Comparison of binding energies of SrcSH2-phosphotyrosyl peptides with structure-based prediction using surface area based empirical parameterization

DENISE A. HENRIQUES, JOHN E. LADBURY, AND RICHARD M. JACKSON

Department of Biochemistry and Molecular Biology, University College, Gower Street, London WC1E 6BT, United Kingdom

(RECEIVED March 20, 2000; FINAL REVISION July 17, 2000; ACCEPTED July 24, 2000)

Abstract

The prediction of binding energies from the three-dimensional (3D) structure of a protein–ligand complex is an important goal of biophysics and structural biology. Here, we critically assess the use of empirical, solvent-accessible surface area-based calculations for the prediction of the binding of Src-SH2 domain with a series of tyrosyl phosphopeptides based on the high-affinity ligand from the hamster middle T antigen (hmT), where the residue in the pY+3 position has been changed. Two other peptides based on the C-terminal regulatory site of the Src protein and the platelet-derived growth factor receptor (PDGFR) are also investigated. Here, we take into account the effects of proton linkage on binding, and test five different surface area-based models that include different treatments for the contributions to conformational change and protein solvation. These differences relate to the treatment of conformational flexibility in the peptide ligand and the inclusion of proximal ordered solvent molecules in the surface area calculations. This allowed the calculation of a range of thermodynamic state functions (ΔC_p , ΔS , ΔH , and ΔG) directly from structure. Comparison with the experimentally derived data shows little agreement for the interaction of SrcSH2 domain and the range of tyrosyl phosphopeptides. Furthermore, the adoption of the different models to treat conformational change and solvation has a dramatic effect on the calculated thermodynamic functions, making the predicted binding energies highly model dependent. While empirical, solvent-accessible surface area based calculations are becoming widely adopted to interpret thermodynamic data, this study highlights potential problems with application and interpretation of this type of approach. There is undoubtedly some agreement between predicted and experimentally determined thermodynamic parameters; however, the tolerance of this approach is not sufficient to make it ubiquitously applicable.

Keywords: calorimetry; free energy relationship; heat capacity; protein–protein interactions; structure-based drug design

With an ever-expanding body of high-resolution structural data on biomolecular complexes, we have increasing amounts of information on the complementary nature of interacting surfaces and the orientation of interacting groups. To rationalize structural detail in terms of energetics, it is necessary to correlate structure with experimentally determined thermodynamic data. Only a complete understanding of this relationship will allow full characterization of biomolecular interactions. Thus, prediction of binding energies directly from three-dimensional (3D) structure is currently a major goal in biophysics and the driving force for structure-based drug design. The immediate challenge remains to develop methods that

are both transferable from complex to complex and that are computationally affordable. Currently, most methods developed for this purpose employ empirical schemes of some kind to calculate binding energies (Vajda et al., 1997). There are a number of atomic level structure-based scoring methodologies suitable for calculating protein–protein and protein–ligand interaction energies from structure (see for example, Horton & Lewis, 1992; Jackson & Sternberg, 1995; Wallqvist et al., 1995; Weng et al., 1996). In addition, there are those methods that relate change in heat capacity (ΔC_p) to the change in solvent exposed surface area on binding (Spolar et al., 1992; Murphy & Friere, 1992).

The recent increase in the extent and reliability of experimentally determined binding energies has further stimulated advances in this field. In particular, binding studies carried out using high sensitivity isothermal titration calorimetry (ITC) provide accurate measurement of association constants ($K_{b\text{obs}}^\circ$) and enthalpies of

Reprint requests to: Richard M. Jackson, Department of Biochemistry and Molecular Biology, University College, Gower Street, London WC1E 6BT, United Kingdom; e-mail: jackson@biochem.ucl.ac.uk.

binding (ΔH_{obs}°) (where $^\circ$ refers the standard state at 1 atm and 298 K). These measured parameters can be used to obtain not only information on the free energy change for the association (ΔG_{obs}°), but also the fundamental components ΔS_{obs}° and $\Delta C_{p,obs}^\circ$, all of which act as further constraints in binding energy calculations. These components also provide detail on the molecular basis of binding affinity based on the enthalpy and entropy contributions. While determining the binding affinity of an interaction is important—alone it will not yield molecular insights into the interaction. For instance, two separate ligands can have similar binding affinities, but differ greatly in terms of their enthalpic and entropic contributions to the change in free energy depending on factors such as hydration, protonation, ion-binding, and conformational changes. Dissecting the free energy of binding into its component parts and describing an interaction in these terms would enable the identification of crucial factors, such as important residues and/or experimental conditions that contribute to the interaction. This would pin-point candidates for mutation/alteration. Therefore, the development of structure-based methods for predicting binding energies that also allow calculation of these additional thermodynamic components are potentially of great use in the design of novel compounds.

The correlation between the change in heat capacity (ΔC_p) and the change in solvent exposed surface area on going from one equilibrium state to another arose from data on the heats of transfer of liquid hydrocarbons and amides and protein folding/unfolding studies (Spolar et al., 1989; Livingstone et al., 1991). Dissecting out contributions from the removal of apolar and polar surface area to the energetics of protein folding led to the formation of predictive relationships between ΔC_p and surface area burial. Indeed, several variations have been proposed (Murphy et al., 1990; Murphy & Friere, 1992; Spolar & Record, 1994) that allow calculation of the individual thermodynamic parameters. In particular, it has been shown that empirical solvent-accessible surface area-based calculations can be applied to the prediction of binding energetics (Baker & Murphy, 1998), protein folding (Murphy & Gill, 1991; Murphy & Friere, 1992), and structure based drug design (Luque et al., 1998). The method has been applied successfully in a number of studies involving the prediction of binding energies including protein-peptide (Murphy et al., 1993; Gómez & Friere, 1995) and protein-protein (Baker & Murphy, 1997) interactions and a comprehensive overview of the practical implementation of the method has been presented by Baker and Murphy (1998). Their method is similar to that of Spolar and Record (1994) in that a complete thermodynamic description of the binding event can be calculated directly from solvent-accessible surface area. Equations are derived that allow calculation of the changes in thermodynamic state functions (ΔC_p , ΔS , ΔH , and ΔG) from changes in accessible surface area derived directly from the structure. These values can be compared with experimentally determined parameters.

The prediction of binding energetics from structure using this empirical model is still an area of some disagreement with conflicting interpretations of experimental data (Karplus & Janin, 1999; Privalov & Tamura, 1999). Indeed, several different empirical models have been proposed to explain the same binding phenomena. To address this, here we assess the influence on the energetics of the assumptions made in the empirical model by looking at the application of several different binding models that might realistically describe protein-peptide energetics. For this, we have adopted a model system based on the interaction of the SH2 domain from the protein Src and tyrosyl phosphopeptides. This model system

satisfies two important criteria required for investigating the tolerance of this type of correlation analysis: (1) the availability of accurate thermodynamic data for the binding energetics (including proton uptake/release during binding), and (2) the existence of high-resolution structural data for the free and complexed forms of the system of interest. A large body of data has been reported on the thermodynamics (Ladbury et al., 1995; Bradshaw et al., 1998; Chung et al., 1998; Bradshaw & Waksman, 1999), and structural detail (Waksman et al., 1992; 1993; Gilmer et al., 1994; Xu et al., 1995) of binding of this SH2 domain to a range of peptides and, indeed, peptidomimetics (Charifson et al., 1997; Plummer et al., 1997a; Xu et al., 1997, 1999). Structural detail of the interaction shows that tyrosyl phosphopeptides interact in a manner likened to a two-pronged plug (Waksman et al., 1992) with the pY residue and the residue three C-terminal of the pY (pY + 3) being buried in deep, well-defined pockets on the surface of the SH2 domain. The crystal structure shows that a hydrogen-bonded network of water molecules mediates the SH2/peptide interaction. This allows us to investigate how these water molecules might effect the surface area free energy relationship in the binding interaction.

The Src protein has been implicated in development of breast and colon carcinomas (Luttrell et al., 1994) and in osteoporosis (Soriano et al., 1991); hence, it is a promising target for drug design (Charifson et al., 1997; Plummer et al., 1997b) making dissemination of the binding energetics all the more important. To test the tolerance of the correlation between thermodynamic data and burial of surface area, we have determined binding data on a range of peptides. We have used a series of peptides based on the proposed high-affinity ligand from the hamster middle T antigen (hmT) where we have changed the residue in the pY + 3 position. This subtle change in the ligand enables the effect of small changes to the functional group of the side chain on the thermodynamic parameters without disrupting the water network (Chung et al., 1998). We have also used two other previously studied peptides that are the C-terminal regulatory site of the Src protein (C-tail) and the platelet-derived growth factor receptor (PDGFR) (Bradshaw et al., 1998). Both of these additional peptides are known ligands of the SH2 domain and have been structurally characterized (Waksman et al., 1992; Xu et al., 1997).

Results

Temperature dependence of thermodynamic parameters for the binding of the specific hmT and variant hmT phosphopeptides to the SrcSH2 protein domain

ITC experiments were used to examine the temperature dependence of the binding enthalpy of the pYEE peptide interactions with the SrcSH2 domain. At all temperatures studied, in all complexes, the binding enthalpies were observed to be negative (exothermic), with their magnitudes increasing with increasing temperature. Figure 1 shows the relationship between the change in the intrinsic enthalpy of binding (ΔH_{int}°) and increasing temperature comparing all the measured peptide-SH2 interactions. Using linear least-squares analysis (correlation coefficient 0.98) the temperature dependence of ΔH_{int}° yields a change in heat capacity $\Delta C_{p,int}^\circ$ for the binding of pYEEI to SrcSH2 of $-713 \pm 67 \text{ J K}^{-1} \text{ mol}^{-1}$. This is in reasonable agreement with the previously reported value of $-820 \pm 105 \text{ J K}^{-1} \text{ mol}^{-1}$ (Bradshaw et al., 1998).

Fig. 1. Comparison of the temperature dependence of ΔH_{int}° for all the pYEEEX-SrcSH2 binding interactions (pYEEI ∇ , pYEED $\blacktriangle$, pYEEE $\bullet$, pYEEV $\diamond$). The solid, dashed, and dotted lines represent linear least-squares fitting of the enthalpy data giving values for $\Delta C_{p,int}^{\circ}$ as -713 ± 67 , -704 ± 230 , -826 ± 66 , $-905 \pm 56 \text{ J mol}^{-1} \text{ K}^{-1}$ for the peptides pYEEI, pYEED, pYEEE, and pYEEV, respectively (see Table 1).

Structure-based thermodynamic calculations

Several authors have used surface area based empirical models to calculate thermodynamic parameters of binding (e.g., Murphy et al., 1990; Murphy & Friere, 1992; Spolar & Record, 1994). In this work, we have used the method of Baker and Murphy (1997). The differences between the variously reported methods are mainly due to variation in the constants applied with respect to polar and apolar surface.

The experimental, intrinsic binding energies ($\Delta C_{p,int}^{\circ}$, ΔH_{int}° , $T\Delta S_{int}^{\circ}$, and ΔG_{int}°) of the tyrosyl phosphopeptides given in Table 1 can be compared directly with the calculated energies for the mutant hmT peptides pYEEEX series given in Table 2. Similarly, those of the

hamster middle T antigen (hmT; sequence EPQpYEEIPIYL), human platelet-derived growth factor receptor (PDGFR; TQpYVP-MLE), and the regulatory tail of human Src (C-tail; PQpYQPGEN) can be compared with Table 3. The five different models make different assumptions regarding calculation of the binding energies and are discussed in detail in Materials and methods. These differences relate to the treatment of conformational flexibility in the peptide ligand and the importance of including proximal ordered solvent molecules in the calculated surface area. The tables clearly illustrate that the adoption of the different models can have a dramatic effect on the thermodynamic parameters calculated.

The best quantitative agreement between the predicted change in free energy of binding (ΔG_{calc}°) and experiment comes from the rigid-backbone + H₂O model for both the mutant hmT peptides pYEEEX series (Table 2) and the different phosphotyrosyl peptides series (Table 3). However, quantitative agreement of the calculated values of the change in heat capacity ($\Delta C_{p,calc}^{\circ}$), enthalpy (ΔH_{calc}°), and entropy of binding (ΔS_{calc}°) with experiment are poor. None of the calculated values predicts both ΔH_{int}° and $T\Delta S_{int}^{\circ}$, regardless of the model. Furthermore, while the approximate magnitude of ΔG_{int}° is maintained the calculations fail to reproduce the observed ranking within either of the two peptide series. Qualitative agreement with the observed ranking is not achieved with any of the models for either of the two series. Structural, thermodynamic, and mass spectral data (Chung et al., 1998) show that the pYEEEX mutant series of peptides interact with the SH2 domain in a similar way, i.e., the peptide adopts a similar conformation and water content in the complexes. Thus, it is somewhat surprising that calculated and experimental data do not show the same trends from one peptide to another.

Experimentally, the hmT pYEEI peptide binds to SrcSH2 with a large negative ΔH_{int}° and large positive $T\Delta S_{int}^{\circ}$ and a large negative $\Delta C_{p,int}^{\circ}$ at 25 °C. These characteristics are not well produced in any of the models studied. The large range in values of ΔG_{calc}° predicted for the peptides can be further broken down into different contributions according to the assumptions made by the model. Focusing on the behavior of the pYEEI peptide (Table 3) show the effects of differences in the models as they relate to the calculated binding

Table 1. Intrinsic thermodynamic parameters for SrcSH2-phosphopeptide binding^a

Peptide	$\Delta C_{p,int}^{\circ}$ ^b (J mol ⁻¹ K ⁻¹)	ΔH_{int}° ^c (kJ mol ⁻¹)	ΔS_{int}° ^d (J mol ⁻¹ K ⁻¹)	$T\Delta S_{int}^{\circ}$ ^d (kJ mol ⁻¹)	ΔG_{int}° ^e (kJ mol ⁻¹)
pYEEI	-713 ± 67	-32.99 ± 4.3	31.56	9.41 ± 4.5	-42.4 ± 0.20
pYEED	-704 ± 230	-21.79 ± 3.1	57.39	17.11 ± 3.5	-38.9 ± 0.40
pYEEE	-826 ± 66	-26.99 ± 2.9	45.31	13.51 ± 3.2	-40.5 ± 0.27
pYEEV	-905 ± 56	-22.89 ± 3.2	64.43	19.27 ± 3.3	-41.1 ± 0.12
C-tail ^f	-665.3 ± 117	-16.35 ± 3.7	40.08	11.95 ± 5.3	-28.3 ± 1.67
PDGFR ^f	-887.0 ± 71	-11.33 ± 3.7	69.26	20.67 ± 4.5	-32.0 ± 0.80

^aCalculated from the experimentally observed values and corrected for proton-linked contributions (see text). All pYEEEX-SrcSH2 experiments were conducted in 20 mM MES, 1 mM DTT, 1 mM EDTA, and 50 mM NaCl, pH 6.0. Experimental values have been corrected for proton release from the PO₃H⁻ group of the pY peptide residue upon binding, under these experimental conditions.

^bDetermined from Figure 1.

^cDetermined from observed data and corrected for proton linkage effects (using Equation 3).

^dCalculated from $\Delta S_{int}^{\circ} = (\Delta H_{int}^{\circ} - \Delta G_{int}^{\circ})/T$.

^eDetermined from $\Delta G_{int}^{\circ} = -RT \ln(K_{int})$, where K_{int} is the intrinsic equilibrium constant of binding corrected for proton linkage effects (using Equation 2).

^fResults taken from Bradshaw et al. (1998).

Table 2. Calculated binding energetics of the SrcSH2 domain interaction with the variant pYEEEX peptides, using the different models^a

Model	X ^b	$\Delta C_{p\text{calc}}^c$			ΔH_{calc}^d			ΔS_{calc}^e			$T\Delta S_{\text{calc}}^e$	ΔG_{calc}^f
		Pol	Apol	Tot	Pol	Apol	Tot	Solv	Conf	Tot ^g	Tot	Tot
Rigid-body	I	376	-1,079	-703	-55.3	55.8	0.5	180.1	-74.0	72.7	21.7	-21.1
	D	465	-892	-426	-68.4	46.1	-22.3	108.6	-75.5	0.3	0.1	-22.4
	E	465	-946	-481	-68.4	48.9	-19.5	123.2	-71.4	18.4	5.5	-24.9
	V	375	-1,033	-659	-55.1	53.5	-1.6	168.7	-66.9	68.4	20.4	-22.0
Rigid-backbone	I	465	-1,054	-588	-68.4	54.5	-13.9	150.7	-76.7	40.6	12.1	-26.0
	D	530	-883	-353	-77.8	45.6	-32.2	90.4	-77.2	-20.2	-6.0	-26.2
	E	551	-905	-354	-80.9	46.8	-34.1	90.7	-73.0	-15.7	-4.7	-29.4
	V	465	-1,008	-542	-68.4	52.1	-16.3	139.0	-69.7	35.9	10.7	-27.0
Flexible-backbone	I	463	-1,119	-656	-68.1	57.9	-10.2	167.6	-170.3	-36.1	-11.8	0.6
	D	520	-950	-430	-76.4	49.1	-27.3	110.0	-178.2	-101.5	-30.3	3.0
	E	543	-970	-427	-79.8	50.2	-29.6	109.2	-179.1	-103.5	-30.7	1.1
	V	461	-1,073	-613	-67.7	55.5	-12.2	157.0	-155.4	-31.8	-9.5	-2.7
Rigid-backbone + H ₂ O	I	636	-1,103	-468	-93.4	57.0	-36.3	119.5	-75.3	11.0	3.3	-39.6
	D	700	-932	-232	-102.8	48.2	-54.6	59.4	-75.9	-49.9	-14.9	-39.8
	E	722	-954	-232	-106.0	49.3	-56.7	59.4	-71.7	-45.7	-13.6	-43.1
	V	636	-1,057	-422	-93.4	54.7	-38.7	108.0	-68.3	6.3	1.9	-40.6
Flexible-backbone + H ₂ O	I	633	-1,168	-535	-93.1	60.4	-32.6	136.9	-174.2	-70.6	-21.0	-11.6
	D	690	-999	-309	-101.3	51.7	-49.7	79.2	-183.3	-137.5	-41.0	-8.7
	E	714	-1,019	-305	-104.8	52.7	-52.2	78.2	-181.4	-136.5	-40.8	-11.3
	V	631	-1,123	-492	-92.7	58.1	-34.7	125.8	-169.0	-76.5	-22.8	-11.9

^aAll calculated parameters are expressed in units of kJ mol⁻¹ except for $\Delta C_{p\text{calc}}$ and ΔS_{calc} , which are listed in units of J K⁻¹ mol⁻¹.

^bBased on Ac-EPQ(pY)EEEXPIYL from the hmT derived peptide.

^cDetermined from the change in apolar and polar surface upon binding using Equation 4.

^dCalculated using $\Delta C_{p\text{calc}}$ and Equations 1 and 5.

^eCalculated using $\Delta S_{\text{assoc}}^{\circ}$ and Equations 6 and 7.

^fCalculated using $\Delta H_{\text{calc}}^{\circ}$ and $T\Delta S_{\text{calc}}^{\circ}$ in Equation 8.

^gAll ΔS_{tot} include the cratic or mixing term $\Delta S_{\text{assoc}}^{\circ}$ (see text).

energies. The rigid-body model commonly used for modeling the energetics of protein–protein interactions (Baker & Murphy, 1997) gives similar results to the rigid-backbone model. Since only the reference states for the free peptides differ, this implies that the side-chain accessibility of the X-ray equilibrium conformation is close to that of the theoretical unfolded state, i.e., the X-ray equilibrium conformation bound to SH2 has similar side-chain accessibility to an unfolded peptide. The difference in binding energy between the rigid-backbone model and the flexible-backbone model is primarily a result of the large change in conformational entropy $\Delta S_{\text{conf}}^{\circ}$ resulting from change in the backbone degrees of freedom. In spite of having different reference states for the unbound peptide conformation, other contributions to $\Delta C_{p\text{calc}}$, $\Delta H_{\text{calc}}^{\circ}$, and $\Delta S_{\text{calc}}^{\circ}$ remain largely unchanged in magnitude. Thus, the fully flexible model for peptide binding underestimates the binding energy. The predicted value of $\Delta G_{\text{calc}}^{\circ}$ is too low, independently of whether binding site waters are included. The inclusion of observed binding site water molecules in the rigid-backbone + H₂O model gives the best agreement between theory and experiment. Relative to the rigid-backbone model, there is an increase in burial of accessible surface area, in particular, the polar accessible surface area that affects $\Delta H_{\text{calc}}^{\circ}$ and $\Delta S_{\text{solv}}^{\circ}$ increasing the agreement with experiment for $\Delta H_{\text{int}}^{\circ}$ but reducing it for $\Delta S_{\text{int}}^{\circ}$.

Comparison of the calculated values for the different peptides pYEEI, C-tail and PDGFR (Table 3) shows that the large differ-

ences in the burial of polar and nonpolar surface can have a dramatic effect on the calculated values of the $\Delta C_{p\text{calc}}$ and $\Delta H_{\text{calc}}^{\circ}$, and $\Delta S_{\text{solv}}^{\circ}$. For example, $\Delta C_{p\text{calc}}$ for the C-tail is predicted to have a much lower value than observed experimentally for any of the models tested. This is primarily because of the large burial of polar surface area. This is reflected in a very large negative $\Delta H_{\text{calc}}^{\circ}$, which is much greater (up to 400% larger) than predicted by experiment, and a large negative $T\Delta S_{\text{calc}}^{\circ}$ at 298 K, which is of the wrong sign. This disagreement with experiment is not strongly influenced by the minor differences in any of the models used.

Contribution of individual residues to binding

The predicted contribution of each residue in the protein–peptide interface to the calculated $\Delta G_{\text{calc}}^{\circ}$ is shown in Figure 2 for the SrcSH2 + pYEEI peptide for both rigid-backbone and flexible-backbone models both with and without water molecules present in the interface. The plots do not include the term $\Delta S_{\text{assoc}}^{\circ}$, which is a global term, applied to the system as a whole. The different models of backbone flexibility for the peptide have a large effect on the contribution of these residues to the free energy of binding. However, contributions from the protein residues remain largely invariant. The pY residue of the peptide is most affected on complex formation, when comparing the rigid-backbone to flexible-backbone model. The pY +3, Ile residue remains largely unaffected

Table 3. Comparison of calculated binding energetics for the SrcSH2 domain and the pYEEI-hmT, C-tail, PDGFR, and EGFR tyrosyl peptides, at 25 °C^a

Model	Peptide ^b	$\Delta C_{p,calc}^c$			ΔH_{calc}^d			ΔS_{calc}^e			$T\Delta S_{calc}^e$	ΔG_{calc}^f
		Pol	Apol	Tot	Pol	Apol	Tot	Solv	Conf	Tot ^g	Tot	Tot
Rigid-body	pYEEI	376	-1,079	-703	-55.3	55.8	0.5	180.1	-74.0	72.7	21.7	-21.1
	C-tail	503	-703	-200	-73.8	36.3	-37.5	51.2	-67.6	-49.8	-14.8	-22.7
	PDGFR	388	-883	-495	-57.0	45.6	-11.4	126.7	-58.7	34.6	10.3	-21.7
Rigid-backbone	pYEEI	465	-1,054	-588	-68.4	54.5	-13.9	150.7	-76.7	40.6	12.1	-26.0
	C-tail	531	-690	-159	-78.0	35.7	-42.3	40.7	-69.1	-61.8	-18.4	-23.9
	PDGFR	428	-850	-422	-62.9	43.9	-19.0	108.0	-60.9	13.7	-4.1	-23.1
Flexible-backbone	pYEEI	463	-1,119	-656	-68.1	57.9	-10.2	167.6	-170.7	-36.1	-11.8	0.6
	C-tail	520	-693	-175	-76.4	35.9	-40.5	44.4	-132.7	-121.8	-36.3	-4.3
	PDGFR	351	-905	-554	-51.6	46.8	-4.8	141.7	-83.9	24.4	7.3	-12.1
Rigid-backbone + H ₂ O	pYEEI	636	-1,103	-468	-93.4	57.0	-36.3	119.5	-75.3	11.0	3.3	-39.6
	C-tail	757	-723	34	-111.1	37.4	-73.8	-8.7	-67.5	-119.7	-32.7	-41.1
	PDGFR	567	-870	-303	-83.3	45.0	-38.3	77.6	-57.6	-13.3	-4.0	-34.3
Flexible-backbone + H ₂ O	pYEEI	633	-1,168	-535	-93.1	60.4	-32.6	136.9	-174.2	-70.6	-21.0	-11.6
	C-tail	746	-726	19	-109.4	37.5	-72.0	-5.0	-153.1	-191.5	-57.1	-15.0
	PDGFR	489	-925	-435	-72.9	47.8	-24.1	111.2	-84.8	-6.8	-2.0	-22.1

^aAll calculated parameters are expressed in units of kJ mol⁻¹ except for $\Delta C_{p,calc}$ and ΔS_{calc} , which are expressed in units of J K⁻¹ mol⁻¹.

^bpYEEI denotes the hmT based peptide Ac-EPQ(pY)EEIPIYL; C-tail denotes a peptide derived from the C-terminal tail of the full human tyrosine kinase Ac-PQ(pY)QPGEN; PDGFR denotes peptide Ac-TQ(pY)VPMLL derived from a sequence found in the beta-subunit of the human platelet derived growth factor.

^cDetermined from the change in apolar and polar surface upon binding using Equation 4.

^dCalculated using $\Delta C_{p,calc}$ and Equations 1 and 5.

^eCalculated using ΔS_{assoc} and Equations 6 and 7.

^fCalculated using ΔH_{calc} and $T\Delta S_{calc}$ in Equation 8.

^gAll ΔS_{tot} include the cratic or mixing term ΔS_{assoc}^o .

going from one model to the other but makes a significant contribution to the calculated ΔG_{calc}^o in both. The pY +1 and +2 Glu residues go from predicted favorable contributions (rigid-backbone model) to unfavorable contributions (flexible-backbone model). This is also observed for the pY -1 Gln residue. However, residues pY +5 to +7 (Ile, Tyr, Ile), range from little or no contribution to significant unfavorable contributions to the ΔG_{calc}^o , relative to the other contributing residues for this model.

Dependence of heat capacity upon the loss of solvent exposed surface area going from the free to the bound state

Figure 3 shows experimentally determined values for $\Delta C_{p,int}^o$ against the loss in water-accessible apolar surface area ΔASA_{apol} . The data set shown as solid black triangles is $\Delta C_{p,int}^o$ taken from the SrcSH2-peptide binding interactions (Table 1) vs. the ΔASA_{apol} obtained using the rigid-body model. For comparison, the data sets shown, as hollow circles, are taken from thermodynamic protein folding experiments (Livingstone et al., 1991, and references therein) and from the heats of solvent transfer of small hydrocarbons and amides (Spolar et al., 1992). The dotted line represents a linear least-squares fit of these data (hollow circles). The SrcSH2-peptide data set all seem to be offset from the linear least-squares fit of the protein folding and heats of solvent transfer data, clustering above the line.

This observation suggests that there is some deficit in $\Delta C_{p,calc}$, which is not accounted for. Previously published work, in which

surface area calculations have been compared with experiment report, an underestimation of experimentally determined heat capacity. It has been suggested that this may reflect the involvement of water molecules in the binding interaction (Guinto & Di Cera, 1996; Morton & Ladbury, 1996).

Discussion

Comparison of the experimental results with those from the surface area based empirical relationship shows little agreement for the interaction of the SrcSH2 domain with a range of tyrosyl phosphopeptides. The lack of correlation between the experimental and calculated data has previously been attributed to the failure to account for structural alterations of the interacting components (Bradshaw et al., 1998). In this case, the protein interacts in a way approximating a rigid body (i.e., structures of the apo- and complexed SH2 domain are essentially identical). Adoption of more sophisticated models that incorporate contributions for conformational change of the peptides still fail to produce any real quantitative agreement with the experimental data.

Structural detail on the SH2 domain with the tyrosyl phosphopeptides shows that water molecules form an intimate part of the bimolecular interface. Surface area calculations are usually done in the absence of interfacial water molecules. Here, we attempt to account for their potential effects on the surface area, however, even including this usually ignored contribution to buried surface area we still see no overall improvement in the agreement between experimental and calculated thermodynamic parameters.

Fig. 2. Contributions from each interface residue in the pYEEI-SrcSH2 binding interaction to the change in free energy ΔG_{calc}° of binding at 25 °C, in kJ mol⁻¹. Plots are shown for the rigid-backbone and flexible-backbone models, with and without the inclusion of interface water molecules. These values do not include the “cratic” or mixing entropy term ΔS_{assoc}° . Protein residues are numbered as in the structure coordinate file and peptide residues are numbered with reference to the pY as 0.

Prior to this study, there had been some reported disagreement between the accessible surface area based empirical model and experimental binding energies (Varadarajan et al., 1992; Kelley & O’Connell, 1993; Kresheck et al., 1995; Pearce et al., 1996). However, in most cases these disagreements could be due to incomplete treatment of proton linkage effects or conformational change involving order/disorder transitions on binding (see Baker & Murphy, 1998). Here, the known proton linkage effects have been treated and comparison of the peptide bound and unbound states of SrcSH2 show conformational change to be minimal on binding (Waksman et al., 1993). While the unbound conformations of the free peptides is unknown, three possible models detailing different levels of conformational change have been studied here. Furthermore, the possible influence of resolved water molecules at the

protein–peptide interface have been taken into account. The level of agreement between calculated and experimental values of the state functions is poor in both of the two sets of peptide series studied here. It can be argued that the model-built structures used for the series of pY + 3 mutants may not accurately model experiment, this is not the case for the experimentally determined structures of the pYEEI peptide, C-tail, and PDGFR. However, it is not the inability of the method to reproduce experiment that gives greatest cause for concern it is the degree to which assumptions made in the modeling have a pronounced effect on calculated values. Differences in the models translate into large differences in the calculated $\Delta C_{p,calc}$, ΔH_{calc}° , and ΔS_{calc}° .

The contribution of individual residues to the free energy of binding in the pYEEI-SH2 domain (Fig. 2) shows that the model

Fig. 3. Plot showing the dependence of the change in heat capacity on the loss in water-accessible apolar surface area (ΔASA_{apol}). The data set shown as black triangles is taken from all the SrcSH2-peptide binding interactions ($\Delta C_{p, int}$, Table 1) and the ΔASA_{apol} obtained using the rigid-body model. For comparison the data sets shown, as circles, are taken from thermodynamic protein folding experiments (Livingstone et al., 1991, and references therein) and from the heats of solvent transfer of small hydrocarbons and amides (Spolar et al., 1992). The dotted line represents a linear least-squares fit of these data (open circles).

for the main-chain entropy has a dramatic effect on the importance of individual residue contributions. In all models, the pY and Ile at pY +3 have the greatest favorable residue contributions to binding, reaffirming the “two-pronged plug two-holed socket” model for the mechanism of binding of the SrcSH2 domain to phosphotyrosyl peptides (Waksman et al., 1993). However, the relative importance of pY changes greatly in going from the rigid-backbone + H₂O to the flexible-backbone + H₂O models. However, the two seemingly important Glu residues at pY +1 and pY +2 have binding contributions of opposite sign depending on the model, i.e., in the flexible-backbone model any favorable contribution from surface area burial is more than offset by the unfavorable contribution from loss in backbone entropy. In the flexible model, only the pY and the Ile and Pro residues at pY +3 and +4, respectively, make sizable favorable contributions to binding, whereas in the rigid model all eight peptide residues make a favorable contribution to binding. These two models give fundamentally different interpretations for individual residue contributions to binding.

The current studies are limited by the experimental error in the binding energies (~10%), and the calculations are limited by the quality and resolution of the crystallographic structural information. In particular, the presence/absence of ordered solvent molecules has a substantial influence on the predicted binding energetics of the peptides studied here (Tables 1, 2). This implies that the quality of the experimental structure will influence the calculated binding energy. Furthermore, it is unclear at what level should observed waters be treated as ordered water molecules or as a solvent continuum for the purpose of the accessibility calculations. Lastly, the level of conformational flexibility in the peptide/protein main chain, in particular, can have a considerable influence on predicted binding energies. As shown in the results, all these factors have a sizable effect on both the sign and the magnitude of the thermodynamic state functions. It is the large magnitude of these effects that make the binding energy highly model dependent.

The predictive power of surface area based empirical methods may lie in the use of experimental structures, which are at equilibrium. However, problems arise when using modeled structures as the method assumes that the model built structure is at equilibrium, which may not be the case. This method is therefore of limited predictive value if it cannot predict structure. To do this, the method must predict the lowest energy for the system as a whole (protein, ligand, and solvent). However, the present method cannot calculate this value, but only that of the lowest energy for the protein–ligand interaction that may not be the observed conformation. This is an inherent limitation making the method of limited utility in structure based drug design.

From the studies conducted here, it would appear that abstractions, such as the parameterized surface area based calculations, are not reliable enough as yet to be able to predict thermodynamic constants for systems such as this. Some qualitatively useful information may however be gained from the analysis of individual residue contributions to binding. The predicted importance of the binding contributions of the pY and pY +3 residues is in agreement with experiment. The parameterized model is being widely adopted to predict thermodynamic data from structure and vice versa; however, clearly there are potential problems with application and interpretation of this type of approach. It should be emphasized that although there is undoubtedly a broad agreement between predicted and experimentally determined thermodynamic parameters, the tolerance of the currently available correlation is not sufficient to make this approach ubiquitously applicable.

Materials and methods

Protein expression and purification

The SH2 domain of the *v-src* gene was expressed in *Escherichia coli* and purified as described previously (Waksman et al., 1992). A final purification step for the protein involved overnight dialysis into 50 mM Tris-HCl, pH 7.5, 1 mM DTT, 1 mM EDTA, and 150 mM NaCl. The dialyzed material was passed through an affinity column of *o*-phosphotyrosine agarose beads (Sigma, St. Louis, Missouri) pre-equilibrated in the same buffer. The protein was eluted with 50 mM Tris-HCl, pH 7.5, 1 mM DTT, 1 mM EDTA, and 1 M NaCl, this step served both to further purify and test affinity. Purity of the SrcSH2 domain was deemed to be >98% by SDS-PAGE. Prior to titration, the protein was dialyzed against 4 L of the standard buffer for the calorimetry experiments; 20 mM MES, pH 6.0, 1 mM DTT, 1 mM EDTA, and 50 mM NaCl, for a minimum of 24 h at 4 °C. Protein concentrations were obtained using the extinction coefficient $\epsilon_{280} = 13,730 \text{ M}^{-1} \text{ cm}^{-1}$.

Tyrosine phosphopeptides

The phosphopeptides used in this study were based on the hamster middle-sized tumor (hmT) specific binding peptide Ac-EPQ (pY)EEIPIYL-NH₂ (subsequently referred to as pYEEI peptide) (Songyang et al., 1993). Three other substituted peptides were also synthesized, in which the pY +3 residue Ile was replaced with an Asp, a Glu, and a Val, and purified by reversed-phase high-performance liquid chromatography as described previously (Pintar et al., 1996). Lyophilized peptides were dissolved in the buffer from the SH2 domain dialysis and the pH checked prior to titration. Peptide concentrations were obtained both by mass and by the

extinction coefficient, $\epsilon_{270} = 1,990 \text{ M}^{-1} \text{ cm}^{-1}$ and $\epsilon_{280} = 1,680 \text{ M}^{-1} \text{ cm}^{-1}$.

Isothermal titration calorimetry

All calorimetric titrations were performed using a MCS ITC (MicroCal Inc., Northampton, Massachusetts). Solutions of dialyzed protein were degassed under vacuum prior to use in the ITC. In a typical titration, 25–30 μM of SrcSH2 domain was placed in the 1.35 mL reaction cell, while 250–400 μM phosphopeptide was loaded in to the 250 μL injection syringe. To obtain the value for the change in heat capacity ($\Delta C_{p \text{ int}}$), titration experiments were performed at a range of temperatures, and the results incorporated into the equation below:

$$\Delta C_{p \text{ int}} = (\Delta H_{T_1}^\circ - \Delta H_{T_2}^\circ) / (T_2 - T_1) \quad (1)$$

where T_1 and T_2 are two experimental temperatures. Data were fitted using the ORIGIN software as described previously (Ladbury & Chowdhry, 1996). Heats associated with the dilution of the peptide into buffer and buffer into protein, determined in separate experiments, were subtracted prior to data fitting.

Correction for the effects of protonation

The direct measurement of enthalpy for the interaction includes contributions from all the enthalpic events that occur on going from free to bound state. Thus, prior to comparing the experimentally determined values with the calculated ones, any enthalpic contributions, which are not accounted for in the calculated values, have to be subtracted from the experimental ones. In this case, studies have shown that proton release from the phosphate occurs in the Src SH2-hmT interaction, at pH 6.0 using MES buffer (Bradshaw & Waksman, 1998). Before the observed and calculated binding energetics can be compared, proton coupling effects must be calculated and the experimental data corrected to give the intrinsic binding energetics. At pH 6.0, a single-proton-linkage model can be applied to the system (Bradshaw & Waksman, 1998), K_{int}° (the intrinsic equilibrium binding constant for the unprotonated ligand) was found by treating K_{obs}° as a function of pH:

$$K_{\text{obs}}^\circ = K_{\text{int}}^\circ (1 + 10^{(\text{p}K_a)_b - \text{pH}} / 1 + 10^{(\text{p}K_a)_f - \text{pH}}) \quad (2)$$

where $10^{(\text{p}K_a)_b}$ and $10^{(\text{p}K_a)_f}$ are the protein binding constants bK_p and fK_p , respectively, for the bound and free forms of the peptide ligand with $(\text{p}K_a)_b$ and $(\text{p}K_a)_f$ reported to be 4.4 ± 0.3 and 6.1 ± 0.1 , respectively. For changes in enthalpy:

$$\Delta H_{\text{obs}}^\circ = \Delta H_{\text{int}}^\circ + ({}^b f) {}^b \Delta H_p^\circ - ({}^f f) {}^f \Delta H_p^\circ + n_H^+ \Delta H_{\text{ion}}^\circ \quad (3)$$

The $\Delta H_{\text{obs}}^\circ$ is the experimentally observed change in enthalpy, $\Delta H_{\text{int}}^\circ$ is the intrinsic heat of binding, $\Delta H_{\text{ion}}^\circ$ the enthalpy or heat of protonation/deprotonation of the buffer, and n_H^+ is the number of protons exchanged with the buffer on binding. The enthalpies of protonation of the bound (${}^b \Delta H_p^\circ$) and free (${}^f \Delta H_p^\circ$) ligand and the fractional saturation of protons at a given pH of the bound (${}^b f$) and free (${}^f f$) ligands are available for the pYEEI peptide (Bradshaw & Waksman, 1998). The free $\text{p}K_a$ of the ionizable group linked to binding ($(\text{p}K_a)_f$) is 6.1 in good agreement with the known $\text{p}K_a$ of pTyr. While the $\text{p}K_a$ of the bound form is 4.4. These values and the

enthalpies of protonation of the bound and free ligands are taken to be the same for all the ligands studied here. This approximation is based on the observation that the environment of the srcSH2 phosphate binding site is conserved in the PDGFR and C-tail peptide complexes, sequestering the phosphate group from solvent. Similarly, the phosphate group will be largely solvated in the free ligands, evidence that is supported by the free $\text{p}K_a$ of the ionizable group linked to binding being in good agreement with the $\text{p}K_a$ of pTyr in solution. $\Delta H_{\text{ion}}^\circ$, the heat of ionization of the buffer MES at 25 $^\circ\text{C}$, is $15.61 \text{ kJ mol}^{-1}$ (Christensen et al., 1976). The number of protons exchanged on binding at pH 6.0, $n(6.0)_H^+$ is -0.48 ± 0.2 for the pYEEI peptide (Bradshaw & Waksman, 1998). For the PDGFR and C-tail peptides, $n(6.0)_H^+$ is taken to be -0.3 ± 0.1 (Bradshaw et al., 1998). $\Delta C_{p \text{ int}}$ is calculated from the temperature dependence of $\Delta H_{\text{int}}^\circ$; we assume here that the proton linkage effects are temperature independent. The observed calorimetric data were corrected to yield the intrinsic binding thermodynamic parameters (shown in Table 1).

Structure-based thermodynamic calculations

The calculation of thermodynamic parameters from changes in accessible surface areas are based on the methods of Baker and Murphy (1998). Here, we give only a brief outline of the methods used. The change in heat capacity $\Delta C_{p \text{ calc}}$ is calculated using Equation 4 and the ΔASA (the change in the solvent accessible surface area) for polar and apolar surface area:

$$\Delta C_{p \text{ calc}} = a \Delta\text{ASA}_{\text{apol}} + b \Delta\text{ASA}_{\text{pol}} \quad (4)$$

where a is $1.82 \text{ J K}^{-1} \text{ mol}^{-1} \text{ \AA}^{-2}$ and b is $-1.09 \text{ J K}^{-1} \text{ mol}^{-1} \text{ \AA}^{-2}$. The contribution to the change in enthalpy ΔH at 60 $^\circ\text{C}$ was calculated using

$$\Delta H_{60^\circ\text{C}}^\circ = c \Delta\text{ASA}_{\text{pol}} + d \Delta\text{ASA}_{\text{apol}} \quad (5)$$

where c is $122 \text{ J mol}^{-1} \text{ \AA}^{-2}$ and d is $-30.4 \text{ J mol}^{-1} \text{ \AA}^{-2}$. Using this value and the previously determined value for $\Delta C_{p \text{ calc}}$, the $\Delta H_{\text{calc}}^\circ$ at 25 $^\circ\text{C}$ was derived from Equation 1. The total entropy $\Delta S_{\text{calc}}^\circ$ was calculated as the sum of three components: entropy of solvation ($\Delta S_{\text{solv}}^\circ$), conformational entropy ($\Delta S_{\text{conf}}^\circ$), and entropy of association ($\Delta S_{\text{assoc}}^\circ$). $\Delta S_{\text{assoc}}^\circ$ is also called the “cratic” or mixing term and is a fixed value reflecting the mixing of solute and solvent molecules. For a bimolecular reaction in a 1 M standard state, this is $-33 \text{ J K}^{-1} \text{ mol}^{-1}$ (Baker & Murphy, 1998). $\Delta S_{\text{solv}}^\circ$ is obtained using Equation 6 and the previously determined value of $\Delta C_{p \text{ calc}}$:

$$\Delta S_{\text{solv}}^\circ = \Delta C_{p \text{ calc}} \ln(298/385) \quad (6)$$

where $\Delta S_{\text{conf}}^\circ$ was calculated from changes in side-chain and backbone conformations upon binding, contributions were dependent on assumptions made regarding the state of the peptide (see structural models section below)

$$\begin{aligned} \Delta S_{\text{conf},i}^\circ = & \sum (\Delta\text{ASA}_{\text{SC},i} / \text{ASA}_{\text{AXA-SC},i}) S_{\text{Bu-Ex}}^\circ \\ & + \sum (\Delta\text{ASA}_{\text{BB},i} / \text{ASA}_{\text{AXA-BB},i}) S_{\text{BB}}^\circ \\ & + \sum (\Delta\text{ASA}_{\text{BB},i} / \text{ASA}_{\text{AXA-BB},i}) S_{\text{Ex-U}}^\circ \end{aligned} \quad (7)$$

where $\Delta ASA_{SC,i}/\Delta ASA_{BB,i}$ are the changes in ASA of side chain/backbone of residue i on binding. $ASA_{AXA-SC,i}/ASA_{AXA-BB,i}$ are the ASA of the side chain/backbone of residue i in a extended Ala-X-Ala tripeptide. Values for the main-chain/backbone entropy (S_{BB}° and S_{EX-U}°) were taken from D'Aquino et al. (1996) and side-chain entropy (S_{BU-EX}°) were from Lee et al. (1994). Finally, ΔG_{calc}° was determined using the calculated values for ΔH_{calc}° , ΔS_{calc}° at 25 °C, and the standard thermodynamic relationship:

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ. \quad (8)$$

Although other investigators have used different constants based on different refinements of the empirical correlation (e.g., Spolar & Record, 1994), we have adopted those based on Baker and Murphy (1998). Initial calibration was performed using the binding of turkey ovomucoid third domain (OMTKY3) to porcine pancreatic elastase (PPE) (coordinates kindly provided by K. Murphy). Results for $\Delta C_{p,cal}$, ΔH_{calc}° , ΔS_{calc}° , and ΔG_{calc}° were within $\pm 5\%$ of those reported by Baker and Murphy (1997) (results not shown).

Structural models

The atomic models were based on the X-ray crystallographic structure, chain C of SrcSH2-hmT (1sps) (Waksman et al., 1993). Models for the pY +3 variants Asp, Glu, and Val were generated by changing the side chain from Ile to the appropriate amino acid. The optimal side-chain conformation was then generated by a mean field multiple copy search of the side-chain conformations (represented in a rotamer library) using the program MULTIDOCK (Jackson & Sternberg, 1998). All atoms were kept fixed with the exception of the pY +3 side-chain conformation to generate an optimal steric placement of the new amino acid with minimum perturbation to the remaining structure. As in previous studies using the methodology (Luque et al., 1998), it is assumed that the structure of the mutated conformation remains essentially unchanged. High resolution X-ray crystallographic structures were also used for SrcSH2 associated with other phosphotyrosyl peptides, human platelet-derived growth factor receptor (PDGFR) (Protein Data Bank (PDB) code, 1sha), and the regulatory tail of human Src (C-tail) (PDB code, 1fmk).

The protein is considered to undergo "rigid-body" body binding (Baker & Murphy, 1997). This is a reasonable assumption based on the similarity of the structures of the free and bound SH2 domains. Three different models were used to represent the behavior of the peptide in the structure-based calculations:

- Rigid-body interaction model (Baker & Murphy, 1997), where the peptide (as well as the protein) is assumed to have a rigid backbone and the same equilibrium side-chain conformations in both the complexed and unbound states.
- Flexible-backbone model (Gómez & Friere, 1995), where peptide is assumed to go from a disordered or unfolded state to the X-ray equilibrium conformations observed in the complexed state.
- Rigid-backbone model; the peptide is assumed to have a rigid backbone; however, side chains are assumed to be disordered in the unbound state.

These models differ in (1) the reference states used for the unbound peptide conformation in the calculation of ΔASA as well as (2) the values of S_{BB}° , S_{EX-U}° used to calculate $\Delta S_{conf,i}^\circ$:

- The reference states used for the side chain and backbone of the unbound peptide conformation ($ASA_{SC,i}$ and $ASA_{BB,i}$) are the equilibrium X-ray structure for the rigid-body model (A) and the Ala-X-Ala peptide conformation for the flexible-backbone model (B). However, in the rigid-backbone model, they are the equilibrium X-ray structure for the backbone and the Ala-X-Ala peptide conformation for the side chains.
- Similarly, the values used to calculate the loss in backbone entropy (S_{BB}° , S_{EX-U}°) are zero in the case of the rigid-body (A) and rigid-backbone (C) models, as only the side-chain entropy contribution is considered.

To analyze the effects of interfacial water molecules on the binding energetics, water molecules were included in the protein surface area calculations (+H₂O). An interface water molecule was defined as being within X Å of both the protein and the peptide according to:

$$X < (3r_{wat} + vdwrA) \quad (9)$$

where r_{wat} is the radius of a water molecule and $vdwrA$ is the van der Waals radius of atom A (see Fig. 4). This includes waters that are simultaneously in the first or second solvation shells of both the peptide and the protein.

Solvent accessible surface areas (ASA) were calculated using the program NACCESS (Hubbard & Thornton, 1993); an implementation of the Lee and Richards algorithm (Lee & Richards, 1971). Default values were used for the water probe sphere (radius; 1.4 Å) and z-slice (width; 0.05 Å) and van der Waals radii (Chothia, 1976). Solvent accessibility was calculated for the complexed structure for the SrcSH2 and hmT peptide and pY +3 variants ($ASA_{SrcSH2+pep}$). The structure of the free protein was extracted from the coordinates of the complexed SrcSH2 domain structure (ASA_{SrcSH2}). The accessibility of the peptide ($ASA_{peptide}$) depends on the model used and is described in detail above. Defined at atomic level, the surfaces of carbons C and sulfur S were classified as apolar (apolar), whereas those of other atoms were classified as polar. The change in the total, polar, and apolar accessible surface area on binding (see Fig. 5A) were obtained using

$$\Delta ASA = ASA_{SrcSH2+pep} - (ASA_{SrcSH2} + ASA_{peptide}). \quad (10)$$

Fig. 4. Schematic representation of the definition of an interface water molecule. Each interface water molecule must be within X Å of both the protein and the peptide. The radius of a water molecule (r_{wat}) is taken as 1.4 Å. $vdwrA$ denotes the van der Waals radius of atom A in the protein or peptide. The van der Waals surface area of the peptide/protein is denoted by (----).

Fig. 5. Schematic showing accessible surface area calculations. **A:** The changes in solvent accessible area were calculated going from the unbound/free state to the bound, using Equation 3. **B:** For the inclusion of interfacial water molecules in the change of total, polar and apolar solvent accessible surface areas required calculations going from the unbound/free state to the bound, using $\Delta ASA = ASA_{SrcSH2+pep+H_2O} - (ASA_{SrcSH2+H_2O} + ASA_{peptide})$.

To observe the effect, if any, that water molecules in the protein-peptide interface have on the calculated thermodynamic parameters, the relative solvent's accessibilities were calculated with interface water molecules being present on the free and bound protein structure (as if an extension of the protein), see Figure 5B.

Acknowledgments

We are grateful to Dr. Gabriel Waksman for the kind gift of a plasmid containing the SrcSH2 domain and to Dr. Kip Murphr any for supplying the unpublished coordinates of the PPE-OMTKY3 structure. This work was supported by a BBSRC Special Studentship to D.A.H. and the Wellcome Trust in the form of a Senior Fellowship to J.E.L. and a Career Development Fellowship to R.M.J.

References

- Baker BM, Murphy KP. 1997. Dissecting the energetics of a protein-protein interaction: The binding of ovomucoid third domain to elastase. *J Mol Biol* 268:557-569.
- Baker BM, Murphy KP. 1998. Prediction of binding energetics from structure using empirical parameterisation. *Methods Enzymol* 295:294-315.
- Bradshaw JM, Gruzca RA, Ladbury JE, Waksman G. 1998. Probing the "two-pronged plug two-holed socket" model for the mechanism of binding of the SrcSH2 domain to phosphotyrosyl peptides: A thermodynamic study. *Biochemistry* 37:9083-9090.
- Bradshaw JM, Waksman G. 1998. Calorimetric investigation of proton linkage by monitoring both the enthalpy and association constant of binding: Application to the interaction of the SrcSH2 domain with a high-affinity tyrosyl phosphopeptide. *Biochemistry* 37:15400-15407.
- Bradshaw JM, Waksman G. 1999. Calorimetric examination of high-affinity SrcSH2 domain-tyrosyl phosphopeptide binding: Dissection of phosphopeptide sequence specificity and coupling energetics. *Biochemistry* 38:5147-5154.
- Charifson PS, Shewchuk LM, Rocque W, Hummel CW, Jordan SR, Mohr C, Pacofsky GJ, Peel, M R, Rodriguez M, Sternbach DD, Consler TG. 1997. Peptide ligands of pp60^{src} SH2 domains: A thermodynamic and structural study. *Biochemistry* 36:6283-6293.
- Chothia C. 1976. The nature of the accessible and buried surfaces in proteins. *J Mol Biol* 105:1-14.
- Christensen JJ, Hansen LD, Izatt RM. 1976. *Handbook of proton ionisation heats and related thermodynamic quantities*. New York: John Wiley and Sons.
- Chung E, Henriques DA, Renzoni D, Zvelebil M, Bradshaw M, Waksman G, Robinson CV, Ladbury JE. 1998. Mass spectrometric and thermodynamic studies reveal the role of water molecules in complexes formed between SH2 domains and tyrosyl phosphopeptides. *Structure* 6:1141-1151.
- D'Aquino JA, Gómez J, Hilser VJ, Lee KH, Amzel LM, Friere E. 1996. The magnitude of the backbone conformational entropy change in protein folding. *Proteins Struct Funct Genet* 25:143-156.
- Gilmer T, Rodriguez M, Jordan S, Crosby R, Alligood K, Gree M, Kimery M, Wagner C, Kinder D, Charifson P, et al. 1994. Peptide inhibitors of SrcSH3-SH2 phosphoprotein interactions. *J Biol Chem* 269:31711-31719.
- Gómez J, Friere E. 1995. Thermodynamic mapping of the inhibitor site of the aspartic protease endothiapepsin. *J Mol Biol* 252:337-350.
- Guinto ER, Di Cera E. 1996. Large heat capacity change in a protein-monovalent cation interaction. *Biochemistry* 35:8800-8804.
- Horton N, Lewis M. 1992. Calculation of the free energy of association for protein complexes. *Protein Sci* 1:169-181.
- Hubbard SJ, Thornton JM. 1993. "NACCESS" computer program. Department of Biochemistry and Molecular Biology, University College London.
- Jackson RM, Sternberg MJE. 1995. A continuum model for protein-protein interactions: Application to the docking problem. *J Mol Biol* 250:258-275.
- Jackson RM, Sternberg MJE. 1998. Rapid refinement of protein interfaces incorporating solvation: Application to the docking problem. *J Mol Biol* 276:265-285.
- Karplus M, Janin J. 1999. Comment on: "The entropy cost of protein association." *Protein Eng* 12:185-186.
- Kelley RF, O'Connell MP. 1993. Thermodynamic analysis of an antibody functional epitope. *Biochemistry* 32:6828-6835.

- Kresheck GC, Vitello LB, Erman JE. 1995. Calorimetric studies on the interaction of horse ferricytochrome *c* and yeast cytochrome *c* peroxidase. *Biochemistry* 34:8398–8405.
- Ladbury JE, Chowdhry BZ. 1996. Sensing the heat: The application of isothermal titration calorimetry to the thermodynamic study of biomolecular interactions. *Chem Biol* 3:791–801.
- Ladbury JE, Lemmon MA, Zhou M, Green J, Botfiel MC, Schlessinger J. 1995. Measurement of the binding of tyrosyl-phosphopeptides to SH2 domains: A reappraisal. *Proc Natl Acad Sci USA* 92:3199–3203.
- Lee B, Richards FM. 1971. The interpretation of protein structures: Estimation of static accessibility. *J Mol Biol* 55:379–400.
- Lee KH, Xie D, Freire E, Amzel LM. 1994. Estimation of changes in side chain configurational entropy in binding and folding; general methods and application to helix formation. *Proteins Struct Funct Genet* 20:68–84.
- Livingstone JR, Spolar RS, Record MT Jr. 1991. Contribution to the thermodynamics of protein folding from the reduction in water-accessible apolar surface area. *Biochemistry* 30:4237–4244.
- Luttrell DK, Lee A, Lansig TJ, Crosby RM, Jung KD, Willard D, Luther M, Rodriguez M, Berman J, Gilmer TM. 1994. Involvement of pp60^{src} with two major signalling pathways in human breast cancer. *Proc Natl Acad Sci USA* 91:83–87.
- Luque I, Gómez J, Semo N, Friere E. 1998. Structure-based thermodynamic design of peptide ligands: Application to peptide inhibitors of the aspartic protease endothiapepsin. *Proteins Struct Funct Genet* 30:74–85.
- Morton CJ, Ladbury JE. 1996. Water mediated protein-DNA interactions: The relationship of thermodynamics to structural detail. *Protein Sci* 5:2115–2118.
- Murphy KP, Friere E. 1992. Thermodynamics of structural stability and co-operative folding behaviour in proteins. *Adv Prot Chem* 43:313–326.
- Murphy KP, Gill SJ. 1991. Solid model compounds and the thermodynamics of protein unfolding. *J Mol Biol* 222:699–709.
- Murphy KP, Privalov PL, Gill SJ. 1990. Common features of protein unfolding and dissociation of hydrophobic compounds. *Science* 247:559–561.
- Murphy KP, Xie D, Garcia KC, Amzel LM, Friere E. 1993. Structural energetics of peptide recognition: Angiotensin II/antibody binding. *Proteins Struct Funct Genet* 15:113–120.
- Pearce KH, Ultsch MH, Kelley RF, de Vos AM, Wells JA. 1996. Structural and mutational analysis of affinity-inert contact residues at the growth hormone-receptor interface. *Biochemistry* 35:10300–10307.
- Pintar A, Hensmann M, Jumel K, Pitkeathely M, Harding SE, Campbell ID. 1996. Solution studies of the SH2 domain from the Fyn tyrosine kinase: Secondary structure, backbone dynamics and protein association. *Eur Biophys J* 24:371–380.
- Plummer MS, Holland DR, Shahripour A, Lunney EA, Fergus JH, Marks JS, McConnell P, Mueller WT, Sawyer TK. 1997a. Design, synthesis, and co-crystal structure of a nonpeptide SrcSH2 domain ligand. *J Med Chem* 40:3719–3725.
- Plummer MS, Lunney EA, Para KS, Shahripour A, Stankovic CJ, Humblet C, Fergus H, Marks JS, Herrera R, Hubbell S, et al. 1997b. Design of peptidomimetic ligands for the pp60^{src} SH2 domain. *Bioorg Med Chem* 5:41–47.
- Privalov PL, Tamura A. 1999. Comments on the Comments. *Protein Eng* 12:187.
- Songyang Z, Shoelson SE, Chaudhuri M, Gish G, Pawson T, Haser WG, King F, Roberts T, Ratnoffsky S, Lechleider RJ, et al. 1993. SH2 domains recognise specific phosphopeptide sequences. *Cell* 72:767–778.
- Soriano P, Montgomery C, Geske R, Bradley A. 1991. Targeted disruption of the c-src proto-oncogene leads to osteoporosis in mice. *Cell* 64:693–702.
- Spolar RS, Ha JH, Record MT Jr. 1989. Hydrophobic effect in protein folding and other noncovalent processes involving proteins. *Proc Natl Acad Sci USA* 86:8382–8385.
- Spolar RS, Livingstone JR, Record MT. 1992. Use of liquid hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of apolar and polar surface from water. *Biochemistry* 31:3947–3955.
- Spolar RS, Record MT. 1994. Coupling of local folding to site-specific binding of proteins to DNA. *Science* 263:777–784.
- Vajda S, Sippl M, Novotny J. 1997. Empirical potentials and functions for protein folding and binding. *Curr Opin Struct Biol* 7:222–228.
- Varadarajan R, Connolly PR, Sturtevant JM, Richards FM. 1992. Heat capacity changes for protein-peptide interactions in the ribonuclease S system. *Biochemistry* 31:1421–1426.
- Waksman G, Kominos D, Robertson SR, Pant N, Baltimore D, Birge RB, Cowburn D, Hanafusa H, Mayer BJ, Overduin M, et al. 1992. Crystal structure of the phosphotyrosine recognition domain SH2 of v-src complexed with tyrosine-phosphorylated peptides. *Nature* 358:646–653.
- Waksman G, Shoelson SE, Pant N, Cowburn D, Kuriya J. 1993. Binding of a high affinity phosphotyrosyl peptide to the SrcSH2 domain: Crystal structures of the complexed and peptide-free forms. *Cell* 72:779–790.
- Wallqvist A, Jernigan RL, Covell DG. 1995. A preference-based free-energy parameterization of enzyme-inhibitor binding. Applications to HIV-1-protease inhibitor design. *Protein Sci* 4:1881–1903.
- Weng Z, Vajda S, Delisi C. 1996. Prediction of protein complexes using empirical free energy functions. *Protein Sci* 5:614–626.
- Xu W, Doshi A, Lei M, Eck MJ, Harrison SC. 1999. Crystal structures of c-src reveal new features of its autoinhibitory mechanism. *Mol Cell* 3:629–638.
- Xu W, Harrison SC, Eck MJ. 1997. Three-dimensional structure of the tyrosine kinase c-src. *Nature* 385:595–602.
- Xu RX, Word JM, Davis DG, Rink MJ, Willard DH Jr, Gampe RT Jr. 1995. Solution structure of the human Pp60^{c-src} Sh2 domain complexed with a phosphorylated tyrosine pentapeptide. *Biochemistry* 34:2107–2121.