Structure and energetics of protein–protein complexes

1 INTRODUCTION

Protein binding can be explained in terms of the funnel-based concept initially developed to describe protein folding (Bryngelson et al., 1995; Camacho and Vajda, 2001; Camacho et al., 1999; Dill, 1999; Elcock et al., 2001; Hunjan, 2008; Tovchirgrenchko and Vakser, 2001; Tsi et al., 1999; Vakser, 1996; Wang and Verkhivker, 2003; Wang et al., 2000; Wolynes, 2005). The concept suggests that unbound proteins are guided by the slope of the rugged energy landscape funnel into the bound state. The nature of the ruggedness and related effects is a subject of active research (Ferreiro et al., 2007; O’Toole and Vakser, 2008; Ruvinsky and Vakser, 2008a; Sutto et al., 2007). Highly frustrated interactions are observed on the protein surface near the binding site (Ferreiro et al., 2007).

© The Author 2009. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oxfordjournals.org
the energy of interaction between two proteins can be written as dependence of the critical cutoffs on the potential power.

In this article, we focus on determination of critical cutoffs for 11 power-type potentials at two thresholds of 5% and 10% of the artificial ruggedness decreases with the increase of the cutoff.

The idealized representation of proteins.

2 METHODS

A simplistic model of a sandwich-like protein complex (Ruvinsky and Vakser, 2008a; see also Laikatsky et al., 2006) can be used to describe the interaction energy between an atom A in Proteins 2 and Protein 1 atoms located at the distance r (Fig. 1) as

\[ U_{ij}(x, r) = \int_0^{\bar{r}} \frac{2\pi \rho_i}{\sin \theta} \left[ \frac{2}{d} \int_0^\rho d\rho_2 \right] \times \left[ \epsilon_0 - \delta_{ij} \right] (r^2 - r_s) dr, \]

where \( \bar{r} \) is the thickness of the spherical layer in Protein 1, \( \rho_i \) is Protein 1 atom density, \( s \) is a distance from the atom A to the surface of Protein 1 and \( \rho_2 = \cos \theta(r) \). For simplicity, we assume that potential \( \epsilon(r) \) does not depend on atom type. The total energy of the atom A is

\[ E_A = \int_0^{\bar{r}} \frac{2\pi \rho_i}{\sin \theta} \left[ \frac{2}{d} \int_0^\rho d\rho_2 \right] \times \left[ \epsilon_0 - \delta_{ij} \right] (r^2 - r_s) dr. \]

where \( R \) is the interaction cutoff. In Equation (2) we assumed that the interface, restricted by polar and azimuth angles, is flat (Fig. 1). Thus, the energy of interaction between two proteins can be written as

\[ E(R) = 2\pi \rho_2 \int_0^{\bar{r}} \frac{2}{d} \int_0^\rho d\rho_2 \times \left[ \epsilon_0 - \delta_{ij} \right] (r^2 - r_s) dr, \]

where \( S(x) \) is the area formed by Protein 2 atoms located at distance \( x \) from the interface, and \( s \) is the minimal distance between the two proteins. It is reasonable to assume that both proteins have equal densities \( \rho = \rho_1 = \rho_2 \) and \( S(x) \) is a weak function of the distance \( x \). Thus, we can rewrite Equation (3) as

\[ E(R) = 2\pi \rho_2 \int_0^{\bar{r}} \frac{2}{d} \int_0^\rho d\rho_2 \times \left[ \epsilon_0 - \delta_{ij} \right] (r^2 - r_s) dr. \]

where \( \epsilon_0 \) is the interaction cutoff. In Equation (2) we assumed that the artificial ruggedness drops below the threshold of 10% or 5% for cutoffs longer than the critical ones.

3 RESULTS AND DISCUSSION

The results of the calculations of the asymptotic behavior of the relative energy change at large cutoffs, of the artificial ruggedness, and the critical cutoffs for different power-type potentials are summarized in Table 1. The relative energy change \( \Delta E(R)/E(R) \) asymptotically approaches zero for \( n \leq 4 \), and approaches a constant \( -(n-4)\mu a/a \) for \( n > 4 \). The artificial ruggedness is a decreasing function of the cutoff for each of 11 potentials. The results show that both critical cutoffs depend non-monotonically on the potential power \( n \) (Fig. 2). They increase up to the maximum at \( n = 3 \pm 4 \) and then decrease with the power increase. The non-monotone character is readily explained by the interplay of the density-related term \( s^2 - rs \) and the energy \( \epsilon(r) \) in the double integral of Equations (4) and (5). The integral is dominated by the density-related term for slow-decreasing potentials \( n < 3 \) and by the energy term for fast decreasing potentials \( n > 4 \). The estimates of the critical cutoff for \( n = 6 \) and 12 are in a good agreement with our previously published results based on use of a soft Lennard–Jones potential on a set of 66 protein complexes (Ruvinsky and Vakser, 2008a). The difference between two cutoffs, which correspond...
the amino acid center of mass. A soft Lennard–Jones potential procedure, we replaced all atoms of each amino acid with a bead at one-bead coarse-grained models were built for the native and near-native of the interface was 1.0 Å. Both values of the average RMSD match relative to the native one was 1.7 Å. The average RMSD average root mean square deviation (RMSD) of the near-native match for each complex was selected for analysis. The closest to generate 5000 matches for each pair of proteins. The closest (http://vakser.bioinformatics.ku.edu/main/resources.php) was used to artificial ruggedness of 10% and 5%, decreases for Asymptotical behavior of the relative energy change for different potentials 

Table 1. Asymptotical behavior of the relative energy change for different potentials

| Types of interatomic interactions | The relative energy change $\delta(R)/E(R)$ | Asymptotical behavior of $\delta(E(R))/E(R)$ at large cutoffs |
|----------------------------------|------------------------------------------|---------------------------------------------------------------|
| $1/r^6 - \text{const}$           | $\frac{\alpha}{6} (\frac{4}{3} + 3 \ln R - \frac{3 \ln R}{2} + \frac{5}{2})$ | $\sim 8 \alpha (\delta R)$                                    |
| $1/r$                            | $-\frac{\alpha}{3} + \frac{\alpha}{8} \ln R$ | $\sim 3 \alpha \delta R$                                    |
| $1/r^2$                          | $-\frac{\alpha}{6} + \frac{\alpha}{8} \ln R$ | $\sim 4 \alpha / R$                                         |
| $1/r^3$                          | $-\frac{\alpha}{12} + \frac{\alpha}{10} \ln R$ | $\sim 2 \alpha (\ln R \alpha)$                              |
| $1/r^4$                          | $-\frac{\alpha}{20} + \frac{\alpha}{16} \ln R$ | $\sim 4 \alpha / (\ln R \alpha)$                           |
| $1/r^5$                          | $-\frac{\alpha}{30} + \frac{\alpha}{24} \ln R$ | $\sim 8 \alpha / (\ln R \alpha)$                           |
| $1/r^6$                          | $-\frac{\alpha}{42} + \frac{\alpha}{32} \ln R$ | $\sim 16 \alpha / (\ln R \alpha)$                          |
| $1/r^{12}$                       | $-\frac{\alpha}{252} + \frac{\alpha}{192} \ln R$ | $\sim 128 \alpha / (\ln R \alpha)$                         |

implemented in the GRAMM-X docking server (Tovchigrichenko and Vakser, 2005) was used to estimate the average ruggedness of all protein–protein complexes selected from a docking benchmark set (Gao et al., 2007). Our protein docking program GRAMM (http://vakser.bioinformatics.ku.edu/main/resources.php) was used to generate 5000 matches for each pair of proteins. The closest near-native match for each complex was selected for analysis. The average root mean square deviation (RMSD) of the near-native match relative to the native one was 1.7 Å. The average RMSD of the interface was 1.0 Å. Both values of the average RMSD play the role of the shift $\nu$ in the analysis above. In addition, one-bead coarse-grained models were built for the native and near-native conformations of each complex. Within the coarse-graining procedure, we replaced all atoms of each amino acid with a bead at the amino acid center of mass. A soft Lennard–Jones potential

$$e(r) = \frac{\varepsilon}{(r^6 + \alpha^6)} - \frac{\alpha}{(r^6 + \alpha^6)^{1/2}}$$

Fig. 2. The critical cutoff as a function of the potential power.
Since protein folding and protein binding are similar processes in terms of the landscape characteristics, including the funnel concept, we may expect that our results have implications to protein folding. Systematic attempts have been undertaken to design pair potentials for protein folding (Tobi and Elber, 2000; Tobi et al., 2000; Vendruscolo and Domany, 1998; Vendruscolo et al., 1999). Using machine learning algorithms, the authors of these studies clearly showed that a set of contact potentials with cutoffs of 8.5 Å or 9 Å, which guarantees the native structure energies lower than those of the decoys, does not exist. Then, using different resolutions of the potential functions, the same learning algorithm, and the 9 Å cutoff, the flexible functional forms of potentials were optimized. Based on the performance of the potentials, it was noted that it is impossible to find a pair potential with the flexible form that recognizes all native folds (Tobi and Elber, 2000; Tobi et al., 2000). Developing contact potentials with the cutoff of 7.5 Å for predicting stability changes in proteins upon mutations, Khutam et al. (2004) note that, 'it is impossible to reach experimental accuracy and derive fully transferable contact parameters using the contact models of potentials'. The choice of the cutoff may partly explain these results and thus encourage new attempts to parameterize potentials for longer ranges. Indeed, the 9 Å range is less than the critical cutoffs of power potentials for $n \leq 6$ and the artificial ruggedness threshold of 5%, or for $n \leq 8$ and the artificial ruggedness threshold of 10% (Table 1). For example, the artificial ruggedness of the energy landscape described by contact or Coulomb potentials cutoff at 8–9 Å is 17–19%. Since substantially frustrated landscapes are not adequate approximations of actual energy profiles due to the principle of minimal frustration, (Bryngelson et al., 1995; Dill, 1999; Miller and Dill, 1997; Tsi et al., 1999; Wolynes, 2005), the above studies had limited chances to detect the actual parameters of the interactions. Our results suggest that using longer cutoffs with such algorithms may improve the potentials.

4 CONCLUSION

Studies of ruggedness of protein–protein energy landscape are important for understanding the connection between protein structure, function and dynamics. We have analyzed energy fluctuations and the artificial contribution to the ruggedness of the protein–protein energy landscape by limited range interactions described by $1/r^n$ potentials. The results show that the undesirable artificial ruggedness exists for short cutoffs and gradually disappears with the cutoff increase. We calculated the critical values of the cutoff for each of 11 popular power-type potentials with $n = 0.1\times 9.12$ and for two thresholds of 5% and 10%. We showed that for both thresholds, the critical cutoff is a non-monotonic function of the potential power $n$. These functions reach the maximum at $n = 3–4$ and then decrease with the increase of the potential power. The difference between the cutoffs for 5% and 10% artificial ruggedness becomes negligible for potentials decreasing faster than $1/r^{12}$. The analytical results were validated on the dataset of 62 protein–protein complexes, with different parameterizations of the soft Lennard–Jones potential and two types of protein representations: all-atom and coarse-grained. The results suggest that the cutoffs larger than the critical ones can be recommended for protein–protein potentials.

Funding: National Institutes of Health (R01 GM074255).

Conflict of Interest: none declared.

REFERENCES

Akberovic,VI et al. (1995) Impact of local and non-local interactions on thermodynamics and kinetics of protein folding. J. Mol. Biol., 282, 460–471.

Alsallaq, R. and Zhou, H.X. (2007) Prediction of protein-protein association rates from electrostatic interactions in macromolecules. Proteins, 69, 845–851.

Allalouf, R. and Zhou, H.X. (2007) Prediction of protein-protein association rates from electrostatic interactions in macromolecules. Proteins, 69, 845–851.

Altman, E. et al. (2000) The use of gene expression profiles to predict laboratory measurements of technical performance. J. Am. Med. Assoc., 283, 328–335.

Anderson, D.L. et al. (1999) Mapping the sequence space of a protein fold. Proc. Natl Acad. Sci. USA, 96, 15234–15239.

Bahar, I. et al. (1997) Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. Fold. Des., 2, 175–181.

Baker, D.A. et al. (1999) How the range of pair interactions governs features of multidimensional potentials. J. Chem. Phys., 93, 7845–7856.

Banks, R.J. et al. (2001) Computer simulation of protein-protein interactions: a synthesis. Proteins, 41, 167–195.

Banks, R.J. and Wolynes, P.G. (1998) Intermediates and barrier crossing in a random energy model (with applications to protein folding). J. Phys. Chem. B, 102, 4913–4926.

Buerger, N. et al. (2004) Orientational potentials extracted from protein structures improve native fold recognition. Proteins, 56, 519–534.

Camacho, C.J. and Vajda, S. (2001) Protein docking along smooth association pathways. Prot. Sci., 10, 166–177.

Camacho, C.J. et al. (1999) Free energy landscapes of encounter complexes in protein-protein association. Biophys. J., 76, 1066–1078.

Camacho, C.J. et al. (2001) Energy landscapes of encounter complexes in protein-protein association. Biophys. J., 79, 1479–1485.

Camacho, C.J. et al. (2001) Free energy landscapes of encounter complexes in protein-protein association. Biophys. J., 79, 1479–1485.

Camps, F. et al. (2007) Docking: A new program for fast protein-protein docking. Nucleic Acids Res., 35, 192–198.

Camps, F. et al. (2007) Docking: A new program for fast protein-protein docking. Nucleic Acids Res., 35, 192–198.

Camacho, C.J. and Vajda, S. (2001) Protein docking along smooth association pathways. Prot. Sci., 10, 166–177.

Carrington, M. et al. (1999) Kinetic simulation of the folding-fusion dynamics of an SH3 domain. Proc. Natl Acad. Sci. USA, 96, 5408–5413.

Carrington, M. et al. (1999) Kinetic simulation of the folding-fusion dynamics of an SH3 domain. Proc. Natl Acad. Sci. USA, 96, 5408–5413.

Carrington, M. et al. (1999) Kinetic simulation of the folding-fusion dynamics of an SH3 domain. Proc. Natl Acad. Sci. USA, 96, 5408–5413.

Carrington, M. et al. (1999) Kinetic simulation of the folding-fusion dynamics of an SH3 domain. Proc. Natl Acad. Sci. USA, 96, 5408–5413.

Carrington, M. et al. (1999) Kinetic simulation of the folding-fusion dynamics of an SH3 domain. Proc. Natl Acad. Sci. USA, 96, 5408–5413.

Carrington, M. et al. (1999) Kinetic simulation of the folding-fusion dynamics of an SH3 domain. Proc. Natl Acad. Sci. USA, 96, 5408–5413.

Carrington, M. et al. (1999) Kinetic simulation of the folding-fusion dynamics of an SH3 domain. Proc. Natl Acad. Sci. USA, 96, 5408–5413.

Carrington, M. et al. (1999) Kinetic simulation of the folding-fusion dynamics of an SH3 domain. Proc. Natl Acad. Sci. USA, 96, 5408–5413.

Carrington, M. et al. (1999) Kinetic simulation of the folding-fusion dynamics of an SH3 domain. Proc. Natl Acad. Sci. USA, 96, 5408–5413.

Carrington, M. et al. (1999) Kinetic simulation of the folding-fusion dynamics of an SH3 domain. Proc. Natl Acad. Sci. USA, 96, 5408–5413.

Carrington, M. et al. (1999) Kinetic simulation of the folding-fusion dynamics of an SH3 domain. Proc. Natl Acad. Sci. USA, 96, 5408–5413.

Carrington, M. et al. (1999) Kinetic simulation of the folding-fusion dynamics of an SH3 domain. Proc. Natl Acad. Sci. USA, 96, 5408–5413.
Gorinshanskaya,S. and Goldstein,R.A. (1995) Optimal local propensities for model proteins. Prot. Sci., 22, 413–418.

Grinell,J.R. et al. (1999) Increasing protein stability by altering long-range Coulombic interactions. Prot. Sci., 8, 1843–1849.

Gromova,M. and Selvaduraj,S. (1999) Importance of long-range interactions in protein folding. Biophys. Chem., 77, 49–68.

Halak,R. and McCammon,J.A. (1998) Correcting for electrostatic cutoffs in free energy simulations: toward consistency between simulations with different cutoffs. Chem. Phys. Lett., 306, 961–9623.

Harrow,S. (1990) Treatment of electrostatic effects in macromolecular modeling. Prot. Sci., 8, 78–82.

Hittner,K. (2000) Structural flexibility in proteins: impact of the crystal environment. Bioinformatics, 24, 521–528.

Hunjan,J. et al. (2008) The size of the intramolecular energy funnel in protein interactions. Prot. Sci., 72, 344–352.

Hysom,C. and Thirumalai,D. (2003) Can energy landscape roughness of proteins and RNA be measured by using mechanical unfolding experiments? Proc. Natl Acad. Sci. USA, 100, 10249–10253.

Leverko,S. et al. (2008) Coarse-graining in interaction space: a systematic approach for replacing long-range electrostatics with short-range potentials. J. Phys. Chem., 112, 4711–4724.

Kaplan,J.G. (2004) Intermolecular Interactions: Physical Picture, Computational Methods and Model Potentials. John Wiley & Sons, Ltd, Chichester.

Khatun,J. et al. (2004) Can contact potentials reliably predict stability of proteins? J. Mol. Biol., 336, 1223–1238.

Klein-Seetharaman,J. et al. (2002) Long-range interactions within a nonnative protein. Science, 295, 1719–1722.

Lafuente,E. et al. (2002) Long-range RNA-RNA interactions between distant regions of the hepatitis C virus internal ribosome entry site element. J. Gen. Virol., 83, 1113–1121.

Lonchampt,R. and Brocks,R.B. (1989) The effects of truncating long-range forces on protein dynamics. Prot. Sci., 8, 52–65.

Lukatsky,D.B. et al. (2006) Statistically enhanced self-attraction of random patterns. Phys. Rev. Lett., 97, 178101.

Miller,D.W. and Dill,K.A. (1997) Ligand binding to proteins: the binding landscape model. J. Mol. Biol., 266, 2166–2179.

Miller,M.A. et al. (1999) Structural relaxation in Morse clusters: energy landscapes. J. Chem. Phys., 110, 328–334.

Mink,D.D.L. et al. (2005) The entropic cost of protein-protein association. A case study on acylcholinesterase binding to fasciculin-2. Biophys. J. Biophys. Lett., 89, 235–247.

Nevo,R. et al. (2005) Direct measurement of protein energy landscape roughness. EMBO Rep., 6, 482–486.

Nerurkar,J. and Nilson,L. (2000) On the truncation of long-range electrostatic interactions in DNA. Biophys. J., 79, 1537–1553.

O’Toole,N. and Vakser,I.A. (2008) Large-scale characteristics of the energy landscape in protein-protein interactions. Prot. Sci., 71, 144–152.

Pappu,R.V. et al. (1999) A potential smoothing algorithm accurately predicts transmembrane helix packing. Nat. Struct. Biol., 6, 50–55.

Ruvinsky,A.M. (2007) Calculations of protein-ligand binding entropy of relative and overall molecular motions. J. Comput. Aided Mol. Des., 21, 361–370.

Ruvinsky,A.M. and Kozintsev,A.V. (2005a) The key role of atom types, reference states, and interaction cutoff radii in the knowledge-based method: new variational approach. Prot. Sci., 58, 845–851.

Ruvinsky,A.M. and Kozintsev,A.V. (2005b) New and fast statistical-thermodynamic method for computation of protein-ligand binding entropy substantially improves docking accuracy. J. Comput. Chem., 26, 1089–1095.

Ruvinsky,A.M. and Vakser,I.A. (2008a) Interaction cutoff effect on ruggedness of protein-protein energy landscape. Prot. Sci., 70, 1498–1505.

Ruvinsky,A.M. and Vakser,I.A. (2008b) Chasing funnels on protein-protein energy landscapes at different resolutions. Biophys. J., 95, 2150–2159.

Sagu,C. and Darden,T.A. (1999) Molecular dynamics simulations of biomolecules: long-range electrostatics effic. Ann. Rev. Biophys. Biomol. Struct., 28, 155–179.

Stillinger,F.H. and Stillinger,D.K. (1990) Cluster optimization simplified by interaction modification. J. Chem. Phys., 93, 6106–6107.

Sutto,L. et al. (2007) Consequences of localized frustration for the folding mechanism of the IAP protein. Proc. Natl Acad. Sci. USA, 104, 19825–19830.

Tetens,M.M. (1996) Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. Phys. Rev. Lett., 77, 1905–1908.

Toh,D. and Elber,R. (2000) Distance-dependent, pair potential for protein folding: Results from linear optimization. Prot. Sci., 41, 40–46.

Toh,D. et al. (2000) On the design and analysis of protein folding potentials. Prot. Sci., 40, 71–85.

Tovchigrechko,A. and Vakser,I.A. (2001) How common is the funnel-like energy landscape in protein-protein interactions? Prot. Sci., 10, 1572–1583.

Tovchigrechko,A. and Vakser,I.A. (2005) Development and testing of an automated approach to protein docking. Prot. Sci., 40, 296–301.

Tovchigrechko,A. et al. (2002) Docking of protein models. Prot. Sci., 11, 1888–1896.

Tsai,C.-J. et al. (1999) Folding funnels, binding funnels, and protein function. Prot. Sci., 8, 1181–1191.

Vakser,I.A. (1996) Long-distance potentials: an approach to the multiple-minima problem in ligand-receptor interaction. Prot. Engng., 9, 37–41.

Vakser,I.A. and Jiang,S. (2002) Strategies for modeling the interactions of transmembrane helices of G-protein coupled receptors by geometric complementarity using the GRAMM computer algorithm. Methods Enzymol., 343, 313–328.

Vakser,I.A. et al. (1999) A systematic study of low-resolution recognition in protein-protein complexes. Proc. Natl Acad. Sci. USA, 96, 8477–8482.

Vemucrola,S.M. and Domany,E. (1996) Pairwise contact potentials are unsuitable for protein folding. J. Chem. Phys., 109, 11011–11018.

Vemucrola,S.M. et al. (1999) Protein folding in contact map space. Phys. Rev. Lett., 82, 856–859.

Wang,J. and Verkhivker,G.M. (2003) Energy landscape theory, funnels, specificity, and optimal criterion of biomolecular binding. Phys. Rev. Lett., 90, 188101.

Wang,J. et al. (2007) Optimal specificity and function for flexible biomolecular recognition. Biophys. J., 92, L109–L111.

Worak,R.J. et al. (1992) Application of the diffusion equation method of global optimization to water clusters. J. Phys. Chem., 96, 5130–5145.

Winston,T.W. and Strahb,J.E. (2002) Gravitational smoothing as a global optimization strategy. J. Comput. Chem., 23, 1100–1102.

Wolynes,P.J. (2005) Recent successes of the energy landscape theory of protein folding and function. Quart. Rev. Biophys., 38, 405–410.