Quantitative Relation between Intermolecular and Intramolecular Binding of Pro-Rich Peptides to SH3 Domains

Huan-Xiang Zhou
Department of Physics and Institute of Molecular Biophysics and School of Computational Science, Florida State University, Tallahassee, Florida

ABSTRACT Flexible linkers are often found to tether binding sequence motifs or connect protein domains. Here we analyze three usages of flexible linkers: 1), intramolecular binding of proline-rich peptides (PRPs) to SH3 domains for kinase regulation; 2), intramolecular binding of PRP for increasing the folding stability of SH3 domains; and 3), covalent linking of PRPs and other ligands for high-affinity bivalent binding. The basis of these analyses is a quantitative relation between intermolecular and intramolecular binding constants. This relation has the form $K_i = K_{e0}p$ for intramolecular binding and $K_e = K_{e01}K_{e02}p$ for bivalent binding. The effective concentration p depends on the length of the linker and the distance between the linker attachment points in the bound state. Several applications illustrate the usefulness of the quantitative relation. These include intramolecular binding to the Itk SH3 domain by an internal PRP and to a circular permutant of the α-spectrin SH3 domain by a designed PRP, and bivalent binding to the two SH3 domains of Grb2 by two linked PRPs. These and other examples suggest that flexible linkers and sequence motifs tethered to them, like folded protein domains, are also subject to tight control during evolution.

INTRODUCTION

Unstructured regions of proteins increasingly are found to play regulatory and signaling roles (1,2). One mode of action is through the binding of a peptide motif within an unstructured region to a target protein domain. Many such peptide motifs have been identified (3). Prominent among these is a proline-rich peptide (PRP) that binds to SH3 domains. X-ray structures of Src, Hck, and Abl kinases (4–8) show that intramolecular binding to the SH3 domain by a PRP within the linker between the SH2 and kinase domains provides an important mechanism for suppression of kinase activity (Fig. 1 *a*). These structures also immediately suggest a mechanism for kinase activation through intermolecular PRP-binding to the SH3 domain, leading to the release of the SH3 domain from the intramolecular PRP (Fig. 1 *b*). Previously we have developed a quantitative relation between intermolecular and intramolecular binding and applied the relation to examine the effects of covalent linking in protein folding stability (9) and in protein-DNA and antibody-antigen binding affinity (10,11) as well as the role of intramolecular binding of the myristoylated N-terminal in Abl kinase regulation (12). In this article we apply the relation to rationalize a number of recent experimental results on intramolecular PRP-SH3 binding in kinase regulation and on enhancement of folding stability and binding affinity through covalent linking of PRP ligands.

One type of intramolecular binding studied here is between a PRP connected to an SH3 domain via a flexible linker (Fig. 1 *c*). In particular, this models the arrangement of PRP and SH3 domains in the Tec family of nonreceptor tyrosine kinases (13–17) and in the diphtheria toxin repressor (18). The quantitative relation between intermolecular and intramolecular binding suggests that, while there is evolutionary pressure to increase the binding affinity of SH3 domains for external PRP from target proteins, it may be advantageous to maintain the affinity for internal PRP at a relatively low level. Hence in Src, Hck, and Abl kinases the sequences of internal PRP deviate significantly from those of high-affinity ligands (19,20). Similarly, a PRP within Btk, a member in the Tec family of kinases, was found to bind with the SH3 domains of Fyn, Hck, and Lyn kinases, but not with Btk's own SH3 domain (21). Indeed, the binding constants of Tec family kinases for their internal PRPs are typically found to be of the order of 10^3 M^{-1} (15–17). In contrast, the affinity of SH3 domains for PRPs on target proteins can reach 10^6 M^{-1} or higher (22–26).

The binding site for PRP is formed only when the SH3 is folded. Therefore in a folding transition PRP-binding can shift the equilibrium from the unfolded to folded state. The presence of an intramolecularly bound PRP is thus expected to increase the folding stability of the SH3 domain. Such stabilization has been observed in SH3 domains with designed intramolecular PRPs (27,28). The quantitative relation between intermolecular and intramolecular binding allows for calculation of the effect of the intramolecular PRPs on the unfolding free energy.

When a protein has multiple binding sites, a high-affinity ligand can be obtained by covalently linking peptide motifs that target the individual binding sites. Such bivalent ligands have been designed to simultaneously bind to the SH3 and SH2 domains of Abl kinase (29,30), the SH2 and kinase domains of Src kinase (31,32), and the two SH3 domains of Grb2 (33–35). In each of these cases there is strong experimental evidence indicating rearrangement between the domains upon ligand binding (6,32,34,36). Previously the

Submitted May 30, 2006, and accepted for publication July 24, 2006.

Address reprint requests to H.-X. Zhou, Tel.: 850-645-1336; E-mail: zhou@sb.fsu.edu.

0006-3495/06/11/3170/12 $2.00

doi: 10.1529/biophysj.106.090258

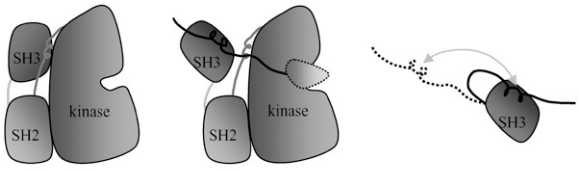

FIGURE 1 PRP-binding to SH3 domains in kinase regulation. (*a*) Suppression of kinase activity through intramolecular binding to the SH3 domain by a PRP in the SH2-kinase linker of a Src-like kinase. (*b*) Kinase activation through intermolecular PRP binding to the SH3 domain. This binding leads to the release of the internal ligand and activation of the kinase. The external PRP may reside in a regulatory protein; upon kinase activation a separate substrate will subsequently bind to the kinase domain. Alternatively, the external PRP may be part of a substrate protein. Binding of the PRP then also serves to anchor the substrate on the kinase. (*c*) Intramolecular binding and unbinding of a PRP to an SH3 domain. The PRP is connected to the SH3 domain by a flexible linker.

affinity-enhancement role of covalent linking has been analyzed for one of these ligands without considering the effect of domain rearrangement (11). Here this effect is incorporated to analyze the affinity of bivalent ligands. The analysis helps dispel misconception regarding bivalent binding that has appeared in the literature and provides an indication on the range of affinity enhancement expected from bivalent binding.

THEORY

Binding of PRP to SH3 domain, with and without flexible linker

When a PRP is attached to an SH3 domain by a flexible linker, the intramolecular binding constant K_i can be related to the intermolecular binding constant K_{e0}, observed when the linker is cleaved (Fig. 2 *a*). Under the assumption that the PRP and the SH3 domain do not interfere with the statistical distribution of the linker, except that the binding restrains the end-to-end distance of the linker to a distance d, it was shown (10,37)

$$K_i = K_{e0}p(d), \tag{1}$$

where $p(d)$ is the probability density for the end-to-end vector of the linker at the distance d. The value $p(d)$ plays the role of effective concentration. It has been found that peptide linkers can be modeled well as wormlike chains (38). For a linker consisting of L peptide bonds, an approximate expression for $p(d)$ is then

$$p_0(d;L) = (3/4\pi l_p l_c)^{3/2}\exp(-3d^2/4l_p l_c)(1 - 5l_p/4l_c + 2d^2/l_c^2 - 33d^4/80l_p l_c^3 - 79l_p^2/160l_c^2 - 329d^2 l_p/120l_c^3 + 6799d^4/1600l_c^4 - 3441d^6/2800l_p l_c^5 + 1089d^8/12800l_p^2 l_c^6), \tag{2}$$

where l_c and l_p are the contour length and persistence length, respectively, of the peptide linker. The contour length can be calculated as bL, where $b = 3.8$ Å is the nearest C_α-C_α distance. The persistence length was found to be ~3 Å for peptide linkers.

Cross-binding and dimerization

An SH3 domain can bind with not only its internal PRP but also the PRP on another molecule. Let us consider the general case of cross-binding between two different types of PRP-SH3 molecules. As illustrated in Fig. 2 *b*, there are a total of seven distinct monomeric and dimeric species: bound and unbound molecule 1 (denoted as M_{b1} and M_{u1}), bound and unbound molecule 2 (denoted as M_{b2} and M_{u2}), two types of singly bound dimers (denoted as D_{bu} and D_{ub}), and doubly bound dimer (denoted as D_{bb}). Higher oligomers can also form, but is not considered here. With the binding constants shown in Fig. 2 *b*, the concentrations of the seven species are related by

$$[M_{b1}] = K_{i1}[M_{u1}] \tag{3a}$$

$$[M_{b2}] = K_{i2}[M_{u2}] \tag{3b}$$

$$[D_{bu}] = K_{e1}[M_{b1}][M_{b2}] \tag{3c}$$

$$[D_{ub}] = K_{e2}[M_{b1}][M_{b2}] \tag{3d}$$

$$[D_{bb}] = K'_{i2}[D_{bu}] = K'_{i1}[D_{ub}]. \tag{3e}$$

Note that Eqs. 3c–e require

$$K_{e1}K'_{i2} = K_{e2}K'_{i1}. \tag{4}$$

When the two types of PRP-SH3 molecules are present at the same total concentration C_t, one finds

$$(1 + K_{i1})[M_{u1}] = (1 + K_{i2})[M_{u2}] \tag{5a}$$

$$= \frac{-1 + (1 + 8K_aC_t)^{1/2}}{4K_dC_t}C_t \equiv f(K_aC_t)C_t, \tag{5b}$$

where $f(K_aC_t)$ is the monomer fraction in a simple monomer-dimer equilibrium

$$2M \overset{K_a}{\rightleftarrows} D.$$

The apparent dimerization constant in the present situation is given by

$$K_a = (K_{e1} + K_{e2} + K_{e1}K'_{i2})/2(1 + K_{i1})(1 + K_{i2}). \tag{6}$$

Suppose that the binding equilibria are monitored by a signal-like chemical shift. For the simple monomer-dimer equilibrium, if the signals of pure monomer and pure dimer are s_m and s_d, then the signal at a total monomer concentration of C_t is

$$s = s_m f(K_aC_t) + s_d[1 - f(K_aC_t)] = (s_m - s_d)f(K_aC_t) + s_d. \tag{7}$$

Let us assume that, for each type of SH3 domain, the bound species all give the same signal $s_{b\alpha}$ ($\alpha = 1$ or 2) and the unbound species all give the same signal $s_{u\alpha}$. At equilibrium the actual signals from the two types of molecules are

$$s_1 = s_{u1}([M_{u1}] + [D_{ub}])/C_t + s_{b1}([M_{b1}] + [D_{bu}] + [D_{bb}])/C_t, \tag{8a}$$

$$s_2 = s_{u2}([M_{u2}] + [D_{bu}])/C_t + s_{b2}([M_{b2}] + [D_{ub}] + [D_{bb}])/C_t. \tag{8b}$$

These can be expressed in the same form as Eq. 7, with the apparent monomer and dimer signals given by

$$s_{m\alpha} = s_{b\alpha} + \frac{s_{u\alpha} - s_{b\alpha}}{(1 + K_{i\alpha})}, \tag{9a}$$

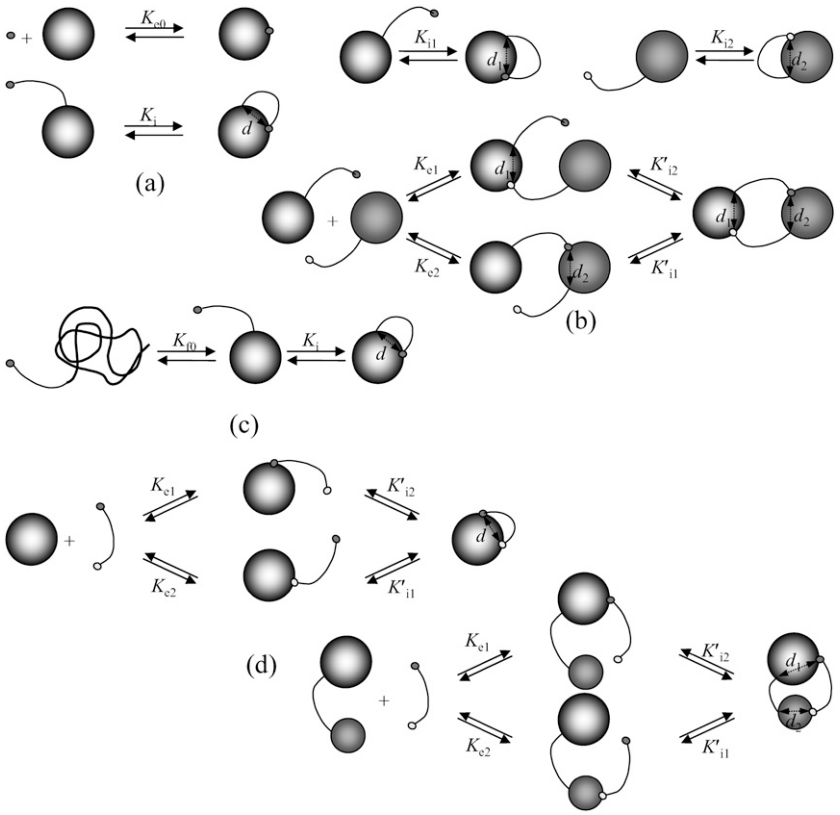

FIGURE 2 Models of intermolecular and intramolecular PRP-SH3 binding. (*a*) PRP-SH3 binding in the absence and presence of a flexible linker. (*b*) Binding of an SH3 domain with its internal PRP or with a PRP attached to another SH3 domain. The latter leads to dimerization of the two SH3 domains. (*c*) Intramolecular binding preceded by a folding transition of the SH3 domain. (*d*) Binding of a bivalent ligand to two sites on the same target or on two targets connected by a flexible linker.

$$s_{d\alpha} = s_{b\alpha} + \frac{(s_{u\alpha} - s_{b\alpha})K_{e1}K_{e2}}{K_{e\alpha}(K_{e1} + K_{e2} + K_{e1}K'_{i2})}, \tag{9b}$$

for $\alpha = 1$ or 2.

The cross-binding model of Fig. 2 *b* has six binding constants. To gain further insight on these binding constants, let us consider a control experiment in which the linkers on the two types of PRP-SH3 molecules are cleaved and bimolecular binding of the two types of PRP to each of the two types of SH3 domains is then observed. Let the binding constants of SH3 domain α for the internal and external PRP be $K_{e\alpha 0}$ and $K_{e'\alpha 0}$, respectively. The cross-binding constant $K_{e\alpha}$ can be approximated by $K_{e'\alpha 0}$. The intramolecular binding constant $K_{i\alpha}$ can be related to $K_{e\alpha 0}$ through Eq. 1,

$$K_{i\alpha} = K_{e\alpha 0}p_0(d_\alpha; L_\alpha). \tag{10a}$$

Here, d_α is the distance between the attachment point of the linker and the PRP binding site on SH3 domain α, and L_α is the linker length. An analogous relation between the intramolecular binding constant K'_{i1}, for forming D_{bb} from $D_{ub,}$ and K_{e10} can be proposed,

$$K'_{i1} = K_{e10}p_1(d_1; L_1, d_2, L_2), \tag{10b}$$

where $p_1(d_1;L_1,d_2,L_2)$ is the probability density for the end-to-end vector of a composite linker, consisting of the flexible peptide (with length L_1) attached to the first SH3 domain, a rigid connector with length d_2, and the flexible peptide (with length L_2) attached to the second SH3 domain. Similarly,

$$K'_{i2} = K_{e20}p_1(d_2; L_2, d_1, L_1). \tag{10c}$$

Since the two SH3 domains will have excluded-volume interactions and perhaps even specific interactions, the last two relations may not be very accurate. It can be shown that $p_1(d_1;L_1,d_2,L_2) = p_1(d_2;L_2,d_1,L_1)$, consequently the last two relations are consistent with the requirement of Eq. 4, $K_{e1}K'_{i2} = K_{e2}K'_{i1}$.

A PRP-SH3 molecule can always cross-bind with another molecule of the same type. The homodimerization process involves four distinct species: bound and unbound monomer (denoted as M_b and M_u), singly-bound dimer (denoted as D_s), and doubly-bound dimer (denoted as D_d). Since occupation of the PRP binding site on either M_u molecule leads to D_s, the binding constant for forming D_s from two M_u molecules is $2K_e$, where $K_e \approx K_{e0}$ is the binding constant if only one of the two identical binding sites were available. By the same token, release of either PRP from D_d leads to D_s, hence the binding constant for forming D_d from D_s is $K'_i/2$, where $K'_i \approx K_{e0}p_1(d;L,d,L)$ is the binding constant if only one of the two identical PRP on D_d were allowed to release. With these considerations, the results derived earlier for heterodimerization can be easily adapted for the present homodimerization process. The apparent dimerization constant is

$$K_a = (2K_e + K_eK'_i)/(1 + K_i)^2. \tag{11}$$

The signals at zero and infinite protein concentration are

$$s_m = s_b + \frac{s_u - s_b}{1 + K_i}; s_d = s_b + \frac{(s_u - s_b)}{2 + K'_i}. \tag{12}$$

Folding of PRP-SH3 molecule

To understand the equilibrium of a PRP-SH3 molecule between the unfolded and folded states, it is convenient to introduce an intermediate (either fictitious or real), in which the SH3 is folded by itself but the PRP is not bound (Fig. 2 *c*). If the folding equilibrium constant for the SH3 domain alone is K_{f0} and the intramolecular PRP binding constant is K_i, then the overall folding equilibrium constant is

$$K_f = K_{f0}K_i. \quad (13)$$

K_{f0} can be mimicked by the folding equilibrium constant of the SH3 with the PRP cleaved. The effect of the PRP on the unfolding free energy is thus

$$\Delta\Delta G_u = k_BT \ln K_f - k_BT \ln K_{f0} = k_BT \ln K_i. \quad (14a)$$

K_i can be obtained, by Eq. 1, from the intermolecular binding constant K_{e0} of the SH3 domain for the cleaved PRP. Hence,

$$\Delta\Delta G_u = k_BT \ln [K_{e0}p(d)]. \quad (14b)$$

Since the SH3 domain autonomously folds and the PRP binding site is presented only after the SH3 domain is folded, it is reasonable to consider the intermediate as real and on-pathway. It can be further suggested that the transition state for the SH3 domain is also the transition state for the PRP-SH3 molecule. Then the overall folding and unfolding rate constants are related to those of the SH3 domain by

$$k_f = k_{f0}, \quad (15a)$$

$$k_u = k_{u0}/K_i. \quad (15b)$$

That is, the PRP only affects the unfolding rate.

Binding of bivalent ligand

Consider two ligands that bind to separate binding sites on the same target protein with binding constants K_{e01} and K_{e02}. If the two ligands are connected by a linker (Fig. 2 *d*), then the bivalent ligand will have affinity

$$K_e = K_{e1}K_{i2} = K_{e2}K_{i1}. \quad (16a)$$

With $K_{e1} \approx K_{e10}$ and Eq. 1 for K_{i2}, one arrives at the result derived previously (10,11),

$$K_e = K_{e10}K_{e20}p(d), \quad (16b)$$

where $p(d)$ is the probability density for the end-to-end vector of the linker at the distance d separating the two binding sites.

When the two binding sites are located on two domains connected by a flexible linker (Fig. 2 *d*), the binding constant between the linked domains and the bivalent ligand is

$$K_e = K_{e1}K_{i'2} = K_{e2}K_{i'1}. \quad (17a)$$

With $K_{e2} \approx K_{e20}$ and Eq. 10c for K'_{i1}, one arrives at

$$K_e = K_{e10}K_{e20}p_1(d_1;L_1,d_2,L_2), \quad (17b)$$

where L_1 and L_2 are the lengths of the linkers connecting the two target domains and the two ligands, d_α is the distance between the attachment point of the interdomain linker and the ligand-binding site on domain α, and the meaning of the probability density $p_1(d_1;L_1,d_2,L_2)$ has been given after Eq. 10b. This formulation can be further extended when the two domains are not connected directly by a single flexible linker but by linkers from both sides to yet a third domain, like the situation between the two SH3 domains in Grb2.

Calculation of probability density

In previous studies of the relation between intermolecular and intramolecular binding (9–12), the probability density $p(d)$ was estimated by Eq. 2. In addition, the peptide linker was allowed to start in any direction off the attachment point on a protein domain or a ligand. To improve on both aspects, here we used conformation sampling of the peptide linker to calculate $p(d)$.

The wormlike chain modeling a peptide linker was represented as a freely rotating chain with a bond-length approaching zero and a bond-angle approaching 180° (39). In our implementation, each chain was represented by $J = 1000$ bonds, with a bond-length $s = l_c/J$ and a bond-angle of $\theta = \text{acos}(s/l_p - 1)$. The C_α-C_α vectors immediately before and immediately after the peptide linker were used to enforce directionality of the start bond (i.e., bond *1*) and end bond (i.e., bond *J*) of the linker. For concreteness, suppose that sequentially the SH3 domain is followed by the linker and then by the PRP. Let the residue number of the last SH3 residue be n_s and the residue number of the first PRP residue be n_e. The C_α-C_α vector $\mathbf{v}_s$ from residue $n_s - 1$ to residue n_s constrains the direction of bond 1 of the linker, and the C_α-C_α vector $\mathbf{v}_e$ from residue n_e to residue $n_e + 1$ constrains the direction of bond J of the linker. In the unbound state, $\mathbf{v}_s$ and $\mathbf{v}_e$ would have arbitrary relative orientations, but in the bound state the relative orientation of $\mathbf{v}_s$ and $\mathbf{v}_e$ is fixed. In addition, the end-to-end vector of the linker, identical to the C_α-C_α vector $\mathbf{d}$ from residue n_s to residue n_e, is also fixed relative to $\mathbf{v}_s$ and $\mathbf{v}_e$. To emphasize that the directionality of the start and end bonds of the linker is constrained, the probability density of the end-to-end vector is now denoted as $p(\mathbf{d})$.

The enforcement of directionality for the start and end bonds and the calculation of $p(\mathbf{d})$ were implemented as follows. A wormlike chain was started at the origin with bond 1 at an arbitrary direction. After growing the chain to bond J, the end-to-end vector of the chain was lined up with $\mathbf{d}$. Upon this lineup, if the angle between bond 1 and vector $\mathbf{v}_s$ and the angle between bond J and vector $\mathbf{v}_e$ were both less than a constraint angle γ_c, then the chain conformation was accepted. For each calculation of $p(\mathbf{d})$, a total of $N_{conf} = 10^7$ chain conformations were started. Of these, the number of chains with start and end bonds satisfying directionality constraints and with end-to-end distances falling between $d - \Delta/2$ and $d + \Delta/2$ was obtained. Let this be $n(d)$. The probability density $p(\mathbf{d})$ was found as $n(d)/4\pi d^2\Delta[(1 - \cos\gamma_c)/2]^2N_{conf}$. The bin size Δ was $10^{-2}l_c$.

The corresponding probability density for a composite linker, with directionally constrained start and end bonds for the flexible components, is now denoted as $p_1(\mathbf{d}_1;L_1,\mathbf{d}_2,L_2)$. The composite linker consists of a flexible chain with length L_1, a rigid connector spanned by vector $\mathbf{d}_2$, and another flexible chain with length L_2. In the bound state, the two flexible chains link up two domains. The directions of the start bond of chain 1 and the end bond of chain 2 are constrained by vectors $\mathbf{v}_{s1}$ and $\mathbf{v}_{e1}$, respectively, which are fixed in the first domain. Similarly, the directions of the end bond of chain 1 and the start bond of chain 2 are constrained by vectors $\mathbf{v}_{e2}$ and $\mathbf{v}_{s2}$, respectively, which are fixed in the second domain.

The procedure for obtaining $p_1(\mathbf{d}_1;L_1,\mathbf{d}_2,L_2)$ was the same as that for $p(\mathbf{d})$, with vector $\mathbf{d}_2$ treated as a special bond. Each flexible chain was represented by $J = 1000$ bonds. After the J^{th} bond of chain 1, vector $\mathbf{v}_{e2}$ was randomly distributed within a cone that spans an angle γ_c around bond J. Vector $\mathbf{d}_2$ was then randomly distributed around vector $\mathbf{v}_{e2}$, keeping the angle between them as found in the second domain. The directions of $\mathbf{v}_{e2}$ and $\mathbf{d}_2$ allowed for unique positioning of $\mathbf{v}_{s2}$. Finally chain 2 was started off randomly within a cone that spans an angle γ_c around $\mathbf{v}_{s2}$. This procedure was extended to Grb2 by an additional flexible linker followed by a rigid domain.

A value of 90° was assigned to the constraint angle γ_c. The persistence length of any flexible chain connected to a PRP ligand was 3 Å. For linkers between two protein domains, the value of l_p was 10 times higher.

RESULTS

Intermolecular and intramolecular PRP binding of Itk, Rlk, and Btk SH3 domains

Andreotti et al. (13) found that the SH3 domain of Itk forms an intramolecular complex with the K^3KPLPPTP^{10} sequence located at the N-terminal (Fig. 3 *a*). Here numerical superscripts are used to denote residue numbers as found in PDB entry No. 1awj. The N-terminal of the SH3 domain is

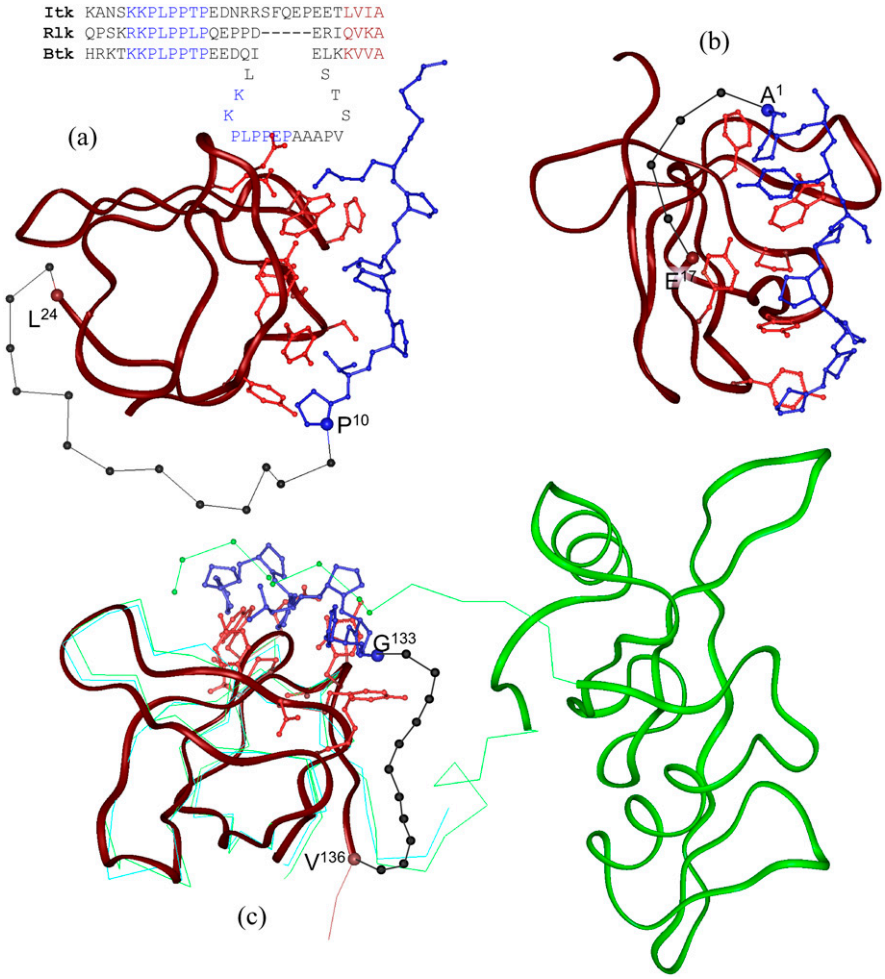

FIGURE 3 Intramolecular PRP-SH3 complexes. SH3 domains are shown as brown ribbon, PRP as blue ball-and-stick, linkers between them as black ball-and-stick, and SH3 residues in contact with PRP as red ball-and-stick. (*a*) Itk SH3 bound with an N-terminal PRP (PDB No. 1awj). Sequence alignment of the N-terminal portions of Itk, Rlk, and Btk are also shown. (*b*) S19P20s bound with P41 (PDB Nos. 1tuc and 1bbz). (*c*) Hck SH3 bound with GAP PRP (PDB No. 5hck). 1wa7 SH3 is shown as cyan trace; 1qcf is in green (SH3 as trace, SH2 as ribbon, PRP in SH2-kinase linker as ball-and-stick).

found on the opposite side of the PRP binding site. The linker has to make a sharp turn to tether the PRP. Laederach et al. (17) measured the binding constant K_{e0} of the cleaved PRP sequence, KNASKKPLPPTP, with the SH3 domain and found $K_{e0} = (0.17 \pm 0.02) \times 10^3\ M^{-1}$.

To apply Eq. 1 to predict the intramolecular binding constant K_i, the two ends of the flexible linker have to be specified. It seems reasonable to define the linker as the sequence, E^{11}DNRRSFQEPEET^{23}, between the PRP and the sharp turn at the N-terminal of the SH3 domain. This linker sequence is shown as black ball-and-stick in Fig. 3 *a*. The end-to-end distance d, measured between the C_α atoms of residues P^{10} and L^{24}, is 22.5 Å. With $L = 14$ peptide bonds, the effective concentration $p(\mathbf{d})$ was found by conformational sampling to be 13.5 mM. Equation 1 then predicts $K_i = 0.17 \times 10^3\ M^{-1} \times 13.5\ mM = 2.3$. This is to be compared with the experimental value of 0.54 ± 0.1 obtained by Laederach et al. (17).

Laederach et al. (17) also studied the effect of deleting five residues, S^{16}FQEP^{20}, of the linker sequence. With L reducing from 14 to 9, we find $p(\mathbf{d})$ lowering to 3 mM and predict $K_i = 0.5$. This is to be compared with the experimental value of 0.11 ± 0.05. The much lower value of $p(\mathbf{d})$ arises from the low probability for the end-to-end vector of the shortened linker to span a large distance of 22.5 Å.

Laederach et al. (17) did not find evidence for Itk PRP-SH3 dimerization at protein concentrations up to 1 mM. Equation 11 gives the apparent dimerization constant K_a. Neglecting the unknown K_i', we find $K_a \approx 2K_{e0}/(1 + K_i)^2 = 0.14 \times 10^3\ M^{-1}$. At $C_t = 1$ mM, the monomer fraction is expected to be 81% Eq. 5b. Apparently, weak intermolecular binding and relatively strong intramolecular binding work together to keep dimerization at a low level for Itk PRP-SH3.

Laederach et al. (17) also studied the binding of PRP sequence QPSKRKPLPPLP in the Rlk SH3 domain. The bimolecular binding constant of this PRP was (1.2 ± 0.1) × $10^3\ M^{-1}$. The intramolecularly bound fraction (i.e., $K_i/(1 + K_i)$) of the PRP was very low, at 0.07 ± 0.1. The corresponding K_i value is 0.075. Without a structure of the intramolecular complex, it is not possible to quantitatively explain the small K_i value. However, Laederach et al. (17) found the intramolecularly bound fraction of the QPSKRKPLPPLP sequence increasing to 0.25 ± 0.05 ($K_i = 0.33$) when five residues were inserted in the linker sequence; this increase

can be qualitatively explained by the expected increase in $p(\mathbf{d})$. From sequence alignment between Rlk and Itk over the PRP and SH3 regions (Fig. 3 *a*), the native linker sequence of Rlk may be identified as QEPPDERI. The insertion of five residues within this sequence increases the linker length from 9 to 14.

In a separate article, Laederach et al. (16) studied PRP binding of the Btk SH3 domain. This SH3 domain is preceded by two PRP sequences (KKPLPPTP and KKPLPPEP; Fig. 3 *a*). Only the first PRP, separated from the SH3 domain by a longer sequence, was found to be capable of intramolecular binding, even though both PRPs were found to be capable of cross-binding to a different SH3 domain. The increase in K_i, from not measurable for the second PRP to 0.29 for the first PRP, with the increase in linker length, is similar to the situations with Itk and Rlk kinases. The bimolecular binding constant K_{e0} can be obtained as half of the apparent dimerization constant of the Pr*PrSH3 mutant, which is incapable of intramolecular binding within either the monomer or the dimer (i.e., $K_i = K'_i = 0$), resulting in $K_{e0} = 0.4 \times 10^3\ M^{-1}$. On the other hand, Patel et al. (40) reported a binding constant of $18 \times 10^3\ M^{-1}$ for the peptide, TKKPLPPTPE, corresponding to the first PRP. The reason for the nearly 50-fold difference in K_{e0} is not clear; possibly the attached SH3 domain interferes with the cross-binding of the PRP to a separate SH3 domain.

In all the Tec family kinases, the bimolecular binding constant K_{e0} is found to be of the order of $10^3\ M^{-1}$. Since typical values of $p(\mathbf{d})$ are of the order of 1 mM, the resulting value of K_i will be ~1. Such a K_i will ensure that the PRP binding site is unoccupied by the internal ligand for a substantial fraction of time, thereby allowing for the displacement of the internal ligand by an external ligand. The $10^3\ M^{-1}$ order of magnitude for K_{e0} values of the Tec family kinases thus seems important for regulatory purposes.

Effects of intramolecular PRP binding on folding of S19P20s and HcK SH3 domains

Matrin-Sierra et al. (28) linked the P41 PRP sequence A^1PSYSPPPPP^{10} to the C-terminal of the circular permutant S19P20s of the α-spectrin SH3 domain and studied the effect on folding stability. The bimolecular binding constant of the PRP to S19P20s was $K_{e0} = 6.3 \times 10^3\ M^{-1}$.

The structure of the P41-S19P20s complex can be modeled after the complex of P41 with the Abl SH3 domain (PDB entry No. 1bbz) (41). The six P41-contacting residues (Y^7, F^9, W^{36}, W^{47}, P^{49}, and Y^{52}) in 1bbz can be superimposed to the corresponding residues (Y^{13}, Y^{15}, W^{41}, F^{52}, P^{54}, and Y^{57}) in S19P20s (PDB entry No. 1tuc) (42) with a backbone RMSD of only 0.2 Å. With this superposition, the C-terminal of S19P20s (E^{17} C_α) and the N-terminal of the P41 N-terminal (A^1 C_α) is 12.9 Å (Fig. 3 *b*). These two residues were connected by the linker sequence SGDN, resulting in a linker length of five peptide bonds. For the above d and L, $p(\mathbf{d})$ is 65 mM. Equation 14b then predicts stabilization of $\Delta\Delta G_u = 14.9$ kJ/mol by the covalently linked P41. This prediction is comparable with the experimental result of 8.3 kJ/mol.

Gmeiner et al. (27) appended the PRP sequence G^{131}GGFPPLPPPPYLPPLGAGL^{150} of human Ras GTPase-activating protein (GAP) to the Hck SH3 domain and studied the effect on unfolding rate by mass spectroscopy after H/D exchange. The particular SH3 construct contains part of the SH3-SH2 linker of Hck, ending with the sequence V^{136}DSLETEE^{143}, but the last five residues appear unstructured (PDB entry No. 5hck) (43). There is also a two-residue (AG) insertion between the SH3 construct and the GAP PRP. The bimolecular binding constant K_{e0} of the GAP PRP for the Hck SH3 domain was measured to be (2–10) $\times 10^4\ M^{-1}$ (43).

We searched the PDB for a template to build a homology model for the Hck SH3-GAP PRP complex and found entry No. 1wa7 for the purpose. This is the complex between the Fyn SH3 domain and the tyrosine-kinase interacting protein (Tip) PRP W^{170}DPGMPTPPLPPRPANLGERQA^{191} (44). There are only two substitutions between the proline-rich segments T^{176}PPLPPRP^{183} of Tip and F^{134}PPLPPPP^{141} of GAP. The backbone RMSD of the two SH3 domains is 1.2 Å after superimposing residues V^{84} to A^{134} of 5hck with residues V^{14} to A^{64} of 1wa7 (Fig. 3 *c*). This RMSD is very close to the value of 1.0 Å obtained between the structures of the Hck SH3 domain in isolation and in intact Hck kinase (PDB entry No. 1qcf) (7). The Tip PRP sequence G^{173}MPTPPLPPRPAN^{185} has a similar binding constant, $8.3 \times 10^4\ M^{-1}$ (44), as the GAP PRP for the Hck SH3 domain.

We take residue V^{136} of the Hck SH3 domain to be the attachment point of the linker to the GAP PRP. After this residue, conformations start to diverge among 5hck, 1wa7, and 1qcf. At the other end, the attachment point of the linker is taken to be G^{133} of the GAP PRP. This residue aligns with P^{175} of the Tip PRP. The resulting linker sequence is DSLETEEAGGG, with length $L = 12$ peptide bonds. The linker end-to-end distance in the bound structure is then measured from V^{136} of 5hck and P^{175} of the Tip PRP in 1wa7 after superposition of the SH3 domains. The result is $d = 21.3$ Å. For these values of L and d, we find $p(\mathbf{d}) = 13.7$ mM. The folding equilibrium constant is then increased by a factor of $K_i = K_{e0}p(\mathbf{d}) = 270$–$1370$ upon covalently linking the GAP PRP for intramolecular binding. If we assume that the folding rate is unaffected by the covalent linking, then the unfolding rate should be reduced by the above factor. By mass spectroscopy Gmeiner et al. (27) found a factor of ~40 for the slowing down of unfolding. It is possible that the experimental value for the decrease in unfolding was underestimated, as mass spectroscopy after H/D exchange might actually probe local instead of global unfolding.

Fig. 3 *c* shows that the internal PRP of Hck kinase is not as closely packed against the SH3 domain as the Tip PRP is. The internal PRP has the sequence K^{249}PQKPWE^{256} and the two lysine residues appear to be responsible for the loose

packing. Lerner et al. (20) found that mutating them into prolines significantly increases the intramolecular binding affinity.

Bivalent binding to Abl SH3-SH2, Src SH2-Kinase, and Grb2 SH3-SH3 domains

Xu et al. (30) constructed a bivalent ligand for the Abl SH3-SH2 domains (SH(32)) by connecting the C-terminal of SH3 ligand 3BP-2 (P^1PAYPPPPVP^{10}) and the N-terminal of SH2 ligand 2BP-1 (P^1V(Y_p)ENV^6; Y_p denotes phosphorylated tyrosine). Separately, the individual ligands have binding constants of $K_{e\alpha 0} = 9.5 \times 10^4$ and 5.0×10^5 M^{-1} (29), for the SH3 and SH2 domains. By connecting P^{10} of 3BP-2 and P^1 of 2BP-1 by an eight-residue linker, Xu et al. (30) found $K_e = 5.3 \times 10^6$ M^{-1} for the bivalent ligand. Fushman et al. (36) observed substantial changes in the overall spatial arrangement of the two domains upon binding a bivalent ligand. It is thus not appropriate to predict K_e by Eq. 16b, which would treat Abl SH(32) as rigid.

To use Eq. 17b to predict K_e, we need the distances between P^{10} of 3BP-2 and the attachment point of the SH3-SH2 linker on the SH3 domain and between P^1 of 2BP-1 and the attachment point of the SH3-SH2 linker on the SH2 domain. We take the two attachment points to be P^{137} and W^{146} (residue numbering according to PDB entry No. 2abl; Fig. 4 *a*) (45). The distances can be measured in the complex of the Abl SH3 domain and a 3BP-2 homolog (A^1PTM-PPPLPP^{10}) (PDB entry No. 1abo) (46) and the complex of Grb2 SH2 and a 2BP-1 homolog [$K^{-1}P^1F(Y_p)VNV^6EF$] (PDB entry No. 1bmb) (47). The results are $d_1 = 18.9$ Å and $d_2 = 21.3$ Å. The linker lengths are $L_1 = L_2 = 9$. The value of $p_1(\mathbf{d}_1;L_1,\mathbf{d}_2, L_2)$ is then 2.1 mM, leading to a prediction of $K_e = K_{e10}K_{e20}p_1(\mathbf{d}_1;L_1,\mathbf{d}_2, L_2) = 10^8$ M^{-1}. This is ~20-fold too high relative to the measured value. It is possible that the orientations of the SH3 and SH2 domains are not totally random, as we have modeled, upon binding the bivalent ligand. Orientational correlation will reduce $p_1(\mathbf{d}_1;L_1,\mathbf{d}_2, L_2)$ and hence the predicted K_e value.

Profit et al. (31,32) connected an SH2-targeting peptide (Y_p)EEIE with a kinase active-site directed inhibitor EE-LL(F_5Phe) by a tether consisting of γ-aminobutyric acid (GABA) residues. The ligands separately have binding constants $K_{e\alpha 0} = 10^6$ and 1.7×10^3 M^{-1} for the Src SH2 and kinase domains. The bivalent inhibitor with the highest affinity, 2.1×10^5 M^{-1}, has F_5Phe and Y_p linked by three GABA residues. In an active conformation of Src kinase in which the SH2 domain is dislocated from the kinase domain (PDB entry No. 1y57) (6), the SH2 binding site and the kinase active site are ~60 Å away (Fig. 4 *b*). The two

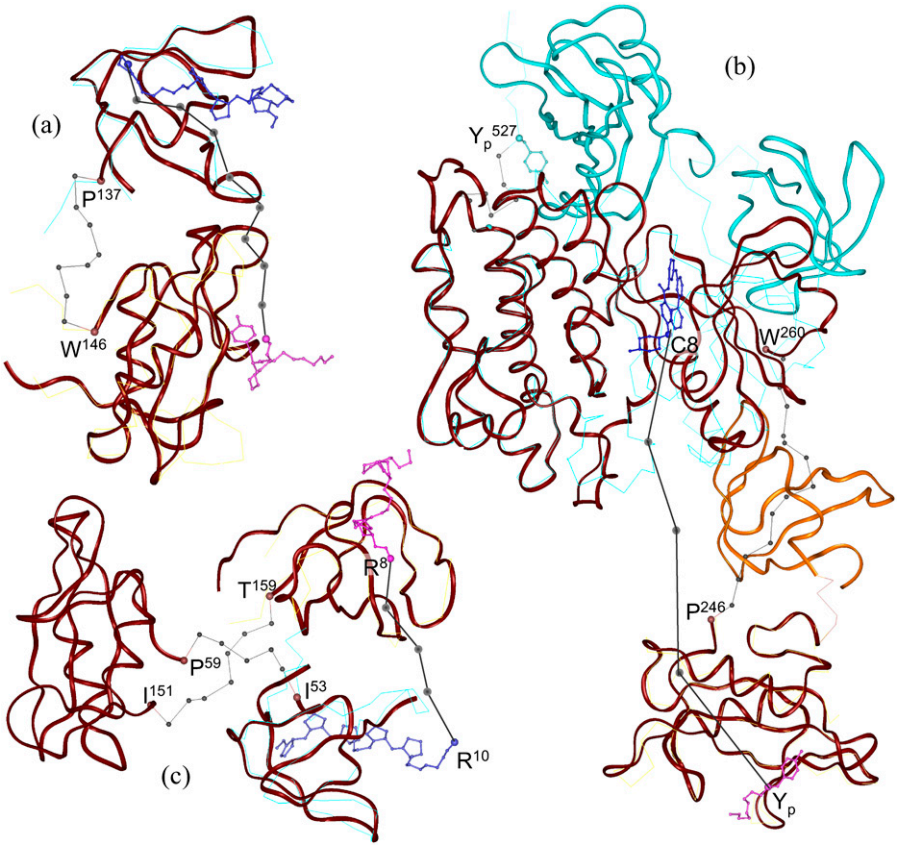

FIGURE 4 Complexes of multidomain proteins with bivalent ligands. Protein domains are shown as brown ribbon, individual ligands as blue or purple ball-and-stick, and linkers between protein domains and between individual ligands as black ball-and-stick. (*a*) Abl SH(32) bound with linked 3BP-2 and 2BP-1 (PDB No. 2abl). PDB Nos. 1abo SH3 and 1bmb SH2 are shown as cyan and yellow traces, respectively. (*b*) Src kinase bound with linked (Y_p)EEIE and EELL(F_5Phe) (PDB No. 1y57). The SH3 domain (gold ribbon) is bound to the PRP in the SH2-kinase linker. PDB No. 1shd SH2 is shown as yellow trace. Also shown in cyan is Src kinase in an inactivated state (PDB No. 2src), with SH3 and SH2 as ribbon and kinase domain as trace. Note the intramolecularly bound Y_p^{527}. (*c*) Grb2 with linked dimer of VPPPVPPRRR (PDB No. 1gri). PDB Nos. 3gbq nSH3 and 1sem cSH3 are shown as cyan and yellow traces, respectively.

domains must rearrange significantly to bring the two binding sites much closer to accommodate the bivalent ligand.

To apply Eq. 17b to predict K_e, we take the linker between the SH2 and kinase domains to be between residues P^{246} and W^{260} (residue numbering according to PDB entry No. 1y57 (6) (Fig. 4 *b*)). The structure of the complex between the SH2 domain and (Y_p)EEIE has been determined (PDB entry No. 1shd) (48). The distance between P^{246} and Y_p in this complex is $d_1 = 22.8$ Å. No structure for EELL(F_5Phe) bound to the kinase domain can be found. We therefore take the inhibitor found in 1y57 as a model of EELL(F_5Phe) for the purpose of calculating the distance d_2 and the orientation of the bound ligand. Specifically, d_2 is measured between P^{246} and atom C8 of the inhibitor in 1y57 and found to be 24.8 Å; the vector from atom C8 to atom C9 of the inhibitor is used to constrain the direction of the linker. With linker lengths of $L_1 = 14$ and $L_2 = 4$, we find $p_1(\mathbf{d}_1;L_1,\mathbf{d}_2, L_2) = 2.7$ mM. From this we predict $K_e = K_{e10}K_{e20}p_1(\mathbf{d}_1;L_1,\mathbf{d}_2, L_2) = 4.6 \times 10^6$ M^{-1}. This again is ~20-fold too high relative to the measured value.

Vidal et al. (35) constructed a bivalent ligand for the two SH3 domains of Grb2 by linking the C-terminals of two copies of V^1PPPVPPRRR^{10}. Individually the PRP has binding constants $K_{e\alpha 0} = 3.8 \times 10^5$ and 2.5×10^4 M^{-1} for the N- and C-terminal SH3 domains (nSH3 and cSH3) of Grb2. Upon linking two copies of the PRP by a single lysine the bivalent ligand has a binding constant of 2.5×10^7 M^{-1}. In the crystal structure of unbound Grb2 (PDB entry No. 1gri) (49), the two PRP binding sites are >25 Å apart (Fig. 4 *c*), and therefore cannot be spanned by a single-residue linker, in contrast to observations of Vidal et al. (35) and Yuzawa et al. (34). The latter authors also found experimental evidence indicating that, in solution, the two SH3 domains of Grb2 move freely relative to each other and the whole protein is less compact than the crystal structure.

The effective concentration for bivalent binding to Grb2 involves three flexible linkers: the first between nSH3 and the SH2 domain, the second between the SH2 domain and cSH3, and the third between the two PRP ligands (Fig. 4 *c*). We take the first linker to be between residues I^{53} and P^{59}, and the second linker to be between I^{151} and T^{159} (residue numbering according to 1gri). The corresponding linker lengths are $L_1 = 6$ and $L_2 = 8$, and the distance between the linker attachment points (residues P^{59} and I^{151}) in the SH2 domain is found to $d_2 = 7.0$ Å in 1gri. The distance d_1 in nSH3 between residue I^{53} and R^{10} of V^1PPPVPPRRR^{10} is measured from PDB entry No. 3gbq, which is the complex between nSH3 and the PRP ligand. The result is $d_1 = 19.2$ Å. The distance d_3 in cSH3 is measured on the complex of the PRP ligand and the *Caenorhabditis elegans* homolog of nSH3 (PDB entry No. 1sem) (50). As R^9 and R^{10} of the ligand are not visible in 1sem, d_3 is measured between the counterpart of Grb2 T^{159} and R^8 of the ligand. This gives $d_3 = 24.5$ Å. The last two residues of the ligand is then treated as part of the third flexible linker, leading to $L_3 = 4$. With these specifications, the effective concentration is found to be 8 mM. The binding constant of the bivalent ligand is then predicted to be 7.6×10^7 M^{-1}, which is comparable to the experimental result of 2.5×10^7 M^{-1}. This calculation is further supported by the experimental result of Vidal et al. (35) for another bivalent ligand, constructed from a peptoid analog of V^1PPPVPPRRR^{10}. For this ligand, the predicted binding constant is 4.4×10^9 M^{-1}, which is the nearly same as the experimental result of 5×10^9 M^{-1}.

Effects of linker length

In addition to the intramolecular binding between the SH3 domain and the PRP within the SH2-kinase linker, the Src family kinases utilize a second mechanism for suppression of kinase activity: binding of a phosphorylated tyrosine on the C-terminal tail to the SH2 domain. Cobb et al. (51) found that the linker length between the kinase domain and the intramolecular ligand, Y_p^{527}, has significant effect on the kinase activity. Both deletion and insertion of two or four residues resulted in kinase activation. If we take D^{518} and Y_p^{527} in the inactivated form of Src kinase (PDB entry No. 2src) (5) as the two attachment points of the linker, then the end-to-end distance is $d = 15.9$ Å and the native linker length is $L = 9$. Fig. 5 *a* shows that the effective concentration $p(\mathbf{d})$ is maximal at the native linker length. Decrease or increase in linker length through deletion or insertion leads to a lower effective

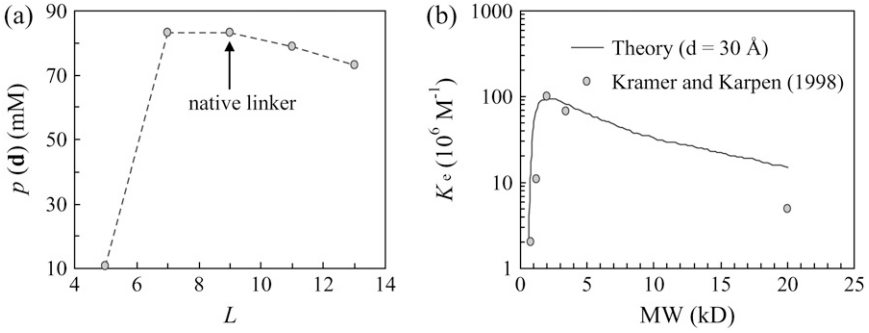

FIGURE 5 Effects of linker length on effective concentration. (*a*) Decrease of $p(\mathbf{d})$ with insertion or deletion in the linker for intramolecular Y_p^{527} binding to the Src kinase domain. (*b*) Variation of the binding affinity of PEG-linked cGMP dimer to olfactory cyclic-nucleotide-gated channel. PEG was modeled as a wormlike chain with persistence length of 3.5 Å (66,67) and contour length given by the number of monomer units (21.5 per 1000 molecular weight (68)) times a unit length of 2.8 Å (66).

concentration and thus a lower bound fraction. As a result the kinase activity would be increased.

Kramer and Karpen (52) developed highly potent polymer-linked ligand dimers for activating cyclic-nucleotide-gated channels and cyclic GMP-dependent protein kinase. They found that each target protein, with a given distance between the two binding sites for a bivalent ligand, has an optimal linker length. As illustrated in Fig. 5 *b* for the binding of PEG-linked cGMP dimer to olfactory cyclic-nucleotide-gated channel, the broad distribution of the binding constant with polymer length is consistent with the effect of linker length on the effective concentration. The structure of the channel is not known. Best fit to the experimental data for binding gives a distance of 30 Å between the two binding sites.

DISCUSSION

We have applied a quantitative relation between intermolecular and intramolecular binding to predict the intramolecular binding constant, the effect of an internal PRP ligand on the folding stability of an SH3 domain, and the enhanced affinity from covalently linking two ligands for separate binding sites. Despite the simplicity of the theoretical model, the predictions generally agree reasonably well with experimental results. These applications lead to several general conclusions, which we now outline.

For an SH3 domain that uses intramolecular PRP binding as a mechanism for regulation, the intramolecular binding constant K_i should probably be $\sim$1; values too high or too low would mean that the regulatory switch is on or off all the time. The applications presented here show that the effective concentration $p(\mathbf{d})$ is in the mM range. This requires that the intermolecular binding constant of the internal ligand is of the order of 10^3 M^{-1}. The requirement for such a low affinity explains why internal PRP ligands often have sequences that deviate significantly from ideal high-affinity sequences, a situation that has been termed ''self-restraint'' (19). This observation also applies to intramolecular binding of Y_p^{527} to the SH2 domain of Src Kinase; the affinity of the native Y_p^{527} sequence has been found to be 10,000-fold lower than that of an ideal SH2 ligand (53). Other examples of intramolecular binding in regulation include binding of the myristoylated N-terminal of Abl kinase to its kinase domain (8,54), binding of the receiver and effector domains of response regulator NarL (55), and binding of the pocket and C-terminal domains of the retinoblastoma protein (56).

Addition of ligands that favor the folded state over the unfolded state is a very common strategy for increasing protein stability. For an SH3 domain, covalently linking a PRP with a bimolecular binding constant of $\sim 10^6$ M^{-1} is predicted to increase the folding stability by 15–20 kcal/mol. As the stability of SH3 domains is often relatively low, one wonders whether intramolecular PRP binding provides an in vivo mechanism for stabilization. The same mechanism can apply to the stabilization of other types of protein domains that recognize different peptide motifs.

Covalent linking of two low-affinity ligands may result in a high-affinity bivalent ligand. It has sometimes been suggested that the affinity K_e of the bivalent ligand is expected to be the product of the affinities K_{e10} and K_{e20} of the separate ligands (29,57). Such a product will have units M^{-2}, inappropriate for a bimolecular binding constant K_e. The correct formulation through Eqs. 16b and 17b predicts K_e as the product of $K_{e10}K_{e20}$ and the effective concentration. The effective concentration depends on the distance between the two binding sites and the linker length. As seen for the binding of PEG-linked cGMP dimer to olfactory cyclic-nucleotide-gated channel, for a given distance between the binding sites, the linker length can be optimized. The effective concentration is in the mM range, thus as long as one ligand has an affinity greater than 10^3 M^{-1}, linking it to another ligand will enhance the affinity of the latter. For example, for two ligands with affinities of 10^5 and 10^6 M^{-1}, if linking them has an effective concentration of 1 mM for the two binding sites, the expected affinity for the bivalent ligand is 10^8 M^{-1}, which is 100- to 1000-fold higher than the affinities of the individual ligands.

The mM range of the effective concentration makes it easy to achieve affinity enhancement through covalent linking. Proteins like Src family kinases and Grb2 may actually utilize the separate binding sites on SH3 and SH2 domains for high-affinity bivalent binding. Flexible linkers between the domains allow the proteins to adapt to different targets. Affinity enhancement of bivalent binding has been observed in many other designed and natural systems (56–65).

Effect of competitor on dissociation kinetics of bivalent ligand

It is of interest to ask, in enhancing binding affinity, whether covalent linking of two separate ligands exerts its influence through the association rate or the dissociation rate. A previous kinetic analysis showed that the affinity enhancement is mainly manifested through decrease of the dissociation rate (11). This theoretical result is supported by experimental observations of Kramer and Karpen (52) and Walker et al. (61).

For a protein that is bound to a monovalent ligand, when a competitor for the ligand is introduced, the rate of exchanging to the competitor is the same as the dissociation rate of the original ligand. If a bivalent ligand behaves the same way, the exchange rate will be exceedingly small, given the effect of covalent linking on the dissociation rate of the bivalent ligand. However, both Kramer and Karpen (52) and Rao et al. (60) observed fast exchange in the presence of excess competitor.

The exchange between a bivalent ligand and a monovalent competitor is illustrated in Fig. 6. With the elementary rate

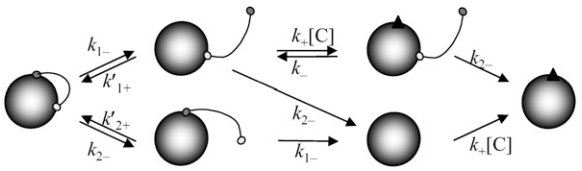

FIGURE 6 Exchange between a bivalent ligand and a monovalent competitor.

constants shown, it can be shown that the overall exchange rate constant is

$$k_{ex} = \frac{k_{1-}\{k_{2-} + k_+[C]/(1 + k_-/k_{2-})\}}{k'_{1+} + k_{2-} + k_+[C]/(1 + k_-/k_{2-})} + \frac{k_{1-}k_{2-}}{k_{1-} + k'_{2+}}, \quad (18)$$

where [C] is the concentration of the competitor. As [C] $\rightarrow \infty$, we have $k_{ex} \rightarrow k_{1-}$, the dissociation rate from one of the two binding sites, which is unaffected by covalent linking. Kramer and Karpen (52) studied the dissociation kinetics of a PEG-linked cGMP dimer from olfactory cyclic-nucleotide-gated channel in the presence of cGMP. The dependence of the dissociation rate on cGMP concentration appears to be in qualitative agreement with Eq. 18.

Further refinement of theoretical model

The model for linkers used here considers only statistical distributions of the end-to-end vector. Effects such as excluded-volume and other types of interactions between linkers and protein domains and conformational preferences of specific sequences may affect the effective concentration. Such detailed effects can be captured by molecular dynamics or Monte Carlo simulations. Given the encouraging results presented here, such further refinements seem well justified and will be pursued in the future.

In summary, based on a quantitative relation between intermolecular and intramolecular binding, we have calculated the intramolecular binding constant of the Itk SH3 domain with an internal PRP, the stabilization of a circular permutant of the α-spectrin SH3 domain by intramolecular binding of a designed PRP, and the affinity enhancement of a bivalent ligand for the two SH3 domains of Grb2. These and other examples suggest that flexible linkers and sequence motifs tethered to them, like folded protein domains, are also subject to tight control during evolution.

This work was supported in part by grant No. GM58187 from the National Institutes of Health.

REFERENCES

1. Oldfield, C. J., Y. Cheng, M. S. Cortese, P. Romero, V. N. Uversky, and A. K. Dunker. 2005. Coupled folding and binding with α-helix-forming molecular recognition elements. *Biochemistry*. 44:12454–12470.
2. Dyson, H. J., and P. E. Wright. 2005. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* 6:197–208.
3. Neduva, V., R. Linding, I. Su-Angrand, A. Stark, F. d. Masi, T. J. Gibson, J. Lewis, L. Serrano, and R. B. Russell. 2005. Systematic discovery of new recognition peptides mediating protein interaction networks. *PLoS Biol.* 3:e405.
4. Williams, J. C., A. Weijland, S. Gonfloni, A. Thompson, S. A. Courtneidge, G. Superti-Furga, and R. K. Wierenga. 1997. The 2.35 Å crystal structure of the inactivated form of chicken Src: a dynamic molecule with multiple regulatory interactions. *J. Mol. Biol.* 274: 757–775.
5. Xu, W., A. Doshi, M. Lei, M. J. Eck, and S. C. Harrison. 1999. Crystal structures of c-Src reveal features of its autoinhibitory mechanism. *Mol. Cell.* 3:629–638.
6. Cowan-Jacob, S. W., G. Fendrich, P. W. Manley, W. Jahnke, D. Fabbro, J. Liebetanz, and T. Meyer. 2005. The crystal structure of a c-Src complex in an active conformation suggests possible steps in c-Src activation. *Structure*. 13:861–871.
7. Schindler, T., F. Sicheri, A. Pico, A. Gazit, A. Levitzki, and J. Kuriyan. 1999. Crystal structure of Hck in complex with a Src family-selective tyrosine kinase inhibitor. *Mol. Cell.* 3:639–648.
8. Nagar, B., O. Hantschel, M. A. Young, K. Scheffzek, D. Veach, W. Bornmann, B. Clarkson, G. Superti-Furga, and J. Kuriyan. 2003. Structural basis for the autoinhibition of c-Abl tyrosine kinase. *Cell.* 112:859–871.
9. Zhou, H.-X. 2001. Single-chain versus dimeric protein folding: thermodynamic and kinetic consequences of covalent linkage. *J. Am. Chem. Soc.* 123:6730–6731.
10. Zhou, H.-X. 2001. The affinity-enhancing roles of flexible linkers in two-domain DNA-binding proteins. *Biochemistry*. 40:15069–15073.
11. Zhou, H.-X. 2003. Quantitative account of the enhanced affinity of two linked scFvs specific for different epitopes on the same antigen. *J. Mol. Biol.* 329:1–8.
12. Zhou, H.-X. 2003. How often does the myristoylated N-terminal latch of c-Abl come off? *FEBS Lett.* 552:160–162.
13. Andreotti, A. H., S. C. Bunnell, S. Feng, L. J. Berg, and S. L. Schreiber. 1997. Regulatory intramolecular association in a tyrosine kinase of the Tec family. *Nature*. 385:93–97.
14. Tzeng, S.-R., Y.-C. Lou, M.-T. Pai, M. L. Jain, and J.-W. Cheng. 2000. Solution structure of the human BTK SH3 domain complexed with a proline-rich peptide from p120cbl. *J. Biomol. NMR*. 16:303–312.
15. Pursglove, S. E., T. D. Mulhern, J. P. Mackay, M. G. Hinds, and G. W. Booker. 2002. The solution structure and intramolecular associations of the Tec kinase Src homology 3 domain. *J. Biol. Chem.* 277: 755–762.
16. Laederach, A., K. W. Cradic, K. N. Brazin, J. Zamoon, D. B. Fulton, X. Y. Huang, and A. H. Andreotti. 2002. Competing modes of self-association in the regulatory domains of Bruton's tyrosine kinase: intramolecular contact versus asymmetric homodimerization. *Protein Sci.* 11:36–45.
17. Laederach, A., K. W. Cradic, D. B. Fulton, and A. H. Andreotti. 2003. Determinants of intra versus intermolecular self-association within the regulatory domains of Rlk and Itk. *J. Mol. Biol.* 329:1011–1020.
18. Wylie, G. P., V. Rangachari, E. A. Bienkiewicz, V. Marin, N. Bhattacharya, J. F. Love, J. R. Murphy, and T. M. Logan. 2005. Prolylpeptide binding by the prokaryotic SH3-like domain of the diphtheria toxin repressor: a regulatory switch. *Biochemistry*. 44:40–51.
19. Nguyen, J. T., and W. A. Lim. 1997. How Src exercises self-restraint. *Nat. Struct. Biol.* 4:256–260.
20. Lerner, E. C., R. P. Trible, A. P. Schiavone, J. M. Hochrein, J. R. Engen, and T. E. Smithgall. 2005. Activation of the Src family kinase Hck without SH3-linker release. *J. Biol. Chem.* 280:40832–40837.
21. Cheng, G., Z. Ye, and D. Baltimore. 1994. Binding of Bruton's tyrosine kinase to Fyn, Lyn, or Hck through a Src homology 3 domain-mediated interaction. *Proc. Natl. Acad. Sci. USA*. 91:8152–8155.
22. Ghose, R., A. Shekhtman, M. J. Goger, H. Ji, and D. Cowburn. 2001. A novel, specific interaction involving the Csk SH3 domain and its natural ligand. *Nat. Struct. Biol.* 8:998–1004.

23. Zarrinpar, A., S.-H. Park, and W. A. Lim. 2003. Optimization of specificity in a cellular protein interaction network by negative selection. *Nature*. 426:676–680.
24. Marles, J. A., S. Dahesh, J. Haynes, B. J. Andrews, and A. R. Davidson. 2004. Protein-protein interaction affinity plays a crucial role in controlling the Sho1p-mediated signal transduction pathway in yeast. *Mol. Cell.* 14:813–823.
25. Mott, H. R., D. Nietlispach, K. A. Evetts, and D. Owen. 2005. Structural analysis of the SH3 domain of β-PIX and its interaction with α-p21 activated kinase (PAK). *Biochemistry*. 44:10977–10983.
26. Bauer, F., K. Schweimer, H. Meiselbach, S. Hoffmann, P. Rosch, and H. Sticht. 2005. Structural characterization of Lyn-SH3 domain in complex with a herpesviral protein reveals an extended recognition motif that enhances binding affinity. *Protein Sci.* 14:2487–2498.
27. Gmeiner, W. H., I. Xu, D. A. Horita, T. E. Smithgall, J. R. Engen, D. L. Smith, and R. A. Byrd. 2001. Intramolecular binding of a proximal PPII helix to an SH3 domain in the fusion protein $SH3_{Hck}$: $PPII_{hGAP}$. *Cell Biochem. Biophys.* 35:115–126.
28. Martin-Sierra, F. M., A. M. Candel, S. Casares, V. V. Filimonov, J. C. Martinez, and F. Conejero-Lara. 2003. A binding event converted into a folding event. *FEBS Lett.* 553:328–332.
29. Cowburn, D., J. Zheng, Q. Xu, and G. Barany. 1995. Enhanced affinities and specificities of consolidated ligands for the Src homology (SH) 3 and SH2 domains of Abelson protein-tyrosine kinase. *J. Biol. Chem.* 270:26738–26741.
30. Xu, Q., J. Zheng, R. Xu, G. Barany, and D. Cowburn. 1999. Flexibility of interdomain contacts revealed by topological isomers of bivalent consolidated ligands to the dual Src homology domain SH(32) of Abelson. *Biochemistry.* 38:3491–3497.
31. Profit, A. A., T. R. Lee, and D. S. Lawrence. 1999. Bivalent inhibitors of protein tyrosine kinases. *J. Am. Chem. Soc.* 121:280–283.
32. Profit, A. A., T. R. Lee, J. Niu, and D. S. Lawrence. 2001. Molecular rulers: an assessment of distance and spatial relationships of Src tyrosine kinase SH2 and active site regions. *J. Biol. Chem.* 276:9446–9451.
33. Cussac, D., M. Vidal, C. Leprince, W.-q. Liu, F. Cornille, G. Tiraboschi, B. P. Roques, and C. Garbay. 1999. A Sos-derived peptidimer blocks the Ras signaling pathway by binding both Grb2 SH3 domains and displays antiproliferative activity. *FASEB J.* 13:31–38.
34. Yuzawa, S., M. Yokochi, H. Hatanaka, K. Ogura, M. Kataoka, K. Miura, V. Mandiyan, J. Schlessinger, and F. Inagaki. 2001. Solution structure of Grb2 reveals extensive flexibility necessary for target recognition. *J. Mol. Biol.* 306:527–537.
35. Vidal, M., W.-Q. Liu, C. Lenoir, J. Salzmann, N. Gresh, and C. Garbay. 2004. Design of peptoid analogue dimers and measure of their affinity for Grb2 SH3 domains. *Biochemistry.* 43:7336–7344.
36. Fushman, D., R. Xu, and D. Cowburn. 1999. Direct determination of changes of interdomain orientation on ligation: use of the orientational dependence of ^{15}N NMR relaxation in Abl SH(32). *Biochemistry.* 38: 10225–10230.
37. Zhou, H.-X. 2004. Polymer models of protein stability, folding, and interactions. *Biochemistry.* 43:2141–2154.
38. Zhou, H. X. 2001. Loops in proteins can be modeled as worm-like chains. *J. Phys. Chem. B.* 105:6763–6766.
39. Cantor, C. R., and P. R. Schimmel. 1980. Biophysical Chemistry. W. H. Freeman, New York.
40. Patel, H. V., S.-R. Tzeng, C.-Y. Liao, S.-H. Chen, and J.-W. Cheng. 1997. SH3 domain of Bruton's tyrosine kinase can bind to proline-rich peptides of TH domain of the kinase and $p120^{cbl}$. *Proteins.* 29: 545–552.
41. Pisabarro, M. T., L. Serrano, and M. Wilmanns. 1998. Crystal structure of the abl-SH3 domain complexed with a designed high-affinity peptide ligand: implications for SH3-ligand interactions. *J. Mol. Biol.* 281:513–521.
42. Viguera, A. R., F. J. Blanco, and L. Serrano. 1995. The order of secondary structure elements does not determine the structure of a protein but does affect its folding kinetics. *J. Mol. Biol.* 247:670–681.
43. Horita, D. A., D. M. Baldisseri, W. Zhang, A. S. Altieri, T. E. Smithgall, W. H. Gmeiner, and R. A. Byrd. 1998. Solution structure of the human Hck SH3 domain and identification of its ligand binding site. *J. Mol. Biol.* 278:253–265.
44. Schweimer, K., S. Hoffmann, F. Bauer, U. Friedrich, C. Kardinal, S. M. Feller, B. Biesinger, and H. Sticht. 2002. Structural investigation of the binding of a herpesviral protein to the SH3 domain of tyrosine kinase Lck. *Biochemistry.* 41:5120–5130.
45. Nam, H. J., W. G. Haser, T. M. Roberts, and C. A. Frederick. 1996. Intramolecular interactions of the regulatory domains of the Bcr-Abl kinase reveal a novel control mechanism. *Structure.* 4:1105–1114.
46. Musacchio, A., M. Saraste, and M. Wilmanns. 1994. High-resolution crystal structures of tyrosine kinase SH3 domains complexed with proline-rich peptides. *Nat. Struct. Biol.* 1:546–551.
47. Ettmayer, P., D. France, J. Gounarides, M. Jarosinski, M.-S. Martin, J.-M. Rondeau, M. Sabio, S. Topiol, B. Weidmann, M. Zurini, and K. W. Bair. 1999. Structural and conformational requirements for high-affinity binding to the SH2 domain of Grb2. *J. Med. Chem.* 42: 971–980.
48. Gilmer, T., M. Rodriguez, S. Jordan, R. Crosby, K. Alligood, M. Green, M. Kimery, C. Wagner, D. Kinder, and P. Charifson. 1994. Peptide inhibitors of Src SH3–SH2-phosphoprotein interactions. *J. Biol. Chem.* 269:31711–31719.
49. Maignan, S., J. P. Guilloteau, N. Fromage, B. Arnoux, J. Becquart, and A. Ducruix. 1995. Crystal structure of the mammalian Grb2 adaptor. *Science.* 268:291–293.
50. Lim, W. A., F. M. Richards, and R. O. Fox. 1994. Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. *Nature.* 372:375–379.
51. Cobb, B. S., D. M. Payne, A. B. Reynolds, and J. T. Parsons. 1991. Regulation of the oncogenic activity of the cellular *src* protein requires the correct spacing between the kinase domain and the C-terminal phosphorylated tyrosine (Tyr-527). *Mol. Cell. Biol.* 11:5832–5838.
52. Kramer, R. H., and J. W. Karpen. 1998. Spanning binding sites on allosteric proteins with polymer-linked ligand dimers. *Nature.* 395: 710–713.
53. Bibbins, K. B., H. Boeuf, and H. E. Varmus. 1993. Binding of the Src SH2 domain to phosphopeptides is determined by residues in both the SH2 domain and the phosphopeptides. *Mol. Cell. Biol.* 13:7278–7287.
54. Hantschel, O., B. Nagar, S. Guettler, J. Kretzschmar, K. Dorey, J. Kuriyan, and G. Superti-Furga. 2003. A myristoyl/phosphotyrosine switch regulates c-Abl. *Cell.* 112:845–857.
55. Eldridge, A. M., H. S. Kang, E. Johnson, R. Gunsalus, and F. W. Dahlquist. 2002. Effect of phosphorylation on the interdomain interaction of the response regulator, NarL. *Biochemistry.* 41:15173–15180.
56. Rubin, S. M., A.-L. Gall, N. Zheng, and N. P. Pavletich. 2005. Structure of the Rb C-terminal domain bound to E2F1–DP1: a mechanism for phosphorylation-induced E2F release. *Cell.* 123:1093–1106.
57. Bachhawat-Sikder, K., and T. Kodadek. 2003. Mixed-element capture agents: a simple strategy for the construction of synthetic, high-affinity protein capture ligands. *J. Am. Chem. Soc.* 125:9550–9551.
58. Neri, D., M. Momo, T. Prospero, and G. Winter. 1995. High affinity antigen binding by chelating recombinant antibodies (CRAbs). *J. Mol. Biol.* 246:367–373.
59. Klemm, J. D., and C. O. Pabo. 1996. Oct-1 POU domain-DNA interactions: cooperative binding of isolated subdomains and effects of covalent linkage. *Genes Dev.* 10:27–36.
60. Rao, J., J. Lahiri, L. Isaacs, R. M. Weis, and G. M. Whitesides. 1998. A trivalent system from vancomycin·D-Ala-D-Ala with higher affinity than avidin·biotin. *Science.* 280:708–711.
61. Walker, D., G. R. Moore, R. James, and C. Kleanthous. 2003. Thermodynamic consequences of bipartite immunity protein binding to the ribosomal ribonuclease colicin E3. *Biochemistry.* 42:4161–4171.
62. Arunkumar, A. I., M. E. Stauffer, E. Bochkareva, A. Bochkarev, and W. J. Chazin. 2003. Independent and coordinated functions of

replication protein A tandem high affinity single stranded DNA binding domains. *J. Biol. Chem.* 278:41077–41082.

63. Lacy, E. R., I. Filippov, W. S. Lewis, S. Otieno, L. Xiao, S. Weiss, L. Hengst, and R. W. Kriwacki. 2004. p27 binds cyclin-CDK complexes through a sequential mechanism involving binding-induced protein folding. *Nat. Struct. Mol. Biol.* 11:358–364.
64. Ferguson, M. R., X. Fan, M. Mukherjee, J. Luo, R. Khan, J. C. Ferreon, V. J. Hilser, R. E. Shope, and R. O. Fox. 2004. Directed discovery of bivalent peptide ligands to an SH3 domain. *Protein Sci.* 13: 626–632.
65. Su, Z., M. J. Osborne, P. Xu, X. Xu, Y. Li, and F. Ni. 2005. A bivalent dissectional analysis of the high-affinity interactions between Cdc42 and the Cdc42/Rac interactive binding domains of signaling kinases in *Candida albicans*. *Biochemistry*. 44:16461–16474.
66. Oesterhelt, F., M. Rief, and H. E. Gaub. 1999. Single molecule force spectroscopy by AFM indicates helical structure of poly(ethylene-glycol) in water. *New J. Phys.* 1:6.1–6.11.
67. Kienberger, F., V. P. Pastushenko, G. Kada, H. J. Gruber, C. Riener, H. Schindler, and P. Hinterdorf. 2000. Static and dynamical properties of single poly(ethylene glycol) molecules investigated by force spectroscopy. *Single Mol.* 1:123–128.
68. Movileanu, L., S. Cheley, and H. Bayley. 2003. Partitioning of individual flexible polymers into a nanoscopic protein pore. *Biophys. J.* 85: 897–910.