Pergamon

Bioorganic & Medicinal Chemistry Letters 8 (1998) 1189–1194

BIOORGANIC & MEDICINAL CHEMISTRY LETTERS

THE FORMATION OF A COVALENT COMPLEX BETWEEN A DIPEPTIDE LIGAND AND THE SRC SH2 DOMAIN

Krystal J. Alligood, Paul S. Charifson, Renae Crosby, Thomas G. Consler, Paul L Feldman, Robert T. Gampe Jr., Tona M. Gilmer, Steven R. Jordan, Mark W. Milstead, Christopher Mohr, Michael R. Peel,* Warren Rocque, Marc Rodriguez, David W. Rusnak, Lisa M. Shewchuk, and Daniel D. Sternbach.

Glaxo Wellcome, Inc., Five Moore Drive, Research Triangle Park, North Carolina 27709, U.S.A.

Received 24 November 1997; accepted 7 April 1998

Abstract. The X-ray crystal structure of the src SH2 domain revealed the presence of a thiol residue (Cys 188) located proximal to the phosphotyrosine portion of a dipeptide ligand. An aldehyde bearing ligand (**1**) was designed to position an electrophilic carbonyl group in the vicinity of the thiol. X-ray crystallographic and NMR examination of the complex formed between (**1**) and the src SH2 domain revealed a hemithioacetal formed by addition of the thiol to the aldehyde group with an additional stabilizing hydrogen bond between the acetal hydroxyl and a backbone carbonyl.

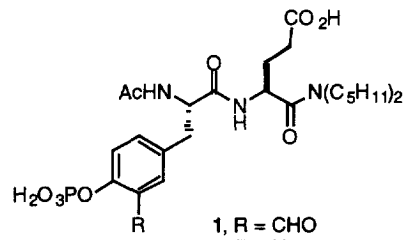
The nonreceptor tyrosine kinase, $pp60^{c\text{-}src}$, has been found to be the major contributor to the elevated tyrosine kinase activity found in cancers of the breast,[1] lung,[2] and colon.[3] It has also been observed that the level of $pp60^{c\text{-}src}$ kinase activity increases with progression of colon cancer from early stage to full metastatic disease.[3] Recently, src has been shown to play a role in mitosis;[4] in linking G-protein coupled receptors to the MAP kinase signalling pathway;[5] and to be involved in a pathway necessary for bone resorption.[6] In addition to the catalytic domain, the src protein contains an approximately 100 amino acid region (SH2 domain) which binds in a sequence dependent manner to phosphotyrosine containing proteins/peptides.[7] The src SH2 domain has been demonstrated to associate with a number of tyrosine phosphorylated proteins, including epidermal growth factor receptor (EGFR),[8] platelet derived growth factor receptor (PDGFR),[9] HER-2/neu,[8,10] and focal adhesion kinase (FAK).[11] Given that src mediated signalling pathways are involved in many proliferative disorders, agents which modulate the ability of src to propagate these signals could be useful tools for research and may have potential as chemotherapeutics.[12]

The relative binding affinity of peptides for the src SH2 domain was shown to depend heavily on the three residues C-terminal to the phosphotyrosine (Y*).[13] The pentamer AcY*EEIE contained the preferred sequence for binding and subsequent structural studies (NMR[14] and X-ray[15]) of this ligand bound to the src SH2 domain revealed a two prong/socket binding mode similar to that found by Waksman[16] for an undecapeptide bound to the v-src SH2 domain. Further investigation of this pentamer led to the identification of tri and dipeptides with comparable binding affinities for the src SH2 domain.[17] Structural analysis of several dipeptide:src SH2 domain complexes revealed a binding motif almost identical to that displayed by the longer sequences in which the phosphate group of the ligand is involved in an extensive network of hydrogen bonds with Arg 158 and Arg 178 of the protein. Particularly provocative was the proximity of the thiol group of Cys 188 to the meta position of the aromatic ring of the phosphotyrosine residue. The demonstration that the src SH2 domain possesses weak phosphatase-like activity[18] together with the suggestion that the thiol group could interact with the phosphotyrosine residue led to the idea that the thiol group might be captured in a (reversible

0960-894X/98/$19.00
PII: S0960-894X(98)00195-4

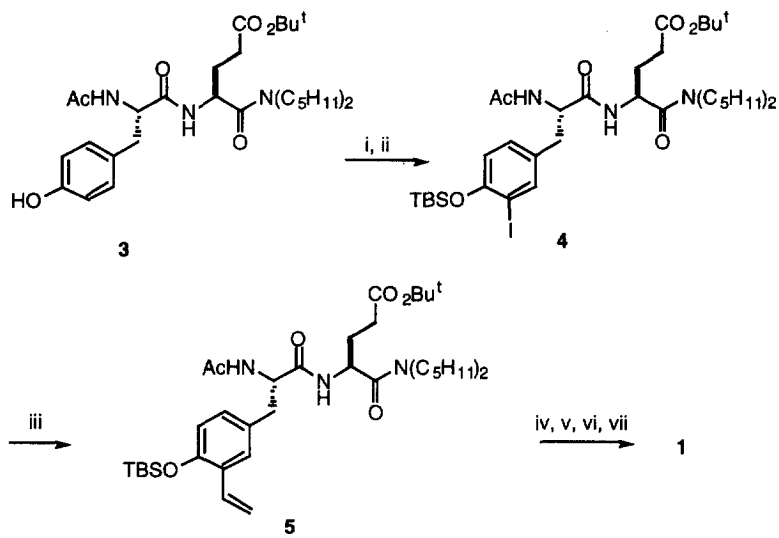
or irreversible) covalent complex. In addition to an expected increase in binding affinity, the formation of a covalent complex with Cys 188 should offer a means to achieve selectivity for the src SH2 domain over SH2 domains of other proteins devoid of similarly positioned cysteines.[19] The use of small, electrophilic groups to engage active site thiols is a well precedented mechanism for inhibition of cysteine proteinases[20] and, in particular, an aldehyde group has been employed frequently to form a hemithioacetal.[21] The o-formylphosphotyrosine dipeptide (**1**) was therefore proposed as a potential selective ligand for the src SH2 domain.

Figure 1



The preparation of **1** was achieved from dipeptide **3** by protection of the phenol followed by iodination to give **4** (Scheme 1). Stille coupling[22] of **4** with vinyl tri-n-butylstannane afforded the styrene **5** in good yield. Sequential deprotection of the phenol, introduction of a protected phosphate group using standard methodology and ozonolysis/reductive work up unmasked the aldehyde group. Finally, acid catalyzed deprotection gave **1**.[23]

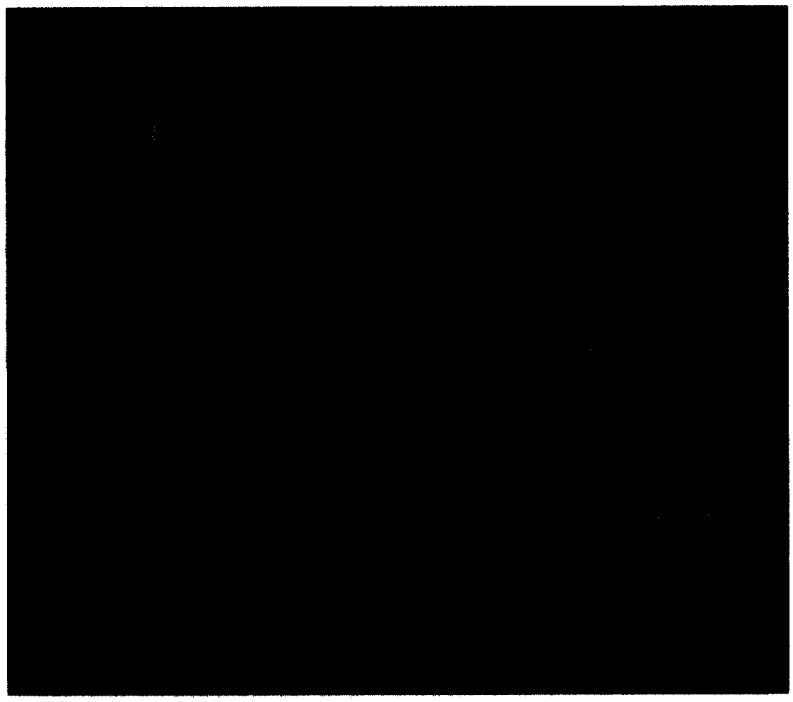
Scheme 1[a]



[a]Reagents: i) ^{t}BuMe$_2$SiCl, Et$_3$N; ii) I$_2$, Hg(OAc)$_2$; iii) VinylSnBu$_3$, Pd(0); iv) Bu$_4$NF; v) NaH, (^{t}BuO)$_2$POCl; vi) O$_3$, Me$_2$S; vii) TFA.

The dipeptide **1** was found to inhibit the binding of the src SH2 domain to activated EGFR with an activity ratio of 3.4 using an ELISA format.[24] This value represents only a twofold improvement in binding affinity when compared to **2** (activity ratio of **2** is 8). These results did not allow firm conclusions to be drawn as to the state of the thiol, however, crystallographic analysis of the complex formed between src SH2 and **1** clearly revealed that the thiol was indeed captured by the aldehyde group to form a hemithioacetal, Figure 2.

Figure 2. X-ray structure of SH2 and **1**



X-ray crystal structure of the covalent thiohemiacetal complex formed between C188 and **1**. The hydrogen bond formed between Y187 and the hydroxyl group of the thiohemiacetal is shown by the dashed line. Details of the crystallization have been described.[25] Data was collected at -170 •C on an RAXIS IIC area detector to 2.0 angstroms resolution. The data was 97% complete, with an Rmerge of 7.1%. The structure was solved by molecular replacement using a previously solved src SH2 structure[25] and refined using XPLOR.[26] The final R factor was 19.8% using data from 6.0-2.0 angstroms. Deviations from ideal bond lengths, angles and torsionals were 0.013 Å, 2.6 degrees, and 23.6 degrees, respectively.

As anticipated, the dipeptide **1** bound to the SH2 domain in a manner virtually identical to AcY*EEIE.[15] In addition to the covalent bond formed between the sulfur atom and the aldehyde carbon, a hydrogen bond is formed between the hemiacetal hydroxyl and the backbone carbonyl of tyrosine 187. Superposition of protein backbone atoms of the complex with **1** with those from non-covalently modified structures reveals an overall

rmsd of 0.5 Å, suggesting that the protein undergoes little, if any, movement upon being covalently bound. In light of the overall similarity of bound and unbound structures, the reason for the very modest improvement in binding affinity becomes unclear. One obvious explanation for the small increase in binding affinity is that the covalent modification of the protein is induced by the crystallization process. To address this suggestion, NMR spectroscopic experiments were conducted to gain further insight into the nature of the complex in solution. Titration of the dipeptide **1** with doubly labelled[27] SH2 domain resulted in a significant decrease in the intensity of the signal corresponding to the aldehyde proton. Furthermore, a proton-carbon correlation map of the fully titrated sample revealed a chemical shift difference for the β-methylene atoms of Cys-188 as compared to a map of the pentapeptide complex. These results, while not definitive, strongly suggest that if an equilibrium exists in solution then the position of this equilibrium lies well toward the covalent complex. These experiments also served to confirm that the aldehyde was indeed present in aqueous solution and was not masked as a hydrate.[28]

In an effort to rationalize the small impact made by covalent bond formation, Charifson and coworkers[25] undertook a detailed thermodynamic study of complexes formed between src SH2 and a variety of ligands. This study suggested that when the covalent complex between SH2 and **1** is compared to the noncovalent complex formed by **2** the ΔH of +8.0 kcal/mol in favour of the covalent complex is consistent with the formation of a new bond. However, the covalent complex was found to be 25.6 EU less favoured entropically. This difference in entropy is unlikely to be the result of subtle desolvation requirements of **1** and it was proposed that this difference results from the inability of the bound phosphotyrosine to transfer lost rotational and translational degrees of freedom into vibrational modes. Thus, any gain realized by formation of a covalent bond has been lost to a high energy complex.

In addition to the challenge of increasing binding affinity for SH2 domains, the recurring problems for ligands designed to interrupt these protein-protein interactions have been the phosphatase lability (of phosphotyrosine containing peptides) and poor cellular penetration of high affinity ligands which bear three negative charges.[28] Compound **1** was found to be devoid of activity when assayed on a variety of cell lines due, presumably, to these same limitations.

The successful use of an electrophilic group to reversibly capture the thiol of Cys188 raises the possibility that this strategy might be employed using different, perhaps uncharged, ligands to afford agents which are not only high affinity ligands for the src SH2 domain but also might penetrate cells effectively.[30]

REFERENCES

1. Ottenhof-Kalff, A. E.; Rijksen, G.; van Beurden, A. C. M.; Hennipman, M. P.; Michels, A. A.; Staal, G. E. J. *Cancer Res.* **1992**, *49*, 4773.
2. Mazurenko, N. N.; Kogan, E. A.; Zborovskaya, I. B.; Kisseljov, F. L. *Eur. J. Cancer* **1992**, *28*, 372.
3. Weber, T. K.; Steele, G.; Summerhayes, I. C. *J. Clin. Inves.* **1992**, *90*, 815.
4. Fumagalli, S.; Totty, N. F.; Hsuan, J. J.; Courtneidge, S. A. *Nature* **1994**, *368*, 871. Pillay, I.; Nakano, H.; Sharma, S. V. *Cell Growth Differ.* **1996**, *7*, 1487.
5. Dikic, I.; Tokiwa, G.; Lev, S.; Courtneidge, S. A.; Schlessinger, J. *Nature* **1996**, *383*, 547.
6. Tanake, S.; Amling, M.; Neff, L.; Peyman, A.; Uhlmann, E.; Levy, J. B.; Baron, R. *Nature* **1996**, *383*, 528.

7. Pawson, T. *Curr. Opin. Struc. Biol.* **1992**, *2*, 432. Margolis, B. *Cell Growth Differ.* **1992**, *3*, 73. Pawson, T.; Gish, G. D. *Cell* **1992**, *71*, 359.
8. Luttrell, D. K.; Lee, A.; Lansing, T. J.; Crosby, R. M.; Jung, K. D.; Willard, D.; Luther, M.; Rodriguez, M.; Berman, J.; Gilmer, T. M. *Proc. Nat. Acad. Sci. U.S.A.* **1994**, *91*, 83.
9. Mori, S.; Ronnstrand, L.; Yokote, K.; Engstrom, A.; Courtneidge, S. A.; Claesson-Welsh, L.; Heldin, C. - H. *EMBO* **1993**, *12*, 2257.
10. Muthuswamy, S. K.; Segel, P. M.; Dankort, D. L.; Webster, M. A.; Muller, W. J. *Mol. Cell. Biol.* **1994**, *14*, 735.
11. Cobb, B. S.; Schaller, M. D.; Leu, T. H.; Parsons, J. T. *Mol. Cell. Biol.* **1994**, *14*, 147.
12 Levitski, A. *Anti-Cancer drug design* **1996**, *11*, 175. Saltiel, A. R.; Sawyer, T. K. *Chem. Biol.* **1996**, *3*, 887.
13. Songyang, Z.; Shoelson, S. E.; Chaudhuri, M.; Gish, G.; Pawson, T.; Haser, W. G.; King, F.; Roberts, T.; Ratnofsky, S.; Lechleider, R. J.; Neel, B. G.; Birge, R. B.; Fajardo, J. E.; Chao, M. M.; Hanafusa, H.; Schaffhausen, B.; Cantley, L. C. *Cell* **1993**, *72*, 767.
14. Xu, R. X.; Word, M.; Davis, D.; Rink, M.J.; Willard, D. H.; Gampe, R. T. *Biochemistry* **1995**, *34*, 2107.
15. Gilmer, T. M.; Rodriguez, M.; Jordan, S.; Crosby, R.; Alligood, K.; Green, M.; Kimery, M.; Wagner, C.; Kinder, D.; Charifson, P.; Hassell, A. M.; Willard, D.; Luther, M.; Rusnak, D.; Sternbach, 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.
16. Waksman, G.; Shoelson, S. E.; Pant, N.; Cowburn, D.; Kuriyan, J. *Cell* **1993**, *72*, 779.
17) Rodriguez, M.; Crosby, R.; Alligood, K.; Gilmer, T. M.; Berman, J. *Lett. Pep. Sci.* **1995**, *2*, 1. Plummer, M. S.; Lunney, E. A.; Para, K. S.; Prasad, J. V. N. V.; Shahripour, A.; Singh, J.; Stankovic, C. J.; Humblet, C.; Fergus, J. H.; Marks, J. S.; Sawyer, T. K. *Drug Design and Discovery* **1996**, *13*, 75. Plummer, M. S.; Lunney, E. A.; Para, K. S.; Shahripour, A.; Stankovic, C. J.; Humblet, C.; Fergus, J. H.; Marks, J. S.; Herrera, R.; Hubbell, S.; Saltiel, A.; Sawyer, T. K. *Bioorg. Med. Chem.* **1997**, *5*, 41.
18. Boerner, R. J.; Consler, T. G.; Gampe, R. T. Jr.; Weigl, D.; Willard, D. H.; Davis, D. G.; Edison, A. M.; Loganzo, F. Jr.; Kassel, D. B.; Xu, R. X.; Patel, I. R.; Robbins, J. S.; Lansing, T.; Gilmer, T. M.; Luther, M. A.; Knight, W. B. *Biochemistry* **1995**, *34*, 15351.
19. SH2 domain sequences published to date indicate that a cysteine residue in the 188 position is unique to src.
20) Rich, D. H. In *Comprehensive Medicinal Chemistry*; Hansch, C.; Sammes, P. G.; Taylor, J. B., Eds.; Pergamon: Oxford, 1990; pp 431-433. Krantz, A. In *Advances in Medicinal Chemistry*; Maryanoff, B. E.; Maryanoff, C. A., Eds.; JAI: Greenwich, 1992; vol. 1, pp 235-261.
21. Woo, J. T.; Sigeizumi, S.; Yamaguchi, K.; Sugimoto, K.; Kobori, T.; Tsuji, T.; Kondo, K. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 1501.
22. Stille, J. K. *Angew. Chem., Int. Ed. Engl.* **1986**, *25*, 508.
23. ^{1}H NMR (400 MHz, DMSO-d_6) • 10.26 (s, 1H), 8.26 (d, 1H, J = 8.4 Hz), 8.07 (d, 1H, J = 8.4 Hz), 7.69 (s, 1H), 7.54 (d, 1H, J = 8.4 Hz), 7.26 (d, 1H, J = 8.8 Hz), 4.67 (m, 1H), 4.50 (m, 1H), 3.4-2.6 (m, 5H), 1.72 (s, 3H), 0.86 (t, 3H, J = 7.2 Hz), 0.87 (t, 3H, J = 7.2 Hz). ^{1}H NMR (400 MHz, MeOH-d_4) • 8.11 (d, 1H), 7.45 (s, 1H), 7.25 (d, 1H, J = 8 Hz), 7.19 (d, 1H, J = 8.8 Hz), 5.64 (s, 1H), 4.58 (dd, 1H, J = 9.2, 4.8 hz), 3.48 (m, 2H), 3.12 (m, 2H), 2.81 (dd, 1H, J = 14, 10 Hz), 2.38 (t, 2H, J = 6.6 Hz), 1.88 (s, 3H), 0.93 (t,

3H, J = 7.2 Hz), 0.89 (t, 3H, J = 7.2 Hz). ESMS 600 (M+H, base). Anal. calcd for $C_{27}H_{42}N_3O_{10}P \cdot H_2O$: C, 52.51; H, 7.18; N, 6.80. Found: C, 52.85; H, 7.19; N, 6.96.

24. Activity ratio is defined as IC_{50} (test compound)/IC_{50} (AcY*EEIE). For details of the ELISA see ref 15.
25. Charifson, P. S.; Shewchuk, L. M.; Rocque, W.; Hummel, C. W.; Jordan, S. R.; Mohr, C.; Pacofsky, G. J.; Peel, M. R.; Rodriguez, M.; Sternbach, D. D.; Consler, T. G. *Biochemistry* **1997**, *36*, 6283.
26. Brunger, A. T. *Acta Crytallogr.* **1990**, *A46*, 46.
27. Labeled protein (^{13}C and ^{15}N) was described in ref 14.
28. The ^1H NMR spectrum of **1** taken in MeOH-d_4 shows the aldehyde masked as an acetal.
29. The problems of phosphatase lability and cell permeability have been investigated by a number of workers. See for example: Burke, T. R.; Smyth, M. S.; Otaka, A.; Nomizu, M.; Roller, P. P.; Wolf, G.; Case, R.; Shoelson, S. E. *Biochemistry*, **1994**, *33*, 6490-6494.
30. Following the completion of this study, Jagoe and co-workers reported that a series of catechols inhibited the binding of src family SH2 domains to high affinity ligands. The proposed mechanism involves initial oxidation of the catechol to an *o*-quinone which may then react with a thiol group such as that found at Cys-188 of the src SH2 domain. Jagoe, C. T.; Kreifels, S. E.; Li, J. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 113.