

# Multicanonical simulation of coupled folding and binding of intrinsically disordered protein using an Ising-like protein model

Katsuyoshi Matsushita<sup>1,2</sup> and Macoto Kikuchi<sup>1,3,4</sup>

<sup>1</sup> Cybermedia Center, Osaka University, Toyonaka 560-0032, Japan

<sup>2</sup> Institute for Protein Research, Osaka University, Suita, 565-0871, Japan

<sup>3</sup> Graduate School of Science, Osaka University, Toyonaka 643-0043, Japan

<sup>4</sup> Graduate School of Frontier Biosciences, Osaka University, Suita, 565-0871, Japan

E-mail: kmatsu@cp.cmc.osaka-u.ac.jp

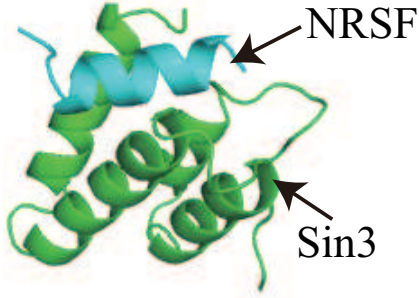
**Abstract.** An intrinsically disordered protein is one that does not spontaneously fold in physiological conditions but only folds when it binds to a target protein. Computer simulation of this coupled folding and binding is one of the central subjects of computational biophysics. Computing the free energy landscape is helpful in understanding coupled folding and binding. For this reason, we developed an Ising-like protein model and a multicanonical simulation in an energy-entropy space. The calculated free energy landscape indicates that coupled folding and binding induces rapid structural switching of the bound target protein.

## 1. Introduction

The multicanonical ensemble method [1] for model proteins has shed light on protein folding and binding. In particular, multicanonical simulations [2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12] provide us with a tool to investigate the free energy landscape, which is a key to understanding protein folding and binding. Recently, Higo *et al.* [13] applied a multicanonical simulation to the coupled folding and binding [14, 15, 16, 17, 18, 19, 20, 21] of intrinsically disordered (ID) proteins [22, 23, 24, 25]. An ID protein is one that is partially or wholly unstructured in physiological conditions, and coupled folding and binding are a folding transition driven by binding between the ID protein and a target protein. Multicanonical simulations succeeded in sampling various low-energy binding structures of an ID protein. However, information only from low-energy structures is not sufficient to achieve a deeper understanding of coupled folding and binding. This is because a protein is a finite size system, and the folding and binding occur at physiological temperatures. We have to clarify the structure of the free energy landscape. Multicanonical simulations to calculate the free energy landscape are difficult not only for elaborate atomic scale models [26, 27] but also for coarse-grained (Gō) models [28, 29], because such simulations require a very large amount of computational power to achieve a thermal equilibrium state. Therefore a simple model expressing the fundamental characteristics of the coupled folding and binding is necessary.

In the present work, we introduce a simple Ising-like protein model which is a variant of the Wako-Saitō-Muñoz-Eaton (WSME) model [30, 31, 32, 33]. The model consists only of binary





**Figure 1.** The binding structure of the binding part of NRSF on that of Sin3 (PDB:2czy).

variables and therefore enables us to easily carry out simulations in thermal equilibrium. We apply a multicanonical simulation method to the model. With this model and method, we investigate the coupled folding and binding of the neuron-restrictive silencer factor (NRSF) [34, 35] and corepressor Sin3. NRSF is an ID protein, and Sin3 is a NRSF-binding target protein. We obtain the free energy landscape representing the coupled folding and binding of NRSF. By comparing it with the free energy landscape of a fictitious binding part which folds in physiological conditions, we show that helical domains of Sin3 are prevented from forming non-cooperatively by the coupled folding and binding. As a result, cooperative forming of the helical domains of Sin3 is promoted. This result implies that the ID protein controls interprotein interaction with the target through the promotion of cooperative folding of the target protein.

## 2. Model and Method

For simplicity, we consider only the binding part of NRSF and that of Sin3, instead of full length proteins. The binding structure of the binding parts was obtained through a NMR experiment, as shown in figure 1 [36]. In the structure, NRSF forms a helix and bridges 4 helices of Sin3. In contrast, the binding part of NRSF without Sin3 is intrinsically disordered. Namely, the binding part has no specific native structure in physiological conditions. Since the structure is not obtained by NMR, it will be natural to consider that the binding part of Sin3 without NRSF is partially structured: the binding part consists of 4  $\alpha$ -helical domains which form independently.

To further gain information for constructing our model, we review a previous numerical study of low energy structures in the ID state of NRSF. The binding part of NRSF consists of 15 residues corresponding to the 43th-57th residues from the N-terminal. The size of the binding part is a tractable size on an atomic scale simulation. In fact, low energy structures were examined by a multicanonical simulation [13]. In the simulation, the binding part of NRSF without Sin3 predominantly adopts hairpin-like structures and adopts an  $\alpha$ -helix structure with a low probability. Furthermore a simulation of binding with Sin3 was also carried out and showed that the binding part of NRSF adopts various binding structures. We note that the structure of the binding part of Sin3 in the simulation was artificially restrained. This is because the number of residues of the binding part, 77, which corresponds to the 31st-107th residues from the N-terminal, is too large.

On the basis of the above observations of low energy structures, we consider the following simple model in order to illustrate the free energy landscape of the coupled folding and binding. We express a state of the binding parts using a set of binary variables,  $m_i$  and  $c_{i,j}$ .  $m_i$  represents the structure of the  $i$ -th residue and takes unity in an  $\alpha$ -helix and zero otherwise.  $c_{i,j}$  represents a contact between the  $i$ -th residue and  $j$ -th one and takes unity in hairpin structures and zero otherwise. The energy of a state expressed by  $\{m_i\}$  and  $\{c_{i,j}\}$  is defined by

$$\mathcal{H}(\{m_i\}, \{c_{i,j}\}) = \varepsilon \mathcal{H}_{\text{helix}}(\{m_i\}) + \eta \mathcal{H}_{\text{hairpin}}(\{m_i\}, \{c_{i,j}\}). \quad (1)$$

The first term in r. h. s. represents the energy stabilizing the helix structures of NRSF and Sin3. The explicit form of  $\mathcal{H}_{\text{helix}}$  is

$$\mathcal{H}_{\text{helix}}(\{m_i\}) = - \sum_{\langle i,j \rangle} \prod_{k=i}^j m_k, \quad (2)$$

where the summation is taken over all residue pairs natively contacting. We adopt the residue pairs from the helix domains of the binding structure obtained by NMR [36]. In this case, as the criterion of the inter-residue contact, we employ  $d_{i,j} < 0.65\text{nm}$  and  $|i - j| > 3$ , where  $d_{i,j}$  is distance between the  $i$ -th residue and  $j$ -th one.

The second term in r. h. s. of (1) is introduced to energetically stabilize hairpin structures of NRSF, which are observed as low-energy structures in the full atomic simulation [13]. Here  $\mathcal{H}_{\text{hairpin}}$  is

$$\mathcal{H}_{\text{hairpin}}(\{m_i\}, \{c_{i,j}\}) = - \sum_{\langle i,j \rangle} c_{i,j} \prod_{k=i}^j (1 - m_k) \quad (3)$$

where the summation is taken over the 2nd-14th, 5th-11th and 7th-9th residue pairs in NRSF. Here the residues are renumbered from N-terminal of the NRSF binding part.  $\varepsilon$  and  $\eta$  denote the coupling constant for inter-residue contacts in the helix structure and that in the hairpin-like structures, respectively.  $\eta$  is an important parameter for controlling disorder in the binding part of NRSF. In fact  $\eta = 0$  corresponds to a fictitious binding part that folds; and  $\eta = 2\varepsilon$  corresponds to the ID binding part in our simulation.

In this model, the state of the helix structure, namely  $m_i = 1$ , is unique and hence has no entropy. In contrast, the state of the hairpin structures or that of the others, namely  $m_i = 0$ , consists of many structures. Therefore we consider their entropy with the following term:

$$\mathcal{S}(\{m_i\}, \{c_{i,j}\}, \{b_{i,j}\}) = \mathcal{S}_{\text{otherwise}}(\{m_i\}, \{b_{i,j}\}) + \mathcal{S}_{\text{hairpin}}(\{m_i\}, \{c_{i,j}\}, \{b_{i,j}\}), \quad (4)$$

where  $\mathcal{S}_{\text{hairpin}}$  represents the entropy for hairpin structures of NRSF, and  $\mathcal{S}_{\text{otherwise}}$  represents that for the other structures of NRSF and Sin3.  $b_{i,j}$  represents a binding state between the  $i$ -th residue in NRSF and the  $j$ -th residue in Sin3 and adopts unity in binding and zero otherwise.

The entropy terms explicitly take the forms

$$\mathcal{S}_{\text{otherwise}}(\{m_i\}, \{b_{i,j}\}) = s \sum_i (1 - \tau_i(\{c_{i,j}\}, \{b_{i,j}\}))(1 - m_i), \quad (5)$$

$$\mathcal{S}_{\text{hairpin}}(\{m_i\}, \{c_{i,j}\}, \{b_{i,j}\}) = (1 - \mu)s \sum_i \tau_i(\{c_{i,j}\}, \{b_{i,j}\})(1 - m_i), \quad (6)$$

where  $\tau_i(\{c_{i,j}\}, \{b_{i,j}\}) = \max_{\langle j,k \rangle \langle p,q;r,s \rangle} \{\{c_{j,k}|j < i < k\}, \{b_{p,q}b_{r,s}|p < i < r\}\}$  for the  $i$ -th residue of NRSF and  $\tau_j(\{c_{i,j}\}, \{b_{i,j}\}) = \max_{\langle p,q;r,s \rangle} \{b_{p,q}b_{r,s}|r < j < s\}$  for the  $j$ -th residue of Sin3. In the maximization in  $\tau_i$ ,  $\langle p,q;r,s \rangle$  is taken over all neighboring binding pair. The neighboring binding pair  $\langle p,q;r,s \rangle$  indicates that there is no other binding site between the  $p$ -th residue and the  $r$ -th one or between the  $q$ -th residue and the  $s$ -th residue. For reproducing the intrinsically disordered state of NRSF, we take  $s = 1$  and  $\mu = 0.25$ . The value of  $\mu$  indicates that the binding part of NRSF loses 25% of the entropy at a cost of forming the hairpin structures. In (5) and (6), we introduce the dependence of  $\tau_i$  on  $b_{i,j}$  in order to consider an entropy effect coming from protein binding. For simplicity, we deal with the entropy effect of the binding by the same way with the entropy loss for the state forming hairpin structures.

We note that  $\varepsilon\mathcal{H}_{\text{helix}}$ ,  $\eta\mathcal{H}_{\text{hairpin}}$ ,  $\mathcal{S}_{\text{hairpin}}$ , and  $\mathcal{S}_{\text{otherwise}}$  described so far are adopted from the WSME model [30, 31, 32, 33, 37, 38] with some modifications for expressing the intrinsic disorder.

We also take into account the binding chemical potential defined on binding sites in NRSF and those in Sin3,  $h\mathcal{G}(\{m_i\}, \{b_{i,j}\})$ , to investigate binding-driven folding.

$$\mathcal{G}(\{m_i\}, \{b_{i,j}\}) = \sum_{\langle i,j \rangle} b_{i,j} m_i m_j \gamma_{i,j}(\{b_{k,l}\}, \{m_p\}), \quad (7)$$

where the summation is taken over all binding site pairs between NRSF and Sin3 and  $h$  denotes the chemical potential gain for each binding  $b_{i,j}$ .  $\gamma_{i,j}(\{b_{k,l}\}, \{m_p\}) = \prod_{\langle k,l \rangle} \{b_{k,l} (\prod_{p=i}^k m_p \prod_{q=j}^l m_q - 1) + 1\}$  and the product over  $\langle k,l \rangle$ 's is taken over the set of the neighboring binding of  $\langle i,j \rangle$ . The binding sites are adopted from the NMR structure [36] which satisfies the same criterion for the native contacts in  $\mathcal{H}_{\text{helix}}$ . The pairs of binding sites are the pair of the 29th residue from the N-terminal in Sin3 binding part and the 11th one from the N-terminal in NRSF binding part, 32nd-8th, 33rd-8th, 36th-5th and 66th-3rd. We introduces  $\gamma_{i,j}$  in order to express cooperativity between the neighboring binding. In  $\mathcal{G}$ ,  $b_{i,j}$  is coupled with a set of  $m_i$  and thereby gives rise to chemical potential gain only in the native binding structure ( $m_i = 1$ ). Therefore the coupling leads to the binding-driven folding.

The probability density of the state  $(\{m_i\}, \{c_{i,j}\}, \{b_{i,j}\})$  is given by

$$P(\{m_i\}, \{c_{i,j}\}, \{b_{i,j}\}) \propto \exp \{-\beta [\mathcal{H}(\{m_i\}) + h\mathcal{G}(\{m_i\}, \{b_{i,j}\})] + \mathcal{S}(\{m_i\}, \{c_{i,j}\}, \{b_{i,j}\})\}. \quad (8)$$

$\beta$  denotes the inverse temperature,  $1/k_B T$ , where  $T$  and  $k_B$  denote a temperature and the Boltzmann constant, respectively.

In order to obtain the free energy landscape, we calculate the density of states

$$n(C_{\text{NRSF}}, C_{\text{Sin3}}, E, S) = \sum_{\{m_i\}, \{c_{i,j}\}, \{b_{i,j}\}} \delta(\mathcal{H}(\{m_i\}) + h\mathcal{G}(\{m_i\}, \{b_{i,j}\}) - E) \\ \times \delta(\mathcal{S}(\{m_i\}, \{c_{i,j}\}, \{b_{i,j}\}) - S) \delta(C_{\text{NRSF}}(\{m_i\}) - C_{\text{NRSF}}) \delta(C_{\text{Sin3}}(\{m_i\}) - C_{\text{Sin3}}),$$

where  $C_{\text{NRSF}}(\{m_i\})$  denotes an order parameter for the  $\alpha$ -helix structure of NRSF, and  $C_{\text{Sin3}}(\{m_i\})$  denotes that of Sin3. They are defined by

$$C_p(\{m_i\}) = \frac{1}{N_p} \sum_{i \in \Omega_p} m_i, \quad (9)$$

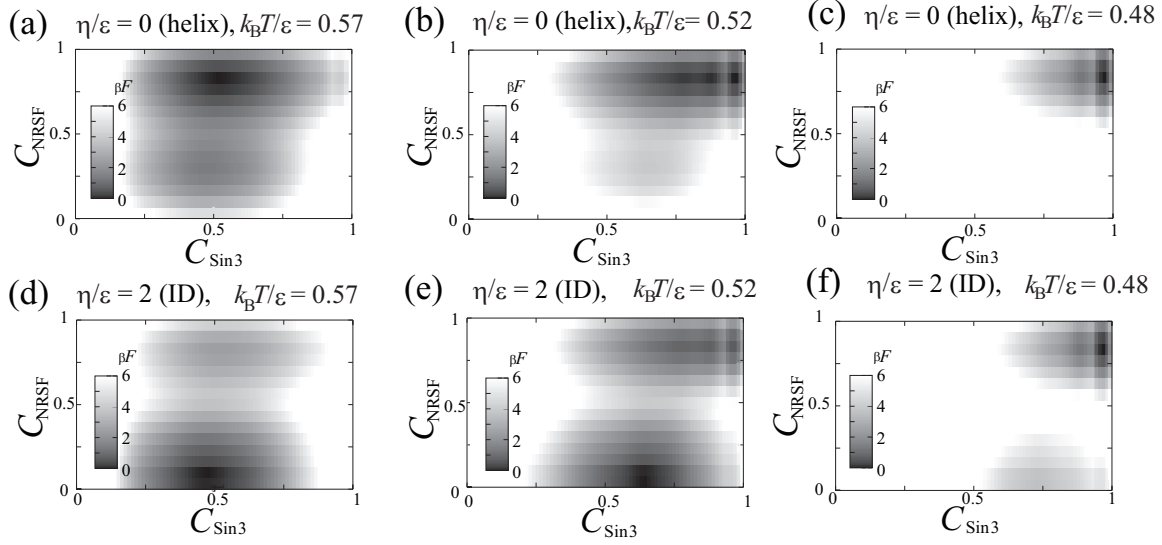
where  $\Omega_p$  and  $N_p$  denote the set and the number of all residues of the binding part of protein  $p$ , respectively. From the density of states, we obtain the free energy landscape

$$\beta F(C_{\text{NRSF}}, C_{\text{Sin3}}) = \ln \sum_{E, S} n(C_{\text{NRSF}}, C_{\text{Sin3}}, E, S) P(E, S), \quad (10)$$

where  $P(E, S) = \exp[-\beta E + S]$ . We calculate  $n(C_{\text{NRSF}}, C_{\text{Sin3}}, E, S)$  by using the Wang-Landau method [39] for the two dimensional space of  $E$  and  $S$ .

### 3. Results and Discussions

Let us compare the free energy landscapes between the case of a fictitious binding part ( $\eta = 0$ ) and that of the ID binding part ( $\eta = 2\varepsilon$ ) in order to examine the effect of the coupled folding and binding of the ID binding part. In figures 2(a)-2(f), we show the free energy landscapes for



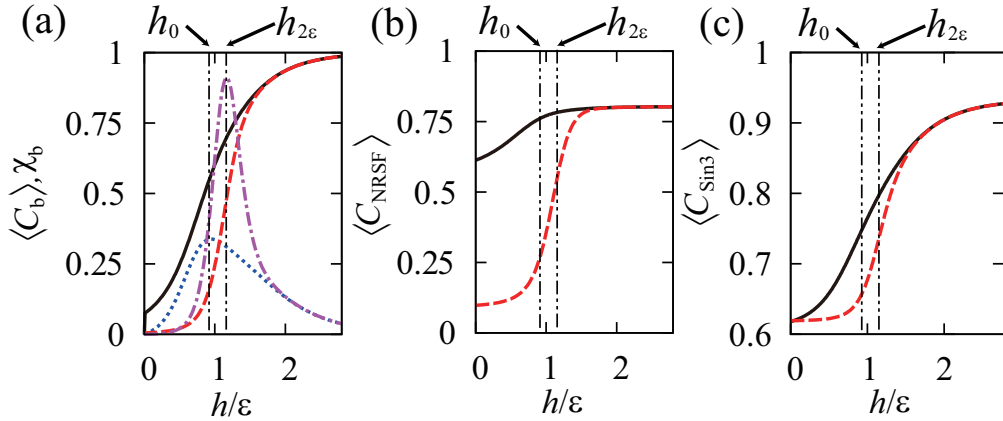
**Figure 2.** Density plots of the free energy landscape  $\beta F(C_{\text{NRSF}}, C_{\text{Sin3}})$ . The data are plotted for (a)  $k_{\text{B}}T = 0.57\epsilon$  and  $\eta = 0$ , (b)  $k_{\text{B}}T = 0.52\epsilon$  and  $\eta = 0$ , (c)  $k_{\text{B}}T = 0.48\epsilon$  and  $\eta = 0$ , (d)  $k_{\text{B}}T = 0.57\epsilon$  and  $\eta = 2\epsilon$ , (e)  $k_{\text{B}}T = 0.52\epsilon$  and  $\eta = 2\epsilon$  and (f)  $k_{\text{B}}T = 0.48\epsilon$  and  $\eta = 2\epsilon$ . Panels (a)-(c) correspond to the binding part folded at  $k_{\text{B}}T = 0.52\epsilon$  and panels (e)-(f) correspond to the ID binding part at  $k_{\text{B}}T = 0.52\epsilon$ .

$h = \epsilon$  at temperatures around  $k_{\text{B}}T = 0.52\epsilon$ , which corresponds to the folding temperature of NRSF for  $\eta = 0$ . At this temperature NRSF is in intrinsic disorder and exhibits binding-driven folding for  $h = 2\epsilon$ .

In the case of  $\eta = 0$ , as shown in figure 2(a), there are two free energy minima corresponding to the folded state ( $C_{\text{NRSF}} = 0.25$ ) and the unfolded one ( $C_{\text{NRSF}} = 0.75$ ) of NRSF at high temperatures. In this case, helical domains of Sin3 independently fold and therefore  $C_{\text{Sin3}}$  takes values around 0.5. In the case of  $k_{\text{B}}T = 0.52\epsilon$ , the unfolded NRSF becomes unstable, as shown in figure 2(b). Hence folding of NRSF is promoted with decreasing temperature. In this case,  $C_{\text{Sin3}}$  of the free energy minimum at  $C_{\text{NRSF}} = 0.75$  increases. Therefore the folding of NRSF promotes that of Sin3. However, Sin3 does not perfectly fold, as indicated by the  $C_{\text{NRSF}}$ , being widely distributed, from 0.5 to 1.0. At low temperatures, Sin3 almost perfectly folds as shown in figure 2(c). These free energy landscapes indicate that non-cooperative helical domain formation of Sin3 is promoted by the binding of NRSF for  $\eta = 0$ .

In the case of the ID binding part, namely,  $\eta = 2\epsilon$ , the free energy landscape in figure 2(d) has two minima similar to those observed for  $\eta = 0$ . At  $k_{\text{B}}T = 0.52\epsilon$ , the intrinsically disordered state of NRSF, corresponding to the minimum for  $C_{\text{NRSF}} \leq 0.25$ , is predominantly remained as shown in figure 2(e), in contrast to the case of  $\eta = 0$ . As a result, the partial folding of Sin3,  $(C_{\text{NRSF}}, C_{\text{Sin3}}) = (0.75, 0.75)$ , is not promoted by the binding of NRSF. Namely the partially-folded state of Sin3 is not predominant even at low temperatures. At  $k_{\text{B}}T = 0.48\epsilon$  only the completely-folded state of Sin3 is realized, as shown in figure 2(f). Thus, at the binding-driven folding transition, the ID binding part does not promote the non-cooperative formation of the helical domains of Sin3, and thereby only helps their cooperative formation.

To see the effects of cooperative helical domain formation on the structural switching of Sin3 in binding with NRSF, we examine the structural order parameter of Sin3,  $C_{\text{Sin3}}$ , as a function of  $h$ . We firstly identify the characteristic chemical potential,  $h_{2\epsilon}$ , where binding-driven folding of NRSF occurs for  $\eta = 2\epsilon$ . At  $h_{2\epsilon}$ , the binding susceptibility of  $C_{\text{b}}$  shows a peak due to the



**Figure 3.** (a)  $\langle C_b \rangle$  for  $\eta = 0$  (solid line) and  $2\varepsilon$  (dashed line) and its susceptibilities for  $\eta = 0$  (dotted line) and  $2\varepsilon$  (dashed-dotted line). (b)  $\langle C_{\text{NRSF}} \rangle$  for  $\eta = 0$  (solid line) and  $2\varepsilon$  (dashed line). (c)  $\langle C_{\text{Sin3}} \rangle$  for  $\eta = 0$  (solid line) and  $2\varepsilon$  (dashed line). The dashed double-dotted lines denote binding chemical potential  $h_0$  for the folded binding part and  $h_{2\varepsilon}$  for ID binding part.

binding. Here,  $C_b$  is a binding order parameter defined by

$$C_b(\{m_i\}, \{b_{i,j}\}) = \frac{\mathcal{G}(\{m_i = 0\}, \{b_{i,j} = 0\}) - \mathcal{G}(\{m_i\}, \{b_{i,j}\})}{\mathcal{G}(\{m_i = 0\}, \{b_{i,j} = 0\}) - \mathcal{G}(\{m_i = 1\}, \{b_{i,j} = 1\})}. \quad (11)$$

We also denote  $h$  with the susceptibility peak for  $\eta = 0$  by  $h_0$ . We show  $\langle C_b \rangle$  and susceptibilities as a function of  $h$  for  $\eta = 0$  and  $2\varepsilon$  in figure 3(a), where  $\langle \dots \rangle$  denotes an ensemble average.  $h_0$  and  $h_{2\varepsilon}$  are indicated by the dashed double-dotted lines.  $\langle C_{\text{NRSF}} \rangle$  for  $\eta = 2\varepsilon$  exhibits rapid switching around  $h_{2\varepsilon}$ , as shown in figure 3(b). Namely, a binding-driven folding transition actually occurs at  $h_{2\varepsilon}$ . Lastly, we show  $\langle C_{\text{Sin3}} \rangle$  as a function of  $h$  in figure 3(c).  $\langle C_{\text{Sin3}} \rangle$  for  $\eta = 2\varepsilon$  around  $h_{2\varepsilon}$  increases rapidly as compared to that for  $\eta = 0$  at  $h_0$ . Thus the coupled folding and binding of the ID binding part enhances sensitivity in the folding response of Sin3 to the change of chemical potential, namely a small change in  $h$  induces a large change in  $\langle C_{\text{Sin3}} \rangle$ .

Finally, we consider the implication of the sensitivity. The chemical potential correlates to the concentration of NRSF and Sin3. The change in chemical potential corresponds to the change in concentration of NRSF and Sin3. Therefore, the high sensitivity is efficient in sensing the concentration change. In addition, the folding of Sin3 controls the interaction between NRSF and Sin3 [40]. Thus ID, which induces high sensitivity, is important in the control of the interaction between NRSF and Sin3.

#### 4. References

- [1] Berg B A and Neuhaus T 1991 *Physics Letters B* **267** 249–253
- [2] Hansmann U H E and Okamoto Y 1993 *J. Comp. Chem* **14** 1333–1338
- [3] Hao M H and Scheraga H A 1994 *J. Chem. Phys* **98** 4940–4948
- [4] Y Okamoto U H E H 1995 *J. Phys. Chem* **99** 11276–11287
- [5] Kolinski A, Galazka W and Skolnick J 1996 *Proteins* **26** 271–287
- [6] Urakami N and Takasu M 1996 *J. Phys. Soc. Jpn* **65** 2694–2699
- [7] Kumar S, Payne P and Vásquez M 1996 *J. Comp. Chem.* **17** 1269–1275
- [8] Nakajima N, Nakamura H and Kidera A 1997 *J. Phys. Chem.* **101** 817–824
- [9] Noguchi H and Yoshikawa K 1997 *Chem. Phys. Lett.* **278** 184–188
- [10] Higo J, Nakajima N, Shirai H, Kidera A and Nakamura H 1997 *J. Comp. Chem.* **18** 2086–2092
- [11] Bartels C and Karplus M 1998 *J. Phys. Chem. B* **102** 865–880

- [12] Iba Y, Chikenji G and Kikuchi M 1998 *J. Phys. Soc. Jpn* **67** 3327–3330
- [13] Higo J, Nishimura Y and Nakamura H 2011 *J. Am. Chem. Soc.* **133** 10448–10458
- [14] Dunker A K, Garner E, Guilliot S, Romero P, Albrecht K, Hart J, Obradovic Z, Kissinger C and Villafranca J E 1998 *Pac. Symp. Biocomput.* **3** 473–484
- [15] Wright P E and Dyson H J 1999 *J. Mol. Biol.* **293** 321–331
- [16] Verkhivker G M, Bouzida D, Gehlhaar D K, Rejto P A, Freer S T and Rose P W 2003 *Proc. Natl. Acad. Sci. USA* **100** 5148–5153
- [17] Gupta N and Irbäck A 2004 *J. Chem. Phys.* **120** 3983
- [18] Hilser V H and Tompson E B 2007 *Proc. Natl. Acad. Sci. USA* **104** 8311–8315
- [19] Turjanski A D, Gutkind J S, Best R B and Hummer G 2008 *PLoS one* **4** e1000060
- [20] Chen H F 2009 *PLoS one* **4** e6516
- [21] Hounig Y and Liu Z 2009 *J. Mol. Bio.* **393** 1143–1159
- [22] Dyson H J and Wright P E 2002 *Curr Opin Struct Biol.* **12** 54–60
- [23] Papoian G A and Wolynes P G 2003 *Biopolymers* **68** 333–349
- [24] Sugase K, Dyson H J and Wright P E 2007 *Nature* **447** 1021–1025
- [25] Wright P E and Dyson H J 2009 *Curr Opin Struct Biol.* **19** 31–38
- [26] McCammon J A 1976 Molecular dynamics study of the bovine pancreatic trypsin inhibitor *Models for Protein Dynamics* (CECAM, Orsay, France) p 137
- [27] McCammon J A, Gelin B R and Karplus M 1977 *Nature* **267** 585–590
- [28] Clementi C, Nymeyer H and Onuchic J N 2000 *J. Mol. Biol* **298** 937–953
- [29] Koga N and Takada S 2001 *J. Mol. Biol.* **313** 171–180
- [30] Wako H and Saitô N 1978 *J. Phys. Soc. Jpn.* **44** 1931–1938
- [31] Wako H and Saitô N 1978 *J. Phys. Soc. Jpn.* **44** 1939–1945
- [32] Muñoz V, Henry E R, Hofrichter J and Eaton W A 1998 *Proc. Natl. Acad. Sci. USA* **95** 5872–5879
- [33] Muñoz V and Eaton W A 1999 *Proc. Natl. Acad. Sci. USA* **96** 11311–11316
- [34] Chong J A, Tapia-Ramirez J, Kim S, Toledo-Aral J J, Yingcong Zheng M C B, Altshuller Y M, Frohman M A, Kraner S D and Mandel G 1995 *Cell* **80** 949–957
- [35] Schoenherr C J and Anderson D J 1995 *Science* **267** 1360–1363
- [36] Nomura M, Uda-Tochio H, Murai K, Mori N and Nishimura Y 2005 *J. Mol. Bio.* **354** 903–915
- [37] Bruscolini P and Pelizzola A 2002 *Phys. Rev. Lett.* **88** 258101
- [38] Itoh K and Sasai M 2010 *Proc. Natl. Acad. Sci. USA* **107** 7775–7780
- [39] Wang F and Landau D P 2001 *Phys. Rev. Lett.* **86** 2050–2053
- [40] Fischer E 1894 *Ber. Dt. Chem. Ges.* **27** 2985–2993