A Molecular Gate which Controls Unnatural ATP Analogue Recognition by the Tyrosine Kinase v-Src

Y. Liu, a K. Shah, a F. Yang, a L. Witucki a and K. M. Shokat a, b,*

a Department of Chemistry, Princeton University, Princeton, NJ 08544, U.S.A.
b Department of Molecular Biology, Princeton University, Princeton, NJ 08544, U.S.A.

Received 21 April 1998; accepted 8 May 1998

Abstract—Engineered proteins with specificity for unnatural substrates or ligands are useful tools for studying or manipulating complex biological systems. We have engineered the prototypical tyrosine kinase v-Src to accept an unnatural ATP analogue N^6-(benzyl) ATP in order to identify v-Src’s direct cellular substrates. Here we have used molecular modeling to analyze the binding mode of N^6-(benzyl) ATP. Based on this modeling we proposed that a new ATP analogue (N^6-(2-phenethyl) ATP might be a better substrate than N^6-(benzyl) ATP for the I338G mutant of v-Src. In fact the newly proposed analogue (N^6-(2-phenethyl) ATP is a somewhat improved substrate for the engineered kinase (k_cat = 0.6 min^{-1}, K_M = 8 µM). We also synthesized and screened three analogues of N^6-(benzyl) ATP: N^6-(2-methylbenzyl), ATP N^6-(3-methylbenzyl), and ATP N^6-(4-methylbenzyl) ATP to further probe the dimensions and shape of the introduced pocket. Results from screening newly synthesized ATP analogues agreed well with our modeling predictions. We conclude that rather than engineering a ‘new’ pocket by mutation of Ile 338 in v-Src to the smaller Ala or Gly residues, the I338G and I338A mutants possess a ‘path’ for the N^6 substituent on ATP to gain access to an existing pocket in the ATP binding site. We expect to be able to extend the engineering of v-Src’s ATP specificity to other kinase families based on our understanding of the binding modes of ATP analogues to engineered kinases. © 1998 Elsevier Science Ltd. All rights reserved.

Introduction

Protein kinases play a central role in controlling many diverse signal transduction pathways in all cells.1–4 These enzymes catalyze the phosphorylation of tyrosine, serine, or threonine residues on proteins using ATP as the phosphodonor. The identification of the cellular substrates of individual protein kinases remains one of the central challenges in the field. We have recently developed a chemical method to tag the direct substrates of two Src family kinases, v-Src and Fyn.5,6 To distinguish the substrates of v-Src from all other kinase substrates we used structure based design and site directed mutagenesis to make v-Src catalyze a unique phosphotransfer reaction not catalyzed by any other protein kinase in the cell. We engineered the ATP binding site of v-Src to uniquely accept an ATP analogue (A*TP) as the phosphodonor substrate. The engineering of v-Src’s active site to accept [γ-32P] A*TP provides a handle by which the direct substrates of v-Src can be specifically radiolabeled in the presence of any number of cellular kinases (Fig. 1).

We have successfully engineered a number of tyrosine kinases to accept unnatural ATP analogues (reference6 and unpublished results). We accomplished this by using a semi-rational approach. We synthesized N^6-substituted derivatives of ATP. We then mutated a bulky residue in the v-Src kinase, Ile 338, to a smaller group, Ala or Gly. Isoleucine 338 was chosen because it appeared to be in close contact with the N^6 amine of ATP (Fig. 2(A) and (B)). By screening a large panel of N^6 substituted ATP analogues containing aliphatic or aromatic substituents we identified the optimal A*TP substrates for the I338A and I338G mutated v-Src kinases.

To our surprise, the best analogue contained a benzyl substituent which is substantially larger than the amino...
We were not the first to combine protein engineering and chemical synthesis to study complex biological systems. Table 1 lists several elegant applications of this strategy to a variety of biological systems. Schreiber and his co-workers at Harvard and Clackson and co-workers at Ariad Pharmaceuticals have both developed non-natural FK506 ligands with designed FK506-binding protein mutant partners for the control of protein dimerization and gene transcription. The engineering...
of uniquely tracable GTPases has proven particularly valuable for answering precise questions about individual GTPases such as Elongation factor-Tu, Rabs, signal recognition particle, as well as others.\textsuperscript{11–13} Amazingly, a single point mutation converts a wide range of GTPases into xanthine triphosphatases (XTPases) without loss of the normal function of the GTPase of interest. A combined chemical and genetic approach has also been applied by Smith and co-workers at Glaxo-Wellcome to dramatically improve antibody directed enzyme immunotherapy (ADEPT) mediated in vitro cell killing of a human colon adenocarcinoma cell line.\textsuperscript{14} In this system human carboxypeptidase A1 was engineered to specifically cleave a

![Figure 2](image_url)

(A) Predicted close contact between the side chain of residue 338 (light blue) in Hck and the benzyl ring of N\textsuperscript{6}-(benzyl) ATP (green). The conformation of the N\textsuperscript{6} benzyl ring was chosen by manually checking multiple rotamers of the N-C and C-Ph bonds in N\textsuperscript{6}-(benzyl) ATP for steric clashes with the residues in the Hck active site using the InsightII program’s “bump” function. (B) Stereo view of the predicted binding conformation of A\textsuperscript{*}TP 1–3 in Hck. The Hck residue atoms within 7 Å are shown in black. Hck with bound quercetin (blue) is superimposed with AMP-PNP (green). Several substituents have been added to the N\textsuperscript{6} position of AMP-PNP, benzyl (green, 2), phenyl (pink, 1), and phenethyl (red, 3). These substituents are added to the N\textsuperscript{6} amine and the rotamer with the least steric clashes determined manually are displayed.
methotrexate-prodrug which is not hydrolyzable by any naturally occurring proteases thus releasing the active chemo- therapeutic, methotrexate, at the tumor.

Nucleic acid based molecular recognition has also been exploited using a similar approach. An example is the use of an unnatural amber codon suppressor tRNA charged with an unnatural amino acid, which has been used by Schultz and co-workers at UC Berkeley to synthesize proteins containing an unnatural amino acid.15,16 The method has also been extended to synthesize membrane proteins in frog egg oocytes by Dougherty and co-workers at Caltech.17,18 Schultz’s group is taking the method even further in an effort to engineer an organism containing a suitably constructed unnatural tRNA synthetase that allows ribosomal protein synthesis utilizing an expanded genetic code.19,20

The variety of systems studied using designed pairs of synthetic substrates and engineered proteins indicates the versatility of the approach. The common feature of all these methods is the design and synthesis of an unnatural substrate or ligand which is not used by any natural enzyme or receptor in the system of interest. Such a molecule is a so-called ‘orthogonal’ substrate or ligand. This term is used to indicate that the synthetic analogue is not recognized by any of the biological targets which recognize the unmodified (natural) substrate. The next step is to redesign the enzyme/receptor of interest to utilize the ‘orthogonal’ analogue. This combined chemical and genetic method essentially breaks down the degeneracy of natural systems. This allows for very precise analysis and control over large systems of proteins that is not afforded by purely genetic or purely biochemical methods.

Results and Discussion

Molecular modeling

Initially we relied on crystal structures of ser/thr kinases, PKA and CDK2,21,22 to guide our engineering of the tyrosine kinase v-Src. No Src family kinase structures were available when we initiated our engineering of v-Src. Crystal structures of PKA and CDK2 which only share approximately 50% homology to v-Src served as a limited guide to our design efforts. Based only on these structures however, we introduced two mutations (V323A, I338A) into v-Src, affording a kinase with good specificity for N6 (cyclopentyl) ATP.5

In the design of second generation v-Src mutants we were guided by the recently solved structures of three tyrosine kinases c-Src, Hck and Lck.23–25 Relying heavily on the Hck-AMP-PNP co-crystal structure and kinase sequence alignments we found that a single bulky

<table>
<thead>
<tr>
<th>Natural system</th>
<th>Engineered system</th>
<th>Biological application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein</td>
<td>Substrate of ligand</td>
<td>Protein</td>
</tr>
<tr>
<td>FKBP12-rapamycin binding domain (FRB)</td>
<td>Rapamycin</td>
<td>Mutant FRB (T2098L, W2101F, K2095P)</td>
</tr>
<tr>
<td>Glutaminyl-tRNA aminoacyl synthetase ((GlnRS)-tRNA\textsubscript{\text{Gln}})</td>
<td>tRNA\textsubscript{\text{Gln}}</td>
<td>Mutated GlnRS enzyme that acylates the engineered tRNA</td>
</tr>
<tr>
<td>Elongation factor -TU (EF-Tu)</td>
<td>Guanosine triphosphate (GTP)</td>
<td>Mutant EF-Tu (D138N)</td>
</tr>
<tr>
<td>Human enzyme carboxypeptidase A1 (hcpA1)</td>
<td>MTX-\text{3-phenylalanine}</td>
<td>Mutant hcpA1 (T268G)</td>
</tr>
</tbody>
</table>
residue at the position corresponding to Ile 338 in v-Src is primarily responsible for restricting the ability of wild-type kinases to accept N6-substituted ATP analogues. To our surprise, N6-(benzyl) ATP was the best substrate for the v-Src single mutant (I338A or I338G) after screening a panel of over 35 unique A*TPs (reference6 and unpublished results). We had initially expected an A*TP with a smaller substituent attached at the N6 position of ATP to better complement the small mutation we introduced.

Although we do not have a crystal structure of the engineered v-Src kinase, we do have several ligand bound structures of wild-type Src family kinases to help in predicting the binding mode of our orthogonal ATP analogue, N6-(benzyl) ATP to the engineered v-Src. When we superimposed the co-crystal structures of Hck/AMP-PNP and Hck/quercetin we noticed the presence of a large binding pocket above the N7 of ATP which is occupied by the catechol ring of quercetin (Fig. 2(B)). This pocket is adjacent but not directly opposite to the site of substitution (N6) on our synthetic ATP substrates. To determine if this existing pocket was close enough to at least partially accommodate the substituents added to the N6-position of ATP we modelled the binding mode of several A*TPs.

To model the binding mode of N6-substituted ATP analogues we first substituted one of the N6 hydrogens of AMP-PNP with a benzyl group in the co-crystal structure of AMP-PNP bound to Hck. The bound conformation of the benzyl ring was chosen by manually checking multiple rotamers of the N-C and C-Ph bonds in N6-(benzyl) ATP for steric clashes with the residues in the Hck active site using the ‘bump’ function in InsightII (InsightII 4.0.0, San Diego: Molecular Simulations). Using this procedure the aromatic ring of the benzyl group makes the least steric clashes with the kinase in the conformation shown in Figure 2(B). The benzyl group is predicted to lie adjacent to the catechol ring of quercetin (the closest distance between two carbon atoms from each of the rings is 0.5 Å).

The one severe steric clash suffered by N6-(benzyl) ATP with the wild-type kinase is with the side chain of residue 338. This appears to be the key residue which blocks access of N6-substituents to the large pocket occupied by the catechol ring of quercetin. Thus, in order for the N6-substituted benzyl group to gain access to the existing pocket adjacent to N7, the side chain of residue 338 (a ‘molecular gate’) needs to be shortened (i.e. I338A or G). We predict the benzyl ring of N6-(benzyl) ATP partially occupies a pocket already present in v-Src after the ‘molecular gate’ which normally restricts access to this pocket in all wild-type kinases is removed by our mutation. This would explain why a larger than expected A*TP (>80 Å3) is the optimal analogue identified in a screen of over 35 A*TPs against I338G v-Src.

From the above analysis we think that the ‘linker’ between the N6 amine and the aromatic ring of N6-(benzyl) ATP is important because it allows the aromatic ring the necessary flexibility to occupy the existing pocket near the N7 nitrogen. If one more methylene group is added to the linker (N6-(2-phenethyl) ATP, 3), it would be expected to allow the aromatic ring more flexibility to reach the existing pocket at N7. From the overlay of the modeled N6-(2-phenethyl) ATP (Fig. 2(B), red) and X-ray structure of bound quercetin (Fig. 2(B), blue), the aromatic ring of N6-(2-phenethyl) ATP is adjacent to the catechol ring of quercetin and the two rings are planar (Fig. 2(B)). The only steric clashes of N6-(2-phenethyl) ATP with the protein come from the aromatic ring of the phenethyl group and the side chain of the T338 residue in Hck (Fig. 2(A), light blue ball-and-stick).

We predicted that removal of the Thr side chain in Hck or the corresponding Ile residue on v-Src would allow this analogue, N6-(2-phenethyl) ATP to be accepted as a substrate. Furthermore, we hoped that it would be a better substrate than our previous optimal analogue, N6-(benzyl) ATP. If the methylene group between the phenyl ring of N6-(benzyl) ATP is removed, the resulting ATP analogue, N6-(phenyl) ATP is not expected to bind to the I338A or I338G mutant of v-Src because the phenyl group will clash with atoms on the protein other than the side chain of I338, (Fig. 2(B)).

We also synthesized methyl substituted analogues of N6-(benzyl) ATP to further probe the pocket into which the phenyl ring is predicted to fit. We reasoned that if we could correctly predict which of the methyl substituted isomers would be the most potent based on modeling them into the N7 pocket, we would have more confidence that our binding mode prediction was correct. Of the 4-, 3- and 2-positions of N6-(benzyl) ATP, we predicted 3-methyl substitution would fit best into the v-Src pocket. Only the meta position points towards an open space in the pocket. Methyl substitution at the
4- or 2-positions is predicted to result in clashes with other residues in the protein.

**Insights from new synthetic A*TPs**

We synthesized two new classes of phenyl substituted N6 derivatives. The first class contains 0, 1 or 2 methylene units as spacers between the phenyl ring and the purine ring of ATP, N6-(phenyl) ATP 1, N6-(benzyl) ATP 2, N6-(2-phenethyl) ATP 3, respectively. The second set of A*TPs were designed to test the dimensions and shape of the pocket occupied by substituted N6-(benzyl) ATP analogue: 4-, 3-, or 2-methyl substituted N6-(benzyl) ATP, 4, 5, and 6, respectively.

As predicted, N6-(2-phenethyl) ATP has the highest binding affinity for the mutant (I338A) and (I338G) v-Src proteins (Fig. 3(A)). The I338A and I338G mutants are inhibited better by N6-(benzyl) ATP than by N6-(phenyl) ATP (Fig. 3(A)). This confirms our

![Figure 3](image-url)

**Figure 3.** (A) Evaluation of the binding efficiency of A*TPs (1–6) to GST-XD4(wild type), GST-XD4(I338A) and GST-XD4(I338G). Inhibition of [γ-32P] ATP-dependent phosphorylation of the peptide substrate IYGEFKKK by GST-XD4(wt), GST-XD4(I338A) and GST-XD4(I338G) with unlabelled ATP and A*TPs 1–6. Percent inhibition (1−v1/v0) is reported as a ratio of v1[cpm in the presence of 100 μM of the indicated triphosphate and 10 μM [γ-32P] ATP(1000 cpm/pmol)]/v0[cpm in the presence of 10 μM [γ-32P] ATP (1000 cpm/pmol)] alone. (B) Structures of A*TP analogs.
modeling method and shows that the linker between N6 and the ring does confer the flexibility to reach the space at N7. The comparison of the inhibition pattern by the three methyl substituted N6-(benzyl) ATPs also confirmed our modeling predictions. The N6-(3-methyl-benzyl) ATP inhibits the mutants almost as well as N6-(benzyl) ATP while the 4-methyl and 2-methyl analogues are considerably poorer inhibitors of the mutant (I338A, or I338G) of v-Src (Fig. 3(A)). The poorer inhibition by these regioisomers suggests that the added methyl groups are encountering significant steric repulsions with other residues in the active site pocket, as we predicted.

The goal of our engineering effort is not identification of good inhibitors of kinases, but a significantly more challenging goal of identifying optimal substrates of the engineered kinases. We tested the catalytic efficiency of our new A*TPs as substrates for the mutated v-Src kinases using a green fluorescent protein engineered with a phosphorylatable tail (GFP-IYGEF) in a phosphotyrosine western blot assay developed in our lab. A*TPs 1–6 are not substrates of wild type kinases, confirming that these are suitable orthogonal substrates to use in our engineered kinase substrate tagging methodology (Fig. 4, lanes 3–8).

From the screen for catalytic activity of the new panel of analogues, 1–6, N6-(2-phenethyl) ATP, 3, is a better substrate \(k_{cat}=0.6 \text{ min}^{-1}, K_M=8 \text{ mM}\) with I338A mutant) than our previous optimal analogue, N6-(benzyl) ATP\(k_{cat}=0.5 \text{ min}^{-1}, K_M=20 \text{ mM}\) with I338A mutant). For comparison, the kinetic constants for the I338A mutant with ATP are \(k_{cat}=1.0 \text{ min}^{-1}, K_M=70 \text{ mM}\). This again confirms our prediction about adding extra flexibility into the substrate serving to orient the orthogonal substrate into the existing pocket above the N7 nitrogen.

The most striking selectivity shown by the I338A and I338G mutants are for N6-(3-methyl) benzyl ATP, 5, over the regioisomers 2-methyl and 4-methyl analogues, 6 and 4, respectively. Just as our modeling analysis suggests the 3-methyl isomer should be accommodated by the geometry of the kinase active site, while the other isomers will not be able to adopt a productive binding orientation. Both the inhibition (Fig. 3) and catalytic activity assays (Fig. 4) support this analysis.

**Conclusion**

In summary, we have mapped the binding site for orthogonal N6-substituted A*TP substrates of mutant v-Src proteins containing I338A or I338G mutations. We found the pocket in the v-Src tyrosine kinase that accommodates the added substituents on ATP analogues 1–3 and 5 is actually present in the wild-type v-Src kinase. It is access to the existing pocket that is controlled by Ile 338 in v-Src. Rather than engineering a ‘new’ pocket by mutation of Ile to the smaller Ala or Gly residues, we believe that the I338G and I338A mutations provides a ‘path’ for the N6-substituent on ATP to gain access to an existing pocket in the ATP binding site. We think this explains the large discrepancy in size between the substituent added to the substrate A*TP and the amino acid side chain.

![Figure 4](image-url)
removed from the kinase. Based on this new appreciation of the binding mode of our A*TP analogues we successfully identified a new analogue, N6-(2-phenethyl) ATP, which is a better substrate for mutated v-Src tyrosine kinase than our previous optimal analogue, N6-(benzyl) ATP. The fact that the N6 phenethyl substituent can adopt more conformations may also mean that it can serve as a better generic analogue to fit into non-Src family kinases containing similar Ala or Gly mutations at the position corresponding to Ile338 in v-Src.

Experimental

Synthesis of ATP analogues. Analogues (1–6) were synthesized as described previously. In brief, analogues 1–6 were prepared by refluxing 6-chloropurine riboside (Aldrich) with aniline, benzylamine, 2-phenylethylamine, 4-methyl-benzylamine, 3-methyl-benzylamine, 2-methyl-benzylamine, respectively, in ethanol overnight. Triphosphate synthesis was carried out as described previously.5

Peptide synthesis. The v-Src substrate peptide, IYGEFKKK, was synthesized as described previously.

Site directed mutagenesis, protein expression, and purification. The site directed mutagenesis, expression and purification of the wild-type and mutant kinases were described previously.

In vitro kinase peptide assay. The inhibition of the wild-type and mutant kinases with ATP and ATP analogues were tested as described previously.

GFP-IYGEF Western blot assay. The non-radioactive assay to test the catalytic efficiency of wild-type and mutant kinases with ATP analogues was carried out as described previously.

Acknowledgements

The authors work was supported by a National Science Foundation Early Career Development Award (MCB-9506929) and by the National Institutes of Health (1R01CA70331-01). K.M.S. is a Pew Scholar in the Biomedical Sciences. The authors thank members of the Shokat Lab for helpful discussion and comments on the manuscript.

References