Springer Theses
Recognizing Outstanding Ph.D. Research

Katrine Kirkeby Skeby

Computational Modelling of the Human Islet Amyloid Polypeptide

Springer

Springer Theses

Recognizing Outstanding Ph.D. Research

Aims and Scope

The series "Springer Theses" brings together a selection of the very best Ph.D. theses from around the world and across the physical sciences. Nominated and endorsed by two recognized specialists, each published volume has been selected for its scientific excellence and the high impact of its contents for the pertinent field of research. For greater accessibility to non-specialists, the published versions include an extended introduction, as well as a foreword by the student's supervisor explaining the special relevance of the work for the field. As a whole, the series will provide a valuable resource both for newcomers to the research fields described, and for other scientists seeking detailed background information on special questions. Finally, it provides an accredited documentation of the valuable contributions made by today's younger generation of scientists.

Theses are accepted into the series by invited nomination only and must fulfill all of the following criteria

- They must be written in good English.
- The topic should fall within the confines of Chemistry, Physics, Earth Sciences, Engineering and related interdisciplinary fields such as Materials, Nanoscience, Chemical Engineering, Complex Systems and Biophysics.
- The work reported in the thesis must represent a significant scientific advance.
- If the thesis includes previously published material, permission to reproduce this must be gained from the respective copyright holder.
- They must have been examined and passed during the 12 months prior to nomination.
- Each thesis should include a foreword by the supervisor outlining the significance of its content.
- The theses should have a clearly defined structure including an introduction accessible to scientists not expert in that particular field.

More information about this series at http://www.springer.com/series/8790

Katrine Kirkeby Skeby

Computational Modelling of the Human Islet Amyloid Polypeptide

Doctoral Thesis accepted by
the Aarhus University, Aarhus C, Denmark

Springer

Author
Prof. Katrine Kirkeby Skeby
Department of Chemistry and the
Interdisciplinary Nanoscience Center
Aarhus University
Aarhus C
Denmark

Supervisor
Prof. Birgit Schiøtt
Department of Chemistry and the
Interdisciplinary Nanoscience Center
Aarhus University
Aarhus C
Denmark

ISSN 2190-5053 ISSN 2190-5061 (electronic)
Springer Theses
ISBN 978-3-319-20039-2 ISBN 978-3-319-20040-8 (eBook)
DOI 10.1007/978-3-319-20040-8

Library of Congress Control Number: 2016939051

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Part of this thesis have been published in the following journal articles:

1. **Protofibrillar Assembly Toward the Formation of Amyloid Fibrils**, Sørensen, J., Periole, X., Skeby, K. K., Marrink, S.-J. and Schiøtt, B. *J. Phys. Chem. Lett.* (2011) 2, 2385–2390
2. **Identification of a Common Binding Mode for Imaging Agents to Amyloid Fibrils from Molecular Dynamics Simulations**, Skeby, K. K., Sørensen, J. and Schiøtt, B. *J. Am. Chem. Soc.* (2013) 135, 15114–15128
3. **The Importance of Being Capped: Terminal Capping of an Amyloidogenic Peptide Affects Fibrillation Propensity and Fibril Morphology**, Andreasen, M., Skeby, K. K., Zhang, S., Nielsen, E. H., Klausen, L. H., Frahm, H., Christiansen, G., Skrydstrup, T., Dong, M., Schiøtt, B. and Otzen, D. *Biochemistry* (2014) 53, 6968–6980
4. **Conformational Dynamics of the Human Islet Amyloid Polypeptide in a Membrane Environment: Toward the Aggregation Prone Form**, Skeby, K. K., Andersen, O. J., Pogorelov, T., Tajkhorshid, E., and Schiøtt, B. *Biochemistry* (2016) 55, 2031–2041

Supervisor's Foreword

It is a great pleasure for me to write this introduction for a truly outstanding Ph.D. graduate Katrine K. Skeby, who graduated at Aarhus University in November 2014. Katrine K. Skeby has been involved in research related to obtaining a much improved understanding of the molecular reason for amyloid diseases and their possible detection through advanced imaging techniques. This class of diseases is emerging as a major societal challenge, with Alzheimer's and Parkinson's diseases being the most prominent age-related neurodegenerative disease and diabetes mellitus type 2 reaching epidemic stages some places around the world. The burden on people affected by these diseases as well as the health-care systems of society is enormous; thus, research into the cause of these detrimental diseases is of utmost importance. So far, the exact cause of the diseases is not known; however, it is known that somehow otherwise well-functioning proteins and peptides misfold and form large aggregates that accumulate in different tissues. However, it is generally thought that it is not the final aggregates that are toxic; rather, it is believed that smaller intermediate forms, which exhibit cytotoxicity toward the cell membranes of the affected tissue, brain tissue in Alzheimer's and Parkinson's disease, and pancreatic tissue in diabetes mellitus type 2. Not only is fundamental understanding at the molecular level of this important class of diseases still rudimentary, but also the detection and diagnosis of, e.g., Alzheimer's disease is extremely challenging, relying mostly on cognitive tests. Only by biopsy of a diseased person, can the cause of dementia be determined with certainty. However, in recent years, much research has focused on developing advanced imaging techniques, such as MR or PET scanning methods, in which a contrast agent that selectively binds the amyloid deposits is injected and visualized.

The scientific contributions from the Ph.D. work of Katrine K. Skeby address both aspects described above, the fundamental cause of the disease as well as modern diagnosis through imaging. She uses advanced computational methods, such as molecular dynamics simulations and free-energy calculations, to describe the process studied. Such methods play an important role in modern research of complex chemical systems, as implied by the Nobel Prize in chemistry in 2013

given to exactly this field. Katrine K. Skeby has worked in a very competitive area of expertise dealing with the discovery of a number of molecular mechanisms related to amyloid disease and detection. Her results include (i) a common binding mode for protein imaging agents to amyloid protofibrils, (ii) an understanding of the intermolecular order in fibrils, (iii) an understanding of the initial phases of peptide interaction with lipid bilayer surrounding a cell, and (iv) understanding of the self-assembly of protofibrils into more mature fibrillary structures. In all of this work, Katrine K. Skeby has used the human islet amyloid polypeptide (hIAPP) as a model system for her simulations. hIAPP is related to the development of diabetes mellitus type 2, where it forms cytotoxic plaques in the pancreas. She has collaborated closely with several experimental groups, providing models that explain experimental observations. Several studies in the thesis answer difficult questions, which will advance future research into the properties of amyloid proteins. Katrine K. Skeby has dealt with an extremely difficult biological subject and has been able to obtain insight into the molecular mechanisms related to what causes the behavior of the hIAPP at a level of detail not previously seen.

The thesis is very well written and contains a comprehensive introduction to amyloid proteins and the computational methods employed. She shows a high degree of understanding of the experimental literature, which is one of the reasons for the successful nature of her research, leading to her contributions being published in high-impact journals such as *J. Am. Chem. Soc.* and *Biochemistry.* The thesis contains a very valuable review of both experimental and theoretical literature regarding amyloid diseases. Most importantly, Katrine K. Skeby diligently describes current structural knowledge in relation to amyloid fibrils, their formation, and detection. Both in the writing of the thesis and in the research phase, Katrine K. Skeby has shown excellent maturity as a computational chemist, dramatically improving the understanding of hIAPP properties and providing important mechanistic information that will, no doubt, help others to improve even further the understanding of the cause and detection of amyloid diseases.

Aarhus C
March 2016

Prof. Birgit Schiøtt

Acknowledgements

First of all, I would like to thank my supervisor, Prof. Birgit Schiøtt, for allowing me to begin my Ph.D. studies in the Biomodelling Group in September 2009, directly following the completion of my bachelor's degree. She has allowed me to conduct independent work under skillful guidance in a comfortable and ambitious working environment. Furthermore, she has encouraged me to participate in international conferences and workshops, which have broadened my competences and network within the fields of computational biophysics, amyloid research, and scientific visualization.

My collaborators also deserve a great acknowledgement. Jesper Sørensen has been a close collaborator and advisor during my Ph.D. studies, even after he was not a part of the Biomodelling Group anymore, for which I am exceedingly grateful. I am also thankful to Prof. Emad Tajkhorshid for allowing me to be a part of his research group at the Beckman Institute, University of Illinois at Urbana-Champaign. In connection, I owe thanks to the entire Computational Structural Biology and Molecular Biophysics Group, and especially Taras Pogorelov, for welcoming me into the group. Furthermore, I would like to thank all of our collaborators in the Center for Insoluble Protein Structures (inSPIN) for fruitful discussions and collaborations which have led to interesting projects and excellent research. Alessandro Laio also deserves special thanks for all his help and excellent guidance during my stay at the Scuola Internazionale Superiore di Studi Avanzati (SISSA) in Trieste, Italy.

A special thanks is given to past and present members of the Biomodelling Group, who have made my time as a Ph.D. student both enjoyable and challenging. The many fun social events are the basis for having a fruitful working environment to which I look forward every day. Sofie Jakobsen deserves a special thanks for always being my partner in crime and a great office buddy.

I am grateful to Jesper Sørensen, Maria Andreasen, Julie Grouleff, Ole Juul Andersen, Sofie Jakobsen, and Heidi Koldsø for proofreading parts of this thesis.

For financial support, I would like to acknowledge the following bodies: the Faculty of Science and Technology and the iNANO Center at Aarhus University;

the Danish Agency for Science, Technology and Innovation; Chr. Hansen A/S; and BioSys. The Hakon Lund Foundation provided me with a travel grant which made it possible for me to visit the Scuola Internazionale Superiore di Studi Avanzati (SISSA) in Trieste, Italy, for which I am grateful. Furthermore, I would like to thank the EMBL ATC Corporate Partnership Programme for a registration fee waiver which allowed me to participate in the EMBO conference "Visualizing Biological Data (VIZBI)" at the EMBL Heidelberg in March 2014. I am also grateful to the Royal Society of Chemistry for a travel bursary which allowed me to participate in the Faraday Discussion 169 on "Molecular Simulations and Visualization" in Nottingham, UK, in May 2014.

Last but definitely not least, I would like to extend a great thank you to my family and friends. My parents have provided me with the values that have allowed me to focus on my studies while keeping a balance with all other important aspects in my life. I am eternally grateful for the help they have given me throughout my life, while teaching me to be independent. My boyfriend, Søren Aabling, also deserves the greatest acknowledgement for putting up with my busy schedule and intermittent break downs. I am grateful for having him in my life. Cecilie Skeby, Sofie Jakobsen, and Henriette Elisabeth Autzen should also know that I look forward to having fun with them almost every day in the gym.

Contents

Abbreviations

Aβ	Amyloid β
AD	Alzheimer's disease
AFM	Atomic force microscopy
AMBER03	Duan et al.'s all-atom point-charge force field
aMD	Accelerated molecular dynamics
C_C	Critical concentration
CD	Circular dichroism
CG	Coarse grained
CR	Congo red
CSF	Cerebrospinal fluid
CV	Collective variable
DCLE	Dichloroethane
DOPC	1,2-dioleoyl-sn-glycero-3-phosphatidylcholine
DOPG	1,2-dioleoyl-sn-glycero-3-phosphatidylglycerol
DOPS	1,2-dioleoyl-sn-glycero-3-phosphatidylserine
DVPC	Divalerylphosphatidylcholine
DVPS	Divalerylphosphatidylserine
EM	Electron microscopy
EPR	Electron paramagnetic resonance
FDA	American Food and Drug Administration
FF	Force field
FRET	Förster resonance energy transfer
FTIR	Fourier transform infrared spectroscopy
GAFF	General AMBER force field
GPU	Graphical processing unit
hIAPP	Human islet amyloid polypeptide/Amylin
HMMM	Highly mobile membrane model
LJ	Lennard-Jones
MD	Molecular dynamics
MM	Molecular mechanics
MM-PBSA	Molecular mechanics Poisson–Boltzmann surface area

MRI	Magnetic resonance imaging
NMR	Nuclear magnetic resonance
NPT	Isothermal–isobaric ensemble
NVE	Microcanonical ensemble
NVT	Canonical ensemble
P	Pressure
PB	Poisson–Boltzmann
PC	Phosphatidylcholine
PDB	RCSB Protein Data Bank
PET	Positron emission tomography
PIB	Pittsburgh compound B
PME	Particle mesh Ewald
PS	Phosphatidylserine
QM	Quantum mechanics
RMSD	Root-mean-square deviation
RMSF	Root-mean-square fluctuation
SA	Surface area
SI	Supplementary information
SN	SNNFGAILSS
ss-NMR	Solid-state nuclear magnetic resonance
STEM	Scanning tunneling electron microscopy
T	Temperature
T2DM	Type 2 diabetes mellitus
TEM	Transmission electron microscopy
ThT	Thioflavin T
vdW	van der Waals

Amyloid and Amyloid Fibrils

Abstract When proteins do not fold correctly, it can lead to very serious diseases. One such group of diseases is the amyloid diseases, of which Alzheimer's disease (AD), Parkinson's disease, and type 2 diabetes mellitus (T2DM) are members. The amyloid diseases are characterized by the aggregation of a specific protein into amyloid fibrils. During this process, a cytotoxic event occurs which can be a serious actor in the evolvement of the disease. This thesis is concerned with elucidating the biological processes concerning amyloid proteins, more specifically, the peptide hormone human islet amyloid polypeptide (hIAPP), which is involved in glucose homeostasis and deposits in the pancreas of T2DM patients.

Amyloid proteins are notoriously difficult to study, as the mechanisms governing their actions have not been perfected during evolution to function robustly. The aggregation mechanism is controlled by a delicate thermodynamic and kinetic balance affected by the slightest change in conditions. The rapid aggregation of the monomer and the insolubility of the fibrils make the proteins difficult to study using conventional experimental techniques. Therefore, it is imperative to combine multiple methods to interpret results correctly. Computational studies and molecular dynamics (MD) simulations in particular have become important tools in the effort to understand biological mechanisms. The strength of these methods is the high resolution in time and space, and the ability to specifically design the system setup.

We have exploited these strengths to study the interactions between an amyloid fibril and amyloid imaging agents. Imaging agents are promising tools for the detection of amyloid deposits in the brain of AD patients. This could aid in the early diagnosis as well as evaluation of new treatments. Using MD simulations we have investigated the binding of 13 different imaging agents to a fibril segment. Using clustering analysis and binding energy calculations we have identified a common binding mode for the 13 agents in the surface grooves of the fibril, which are present on all amyloid fibrils. This information combined with specific knowledge about the AD amyloid fibril is the building block for the design of highly specific amyloid imaging agents.

K.K. Skeby, *Computational Modelling of the Human Islet Amyloid Polypeptide*,
Springer Theses, DOI 10.1007/978-3-319-20040-8_1

We have also used MD simulations to study the interaction between hIAPP and a phospholipid membrane. At neutral pH, we find that the attraction is mainly between the N-terminal part of hIAPP and the lipid head-groups. This is due to positive charges present in the N-terminal part of hIAPP interacting with the anionic lipids. The C-terminal part of hIAPP is unfolded in the solution phase, making this part of hIAPP ready for interaction with other peptides. When the pH is lowered to the value found in the secretory granules storing the peptide hormone prior to release, His18 becomes protonated. This adds an additional positive charge to the peptide, causing the C-terminal part of hIAPP to interact with the membrane as well. This allows the peptide to move deeper into the lipid bilayer. We speculate that this is the mechanism of aggregation inhibition which is seen at low pH.

As the study of amyloid proteins is so difficult, shorter fragments of the amyloid peptides are often used as model systems to study aggregation. However, the implications of using the shorter fragments have not been clarified. We have studied the aggregation and fibril structure of the SNNFGAILSS fragment from hIAPP with different capping groups at the peptide termini. These result in different fibril morphologies as shown by electron microscopy. Using MD simulations we have obtained clues to the mechanisms governing the different physical properties of the peptides and fibrils. Capping the C-terminus of SNNFGAILSS results in twisted antiparallel fibrils caused by the repulsion between the N-terminal charges. Capping the N-terminus as well or removing the C-terminal cap results in flat ribbons which are due to the compatible peptide termini. Capping only the N-terminus abolishes the fibrillation, which is caused by incompatibility of the hydrophobic N-terminus with the anionic C-terminus as well as a lower number of interpeptide hydrogen bonds to overcome the repulsion. This highlights the importance of choosing the right model system, and of considering the consequences of modifying the amyloid system.

We have, furthermore, employed a model amyloid system to study amyloid aggregation using coarse grained MD simulations, where several atoms are grouped into a single particle, reducing the number of particles in the system. Coarse grained MD simulations are necessary to study amyloid aggregation computationally, as the time scale and the system size needed for the process are not currently accessible with atomistic MD simulations. Using the MARTINI coarse grained model, we have simulated the aggregation of 27 small amyloid fibril segments and find that the association primarily occurs at the elongation ends. However, this preference is lost at higher temperatures, which is an example of the delicate nature of the amyloid systems. Furthermore, this study highlights some of the challenges still faced by the community when using coarse grained simulations for studying amyloid aggregation.

Introduction

Proteins are the molecular machines that make cells, organs, and the entire human body function. They are composed of a sequence of amino acids folded into a unique three-dimensional structure, which gives the protein a specific function [1].

However, sometimes proteins do not fold into the proper structure, causing the protein to be inactive or sometimes even toxic. Some proteins misfold and combine to form insoluble aggregates termed amyloid, which can lead to a disease state called amyloidosis. Amyloid disease is the focus of this PhD dissertation; mainly the two related amyloid diseases, type 2 diabetes mellitus (T2DM) and Alzheimer's disease (AD). This chapter describes the current state of knowledge regarding amyloid and amyloid disease.

Amyloid Disease

Amyloid diseases are characterized by the misfolding of specific peptides and proteins followed by self-association into highly ordered and insoluble structures termed amyloid fibrils. More than 40 different amyloid diseases have been characterized, each associated with a specific protein (Table 1) [2].

The amyloid diseases can be divided into three classes dependent on the tissue localization of the misfolding protein [2]. Diseases in which the amyloid formation and deposition occurs in the brain tissue are termed *neurodegenerative amyloidoses*, examples of which are AD and Parkinson's disease. Diseases characterized by the deposition of amyloid in a localized tissue other than the brain are termed *non-neurodegenerative localized amyloidoses*, and include such diseases as T2DM and cataract. Amyloid diseases in which the deposition of misfolded protein occurs in multiple tissues in the body are termed *non-neurodegenerative systemic amyloidoses* [2]. The latter type of disease is less known, but includes e.g. hereditary diseases caused by mutants of lysozyme [2]. Amyloid disease can occur sporadically, like in most cases of AD. Furthermore, transmissible amyloid diseases are known, such as spongiform encephalopathies, and other external factors such as prolonged dialysis treatment can also lead to amyloid disease [2, 3]. Amyloid diseases have very diverse symptoms; however, they are linked by a common pathological mechanism, namely the presence of amyloid deposits.

Table 1 Examples of amyloid diseases and associated amyloid peptides [2]

Disease	Aggregating protein/peptide	No. of Residues
Alzheimer's disease	Amyloid β	40 or 42
Frontotemporal dementia with Parkinsonism	Tau	352–441
Spongiform encephalopathies	Prion protein or fragments thereof	253
Parkinson's disease	α-Synuclein	140
Lysozyme amyloidosis	Mutants of lysozyme	130
Type 2 diabetes mellitus	Islet amyloid polypeptide (hIAPP) (Amylin)	37
Cataract	γ-Crystallins	Variable

Amyloid

The term "Amyloid" was first described as such by the German physician and scientist, Rudolph Virchow, in 1854 [4]. Amyloid means starch-like, and Virchow named it such due to the starch-like behavior of amyloid containing tissue samples, which stained blue upon treatment with iodine [4]. Amyloid has since been defined as "extracellular deposits of protein fibrils with a characteristic appearance in the electron microscope, a common X-ray diffraction pattern, and affinity for Congo Red (CR) with concomitant green birefringence." [5] Amyloid deposits, or plaques, are composed primarily of protein in the form of amyloid fibrils. Amyloid fibrils specifically bind the staining agent CR, which then changes optical properties and absorbance maximum. CR has been used since the 1920s to identify the presence of amyloid in tissue samples [6]. However, the deposits are not composed solely of amyloid fibrils; they also contain a high content of other proteins as well as different lipids, glycosaminoglycans, and heavy metal ions [7]. These other actors are believed to take part in modulating the fibril formation of amyloid proteins.

Amyloid Fibrils

As the major constituent of amyloid deposits, amyloid fibrils have been the focus of study for many years. The general atomic structure of amyloid fibrils was first determined in 1968, and is termed the cross-β structure due to the distinct cross pattern in X-ray fiber diffraction experiments (Fig. 1a) [8, 9].

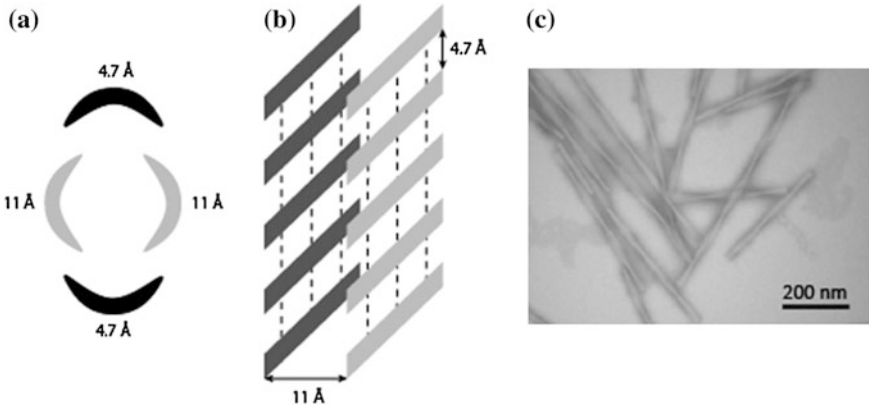


Fig. 1 Structure of amyloid fibrils. **a** Theoretical X-ray fiber diffraction pattern of a cross-β structure. **b** Schematic representation of cross-β fibril structure. **c** EM image of amyloid fibril. This image is of fibrils formed by the SNNFGAILSS-NH_2 peptide fragment studied in Chapter "Effect of Terminal Capping on Aggregation of Peptide Fragments"

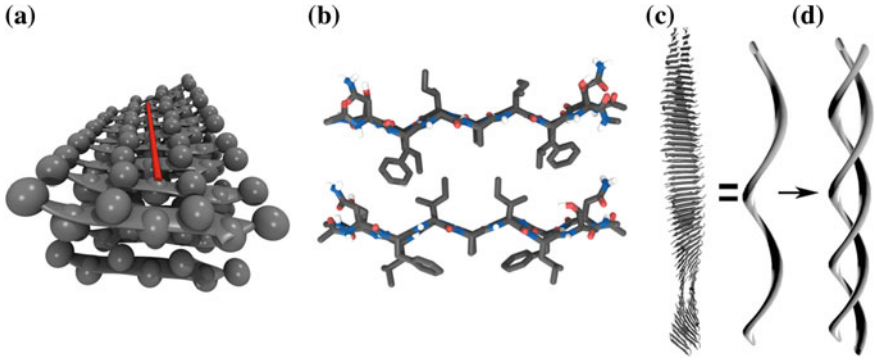


Fig. 2 Fibril structure. **a** Schematic representation of cross-β structure. The *red line* indicates a surface groove created by the adjacent side chains. **b** Steric zipper interactions are formed by interlocking side chains between the two β-sheets. The view is down the long axis of the fibril. Each layer consists of two antiparallel β-strands. **c** Two β-sheets twist to form a protofilament. **d** Intertwined protofilaments join to form a mature fibril

Peaks around 4.7 Å correspond to the regular distance between adjacent strands in a β-sheet, and peaks at 10–11 Å in the perpendicular direction correspond to the distances between stacked β-sheets [10]. The basic structure consists of long β-sheets composed of peptide monomers with the hydrogen bonds running parallel to the long axis of the fibril (Figs. 1b and 2a). The long β-sheets assemble to form a protofilament, which is the smallest structural unit of a fibril (Fig. 2c) [2]. The rope-like supramolecular structure can be seen in electron microscopy (EM) images (Fig. 1c).

The fibrils can become very long, up to several micrometers, and have a diameter of around 5–25 nm [11]. The interface between the β-sheets has interlocking side chains, termed a steric zipper (Fig. 2b), which excludes water [12]. The protofilaments combine laterally to form the mature fibril (Fig. 2d). The number of protofilaments in a fibril is usually around 2–6 [2]. Structural polymorphism can occur in fibrils formed from the same peptide. This means that fibrils formed from the same peptide have different supramolecular or molecular assemblies. This results in different appearances of the fibrils in EM images. The polymorphism can e.g. be a varying number or different packing of the protofilaments. It is common that different conditions during the fibril formation, e.g. pH or salt concentration, will lead to varying polymorphisms of the mature fibrils [11].

Structural Studies

The atomic resolution structures of amyloid fibrils have only recently become available. The X-ray fiber diffraction patterns do not provide detailed structural information about fibrils, and the traditional methods to determine high-resolution protein structures, such as X-ray crystallography and liquid state nuclear magnetic

resonance (NMR) spectroscopy, are not suited for examining amyloid fibrils due to their insoluble nature and large size [13]. X-ray crystal structures of amyloid fibril micro-crystals formed by short stretches of amyloid peptides have been determined, although they do not show the characteristic twist of amyloid fibrils [14–16]. The lack of fibril twist may be an effect of crystal packing interactions, as has been suggested based on molecular dynamics (MD) simulations [17]. Amyloid fibrils are, however, good subjects for solid-state NMR (ss-NMR) spectroscopy, which has made this the most successful method for determining the structure of amyloid fibrils [18]. Several structures of fibrils made from human islet amyloid polypeptide (hIAPP) and from full-length Amyloid β (Aβ), the peptides involved in T2DM and AD, respectively, have been determined [19–24]. Recently, an amyloid fibril structure seeded with in vivo fibrils from an AD patient was determined [24]. Interestingly, the in vivo fibril does not possess β-sheet structure, but does have the regular arrangement of the peptide strands perpendicular to the fibril axis. The surface of amyloid fibrils is characterized by the presence of long surface grooves created by the repeating pattern of side chains. The chemical properties of the grooves depend on the neighboring side chains. The β-sheets in these structures are all parallel, except a recent structure of the Aβ Iowa mutant (D23N) which shows anti-parallel β-sheets [25]. Amyloid fibrils formed by shorter peptides have also been shown to form anti-parallel arrangement of the β-sheets [26–30].

Amyloid Proteins

The proteins forming amyloid deposits exhibit a diverse range of native structures; proteins which are natively unfolded, α-helical proteins, β-sheet proteins, and mixed α-β-proteins have been observed to form amyloid fibrils in vivo [2]. This suggests that the ability to form amyloid fibrils is not connected to the native secondary structure. Furthermore, it has also been found that many proteins not observed to form amyloid in vivo are able to form amyloid fibrils in vitro [31]. It is simply a matter of finding the appropriate conditions. This led to the suggestion that the ability to form amyloid fibrils is a generic property of all peptides and proteins [31]. Since the formation of the cross-β structure only requires the backbone of the peptide, it is not surprising that all peptides seem able to form amyloid fibrils. The propensity for fibril formation is, however, dependent on the primary structure of the protein. It is governed by the balance of hydrophobicity, electrostatics, and secondary structure propensity of the peptides and proteins [32, 33].

Fibril Formation

The prevailing opinion is that the formation of amyloid fibrils occurs through a nucleation-dependent mechanism [34, 35]. This type of mechanism is well-known

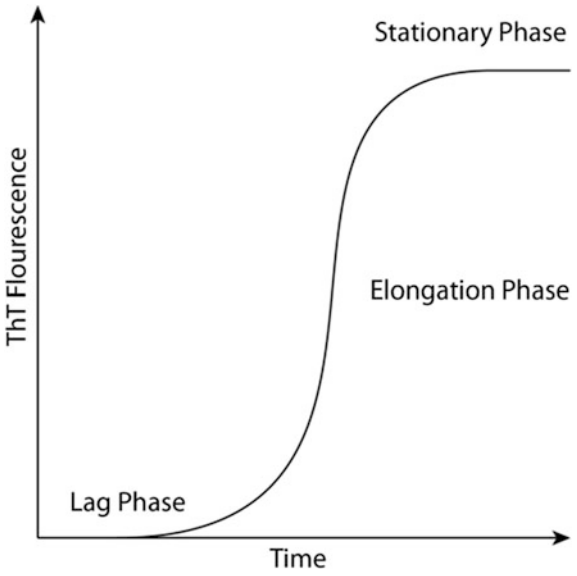


Fig. 3 ThT fluorescence curve of an amyloid fibril formation process

from the study of crystal-growth kinetics. The nucleation-dependent mechanism is evidenced by the sigmoidal curve in a Thioflavin T (ThT) fluorescence experiment probing the formation of amyloid fibrils from a solution of protein monomers [36]. ThT is an amyloid specific dye which shows orders of magnitude increase in fluorescence upon binding to amyloid fibrils [37]. The fibril formation process is composed of three phases; the lag-phase, the elongation-phase, and the stationary-phase (Fig. 3).

Different processes are dominant during the three phases. During the lag-phase the dominating process is the formation of monomers into small assemblies including a possible unfolding of the monomer. It is during this phase that the critical nucleus is formed [2]. The critical nucleus is the assembly needed to begin the elongation of the fibril. Once it is formed, the elongation-phase begins. During this phase the primary event is the elongation of fibrils from the critical nucleus via the addition of monomers (Fig. 4a). When the concentration of monomer in the solution approaches the critical concentration (C_C), below which fibril formation does not occur, the elongation of fibrils slows down, and the stationary-phase begins. Since the formation of the critical nucleus is the rate-limiting step during the lag-phase, the addition of fibril seeds from pre-formed fibrils can eliminate this phase [38].

Another mechanism of fibril formation, termed the double-concerted model, has been proposed (Fig. 4b) [39]. In this model the elongation of fibrils happens via the association of two or more oligomers to form the full fibril. This differs from the nucleation-dependent mechanism in the elongating unit; whereas the elongating unit in the nucleation-dependent model is the fibril monomer, it is an oligomer in

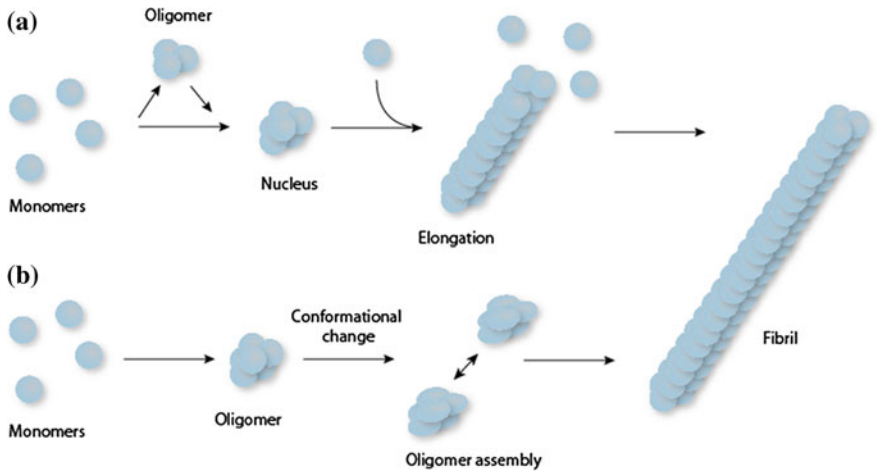


Fig. 4 Mechanisms of fibril formation. **a** Nucleation-dependent fibril formation. In this model, the monomers assemble into a critical nucleus which is the starting point for fibril formation. Monomers then add to the end of the elongating unit to form the full fibril. **b** Double-concerted fibril formation. The monomers associate to form an oligomer, which then undergoes a conformational change. The oligomers then assemble to form the full fibril

the double-concerted model. Evidence for this model comes from experiments showing markedly increased fibril formation rates by α-synuclein under initial conditions stabilizing a granular oligomer and subsequent agitating conditions [39].

It is likely that not only one mechanism is acting during the formation of amyloid fibrils. Multiple mechanisms may be acting in parallel or even cooperatively. This may be part of the reason for the observed fibril polymorphism as well as the difficulty in determining fibril formation mechanisms despite intense efforts.

Amyloid Toxicity

The initial theory regarding the toxicity of amyloid was that the presence of amyloid fibrils was responsible for causing amyloid diseases. This was termed the amyloid cascade hypothesis [40]. This theory was initially formulated for AD, but was later also adopted for the other amyloid diseases. However, the hypothesis was changed when it was realized that soluble oligomeric species of Aβ were much more toxic to cultured neurons than the mature fibrils [41]. During this time it was also shown that the amount of amyloid in AD brains did not correlate well with the progression of the disease [42]; however, the amount of soluble amyloid material in the brain tissue is correlated to the severity of the disease [43]. Furthermore, it was shown that amyloid deposition also occurred in the brain of some healthy individuals [44]. It is now believed that mature amyloid fibrils are a pool of inert protein

material, possibly arising from cellular protective mechanisms [45]. However, it is also very likely that the amyloid fibrils are in equilibrium with the soluble aggregates [46].

In the field of amyloid research the most studied disease is AD. Therefore, many experiments and theories have initially been performed and formulated for Aβ. The same experiments have later been performed for many of the other proteins and peptides linked to amyloid disease [47]. In 2002, this led to the hypothesis that cytotoxicity is a generic property of all amyloid oligomers [48]. Since then, a large effort has been put into finding the molecular entities or mechanisms responsible for the toxicity. Due to the complexity of cellular systems, many of the investigations have been performed in vitro, and the question is what the biological relevance is of the species or event in question. Further complicating the matter, the high polymorphism in structure and mechanism of the amyloid systems has also resulted in many different suggestions as to what the toxic species is.

One of the earliest proposed mechanisms was the formation of membrane pores by aggregating amyloid proteins [49]. This has been supported by observations of circular or tubular structures in EM and atomic force microscopy (AFM) images of oligomeric amyloid species [49–51] (Fig. 5), as well as membrane conductivity [49, 52] and vesicle leakage measurements [53, 54].

Receptor mediated effects of amyloid oligomers have also been proposed to lead to cell death [55], as well as cellular damage from reactive oxygen species following the interaction of Aβ with heavy metal ions [56]. However, it is unclear how this translates to other amyloid proteins as it has been shown that Zn(II) inhibits the fibril formation of hIAPP [57].

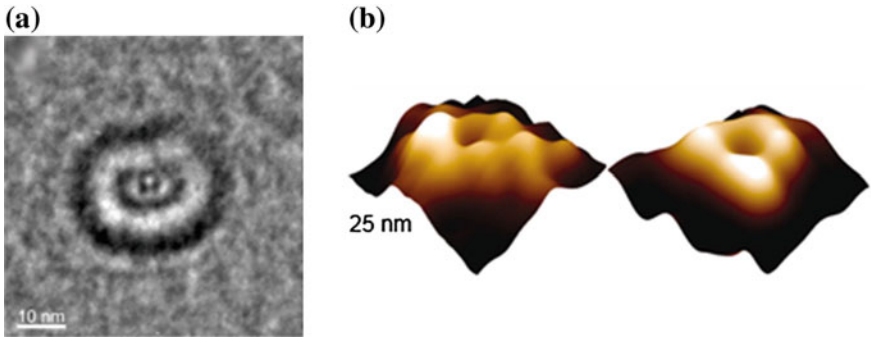


Fig. 5 Ring structures of hIAPP. **a** EM image. Reprinted with permission from Porat et al. [51] Biochemistry, 42, 10971. Copyright 2003 American Chemical Society. **b** AFM image. Reprinted with permission from Quist et al. [49] PNAS, 102, 10427. Copyright © by the National Academy of Sciences

Imaging

Alzheimer's Disease

AD is the main cause of dementia, accounting for between 60 and 80 % of dementia cases [58]. Symptoms of AD include loss of memory functions as the main indicator and reduction in speech, object recognition, motor activities, and abstract thinking as accessory symptoms [58]. The pathology of AD includes the deposition of extracellular Aβ and intracellular tau in the form of amyloid fibrils. This results in both the loss of synapses in the brain and neuronal cell damage and death [58]. It has been shown that the deposition of amyloid aggregates is already present in the early stages of AD before symptoms are detected [59, 60]. The precise pathology of AD has not been elucidated, which also means that no cure is available [2, 9, 61]. Furthermore, symptomatic treatment of the neurodegenerative amyloid diseases has not been very successful [62].

AD is hard to diagnose, since the neurodegenerative symptoms are similar between different forms of dementia [58]. Until recently, a positive diagnosis of AD could only be performed via a brain tissue autopsy of the diseased patient, and the diagnosis of probable AD was performed via observation of cognitive decline and elimination of the possibility of other forms of dementia [63, 64]. Recently, a new guideline for the diagnosis of AD was released by the International Working Group for New Research Criteria for the Diagnosis of Alzheimer's Disease [65]. The guideline includes the presence of an AD biomarker as a criterion in the diagnosis of AD [65]. Mainly two types of biomarkers exist, reduced levels of $A\beta_{1-42}$ and increased levels of tau protein in the cerebrospinal fluid (CSF), or an increased retention of an amyloid fibril imaging agent in a positron emission tomography (PET) scan of the brain [65]. Detection of biomarkers in the CSF is fairly invasive as it requires the insertion of a needle into the spine of the lower back. Therefore, detection of biomarkers using non-invasive imaging techniques such as PET or magnetic resonance imaging (MRI) is an area of intense research. This can help in the early identification of amyloid diseases, as well as in monitoring the progression of disease and the effects of disease modifying therapeutics [66].

In Vivo Amyloid Fibril Detection (Imaging)

The first amyloid binding dye was CR (**1**), which has been used since the 1920s to stain histological samples for the presence of amyloid fibrils (Chart 1). CR specifically binds amyloid fibrils followed by a change in optical properties and absorbance maximum of CR [6]. CR has been reported to affect the fibril formation process, and can therefore not be used as an in situ detection agent [67–70]. ThT (**2**) is the most widely used compound for the detection of amyloid fibrils in vitro and in situ [37] since ThT only affects the fibril formation process to a minor extent

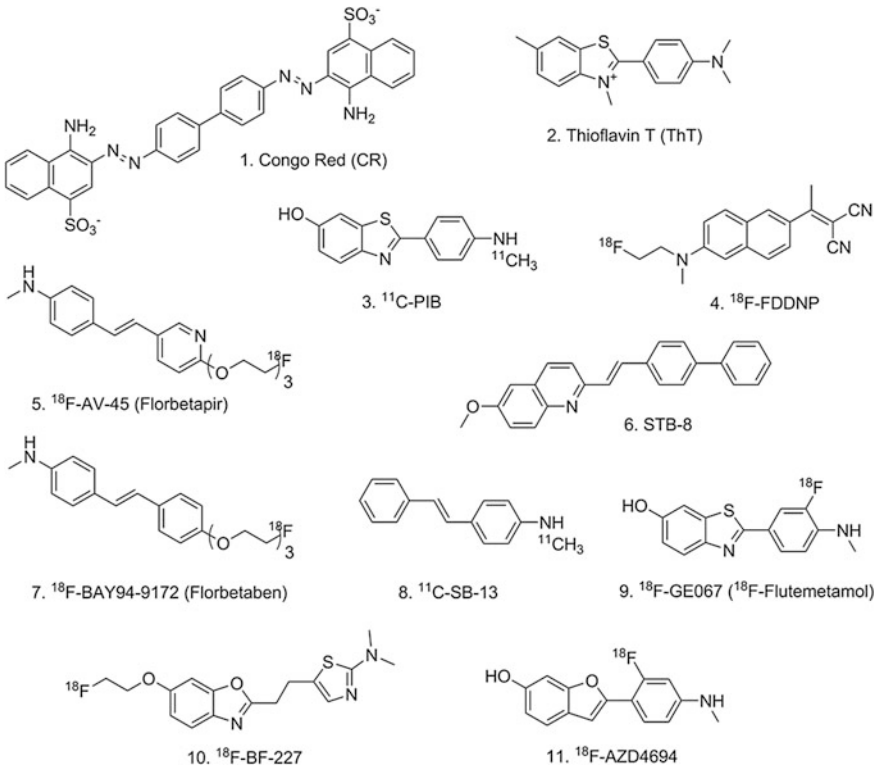


Chart 1 Amyloid imaging agents

[71–73]. Since amyloid specific dyes were already available, it is natural that they have been used as leads in the development of in vivo imaging agents. The requirements of an imaging agent to be used for the detection of brain amyloid are rigorous [74]. First and foremost, it should have high affinity and specificity for Aβ fibrils. It should be hydrophobic enough to cross the blood-brain barrier; however, it should not be so hydrophobic that it has high unspecific binding. Furthermore, the uptake and clearance of non-bound molecules in the brain should be fast to allow the possibility for fast detection which is especially important in PET. Pittsburgh Compound B [75] (^{11}C-PIB) (**3**) is based on the ThT scaffold, and is one of the most widely studied compounds in relation to the imaging of Aβ fibril content in the brain of AD patients [66]. Recently, a European multicenter collaboration has demonstrated the applicability of ^{11}C-PIB in the detection of AD [76]. ^{18}F-FDDNP [77] (**4**) is another compound which has shown promising abilities to bind amyloid fibrils specifically, and moreover, discriminate between the different stages of AD [78]. The stilbene series of compounds also includes some promising and successful compounds, such as ^{18}F-AV-45 (Florbetapir) (**5**) and ^{18}F-BAY94-9172 (Florbetaben) (**7**), which have both recently been approved by the American Food

and Drug Administration (FDA) for the use in PET imaging of brain amyloid content [79, 80]. Furthermore, STB-8 (**6**) has shown promising properties such as excellent blood-brain barrier permeation and specific staining of amyloid fibrils in mouse model studies [81].

Several compounds have been tested in humans and are undergoing clinical trials, including ^{11}C-PIB (**3**), ^{18}F-FDDNP (**4**), ^{11}C-SB-13 (**8**), ^{18}F-GE067 (^{18}F-Flutemetamol) (**9**), ^{18}F-BF-227 (**10**), and ^{18}F-AZD4694 (**11**) (Chart 1) [66, 82–84]. Even though many of these imaging agents differ in the central scaffold, they all contain aromatic moieties, and are elongated and fairly rigid. Due to the difference in half-lives between ^{11}C ($\sim$20 min) and ^{18}F ($\sim$110 min), the fluorine compounds will probably be the future for routine imaging of brain amyloid.

Binding of Imaging Agents

The structural details of the complexes between imaging agents and amyloid fibrils have been difficult to study due to the insoluble nature of amyloid fibrils, although some experimental evidence has been obtained. ThT is so far the most studied compound, and the leading model of ThT binding to amyloid fibrils is the channel model [37, 85]. In this model ThT is placed parallel to the fibril elongation axis in the long surface grooves created by the protruding side chains (red line in Fig. 2a). Based on ss-NMR measurements, ThT has also been suggested to bind parallel to the β-sheet normal [86]. Dimers binding between the protofilaments (such as between the two ribbons in Fig. 2d) were suggested to be the fluorescent species. This theory was based on ThT binding in cavities of different sizes and the subsequent presence or absence of ThT fluorescence [87]. In accordance with several proposed models of ThT binding, the presence of multiple binding sites for the ThT series of compounds has been observed [86, 88].

The binding of CR to amyloid fibrils has also been studied. Evidence for binding parallel to the elongation axis of the fibril analogously to the channel model of ThT has been observed [89, 90]. The distance between the two anionic charges in CR is approximately 19 Å and was shown to be important for binding to amyloid fibrils [91, 92]. The approximate distance between five β-strands in a β-sheet is also 19 Å, which provides the opportunity for the negative charges on CR to interact with positive or polar side chains on strand i and $i + 4$. Stabilization of the charges was shown to be crucial for CR binding to the HET-s prion amyloid fibril. A rationally designed Lys to Ala mutation based on ss-NMR measurements and subsequent docking calculations eliminated the binding of CR to the fibril [93].

Amyloid binding imaging agents can be divided into three different classes based on which molecule they compete with for binding on Aβ fibrils; the CR class, the ThT class, and the FDDNP class of compounds [91]. The CR and ThT classes of compounds have been shown to be non-competitive inhibitors with respect to each other [94]. A distinct binding site for FDDNP has also been observed [95]. The stilbenes have been shown to share the same binding site as ThT [91].

Several MD simulation studies of ligand binding to various amyloid fibrils have been performed during recent years. The studies include dyes [96], imaging agents [97, 98], and potential aggregation inhibitors [99–101]. These studies show that the investigated amyloid dyes and imaging agents bind in the surface grooves, as was also proposed from experiments with ThT and CR. The aggregation inhibitors show some preference for the fibril ends, and may thereby disrupt further elongation of the fibril.

hIAPP

Type 2 Diabetes Mellitus

T2DM, also called late-onset diabetes, is characterized by an abnormal or missing response to an elevated blood-glucose level [102]. In normal glucose regulation, an elevation in blood-glucose, such as after a meal, causes insulin secretion by the pancreatic β-cells. Insulin then acts on various tissues such as brain, muscle, adipose, and liver tissue, to induce the uptake of glucose from the blood [102]. Reduced ability to respond to insulin induces the production and secretion of elevated levels of insulin; however, when the β-cells fail to produce sufficient amounts of insulin, T2DM arises [102]. In diabetic individuals, the response to an elevated blood-glucose level can be defective for mainly two reasons; a resistance to insulin by the cells taking up glucose or a β-cell malfunction, which can include either a reduction in β-cell mass or function [102]. The major hallmark of T2DM is the deposition of amyloid plaques formed mainly by hIAPP in the islets of Langerhans in the pancreas [103]. The number of T2DM patients with pancreatic islet amyloid deposition has been reported to be anywhere from 40 to 100 % [103]. It seems that deposition of islet amyloid is not the main cause of T2DM, but a delayed event in the pathophysiology of T2DM, which could explain the varying occurrence of amyloid in T2DM patients [102].

Physiological Role of hIAPP

In 1987 it was discovered by two independent groups that hIAPP is the major component of islet plaques in T2DM patients [104, 105]. It is a 37-residue peptide hormone produced and secreted by β-cells in the pancreas [106]. It is stored in secretory granules along with insulin and secreted upon elevated glucose in the blood [107]. The physiological function has not been completely elucidated, although it is clear that hIAPP is involved in maintaining glucose homeostasis, suppression of glucagon release, and controlling satiety [103, 108]. Some calcitonin receptors have also been shown to have low affinity for hIAPP. However, this can

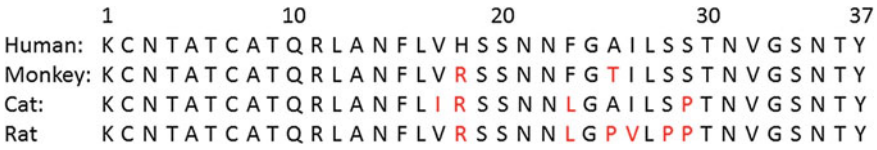


Fig. 6 Peptide sequences of hIAPP and IAPP from several other mammals. Humans, non-human primates, and cats form islet amyloid in vivo. Rats and mice have identical IAPP sequences, and do not develop diabetes [111, 112]

be modulated and increased when the calcitonin receptor is co-expressed with and binds to receptor activating-modifying proteins [109].

hIAPP is translated as a 89-residue pre-pro peptide [110]. A 22-residue signal peptide is then cleaved to form the 67-residue proIAPP. This peptide is then further processed to form the 37-residue hIAPP, which has a C-terminal amidation as well as a disulphide bridge between Cys2 and Cys7 [110]. hIAPP is a member of the calcitonin related family of peptides, and the sequence of hIAPP can be seen in Fig. 6 along with the sequence of IAPP from several other species. IAPP from rat and mouse does not form amyloid fibrils, which has been attributed to the three Pro-substitutions in the 20–29 residue region. This region has been shown to be highly important for fibril formation [111].

hIAPP Monomer

hIAPP is a natively disordered peptide. This does not mean that it is a random coil; it only means that it is not folded into a well-defined globular structure. NMR experiments have shown that a transient helical structure exists for around 40 % of the peptide in solution between residues 6–17, and that stabilization of this helix leads to acceleration of fibril formation [113]. Furthermore, an X-ray crystal structure of hIAPP fused to the maltose binding protein showed a helical dimer with interactions between the two N-terminal helices [114]. Mutations constructed based on the dimer interface showed that disruption of the interface led to slower aggregation kinetics, while mutations predicted to stabilize the interface accelerated the fibril formation [114]. This led to the suggestion that a helical intermediate exists in the fibril formation pathway [113–115].

hIAPP Monomer-Membrane Interactions

When bound to a phospholipid membrane hIAPP displays about 39–43 % helical structure as estimated by circular dichroism (CD) spectroscopy of hIAPP bound to DOPG (1,2-dioleoyl-sn-glycero-3-phosphatidylglycerol) liposomes or 40–50 %

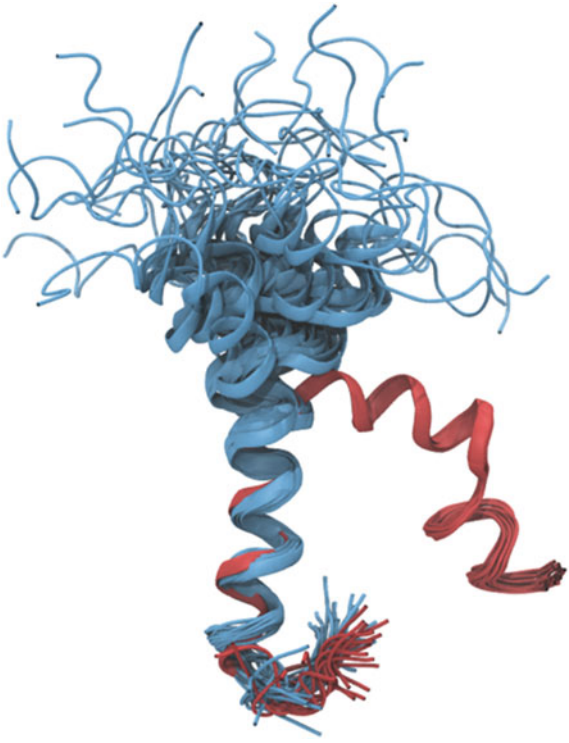


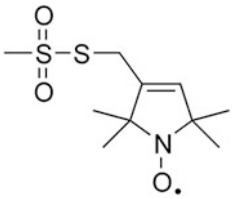
Fig. 7 Ensembles of ss-NMR structures from Patil et al. [119] (PDB-code: 2KB8) (*blue*), and Nanga et al. [118] (PDB-code: 2L86) (*red*). The structures have been aligned on C_α of residues 1–19

when bound to DOPS (1,2-dioleoyl-sn-glycero-3-phosphatidylserine) large unilamellar vesicles [54, 116]. NMR spectroscopy has also provided structural models of $hIAPP_{1–19}$ and the full peptide bound to SDS micelles [117–119]. $hIAPP_{1–19}$ is α-helical with a small kink at the N-terminal disulphide bridge [117]. The full-length peptide also shows mainly α-helical character with a kink at His18 [118]. An older structure from 2009 did not show the kink at His18, and the ensemble of low energy structures released in the study revealed flexibility in the termini of the peptide (Fig. 7) [119]. However, the structure from 2009 was determined at low pH (4.6) and without the N-terminal amidation.

The more recent structure from 2011 was determined at pH 7.3 and the N-terminal amidation was present [118]. The ensemble of structures determined in this study was much more coherent, and did not show much variability (Fig. 7).

Attempts have been made to determine the orientation of the peptide in a membrane. Using site-directed spin-labeling and electron paramagnetic resonance (EPR) spectroscopy it was found that residues 9–20 show higher order than the rest of the peptide [120]. Furthermore, accessibility measurements in the same study indicated the presence of an amphipathic helix spanning from residues 9–22 [120]. One issue with this study is the 80:20 DOPS:DOPC (1,2-dioleoyl-sn-glycero-3-phosphatidylcholine) lipid composition of the membranes used, which is too high a content of anionic (PS) lipids compared to the optimum for fibril formation [116].

Chart 2 MTSL



The high anionic lipid content was used to stabilize the monomer conformation, which questions the usefullness of the results in the context of fibril formation. Furthermore, spin-labeling is performed by mutating each residue to cysteine and allowing the thiol group to react with a spin-label, which in this case was the bulky MTSL (Chart 2).

This drastic change of the peptide is also likely to have an effect on the structure and position of the peptide in the membrane. Using paramagnetic quenching NMR experiments, Nanga et al. showed that the burial of $hIAPP_{1–19}$ in SDS micelles is dependent upon the protonation state of His18 [117]. When His18 is deprotonated, $hIAPP_{1–19}$ is buried within the micelle, while protonation of His18 reorients the peptide to the surface of the micelles. It should be noted though, that a micelle is not an appropriate mimic for a membrane since the curvature of a micelle is much higher than a membrane, and this type of curvature is not likely found in cell membranes. Furthermore, it has been shown that the curvature of the membrane affects the localization and aggregation properties of hIAPP [121, 122].

hIAPP Oligomers

Determination of the exact cause of hIAPP toxicity will enable the design of therapeutic strategies. In 1999 it was shown, by adding freshly dissolved hIAPP monomers or mature fibrils to a β-cell culture, that β-cell toxicity occurred before amyloid fibrils could be detected [123]. Since then, the search for the cytotoxic species has been ongoing. However, the study of amyloid intermediates is very difficult due to the fast aggregation of the peptides. Therefore, the structural information regarding hIAPP aggregation intermediates is very sparse, and much of the information is indirect.

It is important to determine which species is present in the experimental mixture, and also that it is stable. It was shown using light scattering and vesicle permeation assays that the incubation of various initial hIAPP species with vesicles increased the permeation of Ca^{2+} over time [124]. It was shown in this study that the insoluble amyloid fibrils did not show initial permeabilizing activity; however, incubation did increase the permeabilization, suggesting that the mixture of species is changing constantly [124]. Porat et al. showed by a colorimetric assay that the permeabilizing activity for soluble hIAPP was greatest after around 1 h of incubation, and by transmission electron microscopy (TEM) that a circular oligomeric species was

present in the mixture (Fig. 5a) [51]. It was then proposed that the membrane permeabilization occurs through a pore-like mechanism. The pore-like mechanism was further supported by AFM observations of circular oligomeric species for several amyloid peptides including hIAPP, Aβ, and α-synuclein (Fig. 5b) [49].

The size of the oligomeric species, meaning the number of monomers in the oligomer, varies quite significantly and has been reported to be anywhere between 16 and 500 monomers [123–126] The detection of hIAPP oligomers in pancreatic tissue samples has been performed indirectly by the use of an oligomer specific antibody [127]. This antibody was developed by injecting a rabbit with a molecular mimic of an Aβ oligomer. It was subsequently shown that the antibody did not react to Aβ monomers or mature fibrils [128]. It was also shown that the antibody, termed A11, specifically reacts with oligomers from other amyloidogenic peptides, including hIAPP [128]. However, problems with using A11 to detect oligomers in tissue samples have been suggested, as the antibody has also been shown to react with several proteins with a natively globular fold [129]. Therefore, until it has been conclusively shown that the antibody is actually reacting to only the oligomer in pancreatic tissue samples, results obtained using the A11 antibody should be handled with care.

Whether the toxic species is on- or off-pathway is a matter of controversy. It was suggested that the off-pathway species was cytotoxic based on experiments where Rifampicin inhibited the formation of amyloid fibrils [130]. However, it was later realized that Rifampicin does not actually inhibit the formation of fibrils; it merely interferes with the ThT fluorescence assay [131]. This highlights that it is important to verify the presence or absence of amyloid fibrils by more than one method, preferably using a direct method like EM or AFM.

Toxicity of Growing Amyloid Fibrils

Recently, it was suggested that hIAPP toxicity is caused by the growing fibrils during amyloid formation [132]. This was based on observations that changes in experimental conditions which affected fibril formation rate, such as the addition of seeds or changing peptide concentration, also affected the leakage from large unilamellar vesicles in the same manner [132]. Inhibition of fibril formation by insulin also inhibited vesicle leakage, and cryo-EM images showed the association of amyloid fibrils with distorted vesicles, which also indicated that the fibrils grow on the membrane [132]. Lipid extraction from the membrane by the growing fibril was suggested to be the mechanism by which the growth disrupts the membrane. This was further supported by the observation that lipids were present in amyloid fibrils forming in vitro and in amyloid plaques [133, 134]. Support for this hypothesis also came from MD simulations using a very simple amyloid peptide model. The simulation revealed the self-assembly of peptides on a lipid vesicle, which showed leakage of the content upon fibril formation [135].

What Causes Toxicity?

This is probably the most controversial and studied question in amyloid research at the moment.

Evidence supporting a toxic oligomer hypothesis has been presented, and so has evidence for toxicity during amyloid formation. However, one does not rule out the other. Since the complete fibril formation mechanism has not been elucidated, it is not possible to rule out the formation of an oligomer during amyloid aggregation, which would actually constitute a toxic event during the fibril formation. Another explanation for the multitude of different reports regarding toxicity could be that multiple mechanisms are occurring simultaneously. This may very well be the case since amyloid formation is a pathologic event governed by a thermodynamic and kinetic interplay, and not a physiological mechanism which has been optimized during evolution. Several other mechanisms besides oligomers and the actual fibril formation have been suggested as toxic, e.g. endoplasmic reticulum stress, cholesterol accumulation, islet inflammation, and oxidative stress [136].

hIAPP Fibrils

The stability of amyloid fibrils has made them easier to handle than the monomer or intermediate forms of hIAPP. However, they are insoluble, making them difficult to study with traditional methods for structure determination, such as liquid-state NMR and X-ray crystallography. ss-NMR, however, has been successful in determining structural models for many amyloid fibrils. A combination of ss-NMR and restrained Langevin dynamics resulted in a model of an amyloid fibril formed by full-length hIAPP (Fig. 8) [22].

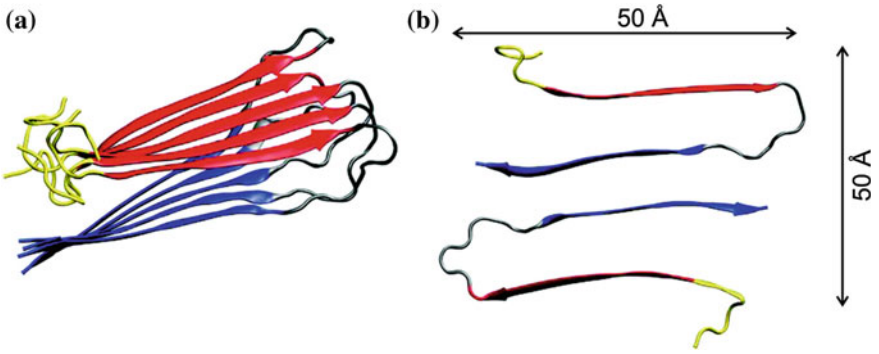


Fig. 8 ss-NMR structural model of hIAPP. **a** The two β-sheet layers are seen in *blue* and *red*. **b** The fibril seen from the end. Reprinted with permission from Luca et al. [22] Biochemistry, 46, 13505. Copyright 2007 American Chemical Society

The peptides in this structure form a double layered parallel β-sheet, with residues 8–17 forming the first sheet (blue), residues 28–37 forming the second sheet (red), and residues 18–27 forming a loop or bend between the two sheets (Fig. 8). Based on mass-per-length scanning tunneling electron microscopy (STEM) measurements, the model was constructed with two peptides in each repeating unit, forming steric zipper interactions between the two units [22]. Another fibril model of full-length hIAPP was proposed based on EPR spectroscopy and a simulated annealing protocol [23]. This model also contains two parallel β-sheets spanning residues 12–19 and 31–36 joined in a hairpin-like double-layer by a loop region spanning residues 20–30. The differences in the residues spanning the β-sheet regions in fibril structures could be a manifestation of the different morphologies seen in EM images of the two structures. The first structure exhibits striated ribbon morphology [22], and the second shows a twisted morphology [23]. The hairpin-like structure may be generic for fibril structures formed by long peptides, and has also been observed for fibrils formed by e.g. Aβ and α-synuclein [18].

Residues 20–29 of hIAPP have long been known to be important for fibril formation, and were originally thought to comprise the fibrillating core of hIAPP [111]. In the two structural models described above, residues 20–29 are placed in the loop region between the two β-sheet regions. Recent reports have suggested that the initial structural feature formed during fibril formation is the turn region [137], which would explain the importance of these residues.

Structural studies of amyloid fibrils formed from smaller peptide fragments have also been extensively performed. Smaller fragments of hIAPP can be grown into microcrystal forms and structurally determined using X-ray crystallography [15]. These fragments usually form a double-layered β-sheet with steric zipper interactions between the two sheets [15].

Conditions Affecting hIAPP Fibril Formation

An important aspect of determining what causes cytotoxicity is elucidating the fibril formation mechanism and the impact of changes in the environment of the peptide.

Anionic lipids accelerate the formation of amyloid fibrils by reducing the lag-phase of fibril formation [138]. It was shown by ThT fluorescence measurements of the fibril formation that the maximum acceleration is obtained at a 25 mol% of PS lipids [116]. This correlates with the finding that a high glucose environment causes an increase in anionic lipid content of the β-cell membrane [139, 140]. Rat pancreatic islet tumor cells showed an anionic lipid content in control cells of 16 mol% while the content in the high glucose cells had risen to 21 mol% [139].

The pH of the environment also affects the fibril formation kinetics. Both in the presence and absence of lipid membranes, a lowering of the pH will inhibit the fibril formation [141, 142]. The pH of the β-cell secretory granules is around 5.5 [143], which will significantly inhibit the fibril formation of hIAPP. It is also known that

insulin inhibits the fibril formation both in the soluble as well as in the crystalline form found in the secretory granules of the β-cells [144, 145]. The low pH and the co-localization of hIAPP and insulin in the β-cell secretory granules probably serve to protect hIAPP from aggregating inside the granules.

Several studies have shown that hIAPP most likely inserts into the membrane as a monomer. This was initially suggested based on an increase in surface tension of a 7:3 mixture of DOPC:DOPS phospholipid monolayer immediately after addition of a fresh solution of hIAPP or rIAPP to the aqueous phase [146]. The increase in surface tension when hIAPP interacts with a lipid monolayer suggests that it is actually inserting, and not just associating with the membrane surface [146]. Quartz Crystal Microbalance with Dissipation Monitoring experiments also indicate that soluble hIAPP rapidly binds to a supported lipid bilayer of a 4:1 DOPC:DOPS mixture after addition to the solution [147]. This study also showed, by cross-linking parts of the membrane, that aggregation occurs at the fluid parts of the membrane, and by photo bleaching experiments that addition of soluble hIAPP to a membrane reduces the fluidity of the membrane [147].

At neutral pH, hIAPP contains two positively charged side chains, Lys1 and Arg11, and a positively charged N-terminus. It has been suggested that the initial interaction between negatively charged lipids and hIAPP is mediated by the positive charges in the N-terminus of the peptide based on surface tension measurements [146]. The increase in surface tension of a phospholipid monolayer was independent of the chosen anionic lipid (DOPG or DOPS), and was also independent of the presence of the C-terminal 20–37 residues. Furthermore, it was shown that $hIAPP_{20–29}$ affects the surface tension to a much lesser extent, and therefore does not insert into the monolayer to the same extent as the full peptide [146]. The fact that an increased salt concentration can slow down the fibril formation in the presence of anionic lipids also points to the importance of electrostatics in the interaction of hIAPP with the membrane [116]. By performing a Lys1Glu mutation, it was shown that the fibril formation was accelerated compared to the wild type protein, and the presence of an anionic lipid membrane did not accelerate the fibril formation of the mutant to the same extent as the wild type [138]. The nature of the anionic lipid is not as important as the actual presence of a negatively charged lipid [122]. The same decrease in lag-time was observed for both DOPG and DOPS lipids, suggesting that it is the electrostatic interaction that is important for the aggregation, and not specific interactions [122]. Evidence suggests that the role of the anionic lipid is to both increase the local concentration of hIAPP and to neutralize the positive charges on hIAPP, allowing the peptides to interact more closely.

The model membrane systems used for the experiments have been probed, and the nature of the model was shown to affect the rate of aggregation. Bicelles, which are disc-like lipid assemblies, and small unilamellar vesicles showed a higher propensity for accelerating the aggregation than large unilamellar vesicles, suggesting that either the size of the membrane model or the curvature of the membrane is important for fibril formation [122]. A study showing that hIAPP favors

binding to regions of high curvature in bicelles, also shows that membrane curvature is highly important for hIAPP [121].

MD Studies of hIAPP-Membrane Interactions

MD simulations have proven useful to study the detailed mechanisms of amyloid fibrils. Since the amyloid peptide systems are so difficult to handle experimentally, MD simulations are an important contributor to understanding the mechanisms of amyloid fibril formation.

Zhao et al. set out to describe the membrane pores observed in AFM experiments (Fig. 5b) [49, 148]. They used the amyloid fibril structure formed by the full-length hIAPP to construct a β-sheet membrane pore composed of hairpin-like peptide monomers. They found that the peptide monomers assemble into multiple domains or subunits resembling the AFM images, and they were able to suggest a possible number of monomers present in one membrane pore [148]. The assumptions about the pore structure made in this study are substantial, as no detailed structural information is available on the secondary structure or the intermolecular assembly of the membrane permeabilizing species of hIAPP.

Other studies have also made assumptions on the membrane bound species of hIAPP, however less severe than the abovementioned example [149, 150]. Zhang et al. used the ss-NMR structure of hIAPP determined at low pH to investigate the interactions of monomer and dimer hIAPP with a POPG membrane [118, 150]. They found that electrostatic interactions play an important role in the hIAPP-membrane interaction, and that the dimer structure perturbed the membrane to a greater extent than the monomer. The peptide was initially inserted with the N-terminus in the membrane and was perpendicular to the membrane layer. This is contrary to the finding by Apostolidou et al. [120] who showed that the N-terminus is located parallel to the membrane layer. Duan et al. [149] however, simulated the parallel membrane binding of the wild-type and the Ser20Gly mutant of hIAPP using a fragment containing residues 1–25. The Ser20Gly mutant is associated with early on-set of T2DM and is found to have greater amyloidogenic and cytotoxic properties than wild-type hIAPP [151, 152]. The initial peptide structure was built with α-helical conformation, and was not an experimentally determined structure. They found that the Ser20Gly mutant adopts an L-shaped motif, which they speculate is the reason for the faster aggregation since it is structurally closer to the hairpin-like motif of hIAPP in the full amyloid fibril than an extended helical structure [149].

The all-atom simulation that is most similar to the common experimental conditions was performed by Jia et al. [153]. They simulated the process of binding of an hIAPP monomer to a POPG membrane bilayer with different initial orientations of the peptide above the membrane layer. The initial protein structure was taken from the ss-NMR study at low pH [119]. They found that the initial attraction between the membrane and hIAPP was through the N-terminal positive residues, and that the peptide oriented in a parallel fashion with respect to the membrane

[153]. A stable helix was present for residues 7–22. The conditions in this study resembled the conditions used for determining properties of the hIAPP monomer in experiments. It is common to use a membrane composed of purely anionic lipids; yet, this has been shown to inhibit the fibril formation [116]. Thus, Jia et al. have probably probed the conformation of the stable hIAPP monomer, and have not actually investigated the aggregation prone monomer conformation.

Coarse-grained (CG) MD simulations have been used to study the aggregation of hIAPP. Xu et al. used the MARTINI force field (FF) [154, 155] to study the self-assembly of cholesterol, lipids, and $hIAPP_{1-19}$ [156]. However, to use the MARTINI protein FF successfully, an elastic network model is used to stabilize the secondary structure of the peptide, because of the missing hydrogen bonds [157]. It is therefore not possible to see protein conformational changes using this model. Xu et al. assumed that hIAPP maintains an α-helical conformation during the assembly process [156]. They find that the presence of cholesterol pushes hIAPP out of the membrane, while the absence of cholesterol allows the formation of a membrane pore.

As this chapter summarizes, the knowledge of amyloid proteins is sparse, and a multitude of hypotheses have been proposed. The coming years will prove important in the pursuit of understanding amyloid proteins, and eventually of developing treatments for the diseases.

References

1. Branden C, Tooze J (1998) In Introduction to protein structure, 2nd edn. Garland
2. Chiti F, Dobson CM (2006) Protein misfolding, functional amyloid, and human disease. Ann Rev Biochem 75:333–366
3. Gorevic PD, Munoz PC, Casey TT, DiRaimondo CR, Stone WJ, Prelli FC, Rodrigues MM, Poulik MD, Frangione B (1986) Polymerization of intact β2-microglobulin in tissue causes amyloidosis in patients on chronic hemodialysis. Proc Natl Acad Sci USA 83:7908–7912
4. Virchow R (1854) Zur cellulose—frage. Virchows arch. Pathol Anat 6:416–426
5. Westermark P, Benson MD, Buxbaum JN, Cohen AS, Frangione B, Ikeda S, Masters CL, Merlini G, Saraiva MJ, Sipe JD (2005) Amyloid: toward terminology clarification report from the nomenclature committee of the international society of amyloidosis. Amyloid 12:1–4
6. Howie AJ, Brewer DB (2009) Optical properties of amyloid stained by Congo red: history and mechanisms. Micron 40:285–301
7. Alexandrescu AT (2005) Amyloid accomplices and enforcers. Protein Sci 14:1–12
8. Geddes AJ, Parker KD, Atkins EDT, Beighton E (1968) "Cross-β" conformation in proteins. J Mol Biol 32:343–358
9. Harrison RS, Sharpe PC, Singh Y, Fairlie DP (2007) Amyloid peptides and proteins in review. Rev Physiol Biochem Pharmacol 159:1–77
10. Sunde M, Serpell L, Bartlam M, Fraser PE, Pepys MB, Blake CCF (1997) Common core structure of amyloid fibrils by synchrotron X-ray diffraction. J Mol Biol 273:729–739
11. Fändrich M (2007) On the structural definition of amyloid fibrils and other polypeptide aggregates. Cell Mol Life Sci 64:2066–2078
12. Greenwald J, Riek R (2010) Biology of amyloid: structure, function, and regulation. Structure 18:1244–1260

13. Shewmaker F, McGlinchey RP, Wickner RB (2011) Structural insights into functional and pathological amyloid. J Biol Chem 286:16533–16540
14. Nelson R, Sawaya MR, Balbirnie M, Madsen AØ, Riekel C, Grothe R, Eisenberg D (2005) Structure of the cross-β spine of amyloid-like fibrils. Nature 435:773–778
15. Sawaya MR, Sambashivan S, Nelson R, Ivanova MB, Sievers SA, Apostol MI, Thompson MJ, Balbirnie M, Wiltzius JJW, McFarlane HT, Madsen AØ, Riekel C, Eisenberg D (2007) Atomic structures of amyloid cross-β spines reveal varied steric zippers. Nature 447:453–457
16. Apostol MI, Sawaya MR, Cascio D, Eisenberg D (2010) Crystallographic studies of prion protein (PrP) segments suggest how structural changes encoded by polymorphism at residue 129 modulate susceptibility to human prion disease. J Biol Chem 285:29671–29675
17. Periole X, Rampioni A, Vendruscolo M, Mark AE (2009) Factors that affect the degree of twist in β-sheet structures: a molecular dynamics simulation study of a cross-β filament of the GNNQQNY peptide. J Phys Chem B 113:1728–1737
18. Tycko R (2011) Solid state NMR studies of amyloid fibril structure. Ann Rev Phys Chem 62:279–299
19. Petkova AT, Ishii Y, Balbach JJ, Antzutkin ON, Leapman RD, Delaglio F, Tycko R (2002) A structural model for Alzheimer's β-amyloid fibrils based on experimental constraints from solid state NMR. Proc Natl Acad Sci USA 99:16742–16747
20. Paravastu AK, Leapman RD, Yau W, Tycko R (2008) Molecular structural basis for polymorphism in Alzheimer's β-amyloid fibrils. Proc Natl Acad Sci USA 105:18349–18354
21. Bertini I, Gonnelli L, Luchinat C, Mao J, Nesi A (2011) A new structural model of $A\beta_{40}$ fibrils. J Am Chem Soc 133:16013–16022
22. Luca S, Yau W, Leapman R, Tycko R (2007) Peptide conformation and supramolecular organization in amylin fibrils: constraints from solid-state NMR. Biochemistry 46: 13505–13522
23. Bedrood S, Li Y, Isas JM, Hegde BG, Baxa U, Haworth IS, Langen R (2012) Fibril structure of human islet amyloid polypeptide. J Biol Chem 287:5235–5241
24. Lu J, Qiang W, Yau W, Schwieters C, Meredith S, Tycko R (2013) Molecular structure of β-amyloid fibrils in Alzheimer's disease brain tissue. Cell 154:1257–1268
25. Qiang W, Yau W, Luo Y, Mattson MP, Tycko R (2012) Antiparallel β-sheet architecture in iowa-mutant β-amyloid fibrils. Proc Natl Acad Sci USA 109:4443–4448
26. Nielsen JT, Bjerring M, Jeppesen MD, Pedersen RO, Pedersen JM, Hein KL, Vosegaard T, Skrydstrup T, Otzen D, Nielsen NC (2009) Unique identification of supramolecular structures in amyloid fibrils by solid-state NMR spectroscopy. Angew Chem Int Ed 48: 2118–2121
27. Balbach JJ, Ishii Y, Antzutkin ON, Leapman RD, Rizzo NW, Dyda F, Reed J, Tycko R (2000) Amyloid fibril formation by $A\beta_{16\text{-}22}$, a seven-residue fragment of the Alzheimer's β-amyloid peptide, and structural characterization by solid state NMR. Biochemistry 39:13748–13759
28. Petkova AT, Buntkowsky G, Dyda F, Leapman RD, Yau WM, Tycko R (2004) Solid state NMR reveals a pH-dependent antiparallel β-sheet registry in fibrils formed by a β-amyloid peptide. J Mol Biol 335:247–260
29. Kammerer RA, Kostrewa D, Zurdo J, Detken A, Garcia-Echeverria C, Green JD, Muller SA, Meier BH, Winkler FK, Dobson CM, Steinmetz MO (2004) Exploring amyloid formation by a de novo design. Proc Natl Acad Sci USA 101:4435–4440
30. Bu Z, Shi Y, Callaway DJE, Tycko R (2007) Molecular alignment within β-sheets in Aβ14-23 fibrils: solid-state NMR experiments and theoretical predictions. Biophys J 92: 594–602
31. Chiti F, Webster P, Taddei N, Clark A, Stefani M, Ramponi G, Dobson CM (1999) Designing conditions for in vitro formation of amyloid protofilaments and fibrils. Proc Natl Acad Sci USA 96:3590–3594

32. DuBay KF, Pawar AP, Chiti F, Zurdo J, Dobson CM, Vendruscolo M (2004) Prediction of the absolute aggregation rates of amyloidogenic polypeptide chains. J Mol Biol 341: 1317–1326
33. Chiti F, Stefani M, Taddei N, Ramponi G, Dobson CM (2003) Rationalization of the effects of mutations on peptide and protein aggregation rates. Nature 424:805–808
34. Naiki H, Hashimoto N, Suzuki S, Kimura H, Nakakuki K (1997) Establishment of a kinetic model of dialysis-related amyloid fibril extension in vitro. Amyloid 4:223–232
35. Serio TR, Cashikar AG, Kowal AS, Sawicki GJ, Moslehi JJ, Serpell L, Arnsdorf MF, Lindquist SL (2000) Nucleated conformational conversion and the replication of conformational information by a prion determinant. Science 289:1317–1321
36. Bhak G, Choe Y, Paik SR (2009) Mechanism of amyloidogenesis: nucleation-dependent fibrillation versus double-concerted fibrillation. BMB Rep 42:541–551
37. Biancalana M, Koide S (2010) Molecular mechanism of Thioflavin-T binding to amyloid fibrils. Biochim Biophys Acta Proteins Proteomics 1804:1405–1412
38. Dobson CM (2003) Protein folding and misfolding. Nature 426:884–890
39. Bhak G, Lee J, Hahn J, Paik SR (2009) Granular assembly of α-synuclein leading to the accelerated amyloid fibril formation with shear stress. PLoS ONE 4:e4177
40. Hardy JA, Higgins GA (1992) Alzheimer's disease: the amyloid cascade hypothesis. Science 256:184–185
41. Lambert MP, Barlow AK, Chromy BA, Edwards C, Freed R, Liosatos M, Morgan TE, Rozovsky I, Trommer B, Viola KL, Wals P, Zhang C, Finch CE, Krafft GA, Klein WL (1998) Diffusible, nonfibrillar ligands derived from $A\beta_{1-42}$ are potent central nervous system neurotoxins. Proc Natl Acad Sci USA 95:6448–6453
42. Terry RD, Masliah E, Salmon DP, Butters N, DeTeresa R, Hill R, Hansen LA, Katzman R (1991) Physical basis of cognitive alterations in Alzheimer's disease: synapse loss is the major correlate of cognitive impairment. Ann Neurol 30:572–580
43. McLean CA, Cherny RA, Fraser FW, Fuller SJ, Smith MJ, Beyreuther K, Bush AI, Masters CL (1999) Soluble pool of Aβ amyloid as a determinant of severity of neurodegeneration in Alzheimer's disease. Ann Neurol 46:860–866
44. Dickson DW, Crystal HA, Mattiace LA, Masur DM, Blau AD, Davies P, Yen S, Aronson MK (1992) Identification of normal and pathological aging in prospectively studied nondemented elderly humans. Neurobiol Aging 13:179–189
45. Caughey B, Lansbury PT (2003) Protofibrils, pores, fibrils, and neurodegeneration: separating the responsible protein aggregates from the innocent bystanders. Ann Rev Neurosci 26:267–298
46. Eisenberg D, Jucker M (2012) The amyloid state of proteins in human diseases. Cell 148:1188–1203
47. Stefani M (2012) Structural features and cytotoxicity of amyloid oligomers: implications in Alzheimer's disease and other diseases with amyloid deposits. Prog Neurobiol 99:226–245
48. Bucciantini M, Giannoni E, Chiti F, Baroni F, Formigli L, Zurdo JS, Taddei N, Ramponi G, Dobson CM, Stefani M (2002) Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. Nature 416:507–511
49. Quist A, Doudevski L, Lin H, Azimova R, Ng D, Frangione B, Kagan B, Ghiso J, Lal R (2005) Amyloid Ion channels: a common structural link for protein-misfolding disease. Proc Natl Acad Sci USA 102:10427–10432
50. Lashuel HA, Hartley D, Petre BM, Walz T, Lansbury PT (2002) Neurodegenerative disease —amyloid pores from pathogenic mutations. Nature 418:291
51. Porat Y, Kolusheva S, Jelinek R, Gazit E (2003) The human islet amyloid polypeptide forms transient membrane-active prefibrillar assemblies. Biochemistry 42:10971–10977
52. Kayed R, Sokolov Y, Edmonds B, McIntire TM, Milton SC, Hall JE, Glabe CG (2004) Permeabilization of lipid bilayers is a common conformation-dependent activity of soluble amyloid oligomers in protein misfolding diseases. J Biol Chem 279:46363–46366

53. Ambroggio EE, Kim DH, Separovic F, Barrow CJ, Barnham KJ, Bagatolli LA, Fidelio GD (2005) Surface behavior and lipid interaction of Alzheimer β-amyloid peptide 1–42: a membrane-disrupting peptide. Biophys J 88:2706–2713
54. Knight JD, Hebda JA, Miranker AD (2006) Conserved and cooperative assembly of membrane-bound α-helical states of islet amyloid polypeptide. Biochemistry 45:9496–9508
55. Soto C (2003) Unfolding the role of protein misfolding in neurodegenerative diseases. Nat Rev Neurosci 4:49–60
56. DeToma AS, Salamekh S, Ramamoorthy A, Lim MH (2012) Misfolded proteins in Alzheimer's disease and type II diabetes. Chem Soc Rev 41:608–621
57. Brender JR, Hartman K, Nanga RPR, Popovych N, de la Salud Bea R, Vivekanandan S, Marsh ENG, Ramamoorthy A (2010) Role of zinc in human islet amyloid polypeptide aggregation. J Am Chem Soc 132:8973–8983
58. Thies W, Bleiler L (2011) Alzheimer's association report 2011 Alzheimer's disease facts and figures. Alzheimer's Dement 7:208–244
59. Braak H, Braak E (1997) Frequency of stages of Alzheimer-related lesions in different age categories. Neurobiol Aging 18:351–357
60. Thal DR, Rub U, Orantes M, Braak H (2002) Phases of Aβ-deposition in the human brain and its relevance for the development of AD. Neurology 58:1791–1800
61. Soto C, Martin Z (2009) Therapeutic strategies against protein misfolding in neurodegenerative diseases. Expert Opin Drug Discov 4:71–84
62. Frisardi V, Solfrizzi V, Imbimbo BP, Capurso C, D'Introno A, Colacicco AM, Vendemiale G, Seripa D, Pilotto A, Capurso A, Panza F (2010) Towards disease-modifying treatment of Alzheimer's disease: drugs targeting β-amyloid. Curr Alzheimer Res 7:40–55
63. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM (1984) Clinical diagnosis of Alzheimer's disease. Neurology 34:939–944
64. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM (2011) Clinical diagnosis of Alzheimer's disease. Neurology 77:333
65. Dubois B, Feldman HH, Jacova C, Hampel H, Molinuevo JL, Blennow K, DeKosky ST, Gauthier S, Selkoe D, Bateman R, Cappa S, Crutch S, Engelborghs S, Frisoni GB, Fox NC, Galasko D, Habert M, Jicha GA, Nordberg A, Pasquier F, Rabinovici G, Robert P, Rowe C, Salloway S, Sarazin M, Epelbaum S, de Souza LC, Vellas B, Visser PJ, Schneider L, Stern Y, Scheltens P, Cummings JL (2014) Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. Lancet Neurol 13:614–629
66. Nordberg A (2010) Amyloid imaging in early detection of Alzheimer's disease. Neurodegener Dis 7:136–138
67. Caughey B, Ernst D, Race RE (1993) Congo red inhibition of scrapie agent replication. J Virol 67:6270–6272
68. Lorenzo A, Yankner BA (1994) β-amyloid neurotoxicity requires fibril formation and is inhibited by Congo red. Proc Natl Acad Sci USA 91:12243–12247
69. Kim YS, Randolph TW, Manning MC, Stevens FJ, Carpenter JF (2003) Congo red populates partially unfolded states of an amyloidogenic protein to enhance aggregation and amyloid fibril formation. J Biol Chem 278:10842–10850
70. Porat Y, Abramowitz A, Gazit E (2006) Inhibition of amyloid fibril formation by polyphenols: structural similarity and aromatic interactions as a common inhibition mechanism. Chem Biol Drug Des 67:27–37
71. Nielsen L, Khurana R, Coats A, Frokjaer S, Brange J, Vyas S, Uversky VN, Fink AL (2001) Effect of environmental factors on the kinetics of insulin fibril formation: elucidation of the molecular mechanism. Biochemistry 40:6036–6046
72. Mauro M, Craparo EF, Podesta A, Bulone D, Carrotta R, Martorana V, Tiana G, San Biagio PL (2007) Kinetics of different processes in human insulin amyloid formation. J Mol Biol 366:258–274
73. Fodera V, Librizzi F, Groenning M, van de Weert M, Leone M (2008) Secondary nucleation and accessible surface in insulin amyloid fibril formation. J Phys Chem B 112:3853–3858

74. Furumoto S, Okamura N, Iwata R, Yanai K, Arai H, Kudo Y (2007) Recent advances in the development of amyloid imaging agents. Curr Top Med Chem 7:1773–1789
75. Mathis CA, Wang YM, Holt DP, Huang GF, Debnath ML, Klunk WE (2003) Synthesis and evaluation of ^{11}C-labeled 6-substituted 2-arylbenzothiazoles as amyloid imaging agents. J Med Chem 46:2740–2754
76. Nordberg A, Carter SF, Rinne J, Drzezga A, Brooks DJ, Vandenberghe R, Perani D, Forsberg A, Langstrom B, Scheinin N, Karrasch M, Nagren K, Grimmer T, Miederer I, Edison P, Okello A, Van Laere K, Nelissen N, Vandenbulcke M, Garibotto V, Almkvist O, Kalbe E, Hinz R, Herholz KA (2013) European multicentre PET study of fibrillar amyloid in Alzheimer's disease. Eur J Nucl Med Mol Imaging 40:104–114
77. Agdeppa ED, Kepe V, Liu J, Flores-Torres S, Satyamurthy N, Petric A, Cole GM, Small GW, Huang SC, Barrio JR (2001) Binding characteristics of radiofluorinated 6-dialkylamino-2-naphthylethylidene derivatives as positron emission tomography imaging probes for β-amyloid plaques in Alzheimer's disease. J Neurosci 21:RC189
78. Small GW, Kepe V, Ercoli LM, Siddarth P, Bookheimer SY, Miller KJ, Lavretsky H, Burggren AC, Cole GM, Vinters HV, Thompson PM, Huang SC, Satyamurthy N, Phelps ME, Barrio JR (2006) PET of brain amyloid and tau in mild cognitive impairment. N Engl J Med 355:2652–2663
79. American Food and Drug Administration (2012) FDA approves ^{18}F-Florbetapir PET agent. J Nucl Med 53:15N
80. FDA PRESS RELEASE: FDA approves second brain imaging drug to help evaluate patients for Alzheimer's disease, dementia. http://www.fda.gov/newsevents/newsroom/pressannouncements/ucm372261.htm. Accessed 21 July 2014
81. Li Q, Min J, Ahn Y, Namm J, Kim EM, Lui R, Kim HY, Ji Y, Wu H, Wisniewski T, Chang Y (2007) Styryl-based compounds as potential in vivo imaging agents for β-amyloid plaques. ChemBioChem 8:1679–1687
82. Koole M, Lewis DM, Buckley C, Nelissen N, Vandenbulcke M, Brooks DJ, Vandenberghe R, Van Laere K (2009) Whole-body biodistribution and radiation dosimetry of ^{18}F-GE067: a radioligand for in vivo brain amyloid imaging. J Nucl Med 50:818–822
83. Joshi AD, Pontecorvo MJ, Clark CM, Carpenter AP, Jennings DL, Sadowsky CH, Adler LP, Kovnat KD, Seibyl JP, Arora A, Saha K, Burns JD, Lowrey MJ, Mintun MA., Skovronsky DM, and the Florbetapir F 18 Study Investigators (2012) Performance characteristics of amyloid PET with Florbetapir F 18 in patients with Alzheimer's disease and cognitively normal subjects. J Nucl Med 53:378–384
84. Cselényi Z, Jonhagen ME, Forsberg A, Halldin C, Julin P, Schou M, Johnstrom P, Varnas K, Svensson S, Farde L (2012) Clinical validation of ^{18}F-AZD4694, an amyloid-β-specific PET radioligand. J Nucl Med 53:415–424
85. Krebs MRH, Bromley EHC, Donald AM (2005) The binding of thioflavin-T to amyloid fibrils: localisation and implications. J Struct Biol 149:30–37
86. Robbins KJ, Liu G, Lin G, Lazo ND (2011) Detection of strongly bound thioflavin T species in amyloid fibrils by ligand-detected ^{1}H NMR. J Phys Chem Lett 2:735–740
87. Groenning M, Olsen L, van de Weert M, Flink JM, Frokjaer S, Jorgensen FS (2007) Study on the binding of thioflavin T to β-sheet-rich and non-β-sheet cavities. J Struct Biol 158: 358–369
88. Lockhart A, Ye L, Judd DB, Merritt AT, Lowe PN, Morgenstern JL, Hong GZ, Gee AD, Brown J (2005) Evidence for the presence of three distinct binding sites for the thioflavin T class of Alzheimer's disease PET imaging agents on β-amyloid peptide fibrils. J Biol Chem 280:7677–7684
89. Jin LW, Claborn KA, Kurimoto M, Geday MA, Maezawa I, Sohraby F, Estrada M, Kaminsky W, Kahr B (2003) Imaging linear birefringence and dichroism in cerebral amyloid pathologies. Proc Natl Acad Sci USA 100:15294–15298

90. Klunk WE, Pettegrew JW, Abraham DJ (1989) Quantitative evaluation of Congo red binding to amyloid-like proteins with a β-pleated sheet conformation. J Histochem Cytochem 37:1273–1281
91. Cai L, Innis RB, Pike VW (2007) Radioligand development for PET imaging of β-amyloid (Aβ)-current status. Curr Med Chem 14:19–52
92. Groenning M (2010) Binding mode of thioflavin T and other molecular probes in the context of amyloid fibrils—current status. J Chem Biol 3:1–18
93. Schütz AK, Soragni A, Hornemann S, Aguzzi A, Ernst M, Boeckmann A, Meier BH (2011) The amyloid-Congo red interface at atomic resolution. Angew ChemInt Ed 50:5956–5960
94. Zhuang ZP, Kung MP, Hou C, Skovronsky DM, Gur TL, Plossl K, Trojanowski JQ, Lee VMY, Kung HF (2001) Radioiodinated styrylbenzenes and thioflavins as probes for amyloid aggregates. J Med Chem 44:1905–1914
95. Agdeppa ED, Kepe V, Shoghi-Jadid K, Liu J, Small GW, Huang S, Petric A, Satyamurthy N, Barrio JR (2003) 2-Dialkylamino-6-acylmalononitrile substituted naphthalenes (DDNP Analogs): novel diagnostic and therapeutic tools in Alzheimer's disease. Mol Imaging Biol 5:404–417
96. Wu C, Wang Z, Lei H, Zhang W, Duan Y (2007) Dual binding modes of Congo red to amyloid protofibril surface observed in molecular dynamics simulations. J Am Chem Soc 129:1225–1232
97. Wu C, Wang Z, Lei H, Duan Y, Bowers MT, Shea J (2008) The binding of thioflavin T and its neutral analog BTA-1 to protofibrils of the Alzheimer's disease $A\beta_{16-22}$ peptide probed by molecular dynamics simulations. J Mol Biol 384:718–729
98. Wu C, Bowers MT, Shea J (2011) On the origin of the stronger binding of PIB over thioflavin T to protofibrils of the Alzheimer amyloid-β peptide: a molecular dynamics study. Biophys J 100:1316–1324
99. Wu C, Lei H, Wang Z, Zhang W, Duan Y (2006) Phenol red interacts with the protofibril-like oligomers of an amyloidogenic hexapeptide NFGAIL through both hydrophobic and aromatic contacts. Biophys J 91:3664–3672
100. Lemkul JA, Bevan DR (2010) Destabilizing Alzheimer's $A\beta_{42}$ protofibrils with morin: mechanistic insights from molecular dynamics simulations. Biochemistry 49:3935–3946
101. Berhanu WM, Masunov AE (2010) Natural polyphenols as inhibitors of amyloid aggregation. Molecular dynamics study of GNNQQNY heptapeptide decamer. Biophys Chem 149:12–21
102. Stumvoll M, Goldstein BJ, van Haeften TW (2007) Pathogenesis of type 2 diabetes. Endocr Res 32:19–37
103. Westermark P, Andersson A, Westermark GT (2011) Islet amyloid polypeptide, islet amyloid, and diabetes mellitus. Physiol Rev 91:795–826
104. Westermark P, Wernstedt C, Wilander E, Hayden DW, O'brien TD, Johnson KH (1987) Amyloid fibrils in human insulinoma and islets of Langerhans of the diabetic cat are derived from a neuropeptide-like protein also present in normal islet cells. Proc Natl Acad Sci USA 84:3881–3885
105. Cooper GJS, Willis AC, Clark A, Turner RC, Sim RB, Reid KBM (1987) Purification and characterization of a peptide from amyloid-rich pancreases of Type-2 diabetic-patients. Proc Natl Acad Sci USA 84:8628–8632
106. Lukinius A, Wilander E, Westermark GT, Engstrom U, Westermark P (1989) Co-localization of islet amyloid polypeptide and insulin in the B cell secretory granules of the human pancreatic-islets. Diabetologia 32:240–244
107. Kahn SE, Dalessio DA, Schwartz MW, Fujimoto WY, Ensinck JW, Taborsky GJ, Porte D (1990) Evidence of cosecretion of islet amyloid polypeptide and insulin by β-cells. Diabetes 39:634–638
108. Lutz TA (2012) Control of energy homeostasis by amylin. Cell Mol Life Sci 69:1947–1965
109. Muff R, Buhlmann N, Fischer JA, Born W (1999) An amylin receptor is revealed following co-transfection of a calcitonin receptor with receptor activity modifying proteins-1 or -3. Endocrinology 140:2924–2927

110. Sanke T, Bell GI, Sample C, Rubenstein AH, Steiner DF (1988) An islet amyloid peptide is derived from an 89-amino acid precursor by proteolytic processing. J Biol Chem 263: 17243–17246
111. Westermark P, Engström U, Johnson KH, Westermark GT, Betsholtz C (1990) Islet amyloid polypeptide: pinpointing amino acid residues linked to amyloid fibril formation. Proc Natl Acad Sci USA 87:5036–5040
112. Betsholtz C, Christmansson L, Engström U, Rorsman F, Svensson V, Johnson KH, Westermark P (1989) Sequence divergence in a specific region of islet amyloid polypeptide (IAPP) explains differences in islet amyloid formation between species. FEBS Lett 251: 261–264
113. Williamson JA, Loria JP, Miranker AD (2009) Helix stabilization precedes aqueous and bilayer-catalyzed fiber formation in islet amyloid polypeptide. J Mol Biol 393:383–396
114. Wiltzius JJW, Sievers SA, Sawaya MR, Eisenberg D (2009) Atomic structures of IAPP (amylin) fusions suggest a mechanism for fibrillation and the role of insulin in the process. Protein Sci 18:1521–1530
115. Abedini A, Raleigh DP (2009) A role for helical intermediates in amyloid formation by natively unfolded polypeptides? Phys Biol 6:015005
116. Jayasinghe S, Langen R (2005) Lipid membranes modulate the structure of islet amyloid polypeptide. Biochemistry 44:12113–12119
117. Nanga RPR, Brender JR, Xu J, Veglia G, Ramamoorthy A (2008) Structures of rat and human islet amyloid polypeptide $IAPP_{1-19}$ in micelles by NMR spectroscopy. Biochemistry 47:12689–12697
118. Nanga RPR, Brender JR, Vivekanandan S, Ramamoorthy A (2011) Structure and membrane orientation of IAPP in its natively amidated form at physiological pH in a membrane environment. Biochim Biophys Acta Biomembr 1808:2337–2342
119. Patil SM, Xu S, Sheftic SR, Alexandrescu AT (2009) Dynamic α-helix structure of micelle-bound human amylin. J Biol Chem 284:11982–11991
120. Apostolidou M, Jayasinghe SA, Langen R (2008) Structure of α-helical membrane-bound human islet amyloid polypeptide and its implications for membrane-mediated misfolding. J Biol Chem 283:17205–17210
121. Smith PES, Brender JR, Ramamoorthy A (2009) Induction of negative curvature as a mechanism of cell toxicity by amyloidogenic peptides: the case of islet amyloid polypeptide. J Am Chem Soc 131:4470–4478
122. Caillon L, Lequin O, Khemtémourian L (2013) Evaluation of membrane models and their composition for islet amyloid polypeptide-membrane aggregation. Biochim Biophys Acta Biomembr 1828:2091–2098
123. Janson J, Ashley RH, Harrison D, McIntyre S, Butler PC (1999) The mechanism of islet amyloid polypeptide toxicity is membrane disruption by intermediate-sized toxic amyloid particles. Diabetes 48:491–498
124. Anguiano M, Nowak RJ, Lansbury PT (2002) Protofibrillar islet amyloid polypeptide permeabilizes synthetic vesicles by a pore-like mechanism that may be relevant to type II diabetes. Biochemistry 41:11338–11343
125. Vaiana SM, Ghirlando R, Yau W, Eaton WA, Hofrichter J (2008) Sedimentation studies on human amylin fail to detect low-molecular-weight oligomers. Biophys J 94:L45–L47
126. Green JD, Goldsbury C, Kistler J, Cooper GJS, Aebi U (2004) Human amylin oligomer growth and fibril elongation define two distinct phases in amyloid formation. J Biol Chem 279:12206–12212
127. Zhao H, Sui Y, Guan J, He L, Gu X, Wong HK, Baum L, Lai FMM, Tong PCY, Chan JCN (2009) Amyloid oligomers in diabetic and nondiabetic human pancreas. Translational Research 153:24–32
128. Kayed R, Head E, Thompson JL, McIntire TM, Milton SC, Cotman CW, Glabe CG (2003) Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. Science 300:486–489

129. Yoshiike Y, Minai R, Matsuo Y, Chen Y, Kimura T, Takashima A (2008) Amyloid oligomer conformation in a group of natively folded proteins. PLoS ONE 3:e3235
130. Meier JJ, Kayed R, Lin C, Gurlo T, Haataja L, Jayasinghe S, Langen R, Glabe CG, Butler PC (2006) Inhibition of human IAPP fibril formation does not prevent β-cell death: evidence for distinct actions of oligomers and fibrils of human IAPP. Am J Physiol Endocrinol Metab 291:E1317–E1324
131. Meng F, Marek P, Potter KJ, Verchere CB, Raleigh DP (2008) Rifampicin does not prevent amyloid fibril formation by human islet amyloid polypeptide but does inhibit fibril thioflavin-T interactions: implications for mechanistic studies β-cell death. Biochemistry 47:6016–6024
132. Engel MFM, Khemtémourian L, Kleijer CC, Meeldijk HJD, Jacobs J, Verkleij AJ, de Kruijff B, Killian JA, Höppener JWM (2008) Membrane damage by human islet amyloid polypeptide through fibril growth at the membrane. Proc Natl Acad Sci USA 105:6033–6038
133. Sparr E, Engel M, Sakharov D, Sprong M, Jacobs J, de Kruijff B, Hoppener J, Killian J (2004) Islet amyloid polypeptide-induced membrane leakage involves uptake of lipids by forming amyloid fibers. FEBS Lett 577:117–120
134. Gellermann GP, Appel TR, Tannert A, Radestock A, Hortschansky P, Schroeckh V, Leisner C, Lütkepohl T, Shtrasburg S, Röcken C, Pras M, Linke RP, Diekmann S, Fändrich M (2005) Raft lipids as common components of human extracellular amyloid fibrils. Proc Natl Acad Sci USA 102:6297–6302
135. Friedman R, Pellarin R, Caflisch A (2009) Amyloid aggregation on lipid bilayers and its impact on membrane permeability. J Mol Biol 387:407–415
136. Abedini A, Schmidt AM (2013) Mechanisms of islet amyloidosis toxicity in type 2 diabetes. FEBS Lett 587:1119–1127
137. Shim S, Gupta R, Ling YL, Strasfeld DB, Raleigh DP, Zanni MT (2009) Two-dimensional IR spectroscopy and isotope labeling defines the pathway of amyloid formation with residue-specific resolution. Proc Natl Acad Sci USA 106:6614–6619
138. Knight JD, Miranker AD (2004) Phospholipid catalysis of diabetic amyloid assembly. J Mol Biol 341:1175–1187
139. Rustenbeck I, Matthies A, Lenzen S (1994) Lipid-composition of glucose-stimulated pancreatic-islets and insulin-secreting tumor-cells. Lipids 29:685–692
140. Farese RV, Dimarco PE, Barnes DE, Sabir MA, Larson RE, Davis JS, Morrison AD (1986) Rapid glucose-dependent increases in phosphatidic-acid and phosphoinositides in rat pancreatic-islets. Endocrinology 118:1498–1503
141. Chargé SBP, Dekoning EJP, Clark A (1995) Effect of pH and insulin on fibrillogenesis of islet amyloid polypeptide in vitro. Biochemistry 34:14588–14593
142. Khemtémourian L, Doménech E, Doux JPF, Koorengevel MC, Killian JA (2011) Low pH acts as inhibitor of membrane damage induced by human islet amyloid polypeptide. J Am Chem Soc 133:15598–15604
143. Hutton JC (1989) The insulin secretory granule. Diabetologia 32:271–281
144. Westermark P, Li Z, Westermark GT, Leckström A, Steiner DF (1996) Effects of beta cell granule components on human islet amyloid polypeptide fibril formation. FEBS Lett 379:203–206
145. Knight JD, Williamson JA, Miranker AD (2008) Interaction of membrane-bound islet amyloid polypeptide with soluble and crystalline insulin. Protein Sci 17:1850–1856
146. Engel MFM, Yigittop H, Elgersma RC, Rijkers DTS, Liskamp RMJ, de Kruijff B, Höppener JWM, Killian JA (2006) Islet amyloid polypeptide inserts into phospholipid monolayers as monomer. J Mol Biol 356:783–789
147. Sasahara K, Hall D, Hamada D (2010) Effect of lipid type on the binding of lipid vesicles to islet amyloid polypeptide amyloid fibrils. Biochemistry 49:3040–3048
148. Zhao J, Luo Y, Jang H, Yu X, Wei G, Nussinov R, Zheng J (2012) Probing ion channel activity of human islet amyloid polypeptide (amylin). biochim. Biophys Acta Biomembr 1818:3121–3130

149. Duan M, Fan J, Huo S (2012) Conformations of islet amyloid polypeptide monomers in a membrane environment: implications for fibril formation. PLoS ONE 7:e47150
150. Zhang Y, Luo Y, Deng Y, Mu Y, Wei G (2012) lipid interaction and membrane perturbation of human islet amyloid polypeptide monomer and dimer by molecular dynamics simulations. PLoS ONE 7:e38191
151. Sakagashira S, Sanke T, Hanabusa T, Shimomura H, Ohagi S, Kumagaye KY, Nakajima K, Nanjo K (1996) Missense mutation of amylin gene (S20G) in Japanese NIDDM patients. Diabetes 45:1279–1281
152. Sakagashira S, Hiddinga HJ, Tateishi K, Sanke T, Hanabusa T, Nanjo K, Eberhardt NL (2000) S20G mutant amylin exhibits increased in vitro amyloidogenicity and increased intracellular cytotoxicity compared to wild-type amylin. Am J Pathol 157:2101–2109
153. Jia Y, Qian Z, Zhang Y, Wei G (2013) Adsorption and orientation of human islet amyloid polypeptide (hIAPP) monomer at anionic lipid bilayers: implications for membrane-mediated aggregation. Int J Mol Sci 14:6241–6258
154. Marrink S-J, Risselada HJ, Yefimov S, Tieleman DP, de Vries AH (2007) The MARTINI force field: coarse grained model for biomolecular simulations. J Phys Chem B 111: 7812–7824
155. Monticelli L, Kandasamy SK, Periole X, Larson RG, Tieleman DP, Marrink S-J (2008) The MARTINI coarse-grained force field: extension to proteins. J Chem Theory Comput 4:819–834
156. Xu W, Wei G, Su H, Nordenskiold L, Mu Y (2011) Effects of cholesterol on pore formation in lipid bilayers induced by human islet amyloid polypeptide fragments: a coarse-grained molecular dynamics study. Phys Rev E Stat Nonlinear Soft Matter Phys 84:051922
157. Periole X, Cavalli M, Marrink S-J, Ceruso MA (2009) Combining an elastic network with a coarse-grained molecular force field: structure, dynamics, and intermolecular recognition. J Chem Theory Comput 5:2531–2543

Computational Theory

Protein Structure and Function

The relationship between protein structure and function is central for understanding how the molecular world works. However, determining both can be very difficult and requires a combination of many different methods, as most methods are focused on determining either one or the other. X-ray crystallography, EM, and to some extent NMR spectroscopy are typical methods used to determine the structure of proteins [1–3]. Function is often coupled to the dynamical properties of the protein, which can be examined using e.g. Förster resonance energy transfer (FRET), variants of NMR spectroscopy, and infrared spectroscopy [4]. Each method has a timescale of which it can provide information ranging from vibrations of individual atoms on a sub-femtosecond timescale, to protein translation by the ribosome on a timescale of seconds [5]. Most of the available techniques provide ensemble properties, and do not report on the dynamics of single proteins.

MD simulation is a computational technique able to provide very high resolution in time and space which is unmatched by any experimental technique. It basically provides a three-dimensional movie of how the atoms in a system move relative to one another. MD is an important complementary technique to experiments, as it can help explain the mechanisms of molecular processes. In principle, simulation of any molecular system imaginable is possible, even if it would not be possible to set it up in practice.

The aim of MD simulations is to determine the motions of atoms in a system. The time-dependent Schrödinger equation describes the behavior of molecules; however, solving this is practically infeasible for large molecular assemblies. The motions of atoms and molecules heavier than the proton are essentially classical. Therefore, classical mechanics is used in MD to propagate the system based on a model, termed a FF, describing the energetics of the system [6].

K.K. Skeby, *Computational Modelling of the Human Islet Amyloid Polypeptide*,
Springer Theses, DOI 10.1007/978-3-319-20040-8_2

Force Fields

The relationship between atomic coordinates and energy is defined in a FF. A FF is composed of a set of functions describing the energy of the interatomic interactions, e.g. bond stretching or angle bending, as well as an associated set of parameters determining the strength of these interactions. The parameters can be derived in various ways, e.g. by fitting to experimental data or quantum mechanical (QM) calculations. FFs rely on transferability, which means that parameters such as bond angles and force constants are assumed to be similar for the same chemical groups in different molecules. This allows fitting parameters to one set of molecules, and applying the parameters to other molecules of similar chemical structure. Transferability also leads to the concept of atom types. An atom type holds information about the element, hybridization, and neighboring atoms of a particle, and is used to define which parameters are used for a specific atom.

Most FFs use the same basic functional to describe the energy, which has not changed considerably for the commonly used FFs since they were first released [7]. The FF energy (E_{FF}) is described as a sum of bonded and non-bonded terms. The bonded contributions are from the stretch (or bond) energy (E_{str}), the bend (or angle) energy (E_{bend}), and the torsional energy (E_{tors}). The non-bonded contributions come from the van der Waals (E_{vdW}) and electrostatic energies (E_{elec}), and are usually included for atoms separated by more than two or three bonds. The vdW energy is usually described by the Lennard-Jones (LJ) potential, and the electrostatic energy is described by the Coulomb potential. The stretch and bend energies are both described by a harmonic function, whereas the torsional energy is described by a sum of cosine functions to account for periodicity.

$$E_{str}(R) = k_{str}(R - R_0)^2 \quad E_{bend}(\theta) = k_{bend}(\theta - \theta_0)^2$$

$$E_{tors}(\omega) = \sum_{n=1} \frac{V_n}{2}[1 + \cos(n\omega - \gamma)]$$

$$E_{vdW}(R) = \varepsilon_{LJ}\left[\left(\frac{R^0_{LJ}}{R}\right)^{12} - 2\left(\frac{R^0_{LJ}}{R}\right)^{6}\right] \quad E_{elec}(R) = \frac{q^A q^B}{\varepsilon R}$$

More terms can be added to the FF for special cases, such as cross terms to account for the coupling between different types of interactions.

The most commonly used FFs were originally fitted to QM calculations as well as experimental data from X-ray structures [7]. Recent advances in both experimental and computational ability have made it both feasible and necessary to reparameterize the FFs to include dynamical experimental data, e.g. from NMR experiments, in the fitting procedure [7]. Comparative studies have shown that the older FFs tend to overestimate the stability of α-helical structures [8–10]. The most recent FFs such as AMBER99SB*−ILDN [11], CHARMM22* [11], and CHARMM36 [12] have refitted torsional parameters providing a better balance between helix, sheet, and coil. A recent comparative study of multiple protein FFs

(not including CHARMM36) showed that the latest FFs are able to reproduce several structural and dynamical properties of small proteins; however, the temperature dependence of the FFs is still lacking, and the folding mechanisms varies from FF to FF [10]. It is unclear whether FF improvements are possible using the current functional, or if it will be necessary to include more sophisticated effects, such as polarizability, to improve the accuracy of the FFs [7].

Molecular Dynamics Simulation

In conventional MD simulations, each atom in the system is described by one particle. The energy of the system is described by the FF. It is possible to determine the force acting on each atom by differentiating the energy with respect to the atomic positions, and then using Newton's second law of motion to obtain the acceleration on the atoms. This allows propagation of the system in time using classical mechanics. Numerical integration is used to determine the trajectory of the system, requiring that the time-step of each integration step is very small. The fastest motions in a system are the bond vibrations, which have frequencies on the order of 10^{14} s^{-1}. It then follows that the time-step needs to be on the order of 1 fs to avoid atoms moving too close, which would produce very large forces that would eventually make the system explode and the simulation crash. This also means that one billion steps are needed to simulate a single μs.

The starting position of all the atoms in the system must be defined before the MD simulation can be initiated. For a protein, the coordinates are usually obtained from an X-ray crystal or NMR structure available from the RCSB Protein Data Bank (PDB) [13]. Starting velocities are usually assigned to each atom from a Maxwell-Boltzmann distribution at the simulation temperature (T).

Performing MD simulations on a system of finite size, e.g. in a cubic box, introduces boundaries to the system. The behavior of molecules close to the boundary is different from the bulk behavior. Periodic boundary conditions can help minimize boundary effects, and can be visualized as replicating the system an infinite number of times in all directions. Imposing the minimum image convention (i.e. no cutoff can be longer than half the length of the shortest box vector) ensures that a particle only interacts with the closest image of another particle, and never with itself. Periodic boundary conditions are also helpful in calculating the electrostatic interactions, as is described below.

MD simulations do not scale well computationally, due to the intrinsic sequential nature of the simulations, and the fact that all atoms in a simulation interact with one another through the non-bonded interactions. In principle, the computational time scales with N^2, with N being the number of particles in the system; however, algorithmic advances in performing Fourier transformations have reduced the scaling with system size to N ln(N), as the long-range electrostatic interactions can be efficiently evaluated using the particle mesh Ewald method (PME) (see below) [6, 14].

The most computationally demanding part of an MD simulation is evaluation of the forces, and particularly the non-bonded forces. To minimize the number of evaluations, the vdW interactions can be truncated at a certain cutoff. This is a reasonable approximation since the vdW energy depends on R^{-6}, and therefore rapidly becomes negligible as the distance between atoms is increased. Truncation of the LJ potential can introduce discontinuities in the energy; therefore, a switching function can be introduced, which reduces the energy smoothly to zero between the switching distance and the cutoff distance. In itself, truncation does not save much time, since the distance between all pairs of atoms still has to be calculated. However, the introduction of a pair-list, which contains the pairs of atoms within a pair-list radius, can reduce the computation of all interatomic distances to every 10–20 steps. During each step, the distances and non-bonded forces are only calculated between the pairs of atoms included in the pair-list, which are within the cutoff distance.

The electrostatic potential depends on the distance as R^{-1}, which means that the distance where the energy becomes close to zero is large compared to the typical system size. Therefore, rather than using a cutoff, other means of reducing the computational time for evaluating the electrostatic forces have been developed, one of which is the use of Ewald sums [15]. Ewald summation divides the electrostatic contribution into two parts, the "near"-field and the "far"-field contributions [6]. The "near"-field contribution is the sum of the atomic partial charges and a screening charge distribution of Gaussian potentials. This screening potential exactly counters the atomic partial charges and is centered on the atomic positions. The sum converges quickly because of the screening field, and can therefore be truncated. The "far"-field contribution derives from a compensating Gaussian charge distribution which is exactly opposite the screening charge distribution. The compensating charge distribution converges rapidly in reciprocal space, and can be evaluated efficiently with Fourier transformation methods. The electrostatic potential at a point is the sum of the partial charges, the screening charge distribution, and the compensating charge distribution. The success of this method depends on efficient ways of evaluating the Fourier transformations; PME [14] uses one such method called Fast Fourier Transformation, which evaluates the transformation in a discrete number of grid points imposed on the system [6].

MD intrinsically samples the microcanonical (NVE) ensemble. Several methods have been developed to maintain the T and/or pressure (P) (near) constant, allowing sampling of other ensembles of interest. The isothermal-isobaric (NPT) ensemble is the most relevant for biomolecular simulations, since this is the condition that most experiments are performed under. Several algorithms exist for sampling constant T, and include scaling of the velocities, including coupling to an external heat bath (Berendsen thermostat); random reassignment of velocities from a Maxwell-Boltzmann distribution (Anderson thermostat); addition of frictional and random forces (Langevin dynamics); and extended system methods (Nosé-Hoover thermostat). It should be noted that scaling of the velocities does not strictly sample the

canonical (NVT) ensemble, as the fluctuations of T are not correct. The same algorithms can be used for constant P simulations by changing the positions of particles rather than the velocities.

Recent Advances in MD

Since all atoms in the system are interacting, the calculation cannot be accelerated by distributing the system to an infinite number of processors; at some point the communication between the different compute nodes will be slower than the actual calculation [7]. However, recent advances in hardware, other than an increase in the number of CPUs, have significantly accelerated the computation of MD trajectories. Graphics processing units (GPUs) originally developed to accelerate rendering of graphics for the gaming industry, have proven particularly useful in MD simulations [16]. The most popular MD engines, NAMD, GROMACS, and AMBER, have already implemented GPUs to calculate the time-consuming non-bonded forces, providing considerable speed-up of the calculations [17–20]. However, this approach is only viable up to a limited system size, as the communication between the GPUs is even worse than between the traditional processors. A more drastic approach has been taken at D.E. Shaw Research, where a special purpose computer chip named Anton has been developed specifically for performing MD simulations [21]. This has allowed simulations to extend into the ms timescale [22].

The current MD simulations being performed have come a long way from the first MD simulation of a folded protein, bovine pancreatic trypsin inhibitor, in 1977 by McCammon et al., which had a simulation length of 8.8 ps [23]. However, the amount of data being produced poses a great challenge in terms of handling, storage, and analysis. Currently, the largest entry in the PDB is an assembly of the HIV-1 capsid with more than 2 million atoms. It was modeled using all-atom MD Flexible Fitting to cryo-EM data [24]. The fully solvated system consisted of 64 million atoms, and was simulated for 100 ns, which is the largest all-atom MD simulation performed to date [25]. Development of new software was needed to handle both the analysis and the visualization of the large scale system, which illustrates that it is not just the desire to go bigger, but actual biological problems that drive the development of new methodologies [26, 27].

Enhanced Sampling Methods

Not everyone has access to a brute force computer such as Anton. It is therefore necessary to devise new methods to simulate rare events within the timescale that most scientists can obtain. Enhanced sampling methods have been developed to encourage the system to explore regions of phase space which are otherwise difficult to investigate. Some methods require knowledge of both the starting and

end-point of the process of interest. For example targeted MD explores the conformational change of a protein going from one state to another by applying a force to minimize the root mean square deviation (RMSD) of the initial structure compared to the final structure [28]. Such methods are not helpful when the end-point structure is not known, and may also move the system along a path, which is not a physically relevant transition path. Steered MD is an enhanced sampling method that does not require knowledge of the end-point structure [29]. A force is applied to one or more atoms while keeping other atoms fixed. This allows e.g. for simulation of the unbinding of ligands or the unfolding of proteins. While the final protein structure is not known before the simulation, a preconceived idea about the direction of change is imposed on the system by the user. This might impose a force on the system making it explore high energy regions of phase space along the path, which are not physically relevant. This is partly avoided in metadynamics, where a history-dependent bias potential is added to the system to discourage it from revisiting already explored areas of phase space [30]. When a potential energy well is filled by the history-dependent potential, the system is free to choose any path to move to another low-energy region of the phase space, which means that the system is not pushed or pulled towards high energy regions. Collective variables (CVs) are used to guide the system, and it is therefore important to choose CVs that cover the important regions of phase space. The advantage of metadynamics is that the phase space is efficiently explored simultaneously with the evaluation of the free energy surface. However, it is also more complicated to execute as the choice of CVs is essential to the success of the simulations, and again, prior knowledge about the system is often necessary. Accelerated MD (aMD) does not require prior knowledge about the system to enhance the sampling of the system. In aMD the potential of the low-energy regions of phase space is raised, which effectively lowers the barriers of transition between minima on the potential energy surface [31]. The amount that the potential energy surface should be raised needs to be optimized. A too aggressive approach leaves the risk of obtaining a flat potential energy surface, which would cause the system to spend a lot of time in regions of phase space which are not of interest.

CG MD simulation is a separate category of enhanced sampling methods in which the number of particles in the system is decreased by combining multiple atoms into a single particle. The main speed-up of these methods does not actually arise from the reduced number of particles, but from the possibility of longer time-steps. The MARTINI CG FF is a popular method for simulating biomolecular systems [32]. It was originally developed for simulating lipid systems, but has been further developed to include models for proteins and carbohydrates [33, 34]. The MARTINI FF combines approximately four heavy atoms into one CG bead, which allows the use of a time-step between 20 and 40 ns [35]. Conformational changes of proteins have proven to be challenging to simulate with MARTINI since hydrogen bonds are not included explicitly in the model, and the secondary structure has to be defined in the bead-type [35]. Gō models use the native interactions of the protein combined with a single particle for each residue to investigate the folding mechanisms of proteins [36]. However, this can thus only be used to

investigate folding into a conformation which is already known. The kinetics of amyloid aggregation has been studied using a phenomenological CG model for which the aggregation propensity can be tuned [37]. A short review and discussion regarding the use of CG models to investigate amyloid aggregation is given in the chapter "Coarse Grained Study of Amyloid Protofibril Aggregation".

Protein-Ligand Binding

A large part of the study of proteins deals with the binding of ligands, e.g. substrates, signaling molecules, or inhibitors, to the protein. Therefore, a large part of the biomolecular modeling research is also devoted to the study of protein-ligand complexes. Software devoted to determining the structure of protein-ligand complexes is termed docking software, and usually contains a scoring function to estimate the strength of the complex [38].

Docking can be used for a variety of applications, from screening of a virtual library containing many thousands or millions of compounds to determining the structure and energy of binding of a single molecule. Most docking software includes flexibility of the ligand, either by generating multiple conformations of the ligand and then docking each as a rigid molecule, or by introducing flexibility into the ligand during the docking procedure. However, flexibility of the protein requires much more computational power, and is not a routine feature. It can be introduced e.g. by using several protein structures obtained from X-ray, NMR, or MD, or it can be a feature of the docking software. Another type of protein flexibility can be introduced using a rotamer library for the side chains in the binding pocket. It is especially important to incorporate protein flexibility when the target ligand does not resemble the ligand present in the binding pocket of the protein structure, or if the protein structure is a model generated based on a structure of a homologous protein [38].

Sometimes the binding site of a protein target is not known, or the goal is to develop a novel drug, which does not have to bind in the substrate binding pocket. Then it becomes necessary to use binding site prediction software, such as Sitemap [39]. This can be used to determine the location of potential binding pockets, and score them based on the likelihood of finding a good drug that will bind in the pocket. Furthermore, Sitemap can also give an idea of which functional groups should be placed in specific areas of the binding pocket to create a high-affinity drug.

However, binding site prediction and docking software fails when the binding site on the protein is not well-defined, such as in amyloid fibrils. First of all, the structural models of amyloid fibrils are usually based on a combination of NMR constraints and other biophysical methods, which do not necessarily give much information on e.g. the side chains of the protein [40]. Secondly, there are no well-defined binding pockets on the surface of the fibril, only shallow crevices which are highly exposed to the solvent. It can, therefore, be necessary to use MD

simulations to investigate possible binding sites of the ligand. A combination of multiple trajectories and free energy calculations of the resulting protein-ligand complexes can be highly effective in these cases [41–44].

Free Energy Calculations

Free energy calculations are of great importance in biomolecular modeling as it allows direct comparison with experiments. The Gibbs free energy is the relevant quantity when comparing simulations to experiments performed with a constant number of particles (N), pressure (P), and temperature (T), i.e. in the NPT ensemble. Unfortunately, the Gibbs free energy is difficult to calculate from MD simulations, as the higher energy regions of phase-space, which make important contributions to the free energy, are not sampled very much [45]. Often, the free energy difference of a reaction, such as the binding energy of a ligand to a protein or the relative energy of two conformational states, is the important property. The energy difference is a thermodynamic property, and therefore only depends on the difference between the two states [45]. This can be computed using rigorous and accurate methods such as free energy perturbation and thermodynamic integration. These methods sample a path between the two states, which make them computationally demanding, and consequently, they can only be performed for small perturbations [45, 46]. End-point models are alternatives to the more demanding methods, which only sample the two states of the reaction, and therefore can be used for a more diverse set of reactions [46].

MM-PBSA

Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) [47–49] is an approximate method to determine free energies. Binding free energies ($\Delta G_{binding}$) for the ligand-protein association (L + P $\rightarrow$ LP) can be evaluated using a thermodynamic cycle [49].

$$\begin{array}{ccc} L_{solv} & + & P_{solv} \xrightarrow{\Delta G_{binding}} LP_{solv} \\ \downarrow -\Delta G^{L}_{solv} & & \downarrow -\Delta G^{P}_{solv} \qquad \uparrow \Delta G^{LP}_{solv} \\ L_{gas} & + & P_{gas} \xrightarrow{\Delta G_{gas}} LP_{gas} \end{array}$$

$$\Delta G_{binding} = -\Delta G^{L}_{solv} - \Delta G^{P}_{solv} + \Delta G_{gas} + \Delta G^{LP}_{solv}$$

ΔG_{gas} is the gas-phase interaction energy and ΔG_{solv} is the free energy of moving a molecule or complex from the gas-phase to the solvated state. The approximation in MM-PBSA lies in determining the energetic components. The gas-phase interaction energy is divided into enthalpic and entropic contributions.

$$\Delta G_{gas} = \Delta H_{gas} - T\Delta S \approx \Delta E_{MM} - T\Delta S_{MM}$$

The enthalpy can be approximated as the Molecular Mechanics (MM) energy (E_{MM}), and the entropy can be approximated using quasi harmonic analysis or normal-mode analysis [47, 49]. However, methods for calculating the entropy are computationally expensive and require extensive sampling to reach convergence. E_{MM} is extracted directly from the FF, and is therefore a sum of the bonded (stretch, bend, and torsion) and the non-bonded (vdW and electrostatic) contributions. The solvation energies are approximated using the PBSA method, which divides the solvation energy into an electrostatic (E_{PB}) and a non-polar contribution (E_{cav}) [47]. The electrostatic interaction between the solute and solvent is calculated using a numerical solution to the Poisson-Boltzmann (PB) equation [47], which is a continuum solvation approach using a combination of the Poisson equation and a Boltzmann distribution of ions [45]. The Poisson equation is a second-order differential equation that relates the change in electrostatic potential to the dielectric constant and the charge density. The Boltzmann equation accounts for the distribution of mobile ionic charges in the solvent [45]. The non-polar contribution is approximately proportional to the surface area (SA) of the solute and can be viewed as arising from two terms: The vdW interactions between the solute and solvent, and the cost of forming the solute cavity in the solvent [45].

The MM-PBSA method requires the determination of the MM and solvation energies of all species of the reaction (L, P, and LP), and the resulting binding free energy is the difference between the energies of the reactants and products.

$$\Delta G_{binding} = G_{LP} - G_P - G_L$$

Ensembles of structures for calculating the energy can be generated by explicit-solvent MD simulations, with subsequent removal of the solvent for the evaluation of the MM-PBSA energy [47]. The ideal situation uses different trajectories to calculate ensemble averages for the complex, protein, and ligand, respectively. However, using a single trajectory to evaluate all contributions can minimize the noise from sampling inconsistencies and the error in the FF and implicit solvation energies [46]. This procedure assumes that the change in structure and conformational freedom upon association of the complex is negligible [46]. It should be noted that using the single-trajectory approach eliminates the bonded contributions to the MM energy. If only relative binding energies of similar ligands are desired, it is common to neglect the entropic contribution, as it is assumed that the change in entropy is comparable for all ligands [49, 50].

References

1. Branden C, Tooze J (1998) In: Introduction to Protein Structure, 2nd edn. Garland
2. Rule GS, Hitchens TK (2006) In: Fundamentals of Protein NMR Spectroscopy: Chapter 17: Protein Structure Determination, Springer
3. Zhou ZH (2008) Towards atomic resolution structural determination by single-particle cryo-electron microscopy. Curr Opin Struct Biol 18:218–228
4. Cooper A (2004) In: Biophysical Chemistry, Royal Society of Chemistry
5. Dror RO, Jensen MØ, Borhani DW, Shaw DE (2010) Exploring atomic resolution physiology on a femtosecond to millisecond timescale using molecular dynamics simulations. J Gen Physiol 135:555–562
6. Jensen F (2007) In: Introduction to Computational Chemistry, 2nd edn, John Wiley & Sons Ltd
7. Dror RO, Dirks RM, Grossman JP, Xu H, Shaw DE (2012) Biomolecular simulation: a computational microscope for molecular biology. Annu Rev Biophys 41:429–452
8. Best RB, Buchete N, Hummer G (2008) Are current molecular dynamics force fields too helical? Biophys J 95:L07–L09
9. Beauchamp KA, Lin Y, Das R, Pande VS (2012) Are protein force fields getting better? A systematic benchmark on 524 diverse NMR measurements. J Chem Theory Comput 8: 1409–1414
10. Lindorff-Larsen K, Maragakis P, Piana S, Eastwood MP, Dror RO, Shaw DE (2012) Systematic Validation of Protein Force Fields Against Experimental Data. PLoS One 7:e32131
11. Piana S, Lindorff-Larsen K, Shaw DE (2011) How robust are protein folding simulations with respect to force field parameterization? Biophys J 100:L47–L49
12. Best RB, Zhu X, Shim J, Lopes PEM, Mittal J, Feig M, MacKerell AD Jr (2012) Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain χ_1 and χ_2 dihedral angles, J Chem Theor Comput 8:3257–3273
13. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank, Nucl Acids Res 28:235–242
14. Darden T, York D, Pedersen L (1993) Particle Mesh Ewald: an N·log(N) method for Ewald sums in large systems. J Chem Phys 98:10089–10092
15. Ewald PP (1921) Die Berechnung Optischer Und Elektrostatischer Gitterpotentiale. Ann Phys 64:253–287
16. Stone JE, Hardy DJ, Ufimtsev IS, Schulten K (2010) GPU-accelerated molecular modeling coming of age. J Mol Graph Model 29:116–125
17. Stone JE, Phillips JC, Freddolino PL, Hardy DJ, Trabuco LG, Schulten K (2007) Accelerating molecular modeling applications with graphics processors. J Comput Chem 28:2618–2640
18. Goetz AW, Williamson MJ, Xu D, Poole D, Le Grand S, Walker RC (2012) Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. generalized born. J Chem Theory Comput 8:1542–1555
19. Salomon-Ferrer R, Goetz AW, Poole D, Le Grand S, Walker RC (2013) Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. generalized born. J Chem Theory Comput 9:3878–3888
20. Abraham M, Apol E, Apostolov R, Berendsen HJC, Buuren Av, Bjelkmar P, Drunen Rv, Feenstra A, Fritsch S, Groenhof G, Junghans C, Hub J, Kasson P, Kutzner C, Lambeth B, Larsson P, Lemkul JA, Marklund E, Meulenhoff P, Murtola T, Páll S, Pronk S, Schulz R, Shirts M, Sijbers A, Tieleman P, Wolf M, Hess B, van der Spoel D, Lindahl E (2010) Gromacs User Manual Version 4.6.5
21. Shaw DE, Deneroff MM, Dror RO, Kuskin JS, Larson RH, Salmon JK, Young C, Batson B, Bowers KJ, Chao JC, Eastwood MP, Gagliardo J, Grossman JP, Ho CR, Ierardi DJ, Kolossváry I, Klepeis JL, Layman T, McLeavey C, Moraes MA, Mueller R, Priest EC, Shan Y, Spengler J, Theobald M, Towles B, Wang SC (2008) Anton, a special-purpose machine for molecular dynamics simulation. ACM 51:91–97

22. Shaw DE, Dror RO, Salmon JK, Grossman JP, Mackenzie KM, Bank JA, Young C, Deneroff MM, Batson B, Bowers KJ, Chow E, Eastwood MP, Ierardi DJ, Klepeis JL, Kuskin JS, Larson RH, Lindorff-Larsen K, Maragakis P, Moraes MA, Piana S, Shan Y, Towles B (2009) Millisecond-scale molecular dynamics simulations on anton, proceedings of the conference on high performance computing networking, Storage and Analysis, 1–11
23. McCammon JA, Gelin BR, Karplus M (1977) Dynamics of folded proteins. Nature 267: 585–590
24. Trabuco LG, Villa E, Mitra K, Frank J, Schulten K (2008) Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. Structure 16:673–683
25. Zhao G, Perilla JR, Yufenyuy EL, Meng X, Chen B, Ning J, Ahn J, Gronenborn AM, Schulten K, Aiken C, Zhang P (2013) Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. Nature 497:643–646
26. Stone JE, Vandivort KL, Schulten K (2013) GPU-accelerated molecular visualization on petascale supercomputing platforms, UltraVis'13 Proceedings of the 8th international workshop on ultrascale visualization, p. 6
27. Stone JE, McGreevy R, Isralewitz B, Schulten K (2014) GPU-Accelerated Analysis and Visualization of Large Structures Solved by Molecular Dynamics Flexible Fitting, Faraday Discuss 169(12)
28. Schlitter J, Engels M, Krüger P (1994) Targeted molecular dynamics: a new approach for searching pathways of conformational transitions. J Mol Graph 12:84–89
29. Park S, Schulten K (2004) Calculating potentials of mean force from steered molecular dynamics simulations. J Chem Phys 120:5946–5961
30. Laio A, Parrinello M (2002) Escaping free-energy minima. Proc Natl Acad Sci USA 99:12562–12566
31. Markwick PRL, McCammon JA (2011) Studying functional dynamics in bio-molecules using accelerated molecular dynamics, Phys Chem Chem Phys 13:20053–20065
32. Marrink S-J, Risselada HJ, Yefimov S, Tieleman DP, de Vries AH (2007) The MARTINI force field: coarse grained model for biomolecular simulations. J Phys Chem B 111: 7812–7824
33. Monticelli L, Kandasamy SK, Periole X, Larson RG, Tieleman DP, Marrink S-J (2008) The MARTINI coarse-grained force field: extension to proteins. J Chem Theor Comput 4:819–834
34. López CA, Rzepiela AJ, de Vries AH, Dijkhuizen L, Hünenberger PH, Marrink S-J (2009) Martini coarse-grained force field: extension to carbohydrates. J Chem Theor Comput 5:3195–3210
35. Periole X, Marrink S-J (2013) The Martini coarse-grained force field. Methods Mol Biol 924:533–565
36. Hills RD Jr, Brooks CL III (2009) Insights from coarse-grained gō models for protein folding and dynamics. Int J Mol Sci 10:889–905
37. Pellarin R, Caflisch A (2006) Interpreting the aggregation kinetics of amyloid peptides. J Mol Biol 360:882–892
38. Höltje HD, Sippl W, Rognan D (2008) In: Molecular Modeling: Basic Principles and Applications: Virtual Screening and Docking, 3rd edn, Wiley-VCH
39. Halgren TA (2009) Identifying and characterizing binding sites and assessing druggability. J Chem Inf Model 49:377–389
40. Tycko R Solid (2011) Solid state NMR studies of amyloid fibril structure. Annu Rev Phys Chem 62:279–299
41. Wu C, Wang Z, Lei H, Zhang W, Duan Y (2007) Dual binding modes of congo red to amyloid protofibril surface observed in molecular dynamics simulations. J Am Chem Soc 129: 1225–1232
42. Wu C, Wang Z, Lei H, Duan Y, Bowers MT, Shea J (2008) The binding of thioflavin T and its neutral analog BTA-1 to protofibrils of the Alzheimer's disease $A\beta_{16\text{-}22}$ peptide probed by molecular dynamics simulations, J Mol Biol 384:718–729

43. Wu C, Bowers MT, Shea J (2011) On the origin of the stronger binding of PIB over thioflavin T to protofibrils of the Alzheimer amyloid-β peptide: a molecular dynamics study, Biophys J 100:1316–1324
44. Skeby KK, Sorensen J, Schiott B (2013) Identification of a common binding mode for imaging agents to amyloid fibrils from molecular dynamics simulations. J Am Chem Soc 135: 15114–15128
45. Leach AR (2001) In: Molecular modelling principles and applications, 2nd edn, Prentice Hall
46. Swanson JMJ, Henchman RH, McCammon JA (2004) Revisiting free energy calculations: a theoretical connection to MM/PBSA and direct calculation of the association free energy, Biophys J 86:67–74
47. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA, Cheatham TEI (2000) Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc Chem Res 33:889–897
48. Srinivasan J, Cheatham TEI, Cieplak P, Kollman PA, Case DA (1998) Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices, J Am Chem Soc 120:9401–9409
49. Wang J, Morin P, Wang W, Kollman PA (2001) Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of Efavirenz by docking and MM-PBSA. J Am Chem Soc 123:5221–5230
50. Hou T, Wang J, Li Y, Wang W (2011) Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. the accuracy of binding free energy calculations based on molecular dynamics simulations, J Chem Inf Model 51:69–82

Imaging Agent Binding to Amyloid Protofibrils

Abstract This section is regarding the study entitled "Identification of a Common Binding Mode for Imaging Agents to Amyloid Fibrils from Molecular Dynamics Simulations", which has been published in the Journal of the American Chemical Society.

Introduction

In this study the binding of 13 different amyloid binding imaging agents (Chart 1) to an amyloid fibril segment was investigated using MD simulations and free energy calculations. The binding of imaging agents to amyloid fibrils is of interest in the context of development of a novel, non-invasive method for diagnosis of AD. Currently, the guidelines for diagnosis of AD from the International Working Group for New Research Criteria for the Diagnosis of Alzheimer's Disease state a positive test for biomarkers for AD as a criterion [1]. These updated criteria were only recently released, and were not available when the present study was performed; however, they highlight the continued importance of the development and understanding of amyloid imaging agents.

The aim of this study is to determine commonalities of binding between the 13 different imaging compounds, as well as investigating which molecular traits are favorable for binding and which are unfavorable.

Experimental Section

An oligomeric segment of an amyloid fibril was used in this study as a model system to study the amyloid fibril binding of 13 different imaging compounds (Chart 1). The oligomeric segment is composed of two layers of antiparallel β-sheets with ten β-strands in each (Fig. 1a). Each peptide is a seven-residue

K.K. Skeby, *Computational Modelling of the Human Islet Amyloid Polypeptide*,
Springer Theses, DOI 10.1007/978-3-319-20040-8_3

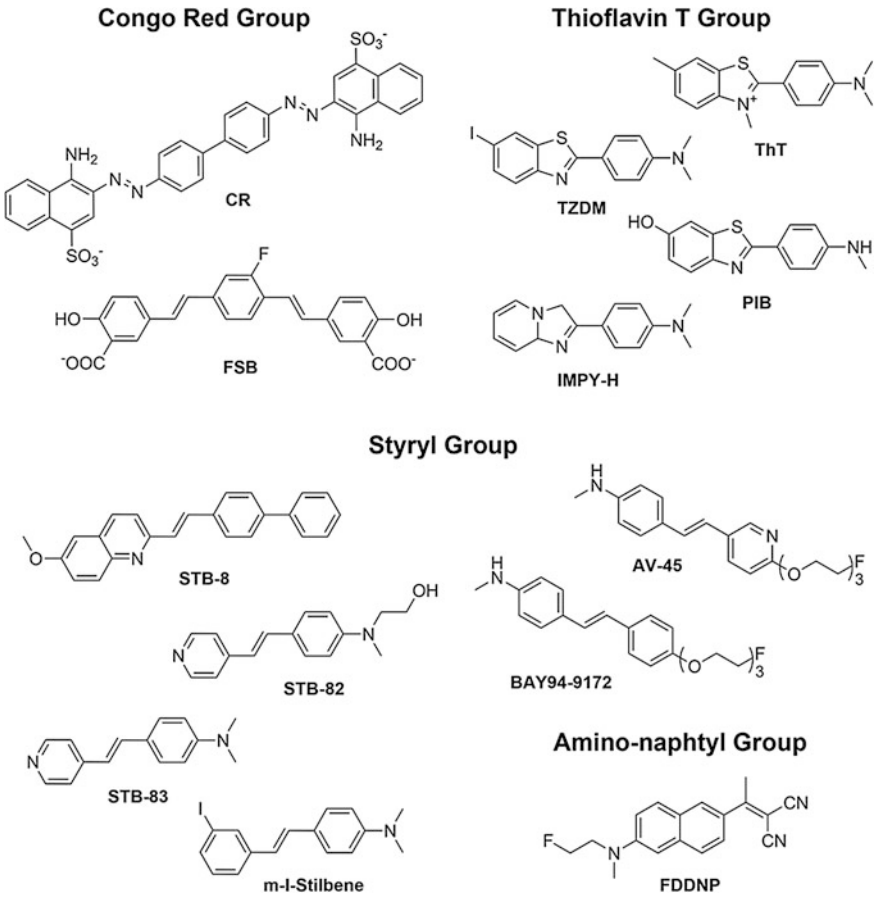


Chart 1 Imaging agents investigated in the present study. Four classes of imaging agents are included; the CR group, the ThT group, the styryl group, and the amino-naphtyl group of compounds. Reprinted with permission from Skeby et al. [2]. Copyright 2013 American Chemical Society

segment (NFGAILS) from hIAPP (residues 22–28), the structure of which was determined using ss-NMR spectroscopy (PDB-code: 2KIB) [3].

The fibril used for the ss-NMR measurements was composed of 10-residue peptide fragments ($hIAPP_{20–29}$), of which only the central NFGAIL was isotopically labeled. $hIAPP_{20–29}$ has been shown to be highly important for fibril formation, and was originally thought to form the core of the hIAPP amyloid fibril [4, 5]. The additional serine residue in the final structural model was included to complete the hydrogen bonding network during the structure calculation. To mimic the presence of additional residues, all peptides in the MD setup were capped with an acetyl group or an amino methyl group at the N-terminus or C-terminus, respectively.

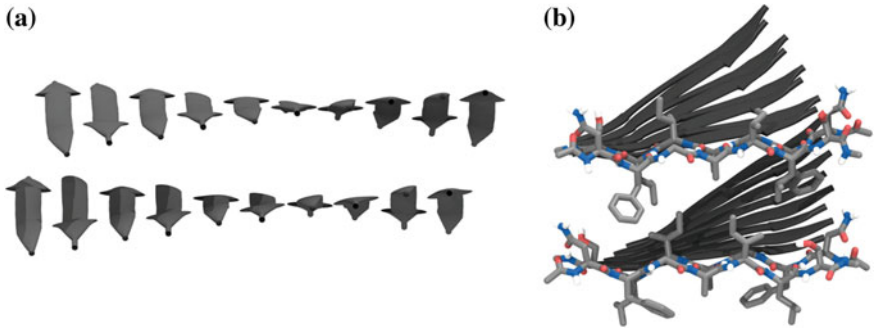


Fig. 1 Amyloid fibril used in the present study viewed from **a** the side and **b** the end. The side chains of the two peptide strands at the end of each sheet (4 in total) have been shown in (**b**). Reprinted with permission from Skeby et al. [2]. Copyright 2013 American Chemical Society

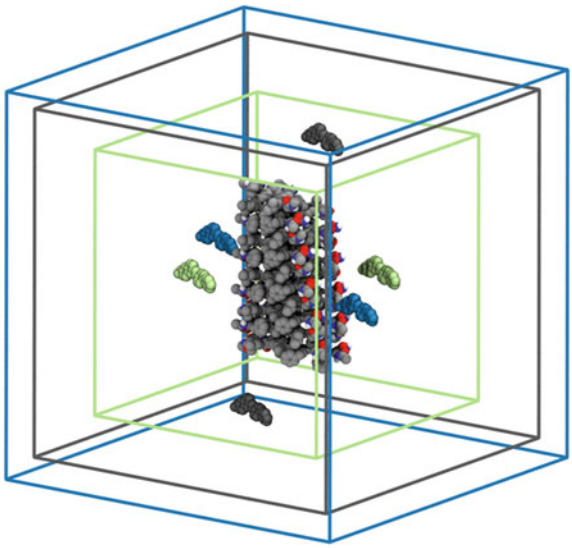


Fig. 2 Setups for simulation **A** (*gray*), **B** (*blue*), and **C** (*green*). Due to the varying placement of the ligands in the initial setups, the sizes of the simulation box differ slightly between the setups. Reprinted with permission from Skeby et al. [2]. Copyright 2013 American Chemical Society

Each setup contains one fibril and two ligands. The presence of two ligands allows greater sampling of the ligand binding process at less computational cost. Each ligand was placed at opposite sides of the fibril at least 12 Å away from any fibril atom. To eliminate bias due to placement of the ligand at any one surface of the fibril, three setups for each ligand were prepared in which the ligands were placed at different locations around the fibril (Fig. 2).

The systems were solvated with TIPS3P [6] water in a cubic box with at least 10 Å from any ligand or fibril atom to the edge of the box. The system was neutralized with either Na^+ or Cl^-. Each system was simulated for 20 ns with the NAMD 2.6 [7] software using the Duan et al. all-atom point-charge FF [8] (AMBER03) for the peptides and the general Amber FF [9] (GAFF) for the ligands.

GAFF was chosen due to the standardized method of generating FF parameters, which would aid in the generation of FF parameters for the 13 different ligands. For comparison, three simulations of the fibril without ligands present were also performed for 80 ns each. Details of the generation of ligand FF parameters and of the MD simulations can be found in the published paper.

Analysis of Trajectories

Each trajectory was split into two, one for each ligand, giving 120 ns of simulation for each ligand. All snapshots in which the fibril and ligand had >20 heavy atom contacts within 5 Å were considered for analysis. A clustering analysis was performed for each ligand trajectory after alignment of the protein C_α atoms using the following clustering algorithm:

1. Combine the two structures with the lowest internal RMSD to initiate a cluster.
2. Find the snapshot with the lowest RMSD relative to the average structure of the current cluster, and add it to the cluster.
3. The cluster is complete when the snapshot with the lowest RMSD with respect to the average of the current cluster has an RMSD value above 5 Å.
4. Restart from 1. to form a new cluster unless the RMSD between all remaining pairs of snapshots is >5 Å.

The representative structure of a cluster is the snapshot with the lowest RMSD to the average positions of the atoms in the cluster. Only clusters with more than 1 % of the total number of snapshots (min. 600 snapshots per cluster) were considered for further analysis.

The resulting clusters were grouped into binding modes based on the position and orientation of the ligand on the fibril. The position on the fibril refers to the different faces of the fibril (Fig. 3). Due to the antiparallel arrangement of the β-strands, the sides of the fibril (at the peptide termini) have the same physical and chemical properties. The ends are also similar, and are characterized by the protrusion of backbone amide groups. However, the two faces of the fibrils are different, and the properties are determined by the protruding side chains (Fig. 1). In the following, the face with Phe, Ala, and Leu will be termed the bottom face, and the face with Asn, Ile, and Ser will be termed the top face. The orientation refers to what the ligand is parallel with: fibril, peptide strands, or β-sheet normal (Fig. 3).

Furthermore, the MM-PBSA binding energy of all clusters was determined using the MM-PBSA method as implemented in AMBER11 [10]. The change in entropy was not considered in the calculation as the assumption that the change in entropy is the same for similar molecules is a commonly used approximation, which has been shown to lead to effective ranking of ligands [11]. The details of the calculations can be found in the published paper.

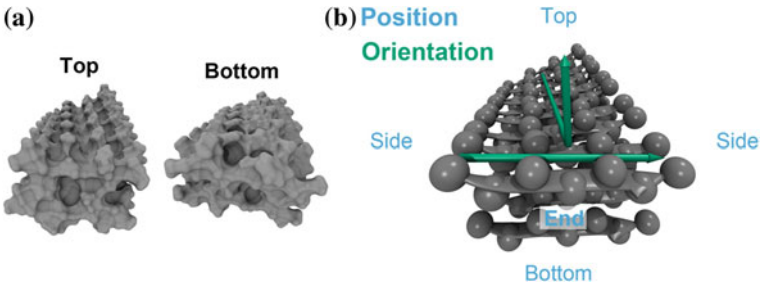


Fig. 3 Characterization of binding modes based on the position and orientation of the ligand. **a** Surface representation of the *top* and *bottom* faces of the fibril. The *top* face contains a large central groove between the Leu side chains, and two minor side grooves. The *bottom* face has no extended linear groove as found on the *top*; rather, a shallow crevice is seen. **b** The position of the ligand refers to the end, side, *top*, or *bottom* face of the fibril. The ligand orientation is described by the ligand axis being parallel with the fibril axis, the peptide strands, or the β-sheet normal, as indicated by the *green arrows*. Reprinted with permission from Skeby et al. [2]. Copyright 2013 American Chemical Society

Results

STB-8 will be used as an example to highlight the trends observed for the ensemble of ligands. The nomenclature for referring to a particular simulation and ligand will be Compound-**Setup**$_{\text{Ligand}}$, e.g. STB-8-**B**$_{\text{lig1}}$ refers to a simulation with STB-8 and the trajectory of ligand 1 as sampled in setup **B**.

The stability of the fibril structure was investigated in three simulations of the fibril without the presence of a ligand. Two different types of motion were identified to cause deformations of the fibril compared with the starting conformation. Two strands at each end of the fibril are not involved in the steric zipper interactions. The only interactions holding them in place are non-bonded interactions, including the hydrogen bonds, to the adjoining peptide in the β-sheet. This makes the end strands very dynamic. In two of the simulations (run1 and run3), this causes an increase in the RMSD of the fibril, which can be seen by measuring the RMSD of the fibril without the end strands (Fig. 4a). In the third simulation a twist of the fibril beyond what is observed in the ss-NMR structure causes the RMSD of the fibril to rise (Fig. 4b). These events are caused by the small size of the fibril and only arise after ∼30–50 ns, and should therefore not influence the fibril structure too much during the 20 ns ligand binding simulations.

The stability of the fibril in the simulations with ligands as measured by the C_α RMSD is comparable to the fibril in the simulations without ligand (Fig. 5a). The RMSD of the fibril is below 3 Å in most of the simulations with ligands, which could suggest that the ligands may actually have a stabilizing effect on the fibril. In only two out of 39 simulations the RMSD significantly exceeds 3 Å, which is approximately the value observed in the simulations without ligands at 20 ns. It occurs before the ligand binds and not as a consequence, which can be realized by

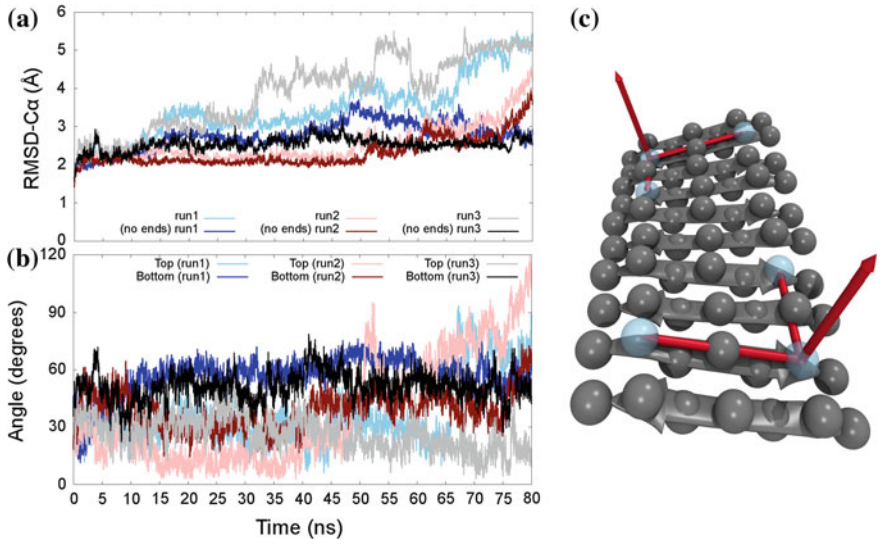


Fig. 4 **a** The C_α RMSD of the fibril simulations (run1, run2, and run3) without ligand present. The RMSD is calculated for the entire fibril, as well as for the fibril excluding the two protruding peptide strands (no ends), and is a running average over 20 ps. **b** The twist angle of the two β-sheets of the fibril calculated as a running average over 20 ps. **c** The twist angle of the β-sheets is measured as the angle between the β-sheet normals at each end of the fibril. The normal is measured at the second peptide strand from the end to avoid the fluctuations from the loose terminal peptides. Reprinted with permission from Skeby et al. [2]. Copyright 2013 American Chemical Society

comparing the RMSD plots with the contact plots. Additional twist of the fibril and the flexible end-peptides is also the reason for the fibril flexibility in the ligand simulations (Fig. 5b).

At least one out of six ligands binds to the fibril for the 13 imaging molecules, which can be seen by the number of ligand-fibril contacts normalized by the number of ligand atoms (Fig. 5c). The number of molecules that bind for a particular ligand depends on both the affinity of the ligand for the fibril as well as the random diffusion of the ligand in the simulation box. The ligands bind to different locations on the fibril, which leads to various degrees of burial and consequently a varying number of contacts even for the same ligand. Furthermore, the number of contacts per heavy atom varies between the ligands; the more hydrophilic ligands such as CR, FSB, and ThT do not form as many contacts per heavy atom as the more hydrophobic ligands such as STB-82. Binding of the ligands is a dynamic process; the ligands do not just stick to the fibril in the first location they encounter, they move around on the fibril. This can be seen by the changes in the number of contacts for the individual ligands. Furthermore, some events of unbinding and binding are observed, e.g. for STB-8-$\mathbf{C}_{\mathrm{lig1}}$ at 14 ns (dark green in Fig. 5c). These

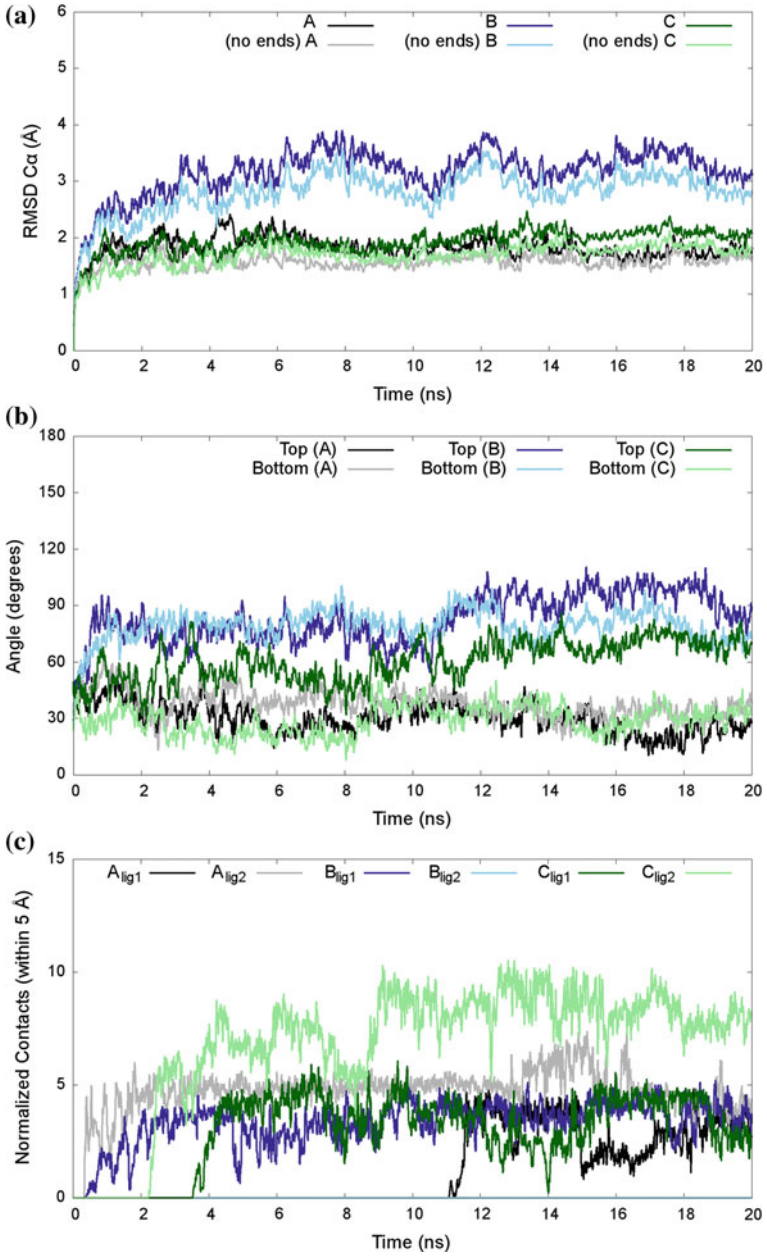


Fig. 5 **a** C_α RMSD of the fibril as a function of time in the three simulations (**A**, **B**, and **C**) with STB-8. The curves are running averages over 20 ps. **b** The twist of the β-sheets is a running average over 20 ps. **c** Atomic contacts of STB-8 with the fibril within 5 Å are plotted for all STB-8 molecules. These are also running averages over 20 ps. Plots for the rest of the ligands can be found in the SI in the published paper. Reprinted with permission from Skeby et al. [2]. Copyright 2013 American Chemical Society

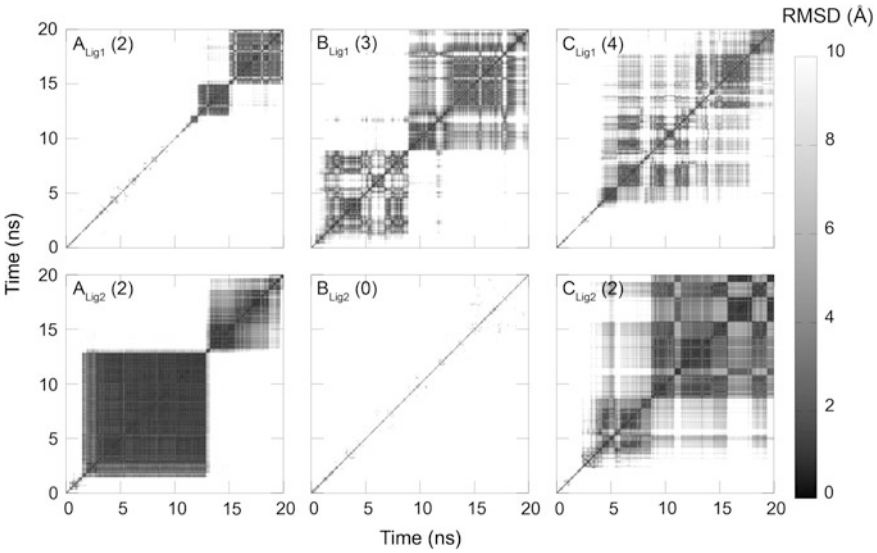


Fig. 6 RMSD matrices for STB-8. The fibril is aligned by the Cα atoms to the starting structure, after which the ligand heavy-atom RMSD is calculated between all pairs of snapshots. A *dark square* indicates that the ligand stays in approximately the same binding mode. The numbers in parenthesis are the number of identified clusters by the clustering analysis. RMSD matrices for the remaining ligands can be found in the SI in the published paper. Reprinted with permission from Skeby et al. [2]. Copyright 2013 American Chemical Society

observations suggest that the more stable fibril-ligand complexes are also specific low-energy complexes and not just sticky encounters.

The fluorescence of ThT has been suggested to arise from an excited dimer [12, 13]; however, dimerization of ThT is not observed in this study, which is most likely due to the positive charge on ThT. It is, however, observed in simulations for other ligands (AV-45-**C**, FDDNP-**A**, FDDNP-**B**, and IMPY-H-**A**). The ligands form a dimer before binding to the fibril probably because the ligands are hydrophobic and aromatic, rendering the interaction in the dimer more favorable than the interaction with water.

The RMSD matrices of the ligands after alignment of the fibrils show the stability and dynamics of the ligands on the fibril (Fig. 6). It is evident that the different ligand trajectories follow different paths. Some ligands find a stable binding mode quickly (STB-8-$\mathbf{A}_{lig2}$), while other ligands move on the surface finding several less stable binding modes (STB-8-$\mathbf{C}_{lig2}$). However, even if the ligand has found a stable binding mode, it is still able to move on the fibril.

This dynamic movement of the ligands on the fibril makes a clustering analysis necessary to determine uniform ligand-fibril clusters. Several clustering algorithms were tested before the algorithm outlined in the experimental section was chosen. This method provides well-defined and spherical clusters, the number of which correlates well with the appearance of the RMSD matrices (Fig. 6).

Further visual categorization of the clusters into binding modes was performed based on the position and orientation of the ligand on the fibril as outlined in the experimental section (Fig. 3). The top and bottom faces of the fibril are characterized by grooves created by protruding side chains. Three grooves are present on the top face; a central groove created by Ile, and two minor side grooves created by Ile on one side and Asp and Ser on the other. The bottom face has only a single minor groove because the large Phe side chains are oriented toward the center of the fibril, closing the two grooves that would otherwise be present; however, the central Ala side chain is small enough that a single minor hydrophobic groove is created. This pattern of elongated grooves on the fibril is present on all amyloid fibrils, with the physical-chemical properties of the grooves determined by the side chains on the adjacent residues. The orientation of the ligand refers to what the ligand is parallel with; the fibril axis, the peptide strands, or the β-sheet normal. Binding modes are referred to using the following convention: $\text{Position}_{Orientation}$. Binding to the top and bottom faces parallel with the fibril was further subdivided into two binding modes, one in which the ligand binds in the central groove, and one where the ligand binds over the side chains.

All four faces of the fibril gave rise to clusters, however, not all orientations were observed for all faces. The populations of all clusters can be seen in Fig. 7 along with the energies of the clusters.

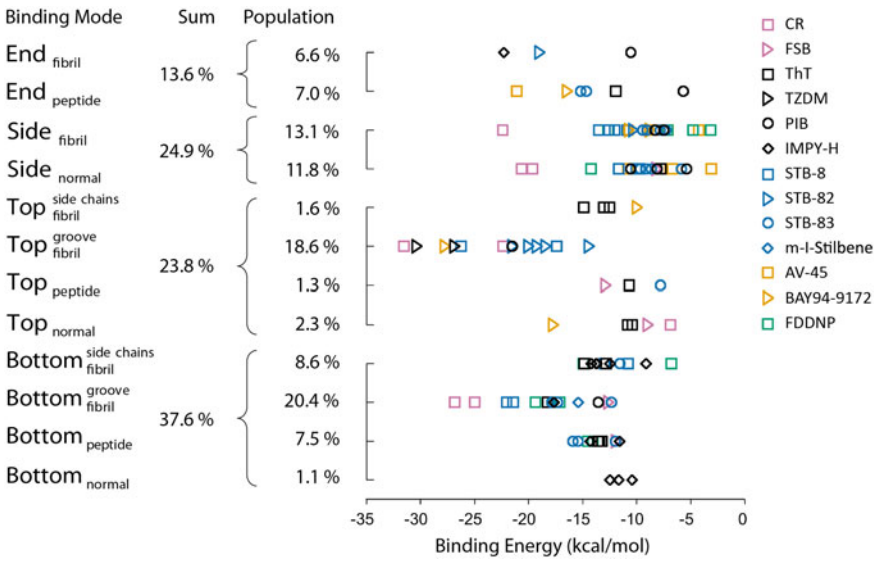


Fig. 7 Population of the identified binding modes and the estimated free energy of binding of the individual clusters. The population of a binding mode (including all clusters belonging to each binding mode) is represented by the fraction of snapshots in the binding mode of the total number of clustered snapshots which is 271,481. Each point represents the MM-PBSA binding free energy of a single ligand cluster. Lower energy means a higher binding affinity. Table S2 in the SI of the published paper contains the data used for creating this figure including standard deviations of the binding energies. Reprinted with permission from Skeby et al. [2]. Copyright 2013 American Chemical Society

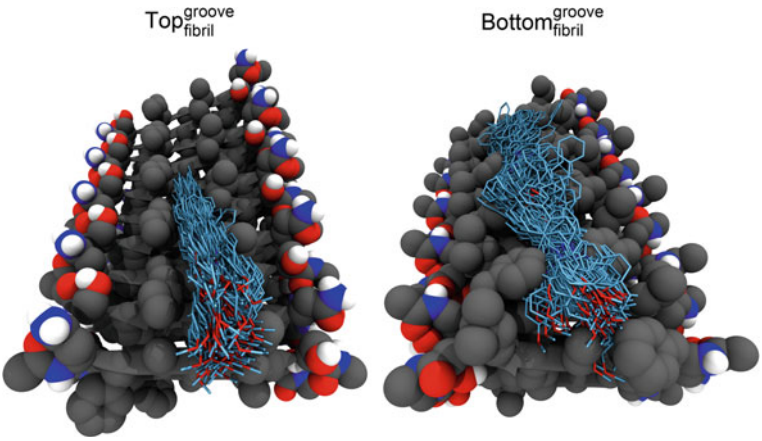


Fig. 8 Example of the Top$^{groove}_{fibril}$ and Bottom$^{groove}_{fibril}$ binding modes for STB-8. Reprinted with permission from Skeby et al. [2]. Copyright 2013 American Chemical Society

Two binding modes have significantly higher populations than the rest; Top$^{groove}_{fibril}$ and Bottom$^{groove}_{fibril}$, each with populations close to 20 % of all complexes. As can be seen from Fig. 8, the ligand in these binding modes is parallel with the fibril axis buried between hydrophobic side chains. Interestingly, no binding occurs in the side grooves on the top face, which could be speculated to be caused by the hydrophilic side chains. Binding to the bottom face in two other binding modes, Bottom$^{side\ chains}_{fibril}$ and Bottom$_{peptide}$, is also quite populated, which is probably due to the possibility of favorable π-π interactions with the aromatic Phe side chains in these binding modes as well as in Bottom$^{groove}_{fibril}$. Contrary to this, the Top$_{normal}$, Top$_{peptide}$, and Top$^{side\ chains}_{fibril}$ binding modes all have neglible populations, which suggests that the top surface groove is easily accessible and has high affinity for the ligands.

The sides of the fibril are characterized by the hydrophobic capping groups, which results in quite high populations in the binding modes on this face. The distribution between the two orientations is even suggesting that no traits are present to favor one orientation over the other. This is most likely linked to the non-specific hydrophobic interactions, which are possible with the capping groups. It is difficult to judge the relevance of the binding modes on the sides of the fibrils, as these probably do not have the same properties as an in vivo or in vitro fibril of the full-length amyloidogenic peptide. The sides of the fibrils of full-length peptides and proteins are usually described as having flexible N- or C–terminal coils which cannot be determined with ss-NMR, making it difficult to say much about binding to the sides of the fibril.

Binding to the ends of the fibril is also not considered highly relevant as the concentration of fibril ends is expected to be low in the in vivo fibrils. In line with this, the populations of the end binding modes are not highly populated; however, there are two quite stable clusters in the End$_{fibril}$ binding mode from STB-82 and

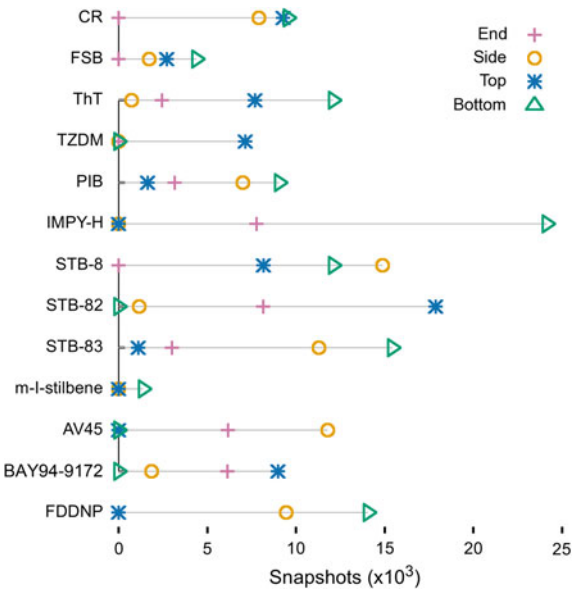


Fig. 9 Number of clustered snapshots for each ligand on the four faces of the fibril. Reprinted with permission from Skeby et al. [2]. Copyright 2013 American Chemical Society

IMPY-H. In these clusters the ligands have pushed in between the two β-sheets, resulting in a binding that is similar to both $\text{Top}_{fibril}^{groove}$ and $\text{Bottom}_{fibril}^{groove}$ simultaneously. To our knowledge, inhibition of fibril formation by STB-82 or IMPY-H has not been observed, however, this binding mode could possibly be relevant for other amyloid fibril formation inhibitors. CR is the only molecule included in our study, which has been shown to influence the fibril formation process, however, inhibition of amyloid aggregation has been observed for other small molecules [14–17].

Eleven out of the total 13 ligands have the highest number of snapshots to either the bottom or the top face (Fig. 9). This correlates well with the top and bottom faces accounting for the highest number of snapshots overall. Based on this, it seems that the faces with the hydrophobic surface grooves are the most favorable for the imaging agents. However, these faces also have a larger surface area than the end and side faces, which could influence the results. It is therefore also important to consider the energetics of binding.

It is also clear from the MM-PBSA binding energies that the $\text{Top}_{fibril}^{groove}$ and $\text{Bottom}_{fibril}^{groove}$ binding modes are the most favorable (Fig. 7). They have no clusters with binding energies above −10 kcal/mol and they have the highest number of clusters with binding energies below −15 kcal/mol. The clusters at the ends of the fibril where the ligand is placed between the two β-sheet layers also have quite low energies, which is not surprising as they also have a high number of contacts between the ligand and the fibril (Figure S3, STB-82-$\mathbf{A}_{\text{lig2}}$ and IMPY-H-$\mathbf{C}_{\text{lig2}}$). The side faces exhibit quite high binding energies suggesting that the binding on the side exhibits unspecific interactions and/or is a consequence of the relatively large surface area. The $\text{Top}_{fibril}^{groove}$ binding mode seems to have slightly lower binding

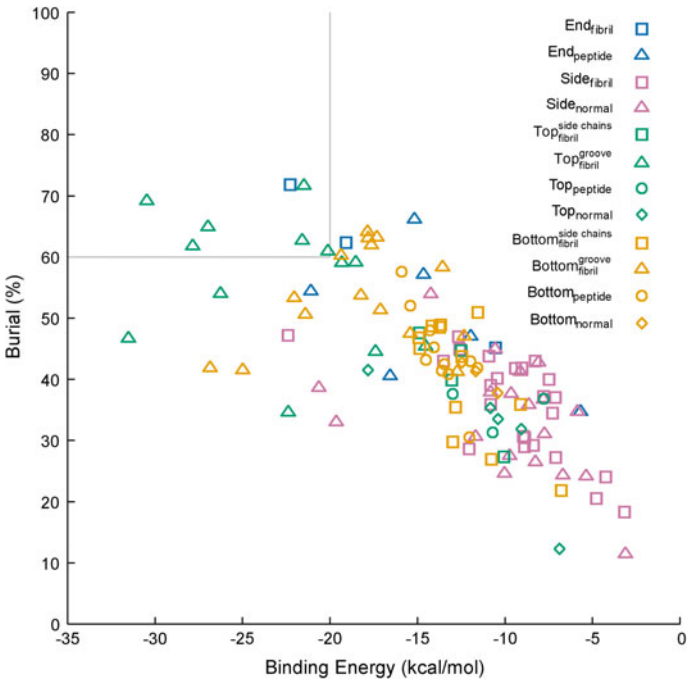


Fig. 10 Degrees of burial for all clusters as a function of the binding energy of the cluster. The burial percentage is calculated as the average ligand surface area that is lost upon association with the fibril; therefore, a burial of 100 % means that the ligand does not have any solvent accessible surface. The exact numbers and standard deviations are available in Table S2 in the SI of the published paper. Reprinted with permission from Skeby et al. [2]. Copyright 2013 American Chemical Society

energies than the $\text{Bottom}^{groove}_{fibril}$ binding mode, even though the $\text{Bottom}^{groove}_{fibril}$ binding mode has the highest number of snapshots. The three clusters with lowest energy in $\text{Bottom}^{groove}_{fibril}$ are CR clusters, and as will be described below, the binding of CR is slightly different from the rest of the ligands.

To understand the underlying mechanisms of what gives rise to the high binding affinity, the average degree of burial for each cluster has been calculated (Fig. 10). The degree of burial is determined as the fraction of the surface area of the ligand that is excluded from water upon contacting the fibril. The total surface area of the ligand and the buried surface area are calculated from the same trajectory snapshot. There is a trend towards lower energies for higher burial percentages, which is also as expected. Many of the clusters with high burial percentages (11 out of 14 with >60 % buried) are from the $\text{Top}^{groove}_{fibril}$ and $\text{Bottom}^{groove}_{fibril}$ due to the complementarity of the surface grooves with the linear shape of the ligands. Of the clusters with high burial and low energies, 6 out of 7 clusters are from the $\text{Top}^{groove}_{fibril}$, which exemplifies that the surface groove on the top face, which is flanked by hydrophobic side chains, is excellent for binding these imaging agents.

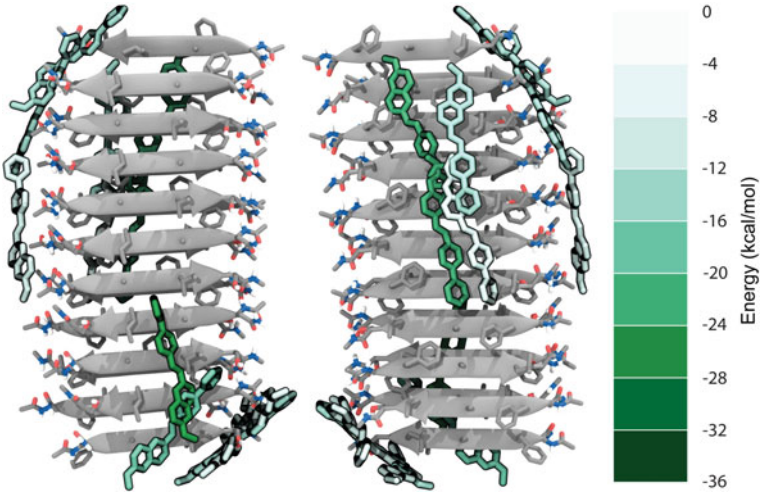


Fig. 11 Representative structures of the STB-8 clusters, colored by the binding energy of the cluster. The *left* figure is a view of the *top* layer and the *right* is a view of the *bottom* layer. The representative structure is the structure with the lowest RMSD to the average structure in the cluster. For clarity, the fibril structure shown is the minimized starting structure. Representative structures for clusters of the 12 other ligands can be found in Figure S6 in the SI of the published paper. Reprinted with permission from Skeby et al. [2]. Copyright 2013 American Chemical Society

The higher affinity of the clusters in the surface grooves becomes obvious when the representative structure for each cluster is colored by the energy of the cluster (Fig. 11). The structures in the grooves on the fibril have a stronger color than the structures binding over the side chains or on the side of the fibril.

We can understand what governs the binding by examining the different contributions to the MM–PBSA binding energy. The binding energy is a combination of a solvation term (a polar term, E_{PB}, and a non-polar term, E_{cavity}) and a term arising from the non-bonded FF interactions (E_{vdW} and $E_{electrostatic}$). The relative contributions of each term to the binding energy of each ligand, i.e. the contribution divided by the total energy, is shown in Fig. 12a, and is a population weighted average over all clusters for the ligand. The non-bonded interactions are not surprisingly favorable upon binding of the ligands, with the electrostatic interactions being more favorable for the two double-charged ligands, CR and FSB. The unfavorable contribution to the binding energy comes from removing the ligand-water interactions (E_{PB}). The change in this term is larger for the polar ligands, CR and FSB, as the energetic cost of removing polar and charged molecules from water is larger than for non-polar molecules.

Combining the polar and non-polar contributions to the binding energy allows us to make some conclusions about which of the two contribute most to the binding energy (Fig. 12b). The polar contribution to the binding energy falls within a quite narrow range. The non-polar contribution to the binding energy shows much more

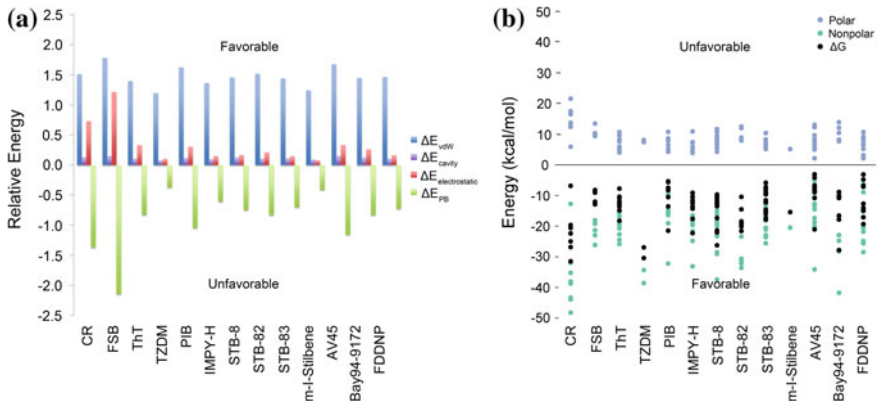


Fig. 12 **a** Average relative contribution to the binding energy by the energetic components. Positive values constitute a favorable contribution to the energy. The gas-phase energetic contributions are the vdW (E_{vdW}) and electrostatic ($E_{electrostatic}$) energies. The solvation energy is divided into a polar term (E_{PB}) and a non-polar term (E_{cavity}). **b** Non-polar (E_{vdW} + E_{cavity}, marked as *green dots*) and polar (E_{PB} + $E_{electrostatic}$, marked as *purple dots*) contributions to the free energy of binding (ΔG, marked with *black dots*) of all clusters for the 13 ligands. Reprinted with permission from Skeby et al. [2]. Copyright 2013 American Chemical Society

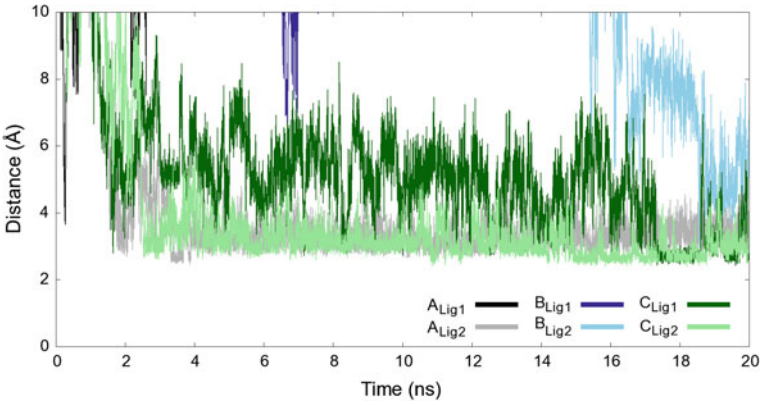


Fig. 13 The shortest distance from any hydrogen bond donor or acceptor heavy-atom in CR to any donor or acceptor heavy-atom in the Ser and Asp side chains of the fibril. Reprinted with permission from Skeby et al. [2]. Copyright 2013 American Chemical Society

variation, which seems to correlate with the variation in the binding energy. This correlates well with the observation that the clusters with higher burial percentage, which provides higher non-polar contact surfaces, have a lower binding energy.

CR has a very low binding energy in almost all clusters due to a highly favorable non-polar energy contribution, which is most likely due to the large size of CR (Fig. 12b). When CR binds on the top face of the fibril, it does not position itself

entirely parallel to the fibril axis; it is positioned diagonally across the surface of the fibril. This enables the two sulfate groups, one at each end of CR, to hydrogen bond with the Asp and Ser residues on the sides of the top face (Figure S6). This is seen for CR-$\mathbf{A}_{\mathrm{Lig2}}$ and CR-$\mathbf{C}_{\mathrm{Lig2}}$ in Fig. 13. When CR binds to the bottom face of the fibril, the Ser and Asp residues of the bottom β-sheet, the side chains of which are involved in steric zipper interactions, break the backbone hydrogen bonds to allow for transient interactions with the CR sulfate groups as seen for CR-$\mathbf{C}_{\mathrm{Lig1}}$.

Discussion

Using MD simulations we have investigated the binding of 13 different imaging agents to a model segment of an amyloid fibril. The fibril contains the necessary structural elements to make general considerations regarding the amyloid fibril binding of the imaging agents covered in this study.

Previously, it has been suggested that three types of binding sites for imaging agents exist on Aβ amyloid fibrils, each of them characterized by the type of ligand which binds to it; CR, ThT, or FDDNP [18–20]. The styryl-based compounds such as AV-45 have been shown to share the same binding site as ThT and analogs [18]. As they share part of the same scaffold structure it is an indication that this is important for the binding.

We also find that the aromatic and rigid scaffold of the ligands plays an important role in the binding. Most of the ligands favored binding in the surface grooves created by the repeating side chains, which is complementary to the elongated and hydrophobic ligands. Due to the small size of the fibril we could not distinguish between different binding sites for the different classes of ligands; however, subtle differences in the binding were observed based on the particular functionalities of the ligands. E.g. CR binds diagonally on the top surface of the fibril in order for the sulfate groups on each end of CR to interact with the hydrophilic Asp and Ser side chains. Previously, the distance between the negative charges on CR has been shown to be important for the binding affinity to amyloid fibrils [21]. Further evidence for this relation comes from a combined docking and mutation study in which the predicted binding of CR to a Lys residue in a prion (HET-s) amyloid fibril led to the subsequent Lys to Ala mutation which eliminated the binding [22].

The ThT group of ligands is the most highly studied in relation to the binding to amyloid fibrils. Multiple binding sites for ThT and analogs on fibrils formed by Aβ have been shown to be present; however, no structural information was available [23]. Structural information was, however, available in a liquid state NMR study, which showed the presence of two distinct binding modes for ThT binding to hIAPP fibrils; a high-affinity binding mode and a low-affinity binding mode [24]. Broadening of specific ThT peaks in the spectrum upon binding to the fibril led to the suggestion of ThT binding perpendicular to the β-sheet surface as in the Top_{normal} binding mode [24]. During certain conditions only a single peak,

consistent with the deeper binding of ThT in the surface grooves such as in the Top_{fibril}^{groove} and $Bottom_{fibril}^{groove}$ binding modes, remained in the spectrum [24]. MD simulations have also previously been used to compare the binding of ThT with neutral analogs. A study by Wu et al. showed the presence of two binding modes for ThT and the neutral analog BTA-1 (2-(4′-methylaminophenyl) benzothiazole), one which is similar to the Top_{fibril}^{groove} and $Bottom_{fibril}^{groove}$ binding modes, and one where the ligand is bound to the ends of the fibril segment [25]. As previously discussed, the end binding mode is most likely an artifact of the small size of the fibril in the simulation, as both in vivo and in vitro fibrils are much longer. Another MD study compared the binding of ThT to fibrils formed by Aβ with the binding of PIB to determine the reasons for the stronger binding of PIB [26]. It was found that PIB inserts deeper into the surface grooves on the fibril than ThT, most likely due to the positive charge on ThT [26]. This is consistent with our observations of the dependence of the binding energy on the degree of burial. Furthermore, the observation in a QM study that the binding of ThT and analogs is mainly driven by dispersion effects is consistent with our results showing that the polar contribution to the binding energy is nearly constant across the ligands, and that the binding energy is determined by the non-polar energy [27].

The recent FDA approval of two agents (^{18}F-florbetaben [28] and ^{18}F-flutemetamol [29]) for the detection of amyloid in brain tissue using PET has sparked a discussion of the usage of amyloid imaging in the diagnosis of AD [30, 31]. The usage of the imaging agents was only approved for ruling out a diagnosis of AD, and not as a positive result for AD diagnosis. This is rooted in mainly two aspects; a technical aspect and a biological aspect [30]. The biological limitation is the uncertainty in the causal relationship between the amyloid load and disease symptoms. The prevailing notion that amyloid in some form is the main cytotoxic actor leads to the natural derivation that the amyloid load must be an indicator for the severity of the disease, which is the main reasoning behind the development of amyloid imaging agents. However, a linear relationship does not necessarily exist between cytotoxicity and amyloid load, which is an aspect that is yet to be elucidated. The technical limitation lies in the abilities of the imaging agents. If these are not specific enough and bind to molecular species besides amyloid, they are not optimally suited for identifying the presence and location of amyloid deposits. The in vivo specificity of the amyloid imaging agents is still to be determined [30], and may actually be the limiting factor in determining whether amyloid load can be used as an indicator of disease severity.

Conclusion

In this study we have identified a common binding mode for imaging agents to amyloid fibrils in the surface grooves created by the repeating pattern of side chains in the cross-β structure. This information can be used as a starting point for the

development of highly specific imaging agents. However, information about the fibril to which the imaging agents should bind is necessary to optimize both the polar and non-polar interactions between the ligand and the fibril. Recently, a ss–NMR structure of an in vivo fibril formed by Aβ was released [32, 33], which could prove to be a huge step in the direction toward the design of highly specific Aβ imaging agents.

References

1. Dubois B, Feldman HH, Jacova C, Hampel H, Molinuevo JL, Blennow K, DeKosky ST, Gauthier S, Selkoe D, Bateman R, Cappa S, Crutch S, Engelborghs S, Frisoni GB, Fox NC, Galasko D, Habert M, Jicha GA, Nordberg A, Pasquier F, Rabinovici G, Robert P, Rowe C, Salloway S, Sarazin M, Epelbaum S, de Souza LC, Vellas B, Visser PJ, Schneider L, Stern Y, Scheltens P, Cummings JL (2014) Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. Lancet Neurol 13:614–629
2. Skeby KK, Sorensen J, Schiott B (2013) Identification of a common binding mode for imaging agents to amyloid fibrils from molecular dynamics simulations. J Am Chem Soc 135:15114–15128
3. Nielsen JT, Bjerring M, Jeppesen MD, Pedersen RO, Pedersen JM, Hein KL, Vosegaard T, Skrydstrup T, Otzen D, Nielsen NC (2009) Unique Identification of Supramolecular structures in amyloid fibrils by solid-state NMR spectroscopy. Angew Chem Int Ed 48:2118–2121
4. Westermark P, Engström U, Johnson KH, Westermark GT, Betsholtz C (1990) Islet amyloid polypeptide: pinpointing amino acid residues linked to amyloid fibril formation. Proc Natl Acad Sci USA 87:5036–5040
5. Shim S, Gupta R, Ling YL, Strasfeld DB, Raleigh DP, Zanni MT (2009) Two-dimensional IR spectroscopy isotope labeling defines the pathway of amyloid formation with residue-specific resolution. Proc Natl Acad Sci USA 106:6614–6619
6. Neria E, Fischer S, Karplus M (1996) Simulation of activation free energies in molecular systems. J Chem Phys 105:1902–1921
7. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kalé L, Schulten K (2005) Scalable molecular dynamics with NAMD. J Comput Chem 26:1781–1802
8. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang J, Kollman PAA (2003) Point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J Comput Chem 24:1999–2012
9. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general amber force field. J Comput Chem 25:1157–1174
10. Case DA, Darden TA, Cheatham TE III, Simmerling CL, Wang J, Duke RE, Luo R, Walker RC, Zhang W, Merz KM, Roberts B, Wang B, Hayik S, Roitberg A, Seabra G, Kolossváry I, Wong KF, Paesani F, Vanicek J, Liu J, Wu X, Brozell SR, Steinbrecher T, Gohlke H, Cai Q, Ye X, Wang J, Hsieh MJ, Cui G, Roe DR, Mathews DH, Seetin MG, Sagui C, Babin V, Luchko T, Gusarov S, Kovalenko A, Kollman PA (2010) AMBER 11. University of California, San Francisco
11. Hou T, Wang J, Li Y, Wang W (2011) Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. J Chem Inf Model 51(1):69–82

12. Groenning M, Olsen L, van de Weert M, Flink JM, Frokjaer S, Jorgensen FS (2007) Study on the binding of thioflavin T to β-Sheet-rich and non-β-sheet cavities. J Struct Biol 158(3):358–369
13. Groenning M, Norrman M, Flink JM, van de Weert M, Bukrinsky JT, Schluckebier G, Frokjaer S (2007) Binding mode of thioflavin T in insulin amyloid fibrils. J Struct Biol 159:483–497
14. Härd T, Lendel C (2012) Inhibition of amyloid formation. J Mol Biol 421:441–465
15. Porat Y, Abramowitz A, Gazit E (2006) Inhibition of amyloid fibril formation by polyphenols: structural similarity and aromatic interactions as a common inhibition mechanism. Chem Biol Drug Des 67:27–37
16. Lorenzo A, Yankner BA (1994) β-amyloid neurotoxicity requires fibril formation and is inhibited by congo red. Proc Natl Acad Sci USA 91:12243–12247
17. Caughey B, Ernst D, Race RE (1993) Congo red inhibition of scrapie agent replication. J Virol 67:6270–6272
18. Cai L, Innis RB, Pike VW (2007) Radioligand development for PET imaging of β-amyloid (Aβ)-current status. Curr Med Chem 14:19–52
19. Agdeppa ED, Kepe V, Liu J, Flores-Torres S, Satyamurthy N, Petric A, Cole GM, Small GW, Huang SC, Barrio JR (2001) Binding characteristics of radiofluorinated 6-dialkylamino-2-naphthylethylidene derivatives as positron emission tomography imaging probes for β-amyloid plaques in alzheimer's disease. J Neurosci 21:RC189
20. Zhuang ZP, Kung MP, Hou C, Skovronsky DM, Gur TL, Plossl K, Trojanowski JQ, Lee VMY, Kung HF (2001) Radio iodinated styryl benzenes thioflavins as probes for amyloid aggregates. J Med Chem 44:1905–1914
21. Ashburn TT, Han H, McGuinness BF, Lansbury PT Jr (1996) Amyloid probes based on congo red distinguish between fibrils comprising different peptides. Chem Biol 3:351–358
22. Schütz AK, Soragni A, Hornemann S, Aguzzi A, Ernst M, Boeckmann A, Meier BH (2011) The amyloid-congo red interface at atomic resolution. Angew Chem Int Ed 50:5956–5960
23. Lockhart A, Ye L, Judd DB, Merritt AT, Lowe PN, Morgenstern JL, Hong GZ, Gee AD, Brown J (2005) Evidence for the presence of three distinct binding sites for the thioflavin T class of Alzheimer's disease PET imaging agents on β-amyloid peptide fibrils. J Biol Chem 280:7677–7684
24. Robbins KJ, Liu G, Lin G, Lazo ND (2011) Detection of strongly bound thioflavin T species in amyloid fibrils by ligand-detected 1H NMR. J Phys Chem Lett 2:735–740
25. Wu C, Wang Z, Lei H, Duan Y, Bowers MT, Shea J (2008) The binding of thioflavin T and its neutral analog BTA-1 to protofibrils of the Alzheimer's disease Aβ16-22 peptide probed by molecular dynamics simulations. J Mol Biol 384:718–729
26. Wu C, Bowers MT, Shea J (2011) On the origin of the stronger binding of PIB over thioflavin T to protofibrils of the Alzheimer amyloid-β peptide: a molecular dynamics study. Biophys J 100:1316–1324
27. Ali-Torres J, Rimola A, Rodriguez-Rodriguez C, Rodriguez-Santiago L, Sodupe M (2013) Insights on the binding of thioflavin derivative markers to amyloid-like fibril models from quantum chemical calculations. J Phys Chem B 117:6674–6680
28. American Food and Drug Administration (2012) FDA approves ^{18}F-Florbetapir PET Agent J Nucl Med 53(6):15 N
29. FDA press release: FDA approves second brain imaging drug to help evaluate patients for Alzheimer's disease, dementia. http://www.fda.gov/newsevents/newsroom/pressannouncements/ucm372261.htm. Accessed 21 July 2014
30. Kepe V, Moghbel MC, Langstrom B, Zaidi H, Vinters HV, Huang S, Satyamurthy N, Doudet D, Mishani E, Cohen RM, Hoilund-Carlsen PF, Alavi A, Barrio JR (2013) Amyloid-β positron emission tomography imaging probes: a critical review. J Alzheimer's Dis 36(4): 613–631
31. Williams SCP (2013) Alzheimer's imaging agents struggle to find a market outside trials. Nat Med 19:1551

32. Lu J, Qiang W, Yau W, Schwieters C, Meredith S, Tycko R (2013) Molecular structure of β-amyloid fibrils in Alzheimer's disease brain tissue. Cell 154:1257–1268
33. Paravastu AK, Qahwash I, Leapman RD, Meredith SC, Tycko R Seeded (2009) Growth of β-amyloid fibrils from Alzheimer's brain-derived fibrils produces a distinct fibril structure. Proc Natl Acad Sci USA 106:7443–7448

Determining the Aggregation Prone Structure of hIAPP

This section presents the results described in the manuscript entitled "Conformational Dynamics of the Human Islet Amyloid Polypeptide in a Membrane Environment: Toward the Aggregation Prone Form". The study was initiated during my stay at the University of Illinois, Urbana-Champaign, and was continued after my return to Aarhus.

Introduction

This study investigates the initial events in hIAPP-membrane binding, the interactions of hIAPP with a mixed anionic-zwitterionic lipid membrane bilayer, and the influence of changes in pH on hIAPP.

Simulating a membrane binding event using a conventional all-atom membrane poses a sampling problem as the diffusion of lipids in the membrane is very slow. Insertion of peptides in the membrane requires the displacement of lipids; therefore, investigating how hIAPP binds to the membrane using a conventional membrane would require a lot of simulation time to obtain adequate sampling.

To overcome some of the challenges involved in simulations of membrane-peptide/protein association, the highly mobile membrane mimetic (HMMM) model has been developed [1]. This model uses short-tail lipids to represent the lipid head-groups and the first part of the hydrophobic center of the membrane. An organic solvent, dichloroethane (DCLE), is used to represent the hydrophobic core of the membrane. This model has the same hydrophobic/hydrophilic profile as a conventional membrane bilayer, but has the advantage of faster lipid diffusion, which enables faster association of proteins with membranes, and has been used to successfully simulate the membrane insertion of the GLA domain of the human coagulation factor VII [1].

K.K. Skeby, *Computational Modelling of the Human Islet Amyloid Polypeptide*,
Springer Theses, DOI 10.1007/978-3-319-20040-8_4

In this study, we have used the HMMM model to simulate the association of hIAPP with a mixed divalerylphosphatidylcholine:divalerylphosphatidylserine (DVPC:DVPS) lipid membrane.

Experimental Section

The protein structure of hIAPP is a ss-NMR structure which contains all 37 residues (PDB-code: 2L86) [2]. It was determined at pH 7.3 and has an amidated C-terminus as is also the case for the physiological form of hIAPP. Three forms of the protein were prepared, one for each protonation state of His18; δ (HSD), ε (HSE), and the positively charged (HSP) form. In the following, a particular peptide form will be referred to as HSD, HSE, or HSP, and the neutral His18 peptides will be referred to collectively as HSD/E. The overall charge of the protein is +3 for HSD/E, and +4 for HSP. To sample the neutral and charged His18 forms equally, two, two, and four systems were set up for HSD, HSE, and HSP, respectively. The initial orientation of the protein with respect to the membrane was also varied; in the A-orientation the side chain of His18 is pointing up, in the B-orientation the protein has been rotated 180° around the y-axis leaving the side chain of His18 pointing toward the membrane. The protein was placed at least 12 Å from the lipid membrane.

Simulations were performed with both the full-length peptide ($hIAPP_{1–37}$) as well as two peptide fragments; $hIAPP_{1–19}$ and $hIAPP_{20–37}$, containing residues 1–19 and 20–37, respectively (Table 1). Eight simulations were also performed with $hIAPP_{1–19}$, since it contains His18, while only four simulations were performed with $hIAPP_{20–37}$. $hIAPP_{1–19}$ was capped with an amino-methyl group at the C-terminus, and $hIAPP_{20–37}$ was capped with an acetyl group at the N-terminus. The simulations with $hIAPP_{1–37}$ were run for 100 ns each, while the peptide fragment simulations were run for 50 ns each.

Table 1 Overview of the all-atom MD simulations of hIAPP and fragments thereof

Peptide	His18	Replicas	Membrane (lipids)	Simulation time (ns)
$hIAPP_{1–19}$	His(δ)/(ε)	4 (2/2)	HMMM (36)	50
$hIAPP_{1–19}$	His(+)	4	HMMM (36)	50
$hIAPP_{20–37}$	–	4	HMMM (36)	50
$hIAPP_{1–37}$	His(δ)/(ε)	4 (2/2)	HMMM (49)	100
$hIAPP_{1–37}$	His(+)	4	HMMM (49)	100
$hIAPP_{1–37}$	His(δ)/(ε)	6 (3/3)	–	50
$hIAPP_{1–37}$	His(+)	4	–	50
$hIAPP_{1–37}$	His(δ)/(ε)	2 (1/1)	DOPC/DOPS (49)	20
$hIAPP_{1–37}$	His(+)	2	DOPC/DOPS (49)	20

The peptide sequence is indicated in the first column with the protonation state of His18 indicated in the second column. The numbers in parenthesis in column three indicate the number of replicas for (δ/ε) protonation of His18. The membrane column indicates the membrane as well as the number of lipids in each leaflet, and the final column indicates the simulation time for each replica

In order to compare the membrane bound peptide structure with the $hIAPP_{1–37}$ monomer structure in solution, three simulations of each of the neutral His18 peptides, and four simulations of the HSP peptide in solution were also performed. Each of these simulations was run for 50 ns (Table 1).

The lipid bilayer is modeled as the HMMM and is composed of a central hydrophobic phase of DCLE, which is separated from water by a 70:30 mixture of two types of short-tailed lipids, zwitterionic DVPC and anionic DVPS (Fig. 1) [1].

This ratio was chosen to match multiple conditions; the lipid composition of most hIAPP-membrane experiments; the lipid composition of the islet β-cell membrane; and the optimum lipid composition for acceleration of hIAPP aggregation [3, 4]. Both the upper and lower leaflet is composed of 49 lipids for the systems with $hIAPP_{1–37}$, and 36 lipids for the systems with $hIAPP_{1–19}$ and $hIAPP_{20–37}$. The lipids were initially placed in a 7 × 7 or 6 × 6 grid for each leaflet, with random orientation around the z-axis. The initial grid position of the DVPC and DVPS lipids was chosen randomly. The separation of the phosphates in the lipid head-groups of the two leaflets can be tuned by the amount of DCLE included between the leaflets. The separation was set to 38 Å to mimic the average bilayer thickness of DOPC and DOPS membranes [5]. The C-1 of the lipid tails were constrained in the z-direction using a weak harmonic potential with a force constant of 0.05 kcal/mol · $Å^2$ to ensure that the short-tail lipids did not fluctuate extensively perpendicular to the

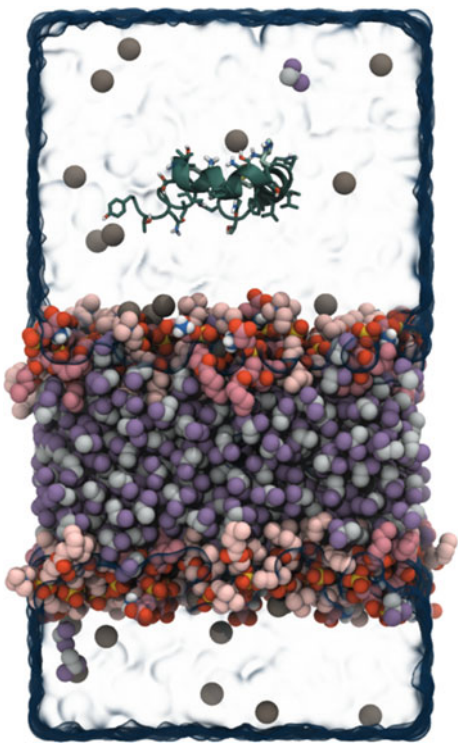


Fig. 1 The simulation setup is composed of one hIAPP peptide shown in dark green placed at least 12 Å above the lipid bilayer. The lipid bilayer is composed of a central hydrophobic core of DCLE shown in *purple* and *gray*, between two layers of DVPC and DVPS lipids shown in *pink colors* at the interface between DCLE and water. The system is neutralized with Na^+ shown in dark gray. Adapted with permission from Skeby et al. Biochemistry (2016) 10.1021/acs.biochem.5b00507. Copyright 2016 American Chemical Society

membrane compared to a membrane composed of long-tail lipids. The systems were solvated with TIP3P [6] water and neutralized with Na^+.

Prior to incorporation of the protein in the system, the membrane was minimized for 1000 steps and equilibrated for 1 ns using the same simulation procedure as the production runs explained below. The protein was incorporated by placing it in the water phase above the equilibrated membrane, and removing overlapping water molecules within 1.0 Å of the protein. Three or four Na^+ ions were removed to neutralize the system, depending on the protonation state of His18, giving a final Na^+ concentration of 0.10–0.12 M.

All systems were simulated in NAMD 2.9 using the CHARMM22* FF [7] for the protein, while the CHARMM36 FF [8] was used to simulate the lipids. Initially, the systems were minimized for 10,000 steps. Then, a short 10 ps equilibration of water was performed in the NVT ensemble while keeping everything but water and ions fixed. Finally, production runs were performed keeping the x-y area fixed to ensure an area per lipid approximately 8 % larger than the average 68 $Å^2$ for DOPC and DOPS lipids [5]. This procedure has been employed previously [9], and enables the protein to insert into the bilayer without creating too much surface tension. The sides (x and y direction) of the simulation box was 51.4 Å for the 36 lipids/leaflet-bilayer and 60 Å for the 49 lipids/leaflet-bilayer, leaving a final area per lipid of $\sim$73.5 $Å^2$. A time-step of 2 fs was used while constraining all bonds to hydrogen. The temperature was held constant at 310 K using a Langevin thermostat with a damping coefficient of 0.5 ps^{-1}. A Nosé-Hoover Langevin piston was used to keep the pressure constant at 1 atm with an oscillation period of 200 fs and a damping coefficient of 200 fs. Periodic boundary conditions were employed to reduce boundary effects, and PME with a grid-size of 1 Å was used to handle the electrostatic interactions in the system [10–12]. The non-bonded interactions were cut off at 12 Å with a switching function working from 10 Å. The pair-list contained pairs of atoms within 13.5 Å and was updated every 40 fs. The non-bonded interactions were calculated every 2 fs, and the full-electrostatics were calculated every 4 fs.

To probe the stability of the resulting membrane bound peptide species, the full-tail lipids were grown in four of the $hIAPP_{1–37}$-HMMM systems (Table 1). The four systems were HSD(A), HSE(A), HSP(A1) and HSP(B2), and were chosen based on the peptide structure at 100 ns. HSD(A) and HSP(A1) were representative for the overall simulations, while HSE(A) and HSP(B2) were outlying structures. To keep the existing interaction between the protein and the lipids, the coordinates of the atoms of the short-tail lipids were retained. A random DOPS or DOPC lipid from a pre-equilibrated membrane was chosen, and the head-group and the first five atoms of the lipid tail were aligned with an existing short-tail lipid. The missing atom coordinates were then adopted from the pre-equilibrated full-tail lipid. This procedure was repeated for each lipid in the system. The system was then minimized for 10,000 steps, which was followed by an equilibration of the newly added atoms along with water, keeping the protein as well as the lipid head-groups and first five atoms of the tails fixed. The systems were then simulated without constraints in the NPT ensemble for 20 ns each.

Results

The simulations of hIAPP binding to the membrane bilayer are presented below. The focus will be on the simulations of the full-length peptide ($hIAPP_{1–37}$) bound to the HMMM membrane, which will be compared to the remaining MD simulations where it is relevant.

In the following, each peptide will be referred to as e.g. $hIAPP_{1–19}$, and specific simulations will be referred to as e.g. $hIAPP_{1–37}$-HSP(A1).

hIAPP Binding to the Membrane

Binding to the membrane occurs within the first 20 ns for the $hIAPP_{1–19}$ and $hIAPP_{1–37}$ peptides (Fig. 2). Subsequently, none of the peptides dissociate from the membrane within the simulation time. Therefore, all analyses have been performed on the remaining 30 or 80 ns of simulation, to ensure that binding has occurred, and to be able to do a straightforward comparison between the different simulations. Four of the $hIAPP_{1–37}$ peptides have His18 in a neutral configuration (HSD/E), and four of the setups have a positively charged His18 (HSP). During the final half of the $hIAPP_{1–37}$ simulations, the $hIAPP_{1–37}$-HSP peptides are moving slightly deeper into the membrane than the $hIAPP_{1-37}$-HSD/E peptides. The $hIAPP_{20–37}$ peptide

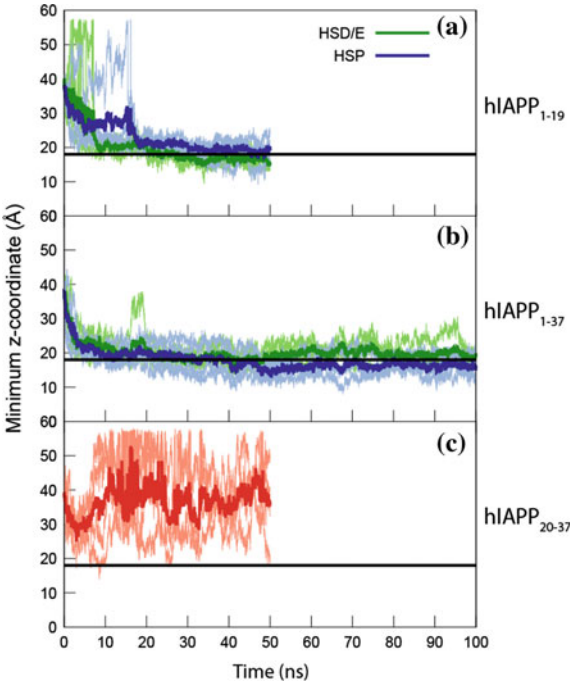


Fig. 2 Minimum z-coordinate of the protein heavy atoms for **a** the $hIAPP_{1–19}$ peptides, **b** the $hIAPP_{1–37}$ peptides, and **c** the $hIAPP_{20–37}$ peptides. The *lighter* colors represent individual peptides while the *darker* colors represent the average. The *black horizontal line* represents the average position of the lipid phosphate layer. Adapted with permission from Skeby et al. Biochemistry (2016) 10.1021/acs.biochem.5b00507. Copyright 2016 American Chemical Society

contacts the membrane, but quickly moves away again (Fig. 2c). It is clear that $hIAPP_{20–37}$ does not have affinity for the membrane, and that the interaction between the peptide and the membrane is a diffusional encounter due to the limited space in the simulation box. The behavior of the full peptide as well as the peptide fragments is in accordance with surface pressure experiments revealing that $hIAPP_{1–37}$ and $hIAPP_{1–19}$ inserts into the head-group region of 70:30 DOPC:DOPS phospholipid monolayers, while $hIAPP_{20–37}$ does not [13].

The N-terminus of $hIAPP_{1–37}$ binds to the membrane first for both the $hIAPP_{1–37}$-HSD/E and the $hIAPP_{1–37}$-HSP peptides (Fig. 3). This can probably be attributed to the positive charges present in this part of the peptide. Lys1 carries two positive charges, one at the N-terminus, and one on the side chain amino group. Arg11 also carries a positive charge. The $hIAPP_{1–37}$-HSD/E peptides have a loose interaction of the C-terminal part of the peptide with the membrane; it contacts the membrane occasionally, but does not stay there. The $hIAPP_{1–37}$-HSP peptides have a close interaction of the C-terminus with the membrane, which is most likely a result of the positive His18 interacting favorably with the anionic lipids, which then pulls the rest of the peptide closer to the membrane. The $hIAPP_{1–37}$-HSP(B2) peptide does not interact as closely with the membrane as the other $hIAPP_{1–37}$-HSP peptides (Appendix A). However, it is possible that this is due to a sampling issue and that the peptide would also bind to the membrane in a similar fashion as the other three $hIAPP_{1–37}$-HSP peptides if the simulation was extended.

The average positions of the side chain center of masses during the final 80 ns of the $hIAPP_{1–37}$ simulations highlights the differences in membrane position and orientation of the $hIAPP_{1–37}$ peptides in the membrane (Fig. 4). Both peptides have an average horizontal orientation with respect to the membrane. It is clear that the $hIAPP_{1–37}$-HSP peptides, except $hIAPP_{1–37}$-HSP(B2), have more of the peptide

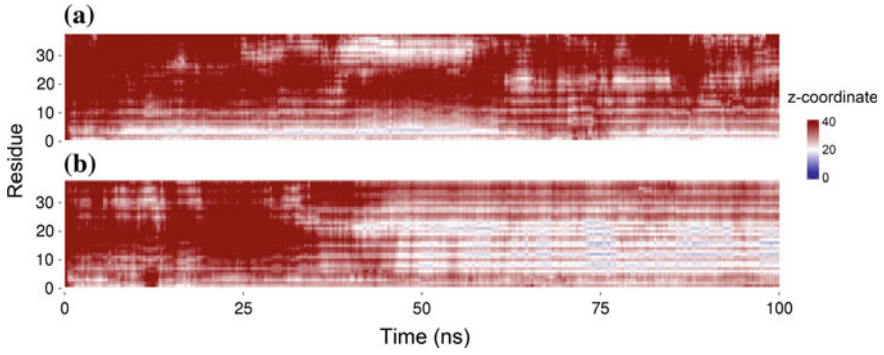


Fig. 3 The z-coordinate of the center of mass of each residue side chain for two representative $hIAPP_{1–37}$ simulations, **a** $hIAPP_{1–37}$-HSD(A) and **b** $hIAPP_{1–37}$-HSP(A1). Plots for all the simulations can be found in Appendix A. The center of the membrane is at z = 0 Å and the average position of the phosphates is at z = 18 Å. Data points with z > 40 Å are the same color as z = 40 Å. Adapted with permission from Skeby et al. Biochemistry (2016) 10.1021/acs.biochem.5b00507. Copyright 2016 American Chemical Society

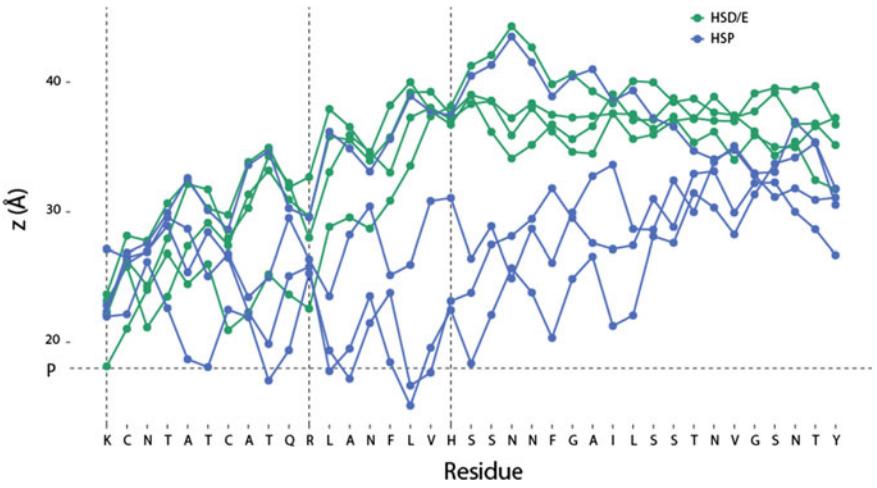


Fig. 4 Average z-coordinate of the peptide side chain center of mass during the final 80 ns of the $hIAPP_{1–37}$ simulations. The *horizontal dashed line* represents the average position of the phosphates, and the *vertical dashed lines* are the positive residues. Adapted with permission from Skeby et al. Biochemistry (2016) 10.1021/acs.biochem.5b00507. Copyright 2016 American Chemical Society

positioned closer to the membrane than the $hIAPP_{1–37}$-HSD/E peptides. Furthermore, the rotation of the helical structure in the N-terminus of the peptide is inverted for the two peptide types. The $hIAPP_{1–37}$-HSD/E peptides have approximately every 3–4 residues, including the positively charged residues, pointing downward so that these can interact with the anionic lipids. The inverse situation exists for the $hIAPP_{1–37}$-HSP peptides, which have the positive residues pointing upward, as this peptide is buried deeper in the head-group region, and the positive residues are interacting with the anionic head-groups as well as the solvent. This trace is characteristic of an α-helix oriented parallel with a membrane bilayer.

The horizontal orientation observed for both peptides, as well as the rotation of the $hIAPP_{1–37}$-HSP peptides, is consistent with accessibility measurements of hIAPP bound to an 80 % anionic membrane [14]. It is unexpected that the experimental rotation matches the $hIAPP_{1–37}$-HSP peptide; since the measurements were performed at neutral pH, we would expect the rotation of the $hIAPP_{1–37}$-HSD/E peptides to match the experiments. However, the lipid membrane in the accessibility experiments has a very high content of anionic lipids. The free form of histidine normally has a pK_a around 6, making the protonation state highly susceptible to changes in the immediate environment. The high anionic lipid content may shift the pK_a value of His18 upwards and make His18 protonated, providing the possibility of a favorable interaction between His18 and the anionic lipids.

The final snapshot of the representative $hIAPP_{1–37}$-HSD(A) and $hIAPP_{1–37}$-HSP (A1) simulations illustrates the difference in the positions of the peptides in the membrane (Fig. 5). $hIAPP_{1–37}$-HSP(A1) has displaced the short tail lipids and is

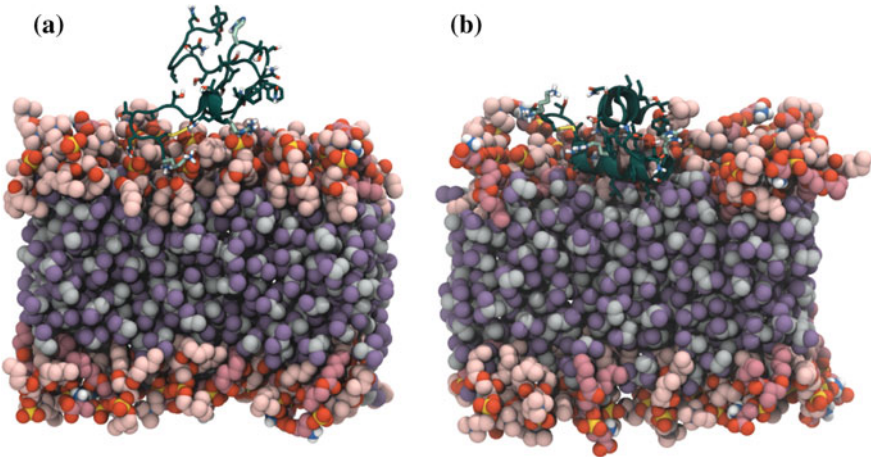


Fig. 5 Final snapshot of **a** $hIAPP_{1–37}$-HSD(A) and **b** $hIAPP_{1–37}$-HSP(A1). The lipids are shown in *red* colors, DCLE is shown in *gray* and *purple*, and the protein is shown in *dark green*. The final frame of all $hIAPP_{1–37}$ peptides bound to the HMMM membrane can be found in Appendix B. Adapted with permission from Skeby et al. Biochemistry (2016) 10.1021/acs.biochem.5b00507. Copyright 2016 American Chemical Society

interacting with the DCLE, while $hIAPP_{1–37}$-HSD(A) has only inserted the three N-terminal residues between the lipids, and the rest of the N-terminal part of $hIAPP_{1–37}$-HSD(A) is binding on top of the membrane. The C-terminus of the $hIAPP_{1–37}$-HSP(A1) peptide is associated closely with the lipids, while the C-terminus of $hIAPP_{1–37}$-HSD(A) is positioned in the solvent and not interacting with the membrane.

hIAPP Conformation and Dynamics

The conformation of the peptide changes significantly during the simulations (Fig. 6). This is not surprising as the starting structure is a ss-NMR structure of the peptide bound to an SDS micelle, whereas in these simulations the micelle is not present, and the peptide is initially placed in the water phase. The RMSD values of the HMMM bound $hIAPP_{1–19}$ and $hIAPP_{20–37}$ fragments plateau around 4–6 Å.

The RMSD of HMMM-bound $hIAPP_{1–37}$ is even higher and plateaus around 8 Å (Fig. 6e). This is mostly arising from the C-terminal part of the peptide (Fig. 6d), as the first 19 residues of the HMMM bound $hIAPP_{1–37}$ have a much lower RMSD around 4 Å (Fig. 6c). The RMSD values of the HMMM bound $hIAPP_{1–37}$ and the solvated peptide are quite similar; however, the RMSD of the first 19 residues of $hIAPP_{1–37}$ in solution (Fig. 6g) is higher than when the peptide is bound to the membrane. This suggests that the membrane stabilizes the structure in the N-terminal part of $hIAPP_{1–37}$, which is consistent with EPR experiments showing

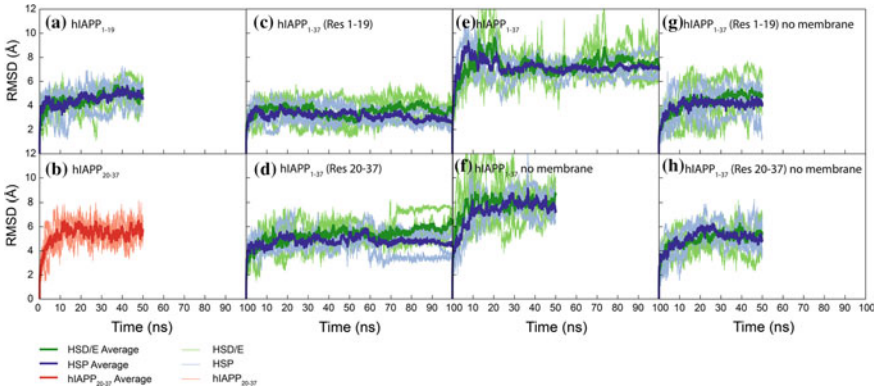


Fig. 6 C_α RMSD of the two peptide fragments, **a** $hIAPP_{1–19}$ and **b** $hIAPP_{20–37}$. The C_α RMSD of the $hIAPP_{1–37}$ peptide bound to the HMMM membrane was calculated for **c** residues 1–19, **d** residues 20–37, and **e** the entire $hIAPP_{1–37}$ peptide. The C_α RMSD of the $hIAPP_{1–37}$ peptide in solvent was calculated for **f** the entire $hIAPP_{1–37}$ peptide, **g** residues 1–19, and **h** residues 20–37. The *lighter* colors represent individual simulations while the *bold* colors represent averages. Adapted with permission from Skeby et al. Biochemistry (2016) 10.1021/acs.biochem.5b00507. Copyright 2016 American Chemical Society

stabilization of residues 9–20 upon membrane binding [14]. Furthermore, the high RMSD of both the free and membrane bound peptide suggests that both structures are different from the SDS bound peptide structure.

hIAPP is a natively-disordered peptide in solution. When it binds to a phospholipid membrane the amount of α-helix structure increases, however, hIAPP still does not fold into a well-defined structure [15]. This suggests that hIAPP should be very dynamic. The amount of flexibility in a protein is usually measured by the per residue root mean squared fluctuation (RMSF), which requires a structural alignment of the protein. A good structural alignment of different conformations of a protein requires that the ensemble is homogenous. However, hIAPP is too disordered to perform a good alignment (Fig. 7).

Therefore, an alternative measure of the protein flexibility is required. It is possible to look at the RMSF of the backbone torsion angles; however, this also poses a problem. Calculating the RMSF requires determining the average of a periodic quantity, which depends highly on the choice of origin. The solution is to convert the angles to Cartesian coordinates, and determine the arithmetic average in a two-dimensional coordinate system. The angle of the vector from the origin to the geometrical average is then the mean angle (Fig. 8a) [16].

The failure to produce a good alignment of the protein as well as the high RMSD value is already a good indicator that the peptide is very flexible, but we need to examine the RMSF to determine variations in flexibility over the peptide sequence. The torsional RMSF (Fig. 8) shows that most of the flexibility is in the C-terminal part of the peptide. This is in accordance with EPR measurements reporting on the

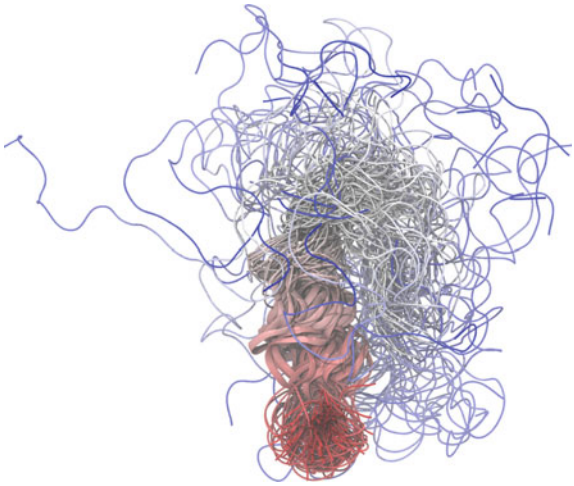


Fig. 7 Structural alignment of the $hIAPP_{1–37}$-HSD(A) peptide. All structures were aligned to the starting structure based on C_α atoms of the first 19 residues. The peptide is colored by sequence with the N-terminus colored *red* and the C-terminus colored *blue*. Structural alignments of all the $hIAPP_{1–37}$ peptides bound to the HMMM membrane can be found in Appendix C. Adapted with permission from Skeby et al. Biochemistry (2016) 10.1021/acs.biochem.5b00507. Copyright 2016 American Chemical Society

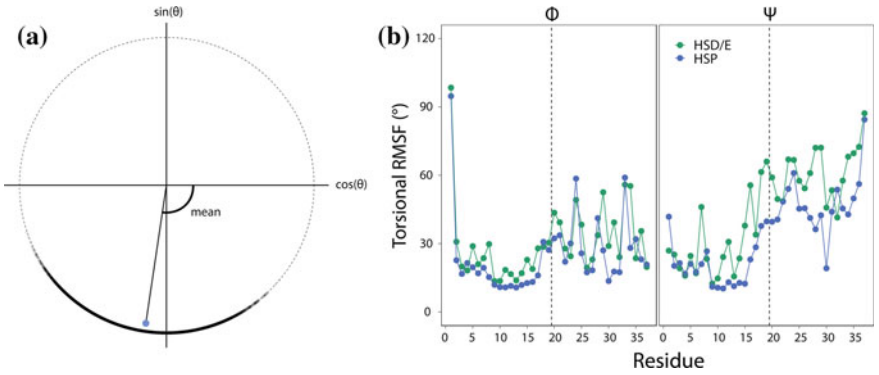


Fig. 8 (**a**) Determination of the mean angle. All angles are plotted on a unit circle, and the geometrical average is determined (*blue dot*). The mean is then the angle of the vector from the origin to the geometrical average. (**b**) Torsional RMSF of the $hIAPP_{1–37}$ HMMM backbone Φ and Ψ angles. Only the final 80 ns of each simulation are included in the analysis. The *vertical line* shows the divide between the 19 N-terminal residues and the 18 C-terminal residues. Adapted with permission from Skeby et al. Biochemistry (2016) 10.1021/acs.biochem.5b00507. Copyright 2016 American Chemical Society

mobility of each residue when the peptide is bound to the membrane, which shows that the C-terminal part of $hIAPP_{1–37}$ is more flexible than the N-terminal part [14].

The variations in the peptide flexibility of the different simulation setups may provide a clue to the mechanism of inhibition of fibril formation by lower pH and acceleration of fibril formation by the membrane. We have therefore looked at the differences in RMSF for the different simulation setups (Fig. 9). The torsional RMSF difference between the two protonation states of the $hIAPP_{1–37}$ peptides bound to the HMMM membrane confirms what we can see from Fig. 8, that $hIAPP_{1–37}$-HSP is not as dynamic as $hIAPP_{1–37}$-HSD/E (Fig. 9a). This trend is also present to some extent when the peptide is not bound to the membrane, however, not nearly as pronounced as for the membrane bound peptides (Fig. 9d).

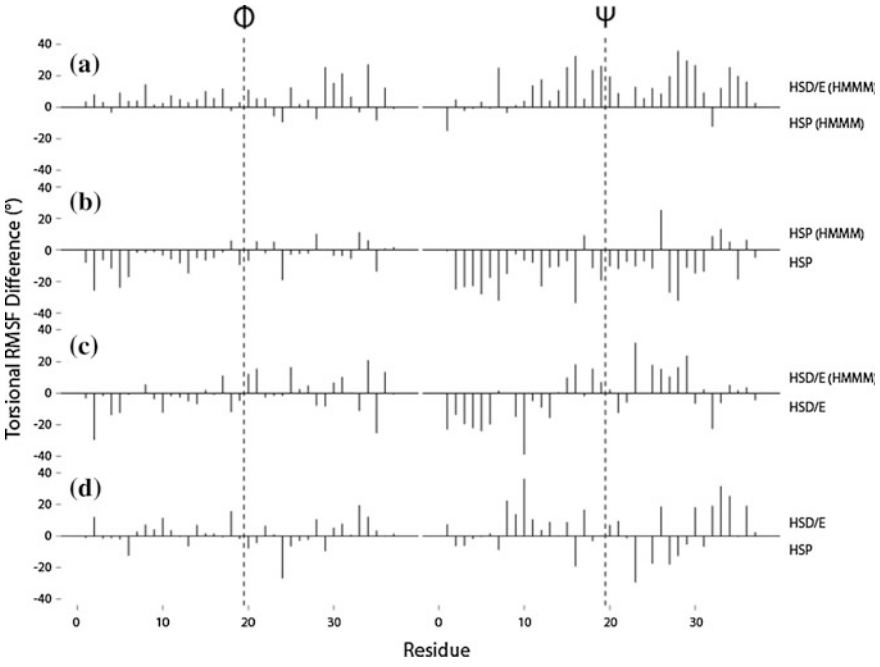


Fig. 9 Torsional RMSF differences. If the difference is positive, the top peptide is most flexible, whereas if the difference is negative, the bottom peptide is most flexible. **a** The difference between HMMM bound $hIAPP_{1–37}$-HSD/E and HMMM bound $hIAPP_{1–37}$-HSP shows the effect of protonating His18. **b** The difference between HMMM bound $hIAPP_{1–37}$-HSP and $hIAPP_{1–37}$-HSP in solution shows the effect of the membrane on the $hIAPP_{1–37}$-HSP peptide. **c** The difference between HMMM bound $hIAPP_{1–37}$-HSD/E and $hIAPP_{1–37}$-HSD/E in solution shows the effect of the membrane on the $hIAPP_{1–37}$-HSD/E peptide. **d** The difference between the $hIAPP_{1–37}$-HSD/E and $hIAPP_{1–37}$-HSP, both in solution, shows the effect of low pH on the free peptide. Only the final 80 ns of simulation were used for the analysis. The vertical line shows the divide between the 19 N-terminal residues and the 18 C-terminal residues. Adapted with permission from Skeby et al. Biochemistry (2016) 10.1021/acs.biochem.5b00507. Copyright 2016 American Chemical Society

The effect of the membrane varies significantly depending on the protonation state of His18 (Fig. 9b, c). The membrane serves to reduce the flexibility of the $hIAPP_{1–37}$-HSP peptide (Fig. 9b). The entire $hIAPP_{1–37}$-HSP peptide binds to the membrane, and is therefore not as disordered as the peptide in solution. For the $hIAPP_{1–37}$-HSD/E peptides, the effect on the N-terminus and C-terminus are opposing; the N-terminus is stabilized by the membrane, while the C-terminus becomes more dynamic when the peptide is bound to the membrane (Fig. 9c). This is a consequence of the affinity of the N-terminus for the membrane and the lack of affinity of the C-terminus for the membrane. When the N-terminus binds to the membrane, it releases the C-terminus from the interaction with the N-terminus leaving it more flexible.

The membrane bound structure of $hIAPP_{1–37}$ is of high interest, as it may provide clues to the initial events in fibril formation as well as the origin of hIAPP toxicity. A highly popular technique to investigate the membrane bound structure is CD spectroscopy, which can determine the amount of secondary structure elements in a protein or peptide. It is possible to determine the amount of secondary structure in the simulations using STRIDE, which evaluates the type of secondary structure based on the backbone dihedral angles as well as the hydrogen bonding pattern (Table 2) [17]. The amount of helical structure is lower at neutral pH than at low pH, which is consistent with CD experiments of $hIAPP_{1–37}$ in the presence of a 70:30 DOPC:DOPS membrane [18]. The amount of helical structure in $hIAPP_{1–37}$ was determined using CD spectroscopy to be 39–43 % when bound to 100 % DOPG liposomes [19], and between 42 and 50 % when bound to large unilamellar vesicles composed of a 75:25 mixture of DOPC:DOPS [20]. The amount of helical structure in the simulations is lower than the experimentally measured amount, which can have several explanations. First, the helical content of $hIAPP_{1–37}$ in buffer was estimated to be $\sim$10 % [19]; thus, the peptide may have unfolded somewhat before binding to the membrane. Second, in spite of the high flexibility of hIAPP and the fast dynamics of the HMMM, it is quite possible that equilibrium has not yet been reached in the simulations. Finally, the estimation of helical structure from CD experiments is based on the assumption that the signal in the

Table 2 Fraction of secondary structure during the simulations

Peptide	His18	Membrane (lipids)	Helix (%)	β (%)	Coil (%)
$hIAPP_{1–37}$	HSD/E	HMMM (49)	28	2	70
$hIAPP_{1–37}$	HSP	HMMM (49)	45	0	55
$hIAPP_{1–37}$	HSD/E	–	24	0	75
$hIAPP_{1–37}$	HSP	–	29	0	70
$hIAPP_{1–19}$	HSD/E	HMMM (36)	23	0	77
$hIAPP_{1–19}$	HSP	HMMM (36)	19	0	81
$hIAPP_{20–37}$	–	HMMM (36)	6	1	93

The helical content comprises both α-, 3_{10}-, and π-helices; β-structure comprises isolated β-bridges and larger β-sheet structures; and coil comprises coil and turn structure. The first 20 ns of each simulation were excluded from the analysis

far-UV region of wavelengths only originates from the amide bonds, which may not be entirely the case as both aromatic side chains and disulphide bonds can contribute to this region of the spectrum [21].

The amount of β-structure is almost non-existent; however, some is present in the C-terminus of the membrane bound hIAPP$_{1-37}$-HSD/E peptide, which is where it has been suggested that the fibril formation is initiated (Fig. 10b) [22]. Yet, the amount is so small that it is difficult to make conclusions based on these simulations. The high fraction of helical structure is present mostly in the N-terminus of the peptide (Fig. 10a). The region around His18 has a very low degree of helical content, while the C-terminus has some helical structure with a dip around Ser29. The effect of the His18 protonation on the membrane bound hIAPP$_{1-37}$ seems to increase the amount of helical structure, particularly in the C-terminus. This is consistent with the reduced flexibility of the hIAPP$_{1-37}$-HSP peptide as measured by the torsional RMSF. Furthermore, if the HMMM bound hIAPP$_{1-37}$-HSP peptide is compared to the solvated hIAPP$_{1-37}$-HSP peptide, it is evident that the membrane also has a stabilizing effect on the helical structure. Furthermore, the membrane

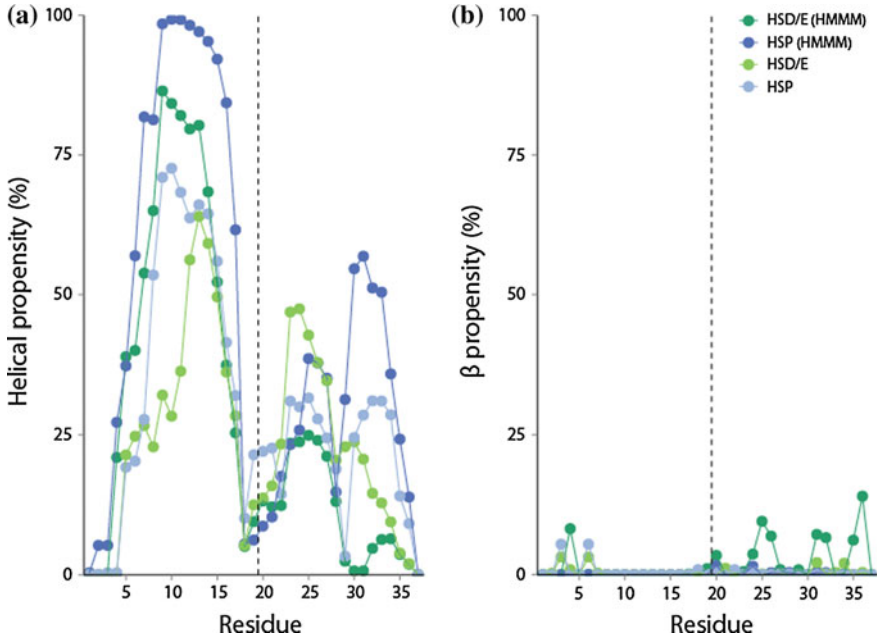


Fig. 10 **a** Helical and **b** β structure propensity of the hIAPP$_{1-37}$ peptide with and without a HMMM membrane in the simulation. A value of 100 % indicates that the residue is in a helical or β-sheet conformation 100 % of the simulation time. The first 20 ns of simulation have not been included in the calculation. The secondary structure was determined using STRIDE.[17] The *vertical line* shows the divide between the 19 N-terminal residues and the 18 C-terminal residues. Adapted with permission from Skeby et al. Biochemistry (2016) 10.1021/acs.biochem.5b00507. Copyright 2016 American Chemical Society

increases the helical structure in the N-terminus of the $hIAPP_{1–37}$-HSD/E peptides, while the helical content in the C-terminus is decreased. The effect of the membrane on the helical structure is consistent with the effect on the torsional RMSF; increased helical structure correlates with decreased flexibility, which makes sense, as a well-defined structure often means less flexibility. In water, the helical content of the $hIAPP_{1–37}$-HSP versus $hIAPP_{1–37}$-HSD/E peptides is very similar, with a slight favor for more helical structure in the $hIAPP_{1–37}$-HSP peptide, which also correlates with the changes in flexibility between these peptides.

From EPR experiments, an amphipathic helix was suggested to be present between residues 9 and 20–22 [14]. This fits reasonably well with the simulations when the differences in conditions are taken into account. The EPR experiments were performed with a higher anionic lipid content (20:80 PC:PS) than the simulations (70:30 PC:PS). Higher anionic lipid content has been shown to induce more helical structure in the peptide [20]. The C-terminus was not observed in the EPR study to exhibit helical structure; however, the estimation of a helix spanning residues 9-22 was based on mobility and accessibility measurements, and as the simulations show, the N-terminus is less flexible than the C-terminus as well as buried deeper in the membrane [14]. It was therefore not possible to determine if helical structure was also present in the C-terminus.

hIAPP-Lipid Interactions

The lipid membrane in these simulations is composed of two types of lipids, 70 % DVPC and 30 % DVPS (Fig. 11). It is well established that the amount of anionic lipid in the membrane has an influence on the fibril formation kinetics of hIAPP. This section describes the detailed interactions between hIAPP and the lipids.

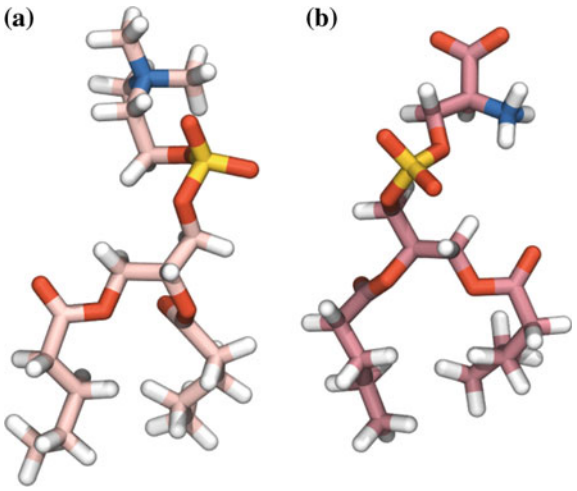


Fig. 11 Structure of **a** DVPC and **b** DVPS. C have been shown in *pink* colors, O is *red*, H is *white*, N is *blue*, and P is *yellow*

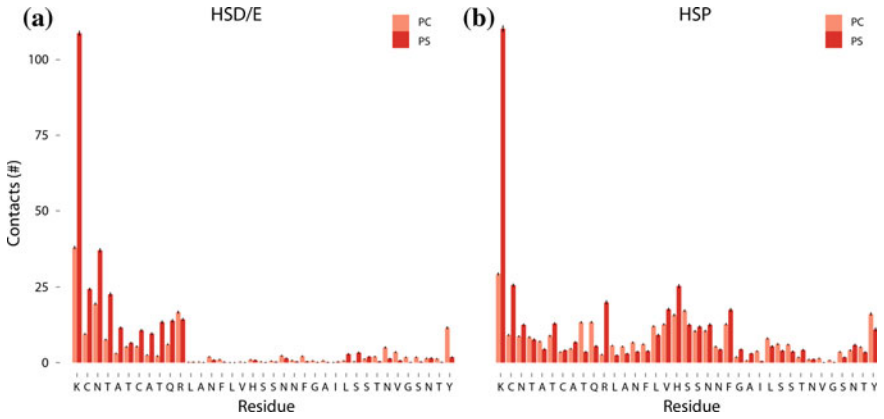


Fig. 12 Average number of heavy atom contacts within 5 Å between the protein and each lipid type for the last 80 ns of simulation. The numbers have been normalized by the number of lipids for each type (70 % PC and 30 % PS), and the error bars represent a 95 % confidence interval. Adapted with permission from Skeby et al. Biochemistry (2016) 10.1021/acs.biochem.5b00507. Copyright 2016 American Chemical Society

The lipid contact profiles for $hIAPP_{1–37}$-HSD/E and $hIAPP_{1–37}$-HSP peptides are quite different (Fig. 12). The $hIAPP_{1–37}$-HSD/E peptides have very little contact with the lipids for residues 12–37. Two hydrophobic stretches are present in this region of the peptide, residues 12–17 and residues 23–26, which do not interact favorably with the hydrophilic head-group region of the membrane. As the $hIAPP_{1–37}$-HSD/E peptides are not inserted between the lipids, the unfavorable interaction between these hydrophobic stretches and the lipids keeps the C-terminus away from the membrane. Residues 27–37, however, contain many hydrophilic residues, which allow some favorable contact between this region and the membrane.

The contact profile for the $hIAPP_{1–37}$-HSP peptides is much more evenly distributed, with peaks around the positively charged residues. This peptide inserts in the interface between the lipid head-groups and the hydrophobic core region, which allows the hydrophobic residues to interact with the hydrophobic tails of the lipids as well as DCLE, while the hydrophilic and charged residues interact with the hydrophilic head-groups (Fig. 13).

The normalized contact profiles from each residue to the two types of lipids are very similar for $hIAPP_{1–37}$-HSP, with a favor for more contacts with the PS lipids for the residues near the positively charged residues (Fig. 12b). The normalized contacts from $hIAPP_{1–37}$-HSD/E are most dominating with the PS lipids, which suggests that the lipids surrounding $hIAPP_{1–37}$-HSD/E may be enriched in PS lipids compared to the overall composition of the membrane. This is in line with the high affinity of hIAPP for anionic membranes, and that the anionic lipids accelerate the rate of fibril formation [4, 20].

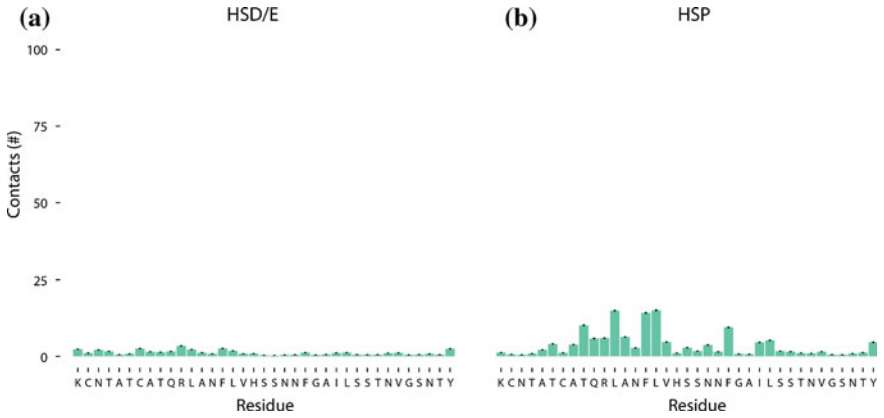


Fig. 13 Average heavy atom contacts within 5 Å between DCLE and the protein residues for the last 80 ns of simulation. The error bars represent a 95 % confidence interval. Adapted with permission from Skeby et al. Biochemistry (2016) 10.1021/acs.biochem.5b00507. Copyright 2016 American Chemical Society

The double-charged Lys1 clearly has the highest number of contacts of all residues to both the phosphatidylcholine (PC) and the PS lipids (Fig. 12). Furthermore, a peak in the number of contacts to Arg11 is present. The $hIAPP_{1–37}$-HSP peptides also have a greater number of contacts to the lipids from His18, which is also expected as this residue is charged. Thus, the positively charged residues have more contacts with the lipids than the rest of the residues, and seem to be anchoring the peptide to the membrane. The higher number of contacts between the lipids and the positively charged residues (Fig. 12) supports the hypothesis that the interaction between hIAPP and the phospholipid membrane is mediated by electrostatic interactions.

The specific interactions between the peptide and the lipids can be evaluated by the number of hydrogen bonds. The length of the hydrogen bond was set to be a maximum of 4 Å to allow for the longer ranging interactions of salt bridges [23] with a maximum 30° deviation from a linear hydrogen bond (Fig. 14). Generally, a higher number of hydrogen bonds are present with the PS lipids. This is likely a combination of two aspects: first, the additional negative charge on the PS lipid gives this lipid a formal charge of −1, while the PC lipid is zwitterionic; and second, PS lipids have two additional hydrogen bond acceptors and three additional hydrogen bond donors compared to PC.

The hydrophilic residues beginning with His18 (residues 18–22) of the $hIAPP_{1–37}$-HSD/E peptides have almost no hydrogen bonds with the lipids. This is somewhat surprising; however, the hydrophobic residues just before and after in the peptide sequence (residues 12–17 and 23–27) may be preventing the interaction of His18 with the lipid head-groups. The number of hydrogen bonds between His18 of the $hIAPP_{1–37}$-HSP peptides is highest to the anionic PS lipid, and the hydrophilic

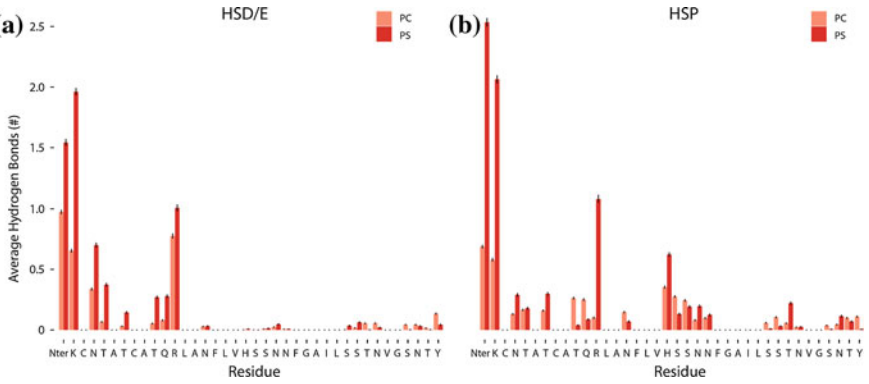


Fig. 14 Average number of hydrogen bonds between the two types of lipid and the peptide side chains for **a** the $hIAPP_{1–37}$-HSD/E peptides and **b** the $hIAPP_{1–37}$-HSP peptides. The first 20 ns of simulation have been excluded from the calculation, and the normalization has taken the differing number of lipids (70 % PC and 30 % PS) into account. The error bars represent a 95 % confidence interval. Adapted with permission from Skeby et al. Biochemistry (2016) 10.1021/acs.biochem.5b00507. Copyright 2016 American Chemical Society

residues just following His18 in the peptide sequence also form hydrogen bonds with both types of lipids.

The high number of hydrogen bonds (which includes salt bridges) between the positively charged groups on the peptide and the anionic PS lipids compared with the PC lipids shows that the affinity for anionic membranes does indeed originate from the favorable electrostatic interactions.

To investigate which groups on the lipid are responsible for the hydrogen bonding with the peptide, the average number of hydrogen bonds between each acceptor group on the lipids (Fig. 15a) and the positively charged groups on the peptide was calculated (Fig. 15b, c). The data was normalized by the number of lipids for each type (70 % PC and 30 % PS), and PS-head was multiplied by 2, as this group only contains two hydrogen bond acceptors, while the other groups contain 4. The number of hydrogen bonds is highest with the lipid groups that are negative and closest to the solvent, with the PS-head group having the highest number. The number of hydrogen bonds from the double-charged Lys1 is, not surprisingly, quite high from both the terminal charge and the side chain amino group. The number of hydrogen bonds between Arg11 of the $hIAPP_{1–37}$-HSD/E peptides and the PS-head group is also quite high, while the number is quite low from the same residue to the PS-P and PC-P groups. Contrary to $hIAPP_{1–37}$-HSD/E, Arg11 and His18 of the $hIAPP_{1–37}$-HSP peptides have hydrogen bonds to both the PS-head and PS-P groups. This difference in the hydrogen bonding pattern between the $hIAPP_{1–37}$-HSD/E and $hIAPP_{1–37}$-HSP peptides is a consequence of the level of insertion of the two peptides. As $hIAPP_{1–37}$-HSP inserts deeper between the lipids, Arg11 of the $hIAPP_{1–37}$-HSP peptides is able to interact with the phosphate groups, while Arg11 of the $hIAPP_{1–37}$-HSD/E peptides is not able to do so to the same extent.

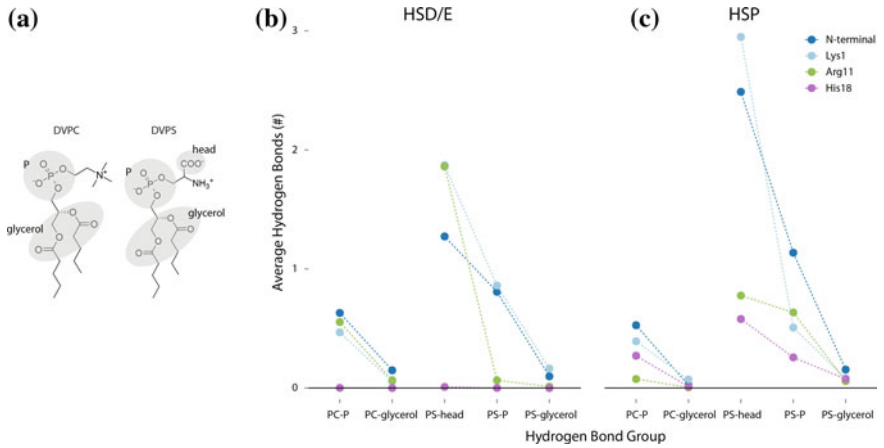


Fig. 15 **a** Hydrogen bond acceptor groups on the two types of lipids. The oxygen atoms in each of the *gray* areas have been used to determine the hydrogen bonds to the positive groups on the peptide. This shows the average number of hydrogen bonds for the **b** $hIAPP_{1–37}$-HSD/E peptides and the **c** $hIAPP_{1–37}$-HSP peptides, between the charged groups in the protein and the lipid acceptor groups. The average is based on the final 80 ns of simulation. The data was normalized by the number of lipids for each type (70 % PC and 30 % PS), and PS-head was multiplied by 2, as this group only contains two hydrogen bond acceptors, while the other groups contain 4

Does the HMMM Influence the Results?

It is imperative to make sure that the use of the HMMM is not affecting the results of the simulations. We have therefore converted the HMMM in four of the systems to a DOPC/DOPS membrane by growing the full lipid tails on the existing short lipids, and extended the simulations by 20 ns. One $hIAPP_{1–37}$-HSD and one $hIAPP_{1–37}$-HSE simulation was chosen. Since all of the simulations are different because of the very flexible peptide, and none of the simulations really stand out, the choice of simulation to extend is arbitrary. Two $hIAPP_{1–37}$-HSP simulations were also chosen to be extended; one that was representative of the $hIAPP_{1–37}$-HSP simulations, $hIAPP_{1–37}$-HSP(A1), and the outlier, $hIAPP_{1–37}$-HSP(B2), where the peptide has not inserted between the lipids.

The peptides all stay bound to the membrane. The minimum z-coordinate of the peptide fluctuates around the level of the phosphates as it also does in the HMMM simulations (Fig. 16a). The structure of the $hIAPP_{1–37}$-HSD/E peptides changes from the final structure in the HMMM simulation (Fig. 16b). However, this is not surprising, as we have seen that this is a very flexible peptide (Fig. 6b, d).

From the alignment of the peptide structures (Fig. 17) it is evident that the $hIAPP_{1–37}$-HSP(A1) peptide, which is embedded in the membrane (Fig. 5b), has a very stable structure, which is also evident from the RMSD. As it is not pushed out of the membrane, the inserted structure of the $hIAPP_{1–37}$-HSP peptide is stable, at least on the timescale of these simulations. The $hIAPP_{1–37}$-HSP(B2) peptide does

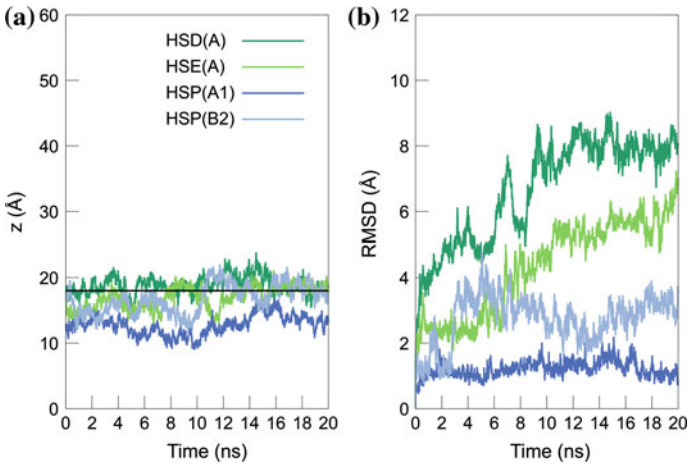


Fig. 16 **a** Minimum z-coordinate and **b** RMSD for the DOPC/DOPS simulations. The *black horizontal line* in (**a**) indicates the position of the lipid phosphates. Adapted with permission from Skeby et al. Biochemistry (2016) 10.1021/acs.biochem.5b00507. Copyright 2016 American Chemical Society

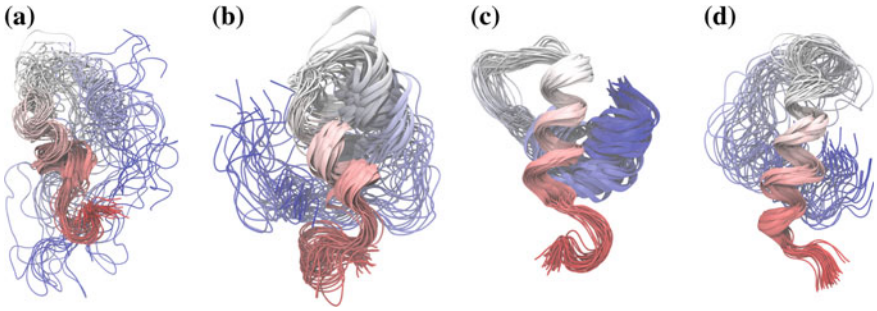


Fig. 17 Alignment of the peptides using C_α atoms of residues 1–19 from the DOPC/DOPS simulations that were extended from the **a** $hIAPP_{1–37}$-HSD(A), **b** $hIAPP_{1–37}$-HSE(A), **c** $hIAPP_{1–37}$-HSP(A1), and **d** $hIAPP_{1–37}$-HSP(B2) HMMM simulations. The peptides are colored by the residue number with the N-terminus colored *red* and the C-terminus colored *blue*. Adapted with permission from Skeby et al. Biochemistry (2016) 10.1021/acs.biochem.5b00507. Copyright 2016 American Chemical Society

not insert in the DOPC/DOPS membrane; however, even though it is only attached to the membrane via the N-terminus and the Lys1 side chain, it does not leave the membrane.

The structures of the $hIAPP_{1–37}$-HSD(A) and $hIAPP_{1–37}$-HSE(A) peptides are more flexible than the $hIAPP_{1–37}$-HSP peptide structures. The flexibility is mainly in the C-terminal part of the peptides, while the N-terminal part binds to the membrane in the same manner throughout the DOPC/DOPS simulation. The flexible behavior of the C-terminus is the same in both membrane models.

The influence of the HMMM on the binding and structure of hIAPP compared to a conventional DOPC/DOPS membrane is minimal based on the limited tests performed here. Although, the greatest confidence in the HMMM simulations comes from the high agreement of the simulations with the available experimental data on the membrane bound conformation of hIAPP. Furthermore, previous demonstration of the ability of the HMMM to self-assemble, to reproduce the density profile of a conventional membrane, as well as to reproduce the insertion of the membrane anchor domain of human coagulation factor VII, also provides a high degree of confidence in the simulations [1].

Discussion

We have investigated the binding of hIAPP to a membrane bilayer. Consistent with surface pressure experiments, all of the peptides in these simulations containing the 19 N-terminal residues bind to the membrane within the first 20 ns of simulation [13]. $hIAPP_{20–37}$ does not bind to the membrane in these simulations, which is consistent with $hIAPP_{20–29}$ not binding in the surface pressure experiments [13]. In accordance with EPR experiments, we find that the most persistent structure is present in the N-terminus while the C-terminus is very disordered [14]. The orientation of the peptide in the simulations is also consistent with EPR experiments if the differing membrane conditions are taken into account [14]. Furthermore, the degree of helicity in the structures in these simulations matches that measured in CD spectroscopy experiments reasonably well [18–20].

The high level of agreement with experiments of these simulations enables the discussion of possible implications of the results presented here regarding the mechanism of binding; the membrane bound structure, as well as the effects of pH and the membrane on the peptide.

Influence of the Membrane on hIAPP

The effect of the membrane on hIAPP depends on the protonation state of His18. When $hIAPP_{1–37}$-HSD/E binds to the membrane, the N-terminal part of the peptide is stabilized and anchored to the membrane, while the C-terminal part is disordered and very flexible. It has already been suggested previously that the acceleration of hIAPP fibril formation by membranes is caused by an increased local concentration of hIAPP [14], however, the present simulations suggest an additional accelerating factor; the unwinding of the C-terminal region. It has been shown that residues 20–29 are important for fibril formation, and that fibril formation is initiated in this region [22, 24]. Unwinding of the C-terminal region would expose the backbone and prime it for interaction with another peptide. Assuming that the unwinding of the peptide is a rate-limiting step, this would ultimately lead to an acceleration of

the fibril formation as the unwinding step in the fibril formation mechanism has been eliminated.

The effect of the membrane on $hIAPP_{1–37}$-HSP is different from the effect on $hIAPP_{1–37}$-HSD/E. When $hIAPP_{1–37}$-HSP binds to the membrane, not only does the N-terminus bind to the membrane, but the C-terminus does as well. This could provide a clue to the inhibiting effect on the fibril formation of high anionic lipid content in the membrane. As discussed above, a higher content of anionic lipids could induce the protonation of His18, which would change the structure and membrane binding of hIAPP from an exposed and dynamic C-terminus to a deeper buried and rigid structure. This hypothesis could potentially be tested using constant pH simulations [25].

As discussed previously, the acceleration of fibril formation by an anionic membrane is most likely due partly to the increased local concentration of hIAPP. However, when the concentration of anionic lipid becomes too high, the fibril formation is actually inhibited [20]. From these simulations we have seen a higher affinity of hIAPP for the anionic PS lipids over the PC lipids, which could hint at an explanation for the inhibition of fibril formation at higher anionic lipid concentration; it is possible that when the concentration of anionic lipids becomes higher, the energetic cost of releasing the lipid from the membrane becomes higher, resulting in slower aggregation. Furthermore, the possible protonation of His18 may also play a role in the inhibition of fibril formation. This is a hypothesis that needs to be tested with further experiments and/or simulations.

Influence of the Positive His18

hIAPP aggregates slower at low pH [18, 26]. The most likely reason is probably the increased electrostatic repulsion between the peptides with the added positive charge on His18 at low pH. Nothing in the present simulations indicate that this should not be the case, as they do not show any drastic differences in either structure or stability between the two protonation states of the peptides. It cannot be supported based on these simulations either, as this would require simulations with multiple peptides in the setup.

When the peptide is bound to the membrane, however, we do see a difference between the two peptide protonation states. When $hIAPP_{1–37}$-HSD/E is bound to the membrane, the N-terminus is interacting with the lipids, while the C-terminus is solvated and not interacting with the membrane. The $hIAPP_{1–37}$-HSP peptide, on the other hand, binds to the membrane with both the N-terminal and the C-terminal parts of the peptide. So, at lower pH, the C-terminus is not free to interact with another peptide, should it encounter any. The electrostatic repulsion probably also plays a role in the membrane bound situation, however, the effect is likely smaller, as the anionic lipids serve to screen the charges on the peptide. We therefore suggest that the inhibiting effect of low pH on the fibril formation could be caused by two different aspects in the presence and absence of the membrane; the

electrostatic repulsion is the main actor in solution, while the interaction of the membrane with the C-terminus of $hIAPP_{1–37}$-HSP is the main actor when the peptide is bound to an anionic membrane.

Specific Lipid Interactions

The analysis of the specific interactions between the residues and the two lipid types has provided possible explanations for the behavior of hIAPP. It is clear that the early attraction between hIAPP and the membrane is an electrostatic interaction that is driven by the affinity of the positively charged residues of hIAPP for the anionic PS lipids, as has also been shown by Infrared Reflection Absorption Spectroscopy [27]. hIAPP binds to the membrane differently depending on the protonation state of His18. When His18 is uncharged, the membrane binding is mediated by residues 1–11, while residues 12–37 have very few contacts with the membrane. When His18 is charged, the C-terminal part is drawn to the membrane, and the entire peptide is buried deeper within the membrane head-group region. The reason for the differences in membrane binding between the two peptide protonation states is the presence of hydrophobic stretches of residues surrounding His18. When His18 is positively charged, the interaction between His18 and the PS lipids is strong enough to overcome the barrier of pushing the peptide between the lipid head-groups to allow the favorable interaction of the hydrophobic residues with the membrane core region. However, when His18 is not positively charged, there is no attracting interaction to pull the C-terminus to the membrane, and the hydrophobic residues are pushed away from the membrane.

Conclusion

The binding of hIAPP and fragments thereof to a mixed zwitterionic/anionic membrane has been investigated using all-atom MD simulations. High agreement of the simulations with experiments has enabled the detailed discussion of mechanisms regarding the influence of pH and membrane on the membrane binding and fibril formation of hIAPP. However, the hypotheses posed in this work need further testing through both simulations and experiments. We have initiated further simulations building on the present work to investigate the association of multiple hIAPP peptides bound to the membrane. Furthermore, it is our hope that these simulations and the hypotheses posed here could be the starting point for future studies concerning the mechanism of hIAPP fibril formation and membrane disruption.

References

1. Ohkubo YZ, Pogorelov TV, Arcario MJ, Christensen GA, Tajkhorshid E (2012) Accelerating membrane insertion of peripheral proteins with a novel membrane mimetic model. Biophys J 102:2130–2139
2. Nanga RPR, Brender JR, Vivekanandan S, Ramamoorthy A (2011) Structure and membrane orientation of IAPP in its natively amidated form at physiological pH in a membrane environment. Biochim Biophys Acta Biomembr 1808:2337–2342
3. Rustenbeck I, Matthies A, Lenzen S (1994) Lipid-composition of glucose-stimulated pancreatic-islets and insulin-secreting tumor-cells. Lipids 29:685–692
4. Knight JD, Miranker AD (2004) Phospholipid catalysis of diabetic amyloid assembly.J Mol Biol 341:1175–1187
5. Petrache HI, Tristram-Nagle S, Gawrisch K, Harries D, Parsegian VA, Nagle JF (2004) Structure and fluctuations of charged phosphatidylserine bilayers in the absence of salt. Biophys J 86:1574–1586
6. Neria E, Fischer S, Karplus M (1996) Simulation of activation free energies in molecular systems. J Chem Phys 105:1902–1921
7. Piana S, Lindorff-Larsen K, Shaw DE (2011) How robust are protein folding simulations with respect to force field parameterization? Biophys J 100:L47–L49
8. Best RB, Zhu X, Shim J, Lopes PEM, Mittal J, Feig M, MacKerell AD Jr (2012) Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain χ_1 and χ_2 dihedral angles. J Chem Theory Comput 8:3257–3273
9. Baylon JL, Lenov IL, Sligar SG, Tajkhorshid E (2013) Characterizing the membrane-bound state of cytochrome P450 3A4: structure, depth of insertion, and orientation. J Am Chem Soc 135:8542–8551
10. Ewald PP (1921) Die Berechnung Optischer Und Elektrostatischer Gitterpotentiale. Ann Phys 64:253–287
11. Darden T, York D, Pedersen L (1993) Particle mesh ewald: an N·log(N) method for ewald sums in large systems. J Chem Phys 98:10089–10092
12. York DM, Wlodawer A, Pedersen LG, Darden TA (1994) Atomic-level accuracy in simulations of large protein crystals. Proc Natl Acad Sci USA 91:8715–8718
13. Engel MFM, Yigittop H, Elgersma RC, Rijkers DTS, Liskamp RMJ, de Kruijff B, Höppener JWM, Killian JA (2006) Islet amyloid polypeptide inserts into phospholipid monolayers as monomer. J Mol Biol 356:783–789
14. Apostolidou M, Jayasinghe SA, Langen R (2008) Structure of α-helical membrane-bound human islet amyloid polypeptide and its implications for membrane-mediated misfolding. J Biol Chem 283:17205–17210
15. Williamson JA, Loria JP, Miranker AD (2009) Helix stabilization precedes aqueous and bilayer-catalyzed fiber formation in islet amyloid polypeptide. J Mol Biol 393:383–396
16. Bishop CM (2006) Pattern recognition and machine learning: Periodic Variables, 1st edn. Springer Science+Business Media, LLC
17. Frishman D, Argos P (1995) Knowledge-based protein secondary structure assignment. Proteins Struct Funct Genet 23:566–579
18. Khemtémourian L, Doménech E, Doux JPF, Koorengevel MC, Killian JA (2011) Low pH acts as inhibitor of membrane damage induced by human islet amyloid polypeptide. J Am Chem Soc 133:15598–15604
19. Knight JD, Hebda JA, Miranker AD (2006) Conserved and cooperative assembly of membrane-bound α-Helical states of islet amyloid polypeptide. Biochemistry 45:9496–9508
20. Jayasinghe S, Langen R (2005) Lipid membranes modulate the structure of islet amyloid polypeptide. Biochemistry 44:12113–12119
21. Kelly SM, Price NC (1997) The application of circular dichroism to studies of protein folding and unfolding. Biochim Biophys Acta Protein Struct Mol Enzymol 1338:161–185

22. Shim S, Gupta R, Ling YL, Strasfeld DB, Raleigh DP, Zanni MT (2009) Two-dimensional IR spectroscopy and isotope labeling defines the pathway of amyloid formation with residue-specific resolution. Proc Natl Acad Sci USA 106:6614–6619
23. Barlow DJ, Thornton JM (1983) Ion pairs in proteins. J Mol Biol 168:867–885
24. Westermark P, Engström U, Johnson KH, Westermark GT, Betsholtz C (1990) Islet amyloid polypeptide: pinpointing amino acid residues linked to amyloid fibril formation. Proc Natl Acad Sci USA 87:5036–5040
25. Mongan J, Case DA, McCammon JA (2004) Constant pH molecular dynamics in generalized born implicit solvent. J Comput Chem 25:2038–2048
26. Chargé SBP, Dekoning EJP, Clark A (1995) Effect of pH and insulin on fibrillogenesis of islet amyloid polypeptide in vitro. Biochemistry 34:14588–14593
27. Lopes DHJ, Meister A, Gohlke A, Hauser A, Blume A, Winter R (2007) Mechanism of islet amyloid polypeptide fibrillation at lipid interfaces studied by infrared reflection absorption spectroscopy. Biophys J 93:3132–3141

Effect of Terminal Capping on Aggregation of Peptide Fragments

This section summarizes and discusses the results from the submitted manuscript entitled "The importance of being capped: Terminal capping of an amyloidogenic peptide affects fibril formation propensity and fibril morphology". The biophysical studies were performed by Maria Andreasen; the synthesis of the peptides was conducted by Erik Holm Nielsen and Heidi Frahm; the microscopy images were produced by Lasse Hyldgaard Klausen; and I performed the MD simulations. The manuscript was prepared mainly by Maria Andreasen, however, a substantial part of the results and discussion was written by me.

Introduction

Studies of amyloid aggregation are often performed using fragments of the full-length amyloidogenic proteins. This simplification of protein structure reduces the complexity of the data, and these modifications are justified by the fact that usually only part of the protein is incorporated into the fibril core. Fragments of the full-length protein have been extensively used to study aggregation of Aβ [1–3], the prion protein [4–6], hIAPP [7–10], huntingtin [11, 12], and transthyretin [13–15]. However, even subtle modifications of the peptide structure can result in changes in fibril formation properties as has been shown for the $hIAPP_{20–29}$ peptide fragment [8, 16]. The effect of changing the terminals of a short peptide has been investigated for the Aβ [16–22] fragment, where the free or capped terminals led to two different fibril morphologies [17]. hIAPP contains a C-terminal amidation in vivo [18], which has been shown to be important for both aggregation propensity and monomer structure; removing the C-terminal amidation results in slower aggregation and a different monomer structure [19–21]. These observations illustrate the importance of determining the effect of terminal modifications on the amyloidogenic properties.

K.K. Skeby, *Computational Modelling of the Human Islet Amyloid Polypeptide*,
Springer Theses, DOI 10.1007/978-3-319-20040-8_5

In this study, we have investigated the effect of different N- and C-terminal modifications on the fibril formation kinetics and fibril morphology using a combination of biophysical and computational techniques.

Experimental and Computational Methods

We have examined the fibril formation properties of five different capping variants of the $hIAPP_{20–29}$ peptide (SN) with the sequence SNNFGAILSS (Table 1).

The full details of the peptide synthesis, ThT fibril formation experiments, Fourier transform infrared (FTIR) spectroscopy, far-UV CD spectroscopy, TEM, and AFM can be found in the published paper.

A 50 ns MD simulation was performed for each of three SN capping variants; SN, SN–NH_2, and Ac–SN. We only included these three capping variants, since they show the most dramatic differences in chemical structure, fibril formation kinetics, and fibril morphology (see experimental results). In our simulations, a fibril consists of two antiparallel β-sheets with 10 peptide strands in each. The fibril structure was obtained from a ss-NMR structure of the SN–NH_2 peptide [22]. This structure is a double-layer of two β-sheets with antiparallel arrangement of the peptide strands, which is probably caused by the N-terminal positive charge. A fibril structure with parallel β-sheets would place the positive charges solely on one side of the β-sheet, which is expected not to be energetically favorable. Parallel β-sheet structures of fragments from hIAPP have been observed using X-ray micro crystallography [23, 24]; however, in a microcrystal, the charged peptide termini pack closely against the termini in the adjacent β-sheet in the crystal lattice, which provides a balance of the charges. This enables the parallel β-sheet conformation to be stabilized. To our knowledge, no atomic resolution amyloid fibril structure composed of peptide fragments in a parallel β-sheet conformation not originating from a microcrystal has been published. Therefore, the assumption that all capping variants used in the present simulations form fibrils with antiparallel β-sheets is

Table 1 Overview of the capping variants of the decapeptide SNNFGAILSS used in the present study

Name	Modification	Sequence
SN	Free N-and C-termini	$^{+}H_3N$–SNNFGAILSS–COO^{-}
Ac–SN	N-terminal acetylation, free C-terminus	H_3CNH–SNNFGAILSS–COO^{-}
SN–NH_2	Free N-terminus, C-terminal amidation	$^{+}H_3N$–SNNFGAILSS–$CONH_2$
Ac–SN–NH_2	N-terminal acetylation, C-terminal amidation	H_3CNH–SNNFGAILSS–$CONH_2$
Ac–SN–NMe	N-terminal acetylation, C-terminal N-methylation	H_3CNH–SNNFGAILSS–CONMe

The terminal capping groups are shown, and the sequence is listed for the fibrillating peptides

reasonable. This also allows a more direct comparison of the effect of the capping groups on the fibril structure.

The ss-NMR fibril structure only contains atomic coordinates for the central NFGAILS residues since only the hexapeptide NFGAIL was isotopically labeled. The N-terminal Ser-Asn and C-terminal Ser were constructed using Maestro 9.2 from the 2011 Schrödinger Suite [25], by continuation of the antiparallel β-sheet conformation of the backbone. The peptide was first extended, after which N-terminal acetyl- and C-terminal methyl amide capping groups were added. This structure was used as a template for all the capping variants to ensure the same starting structure. Each capping variant was created with the psfgen tool for VMD [26].

All systems were solvated using the TIPS3P water model [27] in an 86 Å cubic box, allowing at least 13 Å from any fibril atom to the edge of the box. The system was neutralized to a 50 mM concentration of NaCl to match the ionic concentration of the experiments. All systems were simulated with the CHARMM22* FF [28] in NAMD 2.9 [29]. This FF was chosen since it has optimized backbone torsional parameters, allowing a good balance between α-helix, β-sheet, and random coil. The simulation temperature was kept constant at 310 K using a Langevin thermostat with a damping constant of 0.5 fs^{-1}. The simulation pressure was kept constant at 1 atm using a Langevin piston with a piston period of 100 fs, a piston decay time of 50 fs, and a piston temperature of 310 K. The simulations were run with a 2 fs time-step. Non-bonded interactions were calculated every 2 fs including atoms within 12 Å, and the full electrostatic interactions were calculated every 4 fs. The non-bonded interactions were cut off at 12 Å with a switching function starting at 10 Å. The pair-list was updated every 40 fs and includes atom pairs within 13.5 Å. The systems were simulated using periodic boundary conditions, and PME [30, 31] was used for calculating the long-range electrostatic interactions. Bonds to hydrogen were constrained using the RATTLE algorithm [32].

All systems were initially minimized for 10,000 steps, after which the water and ions were equilibrated for 10 ps in the NVT ensemble, keeping the fibril constrained. Then, a 1 ns equilibration was performed in the NPT ensemble, and finally production runs of 50 ns were performed for each system. Snapshots were stored every 10 ps and were used for analysis.

Results

TEM and AFM images show that the terminal capping affects the fibril morphology. The SN–NH_2 peptide forms long twisted fibrils without branching (Fig. 1), consistent with previous reports on the fibril morphology of this peptide [8, 22]. Ac–SN does not show fibril activity on the microscopy images. The SN, Ac–SN–NH_2, and Ac–SN–NMe peptides form long, straight fibrils with a tendency towards lateral bundling.

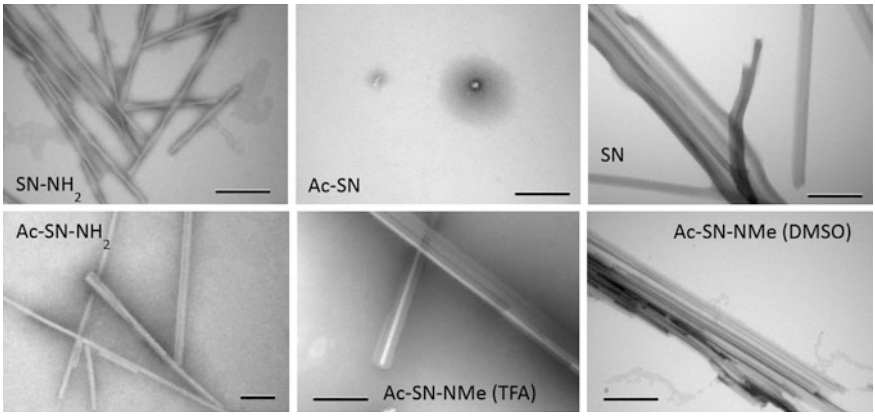


Fig. 1 TEM images of fibrils of terminal capping variants of SN. *Scale bar* equals 200 nm. AFM images of fibrils of terminal capping variants of SN can be found in the published paper. Reprinted with permission from Andreasen et al. [8] Biochemistry, 53, 6968. Copyright 2014 American Chemical Society

The fibril formation of the capping variants of SN was examined using the amyloid binding fluorescent dye ThT [33]. The terminal capping has a profound effect on the fibril formation kinetics of SN (Fig. 2). The lag-time is defined as the time-point when the ThT fluorescence signal has reached 10 % of the maximum signal intensity. The lag-time is a measure of the fibril formation propensity, and reflects the time it takes for the peptide to form the critical nucleus needed to initiate the elongation of the fibril. Based on the lag-times of the fibril formation (Table 2), the capping variants of the SN peptide can be ranked by increasing lag-times;

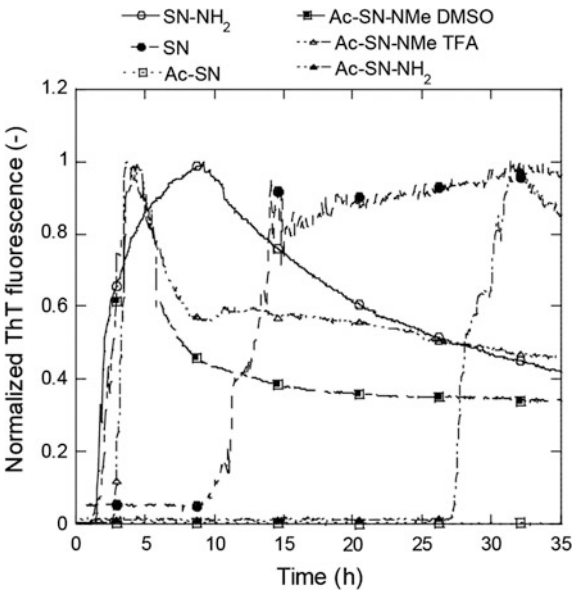


Fig. 2 Fibril formation of terminal capping variants of SN. Normalized ThT fluorescence from triplicate runs plotted against the time of incubation. Reprinted with permission from Andreasen et al. [8] Biochemistry, 53, 6968. Copyright 2014 American Chemical Society

Table 2 Lag-times of the fibril formation kinetics of the terminal capping variants of SN

SN capping variant	Lagtime (h)	Max ThT (A.U.)
SN–NH_2	1.6 ± 0.1	597.3 ± 8.7
Ac–SN–NMe (DMSO stock)	1.8 ± 0.4	1410.5 ± 78.5
SN	11.3 ± 1.3	144.0 ± 28.7
Ac–SN–NH_2	28.3 ± 1.6	129.0 ± 2.6
Ac–SN	N.A.	9.7 ± 0.6

The lag-time is defined as the time-point when the ThT fluorescence signal has reached 10 % of the maximum signal intensity

SN–NH_2 ≈ Ac–SN–NMe < SN < Ac–SN–NH_2. Ac–SN is not listed as it does not display an increase in the ThT fluorescence even after prolonged incubation for up to 80 h. Due to solubility reasons Ac–SN–NMe was synthesized as a depsipeptide intermediate, which contains an ester bond between the carboxylic acid of the backbone and the hydroxyl-group of the adjoining amino acid side chain. The depsipeptide intermediate rearranges into the normal peptide above pH 7. Besides dissolving the depsipeptide in DMSO, which was used for the other peptides, Ac–SN–NMe was also dissolved in TFA, which had been used previously for depsipeptides, to make sure that DMSO did not have an unexpected effect on the rearrangement from the ester to the peptide. Ac–SN–NMe in the two different solvents did not display markedly different fibril formation kinetics.

Besides differences in the lag-time, the capping variants also display differences in the maximum ThT intensity (ThT_{max}) observed during fibril formation (Table 2). The ThT_{max} is influenced by e.g. fibril morphology, the tendency for lateral bundling, and the amount of fibril material in the sample. SN–NH_2 and Ac–SN–NMe in the two stock solutions all display high ThT_{max}, while SN and Ac–SN–NH_2 display very modest ThT_{max}. The decrease in ThT_{max} correlates with the increase in lag-time, except for SN–NH_2, which could be connected to the different morphology of SN–NH_2. The ThT intensity of Ac–SN is indistinguishable from that of ThT in buffer.

We examined the secondary structure of the fibrils using FTIR and CD spectroscopy (Fig. 3). Protein secondary structure elements, such as α-helix, β-sheet, and random coil, give rise to different CD spectra, and deconvolution of a CD spectrum enables the determination of the secondary structure content of a sample. FTIR probes the bond vibrations in a sample, and peaks at specific locations can indicate the presence of e.g. amyloid or regular β-sheet structure. All fibrillating peptides display amyloid structure in the FTIR and CD spectra. Furthermore, the FTIR and CD spectra provide no evidence for β-sheet structure of Ac-SN, and the deconvolution of the FTIR spectrum shows only a random coil contribution to the spectrum (Table 3). Clearly, N-terminal acetylation of SN (with a free C-terminus) abolishes the amyloidogenicity observed for non-capped SN.

All the fibrillating capping variants of SN display amyloid β-sheet components as the major (55–70 %) contributor to the FTIR spectra and comparable amounts of β-turn (20–25 %) (Table 3). SN–NH_2 has a slightly higher level of regular β-sheet

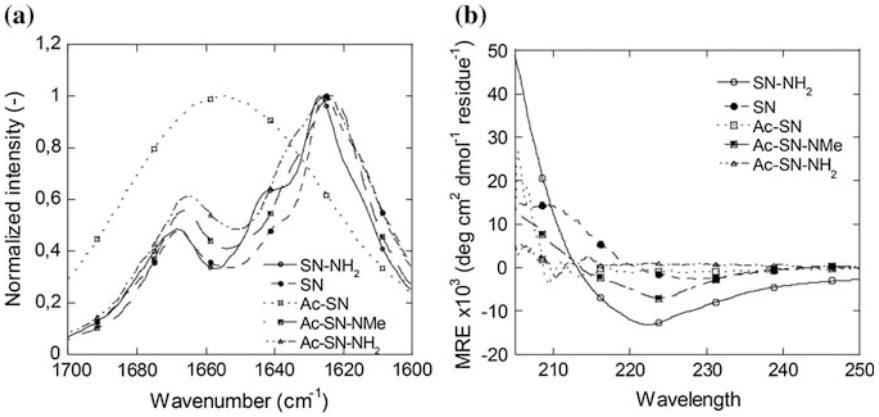


Fig. 3 Secondary structure analysis of the fibrils formed by the terminal capping variants of SN. **a** FTIR analysis in the amide I band of the terminal capping variants of SN (not including the double capped methylated SN from the TFA stock due to the major signal from TFA in the amide I band). **b** Far-UV CD analysis of the terminal capping variants of SN. Reprinted with permission from Andreasen et al. [8] Biochemistry, 53, 6968. Copyright 2014 American Chemical Society

Table 3 Percentages of secondary structural elements of fibrils obtained from deconvolution of FTIR spectra of fibrils

SN variant	Amyloid β-sheet	β-sheet	β-turn	Random	Other
SN	67.8	7.7	24.5	–	–
Ac–SN	–	–	–	100.0	–
SN–NH_2	60.1	19.3	20.6	–	–
Ac–SN–NMe	56.0	9.7	25.7	5.0	3.5
Ac–SN–NH_2	61.7	11.6	26.7	–	–

(and slightly lower levels of β-turn content) than the other fibrillating peptides, while SN has a higher amyloid β-sheet content than the other fibrillating peptides.

The far-UV CD spectra of SN capping variant fibrils seen in Fig. 3b are less informative with regards to the secondary structural elements. The CD spectrum of Ac–SN displays very little signal in the far-UV range consistent with the lack of amyloid structure observed for this capping variant. Despite a positive ThT fibril formation curve and amyloid β-sheet content in the FTIR spectra, Ac–SN–NH_2 does not display a signal in the far-UV range of the CD spectrum, even though a clear fibril pellet is visible after centrifugation. The remaining capping variants (SN–NH_2, Ac–SN–NMe, and SN) all display CD spectra indicative of β-sheet structure with a single minimum at ~218 nm and a positive signal below 210 nm.

We used cross-seeding to examine whether fibril seeds could change the normal morphology formed by a capping variant. SN monomers can form polymorphic fibrils. However, seeding with homogeneous fibril seeds abolishes the polymorphism and leads to a single fibril morphology [34]. Thus polymorphic tendencies can be overruled by specific fibril structures.

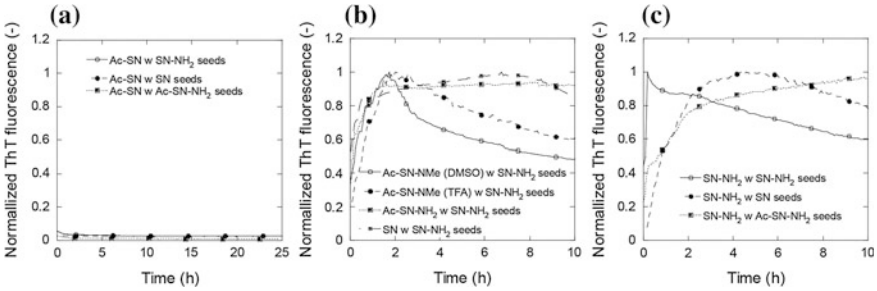


Fig. 4 Kinetics of cross-seeding of the capping variants. **a** Normalized ThT fluorescence for the cross seeding of the non-fibrillogenic Ac–SN seeded with the various capping variants. **b** Normalized ThT fluorescence for the capping variants forming flat fibrils seeded with SN–NH_2 that forms twisted fibrils. **c** Normalized ThT fluorescence for the SN–NH_2 forming twisted fibrils seeded with the various capping variants normally forming flat fibrils. Reprinted with permission from Andreasen et al. [8] Biochemistry, 53, 6968. Copyright 2014 American Chemical Society

Seeding was not able to stimulate fibril formation of the Ac–SN fibril. This is seen both from the lack of increased ThT fluorescence even after prolonged incubation (Fig. 4a); the lack of fibrils observed in TEM and AFM images (Fig. 5) and the lack of amyloid β-sheet signal in the FTIR spectrum (See published paper). The presence of short fibril fragments in the TEM image obtained for Ac–SN seeded with preformed fibrils of SN–NH_2 are most likely the seeds and not structures formed by Ac–SN (Fig. 5).

The fibril kinetics for all other capping variants of SN is affected by the presence of preformed fibril seeds (Fig. 4b–c). Addition of fibril seeds abolishes the lag-time of fibril formation. The fibril morphology did not change in any of the cross-seeding experiments, despite the presence of preformed fibril seeds of the other morphology (Fig. 5).

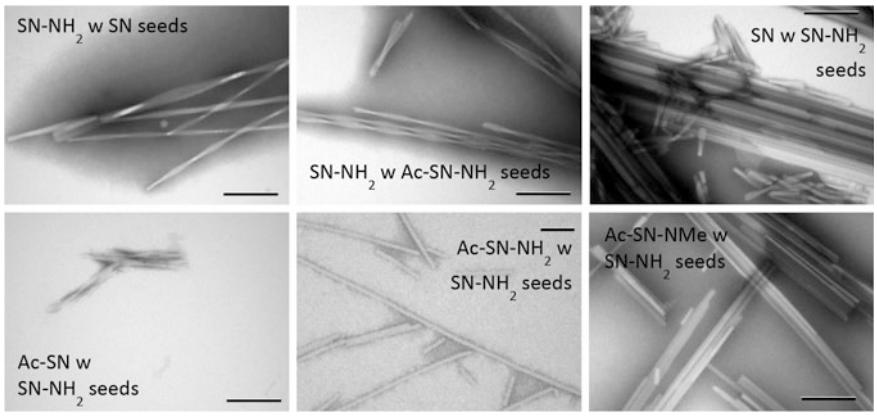


Fig. 5 TEM images of the fibrils formed by cross-seeding the terminal capping variants of SN. *Scale bar* represents 200 nm. AFM analysis of the fibril morphology formed during cross-seeding can be found the published manuscript. Reprinted with permission from Andreasen et al. [8] Biochemistry, 53, 6968. Copyright 2014 American Chemical Society

The preservation of the morphology of the parent peptide in the presence of seeds of the opposite morphology contrasts with previous reports for SN. Seeding has previously led to a homogeneous sample of the same morphology as the initial fibril seed [34]. Since the lag phase disappears, the seeds have interacted with the peptide in solution, and have initiated the fibril formation. However, the intrinsically coded fibril morphology of the peptide in solution overrules the morphology of the seeds.

For most of the fibrils arising from the cross-seeding, the secondary structure, as determined by FTIR does not change compared to the non-seeded fibrils, which is in line with the observations made from the TEM and AFM images. However, the FTIR spectra of SN–NH_2 seeded with Ac–SN–NH_2 and SN seeded with SN–NH_2 differ from the FTIR spectra of the unseeded fibrils, even though no change in the fibril morphology is seen in either TEM or AFM (Fig. 5a). This could indicate a slightly different packing of the individual peptides when seeds from fibrils with a different morphology are present, even though it does not lead to a change in fibril morphology.

To obtain clues to the underlying reasons for the differences in fibril formation properties between these capping variants, we performed MD simulations of three of the capping variants (SN, Ac–SN, and SN–NH_2). We simulated a fibril consisting of two antiparallel β-sheets with 10 strands in each (Fig. 6).

The C_α RMSD relative to the minimized structure of the fibrils quantifies how much the fibril structure changes from the original conformation during the simulation (Fig. 7a). The SN fibril is the most stable with an RMSD around 3 Å, which is reasonable as this structure is a small part taken from a larger fibril assembly, making the peptide strands at the fibril termini quite flexible. However, the RMSD curves for the Ac–SN and SN–NH_2 did not reach a plateau, suggesting that these peptides may not have equilibrated completely. Despite this, the present simulations

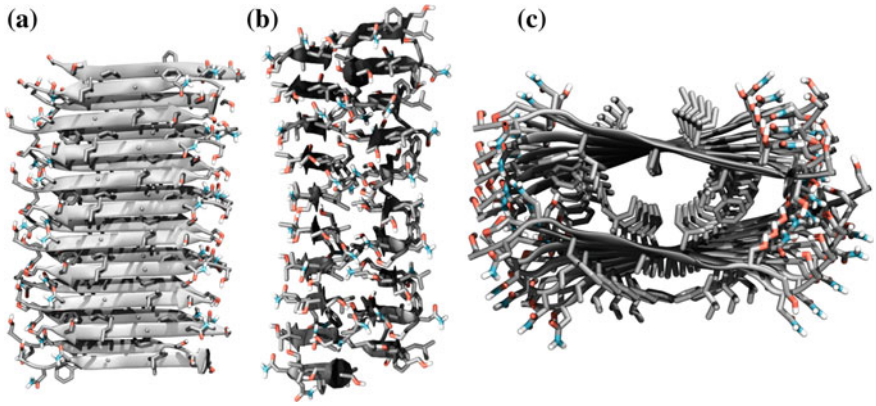


Fig. 6 Initial structure of the fibril used in the simulations. The fibril is viewed from different orientations; **a** top view, **b** side view, and **c** view from the fibril end. Reprinted with permission from Andreasen et al. [8] Biochemistry, 53, 6968. Copyright 2014 American Chemical Society

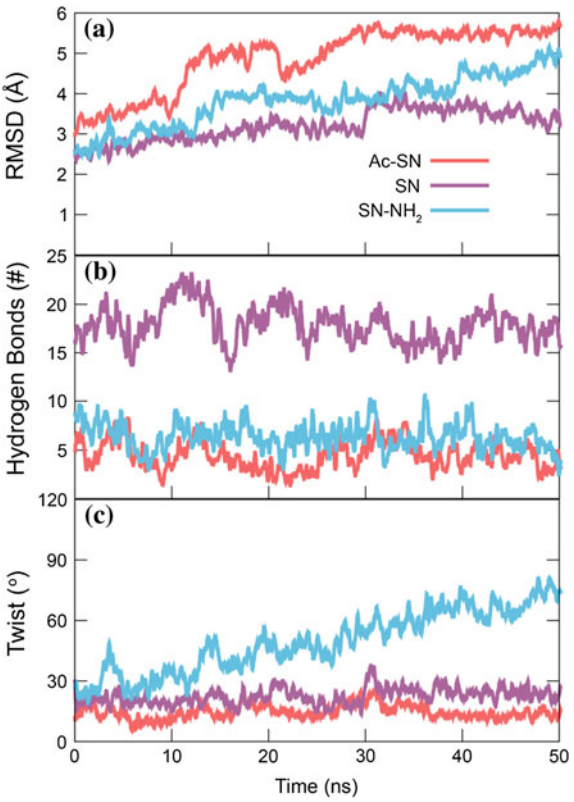


Fig. 7 Quantification of the MD simulations. All traces are running averages over 200 ps. **a** The RMSD was calculated for all C_α atoms. **b** The number of hydrogen bonds between the ends of the peptides was calculated using the atoms displayed in Fig. 9. **c** The twist along the fibril axis is the angle between two normals defined at each end of the β-sheets, and is the average of the two sheets. Adapted with permission from Andreasen et al. [8] Biochemistry, 53, 6968. Copyright 2014 American Chemical Society

do provide clues to the mechanisms governing the differences observed for the capping variants.

From the fibril structure at the end of the simulation (Fig. 8) it is evident that the overall cross-β structure has been retained during the simulation. The RMSD of SN–NH_2 steadily rises during the simulation and ends around 5 Å. This is due to an observed additional twisting of the fibril (Fig. 7c). The RMSD of Ac–SN is higher than for the other two peptides, reaching 5.5 Å at 30 ns and retains this value for the remainder of the simulation. This is due, in part, to very flexible peptide strands at the ends of the fibril which break the β-sheet structure; and in part to the capped peptide termini on the sides of the β-sheets which fold up to interact with the hydrophobic side chains on the surface of the fibril (Fig. 8c). In accordance with the inability of Ac–SN to form fibrils in vitro, this is the most unstable fibril. The time scale of the MD simulations, however, is too short to observe actual fibril dissociation.

The optimal hydrogen-bonding pattern for the peptide termini at the sides of the β-sheet in the antiparallel arrangement present in the 2KIB PDB-structure is schematically represented in Fig. 9. The possible number of backbone hydrogen

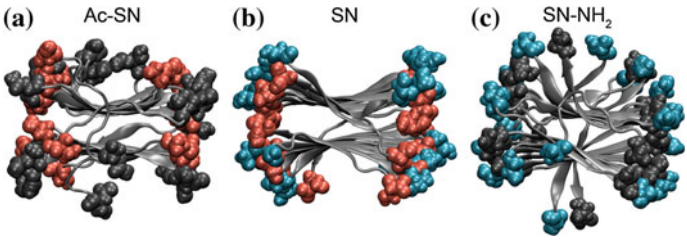


Fig. 8 The last frame from each MD simulation, shown looking down the fibril axis. Each peptide is represented in cartoon, and the first and last residue is represented with vdW spheres. *Red* or *blue* color indicates a residue/terminus with a negative or positive charge, respectively. **a** The neutral N-termini (*gray spheres*) of Ac–SN, which are curling up towards the side chains on the surface of the fibril. **b** This shows the stability of the SN fibril, which has not changed much from the beginning of the simulation. **c** The additional twist of the SN–NH_2 fibril is evident in this image

bonds between the termini is two for SN–NH_2, while Ac–SN and SN can make one hydrogen bond for each terminus (Fig. 9). The actual number of hydrogen bonds between the peptide termini in the simulation can be seen in Fig. 7b. The atoms depicted in Fig. 9 have been used for the summation, and a hydrogen bond is estimated to be present if the two hetero atoms are within 4 Å of each other, and the angle between donor atom, hydrogen atom, and acceptor atom deviates less than 30° from the ideal 180°. This means that the side chain from Ser1 has also been included in the calculation, since this side chain would most likely also participate in the hydrogen bonding pattern of the fibril in vitro.

The SN fibril has a very high number of hydrogen bonds, fluctuating around 18, due to the positive charges on the N-termini and the negative charges on the C-termini alternating on the side of the β-sheet. These favorable electrostatic

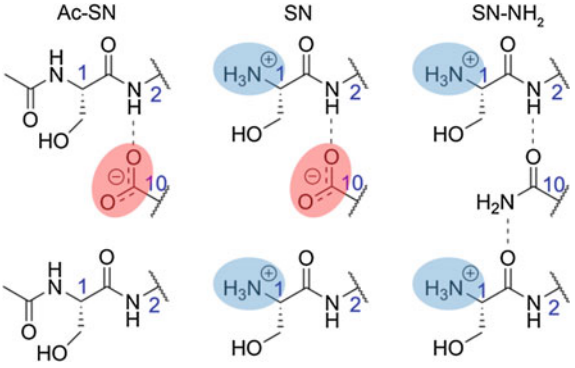


Fig. 9 Possible backbone hydrogen bonding networks in different fibril variants. Networks are shown as dotted lines at the sides of the β-sheets for the Ac–SN, SN, and SN–NH_2 capping variants. The *blue* numbers indicate the residue number, and the *red* and *blue circles* indicate a negative and positive charge, respectively. Reprinted with permission from Andreasen et al. [8] Biochemistry, 53, 6968. Copyright 2014 American Chemical Society

interactions keep the fibril stable. The number of hydrogen bonds for the Ac–SN variant fluctuates around 4 and the number of hydrogen bonds for the SN–NH_2 variant fluctuates around 6. While the difference between these two peptides is small, there are two different reasons for the low number of hydrogen bonds in the two fibrils. In SN–NH_2 the additional twisting of this fibril (see below) is probably the cause. Whereas for Ac–SN, the capped N-termini approach the hydrophobic side chains on the surface of the fibril, thereby breaking the hydrogen bonding pattern on the side of the fibril. During the simulation of the Ac–SN fibril (Fig. 8a), the N–termini with the acetyl capping groups (gray) move away from the negative charges on the C–termini (red). It is more favorable for the hydrophobic acetyl-cap to interact with the hydrophobic side chains present on the surface of the fibril than with the negatively charged C-termini or water. Furthermore, only one hydrogen bond per peptide terminus is present to keep the N-terminus of Ac–SN in place. The hydrophobicity of the capping group along with the low number of hydrogen bonds between the termini disrupts the β-sheet structure. This shows that a fibril composed of Ac–SN peptides is not stable even for as short a time as 50 ns, and the reason seems to be that the negative C–terminal carboxylate does not have enough hydrogen bonding partners to stabilize the cross-β arrangement. This is in accordance with the experimental evidence that this fibril does not form fibrils in vitro.

The differences in fibril morphology between the capping variants can also be observed in the MD simulations. The twist of the fibrils is calculated as in Fig. 4c in the chapter "Imaging Agent Binding to Amyloid Protofibrils". The twist of the SN–NH_2 (Figs. 7c and 8c) fibril becomes more pronounced during the simulation, while the twist of the Ac–SN and SN fibrils stays at $\sim 20°$ as in the original ss-NMR structure. This is in accordance with the SN–NH_2 capping variant being observed to produce twisted fibrils, while the SN fibril produces flat ribbons. The twist in the structure is a consequence of the single positive charge of the fibril being placed at one end of the peptide. The twist allows a greater distance between the individual charges on the side of the fibril. The peptides with complementary capping groups (SN, Ac–SN–NH_2, and Ac–SN–NMe) at each terminus of the peptide produce flat fibrils since they have no unfavorable interactions on one side of the β-sheets which need to be relieved by a twist. The lack of additional twist in the Ac–SN fibril could be a consequence of the C-termini being able to move away from one another, since they are not anchored to the same extent by hydrogen bonds as the N-termini in the SN–NH_2 fibril.

Discussion

We have shown that modification of the free termini of the amyloidogenic peptide SNNFGAILSS has profound impact on the fibril formation propensity, kinetics and morphology. This is manifested both in vitro and in silico, where modification of the peptide affects the hydrogen bonding network as well as the overall electrostatics.

Clearly, the termini play a major role in determining peptide self-assembly properties. Empirical support for this has also been provided by others. In a comparative study of the fibril formation of non-capped and double-capped Aβ [16–22] (KLVFFAE and Ac–KLVFFAE–NH_2) (which admittedly has a higher level of charged side chains than SNNFGAILSS), Tao and coworkers observed that the capped version of Aβ [16–22] formed flat nanotapes while the non-capped version formed twisted fibrils [17]. The twisted fibrils formed by the non-capped version of Aβ [16–22], displayed weaker hydrogen bonding and weaker π-π stacking of the aromatic residues, both of which are likely due to the twisting of the β–sheets in the fibrils as compared to the flat β-sheets, which make up the flat nanotapes. The twist in the β-sheets is most likely caused by electrostatic repulsion of two positive charges at the N-termini or two negative charges at the C-termini, in experiments conducted at pH 2 or pH 12, respectively. One of the charges is present on the terminus and one is present on the first or last residue side chain. Capping the termini of the peptides reduces the electrostatic repulsion and allows for the formation of flat fibrillar structures. The charge on the side chain is still present (in contrast to SNNFGAILSS); however, since the charge is located at the end of the side chain and not the backbone, the charges on adjacent residues are able to be farther apart without disrupting the backbone hydrogen bonding pattern.

The twisted morphology of the SN–NH_2 peptide shows that if the single charge is present on the backbone, the ability to form flat fibrils is abolished. The Ac–SN has a single charge at the C-termini, which, following the previous rationale should also result in twisted fibril morphology. However, the Ac–SN peptide does not form fibrils at all, most likely due to the lower number of possible hydrogen bonds which can be formed between the termini of the peptides. All other variants of SN form flat ribbon-like fibrils. This correlates with these peptides having complementary capping groups at the N- and C-termini, i.e. no repulsive charges acting on the fibril morphology, or a higher possible number of hydrogen bonds keeping the terminus together in the cross-β arrangement.

hIAPP is C-terminally amidated in vivo, and the role of this amidation on the amyloidogenic behavior of hIAPP has previously been examined. C-terminally amidated hIAPP fibrillates faster than non-amidated hIAPP, but has a lower propensity to form cytotoxic species during the fibril formation [35, 36]. This suggests that even in the full-length hormone, capping of the termini can have a dramatic effect on the amyloidogenic behavior of a protein.

Conclusion

The simulations in the present study have provided hints to the underlying reasons for the differences in fibril formation properties of the capping variants. It is obvious from the present work that slight modifications in the peptide can change the physical and chemical properties of the peptide greatly, and thereby affect fibril formation ability, kinetics, and morphology. This has previously been disregarded

when designing protein fragments as model systems of a full-length protein. However, this aspect has to be considered in future fibril formation studies using model peptides. Additionally, knowledge about which factors are involved in fibril formation kinetics, propensity, and stability is of utmost importance in the pursuit of treatments for amyloid diseases.

References

1. Benzinger TLS, Gregory DM, Burkoth TS, Miller-Auer H, Lynn DG, Botto RE, Meredith SC (2000) Two-dimensional structure of β-amyloid(10-35) fibrils. Biochemistry 39:3491–3499
2. Kirschner DA, Inouye H, Duffy LK, Sinclair A, Lind M, Selkoe DJ (1987) Synthetic peptide homologous to β protein from alzheimer disease forms amyloid-like fibrils in vitro. Proc Natl Acad Sci USA 84:6953–6957
3. Krysmann MJ, Castelletto V, Kelarakis A, Hamley IW, Hule RA, Pochan DJ (2008) Self-assembly and hydrogelation of an amyloid peptide fragment. Biochemistry 47:4597–4605
4. Brown DR (2000) Prion protein peptides: optimal toxicity and peptide blockade of toxicity. Mol Cell Neurosci 15:66–78
5. Hope J, Shearman MS, Baxter HC, Chong A, Kelly SM, Price NC (1996) Cytotoxicity of prion protein peptide ($PrP^{106-126}$) differs in mechanism from the cytotoxic activity of the alzheimer's disease amyloid peptide, Aβ25–35. Neurodegeneration 5:1–11
6. van der Wel PCA, Lewandowski JR, Griffin RG (2007) Solid-state NMR study of amyloid nanocrystals and fibrils formed by the peptide GNNQQNY from yeast prion protein Sup35p. J Am Chem Soc 129:5117–5130
7. Westermark P, Engström U, Johnson KH, Westermark GT, Betsholtz C (1990) Islet amyloid polypeptide: pinpointing amino acid residues linked to amyloid fibril formation. Proc Natl Acad Sci USA 87:5036–5040
8. Andreasen M, Nielsen SB, Mittag T, Bjerring M, Nielsen JT, Zhang S, Nielsen EH, Jeppesen M, Christiansen G, Besenbacher F, Dong M, Nielsen NC, Skrydstrup T, Otzen DE (2012) Modulation of fibrillation of hIAPP core fragments by chemical modification of the peptide backbone. Biochim Biophys Acta, Proteins Proteomics 1824:274–285
9. Tenidis K, Waldner M, Bernhagen J, Fischle W, Bergmann M, Weber M, Merkle ML, Voelter W, Brunner H, Kapurniotu A (2000) Identification of a penta- and hexapeptide of islet amyloid polypeptide (IAPP) with amyloidogenic and cytotoxic properties. J Mol Biol 295:1055–1071
10. Sørensen J, Periole X, Skeby KK, Marrink S-J, Schiøtt B (2011) Protofibrillar assembly toward the formation of amyloid fibrils. J Phys Chem B 2:2385–2390
11. Marsh JL, Walker H, Theisen H, Zhu YZ, Fielder T, Purcell J, Thompson LM (2000) Expanded polyglutamine peptides alone are intrinsically cytotoxic and cause neurodegeneration in drosophila. Hum Mol Genet 9:13–25
12. Raspe M, Gillis J, Krol H, Krom S, Bosch K, van Veen H, Reits E (2009) Mimicking proteasomal release of polyglutamine peptides initiates aggregation and toxicity. J Cell Sci 122:3262–3271
13. Jarvis JA, Craik DJ, Wilce MCJ (1993) X-ray-diffraction studies of fibrils formed from peptide fragments of transthyretin. Biochem Biophys Res Commun 192:991–998
14. Jaroniec CP, MacPhee CE, Astrof NS, Dobson CM, Griffin RG (2002) Molecular conformation of a peptide fragment of transthyretin in an amyloid fibril. Proc Natl Acad Sci USA 99:16748–16753
15. Sørensen J, Hamelberg D, Schiøtt B, McCammon JA (2007) Comparative MD analysis of the stability of transthyretin providing insight into the fibrillation mechanism. Biopolymers 86: 73–82

16. Zhang S, Andreasen M, Nielsen JT, Liu L, Nielsen EH, Song J, Ji G, Sun F, Skrydstrup T, Besenbacher F, Nielsen NC, Otzen DE, Dong M (2013) Coexistence of ribbon and helical fibrils originating from hIAPP$_{20–29}$ revealed by quantitative nanomechanical atomic force microscopy. Proc Natl Acad Sci USA 110:2798–2803
17. Tao K, Wang J, Zhou P, Wang C, Xu H, Zhao X, Lu JR (2011) Self-assembly of short Aβ(16-22) peptides: effect of terminal capping and the role of electrostatic interaction. Langmuir 27:2723–2730
18. Roberts AN, Leighton B, Todd JA, Cockburn D, Schofield PN, Sutton R, Holt S, Boyd Y, Day AJ, Foot EA (1989) Molecular and functional characterization of amylin, a peptide associated with type 2 diabetes mellitus. Proc Natl Acad Sci USA 86:9662–9666
19. Yonemoto IT, Kroon GJA, Dyson HJ, Balch WE, Kelly JW (2008) Amylin proprotein processing generates progressively more amyloidogenic peptides that initially sample the helical state. Biochemistry 47:9900–9910
20. Nanga RPR, Brender JR, Vivekanandan S, Ramamoorthy A (2011) Structure and membrane orientation of IAPP in its natively amidated form at physiological pH in a membrane environment. Biochim Biophys Acta, Biomembr 1808:2337–2342
21. Patil SM, Xu S, Sheftic SR, Alexandrescu AT (2009) Dynamic α-helix structure of micelle-bound human amylin. J Biol Chem 284:11982–11991
22. Nielsen JT, Bjerring M, Jeppesen MD, Pedersen RO, Pedersen JM, Hein KL, Vosegaard T, Skrydstrup T, Otzen D, Nielsen NC (2009) Unique identification of supramolecular structures in amyloid fibrils by solid-state NMR spectroscopy. Angew Chem, Int Ed 48:2118–2121
23. Wiltzius JJW, Sievers SA, Sawaya MR, Cascio D, Popov D, Riekel C, Eisenberg D (2008) Atomic structure of the cross-β spine of islet amyloid polypeptide (amylin). Protein Sci 17:1467–1474
24. Wiltzius JJW, Landau M, Nelson R, Sawaya MR, Apostol MI, Goldschmidt L, Soriaga AB, Cascio D, Rajashankar K, Eisenberg D (2009) Molecular mechanisms for protein-encoded inheritance. Nat Struct Mol Biol 16:973–978
25. Schrödinger LLC, New York NY (2011) Maestro, version 9.2
26. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. J Mol Graph 14:33–38
27. Neria E, Fischer S, Karplus M (1996) Simulation of activation free energies in molecular systems. J Chem Phys 105:1902–1921
28. Piana S, Lindorff-Larsen K, Shaw DE (2011) How robust are protein folding simulations with respect to force field parameterization? Biophys J 100:L47–L49
29. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kalé L, Schulten K (2005) Scalable molecular dynamics with NAMD. J Comput Chem 26 1781–1802
30. Darden T, York D, Pedersen L (1993) Particle mesh ewald: an N·log(N) method for ewald sums in large systems. J Chem Phys 98:10089–10092
31. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A smooth particle mesh ewald method. J Chem Phys 103:8577–8593
32. Andersen HC (1983) Rattle: a "velocity" version of the shake algorithm for molecular dynamics calculations. J Comput Phys 52:24–34
33. LeVine H (1995) Thioflavine T interaction with amyloid β-sheet structures. Amyloid 2:1–6
34. Madine J, Jack E, Stockley PG, Radford SE, Serpell LC, Middleton DA (2008) Structural insights into the polymorphism of amyloid-like fibrils formed by region 20–29 of amylin revealed by solid-state NMR and X-ray fiber diffraction. J Am Chem Soc 130:14990–15001
35. Chen M, Zhao D, Yu Y, Li W, Chen Y, Zhao Y, Li Y (2013) Characterizing the assembly behaviors of human amylin: a perspective derived from C-terminal variants. Biochim Biophys Acta, Biomembr 49:1799–1801
36. Tu L, Serrano A, Zanni M, Raleigh D (2014) Mutational analysis of preamyloid intermediates: the role of His-Tyr interactions in islet amyloid formation. Biophys J 106:1520–1527

Coarse Grained Study of Amyloid Protofibril Aggregation

This study is described in the manuscript entitled "Protofibrillar Assembly Toward the Formation of Amyloid Fibrils". The study was mainly conducted by Jesper Sørensen. My part was in building the 10-residue fibril fragment from the 7-residue fibril structure obtained by ss-NMR [1] and performing the all-atom simulations. Furthermore, I participated in the analysis and discussion of the work. The motivation behind the study and the main conclusions will be outlined in this section, followed by a discussion of the use of CG simulations to investigate amyloid aggregation.

Introduction

Elucidating the mechanism of amyloid protein aggregation is a step in the direction toward understanding and ultimately preventing diseases such as AD and T2DM. Therefore, this is a topic under intense scrutiny, which is unfortunately highly challenged by the properties of the process itself. Conventional techniques for studying globular proteins often require high concentration of protein. A solution with high concentration of amyloid protein, aggregates rapidly, which makes the study of the actual process very difficult. Computational studies can help in this regard by simulating the aggregation events at high resolution in both time and space. However, as of yet, we are still limited in the computational resources that we have available. Therefore, computational methods and models have been developed to enable a speed-up of the molecular events. An example of this is enhanced sampling methods applied to atomistic simulations, which aim to keep the high resolution in space provided by all-atom FFs. Another approach is using CG FFs where the number of degrees of freedom in the system has been reduced by a decreased number of particles.

In this study the CG approach has been used to investigate the process of fibril formation; in particular, the process of elongation of fibrils from preformed

K.K. Skeby, *Computational Modelling of the Human Islet Amyloid Polypeptide*,
Springer Theses, DOI 10.1007/978-3-319-20040-8_6

protofibrils. This was motivated by studies showing the significant increase in aggregation rate by the use of oligomers as seeds for the formation of amyloid fibrils [2].

Experimental Section

CG MD simulations were performed of the self-assembly of 27 protofibril fragments. Each fragment consists of two antiparallel β-sheets formed by 10 decapeptide fragments of hIAPP (hIAPP [20–29]; SNNFGAILSS) (Fig. 1a). The protofibril fragments were placed on a 3 × 3 × 3 grid with initial random orientations (Fig. 1b). The MARTINI CG FF [3, 4] was used to model the system, and GROMACS 4.0.7 [5] was used to perform the simulations. The system was initially simulated at three different temperatures (300, 400, and 500 K) for 10 μs*. To increase statistics at 300 K, an additional 10 simulations of 2 μs* were performed. The asterisk denotes effective time accounting for the speed-up of the dynamics by the MARTINI FF [3, 6]. The ElNeDyn approach [7] was applied to each individual β-sheet to maintain the secondary structure of the protein, and was calibrated against all-atom simulations of the fibril. The all-atom simulations were performed with the AMBER03 [8] FF with the NAMD 2.7 [9] software. See further details of the simulations in the published paper.

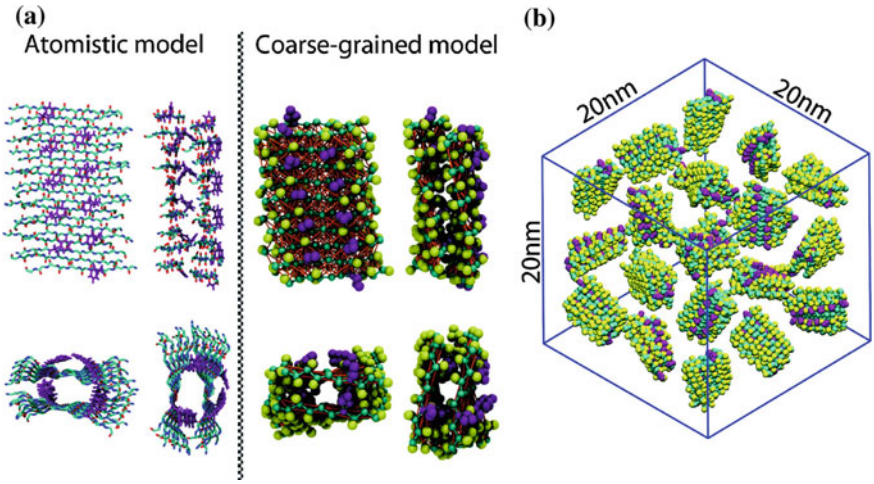


Fig. 1 **a** Atomistic and CG model of the protofibril fragment. **b** Initial placement of protofibril fragments in the simulation box. Reprinted with permission from Ref. [10]. Copyright 2011 American Chemical Society

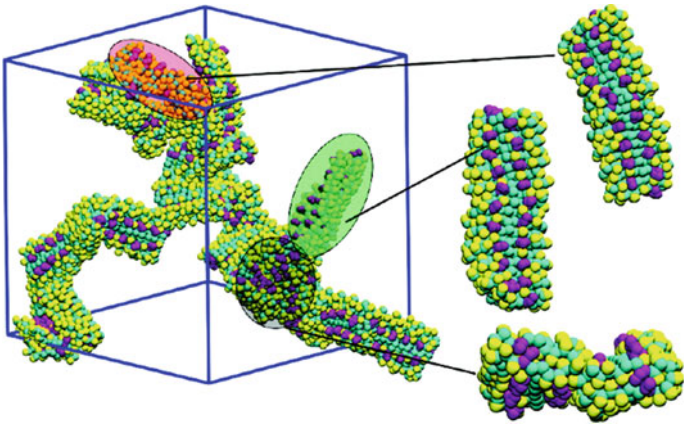


Fig. 2 Aggregation after 2 μs*. The seemingly random aggregation has elements of ordered assembly. Reprinted with permission from Ref. [10]. Copyright 2011 American Chemical Society

Results

A visual inspection of the system at 300 K after 2 μs* gives the impression of random assembly; however, elements of ordered assembly are present (Fig. 2). Two protofibrils assembling end–to–end is seen twice in the snapshot, and lateral assembly of two protofibrils at the sides is also observed.

The aggregation of the protofibril segments is followed by the burial percentage, which is the amount of surface area, which is buried upon association of the segments (Fig. 3a). Within 2 μs*, the burial percentage reaches a plateau level of around 35 and 40 % for 300 and 400 K, respectively. The simulation at 500 K takes longer to reach a plateau of around 60 %. This behavior is representative of

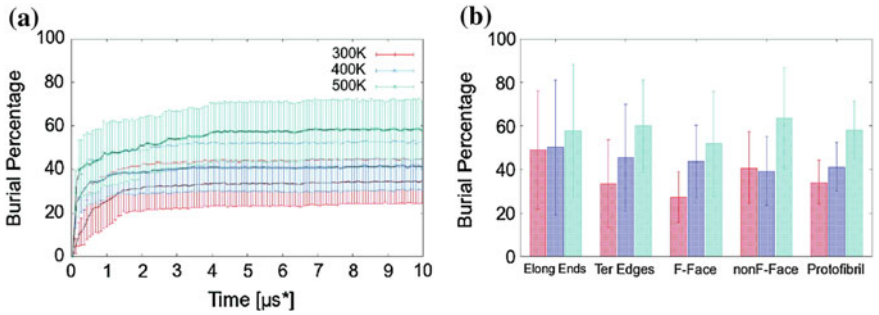


Fig. 3 **a** Burial percentage of the three 10 μs* simulations at different temperatures. **b** The average burial degree of each protofibril surface during the final 5 μs* of simulation. The error bars are standard deviations. Reprinted with permission from Ref. [10]. Copyright 2011 American Chemical Society

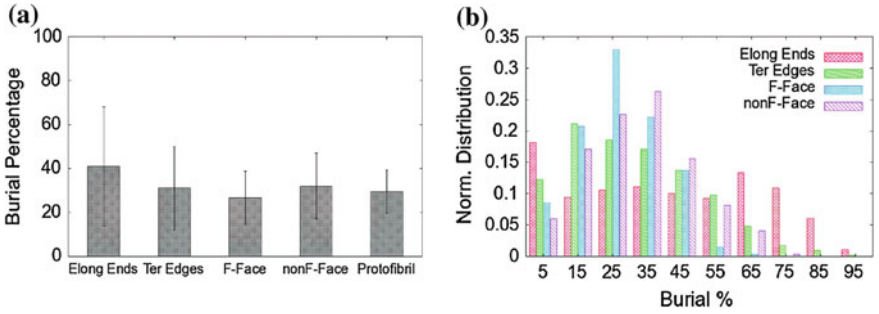


Fig. 4 **a** Average burial percentage for each face of the fibril for the 27 protofibrils during the last 500 ns* of the ten 2 μs* simulations. **b** Distribution of burial percentages for the 27 protofibrils during the last 500 ns* of the ten 2 μs* simulations. Reprinted with permission from Ref. [10]. Copyright 2011 American Chemical Society

the rapid formation of an encounter-complex followed by the slow optimization of the protein-protein interfaces [11].

Each protofibril has six faces: two ends where elongation of the fibril would occur; two sides; a Phe face; and a non-Phe face. The average burial percentage of each face during the final 5 μs* reveals that at 300 K the elongation ends have a tendency for higher burial than the other faces (Fig. 3b). This trend is lost at higher temperatures.

To enhance the sampling, ten additional simulations of 2 μs* at 300 K were performed with different random orientations of the 27 protofibrils. The average burial during the last 500 ns* is reported in Fig. 4a. During these simulations the preference for association of the elongation ends is also apparent. The distribution of the burial (Fig. 4b) shows that the elongation ends have a significant amount of faces which are more than 60 % buried. The burial of a perfectly matched end-to-end assembly is 80 %.

The preference for association of the elongation ends seen here suggests a preference for elongation by protofibril oligomers in this size regime. AFM studies have shown that lateral growth occurs for oligomers of 23 nm in length before elongation occurs [12]. This is consistent with the present results, as the protofibril segments in the simulations are 4.5 nm in length, and therefore should elongate before lateral growth is initiated. The loss of preferential assembly of the elongation ends at higher temperatures may be an example of the varying fibril morphologies that occur with different growth conditions [13].

Discussion

It is important to understand the limitations of the model to be able to understand and interpret the results. The obvious limitation of the MARTINI CG model is the inability to change the secondary structure of the protein. This would seem like a

complication, which could eliminate the use of this model for the study of protein aggregation. However, it is possible to get insight into the later stages of aggregation when the secondary structure of the protein has already changed to β-sheet.

Another limitation in the MARTINI FF is the lack of directionality of hydrogen bonds. This creates a problem in forming backbone secondary structures, which is to some degree compensated for by the preferential interaction of backbone beads with other backbone beads over solvent when it is involved in secondary structure [4]. This may have underestimated the elongation at the ends as the backbone-backbone interactions are not well described.

Studies have indicated that the MARTINI 2.1 FF overestimates the hydrophobicity of aromatic residues and underestimates polar and charge interactions [14]. This further supports the notion that the association at the elongation ends might have been underestimated in the present study. These problems have been addressed in the newest version of the protein parameters for the MARTINI FF [14] However, the problem of the missing directional hydrogen bonds still persists, although investigations into addressing this, e.g. by polarization of the backbone beads, are being performed [15].

The present study seeks to overcome some of the challenges that experimental investigation of protein aggregation faces, especially the limitation of resolution in time and size that experimental techniques are currently capable of achieving. Using atomic resolution simulations it is possible to examine the initial and final points in protein aggregation, namely the monomer and fibril structures. The gap in between can be investigated using coarser models.

Several types of CG models exists, each developed to study a specific aspect of the fibril formation mechanism (Fig. 5) [16]. The level of coarse graining also determines the level at which events can be studied. Higher resolution models, such as the MARTINI model (Fig. 5a) and the OPEP model [17], are parameterized based on all-atom simulations and experimental data. They rely on transferability

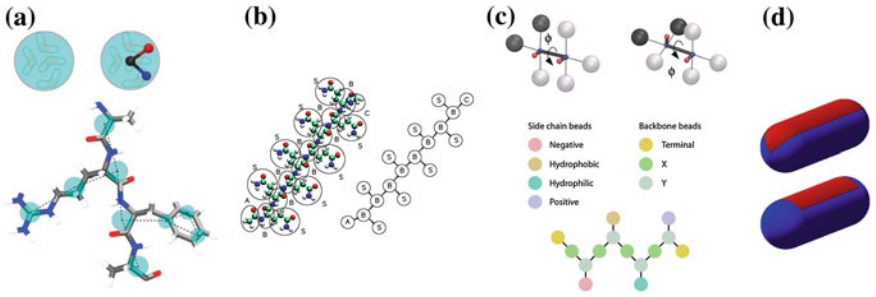


Fig. 5 Levels of coarse graining used in studying protein aggregation. The examples were assembled with inspiration from Wu et al. [16]. **a** MARTINI Model. Reprinted with permission from Ref. [20]. Copyright 2013 Royal Society of Chemistry. **b** Systematic coarse graining. Reprinted with permission from Ref. [21]. Copyright 2010 American Chemical Society. **c** Phenomenological models by Shea et al. [22] (*bottom*) and Caflisch et al. (*top*) Reprinted from Ref. [23]. **d** Simple model. Reprinted from Ref. [24]

meaning that the models can be used for other proteins than were used in the parameterization. The OPEP model has a united-atom backbone and represents the side chain as a single CG particle, which enables the directionality of backbone hydrogen bonds, and thereby also changes in secondary structure. Both the MARTINI and the OPEP FFs have been used to study aggregation of hIAPP. The MARTINI FF was used to assemble membrane pore structures of an α-helical conformation of hIAPP in a lipid membrane, followed by reverse-mapping of the structure into an all-atom representation [18]. The OPEP FF was used to study the dimerization of both hIAPP and rIAPP in solution, showing β-sheet structure in the C-terminus of hIAPP which was not present in rIAPP [19].

Systematic coarse graining attempts to lower the resolution even further while still retaining sequence information. Specific protein sequences are modeled with typically 2–3 beads per residue, one backbone bead, and one side chain bead (Fig. 5b). The parameters for the model are fitted for each system independently. This requires a lot of atomistic sampling; however, it allows the study of the interactions between many copies of a single protein, extending the system size beyond what is possible in both atomistic simulations and the MARTINI and OPEP models. Investigations of the aggregation of polyglutamine [21] and polyalanine [25] have been performed with these types of models, and have examined the fluctuations in structural assemblies as well as the nucleation mechanism.

Phenomenological models typically investigate the underlying aggregation mechanisms instead of the aggregation of a specific protein or peptide. They have a low resolution with one or two beads per residue and a tunable parameter to modify the aggregation propensity of the peptide. Examples include the three-bead model by Shea et al. [22] and the two-bead model by Caflisch et al. (Fig. 5c) [23, 26], which both showed that the aggregation mechanism depends critically on the aggregation propensity of the peptide. The Caflisch model has further been extended to also study the aggregation of amyloid peptides in the presence of a lipid vesicle, showing the disruption of the lipid bilayer and subsequent release of the vesicle contents [27].

Simplistic models have a very low resolution, representing each peptide as a single unit, and similar to phenomenological models, simplistic models investigate the aggregation mechanisms instead of specific proteins (Fig. 5d). Examples of this type of model are a cuboid model [28], a rod-like model [24], and an orientable stick model [29], which all reproduce the sigmoidal kinetics profile observed in simulations.

Conclusion

It is clearly evident that atomistic simulation is not going to reach the time and length scales required to simulate the full aggregation of amyloid peptides within the foreseeable future. Using a CG representation of a small oligomeric assembly of an amyloid protein, we have been able to characterize a small part of the

aggregation pathway. The current range of CG models available each have strengths and weaknesses; however, it is clear that multiple levels of coarse graining are needed to investigate the entire aggregation pathway. Furthermore, a close collaboration and comparison with experiments is needed to validate the simulations as well as explain the phenomena observed in the experiments. Hopefully, this will provide valuable insight in the mechanism of amyloid aggregation in the near future.

References

1. Nielsen JT, Bjerring M, Jeppesen MD, Pedersen RO, Pedersen JM, Hein KL, Vosegaard T, Skrydstrup T, Otzen D, Nielsen NC (2009) Unique identification of supramolecular structures in amyloid fibrils by solid-state NMR spectroscopy. Angew Chem Int Ed 48:2118–2121
2. Kayed R, Bernhagen J, Greenfield N, Sweimeh K, Brunner H, Voelter W, Kapurniotu A (1999) Conformational transitions of islet amyloid polypeptide (IAPP) in amyloid formation in vitro. J Mol Biol 287:781–796
3. Marrink S-J, Risselada HJ, Yefimov S, Tieleman DP, de Vries AH (2007) The MARTINI force field: coarse grained model for biomolecular simulations. J Phys Chem B 111:7812–7824
4. Monticelli L, Kandasamy SK, Periole X, Larson RG, Tieleman DP, Marrink S-J (2008) The MARTINI coarse-grained force field: extension to proteins. J Chem Theor Comput 4:819–834
5. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theor Comput 4:435–447
6. Marrink S-J, de Vries AH, Mark AE (2004) Coarse grained model for semiquantitative lipid simulations. J Phys Chem B 108:750–760
7. Periole X, Cavalli M, Marrink S-J, Ceruso MA (2009) Combining an elastic network with a coarse-grained molecular force field: structure, dynamics, and intermolecular recognition. J Chem Theor Comput 5:2531–2543
8. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang J, Kollman PA (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J Comput Chem 24:1999–2012
9. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kalé L, Schulten K (2005) Scalable molecular dynamics with NAMD. J Comput Chem 26:1781–1802
10. Sørensen J, Periole X, Skeby KK, Marrink S-J, Schiøtt B (2011) Protofibrillar assembly toward the formation of amyloid fibrils. J Phys Chem B 2:2385–2390
11. Gabdoulline RR, Wade RC (2002) Biomolecular diffusional association. Curr Opin Struct Biol 12:204–213
12. Green JD, Goldsbury C, Kistler J, Cooper GJS, Aebi U (2004) Human amylin oligomer growth and fibril elongation define two distinct phases in amyloid formation. J Biol Chem 279:12206–12212
13. Fändrich M, Meinhardt J, Grigorieff N (2009) Structural polymorphism of alzheimer Aβ and other amyloid fibrils. Prion 3:89–93
14. de Jong DH, Singh G, Bennett WFD, Arnarez C, Wassenaar TA, Schäfer LV, Periole X, Tieleman DP, Marrink S-J (2013) Improved parameters for the martini coarse-grained protein force field. J Chem Theor Comput 9:687–697

15. Yesylevskyy SO, Schäfer LV, Sengupta D, Marrink S-J (2010) Polarizable water model for the coarse-grained MARTINI force field. PLoS Comput Biol 6:e1000810
16. Wu C, Shea J (2011) Coarse-grained models for protein aggregation. Curr Opin Struct Biol 21:209–220
17. Chebaro Y, Pasquali S, Derreumaux P (2012) The coarse-grained OPEP force field for non-amyloid and amyloid proteins. J Phys Chem B 116:8741–8752
18. Pannuzzo M, Raudino A, Milardi D, La Rosa C, Karttunen M (2013) A-helical structures drive early stages of self-assembly of amyloidogenic amyloid polypeptide aggregate formation in membranes. Sci Rep 3:2781
19. Laghaei R, Mousseau N, Wei G (2011) Structure and thermodynamics of amylin dimer studied by hamiltonian-temperature replica exchange molecular dynamics simulations. J Phys Chem B 115:3146–3154
20. Marrink S-J, Tieleman DP (2013) Perspective on the martini model. Chem Soc Rev 42:6801–6822
21. Wang Y, Voth GA (2010) Molecular dynamics simulations of polyglutamine aggregation using solvent-free multiscale coarse-grained models. J Phys Chem B 114:8735–8743
22. Bellesia G, Shea J (2007) Self-assembly of β-sheet forming peptides into chiral fibrillar aggregates. J Chem Phys 126:245104
23. Pellarin R, Guarnera E, Caflisch A (2007) Pathways and intermediates of amyloid fibril formation. J Mol Biol 374:917–924
24. Vácha R, Frenkel D (2011) Relation between molecular shape and the morphology of self-assembling aggregates: a simulation study. Biophys J 101:1432–1439
25. Carmichael SP, Shell MS (2012) A new multiscale algorithm and its application to coarse-grained peptide models for self-assembly. J Phys Chem B 116:8383–8393
26. Pellarin R, Caflisch A (2006) Interpreting the aggregation kinetics of amyloid peptides. J Mol Biol 360:882–892
27. Friedman R, Pellarin R, Caflisch A (2009) Amyloid aggregation on lipid bilayers and its impact on membrane permeability. J Mol Biol 387:407–415
28. Zhang J, Muthukumar M (2009) Simulations of nucleation and elongation of amyloid fibrils. J Chem Phys 130:035102
29. Irback A, Jonsson SAE, Linnemann N, Linse B, Wallin S (2013) Aggregate geometry in amyloid fibril nucleation. Phys Rev Lett 110:058101

Conclusion and Perspectives

This dissertation is focused on the investigation of biological mechanisms related to hIAPP. This peptide is involved in T2DM where it misfolds and subsequently aggregates within the pancreas in the form of amyloid fibrils. The amyloid field of research spans various length and timescales from the ns/Å-scale of ligand binding to the day/μm-scale of mature fibril formation and behavior. We have used computational studies to investigate small segments of this very complex subject.

In AD, Aβ deposits in the brain tissue as amyloid fibrils. Diagnosis of AD is complicated and requires distinguishing the disease from other forms of dementia relying heavily on cognitive evaluation. Establishing a reliable method of biomarker detection specific for AD could aid greatly in distinguishing AD from other forms of dementia, in early detection, and in evaluation of treatments. A promising, non-invasive approach to biomarker detection is the imaging of brain amyloid with either PET or MRI using an imaging agent selective for Aβ amyloid. In the chapter "Imaging Agent Binding to Amyloid Protofibrils", the binding of amyloid imaging agents to an amyloid fibril was investigated using MD simulations. Thirteen different imaging agents were included in the study, spanning a range of scaffold structures and attached functional groups. An amyloid fibril segment formed from a peptide fragment from hIAPP was used as a model fibril, as it contains the characteristics of an amyloid fibril while keeping the size of the system to a minimum. This enables us to extract common features of binding of imaging agents to amyloid fibrils from the simulations. We identified a common binding mode for the imaging agents in the surface grooves on the fibril created by the repeating pattern of side chains. The affinity of the ligand for a binding site depends on the peptide side chains as well as the chemical properties of the ligand. Recently, a ss-NMR structure of an in vivo Aβ amyloid fibril was determined [1]. This could be a great step in the direction towards intelligent design of a high-affinity and high-specificity amyloid ligand, which can be used in the in vivo detection of amyloid.

Amyloid proteins are cytotoxic; however, the precise mechanism of cytotoxicity has not been elucidated. Several theories exist, including the formation of toxic oligomers which disrupt the cell membrane during fibril growth. The experimental

K.K. Skeby, *Computational Modelling of the Human Islet Amyloid Polypeptide*,
Springer Theses, DOI 10.1007/978-3-319-20040-8_7

study of cytotoxicity is complicated by the rapid aggregation of the amyloid proteins making it difficult to isolate single species from the aggregation process. However, it is established that hIAPP interacts with phospholipid membranes, and anionic membranes in particular. In the chapter "Determining the Aggregation Prone Structure of hIAPP", MD simulations of the interaction between hIAPP and a phospholipid membrane were presented. The aim of the study was to understand the initial steps in the process of hIAPP-membrane interaction. Using a highly mobile membrane model we were able to show that the attraction of the two positively charged residues (Lys1 and Arg11) and the N-terminal charge with the anionic lipids was responsible for attracting the N-terminal part of the peptide to the membrane. The C-terminus was not associated with the membrane, but was unfolded and very dynamic in the solvent phase. When His18 is protonated, the aggregation of hIAPP is inhibited. In the solvent, this is most likely due to the electrostatic repulsion between the positively charged peptides. We showed that the protonation of His18 results in deeper binding of the peptide in the membrane and causes the C-terminus to be immobilized. This could be the mechanism by which low pH (<6) inhibits the aggregation when the peptide is bound to the membrane. The differences in membrane binding of the two protonation states of the peptide may be due to two hydrophobic segments of the peptide sequence flanking His18. They do not interact favorably with the lipid head-groups. However, when His18 is protonated, the favorable interaction of this residue with the anionic lipids enables these hydrophobic residues to cross the head-group region. This allows the favorable interaction of the hydrophobic residues with the hydrophobic center of the membrane. The hypotheses proposed in this study require further investigation, as we cannot make conclusions about the aggregation of peptides when we only have a single peptide in the simulation. Therefore, we have initiated further simulations of hIAPP aggregation at the membrane, which are built upon the present simulations.

Due to the difficulties of studying amyloid fibrils, model systems are often employed to simplify systems and possibly make the interpretation of results easier. An example of this is the use of smaller peptide fragments as model systems for larger peptides, as these fibrillate with the same kinetics as larger peptides, as well as form cross-β structures. An example of this is present in this thesis in the chapter "Imaging Agent Binding to Amyloid Protofibrils", were an amyloid fibril segment formed by 7-residue peptide fragments has been used as a model system for a larger amyloid fibril. However, it is important to consider the effects of using only a part of the full-length peptide when interpreting the results. In the chapter "Effect of Terminal Capping on Aggregation of Peptide Fragments", the effects of various capping groups on the peptide termini are investigated. This is relevant as the peptide fragment will often be a sequence extracted from a full-length amyloid peptide, and not capping the termini will introduce a charge at each end of the peptide fragment, which is not present at that position in the full-length peptide. We find that the nature of the capping group changes both the fibril formation kinetics and the resulting fibril morphology. Capping both termini or leaving both termini uncapped results in flat fibrils. Capping the N-terminal abolishes fibril formation, while capping the C-terminus results in twisted fibrils.

MD simulations provide clues to the reasons for the behavior of the peptides. In the simulation of a fibril with N-terminal acetyl-caps (Ac–SN), the hydrophobic acetyl-groups fold up to interact with the hydrophobic side chains, which disrupts the hydrogen bonding network between the peptide termini. The simulation of the C-terminally amidated fibril (SN–NH_2) reveals that the twisted morphology is caused by charge repulsion of the N-termini. The results highlight the importance of choosing a model system with care as amyloid fibrils are a misfolded state, which is highly susceptible to changes in the environment and peptide sequence.

Aggregation of amyloid fibrils is a process that is difficult to study on a molecular scale due to the rapid nature of the process. It is therefore necessary to combine the strengths of different methods to elucidate the mechanism. Atomistic MD simulations are an important complementary technique to experimental studies as they provide very high resolution at time scales of <1 μs, enabling the detailed analysis of parts of the aggregation process. However, the limitations in sampling speed and system size necessitate the use of methods such as coarse graining for studying systems with many peptide monomers. In the chapter“ Coarse Grained Study of Amyloid Protofibril Aggregation”, the study of aggregation of amyloid protofibril oligomers is presented. Several assembly simulations using the MARTINI FF were performed of a system composed of 27 protofibrils randomly oriented on a 3 × 3 × 3 grid. The simulations showed a preference for association at the elongation ends, which suggests that elongation is the main aggregation process at the 4.5 nm–scale of these protofibrils. This preference was lost at higher temperatures exemplifying the susceptible nature of amyloid proteins. However, these simulations also highlight some of the challenges that are faced by CG models. The MARTINI FF does not contain a representation of directional hydrogen bonds, which are highly important in amyloid fibrils. It is therefore not possible to study conformational change of secondary structure. This is a situation that the developers of the MARTINI FF are trying to improve. A discussion in the chapter “Coarse Grained Study of Amyloid Protofibril Aggregation” of the various types of CG models which have been used to study amyloid aggregation highlight the importance of the choice of model for studying a particular part of the aggregation process.

Recent developments in computational speed via the use of e.g. GPUs and specialized hardware have brought us even further in the pursuit of understanding the molecular mechanisms of life through computational studies. Although, many new challenges still lie ahead such as the handling and analysis of the large amounts of data being produced as a consequence of the increase in sampling ability. Furthermore, it is possible that we have reached a point where FFs cannot be improved further by reparameterization, and that different strategies, such as polarization effects, are needed to overcome some of the limitations currently seen in the most widely used all-atom FFs. Amyloid aggregation is a perfect example of a subject where simulation and experiment need and complement each other. We are still a long way from simulating the entire aggregation process at an atomistic level, and it is unlikely that this will ever be possible; however, the studies presented here show how computational studies can complement experimental studies and provide essential insight within the field of amyloid research.

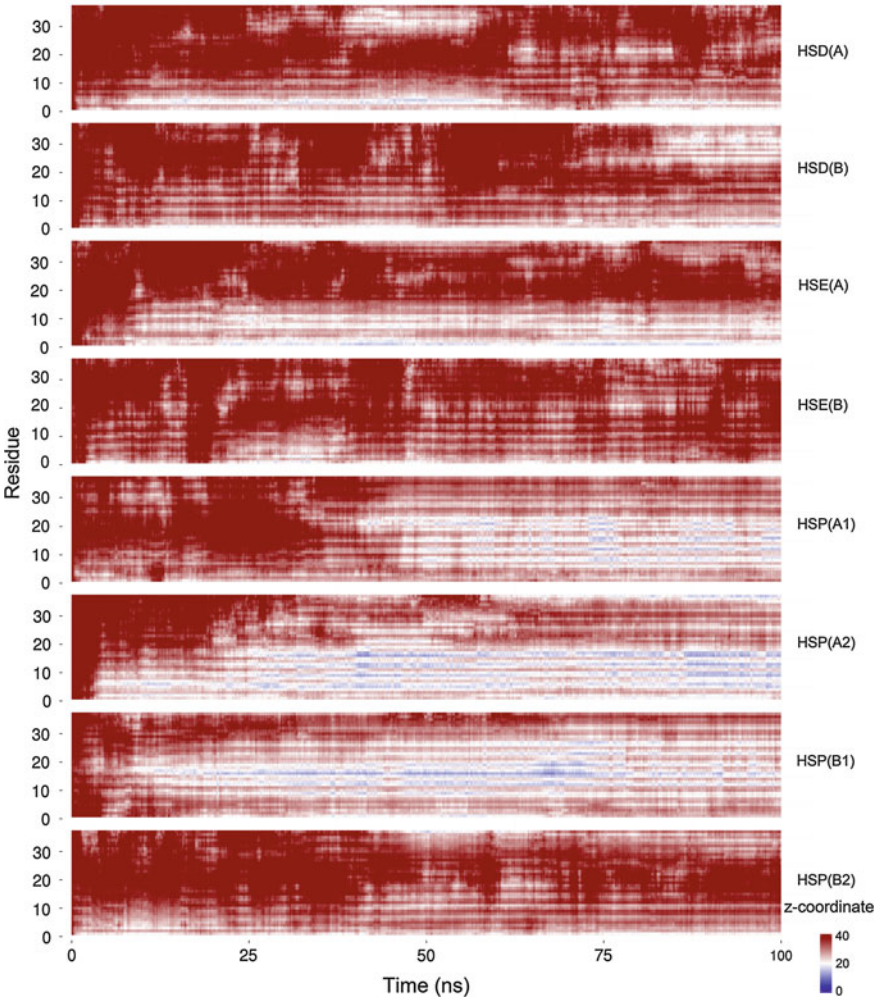
Reference

1. Paravastu AK, Qahwash I, Leapman RD, Meredith SC, Tycko R (2009) Seeded growth of β-amyloid fibrils from alzheimer's brain-derived fibrils produces a distinct fibril structure. Proc Natl Acad Sci USA 106:7443–7448

Appendix A

The z-coordinate of the center of mass of each residue side chain for the $hIAPP_{1-37}$ simulations. The center of the membrane is at z = 0 Å and the average position of the phosphates is at z = 18 Å. Data points with z > 40 Å are the same color as z = 40 Å.

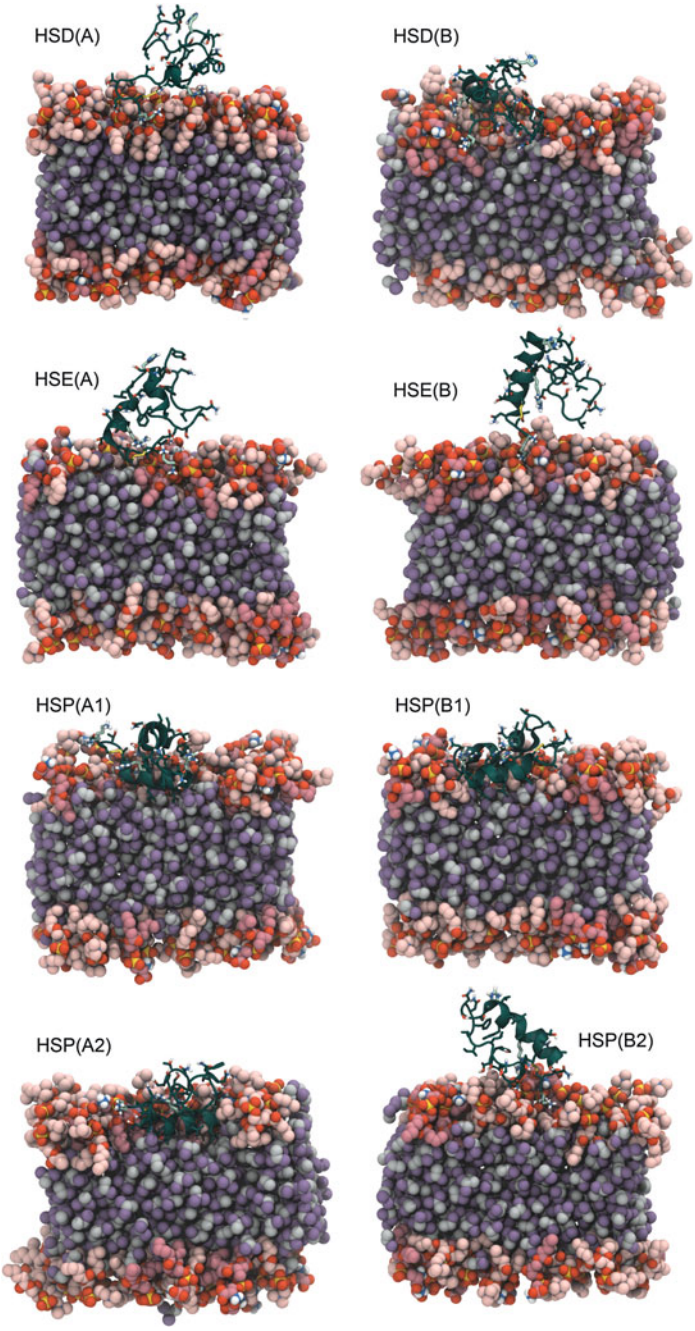
K.K. Skeby, *Computational Modelling of the Human Islet Amyloid Polypeptide*,
Springer Theses, DOI 10.1007/978-3-319-20040-8


Residue
30
20
10
0
HSD(A)
HSD(B)
HSE(A)
HSE(B)
HSP(A1)
HSP(A2)
HSP(B1)
HSP(B2)
0
25
50
75
100
Time (ns)
z-coordinate
40
20
0

Appendix B

Final snapshot of all $hIAPP_{1-37}$ peptides bound to the HMMM membrane. The lipids are shown in red colors, DCLE is shown in gray and purple, and the protein is shown in dark green.

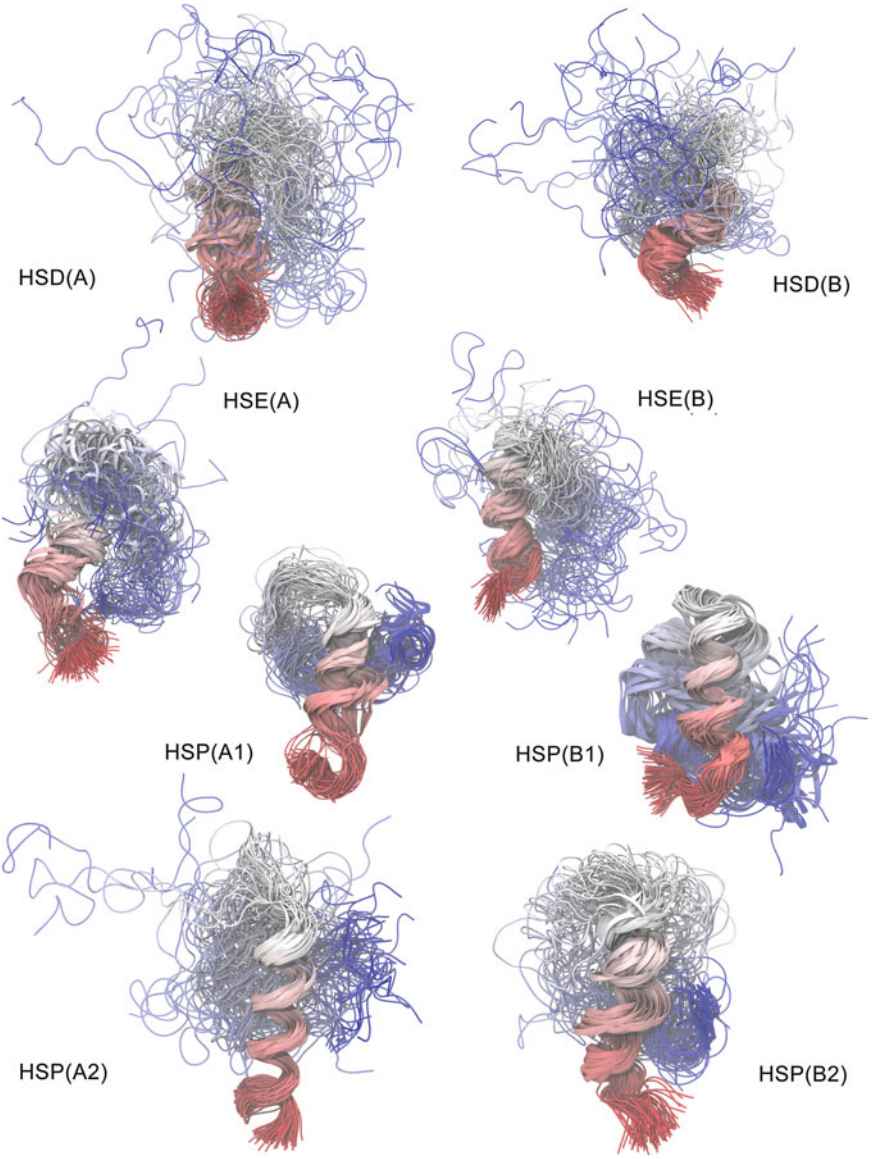
K.K. Skeby, *Computational Modelling of the Human Islet Amyloid Polypeptide*,
Springer Theses, DOI 10.1007/978-3-319-20040-8


HSD(A)
HSD(B)
HSE(A)
HSE(B)
HSP(A1)
HSP(B1)
HSP(A2)
HSP(B2)

Appendix C

Structural alignment of all $hIAPP_{1-37}$ peptides from simulations with a HMMM. All structures were aligned to the starting structure based on C_α atoms of the first 19 residues. The peptide is colored by sequence with the N-terminus colored red and the C-terminus colored blue.

K.K. Skeby, *Computational Modelling of the Human Islet Amyloid Polypeptide*,
Springer Theses, DOI 10.1007/978-3-319-20040-8


HSD(A)
HSD(B)
HSE(A)
HSE(B)
HSP(A1)
HSP(B1)
HSP(A2)
HSP(B2)