VIP Protein–Protein Interactions

DOI: 10.1002/anie.201208683

Directional Selection Precedes Conformational Selection in Ubiquitin–UIM Binding**

*Dong Long and Rafael Brüschweiler**

In recent years, the mechanistic description of proteins under physiological conditions has evolved from a static view of their average structures[1] to an energy landscape-based view of dynamic ensembles[2–6] that accounts for the intrinsic dynamic nature of protein molecules.[7–9] In the case of a protein–protein recognition event, the conformational ensembles of both binding partners are often skewed toward the ensembles in their bound state, a process known as conformational selection or population shift.[10,11] This process has been studied recently for ubiquitin for which realistic molecular ensembles of its free state have been determined by the combination of NMR data with computational modelling,[12,13] accelerated molecular dynamics simulations (AMD),[14] and extended MD.[15] Such ensembles were shown to reproduce the structural heterogeneity displayed by ubiquitin when bound to various different proteins, as found in crystal structures, in general support of a conformational selection mechanism.[10,11] During the approach of a ubiquitin-interacting motif (UIM),[16] which is a conserved short α-helical motif for ubiquitin recognition by Hrs protein, the gradual reshaping of the energy landscape of ubiquitin from its free state to the UIM-bound state was found to set in already at nanometer distance, using an energy-based reweighting method.[15]

Protein–protein recognition proceeds through several stages that cause the two proteins undergoing independent overall translational, rotational and internal motions to form a composite unit, whose parts move in a highly concerted fashion. To better understand the factors that govern the efficiency of this process, we address herein the following questions: 1) is there a directional preference of UIM with respect to ubiquitin over a certain distance range? 2) In addition, is there an orientational preference of UIM? 3) Which parts of ubiquitin dominate the UIM-induced population shift at the individual residue level?

In the molecular reference frame of ubiquitin, the position and orientation of UIM is uniquely specified by six parameters, r, θ, ϕ, α, β, γ, where the spherical coordinates (r, θ, ϕ) define the direction and distance of the center of UIM with respect to the center of the ubiquitin and the three Euler angles (α, β, γ) describe the orientation of UIM, defined by a rotation about its center when fixed at location (r, θ, ϕ) (see the Supporting Information, Figure S1 for an illustration of the six parameters). Although this parameterization depends on the choice of the molecule used as the reference system, which is ubiquitin herein, analogous conclusions can be obtained when choosing UIM as the system of reference. In the presence of an intermolecular interaction V_{inter}, the probability density $p_{\text{dist}}(\theta,\phi)$ of a given direction (θ, ϕ) of UIM with respect to ubiquitin at a fixed distance r follows the Boltzmann relationship, as shown in Equation (1):

$$p_{\text{dist}}(\theta,\varphi) \propto \int_{-\pi}^{\pi}\int_{0}^{\pi}\int_{-\pi}^{\pi} e^{-V_{\text{inter}}(r,\theta,\varphi,\alpha,\beta,\gamma)/k_{\text{B}}T} d\alpha \sin(\beta) d\beta d\gamma \quad (1)$$

where k_{B} is the Boltzmann constant and T is the absolute temperature (see the Supporting Information for the complete expression of [Eq. (1)]). Equation (2) is the electrostatic potential between ubiquitin and UIM, dominating the intermolecular interaction energy at large distances.

$$V_{\text{inter}} = \sum_{i,j} f q_i q_j / (\varepsilon_r r_{ij}) \quad (2)$$

q_i and q_j are the partial charges of individual atoms of ubiquitin and UIM, respectively, and r_{ij} is the corresponding distance between the two atoms. f is the electric conversion factor and ε_r is the dielectric constant (see the Supporting Information for further details). The integral in Equation (1) represents a Boltzmann-weighted average over all possible internal orientations of UIM at the fixed distance r and direction (θ, ϕ). Figure 1 depicts $p_{\text{dist}}(\theta,\phi)$ of Equation (1) by color-coded spheres, showing the directional preferences at a range of distances. At distances $r\rightarrow\infty$, intermolecular interactions essentially vanish and all directions (θ, ϕ) will assume equal probability. On the other hand, when UIM approaches ubiquitin within approximately 10 nm distance, the intermolecular electrostatic interaction energy V_{inter} between UIM and ubiquitin becomes non-negligible compared to $k_{\text{B}}T$ and starts to bias the directional distribution of UIM. As a consequence, the directional distribution of UIM relative to ubiquitin becomes more and more anisotropic, as can be seen in Figure 1 and Figure 2 for the distance range from 15 nm to 4.5 nm.

The extent of the directional anisotropy can be quantified by the information-theory-based collectivity measure κ or the geometric anisotropy measure σ (see the Supporting Information). κ ([Eqs. (S5–S7)]), which can take values between 0 (maximal anisotropy) and 100% (isotropic distribution), is

[*] Dr. D. Long, Prof. Dr. R. Brüschweiler
Department of Chemistry and Biochemistry & National High Magnetic Field Laboratory, Florida State University
Tallahassee, FL 32306 (USA)
E-mail: bruschweiler@magnet.fsu.edu
Homepage: http://spin.magnet.fsu.edu

[**] This work was supported by the National Science Foundation (grant MCB-0918362).

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201208683.

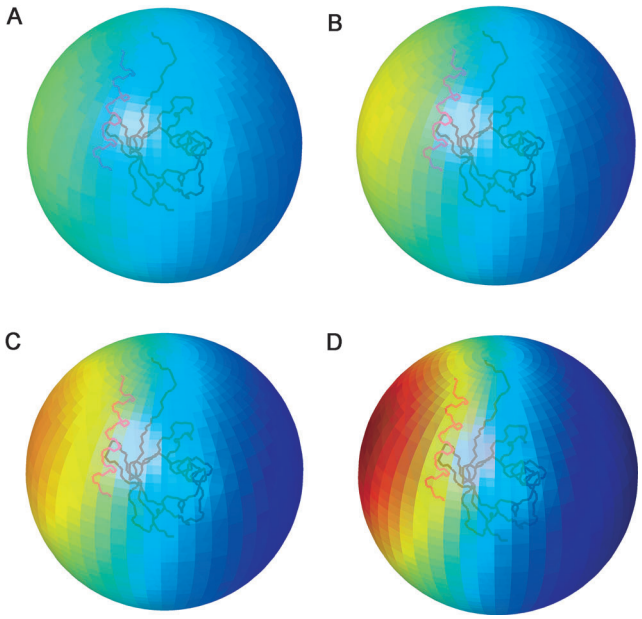

Figure 1. Probability densities [Eq. (1)] of the directional distribution of the ligand UIM with respect to ubiquitin at distances A) $r = 8$ nm, B) 6.5 nm, C) 5.5 nm, and D) 4.5 nm. The backbone traces of ubiquitin (black; PDB ID: 1UBQ) and UIM (magenta; PDB ID: 2D3G), shown as solid lines, indicate the relative positions of ubiquitin and UIM in the crystal structure of the complex (PDB ID: 2D3G). Dark blue on the sphere represents directions with the lowest probability density (p_{min}) and dark red with maximal probability density (p_{max}), where $p_{max}/p_{min} = 7.6$.

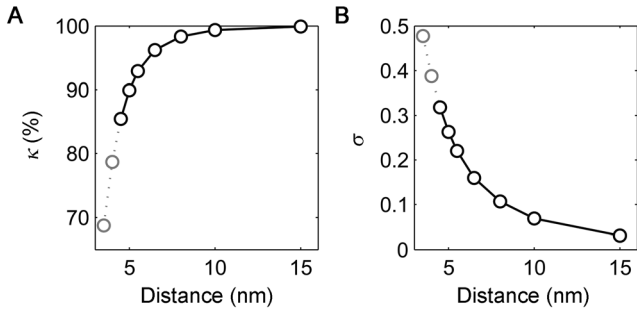

Figure 2. Directional anisotropy described by parameters A) κ [Eq. (S5)] and B) σ [Eq. (S4)] as a function of the UIM–ubiquitin distance r. For $r < 4.5$ nm, the simplified form of the interaction energy [Eq. (S1)] is no longer accurate, and the corresponding σ and κ values (gray dotted lines) are given for qualitative purposes to show the trend of the change.

the effective percentage of directions sampled by UIM relative to ubiquitin and σ (see [Eq. (S4)]) is the length of a vector representing the weighted average of all UIM directions, each represented by a unit vector; $\sigma = 0$ if the directional distribution is isotropic and $\sigma = 1$ if it is fully anisotropic. As shown in Figure 2, σ increases from 0.03 at 15 nm to 0.32 at 4.5 nm, while κ drops from 99.86% to 85.54%, over the same distance range.

Directional selection of UIM is thus a long-range process that starts to be noticeable when the ligand is as far as 10 nm away, and it becomes significant when UIM and ubiquitin are separated by 5 nm or less. Remarkably, precisely those directions are favored by the intermolecular interactions that face the native binding interface of ubiquitin. Figure 1 shows that the associated "directional funnel" is very smooth, facilitating ligand approach from those directions that are most productive for correct binding to the binding interface of ubiquitin. The probability distribution resembles the one of a systematically scaled charge-dipole (Figure S3,S4). Because at 4.5 nm distance the population shift of internal conformations of ubiquitin is still marginal,[15] the directional selection process represents the starting point of the ligand-induced population shift by preceding the internal conformational selection, eventually leading to the final complex.

Not only the direction of UIM, but also its orientation (α, β, γ) may affect the intermolecular interaction energy V_{inter} and thereby gives rise to a non-uniform probability density with respect to (α, β, γ). To illustrate the dependence of V_{inter} on the orientation of UIM, we rotated UIM independently about the three axes (Figure S2). For this purpose, the center of UIM was placed along the most favorable direction facing the binding interface of ubiquitin at the fixed distance $r =$ 4.5 nm, followed by rotations of UIM about one axis at a time. The resulting probability distributions $p(\xi)$ show very little variation for rotations about axis 1, which connects the centers of ubquitin and UIM, and axis 2, which is the axis of the α-helix of UIM. A slightly larger variation is observed for the rotation about axis 3 that is perpendicular to both axis 1 and axis 2. The magnitudes of changes are $p_{max}/p_{min} = 1.01$, 1.14, and 1.62 for rotations about axes 1, 2, and 3, respectively, which however are much smaller than the directional effect $(p(\theta,\phi)_{max}/p(\theta,\phi)_{min} = 7.6)$ at this distance ($r = 4.5$ nm). This suggests that, the effect of the orientational selection is marginal at this distance range, while at the same time, directional selection of UIM plays a significant role during these early stages of molecular recognition. Orientational selection of UIM, however, is expected to become important at shorter distances, that is, at later stages of the recognition process, when interactions other than electrostatic interactions, such as van der Waals interactions, become important.

It was previously observed that the presence of UIM at nanometer distances from the binding interface could effectively induce a conformational population shift toward the ensemble of the final bound state of ubiquitin.[15] This behavior was found to be robust with respect to the precise direction of the ligand as long as it faces the binding interface of ubiquitin.[15] However, the contributions of individual protein residues to this conformational population shift are unknown. For this purpose, we delineated the residue-specific contributions to the internal population shift by placing UIM at 12 different positions facing the ubiquitin binding interface (Figure S5). To evaluate the contribution of each individual residue to the conformational population shift along the pincer-like motional coordinate x, we calculated the Kullback–Leibler divergence (D_{KL}),[17] which is a general measure for the difference between two distributions in Equation (3):

$$D_{KL,k} = \int_{-\infty}^{\infty} p_{all}(x) \log\left[\frac{p_{all}(x)}{p_k(x)}\right] dx \qquad (3)$$

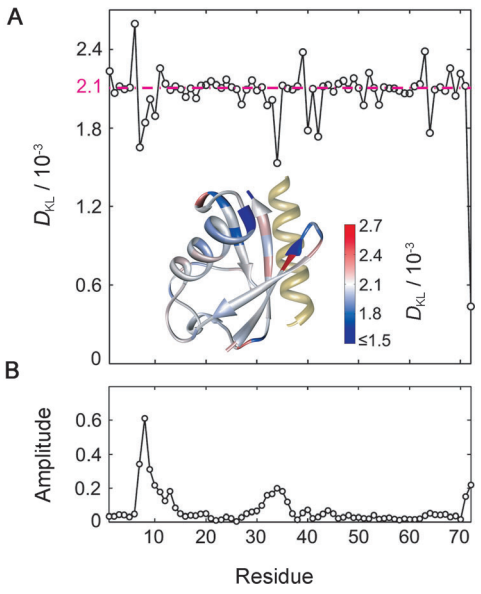

Figure 3. A) Delineation of residue-specific contributions to the conformational population shift of ubiquitin for a distance $r = 3.4$ nm between the centers of ubiquitin and UIM and a direction of UIM facing the binding interface of ubiquitin (see Figure S5 A). The Kullback–Leibler divergence $D_{KL,k}$ was calculated for each residue according to Equation (3). The magenta dashed line represents the D_{KL} reference value for a non-interacting, neutral residue. Inset: ribbon diagram of ubiquitin (PDB ID: 2D3G, chain A) color-coded by amino-acid specific $D_{KL,k}$ values. The yellow helix represents UIM in the crystal structure (2D3G, chain P). B) Relative fluctuation amplitudes of Cα atoms along the largest principal component mode.

where $p_{all}(x)$ is the probability distribution of the conformational ensemble reweighted according to the interaction between UIM and the entire ubiquitin (except for the tail residues 73–76) and $p_k(x)$ is the corresponding distribution when only the contribution of a specific residue k of ubiquitin is included. x is the coordinate along the largest principal component of the free ubiquitin ensemble (see the Supporting Information for details). As shown in Figure 3, residues with a $D_{KL,k}$ below the reference value 2.1×10^{-3} for non-interacting residues (defined in [Eq. (S17)]), show positive energetic contributions to the overall population shift pattern. It is interesting to note that Glu34, which belongs to helix α1 and makes a large favorable contribution to the recognition, is involved in the pincer-like motion, but it is not part of the binding interface. Other residues, such as Lys11, with $D_{KL} > 2.1 \times 10^{-3}$ contribute negatively to the overall population shift. Indeed, we find that the amplitude of the overall population shift may be decreased, owing to the significant opposing effect of Lys11 along certain directions (Figure S5 C). While the effect of individual residues on the net population shift varies as a function of the relative position of UIM, the general trends of the contributing residues, as well as their cumulative effect, are conserved (Figure S5).

In conclusion, the early-stage interplay between ubiquitin and UIM is dominated by long-range electrostatic interactions at 5–10 nanometer distances. The first effect in a ubiquitin-UIM encounter is directional selection, which guides UIM toward the binding interface of ubiquitin as the two molecules approach each other, enabling the subsequent internal conformational selection process. Interestingly, at this distance range the orientational effect of UIM is notably weak and the interplay is dominated by the directional effect. The orientational selection of UIM is expected to be controlled by van der Waals interactions in the late stages of the binding process. A detailed analysis of the role of individual residues for the internal conformational population shift of ubiquitin revealed that the residues whose positions are modulated by the pincer-like motion of ubiquitin contribute the largest energetic contributions, which can be both positive and negative. They also include residues that are not part of the binding interface. These results suggest that binding of UIM to ubiquitin follows, on average, a specific sequence of stochastic events starting with directional selection, followed by rotational selection accompanied by a conformational population shift. Together these processes promote the formation of the final complex of this model system. It is expected that complex formation between ubiquitin and other proteins also follows a sequence of selection events that are specific to the charge distributions, shapes, and flexibilities of the binding partners. Together these factors determine the overall efficiency of the protein–protein binding process.

Received: October 29, 2012
Revised: January 15, 2013
Published online: February 20, 2013

Keywords: computational chemistry · directional selection · long-range effects · protein-protein interactions · ubiquitin

[1] E. Fischer, *Ber. Dtsch. Chem. Ges.* **1894**, *27*, 2985–2993.
[2] H. Frauenfelder, S. G. Sligar, P. G. Wolynes, *Science* **1991**, *254*, 1598–1603.
[3] C. J. Tsai, B. Ma, R. Nussinov, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 9970–9972.
[4] P. Csermely, R. Palotai, R. Nussinov, *Trends Biochem. Sci.* **2010**, *35*, 539–546.
[5] G. Kar, O. Keskin, A. Gursoy, R. Nussinov, *Curr. Opin. Pharmacol.* **2010**, *10*, 715–722.
[6] R. B. Fenwick, S. Esteban-Martin, X. Salvatella, *Eur. Biophys. J.* **2011**, *40*, 1339–1355.
[7] D. D. Boehr, R. Nussinov, P. E. Wright, *Nat. Chem. Biol.* **2009**, *5*, 789–796.
[8] K. Henzler-Wildman, D. Kern, *Nature* **2007**, *450*, 964–972.
[9] S. R. Tzeng, C. G. Kalodimos, *Nature* **2012**, *488*, 236–240.
[10] B. Ma, S. Kumar, C. J. Tsai, R. Nussinov, *Protein Eng.* **1999**, *12*, 713–720.
[11] C. J. Tsai, S. Kumar, B. Ma, R. Nussinov, *Protein Sci.* **1999**, *8*, 1181–1190.
[12] O. F. Lange, N. A. Lakomek, C. Fares, G. F. Schroder, K. F. Walter, S. Becker, J. Meiler, H. Grubmuller, C. Griesinger, B. L. de Groot, *Science* **2008**, *320*, 1471–1475.
[13] R. B. Fenwick, S. Esteban-Martin, B. Richter, D. Lee, K. F. Walter, D. Milovanovic, S. Becker, N. A. Lakomek, C. Griesinger, X. Salvatella, *J. Am. Chem. Soc.* **2011**, *133*, 10336–10339.
[14] P. R. Markwick, G. Bouvignies, L. Salmon, J. A. McCammon, M. Nilges, M. Blackledge, *J. Am. Chem. Soc.* **2009**, *131*, 16968–16975.
[15] D. Long, R. Brüschweiler, *PLoS Comput. Biol.* **2011**, *7*, e1002035.
[16] S. Hirano, M. Kawasaki, H. Ura, R. Kato, C. Raiborg, H. Stenmark, S. Wakatsuki, *Nat. Struct. Mol. Biol.* **2006**, *13*, 272–277.
[17] S. Kullback, R. A. Leibler, *Ann. Math. Stat.* **1951**, *22*, 79–86.

