Reassessing a sparse energetic network within a single protein domain

Celestine N. Chi*, Lisa Elfström*, Yao Shi*, Tord Snäll†, Åke Engström*, and Per Jemth*‡

*Department of Medical Biochemistry and Microbiology, Uppsala University Biomedical Centre, Box 582, SE-751 23 Uppsala, Sweden; and †Department of Ecology, Swedish University of Agricultural Sciences, P.O. Box 7044, SE-750 07 Uppsala, Sweden

Edited by Alan R. Fersht, University of Cambridge, Cambridge, United Kingdom, and approved January 22, 2008 (received for review December 13, 2007)

Understanding the molecular principles that govern allosteric communication is an important goal in protein science. One way allostery could be transmitted is via sparse energetic networks of residues, and one such evolutionary conserved network was identified in the PDZ domain family of proteins by multiple sequence alignment [Lockless SW, Ranganathan R (1999) *Science* 286:295–299]. We have reassessed the energetic coupling of these residues by double mutant cycles together with ligand binding and stability experiments and found that coupling is not a special property of the coevolved network of residues in PDZ domains. The observed coupling for ligand binding is better explained by a distance relationship, where residues close in space are more likely to couple than distal residues. Our study demonstrates that statistical coupling from sequence analysis is not necessarily a reporter of energetic coupling and allostery.

allostery | coupling energy | dynamics | energetic network of residues | PDZ domain

Predicting allosteric mechanisms from protein sequence is an important goal, yet very difficult because the molecular basis for allostery in proteins is still not fully understood. Allosteric regulation is the alteration of protein function through the binding of an effector molecule elsewhere within the same protein. This definition can be expanded to include amino acid point mutations distal from the active site that can effectively alter protein function (1). Allosteric regulation is well established in multidomain proteins (2, 3), but allosteric behavior in single protein domains is less documented (for example, ref. 4), and its mechanisms unclear. Nevertheless, the idea of allostery without a (well defined) conformational change (5) has become popular, and it has even been proposed that all dynamic proteins are allosteric (6).

Lockless and Ranganathan (7) set out to investigate whether allostery could be predicted from sequence conservation. They found a coevolved network of residues in the PDZ domain family of proteins, and these statistically coupled residues were confirmed by experiments on PSD95 PDZ3 to be energetically coupled (7). This study has served as a classic example of allostery in a single protein domain, with a "sparse network" of energetically linked positions that could affect function, in this case ligand binding. However, pathways of energetic connectivity inferred from statistical analysis have been questioned because the correlated mutation algorithm that was used finds pairs of residues that are close in physical space in the protein structure and it is therefore not surprising that the residues actually do couple (8). Furthermore, other issues regarding the validity of the statistical method used by Lockless and Ranganathan have been raised: the algorithm used is not symmetric (9), does not incorporate evolutionary noise (10), and performs less well than other algorithms (8, 9, 11). From an experimental point of view, previous work on PSD95 PDZ3 in our laboratory suggested that this PDZ domain is conformationally rigid upon ligand binding (12) in agreement with the crystal structure (13). Finally, for PTP-BL PDZ2 the energetic coupling between two of the network residues was found to be absent (14). In light of these uncertainties, we here reassess the network of residues proposed to form a pathway of energetic connectivity in PDZ domains. Our results reveal that the sparse energetic network in PSD95 PDZ3 is not affecting ligand binding more than non-network residues. The observed coupling energies for binding can be explained by a model where distance is correlated with coupling (8), as shown previously for the barnase–barstar complex (15) and Staphylococcal nuclease (8, 16). Thus, statistical coupling inferred from multiple sequence alignment is not necessarily a true reporter of functional coupling.

Results

To test whether the coevolved network of residues in PDZ domains (7) is energetically coupled we performed double-mutant cycles (17–19) in conjunction with ligand binding and stability experiments. In such experiments, two positions are mutated singly and in concert. The effect of the first mutation on, for example, ligand binding is measured first against the wild-type background and second against the second mutation. The different effect (in free energy) of the mutation on the two backgrounds is the coupling energy, $\Delta\Delta\Delta G_C$, between the two mutated residues (see *Materials and Methods*). In PSD95 PDZ3, H372 is a network residue (7) that makes a hydrogen bond with the ligand peptide (13). Therefore, this residue was chosen as the reference position to probe the energetic connectivity of the conserved network (7). We chose to use the same reference mutation, H372Y, as in the previous study (7). The decrease in affinity for the H372Y mutation (two orders of magnitude) was similar to that for the deletion mutation H372A (20), suggesting that the hydrogen bond present in the wild-type complex (13) is not formed between the mutant H372Y and the peptide. It is clear that the PDZ domain pays a thermodynamic penalty for this functional His residue because mutation to either Ala or Tyr stabilizes the protein by $\approx$2 kcal·mol^{-1}. Double-mutant cycle analyses were performed for 10 positions using conservative mutations (21) in PDZ3 (Fig. 1), with the His $\rightarrow$ Tyr at position 372 as the reference mutation. Six of the residues mutated in the present study were from the 10 evolutionarily conserved network positions previously identified (7), and four residues were non-network positions in the protein core. In the binding studies, a dansylated peptide was titrated with PDZ3 and the resulting change in fluorescence was monitored. Fig. 2 illustrates two examples of double mutant cycle binding experiments, for positions 340 and 380: position 340 is a non-network position whereas 380 is a network position, but none of the residues couple with H372. The three positions 329, 362, and 376 yielded

Author contributions: C.N.C., L.E., Y.S., and P.J. designed research; C.N.C., L.E., Y.S., and Å.E. performed research; C.N.C., L.E., Y.S., T.S., Å.E., and P.J. analyzed data; and C.N.C., T.S., and P.J. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

‡To whom correspondence should be addressed. E-mail: per.jemth@imbim.uu.se.

This article contains supporting information online at www.pnas.org/cgi/content/full/0711732105/DC1.

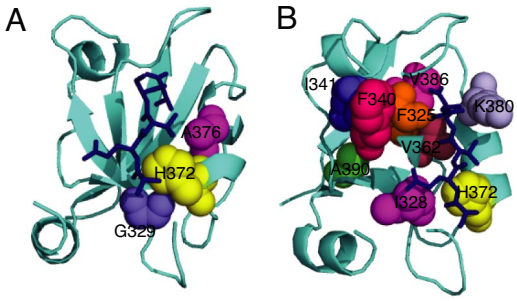

Fig. 1. The crystal structure of PSD95 PDZ3 bound to its peptide ligand (thick blue sticks) [PDB entry 1BE9 (Doyle)]. Residues mutated in the present study are shown as spheres. (*A*) Residues located close and making direct contact with H372. (*B*) Residues that are distal to H372 and not making direct contact. Note that residue 328 is an Ile in this crystal structure but that in our construct we have the wild-type residue, Val. This figure was drawn in PyMOL (47).

coupling energies of 1.5 ± 0.2, 0.6 ± 0.3, and 0.7 ± 0.4 kcal·mol^{-1}, respectively, whereas the remaining seven positions had coupling energies close to zero (Fig. 3*A* and Table 1). The high coupling energy for position 329 is reasonable because the backbone carbonyl of Gly makes a hydrogen bond to a side chain nitrogen of H372. Mutation to Ala at this position prevents the backbone from adopting the required conformation for formation of the hydrogen bond to H372. The hydrogen bond presumably stabilizes the peptide binding conformation of the H372 side chain, and mutation abolishes this preorganization of the binding site. Position 362 is located in a turn on the opposite side of helix 2, compared with H372. The coupling displayed may be explained by loss of hydrophobic interactions with helix 2. Removal of two methyl groups possibly enhances the flexibility of the helix and thus distorts the position of H372, situated at the end of helix 2. The loss of hydrophobic interaction by the Val → Ala mutation is also reflected as a destabilization of this protein by 0.7 kcal·mol^{-1} (Table 1). Position 376 is situated in the middle of helix 2 only one turn away from H372. The coupling free energy calculated for position 376 depended on the type of mutation: the (nonconservative) A376V mutation yielded a coupling free energy of −0.2 ± 0.2 kcal·mol^{-1} and the A376G mutation yielded a coupling free energy of 0.7 ± 0.4 kcal·mol^{-1}. Alanine has the highest helix propensity of all residues whereas glycine has a very low helix propensity (21). The Ala → Gly mutation at this position destabilized the helix by 1.0 kcal·mol^{-1} (Table 1), and a likely cause of the coupling is that the backbone nitrogen makes a weaker interaction with the backbone oxygen of the H372, leading to a more flexible His side chain and again a destabilization of the preorganized binding site. The helix is not destabilized by the A376V mutation despite Val's having lower general helix propensity than Ala (Table 1). This result emphasizes how the choice of mutation affects the measured coupling energies (0.7 and −0.2 kcal·mol^{-1} for A → G and A → V, respectively). Except for this A376V mutation (and the reference H → Y), we used deletion mutations in this work because other mutations will complicate the analysis (22). For binding coupling energies, residues far away from H372 displayed less coupling energy than close residues (Fig. 3*B*). We found no correlation between the binding coupling energy ($\Delta\Delta\Delta G_C$) and the stability coupling energy ($\Delta\Delta\Delta G_{D\text{-}N}$) (Fig. 3*C*). In other words, residues that affect His-372's contribution to stability cannot in general be predicted to affect its contribution to the ligand binding energy. To see whether there was any correlation between our experimental coupling energies and the previously calculated statistical coupling energies, we plotted the binding coupling energy ($|\Delta\Delta\Delta G_C|$, absolute numbers) as a function of statistical coupling energy ($\Delta\Delta G_{STAT}$) extracted from figure 2b of Lockless and Ranganathan (7). We found a very weak linear correlation ($R^2 = 0.18$) between experimental coupling for ligand binding and statistical coupling in the range 0.1–1.5 kcal·mol^{-1}. The correlation between coupling for stability and statistical coupling was also weak ($R^2 = 0.21$), and statistical coupling can therefore not be used to predict real thermodynamic coupling. Finally, we investigated (*i*) the probabilities that the coupling energies are greater than zero (Table 1) and (*ii*) the probability that the coupling energies of distal network residues are different from those of the tested non-network residues. The G329 residue is clearly coupled to His-372 and was excluded from the

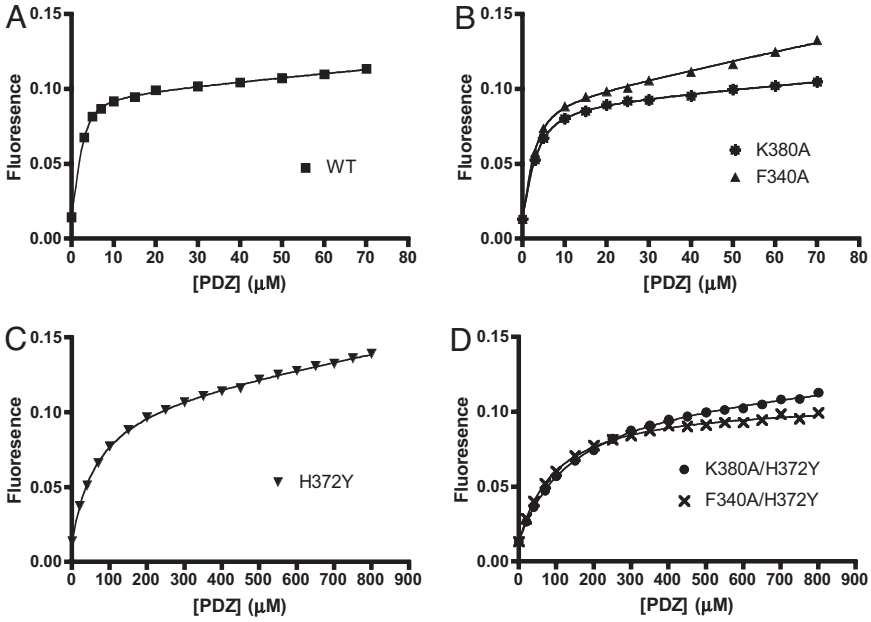

Fig. 2. Equilibrium ligand binding experiments. Shown are equilibrium titrations of the wild-type PDZ3 (*A*), single mutants F340A (filled triangles) and K380A (✱) (*B*), H372Y mutant (*C*), and double mutants F340A/H372Y (×) and K380A/H372Y (filled circles) (*D*). K380 is a network residue, whereas F340 is a non-network residue (7). (See *Materials and Methods* for experimental conditions.)

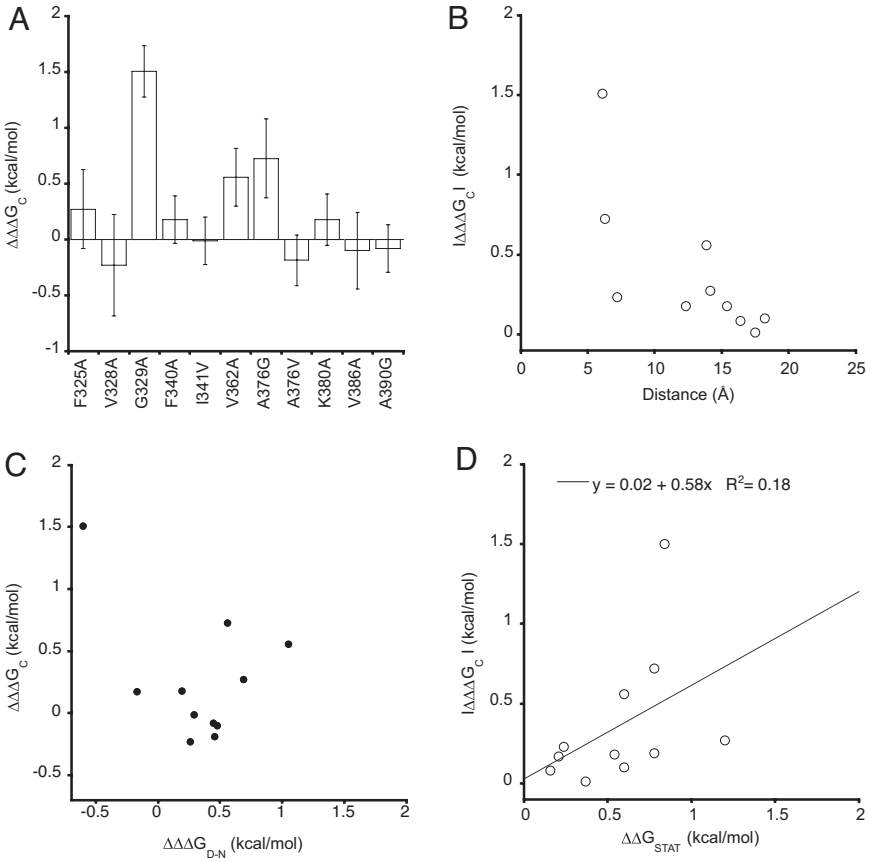

Fig. 3. Plots of measured coupling energies. The coupling free energies were obtained by fitting K_D values to Eq. **3**. (*A*) Binding coupling energies in kcal·mol^{-1}. Error bars are propagated standard errors determined from at least three independent measurements of each K_D (for V328A, A376G, and K380A) or propagated fitting errors and standard errors of independent measurements. Note that the error bars thus depend on the individual errors of each of the four K_D values used to calculate the coupling energy. (*B*) C_α–C_α distance (Å) as a function of binding coupling energy. (*C* and *D*) Binding coupling energy $\Delta\Delta\Delta G_C$ as a function of stability coupling energy $\Delta\Delta\Delta G_{D\text{-}N}$ (*C*) and statistical coupling energy $\Delta\Delta G_{STAT}$ (*D*). Statistical coupling energies were extracted from figure 2b of ref. 7. The units of $\Delta\Delta G_{STAT}$ are arbitrary. Note that the absolute values for the coupling energies are plotted in *B* and *D*.

analysis. A376 is also close to His-372 (backbone hydrogen bond), but the probability that the distal network residues are different from the tested non-network residues is 79% with and 73% without this residue. Hence, there is a low probability that network and non-network residues have different coupling energies. On the residue level, the probabilities for $\Delta\Delta\Delta G_C > 0$ decrease almost linearly with distance between 5 and 20 Å (Table 1) ($R^2 = 0.74$; data not shown), if the outlier V328 is omitted. This residue affects binding possibly through favorable van der Waals interactions with the Gln(-4) of the peptide ligand but has a very low coupling energy with His-372.

Discussion

We demonstrate here for a PDZ domain that the statistical coupling energies of a proposed allosteric evolutionary conserved network of residues, deduced from statistical coupling analysis (7, 23), cannot predict coupling energies from ligand binding and stability experiments. Instead, residues that are close in space tend to couple energetically. Allostery in single protein domains is often discussed in the recent literature in terms of an important functional property (24–26), but experimental data are scarce. Indeed, all proteins including single domains are dynamic on the picosecond-to-microsecond time scale as shown by NMR experiments (27), and sometimes this property plays a functional role. It has, for example, been shown by NMR relaxation experiments that microsecond-time-scale backbone dynamics correlate with enzyme activity for the peptidyl-prolyl *cis-trans* isomerase Cyclophilin A (28). Another recent example is adenylate kinase, where multiple experimental techniques together with molecular dynamics simulations demonstrate that motions on the nano-millisecond time scales promote catalysis (29, 30). There are also rare examples of allostery without a conformational change. Binding of the first metal ion to calbindin D_{9k} affects the affinity for the second ion via a stiffening of the protein backbone (31). A second example is binding of cAMP to catabolite activator protein, which displays dynamics-driven negative cooperativity (32). However, in most cases, the relevance of allostery for the function of isolated domains appears negligible unless it is manifested as a well defined conformational change, such as for the p38 MAP kinase (two domains) (33), the signaling protein NtrC (4), and myoglobin (34). One of the best examples of an allosteric protein domain was found in the PDZ domain family where an evolutionary conserved network of residues was identified using multiple sequence alignment and confirmed by experiment to be a "sparse energetic network" (7, 23). Our present results, however, demonstrate that this network of residues does not display energetic coupling beyond a straightforward distance relationship. Allostery might be an intrinsic property of all proteins (6), but its role in protein function should not be exaggerated in the absence of solid experimental data. Instead, it appears likely that allostery in single-domain proteins often works on short distances (8, 15, 16). Energetic coupling and allostery in its widest meaning (35) is best tested experimentally by double mutant

Table 1. Binding constants and probability estimation of PDZ-mutants and peptide ligand

PDZ	K_D,* μM	k_{on},† $\mu M^{-1}\cdot s^{-1}$	$\Delta\Delta\Delta G_C$,‡ kcal·mol^{-1}	Probability§	$\Delta\Delta G_{D\text{-}N}$,¶ kcal·mol^{-1}
WT	0.5 (0.1)	10 (0.3)			
F325A	0.3	10 (0.2)			3.0 (0.1)
V328A	1.7 (0.1)	9 (0.1)			0.2 (0.2)
G329A	46 (4)	9 (0.7)			0.5 (0.1)
F340A	1.0	10 (0.1)			0.2 (0.1)
I341V	0.5	10 (0.1)			−0.2 (0.1)
V362A	0.7 (0.1)	7 (0.1)			0.7 (0.1)
A376G	2.5 (0.7)	7 (0.1)			1.0 (0.1)
A376V	0.3 (0.02)	8 (0.1)			−0.7 (0.1)
K380A	1.5 (0.1)	7 (0.1)			−1.0 (0.1)
V386A	0.5	8 (0.1)			1.2 (0.1)
A390G	0.4	10 (0.1)			2.2 (0.1)
H372Y	67 (4)	nd‖			−2.1 (0.1)
F325A/H372Y	28 (8)	nd	0.3	0.79	3.6 (0.1)
V328A/H372Y	330 (130)	nd	−0.2 (0.5)	0.38	0.4 (0.1)
G329A/H372Y	490	nd	1.5		−0.1 (0.1)
F340A/H372Y	98	nd	0.2	0.78	0.0 (0.1)
I341V/H372Y	70	nd	0.0	0.60	0.1 (0.1)
V362A/H372Y	36	nd	0.6	0.95	1.8 (0.1)
A376G/H372Y	100 (5)	nd	0.7 (0.4)	0.97	1.5 (0.1)
A376V/H372Y	48	nd	−0.2		−1.0 (0.1)
K380A/H372Y	150 (6)	nd	0.2 (0.2)	0.76	−0.8 (0.1)
V386A/H372Y	79 (21)	nd	−0.1	0.54	1.7 (0.1)
A390G/H372Y	54	nd	−0.1	0.53	2.6 (0.1)

*K_D values were obtained by fitting Eq. **1** to the equilibrium binding data. Errors are shown in parenthesis and were calculated from at least three independent measurements.

†The association rate constant k_{on} was determined by fitting Eq. **2** to the plots of observed rate constants versus PDZ concentrations. Errors shown in parenthesis are fitting errors.

‡$\Delta\Delta\Delta G_C$ is the coupling energy for ligand binding between tested positions and H372. Errors shown in parenthesis are propagated errors in the K_D values.

§The estimate of probability $\Delta\Delta\Delta G_C > 0$ is based on the posterior distribution of the mean among one to five measurements of independent K_D measurements. The model fitted is described in *Materials and Methods*.

¶$\Delta\Delta\Delta G_{D\text{-}N}$ is the difference in stability between the wild-type and the single mutants or between the H372Y mutant and the double mutants. Errors shown in parenthesis are estimated from fitting errors of the denaturation constants.

‖nd (not determined), no association rate constants were determined for the double mutants. Observed rate constants were too high for the stopped-flow machine because of large off-rate constants.

cycles (17–19). In the double mutant cycles performed in our study, we chose to make as conservative mutations as possible; for example, Phe-325 was mutated to Ala, and Ala-376 was mutated to Gly (rather than to Lys and Val, respectively, as in the study by Lockless and Ranganathan). The reason is to avoid adding new functional groups to the protein or changing the structure by inserting a larger residue, because such mutations will complicate analysis of the results (21, 22, 36). For most positions tested, we obtained different coupling energies from those in the previous study, even in two cases when we made identical mutations (K380A and the nonconservative A376V). The basis for this difference is unclear, but it stresses an important point: depending on the type of experiment performed, the experimental error in measured coupling energies could be high and the threshold for a reliable coupling energy should therefore also be high. In our study, we used 12–20 data points per binding isotherm (Fig. 2) including independent measurements to account for the measurement error in the dissociation constant K_D. Furthermore, we performed a comprehensive statistical analysis of all experimental data. This analysis showed that one can question strong conclusions from coupling energies unless there is a clear coupling between residues, such as the one found for G329 and H372, with a $\Delta\Delta\Delta G_C$ of 1.5 kcal·mol^{-1}.

Thus, PSD95 PDZ3 is a rigid protein–protein interaction module, which cannot be used as a paradigm for allostery mediated by an evolutionary conserved network of residues. For PSD95 PDZ3, we did not observe the correlation between the energetic barrier for association of ligand and overall stability as for PTP-BL PDZ2 (37), but it should be mentioned that there are several examples of conformational changes that relate to function in the PDZ domain family. The CRIB domain induces a conformational change in its adjacent PDZ domain in Par6 (38). In PTP-BL PDZ2, a small conformational change involving helix 2 was found to be associated with ligand binding (12) and a similar reorientation of helix 2 was found to depend on the nature of the ligand for Syntenin PDZ2 (39). Note that none of these examples relates to the evolutionarily conserved network of residues in the PDZ family. Identifying coevolving residues using multiple sequence alignment is an interesting tool (7, 23, 40), but for PDZ domains, this evolutionarily conserved network of residues appears not to play a particular functional role. It is clear from the results presented here that statistical coupling is not necessarily a reporter of energetic coupling.

Materials and Methods

Cloning, Expression, and Purification of Proteins. The PSD95 PDZ3 used in this study was similar to that in Chi *et al.* (20). The numbering of residues in this article is in accordance with the crystal structure of PSD95 PDZ3 (13) (PDB entry 1BE9). Eleven single mutants and double mutants were generated by inverted PCR using Pfu Turbo polymerase (Stratagene) from the pseudo-wild-type PDZ3 cDNA (containing a Trp at position 331 as probe for folding and binding) (Fig. 1). All double mutants contained H372Y as the reference mutation. Protein expression and purification were as described in ref. 20. All purified samples were dialyzed against 25 mM Tris·HCl (pH 7.5). MALDI-TOF mass spectrometry was used to

confirm the identity of the purified proteins, and their concentrations were determined by absorbance at 280 nm with an extinction coefficient determined from amino acid analysis. All mutants were well folded and stable as judged by urea denaturation experiments (data not shown).

Equilibrium and Pre-Steady-State Binding Experiments. All equilibrium-binding measurements were performed at 25°C in 50 mM potassium phosphate (pH 7.5) in an SLM 4800 spectrofluorimeter (SLM Instruments). The peptides used in the experiments, Dansyl-YKQTSVcoo- and Dansyl-KQSTVcoo-, correspond to the C terminus of CRIPT and were of high grade (purity >95%) (JPT Peptide Technologies). To determine the equilibrium constants for the PDZ3–peptide interaction, the concentration of PDZ wild-type and mutants were varied at constant concentration of peptide (3 μM), and the change in Dansyl fluorescence was measured. Excitation was at 345 nm, and emission was at 551 nm. The pre-steady-state binding kinetics were measured at 10°C on an SX-20 MV spectrometer (Applied Photophysics) as described in ref. 41. All stability measurements were performed by urea-induced denaturation experiments at 25°C in 0.1 M sodium acetate (pH 2.85) (giving reliable denatured and native baselines for most mutants) except for the F325A and F325A/H372Y mutants, which were determined in 50 mM potassium phosphate (pH 7.5) to obtain a good native baseline for F325A in the curve fitting. The change in tryptophan fluorescence upon denaturation and subsequent stability estimation was as described in ref. 42.

Data Analysis. Data from equilibrium binding experiments were fitted to Eq. **1**:

$$F = [[([\mathrm{A}]_0 + K_\mathrm{D} + n)/2 - \mathrm{sqrt}(([\mathrm{A}]_0 + K_\mathrm{D} + n)^2/4 - [\mathrm{A}]_0 n) + n[\mathrm{A}]_0] \times \mathrm{B}] + \mathrm{C} + \mathrm{D} \times [\mathrm{A}]_0, \quad [1]$$

where F is the observed fluorescence signal, n and $[\mathrm{A}]_0$ are the total concentrations of the nonvaried and varied species, respectively, and K_D is the equilibrium dissociation constant. D is the correction factor that takes into account the background fluorescence due to increase in protein concentration. B and C are the constants that take into account the total fluorescence change and fluorescence at $[\mathrm{A}]_0 = 0$, respectively (41).

Data from pre-steady-state binding kinetics were fitted to the Eq. **2** (43):

$$k_\mathrm{obs} = \mathrm{sqrt}(k_1^2(n - [\mathrm{A}]_0)^2 + k_{-1}^2 + 2k_1k_{-1}(n + [\mathrm{A}]_0)), \quad [2]$$

where k_obs is the observed rate constants, n and $[\mathrm{A}]_0$ are as mentioned above, and k_1 and k_{-1} are the on-rate constant and off-rate constant, respectively.

Coupling Free Energies. For calculation of the coupling free energy between two positions, a double-mutant cycle was constructed. The effect of mutation m1 is measured on two different species: first, against the wild-type background ($\Delta\Delta G_{\mathrm{m1}}$) and second against a reference mutation m2 ($\Delta\Delta G_{\mathrm{m1/m2}}$). The difference between the two energies ($\Delta\Delta G_{\mathrm{m1}} - \Delta\Delta G_{\mathrm{m1/m2}}$) is the coupling energy $\Delta\Delta\Delta G_\mathrm{C}$ between the two sites 1 and 2. The equation for the calculation of the coupling energy for binding is shown in Eq. **3**:

$$\Delta\Delta\Delta G_\mathrm{C} = RT \times \ln[(K_\mathrm{D}^\mathrm{WT} \times K_\mathrm{D}^\mathrm{m1/m2})/(K_\mathrm{D}^\mathrm{m1} \times K_\mathrm{D}^\mathrm{m2})]/1{,}000(\mathrm{kcal{\cdot}mol^{-1}}), \quad [3]$$

where R is the gas constant and T is the temperature in Kelvin. K_D^WT, K_D^m1, K_D^m2, and $K_\mathrm{D}^\mathrm{m1/m2}$ are the equilibrium dissociation constants for the wild-type and respective mutants. To calculate the coupling energy for stability, the K_D values are replaced by the respective equilibrium constants for denaturation.

Statistical Analysis. We calculated the probability that the two coupling energies for network (group 1) and non-network (group 2) residues [supporting information (SI) Table 2] are different based on the difference in the probability distributions of the mean coupling energies among the two groups. We used the Bayesian statistical modeling approach because it provides the full (posterior) probability distribution of the parameters, and it is convenient for fitting models based on data with missing observations. For details on Bayesian model fitting and inference, see refs. 44–46. Briefly, the Bayesian statistical modeling approach is different from the classic, so-called ''frequentist'' approach. One main difference is the fitting procedure. Bayesian models are typically fitted using numerical approximation methods such as MCMC (e.g., ref. 46). This procedure essentially means finding the right part of parameter space (the burn-in) and then sampling this part of parameter space. The sample from parameter space constitutes the posterior distribution of the parameters, and this distribution is used for inference. The burn-in is excluded. The resulting posterior distribution is based on both the data and the á-priori assumption about the distribution of the parameters, the ''prior distribution'' (SI Table 3). Therefore, one can speak of distributions even with a small set of data. With decreasing amounts of data, the influence of the prior distribution on the resulting posterior distribution increases. We had no prior information about plausible parameter values, which led us to use so-called noninformative or ''vague'' priors. This is typical for Bayesian analysis. We fitted the model by running three chains: after a ''burn-in'' of 1,000 iterations, every 100th iteration was taken from a total of 200,000 iterations. Convergence was assessed by eye. For prior distributions and estimates of parameters, see SI Table 3.

The probability that the group-specific $\Delta\Delta\Delta G_\mathrm{C}$ are different were calculated as the proportion of the probability distribution of the difference, $\overline{\Delta\Delta\Delta G_\mathrm{C}diff}$, that is >0. For a description of this statistical testing approach, see, e.g., page 20 in ref. 45. This probability distribution of the difference in means between the group-specific $\Delta\Delta\Delta G_\mathrm{C}$ was calculated as follows. For each coupling energy for network residues, $i = 1, \ldots, 5$, and non-network residues, $i = 6, \ldots, 9$,

$$\overline{\Delta\Delta\Delta G_\mathrm{c}}\mathrm{diff} = \overline{\Delta\Delta\Delta G_\mathrm{c}[1{-}5]} - \overline{\Delta\Delta\Delta G_\mathrm{c}[6{-}9]}.$$

For each coupling energy, i,

$$\Delta\Delta\Delta G_\mathrm{C}[i] = \Delta\Delta G_\mathrm{m1}[i] - \Delta\Delta G_\mathrm{m1/m2}[i],$$

where

$$\Delta\Delta G_\mathrm{m1}[i] = -RT\ln(\overline{K_\mathrm{D}^\mathrm{WT}}/\overline{K_\mathrm{D}^\mathrm{m1}[i]})$$

and

$$\Delta\Delta G_{\mathrm{m1}/m2}[i] = -RT\ln(\overline{K_\mathrm{D}^\mathrm{m1/m2}}/\overline{K_\mathrm{D}^\mathrm{m2}[i]}).$$

$\overline{K_\mathrm{D}^\mathrm{m1}[i]}$, $\overline{K_\mathrm{D}^\mathrm{m2}[i]}$, $\overline{K_\mathrm{D}^\mathrm{WT}}$, and $\overline{K_\mathrm{D}^\mathrm{m1/m2}}$ are mean parameters.

$K_\mathrm{D}^\mathrm{m1}[i]$ was estimated based on j measurements, $j = 1$, 2, or 3 as

$$\ln(K_\mathrm{D}^\mathrm{m1}[i, j]) \sim Normal(\ln(\overline{K_\mathrm{D}^\mathrm{m1}[i]}),\ \tau_{K_\mathrm{D}^\mathrm{m1}[i]}),$$

where *Normal*(*a*, *b*) is a normal distribution with mean a and precision b (= variance^{-1}). For mutants where less than three measurements j were available, we estimated $K_\mathrm{D}^\mathrm{m1}[i, j]$ (44). These estimates were based on modeling the precision as

$$\tau_{K_\mathrm{D}^\mathrm{m1}[i]} \sim Gamma(\exp(\phi_{K_\mathrm{D}^\mathrm{m1}[i]}),\ \theta_{K_\mathrm{D}^\mathrm{m1}}),$$

where *Gamma*() is a gamma distribution with shape a and scale b. We assumed that $\phi_{K_\mathrm{D}^\mathrm{m1}}[i]$ increased linearly with $\overline{K_\mathrm{D}^\mathrm{m1}[i]}$; hence,

$$\phi_{K_\mathrm{D}^\mathrm{m1}[i]} = \beta^\mathrm{m1}(\overline{K_\mathrm{D}^\mathrm{m1}[i]} - \mathrm{mean}(\overline{K_\mathrm{D}^\mathrm{m1}[]})) + \beta_0^\mathrm{m1},$$

where β^m1 and β_0^m1 are regression parameters.

$\overline{K_\mathrm{D}^\mathrm{m2}[i]}$ was estimated using the same principle but based on k measurements, $k = 1$, 2, 3, 4, or 5, as

$$\ln(K_\mathrm{D}^\mathrm{m2}[i, k]) \sim Normal(\ln(\overline{K_\mathrm{D}^\mathrm{m2}[i]}),\ \tau_{K_\mathrm{D}^\mathrm{m2}[i]}).$$

For mutant pairs where less than five measurements k were available, we estimated $K_\mathrm{D}^\mathrm{m2}[i, k]$. These estimates were based on modeling the precision as

$$\tau_{K_\mathrm{D}^\mathrm{m2}[i]} \sim Gamma(\exp(\phi_{K_\mathrm{D}^\mathrm{m2}[i]}),\ \theta_{K_\mathrm{D}^\mathrm{m2}}),$$

where

$$\phi_{K_\mathrm{D}^\mathrm{m2}[i]} = \beta^\mathrm{m2}(\overline{K_\mathrm{D}^\mathrm{m2}[i]} - \mathrm{mean}(\overline{K_\mathrm{D}^\mathrm{m2}[]})) + \beta_0^\mathrm{m2},$$

where β^m2 and β_0^m2 are parameters.

We had problems with model convergence in estimating $\phi_{K_\mathrm{D}^\mathrm{m2}}[i]$ because of a very high value of $K_\mathrm{D}^\mathrm{m2}[V328,1]$. This made it difficult to estimate $K_\mathrm{D}^\mathrm{m2}[V328,3{-}5]$. We therefore estimated $\tau_{K_\mathrm{D}^\mathrm{m2}}[V328]$ as the precision among the two values $K_\mathrm{D}^\mathrm{m2}[V328,1{-}2]$.

We estimated $\overline{K_\mathrm{D}^\mathrm{WT}}$ and $\overline{K_\mathrm{D}^\mathrm{m1/m2}}$ based on m measurements, $m = 1$, 2, 3, as

$$\ln(K_\mathrm{D}^\mathrm{WT}[m]) \sim Normal(\ln(\overline{K_\mathrm{D}^\mathrm{WT}}),\ \tau_{K_\mathrm{D}^\mathrm{WT}})$$

BIOPHYSICS

and

$$\ln(K_{\mathrm{D}}^{m1/m2}[m]) \sim Normal(\ln(\overline{K_{\mathrm{D}}^{m1/m2}}),\ \tau_{K_{\mathrm{D}}^{m1/m2}}),$$

respectively.

ACKNOWLEDGMENTS. This work was supported by grants from the Swedish Research Council (to P.J.); the Magnus Bergvall Foundation (to P.J.); the Linné Foundation for Medical Research (to P.J.); the O. E. and Edla Johansson Foundation (to P.J.); the Carl Trugger Foundation (to P.J.); the Medical Faculty, Uppsala University (to P.J.); and FORMAS (to T.S.).

1. Gagne SM, Li MX, Sykes BD (1997) Mechanism of direct coupling between binding and induced structural change in regulatory calcium binding proteins. *Biochemistry* 36:4386–4392.
2. Monod J, Wyman J, Changeux JP (1965) On the nature of allosteric transitions: A plausible model. *J Mol Biol* 12:88–118.
3. Koshland DE, Jr, Nemethy G, Filmer D (1966) Comparison of experimental binding data and theoretical models in proteins containing subunits. *Biochemistry* 5:365–385.
4. Volkman BF, Lipson D, Wemmer DE, Kern D (2001) Two-state allosteric behavior in a single-domain signaling protein. *Science* 291:2429–2433.
5. Cooper A, Dryden DT (1984) Allostery without conformational change. A plausible model. *Eur Biophys J* 11:103–109.
6. Gunasekaran K, Ma B, Nussinov R (2004) Is allostery an intrinsic property of all dynamic proteins? *Proteins* 57:433–443.
7. Lockless SW, Ranganathan R (1999) Evolutionarily conserved pathways of energetic connectivity in protein families. *Science* 286:295–299.
8. Fodor AA, Aldrich RW (2004) On evolutionary conservation of thermodynamic coupling in proteins. *J Biol Chem* 279:19046–19050.
9. Fodor AA, Aldrich RW (2004) Influence of conservation on calculations of amino acid covariance in multiple sequence alignments. *Proteins* 56:211–221.
10. Noivirt O, Eisenstein M, Horovitz A (2005) Detection and reduction of evolutionary noise in correlated mutation analysis. *Protein Eng Des Sel* 18:247–253.
11. Fuchs A, *et al.* (2007) Co-evolving residues in membrane proteins. *Bioinformatics* 23:3312–3319.
12. Gianni S, *et al.* (2006) Demonstration of long-range interactions in a PDZ domain by NMR, kinetics, and protein engineering. *Structure* 14:1801–1809.
13. Doyle DA, *et al.* (1996) Crystal structures of a complexed and peptide-free membrane protein-binding domain: Molecular basis of peptide recognition by PDZ. *Cell* 85:1067–1076.
14. Fuentes EJ, Gilmore SA, Mauldin RV, Lee AL (2006) Evaluation of energetic and dynamic coupling networks in a PDZ domain protein. *J Mol Biol* 364:337–351.
15. Schreiber G, Fersht AR (1995) Energetics of protein–protein interactions: Analysis of the barnase-barstar interface by single mutations and double mutant cycles. *J Mol Biol* 248:478–486.
16. Green SM, Shortle D (1993) Patterns of nonadditivity between pairs of stability mutations in staphylococcal nuclease. *Biochemistry* 32:10131–10139.
17. Ackers GK, Smith FR (1985) Effects of site-specific amino acid modification on protein interactions and biological function. *Annu Rev Biochem* 54:597–629.
18. Carter PJ, Winter G, Wilkinson AJ, Fersht AR (1984) The use of double mutants to detect structural changes in the active site of the tyrosyl-tRNA synthetase (Bacillus stearothermophilus). *Cell* 38:835–840.
19. Horovitz A, Fersht AR (1990) Strategy for analysing the co-operativity of intramolecular interactions in peptides and proteins. *J Mol Biol* 214:613–617.
20. Chi CN, Engström Å, Gianni S, Larsson M, Jemth P (2006) Two conserved residues govern the salt and pH dependencies of the binding reaction of a PDZ domain. *J Biol Chem* 281:36811–36818.
21. Fersht A (1999) *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding* (Freeman,).
22. Faiman GA, Horovitz A (1996) On the choice of reference mutant states in the application of the double-mutant cycle method. *Protein Eng* 9:315–316.
23. Dima RI, Thirumalai D (2006) Determination of network of residues that regulate allostery in protein families using sequence analysis. *Protein Sci* 15:258–268.
24. Kern D, Zuiderweg ER (2003) The role of dynamics in allosteric regulation. *Curr Opin Struct Biol* 13:748–757.
25. Swain JF, Gierasch LM (2006) The changing landscape of protein allostery. *Curr Opin Struct Biol* 16:102–108.
26. Rousseau F, Schymkowitz J (2005) A systems biology perspective on protein structural dynamics and signal transduction. *Curr Opin Struct Biol* 15:23–30.
27. Mittermaier A, Kay LE (2006) New tools provide new insights in NMR studies of protein dynamics. *Science* 312:224–228.
28. Eisenmesser EZ, Bosco DA, Akke M, Kern D (2002) Enzyme dynamics during catalysis. *Science* 295:1520–1523.
29. Henzler-Wildman KA, *et al.* (2007) A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature* 450:913–916.
30. Henzler-Wildman KA, *et al.* (2007) Intrinsic motions along an enzymatic reaction trajectory. *Nature* 450:838–844.
31. Akke M, Bruschweiler R, Palmer AG, III (1993) Effects of ion binding on the backbone dynamics of calbindin D9k determined by 15N NMR relaxation. *J Am Chem Soc* 115:9832–9833.
32. Popovych N, Sun S, Ebright RH, Kalodimos CG (2006) Dynamically driven protein allostery. *Nat Struct Mol Biol* 13:831–838.
33. Pargellis C, *et al.* (2002) Inhibition of p38 MAP kinase by utilizing a novel allosteric binding site. *Nat Struct Mol Biol* 9:268–272.
34. Schotte F, *et al.* (2003) Watching a protein as it functions with 150-ps time-resolved x-ray crystallography. *Science* 300:1944–1947.
35. Forsen S, Linse S (1995) Cooperativity: Over the Hill. *Trends Biochem Sci* 20:495–497.
36. Fersht AR, Sato S (2004) Phi-value analysis and the nature of protein-folding transition states. *Proc Natl Acad Sci USA* 101:7976–7981.
37. Jemth P, Gianni S (2007) PDZ domains: folding and binding. *Biochemistry* 46:8701–8708.
38. Peterson FC, Penkert RR, Volkman BF, Prehoda KE (2004) Cdc42 regulates the Par-6 PDZ domain through an allosteric CRIB-PDZ transition. *Mol Cell* 13:665–676.
39. Grembecka J, *et al.* (2006) The binding of the PDZ tandem of syntenin to target proteins. *Biochemistry* 45:3674–3683.
40. Suel GM, Lockless SW, Wall MA, Ranganathan R (2003) Evolutionarily conserved networks of residues mediate allosteric communication in proteins. *Nat Struct Biol* 10:59–69.
41. Gianni S, *et al.* (2005) The kinetics of PDZ domain-ligand interactions and implications for the binding mechanism. *J Biol Chem* 280:34805–34812.
42. Chi CN, *et al.* (2007) A conserved folding mechanism for PDZ domains. *FEBS Lett* 581:1109–1113.
43. Malatesta F (2005) The study of bimolecular reactions under non-pseudo-first order conditions. *Biophys Chem* 116:251–256.
44. Gelman A, Carlin JB, Stern HS, Rubin DB (2004) *Bayesian Data Analysis* (Chapman & Hall/CRC, Boca Raton, FL).
45. Gelman A, Hill J (2007) *Data Analysis Using Regression and Multilevel/Hierarchical Models* (Cambridge Univ Press, Cambridge, UK).
46. Gilks WR, Richardson S, Spiegelhalter DJ (1996) *Markov Chain Monte Carlo in Practice* (Chapman & Hall, London).
47. DeLano WL (2002) PyMOL (DeLano Scientific, San Carlos, CA).