John von Neumann Institute for Computing

Correlation Effects in Protein-Protein Recognition

H. Behringer, F. Schmid

published in

From Computational Biophysics to Systems Biology (CBSB08),
Proceedings of the NIC Workshop 2008,
Ulrich H. E. Hansmann, Jan H. Meinke, Sandipan Mohanty,
Walter Nadler, Olav Zimmermann (Editors),
John von Neumann Institute for Computing, Jülich,
NIC Series, Vol. **40**, ISBN 978-3-9810843-6-8, pp. 165-168, 2008.

http://www.fz-juelich.de/nic-series/volume40

Correlation Effects in Protein-Protein Recognition

Hans Behringer and Friederike Schmid

Fakultät für Physik, Universität Bielefeld, D-33615 Bielefeld, Germany
E-mail: {behringer, schmid}@physik.uni-bielefeld.de

Correlation effects in the distribution of hydrophobic and polar residues are investigated within an idealised coarse-grained model for the recognition of two rigid biomolecules such as proteins. To this end, a two-stage approach is adopted where the biomolecules are first optimised with respect to each other and afterwards their selectivity is tested in the presence of other molecules. Correlations lead to different optimum characteristic lengths of the hydrophobic and polar patches for the design of the two biomolecules on the one hand and their selectivity in the presence of other molecules on the other hand.

1 Introduction

Biomolecular recognition, that is the ability of a biomolecule to interact specifically with another molecule in an heterogeneous environment of structurally similar rival molecules, is an essential component in biological systems. The recognition process between two molecules is governed by a complicated interplay of non-covalent interactions of strengths comparable to the thermal energy[1]. This implies that the study of idealised models with methods from statistical physics might lead to valuable insight into this problem.

2 Model and General Approach

In this work we consider protein-protein recognition on a coarse-grained level in the framework of idealised models. The biomolecules are assumed to undergo no refolding during the association process which is a justified assumption for most protein-protein recognition processes[1]. Motivated by the observation that hydrophobicity is the major driving force in molecular recognition[1] we describe the type of the residue at the position $i = 1, \ldots, N$ of the interface by a binary variable $\sigma_i \in \{\pm 1\}$ for the target molecule and by $\theta_i \in \{\pm 1\}$ for the interaction partner[2]. We then model the energetics at the two-dimensional interface by

$$\mathcal{H}(\sigma, \theta; S) = -\varepsilon \sum_{i=1}^{N} \frac{1 + S_i}{2} \sigma_i \theta_i \tag{1}$$

as a direct contact interaction of strength ε. The variable S_i takes on the two values ± 1 and describes the local fit of the shape of the molecules at the interface resulting from a rearrangement of the amino acid side chains when the complex is formed[1].

To study the recognition process between two rigid biomolecules we adopt a two-stage approach. For a fixed target sequence $\sigma^{(\mathrm{t})} = (\sigma_1^{(\mathrm{t})}, \ldots, \sigma_N^{(\mathrm{t})})$ we first design an ensemble of probe molecules θ at a design temperature $1/\beta_{\mathrm{D}}$ leading to the distribution $P(\theta|\sigma^{(\mathrm{t})}) = \frac{1}{Z_{\mathrm{D}}} \sum_S \exp\left(-\beta_{\mathrm{D}} \mathcal{H}(\sigma^{(\mathrm{t})}, \theta; S)\right)$. In a second step the free energy difference of association at temperature $1/\beta$ is calculated for the interaction of the probe ensemble with the target molecule $\sigma^{(\mathrm{t})}$ and a structurally different rival molecule $\sigma^{(\mathrm{r})}$. In this step the free energy of

the interaction of the molecule $\sigma^{(\alpha)}$ with a particular probe sequence θ has to be averaged with respect to the distribution $P(\theta|\sigma^{(\mathrm{t})})$ giving finally the selectivity $\Delta F = F_{\mathrm{target}} - F_{\mathrm{rival}}$. A negative ΔF then signals recognition of the target.

For the majority of real protein-protein complexes the appearance of extended but fairly small patches of residues of the same type has been reported[3]. Biomolecular binding seems also to be strongly influenced by small-scale structures[4]. We therefore consider molecules which have correlated recognition sites at the interface with extended patches of residues of the same type. This can be incorporated into our model by adding additional correlation terms like

$$\mathcal{H}_{\mathrm{cor}} = -\gamma \sum_{\langle i,j \rangle} \theta_i \theta_j - \mu \sum_i \theta_i. \tag{2}$$

to the Hamiltonian of the system. The correlation parameters γ and μ (for the different types of molecules) are then used to fix the hydrophobicity and correlation length on the recognition sites. The average extension of the patches of residues from the same class is used as a measure for the correlation length of the finite system. Introducing the complementarity parameter $K = \sum_i \sigma_i \theta_i$ which measures the structural fit of the two biomolecules σ and θ at the interface, the selectivity averaged over all targets and rivals with the same correlation properties turns out to be the negative difference between the averaged complementarity with the target and the rival molecules, respectively[5].

3 Results

In this section molecular recognition for target and rival molecules with a fixed average hydrophobicity $h = 0.4$ per residue and a fixed (to length unity normalised) correlation length $\lambda = 0.263$ is considered within the model discussed above. The selectivity is studied as a function of the correlation length of the recognition site of the probe molecules. Figure 1

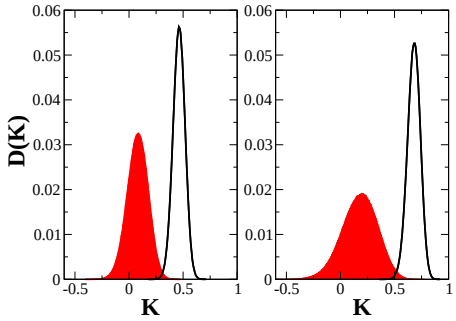

Figure 1. Distribution of the complementarity of the probe molecules with the target molecules (solid curve) and the rival molecules (shaded curve) for uncorrelated (left) and correlated probe molecules (with correlation length $\lambda_{\mathrm{p}} = 0.25$, right).

shows the distributions of the complementarities. For uncorrelated probe molecules the distribution for the complementarity with the target molecules is clearly separated from

the one with the rival molecules and shifted to larger values. This indicates the overall recognition ability of the system. A moderate increase of the correlation length on the probe molecules shifts the distribution to larger values of the complementarity so that an increased selectivity is expected. The first moments of the distributions are shown in figure 2. The average complementarity of the probe molecules with the target is always larger than that of the probe molecules with the rival. In the extreme cases where the correlation length tends to the minimum and maximum possible values the two averages become identical indicating that selectivity is lost as the probe molecules are not structured any more with respect to a particular molecule. The selectivity as shown in figure 2 has an optimum at a correlation length that is shifted below the value corresponding to the optimum of the complementarity with the target molecules. A smaller correlation length implies the appearance of an increased number of smaller patches on the recognition site of the probe molecule and hence an entropic profit for the interaction with the target due to more possible ways to align each other favourably. This effect only influences the contributions from the target-probe interactions due to the optimisation during the design step. The rival-probe interaction is not optimised and hence it is insensitive to a matching of structure elements.

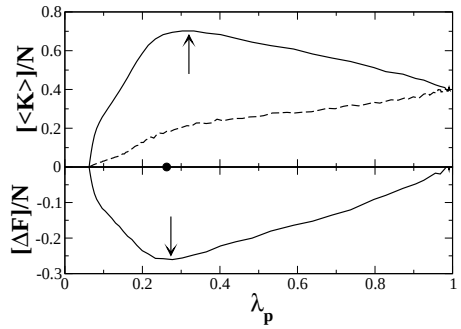

Figure 2. Upper part: Complementarity with the target molecules (solid curve) and the rival molecules (dashed curve) as a function of the (normalised) correlation length of the probe molecules (the fixed correlation length of the target and rival molecules is shown by the circle). Lower part: Resulting selectivity of the system.

Acknowledgments

We thank the Deutsche Forschungsgemeinschaft (SFB 613) for financial support.

References

1. C. Kleanthous, ed., *Protein-Protein Recognition* (Oxford University Press, Oxford, 2000).
2. H. Behringer, A. Degenhard, and F. Schmid, Phys. Rev. Lett. **97**, 128101, 2006.
3. T. A. Larsen, A. J. Olsen, and D. S. Goodsell, Structure **6**, 421, 1998.
4. T. Bogner, A. Degenhard, and F. Schmid, Phys. Rev. Lett. **93**, 268108, 2004.
5. H. Behringer and F. Schmid, in preparation.