Tetrahedron Letters, Vol. 33, No. 31, pp. 4453-4456, 1992
Printed in Great Britain

0040-4039/92 $5.00 + .00
Pergamon Press Ltd

SYNTHESIS OF A POTENTIAL "SUICIDE SUBSTRATE" OF THE HIV-1 PROTEASE, INCORPORATING A 3-ACETOXY-Δ-4,5-(L)-PIPECOLIC ACID AS PROLINE SUBSTITUTE.

Jean-Paul Mazaleyrat*, Isabelle Rage, Juan Xie, Jaroslav Savrda and Michel Wakselman
CNRS-CERCOA, 2, rue Henri Dunant, F-94320 Thiais, France

Summary: Coupling of (D,L)-baikiain methyl ester with Boc-(L)-phenylalanine, followed by saponification, iodolactonization and aminolysis of the diastereoisomeric iodolactones by (L)-isoleucyl-(L)-valine methyl ester gave two diastereoisomeric peptides incorporating either a 3*S*-hydroxy-4*S*-iodo-(L)-pipecolic acid or a 3*R*-hydroxy-4*R*-iodo-(D)-pipecolic acid residue, which were separated by thin layer chromatography. Assignment of configuration was unambiguous when the synthesis was repeated with (L)-baikiain as starting material. All compounds exhibited conformational isomerism in their ^{1}H NMR spectra, attributed to the existence of both *s-cis* and *s-trans* configurations of the Phe-NR_2 peptide bond. Acetylation of each of the two separated diastereoisomers gave Boc-(L)-Phe-3*S*-OAc-4*S*-I-(L)-Pip-(L)-Ile-(L)-Val-OMe and Boc-(L)-Phe-3*R*-OAc-4*R*-I-(D)-Pip-(L)-Ile-(L)-Val-OMe. N-deprotection and coupling of the former with Boc-(L)-Ser-(L)-Ala-(L)-Ala-OH by the DCC/HOBt method or stepwise elongation of the peptidic chain, gave Boc-(L)-Ser-(L)-Ala-(L)-Ala-(L)-Phe-3*S*-OAc-4*S*-I-(L)-Pip-(L)-Ile-(L)-Val-OMe. Dehydroiodination of this compound on treatment with DBU gave Boc-(L)-Ser-(L)-Ala-(L)-Ala-(L)-Phe-3*S*-OAc-Δ-4,5-(L)-Pip-(L)-Ile-(L)-Val-OMe, a potential "suicide substrate" of the HIV-1 protease.

The HIV aspartic protease catalyzes the hydrolytic cleavage at P_1-P_1' (Schechter and Berger notation)[1] peptide bonds such as Phe-Leu, Leu-Ala, Leu-Phe, Met-Met and particularly Phe-Pro and Tyr-Pro (type I cleavages in the Henderson classification),[2] in the Pr_{55}^{gag} and $Pr_{160}^{gag\text{-}pol}$ polyproteins present in the immature virion. The proteins resulting from these cleavages are enzymes (the protease itself -autocatalytic process-, reverse transcriptase and endonuclease) and structural proteins of matrix (p17), capsid (p24) and nucleocapsid (p9-p6). Therefore, the HIV protease plays a crucial role in the process of virion maturation. As such it is a very important target in the aim of AIDS therapy, and a great deal of effort has been devoted to the elaboration of synthetic inhibitors of this enzyme. So far, mainly transition state analogs in which the P_1-P_1' peptide bond of substrates of the HIV protease has been replaced by hydroxyethylene isosteres, hydroxyethylamine analogs, statine or difluorostatine analogs, and many others, have been developed with success.[3-9] To date, the best "transition state analog" inhibitors *in vitro* have K_i's in the nanomolar range and less, but the *inherent reversible character* of such type of inhibition may question their long term efficiency *in vivo*.[9] Thus, in parallel with the currently searched structural optimizations of the known families of transition state analogs, the still wide open field of mechanism-based inactivation of the HIV protease is of the greatest interest. Inhibition by mechanism-based inactivators is irreversible *by principle* and, since the reactive function is unmasked only in the active site by the action of the target protease itself, it is expected to be highly selective.[10] From these considerations, and in spite of the obstacle represented by the possible fast diffusion of the reactive species in the case of an aspartic protease where no acyl-enzyme formation occurs, we have planed to synthesize potential "suicide substrates" of the HIV protease. Taking into account the known Phe-Pro site of cleavage in synthetic heptapeptide substrates of the HIV-1 protease, we designed substrates analogs of type **A** (Fig. 1) in which the proline P_1' residue is replaced by a cyclic α-ene-amino acid having a leaving group X in γ (relative to the nitrogen atom) allylic position, as good candidates for mechanism-based inactivation.

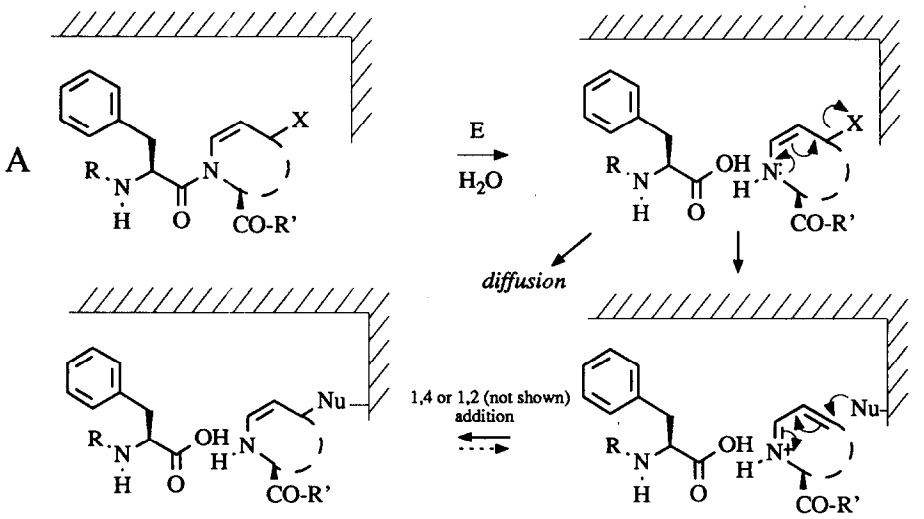
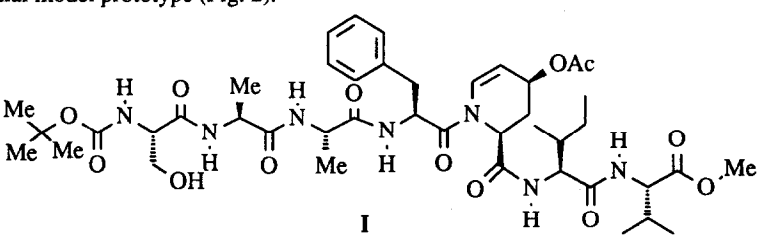


Figure 1

Indeed, if the HIV protease agrees such a substitution of the prolyl residue and cleaves the Phe-P_1' peptide bond, the initial γ-X-α-eneamide function becomes a γ-X-α-eneamine. Rapid elimination of X^- should result, followed by alkylation of an active-site nucleophile by the demasked α-eneimmonium,[10-12] hopefully before diffusion of the reactive species. The peptide **I**, in which the Ser-Gln-Asn-Phe-Pro-Ile-Val sequence of known synthetic substrates of the HIV-1 protease [14] is replaced by Ser-Ala-Ala-Phe-3*S*-acetoxy-Δ-4,5-(L)-Pip-Ile-Val, was chosen as an initial model prototype (Fig. 2).

Figure 2



I

Our choice of Ala(P_3)-Ala(P_2) instead of Gln-Asn is allowed, since it has been shown that a same change in Ac-Ser-P_3-P_2-Tyr-Pro-Val-Val-NH_2 resulted in a decrease of the k_{cat}/K_M constant by a factor of only 9.[8] As proline analogue, we considered a conveniently functionalized pipecolic acid P_1' residue,[15] because we knew from the work of Hanson and Russel [17] its possible access from baikiain aminoacid. Thus, racemic baikiain methyl ester (D,L)-**1** [18] (Fig. 3) was coupled with Boc-(L)-phenylalanine by the DCC/HOBt method to give the dipeptide **2** (54 %) [19] then, after saponification, **3** (89 %),[19] both obtained as a mixture of two diastereoisomers **2a** + **2b** and **3a** + **3b**, respectively. The mixture **3** was dissolved in 0.5 M aq. $NaHCO_3$ (3 equiv.) and iodolactonized [17] on treatment with an aqueous solution (3 ml water / mmol **3**) of I_2 (2 equiv.) and KI (6 equiv.), and standing at room temperature in the dark for 3 days. The resulting diastereoisomeric mixture of iodolactones **4** (58 %) [19] was treated with (L)-isoleucyl-(L)-valine methyl ester (3 equiv.) in dichloromethane, at room temperature for 12 h, to give the two diastereomers **5a** [19] and **5b**,[19] which were separated on preparative t.l.c. plates (SiO_2; CH_2Cl_2/MeOH 97:3) and obtained with 56 % and 32 % yield, respectively. No other spots than the ones corresponding to **5a** and **5b** were seen by t.l.c. of the crude reaction mixture, demonstrating that

displacement of iodine was not competing with the desired aminolysis of the lactone function. Acetylation of **5a** and **5b** (1.2 equiv. Ac_2O, 0.2 equiv. DMAP, CH_2Cl_2, rt, 12 h) gave **6a** (90 %) [19] and **6b** (93 %),[19] respectively.

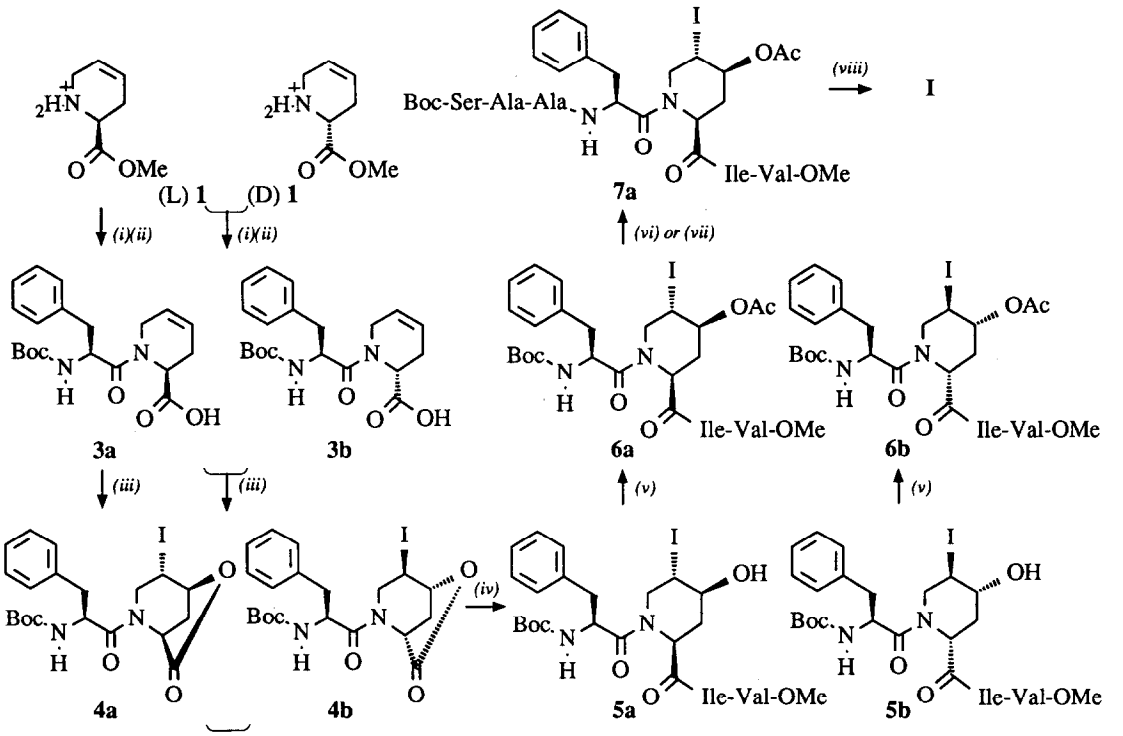


Figure 3. *(i)* Boc-(L)-Phe-OH / NMM / DCC / HOBt (**2a** + **2b** and **2a** isolated) *(ii)* 1. KOH / MeOH 2. H^+ *(iii)* $NaHCO_3$ / I_2 / KI / H_2O *(iv)* H-(L)-Ile-(L)-Val-OMe / CH_2Cl_2 *(v)* Ac_2O / DMAP cat / CH_2Cl_2 *(vi)* 1. HCl / AcOEt 2. Boc-(L)-Ala-OH / NMM / DCC / HOBt 3. HCl / AcOEt 4. Boc-(L)-Ala-OH / NMM / DCC / HOBt 5. HCl / AcOEt 6. Boc-(L)-Ser-OH / NMM / DCC / HOBt *(vii)* 1. HCl / AcOEt 2. Boc-(L)-Ser-(L)-Ala-(L)-Ala-OH / NMM / DCC / HOBt *(viii)* DBU / DMF / 85 °C / 0.5 h.

The ^1H NMR spectra (300 MHz) of all the diastereoisomeric mixtures **2a** + **2b**, **3a** + **3b** and **4a** + **4b** were inextricable. After separation of **5a** and **5b**, it became clear that each diastereomer could be present in solution as a mixture of conformers exchanging slowly on the NMR time scale, with either a *s-cis* or a *s-trans* configuration at the Phe-NR_2 bond. In $CDCl_3$, the compound **5b** exhibited a single conformer while **5a** existed in the form of two nearly equal percent conformers (*ca* 60/40). Surprisingly, acetylation of the 3*S* hydroxy function of **5a** was enough to result in a dramatic change in the ratio of the two conformers, which became *ca* 95/5 in **6a.** Similar complications of the ^1H NMR spectra of N-acyl-proline derivatives have led to misunderstanding of stereochemical outcomes, recently elucidated.[20] The problem of the absolute configuration of the isolated diastereoisomers **5a**, **5b** and **6a**, **6b** was solved by repetition of the synthesis starting from natural (L)-baikiain, isolated from Rhodesian teck wood, which was given to us by M. J. O. Anteunis.[21] Coupling of (L)-**1** with Boc-(L)-Phe-OH gave **2a** (61 %,[19] mixture of two conformers present in different ratios in $CDCl_3$, CD_3OD and DMSO-d_6) which, after saponification to **3a** (98 %,[19] mixture of two conformers) and iodolactonization, gave **4a** (61 %,[19] mixture of two conformers). Aminolysis of **4a** gave **5a** (86 %), then **6a** after acetylation, both having a strictly identical ^1H NMR spectrum in $CDCl_3$ (all sets of signals corresponding to both conformers of **5a** were

attributed) as the diastereoisomer **5a**, then **6a**, previously obtained from (D,L)-baikiain. In each of these reactions, a *single product* (by t.l.c.) was formed. Chain elongation of the deprotected peptide **6a** was performed either stepwise or by segment coupling with a pre-synthesized Boc-(L)-Ser-(L)-Ala-(L)-Ala-OH fragment. Both methods gave **7a** [19] with (6 %) and (27 %) overall yield, respectively. Treatment of **7a** with DBU in DMF at 85 °C for 0.5 h [17] gave **I** (68 %) [19] as single product.

Acknowledgement. This work was supported by Agence Nationale de Recherches sur le Sida (A.N.R.S.). We warmly thank Dr Anteunis for his generous gift of a sample of (L)-baikiain.

References and Notes

1. Schechter, I.; Berger, A., *Biochem. Biophys. Res. Comm.*, **1967**, *27*, 157-162
2. Henderson, L. E.; Benveniste, R. E.; Sowder, R.; Copeland, T. D.; Schultz, A. M.; Oroszlan, S., *J. Virol.*, **1988**, *62*, 2587-2595.
3. Jaskolski, M.; Tomasselli, A. G.; Sawyer, T. K.; Staples, D. G.; Heinrikson, R. L.; Schneider, J.; Kent, S. B. H.; Wlodawer, A., *Biochemistry*, **1991**, *30*, 1600-1609.
4. Hyland, L. J.; Tomaszek Jr, T. A.; Meek, T. D., *Biochemistry*, **1991**, *30*, 8454-8463.
5. Meek, T.D.; Lambert, D. M.; Metcalf, B. W.; Petteway Jr, S. R.; Dreyer, G. B., in E. De Clercq *"Design of anti-AIDS drugs"*, Elsevier, Amsterdam, 1990; pp 225-255.
6. Tomasselli, A. G.; Howe, W. J.; Sawyer, T.K.; Wlodaver, A.; Heinrikson, R. L., *Chem. Today*, **1991**, *9*, 6-27.
7. Erickson, J.; Neidhart, D. J.; Van Drie, J.; Kempf, D. J.; Chun Wang, X.; Norbeck, D. W.; Plattner, J. J.; Rittenhouse, J. W.; Turon, M.; Wideburg, M.; Kohlbrenner, W. E.; Simmer, R.; Helfrich, R.; Paul, D. A.; Knigge, R., *Science*, **1990**, *249*, 527-533.
8. Dreyer, G. B.; Metcalf, B. W.; Tomaszek Jr., T. A.; Carr, T. J.; Chandler III, A. C.; Hyland, L.; Fakhoury, S. A.; Magaard, V. W.; Moore, M. L.; Strickler, J. E.; Debouck, C.; Meek, T. D., *Proc. Natl. Acad. Sci. US.*, **1989**, *86*, 9752-9756.
9. Blundell, T. L.; Lapatto, R.; Wilderspin, A. F.; Hemmings, A. M.; Hobart, P. M.; Danley, D. E.; Whittle, P. J., *Trends Biochem. Sci.*, **1990**, *15*, 425-430.
10. Silverman, R. B., *Mechanism-based Enzyme Inactivation: Chemistry and Enzymology*", CRC Press, Boca Raton, Florida, USA, 1988, Vol I.
11. Navia, M. A.; Springer, J.P.; Lin, T. Y.; Williams, H. R.; Firestone, R. A.; Pisano, J. M.; Doherty, J. B.; Finke, P. E.; Hoogsteen, K., *Nature*, **1987**, *327*, 79-82.
12. A similar mechanism has been postulated for the inactivation of serine proteinases by cyclopeptides incorporating a latent quinonimmonium methide electrophile.[13]
13. Wakselman, M.; Mazaleyrat, J. P.; Xie, J.; Montagne, J. J.; Vilain, A. C.; Reboud-Ravaux M., *Eur. J. Med. Chem.*, **1991**, *26*, 699-707.
14. Billich, A.; Winkler, G., *Archiv. Biochem. Biophys.*, **1991**, *290*, 186-190.
15. Substitution of proline by pipecolic acid in a heptapeptide substrate of the HIV protease, has been shown to result in competitive inhibition of this enzyme,[16] but the behaviour of peptides incorporating a Δ-4,5-(L)-pipecolyl or a Δ-3,4-(L)-pipecolyl (baikiain) P_1' residue is not known.
16. Copeland, T. D.; Wondrak, E. M.; Tozser, J.; Roberts, M. M.; Oroszlan, S., *Biochem. Biophys. Res. Comm.*, **1990**, *169*, 310-314.
17. Hanson, G. J.; Russel, M. A., *Tetrahedron Letters*, **1989**, *30*, 5751-5754.
18. Burgstahler, A. W.; Aiman, C. E., *J. Org. Chem.*, **1960**, *25*, 489-492.
19. All compounds gave satisfactory analytical data (^{1}H NMR at 300 MHz, C,H,N analysis and/or FAB spectroscopy). The complete details of synthesis and NMR work will be published elsewhere in a full account of this study. The given yields correspond to isolated pure compounds. **I**: $[\alpha]^{25}$(c 0.15 ; MeOH) = - 121 (546 nm). MS (FAB$^+$): 888 $(MH)^+$. ^{1}H NMR (300 MHz; CD_3OD) (a single conformer present): 7.24, m, 5H (ArH Phe); 6.95, d (7.9), 1H (Hε Pip); 5.18, m, 5.15, m and 5.10, m, 3H (respectively Hα Phe, Hγ Pip and Hδ Pip); 5.00, m, 1H (Hα Pip); 4.37, d (8.1), 1H (Hα Ile); 4.27, d (6.2), 1H (Hα Val); 4.18, m, 2H (Hα Ala^1 and Hα Ala^2); 4.14, m, 1H (Hα Ser); 3.83, dd (5.4; 10.8), 1H (Hβ Ser); 3.72, dd (partly masked), 1H (Hβ Ser); 3.71, s, 3H (OMe); 3.24, dd (5.0; 13.9), 1H (Hβ Phe); 2.93, dd (9.7; 13.9), 1H (Hβ Phe); 2.64, d (14.6), 1H (Hβ Pip); 2.19, m (partly masked), 1H (Hβ Pip); 2.12, m, 1H (Hβ Val); 1.91, s, 3H (OAc); 1.87, m, 1H (Hβ Ile); 1.60, m, 1H (Hγ Ile); 1.47, s, 9H (Boc); 1.35, d (7.3), 3H (CH_3 Ala^1); 1.16, d (7.3) and m (masked), 4H (CH_3 Ala^2 and Hγ Ile); 0.96, m and 0.92, m (t-like), 12H (2 $CH_3\gamma$ Val, $CH_3\gamma$ Ile and $CH_3\delta$ Ile).
20. Avent, G. A.; Bowler, A. N.; Doyle, P. M.; Marchand, C. M.; Young, D. W., *Tetrahedron Letters*, **1992**, *33*, 1509-1512.
21. Callens, R.E.A.; Anteunis, M. J. O.; Reyniers, F., *Bull. Soc. Chim. Belg.*, **1982**, *91*, 713-723.

(Received in France 21 April 1992)