

# Computational Binding Studies of Human pp60<sup>c-src</sup> SH2 Domain with a Series of Nonpeptide, Phosphophenyl-Containing Ligands

Daniel J. Price and William L. Jorgensen\*

Department of Chemistry, Yale University, New Haven, CT 06500-8107, USA

Received 18 May 2000; accepted 30 June 2000

**Abstract**—Monte Carlo/free energy perturbation (MC/FEP) simulations were performed on a series of nonpeptide ligands of the human pp60<sup>c-src</sup> SH2 domain in order to calculate relative free energies of binding for each compound and to understand the structural requirements for high affinity binding. The amido compound, exhibiting the highest experimental affinity, takes advantage of an interaction with a previously unobserved structural water. © 2000 Elsevier Science Ltd. All rights reserved.

Activation of pp60<sup>c-src</sup> (Src) tyrosine kinase is dependent on membrane localization and recognition, via its SH2 domain, of a phosphorylated tyrosine (Y\*) located on the cytoplasmic side of a stimulated surface receptor. Considering that various mitogenic signals have been shown to be transduced by Src-mediated pathways (for review, see ref 1) and that elevated levels of Src have been implicated in numerous cancers, particularly breast and colon carcinomas,<sup>2,3</sup> there has been considerable interest in developing synthetic Src SH2 ligands as anti-proliferative agents. The complexity of this task can be appreciated when results from knockout studies in murine models are considered. Instead of observing the expected cell-lethal phenotype in *src*<sup>-/-</sup> mice, the mice suffered predominantly from osteopetrosis.<sup>4</sup> This and subsequent studies showing changes in the morphology and bone-resorptive properties of osteoclasts<sup>5,6</sup> have identified Src as a candidate target in the treatment of osteoporosis. The knockout studies emphasize that the redundancies in these signaling pathways make predicting the effects of targeting Src alone difficult. A set of high affinity ligands that modulate Src and related proteins' activities may help unravel the complex circuitry involved in intracellular signaling, and may provide novel therapeutic options for several diseases.

Src's SH2 domain negotiates Y\* recognition in a sequence-specific manner, binding the sequence, Y\*EEIE, with submicromolar affinity.<sup>7</sup> Structural studies of SH2-peptide complexes have characterized the SH2 binding site as containing an electropositive Y\* binding pocket, a hydrophobic cleft for binding the  $\beta$ -branched amino

acid 3-residues C-terminal to the Y\*, and a glutamic acid binding region between them.<sup>7-9</sup> A team at Parke-Davis has reported a set of ligands of the form 4-([4-(cyclohexylmethoxy) benzyl]amino)carbonyl)phenyl phosphate where functionalization at the 3-position of the central ring targets side-chain interactions in the glutamic acid binding site.<sup>10</sup> In the present study, MC/FEP simulations were used to examine several previously synthesized and several novel ligands involving functionalization at the 3-position (Table 1).

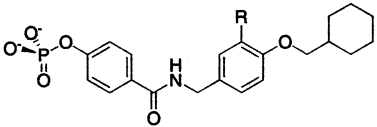
## Computational Methods

### Geometry of protein–ligand complexes and unbound ligands

Protein coordinates were taken from an X-ray crystal structure (PDB entry: 1SHD) of the SH2 domain bound to a consensus pentamer (Y\*EEIE). Coordinates for four residues, unresolved in the crystal structure, were taken from an average NMR structure of the same complex (PDB entry: 1HCS). The ligand with the highest experimental affinity, **1**, was modeled into the binding site based on the coordinates of the consensus peptide and on qualitative structural comments made about an unavailable crystal structure of a similar compound.<sup>10</sup> The complex was then minimized in a distance-dependent dielectric ( $\epsilon=r$ ) with a conjugate gradient algorithm to alleviate poor steric contacts.

All bonds not directly involved in the perturbation and all internal degrees of freedom of the entire protein backbone and residues 18 Å or farther from the center of the ligand were held rigid during all subsequent

\*Corresponding author. Fax: +1-203-432-6299; e-mail: bill@adrik.chem.yale.edu

**Table 1.** Available experimental IC<sub>50</sub> data


Compound	R	IC <sub>50</sub> (μM)
1	-CONH <sub>2</sub>	6.5 <sup>a</sup>
2	-CONH(CH <sub>3</sub> )	30 <sup>a</sup>
3	-COCH <sub>3</sub>	69 <sup>b</sup>
4	-NH <sub>2</sub>	
5	-CH <sub>3</sub>	
6	-Cl	
7	-H	>100 <sup>a</sup>

<sup>a</sup>Ref 10.<sup>b</sup>Lunney, E. A. Personal communication.

simulations. A 20-Å sphere of pre-equilibrated, rigid, TIP4P<sup>11</sup> water was centered on the ligand in the binding site. A weak half-harmonic potential (force constant of 1.5 kcal/mol-Å<sup>2</sup>) was applied to any water that leaves the sphere. Charged protein residues beyond the 18-Å limit were neutralized, changing the net charge of the complex from +5 to +3.

For the unbound ligands, conformational searching using GB/SA<sup>12</sup> as a continuum solvation model was performed. The lowest energy structure was used as a starting point for the MC/FEP simulations. As the predicted ligand structure was considerably more extended when unbound, a sphere of water with a 22-Å radius was necessary to ensure complete solvation.

### Simulation details

All simulations were carried out at 25 °C. Monte Carlo moves were done on a per-residue basis,<sup>13</sup> but otherwise were done in the standard manner.<sup>14</sup> Specifically, ligands were divided into three artificial, uncharged residues to improve sampling. A residue-based cutoff of 10 Å was used.<sup>13</sup> Water-only equilibration was performed for 10 million (M) configurations where waters close to the ligand were preferentially sampled.<sup>15</sup> At least 10 M configurations of equilibration followed, where all variable degrees of freedom were sampled. Averaging periods are noted below.

FEP protocols for MC simulation have been previously discussed in detail.<sup>16</sup> Briefly, the group of interest is mutated from one functionality to another in a series of steps, or windows. All FEPs here involved 10 windows of double-wide sampling, where each window was equilibrated as described above. The total change in free energy of the mutation is the sum of the results of each step. The difference between the ΔG calculated for the mutation in the binding site and the mutation when the ligand is free in solution equals the relative free energy of binding (ΔΔG<sub>b</sub>).<sup>16</sup>

All parameters are taken from the OPLS-AA force field.<sup>17</sup> All FEPs were performed with MCPRO, version 1.6 or greater.<sup>18</sup> Torsional scans and gas-phase mini-

mizations were done with BOSS, version 4.0 or greater.<sup>19</sup>

## Results and Discussion

### FEP results

The computed ΔΔG<sub>b</sub> for each perturbation is shown in Table 2. Convergence of the ΔG's depended highly on the protocol used for disappearing atoms and will be discussed elsewhere. Reported uncertainties are calculated based on normal batch statistics<sup>20</sup> where each batch consisted of 0.25 M configurations. Due to low-frequency fluctuations in ΔG, low uncertainties did not always indicate strong convergence. Consequently, convergence was monitored on a case-by-case basis and averaging periods were therefore variable, ranging from 4 to 10 M configurations for bound simulations and 20 to 40 M configurations for unbound simulations. Closure of thermodynamic cycles is probably a better indicator of convergence; ΔΔG's of 0.2 kcal/mol for 1→2→1 and 1.1 kcal/mol for 1→4→6→7→5→1 are reasonable.

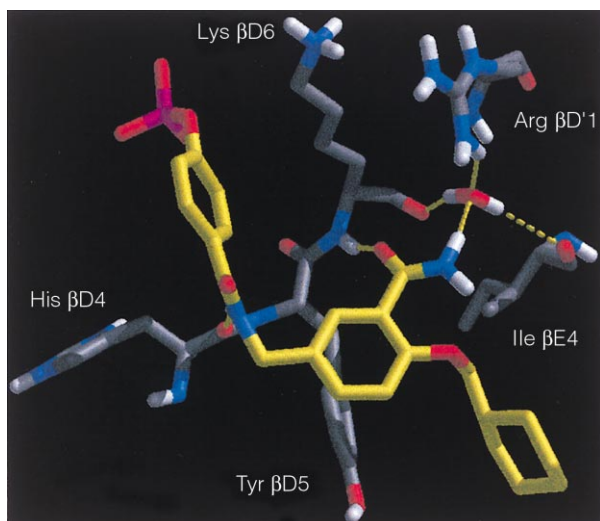
The computations predict similar binding affinities for 1 and 2, while 2 is observed to be the weaker binder by 0.9 kcal/mol. Solution- and gas-phase FEPs for the model system, acetamide to *N*-methyl acetamide, predict acetamide as too hydrophilic by ca. 1.7 kcal/mol.<sup>21</sup> These results indicate that the disparity with experiment in this mutation owes to 1 being too well solvated in the unbound leg. As these compounds are bound at the surface of the protein and have some water contacts, it is unclear how much of the apparent heightened hydrophilicity of 1 is also reflected in the bound mutation, thus canceling some of the error in the unbound leg. Regardless, the hydration studies put a bound on how much the relative binding energy of 2 may be corrected with improved parameters.<sup>21</sup>

The binding affinity of the unsubstituted compound, 7, was beyond the sensitivity of the experimental assay, consequently its relative free energy of binding is shown here as a lower limit. The calculated relative free energies

**Table 2.** Computed versus experimental relative free energies of binding

Perturbation	Computed ΔΔG <sub>b</sub> (kcal/mol)	Exptl. ΔΔG <sub>b</sub> (kcal/mol) <sup>a</sup>
1→2	0.0±0.2	0.9
2→1	-0.2±0.3	-0.9
1→3	0.4±0.3	1.4
1→4	3.5±0.3	
1→5	2.8±0.3	
4→6	-2.2±0.1	
5→7	1.8±0.1	
6→7	2.17±0.09	
1→5→7	4.6±0.3	>1.6
1→4→6→7	3.5±0.3	>1.6

<sup>a</sup>Experimental values calculated assuming that the ratio of IC<sub>50</sub> values equals the ratio of dissociation constants. Error bars for the IC<sub>50</sub> values, and consequently the experimental ΔΔG<sub>b</sub>'s, are unavailable.



**Figure 1.** Compound **1** bound to residues in and near the glutamate-binding site. Ligand carbons are yellow; aliphatic hydrogens have been removed for clarity. Graphics prepared with MidasPlus.<sup>22</sup>

of binding agree well with this result regardless of the pathway used to determine it (i.e., the **1**→**7** mutation was not explicitly performed, but can be calculated from the sum of several transformations, **1**→**5**→**7** for instance). Overall, the predicted order of binding is correct, though the selectivity for **1**, **2**, and **3** is compressed. The amino and methyl derivatives are predicted to be poor binders, whereas the chloro analogue is intermediate between **3** and **7**.

### Structural details of the glutamic acid binding site

The exocyclic group of the best binder, **1**, makes a hydrogen bond to a backbone amide hydrogen of LysβD6 and to a water molecule which is also bound to LysβD6, ArgβD'1, and IleβE4, as shown in Figure 1. Up to three waters had been previously shown to mediate glutamate binding in structures of complexes with peptides,<sup>8</sup> but no structural waters had been noted in the comments on the crystal structure of the dimethylphenylmethoxy analogue of this series.<sup>10</sup> This water molecule can be found in every snapshot (recorded every 0.25 M configurations) of every simulation.

Methylation of the amide, **2**, in the *trans* position sacrifices the hydrogen bond to the bridging water molecule. As the *cis* amide geometry would maintain this interaction, a gas-phase torsional energy scan using OPLS-AA (data not shown) of the model compound, *o*-methoxy, *N*-methyl benzamide, was used to probe whether the methylamide might bind the protein in a *cis* conformation. This rotation would incur a 7.5 kcal/mol internal energy penalty that includes loss of favorable electrostatic interactions between the *trans* NH and ether oxygen. Thus, the *cis* conformation is highly unlikely.

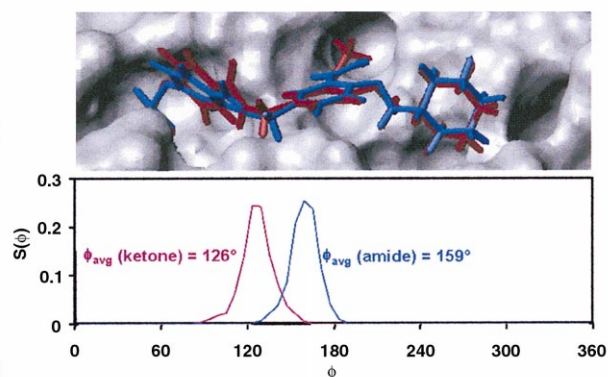
The keto compound, **3**, also loses the hydrogen bond to the bridging water molecule. In addition, gas-phase interaction studies (bimolecular minimizations using OPLS-AA) on model systems, benzamide and aceto-

phenone, revealed that the aryl ketone is a weaker hydrogen-bond acceptor at the carbonyl oxygen than the aryl amide. Indeed, benzamide's hydrogen bond to water was 0.9 kcal/mol stronger than acetophenone's, and 1.2 kcal/mol stronger when hydrogen bonding with *N*-methyl acetamide instead of water. It may be interesting to note that unfavorable interactions between the methyl group and the *ortho*-ether substituent cause the ketone to be forced farther out of the ring plane, when bound (Fig. 2) and unbound. Gas-phase dihedral scans of acetophenone and benzamide predict that there is a ca. 1 kcal/mol penalty for rotating the ketone out of the plane of the ring by the additional 30°. However, since this more twisted form occurs for both the bound and unbound states, it does not affect the binding affinity.

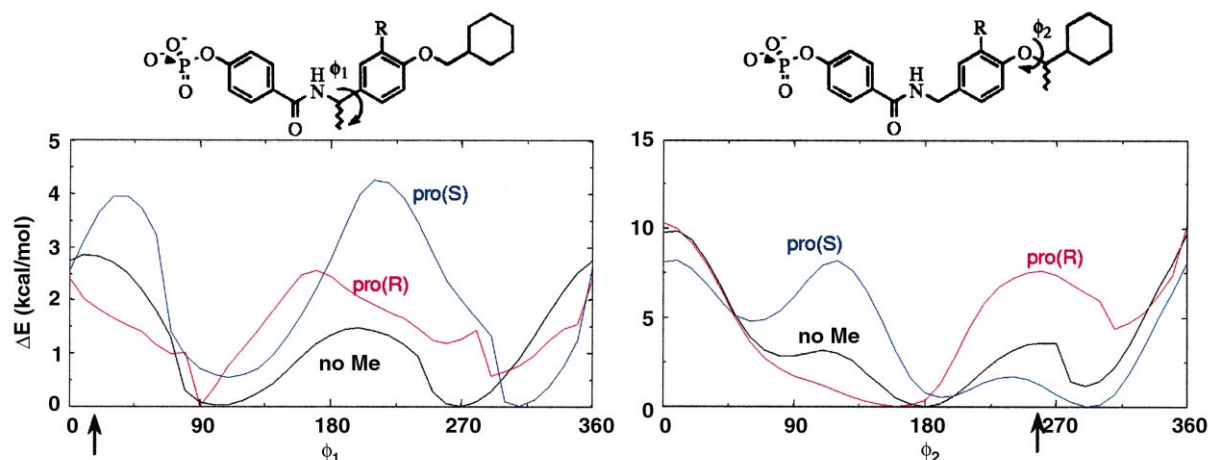
Compounds **5**–**7** are unable to hydrogen bond to either the backbone of Lys βD6 or to the bound water molecule. In addition, they prevent the solvation of the amide hydrogen (HN) of Lys βD6, thus increasing the desolvation penalty for the protein. Compound **4** seems to alleviate potentially severe electrostatic repulsion between the amino hydrogen and the HN of Lys βD6 by either rotation of the amine or translation of the entire ligand further towards the bound water molecule. Rotation of the amine allows the HN of Lys βD6 to hydrogen bond to the amino nitrogen and directs one amino hydrogen towards solvent. The other hydrogen points towards the side chain of Ile βE4, making its solvent contact nonlinear. Translation of the ligand entails solvent H-bonds to the amino nitrogen and one hydrogen, but no favorable contacts within 3 Å for either the second hydrogen or the HN of Lys βD6.

### Ligand binding geometry

Examination of torsional energy scans reveals that several dihedrals are not in their solution-phase minima when the ligand is bound. Stereospecific methylation at the methylenes adjacent to two of these dihedrals can affect their solution preferences, as the torsional scans indicate in Figure 3. Lunney and co-workers noted that a racemic mixture of **1** methylated at the methylene adjacent to  $\phi_1$  improves binding.<sup>10</sup> The analysis here predicts that either enantiomer may improve binding by virtue of lowering the relative energy of the conformation



**Figure 2.** (top) Overlay of compounds **1** (blue) and **3** (red), showing the twisting of the exocyclic methyl ketone out of the ring plane. (bottom) Dihedral distributions for the aryl carbonyl torsion (CCCO) in **1** and **3**. Molecular graphics prepared with MOLMOL.<sup>23</sup>



**Figure 3.** (left) Results of unbound, semi-constrained torsional scans in a distance-dependent dielectric from rotation around  $\phi_1$  (refer to structure). (right) Results of identical calculations for  $\phi_2$ . Arrows indicate dihedral value when bound. Two arrows indicate distribution is bimodal.

required for binding. The pro(R) methyl group faces solvent and would not disrupt binding, while the pro(S) methyl group fits a small pocket near Tyr  $\beta$ D5.

Pro(S) methylation adjacent to  $\phi_2$  could also aid in preorganization of the ligand. Likewise, substitution of the ether oxygen with a sulfur would alter the adjacent C=C–O–C torsional minimum from planar ( $0^\circ$ ) to the binding geometry of  $90^\circ$ . However, modeling of these modifications predicts steric clashes with the binding site in the absence of any backbone adjustments.

### Acknowledgements

Gratitude is expressed to the National Institute of Health (GM32136) for support of this research.

### References

1. Brown, M. T.; Cooper, J. A. *Biochim. Biophys. Acta* **1996**, *1287*, 121.
2. Mao, W.; Irby, R.; Coppola, D.; Fu, L.; Wloch, M.; Turner, J.; Yu, H.; Garcia, R.; Jove, R.; Yeatman, T. *J. Oncogene* **1997**, *15*, 3083.
3. Verbeek, B. S.; Vroom, T. M.; Adriaansen-Slot, S. S.; Ottenhoff-Kalf, A. E.; Geertzema, J. G. N.; Hennipman, A.; Rijksen, G. *J. Pathol.* **1996**, *180*, 383.
4. Soriano, P.; Montgomery, C.; Geske, R.; Bradley, A. *Cell* **1991**, *64*, 693.
5. Boyce, B. F.; Yoneda, T.; Lowe, C.; Soriano, P.; Mundy, G. R. *J. Clin. Invest.* **1992**, *90*, 1622.
6. Lowe, C.; Yoneda, T.; Boyce, B. F.; Chen, H.; Mundy, G. R.; Soriano, P. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 4485.
7. Gilmer, T.; Rodriguez, M.; Jordan, S.; Crosby, R.; Allgood, K.; Green, M.; Kimery, M.; Wagner, C.; Kinder, D.; Charifson, P.; Hassell, A. M.; Willard, D.; Luther, M.; Rusnak, D.; Sternback, D. D.; Mehrotra, M.; Peel, M.; Shampine, L.; Davis, R.; Robbins, J.; Patel, I. R.; Kassel, D.; Burkhart, W.; Moyer, M.; Bradshaw, T.; Berman, J. *J. Biol. Chem.* **1994**, *269*, 31711.
8. Waksman, G.; Shoelson, S. E.; Pant, N.; Cowburn, D.; Kuriyan, J. *Cell* **1993**, *72*, 779.
9. Xu, R. X.; Word, J. M.; Davis, D. G.; Rink, M. J.; Willard, D. H. J.; Gampe, R. T. *J. Biochemistry* **1995**, *34*, 107.
10. Lunney, E. A.; Para, K. S.; Rubin, J. R.; Humblet, C.; Fergus, J. H.; Marks, J. S.; Sawyer, T. K. *J. Am. Chem. Soc.* **1997**, *119*, 12471.
11. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D. *J. Chem. Phys.* **1983**, *79*, 926.
12. Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127.
13. Essex, J. W.; Severence, D. L.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **1997**, *101*, 9663.
14. Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford Science Publications: Oxford, 1987.
15. Jorgensen, W. L. *J. Phys. Chem.* **1983**, *87*, 5304.
16. Jorgensen, W. L. *Free Energy Changes in Solution*. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Ed.; John Wiley & Sons: Chichester, 1998.
17. Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225. Atomic parameters for phosphate, methylene of *N*-benzyl amide, and neutralized Lys and Arg are awaiting publication.
18. Jorgensen, W. L. *MCPRO*, Version 1.60, 1.65, Yale University: New Haven, 1999.
19. Jorgensen, W. L. *BOSS*, Version 4.0, 4.1, 4.2, Yale University: New Haven, 1999.
20. Jorgensen, W. L. *Monte Carlo Simulations for Liquids*. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Ed.; John Wiley & Sons: Chichester, 1998.
21. Pearlman, S.; Jorgensen, W. L. unpublished results.
22. Ferrin, T. E.; Huang, C. C.; Jarvis, L. E.; Langridge, R. *J. Mol. Graphics* **1988**, *6*, 13–27, 36–37.
23. Koradi, R.; Billeter, M.; Wüthrich, K. *J. Mol. Graphics* **1996**, *14*, 51.

