

**Investigation of the interaction between Alzheimer's
abeta peptide and aggregation inhibitors using
molecular simulations**

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der
RWTH Aachen University zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Olujide O. Olubiyi,

M.Sc. in Drug Discovery (London)

aus Efon-Alaaye, Nigeria

Berichter:

Jun.- Professor Dr. Birgit Strodel

Universitätsprofessor Ph.D. Paolo Carloni

Tag der mündlichen Prüfung: 08.07.2013

Diese Dissertation ist auf den Internetseiten der Hochschulbibliothek online verfügbar.

Acknowledgments

It is my delight to appreciate Jun.-Prof Dr Birgit Strodel, my advisor and under whose guidance I have been able to complete my doctoral training. I am in particular grateful for the great care with which she revised my thesis and for her readiness to explain scientific concepts encountered during this PhD studentship. My acknowledgments also go to advisors, Prof Paolo Carloni and Prof Arne Lüchow of the Faculty of Mathematics Computer Science and Natural Sciences of RWTH Aachen University.

I thank Prof Dieter Willbold and Dr Aileen Funke who direct the Alzheimer's research project at the Institute of Complex Systems 6, Forschungszentrum Jülich, with whom we held extensive discussions with respect to study design and harmonising their experimental and our computation data.

This work was funded by the German Research School for Simulation Sciences, Forschungszentrum Jülich, to whom I express my unreserved gratitude. And the Jülich Supercomputing Center for providing and maintaining the computing resources employed in this work. I also thank Prof. Michael G. Zagorski and Dr. Nikolas G. Sgourakis who provided us with the needed experimentally determined chemical shifts and J-coupling data, respectively. I acknowledge Dr Ioan Vancea and Dr Kumaran Baskaran for discussions and their assistance with certain aspects of programming. The fellow members of the Multiscale Modelling group, Chetan Poojari, Falk Hoffmann, Dr Bogdan Barz, Qinghua Liao, Phillip Kynast, and friends and colleagues Dr Hoa Do Quing, Dr Pallavi Thiagarajan-Rosenkranz, Dr Tobias Rosenkranz, Dr Amine Aladag, my landlord and landlady Mr and Mrs Peters

...it certainly is impossible to quantify how much you have contributed to my doctoral experience.

I appreciate my wonderful wife, Ifeoluwa Olubiyi for your support, friendship and encouragement. I appreciate my Dad and Mum, Mr and Mrs Olubiyi, you will always be my heroes. Also Mr and Dr(Mrs) Jemilugba and my brothers and sisters, I cannot adequately thank you for your prayers and support.

Above all, I am grateful to God for His mercies and goodness to me. Thank You!

Contents

Acknowledgments	3
1 Introduction	9
2 Molecular dynamics	13
2.1 Introduction	13
2.2 The molecular dynamics algorithm	15
2.2.1 Initial set of conditions—positions and velocities	16
2.2.2 Computing forces between interacting atoms	18
2.2.3 Integrating Newton’s equation of motion	19
2.2.4 Data analysis	21
2.3 The potential energy function	22
2.3.1 Bonded interactions	24
2.3.2 Nonbonded interactions	27
2.3.3 Force field parameterisation	30
3 Amyloid beta peptide	33
3.1 Introduction	33
3.2 The amyloid hypothesis	34
3.3 From structure to dysfunction: the mechanism of toxicity	37
3.3.1 A β peptide	38
	5

3.4	$A\beta$ conformational sampling	40
3.4.1	Initial configurations	40
3.4.2	Simulation details	42
3.4.3	The structure of $A\beta_{40}$	45
3.4.4	The structure of $A\beta_{42}$	49
3.4.5	Comparison of $A\beta_{40}$ and $A\beta_{42}$	51
3.4.6	Experimental validation using chemical shifts	54
3.5	The effect of pH on $A\beta$	59
3.5.1	Simulation setup, models and analyses	60
3.5.2	Effect of pH on $A\beta_{42}$ secondary structure	61
3.5.3	Role of histidine residues	64
3.6	Conclusion	66
3.7	Supplementary information	69
4	Alzheimer's disease: In search of treatment	71
4.1	AD drug discovery and mechanisms of toxicity	71
4.2	Current treatment strategies	72
4.2.1	$A\beta$ peptide production from APP	75
4.2.2	$A\beta$ aggregation inhibitors	76
	The $A\beta$ self-recognition unit: sequence	78
	The $A\beta$ self-recognition: hydrophobicity	79
	The $A\beta$ self-recognition: disruption of amyloid binding pattern	79
	The $A\beta$ self-recognition: stereochemistry	80
4.3	Aggregation inhibition with D3	81
5	Computational study of D3's anti-amyloid properties	83
5.1	Overview	83

5.2	Methodology	84
5.2.1	Peptide models	84
5.2.2	Generation and optimisation of A β complexed with D3	88
	Basin-hopping (BH) simulations	88
	A β monomer/D3	89
	A β pentamer/D3	90
5.2.3	Calculation of binding energies	91
5.2.4	Explicit solvent MD	92
5.3	Results and discussion	93
5.3.1	A β_{42} monomer/D3	93
5.3.2	A β_{42} pentamer/D3	98
5.3.3	Conclusions	102
6	D3 Modifications	105
6.1	Introduction	105
6.2	D3 sequence reshuffling: RD2 and progenies	106
6.2.1	Simulation protocol	108
6.2.2	Results and discussion	109
6.3	D3 mutations with natural and non-natural amino acids	117
6.3.1	Mutation studies with A β_{42} monomer	117
6.3.2	Mutation studies with A β_{42} pentamer	119
6.3.3	Results and discussion	121
6.3.4	Conclusion	123
6.4	Supplementary information: Force field parameterisation	126
7	Summary and Future perspectives	129
	Appendix I: List of abbreviations	133

7.1 List of abbreviations	134
Appendix II: Amino acids codes	135
7.2 Amino acids codes	135
Appendix II: Article reuse permission	136
7.3 Article reuse permission	136
Bibliography	137
Curriculum Vitae	159

Chapter 1

Introduction

Protein misfolding has long been known to constitute an important class of disease initiating factors. Of special significance in this group is Alzheimers disease (AD) in which the aggregation of misfolded small molecular weight amyloid β peptides ($A\beta$) triggers a host of biochemical anomalies that destroy brain neuronal processes. AD is recognised as the most common form of dementia in the elderly and characterised by a progressive loss of brain neural and synaptic functions. While a number of cellular processes have been implicated in AD, the overwhelming majority of research evidences support $A\beta$ as being central in Alzheimer's pathogenesis. Alzheimer's disease was first described by the German neuropathologist, Alois Alzheimers in 1906. However, in spite of the enormous efforts invested into AD research over the past century, there still remain crucial questions that are yet to be answered, in particular with respect to AD treatment. In other words, after 100 years of research, AD remains incurable. Even though a number of drugs exists for the relief of associated symptoms, there is at the moment no definitive treatment that halts the progressive neurological decline in AD patients. An explanation for this may be found in the peculiarities of the $A\beta$ peptide, the molecular target commonly targeted for drug development.

$A\beta$ is produced via post-translational cleavage of the transmembrane amyloid precur-

Introduction

sor protein (APP) followed by its release into the extracellular medium. Unlike most other protein drug targets however, $A\beta$ both lacks a regular three dimensional fold and possesses a significantly high aggregation propensity under physiological conditions. In healthy individuals, the production of $A\beta$ is delicately balanced by its clearance, and when this balance is thrown out of sync, as observed in AD, aggregating $A\beta$ initiates the neuronal and synaptic damage characterising AD. However, $A\beta$'s extremely high aggregation tendency renders most available experimental structure determination tools to an extent unable to determine its physiologically relevant conformations. Attempts to address this challenge includes the use of nonphysiological solubilising conditions, which at the same time compromises the usefulness of such models for $A\beta$ -directed drug discovery. This, however, is just one of the several dimensions of the challenge associated with the $A\beta$ peptide.

Also unlike most drug targets, the $A\beta$ peptide exhibits a high level of structural heterogeneity involving different truncated forms of the 39 to 43-residue $A\beta$ monomer, each exhibiting unique structural and toxicity properties. Added to this, $A\beta$ monomer aggregates into differently sized, conformationally and toxicologically diverse oligomers and fibrils. In other words, the number of structural states that $A\beta$ can adopt is so immense that it constitutes a potent challenge to therapy development. It is therefore obvious that the search for definitive treatment of AD will significantly benefit from an in-depth understanding of $A\beta$'s structural dynamics. This will necessarily depend on first finding a means of studying $A\beta$ conformations under physiological conditions. In this work, we have employed a number of molecular simulation approaches to address these challenges. We study and describe the structural dynamics of the two physiologically dominant $A\beta$ species— $A\beta_{40}$ and $A\beta_{42}$ monomers.

Using multiple molecular dynamics (MD) simulations on microsecond time scale, we studied the conformations of $A\beta_{40}$ and $A\beta_{42}$ in explicit water and under simulation conditions

Introduction

mimicing physiological conditions. To validate the obtained results, we employed chemical shift calculations which we compared with Nuclear Magnetic Resonance (NMR) chemical shifts, enabling us identify the force field that correctly models experimentally relevant $A\beta$ structural ensembles. An explanation is advanced for the differences observed in force fields, especially with respect to natively disordered peptides of which $A\beta$ is a type. We provide an atomistic explanation for certain important aspects of the $A\beta$ structure, such as the observed differences between $A\beta_{40}$ and $A\beta_{42}$ monomers, and the intramolecular factors that drive conformational behaviour of $A\beta$ under different pH conditions.

Our goal for analysing $A\beta$'s conformation is to obtain structural ensembles closely resembling the physiological state, for use in investigating $A\beta$'s interaction with aggregation inhibitors currently investigated in the group of Prof. Dr. Willbold (ICS-6 Forschungszentrum, Jülich). These aggregation inhibitors were discovered using the mirror image phage display technique. They showed promising anti-amyloid activities against $A\beta_{42}$ in *in vitro* tests and improved behaviour indices in transgenic mice. The inhibitors abolished $A\beta$'s toxicity in a dose-dependent manner, but their mechanism(s) of action, to a large extent, remains unknown. The aim of our study is thus to provide explanation for the mode of action for the observed anti-amyloid effects of the D-peptide inhibitors. Our *in silico* investigation is based on an hypothesis linking $A\beta$'s neurotoxicity to its structure, in particular the β -sheet content. The evidence for this came from an observation that the D-peptide- treated solutions of $A\beta$ lose their toxicity to cultured cell and at the same time their β -sheet contents. The structural effects obtained in our calculations thus serve, both as a means of approximating and comparing the D-peptides' anti-amyloid effects, as well as provide a way of validating the correctness of our results with respect to experimental observation. Our calculations reveal the possible mechanism of action of the D-peptides, and at the same time provide a detailed description of their effects on $A\beta$'s secondary structure. We also performed exhaustive point mutations on the D-peptides'

Introduction

sequences using both natural amino acids and some non-standard residues. The results suggest possible modifications that can be performed on the original D-peptides' amino acid sequences that can help modify their selectivity for different $A\beta$ oligomer sizes.

Chapter 2

Molecular dynamics

2.1 Introduction

”Given for one instant an intelligence which could comprehend all the forces by which nature is animated and the respective situation of the beings who compose it—an intelligence sufficiently vast to submit this data to analysis—it would embrace in the same formula the movements of the greatest bodies of the universe and those of the lightest atom; for it, nothing would be uncertain and the future, as the past, would be present in its eyes.”[1]

—Pierre-Simon Laplace writing on causal determinism (1812-1820)

Dynamics simply put describes the motion and at the same time the property of a particle or system of particles undergoing motion, which in biological systems represents an inseparable aspect of biomolecular processes. Fairly recently the study of the dynamic properties of biological systems from different perspectives (chemist, biologist and physicist’s) have assumed significant proportion, and that not unnecessarily given the critical roles played by dynamic processes in living systems. The microscopic scale characterising most biological processes in some cases has caused biomolecular dynamics to elude reckoning. For instance, biomolecular processes had for some time been approached as involving still

molecules (an example is the lock-and-key model used to describe the interaction of drug molecules to their biological receptors). The realisation that a description of biochemical functions remains incomplete without accounting for structural dynamics of constituent particles (e.g., atoms and molecules) is a major driving factor spurring the rapid growth in the study of molecular scale dynamics. Added to this are a plethora of powerful experimental tools such as X-ray crystallography, nuclear magnetic resonance, atomic force microscopy, circular dichroism (CD), electron microscopy and fluorescence assays used in biomolecular structure determination and limited aspects of structure dynamics, for instance, using solution Nuclear Magnetic Resonance (NMR) techniques. Molecular scale motions, however, traverse such broad time and distance scales (table 2.1) typically ranging from the fast femtosecond-scale vibrational motions featuring small amplitudes ($\approx 0.01 \text{ \AA}$) and energetics ($\approx 0.1 \text{ kcal/mol}$) to the much slower domain and global rearrangements occurring on hour scale characterised by much larger amplitudes ($\approx 100 \text{ \AA}$) and energies ($\approx 100 \text{ kcal/mol}$)[2]. Within the two extremes lie a host of intermediate scales each of which has specific biological relevance. These different scales do not exist in isolation relative to one another, instead they are interdependent—*coupled*—in that the smaller scale motions determine the dynamics of processes existing on larger scales in such a way that the largest dynamics (both distance- and time-wise) is indirectly dictated by the smallest motion [2].

The dynamic properties of biomolecular systems are often inseparable from biological

Molecular motion	Typical time-scale (s)	Amplitude (\AA)	Accessibility
Bond vibrations	$10^{-15} - 10^{-12}$	< 1	Fully
Loop motions	$10^{-9} - 10^{-6}$	1-5	Fully
Domain movement	$10^{-6} - 10^{-3}$	5-10	Fully
Protein folding and interaction	$10^{-3} - 10^4$	>10	Partly

Table 2.1: Time-scales of biological systems[2].

functions; however, such motions usually occur on scales that put them beyond the scope of many conventional experimental techniques. Fortunately many of these biologically relevant dynamics happen on time-scales that have become amenable to molecular scale

simulation methods, such as deterministic methods making use of empirical potentials in which either the Newtonian or the Lagrangian equation of motion is integrated over time to obtain the positions and momenta of particles (idealised atoms) of the system of interest[3]. Added to the recent tremendous progress in computer ware and algorithm development, the investigation of biomolecular dynamic processes have increasingly been brought within the domain of computational methods, thus providing a powerful and increasingly indispensable high resolution details to complement experimental investigation.

2.2 The molecular dynamics algorithm

A molecular dynamics simulation involves solving the Newton's second-order equation of motion for atom i having mass m_i experiencing a net force \mathbf{F}_i and typically within a system featuring other interacting particles.

$$\mathbf{F}_i = m_i \mathbf{a}_i \tag{2.1}$$

$$= m_i \ddot{\mathbf{r}}_i \tag{2.2}$$

$$= m_i \frac{\partial^2 \mathbf{r}_i}{\partial t^2} \tag{2.3}$$

Although the specific implementation of the code varies with parameterisation philosophy, the fundamental sequence of steps comprising an MD algorithm include: 1) the specification of the initial conditions, 2) the calculation of atomic forces, 3) configuration update involving the integration of the equation of motion, and 4) data (i.e., output) collection and analysis [4].

2.2.1 Initial set of conditions—positions and velocities

The molecular dynamics algorithm is a deterministic simulation method. Deterministic in the sense that starting with a specified set of initial system conditions, the method can be used to provide exact solutions for state properties such as positions and velocities (i.e., the phase space) at later time-points[5]. An MD simulation typically commences with the specification of the starting coordinates $\mathbf{R}(t = 0)$ at time zero, consisting of coordinates $\mathbf{r}_{1,2,3,\dots,N}$ for N atoms. The degrees of freedom represent the $6N$ -dimensional phase space corresponding to the x , y and z components of the positional vectors \mathbf{r}_i and the momentum vectors \mathbf{p}_i for each atom i . When completely specified for all atoms, the phase space vectors \mathbf{R} and \mathbf{P} represent together a microstate in the phase space[6].

With \mathbf{r}_i known, a value for the initial interaction $U(t = 0)$ (with velocities known) and the initial velocities $\mathbf{v}_{1,2,3,\dots,N}$ (may be optional in some MD codes; with $U(t = 0)$ known) of the N atoms can be calculated [4]. The starting coordinates, $\mathbf{R}(t = 0)$, for all atoms are usually obtained from experimental techniques such as X-ray crystallography and NMR, and in some cases from homology modelling[2]. Often times however, the initial structures require some pretreatment and structural refinement before they are suitable for use in molecular dynamics simulation. Hydrogen atoms, for instance, being too light for X-ray scattering often have their coordinates missing in X-ray structures; while the high structural flexibility characterising residues (both protein and nucleic acids) in loop segments hamper their determination by both NMR and crystallographic methods. In certain instances also, the missing residues result from deliberate removal as part of the X-ray crystallographic procedure. Fortunately however, a number of computational techniques exist to preprocess such starting coordinates and add in the missing atoms and residues even up to whole absent loop regions.

Once the starting atomic positions have been fully specified there arises next the need to specify the initial velocities which in most cases can not be obtained empirically. With

the system temperature specified (typically but not always 300 K), however, the initial velocities may be assigned according to a Maxwellian distribution, conducted initially in an *ad hoc* manner,

$$P(v) = \sqrt{\frac{m}{2\pi kT}} e^{-\frac{mv^2}{2kT}} \quad (2.4)$$

where $P(v)$ is the probability, m and v are respectively the atomic mass and velocity, while k is the Boltzmann constant. The initial system setup is however sometimes far from equilibrium partly resulting from the random manner in which velocities were assigned. This may also result from factors associated with the specific details of the method employed in generating the starting set of coordinates. For this reason further fine-tuning of the assigned velocities may be necessary. Starting with low temperature values where the system atoms feature wildly fluctuating kinetic energies (and thus velocities, equation (2.5)), [2] the temperature is slowly increased (in principle) until thermal equilibrium is achieved. From the equipartition theorem (equation (2.6)), it is possible to compute the system average kinetic energy and thus the velocities. The new velocities corresponding to thermal equilibrium can then be assigned. A second workaround involves iteratively assigning the velocities from the Maxwellian distribution until certain predefined convergence criteria are met.

The equipartition theorem allows to connect the system temperature T with the atomic velocities v_i :

$$E_{kin} = \frac{1}{2}mv^2 \quad (2.5)$$

$$T = \left(\frac{1}{k3N}\right) \sum_{i=1}^N \sum_{a=x,y,z} m_i v_{i,a}^2 \quad (2.6)$$

where $v_{i,a}$ refers to the velocity of atom i in either x , y or z direction.

2.2.2 Computing forces between interacting atoms

The level of detail being modelled will necessarily affect the manner in which the force calculation is conducted. This as well affects how atomic positions and momenta are subsequently computed. [3] For systems for which less-than-atomic resolution is acceptable some other equations of motion are solved to obtain forces: for instance the Lagrangian equation employed for studying whole domain motion, while solution of the Langevin equation is more appropriate for stochastic processes [3] In systems requiring atomistic detail description, numerical solutions of Newton's equation of motion are required for every atom and in each x, y and z direction. Solution to the Newtonian equation for atom i in equations (2.1) to (2.3) yields the atomic coordinates \mathbf{r}_i at time t . In order to obtain the atomic forces, from which positions and momenta are subsequently computed, the gradient of the atomic potential, $-\nabla_i U$, is taken with respect to the position vector of atom i :

$$\mathbf{F}_i = -\nabla_i U \quad (2.7)$$

$$= -\frac{\partial U(\mathbf{R})}{\partial \mathbf{r}_i} \quad (2.8)$$

That is, forces are only computed from the interaction part of the Hamiltonian H which is equivalent to the total energy and comprises the kinetic energy which is a function of the atomic momenta ($\mathbf{p}_i = m_i \dot{\mathbf{r}}_i$), as well as the potential (interaction) energy part which is a function of the atomic coordinate \mathbf{r}_i :

$$H(\mathbf{P}, \mathbf{R}) = E_{kin}(\mathbf{P}) + U(\mathbf{R}) \quad (2.9)$$

$$= \sum_i \left(\frac{1}{2} \right) m_i \dot{\mathbf{r}}_i^2(\mathbf{p}) + U(\mathbf{r}) \quad (2.10)$$

2.2.3 Integrating Newton's equation of motion

An MD simulation involves sampling the phase space by tracking the time-dependent evolution of the system via time discretisation and choice of an appropriate time-step to integrate the equation of motion. The molecular dynamics simulation of biomolecular systems studied in this work is based on the numerical solution of Newton's second law of motion for a system of interacting particles. Knowing the system coordinates, velocities and forces at time t , the task is thus to calculate subsequent positions of the interacting atoms following a step taken in time space. The resulting sequence of positions as a function of time are referred to as a trajectory. In principle, what is done is to calculate the atomic forces, using a force field described in 1.3.3, from which the subsequent velocities and positions can be obtained.

Molecular dynamics employs the finite-difference method for the numerical solution of the equation of motion[2]. To this end, the time space is discretized,

$$\tau = t_0 + n \sum \Delta t \quad (2.11)$$

where τ is the length of the simulation, t_0 is the initial time (often zero), Δt is the time-step, and n is the number of such steps taken during the simulation. In the finite-difference method, the atomic velocities and positions after taking a Δt time-step (i.e., at $t_{new} = t + \Delta t$) are calculated using the corresponding values before the time-step. In practice, however, this is only possible when the value chosen for Δt is sufficiently small (usually between 1–2 fs) in order to capture the fastest motion of the system, which is hydrogen bond vibration (table 2.1). Otherwise, the magnitude of accumulated error would become too large.

Three popular integration algorithms exist: the Verlet algorithm, the Leap-frog algorithm and the Velocity-Verlet algorithms with the last being the most accurate especially with

respect to velocity calculation, and the first being the least. They are all based on Taylor expansion of the atomic positions:

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \left(\frac{1}{2m}\right) \mathbf{F}(t)\Delta t^2 + \dots \quad (2.12)$$

The equation calculates the new coordinate at $(t + \Delta t)$ from the current coordinates (first term on the RHS), the current velocities (second term on the RHS), the current acceleration (third term on the RHS) and higher order terms. In practice the higher orders are usually ignored while evaluating the expansion. The Verlet algorithm is based on two such Taylor expansions, one for the future $(t + \Delta t)$, equation (2.13) and one for the past $(t - \Delta t)$, equation (2.14). By summing up the two expansions all terms with odd numbered orders $(\Delta t^1, \Delta t^3, \Delta t^5 \dots)$ cancel out, and the coordinates at $t + \Delta t$ can then be calculated from the current coordinates and forces, and the previous coordinates[2]:

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \left(\frac{1}{2m}\right) \mathbf{F}(t)\Delta t^2 + \mathcal{O}(\Delta t^3) \quad (2.13)$$

$$\mathbf{r}(t - \Delta t) = \mathbf{r}(t) - \mathbf{v}(t)\Delta t + \left(\frac{1}{2m}\right) \mathbf{F}(t)\Delta t^2 - \mathcal{O}(\Delta t^3) \quad (2.14)$$

$$\mathbf{r}(t + \Delta t) = 2\mathbf{r}(t) - \mathbf{r}(t - \Delta t) + \frac{\mathbf{F}(t)}{m}\Delta t^2 + \mathcal{O}(\Delta t^4) \quad (2.15)$$

$$\mathbf{v}(t) = \frac{\mathbf{r}(t + \Delta t) - \mathbf{r}(t - \Delta t)}{2\Delta t} + \mathcal{O}(\Delta t^2) \quad (2.16)$$

When equations (2.13) and (2.14) are subtracted the value of the velocity is obtained (equation (2.16)). The main advantages of the Verlet algorithm are its associated low computational cost deriving mostly from the use of single force calculations per time-step and its time-reversibility [2]. The disadvantages associated with the algorithm includes the fact that velocities at t can only be calculated when the positions $(\mathbf{r}(t + \Delta t))$ are known. Another disadvantage is the integration error for velocities being in the order of

$\mathcal{O}(\Delta t^2)$. This, however, has been reduced by the more recent modifications to the Verlet algorithm, such as the Leap-frog algorithm, which evaluates velocities at half time-steps, and the Velocity-Verlet algorithms. A detailed description of the two algorithms can be found in reference [2].

2.2.4 Data analysis

The classical equation of motion as described above is solved at every time-step for the desired length of simulation (τ), yielding a trajectory. The coordinates, velocities, potential energy, pressure, etc can then be used to compute the desired time-averaged equilibrium properties [7]. Examples include the root-mean-squared deviation (RMSD) which can either be averaged over selected coordinates to temporally trace the structural fluctuation, or over time in which case the relative fluctuation of the different subsets (e.g., a loop region) of the system can be determined. Generally, analysis of MD-generated trajectories typically involves calculating the time average of a desired quantity A over the length of the simulation T :

$$\langle A(\mathbf{r}, \mathbf{p}) \rangle_{time} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{t=0}^T A(\mathbf{r}(t), \mathbf{p}(t)) dt \quad (2.17)$$

In other words, $\langle A(\mathbf{r}, \mathbf{p}) \rangle_{time}$ represents the average, which is a function of the position \mathbf{r} and momentum \mathbf{p} , calculated for a single point over the several time steps making up the total simulation time. MD-derived average contrasts with what obtains at the macroscopic level, where an ensemble average (equation (2.18)) represents several multiple points existing at the same time:

$$\langle A(\mathbf{r}, \mathbf{p}) \rangle_{ensemble} = \iint A(\mathbf{r}, \mathbf{p}) \rho(\mathbf{r}, \mathbf{p}) dr dp \quad (2.18)$$

where $\rho(\mathbf{r}, \mathbf{p})$ is the probability density. According to the ergodic hypothesis [8], at the limit of an infinitely long MD trajectory, $\langle A(\mathbf{r}, \mathbf{p}) \rangle_{time}$ converges to $\langle A(\mathbf{r}, \mathbf{p}) \rangle_{ensemble}$. Rather than conducting infinite MD sampling, in practice, what is done is to generate trajectories that are long enough for the simulated system to have equilibrated. Subsequent trajectory analysis then involves calculating the time averages of the desired quantities ($\langle A(\mathbf{r}, \mathbf{p}) \rangle_{time}$) and approximating these as representing the ensemble average ($\langle A(\mathbf{r}, \mathbf{p}) \rangle_{ensemble}$). To a large extent the validity of this assumption will depend on the degree to which the phase space has been sampled. And generally, the longer the simulation the larger the part of sampled phase space. There are different types of properties that can be calculated from an MD trajectory, but the analyses eventually performed will depend on the question the simulation seeks to address. The properties calculated in the present work are discussed in the relevant sections.

2.3 The potential energy function

The potential energy function represents a set of mathematical expressions that allow a atomic system's potential energy to be obtained from its 3-dimensional structure [2]. The mathematical models employed efficiently capture the various physicochemical interactions contained in the system[2]. Such potential energy functions are often typically designed for a defined group of molecular systems such as proteins, and are commonly referred to as force field. Individual force fields are developed with various objectives in mind, but practical applicability often demands that the employed forms of the mathematical expressions aim at achieving a balance between accuracy, implementation efficiency and practical applicability [2]. It can as well be argued that the emergence of force field methods already represents a major compromise between model accuracy and applicability compared with quantum mechanical approaches in which electrons are explicitly

treated. The advantage of accuracy obtainable in quantum mechanical (QM) treatments is, however, at the same time a major limitation to their relevance in investigating most biomolecular systems of interest. This is mostly because the high-level quantum chemical calculations employed in QM treatment are in reality only able to study systems containing less than 100 atoms and quickly requiring beyond available computational resources for larger systems. Nonetheless, force field development has immensely benefited from quantum mechanics for the derivation of force field parameters. Furthermore the Born-Oppenheimer approximation which *decouples* the nuclear and the electronic wavefunctions serves as an important foundation for the classical treatment employed in force field methods.

The potential energy, written as $U(\mathbf{R})$ in equation (2.8) specifies the interaction energies between the atoms composing the system but can also includes contributions to the energy function resulting from the use of special restraining forces during an MD protocol. An instance is the use of restraints in biasing defined inter-atomic distances or angles[9]. These *special* interactions must then be accounted for as part of the potential energy while computing forces.

The mathematical description of the potential energy function comprises two core ingredients: bonding terms which specify the energy contributed by covalent bond terms, and the nonbonding terms specifying contributions between atoms that are not via covalent connections:

$$U = U_b + U_{nb} \quad (2.19)$$

where both U_b and U_{nb} are in turn composed of separate terms the specific details of which depend on the adopted functional form.

2.3.1 Bonded interactions

The U_b term is obtained as a sum over all covalent bond lengths, bond angles, bond torsions (or proper dihedral angles, often called dihedrals for short), and an improper dihedral term:

$$U_b = U_{Bonds} + U_{Angles} + U_{Torsions} + U_{Improper} \quad (2.20)$$

Based on Fig. 2.1 an example for each of these four contributions is given:

- The **covalent bond** formed between atoms i and j : e.g., the **C—CA** bond
- The **bond angle** formed by atoms i , j and k and centred around j : e.g., the **C—CA—N** angle
- The **torsional angle** involving successive atoms i , j , k , l : e.g., the angle defined around the bond **CA—N** between the two planes (shown as gray triangles) involving atoms **C**, **CA**, **N** and atoms **CA**, **N**, **H**.
- The **improper dihedral** involving four atoms i , j , k , l where j , k , l are all directly bound to i . The improper torsion is also known as the out-of-plane bending term. It is especially useful in cases where it is necessary to enforce a particular configuration, such as planarity or chirality and tetrahedral centres[9]. In figure 2.1 the planarity of the ring is enforced by an improper dihedral angle defined as the angle between the planes spanned by **Cb**, **CD1**, **CD2** and **Cb**, **CG**, **CD1**

The separate bonded terms are often treated either as harmonic or trigonometric functions. For the harmonic treatment, the bond length, bond angle and improper dihedral terms are calculated respectively as sum over the total number of covalent bonds ($i_b = 1, 2, 3, \dots, N_b$), bond angles ($i_\theta = 1, 2, 3, \dots, N_\theta$), and improper dihedral angles

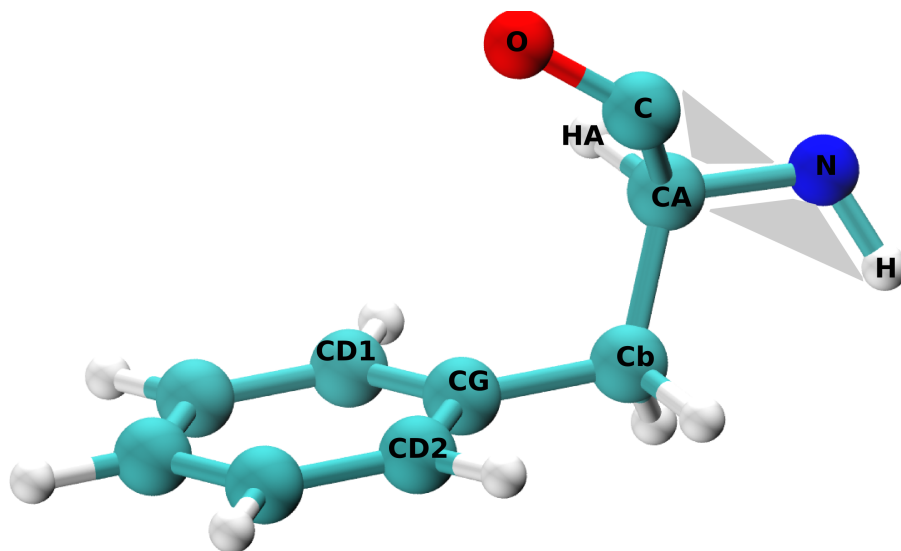


Figure 2.1: Phenylalanine for the illustration of bonded interactions.

($i_\omega = 1, 2, 3, \dots, N_\omega$) according to the Hookean spring equation:

$$U_\alpha = \sum_{i_\alpha=1}^{N_\alpha} \frac{1}{2} K_\alpha [\alpha - \alpha_{eq}]^2 \quad (2.21)$$

where α is any of b , θ and ω , and K_α represents the harmonic force constant for the motion in question, and α_{eq} is the equilibrium value of the bond or (improper dihedral) angle. The force constants, usually obtained from spectroscopic measurements of related compounds or from QM calculations, [2, 9] are typically defined for specific types of bonded interactions. These are then supplied as parameters in the mathematical formulation of the potential function. As an illustration, the value of K_b for a $C-sp^3-C-sp^3$ (σ -bond) bond is expected to be different from the value for either of a $C-sp^2-C-sp^2$ (π -bond) or a $C-sp^3-N-sp^3$ bond. The equilibrium value α_{eq} is, like the force constant, supplied as input parameter for different bond types (e.g., $C-sp^3-C-sp^3$ versus $C-sp^2-C-sp^2$), bond angle types (e.g. a tetragonal centred around sp^3-C versus a trigonal centred around sp^2-C), and improper angles. In addition, the equilibrium values for both bonds and angles (α_{eq}) can also be obtained from experimental measurements (X-ray diffracton in particular)[9] and QM calculations.

An harmonic function is generally sufficient in capturing the properties of covalent connections which generally oscillate around the equilibrium values under commonly encountered conditions. It is however important to note that the adopted functional forms of each of these models also crucially depend on the need to simplify and reduce the number of computational operations that need to be evaluated while computing the potential. For instance, the new generation GROMOS force fields employ a non-harmonic function for the bonds:

$$U_b = \sum_{i_b=1}^{N_b} \frac{1}{4} K_{b,anhar} [b^2 - b_{eq}^2]^2 \quad (2.22)$$

This form allows a reduction in the number of mathematical operations requiring square roots evaluation and by so doing simplifies the calculations. Similarly, a cosine expression is employed (GROMOS) for the bond angles (equation (2.23)), evaluating $\cos \theta$ rather than θ , which reportedly improves *numerical stability* and allows the operations to be performed more efficiently [9].

$$U_\theta = \sum_{i_\theta=1}^{N_\theta} \frac{1}{2} K_\theta [\cos \theta - \cos \theta_{eq}]^2 \quad (2.23)$$

In the case of the torsion angle a different function capable of adequately capturing the unique features of the angle is required. For these, the cosine function is used either in a Fourier expansion (equation (2.24)) or Ryckaert-Bellemans formulation (equation 2.25) which more satisfactorily (compared with an harmonic function) models the periodic nature of a torsion angle moving through angles ranging from 0° – 360° (or -180° depending on adopted the angular definition):

$$U_\phi = \sum_{i_\phi=1}^{N_\phi} \sum_{m=1} K_{\phi,m} [1 + \cos(m\phi - \delta)] \quad (2.24)$$

$$U_{\phi,RB} = \sum_{i_\phi=1}^{N_\phi} \sum_{m=0}^5 C_m [\cos(\phi - 180^\circ)]^m \quad (2.25)$$

The cosine function, in addition, better models the uniquely small energy barriers that exist between multiple stable states, which partly result from the influence of nonbonded interactions on the torsional potential [4, 11]. As a result of the dependence of the torsional potential on the nonbonded interaction, the Ryckaert-Bellemans function ignores the through-space interaction between the two outer atoms (i.e, 1-4 interaction) during nonbonded interaction calculation while in the implementation in GROMACS, for instance, special dispersion parameters are employed for the two atoms.

2.3.2 Nonbonded interactions

The nonbonded component of the interactions involve pairwise interactions transmitted via *space* and not via the covalent network of bonds and angles[10]. There are two types commonly included in most force fields: the electrostatic interaction resulting from explicit charges (e.g., attraction potential between oppositely charged aspartate and lysine sidechains) and the van der Waals contribution (vdW, dispersion) resulting from uneven and instantaneous electron distribution:

$$U_{nb} = U_{vdW} + U_{Electrostatic} \quad (2.26)$$

In the calculation of nonbonded interactions covalently bonded atoms are generally excluded. This may take the form of an exclusion list of all atoms connected via 1 or 2 bonds for which nonbonded calculations are not performed.

Compared with the electrostatic interaction (attractive and repulsive) which is a long-range interaction as it scales with $\frac{1}{r}$ where r is the distance between the interacting atoms, the vdW interaction is a short-range interaction. The van der Waals potential becomes attractive as the interaction distance decreases to an intermediate value, resulting from an electron redistribution also known as induced dipoles. On further decreasing

the distance, electron clouds from different atoms get too close and the potential turns repulsive[11]. This unique distance-dependent behaviour can be conceived of as arising from the electrons *getting in each other's way* [12] at distances smaller than the sum of the interacting atomic radii where the interaction is dominated by high nuclear and electron repulsion. The employed functional forms thus seek to adequately capture three core aspects of the potential. That is: repulsive at short distances, slightly attractive at intermediate distances, and vanishing at large distances. The different forms employed usually only differ in the repulsive component. While the Buckingham [13] and the Morse potentials [14] yield better results, the Lennard-Jones (LJ) potential is more commonly preferred because of its associated lower computational cost:

$$U_{vdW} = U_{ij} - \frac{C_{ij}}{r_{ij}^6} \quad (2.27)$$

$$U_{Buckingham} = A_{ij} \exp^{B_{ij}r_{ij}} - \frac{C_{ij}}{r_{ij}^6} \quad (2.28)$$

$$U_{Morse} = D_{ij} \left[1 - \exp\left(-\sqrt{\frac{k}{2D_{ij}}}r_{ij}\right) \right] \quad (2.29)$$

$$U_{LJ} = 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^6 \right] \quad (2.30)$$

where $C_{ij}(vdW) \neq C_{ij}(Buckingham)$. Equation (2.27) captures the essential ingredients for modelling van der Waals interactions. The choice of exponential functions by the Buckingham (2.28) and Morse (2.29) functional forms more accurately capture the distance dependence than the LJ function (2.30). The constants A_{ij} , B_{ij} , C_{ij} , and D_{ij} , are parameters defining the interaction of atoms i and j . The repulsive part defined in the Lennard-Jones formulation by A_{ij}/r_{ij}^{12} is chosen for no obscured reasons than to ensure that it vanishes faster than the attractive part (dispersion) as r_{ij} increases [11]. The choice of 12 for the exponent is thus mainly for the ease of computation ($r_{ij}^{12} = (r_{ij}^6)^2$) for which reason other force fields can be found using values other than 12. A_{ij} and B_{ij}

are values depending on the atomic radii of the atoms involved in interaction and are supplied as parameters for the particular atom types [9] being treated. Equation (2.30) shows the functional form of the LJ potential employed by MD codes like GROMACS. The parameter σ_{ij} is the value of r_{ij} at the minimum energy while ϵ_{ij} is the depth of the potential minimum energy.

The electrostatic interaction of charged groups (atoms and group of atoms) within the system is usually modelled by the Coulombic function where partial charges are assigned to the concerned atoms:

$$U_{Coul.} = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{\epsilon_r r_{ij}} \quad (2.31)$$

where q_i and q_j are respectively the partial charges of atoms i and j , while ϵ_r and ϵ_0 are respectively the relative dielectric constant and a distance dependence of the ϵ_r included to include the screening effect of the medium. While functions including polarisation effects and higher multipoles have proved more accurate at modelling the electrostatic potential, such functions often come with high computation cost which has continued to make the Coulombic model an attractive choice. For instance, evaluation of all pairwise electrostatic interactions is associated with a $\mathcal{O}(N^2)$ complexity. In practice, schemes using spherical cutoffs such as the switching and shifting functions to a significant degree reduce the discontinuity in energies and forces resulting from the use of simple truncation, and have been used extensively in computing nonbonded interactions. A better treatment of the long-range component of the electrostatic interaction in biomolecular simulation is achieved with the particle-mesh Ewald method (PME), an Ewald sum methods which computes electrostatic interaction using a lattice with periodic boundary conditions (PBC). With the PME the evaluation complexity is reduced to $\mathcal{O}(N \log N)$.

The overall potential energy function for the system can therefore be written in the form,

$$\begin{aligned}
 U = & \sum_{i_b=1}^{N_b} \frac{1}{2} K_b [b - b_{eq}]^2 + \sum_{i_\theta=1}^{N_\theta} \frac{1}{2} K_\theta [\theta - \theta_{eq}]^2 \\
 & + \sum_{i_\phi=1}^{N_\phi} \sum_{m=1} K_{\phi,m} [1 + \cos(m\phi - \delta)] + \sum_{i_\omega=1}^{N_\omega} \frac{1}{2} K_\omega [\omega - \omega_{eq}]^2 \\
 & + \sum_{i < j}^N 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \sum_{i < j}^N \left[\frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{\epsilon_r r_{ij}} \right]
 \end{aligned} \tag{2.32}$$

The discussion of the implementation details and the specific mathematical treatment employed by different force fields and MD codes is beyond the scope of this thesis, which aims at providing an overview, though incomplete but sufficiently broad to highlight the most important contributory terms typically employed in a force field.

The actual force calculation is performed by taking the first derivative of the calculated potential energy with respect to the atomic coordinate. As an illustration, the calculation of forces coming from the nonbonded terms of the potential can given as[4]:

$$\mathbf{F}_{nb}(r_{ij}) = \left\{ \left(\frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{\epsilon_r r_{ij}^2} \right) + 4\epsilon_{ij} \left[12 \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{13} - 6 \left(\frac{\sigma_{ij}}{r_{ij}} \right)^7 \right] \right\} \frac{\mathbf{r}_{ij}}{r_{ij}} \tag{2.33}$$

2.3.3 Force field parameterisation

Currently employed force fields usually have all the necessary parameters required to perform molecular dynamics simulations on proteins, nucleic acids, lipids, and carbohydrates which represent the most commonly encountered biomolecular systems. For this reason force field parameter development focuses primarily on the building blocks of these macromolecules as standard entries. From time to time however, there arises the need to study nonstandard groups such as small molecule ligands, chemically derivitised forms of standard groups of interest, or non-biological polymer systems. The nontrivial process

of obtaining force field parameters for such nonstandard groups is what is referred to as force field parameterisation, and it represents an advanced field in molecular simulation. The parameterisation effort commences with first identifying the existing force field description to which the new parameters should agree. This ensures compatibility with the existing potential description and eliminates unreasonable behaviour. Subsequently parameters that already exist in the force field having the same (or in certain cases similar) atomic properties, bond hybridisation, and chemical environment as existing in the nonstandard molecules should be identified. This approach is known as parameterisation by analogy and benefits from a good chemical intuition and knowledge of organic chemistry. In a great number of cases the required parameters already exist but not necessarily in a single force field chemical entry. As illustration, the phenyl ring parameters for phenylglycine, a nonstandard amino acid residue, maybe wholly sourced from the related phenylalanine, but not from tyrosine because the hydroxyl group will have altered the partial charges on the aromatic ring. In some cases, not all of the required parameters that are available in the force field, making (high-level) QM calculations necessary for parameter generation. This typically involves geometry optimisation at the appropriate level of theory, followed by potential energy scanning along the internal modes of interest, such as bonds or bond angles. The resulting energy profiles can then be used to determine the equilibrium values and force constants via fitting to the specific potential energy function employed by the force field (equation (2.21)).

Chapter 3

Amyloid beta peptide

3.1 Introduction

Amyloid beta ($A\beta$) polypeptide, a 4-kDa metallopeptide and principal component of amyloid plaques found in the brains of Alzheimer's patients, derives its name *amyloid* from an initial description of it as a carbohydrate product, and *beta* to indicate its secondary structural signature of β -pleated sheets. The self-aggregating peptide is an enzymatic product of two secretases—the β and the γ secretases—acting on the parent amyloid precursor peptide, a type-I transmembrane protein. The $A\beta$ peptide is perhaps the most important peptide involved in neurodegeneration, a reputation earned by its role in the pathogenesis of Alzheimer's disease. In 1987, the peptide was identified by Kang *et al.* as the principal component of the amyloid plaque deposited in the brain of AD patients [15]. It was later accepted widely as the single most important hallmark of the disease. At that time the peptide was referred to as A4 protein. This, however, was not its first discovery: four years earlier Glenner and Wong had reported the same peptide as the main constituent of a vascular amyloid system[16]. Its distribution, however, is not exclusively limited to the compact or diffuse senile plaques typically found in the hippocampal and neocortical compactments of the brain, the peptide has also been isolated from the vascular

tissues and generally in the central nervous system especially in the cerebrospinal fluid [17–19].

3.2 The amyloid hypothesis

Initially different explanations for the development of Alzheimer’s disease were put forward but it was only the amyloid-beta based hypothesis that best accounted for the different pathological and biochemical events characterising the disease[20–26]. In the hypothesis, elevated levels of the peptide in the cerebral cortex is considered as the trigger factor for the development of AD pathology [27]. Increase in cerebral $A\beta$ peptide results from a disturbance in the fine balance between the generation and clearance of the peptide. In other words, any event capable of altering $A\beta$ peptide production or its clearance is potentially a candidate in the initiation of such disturbance. And in fact, both aspects are considered in attempts to design therapeutic agents aimed at rectifying the $A\beta$ peptide concentration anomaly in AD. The disruption of $A\beta$ peptide level regulation in the cerebral system (mainly though not exclusively) soon sets off a host of biochemical events which ultimately results in the clinically observed systems typical of AD—cognition and memory impairment[28].

The amyloid hypothesis has been widely accepted in explaining the specific details associated with AD mostly because it represents the model that most accurately fits the different biochemical and functional changes accompanying the development of the disease. Some of the strongest evidence supporting the validity of the amyloid hypothesis have come from $A\beta$ toxicity studies, from the correlation of genetic mutations in the APP and presenilin-1 and 2 genes, and from transgenic mice experiments. $A\beta$ peptide was shown to possess an aggregation-dependent neurotoxicity in cell culture studies, believed to be an indication of the peptides toxicity in AD [29, 30] The mutations associated with Alzheimer’s disease pathology involve three main peptide systems—the amyloid precursor

peptide, the presenilin-1 (chromosome 14) and presenilin-2 (chromosome 1) and finally the apolipoprotein E lipoprotein (chromosome 19). Mutations in the amyloid precursor peptide gene, though representing only a fraction of familial Alzheimer's disease, have been observed to involve sequence regions within the APP identified to correspond to the $A\beta$ peptide. An example is the London mutation in which a point mutation of codon 717 leads to an increased generation of $A\beta_{42}$, an alloform of the amyloid peptide associated with higher aggregation and neurotoxicity but normally produced in much lower titres than what is seen in the London mutation of familial Alzheimer's disease [31,32]. Another strong evidence implicating $A\beta$ as the main player in AD development comes from Down Syndrome (DS) in which people with DS invariably develop AD if they live up to the 40 years [33]. The link between the two conditions arises from a triplication in DS of the Chromosome 21 encoding the APP from which $A\beta$ peptide is produced.

The strong correlation of amino acid substitutions involving the $A\beta$ peptide sequence with the pathology of Alzheimer's disease further strongly establishes the causative role of the peptide in disease development. Of special importance are point mutations involving glutamic and aspartic acids respectively at positions 22 and 23 of the $A\beta$ peptide sequence. Both assembly structure and clinical presentation of AD are strongly affected by this group of mutations, which on a general note are associated with a faster aggregation kinetics and more aggressive symptom presentation. The Arctic mutation, involving a change of Glu22 to a glycine (E22G) was reported to significantly increase $A\beta$ prefibrillar structures and is associated with an early-onset pathology with a symptom complex identical to the later-onset disease [34–37]. This is unlike the other mutations involving Glu22 and Asp23 whose distinct structural and pathological signatures separate them from the wild-type disease. Mutating Glu22 to either a lysine (E22K) as seen in the Italian mutation or to a glutamine (E22Q) of the Dutch mutation results in a symptom complex involving significant cerebral haemorrhage[38, 39]. Either the haemorrhage in

the Dutch mutation, or the reported cerebral amyloid angiopathy [40] causes episodes of cerebrovascular accidents (i.e. ischaemic or haemorrhagic stroke) with the attendant loss of function of the area of the brain so affected. Both mutations are associated with the formation of diffuse plaques and intensely increased polymerization rates; for instance, the Italian mutant has been reported to be associated with twice as fast aggregation kinetic compared with the wild-type $A\beta$ peptide [39, 41]. Lastly, diffuse plaque formation as well as nonhaemorrhagic cerebral amyloid angiopathy have similarly been observed in the Iowa mutation with aspartic acid 23 changed to asparagine (D23N)[38, 42, 43].

Presenilin-1 and 2 are part of the γ secretase enzyme complex that cleaves the APP in regions corresponding to the $A\beta$ C-terminus. While mutations involving the $\epsilon 4$ variants of apolipoprotein E is associated with increased amyloid deposition in late-onset Alzheimer's disease, mutations in the presenilin-1 and 2 genes have been observed to shift the C-terminal processing activities of γ -secretase in favour of $A\beta_{42}$ over other less toxic alloforms in early-onset AD[44, 45]. Lastly, features akin to human AD such as gliosis, loss of cerebral synapses and behavioural dysfunction have also been observed in transgenic mice expressing the APP gene, but much less in mice lacking the apoE gene[45–47].

Criticism: At the turn of the century, when the amyloid hypothesis was just proposed for explaining the pathogenesis of AD, the insoluble deposits were believed to be the principal causative of the neurodegeneration associated with the disease. However, there is a noticeable absence of the direct proof that amyloid plaque deposition is the major factor that culminates in AD development. [48] And crucially, the observation that neurodegeneration occurs prior to the detection of amyloid plaques and the detection of the deposit in nonsymptomatic or healthy persons represents a major disconnect between the amyloid plaque-based hypothesis and clinical reality of the disease [49]. The available body of evidence for the disease better points at the water-soluble prefibrillar aggregates as the clinically relevant neurotoxic species rather than amyloid plaques whose deposition fail to

sufficiently correlate with AD pathogenicity and progression[50]. This then led to a major revision of the hypothesis to account for the observation that the water soluble oligomers of $A\beta$ are the main neurotoxic component[25]. The soluble $A\beta$ oligomers not only demonstrate significantly higher toxicities in cultured cells than the insoluble plaques, they also reveal toxicity patterns that correlate better with AD pathology[51, 52].

3.3 From structure to dysfunction: the mechanism of toxicity

Following enzymatic cleavage from APP the generated small molecular weight amyloid- β peptide undergoes a set of structural changes involving both intramolecular and intermolecular transitions, which two kinds of structural remodelling do not necessarily proceed independently. Prior to release from APP, the $A\beta$ peptide adopts a predominantly helical state inside the hydrophobic transmembrane environment, and after cleavage and release into the aqueous extracellular medium it dynamically undergoes structural conversions in what is commonly believed to be an unstructured state in the monomer. At this particular level attempts to unambiguously characterize the structural states of the peptide have often been less than successful, largely a consequence of the rapid intrinsic aggregation properties of the peptides under physiological conditions. While this has been partly circumvented by employing artificial and nonphysiological experimental conditions, including the use of nonpolar solvents, the obtained results cannot be safely extrapolated to physiologic conditions. And this is where computational experiments have amply filled in the gap.

3.3.1 $A\beta$ peptide

Based on intrinsic toxicity $A\beta_{42}$ represents the most clinically important alloform of the proteolytic products of β and γ secretases. The extra two residues in the C-terminus confers subtle but important structural and toxicity properties on the peptide compared to the most abundant but less amyloidogenic $A\beta_{40}$ variant. The primary sequence of the full length $A\beta$ peptide features neighbouring subregions with sharply differing polar/solubility properties, with the N-terminal portion largely composed of polar and hydrophilic residues and the C-terminus lined up with hydrophobic residues. Within the mid-section is a short hydrophobic stretch referred to as the central hydrophobic core (CHC). The very detail of the ordering of the $A\beta$ peptide sequence represents a vital driving force in the peptides structural identity and by extension, its toxicity:



The unique structural features of the peptide partly derives from its sequence constitution composed of six negative, three positive, and three histidine residues mostly in the N-terminal half, and a predominantly hydrophobic C-terminal half. The peptide has an isoelectric pH of 5.3 and under most experimental conditions has a negative charge surplus at physiological pH values. Following cleavage by β - and γ -secretases and release into the aqueous extracellular compartment, the peptide undergoes a sequence of conformational transitions from being largely unstructured, to rapidly interconverting secondary structural units, and assembly into water-soluble structurally heterogeneous oligomers, and ultimately to β -sheet-rich protofibrils and amyloid fibrils. The level of β -sheet formed appears to be a characteristic feature that increases with aggregate size and stage.

The monomer is widely considered as an intrinsically disordered peptide. But rather than being altogether unstructured, studies have shown that the monomers sample different features in different parts of the sequence. For instance, studies by Yang and Teplow revealed the two major variants as possessing a "unique statistical coil" bearing identi-

fiable secondary structure units, which are separated by turn structures[53]. NMR and MD studies aimed at distinguishing between the two peptides revealed a higher rigidity in $A\beta_{42}$'s C-terminus[54, 55]. This was used to explain the higher amyloidogenicity and neurotoxic potencies of $A\beta_{42}$ compared with the $A\beta_{40}$ [54, 55]. Other studies comparing $A\beta_{40}$ and $A\beta_{42}$ have also revealed distinctions in their structures and aggregation properties[56, 57]. Nuclear Overhauser effect (NOE) measurements indicate both peptides as exhibiting high conformational interconversion at the pico- to nanosecond time-scale[54]. The NOE values for $A\beta_{42}$ C-terminus, however, suggested a higher rigidity compared with $A\beta_{40}$ [54]. This was believed to enforce some sort of conformational restriction that keeps $A\beta_{42}$'s C-terminal segment in a β -strand competent conformation. This in addition serves an internal seeding function that translates into higher aggregation propensities in the longer $A\beta$ alloform. For instance, Yang and Teplov reported an increased C-terminal residue contact with the CHC of $A\beta_{42}$, which effect serves to better stabilize the β -sheet structure more than in the shorter peptide, where the CHC would rather interact with the N-terminus. Apart from the difference in primary amino acid sequences of the two extra hydrophobic residues, Urbanc and co-workers identified a turn structure at sequence positions Gly37-Gly38 in $A\beta_{42}$ as another feature distinguishing it from the shorter peptide[56, 57]. The Gly37-Gly38 turn has been suggested as an explanation for the relative preference of $A\beta_{42}$ for pentameric and hexameric oligomers, compared with the dimeric assemblies more preferentially favored by $A\beta_{40}$ [56, 58, 59]. A β -strand located at positions Ala2-Phe4 has also only been exclusively identified in the $A\beta_{40}$ [56].

The aggregation of $A\beta$ monomers into toxic β -sheet-rich oligomer structures is believed to depend to a significant extent on the sampled conformational state of the monomer and factors influencing it. Metal ions, such as Cu^{2+} , Zn^{2+} , and Fe^{2+} , lipid membranes, the presence of preformed oligomers, hydrogen ion homeostasis as well as sample preparation conditions have been found to strongly influence both folding and the aggregation

kinetics[60–65]. A weakly acidic pH (pH range 3.5-6.5) was suggested to be an important factor promoting amyloid aggregation[60]. The protonation state of A β histidine residues (His6, His13, and His14) is strongly influenced by pH changes with a change in net molecular charge from -3 to 0 associated with a change from neutral to slightly acidic pH values. It was found that pH-controlled structural changes can also be attributed to metal ion-binding and to changes in the electrostatic charge distribution of the molecule[62]. For instance, the binding of Cu $^{2+}$ by A β_{28} was observed to be pH-dependent, with the divalent metal ion inducing aggregation only at pH values lower than the physiological pH[61]. Whether it is by promoting metal coordination or by intrinsic redistribution of electrostatic charges, acidic pH values are likely to promote amyloid aggregation by decreasing the hydrophilicity of the N-terminal portion of A β . Given the slightly acidic nature of accompanying inflammatory response in AD,[66] acidic pH-facilitated aggregation is likely to be of importance in disease development. It has been shown that brain from patients who die from AD are more acidic than brains from patients who die suddenly with no brain disease[67].

3.4 A β conformational sampling

The current section, written by me and corrected by Jun.-Prof Dr. Birgit Strodel, has been published in the Journal of Physical Chemistry B[68] and has been cited here.

3.4.1 Initial configurations

The starting structures of the monomeric A β peptides, A β_{40} and A β_{42} were taken from the RCSB Protein Data Bank (www.rcsb.org). The structure of A β_{40} (figure (3.1(a)))

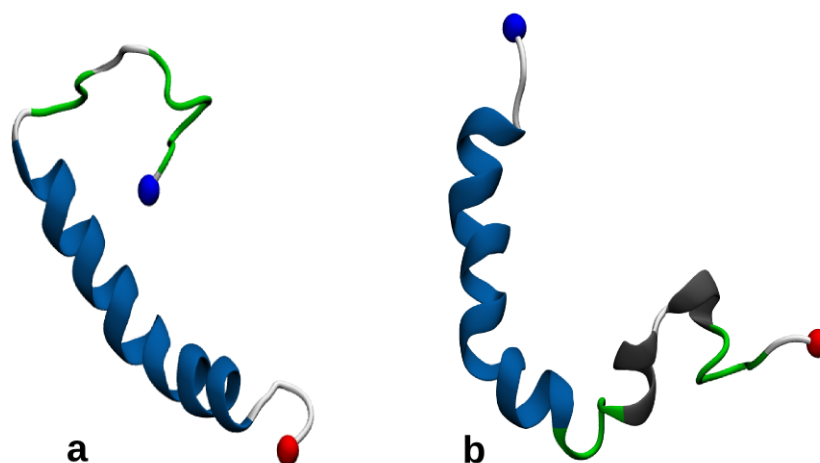


Figure 3.1: $A\beta$ starting structures: (a) $A\beta_{40}$ (PDB 1BA4), (b) $A\beta_{42}$ (PDB 1Z0Q). The N and C termini are represented as blue and red beads, respectively.

with accession code PDB 1BA4 [69] was solved in 1998 by solution NMR at a moderately acidic pH 5.1 and in a micellar system composed of water–sodium dodecyl sulfate mixture. It features largely unstructured N-terminal residues extending up to His14, after which an α -helix involving a Gly25-Ser26-Asn27 hinge is adopted up to Val36. The employed conditions more closely mimic a membrane or a membrane-aqueous system rather a predominantly aqueous system as modelled in the present work. The first of the ten NMR structures deposited in the data bank was selected as starting configuration for MD in explicit solvent environment. The simulation was designed to grant insight into the likely conformational behaviour of the peptide in completely aqueous milieu, akin to what is experienced *in vivo* after enzymatic cleavage, after which the peptide is released in the extracellular compartment where aggregation into toxic oligomeric structures takes place. Importantly, the choice of aqueous simulation conditions would provide an understanding of the conformational profile of the monomer which expectedly serves as the foundation for higher order structural organisation, and at the same time provide a structural reference for rationalising the observed differences in the closely related $A\beta_{42}$ monomer whose possession of two additional C-terminal residues has been proposed as the explanation for the observed higher aggregation and neurotoxicity propensities[25, 54].

The starting structure for A β_{42} was also obtained from solution NMR-determined configuration obtained from the RCSB Protein Data Bank (figure (3.1(b)), accession code PDB 1Z0Q)[70]. The structural determination was performed in an apolar organic solvent, hexafluoroisopropanol (HFIP) and the A β_{42} monomer was observed to adopt two helical subunits, the shorter of which was sampled in the C-terminal segment of the peptide. Using a combination of CD and NMR spectroscopies, the effect of increasing water concentration up to 99%v/v was probed and observed to involve reversible conversion of the predominantly C-terminal helix at high HFIP concentration to β -sheet-rich conformation at water levels greater than 80%v/v[70]. As long as the apolar HFIP was present however, the N-terminal helix was retained leaving the important question of what conformations would be adopted in a completely aqueous environment. The significantly high aggregation propensity of the peptide has, however, prevented conventional experiment from being able to address the question. And this is where molecular dynamics simulation provides a powerful method of investigating the peptide's conformational experience in 100%v/v aqueous system: the very fact that a single molecule can be studied, immediately circumvents the high aggregation propensity challenge.

3.4.2 Simulation details

The relevance of a force field crucially depends on its parameterisation scheme and employing it for systems for which it was not originally designed should be done with caution. This, for example, applies to most protein force fields whose parameter development were based on small non-protein compounds. One way this work has selected to reduce force field peculiarities is by employing two different force fields for each protein system—the GROMOS force fields ffG43a2 parameterised to reproduce experimental heat of vaporisation of small molecules [71], and ffG53a6 parameterised to reproduce experimental

solvation enthalpies for small molecules. [9]. Both are united-atom force fields which ensures sufficient similarity in overall force field structure, and the simulation results from each were validated against experimental NMR chemical shift data (discussed below). A cubic simulation box was set up for each of the A β_{40} and A β_{42} monomers, and the peptide centered inside with a 10 Å minimum distance allowed between each atom and the edges of the box giving a box length of ≈ 65 Å. A rather large box was chosen so as to avoid artefacts resulting from the constraining effect of smaller box types preventing protein unfolding[72]. Artefacts from edge effects (unnatural edge truncation) resulting from the use of a finite system, however, remains and in order to reduce this effect, a periodic boundary condition (PBC) was applied across the box edges. This entails surrounding the simulation box with the translated copies (periodic images) of itself in the xyz directions closely resembling a boundless situation[4]. There is a special implication to this, the possibility of simulated particles to be surrounded by multiple copies of itself which calls for a special way of handling nonbonded interactions. In GROMACS, the minimum image convention method is employed, that involves computing the short-range component of the nonbonded interaction for only the closest periodic image. The choice of a cut-off radius smaller than half the box vector for a cubic box (the smallest of the box vectors for noncubic boxes) ensures that a maximum of one copy of the particle is considered for short-range interactions. The Particle Mesh-Ewald method, a lattice sum approach, was employed for calculating the long-range component of the coulombic interactions. A 14 Å cutoff was used for truncating the short-range electrostatic (also van der Waals) forces. Following steepest descent energy minimization of A β_{40} (and A β_{42}) *in vacuo*, sufficient water molecules, using the simple point-charge (SPC) water model were added to fill the simulation box. The choice of protonation states for the ionizable residues to mimic the near-neutral physiological pH of 7.4 produced a -3 net charge per molecule. This involves protonating basic groups – Arg5, Lys16, Lys28 and the N-terminus, while deprotonating

the acidic groups Asp1, 7, 23, Glu3, 11, 22 and the C-terminus. Counter ions, Na⁺ and Cl⁻, were randomly added to neutralise the net molecular charge while at the same time achieving a NaCl concentration of 150 mM. The addition of solvent often leaves pockets of empty spaces within the system, while at the same time there is the possibility of placing a solvent molecule too close to the peptide (that is, violating the van der Waals radius). Both phenomena results in high system energy, and it is crucial to first bring the system as close to equilibrium as possible before performing the production MD run. For this reason, the solvated system was first subjected to 5000 steps of steepest descent minimisation followed by 5000 steps of conjugate gradient minimization. This was followed by position-restrained dynamics performed for 1 ns, during which a harmonic restraining force of 239 kcal/mol was placed on each coordinate of the peptide's heavy (i.e., non-hydrogen) atoms, allowing the solvent molecules to relax around the restrained peptide. This serves as a very efficient means of removing *bad contacts* remaining after the energy minimisations steps, thus bringing the system near equilibrium conditions. The position-restrained dynamics was performed under constant temperature (300 K) and pressure (1 bar) conditions (i.e., NPT ensemble) involving a coupling of the system to a Berendsen thermostat and barostat, respectively. After turning off the position restraints, and replacing the temperature and pressure coupling respectively with the Nosé-Hoover thermostat and a Parrinello–Rahman barostat, the system was subjected to a 1.5 μ s MD production run. Neighbour list generation, that is the *dynamic* (static over the number of steps used in updating the list) list defining neighbouring atoms for computing nonbonded interactions, was updated every 10 ps using a cutoff radius of 1.4 nm. Coordinates were saved every 20 ps, and after completion the trajectories were analysed.

System in the simulation setup description stands for each of A β_{40} -ffG43a2, A β_{40} -ffG53a6, A β_{42} -ffG43a2, and A β_{42} -ffG53a6 peptide-force field combinations. That is, an aggregate of 6.0 μ s MD simulation in explicit solvent was performed for a comparative study of the

conformational behaviour of the two A β monomers as well as for determining the accuracy of the force fields at describing the two A β peptides' dynamics. Each simulation was performed using 64 cores of the Jülich Supercomputer Centre's (JSC) JUROPA computer cluster. The use of virtual hydrogen sites allowed the use of a simulation timestep of 5 fs, and a simulation time of approximately fifteen days for each system.

3.4.3 The structure of A β_{40}

In order to monitor the transition of the peptide's secondary structure, the program DSSP [73] that employs hydrogen bond definition in addition to information from geometric pattern analysis, was employed. Interfacing of GROMACS with DSSP allows the secondary structure transition to be obtained as a function of simulation time. The DSSP plot for ffG43a2 (figure 3.2) shows a retention of the N-terminal α -helix, while the C-terminal-ward edge underwent a conversion into turns and β -strands within the first 20 ns. Between 20–90 ns the α -helix was partly converted to π -helix involving residues 13–25, which between 140 and 340 ns reduced to residues 13–18. Relative to the other residues of the peptide and according to the force field, residues 13–25 seem to possess a high helical propensity, which gradually converts between the α -helical form obtained in apolar medium (HFIP) to π -helical in 100% completely aqueous medium. Conversion between the helical forms represents gradual destabilisation of the helix which eventually was converted into turns between 900–1100 ns. The N-terminal residues 1–14 on the other hand remained largely disordered throughout the 1.5 μ s-long trajectory, only fleetingly sampling α -helix (for instance, the boundary residue 13) and β -bridge in the second half of the simulation. β -structures were also sampled, involving the C-terminal residues 28–39, separated from the helical CHC by a coil structure involving Gly25-Ser26-Asn27. Residues 30–31 and 34–35 existed as turns.

Simulation with the ffG53a6 parameter set produced certain differences in the confor-

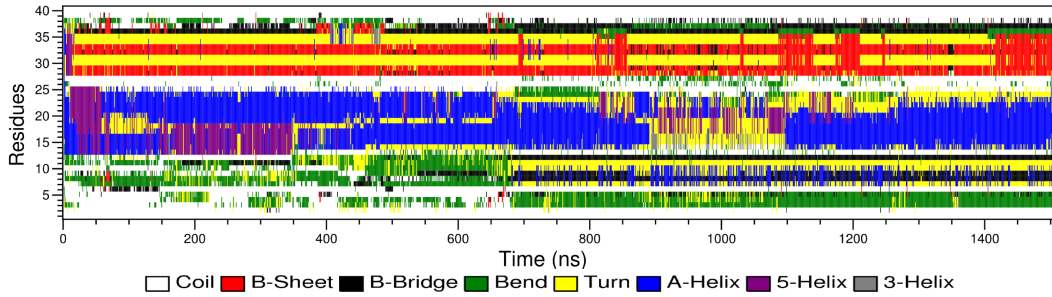


Figure 3.2: DSSP plot showing secondary structure transition in $A\beta_{40}$ as modelled by ffG43a2. The structural pattern involves three distinct segment-based conformations with N-terminal up to residue 14 sampling turns and bends; a stable α -helix is sampled involving residues 16 to 24 of the peptide; and residues in the C-terminal segment forms β -sheet structure.

mations sampled by $A\beta_{40}$ (figure 3.3). With ffG53a6, the helix in the CHC completely disappeared after 600 ns. Also notable is the fact that participation of residues in the helix was limited to just residues 10–14. The initial 200 ns saw the C-terminal section undergo structural transitions involving helix, β -strands and disordered conformations. From 200 ns on a β -strand was more stably sampled between residues 19–23 and 32–37, while 24–31 for the most part remained disordered. To a large extent the N-terminal residues did not form a β -strand: between $t=0$ and $t=600$ ns residues up to 10 were disordered, which for the rest of the simulations increased to residues 1–17. In general, the appearance of the DSSP-obtained secondary structure plots for $A\beta_{40}$ indicates a higher structural flexibility for ffG53a6 than for ffG43a2 in which sampled structures generally persisted longer.

To better describe the structural states sampled by the force fields, cluster analysis was

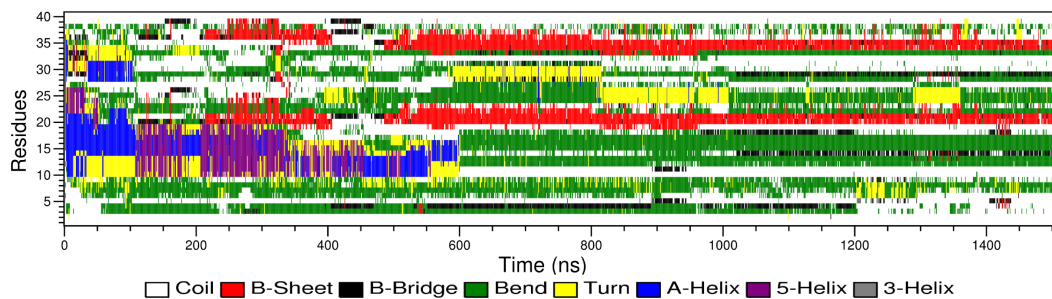


Figure 3.3: DSSP plot showing secondary structure transition in $A\beta_{40}$ as modelled by ffG53a6.

performed on each trajectory using the method by Daura et al. [74]. This involves the

pooling together of all the sampled configurations in the trajectories, and counting the number of structures (neighbours) within a predetermined cutoff radius. The structure with the highest number of neighbours which together with its neighbours constitute a cluster, is then eliminated from the pool alongside its neighbours. The process is repeated until all the structures have been assigned to clusters. Using a 2 Å cutoff for the backbone atoms, and the last 1.0 μ s of the trajectories (chemical shift analysis, discussed below revealed the trajectories as having equilibrated after $t=500$ ns), cluster analysis was performed. Figure 3.4 presents the central structures for the five largest clusters for each force field. Before discussing the result from the cluster analysis, it should be noted that an instance of secondary structure assignment conflict occurred in figure 3a, involving a π -helix assigned by STRIDE (the secondary structure assignment algorithm used by VMD, the employed graphic programme) for residues 28–36 to which DSSP assigned a strand-turn-strand-turn structure. No other instance of assignment conflict was observed and for consistency the assignment by DSSP has been adopted in this work.

Results of the cluster analysis significantly capture the structural changes featured in

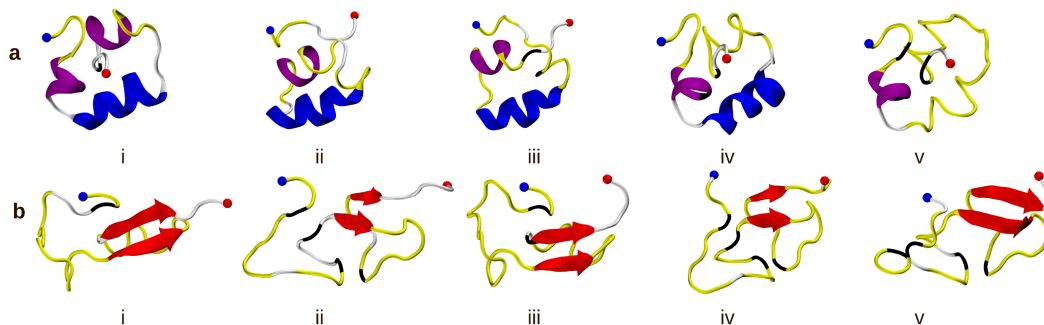


Figure 3.4: A β_{40} structures for the centers of the five most populated clusters obtained from the last 1000 ns of the MD simulations using ffG43a2(**a**) and ffG53a6(**b**) with the cluster size decreasing from (i) to (v). The blue and the red spheres respectively show the N- and C-termini. π helix, β -sheet, and β -bridge regions are shown in purple, red and black respectively, while the yellow and white regions are unstructured.

the DSSP plots (figures 3.2 and 3.3). In the case of ffG43a2 the five largest clusters represent 94% of all the sampled configurations while for ffG53a6 only 26% of the sampled configurations were contained in the five clusters, indicating a higher conformational flexibility in the latter force field compared to the former. The DSSP plots suggest that

the N-terminal region contributes significantly to this flexibility difference. An overlay of the root mean square fluctuations of the backbone atoms in figure 3.5 indicates a much higher structural flexibility in the peptide when simulated with ffG53a6. This may also

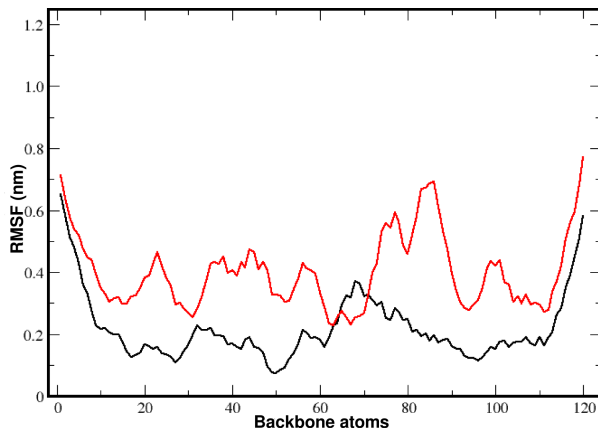


Figure 3.5: Root mean square fluctuation of the backbone atoms of $A\beta_{40}$ showing higher peptide flexibility with ffG53a6 (red) compared with ffG43a2 (black).

be described as ffG53a6 favouring disordered states for $A\beta_{40}$, which better correlates with the expectation for $A\beta$ monomers. The cluster centres for ffG53a6 feature a largely disordered N-terminal segment and a β -strand involving two sequence stretches, $^{19}\text{FAED}^{23}$ of the CHC self-recognition unit, and $^{32}\text{IGLMVG}^{37}$ belonging to the C-terminal hydrophobic patch. The terminal three residues, $^{38}\text{GVV}^{40}$, exist as disordered in all five clusters. This partly agrees with the structural model proposed by Danielsson *et al.* from ^{15}N relaxation data.[75]. Using persistence lengths obtained at temperatures ranging from 3 to 18°C, a model composed of six structurally distinct units was proposed for $A\beta_{40}$ involving a π -helix for residues 1–4; a coil structure between residues 5–10; a second π -helix for residues 11–15; a β -strand for residues 16–24; a second coil involving residues 25–30; and finally a β -strand involving C-terminal hydrophobic residues 31–40. In the ffG53a6 simulation, the observed unstructured region involving residues 1–18 and 24–31 and the β -strand of residues 19–23 (16–24 in the experimental model) and 32–37 (31–40 in the experimental model) are in agreement with the proposed NMR relaxation model. The exception is the π -helix involving residues 1–4 and 11–15, which was not found to be stable in the ffG53a6

simulation of $A\beta_{40}$.

3.4.4 The structure of $A\beta_{42}$

As presented in figure 3.6, the plots of the secondary structure of $A\beta_{42}$ obtained with ffG43a2 reveals a pattern similar to that obtained for the shorter peptide, especially in the formation of a relatively stable N-terminal helix. However, in the present case, a π -helix (rather than an α -helix) between residues 5–15 is observed. Between ca 370 and 840 ns, the mid-section region spanning residues 20–30 support a strand–turn–strand conformation. This structure was subsequently replaced by a disordered state, while sequences spanning residues 33–40 and 2–4 sampled β -strands.

With the ffG53a6 force field, the two helical units present in the NMR starting structure

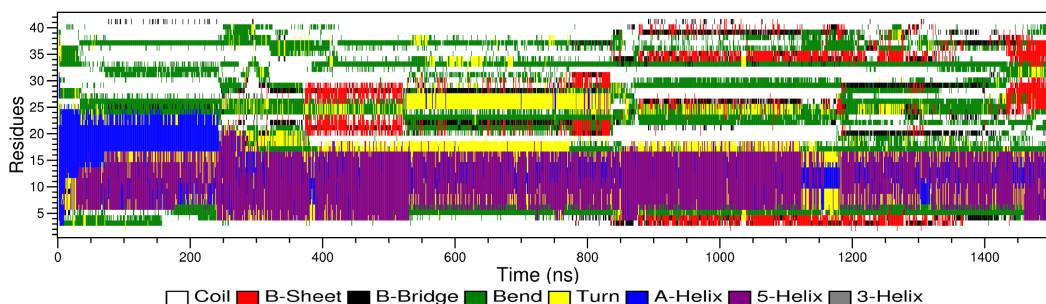


Figure 3.6: DSSP plot showing secondary structure transition in $A\beta_{42}$ as modelled by ffG43a2.

were completely lost within 100 ns (figure 3.7), suggesting a much higher conformational flexibility than observed with ffG43a2. This, however, does not indicate inability to sample specific conformations. Based on sampled β -strand–coil features, the trajectory may be viewed as being temporally divided into four stages. In the first 400 ns residues 24–40 intermittently but continually sample β -strand structure while residues 1–23 remain largely disordered. During the next 325 ns, the N-terminal section stably sampled β -strand while the remaining residues (i.e. covering the CHC and the C-terminal sections) adopted disordered states. Between 725 ns and 1250 ns, the peptide again adopts a

structure characterised by the C-terminal region adopting β -sheet state and disordered N-terminal residues. The first 1250 ns thus sees β -sheets sampled interchangeably and with a mutual exclusivity by the two half segments of the peptide. During the last 250 ns, however, the β -sheet additionally extends to the N-terminal residues 6–13.

As for A β_{40} , cluster analysis was performed on the last microsecond of the MD tra-

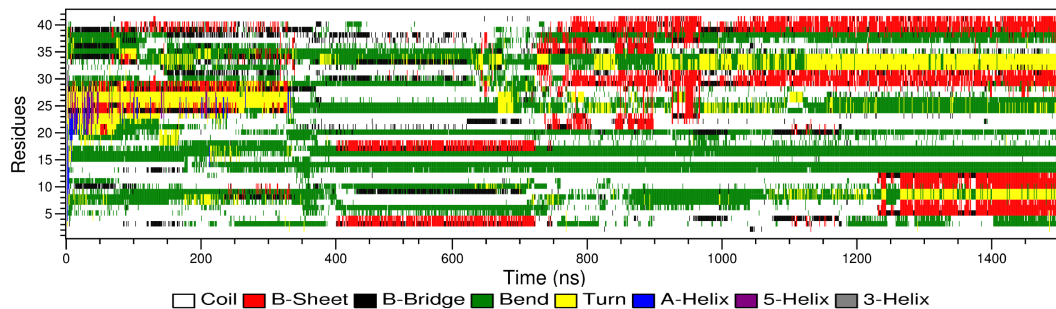


Figure 3.7: DSSP plot showing secondary structure transition in A β_{42} as modelled by ffG53a6.

jectories with the same 2 Å RMSD cutoff. In agreement with the secondary structure plot for ffG43a2, the five cluster centres (Figure 3.8) adopt an extensive N-terminal π -helix spanning residues 4–16. The five largest clusters for ffG43a2 represent 76% of all sampled configurations indicating a higher conformational rigidity compared with ffG53a6, in which 48% of all structures are captured in the five clusters. In the case of ffG53a6, the cluster centers mostly feature β -sheets involving residues ³EFRHDSG⁹, residues ¹⁷LVFFAE²² – the CHC self-recognition unit – and the C-terminal hydrophobic stretch ³⁰AIIGLMVGGVV⁴⁰ where the italicised residues are disordered. In 2004, Urbanc *et al.* identified the disordered ³⁶VGG³⁸ as distinctly characteristic of A β_{42} where it forms a hinge[56]. In agreement, the ffG53a6 force field correctly identified the hinge at ³⁶VGG³⁸ in A β_{42} .

From the relative size of the five most populated clusters, it becomes obvious that the observed pattern of conformational flexibility arises from the force field rather than solely from the simulated peptides. As observed above for the shorter peptide, A β_{42} monomer also undergoes a higher structural dynamics when simulated with the ffG53a6 as shown

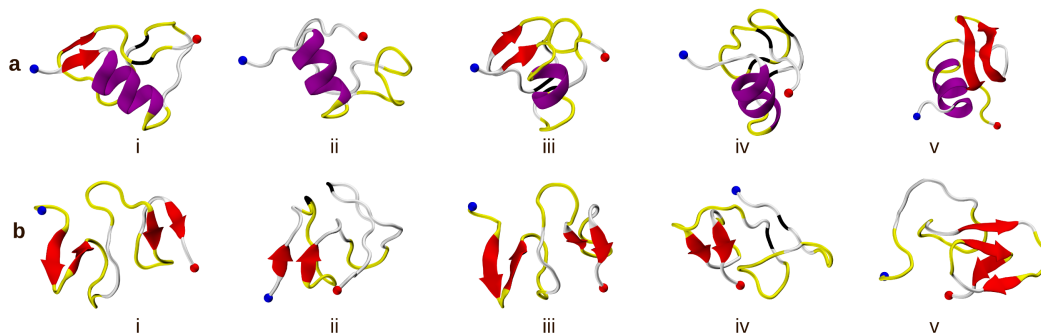


Figure 3.8: $A\beta_{42}$ structures for the centers of the five most populated clusters obtained from the last 1000 ns of the MD simulations using ffG43a2(**a**) and ffG53a6(**b**) with the cluster size decreasing from (i) to (v). The blue and the red spheres respectively show the N- and C-termini. π helix, β -sheet, and β -bridge regions are shown in purple, red and black respectively, while the yellow and white regions are unstructured.

in the backbone RMSF plot in figure 3.9.

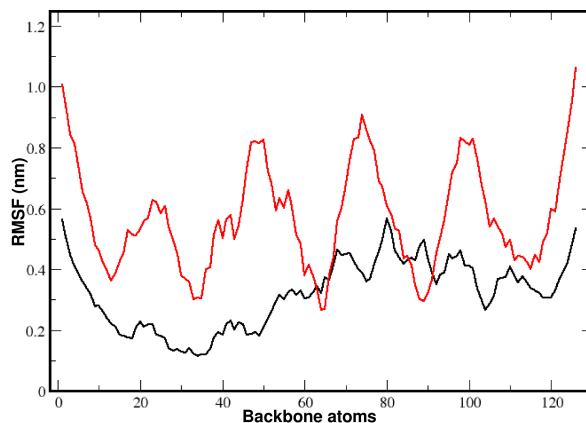


Figure 3.9: Root mean square fluctuation of the backbone atoms of $A\beta_{42}$ showing higher peptide flexibility with ffG53a6(red) compared with ffG43a2 (black).

3.4.5 Comparison of $A\beta_{40}$ and $A\beta_{42}$

The extra two residues distinguishing the primary sequence of $A\beta_{40}$ and $A\beta_{42}$ have been indicated to confer significant differences on both the conformation and aggregation properties of the two peptides[56, 57]. Both peptides were shown by NOE measurements to undergo a high conformational interconversion at the ps-to-ns time-scale; however, higher NOE values were obtained for the C-terminus of $A\beta_{42}$ indicating a relatively more rigid C-terminus than for the shorter peptide[54]. This was proposed to maintain $A\beta_{42}$'s C-terminus in a β -strand conformation which acts as an internal seed for amyloid aggre-

gation. The diglycine turn (Gly37-Gly38) turn was only identified in A β_{42} , which was proposed as a distinguishing feature of A β_{42} since it was not found in A β_{40} [56, 57]. This was used to explain the preference of A β_{42} for pentamer and hexamer formation, compared with A β_{40} which would rather aggregate into dimeric structures [56, 58, 59]. A β -strand was also uniquely identified between Ala2–Phe4 in A β_{40} [56].

Our simulations present the two peptides as largely unstructured in their first 10 residues (N-terminal). In comparative terms however, A β_{40} appears to be more disordered in this segment than A β_{42} , which additionally samples either π -helix and β -strands in its N-terminal residues. Both force fields successfully predict β -strands in the β -sheet-forming self-recognition sequence ¹⁷LVFFAE²² of A β_{42} . In A β_{40} , only ffG53a6 predicts β -strand and only in the vicinity of the self-recognition unit involving ²⁰FAED²³. The ¹⁶KLVFF²⁰ motif is widely accepted as the first key motif for nucleating A β aggregation [76], being one of two hydrophobic stretches present—¹⁷LVFFA²¹ and ³⁰AIIGLMVGGVV⁴⁰IA⁴². The simulations for both peptides additionally reveal Gly25–Asn27 as constituting some sort of a bridge, which conformationally separates the two hydrophobic patches. In agreement with the suggestion that a conformational transition from helix to coil precedes the characteristic β -sheet-dependent aggregation of A β [70], both force fields predict the complete loss of the helix in the second hydrophobic patch of both peptides. The C-terminal segments of both peptides sampled a β -strand conformation; however in agreement with earlier studies the hinge ³⁷GG³⁸ identified as characteristic of the longer peptide [56, 57] was sampled only in A β_{42} .

Furthermore, the ffG53a6 simulation showed A β_{42} as capable of sampling β -strand in its N-terminal, mid-sequence, and the C-terminal segments, although the last two segments were significantly better able to do this. This compares with A β_{40} in which β -strands were largely limited to the last two segments, and that to a much lesser degree. The five clusters obtained from ffG43a2 for A β_{42} feature an extensive π -helix spanning sequence

spanning residues 4–16, which slightly contrasts with the preferential sampling of α -helix in $A\beta_{40}$, but strongly contrasts with the stabilisation of disordered structure and β -strands predicted by ffG53a6. Both force fields however indicate a higher β -sheet propensity for $A\beta_{42}$ than $A\beta_{40}$. Thus, by using two force fields based on different parameterisation philosophies we have been able to identify peptide-based structural properties.

The terminal residues of both peptides adopt a coil structure. However, while these in-

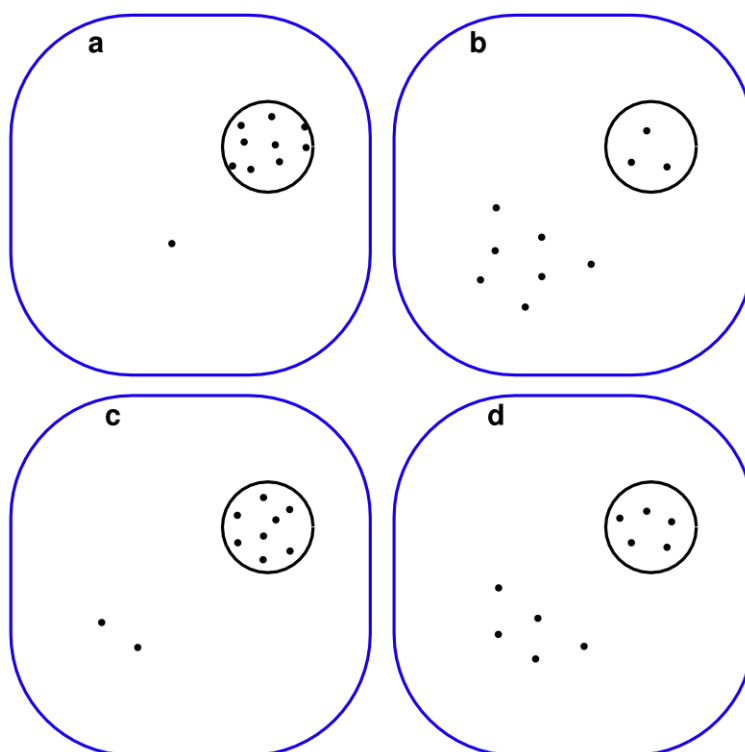


Figure 3.10: A pictorial representation of the relative structural flexibilities of $A\beta_{40}$ and $A\beta_{42}$ and the influence of force fields: (a) $A\beta_{40}$ with ffG43a2, (b) $A\beta_{40}$ with ffG53a6, (c) $A\beta_{42}$ with ffG43a2, (d) $A\beta_{42}$ with ffG53a6. For a description of the figure, see text below.

volve three residues of $A\beta_{40}$, only two of $A\beta_{42}$ exist as coil. This perhaps is an indication of a higher C-terminal stiffness reported for the longer peptide [54]. Comparing the percentage representations obtained in the cluster analyses of the two peptides is instructive in further capturing the influence of the extra two residues in $A\beta_{42}$. With ffG53a6, 48% of the sampled frames were captured in the five biggest clusters of $A\beta_{42}$ and 26% for $A\beta_{40}$, which demonstrates $A\beta_{42}$ as being more rigid than $A\beta_{40}$. This supports the findings of Yan and Wang, who, through experimental measurements, found $A\beta_{42}$ to be more rigid

at its C-terminus than $A\beta_{40}$ [54]. On the other hand, ffG43a2 supports a more rigid $A\beta_{40}$, with 94% structures within the five largest clusters, compared with 76% in $A\beta_{42}$. An estimate of the relative rigidity can be determined from the cluster analysis results. Figure 3.10 shows a pictorial representation using information from cluster analyses to distinguish between the observed peptide- and force field-based differences. Each of the four peptide-force field systems is represented by a blue boundary (dimensionless) within which all sampled conformations are represented by the 10 black dots. The conformations falling within the five largest clusters are shown inside the black circles. The number of conformations within and without the black cycles provides a simple means of comparing the behaviour of both the peptides and the force fields. The degree of force field bias for the folded state, for example, can be estimated using the number of clusters falling outside the black circles. For instance, ffG43a2 is about 7 times and 2.5 times less likely to sample disordered conformations compared with ffG53a6 for $A\beta_{40}$ (figure 3.10(a) vs (b)) and $A\beta_{42}$ (figure 3.10(c) vs (d)) respectively. Vertical comparison on the other hand presents the differences between the two peptides, with ffG43a2 showing $A\beta_{40}$ to be about 2 times more rigid than $A\beta_{42}$ (figure 3.10(a) vs (c)) against experimental findings[54]. In figure 3.10(b) and (d) on the other hand and in agreement with experiment findings[54], ffG53a6 supports a slightly higher $A\beta_{42}$ conformational rigidity, about 1.4 that obtained for $A\beta_{40}$.

3.4.6 Experimental validation using chemical shifts

The simulation results presented above suggest there are significant force field biases in the conformational study of the two $A\beta$ peptides. One possibility of identifying which force field is biased towards which type of secondary structure (e.g., helical, coil and β -sheet) is by assessing the convergence of the trajectories with respect to experimental data. For

this we have chosen solution NMR chemical shifts which were determined in the group of Prof. Zagorski at the Department of Chemistry, Case Western Reserve University, Cleveland, USA[77]. It should be noted that the determination of the NMR chemical shifts at 278 K rather than at 300 K as employed by our simulations, may slightly but not greatly affect the results since both studies employ the same pH and aqueous conditions.

We calculated the N, HN, C α , and C β chemical shifts (δ^{calc}) using the programme CamShift that employs the dependence of chemical shift on conformation as a function of intramolecular atomic distances expressed as a polynomial expansion[78]. See section 3.7 for additional information on CamShift. CamShift takes individual coordinate files as input for the prediction, which we adapted to work on trajectories by taking snapshots at every 50 ps and performing the shift calculation. We then calculated the deviation $\Delta\delta$ of δ^{calc} from the δ^{exp} for the entire 1.5 μ s trajectories for each A β peptide. Figure 3.11 shows the time evolution of δ^{calc} and $\Delta\delta$ for the C α atoms, while table 3.1 presents the $\Delta\delta$ for each 500 ns time-block for the selected four atom types. We base the acceptability of the calculated $\Delta\delta$ values on the published values for a test set of protein in which CamShift was shown to be of comparable accuracy as other chemical shift predictors[78]. The reference $\Delta\delta$ data have also been presented in table 3.1.

The data obtained for the four chemical shift types (N, HN, C α , and C β) show the $\Delta\delta$ values as being within the expected deviations, in particular for the force field ffG53a6. The discussion of the results, from this point on until the end of italicization, is cited as excerpt from our work, Olubiyi OO, Strodel B, Journal of Physical Chemistry 2012, Volume 116, pages 3280-91, 2012:

"The $\Delta\delta$ for A β_{40} and A β_{42} modeled with ffG43a2 show that this force field is not fully able to reproduce the experimentally determined shift values. Although δ^{calc} for HN and C β are generally consistent with the experimental values (ffG43a2 and ffG53a6), the $\Delta\delta$ for the amide N and C α shifts reveal that these shifts really converged to the experimental

values only for the simulations with ffG53a6. With ffG53a6, the average $\Delta\delta$ for the N and C α shifts are generally below the expected deviation for simulation times longer than 500 ns. Only for A β_{40} with ffG53a6, $\Delta\delta$ for the N chemical shift is slightly above 3.01 ppm between 500 and 1000 ns; it decreases to values below 3.01 ppm for simulation times above 600 ns. This decrease in $\Delta\delta$ coincides with the conversion from a helical to coil state in the N-terminal segment around 600 ns (figure 3.3).

In A β_{42} (ffG43a2), a transition from coil to β -strand is observed for some N-terminal

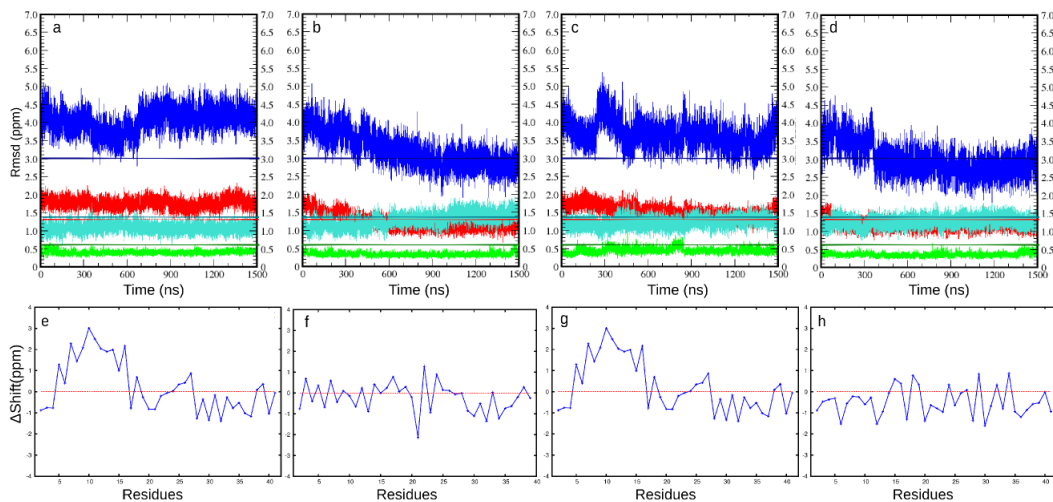


Figure 3.11: $\Delta\delta$ values between δ^{calc} and δ^{exp} plotted against time for (a) A β_{40} with ffG43a2, (b) A β_{40} with ffG53a6, (c) A β_{42} with ffG43a2, (d) A β_{42} with ffG53a6. Values for the amide N, C α , C β and HN atoms are shown in blue, red, turquoise and green with the CamShift published deviations for the protein test set represented with horizontal lines. Deviations of the δ^{calc} from δ^{exp} for the C α atoms (i.e., $\Delta\delta = \delta^{calc} - \delta^{exp}$) using the last 1000 ns are shown for (e) A β_{40} with ffG43a2, (f) A β_{40} with ffG53a6, (g) A β_{42} with ffG43a2, (h) A β_{42} with ffG53a6.

residues around 400 ns (figure 3.7), which causes $\Delta\delta$ to decrease considerably at this time, that is, better agreeing with experimental data. This finding supports our approach to use 500 ns as the boundary between the equilibration and the production phases for the analysis of our MD simulations. For the ffG43a2 simulations, however, it becomes evident that the $\Delta\delta$ values do not decrease during the last 1000 ns of the MD trajectories; in some cases, they even increase, showing that this particular force field does not correctly sample the experimentally observed structures.

To get insight into which of the residues cause the deviation of δ^{calc} from δ^{exp} , we computed for the last 1000 ns of the MD trajectories, the average δ^{calc} values for C α , $\langle\delta^{calc}\rangle$, and plotted these against δ^{exp} in Figure 3.12. And we calculated the difference between $\langle\delta^{calc}\rangle$ and δ^{exp} : $\Delta\delta = \langle\delta^{calc}\rangle - \delta^{exp}$. The results for $\Delta\delta$ are shown in figure 3.11. For the simulation of A β_{40} with ffG43a2, it becomes obvious that large deviations between

$\langle \delta^{calc} \rangle$ and δ^{exp} ($|\Delta\delta| > 1$ ppm) occur for almost all residues between residues 7 and 35. Particularly striking is the overestimation of $\langle \delta^{calc} \rangle$ between residues 13 and 23, whereas for the other residues, we observe an alternation of over- and underestimation of $\langle \delta^{calc} \rangle$, for example, from residues 7 to 12 and from 28 to 35.

Since C α atoms experience a relative upfield shift when the residue in question is in-

Force field	Peptide	Time (ns)	N(3.01)	HN(0.56)	C α (1.30)	C β (1.36)
ffG43a2	A β_{40}	0 \rightarrow 500	3.95	0.40	1.79	1.09
	A β_{40}	500 \rightarrow 1000	4.01	0.40	1.76	1.09
	A β_{40}	1000 \rightarrow 1500	4.19	0.41	1.77	1.07
	A β_{42}	0 \rightarrow 500	3.87	0.43	1.69	1.17
	A β_{42}	500 \rightarrow 1000	3.64	0.46	1.49	1.20
	A β_{42}	1000 \rightarrow 1500	3.51	0.44	1.43	1.26
ffG53a6	A β_{40}	0 \rightarrow 500	3.71	0.35	1.46	1.13
	A β_{40}	500 \rightarrow 1000	3.08	0.33	1.16	1.27
	A β_{40}	1000 \rightarrow 1500	2.87	0.34	1.12	1.44
	A β_{42}	0 \rightarrow 500	3.40	0.36	1.24	1.22
	A β_{42}	500 \rightarrow 1000	2.68	0.35	1.30	1.19
	A β_{42}	1000 \rightarrow 1500	2.74	0.37	1.18	1.30

Table 3.1: $|\Delta\delta|$ table for the amide N, C α , C β and HN atoms. For comparison, $\Delta\delta$ values published for CamShift using a test set of proteins are presented in parentheses.

corporated into an α - or π -helix [79, 80] or a turn region,[79] we conclude that in the regions with $\Delta\delta > 1$ ppm, one of these secondary structures is overstabilized by ffG43a2. Indeed, the DSSP plot in figure 3.2 shows a rather stable α -helix between residues 13 and 23 in A β_{40} . From ^{15}N relaxation data, a structural propensity for π -helix was predicted for residues 11-15 in A β_{40} . However, from the comparison of the C α shift results, it becomes obvious that ffG43a2 overestimates the N-terminal helical stability in A β_{40} . This helical stability leads to a reduction of the conformational flexibility of the A β peptides, as the high population of the first five clusters from the ffG43a2 simulations revealed. A too-small computed C α shift is indicative of an overstabilization of β structure as C α protons experience a relative downfield shift when incorporated into a β -sheet[79]. The alternation of too-high and too-low shifts between residues 7-12 and 28-35 is thus a result of the overestimation of successive turn and β conformations, respectively. This conclusion is underpinned by the DSSP plot in figure 3.2.

In summary, ffG43a2 does not provide a satisfactory model for the conformational dynamics of A β_{40} . For A β_{42} , the ffG43a2 result is better, especially for the C-terminal part from residue 17 onward, for which most $|\Delta\delta|$ values are < 1 ppm. However, for residues 5-16, we observe an overstabilization of the N-terminal π -helix (figure 3.6), leading to $\Delta\delta > 1$ ppm in this region. According to our chemical shift analysis, ffG53a6 provides a better description for both A β_{40} and A β_{42} . Most of the $|\Delta\delta|$ values are < 1 ppm. Note-

worthy exceptions are residue 22 in A β_{40} and residues 6, 12, and 30 in A β_{42} , for which the β content is slightly overestimated. Our comparison between experimental and simulated chemical shifts is of similar performance as the same kind of analysis in ref [81], in which the simulated shifts were obtained from replica exchange MD simulations of A β_{40} and A β_{42} using the AMBER99SB force field[81] with implicit solvation. The representative structures obtained by Yang and Teplow also show similarities to the most important cluster structures, which we obtained with ffG53a6 [53].

In summary, from the comparison between calculated and experimental chemical shifts,

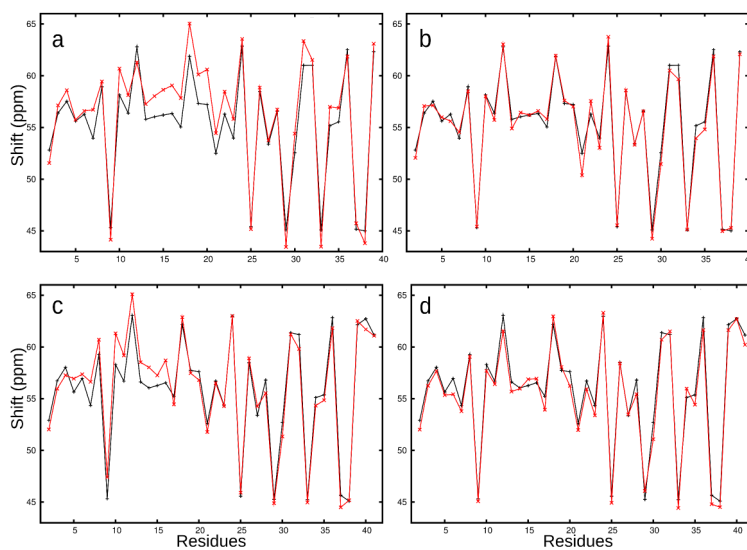


Figure 3.12: Comparison of $|\Delta\delta|$ (red) and δ^{exp} (black) using C α atoms for (a) A β_{40} with ffG43a2, (b) A β_{40} with ffG53a6, (c) A β_{42} with ffG43a2, and (d) A β_{42} with ffG53a6.

we can conclude that ffG53a6 is better able to model the intrinsically disordered A β peptide than ffG43a2, which overstabilizes the helical state in the N-terminal half of the sequence and underestimates the overall conformational flexibility of the peptide. This finding is contrary to the performance of ffG43a2 and ffG53a6 for folded proteins, in which ffG43a2 was found to be better able to maintain folded structures and generally provides a good balance between helical and β folds[82]. The superiority of ffG53a6 for modeling A β may be due to the fact that this force field was solely parametrized to reproduce the free enthalpies of hydration and apolar solvation for a range of compounds[83].

The relative free enthalpy of solvation between polar and apolar environments is a key determinant in many biomolecular processes of interest, such as protein folding and aggregation or membrane formation and transport over membranes. Thus, the folded protein state in ffG53a6 results from the choice of the parameter set and not vice versa. In a re-

cent study, it was shown that the common protein force fields (ffG53a6 was not included) are generally able to correctly predict the rate of folding and the structure of the native state[84]. The folding mechanism and the properties of the unfolded state, however, were found to depend substantially on the choice of force field. This finding, as well as other force field studies for A β peptides,[55] demonstrate the need for proper benchmarking of the protein force fields for unfolded and intrinsically disordered proteins.”

3.5 The effect of pH on A β

Various studies conducted on the effect of pH variation on the structure of A β have indicated that the protonation states of the three histidine residues in A β (His6, His13 and His14) are crucial in explaining the role of pH [61, 85, 86]. And in fact, histidine residues have been suggested to sensitise the peptide to metal ions binding, as revealed in mutation studies in which His substitution resulted in reduced A β aggregation in the presence of aggregation-promoting Cu²⁺ and Zn²⁺ [61, 85, 86]. In this section the effect of pH variation on the structural dynamics of the two A β monomers is discussed.

The pH of a system can be defined as the negative logarithm of the molar concentration of protons present:

$$pH = -\log[H^+] \quad (3.1)$$

Since subatomic units like electrons and protons may only be accurately captured using quantum descriptions (and this not without a prohibitive cost in terms of computational resources for modelling systems containing 30,000 atoms or more), we have employed a classical description in modelling different pH conditions. This essentially involves the choice of the ionisation states of Arg5, Lys16, Lys28, Asp1, Asp7, Asp23, Glu3, Glu11, Glu22, His6, His13, His14 and the N- and C-termini to mimic four pH conditions: an

acidic pH, an isoelectric pH, the physiologic pH and a basic pH (table 3.2).

pH	2.0	5.0	7.4	8.0
R5	+1	+1	+1	0
K16	+1	+1	+1	0
K28	+1	+1	+1	0
D1	0	-1	-1	-1
D7	0	-1	-1	-1
D23	0	-1	-1	-1
E3	0	-1	-1	-1
D11	0	-1	-1	-1
D22	0	-1	-1	-1
H6	+1	+1	0	0
H13	+1	+1	0	0
H14	+1	+1	0	0
N-Term.	+1	+1	+1	0
C-Term.	0	-1	-1	-1
Net Charge	+7	0	-3	-7

Table 3.2: Modelled ionisation states for residues and termini of A β ₄₂.

3.5.1 Simulation setup, models and analyses

Histidine is neutral (also referred to as deprotonated) when protonated only at the δ nitrogen of the imidazolyl sidechain, and positively charged (protonated) when the ϵ nitrogen is additionally protonated. Using an MD protocol similar to that described in 2.4.2, 1 μ s production runs were performed for each modelled pH and for each peptide-force field combination. That is, an additional 6 μ s of simulations time was cumulated; while the first 1 μ s from the 1.5 μ s trajectories discussed above were used for the physiologic pH. All simulations were performed at 300 K and 1 bar, using an explicit solvent representation and an isothermal-isobaric ensemble that closely reflects standard laboratory conditions. Although the simulations were performed with both ffG43a2 and ffG53a6, only results for the latter force field are discussed since ffG43a2 was already shown (above) to suffer from a significant structural bias. Furthermore, test simulations for ffG43a2 revealed that the pattern of variations as a result of pH changes, was captured in a similar manner to that of ffG53a6 [68].

To provide an explanation for the differences in the observed dynamics of the peptide under the considered pH conditions we performed contact map analyses for the four tra-

jectories with a truncation distance of 15 Å at 0 ns, 300 ns, 600 ns and 1 μ s. In order to account for sidechain as well as backbone contributions, all the atoms in the protein were included while constructing the maps. Figure 3.13 presents the contact maps. Each resulting pattern in figure 3.13 may be taken as a fingerprint of A β_{42} monomer's tertiary structure for the particular pH state. In this regard, helical and β structures are the easiest to identify. Helical structures present a diagonal pattern involving residues $i + n$ running parallel to the self-contact axis. For 3_{10} -, α -, and π -helices, n is respectively 3, 4 and 5. Thus helices appear in the map as interactions involving a maximum of 5 neighbouring residues. Extended antiparallel β -sheet structures on the other hand involves off-diagonal interaction, the length and number of which respectively signify the extent and number of the β -sheet. Random coil lacks the discernible patterns of helical and β structures. To better compare the sampled structures we performed DSSP analyses which we have restricted to the final 500 ns of each trajectory at which time the peptide is considered to have equilibrated.

3.5.2 Effect of pH on A β_{42} secondary structure

For all four systems, at time $t = 0$ the peptide features a pattern representing helical conformation derived from the starting structure. By 300 ns, the contact maps have become substantially different for each pH. The peptide has almost entirely lost its helices at pH 5.4 and 7.4 while they continued to be sampled at pH 2.0 and 8.0. This appears to be a separation in dynamics based on A β_{42} 's net-charge, whereby the peptide's helices are retained under high charges of ± 7 and lost at lower charges of 0 (pH 5.4) and -3 (pH 7.4). Visual inspection of the structures in figure 3.15 further shows that at 300 ns, pH 5.4 and 7.4 the helices have been replaced by coils and β -sheets. While the peptide is able to form β -sheets under all four modelled pH values as shown in the DSSP plot in figure

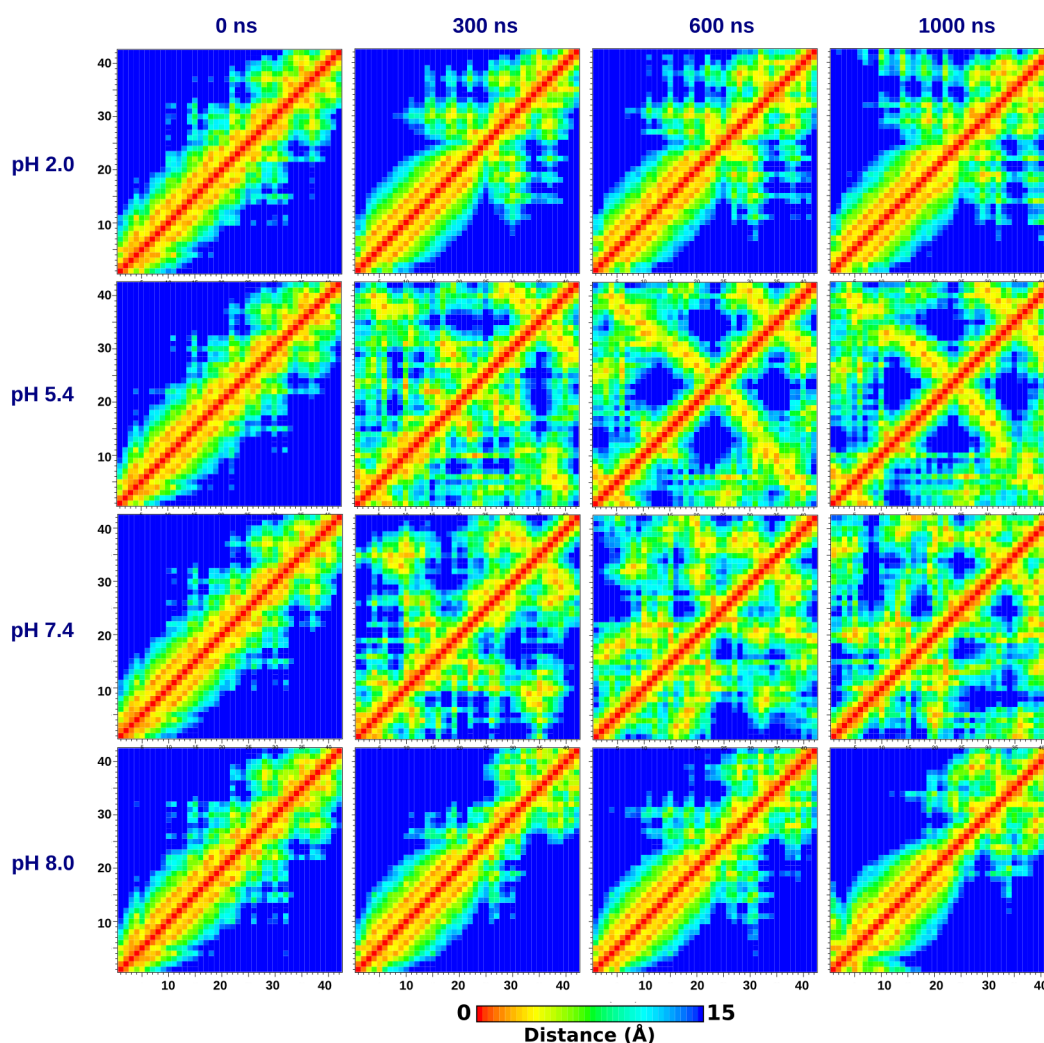


Figure 3.13: Contact maps for A β_{42} monomer at pH 2.0, 5.4, 7.4 and 8.0 using conformations sampled at 0 ns, 300 ns, 600 ns, and 1000 ns.

3.14, β -sheets at pH 2.0 and 8.0 are mostly limited to residues 33 to 42 with a turn at 29 GAI I 32 (pH 2.0) or 27 NKGAI I 32 (pH 8.0). Residues 4–28 and 1–26, respectively at pH 2.0 and 8.0, form helices (figures 3.13 and 3.15). At pH 2.0 and 100 ns, the N-terminal helix formed at 6 HDSGYEVHHQKLVFFAED 23 contains all the ionisable residues (except Lys28). By 200 ns, the helix has reduced to 10 YEVHHQKLVFF 20 . Towards the end of the trajectory at 800 ns, the β -sheet structure increased slightly, accompanied by a decrease in helical structure. A similar trend was obtained for the peptide at pH 8.0. with the difference being in the relative percentage of structures sampled as presented in table 3.3. At pH 5.4 and 7.4, there exist more far-reaching contacts with β -sheets between residues

1–7 and 11–18 at pH 7.4, and between residues 2–23 and 25–42 at pH 5.4. However, β content at pH 7.4 is the lowest.

Our observations can be summed up as follows: at high net charge, A β_{42} 's N-terminal segment forms helical structures while a *separated* C-terminal segment forms β -sheets and random coil. At physiological pH, β -sheet formation is confined to the N-terminal and mid-sequence segments, while at the isoelectric pH, all the peptide residues are involved in β -sheets. For the most part these distinctions remain up to 1000 ns. The structure of A β_{42} at 7.4 continues to evolve with β -sheet alternatively sampled between the N-terminal versus mid-sequence segments and the C-terminal versus mid-sequence segments.

The most significant differences in secondary structure involve the sampled coil, β , and helical structures. The differences are however more poignant when comparing pH 5.4 and 7.4. The latter pH supports a relatively higher coil structure (51%) but lower β -sheet (15%) contents compared to pH 5.4 (31% coil and 39% β -sheet). The highest content of random coil structure was observed at pH 7.4 in agreement with the belief that A β_{42} is unstructured under physiological conditions. The highest percentage of β -sheet was found at pH 5.4 in agreement with CD results [87]. This result suggests a strong role for pH conditions in determining A β_{42} 's structure. Importantly, it indicates the need for caution when generalising the results obtained for the peptide under a specific experimental pH condition. Furthermore, our results revealing A β_{42} 's ability to form β -sheet structure at all four pH conditions, provide an explanation for the observation by Guo *et al.*[63] that A β is able to polymerise regardless of the charge state. The significant differences in the extent to which the different structures are sampled, however, suggests possible differences in aggregation rate, a possible explanation for the role of mutations in AD pathology. For instance, most of the clinically observed mutations, including Glu22→Gly (Arctic), Glu22→Gln (Dutch), Glu22→Lys (Italian), and Asp23→Asn (Iowa), lower the net charge of the peptide and can explain the peptide's associated enhanced aggregation

kinetics.

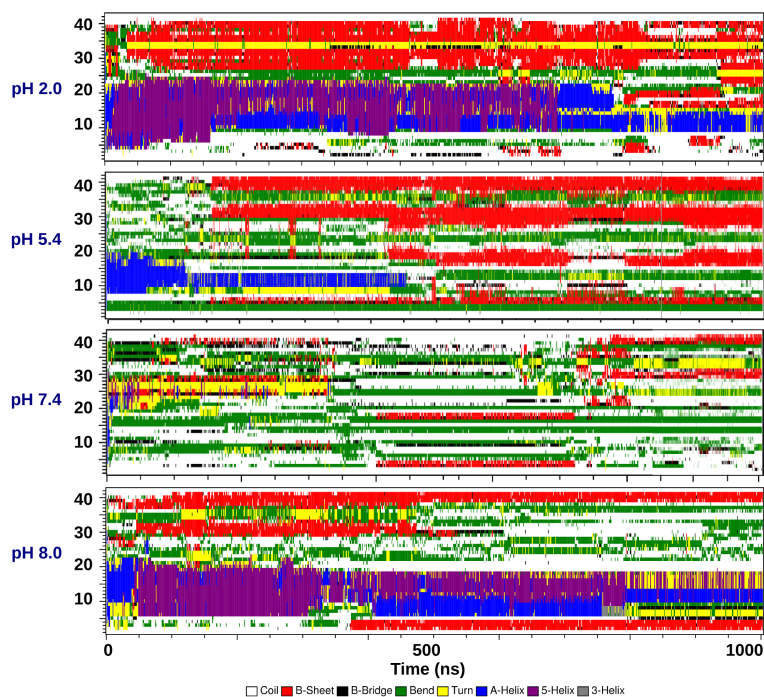


Figure 3.14: DSSP plot showing the secondary structure evolution of $A\beta_{42}$ monomer under different pH conditions.

Structure (%)	pH 2.0 (%)	pH 5.4 (%)	pH 7.4 (%)	pH 8.0 (%)
Coil (14)	29	31	51	38
Beta (0)	29	39	15	21
Bend (7)	11	27	29	20
Turn (19)	10	3	5	8
Helix (64)	21	0	0	14

Table 3.3: The effect of pH on the secondary structure of $A\beta_{42}$ monomer averaged over the last 500 ns of the trajectories. The values in parentheses represent the secondary structure at $t = 0$.

3.5.3 Role of histidine residues

Since the only difference between the $A\beta_{42}$ models employed in modelling the pH 5.4 and 7.4 is the protonation state of His6, His13 and His14, the differences in the structural states observed for the two pH conditions may be directly attributed to the protonation states of the histidine residues. We therefore examined the interactions involving the three histidine residues for these two pH values. At pH 5.4 and as early as 100 ns the three histidine residues (and in particular His6) were observed to be involved in extensive

intermolecular contacts with both the N- and C-terminus of A β_{42} . Intermolecular contacts formed by His6 in the early part of the trajectory include residues Asp1, Glu3 and Ala42. The contact with Ala42 was soon lost, after which His6 employs its protonated imidazolyl sidechain as a surface promoting intramolecular contact between residues from different regions of the peptide (figure 3.16). For this to occur His6 adopted an arrangement that allows all interaction surfaces of the imidazole ring to be available. The δ -N surface forms interaction with residues in the N-terminal segment, while the ϵ -N surface promotes contact with the C-terminal residues. As early as 200 ns $^4\text{FR}^5$ and $^{17}\text{LV}^{18}$ have begun to adopt β -sheet competent alignment. The underside of the imidazolyl ring forms a π -stacking contact with Tyr10, which later recruits $^{17}\text{LVF}^{19}$. His6 was observed to form the nucleus of the inter-domain contacts responsible for the higher β -sheet of A β_{42} at pH 5.4 compared with pH 7.4 under which condition the histidine residues were not observed to serve the described anchoring function.

Instead, at pH 7.4, the N- and C-terminal segments avoid interacting with each other,

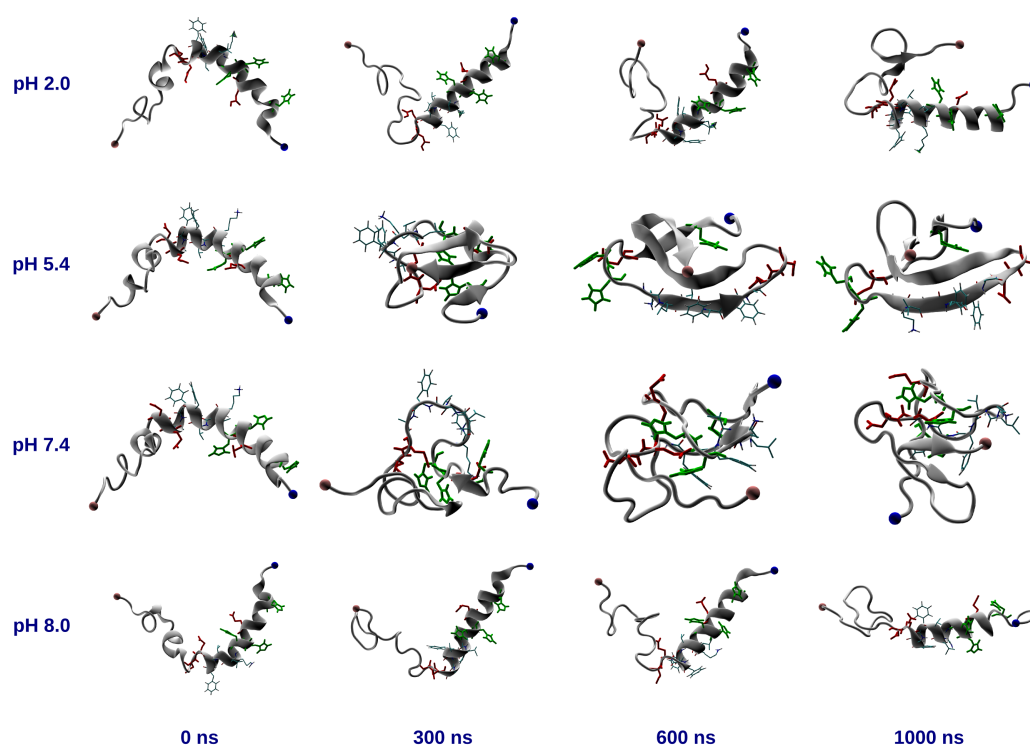


Figure 3.15: A β_{42} monomer structures at different pH and time. Histidine and acidic residues are shown in green and red, respectively, while the KLVFFA sequence is shown in liquorice coloured by atom names. The N- and C-termini are respectively shown as blue and red beads.

which is likely a consequence of charge incompatibility between the hydrophobic C-terminal and the predominantly hydrophilic N-terminal. As can be observed in the distance maps for pH 7.4 (figure 3.13), $A\beta_{42}$ exhibit an interaction pattern mostly suggestive of a disordered state and limited β -sheet formation. The net charge removal when moving from pH 7.4 ($A\beta_{42}^{-3}$) to pH 5.4 ($A\beta_{42}^0$) thus promotes hydrophobic interactions between the different segments, which in turn leads to significant increase in β -sheet formation. We believe this reduction in electrostatic repulsion between the N- and C-termini to be an important driving force for the fast aggregation kinetics observed for $A\beta$ under slightly acidic pH [60–62].

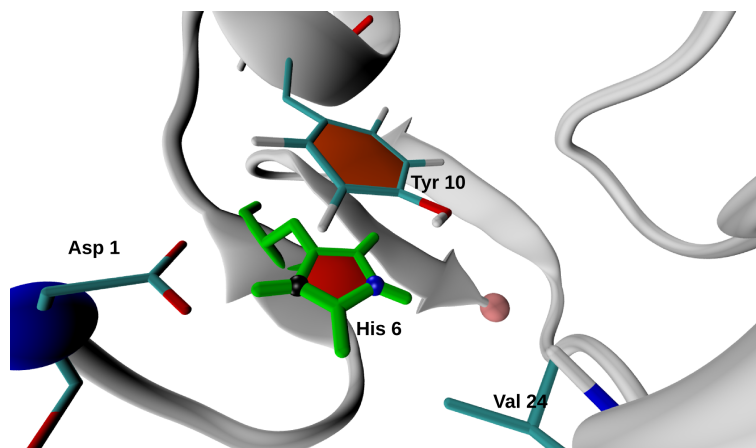


Figure 3.16: The role of His6 in promoting β -sheet in $A\beta_{42}$ monomer at pH 5.4. Contacts formed by His6 (green), Asp1, Tyr10 and Val24 are shown. The δ and ϵ -N atoms of His6 are shown as black and blue beads, respectively.

3.6 Conclusion

Using explicit solvent MD simulations on the microsecond time-scale and two GROMOS96 force fields, we investigated different influences affecting the secondary structure of the $A\beta$ peptide. We conducted a comparative study of $A\beta_{40}$ and $A\beta_{42}$ and also present a perspective on the intramolecular effects of the histidine protonation state on $A\beta_{42}$, explaining how this leads to the experimentally observed increase in aggregation kinetics at

acidic pH[60–63]. Our simulation results revealed that both peptides are mainly unstructured in the first 10 N-terminal residues. However, $A\beta_{40}$ is more disordered in this region than $A\beta_{42}$, which also samples either a π -helix or β -sheet in the N-terminal residues. The β -sheet-rich self-recognition motif $^{16}\text{KLVFF}^{20}$ was sampled in both peptides in our simulations but to a higher degree in $A\beta_{42}$. Both force fields reveal a loss of the C-terminal helix for both peptides, replaced by random coil and β -sheets. We observed the Gly37-Gly38 hinge structure in $A\beta_{42}$, which was identified as an important feature distinguishing $A\beta_{42}$ from its C-terminal truncated relative[56,57]. To validate the force field models, we calculated NMR chemical shifts using CamShift [78] and compared them to experimentally determined chemical shifts[65]. For the ffG53a6 simulations of $A\beta_{40}$ and $A\beta_{42}$, we found that the conformational sampling converged to an ensemble that is representative of the experimental data after 500 ns of simulation time. In addition, only the ffG53a6 suitably capture the rapid structural conversion expected for an intrinsically disordered peptide like $A\beta$. The ffG43a2 simulations, on the other hand, fail to completely reproduce the experimental NMR chemical shifts. The largest deviations were observed for the N-terminal half of $A\beta_{40}$ and $A\beta_{42}$, for which ffG43a2 predicts a rather stable α - and π -helix, respectively. The comparison between calculated and experimental chemical shifts allows us to conclude that in this region, ffG43a2 overestimates the stability of the N-terminal helix and generally underestimates the conformational flexibility of the $A\beta$ peptides.

We subsequently investigated the effect of pH on the structural dynamics of $A\beta_{42}$ monomer. Previous experiment had indicated the structure and aggregation of $A\beta$ are largely affected by pH. For instance by mutating selected residues in $A\beta_{40}$, a significant correlation was observed between net molecular charge and aggregation[63], while in the case of $A\beta_{42}$, a higher aggregation and toxicity was reported at pH 5.8 than at physiological pH[88]. In our *in silico* study, we modelled four pH states by selecting the ionisable states of $A\beta_{42}$ residues as expected under corresponding pH conditions. Our results show that,

under extreme pH conditions (pH 2.0 and pH 8.0) $A\beta_{42}$ structural dynamics suggests a two-domain interaction system with the highly polar N-terminal domain folded into helices, while the hydrophobic C-terminal segment samples β -sheet and coil structures. At both the isoelectric (5.4) and physiological (7.4) pH, on the other hand, $A\beta_{42}$ appears to be composed of a 3-folding units. At pH 5.4, the N- and the C-terminal domains cooperatively form an extensive network of β -sheets via His6 with the CHC. The double protonation of His6, His13 and His14 at this pH reduces the net peptide charge from -3 (pH 7.4) to zero, increasing intramolecular interactions by enhancing hydrophobic contacts. That is, the N-terminal segment becomes better able to interact with the highly hydrophobic C-terminal segment as a result of a lowering of electrostatic repulsion. The absence of these factors at pH 7.4, and the two other pH values, explains why $A\beta_{42}$ forms significantly high percentage of β -sheet only at the isoelectric pH.

3.7 Supplementary information

Chemical shift calculation with CamShift

Chemical shifts, conventionally obtained from NMR experiment, are powerful tools for determining the structure of organic compounds in particular[89,90]. The delicate dependence on the chemical environment, molecular geometry, and chemical bond types make them highly applicable in biological structure investigation for determining molecular conformation. Computational methods are increasingly exploiting this dependence on local configurational parameters to predict chemical shifts for biomolecular systems. Both *ab initio* and semi-empirical quantum chemical calculations have yielded good accuracy in predicting chemical shifts [91–97]. The high accuracy associated with quantum chemical approaches notwithstanding, their restriction to systems containing fewer number of atoms than typically encountered in proteins and nucleic acids precludes them from routine use in protein structure prediction.

To calculate chemical shifts, CamShift employs the intricate network of contributions within the 3-dimensional protein structure using a polynomial expansion of the interatomic distances[78]. Equation (3.2) depicts the different contributions used in approximating the dependence of the chemical shift on local configuration.

$$\delta_i^{calc} = \delta_i^{rc} + \delta_i^{sc} + \delta_i^{bb} + \delta_i^{\phi} + \delta_i^{nb} + \delta_i^{RC} + \delta_i^{HB} + \delta_i^{i-1} + \delta_i^{i+1} \quad (3.2)$$

The total chemical shift of atom i (δ_i^{calc}) involves separate terms representing contributions from the sidechain atoms (δ_i^{sc}), backbone atoms (δ_i^{bb}), the dihedral angle (δ_i^{ϕ}), nonbonded interactions (δ_i^{nb}), ring current effect (δ_i^{RC}), hydrogen bonds (δ_i^{HB}), neighbouring atoms (δ_i^{i-1} and δ_i^{i+1}), while δ_i^{rc} represents the chemical shift of atom i in random coil confor-

mation as obtained from NMR database. Equation (3.2) can be re-written as

$$\delta_i^{calc} = \delta_i^{rc} + \sum_{j,k} \alpha_{j,k} d_{j,k}^{\beta_{j,k}} \quad (3.3)$$

where $\alpha_{j,k}$ and $\beta_{j,k}$ (the value of which is 1 if covalent bond exists, otherwise 1 and -3) are parameters determined by the atom and residue under consideration. $d_{j,k}$ is the distance between atoms j and k . Experimental shifts from NMR database were subsequently employed for fitting each parameter. Benchmarking calculations using different protein test sets were reported to generate root mean square deviations from experimental data (table 3.4) that are comparable with the values obtained for SPARTA[97] and ShiftX[94], two leading chemical shift predictors[78].

Atoms	CamShift	SPARTA	ShiftX
N	3.01 (2.78)	2.87 (2.66)	2.87 (2.66)
HN	0.56 (0.56)	0.58 (0.53)	0.59 (0.55)
HA	0.28 (0.26)	0.30 (0.26)	0.30 (0.28)
C α	1.30 (1.22)	1.19 (1.03)	1.30 (1.14)
C β	1.36 (1.19)	1.30 (1.07)	1.42 (1.25)
C'	1.38 (1.12)	1.39 (1.05)	1.48 (1.19)

Table 3.4: How $\text{RMS}\delta^{calc}$ values for CamShift compare with the quantum chemical-based SPARTA[97], and ShiftX[94]. The predictions employed a test set comprising 28 proteins; the values appearing in parenthesis were obtained from 7 proteins.

Chapter 4

Alzheimer's disease: In search of treatment

4.1 AD drug discovery and mechanisms of toxicity

$A\beta$ toxicity features a plethora of biochemical process disruption that eventually culminates in AD. This perhaps can be attributed to the high structural diversity (both the aggregation and conformational states) of the $A\beta$ peptide. Different $A\beta$ assemblies are associated with different degrees of cellular toxicities: this suggests the existence of either toxicity mechanisms varying from one $A\beta$ assembly to the other, or simply a varying degree of the same toxicity pattern. The exceedingly high structural complexity associated with $A\beta$ notwithstanding, damage of neuronal and synaptic structures is recognised as the common significant theme characterising $A\beta$'s involvement in AD. This results from a diverse disruptive action of the amyloid peptide on subcellular processes. On cell membranes and membrane channels electrophysiological studies including the long-term potentiation (LTP) measurements [98, 99] revealed a Ca^{2+} (in particular) homeostasis disruptive effect resulting in increased Ca^{2+} influx, excitotoxicity and eventually synaptoneuronal death[100–103]. These effects are mediated both by the structural interaction

of $A\beta$ with membrane lipids and via the action of Ca^{2+} -, K^{1+} - and glutamate (NMDA) receptors[100,104]. Additional mechanisms include induction of oxidative stress and formation of highly toxic radicals, and a caspase-derived apoptotic process which amongst others result from high Ca^{2+} influx and mitochondrial destruction[105,106]. The pathology of Alzheimer's disease is characterised by a host of disease influencing factors and many more underlying biochemical components of the associated neuronal degeneration. The implication of this is in two respects. First, the complexity of the disease mechanism implies the possibility of using different pharmacological agents in the management of the disease. These drugs are typically aimed at *correcting* defined aspects of the toxicity complex. While these are currently employed for symptomatic management, the complexity of the pathology implies the likelihood of pharmacotherapy burden on the patient since different agents are required to manage different aspects of the disease. Secondly it implies the presence of multiple targets that can be considered in the search for definitive cure. Drug development efforts are currently being undertaken on two fronts including the search for biomarkers needed for facilitating early diagnosis, and the search for compounds capable of altering the identified underlying molecular causes of AD.

4.2 Current treatment strategies

For the pharmacological management of AD it is often necessary to make a distinction between factors immediately responsible for disease signs and symptoms from those initiating the entire pathology. While the latter represent the ideal in AD management, the former are currently of crucial immediate importance for initiating treatment plans that assuages patients discomfort. In this regard, Terry *et al.* suggested in 1991 that the degree of synaptic function loss should be considered in staging Alzheimer's disease with respect to cognition loss rather than the two accepted hallmarks of the disease, that

is, senile plaques and neurofibrillary tangles [107]. The cholinergic, glutaminergic and GABAergic synaptic neuron networks have been observed to be significantly impaired as a result of the $A\beta$ -induced neurotoxicity [108, 109]. Following the establishment of a relationship between AD and cholinergic function loss, the cholinergic hypothesis was proposed in the early 1980s [110]. According to the hypothesis, cognitive decline observed in AD pathology is a consequence of compromised cholinergic neurotransmission in both the cerebral cortex and the forebrain. This belief then prompted the search for compounds with cholinesterase-blocking activities to halt the depletion of the CNS levels of acetylcholine.

The major factor in the advent of treatment with cholinesterase inhibitors was the *post mortem* detection of a significant reduction of the levels of biomarkers associated with the cholinergic neurons in severe Alzheimer's disease brains. Clinical trials [111–114] were conducted starting with the natural anticholinesterase agent physostigmine obtained from the plant *Physostigma venenosum* Balf [111]. The initial trials were plagued with significant gastrointestinal adverse drug reactions especially in the case of the first-generation cholinesterase inhibitors, physostigmine and tetrahydroaminoacridine (tacrine), and additional hepatic reactions observed in a small number of patients in the case of the second agent. In 1993, Tacrine became the first drug treatment to be approved by the Food and Drug Administration (FDA) for the management of Alzheimer's disease [114–117]. Three second-generation cholinesterase inhibitors are currently approved for the treatment of mild-to-moderate AD, and these are donepezil, galantamine, and rivastigmine. They all act either as reversible (donepezil, the first drug in this category to be licensed for use in AD, and galantamine) or *pseudoreversible* [117] (rivastigmine) inhibitors of acetylcholinesterase and butyrylcholinesterase (rivastigmine). Huperzine-A is additionally available for use in China. Memantine, licensed for the treatment of moderate-to-severe Alzheimer's disease, regulates the central nervous system's (CNS) concentration of the

neurotransmitter glutamate by blocking the NMDA (N-methyl D-aspartic Acid) receptor. The concurrent administration of both memantine and the cholinesterase inhibitors has increasingly been employed to take advantage of their complementary activities.

The symptomatology of Alzheimer's disease is composed of both cognitive and noncognitive symptomatic presentations. While the greater percentage of drug trials are directed at evaluating the efficacy of cholinesterase inhibitors on cognitive presentations, the noncognitive symptoms—especially in the advanced disease—are predominantly responsible for hospital admissions. Clinical management of the symptoms of Alzheimer's disease with the cholinesterase inhibitors and memantine, while it in many cases has modestly improved the quality of life of sufferers, it has conspicuously left the course of the disease unaffected. Literally, the import of this is that once diagnosed of AD the individual progresses through phases of progressive cognitive and neurological decline and ultimately to death with no treatment capable of slowing down, let alone aborting the course. The past few decades have witnessed increasing research efforts both in the academia and in the industry specifically focussed on designing definitive therapies capable of modifying the disease path. Some of such efforts have advanced to preclinical and clinical phases of testing, of which a significant percentage have failed either as a result of insufficient proof of efficacy or serious toxicities.

The current search for a definitive treatment of Alzheimer's disease is principally based on the amyloid cascade hypothesis and biochemical events secondarily triggered as a result. Drug development focussing on the $A\beta$ -centric pathway maybe broadly categorised into compounds directed at:

1. $A\beta$ peptide production from APP
2. $A\beta$ misfolding and aggregation
3. $A\beta$ clearance.

The first two points are discussed in detail.

4.2.1 $A\beta$ peptide production from APP

The enzymatic cleavage of the amyloid precursor peptide into $A\beta$ is widely accepted as the foremost step in the subsequent neurotoxicity complex that characterises Alzheimer's disease and designing drug candidates that target this step has all along constituted a viable drug design approach. The β -secretase which cleaves the N-terminal segment of $A\beta$ and the γ -secretase that cleaves APP at $A\beta$'s C-terminal segment thus represent an attractive target for drug design. As a further indication of the flurry of interest in this drug design school of thought, the past few decades have witnessed significant research efforts involving both academia and pharmaceutical industries out of which some inhibitors have reportedly been identified with inhibitory activities[118–124]. Targeting the secretases in developing Alzheimer's disease treatment is however not without serious drawbacks. In the first instance, there are strong indications that both enzymes are involved in other biological processes making the business of inhibiting the secretases a very tricky one. Apart from APP cleavage, the enzymatic activity of the γ -secretase, for example, also includes the transmembrane Notch growth hormone receptor, the voltage-gated Na channels, the vascular endothelial growth factor (VEGF), the CD44, and a number of other endogeneous substrates[125–129]. Inhibiting γ -secretase thus carries the potential risk of disrupting important biological functions. In the case of β -secretase, the solved structure of the enzyme revealed an extensive binding site that in practical terms precludes the design of small-molecule non-peptidic inhibitors capable of disrupting the APP-processing function of the enzyme[130]. This however did not completely discourage search into prospective BACE-1 inhibitors, some of which reportedly inhibit APP processing in animal models[131, 132].

4.2.2 $A\beta$ aggregation inhibitors

Alzheimer's disease pathology of synaptoneuronal loss is characterised by the presence of amyloid plaque (mostly extracellular) and neurofibrillary tangles (intra-axonal) in the brain of the sufferers of the disease. Of the two causative peptides, $A\beta$ currently takes the centre stage in research spending, both facilitated by and resulting in the availability of a large body of data correlating structural features (e.g., the formation of soluble oligomer) with neurotoxicity. The $A\beta$ composite conformational system is characterised by the presence of a dynamic structural equilibrium between the different aggregation states of the peptide consisting of monomers, soluble oligomeric intermediates and the insoluble fibrils[133,134]. While the specific details of the structural events accompanying the conversion from the intrinsically disordered monomer to the β -sheet-dominated higher order aggregates remain unclear, there are evidences that the neurotoxicity cascade characterising AD is linked to the formation of soluble $A\beta$ oligomers[135]. This however does not singularly attribute neurotoxicity to β -sheet formation alone, otherwise the original amyloid hypothesis identifying the high- β -sheet-containing fibrils as the basic neurotoxic unit in AD would have not required revision. Instead, AD neurotoxicity appears to result from a highly complex interplay between $A\beta$ solubility, β -sheet content, peptide sequence, as well as prevailing factors in the local brain environment.

The paucity of knowledge on the aggregation pathway notwithstanding, the misfolding and aggregation of $A\beta$ peptides provide a viable mechanism being explored in AD drug discovery [136] In principle, the gamut of molecular events surrounding the $A\beta$ peptide's involvement in Alzheimer's disease present possible targets that may be considered for treatment. And the conformational aspects of the peptide have been extensively explored[137,138]. Such efforts have included the in vitro screening of large chemical libraries [139], an approach traditionally employed in pharmaceutical industries which has the benefit of generating viable hits with desirable physicochemical properties. This ap-

proach successfully identified small molecule aggregation inhibitors belonging to different chemical classes including curcumin[140], melatonin[141], tannic acid [142], trehalose[143], nicotine[65], and dapsone[144], to mention a few. While the small size of this category of chemical compounds make them ideal for optimisation, this feature at the same time is what disqualifies most of them from progressing beyond the initial in vitro studies onto clinical testings. This largely results from the absence of binding specificity, a challenge further attributed to the fact that the $A\beta$ peptide systems represents a structurally ill-defined target, with such dynamic conformational aspects and polymorphism that it is practically impossible for such small molecules to effectively arrest aggregation.

One approach that circumvents the challenge with small molecular inhibitors is the immunotherapeutic method employing both active and passive immunisation for clearing endogenous $A\beta$ [45, 145]. In active immunisation sequence fragments of the $A\beta$ peptide are injected into subjects to stimulate antibody production, while passive immunization involves the injection of prepared anti- $A\beta$ antibodies [45, 145]. Immunotherapy eliminates the size-dependent lack of binding specificity associated with small molecule drugs. And both cognition and plaque deposition were reported to be improved in trials involving APP transgenic mouse models. Plaque clearance, however, appeared to be significantly compensated for in the brain by a tripling and quadrupling of the quantity of cerebral amyloid vascular deposits[146]. Occurrence of microhaemorrhages in a small but significant amount of human subjects subsequently led to the stoppage of clinical testing. However, reports emerging from a subset of treated patients strongly suggest a relative improvement of the cognitive decline profile[147, 148].

Lastly, an entirely different AD drug discovery philosophy known as the rational drug discovery approach, makes use of insight into $A\beta$'s structure and aggregation dynamics in the design of aggregation inhibitors. At the most fundamental level, this approach utilizes knowledge of the identified critical *folding* ($A\beta$ being a disordered peptide cannot be said

to form a unique 3D structure) units of the monomers which, importantly, are also critical for oligomeric assembly. The various research efforts in this area are discussed under the following subcategories that should not be taken as illustrative and not as an exhaustive treatment of the topic.

The A β self-recognition unit: sequence

The self-recognition unit in the A β peptide is made up of the CHC residues ¹⁶KLVFF²⁰ which is crucial in the amyloid misfolding event and believed to kick-start the entire aggregation-dependent neuropathology [138]. The first attempt to target the self-recognition unit was by Tjernberg et al in a report published in 1996 [76]. This pioneering work paved way to subsequent designs employing variants of this approach to design aggregation inhibitors, also known as β breakers. In the work by Tjernberg *et al.*, the self-aggregating self-recognition unit was synthesized as part of a peptide that was demonstrated to inhibit fibrillation of the full peptide. Expectedly, the strong aggregation propensity of this peptide coupled with its susceptibility to peptidase degradation prevented it from being considered for further development. However, this work served to demonstrate the viability of pursuing this mechanism in the search for amyloid aggregation inhibitors. In addition, a small number of peptide-based inhibitors were also reported, where the proposed mechanism of action is the stabilization of the peptide in a helical conformation via binding to residues 13–23, which includes the self-recognition unit [149]. These compounds were reported to frustrate assembly into neurotoxic aggregates of the amyloid peptide [149].

The A β self-recognition: hydrophobicity

Since the high intrinsic hydrophobicity of the A β peptide is an important factor in driving the undesirable aggregation, it was hypothesized that by attaching a charged unit to the self-recognition unit, the effective hydrophobicity of the self-recognition unit could be reduced. The self-recognition unit provided a good template for in this instance, mostly because of its critical role as the *recognition element* in fibril assembly[150, 151]. To confirm this hypothesis, Ghanta *et al.* incorporated a poly-K(i.e, polylysine) fragment to the C-terminus of A β _{15–25}, resulting in significant reduction in A β ₃₉'s amyloidogenic properties [150]. Inhibitors with poly-K units inserted via the N-terminus of the self-recognition unit were reported to be ineffective or mildly amyloidogenicity enhancing. In addition to this outcome, in 1999, Pallitto *et al.* published their findings on the activities of hybrid peptides containing poly-K units at the C-termini of a number of short self-recognition unit-containing A β peptides[152]. They observed that the hybrid peptide's effect on A β 's toxicity and aggregation profile are related. Interestingly, however, these compounds also significantly increased A β 's aggregation kinetics leading to the suggestion that the observed lowering of A β toxicity might have resulted from a facilitated conversion of toxic A β oligomers into relatively less toxic fibrils. In these studies, however, a minimum of three lysine residues were required for the observed anti-amyloid activities.

The A β self-recognition: disruption of amyloid binding pattern

Another way to disrupt the amyloid aggregation process of the self-recognition unit is by incorporating bulky groups capable of sterically interfering with intermolecular contacts necessary for cross- β structure formation. Findeis *et al.* in 1999, designed a range of hybrid aggregation inhibitors incorporating the bulky cholic acid at the N-terminus of the CHC peptide LVFFA. *In vitro* assays revealed marked lowering of A β ₄₀ and A β ₄₂'s aggregation and toxicity [153, 154]. However, administration of the inhibitors to rats

revealed significant hepatic clearance, which effect was pronounced in inhibitors with L-amino acids. The D-analogues were observed to possess higher hepatic stability [153,154]. Another bulky group that has been used involves the methylation of the amide *N* on alternate residues of $A\beta$ peptide fragments which both inhibited fibril formation as well as dissolved mature fibrils [155,156]. The incorporation of the amino acid proline within the self-recognition unit sequence was also shown to produce inhibitors capable of interfering with $A\beta$ fibrillization [157]. Since both N-methylation and proline lack the amide *N*, they most likely produce the observed inhibitory effect via a disruption of intermolecular hydrogen bonds.

The $A\beta$ self-recognition: stereochemistry

As promising as these peptide-based molecules are, they generally suffer from a drawback relating to their rapid clearance from the circulation, especially by endogenous peptidases. This drawback is partly mitigated by the incorporation of bulky non-peptide groups in the peptides; by doing so this confers on the inhibitors a measure of resistance to protease activities. A second approach employs inhibitors based on D-enantiomeric amino acid residues. Certain D-Peptides have been found to inhibit both $A\beta$ aggregation and neurotoxicity [158–164]. These include D-peptide sequences constructed based on the key amyloidogenic motif (the self-recognition unit), $^{16}\text{KLVFFA}^{21}$, which in addition to the desired fibril-disrupting abilities and protease-resistance, also exhibit little to no immunogenicity compared with the L-peptide-based counterparts [161]. This is the approach being pursued in the present work, and the design aim is to realise aggregation inhibitory effects, protease resistance, as well as an non-immunogenicity in the same compounds.

Some years ago, a 12-residue D-peptide, code-named D3, and with the sequence Arg-Pro-Arg-Thr-Arg-Leu-His-Thr-His-Arg-Asn-Arg, was designed using mirror-image phage dis-

play selection in the group of Prof. Dieter Willbold, the head of the Institute of Complex Systems 6, Forschungszentrum Jülich [162]. The peptide demonstrated amyloid aggregation inhibitory effect, a dose-dependent reduction in plaque load, as well as improvement in cognitive function in transgenic mice [162,163]. Rather than merely disassemble preformed fibrils, the peptide is believed to drive the aggregation path through a nontoxic pathway in which D3 precipitates the formation of nontoxic amorphous Thioflavin-negative aggregates [162].

4.3 Aggregation inhibition with D3

In mirror-image phage display, a technique described by Mayr *et al.*, the target protein is converted into its D-enantiomer via chemical synthesis followed by phage libraries screening for strongly binding L-peptide ligands [165]. Synthesis of the D-enantiomeric form of the selected ligands should then produce D-peptide ligands with desirable binding properties for the original target L-protein. The D3 peptide was discovered in a similar fashion [162]. Using highly dilute solutions of synthetic D-A β_{42} (2 nM under which solution condition low order A β_{42} assemblies are expected to exist) as target protein, an entire phage library of more than one trillion 12-residue peptides was screened for high-binding ligands, a biopanning process that was conducted six times thereby enriching the phage. Sequence determination of the predominant peptide in the enriched phage identified D3 as constituting about 40% of randomly selected phage peptides, with the remaining high-affinity ligands displaying greater than 75% sequence identity to D3[163].

Both *in vitro* and *in vivo* experiments were performed to determine the effect of D3 on A β_{42} aggregation and toxicity. *In vitro* determination of amyloidogenic properties of amyloid proteins in most cases utilises Thioflavin T (ThT) fluorescence assay in which amyloid property is determined by the characteristic fluorescence enhancement and red shift exhibited by Thioflavin T, the chloride of 4-(3,6-dimethyl-1,3-benzothiazol-3-ium-2-yl)-

N,N-dimethylaniline, upon binding to ordered β -sheet-rich amyloid structures [166–168]. In addition, an apple-green birefringence upon staining a peptide system with Congo Red (Sodium 3,3'-([1,1'-biphenyl]-4,4'-diyl)bis(4-aminonaphthalene-1-sulfonate)) is also indicative of amyloid properties. Both ThT and Congo Red assays revealed significant and concentration-dependent reduction in $A\beta$ amyloidogenic properties when treated with D3 [162, 163].

Following treatment with D3 using both the oral and direct hippocampal administration routes, transgenic mice expressing human APP and presenilin-1 genes were reported to show significant improvement in cognition over the untreated group as determined by the Morris water maze experiment. Histological assessment using the frontal cortical and hippocampal brain sections further revealed significant reduction in both plaque load and plaque-dependent inflammation in the treatment group [164]. Follow-up measurements aimed at determining the possible mechanisms of D3/ $A\beta$ aggregation inhibition included $A\beta$ particle size determination by dynamic light scattering (DLS). These experiments revealed an increased average particle size in $A\beta_{42}$ solution systems from amyloidogenic 300 Å hydrodynamic radius (without D3) to large amorphous aggregates of about 800–7000 Å hydrodynamic radius (with D3) that were both ThT- and Congo Red-negative. This result was further confirmed by electron microscopy (EM). In other words, treatment with D3 peptide abolishes $A\beta$ toxicity and converts the neurotoxic water-soluble oligomers into non-toxic, non-amyloidogenic and non-fibrillar aggregates but without increasing the concentration of $A\beta$ monomer.

Chapter 5

Computational study of D3's anti-amyloid properties

5.1 Overview

The simulations performed in this section rest on the background of insight obtained from the conformational study of $A\beta$ discussed in chapter 2.

A general outline of the study plan is first presented here. The overarching aim of the studies conducted in this section is to unravel the mode of action. Two aspects of the D-peptide's inhibitory activities are important: D3 binding and binding *efficacy*. Investigation into the strength and specific nature of the $A\beta$ /D3 binding requires identifying residues within both peptides that are responsible for the interaction. Here, a global optimisation method implemented in the GMIN programme [169], turned out to be particularly useful in this respect because it allows for the generation and optimisation of protein complexes, taking into account translational, rotational and conformational degrees of freedom (flexible docking). The methodology description below includes a detailed introduction to the GMIN simulations. The optimised $A\beta$ -D3 complexes were subjected to a rigorous scoring, in which the binding energy for each complex was computed, and

then ranked. The most favourable complexes based on binding energy were selected for further processing using explicit solvent MD.

The term binding *efficacy*, used as a general analogy to pharmacological efficacy, describes the effect of D3 binding on A β . In order to have an idea of how well the calculations agree with experimental measurements, changes in secondary structure upon D3 binding have been monitored. This makes it possible to take the description of binding event beyond mere binding to what may be described as *effective binding*. This becomes particularly useful in the section dealing with the investigation of different lengths of a close relative of D3 where strong binding was observed to not automatically translate to efficacy. We have employed explicit solvent MD simulations together with appropriate control systems to quantify the effect of binding on A β . The applicability of explicit solvent MD is however not limited to quantifying efficacy, it is also used for decomposing the interaction energy into residue–residue contributions. The resulting energy matrix provides an indication of the stability of the system, and more importantly of the relative significance of the individual residues in driving the inhibitor-A β peptide interactions.

5.2 Methodology

5.2.1 Peptide models

D3 peptide

D3 is a twelve-residue peptide with the sequence ¹RPRTLHTHRNR¹⁰ containing all amino acid residues in the D-enantiomeric configuration. To obtain suitable GROMOS53a6 [9] and CHARMM22 [170] force field models, for each residue the chirality of the backbone centred around the C $_{\alpha}$ atoms was inverted in congruence with each force fields topological convention. In GROMOS53a6 this entailed inverting the order for the improper dihedral

entry from "CA N C CB" for L-amino acids to "CB N C CA" for the D-enantiomer. In CHARMM22 it was required to invert the sign for the improper dihedral angle formed by "N C CA HA" from -117° to $+117^\circ$. New D-amino acid entries were thus created in both force field for all twenty amino acid residues using new residue names essentially involving prefixing the old names with a "D" and dropping the last letter. For instance, D-leucine has the name "DLE" in our model.

Starting with a fully extended peptide configuration for D3 generated with the CHARMM simulation package [171], the conformational space of the peptide was studied in a 100 ns-MD simulation conducted with GROMACS [83] using the ffG53a6 force field [9]. The extended peptide was first subjected to a cycle of *in vacuo* steepest descent minimization and then inserted in a cubic water box (SPC water model) using a 10 Å separation distance from the peptide to the box edges. Boundary treatment involves surrounding the simulation box with its periodic images (i.e., periodic boundary condition). A 14 Å cutoff was used in truncating short-range forces, while the long-range component of the Coulombic forces was treated using the particle mesh Ewald method. Sufficient numbers of Na^+ and Cl^- were included to neutralise the net +5 peptide charge emanating from the five arginine residues and at the same time to achieve a 150 mM NaCl concentration. After further minimisation steps involving steepest descent followed by conjugate gradient minimisation, an equilibration dynamics in the NPT (24833 atoms, 1 atm, 300K) ensemble was performed using a 239 kcal/mol restraining force on peptide heavy atoms. This lasted 1 ns, and it served to relax the water molecules around the peptide in order to remove high forces resulting from atomic clashes while at the same time removing pockets of empty space resulting from the random process by which the water molecules were added to the box. After the position restraints were turned off, a 100 ns production MD run was performed under simulation conditions essentially the same as in the equilibration step, except that a Nosé-Hoover thermostat was used instead of a Berendsen

thermostat. In addition pressure coupling made use of the Parrinello-Rahman barostat in place of the Berendsen barostat. The neighbour list was updated every 10 ps while new coordinates and energies were saved every 20 ps. The resulting trajectory, containing

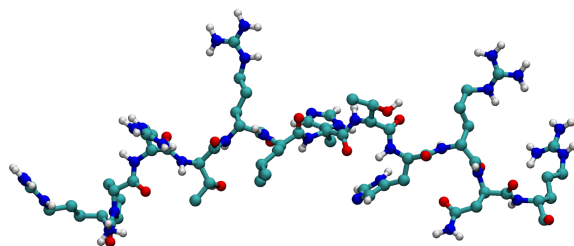


Figure 5.1: Most populated D3 structure.

5000 configurations, was subjected to cluster analysis using a 20 Å RMSD cut-off. The largest cluster, for which a representative configuration is shown in figure 5.1, features an extended structure. The adopted conformation of D3 is driven by an effort to minimise electrostatic repulsion, which requires a careful spatial arrangement of the sidechain groups. This is however not unexpected, especially given that the peptide contains no more than twelve residues out of which five are positively charged. The central structure of the largest cluster (figure 5.1) was employed for all subsequent interaction studies involving $A\beta$ peptides and is hereafter referred to as "D3" peptide.

$A\beta$ monomer

The five largest cluster centres obtained from the 1.5 μ s MD simulation of $A\beta_{42}$ (force field ffG53a6) were independently used as the starting structure for studying the interaction of D3 with $A\beta_{42}$ monomer. The choice of five conformationally different $A\beta$ structures (figure 3.8) ensures structural variability to account for the structural flexibility of the peptide in solution. This helps to limit the error of conformationally biasing the binding had a single conformation been employed for $A\beta_{42}$. In addition, the five cluster centres

were obtained from the equilibrated part of the trajectory as determined by experimental chemical shift validation.

A β pentamer

The coordinate file PDB 2BEG of an A β_{42} pentamer solved by solid state NMR was taken from the www.rcsb.org database [172]. The structure resulted from solid-state NMR investigation of A β_{42} fibril which suggested the pentamer as the β -sheet structural unit of the fibril. In this structure, A β_{42} samples a strand-turn-strand conformation with residues 18–26 and 31–42 forming an in-register β -sheets. The N-terminal residues 1–17 were, however, disordered and lacked coordinates. Starting from 2BEG, residues 1–17 taken from the A β_{42} monomer MD simulation and sampling a coil conformation were *stitched* onto each of the five peptide units composing the pentamer. A coil conformation was decided for the attached N-terminal residues 1–17 based on the experimental observation that this section of the peptide is disordered [172]. Furthermore, the generated structural construct was thereafter subjected to global optimisation with basin-hopping in which the positions of residues 18–42 were fixed while the added residues 1-17 of each peptide were perturbed until all atomic clashes were resolved and a low-energy structure was obtained. Several cycles of basin-hopping were conducted before a sufficiently disordered N-terminus was obtained in each of the five chains. This also necessitated the alternate freezing of the peptide chains. For example, the freezing of the coordinates of chains 1, 3 and 5 allowed chains 2 and 4 to be perturbed until they were adequately separated from the frozen chains. Then chains 2 and 4 were frozen while chains 1, 3, and 5 were perturbed. This was repeated until the total energy of the entire system converged. Thereafter, 500 ns explicit solvent MD using ffG53a6 was performed on the full-length A β pentamer at 300 K and 350 K. The simulations were performed in the absence of any form of restraint so as to determine whether or not the new model would remain stable. The full-length

pentamer remained stable at both temperatures. A clustering analysis with a 1 Å RMSD cut-off was subsequently performed on the MD trajectory at 350 K using the C $_{\alpha}$ atoms of residues 18–42 of each peptide for least squares fitting. The representative structure of the most populated cluster, hereafter referred to as *A β pentamer* or simply *pentamer*, is shown in figure 5.2. This structure was then selected as the starting structure for the A β_{42} pentamer-D3 interaction studies using GMIN [169].

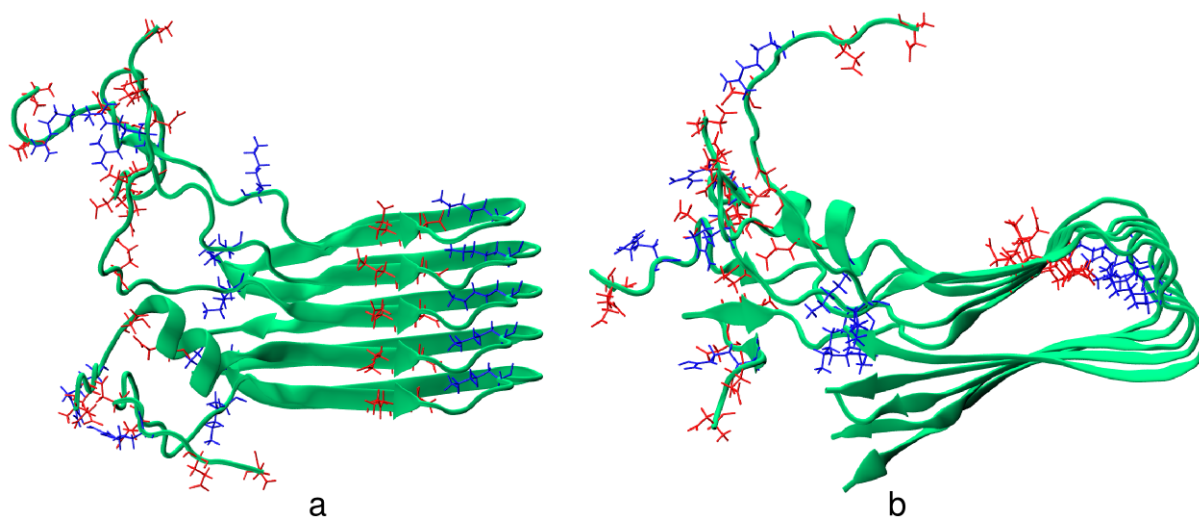


Figure 5.2: Full-length A β_{42} pentamer in VMD green NewCartoon representation showing the top view (a) and the sideview (b). The sidechain groups of the charged amino acid residues are represented with liquorice coloured according to charges—red and blue respectively for acidic and basic amino acid residues. The newly modelled residues 1–17 sample a disordered state and the fibril-stabilising salt bridges involving Asp23 and Lys²⁸ are visible in (b) while Glu22 lies above the fibril plane.

5.2.2 Generation and optimisation of A β complexed with D3

Basin-hopping (BH) simulations

The basin-hopping (BH) approach to global optimization [174, 175] is equivalent, in principle, to the Monte Carlo-minimization method [173]. It involves proposing moves by perturbing the current geometry, which are either accepted or rejected depending on the difference between the energy of the local minimum of the instantaneous configuration following energy minimisation, and the energy of the previous local minimum in the chain. What results in essence is that the typically complex potential energy surface (PES) is

transformed into what has been referred to as *the basins of attraction* [174, 175] consisting of all the local energy minima. That is, the energy for configuration \mathbf{R} can now be expressed as,

$$\tilde{E}(\mathbf{R}) = \min\{E(\mathbf{R})\} \quad (5.1)$$

where *min* signifies energy minimisation. With the transformed energy surface, it becomes feasible to take large steps, especially since the aim is to step between the energy minima in search of the global configuration. In addition, the need to maintain detailed balance while taking the steps between the minima vanishes since the BH method is intended for locating the global minimum, and not for sampling thermodynamic properties. The BH algorithm has been implemented in the GMIN program [169] employed in the present work for finding the global minimum of D-peptide- $A\beta$ complexes. In all the interaction studies performed using GMIN, the $A\beta$ peptide(s) represent the *receptor* molecule(s) whose coordinates were fixed while the D-peptides represent the *ligand* molecules.

$A\beta$ monomer/D3

With the aid of the oligomer-generation procedure [176] implemented into GMIN, 1000 $A\beta_{42}$ -D3 binary complexes were generated for each of the five $A\beta_{42}$ starting structures, i.e. a total of 5000 $A\beta_{42}$ -D3 complexes. The structure of D3 was taken from a 100 ns explicit solvent MD simulation described above. The C-terminus was amidated to mimic the employed experimental condition in which D3's C-terminus was attached to a surface (personal communication with Prof Dr. D. Willbold).

The generated complexes were subsequently optimized using 300 BH steps in which dihedral angle moves [177], small rigid body rotations and translations [178] were applied to D3 peptide. This involves a twisting of the dihedral angles to a maximum of 30 ° with a 30% twisting probability while the rigid body rotation and translation were performed,

respectively, to a maximum of 80° and 2 \AA . Parameters from the CHARMM22 force field were employed for treating interacting systems [170], while the generalised Born solvent model FACTS [179] was used in modelling the aqueous environment. The reason for using the CHARMM force field for the GMIN simulations, rather than the GROMOS force field employed in the MD runs, was because the latter force fields are yet to be interfaced to the GMIN programme. In addition, the CHARMM force fields had not yet been implemented in GROMACS at the time of commencement of the study.

A total of 5000 $A\beta_{42}$ monomer-D3 complexes were generated and optimised, after which the binding energies ($\Delta\Delta G_{bind}$, section 5.2.3) were calculated and used in ranking the predicted complexes. The top-ranking 100 complexes were subsequently submitted to explicit solvent MD simulations described below.

$A\beta$ pentamer/D3

A protocol similar to that described for the $A\beta_{42}$ monomer-D3 system was employed for the study of the full-length $A\beta_{42}$ pentamer complexed with D3 (stoichiometry $A\beta_{42}:D3 = 5:1$). The employed simulation parameters were also similar, except few changes were necessary to account for the bigger size of the system. To this end, a total of 400 BH steps were performed. Rigid body rotation (up to a maximum of 90°) and translation (up to a maximum of 2 \AA) were applied to D3 only at every BH step. In addition, D3's dihedral angles were also twisted with a probability of 30% up to a maximum of 30° . A total of 4000 $A\beta_{42}$ pentamer-D3 complexes were generated and optimised. The predictions were subsequently ranked according to $\Delta\Delta G_{bind}$ (calculated as described in section 5.2.3). The 100 predictions with the best $\Delta\Delta G_{bind}$ values were then selected for explicit solvent MD simulation.

5.2.3 Calculation of binding energies

To estimate the binding energy ($\Delta\Delta G_{bind}$), the idea of the MM/GBSA formalism was employed with force field parameters derived from the CHARMM22 potentials, and the FACTS implicit solvent parameters employed for the GBSA part:

$$\Delta\Delta G_{bind} = \Delta\Delta G_{int} + \Delta\Delta G_{Sol} \quad (5.2)$$

The $\Delta\Delta G_{bind}$ was calculated as a sum of contributions from interpeptide interactions and solvation terms. The interaction energy includes van der Waals and electrostatic interactions between A β and D3, while the solvation term contains the polar (electrostatic) and nonpolar solvation terms:

$$\Delta G_{int} = \Delta G_{int}^{vdw} + \Delta G_{int}^{ele} \quad (5.3)$$

$$\Delta\Delta G_{Sol} = \Delta\Delta G_{Sol}^{pol} + \Delta\Delta G_{Sol}^{npol} \quad (5.4)$$

$\Delta\Delta G_{Sol}^{pol}$ and $\Delta\Delta G_{Sol}^{npol}$ were each computed by removing the individual terms for A β ₄₂ and D3 from the energies of the complex:

$$\Delta\Delta G_{Sol}^{pol} = \Delta G_{AB/D3}^{pol} - \Delta G_{AB}^{pol} - \Delta G_{D3}^{pol} \quad (5.5)$$

$$\Delta\Delta G_{Sol}^{npol} = \Delta G_{AB/D3}^{npol} - \Delta G_{AB}^{npol} - \Delta G_{D3}^{npol} \quad (5.6)$$

$\Delta\Delta G_{bind}$ was computed for each of the optimised 5000 A β monomer-D3 complexes and 4000 A β pentamer/D3 complexes, and the obtained data used in ranking the predictions. The top one hundred with the most favourable energies $\Delta\Delta G_{bind}$ were selected for explicit solvent MD simulation as discussed below. Within MM/GBSA, $\Delta\Delta G_{bind}$ is commonly referred to as the *binding free energy*, which use has also been largely adopted in this

thesis. However, it is important to point out that only $\Delta\Delta G_{sol}$ is a free energy, whereas $\Delta\Delta G_{int}$ is a potential energy (also in MM/GBSA).

5.2.4 Explicit solvent MD

The final explicit solvent molecular dynamics simulations performed with the GMIN-predicted A β -D3 complexes serve different purposes. The first is confirming the stability of the bound configurations, which is tested by calculating the backbone RMS displacement from the starting (GMIN) structure. The MD trajectory also provides a means of obtaining a time-averaged description for the energetics of the bound systems. Lastly, it enables the assessment of the effect of binding (i.e., the binding efficacy) on peptide structural features. Two groups of MD simulations were conducted: simulations of A β -D3 complexes and control simulations of A β alone. All MD simulations were performed with the SPC water model, employing the ffG53a6 force field and a protocol similar to the one described for D3 conformational sampling above in section 5.2.1. In the simulations of A β monomer-D3, each of the 100 GMIN-generated configurations with the lowest $\Delta\Delta G_{bind}$ values was subjected to a 10 ns MD run. With respect to the A β pentamer-D3 systems, the 100 best $\Delta\Delta G_{bind}$ configurations were first parsed through a clustering procedure to reduce the geometric redundancies, as D3 binds more specifically to the ordered pentamer structure than the flexible monomer A β . As an example, there are a number of A β -D3 complexes in which D3 binds to similar sites in the disordered N-terminal section of the pentamer. The cluster analysis involves a least-square fitting to the A β_{42} pentamer atoms and a clustering of D3 placements around the pentamer. The choice of a 2.19 Å RMSD cutoff distance yielded 15 clusters in which redundancies in D3 binding to A β pentamer were largely eliminated. 50 ns MD simulation was initially performed on one the cluster centres after which this was extended to 100 ns. After it was observed that extending the simulation to 100 ns did not grant additional information, 5 other cluster centres were

each subjected to 50 ns MD simulation.

The second set of MD simulations are control simulations of $A\beta$ peptides in the absence of D3. This, in the case of $A\beta$ monomer, entails independently performing a hundred 10 ns MD simulations. The starting structures employed were obtained by stripping each $A\beta_{42}$ monomer-D3 complex off its D3 binding partner. The unbound $A\beta$ were then subjected to explicit solvent MD simulation following the setup described in section 2. In the case of $A\beta_{42}$ pentamer, a single control MD simulation (50 ns) the $A\beta$ pentamer (figure 5.2) was performed. This was deemed sufficient because the same starting structure was employed for generating the 4000 $A\beta_{42}$ pentamer-D3 complexes.

5.3 Results and discussion

5.3.1 $A\beta_{42}$ monomer/D3

Before discussing the results obtained for the $A\beta_{42}$ monomer-D3 complexes it is useful to note that the choice of five different $A\beta_{42}$ configurations for the global optimisation was a necessary attempt to capture the high conformational flexibility exhibited by $A\beta_{42}$ monomer as discussed in chapter 2. Unlike most folded peptides whose conformational space may be represented by an average structure with essentially invariable backbone architecture, the $A\beta_{42}$ monomer conformational space is characterised by its ability to rapidly interconvert between different secondary structures. This property is factored into the present setup by employing five different representative structures (cluster centres). The employment of a single D3 configuration is deemed sufficient since all sampled states principally consisted of extended conformations driven by the need to increase the distance between the five arginine residues and thereby minimising electrostatic repulsion. Finally, the choice of one hundred binding poses for more detailed analysis serves to mimic

experimental situations where thousands of molecules are involved in peptide-peptide association. Ideally, the entire set comprising 5000 D3- $A\beta_{42}$ binary complexes obtained from GMIN should have been subjected to MD. This however, would have increased the required computational time and resources by almost two orders of magnitude. The choice of 100 poses therefore reflects a compromise between the computational resources available for this project (provided by the Jülich Supercomputing Centre, Forschungszentrum Jülich) and the requirement to make the analyses as descriptive (and predictive) of experimental observations as possible.

The top one hundred favourable $\Delta\Delta G_{bind}$ values range from -170.75 kcal/mol to -120.05

$\Delta\Delta G_{bind}$ (kcal/mol)				
-170.75	-140.83	-132.94	-129.22	-123.01
-170.66	-140.23	-132.65	-128.84	-122.96
-170.17	-139.82	-132.53	-128.58	-122.75
-168.53	-139.63	-132.01	-128.07	-122.54
-161.31	-139.28	-131.72	-128.05	-122.42
-161.20	-138.83	-131.46	-127.05	-122.30
-160.58	-138.78	-131.34	-127.02	-122.22
-154.59	-138.34	-131.25	-126.72	-122.21
-153.43	-138.30	-131.14	-126.45	-122.20
-151.41	-137.87	-131.14	-126.02	-121.99
-146.58	-137.40	-131.04	-125.51	-121.86
-146.11	-136.71	-130.91	-125.29	-121.78
-145.95	-136.37	-130.89	-125.10	-121.48
-145.90	-135.70	-130.65	-124.74	-121.43
-144.68	-135.27	-130.52	-124.36	-120.93
-144.47	-135.00	-130.42	-124.21	-120.89
-144.43	-134.86	-130.04	-123.65	-120.86
-143.49	-133.98	-129.80	-123.45	-120.63
-142.14	-133.67	-129.52	-123.33	-120.08
-141.17	-133.47	-129.44	-123.03	-120.05

Table 5.1: Binding energies (kcal/mol) of the one hundred $A\beta$ monomer-D3 complexes with the lowest $\Delta\Delta G_{bind}$.

kcal/mol with an average of -133.87 kcal/mol. The full list has been presented in table 5.1. These were each subjected to 10 ns of MD simulation, making an aggregate of 1 μ s MD simulation. For each of the MD simulations, the minimum separation between $A\beta_{42}$ and D3 was calculated over the 10 ns MD trajectories and then averaged over the one hundred systems. This serves as a check to verify the overall stability of the predicted complexes, and if the switch from implicit solvent in the GMIN simulations to explicit solvent in the MD simulations has any significant effect on the stability of the systems. The obtained distance profile in figure 5.3 shows that the average separation distance stays between 2.8 Å and 3.0 Å throughout the trajectories. Thus the two peptides remain tightly bound. The slight fluctuation observed might have resulted from the system’s re-

sponse to the space-filling effect of the explicit water and the peptides' thermal vibrations within the allowed limits of the bound systems. In any case, the magnitude of the fluctuation is sufficiently small to render it incapable of disrupting the tight binding between the peptides.

Figure 5.4 shows the residue-residue contributions to ΔG_{int} generated from the se-

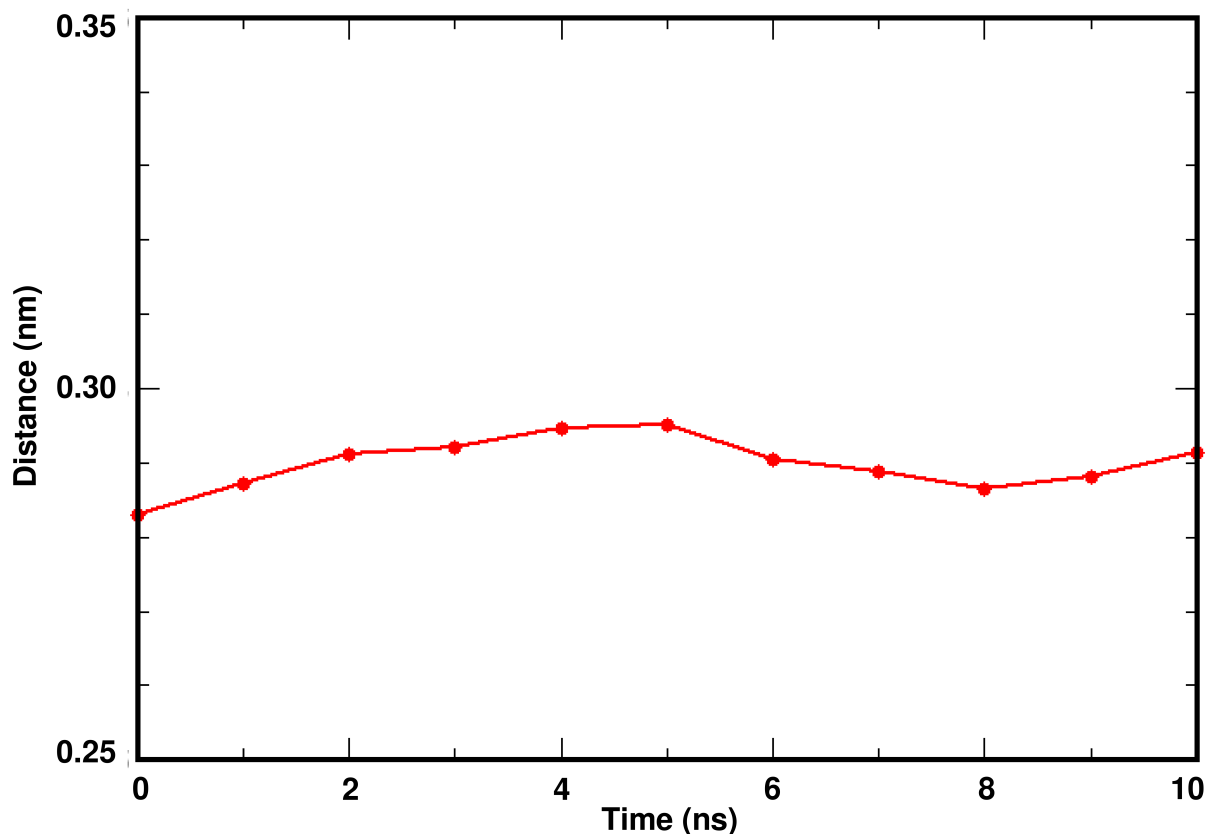


Figure 5.3: Average minimum separation between D3 and $A\beta_{42}$ as a function of time in the 100 $A\beta_{42}$ monomer-D3 complexes.

lected 100 binding configurations. Each of the selected $A\beta_{42}$ -D3 complexes had been subjected to MD simulation and from each resulting trajectory energy the residue($A\beta_{42}$)-residue(D3) interaction energy ΔG_{int} composed of the nonbonded energies (Lennard-Jones and Coulombic) was calculated. This resulted in 540 time-averaged energy values for each binding pose (i.e., 42×12). Subsequently an averaging over the one hundred binding poses was performed to produce the energy matrix shown in figure 5.4. In figure 5.4 the interaction energy map summarises the energetic contributions showing the relative significance

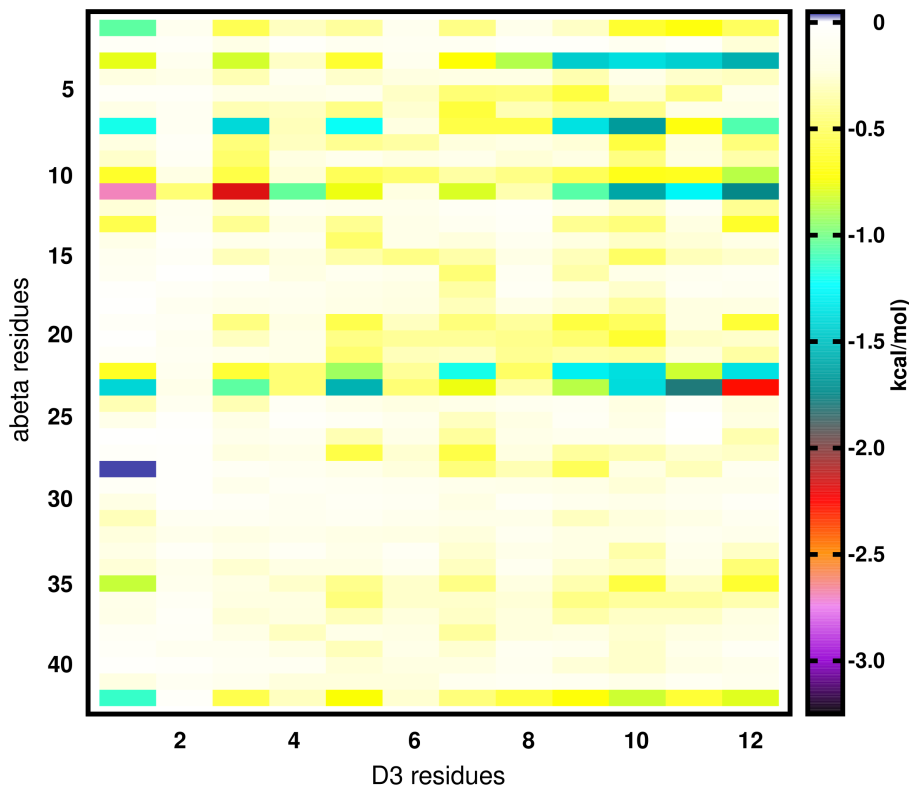


Figure 5.4: D3- $A\beta_{42}$ monomer explicit solvent molecular dynamics-generated interaction energies decomposed into residue-residue interactions. The interaction map represents a mean over one hundred such maps separately calculated for each unique D3- $A\beta_{42}$ GMIN-predicted binding poses with the best $\Delta\Delta G_{bind}$. The map reveals residues of the D-peptide preferentially interacting $A\beta$ hydrophilic residues in sequence regions 3 to 11 and 22 to 23 while the two hydrophobic stretches—the CHC and the C-terminal hydrophobic residues—are avoided.

of each residue pair interaction to the association event. The pattern shows D3 as near exclusively interacting with the N-terminal residues of $A\beta_{42}$, in particular Glu3, Asp7, Glu11, Glu22, Asp23. The fact that these five $A\beta_{42}$ residues are all negatively charged suggests that interaction with D3 is mediated by electrostatic attraction involving D3’s five arginine residues. The averaged interaction energies between D3 and $A\beta_{42}$ ’s negatively charged residues range between -0.78 to -1.13 kcal/mol, in sharply contrasts with the -0.09 to -0.20 kcal/mol for $A\beta$ residues in the CHC (residues 13 to 21) and the C-terminal hydrophobic (residues 25 to 42) sequences. The pattern reveals that D3 preferentially binds $A\beta_{42}$ monomer’s N-terminal segment. D3 on the other hand shows little residual selectivity for its interaction with the amyloid peptide. With averaged interaction energies ranging from -0.35 to -0.46 kcal/mol, the five arginine residues shows

a similar preference for interaction with $A\beta$ monomer which suggests that the reported amyloid inhibitory activity of D3 involves an electrostatic steering that directs D3 to the N-terminal half of $A\beta$. Interestingly the C-terminal end of D3, made of the sequence HRNR, demonstrates a mildly superior average ΔG_{int} in partial agreement with the work of Ghanta *et al.* [150,152]. Figure 5.4 shows that D3 binding to the negative residues in $A\beta$ brings D3 to $A\beta$'s self-recognition sequence, which is expected to significantly affect amyloid aggregation. This can also be seen in figure 5.5 presenting ten complexes

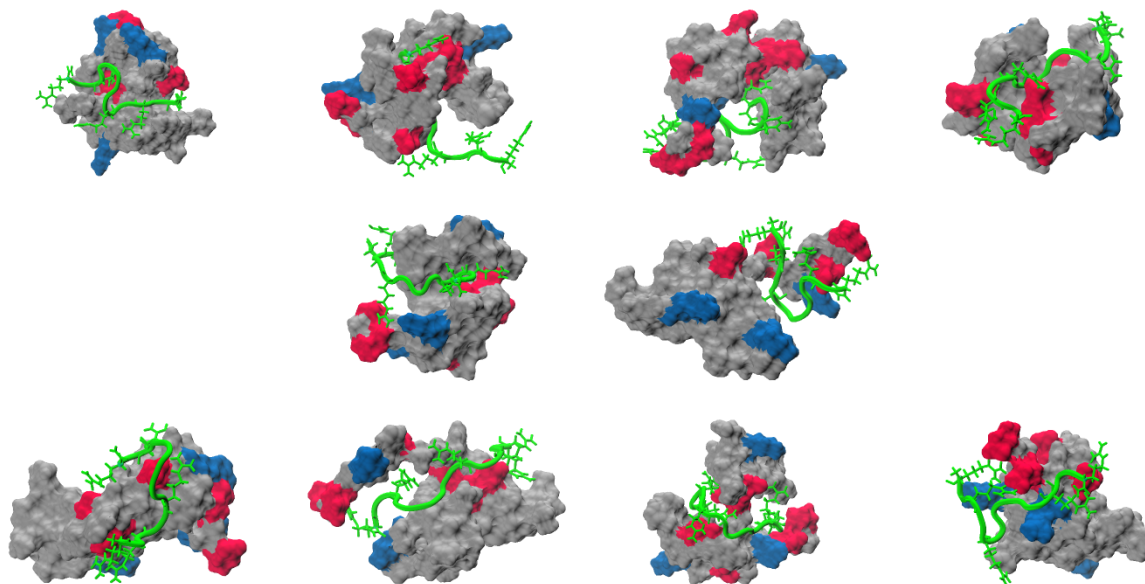


Figure 5.5: VMD rendering of some of the one hundred studied $A\beta_{42}$ monomer-D3 complexes. D3 is shown as green tube with the sidechain group of arginine residue shown in liquorice. $A\beta_{42}$ monomer is rendered using a surface representation coloured according to residue types—red, blue and silver respectively for acidic, basic and uncharged residues.

randomly selected from the 100 investigated systems. It therefore appears that amyloid inhibitory mechanism of D3 involves electrostatic-driven attraction to $A\beta$'s negatively charged residues, which then prevents aggregation into toxic oligomers via a shielding of $A\beta$'s aggregation nucleating sequence.

It then became crucial to determine the effect (if any) of D3 binding on the secondary

Structure	A β alone	A β with D3	% Δ
Coil	0.484	0.539	+11.36
Beta	0.162	0.125	-22.84
Bend	0.307	0.302	-1.63
Turn	0.037	0.025	-32.43

Table 5.2: The effect of D3 binding on A β_{42} monomer's secondary structure. Given are the populations of each secondary structure element in A β_{42} obtained from 100 10 ns MD simulations (averaged over time).

structure of A β_{42} monomer. In order to achieve this, the secondary structure content of A β_{42} monomer in the bound and unbound states was calculated using the secondary structure prediction programme DSSP [73] interfaced to GROMACS. The secondary structure results obtained as a function of time were then first averaged over time (i.e., 10 ns) and subsequently over the 100 simulated systems. Table 5.2 presents the values obtained for each secondary structural element. Comparing these values to those obtained from A β_{42} alone, it becomes obvious that the interaction of D3 peptide goes beyond a mere strong binding to A β_{42} monomer. It also has a noticeable effect on the structure of the amyloid peptide, which includes a significant reduction in β -sheet (-22.84%) and turn (-32.43%) content accompanied by an increase in coil structure ($+21.4\%$). D3 thus dissolves regular A β_{42} structure into unstructured states, in agreement with staining experiment that reported the absence of fibrillogenic property (i.e. β -sheet structure) in A β_{42} treated with D3 [163, 164]. This result partly explains the experimentally observed amyloid inhibitory activity of the D-peptide involving the conversion of toxic fibrillogenic A β structure into nontoxic nonfibrillogenic amorphous assemblies [162–164]. It thus further lends credence to the belief linking A β structure and toxicity—that is, destroying *toxic* amyloid structures provides a viable mechanism for Alzheimer's disease therapy.

5.3.2 A β_{42} pentamer/D3

Using basin-hopping, 4000 A β_{42} pentamer-D3 complexes were generated and optimised, and the binding energies $\Delta\Delta G_{bind}$ were calculated for the resulting complexes. The one

hundred lowest $\Delta\Delta G_{bind}$ values (table 5.3) range from -222.22 kcal/mol to -160.42 kcal/mol with an average of -175.30 kcal/mol. This represents a gain in $\Delta\Delta G_{bind}$ of -41.44 kcal/mol compared with the average value of -133.87 kcal/mol for $A\beta_{42}$ monomer-D3. This does suggest a stronger affinity of D3 for higher-order $A\beta$ oligomers (pentamer in this case) than for the monomer.

The higher affinity of D3 to pentameric $A\beta$ may partly be explained as resulting from

$\Delta\Delta G_{bind}$ (kcal/mol)				
-222.22	-183.54	-175.08	-169.42	-163.82
-205.93	-183.15	-174.78	-169.35	-163.78
-203.74	-182.69	-174.75	-169.34	-163.48
-203.37	-182.41	-174.55	-169.13	-163.23
-203.14	-181.89	-173.40	-168.79	-163.23
-201.75	-181.56	-173.33	-168.79	-163.09
-199.29	-181.47	-172.74	-168.63	-163.08
-198.42	-180.57	-172.71	-167.50	-163.08
-196.45	-179.84	-172.13	-166.58	-162.94
-194.43	-179.74	-172.00	-166.47	-162.76
-193.64	-179.15	-171.95	-166.26	-162.74
-192.98	-178.90	-171.65	-165.90	-162.70
-191.98	-178.13	-171.42	-165.36	-162.40
-188.70	-178.00	-171.27	-165.25	-162.39
-188.47	-177.41	-170.57	-165.22	-161.35
-188.26	-177.07	-170.51	-165.21	-161.27
-187.84	-176.56	-170.47	-164.94	-161.09
-184.51	-176.18	-170.30	-164.79	-160.86
-184.49	-175.96	-170.16	-163.90	-160.44
-184.46	-175.61	-169.96	-163.89	-160.42

Table 5.3: Binding energies (kcal/mol) of the one hundred $A\beta$ pentamer-D3 complexes with the lowest $\Delta\Delta G_{bind}$.

the bigger oligomeric size, which allows D3 to form more contacts with $A\beta_{42}$ pentamer. To provide a graphical representation of D3 binding, the D3 centre-of-mass was calculated in each of the 100 systems:

$$\mathbf{r}_{com} = \sum_i \frac{m_i \mathbf{r}_i}{N} \quad (5.7)$$

where m_i and \mathbf{r}_i are atomic masses and coordinates, respectively. This transforms the representation from atomic to molecular resolution, which becomes advantageous for representing several interacting systems at the same time. In figure 5.6 D3’s centres-of-mass in the one hundred systems are shown together with the $A\beta_{42}$ pentamer to highlight the interaction poses at a glance. The regular scaffold formed by the pentamer, unlike monomeric $A\beta_{42}$, enforces a more directed D3 binding. As would be expected, a large proportion of the interaction is confined to the disordered but negatively charged sections of the $A\beta_{42}$ peptide, reminiscent of D3- $A\beta_{42}$ monomer binding discussed above. A smaller

proportion involves D3 interacting, at the same time, with the disordered and the β -sheet ordered sections of the pentamer. Interestingly, in about 10 % of the observed complexes D3 is able to bind within A β 's U-shaped β -sheet segment and interact directly with the Asp²³-Lys²⁸ salt-bridge with the possibility to affect ordered fibril formation. The differ-

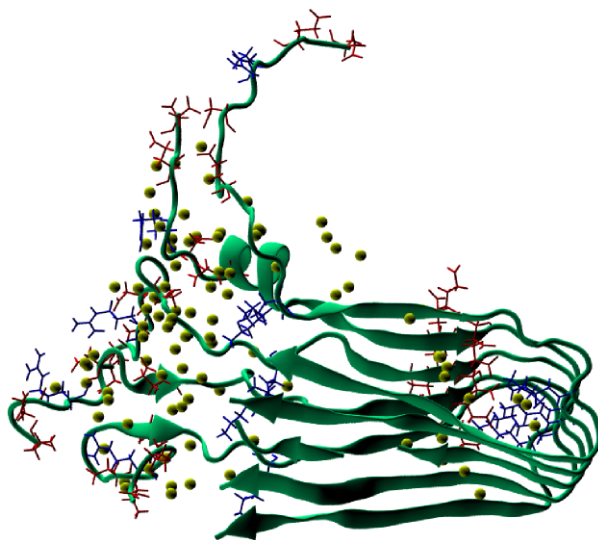


Figure 5.6: Overview of the interaction of D3 with A β_{42} pentamer. In yellow VDW radius the centres-of-mass of D3 in the one hundred lowest $\Delta\Delta G_{bind}^*$ free energy configurations are shown superposed on A β_{42} pentamer represented in green NewCartoon with acidic and basic residues shown in red and blue liquorice, respectively.

ent D3 binding regions must be independently considered as a variation in the binding efficacy (i.e., the effect on structure) between different binding modes as against what was obtained for A β monomer. Therefore, a cluster analysis was performed on the positions of D3 around A β_{42} pentamer with the aim to identify regular patterns in the binding modes (detail about the clustering in section 5.2.4). Fifteen clusters were obtained. After visual inspection, six clusters were selected (figure 5.7) and subjected to 50 ns explicit solvent MD simulation. Clusters 1, 4, 5, 6, 11 and 12 were selected based on the uniqueness of binding poses, and involve D3 binding to both or either of the two termini residues of the pentamer. Clusters 1, 5, 11, and 12 have the D3 residues at least partially piercing through two adjacent chains of the pentamer in the N-terminal segment (cluster 1) or the β -strand-forming segments (clusters 5, 11 and 12). And in clusters 4 and 6, D3 bridges the length of the pentamer and interacts with both termini.

As for the monomer, we analysed the effect of D3 binding on $A\beta_{42}$ secondary structure

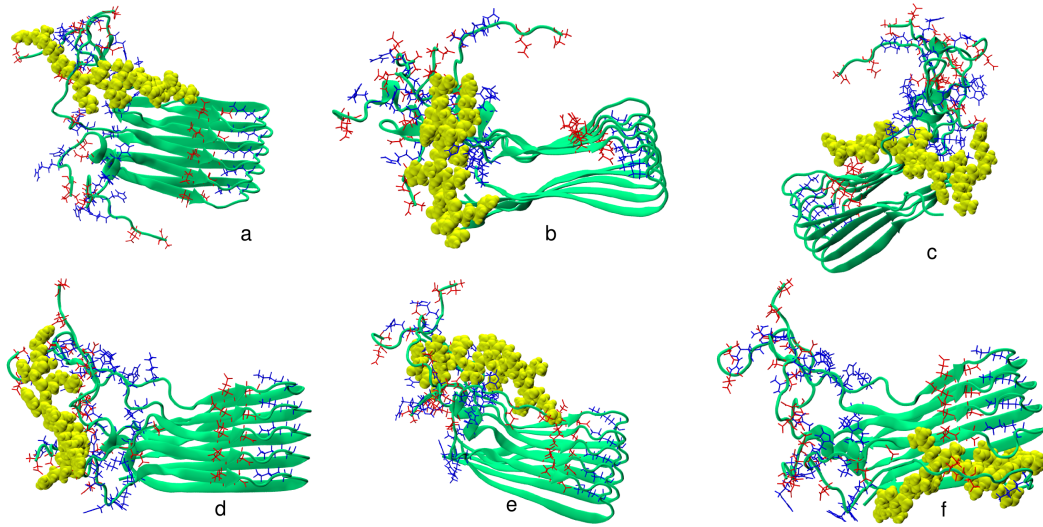


Figure 5.7: Six unique $A\beta_{42}$ pentamer-D3 binding modes obtained after clustering. The yellow VDW representation shows D3, bound to $A\beta_{42}$ pentamer represented in green NewCartoon with acidic and basic residues shown in red and blue liquorice, respectively. (a) cluster 1, (b) cluster 4, (c) cluster 5, (d) cluster 6, (e) cluster 11, and (f) cluster 12.

in the pentamer. The observed changes in $A\beta_{42}$ helical structures for the six binding modes are, for all purposes, insignificant. Although this is not surprising since helical structures represent less than 3 % of the sampled structure. The situation, however, becomes drastically different in the case of β -sheet and random coil structures, with all six binding poses being able to induce noticeable changes in the secondary structure of the pentamer. All binding modes increased the coil content while reducing β -sheet structure, but to greatly differing degrees. Cluster 6, where D3 binding exclusively to the disordered region of $A\beta$, records the least changes in both coil and β contents. On the other hand, binding poses with D3 simultaneously interacting with $A\beta$'s N-terminal and U-shaped regions significantly reduce the β -sheet content. The strongest effect is observed for cluster 12, in which D3 wedges between adjacent $A\beta_{42}$ chains in the U-shaped section of the pentamer. This leads to a disruption of the in-register backbone interaction (hydrogen bonds and sidechain packing) as well as the Asp²³-Lys²⁸ salt-bridge, which can account for the *deletion* of β -sheet. In clusters 1, 5 and 11 D3 concurrently interacts with $A\beta$'s disordered end and the Glu²² residues oriented above the fibril plain, while in cluster 4

the interaction involves the pentamer’s disordered region and Ile⁴¹ and Ala⁴² at the C-termini. This diversity in D3 interaction with A β suggests the existence of many fruitful binding modes in solution—fruitful because they all induce the desired reduction in A β ’s secondary structure. While it will be instructive to study the relative importance of the different binding modes, such detail is unlikely to increase the understanding of the association mechanism, especially, given the high degree of conformational plasticity and heterogeneity associated with the oligomeric peptide A β . In other words, several of these binding modes are likely to be concurrently significant, which will also partly explain the transformation of β -sheet-rich A β into an amorphous mass rather than a single structure.

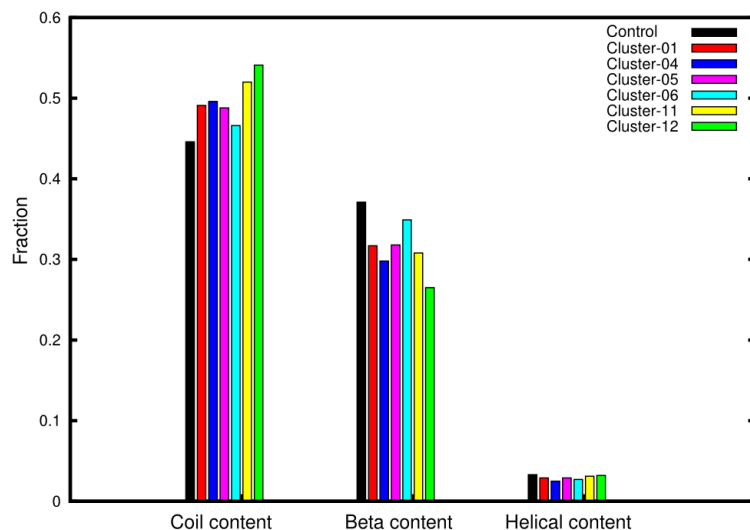


Figure 5.8: Secondary structure of A β_{42} in the pentamer with and without D3. The black histogram represents the reference simulation, i.e., A β_{42} pentamer in the absence of D3.

5.3.3 Conclusions

Our study of the interaction between D3 and both monomeric and pentameric A β_{42} reveals important structural effects in both A β species that can explain the experimentally observed inhibitory activities of D3, as well as structural aspects of amyloid toxicity inhibition. The performed simulations reveal that D3 strongly binds to A β_{42} monomer and pentamer, and in the process forms very stable complexes. Binding energy calculations

reveal a stronger interaction with the pentamer. Binding of D3, in both $A\beta$ species, results in the disruption of regular secondary structure, β -sheet in particular, in favour of random coil structure. This corroborates the experimental observation of $A\beta$ structural conversion into amorphous Tht-negative aggregates after treatment of a Tht-positive $A\beta$ solution with D3. Thus, the structural aspect of the inhibitory activity of D3 can be explained based on the outcome of the present study. Our results reveal an electrostatic-driven association between D3 and $A\beta_{42}$ which possibly brings D3 to the self-recognition unit of $A\beta_{42}$. Our *in silico* study captures well salient structural aspects of an otherwise complex *in vitro* situation, and even a more complex *in vivo* environment. In experiments, the simultaneous presence of several thousand molecules of D3 and an heterogenous mix of $A\beta$ assemblies ranging from monomers to protofibrils and fibrils is likely to influence the activity of D3. Under such conditions, D3 binding alone has the capability of significantly altering $A\beta$ size structure with each D3 molecule simultaneously binding different $A\beta$ units. The arginine-rich D3 may act as "glue" onto which $A\beta$ units stick. This alone immediately destroys the intricate amyloid aggregation process that is crucially dependent on a self-recognition-driven assembly, backbone hydrogen bond formation and sidechain packing.

So far we have learned how D3 can bind to preformed β -sheets and start to disrupt them. Future simulation will study the aggregation process of $A\beta_{42}$ with and without D3 to learn how D3 inhibits amyloid aggregation in the first place.

Chapter 6

D3 Modifications

6.1 Introduction

An overarching aim of the lead identification stage of pharmaceutical drug discovery is the identification of active compounds with the needed physicochemical profile such as molecular size, lipophilicity, number of rotatable bonds, hydrogen bond-donating and -acceptor groups. This allows the compound to possess the necessary pharmacokinetic attributes, such as bioavailability, without which candidate compounds cannot progress through subsequent stages of the discovery pipeline. In addition, lead optimisation is conducted (which is more of an iterative process) in which the desired activity of the compound of interest is improved upon (optimised) mainly via structural modifications relying on a knowledge of the structure-activity relationship (SAR) of the compound. Most peptide drugs have properties that frustrate their blood brain barrier (BBB) permeation, which constitutes a major reason why peptide-based systems sometimes represent a poor choice in the diagnostic and treatment of neurological pathologies. For instance, D3's molecular weight of 1.6 kD and other molecular indices violating acceptable thresholds for druggability, should either make it unfit for development, or then call for extensive chemical derivitisation to obtain agents with a better pharmacokinetic profile. For one, the preponderance

of arginine residues in D3 (isoelectric point of 13.1) would normally make unassisted BBB permeation difficult under physiological conditions. However, results from oral administration of D3 to transgenic mice indicates that D3 is able to sufficiently permeate the BBB and elicit its actions in the brain, suggesting the involvement of an assisted transport mechanism [164]. In 2010, using a model consisting of rat's brain microvascular endothelial cells and astrocytes, Liu *et al.* demonstrated [180] that D3's permeation of the BBB involves a partly saturable BBB transport via the adsorptive-mediated transcytosis mechanism [181] similar to that employed for ferrying strongly basic and arginine-rich peptides across membrane barriers [181,182]. The permeability, which was found to be higher than that obtained for another second 12-residue D-peptide (codenamed D1, sequence QSHYRHHISPAQV) with better lipophilicity than D3, showed a strong inhibition by cationic agents such as polylysine, further lending supporting to the presence of a facilitated BBB transport [180]. The result further highlights the importance of D3's five arginine residues in ensuring BBB permeation, as D1 with a single arginine residue showed reduced permeability. The five arginine residues of D3 thus confer upon it the ability to pass through the BBB to the brain where its actions are needed.

In the work reported in this chapter, we decided to determine if the desirable properties of D3 can be improved upon. We performed *in silico* modifications of D3 involving minor alterations of the D3 molecule. This way we did not risk losing the desirable features of D3 and at the same time the ability to evaluate the properties of modified molecules against those of D3.

6.2 D3 sequence reshuffling: RD2 and progenies

The first modification to the D3 structure involves a simple sequence reshuffling. The fact that D3 demonstrates no pattern involving a huge relative importance for any of the

non-arginine amino acid sequence suggests that the anti-amyloid activity is expected to be retained as long as the arginines are present. The sequence reshuffling involved pooling together all the arginine residues at the C-terminus while the remaining residues are in the N-terminus. This resulted in RD2, a D-peptide bearing 100% residue identity with D3 peptide but having a different D-amino acid sequence:

PTLHTHNRRRRR

The construction of RD2 aims at designing a template molecule having the five arginine residues as molecular *arrowhead* for targeted binding to A β , while small molecules possessing β -sheet-breaking activities may be incorporated into the N-terminal via substitution. In our simulation studies of RD2, the peptide was subjected to the same procedures (conformational search and interaction with both A β_{42} monomer and pentamer) as described for D3 in the previous chapter. The starting structure, generated with CHARMM with backbone ϕ and ψ angles set to 180° (figure 6.1(a)), was subjected to a 100 ns explicit solvent MD run after which a cluster analysis revealed the structure in figure 6.1(b) as the dominant conformation. In order to minimise repulsion, the sidechain groups of the five successive arginine residues are arranged in whorl-like manner roughly spanning 360° around the backbone.

Two additional D-peptide models were obtained from the RD2 sequence, HN5 rep-

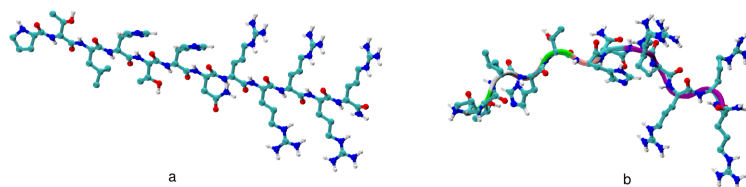


Figure 6.1: RD2 peptide model. (a) extended conformation before MD simulation. (b) dominant structure after 100 ns explicit solvent MD simulation with backbone coloured according to residue types with purple representing arginine residues.

resenting the last seven residues with sequence **HNRRRRRR**, and 5RS which is the

penta-D-arginine group. Each peptide was subjected to explicit solvent MD simulation as performed for RD2 and the dominant structures after cluster analysis of the resulting trajectories determined (figure 6.2). All three D-peptides had an amidated carboxy terminal. Adopting the study procedure described for D3, the interaction of each D-peptide with $A\beta_{42}$ monomer (employing five representative structures from the largest five clusters obtained from $A\beta_{42}$ conformational sampling in a 1.5 μs MD simulation) was investigated using a combination of global optimisation, binding free energy calculations, and explicit solvent MD simulations.

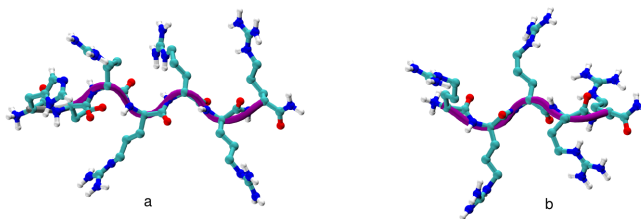


Figure 6.2: The two RD2 peptide progenies obtained after explicit solvent MD simulation: (a) HN5 and (b) 5RS. The backbone segment corresponding to arginine sequence is shown as purple tube.

6.2.1 Simulation protocol

With the aid of the oligomer-generation procedure [176] implemented into GMIN, 1000 $A\beta_{42}$ -D-peptide binary complexes were generated for each of the five $A\beta_{42}$ starting structures, which were then subjected to 300 BH steps to optimise them employing dihedral angle moves [177] and small rigid body rotations and translations [178] applied to the D-peptides. Dihedral perturbation was performed with a 0.3 probability with a maximal angle change of 30° , while rigid body translation and rotation were performed with a 0.8 probability and a maximum displacement of 2 \AA and 80° , respectively. 5000 lowest energy configurations were collected for each $A\beta_{42}$ -D-peptide system (1000 structures for each of the five $A\beta_{42}$ starting configurations) for which $\Delta\Delta G_{bind}$ was computed and employed

for ranking the binding poses. Each of them was subjected to a 10 ns explicit solvent MD simulation. The trajectories were analysed by averaging over time and the 100 systems, to obtain interaction energy maps between $A\beta_{42}$ residues and D-peptide residues. The simulations were then analysed to determine the effect of D-peptide binding on $A\beta_{42}$ monomer's secondary structure.

The interaction of RD2 with the full-length $A\beta_{42}$ pentamer was subsequently studied as for D3, i.e. 4000 $A\beta_{42}$ pentamer-RD2 complexes were generated and optimised in 400 basin-hopping steps. Here, a maximum angle of 90° was allowed for rigid body rotation with 1.0 probability and a maximum translation of 2 \AA was applied to RD2 at every BH step. The lowest energy $A\beta_{42}$ -RD2 complexes were collected and ranked using calculated $\Delta\Delta G_{bind}$ (as described in section 2.2.1). The distribution of RD2 around $A\beta_{42}$ pentamer was then determined by calculating the centre-of-mass of RD2 in each of the 100 binding poses having the most favourable $\Delta\Delta G_{bind}$ values. Based on the RD2 position in the complexes, a cluster analysis was performed with a 2.9 \AA RMSD cutoff yielding 15 clusters. Representative structures for six clusters with sufficiently different binding patterns were selected and each subjected to a 50 ns explicit solvent MD simulation using the same simulation parameters as for D3- $A\beta_{42}$ pentamer. The outcome were then compared to the 50 ns reference simulation of unbound $A\beta_{42}$ pentamer.

6.2.2 Results and discussion

The top 100 $A\beta_{42}$ monomer-RD2 binding poses have $\Delta\Delta G_{bind}$ values ranging from -177.4 to -114.3 kcal/mol, representing an improvement in binding for the complex of -7 kcal/mol compared with D3 (lowest $\Delta\Delta G_{bind}$ of -170.8 kcal). The situation between the two D-peptides however reverses when considering the average binding energy values ($\Delta\Delta\bar{G}_{bind}$), -128.0 kcal/mol and -133.9 kcal/mol for RD2 and D3, respectively. This

seems to indicate that both RD2 and D3 generally display a similar pattern of binding to monomeric $A\beta_{42}$. However, this is not surprising given that RD2 and D3 share 100 % residue similarity, the only difference being the sequence, which is not expected to significantly affect the electrostatic-driven association between the D-peptides and $A\beta_{42}$ monomer.

However, moving on to the shorter seven-residue HN5 and the five-residue 5RS peptides, a second factor influencing interaction with $A\beta_{42}$ monomer seems to emerge. Despite the fact that HN5 and 5RS also contain five arginine residues as RD2 and D3, they exhibit much higher $\Delta\Delta\bar{G}_{bind}$ values than the 12-residue D-peptides (table 6.1). Additionally, HN5 is better bound to $A\beta_{42}$ monomer than 5RS, which suggests the larger interaction surface in the longer D-peptides as well as additional contributions from non-arginine residues are important for binding of the D-peptides to $A\beta_{42}$. Normalising $\Delta\Delta\bar{G}_{bind}$ to the number of residues,

$$\Delta\Delta\bar{G}_{res} = \frac{\Delta\Delta\bar{G}_{bind}}{N_{res}} \quad (6.1)$$

where N_{res} is the number of residues in the D-peptide. We find that $\Delta\Delta\bar{G}_{res}$ values for HN5 (-14.8 kcal/mol) and 5RS (-18.6 kcal/mol) are more favourable than the values for RD2 (-10.7 kcal/mol) and D3 (-11.2 kcal/mol). However, this appears to result from the increasing relative number of D-arginine residues going from the 12-residue RD2 and D3 (41.7 %) to the penta-D-arginine 5RS (100 %). This further reveals the relative importance of arginine residues in driving the interaction. $\Delta\Delta\bar{G}_{res}$ for 5RS is markedly different from the other three systems, principally because the other systems contain non-arginine residues unlike 5RS. In order to properly describe the contributions of the different residues, especially in D-peptides containing non-Arginine residues, it is necessary to calculate $\Delta\Delta\bar{G}_{Arg}$ and $\Delta\Delta\bar{G}_{X,X\neq Arg}$, respectively representing the averaged contributions of arginine and non-arginine residues to $\Delta\Delta\bar{G}_{int}$ (table 6.1).

$A\beta$	System	% Arg	$\Delta\Delta G_{bind}$	$\Delta\Delta G_{res}$	$\Delta\Delta G_{Arg}^*$	$\Delta\Delta G_{X,X\neq Arg}^*$
$A\beta_{42}$ monomer	D3	41.7	-133.9	-11.2	-16.5	-11.1
	RD2	41.7	-128.0	-10.7	-18.8	-9.6
	HN5	71.4	-103.9	-14.8	-16.8	-11.6
	5RS	100.0	-93.1	-18.6	-18.1	-
$A\beta_{42}$ pentamer	D3	41.7	-175.3	-14.6	-	-
	RD2	41.7	-182.2	-15.2	-	-

Table 6.1: Binding free energies for D-peptide $A\beta_{42}$ complexes in kcal/mol. The $\Delta\Delta\bar{G}^*$ values were obtained using GROMACS and ffG53a6.

Following the explicit solvent MD simulations of the selected 100 binding poses for RD2- $A\beta_{42}$ monomer, HN5- $A\beta_{42}$ monomer and 5RS- $A\beta_{42}$ monomer, a decomposition of the time-averaged interaction energy into inter-residue contributions was performed and averaged over the 100 complexes. This generated the interaction energy maps presented in figure 6.3 (the map for D3 has been included for comparison). While a pattern of electrostatic-driven binding similar to that for D3- $A\beta_{42}$ monomer is also obtained for RD2, HN5, and 5RS, the three D-peptides display a generally higher preference for the acidic residues of $A\beta_{42}$. The number of residue-residue contacts with interaction energies lower than -2.00 kcal/mol is much higher for RD2 than for D3. As the relative D-arginine content increases from RD2 through 5RS so does the number of interactions with energies close to -3.00 kcal/mol. It is however interesting to note the increased preference for $A\beta_{42}$'s N-terminal residues 1 \rightarrow 11 with increasing relative D-arginine content. 5RS almost exclusively binds $A\beta_{42}$ monomer at residues 1 \rightarrow 11; completely avoiding Glu22 while the strength of interaction with Asp23 has also reduced to ≈ -1.40 kcal/mol. This is partly a result of electrostatic repulsion between the penta-D-arginine and Lys28 of $A\beta$ as the interaction map for 5RS shows. As the number of non-arginine residues increase in the RD2 family, so does the interaction strength with Glu22-Asp23 of $A\beta$ (≈ -2.20 kcal/mol and -3.00 kcal/mol respectively in HN5 and RD2). RD2 containing seven non-arginine residues also interacts mildly with the hydrophobic C-terminal half of $A\beta$. Compared to D3, RD2 more selectively binds Glu11 and Asp23 of $A\beta_{42}$ monomer. $\Delta\Delta\bar{G}_{Arg}$ and $\Delta\Delta\bar{G}_{X,X\neq Arg}$ values for D3 and RD2 in table 6.1 suggest a probable cause for this. With the five D-arginine residues pooled in the C-terminal half of the peptide, RD2 demon-

strates a more favourable $\Delta\Delta\bar{G}_{Arg}$ of -18.79 kcal/mol compared with -16.49 kcal/mol for D3. This partly explains while RD2 demonstrates higher selectivity for A β 's Glu11 and Asp23. There is however a trade-off with respect to a less binding favourability of the non-arginine residues in RD2 ($\Delta\Delta\bar{G}_{X,X\neq Arg} -9.60$ kcal/mol) compared with D3 ($\Delta\Delta\bar{G}_{X,X\neq Arg} -11.06$ kcal/mol). This can also be seen in figure 6.3(c) where D3 is seen to form more extensive interactions with non-charged residues of A β than RD2.

$\Delta\Delta\bar{G}_{bind}$ value obtained for the binding of RD2 to A β_{42} pentamer differs by -6.9

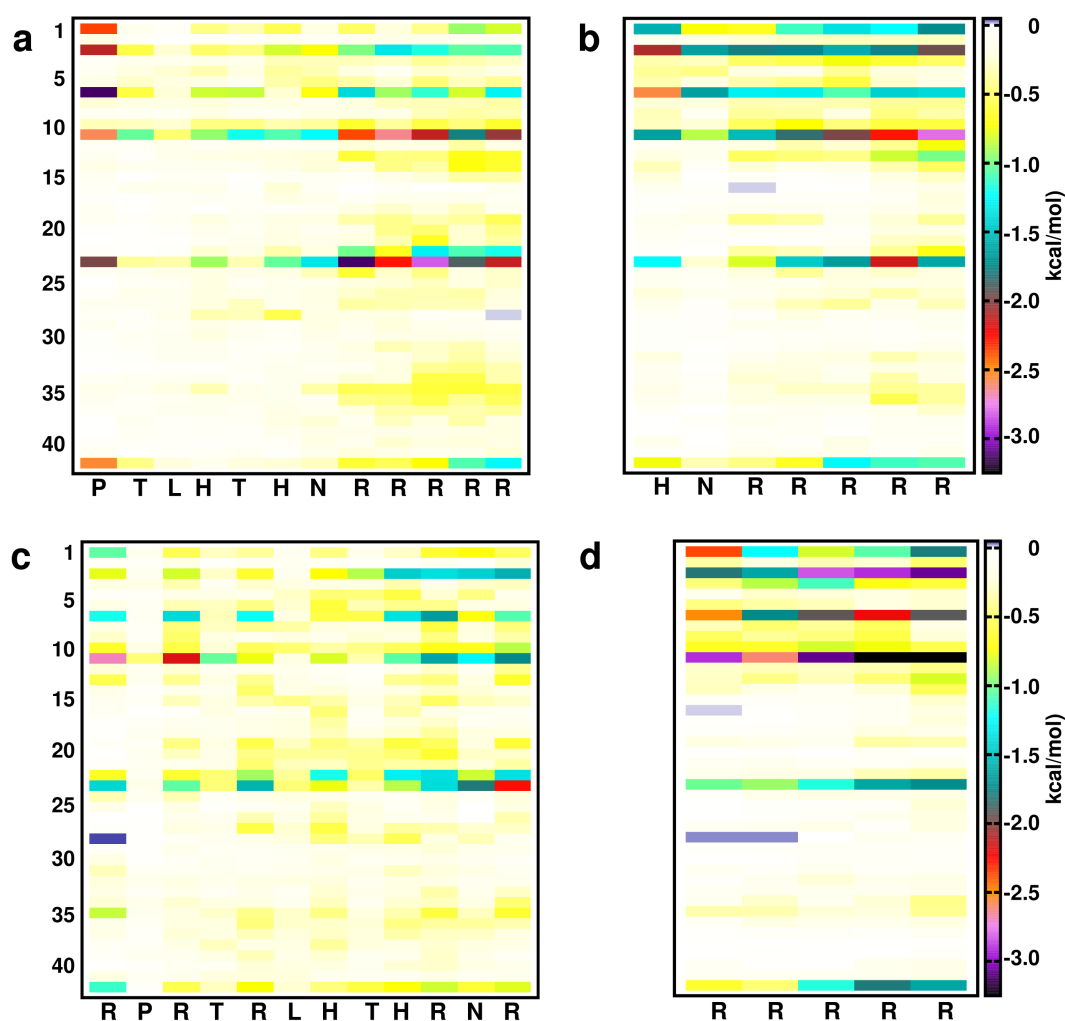


Figure 6.3: Interaction energy maps for A β_{42} monomer residues and (a) RD2 residues, (b) HN5 residues, (c) D3 residues, and (d) 5RS residues.

kcal/mol from that of D3 (table 6.1). RD2 thus exhibits slightly stronger binding than D3 to A β_{42} pentamer. In addition, the lowest $\Delta\Delta G_{bind}$ value of -272.52 kcal/mol obtained for RD2 is substantially lower than the corresponding value of -222.22 kcal/mol for

D3. The presence of a better structurally defined binding partner in $A\beta_{42}$ pentamer, compared with the largely *unfolded* $A\beta_{42}$ monomer, might hold a special significance for this relatively large disparity in the binding of RD2 and D3. The ordered stacks of five Asp23 residues within the fibril core of $A\beta_{42}$ pentamer and the five Glu22 residues immediately above it (figure 6.4) possibly provide a better binding surface for the penta-D-arginine groups of RD2 which are absent in D3. This is captured in figure 6.5 which presents the distribution of RD2 and D3's centers-of-mass around $A\beta_{42}$ pentamer.

RD2 binding modes are divided into two classes in which interaction is either restricted

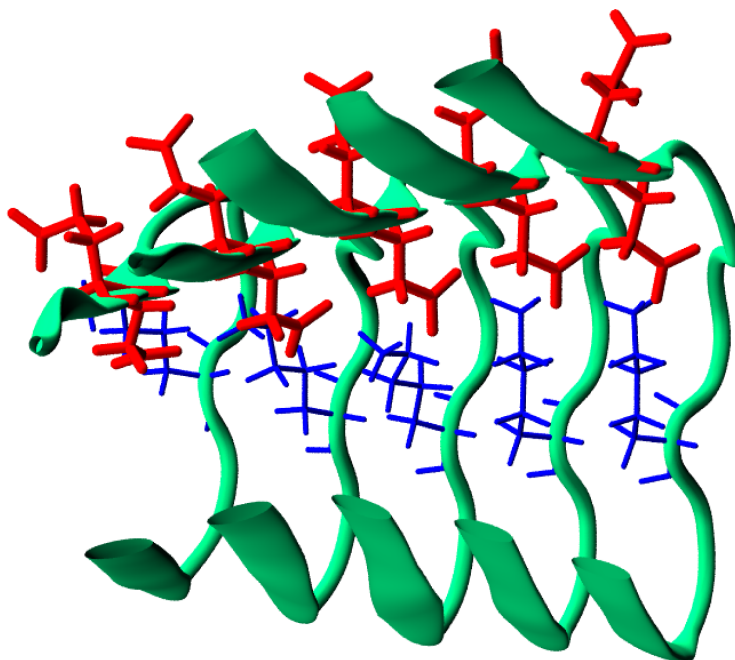


Figure 6.4: A view of the frontal section of $A\beta_{42}$ pentamer showing Phe20 \rightarrow Gly33 in green cartoon. The five Asp23 residues inside the fibril core and the Glu22 above appear in red liquorice, and Lys28 residues in blue liquorice.

to $A\beta_{42}$ pentamer's disordered segment, or to the hydrophobic U-shaped fibril core bearing Asp23 and Glu22 residues. This sharply contrasts with D3 which is distributed more widely around the $A\beta_{42}$ pentamer. Of the top 100 binding poses, RD2 centres-of-mass within the fibril core can be found 41 times, compared with 16 for D3. In other words, RD2 is ≈ 2.6 times better at targeting $A\beta_{42}$ pentamer's fibril core compared to D3. This thus suggests that the reshuffling of D3 residues into RD2 with the five D-arginine con-

centrated at the C-terminus significantly enhances its ability for focussed binding to the fibril core of the $A\beta_{42}$ pentamer.

This brings us to the important question: how does the observed binding pattern trans-

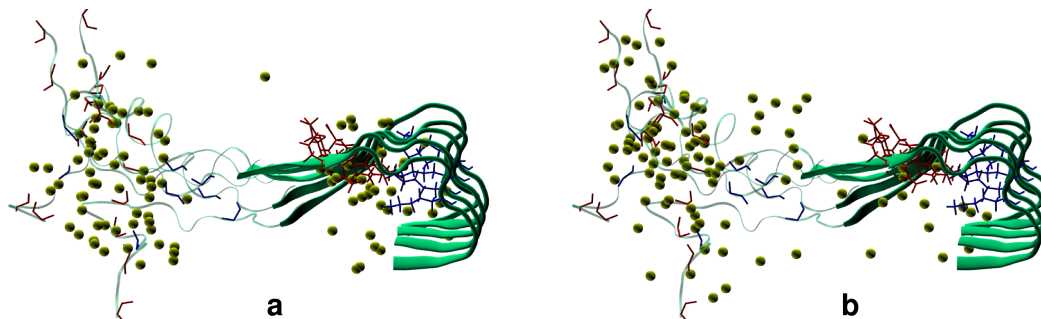


Figure 6.5: Binding of (a) RD2 and (b) D3 to $A\beta_{42}$ pentamer. The centre-of-mass of RD2 and D3 in each of the selected 100 binding poses is shown as yellow balls, residues 1–20 of $A\beta_{42}$ pentamer are represented with transparent ribbon, while residues Phe20–Gly33 have been shown in green NewCartoon. The fibril core's five Asp23, Glu22 (red) and Lys28 residues (blue) are represented in liquorice. The backbone groups of the acidic (red) and basic (blue) residues within $A\beta_{42}$'s disordered segment are also shown as liquorice.

late into structural efficacy in monomeric and pentameric $A\beta_{42}$ peptides? To this end, the trajectories for the 100 $A\beta_{42}$ monomer-D-peptide systems were analysed for secondary structure composition and the average over the 100 systems calculated. This was repeated for the unbound $A\beta_{42}$ from each of the predictions to provide a suitable control analysis. The changes in secondary structure between the bound and unbound $A\beta_{42}$ structures were then calculated and expressed as percentage change as shown in table 6.2. The percentage changes were calculated in such a way that a negative value represents a decrease in the respective secondary structure element while a positive value represents an increase. Out of the four D-peptides, D3 and RD2 binding induce the largest changes in $A\beta_{42}$ monomer structure consisting mostly in β structure destruction and increased coil content. However, the effect of D3 is conspicuously higher than that of RD2, which is likely to be a result of the higher level of D3 interaction with the hydrophobic residues of $A\beta_{42}$ monomer (figure 6.3). Figure 6.3 shows D3 as relatively more readily interacting with the two stretches of hydrophobic residues (17–21 and 30–42) with interaction energies in the order of -0.72 kcal/mol compared with > -0.24 kcal/mol for the RD2 family. Both HN5 and 5RS only

mildly affect the $A\beta$ structure. In conclusion, the RD2 family of D-peptides' increased targeting of $A\beta_{42}$ monomer's acidic residues fails to translate into better structural effect in $A\beta_{42}$ monomer.

RD2 binding positions on $A\beta_{42}$ pentamer were subjected to geometric clustering that

D-peptide	Beta (% Δ)	Coil (% Δ)	Bend (% Δ)	Turn (% Δ)
D3	-22.84	11.36	-1.63	-32.43
RD2	-15.09	5.53	1.95	-20.00
HN5	-3.66	0.60	1.68	-8.33
5RS	-4.07	1.00	2.03	-15.38

Table 6.2: Percentage change in $A\beta_{42}$ monomer's secondary structure following D-peptide binding

generated 15 clusters (in line with the number obtained for D3). Six of these clusters (figure 6.6) were each subjected to 50 ns MD simulation in explicit solvent. The trajectories were analysed and the sampled secondary structure contents were determined and compared to the reference (control MD simulation of unbound $A\beta_{42}$ pentamer). As can be observed in figure 6.6 RD2 binding, especially in clusters 2, 7 and 13, appears to induce much larger disruption of the $A\beta_{42}$ pentamer structure than did D3. Figure 6.7 shows secondary structure analysis and how the selected representative cluster centres compare with the D3- $A\beta_{42}$ pentamer complexes.

On a general note, RD2 more strongly inhibits β structure while at the same time increases coil content compared with D3. When RD2 binds along the fibril axis (i.e. perpendicular to the β -sheet) as in clusters 8 and 9, the effect on the $A\beta_{42}$ secondary structure is maximised. This represents a deviation from the pattern obtained for $A\beta_{42}$ monomer in which D3, as a result of the ability to form more widespread interaction with $A\beta$ generated higher structure-disrupting influence on the amyloid peptide. Experimental measurements on different D-peptides including D3 and RD2 reported RD2 as possessing higher amyloid-inhibiting effect than D3 peptide (Oleksandr Brener, personal communication). However, the relative amyloid-inhibitory effect of the two D-peptides under experimental conditions will very likely also depend on the relative abundance of different $A\beta$ species, with D3 likely to have a higher effect on monomeric and less structured $A\beta$

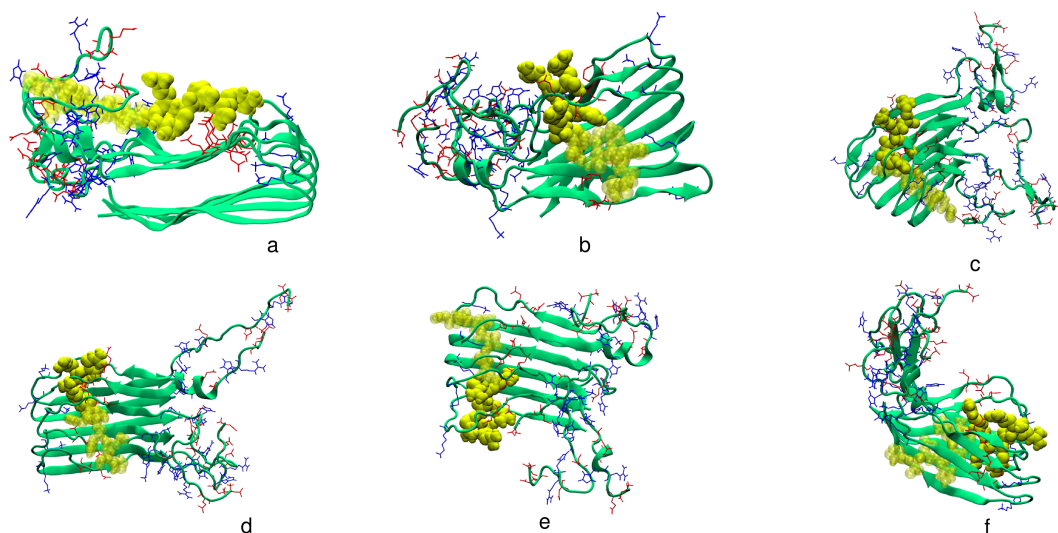


Figure 6.6: Six unique $A\beta_{42}$ pentamer-RD2 binding modes obtained from cluster analysis. The yellow VDW representation shows RD2 (the penta-D-arginine can be distinguished by its non-transparent representation), bound to $A\beta_{42}$ pentamer, which is shown in green NewCartoon with acidic and basic residues shown in red and blue liquorice, respectively. (a) cluster 1, (b) cluster 2, (c) cluster 7, (d) cluster 8, (e) cluster 9, and (f) cluster 13.

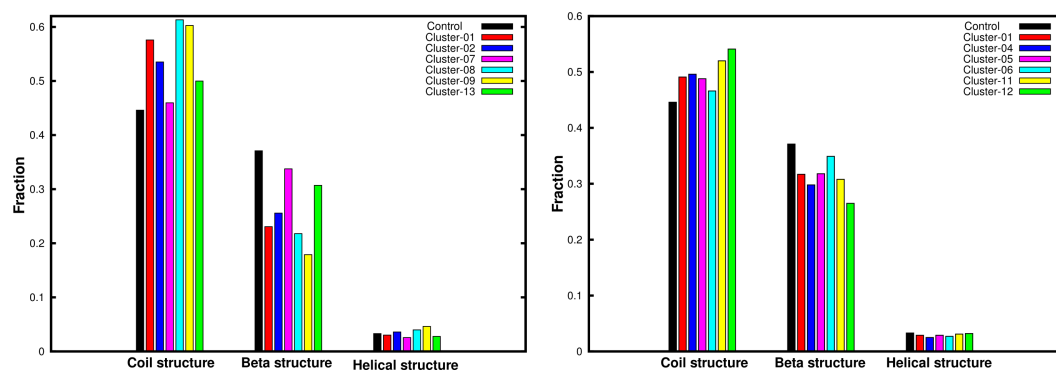


Figure 6.7: Effect of D-peptide binding on $A\beta_{42}$ pentamer secondary structure. On the left is the plot of changes resulting from RD2 binding, and on the right is a plot of changes from D3 binding. Generally, RD2 demonstrates much stronger effect $A\beta_{42}$ pentamer structure than D3.

species, while RD2 preferentially inhibits more structured $A\beta$ oligomers.

6.3 D3 mutations with natural and non-natural amino acids

6.3.1 Mutation studies with $A\beta_{42}$ monomer

The second category of modifications performed involve replacement of each amino acid sequence of D3 with the remaining nineteen possibilities. These mutations were performed for each of the 100 $A\beta_{42}$ monomer-D3 complexes with the lowest $\Delta\Delta G_{bind}$ values. That is, with each of the 100 complexes as starting structure, each of the D-amino acid residues of D3 was replaced with the other 19 D-amino acid residues to produce new complexes bearing the modified D-peptides. Glycine and proline residues were not substituted into the first sequence position; accounting for these two exceptions, the total number of amino acid substitutions is:

$$N_{mut} = (17 + 11 \times 19) \times 100 = 22,600 \quad (6.2)$$

We used the MMTSB tool set [183] to carry out the mutations using parameters from the CHARMM22 force field for topology definition [170]. Each of the resulting mutant system was first subjected to energy minimisation to eliminate atomic clashes resulting from the substitutions. Using the binding free energy calculation method described in the previous chapter for D3, $\Delta\Delta G_{bind}^{mut}$ was subsequently calculated for each of the 22,600 mutants. The deviation of $\Delta\Delta G_{bind}^{mut}$ from $\Delta\Delta G_{bind}^{orig}$ was then calculated:

$$\Delta\Delta G_{bind}^{diff} = \Delta\Delta G_{bind}^{mut} - \Delta\Delta G_{bind}^{orig} \quad (6.3)$$

where $\Delta\Delta G_{bind}^{orig}$ is the energy for the corresponding precursor D3- $A\beta_{42}$ monomer complex. From equation (6.3) it follows that any mutation improving the binding for a particular

amino acid position will have a negative $\Delta\Delta G_{bind}^{diff}$, and vice versa. Lastly, the average for each mutation was calculated over the 100 different binding poses and the standard deviations determined.

Out of 226 point mutations presented in figure 6.8, only ten show slight improvement

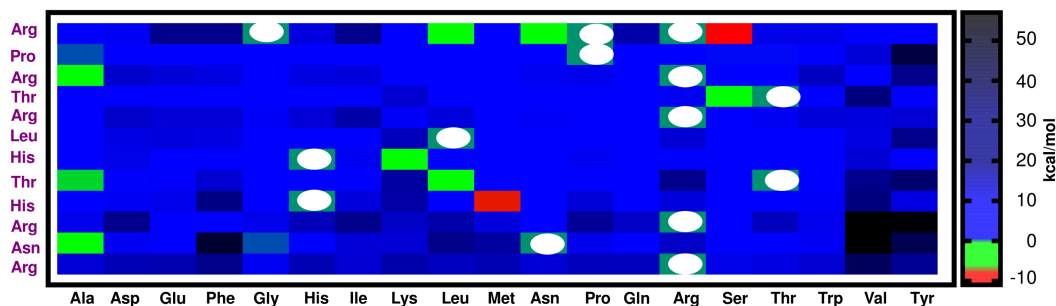


Figure 6.8: D3 substitution with natural amino acids for binding to $A\beta_{42}$ monomer. D3 residues appear on the vertical axis while the replacement D-amino acids are on the horizontal axis. The white circles represent non applicable substitutions involving self-substitutions (e.g., Arg replacing itself) and the omitted Pro and Gly at the N-terminus.

(i.e., $\Delta\Delta G_{bind}^{diff} < 0$) over the original D3. This suggests that D3 represents a highly optimised sequence in particular with respect to its $A\beta$ monomer binding activity. Four of the substitutions have $\Delta\Delta G_{bind}^{diff}$ values only slightly lower than zero and these are Thr8→Leu, Arg1→Asn, His7→Lys and Thr8→Ala. Replacing Arg1 with Ser brought about the highest observed improvement with a $\Delta\Delta G_{bind}^{diff}$ value of -10.5 ± 1.18 kcal/mol, followed by His9→Met and Arg3→Ala with -7.60 ± 1.09 kcal/mol and -5.31 ± 1.15 kcal/mol respectively. Changing Arg1→Leu and Thr4→Ser resulted in -4.50 ± 1.21 kcal/mol and -4.08 ± 0.92 kcal/mol, respectively. The binding enhancing mutations may be put together to produce the sequence **(A)**,

(A): SPASRLKLMRAR
(B): RPRTRLHTHRNR

where Pro2, Arg5, Leu6, Arg10 and Arg12 of D3 are all conserved. The sequence **(B)** is the D3 peptide. The introduced 'KLM' group vaguely resembles $A\beta_{42}$'s self-recognition unit that is critical for fibril assembly. Interestingly, the proposed sequence changes sup-

port a more hydrophobic N-terminal segment for the D-peptide which mildly echoes the sequence pattern in RD2. In this regard, the mutant peptide is likely to improve D3's affinity for Glu22-Asp23. This, however, may not directly translate into structural efficacy considering that D3 demonstrated better destruction of the β content in $A\beta_{42}$ monomer than RD2. However, considering the fact that the mutations increase the peptide's net hydrophobicity from 5 charged and 2 hydrophobic residues in D3 to 4 charged and 5 hydrophobic residues in the mutant, increased hydrophobic interactions ($\Delta\Delta\bar{G}_{X,X\neq Arg/Lys}$) in the new sequence with $A\beta_{42}$ monomer will likely improve binding efficacy.

Another interesting observation is that the three aromatic amino acids, Phe, Tyr, and Trp as well as the acidic amino acids, Asp and Glu, are strictly forbidden, i.e., they increased $\Delta\Delta G_{bind}$. Substitution with Asp or Glu is likely to interfere with the crucial role of D3's arginine residues via repulsive interaction with the acidic residues of $A\beta_{42}$. The avoidance of the aromatic residues points to a likely influence of chemical structure, in particular steric hindrance disallowing the replacement of D3's residues with large aromatic side-chains. In addition, the Thr \rightarrow Ser mutation suggests that the 4th D-peptide position is better occupied by a hydroxymethyl group of serine than threonine's 1-hydroxyethyl group. On the other hand, methionine's $-\text{CH}_2\text{SCH}_3$ side-chain is preferred to histidine's imidazole ring at position 9.

6.3.2 Mutation studies with $A\beta_{42}$ pentamer

The same kind of mutation study was repeated for $A\beta_{42}$ pentamer, but with three important modifications: 1) Some selected **non-natural** amino acids were included; 2) the mutations were performed on both D3 and RD2 sequences and ; 3) the programme for performing the point mutations was changed to Modeller [184] (used with the CHARMM22 force field [170]) that allows the definition of non-standard molecules. This represents an

initial attempt at diversifying the molecular structure of the two D-peptides and incorporating wider chemical classes. Starting with the top 100 D3- $A\beta_{42}$ pentamer and RD2- $A\beta_{42}$ pentamer structures used for the above described interaction study, each residue of D3 and RD2 was mutated to one of the other 19 natural and the seven non-natural amino acids. The structures of the non-natural residues, after geometry optimisation at the semi-empirical level of theory (PM6)[185], are shown in figure 6.9.

Binding free energy calculations, as described under section 6.3 were performed and

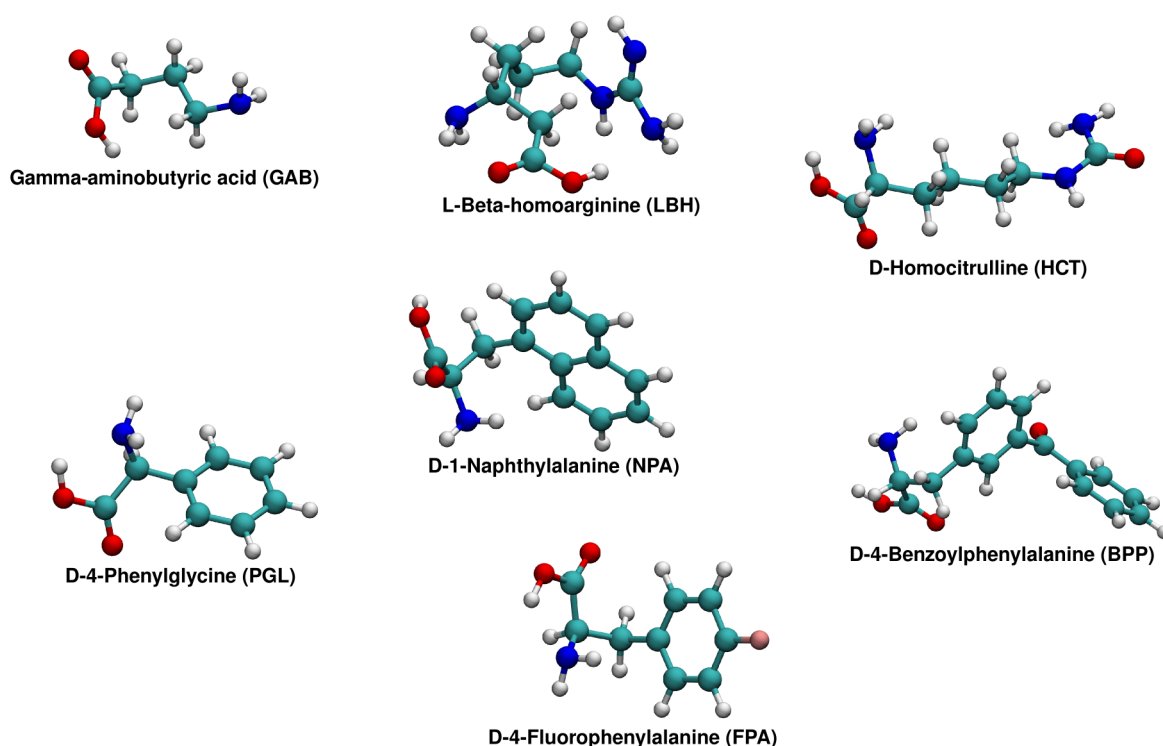


Figure 6.9: Structures of the non-natural amino acids employed for mutation. Structure generation and geometry optimisation were performed with Spartan'10 [12].

$\Delta\Delta G_{bind}^{diff}$ of resulting mutant D3 calculated for the $A\beta$ -D3 bound systems (total of $12 \times 24 \times 100 = 28,800$ mutations). In the case of RD2, the penta-D-arginine group shown to be crucial in RD2's interaction with $A\beta_{42}$ was kept, so that the substitutions were restricted to the seven non-arginine residues producing a total of $7 \times 24 \times 100 = 16,800$ mutations.

6.3.3 Results and discussion

The obtained $\Delta\Delta G_{bind}^{diff}$ values for the D-peptide– $A\beta_{42}$ pentamer differ from the values obtained for the $A\beta$ monomer mainly as a result of the differences in the extent of the D-peptide’s interaction with the $A\beta$ peptides. Interaction with $A\beta_{42}$ monomer involves less extensive peptide contacts than with $A\beta_{42}$ pentamer; this, perhaps, partially accounts for the fewer number of free-energy-improving substitutions obtained in the former. The higher extent of contact in the D-peptide– $A\beta$ pentamer, however, comes with certain implementation challenges in which some of the systems returned very high $\Delta\Delta G_{bind}$ values resulting from serious clashes involving residues with large sidechain groups. Examples include BPP with a benzoylphenyl sidechain, and NPA with a naphthalyl sidechain, which as seen in figures 6.10 and 6.11, are too large to be accommodated within the D-peptide– $A\beta_{42}$ pentamer tightly bound complexes. Nevertheless, their energies were not excluded since they provide useful information about unfavourable energetics of the concerned residues.

In other words, BPP and NPA are disfavoured in both D3 and RD2 peptides mostly

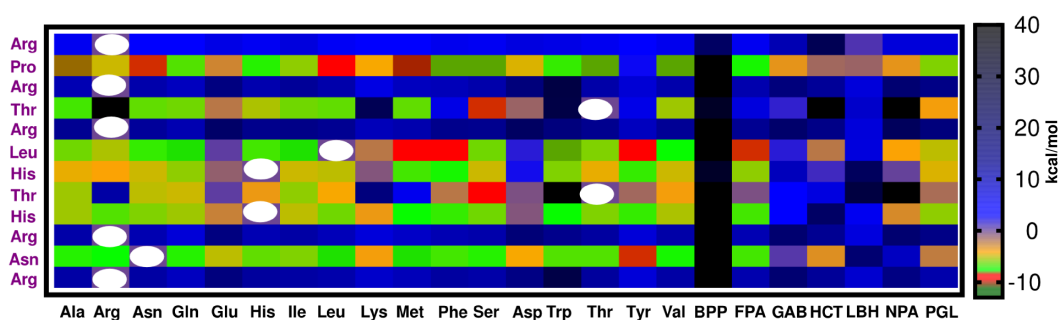


Figure 6.10: D3 optimisation by amino acids substitutions for binding to $A\beta_{42}$ pentamer. D3 residues are listed on the vertical axis while the substituting D-amino acids are on the horizontal axis. The white circles represent zero $\Delta\Delta G_{bind}^{diff}$ for self-substitution, e.g., replacing Arg residue with itself.

because of unfavourable size. In the case of D3, however, the picture changes significantly going from BPP and NPA, with two-ring aromatic sidechains, to one-ring aromatic sidechain where changing Pro2→Phe, Leu6→Trp, Leu6→Phe, Leu6→Tyr, Leu6→FPA, and Asn11→Tyr result in some of the most favoured substitutions. While the replace-

ment of Pro2 and Leu6 with hydrophobic and aromatic residues represents a significant majority of the favourable substitutions, D3's arginine residues are all strictly conserved. This tends to suggest that the arginine residues play a more crucial role in binding $A\beta_{42}$ pentamer than they do in $A\beta_{42}$ monomer. Also, the fact that aromatic substitutions are disallowed in D3- $A\beta_{42}$ monomer complexes while they are allowed in D3- $A\beta_{42}$ pentamer complexes suggests that aromatic substitutions may be important for designing D-peptide inhibitors with increased specificity for $A\beta_{42}$ oligomers.

The pattern observed in RD2 substitutions is similar to that of D3, especially in the

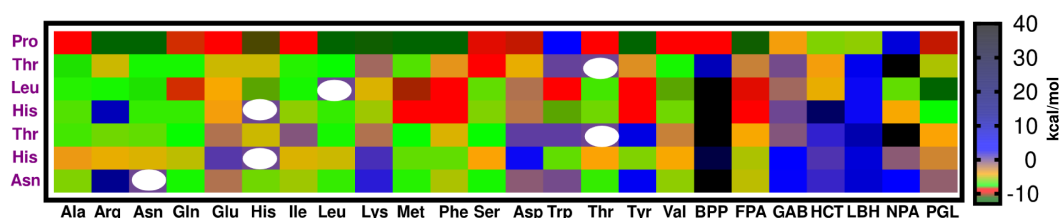


Figure 6.11: RD2 optimisation by amino acids substitutions for binding to $A\beta_{42}$ pentamer. D3 residues are listed on the vertical axis while the substituting D-amino acids are on the horizontal axis. The white circles represent zero $\Delta\Delta G_{bind}^{diff}$ for self-substitution, e.g., replacing Arg residue with itself.

preference for hydrophobic and aromatic substitutions. However, many more favourable amino acid substitutions are possible in RD2 than in D3 peptide. For example, RD2's proline in position 1 accommodates most of the substitutions with the exception of Trp, GAB, HCT, LBH, and NPA. The higher chances of substitution observed for RD2 compared with D3 may have resulted from the specific nature of RD2's binding to $A\beta_{42}$ pentamer. By mostly utilising its penta-D-lysine group for interacting with $A\beta_{42}$ pentamer, RD2 makes its seven non-arginine residues relatively more available for substitution than D3 where the staggering of the five D-arginine residues along the entire sequence makes it difficult to introduce substitutions that do not affect interaction. From the non-natural residues, only FPA replacement of Leu6 produces favourable $\Delta\Delta G_{bind}^{diff}$ for D3- $A\beta_{42}$ pentamer, while for RD2 FPA replacement of Pro1, Leu3, His4, as well as PGL replacement of Pro1 and Leu3 improve $A\beta_{42}$ pentamer binding. That is, RD2 provides a better molecular template for chemical modifications than D3.

A closer look at the substitutions involving single-ring aryl sidechains such as Phe, Tyr, FPA (4-fluorophenylalanine), and PGL (phenylglycine) in figure 6.11, seems to suggest a distinct pattern. Replacing Thr5 and Asn7 with Tyr is unfavourable, suggesting a disallowing of the aromatic OH group at this position. Interestingly however, replacing Tyr's hydroxyl group with a flouride (FPA) at this position produces a favourable binding energy. Phe, Tyr, and FPA produce an identical effect when replacing Leu3 and Thr5 ($\Delta\Delta G_{bind}^{diff} \approx -10$ kcal/mol) and Pro1 ($\Delta\Delta G_{bind}^{diff} \approx -12$ kcal/mol), suggesting the effect here mostly results from the phenyl ring. However, with the sidechain shortened by a methylene group, as seen in PGL, the number of favourable substitutions reduced from 3, in Phe, Tyr and FPA, to 2 in PGL where substitution is only allowed for Leu3 ($\Delta\Delta G_{bind}^{diff} \approx -12$ kcal/mol) and Pro1 ($\Delta\Delta G_{bind}^{diff} \approx -10$ kcal/mol). Only Leu3→PGL can be considered seriously given the fact that Pro1 is generally open to all substitutions including BPP. Apart from PGL, valine and histidine also successfully replace Leu3 which seems to suggest this position in RD2 prefers sidechain groups that are slightly smaller than a benzyl group. Lastly, the incorporation of FPA provides a means of improving the lipophilicity of the molecule, should that be desired, since flourine in the para position has the ability of improving both the lipophilicity and stability of the aromatic ring [186].

6.3.4 Conclusion

Using sequence shuffling and about 700 hundred different point mutations, we examined the effect of amino acid sequence modifications on the binding free energies of D3 interacting with $A\beta_{42}$ monomer and pentamer. Pooling the five D-arginine residues present in D3 to the C-terminus produces RD2, a reshuffled form of D3 with a penta-D-arginine sequence expected to facilitate targeted binding of the D-peptide to $A\beta_{42}$'s acidic residues. We performed an assessment of the binding free energy as well as the effect of binding

on $A\beta_{42}$ monomer and pentamer structures. A comparison of RD2 with D3 resulted in an identical pattern of binding to $A\beta_{42}$ monomer dominated by electrostatic attraction between the D-arginine residues of D3 and RD2 and $A\beta$'s acidic amino acid residues. However, while D3 prefers to interact with $A\beta_{42}$'s Glu22 and Asp23, RD2 demonstrates a slightly higher preference for Glu11 and Asp23. A deconvolution of the binding energy shows that the five D-arginine residues are more efficiently employed in RD2 than in D3. However, while interacting with $A\beta_{42}$, D3 additionally utilises several energetically weak hydrophobic contacts. While both D-peptides demonstrate β -sheet breaking effect on $A\beta_{42}$ monomer, the fact that D3 more significantly affects $A\beta$'s structure demonstrates the importance of hydrophobic contacts to overall binding.

Sequential removal of RD2's non-arginine residues almost completely abolishes the β -sheet breaking effect, further supporting the contribution of hydrophobic interactions. Visual examination of D3- $A\beta_{42}$ monomer complexes and interaction energies suggests the employment of binding poses which allows D3 to interact with $A\beta_{42}$'s hydrophobic sequences that are often involved in β -sheet formation. This explains the higher structural effect obtained for D3 compared with RD2 in the $A\beta_{42}$ monomer complexes.

The pattern, however, reverses in $A\beta_{42}$ pentamer systems. Interestingly, the penta-D-arginine group of RD2 renders it better able to interact with Asp23 existing deep within $A\beta_{42}$ pentamer's hydrophobic core. In this regards, compared with D3, we have shown that RD2 is approximately thrice as likely to penetrate into $A\beta_{42}$ pentamer's fibril core. Drawing inference from the energetic pattern seen in $A\beta_{42}$ monomer, RD2's binding may be a consequence of the increased binding effectiveness observed for the D-arginine residues in RD2. In addition, it explains the higher destruction of $A\beta_{42}$ pentamer's β -sheet obtained for RD2 than for D3. We further observed that the structural effect of D-peptide binding in both RD2 and D3 correlates well with the binding site within $A\beta_{42}$ pentamer. The two peptides produce the largest effect on $A\beta$ structure when bound to the β -sheet-

forming regions than when bound to the disordered N-termini. However, it may be possible that the *in vitro* and *in vivo* anti-amyloid effects reported for the two D-peptides are mediated via both binding modes.

Results from the amino acid substitutions reveal D3 as already highly optimal for binding $A\beta_{42}$ monomer. Replacing Arg1 and His9 with Ser and Met, respectively, produced the highest improvement in binding free energy. We also observed that substitution of D3's Pro1, Arg5, Leu6, Arg10 and Arg12 is strongly disallowed, and that aromatic residues are generally not favoured over existing D3 residues. This pattern, however, changes slightly with D3- $A\beta_{42}$ pentamer and RD2- $A\beta_{42}$ pentamer systems where more substitutions are allowed than for the $A\beta_{42}$ monomer. This further supports the position that D3 sequence is highly optimised for monomeric $A\beta$. Using substitutions with natural and seven non-natural amino acids, we observe that certain amino acid positions in the D3 sequence demonstrate selectivity for a particular chemical class. The fact that single-ring aromatic substitutions are favoured over other groups suggests the possibility of using these groups in designing inhibitors with improved selectivity for $A\beta_{42}$ oligomers.

6.4 Supplementary information: Force field parameterisation

In the quantum chemical description of molecular systems, molecules are modelled as collections of interacting electrons and nuclei. Various levels of approximations are used in order to solve the molecular Schrödinger equation, which expresses the total energy of an electronic system as an eigen-equation:

$$\hat{H}\Psi = E\Psi \quad (6.4)$$

Here, Ψ is a wavefunction representing the electronic state, and when operated upon by the Hamilton operator, \hat{H} , the system energy is obtained. For the simplest case, which is the hydrogen atom—the only system for which the equation has an exact solution—the equation expands (in atomic units) into,

$$\left(-\frac{1}{2}\nabla^2 - \frac{Z}{r}\right)\Psi(\mathbf{r}) = E\Psi(\mathbf{r}) \quad (6.5)$$

which captures both the electronic kinetic (first term in the equation) and potential energy (second term in the equation). The distance r is the separation between the electron and the nucleus having a nuclear charge Z ($Z = 1$ for H) and \mathbf{r} is the 3D coordinate of the electron. The Schrödinger equation for polyelectronic systems includes terms for electron-electron, electron-nucleus, and nucleus-nucleus interactions:

$$\hat{H}\Psi(\mathbf{r},\mathbf{R}) = E\Psi(\mathbf{r},\mathbf{R}) \quad (6.6)$$

where the Hamiltonian operator is given by:

$$\begin{aligned} \hat{H} = & -\frac{1}{2} \sum_i^{\text{electrons}} \nabla_i^2 + \sum_i^{\text{electrons}} \sum_{j>i}^{\text{electrons}} \frac{1}{r_{ij}} - \frac{1}{2} \sum_A^{\text{nuclei}} \frac{1}{M_A} \nabla_A^2 \\ & + \sum_A^{\text{nuclei}} \sum_{B>A}^{\text{nuclei}} \frac{Z_A Z_B}{R_{AB}} - \sum_i^{\text{electrons}} \sum_A^{\text{nuclei}} \frac{Z_A}{r_{iA}} \end{aligned} \quad (6.7)$$

where r_{ij} is the distance between electrons i and j , R_{AB} is the distance between nuclei A and B , and r_{iA} is the separation between electron i and nucleus A . This equation cannot be solved analytically. Approximations introduced to make it numerically tractable include the Born-Oppenheimer approximation that neglects the nuclear motion removing the third term in equation 6.7 and turning the 4th term into a constant. The Hartree-Fock (HF) model derives from the similarly named approximation in which the wave function of the system can be approximated by a single Slater determinant, which considers Ψ for a collection of electrons, each with a wave function known as spin orbital. By invoking the variational method, one can derive a set of N_{spin} -coupled equations for the N_{spin} spins. The Hartree-Fock system gave birth to semi-empirical models which neglect non-valence electrons and use much fewer basis sets than the HF models. Finally but importantly, semi-empirical models also employ approximations that allow the incorporation of empirical terms, which coupled with its computational efficiency and the availability of basis sets for calculations relating to transition metals make semi-empirical method a good choice for systems containing up to few hundreds of atoms. In this work, the semi-empirical method PM6 [185] as implemented in the quantum chemistry package Spartan'10 [12] was employed to derive force field parameters for 7 non-natural amino acids.

Following energy minimisation, the equilibrium geometries for the seven non-natural amino acids were used to determine the molecular mechanics parameters compliant with the

CHARMM27 potentials. With the newly generated CHARMM parameters a 10 ns MD

simulation in explicit solvent using the MD programme NAMD [187] was performed for each molecule to confirm that the parameters did not sample *unreasonable* conformations. The parameters were thereafter employed for conducting the amino acid substitution in the bound D3- $A\beta_{42}$ pentamer and RD2- $A\beta_{42}$ pentamer systems.

Chapter 7

Summary and Future perspectives

The past decades have witnessed a great deal of research activities into the role of $A\beta$ peptide in the development of Alzheimer's disease. While increasing number of biochemical pathways have been discovered to be involved in AD development, research findings have also increasingly supported the involvement of the $A\beta$ peptide in many of these pathways. For instance, there is a good documentation of the neuronal toxicities of different $A\beta$ alloforms and oligomeric sizes, which understanding has also motivated targeting of $A\beta$ associated pathways in Alzheimer's drug design. However, the very fact that no definitive treatment currently exists for Alzheimer's disease highlights the associated level of complexity. Unlike other drug targets, the small oligomers of the $A\beta$ peptide, which are rightly believed to be the main toxic species, lack a unique three-dimensional structure. In addition, their significantly high aggregation rates make it difficult to employ powerful tools like solution NMR, for $A\beta$ structure determination. This has prompted the employment of artificial solubilising conditions such as the use of organic solvents in structure determination. The problem with this approach is that the obtained structures, which are often helical, do not exactly represent the physiological state under which conditions $A\beta$ is believed to mostly sample β -sheet and random coil structures.

Using the ability of molecular simulations to focus attention on a single molecule, we

studied the conformation dynamics of A β peptides. To limit force field bias, we employed long MD trajectories and validation with NMR chemical shifts to select the force field that best models A β_{40} and A β_{42} monomers' physiologic conformations. Only the force field ffG53a6 sufficiently reproduced the NMR chemical shifts for both A β_{40} and A β_{42} monomer, and captured the fast conformational transitions associated with an intrinsically disordered peptide (IDP) like A β . The force field shows both peptide as being essentially unstructured in the N-terminal segment, but with A β_{40} slightly more disordered than A β_{42} . In addition, it shows the self-recognition unit sequence ¹⁶KLVFF²⁰ as forming β -sheet in both peptides but to a slightly higher extent in A β_{42} , in which a Gly37-Gly38 turn thought to distinguish A β_{40} from A β_{42} was also correctly sampled. Cluster as well as secondary structure analyses revealed A β_{42} as more rigid and a better β -sheet former than A β_{40} . Importantly, our results agree well with experimental findings, suggesting that the behaviour of the two peptides in a laboratory set up involving thousands of molecules, can be explained at a single molecule level using molecular simulation. This is also true with respect to the conformations sampled by A β_{42} under the different pH conditions we modelled based on amino acid charged states. Previous experiments had indicated that the structure and aggregation of A β are largely affected by pH. In our *in silico* experiment, we observed the A β_{42} conformation to depend on the protein net charge. The observed structural effect is most dramatic at the isoelectric pH 5.4 where the highest percentage of β -sheet is formed resulting from a minimisation of electrostatic repulsion. His6 was observed to play a critical role in ordering intramolecular contacts and facilitating secondary structure formation. We believe this explains, at a single molecule level, the high aggregation propensity associated with the peptide at slightly acidic pH. The weakness of our approach is however also associated with the ability to single out individual molecules for investigation. Atomistic simulations are still unable to handle system sizes comparable to what can be achieved in experiments. In addition, there is

the risk of mismatching protein system and force field capabilities mainly because there is still no universal protein force field. Instead the design of the currently available force fields restricts their very use to systems not too much different from those for which they were designed. For instance, the factors driving the conformational properties of IDPs are sufficiently different from most globular proteins, which probably explains why most existing force fields have not been as successful with IDPs as they have been with natively folded proteins. Most of the existing protein force fields have been observed to correctly predict protein folding rates as well as native state conformations, while the performance with respect to the features of the unfolded state and the folding mechanism depends heavily on the choice of force field. It therefore becomes evident, that the inability of the force field ffG43a2 to sample the correct structural ensemble for $A\beta$ emanates from a restriction of its applicability to folded protein systems for which it has been largely successful. On the other hand, the success of the ffG53a6 force field for modelling $A\beta$ is likely because it was parametrised to reproduce both the free enthalpies of hydration and apolar solvation which are important factors in most biomolecular processes. In other words, $A\beta$ conformational states obtained for ffG53a6 result from the choice of the parameter set and not vice versa. Our results further highlights the need for a proper benchmarking of the protein force fields, and in particular for the need to broaden the selection of protein test sets beyond natively folded proteins to include IDPs. In addition, the benchmarking may include validation with experimental observables such as chemical shifts.

Using a combination of global optimisation, binding free energy calculation and MD simulation methods, we investigated the interaction of amyloid aggregation-inhibiting D-peptide inhibitors on $A\beta_{42}$ monomer and pentamer structures. The D-peptides, especially the 12-residue D3 peptide, interact via electrostatic interactions with $A\beta_{42}$'s Asp23, Glu22, and Asp11 residues, to produce significant reduction in $A\beta$'s β -sheet content. While RD2, a reshuffled form of D3 with a C-terminal penta-D-arginine sequence, presents similar

binding pattern as D3 to $A\beta_{42}$ monomer, it is slightly not as efficient at destroying $A\beta_{42}$ monomer's secondary structure; an effect which results from D3's better ability to additionally interact with $A\beta_{42}$ monomer's hydrophobic residues. Truncated forms of RD2 including a penta-D-arginine peptide, bind strongly to $A\beta_{42}$ monomer but failed to reproduce the amyloid-disrupting effect of RD2 and D3. Interestingly, with the $A\beta_{42}$ pentamer, RD2 produces a much greater β -sheet breaking effect than D3. This was observed to result from RD2's greater ability to interact within $A\beta_{42}$ pentamer's fibril core. We suggest this to be attributable to a difference in the $A\beta$ oligomer specificity of the two D-peptides and the possibility of deriving, via co-administration, a synergistic advantage from D3's effect on $A\beta_{42}$ monomer and RD2's on small $A\beta_{42}$ oligomers. Furthermore, we performed exhaustive point mutations of D3 and RD2's sequences and assessed the resulting binding free energies. The obtained results indicate that D3 is already highly optimised for $A\beta_{42}$ monomer in agreement with the original phage display selection that identified D3 under a set of solution conditions where monomeric $A\beta_{42}$ was expected to be the dominant species. On the other hand, substitutions with natural and non-natural amino acids with single aromatic sidechains were observed to improve binding to $A\beta_{42}$ pentamer.

While our use of representative hundreds of $A\beta_{42}$ -D-peptide systems successfully captured important aspects of the interaction, it is still important to acknowledge that this number represents only a very small fraction of the number of the interacting partners present in a typical test tube. In particular, by selecting the top 100 complexes in each system our results have become more narrowed to the set of systems analysed. The analyses would doubtlessly have better modelled the ensemble averaged properties existing *in vitro* had we included the entire 5000 GMIN-generated complexes. This, however, is currently beyond the computational resources available for the project. Also, the mutational analyses employed only the binding configurations of the selected top D3- $A\beta$ complexes. For instance, the global energy minimum of each mutant may be expected to differ from that

of the parent D-peptide-A β complex, and by ignoring this difference the mutation study could not completely account for all energetic and configurational possibilities. In other words, configurational entropy of the interacting peptide and inhibitors was not included. On the other hand, fully sampling each of the ≈ 700 point mutations studied in this work would require basin-hopping optimisation, binding energy calculation, and explicit solvent MD simulations for about 3,500,000 different configurations!

Future work should consider the full optimisation of a number of point mutations selected based, for instance, on the binding energies reported in this study. With this, systematic chemical modifications may be employed for designing D-peptide-derived inhibitors with specificity for different A β oligomer size. In addition, the effect of varying the A β -D-peptide stoichiometric ratio will also be conducted which will help to determine the optimal number of D-peptides necessary to produce the largest structural effect on A β . This study will require the profiling of A β aggregation process using atomistic MD simulations of up to 20 for each of A β_{40} and A β_{42} peptides.

7.1 List of abbreviations

Aβ	Amyloid beta
AD	Alzheimer's disease
AFM	Atomic force microscopy
APP	Amyloid precursor peptide
BH	Basin hopping
BBB	Blood-brain barrier
CHC	Central hydrophobic core
COM	Centre of mass
CSF	Cerebrospinal fluid
LTP	Long-term potentiation
MD	Molecular dynamics
NMR	Nuclear magnetic resonance
QM	Quantum mechanics
SRU	Self-recognition unit
ThT	Thioflavin T

7.2 Amino acids codes

Ala (A)	Alanine
Arg (R)	Arginine
Asn (N)	Asparagine
Asp (D)	Aspartic acid
Cys (C)	Cysteine
Gln (Q)	Glutamine
Glu (E)	Glutamic acid
Gly (G)	Glycine
His (H)	Histidine
Ile (I)	Isoleucine
Leu (L)	Leucine
Lys (K)	Lysine
Met (M)	Methionine
Phe (F)	Phenylalanine
Pro (P)	Proline
Ser (S)	Serine
Thr (T)	Threonine
Trp (W)	Tryptophan
Tyr (Y)	Glutamic acid
Val (V)	Valine

7.3 Article reuse permission



RightsLink®

[Home](#)[Create Account](#)[Help](#)

ACS Publications Title:
High quality. High impact.

Structures of the Amyloid β -Peptides A β 1–40 and A β 1–42 as Influenced by pH and a d-Peptide

Author: Olujide O. Olubiyi and Birgit Strodel

Publication: The Journal of Physical Chemistry B

Publisher: American Chemical Society

Date: Mar 1, 2012

Copyright © 2012, American Chemical Society

User ID
<input type="text"/>
Password
<input type="text"/>
<input type="checkbox"/> Enable Auto Login
<input type="button" value="LOGIN"/>
Forgot Password/User ID?
If you're a copyright.com user, you can login to RightsLink using your copyright.com credentials. Already a RightsLink user or want to learn more?

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

[BACK](#)[CLOSE WINDOW](#)

Copyright © 2013 [Copyright Clearance Center, Inc.](#) All Rights Reserved. [Privacy statement.](#)
Comments? We would like to hear from you. E-mail us at customercare@copyright.com

Bibliography

- [1] Laplace PS, *Philosophical essay on probabilities*, Dover, New York, 1951.
- [2] Becker OM, MacKerell Jr AD, Roux B, Watanabe M, *Computational biochemistry and biophysics*, Marcel Dekker, Inc. 270 Madison Avenue, NY, 2001.
- [3] Scheraga HA, Khalili H, Liwo A, *Protein-Folding Dynamics: Overview of Molecular Simulation Techniques*, *Annu Rev Phy Chem.*,58: 57-83, 2007.
- [4] van der Spoel C, Lindahl E, Hess B, Kutzner C, van Buuren AR, et al., *GROMACS user manual version 4.5*, *J Comput Chem*, 25:1656-76, 2004.
- [5] Ippoliti E, *Parallel Algorithms for Short-Range Molecular Dynamics What is Molecular Dynamics? Biophysics Application Project*, September 14, 2011.
- [6] M.S. Shell, *Molecular Dynamics: CHE210D Lecture notes*, 2009.
- [7] Tsai CS, *An introduction to computational biochemistry*, Wiley-Liss, Inc., New York, 2002.
- [8] Boltzmann L, *Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen (Further studies on the temperature equilibrium of gas molecules)*, *Wissenschaftliche Abhandlungen*, 1:316-402, 1872.

- [9] Oostenbrink C, Villa A, Mark, AE, van Gunsteren WF, *A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6*, J Comput Chem, 25:1656-76, 2004.
- [10] Liljefors T, *Computational Medicinal Chemistry for Drug Discovery*, edited by Bultinck, Marcel Dekker, New York, 2004, pp1-284.
- [11] Jensen F, *Introduction to Computational Chemistry Second Edition*, John Wiley & Sons Ltd, West Sussex, England, 2007.
- [12] *Spartan'10 for Windows, Macintosh, and Linux:tutorial and user's guide*, Wavefunction, Inc. 18401, Irvine, CA 92612, USA. 2011.
- [13] , Buckingham RA, *The Classical Equation of State of Gaseous Helium, Neon and Argon*, P Roy Soc Lond A Mat, 168:264-83, 1938.
- [14] Morse PM, *Diatomic molecules according to the wave mechanics. II. vibrational levels*, Phys Rev 34:57-64, 1929.
- [15] Kang J, Lemaire HG, Unterbeck A, et al., *The precursor of Alzheimer's disease amyloid A₄ protein resembles a cell-surface receptor*, Nature, 325:733-6, 1987.
- [16] Glenner GG, Wong CW, *Alzheimer's disease: initial report of the purification and characterization of a novel cerebrovascular amyloid protein*, Biochem Biophys Res Commun, 120:885-90, 1984.
- [17] Seubert P, Vigo-Pelfrey C, Esch F, et al., *Isolation and quantification of soluble Alzheimer's beta-peptide from biological fluids*, Nature, 359:325-7, 1992.
- [18] Shoji M, Golde TE, Ghiso J, et al., *Production of the Alzheimer amyloid beta protein by normal proteolytic processing*, Science, 258:126-9, 1992.

- [19] Haas C, Koo EH, Mellon A, et al., *Targeting of cell-surface beta-amyloid precursor protein to lysosomes: alternative processing into amyloid-bearing fragments*, Nature, 357:500-3, 1992.
- [20] Bartus RT, Dean RL, 3rd, Beer B, Lippa AS, *The cholinergic hypothesis of geriatric memory dysfunction*, Science, 217:408-14,1982.
- [21] Masters CL, and Beyreuther K, *Molecular neuropathology of Alzheimer's disease*, Arzneimittelforschung, 45:410-2,1995.
- [22] Masters CL, and Beyreuther K, *Alzheimer's disease*, Br Med J, 316:446-448,1998.
- [23] Bartus RT, and Emerich DF, *Cholinergic markers in Alzheimer's disease*, J Am Med Assoc, 282:2208-9,1999.
- [24] Selkoe DJ, *Toward a comprehensive theory for Alzheimer's disease. Hypothesis: Alzheimer's disease is caused by the cerebral accumulation and cytotoxicity of amyloid beta-protein*, Ann N Y Acad Sci, 924:17-25,2000.
- [25] Selkoe DJ, *Alzheimer's disease is a synaptic failure*, Science, 298:789-91,2002.
- [26] Masters CL, Beyreuther K, Henryk M, Wisniewski and the amyloid theory of Alzheimer's disease, J Alzheimers Dis,3:83-6,2001.
- [27] Villemagne VL, Cappai R, Barnham KJ, Cherny RA, Opazo C, et al., *The A β centric pathway to Alzheimer's disease*, Abeta Peptide and Alzheimer's Disease,pp 5-19. Springer-Verlag London Limited, 2007.
- [28] Fagarasan MO, Efthimiopoulos S, *Mechanism of amyloid beta-peptide (1-42) toxicity in PC12 cells*, Mol Psychiatry, 1:398-403, 1996.
- [29] Yankner BA, Dawes LR, Fisher S, et al., *Neurotoxicity of a fragment of the amyloid precursor associated with Alzheimer's disease*, Science, 245:417-20, 1989.

- [30] Pike CJ, Walencewicz-Wasserman AJ, Kosmoski J, et al., *Structure-activity analyses of beta-amyloid peptides: contributions of the beta 25-35 region to aggregation and neurotoxicity*, J Neurochem, 64:253-265, 1995.
- [31] Scheuner D, Eckman C, Jensen M, et al., *Secreted amyloid beta-protein similar to that in the senile plaques of Alzheimer's disease is increased in vivo by the presenilin-1 and 2 and APP mutations linked to familial Alzheimer's disease*, Nat Med, 2:864-70, 1996.
- [32] Jarrett JT, Lansbury PT, Jr, *Seeding "one-dimensional crystallization" of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie?*, Cell, 73:1055-8, 1993.
- [33] Lott IT and Head E, *Alzheimer disease and Down syndrome: factors in pathogenesis*, Neurobiology of Aging 26:383-9, 2005.
- [34] Kamino K, et al., *Linkage and mutational analysis of familial Alzheimer disease kindreds for the APP gene region*, Am J Hum Genet, 51:998-1014, 1992.
- [35] Nilsberth C, Westlind-Danielsson A, Eckman CB, Condron MM, Axelman K, et al., *The 'Arctic' APP mutation (E693G) causes Alzheimer's disease by enhanced A β protofibril formation*, Nat Neurosci, 4:887-93, 2001.
- [36] Päiviö A, Jarvet J, Gräslund A, Lannfelt L, Westlind-Danielsson A, *Unique physico-chemical profile of β -amyloid peptide variant A β 1-40 E22G protofibrils: Conceivable neuropathogen in Arctic mutant carriers*, J Mol Biol, 339:145-59, 2004.
- [37] Whalen B, Selkoe D, Hartley D, *Small non-fibrillar assemblies of amyloid β -protein bearing the Arctic mutation induce rapid neuritic degeneration*, Neurobiol Dis 20:254-66, 2005.

- [38] Grabowski TJ, Cho HS, Vonsattel JPG, Rebeck GW, Greenberg SM, *Novel amyloid precursor protein mutation in an Iowa family with dementia and severe cerebral amyloid angiopathy*, Ann Neurol 49:697-705, 2001.
- [39] Murakami K, Irie K, Morimoto A, Ohigashi H, Shindo M, et al *Neurotoxicity and Physicochemical Properties of A Beta Mutant Peptides From Cerebral Amyloid Angiopathy - Implication For the Pathogenesis of Cerebral Amyloid Angiopathy and Alzheimer's Disease*, J Biol Chem, 278:46179-87, 2003.
- [40] Levy E, Carman MD, Fernandez-Madrid IJ, Power MD, Lieberburg I, et.al, *Mutation of the Alzheimer's disease amyloid gene in hereditary cerebral hemorrhage, Dutch-type*, Science, 248:1124-6, 1990.
- [41] Walsh DM, Lomakin A, Benedek GB, Condron MM, Teplow DB, *Amyloid β -protein fibrillogenesis. Detection of a protofibrillar intermediate*, J Biol Chem, 272:22364-72, 1997.
- [42] Van Nostrand W, Melchor J, Cho H, Greenberg S, Rebeck G, *Pathogenic effects of D23N Iowa mutant amyloid β -protein*, J Biol Chem, 276:32860-6, 2001.
- [43] Tomidokoro Y, Rostagno A, Neubert TA, Lu Y, Rebeck G, et al, *Iowa Variant of Familial Alzheimer's Disease - Accumulation of Post-translationally Modified A β D23N in Parenchymal and Cerebrovascular Amyloid Deposits*, Am J Pathol, 176:1841-54, 2010.
- [44] Saunders AM, Strittmatter WJ, et al., *Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease*, Neurology, 43:1467-72, 1993.

- [45] Schenk D, Barbour R, Dunn W, Gordon G, Grajeda H, et al., *Immunization with amyloid-beta attenuates Alzheimer-disease-like pathology in the PDAPP mouse*, Nature 1999, 400:173-7, 1999.
- [46] Hock BJ Jr., Lamb BT, *Transgenic mouse models of Alzheimer's disease*, Trends Genet, 17:S7-12, 2001.
- [47] Lewis J, Dickson DW, Lin WL, et al., *Enhanced neurofibrillary degeneration in transgenic mice expressing mutant tau and APP*, Science, 293:1487-91, 2001.
- [48] Barrow CJ, and Small DH, *A brief introduction to the history of the β -amyloid protein ($A\beta$) of Alzheimer's disease*, Abeta Peptide and Alzheimer's Disease, pp 1-3. Springer-Verlag London Limited, 2007.
- [49] Braak H and Braak E, *Neuropathological staging of Alzheimer-related changes*, Acta Neuropathol (Berl), 82:239-59, 1991.
- [50] Selkoe DJ, *Alzheimer's disease: genes, proteins, and therapy*, Physiol Rev, 81:741-66, 2001.
- [51] Bucciantini M, Giannoni E, Chiti F, Baroni F, Formigli L, et al., *Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases*, Nature, 416:507-11, 2002.
- [52] Caughey B and Lansbury PT, *Protofibrils, pores, fibrils, and neurodegeneration: separating the responsible protein aggregates from the innocent bystanders*, Ann Rev Neurosci 26:267-98, 2003.
- [53] Yang M and Teplow DB, *Amyloid β -Protein Monomer Folding: Free-Energy Surfaces Reveal Alloform-Specific Differences*, J Mol Biol, 384:450-64, 2008.

- [54] Yan Y, and Wang C, *A β_{42} is More Rigid than A β_{40} at the C Terminus: Implications for A β Aggregation and Toxicity*, J Mol Biol, 364:853-62, 2006.
- [55] Sgourakis NG, Yan Y, McCallum SA, Wang C, Garcia A, *The Alzheimer's Peptides A β_{40} and 42 Adopt Distinct Conformations in Water: A Combined MD / NMR Study*, J Mol Biol, 368:1448-57, 2007.
- [56] Urbanc B, Cruz L, Yun S, Buldyrev SV, Bitan G, Teplow DB, Stanley HE, *In silico study of amyloid β -protein folding and oligomerization*, Proc Natl Acad Sci USA, 101:17345-50, 2004.
- [57] Yun S, Urbanc B, Cruz L, Bitan G, Teplow DB, Stanley HE, *Role of Electrostatic Interactions in Amyloid β -Protein (A β) Oligomer Formation: A Discrete Molecular Dynamics Study*, Biophys J, 92:4064-77, 2007.
- [58] Bitan G, Vollers SS, Teplow DB, *Elucidation of Primary Structure Elements Controlling Early Amyloid β -Protein Oligomerization*, J Biol Chem, 278:34882-9, 2003.
- [59] Bitan G, Kirkitadze MD, Lomakin A, Vollers SS, Benedek GB, Teplow DB, *Amyloid β -protein (A β) assembly: A β_{40} and A β_{42} oligomerize through distinct pathways*, Proc Natl Acad Sci USA, 100:330-5, 2003.
- [60] Burdick D, Soreghan B, Kwon, M, Kosmoski J, Knauer M, Henschen A, Yates J, Cotman C, Glabe C, *Assembly and aggregation properties of synthetic Alzheimer's A β /beta amyloid peptide analogs*, J Biol Chem, 267:546-554,1992.
- [61] Atwood CS, Moir RD, Huang X, Scarpa RC, Bacarra NM, Romano DM, Hartshorn MA, Tanzi RE, Bush AI, *Dramatic Aggregation of Alzheimer A β by Cu(II) Is Induced by Conditions Representing Physiological Acidosis*, J Biol Chem, 273:12817-26,1998.
- [62] Syme CD, Nadal RC, Rigby SE, Viles JH, *Copper Binding to the Amyloid- β (A β) Peptide Associated with Alzheimer's Disease: folding, coordination geometry, pH de-*

- pendence, stoichiometry, and affinity of A β -(128): insights from a range of complementary spectroscopic techniques*, J Biol Chem, 279:18169-77,2004.
- [63] Guo M, Gorman PM, Rico M, Chakrabartty A, Laurents DV, *Charge substitution shows that repulsive electrostatic interactions impede the oligomerization of Alzheimer amyloid peptides*, FEBS Lett, 579:3574-8, 2005.
- [64] Vivekanandan S, Brender JR, Lee SY, Ramamoorthy A, *A partially folded structure of amyloid-beta(1-40) in an aqueous environment*, Biochem Biophys Res Commun, 411:312-6, 2011.
- [65] Salomon AR, Marcinowski KJ, Friedland RP, Zagorski MG, *Nicotine inhibits amyloid formation by the beta-peptide*, Biochemistry, 35:13568-78, 1996.
- [66] Ibarreta D, Urcelay E, Parrilla R, Ayuso MS, *Distinct pH homeostatic features in lymphoblasts from Alzheimer's disease patients*, Ann Neurol, 44:216-222, 1998.
- [67] Yates CM, Butterworth J, Tennant MC, Gordon AJ, *Enzyme Activities in Relation to pH and Lactate in Postmortem Brain in Alzheimer-Type and Other Dementias*, Neurochem, 55:1624-30, 1990.
- [68] Olubiyi OO, Strodel B, *Structures of the Amyloid β -Peptides A β ₄₀ and A β ₄₂ as Influenced by pH and a D-Peptide*, J Phys Chem B, 116:3280-91, 2012.
- [69] Coles M, Bicknell W, Watson AA, Fairlie DP, Craik DJ, *Solution structure of amyloid beta-peptide(1-40) in a water-micelle environment. Is the membrane-spanning domain where we think it is?*, Biochemistry, 37:11064-77, 1998.
- [70] Tomaselli S, Esposito V, Vangone P, van Nuland NA, Bonvin AM, et al., *The alpha-to-beta conformational transition of Alzheimer's A β (1-42) peptide in aqueous media is reversible: a step by step conformational analysis suggests the location of beta conformation seeding*, ChemBioChem, 7:257-67,2006.

- [71] Daura X, Mark AE, van Gunsteren WF, *Parametrization of aliphatic CH_n united atoms of GROMOS96 force field*, J Comput Chem, 19:535-47, 1998.
- [72] Ping G, Yuan JM, Vallieres M, Dong H, Sun Z, et al., *Effects of confinement on protein folding and protein stability*, J Biol Chem, 271:8545-8, 1996.
- [73] Kabsch W, Sander C, *Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features*, Biopolymers, 22:2577-37, 1983.
- [74] Daura X, Gademann K, Jaun B, Seebach D, van Gunsteren W, et al., *Peptide Folding: When Simulation Meets Experiment*, Angew Chem Int Ed, 38:236-40, 1999.
- [75] Danielsson J, Andersson A, Jarvet J, Graslund A, *¹⁵N relaxation study of the amyloid β -peptide: structural propensities and persistence length*, Magn Reson Chem, 44:S114-21, 2006.
- [76] Tjernberg LO, Naslund J, Lindqvist F, Johansson J, Karlstrom AR, et al., *Arrest of β -Amyloid Fibril Formation by a Pentapeptide Ligand*, J Biol Chem 271:8545-8, 1996.
- [77] Hou L, Shao H, Zhang Y, Li H, Menon NK, et al., *Solution NMR studies of the A β (1-40) and A β (1-42) peptides establish that the Met35 oxidation state affects the mechanism of amyloid formation*, J Am Chem Soc, 126:1992-2005, 2004.
- [78] Kohlhoff KJ, Robustelli P, Cavalli A, Salvatella X, Vendruscolo M, *Fast and Accurate Predictions of Protein NMR Chemical Shifts from Interatomic Distances*, J Am Chem Soc, 131:13894-5, 2009.
- [79] Wood GPF, Rothlisberger U, *Secondary Structure Assignment of Amyloid- β Peptide Using Chemical Shifts*, J Chem Theory Comput, 7:1552-63, 2011.

- [80] Dalgarno DC, Levine BA, Williams R, *Structural information from NMR secondary chemical shifts of peptide alpha C-H protons in proteins*, Biosci Rep, 3:443-52, 1983.
- [81] Hornak V, Abel R, Okur A, Strockbine B, Roitberg, A, *et al.*, *Comparison of multiple Amber force fields and development of improved protein backbone parameters*, Proteins, 65:712-25, 2006.
- [82] Matthes D, de Groot BL, *Secondary structure propensities in peptide folding simulations: a systematic comparison of molecular mechanics interaction schemes*, Biophys J, 97:599-608, 2009.
- [83] Hess B, Kutzner C, van der Spoel D, Lindahl E., *GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation*, J Chem Theory Comput, 4:435-47, 2008.
- [84] Piana S, Lindorff-Larsen K, Shaw DE, *How robust are protein folding simulations with respect to force field parameterization?*, Biophys J, 100:L47-9, 2011.
- [85] Raffa DF, Rauk AJ, *Molecular dynamics study of the beta amyloid peptide of Alzheimer's disease and its divalent copper complexes*, J Phys Chem B, 111:3789-99, 2007.
- [86] Pagel K, Seri T, von Berlepsch H, Griebel J, Kirmse R, *et al.*, *How metal ions affect amyloid formation: Cu²⁺- and Zn²⁺-sensitive peptides*, ChemBioChem, 2008, 9, 531-536, 2008.
- [87] Barrow CJ, Yasuda A, Kenny PT, Zagorski MG, *Solution conformations and aggregational properties of synthetic amyloid beta-peptides of Alzheimer's disease. Analysis of circular dichroism spectra*, J Mol Biol, 225:1075-93, 1992.
- [88] Su Y, Chang PT, *Acidic pH promotes the formation of toxic fibrils from beta-amyloid peptide*, Brain Res, 893:287-91, 2001.

- [89] Wishart DS, Case DA, *Use of chemical shifts in macromolecular structure determination*, Methods Enzymol, 338:3-34, 2001.
- [90] Cornilescu G, Delaglio F, Bax A, *Protein backbone angle restraints from searching a database for chemical shift and sequence homology*, J Biomol NMR, 13:289-02, 1999.
- [91] Xu XP, Case DA, *Probing multiple effects on ^{15}N , $^{13}\text{C}\alpha$, $^{13}\text{C}\beta$, and $^{13}\text{C}'$ chemical shifts in peptides using density functional theory*, 2002, Biopolymers 65:408-23, 2002.
- [92] Mauri F, Pfrommer BG, Louie SG, *Ab initio theory of NMR chemical shifts in solids and liquids*, Phy Rev Lett, 77:5300-03, 1996.
- [93] Sebastiani D, Parrinello M, *Ab-initio study of NMR chemical shifts of water under normal and supercritical conditions*, ChemPhysChem 3:675-9, 2002.
- [94] Neal S, Nip AM, Zhang H, Wishart DS, *Rapid and accurate calculation of protein ^1H , ^{13}C and ^{15}N chemical shifts*, J Biomol NMR, 26:215-40, 2003.
- [95] Xu XP, Case DA, *Automated prediction of ^{15}N , $^{13}\text{C}\alpha$, $^{13}\text{C}\beta$ and $^{13}\text{C}'$ chemical shifts in proteins using a density functional database*, J Biomol NMR, 21:321-33, 2001.
- [96] Meiler J, *PROSHIFT: protein chemical shift prediction using artificial neural networks*, J Biomol NMR, 26:25-37, 2003.
- [97] Shen Y, Bax A, *Protein backbone chemical shifts predicted from searching a database for torsion angle and sequence homology*, J Biomol NMR, 38:289-302, 2007.
- [98] Smith DL, Pozueta J, Gong B, Arancio O, Shelanski M, *Reversal of long-term dendritic spine alterations in Alzheimer disease models*, Proc Natl Acad Sci USA, 106:16877-82, 2009.

- [99] Reymann KG, Frey JU, *The late maintenance of hippocampal LTP: requirements, phases, synaptic tagging, late-associativity and implications*, Neuropharmacology, 52:24-40, 2007.
- [100] Demuro A, Parker I, Stutzmann GE, *Calcium signaling and amyloid toxicity in Alzheimer disease*, J Biol Chem, 285:12463-8, 2010.
- [101] Butterfield SM, Lashuel HA, *Amyloidogenic protein-membrane interactions: mechanistic insight from model systems*, Angew Chem, Int Ed, 49:5628-54, 2010.
- [102] Eliezer D, *Amyloid ion channels: a porous argument or a thin excuse*, J Gen Physiol, 128:631-3, 2006.
- [103] Furukawa K, Barger SW, Blalock EM, Mattson MP, *Activation of K⁺ channels and suppression of neuronal activity by secreted β -amyloid-precursor protein*, Nature, 379:74-8, 1996.
- [104] Ittner LM, Götz J, *Amyloid- β and tau—a toxic pas de deux in Alzheimer’s disease*, Nat Rev, Neurosci, 12:65-72, 2011.
- [105] Du H, Guo L, Fang F, Chen D, Sosunov AA, et al, *Cyclophilin D deficiency attenuates mitochondrial and neuronal perturbation and ameliorates learning and memory in Alzheimer’s disease*, Nat Med, 14:1097-1105, 2008.
- [106] Jakob-Roetne R, H. Jacobsen H, *Alzheimer’s disease: from pathology to therapeutic approaches*, Angew Chem, Int Ed Engl, 48: 3030-59, 2009.
- [107] Terry RD, Masliah E, Salmon DP, Butters N, DeTeresa R, et al, *Physical basis of cognitive alterations in Alzheimer’s disease: Synapse loss is the major correlate of cognitive impairment*, Annals of Neurology, 30:572-80, 1991.

- [108] Bell KF, Ducatenzeiler A, Ribeiro-da-Silva A, Duff K, Bennett DA, Cuellar AC, *The amyloid pathology progresses in a neurotransmitter-specific manner*, *Neurobiology of Aging*,27:1644-57, 2006.
- [109] Bell KF, and Cuellar AC, *Altered synaptic function in Alzheimer's disease*, *Eur J Pharmacology*, 545:11-21, 2006.
- [110] Bartus RT, Dean RL III, Beer B, Lippa AS, *The cholinergic hypothesis of geriatric memory dysfunction*, *Science*,217:408-14, 1982.
- [111] Davis KL, Mohs RC, Tinklenberg JR, *Enhancement of memory by physostigmine*, *New Engl J Med*, 301:946, 1979.
- [112] Asthana S, Raffaele KC, Berardi A, Greig NH, Haxby JV, et al, *Treatment of Alzheimer's disease by continuous intravenous infusion of physostigmine*, *Alzheimer Dis Assoc Disord*, 9:223-32, 1995.
- [113] Thal LJ, Ferguson JM, Mintzer J, Raskin A, Targum SD, *A 24-week randomized trial of controlled-release physostigmine in patients with Alzheimer's disease*, *Neurology*, 52:1146-52, 1999.
- [114] Summers WK, Majovski LV, Marsh GM, Tachiki K, Kling A, *Oral tetrahydroaminoacridine in long-term treatment of senile dementia, Alzheimer's disease type*, *New Engl J Med*, 315:1241-5, 1986.
- [115] Watkins PB, Zimmerman HJ, Knapp MJ, Gracon SI, Lewis KW, *Hepatotoxic effects of tacrine administration in patients with Alzheimer's disease*, *Am J Med*, 271:992-998, 1994.
- [116] Qizilbash N, Birks J, Lopez-Arrieta J, Lewington S, Szeto S, *Tacrine for Alzheimer's disease [Review]*, *Cochrane Database for Systematic Reviews*, CD000202.

- [117] Wilcock GK, *Book chapter: The pharmacological Treatment of Alzheimer's disease with cholinesterase inhibitors and memantine*, Pharmacological Mechanisms in Alzheimer's Therapeutics, ed Cuello AC, Springer 2008, pp36-49.
- [118] Dovey HF, John V, Anderson JP, Chen LZ, de Saint Andrieu P, et al, *Functional gamma-secretase inhibitors reduce beta-amyloid peptide levels in brain*, J Neurochem, 76:173-81, 2001.
- [119] Lanz TA, Himes CS, Pallante G, Adams L, Yamazaki S, et al, *The gamma-secretase inhibitor N-[N-(3,5-difluorophenacetyl)-L-alanyl]-S-phenylglycine t-butyl ester reduces A beta level in vivo in plasma and cerebrospinal fluid in young (plaque-free) and aged (plaque-bearing) Tg2576 mice*, J Pharmacol Exp Ther, 305:864-71, 2003.
- [120] Anderson JJ, Holtz G, Baskin PP, Turner M, Rowe B, et al, *Reductions in beta-amyloid concentrations in vivo by the gamma-secretase inhibitors BMS-289948 and BMS-299897*, Biochem Pharmacol, 69:689-98, 2005.
- [121] Barten DM, Guss VL, Corsa JA, Loo A, Hansel SB, et al, *Dynamics of beta-amyloid reductions in brain, cerebrospinal fluid and plasma of beta-Amyloid Precursor Peptide transgenic mice treated with a gamma-secretase inhibitor*, J Pharmacol Exp Ther, 27:27, 2004.
- [122] Best JD, Jay MT, Out F, Ma J, Nadin A, et al, *Quantitative measurement of changes in A β (40) in the rat brain and cerebrospinal fluid following treatment with the gamma-secretase inhibitor LY-411575 [N2-[(2S)-2(3,5-difluorophenyl)-2-hydroxyethanoyl]-N1-[(7S)-5-methyl-6-oxo-6,7-dihydro-5H-dibenzo[b,d]azepin-7-yl]-L-alaninamide]*, J Pharmacol Exp Ther, 313:902-8, 2005.

- [123] Siemers E, Skinner M, Dean RA, Gonzales C, Satterwhite J, et al, *Safety, tolerability, and changes in A β concentrations after administration of a gamma-secretase inhibitor in volunteers*, Clin Neuropharmacol, 28:126-32, 2005.
- [124] Siemers ER, Quinn JF, Kaye J, Farlow MR, Porsteinsson A, et al, *Effects of gamma-secretase inhibitor in randomized study of patients with Alzheimer's disease*, Neurology, 66:602-4, 2006.
- [125] Artavanis-tsakonas S, Matsuno K, Fortini ME, *Notch signalling*, Science, 268:225-32, 1995.
- [126] Wong HK, Sakurai T, Oyama F, Kaneko K, Wada K, et al, *Beta subunits of voltage-gated sodium channels are novel substrates of beta-site Amyloid Precursor Peptide-cleaving enzyme (BACE1) and gamma-secretase*, J Bio Chem, 280:23009-17, 2005.
- [127] Kim DY, Ingano LAM, Carey BW, Pettingell WH, Kovacs DM, *Presenilin/gamma-secretase-mediated cleavage of the voltage-gated sodium channel beta2-subunit regulates cell adhesion and migration*, J Bio Chem, 280:8589-95, 2005.
- [128] Cowan JW, Wang X, Guan R, He K, Jiang J, et al, *Growth hormone receptor is a target for presenilin-dependent gamma-secretase cleavage*, J Bio Chem, 280:19331-42, 2005.
- [129] Cai J, Jiang WG, Grant MB, Boulton M, *Pigment epithelium-derived factor inhibits angiogenesis via regulated intracellular proteolysis of vascular endothelial growth factor receptor 1*, J Biol Chem, 281:3604-13, 2006.
- [130] Hong L, Koelsch G, Lin X, Wu S, Terzyan S, et al, *Structure of the protease domain of memapsin 2 (beta-secretase) complexed with inhibitor*, Science, 290:150-3, 2000.
- [131] Hussain I, *The potential for BACE-1 inhibitors in the treatment of Alzheimer's disease*, IDrugs, 7:653-8, 2004.

- [132] Citron M, *Beta-secretase inhibition for the treatment of Alzheimer's disease—promise and challenge*, Trends Pharmacol Sci, 25:92-7, 2004.
- [133] Teplow DB, *Structural and kinetic features of amyloid beta-protein fibrillogenesis*, Amyloid, 5:121-42, 1998.
- [134] Levine H, III, *Soluble multimeric Alzheimer beta(1-40) pre-amyloid complexes in dilute solution*, Neurobiology of Aging, 16:755-64, 1995.
- [135] Klein WL, Krafft GA, Finch CE, *Targeting small A β oligomers: the solution to an Alzheimer's disease conundrum?*, Trends Neurosci, 24:219-24, 2001.
- [136] Soto C, *Plaque busters: Strategies to inhibit amyloid in Alzheimer's disease*, Mol Med Today, 5:343-350, 1999.
- [137] Soto P, Griffin MA, Shea JE, *New insights into the mechanism of Alzheimer amyloid-beta fibrillogenesis inhibition by N-methylated peptides*, Biophys J, 93:3015-25, 2007.
- [138] Scherzer-Attali R, Pellarin R, Convertino M, Frydman-Marom A, Egoz-Matia N, et al., *Complete phenotypic recovery of an Alzheimer's disease model by a quinone-tryptophan hybrid aggregation inhibitor*, PLoS One, 5:1-15, 2010.
- [139] Estrada LD, Lasagna C, Soto C, *Design of Inhibitors of Amyloid- β misfolding and aggregation for Alzheimer's therapy*, Book chapter: Pharmacological mechanisms in Alzheimer's therapeutics, Springer Science+Business Media, LLC, 2008.
- [140] Ono K, Hasegawa K, Naiki H, Yamada M, *Curcumin has potent anti-amyloidogenic effects for Alzheimer's beta-amyloid fibrils in vitro*, J Neurosci Res, 75:742-50, 2004.
- [141] Pappolla M, Bozner P, Soto C, Shao H, Robakis NK, et al., *Inhibition of Alzheimer's beta-fibrillogenesis by melatonin*, J Bio Chem, 273:7185-8, 1998.

- [142] Ono K, Yoshiike Y, Takashima A, Hasegawa K, Naiki H, et al., *Potent anti-amyloidogenic and fibril-destabilizing effects of polyphenols in vitro: implications for the prevention and therapeutics of Alzheimer's disease*, J Neurochem, 87:172-81, 2003.
- [143] Liu R, Barkhordarian H, Emadi S, Park CB, Sierks MR, et al., *Trehalose differentially inhibits aggregation and neurotoxicity of beta-amyloid 40 and 42*, Neurobiol Dis, 20:74-81, 2005.
- [144] Tomiyama T, Shoji A, Kataoka K, Suwa Y, Asano S, et al., *Inhibition of amyloid beta protein aggregation and neurotoxicity by rifampicin. Its possible function as a hydroxyl radical scavenger*, J Biol Chem, 271:6839-44, 1996.
- [145] Bard F, Cannon C, Barbour R, Burke RL, Games D, et al., *Peripherally administered antibodies against amyloid beta-peptide enter the central nervous system and reduce pathology in a mouse model of Alzheimer disease*, Nat Med, 6:916-9, 2000.
- [146] Pfeifer M, Boncristiano S, Bondolfi L, Stalder A, Deller T, et al., *Cerebral hemorrhage after passive anti-Abeta immunotherapy*, Science, 298:1379, 2002.
- [147] Orgogozo JM, Gilman S, Dartigues JF, Laurent B, Puel M, et al., *Subacute meningoencephalitis in a subset of patients with AD after Abeta42 immunization*, Neurology, 61:46-54, 2003.
- [148] Hock C, Konietzko U, Streffer JR, Tracy J, Signorell A, et al., *Antibodies against beta-amyloid slow cognitive decline in Alzheimer's disease*, Neuron, 38:547-54, 2003.
- [149] Nerelius C, Sandegren A, Sargsyan H, Raunak R, Leijonmarck H, et al., *Alpha-helix targeting reduces amyloid-beta peptide toxicity*, J Proc Natl Acad Sci USA, 106, 9191-6, 2009.
- [150] Ghanta J, Shen CL, Kiessling LL, Murphy RM, *A strategy for designing inhibitors of beta-amyloid toxicity*, J Biol Chem, 271:29525-8, 1996.

- [151] Wood SJ, Wetzel R, Martin JD, Hurle MR, *Prolines and Amyloidogenicity in Fragments of the Alzheimer's Peptide .beta./A4*, *Biochemistry*, 34:724-30, 1995.
- [152] Pallitto MM, Ghanta J, Heinzelman P, Kiessling LL, Murphy RM, *Recognition sequence design for peptidyl modulators of beta-amyloid aggregation and toxicity*, *Biochemistry* 38:3570-8, 1999.
- [153] Findeis MA, Lee JJ, Kelley M, Wakefield JD, Zhang MH, et al., *Characterization of choly-leu-val-phe-phe-ala-OH as an inhibitor of amyloid beta-peptide polymerization*, *Amyloid*, 8:231-41, 2001.
- [154] Findeis MA, Musso GM, Arico-Muendel CC, Benjamin HW, Hundal AM, et al., *Modified-peptide inhibitors of amyloid beta-peptide polymerization*, *Biochemistry*, 38:6791-800, 1999.
- [155] Hughes E, Burke RM, Doig AJ, *Inhibition of toxicity in the beta-amyloid peptide fragment beta-(25-35) using N-methylated derivatives: A general strategy to prevent amyloid formation*, *J Bio Chem*, 275:25109-115, 2000.
- [156] Gordon DJ, Sciarretta KL, Meredith SC, *Inhibition of beta-amyloid(40) fibrillogenesis and disassembly of beta-amyloid(40) fibrils by short beta-amyloid congeners containing N-methyl amino acids at alternate residues*, *Biochemistry*, 40:8237-45, 2001.
- [157] Soto C, Kindy MS, Baumann M, Frangione B, *Inhibition of Alzheimer's amyloidosis by peptides that prevent b-sheet conformation*, *Biochem Biophys Res Commun*, 226:672-80, 1996.
- [158] Blanchard BJ, Hiniker AE, Lu CC, Margolin Y, Yu AS, et al., *Elimination of Amyloid beta Neurotoxicity*, *J Alzheimers Dis*, 2:137-49,2000.

- [159] Blanchard BJ, Konopka G, Russell M, Ingram VM, *Mechanism and prevention of neurotoxicity caused by beta-amyloid peptides: relation to Alzheimer's disease*, Brain Res, 776:40-50, 1997.
- [160] Chalifour RJ, McLaughlin RW, Lavoie L, Morissette C, Tremblay N, et al., *Stereoselective interactions of peptide inhibitors with the beta-amyloid peptide*, J Biol Chem, 278:34874-81, 2003.
- [161] Tjernberg LO, Lilliehook C, Callaway DJ, Naslund J, Hahne S, et al., *Controlling amyloid beta-peptide fibril formation with protease-stable ligands.*, J Biol Chem, 272:12601-5, 1997.
- [162] Wiesehan K, Buder K, Linke RP, Patt S, Stoldt M, et al., *Selection of D-amino-acid peptides that bind to Alzheimer's disease amyloid peptide abeta1-42 by mirror image phage display*, ChemBioChem, 4:748-53, 2003.
- [163] van Groen T, Wiesehan K, Funke SA, Kadish I, Nagel-Steger L, et al., *Reduction of Alzheimer's Disease amyloid plaque load in transgenic mice by D3, a d-enantiomeric peptide identified by mirror image phage display*, ChemMedChem, 3:1848-52, 2008.
- [164] Funke SA, van Groen T, Kadish I, Bartnik D, Nagel-Steger, L, et al., *Oral treatment with the d-enantiomeric peptide D3 improves the pathology and behavior of Alzheimer's Disease transgenic mice*, ACS Chem Neurosci, 1:639-48, 2010.
- [165] Schumacher TNM, Mayr LM, Minor Jr DL, Milhollen MA, Burgess MW, et al., *Identification of D-peptide ligands through mirror-image phage display*, Science, 271:1854-7, 1996.
- [166] LeVine H., III *Thioflavine T interaction with synthetic Alzheimer's disease beta-amyloid peptides: Detection of amyloid aggregation in solution*, Protein Sci, 2:404-10, 1993.

- [167] Wolfe LS, Calabrese MF, Nath A, Blaho DV, Miranker AD, et al., *Protein-induced photophysical changes to the amyloid indicator dye thioflavin T*, Proc Natl Acad Sci USA, 107: 16863-8, 2010.
- [168] Khurana R, Coleman C, Ionescu-Zanetti C, Carter SA, Krishna V, *Mechanism of thioflavin T binding to amyloid fibrils*, J. Struct. Biol. 151:229-38, 2005.
- [169] GMIN: A program for basin-hopping global optimization. Wales, D. J. <http://www.wales.ch.cam.ac.uk/software.html>.
- [170] MacKerell AD, Jr. Bashford D, Bellott M, Dunbrack RL, Jr. Evanseck JD, et al., *All-atom empirical potential for molecular modeling and dynamics study of protein*, J Phys Chem B, 102:3586-616, 1998.
- [171] Brooks BR, Bruccoleri RE, BD Olafson, States DJ, Swaminathan S, et al., *CHARMM: A program for macromolecular energy, minimization, and dynamics calculations*, J Comput Chem, 4:187-217, 1983.
- [172] Lührs T, Ritter C, Adrian M, Riek-Loher D, Bohrmann B, et al., *3D structure of Alzheimer's amyloid- β (1-42) fibrils*, Proc Natl Acad Sci USA, 102:17342-7, 2005.
- [173] Li Z, Scheraga HA, *Monte Carlo-minimization approach to the multiple-minima problem in protein folding*, Proc Natl Acad Sci USA, 84:6611-5, 1987.
- [174] Wales JD, Doye JPK, *Global optimisation by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms*, J Phys Chem A 101:5111-6, 1997.
- [175] Wales DJ, Scheraga HA, *Global optimization of clusters, crystals, and biomolecules*, Science, 285:1368-72, 1999.

- [176] Strodel B, Lee JW, Whittleston CS, Wales DJ, *Transmembrane structures for Alzheimer's A β (1-42) oligomers*, J Am Chem Soc, 132:13300-12, 2010.
- [177] Mortenson PN, Wales DJ, *Energy landscapes, global optimization and dynamics of the polyaniline Ac(ala)₈NHMe*, J Chem Phys, 114:6443-54, 2001.
- [178] Strodel B, Wales DJ, *Implicit Solvent Models and the Energy Landscape for Aggregation of the Amyloidogenic KFFE Peptide*, J Chem Theory Comput, 4:657-72, 2008.
- [179] Haberthur U, Caffisch A, *FACTS: fast analytical continuum treatment of solvation*, J Comput Chem, 29:701-15, 2008.
- [180] Liu H, Funke SA, Willbold D, *Transport of Alzheimer's disease amyloid- β -binding D-amino acid peptides across an in vitro blood-brain barrier model*, Rejuvenation Research, 13:210-3, 2010.
- [181] Tanabe S, Shimohigashi Y, Nakayama Y, Makino T, Fujita T, et al., *In vivo and in vitro evidence of blood-brain barrier transport of a novel cationic arginine-vasopressin fragment 4-9 analog*, J Pharmacol Exp Ther, 290:561-8, 1999.
- [182] Drin G, Cottin S, Blanc E, Rees AR, Temsamani J, *Studies on the internalization mechanism of cationic cell-penetrating peptides*, J Biol Chem, 278:31192-201, 2003.
- [183] www.mmts.org.
- [184] Eswar N, Marti-Renom MA, Webb B, Madhusudhan MS, Eramian D, et al, *Comparative Protein Structure Modeling With MODELLER. Current Protocols in Bioinformatics*, John Wiley & Sons, Inc., Supplement 15, 5.6.1-5.6.30, 2006.

-
- [185] Stewart JJP, *Optimization of Parameters for Semiempirical Methods V: Modification of NDDO Approximations and Application to 70 Elements*, J Mol Modeling (Springer), 13:1173-213.
- [186] Swinson J, *Fluorine – A vital element in the medicine chest*, B5 srl Via Cesare da Sesto, 10 20123 Italy, Pharmaceutical Chemistry, 26-27, 2005.
- [187] Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, *et al.*, *Scalable molecular dynamics with NAMD*, J Comput Chem, 26:1781-802, 2005.

Curriculum Vitae

Personal data

Last name: Olubiyi
First name: Olujide
Middle name: Oludayo
Date of birth: 4th June, 1981
Place of birth: Efon-Alaaye, Nigeria
Nationality: Nigeria

Qualifications

PhD: RWTH Aachen University, Germany (2010–2013).
Drug Discovery (M.Sc.): London School of Pharmacy, United Kingdom (2008–2009).
Pharmacy (BPharm): Obafemi Awolowo University, Nigeria (2000–2006).

Signature: