Experimental X-ray crystallography, NMR (Nuclear Magnetic Resonance) spectroscopy, dual polarization interferometry, etc are indeed very powerful tools to determine the 3-Dimensional structure of a protein (including the membrane protein); theoretical mathematical and physical computational approaches can also allow us to obtain a description of the protein 3D structure at a submicroscopic level for some unstable, noncrystalline and insoluble proteins. X-ray crystallography finds the X-ray final structure of a protein, which usually need refinements using theoretical protocols in order to produce a better structure. This means theoretical methods are also important in determinations of protein structures. Optimization is always needed in the computer-aided drug design, structure-based drug design, molecular dynamics, and quantum and molecular mechanics. This paper introduces some optimization algorithms used in these research fields and presents a new theoretical computational method - an improved LBFGS Quasi-Newtonian mathematical optimization method - to produce 3D structures of Prion AGAAAAGA amyloid fibrils (which are unstable, noncrystalline and insoluble), from the potential energy minimization point of view. Because the NMR or X-ray structure of the hydrophobic region AGAAAAGA of prion proteins has not yet been determined, the model constructed by this paper can be used as a reference for experimental studies on this region, and may be useful in furthering the goals of medicinal chemistry in this field.

Keywords: Protein 3D Structure; Computational Approaches; Optimization Method; Molecular Modelling; Prion AGAAAAGA Amyloid Fibrils.

1

Highlights: ▶ Determinations of 3D structures of unstable, noncrystalline and insoluble proteins. ▶ Theoretical computer, mathematical/physical computational approaches and concepts. ▶ Time-consuming, costly X-ray crystallography and NMR spectroscopy powerful tools. ▶ Mathematical local search optimization methods to solve three-body physical problems. ▶ Using computational chemistry Amber/Gromacs packages to refine molecular models.

1 Introduction

Neurodegenerative diseases including Parkinson's, Alzheimer's, Huntington's, and Prion's were found they all featured amyloid fibrils [1, 2, 3, 4, 5, 6]. Amyloid is characterized by a cross-β sheet quaternary structure and recent X-ray diffraction studies of microcrystals revealed atomistic details of core region of amyloid [7, 8]. All the quaternary structures of amyloid cross-β spines can be reduced to one of the 8 classes of steric zippers of [8], with strong van der Waals (vdW) interactions between β-sheets and hydrogen bonds (HBs) to maintain the β-strands. A new era in the structural analysis of amyloids started from the 'steric zipper'- β-sheets [7]. As the two β-sheets zip up, Hydrophobic Packings (HPs & vdWs) have been formed. The extension of the 'steric zipper' above and below (i.e. the β-strands) is maintained by Hydrogen Bonds (HBs) (but usually there is no HB between the two β-sheets). This is the common structure associated with some 20 neurodegenerative amyloid diseases.

We first do some mathematical analysis for the common structure. Let r be the distance between two atoms, the vdW contacts of the two atoms are described by the Lennard-Jones (LJ) potential energy:

$$V_{LJ}(r) = 4\epsilon \left((\frac{\sigma}{r})^{12} - (\frac{\sigma}{r})^6 \right), \tag{1.1}$$

where ϵ is the depth of the potential well and σ is the atom diameter; these parameters can be fitted to reproduce experimental data or deduced from results of accurate quantum chemistry calculations. The $(\frac{\sigma}{r})^{12}$ item describes repulsion and the $-(\frac{\sigma}{r})^6$ item describes attraction (Figure 1). If we introduce the coordinates of the atoms whose number is denoted by N and let $\epsilon = \sigma = 1$ be the reduced units, then, for N atoms, Eq. 1.1 is

$$f(x) = 4 \sum_{i=1}^{N} \sum_{j=1, j<i}^{N} \left(\frac{1}{\tau_{ij}^6} - \frac{1}{\tau_{ij}^3} \right), \tag{1.2}$$

where $\tau_{ij} = (x_{3i-2} - x_{3j-2})^2 + (x_{3i-1} - x_{3j-1})^2 + (x_{3i} - x_{3j})^2$, and $(x_{3i-2}, x_{3i-1}, x_{3i})$ is the coordinates of atom i, $N \geq 2$. The minimization of LJ potential energy $f(x)$
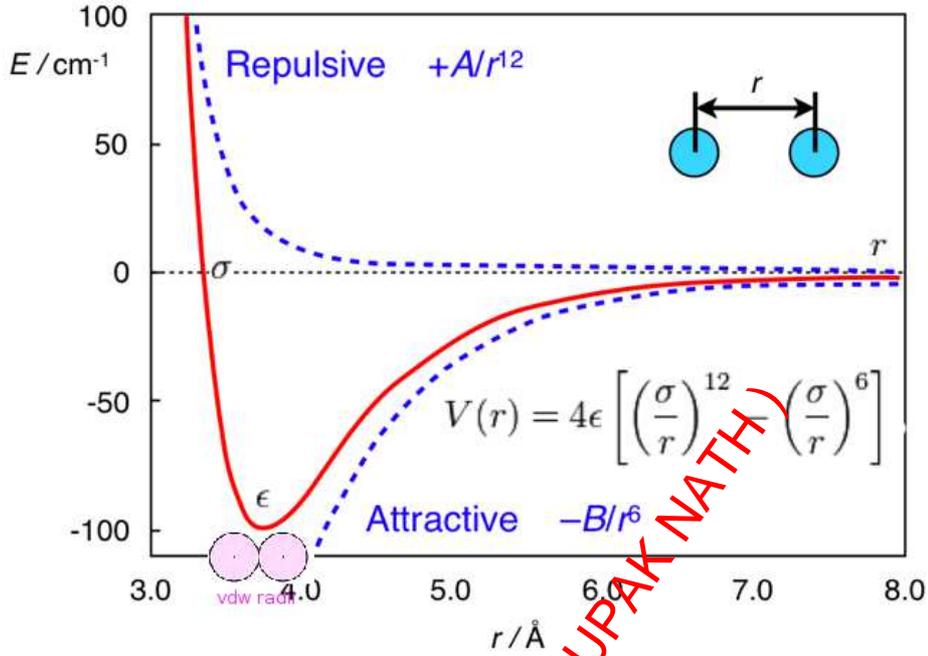
Figure 1: *The LJ potential energy (Eqs. 1.1 and 1.4) (can be seen in Figure 1 of [9]).*

on $\mathbb{R}^n$ (where $n = 3N$) is an optimization problem

$$\min_{x \in \mathbb{R}^n} f(x), \qquad (1.3)$$

which is a well-known and challenging test problem for global optimization (see http://www-wales.ch.cam.ac.uk/CCD.html and its recent references such as [10, 11, 12, 13, 14]). It is very hard for global optimization to directly solve Eq. 1.3 even without a large number of atoms. Similarly as Eq. 1.1 - i.e.

$$V_{LJ}(r) = \frac{A}{r^{12}} - \frac{B}{r^6}, \qquad (1.4)$$

the potential energy for the HBs between β-strands is

$$V_{HB}(r) = \frac{C}{r^{12}} - \frac{D}{r^{10}}, \qquad (1.5)$$

where A, B, C, D are given constants and usually most of the HBs are still kept during the phase of molecular modeling. Thus, the amyloid fibril molecular modeling problem can be reduced to solve the optimization problem Eq. 1.3 though it

3

is not easy to accurately solve Eq. 1.3 for a large molecule.

Alternatively, we have found another way to solve Eq. 1.3 [15]. Seeing Figure 1, we may know that the optimization problem Eq. 1.3 reaches its optimal value at the bottom of the LJ potential well, where the distance between two atoms equals to the sum of vdW radii of the two atoms. Hence, the amyloid fibril molecular modeling problem can be looked as a molecular distance geometry problem (MDGP) [16]. As an example to explain MDGP, the problem of locating sensors in telecommunication networks is a DGP. In such a case, the positions of some sensors are known (which are called anchors) and some of the distances between sensors (which can or cannot be anchors) are known. The DGP is to locate the positions of all the sensors. The MDGP looks sensors as atoms and their telecommunication network as a molecule. In mathematics, the following Eqs. 1.6~1.8 can express the MDGP for Eq. 1.3. The 3D structure of a molecule with N atoms can be described by specifying the 3D coordinate positions $x_1, x_2, \ldots, x_N \in \mathbb{R}^3$ of all its atoms. Given bond lengths d_{ij} between a subset S of the atom pairs, the determination of the molecular 3D structure is

$$(\mathcal{P}_0) \quad to \quad find \quad x_1, x_2, \ldots, x_N \quad such \ that \quad ||x_i - x_j|| = d_{ij}, (i,j) \in S, \tag{1.6}$$

where $|| \cdot ||$ denotes a norm in a real vector space and in this paper it is calculated as the Euclidean distance 2-norm. Eq. 1.6 can be reformulated as a mathematical global optimization problem

$$(\mathcal{P}) \quad \min P(X) = \sum_{(i,j) \in S} w_{ij}(||x_i - x_j||^2 - d_{ij}^2)^2 \tag{1.7}$$

in the terms of finding the global minimum of the function $P(X)$, where $w_{ij}, (i,j) \in S$ are positive weights, $X = (x_1, x_2, \ldots, x_N)^T \in \mathbb{R}^n$ [17] and usually S has fewer elements than $N^2/2$ due to the error in the theoretical or experimental data [18, 16]. Even there may not exist any solution $x_1, x_2, \ldots, x_N$ to satisfy the distance constraints in Eq. 1.6, for example when data for atoms $i, j, k \in S$ violate the triangle inequality; in this case, we may add a perturbation item $-\varepsilon^T X$ to $P(X)$:

$$(\mathcal{P}_\varepsilon) \quad \min P_\varepsilon(X) = \sum_{(i,j) \in S} w_{ij}(||x_i - x_j||^2 - d_{ij}^2)^2 - \varepsilon^T X, \tag{1.8}$$

where $\varepsilon \geq 0$. In some cases, instead exact values $d_{ij}, (i,j) \in S$ can be found, we can only specify lower and upper bounds on the distances: $l_{ij} \leq ||x_i - x_j|| \leq u_{ij}, (i,j) \in S$; in such cases we may penalize all the unsatisfied constraints into the objective function of $(\mathcal{P}_\varepsilon)$ by adding $\sum_{(i,j) \in S} \left(\max \left\{ l_{i,j}^2 - ||x_i - x_j||^2, 0 \right\} \right)^2 +$

$\left(\max\left\{||x_i - x_j||^2 - u_{i,j}^2, 0\right\}\right)^2$ into $P_\varepsilon(X)$ [18, 16], where we may let d_{ij} be the interatomic distance (less than 6 angstroms) for the pair in successive residues of a protein and set $l_{ij} = (1 - 0.05)d_{ij}$ and $u_{ij} = (1 + 0.05)d_{ij}$ [16]. In this paper, we aim to solve Eq. 1.8 (or Eq. 1.3) for modeling amyloid fibril molecular 3D structures.

Neurodegeneration is the progressive loss of structure or function of neurons, including death of neurons. A prion is a misshapen protein that acts like an infectious agent (but not requiring either DNA, RNA, or both) to cause a number of fatal diseases. Prion diseases are rich in β-sheets (compared with the normal prion protein PrP^C in rich of α-helices) and are so-called "protein structural conformational" diseases. The normal hydrophobic region 113–120 AGAAAAGA peptide of prion proteins is an inhibitor/blocker of prion diseases. PrP lacking this palindrome could not convert to prion diseases. Brown et al. pointed out that the AGAAAAGA peptide was found to be necessary (though not sufficient) for blocking the toxicity and amyloidogenicity of PrP 106–126, and the peptide AGAA does not form fibrils [19]. The minimum sequence necessary for fibril formation should be AGAAA, AGAAAA, AGAAAAG, AGAAAAGA and GAAAAGA, but the molecular structures of these fibrils have not known yet. This paper addresses an important problem on modeling the 3D molecular structures of prion AGAAAAGA amyloid fibrils of neurodegenerative diseases. The rest of this paper is arranged as follows. In the next section, i.e. Section 2, an improved LBFGS Quasi-Newtonian method is presented for solving Eq. 1.8. Section 3 implements this Quasi-Newtonian method by constructing an 3D molecular structure of prion AGAAAAGA amyloid fibrils of neurodegenerative prion diseases. Numerical results of computations show that the method designed in Section 2 is very effective and successful. This concluding remark will be made in the last section, i.e. Section 4.

2 Methods

In a (macro)molecular system, if it is very far from equilibrium, then the forces may be excessively large, a robust energy minimization (EM) is required; another reason to perform an EM is the removal of all kinetic energy from the system: EM reduces the thermal noise in the structures and potential energies [20]. EM, with the images at the endpoints fixed in space, of the total system energy provides a minimum energy path. EM can be done using steepest descent (SD), conjugate gradient (CG), and Limited-memory Broyden-Fletcher-Goldfarb-Shanno (LBFGS) methods.

Three kinds of possible EM methods are: (1) derivative-free methods - that require only function evaluations, e.g. the simplex method and its variants; (2) derivative information methods - the partial derivatives of the potential energy with respect to all coordinates are known and the forces are minimized, e.g. SD, CG methods; and (3) second derivative information methods, e.g. LBFGS method. "SD is based on the observation that if the real-valued function $f(x)$ is defined and differentiable in a neighborhood of a point x_0 then $f(x)$ decreases fastest if one goes from x_0 in the direction of the negative gradient of $f(x)$ at x_0 and SD local search method converges fast [21]. SD is robust and easy to implement but it is not most efficient especially when closer to minimum; at this moment, we may use the efficient CG. CG is slower than SD in the early stages but more efficient when closer to minimum. CG algorithm adds an orthogonal vector to the current direction of the search, and then moves them in another direction nearly perpendicular to this vector. The hybrid of SD-CG will make SD or CG more efficient than SD or CG alone. However, CG cannot be used to find the EM path, for example, when "forces are truncated according to the tangent direction, making it impossible to define a Lagrangian" [22, 23]. In this case, the powerful and faster quasi-Newtonian method (e.g. the LBFGS quasi-Newtonian minimizer) can be used [22, 24, 25, 26, 27, 28]. We briefly introduce the LBFGS quasi-Newtonian method as follows.

Newton's method in optimization explicitly calculates the Hessian matrix of the second-order derivatives of the objective function and the reverse of the Hessian matrix [29]. The convergence of this method is quadratic, so it is faster than SD or CG. In high dimensions, finding the inverse of the Hessian is very expensive. In some cases, the Hessian is a non-invertible matrix, and furthermore in some cases, the Hessian is symmetric indefinite. Quasi-Newton methods thus appear to overcome all these shortcomings.

Quasi-Newton methods (a special case of variable metric methods) are to approximate the Hessian. Currently, the most common quasi-Newton algorithms are the SR1 formula, the BHHH method, the widespread BFGS method and its limited/low-memory extension LBFGS, DFP, MS, and Broyden's methods [30, 31, 32, 33]. In Amber [23] and Gromacs [20], LBFGS is used, and the hybrid of LBFGS with CG - a Truncated Newton linear CG method with optional LBFGS Preconditioning [25] - is used in Amber [23].

For BFGS method, whether it converges at all on nonconvex problems is still an open problem. In fact, Powell (1984) gave a counter-example that shows that BFGS with an inexact line-search search may fail to converge [34, 35, 36]. Li and Fukushima (2001) proposed a modified BFGS method for nonconvex objec-

tive function [37]. Basing on [37, 24, 38, 25, 39], in this paper we present an improved LBFGS method described as follows [40] - which presents the non-monotone line search technique [41, 42] for the Wolfe-type search. The improved LBFGS method presented in this paper is much better than the standard BFGS method in view of the CPU time (see Figure 2) tested through more than 30 non-linear programming problems (where each selected problem is regular, that is, its first and second derivatives exist and are continuous everywhere, and each problem is with different dimensions, i.e., 100, 500, 1000 and 10000 dimensions) and its mathematical theory to support this algorithm can be seen from the Supplementary Materials [40] listed at the end of this paper. This paper implements the Wolfe-type search by the approximation technique of piecewise linear/quadratic function [43].

Algorithm 1 An Improved LBFGS Method for minimizing nonconvex function [40]

Step 0: Choose an initial point $x_0 \in \mathbb{R}^n$, an initial positive definite matrix H_0, and choose constants σ_1, σ_2 such that $0 < \sigma_1 < \sigma_2 < 1$, and choose an positive integer m_1. Let $k = 0$.

Step 1: If $\|g_k\| = 0$, then output x_k and stop; otherwise, go to Step 2.

Step 2: Solve the following linear equation to get d_k:

$$d_k = -H_k g_k,$$

Step 3: Find a step-size $\lambda_k > 0$ satisfying the Wolfe-type line search conditions:

$$f(x_k + \lambda_k d_k) \leq f(x_k) + \sigma_1 \lambda_k g_k^T d_k, \tag{2.1}$$

$$g(x_k + \lambda_k d_k)^T d_k \geq \sigma_2 g_k^T d_k. \tag{2.2}$$

Step 4: Let the next iterate by $x_{k+1} = x_k + \lambda_k d_k$. Calculate g_{k+1} and $\|g_{k+1}\|$.

Step 5: Let $s_k = x_{k+1} - x_k = \lambda_k d_k, y_k = g_{k+1} - g_k, \gamma_k = \|g_k\|$, then $y_k^* = y_k + \gamma_k s_k$.

Step 6: Let $\bar{m} = \min\{k+1, m_1\}$. Update H_k following the formula

$$
\begin{aligned}
H_{k+1} &= (V_k^{*T} \cdots V_{k-\bar{m}+1}^{*T}) H_{k-\bar{m}+1} (V_{k-\bar{m}+1}^* \cdots V_k^*) \\
&+ \omega_{k-\bar{m}+1}^* (V_{k-1}^{*T} \cdots V_{k-\bar{m}+2}^{*T}) s_{k-\bar{m}+1} s_{k-\bar{m}+1}^T (V_{k-\bar{m}+2}^* \cdots V_{k-1}^*) \\
&+ \cdots + \omega_k^* s_k s_k^T,
\end{aligned}
\tag{2.3}
$$

where $\omega_k^* = \frac{1}{s_k^T y_k^*}$, $V_k^* = I - \omega_k^* y_k^* s_k^T$ (when $k = 0$, $H_0 = \frac{y_0^T s_0}{\|y_0\|^2}$, $V_0^* = I - \frac{y_0^* s_0^T}{s_0^T y_0^*}$, $H_1 = (I - \frac{y_0^* s_0^T}{s_0^T y_0^*})^T H_0 (I - \frac{y_0^* s_0^T}{s_0^T y_0^*}) + \frac{1}{s_0^T y_0^*} s_k s_k^T$).

Step 7: Let $k = k + 1$ and go to Step 1.

In Algorithm 1, g_k denotes the gradient of $f(x)$ at x_k, and the convergence and the R-linear convergent rate of Algorithm 1 are guaranteed by the following assumptions: (A1) f is twice continuously differentiable, (A2) f has Lipschitz continuous gradients and Hessians, (A3) the Hessian at the stationary point is always positively definite, and (A4) the level set of f is bounded. The detailed proof of the convergence and R-linear convergent rate of Algorithm 1 can be found in the Supplementary Materials supplied at the end of this paper.
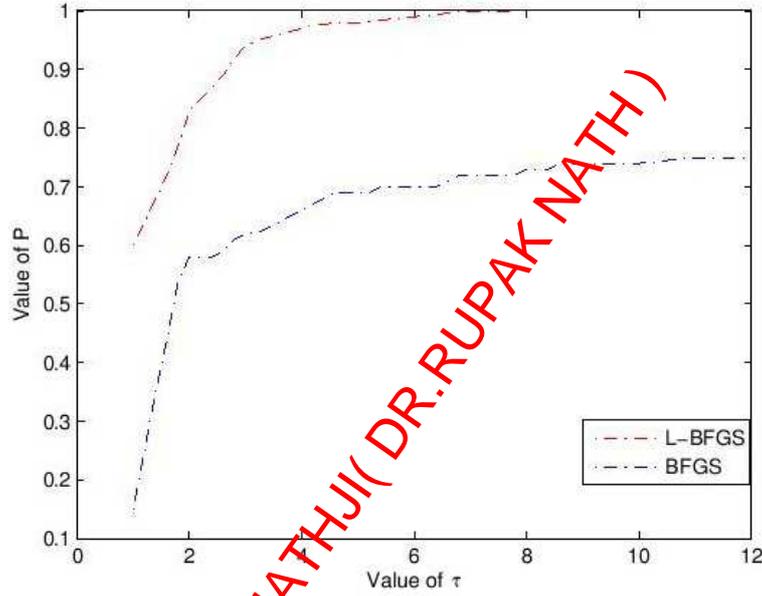
Figure 2: *Performance based on CPU time, where the definitions of P and τ can be found in [44].*

In [40], the following nonmonotone line search technique for stepsize λ_k is used:

$$f(x_k + \lambda_k d_k) \leq \max_{0 \leq j \leq M_0} f(x_{k-j}) + \sigma_1 \lambda_k g_k^T d_k,$$

$$g(x_k + \lambda_k d_k)^T d_k \geq \max\{\sigma_2, 1 - (\lambda_k ||d_k||)^p\} g_k^T d_k,$$

where $p \in (-\infty, 1)$, and M_0 is a nonnegative integer. This is a difference between the algorithm of [40] and this paper. All in all, it is well known that quasi-Newton method is an efficient solution method for unconstrained and continuously differentiable minimization problem [45, 46, 47]. However, it needs computing and storage of the updated matrix which is an approximation to the Hessian matrix at each iteration of the method. Hence, its efficiency may decrease when it is applied to large scale optimization problem. To overcome the drawback, limited memory

8

quasi-Newton method is proposed [48]. The main ideal of this method is nearly identical to that of the standard BFGS method, and the only difference is that the inverse Hessian approximation is not formed explicitly, but defined by a small number, say $\bar{m}$, of BFGS updates. This technique received much attention in recent years and numerical experiments show that it is very competitive [49, 24], and its global convergence and R-linear convergence rate with Wolfe line search are established for the uniformly convex case [26, 24]. Since the limited memory BFGS method may suffer from ill-conditions for small value of $\bar{m}$, Al-Baali (1999) [50] made some modifications to the method and establish its global convergence based on the same assumptions, and Byrd et al. (1994) [51] derives new representation of limited memory quasi-Newton matrices for the benefit of computing the updated matrix. Recently, a non-monotone line search is introduced, see e.g., [41, 42]. Then it is showed to be more competitive and practical for solving nonlinear optimization problems, and [52] established the global convergence of this line search applied to limited memory BFGS method based on the uniformly convex assumption. Motivated by the above observation, it turn out that in two respects the limited memory BFGS method is much less effective. First, we note that the convergence analysis of these method are focused on the uniformly convex assumption and little is known for nonconvex case. Second, numerical experiments have suggested the main weakness of limited memory method is that it may converge very slowly in terms of number of iterations for ill-conditioned problems. The purpose of the above Algorithm 1 is to reduce these defects and Figure 2 shows the effectiveness of the proposed algorithm. We will apply it into the molecular modeling of prion AGAAAAGA amyloid fibrils in the next section.

3   Results and Discussion

From their research of prion, scientists found that the cross-β structure of peptides is with the nature of self-aggregation, the self-aggregating to form fibers. This provides us a new research idea for nanomaterials. HBs can be formed between peptide β-strands, and one peptide monomer connects together with another in accordance with the specific structure to form fibers. Many laboratories in the world are synthesizing peptides that can self-aggregate to form fibers, and want to be able to control the growth of the fiber to find out new functional materials [53, 54]. The studies of this paper not only benefit nanometerials research, but also benefit the research on neurodegenerative amyloid fibril diseases. Prion AGAAAAGA peptide has been reported to own an amyloid fibril forming property (initially described in 1992 by Gasset et al. of Prusiner's Group) [19, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78], but there has not been experimental structural bioinformatics for this segment yet due to the

unstable, noncrystalline and insoluble nature of this region. Furthermore, Zhang (2011) did accurate calculations to confirm the amyloid fibril property at this region (Figure 3) [79].
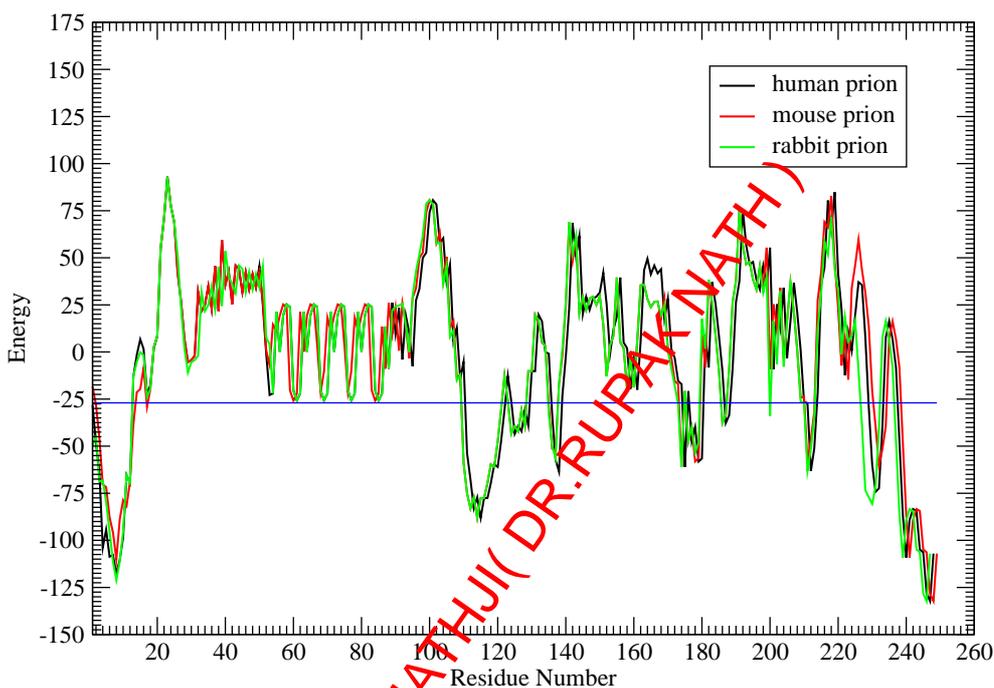
Amyloid fibril identification for prions

Figure 3: *Prion AGAAAAGA (113-120) segment is clearly and surely identified as the amyloid fibril formation region, because its energy is less than the amyloid fibril formation threshold energy -26 kcal/mol [80].*

3.1 Material for the Molecular Modeling

This paper uses a suitable pdb file template 2OMP.pdb (the LYQLEN peptide derived from human insulin residues 13-18 [8]) from the Protein Data Bank to build an 8-chain AGAAAAGA prion amyloid fibril molecular model to illuminate Algorithm 1 works very well. To choose 2OMP.pdb (Figure 4) as the modeling template is due to it can pass all the long procedures of SDCG-SA (equilibrations & productions)-SDCG of [79]. By observations of Figure 4 and the 2nd column of coordinates of 2OMP.pdb, we know that E(F) chains can be calculated on the XZ-plane from A(B) chains by Eq. 3.1 and other chains can be got by a parallel
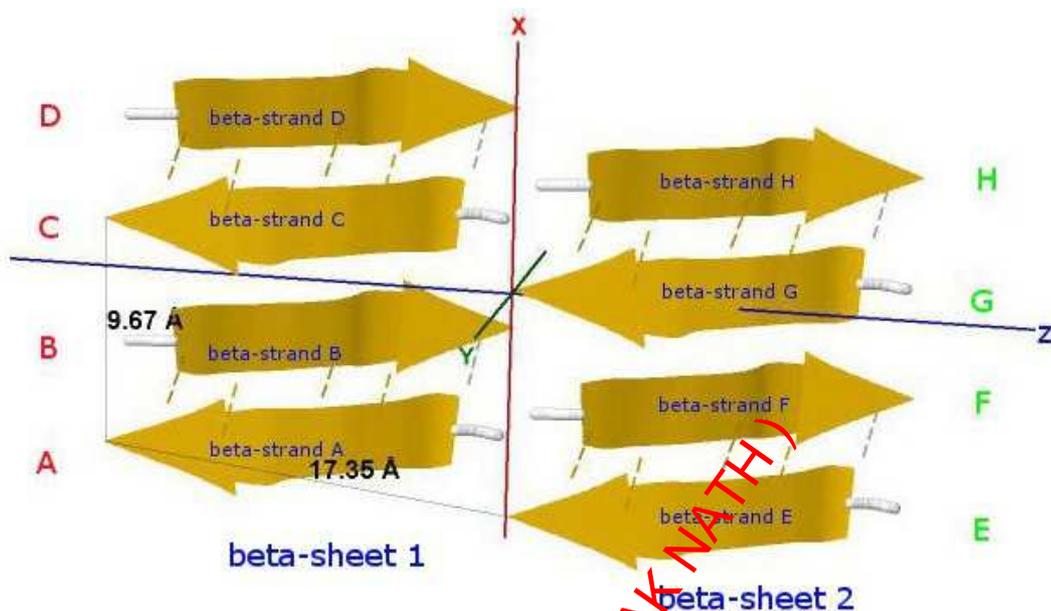
Figure 4: *Protein fibril structure of human insulin LYQLEN (13-18) (PDB id: 2OMP). The dashed lines denote the HBs between the pairs of β-strands. A, B, C, D, E, F, G, H denote the chains of the fibril. The pair of β-sheets 1 & 2 forms a completely dry interface by vdWs, and between many pairs of β-sheets wet interfaces are formed with water molecules.*

up (or down) along the X-axis by Eqs. 3.2∼3.3:

$$E(F) = A(B) + (-1.885, 0, 17.243), \tag{3.1}$$

$$C(D) = A(B) + (9.666, 0, 0), \tag{3.2}$$

$$G(H) = E(F) + (9.666, 0, 0). \tag{3.3}$$

3.2   New Molecular Modeling Homology Model

Basing on the template 2OMP.pdb from the Protein Data Bank (Figure 4), the AGAAAAGA palindrome amyloid fibril model of prions (denoted as Model 1) will be constructed. Chains AB of Model 1 will be got from AB Chains of 2OMP.pdb using the mutate module of the free package Swiss-PdbViewer (SPDBV Version 4.01) (http://spdbv.vital-it.ch). It is pleasant to see that some HBs are still kept after the mutations; thus we just need to consider the vdWs only. Making mutations for EF Chains of 2OMP.pdb, we can get the EF Chains of Model 1. Then we add GLY and ALA residues by XLEaP module of Amber 11. However, the vdWs between Chain A and Chain E, between B Chain and F Chain are too

11

far at this moment (Figure 5, where the twice of the vdW radius of CB atom is 3.4 angstroms).

In [79] the commercial package InsightII (http://accelrys.com/) is used to build models. Instead of InsightII, because this package is not available by the authors, this paper uses Algorithm 1 to build and optimize Model 1. In "Zipper 1", fixing the coordinates of A.ALA3.CB, A.ALA1.CB and letting the coordinates of E.ALA6.CB, E.ALA4.CB be variables, we may get an optimization problem:

$$
\begin{aligned}
\min 4 & \left\{ \frac{1}{[(x_{11}-1.071)^2+(x_{12}-2.986)^2+(x_{13}-1.888)^2]^6} - \frac{1}{[(x_{11}-1.071)^2+(x_{12}-2.986)^2+(x_{13}-1.888)^2]^3} \right\} \\
+4 & \left\{ \frac{1}{[(x_{21}-1.071)^2+(x_{22}-2.986)^2+(x_{23}-1.888)^2]^6} - \frac{1}{[(x_{21}-1.071)^2+(x_{22}-2.986)^2+(x_{23}-1.888)^2]^3} \right\} \\
+4 & \left\{ \frac{1}{[(x_{21}-1.135)^2+(x_{22}+0.763)^2+(x_{23}-7.209)^2]^6} - \frac{1}{[(x_{21}-1.135)^2+(x_{22}+0.763)^2+(x_{23}-7.209)^2]^3} \right\}
\end{aligned}
\tag{3.4}
$$

or

$$
\begin{aligned}
\min \tfrac{1}{2} & \left\{ (x_{11}-1.071)^2 + (x_{12}-2.986)^2 + (x_{13}-1.888)^2 - 3.4^2 \right\}^2 \\
+\tfrac{1}{2} & \left\{ (x_{21}-1.071)^2 + (x_{22}-2.986)^2 + (x_{23}-1.888)^2 - 3.4^2 \right\}^2 \\
+\tfrac{1}{2} & \left\{ (x_{21}-1.135)^2 + (x_{22}+0.763)^2 + (x_{23}-7.209)^2 - 3.4^2 \right\}^2 \\
& -0.05 \left\{ x_{11} + x_{12} + x_{13} + x_{21} + x_{22} + x_{23} \right\}
\end{aligned}
\tag{3.5}
$$

with an initial solution (-0.067, 5.274, 7.860; -1.119, 1.311, 13.564). Similarly, in "Zipper 2", fixing the coordinates of B.ALA4.CB, B.ALA6.CB and letting the coordinates of F.ALA1.CB, F.ALA3.CB be variables, we get another optimization problem:

$$
\begin{aligned}
\min 4 & \left\{ \frac{1}{[(x_{11}-5.446)^2+(x_{12}-2.796)^2+(x_{13}-2.662)^2]^6} - \frac{1}{[(x_{11}-5.446)^2+(x_{12}-2.796)^2+(x_{13}-2.662)^2]^3} \right\} \\
+4 & \left\{ \frac{1}{[(x_{21}-5.446)^2+(x_{22}-2.796)^2+(x_{23}-2.662)^2]^6} - \frac{1}{[(x_{21}-5.446)^2+(x_{22}-2.796)^2+(x_{23}-2.662)^2]^3} \right\} \\
+4 & \left\{ \frac{1}{[(x_{21}-5.201)^2+(x_{22}+1.125)^2+(x_{23}-7.873)^2]^6} - \frac{1}{[(x_{21}-5.201)^2+(x_{22}+1.125)^2+(x_{23}-7.873)^2]^3} \right\}
\end{aligned}
\tag{3.6}
$$

or

$$
\begin{aligned}
\min \tfrac{1}{2} & \left\{ (x_{11}-5.446)^2 + (x_{12}-2.796)^2 + (x_{13}-2.662)^2 - 3.4^2 \right\}^2 \\
+\tfrac{1}{2} & \left\{ (x_{21}-5.446)^2 + (x_{22}-2.796)^2 + (x_{23}-2.662)^2 - 3.4^2 \right\}^2 \\
+\tfrac{1}{2} & \left\{ (x_{21}-5.201)^2 + (x_{22}+1.125)^2 + (x_{23}-7.873)^2 - 3.4^2 \right\}^2 \\
& -0.05 \left\{ x_{11} + x_{12} + x_{13} + x_{21} + x_{22} + x_{23} \right\}
\end{aligned}
\tag{3.7}
$$

with an initial solution (4.714, 4.878, 8.881; 4.170, 1.360, 14.292). Next, we solve Eqs. 3.5 and 3.7 by Algorithm 1.

We first solve Eq. 3.5 in the use of Algorithm 1. We set $\sigma_1 = 10^{-4}$, $\sigma_2 = 0.1 \sim 0.9$, $m_1 = 3 \sim 7$, take the initial solution $x_0 =$(-0.067, 5.274, 7.860; -1.119, 1.311, 13.564) and calculate its gradient g_0=(-69.7747, 140.135, 365.852,
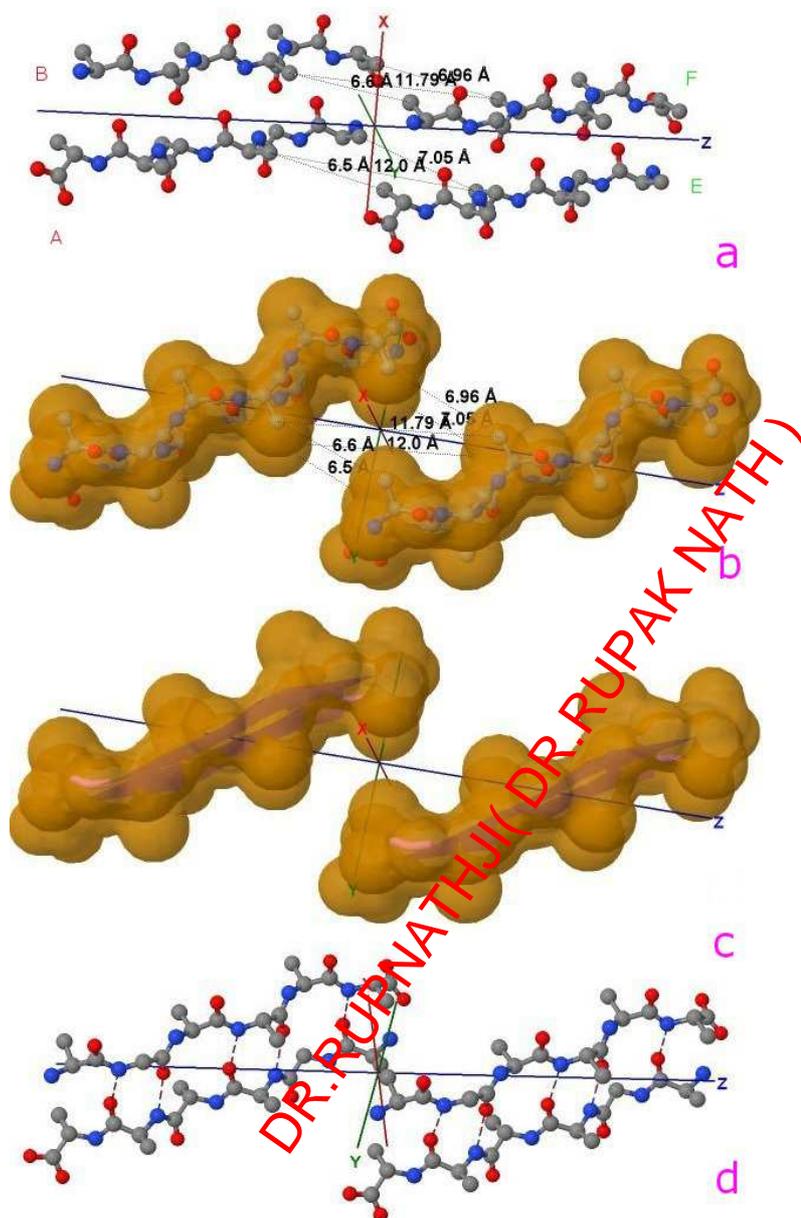
Figure 5: *5a shows the distances of "Zipper 1"-E.ALA6.CB-A.ALA3.CB-E.ALA4.CB-A.ALA1.CB are 6.5, 12.0 and 7.05 angstroms respectively, and the distances of "Zipper 2"-F.ALA1.CB-B.ALA4.CB-F.ALA3.CB-B.ALA6.CB are 6.6, 11.79, 6.96 angstroms respectively. 5b shows the far vdW surface. 5c shows the violet colored ABEF Chains of Figure 3. 5d shows HBs: A/E.ALA5.O-B/F.GLY2.N, A/E.ALA5.N-B/F.GLY2.O, A/E.ALA3.O-B/F.ALA4.N, A/E.ALA3.N-B/F.ALA4.O, A/E.ALA1.O-B/F.ALA6.N.*

13

-752.075, -285.005, 3576.69) and its Hessian matrix $H_0 =$

$$
\begin{pmatrix}
66.4497 & -10.415 & -27.1845 & 0 & 0 & 0 \\
-10.415 & 82.2093 & 54.6557 & 0 & 0 & 0 \\
-27.1845 & 54.6557 & 203.929 & 0 & 0 & 0 \\
0 & 0 & 0 & 380.664 & -4.02618 & -159.578 \\
0 & 0 & 0 & -4.02618 & 369.586 & -25.5081 \\
0 & 0 & 0 & -159.578 & -25.5081 & 1048.02
\end{pmatrix}, \quad (3.8)
$$

which is a positive definite matrix with eigenvalues (1085.02, 372.093, 341.157, 230.049, 61.2695, 61.2695). Then Algorithm 1 hybridized with simulated annealing global optimal search (in order to bring local optimal solutions to jump out of local traps, replacing the discrete gradient local optimal search method in Algorithm 1 of [14] by the Algorithm 1 of this paper) is executed and the optimal solution (3.027, 4.954, 3.856; 1.679, 1.777, 5.011) for Eq. 3.5 is got.

Similarly, for Eq. 3.7, we take the initial solution $x_0 =$(4.714, 4.878, 8.881; 4.170, 1.360, 14.292) and calculate its gradient g_0=(-46.8782, 133.142, 397.798, -401.192, -182.604, 3436.46) and its Hessian matrix $H_0 =$

$$
\begin{pmatrix}
66.1163 & -6.0961 & -18.2092 & 0 & 0 & 0 \\
-6.0961 & 81.3119 & 51.7918 & 0 & 0 & 0 \\
-18.2092 & 51.7918 & 218.677 & 0 & 0 & 0 \\
0 & 0 & 0 & 339.302 & -2.9188 & -85.8315 \\
0 & 0 & 0 & -2.9188 & 361.487 & -2.99786 \\
0 & 0 & 0 & -85.8315 & -2.99786 & 1034.38
\end{pmatrix}, \quad (3.9)
$$

which is a positive definite matrix with eigenvalues (1044.83, 361.8, 328.538, 238.159, 63.973, 63.973). The optimal solution (7.412, 4.760, 4.624; 5.887, 1.451, 5.757) for Eq. 3.7 is got.

We set (3.027, 4.954, 3.856; 1.679, 1.777, 5.011) as the coordinates of E.ALA6.CB and E.ALA4.CB, (7.412, 4.760, 4.624; 5.887, 1.451, 5.757) as the coordinates of F.ALA1.CB and F.ALA3.CB, and taking the average value we get

$$
E(F) = A(B) + (1.0335, 1.0823, 0.9723). \quad (3.10)
$$

By Eq. 3.10 we can get very close vdW contacts between A (B) chains and E(F) chains (Figure 5(b)), and other chains of Model 1 can be got by Eqs. 3.2~3.3 and Eq. 3.11:

$$
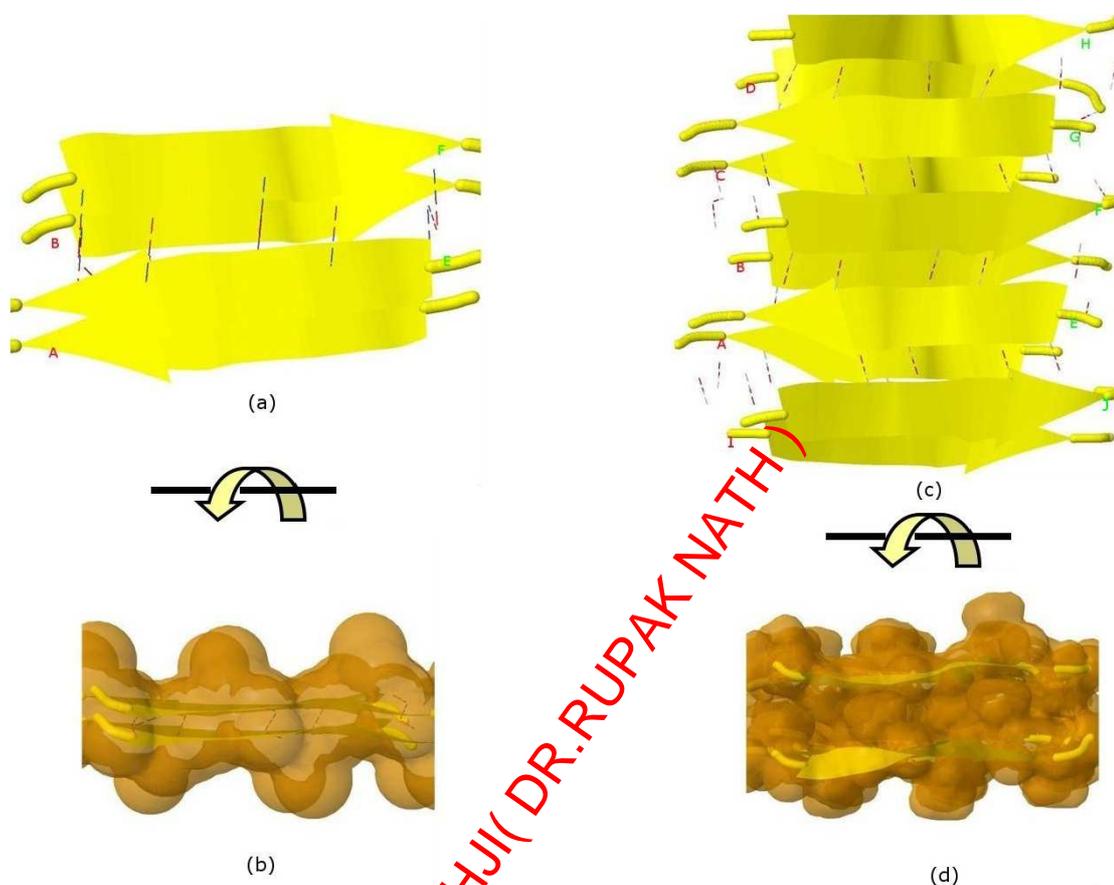I(J) = A(E) - (9.666, 0, 0). \quad (3.11)
$$

Figure 6: *Left: The close vdW surface contacts between Chains A(B) and E(F) of Model 1 after solving Eqs. 3.5 and 3.7; Right: The constructed Model 1 with 10 chains. The dashed lines denote HBs.*

The initial structure of Model 1 illuminated in Figure 6(a)∼6(b) is not the optimal structure with the lowest total potential energy. The initial structure also has no hydrogen atoms (so no hydrogen bonds existed) and water molecules added. For each Chain, the C-terminal and N-terminal atoms also have problems. Clearly there are a lot of close/bad contacts between β-strand atoms as illuminated in Figure 6(a)∼6(b). We used the ff03 force field of AMBER 11,in a neutral pH environment. The amyloid fibrils were surrounded with a 8 angstroms layer of TIP3PBOX water molecules using the XLEaP module of AMBER 11. 1944 waters and 408 hydrogen atoms were added for Model 1 byt he XLEaP module. solvated amyloid fibril was inimized by the method.

The LJ potential energy of atoms' vdW interactions is just a part of the total potential energy of a protein, and by observations from Model 1 computed by Eqs. 3.10, 3.2∼3.3 and 3.11 we can see that the contacts between β-strand

atoms and β-sheet atoms are too close/bad. Thus, we need to relax Model 1 computed. The relaxation is done in the use of local search LBFGS Quasi-Newton method (lbfgs_memory_depth=3) within AMBER 11 [23]. The relaxed/optimized Model 1 is illuminated in Figure 6(c). Seeing Figure 6(d) compared with Figure 6(b), we may know the vdW interactions between the two β-sheets are very relaxed/optimized now. Figure 6(c) shows the Model 1 of optimal molecular structure for prion AGAAAAGA amyloid fibrils.

4   Conclusion

In a (macro)molecular system, a robust energy minimization is very necessarily required. Mathematical optimization minimization methods find a place to apply in these systems. Because in physics the (macro)molecular system usually is not a simple two-body problem of system, local search optimization methods are very useful in the applications to the (macro)molecular system. On anther sense, when a protein is unstable, noncrystalline or insoluble and very difficult to detect its 3D structure by the expensive and costly NMR and X-ray, theoretical mathematical or physical computational method can be used to produce the 3D structure of the protein. Moreover, even the X-ray crystallography finds the X-ray final structure of a protein, we still need refinements using theoretical protocols in order to produce a better structure. The theoretical computational method - an improved LBFGS Quasi-Newtonian mathematical optimization method - presented in this paper and other mathematical optimization methods mentioned in this paper should be very useful in the protein molecular modeling research field.

This paper also shows the effectiveness of the improved LBFGS mathematical optimization method presented. Prion AGAAAAGA amyloid fibrils have not much structural information. This paper presents some bioinformatics on the molecular structures of prion AGAAAAGA amyloid fibrils in the sense of theoretical emphasis. The structures may be helpful in the advance in the biochemical knowledge of prion protein misfolding or instability and in the future applications for therapeutic agent design.

Acknowledgments

References

[1] Chiang PK, Lam MA, Luo Y (2008) The many faces of amyloid beta in Alzheimer's disease. Curr Mol Med 8(6): 580–584.

[2] Irvine GB, El-Agnaf OM, Shankar GM, Walsh DM (2008) Protein aggregation in the brain: the molecular basis for Alzheimer's and Parkinson's diseases. Mol med Cambridge Mass 14(7-8): 451-64.

[3] Ferreira ST, Vieira MN, De Felice FG (2007) Soluble protein oligomers as emerging toxins in Alzheimer's and other amyloid diseases. IUBMB life 59(4–5): 332-45.

[4] Nature Editorial (2001) More than just mad cow disease. Nature Struct Biol 8(4): 281. doi:10.1038/86132

[5] Truant R, Atwal RS, Desmond C, Munsie L, Tran T (2008) Huntington's disease: revisiting the aggregation hypothesis in polyglutamine neurodegenerative diseases. The FEBS J 275(17): 4252-4262.

[6] Weydt P, La Spada AR (2006) Targeting protein aggregation in neurodegeneration–lessons from polyglutamine disorders. Expert opinion on therapeutic targets 10(4): 505-513.

[7] Nelson R, Sawaya MR, Balbirnie M, Madsen A, Riekel C, Grothe R, Eisenberg D (2005) Structure of the cross-beta spine of amyloid-like fibrils. Nature 435(7043): 773–778.

[8] Sawaya MR, Sambashivan S, Nelson R, Ivanova MI, Sievers SA, Apostol MI, Thompson MJ, Balbirnie M, Wiltzius JJ, McFarlane HT, Madsen A, Riekel C, Eisenberg D (2007) Atomic structures of amyloid cross-beta spines reveal varied steric zippers. Nature 447(7143): 453–457.

[9] Zhang JP (2011) The Lennard-Jones potential minimization problem for prion AGAAAAGA amyloid fibril molecular modeling. arXiv:1106.1584.

[10] Kolossvry I, Bowers KJ (2010) Global optimization of additive potential energy functions: predicting binary Lennard-Jones clusters. Phys Rev E Stat Nonlin Soft Matter Phys 82(5 Pt 2): 056711.

[11] Sicher M, Mohr S, Goedecker S (2010) Efficient moves for global geometry optimization methods and their application to binary systems. J Chem Phys 134(4): 44–106.

[12] Strodel B, Lee JW, Whittleston CS, Wales DJ (2010) Transmembrane structures for Alzheimer's $A\beta$(1-42) oligomers. J Am Chem Soc 132(38): 13300–13312.

[13] Ye T, Xu R, Huang W (2011) Global optimization of binary Lennard-Jones clusters using three perturbation operators. J Chem Inf Model 51(3): 572–577.

[14] Zhang JP, Sun J, Wu CZ (2011) Optimal atomic-resolution structures of prion AGAAAAGA amyloid fibrils. J Theor Biol 279(1): 17–28.

[15] Zhang JP, Gao DY, Yearwood J. (2011) A novel canonical dual computational approach for prion AGAAAAGA amyloid fibril molecular modeling. J Theor Biol 284(1): 149–157.

[16] Grosso A, Locatelli M, Schoen F (2009) Solving molecular distance geometry problems by global optimization algorithms. Comput Optim Appl 43: 23–37.

[17] More JJ, Wu ZJ (1997) Global continuation for distance geometry problems. SIAM J Optim 7: 814-836.

[18] Zou ZH, Bird RH, Schnabel RB (1997) A stochastic/perturbation global optimization algorithm for distance geometry problems. J Glob Optim 11: 91-105.

[19] Brown DR (2000) Prion protein peptides: optimal toxicity and peptide blockade of toxicity. Mol Cell Neurosci 15: 66–78.

[20] van der Spoel D, Lindahl E, Hess B, van Buuren AR, Apol E, Meulenhoff PJ, Tieleman DP, Sijbers ALTM, Feenstra KA, van Drunen R, Berendsen HJC (2010) Gromacs User Manual version 4.5.4, www.gromacs.org.

[21] Snyman JA (2005) Practical Mathematical Optimization: An Introduction to Basic Optimization Theory and Classical and New Gradient-Based Algorithms. Springer Publishing, ISBN 0-387-24348-8.

[22] Chu, J, Trout BL, Brooks BR (2003) A super-linear minimization scheme for the nudged elastic band method. J Chem Phys 119: 12708-12717.

[23] Case DA, Darden TA, Cheatham TE, Simmerling III CL, Wang J, Duke RE, Luo R, Walker RC, Zhang W, Merz KM, Roberts BP, Wang B, Hayik S, Roitberg A, Seabra G, Kolossvary I, Wong KF, Paesani F, Vanicek J, Liu J, Wu X, Brozell SR, Steinbrecher T, Gohlke H, Cai Q, Ye X, Wang J, Hsieh M-J, Cui G, Roe DR, Mathews DH, Seetin MG, Sagui C, Babin V, Luchko T, Gusarov S, Kovalenko A, Kollman PA (2010) *AMBER 11*, University of California, San Francisco.

[24] Liu DC, Nocedal J (1989) On the limited memory method for large scale optimization. Math Programming B 45: 503-528.

[25] Nocedal J, Morales J (2000) Automatic preconditioning by limited memory quasi-Newton updating. SIAM J Opt 10: 1079–1096.

[26] Byrd RH, Nocedal J, Zhu C (1995) Towards a discrete Newton method with memory for large-scale optimization, Optimization Technology Center Report OTC-95-1, EEC Department, North-western University.

[27] Byrd RH, Lu P, Nocedal J (1995) A limited memory algorithm for bound constrained optimization. SIAM J Scientif Statistic Comput 16: 1190–1208.

[28] Zhu C, Byrd RH, Nocedal J (1997) L-BFGS-B: Algorithm 778: L-BFGS-B, FORTRAN routines for large scale bound constrained optimization. ACM Trans Math Softw 23: 550–560.

[29] Dennis JE, Robert JR, Schnabel B (1996) Numerical Methods for Unconstrained Optimization and Nonlinear Equations, SIAM.

[30] Nocedal J, Wright SJ (1999) Numerical Optimization, Springer-Verlag, ISBN 0-387-98793-2 & Numerical Optimization (2nd ed.), Berlin, New York: Springer-Verlag, ISBN 978-0-387-30303-1

[31] Berndt E, Hall B, Hall R, Hausman J (1974) Estimation and Inference in Nonlinear Structural Models. Annals of Economic and Social Measurement 3: 653-665.

[32] Luenberger DG, Ye YY (2008) Linear and nonlinear programming (3rd edition), International Series in Operations Research & Management Science 116, New York: Springer, ISBN 978-0-387-74502-2.

[33] Wang YJ, Xiu NH (2012) Nonlinear Optimization Theory and Methods, Science Press, Beijing, ISBN 9787030330284.

[34] Powell MJD (1984) Nonconvex minimization calculations and the conjugate gradient method. In: Griffiths D.F. ed., Numerical Analysis, Lecture Notes in Mathematics 1066, Springer Verlag, Berlin: 122-141.

[35] Mascarenhas WF (2004) The BFGS method with exact line searches fails for non-convex objective functions. Math Program Ser A 99: 49-61.

[36] Dai YH (2002) Convergence properties of the BFGS Algorithm. SIAM J Optim 13(3): 693-701.

[37] Li DH, Fukushima M (2001) A modified BFGS method and its global convergence in nonconvex minimization. J Comput Math Appl 129: 15–35.

[38] Xiao YH, Wei ZX, Wang ZG (2008) A limited memory BFGS-type method for large-scale unconstrained optimization. J Comput Math Appl 56(4): 1001-1009.

[39] Yang YT, Xu CX (2007) A compact limited memory method for large scale unconstrained optimization. Eur J Oper Res 180(1): 48-56.

[40] Hou YT, Wang YJ (2012) The modified limited memory BFGS method for large-scale optimization, MSc degree thesis, Qufu Normal University, China.

[41] Grippo L, Lampariello F, Lucidi S (1986) A nonmonotone line search technique for Newton's method. SIAM J Numer Anal 23: 707–716.

[42] Han JY, Liu GH (1997) Global convergence analysis of a new nonmonotone BFGS algorithm on convex objective Functions. Compu Optim Appl 7: 277–289.

[43] Bagirov AM, Karaszen B, Sezer M (2008) Discrete gradient method: derivative-free method for nonsmooth optimization. J Optim Theor Appl 137(2): 317–334.

[44] Dolan ED, Mor JJ (2002) Benchmarking optimization software with performance profiles. Math Program Ser A 91: 201-213.

[45] Byrd RH, Nocedal J (1989) A tool for the analysis of quasi-Newton methods with application to unconstrained minimization. SIAM J Numer Anal 26: 727–739.

[46] Byrd RH, Nocedal J, Yuan YX (1987) Global convergence of a class of quasi-Newton methods on convex problems. SIAM J Numer Anal 24: 1171–1189.

[47] Dennis JE, More JJ (1977) Quasi-Newton method, motivation and theory. SIAM Rev 19: 46–89.

[48] Nocedal J (1980) Updating quasi-Newton matrices with limited storage. Math Comput 35: 773–782.

[49] Gilbert JC, Lemarichal C (1989) Some numerical experiments with variable storage quasi-Newton algorithms. Math Programming 45: 407–435.

[50] Al-Baali M (1999) Improved Hessian approximations for the limited memory BFGS method. Numer. Algorithms 22: 99–112.

[51] Byrd RH, Nocedal J, Schnabel RB (1994) Representations of quasi-Newton matrices and their use in limited memory methods. Math Programming 63: 129–156.

[52] Yuan GL, Wei ZX, Wu YL (2010) Modified limited memory BFGS method with nonmonotone line search for unconstrained optimization. J Korean Math Soc 47: 767–788.

[53] Alper T, Cramp W, Haig D, Clarke M (1967) Does the agent of scrapie replicate without nucleic acid?. Nature 214(5090): 764–766.

[54] Griffith J (1967) Self-replication and scrapie. Nature 215(5105): 1043-1044.

[55] Brown DR (2001) Microglia and prion disease. Microsc Res Tech 54: 71–80.

[56] Brown DR, Herms J, Kretzschmar HA (1994) Mouse cortical cells lacking cellular PrP survive in culture with a neurotoxic PrP fragment. Neuroreport 5: 2057–2060.

[57] Cappai R, Collins SJ (2004) Structural biology of prions, in Prions A Challenge for Science, Medicine and the Public Health System (Rabenau H.F., Cinatl J., Doerr H.W. (eds)), Basel, Karger 11: 14–32.

[58] Chabry J, Caughey B, Chesebro B (1998) Specific inhibition of in vitro formation of protease-resistant prion protein by synthetic peptides. J Biol Chem 273(21): 13203–13207.

[59] Cheng HM, Tsai TWT, Huang WYC, Lee HK, Lian HY, Chou FC, Mou Y, Chu J, Chan JC (2011) Steric zipper formed by hydrophobic peptide fragment of Syrian hamster prion protein. Biochem 50(32): 6815–6823.

[60] Gasset M, Baldwin MA, Lloyd DH, Gabriel JM, Holtzman DM, Cohen F, Fletterick R, Prusiner SB (1992) Predicted alpha-helical regions of the prion protein when synthesized as peptides form amyloid. Proc Natl Acad Sci USA 89(22): 10940–10944.

[61] Haigh CL, Edwards K, Brown DR (2005) Copper binding is the governing determinant of prion protein turnover. Mol Cell Neurosci 30(2): 186–196.

[62] Harrison CF, Lawson VA, Coleman BM, Kim YS, Masters CL, Cappai R, Barnham KJ, Hill AF (2010) Conservation of a glycine-rich region in the prion protein is required for uptake of prion infectivity. J Biol Chem 285(26): 20213–20223.

[63] Holscher C, Delius H, Burkle A (1998) Overexpression of nonconvertible PrP^C delta114-121 in scrapie-infected mouse neuroblastoma cells leads to trans-dominant inhibition of wild-type PrP^{Sc} accumulation. J Virol 72: 1153–1159.

[64] Jobling MF, Huang X, Stewart LR, Barnham KJ, Curtain C, Volitakis I, Perugini M, White AR, Cherny RA, Masters CL, Barrow CJ, Collins SJ, Bush AI, Cappai R (2001) Copper and Zinc binding modulates the aggregation and neurotoxic properties of the prion peptide PrP 106–126. Biochem 40: 8073–8084.

[65] Jobling MF, Stewart LR, White AR, McLean C, Friedhuber A, Maher F, Beyreuther K, Masters CL, Barrow CJ, Collins SJ, Cappai R (1999) The hydrophobic core sequence modulates the neurotoxic and secondary structure properties of the prion peptide 106-126. J Neurochem 73: 1557–1565.

[66] Jones EM, Wu B, Surewicz K, Nadaud PS, Helmus JJ, Chen S, Jaroniec CP, Surewicz WK (2011) Structural polymorphism in amyloids: new insights from studies with Y145Stop prion protein fibrils. J Biol Chem 286(49): 42777–42784.

[67] Kourie JI (2001) Mechanisms of prion-induced modifications in membrane transport properties: implications for signal transduction and neurotoxicity. Chem Biol Interact 138(1): 1–26.

[68] Kourie JI, Kenna BL, Tew D, Jobling MF, Curtain CC, Masters CL, Barnham KJ, Cappai R (2003) Copper modulation of ion channels of PrP[106-126] mutant prion peptide fragments. J Membr Biol 193(1): 35–45.

[69] Kuwata K, Matumoto T, Cheng H, Nagayama K, James TL, Roder H (2003) NMR-detected hydrogen exchange and molecular dynamics simulations provide structural insight into fibril formation of prion protein fragment 106-126. Proc Natl Acad Sci USA 100(25): 14790–14795.

[70] Laganowsky A, Liu C, Sawaya MR, Whitelegge JP, Park J, Zhao M, Pensalfini A, Soriaga AB, Landau M, Teng PK, Cascio D, Glabe C, Eisenberg D (2012) Atomic view of a toxic amyloid small oligomer. Science 335: 1228–1231.

[71] Lee SW, Mou Y, Lin SY, Chou FC, Tseng WH, Chen CH, Lu CY, Yu SS, Chan JC (2008) Steric zipper of the amyloid fibrils formed by residues 109-122 of the Syrian hamster prion protein. J Mol Biol 378(5): 1142–1154.

[72] Ma BY, Nussinov R (2002) Molecular dynamics simulations of alanine rich β-sheet oligomers: insight into amyloid formation. Protein Sci 11: 2335–2350.

[73] Norstrom EM, Mastrianni JA (2005) The AGAAAAGA palindrome in PrP is required to generate a productive PrP^{Sc}-PrP^{C} complex that leads to prion propagation. J Biol Chem 280: 27236–27243.

[74] Sasaki K, Gaikwad J, Hashiguchi S, Kubota T, Sugimura K, Kremer W, Kalbitzer HR, Akasaka K (2008) Reversible monomer-oligomer transition in human prion protein. Prion 2(3): 118–122.

[75] Wagoner VA (2010) Computer simulation studies of self-assembly of fibril forming peptides with an Intermediate resolution protein model. PhD thesis, North Carolina State University, Raleigh, North Carolina.

[76] Wagoner VA, Cheon M, Chang I, Hall CK (2011) Computer simulation study of amyloid fibril formation by palindromic sequences in prion peptides. Proteins Struct Funct Bioinf 79(7): 2132–2145.

[77] Wegner C, Romer A, Schmalzbauer R, Lorenz H, Windl O, Kretzschmar HA (2002) Mutant prion protein acquires resistance to protease in mouse neuroblastoma cells. J Gen Virol 83: 1237–1245.

[78] Zanuy D, Ma B, Nussinov R (2003) Short peptide amyloid organization: stabilities and conformations of the islet amyloid peptide NFGAIL. Biophys J 84(3): 1884–1894.

[79] Zhang JP (2011) Optimal molecular structures of prion AGAAAAGA palindrome amyloid fibrils formatted by simulated annealing. J Mol Model 17(1): 173–179.

[80] Zhang ZQ, Chen H, Lai LH (2007) Identification of amyloid fibril-forming segments based on structure and residue-based statistical potential. Bioinf 23: 2218–2225.

Supplementary Materials [40]

Mathematical Assumptions (A1)-(A4):

(A1) the objective function $f : \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable, i.e. $f \in C^2$;

(A2) f has Lipschitz continuous gradients and Hessians, i.e., there exist constants $L, M_2 > 0$ such that

$$\|g(x) - g(y)\| \le L\|x - y\|, \quad \forall x, y \in \mathbb{R}^n,$$

$$\|G(x) - G(x^*)\| \le M_2\|x - x^*\|,$$

where G denotes the Hessian and x^* is a stationary point (i.e. x_k converges to x^* where $g(x^*) = 0$);

(A3) $G(x^*)$ is positive-definite; and

(A4) the level set of f

$$\Omega = \{x \in \mathbb{R}^n \mid f(x) \le f(x_0)\},$$

is bounded, where $x_0 \in \mathbb{R}^n$ is an initial point.

Mathematical Proof of the Algorithm Convergence and Convergent Rate:

In order to establish the convergence for Algorithm 1, we represent Eq. 2.3 as

$$B_k^{l+1} = B_k^l - \frac{B_k^l s_l s_l^T B_k^l}{s_l^T B_k^l s_l} + \frac{y_l^* y_l^{*T}}{y_k^{*T} s_l}, \quad l = k - \bar{m} + 1, \cdots, k, \qquad (5.1)$$

where $B_k = H_k^{-1}$, $s_l = x_{l+1} - x_l$, $y_l^* = y_l + \gamma_l s_l$, and $B_k^{k-\bar{m}+1} = B_0$ for all k.

Theorem 1   *Let Assumptions (A1)~(A4) hold, and $\{x_k\}$ denotes the sequence generated by Algorithm 1. Then $\displaystyle\liminf_{k\to\infty} \|g_k\| = 0$; moreover, there exits a constant $t \in [0, 1)$ such that*

$$f(x_k) - f(x^*) \leq t^k[f(x_0 - f(x^*)],$$

which is to say that x_k converges to the minimum R-linearly.

Proof.   Denote by M_1 an upper bound of $\|g_k\|$ on Ω and by $y_k^* = y_k + \gamma_k s_k$, we can prove that

$$\frac{\|y_k^{*T}\|^2}{y_k^{*T} s_k} \leq \frac{(L + M_1)^2}{\gamma},$$

where γ is a sufficiently small positive number. By this formula and taking the trace operation in both sides of Eq. 5.1, we get

$$
\begin{aligned}
tr(B_{k+1}) &= tr(B_k^{k-\bar{m}+1}) - \sum_{l=k-\bar{m}+1}^{k} \frac{\|B_l s_l\|^2}{s_l^T B_l s_l} + \sum_{l=k-\bar{m}+1}^{k} \frac{\|y_l^*\|^2}{y_l^{*T} s_l} \\
&= \cdots \\
&\leq tr(B_0) - \sum_{l=0}^{k} \frac{\|B_l s_l\|^2}{s_l^T B_l s_l} + \sum_{l=0}^{k} \frac{\|y_l^*\|^2}{y_l^{*T} s_l} \\
&\leq tr(B_0) - \sum_{l=0}^{k} \frac{\|B_l s_l\|^2}{s_l^T B_l s_l} + \frac{(L+M_1)^2}{\gamma}(k+1) \\
&\leq M_3,
\end{aligned}
\tag{5.2}
$$

where M_3 is a positive constant. We also take the determinant and by Eq. 5.2, we have

$$
\begin{aligned}
det(B_{k+1}) &= det(B_k^{k-\bar{m}+1}) \prod_{l=k-\bar{m}+1}^{k} \frac{y_l^{*T}}{s_l^T B_l s_l} \\
&= \cdots = det(B_0) \prod_{l=0}^{k} \frac{y_l^{*T}}{s_l^T B_l s_l} \geq det(B_0)(\frac{\gamma}{M_3})^{k+1} \\
&\geq M_4,
\end{aligned}
\tag{5.3}
$$

where M_4 is a positive constant. Combining Eqs. 5.2 and 5.3, we obtain a constant $\delta > 0$ such that $\cos(\theta_k) \geq \delta$, where $\cos(\theta_k) = \frac{s_k^T B_k s_k}{\|s_k\|\|B_k s_k\|}$.

Because $g(x)$ is Lipschitz continuous and by the Wolfe-type line search condition Eq. 2.2, we have

$$(\sigma_2 - 1)g_k^T d_k \leq (g_{k+1} - g_k)^T d_k \leq \lambda_k L \|d_k\|^2,$$

which implies

$$\lambda_k \geq \frac{(\sigma_2 - 1)}{L} \frac{g_k^T d_k}{\|d_k\|^2}. \tag{5.4}$$

Using the Wolfe-type line search condition Eq. 2.1 and Eq. 5.4, we have

$$
\begin{aligned}
f_{k+1} &\leq f_k + \sigma_1 \lambda_k g_k^T d_k \leq f_k + \sigma_1 \frac{(\sigma_2 - 1)}{L} \frac{(g_k^T d_k)^2}{\|d_k\|^2} \\
&= f_k + \sigma_1 \frac{(\sigma_2 - 1)}{L} \|g_k\|^2 \cos^2 \theta_k.
\end{aligned}
$$

Then using the monotonicity and the bound of $\{f_k\}$ on Ω, we obtain

$$\sum_{k=1}^{\infty} \|g_k\|^2 \cos^2 \theta_k < \infty.$$

Combining $\cos \theta_k \geq \delta$, we get

$$\lim_{k \to \infty} \inf \|g_k\| = 0$$

Under the Assumptions (A2)~(A3), there is a neighbourhood $U(x^*)$ of x^* such that for all $x \in U(x^*)$,

$$\|g(x)\| \geq \|g(x) - g(x^*)\| \geq m\|x - x^*\|, \tag{5.5}$$

and for all $d \in R^n$,

$$d^T G(x) d \geq m\|d\|^2, \tag{5.6}$$

where m is a positive constant got according to the Assumptions. Hence, Eqs. 5.5~5.6 hold with $x = x_k$ for all k sufficiently large.

Since

$$
\begin{aligned}
y_k^{*T} s_k &= (y_k + \gamma_k s_k)^T s_k \geq y_k^T s_k \\
&\geq -(1 - \sigma_2) g_k^T s_k \\
&= (1 - \sigma_2)\|g_k\|\|s_k\| \cos \theta_k,
\end{aligned}
$$

and $y_k^{*T} s_k \leq (L + M_1)^2 \|s_k\|^2$, if denoting $\alpha_1 = \frac{1 - \sigma_1}{(L + M_1)^2}$, we can obtain

$$\|s_k\| \geq \frac{1 - \sigma_1}{(L + M_1)^2} \|g_k\| \cos \theta_k = \alpha_1 \|g_k\| \cos \theta_k. \tag{5.7}$$

Combining Eq. 5.7 and the Wolfe-type line search condition Eq. 2.1, we get

$$
\begin{aligned}
f_{k+1} - f^* &\leq f_k - f^* + \sigma_1 g_k^T s_k \\
&= f_k - f^* - \sigma_1 \|g_k\|\|s_k\| \cos \theta_k \\
&\leq f_k - f^* - \sigma_1 \alpha_1 \|g_k\|^2 \cos^2 \theta_k.
\end{aligned}
$$

26

Using Eq. 5.5, then Eq. 5.7 implies for all k sufficiently large,

$$
\begin{aligned}
f_{k+1} - f^* &\leq f_k - f^* - \sigma_1 \alpha_1 \|g_k\|^2 \cos^2 \theta_k \\
&\leq f_k - f^* - \sigma_1 \alpha_1 m^2 \|x_k - x^*\|^2 \cos^2 \theta_k.
\end{aligned}
\tag{5.8}
$$

By Taylor's expansion, then there exists a positive number $M_5 > 0$ such that

$$
f_k - f^* \leq M_5 \|x_k - x^*\|^2
$$

for all k sufficiently large.

Let $\alpha_2 = \frac{\sigma_1 \alpha_1 m^2}{M_5}$, indeed Eq. 5.8 implies that

$$
\begin{aligned}
f_{k+1} - f^* &\leq f_k - f^* - \sigma_1 \alpha_1 m^2 \|x_k - x^*\|^2 \cos^2 \theta_k \\
&\leq f_k - f^* - \alpha_2 (f_k - f^*) \cos^2 \theta_k \\
&= (1 - \alpha_2 \cos^2 \theta_k)(f_k - f^*)
\end{aligned}
$$

for all sufficiently large k.

Hence, there is a constant $t = 1 - \alpha_2 \cos^2 \theta_k \in [0, 1)$ such that

$$
f(x_k) - f(x^*) \leq t^k [f(x_0) - f(x^*)].
\tag{5.9}
$$

By Eq. 5.6, we have

$$
f_k - f^* \geq \frac{1}{2} m \|x_k - x^*\|^2,
$$

which, together with Eq. 5.9, implies

$$
\|x_k - x^*\| \leq t^{k/2} \Big[\frac{2(f(x_0) - f(x^*))}{m} \Big]^{1/2}.
$$

This is to say the sequence $\{x_k\}$ is R-linearly convergent. The proof of this theorem is complete. ∎