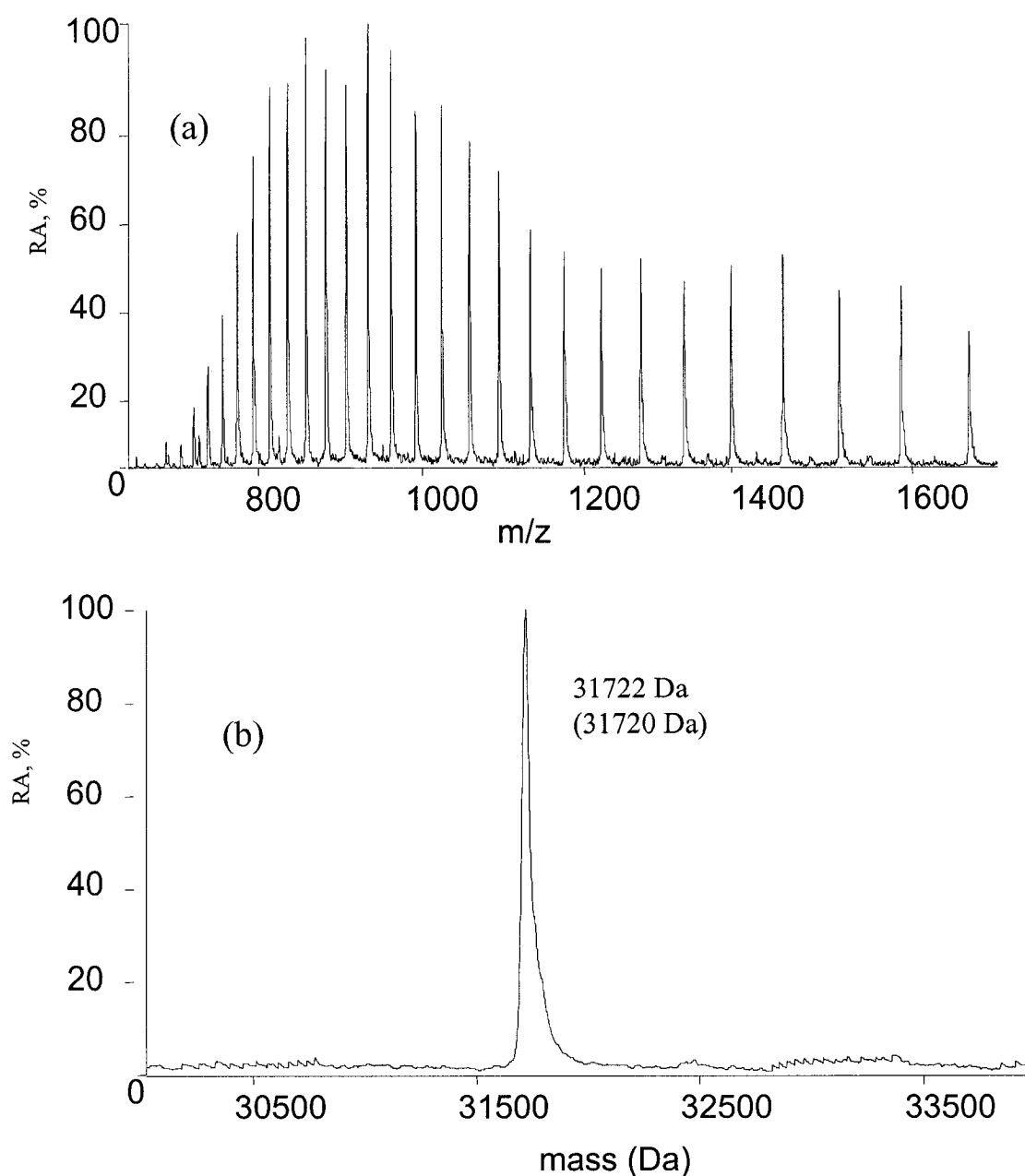


Figure 2.5. MS analysis of purified His-calb. (a) ESI spectrum of purified His-calb corresponding to the sample in Lanes 10 and 11 of Figure 2.4. (b) Deconvoluted spectrum of (a). His-calb (1 μ M) in 50% ACN/50% H₂O/0.2% FA was directly infused into a Waters Micromass Q-ToF MS with a flow rate of 1 μ L/min. The voltage settings of the mass spectrometer were: capillary: 3.8 kV, cone: 45 V, multiplier: 550 V, MCP: 2.1 kV. RA: relative abundance. The observed and theoretical mass (in parentheses) of His-calb are given. The theoretical mass was calculated by adding 1714 Da for the mass of the MATSH₆IEGRAS tag to the mass of human calbindin D_{28k} based on its published sequence (6, 7).

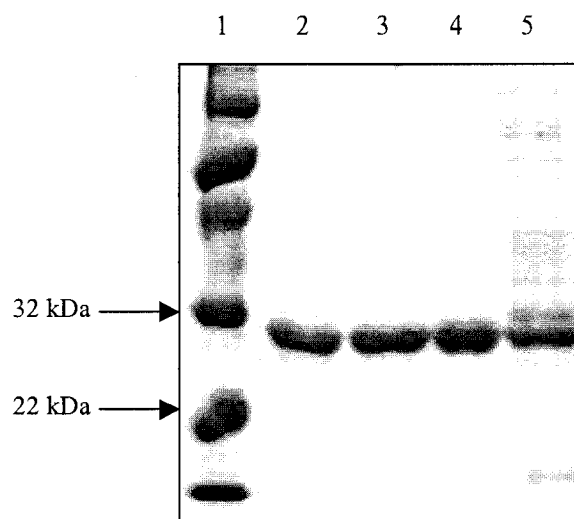


Figure 2.6. SDS-PAGE analysis of His-tag cleavage in His-calb. The cleavage was carried out in 20 mM TrisHCl pH 7.4, 100 mM NaCl at 20°C for 16 h with a factor Xa:calbindin ratio of 1:250 (w:w). Lane 1, protein molecular weight markers; Lanes 2 and 3 contain the flow through of a His-calb solution treated with factor Xa. Lane 4 is the same as Lanes 2 and 3 but under non-reducing conditions (no β -mercaptoethanol in the SDS-PAGE sample buffer). Lane 5 contains the 500 mM imidazole eluate.

2.3.3. Analysis of thiol reactivity in His-calb vs. untagged calbindin D_{28k}

Calbindin D_{28k} contains 5 free cysteine residues. These are of particular interest because they have been shown to be S-nitrosated by S-nitrosoglutathione and S-nitrosocysteine (41). The reactivity of the thiol groups of His-calb has been evaluated previously by titration with N-ethylmaleimide (NEM) and 2,2'-dithiobis-5,5'-nitrobenzoic acid (DTNB) (36). The reduction of DTNB by free thiol produces 5-thio-2-nitrobenzoate, which can be quantitated spectrophotometrically. In the present study, the reactivities of the thiols in recombinant calbindin D_{28k} and His-calb were compared.

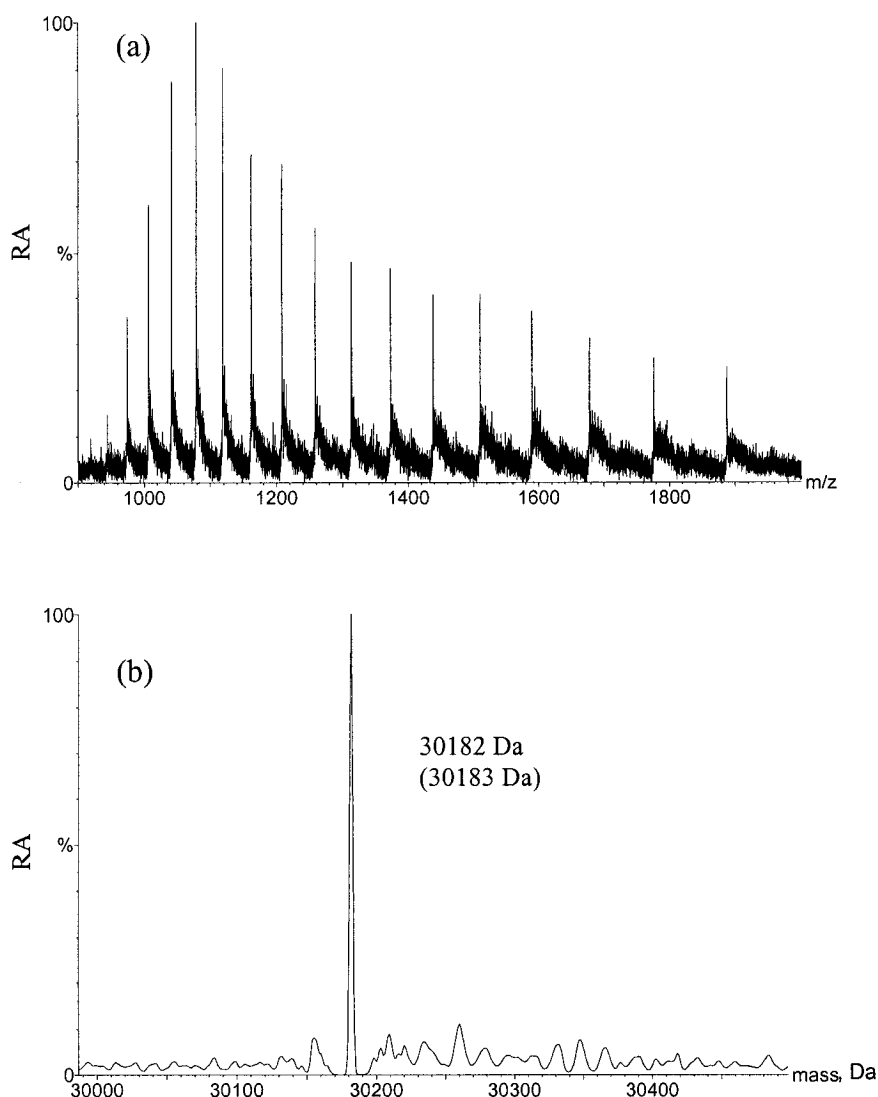


Figure 2.7. MS analysis of purified recombinant calbindin D_{28k}. (a) ESI mass spectrum of 1 μ M calbindin D_{28k} (from Figure 26, Lanes 2 and 3). (b) Deconvoluted spectrum of (a). Calbindin D_{28k} in 50% ACN/50% H₂O/0.2% FA was directly infused into a Waters Micromass Q-ToF MS at a flow rate of 1 μ L/min. The voltage settings of the mass spectrometer were: capillary: 3.8 kV, cone: 45 V, multiplier: 550 V, MCP: 2.1 kV. RA: relative abundance. The observed and theoretical mass (in parentheses) of recombinant calbindin D_{28k} is given. The theoretical mass was calculated by adding 157 Da for the mass of the additional AS N-terminal residues to the mass of human calbindin D_{28k} calculated based on its published sequence (6, 7).

Calbindin D_{28k} and His-calb were exchanged into 100 mM TrisHCl (pH 8.3) with 0.5 M EGTA using centrifugal ultrafiltration with 10 kDa molecular weight cutoff filters (Millipore). A stock solution of 4 mM NEM was prepared fresh and a portion was added

to a solution of 13 μM calbindin $\text{D}_{28\text{k}}$ or His-calb to give a NEM to cysteine molar ratio of 10:1. The reaction was allowed to proceed for 30 min. Unreacted NEM was removed by washing the sample with water in a 10 kDa molecular weight cutoff centrifugal filter. ESI-MS was used to characterize the products.

Calbindin $\text{D}_{28\text{k}}$ and His-calb showed significant differences in their reactivity towards NEM (Figure 2.8). Almost all cysteine sites of His-calb are modified, consistent with the

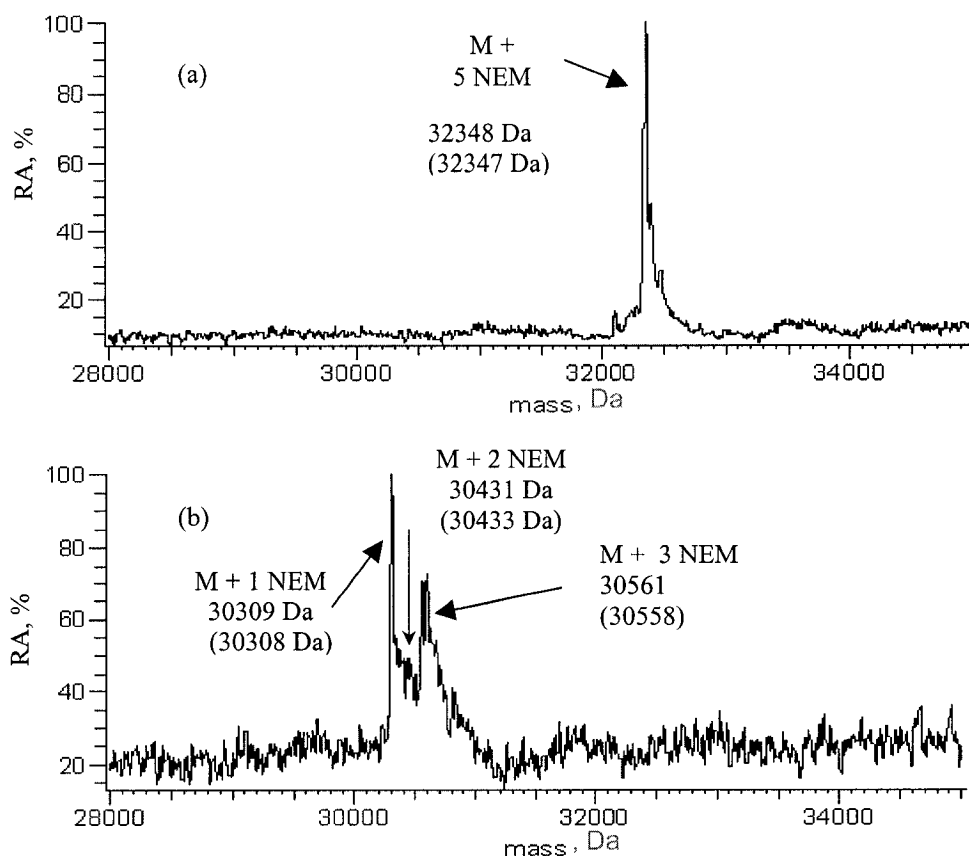


Figure 2.8. Deconvoluted ESI-MS spectra of NEM-modified (a) His-calb and (b) calbindin $\text{D}_{28\text{k}}$. The NEM-modified proteins were prepared as described in Section 2.3.3. NEM-modified His-calb and calbindin $\text{D}_{28\text{k}}$ (10 μM) in 50 % ACN/50% H_2O /0.1 % TFA were directly infused at a flow rate of 5 $\mu\text{L}/\text{min}$ into the ESI source of a ThermoFinnigan SSQ7000 MS. The source and capillary temperatures were set at 70°C and 250°C, respectively. Nitrogen sheath gas was applied at 35 psi, and ion scans were performed in the positive-ion mode between m/z 600 to 1500 at a scan time of 3 s. The observed and theoretical (in parentheses) masses are given on the spectra. The theoretical masses were calculated by adding 125 Da per NEM label to the calculated mass of the proteins (31722 Da and 30183 Da for His-calb and calbindin $\text{D}_{28\text{k}}$, respectively).

findings of Tao *et al.* (36). However, for calbindin D_{28k}, the predominant species is the singly NEM modified protein. The difference in reactivity between His-calb and calbindin D_{28k} is consistent with the number of free cysteines as determined by DTNB titration. DTNB titration revealed 4.2 reactive cysteines in His-calb and only 1.4 in calbindin D_{28k} (Limei Tao, personal communication).

The reasons for the differences in the thiol reactivity of His-calb and calbindin D_{28k} are not clear. It is conceivable that calbindin D_{28k} may have formed intramolecular disulfide bridges in the process of His-tag removal. However, SDS-PAGE analysis of the protein under non-reducing conditions shows that there were little or no intermolecular disulfide bridges formed in the samples (Figure 2.6, Lane 4).

2.4. Discussion

The reconstruction of the pTrc99A-His-calb construct from pGYMX-His-calb was described. In addition, the MAC purification method was modified from the original protocol. The published protocol employed by Tao *et al.* made use of a single MAC purification step (36). The fusion protein was not treated with factor Xa so the His-tag remained tethered to the protein. As shown in Section 2.3.3, His-calb and untagged calbindin D_{28k} show differences in their reactivity towards NEM and DTNB. It is not clear if the two forms are structurally different due to the presence of the N-terminal tag in His-calb or, as is more likely, the cysteine residues have been oxidized during removal of the His-tag (Limei Tao, personal communication).

Removal of the His-tag yields a calbindin D_{28k} product with less impurity than His-calb. The yield is 12-15 mg of purified His-calb per litre of culture with

approximately 84% purity based on SDS-PAGE analysis. Following cleavage and removal of the tag, the yield is approximately 9 mg/L with a purity of 93%. The removal of the tag directly accounts for a decrease of approximately 0.8 g/L. Also, as shown in the SDS-PAGE gel of Figure 2.6 (Lane 4), factor Xa does not completely remove the His-tag and some uncleaved His-calb remains on the Ni-NTA column.

With the pTrc99A-His-calb construct a dialysis step was necessary to remove imidazole, which is known to inhibit factor Xa (42 Factor Xa kits (2001)). Moreover, the manipulations were performed in the absence of reducing agent. Tao evaluated the propensity of calbindin D_{28k} to form disulfide bridges. Solutions of His-calb and calbindin D_{28k} were left at room temperature for 24 h. When analyzed by SDS-PAGE, there was no evidence for dimerization (Limei Tao, personal communication).

There is interest in our lab in further studying wild-type and mutant forms of calbindin D_{28k}. Hence a more efficient expression system was desired. Recently, Tao *et al.* engineered a pET15b-His-calb(GSH) plasmid (GSH refers to the additional glycine-serine-histidine sequence present at the N-terminal of the recombinant protein), which was used to transform BL21(DE3)pLysS *E. coli* cells (43). This system produced approximately 34 mg of fusion protein per litre of culture with approximately 95% purity. The two expression systems are compared in Table 2.1. In addition to the increased yield, the pET15b plasmid contains a thrombin rather than a factor Xa cut site between the His-tag and the N-terminal of calbindin D_{28k}. The advantage of a thrombin cut site is that this protease remains active in the presence of reductants such as β -mercaptoethanol and in the presence of imidazole. Thus, His-tagged calbindin D_{28k} can be cleaved directly after MAC purification without the need for a dialysis step. Also, since thrombin can be

used in the presence of β -mercaptoethanol, there is less concern that the cysteine residues of calbindin D_{28k} may become oxidized. This is important because disulfide bridge formation is prevented under the reducing conditions found in the cytosol (19).

Although the pET15b-His-calb(GSH) system is more efficient, it was not available at the time that the cross-linking experiments were to be performed. The calbindin D_{28k} used in the cross-linking study was produced using the pTrc99A-His-calb system, and the His-tag was cleaved before use. As described in Chapter 3, the techniques used to characterize calbindin D_{28k} cross-linked with a bifunctional reagent are SDS-PAGE, ESI-MS and MALDI-MS. These techniques are sensitive and require only microgram amounts of sample. The 9 mg yield of pure recombinant calbindin D_{28k} per litre of cell culture is sufficient for the intended purposes.

Table 2.1. Recombinant calbindin D_{28k} expression systems ^a

Vector	Host cell	Tag cut site	Purity	Yield (mg/L culture)
pTrc99A-His-calb	JM105	factor Xa	93%	9
pET15b-His-calb(GSH)	BL21(DE3)pLysS	thrombin	95%	34

^a This table is adapted from Reference (43).

3.0. Reactions of myoglobin and calbindin D_{28k} with the homocross-linker bis(sulfosuccinimidyl) suberate

3.1. Introduction

Protein cross-linking involves the linkage of two reactive sidechain atoms through a covalent bond. Reactive groups in naturally occurring amino acid sidechains include the carboxylate group of aspartate and glutamate, the amino group of lysine and the thiol group of cysteine. Cross-linkers that are specific to the above mentioned amino acids are commercially available. When choosing a cross-linker for structural studies, characteristics that need to be considered are its arm length, solubility, specificity, functionality (homo or hetero) and the abundance of the amino acid(s) of interest in the protein.

Calbindin D_{28k} is a water-soluble protein containing 5 cysteine, 25 lysine, 23 aspartic and 25 glutamic acid residues (the full primary structure and amino acid composition of calbindin D_{28k} can be found in Appendix B). 4-(*p*-Azidosalicylamido)butylamine is an example of a commercially available chemical cross-linker (Pierce) that reacts with both aspartic and glutamic residues. Thus, there are 48 carboxylate groups within the calbindin D_{28k} that can react with this cross-linker. This translates into 1128 cross-links that can potentially form. A large distribution of possible products would lead to low yields of specific species, which would likely complicate the detection of any cross-linked peptides. Also, with increasing numbers of possible modifications, the probability that a particular *m/z* can be uniquely assigned decreases. In contrast, given the low number (five) of cysteine residues in calbindin D_{28k}, there is a

risk that too few cross-links would form. More importantly, none of the cysteine residues appear in EF1 or EF2 (rows 1 and 2 in Figure 1.2). Thus cross-linking experiments targeting cysteine residues will not give any information on the spatial position of these two subdomains.

Calbindin D_{28k} contains 25 lysine residues distributed throughout the primary structure (Figure 1.2). BS3 (Figure 3.1) is a homobifunctional cross-linker that specifically targets primary amino groups. It can react with lysine sidechains as well as the N-terminal amino group of the protein backbone. The sulfate groups in the cross-linker are incorporated to increase its water solubility. The linker arm is approximately 11 Å when the hydrocarbon chain is fully extended. The lysine sidechain is flexible and dynamic, with a span of 6 Å (from C_α to N_ε) when fully extended. Since a lysine sidechain can rotate around its C_α, it is more informative to consider the distance separating the C_α atoms. The separation between the C_α atoms of two lysine residues involved in a BS3 cross-link can be no further than 24 Å. This is the distance constraint

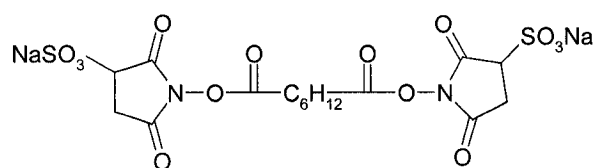


Figure 3.1. Stick diagram of bis(sulfosuccinimidyl) suberate (BS3). The separation between the carbon atoms of the ester groups is 11 Å.

that was used when constructing the protein models. BS3 was chosen for the chemical cross-linking experiments described here. The characterization of the intact protein product from BS3 reactions was performed by SDS-PAGE and ESI-MS analysis. SDS-PAGE is a rapid method of detecting any unwanted intermolecular cross-links since these

would appear as multimers at higher masses on the gel. ESI-MS is a powerful tool for accurate mass determination of large proteins and can give a detailed picture of the reaction products that are formed. The mass increase due to intramolecular cross-linking of a protein by BS3 is 138 Da, which can be easily resolved by ESI-MS but not by SDS-PAGE.

To identify the specific residues that are cross-linked, the modified protein is first digested using proteolytic enzymes. The two enzymes used in this work are trypsin and Glu-C. Trypsin cleaves at the carboxyl terminal of lysine and arginine residues, while Glu-C in phosphate buffer cleaves at the carboxyl terminal of glutamate and aspartate residues. The resulting peptide mixtures can be analyzed by ESI-MS. However, because of multiple charging of the peptides the mass spectrum of a protein digest is complex. A separation procedure, typically high performance liquid chromatography (HPLC), is often used in the analysis of digests. MALDI-MS is ideally suited for the analysis of mixtures since ionization by laser desorption predominantly produces singly charged ions. The spectra obtained from mixtures are thus greatly simplified. Moreover, with the advent of time-lag-focusing and reflectron time-of-flight mass analysers, resolution obtained using MALDI-MS is on par with that using ESI-MS.

The objective is to apply MALDI-MS analysis of intramolecular cross-linked calbindin D_{28k} to obtain structural information in the form of distance constraints. The feasibility of this approach is first evaluated in this work using myoglobin as a model protein. The crystal structure of horse heart myoglobin is known at high resolution (PDB ID, 1dwr) (44), it contains 19 lysine residues and is 17 kDa in size, approximately half the size of calbindin D_{28k}. The sequence and amino acid composition of myoglobin are

given in Appendix B. Because of the smaller number of potential modification sites and proteolytic cleavage sites in myoglobin, the protein digests should be less complex, which should simplify the data interpretation.

3.2. Materials and methods

3.2.1. Materials

Horse heart myoglobin (M-1882), ACTH(A0673), renin (R-8129), angiotensin (A-9650), endoproteinase Glu-C (P6181-1VL) and dithiothreitol (DTT, D-9163) were from Sigma. Trifluoroacetic acid (TFA, 200-929-3), 4-hydroxy- α -cyanocinnamic-acid (HCCA, 47,657-0), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES, 23,388-9) were from Aldrich. N-octylglucopyranosid (NOG, 1359 088) and modified, sequencing grade trypsin (T-1443) were purchased from Roche. Formic acid (FA, AC-3456-1) was obtained from Anachemia Science. HPLC grade acetonitrile was from Fisher (A998-4). Tris (TRS001) and SDS (SDS001) are supplied by Bioshop. Recombinant calbindin D_{28k} was prepared as outlined in Chapter 2 and the His-tag was removed with factor Xa before use (Section 2.3.2).

3.2.2. BS3 Cross-linking reactions

A stock solution of 60 μ M myoglobin was prepared by weight and diluted to 15 μ M in 20 mM HEPES buffer (pH 7.0). Ammonium and Tris buffers should be avoided (as suggested by the supplier) because they contain amino groups, which are reactive towards BS3. A 10 mM stock BS3 solution was prepared by dissolving 5 mg of the

powder in 1 mL of 0.5 mM sodium citrate buffer, pH 5.0. The slightly acidic pH of the sodium citrate buffer reduces the occurrence of base catalyzed hydrolysis of the cross-linker prior to the reaction with the protein. Between 0.3 to 5 molar excess of BS3 was mixed with myoglobin. The reaction was carried out at room temperature for 30 min and quenched by adding Tris buffer (pH 8.0) to a final concentration of 20 mM. Tris contains an amino group that attacks the carbonyl carbon of the cross-linker (Figure 3.1) and forms an amide bond in an analogous manner to lysine residues. Controls were prepared by adding sodium citrate (without BS3) and Tris to the protein solution. Calbindin D_{28k} ($\epsilon_{280} = 28037 \text{ M}^{-1}\text{cm}^{-1}$ for the calcium-loaded form) (36) was dissolved in 20 mM HEPES buffer (pH 7.4) with 0.1 mM CaCl₂ to a final concentration of 15 μM , and incubated with 0.5 to 10 molar excess of BS3 as outlined above for myoglobin.

3.2.3. ESI-MS of intact myoglobin and calbindin D_{28k} from BS3 cross-linking reactions

The undigested reaction products from the cross-linking reactions were characterized by ESI-MS. Excess salt and buffer were removed by washing the reaction solution with H₂O in a 5 kDa molecular weight cutoff centrifugal size-exclusion filter (Millipore, UFV5BG00). ACN, H₂O and TFA were added to obtain a 10 μM myoglobin solution in 50% ACN/50% H₂O/0.1% TFA prior to MS analysis. Likewise, calbindin D_{28k} was diluted with ACN, H₂O and FA to give a 1 μM protein solution in 50% ACN/50% H₂O/0.2 % FA prior to MS analysis.

Two ESI-MS instruments were used in this study. The first MS used was a ThermoFinnigan SSQ7000 single quadrupole instrument. A solution of 10 μM cross-

linked myoglobin was directly infused into the source at a flow rate of 5 $\mu\text{L}/\text{min}$ using a Harvard Apparatus 980532 syringe pump. The source and capillary temperatures were set at 70°C and 250°C, respectively. Nitrogen sheath gas was applied at 35 psi, and ion scans were performed in the positive-ion mode between m/z 600 and 1500 at a scan time of 3 s. The observed spectra were deconvoluted using BioBrowser 1.0 software from ThermoFinnigan. The instrument was mass calibrated with L-methionyl-arginyl-phenylalanyl-alanine acetate (Research Plus Inc.) and apo-myoglobin (Sigma, A-8673) as external calibrants.

The second instrument used was a Waters Micromass Q-ToF-2 MS. The capillary, cone and ToF voltages were set to 3500 V, 50 V and 9100 V, respectively. The multiplier and the multichannel plate detector voltages were set at 550 V and 2000 V, respectively. The source block and desolvation temperatures were set at 80°C and 30°C, respectively. The scan range was from m/z 700 to 2000. Calibration was achieved using [glu1]-fibrinopeptide B (Sigma, F-3261). A 1 μM solution of cross-linked calbindin D_{28k} was directly infused into the source at a flow rate of 1 $\mu\text{L}/\text{min}$.

3.2.4. SDS-PAGE analysis of cross-linked myoglobin and cross-linked calbindin D_{28k} solutions

Refer to Section 2.2 for the analysis of protein molecular weights by SDS-PAGE.

3.2.5. Digestion of the myoglobin and calbindin D_{28k} products from the BS3 cross-linking reaction

Two proteolytic enzymes were employed in this study. The myoglobin products were digested with trypsin (0.5 $\mu\text{g}/\mu\text{L}$ trypsin in 1% acetic acid) for 14 to 16 h at 37°C with a 1:20 (w:w) enzyme to myoglobin ratio in 100 mM TrisHCl (pH 7.5) with 0.01% SDS, which was added as a denaturant to aid in myoglobin digestion by denaturing the protein and exposing cleavage sites. The calcium-loaded calbindin D_{28k} products were digested with GluC (0.5 $\mu\text{g}/\mu\text{L}$ in H₂O) at 22°C for 14 to 18 h at a 1:20 (w:w) enzyme to calbindin ratio in 5 mM phosphate (pH 7.8) with 0.001% SDS. A double digest was also performed on the calbindin D_{28k} sample. First the protein was treated with Glu-C under the conditions outlined above. The sample was then heated at 95°C for 5 min to inactivate GluC, and trypsin (0.5 $\mu\text{g}/\mu\text{L}$ trypsin in H₂O) was added to further digest the peptides. The tryptic digest was carried out for 14 to 18 h at 37°C with a 1:20 (w:w) enzyme to protein ratio.

Samples digested in 100 mM TrisHCl were desalted using C18 ZipTips (Millipore) as per the manufacturer's instructions (45 2002)). This step was necessary because peptide-salt adducts are formed at high salt concentrations, which suppresses the peptide signals in the MS analysis. Samples in less than 10 mM buffer salt did not require desalting prior to MS analysis.

3.2.6. MALDI-MS analysis of the cross-linked peptides

For MALDI-ToF analysis, approximately 1 μL of 1 $\mu\text{g}/\mu\text{L}$ of digest was mixed with 1 μL of matrix solution. The matrix was either HCCA (10 mg/mL in 50% ACN/50% H₂O) or a 2:3 (v:v) mixture of NOG (6 mg/mL in H₂O) and HCCA (10 mg/mL in 50% ACN/50% H₂O). One μL of the each digest/matrix mixture was spotted

on a 96 well stainless steel MALDI target plate (Waters, Micromass). The samples were analyzed on a Waters Micromass M@LDI-ToF-LR mass spectrometer equipped with a nitrogen laser (emitting at 337 nm). The instrument was operated in the positive-ion, reflectron mode. The laser was pulsed at a rate of 5 Hz at a voltage between 2.4 to 2.5 kV. The source voltage and reflectron voltages were set at 15 kV and 2 kV, respectively. A microchannel plate detector set at 1.8 kV was used to detect the ions. Each spectrum represents an average of at least 100 accumulated shots. The instrument was mass calibrated with the external standards angiotensin (1296.685 m/z), renin (1758.933 m/z) and 18-39 clip ACTH (2465.199 m/z), and the data were analyzed using Waters MassLynx 4.0 software. Internal mass lock corrections were performed based on peaks that correspond to unmodified peptides whose ions exhibited high relative abundance. The MALDI mass spectra of the cross-linked samples were compared to those of the controls. Peaks appearing in the samples treated with BS3 but not in the controls were submitted to the Automated Spectrum Assignment Program, or ASAP (31).

3.3. Results

3.3.1. Cross-linking of myoglobin

Myoglobin was reacted with varying amounts of BS3 cross-linker and analyzed by SDS-PAGE. The gel indicated that there was little intermolecular cross-linking when the BS3-to-protein molar ratio was less than 3:1 (Figure 3.2). At higher ratios of cross-linker faint bands corresponding to myoglobin dimer and trimer become more intense. Intermolecular cross-linking should be avoided in this type of experiment because

residues are cross-linked with no dependence on their spatial separation within a molecule. The band in the gel appearing below the unmodified protein (Figure 3.2) corresponds to intramolecular cross-linked product.

ESI-MS analysis of the reaction mixture indicates the presence of a number of species and structures of the observed products and the corresponding mass increases are

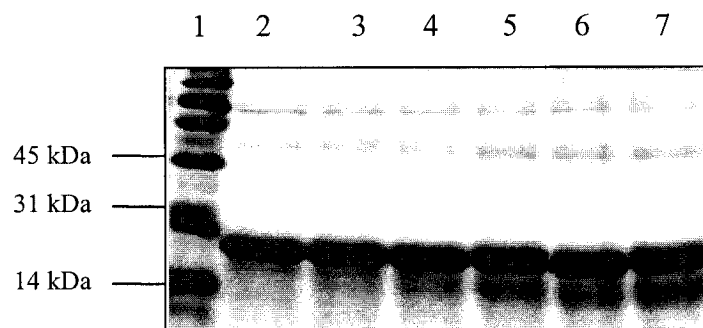


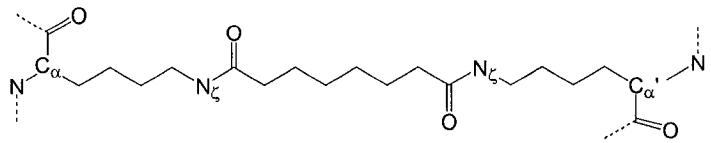
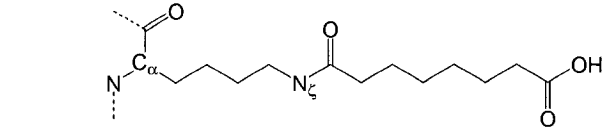
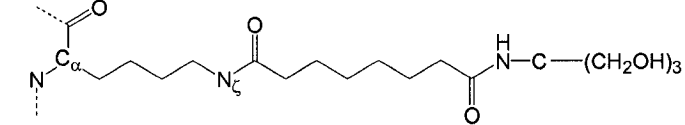
Figure 3.2. 12% SDS-PAGE analysis of myoglobin from the BS3 cross-linking reactions. Myoglobin (15 μ M) in 20 mM HEPES buffer (pH 7.0) was reacted with BS3 at room temperature for 30 min in the following BS3:myoglobin molar ratios: (2) 0:1; (3) 0.3:1; (4) 0.5:1; (5) 1:1; (6) 3:1; (7) 5:1. The reactions were quenched by adding 20 mM Tris buffer (pH 8.0). Ten μ g of myoglobin was added to each well and a 30 mA current was applied across the gel. Lane (1) contains the protein molecular weight markers.

shown in Table 3.1. For example, using a 3:1 ratio of BS3 to myoglobin, the unmodified protein appears as the base peak (Figure 3.3). Myoglobin with one intramolecular cross-link (Mb-BS3) appears at $M+138$ Da with 40% relative intensity. In addition to the target molecule, Mb-BS3, two side products are also observed. These were labeled at lysine residues with BS3-OH or BS3-Tris. Tris quenching is important to ensure that there is no reactive cross-linker remaining before proceeding with the digestion. Very little protein contained multiple modifications on a single molecule.

At a ratio of 5:1, the cross-linked species was more abundant than unmodified myoglobin (Figure 3.4). However, up to three modifications [Mb-(BS3)(BS3OH)₂] on a

single protein molecule were also observed. Multiple modifications should be avoided because there is concern that this may alter the protein structure (31, 32). The optimized ratio of BS3-to-myoglobin was, therefore, 3:1 despite the lower yield of the desired target molecule, Mb-BS3.

Table 3.1. Possible products following modification of protein by BS3

Modification		Mass shift, Da
BS3 ^a		138.068
BS3-OH ^b		156.079
BS3-Tris ^c		259.142

^a Both sulfosuccinimidyl ester groups of BS3 reacted with lysine sidechains. Note the C_α - N_ζ and C_α - C_α separations are 6 and 24 Å respectively.

^b One sulfosuccinimidyl ester group reacted with a lysine sidechain and the other ester group of BS3 was hydrolyzed by the solvent.

^c One sulfosuccinimidyl ester group reacted with a lysine sidechain and the other ester group of BS3 was quenched with Tris.

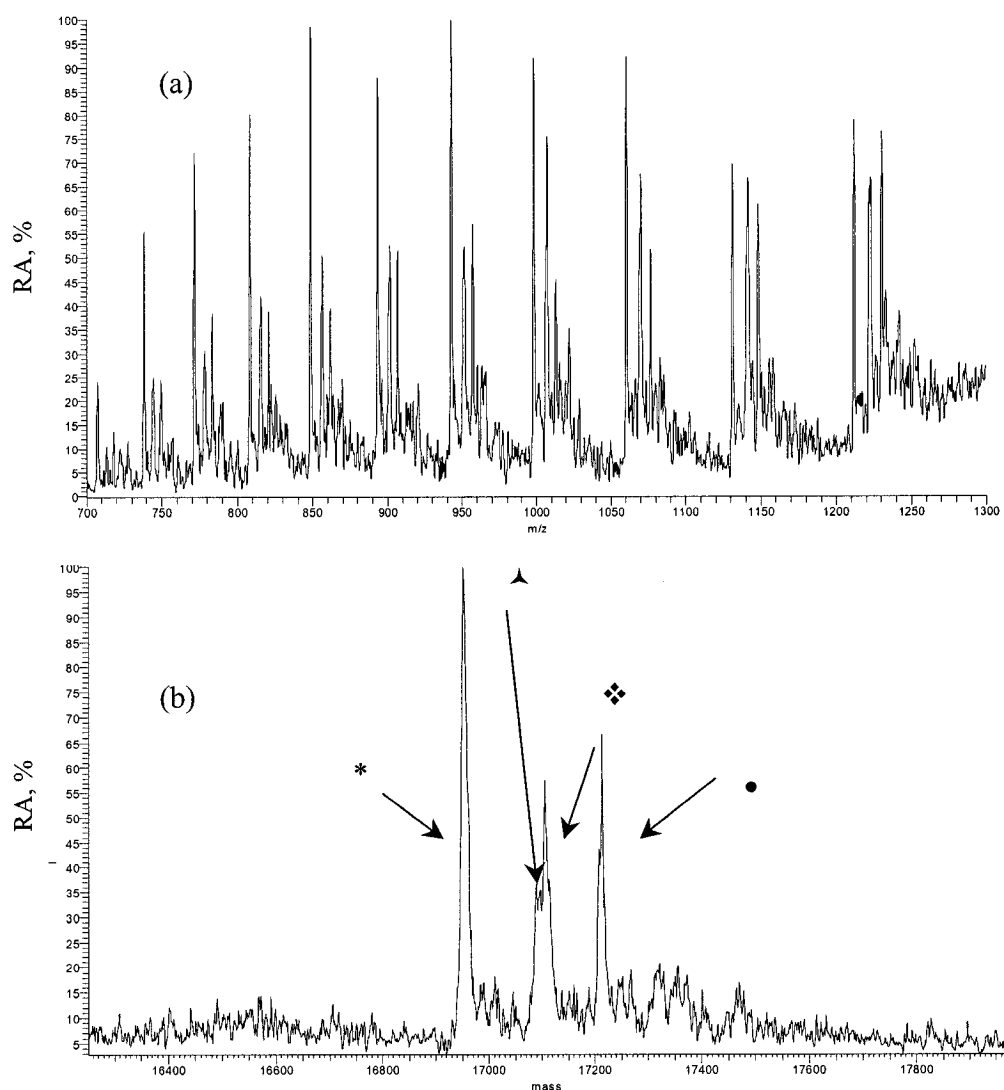


Figure 3.3. ESI-MS spectrum of myoglobin reacted with 3-fold molar excess of BS3. (a) Observed spectrum. (b) Deconvoluted spectrum. (*) unreacted myoglobin, (▲) cross-linked myoglobin (Mb-BS3), (❖) myoglobin modified by BS3-OH, (●) myoglobin modified by BS3-Tris. Myoglobin from the BS3 reaction was diluted to 10 μM into 50% ACN/50% H_2O /0.1% TFA and directly infused at a flow rate of 5 $\mu\text{L}/\text{min}$ into the ESI source of a ThermoFinnigan SSQ7000 MS. The source and capillary temperatures were set at 70°C and 250°C, respectively. Nitrogen sheath gas was applied at 35 psi, and ion scans were performed in the positive-ion mode between m/z 600 and 1500 at a scan time of 3 s. RA is the relative abundance of the ions.

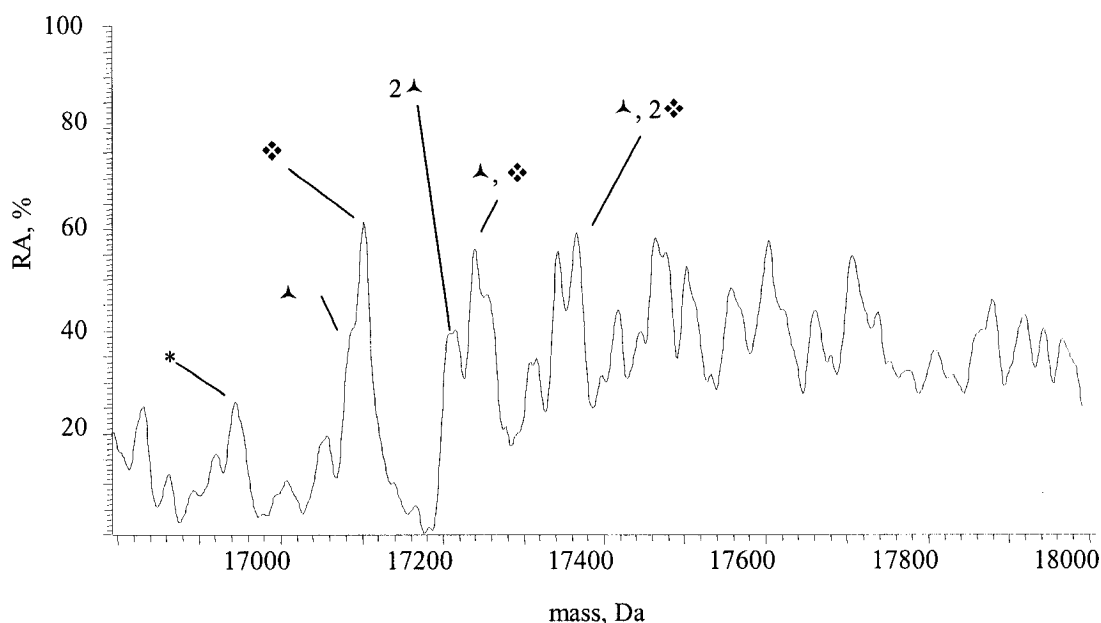


Figure 3.4. Deconvoluted ESI-MS spectrum of myoglobin reacted with 5-fold molar excess of BS3. (*) unreacted myoglobin, (▲) cross-linked myoglobin (Mb-BS3), (♦) myoglobin modified by BS3-OH, (●) myoglobin modified by BS3-Tris. See the legend of Figure 3.3 for the experimental conditions. RA is the relative abundance of the ions.

3.3.2. Cross-linking of calcium-loaded calbindin D_{28k}

Cross-linking of calbindin D_{28k} was also optimized to avoid BS3 intermolecular cross-linking as assessed by SDS-PAGE (Figure 3.5). Under the non-reducing conditions employed, a weak band in the control (Lane 2, no BS3) appears just below the 66 kDa marker. This probably corresponds to dimers formed by intermolecular disulfide linkages. As the amount of BS3 added is increased a second band, also just below the 66 kDa marker, increases in intensity. This presumably corresponds to intermolecular BS3-cross-linked calbindin D_{28k}. A 5:1 cross-linker to protein ratio was chosen to be the optimal condition as assessed from the gel (Lane 6). ESI-MS analysis of the mixture (Figure 3.6) indicates that the major product from the reaction is the intramolecular cross-

linked species, calb-BS3. BS3-OH modified calbindin D_{28k} is also observed but no calb-BS3-Tris species.

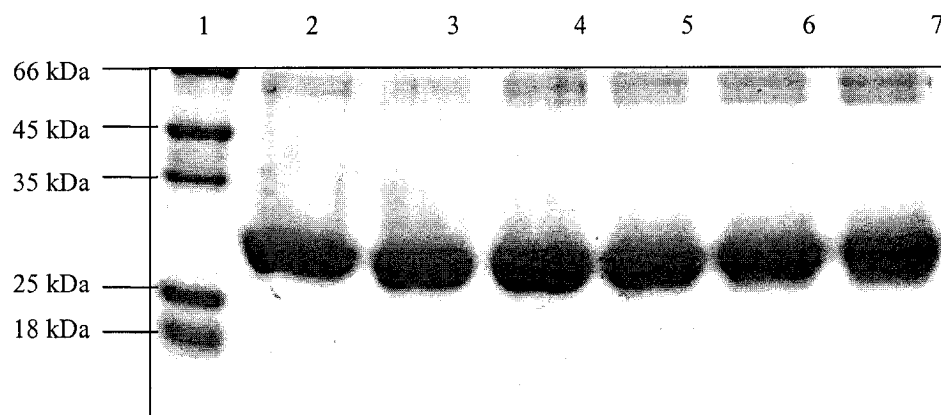


Figure 3.5. 15% SDS-PAGE analysis of calbindin D_{28k} from the BS3 cross-linking reactions. Calbindin D_{28k} (10 μ M) in 20 mM HEPES buffer (pH 7.4) was reacted with BS3 at room temperature for 30 min in the following BS3:calbindin molar ratios: (2) 0:1; (3) 0.5:1; (4) 1:1; (5) 3:1; (6) 5:1; (7) 10:1. The reactions were quenched by adding 20 mM Tris buffer (pH 8.0). Ten μ g of calbindin D_{28k} was added to each well and a 30 mA current was applied across the gel. Lane 1 contains the protein molecular weight markers. The sample buffer used was non-reducing (did not contain β -mercaptoethanol).

3.3.3. Identification of cross-linked sites by MALDI-ToF-MS

The digests of myoglobin and calbindin D_{28k} from the BS3 cross-linking reactions were analyzed by MALDI-ToF. The percent sequence coverages of the digests were similar for the samples treated with the cross-linker and for the controls.

The MALDI peptide mass map of the tryptic digest of the myoglobin products from a 3:1 cross-linking reaction is shown in Figure 3.7. The sequence coverage from this map was 71%. In addition to the in-solution digest, the peak containing cross-linked myoglobin (MB-BS3) was cut from the SDS-PAGE gel and in-gel digested. The resulting MALDI peptide map of the digest is shown in Figure 3.8 (62% sequence coverage). Figures 3.9 and 3.10 show the maps of a GluC digest and a tandem digest of the calbindin D_{28k}

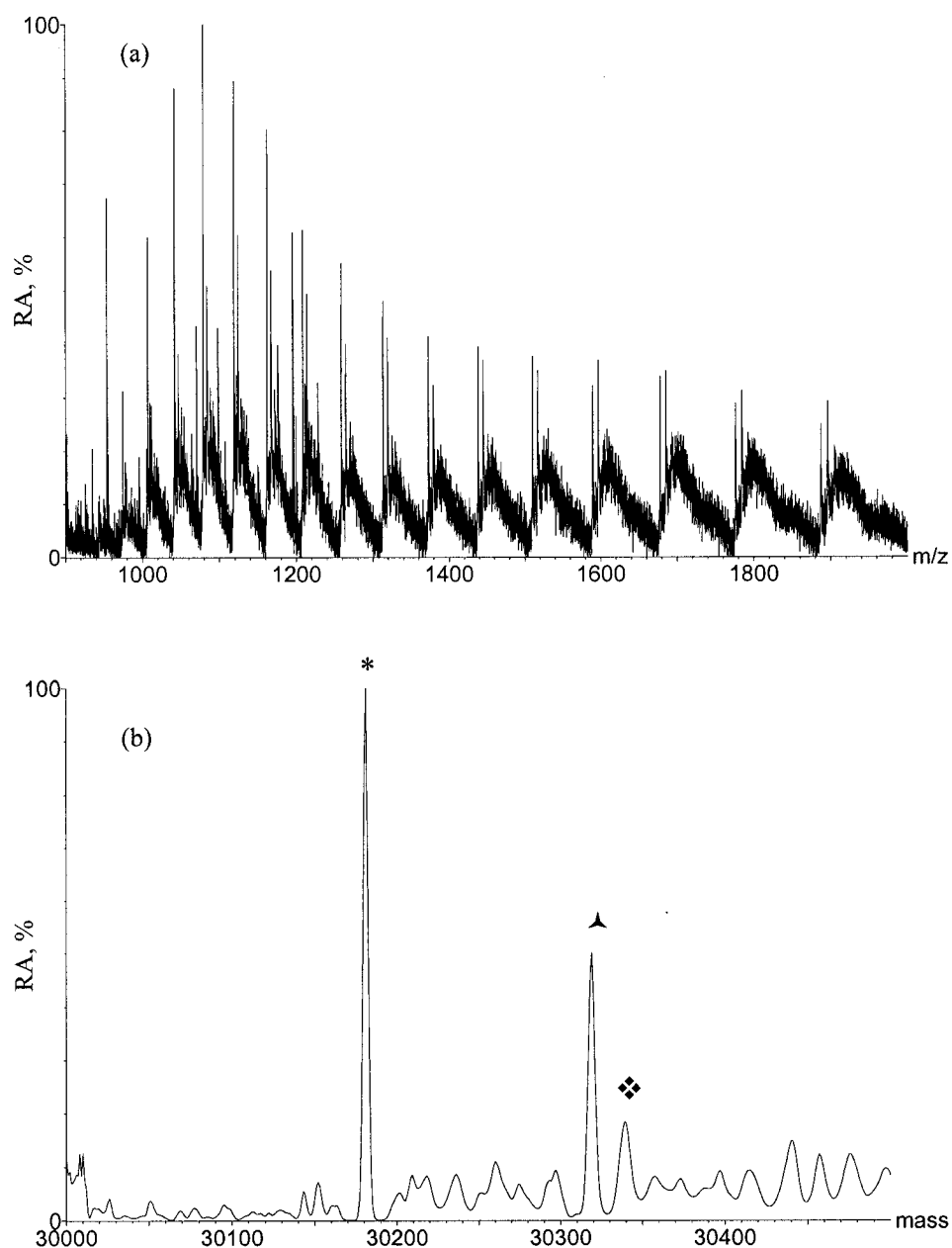


Figure 3.6. ESI-MS spectrum of calbindin D_{28k} reacted with 5-fold molar excess of BS3. (a) Observed spectrum. (b) Deconvoluted spectrum. (*) unreacted calbindin D_{28k}, (▲) cross-linked calbindin D_{28k}, (◆) calbindin D_{28k} modified by BS3-OH. His-calb (1 μ M) in 50% ACN/50% H₂O/0.2% FA was directly infused into a Waters Micromass Q-ToF MS with a flow rate of 1 μ L/min. The voltage settings of the mass spectrometer were: capillary: 3.8 kV, cone: 45 V, multiplier: 550 V, MCP: 2.1 kV. RA is the relative abundance of the ions.

products, respectively. The sequence coverage of the double digest map (40%) is lower than that of the single GluC digest (63%). This probably arises because the peptides are

smaller in the double digest and appear at $m/z < 700$ in the spectrum, a region where the matrix interferes with peptide signals. All ions detected were singly charged as demonstrated by the integral m/z difference separating the isotopic peaks.

The MALDI peptide mass maps of the control and of the BS3 treated samples were compared. Peaks appearing in the latter but not in the control were submitted to the online program ASAP (<http://roswell.ca.sandia.gov/~mmyoung/asap.html>) (31). This program matches the observed m/z with the theoretical masses of peptides, and allows for modifications such as those outlined in Table 3.1. A tolerance of 100 ppm between the theoretical and observed m/z values was used because of the low signal-to-noise ratios of the MS peaks, which can decrease the accuracy of the measurements.

The observed peaks that correspond to modified peptides of myoglobin are summarized in Table 3.2 with the positions of the lysines (K) assumed to be involved given as subscripts. Six peaks were uniquely identified as cross-linked peptides. Three peaks could not be assigned to a unique species. For example, the peak at m/z 1200.650 corresponds to the mass of a BS3-cross-linked peptide between K₄₅ and K₄₇ or a BS3-Tris modified peptide at K₁₄₇. Also, the peak at m/z 2328.378 corresponds to two isobaric species (FDK₄FK – K₆₃HGTVVLTALGGILK cross-linked between K₄ and K₆₃, and FDK₄₅FK – HGTVVLTALGGILK₇₇K cross-linked between K₄₅ and K₇₇). There were four BS3-Tris-modified peptides detected by MALDI-MS but no BS3-OH peptides, although the ESI mass spectrum of the intact protein (Figure 3.2) indicates the presence of BS3-OH labeled myoglobin.

The band corresponding to cross-linked myoglobin (Mb-BS3) was cut from the SDS-PAGE gel. In-gel digestion was performed and the digest analyzed by MALDI-MS

(Figure 3.8). The only cross-linked peptides detected were LFTGHPETLEK₄₂FDK₄₅FK (m/z 2075.085) and LFTGHPETLEK₄₂FDK–HK₉₈IPIK (m/z 2534.426), which were also observed in the in-solution digest of myoglobin products from the BS3 reaction (Table 3.2).

Cross-linked peptides of calbindin D_{28k} were more elusive and required a trial-and-error approach in the sample preparation protocol. MS analysis of a tryptic digest of the calbindin D_{28k} products from the 5:1 BS3 reaction did not show any cross-linked species. Four cross-linked peptides were found in a GluC digest when NOG was added to the matrix. NOG is a detergent that is thought to solubilize high molecular weight peptides, allowing them to be co-crystallized with the matrix (46). Tandem digestion by GluC followed by trypsin allowed three additional cross-linked peptides to be detected. The modified peptides found in the calbindin D_{28k} digests are summarized in Table 3.3.

3.4. Discussion

3.4.1. Myoglobin and calbindin D_{28k} cross-linked products

The characterization of cross-linked myoglobin and calbindin D_{28k} was accomplished through SDS-PAGE and mass spectral analysis. The SDS-PAGE analysis provides a simple method for assessing the extent of intermolecular cross-linking. It is also able to resolve intramolecularly cross-linked proteins, presumably because they are more compact than a true random coil and therefore migrate through the gel faster. This is evident in the case of myoglobin where, with the addition of the cross-linker, a new band appears below that of unmodified myoglobin (Figure 3.2). Thus, SDS-PAGE

provides evidence of intramolecular cross-linking based on the migration of the modified protein. Cross-linked calbindin D_{28k} is not as well resolved by SDS-PAGE (Figure 3.5). Nonetheless, the band in the SDS-PAGE corresponding to calbindin D_{28k} from the cross-

Table 3.2. Modified peptides detected in the MALDI mass fingerprint of horse heart myoglobin from a BS3 reaction

m/z, observed ^a	Peptide(s) ^b	Modifi- cation ^c	Error, ppm ^d	Lys C _α – Lys C _α Å ^e
1200.650	YK ₁₄₇ ELGFQG	BS3-Tris	29	-
	FDK ₄₅ FK ₄₇ HLK	x-link	23	6.0
1766.072	HGTVVLTALGGILK ₇₇ K	BS3-Tris	4	-
	K ₆₃ GTVVLTALGGILK	BS3-Tris	4	-
2075.085	LFTGHPETLEK ₄₂ FDK ₄₅ FK	x-link	10	8.9
2113.017	GHHEAELK ₈₇ PLAQSHATK	BS3-Tris	41	-
2257.199	GHHEAELK ₈₇ PLAQSHATK ₉₆ HK	x-link	16	13.5
2288.329	ASEDLK ₆₂ K ₆₃ HGTVVLTALGGILK	x-link	12	3.8
2328.378	FDK ₄ FK – K ₆₃ HGTVVLTALGGILK	x-link	13	12.0
	FDK ₄₅ FK – HGTVVLTALGGILK ₇₇ K	x-link	13	29.3
2471.347	LFTGHPETLEK ₄₂ FDK – FK ₄₇ HLK	x-link	15	8.1
2534.426	LFTGHPETLEK ₄₂ FDK – HK ₉₈ IPIK	x-link	19	7.6
3560.010	HLK ₅₀ TEAEMK – IPIK ₁₀₂ YLEFISDAIIHVLHKS	x-link	14	23.0
3663.183	GLSDGEWQQVLNVWGK ₁₆ VEADI AGHGQEVLR	BS3-Tris	82	-

^a Observed m/z from MALDI mass fingerprint (Figure 3.7).

^b Modified peptide with a mass corresponding to the observed mass. The positions in the sequence of the lysines (K) assumed to be cross-linked or modified are given as subscripts. A cross-link between lysines on different tryptic peptides is indicated by a dash (–).

^c Lysine modification as defined in Table 3.1.

^d Difference between the observed m/z given in column 1 and the theoretical m/z of the peptide multiplied by 10⁶.

^e Distance separating the C_α's of the cross-linked lysine residues based on the crystal structure (PDB ID: 1dwr).

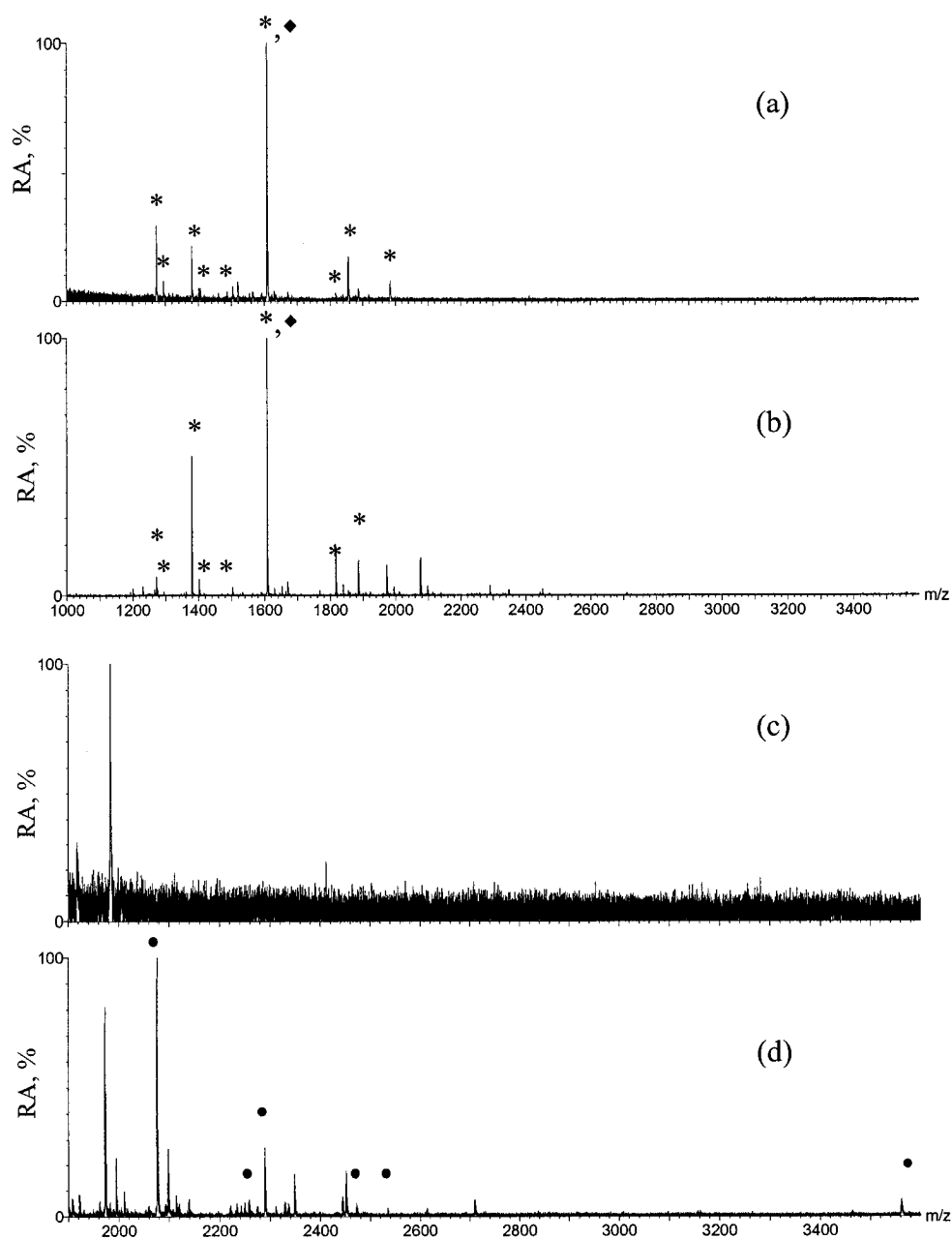


Figure 3.7. MALDI mass fingerprint of tryptic digests of (a) unmodified myoglobin and (b) myoglobin reacted with a 3-fold excess of BS3. The digestion was carried out at a 1:20 (w:w) protein-to-trypsin ratio in 100 mM TrisHCl (pH 8.5) with 0.01% SDS at 37°C for 16 -18 h. One μg of digest was mixed with 1 μL of 10 mg/mL HCCA matrix and spotted on a MALDI target plate. (c) and (d) Expansion of region of (a) and (b), respectively, to observe the ions due to the cross-linked peptides. (*) unmodified peptides, (♦) peak used as an internal mass correction, (•) cross-linked peptides. The spectra were acquired with a Waters Micromass M@LDI analyzer operated in positive ion, reflectron mode. The spectra are averages of at least 100 laser shots. RA is the relative abundance of the ions.

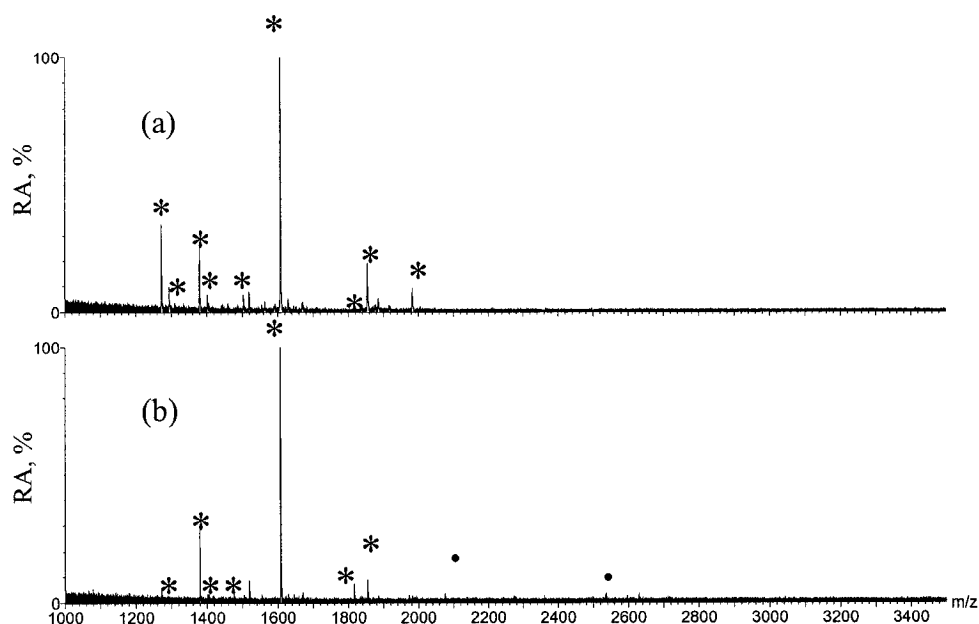


Figure 3.8. MALDI mass fingerprint of the in-gel tryptic digest of (a) unmodified myoglobin and (b) BS3-cross-linked (Mb-BS3). Myoglobin and cross-linked myoglobin were isolated from a 15% SDS-PAGE gel. Gel pieces (1 mm²) were incubated with three different wash solutions for 15 min each. The wash solutions were: (1) 100 μ L of 50% ACN, (2) 100 μ L of 50% ACN in 50 mM ammonium bicarbonate and (3) 100 μ L of 50% ACN in 15 mM N-ethyl morpholine plus 5 mM acetic acid. The gel pieces were dried in a speed vac, and in-gel digestion of the protein was carried out for 16-18 h at 30°C with 10 μ L of trypsin digest solution (15 mM N-ethyl morpholine, 42 mM acetic acid, 8.3 ng/ μ L trypsin and 2 mM CaCl₂). Peptides were extracted from the gel with 60% ACN/0.5% formic acid, dried using the speed vac, resuspended in 10 μ L of 5% methanol, and washed with C18 ZipTips. The desalted peptides were eluted with 60% ACN/0.1% TFA. One μ L of the eluted digest was mixed with 1 μ L of 10 mg/mL HCCA matrix and spotted on a MALDI target plate. (*) unmodified peptides, (•) cross-linked peptides. The spectra were acquired with a Waters Micromass M@LDI analyzer operated in positive ion, reflectron mode. The spectra are averages of at least 100 laser shots. RA is the relative abundance of the ions. RA is the relative abundance of the ions.

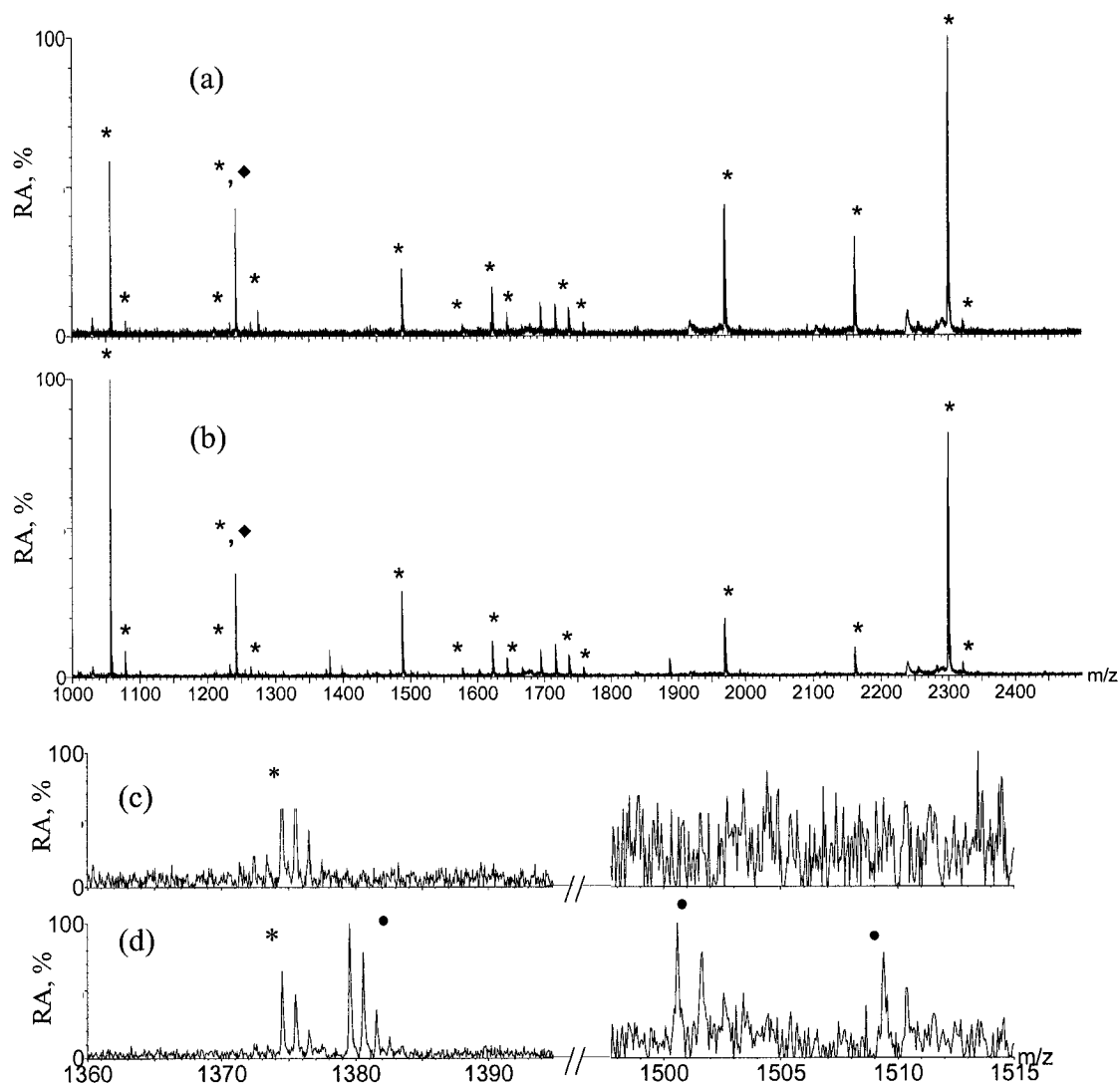


Figure 3.9. MALDI mass fingerprint of a GluC digest of (a) unmodified calbindin D_{28k} and (b) calbindin D_{28k} reacted with 5-molar excess of BS3. The digestion was carried out at a protein-to-GluC ratio of 1:20 (w:w) in 5 mM phosphate buffer (pH 7.8) with 0.01% SDS at 22°C for 14 h. One μ g of digest was mixed with 1 μ L of a 5 mg/mL NOG plus 6 mg/mL HCCA matrix mixture and spotted on a MALDI target plate. (c) and (d) Expansion of region of (a) and (b), respectively, to observe the ions due to the cross-linked peptides. (*) unmodified peptides, (♦) peak used as an internal mass correction, (•) cross-linked peptides. The spectra were acquired with a Waters Micromass M@LDI analyzer operated in positive ion, reflectron mode. The spectra are averages of at least 100 laser shots. RA is the relative abundance of the ions.

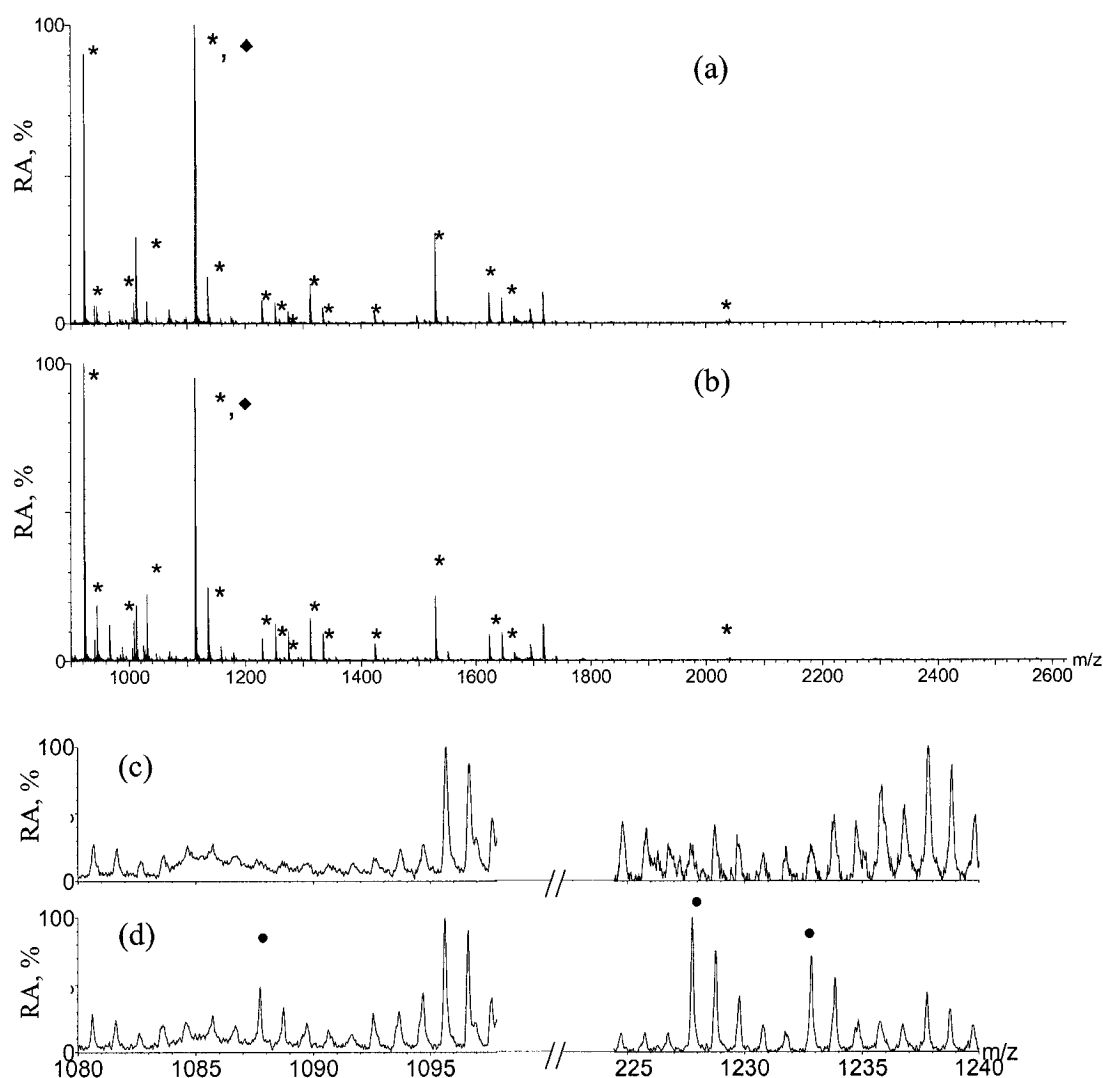


Figure 3.10. MALDI mass fingerprint of a trypsin and GluC double digest of (a) unmodified calbindin D_{28k} and (b) calbindin D_{28k} reacted with 5-molar excess of BS3. The digestion was carried out at a protein-to-GluC ratio of 1:20 (w:w) in 5 mM phosphate buffer (pH 7.8) with 0.01% SDS at 22°C for 14 h. The digest mixture was then further digested with peptide-to-trypsin 1:20 (w:w) 37°C for 16 h in the phosphate buffer. One µg of digest was mixed with 1 µL of a 10 mg/mL HCCA matrix mixture and spotted on a MALDI target plate. (c) and (d) Expansion of region of (a) and (b), respectively, to observe the ions due to the cross-linked peptides. (*) unmodified peptides, (♦) peak used as an internal mass correction, (•) cross-linked peptides. The spectra were acquired with a Waters Micromass M@LDI analyzer operated in positive ion, reflectron mode. The spectra are averages of at least 100 laser shots. RA is the relative abundance of the ions.

Table 3.3. Modified peptides detected in the MALDI mass fingerprint of calbindin D_{28k} from a BS3 reaction

m/z Observed	Peptide(s) ^c	Modification ^d	Error ^e , ppm
1087.64 ^a	TK ₁₄₂ LAE – K ₂₂₁ NK	x-link	1
1227.69 ^a	K ₄₈ K – LLEK ₁₃₃ ANK	x-link	6
1232.74 ^a	GK ₃₄ E – LK ₁₂₄ NFLK	x-link	2
1374.735 ^b	TEELK ₁₂₄ NFLK ₁₂₈ D	x-link	14
	GK ₃₄ E – ALLK ₂₁₆ DLCE	x-link	12
1379.778 ^b	GK ₃₄ E – G GK ₂₄₆ LYRTD	x-link	44
1500.876 ^b	GK ₃₄ E – NELDALLK ₂₁₆ D	x-link	64
1509.611 ^b	MK ₅₉ TFVD – K ₂₂₁ NK ₂₂₃ QD	x-link	100

^a Peptide ions observed following a tandem GluC and trypsin digest of cross-linked calbindin D_{28k}.

^b Peptide ions observed following a GluC digest of cross-linked calbindin D_{28k}.

^c Modified peptide with a mass corresponding to the observed mass. The positions in the sequence of the lysines (K) assumed to be cross-linked or modified are given as subscripts. A cross-link between lysines on different tryptic peptides is indicated by a dash (–).

^d Lysine modification as defined in Table 3.1.

^e Difference between the m/z observed given in column 1 and the theoretical m/z of the peptide multiplied by 10⁶.

linking reaction is broadened compared to the control indicating that a number of intramolecular cross-linked species are present.

Calbindin D_{28k} appears to be more reactive towards BS3 than does myoglobin. The relative abundance of the ions corresponding to the singly cross-linked products are 40% and 50% for myoglobin and calbindin D_{28k}, respectively while the corresponding abundance of the BS3-OH ions are 60 and 20%, respectively (Table 3.4). Thus, when one succinimidyl ester group reacts with calbindin D_{28k} the second is more prone to react with another lysine residue than a water molecule. No calbindin-BS3-Tris species are

detected by ESI-MS (Figure 3.6) suggesting that no BS3 remained after 30 min incubation to react with Tris. This is further evidence that calbindin reacts faster with BS3 than myoglobin. In contrast, the relative abundance of myoglobin-BS3-Tris ion was 65% (Table 3.4). The higher reactivity of BS3 with calbindin D_{28k} compared to myoglobin may be in part related to the higher abundance of lysine residues in calbindin D_{28k} (25 lysines) vs. myoglobin (19 lysines).

Table 3.4. Relative abundance of ions arising from modified proteins in the ESI mass spectra

Modification ^a	Myoglobin ^b	Calbindin D _{28k} ^c
BS3	40	50
BS3-OH	60	20
BS3-Tris ^d	65	-

^a See Table 3.1.

^b See Figure 3.3.

^c See Figure 3.6.

^d Formed on quenching with 20 mM Tris (see Section 3.2.).

3.4.2. Cross-linked peptides in myoglobin

Following the cross-linking reaction, myoglobin was digested and analyzed by MALDI-ToF-MS. All the cross-linked peptides in myoglobin identified in Table 3.2 contain at least three lysine residues. It is assumed that the C-terminal lysines are not involved in cross-linking. Modified lysine residues are not recognized by trypsin (31) so that in the peak at m/z 2075.085, for example, K47 is the trypsin cleavage site and BS3 must cross-link K42 with K45.

Not all peaks in the myoglobin digest could be uniquely identified within the 100 ppm tolerance in m/z . For example, the peak at m/z 1200.650 may correspond to a BS3

cross-linked peptide or a BS-Tris-modified peptide (Table 3.2). The peak at m/z 2328.378 presents another signal that is difficult to assign. Two isobaric cross-linked peptides can be assigned to this mass (Table 3.2) so sequencing is necessary to identify this species.

Six cross-linked myoglobin peptides were uniquely identified (Table 3.2). The lysine separations of the cross-linked peptides were determined from the crystal structure and all were within the C_{α} - C_{α} distance of 24 Å dictated by the cross-linker (Table 3.1). The crystal structure of myoglobin with the identified cross-links is shown in Figure 3.11.

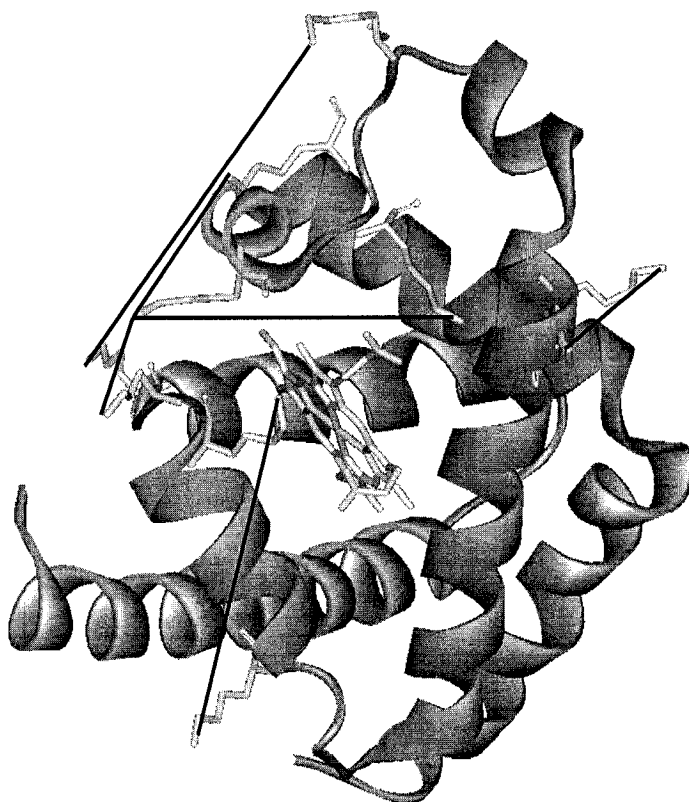


Figure 3.11. Ribbon diagram of the NMR derived structure of myoglobin showing the six observed BS3 cross-links. The cross-links between lysine N_{ϵ} 's are shown by solid lines. The structure was taken from PDB ID: 1dwr and the uniquely assigned cross-links detected by MALDI mass fingerprinting are listed in Table 3.2.

From an examination of the location of the lysine residues in the primary structure (Table

3.2) it is clear that not all cross-linked peptides yield useful information. For example, the peak at m/z 2288.329 corresponds to a peptide where K62 and K63, adjacent lysines residues, are cross-linked. Generally, when any neighboring residues are positioned such that their C_α atoms are maximally separated, these atoms are up to 4 Å apart (31, 32). Therefore, any two lysine residues that are six residues apart or less in the primary structure must have C_α atoms within 24 Å (31, 32). The peaks appearing at m/z 2471.347 and 2075.085 correspond to products from cross-links between lysines that are 5 and 3 residues apart, respectively. Therefore, only 3 (m/z 2257.199, 2534.426 and 3560.010) of the 6 uniquely identified species yield useful structural information.

3.4.3. Factors influencing the cross-linking of myoglobin

Proximity appears to be an important factor in the occurrence of a BS3 cross-link between lysine residues. Considering five of the six uniquely identified cross-linked peptides (Figure 3.11), the distances separating their lysine C_α atoms range from 4 to 14 Å in the crystal structure of myoglobin. This is much closer than the 24 Å maximum dictated by the cross-linker structure (31). The cross-linked peptide HLK₅₀TEAEMK – IPIK₁₀₂YLEFISDAIIHVLHISK (m/z 3560.010) where the C_α atoms of K₅₀ and K₁₀₂ are separated by 23 Å appears to be an exception. Involvement in salt bridges does not necessarily inhibit cross-linking of a lysine. Based on the crystal structure, of the 12 lysine residues cross-linked, four were found to be in salt bridges (K42, K50, K62 and K98).

3.4.4. Cross-linked peptides in calbindin D_{28k}

Cross-linked calbindin D_{28k} was predominantly produced with little side products (Table 3.4). Thus, only cross-linked peptides were detected and no BS3-Tris or BS3-OH modified peptides were seen in the peptide maps (Table 3.3). Furthermore, all the cross-linked lysines are more than 6 residues apart in the primary structure of calbindin D_{28k}. There are two peaks at m/z 1374.735 and 1509.611 (Table 3.3) which arise from modified peptides that could not be uniquely identified. The latter involves either K221 or K223 linked to K59. Since K221 and K223 are in close proximity, this cross-linked peptide can still be useful in defining the orientation of the subdomains. A model structure predicting that the C $_{\alpha}$ atoms of both K59-K221 and K59-K223 are separated by more than 24 Å can be rejected.

3.4.5. Stability of the BS3 cross-linked peptides

Throughout the experimental manipulation, the cross-linked species were subjected to thermal and chemical stress. For example, prior to SDS-PAGE analysis, cross-linked myoglobin and calbindin D_{28k} were heated to 95°C for 5 min. Therefore, the cross-links must be thermally stable in order for the modified proteins to be detected in the gels (Figures 3.2 and 3.5). Moreover, myoglobin and calbindin D_{28k} from the cross-linking reactions were dissolved in 0.1% TFA and 0.2% FA, respectively, prior to ESI analysis (Figures 3.3 and 3.6), so the cross-links observed must be stable under the acidic conditions used in the MS analysis. The cross-links are also resistant to enzymatic digestion and trypsin does not cleave at BS3-labelled lysine residues (31). The detection of cross-linked peptides in the MALDI mass maps (Tables 4.2 and 4.3) confirms that the BS3 cross-links are resistant to proteolysis by trypsin.

3.4.6. Abundance of the cross-linked peptides

Although it is clear from the MS (Figures 3.3 and 3.6) and SDS-PAGE (Figures 3.2 and 3.5) analyses that ions from cross-linked calbindin D_{28k} and cross-linked myoglobin are produced in high relative abundance few cross-linked peptides were detected (Figures 3.7, 3.8, 3.9 and 3.10). The low abundance of the ions corresponding to the cross-linked peptides could arise from a large distribution of possible products. Disregarding distance restraints between lysine residues, the number of combinations of links using a homobifunctional cross-linker can be expressed as $(n^2-n)/2$ where n is the abundance of reactive residues. For myoglobin, which contains 19 lysines, 171 combinations are possible. Based on the separation of lysine residues determined from the crystal structure of myoglobin, there are actually 72 different K-K (C_{α} - C_{α}) separations within 24 Å. Peaks corresponding to the cross-linked peptides of calbindin D_{28k} show lower signal-to-noise ratios compared to those observed for myoglobin (Figure 3.7 vs. 3.9 and 3.10). Calbindin D_{28k} contains 25 lysine residues so 300 different cross-links are possible if distance constraints are ignored.

It was suspected that the cross-linked species might not have been digested and that the peaks observed in the MALDI spectra are from unmodified protein. This argument was rejected because when the digests were analyzed by SDS-PAGE no bands due to undigested protein were detected (data not shown). Moreover, in-gel digestion of the isolated cross-linked myoglobin (Mb-BS3) resulted in 62% sequence coverage, but no additional cross-linked species were detected.

The desalting step following proteolytic digestion was considered a sink for cross-linked peptides. A reversed-phase absorbent matrix on a pipette tip (C18 ZipTip) was used in the desalting procedure. It was assumed that the peptides bound to the C18 matrix and, after washing out the salts, the peptides were eluted. To test the possibility that the cross-linked peptides were either not binding to, or not eluting from the C18 matrix the desalting step was omitted. To circumvent the desalting step, 10 mM buffer salts were used in the digestion since this did not show interference with the mass spectral analysis, but maintained enough buffering capacity to keep the proteolytic enzymes active. Even without the desalting step no additional cross-linked peptides were detected.

In our study, only six cross-linked peptides were found for both myoglobin and calbindin D_{28k}. Other research groups also reported low numbers (0-3) of cross-linked peptides. Pearson and coworkers examined intramolecular cross-linking of horse skeletal muscle myoglobin, bovine heart cytochrome c and bovine pancreas ribonuclease A using disuccinimidyladipate (DSA) as a cross-linker (32). The structure of DSA in Figure 3.12 shows that it differs from BS3 (Figure 3.1) in that it possesses a 4 vs. a 6 carbon linker

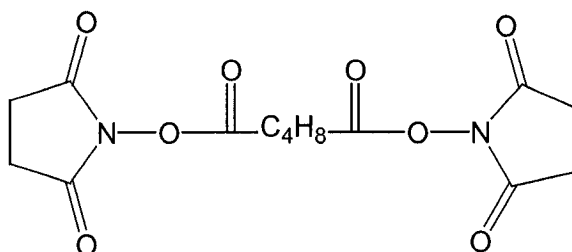


Figure 3.12. Stick diagram of disuccinimidyladipate (DSA).
The separation between the carbon atoms of the ester groups is 6 Å.

and no sulfate groups on the succinimidyl ring. MALDI mass fingerprint analysis of myoglobin reacted with 4 molar excess of DSA showed only Mb-DSA-OH and the

ammonium-quenched, DSA-modified (Mb-DSA-NH₂) protein. Three cross-linked peptides were detected for cytochrome c. Considering the 6 Å linker of DSA, 88 cross-links are possible based on the NMR structure of cytochrome c. Three cross-linked peptides were found in ribonuclease A. Pearson *et al.* also employed LC-ESI-MS to analyze tryptic digests for cross-linked peptides. One and three additional cross-links were found for ribonuclease A and cytochrome c, respectively.

Young and coworkers (31) were more successful in their work on FGF discussed in Section 1.4. Cross-linking, digestion and MALDI analysis of the protein revealed 18 BS3-cross-linked peptides. Young *et al.* also investigated the use of tandem mass spectrometry to identify the cross-linked sites in an intact cross-linked protein (33). Ubiquitin was cross-linked using a 2 molar excess of dissuccinimidyl suberate (DSS), which has the same structure as BS3, but is not sulfonated on the succinimidyl rings. The reaction mixture was directly infused into an ESI Fourier transform MS. Unmodified and singly cross-linked ubiquitin were isolated in separate experiments and fragmented using collision induced dissociation (CID). A list of the resulting fragments was compared to a theoretical library of all possible cross-linked products. Only two cross-linked peptides were reported using this approach.

4.0. Molecular modeling

4.1. Introduction

4.1.1. Building the calbindin D_{28k} models

Calbindin D_{28k} is comprised of 261 residues and contains approximately 4000 atoms. An efficient strategy to construct initial models of calbindin D_{28k} is therefore needed. This challenging task is simplified if the unknown protein is expected to be similar in structure to other proteins that have been previously characterized. In the case of calbindin D_{28k}, the protein is composed of six EF-hand subdomains (7). Because a large number of proteins of known structure contain EF-hand subdomains, a basis for constructing the subdomains of calbindin D_{28k} is available. The backbone of calbindin D_{9k} is used here as a template for the EF-hands of calbindin D_{28k}. Of a total of 76 residues in calbindin D_{9k}, 59 (78%) are involved in EF-hand formation (PDB ID: 1b1g). The remaining 17 residues form linker loops connecting the two EF-hands and the amino- and carboxyl-terminals.

The six EF-hands of calbindin D_{28k} can be arranged in many different configurations of the subdomains. The protein could conceivably have one to six separate domains containing one to six EF-hands each, and each EF-hand could pair with any of the other five EF-hands. Fortunately, experiments have been performed to probe domain organization within calbindin D_{28k} (Section 1.3) and the NMR structure of the EF1EF2 fragment of calbindin D_{28k} has been solved (47). Moreover, a number of

reviews have been reported pertaining to the structures of EF-hand containing proteins based on high-resolution structural data (11, 20, 48, 49).

Calbindin D_{28k} is known to fold into a single globular structure as described in Section 1.3. EF-hands very often occur in pairs with the calcium-binding loops forming a three residue anti-parallel β -sheet region (11, 20). The only exception appears to be parvalbumin, which contains one set of paired EF-hands and a single unpaired EF-hand (20). In the case of calbindin D_{28k}, EF1, EF3, and EF5 possess a closely related sequence, as do domains EF2, EF4 and EF6 (20). Thus, it is believed that the six EF-hands arose from two gene duplications and fusion of a two-domain precursor (20, 50). As for the association of EF-hands, Linse *et al.* note that “*In proteins with multiple globular domains, each domain is formed by a contiguous segment of polypeptide, and this rule is fulfilled by all EF-hand proteins with known three-dimensional structures*” (22). The above hypothesis of gene duplication coupled with this observation point toward a structural arrangement whereby EF1 is paired with EF2, EF3 with EF4, and EF5 with EF6. These EF-hand pairs, or subdomains, will be abbreviated as EF12, EF34 and EF56.

In the current structural study, the EF-hands are paired in this manner. Thus the challenge of building a tertiary structure of calbindin D_{28k} is reduced to determining how the EF12, EF34 and EF56 are oriented with respect to each other. The six experimentally derived distance restraints (Table 3.3) are used to eliminate certain configurations or arrangements of the EF-hand subdomains. Finally, a proposed structure is refined using a molecular dynamics (MD) simulation for energy minimization of the system.

4.1.2. Calcium-binding in calbindin D_{28k}

The conformation and stability of calbindin D_{28k} appear to be dependent on whether or not it contains bound calcium (12, 13, 19). The calcium-loaded form is more stable towards urea denaturation (12). Moreover, in fragment complementation studies, it was found that the fragments only interact to form EF-hand-like structures when in the presence of calcium (24). Thus, the current study focuses on the calcium-loaded form of calbindin D_{28k} since the intramolecular forces that govern EF-hand formation are stronger in this form. The methodology used to model the calcium-binding sites is first evaluated with calbindin D_{9k} as a model system, since its structure has been solved in both the calcium-loaded (51) and calcium-free states (52). Calbindin D_{9k} is essentially comprised of 2 EF-hands and simulations on this small protein (76 residues, 1195 atoms) are computationally less demanding than on calbindin D_{28k}.

4.2. Methods

4.2.1. Molecular modeling software

The energy minimization and MD simulations are performed with molecular mechanics in the AMBER 6 package. Molecular mechanics refers to a modeling method that employs classical, or Newtonian mechanics. The interaction energy of a molecular system is described by a so-called force field. AMBER, which is an acronym for Assisted Model Building with Energy Restraints, refers to both the modeling package including a set of programs (53) and a force field. In the AMBER 6 package used in this work the force field comprises bond stretching, bending and torsion, as well as van der

Waals, electrostatic and polarization terms. Hydrogen bonding is treated by modified van der Waals term, though the hydrogen bond term was considered separately in previous versions. Using the AMBER force field, it is possible to refine or “minimize the energy” of a structure by subjecting the system to energy minimization via MD simulations. During a MD simulation the atomic motion is propagated in time, which relaxes unfavorable interactions.

Prior to simulations using AMBER, protein coordinate files were formatted to be recognizable to AMBER. Coordinate files in pdb format were obtained from either the National Center for Biotechnology Information database (<http://www.ncbi.nih.gov>) or from other modeling packages such as HyperChem or DS Viewer Pro. First, the hydrogen atoms contained in the proteins’ pdb files were added and/or checked for correct nomenclature with *PROTONATE*. *TLEAP* was then used to convert the pdb files into topology and coordinate files, which are recognized by *SANDER*, the main module in AMBER used to perform minimization and molecular dynamics simulations. The topology file contains information on the connectivity of atoms, atom names, types and charge as well as residue name of the protein whereas the coordinate file contains only Cartesian coordinates. The *solvateBox* command in *TLEAP* was used to solvate the protein with transferable intermolecular potential 3 point (TIP3P) water molecules (54) in a rectangular box. These are also included in the topology and coordinate files. Parameters for calcium were obtained from <http://pharmacy.man.ac.uk/amber>. These parameters were originally used in the simulation of concanavalin A, another calcium-binding protein (55). All other parameters were from the *all_amino94.lib*,

all_aminoc94.lib and all_aminont94.lib residue descriptions and the Parm99.dat parameter file.

Energy minimization is performed prior to, and following MD simulations using *SANDER*. The first 10 steps of the minimization employed the steepest descent method and the remaining steps were performed with the conjugate gradient method. MD simulations were carried out for 100 or 1000 ps. Each structure was treated with a single MD constant temperature run where the systems were heated from 0 to 300 K, then cooled back to 0 K. Trajectories were propagated using the leap-frog Verlet algorithm. No periodicity was applied during the course of the simulations and *SHAKE* (which removes the bond stretching freedom thus, allowing for larger time steps to be taken) was not employed on the protein. Following the simulations, the output coordinate file generated by *SANDER* and the topology file were used to obtain a pdb format file using *CARNAL*.

4.2.2. Modeling calcium-binding in calbindin D_{9k}

The NMR derived structure of calcium-free calbindin D_{9k} was obtained from PDB ID: 1clb (52). A calcium ion was placed in the calcium-binding loop of EF2. Although EF1 of calbindin D_{9k} binds calcium, it was not loaded in the simulation since the 12 residues making up its calcium-binding loop deviate from the consensus sequence at positions 1(D), 3(N) and 6(G) (Figure 1.3 vs. 4.1). The sidechains of 1 and 3 are calcium ligands (Figure 1.3). The system was solvated with approximately 1100 TIP3P, which represents three solvation shells around the protein. Discrete water molecules rather than a polarizable continuum model were used because the calcium ions in EF-hands can have

0 to 4 water ligands (11). The water molecules were pre-equilibrated with a minimization prior to the MD simulation. A 100 ps MD simulation with a 1 fs time step was then performed. The system was heated from 0 to 300 K over the first ps and cooled to 0 K over the last 30 ps. Increasing the temperature provides more kinetic energy to allow the system to overcome potential energy barriers between conformations. As a result, the system samples more conformational space. A presumably optimal conformation is then obtained by slowly reducing the temperature to 0 K.

4.2.3. Building the EF-hand pairs of calbindin D_{28k}

The residues of each EF-hand of calbindin D_{28k} were substituted onto the backbone of EF1 (residues S2-F36) of calbindin D_{9k} (Figure 4.1). The six EF-hands were constructed separately using HyperChem Office 5.0 software from Hypercube Inc. EF2 of calbindin D_{9k} was not chosen as a template because its amino terminal α -helix is shorter than those of calbindin D_{28k} (Figure 4.1). The sequence of the EF-hands of calbindin D_{28k} as well as their predicted secondary structures are compared in Figure 4.1 to EF1 and EF2 of calbindin D_{9k}. Note that the N-terminal tail (residues 1-15) and the linker loops (Figure 1.3) of calbindin D_{28k} were omitted from the simulations at this point.

The EF-hands were paired with their expected partners to construct EF12, EF34 and EF56 by manually docking the three predicted subdomains using DS Viewer Pro 5.0 from Accelrys Inc. The EF-hands in each pair were linked by inserting the residues of the linker loops to yield the constructs shown in Figure 4.3. Calcium ions were placed in the calcium-binding loops of each EF-hand except for EF2, which does not bind calcium

Calbindin, subdomains	α -helix (E)	Ca ²⁺ binding-loop	α -helix (F)
D _{28k} , EF1:	ASQFFEIWLHF	DADGSGYLEGKE	LQNLIQEL
D _{28k} , EF2:	EMKTFVDQY	GQRDDGKIGIVE	LAHVLPTTE
D _{28k} , EF3:	EFMKTWRKY	DTDHSGFIETEE	LKNFLKDLL
D _{28k} , EF4:	YTDLMLKLF	DSNNDGKLELTE	MARLLPVQE
D _{28k} , EF5:	EFNKAFELY	DQDGNGYIDENE	LDALLKDL
D _{28k} , EF6:	ITTYKKNIM	ALSDGGKLYRTD	LALILCAGD
D _{9k} , EF1:	SPEELKGIFEKYA	AKEGDPNQLSKE	ELKLLLQTEF
D _{9k} , EF2:	LFEEL	DKNGDGEVSFEE	FQVLVKKIS

Figure 4.1. Residues used in modeling the individual EF-hands of calbindin D_{28k}. The corresponding residues in EF-hands of calbindin D_{9k} are also shown.

(14, 24) and EF6, which binds calcium with low affinity (14). The three EF pairs were separately solvated with approximately 1900 water molecules, which represents three solvation shells around each EF-hand pair. The three systems were refined separately using 1000 ps MD simulation with a 1.5 fs time step. The temperature of each system was raised from 0 to 300 K for the first 10 ps and cooled back to 0 K over the last 100 ps.

Following the refinement of the three EF-hand pairs, the full-length protein was constructed. The fragments were positioned in different configurations using DS Viewer Pro, and the experimentally derived distance restraints (Table 3.3) were monitored during this process. The K59-K221 and K59-K223 distance constraints represent a special case, because it is not clear from the MS analysis which set of lysines is cross-linked. Accordingly, proposed structures were rejected only if both the K59-K221 and K59-K223 distances exceeded the 24 Å constraint. First, possible EF12 and EF56 configurations were created and five of these respected the constraints. These five structures were then coupled with EF34 to generate configurations of the full protein. All configurations that were considered can be found in Appendix C.

Finally, the N-terminal tail and the linker loops that connect the EF-hand pairs were added using HyperChem. The full-length calbindin D_{28k} was also subjected to MD refinement. Because simulations on solvated systems of this size are computationally taxing, a vacuum calculation was performed. Only water molecules that were observed to ligate to the calcium ions were considered. Refinement of the structure was performed using a 1000 ps MD simulation with a 1.5 fs time step. The temperature of the system was raised from 0 to 300 K for the first 50 ps and cooled back to 0 K over the last 100 ps. Over the course of a MD simulation, unfavorable steric interactions sometimes occur resulting in unnaturally high velocity of atomic particles. A velocity limit (VLIMIT) for atomic trajectories of 20 Å/ps was therefore imposed (the average velocity as described by the Maxwell-Boltzmann relationship is approximately 5 Å/ps).

4.3. Results

4.3.1. Calcium-binding in calbindin D_{9k}

A calcium ion was placed into the calcium-binding loop of EF2 of calbindin D_{9k}. This system was refined by minimization and MD simulation and compared to the crystal structure of the calcium-loaded form of the protein (PDB ID: 4icb) (51). The predicted coordination sphere is shown in Figure 4.2, and the observed and predicted calcium–oxygen distances are listed in Table 4.1. Calcium and oxygen atoms separated by more than 3.0 Å are not considered to interact with each other in the calcium-binding complex (11). The modeled structure shows only small differences with the crystal structure, except that the backbone carbonyl oxygen ligating the calcium ion is predicted to be that

of the conserved G57 rather than N56 found in the crystal structure. Thus, the predicted calcium-binding model does not agree with the consensus binding mode shown in Figure 1.3.

Table 4.1. Comparison between predicted and X-ray derived calcium-oxygen distances in EF2 of calbindin D_{9k}

Residue	X-ray ^a <i>r</i> , Å	Model <i>r</i> , Å
D54	2.4	2.2
N56	2.3 ^b	6.3
G57	6.5	2.3 ^b
D58	2.4	2.5
E60	2.4 ^b	2.3 ^b
E65	2.5	2.3
E65	2.5	2.3
H ₂ O	2.5	2.4

^a X-ray data based on the 1.6 Å resolution structure of calbindin D_{9k} (PDB ID: 4icb).

^b The coordinating oxygen is that of the backbone carbonyl.

4.3.2. Modeled EF-hand pairs of calbindin D_{28k}

EF-hands were constructed individually using the backbone of EF1 of calbindin D_{9k} (Figure 4.1) as a template. A calcium ion was placed in the binding pockets of EF1, EF3, EF4 and EF5. The literature appears to be divided on the calcium-binding capabilities of EF6 (14, 23, 24). The primary structure of the calcium-binding loop of EF6 does not respect the consensus sequence of an EF-hand (Figures 1.2 and 1.3) and so the oxygen atoms are presumably not ideally positioned to chelate calcium. Thus, the ion was omitted from EF6 in the simulations of full-length calbindin D_{28k}. The EF-hands were paired and subjected to MD treatment for refinement. The resulting structures are shown in Figure 4.3.

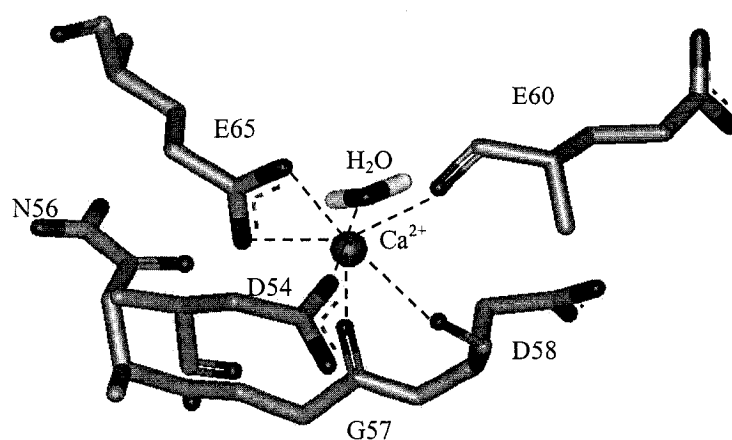


Figure 4.2. Predicted calcium coordination in EF2 of calbindin D_{9k}. Note that the backbone carbonyl of N56 is found to coordinate the calcium ion in the X-ray structure whereas the MD simulation predicts coordination to the carbonyl of G57 as shown here.

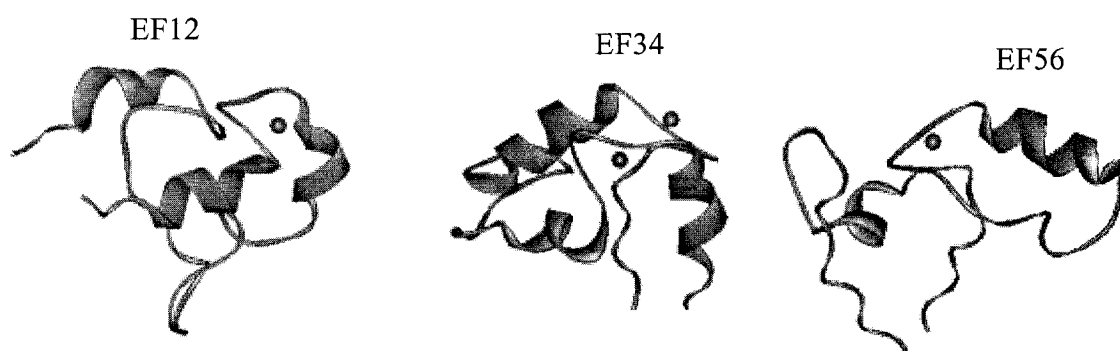


Figure 4.3. Ribbon diagram of predicted structures of the three EF-hand pairs of calbindin D_{28k}. The RMSD between the initial and final structures are 1.3 Å, 1.8 Å and 2.1 Å for EF12 (residues 13-86), EF34 (residues 102-175) and EF56 (residues 190-261), respectively.

4.3.3. Modeled full-length calbindin D_{28k}

The three free EF-hand pairs were positioned in a number of different configurations. The relative orientations of the EF-hands are represented by simplified diagrams as illustrated in Figure 4.4.

Figure 4.5 shows all structures that respect the distance constraints from the cross-linking experiments (Table 3.3). It is not possible to identify a single orientation of the

EF-hands based on the six cross-links found. K72, K235 and K246 are solvent accessible in the structures in Figure 4.5 consistent with the published limited proteolysis work (Section 1.4). Since time did not permit a thorough investigation of all the structures in Figure 4.5, a representative structure [Figure 4.5 (a)], was chosen for refinement but, based on all available experimental data, the other 13 structures are equally valid.

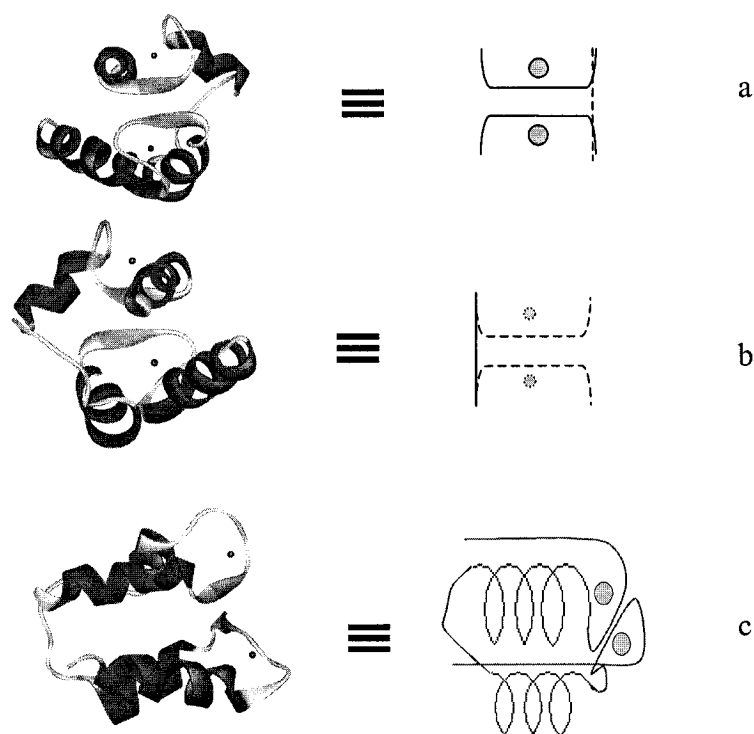


Figure 4.4. Representations of EF-hands. The two EF-hands in a pair are related by an approximate two-fold rotation axis at the center of the pair (11,49) such that both calcium-binding loops are in the same plane. (a) Calcium-binding loops above the plane of the page. The calcium ion is outlined by a solid line and the linker is shown as a dashed line. (b) Calcium-binding loops below the plane of the page. The calcium ion is outlined by a dashed line and the linker loop shown as a solid line. (c) Calcium-binding loop in the plane of the page.

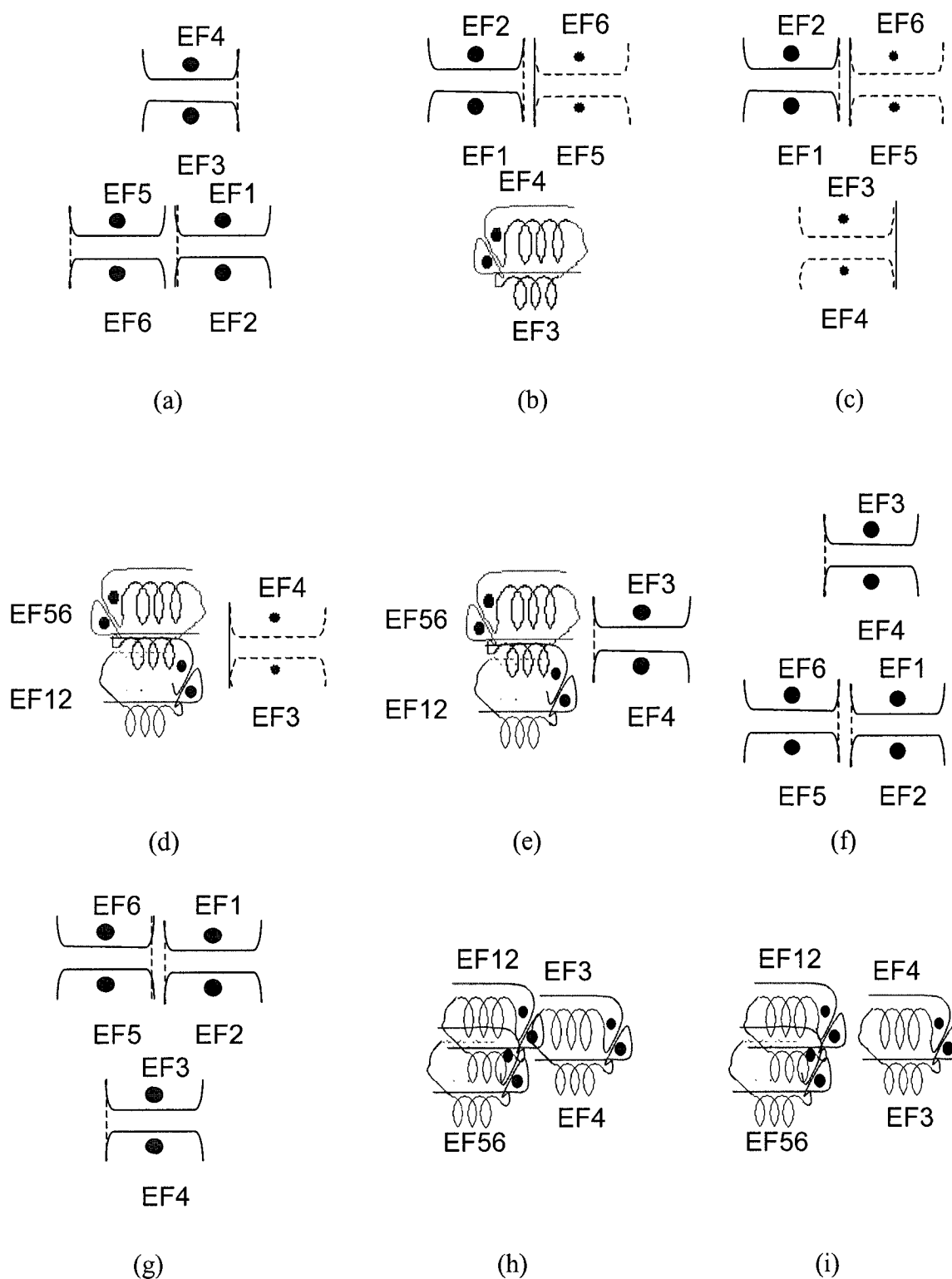


Figure 4.5. The 14 EF-hand configurations that are consistent with the experimentally derived distance restraints. See Figure 4.4 for notation. These configurations were constructed using DS Viewer Pro 5.0. Configuration (a) was selected for further refinement using AMBER. The atomic coordinates for these models are available in the attached CD.

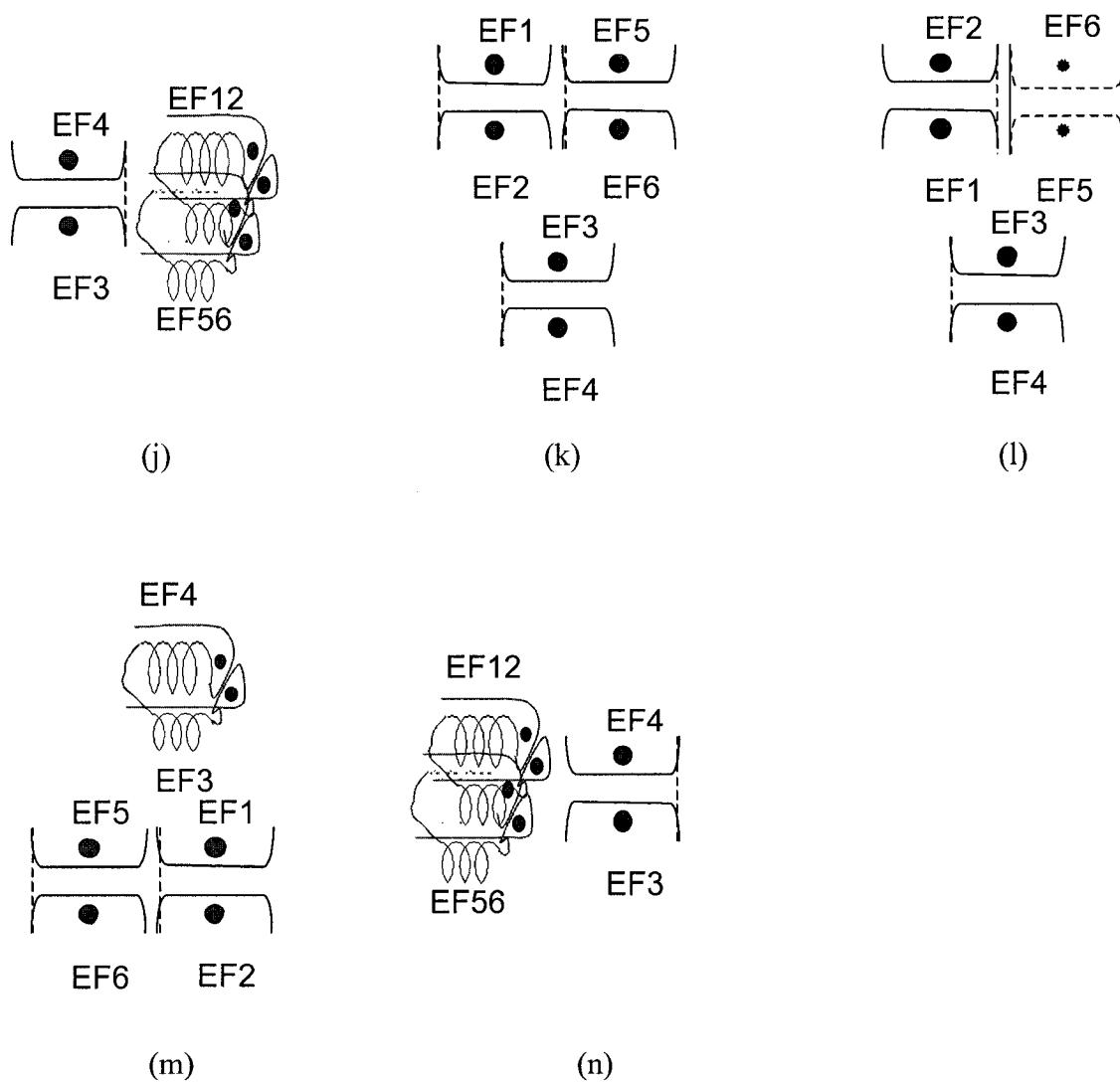


Figure 4.5. Continued.

4.3.4. Refinement of a full-length calbindin D_{28k} structure

The energetics of the system was monitored over the course of the MD run

(Figure 4.6). The kinetic energy is given by $E_{\text{kin}} = \frac{mv^2}{2}$ and the potential energy by

$$E_{\text{pot}} = E_{\text{stretch}} + E_{\text{bend}} + E_{\text{torsion}} + E_{\text{van der Waals}} + E_{\text{electrostatic}} + E_{\text{polarization}}.$$

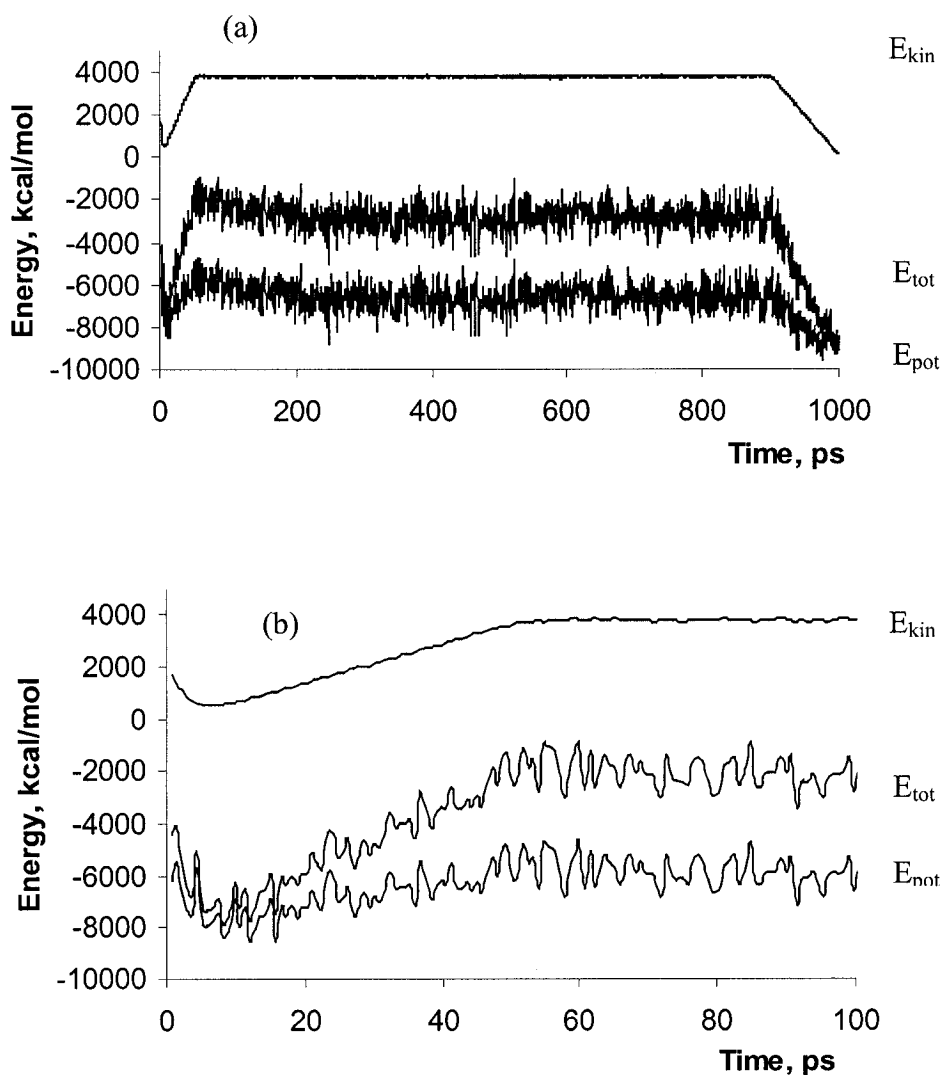


Figure 4.6. Energy fluctuation of calbindin D_{28k} during the 1000 ps MD simulation
 (a). (b) An expanded region of the first 100 ps. The system was heated from 0 to 300 K over a period of 50 ps and cooled to 0 K for the last 100 ps. The decrease in energies during the first 15 ps indicates that the initial structure results in a number of unfavorable interactions which are relaxed very early in the simulation.

The structure of the protein at given time steps was compared (Figure 4.7). The distances separating the lysine residues that were observed to be cross-linked are compared in the initial and final structures in Table 4.2, while the calcium coordination spheres of the final (refined) structure are described in Table 4.3.

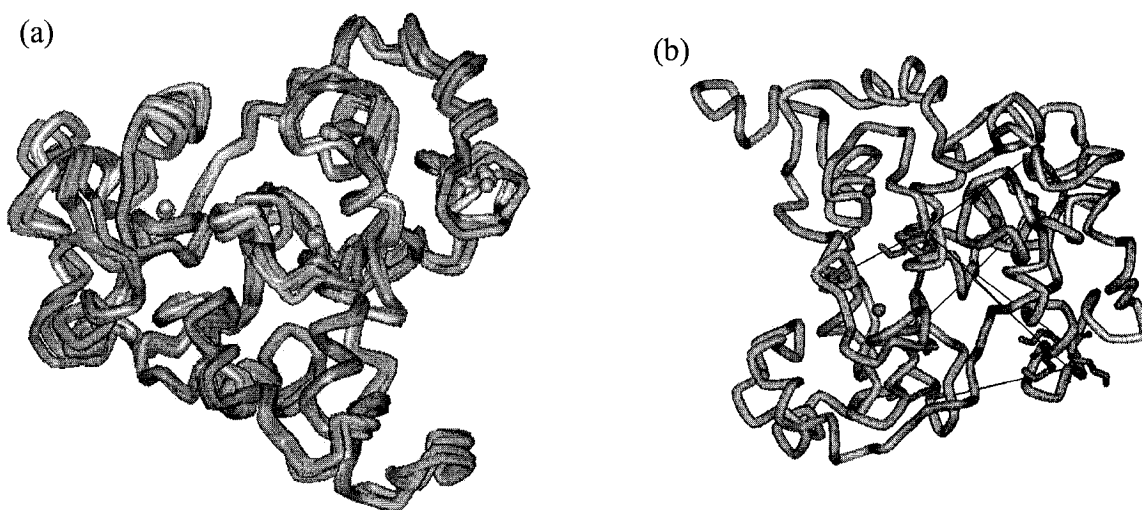


Figure 4.7. Refined structure of calbindin D_{28k}. (a) Superimposed backbone snapshots of a calbindin D_{28k} structure during the MD simulation. Starting with the model of calbindin D_{28k} shown in Figure 4.5(a), structures were recorded every 200 ps over the course of the 1 ns simulation. The initial and final structures are within 1.9 Å RMSD. Most variation occurs in the linker loops. (b) Refined structure of calbindin D_{28k} showing the cross-linked lysine residues. The atomic coordinates for the refined model is available in the attached CD.

Table 4.2. C_α separations of lysine residues of calbindin D_{28k} that were observed to be cross-linked

Cross-linked residues ^a	Lysine C _α separations (Å) ^b		Δ (Å) ^d	Regions cross-linked ^c
	Initial structure ^b	Refined structure ^c		
K34-K124	15	15	0	Ca1- α F3
K34-K216	22	20	-2	Ca1- α F5
K34-K246	16	16	0	Ca1-Ca6
K48-K133	15	19	+4	L12-L34
K59-K221	27	27	0	α E2-L56
K59-K223	23	23	0	α E2-L56
K142-K221	17	22	+5	L34-L56

^a From Table 3.3.

^b Separation between C_α of BS3-cross-linked lysines in the initial structure are within the required 24 Å except for K59-K221, which may not be a cross-link (See Table 3.3).

^{c,d} Separations following MD simulation and the difference from the initial values.

^c CaX, calcium-binding loop X; α EX, α -helix E in EFX; α FX, α -helix F in EFX; LXY, linker between EFX and EFX.

Table 4.3. Predicted coordination sphere and calcium-oxygen distances (r) in the modeled calcium-binding loops of calbindin D_{28k}

EF1		EF3		EF4		EF5	
Residue	r , Å	Residue	r , Å	Residue	r , Å	Residue	r , Å
D24	2.2	D111	2.2	D155	2.3	D199	2.2
D24 ^a	2.3	D113	2.2	D155	2.3	D201	2.2
D26	2.2	I118 ^a	2.4	N157	2.4	D201	2.4
L31	2.4	E122	2.3	D159	2.2	N203	2.4
E35	2.2	H ₂ O	2.3	L162 ^a	2.3	Y205 ^a	2.3
E35	2.4			E166	2.2	E210	2.4
H ₂ O	2.4			E166	2.4	E210	2.3

^a Coordinating oxygen is on the backbone carbonyl

4.4. Discussion

4.4.1. Screening the calbindin D_{28k} structures

The six EF-hands of calbindin D_{28k} were constructed using the backbone of EF1 of calbindin D_{9k} as a template. EF1 of calbindin D_{28k} was then paired with EF2, EF3 with EF4 and EF5 with EF6 so that the calcium-binding loops were juxtaposed. The structures of the three isolated EF-hand pairs were refined by MD simulations (Figure 4.3).

The only available structure to compare with the models is the NMR-derived structure of residues 1-93 representing EF12 of calbindin D_{28k} (47). The RMSD, calculated with Deep View / Swiss Pdb Viewer 3.7, GlaxoSmith Kline (<http://www.expasy.org/spdbv>), between the modeled and NMR structure is 6.9 Å for residues A13-L43, E57-E86. Most deviation occurs in the F helix of EF2. However, a detailed comparison between the NMR structure and the modeled structure may not be meaningful. The atomic coordinates for residues 1-93 obtained from Dr. Werner Klaus

(attached CD) for the NMR structure are not those used in Reference (47), and it is not known if the NMR of the calcium-loaded or calcium-free form was recorded. Moreover, the structure of the free EF12 subdomain may vary from that in the native full-length protein.

The refined EF-hand pairs were used to assemble full-length calbindin D_{28k}. Linear structural arrangements (Figure 4.8) violated at least three of the lysine distance restraints in Table 3.3. This is in agreement with the experimental findings of Linse *et al.* (19, 21, 22), which showed that each EF-hand interacts with several others in the protein (Section 1.3).

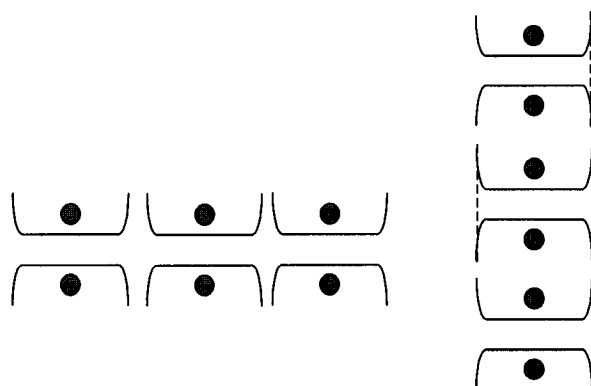


Figure 4.8. Linear arrangements of the EF-hand pairs of calbindin D_{28k}.

4.4.2. Refinement of the proposed structure of human calbindin D_{28k}

Of the 103 configurations of the three EF-hand pairs in Appendix C, 14 respected the distance restraints (Figure 4.5), and one of these was chosen for refinement. In the chosen structure, K59 and K221 are separated by more than 24 Å (Table 4.3) but this does not necessarily mean that the structure is invalid as mentioned in Section 4.2.3. The refinement resulted in little or no changes in the separation of cross-linked lysine residues

except those involved in cross-linking two linker regions of the protein. The greatest difference was observed in the K48-K133 and K142-K221 separations, which increased by 4 and 5 Å, respectively. Because these four residues are in linker regions L12, L34, L34 and L56 (Table 4.2), their movement within the 1 ns simulation indicates that these regions are highly flexible.

MD simulation results show that the system is initially high in energy. During the first 8 ps of the simulation, unfavorable interactions in the initial structure are relaxed, resulting in a total energy decrease of 3300 kcal/mol (Figure 4.6). The increase in energy after 15 ps is due to the ramping of the temperature from 0 to 300 K, and the decrease in energy after 900 ps is due to cooling of the system to 0 K. The system otherwise exhibited little change in energy over the course of the MD simulation. The MD refinement resulted in a 5800 kcal/mol reduction in total energy, corresponding to an energy reduction of 1.4 kcal/mol per atom. From visual inspection, the superimposed structures of the protein in Figure 4.7 do not show significant fluctuations. The initial and the final structures differ by a RMSD of 1.9 Å. The small structural change of the tertiary structure over the 1 ns simulation however is not surprising given that the time required to fold a protein is in the 0.1 to 1000 s timescale (56).

One focus of our research group is the S-nitrosation of the cysteine residues of calbindin D_{28k} (Section 2.4). Thus, the environment of the cysteine residues and how this may affect their reactivity are of interest. In the refined model, all cysteines are solvent exposed. From Figure 1.2, C257 is expected to be in α -helix F of EF6, and C94, C100, C187 and C219 are found in linker regions L23, L23, L45 and L56, respectively. The cysteines are too far apart to form intramolecular disulfide bridges based on the S-S

distances given in Table 4.4. However, it would be premature to draw conclusions from the modeled structure as to the feasibility of intramolecular disulfide bridges, which have been reported (17).

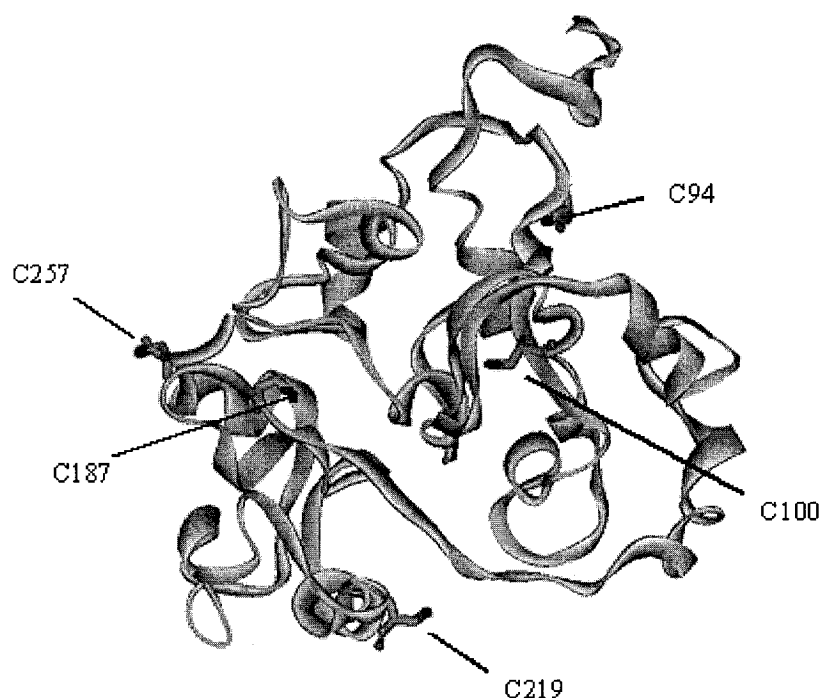


Figure 4.9. Ribbon diagram of refined structure of calbindin D_{28k} showing the positions of the cysteine residues. The SH groups are in black.

Table 4.4. Distances (Å) separating the sulfur atoms of the cysteine residues in the modeled structure of calbindin D_{28k}^a

Cys ^b	C100	C187	C219	C257
C94	16	25	33	33
C100		16	20	32
C187			20	18
C219				34

^a Structure shown in Figure 4.9.

^b C94 and C100 are in L23; C187 is in L45; C219 is in L56 and C257 is in α -helix F of EF6 (Figure 1.2). Here LXY refers to the linker region between EFX and EFY.

4.4.3. Modeling the calcium coordination sites

Calcium-binding in calbindin D_{9k}

To evaluate the feasibility of modeling the calcium-binding sites of calbindin D_{28k}, the method to be used was first tested on calbindin D_{9k}. A single calcium ion was placed in EF2 of calbindin D_{9k}. The simulations were performed without the use of restraints to avoid biasing the modeled coordination sphere toward the expected geometry. The modeled structure is in good agreement with the X-ray structure of the calcium-loaded form of the protein except that the backbone carbonyl of G57 acts as a ligand rather than that of N56 as observed in the crystal structure (Table 4.1).

Ligation of calcium to the backbone carbonyl of G57 was also observed by Kordel *et al.* who performed a similar MD simulation on calbindin D_{9k} using NMR derived coordinates (57). Restricted MD simulations were performed using constraints derived from nuclear Overhauser effect (NOE) values obtained from the NMR analysis. NOE data give throughspace distances between nuclei. In the NMR analysis of calbindin D_{9k}, more constraints were available for the calcium-binding loop of EF1 compared to EF2, and the modeled calcium coordination sphere of EF1 contained the same ligands as the crystal structure. The discrepancy in the binding loop of EF2 was therefore attributed to a lack of NMR derived restraints. This points to the importance of using a well-defined scaffold to accurately model the calcium coordination sphere of an EF-hand.

The model agreed well with the X-ray structure (Table 4.1) for the other six chelating oxygens of calbindin D_{9k} EF2. Thus, the approach of placing a calcium ion into the binding pocket prior to MD simulation was employed to model the calcium-binding sites of the EF-hands of calbindin D_{28k}.

Calcium-binding in EF-hand pairs of calbindin D_{28k}

The isolated calcium-loaded EF-hand pairs of calbindin D_{28k} were subjected to a MD simulation in the presence of explicit water molecules. Only the water molecules that were observed to ligate calcium were retained for the vacuum simulation of full-length calbindin D_{28k}. Following a 1000 ps MD simulation of full-length calbindin D_{28k}, the coordination spheres around calcium were analyzed. All calcium-oxygen distances in the model were found between 2.2 and 2.4 Å (Table 4.3). Observed calcium-oxygen distances in EF-hands are in the range of 2.0 to 3.0 Å with an average of 2.4 Å (11). Heptacoordination was observed for EF1, EF4 and EF5, and pentacoordination for EF3. Penta- to octacoordinated calcium is observed in EF-hands, but heptacoordination is most common (11).

The low coordination number may give rise to the low calcium affinity of EF3 found by NMR spectroscopy (14). It is noteworthy that EF3 does not follow the consensus EF-hand sequence (Figures 1.2 and 1.3). Residue 5 of the calcium-binding loop of EF3 is serine rather than the conserved asparagine or aspartic acid (Figure 1.3 and Table 4.3). The sidechain of residue 5 is a consensus calcium ligand, and that of serine (approximately 2.4 Å vs. 3.2 Å for asparagine and aspartic acid) in EF3 may be too short for its oxygen to interact with calcium. Moreover, residue 4 of the EF3 calcium-binding loop is histidine, whereas glycine is found in the consensus sequence at this position (Figure 1.3) and is believed to give the loop more flexibility, allowing for its proper folding (4). The lower than expected coordination number in EF3 may also be partly due

to an artifact in the simulation protocol. Since a vacuum calculation was performed, no additional water molecules were available to coordinate to the calcium ions.

EF1 follows the consensus sequence except that serine is also substituted for asparagine/aspartic acid at position 5 (Figure 1.3 and Table 4.5). Like EF3, the predicted coordination sphere does not show serine coordination to the calcium. The heptacoordinated sphere is completed by D24, which binds calcium *via* its carboxylate sidechain and backbone carbonyl.

Table 4.5. Predicted calcium coordination spheres of calbindin D_{28k} vs. consensus ligands

Position ^a	1	2	3	4	5	6	7	8	9	10	11	12
Consensus ^{b,c} :	<u>Asp</u>	-	<u>Asx</u>	-	<u>Asx</u>	Gly	<u>I/L/E</u>	-	-	-	-	<u>Glu</u>
EF1:	<u>Asp₂₄</u>	Ala	<u>Asp</u>	Gly	<i>Ser</i>	Gly	Tyr	<u>Leu</u>	Glu	Gly	Lys	<u>Glu</u>
EF2:	<i>Gly₆₈</i>	Met	<i>Lys</i>	Thr	Asp	Gly	<i>Lys</i>	Ile	Gly	Ile	Val	Glu
EF3:	<u>Asp₁₁₁</u>	Thr	<u>Asp</u>	His	<i>Ser</i>	Gly	Phe	<u>Ile</u>	Glu	Thr	Glu	<u>Glu</u>
EF4:	<u>Asp₁₅₅</u>	Ser	<u>Asn</u>	Asn	Asp	Gly	<i>Lys</i>	<u>Leu</u>	Glu	Leu	Thr	<u>Glu</u>
EF5:	<u>Asp₁₉₉</u>	Gln	<u>Asp</u>	Gly	Asn	Gly	<u>Tyr</u>	Ile	Asp	Glu	Asn	Glu
EF6	<i>Ala₂₄₂</i>	Leu	<i>Ser</i>	Asp	<i>Gly</i>	Gly	<i>Lys</i>	Leu	Tyr	Arg	Thr	<i>Asp</i>

^a Position of residue in calcium-binding loop

^b Residues that are underlined are observed to coordinate calcium as shown in Table 4.3.

^c Residues in italics are not consistent with the consensus sequence of a calcium-binding site.

The calcium binding loops of EF4 and EF5 possess the consensus sequence (Figure 1.3 and Table 4.3). The ligating residues predicted by the consensus sequence coordinate calcium in EF5. In the model, EF4 coordinates calcium *via* residue 8 rather than 7 or 9 as predicted by the consensus sequence. Again, the differences in the predicted and observed calcium coordination spheres indicate that the binding loop conformations do not reflect those found in calbindin D_{28k}. Kordel *et al.* (57) noted that defects in the

modeled binding site of calbindin D_{9k} result from lack of structural constraints in defining the loop conformation.

5.0. Summary and suggested further studies

5.1. Summary

Chapter 2

The JM105-pTrc99A-His-calb system was successfully reconstituted starting from the pET15b-His-calb plasmid. The system produced 12 to 15 mg of His-calb per litre of JM105 bacterial culture. MAC purification was employed to obtain approximately 84% purity. Following the removal of the His-tag, the yield of recombinant calbindin D_{28k} was 9 mg per litre of JM105 culture. The purity following a second MAC purification treatment was approximately 93%.

Chapter 3

Myoglobin was used as a model protein to study chemical cross-linking with BS3. The cross-linking reaction was optimized to minimize dimerization and to obtain a single BS3 modification per myoglobin molecule. The reaction products consisted of a mixture of unmodified myoglobin, cross-linked myoglobin and myoglobin modified at a single lysine residue. Tryptic digest of the reaction products followed by MALDI-MS analysis yielded six uniquely identified cross-linked peptides. The crystal structure of myoglobin shows that the six pairs of cross-linked lysine residues are sufficiently close to form cross-links.

Recombinant calbindin D_{28k} was also cross-linked with BS3. Again, little dimerization and a single modification per calbindin D_{28k} molecule was obtained. GluC digestion and GluC-trypsin double digestion of the reaction product followed by MS

analysis revealed six cross-linked peptides. The MS signals corresponding to the cross-linked peptides exhibited low signal-to-noise ratios, probably because of the low abundance of individual species.

Chapter 4

The EF-hands of calbindin D_{28k} were constructed based on the backbone of EF2 of calbindin D_{9k}, and paired with their expected natural partner. Models of full-length calbindin D_{28k} were constructed and the distance constraints from the chemical cross-linking experiments were monitored. Of the 103 configurations considered (Appendix C), only 14 respected the distance restraints. One representative model of the 14 was chosen and refined by computer simulation of full-length calbindin D_{28k}. Little change in the tertiary structure was observed following a 1 ns MD simulation.

The ability of MD simulations to predict the correct insertion of a calcium ion into an empty calcium-binding loop was evaluated with calbindin D_{9k}. The modeled coordination sphere around the calcium ion agreed reasonably well with the crystal structure. However, there was a discrepancy between the predicted calcium ligation through G57 and N56 observed in the crystal structure. The refinement of the calbindin D_{28k} structure was performed on the calcium-loaded form. The calcium coordination spheres of calbindin D_{28k} showed 3 heptacoordinated and 1 pentacoordinated calcium ions.

5.2. Future study

1. To thoroughly exploit the cross-linking reaction more distance restraints in the form of cross-linked peptides could be obtained. LC-ESI-MS of a digest of cross-linked protein should be obtained. A possible reason for the low detection of different cross-linked species is that peptides are not efficiently ionized by MALDI. ESI relies on a different ionization mechanism and has been shown by others to detect cross-linked peptides that were not observed by MALDI (32).
2. Another alternative is to use a library of cross-linkers that are of different length or specificity.
3. Not all MALDI-MS peaks could be uniquely identified and so no distance restraint information could be gained. Tandem mass spectrometry, either LC-ESI-MS/MS or MALDI-post-source-decay should be used to confirm the sequence of the peptides.
4. The 14 proposed calbindin D_{28k} structures that respect all distance restraints obtained from the cross-linking experiments differ in the solvent accessibility of the protein residues. H/D exchange, an approach that probes surface accessibility of residues, could be performed. The data could be coupled with the cross-linking data, which should narrow down the number of possible configurations of the EF-hands.
5. Although there is interest in applying chemical cross-linking to obtain structural information, it is not clear if and how the cross-linker affects the protein structure. To date, it has been assumed that a single modification will not significantly perturb the structure of the protein, but this has not yet been shown experimentally. To demonstrate the effects of cross-linking, fluorescence and near UV-CD spectroscopies could be used to probe changes in the environment around aromatic

residues. Alternatively, MD simulation of a cross-linked protein could be performed. From this, it may be possible to determine if the cross-linker remains on the protein surface or if it gets buried in the core of the protein, which would conceivably distort the protein structure.

Appendix A

Solutions used for agarose analysis

5 X Tris-borate-EDTA(TBE) buffer

0.5 M EDTA (pH 8.0)	20 mL
Tris base	54 g
Boric acid	27.5 g
H ₂ O	Fill to 1 L

Solutions used for SDS-PAGE analysis

4 % stacking gel

30 % acrylamide/Bis	0.66 mL
0.5 M Tris HCl, pH 6.8	1.26 mL
10 % SDS	50 μ L
H ₂ O	3 mL
TEMED	5 μ L
10 % APS	25 μ L

12 % resolving gel

30 % acrylamide/Bis	4 mL
1.5 M Tris HCl, pH 8.8	2.5 mL
10 % SDS	100 μ L
H ₂ O	4 mL
TEMED	5 μ L
10 % APS	50 μ L

10 X glycine buffer, pH 8.3

Glycine	144 g
Tris base	30 g
SDS	10 g
H ₂ O	1 L

Sample buffer

H ₂ O	4.8 mL
0.5 M Tris HCl, pH 6.8	1.2 mL
Glycerol	1 mL
10 % SDS	2 mL
0.1 % Bromophenol blue	0.5 mL

Appendix B

Primary structure of recombinant human calbindin D_{28k} (6, 7). The residues marked in italics are not present in the native protein. The 25 lysine residues (K) are underlined.

ASMAESHLQSSLITASQFFEIWLFHDADGSGYLEGKELQNLIQELQQARKKAGLE
LSPEMKTFVDQYGQRDDGKIGIVELAHVLPTEENFLLLFRCCQLKSC~~EE~~FMKTW
RKYDTDHSGFIETEELKNFLKDLLEKANKTVDDTKLAEYTDLMLKLFDSNNDGK
LELTEMARLLPVQENFLLKFQGIKMCGKEFNKAFELYDQDGNGYIDENELDALL
KDLCEKNKQDLNINITYKNIMALSDGGKLYRTDLALILCAGDN

Amino acid composition of recombinant human calbindin D_{28k}^a

Amino acid	Number of residues
Alanine	15
Arginine	6
Asparagine	15
Aspartic acid	23
Cysteine	5
Glutamine	14
Glutamic acid	25
Glycine	16
Histidine	4
Isoleucine	12
Leucine	40
Lysine	25
Methionine	7
Phenylalanine	14
Proline	3
Serine	11
Threonine	13
Tryptophan	2
Tyrosine	8
Valine	5

^a Includes the additional residues (AS) at the N-terminus that are not present in the native protein. Data from References (6, 7)

Primary structure of horse heart myoglobin (44). The 19 lysine residues (K) are underlined.

GLSDGEWQQVLNVWGKVEADIAGHGQEVLIRLFTGHPETLEKFDKFKHLKTEAE
 MKASEDLKKHGTVVLTALGGILKKKGHHEAELKPLAQSHATKHKIPKYLEFISD
 AIIHVLHSKHPGDFGADAQGAMTKALELFRNDIAAKYKELGFQG

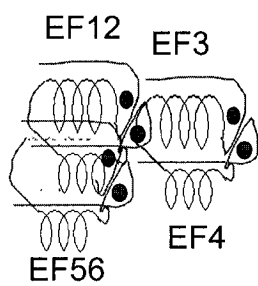
Amino acid composition of horse heart myoglobin^a

Amino acid	Number of residues
Alanine	15
Arginine	2
Asparagine	2
Aspartic acid	8
Cysteine	0
Glutamine	6
Glutamic acid	13
Glycine	15
Histidine	11
Isoleucine	9
Leucine	17
Lysine	19
Methionine	2
Phenylalanine	7
Proline	4
Serine	5
Threonine	7
Tryptophan	2
Tyrosine	2
Valine	7

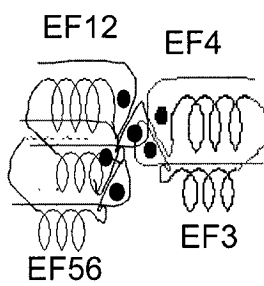
Data from Reference (44).

Appendix C

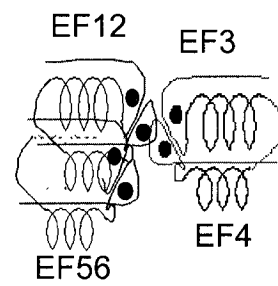
The three EF-hand pairs of human calbindin D_{28k} were arranged in the configurations shown below. The number of violations refers to the number of lysine-lysine separations that are greater than the 24 Å dictated by the six distance constraints given in Table 3.3.



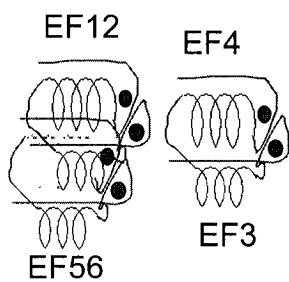
0 violation



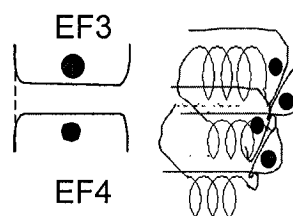
2 violations



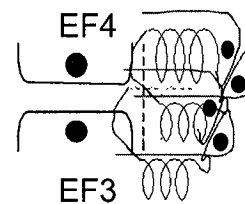
2 violations



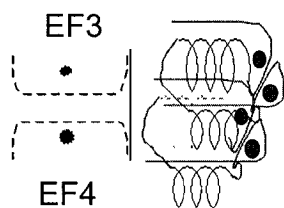
0 violation



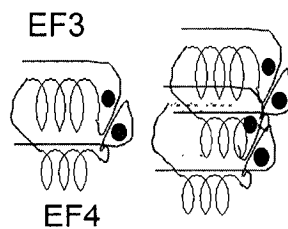
2 violations



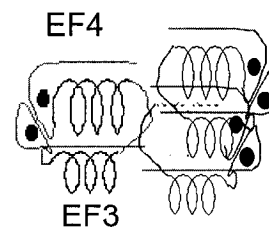
0 violation



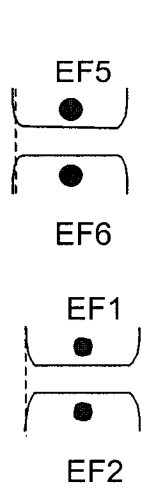
2 violations



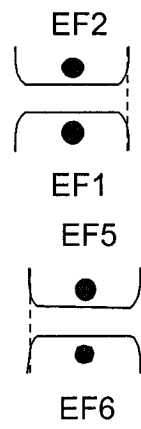
3 violations



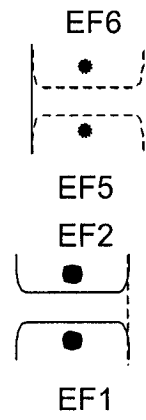
1 violation



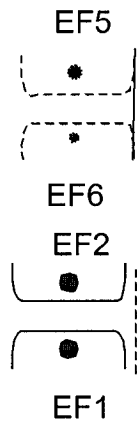
2 violations



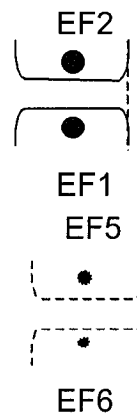
0 violation



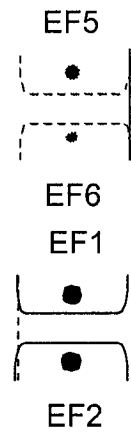
2 violations



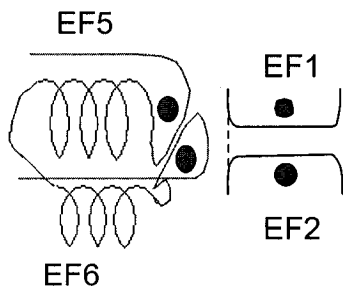
3 violations



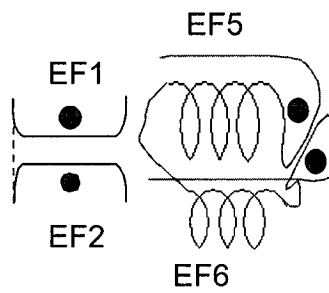
2 violations



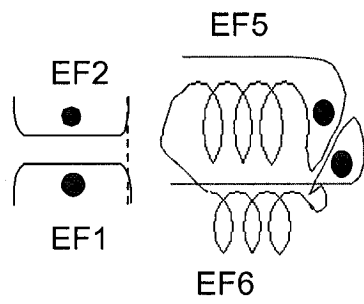
2 violations



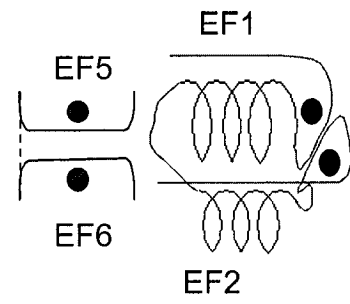
1 violation



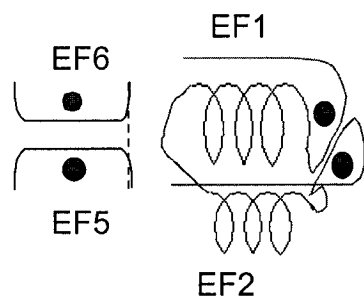
2 violations



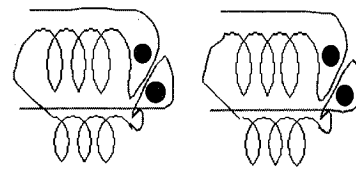
1 violation



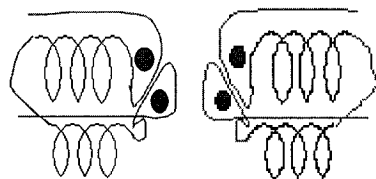
2 violations



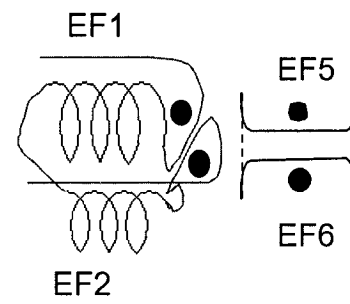
2 violations



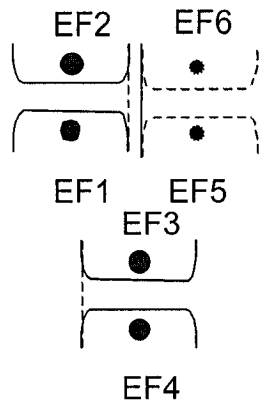
≥ 1 violation



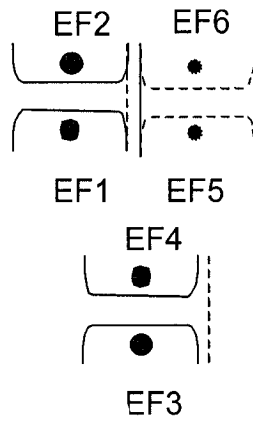
≥ 2 violations



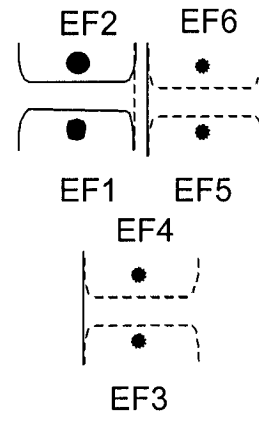
1 violation



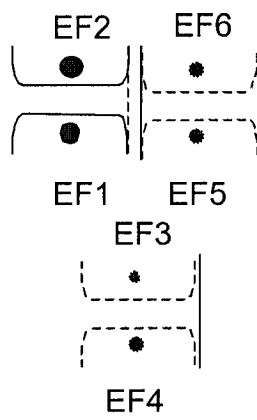
0 violation



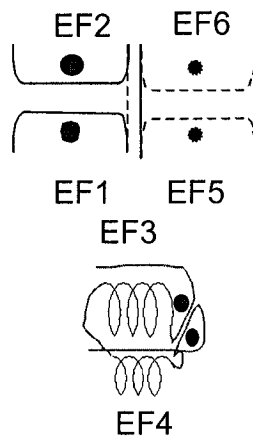
3 violations



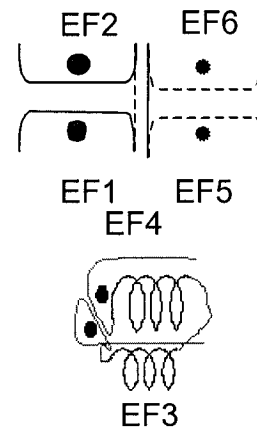
1 violation



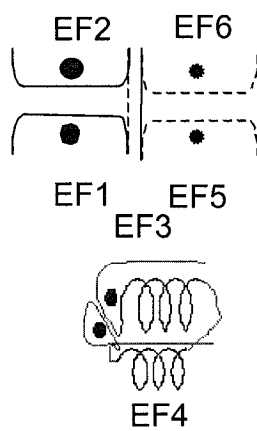
0 violation



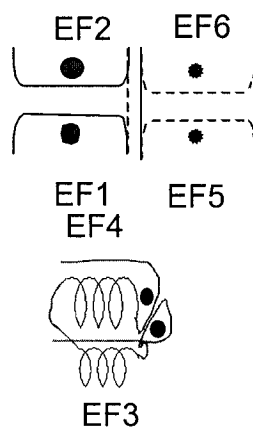
1 violation



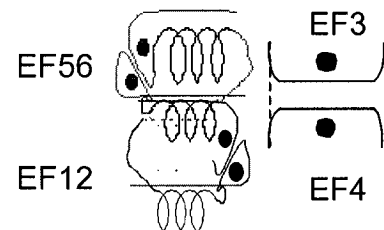
0 violation



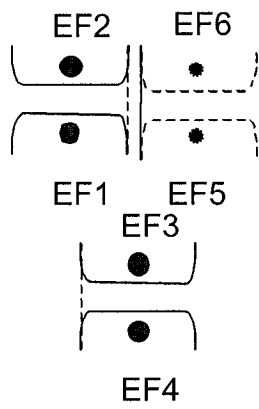
2 violations



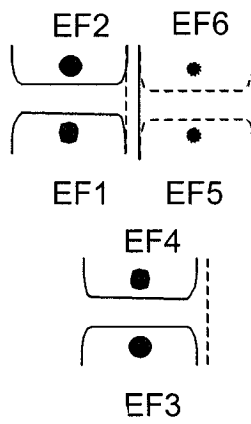
2 violations



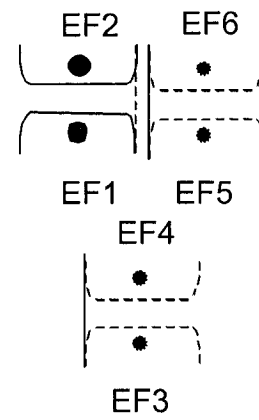
0 violation



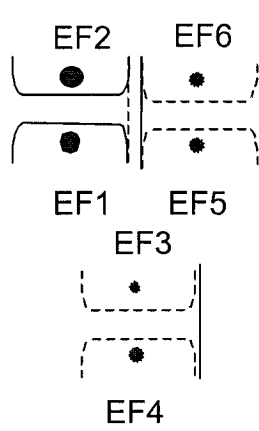
0 violation



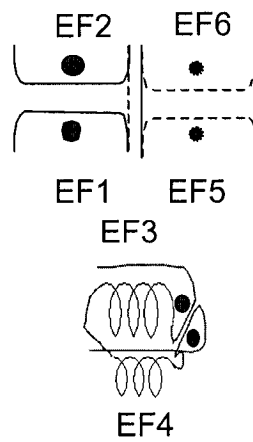
3 violations



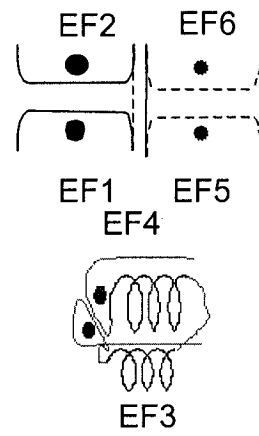
1 violation



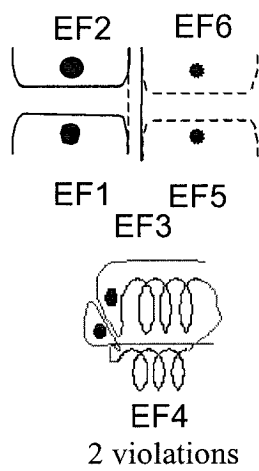
0 violation



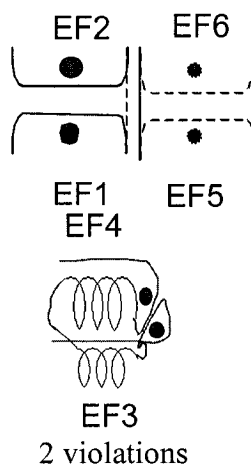
1 violation



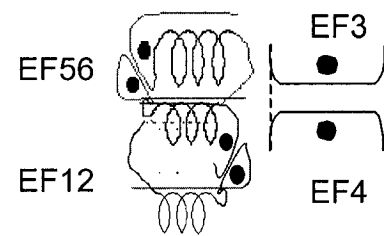
0 violation



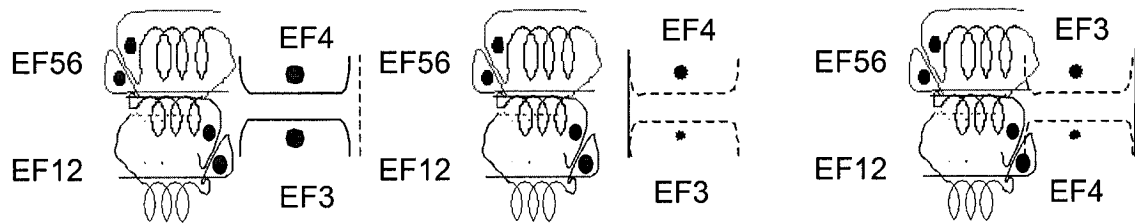
2 violations



2 violations



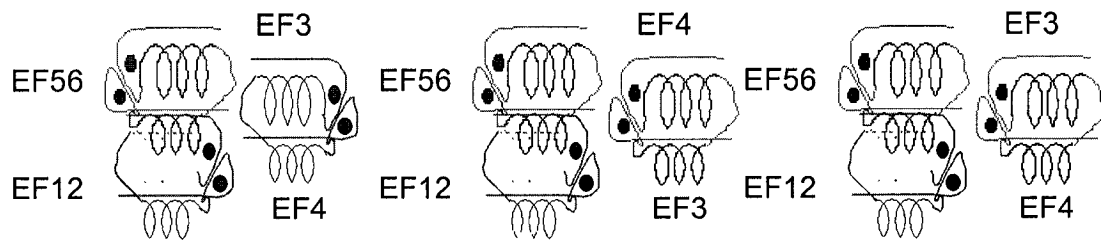
0 violation



1 violation

0 violation

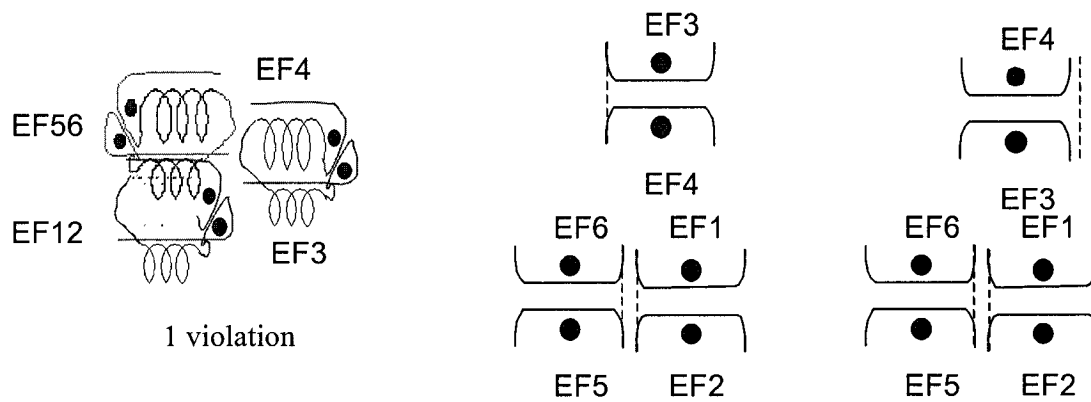
1 violation



1 violation

2 violations

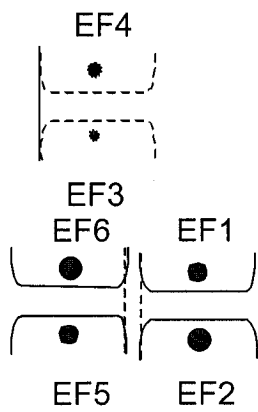
2 violations



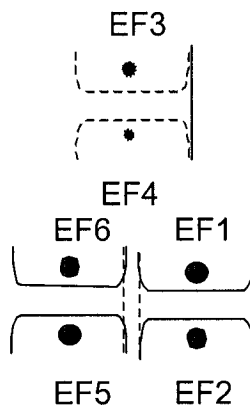
1 violation

0 violation

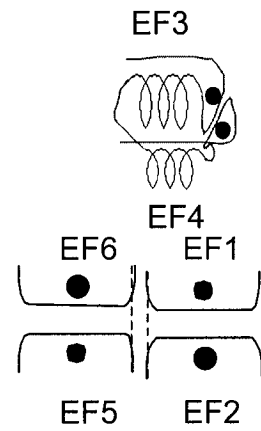
2 violations



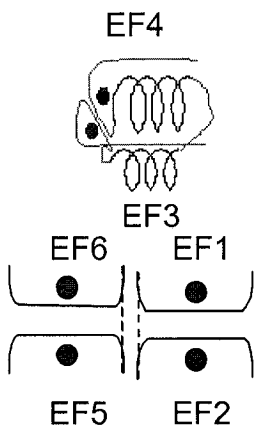
1 violation



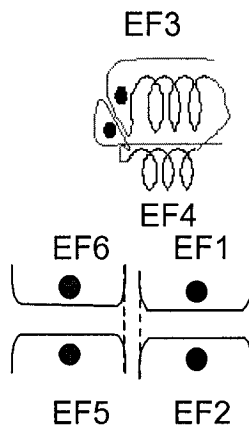
3 violations



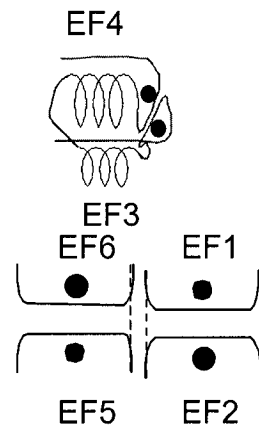
2 violations



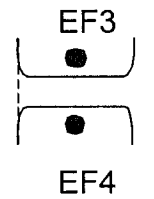
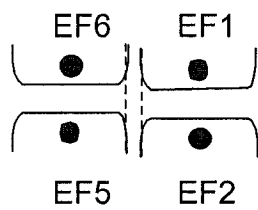
2 violations



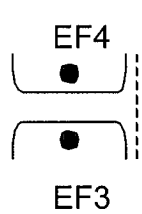
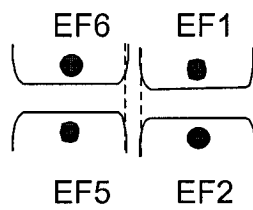
1 violation



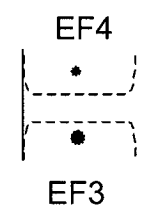
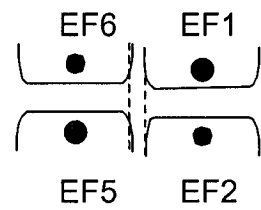
1 violation



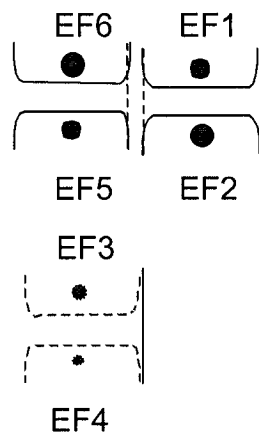
0 violation



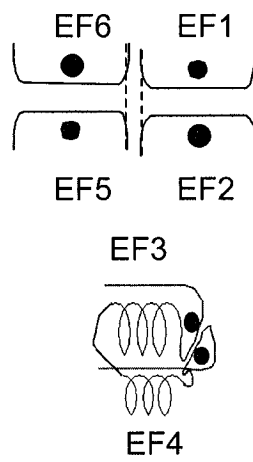
3 violations



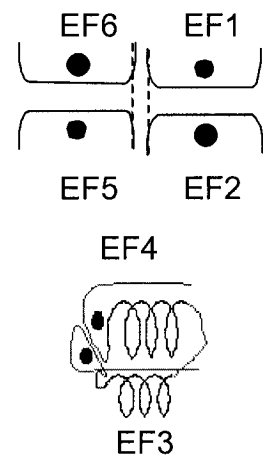
3 violations



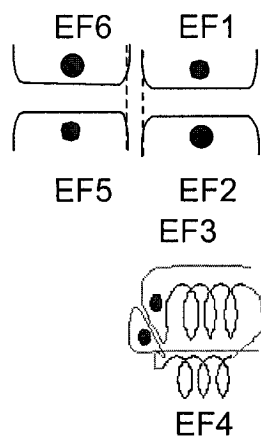
3 violations



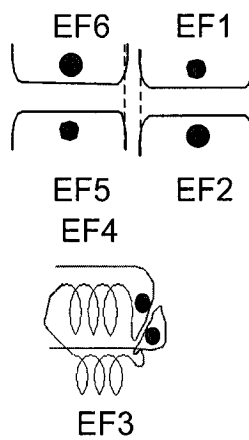
1 violation



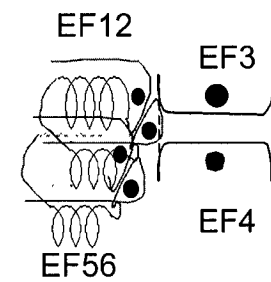
1 violation



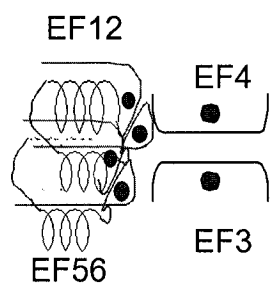
1 violation



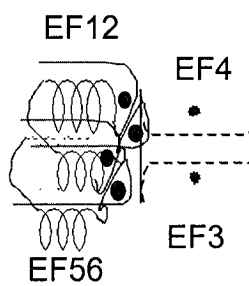
1 violation



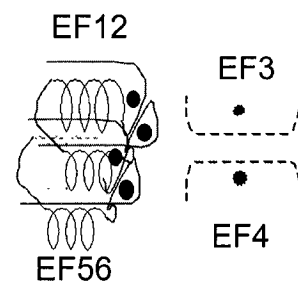
1 violation



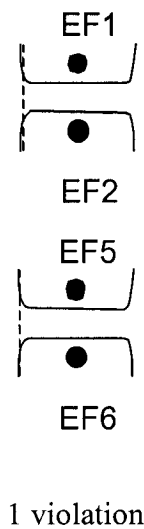
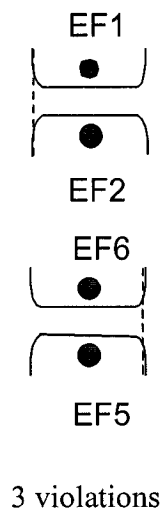
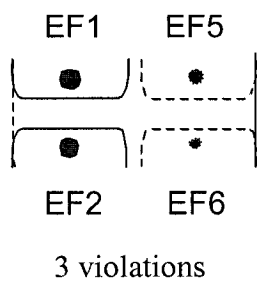
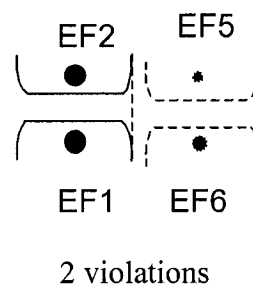
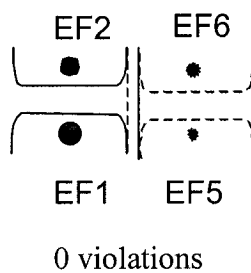
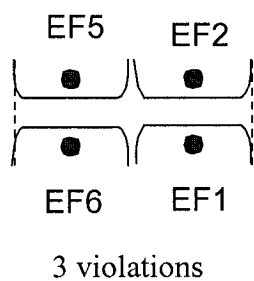
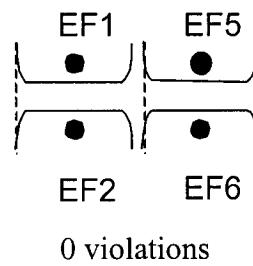
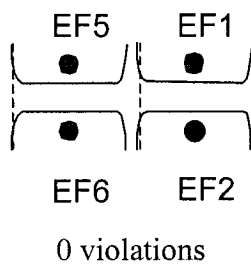
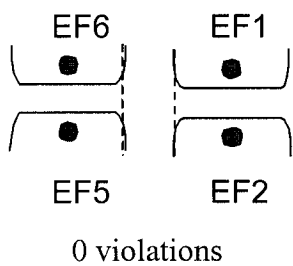
1 violation

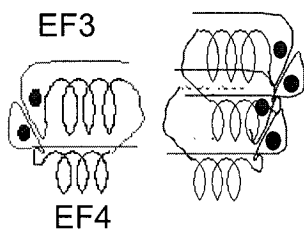


2 violations

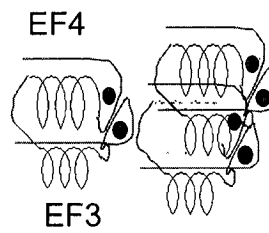


1 violation

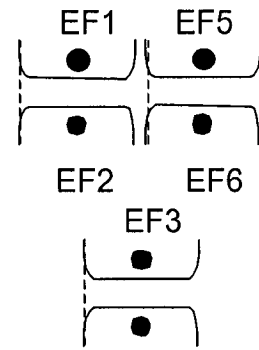




1 violation

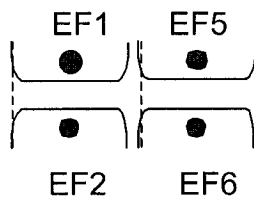


1 violation

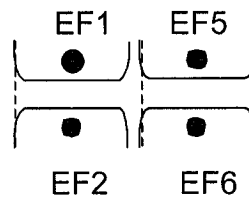


EF4

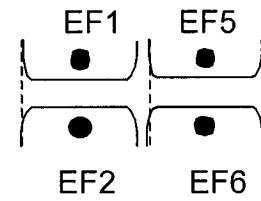
0 violation



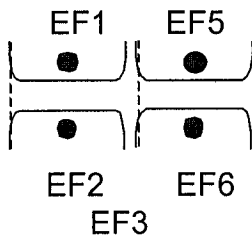
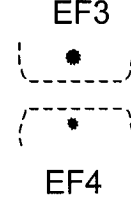
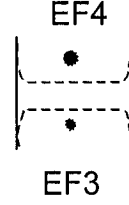
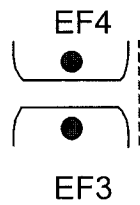
2 violations



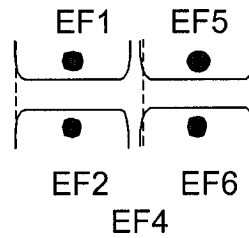
3 violations



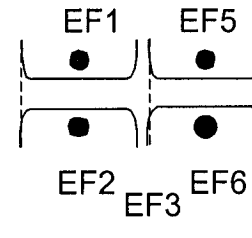
2 violations



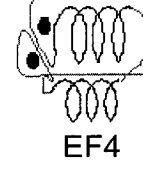
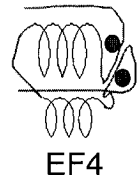
1 violation

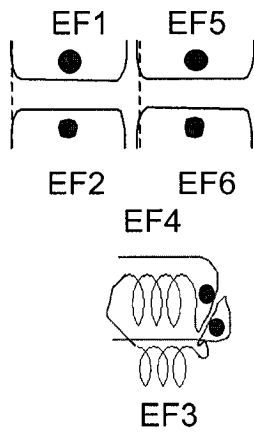


3 violations

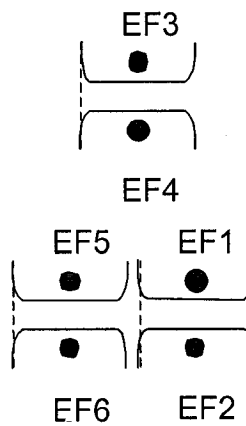


2 violations

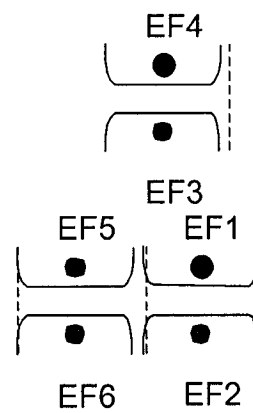




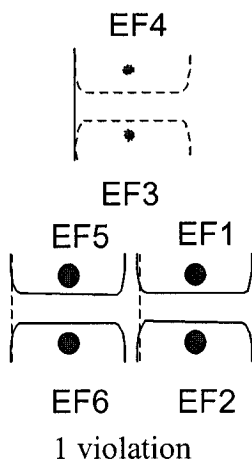
3 violations



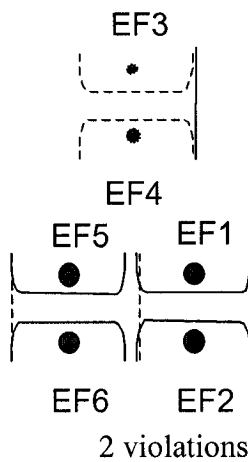
2 violations



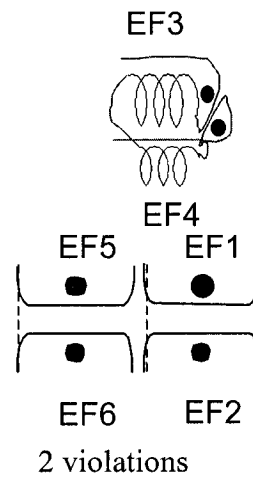
0 violation



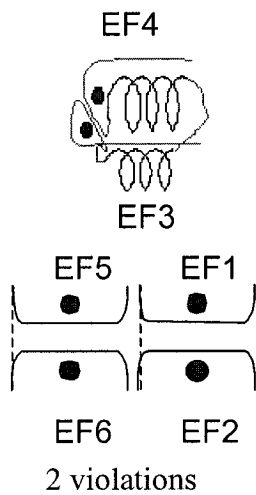
1 violation



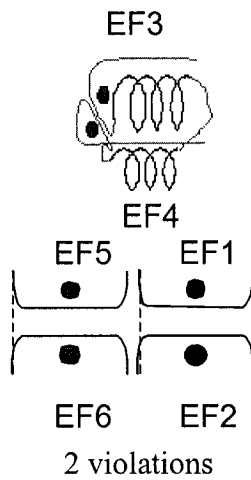
2 violations



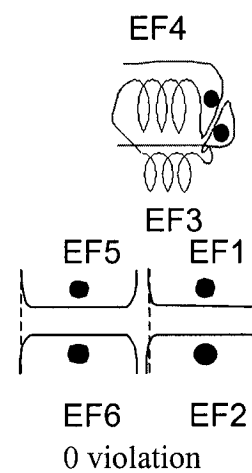
2 violations



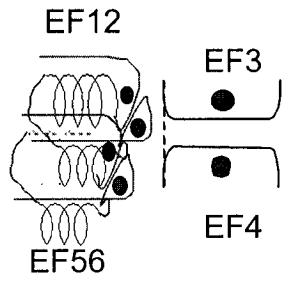
2 violations



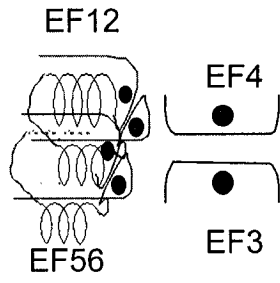
2 violations



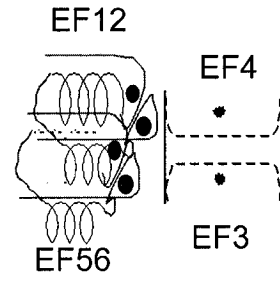
0 violation



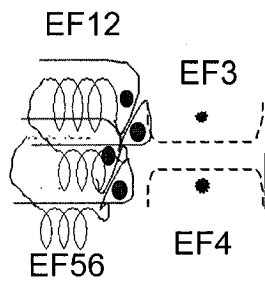
1 violation



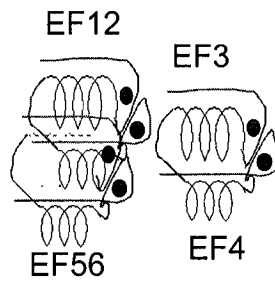
0 violation



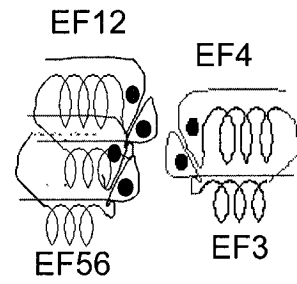
2 violations



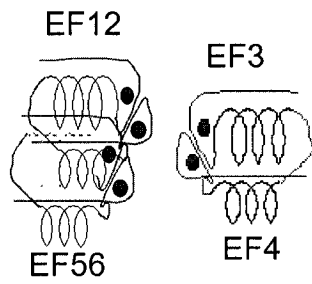
2 violations



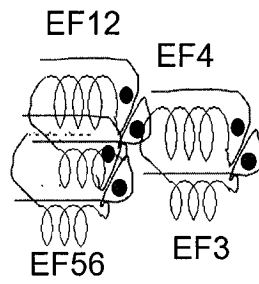
2 violations



2 violations



2 violations



2 violations

References

1. Wasserman, R. H., and Taylor, A. N., *Vitamin D₃-induced calcium binding protein in chick intestinal mucosa*, Science, **1966**, 152, 791-793.
2. Steffensen, B., Wallon, U. M., and Overall, C. M., *Extracellular matrix binding properties of recombinant fibronectin type II-like modules of human 72kDa gelatinase/Type IV collagenase*, Journal of Biological Chemistry, **1995**, 270, 11555-11566.
3. Baimbridge, K. G., Celio, M. R., and Rogers, J. H., *Calcium-binding proteins in the nervous system*, Trends in Neurosciences, **1992**, 15.
4. Gariepy, J., and Hodges, R. S., *Primary sequence analysis and folding behavior of EF hands in relation to the mechanism of action of troponin C and calmodulin*, Federation of European Biochemical Societies Letters, **1983**, 160, 1-6.
5. Meir, T. J., Ho, D. Y., and Sapolsky, R. M., *Increased expression of calbindin D_{28k} via herpes simplex virous amplicon vector decreases calcium ion mobilization and enhances neuronal survival after hypoglycemic challenge*, Journal of Neurochemistry, **1997**, 69, 1039-1047.
6. Parmentier, M., Lawson, D. E. M., and Vassart, G., *Human 27-kDa calbindin complementary DNA sequence*, European Journal of Biochemistry, **1987**, 170, 207-215.
7. Fullmer, C. S., and Wasserman, R. H., *Chicken intestinal 28-kilodalton calbindin-D: Complete amino acid sequence and structural considerations*, Proceeding of the National Academy of Sciences, **1987**, 84, 4772-4776.
8. Vyas, S., Michel, P. P., Copin, M.-C., Biguet, N. F., Thomasset, M., and Agid, Y., *Induction of calbindin-D28K gene and protein expression by physiological stimuli but not in calcium-mediated degeneration in rat PC12 pheochromocytoma cells*, Federation of European Biochemical Societies, **1994**, 351, 53-57.
9. Guo, Q., Cristakos, S., Robinson, N., and Mattson, M. P., *Calbindin D28k blocks the proapoptotic actions of mutant presenilin1: Reduced oxidative stress and preserved mitochondrial function*, Proceeding of the National Academy of Sciences, **1998**, 95, 3227-3232.
10. Yap, K. L., Ames, J. B., Swindells, M. B., and Ikura, M., *Diversity of conformational states and changes within the EF-hand protein superfamily*, Proteins, **1999**, 37, 499-507.
11. Strynadka, N. C. J., and James, M. N. G., *Calcium-binding proteins*, Encyclopedia of Inorganic Chemistry, **1994**.

12. Berggård, T., Silow, M., Thulin, E., and Linse, S., *Ca²⁺ -and H⁺-dependent conformational changes of calbindin D_{28K}*, *Biochemistry*, **2000**, 39, 6864-6873.
13. Berggård, T., Miron, S., Onnerfjord, P., Thulin, E., Åkerfeldt, K. S., Enghild, J. J., Akke, M., and Linse, S., *Calbindin D_{28K} exhibits properties characteristic of a Ca²⁺ sensor*, *Journal of Biological Chemistry*, **2002**, 277, 16662-16672.
14. Venters, R. A., Benson, L. M., Craig, T. A., Paul, K. H., Kordys, D. R., Thompson, R., Naylor, S., Kumar, R., and Cavanagh, J., *The effects of Ca²⁺ binding on the conformation of calbindin D_{28K}: A nuclear magnetic resonance and microelectrospray mass spectrometry study*, *Analytical Biochemistry*, **2003**, 317, 59-66.
15. Berggård, T., Szczepankiewicz, O., Thulin, E., and Sara Linse, *myo-Inositol monophosphatase is an activated target of calbindin D_{28k}*, *Journal of Biological Chemistry*, **2002**, 277, 41954-41959.
16. Tao, L., and English, A. M., *Mechanisms of the S-nitrosation of recombinant human brain calbindin D_{28K}*, *Biochemistry*, **2003**, 42, 3326-3334.
17. Helgstrand, M., Vanbelle, C., Thulin, E., Linse, S., and Akke, M., *Sequential ¹H, ¹⁵N and ¹³C NMR assignment of human calbindin D28k*, *Journal of Biomolecular NMR*, **2004**, 28, 305-306.
18. Kretsinger, R., *Calcium-binding proteins*, *Annual Review of Biochemistry*, **1976**, 45.
19. Berggård, T., Thulin, E., Åkerfeldt, K., and Linse, S., *Fragment complementation of calbindin D_{28k}*, *Protein Science*, **2000**, 9, 2094-2108.
20. Kawasaki, H., Nakayama, S., and Kretsinger, R. H., *Classification and evolution of EF-hand proteins*, *BioMetals*, **1998**, 11, 277-295.
21. Berggård, T., Julenius, K., Ogard, A., Drakenberg, T., and Linse, S., *Fragment complementation studies of the stabilization by hydrophobic core residues*, *Biochemistry*, **2001**, 40, 1257-1264.
22. Linse, S., Thulin, E., Gifford, L. K., Radzewsky, D., Hagan, J., Wilk, R. R., and Åkerfeldt, K. S., *Domain organization of calbindin D_{28k} as determined from the association of six synthetic EF-hand fragments*, *Protein Science*, **1997**, 6, 2385-2396.
23. Veenstra, T. D., Johnson, K. L., Tomlinson, A. J., Naylor, S., and Kumar, R., *Determination of calcium-binding sites in rat brain calbindin D28k by electrospray ionization mass spectrometry*, *Biochemistry*, **1997**, 36, 3535-3542.
24. Åkerfeldt, K. S., Coyne, A. N., Wilk, R. R., Thulin, E., and Linse, S., *Ca²⁺ binding stoichiometry of calbindin D_{28k} as assessed by spectroscopic analyses of synthetic peptide fragments*, *Biochemistry*, **1996**, 35, 3662-3669.

25. Simmons, D. A., Dunns, S. D., and Konerman, L., *Conformational dynamics of partially denatured myoglobin studied by time-resolved electrospray mass spectrometry with online hydrogen-deuterium exchange*, *Biochemistry*, **2003**, 42, 5896-5905.
26. Konerman, L., and Douglas, D. J., *Unfolding of Proteins Monitored by Electrospray Ionization Mass Spectrometry: A Comparison of Positive and Negative Ion Modes*, *Journal of the American Society for Mass Spectrometry*, **1998**, 9, 1248-1254.
27. Furukawa, Y., Matsuda, F., Ishimori, K., and Morishima, I., *Investigation of the electron-transfer mechanism by cross-linking between Zn-substituted myoglobin and cytochrom b₅*, *Journal of the American Chemical Society*, **2002**, 124, 4008-4019.
28. Yang, T., Horejsh, D. R., Mahan, K. J., Zaluzec, E. J., Watson, T. J., and Gage, D. A., *Mapping cross-linking sites in modified proteins with mass spectrometry: An application to cross0linked hemoglobins*, *Analytical Biochemistry*, **1996**, 242, 55-63.
29. Rappsilber, J., Siniosoglou, S., Hurt, E. C., and Mann, M., *A generic strategy to analyze the spatial organization of multi-protein complexes by cross-linking and mass spectrometry*, *Analytical Chemistry*, **2000**, 72, 267-275.
30. Pappa, H. S., and Poulos, T. L., *Site-specific cross-linking as a method for studying intramolecular electron transfer*, *Biochemistry*, **1995**, 34, 6573-6580.
31. Young, M. M., Tang, N., Hempel, J. C., Oshiro, C. M., Taylor, E. W., Kuntz, I. D., Gibson, B. W., and Dollinger, G., *High throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry*, *Proceedings of the National Academy of Sciences*, **2000**, 97, 5802-5806.
32. Pearson, K. M., Pannell, L. K., and Fales, H. M., *Intramolecular cross-linking experiments on cytochrome C and ribonuclease A using an isotope multiplet method*, *Rapid Communications in Mass Spectrometry*, **2002**, 16, 149-159.
33. Kruppa, G. H., Schoeniger, J., and Young, M. M., *A top down approach to protein structural studies using chemical cross-linking and Fourier transform mass spectrometry*, *Rapid Communications in Mass Spectrometry*, **2003**, 17, 155-162.
34. Gross, M. D., Kumar, R., and Hunziker, W., *Expression in Escherichia coli of full-length and mutant rat brain calbindin D28k*, *Journal of Biological Chemistry*, **1988**, 263, 14426-14432.
35. Thulin, E., and Linse, S., *Expression and purification of human calbindin D28K*, *Protein Expression and Purification*, **1999**, 15, 265-270.
36. Tao, L., Murphy, M. E., and English, A. M., *S-nitrosation of Ca(2+)-loaded and Ca(2+)-free recombinant calbindin D(28K) from human brain*, *Biochemistry*, **2002**, 41, 6185-6192.

37. Amersham Pharmacia Biotech. Catalogue, **1998**.
38. Sambrook, J., Fritsch, E. F., and Maniatis, T., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, **1989**, 2nd Ed.
39. QIAquick Spin Handbook, **1999**.
40. Sambrook, J., and Russell, D., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, **2000**, 3rd Ed.
41. Tao, L., and English, A. M., *Protein S-glutathiolation triggered by decomposed S-nitrosoglutathione*, *Biochemistry*, **2004**, 43, 4028-38.
42. Novagen, Factor Xa kits, **2001**.
43. Tao, L., Ph.D. Thesis, Concordia University, Montreal, **2004**.
44. Chu, K., Vojtechovský, J., McMahon, B. H., Sweet, R. M., Berendzen, J., and Schlichting, I., *Structure of a ligand-binding intermediate in wild-type carbonmonoxy myoglobin*, *Nature*, **2000**, 403, 921-923.
45. User guide for reversed-phase ZipTip Millipore, **2002**.
46. Cohen, S. L., and Chait, B. T., *Influence of matrix solution conditions on the MALDI-MS analysis of peptides and proteins*, *Analytical Chemistry*, **1996**, 68, 31-37.
47. Klaus, W., Grzesiek, S., Labhardt, A. M., Buchwald, P., Hunziker, W., Gross, M., and Kallick, D., *NMR investigation and secondary structure of domains I and II of rat brain calbindin D_{28k} (1-93)*, *European Journal of Biochemistry*, **1999**, 262, 933-938.
48. Pidcock, E., and Moore, G. R., *Structural characteristics of protein binding sites for calcium and lanthanide ions*, *Journal of Biological Inorganic Chemistry*, **2001**, 6, 479-89.
49. Strynadka, N. C., and James, M. N., *Crystal structures of the helix-loop-helix calcium-binding proteins*, *Annu Rev Biochem*, **1989**, 58, 951-98.
50. Nakayama, S., and Kretsinger, R. H., *Evolution of the EF-hand family of proteins*, *Annual Review of Biophysics Biomolecular Structure*, **1994**, 23, 473-507.
51. Svensson, L. A., Thulin, E., and Forsen, S., *Proline cis-trans isomers in calbindin D9k observed by X-ray crystallography*, *Journal of Molecular Biology*, **1992**, 223, 601-6.

52. Skelton, N. J., Kordel, J., and Chazin, W. J., *Determination of the solution structure of Apo calbindin D9k by NMR spectroscopy*, Journal of Molecular Biology, **1995**, 249, 441-62.
53. Case, D. A., Pearlman, D. A., Caldwell, J. W., III, T. E. C., Ross, W. S., Simmerling, C. L., Darden, T. A., Merz, K. M., Stanton, R. V., Cheng, A. L., Vincent, J. J., Crowley, M., Tsio, V., Radmer, R. J., Duan, Y., Pitera, J., Massova, I., Seibel, G. L., Singh, U. C., Weiner, P. K., and Kollman, P. A., *AMBER 6*, University of California, San Francisco, **1999**.
54. Jorgensen, W. I., *Transferable Intermolecular Potential Functions for Water, Alcohols, and Ethers. Application to Liquid Water*, Journal of the American Chemical Society, **1981**, 103, 335-340.
55. Bradbrook, G. M., Gleichmann, T., Harrop, S. J., Habash, J., Raftery, J., Kalb, J., Yariv, J., Hillier, I. H., and Helliwell, J. R., *X-ray and molecular dynamics studies of concanavalin-A glucoside and mannoside complexes - Relating structure to thermodynamics of binding*, Journal of the Chemical Society-Faraday Transactions, **1998**, 94, 1603.
56. Creighton, T. E., *Proteins, Structures and molecular properties*, 2nd ed., (1993), W.H. Freeman and Company, New York.
57. Kordel, J., Pearlman, D. A., and Chazin, W. J., *Protein solution structure calculations in solution: solvated molecular dynamics refinement of calbindin D_{9k}*, Journal of Biomolecular NMR, **1997**, 10, 231.