Bioorganic & Medicinal Chemistry Letters 18 (2008) 4554–4557

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

Crystal structure of chemically synthesized HIV-1 protease and a ketomethylene isostere inhibitor based on the p2/NC cleavage site

Vladimir Yu. Torbeev [a,b,*], Kalyaneswar Mandal [a,c], Valentina A. Terechko [a,c], Stephen B. H. Kent [a,b,c]

[a] Institute for Biophysical Dynamics, The University of Chicago, 929 East 57th Street, Chicago, IL 60637, USA
[b] Department of Chemistry, The University of Chicago, 929 East 57th Street, Chicago, IL 60637, USA
[c] Department of Biochemistry & Molecular Biology, The University of Chicago, 929 East 57th Street, Chicago, IL 60637, USA

ARTICLE INFO

Article history:
Received 16 June 2008
Revised 8 July 2008
Accepted 10 July 2008
Available online 15 July 2008

Keywords:
HIV-1 protease
Ketomethylene isostere
Aspartic proteases
Chemical protein synthesis

ABSTRACT

Here we report the X-ray structures of chemically synthesized HIV-1 protease and the inactive [D25N]HIV-1 protease complexed with the ketomethylene isostere inhibitor Ac–Thr–Ile–Nleψ[CO–CH_2]Nle–Gln–Arg.amide at 1.4 and 1.8 Å resolution, respectively. In complex with the active enzyme, the keto-group was found to be converted into the hydrated gem-diol, while the structure of the complex with the inactive D25N enzyme revealed an intact keto-group. These data support the general acid–general base mechanism for HIV-1 protease catalysis.

Three distinctly different chemical mechanisms have been proposed for catalysis of peptide bond cleavage by aspartic proteases. In the first one, a nucleophilic aspartic acid side chain carboxylate attacks the carbonyl-group of the peptide bond, forming a covalent enzyme-substrate tetrahedral intermediate, followed by expulsion of the amine component.[1] Second, is the general acid–general base mechanism, where one catalytic aspartate side chain carboxylate (COO^-) acts as a general base to remove a proton from the water molecule nucleophile, while another aspartic acid side chain carboxyl (COOH) general acid donates a proton to the carbonyl oxygen atom of the scissile peptide bond.[2,3] In the third 'kinetic isomechanism', a 10-membered cyclic structure is formed, involving the two aspartic acid side chain carboxyl groups (COO^-), with a proton between them, and the water molecule nucleophile; this mechanism allows for energy-inexpensive proton shuffling within the cyclic structure along the reaction coordinate.[4] The last two mechanisms also invoke a low-barrier hydrogen bond (LBHB); in the general acid–general base mechanism the LBHB would stabilize the transition state,[5,6] while in the kinetic isomechanism it allows for hydrogen tunneling.[4]

X-ray structures of aspartic protease enzymes cocrystallized with inhibitors that are available in the Protein Data Bank do not contain the lytic water molecule. Thus, the hypothesis that the lytic water is not initially present in the active site and that catalysis occurs via a covalent enzyme-substrate tetrahedral intermediate cannot be ruled out. A ketomethylene isostere, in which the scissile peptide bond [C(O)NH] is substituted by a ψ[$C(O)CH_2$] linkage, is the most suitable substrate surrogate to test the 'covalent intermediate' hypothesis. If a water molecule is initially present in the active site, then the active enzyme should catalyze the hydration of the keto group to form the gem-diol; on the other hand, direct nucleophilic attack by an ionized aspartic acid side chain would lead to a covalent adduct with the inhibitor that would be readily observed by X-ray crystallography.

Chemical synthesis of substrate-derived ketomethylene inhibitors and their tight binding to HIV-1 protease was reported previously (IC_{50} down to 4.6 nM).[7] We have reproduced the chemical synthesis for the selected ketomethylene isostere (IC_{50} 6.3 nM,[7] mimicking the p2/NC cleavage site) and cocrystallized the resulting substrate-derived inhibitor with wild-type HIV-1 protease (based on the SF2 isolate) and with the inactive [D25N]HIV-1 protease analogue.

Total chemical synthesis of the HIV-1 protease[8] was based on a two segment native chemical ligation (Fig. 1a).[9] Both segments were prepared by in situ neutralization Boc chemistry SPPS,[10] and after ligation the Cys41 residue at the ligation site was alkylated with 2-bromoacetamide to form a *ψ-homo*-Gln41 residue. After removal of the formyl groups from Trp6 and Trp42, the (1-99)-polypeptide was purified by reverse-phase HPLC and folded by two-step dialysis to form fully active enzyme (k_{cat} 23.4 ± 0.4 s^{-1}, K_m 25.1 ± 1.2 μM). The inactive [D25N]HIV-1 protease analogue was synthesized according to the same strategy.

* Corresponding author. Fax: +1 773 702 0439.
E-mail address: torbeev@uchicago.edu (V.Yu. Torbeev).

0960-894X/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2008.07.039

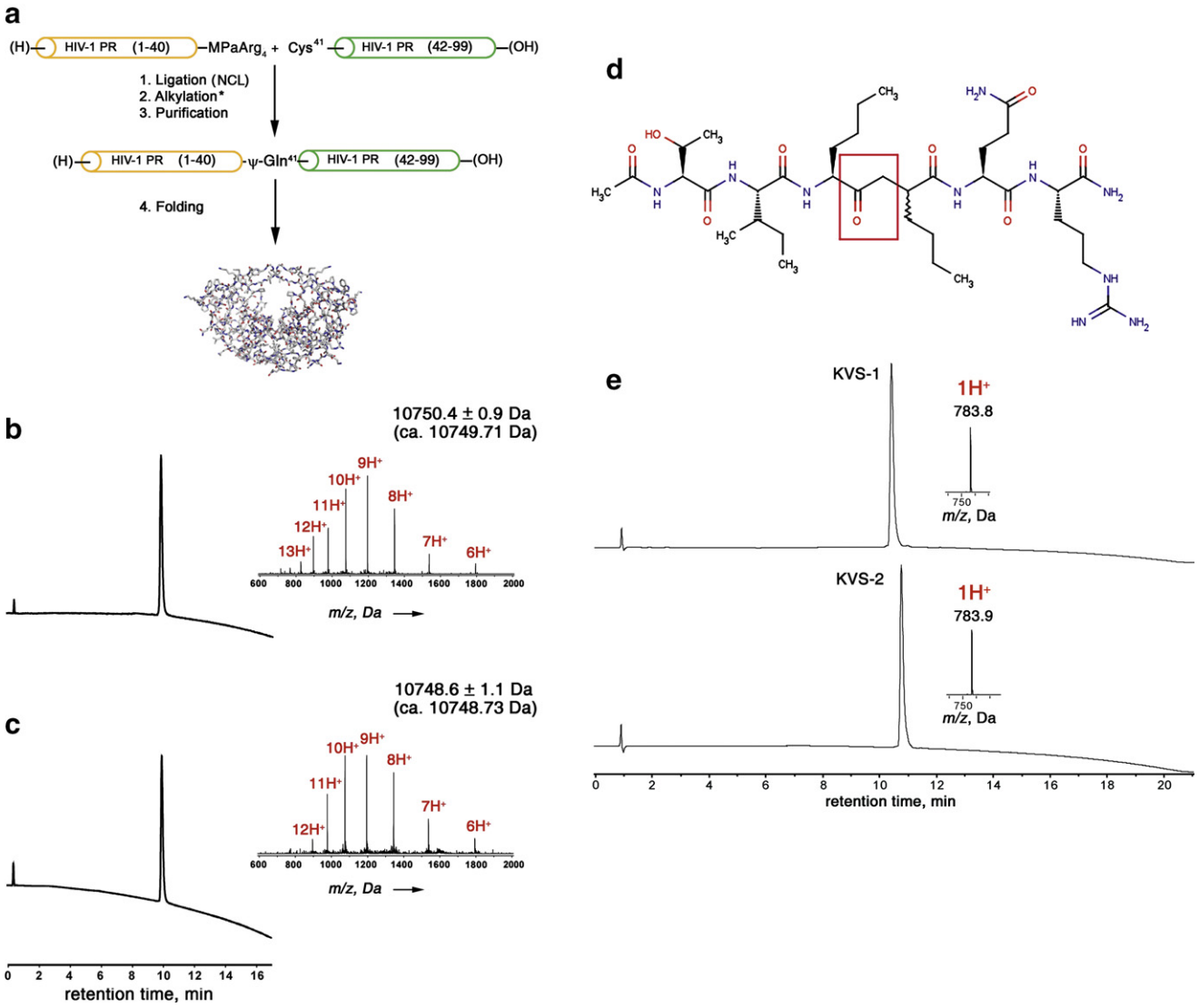

Figure 1. (a) Total chemical synthesis of wild-type HIV-1 PR and the [D25N]HIV-1 PR analogue. Native chemical ligation of two unprotected synthetic peptides was followed by alkylation with 2-bromoacetamide to convert Cys41 to *ψ-homo*-Gln41. After deformylation of Trp6 and Trp42 and reverse phase HPLC purification, the (1-99)-polypeptide was folded by two-step dialysis against 10 mM NaOAc, pH 5.6. $MPaArg_4$ = 3-mercaptopropionate tetraarginine.amide (i.e. the thioester leaving group). Analytical HPLC traces (λ = 214 nm) and ESI-MS of purified wild-type HIV-1 protease (b) and the [D25N]HIV-1 PR analogue (c). In (d), the structure of the ketomethylene isostere inhibitor is shown. After the synthesis it exists as a mixture of diastereomers (with respect to the carbon atom corresponding to the α-carbon of Nle4), which were separated by reverse phase HPLC. (e) Analytical HPLC traces (λ = 214 nm) and ESI-MS of the (*R*)-'Nle4' diastereomer (KVS-1) and the (*S*)-'Nle4' diastereomer (KVS-2). See Supplementary Data for more information.

Chemical synthesis of the ketomethylene isostere Ac–Thr–Ile–Nleψ[CO–CH_2]Nle–Gln–Arg.amide (Fig. 1d) gave two diastereomers (abbreviated as KVS-1 and KVS-2), differing in their stereochemistry at the tertiary carbon corresponding to the α-carbon of the Nle4 residue. The diastereomers could be readily separated by reverse phase HPLC on C18, 10 × 250 mm column (see Fig. 1e; and Supplementary Data), and their stereochemical identity was inferred on the basis of their cocrystallization behavior with HIV-1 protease. For the KVS-1 diastereomer, crystals could be obtained within 1–2 days using standard crystallization conditions[11] developed for HIV-1 protease/inhibitor complexes.

Use of a ketone functionality in aspartic protease inhibitors was originally suggested for ketone analogues of statines (statones) to mimic the presumed tetrahedral intermediate.[12,13] It was further discovered that statones are weaker binders than their hydroxy counterparts; this was attributed to the unfavorable equilibrium for hydration of the ketone function.[14] The [$C(O)CH_2$] moiety was thus substituted by 2,2-difluorostatone moiety [$C(O)CF_2$] which is readily hydrated in water to give the gem-diol.[15] The 2,2-difluorostatones were found to be 50- to 1000-fold better inhibitors than non-fluorinated analogues.[16] In contrast to statones, peptide substrate-derived non-fluorinated ketomethylene isosteres are 5-10 times more potent inhibitors of HIV-1 protease than their hydroxyethyl counterparts.[7]

In the work reported here, we have determined high resolution X-ray structures for a substrate-derived ketomethylene isostere complexed with fully active ('wild type') HIV-1 protease, and for the complex of the same ketomethylene inhibitor with inactive [D25N]HIV-1 protease (Table 1). In the complex with inactive [D25N]HIV-1 protease, the ketone group was found to be intact indicating that the ketone exists predominantly in the non-hydrated form in water solution (see Fig. 2a and c), which is in agreement with previous studies of the variety of carbonyl compounds.[17] The complex with the wild type HIV-1 protease, however, clearly demonstrated the tetrahedral geometry of an hydrated ketone group (Fig. 2b and d). This result unambiguously demonstrates that hydration of the ketomethylene isostere is due to the first step of the proteolytic catalytic mechanism of HIV-1 protease.

Table 1
Crystal parameters, data collection, and refinement statistics

Complex	[D25N]HIV-1 PR + KVS-1	wt HIV-1 PR + KVS-1
Data collection		
Beamline	APS 23ID-D	APS 23ID-B
Wavelength, Å	0.97934	1.03320
Space group	$P\,2_12_12_1$	$P\,2_12_12_1$
Cell dimensions		
a (Å)	51.028	51.200
b (Å)	58.551	58.077
c (Å)	61.584	61.658
$\alpha = \beta = \gamma$ (°)	90.00	90.00
Resolution (Å)	50.00–1.80	50.00–1.40
R_{merge}	0.089 (0.590)[a]	0.084(0.384)
$I/\sigma I$	31.3 (3.9)	23.9 (6.3)
Redundancy	7.8	6.8
Refinement		
Resolution (Å)	20.00 – 1.80	20.00 – 1.40
Completeness (%)	98.61	99.66
No. reflections		
Work/free set	16491/882	35053/1846
R_{work}/R_{free}	0.194/0.223	0.204/0.215
No. protein atoms	1518	1518
No. inhibitor atoms	55	56
No. water atoms	83	113
Average B-factor (Å^2)	24.05	11.995
R.m.s. deviations		
Bond lengths (Å)	0.017	0.017
Bond angles (°)	1.595	1.684
PDB ID	3DCK	3DCR

[a] Highest resolution is in parentheses.

Recently, Kumar et al.[18] have been able to directly view by X-ray crystallography an actual tetrahedral intermediate by soaking peptide substrate into a crystal of apo-HIV-1 protease. In the crystal, the enzyme is locked in a 'closed' conformation and, presumably, this inflexibility prevents dissociation of the tetrahedral intermediate to release products. With an alternative approach, Kovalevsky et al.[19] successfully crystallized the tetrahedral intermediate by incorporating Thr at the P1′ position, which is known to greatly diminish the ability of protease to hydrolyze the peptide bond. Figure 3 depicts the superposition of the active site region in the HIV-1 protease and hydrated KVS-1 (this work), and tetrahedral intermediates for wild-type and I54V mutants, determined at 1.46 and 1.50 Å resolution, respectively. Although the substrate peptide sequences are not the same, the methylene unit serves as a very natural replacement for NH-moiety. A number of short H-bonds were observed in the complex of wild-type HIV-1 protease with KVS-1 inhibitor (Fig. 2d). However, whereas in the structures of actual tetrahedral intermediates very short, strong hydrogen bond has been observed between one of the aspartates and one hydroxyl group of the gem-diol moiety (2.3 Å in PDB ID:3B7V and 2.5 Å PDB ID:3B80),[19] in the structure with KVS-1 the identical H-bond is loosened to 2.7 Å (Fig. 2d).

Our present data do not support the 'covalent intermediate' mechanism of catalysis in HIV-1 protease; if direct attack had taken place one would expect for this inhibitor (KVS-1) to be covalently bound to one of the aspartates. On the other hand, the structure of the inhibitor mimics a natural tetrahedral intermediate very closely

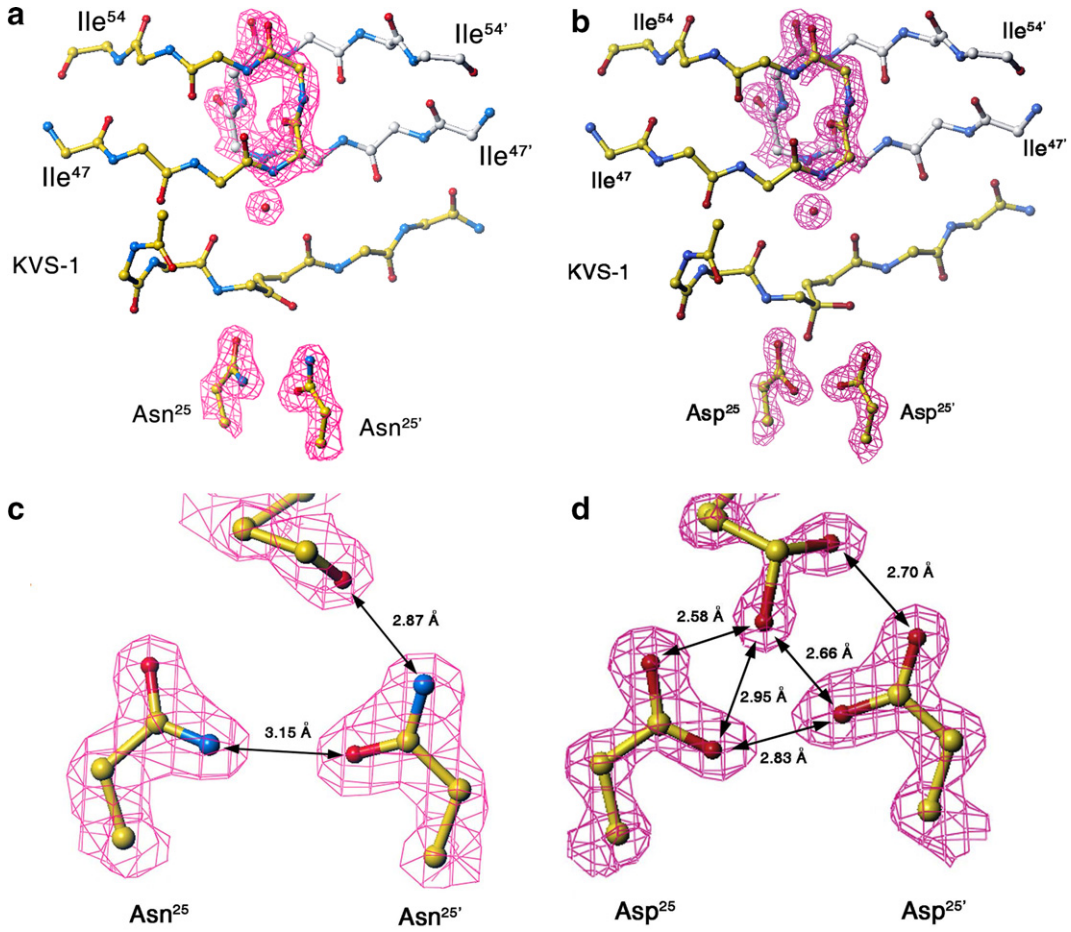

Figure 2. Active site region in complexes of: (a) inactive D25N chemical analogue of HIV-1 PR and intact keto-form of KVS-1 inhibitor; and, (b) wild type HIV-1 PR and hydrated 'gem-diol' KVS-1 inhibitor. (c) Residues Asn^{25} and Asn$^{25'}$ hydrogen-bonded to the ketomethylene isostere in the inactive complex; and, (d) catalytic Asp^{25} and Asp$^{25'}$ hydrogen-bonded to the gem-diol, in the wild type complex. The |2Fo-Fc| electron density map for selected residues was contoured at 2.0σ (in a and b) and 1.8σ (in c and d).

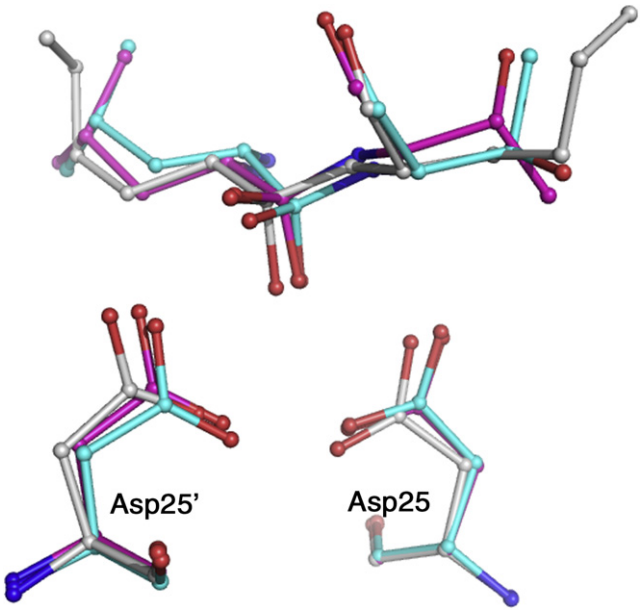

Figure 3. Superposition of the active site regions of wild-type HIV-1 protease complexed with KVS-1 inhibitor (gray) onto structures of tetrahedral peptide intermediates of wild-type HIV-1 protease (PDB ID:3B7V, in cyan) and I54V mutant (PDB ID:3B80, in magenta).

and hence may be of use in advanced studies by X-ray crystallography and neutron diffraction to elucidate the nature of short, strong hydrogen bonds in the active sites of aspartic acid proteases.

Acknowledgments

This research was supported by the Office of Science (BER), U.S. Department of Energy, Grant No. DE-FG02-07ER64501. Use of the Advanced Photon Source (Argonne, IL) was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under contract No. DE-AC02-06CH11357. GM/CA CAT has been funded in whole or in part with Federal funds from the National Cancer Institute (Y1-CO-1020) and the National Institute of General Medical Science (Y1-GM-1104).

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmcl.2008.07.039.

References and notes

1. Dunn, B. M.; Fink, A. L. *Biochemistry* **1984**, *23*, 5241.
2. Hyland, L. J.; Tomaszek, T. A.; Meek, T. D. *Biochemistry* **1991**, *30*, 8454.
3. Veerapandian, B.; Cooper, J. B.; Šali, A.; Blundell, T. L.; Rosati, R. L.; Dominy, B. W.; Damon, D. B.; Hoover, D. J. *Protein Sci.* **1992**, *1*, 322.
4. Northrop, D. B. *Acc. Chem. Res.* **2001**, *34*, 790.
5. Trylska, J.; Crochowski, P.; McCammon, J. A. *Protein Sci.* **2004**, *13*, 513.
6. Coates, L.; Tuan, H.-T.; Tomanicek, S.; Kovalevsky, A.; Mustyakimov, M.; Erskine, P.; Cooper, J. *J. Am. Chem. Soc.* **2008**, *130*, 7235.
7. Marinier, A.; Toth, M. V.; Houseman, K.; Mueller, R.; Marshall, G. R. *Bioorg. Med. Chem.* **1994**, *2*, 919.
8. PQITLWKRPL[10]VTIRIGGQLK[20]EALL*Xaa*TGADD[30]TVIEE*Nle*NLPG[40]*ψ-Gln*WKPK*Nle* IGGI[50]GGFIKVRQYD [60]QIPVEI*Abu*GHK[70]AIGTVLVGPT[80]PVNIIGRNLL[90]TQIG*Abu* TLNF[99]. Unnatural amino acids are in italics in three-letter code. *Nle* = norleucine, *Abu* = α-aminobutyric acid, *ψ-Gln* = pseudo-homoglutamine, *Xaa* = either Asp or Asn. Ligation site is underlined.
9. Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776.
10. Schnolzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. H. *Int. J. Pept. Res. Therap.* **2007**, *13*, 31.
11. For crystallization conditions and structure solution and refinement see Supplementary Data.
12. Rich, D. H.; Bernatowicz, M. S.; Schmidt, P. G. *J. Am. Chem. Soc.* **1982**, *104*, 3535.
13. Holladay, M. W.; Salituro, F. G.; Rich, D. H. *J. Med. Chem.* **1987**, *30*, 374.
14. Rich, D. H.; Bernatowicz, M. S.; Agarwal, N. S.; Kawai, M.; Saituro, F. G. *Biochemistry* **1985**, *24*, 3165.
15. Gelb, M. H.; Svaren, J. P.; Abeles, R. H. *Biochemistry* **1985**, *24*, 1813.
16. Sani, M.; Sinisi, R.; Viani, T. *Curr. Top. Med. Chem.* **2006**, *6*, 1545.
17. Guthrie, J. P. *Acc. Chem. Res.* **1983**, *16*, 122.
18. Kumar, M.; Prashar, V.; Mahale, S.; Hosur, M. V. *Biochem. J.* **2005**, *389*, 365.
19. Kovalevsky, A. Y.; Chumanevich, A. A.; Liu, F.; Louis, J. M.; Weber, I. T. *Biochemistry* **2007**, *46*, 14854.