Counterbalance of ligand- and self-coupled motions characterizes multispecificity of Ubiquitin

Bhaskar Dasgupta,* Haruki Nakamura, and Akira R Kinjo

Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan

Received 26 July 2012; Revised 15 October 2012; Accepted 9 November 2012

DOI: 10.1002/pro.2195

Published online 20 November 2012 proteinscience.org

Abstract: Date hub proteins are a type of proteins that show multispecificity in a time-dependent manner. To understand dynamic aspects of such multispecificity we studied Ubiquitin as a typical example of a date hub protein. Here we analyzed 9 biologically relevant Ubiquitin-protein (ligand) heterodimer structures by using normal mode analysis based on an elastic network model. Our result showed that the self-coupled motion of Ubiquitin in the complex, rather than its ligand-coupled motion, is similar to the motion of Ubiquitin in the unbound condition. The ligand-coupled motions are correlated to the conformational change between the unbound and bound conditions of Ubiquitin. Moreover, ligand-coupled motions favor the formation of the bound states, due to its in-phase movements of the contacting atoms at the interface. The self-coupled motions at the interface indicated loss of conformational entropy due to binding. Therefore, such motions disfavor the formation of the bound state. We observed that the ligand-coupled motions are embedded in the motions of unbound Ubiquitin. In conclusion, multispecificity of Ubiquitin can be characterized by an intricate balance of the ligand- and self-coupled motions, both of which are embedded in the motions of the unbound form.

Keywords: date hub protein; Ubiquitin; ligand multispecificity; normal mode analysis; coupling of receptor-ligand motions; functional modes in unbound condition

Introduction

Proteins perform their biological functions through specific interactions with other proteins. To do this, one protein (receptor) must recognize its partner protein (ligand) out of all other molecules present in that cellular environment.¹ The specificity in such a recognition process is mediated by the protein-protein interaction interfaces.^{2,3} Doubtlessly, both static and dynamic properties of protein-protein interface are essential to specific binding.³ Although there has been already a large number of studies aimed at the

systematic analyses of static snapshots of the protein-protein interfaces,^{2,4,5} a growing number of studies indicate dynamics play a key role in mediating protein-protein interactions.⁶⁻⁹

A high specificity in protein-protein recognition is observed in general, nevertheless there are many proteins that bind to more than one partners. Such promiscuous protein-protein recognition leads to the multifunctional property of the receptor protein,¹⁰ and helps to make protein-protein interaction network coordinated. The promiscuous proteins are often termed as hub proteins^{11,12}; and depending on the temporal nature of interactions they are classified into date hub proteins (asynchronous binding) or party hub proteins (synchronous binding).¹² In many cases, the promiscuity of date hub proteins is characterized by transient interactions, and they bind their ligands by using an overlapping region of the interface.¹³⁻¹⁵ Such an overlapping region selects

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: JSPS Grant-in-Aid for Scientific Research B; Grant number: 23370071.

*Correspondence to: Bhaskar Dasgupta, Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan. E-mail: bhaskar@protein.osaka-u.ac.jp

Figure 1. (a) The superimposition of nine Ubiquitin chains in holo conditions to the apo condition (in yellow), shown in tube representations. The Ubiquitin chains in holo conditions include definitions of the OV (in blue) and Non-OV (in red) regions as obtained from our previous analysis.¹⁶ The OV and Non-OV regions constitute multiligand interface and the corresponding noninterface region is shown in gray. The most-essential surface patch of Ubiquitin comprising of Leu8, Ile44, and Val70 residues are included in the OV region,³⁴ whereas second-most essential residue cluster (including Gln2, Phe4, and Thr12) and polar patch (including Thr55, Asp58, Asn60)³⁷ are included in the Non-OV region. (b) Secondary structure assignment⁵⁵ of Ubiquitin. In top of the sequence, secondary structures were further annotated and numbered. The coil region occurring between the “H1” and “H2” helices is annotated as “C”

multiple ligands out of all other molecules present in that cellular environment. However, at a given moment the formation of the specific receptor-ligand complex is achieved by simultaneous use of the overlapping region and a particular nonoverlapping region. For example, Ubiquitin as a receptor form a large number of specific complexes with other proteins sharing an identical common binding region (in terms of sequence).¹⁶ To demonstrate such a phenomenon of ligand recognition, previously we defined a multiligand interface by integrating all receptor-ligand interfaces on the receptor that shared a common binding region. Moreover, we defined an overlapping region ("OV") and a nonoverlapping ("Non-OV") region of a multiligand interface on the basis of promiscuity of interface sites [Fig. 1(a), left]. We observed that the OV and Non-OV regions are different in terms of residue and atomic burial, side-chain torsion angle diversity, residue and secondary structure compositions. In this work, we analyzed dynamic characteristics of the OV and Non-OV regions that are coupled to the specificity and selectivity¹⁶ of the ligand recognition. As in the previous study we defined selectivity as the ability to recog-

nize multiple ligands out of nonligands, and specificity as the ability to bind one ligand out of all other ligands. An analysis of differential dynamics between the OV and Non-OV regions would show how a receptor molecule recognizes its ligands using multiligand interface. The dynamics of multiligand interface is influenced by global dynamics of the receptor–ligand complex. However, in this study we focus on interfacial dynamics, because the global dynamics is implicitly considered due to the collectivity of motion.

A previous bioinformatic analysis of the static structures of Ubiquitin in different bound (holo) and unbound (apo) conditions revealed many essential variabilities of binding regions.¹⁷ However, how promiscuous yet specific interaction is related to the ligand-dependent dynamics of the receptor is still unclear. One approach to understand conformational changes due to the binding is to perform molecular dynamics simulations.¹⁸ On the other hand, it is already established that such a conformational change can also be described by the low-frequency vibrational modes derived from the normal mode analysis (NMA).^{19,20} It has been found that such soft modes match well with the essential modes obtained from the molecular dynamics calculations,^{21,22} and they help describe functionally relevant motions.^{19,23,24} Therefore, it is possible to obtain dynamic characteristics of the binding process by analyzing the low-frequency modes of the receptor in its apo and holo conditions.^{25,26} In the current analysis, we related ligand-dependent dynamics of the receptor to its ligand-free dynamics. For the sake of fast and robust computation, we employed an all-atom elastic network model (ENM), where a protein structure is assumed to be composed of coupled harmonic oscillators with its native conformation being the ground state.²⁷

From a protein-protein complex including N atoms NMA yields $3N-6$ vibrational modes by keeping the whole complex in its Eckart frame.²⁸ However, in this case the constituent receptor (including N^R atoms) and ligand (including N^L atoms) molecules are not in their Eckart frame. Therefore, total motion of the receptor in bound condition (included in $3N^R$ normal modes) includes internal vibrations of the receptor (included in $3N^R-6$ normal modes) as well as external motions (6 modes) of the receptor with respect to the ligand. In this study we are interested in how motions of the receptor are influenced by the motion of the ligand (Fig. S1 in the Supporting Information). To study internal motions of the receptor an accepted strategy is to calculate effective force-constant matrix of the receptor obtained from the original force-constant matrix by a first-order perturbation due to the ligand motion.^{24,29} However, to understand the dynamics in more detail we were not only interested in the internal motions of the receptor, but also to the combinations of internal and external motions of the receptor

and ligand. In this analysis we employed a method similar to that of Ishida *et al.*,³⁰ who separated the external and internal modes of the interacting subunits. Primarily, we analyzed the coupling of internal motion of the receptor to the internal motion of the ligand and also the coupling of total motion of the receptor to the total motion of the ligand.

Here, we took Ubiquitin as a model case for the date hub protein (Fig. 1) because it was found to be most promiscuous among the date hub proteins¹⁶ available in the Protein Data Bank (PDB).³¹ Ubiquitin is a small regulatory protein with the β -grasp fold.³² It is ubiquitously found in all tissues of eukaryotes and labels other proteins for degradation. The sequence of Ubiquitin is highly conserved between human and yeast (96%).³³ Moreover, recent studies also confirmed conformational changes associated to Ubiquitin recognition.^{17,18} A surface patch of Ubiquitin residues (Ile44-surface) was found to be related to the specific binding to Ubiquitin.³⁴ Here, we studied how the ligand-dependent dynamics of Ubiquitin realizes its multispecific characteristics by using its OV and Non-OV regions.

Results

Data set of protein–protein complexes and normal mode analysis

To understand dynamic aspect of promiscuous binding of Ubiquitin (Supporting Information Fig. S1) to various types of ligands we analyzed 9 Ubiquitin complexes (Fig. 1, Supporting Information Table S2) (see also Materials and Methods).¹⁶ The dynamics of Ubiquitin were probed by performing all-atom normal mode analysis (AA-NMA) of an ENM. The AA-NMA of the complex yielded the couplings between total motions (referred to as total coupling) of Ubiquitin and the ligand. The subsequent separation of internal and external motions (Supporting Information Fig. S1) provided the information regarding couplings between all combinations of internal and external motions (referred to as partial couplings in the following text). To study such coupled motions we analyzed by singular value decomposition (SVD) the rectangular part of the covariance matrices that represents receptor–ligand coupling obtained by AA-NMA.

Distinct roles of the OV and Non-OV regions regarding ligand-coupled and self-coupled vibrations of Ubiquitin

The motion of Ubiquitin in a complex can be interpreted as a combination of its ligand-coupled (from the “RL” part, S1) and self-coupled motions (from the “RR” part). From the ligand-coupled and self-coupled motions we extracted only their vibrational parts by the separation of internal and external motions. Here, we compare the differential characteristics of the ligand-coupled and self-coupled vibrations in the

selectivity-determining OV and specificity-determining Non-OV regions. To understand how internal motions of different ligands influenced the internal motions of Ubiquitin, we analyzed the left singular vectors of the rectangular covariance matrices (“RL” part, Supporting Information Fig. S1) that represent Ubiquitin–ligand internal–internal coupled motions for nine complexes. Each Ubiquitin singular vector can be interpreted as a set of initial velocities of Ubiquitin atoms (Supporting Information S3d). On average the atomic velocity is higher in the OV region than in the Non-OV region (and in the multiligand interface than in the surface) (Supporting Information Fig. S4). To compare the diversities of the atomic motions in the OV and Non-OV regions we defined complex-averaged order parameter, which is the sum of the normalized atomic initial velocities in the OV and Non-OV regions averaged over complexes (Equation 14 in the Supporting Information S3f). Such complex-averaged order parameter represents collectivity of motion in OV and Non-OV regions. We observed that the complex-averaged order parameter in the OV region was smaller than that in the Non-OV region [Fig. 2(a)]. This indicates that in a receptor–ligand complex the motion of the atoms in the Non-OV region are more directed. Moreover, more random motion in the OV region indicated that the flexibility of this region was greater than that of the Non-OV region. This result is consistent with our previous analysis showing that the conformational variability in the OV region was greater than in the Non-OV region.¹⁶

On the other hand, the self-coupled motion of the receptor exhibited different characteristics compared to the ligand-coupled motion. In the self-coupled motion the OV region showed more directed motion than the Non-OV region [Fig. 2(b)]. We observed that on average the characteristics of the self-coupled motion, rather than the characteristics of the ligand-coupled motion, in the OV and Non-OV regions were similar to the motion of the Ubiquitin in the unbound condition (Supporting Information Fig. S5c). Namely, the more directed self-coupled motion of the OV region is similar to the collective motion in the unbound condition. Moreover, we observed that on average the mean square fluctuation (MSF) of the atoms at the multiligand interface obtained from the ligand-constrained self-coupled total motion of the receptor ($1.34 \text{ \AA}^2$) is significantly less than that in the unbound condition ($4.21 \text{ \AA}^2$, P -value 7.2×10^{-125} , Fig. 3). Such attenuation of atomic motions (Supporting Information Fig. S6) indicate a loss in conformational entropy due to binding at the interface. This also implies that self-coupled motions disfavor the complex formation and may be related to the formation of transient complexes by Ubiquitin.

Figure 2. The diversity in atomic displacements in OV (in blue) and Non-OV (in red) regions defined from complex-averaged order parameters from different regions. (a) First 50 internal-internal ligand-coupled motions of the receptor were considered, and (b) First 200 self-coupled motion of the receptor were considered. The errorbars were calculated by bootstrapping.

Ligand-coupled vibrations of Ubiquitin correlated to conformational change upon binding

In a complex, the opposite directional characteristics of motions in the OV and Non-OV regions for the self-coupled and ligand-coupled motions of Ubiquitin (Fig. 2) would counterbalance each other. The optimized directional characteristics of motions in the OV and Non-OV regions were embedded in the conformational change between the unbound and the bound Ubiquitin structures. We superimposed bound Ubiquitin structures to its unbound counterpart (Fig. 1) and defined atomic site-specific order parameters from the displacement of the equivalent atomic sites between the unbound and bound structures (equation 16 in Supporting Information S3f). Such atomic site-specific order parameter represents diversity of motion of an atom due to the conformational change. We compared the atomic site-specific order parameters averaged over the atoms in OV (0.62) and Non-OV (0.70) regions (P -value 5.82×10^{-6}) (equation 18 in Supporting Information S3f). This concretely reflects that the OV region is more flexible than the Non-OV region, which agrees with our previous analysis.¹⁶

The Ubiquitin atoms with atomic site-specific order parameters higher or lower than a threshold

Figure 3. MSF of $C\alpha$ -atoms of Ubiquitin residues in apo and holo condition. MSF in holo condition describing self-coupled motion was calculated from the average MSF from covariance matrices ("RR" part) obtained from AA-NMA of 9 heterodimers present in the data set. Large suppression of self-coupled motion was observed in the region of the residues 8–12 [part of loop L1, see Fig. 1(b)], 46–49, 57–67. The residues 8–12 are flexible in both apo and holo conditions (and in different conformers of apo Ubiquitin, see Supporting Information Fig. S5), and may be they are responsible for selecting a set of ligands.

Table I. Number of atomic sites showing coherent and incoherent motions associated to different regions^a

	Number of atomic sites in the OV region	Number of atomic sites in the Non-OV region	Total number of atomic sites showing the coherent or incoherent motions
Number of atomic sites showing coherent motion	19	70	89 (109 ^b)
Number of atomic sites showing incoherent motion	39	42	81 (88)
Total number of atomic sites in the OV or Non-OV region	58	112	Total number of atomic sites = 170 (197)

^a Here, the numbers of atomic sites in the OV and Non-OV regions included atomic sites that showed the coherent and incoherent motions. The remaining atomic sites were not considered in the calculation of probabilities (main text).

^b Values in the parentheses indicate total number of atom-sites including noninterface atoms.

(from the average atomic site-specific order parameters over all the atoms, 0.68 ± 0.18) would show coherent (i.e., directed) or incoherent (i.e., undirected) motions, respectively. We identified Ubiquitin atoms that were responsible for coherent and incoherent motions. 109 atomic sites showed coherent motions (average order parameter 0.91 ± 0.03) and 88 atomic sites showed incoherent motions (average order parameter 0.38 ± 0.12) (Table I, Supporting Information Fig. S7 and Table S8). From Table I, we obtained that the probability of an atomic site showing coherent motion ($70/112 = 0.625$) was higher than that showing incoherent motion ($42/112 = 0.375$) given that the atomic site was included in the Non-OV region. On the contrary, we obtained that the probability of an atomic site showing incoherent motion ($39/58 = 0.67$) was higher than that showing coherent motion ($19/58 = 0.33$) given that the atomic site was included in the OV region. The above observations demonstrate the dominance of coherent and incoherent motions in the Non-OV and OV regions, respectively. This also indicates that the ligand-coupled motions were more correlated to the directions of atomic motions than the self-coupled motions [Fig. 2(a,b)].

Correlation of motion at the interface

To achieve high specificity of interactions, the interface atoms defined from the intersubunit atomic contacts ($<5 \text{ \AA}$)³⁵ were expected to show complementary

motions. Intuitively, the complementarity of motion at the interface would favor the formation of the complex. The dynamic aspect of the complementary motions can be obtained from the correlations of motions of the interface atoms from Ubiquitin and the ligands. In principle, motion of receptor and ligand atoms that are in contact can be positively correlated (i.e., in-phase motions of the atoms) or they can be negatively correlated (i.e., out-of-phase motions of the atoms). In case of the in-phase motions the interacting atoms are vibrating like a single unit held by an attractive force, and this will avoid any colliding motion that would disrupt the complementarities.

We observed that when total coupling was considered almost all the receptor-ligand atom-pairs showed positive correlation (Table II). Moreover, when internal-internal, internal-external, and external-internal couplings were considered, a majority of the atom-pairs showed positive correlation. This is in accordance with a recent study by Tsuchiya *et al.* which showed that on average intersubunit contacts across homodimeric interfaces involved positive correlations.³⁶ It is previously demonstrated that the in-phase motions favor the formation of bound state.³⁰ Our result indicates that the partial couplings (except external-external coupling) favor the ligand binding process by showing complementary dynamics at the respective interfaces. The asynchronous binding of Ubiquitin to the ligands assumes that the complexes are transient. We

Table II. Correlation in atomic motions of the receptor-ligand contacting atoms averaged over 9 complexes^a

Partial couplings	Average positive correlation ^b	Average negative correlation ^b	Fraction of number of contacting atoms with positive correlation averaged over all complexes
Receptor-Ligand			
Total-Total	0.0492 (0.020)	-0.0008 (0.000)	1.0000 (0.001)
Internal-Internal	0.0133 (0.009)	-0.0046 (0.004)	0.8873 (0.103)
External-External	0.1299 (0.073)	-0.2458 (0.128)	0.3957 (0.269)
Internal-External	0.0643 (0.035)	-0.0071 (0.004)	0.9585 (0.050)
External-Internal	0.1044 (0.044)	-0.0086 (0.006)	0.9531 (0.074)

^a Values in the parentheses are standard deviations.

^b Average positive or negative correlations were defined as, $\text{Corr}_{+/-} = \frac{1}{N_c} \sum_c \frac{1}{N_c^{\text{cont}}} \sum_I^{\text{cont}} \text{Corr}_{c,+/-}^I$, where N_c is the number of complexes, N_c^{cont} is number of interface contacts in complex c , and $\text{Corr}_{c,+/-}^I$ is the positive or negative correlation for I -th contact in complex c .

Figure 4. Dynamics of apo-Ubiquitin due to the modes that are conserved in all the complexes. (a) 2nd apo-mode for Ubiquitin is shown with atomic displacements of $C\alpha$ -atoms as black lines describing the directions of motion. The blue and red colors refer to the OV and Non-OV region, respectively. (b) The frequency of occurrence of apo-modes that are embedded in total ligand-coupled motions of nine complexes. The frequency of occurrence of an apo-mode was obtained from the sets of apo-modes that showed maximum similarity with the ligand-coupled motions (see Supporting Information S3c). The apo-modes annotated at the top are conserved in all the complexes (Modes 2, 3, 7, and 8) or in 8 out of 9 complexes (Modes 35, 37, and 45).

observed that the external–external coupling shows out-of phase overlap for a considerable number of contacting atoms (Table II, Column 4). Therefore, external–external coupling counter-balances the in-phase and out-of-phase motions across the contacting atoms, and hence may influence the transient nature of binding. In summary, we conclude that the combinations of rigid-body and vibrational motions optimize specific interactions between Ubiquitin and the ligands across the interface.

Ligand-coupled motions are embedded in the dynamics of apo-Ubiquitin

Presently, a large number of analyses that compared the dynamics of apo and holo states of a protein indicated that the functionally relevant motion of a protein in the bound form can be demonstrated from the motion in the unbound form.^{25,26} Our result showing similarity between apo and ligand-constrained self-coupled motions is consistent with this. It may be pertinent to ask whether the ligand-coupled motions of Ubiquitin can be realized from the motions of the apo Ubiquitin (or apo-modes). For each complex showing different types of coupled motions we identified a set of apo-modes (Supporting Information S3c outlines the methods) that embed ligand-coupled motions (Supporting Information Table S9). From the first 50 lowest-frequency apo-modes we observed that four low frequency (modes of order less than 10) and three relatively higher frequency (modes of order greater than 10) apo-modes were conserved in all the nine complexes considering

total-total couplings. The above four low-frequency apo-modes are the 2, 3, 7, and 8th modes and the three relatively higher frequency modes are the 35, 37, and 45th modes [Fig. 4(b) and Supporting Information Fig. S10]. These conserved modes facilitate a motion in Ubiquitin such that it recognizes all the ligands in a similar way. This might lead to selectivity of Ubiquitin towards its ligands. However, it is not certain whether these conserved modes can discriminate between ligands and nonligands.

From the mapping between holo-modes and apo-modes of the Ubiquitin we observed that there were 10 total-total coupling modes that cannot be matched with any single apo-modes (Supporting Information Table S9). They are all low-frequency modes (order less than 5). The external motions of Ubiquitin compared to its internal vibrational modes largely influenced the motions in these modes. Apart from these 10 modes out of 270 analyzed, the mappings between the ligand-coupled vibrational modes and the apo-modes indicate that the ligand-coupled vibrations are embedded in the motion of the unbound Ubiquitin.

Functionally relevant dynamics of apo Ubiquitin

In the previous section, we identified conserved motions of apo-Ubiquitin that may help recognition of all the ligands. In these modes [Fig. 4(a) and Supporting Information Fig. S10] we observed that β -turns involving residues 7–10 [L1, Fig. 1(b)] and 45–48 (L3) and coil region between α -helix and 3_{10} -helix [residues 35–37 region, annotated as “C”, Fig. 1(b)]

were showing large motions. In the 2nd apo-mode the “pincer-like” motion of Ubiquitin involving L1 and L3 turns were accompanied by opening or closing of functional residues Ile44 and Val70, whereas in the 3rd apo-mode a large displacement of an antiparallel β -sheet (strand β 3 and β 4, together with L3 turn) together with the motion in L1 turn increases the accessibility of β 3-strand (Supporting Information Fig. S10). In relatively higher frequency conserved modes (35, 37, and 45th apo-modes, Supporting Information Fig. S10d–f) we observed local motions in residues included in L3 turn and residues Asp32, Asn60, and Gln62. Interestingly, the later three residues were in the Non-OV region. The presence of significant motions in Asn60 and Gln62 were consistent with a previous NMR analysis indicating significant perturbation of chemical-shift indexes for these residues.³⁷ The dynamics in the moderately higher-frequency modes may be related to the specific interactions involving atoms in the Non-OV region.

Discussion

The purpose of this analysis is to discern differential dynamics at the multiligand interface, because such motions are primarily related to the recognition of multiple ligands. To demonstrate such differential dynamics here we have used previously characterized two distinct regions of multiligand interface, namely OV and Non-OV regions.¹⁶ Previously we showed that the OV region is more flexible than the Non-OV region by comparing different bound structures of date hub proteins. In this study we have used harmonic dynamics to probe such differential flexibility for a model date hub protein, Ubiquitin. However, it is possible that such experimental structures span a very wide conformational space comprising of multiple minima, and therefore the dynamics is anharmonic. For Ubiquitin this has been already shown in a long-term molecular dynamics simulation by Long and Brüschweiler.¹⁸ The use of a multidimensional parabolic energy surface (as in this study) instead of actual rugged energy surface³⁸ can be validated in the following way. (1) It was observed that the low-frequency modes of NMA capture large-scale conformational change even in case of F1-ATPase,^{39,40} that show a remarkable conformational change.⁴¹ There are numerous other examples in which low-frequency normal modes agree well with conformational changes.^{20,21,25,26,42,43} (2) Ahmed *et al.*²¹ compared essential dynamics (ED) modes from molecular dynamics simulation and coarse-grained normal modes (C α -based). Their analysis showed good correlation between essential and normal modes for α - β CATH^{44,45} class of proteins, which include Ubiquitin. (3) In this study we have compared MSF of Ubiquitin in all pairs of the complexes. The average correlation coefficient was found

to be 0.68. When we have compared MSF of Ubiquitin in the bound forms to the unbound form the average correlation coefficient was 0.65. Such significant correlation coefficients confirm that it is possible to use NMA under ENM to find the general rules of dynamics in case of Ubiquitin.

Next, we considered the validity of using dynamics at the multiligand interface instead of global dynamics to understand specificity and selectivity. (1) Zen *et al.*⁴⁶ demonstrated the importance of interfacial dynamics from NMA of a data set of 22 structures. They observed that the flexibility at the interface is less than that at the surface. This is in accordance with our observation (Supporting Information Fig. S6). Moreover, we observed that the magnitude of atomic displacement obtained from the ligand-coupled motions at the multiligand interface and surface are significantly different (Supporting Information Fig. S4). These observations indicate that the dynamic characteristics at the interface are different from surface, and such differential characteristics may encode specificity of recognition. In fact, Haliloglu *et al.*⁴⁷ showed that low-frequency anticorrelated motions in the unbound ligands from 15 enzyme-inhibitor and antigen-antibody complexes successfully predicted binding-sites. (2) The dynamic characteristics of multiligand interface implicitly takes into account global dynamics in the slowest modes. To demonstrate that the dynamics of multiligand interface is in conformity to the global dynamics we have compared 100 low-frequency modes due to the ligand- and self-coupled motions by using all the Ubiquitin atoms (gray boxes), and by using atoms at the multiligand interface of Ubiquitin (black boxes) (Fig. 5). For all the complexes we have observed that the patterns shown by using all atoms and multiligand interface atoms are highly correlated. This indicated that the analyses of dynamics of the multiligand interface capture essential details of the global dynamics.

The dynamics of apo Ubiquitin previously revealed that functionally relevant residues Ile44 and His68 are relatively rigid.⁴⁸ This is consistent with our result (Fig. 3). Moreover, we detected some functionally relevant apo modes of Ubiquitin that showed collective pincer-like motions around the pivotal Ile44 residue [Fig. 4(a)]. The pincer-like motion of Ubiquitin was also observed previously.⁴⁸ The counterbalancing motions in the OV and Non-OV regions due to the self-coupled and ligand-coupled motions indicate that there exists a fine-tuning mechanism of ligand-bound state [Fig. 2(a,b)]. The presence of fine-tuning motions around binding sites in Ubiquitin complexes was proposed in the work of Wlodarski and Zagrovic⁴⁹ and this study might be helpful in extending that idea to characterize such local motions in atomic detail (Supporting Information Table S8).

Figure 5. Comparison of first 100 self-coupled and ligand-coupled modes for the complex Ubiquitin-Vacuolar protein sorting protein 36 (PDB code 2HTH, chain A and B for Ubiquitin and the ligand, respectively). The singular vectors from the self-coupled and ligand-coupled modes are compared by using absolute value of dot-product [gray and black boxes in (a) and (b)]. The cut-off value of the dot-products was 0.1. (a) All atoms of Ubiquitin (N^R) were used. Therefore, the singular vectors were $3N^R$ -dimensional. (b) Atoms at the multiligand interface (N^R_{ml}) were used. Therefore, the singular vectors were $3N^R_{ml}$ -dimensional. Here, N^R and N^R_{ml} are 574 and 464, respectively. Note that the two plots are almost identical.

In a recent analysis Long and Brüschweiler¹⁸ compared the dynamics of apo Ubiquitin to one Ubiquitin complex by molecular dynamics simulations. As far as the C α -atom MSF of apo Ubiquitin is concerned we observed a trend similar to that of Long and Brüschweiler (Fig. 3 and Supporting Information Fig. S4e). In their analysis of all molecular dynamics snap-shots, Long and Brüschweiler showed that the binding region of Ubiquitin in Ubiquitin-Hrs-UIM complex is indistinguishable from the rest of the protein in terms of flexibility. However, based on residual dipolar coupling analysis of apo-Ubiquitin structure and 46 X-ray Ubiquitin complex structures, Lange *et al.*⁴⁸ observed that the flexibility of a residue was positively correlated with its promiscuity of binding. The bound Ubiquitin structures include different types of ligand in the work of Lange *et al.* Therefore, the binding site in their work is similar to the multiligand interface in our study. We observed that the flexibility of a residue (defined by the average flexibility over the atoms) in the multiligand interface is higher than that in the noninterface (P value 7.52×10^{-4}), consistent with Lange *et al.* The apparently contradictory result of Long and Brüschweiler suggests that it is essential to consider multispecific nature of binding to understand the recognition dynamics of Ubiquitin.

We also checked whether the correlation between flexibility and promiscuity is robust against a change in the conformations of apo Ubiquitin. It is known from the variable pressure NMR experiments that the apo Ubiquitin has a more compact conformer (probed at a high-pressure), which remains in an equilibrium with the native conformer.^{50,51} We

performed AA-NMA of a high-pressure Ubiquitin structure (PDB identifier: 2ZCC, chain C) and we observed that the trend in MSF is similar to the native Ubiquitin (Supporting Information Fig. S5f). In both the conformers the conserved flexible regions are included primarily in the OV region. Therefore, we conclude that the unbound Ubiquitin robustly included flexibility in the multiligand interface.

Perica and Chothia¹⁷ studied variable and conserved regions in Ubiquitin, which are related to the recognition of the ligands. They indicated that the residues that make most of the contacts (Leu8, Arg42, Ile44, Gly47, His68, and Val70) are responsible for the specificity and selectivity, although dynamic aspects of the recognition were not studied. Those residues are part of the OV region in present study [Fig. 1(a)] and current result indicates that they undergo diverse ligand-coupled vibrational motions [Fig. 2(a)]. A comparison of ligand-bound and ligand-free dynamics of Ubiquitin showed that MSF averaged over atoms in the OV region is higher in unbound condition ($4.59 \text{ \AA}^2$ as opposed to $1.14 \text{ \AA}^2$) and therefore ligand recognition may be triggered by attenuation of atomic motion in the OV region (Fig. 3 and Supporting Information Fig. S5f).

The main goal of this analysis is to understand how the interplay of ligand-coupled and ligand-constrained self-coupled motions of Ubiquitin is related to its multispecificity. The ligand-coupled motions show direct influence of the ligands on the motions of Ubiquitin and correlated to the conformational change on ligand binding. Such motions directly include dynamic aspects of the multiligand interface related to the multispecificity of Ubiquitin. We

mention that the ligand- and self-coupled motions of Ubiquitin are not independent of each other. This is reflected in the Figure 5 that compared ligand-coupled and self-coupled motions. However, Figure 5 also indicates that the characteristics of ligand- and self-coupled motions are different (because a large number of self-coupled modes are required to characterize one ligand-coupled mode). The similarity between the self-coupled motion of Ubiquitin in complex and its vibrational motion in the unbound condition is consistent with the recent work by Wako and Endo.²⁶ In this work we showed how the self-coupled motion in holo Ubiquitin is similar to the motion in apo condition.

This study helps better understand the motions of Ubiquitin in presence of the ligand. We pointed out that in the complex the motions of Ubiquitin are an intricate combination of ligand-constrained self-coupled motions and ligand-coupled motions of Ubiquitin. The ligand-constrained self-coupled vibrational motion is similar to the motion obtained from the effective Hessian matrix proposed by Zheng and Brooks.²⁹ However, here we directly analyzed the coupling between the receptor and the ligand to explicitly capture the influence of the ligands on receptor motions.

The results of this study may indicate the importance of analyzing ligand-coupled motions to understand the transient nature of date hub protein complexes. We observed that the promiscuity (OV and Non-OV regions) and flexibility are positively correlated, which is consistent with the experiments. In summary, the receptor motions that are directly influenced by ligand motions include essential dynamic details of multispecificity and are robustly embedded in the motions of the unbound condition.

Materials and methods

Preparation of the data set

We selected nine Ubiquitin-protein complexes (Supporting Information Table S2) where Ubiquitin is bound to diverse types of ligands. These nine protein-protein complexes were obtained from our previous data set of 16 types of the date hub proteins.¹⁶ From all types of complexes that occur in the Ubiquitin receptor cluster we selected representative nine heterodimers that were solved by X-ray crystallography and the corresponding protein-protein interfaces were annotated as biological heterodimers in the PDB. Moreover, we confirmed that there were no covalent interactions between the receptor and ligand and there were no chain breaks and missing side-chain atoms in the receptor and ligands. However, in some cases Ubiquitin chains do not contain X-ray coordinates for the last four residues in the C-terminal end. In those cases we obtained gaps (starting from residue 73) in the multiple sequence align-

ment.⁵² Therefore, we removed last four residues from the Ubiquitin sequences so that each Ubiquitin chains were of length 72 (residues 1–72), including 574 atoms. To analyze the unbound condition of Ubiquitin, we have used a monomer structure from the PDB code 1UBQ. All the Ubiquitin chains from nine complexes were superimposed to the Ubiquitin apo-structure [Fig. 1(a)]⁵³ before performing all-atom NMA (AA-NMA).

Definitions of multiligand interface, OV, and Non-OV regions

The multiligand interface on Ubiquitin receptor was defined by multiple sequence alignment⁵² of the receptor and integration of the interface residue sites from the 87 protein-protein interfaces (including nonrepresentatives) present in the Ubiquitin receptor cluster. The overlapping (OV) and nonoverlapping (Non-OV) region of the multiligand interface was defined from the median of promiscuity of each residue sites of Ubiquitin [Fig. 1(a)]. The residues in the OV region were more promiscuous and the residues in the Non-OV region were less promiscuous. The atoms in the OV and Non-OV regions were defined from the atoms obtained from the residues in the OV and Non-OV regions, respectively. There were 22 residues in OV region (178 atoms) (in our initial multiple sequence alignment of the sequences in Ubiquitin cluster there were five receptor sequences where Lys63 was mutated to Arg resulting in a gap; however, in the present set of nine heterodimers Lys63 is not mutated and included in the interfaces in most of the cases, therefore we defined this residue as part of the OV region) and 36 residues in Non-OV region (286 atoms). The rest of the residues (14 in number, corresponds to 110 atoms) were defined to be in the noninterface part of the Ubiquitin, of which 6 residues (48 atoms) were in surface.⁵⁴

Normal Mode analysis using Elastic network model (ENM)

We performed all-atom normal mode analysis (NMA) using elastic network model (ENM)²⁷ for the nine complexes in our data set (Supporting Information Table S2) using our home-made programs. The details of ENM with parameters are given in the Supporting Material S3a. For a complex, NMA will give a covariance matrix (C) of dimension $3N \times 3N$, where N is the number of atoms in the complex. Such a covariance matrix was further split into M_{RR} and M_{LL} square submatrices, M_{RL} and M_{LR} rectangular sub-matrices (“R” represents receptor and “L” represents ligand, Supporting Information Fig. S1c). The M_{RR} and M_{RL} submatrices were analyzed in detail in this study (Supporting Information Fig. S1b). The M_{RR} matrix represents how motions of the receptor atoms are influenced by itself in the presence of a ligand. Such motions of the receptor are

defined as “ligand-constrained self-coupled motion” and are referred to as “self-coupled motion” in the result section. The M_{RL} matrix represents how the motions of the ligand atoms influence motions of the receptor atoms. Such motions of the receptor are defined as “ligand-coupled motion”.

Separation of external and internal motion from the total motion of Ubiquitin or ligands

The covariance matrix C (see above and Supporting Information S3a) indicates how Ubiquitin and the corresponding ligand are fluctuating in the Eckart frame of the whole complex.²⁸ In this case, the fluctuations of the receptor and ligand were due to the total motions of the receptor and ligand, respectively. Such total motion of a sub-unit is composed of internal vibrational motion of that subunit and external rotational, translational movement of one subunit with respect to the other.

The receptor subunit contains $3N^R$ degrees of freedom, and there will be $(3N^R-6)$ vibrational modes, which we defined as the internal modes of the receptor. The remaining six modes corresponding to the rigid body motions of the receptor in its non-Eckart frame are the external modes of the receptor (with respect to the complex). The internal and external modes of receptor (or ligand) were obtained by keeping only the receptor (or ligand) in its Eckart and non-Eckart frames, respectively. In order to do this, projection matrices for the receptor (P_R) and ligand (P_L) can be defined (Supporting Information S3b). For example, to put only the receptor subunit in the Eckart frame, we projected the sub-matrix of the covariance matrix C corresponding to the receptor ($M_{RR}^{(t-t)}$, a $3N^R \times 3N^R$ matrix, where “ t ” represents total motion) by applying a projection matrix P_R , i.e.,

$$M_{RR}^{(i-i)} = P_R M_{RR}^{(t-t)} P_R \quad (1a)$$

where $M_{RR}^{(i-i)}$ indicates covariance matrix due to internal motions of the receptor (“ i ” represents internal motions). The covariance matrix due to external motions of the receptor ($M_{RR}^{(e-e)}$, where “ e ” represents external motions) was obtained by,

$$M_{RR}^{(e-e)} = (1 - P_R) M_{RR}^{(t-t)} (1 - P_R) \quad (1b)$$

To obtain different combination of internal and external motions for ligand-coupled motion similarly we projected the rectangular submatrix $M_{RL}^{(t-t)}$ of C .

To analyze the coupling between internal, external, and total motions of the receptor and ligand we performed singular value decomposition (SVD) of the M_{RR} and M_{RL} matrices (Supporting Information S3b). The AA-NMA and the separation of internal and external motions provided the information regarding couplings between all combinations of in-

ternal and external motions (referred to as partial couplings), as well as the coupling between total motions (referred to as total coupling) of Ubiquitin and the ligands.

References

1. Deeds E, Ashenberg O, Gerardine J, Shakhnovich E (2007) Robust protein-protein interactions in crowded cellular environments. *Proc Natl Acad Sci USA* 104: 14952–14957.
2. Cho K, Lee K, Lee KH, Kim D, Lee D (2006) Specificity of molecular interactions in transient protein-protein interaction interfaces. *Proteins* 65:593–606.
3. Vakser IA (2004) Protein-protein interfaces are special. *Structure* 12:910–912.
4. Keskin O, Gursoy A, Ma B, Nussinov R (2008) Principles of protein-protein interactions: what are the preferred ways for proteins to interact? *Chem Rev* 108: 1225–1244.
5. Bahadur RP, Chakrabarti P, Rodier F, Janin J (2004) A dissection of specific and non-specific protein-protein interfaces. *J Mol Biol* 336:943–955.
6. Bai H, Ma W, Liu S, Lai L (2008) Dynamic property is a key determinant for protein-protein interactions. *Proteins* 70:1323–1331.
7. Henzler-Wildman K, Kern D (2007) Dynamic personalities of proteins. *Nature* 450:964–972.
8. Miller DW, Agard DA (1999) Enzyme specificity under dynamic control: a normal mode analysis of alpha-lytic protease. *J Mol Biol* 286:267–278.
9. Wells JA, McClendon CL (2007) Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* 450:1001–1009.
10. Schreiber G, Keating AE (2011) Protein binding specificity versus promiscuity. *Curr Opin Struct Biol* 21:50–61.
11. Jeong H, Mason SP, Barabasi A-L, Oltvai ZN (2001) Lethality and centrality in protein networks. *Nature* 411:41–42.
12. Han J-DJ, Bertin N, Hao T, Goldberg DS, Berriz GF, Zhang LV, Dupuy D, Walhout AJM, Cusick ME, Roth FP, Vidal M (2004) Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature* 430:88–93.
13. Kim PM, Lu LJ, Xia Y, Gerstein MB (2006) Relating three-dimensional structures to protein networks provides evolutionary insights. *Science* 314:1938–1941.
14. Kim PM, Sboner A, Xia Y, Gerstein M (2008) The role of disorder in interaction networks: a structural analysis. *Mol Syst Biol* 4:179.
15. Tyagi M, Shoemaker BA, Bryant SH, Panchenko AR (2009) Exploring functional roles of multibinding protein interfaces. *Protein Sci* 18:1674–1683.
16. Dasgupta B, Nakamura H, Kinjo AR (2011) Distinct roles of overlapping and non-overlapping regions of hub protein interfaces in recognition of multiple partners. *J Mol Biol* 411:713–727.
17. Perica T, Chothia C (2010) Ubiquitin—molecular mechanisms for recognition of different structures. *Curr Opin Struct Biol* 20:367–376.
18. Long D, Bruschweiler R (2011) In silico elucidation of the recognition dynamics of ubiquitin. *PLoS Comput Biol* 7:9.
19. Sanjouand YH (1996) Normal-mode analysis suggests important flexibility between the two N-terminal domains of CD4 and supports the hypothesis of a conformational change in CD4 upon HIV binding. *Protein Eng Des Sel* 9:671–677.

20. Tama F, Sanejouand YH (2001) Conformational change of proteins arising from normal mode calculations. *Protein Eng Des Sel* 14:1–6.
21. Ahmed A, Villinger S, Gohlke H (2010) Large-scale comparison of protein essential dynamics from molecular dynamics simulations and coarse-grained normal mode analyses. *Proteins* 78:3341–3352.
22. Kitao A, Go N (1999) Investigating protein dynamics in collective coordinate space. *Curr Opin Struct Biol* 9: 164–169.
23. Zheng W, Brooks BR, Thirumalai D (2006) Low-frequency normal modes that describe allosteric transitions in biological nanomachines are robust to sequence variations. *Proc Natl Acad Sci USA* 103: 7664–7669.
24. Marcos E, Crehuet R, Bahar I (2011) Changes in dynamics upon oligomerization regulate substrate binding and allostery in amino acid kinase family members. *PLoS Comput Biol* 7:e1002201.
25. Meireles L, Gur M, Bakan A, Bahar I (2011) Pre-existing soft modes of motion uniquely defined by native contact topology facilitate ligand binding to proteins. *Protein Sci* 20:1645–1658.
26. Wako H, Endo S (2011) Ligand-induced conformational change of a protein reproduced by a linear combination of displacement vectors obtained from normal mode analysis. *Biophys Chem* 159:257–266.
27. Tirion M (1996) Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Phys Rev Lett* 77:1905–1908.
28. Wilson EB, Decius JC, Cross PC (1955) *Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra*. McGraw-Hill, New York; p. 394.
29. Zheng W, Brooks BR (2005) Probing the local dynamics of nucleotide-binding pocket coupled to the global dynamics: myosin versus kinesin. *Biophys J* 89:167–178.
30. Ishida H, Jochi Y, Kidera A (1998) Dynamic structure of subtilisin-eglin c complex studied by normal mode analysis. *Proteins* 32:324–333.
31. Standley DM, Kinjo AR, Kinoshita K, Nakamura H (2008) Protein structure databases with new web services for structural biology and biomedical research. *Briefings Bioinf* 9:276–285.
32. Murzin AG, Brenner SE, Hubbard T, Chothia C (1995) SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol* 247:536–540.
33. Punta M, Coghill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer ELL, Eddy SR, Bateman A (2011) The Pfam protein families database. *Nucleic Acids Res* 40:290–301.
34. Sloper-Mould KE, Jemc JC, Pickart CM, Hicke L (2001) Distinct functional surface regions on ubiquitin. *J Biol Chem* 276:30483–30489.
35. Kinjo AR, Nakamura H (2010) Geometric similarities of protein-protein interfaces at atomic resolution are only observed within homologous families: an exhaustive structural classification study. *J Mol Biol* 399: 526–540.
36. Tsuchiya Y, Kinoshita K, Endo S, Wako H (2012) Dynamic features of homodimer interfaces calculated by normal-mode analysis. *Protein Sci* 21:1503–1513.
37. Garner TP, Strachan J, Shedden EC, Long JE, Cavey JR, Shaw B, Layfield R, Searle MS (2011) Independent interactions of ubiquitin-binding domains in a ubiquitin-mediated ternary complex. *Biochemistry* 50: 9076–9087.
38. Ma J (2005) Usefulness and limitations of normal mode analysis in modeling dynamics of biomolecular complexes. *Structure* 13:373–380.
39. Ikeguchi M, Ueno J, Sato M, Kidera A (2005) Protein structural change upon ligand binding: linear response theory. *Phys Rev Lett* 94:078102.
40. Zheng W, Doniach S (2003) A comparative study of motor-protein motions by using a simple elastic-network model. *Proc Natl Acad Sci USA* 100:13253–13258.
41. Koga N, Takada S (2006) Folding-based molecular simulations reveal mechanisms of the rotary motor F1-ATPase. *Proc Natl Acad Sci USA* 103:5367–5372.
42. Dobbins SE, Lesk VI, Sternberg MJE (2008) Insights into protein flexibility: The relationship between normal modes and conformational change upon protein-protein docking. *Proc Natl Acad Sci USA* 105: 10390–10395.
43. Gibrat JF, Go N (1990) Normal mode analysis of human lysozyme: study of the relative motion of the two domains and characterization of the harmonic motion. *Proteins* 8:258–279.
44. Cuff AL, Sillitoe I, Lewis T, Clegg AB, Rentzsch R, Furnham N, Pellegrini-Calace M, Jones D, Thornton J, Orengo CA (2011) Extending CATH: increasing coverage of the protein structure universe and linking structure with function. *Nucleic Acids Res* 39:D420–D426.
45. Orengo CA, Michie AD, Jones S, Jones DT, Swindells MB, Thornton JM (1997) CATH—a hierarchic classification of protein domain structures. *Structure* 5: 1093–1108.
46. Zen A, Micheletti C, Keskin O, Nussinov R (2010) Comparing interfacial dynamics in protein-protein complexes: an elastic network approach. *BMC Struct Biol* 10:26.
47. Haliloglu T, Seyrek E, Erman B (2008) Prediction of binding sites in receptor-ligand complexes with the Gaussian Network Model. *Phys Rev Lett* 100:228102.
48. Lange OF, Lakomek N-A, Farès C, Schröder GF, Walter KFA, Becker S, Meiler J, Grubmüller H, Griesinger C, De Groot BL (2008) Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. *Science* 320:1471–1475.
49. Wlodarski T, Zagrovic B (2009) Conformational selection and induced fit mechanism underlie specificity in noncovalent interactions with ubiquitin. *Proc Natl Acad Sci USA* 106:19346–19351.
50. Kitahara R, Yokoyama S, Akasaka K (2005) NMR snapshots of a fluctuating protein structure: ubiquitin at 30 bar–3 kbar. *J Mol Biol* 347:277–285.
51. Kitahara R, Yamada H, Akasaka K (2001) Two folded conformers of ubiquitin revealed by high-pressure NMR. *Biochemistry* 40:13556–13563.
52. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) ClustalW and ClustalX version 2. *Bioinformatics* 23:2947–2948.
53. Diamond R (1988) A note on the rotational superposition problem. *Acta Cryst.* 44:211–216.
54. Hubbard SJ, Thornton JM (1993) “NACCESS”, Computer Program.
55. Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22:2577–2637.