Pergamon

Bioorganic & Medicinal Chemistry Letters 12 (2002) 3223–3227

BIOORGANIC & MEDICINAL CHEMISTRY LETTERS

Conformational Restraint is a Critical Determinant of Unnatural Nucleotide Recognition by Protein Kinases

Scott M. Ulrich,[a] Nathan A. Sallee[b] and Kevan M. Shokat[a,c,*]

[a]*Department of Cellular and Molecular Pharmacology, University of California, San Francisco, CA 94143-0450, USA*
[b]*Program in Chemistry and Chemical Biology, University of California, San Francisco, CA 94143-0450, USA*
[c]*Department of Chemistry, University of California, Berkeley, CA 94720, USA*

Received 11 April 2002; accepted 27 April 2002

Abstract—This report describes the synthesis of N^4-(benzyl) AICAR triphosphate, a conformationally restrained analogue of N^4-(benzyl) ribavirin triphosphate. Both of these nucleotides were evaluated as phosphodonors for wild-type p38 MAP kinase and T106G p38 MAP kinase, a designed mutant with expanded nucleotide specificity. The conformationally restrained nucleotide, N^4-(benzyl) AICAR triphosphate, is orthogonal to (not accepted as a substrate by) wild-type p38 MAP kinase, in contrast to N^4-(benzyl) ribavirin triphosphate. Furthermore, N^4-(benzyl) AICAR triphosphate, is accepted as a substrate by T106G p38 MAP kinase, in contrast to N^4-(benzyl) ribavirin triphosphate. We hypothesize that the presence of an internal hydrogen bond in N^4-(benzyl) AICAR and its absence in N^4-(benzyl) ribavirin triphosphate is the main determinant for their differing structure–activity relationships.

Protein kinases are a large enzyme family that catalyze phosphoryl group transfer from ATP to tyrosine or serine/threonine residues on cellular proteins. Protein phosphorylation is a central mechanism of cellular signal transduction and controls pathways such as insulin signaling, T- and B-cell activation, cellular differentiation, mitogenic signaling, and many others.[1–3]

Identification of the direct substrate(s) of an individual protein kinase within a complex signaling cascade involving many kinases has been difficult to accomplish with existing experimental techniques. Efforts to elucidate protein kinase substrates using peptide libraries[4–7] do not account for the secondary and tertiary structure of substrate proteins, known to be a critical determinant of protein kinase substrate specificity.[8] Screening cDNA libraries expressed on phage for protein kinase substrates has also been used to identify candidate substrates of MAP kinases[9] and Src.[10] However, it is clear that these experimental conditions are vastly different from conditions found in the cell, where protein–protein interactions and cellular localization are both critical determinants of protein kinase substrate specificity.[1]

Our laboratory has developed a chemical-genetic[11–13] strategy to identify the direct substrates of protein kinases. This method involves engineering a protein kinase to utilize a synthetic analogue of ATP [N^6-(benzyl) ATP] that is orthogonal to, or a poor substrate of, all wild type kinases. The benzyl substituent of N^6-(benzyl) ATP contacts I338 in v-Src, causing a severe steric clash.[14] Truncating the I338 residue to glycine in v-Src[15] (or the corresponding mutation in other kinases[16,17]) allows N^6-(benzyl) ATP to bind and support phosphoryl group transfer. Adding γ-[^{32}P]-N^6-(benzyl) ATP to a complex cell preparation such as a cell lysate or immunoprecipitation containing the mutant kinase results in specific labeling of the direct substrates of only the mutant kinase, even in the presence of all cellular wild type kinases.[15]

Live or intact cells are the best context in which to identify bona fide kinase substrates, since subcellular localization and protein–protein interactions are key determinants of kinase signaling specificity.[1] Labeling kinase substrates in vivo would require transient permeablization of the cell membrane to allow impermeant ATP analogues to enter.[18] This limited amount of γ-[^{32}P] labeled ATP analogue would then compete with the high (millimolar) cellular concentration of ATP for the mutant kinase active site. In this case, it would be

*Corresponding author. Tel.: +1-415-514-0472; fax: +1-415-502-8644; e-mail: shokat@cmp.ucsf.edu

0960-894X/02/$ - see front matter © 2002 Elsevier Science Ltd. All rights reserved.
PII: S0960-894X(02)00616-9

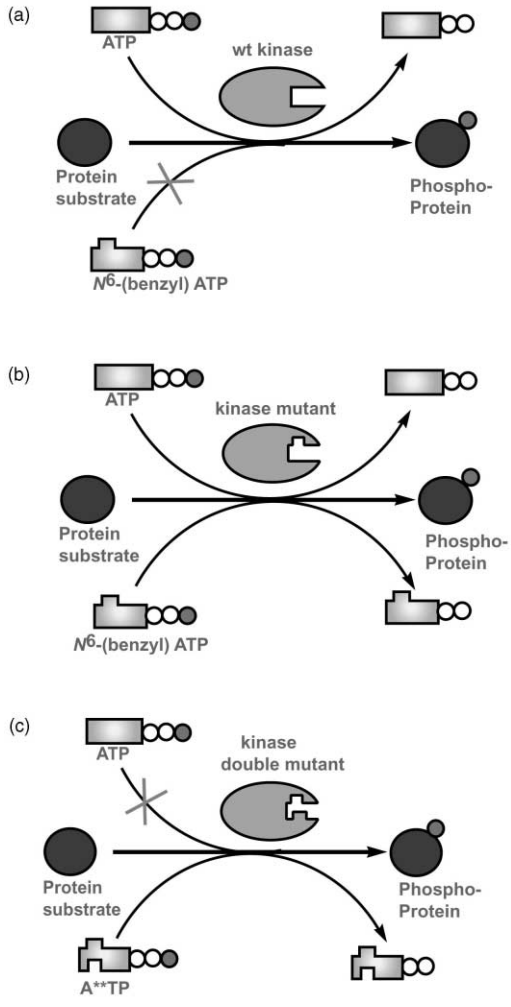


Figure 1. Design goals of natural and unnatural nucleotides with wild type and mutant protein kinases.

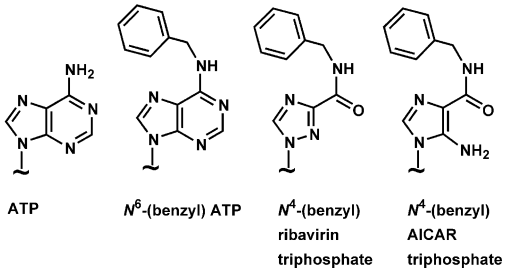


Figure 2. Nucleotides discussed in this report.

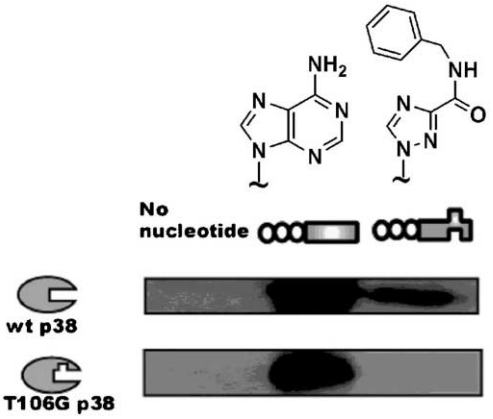


Figure 3. Wild type p38 (Row 1) and T106G p38 (Row 2) kinase assay with no nucleotide (Column 1), ATP (Column 2) and N^4-(benzyl) ribavirin triphosphate (Column 3), using a western blot assay which detects phospho-ATF-2, a substrate of p38.

desirable to engineer a kinase allele that *solely* uses analogues of ATP, but not ATP, GTP, or any other naturally occurring nucleotide triphosphate. A kinase that is insensitive to ATP will provide a higher signal-to-noise ratio of substrate labeling, which is likely to be critical for the identification of low abundance kinase substrates in intact cells.

Here we report progress towards an ATP analogue that is suited for a protein kinase that is insensitive to ATP according to the following design strategy. N^6-(benzyl) ATP is a poor substrate for wild type kinases, because it is sterically excluded from the active site (Fig. 1a). A complementary space-creating mutation (I338G in v-Src) is known to allow recognition and catalysis with N^6-(benzyl) ATP (Fig. 1b).[19,20] We reasoned that making small-to-large amino acid substitutions in the kinase active site would sterically exclude ATP from the active site. Synthetic ATP analogues that have atoms 'deleted' from the adenine scaffold could avoid a steric clash with this second mutation and bind the pocket. Additionally, in order to label substrates specific to a mutant kinase, such ATP analogues would have to be orthogonal to the set of wild type protein kinases in the cell; thus the original N^6-benzyl substituent (and corresponding I338G kinase mutation) should, if possible, remain in place (Fig. 1c).

We previously described[21] a nucleotide that was designed to meet these criteria based on ribavirin, a clinical antiviral agent.[22] Ribavirin has the structural features that we require to fit a sterically occluded protein kinase active site, yet is capable of presenting a benzyl substituent to the I338G pocket and satisfies hydrogen bonding patterns seen in most kinase/ligand crystal structures (Fig. 2). We synthesized N^4-(benzyl) ribavirin triphosphate and tested this compound as a phosphodonor for wild type p38 mitogen-activated protein kinase (MAP kinase)[23] and T106G p38[†] MAP kinase. We expressed p38 and T106G p38 carrying an N-terminal 6-histidine purification tag from bacteria as described.[24] Activating transcription factor-2 (ATF-2), a known p38 substrate[25] was used in conjunction with a phospho-ATF-2 specific antibody (Cell Signaling, Beverly, MA)[‡] to detect p38-dependent phosporylation of ATF-2 by western blotting (Fig. 3).

In agreement with similar experiments using v-Src and I338G v-Src,[21] the data indicate that N^4-(benzyl) ribavirin

[†]T106 in p38 corresponds to I338 in v-Src by sequence alignment.

[‡]Purified p38 or T106G p38 (100 ng), ATF-2 (2 μg), were incubated with the appropriate nucleotide (100 μM) for 20 min at 23 °C in kinase buffer (50 mM Tris pH 8, 10 mM $MgCl_2$, 100 mM NaCl) in a final volume of 30 μL. Proteins were resolved by PAGE, transferred to a nitrocellulose membrane and probed by Western blotting.

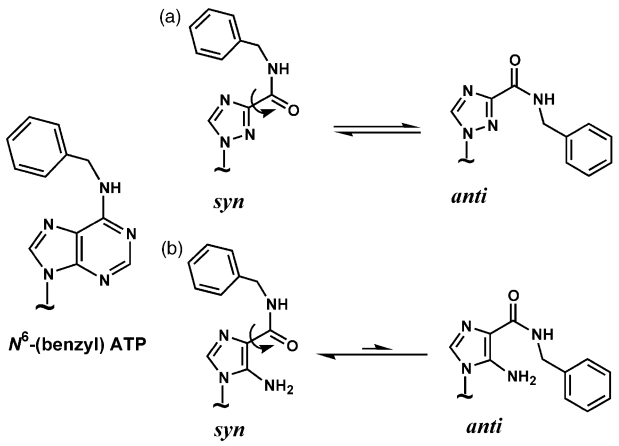


Figure 4. N^4-(benzyl) ribavirin triphosphate (a) can readily interconvert between two conformers with respect to the amide-triazole bond. *Syn* N^4-(benzyl) ribavirin triphosphate mimics N^6-(benzyl) ATP, yet binds T106G p38 (a kinase that efficiently uses N^6-(benzyl) ATP) in a non-productive conformation. N^4-(benzyl) ribavirin triphosphate, however, supports catalysis with wild type p38, presumably through the *anti* conformation. Both of these facts are in contrast to the more rigid nucleotide, N^6-(benzyl) ATP. We therefore designed N^4-(benzyl) AICAR triphosphate (b), an analogue of N^4-(benzyl) ribavirin triphosphate that is restricted to the *syn* rotamer. Therefore, N^4-(benzyl) AICAR triphosphate is a nucleotide that should replicate the phosphodonor properties of N^6-(benzyl) ATP, yet share with N^4-(benzyl) ribavirin triphosphate the ability to bind a mutant kinase active site that excludes ATP.

triphosphate is not orthogonal to wild type p38. Furthermore, N^4-(benzyl) ribavirin triphosphate is not a competent substrate for T106G p38. This is the exact opposite of the structure–activity relationship we desired.

We hypothesized that N^4-(benzyl) ribavirin triphosphate failed to meet our design criteria because the amide bearing the benzyl substituent is conformationally flexible, and can adopt a conformation (*anti* in Fig. 4a) that does not clash with the T106 active site residue as we observe in the case of N^6-(benzyl) ATP.[14] Thus, N^4-(benzyl) ribavirin triphosphate is not orthogonal to wild type p38. The second problem is that N^4-(benzyl) ribavirin triphosphate is not a viable substrate for T106G p38. Although N^4-(benzyl) ribavirin triphosphate can also adopt a conformation similar to N^6-(benzyl) ATP (*syn* in Fig. 4a), we found it binds the pocket exposed by the I338G mutation in v-Src much more tightly than N^6-(benzyl) ATP does, and apparently in a non-productive orientation.[21] Previous studies on the base pairing preferences of deoxyribavirin incorporated into a DNA duplex utilizing NMR measurements as well as semi-empirical calculations of the amide conformation[26,27] indicate that rotation about the amide bond is extremely facile, and furthermore, there is no significant energetic difference between the *syn* and *anti* rotamers.

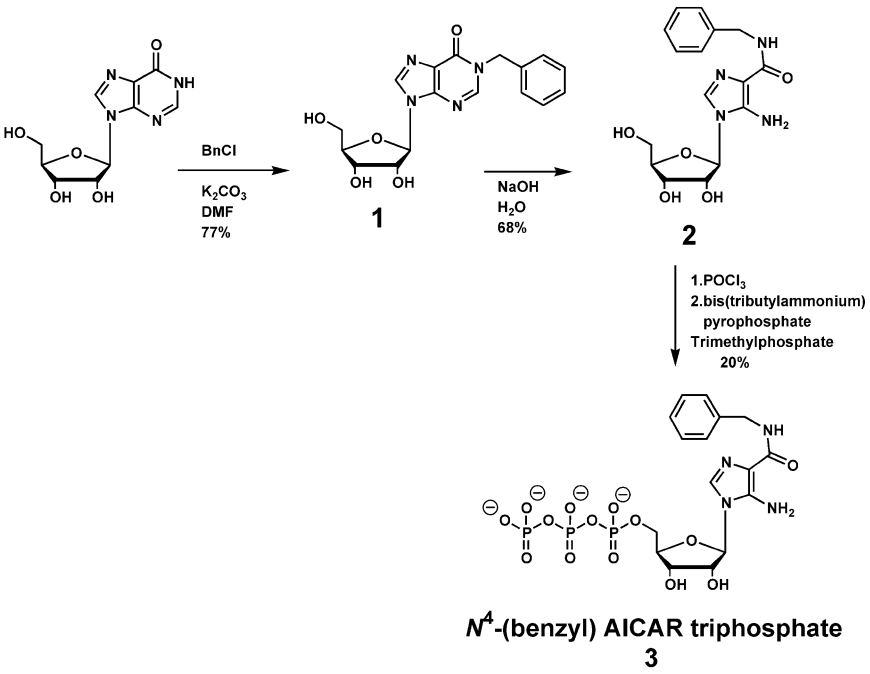
To remedy these issues, we designed an analogue of N^4-(benzyl) ribavirin triphosphate that is a benzyl substituted derivative of AICAR triphosphate (5-aminoimidazole-4-carboxamide ribotide), an intermediate in purine biosynthesis (Fig. 2).[28] N^4-(benzyl) AICAR triphosphate was specifically designed to conformationally restrain the benzyl substituent and properly orient it towards the T106G pocket, rendering it orthogonal to wild type p38 and a competent substrate for T106G p38 (Fig. 1c). The 5-amino group, present in AICAR and absent in ribavirin, should constrain the benzyl substituent to a conformation similar to N^6-(benzyl) ATP by steric hindrance as well as through an intramolecular hydrogen bond to the amide carbonyl group (Fig. 4b). The crystal structure of AICAR shows this hydrogen bond between the amide carbonyl and 5-amino group, supporting this idea.[29] Furthermore, studies on the mechanism of AICAR transformylase showed that N^4-(allyl) AICAR and N^4-(methyl) AICAR can bind enzyme active sites that require the *syn* rotamer, but not enzymes that require the *anti* rotamer (Fig. 4b).[30] Therefore, we believe that N^4-(benzyl) AICAR triphosphate should similarly adopt primarily the *syn* conformation shown in Fig. 4. This is the conformation that resembles N^6-(benzyl) ATP, and should make N^4-(benzyl) AICAR triphosphate orthogonal to wild type kinases and a substrate of kinases with a mutation corresponding to T106G in p38. However, N^4-(benzyl) AICAR triphosphate is similar to N^4-(benzyl) ribavirin triphosphate in that it may accommodate mutations that block ATP binding, our ultimate design goal.[21]

We synthesized the AICAR nucleoside by benzylation of inosine[31] followed by deformylation in base[32] to yield compound **2**. Addition of the triphosphate moiety was carried out according to the procedure of Ludwig[33] (Scheme 1).[34]

We then evaluated N^4-(benzyl) AICAR triphosphate as a phosphodonor for wild type p38 and T106G p38 (Fig. 5) using the same assay used for N^4-(benzyl) ribavirin triphosphate described above. The result of this experiment indicates that the 5-amino group of N^4-(benzyl) AICAR triphosphate has a profound effect on the phosphodonor properties with respect to wild type p38 and T106G p38. N^4-(benzyl) AICAR triphosphate is not utilized by wild type p38 (Fig. 5, Row 1), yet is a substrate for T106G p38 (Fig. 5, Row 2). Densitometry measurements of the relative amounts of phospho-ATF2 indicate that ATP is 12-fold more efficient than N^4-(benzyl) AICAR triphosphate as a T106G p38 substrate.[§] This result confirms our hypothesis that constraining the conformation of the benzyl substituent is a key factor in generating nucleotides that are: (1) orthogonal to wild type kinases and (2) competent substrates for mutant kinases such as I338G v-Src and T106G p38.

In conclusion, we have demonstrated that N^4-(benzyl) AICAR triphosphate is orthogonal to wild type kinases and is a substrate of T106G p38, although it is utilized less efficiently than ATP. This structure–activity relationship is critically dependent on the conformational rigidity of N^4-(benzyl) AICAR triphosphate. In contrast, N^4-(benzyl) ribavirin triphosphate, a related nucleotide that is conformationally mobile, is neither orthogonal to wild type kinases nor a substrate of

[§]Measured on an Alpha Innotech Chemimager 5500.



Scheme 1. Synthesis of N^4-(benzyl) AICAR triphosphate.

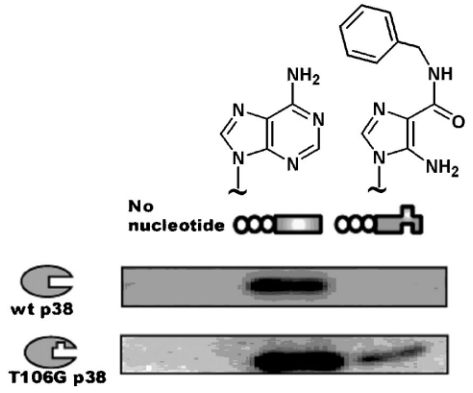


Figure 5. Wild type p38 (Row 1) and T106G p38 (Row 2) kinase assay with no nucleotide (Column 1), ATP (Column 2), and N^4-(benzyl) AICAR triphosphate (Column 3), using a western blot assay which detects phospho-ATF-2, a substrate of p38.

kinases with mutations corresponding to T106G in p38. N^4-(benzyl) AICAR triphosphate is a good candidate substrate for a kinase which is insensitive to ATP because it can potentially bind a kinase active site that sterically excludes ATP. Thus, N^4-(benzyl) AICAR triphosphate represents the first step in our ultimate goal of designing a unique nucleotide triphosphate/protein kinase pair. This will extend the scope of experiments designed to label the direct substrates of protein kinases to include most relevant context, the intact cell.

Acknowledgements

This work was supported by the National Institutes of Health (CA70031) and the National Science Foundation (MCB-9874587). We thank Sasha Buzko for assistance in computational conformational analysis, and members of the Shokat laboratory for critical reading of the manuscript.

References and Notes

1. Pawson, T.; Nash, P. *Genes. Dev.* **2000**, *14*, 1027.
2. Hunter, T. *Cell* **2000**, *100*, 113.
3. Cohen, P. *Trends. Biochem. Sci.* **2000**, *25*, 596.
4. Songyang, Z.; Blechner, S.; Hoagland, N.; Hoekstra, M. F.; Piwnica-Worms, H.; Cantley, L. C. *Curr. Biol.* **1994**, *4*, 973.
5. Songyang, Z.; Lu, K. P.; Kwon, Y. T.; Tsai, L. H.; Filhol, O.; Cochet, C.; Brickey, D. A.; Soderling, T. R.; Bartleson, C.; Graves, D. J.; DeMaggio, A. J.; Hoekstra, M. F.; Blenis, J.; Hunter, T.; Cantley, L. C. *Mol. Cell. Biol.* **1996**, *16*, 6486.
6. Songyang, Z.; Cantley, L. C. *Methods Mol. Biol.* **1998**, *87*, 87.
7. Obata, T.; Yaffe, M. B.; Leparc, G. G.; Piro, E. T.; Maegawa, H.; Kashiwagi, A.; Kikkawa, R.; Cantley, L. C. *J. Biol. Chem.* **2000**, *275*, 36108.
8. Hawkins, J.; Zheng, S.; Frantz, B.; LoGrasso, P. *Arch. Biochem. Biophys.* **2000**, *382*, 310.
9. Fukunaga, R.; Hunter, T. *Embo. J.* **1997**, *16*, 1921.
10. Lock, P.; Abram, C. L.; Gibson, T.; Courtneidge, S. A. *EMBO J.* **1998**, *17*, 4346.
11. Alaimo, P. J.; Shogren-Knaak, M. A.; Shokat, K. M. *Curr. Opin. Chem. Biol.* **2001**, *5*, 360.
12. Shogren-Knaak, M. A.; Alaimo, P. J.; Shokat, K. M. *Annu. Rev. Cell. Dev. Biol.* **2001**, *17*, 405.
13. Bishop, A.; Buzko, O.; Heyeck-Dumas, S.; Jung, I.; Kraybill, B.; Liu, Y.; Shah, K.; Ulrich, S.; Witucki, L.; Yang, F.; Zhang, C.; Shokat, K. M. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 577.
14. Witucki, L. A.; Huang, X.; Shah, K.; Liu, Y.; Kyin, S.; Eck, M. J.; Shokat, K. M. *Chem. Biol.* **2002**, *9*, 25.
15. Shah, K.; Shokat, K. M. *Chem. Biol.* **2002**, *9*, 35.

16. Habelhah, H.; Shah, K.; Huang, L.; Burlingame, A. L.; Shokat, K. M.; Ronai, Z. *J. Biol. Chem.* **2001**, *276*, 18090.
17. Habelhah, H.; Shah, K.; Huang, L.; Ostareck-Lederer, A.; Burlingame, A. L.; Shokat, K. M.; Hentze, M. W.; Ronai, Z. *Nat. Cell. Biol.* **2001**, *3*, 325.
18. Chitu, V.; Demydenko, D.; Toth, G. K.; Hegedus, Z.; Monostori, E. *Biotechniques* **1999**, *27*, 435.
19. Liu, Y.; Shah, K.; Yang, F.; Witucki, L.; Shokat, K. M. *Chem. Biol.* **1998**, *5*, 91.
20. Shah, K.; Liu, Y.; Deirmengian, C.; Shokat, K. M. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 3565.
21. Ulrich, S. M.; Buzko, O.; Shah, K.; Shokat, K. M. *Tetrahedron* **2000**, *56*, 9495.
22. Sidwell, R. W.; Robins, R. K.; Hillyard, I. W. *Pharmacol. Ther.* **1979**, *6*, 123.
23. English, J. M.; Cobb, M. H. *Trends Pharmacol. Sci.* **2002**, *23*, 40.
24. Khokhlatchev, A.; Xu, S.; English, J.; Wu, P.; Schaefer, E.; Cobb, M. H. *J. Biol. Chem.* **1997**, *272*, 11057.
25. Gupta, S.; Campbell, D.; Derijard, B.; Davis, R. J. *Science* **1995**, *267*, 389.
26. Klewer, D. A.; Zhang, P.; Bergstrom, D. E.; Davisson, V. J.; LiWang, A. C. *Biochemistry* **2001**, *40*, 1518.
27. Zhang, P.; Johnson, W. T.; Klewer, D.; Paul, N.; Hoops, G.; Davisson, V. J.; Bergstrom, D. E. *Nucleic Acids Res.* **1998**, *26*, 2208.
28. Voet, D.; Voet, J. G. *Biochemistry*; 2nd ed.; Wiley: New York, 1995.
29. Adamiak, D. A.; Saenger, W. *Acta Crystallogr.* **1979**, *B35*, 924.
30. Wall, M.; Shim, J. H.; Benkovic, S. J. *Biochemistry* **2000**, *39*, 11303.
31. Szarek, W. A.; Pinto, B. M.; Iwakawa, M. *Can. J. Chem.* **1985**, *63*, 2149.
32. Fujii, T.; Saito, T.; Hisata, H.; Shinbo, K. *Chem. Pharm. Bull.* **1990**, *38*, 3326.
33. Ludwig, J. *Acta Biochim. Biophys. Acad. Hung.* **1981**, *16*, 131.
34. **Compound 2**: ^{1}H NMR (400 MHz, DMSO-d_6) δ 7.89 ppm (t, 1H, J = 6.4 Hz), 7.29 ppm (s, 1H), 7.23 ppm (m, 5H), 5.88 ppm (br s, 2H), 5.43 ppm (d, 1H, J = 6.4 Hz), 5.32 ppm (d, 1H, J = 6.5 Hz), 5.19 ppm (t, 1H, J = 4.8 Hz), 5.11 ppm (d, 1H, J = 4.4 Hz), 4.33 ppm (d, 2H, J = 6.4 Hz), 4.24 ppm (m, 1H), 4.00 ppm (m, 1H), 3.85 ppm (m, 1H), 3.54 ppm (m, 2H).

^{13}C NMR (400 MHz, DMSO-d_6) δ 165.22, 143.27, 141.29, 129.42, 128.79, 127.81, 127.12, 113.33, 88.32, 86.00, 73.55, 70.90, 61.78, 41.98 ppm. HRMS calcd for $C_{16}H_{20}N_4O_5$: 348.1434. Found 348.1436.

Compound 3: ^{31}P NMR (161 MHz, D_2O) δ = −5.7 ppm (d, 1P, J = 19.61 Hz), −10.80 ppm (d, 1P, J = 19.81 Hz), −21.94 ppm (vt, 1P, J = 19.62 Hz). MS (FAB+) calcd for $C_{16}H_{23}N_4O_{14}P_3$: 588. Found 588, 610 (M−H+Na).