The Lennard-Jones Potential Minimization Problem for Prion AGAAAAGA Amyloid Fibril Molecular Modeling

Jiapu Zhang

Centre in Informatics and Applied Optimization & Graduate School of ITMS, University of Ballarat, Mount Helen, VIC 3353, Australia. Emails: j.zhang@ballarat.edu.au, jiapu_zhang@hotmail.com, Phones: 61-3 5327 9809, 61-423 487 360 (mobile).

Abstract: The simplified Lennard-Jones (LJ) potential minimization problem is

$$\text{minimize} \quad f(x) = 4 \sum_{i=1}^{N} \sum_{j=1, j<i}^{N} \left(\frac{1}{\tau_{ij}^6} - \frac{1}{\tau_{ij}^3} \right) \quad \text{subject to} \quad x \in \mathbb{R}^n,$$

where $\tau_{ij} = (x_{3i-2} - x_{3j-2})^2 + (x_{3i-1} - x_{3j-1})^2 + (x_{3i} - x_{3j})^2$, $(x_{3i-2}, x_{3i-1}, x_{3i})$ is the coordinates of atom i in $\mathbb{R}^3$, $i, j = 1, 2, \ldots, N (\geq 2 \quad \text{integer})$, $n = 3N$ and N is the whole number of atoms. The nonconvexity of the objective function and the huge number of local minima, which is growing exponentially with N, interest many mathematical optimization experts. The global minimizer should be just at the point of the bottom of the LJ potential well. Based on this point, this paper tackles this problem illuminated by amyloid fibril molecular model building. The 3nhc.pdb, 3nve.pdb, 3nvf.pdb, 3nvg.pdb and 3nvh.pdb of PDB bank are used for the successful molecular modeling.

Key words: Lennard-Jones Potential Minimization Problem, Lennard-Jones Potential Well, van der Waals radii.

1 At the Bottom of the L-J Potential Well

Neutral atoms are subject to two distinct forces in the limit of large distance and short distance: a dispersion force (i.e. attractive van der Waals (vdw) force) at long ranges, and a repulsion force, the result of overlapping electron orbitals. The Lennard-Jones (L-J) potential represents this behavior (http://en.wikipedia.org/wiki/Lennard-Jones_potential, or [Locatelli and Schoen 2008] and references therein). The L-J potential is of the form

$$V(r) = 4\varepsilon \left[(\frac{\sigma}{r})^{12} - (\frac{\sigma}{r})^6 \right], \tag{1}$$

where r is the distance between two atoms, 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 $(\frac{\sigma}{r})^{12}$ term describes repulsion and the $(\frac{\sigma}{r})^6$ term describes attraction (Fig. 1). In Fig. 1 we may
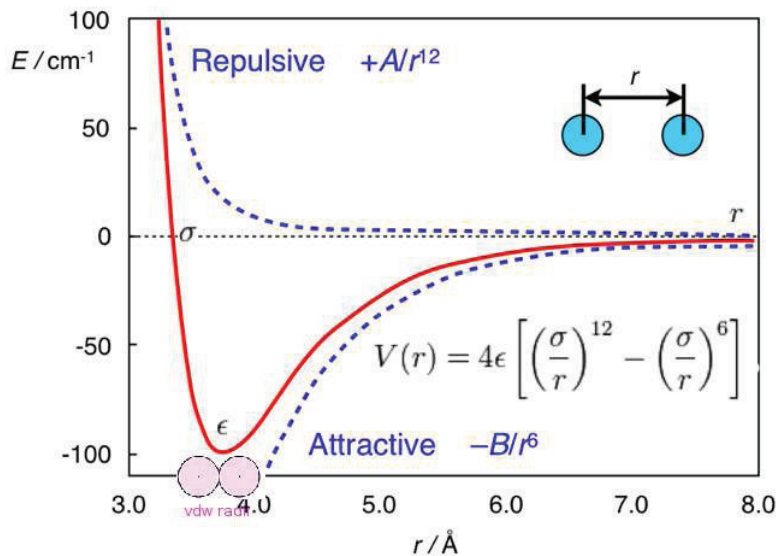
Figure 1: The Lennard-Jone Potential (formulas (1) and (7)) (This Figure can be found in website http://homepage.mac.com/swain/CMC/DDResources/mol_interactions/molecular_interactions.html).

see two points: (I) $V(r) = 0$ (but the value of $V(r)$ is not the minimal value) when $r = \sigma$ (i.e. the distance between two atoms equals to the sum of *atom radii* of the atoms); and (II) when $r = 2^{1/6}\sigma$ (i.e. the distance between two atoms equals to the sum of *vdw radii* of the atoms), the value of $V(r)$ reaches its minimal value $-\varepsilon$ (i.e. the bottom of the potential well; the force between the atoms is zero at this point). This paper is written based on (II). If we introduce the coordinates of the atoms whose number is denoted by N and let $\varepsilon = \sigma = 1$ be the reduced units, the form (1) becomes

$$f(x) = 4 \sum_{i=1}^{N} \sum_{j=1, j<i}^{N} \left(\frac{1}{\tau_{ij}^6} - \frac{1}{\tau_{ij}^3} \right), \tag{2}$$

where $\tau_{ij} = (x_{3i-2} - x_{3j-2})^2 + (x_{3i-1} - x_{3j-1})^2 + (x_{3i} - x_{3j})^2 = \|X_i - X_j\|_2^2$ and $(x_{3i-2}, x_{3i-1}, x_{3i})$ is the coordinates of atom i, $i, j = 1, 2, \ldots, N(\geq 2)$. The minimization of L-J potential $f(x)$ on $\mathbb{R}^n$ (where $n = 3N$) is an optimization problem:

$$\min_{s.t. x \in \mathbb{R}^{3N}} f(x). \tag{3}$$

For (3), when its global optimization solution is reached, the value r in (1) should be the sum of two *vdw radii* of the two atoms interacted. The three dimensional structure of a molecule with N atoms can be described by specifying the 3-Dimensional coordinate positions $X_1, X_2, \ldots, X_N \in \mathbb{R}^3$ of all its atoms. Given bond lengths r_{ij} between a subset S of the atom pairs, the determination of the molecular structure is

$$(\mathcal{P}_0) \quad to \quad find \quad X_1, X_2, \ldots, X_N \quad s.t. \quad \|X_i - X_j\| = r_{ij}, (i, j) \in S, \tag{4}$$

where $|| \cdot ||$ denotes a norm in a real vector space and it is calculated as the Euclidean distance 2-norm in this paper. (4) can be reformulated as a mathematical global optimization problem (GOP)

$$(\mathcal{P}) \qquad \min P(X) = \sum_{(i,j) \in S} w_{ij}(||X_i - X_j||^2 - r_{ij}^2)^2 \qquad (5)$$

in the terms of finding the global minimum of the function $P(X)$, where $w_{ij}, (i,j) \in S$ are positive weights, $X = (X_1, X_2, \ldots, X_N)^T \in \mathbb{R}^{N \times 3}$ [More et al. 1997] and usually S has many fewer than $N^2/2$ elements due to the error in the theoretical or experimental data [Zou et al. 1997, Grosso et al. 2009]. There may even not exist any solution $X_1, X_2, \ldots, X_N$ to satisfy the distance constraints in (4), for example when data for atoms $i, j, k \in S$ violate the triangle inequality; in this case, we may add a perturbation term $-\epsilon^T X$ to $P(X)$:

$$(\mathcal{P}_\epsilon) \qquad \min P_\epsilon(X) = \sum_{(i,j) \in S} w_{ij}(||X_i - X_j||^2 - r_{ij}^2)^2 - \epsilon^T X, \qquad (6)$$

where $\epsilon \geq 0$. Thus, the L-J potential optimization problem (3) is rewritten into the optimization problem (6). In this paper we elegantly solve 6 (thus solve (3)) illuminated by the prion AGAAAAGA amyloid fibril molecular modeling.

2 L-J Potential in Prion AGAAAAGA Amyloid Fibril Molecular Modelling

In 2007, Sawaya et al. got a breakthrough finding: the atomic structures of all amyloid fibrils revealed steric zippers, with strong vdw interactions (LJ) between β-sheets and hydrogen bonds (HBs) to maintain the β-strands [Sawaya et al. 2007]. Similarly as (1), i.e. the potential energy for the vdw interactions (Fig. 1) between β-sheets:

$$V_{LJ}(r) = \frac{A}{r^{12}} - \frac{B}{r^6}, \qquad (7)$$

the potential energy for the HBs between the β-strands has a similar formula

$$V_{HB}(r) = \frac{C}{r^{12}} - \frac{D}{r^{10}}, \qquad (8)$$

where A, B, C, D are constants given. Thus, the amyloid fibril molecular model building problem is reduced to well solve the optimization problem (3) or (6).

In this section, we will use suitable templates 3nvf.pdb, 3nvg.pdb and 3nvh.pdb from the Protein Data Bank (http://www.rcsb.org/) to build some amyloid fibril models. The models were built by the mutations in the use of SPDBV.4.01.PC (which make all the vdw contacts between the two β-sheets are very far), any Optimization Solver (which will remove the bad vdw/HB contacts) and Optimization program of Amber 11 (which furthermore refines the models and removes all the bad contacts and relax the models into a perfect way).

The amyloid fibril models of prion AGAAAAGA segment lastly refined by Amber 11 are illuminated in Fig.s 2-6. All these models are without any bad contact now (checked by package Swiss-PdbViewer), and the vdw interactions between the two β-sheets are in a very perfect way now. All the initial structures after SPDBV.01.PC
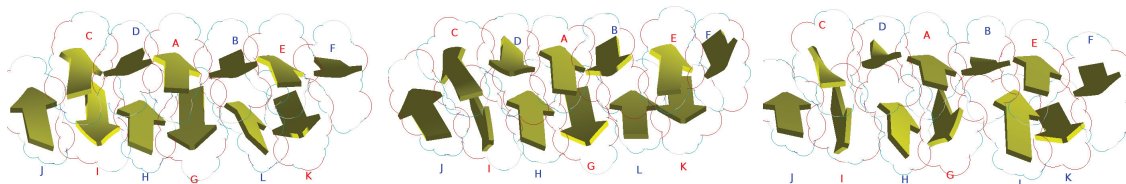

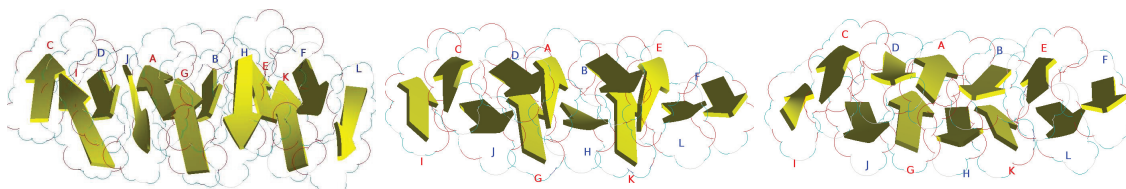Figure 2: Perfect 3nhc-Models 1-3 (from left to right respectively) for prion AGAAAAGA segment 113-120.

Figure 3: Perfect 3nve-Models 1-3 (from left to right respectively) for prion AGAAAAGA segment 113-120.
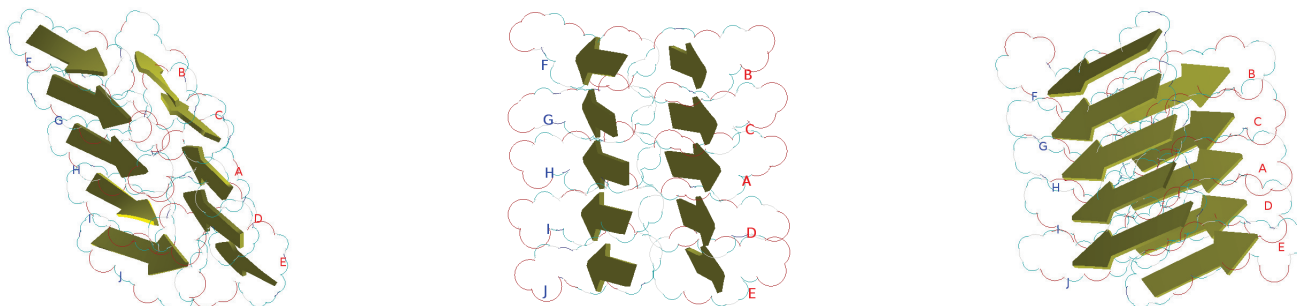

Figure 4: Perfect 3nvf-Models 1-3 (from left to right respectively) for prion AGAAAAGA segment 113-120.

mutations have very far vdw contacts between the two β-sheets. Now the vdw contacts come closer and reach a state with the lowest potential energy, which has perfect vdw
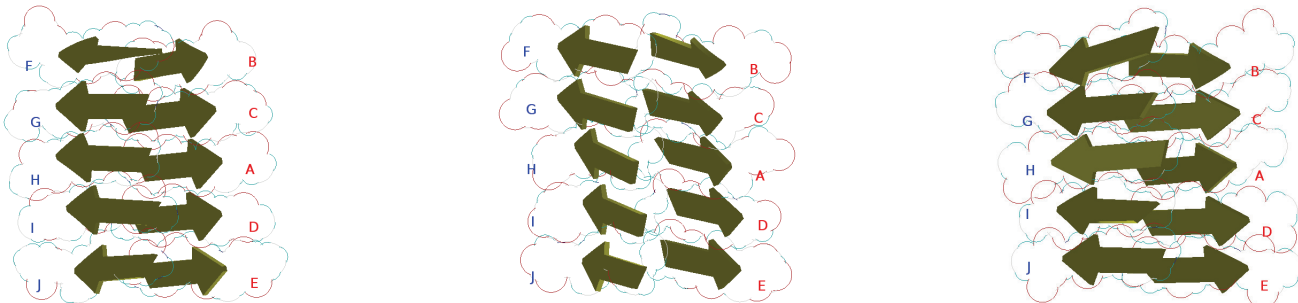
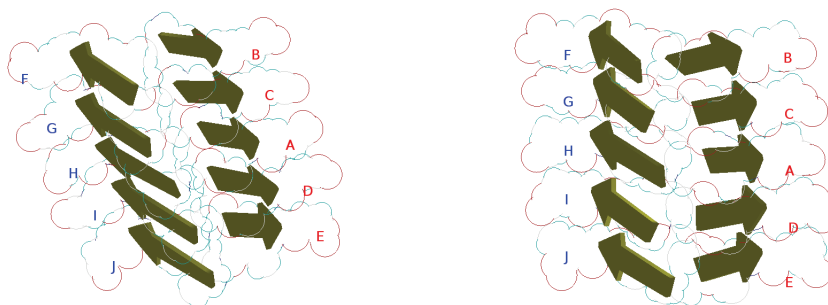Figure 5: Perfect 3nvg-Models 1-3 (from left to right respectively) for prion AGAAAAGA segment 113-120.

Figure 6: Perfect 3nvh-Models 1-2 (from left to right respectively) for prion AGAAAAGA segment 113-120.

contacts as shown in Fig.s 4-6.

References

[Apostol et al. 2011] Apostol M.I., Wiltzius J.J., Sawaya M.R., Cascio D., Eisenberg D., 2011. Atomic structures suggest determinants of transmission barriers in mammalian prion disease. Biochemistry 50: 2456–2463.

[Case et al. 2010] Case A.D., Darden T.A., Cheatham T.E., Simmerling III C.L., Wang J., Duke R.E., Luo R., Walker R.C., Zhang W., Merz K.M., Roberts B., Wang B., Hayik S., Roitberg A., Seabra G., Kolossvry I., Wong K.F., Paesani F., Vanicek J., Wu X., Brozell S.R., Steinbrecher T., Gohlke H., Cai Q., Ye X., Wang J., Hsieh M.-J., Cui G., Roe D.R., Mathews D.H., Seetin M.G., Sagui C., Babin V., Luchko T., Gusarov S., Kovalenko A., Kollman P.A., 2010. AMBER 11, University of California, San Francisco.

[Grosso et al. 2009] Grosso A., Locatelli M., Schoen F., 2009. Solving molecular distance geometry problems by global optimization algorithms. Comput. Optim. Appl. 43, 23-37.

[Locatelli and Schoen 2008] Locatelli, M., Schoen, F., 2008. Structure prediction and global optimization. Optima Mathematical Programming Society Newsletter USA 76, 1-8.

[More et al. 1997] More J.J., Wu Z.J., 1997. Global continuation for distance geometry problems. SIAM J. Optim. 7, 814-836.

[Sawaya et al. 2007] Sawaya M.R., Sambashivan S., Nelson R., Ivanova M.I., Sievers S.A., Apostol M.I., Thompson M.J., Balbirnie M., Wiltzius J.J., McFarlane H.T., Madsen A., Riekel C., Eisenberg D., 2007. Atomic structures of amyloid cross-beta spines reveal varied steric zippers. Nature 447, 453–7.

[Zou et al. 1997] Zou Z.H., Bird R.H., Schnabel R.B., 1997. A stochastic/perturbation global optimization algorithm for distance geometry problems. J. Glob. Optim. 11, 91-105.