Interactions driving the collapse of islet amyloid polypeptide: implications for amyloid aggregation

by

Stephanie M. Cope

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved November 2013 by the
Graduate Supervisory Committee:

Sara M. Vaiana, Chair
Giovanna Ghirlanda
Robert Ros
Stuart M. Lindsay
Sefika Ozkan

ARIZONA STATE UNIVERSITY

December 2013

ABSTRACT

Human islet amyloid polypeptide (hIAPP), also known as amylin, is a 37-residue intrinsically disordered hormone involved in glucose regulation and gastric emptying. The aggregation of hIAPP into amyloid fibrils is believed to play a causal role in type 2 diabetes. To date, not much is known about the monomeric state of hIAPP or how it undergoes an irreversible transformation from disordered peptide to insoluble aggregate.

IAPP contains a highly conserved disulfide bond that restricts hIAPP(1-8) into a short ring-like structure: N_loop. Removal or chemical reduction of N_loop not only prevents cell response upon binding to the CGRP receptor, but also alters the mass per length distribution of hIAPP fibers and the kinetics of fibril formation. The mechanism by which N_loop affects hIAPP aggregation is not yet understood, but is important for rationalizing kinetics and developing potential inhibitors. By measuring end-to-end contact formation rates, Vaiana et al. showed that N_loop induces collapsed states in IAPP monomers, implying attractive interactions between N_loop and other regions of the disordered polypeptide chain[1]. We show that in addition to being involved in intra-protein interactions, the N_loop is involved in inter-protein interactions, which lead to the formation of extremely long and stable β-turn fibers. These non-amyloid fibers are present in the 10 μM concentration range, under the same solution conditions in which hIAPP forms amyloid fibers. We discuss the effect of peptide cyclization on both intra-

[1] Vaiana, S.M., Best, R.B., Eaton, W.A., and Hofrichter, J. 2009. Evidence for a partially structured state of the amylin monomer. Biophys J. 97:2948-2957.

and inter-protein interactions, and its possible implications for aggregation. Our findings indicate a potential role of N_loop-N_loop interactions in hIAPP aggregation, which has not previously been explored.

Though our findings suggest that N_loop plays an important role in the pathway of amyloid formation, other naturally occurring IAPP variants that contain this structural feature are incapable of forming amyloids. For example, hIAPP readily forms amyloid fibrils in vitro, whereas the rat variant (rIAPP), differing by six amino acids, does not. In addition to being highly soluble, rIAPP is an effective inhibitor of hIAPP fibril formation[2]. Both of these properties have been attributed to rIAPP's three proline residues: A25P, S28P and S29P. Single proline mutants of hIAPP have also been shown to kinetically inhibit hIAPP fibril formation. Because of their intrinsic dihedral angle preferences, prolines are expected to affect conformational ensembles of intrinsically disordered proteins. The specific effect of proline substitutions on IAPP structure and dynamics has not yet been explored, as the detection of such properties is experimentally challenging due to the low molecular weight, fast reconfiguration times, and very low solubility of IAPP peptides. High-resolution techniques able to measure tertiary contact formations are needed to address this issue. We employ a nanosecond laser spectroscopy technique to measure end-to-end contact formation rates in IAPP mutants. We explore

[2] Cao, P., Meng, F., and Raleigh, D.P. 2010. The ability of rodent islet amyloid polypeptide to inhibit amyloid formation by human islet amyloid polypeptide has important implications for the mechanism of amyloid formation and the design of inhibitors. Biochemistry. 49:872–881.

the proline substitutions in IAPP and quantify their effects in terms of intrinsic chain stiffness. We find that the three proline mutations found in rIAPP increase chain stiffness. Interestingly, we also find that residue R18 plays an important role in rIAPP's unique chain stiffness and, together with the proline residues, is a determinant for its non-amyloidogenic properties. We discuss the implications of our findings on the role of prolines in IDPs.

ACKNOWLEDGEMENTS

> “There is no such thing as a 'self-made' man. We are made up of thousands of others. Everyone who has ever done a kind deed for us, or spoken one word of encouragement to us, has entered into the make-up of our character and of our thoughts, as well as our success."
>
> - George Burton Adams

Foremost, I would like to express my deepest thanks to my advisor, Dr. Sara Vaiana, for all the time and energy she has put into this work. Without all of our productive discussions, I would have never made it to this point. I am truly indebted to her for helping me to think more critically and for giving me such an exciting project to work on. Thank you for giving me the freedom to explore new problems and shaping me into the scientist that I am today.

I am also grateful to all of my labmates over the past five years, including Sara Sizemore, Danmei Bien, Alejandro Solis and Sebastian Haefner. Sara has been extremely helpful in catching my mistakes, and helping me think through problems. Her assistance with all the MatLab codes has been invaluable.

All of my collaborators have made this work possible. Firstly, I am extremely grateful to Dr. Giovanna Ghirlanda to opening up her lab to me since day one and supporting me all along the way. Thank you for entrusting your chemistry lab to a physicist. I am extremely grateful for having learned to synthesize and purify proteins, and I owe this knowledge to Giovanna’s students: Anindya Roy and Dr. Sandip Shinde. Thank you both for teaching

me every ounce of chemistry I know and for being walking organic chemistry textbooks for all of my questions.

The work presented in this thesis includes numerous experimental techniques, and I could have never mastered so many techniques without the expertise and guidance of so many scientists. Thank you to Dr. John Lopez and Dr. Zach Laughrey for their help in teaching me HPLC, MALDI-TOF and CD. Thank you to Dr. Brian Cherry for his patience in training me in NMR. Thank you to Dr. Marina Faiella for offering help in NMR data analysis. Thank you to Dr. Arnab Dutta and Dr. Anne Jones for training me in FTIR and allowing me to use their lab. Thank you to David Lowry for teaching me TEM sample preparation, and training me to use the SOLS' Phillips CM12. Thank you to my advisor, Alejandro Solis and Sebastian Haefner for helping me understand all the intricacies of our lasers and optics, and for allowing me to be a part of the construction and calibration of the new instrument.

I am also extremely grateful for being asked to work on collaborative projects. Thank you to Dr. Stuart Lindsay and Dr. Parminder Kaur for entrusting me with your light scattering experiments and letting me be a part of your exciting work. Thank you, Dr. Frank Tsen, for allowing me to do spectroscopy for your M13 virus studies. Thank you to Dr. Robert Best for providing a structural model of the N_loop through MD and for suggesting many ideas for the NMR analysis.

Thank you to my committee members for supporting me along the way. I have appreciated all of the useful feedback on my project and also future career plans.

Thanks to Deanna Clarke and Jill Kolp for being wonderful administrators: they have gone out of their way on a consistent basis to make my life easier.

I am extremely grateful for all of the financial support I have received from the following sources over the past five years: ASU Dept. of Physics, ASU Graduate College, Arizona Board of Regents, Biophysical Society, Graduate Professional and Student Association Grant, Mr. and Mrs. Wally Stoelzel, and the U.S. Department of Education.

Finally, thank you to my friends and family for being a wonderful support system over the past five years.

PROLOGUE

Diseases desperate grown
By desperate appliances are relieved,
Or not at all.

- Claudius, King of Denmark
Hamlet, Act IV, Scene III
William Shakespeare

Protein folding is at the heart of life: it is what allows organisms to function with remarkable complexity. And yet, when it goes awry, the results are debilitating. Perhaps a testimony to the power of protein folding is protein aggregation, the core process in numerous degenerative diseases.

Amyloid diseases were first studied in the 1960s following the discovery that transmissible spongiform encephalopathy is related to the presence of proteins, and not nucleic acids. Before this finding, scientists and physicians had considered diseases as infecting persons by two means: hyperplasia (as with cancers) or DNA replication (as with viruses). This marked a new class of diseases that infected host organisms by changing the morphology of native proteins, in a manner analogous to XLF mutations. At first, amyloid diseases were acknowledged primarily through histopathological assays. Physicians became very good at diagnosing these diseases, but the community lacked an understanding of the underlying driving forces that caused protein aggregation.

Biological physicists entered this new field, utilizing the fundamental principles of, among many areas, polymer physics, statistical mechanics, condensed-matter physics,

and spectroscopy. With interdisciplinary collaborations, biophysicists tackle fundamental questions that aimed to understand what causes soluble proteins to spontaneously nucleate into amyloid fibrils. Does a protein's structure and dynamics influence its propensity to aggregate? What are the competing inter-protein and intra-protein interactions? How can subtle sequence mutations or solution conditions alter intra-protein interactions and promote solubility?

In this work, we apply a host of techniques from the biophysicist's toolbox in order to address these questions, among others, for islet amyloid polypeptide, an intrinsically disordered hormone implicated in type 2 diabetes.

TABLE OF CONTENTS

Page

APPENDIX Page

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

Symbol Page

Symbol Page

INTRODUCTION

Intrinsically disordered proteins (IDPs) are a newly discovered class of proteins that lack tertiary structure, yet perform complex functions in vivo. In Chapter 1 we discuss the emerging importance of understanding IDPs in order to rationalize eukaryotic tasks, such as hormonal regulation and neural signaling pathways. Because of the overwhelming presence of IDPs in biological systems, it is important to understand how IDPs behave in solution. We discuss cases in which disorder has been shown to be both essential and advantageous for function. We know from the literature that IDPs have less hydrophobic and more charged residues than their folded counterparts. Furthermore, IDPs populate more compact states than natively folded proteins in denaturants, and subtle sequence mutations have been shown to greatly affect the dimensions of disordered proteins. To date, there is little experimental data that characterizes the structure and dynamics of IDPs, a consequence of the experimental challenges faced in characterizing systems that are so dynamic, and fluctuate on such quick time scales.

IDPs are prevalent in human diseases, and are notorious for their involvement in amyloid aggregations. We focus our efforts on understanding islet amyloid polypeptide (IAPP), an IDP implicated in type 2 diabetes. IAPP is a hormone co-secreted with insulin, and responsible for mammalian satiety. Human IAPP is perhaps the most amyloidogenic of all peptides, and was chosen due to its unique structural properties and importance in type 2 diabetes. Due to its size (3.8 kDa), IAPP can be readily synthesized and simulated. Though IAPP has been extensively studied over the past 2 decades, little is known about

the monomeric states of IAPP in solution. Furthermore, though researchers have focused on measuring the kinetics of IAPP aggregation, the mechanisms driving aggregation remain elusive. In Chapter 2 we discuss what is known about the structure of IAPP, and focus our discussion on current hypotheses that aim to explain the mechanisms of aggregation.

Novel experimental techniques are essential to understanding the structure and dynamics of IDPs. IAPP is a short peptide, requiring high temporal resolution. In Chapter 3 we discuss the parameters essential for characterizing an IDP's ensemble of conformations: experimental techniques must be able to probe dynamics occurring on times ranging from 100ns to 10 μs. We present a host of experimental techniques currently used to measure the structure and dynamics of IDPs, and discuss their advantages and limitations. We then introduce the experimental set-up employed in the Vaiana Research Group: tryptophan triplet quenching (TTQ) to measure end-to-end contact formation rates. This experimental technique allows us to monitor both equilibrium properties of IAPP: such as the end-to-end equilibrium distribution and also dynamic properties: such as the time for chain reconfiguration in solution. TTQ enables comprehensive characterization without prosthetic dyes or high concentrations. Finally, we discuss how chain configurations for IDPs in solution can be conceptualized in terms of polymer models.

In Chapters 4, we show that N_loop forms very long, stable β-turn fibers. We find through CD, FTIR, NMR and MD simulations that N_loop is a very rigid structure with little to no conformational freedom. We find that aggregation results from the small

entropic cost of forming inter-protein interactions. This suggests that the rigidity of N_loop could play a role in the pathological aggregation of IAPP, a possibility that has not yet been proposed. We propose that N_loop-N_loop interactions could be exploited to develop novel inhibitors and/or therapeutics for hIAPP aggregation.

In Chapter 5 we explore the role of prolines in determining rIAPP's unique structure, an analog of hIAPP that does not form amyloid aggregates and inhibits hIAPP aggregation. We find that the 3 prolines intrinsic to rIAPP add to increased chain stiffness, but are not the sole contributor. The R18H substitution found in rIAPP appears to substantially alter the monomeric state, and could play a role in aggregation kinetics.

The focus of this dissertation has been on the use of TTQ to study the dynamics of IAPP: a 37 amino acid peptide. TTQ is a powerful tool to characterize both the structure and dynamics of short peptides and proteins, but cannot be used to study larger biomolecules due to limitations imposed by the triplet state of tryptophan. For larger polymers such as DNA, we must employ an experimental technique that probes time scales from 0.1 μs to 0.1 ms. In Chapter 6, we discuss the internal motions of methylated and unmethylated DNA, measured via multi-angle dynamic light scattering. This novel technique allows us to characterize DNA in terms of polymer models, and quantify both the persistence length and radius of gyration. This work, completed in collaboration with Dr. Stuart Lindsay's group, was essential in understanding how the hydrophobicity of DNA alters its behavior in solution versus interfaces, and lead to the proposal of a new explanation

for how methylated DNA compacts chromatin in vivo, an essential mechanism in epigenetics and gene silencing.

CHAPTER 1

INTRINSICALLY DISORDERED PROTEINS

1.1. REDEFINING THE STRUCTURE-FUNCTION PARADIGM

The past two decades have shed light on the presence of a new class of proteins that challenge the traditional notion that structure determines function. For the majority of proteins, we equate biological function with a unique tertiary structure. Enzymes, for example, maintain a specialized 3-dimensional conformation that is specifically designed to catalyze a unique chemical reaction. Intrinsically Disordered Proteins (IDPs) lack 3-dimensional structure yet carry out biological functions. Currently, more than 300 IDPs have been experimentally identified. They account for an estimated 50% of mammalian proteins, becoming more populous with an organism's complexity[1,2]. Certain IDPs have even been shown to exhibit high evolutionary rates, becoming more specialized throughout the lifetime of their host organism[3]. IDPs populate an ensemble of conformations, which allows them to bind with multiple partners and perform more complex and diverse functions in vivo[4]. In contrast with natively folded proteins that often function as enzymes, IDPs primarily carry out signaling and regulatory functions. The majority of IDPs take on a stable secondary structure upon binding, though this is not obligatory for function[5]. Proteins that maintain disorder in the bound state form so-called "fuzzy" complexes. In these cases disorder seems to be a property inherent to their functionality[6]. IDPs appear to be increasingly important for eukaryotes, yet little is known about how protein function arises from disorder.

It has recently been proposed that the presence of disordered states actually increases protein efficacy[7]. It has been postulated that IDPs are advantageous for signaling, as they require shorter amino acid sequences due to their broader spatial arrangements: that is, fewer amino acids are required in an IDP than a NFP to cover the same distance[8]. In addition to simple geometrical arguments, more quantitative models have been developed to explain how unstructured proteins may enhance their binding kinetics. It has been experimentally validated that IDPs do not necessarily bind through an induced fit mechanism, as the reaction is not diffusion-limited[9]. Shoemaker et al. first proposed the "fly-casting mechanism" in which the unfolded state of a protein is able to weakly bind to its partner from large distances[10]. As the protein encounters the binding site, folding commences. In this protein binding schematic, an unfolded domain would be able to interact with the binding site from a larger distance, or so-called "capture radius", subsequently increasing the on-rate of binding. Turjanski and coworkers have shown that for the phosphorylated KID domain of the transcription factor CREB, a disordered region essential to transcriptional regulation, that the amount of structure present in the unbound state is anti-correlated with binding rates[11]. Thus, disorder appears essential for efficient binding in vivo. Furthermore, experimentally characterizing the dynamics of disordered regions remains essential to validating these hypotheses.

In addition to the fly-casting mechanism, other models have lead to insightful discussions on how disordered structures mediate function. It has been suggested that structural flexibility permits a given IDP to adopt different conformations depending on the

intended binding partner. This intrinsic flexibility enhances protein promiscuity, as is seen the tumor-suppressing protein p53, which boasts over 80 binding partners[12,13]. Furthermore, it has been proposed that the conformational agility of IDPs allows them to function as hubs in large, multi-pathway networks[14,15]. As is evident from the literature, disordered domains are critical for higher-order, eukaryotic function. And while many hypotheses exist to elucidate the mechanisms by which IDPs function, few have been validated experimentally.

1.2. INTRINSICALLY DISORDERED PROTEINS AND DISEASE

Intrinsically disordered proteins are essential for the functionality of numerous transcriptional activators and cellular signal regulators. Perhaps intrinsic to the complexity of such biological networks, there is a price to pay when these cellular machineries go array. Dyson and Wright showed that disordered regions are associated with several aberrant behaviors in the cell, including chromosomal translocations and protein aggregation[4]. These phenomena are thought to be a result of the structural reorganization of IDPs, which proliferate disease when undetected by the cell. When natively folded proteins mutate or misfold, they are rapidly degraded by cellular machinery. Conversely, IDPs are able to avoid proteostasis. In addition, low sequence complexity, an inherent property of many of IDPs, is strongly correlated with diseases. Such is the case with polyglutamine (polyQ) and its antagonistic role in Huntington's disease[16]. Table 1 highlights a select number of IDPs implicated in human disease:

IDP	Associated pathology
p53	Various cancers
BRCA-1	Breast cancer
α-Synucleain	Parkinson's disease
Prion Proteins	Transmissible spongiform encephalopathy
Amyloid-β	Alzheimer's disease
Hirudin and Thrombin	Cardiovascular disease
Islet amyloid polypeptide	Type 2 diabetes mellitus
Polyglutamine	Huntington's disease
Tau protein	Alzheimer's disease
HPV proteins	Cervical Cancer
α-Fetoproteins	Various cancer

Table 1.1. A select list of intrinsically disordered proteins implicated in human disease[17].

To understand the delicate interplay between disorder and disease, we must first gain insight into the physical mechanisms driving the functionality of IDPs.

1.3. WHY DON'T INTRINSICALLY DISORDERED PROTEINS FOLD OR AGGREGATE

It is clear that disorder is essential for the biological function of several proteins, particularly those found in eukaryotes. And despite the role IDPs play in disease, the overwhelming majority of IDPs competently function in the cell without aggregating or alerting proteostasis. Furthermore, denatured or misfolded NFPs are extremely unstable in the cell. So how are IDPs able to remain soluble in the cell when their thermodynamic free energy minimum is not an ordered state?

To resolve this conflict we must first understand why intrinsically disordered proteins remain unfolded in physiological conditions.

IDPs are highly dynamic in solution; they fluctuate between a large number of different conformations on nanosecond to microsecond timescales. On average, IDPs have a higher net charge and lower content of hydrophobic residues than natively folded proteins (Figure 1)[18].

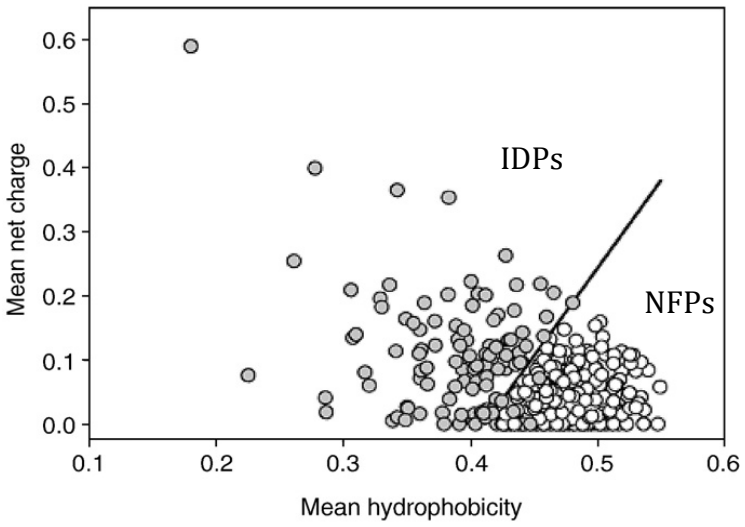

Figure 1.1. The Uverksy Plot (Uverksy, Proteins 2000). Select natively folded (open circles) and intrinsically disordered (closed circles) proteins plotted as a function of their mean net charge and hydrophobicity.

At first glance, one might assume that Figure 1 indicates that IDPs' structures and dynamics are determined by electrostatics, and thus void of attractive interactions. Apart from highly charged proteins, IDPs generally populate compact states. Marsh and Forman-Kay recently analyzed a number of IDPs to determine how hydrodynamic radius scales with the number of constituent amino acids. They found that the majority of IDPs populate states that are more compact than natively folded proteins (NFPs) in denaturing conditions, but more expanded than NFPs in physiological conditions[19]. One can look to the protein folding field to closer analyze the interactions present in IDPs. Hydrophobic collapse, especially due to nonpolar residues, is an essential component of folding. This raises the question of what drives collapse in IDPs, which generally lack hydrophobic clusters. To address this issue, Neuweiler and coworkers measured the intra-chain diffusion of polyglycine peptides (ie model polypeptides devoid of side chains), as well as disordered proteins with side chains of varying hydrophobicity. Their findings

suggested that the presence of side chains increases intra-chain diffusion times, suggesting that backbone-backbone interactions drive collapse[20]. Recent work from the Pappu Lab further supports this hypothesis, showing that polyQ populates a collapsed state despite the absence of hydrophobic residues[21]. It is interesting to note that residual secondary structure content does not correlate with IDP chain compaction, further supporting the view that transient backbone-backbone interactions play an important role in IDP conformations[16]. This does not, however, indicate that sidechain interactions are not important in modulating the dimensions of IDPs in solution: sequence-specific effects are responsible for the wide spectrum of chain dimensions seen for IDPs (Figure 1.2).

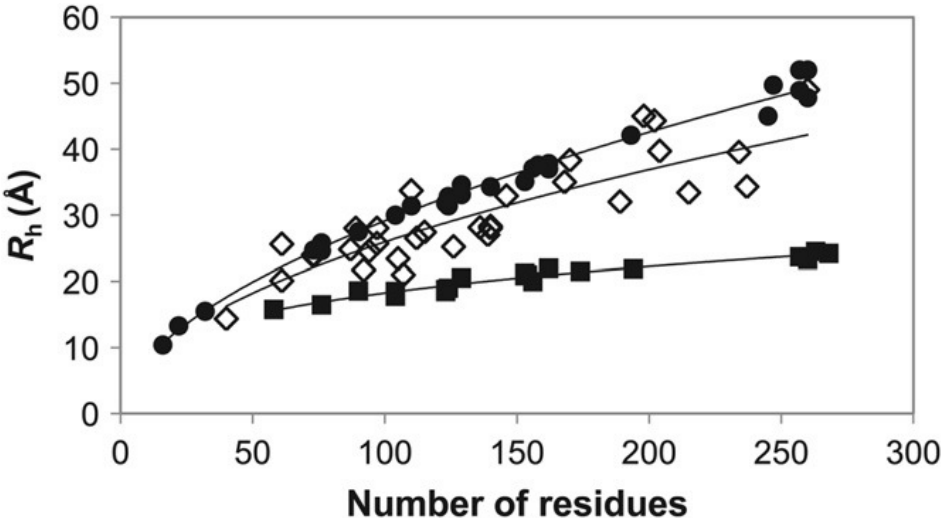

Figure 1.2. Compaction of IDPs in solution. Number of residues versus r_H for 20 folded (solid squares) and 27 chemically denatured (solid circles), and 32 intrinsically disordered (open diamonds) proteins. (Marsh, J.A. and Forman-Kay, J.D. 2010. Sequence Determinants of Compaction in Intrinsically Disordered Proteins. Biophys. J. 98:2383-2390.)

If IDPs occupy compact states and contain several interaction-prone regions required for function, how do they remain soluble and avoid non-specific binding in physiological conditions? In the cell, certain protective mechanisms help stabilize IDPs, but even in

vitro IDPs are generally more soluble than the denatured states of NFPs[4]. Specific protein-protein interactions must be preventing aggregation and non-specific binding, while simultaneously promoting solubility. Currently, there are several hypotheses of how IDPs function while avoiding non-native targets in the cell. Uversky proposes that IDPs escape aggregation and non-specific binding partners through "functional misfolding": the participation in "non-native" intramolecular interactions to maintain functionality[22]. IDPs may sequester specific binding regions from the solvent and target molecules, but remain dynamic enough to expose the binding region when a native binding partner is available. If this model holds, these non-native contacts must be short lived, as they have not yet been observed experimentally. This emphasizes the need for more sensitive experimental techniques able to probe long-range, transient contact formation. As the importance of intrinsically disordered proteins becomes increasingly recognized, so does the demand to experimentally observe non-native, transient tertiary contacts in disordered proteins in order to quantify their structure and dynamics.

REFERENCES

[1] Mohan A., Sullivan, W.J., Radivojac, P., Dunker, A.K., and Uversky, V.M. 2008. Intrinsic disorder in pathogenic and non-pathogenic microbes: discovering and analyzing the unfoldomes of early-branching eukaryotes. Molecular Biosystems. 4:328-340.

[2] Fink, A.L. 2005. Natively unfolded proteins. Structural Biology. 15:35-41.

[3] Dunker, K. 2009. conversation (unpublished)

[4] Dyson, H.J. and Wright P. Intrinsically unstructured proteins and their functions. 2005. Nature. 6:197-208.

[5] Frauenfelder, H., Sligar, S.G., and Wolynes, P.G., The energy landscapes and motions of proteins. 1991. Science. 254:1598-1603.

[6] Fuxreiter, M. and Tompa, P. 2012. Fuzzy Complexes: A More Stochastic View of Protein Function. Advances in Experimental Medicine and Biology. 725:1-14.

[7] Tompa, P. 2005. The interplay between structure and function in intrinsically unstructured proteins. Federation of European Biochemical Societies Letters. 579:3346-3354.

[8] Gunasekaran, K., Tsai, C.J., Kumar, S., Zanuy, D., and Nussinov, R. 2003. Extended disordered proteins: targeting function with less scaffold. Trends in Biochemistry. 28:81-85.

[9] Rogers, J.M., Steward, A., Clarke, J. 2013. Folding and Binding of an Intrinsically Disordered Protein: Fast, but Not 'Diffusion-Limited'. J. Am. Chem. Soc. 135:1415–1422.

[10] Shoemaker, B.S., Portman, J.J., and Wolynes, P.G. 2000. Speeding molecular recognition by using the folding funnel: The fly-casting mechanism. Proc. Nat. Aca. Sci. 97:8868-8873.

[11] Turjanski, A.G., Gutkind, J.S., Best, R.B., Hummer, G. 2008. Binding-Induced Folding of a Natively Unstructured Transcription Factor. PLoS Comput. Biol. 4: e1000060

[12] Oldfield, C.J., Meng, J., Yang, J.Y., Yang, M.Q., Uversky, V.N., and Dunker, A.K. 2008. Flexible nets: disorder and induced fit in the associations of p53 and 14-3-3 with their partners. BMC Genomics 9:S1.

[13] Zhang, W., Ganguly, D., and Chen, J. 2012. Residual structures, conformational fluctuations, and electrostatic interactions in the synergistic folding of two intrinsically disordered proteins. PLoS Comput. Biol. 8:e1002353.

[14] Uversky, V.N., Oldfield, C.J., and Dunker, A.K. 2005. Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling. J. Mol. Recognit. 18:343–384.

[15] Barbar, E. 2008. Dynein Light Chain LC8 is a dimerization hub essential in diverse protein networks. Biochemistry 47: 503–508.

[16] Karlin, S., Brocchieri, L., Bergman, A., Mrazek, J. and Gentles, A. J. 20002. Amino acid runs in eukaryotic proteomes and disease associations. Proc. Nat. Acad. Sci. 99:333–338.

[17] Uversky, V.N., Oldfield, C.J., and Dunker, A.K. 2008. Intrinsically Disordered Proteins in Human Diseases: Introducing the D^2 Concept
Annual Review of Biophysics. 37: 215-246.

[18] Uversky, V.N. 2000. Why are natively unfolded proteins unstructured under physiologic conditions. Proteins. 41:415-427.

[19] Marsh, J.A. and Forman-Kay, J.D. 2010. Sequence Determinants of Compaction in Intrinsically Disordered Proteins. Biophys. J. 98:2383-2390.

[20] Teufel, D.P., Johnson, C.M., Lum, J.K., and Neuweiler, H. 2011. Backbone-driven collapse in unfolded protein chains. Journal of Mol. Biol. 409:250-262.

[21] Crick, S.L., Jayarman, M., Frieden, C., Wetzel, C., and Pappu, R.V. 2006. Fluorescence correlation spectroscopy shows monomeric polyglutamine molecules form collapsed structures in aqueous solutions. Proc. Nat. Acad. Sci. 103:16764-16769

[22] Uversky, V.N. 2011. Intrinsically disordered proteins may escape unwanted interactions via functional misfolding. Biochim Biophys Acta. 1814:693-712.

CHAPTER 2

ISLET AMYLOID POLYPEPTIDE (IAPP)

2.1. MOTIVATION

IDPs are notorious for their role in amyloid aggregation. Though not all amyloid deposits are composed of proteins that are natively unstructured, several are. As discussed in Chapter 1, we are interested in characterizing the conformation and dynamics of monomeric IDPs to gain insight as to how they function, and what makes certain proteins more susceptible to aggregation. A protein's aggregation propensity is determined by the competition of intra- and inter-protein interactions. Proteins that remain soluble in vivo must have intra-protein interactions that stabilize the monomeric state. Moreover, changes to a protein's hydrophobicity, electrostatics, or hydrogen bonding networks can alter this delicate balance between intra- and inter-protein interactions and favor aggregation over solubility. We aim to quantify how sequence specificities and structural properties determine aggregation propensity.

To address this problem, we present a study of islet amyloid polypeptide (IAPP), an intrinsically disordered hormone which forms amyloid fibrils in persons with type 2 diabetes. IAPP has several unique features that make it an ideal candidate for studying the interplay of intra- and inter-protein interactions. IAPP is a relatively short protein, making it feasible for both solid-state synthesis and computational simulations. Furthermore, IAPP is a member of the (Ct) peptide superfamily: a family of proteins that

have high sequence homology, yet varying levels of aggregation propensity. We know that subtle changes in IAPP sequence can have large effects in aggregation propensity: for example, even though all mammals produce IAPP, type 2 diabetes has only been reported in primates and cats. From a practical standpoint, the rat variant of IAPP (rIAPP) is a suitable control peptide, as it has never been observed to form amyloid aggregates in vivo or in vitro. Rats do not form type 2 diabetes, unless mutated to produce the human variant of IAPP (hIAPP)[1]. This makes rIAPP is an ideal control peptide for experimental characterization.

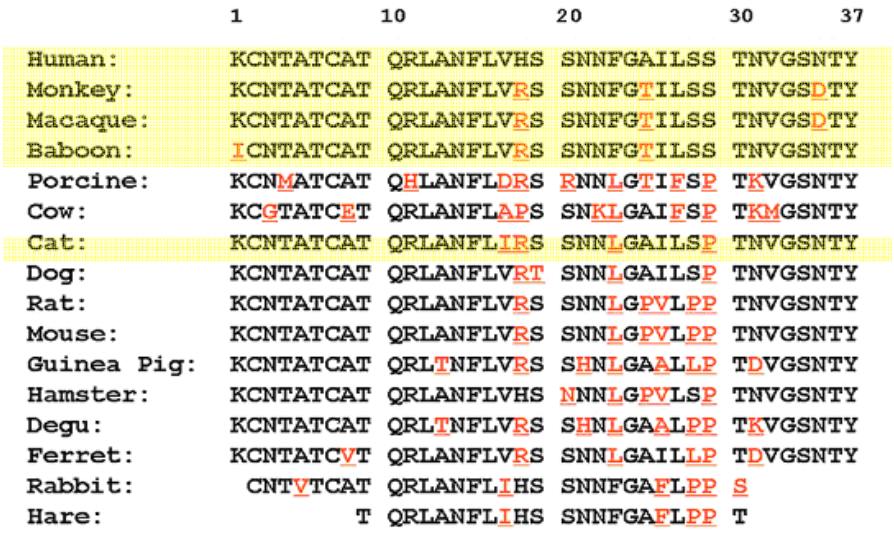

Figure 2.1 Amino acid sequences of IAPP for different species. Mutations from the hIAPP variant are highlighted in red. Yellow boxes indicate sequences known to form amyloids. Two unique features of the sequences include a disulfide bond from Cys2 to Cys7 and the presence of prolines in a majority of variants. (Cao, P., Abedini, A., amd Raleigh, D. 2013. Aggregation of islet amyloid polypeptide: from physical chemistry to cell biology. Current Opinion in Structural Biology. 23:82–89.)

Furthermore, experimental data on the monomeric state of IAPP is sparse. As a consequence, IAPP and its role in type 2 diabetes are not well understood. IAPP has

unique structural properties, such as an N-terminal disulfide bond: N_loop, and distinct amino acid mutations, such as the presence of proline residues in every soluble variant. These two features make this peptide an ideal model system for studying the physical and chemical properties that determine aggregation propensity.

2.2. THE BIOLOGICAL ROLE OF IAPP

Islet amyloid polypeptide (IAPP), also known as amylin, is a 37-residue intrinsically disordered hormone that is co-secreted with insulin in the beta cells of the pancreas, and is involved in glucose regulation and gastric emptying[2345]. IAPP is implicated in the pathogenesis of type 2 diabetes, due to its deposition in the form of amyloid fibers in the beta cells of the pancreas[6]. The process of amyloid aggregation leads to beta cell dysfunction and death[7]. Though IAPP amyloids are not the cause of type 2 diabetes, they are a contributing factor to beta cell failure, and the subsequent decrease in insulin production leading to replacement therapy[8]. Of all the amyloid forming proteins studied thus far, IAPP is arguably the most amyloidogenic, forming aggregates in the 10 μM concentration range.

IAPP is a member of the calcitonin peptide family, which includes calcitonin (Ct), calcitonin gene related peptide (CGRP), adrenomedullin (Ad), and intermedin[9]. These are structurally and genetically related intrinsically disordered hormones, with sequence homology ranging between 20% and 50%[10]. All members of the Ct family contain a functional disulfide bond that confers a short ring-like structure (N_loop) to the N-

terminus of the peptide (with the exception of Ad, in which the disulfide is located between residues 16 and 21).

As mentioned above, subtle difference in IAPP sequence equate to large differences in aggregation propensity. Rat islet amyloid polypeptide (rIAPP) differs from human IAPP (hIAPP) by only six of its 37 constituent amino acids and yet, to date, all evidence suggests that full-length rIAPP is unable to form amyloid fibrils. Many believe that rIAPP's remarkable solubility can be attributed to the presence of three proline mutations, A25P, S28P and S29P, as proline is a β-breaking residue due to the confinement of its backbone dihedral angles.

In this study, we will focus on two features of IAPP: (i) the N_loop and (ii) proline mutations, to study how physiochemical properties play a role in determining structure and dynamics, and subsequently, aggregation propensity.

2.3. THE ROLE OF THE N_LOOP ON IAPP STRUCTRURE AND AGGREGATION

To date, the structural specificities of monomeric IAPP remain ambiguous. Because IAPP is natively unstructured and readily aggregates at very low concentrations in vitro, structural data on IAPP are sparse. In its native state, the C-terminus of IAPP appears to be completely disordered[11]. As seen from NMR and simulations, N-terminal residues 8-20 transiently populate an α-helical structure[12,13,14,15]. Computational work suggests that this helix is nucleated and stabilized by N_loop: a short, ring-like structure formed by a

disulfide bond that constrains N-terminal residues hIAPP(1-8). Reduction or removal of the disulfide bond in N_loop renders IAPP incapable of activating cell response upon binding to the CGRP receptor, indicating that the N_loop is essential for function[16]. Vaiana et al.[17], by measuring end-to-end contact formation rates, showed that the N_loop drives collapse in monomeric IAPP. Based on MD simulations, they suggested that N_loop populates collapsed states by forming hydrogen bonds with disordered regions of the polypeptide chain[17]. In addition to playing a key role in the monomer structure of IAPP, N_loop has also been shown to affect aggregation. Chemical reduction or complete removal of the N_loop results in slower aggregation kinetics and altered structural properties of mature amyloid fibrils[18]. So far, it is unclear what role N_loop plays in IAPP aggregation. The current hypothesis is that the N_loop alters the monomer conformations populated by hIAPP in solution, indirectly affecting aggregation kinetics[16,19,20]. A direct role of the N_loop in the aggregation of hIAPP has not yet been explored. In chapter 4, we characterize N_loop-N_loop interactions, and explore what role N_loop may play in amyloid aggregation.

2.4. THE ROLE OF SEQUENCE MUTATIONS ON THE STRUCTURE AND DYNAMICS OF MONOMERIC IAPP

Much of the work on IAPP has been devoted to studying the effect of specific mutations in its disordered region on aggregation kinetics. As mentioned before, rIAPP differs from hIAPP by only six amino acids, yet is unable to form amyloid fibrils. Five of the six substitutions are found between residues 20 and 29. Historically, researchers studied fragments of IAPP to understand which regions were most amyloidogenic. It was found

that the ten amino acid fragment, hIAPP(20-29), forms amyloid fibers independently in solution while the rIAPP(20-29) analog does not[21]. These observations, along with others, led many to believe the "FGAIL" region (hIAPP (23-37)) was responsible for aggregation. Later studies shifted their attention on full-length hIAPP, and concluded that fibrils formed by peptide fragments exhibited very different structural and kinetic properties than the full-length peptide. Perhaps the most studied aspect of full-length hIAPP has been how sequence mutations alter aggregation kinetics. It has been shown that the presence of single proline mutations (I26P, G24, A13P) both hinders and inhibits fibril formation[22,23]. The focus on proline residues is an obvious choice, but in contrast, rIAPP variants that maintain prolines at position 25, 28 and 29 have been shown to be capable of forming an amyloid structure[24]. This suggests that subtle sequence variations greatly alter IAPP's propensity to aggregate. And though prolines appear to play an important role in amyloid formation, it is evident that several interactions are at play. In contrast to the large number of studies on the role of prolines in IAPP aggregation kinetics, experimental studies on the effect of prolines on the monomeric structure of IAPP (such as chain stiffness, hydrogen bonding, and steric effects) are scarce. In this study, we investigate the role of prolines on the structure and dynamics of the IAPP monomer in solution. As explained in Chapter 1, we expect that these properties will play an important role in understanding whether or not a protein is prone to aggregate. In Chapter 5 we address this problem by employing a nanosecond laser spectrometer to characterize the monomeric states of three IAPP variants, and discuss the implications for amyloid aggregation.

REFERENCES

[1] Höppener, J.W., Ahrén, B., and Lips, C.J. 2000. Islet amyloid and type 2 diabetes mellitus. The New England Journal of Medicine 343: 411–419.

[2] Cooper, G.J.S., Leighton, B., Dimitriadis, G.D., Parrybillings, M., Kowalchuk, J J.M., Howland, K., Rothbard, J.B., Willis, A. C., and Reid, K. B. M. 1988. Amylin Found in Amyloid Deposits in Human Type-2 Diabetes-Mellitus May Be a Hormone That Regulates Glycogen-Metabolism in Skeletal-Muscle. Proceedings of the National Academy of Sciences. 85:7763-7766.

[3] Reidelberger, R. D., U. Arnelo, L. Granqvist, and J. Permert. 2001. Comparative effects of amylin and cholecystokinin on food intake and gastric emptying in rats. American Journal of Physiology-Regulatory Integrative and Comparative Physiology. 280:R605-R611.

[4] Kahn, S. E., D. A. Dalessio, M. W. Schwartz, W. Y. Fujimoto, J. W. Ensinck, G. J. Taborsky, and D. Porte. 1990. Evidence of Cosecretion of Islet Amyloid Polypeptide and Insulin by Beta-Cells. Diabetes 39:634-638.

[5] Westermark, P., A. Andersson, and G. T. Westermark. 2011. Islet Amyloid Polypeptide, Islet Amyloid, and Diabetes Mellitus. Physiological Reviews 91:795-826.

[6] Cooper, G. J. S., A. C. Willis, A. Clark, R. C. Turner, R. B. Sim, and K. B. M. Reid.1987. Purification and Characterization of a Peptide from Amyloid-Rich Pancreases of Type-2 Diabetic-Patients. Proceedings of the National Academy of Sciences. 84:8628-8632.

[7] Westermark, G.T., Westermark, P., Berne, C., Korsgren, O., and Transpla, N.N. 2008. Widespread amyloid deposition in transplanted human pancreatic islets. N Engl J Med. 359:977-979.

[8] Potter, K.J., Abedini, A., Marek, P., Klimek, A.M., Butterworth, S., Driscoll, M., Baker, R., Nilsson, M.R., Warnock, G.L., and Oberholzer, J. 2010 Islet amyloid deposition limits the viability of human islet grafts but not porcine islet grafts. Proc Natl Acad Sci.107:4305-4310.

[9] Westermark, P., C. Wernstedt, E. Wilander, and K. Sletten. 1986. A Novel Peptide in the Calcitonin Gene Related Peptide Family as an Amyloid Fibril Protein in the Endocrine Pancreas. Biochemical and Biophysical Research Communications 140:827-831.

[10] Wimalawansa, S. J. 1997. Amylin, calcitonin gene-related peptide, calcitonin, and adrenomedullin: A peptide superfamily. Critical Reviews in Neurobiology. 11:167-239.

[11] Nanga, R.P., Brender, J.R., Xu, J., Hartman, K., Subramanian, V. and Ramamoorthy, A. 2009. Three-dimensional structure and orientation of rat islet amyloid polypeptide protein in a membrane environment by solution NMR spectroscopy. J. Am. Chem. Soc. 131: 8252–8261.

[12] Williamson, J.A., and Miranker, A.D. 2007. Direct detection of transient alpha-helical states in islet amyloid polypeptide. Protein Sci. 16:110-117.

[13] Andrews, M.N. and Winter, R. 2011. Comparing the structural properties of human and rat islet amyloid polypeptide by MD computer simulations. Biophysical Chemistry. 156:43-50.

[14] Khemtemourian, L., Engel, M.F.M., Kruijtzer, J.A.W., Hoppener, J.W.M., Liskamp, R.M.J., and Killian, J.A. 2010. The role of the disulfide bond in the interaction of islet amyloid polypeptide with membranes. European Biophys J. 39: 1359-1364.

[15] Laghaei R, Mousseau N, and Wei G. 2011. Structure and thermodynamics of amylin dimer studied by Hamiltonian-temperature replica exchange molecular dynamics simulations.J Phys Chem B. 115:3146-54.

[16] Koo, B.W., Miranker, A.D. 2005. Contribution of the intrinsic disulfide to the assembly mechanism of islet amyloid. Protein Sci. 14:231–239.

[17] Vaiana, S. M., R. B. Best, W. M. Yau, W. A. Eaton, and J. Hofrichter. 2009. Evidence for a Partially Structured State of the Amylin Monomer. Biophysical Journal 97:2948-2957.

[18] Conner, A.C., Simms, J., Barwell, J., Wheatley, M., and Poyner, D.R. 2007. Ligand binding and activation of the CGRP receptor. Biochemical Society Transactions 35:729-732.

[19] Yonemoto, I. T., G. J. A. Kroon, H. J. Dyson, W. E. Balch, and J. W. Kelly. 2008. Amylin Proprotein Processing Generates Progressively More Amyloidogenic Peptides that Initially Sample the Helical State. Biochemistry 47:9900-9910.

[20] Williamson, J. A., J. P. Loria, and A. D. Miranker. 2009. Helix Stabilization Precedes Aqueous and Bilayer-Catalyzed Fiber Formation in Islet Amyloid Polypeptide. Journal of Molecular Biology 393:383-396.

[21] Tenidis, K., Waldner, M., Bernhagen, J., Fischle, W., Berhmann, M., Weber, M., Merkle, M.L., Voelter, W., Brunner, H., and Kapurniotu, A. 2000. Identification of a penta- and hexapeptide of islet amyloid polypeptide (IAPP) with amyloidogenic and cytotoxic properties. Journal of Molecular Biology. 295: 1055–1071.

[22] Abedini, A. and Raleigh, D. 2006. Destabilization of human IAPP amyloid fibrils by proline mutations outside of the putative amyloidogenic domain: is there a critical amyloidogenic domain in human IAPP? J. Mol. Biol. 355, 274–281.

[23] Abedini, A., Meng, F. and Raleigh, D. 2007. A single-point mutation converts the highly amyloidogenic human islet polypeptide into a potent fibrillization inhibitor. J. Am. Chem. Soc. 129:11300–11301.

[24] Green, J., Goldsbury, C., Mini, T., Sunderji, S., Frey, P., Kistler, J., Cooper, G. and Aebi, U. 2003. Full-length rat amylin forms fibrils following substitution of single residues from human amylin. J Mol Biol. 326:1147-1156.

CHAPTER 3

USING NANOSECOND LASER SPECTROSCOPY TO MEASURE THE STRCTURE AND DYNAMICS OF IDPS

3.1. INTRODUCTION

The disordered states of proteins are experimentally challenging to study due to their intrinsic heterogeneity and rapid interconversions between substantially different conformations. IDPs are by definition proteins that populate unfolded states. Up until the late 1990s, methods used to probe disordered states of proteins involved exposure to harsh organic solvents, high temperatures, and denaturants: such as Guanadinium Chloride (GdmCl) and Urea. Characterizing IDPs involved classifying them as structured or unstructured, with little quantitative analysis on what the unstructured state looked like. In this chapter, we will explore both the advantages and limitations of new techniques used to study IDPs.

3.1.1. HIGH RESOLUTION TECHNIQUES

In order to characterize IDPs, traditional spectroscopy needs to be used in conjunction with novel, high-resolution techniques. Traditional spectroscopic techniques, such as Circular Dichroism (CD), Fourier Transform infrared spectroscopy (FTIR) and fluorescence, reveal information such as percent secondary structure and solvent exposure of aromatic residues. Though this information is valuable, it does not offer

detailed information on conformation and dynamics. Over the past decade, substantial progress has been made in measuring both global properties and transient interactions in IDPs. Local contact formation and transient tertiary structure offer information on how disordered regions behave in solution, ultimately providing insight into how IDPs may function. Measuring these properties experimentally is challenging due to the timescale on which IDPs fluctuate between conformers. In order to detect long-range contacts and intrachain dynamics of IDPs, experimental techniques must be able to probe interactions occurring between 10^{-8} to 10^{-5} seconds.

Perhaps the most common experimental technique used to studying IDPs is Nuclear Magnetic Resonance (NMR). NMR allows for immediate identification of disordered regions, as IDPs show wide and dispersed chemical shifts. Nuclear Overhauser effect spectroscopy (NOESY) can be used to detect hydrogen-bonding networks, by measuring strength of NOEs, which is directly related to spatial separations. Furthermore, information on secondary structure propensities can be detected by comparing measured chemical shifts to known libraries of the chemical shifts of random coil structures. Though this method is widely used, it comes with its limitations, as disordered proteins are never truly random coil. Paramagnetic Relaxation Enhancement (PRE), a more recent NMR technique, allows one to measure tertiary contacts between 1-3.5 nm[1]. This technique requires labeling with prosthetic dyes, which may perturb the conformations of IDPs. Also, the distance dependence of the relaxation enhancement factor for PRE dyes goes as $1/r^6$, making this technique more sensitive to larger distances rather than short contact formations. NMR relaxation rates can offer general information on the overall

dynamics of the system at hand[2,3]. The largest limitation of NMR is protein concentrations; NMR typically requires mM range concentrations. IDPs can aggregate at concentrations as low as 10 μM, making NMR in aqueous conditions extremely challenging or implausible.

There are several experimental techniques used to characterize global properties of IDPs, such as the hydrodynamic radius, r_H. Pulse field gradient (PFG) NMR can be used to measure the translational diffusion coefficient of an IDP, though it remains very sensitive to the hydration of the protein. Small-angle X-ray scattering and dynamic light scattering are used to measure r_G and r_H, respectively. Since small-angle scattering is most sensitive to long-range interactions, these techniques are often infeasible for small, aggregation-prone IDPs. Fluorescence correlation spectroscopy (FCS) is also used to measure r_H, and when used in combination with single molecule Förster Resonance Energy Transfer (smFRET), has the ability to measure long-range structural properties of IDPs. Again, as with PRE-NMR, smFRET requires the addition of large dyes and linkers and is most sensitive to larger distances on the order of nanometers. Though all these techniques are powerful in their own right, their limitations have so far hindered the study of the structure and dynamics of monomeric IAPP in solution.

It is important that we take into consideration IAPP's unique properties when choosing an experimental technique. IAPP readily aggregates at 100 μM, and thus we need a technique that does not require high concentrations. Because small perturbations in IAPP can cause drastic changes in its aggregation propensity, we need a technique that does not

require prosthetic dyes. Finally, IAPP has relatively fast intrachain dynamics that require an experimental technique able to measure interactions occurring on a timescale of 10 ns to 10 μs. For these reasons, we employ a relatively new experimental technique to measure contact formation rates between tryptophan and cysteine in IAPP. This techniques was originally developed by Lapidus et al. in the laboratory of W. A. Eaton at NIH, following important work by Strambini and Gonnelli on the phosphorescence of tryptophan, and was later employed to study IDPs [4,5,6,7,8,9,13]. By optically exciting tryptophan to its triplet state, which can be quenched by cysteine, we are able to measure the time it takes for the two residues to come into contact. This technique provides information on both the end-to-end distance probability distribution as well as the intrachain diffusion. The advantage of this technique is that we can gain information on both the structure and the dynamics, without significantly perturbing IAPP's native state.

3.2. END-TO-END CONTACT FORMATION MEASUREMENTS

3.2.1. PHOSPHORESCENCE OF TRYPTOPHAN

As mentioned above, typical end-to-end contact formation for intrinsically disordered proteins is between 10 ns and 10 μs. In order to monitor contact formation rates, one needs a probe that is long lived. Tryptophan's triplet state has an absolute lifetime of between 5 and 7 seconds[8]. When tryptophan is present in a polypeptide, in the absence of hemes and coenzymes, this lifetime decreases to about 1 ms. Under normal solution conditions, where impurities in solvents and samples are inevitable, the average lifetime

is approximately 40 μs. This is sufficiently long relative to the fast intra-protein motions of a disordered protein. Beginning in the 1980s, this intrinsic property of tryptophan has been exploited to study the dynamics and flexibility of proteins in solution[9101112].

3.2.2. CONTACT FORMATION MEASUREMENTS VIA CYSTEINE QUENCHING

In order to measure the rate of formation for a specific contact in a polymer chain, we exploit the fact that the triplet state of tryptophan is efficiently quenched by contact with cysteine or cystine, but not with other naturally occurring amino acids (with the exception of tryptophan)[6]. The mechanism of quenching is most likely due to electron transfer, as the phenomenon decays exponentially with distance, with a contact distance of 4 Å[13]. This unique property of cysteine or cystine can be used to explore contact formation in disordered proteins.

3.2.3. THEORETICAL FRAMEWORK

This technique works by first placing tryptophan and cysteine (or cystine) on opposite termini of a polypeptide chain. Tryptophan is optically excited to its first excited state via a pulsed UV laser source. From here, a small number of molecules undergo a transition to a triplet state through intersystem crossing (Figure 3.1).

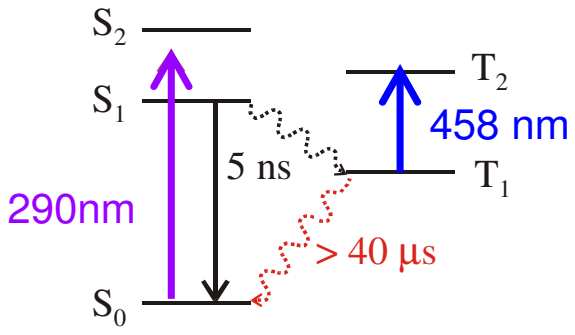

Figure 3.1. Jablonski diagram of Tryptophan's electronic energy states.

The lifetime of this state is long lived relative to the time it takes for the two ends of the polypeptide chain to come into contact through intrachain diffusion. While in solution, the disordered polypeptide will reconfigure, giving rise to the probability that cysteine and tryptophan will encounter each other at a rate of k_{D+}(Figure 3.3). If cysteine and tryptophan come into contact for a long enough time or for a sufficient number of times, quenching of the excited state will occur. If the two ends do not stay in contact for long enough, they will diffuse away without quenching at a rate of k_{D-}(Figure 3.3). During the measurement, the first triplet state is continuously being pumped to the second triplet state via a continuous wave (CW) laser tuned to 458 nm (see Figure 3.1). The absorbance at 458 nm is proportional to the population of molecules in the first triplet state. Thus, when quenching occurs, the absorbance value decreases. Absorbance as a function of time yields the average rate of observed quenching, k_{obs}.

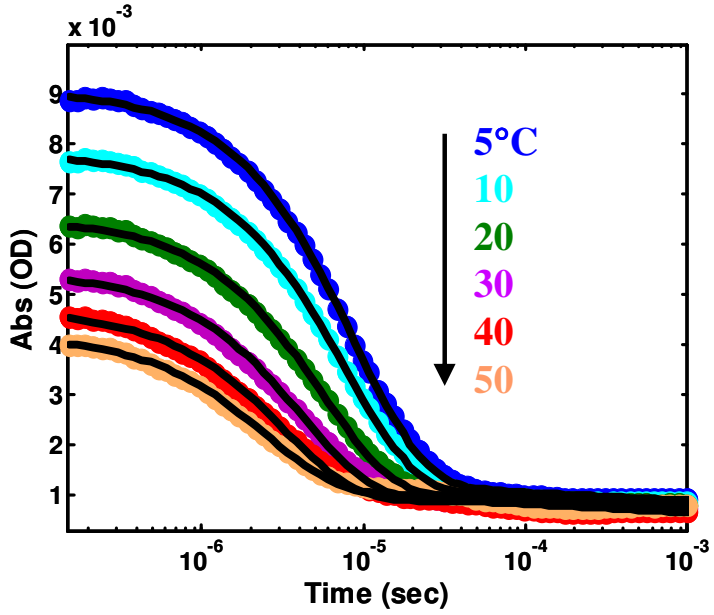

Figure 3.2. Time-resolved absorbance of tryptophan triplet→triplet absorbance at 458 nm. These data, corresponding to the temperatures listed in the graph, were fit to an exponential decay, together with a second offset function which varied linearly with log(t) for times greater than 4e-5 s (to account for observed decays at long times, due to a radical photoproduct) and an overall offset constant[10]. The exponential decay has a characteristic time, τ_{OBS}: the inverse of k_{OBS}. The data shown here is for Pramlintide Y37W in 6M GdmCl, pH=4.9.

Figure 3.3. Schematic for the reaction of tyrpotophan (red molecule) quenching via contact formation with cystine (yellow molecule).

The observed quenching rate, k_{obs}, is equal to the rate at which contact formation occurs, k_{D+}, times the rate at which quenching occurs for the encounter complex, q.

$$k_{obs} = k_{D+}\frac{q}{k_{D-}+q} \tag{1}$$

This observed rate, written for a two-state model, can be re-written in terms of a diffusion-limited rate and a reaction-limited rate.

$$\frac{1}{k_{obs}} = \frac{k_{D-}}{k_{D+}q} + \frac{1}{k_{D+}}, \qquad \frac{1}{k_{obs}} = \frac{1}{k_R} + \frac{1}{k_{D+}} \tag{2}$$

where

$$k_R = K_{eq}q, \qquad \text{and} \qquad K_{eq} = k_{D+} / k_{D-} \tag{3}$$

The diffusion limited rate, k_{D+}, will be the rate for which $q>2k_{D-}$, or the condition in which diffusion is sufficiently slow, such that every time cysteine and tryptophan come into contact, quenching occurs. The reaction-limited rate, k_R, is the case in which $k_{D-}>>q$, where diffusion is so fast such that the quenching reaction sees an equilibrium distribution of open versus closed states. It's important to note that by writing the reaction in terms of these two rates, we separate the variables that depend only on structure (k_R) from the variables that depend on dynamics (k_{D+}). This is not the case for purely diffusion-limited quenchers, which are used in other experimental techniques. Experimentally, we can go from one limiting case to the other by changing the solvent viscosity, thus obtaining both k_R and k_{D+}. In aqueous solution we are close to the reaction-limited rate, and at very high viscosities, we are close to the diffusion-limited rate.

Though we have written a simple two-state model, in reality quenching is a distance-dependent phenomenon. Thus, mathematically, q is a function of distance: q(r). Szabo, Schulten, Schulten showed that for such a reaction, you can approximate contact

formation as a diffusive process in one dimension, that is, along the end-to-end distance coordinate. Using a time-correlated sink-sink function, the introduction of q(r) into the rate equations shown above yields an analytical expression for k_{obs} in terms of k_R and k_{D+}:

$$k_R = \int_{a_0}^{l_c} p(r)q(r)dr, \tag{5}$$

$$k_{D+} = (\int_{a}^{lc} (D(x)\, p_{eq}(x))^{-1} [\int_{x}^{lc} \delta q(y)\, p_{eq}(y) dy]^2 dx)^{-1} \tag{6}$$

With this assumption, the reaction- and diffusion-limited rates can be related to the end-to-end distance distribution, P(r), which contains structural information, and the intrachain diffusion coefficient, D, respectively[14].

3.2.4. OBTAINING REACTION- AND DIFFUSION-LIMITED RATES EXPERIMENTALLY

For each sample, data is collected at T=5°, 10°, 20°, 30°, 40°, and 50°C. A set of 5 measurements at each temperature, resulting from 256 laser pulses, are averaged to obtain the curves shown in Figure 3.2. Each curve is fit to an exponential decay, plus a slowly decaying function which varied linearly with log(t) for times greater than 4e-5 s to account for known photoproducts and electron recombination[10]. A constant offset is also added to this function. The initial exponential decay represents the triplet lifetime and the inverse of this characteristic time of decay is taken as the observed relaxation rate, k_{OBS}.

To obtain reaction-limited and diffusion-limited rates, we measure k_{OBS} as a function of solvent viscosity. For 6M GdmCl measurements, we alter the solvent viscosity by adding

between 0%-32% sucrose in solution. For viscosity-dependence of k_{OBS}, we perform Singular Value Decomposition (SVD) analysis, and fit only the sum of the first 12 components. These are then globally fit using the function described above. By plotting $1/k_{OBS}$ as a function of solution viscosity, we are able to obtain k_R and ηk_{D+}.

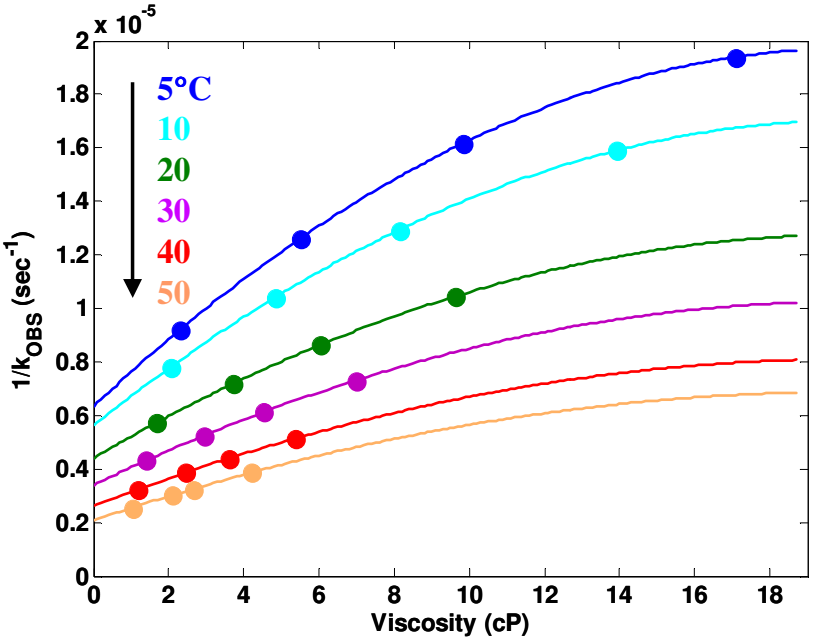

Figure 3.4. Viscosity dependence of the observed quenching rates. By measuring observed quenching rates as a function of solvent viscosity, we can separate effects due to intra-chain dynamics from those due to overall chain dimensions. Viscosity dependent measurements yield values of k_R and k_{D+} at various temperatures. In the graph above, 1/kR is the intercept and 1/ ηkD+ is the initial slope at η→0. These data are for Pramlintide in 6M GdmCl with varying %sucrose, pH=4.9

We then assume that k_R depends solely on the temperature and k_{D+} depends on both the temperature, T and the viscosity, η. Because k_{OBS} does not necessarily change linearly with solvent viscosity, we fit the curve to the following empirical function:

$$k_{D+} = \eta(1 - A\eta)^{-1} k_{D+}(T) \tag{7}$$

In terms of Figure 3.4, $1/k_R$ is taken as the y-intercept derived from the fitting dunction and $1/\eta k_{D+}$ is the slope at zero viscosity.

3.2.5. HOME-BUILT NANOSECOND TIME RESOLVED SPECTROMETER

The instrumentation used to measure contact formation in IDPs was home-built in Dr. Sara M. Vaiana's Laser Spectroscopy and Biophysics Lab (LSBL) in the Department of Physics and Center for Biological Physics at Arizona State University. The set-up was constructed over a two-year time frame. Planning, construction and calibration was done by Alejandro Solis, Sebastian Hoeffner, Sara M. Sizemore, and Stephanie Cope, under the supervision of Dr. Sara. M. Vaiana.

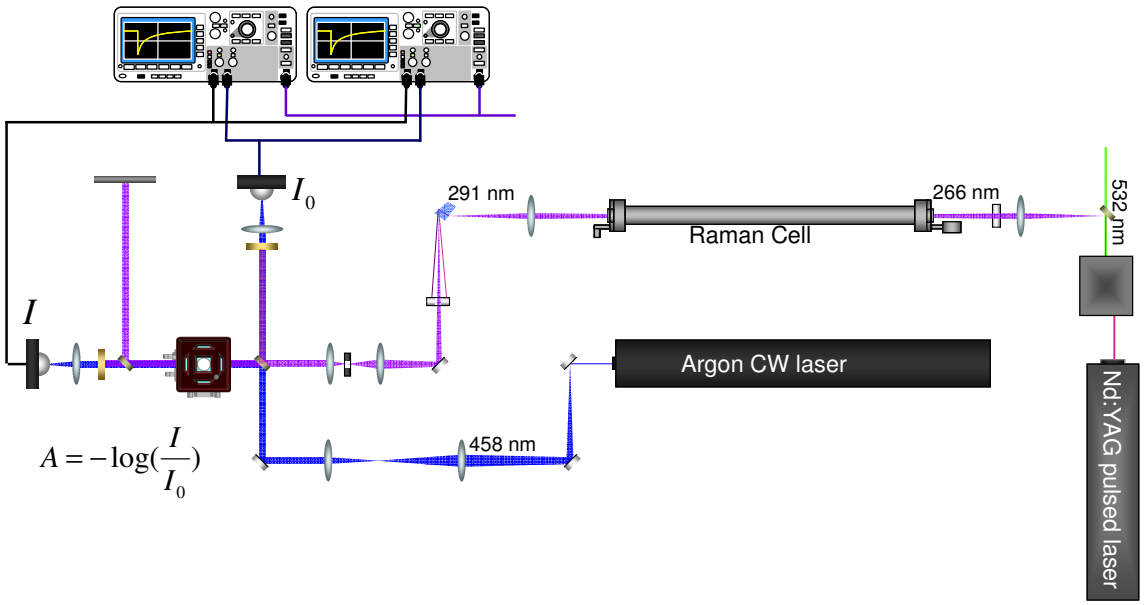

Figure 3.5. Homebuilt nanosecond-resolved pump-probe spectometer. Nd:YAG laser is used to pump trypotphan to its first excited triplet state and Argon laser is used to probe the population of tryptophan's first triplet state in time.

3.3. POLYMER MODELS

It is important to note that though IDPs remain unfolded in all solvents, the degree to which their conformations are expanded or compact can differ greatly from one protein to another, and in varying solvent conditions. We can use basic polymer models to quantify the end-to-end distance of a protein and study how this measurement changes with solvent. The simplest example is an ideal chain. Monomer segments are assumed to be of equal lengths, b. The monomers are uncorrelated such that the polymer mimics a random walk. The root mean square of the polymer's end-to-end distance is given by:

$$\sqrt{\langle R^2 \rangle} = N^{1/2} b$$

where N is the number of monomers making up the poymer[9]. The ideal chain allows polymerization units to loop back onto itself. Since occupying the same position in space is a physical impossibility, overlapping positions in space must be excluded from the ideal model. This leads to the excluded volume model, in which the end-to-end distance follows a self-avoiding random walk. The average end-to-end distance therefore takes the form:

$$\sqrt{\langle R^2 \rangle} = N^{3/5} b$$

Flory pioneered this theoretical work and showed that polymers exhibit scaling behavior of the form:

$$\sqrt{\langle R^2 \rangle} \propto N^{\gamma}$$

where the exponent γ is a function of the model used to describe the interactions within the polymer chain. Because this depends on solvent, Flory grouped solvents into one of three classes: good, bad or theta solvents. A polymer in "good" solvent is one in which the polymer behaves as in the excluded volume model, where $\gamma_{theoretical}=3/5$. A polymer in a 'theta' solvent is one in which effective attractive solvent interactions counteract the excluded volume interactions and the polymer behaves as an ideal chain with $\gamma_{theoretical}=1/2$. A "bad" solvent is one in which attractive interactions are stronger than excluded volume interactions leading to $\gamma_{theoretical}<1/2$. We note that a maximally compact polymer will scale as a sphere with $\gamma_{theoretical}=1/3$. Characterization of polymers has been extended to explain proteins in different solvents. Flory's polymer models were applied to proteins in the experimental work of Tanford, followed later by Kuhn et al. Both groups found that the end-to-end distance of proteins in highly denaturing solvents

(i.e. 6M GdmCl) scales with the number of constituent amino acids, and therefore chain length, with a scaling factor: $\gamma_{empirical}$=0.59, very close to the theoretical value for the excluded volume model[11,12].

We can use these models to predict conformational changes for IDPs in aqueous solvent and GdmCl. Aqueous conditions mimic a theta or bad solvent, favoring a compact protein structure, while GdmCl acts as a good solvent, promoting an extended conformation. In conjunction with spectroscopic techniques, we will use these polymer models to quantify changes in IDP conformation caused by amino acid substitutions or solvent conditions.

REFERENCES

[1] Clore, G. M., and Iwahara, J. 2009. Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes, Chem. Rev. 109:4108-4139.

[2] Mittag, T., Marsh. J, Grishaev, A., Orlicky, S., Lin, H., Sicheri, F., Tyers, M., and Forman-Kay, J.D. 2010. Structure/Function Implications in a Dynamic Complex of the Intrinsically Disordered Sic1 with the Cdc4 Subunit of an SCF Ubiquitin Ligase. Structure. 18:494-506.

[3] Mittag, T. and Forman-Kay, J.D. 2007. Atomic-level characterization of disordered protein ensembles. Current Opinion in Structural Biology. 17:3-14.

[4] L.J. Lapidus, P.J. Steinbach, W.A. Eaton, A. Szabo and J. Hofrichter. 2002 Effects of chain stiffness on the dynamics of loop formation in polypeptides. Appendix: Testing a 1-dimensional diffusion model for peptide dynamics. J. Phys. Chem.B 106:11628.

[5] L.J. Lapidus, W.A. Eaton and J. Hofrichter. 2002. Measuring dynamic flexibility of the coil state of a helix-forming peptide. J. Mol. Biol. 319:19.

[6] L.J. Lapidus, W.A. Eaton and J. Hofrichter. 2000. Measuring the rate of intramolecular contact formation in polypeptides. PNAS. 97:7220).

[7] Vaiana, S. M., R. B. Best, W. M. Yau, W. A. Eaton, and J. Hofrichter. 2009. Evidence for a Partially Structured State of the Amylin Monomer. Biophysical Journal 97:2948-2957.

[8] Gonnelli, M. and G. B. Strambini, G.B. Phosphorescence lifetime of tryptophan in proteins.1995. Biochemistry. 34:13847-13857.

[9] Strambini GB, Cioni P, Felicioli RA. 1987. Characterization of tryptophan environments in glutamate dehydrogenases from temperature-dependent phosphorescence. Biochemistry. 26:4968-75.

[10] Cioni, P. and Strambini, G.B. 1994. Pressure Effects on Protein Flexibility Monomeric Proteins. Journal of Molecular Biology. 242:291-301.

[11] Gabellieri, E. and Strambini G.B. 1989. Phosphorescence properties of Trp-84 and Trp-310 in glyceraldehyde-3-phosphate dehydrogenase from Bacillus stearothermophilus. Biophys Chem. 33:257-64.

[12] Calhoun, D.B., Englander, S.W., Wright, W.W., and Vanderkooi, J.M. 1988. Quenching of room temperature protein phosphorescence by added small molecules. Biochemistry. 27:8466-74.

[13] Lisa J. Lapidus, William A. Eaton, and James Hofrichter. 2001. Dynamics of intramolecular contact formation in polypeptides: distance dependence of quenching rates in a room-temperature glass. PRL 87:258011-258014.

[14] Szabo, A., Schulten, K. and Schulten, Z. 1980. First passage time approach to diffusion controlled reactions. J. Chem. Phys. 72: 4350-4357.

CHAPTER 4

CYCLIC N-TERMINAL LOOP OF AMYLIN FORMS NON-AMYLOID FIBERS

The results reported in this chapter have been published in Biophysical Journal, 105:7, Cope, S.M., Shinde, S., Best, R.B., Ghirlanda, G., and Vaiana, S.M., Cyclic N-terminal loop of amylin forms non amyloid fibers, Copyright (2013). The results presented here have been reprinted with permission from Elsevier Publishing. This work was done in collaboration with Dr. G. Ghirlanda in the Department of Chemistry and Biochemistry at ASU and Dr. R. B. Best in the laboratory of Chemical Physics at NIDDK, NIH. My contribution to this work consisted of all the experimental data presented below, including FTIR, CD, TEM, ThT fluorescence, optical microscopy, and solution state NMR.

4.1. INTRODUCTION

As mentioned in Chapter 2, the sequence of the N_loop is highly conserved among naturally occurring variants of IAPP (see Figure 4.1). While some members of this family are prone to aggregate, others are remarkably soluble. This is the case with rIAPP, which has also been found to inhibit hIAPP amyloid formation and was recently shown to interact with hIAPP during fibril formation through its N-terminal region[1]. Although the nature of this interaction has not been identified, association through alpha helical intermediates had originally been proposed[2]. According to the most recent structural model of hIAPP amyloid fibrils, derived from 2D NMR restraints, monomers are aligned at the N-terminus and form in register parallel β-sheets[3]. In this structure, shown in

Figure 4.2, the N_loop is not involved in direct inter-molecular interactions. Removal of the N_loop, however, alters both the mass per length distributions of hIAPP fibers and the kinetics of fibril formation, with effects that vary with experimental conditions[4-6].

```
Human     KCNTATCATQ-RLANFLVHSS-NNFGAILSST-NVGSNTY
Rat       KCNTATCATQ-RLANFLVRSS-NNLGPVLPPT-NVGSNTY
N_loop    KCNTATCA
```

Figure 4.1. Sequences of human IAPP (hIAPP), rat IAPP (rIAPP) and N_loop. The disulfide bond between residues 2 and 7, forming a loop at the N-terminus of the sequences, is indicated in blue. All sequences have a free N-terminus and amidated C-terminus.

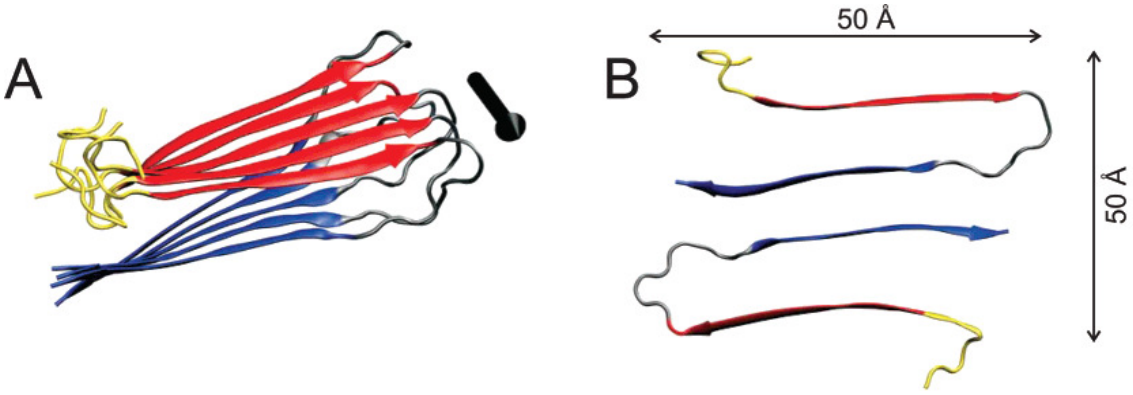

Figure 4.2. Currently accepted model of hIAPP amyloid fibrils, derived from solid state 2D-NMR experiments by Luca et al. (A) Ribbon representation of one cross-β-molecular layer, with N- and C-terminal B-strand segments in red and blue, respectively. The black arrow indicates the fibril axis. (B) Cross-sectional view of two amylin molecules in the protofilament. (From Luca, S., W. M. Yau, R. Leapman, and R. Tycko. 2007. Peptide conformation and supramolecular organization in amylin fibrils: Constraints from solid-state NMR. Biochemistry 46:13505-13522.)

The mechanism by which the N_loop affects hIAPP aggregation is not yet understood, but it is important for rationalizing the kinetics and potentially developing inhibitors[7,8].

So far the leading hypothesis has been that the N_loop alters hIAPP aggregation kinetics indirectly, by affecting the conformational preferences of the monomer in solution[5,6,9]. For example, ring opening of the N_loop by chemical reduction decreases the helical content of rIAPP (assumed to be a good structural model of hIAPP) [9,10]. Early work by Vaiana et. al using tryptophan triplet quenching (as described in Chapter 3) had shown that N_loop induces the collapse of rIAPP, hIAPP and a model hydrophilic peptide in aqueous solution[11]. Molecular simulations in that work indicated that the N_loop has a rigid structure. The authors suggested that it may act as a nucleus driving compaction[11]. Though clearly the N_loop affects the conformational preferences of the hIAPP monomer in solution, a direct role of N_loop in the aggregation of hIAPP and in the inhibition of hIAPP by rIAPP has not yet been explored. The findings presented here indicate that direct association of the N_loop alone is a favorable process at μM concentrations, the same concentrations at which hIAPP forms amyloid fibers.

4.2. METHODS

MATERIALS

Fmoc(9-fluorenylmethoxycarbonyl)-protected amino acids were purchased from Novabiochem. HOBt (N-hydroxy benzotriazole) and HBTU (O-Benzotriazole-N, N, N', N'-tetramethyl-uronium-hexafluoro-phosphate) were purchased from Genscript. N,N-iisopropylethylamine, or Hünig's base (DIPEA), and N-Methyl-2-pyrrolidone (NMP), used as base in solid phase peptide synthesis, were purchased from Sigma Aldrich. Piperidine (Sigma-Aldrich) was used for deprotection. Rink Amide ChemMatrix® was

purchased from Matrix innovations. Dimethyl formamide (DMF), Dichloromethane (DCM) and Acetonitrile were purchased from Fisher Scientific and were used without further purification.

PEPTIDE SYNTHESIS AND PURIFICATION

N1-8 was synthesized on a 0.25 mmol scale using a CEM Liberty Automated Microwave Peptide Synthesizer and PALChem Matrix resin. After synthesis, the peptide was thoroughly washed 5 times by DMF followed by DCM. After washing, the peptide was stored on the resin at -20 °C. For deprotection, the peptide was shaken for one hour in 20% piperdine, 0.1 M HOBt in DMF. The cleavage cocktail consisted of 95% trifluoroacetic acid (TFA) + 2.5% Water + 2.5% triisopropylsilane (TIS) at the ratio of 150 µL/ 10 mgs of resin. N1-8 was purified using Reverse Phase High performance liquid chromatography (HPLC) on a Waters 600E system. Crude peptide was purified on a C18 semi-preparative column (Vydac/Grace Deerfield, IL) at a gradient of 0-14% Acetonitrile with 0.1% (v/v) TFA over 45 minutes.

DISULFIDE FORMATION AND OXIDIZED PEPTIDE PURIFICATION

1.0 mM of lyophilized peptide was dissolved in 30% DMSO and 3% Acetic Acid. The sample was stirred with a magnetic stir bar at 1100 rpm. During this time, the formation of the intra-molecular disulfide bonds was monitored via HPLC on a C18 analytical column. The reaction was deemed complete when the reduced peptide's HPLC peak was no longer apparent: approximately 10 hours. After this time, the sample was frozen

and lyophilized. Oxidized peptide was re-purified on a C18 semi-preparative column to separate any un-reacted peptide.

Calibrated MALDI-TOF mass spectrometry indicated the presence of a -2 Da species (above the margin of error the instrument), corresponding to the oxidized form of the peptide, N_loop. To further support these findings, a maleimide sulfhydryl reaction was performed. 1 mg 3-(N-Maleimidopropionyl)-biocytin (Cayman Chemical Company Ann Arbor, MI) was dissolved in 180 µL 20 mM PBS buffer, pH=7.0 and 20 µL acetonitrile. 1 mg pure, oxidized N_loop was dissolved in 90 µL 20 mM PBS buffer, pH=7.0 and 10 µL acetonitrile. This solution was combined with the 3-(N-Maleimidopropionyl)-biocytin solution and shaken for four hours. At 1 hour increments, 20 µL aliquot was removed and frozen in dry ice. After four hours, the four aliquots were analyzed via MALDI-TOF MS.

HPLC peaks were analyzed by a Voyager Systems 4320 (Applied Biosystems) matrix assisted laser desorption/ionization-time of flight mass spectrometer (MALDI-TOF MS). The peak corresponding to a molecular weight of 807 Da (corresponding to the amidated and oxidized form of the peptide) was analyzed further for purity by analytical HPLC, using a reverse phase C18 analytical column 214TP54 (Length 250 mm × ID 4.6 mm) particle size 5 µm using the same gradient conditions with 0.9 mL/min flow rate. A single peak eluting at a gradient corresponding to the hydrophobicity of N_loop was collected, immediately frozen in liquid nitrogen, lyophilized, and kept at -20 °C.

FIBER PREPARATION

For each measurement, solutions of N_loop were prepared fresh at T=23 °C. The N1-8 samples were prepared in deoxygenated conditions and in the presence of a 2 time molar excess of TCEP-HCl (Sigma Aldrich). Buffer was 50 mM sodium acetate at pH=4.9. After adding the buffer to lyophilized peptide in a glass vial, samples were stirred with a glass stirring rod until all peptide was dissolved. The concentration of peptide was measured by absorbance at 205 nm, using an extinction coefficient of ε=11,040 $M^{-1}cm^{-1}$. This was obtained from the linear dependence of absorbance versus concentration measured for solutions of N1-8 in the 60 μM - 200 μM range (solutions were prepared by direct dilution from a 200 μM stock solution).

CIRCULAR DICHROISM

CD Spectra of the samples were measured in quartz cuvettes (Starna Cells) ranging in pathlength from 10 to 0.1 mm. Stock solutions were diluted with Millipore H_2O, resulting in final buffer concentrations ranging from 10-25 mM NaAc. A Jasco J-710 spectropolarimeter (Jasco Company) was used with a 1 nm bandwidth. For each sample, eight spectra taken with a 0.2 nm pitch at a 50 nm/min scan speed were averaged. Before data analysis, spectra were buffer subtracted.

TRANSMISSION ELECTRON MICROSCOPY (TEM)

Carbon Type B (15-25 nm) 200 mesh grids with Formvar coating (Ted Pella, Inc. Redding, CA) were glow discharged using a plasma cleaner immediately before applying

sample. N_loop sample was diluted to a final concentration of 7-25 mM, and allowed to adsorb to the grid for 5 minutes. Sample was blotted and rinsed with Millipore water using Whatman filter paper, No. 1. Immediately after blotting, 5 μL of freshly filtered 1% uranyl acetate was applied for 2 minutes, then blotted. Grids were imaged on a Philips CM 12 Scanning Transmission EM, operated at 80 kV.

THIOFLAVIN-T ASSAY

Thioflavin-T (ThT) binding assays were used to detect the presence of amyloid fibrils within N_loop samples. Final solutions contained 10 μM ThT, 12.5 mM KCNTATCA in 50 mM sodium acetate pH=4.9. All thioflavin-T experiments were conducted in a QuantaMaster 40 (Photon Technology International Birmingham, NJ). Emission spectra were monitored at an excitation wavelength of both 350 nm and 450 nm, while excitation spectra were monitored at emission wavelengths of 450 nm and 480 nm. Excitation and emission slits were kept at 2 nm. Thioflavin-T was purchased from Pierce, Thermo Fisher Scientific (Rockford, IL).

FOURIER TRANSFORM INFRARED SPECTROSCOPY (FTIR)

N_loop fiber solutions were lyophilized after 1 week aging at 4 °C. Monomeric N_loop samples were lyophilized immediately after HPLC purification. To minimize the possibility of aggregates forming during the lyophilization process, small fractions corresponding to pure N_loop were instantly frozen in liquid nitrogen and maintained in the solid (frozen) phase during the entire lyophilization process. HPLC analysis of

samples pre- and post lyophilization show conservation of mass. This observation supports the lack of insoluble aggregate formation during lyophilization. Lyophilized peptide was mixed with oven-dried KBr using a mortar and pestle. The sample was further dried under vacuum for 1 min, and pressed into a pellet. FTIR spectra were recorded on a Bruker Vertex 70 series instrument. The optical chamber was flushed with Nitrogen for 15 minutes before each scan. The interferograms were averaged over 512 scans with a resolution of 1 cm^{-1}. Atmospheric compensation (for residual H_2O vapor and CO_2) and baseline correction (using 10 iterations of concave rubberband correction) were performed using Opus V6.5 software. Control KBr spectra were subtracted from peptide spectra. In data shown, the fiber absorbance value was rescaled by a constant such that the magnitude of the Amide I band matched the absorbance value for the monomer pellet.

OPTICAL MICROSCOPY

All imaging was performed on an Olympus BH2-UMA optical microscope. A quarter waveplate was aligned to 45 degrees for birefringence imaging. All measurements were made at 10x magnification.

NMR SPECTROSCOPY

Lyophilized peptide was dissolved in 100% deuterated dimethyl sulfoxide (DMSO)(Cambridge Isotopes; Andover, MA) for a final peptide concentration of 1.3 mM. All spectra were acquired at 25 °C on a 500 MHz Varian spectrometer. 1D NMR

spectra of N_loop were recorded before each experiment and did not change, indicating that no significant conformational change or aggregation occurred for the duration of the experiments. Spectra were calibrated to the residual solvent peak for DMSO: 2.49 ppm. Two-dimensional COSY and TOCSY spectra were recorded using a mixing time of 100 ms. ^{1}H chemical shifts were determined from the TOCSY spectrum together with the NOESY spectrum. Secondary chemical shifts, used to identify secondary and residual structure in IDPs, were calculated by subtracting reference random coil values from De Simone et al. from our measured chemical shifts (Figure 4.9) [12,13]. Comparison with alternative random coil libraries is reported in Figure 4.10[12,14,15]. Two-dimensional NOESY experiments were measured for both 400 ms and 80 ms mixing times. NOEs were exported from VNMRJ (Varian) and further analyzed in CARA (Keller, Rochus, The Computer Aided Resonance Assignment Tutorial). NOE peak volumes were converted into relative strengths by calibrating them to the crosspeak of the two ^{1}Hβ protons of cysteine 2, assumed to be at a fixed distance of 1.8 Å. NOE strengths were classified into weak (<4.0 Å), medium (<3.3 Å) and strong (<2.6 Å) using cutoffs from reference (26).

MOLECULAR DYNAMICS SIMULATIONS

Replica exchange molecular dynamics (REMD) simulations of the amylin monomer were carried out in explicit TIP3P water, with the Amber ff03* force field for the protein and ions[17]. Periodic boundary conditions with a 3.5 nm truncated octahedron cell, containing the peptide and 1033 water molecules, four sodium and 6 chloride ions, were used. Long range electrostatic interactions were computed using particle mesh Ewald with a real

space cutoff of 0.9 nm and a grid spacing of 0.1 nm. All bond lengths were constrained to their reference values using the LINCS algorithm. This system was equilibrated at constant pressure and a temperature of 295 K, using a Parrinello-Rahman Barostat and Langevin thermostat with a friction coefficient of 1 ps^{-1} for 200 ps, with an integration time step of 2 fs. Following the equilibration, REMD simulations were performed at constant volume, with 32 replicas spanning a range of temperature from 278 to 595 K, using a Langevin thermostat to control the temperature of each replica, and with exchange attempts every 2 ps. Structures from the 295 K replica were clustered using the g_cluster utility with the single linkage method and a cutoff of 0.08 nm. Additional REMD simulations of a pair of amylin monomers were performed using the same simulation parameters, and all structures forming direct contacts (defined as a minimum atom-atom distance between monomers of less than 0.2 nm) were analyzed.

4.3. RESULTS

N_LOOP FIBER FORMATION

We set out to investigate the properties of the isolated N_loop (residues 1-8 of hIAPP) in the oxidized and reduced form. We prepared the linear, reduced form of hIAPP(1-8), (N1-8) by solid phase peptide synthesis. We obtained the oxidized, cyclic form (N_loop) by oxidation in DMSO of N1-8 followed by purification and lyophilization (see Methods). Remarkably, we found that the cyclic peptide immediately forms fibers visible by eye when dissolved in aqueous buffer at peptide concentrations ranging from 0.65 mM to 50 mM (Figure 4.3a). By contrast, the linear peptide N1-8 did not form fibers under

the same experimental conditions. Light microscopy imaging of the fibers reveals strong birefringence (Figure 4.3b), indicating a large degree of supra-molecular ordering in the fibers. Consistently, TEM images show large bundles of aligned microfibers (Figure 4.3c). Fibers of the same morphology where observed in TEM within 1 hour of sample preparation, at concentrations as low as 57 μM (Figure 4.4). Fibers did not appear to dissolve under highly acidic conditions (pH<2) or after heating to 80ºC. Based on Thioflavin T (ThT) binding assay the fibers do not appear to be amyloid in nature (Figure 4.5) [18]. This is further confirmed by the absence of β-sheet features in both CD and FTIR spectra presented below.

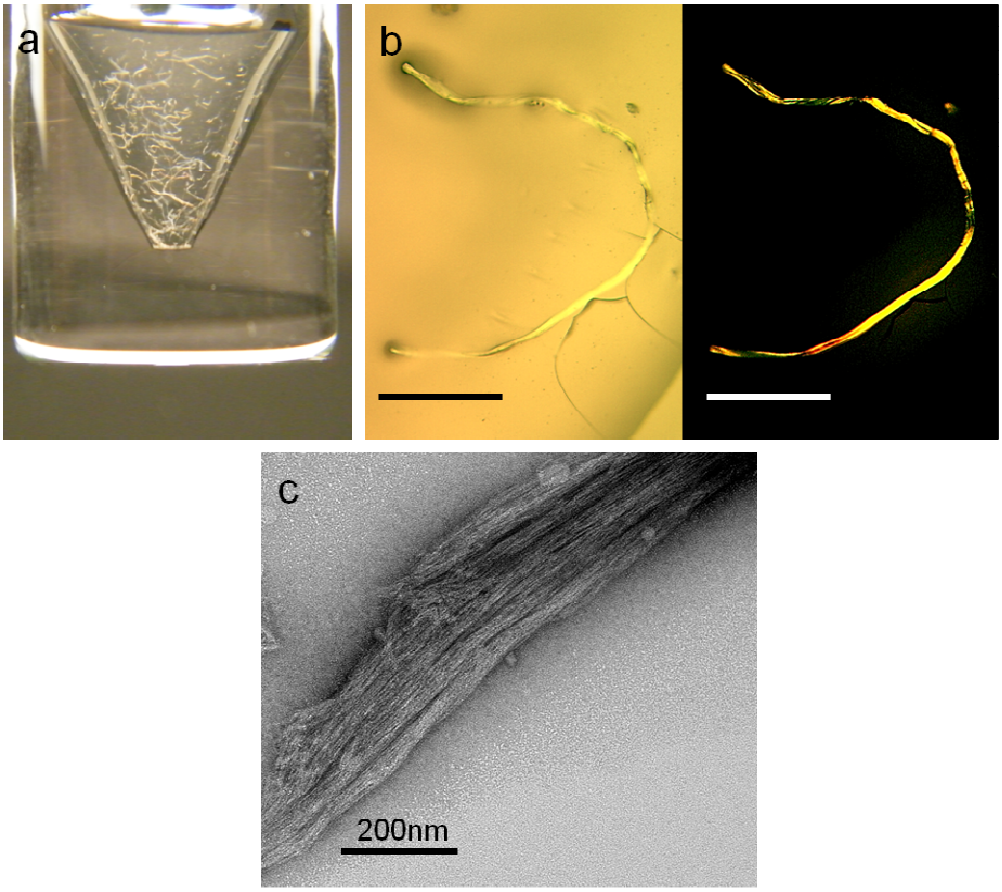

Figure 4.3. (a) Oxidized N_loop fibers, visible immediately upon dissolving peptide in buffer (50 mM peptide in 50 mM sodium acetate,

pH=4.9). (b) Optical microscopy of N_loop fibers in the absence (right) and presence (left) of crossed polarizers, revealing fiber birefringence (scale bar: 1 mm) (c) TEM image of fibers prepared at 25 mM (scale bar: 200 nm). Fibers of same morphology were observed at concentrations as low as 57 μM (below: Figure 4.4).

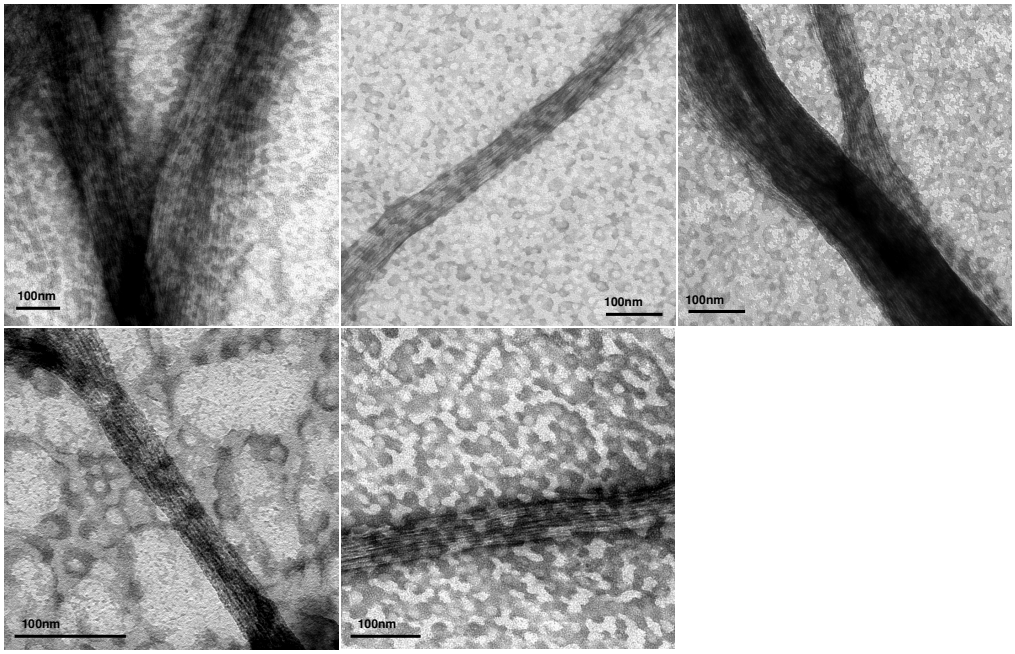

Figure 4.4. TEM images of N_loop fibers formed at different peptide concentrations. From top left to bottom right: 50mM, 1.4mM, 1mM, 0.15mM, 57μM. All scale bars are 100nm. Lyopholized peptide was dissolved in 50mM NaAc, pH=4.9. Samples were deposited onto grids within 1 hour of sample preparation. Fibers are evident at concentrations as low as 57μM.

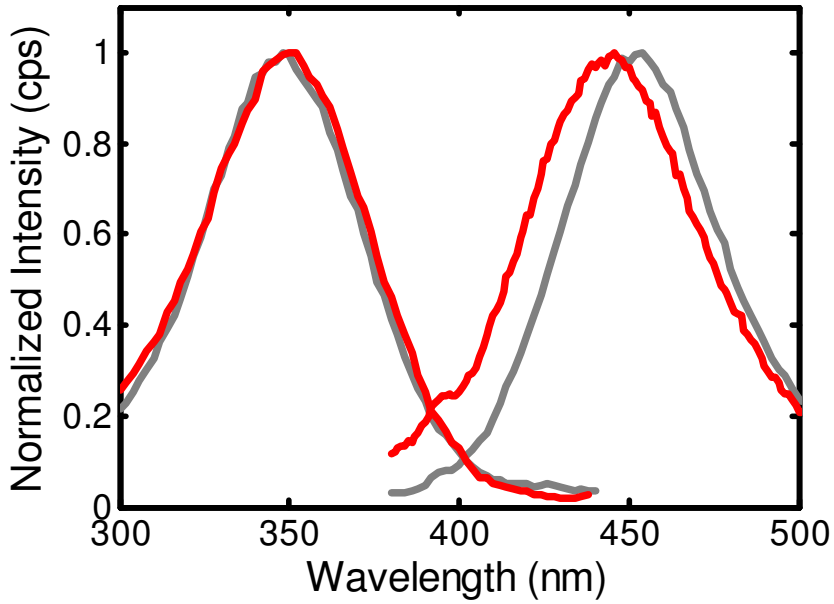

Figure 4.5 Thioflavin T binding assay of N_loop fibers. Fluorescence excitation (λ_{em}=450nm) and emission spectra (λ_{ex}=350nm) of 10uM ThT solutions in 50mM NaAc pH=4.9, in the absence (red) and presence (grey) of 12.5mM N_loop fibers. Only a slight shift in the emission spectrum of ThT was detected, in contrast to the large shift of emission and excitation peaks (λ_{ex} =450nm and λ_{em}=480nm) typically observed in the presence of amyloid fibers.

The sequence of N_loop (Figure 4.1) does not contain particularly hydrophobic or insoluble side-chains that would suggest such aggregation. We therefore hypothesize that structural features associated with peptide cyclization are conducive to fibril formation. To investigate such features we analyzed the linear and cyclic monomer as well as the fibers, using circular dichroism (CD) and Fourier transform infrared spectroscopy (FTIR) (Figure 4.6).

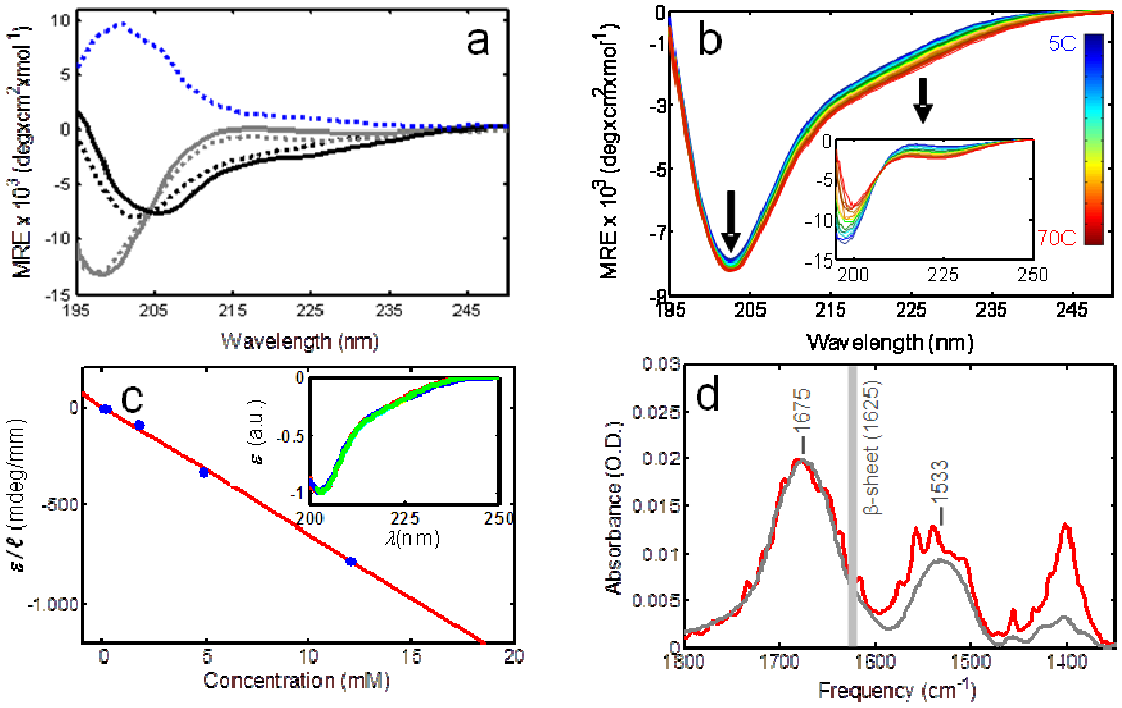

Figure 4.6. (a) CD spectra of linear N1-8 (gray line) and cyclic N_loop (black line) peptides at 5 ºC compared to pure component spectra for disordered coil (dotted grey line), type I β-turn (dotted black line) and type II β-turn (dotted blue line) (redrawn from ref(20)); MRE: Mean Residue Ellipticity. (b) Temperature dependence of CD spectra from 5 to 70 ºC of N_loop and N1-8 (inset) (660 μM peptide in 50 mM sodium acetate pH=4.9). (c) concentration dependence of N_loop CD signal amplitude at minimum, ε, rescaled by pathlength l (closed circles) and linear fit (line); renormalized CD spectra do not vary with concentration (inset) (20 °C, pH=4.9 in 10-25 mM sodium acetate buffer). (d) FTIR spectra of KBr pellets of HPLC purified N_loop peptide (gray) versus fibers (red) in the Amide region. N_loop samples were lyophilized immediately after HPLC purification. N_loop fiber samples were lyophilized after 1 week aging at 4 °C (7 mM peptide in 50 mM NaAc pH=4.9). Peak positions for the Amide I (1675 cm^{-1}) and II (1533 cm^{-1}) bands are given for the N_loop spectrum. The vertical gray line indicates the characteristic Amide I peak of IAPP amyloid fibrils. All spectra are atmospheric compensated, baseline corrected, and KBr subtracted.

The CD spectrum of the linear peptide, shown in Figure 4.6a, is typical of linear disordered coils[19]. In contrast, the spectrum of the cyclic peptide deviates significantly from disordered coils and contains distinct features. We compared it to the pure component spectra obtained by spectral deconvolution of model cyclic peptides of similar length, in which the structure was determined by X-ray crystallography and solution NMR [20,21]. Figure 4.6a shows the overlay of the type II (blue line) and type I β-turn spectrum (black dotted line) with the N_loop spectrum (black line) [20]. The resemblance with the type I β-turn spectrum is striking, considering that in short cyclic peptides CD is sensitive to very small backbone conformational distortions [21]. We therefore attribute the CD spectral features of cyclic N_loop to a type I β-turn structure.

We derived further information on the structural rigidity of the peptides by monitoring the change in CD spectra as a function of temperature (Figure 4.6b). While the spectrum of N1-8 displays significant changes upon increasing temperature, consistent with a non-cooperative loss of polyproline II structure (as observed in other disordered proteins, N_loop remains almost unchanged[22,23]. We conclude that cyclization imparts rigidity to the N_loop rendering the type I β-turn stable to thermal denaturation.

To gain insight into the structural changes occurring upon fiber formation we measured the CD spectra as a function of peptide concentration in the 10 μM to 12 mM range, which includes concentrations at which fibers were clearly visible. Surprisingly, the signal amplitude followed a simple linear dependence with concentration (Figure 4.6c)

and no spectral changes were observed (Figure 4.6c inset). By contrast, significant spectral changes are typically observed during amyloid fibril formation, as peptides convert from disordered structures to β-sheet structures. Two explanations are possible: either the N-loop monomer does not change conformation upon aggregation or the fibers do not contribute to the CD signal. In the latter case we would expect to observe a saturation of the signal amplitude at increasing peptide concentrations, as the monomers approach the solubility limit. The linear dependence of the signal at concentrations beyond those needed for fiber formation suggests instead that the monomer structure does not change upon fiber formation, and that the fiber structure does not introduce appreciable supra-molecular chirality. There remains a third possibility that, even at 10 μM, the N_loop is present as a very low molecular weight oligomer (dimer/trimer/...) rather than as a monomer. In this event, our results would mean that the N_loop monomer has the same conformation in the low molecular weight oligomers (rather than monomers) and in the fibers. The question is then whether the structure of the monomer would be much affected by the formation of oligomers. Further evidence provided below from REMD simulations support the contention that the monomer structure is rigid and is not significantly affected by aggregation.

To further test these conclusions, we compared the FTIR spectra of the HPLC purified peptide and of the fibers in the amide region (Figure 4.6d). The appearance of fine structure in the fiber spectrum (red curve) indicates quasi-crystalline order in the fibers. These data confirm the high degree of microscopic order revealed by birefringence (Figure 4.3b). Further, the observed increase in intensity of the near-amide III bands

(1300-1500 cm-1), which reflects contributions from methylene (CH2) and methyl (CH3) bending vibrations of aliphatic side-chains, is consistent with packing of these side-chains within the fibers[24]. In contrast, the positions of the amide I and II peaks, which reflect backbone secondary structure, do not change[25]. Thus, there is no significant backbone structural rearrangement when the N_loop is incorporated into the fibrils. The FTIR spectrum also confirms the absence of β-sheet structures in the fibers, which would appear as a distinct Amide I peak at (1620-1630 cm^{-1}) [26-28]. The position of the Amide I peak at 1675 cm^{-1} is consistent with a β-turn structure (observed in model linear and cyclic β-turn peptides at 1672-1674cm^{-1}), in agreement with our interpretation of the CD spectrum of the N_loop (Figure 4.6a) [29-32].Taken together the data presented here indicate that the disulfide bond constrains the N_loop into a fairly rigid, well defined structure containing a type I β-turn. This structure highly favors the formation of stable inter-peptide interactions resulting in the observed formation of fibers.

SOLUTION NMR

To obtain more detailed structural information on N_loop we measured proton chemical shifts and NOEs by 2D-NMR solution spectroscopy. Given the low solubility of N_loop in aqueous solvent we performed these measurements in 100% deuterated DMSO. We calculated proton secondary chemical shifts by subtracting residue specific random coil values from our measured chemical shifts[13]. These values are typically used to identify propensities to populate alpha helix versus β-sheet dihedral angles in intrinsically disordered proteins[12]. Figure 4.9 shows a comparison between values obtained here for N_loop and ones previously reported by Yonemoto et al. and Williamson et al. for the

corresponding residues of full length hIAPP and rIAPP respectively[6,33]. These data do not indicate significant structural changes between the N_loop and the corresponding residues of full length IAPP, considering the difference in solvent conditions and the variations originating from the use of different random coil libraries. This comparison between libraries is shown in Figure 4.8[13-15]. In Appendix A, we also consider the effect of aqueous dilutions of the DMSO, Appendix A: Figures 1, 2).

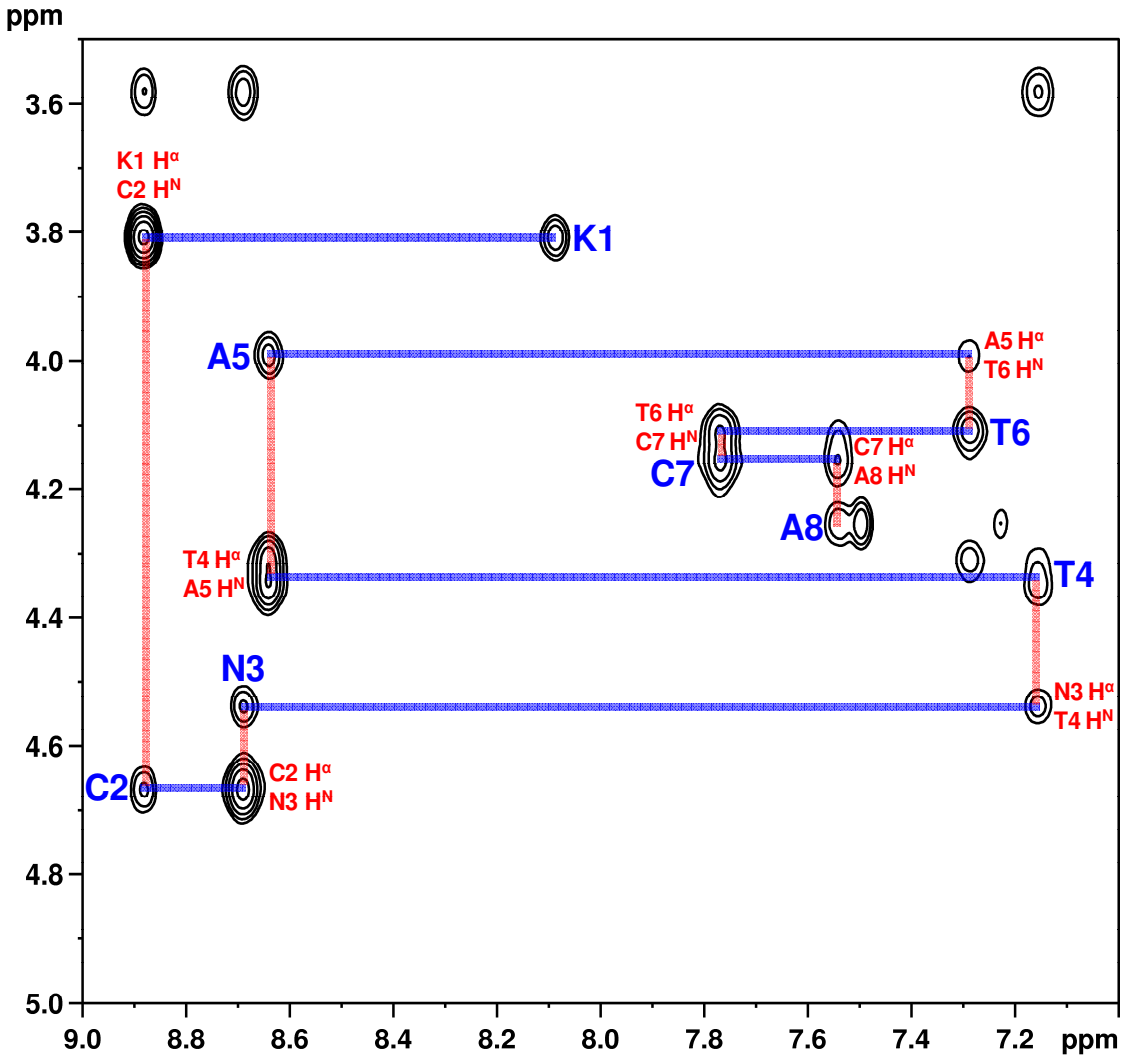

Figure 4.7. Backbone walk (400 ms NOESY)

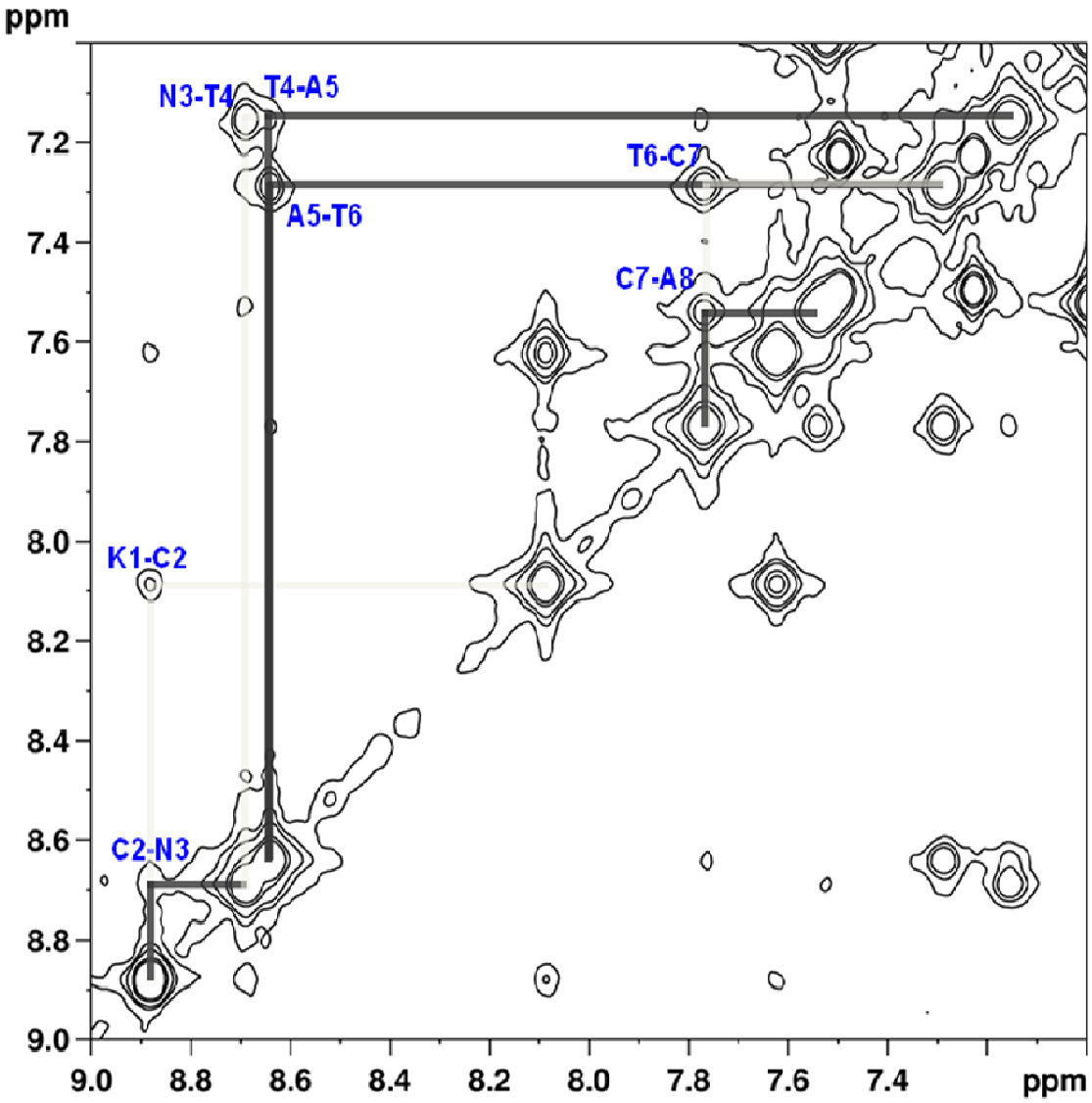

Figure 4.8. NMR spectrum of i->i+1 in NH-NH region (400 ms NOESY)

K1	N 8.086	α 3.808
C2	N 8.881	α 4.667
N3	N 8.687	α 4.534
T4	N 7.152	α 4.346
A5	N 8.636	α 3.985
T6	N 7.285	α 4.107
C7	N 7.768	α 4.148
A8	N 7.54	α 4.255

Table 4.1. $^1H^N$ and $^1H^\alpha$ Chemical Shifts for N_loop in 100% DMSO

Residues	NOE	Simulation dist. [Å]
1-2	Medium	4.05
2-3	Medium	4.44
3-4	Strong	3.03
4-5	Medium	3.13
5-6	Strong	2.73
6-7	Strong	2.40
7-8	Strong	2.88

Table 4.2. Comparison between measured NOEs and distances from simulations: Amide-amide sequential NOE strengths and averaged distances computed as $\langle r^3 \rangle^{-1/3}$ from the REMD simulation replica at 295 K. Additional NOEs and comparison to simulations are shown in Appendix A: Figure 3.

In Table 4.2 we report the strength of amide-amide sequential NOEs obtained from 2D-NOESY experiments, which we have classified into weak, medium and strong according to cutoffs of reference (16). As an internal calibration standard we used the known distance (1.8 Å) between the two β protons of C2. Analysis of the 80 msec and 400 msec

mixing time spectra yielded similar results (Figure 4.8). We note that direct comparison between these values and those reported by Yonemoto et al. would require the use of robust and self-consistent internal calibration standards. Generally speaking though, our NOE peaks are consistent with those observed by Yonemoto et al., with the exception of the presence of a T4-A5 crosspeak (absent in reference (6)), which appears instead of the T4-A8 crosspeak observed in reference (6). The proximity of T4 and A8 observed in reference (6) was attributed to a stabilizing effect that the N_loop would have on the helical conformational ensemble sampled by residues 8-18 of full length hIAPP. The absence of the T4-A8 crosspeak in N_loop, where residues 9-37 are missing, is consistent with this interpretation. In summary these data do not indicate significant backbone structural changes between the isolated N_loop and the corresponding residues of full length hIAPP. While this is not generally expected for linear fragments of full length proteins, it is not surprising in the case of short, relatively rigid cyclic peptides such as N_loop.

MOLECULAR DYNAMICS SIMULATIONS

To provide a molecular description of the possible location of the type I β-turn in the N_loop monomer, R. B. Best ran replica exchange molecular dynamics simulations of the N_loop sequence studied here (residues 1-8). The N_loop conformations obtained are similar to those found in our earlier work[11]. Because of the constraint introduced by the linker, the N_loop only adopts a limited number of backbone conformations. This rigidity is evident in the Cα root-mean-square fluctuations of the backbone which are only 1.0 Å at 278 K, and increase to 1.15 Å at the highest REMD temperature of 595 K.

Conformational clustering with a 0.8 Å cut-off yields two major conformations with populations of 64.0 % and 33.3 %, with the remaining conformations having very low population. Secondary structure assignment using STRIDE indicates that the most populated cluster (Figure 4.11a) has a Type I turn between residues 3 and 6 and a Type IV turn between residues 4 and 7, while the second most populated cluster (Figure 4.11b) has a Type I turn from residues 4-7. Thus there is expected to be a large population of Type I turn, consistent with the experimental findings. Type I β-turns are geometrically defined as a stretch of four residues (i,i+3), with alpha carbons of residues i and i+3 within 7 Å and internal residues having (ψ,ϕ) angles: i+1:(-30,-60) i+2:(0,-90) (44). Based on these geometric criteria (Appendix A: Figure 4), it was found that a type I β-turn is present at residues 3-6 most of the time, with a significant fraction of type I turn also present at residues 4-7, in agreement with the analysis of the most populated clusters. The results of the simulations were also validated by comparing against the available NMR data. Calculation of average $^{1}H^{\alpha}$ and $^{1}H^{N}$ chemical shifts using the SPARTA+ shift prediction algorithm results in reasonable agreement with experimental data from several groups, including the ones presented here, considering the uncertainty in the shift prediction (Figure 4.9-10)[25]. It was found that 75% of the predictions lie within the (one standard deviation) error bars for the H^{N} shifts and 63% for the H^{α} shifts. The linear correlation coefficients are 0.84 and 0.1 for the H^{N} and H^{α} shifts respectively. Furthermore, the distances between amide protons measured in the simulation correlate well with the intensities of the NOEs measured for N_loop (Table 4.2 and Appendix A: Figure 2).

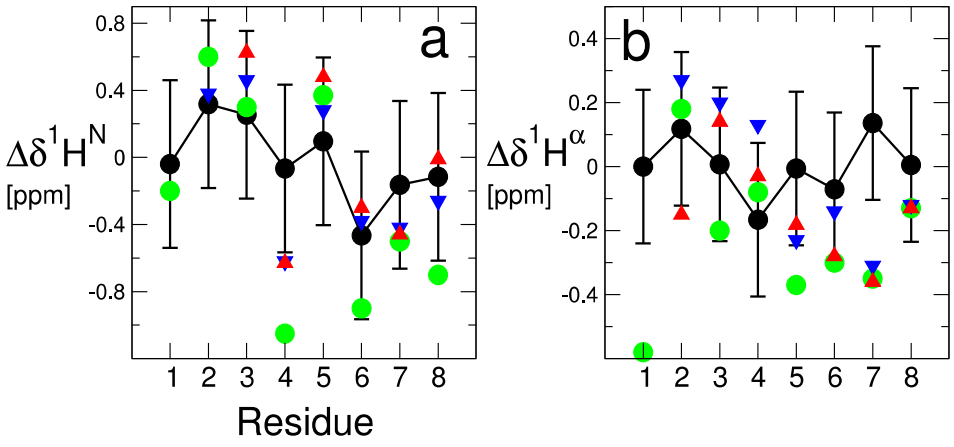

Figure 4.9. Comparison of secondary chemical shifts measured for N_loop (green), using random coil values of reference (13), and values reported in reference (6) and (33), for residues 1-8 of full length hIAPP (red) and rIAPP (blue), compared to computed values from REMD simulations using SPARTA+ (black) are also compared with experimental data sets. Error bars on the calculated shifts represent the R.M.S. error in the shift prediction algorithm.

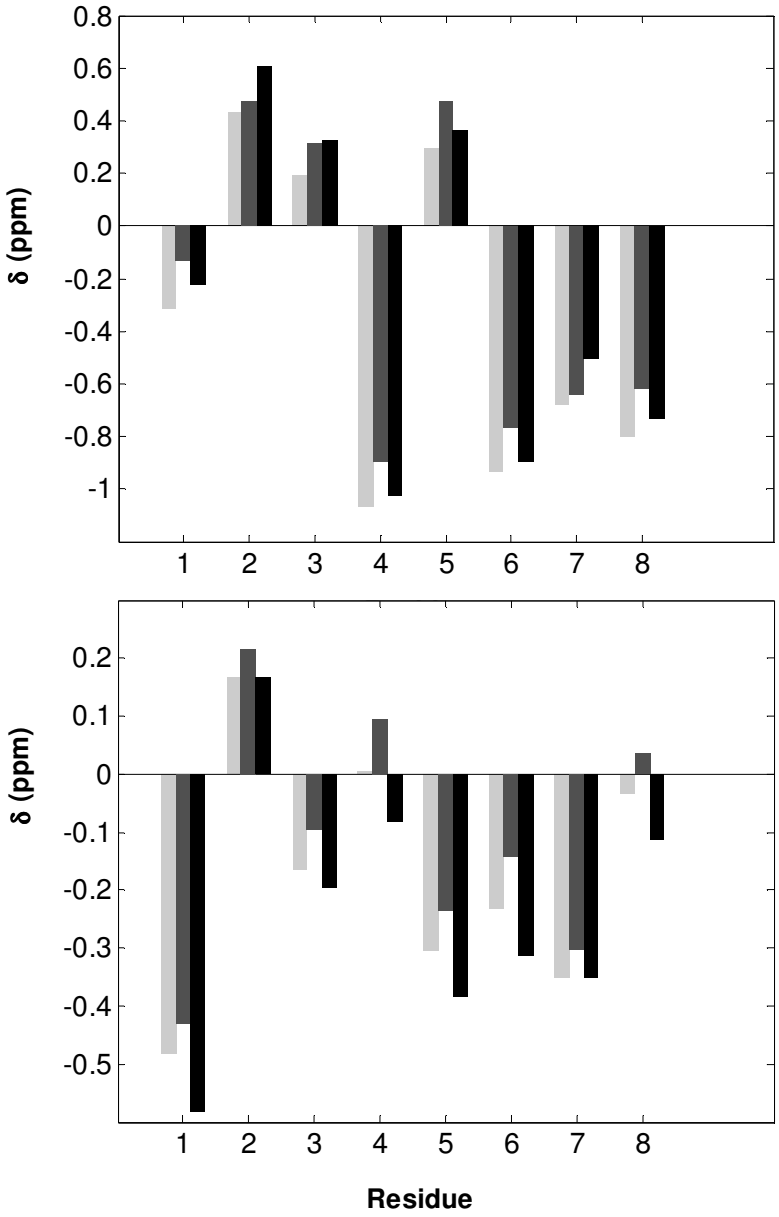

Figure 4.10. $^{1}H^{N}$ *(top) and* $^{1}H^{\alpha}$ *(bottom) secondary chemical shifts for N_loop, calculated using reference random coil chemical shifts from different libraries: Kjaergaard and Poulsen (light gray), Tamiola et al. (dark gray), and De Simone et al. (black).*

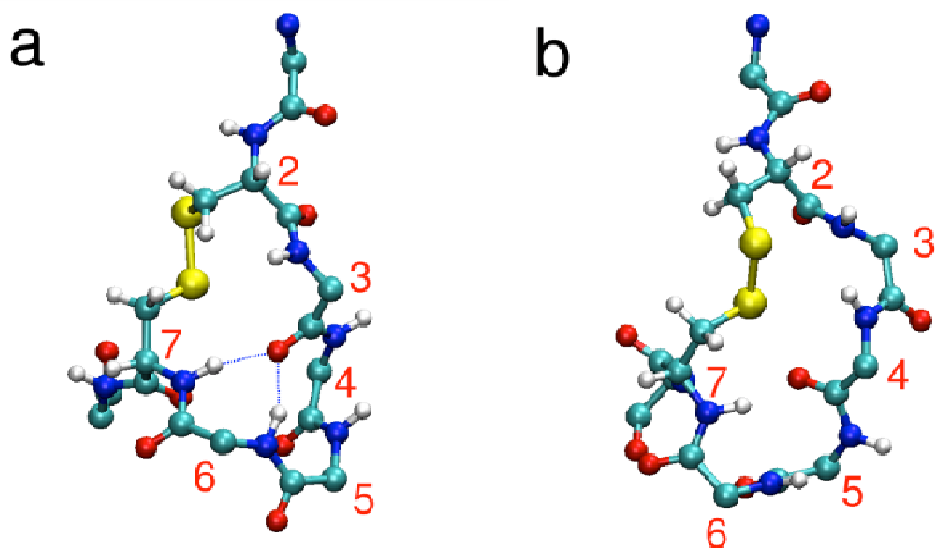

Figure 4.11. Structure of N_loop from R. Best's molecular simulations in CPK representation. (a) Most populated conformation (64.0%) (b) next most populated conformation (33.3%). CPK representation, with type I turns between residues 3-6, and between residues 4-7. (b) Ensemble of structures representing N_loop fluctuations.

In order to investigate how much the structure of the monomer might be affected by self-association, R. Best performed REMD simulations of a pair of monomers at ~100 mM concentration, and analyzed the configurations where the two molecules are in direct contact. The structures of the monomers in the associated configurations were generally very similar to the isolated peptides, with 70% of the monomers having a C^{α} RMSD to the central structure of the largest cluster of the monomer simulations of less than 1.0 Å. Representative clusters of the associated pairs of monomers are given in Appendix A: Figure 5. These results also shed some light on the type of interactions which may drive aggregation – the peptides associate through a variety of polar interactions, but clearly not through a classical amyloid structure. While there are certainly some intermolecular

interactions involving the backbone, because the amide groups are not aligned in the same direction, conventional amyloid structure is not formed.

4.4. DISCUSSION

Our experimental observations, in conjunction with REMD simulations, concur in showing that the fragment (1-8) of hIAPP is constrained into a well-defined, rigid structure containing a Type I β-turn. This structure highly favors the formation of stable inter-peptide interactions resulting in the observed formation of fibers. Cyclization has the effect of significantly decreasing the conformational entropy cost required for fiber formation, making it much more favorable for the N_loop to aggregate compared to the linear peptide, N1-8. This mechanism is similar to that exploited in the design of cyclized, stable β-hairpin peptides that readily self-assemble into highly ordered nanofibers, rodlike crystals and other β-sheet aggregates[25]. To our knowledge though, this is the first time this mechanism is observed for a non-β-sheet fiber, formed by Type I β-turn structures. Molecular simulations suggest that the amide groups of the N_loop are not correctly aligned for formation of conventional amyloid structure, and simulations of a pair of N_loops did not reveal any amyloid-like association. The rigidity of N_loop coupled with the relatively small number of observed modes of interaction, may explain the high propensity to form fibrils.

Our observation of stable N_loop fibers at concentrations as low as 57 μM suggests a direct role of the N_loop in IAPP aggregation. According to structural models of hIAPP amyloid fibers, these fibers are formed by intermolecular beta-sheets involving residues ~25 to 37 (C-terminal sheet) and 8 to ~17 (N terminal sheet), and the N_loop (residues 1-7) is not involved in direct inter-molecular interactions[3]. However, removal of the N_loop alters both the mass per length distributions of hIAPP fibers and the kinetics of fibril formation[4-6]. Based on the observed strong tendency of the N_loop to form inter-peptide interactions, we propose that N_loop-N_loop interactions occurring at the N-terminus of full length IAPP may drive initial association of IAPP, prior to β-sheet formation. This would be consistent with both experimental data on hIAPP amyloid fiber formation and on the disruption of β-sheet interactions in hIAPP fibers in the presence of rIAPP, as explained below.

Initial association of IAPP by N_loop-N_loop interactions could occur at a relatively low entropy cost, unlike the direct association of β-sheets. In sequences that allow extensive β-sheet formation (e.g. in hIAPP and not in rIAPP) such interactions would favor aggregation into ordered fibrils by aligning the peptides at the N-terminus, consistent with accepted models of hIAPP fibrils[3]. This mechanism would be consistent with recent 2D-IR studies indicating that β-sheet structures form in the later stages of hIAPP aggregation, and with p-Cyanophenilalanine fluorescence studies indicating that side chains of residues 15, 23 and 37 remain partially exposed to solvent during the lag phase, until significant β-sheet structure has formed[43,44]. In addition, it would explain the observed changes to nucleation processes in the absence of the intact N_loop [5,6]. We note

that N_loop fibers (i.e. fibers formed by N_loop-N_loop interactions) have not been observed so far at equilibrium in full length hIAPP. Two factors would interfere with the formation of N_loop-based fibers in full length hIAPP: first, intra-molecular interactions between the N_loop and the linear portion of the chain (residues 8-37) will contribute to stabilize the monomer, and compete with N_loop-N_loop interactions. Such intra-molecular interactions have been proposed by Vaiana et al. to explain the large degree of compaction observed in full length hIAPP, rIAPP and in a model hydrophilic sequence that contained the intact N_loop[11]. This model is supported by molecular simulations[11]. Second, the distance between the N_loops in the NMR structures of the fiber suggest that formation of N_loop-based fibers may be in competition with the extensive inter-molecular hydrogen bonding of the amyloid fibrils.

Our proposed model would also be consistent with the ability of rIAPP to disrupt inter-molecular β-sheets in hIAPP amyloid fibers. According to our model in fact direct N_loop-N_loop interactions would also occur in solutions containing mixtures of IAPP variants. Variants of hIAPP containing proline substitutions in the C-terminal region but with identical N-terminal sequences and intact N_loop (e.g. rIAPP, I26P and G24P), have been shown to inhibit hIAPP amyloid fiber formation, implying a direct interaction between hIAPP and the inhibitor[45,46]. Recent results show that interactions between rIAPP and hIAPP initiate in the N-terminal region of the peptides, and disrupt the "native" β-sheet structure of hIAPP amyloid fibers[2]. Our present findings suggest that such interactions could be mediated by N_loop-N_loop association, which could act in

addition or as an alternative to the originally proposed mechanism of aggregation via alpha helical or β-sheet intermediates[2,9,10,47].

Our proposed role of the N_loop in mediating intermolecular interactions of IAPP may in fact be related to the role of this highly conserved sequence in the biological activity of Ct family peptides. While receptor binding affinity is mainly determined by residues in the 8-37 region, the N_loop is required for Ct peptides to activate cell response when binding to their receptors, possibly via interactions between the N_loop and the extramembrane region of the CGRP receptor[48-50]. We propose that the rigid Type I β-turn structure reported here mediates such interactions. This would be consistent with early work showing that modifications of the disulfide bridge topography (by substitution of cysteinyl residues with penicillamine) can greatly affect CGRP receptor activation upon ligand binding[51].

4.5. SUMMARY AND CONCLUSIONS

In conclusion, data reported here demonstrate for the first time that the N_loop of IAPP forms stable fibers containing a type I β-turn structure and not β-sheets. We propose that the rigid β-turn structure plays a role in both the biological activity of Ct family peptides and in the pathological aggregation of IAPP, a possibility that has not yet been explored. We note that N_loop-N_loop interactions could be exploited to develop novel inhibitors of hIAPP amyloid formation.

REFERENCES

_____

[1] Middleton, C. T., P. Marek, P. Cao, C. C. Chiu, S. Singh, A. M. Woys, J. J. de Pablo, D. P. Raleigh, and M. T. Zanni. 2012. Two-dimensional infrared spectroscopy reveals the complex behaviour of an amyloid fibril inhibitor. Nature Chemistry 4:355-360.

[2] Cao, P., F. Meng, A. Abedini, and D. P. Raleigh. 2010. The Ability of Rodent Islet Amyloid Polypeptide To Inhibit Amyloid Formation by Human Islet Amyloid Polypeptide Has Important Implications for the Mechanism of Amyloid Formation and the Design of Inhibitors. Biochemistry 49:872-881.

[3] Luca, S., W. M. Yau, R. Leapman, and R. Tycko. 2007. Peptide conformation and supramolecular organization in amylin fibrils: Constraints from solid-state NMR. Biochemistry 46:13505-13522.

[4] Goldsbury, C., K. Goldie, J. Pellaud, J. Seelig, P. Frey, S. A. Muller, J. Kistler, G. J. S. Cooper, and U. Aebi. 2000. Amyloid fibril formation from full-length and fragments of amylin. Journal of Structural Biology 130:352-362.

[5] Koo, B. W., and A. D. Miranker. 2005. Contribution of the intrinsic disulfide to the assembly mechanism of islet amyloid. Protein Science 14:231-239.

[6] Yonemoto, I. T., G. J. A. Kroon, H. J. Dyson, W. E. Balch, and J. W. Kelly. 2008. Amylin Proprotein Processing Generates Progressively More Amyloidogenic Peptides that Initially Sample the Helical State. Biochemistry 47:9900-9910.

[7] Ferrone, F. A. 2006. Nucleation: The connections between equilibrium and kinetic behavior. Methods in Enzymology 412:285-299.

[8] Eichner, T., and S. E. Radford. 2011. A Diversity of Assembly Mechanisms of a Generic Amyloid Fold. Molecular Cell 43:8-18.

[9] Williamson, J. A., J. P. Loria, and A. D. Miranker. 2009. Helix Stabilization Precedes Aqueous and Bilayer-Catalyzed Fiber Formation in Islet Amyloid Polypeptide. Journal of Molecular Biology 393:383-396.

[10] Abedini, A., and D. P. Raleigh. 2009. A critical assessment of the role of helical intermediates in amyloid formation by natively unfolded proteins and polypeptides. Protein Engineering Design & Selection 22:453-459.

[11] Vaiana, S. M., R. B. Best, W. M. Yau, W. A. Eaton, and J. Hofrichter. 2009. Evidence for a Partially Structured State of the Amylin Monomer. Biophysical Journal 97:2948-2957.

[12] Kjaergaard, M., and F. M. Poulsen. 2012. Disordered proteins studied by chemical shifts. Progress in Nuclear Magnetic Resonance Spectroscopy 60:42-51.

[13] De Simone, A., A. Cavalli, S. T. D. Hsu, W. Vranken, and M. Vendruscolo. 2009. Accurate Random Coil Chemical Shifts from an Analysis of Loop Regions in Native States of Proteins. Journal of the American Chemical Society 131:16332-+.

[14] Tamiola, K., B. Acar, and F. A. A. Mulder. 2010. Sequence-Specific Random Coil Chemical Shifts of Intrinsically Disordered Proteins. Journal of the American Chemical Society 132:18000-18003.

[15] Kjaergaard, M., and F. M. Poulsen. 2011. Sequence correction of random coil chemical shifts: correlation between neighbor correction factors and changes in the Ramachandran distribution. Journal of Biomolecular Nmr 50:157-165.

[16] Nilges, M., G. M. Clore, and A. M. Gronenborn. 1990. H-1-Nmr Stereospecific Assignments by Conformational Database Searches. Biopolymers 29:813-822.

[17] Best, R. B., and G. Hummer. 2009. Optimized Molecular Dynamics Force Fields Applied to the Helix-Coil Transition of Polypeptides. Journal of Physical Chemistry B 113:9004-9015.

[18] Kudva, Y. C., C. Mueske, P. C. Butler, and N. L. Eberhardt. 1998. A novel assay in vitro of human islet amyloid polypeptide amyloidogenesis and effects of insulin secretory vesicle peptides on amyloid formation. Biochemical Journal 331:809-813.

[19] Greenfield, N. J. 2004. Analysis of circular dichroism data. Numerical Computer Methods, Pt D 383:282-317.

[20] Perczel, A., and G. D. Fasman. 1992. Quantitative-Analysis of Cyclic Beta-Turn Models. Protein Science 1:378-395.

[21] Manning, M. C., M. Illangasekare, and R. W. Woody. 1988. Circular-Dichroism Studies of Distorted Alpha-Helices, Twisted Beta-Sheets, and Beta-Turns. Biophysical Chemistry 31:77-86.

[22] Kjaergaard, M., A. B. Norholm, R. Hendus-Altenburger, S. F. Pedersen, F. M. Poulsen, and B. B. Kragelund. 2010. Temperature-dependent structural changes in intrinsically disordered proteins: Formation of alpha-helices or loss of polyproline II? Protein Science 19:1555-1564.

[23] Uversky, V. N. 2009. Intrinsically Disordered Proteins and Their Environment: Effects of Strong Denaturants, Temperature, pH, Counter Ions, Membranes, Binding Partners, Osmolytes, and Macromolecular Crowding. Protein Journal 28:305-325.

[24] Barth, A. 2000. The infrared absorption of amino acid side chains. Progress in Biophysics & Molecular Biology 74:141-173.

[25] Susi, H., Timashef.Sn, and L. Stevens. 1967. Infrared Spectra and Protein Conformations in Aqueous Solutions .I. Amide I Band in H2o and D2o Solutions. Journal of Biological Chemistry 242:5460-&.

[26] Zandomeneghi, G., M. R. H. Krebs, M. G. Mccammon, and M. Fandrich. 2004. FTIR reveals structural differences between native beta-sheet proteins and amyloid fibrils. Protein Science 13:3314-3321.

[27] Nilsson, M. R., and D. P. Raleigh. 1999. Analysis of amylin cleavage products provides new insights into the amyloidogenic region of human amylin. Journal of Molecular Biology 294:1375-1385.

[28] Mukherjee, S., P. Chowdhury, and F. Gai. 2009. Effect of Dehydration on the Aggregation Kinetics of Two Amyloid Peptides. Journal of Physical Chemistry B 113:531-535.

[29] Okabayashi, H., M. Ishida, H. Tamaoki, H. Masuda, and C. J. O'Connor. 2002. Fourier transform IR study of aggregational behavior of N-acetyl-L- and N-butyloxycarbonyl-L-glutamic acid oligomeric benzyl esters in dioxane and benzene: beta-turn -> antiparallel beta-sheet transition. Biopolymers 65:129-141.

[30] Krimm, S., and J. Bandekar. 1986. Vibrational Spectroscopy and Conformation of Peptides, Polypeptides, and Proteins. Advances in Protein Chemistry 38:181-364.

[31] Mantsch, H. H., D. J. Moffatt, and H. L. Casal. 1988. Fourier-Transform Methods for Spectral Resolution Enhancement. Journal of Molecular Structure 173:285-298.

[32] Vass, E., M. Kurz, R. K. Konat, and M. Hollosi. 1998. FTIR and CD spectroscopic studies on cyclic penta- and hexa-peptides. Detailed examination of hydrogen bonding in beta- and gamma-turns determined by NMR. Spectrochimica Acta Part a-Molecular and Biomolecular Spectroscopy 54:773-786.

[33] Williamson, J. A., and A. D. Miranker. 2007. Direct detection of transient alpha-helical states in islet amyloid polypeptide. Protein Science 16:110-117.

[34] Hutchinson, E. G., and J. M. Thornton. 1994. A Revised Set of Potentials for Beta-Turn Formation in Proteins. Protein Science 3:2207-2216.

[35] Shen, Y., and A. Bax. 2010. SPARTA plus : a modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network. Journal of Biomolecular Nmr 48:13-22.

[36] Valery, C., E. Pouget, A. Pandit, J. M. Verbavatz, L. Bordes, I. Boisde, R. Cherif-Cheikh, F. Artzner, and M. Paternostre. 2008. Molecular origin of the self-assembly of lanreotide into nanotubes: A mutational approach. Biophysical Journal 94:1782-1795.

[37] Ishihara, Y., and S. Kimura. 2010. Nanofiber formation of amphiphilic cyclic tri-beta-peptide. Journal of Peptide Science 16:110-114.

[38] Joshi, K. B., and S. Verma. 2006. Ordered self-assembly of a glycine-rich linear and cyclic hexapeptide: Contrasting ultrastructural morphologies of fiber growth. Supramolecular Chemistry 18:405-414.

[39] Clark, T. D., J. M. Buriak, K. Kobayashi, M. P. Isler, D. E. McRee, and M. R. Ghadiri. 1998. Cylindrical beta-sheet peptide assemblies. Journal of the American Chemical Society 120:8949-8962.

[40] Kobayashi, K., J. R. Granja, and M. R. Ghadiri. 1995. Beta-Sheet Peptide Architecture - Measuring the Relative Stability of Parallel Vs Antiparallel Beta-Sheets. Angewandte Chemie-International Edition in English 34:95-98.

[41] Gibbs, A. C., L. H. Kondejewski, W. Gronwald, A. M. Nip, R. S. Hodges, B. D. Sykes, and D. S. Wishart. 1998. Unusual beta-sheet periodicity in small cyclic peptides. Nature Structural Biology 5:284-288.

[42] Khakshoor, O., and J. S. Nowick. 2009. Use of Disulfide "Staples" To Stabilize beta-Sheet Quaternary Structure. Organic Letters 11:3000-3003.

[43] Shim, S. H., R. Gupta, Y. L. Ling, D. B. Strasfeld, D. P. Raleigh, and M. T. Zanni. 2009. Two-dimensional IR spectroscopy and isotope labeling defines the pathway of amyloid formation with residue-specific resolution. Proceedings of the National Academy of Sciences of the United States of America 106:6614-6619.

[44] Marek, P., S. Mukherjee, M. T. Zanni, and D. P. Raleigh. Residue-Specific, Real-Time Characterization of Lag-Phase Species and Fibril Growth During Amyloid Formation: A Combined Fluorescence and IR Study of p-Cyanophenylalanine Analogs of Islet Amyloid Polypeptide. Journal of Molecular Biology 400:878-888.

[45] Meng, F. L., D. P. Raleigh, and A. Abedini. 2010. Combination of Kinetically Selected Inhibitors in Trans Leads to Highly Effective Inhibition of Amyloid Formation. Journal of the American Chemical Society 132:14340-14342.

[46] Abedini, A., F. L. Meng, and D. P. Raleigh. 2007. A single-point mutation converts the highly amyloidogenic human islet amyloid polypeptide into a potent fibrillization inhibitor. Journal of the American Chemical Society 129:11300-+.

[47] Dupuis, N. F., C. Wu, J. E. Shea, and M. T. Bowers. 2011. The Amyloid Formation Mechanism in Human IAPP: Dimers Have beta-Strand Monomer-Monomer Interfaces. Journal of the American Chemical Society 133:7240-7243.

[48] Hay, D. L., G. Christopoulos, A. Christopoulos, D. R. Poyner, and P. M. Sexton. 2005. Pharmacological discrimination of calcitonin receptor: receptor activity-modifying protein complexes. Molecular Pharmacology 67:1655-1665.

[49] Poyner, D. 1995. Pharmacology of receptors for calcitonin gene-related peptide and amylin. Trends in Pharmacological Sciences 16:424-428.

[50] Conner, A. C., J. Simms, J. Barwell, M. Wheatley, and D. R. Poyner. 2007. Ligand binding and activation of the CGRP receptor. Biochemical Society Transactions 35:729-732.

[51] Saha, S., D. J. J. Waugh, P. Zhao, P. W. Abel, and D. D. Smith. 1998. Role of conformational constraints of position 7 of the disulphide bridge of h-alpha-CGRP derivatives in their agonist versus antagonist properties. Journal of Peptide Research 52:112-120.

CHAPTER 5

EFFECT OF CHAIN STIFFNESS ON THE CONFORMATION AND DYNAMICS OF MONOMERIC IAPP

5.1. INTRODUCTION

Only primates and cats are known to form islet amyloids derived from IAPP. As discussed in Chapter 2, several mammalian IAPP do not form amyloid aggregates, even though they are largely homologous to human IAPP (hIAPP) in both sequence and structure. A naturally occurring variant of hIAPP, rat IAPP (rIAPP) garners much attention in IAPP aggregation studies. In fact, rIAPP does not form amyloid fibrils, even in vitro, despite its 84% sequence homology. Thus far, rIAPP's remarkable solubility has been attributed to the three proline residues present in the disordered region of the C-terminus: P25, P28 and P29 (Figure 5.1). This unique property of rIAPP could be responsible for its inability to form amyloid fibrils. Proline residues are known to break β-sheet structures and to energetically favor disordered and turn states[1]. Furthermore, based on experimental work and simulations, several groups have proposed that IAPP aggregation may occur through a pathway which includes a stable dimer intermediate[2,3,4,5]. These dimers would form through N-terminus self-association. If this hypothesis is correct, then the presence of a proline mutant in the C-terminal region would disfavor the transition into the β-sheet structure characteristic of protofibers. The Raleigh group has succeeded in developing novel inhibitory analogs for IAPP through the addition of proline residues. They have found that single point mutations at sites I26,

G24 and A13 inhibit hIAPP aggregation to varying degrees[678]. Furthermore, these mutants form amyloid fibers on much slower timescales[6-9]. It appears thought that these single site mutations mainly impact the kinetics of aggregation. Green et al. have in fact shown that substituting any of the 3 non-proline residues in rIAPP with the corresponding residues from hIAPP causes it to form amyloid aggregates, but on much longer timescales than hIAPP. Most recently, Amylin Pharmaceuticals has developed Pramlintide, an hIAPP analog that contains all three of rIAPP's proline substitutions. Pramlintide is the only FDA-approved, non-insulin drug found effective in both type 1 and type 2 diabetes[1011]. In addition, Exenatide, a drug prescribed for persons with type 2 diabetes, contains 4 proline residues in its C-terminus[12]. It is evident that prolines have a significant effect on the kinetics of aggregation, but the manner in which they mechanistically do so remains unknown. While most studies have been focusing on the possible role of prolines in disrupting the amyloid fiber structure, very little is known about their effect on the monomeric state of IAPP. To truly grasp the role of prolines in aggregation, one must understand how they alter the energetic landscape of the monomer state.

The goal of this work is to understand how the presence of proline and non-proline mutations in rIAPP affects monomer structure and dynamics, and to indicate what implications this could have on aggregation propensity. In order to better grasp the impact that proline substitutions may have on a polypeptide chain, we first present a survey of the existing literature. The simplest view of the contribution of prolines to the conformational space occupied by a given protein is to consider Flory's isolated-pair

hypothesis[13]. Proline residues may drastically reduce the number of microstates a disordered protein can populate because of the unique ϕ-φ angles that this residue can populate. Based on this simple isolated-pair hypothesis, such effects should only be local and should not extend beyond nearest-neighbors. For long enough polymers, prolines should therefore not affect the accessible conformational space. But, as shown by Pappu and Rose, such simplifications never accurately depict reality. Even in the case of polyalanyl chains, which have relatively small excluded volumes, computer simulations showed that local steric effects extend beyond nearest-chain neighbors and significantly alter the accessible conformational space[14]. In the case of prolines, these effects can be long-range[15]. Several groups have attempted to quantify these long-range effects with varying success. Keifhaber and coworkers found that introducing a proline into a linear polymer containing 10 or less amino acids substantially increased the rate of loop formation, i.e. the rate of forming a closed contact between two ends of the polypeptide chain[16]. They found that by preferentially introducing proline residues in the cis conformation, the rate of observed loop formation increased. They interpreted this increase in rates to mean shorter end-to-end distances, suggesting that the cis conformation causes short peptides to populate more compact states. This was also supported by computer simulations. Interesting, they saw no difference in observed rates for loops larger than 10 residues. By contract, recent experimental work measuring histidine-heme loop formation showed that introducing a single proline residue to a 22-residue protein in denaturant (3M GdmCl) can increase the Flory characteristic rate of a peptide by 10-15%, indicating long-range expansion of the protein[17]. This result was used to support the claim that prolines promote residual structure in the denatured state

ensembles (DSE) of functional proteins. How prolines affect overall conformation sampling for IDPs remains elusive, and evidence suggests that it is sequence-specific.

As discussed above, understanding the role of prolines in IAPP has been studied mainly in the context of aggregation. Though much work has focused on developing proline mutants to design inhibitors and possible therapeutics, little is known about how the monomeric conformations of IAPP change with the addition of proline residues. Most of the work done so far on the monomeric state of IAPP with proline substitutions has been computational. Recent simulations have suggested that the presence of prolines changes the ensemble of populated states to include less β-sheet conformers[18]. Experimentally, Vaiana et al.[19] reported that rIAPP occupies a more expanded conformation than hIAPP in denaturant (6M GdmCl), which was attributed to an increase in stiffness, most likely due to the presence of prolines. We aim to understand if proline residues in IAPP are responsible for altering the conformational sampling of IAPP in solution. To probe the intrinsic polymeric properties of IAPP in the absence of sequence-specific, water-induced interactions, we consider proline mutants in 6M GdmCl. 6M GdmCl is a "good" solvent where we can assume that the only interactions affecting chain dimensions are excluded volume and intrinsic chain stiffness. This is supported by experimental findings on the denatured states of proteins[22,23] and model peptide sequences[24], which show that in 6M GdmCl they behave like worm-like chains with excluded volume. It should be noted that in polymer terms, intrinsic stiffness, usually represented by a given persistence length, encapsulates all effects due to both the presence of local structural preferences, and therefore of local residual structure in the denatured state. This work, in 6M GdmCl,

should be considered a first step in understanding the effect of prolines on IAPP conformations and dynamics, as it should later be extended to aqueous solution conditions.

To understand the structure and dynamics of IAPP mutants in good solvent, we measure end-to-end contact formation rates of the monomer. As described in Chapter 3, these measurements give us information on both P(r), the end-to-end distance distribution of the polypeptide chain, and D, the intrachain diffusion coefficient. Using our homebuilt nanosecond laser spectrometer (Figure 3.5), we performed these measurements for the following IAPP mutants, and compared them to data previously reported for hIAPP. For all peptides in this study, the last residue (Y37W) was substituted with a tryptophan (W37) and the C-terminus was amidated, as in previous studies[19].

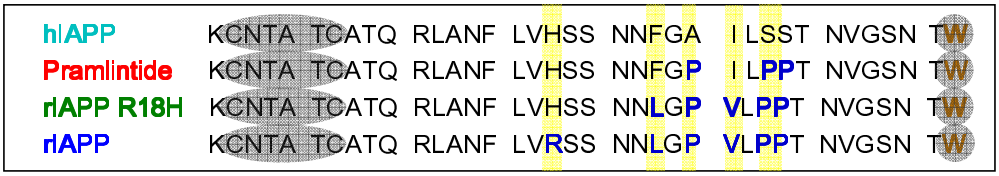

Figure 5.1 IAPP mutants used to study the effects of proline mutants on IAPP structure and dynamics. Our experiments measure the rate of contact formation between the C-terminal tryptophan (W37, gray circle) and the N-terminal disulfide bond (C2-C7, grey oval).

As mentioned above, rIAPP differs from hIAPP by 6 amino acids, three of which are proline residues. We set out to understand if the expanded conformation seen for rIAPP by Vaiana et al.[19] was due solely to the proline residues. To do this, we first studied end-to-end contact formation rates in Pramlintide, an hIAPP analog that contains all three of the proline residues found in rIAPP (Figure 5.1).

5.2. METHODS

MATERIALS

Fmoc(9-fluorenylmethoxycarbonyl)-protected amino acids were purchased from Novabiochem. HOBt (N-hydroxy benzotriazole) and HBTU (O-Benzotriazole-N, N, N', N'-tetramethyl-uronium-hexafluoro-phosphate) were purchased from Genscript. N,N-iisopropylethylamine, or Hünig's base (DIPEA), and N-Methyl-2-pyrrolidone (NMP), used as base in solid phase peptide synthesis, were purchased from Sigma-Aldrich. Piperidine (Sigma-Aldrich) was used for deprotection. Rink Amide-ChemMatrix® was purchased from Matrix innovations. Dimethyl formamide (DMF), Dichloromethane (DCM) and Acetonitrile were purchased from Fisher Scientific and were used without further purification.

PEPTIDE SYNTHESIS AND PURIFICATION

rIAPP Y37W was synthesized on a CEM Liberty Automated Microwave Peptide Synthesizer using PALChem Matrix resin. After synthesis, the peptides were thoroughly washed with DMF followed by DCM. For deprotection, the peptides were shaken for one hour in 20% piperdine, 0.1M HOBt in DMF. The Cleavage cocktail consisted of 94% Trifluoroacetic acid (TFA) + 2.5% Water + 2.5% 1,2-Ethanedithiol + 1% Triisopropylsilane (TIS) at the ratio of 150 µL/ 10 mgs of resin. rIAPP Y37W was purified using Reverse Phase High performance liquid chromatography (HPLC) on a Waters 600E system. Crude peptide was first purified on a C4 column at a gradient of 20-50% Acetonitrile with 0.1% TFA over 30 minutes. All the peaks were analyzed by a

Voyager Systems 4320 (Applied Biosystems) matrix assisted laser desorption/ionization-time of flight mass spectrometer (MALDI-TOF MS).

DISULFIDE FORMATION AND OXIDIZED PEPTIDE PURIFICATION

1.3 mM of lyophilized peptide was dissolved in 30% DMSO and 3% Acetic Acid. Sample was stirred with a magnetic stir bar at 1100 rpm. During this time, the formation of the intramolecular disulfide bonds was monitored via a C4 analytical column. The reaction was deemed complete when the reduced peptide's HPLC peak was no longer visible, about 10 hours. After this time, the sample was immediately dissolved in water for a final peptide concentration of 100uM, frozen, and lyophilized. Oxidized peptide was re-purified by analytical HPLC, using a reverse phase C4 column (Length 250mm × ID 4.6mm) particle size 5µm using the same gradient conditions with 0.9 mL/min flow rate. A single peak eluting at a gradient corresponding to the hydrophobicity of rIAPP Y37W was collected, frozen, lyophilized and analyzed before being used in characterization experiments. Peak integration deemed purity to be >99%. Fractions were immediately frozen in liquid nitrogen and lyophilized.

SAMPLE PREPARATION

Pure, lyophilized IAPP mutants were purchased by GenScript (Piscataway, NJ) or synthesized and purified as described above. To probe the end-to-end contact formation rates in rIAPP, rIAPP R18H and Pramlintide, we mutated residue 37 from a tyrosine to a tryptophan. This substitution provides us with a probe that is nonintrusive and has been shown not to alter the aggregation properties of hIAPP[19].

Hours before each experiment, peptides were dissolved in 50mM NaAc, 6M GdmCl, pH=4.9. Samples were deoxygenated with N_2O before measurements.

END-TO-END CONTACT FORMATION MEASUREMENTS AND ANALYSIS

End-to-end contact formation rates were measured by following the protocol described in Chapter 2. For each sample, measurements were performed from 5° to 50° C. Data was subsequently fit to single exponential decays. Because of photoproducts present at longer times, the data was additionally fit to a linear function for times > 5e-5 seconds. Viscosity dependent measurements were performed by preparing buffers at varying sucrose concentrations. Reaction-limited rates were fit to Thirumalai's distribution function for a worm-like chain model (Equation 1). Since the reaction-limited rate also depends on the distance dependence of the quenching rate, we needed to determine the value of q(r) for our quencher, N_loop. We assumed that the effective polymer length of IAPP was N=31, due to the rigidity between Cys2-Cys7. Using Ref 23, we found the $<R^2>$ value for a 31mer in 6M GdmCl to be 12Å. This model peptide was measured to have a reaction-limited rate of 0.79 μsec-1. We assumed this to be the reaction-limited rate for hIAPP in 6M GdmCl if the quencher was solely cysteine, and not N_loop. We assumed the same exponential behavior of q(r) as previously reported, and assumed a rescaling factor[20]. We found that the reaction-limited rates were 0.18 of the original q(r), accounting for the steric effects of N_loop. This allowed us to model the P(r) for each reaction-limited rate, and determine a persistence length (Table 5.2).

5.3. RESULTS

5.3.1. OBSERVED RELAXATION RATES FOR PROLINE MUTANTS

To quantify the effect of the proline residues on the end-to-end distance of IAPP, we measured time resolved triplet-triplet absorption of each peptide as a function of temperature. Figure 5.2 shows representative data for Pramlintide at varying temperatures. These data were fit to an exponential decay, corresponding to the triplet state decay, and a smaller slowly varying decay which corresponds to a photoproduct[19,20]. This slower decay was fit to an empirical function which varied linearly with log(t) for times greater than 4e-5 s (as described in Chapter 3 and Ref. 19). The exponential decay, corresponding to the triplet state decay, has a characteristic time, τ_{OBS}, which is equal to the inverse of the observed quenching rate, k_{OBS}.

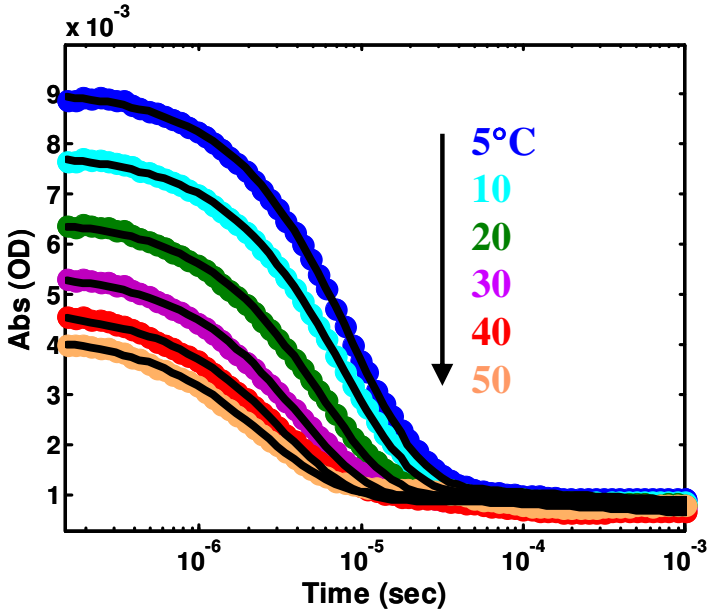

Figure 5.3 Time resolved triplet-triplet absorbance after nanosecond UV excitation of Pramlintide in 6M GdmCl, pH=4.9 at varying temperatures (see legend). Data were individually fit as described above. The exponential decay, corresponding to the triplet state lifetime, has a characteristic time, $\tau_{OBS}=1/k_{OBS}$.

We first directly compared the relaxation rates for each peptide. Figure 5.4 is an Arrhenius plot of k_{OBS} as a function of temperature.

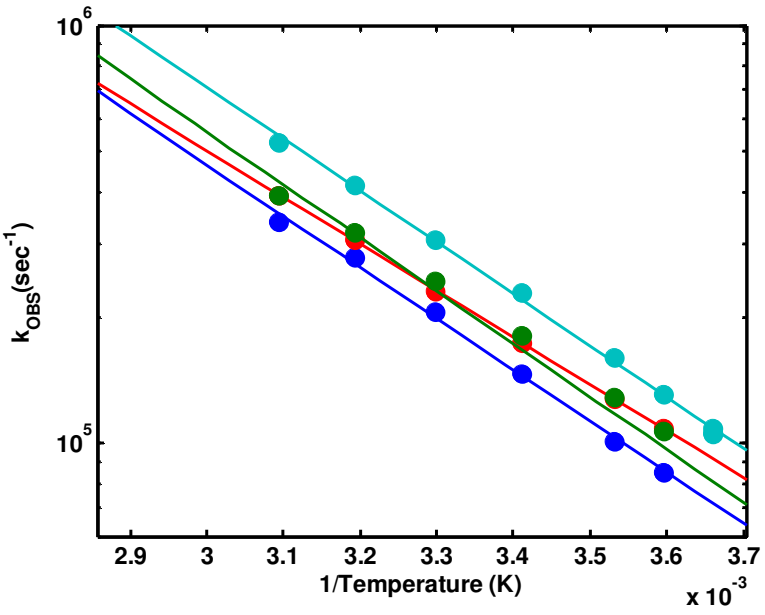

Figure 5.2 Arrhenius plots of the observed relaxation rates, $k_{OBS}=1/\tau_{OBS}$, for Pamlintide (red), rIAPP (blue), and the mutant rIAPP R18H, plotted with previously published data for hIAPP (cyan) from Ref 19[a].

As was seen by Vaiana et al.[19], rIAPP (blue) has slower relaxation rates compared to hIAPP (cyan), indicating a larger end-to-end distance, which for simplicity we will describe as more expanded. Our data for Pramlintide (red) clearly shows that the addition of 3 prolines causes the rates to decrease. This indicates that indeed the three prolines induce more expanded conformations in IAPP. However, Figure 5.2 shows that the 3 prolines alone are not sufficient to explain the entire effect seen in rIAPP. This indicates that the remaining three non-proline mutations of rIAPP (I26, F23, R18) also contribute to the chain expansion. Because in 6M GdmCl we expect the main contribution to chain stiffness to be excluded volume, we calculated the van der Waals volume for the three amino acids to determine which one would most likely yield an increase in chain dimensions. Both the I26V and F23L substitutions account for a decrease in the van der

Waals volume of 15% and 8%, respectively. We therefore first looked at the H18R substitution, which results in a 30% increase in the van der Waals volume, as calculated from the Darby and Creighton data[21]. Our data for rIAPP R18H, the peptide containing L23 and V26 in addition to the 3 prolines, but lacking the R18 of rIAPP (Figure 5.2, green) show that the observed end-to-end contact rates are identical to that of Pramlintide (red). These unexpected results indicate that the single point mutation R18H, in the presence of the three prolines and remaining two mutations L23 and V26, causes the peptide to populate conformations with larger end-to-end distances in 6M GdmCl, highly denaturing conditions.

As mentioned previously, differences in observed relaxation rates can reflect differences in the structure and/or dynamics of a peptide. To fully understand why rIAPP has slower end-to-end contact formation rates than rIAPP R18H and Pramlintide, we performed a full analysis on the viscosity-dependent data to find the reaction- and diffusion-limited rates.

5.3.2. SEPARATING STRUCTURE AND DYNAMICS

An advantage of our technique over similar techniques that are diffusion-limited[16] is that we are able to distinguish whether large end-to-end contact formation rates are actually due to a shorter end-to-end distance (P(r)) or due to faster intra-chain diffusion (D). To do this and quantify our observations in terms of an effective persistence length, we performed viscosity- and temperature-dependent experiments and analysis. The

relaxation rates for both rIAPP R18H and Pramlintide are approximately 25% faster than rIAPP in denaturing conditions. As described in Chapter 3, the observed relaxation rate depends on both the reaction-limited rate and the diffusion-limited rate. Therefore, an increase in observed relaxation rates can be due to (i) an overall decrease in the average end-to-end distance or (ii) an increase in the intrachain diffusion of the peptide. Because cystine is not a diffusion-limited quencher, our technique allows us to experimentally separate these two effects. In order to do so, we measure k_{OBS} as a function of solvent viscosity. By measuring triplet-triplet absorbance for sucrose concentrations ranging from 0% to 32% w/v we are able to separate effects that solely depend on changes in the reaction-limited rate, k_R, from the diffusion-limited rate, k_{D+}. Recall from Chapter 3 that k_R depends solely on the structural properties of the peptide, such as P(r), where k_{D+} also depends on the intrachain diffusion. Figure 5.3 shows the viscosity dependence of the observed relaxation times for each peptide, from 5° to 50° C.

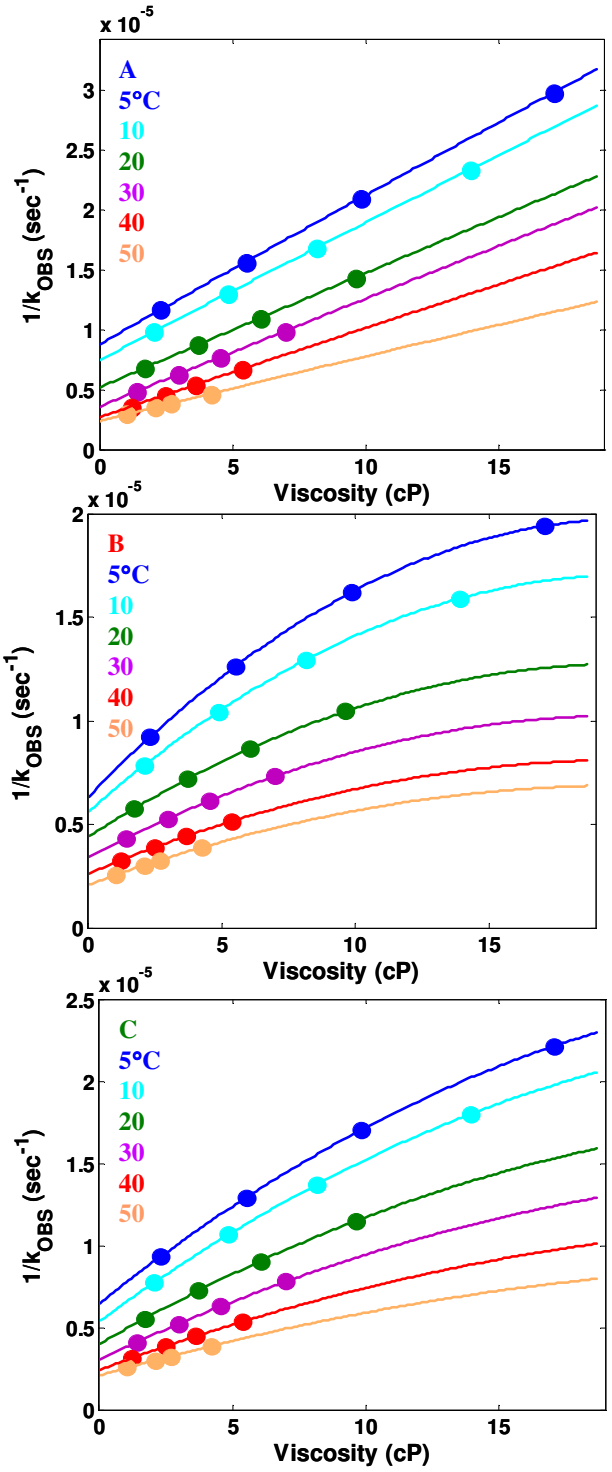

Figure 5.3 Viscosity dependence of the observed end-to-end contact rates for rIAPP(A), Pramlintide(B) and rIAPP R18H (C). The reaction-limited and diffusion-limited rates can be obtained by a global fitting this data as described in Chapter 3.

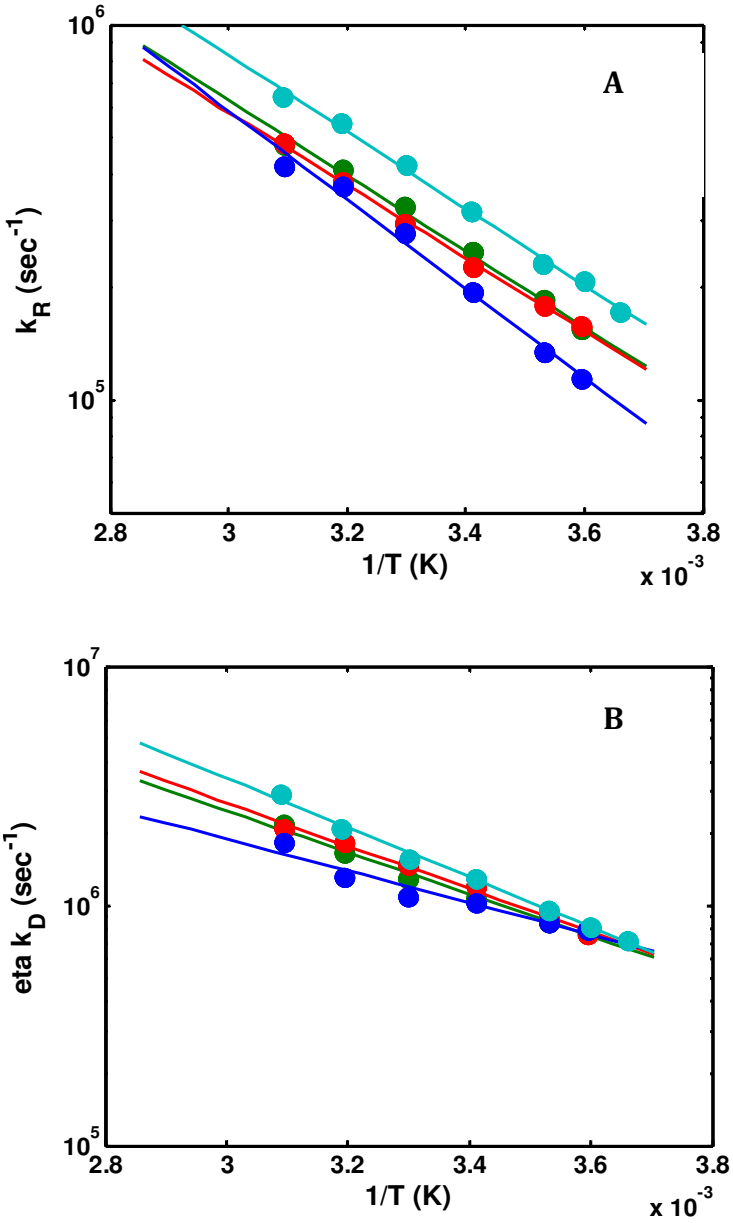

Figure 5.4 Arrhenius behavior of the diffusion-limited (A) and reaction-limited (B) rates for Pramlintide (red), rIAPP R18H (green), rIAPP (blue), and previously published data for hIAPP (cyan) [a].

Below are the tabulated results reported for 20° C measurements, obtained from Figure 5.4:

Peptide	k_{OBS} (μsec^{-1})	k_R (μsec^{-1})	ηk_D (μsec^{-1})
hIAPP	0.23 ± 0.02	0.32 ± 0.03	1.3 ± 0.2
Pramlintide	0.18 ±.02	0.23 ±.02	1.15 ±.2
rIAPP R18H	0.18 ±.02	0.24 ±.02	1.10 ±.2
rIAPP	0.15 ±.01	0.19 ±.02	1.02 ±.2

Table 5.1 Observed, reaction-limited and diffusion-limited rates for all peptides at 20° C. Pramlintide and rIAPP R18H exhibit very similar rates, while rIAPP is consistently slower. hIAPP rates are taken from Ref 19[a].

The data in Table 5.1 and Figure 5.4a show that reaction-limited rates for Pramlintide and rIAPP R18H are approximately 25% faster than rIAPP. This confirms our conclusions that, on average, rIAPP populates a more expanded conformation. Figure 5.4b shows that the diffusion-limited rates, k_{D+} (corrected for intrinsic viscosity dependence), simply reflect the change in k_R, and thus a change in the equilibrium end-to-end distribution. To quantify how much more expanded rIAPP is in denaturant, we can assume a simple polymer model. In 6M GdmCl, it has been shown that a worm-like chain with excluded volume is a good model for the dimensions of disordered proteins. This has been experimentally validated for the unfolded states of a host of naively folded proteins, and has also been verified for the TTQ technique using model peptides and IDPs[22,23,24,25].

Below is the equilibrium end-to-end distance probability distribution for all IAPP mutants assuming Thirumalai's distribution function for a worm-like chain model[26]:

$$P(r) = \frac{4\pi r^2 N}{L_c^2\left[1-(r/L_c)^2\right]^{9/2}} \exp\left(\frac{-3L_c}{4\xi\left[1-(r/L_c)^2\right]}\right) \quad , (1)$$

which has been shown to be a good model for unfolded proteins in denaturant.

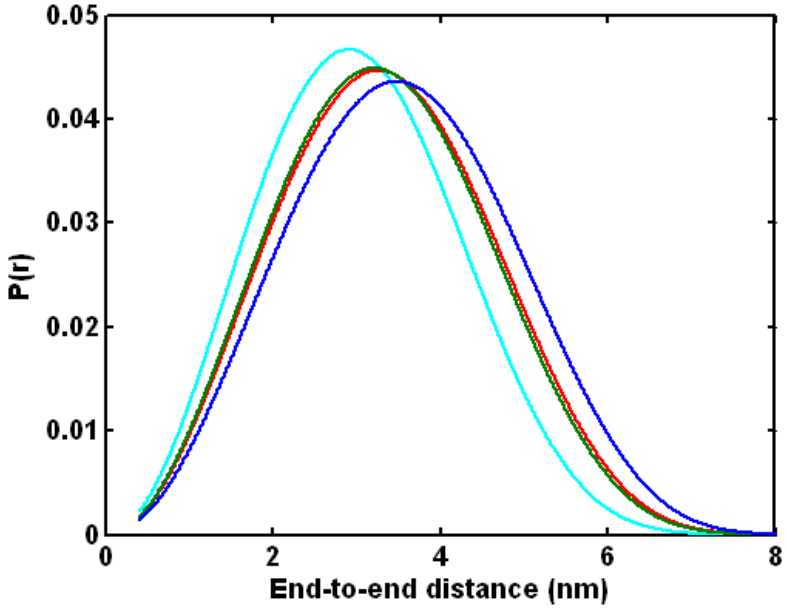

Figure 5.5 The equilibrium end-to-end distance probability distribution for hIAPP(cyan), Pramlintide (red), rIAPP R18H (green), and rIAPP (blue) assuming a worm-like chain distribution for a polymer of length N=31. Details on the analysis can be found in methods.

Peptide	ξ (Å)
hIAPP	4.9
Pramlintide	5.9
rIAPP R18H	5.8
rIAPP	6.6

Table 5.2 Persistence lengths for IAPP mutants assuming the distribution function in Equation 1.

We therefore conclude that the main contribution from amino acid substitutions is to change the equilibrium end-to-end distance distribution, P(r), and not significantly alter the intrinsic dynamics of the peptide, that is, the intrachain diffusion coefficient, D.

5.4. DISCUSSION

These results clearly indicate that the presence of 3 proline mutations significantly alters the P(r) of IAPP in 6M GdmCl, causing it to populate significantly more expanded states. Interestingly, however, the 3 prolines are not the only factor causing rIAPP to expand relative to hIAPP. Further substituting L23F and I26V in Pramlintide appear to have no noticeable effects on either the structure of dynamics of the peptide. However, substituting residue 18 from histidine to arginine (H18R), in the presence of the three proline and L23 and V26 mutations, accounts for a 25% increase in the reaction-limited rate.

It is well known from the literature that the R18H mutant has significant implications for aggregation. Green et al. showed that this single mutation is enough to render rIAPP capable of forming amyloid aggregates[27]. Furthermore, Forman-Kay has shown that the presence of histidine is correlated with a smaller hydrodynamic radius, r_H, for IDPs[28]. It is also well known that polyHis tags in PFG-NMR significantly alter the dimensions of IDPs[28]. The most obvious effect of the R18H mutation is the pH-dependence. Histidine's side chain titrates near physiological conditions (the pK_a is highly dependent on the environment), while arginine is positively charged in neutral, acidic and even most basic environments. pH induced changes are extremely relevant to IAPP as it is found in solution conditions ranging from pH=5.5 to pH=7.4 in vivo. While the in vivo solubility of hIAPP has been attributed to its interactions with insulin, in vitro hIAPP aggregation is considerably slower at pH=5 compared to pH=8. Charge has been shown to affect the monomer conformations of IDPs[25], and the Vaiana group has recently shown that this is the case for CGRP, a member of the Ct family which shares a high sequence homology with IAPP[29]. We note, however, that our data indicate an effect by the R18H substitution at pH=4.9 in 6M GdmCl, an ionic denaturant which is expected to completely screen charges. Therefore, our observed differences cannot be attributed to differences in charge between R and H, indicating that intra-chain interactions lead to a more collapsed state. In 6M GdmCl, this would indicate that rIAPP R18H is not devoid of short-range interactions, which may arise from residual secondary structure While end-to-end distances of peptides in denaturant have been shown to scale as a polymer with excluded volume, data show that this does not contradict the presence of local structure[30,31,32,33,34].

5.5. CONCLUSIONS

Our present findings show that three proline substitutions in hIAPP not only affect its ability to form amyloid fibrils, but also directly affect the monomer conformation by increasing chain stiffness. This chain stiffness can be equated to aggregation propensity, as IAPP must populate a short end-to-end distance on its pathway to amyloid formation.

Future work is needed to determine whether or not the single-point mutation R18H is responsible for the increase in end-to-end distances. This chain expansion could be due to a non-additive effect between R18H and the other non-proline substitutions. To fully understand why sequence specificities are affecting the structure and dynamics of IAPP, we need to further measure the end-to-end contact formation rates for all combinations of the non-proline mutations.

[a] Reprinted from Biophysical Society, Vol 97, Sara M. Vaiana,Robert B. Best,Wai-Ming Yau,William A. Eaton,James Hofrichter. Evidence for a Partially Structured State of the Amylin Monomer, Pages No. 2948-2957, Copyright 2009, with permission from Elsevier. License number 3243820201229.

REFERENCES

[1] Hayashi, T., Asai, T., and Ogoshi, H.1997. Conformational analysis of β-turn structure in tetrapeptides containing proline or proline analogs. Tetrahedron Letters. 38:3039–3042.

[2] Abedini, A. and Raleigh. D. 2009. A Role For Helical Intermediates in Amyloid Formation By Natively Unfolded Polypeptides? Phys. Biol. 6:015005.

[3] Abedini, A. and Raleigh. D. 2009. A critical assessment of the role of helical intermediates in amyloid formation by natively unfolded proteins and polypeptides. Protein Eng Des Sel. 22:453-459

[4] Wiltzius, J.J., Sievers, S.A., Sawaya, M.R., and Eisenberg, D. 2009. Atomic structures of IAPP (amylin) fusions suggest a mechanism for fibrillation and the role of insulin in the process. Protein Sci. 18:1521-30

[5] Dupuis, N. F., C. Wu, J. E. Shea, and M. T. Bowers. 2011. The Amyloid Formation Mechanism in Human IAPP: Dimers Have beta-Strand Monomer-Monomer Interfaces. Journal of the American Chemical Society 133:7240-7243.

[6] Abedini, A., Meng, F., and Raleigh, D.P. 2007. A single-point mutation converts the highly amyloidogenic human islet amyloid polypeptide into a potent fibrillization inhibitor. J Am Chem Soc 129: 11300–11301.

[7] Meng, F., Abedini, A., and Raleigh, D.P. 2010. The Combination of Kinetically Selected Inhibitors in Trans Leads to the Highly Effective Inhibition of Amyloid Formation. J Am Chem Soc 132: 14340-14342.

[8] Cao, P., Meng, F., and Raleigh, D.P. 2010. The Ability of Rodent Islet Amyloid Polypeptide to Inhibit Amyloid Formation by Human Islet Amyloid Polypeptide Has Important Implications For the Mechanism of Amyloid Formation and the Design of Inhibitors. Biochemistry 49:872-881.

[9] Cope, S.M., Shinde, S., Ghirlanda, G. and Vaiana S.M. 2012. Effect of a Single-Point Mutation on the Conformation and Dynamics of Islet Amyloid Polypeptide from Nanosecond-Resolved Intramolecular Contact Formation. Biophy Soc Mtg 3206-Pos

[10] Ryan, G.J., Jobe, L.J., and Martin, R. 2005. Pramlintide in the treatment of type 1 and type 2 diabetes mellitus. Clinical therapeutics. 27: 1500–12.

[11] Edelman, S., Maier, H., amd Wilhelm, K. 2008. Pramlintide in the Treatment of Diabetes Mellitus". BioDrugs. 22: 375–386.

[12] Jones, M.C. 2007. Therapies for diabetes: pramlintide and exenatide. American Family Physician. 75: 1831–5.

[13] Flory, P. J. 1969. Statistical Mechanics of Chain Molecules. Wiley, New York. p252.

[14] Pappu, R.V., Srinivasan, R., and Rose, G.D. 2000. The Flory isolated-pair hypothesis is not valid for polypeptide chains: implications for protein folding.
Proc Natl Acad Sci. 97:12565-70.

[15] Wedemeyer, W.J., Welker, E., Scheraga, H.A. 2001. Proline Cis–Trans Isomerization and Protein Folding. Biochemistry. 41:14637–14644.

[16] Krieger, F., Moglich, A., and Kiefhaber, T. 2005. Effect of proline and glycine residues on dynamics and barriers of loop formation in polypeptide chains. Keifhaber. J Am Chem Soc. 127:2246-3352.

[17] Finnegan, M.L. and Bowler, B.E. 2010. Propensities of Aromatic Amino Acids versus Leucine and Proline to Induce Residual Structure in the Denatured-State Ensemble of Iso-1-cytochrome c. Journal of Molecular Biology. 403:495–504.

[18] Chiu, C., Singh, S., and de Pablo, J. J. 2013. Effect of Proline Mutations on the Monomer Conformations of Amylin. Biophys J. 105: 1227-1235.

[19] Vaiana, S. M., R. B. Best, W. M. Yau, W. A. Eaton, and J. Hofrichter. 2009. Evidence for a Partially Structured State of the Amylin Monomer. Biophysical Journal 97:2948-2957.

[20] Lapidus, L. J., Eaton, W. A. and Hofrichter, J. 2001. Dynamics of intramolecular contact formation in polypeptides: distance dependence of quenching rates in a room-temperature glass. Phys. Rev. Letters. 87:258101–258104.

[21] Darby, N.J. and Creighton, T.E. 1993. Protein Structure. Oxford University Press, New York.

[22] Kohn J.E., Millett, I.S., Jacob, J., Zagrovic, B., Dillon, T.M., Cingel, N., Dothanger, R.S., Seifert, S., Thiyagarajan, P., Sosnick, T.R., Hasan, M.Z., Pande, V.S., ruczinski, I., Doniach, S., Plaxco, K.W. 2004. Random-coil behavior and the dimensions of chemically unfolded proteins. Proc Natl Acad Sci. 101:12491–12496.

[23] Tanford, C. 1970. Protein denaturation. Adv Protein Chem 24:1–95.

[24] Buscaglia, M., Lapidus, L., W. A. Eaton, and J. Hofrichter. Effects of Denaturants on the Dynamics of Loop Formation in Polypeptides. Biophys J. 91: 276–288.

[25] Müller-Späth, S., Soranno, A., Hirschfeld, V., Hofmann, H., Rüegger, S., Reymond, L., Nettels, D., and Schuler, B. 2010. Charge interactions can dominate the dimensions of intrinsically disordered proteins. Proc Natl Acad Sci. 107:14609–14614.

[26] Ha, B.Y. and Thirumalai, D. 1997. Semiflexible Chains Under Tension. J. Chem. Phys. 106:4243-4247.

[27] Green, J., Goldsbury, C., Mini, T., Sunderji, S., Frey, P., Kistler, J., Cooper, G. and Aebi, U. 2003. Full-length rat amylin forms fibrils following substitution of single residues from human amylin. J Mol Biol. 326:1147-1156.

[28] Marsh, J.A. and Forman-Kay, J.D. 2010. Sequence Determinants of Compaction in Intrinsically Disordered Proteins. Biophys. J. 98:2383-2390.

[29] Sizemore, S., Cope, S.M., Roy, A., Ghirlanda, G., Vaiana, S.M. Charge modulates degree of compaction for CGRP (in preparation)

[30] Klein-Seetharaman, J., Oikawa, M., Grimshaw, S. B., Wirmer, J., Duchardt, E., Ueda, T., Smith, L.J., Dobson, C.M., and Schwalbe, H. 2002. Long-range interactions within a nonnative protein. Science. 295:1719–1722.

[31] Neri, D., Billeter, M., Wider, G. and Wuthrich, K. 1992.NMR determination of residual structure in a urea-denatured protein, the 434-repressor. Science. 257:1559–1563.

[32] Demarest, S. J., Fairman, R. and Raleigh, D. P. 1998. Peptide models of local and long-range interactions in the molten globule state of human alpha-lactalbumin. J. Mol. Biol. 283:279–291.

[33] Fitzkee, N. C. and Rose, G. D. 2004. Reassessing random-coil statistics in unfolded proteins. Proc. Natl Acad. Sci. 101:12497–12502.

[34] Dill, K. A. and Shortle, D. 1991. Denatured states of proteins. Annu. Rev. Biochem. 60: 795–825.

CHAPTER 6

CHARACTERIZATION OF LARGER POLYMERS WITH SLOW INTERNAL MOTION

The majority of the results reported in this chapter have been published in Physical Biology, 9:6, Kaur P., Plochberger B., Costa P., Cope S.M., Vaiana S.M., and Lindsay S., Hydrophobicity of methylated DNA as a possible mechanism for gene silencing, Copyright (2012). The results presented here have been reprinted with permission from IOP Publishing. This work was done in collaboration with Dr. Stuart Lindsay's group, and originated from their interest in studying the effect of methylation on DNA. My contribution to this work consisted of the light scattering measurements and analysis, as presented here.

6.1. INTRODUCTION

In Chapter 5 we introduced an application of tryptophan triplet quenching (TTQ) to study the internal motions of IAPP, a 37 residue intrinsically disordered protein. TTQ is a powerful tool to characterize both the structure and dynamics of peptides and proteins that move on fast time scales, not accessible by other techniques. To study longer and more soluble biomolecules, which diffuse in solution at times longer than microseconds, we can use more conventional techniques, such as light scattering.

The scattering of electromagnetic waves can offer valuable information to characterize the system at hand. In particular, light scattering can be used to monitor the changes in the hydrodynamic radius, radius of gyration and molecular weight of proteins in solution. Advanced data analysis can give further information about the polymer properties of the system, such as the persistence length. In most light scattering instruments, particles up to 500 nm can be fully characterized. This makes light scattering a reasonable option to characterize the DNA molecules discussed in this chapter.

6.2. AFM REVEALS INCREASED STIFFNESS OF METHYLATED DNA IN THE PRESENCE OF CHROMATIN

AFM images performed by P. Kaur and the Lindsay Group showed that chromatin reconstituted on methylated DNA (meDNA) is compact when imaged under water[1]. Chromatin reconstituted on unmethylated DNA is less compact and less sensitive to hydration. These AFM images of methylated DNA appear stiffer based on the observation that contours lack the rapid fluctuations seen at short distances in the control DNA. Kaur et al. quantified these differences by fitting DNA contours in the images to a wormlike chain model[2]. They found that the persistence length of the unmethylated DNA was 47±9.5 nm, in line with the commonly reported value of 50 nm[3]. This value increases when half the possible sites are methylated, reaching 92.5±4 nm when the sample is fully methylated (i.e., 9% of all bases). These differences must reflect changes in the physical properties of DNA upon methylation, but prior studies have not revealed large differences between methylated and unmethylated DNA[4,5,6,7,8,9].

To resolve this issue, the Lindsay group asked us to measure the hydrodynamic radii of methylated DNA and unmethylated DNA in bulk solution. We carried out these measurements in order to understand if the difference in the conformation of methylated DNA (meDNA) was a property of DNA unique to interfaces, or rather an intrinsic property of meDNA. Measurements characterizing the persistence length of methylated and unmethylated DNA in bulk solution were done using angle-resolved quasi-elastic light scattering (QELS).

6.3. MULTI-ANGLE QELS OF DNA

6.3.1. MATERIALS AND METHODS

INSTRUMENTATION

To perform simultaneous multi-angle static (MALS) and dynamic light scattering experiments, we adapted a conventional Wyatt instrument to perform multi-angle dynamic light scattering, while maintaining absolute intensity calibrations for MALS. Multi-angle quasi-elastic dynamic light scattering (QELS) data were collected at 25°C at angles ranging from 26° to 100°, using a Peltier temperature-controlled Wyatt Technology Dawn Heleos II instrument for MALS, equipped with a fiber optic connection at one angle to a DynaPro NanoStar multi-tau correlator with a 100 ns sampling time, a 658 nm, 120 mW GaAs linearly polarized laser and a 70 μL fused silica flow cell. A Razel R99-EJ syringe pump system was used to deliver small sample

volumes into the flow cell minimizing the introduction of air bubbles. The system was initially calibrated with toluene to obtain absolute scattered intensities at 90°. The fiber optic, typically 90°, was mounted at different angles to obtain multi-angle QELS data. At each angle, the fiber optic coupling was optimized by maximizing the scattered intensity of buffer. For each position of the fiber optic, 1 mg/mL bovine serum albumin (BSA) in 1/10 PBS buffer was then injected as a control and to normalize photodiodes to 90°, prior to injecting 200 μL of 0.03-0.06 mg/mL (dilute conditions) DNA for QELS data collection. This ensured an absolute intensity calibration, allowing for simultaneous collection of MALS data and QELS data at each angle. Data were first analyzed using Wyatt Technology ASTRA software and subsequently exported into MatLab for angle dependent analysis as described below.

DATA ACQUISITON AND ANALYSIS

DNA sample concentrations (0.03 mg/ml methylated DNA and 0.06 mg/ml unmethylated DNA) were chosen to ensure dilute conditions while yielding scattering intensities at least 10 times larger than buffer at 90° and an autocorrelation function amplitude greater than 1.2. Because the presence of small amounts of dust particles, air bubbles and possible aggregates can greatly affect the signal at low angles, data filtration procedures were applied. At each angle, individual autocorrelation functions were acquired for short, 5-second intervals for a total time of 4-10 minutes, and correlation functions that did not meet specific criteria were rejected (the total acquisition time was limited by eventual leakage of sample from the flow cell, due to the small sample volumes used). Initially, a consecutive time window in which scattered intensities did not exceed 10% of the

average was selected. Within this time window, individual correlation functions with baselines greater than 1.01 were rejected. This resulted in an effective acquisition time of 0.8-7.1 min, corresponding to a minimum of 9 correlation functions per angle. These correlation functions were averaged and fit to a cumulant algorithm (apparent polydispersity caused deviations from single exponential decay and the signal to noise ratio was insufficient for reliable regularization fits). Data sets, best fit functions and fitting parameters provided by ASTRA software for each angle were exported into MatLab for further processing. Hydrodynamic radii reported by ASTRA at each angle were re-converted to a raw correlation time, τ, using the expression $\tau = (6\pi\eta r_H)/(q^2 k_B T)$, where q is the scattering vector. We measured the following q values in our configuration: 5.70e4, 9.29e4, 1.27e5, 1.63e5, 1.96e5 cm-1. These values were used to create plots in Figure 6.2. To calculate the error, each time window selected for analysis was split into 5 equal time intervals. The correlation time, τ, was calculated for each data slice from the reported hydrodynamic radius and the error calculated as the standard deviation of the mean. Because the hydrodynamic radius does not scale linearly with the persistence length, ξ, (Equations 5), errors on ξ were determined by calculating ξ for the upper and lower bounds of r_H. For correlation times due to pure translational diffusion (i.e. in the absence of rotational diffusion or internal motions) $\tau = ((D*q^2)^{-1})$, where D is the diffusion coefficient, related to the hydrodynamic radius by Stoke's-Einstein (Equation 4). In the case of long polymers (e.g. DNA above 1Kbp) internal motions/rotational diffusion give rise to additional correlation relaxation times which can overlap with the translational correlation time at large angles[10]. Plots of $1/\tau$ versus q^2 that deviate from linearity at high angles are a typical signature of this effect. In this case, the

actual translational diffusion coefficient D is obtained by fitting $1/\tau$ versus q^2 data points to a second order polynomial and taking the slope at q=0 (zero angle extrapolation), as in Figure 6.2. Data in Figure 6.2 were fit to a second order polynomial of the form $y=ax^2+bx$ and D was calculated from the fitting parameter b. Error on the diffusion coefficients was taken from the 95% confidence interval of the fitting.

6.3.2. RESULTS

In order to determine if control DNA and meDNA had the same persistence length in bulk solution, we first looked at the simplest case by comparing their "effective" hydrodynamic radius at 90°, that is considering that the time, τ, was the intrinsic time of translational diffusion. As seen in Figure 6.1, the data do not follow a single exponential decay, but rather exponential decays with a Gaussian distribution of relaxation times (cumulant fit), τ_R. This indicates that there is apparent polydispersity. Verification by gel filtration showed that the sample was highly pure and monodisperse. Apparent polydispersity results from the timescales of internal motions/rotational diffusion, which overlap with the translational correlation time at large angles[9]. This effect is expected for long, stiff polymers, and thus we verified it by performing measurements as a function of angles.

Figure 6.1 shows the scattering intensity, I(t), correlation function, $g^2(q,\tau)$, which is the second order correlation function of the scattered field as a function of time, τ.

$$g^2(q,\tau) = \frac{\langle I(t), I(t+\tau)\rangle}{\langle I(t)\rangle^2} \qquad (1)$$

and q is the scattering vector, obtained from the scattering angle, θ, light wavelength, λ, and refractive index of the solvent, n, via

$$q = \frac{4\pi n}{\lambda}\sin\left(\frac{\theta}{2}\right). \qquad (2)$$

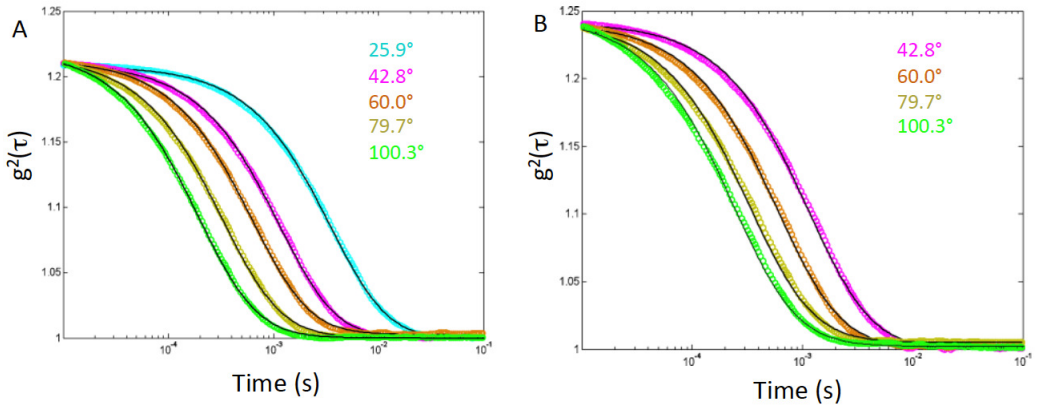

Figure 6.1. QELS data for the intensity correlation function, $g^2(q,\tau)$ as a function of time, τ, for control DNA (A) and meDNA (B) at different scattering angles (as indicated). The solid lines are cumulant fits using an exponential decay with a Gaussian distribution of relaxation times, τ_R. The average value is used in subsequent analysis.

Hydrodynamic radii for control DNA and meDNA reported by ASTRA at each angle were re-converted to raw correlation times, τ, using the expression

$$\frac{1}{\tau_R} = Dq^2 \qquad (3)$$

,

where D is the diffusion constant of the molecule which is given by

$$D = \frac{k_B T}{6\pi\eta r_h}. \qquad (4)$$

As explained in the methods, data in Figure 6.2 were fit to a second order polynomial of the form $y=ax^2+bx$ and D was taken to be the fitting parameter b.

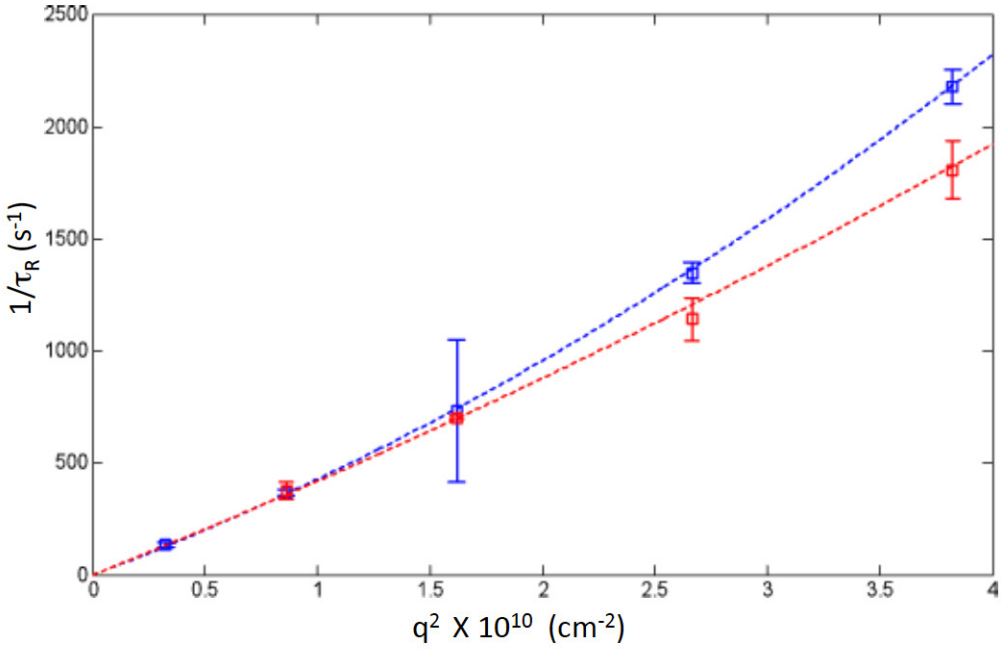

Figure 6.2: Inverse relaxation time plotted versus the square of the scattering vector for meDNA (red data points) and control DNA (blue data points). Departure from linearity, owing to internal fluctuations, is most evident for the control DNA. The lines are fits to a quadratic in q^2. The coefficient of the linear term (in q^2) is the diffusion constant for the molecule.

Both sets of data converge at small values of *q* as expected for large polymers, yielding D=3.97±0.61×10^{-8} cm^2s^{-1} for meDNA and 3.76±0.44×10^{-8} cm^2s^{-1} for the control DNA. These diffusion constants are equivalent, within error, indicating equal sizes for the molecules, and thus equal persistence lengths for the two types of DNA in solution.

Note, however, the smaller departure of meDNA (red dashed line) from ideal (linear) behavior indicates less structural fluctuation. This is consistent with molecular dynamics simulations that suggest that methylation leads to suppression of such fluctuations[8,11].

These results can be compared to our data for persistence length via the ratio of the hydrodynamic radii for the two molecules. The Stokes-Einstein relation (Equation 4) yields the hydrodynamic radii, r_H. Given that the radius of gyration, r_G, and the hydrodynamic radius are related by a constant factor[12], the light scattering results yields the ratio of the radii of gyration of the methylated molecule to that of the control molecule as

$$\frac{r_G(me)}{r_G(c)} = 0.95 \pm 0.3.$$

In the worm-like chain model, the radius of gyration and persistence length are related by

$$r_G^2 = \xi^2 \left[\frac{L_C}{3\xi} - 1 + \frac{2\xi}{L_C} - \frac{1 - \exp\left(-L_C/\xi\right)}{\left(L_C/\xi\right)^2} \right] \qquad (5)^{12}$$

where L_C is the contour length of the molecule. Taking the control DNA to have a persistence length of 50 nm and the methylated DNA to have a persistence length of 92.5±4 nm, equation 5 yields

$$\frac{r_G(me)}{r_G(c)} = 1.28 \pm 0.01.$$

Thus, although the hydrodynamic radius is rather insensitive to changes in persistence length, the difference between the light scattering result in solution and the AFM result

on a surface is significant. The large persistence length observed in the AFM measurements is unlikely to apply in solution, though the uncertainties in the light scattering data do not rule out the possibility of a somewhat larger persistence length for meDNA in solution.

A more robust way of characterizing persistence length is through direct measurement of the radius of gyration. Conversion between hydrodynamic radius and radius of gyration can be challenging, and so we set out to measure r_G directly using static multi-angle light scattering[12]. The presence of small amounts of dust particles, air bubbles and possible aggregates can greatly affect the signal at low angles. This notoriously limits the possibility of directly measuring r_G for DNA samples. Our set-up allowed us to simultaneously measure MALS and QELS, offering us a convenient method by which we could filter our data from contributions due to dust or air bubbles. We used the correlation functions to determine windows over which we would accept static intensity measurements. Since static scattering intensities were measured concurrently with QELS, the method used to filter correlation functions (explained above) was also used to filter static intensities.

From these measurements, we measured the static intensity of scattered light as a function of angle. This data, shown in Figure 6.3, was then used to obtain r_G from both a model-free analysis utilizing the Guinier Approximation for small angles and a random coil model[13].

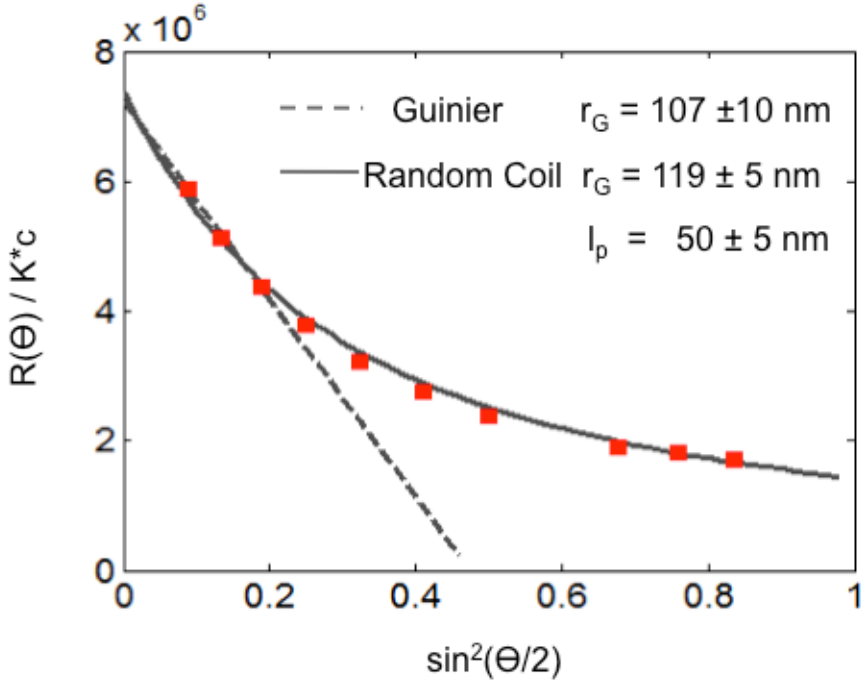

Figure 6.3: Multi-Angle Static Light Scattering measurements of meDNA achieved by using correlation functions to filter data. The angular dependence of the scattered light was used to find r_G by both a model-free analysis, employing the Guinier Approximation, and a random coil model.

We then used Equation 5 to calculate the persistence length for each r_G. Using both analyses, we found a persistence length of 50nm ± 5 nm for meDNA, corresponding to the known persistence length of unmethylated DNA in solution. These two methods, multi-angle QELS and multi-angle static light scattering, confirmed that, within the limitations of our experiment, there is no difference in the persistence length of meDNA and control DNA in solution.

6.4. DISCUSSION

The light scattering results for the ratio of hydrodynamic radii, taken together with the lack of an observed conformational transition in solution are evidence of a significant interfacial effect driving the stiffening of meDNA at the interface, an effect that would not occur for isolated DNA in solution. As explained at length in Ref. 1, this suggests that the hydrophobicity of DNA, as determined from analyzing the contour length and persistence length, could act as a simple mechanism for gene silencing, as the stiffer meDNA is more difficult to remove from nucleosomes.

REFERENCES

[1] Kaur P., Plochberger B., Costa P., Cope S.M., Vaiana S.M., and Lindsay S., 201. Hydrophobicity of methylated DNA as a possible mechanism for gene silencing. Physical Biology. 9:6.

[2] Wiggins, P. and Nelson, P. 2006 Generalized theory of semiflexible polymers. Phys. Rev.17:2709–18.

[3] Perez, A. 2012. Impact of methylation on the physical properties of DNA. Biophys. J. 102:2140–8.

[4] Hodges-Garcia, Y. and Hagerman, P.J. 1995. Investigation of the influence of cytosine methylation on DNA flexibility. J. Biol. Chem. 270:197–201.

[5] Nathan, D. and Crothers, D.M. 2002. Bending and flexibility of methylated and unmethylated EcoRI DNA. J. Mol. Biol. 316:7–17.

[6] Hodges-Garcia, Y. and Hagerman, P.J. 1992. Cytosine methylation can induce local distortions in the structure of duplex DNA. Biochemistry. 31:7595–9.

[7] Banyay, M. and Gräslund, A. 2002. Structural effects of cytosine methylation on DNA sugar pucker studied by FTIR. J. Mol. Biol. 324:667–76.

[8] Severin, P.M.D., Zou X., Gaub, H.E. and Schulten, K. 2011. Cytosine methylation alters DNA mechanical properties. Nucleic Acids Res. 39:8740–51.

[9] Perez, A. 2012. Impact of methylation on the physical properties of DNA. Biophys. J. 102:2140–8.

[10] Kuil, M.E., van Mourik, F., Burger, W., and van Grondelle, R. 1988. The internal dynamics of gene 32 protein-DNA complexes studied by quasi-elastic light scattering. Biophysical chemistry. 32: 211–227.

[11] Severin, P.M.D, Zou, X., Gaub, H.E., and Schulten, K. 2011. Cytosine methylation alters DNA mechanical properties Nucleic Acids Res. 39:8740–51.

[12] Sorlie, S.S. and Pecora, R. 1990. A dynamic light scattering study of four DNA restriction fragments. Macromolecules. 23:87–97.

APPENDIX A

SUPPORTING INFORMATION FOR CHAPTER 4

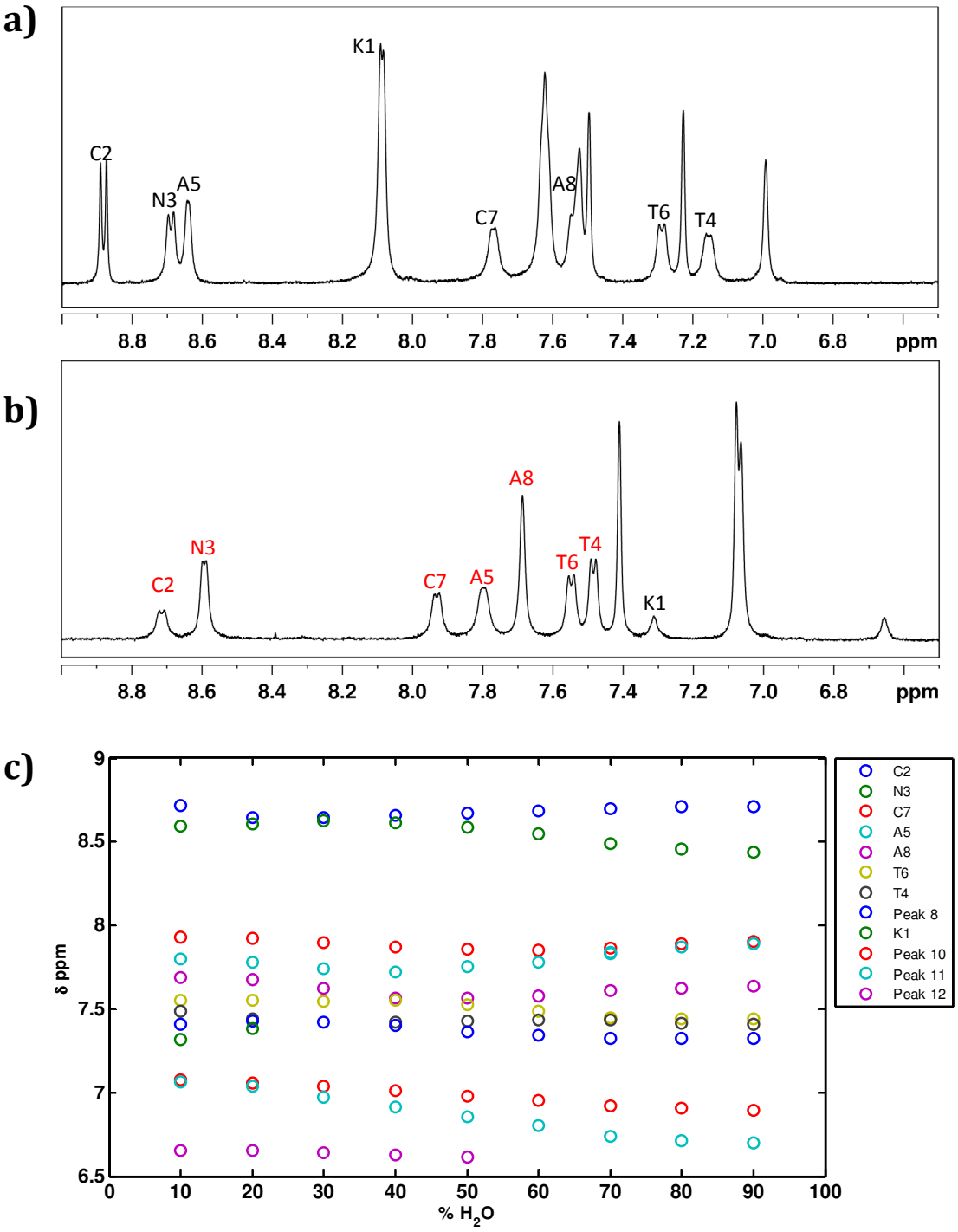

Figure A1.1. To investigate the effect of solvent on N_loop structure we obtained ^{1}H-1D NMR spectra of N_loop in 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20% and 10% DMSO by directly diluting a 100%DMSO, 1.3 mM peptide sample in non-

deuterated buffer (50mM NaAc pH=4.9, prepared with H_2O and deuterated salts/acid). (a) 100% DMSO spectrum (b) 90% DMSO spectrum. (c) peak positions as a function of dilution, from 90%DMSO to 10%DMSO. As expected, an evident change in spectrum occurs between the 100% and 90% DMSO samples, reflecting the change from aprotic to protic solvent conditions (a,b). By contrast, the peak positions hardly change after further dilutions, all the way down to 10% DMSO (c). The 90% DMSO peaks labeled in red (b) were assigned by assuming that each doublet/singlet shifted minimally respect to the 100%DMSO spectrum. This was based on the observed differences in chemical shifts between our 100% DMSO sample and those of both Yonemoto et al. and Williamson et al. in buffer (Figure 4.9). The K1 peak was assigned directly from TOCSY analysis of the 90% DMSO spectrum. All spectra were acquired at 25°C on a 500 MHz Varian spectrometer. Presaturation was used to suppress water in all measurements containing H_2O. Comparison between secondary chemical shifts obtained here and those shown in Figure 4.9 are shown below.

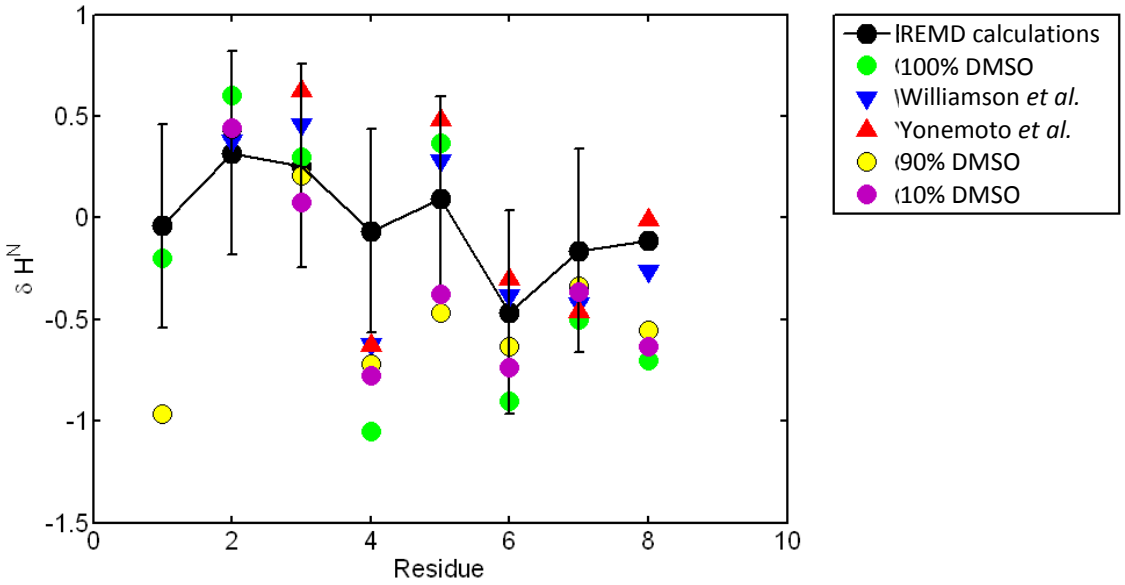

Figure A1.2 Comparison between secondary chemical shifts for N_loop in 100% DMSO (Figure 4.9 in the main text) and in 90% and 10% DMSO (obtained from 1D NMR of Figure A1.1, above). Apart from K1, which changes protonation state in water versus DMSO and is absent in the previous publications (due to proton exchange with the solvent), the observed chemical shifts do not show significant changes from DMSO to water-like solvent. This supports the conclusion that the DMSO has very little effect on the structure of the N_loop and that this structure is maintained in full length IAPP.

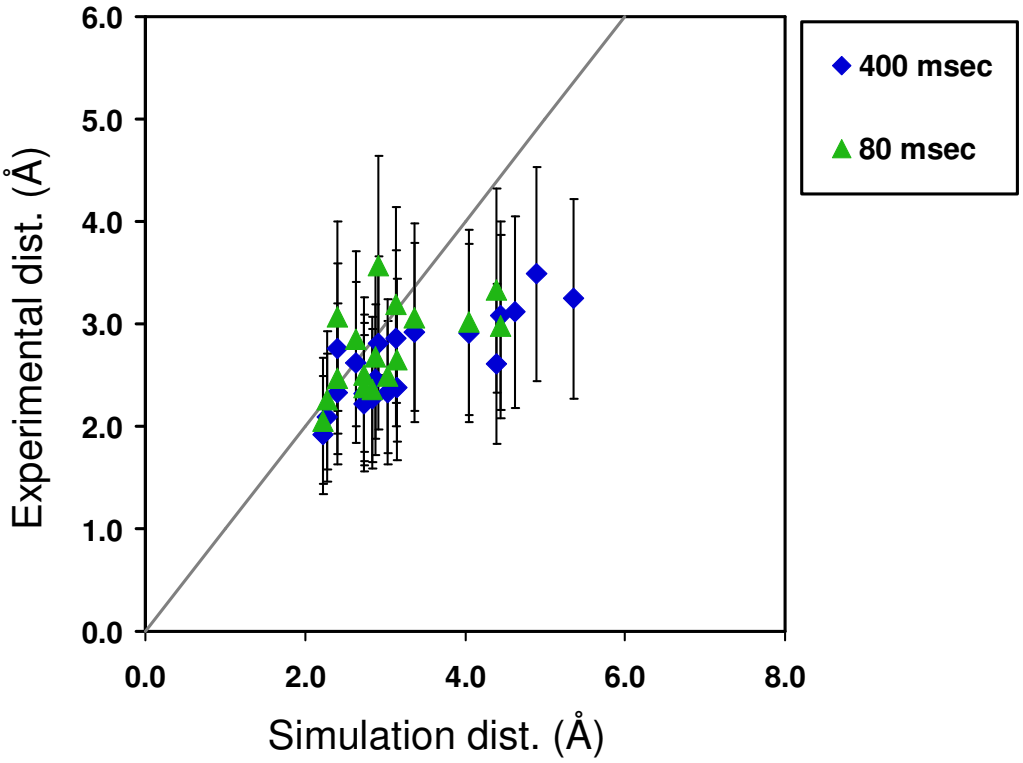

Figure A1.3. Distances estimated from NOE peak intensities of all assignable crosspeaks visible in NOESY spectrum, compared to average distances obtained from REMD simulations. The straight line corresponds to y(x)=x. There is a good correlation between the NOEs reported and the average MD distances. The quantitative agreement is lost at larger distances, most likely because of spin diffusion effects.

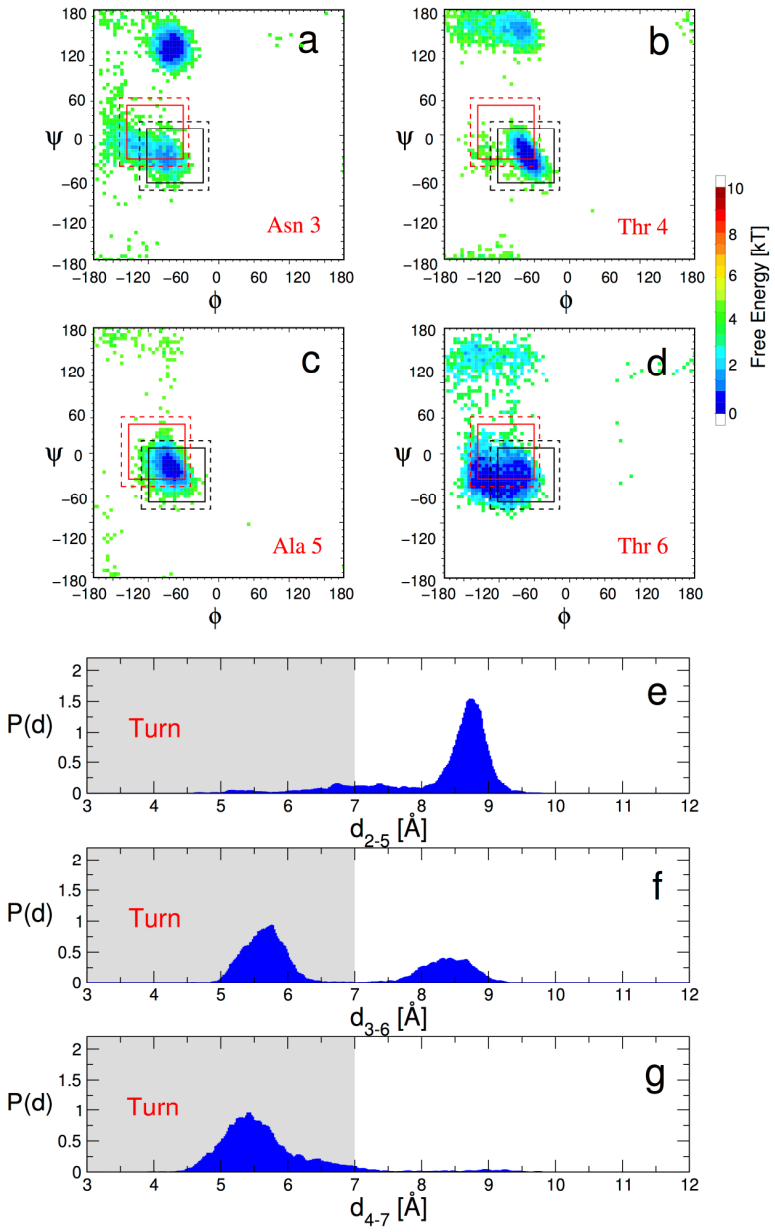

Figure A1.4. Geometric criteria for type I turn formation in N_loop: (a-d) Ramachandran maps of interior residues of N_loop. For a type I β-turn from residues i to i+3, the allowed Ramachandran angles of residues i+1 and i+2 are shown by black and red boxes respectively. Solid and broken lines indicate a range of 40 or 50 degrees respectively, with respect to the

ideal turn angles[1]*. (e-g) Distance between alpha carbon atoms of residues i,i+3 for different i; this distance should be less than 7 Angstroms for a turn to be defined. Based on these geometric criteria, a type I β-turn is present at residues 3-6 most of the time, with a significant fraction of type I turn also present at residues 4-7. The criteria for a type II turn, for which the ideal Ramachandran angles are i: (-60,120); i+1: (80, 0), are clearly not satisfied.*

[1] Hutchinson, E. G., and J. M. Thornton. 1994. A Revised Set of Potentials for Beta-Turn Formation in Proteins. Protein Science 3:2207-2216.

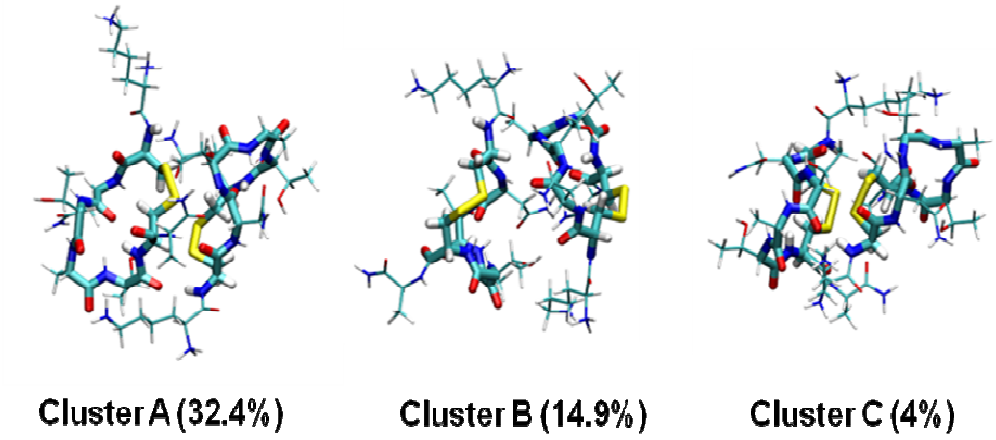

Figure A1.5. Major clusters (of associated states) populated by the N_loop dimer in 298 K replica of REMD simulations.

APPENDIX B

PERMISSION TO REPRODUCE COPYRIGHTED AND CO-AUTHORED WORK

All co-authors have granted permission for use of the material in Chapters 4 and 6 for the purpose of this dissertation

The results reported in Chapter 4 are reprinted from Biophysical Journal, 105:7, Cope, S.M., Shinde, S., Best, R.B., Ghirlanda, G., and Vaiana, S.M., Cyclic N-terminal loop of amylin forms non amyloid fibers, Copyright (2013), with permission from Elsevier Publishing (VAT GB 494 6272 12).

The results reported in Chapter 6 are reprinted from Physical Biology, 9:6, Kaur P., Plochberger B., Costa P., Cope S.M., Vaiana S.M., and Lindsay S., Hydrophobicity of methylated DNA as a possible mechanism for gene silencing, Copyright (2012), with permission from IOP Publishing.