J|A|C|S ARTICLES

Published on Web 11/03/2009

Thermodynamic and Structural Effects of Conformational Constraints in Protein−Ligand Interactions. Entropic Paradoxy Associated with Ligand Preorganization

John E. DeLorbe, John H. Clements, Martin G. Teresk, Aaron P. Benfield, Hilary R. Plake, Laura E. Millspaugh, and Stephen F. Martin*

Department of Chemistry and Biochemistry, The Institute of Cellular and Molecular Biology, and The Texas Institute of Drug and Diagnostic Development, The University of Texas, Austin, Texas 78712

Received June 16, 2009; E-mail: sfmartin@mail.utexas.edu

Abstract: Succinate- and cyclopropane-derived phosphotyrosine (pY) replacements were incorporated into a series of Grb2 SH2 binding ligands wherein the pY+1 residue was varied to determine explicitly how variations in ligand preorganization affect binding energetics and structure. The complexes of these ligands with the Grb2 SH2 domain were examined in a series of thermodynamic and structural investigations using isothermal titration calorimetry and X-ray crystallography. The binding enthalpies for all ligands were favorable, and although binding entropies for all ligands having a hydrophobic residue at the pY+1 site were favorable, binding entropies for those having a hydrophilic residue at this site were unfavorable. Preorganized ligands generally bound with more favorable Gibbs energies than their flexible controls, but this increased affinity was the consequence of relatively more favorable binding enthalpies. Unexpectedly, binding entropies of the constrained ligands were uniformly disfavored relative to their flexible controls, demonstrating that *the widely held belief that ligand preorganization should result in an entropic advantage is not necessarily true.* Crystallographic studies of complexes of several flexible and constrained ligands having the same amino acid at the pY+1 position revealed extensive similarities, but there were some notable differences. There are a greater number of direct polar contacts in complexes of the constrained ligands that correlate qualitatively with their more favorable binding enthalpies and Gibbs energies. There are more single water-mediated contacts between the domain and the flexible ligand of each pair; although fixing water molecules at a protein−ligand interface is commonly viewed as entropically unfavorable, entropies for forming these complexes are favored relative to those of their constrained counterparts. Crystallographic b-factors in the complexes of constrained ligands are greater than those of their flexible counterparts, an observation that seems inconsistent with our finding that entropies for forming complexes of flexible ligands are relatively more favorable. This systematic study highlights the profound challenges and complexities associated with predicting how structural changes in a ligand will affect enthalpies, entropies, and structure in protein−ligand interactions.

Introduction

A major goal in contemporary bioorganic and medicinal chemistry is the design of small molecules that bind with high affinities to proteins. Toward this goal, numerous computational methods and scoring functions have been developed to predict association constants, and when there are well parametrized training sets, some of these are reasonably good as qualitative predictors of protein−ligand affinities; however, none are reliable as quantitative tools for ordering relative affinities for a range of different biological systems.[1] Indeed, despite the increasing availability of data relating to the energetics of protein−ligand interactions,[2] the ability to predict accurately the binding affinity of a given ligand and derivatives thereof for a target protein remains the Holy Grail for medicinal chemists and those studying molecular recognition in biological systems.[3]

The binding affinity, K_a, for a protein−ligand interaction is reflected by the net change in Gibbs energy, $\Delta G°$, which comprises enthalpic, $\Delta H°$, and entropic, $\Delta S°$, contributions (i.e., $\Delta G° = \Delta H° - T\Delta S°$), that occurs upon complexation according to the expression $\Delta G° = -RT \ln K_a$. The optimization of protein−ligand interactions thus requires modifying the structure of a ligand in a way that results in a more negative binding enthalpy and a more positive binding entropy.

(1) For lead references on the limitations of computational methods for predicting binding affinities, see: (a) Warren, G. L.; Andrews, C. W.; Capelli, A.-M.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Wall, I. D.; Woolven, J. M.; Peishoff, C. E.; Head, M. S. *J. Med. Chem.* **2006**, *49*, 5912–5931. (b) Klebe, G. *Drug Discovery Today* **2006**, *11*, 580–594. (c) Damm, K. L.; Carlson, H. A. *J. Am. Chem. Soc.* **2007**, *129*, 8225–8235.

(2) Liu, T.; Lin, Y.; Wen, X.; Jorissen, R. N.; Gilson, M. K. *Nucleic Acids Res.* **2007**, *35*, Database issue D198-D201.

(3) For excellent and comprehensive reviews of energetics in bimolecular interactions, see: (a) Houk, K. N.; Leach, A. G.; Kim, S. P.; Zhang, X. *Angew. Chem., Int. Ed.* **2003**, *42*, 4872–4897. (b) Zhou, H.-X.; Gilson, M. K. *Chem. Rev.* **2009**, *109*, 4092–4107.

10.1021/ja904698q CCC: $40.75

Obtaining a more favorable enthalpy of binding typically entails increasing the noncovalent associations that arise from polar, van der Waals, and charge−dipole interactions.[4] This objective is not easily achieved, however, because desolvation of polar groups on a ligand is energetically more costly than desolvating nonpolar groups.[5] Further complicating the task is the difficulty associated with controlling the geometries of individual polar interactions between the ligand and the protein, the energies of which are both distance and angle dependent.

A more favorable binding entropy may be achieved by reducing unfavorable solvation and conformational parameters associated with the ligand.[6,7] One common tactic to enhance the entropy of solvation involves increasing the hydrophobicity of the ligand, and although this often leads to an improved binding affinity, formation of the resultant complex is not necessarily accompanied by the expected entropy-driven signature characteristic of the hydrophobic effect.[8] For example, adding methylene groups to increase the hydrophobicity of a ligand can lead to a *less* favorable binding entropy and a *more* favorable binding enthalpy.[9] Indeed, a search of the literature reveals numerous examples wherein binding of two nonpolar molecules is characterized by a large enthalpic driving force rather than an entropic one.[10,11]

Constraining a flexible ligand in the three-dimensional shape it adopts when bound to a receptor, namely its biologically active conformation, can also result in increased association constants.[12−14] This enhanced affinity has been commonly attributed to the more favorable configurational entropy of binding that is expected from reducing the dynamic motion of a ligand prior to its complexation with the protein (i.e., $\Delta S^\circ < \Delta S^{\circ\prime}$) (Figure 1). An implicit assumption in this reasoning is that solvent and protein interact in the same way with both the flexible and constrained ligands so that no significant change in binding enthalpy is expected (i.e., $\Delta H^\circ \approx \Delta H^{\circ\prime}$). However, increases in potencies accompanying ligand preorganization are often less than 10-fold, an amount somewhat less than what would be expected based upon the accepted energetic estimates of 0.7−1.6 kcal mol^{-1} (i.e., ∼2.3−5.3 eu at 25 °C) for completely restricting an independent rotor.[15] Indeed, the small energetic benefits observed for constraining rotors in the arena of host−guest chemistry led Schneider to question whether reducing the

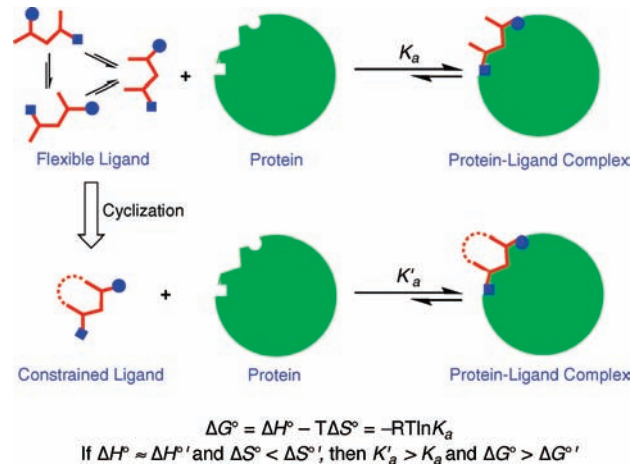


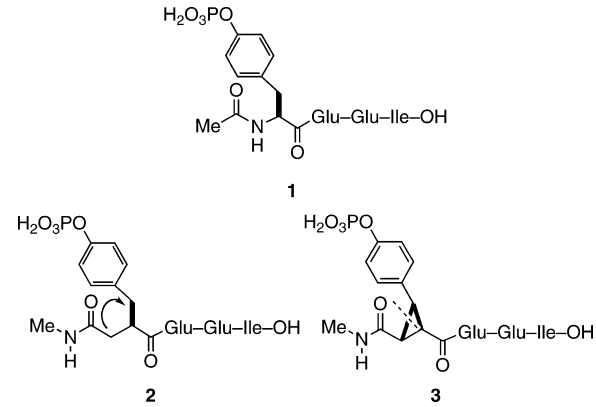
Figure 1. Simple model of the putative energetic effects associated with ligand preorganization. Cyclization of a flexible ligand limits the degrees of freedom and reduces the number of conformational isomers in solution, so the probability that the ligand adopts its biologically active conformation in solution is enhanced, thereby resulting in a more favorable entropy of binding. This analysis presumes that the two ligands interact similarly with the solvent and protein so that the binding enthalpies for the two are approximately the same.

(4) For examples of parametrizing binding enthalpy, see: Luque, I.; Freire, E. *Proteins: Struct., Funct., Genet.* **2002**, *49*, 181–190.
(5) Cabani, S.; Gianni, P.; Mollica, V.; Lepori, L. *J. Solution Chem.* **1981**, *8*, 563–595.
(6) (a) Ruben, A. J.; Kiso, Y.; Freire, E. *Chem. Biol. Drug Des.* **2006**, *67*, 2–4. (b) Lafont, V.; Armstrong, A. A.; Ohtaka, H.; Kiso, Y.; Amzel, L. M.; Freire, E. *Chem. Biol. Drug Des.* **2007**, *69*, 413–422.
(7) For detailed analyses of intermolecular interactions, see: (a) Gohlke, H.; Klebe, G. *Angew. Chem., Int. Ed.* **2002**, *41*, 2644–2676. (b) Hunter, C. A. *Angew. Chem., Int. Ed.* **2004**, *43*, 5310–5324. (c) Williams, D. H.; Stephens, E.; O'Brien, D. P.; Zhou, M. *Angew. Chem., Int. Ed.* **2004**, *43*, 6596–6616.
(8) For some leading articles on hydrophobic effects associated with nonpolar and polar groups, see: (a) Bartlett, P. A.; Yusuff, N.; Rico, A. C.; Lindvall, M. K. *J. Am. Chem. Soc.* **2002**, *124*, 3853–3857. (b) Southall, N. T.; Dill, K. A.; Haymet, A. D. J. *J. Phys. Chem. B* **2002**, *106*, 521–533. (c) Kyte, J. *Biophys. Chem.* **2003**, *100*, 193–203. (d) Chalikian, T. V. *Biopolymers* **2003**, *70*, 492–496.
(9) Malham, R.; Johnstone, S.; Bingham, R. J.; Barratt, E.; Phillips, S. E. V.; Laughton, C. A.; Homans, S. W. *J. Am. Chem. Soc.* **2005**, *127*, 17061–17067.
(10) For a review of so-called nonclassical hydrophobic effects, see: Meyer, E. A.; Castellano, R. K.; Diederich, F. *Angew. Chem., Int. Ed.* **2003**, *42*, 1210–1250.
(11) For some leading references to enthalpy driven hydrophobic associations, see: (a) Chervenak, M. C.; Toone, E. J. *J. Am. Chem. Soc.* **1994**, *116*, 10533–10539. (b) Carey, C.; Cheng, Y.-K.; Rossky, P. *J. Chem. Phys.* **2000**, *258*, 415–425. (c) Bingham, R. J.; Findlay, J. H. B. C.; Hsieh, S.-Y.; Kalverda, A. P.; Kjellberg, A.; Perazzolo, C.; Phillips, S. E. V.; Seshadri, K.; Trinh, C. H.; Turnbull, W. B.; Bodenhausen, G.; Homans, S. W. *J. Am. Chem. Soc.* **2004**, *126*, 1675–1681. (d) Homans, S. W. *Drug Discovery Today* **2007**, *12*, 534–539.
(12) For reviews, see:(a) Kessler, H. *Angew. Chem., Int. Ed. Engl.* **1982**, *21*, 512–523. (b) Cram, D. J. *Angew. Chem., Int. Ed. Engl.* **1986**, *25*, 1039–1057. (c) Böhm, H.-J.; Klebe, G. *Angew. Chem., Int. Ed.* **1996**, *35*, 2588–2614. (d) Hanessian, S.; McNaughton-Smith, G.; Lombart, H.-G.; Lubell, W. D. *Tetrahedron* **1997**, *53*, 12789–l2854. (e) Mann, A. In *The Practice of Medicinal Chemistry*, 2nd ed.; Wermuth, C. G., Ed.; Academic Press: London, U.K., 2003; pp 233−250. (f) Nakanishi, H.; Kahn, M. Design of Peptidomimetics. In *The Practice of Medicinal Chemistry*, 2nd ed.; Wermuth, C. G., Ed.; Academic Press: London, U.K., 2003; pp 477−500. (g) Loughlin, W. A.; Tyndall, J. D. A.; Glenn, M. P.; Fairlie, D. P. *Chem. Rev.* **2004**, *104*, 6085–6118.
(13) For some comparisons wherein the flexible and constrained ligands have the same number of heavy atoms, see: (a) Widlanski, T.; Bender, S. L.; Knowles, J. R. *J. Am. Chem. Soc.* **1989**, *111*, 2299–2300. (b) Morgan, B. P.; Holland, D. R.; Matthews, B. W.; Bartlett, P. A. *J. Am. Chem. Soc.* **1994**, *116*, 3251–3260. (c) Ettmayer, P.; France, D.; Gounarides, J.; Jarosinski, M.; Martin, M. S.; Rondeau, J. M.; Sabio, M.; Topiol, S.; Weidmann, B.; Zurini, M.; Bair, K. W. *J. Med. Chem.* **1999**, *42*, 971–980. (d) Hansen, K. K.; Grosch, B.; Greiveldinger-Poenaru, S.; Bartlett, P. A. *J. Org. Chem.* **2003**, *68*, 8465–8470. (f) Tsantrizos, Y. S.; Bolger, G.; Bonneau, P.; Cameron, D. R.; Goudreau, N.; Kukolj, G.; LaPlante, S. R.; Llinas-Brunet, M.; Nar, H.; Lamarre, D. *Angew. Chem., Int. Ed.* **2003**, *42*, 1355–1360.
(14) For some comparisons wherein the flexible and constrained ligands differ by one or two heavy atoms, see: (a) Meyer, J. H.; Bartlett, P. A. *J. Am. Chem. Soc.* **1998**, *120*, 4600–4609. (b) Smith, W. W.; Bartlett, P. A. *J. Am. Chem. Soc.* **1998**, *120*, 4622–4628. (c) Marquis, R. W.; et al. *J. Med. Chem.* **2001**, *44*, 1380–1395. (d) Dekker, F. J.; de Mol, N. J.; Fischer, M. J. E.; Kemmink, J.; Liskamp, R. M. J. *Org. Biomol. Chem.* **2003**, *1*, 3297–3303. (e) Nam, N.-H.; Ye, G.; Sun, G.; Parang, K. *J. Med. Chem.* **2004**, *47*, 3131–3141. (f) Ghosh, A. K.; Swanson, L. M.; Cho, H.; Leshchenko, S.; Hussain, K. A.; Kay, S.; Walters, D. E.; Koh, Y.; Mitsuya, H. *J. Med. Chem.* **2005**, *48*, 3576–3585.
(15) For leading references to energetic contributions to binding of freezing rotors, see: (a) Gerhard, U.; Searle, M. S.; Williams, D. H. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 803–808. (b) Khan, A. R.; Parrish, J. C.; Fraser, M. E.; Smith, W. W.; Bartlett, P. A.; James, M. N. G. *Biochemistry* **1998**, *37*, 16839–16845. (c) Hossain, M. A.; Schneider, H.-J. *Chem.—Eur. J.* **1999**, *5*, 1284–1290. (d) See also in: Schneider, H.-J. In *Protein-Ligand Interactions: From Molecular Recognition to Drug Design*; Böhm, H.-J., Schneider, G., Eds.; Wiley-VCH: Weinheim, 2003; pp 21−50.

number of single bonds in synthetic hosts was actually a viable strategy to enhance the binding affinity.[16]

Notwithstanding some notable exceptions,[13a,15b] ligand preorganization has not provided the magnitude of increased binding affinity that would be expected from the number of rotors being restricted; however, understanding the origin(s) of this shortcoming is problematic. For example, the lack of structural information for complexes of flexible and constrained ligand pairs makes it difficult to determine whether the two ligands interact similarly with the protein. Furthermore, because of the dynamic nature of protein−ligand interactions, it is not easy to assess the extent to which individual rotors of a flexible molecule are restricted upon binding. A related complicating factor is that ligands typically bind to proteins in conformations that are higher in energy than their global minima in solution,[17] and the effects of differential conformational strain energies upon the relative binding affinities of related compounds are poorly understood. The challenges associated with evaluating the specific energetic consequences of ligand preorganization are further exacerbated by the common failure to compare the affinity of a constrained molecule with an appropriate flexible control having the same number and type of heavy atoms, the same functionality, and the same number of hydrogen bond donors and acceptors.[13] Another impediment to understanding how ligand preorganization affects binding affinity is that potencies are typically reported as K_i's or IC_{50}'s, and the specific contributions to $\Delta S°$ and $\Delta H°$ of binding are rarely determined. Exceptions to this generalization are found in reports from our group,[18,19] as well as in work disclosed by Spaller, who has conducted a detailed analysis of thermodynamic parameters for the binding of a series of linear and macrocyclic ligands to a PDZ3 domain.[20] These and other studies further reveal how favorable changes in binding enthalpies or entropies are attenuated by enthalpy−entropy compensation, an ubiquitous phenomenon in bimolecular interactions.[21,22]

Our first venture toward correlating ligand preorganization and energetics in protein−ligand interactions involved introducing cyclopropane-derived amino acid replacements into analogs of known enzyme inhibitors.[23]Although these studies led to the discovery of potent inhibitors of a number of enzymes, we, like others, found that the benefits of ligand preorganization fell short of expectations. These results inspired us to initiate studies that would explicitly elucidate the energetic and structural effects of preorganizing pseudopeptide ligands by incorporating cyclopropane rings as conformational constraints in flexible controls. Toward this objective, the phosphotyrosine residue of Ac−pTyr-Glu-Glu-Ile−OH (Ac−pY-E-E-I−OH) (**1**), a tetrapeptide sequence present in peptides that bind to the Src SH2 domain,[24] was replaced with substituted succinyl and cyclopropane dicarboxyl moieties leading to the pseudopeptides **2** and **3**.[18] Cyclization of **2** to give the cyclopropane-derived ligand **3** completely restricts two rotors, and the three rotors associated with the carbon−carbon single bonds joining the two carbonyl and the aryl substituents on the cyclopropane ring are partially restricted. Significantly, compounds **2** and **3** have the same number and type of heavy atoms, identical functional groups, and the same number of hydrogen bond donors and acceptors; they differ by only two hydrogen atoms arising from formation or scission of a carbon−carbon single bond as shown in **2** and **3**. Six related pseudopeptides were also prepared in which each Glu residue in **2** and **3** was replaced with an Asp residue, and the Ile residue was substituted with a Val residue.[25]



The thermodynamic parameters for complexation of each of these ligands with recombinant Src SH2 domain were determined by isothermal titration calorimetry (ITC). The constrained ligand **3** and its derivatives containing Asp or Val replacements at the pY+1 − pY+3 positions invariably bound with more favorable entropies than their flexible counterparts, and the observed average entropic advantage of ∼7.5 eu correlated modestly with the total number of rotors restricted.[15,25] However, this expected entropic advantage was always balanced by an enthalpic penalty that eventuated in comparable binding affinities for each ligand in a flexible/constrained pair. Because of this balancing enthalpy−entropy compensation, no net energetic advantage attended ligand preorganization. We determined the structure of the complex of **3** with the domain, and comparing this structure with that of an 11-mer peptide containing the pTyr-

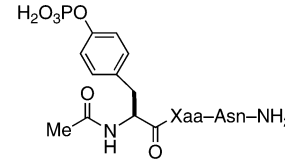
(16) Eblinger, F.; Schneider, H.-J. *Angew. Chem., Int. Ed.* **1998**, *37*, 826–829.
(17) Perola, E.; Charifson, P. S. *J. Med. Chem.* **2004**, *47*, 2499–2510.
(18) Davidson, J. P.; Lubman, O.; Rose, T.; Waksman, G.; Martin, S. F. *J. Am. Chem. Soc.* **2002**, *124*, 205–215.
(19) For a preliminary account of some of the results presented in this account, see: Benfield, A. P.; Teresk, M. G.; Plake, H. R.; DeLorbe, J. E.; Millspaugh, L. E.; Martin, S. F. *Angew. Chem., Int. Ed.* **2006**, *45*, 6830–6835.
(20) (a) Li, T.; Saro, D.; Spaller, M. R. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 1385–1388. (b) Udugamasooriya, D. G.; Spaller, M. R. *Biopolymers* **2008**, *8*, 653–667, and references therein.
(21) For reviews, see: (a) Sharp, K. *Protein Sci.* **2001**, *10*, 661–667. (b) Liu, L.; Guo, Q.-X. *Chem. Rev.* **2001**, *101*, 673–695.
(22) For some leading references of enthalpy−entropy compensation, see: (a) Gilli, P.; Ferretti, V.; Gillli, G.; Borea, P. A. *J. Phys. Chem.* **1994**, *98*, 1515–1518. (b) Dunitz, J. D. *Chem. Biol.* **1995**, *2*, 709–712. (c) Calderone, C. T.; Williams, D. H. *J. Am. Chem. Soc.* **2001**, *123*, 6262–6267. (d) Cooper, A.; Johnson, C. M.; Lakey, J. H.; Nöllmann, M. *Biophys. Chem.* **2001**, *93*, 215–230. (e) Talhout, R.; Villa, A.; Mark, A. E.; Engberts, J. B. F. N. *J. Am. Chem. Soc.* **2003**, *125*, 10570–10579. (f) Williams, D. H.; Stephens, E.; O'Brien, D. P.; Zhou, M. *Angew. Chem., Int. Ed.* **2004**, *43*, 6596–6616. (g) Krishnamurthy, V. M.; Bohall, B. R.; Semetey, V.; Whitesides, G. M. *J. Am. Chem. Soc.* **2006**, *128*, 5802–5812.
(23) For a review, see: (a) Reichelt, A.; Martin, S. F. *Acc. Chem. Res.* **2006**, *39*, 433–442. See also: (b) Martin, S. F.; Austin, R. E.; Oalmann, C. J.; Baker, W. R.; Condon, S. L.; deLara, E.; Rosenberg, S. H.; Spina, K. P.; Stein, H. H.; Cohen, J.; Kleinert, H. D. *J. Med. Chem.* **1992**, *35*, 1710–1721. (c) Martin, S. F.; Dorsey, G. O.; Gane, T.; Hillier, M. C.; Kessler, H.; Baur, M.; Matha, B.; Erickson, J. W.; Bhat, T. N.; Munshi, S.; Gulnick, S. V.; Topol, I. A. *J. Med. Chem.* **1998**, *41*, 1581–1597. (d) Martin, S. F.; Oalmann, C. J.; Liras, S. *Tetrahedron* **1993**, *49*, 3521–3532. (e) Reichelt, A.; Gaul, C.; Frey, R.; Kennedy, A.; Martin, S. F. *J. Org. Chem.* **2002**, *67*, 4062–4075.
(24) For reviews of SH2 domains, see: (a) Cody, W. L.; Lin, Z.; Panek, R. L.; Rose, D. W.; Rubin, J. R. *Curr. Pharm. Des.* **2000**, *6*, 59–98. (b) Bradshaw, J. M.; Waksman, G. *Adv. Protein Chem.* **2002**, *61*, 161–210. (c) Machida, K.; Mayer, B. J. *Biochim. Biophys. Acta* **2005**, *1747*, 1–25.
(25) Davidson, J. P. Ph.D. Dissertation, The University of Texas, 2001.

Glu-Glu-Ile sequence revealed that all interatomic distances between the domain and the ligand in each of the two structures were in close agreement.[18] Although this structural study suggested that the cyclopropane-derived pY replacement in **3** is a viable rigid mimic of the bound conformation of phosphotyrosine in Src SH2 binding ligands, we were unable to identify the origin of the enthalpic penalty associated with preorganizing **2**.

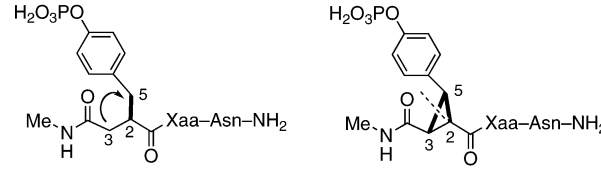
To our knowledge, these key experiments were the first to demonstrate explicitly that introducing a conformational constraint into a flexible ligand resulted in the expected entropic advantage for forming protein−ligand complexes. It was nevertheless perplexing that preorganizing flexible Src SH2 binding ligands did not afford analogs having higher affinities, despite their having more favorable binding entropies. To further explore the effects of preorganization upon energetics in protein−ligand interactions, we extended the scope of our inquiry to the binding of phosphotyrosine-containing peptide analogs to the SH2 domain of growth receptor binding protein 2 (Grb2), a 25 kDa cytosolic adapter protein that participates in the Ras signal transduction pathway.[24] The binding of pY residues on receptor tyrosine kinases (RTKs) to the SH2 domain of Grb2 leads to Ras activation and consequent cell growth and differentiation. There has thus been considerable interest in identifying compounds that selectively and potently inhibit binding of the Grb2 SH2 domain to RTKs as a strategy to modulate Ras signaling and to discover potential anticancer agents.[13c,26,27] The SH2 domain of Grb2 recognizes and binds pY peptides containing the amino acid sequence pTyr-Xaa-Asn (pYXN). Although the pTyr and the Asn residues are essential for high affinity binding to the Grb2 SH2 domain, there is considerable flexibility in the nature of the residue at the pY+1 position; it is usually hydrophobic, but Gln, Glu, and Lys residues are found at this position in potent Grb2 SH2 binding ligands.[28]

Results and Discussion

Ligand Design and Synthesis. Given the known structure−activity relationships of Grb2 SH2 binding ligands, our initial studies were focused upon derivatives of Ac-pTyr-Xaa-Asn-NH_2 (**4**).[19] Inspection of crystal structures of complexes of the Grb2 SH2 domain with different phosphotyrosine-containing peptides revealed that the interactions between the pY residue and the Grb2 SH2 domain were closely similar to those observed between the pY residues of **1** and **3** and the Src SH2 domain.[13c,26] It thus occurred to us that the flexible and constrained pY mimics found in **2** and **3**, respectively, might again serve as efficacious replacements of the pY residue in derivatives of **4**. Accordingly, we targeted the series of flexible ligands **5**−**10**, in which a benzyl succinyl moiety serves as a flexible pY replacement, and the corresponding series of constrained ligands **11**−**16**, in which a cyclopropane ring is a rigid mimic of the pY residue. Any flexible/constrained ligand pair would have enabled us to evaluate the energetic effects of preorganization; however, we reasoned that a broader study



4: Ac–pYXN

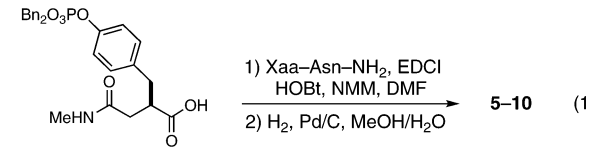


5: Xaa = Val (fpYVN)
6: Xaa = Ile (fpYIN)
7: Xaa = Leu (fpYLN)
8: Xaa = Gln (fpYQN)
9: Xaa = Glu (fpYEN)
10: Xaa = Lys (fpYKN)

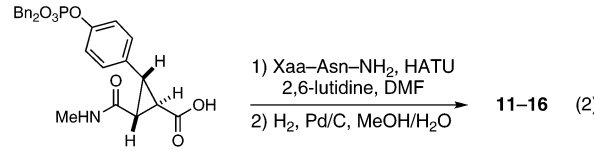
11: Xaa = Val (cpYVN)
12: Xaa = Ile (cpYIN)
13: Xaa = Leu (cpYLN)
14: Xaa = Gln (cpYQN)
15: Xaa = Glu (cpYEN)
16: Xaa = Lys (cpYKN)

wherein the polar and functional nature of the side chain of the pY+1 residue was varied might provide additional insights relative to how changing the hydrophobicity and charge of a ligand might affect binding energetics.

The syntheses of fpYVN (**5**) and cpYVN (**11**) by coupling the known acids **17** and **18**,[18] respectively, with a Val-Asn-NH_2 dipeptide, followed by hydrogenolysis of the *O*-benzyl groups, have been described,[29] and compounds **6**−**10** and **12**−**16** were prepared analogously (eqs 1 and 2).



17



18

Thermodynamic Properties of Flexible and Constrained Ligands. The thermodynamic parameters (K_a, $\Delta G°$, $\Delta H°$, $\Delta S°$) for binding of **5** and **11** to the Grb2 SH2 domain were first determined by titration studies using ITC.[30] Briefly, a solution

(26) For some structural studies, see: (a) Rahuel, J.; Gay, B.; Erdmann, D.; Strauss, A.; Garcia-Echeverria, C.; Furet, P.; Caravatti, G.; Fretz, H.; Schoepfer, J.; Grütter, M. G. *Nat. Struct. Biol.* **1996**, *3*, 586–589. (b) Ogura, K.; Tsuchiya, S.; Terasawa, H.; Yuzawa, S.; Hatanaka, H.; Mandiyan, V.; Schlessinger, J.; Inagaki, F. *J. Biomol. NMR* **1997**, *10*, 273–278. (c) Rahuel, J.; Garcia-Echeverria, C.; Furet, P.; Strauss, A.; Caravatti, G.; Fretz, H.; Schoepfer, J.; Gay, B. *J. Mol. Biol.* **1998**, *279*, 1013–1022. (d) Nioche, P.; Liu, W.-Q.; Broutin, I.; Charbonnier, F.; Latreille, M.-T.; Vidal, M.; Roques, B.; Garbay, C.; Ducruix, A. *J. Mol. Biol.* **2002**, *315*, 1167–1177. (e) Ogura, K.; Shiga, T.; Yokochi, M.; Yuzawa, S.; Burke, T. R., Jr.; Inagaki, F. *J. Biomol. NMR* **2008**, *42*, 197–207.

(27) For a review, see: (a) Fretz, H.; Furet, P.; Garcia-Echeverria, C.; Rahuel, J.; Schoepfer, J. *Curr. Pharm. Des.* **2000**, *6*, 1777–1796. See also: (b) McNemar, C.; Snow, M. E.; Windsor, W. T.; Prongay, A.; Mui, P.; Zhang, R.; Durkin, J.; Le, H. V.; Weber, P. C. *Biochemistry* **1997**, *36*, 10006–10014. (c) Garcia-Echeverria, C.; Furet, P.; Gay, B.; Fretz, H.; Rahuel, J.; Schoepfer, J.; Caravatti, G. *J. Med. Chem.* **1998**, *41*, 1741–1744. (d) Liu, W.-Q.; Vidal, M.; Olszowy, C.; Million, E.; Lenoir, C.; Dhôtel, H.; Garbay, C. *J. Med. Chem.* **2004**, *47*, 1223–1233.

(28) Kessels, H. W. H. G.; Ward, A. C.; Schumacher, T. N. M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 8524–8529.

(29) (a) Davidson, J. P.; Martin, S. F. *Tetrahedron Lett.* **2000**, *41*, 9459–9564. (b) Plake, H. R.; Sundberg, T. B.; Woodward, A. R.; Martin, S. F. *Tetrahedron Lett.* **2003**, *44*, 1571–1574.

(30) For a review of the use of ITC, see: Ladbury, J. E. *Thermochim. Acta* **2001**, *380*, 209–215.

Table 1. Thermodynamic Parameters for Complex Formation between the Grb2 SH2 Domain and Pseudopeptides **5**−**16**[a]

ligand	K_a (M^{-1})	$\Delta G°$ (kcal mol^{-1})	$\Delta H°$ (kcal mol^{-1})	$\Delta S°$ (cal mol^{-1} K^{-1})	$-T\Delta S°$ (kcal mol^{-1})
5 (fpYVN)	$(4.5 \pm 0.12) \times 10^5$	-7.7 ± 0.02	-5.4 ± 0.14	7.9 ± 0.22	-2.4 ± 0.07
11 (cpYVN)	$(2.8 \pm 0.10) \times 10^6$	-8.8 ± 0.02	-7.9 ± 0.29	3.0 ± 0.30	-0.9 ± 0.09
6 (fpYIN)	$(4.0 \pm 0.15) \times 10^5$	-7.7 ± 0.02	-5.5 ± 0.20	7.4 ± 0.30	-2.2 ± 0.09
12 (cpYIN)	$(2.1 \pm 0.08) \times 10^6$	-8.6 ± 0.02	-8.3 ± 0.30	1.3 ± 0.30	-0.4 ± 0.09
7 (fpYLN)	$(1.7 \pm 0.06) \times 10^5$	-7.1 ± 0.02	-4.6 ± 0.17	8.6 ± 0.30	-2.6 ± 0.09
13 (cpYLN)	$(7.1 \pm 0.27) \times 10^5$	-8.0 ± 0.02	-6.0 ± 0.22	6.6 ± 0.30	-2.0 ± 0.09
8 (fpYQN)	$(5.6 \pm 0.15) \times 10^5$	-7.8 ± 0.02	-8.7 ± 0.23	-2.8 ± 0.22	0.8 ± 0.07
14 (cpYQN)	$(1.2 \pm 0.06) \times 10^6$	-8.3 ± 0.01	-9.8 ± 0.20	-5.2 ± 0.18	1.5 ± 0.05
9 (fpYEN)	$(3.0 \pm 0.08) \times 10^5$	-7.5 ± 0.02	-8.8 ± 0.23	-4.3 ± 0.22	1.3 ± 0.07
15 (cpYEN)	$(3.6 \pm 0.10) \times 10^5$	-7.6 ± 0.02	-10.3 ± 0.27	-9.0 ± 0.22	2.7 ± 0.07
10 (fpYKN)	$(9.8 \pm 0.23) \times 10^4$	-6.8 ± 0.02	-7.7 ± 0.20	-3.0 ± 0.21	0.9 ± 0.07
16 (cpYKN)	$(5.5 \pm 0.15) \times 10^5$	-7.8 ± 0.02	-9.2 ± 0.24	-4.6 ± 0.22	1.4 ± 0.07

[a] ITC experiments were conducted at 25 °C in duplicate in 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid (HEPES, 50 mM) with NaCl (150 mM) at pH 7.45 ± 0.05.[31] A value of 5.2% for the error in ligand concentration was used, and the error arising therefrom was propagated accordingly.[22g]

of monomeric Grb2 SH2 domain in buffer (50 mM HEPES, 150 mM NaCl, pH 7.45 ± 0.05) at 25 °C was titrated with ligand in the same buffer, and the raw data were processed to give K_a and $\Delta H°$. The $\Delta G°$ term was calculated indirectly by applying the modified Arrhenius equation, $\Delta G° = -RT \ln K_a$, and $\Delta S°$ was calculated using the Gibbs relationship $\Delta G° = \Delta H° - T\Delta S°$. These experiments revealed that the constrained ligand **11** bound approximately 6-fold better than its flexible counterpart **5** (Table 1). This increased affinity is qualitatively in accord with the *net* thermodynamic benefit that would be expected from ligand preorganization; however, *the entropy of binding for the constrained pseudopeptide* **11** was 4.9 eu less favorable than that for its flexible counterpart **5**. Hence, the increased binding affinity of **11** relative to **5**, which corresponds to a $\Delta\Delta G°$ of 1.1 kcal mol^{-1}, arose from an enthalpic advantage of 2.5 kcal mol^{-1} and *not* the more favorable entropy of binding that is commonly associated with restricting rotors.[15]

Because the results of these studies were inconsistent with the conventional wisdom regarding the putative energetic, especially entropic, effects of ligand preorganization, the thermodynamic parameters for the binding of **6**−**10** and **12**−**16** to the Grb2 SH2 domain were determined by ITC to ascertain whether other flexible/constrained ligand pairs exhibited similar thermodynamic behavior; these results are collected in Table 1 and Figure 2. Each of the constrained pseudopeptides **12**−**16** bound with $\Delta G°$'s that were more favorable than those of their respective flexible controls **6**−**10** by 0.1−1.0 kcal mol^{-1}. Consistent with observations for the flexible/constrained ligand pair **5** and **11**, the increased affinities for the preorganized ligands **12**−**16** were uniformly the consequence of binding enthalpies that were more favorable by 1.1−2.8 kcal mol^{-1}. *Without exception the binding entropies for the preorganized ligands* **12**−**16** were unfavorable relative to their flexible counterparts **6**−**10** by 1.6−6.1 eu or 0.5−1.8 kcal mol^{-1} at 25 °C.

Protein−ligand binding of all ligands **5**−**16** was primarily enthalpy driven as binding enthalpies contributed more to the overall energetics of complex formation than binding entropies. Depending upon the side chain at the pY+1 position, binding

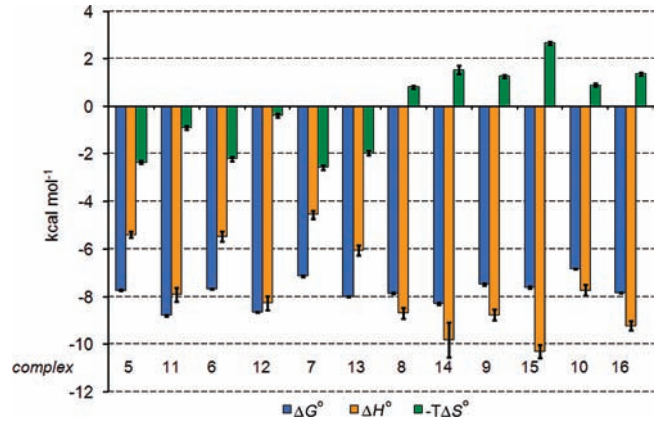


Figure 2. Graphical representation of the thermodynamic parameters for complex formation between the Grb2 SH2 domain and pseudopeptides **5**−**16** as determined by ITC.

entropies were found to be either favorable or unfavorable. For example, the $\Delta S°$ term for **5**−**7** and **11**−**13**, each of which has a hydrophobic residue at the pY+1 site, is favorable, whereas the $\Delta S°$ term for **8**−**10** and **14**−**16**, wherein this side chain is either polar (i.e., $-CONH_2$) or charged at pH ~7.5 (i.e., $-CO_2^-$ or $-NH_3^+$), is unfavorable.

The stunning finding that preorganization is not necessarily accompanied by the anticipated favorable change in entropy but rather by an enthalpic advantage prompted us to query why. The overall energetics of complex formation can be broadly partitioned into the thermodynamics associated with (1) proton transfer and desolvation of the ligand and the binding site of the protein; (2) new nonbonded interactions, including polar and van der Waals, that form between the ligand and the protein; and (3) conformational and dynamic changes of both ligand and protein. Each of these factors was probed to some degree in the following experimental studies.

Effect of Proton Transfer on Binding Energetics. Protein binding of ligands having ionizable groups such as a phosphate group, which plays a significant role in the overall binding energetics of **5**−**16**, may be accompanied by proton transfer. Values determined for $\Delta H°$ will then comprise not only an enthalpy term for binding, $\Delta H°_{bind}$, but also an enthalpy term that is associated with proton exchange. The contribution from this term is dependent upon the heat of ionization of the buffer,

(31) The values of the thermodynamic parameters reported herein for **5**−**7** and **11**−**13** differ slightly from those reported in ref 18 because ITC experiments were conducted using a different protocol for drying ligands that led to *n* values closer to 1.0.

$\Delta H°_{ion}$, and the number of protons, n, transferred according to eq 3.[32,33]

$$\Delta H° = \Delta H°_{bind} + n\Delta H°_{ion} \quad (3)$$

To ascertain the extent to which the binding enthalpies included contributions from proton transfers, we determined $\Delta H°$ for the pair of flexible and constrained ligands **5** and **11** in PIPES, HEPES, and imidazole buffers, which have corresponding heats of ionization of 2.7, 5.0, and 8.8 kcal mol^{-1}.[27b] The slopes of the plots of $\Delta H°$ values against $\Delta H°_{ion}$ for **5** and **11** in these buffers reveal that the number of protons transferred upon complexation was 0.27 and 0.20 for **5** and **11**, respectively (Figure 3). This corresponds to a contribution to $\Delta H°$ from proton transfer processes of ~1.0−1.4 kcal mol^{-1} in HEPES, the buffer used in the ITC studies. Because the number of proton equivalents transferred is nearly the same for **5** and **11**, the contribution from $\Delta H°_{ion}$ to $\Delta H°$ of binding for each is within experimental error of being identical. Consequently, *the enhanced $\Delta H°$ for binding associated with preorganizing **5** does not appear to arise from any significant difference in proton transfer processes.* We posit that this conclusion holds for the other flexible/constrained ligand pairs in this study, although the number of protons transferred might be different owing to the presence of acidic and basic groups on the side chains of the Glu- and Lys-derived ligands **9** and **15** and **10** and **16**, respectively.

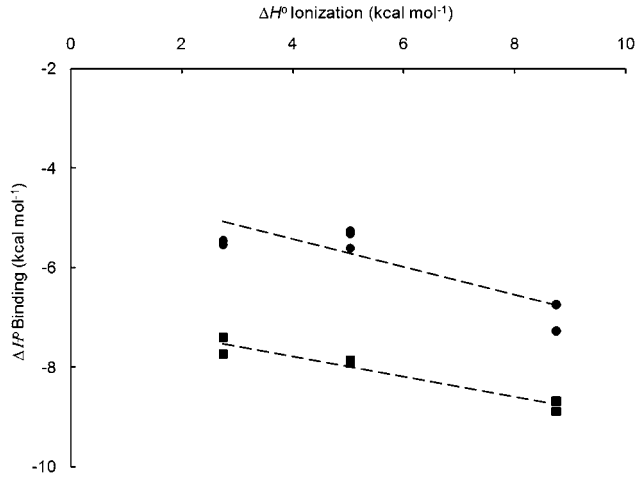


Figure 3. Dependence of the measured $\Delta H°$ of binding of **5** (●) and **11** (■) at pH 7.5 and 25 °C upon the heat of ionization of the buffer: PIPES, $\Delta H°_{ion}$ = 2.7 kcal mol^{-1}; HEPES, $\Delta H°_{ion}$ = 5.0 kcal mol^{-1}; imidazole, $\Delta H°_{ion}$ = 8.8 kcal mol^{-1}. The slope of each line is proportional to the contribution of proton transfer to $\Delta H°$ of binding.

Temperature Dependence of Binding Thermodynamics. Temperature dependent studies of $\Delta H°$ for bimolecular interactions give the heat capacity changes, ΔC_p, for complex formation as defined in eq 4. The significance of this parameter has been ascribed to a number of phenomena,[34] but in biological systems, ΔC_p serves primarily as a barometer of the hydrophobic effect. A negative value of ΔC_p signifies the burial of nonpolar surfaces, whereas a positive value is attributed to the burial of polar surfaces.[35] Hence, in protein−ligand interactions, a comparison of ΔC_ps for the binding of closely related ligands to a protein provides an indication of whether desolvation effects might be involved in differential binding affinities.

$$\Delta C_p = \left(\frac{\partial(\Delta H^o)}{\partial T}\right)_p = T\left(\frac{\partial(\Delta S^o)}{\partial T}\right)_p \quad (4)$$

The ΔC_ps for **5** and **11**, each of which binds with a favorable entropy, and **8** and **14**, each of which binds with an unfavorable entropy, were determined by obtaining binding enthalpies for these ligands at three different temperatures within the range 15−35 °C. This study was designed to ascertain how ΔC_ps might vary depending upon whether binding entropies are positive or negative. The values of ΔC_p for **5** and **11** derived from plots of $\Delta H°$ versus temperature, T, were found to be −158.5 ± 8.6 and −167.5 ± 6.2 cal mol^{-1} K^{-1}, respectively, whereas the corresponding values for **8** and **14** were −133.0 ± 15.1 and −125.8 ± 29.5 cal mol^{-1} K^{-1} (Figure 4). The negative values of these ΔC_ps are in accord with the burial of nonpolar surfaces upon binding, and the less negative values for **8** and **14** relative to **5** and **11** are qualitatively consistent with the presence of a more polar side chain at the pY+1 position. More important from the perspective of evaluating the energetic effects of preorganization, the values of ΔC_p for each ligand in a flexible/constrained pair are the same within experimental error, suggesting that *differences in the binding free energies of pairs of flexible and constrained ligands having the same pY+1 residue do not appear to arise from desolvation or hydrophobic effects*.

Although the thermodynamic parameter ΔC_p provides some information relative to desolvation effects in protein−ligand interactions, it does not give a direct measure of differences in ligand hydrophobicity. Such information may be obtained by comparing solvation free energies,[36] but the nonvolatile nature of **5**−**16** precluded a direct experimental determination of solvation free energy. The contribution of ligand desolvation to binding can also be estimated by transferring the ligand from an aqueous buffer to an organic solvent such as octanol.[37] However, owing to the presence of the ionized phosphate group, none of the ligands of this study were sufficiently soluble in octanol at pH 7.5, the pH at which ITC experiments were performed, to determine the partition coefficient.

Effects of Enthalpy−Entropy Compensation in Binding Energetics. Examination of the data in Table 1 reveals that $\Delta G°$s for forming complexes between the Grb2 SH2 domain and the constrained ligands **11**−**16** are more favorable than for their respective flexible controls **5**−**10** by up to 1.1 kcal mol^{-1}. The variation in binding enthalpy for a pair of flexible and constrained ligands is as large as 2.8 kcal mol^{-1}, but this enthalpic advantage is partially offset by an unfavorable binding entropy of as much as 1.8 kcal mol^{-1}. Hence, *enthalpy−entropy compensation for forming these complexes of corresponding*

(32) For example, see: (a) Baker, B. M.; Murphy, K. P. *Biophys. J.* **1996**, *71*, 2049–2055. (b) Fukuda, H.; Takahashi, K. *Proteins: Struct., Funct., Genet.* **1998**, *33*, 159–166. (c) Bradshaw, J. M.; Waksman, G. *Biochemistry* **1998**, *37*, 15400–15407.
(33) For example, see: (a) Bradshaw, J. M.; Waksman, G. *Biochemistry* **1998**, *37*, 15400–15407.
(34) (a) Sturtevant, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 2236–2240. (b) Spolar, R. S.; Ha, J. H.; Record, M. T., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 8382–8385.
(35) For a review, see: Southall, N. T.; Dill, K. A.; Haymet, A. D. J. *J. Phys. Chem. B* **2002**, *106*, 521–533.
(36) For tables of experimental solvation Gibbs energies and computional methods for their calculation, see: (a) Wang, J.; Wang, W.; Huo, S.; Lee, M.; Kollman, P. A. *J. Phys. Chem. B* **2001**, *105*, 5055–5067. (b) Chuman, H.; Mori, A.; Tanaka, H. *Anal. Sci.* **2002**, *18*, 1015–1020. (c) Gallicchio, E.; Zhang, L. Y.; Levy, R. M. *J. Comput. Chem.* **2002**, *23*, 517–529.

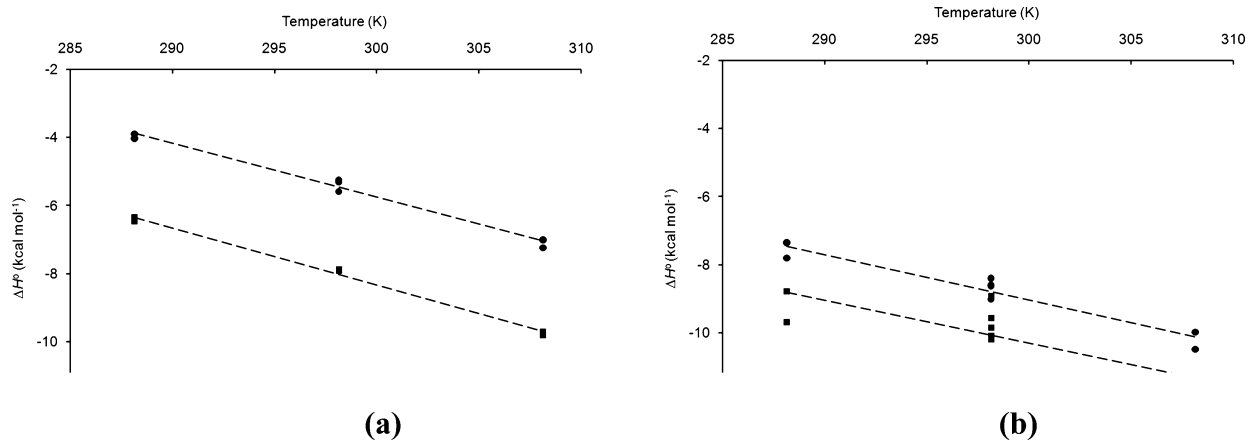


Figure 4. $\Delta H°$ of ligand binding as a function of temperature. Measurements were performed in 50 mM HEPES buffer (150 mM NaCl, pH 7.45 ± 0.05) at three different temperatures. At least two measurements were performed at each temperature. ΔC_p values for the binding of each ligand were obtained from the slope of the plots with the error in ΔC_p being the standard error in the slope. (a) Plot for the flexible ligand **5** (●) and the constrained ligand **11** (■). (b) Plot for the flexible ligand **8** (●) and the constrained ligand **14** (■).

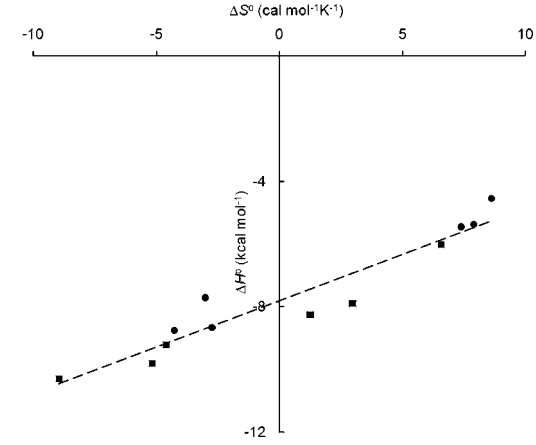


Figure 5. Correlation between $\Delta H°$ and $\Delta S°$ of binding as obtained from ITC measurements for flexible ligands (●) and constrained ligands (■).

flexible and constrained ligands is not balancing, and relative changes in $\Delta H°$ dominate those in $\Delta S°$.

The range of $\Delta G°$ values for the set of ligands **5**−**16** was 2.0 kcal mol^{-1}, whereas the values of $\Delta H°$ varied by 5.7 kcal mol^{-1} and $T\Delta S°$ differed by 5.3 kcal mol^{-1}. It follows directly from the Gibbs relationship, $\Delta G° = \Delta H° - T\Delta S°$, that when variations in $\Delta G°$ are small, a change in $\Delta H°$ will be offset by a compensating change in $\Delta S°$. A plot of $\Delta H°$ versus $\Delta S°$ will then be linear, and the slope of the line thus obtained is the compensation temperature, T_c. Based upon such a plot for the binding of **5**−**16** (Figure 5), T_c was found to be 296 K. The test of Krug (eq 5),[38] where σ is the standard error in the slope obtained from linear regression analysis and T is the temperature at which the titration experiments were conducted, was applied to the data to ascertain whether the difference between T_c and T was statistically significant at a 95% confidence level.[21a] Because application of this relationship gives a range for T of

$$T_c - 1.81\sigma < T < T_c + 1.81\sigma \quad (5)$$

243−350 K for experiments that were performed at 298 K, the observed enthalpy−entropy compensation for **5**−**16** appears to be a consequence of the Gibbs free energy relationship rather than some extra thermodynamic phenomenon.

Structural Studies of Complexes of Flexible and Constrained Ligands with the Grb2 SH2 Domain. In the next stage of the inquiry, we undertook X-ray crystallographic investigations of complexes of flexible and constrained ligands having the same amino acid residue at the pY+1 position. We recognize, however, that interpreting such structural data is subject to the caveats that differences in resolution necessarily lead to uncertainties in structure and contact distances and angles and that the static interactions observed in the solid state may not accurately reflect the dynamic interactions in solution. Nevertheless, such studies are commonly relied upon to ascertain whether structurally related ligands bind similarly to proteins.[39] By comparing the binding modes and interactions for each member of different flexible/constrained ligand pairs, we hoped, perhaps naively, to identify structural differences that might be correlated with differential binding energetics.

Suitable crystals for X-ray analysis were obtained for complexes of the Grb2 SH2 domain with **5**, **6**, **8**, **11**, **12**, and **14**.[40] Data for the complexes of **5** and **11** were collected to 1.7 and 1.9 Å resolution, respectively, and the structures were solved by molecular replacement using a known structure.[26d] Data for the complexes with **6**, **8**, **12**, and **14** were collected to 1.7, 1.8, 2.0, and 1.7 Å resolution, respectively, and the structures were solved similarly. There was one complex in the asymmetric unit in the structures of **5** and **6**, but there were two complexes in the asymmetric unit for **8**, **11**, **12**, and **14**. Having multiple copies of a protein−ligand complex in the asymmetric unit is useful because it affords independent snapshots of approximately isoenergetic structures, thereby providing a benchmark for interpreting the importance of any observed structural differences.

Comparing the structures of the complexes of **5**, **6**, and **8** with those of their corresponding cyclopropane-derived ligands **11**, **12**, and **14** reveals a number of similarities as well as some

(37) For example, see: (a) Radzicka, A.; Wolfenden, R. *Biochemistry* **1988**, *27*, 1664–1670. (b) Wimley, W. C.; Creamer, T. P.; White, S. H. *Biochemistry* **1996**, *35*, 5109–5124.
(38) Krug, R.; Hunter, W.; Grieger, R. *Nature* **1976**, *261*, 566–567.
(39) A number of structurally similar ligands are known to bind differently to the same protein. See: Boström, J.; Hogner, A.; Schmitt, S. *J. Med. Chem.* **2006**, *49*, 6716–6725.
(40) Complexes of the Grb2 SH2 domain with ligands **6**, **8**, **12**, and **14** were deposited into the Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank under entries 3IN8, 3IMD, 3IMJ, and 3IN7, respectively. Complexes with ligands **5** and **11** were deposited previously under entries 3C7I and 2HUW, respectively.

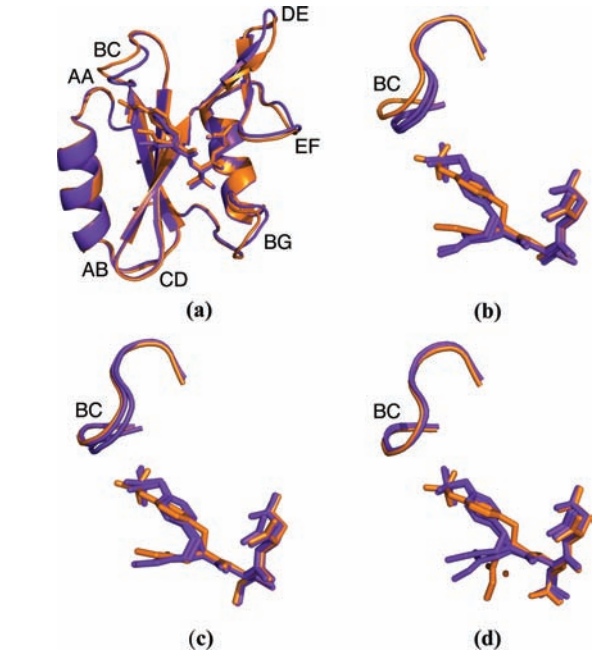


Figure 6. Overlay of Grb2 SH2 domain complexed with flexible and constrained ligands. Loops are labeled, and residues are represented as ribbons and the bound ligands as sticks. In images (b)−(d), only residues SerβB7 to AspβC1 of the domain, which includes the BC loop (GluBC1-ProBC4), are shown for all complexes in the asymmetric unit. (a) Overlay of complex of **5** (orange) and one complex in the asymmetric unit of **11** (purple), showing the complete domain. (b) Overlay of Grb2 SH2 in complex with **5** (orange) and **11** (purple). (c) Overlay of Grb2 SH2 in complex with **6** (orange) and **12** (purple). (d) Overlay of Grb2 SH2 in complex with **8** (orange) and **14** (purple) showing water molecules involved in hydrogen bond network between the *N*-terminal carbonyl oxygen atom, the amide moiety in the Gln side chain, and HisβD4.

differences (Figure 6). Flexible and constrained ligands bind to the domain in a β-turn-like conformation featuring an intramolecular hydrogen bond between the pY+1 carbonyl oxygen atom and the *C*-terminal amide. The alignment of the complex of **5** (orange) and one of the complexes in the asymmetric unit of **11** (purple) (Figure 6a) shows that the helical and the β-strand secondary structural elements of the domain in the two complexes are virtually identical; however, there are some notable differences in the loop regions, especially in the BC loop. A comparison of the structures of other complexes of flexible and constrained ligand pairs reveals similar features. For example, the backbone atoms of the Grb2 SH2 domain in its complexes with **5**, **6**, and **8** align with the backbone atoms in the corresponding complexes with **11**, **12**, and **14** with a root-mean-square deviation (rmsd) of 0.37−0.56 Å. The rmsd values for these alignments are comparable to those obtained from aligning the backbone atoms in each of the two complexes in the asymmetric unit for **8** (0.07 Å), **11** (0.29 Å), **12** (0.42 Å), and **14** (0.42 Å). Hence, *excepting some variations in the flexible loop regions, there are no significant conformational differences for the backbone atoms of the Grb2 SH2 domain in its complexes with flexible and constrained ligands*.

The aforementioned dissimilarities in the orientations of the backbone atoms of the residues in the BC loop (GluBC1-ProBC4) in complexes of flexible and constrained ligand pairs are apparent in the overlays in Figure 6b−d. These differences do not result from variations in the conformation of the backbone atoms in the BC loops in the complexes of the constrained ligands. Rather they arise from variable orientations of these

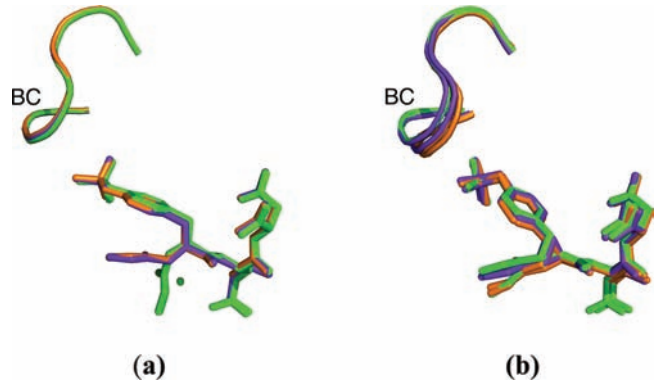


Figure 7. Overlay of residues SerβB7 to AspβC1, which includes the BC-loop (GluBC1-ProBC4), of the Grb2 SH2 domain complexed with flexible and constrained ligands showing water molecules involved in hydrogen bond network between the *N*-terminal carbonyl oxygen atom, the amide moiety in the Gln side chain, and HisβD4. Residues are represented as ribbons, and the bound ligands as sticks. (a) The set of flexible ligands **5** (orange), **6** (purple), and **8** (green). (b) The set of constrained ligands **11** (orange), **12** (purple), and **14** (green).

loops relative to the backbone of the domains and the oxygen atoms of the phosphate groups (*cf*. Figure 7b); the relative relationship between the BC loops and the phosphate groups of the flexible ligands are virtually identical (*cf*. Figure 7a). Inspection of Figure 6b reveals that the backbone atoms in the BC loops in both complexes in the asymmetric unit of **11** are packed closer to the phosphate moiety than in the complex with **5**. On the other hand, there is a notable difference in the relationship between the phosphate groups and the BC loops in the two complexes in the asymmetric unit of **12**, and the spatial relationship in one of these is similar to that found in the complex of its flexible counterpart **6** (Figure 6c). The orientations of the phosphate groups relative to the backbone atoms of the BC loops in the two complexes of **8** and in the two complexes of its constrained analog **14** are similar (Figure 6d). In the context of making structural comparisons, it is noteworthy that *variations in the dispositions of the backbone atoms in the BC loops relative to the phosphate groups of flexible ligands and their constrained counterparts are generally comparable to differences found in multiple copies of the same complex in some of the asymmetric units, especially those of **12***. Accordingly, it is difficult to assess the extent to which these differences in relative orientations are energetically significant.

Structural alignments of the pYVN analogs **5** and **11** (Figure 6b), the pYIN analogs **6** and **12** (Figure 6c), and the pYQN analogs **8** and **14** (Figure 6d) reveal that the pY+1 and pY+2 residues in each are virtually identical. Indeed, the positions of all atoms in these residues align with rmsd values of less than 0.24 Å, a value that is within the rmsd cutoff of 0.25 Å commonly used to define "identical" structures found by molecular dynamics calculations.[41] These superimpositions do reveal, however, significant differences in the positions of atoms in the flexible and constrained Ac-pY replacements. Namely, atoms in these replacements in the complexes of **5** and **11**, **6** and **12**, and **8** and **14** align with average rmsd values of 0.88, 0.77, and 1.08 Å, respectively, with the greatest variations being in the relative positions of the bridging phosphate oxygen atoms and the atoms of the *N*-acetyl moieties. Most notably, the

(41) See: *Macromodel*, version 9.0; Schrödinger, LLC: New York, NY, 2005.

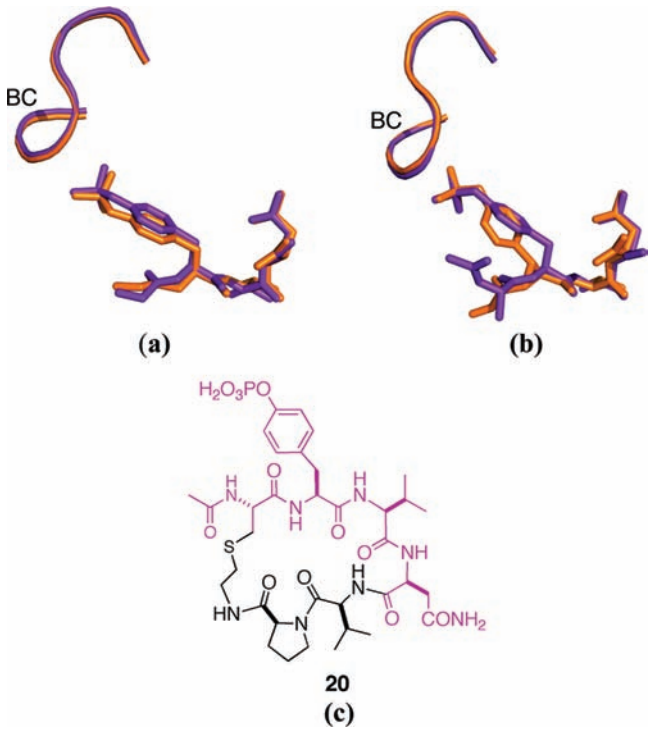


Figure 8. Complexes of ligands with the Grb2 SH2 domain showing relationship between the ligands and residues SerβB7-AspβC1, which includes the BC-loop (GluBC1-ProBC4), of the domain. (a) Overlay of structures of complexes of **5** (orange) with the linear nonapeptide **19** (purple),[26d] showing only the Ser-pTyr-Val-Asn segment of **19** for clarity. (b) Alignment of one complex in the asymmetric unit of **11** (orange) with the structure of the complex of macrocycle **20** (purple),[13c] showing only the portion of **20** that is purple for clarity. (c) Structure of **20**.

N-terminal amide of **8** is oriented so that the carbonyl oxygen atom can participate in a water-mediated hydrogen bond network with the side chain of the glutamine residue at the pY+1 site (Figure 6d). Because the corresponding carbon−carbon bond in **14** is incorporated in the cyclopropane ring, this conformation is not accessible to **14**.

A comparison of structural alignments for all complexes reveals that the set of flexible ligands **5**, **6**, and **8** (Figure 7a) bind and interact with the Grb2 SH2 domain with more consistency than does the set of their corresponding constrained ligands **11**, **12**, and **14** (Figure 7b). Excepting the orientation of the *N*-terminal amide group in **8**, all of the corresponding atoms in the flexible ligands align with rmsd values ≤0.12 Å, and the maximum displacement of the backbone atoms of the BC loops is 0.3 Å. Although the corresponding atoms in the pY+1 and pY+2 residues of the constrained ligands **11**, **12** and **14** also align closely (rmsd values ≤0.16 Å), superimposition of the atoms of the cyclopropane-derived Ac-pY replacement is not quite as good (rmsd values ≤0.26 Å) because of variable orientations of the *N*-acetyl group. The maximum displacement of the backbone atoms of the BC loops in these complexes is 1.8 Å, suggesting that the BC loop can readily adopt a number of orientations so the amino acid side chains can achieve optimal interactions with the phosphate group.

We then compared the structures of the complexes of the Grb2 SH2 domain with **5** and the linear nonapeptide Ala-Pro-Ser-pTyr-Val-Asn-Val-Gln-Asn (**19**)[26d] (Figure 8a) and of the domain with **11** and the macrocycle **20**[13c] (Figure 8b). The backbone atoms of the domain in the complexes of **5** and **19** align with a rmsd of 0.5 Å, whereas the corresponding atoms in the ligands **5** and **19** align with a rmsd of 0.3 Å. There are slight variations in the pairwise polar interactions and contact distances between the ligands and the domain in these two complexes; however, because these differences are similar to those observed in structures of other complexes of phosphotyrosine-containing ligands bound to the Grb2 SH2 domain,[13c,26] attaching significance to these dissimilarities is difficult. The backbone atoms of the domain in the two complexes of **11** and the complex of **20** align with a rmsd of 0.35 Å, and the corresponding atoms in **11** and **20** align with a rmsd of 0.6 Å. Despite these differences, the pairwise polar interactions and contact distances between the BC loop of the domain and the phosphate groups of **11** and **20** in these complexes are closely comparable. This analysis suggests that *the succinyl- and cyclopropane-derived pY replacements in **5** and **11** are not uniquely responsible for any significant structural variations in their respective complexes with the Grb2 SH2 domain.* The pY surrogates in **5** and **11** thus appear to be good mimics of the corresponding pY residues in the linear peptide **19** and the macrocycle **20**.

That there are discernible structural differences in complexes of corresponding pairs of flexible and constrained ligands is not surprising, as Boström has shown that like molecules often do not bind to and interact with a given target in a similar manner.[39] This, however, begs the question of whether there are any dissimilarities in the interactions between the domain and pairs of flexible and constrained ligands that may be correlated with the observed differences in binding enthalpies or entropies. To address this issue, the polar contacts, which may be categorized as direct or single water-mediated, between the domain and flexible and constrained ligands having the same residue at the pY+1 position were compared using contact diagrams such as those in Figure 9 (see Supporting Information for contact diagrams for **6**, **12**, **8**, and **14**). The only direct contacts considered in this analysis are those wherein the distance between an electronegative atom on the ligand and an electronegative atom on the protein is within 2.5−3.4 Å. The same criteria were applied to contacts that are mediated by a single water molecule; protein−ligand interactions that are mediated by more than one water molecule are not considered. Because of differences in the resolution of the crystallographic data, we made no attempt to further distinguish polar interactions based upon their measured contact distances or angles.

As highlighted in Figure 9a, there are seven direct contacts between the domain and the Ac-pY replacement that are conserved in all complexes of flexible ligands, whereas there are eight direct contacts involving the Ac-pY replacement that are conserved in all complexes of constrained ligands (Figure 9c). The contact between SerβC3 and the bridging phosphate oxygen atom that is conserved in the complexes of the constrained ligands is not found in any of the complexes of their flexible counterparts owing to the different orientation of this atom in these latter complexes. There are two direct contacts between the *N*-terminal carbonyl oxygen atom and the side chain of ArgαA2 that are conserved in the complexes of **5**, **6**, and all of the constrained ligands. On the other hand, the *N*-terminal carbonyl oxygen atom of **8** makes one contact with the side chain of ArgαA2 and a second contact with a water molecule involved in a hydrogen bond network with the Gln side chain of the ligand. There are two conserved, single water-mediated interactions involving W1 and W2, the domain, and the phosphate group in the flexible ligands **5**, **6**, and **8**. In one of these, W1 participates in interactions with the backbone N−H

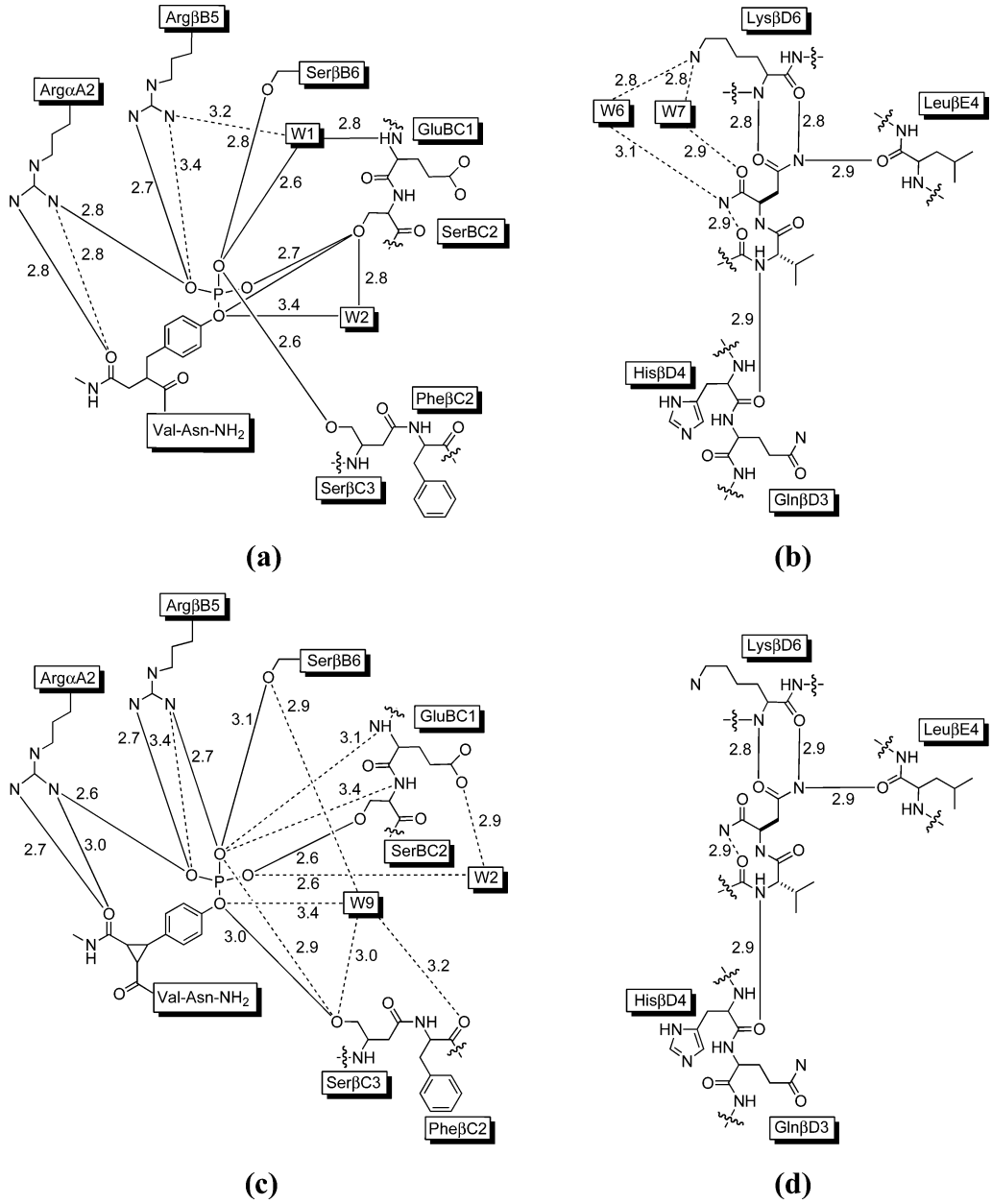


Figure 9. Polar interactions in the range of 2.5−3.4 Å in complexes of Grb2 SH2 domain with **5** and one complex in the asymmetric unit of **11**. All labile hydrogen atoms have been omitted for clarity except those on protein backbone nitrogen atoms. Only those water molecules that mediate a contact between the domain and the ligand are shown, and these are numbered so that water molecules that are conserved in at least two complexes have the same number. Solid lines in (a) and (b) indicate those polar contacts that are conserved for all complexes of flexible ligands, and solid lines in (c) and (d) indicate those polar contacts that are conserved for all complexes of constrained ligands. Dotted lines represent contacts that are not conserved within the set of flexible or constrained ligands. (a) Interactions between pY replacement of **5** and domain. (b) Interactions between VN of **5** and domain. (c) Interactions between pY replacement of **11** and domain. (d) Interactions between VN of **11** and domain.

of GluBC1 and one of the nonbridging phosphate oxygen atoms of the Ac-pY replacement, and in the other, W2 interacts with the side chain hydroxyl group of SerBC2 and the bridging phosphate oxygen atom of the ligand (*cf*. Figure 9a). There are no water-mediated, protein−ligand interactions in the Ac-pY binding pocket that are conserved in the set of constrained ligands.

There are four direct contacts between the domain and the pY+1−N subunit of the ligands that are conserved in all complexes (*cf*. Figure 9b and 9d). There are more single water-mediated contacts in the pY+1−N binding pocket in the complexes of the flexible ligands, but there are no water-mediated protein−ligand interactions in this binding pocket that are conserved in either set of complexes.

A summary of direct and single water-mediated contacts for all complexes is presented in Table 2. An inventory of the interactions for ligands having the same amino acid at the pY+1 site reveals that there are generally more direct contacts between the domain and the constrained member of a pair, whereas there are more single water-mediated contacts involving its flexible counterpart. For each pair of ligands, the larger number of *direct* contacts correlates qualitatively with the more favorable binding enthalpies and Gibbs energies that are observed, although there is no such correlation for the total number of polar contacts. A

Table 2. Direct and Single Water-Mediated Polar Protein-Ligand Contacts As Determined by X-Ray Crystallography[a]

Complex	Direct Contacts of Ac–pY	Single Water-Mediated Contacts of Ac–pY	Direct Contacts of pY+1–N	Single Water-Mediated Contacts of pY+1–N	Total Direct Contacts	Total Single Water-Mediated Contacts	Total Contacts
5 (fpYVN)	9	2	4	2	13	4	17
11a (cpYVN1)	12	2	4	0	16	2	18
11b (cpYVN2)	12	2	4	1	16	3	19
6 (fpYIN)	9	2	4	2	13	4	17
12a (cpYIN1)	9	1	4	1	13	2	15
12b (cpYIN2)	12	0	4	0	16	0	16
8a (fpYQN1)	8	3[b]	4	3[b]	12	6[b]	18
8b (fpYQN2)	8	3[b]	4	3[b]	12	6[b]	18
14a (cpYQN1)	9	1	4	1	13	2	15
14b (cpYQN2)	10	2	4	2	14	4	18

[a] See Supporting Information for contact diagrams of **6**, **8**, **12** and **14**. [b] One water molecule makes a contact with the backbone nitrogen atom of HisβD4 and two contacts with the ligand, one with the *N*-Ac moiety of the pY replacement and the other with the Gln side chain; each are counted independently.

closer examination of the summary shown in Table 2, however, underscores the difficulties associated with trying to correlate polar protein−ligand contacts with binding energetics because the *variations in the number of direct and water mediated contacts involving similar ligands may be comparable to differences found in multiple copies of a complex in an asymmetric unit* as exemplified by analyzing **12a** and **12b** (Table 2).

It is evident that the interfacial interactions involving water molecules vary for the flexible and constrained ligands of a given pair, but assessing the detailed role of water in the energetics of protein−ligand interactions is a daunting challenge.[42] Although such a discussion is beyond the scope of the present work, several points are worth brief mention. Fixing a water molecule at a protein−ligand interface is generally regarded as being entropically unfavorable.[43] Because there are more single-water mediated contacts in the complexes of the flexible ligands, it is somewhat perplexing that the binding entropies for forming these complexes are *more* favorable. Other contributions to binding entropy must overcome the entropic penalty associated with these bound water molecules. It has also been shown that the release of water molecules that are not optimally hydrogen bonded can have favorable enthalpic consequences.[11] This finding raises the interesting question of whether the more favorable enthalpies observed for forming complexes of the preorganized ligands, which have fewer water molecules bound at the protein−ligand interface, might arise from the release of such water molecules from the domain. Unfortunately, the Grb2 SH2 domain invariably crystallizes as a domain swapped dimer in the absence of ligand,[44] so the details of the hydration of the ligand binding site of the uncomplexed domain and the fate of these water molecules upon binding are unknown.

The van der Waals contacts in complexes of the Grb2 SH2 domain with the flexible and constrained ligands were then examined. Perhaps the most notable variation in the different complexes is the orientation and packing of the side chain of LysβD6, which may form a van der Waals contact either with the aromatic ring of the pY residue or with the *C*-terminal amide moiety of the ligand. These differences would seem to be approximately offsetting. There are a number of other dissimilarities in the van der Waals contacts in the structures of flexible and constrained ligands, but these generally do not appear significant relative to the variations that occur in different complexes in the same asymmetric unit. Based upon this analysis, *the van der Waals contacts in complexes of flexible and constrained ligands seem comparable and cannot be easily correlated with any differences in binding enthalpies*.

Crystallographic b-factors have been used to probe changes in protein flexibility in protein−ligand complexes.[45] This practice is, however, subject to numerous caveats because these thermal parameters are influenced by crystal packing, resolution, temperature of data acquisition, methods used to solve the structure, and a number of other effects that are not directly related to changes in protein flexibility.[46] Atomic b-factors also reflect a static, solid-state picture of nonbonded interactions that are much more dynamic in solution owing to greater thermal fluctuations in atomic positions. These issues notwithstanding, comparisons of b-factors can provide a sense of the relative magnitudes of motion in the solid state, *if* all of the b-factors in the data sets are adjusted by a constant value for the data set so the lowest value in each set corresponds to the lowest value in the data set having the highest overall displacement.[46a]

Applying this protocol to the crystallographic data sets reveals that the average adjusted b-factors for all atoms in residues 59−151 of the Grb2 SH2 domain in its complexes with the flexible ligands **5**, **6**, and **8** are 3.7−6.6 Å^2 less than those in the corresponding set of complexes of the constrained ligands **11**, **12**, and **14** (Figure 10). Although there are differences in the magnitudes of these b-factors throughout the complexes, the most notable variations are in the loop regions, especially in the four residues in the BC loop, which is involved in

(42) For some leading, general references, see: (a) Ladbury, John E. *Chem. Biol.* **1996**, *3*, 973–980. (b) Levy, Y.; Onuchic, J. N. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 389–415. (c) Li, Z.; Lazaridis, T. *Phys. Chem. Chem. Phys.* **2007**, *9*, 573−581. See also ref 11d.

(43) The maximal value for transferring a water molecule from bulk water to a protein binding site has been suggested to be ~2 kcal mol^{-1}. See: Dunitz, J. D. *Science* **1994**, *264*, 670.

(44) Benfield, A. P.; Whiddon, B. B.; Clements, J. H.; Martin, S. F. *Arch. Biochem. Biophys.* **2007**, *462*, 47–53.

(45) For a recent discussion of using b-factors to compare protein flexibility in protein−ligand complexes, see: Yang, C.-Y.; Wang, R.; Wang, S. *J. Med. Chem.* **2005**, *48*, 5648–5650.

(46) For leading references, see: (a) Ringe, D.; Petsko, G. A. *Methods Enzymol.* **1986**, *131*, 389–433. (b) Bhalla, J.; Storchan, G. B.; MacCarthy, C. M.; Uversky, V. N.; Tcherkasskay, O. *Mol. Cell. Proteomics* **2006**, *5*, 1212–1223.

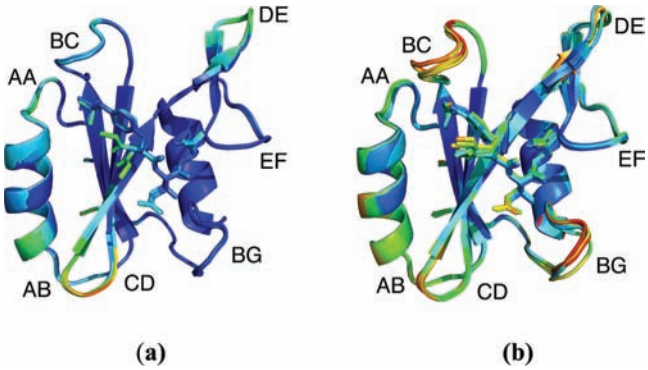


Figure 10. Adjusted atomic b-factors in complexes of the Grb2 SH2 domain (residues 59−151) with flexible and constrained ligands. Backbone atoms of residues are color-coded from blue to green to red for atoms having values of 20, 30, and 40 Å^2. (a) Complexes of flexible ligands **5**, **6**, and **8**. (b) Complexes of constrained ligands **11**, **12**, and **14**.

phosphate binding, and in the seven residues in the BG loop (SerBG1-GlnBG7), which makes single water-mediated contacts with the pY+1 residue. Indeed, the average b-factor for all backbone atoms in the BC loops in the complexes of **5**, **6**, and **8** is approximately the same as the domain average, whereas the average b-factor for the same backbone atoms in the complexes of **11**, **12**, and **14** is nearly one standard deviation above the domain average.

It is known, but not widely appreciated, that forming a protein−ligand complex may result in either an overall decrease or increase in protein flexibility and dynamics.[22f,47] Such changes in flexibility will be accompanied by changes in nonbonded interactions and order throughout the complex that will eventuate in enthalpic and entropic consequences. It is difficult, however, to evaluate how such changes in dynamics will affect the relative magnitudes of the compensating enthalpies and entropies of binding. For example, one might infer from the preceding analysis of b-factors that the thermal motions, at least in the solid state, in the complexes of the constrained ligands are greater than those in the complexes of their more flexible counterparts. Considering the contributions of the protein to the thermodynamic binding parameters, one might then predict that the entropies, especially the configurational entropies, for forming complexes of the constrained ligands would be *more* favorable than those for their flexible analogs, a prediction that is opposite the experimental observation.

Crystal packing in the different complexes of the Grb2 SH2 domain with the flexible ligands **5**, **6**, and **8** is virtually identical, whereas the crystal packing in the complexes of the constrained ligands **11**, **12**, and **14** varies from one to another and from their flexible counterparts. Despite these differences, however, variations in crystal packing do not seem to correlate with structural changes. For example, crystal packing in one complex of **11** (cpYVN1) and in one complex of **12** (cpYIN1) in their respective asymmetric units is similar, yet the relative orientations of the BC loops in each of these complexes differ markedly. There are also significant dissimilarities in the crystal packing for each of the two complexes in the asymmetric unit of **14** (cpYQN1 and cpYQN2), but the structures of the two complexes are comparable. Moreover, the crystal packing in these two complexes is different from either of the two complexes in the asymmetric unit of **8** (fpYQN1 and fpYQN2), even though the structures of all of these complexes, with the notable exception of the orientation of the *N*-terminal acetyl group in **8**, are similar. Based upon these observations, *it is not possible to correlate differences in crystal packing with variations in contact distances, b-factors, or other structural details*.

Summary and Conclusions

There is ample evidence that ligand preorganization may provide compounds having improved target selectivity,[48] bioavailability,[49] and higher binding affinity.[13,14] The present investigations also demonstrate that constraining ligands in their biologically active conformations can lead to more potent analogs. The higher affinities observed for the conformationally constrained ligands **11**−**16** arise, however, because they benefit from more favorable binding enthalpies than their respective flexible controls **5**−**10**. Indeed, binding entropies for these preorganized ligands are uniformly disfavored relative to their flexible counterparts. That this surprising finding is not simply a consequence of the peptide replacements used in these studies is evident from our previous work with Src SH2 domain binding ligands where we found that compounds having cyclopropane-derived pY replacements did exhibit more favorable binding entropies than their flexible succinate analogs. Moreover, Spaller has recently found that preorganization of a linear peptide by macrocyclization can lead to increased affinity for binding to a PDZ3 domain because a relative enthalpic advantage dominates a corresponding entropic penalty.[20b] Based upon several independent studies, *it is thus evident that the widely held assumption that ligand preorganization should be accompanied by an entropic advantage is not always true*.[50]

A search for the origin of this unexpected discovery has, however, not yet led to concrete answers. We found that variations in binding enthalpies and entropies for the flexible and constrained ligands did not appear to arise from differences in either proton exchange or desolvation phenomena. The Gibbs energy relationship accounts for the observed enthalpy−entropy compensation that accompanies preorganization and changes in the nature of the pY+1 residue. Crystallographic studies of complexes of flexible and constrained

(47) For some leading references, see: (a) Anderegg, R. J.; Wagner, D. S. *J. Am. Chem. Soc.* **1995**, *117*, 1374–1377. (b) Kay, L. E.; Muhandiram, D. R.; Wolf, G.; Shoelson, S. E.; Forman-Kay, J. D. *Nat. Struct. Biol.* **1998**, *5*, 156–163. (c) Engen, J. R.; Gmeiner, W. H.; Smithgall, T. E.; Smith, D. L. *Biochemistry* **1999**, *38*, 8926–8935. (d) Wang, F.; Miles, R. W.; Kicska, G.; Nieves, E.; Schramm, V. L.; Angelettin, R. H. *Protein Sci.* **2000**, *9*, 1660–1668. (e) Stone, M. J. *Acc. Chem. Res.* **2001**, *34*, 379–388. (f) Wang, F.; Shi, W.; Nieves, E.; Angeletti, R. H.; Schramm, V. L.; Grubmeyer, C. *Biochemistry* **2001**, *40*, 8043–8054. (g) Ferreon, J. C.; Hilser, V. J. *Protein Sci.* **2003**, *12*, 982–996. (h) de Mol, N. J.; Catalina, M. I.; Fischer, M. J. E.; Broutin, I.; Maier, C. S.; Heck, A. J. R. *Biochim. Biophys. Acta* **2004**, *1700*, 53–64. (i) MacRaild, C.; Daranas, H.; Bronowska, A.; Homans, S. W. *J. Mol. Biol.* **2007**, *368*, 822–832.

(48) For a review, see: (a) Tyndall, J. D. A.; Pfeiffer, B.; Abbenante, G.; Fairlie, D. P. *Chem. Rev.* **2005**, *105*, 793–826. For some selected examples, see: (b) Kawai, M.; Horikawa, Y.; Ishihara, T.; Shimamoto, K.; Ohfune, Y. *Eur. J. Pharmacol.* **1992**, *211*, 195–202. (c) Liao, S.; Alfaro-Lopez, J.; Shenderovich, M. D.; Hosohata, K.; Lin, J.; Li, X.; Stropova, D.; Davis, P.; Jernigan, K. A.; Porreca, F.; Yamamura, H. I.; Hruby, V. J. *J. Med. Chem.* **1998**, *41*, 4767–4776. (d) Suich, D. J.; Mousa, S. A.; Singh, G.; Liapakis, G.; Reisine, T.; DeGrado, W. F. *Bioorg. Med. Chem.* **2000**, *8*, 2229–2241. (e) Ying, J.; Gu, X.; Cai, M.; Dedek, M.; Vagner, J.; Trivedi, D. B.; Hruby, V. J. *J. Med. Chem.* **2006**, *49*, 6888–6896.

(49) Veber, D. F.; Johnson, S. R.; Cheng, H.-Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. *J. Med. Chem.* **2002**, *45*, 2615–2623.

(50) For a recent computational study of the energetic consequences associated with the flexibility of hosts and guests, see: Moghaddam, S.; Inoue, Y.; Gilson, M. K. *J. Am. Chem. Soc.* **2009**, *109*, 4012–4021.

ligands having the same pY+1 residue reveal widespread similarities, and many of the observed differences are comparable to those seen in multiple copies of the same complex in some asymmetric units. The more notable dissimilarities in the complexes of a flexible/constrained pair are found in the number and type of polar interactions between the BC loop and the AcpY replacements in the ligands. Specifically, the constrained ligand of a given pair makes more direct contacts with the domain, so there is a qualitative correlation between the total number of direct protein−ligand contacts and the relative binding enthalpies and Gibbs energies for each ligand pair. Conversely, the more flexible ligand of a pair makes more single water-mediated contacts with the domain. Inasmuch as the entropies for forming complexes of the flexible ligands are more favorable, this finding seems incompatible with the prevailing view that there is an entropic cost associated with fixing water molecules at a protein−ligand interface. There are no significant differences in the van der Waals contacts for a given flexible/constrained ligand pair. An analysis of crystallographic b-factors suggests that thermal motions in complexes of the constrained ligands are generally greater than those in the corresponding complexes of their flexible controls. If these motions in the solid state actually reflect more disorder, and hence more favorable configurational entropies, in the complexes of constrained ligands, one might anticipate that the binding entropies for their formation would be greater than those for their flexible counterparts; however, this prediction is inconsistent with our results.

These systematic and extensive studies clearly demonstrate our lack of understanding of energetics in protein−ligand interactions, even in biological systems that are well-characterized by thermodynamic and structural studies. They also underscore the difficulty of correlating the number or type of protein−ligand contacts or the number of water molecules at the protein−ligand interface with specific contributions to binding enthalpies and entropies. One shortcoming of many studies of protein−ligand interactions is that they focus upon the direct interactions between the protein and the ligand, perhaps including regions of the protein proximal to the ligand binding site; the effects of ligand binding upon protein dynamics and structure distal to the binding site are rarely considered. Indeed, there is a profound lack of experimental data that correlates the energetic effects associated with the net conformational and dynamic changes that occur in the protein and the ligand upon complexation.

We believe that future models of protein−ligand interactions must include explicit consideration of the enthalpic and entropic contributions arising from changes in nonbonded interactions and order that occur upon complex formation and how these changes vary as a function of ligand structure. Ligands may bind to proteins with high affinity in conformations that are energetically higher than their global minima in solution, so the relationship between the conformational strain energy of a ligand in its bound conformation and its affinity must be better understood. Moreover, the detailed interactions of the protein, ligand, and the protein−ligand complex with water molecules as well as the effects of solvent reorganization must be incorporated in the analysis. Finally, to optimize protein−ligand interactions, means must be found to modify the structure of a ligand in a way that minimizes enthalpy−entropy compensation. We are beginning to address these and other difficult questions in ongoing studies of energetics and structure in protein−ligand interactions that will include the use of NMR and computational methods. These results will be reported in due course.

Acknowledgment. We thank the National Institutes of Health (GM 84965), the National Science Foundation (CHE 0750329), The Robert A. Welch Foundation (F-652), the Norman Hackerman Advanced Research Program, The Texas Institute for Drug and Diagnostic Development through a Welch Foundation Grant (#H-F-0032) for their generous support of this research. J.E.D. thanks the ACS Division of Medicinal Chemistry and Pfizer Inc. as well as the Dorothy B. Banks Foundation for predoctoral fellowships. We are also thankful to Dr. Stuart Black (Schering-Plough Research Institute) for the Grb2 SH2 expression vector and helpful discussions and to Dr. James Davidson for helpful discussions. We are also grateful to Brahm Carlson, David Cramer, Greg Frayne, Monica Lotz, Nicole Martin, James Myslinski, Naveed Nosrati, Nicole Pulliam, Adithya Rangarajan, and Benjamin Whiddon for their help in numerous aspects of these investigations.

Supporting Information Available: Complete ref 14c, experimental procedures and characterization data for all new compounds, methods and materials for ITC and X-ray crystallographic experiments, plots of ITC data, and contact diagrams for complexes of **6**, **8**, **12**, and **14**. This material is available free of charge via the Internet at http://pubs.acs.org.

JA904698Q