International Journal of Pure and Applied Mathematics
Volume 42 No. 4 2008, 481-488

IMPLEMENTATION OF A MATHEMATICAL PROTEIN FOLDING MODEL

Yi Fang[1] §, Junmei Jing[2]
[1,2]Centre for Bioinformation Science
Mathematical Sciences Institute
Australian National University
Canberra, ACT 0200, AUSTRALIA
[1]e-mail: yi@maths.anu.edu.au
[2]e-mail: junmei@maths.anu.edu.au

AMS Subject Classification: 92C45, 92C35
Key Words: amino acids, protein folding model

1. Introduction

Proteins are linear polymers built of monomer units called *amino acids*. There are 20 different amino acids. A protein molecule can form infinite shapes in $\mathbb{R}^3$, but the protein performs its biological function only in one shape, its *native structure*. The consensus is that the native structure of a protein only depends on its amino acids sequence $U = A_1A_2 \cdots A_n$, see [1].

Based on well-known global geometric features of native structures of globular proteins, Yi Fang [3] suggested a geometric model for protein folding. Let $B(\mathbf{a}, r) \subset \mathbb{R}^3$ be the closed ball with center $\mathbf{a} \in \mathbb{R}^3$ and radius $r > 0$. The idea is that for each conformation $P = \cup_{i=1}^{N} B(\mathbf{a}_i, r_i)$ of a given amino acid sequence U, we can enclose P by any Lipschitz surface S and then to minimize an energy function $E(V(S), A(S), W(S))$ among all P and all S, where $V(S)$ is the volume enclosed by S, $A(S)$ is the area of S, and $W(S)$ is the area of hydrophobic part W of S, which is defined by a Lipschitz function $f_P : \mathbb{R}^3 \to \mathbb{R}$, and $W = \{\mathbf{x} \in S : \ f_P(\mathbf{x}) \leq 0\}$. A conformation P should satisfy the following

Received: August 17, 2007

§Correspondence author

steric conditions

$$\begin{aligned} r_i + r_j > \Delta_{ij} \geq |\mathbf{a}_i - \mathbf{a}_j| &\geq \delta_{ij} > 0, \text{ if chemically bonded;} \\ |\mathbf{a}_i - \mathbf{a}_j| &\geq \epsilon_{ij} > 0, \text{ if chemically not bonded.} \end{aligned} \tag{1.1}$$

Let $\mathcal{P}$ be the set of all possible conformations and $\mathcal{S}(P)$ be the set of all enclosing Lipschitz surfaces of P, the mathematical model is a constrained minimization problem, finding a $Q \in \mathcal{P}$ and a $\Sigma \in \mathcal{S}(Q)$ such that

$$E(V(\Sigma), A(\Sigma), W(\Sigma)) = \inf_{P\in\mathcal{P}, S\in\mathcal{S}(P)} E(V(S), A(S), W(S)). \tag{1.2}$$

The hypothesis then is that this Q is the native structure of the globular protein U. To implement the model, we simplify the model. Rolling a probe sphere of radius r along a conformation P generates a unique *molecular surface* M_P, see [2]. We then define f_P on M_P by (here $\bullet$ is the inner product in $\mathbb{R}^3$)

$$f_P(\mathbf{x}) = \mathbf{x}\bullet\mathbf{n}_i,\ \mathbf{x} \in A_i, \quad \mathbf{n}_i \in \mathbb{R}^3,\ |\mathbf{n}_i| = 1; \quad \cup_{i=1}^M A_i = \mathbb{R}^3. \tag{1.3}$$

The subsets A_i's are uniquely determined by P. In fact, $\mathbf{n}_i$ is the directed unit normal to the plane between two surface hydrophobic and hydrophilic atoms and passing through a probe center. Then the variational problem is simplified as finding a $Q \in \mathcal{P}$, such as

$$E(V(M_Q), A(M_Q), W(M_Q)) = \inf_{P\in\mathcal{P}} E(V(M_P), A(M_P), W(M_P)).$$

Here E is C^1 and is non-decreasing in each of its three variables.

To find the minimum structure Q, we have to change the conformation P to conformations with smaller E value. Let b_i, $1 \leq i \leq L$ be rotatable chemical bonds of the molecule U. The main method of changing the conformation P is rotating around these bonds. The ideal way of finding a minimum conformation is simultaneously rotating around these bonds with angles proportional to $-\bigtriangledown E = -\left(\frac{\partial E}{\partial b_1}, \cdots, \frac{\partial E}{\partial b_i}, \cdots, \frac{\partial E}{\partial b_L}\right)$, if we can prove that the final conformation is independent of the rotation order and if we can supply the analytic formulae of the these partial derivatives.

2. Rotation Order

Let $P = \cup_{i=1}^N B(\mathbf{a}_i, r_i)$ and $\mathbf{a}_{\alpha 0}$ and $\mathbf{a}_{\alpha 1}$ be bonded by b_α, the rotation line of b_α is $\mathbf{a}_{\alpha 0} + t\frac{\mathbf{a}_{\alpha 1}-\mathbf{a}_{\alpha 0}}{|\mathbf{a}_{\alpha 1}-\mathbf{a}_{\alpha 0}|} = \mathbf{a}_{\alpha 0} + t\mathbf{b}_\alpha$. Each b_α divides $\{\mathbf{a}_1, \cdots, \mathbf{a}_N\}$ into two groups F_α and R_α, we will rotate balls in R_α while fix balls in F_α. Note that these partitions are independent of P, they only depend on the chemical bonding

relations of the molecule. Let M_α be this rotation-fixization, we will show that

$$M_\alpha \circ M_\beta(X) = M_\beta \circ M_\alpha(X), \quad X \in (\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_N), \quad 1 \le \alpha, \beta \le L. \tag{2.1}$$

The formula of rotating a point X around a line $L : \mathbf{x} = \mathbf{a} + t\mathbf{b}$ ($|\mathbf{b}| = 1$) by an angle ω is $R(X) = \mathbf{a} + \mathbf{A}(\omega)(X - \mathbf{a})$. Let I be the identity matrix, $B = \mathbf{b}\mathbf{b}^T$ and Z_b the matrix such that the outer product $\mathbf{b} \wedge X = Z_b X$, then the orthonormal matrix $\mathbf{A}(\omega) = (1 - \cos\omega)B + \cos\omega I + \sin\omega Z_b$.

The topology of a protein molecule guarantees that if two bonds b_α and b_β such that $R_\alpha \subset R_\beta$, then $\{\mathbf{a}_{\alpha 0}, \mathbf{a}_{\alpha 1}\} \subset R_\beta$. Let b_1 and b_2 be two bonds and $L_1 : \mathbf{x} = \mathbf{a}_1 + t\mathbf{b}_1$ and $L_2 : \mathbf{x} = \mathbf{a}_2 + t\mathbf{b}_2$ be the two rotating lines and $X \in P = (\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_N)$. To prove (2.1), there are only two cases to consider: $R_1 \subset R_2$ and $R_1 \cap R_2 = \emptyset$. In any case, if $X \in F_1 \cap F_2$, then $M_1 \circ M_2(X) = M_2 \circ M_1(X) = X$. If $X \in R_1 \subset R_2$, then

$$M_2 \circ M_1(X) = \mathbf{a}_2 + \mathbf{A}_2(\omega_2)(\mathbf{a}_1 - \mathbf{a}_2) + \mathbf{A}_2(\omega_2)\mathbf{A}_1(\omega_1)(X - \mathbf{a}_1). \tag{2.2}$$

On the other hand b_1 and hence L_1 itself will be rotated by M_2, $L_3 = M_2(L_1) = \mathbf{a}_3 + t\mathbf{b}_3$, where $\mathbf{a}_3 = \mathbf{a}_2 + \mathbf{A}_2(\omega_2)(\mathbf{a}_1 - \mathbf{a}_2)$, $\mathbf{b}_3 = \mathbf{A}_2(\omega_2)\mathbf{b}_1$. Since $X \in R_1 \subset R_2$ and $M_2(X) \in R_1$ (in the new conformation $M_2(P)$ where rotation around b_1 is rotation around L_3), $M_1 \circ M_2(X)$ will be the rotation R_3 around L_3 of $M_2(X)$ by angle ω_1, thus

$$M_1 \circ M_2(X) = \mathbf{a}_2 + \mathbf{A}_2(\omega_2)(\mathbf{a}_1 - \mathbf{a}_2) + \mathbf{A}_3(\omega_1)\mathbf{A}_2(\omega_2)(X - \mathbf{a}_1). \tag{2.3}$$

Let $\mathbf{v} \in \mathbb{R}^3$ be an arbitrary vector, writing $\mathbf{A}_1(\omega_1) = \mathbf{A}_1$, $\mathbf{A}_2(\omega_2) = \mathbf{A}_2$, and $\mathbf{A}_3(\omega_1) = \mathbf{A}_3$, then

$$\mathbf{A}_2\mathbf{A}_1\mathbf{v} = (1 - \cos\omega_1)(\mathbf{b}_1 \bullet \mathbf{v})\mathbf{A}_2\mathbf{b}_1 + \cos\omega_1 \mathbf{A}_2\mathbf{v} + \sin\omega_1 \mathbf{A}_2(\mathbf{b}_1 \wedge \mathbf{v}). \tag{2.4}$$

For any orthonormal matrix O, $(O\mathbf{b}_1) \bullet (O\mathbf{v}) = \mathbf{b}_1 \bullet \mathbf{v}$, $O(\mathbf{b}_1 \wedge \mathbf{v}) = (O\mathbf{b}_1) \wedge (O\mathbf{v})$. Then by $\mathbf{b}_3 = \mathbf{A}_2(\omega_2)\mathbf{b}_1$,

$$\begin{aligned} \mathbf{A}_3\mathbf{A}_2\mathbf{v} &= (1 - \cos\omega_1)[\mathbf{b}_3 \bullet (\mathbf{A}_2\mathbf{v})]\mathbf{b}_3 + \cos\omega_1(\mathbf{A}_2\mathbf{v}) + \sin\omega_1 \mathbf{b}_3 \wedge (\mathbf{A}_2\mathbf{v}) \\ &= (1 - \cos\omega_1)(\mathbf{b}_1 \bullet \mathbf{v})\mathbf{A}_2\mathbf{b}_1 + \cos\omega_1 \mathbf{A}_2\mathbf{v} + \sin\omega_1 \mathbf{A}_2(\mathbf{b}_1 \wedge \mathbf{v}). \end{aligned} \tag{2.5}$$

Since $\mathbf{v}$ was arbitrary, (2.2) to (2.5) show (2.1) is true.

If $R_1 \cap R_2 = \emptyset$ and $X \in R_2$, then X and $M_2(X) \in F_1$ hence $M_1 \circ M_2(X) = M_2(X) = M_2 \circ M_1(X)$.

3. Gradient Formulae

The rotation around b_α induces a family of conformations $P_t \in \mathcal{P}$ and a one family homeomorphism ϕ_t such that $P_t = \phi_t(P)$ and $M_{P_t} = \phi_t(M_P)$. The extended vector field then is $X_\alpha(\mathbf{x}) = \left.\frac{\mathrm{d}\phi_t(\mathbf{x})}{\mathrm{d}t}\right|_{t=0}$. For our rotation-fixation moving

of conformation, we have the induced vector fields X_α on the balls as follows: if $\mathbf{a}_j \in F_\alpha$ then $\dot{\mathbf{a}}_j = X_\alpha(\mathbf{a}_j) = \mathbf{0}$; if $\mathbf{a}_j \in R_\alpha$ then $\dot{\mathbf{a}}_j = X_\alpha(\mathbf{a}_j) = \mathbf{b}_i \wedge (\mathbf{a}_j - \mathbf{a}_{\alpha 0})$. Let $\mu_i = \frac{\partial E}{\partial x_i}$, i = 1, 2, 3, then by classical variational formulae and the fact that W_t is defined via f_{P_t}, we have

$$\frac{\partial E}{\partial b_\alpha}(V(M_P), A(M_P), W(M_P)) = \mu_1 \frac{\partial V}{\partial b_\alpha} + \mu_2 \frac{\partial A}{\partial b_\alpha} + \mu_3 \frac{\partial W}{\partial b_\alpha}, \tag{3.1}$$

$$\frac{\partial V}{\partial b_\alpha} = -\int_{M_P} X_\alpha \bullet N \mathrm{d}\mathcal{H}^2, \quad \frac{\partial A}{\partial b_\alpha} = -2\int_{M_P} (X_\alpha \bullet N) H \mathrm{d}\mathcal{H}^2, \tag{3.2}$$

$$\frac{\partial W}{\partial b_\alpha} = \int_{\partial W} \left[X_\alpha \bullet \eta - \frac{\frac{\mathrm{d}f_P}{\mathrm{d}t}}{|\bigtriangledown_{M_P} f_P|} \right] \mathrm{d}\mathcal{H}^1 - 2\int_W (X_\alpha \bullet N) H \mathrm{d}\mathcal{H}^2, \tag{3.3}$$

where $\frac{\mathrm{d}f_P}{\mathrm{d}t} = \frac{\partial f_P}{\partial t} + X_\alpha \bullet \bigtriangledown f_P$, N and H are the inward unit normal and the mean curvature of M_P, η the unit outward conormal to ∂W, $\bigtriangledown_{M_P}$ the projection on tangent spaces of M_P of the gradient $\bigtriangledown f_P$. The molecular surface M_P consists of three kinds of faces, convex, concave, and saddle. We can do the integrals face by face such that all data needed are the coordinates of $\mathbf{a}_i$, i.e., the conformation P.

3.1. Convex Face

A convex face is a piece of spherical surface lying on some $S_i = \partial B(\mathbf{a}_i, r_i)$ and bounded by circular arcs γ_ν, $\nu = 1, \cdots, n_F$, let $\mathbf{v}_\nu^0$ and $\mathbf{v}_\nu^1$ be γ_ν's vertices and $\mathbf{c}_\nu$ and r_ν the center and radius of γ_ν's circle, $r_\nu \phi_\nu$ the arc length of γ_ν, $\mathbf{e}_3^\nu = (z_{\nu 1}, z_{\nu 2}, z_{\nu 3})$ be the unit vector in the direction of $(\mathbf{v}_\nu^0 - \mathbf{c}_\nu) \wedge (\mathbf{v}_\nu^1 - \mathbf{c}_\nu)$, $d_\nu = \mathbf{e}_3^\nu \bullet (\mathbf{c}_\nu - \mathbf{a}_i)$, $\mathbf{e}_1^\nu = \frac{\mathbf{v}_\nu^0 - \mathbf{c}_{ij_\nu}}{r_\nu} = (x_{\nu 1}, x_{\nu 2}, x_{\nu 3})$, $\mathbf{e}_2^\nu = \mathbf{e}_3^\nu \wedge \mathbf{e}_1^\nu = (y_{\nu 1}, y_{\nu 2}, y_{\nu 3})$, $1 \le \nu \le n_F$. A point $\mathbf{x}$ on F has the form $\mathbf{x} = \mathbf{a}_i - r_i N$ and $X_\alpha(\mathbf{x}) = \dot{\mathbf{a}} - r_i \dot{N}$, by $N \bullet \dot{N} \equiv 0$ and the general divergence formula on sphere we have

$$r_i \int_F (X_\alpha \bullet N) H \, \mathrm{d}\mathcal{H}^2 = \int_F X_\alpha \bullet N \, \mathrm{d}\mathcal{H}^2 = \frac{-1}{r_i} \dot{\mathbf{a}}_i \bullet \sum_{\nu=1}^{M} (X_\nu, Y_\nu, Z_\nu), \tag{3.4}$$

where

$$\begin{aligned} X_\nu = {} & \frac{r_\nu^2}{2} \{ \phi_\nu z_{\nu 1} + \sin\phi_\nu [\cos\phi_\nu (x_{\nu 2} y_{\nu 3} + x_{\nu 3} y_{\nu 2}) + \sin\phi_\nu (y_{\nu 2} y_{\nu 3} - x_{\nu 2} x_{\nu 3})] \} \\ & + r_\nu d_\nu z_{\nu 2} \left[y_{\nu 3} \sin\phi_\nu - x_{\nu 3} (1 - \cos\phi_\nu) \right], \end{aligned} \tag{3.5}$$

$$\begin{aligned} Y_\nu = {} & \frac{r_\nu^2}{2} \{ \phi_\nu z_{\nu 2} + \sin\phi_\nu [\cos\phi_\nu (x_{\nu 3} y_{\nu 1} + x_{\nu 1} y_{\nu 3}) + \sin\phi_\nu (y_{\nu 1} y_{\nu 3} - x_{\nu 1} x_{\nu 3})] \} \\ & + r_\nu d_\nu z_{\nu 3} \left[y_{\nu 1} \sin\phi_\nu - x_{\nu 1} (1 - \cos\phi_\nu) \right], \end{aligned} \tag{3.6}$$

$$Z_\nu = \frac{r_\nu^2}{2}\{\phi_\nu z_{\nu 3} + \sin\phi_\nu[\cos\phi_\nu(x_{\nu 1}y_{\nu 2} + x_{\nu 2}y_{\nu 1}) + \sin\phi_\nu(y_{\nu 1}y_{\nu 2} - x_{\nu 1}x_{\nu 2})]\}$$
$$+ r_\nu d_\nu z_{\nu 1}\left[y_{\nu 2}\sin\phi_\nu - x_{\nu 2}\left(1-\cos\phi_\nu\right)\right]. \tag{3.7}$$

It is true that either $F \cap W = \emptyset$ or $F \subset W$, and $\partial W \cap \overline{F} = \emptyset$.

3.2. Concave Face

A concave face F is a spherical polygon on the probe sphere S when S is simultaneously tangent to 3 balls $B(\mathbf{a}_i, r_i)$, $1 \le i \le 3$. We express F by parameters $t_i \ge 0$, $i = 1, 2, 3$,

$$\mathbf{x} = \mathbf{p} + rN = \mathbf{p} + r\frac{t_1\mathbf{a}_1 + t_2\mathbf{a}_2 + t_3\mathbf{a}_3 - \mathbf{p}}{|t_1\mathbf{a}_1 + t_2\mathbf{a}_2 + t_3\mathbf{a}_3 - \mathbf{p}|}, \quad t_1 + t_2 + t_3 = 1. \tag{3.8}$$

$$\phi_t(\mathbf{x}) = \mathbf{p}(t) + rN(t) = \mathbf{p}(t) + r\frac{t_1\mathbf{a}_1(t) + t_2\mathbf{a}_2(t) + t_3\mathbf{a}_3(t) - \mathbf{p}(t)}{|t_1\mathbf{a}_1(t) + t_2\mathbf{a}_2(t) + t_3\mathbf{a}_3(t) - \mathbf{p}(t)|},$$

$X_\alpha(\mathbf{x}) = \left.\frac{\mathrm{d}\phi_t(\mathbf{x})}{\mathrm{d}t}\right|_{t=0} = \dot{\mathbf{p}} + r\dot{N}$. Using $|\mathbf{p}(t) - \mathbf{a}_i(t)| = r_i + r = \text{constant}$, let $b_i = (\mathbf{a}_i - \mathbf{p})\bullet\dot{\mathbf{a}}_i$, $\mathbf{b} = (b_1, b_2, b_3)^T$, $A = \begin{pmatrix} \mathbf{a}_1 - \mathbf{p} \\ \mathbf{a}_2 - \mathbf{p} \\ \mathbf{a}_3 - \mathbf{p} \end{pmatrix}$, then $\det A \ne 0$, $\dot{\mathbf{p}} = A^{-1}\mathbf{b}$.

By $X_\alpha \bullet N = \dot{\mathbf{p}} \bullet N$,

$$r\int_F (X_\alpha \bullet N) H\,\mathrm{d}\mathcal{H}^2 = -\int_F X_\alpha \bullet N\,\mathrm{d}\mathcal{H}^2 = \frac{1}{r}\dot{\mathbf{p}} \bullet \sum_{i=1}^{3}(X_i, Y_i, Z_i). \tag{3.9}$$

Here the X_i, Y_i, and Z_i are the same as in (3.5) to (3.7).

Assume that $\mathbf{a}_1$ has different water association with $\mathbf{a}_2$ and $\mathbf{a}_3$, let $R_i = r_i + r$, $d_{ij} = |\mathbf{a}_i - \mathbf{a}_j|$, $y_{ij} = (R_i^2 - R_j^2)/2d_{ij}$. Then $f_P(\mathbf{x}) = (\mathbf{x} - \mathbf{p}) \bullet \mathbf{n}_k$, where $\mathbf{n}_k = (\mathbf{a}_k - \mathbf{a}_1)/d_{1k}$ is the directed unit normal of the dividing plane P_k (passing through $\mathbf{p}$ and $\mathbf{t}_{1k} = \frac{1}{2}(\mathbf{a}_1 + \mathbf{a}_k) + y_{1k}\mathbf{n}_k$ and perpendicular to it), $k = 2, 3$. The projection of $\partial W \cap F$ on the $\mathbf{a}_1\mathbf{a}_2\mathbf{a}_3$ plane is in the form of one or two curves γ_k, ($\{j, k\} = \{2, 3\}$)

$$t_k = A_k t_j + B_k,\ 0 \le t_j \le z_j, \quad A_k = \frac{d_{1j}\cos\omega}{-d_{1k}}, \quad B_k = \frac{d_{1k} + 2y_{1k}}{2d_{1k}}, \tag{3.10}$$

where $\cos\omega = \frac{(\mathbf{a}_2 - \mathbf{a}_1)\bullet(\mathbf{a}_3 - \mathbf{a}_1)}{d_{12}d_{13}}$. $F \cap W$ is a spherical polygon with arcs γ_ν, $1 \le \nu \le n$, including some γ_k as above, so $\int_{W\cap F}(X_\alpha \bullet N)H\mathrm{d}\mathcal{H}^2$ has the similar form as that in (3.9).

Let $\mathbf{A}_k = \mathbf{a}_j - \mathbf{a}_1 + A_k(\mathbf{a}_k - \mathbf{a}_1)$, $\mathbf{B}_k = B_k(\mathbf{a}_k - \mathbf{a}_1) + (\mathbf{a}_1 - \mathbf{p})$, $\mathbf{C}_k = \mathbf{B}_k \wedge \mathbf{A}_k$. Treating A_k and B_k as constants and setting $H_k = \dot{\mathbf{p}} \bullet \mathbf{C}_k$, $J_k = \dot{\mathbf{A}}_k \bullet \mathbf{C}_k$, and

$K_k = \dot{\mathbf{B}}_k \bullet \mathbf{C}_k$. Let $a_k t_j^2 + b_k t_j + c_k = |\mathbf{A}_k t_j + \mathbf{B}_k|^2 > 0$, then $\Delta_k = 4a_k c_k - b_k^2 > 0$. By $\eta = N'_{t_j} \wedge N / |N'_{t_j}|$ and $\mathrm{d}\mathcal{H}^1 = r|N'_{t_j}|\mathrm{d}t_j$,

$$\begin{aligned}\int_{(\partial W \cap F)\cap\gamma_k} X_\alpha \bullet \eta \ \mathrm{d}\mathcal{H}^1 =& \frac{2rH_k}{\sqrt{\Delta_k}} \left(\arctan \frac{2a_k z_j + b_k}{\sqrt{\Delta_k}} - \arctan \frac{b_k}{\sqrt{\Delta_k}} \right) \\ &+ \frac{2r^2 J_k}{\Delta_k} \left(2\sqrt{c_k} - \frac{b_k z_j + 2c_k}{\sqrt{a_k z_j^2 + b_k z_j + c_k}} \right) \\ &+ \frac{2r^2 K_k}{\Delta_k} \left(\frac{2a_k z_j + b_k}{\sqrt{a_k z_j^2 + b_k z_j + c_k}} - \frac{b_k}{\sqrt{c_k}} \right).\end{aligned} \tag{3.11}$$

Let $U_k = (\mathbf{A}_k \dot{\bullet} \mathbf{n}_k)$, $V_k = (\mathbf{B}_k \dot{\bullet} \mathbf{n}_k)$, $W_k = |\mathbf{C}_k \bullet \mathbf{n}_k| > 0$, then

$$\begin{aligned}\int_{\gamma_k} \frac{\frac{\mathrm{d}f_P}{\mathrm{d}t}}{|\bigtriangledown_{M_P} f_P|} \ \mathrm{d}\mathcal{H}^1 \ =& \frac{\pm 2r^2}{W_k} \left(\frac{(2a_k z_j + b_k)V_k}{\sqrt{a_k z_j^2 + b_k z_j + c_k}} - \frac{b_k V_k}{\sqrt{c_k}} \right. \\ &+ \left. 2\sqrt{c_k} U_k - \frac{(b_k z_j + 2c_k)U_k}{\sqrt{a_k z_j^2 + b_k z_j + c_k}} \right),\end{aligned} \tag{3.12}$$

where the sign is determined by orientation.

3.3. Saddle Face

A saddle face F is generated when the probe S simultaneously tangents to two balls $B(\mathbf{a}_1, r_1)$ and $B(\mathbf{a}_2, r_2)$, and rolls around the axis $\mathbf{e}_2 = \frac{\mathbf{a}_2 - \mathbf{a}_1}{d_{12}}$. The starting and stopping positions of the probe center is $\mathbf{p}$ and $\mathbf{q}$. Let $y = [(r_1 + r)^2 - (r_2 + r)^2]/2d_{12}$ and $\mathbf{t} = \frac{1}{2}(\mathbf{a}_1 + \mathbf{a}_2) + y\mathbf{e}_2$, $R = |\mathbf{p} - \mathbf{t}|$, $\mathbf{e}_1 = (\mathbf{p} - \mathbf{t})/R$, $\mathbf{e}_3 = \mathbf{e}_1 \wedge \mathbf{e}_2$, then F is parametrized by $0 \le \psi \le \psi_s$, $\theta_1 \le \theta \le \theta_2$,

$$\mathbf{x}(\psi, \theta) = \mathbf{t} + (R - r\cos\theta)(\cos\psi \mathbf{e}_1 + \sin\psi \mathbf{e}_3) + r\sin\theta \mathbf{e}_2, \tag{3.13}$$

where let $\omega_s = \arccos[(\mathbf{p} - \mathbf{t}) \bullet (\mathbf{q} - \mathbf{t})/R^2]$, then $\psi_s = \omega_s$ or $2\pi - \omega_s$. $\theta_1 = \arctan[-(d_{12} + 2y)/2R]$, $\theta_2 = \arctan[(d_{12} - 2y)/2R]$. These data are uniquely determined by the conformation P, see [2]. Let $\theta_k(t)$ and $\phi_s(t)$ be similarly defined for the conformation P_t, we can define $\phi_t(\psi) = \frac{\psi\psi_s(t)}{\psi_s}$, $\phi_t(\theta) = \frac{\theta_1(t)(\theta_2 - \theta) + \theta_2(t)(\theta - \theta_2)}{\theta_2 - \theta_1}$, and $U(\psi) = \cos\psi \mathbf{e}_1 + \sin\psi \mathbf{e}_3$, then for the same $0 \le \phi \le \phi_s$ and $\theta_1 \le \theta \le \theta_2$,

$$\phi_t(\mathbf{x}) = \mathbf{t}(t) + [R - r\cos\phi_t(\theta)]U(\phi_t(\psi)) + r\sin\phi_t(\theta)\mathbf{e}_2(t). \tag{3.14}$$

Let $\dot{U} = \cos\psi\dot{\mathbf{e}}_1 + \sin\psi\dot{\mathbf{e}}_3$, $U' = -\sin\psi\mathbf{e}_1 + \cos\psi\mathbf{e}_3$, then

$$\begin{aligned} X_\alpha(\mathbf{x}) &= \dot{\mathbf{t}} + (\dot{R} + r\dot{\theta}\sin\theta)U + (R - r\cos\theta)(\dot{U} + \dot{\psi}U') \\ &\quad + r\dot{\theta}\cos\theta\mathbf{e}_2 + r\sin\theta\dot{\mathbf{e}}_2. \end{aligned} \tag{3.15}$$

On F, $N = -\cos\theta U(\psi) + \sin\theta\mathbf{e}_2$, $\mathrm{d}\mathcal{H}^2 = r(R - r\cos\theta)\mathrm{d}\theta\mathrm{d}\psi$, $2H = \frac{R-2r\cos\theta}{r(R-r\cos\theta)}$. Let $J = J(\psi_s) = \int_0^{\psi_s} U(\psi)\mathrm{d}\psi$, then

$$\begin{aligned} 4\int_F X_\alpha \bullet N \ \mathrm{d}\mathcal{H}^2 &= 4rR(\phi_s\dot{\mathbf{t}}\bullet\mathbf{e}_2 - RJ\bullet\dot{\mathbf{e}}_2)(\cos\theta_1 - \cos\theta_2) \\ &+ 4rR(\phi_s\dot{R} + \dot{\mathbf{t}}\bullet J)(\sin\theta_1 - \sin\theta_2) \\ &- r^2(\phi_s\dot{\mathbf{t}}\bullet\mathbf{e}_2 - RJ\bullet\dot{\mathbf{e}}_2)(\cos 2\theta_1 - \cos 2\theta_2) \\ &+ r^2(\phi_s\dot{R} + \dot{\mathbf{t}}\bullet J)[2(\theta_1 - \theta_2) + \sin 2\theta_1 - \sin 2\theta_2], \end{aligned} \tag{3.16}$$

$$\begin{aligned} 2\int_F (X_\alpha \bullet N)H \ \mathrm{d}\mathcal{H}^2 &= 2R(\phi_s\dot{\mathbf{t}}\bullet\mathbf{e}_2 - R\bullet\dot{\mathbf{e}}_2)(\cos\theta_1 - \cos\theta_2) \\ &+ 2R(\phi_s\dot{R} + \dot{\mathbf{t}}\bullet J)(\sin\theta_1 - \sin\theta_2) \\ &- r(\phi_s\dot{\mathbf{t}}\bullet\mathbf{e}_2 - RJ\bullet\dot{\mathbf{e}}_2)(\cos 2\theta_1 - \cos 2\theta_2) \\ &+ r(\phi_s\dot{R} + \dot{\mathbf{t}}\bullet J)[2(\theta_1 - \theta_2) + \sin 2\theta_1 - \sin 2\theta_2]. \end{aligned} \tag{3.17}$$

Assume that $\mathbf{a}_1$ is hydrophobic and $\mathbf{a}_2$ is not, then the dividing plane P passing through $\mathbf{p}$ and $\mathbf{t}$ and is perpendicular to $\mathbf{e}_2$. The curve $\partial W \cap F$ is given by $\mathbf{x}(\psi) = \mathbf{t} + (R - r)U(\psi)$, $0 \le \phi \le \phi_s$, on which $\mathrm{d}\mathcal{H}^1 = (R - r)\mathrm{d}\phi$. The hydrophobic surface integral on F then is the same as in (3.17), except $\theta_1 = 0$. Since on $\partial W \cap F$, $\eta = N' \wedge N = \mathbf{e}_2$, $\frac{\mathrm{d}\theta(t)}{\mathrm{d}t}|_{t=0} = \dot{\theta}_0 = \frac{\dot{\theta}_1\theta_2 - \dot{\theta}_2\theta_1}{\theta_2 - \theta_1}$, by (3.15),

$$\int_{\partial W\cap F} X_\alpha \bullet \eta \, \mathrm{d}\mathcal{H}^1 = (R - r)\phi_s(r\dot{\theta}_0 + \dot{\mathbf{t}}\bullet\mathbf{e}_2) - (R - r)^2\dot{\mathbf{e}}_2\bullet J\,. \tag{3.18}$$

Let $\mathbf{n}_j = \mathbf{e}_2$, then $f_{P_t}(\phi_t(\mathbf{x})) = [\phi_t(\mathbf{x}) - \mathbf{t}(t)]\bullet\mathbf{n}_j(t)$, $|\bigtriangledown_{M_P} f_P| = \mathbf{n}_j \bullet \eta = 1$, and $\frac{\mathrm{d}f_{P_t}(\phi_t(X))}{\mathrm{d}t}|_{t=0} = [\phi_t(X)\dot{} - \mathbf{t}(t)]\bullet\mathbf{e}_2 + [(R - r)U]\bullet\dot{\mathbf{e}}_2 = r\dot{\theta}_0$

$$\int_{\partial W_P\cap F} \frac{\frac{\mathrm{d}f_P}{\mathrm{d}t}}{|\bigtriangledown_{M_P} f_P|}\mathrm{d}\mathcal{H}^1 = (R - r)\phi_s r\dot{\theta}_0. \tag{3.19}$$

References

[1] C.B. Anfinsen, Principles that govern the folding of protein chains, *Science*, **181** (1973), 223-230.

[2] M.L. Connolly, Analytical molecular surface calculation, *J. Appl. Cryst.*, **16** (1983), 548-58.

[3] Y. Fang, Mathematical protein folding problem, In: *Global Theory of Minimal Surfaces*, In: *Proceedings of the Clay Mathematics Institute* (Ed. David Hoffman), **2** (2005), 611-22.