Modeling chaperone-substrate interactions of alpha crystallin from the ocular lens

Lisa Marie S. Ramirez
University at Albany, State University of New York
MODELING CHAPERONE-SUBSTRATE INTERACTIONS OF ALPHA CRystALLIN FROM THE OCULAR LENS

by

Lisa Marie S. Ramirez

A Dissertation
Submitted to the University at Albany, State University of New York
in Partial Fulfillment of
the Requirements for the Degree of
Doctor of Philosophy

College of Arts & Sciences
Department of Chemistry

2019
Preface

Parts of this dissertation have been published as peer-reviewed journal articles. The study on the recombinant expression, purification, and structural characterization of the model substrate, melittin (discussed in Chapters 3-4a) was published with the following citation: Biochemistry 2018, Vol. 57, Ramirez et al., “Nuclear Magnetic Resonance-Based Structural Characterization and Backbone Dynamics of Recombinant Bee Venom Melittin”, pp. 2775-2785. In addition, the three-dimensional atomic-level structure of melittin (discussed in Chapter 4b) was published with the following citation: The Journal of Physical Chemistry B, Vol. 123, Issue 2, Ramirez et al., “Helical Structure of Recombinant Melittin”, pp. 356-368. These journal articles are included as full reprints in the appendix section of the dissertation (Appendix A1 – Biochemistry article, and Appendix B1 – The Journal of Physical Chemistry B article). Permissions (Appendices A2-B2) were granted to me, the dissertation author, by the American Chemical Society, to reproduce these journal articles in their entirety for inclusion in printed and electronic copies of this dissertation. I served as the first author for these journal articles, and the lead researcher for these projects. These articles are included in this dissertation (in revised wording for the text of the manuscript, and published as full reprints in Appendices A1 and B1), because they reflect the line of research that constitute the dissertation.

The research described in Chapters 5 and 7 of this dissertation is being included in a manuscript to be submitted for publication.
Permissions for Use of Figures, Tables, and Schemes

The following figures, tables, and schemes were reproduced from previous publications with permission.

**Figure 1** was adapted from *Dev. Biol.* Vol. 338, Issue 2, Sugiyama *et al.*, “Secreted frizzled-related protein disrupts PCP in eye lens fiber cells that have polarised primary cilia”, pp. 193-201, Copyright (2010), with permission, Elsevier. The license is included in Appendix C.

**Figure 2** was adapted from *Biochim. Biophys. Acta. Gen. Subj.* Vol. 1860, Haslbeck *et al.*, “Structure and function of alpha-crystallins: Traversing from in vitro to in vivo”, pp. 149-166, Copyright (2016), with permission from Elsevier. The license is included in Appendix D.

**Figure 3** was adapted from *J. Phys. Chem. B*, Vol. 119, Issue 33, Liao *et al.*, “Melittin aggregation in aqueous solutions: Insights from molecular dynamics simulations”, pp. 10390-10398, Copyright (2015) with permission from the American Chemical Society. The license is included in Appendix E.

**Scheme 2, Figure 4, Figure 5, Figure 6A, Figure 7, Figure 8, and Figure 9** were adapted from *Biochemistry* 2018, Vol. 57, Ramirez *et al.*, “Nuclear Magnetic Resonance-Based Structural Characterization and Backbone Dynamics of Recombinant Bee Venom Melittin”, pp. 2775-2785, Copyright (2018) with permission from the American Chemical Society. The license is included in Appendix A2.

**Figure 6B** was adapted from *J. Am. Chem. Soc.* Vol. 138, Issue 35, Ishida *et al.*, “Overexpression of antimicrobial, anticancer, and transmembrane peptides in *Escherichia coli* through a calmodulin peptide fusion system”, pp. 11318-11326, Copyright (2016) with permission from the American Chemical Society. The license is included in Appendix F.
Table 1, Figure 10, Figure 11, and Figure 12 were adapted from *J. Phys. Chem. B*. Vol. 123, Issue 2, Ramirez *et al.*, “Helical structure of recombinant melittin”, pp. 356-368, Copyright (2019) with permission from the American Chemical Society. The license is included in Appendix B2.
Acknowledgements

I am deeply grateful for the guidance and support of the people who have accompanied me in my pursuit of a doctorate degree in Chemistry.

To Dr. Jayanti Pande, my research advisor, thank you for believing in me even when I doubted myself. Thank you for giving me room to grow, to develop critical thinking and level-headedness, and to hone skills that would help me become an independent researcher.

To Dr. Alexander Shekhtman, thank you for very generously teaching me how to use different research tools, especially NMR spectroscopy. Thank you for your helpful insight into improving my research projects, and thank you for also being supportive of my efforts to pursue a career in the field of structural biology.

To Dr. Ajay Pande, thank you for mentoring me on the proper use and care of various lab equipment, and for consistently providing protein/peptide samples for my projects. Thank you for your guidance in my preparation of manuscripts and presentations.

To my proposal and dissertation committee members, Dr. Li Niu, Dr. John Welch, Dr. Jia Sheng, and Dr. Chunyu Wang, thank you for your critical assessment of my research projects. I am very grateful for your insight and your guidance.

To the current and former members of the Pande lab, especially Wenjuan and Alok: thank you for keeping the lab enjoyable. To the current and former members of the Shekhtman lab, especially Sergey, Lenny, Chris, Steve, Jianchao and Nick: thank you for welcoming me into your workspace. I very much appreciate your help and your friendship. To Kim DeWeerd and Vladimir Ermolenkov, thank you for training me on the use of Molecular Core and Chemistry Core facilities.
To my friends, especially Danjo, Leslie, Nidhi, Malvin, Bea, Missy, Ginny, Carms, Zarah,- I am very grateful for your guidance. Thank you for keeping me grounded and for supporting my aspirations. To my family, thank you for being my inspiration. Thank you especially to Tita Letty, Tita Nina, Tito Manny, Leica, Minolta, Konika, Alex, and Yvette for making me feel at home in NY.
Abstract

The α-crystallins (αA- and αB-crystallin isoforms) from the ocular lens are small heat shock proteins and molecular chaperones, widely believed to play a significant role in protecting the lens from cataract. Mutations and modifications on human αA- and αB-crystallin (HAA and HAB, respectively) are linked to a variety of diseases, including cataract formation, neuropathological protein folding disorders, and many others. It is believed that HAA and HAB prevent the aggregation of other lens proteins. However, the structural and thermodynamic details of the chaperone-substrate (aka “client”) interaction are sparse. Therefore, the main objective of this dissertation is to reveal structural and thermodynamic aspects of such interaction, using a model substrate, the 26-residue peptide melittin.

HAA is primarily lenticular while HAB is more widespread in the human body. To date, more structural models are available for αB-crystallin, which has been more extensively investigated. In this dissertation, the focus is on chaperone-substrate interactions involving αA-crystallin which is – for the most part – lens specific, and therefore plays a dominant role in lens health and disease.

Melittin, a peptide from honey bee venom, has been previously used by other researchers to identify a putative substrate-binding region on α-crystallin. Structural studies on this model substrate have been carried out in the last few decades in an effort to understand melittin’s antimicrobial/antibacterial and hemolytic properties, however atomic level details of melittin’s structure while bound to its biological targets (including membranes, polysaccharides, and proteins) have been limited.

In order to facilitate high resolution nuclear magnetic resonance (NMR) structural studies, we therefore developed a protocol for the recombinant expression of the model substrate with stable isotope labels. Using NMR, we characterized the coil to helix transition of the isotope-labeled
melittin in aqueous solution and 2,2,2-trifluoroethanol (TFE)/water mixtures, and studied the effect of proline cis-trans isomerization on its secondary structure and backbone dynamics. We found that in the presence of TFE, a helix-inducing agent, the cis conformer of melittin forms one N-terminal helix, while the trans conformer forms distinct N- and C-terminal helices separated by an unstructured ‘hinge’. We also determined the atomic level structure of the trans conformer in TFE-containing medium using a workflow for optimizing distance restraint collection for small peptides. With our structural studies on melittin, we were able to provide insight into the conformational transitions of melittin relevant to its interactions with other biological partners.

We explored the interaction of melittin with the mini-αA-crystallin chaperone called “MAC” (a 19-residue peptide corresponding to a putative substrate binding domain on αA-crystallin), and αA-G98R, a cataract-associated mutant of αA-crystallin. The MAC peptide was previously reported to have chaperone properties comparable to that of the parent α-crystallin protein. Studying the interaction between MAC and the model substrate melittin is essential to elucidating structural aspects of client binding to α-crystallin. Titrations of melittin with MAC monitored by NMR spectroscopy and fluorescence emission spectroscopy indicated that melittin binds to MAC at low affinity (with a dissociation constant in the range ~10-100 μM), with positive cooperativity. The model of the melittin-MAC complex suggested that hydrophobic contacts and electrostatic interactions help in stabilizing the complex.

With regard to the mutant, previous reports indicated that αAG98R preferentially forms monomers under certain experimental conditions, while still retaining considerable chaperone activity. We reasoned that this property may facilitate the acquisition of NMR data, considering that larger, multimeric assemblies of αA-crystallin are beyond the detection limit of solution NMR spectroscopy. We therefore characterized the mutant by solution NMR spectroscopy and screened
conditions with the aim of attaining high resolution data for structure calculations. Titrations of melittin with αAG98R suggested that the N- and C-terminal flexible regions on αAG98R interacted with melittin in an intermediate exchange process.

Finally, we examined the chaperone-substrate interaction between parent αA-crystallin and melittin. Using NMR spectroscopy, we determined the residues on melittin interacting with αA-crystallin. From ITC, we found that the melittin-αA-crystallin binding event is endothermic and entropy-driven. NMR and ITC experiments indicated that melittin binds to αA-crystallin at low affinity ($K_d \sim 20 \mu M$), and that hydrophobic residues on the C-terminal region of melittin are implicated in the binding event. From saturation transfer difference NMR experiments, we found that residues Ile17 and Ile20 on melittin mainly mediated the interaction between melittin and αA-crystallin. We provided a model of the chaperone-substrate complex given residue-specific data derived from our NMR studies and previously published cross-linking data. The resulting model shows hydrophobic contacts and salt bridges that form the binding interface in the melittin-αA-crystallin complex.
Table of Contents

Preface .......................................................................................................................... ii
Permissions for Use of Figures, Tables, and Schemes ............................................. iii
Acknowledgements ....................................................................................................... v
Abstract ........................................................................................................................ vii
Table of Contents ......................................................................................................... x
List of Abbreviations and Symbols Used .................................................................... xii

1.1. Function of α-Crystallin in the Human Lens ......................................................... 1
1.2. Structural Characterization of the α-Crystallin Molecular Chaperone: Progress and Challenges ................................................................................................................. 4
1.3. Interactions between Client/Substrate Proteins and the α-Crystallin Molecular Chaperone ......................................................................................................................... 7
1.4. Chaperone Activity of the Mini-αA-Crystallin Molecular Chaperone ..................... 8
1.5. Chaperone Activity of the Cataract-Associated Mutant of αA-Crystallin, αA-G98R .... 8
1.6. Melittin Is a Model Substrate for the α-Crystallin Molecular Chaperone ............... 9
1.7. Melittin Is a Peptide Model for Studies on Protein Folding and Protein-Membrane Interactions .................................................................................................................. 11

Chapter 2 - Materials and Methods ........................................................................... 14
2.1 Preparation of Protein and Peptide Samples ......................................................... 14
2.2 Overview of NMR Experiments and NMR Data Analysis .................................... 15
2.3 Liquid Chromatography Mass Spectrometry (LC-MS) ........................................ 15
2.4 Identification of Putative Interaction Sites on Melittin by NMR Line Broadening Analysis ............................................................................................................................ 16
2.5 Expression and Purification of the G98R Mutant of αA-Crystallin ....................... 16
2.6 Isothermal Titration Calorimetry for Characterizing the Binding Thermodynamics of the Chaperone-Substrate Interaction .............................................................................. 17

Chapter 3 – Cloning, Expression, and Purification of Melittin .................................. 19
3.1 Introduction ............................................................................................................. 19
3.2 Materials and Methods ......................................................................................... 20
3.3 Results ................................................................................................................... 22
3.4 Discussion ............................................................................................................. 28
Chapter 4A – Structural Characterization and Backbone Dynamics of Melittin ............................................... 30
  4A.1 Introduction ........................................................................................................................................... 30
  4A.2 Materials and Methods ......................................................................................................................... 32
  4A.3 Results .................................................................................................................................................. 33
  4A.4 Discussion ............................................................................................................................................ 37
Chapter 4B – High Resolution Structure of Helical Recombinant Melittin ......................................................... 39
  4B.1 Introduction ........................................................................................................................................... 39
  4B.3 Results .................................................................................................................................................. 42
  4B.4 Discussion ............................................................................................................................................ 47
Chapter 5 - Examining Peptide-Peptide Interactions between MAC (a 19-Residue Peptide of α-Crystallin) and Melittin ................................................................................................................................. 49
  5.1 Introduction ........................................................................................................................................... 49
  5.2 Materials and Methods ............................................................................................................................ 50
  5.3 Results ................................................................................................................................................... 53
  5.4 Discussion ............................................................................................................................................. 62
Chapter 6 - Examining the Interaction between ............................................................................................... 64
  the Cataract-Associated Mutant of αA–Crystallin, αA– G98R, and Melittin ..................................................... 64
  6.1. Introduction ........................................................................................................................................... 64
  6.2. Materials and Methods .......................................................................................................................... 64
  6.3. Results .................................................................................................................................................. 66
  6.4 Discussion ............................................................................................................................................. 73
Chapter 7 - Examining the Chaperone-Substrate Interaction between αA–Crystallin and Melittin ....................... 74
  7.1. Introduction ........................................................................................................................................... 74
  7.2. Materials and Methods .......................................................................................................................... 75
  7.3 Results ................................................................................................................................................... 81
  7.4 Discussion ............................................................................................................................................. 94
Summary and Conclusions ............................................................................................................................... 100
List of Publications and Presentations ............................................................................................................... 103
Bibliography .................................................................................................................................................... 105
Appendix: Publications and Licenses ............................................................................................................... 117
List of Abbreviations and Symbols Used

ACD    \( \alpha \)-crystallin domain
AEBSF  4-benzenesulfonyl fluoride hydrochloride
\( \alpha \)-G98R the cataract-associated Gly98 to Arg mutant of HAA
Bis-ANS 4,4’-dianilino-1,1’-binaphthyl-5,5’-disulfonic acid
CARA  computer aided resonance assignment
EDTA  ethylenediaminetetraacetic acid
ESI-MS electrospray ionization mass spectrometry
HAB  human \( \alpha \B-crystallin
HAA  human \( \alpha \A-crystallin
HADDOCK high ambiguity driven biomolecular docking
HSQC heteronuclear single quantum coherence
ITC  isothermal titration calorimetry
IPTG isopropyl 1-thio-D-galactopyranoside
LB  Luria-Bertani or lysogeny broth
LC-MS liquid chromatography-mass spectrometry
MAC mini- \( \alpha \A-crystallin chaperone
\( \mu \)M  micromolar
NMR  nuclear magnetic resonance
NTD  N-terminal domain
Ni-NTA nickel nitrilotriacetic acid
NOE  nuclear Overhauser effect
NOESY nuclear Overhauser effect spectroscopy
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDB</td>
<td>Protein Data Bank</td>
</tr>
<tr>
<td>PNK</td>
<td>polynucleotide kinase</td>
</tr>
<tr>
<td>RP-HPLC</td>
<td>reversed phase-high performance liquid chromatography</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>sodium dodecyl sulfate-polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>sHSPs</td>
<td>small heat-shock proteins</td>
</tr>
<tr>
<td>SOFAST-HMQC</td>
<td>band-selective optimized-flip-angle short-transient heteronuclear multiple quantum coherence</td>
</tr>
<tr>
<td>STD-NMR</td>
<td>saturation transfer difference-nuclear magnetic resonance</td>
</tr>
<tr>
<td>TFE</td>
<td>2,2,2-trifluoroethanol</td>
</tr>
<tr>
<td>TOCSY</td>
<td>total correlation spectroscopy</td>
</tr>
<tr>
<td>TrpLE</td>
<td>tryptophan leader peptide</td>
</tr>
</tbody>
</table>
Chapter 1 – Introduction

The main goal of this thesis is to examine the nature of the interaction between the small heat-shock protein and molecular chaperone, αA-crystallin from the ocular lens, with a biologically active peptide, melittin, which serves as a client mimic for the chaperone. This goal was approached by applying the following strategies;

1. Recombinant expression, purification, and structural characterization of the isotope-enriched “model” client, melittin (Chapter 3-4A)

2. Determination of the high-resolution structure of the model client, melittin (Chapter 4B)

3. Investigation of the interactions of the model client with a model chaperone peptide derived from the parent protein (Chapter 5)

4. Investigation of the interactions of the model client with a cataract-associated mutant of the parent protein that retains chaperone activity (Chapter 6)

5. Investigation of the interactions of the model client with the parent protein (Chapter 7)

1.1. Function of α-Crystallin in the Human Lens

The vertebrate lens can be described as a stratified assembly of crescent-shaped fiber cells (comprising more than 95% of the tissue volume) (Figure 1), arranged radially around a common axis.\textsuperscript{10,11} The fiber cells are arranged such that oldest fiber cells (called nuclear fibers, NF) are in the central ‘nucleus’ of the lens, and the age of the fiber cells decreases with each layer moving outward.\textsuperscript{12,13} A monolayer of epithelial cells surrounds the fiber cells.\textsuperscript{11} Growth of the lens tissue involves the mitosis of epithelial cells in the germinative zone (GZ).\textsuperscript{11} Progeny of these cells can differentiate into lens fiber cells, in a process that involves cell elongation. Fiber cells eventually
mature, in a process that involves degradation of nuclei and loss of other organelles. It is believed that the organelle loss in fiber cells contributes to lens clarity by minimizing light scattering, given that organelles in the fiber cell would scatter light.\textsuperscript{11, 13} Fiber cells form throughout life, however the rate of fiber cell formation is reduced in adulthood.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{mouse-eye-diagram.png}
\caption{Diagram of the mouse eye. A) Sagittal-section. B-C) Horizontal sections. Abbreviations: CE, central epithelium; GZ, germinative zone; TZ, transitional zone; OCF, outer cortical fibers; ICF, inner cortical fibers; NF, nuclear fibers. Adapted from \textit{Dev. Biol.} Vol. 338, Issue 2, Sugiyama et al.,\textsuperscript{14} “Secreted frizzled-related protein disrupts PCP in eye lens fiber cells that have polarised primary cilia”, pp. 193-201, Copyright (2010), with permission. Elsevier.}
\end{figure}

The crystallin proteins (\(\alpha\)-, \(\beta\)-, and \(\gamma\)-crystallins) constitute about 90\% of all lens proteins.\textsuperscript{2, 6, 11, 15} It was suggested that short-range liquid-like order assumed by the crystallin proteins in the fiber cell also accounts for lens transparency by minimizing spatial fluctuations in the refractive index.\textsuperscript{16}

To avoid clouding of the lens, lens proteins must retain solubility even at high protein concentrations (up to 450 mg/mL protein in the fiber cell).\textsuperscript{17} It is believed that among the crystallin proteins, the \(\alpha\)-crystallins assume the role of maintaining the solubility of other lens proteins and protecting against unfolding and aggregation.\textsuperscript{1, 2} This is partly because the \(\alpha\)-crystallins appear to be small heat-shock proteins (sHSPs),\textsuperscript{18} which, as shown later, also function as molecular chaperones. Studies \textit{in vitro} have established the molecular chaperone functions of the \(\alpha\)-crystallin
hetero-oligomeric mixture, as well as homo-oligomeric assemblies formed by αA-crystallin or αB-crystallin.\textsuperscript{3, 5}

The α-crystallin family is composed of homologous proteins, αA-crystallin and αB-crystallin. αA-crystallin is composed of 173 residues, while αB is composed of 175 residues. The two proteins share about 57\% sequence homology.\textsuperscript{1} In humans, the αA gene is found on chromosome 21 while that of αB is on chromosome 11.\textsuperscript{19, 20} αA-crystallin is predominantly found in lenticular tissue, with minute amounts in the spleen, heart, brain and liver.\textsuperscript{21, 22} In contrast, αB-crystallin is more widely distributed in the body, and, in addition to the lens, is expressed in the lung, spleen, brain, kidney, heart, and skeletal muscle.\textsuperscript{23-25} When isolated from adult mammalian lens, α-crystallin is composed of approximately a 3:1 ratio for αA to αB.\textsuperscript{3} The α-crystallin mixture is described as a polydisperse high molecular weight assembly with a broad size distribution (masses within the range \~300 - 1200 kilodaltons).\textsuperscript{3, 15} It must be noted that the size of the assembly varies with the source organism, solution conditions, and the method by which the mass was determined.\textsuperscript{1, 15, 26, 27} Rapid subunit exchange within the dynamic α-crystallin assembly may also be important in the chaperone function of α-crystallin.\textsuperscript{28, 29}

As stated above, the α-crystallins are members of the small heat shock protein (sHSP) family. Small heat shock proteins are generally considered as a first line of defense against protein aggregation in cells.\textsuperscript{5} sHSPs are predominantly, ATP-independent molecular chaperones that function in stabilizing protein conformations, mediating and stabilizing correct oligomerization of proteins, protein translocation, and protection from stresses such as heat denaturation and osmotic shock.\textsuperscript{3, 5} The \textit{in vitro} study conducted by Horwitz\textsuperscript{3} established the molecular chaperone function of α-crystallin by demonstrating its ability to suppress thermal aggregation of the natural targets in the lens: namely, the β- and γ-crystallins.
Mutations on αA-crystallin and αB-crystallin are implicated in cataract formation in humans.\textsuperscript{30-35} Interestingly however, αB-crystallin is also associated with a variety of other diseases, notably those involving protein folding defects such as Alzheimer’s disease, Parkinson’s disease, Creutzfeldt-Jakob disease, and Alexander’s disease.\textsuperscript{36-40} Considering the dearth of knowledge on the detailed structure of the intact α-crystallins, it is not surprising that the molecular mechanisms behind such pathology are still poorly defined. Therefore, there is a need for detailed structural and thermodynamic studies on the α-crystallin molecular chaperones that may facilitate a molecular understanding of disease mechanisms in such cases.

### 1.2. Structural Characterization of the α-Crystallin Molecular Chaperone: Progress and Challenges

The structured region in α-crystallins is now known to consist of a 80-100 residue-long segment, named appropriately as the α-crystallin domain (ACD).\textsuperscript{41} This β-sandwich domain is present with minor variations in most sHSPs known so far.\textsuperscript{5,42} The ACD is flanked by N- and C-terminal regions, both of which are highly disordered in α-crystallins.\textsuperscript{5} It has been suggested that the lack of defined structure of the N-terminal domain may explain the ability of α-crystallin and other sHSPs to bind to a variety of substrates.\textsuperscript{5,43,44} It was also proposed that the C-terminal tail may be involved in solubilizing the molecular chaperone.\textsuperscript{44,45} The N-terminal arm is also thought to assume certain functions in oligomer organization of the molecular chaperone.\textsuperscript{5}

High resolution structure determination of the α-crystallin molecular chaperone has proven to be difficult, primarily owing to the dynamic and heterogeneous nature of the α-crystallin assembly.\textsuperscript{44} The tertiary and quaternary structure of human α-crystallin are still not known, and the structural basis of its chaperone activity is not well-understood.\textsuperscript{5,44} Attempts to determine the
X-ray crystal structure of α-crystallin are reportedly hindered by the flexible N- and C-terminal regions of the molecular chaperone. The large size of α-crystallin also precludes structure determination by conventional solution NMR approaches. Solution NMR studies on α-crystallin carried out by the group of Carver have primarily focused on the unstructured C-terminal tails, which can be observed in NMR spectra due to the extreme flexibility of the tail.

Considering the above-mentioned challenges, several research groups have, in addition to or in lieu of full-length α-crystallin, utilized alternative constructs for the structural characterization of α-crystallin. Selected examples are discussed here. Laganowsky et al. determined the crystal structures of truncated α-crystallins containing the ACD (bovine αA59-163 in Figure 2, right panel, and human αB68-162) but it is worth noting that the truncated proteins do not show chaperone activity. Bagneris et al. determined the 2.9-Angstrom crystal structure of the homodimer formed by excised ACD from human αB-crystallin (αB67-157), although the chaperone function of this construct was not characterized by the authors. Solid-state NMR studies on full-length human αB-crystallin conducted by Jehle et al. also provided a dimeric model of residues 69-150 from the ACD (Figure 2, left panel). Taken together, these studies indicate that the ACD dimer of αA and αB each adopt an immunoglobulin-like β-sandwich fold.
Figure 2. Representative structures of the ACD dimers of α-crystallin (X-ray structure, PDB ID: 3L1F, bovine αA\textsubscript{59-163})\textsuperscript{48} and of αB-crystallin (solid state NMR, PDB ID: 2KLR)\textsuperscript{49} human αB\textsubscript{69-150}. The β-sandwich fold is formed by two antiparallel layers of three (β3- β9- β8) and four (β4- β5- β6+7) β-strands connected by a loop. Adapted from Biochim. Biophys. Acta. Gen. Subj. Vol. 1860, Haslbeck et al.,\textsuperscript{2} “Structure and function of alpha-crystallins: Traversing from in vitro to in vivo”, pp. 149-166, Copyright (2016), with permission from Elsevier.

Notably, the three-dimensional (3D) structure of full-length, native human α-crystallin has not been determined,\textsuperscript{2} whereas the 3D structure of bovine α-crystallin has been recently modeled by electron microscopy.\textsuperscript{50} Hybrid approaches using solid-state NMR and solution-state NMR together with electron microscopy have also been proposed to investigate the quaternary structure of α-crystallin.\textsuperscript{51} Several crude models have been proposed for oligomeric human αB-crystallin based on cryo-electron microscopy\textsuperscript{26, 52} and NMR.\textsuperscript{49} The structure of oligomeric human αA-crystallin was recently modeled by Swiss Model,\textsuperscript{53-55} based on its homology with other sHSPs.
1.3. Interactions between Client/Substrate Proteins and the α-Crystallin Molecular Chaperone

As mentioned in Section 1.1, the molecular chaperone function of α-crystallin is reflected in its ability to suppress the aggregation of ‘client’ or ‘substrate’ proteins, especially, the lens proteins β- and γ-crystallins, which are its natural targets. Enzymes such as alcohol dehydrogenase, citrate synthase, and enolase have also been used as model substrates for studying the molecular chaperone activity of α-crystallin. The aggregation of the model substrates in vitro is typically determined in a ‘chaperone assay’, in which the molecular chaperone is incubated with a partly unfolded model substrate, and apparent absorption due to light scattering is measured over time, typically at 360 nm. Using this type of assay, a multitude of in vitro studies on the ability of α-crystallin to ‘chaperone’ model substrates have been carried out, although little is understood about the specific residues and surface patches on α-crystallin that directly participate in interactions with α-crystallin’s natural targets.

Sequences on αB-crystallin thought to serve as ‘interactive domains’ between αB-crystallin and other proteins (including lens proteins and regulatory proteins) were identified by Ghosh et al. using protein pin arrays. Using cross-linking and peptide-mapping methods with melittin and 4,4’-dianilino-1,1’-binaphthyl-5,5’-disulfonic acid (bis-ANS) as model substrates, the group of Sharma et al. identified a 19-residue sequence on αA-crystallin as a putative substrate-binding site. The peptide corresponding to this fragment (70DFVIFLDVKHFSPEDLTVK88) is known as the mini- αA-crystallin chaperone (MAC) (Section 1.4). The same research group identified a functional element (with the sequence 73DRFSVNLVDVKHFSPPEELVK92) on αB-crystallin, and the peptide corresponding to this sequence was also reported to show molecular chaperone activity.
1.4. Chaperone Activity of the Mini-αA-Crystallin Molecular Chaperone

The mini-αA-crystallin chaperone (MAC), introduced in Section 1.3, is a 19-residue fragment (70DFVIFLDVKHFSPEDLTVK88), which is part of the α-crystallin domain (ACD) of the αA-crystallin parent protein.\(^8\) This sequence corresponds to two β-strands on the ACD that are separated by a loop region.\(^7\) Previously, it was determined that this peptide possesses chaperone properties similar to that of αA-crystallin in vitro.\(^7, 8, 60\) Specifically, MAC has been found to exhibit chaperone activity toward human γD-crystallin (HGD), alcohol dehydrogenase, and α-lactalbumin.\(^7, 8, 61\) In other reports, MAC was found to suppress amyloid fibril formation of amyloid-β protein.\(^60, 62\) Thus, it appears that MAC may act as a suitable model that could mimic the molecular chaperone properties of intact α–crystallin. Given these intriguing properties, we hypothesized that studying the interaction between melittin (a model substrate for αA–crystallin) and MAC may provide insights into the interactions between the substrate model and the full-length αA–crystallin protein. In Chapter 5, the interaction between melittin and MAC is examined by fluorescence titrations, NMR spectroscopy, and computational docking.

1.5. Chaperone Activity of the Cataract-Associated Mutant of αA-Crystallin, αA-G98R

The G98R mutation in αA-crystallin was found to be associated with early-onset cataracts.\(^33\) Previous studies have shown that the mutation significantly alters the protein’s structure and chaperone activity.\(^9\) Interestingly, studies on recombinant G98R mutant protein have revealed that the mutant’s molecular chaperone properties may vary in a substrate dependent manner: the mutant was found to have diminished chaperone activity toward α-lactalbumin, but prevented aggregation of citrate synthase and alcohol dehydrogenase.\(^63\) Moreover, the oligomeric state of the mutant seems to affect its chaperone function.\(^9\) The work of Raju et al.\(^9\) showed that the mutant favors the
monomeric form under certain solution conditions in which the wild-type αA-crystallin favors the oligomeric form. Using chaperone assays with the monomer form of the G98R mutant, the researchers found that the monomeric subunit exhibits chaperone function toward citrate synthase, ovotransferrin, βB2-crystallin, and alcohol dehydrogenase comparable to that of wild-type αA-crystallin. In contrast, a different study showed that large oligomers of the mutant lacked chaperone activity.

As mentioned in Section 1.2, structural characterization of the HAA oligomer has proved to be challenging owing to its large size and polydispersity. In view of the fact that, unlike the wild-type HAA, the mutant αA-G98R shows a higher propensity to form monomers in solution, we decided that structural studies on monomeric αA-G98R may provide insight into the structural basis of HAA molecular chaperone activity. The present study thus involves the preparation and characterization of αA-G98R. In addition, the interaction between melittin and αA-G98R is examined using NMR (Chapter 6).

1.6. Melittin Is a Model Substrate for the α-Crystallin Molecular Chaperone

Melittin is a small peptide consisting of 26 amino acid residues. It has been extensively studied, but because of its versatility, melittin continues to be the focus of ongoing investigations in many areas. Melittin is also known to possess antibacterial, antiviral, and anticancer properties. Melittin at low concentrations (~0.035 mM) has also been shown to cause complete lysis of erythrocytes. It is believed that melittin’s toxicity stems from its interaction with membranes of bacteria, cancer cells, and erythrocytes.
Melittin is the primary component of honey bee venom, and it is known that bee stings – particularly those that directly affect the eye – often lead to cataract development.\textsuperscript{71,72,73} Melittin is highly basic, with an isoelectric point of about 12. It possesses some structural features typical of cationic detergents – the N-terminal end is predominantly hydrophobic, while the C-terminal end is highly hydrophilic and basic (Figure 3).\textsuperscript{71,74} Numerous investigations into the characterization of the structure of melittin and its role as a model peptide for different systems have been carried out over the last few decades.\textsuperscript{65,75}

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure3.png}
\end{figure}

The selection of melittin as a model substrate for $\alpha$-crystallin is mainly based upon a previous report by Sharma et al\textsuperscript{7} in which they used melittin binding to map an important substrate-binding site on $\alpha$-crystallin. The authors found that melittin binding reduced the molecular chaperone activity of $\alpha$-crystallin toward partially unfolded alcohol dehydrogenase.\textsuperscript{7}

The interaction of melittin with $\alpha$-crystallin \textit{in vitro} has been studied using electron paramagnetic resonance (EPR) spectroscopy.\textsuperscript{77} Equilibrium binding studies for spin-labeled melittin and $\alpha$-crystallin yielded a dissociation constant of 7.3 $\mu$M and a binding stoichiometry of
1:1 melittin per α-crystallin monomer.\textsuperscript{77} One should note that this binding study used whole bovine α-crystallin composed of a mixture of the αA and αB isoforms. The dissociation constant for each α-crystallin isoform was not determined in the said study. The authors further reported that, in the bound state, melittin may form a helix or a β-strand.

The use of a small peptide as a ‘probe’ for α-crystallin has advantages in the context of acquiring high-quality NMR spectra. In general, it is difficult to obtain well-resolved spectra (by conventional NMR methods) of large protein assemblies,\textsuperscript{78} while NMR spectra for small peptides and proteins can be readily measured. Based on these considerations, we developed a melittin probe labeled with NMR-active isotopes, to facilitate the observation of NMR signals on melittin in a mixture of melittin and αA-crystallin. These studies are reported in Chapter 3.

\subsection*{1.7. Melittin Is a Peptide Model for Studies on Protein Folding and Protein-Membrane Interactions}

As mentioned in Sections 1.3 and 1.6, it is known that melittin binds to α-crystallin, and previous reports have suggested that melittin may undergo a conformational change upon binding.\textsuperscript{7, 77} Structure determination of melittin (focusing on the conformation of melittin in solution, under near-physiological conditions) is an important step in modeling the binding of melittin to α-crystallin. In the present study (Chapter 4A-4B), the conformational transitions of melittin in solution are characterized, and the high-resolution structure of melittin in solution is determined. An added advantage of these structural studies is that they provide insight into the detailed mechanisms of other biological processes, including protein folding, and associations between proteins and lipids, as explained below.
There is general agreement that melittin may be in a random coil or α-helical conformation depending on solution conditions such as pH, ionic strength, and peptide concentration.\textsuperscript{75, 79-81} Lauterwein et al.\textsuperscript{80} reported that monomeric melittin in a random coil configuration was the predominant form of melittin in dilute solutions, at low ionic strength, and at acidic pH. The study of Brown et al.\textsuperscript{81} examined solution conditions that favored the self-association of melittin into its aggregated (possibly tetrameric) form including high melittin concentrations, basic pH, and high ionic strength. Bazzo et al.\textsuperscript{82} reported on the structure of melittin in methanol determined by $^1$H-NMR. The findings show that melittin also assumes a helical conformation in methanol, and is monomeric.\textsuperscript{82} The ability of melittin to change secondary structure as a way of adapting to different solution conditions has prompted researchers to use melittin as a peptide model for gaining insight into the protein folding process.

The membrane-active nature of melittin has been extensively studied in natural and artificial membrane systems - thus melittin is widely considered to be a useful model for protein-lipid interactions in biological membranes.\textsuperscript{75, 83, 84} It has been suggested that melittin forms a helical structure in membrane bilayers.\textsuperscript{75} Several mechanisms have been proposed to explain how melittin traverses (or permeates through) the membrane, including the toroidal pore model.\textsuperscript{71, 74, 85-89} It is worth noting that most structural studies available for melittin have not emphasized a clear, structural distinction between \textit{cis-} and \textit{trans-} conformations of the Leu13-Pro14 bond in melittin, which may have – as yet unknown – implications for a variety of applications. The study of Lauterwein et al.\textsuperscript{80} simply mentioned that a melittin sample in aqueous solution shows two populations, with the low-abundance form possibly arising from the \textit{cis-trans} isomerization of Leu13-Pro14. The study of Bazzo et al.\textsuperscript{82} notes that no \textit{cis-} isomer was observed for melittin in methanolic solution. Lack of structural studies on this isomerization may be attributed to the
relatively low amount of the cis- isomer compared to the trans- isomer, and the difficulty in resolving overlapping NMR signals of melittin samples lacking NMR-active isotopes. In the present study (Chapter 4A), we distinguish between the cis- and trans- forms of the melittin peptide using high-resolution NMR and the isotope-labeled melittin probe, and we discuss the structural implications of this isomerization.
Chapter 2 - Materials and Methods

2.1 Preparation of Protein and Peptide Samples

The concentrations of melittin, HAA, and αA-G98R samples were determined from absorbance at 280 nm. The extinction coefficient used for melittin was 5,500 M⁻¹cm⁻¹. HAA and αA-G98R concentrations were calculated using an extinction coefficient of 14,564 M⁻¹cm⁻¹ (extinction coefficients were calculated using the EXPASY ProtParam tool). The concentration of mini αA-crystallin chaperone (MAC) was determined as described previously from the absorption of the 3 Phe residues at 258 nm (195 cm⁻¹M⁻¹ per Phe residue). The absorption spectra of protein and peptide samples were acquired using a Cary 50 UV-Vis spectrophotometer.

Recombinant, isotope-labeled melittin was expressed and purified using a published procedure. The development of this procedure is discussed in Chapter 3. Samples of recombinant melittin for NMR experiments were typically prepared using [U⁻¹³C,¹⁵N]-melittin at physiologically relevant concentrations (0.05 mM-0.09 mM) in “NMR buffer” (10 mM KH₂PO₄/K₂HPO₄ buffer at pH 7.0 with 10% (v/v) D₂O). Melittin samples used for 3D structure determination were supplemented with 10% w/v deuterated glycerol and 30% v/v deuterated 2,2,2-trifluoroethanol (from Cambridge Isotope Laboratories).

Recombinant human αA-crystallin was prepared using a previously described method.

Samples of MAC (unlabeled and ¹⁵N-Val-labeled ) were prepared by sonicating the lyophilized peptide (>95% purity, purchased from Genscript) in NMR buffer at 298 K and centrifuging the mixture at 15,000xg for one minute to remove large MAC aggregates.
2.2 Overview of NMR Experiments and NMR Data Analysis

Typically, NMR experiments (including $^1$H-$^{13}$C heteronuclear single quantum coherence or HSQC, $^1$H-$^{15}$N HSQC, 3D total correlation spectroscopy or $^1$H-$^{15}$N TOCSY, 3D nuclear Overhauser effect spectroscopy or $^1$H-$^{15}$N NOESY) were performed at 298 K using 700 MHz Bruker Avance II or 600 MHz Bruker Ascend spectrometers equipped with ultrasensitive TXI or QXI cryoprobe, respectively. The WATERGATE water suppression pulse sequence was used to attenuate the NMR signal of water at ~4.7 ppm. The $^1$H, $^{13}$C, and $^{15}$N backbone resonances were assigned using standard triple resonance experiments (HNCA, HNCOCA, HNCACB, and CBCACONH) NMR spectra were processed using Topsin 2.1 (Bruker, Inc.). Assignments were done manually using CARA.

2.3 Liquid Chromatography Mass Spectrometry (LC-MS)

Electrospray ionization mass spectrometry (ESI-MS) linked to an LC system was used to evaluate purity of protein and peptide samples, and to determine molecular weights. In the case of $[^{15}$C,$^{15}$N]-melittin and $[^{15}$N]-melittin, the molecular weights determined from deconvoluted mass spectra ($M_X$) were used to calculate the percent isotope enrichment (%IE), as shown in Equation 1:

$$%IE = 100 - \left[100 \times \left(\frac{M_{CL} - M_X}{N}\right)\right]$$

Where $M_{CL}$ is the theoretical molecular weight of completely labeled melittin, and $N$ the theoretical number of $^{15}$N and/or $^{13}$C nuclei in the sample: 38 for $[^{15}$N]-melittin, or 169 for $[^{13}$C,$^{15}$N]-melittin.
2.4 Identification of Putative Interaction Sites on Melittin by NMR Line Broadening Analysis

The interactions between [U-15N]-melittin and unlabeled recombinant HAA or MAC were probed by performing titrations (melittin as titrand, αA or MAC as titrant) while monitoring the 2D 15N-HSQC spectra of melittin (Chapter 5, Figure 14 and Chapter 7 Figure 23). Samples of 0.040 mM [U-15N]-melittin were prepared in NMR buffer. HAA or MAC from a stock solution (0.287 mM HAA or 0.300 mM MAC in NMR buffer) was added successively such that mole ratios of melittin: αA monomer from 4:1 to 1:1 were covered in the titration. The 2D 15N-HSQC spectrum of free [U-15N]-melittin (0.040 mM) was used as a control.

Similarly, the interaction between [U-15N]-αAG98R and unlabeled melittin was investigated by performing titrations with αAG98R as titrand and melittin as titrant while monitoring the 2D 15N-HSQC spectra of αAG98R (Chapter 6, Figure 21).

2.5 Expression and Purification of the G98R Mutant of αA-Crystallin

E. coli BL21 (DE3) Codon Plus cells transformed with the plasmid for G98R-αA-crystallin were grown overnight in Luria Broth at 37°C. The overnight culture was diluted with LB to an OD₆₀₀ of 0.1, and grown to an OD₆₀₀ of 0.8. The cell pellet was collected by centrifugation at 3000xg and was resuspended in M9 medium (containing 15NH₄Cl as the sole nitrogen source) to produce [U-15N]-labeled protein. Overexpression was induced by addition of 1mM IPTG. After 4h of incubation at 37°C, the cell pellet was collected. Cells were lysed by sonication in lysis buffer (50 mM Tris, 1mM EDTA, 50 mM dithiothreitol, 1 w/v % Triton-X, pH 7.2). The soluble and
insoluble (inclusion body) fractions of the cell lysate were separated by centrifugation at 9000xg. The inclusion body fraction was washed with the following buffers, in sequence: A) 2 w/v % Triton-X, 50 mM Tris, 1 mM EDTA, 50 mM DTT at pH 7.2, B) 25 mM Tris, 0.5 mM EDTA 50 mM DTT, 1 M NaCl at pH 7.2 (2 washes), C) 50 mM Tris, 1 mM EDTA, pH 7.2. The washed inclusion body, which was ~%90 pure (from SDS-PAGE analysis) was stored at -80°C until use. To refold the protein, the inclusion body was dissolved in 6M urea with 2 mM DTT, 50 mM Tris, 1 mM EDTA at pH 7.2. The solution was loaded onto a Q-column (Q-Sepharose Fast Flow, 60 mL resin) pre-equilibrated with 50 mM Tris, 1mM EDTA at pH 7.2. After loading, the column was washed with 1 column volume of 150 mM NaCl, 50 mM Tris, 1 mM EDTA pH 7.2. This step removed impurities. A second wash with 1 column volume of 350 mM NaCl, 50 mM Tris, 1 mM EDTA pH 7.2 eluted the pure αA-G98R, enriched in the monomeric form (Chapter 6, Figure 19). The eluted protein was characterized by NMR, circular dichroism spectroscopy, spectrofluorometry, and size exclusion chromatography (Chapter 6, Figures 19-21).

2.6 Isothermal Titration Calorimetry for Characterizing the Binding Thermodynamics of the Chaperone-Substrate Interaction

Recombinant HAA (>95% purity assessed from SDS-PAGE and LC-MS) and synthetic melittin (>95% purity, purchased from Genscript) were prepared in identical buffers (10 mM KH₂PO₄/K₂HPO₄ buffer at pH 7.0). Protein concentrations were determined from absorbance measurements at 280 nm using a Cary 50 UV-Vis spectrophotometer, and extinction coefficients of 14,564 M⁻¹cm⁻¹ for HAA and 5,500 M⁻¹cm⁻¹ for melittin, obtained using the EXPASY ProtParam tool. All ITC experiments were performed on a Low-Volume Nano ITC (TA instruments) with a cell volume of 167 µL and syringe volume of 50 µL. The stirring speed was
kept at 250 rpm. The volume of the first injection was set at 0.5 μL, and the volume of the rest of the injections (19 injections) was 2.5 μL. Control experiments with deionized water were performed prior to every titration. Titrations were performed with HAA as the titrant (0.603-0.680 mM) and melittin (0.071 mM) as the titrand. To determine the heat of dilution of αA-crystallin, a dilution control experiment was performed in which HAA was the titrant and the potassium phosphate buffer was the titrand. The thermograms were processed and analyzed using the NanoAnalyze software from TA instruments. For HAA binding to melittin, an independent binding sites model was used in fitting the isotherm to obtain thermodynamic parameters (See Chapter 7 Figure 22).
Chapter 3 – Cloning, Expression, and Purification of Melittin

3.1 Introduction

A model substrate completely labeled with NMR-active isotopes is required for performing high resolution NMR experiments on chaperone-substrate models. One of the aims of this project was to develop a protocol for recombinant expression of the melittin peptide with uniform incorporation of $^{13}$C- and $^{15}$N- isotopes. Overexpression of melittin has been attempted by several research groups although most of these studies have not comprehensively explored the application of their recombinant techniques to isotope-labeling of recombinant melittin for NMR spectroscopic applications.94-97

To our knowledge, only the group of Ishida et al.94 reported on the production of $^{[U,^{13}C,^{15}N]}$-melittin using recombinant techniques, but their melittin sample contained two extra residues which may affect the higher order structure of melittin, and their work did not extend to resonance assignment of melittin. Here we describe the development of a strategy to produce $^{[U,^{13}C,^{15}N]}$-melittin without additional residues, with a yield of ~0.1 mg per liter of E. coli and at least 95% purity. In this chapter, we present the protocol for cloning, expression, and purification of uniformly isotope-labeled recombinant melittin. This chapter also covers the characterization of recombinant melittin by mass spectrometry and NMR spectroscopy.
3.2 Materials and Methods

Cloning

Directional cloning was performed using standard protocols. The amino acid sequence of melittin was designed as previously reported (GIGAVLKVLTGLPALISWIKRKRQQ). The melittin gene was ligated into a pTM vector (restriction map shown in Scheme 1).

Scheme 1. Restriction map of the pTM vector.
The resulting expression plasmid, “pTM-melittin,” utilizes a T7 promoter and is designed to express a 9x-histidine tag (“9x-His”) followed by a portion of the TrpLE 1413 peptide, abbreviated as “TrpLE,” fused to the translation product of interest, the 26-residue melittin peptide. The fusion protein (~16 kD) expressed by this construct is thus called “9x-His-TrpLE-melittin.” The scheme for the pTM-melittin expression system is shown in Scheme 2.

![Scheme 2. Design of the pTM-melittin expression system.](image)

**Scheme 2. Design of the pTM-melittin expression system.** The open reading frame of the pTM-melittin plasmid codes for a 9x-histidine tag, followed by the modified TrpLE sequence, a methionine residue, and finally the melittin sequence. Adapted from *Biochemistry* Vol. 57, Ramirez et al., 90 “Nuclear magnetic resonance based structural characterization and backbone dynamics of recombinant bee venom melittin”, pp. 2775-2785, Copyright (2018) with permission from the American Chemical Society.

The DNA sequence of pTM-melittin was confirmed at the Molecular Core Facility at the University of Albany, SUNY.

*Expression and Purification of Melittin*

The expression and purification of recombinant uniformly isotope-labeled melittin were performed as described in our published work.90 Briefly, *Escherichia coli* (*E. coli*) strain C41(DE3) was transformed with the pTM-melittin expression plasmid. Colonies expressing the melittin-containing fusion protein were cultured and grown in M9 medium containing $^{15}$NH$_4$Cl.
and \([U-^{13}C]\)-glucose to produce \([U-^{13}C,^{15}N]\)-labeled fusion protein. Overexpression was induced using 1 mM isopropyl-\(\beta\)-D-1-thiogalactopyranoside (IPTG). Following expression, cells were pelleted and lysed. The fusion protein (9x-His-TrpLE-melittin) from the lysate was purified by nickel affinity chromatography under denaturing conditions. The fusion protein was cleaved using cyanogen bromide (CNBr) to release the recombinant melittin fragment. The melittin peptide from the cleavage mixture was purified by reversed phase high performance liquid chromatography (RP-HPLC). Recombinant melittin with at least 95% purity and 90% isotope enrichment (as determined by LC-MS using the procedure in Chapter 2, Section 2.3) was lyophilized and stored at −80°C until it was used in NMR experiments.

**Chemical Shift Resonance Assignment**

The backbone \(^1\text{H},\ ^{15}\text{N},\text{ and }^{13}\text{C}\) nuclei on \([U-^{13}C,^{15}N]\)-melittin in NMR buffer were assigned using the procedure in Chapter 2 Section 2.2.

**3.3 Results**

**Expression and Purification of Isotope-Labeled Recombinant Melittin Using a Novel Protocol**

Using the procedure outlined above, melittin was cloned into a pTM vector, expressed as a fusion protein (“9x-His-TrpLE-melittin” fusion protein) in *E. coli* cells, and purified. For recombinant expression, we used the *E. coli* strain C41(DE3), instead of the commonly used BL21(DE3), because the former is known to be more tolerant to expression of proteins that are host-toxic. Figure 4 shows the results of the overexpression of the pTM melittin fusion protein.
in *E. coli*, extraction of the fusion protein from the cell lysate, cleavage of the fusion protein, and final purification of the melittin peptide after cleavage of the fusion protein.

**Figure 4. Expression and purification of recombinant melittin.** 15% acrylamide gel with protein ladder in Lane 1, un-induced culture in Lane 2, 4h-induced culture in Lane 3, fusion protein (9x-His-TrpLE-melittin) in Lane 4, cyanogen bromide cleavage reaction mixture in Lane 5, purified recombinant [U-13C,15N]-melittin in Lane 6. The positions of the fusion protein (blue arrow), 9x-His-TrpLE fragment (yellow arrow), and purified recombinant melittin (red arrow) are indicated. Adapted from *Biochemistry* Vol. 57, Ramirez et al.,90 “Nuclear magnetic resonance based structural characterization and backbone dynamics of recombinant bee venom melittin”, pp. 2775-2785, Copyright (2018) with permission from the American Chemical Society.

The SDS-PAGE results shown in **Figure 4** indicate that recombinant melittin was overexpressed as the ~16-kD 9x-His-TrpLE-Melittin fusion protein (Lanes 3 and 4). Melittin, being an antimicrobial peptide, has intrinsic toxicity toward *E. coli*.100 therefore it was necessary to use a fusion partner that could ‘neutralize’ its toxicity during protein expression. Here, we use TrpLE, an ~11 kD protein, as a fusion partner for melittin. TrpLE has been previously shown to reduce solubility of proteins when used as a fusion partner, and to promote inclusion body formation in *E. coli*.98 Importantly, the TrpLE sequence was modified such that two cysteine and six methionine residues were replaced with alanine and leucine residues respectively, therefore the fusion protein contains only 2 methionine residues, one at the start codon and another preceding
the melittin sequence. The modified TrpLE sequence allowed for His tag-based affinity purification and CNBr cleavage of the fusion protein.

The purified 9x-His-TrpLE-Melittin fusion protein was incubated in a solution of cyanogen bromide (at least 100-fold molar excess) in 70% (v/v) formic acid to facilitate cleavage at the C-terminal end of the methionine residues in the fusion protein. Because a methionine residue precedes the melittin target sequence, the CNBr cleavage resulted in the release of the melittin fragment with the sequence GIGAVLKVLTTGLPALISWIKRKRQQ, which is nearly identical to that of the wild type (i.e. melittin isolated from bee venom), except that the C-terminal -CONH$_2$ group in the wild type is replaced by a -COO$^-$ group. Notably, Gly-1 in some batches of melittin isolated from bee venom was reported to be formylated. The cleavage mixture was then subjected to RP-HPLC and ultrafiltration to separate the recombinant melittin from other components of the mixture (uncleaved fusion protein, TrpLE fragment as shown in Figure 4 Lanes 5-6). The protocol described in this chapter and in our published work yielded recombinant melittin with at least 95% purity. The overall yield for recombinant melittin was ~0.1 mg melittin per liter of *E. coli*. We note that the recombinant melittin sequence is identical to that of the synthetic melittin (product number RP20415) purchased from Genscript (Piscataway, NJ).

*Characterization of Isotope-Labeled Recombinant Melittin by MS and NMR*

Typically, for acquisition of high-quality NMR spectra, ~100% isotope enrichment for $^{13}\text{C}$ and $^{15}\text{N}$ are required. The extent of isotope enrichment of purified recombinant melittin was determined by ESI-MS. Figure 5 shows the ESI-MS results for characterizing recombinant melittin in its unlabeled, [$U-^{15}\text{N}$]-labeled, and [$U-^{13}\text{C},^{15}\text{N}$]-labeled forms.
The m/z ratios in the ESI mass spectrum for recombinant unlabeled melittin (Figure 5A) were found to be consistent with previously published data.\textsuperscript{102} The percent isotope enrichment values were calculated using deconvoluted masses of isotope-labeled samples (Figure 5B-C), and were determined to be 93.2\% and 90.1\% for [\textit{U-}^{15}\text{N}] and [\textit{U-}^{13}\text{C},^{15}\text{N}]-labeled melittin, respectively.

\textbf{Figure 5. ESI mass spectra of recombinant melittin.} A) unlabeled melittin B) [\textit{U-}^{15}\text{N}]-melittin and C) [\textit{U-}^{13}\text{C},^{15}\text{N}]-melittin. Mass-to-charge ratios (left panels) and deconvoluted masses (right panels) are shown. Adapted from Biochemistry Vol. 57, Ramirez et al.,\textsuperscript{90} “Nuclear magnetic resonance based structural characterization and backbone dynamics of recombinant bee venom melittin”, pp. 2775-2785, Copyright (2018) with permission from the American Chemical Society.
The efficient incorporation of isotope labels on recombinant melittin allowed for acquisition of high-resolution NMR spectra of melittin using triple resonance experiments, from which complete backbone resonance assignments for melittin were obtained. The resonance assignments (backbone $^1$H, $^{13}$C, and $^{15}$N) are included in our recently published work. The assigned $^1$H-$^{15}$N HSQC spectrum of melittin is shown in Figure 6.

**Figure 6.** $^1$H-$^{15}$N HSQC spectra of recombinant melittin A) Uniformly labeled recombinant melittin in NMR buffer, expressed and purified using our protocol. Spin systems for all residues
except for G1, I2, and P14 are assigned. Peaks enclosed in Box (a) correspond to the amide side chains of Q25 and Q26 whereas those in Box (b) are assigned to the side chain indole of W19. The subscript “c” indicates the assignments from the cis conformation. For residues A15, L16, and L13, two peaks were found; the less intense peaks (labeled A15c, L16c, and L13c) are from cis-Pro melittin. Adapted from Biochemistry Vol. 57, Ramirez et al.,90 “Nuclear magnetic resonance based structural characterization and backbone dynamics of recombinant bee venom melittin”, pp. 2775-2785, Copyright (2018) with permission from the American Chemical Society.

B) Uniformly labeled melittin (with Gly-Thr at the N-terminus) from the work of Ishida et al.94 The spectrum is not assigned by the authors and shows extreme broadening of glycine resonances. 6.5 mM melittin in 90% water, 10% D2O, with 0.5mM 2,2-dimethyl-2-silapentanesulfonic acid and 0.03% sodium azide. Adapted from J. Am. Chem. Soc. Vol. 138, Issue 35, Ishida et al.,94 “Overexpression of antimicrobial, anticancer, and transmembrane peptides in Escherichia coli through a calmodulin-peptide fusion system”, pp. 11318-11326, Copyright (2016) with permission from the American Chemical Society.

It is worth noting that the secondary structure (e.g. unstructured or helical) and quaternary structure of melittin (e.g. tetrameric melittin) are highly dependent on solution conditions such as pH, temperature, peptide concentration, and ionic strength.80, 81, 103 The 1H-15N HSQC spectrum of uniformly labeled melittin in Figure 6A shows a narrow chemical shift dispersion along the 1H dimension, which suggests an unstructured form of the peptide. This is in accordance with previous NMR studies on unlabeled melittin, which indicate that melittin at low (micromolar) concentrations in aqueous media, room temperature, and low ionic strength (e.g. <0.1 M) assumes a random coil configuration.80

The spectrum of our recombinant melittin sample also appears to be better resolved than a previously published, unassigned 1H-15N HSQC spectrum (Figure 6B, from the study of Ishida et al.94) which used a melittin sample with a with an additional Gly-Thr dipeptide at the N-terminus. Some amide resonances in their spectrum (Figure 6B) appear to be severely broadened, unlike ours in Figure 6A. The spectrum of Ishida et al.94 was also acquired with higher melittin concentration (6.5 mM) and was prepared in different solution conditions (90% v/v water, 10% v/v D2O, with 0.5mM 2,2-dimethyl-2-silapentanesulfonic acid and 0.03% w/v sodium azide). We
note that the authors attribute the signal broadening in their spectrum to melittin tetramerization. Previous studies have suggested that melittin in a tetrameric system is helical.\textsuperscript{85,104} However, the obvious similarities in the general pattern of NH peaks of the spectra in Figure 6A-B suggest that in both cases, melittin is unstructured, not helical. Therefore, while our analysis of the $^1$H-$^{15}$N HSQC spectrum in Figure 6A is consistent with findings from previous studies,\textsuperscript{80} the same cannot be said of the spectrum in Figure 6B.

Notably, in our analysis of $[U-^{13}$C,$^{15}$N]-labeled recombinant melittin, we found that two populations of melittin are present in solution (10 mM phosphate buffer at pH 7.0), which arise from the cis-trans isomerization of the Leu13-Pro14 peptide bond. The population of the cis was much smaller than that of the trans, as indicated by the lower intensity of peaks labeled A15\textsubscript{c}, L16\textsubscript{c}, and L13\textsubscript{c} of the $^1$H-$^{15}$N HSQC spectrum in Figure 6A.

3.4 Discussion

The recombinant expression of antimicrobial peptides is generally deemed to be challenging, in part because of the inherent toxicity of such peptides towards bacterial hosts.\textsuperscript{100} The incorporation of a fusion partner within the reading frame of the expression system for melittin was suggested by previous work as a means of attenuating the toxicity of melittin towards E. coli.\textsuperscript{94,95} However, previous studies using X-melittin fusion systems (with X as the fusion peptide sequence) have not produced uniformly isotope-labeled melittin suitable for triple resonance NMR experiments. Therefore, our published protocol\textsuperscript{99} is the first to be used for the production of $[U-^{13}$C,$^{15}$N]-labeled recombinant melittin with a high percent isotope enrichment (>90% for both $^{13}$C and $^{15}$N), as well as high purity (~95%) required for NMR structural characterization.
It is noteworthy that most structural studies available for melittin have not emphasized a clear, structural distinction between cis- and trans- conformations of the Leu13-Pro14 bond in melittin, which may have – as yet unknown – implications for a variety of applications. Previous structural studies (by NMR and X-ray crystallography) on melittin have mostly reported on the trans conformer, and only little was known about the secondary structure of the cis conformer of melittin. Lack of detailed studies on the cis-trans isomerization may be attributed to the relatively low amount of the cis- isomer compared to the trans- isomer, and the difficulty in resolving overlapping NMR signals of melittin samples lacking NMR-active isotopes. Bazzo et al. note that no cis- isomer was observed for melittin in methanolic solution. On the other hand, in the solution NMR study of Lauterwein et al. the authors observed the presence of two populations of melittin in aqueous solution, and these populations were presumed to arise from the cis-trans isomerization of Leu13-Pro14, with a cis conformer : trans conformer ratio of roughly 0.05-0.1. Thus, our observation of two conformations for residues Leu13, Ala15, and Leu16 is consistent with their study.
Chapter 4A – Structural Characterization and Backbone Dynamics of Melittin

4A.1 Introduction

As discussed in Chapter 1, Section 1.7, melittin may adopt a random coil or helical conformation depending on solution conditions such as pH, ionic strength, and peptide concentration. Lauterwein et al. reported that monomeric melittin in a random coil configuration was the predominant form of melittin in dilute solutions, at low ionic strength, and at acidic pH. The study of Brown et al. examined solution conditions that favored the self-association of melittin into its aggregated (possibly tetrameric) form including high melittin concentrations, basic pH, and high ionic strength. Bazzo et al. reported on the structure of monomeric melittin in methanol determined by homonuclear NMR spectroscopy, which shows an α-helical conformation. In the crystalline state, melittin forms tetramers, with each monomeric unit in α-helical form. The ability of melittin to change secondary structure as a way of adapting to different solution conditions has prompted researchers to use melittin as a peptide model for providing insight into the protein folding process.

It is important to point out that most previous structural studies have focused on melittin in the high concentration regime (~1 mM to 40 mM melittin), despite the fact that even micromolar amounts of the peptide are already sufficiently toxic towards bacteria and erythrocytes. Arguably, the lack of structural studies on melittin in the physiologically relevant, low concentration (micromolar) regime is related to the absence of uniformly isotope-labeled melittin samples, which were only made available recently using our published protocol. This has enabled us, for the first time to probe the structural aspects of melittin in the low
concentration (i.e. micromolar), regime. Therefore, in this chapter, we present a detailed characterization of the coil to helix transition of recombinant, uniformly labeled melittin in this physiologically relevant concentration regime. Such a study is important in describing the structural propensities of the model substrate, even in the absence of a molecular chaperone binding partner, and helps to better understand the structural transitions following binding to various biological partners. Here we also distinguish between the cis- and trans- forms of the helical melittin peptide using high-resolution NMR and the isotope-labeled melittin probe, and discuss the structural implications of this isomerization.

NMR-based structural characterization of the model substrate, melittin, was carried out under two main conditions: a) ‘free’ form of melittin in aqueous solution, and b) melittin in a mixture of water and 2,2,2-trifluoroethanol (TFE). TFE is known to induce helical structure in peptides and proteins, and has been used as a ‘membrane mimic’ in modeling protein-membrane interactions.\textsuperscript{105,106} Moreover, TFE is also used as a mimic for the hydrophobic interiors of proteins.\textsuperscript{106} In Section A of Chapter 4, we present the NMR-based characterization of melittin’s backbone dynamics and secondary structure under the stated conditions, with a comparison between trans and cis isomers. In Section B, we show the three-dimensional (3D) atomic resolution solution NMR structure of trans-Pro14 melittin in a TFE/water mixture, corresponding to the helical form of melittin that may represent the conformation of melittin when bound to biological partners such as proteins and lipids.\textsuperscript{75,77}
4A.2 Materials and Methods

Overview of NMR Experiments and NMR Data Analysis

Uniformly isotope-labeled recombinant melittin was prepared using the protocol for recombinant melittin expression and purification discussed in Chapter 3. The NMR experiments and data analysis followed the procedures in Chapter 2, Section 2.2, with a few additions and modifications: NMR experiments were performed on \([U^{13}C, ^{15}N]\)-melittin (0.050 mM) in the NMR buffer with 30% (v/v) deuterated TFE.

Chemical Shift Index Analysis

Using the resonance assignments for melittin, the secondary structures of various segments of the peptide chain were predicted and categorized as α-helix, random coil, or β-sheet. The chemical shift index (CSI) scores assigned to these secondary structures are -1, 0, and +1, respectively.\(^\text{107}\) (See Figure 8)

Backbone Dynamics

Backbone dynamics data (see Figure 9) were obtained for melittin samples in the presence (30% v/v) and absence of TFE according to published procedure.\(^\text{99}\) Briefly, the \(^1H-^{15}N\) steady-state heteronuclear NOE values were obtained by recording interleaved spectra with and without applying \(^1H\) saturation (“sat” and “unsat” respectively). The steady-state heteronuclear NOEs were calculated as the ratio of peak amplitudes, \(I_{\text{sat}}/I_{\text{unsat}}\), and the standard deviation of the heteronuclear NOEs were determined using the method of Farrow et al.\(^\text{108}\) Positive heteronuclear NOE values (approaching 1) correspond to segments with restricted flexibility (e.g. helical regions) while
lower, more negative values are correlated with highly flexible segments (e.g. unstructured regions).

4A.3 Results

**Backbone Resonance Assignment of Melittin**

The backbone $^1$H, $^{13}$C, and $^{15}$N were assigned for melittin in thr NMR buffer at 298 K, with 30% (v/v) TFE. **Figure 7** shows the corresponding, assigned $^1$H-$^{15}$N HSQC spectrum for the melittin probe in the presence of TFE.

![Figure 7](image)

**Figure 7.** $^1$H-$^{15}$N HSQC spectrum of melittin in 30% (v/v) TFE (helical state). Spin systems for all residues except for G1, I2, and P14 are assigned. Peaks enclosed in Box (a) correspond to the amide side chains of Q25 and Q26 whereas those in Box (b) are assigned to the side chain indole of W19. The subscript “c” indicates the assignments from the cis conformation. Adapted from Biochemistry Vol. 57, Ramirez et al., $^{90}$ “Nuclear magnetic resonance based structural characterization and backbone dynamics of recombinant bee venom melittin”, pp. 2775-2785, Copyright (2018) with permission from the American Chemical Society.

As discussed in our published report$^{90}$ and in **Chapter 2**, the observation of two sets of peaks in the $^1$H-$^{15}$N HSQC spectrum of **Figure 7** is attributed to the *cis-trans* isomerization of the Leu13-
Pro14 bond in recombinant melittin. The $^1$H-$^{15}$N HSQC spectrum in Figure 6A (no TFE) shows a more limited $^1$H chemical shift dispersion (backbone H$^N$ resonate in the range ~7.6-8.6 ppm) than the corresponding spectrum in Figure 7 (with 30% v/v TFE, showing backbone HN resonating in the range ~7.0-8.8 ppm), which indicates that the melittin peptide (mixture of cis and trans conformers) underwent a structural transition, from unstructured (Figure 6A) to structured (Figure 7), as a result of adding TFE.

**Chemical Shift Index Analysis**

To verify that the ‘structured’ form of melittin in TFE/water (corresponding to Figure 7) was $\alpha$-helical, CSI analysis was performed. CSI designates secondary structural elements (-1 for $\alpha$-helix, 0 for random coil, and +1 for $\beta$-sheet) for every residue in a peptide sequence.$^{107}$ The predicted secondary structures for melittin are shown in Figure 8.
**Figure 8. Secondary structure of melittin in 0% (v/v) TFE and 30% (v/v) TFE as predicted by CSI**\textsuperscript{107} version 3.0. A) Cartoon showing the secondary structures adopted by melittin under these different conditions. Melittin is initially unstructured, but after addition of TFE, two helices form for the \textit{trans} conformer while only one stable helix remains in the \textit{cis} conformer. B) Consensus CSI values (-1 for \textit{α}-helix, 0 for random coil, and +1 for \textit{β}-sheet) were plotted against residue number. Adapted from Biochemistry Vol. 57, Ramirez et al.,\textsuperscript{90} “Nuclear magnetic resonance based structural characterization and backbone dynamics of recombinant bee venom melittin”, pp. 2775-2785, Copyright (2018) with permission from the American Chemical Society.

**Figure 8** shows that melittin in aqueous solution without TFE assumes an extended, unstructured form. In contrast, melittin in 30% v/v TFE contains \textit{α}-helical segments: 2 helical segments - one each in the N- and C-terminal domains - for \textit{trans}-Pro-melittin, and one N-terminal helix for \textit{cis}-Pro-melittin. The presence of 2 helices has been alluded to in previous structural studies on melittin,\textsuperscript{65} however this marks the first report of melittin having only one C-terminal helix as a consequence of the \textit{trans to cis} isomerization of the Leu13-Pro14 peptide bond.
**Backbone Dynamics of Melittin**

Heteronuclear $^{1}\text{H}-^{15}\text{N}$ NOE experiments were performed to characterize the dynamics of melittin in a pico- to nanosecond time scale. The size and magnitude of the heteronuclear $^{1}\text{H}-^{15}\text{N}$ NOE signals allow for distinguishing between flexible and non-flexible segments on melittin. **Figure 9** shows heteronuclear $^{1}\text{H}-^{15}\text{N}$ NOEs observed for melittin in 0% v/v TFE (free state) and in 30% v/v TFE.

![Figure 9. Backbone dynamics comparison between melittin in solution without TFE and with 30% v/v TFE. Steady-state heteronuclear $^{1}\text{H}-^{15}\text{N}$ NOE for the amide backbone of melittin in NMR buffer (A) without TFE and (B) with 30% v/v TFE. Values for Pro14 are not shown. Residues at the N- and C- termini of melittin were disregarded in the analysis because the extreme flexibility demonstrated by these resulted in unreliable estimates of heteronuclear NOE intensities. Adapted from *Biochemistry* Vol. 57, Ramirez et al., “Nuclear magnetic resonance based structural characterization and backbone dynamics of recombinant bee venom melittin”, pp. 2775-2785, Copyright (2018) with permission from the American Chemical Society.](image)
Segments of the protein that do not participate in the folded structure and are relatively flexible have small or negative heteronuclear $^1$H-$^{15}$N NOEs, while those with large, positive heteronuclear $^1$H-$^{15}$N NOEs are usually associated with rigid, $\alpha$-helical structures. For free melittin (0% v/v TFE) shown in Figure 9A, most heteronuclear $^1$H-$^{15}$N NOEs were negative, indicative of highly flexible regions that show faster motions relative to the overall tumbling of the peptide. No significant difference is found in the heteronuclear $^1$H-$^{15}$N NOEs for cis-Pro melittin and trans-Pro melittin. For both conformers, a small C-terminal segment formed by residues 13-20 had low, positive heteronuclear $^1$H-$^{15}$N NOE values, indicating that the C-terminal segment may be less flexible than other regions of the peptide, although overall this segment cannot be classified as helical because the $^1$H-$^{15}$N NOE values are much lower than 1.0. On the other hand, for melittin in 30% v/v TFE, the heteronuclear $^1$H-$^{15}$N NOEs (Figure 9B) are generally higher (more positive) than those of the free melittin, and this suggests an overall reduction in flexibility attributed to helix formation. The trans conformer was generally less flexible (and more helical) than the cis conformer.

4A.4 Discussion

Trifluoroethanol (TFE) is a co-solvent known to induce $\alpha$-helical structures in proteins and peptides. Previous studies have proposed that TFE induces $\alpha$-helical structure through mechanisms such as promotion of intramolecular hydrogen bonding, molecular crowding, and disruption of hydrophobic interactions. In our structural studies on melittin in a TFE/water environment, we observed that melittin undergoes a structural transition from random coil to helix as a result of binding with TFE. Our CSI analysis of backbone chemical shifts on melittin (trans- and cis- forms) in the absence of TFE predicts that melittin assumes a random coil
conformation. This is supported by our backbone dynamics data for melittin without TFE. On the other hand, in the presence of 30% v/v TFE, \textit{trans}-Pro-melittin was predicted to form two \(\alpha\)-helical segments: residues 2-9 and 13-22. The backbone dynamics data indicate that overall, \textit{trans}-Pro-melittin has restricted mobility, which is characteristic of helical segments. The predominantly helical secondary structure of melittin was also observed in the X-ray crystal structure of tetrameric melittin,\textsuperscript{85, 104, 114} as well as structural studies of melittin in methanolic solution.\textsuperscript{82, 115} Moreover, this secondary structure is also adopted by melittin in membrane models,\textsuperscript{116, 117} in the polysaccharide-bound state,\textsuperscript{118} and in a number of protein-melittin complexes,\textsuperscript{119, 120} thus underscoring the physiological relevance of the helical conformation.

Interestingly, the CSI and backbone dynamics results for \textit{cis}-Pro-melittin markedly deviate from those of the \textit{trans}-Pro conformer. Based on CSI analysis, \textit{cis}-Pro-melittin forms only one helix, situated in the N-terminal region of the peptide (residues 2-11), while the C-terminal region remains unstructured. This appears to be consistent with the backbone dynamics data, which show that the N-terminal region of \textit{cis}-Pro-melittin is more rigid than the C-terminal counterpart. To our knowledge, this is the first report of the secondary structure adopted by \textit{cis}-Pro melittin in solution. This study also demonstrates that the \textit{cis}-\textit{trans} isomerization of the Leu13-Pro14 peptide bond on melittin has significant influence on the secondary structure and backbone dynamics of melittin. Previous studies have also mentioned the importance of Pro14 on determining the overall electrostatic properties of melittin and the angle of rotation formed by the C-terminal helix relative to the N-terminal helix.\textsuperscript{121, 122} The biological implications of these findings, if any, could be addressed by further experiments.
Chapter 4B – High Resolution Structure of Helical Recombinant Melittin

4B.1 Introduction

The helical conformation of melittin is thought to be adopted by melittin when associated with target proteins such as staphylokinase,\textsuperscript{123} centrin,\textsuperscript{119} and calmodulin.\textsuperscript{120} The EPR study of Farahbakhsh et al. \textsuperscript{77} suggests that melittin may form either an $\alpha$-helix or a $\beta$-strand upon binding to $\alpha$-crystallin. Recognizing that the helical form of melittin may be involved in its interaction with the $\alpha$-crystallin molecular chaperone in solution, we sought to determine the solution NMR structure of helical melittin, specifically trans-Pro melittin, in a TFE/water/glycerol environment. This solvent system was carefully selected for a number of reasons. The TFE component induces helical structure in melittin, as discussed in Chapter 4A. Glycerol and TFE increase the viscosity of the medium, which helps to minimize the conformational flexibility of melittin and to enhance the collection of nuclear Overhauser effect (NOE)-derived distance constraints for structure calculations. The workflow we developed for optimizing solution conditions for structure calculation purposes has been published,\textsuperscript{124} and can also be applied to other small peptides studied by solution NMR. With this workflow, we determined the 3D structure of melittin with atomic precision ($\sim0.7$ Angstrom). The structural coordinates are currently available in the Protein Data Bank (PDB) with the PDB ID: 6DST.
4B.2 Materials and Methods

Optimizing the Transverse Relaxation Rate of Melittin

The main strategy for solving a 3D model of melittin in the solution state requires high-resolution 3D $^1$H-$^{13}$C NOESY and 3D $^1$H-$^{15}$N NOESY with complete backbone and side chain assignments. These spectral data are used to generate distance, which are analyzed by CYANA\textsuperscript{125} (version 3.98) for structure calculation. Our attempts at collecting 3D $^1$H-$^{15}$N NOESY spectra for $[U-^{13}$C,$^{15}$N]-melittin in NMR buffer with 30\% (v/v) deuterated TFE alone at 298K were unsuccessful because cross peaks associated with $^N$H(i)-$^N$H(i+1) connectivities were not observed, due to the unfavorable transverse ($T_2$) relaxation of melittin.\textsuperscript{78} To circumvent this problem, we developed an approach\textsuperscript{124} to minimize $T_2$ by increasing solvent viscosity by adding a viscous component (e.g. deuterated glycerol) and decreasing the temperature of the solvent. The minimum $T_2$ was achieved when 10\% (w/v) deuterated glycerol was present in the melittin sample and the sample was cooled to 285 K. Therefore these solution conditions were used in acquiring all NOESY data for structure calculations.

3D-Structure Determination of Melittin

We used CYANA\textsuperscript{125} (version 3.98) to calculate the 3D structure of melittin (Chapter 4B, Figure 10 and Table 1). The input consisted of a chemical shift list obtained from the resonance assignment, a $^1$H-$^{15}$N and $^1$H-$^{13}$C NOESY peak lists, and lower distance limits that account for the lack of NOEs between W19 ring protons and non-labile protons on neighboring residues. Backbone assignments were determined using the procedure in Chapter 2.2, while complete side chain assignments were determined using HCCH-TOCSY correlation spectroscopy. The 3D $^1$H-
$^{13}$C NOESY and $^1$H-$^{15}$N NOESY experiments were performed at 285 K with a mixing time of 200 ms. The melittin sample was 0.050 mM $[U-^{13}$C,$^{15}$N]-melittin in 30% (v/v) TFE with 10% (w/v) glycerol in the NMR buffer. The NOESY peaks were assigned automatically using the standard CYANA protocol with seven cycles of NOE assignment and simulated annealing in torsion angle space. Backbone $\Phi$ and $\Psi$ dihedral angle constraints were determined using TALOS+. The 20 conformers with the lowest residual CYANA target function values obtained in the final cycle of the structure calculation were chosen to represent the NMR structure of melittin. MOLMOL was used to calculate the helix bend angles in the structures. Swiss Protein Data Bank (PDB) Viewer (version 4.1.0) was used to generate the molecular surface and electrostatic potential map of melittin. Figures showing the 3D models of the NMR structure were produced using YASARA. The programs MOLMOL and Swiss-PDB Viewer (version 4.1.0) together with the web-based MolProbity server were used to determine structural statistics.

**Analysis of 3D Structures of Melittin**

The 3D structures of melittin from the following PDB entries were compared, in accordance with our published method: 2MLT (tetrameric melittin in the crystalline state), 2MW6 (solution NMR structure of melittin modified with an organometallic fragment, and solved in methanol), 6DST (solution NMR structure of melittin determined in this chapter, and solved in TFE/water/glycerol), and 3QRX (crystal structure of melittin in a melittin-centrin complex). The Dictionary of Protein Secondary Structure (DSSP) was used to assign secondary structural elements and to determine hydrogen bond patterns in the peptide backbone (e.g. CO$_i$-NH$_{i+3}$ bonds are characteristic of $\beta$-helices, while CO$_i$-NH$_{i+4}$ bonds are for $\alpha$-helices). The categories for secondary structural elements in DSSP are unstructured, $\alpha$-helix, $\beta_{10}$-helix, $\pi$-helix, hydrogen-
bonded turn, bend, extended strand, and β-bridge),\textsuperscript{131} however DSSP prioritizes the assignment of α-helix over $3_{10}$-helix in the case of overlapping helical elements. Recognizing that our 3D melittin structures have a combination of CO$_i$-NH$_{i+3}$ and CO$_i$-NH$_{i+4}$ backbone hydrogen bonds and therefore have overlapping $3_{10}$-/α-helical character, we adopted and appended the criterion used by Cao and Bowie\textsuperscript{132} for distinguishing between a canonical structural element from a non-canonical one. Specifically, we classified a peptide segment as ‘α-helix’ when it contained CO$_i$-NH$_{i+4}$ hydrogen bonds with bond energies < -1.0 kcal mol$^{-1}$, and as ‘non-canonical $3_{10}$-helix’ when the segment was assigned as α-helix by DSSP but contained a combination of CO$_i$-NH$_{i+3}$ and CO$_i$-NH$_{i+4}$ backbone hydrogen bonds with bond energies < -1.0 kcal mol$^{-1}$.

4B.3 Results

3D Solution Structure of Helical Melittin at Atomic Resolution

The atomic coordinates of the ensemble of solution NMR structures of melittin have been deposited in the Protein Data Bank (PDB ID: 6DST). The three-dimensional structure of melittin was solved in a TFE/water/glycerol environment using CYANA.\textsuperscript{125} For simplicity, only the structure of the trans-Pro conformer of melittin was determined. For the NMR structure calculation of melittin in CYANA, 369 NOE upper distance limits and 252 dihedral angle constraints were identified, which resulted in a high-quality structure determination (Table 1).
Table 1. Structural Statistics of the Ensemble of 20 Melittin Conformers\textsuperscript{124} Adapted from \textit{J. Phys. Chem. B.} Vol. 123, Issue 2, Ramirez et al.,\textsuperscript{124} “Helical structure of recombinant melittin”, pp. 356-368, Copyright (2019) with permission from the American Chemical Society

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOE upper distance limits\textsuperscript{a}</td>
<td>369</td>
</tr>
<tr>
<td>Intra residual (i = j)</td>
<td>104</td>
</tr>
<tr>
<td>Short-range [(i – j) = 1 ]</td>
<td>154</td>
</tr>
<tr>
<td>Medium-range [1 &lt; (i – j) ≤ 5]</td>
<td>111</td>
</tr>
<tr>
<td>Long-range</td>
<td>0</td>
</tr>
<tr>
<td>Dihedral angle ($\psi$ and $\phi$) constraints\textsuperscript{a}</td>
<td>252</td>
</tr>
<tr>
<td>Residual target function value ($\text{Å}^2$)\textsuperscript{a}</td>
<td>$1.19 \pm 0.03$</td>
</tr>
<tr>
<td>Residual NOE violations\textsuperscript{a}</td>
<td></td>
</tr>
<tr>
<td>Number ≥ 0.1 \text{Å}</td>
<td>$1 \pm 1$</td>
</tr>
<tr>
<td>Maximum (\text{Å})</td>
<td>$0.25 \pm 0.03$</td>
</tr>
<tr>
<td>Residual dihedral angle violations\textsuperscript{a}</td>
<td></td>
</tr>
<tr>
<td>Number ≥ 5\textdegree</td>
<td>$1 \pm 1$</td>
</tr>
<tr>
<td>Maximum (\textdegree)</td>
<td>$6.28 \pm 0.35$</td>
</tr>
<tr>
<td>Residual van der Waals violations\textsuperscript{a}</td>
<td></td>
</tr>
<tr>
<td>Number ≥ 0.2 \text{Å}</td>
<td>$2 \pm 0$</td>
</tr>
<tr>
<td>Maximum (\text{Å})</td>
<td>$0.26 \pm 0.01$</td>
</tr>
<tr>
<td>RMSD to the mean coordinates (\text{Å})\textsuperscript{a}</td>
<td></td>
</tr>
<tr>
<td>Backbone (N, C\textsubscript{a}, C')</td>
<td>$0.66 \pm 0.23$</td>
</tr>
<tr>
<td>heavy atoms</td>
<td>$1.17 \pm 0.22$</td>
</tr>
<tr>
<td>Ramachandran plot statistics (%)\textsuperscript{a}</td>
<td></td>
</tr>
<tr>
<td>Most favored regions</td>
<td>92.1</td>
</tr>
<tr>
<td>Additional allowed regions</td>
<td>7.9</td>
</tr>
<tr>
<td>Generally allowed</td>
<td>0.0</td>
</tr>
<tr>
<td>Disallowed</td>
<td>0.0</td>
</tr>
<tr>
<td>Clash score\textsuperscript{b}</td>
<td>8</td>
</tr>
</tbody>
</table>

\textsuperscript{a) The average values for the 20 conformers with the lowest residual CYANA target function with corresponding standard deviation are given. These values were calculated using CYANA\textsuperscript{125} (version 3.98) following the 7\textsuperscript{th} cycle of the automated NOE assignment. RMSD: root-mean-square deviation.

\textsuperscript{b) Calculated using the PDB validation server (https://validate-rcsb.wwpdb.org/)
Figure 10 shows that the bundle of 20 CYANA conformers has a helical secondary structure. In all structures, one continuous bent helix is defined by residues 4-23.

![Stereoview of the C-alpha traces of the 20 conformers. The C-alpha traces are mostly overlapping for residues ~4-23. The termini of the peptide appear to be ‘frayed’.](image)

**Figure 10. 20 conformers representing the atomic resolution structure of melittin.** A) Stereoview of the C-alpha traces of the 20 conformers. The C-alpha traces are mostly overlapping for residues ~4-23. The termini of the peptide appear to be ‘frayed’. B) Stereoview of the ribbon diagrams of conformers shown in Panel A. Structured regions are rendered in thick ribbons while unstructured regions are shown as cyan-colored tubes. Secondary structural elements were identified by DSSP \(^{131, 133}\) and are coded by color: blue for helix, green for turns, and yellow for bends. For 100% of the conformers, the secondary structure assignments are as follows: residues 4-23 comprise a helix, unstructured regions are assigned to residue 1, 25-26, and a turn is assigned to residue 3. Residue 2 may form a turn or be part of an unstructured region, while residue 24 is assigned either to a bend or a helix. The average bend angle of the NMR structure is 131 ± 8°, calculated using MOLMOL.\(^{127}\) Adapted from *J. Phys. Chem. B.* Vol. 123, Issue 2, Ramirez et al.,\(^{124}\) “Helical structure of recombinant melittin”, pp. 356-368, Copyright (2019) with permission from the American Chemical Society

The final ensemble of melittin conformers features a ‘bent’ helical structure with an angle of 131 ±8°. The NOESY data relevant to the determination of the magnitude of this bend angle are shown in Figure 11.
Figure 11. NOEs from $^1$H-$^{15}$N NOESY and 3D $^1$H-$^{13}$C NOESY spectra that constrain the geometry of melittin's 'hinge' region (residues 7-13). A) Cartoon and 3D model of the representative structure of melittin showing the hinge region enclosed in a black box, and inter-residual NOEs represented as black lines. Secondary structure is assigned by DSSP and indicated by color: green for turns, yellow for bends, and cyan for unstructured ribbons. The helical regions are either colored blue (for $\alpha$-helix, containing only CO$_i$-NH$_{i+4}$ backbone hydrogen bonds) or red (for noncanonical 3-$\alpha$-10 helix, containing both CO$_i$-NH$_{i+3}$ and CO$_i$-NH$_{i+4}$ backbone hydrogen bonds). P14 is shown as a stick model, colored magenta. The representative NMR structure is taken from the ensemble of 20 melittin conformers in Figure 8. B) Selected $\omega_2(^1$H$_N$)-$\omega_3(^1$H$_{NOE}$) strips at the $\omega_1(^{15}$N$_H$) chemical shifts of T10, T11, and G12 are shown and the corresponding positions of amide protons are indicated in the 3D model. The d$_{NN}$ connectivities in the NOESY strips are indicated by red broken lines. C) The $\omega_2(^1$H$_N$)-$\omega_3(^1$H$_{NOE}$) strip at the T11 $\omega_1(^{15}$N$_H$) chemical shift shows a medium range NOE between H$_N$ of T11 and H$_\alpha$ of V8. The $\omega_2(^3$H$_\beta$)-$\omega_3(^1$H$_{NOE}$) strip at the T11 $\omega_1(^{13}$C$_\beta$) chemical shift also shows a medium range NOE between H$_\beta$ of T11 and H$_\alpha$ of V8. The $\omega_2(^1$H$_\beta$)-$\omega_3(^1$H$_{NOE}$) strip at the T11 $\omega_1(^{13}$C$_\alpha$) chemical shift shows medium range NOEs involving P14 protons and H$_\alpha$ of T11. Adapted from J. Phys. Chem. B. Vol. 123, Issue 2, Ramirez et al., "Helical structure of recombinant melittin", pp. 356-368, Copyright (2019) with permission from the American Chemical Society.
It is evident that medium-range NOEs between V8 and T11 as well as T10 and G12 (Figure 11) contribute to the bent geometry of the melittin helix in the central region of the peptide. Notably, the segment spanning residues 7-13 on melittin is classified as a non-canonical 3_10-helix, using the criteria defined in Section 4B.2, Analysis of 3D Structures. This non-canonical helix is not found in other PDB structural models of melittin (Figure 12).

**Figure 12. Comparison of hydrogen bonding patterns and secondary structural elements in different PDB structures of melittin.** The Pro14 residue is shown as a magenta stick model. The hydrogen bonds (CO_i-NH_{i+3}, CO_i-NH_{i+4}, and CO_i-NH_{i+5}) in representative structures of the PDB entries are classified by color. Secondary structural elements are color-coded as shown. Adapted from *J. Phys. Chem. B.* Vol. 123, Issue 2, Ramirez et al., “Helical structure of recombinant melittin”, pp. 356-368, Copyright (2019) with permission from the American Chemical Society

Figure 12 shows that melittin structures deposited in the PDB, to date, feature a variety of secondary structural elements and hydrogen bonding patterns. Of particular interest is the N-terminal region of the peptide, which may adopt α-, π-, and non-canonical 3_10-helical conformations.
4B.4 Discussion

Structure determination of small peptides in solution is generally challenged by the inherent conformational flexibility of such peptides. In this study, we devised a strategy to restrict the conformational flexibility of the melittin peptide by increasing solvent viscosity, through the addition of TFE and glycerol, and by lowering the temperature of the medium from 298 K to 285 K. In addition to increasing viscosity of the medium, TFE also induced helicity in melittin. Altogether, these steps allowed us to acquire high-resolution NOESY spectra from which we obtained distance restraints for the structure calculation of melittin. The 3D structure of helical recombinant melittin was determined with atomic precision (~0.7 Angstrom) and the coordinates of the ensemble of calculated NMR structures are available in the Protein Data Bank. The ensemble of structures shows a variety of bend angles (131 ± 8°). The segment defined by residues ~4-23 forms a ‘helix’. However, upon closer examination using the secondary structure classification scheme described earlier, this ‘helix’ is made up of N-(4-6) and C-(14-23) terminal α-helices flanking a non-canonical 3_10-helix (residues 7-13).

We observed that the bent conformation of melittin in TFE/water/glycerol is reminiscent of the ‘kinked’ conformation of melittin in the tetrameric, crystalline state. The ‘bend’ or ‘kink’ in melittin’s structure is thought to arise from the proline residue situated roughly in the center of the melittin sequence. Proline cannot form hydrogen bonds to backbone carbonyl oxygen atoms, and therefore the incorporation of Pro in a peptide sequence can interrupt the backbone hydrogen bonding pattern required for an α-helix. Interestingly, in the structure reported herein (PDB ID: 6DST), Pro14 is still part of the C-terminal α-helix (residues 14-23) in melittin. Preceding this residue is a non-canonical 3_10-helix, which may have been formed instead of an α-
helix owing to the restriction imposed by Pro14 on the backbone hydrogen bonding pattern. Indeed, helices with a combination of α- and 3_{10}-helical character have been found in other proteins,\textsuperscript{132, 137, 138} and it is believed that helix kinks may stem from contiguous 3_{10}- and α-helical segments.\textsuperscript{137} Both 3_{10}- and α-helical conformations are also important in the protein folding pathway – the 3_{10}-helix may act as an intermediate in the folding of a α-helix.\textsuperscript{139, 140} Altogether, our findings indicate that the structural plasticity of melittin encompasses the formation of non-canonical 3_{10}-helical structures, which are not observed in previous structural studies on melittin. The structure reported herein may also be relevant in understanding the folding pathway of melittin, and in describing its conformational transitions while binding to its biological targets.
Chapter 5 - Examining Peptide-Peptide Interactions between MAC (a 19-Residue Peptide of α-Crystallin) and Melittin

5.1 Introduction

As stated in Chapter 1, Sections 1.3-1.4, a previous study conducted by Sharma et al.\(^7\) identified a 19-residue sequence in αA-crystallin (residues 70-88), which is thought to be a putative substrate-binding site in αA-crystallin. The amino acid sequence of this functional region is highly conserved among small heat shock proteins.\(^{141, 142}\) The peptide corresponding to this sequence, called the mini-chaperone (MAC), was found to suppress the aggregation of model substrates under denaturing conditions. These model substrates include alcohol dehydrogenase,\(^7\) insulin,\(^{143}\) α-lactalbumin,\(^{61}\) and also human gamma crystallin (the natural substrate of α-crystallin in the lens).\(^8, 143\) MAC was also shown to inhibit fibril formation of amyloid β (Aβ) peptides.\(^62\)

The small peptide melittin was used as a ‘probe’ or model substrate to map the MAC region on the αA-crystallin parent protein in cross-linking studies conducted by Sharma et al.\(^7\) However, previous reports have not identified the residues involved in the binding interaction between melittin and MAC. In this study, we model the melittin-MAC complex using interaction data from NMR and fluorescence experiments. The model provides insights into the forces driving the chaperone function of MAC towards client proteins.
5.2 Materials and Methods

In the previous report of Sharma et al.⁷ cross-linking experiments with melittin and α-crystallin identified the sequence ⁷⁰KVFIFLDVKHFSPEDLTVK⁸⁸ in αA-crystallin as a putative substrate-binding site. The peptide fragment corresponding to this sequence, as well as its mutant (K70D) have historically been called the mini- αA-crystallin chaperone (MAC).⁷, ⁶¹, ¹⁴⁴, ¹⁴⁵ The K to D mutation was introduced by Sharma et al.⁷ to increase the solubility of the peptide in aqueous solution, while retaining its molecular chaperone activity toward model client proteins. For these reasons, the work reported here was carried out using the K70D mutant and we use the term “MAC” to refer to it. Throughout this chapter, we also use the following numbering scheme for the MAC sequence: ¹⁸DFVIFLDVKHFSPEDLTVK¹⁹.

Chemical Shift Resonance Assignment of MAC Valine Residues

2D ¹H-¹H TOCSY and 2D ¹H-¹H NOESY of unlabeled MAC (200 μM, prepared as described in Chapter 2, Section 2.1) were recorded at 298 K in 10 mM potassium phosphate buffer, pH 7.0 with and without 25% (v/v) deuterated dimethyl sulfoxide (d₆-DMSO). NMR samples were supplemented with DMSO to prevent precipitation of MAC during long acquisitions (>1 day). Backbone H N of valine residues 3, 8, and 18 on MAC were identified via Hα-H N connectivities (Val8-Lys9, Val18-Lys19).

Analysis of NMR Line Broadening

The residues in melittin that potentially participate in the binding interaction were identified by monitoring the ¹H-¹⁵N band-Selective Optimized-Flip-Angle Short-Transient Heteronuclear
Multiple Quantum Coherence (SOFAST-HMQC) and $^1$H-$^{15}$N HSQC spectra of $[U^{-15}\text{N}]$-labeled melittin upon addition of unlabeled MAC. For $^1$H-$^{15}$N SOFAST-HMQC experiments, MAC (final concentrations range from 0 to 30 μM) was added to 10 μM melittin. For $^1$H-$^{15}$N HSQC experiments, unlabeled MAC (0 to 40 μM) was added to 40 μM $[U^{-15}\text{N}]$-melittin. Peak amplitudes in the $^1$H-$^{15}$N SOFAST-HMQC and $^1$H-$^{15}$N HSQC spectra were normalized with respect to the sidechain amide peak of one of the glutamine residues.

The $^1$H-$^{15}$N HSQC spectra of $^{15}$N-Val-labeled MAC (40 μM) were also monitored upon titration with 0-40 μM unlabeled melittin.

**Fluorescence Titrations**

Changes in the tryptophan fluorescence emission spectra of melittin were monitored upon the addition of unlabeled MAC (0 to 128 μM) to unlabeled melittin (10 μM) in 10 mM potassium phosphate buffer at pH 7.0 and at 25°C. Titrations were performed in triplicate. The spectra were recorded using a Fluorolog spectrofluorometer (HORIBA Jobin Yvon) equipped with a temperature-controlled cuvette holder. Fluorescence emission was recorded in the range 300-400 nm (excitation wavelength at 280 nm) with 2 nm slit widths. Neither melittin nor MAC contained tyrosine residues, therefore the fluorescence emission measured by exciting at 280 nm was solely due to W19 in melittin. The absorption of melittin at 280 nm was maintained low (<0.05 OD) to avoid inner filter effects.

Fluorescence emission spectra were normalized by dividing the fluorescence emission intensity (F) by the maximum ($F_{\text{max}}$) value. The normalized spectra were corrected by subtracting the corresponding normalized spectra of MAC alone. Quenching of W19 fluorescence at 346 nm
was plotted against MAC concentration and the dissociation constant was determined by fitting titration data to a specific binding isotherm with a Hill slope (Graphpad Prism).

**Modeling the Melittin-MAC Complex**

HADDOCK\textsuperscript{148, 149} (high ambiguity driven protein-protein docking) version 2.2 was used to model the melittin-MAC complex. HADDOCK is a docking approach driven by interaction data taken from experiments, including NMR titrations, mutagenesis, and other biophysical methods. User input is in the form of PDB coordinate files of the interacting molecules and a list of potentially interacting residues. Melittin coordinates were taken from the representative member of the NMR ensemble in PDB entry 6DST (this work), corresponding to the solution structure of recombinant melittin in a trifluoroethanol/glycerol/water solvent system. MAC (residues 2-19) coordinates were taken from the crystal structure of bovine truncated αA-crystallin in PDB entry 3L1F. Notably, in 3L1F, residue 70 is Lys instead of Asp, therefore the first residue of MAC was omitted. In HADDOCK, the ‘active’ and ‘passive’ residues, i.e. residues that are significantly perturbed in the process of complex formation, are converted into ambiguous interaction restraints (AIRs) that drive the structure calculation. HADDOCK performs structure calculation for the complex and ranks results based on a combination of electrostatic, van der Waals, and AIR energy terms. The structure calculation protocol includes three major stages: 1) randomization of spatial orientations and rigid body energy minimization, 2) semi-rigid simulated annealing in torsion angle space, and 3) refinement in explicit solvent (water).\textsuperscript{148} In stages 2 and 3, the residues allowed to move to optimize the interface packing are constrained by the AIRs. In the MAC-melittin docking, the active residue in melittin is 19 (on the basis of the observed Trp19 blue shift in fluorescence experiments and the consistent NMR signal broadening on the W19 indole upon
addition of MAC to melittin) and the passive residues are defined automatically (i.e. surface residues within 6.5 Angstroms of the active residues). For MAC, the active residue list includes residues 5, 7, 9, 10-19 (these are residues in the bovine αA-crystallin structure with at least 10% relative accessibility as determined by Swiss Protein Data Bank viewer\textsuperscript{128} ) and the passive residues are also defined automatically.

5.3 Results

Tryptophan Fluorescence Blue Shift

To monitor the interaction between melittin and MAC, we examined the changes in the tryptophan fluorescence emission spectra of melittin as a function of MAC concentration (Figure 13A).
Figure 13. Binding of MAC to melittin monitored by tryptophan fluorescence emission. A) Representative intrinsic tryptophan fluorescence spectra (from 3 trials) monitored upon addition of MAC to (0-128 μM) to 10 μM melittin. The wavelength of maximum fluorescence emission was 346 nm in the absence of MAC, and 332 nm at saturating amounts of MAC. The excitation wavelength was 280 nm for all spectra. B) Binding isotherm of MAC to melittin. The extent of quenching at 346 nm (F-F₀)/F₀ was plotted versus the concentration of MAC. The MAC concentration needed to achieve a half-maximum binding at equilibrium is 13.7 ± 0.4 μM and the Hill coefficient is 2.9 ± 0.3.
Because MAC lacks a tryptophan residue in its sequence, the intrinsic tryptophan fluorescence measurement serves as a qualitative assessment of the environment of W19 on melittin alone as melittin interacts with MAC. The blue shift (Figure 13A) from 346 nm to 332 nm observed in the fluorescence emission spectra suggests that W19 on melittin becomes less exposed to the polar solvent as it binds to hydrophobic surfaces on MAC.

The extent of tryptophan fluorescence quenching at 346 nm was taken as a measure of the fraction of MAC bound to melittin (Figure 13B). The binding isotherm for the MAC-melittin interaction has a slightly sigmoidal shape, which may suggest binding with positive cooperativity. Fitting the titration data to a specific binding model with a Hill slope yielded an apparent dissociation constant (Kd) of ~14 μM and a Hill coefficient of ~3. In the simplest way, positive cooperativity in the melittin-MAC system may be interpreted as follows: the binding of one molecule of MAC enhances the binding affinity of subsequent MAC molecules to the same melittin molecule.

Typically, 335 nm corresponds to a partially buried configuration of tryptophan in relatively hydrophobic interiors or interfaces of a protein. Therefore, the blue shift observed in W19 fluorescence emission may be consistent with sequestration of W19 by MAC upon binding. Blue shift of W19 fluorescence has also been observed in the interaction of melittin with lipid vesicles, calmodulin, and ganglioside monosialylated type 1 (GM1). Under solution conditions that favor the helical, tetrameric form of melittin, W19 fluorescence emission is also blue-shifted.
Melittin Peak Broadening Upon Binding to MAC

We monitored the $^1$H-$^{15}$N HSQC and $^1$H-$^{15}$N SOFAST-HMQC spectra of $^{15}$N-labeled melittin at different mole ratios of MAC:melittin to identify potential residue-specific interactions between melittin and MAC. The titration monitored by $^1$H-$^{15}$N HSQC was completed in under 8 hours, while the titration monitored by $^1$H-$^{15}$N SOFAST-HMQC took ~2 hours. It was important to limit the acquisition time for the following reasons. First, beyond 20 hours, the melittin-MAC mixture precipitates at 298 K. Second, the Pro residues on melittin and MAC (initially in the trans-Pro form) undergo isomerization over time, and binding of multiple forms of each peptide may complicate the analysis of spectra. For the melittin and MAC peptides, we observed that >85% of each peptide was in the trans-Pro isomerization state after incubation in the NMR buffer at 298 K for one day, therefore the NMR titration experiments performed in under 8 hours mainly involve the trans-Pro form of each peptide, and the contribution of the cis-Pro form is neglected in the data analysis. We found that the melittin spectra did not exhibit significant changes (>0.01 ppm) in the chemical shifts of melittin H$^N$ peaks upon addition of MAC, however we observed peak broadening and/or sharpening for most melittin residues (Figure 14 A-C). We monitored the titration by $^1$H-$^{15}$N SOFAST-HMQC (Figure 14A) using a melittin concentration close to that used in the fluorescence titration (10 μM melittin). We observed anomalous peak broadening and peak sharpening for most residues on melittin, which may be interpreted as complicated association-dissociation of the melittin-MAC complex. The titrations monitored by $^1$H-$^{15}$N HSQC (Figure 14B-C) were performed using a concentration of melittin (40 μM melittin) higher than that used in fluorescence titrations, to facilitate rapid acquisition of $^1$H-$^{15}$N HSQC spectra with good signal-to-noise (S/N > 20). The broadening/sharpening of amide backbone peaks from $^1$H-$^{15}$N HSQC titrations are consistent with those of the $^1$H-$^{15}$N SOFAST-HMQC titrations. The
$^1$H-$^{15}$N HSQC titration also allowed for monitoring the broadening of the W19 indole $^1$H (Figure 14C).

**Figure 14.** MAC interaction with melittin monitored by NMR A) MAC was titrated into a solution of 10 μM [U-$^{15}$N]-melittin (C), and peak broadening in the $^1$H-$^{15}$N SOFAST-HMQC spectra was monitored. The % broadening of backbone amide cross-peaks in $^1$H-$^{15}$N SOFAST-HMQC spectra was plotted against melittin residue number. B) MAC was titrated into a solution of 40 μM [U-$^{15}$N]-melittin. The % broadening of backbone amide cross-peaks in $^1$H-$^{15}$N HSQC spectra was plotted against melittin residue number. For Panels A and B, Gly1, Pro14, Ala4, Lys21, Gln25, and Gln26 were omitted from the analysis. The $^1$H peak of Gly1 was not visible in all the $^1$H-$^{15}$N SOFAST-HMQC and $^1$H-$^{15}$N HSQC spectra due to chemical exchange of the terminal amino group with water. Pro14 lacks an amide group. A4, K21, Q25, Q26 had overlapping $^1$H peaks. C) Broadening of the W19 indole in melittin upon addition of MAC was monitored by $^1$H-$^{15}$N HSQC.
It was observed that as the number of equivalents of MAC increased, so did the extent of broadening of the indole H^N peak. This points to a possible interaction between MAC and the indole ring of W19 on melittin, and are consistent with the data from the fluorometric titration of melittin with MAC (Figure 13).

Slow tumbling or exchange can give rise to signal broadening in NMR. Signal broadening in the melittin spectra may be due to a number of factors, including intermediate exchange (K_d \approx 10-100 \mu M) between free and MAC-bound states, structural rearrangement of the melittin/MAC complex after binding, or oligomerization before or after binding. We observed that when we increased the amount of MAC (up to 10 equivalents of MAC), all melittin signals in the ^1H-^15N HSQC and ^1H-^15N SOFAST-HMQC disappeared. This may be due to MAC-melittin binding and/or the precipitation of the sample at such high amounts of MAC. We did not observe significant, reproducible differential broadening in the backbone amide signals from the ^1H-^15N HSQC and ^1H-^15N SOFAST-HMQC titrations of melittin with MAC (Figure 14A-B), therefore our NMR data cannot distinguish which backbone amides on melittin interact specifically with MAC. Most melittin H^N peaks underwent broadening and sharpening at various stages of the titration, indicating complicated phenomena. Interestingly however, despite the unusual behavior exhibited by melittin backbone amide signals upon addition of MAC, the H^N proton on the indole group of W19 on melittin does reveal a clear trend, and the signal shows increasing broadening as the MAC concentration is increased (Figure 14C). This is also consistent with the blue shift in the tryptophan fluorescence emission spectra (Figure 13A-B) of melittin upon addition of MAC. The backbone H^N of W19 also shows broadening to a slightly higher extent than other backbone H^N peaks (except for Gly3), when we titrated with a mole ratio of 2:1 MAC:melittin (Figure 14B).
In order to ascertain whether the binding of MAC to melittin may involve specific hydrophobic residues of MAC, we also monitored the $^1$H-$^{15}$N HSQC of $^{15}$N-Valine-labeled ($^{15}$N-Val) MAC at different mole ratios of melittin:MAC. Since the MAC peptide has three Val residues that are fairly evenly distributed throughout the peptide sequence, we reasoned that these residues would be potentially well suited for probing MAC-melittin interactions with NMR titration experiments. However, to our surprise we found that the three valine residues broaden to approximately the same extent upon addition of melittin. Therefore, NMR titration with $^{15}$N-Val MAC also could not help identify which specific MAC residues bind to melittin. This led us to model the melittin-MAC complex using W19 on melittin alone, and the residues on MAC with at least 10% relative solvent accessibility-designated as active residues.

**Modeling the Melittin-MAC Complex**

As stated above, considering the significant peak broadening observed for the indole H$^N$ and backbone H$^N$ of W19, we constrained the melittin-MAC docking by defining W19 on melittin as an ‘active’ residue in the melittin-MAC binding interface.

The 149 structures generated by HADDOCK (representing 74.5% of the water-refined models calculated) were categorized into 10 clusters based on interface-ligand RMSDs (Figure 15). The most reliable docking solutions are in Cluster 2, consisting of 28 models for the melittin-MAC complex. In HADDOCK, the “most reliable” or “best” docking solution is defined as a cluster of structures with the lowest HADDOCK score.
Figure 15. Docking statistics for the melittin-MAC complex. Energy terms (electrostatics in Panel A, ambiguous interaction restraints or AIRs in Panel B, and van der Waals in Panel C) against i-RMSD. iRMSD is defined as the backbone root mean square deviation at the interface defined automatically by HADDOCK based on all observed contacts. Cluster 2 (the average is designated by the blue triangle) is identified as the most reliable set of docking solutions for the melittin-MAC complex, consisting of 28 models. In all panels, i-RMSD of cluster 2 is well below the 10-Angstrom cut-off.

The representative model of the melittin-MAC complex from Cluster 2 (Figure 16) shows that the MAC peptide wraps around the C-terminal region of melittin. Electrostatic interactions are observed between the negatively charged α-carboxylate group on Asp15 of MAC, and the positively charged ε-amino group on Lys23 of melittin (Figure 16A). The negatively charged terminal carboxylate group (COO-) on MAC is also in close contact with the positively charged guanidinium (-C-(NH$_2$)$_2$+) side chain of R22 on melittin (Figure 16A). The indole group of W19 in
melittin is embedded in a ‘hydrophobic pocket’ formed by the hydrophobic side chains of Leu6, Val8, Phe11, Leu16, and Val18 on MAC (Figure 16B).

Figure 16. Representative structure of the melittin-MAC complex modeled by HADDOCK\textsuperscript{148,149} Panel A shows the electrostatic potentials mapped onto the surfaces of melittin (Mel) and MAC using Swiss PDB viewer.\textsuperscript{128} The interacting charged groups are indicated by arrows. Panel B shows the ‘hydrophobic pocket’ formed by hydrophobic side chains of L6, V8, L16, V18, and F11 on MAC (shown as a surface), which envelop the indole group of W19 in melittin (magenta-colored stick model). The ribbon diagram of the melittin backbone is shown in green.
5.4 Discussion

The cross-linking study of Sharma et al\textsuperscript{7} identified the segment defined by residues 71-88 on bovine $\alpha$A-crystallin as one of the binding sites of melittin (the other site being defined by residues 13-21). In this study, we focus on the interaction between melittin and MAC, a peptide whose amino acid sequence corresponds to residues 71-88 on the parent protein.

Our NMR and fluorescence experiments have shown that W19 on melittin is prominently involved in the binding interaction (Figures 13-14). In particular, the blue shift of W19 fluorescence emission upon addition of MAC to melittin indicates that W19 is partially buried in the melittin-MAC complex. The best docking solution generated by HADDOCK\textsuperscript{148, 149} shows a partially buried W19 melittin residue in a hydrophobic environment formed by the side chains L6, V8, F11, L16, and V18 on MAC (Figure 16), thus suggesting that the hydrophobic interaction is one of the driving forces for melittin-MAC association. Notably, F11, L16 and V18 are part of the segment spanning residues 10-19 of MAC (HFSPEDLTVK), which is also considered to be a bis-ANS binding site on $\alpha$A-crystallin, in addition to being a melittin-binding site.\textsuperscript{7, 158} Bis-ANS is a well-known probe for hydrophobic surfaces on proteins, and has historically been used to study the hydrophobicity of $\alpha$-crystallin in relation to the chaperone function of $\alpha$-crystallin.\textsuperscript{159, 160} The binding of bis-ANS to $\alpha$-crystallin is known to decrease the molecular chaperone activity of $\alpha$-crystallin,\textsuperscript{161} therefore the bis-ANS binding site on MAC, especially residues F11, L16, and V18, may be important for the molecular chaperone function of MAC.

Previous studies on deletion variants of MAC have shown that removal of D1 on MAC only results in about 25% loss of chaperone activity (towards alcohol dehydrogenase), while deletion of both V18 and K19 (which are V87 and K88 on HAA) did not significantly affect the molecular
chaperone function toward the same model substrate.\textsuperscript{145} Deletion of the last 6 residues on MAC (E14, D15, L16, T17, V18, K19 on MAC corresponding to residues 83-88 on HAA), however resulted in complete loss of chaperone activity.\textsuperscript{145} These findings may be consistent with our observation by modeling that the sequence FSPEDLT in MAC is involved in the binding of MAC to model substrates, possibly by forming a hydrophobic binding pocket as modeled in Figure 16.

The melittin-MAC complex (Figure 16) notably involves electrostatic interaction between Asp15 on MAC (which is also part of the FSPEDLT sequence) and Lys23 on melittin, which suggests that electrostatic interactions may accompany hydrophobic associations in the formation of the complex.

We should point out that within the MAC sequence, residues outside of the FSPEDLT segment may also be involved in molecular chaperone function. For instance, the F71G mutation in the parent αA-crystallin (corresponding to F2G on MAC) resulted in the loss of molecular chaperone activity toward insulin, citrate synthase, and alcohol dehydrogenase.\textsuperscript{162} In our representative model of the melittin-MAC complex (Figure 16), F2 on MAC is not part of the hydrophobic pocket surrounding the W19 indole of melittin. Additionally, I4P mutation on the MAC peptide in a previous study also resulted in loss of chaperone function of the peptide toward alcohol dehydrogenase.\textsuperscript{145}
Chapter 6 - Examining the Interaction between the Cataract-Associated Mutant of $\alpha$A–Crystallin, $\alpha$A– G98R, and Melittin

6.1. Introduction

As explained in Chapter 1, Section 1.2, structure determination of the $\alpha$-crystallins using solution NMR spectroscopy poses special challenges due to the hetero-oligomeric nature of the chaperone. Therefore, recent reports$^9$ of a cataract-associated mutant (G98R) of $\alpha$A-crystallin that could exist as a monomer in contrast to the native protein was highly exciting. Moreover, the mutant in monomeric form was found to retain chaperone activity.$^9$ Based on these reports we hypothesized that it may be feasible to record well-resolved NMR spectra for the mutant in the presence and absence of melittin. The structural information obtained from such NMR studies may provide residue-specific details of the binding interaction and enable us to map the substrate-binding site on $\alpha$A-crystallin.

6.2. Materials and Methods

Preparation of Protein Samples

The expression and purification of $[U-^{15}$N]$- \alpha$A-G98R is described in Chapter 2, Section 2.5.
NMR Experiments

$^1$H-$^{15}$N HSQC experiments were performed on 20 µM $[U-^{15}$N]-αAG98R at various temperatures (285 K to 301 K) using a 700 MHz Bruker Avance II spectrometer equipped with a cryoprobe. Different buffer conditions were tested (see Results section). Experiments were performed with WATERGATE water suppression.\(^{92}\)

The titration of 20 µM $[U-^{15}$N] αAG98R with unlabeled melittin, spanning mole ratios of 10:1 to 1:1 (αAG98R-to-melittin ratio), was monitored by $^1$H-$^{15}$N HSQC experiments at 298 K. Samples were prepared in NMR buffer at pH 7.0, supplemented with 1 mM EDTA and 1 mM AEBSF.

Size Exclusion Chromatography

The size of αA-G98R in various solution conditions was characterized by size exclusion chromatography using a Superdex 200 column (24 mL column volume) operated at a flow rate of 0.75 mL per minute. Size exclusion standards were obtained from Bio-Rad.

Circular Dichroism Studies

The far-ultraviolet circular dichroism spectra of αA-G98R was recorded at 298 K using a JASCO J-815 CD spectrometer, a cuvette with a path length of 0.1 cm, and a protein concentration of 0.14 mg/mL. Spectra were the average of 3 accumulations. The buffer was 50 mM Tris, 100 mM NaCl, 1 mM EDTA (pH 7.2).
**Intrinsic Tryptophan Fluorescence**

The fluorescence measurements for 0.14 mg/mL of αA-G98R in 50 mM Tris, 100 mM NaCl, 1 mM EDTA (pH 7.2) were made using a Fluorolog spectrofluorometer (HORIBA Jobin Yvon) with a path length of 0.5 cm. Fluorescence emission was recorded in the range 300-400 nm (excitation wavelength at 280 nm) with 2 nm slit widths. Spectra were the average of 3 accumulations.

**6.3. Results**

*Expression and Characterization of the G98R Mutant of αA-Crystallin*

We expressed the \([U-^{15}N]\) G98R mutant in minimal medium (Figure 17), and characterized the mutant protein using ESI-MS.

![Figure 17. Expression and characterization of the G98R mutant of αA-crystallin. A) acrylamide gel with protein ladder in Lane 1, 4h-induction at 4°C in Lane 2, 22h-induction at 4°C in Lane 3, 4h-induction at 30°C in Lane 4, 22h-induction at 30°C in Lane 5, 4h-induction at 37°C in Lane 6, 22h-induction at 37°C in Lane 7, and no induction in Lane 8. The bands corresponding to mutant G98R are indicated by an arrow at ~20 kD. B) acrylamide gel showing whole lysate in Lane 1, soluble (s) and insoluble (in) fractions of lysate from 4h-induction at 37°C in Lanes 2-3, soluble (s) and insoluble (in) fractions of lysate from 4h-induction at 30°C in Lanes 4-5, soluble (s) and insoluble (in) fractions of lysate from 22h-induction at 4°C in Lanes 6-7, and lysozyme in Lane 8. The bands corresponding to G98R (~20 kD) and lysozyme (~13 kD) are indicated by arrows.*
It is evident in Figure 17 that the overexpression of the mutant was accomplished under the conditions tested. In Figure 17A, the results of the SDS-PAGE analysis indicate that the mutant was expressed at temperatures 4°C, 30°C, and 37°C, and that the expression level was better at the higher temperatures used. In Figure 17B, the SDS-PAGE analysis shows that the G98R mutant was preferentially expressed in insoluble inclusion bodies under all temperatures tested. This is consistent with the observations of Singh et al. and Raju et al.9

In Figure 18A, the mass spectrometric analysis shows that the unlabeled mutant has the expected mass, 20,006.2 ±0.68 Da, which is close to the 20,008.5 Da predicted by the EXPASY Protparam tool. The difference in mass may be attributed to the oxidation of cysteines in the mutant protein sample under the conditions used in acquiring the mass spectrum (0.1% formic acid in water). The mass spectrum of [U-15N]-αA G98R crystallin in Figure 18B corresponds to a deconvoluted mass of 20,239.9 ±0.82 Da, which corresponds to a 94% isotope enrichment for 15N.
Figure 18 ESI mass spectra of A) unlabeled αA G98R crystallin and B) [$^{15}$N]-αA G98R crystallin.
The inclusion body of the mutant protein was refolded on a Q Sepharose Fast Flow column, and eluted at a salt concentration of ~250 mM NaCl. The refolded mutant was subjected to size exclusion chromatography (Figure 19) to determine the size distribution.

**Figure 19. Characterization of αA-G98R by size exclusion chromatography.** Chromatography was performed at room temperature, using 10 mM potassium phosphate at pH 7.0 with 150 mM NaCl. Size exclusion chromatogram showing that αA-G98R is initially enriched with the monomeric form (~20 kD) of the mutant (~2.5 h after on-column refolding, blue trace), but eventually forms a large oligomer >600 kD (orange trace) after prolonged storage (e.g. 1 day). Grey traces correspond to gel filtration standards.

Size exclusion chromatography results indicate that the mutant has a molecular weight close to 17 kD, although the peak is rather broad and may represent a monomer-dimer distribution/equilibrium. However, this monomeric form is eventually lost over time (after 1 day of storage at room temperature, and even faster, under ~2h, after subjecting to 4°C). The oligomeric size of the mutant is >600 kD.

The secondary structure of the mutant was determined by measuring circular dichroism spectra in the far-ultraviolet region (190-250 nm, Figure 20A). CD spectra in the far-ultraviolet region are useful for the qualitative characterization of the secondary structure content of peptides and
proteins. The far-UV CD spectrum of the G98R oligomer (>600 kD) shows negative ellipticity at 210 nm, which is close to the value of 212 nm observed by Singh et al. In the work of Raju et al. the negative ellipticity spans the range ~208-210 nm. Negative ellipticity bands near 217-208 nm are indicative of β-sheet character.

Intrinsic tryptophan fluorescence emission of oligomeric G98R (>600kD) was also measured (Figure 20B). G98R has one tryptophan residue (W9), which is sensitive to the polarity/hydrophobicity of its environment, and thus serves as a probe of the tertiary or quaternary structure of the mutant protein. Upon unfolding, W becomes more exposed to solvent. Tryptophan fluorescence can range from 309 nm (corresponding to extensively buried in a hydrophobic interior or hydrophobic interface), to 335 nm (typical for a partially buried W), and 355 nm (fully exposed to water). The fluorescence spectrum of the oligomeric form (>600 kD) is comparable to the published spectrum of Singh et al. which has an emission maximum at 338 nm, while our spectrum shows a λ\text{max} of 334 nm in the same buffer. This corresponds to a partially buried W9.

Figure 20. Characterization of αA-G98R by circular dichroism and fluorescence emission spectroscopy at 298K. The mutant protein concentration was 0.14 mg/mL in a solution buffered with 50 mM Tris, 100 mM NaCl, 1 mM EDTA (pH 7.2) A) Far-UV CD spectrum of αA-G98R indicating enrichment of β-sheet structure B) Intrinsic tryptophan fluorescence of αA-G98R.
**NMR Measurements**

The $^1$H-$^{15}$N HSQC spectrum of the mutant (20 uM $[U-^{15}$N] $\alpha$A-G98R crystallin), which corresponds to the oligomer peak in **Figure 19**, was obtained at 298 K in NMR buffer: 10 mM potassium phosphate, pH 7.0, with added 1 mM EDTA and 1 mM AEBSF (**Figure 21A**). This spectrum is identical to that of the monomer peak in **Figure 19**. The spectrum shows severely broadened peaks in the amide region, indicating that the mutant undergoes intermediate exchange processes between various oligomers of $\alpha$A-G98R crystallin, which preclude the acquisition of high-quality $^1$H-$^{15}$N HSQC spectra. The spectrum also shows a narrow chemical shift dispersion in the amide region of $^1$H dimension, pointing to the possibility that the observed peaks correspond to residues belonging to unstructured regions of $\alpha$A-G98R. Multiple attempts were made – unsuccessfully – to stabilize the monomeric form during NMR characterization: this included using different pH values (pH 3.5-7), varying the concentration of the mutant protein (5-20 uM, concentration determined on a monomer basis), varying salt concentration (0-1 M NaCl, 0-1 M Guanidinium chloride), and temperature (285 K – 305 K). Addition of stabilizing agents such as dimethyl sulfoxide (DMSO), ethylenediaminetetraacetic acid (EDTA), arginine, hexamethylphosphoramidate (HMPA) and dithiothreitol (DTT) did not improve the quality of the spectra. In all conditions tested, the $^1$H-$^{15}$N HSQC spectra did not show peaks corresponding to folded regions of the protein. The $^1$H-$^{15}$N HSQC spectrum obtained at pH 3.5 with 20 uM $[U-^{15}$N]-$\alpha$A-G98R in 10 mM potassium phosphate shows an increased number of peaks with narrower linewidth (**Figure 21B**), however the chemical shift dispersion is still limited (~7.5-8.5 ppm in the $^1$H dimension) which suggests that the unstructured regions of the mutant are observed on the spectrum, while the folded regions are not.
Figure 21. NMR characterization of αA-G98R. A) ¹H-¹⁵N HSQC spectrum of the G98R oligomer (corresponding to the orange peak in Figure 19) in 10 mM potassium phosphate buffer at pH 7.0 with 1 mM EDTA and 1 mM AEBSF. The spectrum was taken at 298 K, using 20µM [U-¹⁵N]-G98R. The spectrum is identical to that of the monomeric G98R at the same pH. B) ¹H-¹⁵N HSQC spectrum of the G98R monomer in 10 mM potassium phosphate buffer at pH 3.5, 298 K. C) Titration of the G98R oligomer with melittin (mole ratios of 10:1 to 1:1 G98R:melittin) resulted in broadening and sharpening of peaks presumably belonging to the N- and C-terminal unstructured regions of G98R.

We monitored the change in intensity of H² peaks in the ¹H-¹⁵N HSQC spectra of αA-G98R (corresponding to unstructured regions of αA-G98R) upon addition of melittin (Figure 21C). We observed that the amide protons belonging to the unstructured regions broadened and sharpened at different stages of the titration, indicative of complex association-dissociation phenomena.
6.4 Discussion

The heterogeneity and dynamic nature of HAA has foiled previous attempts to determine its high-resolution structure. We hypothesized that, in light of the evidence that the G98R mutant of αA-crystallin forms monomers in solution, it may be feasible to obtain high-resolution NMR data on the mutant protein that would aid our structural characterization of the parent HAA. In this report, we ascertained that αA-G98R forms monomers under certain solution conditions (e.g., in pH 3.0 for ~2 days, and for several hours in a pH 7.0 solution with ~250 mM NaCl), however the monomeric form still does not provide high-quality NMR spectra. Only H^N peaks corresponding to unstructured regions of the protein were observable under the various solution conditions tested (Figure 21A-B). The addition of the model substrate melittin to αA-G98R resulted in sharpening and broadening of H^N peaks from unstructured regions of the mutant protein, indicating that melittin binds to these regions on αA-G98R through complex binding events. The observed broadening of some αA-G98R peaks upon addition of melittin may point to binding in the intermediate exchange regime. However, the data at this point do not provide residue-specific information of the binding interaction. Therefore, further exploration of solution conditions and possibly other isotope-labeling strategies may be effective for measuring NMR spectra from which high-resolution structural information could be extracted for αA-G98R in its free and melittin-bound forms.
Chapter 7 - Examining the Chaperone-Substrate Interaction between \( \alpha \text{A–Crystallin} \) and Melittin

7.1. Introduction

As mentioned in Chapter 1 Section 1.1, the molecular chaperone \( \alpha \text{A-crystallin} \) is believed to protect the lens from light scattering, by suppressing the aggregation of other lens proteins. Our main goal in the present study was to provide structural and thermodynamic insights into the ability of human \( \alpha \text{A-crystallin} \) (HAA) to bind to its partially unfolded clients in the lens. For the reasons detailed in the earlier chapters (Chapter 5 and Chapter 6), this goal could only be partially achieved. Using isothermal titration calorimetry (ITC), we determined the thermodynamic parameters for the binding process between native HAA and a substrate model, melittin. We identified the amino acids on melittin important for binding to HAA by saturation-transfer difference (STD) nuclear magnetic resonance (NMR) experiments, and analyzed NMR line broadening upon titration of melittin with HAA. We also modeled the putative melittin-binding interface of HAA.\(^7\) The information obtained from these experiments were applied in generating structural models of the melittin-HAA complex by molecular docking with high-ambiguity driven docking (HADDOCK). Structural models of the melittin-HAA complex reveal important principles underlying the interaction of HAA with its clients. Thus, here we describe the structural and thermodynamic parameters of the HAA-melittin interaction using a “melittin (model substrate)-centered” approach.
7.2. Materials and Methods

Isothermal Titration Calorimetry

The procedure for ITC is outlined in Chapter 2, Section 2.6. Titrations were performed with HAA as ‘ligand’ because the reverse orientation resulted in immediate precipitation.

NMR Experiments

NMR data were acquired at 25°C on a 700 MHz Bruker Avance II NMR spectrometer with a TXI cryoprobe. The WATERGATE pulse sequence was used for water suppression. All protein and peptide samples were prepared in the NMR buffer consisting of 10 mM potassium phosphate buffer at pH 7.0. Melittin resonance assignments were obtained previously. All [U-15N]-labeled recombinant melittin samples mostly contained the trans-Pro isomer, comprising approximately 90%, based on peak amplitude analysis of backbone amide protons from the trans- and cis- forms. Therefore, all NMR data on melittin correspond to the trans-Pro isomer and the contribution of the cis- isomer is neglected.

Residues on melittin that potentially participate in the binding interaction with HAA were identified by monitoring the 1H-15N Heteronuclear Single Quantum Coherence (HSQC) spectra of [U-15N]-labeled melittin (40 μM melittin) upon addition of unlabeled HAA (0 to 40 μM).

To identify melittin backbone amide protons that potentially form contacts with HAA, Saturation Transfer Difference (STD)-15N-HSQC NMR was performed with [U-15N]-labeled melittin as the ligand and HAA as the protein target. The on resonance frequency of the selective saturation Gauss pulse (number of pulses, n =40) was set to 0.2 ppm. The off resonance frequency was 28.0 ppm. The length of each selective pulse was 50 ms, giving 2s total saturation time. To build a binding isotherm, the STD amplification factor was determined from samples with the
following mole ratios of melittin-to-HAA: 40:1, 91:1, 169:1, and 219:1. The STD amplification factor, $A_{STD}$ was calculated using Equation 2:

$$A_{STD} = [(I_0-I_{\text{sat}})/I_0]*(L_T/P_T) \quad (2)$$

In Equation 2, $I_0$ is the peak amplitude of one unsaturated proton, $I_{\text{sat}}$ is the peak amplitude of one proton after saturation. $L_T$ and $P_T$ correspond to the total number of moles of the ligand and protein, respectively.

$I_0$-$I_{\text{sat}}$ is the peak amplitude of a proton in the difference (STD) spectrum, while $I_0$ is the peak amplitude of a proton in the reference $^1$H-$^{15}$N HSQC spectrum. Difference spectra were recorded with 256 scans while reference spectra were recorded with 32 scans. The apparent dissociation constant ($K_d$) was determined by fitting data into Equation 3:

$$A_{STD} = \alpha_{STD}[L]/(K_d+[L]) \quad (3)$$

$\alpha_{STD}$ represents the maximum STD amplification of the monitored signal. Hyperbolic curve fitting on Graphpad Prism yielded the parameters $\alpha_{STD}$ and $K_d$.

**Docking Studies**

HADDOCK (high ambiguity driven docking) version 2.2 \cite{148, 149} was used to model the melittin-HAA complexes. Melittin coordinates were taken from the representative member of the NMR ensemble in PDB entry 6DST, corresponding to the solution structure of recombinant melittin in a trifluoroethanol/glycerol/water solvent system. HAA coordinates were taken from the homology model of the homo-2-mer generated by SwissModel.\cite{53-55, 166} The active and passive residues on melittin were selected (Table 2) based on STD-NMR criteria and relative solvent accessibility:
Table 2. Active and Passive Residues of Melittin and HAA, Round 1

<table>
<thead>
<tr>
<th>Melittin</th>
<th>HAA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Active</strong>&lt;sup&gt;a&lt;/sup&gt;</td>
<td><strong>Passive</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Ile&lt;sub&gt;17&lt;/sub&gt;, Ile&lt;sub&gt;20&lt;/sub&gt;</td>
<td>Leu&lt;sub&gt;6&lt;/sub&gt;, Leu&lt;sub&gt;9&lt;/sub&gt;, Leu&lt;sub&gt;13&lt;/sub&gt;, Ala&lt;sub&gt;15&lt;/sub&gt;, Leu&lt;sub&gt;16&lt;/sub&gt;</td>
</tr>
</tbody>
</table>

a) These residues on melittin have backbone amide protons showing >60% relative A<sub>STD</sub> measured in the <sup>1</sup>H<sup>-15</sup>N STD-HSQC experiments, and are solvent accessible (>10% relative solvent accessibility)

b) Passive residues have backbone amide protons with ~45-60% relative A<sub>STD</sub> and >10% relative solvent accessibility.
c) Active residues belong to the MAC sequence (residues 71-88 on HAA), and with >10% relative solvent accessibility

d) Passive residues are surface residues within a radius of 6.5 Angstroms around the active residues

The HAA binding interface from the best cluster of docking solutions for round 1 was assumed to contain the largest amount of intermolecular (hydrophobic) contacts. This type of contact is identified by HADDOCK when two carbon atoms are at most 3.9 Angstrom apart. The binding interface identified in round 1 was used to generate ambiguous interaction restraints (AIRs) in round 2 (Table 3):

**Table 3.** Active and Passive Residues of Melittin and HAA, Round 2

<table>
<thead>
<tr>
<th>Melittin</th>
<th>HAA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active</td>
<td>Passive</td>
</tr>
<tr>
<td>Ile17,</td>
<td>Leu6,</td>
</tr>
<tr>
<td>Ile20</td>
<td>Leu9,</td>
</tr>
<tr>
<td></td>
<td>Leu13,</td>
</tr>
<tr>
<td></td>
<td>Ala15,</td>
</tr>
<tr>
<td></td>
<td>Leu16</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
a) These residues on melittin have backbone amide protons showing >60% relative $A_{STD}$ measured in the $^1$H-$^{15}$N STD-HSQC experiments, and are solvent accessible (>10% relative solvent accessibility)

b) Passive residues have backbone amide protons with ~45-60% relative $A_{STD}$ and >10% relative solvent accessibility.

c) Active residues belong to the MAC sequence (residues 71-88 on HAA), and form the largest number of non-bonded contacts with melittin in Round 1 of the docking procedure.

d) Passive residues are surface residues within a radius of 6.5 Angstroms around the active residues.

In addition to AIRs, unambiguous restraints were applied based on the $A_{STD}$ values of melittin residues. **Table 4** shows the upper limits defined by residues which showed high $A_{STD}$ values.

**Table 4. Unambiguous Restraints (Upper Limits) Involving Melittin Residues with High $A_{STD}$**

<table>
<thead>
<tr>
<th>Melittin Residue, Atom Name&lt;sup&gt;a&lt;/sup&gt;</th>
<th>HAA Residue&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Distance, Lower Bound Correction, Upper Bound Correction</th>
</tr>
</thead>
<tbody>
<tr>
<td>17, $H^N$</td>
<td>78, 81, 82, 83</td>
<td>5.0, 5.0, 0.0</td>
</tr>
<tr>
<td>20, $H^N$</td>
<td>78, 81, 82, 83</td>
<td>5.0, 5.0, 1.4</td>
</tr>
</tbody>
</table>

a) Backbone amide protons ($H^N$) of melittin residues with >60% relative STD amplification are placed within 5 Å of ‘active residues’ on chain A of HAA.
b) Active residues on HAA belong to the MAC sequence (residues 71-88 on HAA), and have the largest number of non-bonded contacts in Round 1 of docking

**Table 5** shows the lower limits defined by residues that showed negligible $A_{STD}$.

**Table 5.** Unambiguous Restraints (Lower Limits) Involving Melittin Residues with Negligible $A_{STD}$

<table>
<thead>
<tr>
<th>Melittin Residue, Atom Name $^a$</th>
<th>HAA residue $^b$</th>
<th>Distance, lower bound correction, upper bound correction</th>
</tr>
</thead>
<tbody>
<tr>
<td>10, $H^N$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11, $H^N$</td>
<td>63-141, 263-341</td>
<td>6.0, 1.0, 72.0</td>
</tr>
<tr>
<td>18, $H^N$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19, HE1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23, $H^N$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a) Protons ($H^N$ - backbone amide, HE1- tryptophan indole side chain) of melittin residues with <20% relative STD amplification are placed 5-78 Å from all residues on the model of the HAA dimer. 78 Å is the longest distance between 2 atoms on the HAA dimer.

b) HAA residues on chain A of the dimer are numbered 63-141, while those on chain B are numbered 263-341.
For the last round of docking, the homology model of the 32-subunit homo-oligomer of HAA generated by SwissModel\textsuperscript{53-55,166} was used, and restraints from round 2 were applied.

Flexible termini of each subunit in the homology model were removed, leaving only residues 63 – 141 per subunit. Figures of the docked melittin-HAA complex were made using Chimera.\textsuperscript{167} Alignment of structures were accomplished using Chimera MatchMaker.\textsuperscript{167} Relative solvent accessibilities were calculated using Swiss PDB viewer.\textsuperscript{128} Electrostatic potential maps were generated by DelPhi webserver (http://compbio.clemson.edu/sapp/delphi_webserver).

7.3 Results

*Human αA-Crystallin Binds to Melittin in an Endothermic, Entropy-Driven Process*

Isothermal titration calorimetry was used to determine thermodynamic parameters for the binding interaction between melittin and HAA. ITC measures stepwise changes in the heat of reaction during the course of a titration experiment, and does not require any modification or labeling of the ligand/receptor under analysis.\textsuperscript{168} We characterized the binding enthalpy (ΔH), binding entropy (ΔS) and the dissociation constant (K\textsubscript{d}) at 298 K. **Figure 22** shows the thermogram and binding isotherm associated with isothermal titration of HAA (titrant or “ligand”) into melittin (titrand or “receptor”).
Figure 22. Calorimetric titration profile for the binding of HAA to melittin at 298K. A) Heat associated with injections of HAA (0.603-0.680 mM) into melittin (0.071mM), after subtracting heat of dilution. B) Binding isotherm corresponding to the data in Panel A. The first injection was omitted from data analysis. Curve fitting was done using the independent binding sites model on NanoAnalyze (TA Instruments).

The thermodynamic parameters obtained from curve fitting of the data in Figure 22 are shown in Table 6.
Table 6. Thermodynamic Parameters of the Binding Interaction Between Melittin and HAA

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_a$ (M$^{-1}$) $\times 10^{-4}$</td>
<td>$4.5 \pm 1$</td>
</tr>
<tr>
<td>$n$ (HAA monomer : melittin)</td>
<td>$1.5 \pm 0.01$</td>
</tr>
<tr>
<td>$\Delta H$ (kJ/mol)</td>
<td>$19 \pm 3$</td>
</tr>
<tr>
<td>$K_d$ (M) $\times 10^5$</td>
<td>$2.3 \pm 0.7$</td>
</tr>
<tr>
<td>$\Delta S$ (J/K mol)</td>
<td>$150 \pm 8$</td>
</tr>
</tbody>
</table>

The Wiseman parameter$^\text{169}$ or $c$ value for the data set is ~4.8. This parameter is calculated as the product of the binding constant ($K_a$), number of binding sites ($n$), and receptor (melittin) concentration. Conventionally, $c$ values outside the range $10 \leq c \leq 500$ indicate sub-optimal curve fitting,$^\text{170}$ therefore, the thermodynamic parameters extracted from the isotherm in Figure 22 should be treated as ‘apparent’ values. Nevertheless, the signs of $\Delta H$ and $\Delta S$ remain useful in providing a qualitative assessment of the thermodynamic signature of the binding event. The apparent binding stoichiometry was 1.5:1 or 3:2 (HAA monomer:melittin), which is different from the previously reported 1:1 binding stoichiometry ($\alpha$-crystallin monomer:melittin) by Farahbakhsh et al,$^\text{77}$ although it must be noted that the previous report titrated melittin with bovine $\alpha$-crystallin (mixture of $\alpha$A- and $\alpha$B- isoforms). The said study obtained a $K_d$ of 7.3 $\mu$M. Our apparent $K_d$ is higher (~20 $\mu$M). Our ITC results (Table 6) show that the change in enthalpy ($\Delta H$) of the binding process is positive, indicating an endothermic reaction, and the change in entropy ($\Delta S$) is also positive. The fact that both $\Delta H$ and $\Delta S$ are positive indicates that the binding interaction is entropy-driven. Hydrophobic interactions associated with a loss of water in hydrophobic pockets on a receptor upon ligand binding have generally been considered entropically favorable.$^\text{171}$
Melittin Residues with Long Hydrophobic Sidechains are Prominently Involved in the Binding Interaction with Alpha-A Crystallin

Spectral changes in the $^1$H-$^{15}$N correlation spectrum can be used to monitor the binding surface on an $^{15}$N-labeled ‘receptor’ upon titration with an unlabeled ‘ligand’. We titrated $^{15}$N-labeled recombinant melittin (receptor) with unlabeled HAA (ligand) (Figure 23), using mole ratios of 1:1 to 1:4 ($\alpha$A monomer: melittin) to identify residues on melittin that potentially participate in the binding surface. We observed a general decreasing trend for the peak amplitude of backbone amide protons within the mole ratio range of 1:1-1:4. Signal broadening in the melittin spectra may be due to a number of reasons, including intermediate exchange between free and HAA-bound states, structural rearrangement of the melittin/HAA complex after binding, or oligomerization before or after binding. Notably, the intermediate exchange regime in NMR is characterized by a dissociation constant, $K_d$, within the range ~10-100µM. The observed broadening of melittin signals may therefore corroborate the data from ITC (Figure 22, Table 6), which gives an apparent $K_d$ within this range. At high $\alpha$A monomer:melittin mole ratios (above 2:1), we observed precipitate formation, and most amide proton peaks eventually disappeared at a mole ratio of 4:1.
Figure 23. Amide proton peaks on melittin show differential peak broadening upon addition of HAA. $^{15}$N-HSQC spectra of $[^{15}$N]-melittin in NMR buffer were acquired for various mole ratios of αA-crystallin monomer to melittin (1:1 to 1:4). The percent decrease in peak amplitude (% broadening) of backbone amide cross-peaks was plotted against melittin residue number. Gly1 and Pro14 are omitted from the interaction map because the Gly1 H$_N$ peak was not visible in the $^{15}$N-HSQC spectra due to chemical exchange of the terminal amino group. Pro14 lacks an amide group, therefore it did not show a peak in the spectra. Data for overlapping peaks (A4/K21, and Q25/Q26) were omitted. The $^{15}$N-HSQC spectrum of $[^{15}$N]-melittin without αA-crystallin was used as control. Residues Ile17, and Ile20 showed the most prominent decrease in peak intensity (>60%, as indicated by the dashed line) upon addition of HAA. The amide proton peaks for Gly3 and Lys23 were broadened and sharpened in different stages of the titration, suggesting complex association and dissociation phenomena.$^{157}$

The melittin residues that showed the highest signal broadening upon addition of HAA were Ile17, and Ile20, with above 60% decrease in peak amplitude (Figure 23). Next to these are Leu13, Ala15, Leu16, Ser18, and Trp19 which showed at least 50% decrease in peak amplitude. Most of these residues have hydrophobic side chains, which clearly suggests that hydrophobic interaction between αA-crystallin residues and melittin residues is involved in the binding event, consistent with our ITC findings.

The interaction between HAA and melittin was further investigated by STD-NMR spectroscopy. STD-NMR enables identification of proton resonances of a ligand (melittin) in close contact with a target protein (HAA).$^{164,172}$
**Figure 24. Residue-specific contacts of melittin with HAA mapped by STD-NMR.** A) Overlaid $^1$H-$^{15}$N STD-HSQC (blue) and reference $^1$H-$^{15}$N HSQC (red) spectra of $[U^{15}$N$]$-melittin with HAA (1:219 HAA:melittin monomer mole ratio). Amide backbone protons showing significant STD amplification (~40-100% relative A$_{STD}$) are labeled. B) Relative A$_{STD}$ of amide protons on the melittin backbone are plotted against residue number. The dotted line corresponds to a 45% cut-off. Residues 4, 21, 25, and 26 are omitted from the analysis because of overlapping peaks. Gly1 H$^N$ is not visible in all spectra due to exchange with the solvent and Pro14 does not have an amide proton.

In STD-NMR, saturation is transferred from the receptor (HAA) to the ligand (melittin) protons when the ligand binds to the receptor. The melittin protons nearest to the HAA binding surface are most likely saturated to the highest extent, and therefore would show the strongest STD
amplification, whereas those protons located farther away are saturated to a lower extent and their STD signals are weaker. **Figure 24** shows that amide protons from hydrophobic residues Ile17 and Ile20 exhibited the largest STD amplification (relative $A_{\text{STD}} > 60\%$). Relative $A_{\text{STD}}$ values in the range $\sim 45$-60% correspond to $H^N$ of other hydrophobic residues: Leu6, Leu9, Leu13, Ala15, Leu16. We monitored the increase in STD signal as ligand (melittin) concentration was increased, until a 1:219 monomer mole ratio of HAA:melittin was reached (**Figure 25**).

**Figure 25.** Titration plots of melittin to HAA (1 µM) monitoring the increase of the STD amplification ($A_{\text{STD}}$) of backbone amide protons versus the melittin concentration ($t_{\text{sat}} = 2s$). The STD amplification factor is calculated using Equation 2. A) Backbone amide protons on hydrophobic residues of melittin show high $A_{\text{STD}}$ (Ile 17 and Ile20) and moderate $A_{\text{STD}}$ (Leu 6, Leu9, Leu13, Ala15, Leu16). B) Backbone amide protons of other melittin residues showed relatively weaker STD effects C) Fitting of titration data for the backbone amide proton of S18 to a one-site specific-binding isotherm results in an apparent dissociation constant of $\sim 147\mu$M. D) Side chain $H^N$ protons from glutamine residues (Qa and Qb, belonging to either Q25 or Q26) and the HE1 proton from the indole group of W19 show weak STD effects.
We classified all $A_{\text{STD}}$ signals into four categories, based on the relative $A_{\text{STD}}$ signal at the 1:219 monomer mole ratio of HAA:melittin: high ($\geq 60\%$), moderate (45-59\%), low (16-44\%), and negligible (0-15\%). We did not further increase the amount of melittin to avoid the possible tetramerization of the peptide, a process which is favored at high peptide concentrations.\textsuperscript{81} Formation of the melittin tetramer may complicate the analysis of STD-NMR data. Previously, we reported on the structural characterization of recombinant, monomeric melittin by NMR.\textsuperscript{90} The control $^1$H-$^{15}$N HSQC spectra collected for the sample with 1:219 HAA:melittin mole ratio in this report matches that of the monomeric melittin, and did not show new peaks that may be attributed to tetramer formation, therefore we conclude that the observed STD signals correspond to monomeric melittin. The apparent dissociation constant obtained from the relationship between $A_{\text{STD}}$ of the amide proton of S18 and melittin concentration was $\sim 147 \mu M$ (Figure 25C), which is higher than that calculated from ITC data.

We applied the residue-specific data from STD-NMR to the HADDOCK docking protocol to generate crude models of the HAA/melittin complex in three rounds.\textsuperscript{149, 173} The goal of the first round of docking was to identify the interaction surface on HAA that included the MAC sequence, the putative substrate-binding region on bovine $\alpha$A-crystallin.\textsuperscript{7} Active residues on HAA were defined to be part of the MAC segment, with at least 10\% relative solvent accessibility. For simplicity, we assumed a 1:1 binding stoichiometry (monomeric HAA:melittin), although our ITC data also suggest a 3:2 stoichiometry. The first docking procedure generated numerous possible docking solutions - eleven clusters were identified, corresponding to 71.5\% of all water-refined models generated by HADDOCK. The most reliable cluster of docking solutions (cluster 2) had the lowest HADDOCK score ($-74.9 \pm 4.3$ a.u.) and Z-score (-2.0). This marks the second most populated cluster (34 models), next to cluster 1 (45 models, constituting the majority). The best
four structures of the most reliable cluster (cluster 2) have average AIR energies of 108.9 ± 40.23 kcal·mol$^{-1}$ and buried surface areas of 1157.4 ± 40.9 Å$^2$. We observed that residues 78, 79, 81, 82, and 83 on HAA formed the highest number of non-bonded contacts, therefore these residues form the interaction surface on HAA.

In the second round of docking, we set these HAA residues (78, 79, 81, 82, 83) as active, while passive residues were those within 6.5 Å of the active residues. We also retained the designation of active and passive residues on melittin. In addition to ambiguous restraints, we provided unambiguous restraints in the form of upper and lower distance limits, which reflect our STD-NMR findings (Tables 4-5). Melittin backbone amide protons with the highest A$_{STD}$ (Ile17, Ile20) were used to generate upper limits. Ile17 H$^N$, which showed the highest A$_{STD}$, was constrained to be at most 5 Å away from active residues on HAA. On the other hand, Ile20 H$^N$, which generated A$_{STD}$ signals about 72% of the corresponding Ile17 H$^N$ A$_{STD}$ signal, was constrained to be at most 6.4 Å away from the active residues on HAA, so that the difference between the upper limit of Ile17 H$^N$ and Ile20 H$^N$ was 28% (i.e. the difference between 100% and 72%). Lower limits for the distance between melittin and HAA were generated as follows: protons with the lowest A$_{STD}$ signals (backbone amide protons of Thr10, Thr11, Ser18, and Lys23, as well as the HE1 proton of Trp19) were set to be at least 5 Å away from any residue on HAA.

The second round of docking provided a well-converged result, with only one cluster containing 100% of the water-refined models generated by HADDOCK. The HADDOCK score for this cluster was -61.3 ± 2.1 a.u., and the Z-score was 0.0. The ensemble of 10 models that best represent the HAA/melittin complex is shown in Figure 26. Melittin adopts a mostly helical structure in its bound form, and the longitudinal axis of the C-terminal helix of melittin is roughly aligned with loop region of the MAC sequence on HAA. The interaction surface on melittin
involves mostly hydrophobic residues, as identified by our STD-NMR experiments: Ile17 and Ile20 had high $A_{\text{STD}}$ signals for their amide protons and therefore formed the majority of non-bonded contacts with HAA. Leu16 and Leu13 had moderate $A_{\text{STD}}$ values, and thus formed fewer contacts. Notably, the model suggests that Lys21 is part of the interaction surface, although $A_{\text{STD}}$ of Lys21 could not be reliably quantified because the amide peak of Lys21 overlapped with that of Ala4 in the $^1\text{H}-^1\text{N}$ HSQC and STD-$^1\text{H}-^1\text{N}$ HSQC spectra.
Figure 26. Structural model of the melittin-HAA complex from docking round 2. Melittin residues are labeled with superscript “m” while HAA residues are labeled with superscript “a”. The molecular models are generated in Chimera.\textsuperscript{167} A) Ensemble of 10 best models from Round 2 of docking. The melittin chain is colored green. Thick ribbons on melittin indicate helical regions. Chain A of the HAA dimer is colored red, except for the MAC sequence (residues 71-88) which is shown in cyan. Chain B is not shown. Thick ribbons on HAA indicate β-strands. B) Residues belonging to the interaction surface on melittin. Beige surfaces indicate ‘active’ residues with high amide proton A\textsubscript{STD} signals (Ile17 and Ile20). Green surfaces indicate ‘passive’ residues with moderate amide proton A\textsubscript{STD} signals (Leu13 and Leu16). A\textsubscript{STD} of Lys21 (sky blue surface) is not known because the amide peak of Lys21 on the STD \textsuperscript{1}H-\textsuperscript{15}N-HSQC spectra overlaps with that of A4. C) Residues belonging to the putative interaction surface on HAA. D) Orientation of the C-terminal segment of melittin with respect to the interaction surface on HAA. Side chains of residues on the interaction surface of melittin are shown in stick models (Ile17, Ile20, Leu13, Leu16, Lys21). The residues labeled with an asterisk (*) showed moderate A\textsubscript{STD} signals, however they are not bound to the ACD of HAA. E) Electrostatic contacts in the HAA-melittin complex. Lys21 and Lys23 of melittin form contacts with Glu102 and Asp136 of HAA, respectively.

The interaction surface on HAA (Figure 26C) consists of His79, Phe80, Ser81, Pro82, Glu83, Glu102, and Asp136. Among these, residues 79-83 are part of the MAC sequence. Residues with moderate A\textsubscript{STD} that are not part of the binding interface are Ala15, Leu9, and Leu6 (Figure 26D). In addition to hydrophobic contacts, electrostatic contacts were also observed involving the charged side chains of the following residue pairs (melittin-HAA): Lys21-Glu102, Lys23-Asp136 (Figure 26E).

The final round of docking involved binding of one melittin molecule to ten chains (chains A,B,C,D,M,N,O,P, 0, and 1) of the HAA oligomeric model generated by SwissModel (Figure 27).\textsuperscript{53-55, 166} The purpose of this docking approach was to visualize the possible orientation of melittin in the oligomeric form of HAA. The melittin binding site was constrained only on chain A of the oligomer. Previous studies on the recombinant HAA have suggested that the oligomeric assembly has a variable size, and may consist of 24 subunits (similar to the apparent size of HAB), although smaller and larger clusters are possible.\textsuperscript{774} Taking into account the SwissModel database for evolutionarily related proteins, SwissModel generated a homology model of HAA with 32
subunits (https://swissmodel.expasy.org, generated using the SWISS-MODEL template library version 2019-04-11 and Protein Data Bank release 2019-04-05), and therefore this model was used in the docking procedure. The best cluster (Cluster 1) obtained by docking had a HADDOCK score of -85.3 ± 3.5 a.u. and a Z-score of -1.0. The 10 best structures on cluster 1 are shown in Figure 27A. The best model (Figure 27B) shows that melittin may occupy the interior of the oligomer, and may have more than one binding interface (Figure 27C-E).
Figure 7. Possible melittin location in the homology model of the HAA 32-subunit oligomer. A) Ensemble of 10 docking solutions representing melittin bound to the interior of the HAA oligomer. HAA chains are colored red, except for the MAC sequence (residues 71-88 in each subunit), which are rendered in cyan. N- and C-terminal regions of melittin are indicated by ‘N’ and ‘C’, respectively. B) Best model of the HAA-melittin complex in the oligomer context. The helix axis of melittin is roughly aligned with the loop on the MAC sequence (colored cyan). C) Residues on the binding surface of HAA chain A that form nonbonded contacts with melittin. D) Residues on melittin that form non-bonded contacts with HAA chain A. Ile17 and Ile20 (colored beige) were designated as active residues on melittin, and showed high $A_{STD}$. Leu13 and Leu16 (colored green) were designated as passive, and showed moderate $A_{STD}$. Lys21 and Arg24 (colored sky blue) also form non-bonded contacts. Lys21 $A_{STD}$ was not quantified due to spectral overlap in the STD-NMR experiments. Arg24 had low $A_{STD}$. E) The N-terminal side of melittin (Leu6 and Lys7) forms contacts with His100 and Glu83 in HAA chain C. Leu6 was regarded as a passive residue, with moderate $A_{STD}$. Lys7 had low $A_{STD}$.

The model suggests that in the oligomeric form of HAA, one melittin molecule may form contacts with two subunits at the same time. The main binding interface between melittin and HAA is similar to that determined in round 2 of the docking protocol (Figure 26C-D), and involves the C-terminal region of melittin and the loop region on the MAC sequence of HAA chain A. A secondary binding interface may form between the N-terminal residues of melittin (Leu6 and Lys7) with a different HAA subunit (Figure 27E), involving Glu83 and His100. Notably, Glu83 on chain C of the HAA homo-32-mer belongs to the MAC sequence as well.

7.4 Discussion

Hydrophobic interactions are implicated in a number of molecular chaperone-substrate interactions. The molecular chaperone Hsp70 is believed to bind to client proteins through a mechanism involving the anchoring of a hydrophobic residue. The molecular chaperone Spy, a periplasmic protein of E. coli, reportedly binds to its client Im7 through a combination of long-range electrostatic interactions and short-range hydrophobic interactions. Through the use of
NMR techniques, it was discovered that the apical domain of bacterial chaperonin GroEL binds to the hydrophobic side of a helical peptide ligand\(^{176}\) and that the binding of unfolded alkaline phosphatase (PhoA) to bacterial trigger factor (TF) involves hydrophobic contacts between the substrate and the molecular chaperone.\(^{177}\) In this report, we aimed to characterize the forces that drive the chaperone-substrate interaction between HAA and melittin. Using ITC, we probed the thermodynamic signatures of the binding event, and through NMR, we determined the residues on melittin that are putatively involved in the binding surface.

Our ITC findings suggest that hydrophobic associations are involved in the chaperone-substrate interaction, as indicated by the positive values of $\Delta H$ and $\Delta S$ (Figure 22, Table 6). The association of hydrophobic groups is classically recognized as entropy-driven.\(^{171}\) This conventional view of hydrophobic association implicates release of ‘more ordered’ water molecules near the hydrophobic surface of the ligand to the bulk water, upon binding of the ligand to a cavity in the receptor, resulting in net entropic gain.\(^{179,181}\) In the HADDOCK-generated model of the melittin-HAA complex (round 2 of docking, Figure 26), a mostly hydrophobic, neutral surface on melittin is in close contact with a hydrophobic surface on HAA. Thus, the model reveals possible hydrophobic associations that provide an entropic contribution. Notably, in addition to hydrophobic contacts, there are electrostatic contacts observed in the complex, involving the positively charged side chains on melittin and negatively charged side chains on the surface of HAA (Figure 26E). In a previous study by Kumar et al\(^{182}\) hydrophobic sites on recombinant $\alpha$A-crystallin and $\alpha$B-crystallin were quantified by ITC using ANS as a hydrophobic probe. These authors found that two types of binding sites were present on both $\alpha$A and $\alpha$B, one with higher affinity ($K_{d} \sim 1-4 \ \mu M$, in the temperature range 15-30 °C), and another with lower affinity ($K_{d} \sim 5-13 \ \mu M$, in the temperature range 15-30 °C). Binding to both types of sites was characterized by a
negative change in enthalpy, and positive change in entropy. In contrast, our findings on melittin-HAA binding point to a positive change in both enthalpy and entropy, with lower affinity (Kd~20 µM at 25 °C). This suggests that the thermodynamic signature of chaperone-substrate binding by HAA is substrate-dependent.

Literature evidence suggests the existence of models in which the hydrophobic interaction in cavity-ligand binding is enthalpy-driven,\textsuperscript{183} interpreted as binding associated with release of ‘less ordered’ water molecules to the bulk. Kinoshita and Hayashi\textsuperscript{171} have argued that the structural difference between the water near the binding surfaces and within the space between the binding surfaces determines whether the binding is entropically favorable or unfavorable. In contrast, Baron et al.\textsuperscript{184} suggested that cavity-ligand binding is driven by entropy when either the cavity or the ligand binding surface carries a charge (positive or negative), whereas it is entropically unfavorable when they both carry unlike charges. Perhaps, the architecture of the binding surface of the substrate dictates whether the binding would be dominated by enthalpic or entropic gain.

Residues 79, 81, 82, 83 are common to the binding surfaces on HAA determined in docking round 2 and round 3 (\textbf{Figures 26 and 27}). These residues are part of the loop region of the MAC sequence in the ACD. Notably, in the first round of docking, residues belonging to the two beta-strand regions and the loop on MAC were considered as active residues, however the most reliable docking solutions show contacts mainly involving the loop region. Therefore, the docking procedure we used may be useful in identifying which parts of the MAC sequence on the ACD of HAA preferentially form contacts with melittin, in the absence of structural information for domains in HAA other than the ACD. Interestingly, residues 79, 81, 82, and 83 on the ACD also belong to a region on αA-crystallin that was found to bind to the hydrophobic dye with 1,1’-bis(4-anilino)naphthalene-5,5’-disulfonic acid or bis-ANS (residues 79-99).\textsuperscript{158} A previous study on
spin-labeled full-length rat αA-crystallin reported evidence that residues 75-80 are buried in the oligomer. Thus, as others have hypothesized, it is likely that the substrate-binding surface on the ACD is sequestered in the free form of αA-crystallin (possibly by the N- and C-terminal domains), but binding to substrates exposes this region. We note that the docking procedure employed in modeling the HAA-melittin complex does not take into account the contributions of N- and C-terminal regions on the binding event, and the solvent accessibility of residues in the HAA models used for docking do not consider possible sequestration due to subunit interactions.

Another melittin-binding site mapped by Sharma et al on αA-crystallin involved the N-terminal arm (12RTLGPFYPSR21) however, due to lack of structural coordinates for the N-terminal domain of αA-crystallin, this report does not model the melittin-N-terminal domain interaction. We hypothesize that while bound to the ACD, melittin may also bind to the NTD, following our observation that melittin in the proposed models (Figure 26 and 27) has other surfaces that include residues showing moderate ASTD, but were not part of the binding interface on the ACD in the HADDOCK-generated models. We previously determined that melittin is unstructured in its free form, in the buffer system used for this report, but interaction with molecules such as trifluoroethanol, which contains hydrophobic surfaces, can induce helix formation. To keep melittin in the mostly helical state modeled in Figures 26-27, it is possible that the N-terminal segment of melittin associates with hydrophobic groups that are not part of the ACD. Notably, the melittin-binding sequence (12RTLGPFYPSR21) identified by Sharma et al includes hydrophobic residues. Moreover, the N-terminal domain of αA-crystallin contains the sequence SRLFDQFFG which has an abundance of hydrophobic side chains. This sequence is also conserved in the sHSP family and in α-crystallins, and deletion of this sequence has been correlated with decrease in oligomeric size and increased chaperone activity of αA-crystallin and αB-crystallin.
The role of the C-terminal extension in melittin-HAA binding was also not considered in our docking approach, due to unavailability of the model of the full-length C-terminal segment (last 34 residues on αA-crystallin). The C-terminal region of αA-crystallin is implicated in subunit interactions and chaperone activity (reviewed previously) and is proposed to promote the solubility of αA-crystallin. In the crystal structure of truncated bovine αA-crystallin, part of the C-terminal tail is bound to the groove formed by β4 and β8 on the ACD (β4 contains LTVK from the MAC sequence). Based on our models of HAA-melittin, the LTVK segment on HAA is spared from melittin binding (Figures 26-27), in spite of being identified as a melittin-binding sequence by mass spectrometric analysis of cross-linked melittin and αA-crystallin.

In docking round 3, we proposed a model of the melittin molecule bound to the HAA oligomer (Figure 27). Considering that there are multiple sites in the oligomer containing the MAC sequence, and that some of these sites (colored cyan in Figure 27A) are on the surface of the oligomeric model, it must be noted that melittin may not only be found in the cavity of the oligomer, but it may also be protruding onto the surface of HAA. Therefore, when melittin binds to oligomeric HAA, it may assume a wide variety of conformations, and our proposed models might represent a few of these conformations. The line broadening of melittin amide protons that we observed in our 1H-15N HSQC titrations of melittin with HAA (Figure 23) may reflect the intermediate exchange not only between free and bound forms of melittin, but also interconversion among various melittin conformations within the HAA oligomeric assembly.

The results of this investigation mark a step forward in the understanding of the function of αA-crystallin, specifically in binding to its clients. The significant role of hydrophobic interactions in HAA-client binding is suggested by the observed thermodynamic signatures of HAA-melittin binding. HAA, however, is able to discriminate between various solvent-exposed hydrophobic side
chains on an unfolded client (represented by melittin, which is unstructured in its unbound state), as indicated by the different extents of saturation transfer observed for hydrophobic residues on melittin (Figure 24). For decades, the prevailing paradigm regarding client binding by sHsps and the related α-crystallins was that these molecular chaperones recognize the partially unfolded clients by exposed hydrophobic patches. Our findings in this report suggest that, although hydrophobic residues may mediate client binding to HAA, there are factors other than hydrophobicity that come into play in client recognition by HAA.
Summary and Conclusions

The main goals of the present work were to provide structural and thermodynamic details underlying the molecular chaperone activity of α-crystallin, with an emphasis on human alphaA-crystallin. Given the number of difficulties encountered in the structure elucidation of HAA, detailed in Chapter 1, (Section 1.2), this task proved to be difficult. We therefore devised an alternative strategy involving the use of a model substrate, melittin, which had been previously been used by others as a client mimic, to probe HAA. With the aid of this model substrate, we collected significant data on the putative substrate-binding interface of HAA despite the lack of a high-resolution structure of the chaperone. In Chapters 3-4A, we characterized the secondary structure and backbone dynamics of melittin, after expressing a uniformly isotopically-labeled form of the model substrate in solution. We found that, in agreement with previous literature, melittin can be unstructured or helical. However, our work additionally provided residue-level information on the cis-Pro to trans-Pro structural transition, and, for the first time, revealed the structural distinction between the two isomeric forms of melittin under near-physiological solution conditions.

In Chapter 4B, we described the determination of a high-resolution structure of the model substrate in solution (showing the helical, bound state of melittin), which facilitated our docking studies on the melittin-HAA complex. Arguably, the structure determined in this work will be a distinct advantage in revealing the conformational transitions of melittin as it binds to lipids, proteins, and polysaccharides, and will be useful in other binding studies involving this versatile peptide.
We also investigated the interaction between the model substrate and a 19-residue peptide fragment of HAA (called MAC) which shows molecular chaperone properties comparable to those of native HAA. These binding studies modeled the substrate-binding region of HAA (Chapter 5). We found that melittin binds to the MAC peptide with low affinity and positive cooperativity, although our NMR studies on the melittin-MAC interaction could not provide information on the MAC residues that specifically formed contacts with melittin. This was primarily because this interaction took place in the intermediate exchange regime of NMR, which precluded the direct determination of the structure of MAC-bound melittin. Therefore, guided by our NMR titration data and fluorescence spectroscopy measurements we took a computational docking approach to model the melittin-MAC interaction.

In Chapter 6, we presented an approach to studying HAA structure using the cataract-associated G98R mutant of HAA, which formed monomers under certain solution conditions. Again however, the monomeric form of the mutant protein did not produce high-quality NMR spectra under the conditions tested, presumably because of chemical exchange undergone by backbone amide protons of the protein in the intermediate exchange regime of NMR. Further experiments that explore other isotope-labeling strategies may need to be performed to circumvent this problem.

Finally, in Chapter 7, we presented a detailed investigation of the melittin-HAA binding interaction, which included determination of thermodynamic parameters of the process, and provided a model of the melittin-HAA complex. We found that several hydrophobic residues on melittin were implicated in the melittin-HAA binding interaction, and that the interaction was entropy-driven and of low affinity. Our findings suggested that hydrophobic associations may mediate the chaperone-substrate interaction. The work also provides evidence that HAA may be
able to discriminate between various solvent-exposed hydrophobic groups of the model substrate, and thus suggests a complex mechanism of substrate recognition by HAA.
List of Publications and Presentations

Publications:


Manuscript in Preparation:


Presentations:

(1) Oral presentation


(2) Poster presentation


(3) Poster presentation

(4) Oral presentation


(5) Oral presentation

Ramirez L., Shekhtman A., Pande J. Modeling a Chaperone-Client Interaction: Melittin Binding to Alpha Crystallin, presented at the 8th Annual Life Sciences Research Symposium (2016) at the University at Albany, SUNY (November 4, 2016)

(6) Oral presentation

Ramirez L., Shekhtman A., Pande J. Modeling a Chaperone-Client Interaction: Melittin Binding to Alpha Crystallin, presented at the 2016 American Chemical Society Northeast Regional Meeting in Binghamton, NY (October 7, 2016)
Bibliography


structural basis for the activation of alphaB-crystallin oligomers, Nat. Struct. Mol. Biol. 17 (9), 1037-1042. DOI: 10.1038/nsmb.1891


Appendix: Publications and Licenses are available as Supplemental Documents