Pergamon

Tetrahedron, Vol. 53, No. 12, pp. 4231-4238, 1997

0040-4020/97 $17.00 + 0.00

PII: S0040-4020(97)00150-6

NMR and Molecular Modeling Characterization of HIV-1 Protease Inhibitors: Difunctional Enols of N-Protected Phenylalanine

Yvan Boulanger*, Alain Larocque, Abdesslem Khiat,

INRS-Santé, Université du Québec, 245 boul. Hymus, Pointe-Claire, Québec, Canada H9R 1G6;

François Deschamps and Gilles Sauvé

Institut Armand-Frappier, Université du Québec, 531 boul. des Prairies, Laval, Québec, Canada H7N 4Z3

Abstract: The three-dimensional structure of five HIV protease inhibitors of the N-*tert*-butoxycarbonylphenylalanyl enol family have been investigated by NMR and molecular modeling. Complete ^{1}H and ^{13}C chemical shift assignments were obtained. Both the *E* and *Z* forms were observable, the percentage of *Z* form being of the order of 20-30%. Coupling constant measurements indicated that the NH-CHα dihedral angles are different for the *E* and *Z* forms. Docking calculations with the HIV protease yield minimal energies which correlate with their experimental inhibition constants.

The spread of the AIDS virus, for which no cure exists, and its dramatic consequences on human health has lead to a considerable research effort in different fields. The main strategy consists in an inhibition of enzymes involved in the replication of the virus such as the reverse transcriptase, the protease and the integrase. In the case of the HIV-1 protease, a number of inhibitors have been designed and the structures of many inhibitor-enzyme complexes have been characterized by X-ray diffraction, NMR and molecular modeling.[1,2] Some of these inhibitors are highly active but are of little practical use due to their toxicity, lack of specificity or chemical instability. Inhibitors of peptidic nature are especially sensitive to hydrolysis.

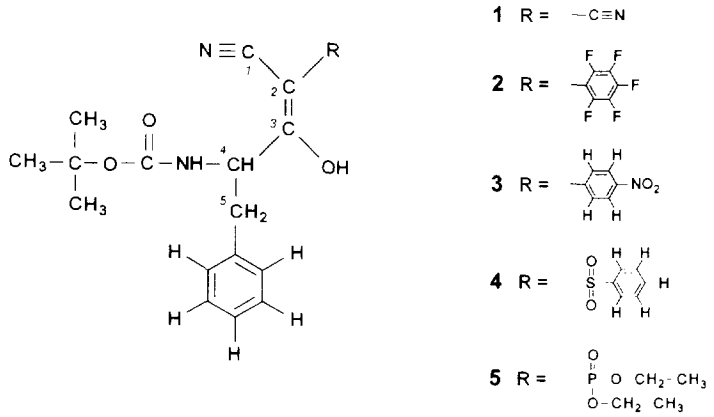

Fig. 1. Structure of Boc-Phe-enols (*E* form).

In an attempt to design simple and chemically stable inhibitors, difunctional enols of N-protected amino acids were synthesized based on their topographical similarities with other protease inhibitors such as N-*tert*-butoxycarbonylstatine.[3,4] Some of these enols were found to be relatively good inhibitors of the HIV-1 protease, the size and the electron-withdrawing character of the enol substituents being important factors for the inhibitory activity.[5] Only *E* isomers were found to be active.

In this work, NMR and molecular modeling were used to characterize the structure of five inhibitors of the N-*tert*-butoxycarbonylphenylalanyl (Boc-Phe) enol family. A complete analysis of the ^{1}H and ^{13}C NMR data could be performed. In addition, docking experiments to determine the fitting ability of each inhibitor into the binding site of the HIV-1 protease were conducted and the minimal intermolecular energies were found to correlate with the inhibition coefficients.

MATERIALS AND METHODS

Synthesis of Inhibitors

Difunctionalized enols of Boc-Phe with the R group on the enol being a pentafluorophenyl group ((*E,Z*)-(4*S*)-4-*tert*-butoxycarbonylamino-3-hydroxy-2-pentafluorophenyl-2-pentenenitrile **2**) or a 4-nitrophenyl group ((*E,Z*)-(4*S*)-4-*tert*-butoxycarbonylamino-3-hydroxy-2-(4-nitrophenyl)-2-pentenenitrile **3**) were synthesized as previously described.[4] The synthesis of the Boc-Phe-enol inhibitors with the R group being a cyano ((4*S*)-4-*tert*-butoxycarbonylamino-2-cyano-3-hydroxy-2-pentenenitrile **1**), a phenylsulfoxy ((*E,Z*)-(4*S*)-4-*tert*-butoxycarbonylamino-3-hydroxy-2-phenylsulfoxy-2-pentenenitrile **4**) or a diethylphosphonyl ((*E,Z*)-(4*S*)-4-*tert*-butoxycarbonylamino-2-diethylphosphonyl-3-hydroxy-2-pentenenitrile **5**) group will be reported separately (Vaillancourt, M. and Sauvé, G., to be published).

NMR Measurements

Samples (30-40 mg) of the enol inhibitors **1-5** were dried under vacuum at 40°C for several hours and dissolved in DMSO-d_6 to obtain final concentrations of 140-210 mM. NMR spectra were recorded on a Bruker AMX2 500 spectrometer (Bruker Spectrospin, Milton, Canada) operating at 500.13 MHz for ^{1}H and at 125.77 MHz for ^{13}C. All spectra were recorded at 300 K. One-dimensional ^{1}H and ^{1}H-decoupled ^{13}C NMR spectra as well as two-dimensional DQFCOSY, TOCSY, NOESY (mixing times 100-500 ms) and HMQC spectra with time-proportional phase incrementation were recorded. In order to determine whether the Phe NH group is in fast or slow exchange, its chemical shift was monitored by recording ^{1}H NMR spectra at temperatures of 300, 305, 310, 315, 320, 325 and 330 K. NMR spectra were processed with FELIX 2.30 software (Molecular Simulations, Inc., San Diego CA) operating on a Silicon Graphics Indigo workstation (Silicon Graphics Inc., Montreal, Canada).

Measurement of Inhibition Constants

The apparent inhibition constants $K_{i\,app}$ of the enol inhibitors were determined in the presence of the HIV-1

protease using the Dixon plot analysis as previously described.[5] Pepstatin was used as control.

Molecular Modeling

Molecular modeling calculations were performed on the Silicon Graphics workstation using softwares (INSIGHT II, DISCOVER, DGII, DOCKING, SORPTION) purchased from Molecular Simulations Inc. The chemical structures of the inhibitors were constructed using the Molecular Simulations consistent valence forcefield (CVFF) and submitted to a structural optimization using 1000 steps of energy minimization with the steepest descent and conjugate gradients algorithms.The structure of the HIV-1 protease, determined by X-ray crystallography in the presence of the hydroxyethylene pseudopeptide inhibitor Ala-Ala-Phe-ψ(CHOH-CH_2)-Ala-Val-Val-OMe,[6] was obtained from the Protein Data Bank. Docking of the inhibitor was first performed by positioning the hydroxyl group between the carboxyl groups of the aspartic acid residue 25 of the two monomers. The intermolecular energy was calculated between the inhibitor and a molecular subset of the protease consisting of a 10 Å sphere around the inhibitor. The inhibitor was first manually displaced to obtain an initial minimal intermolecular energy. When reasonable energies could be obtained (less than 10^7 kcal/mol), the enzyme-inhibitor complex was submitted to 2000 steps of docking using the SORPTION module. This procedure was performed at 500 K allowing a maximal displacement of 0.75 Å and a maximal rotation of 15°. The procedure ended when no further energy reduction could be achieved for 100 consecutive steps.

RESULTS AND DISCUSSION

NMR Data

Proton and ^{13}C chemical shifts for all five enol inhibitors **1-5** are given in Table 1 and Table 2, respectively. For most protons of inhibitors **2-5**, two series of chemical shift values are observed, the more intense signals being observed for the more stable *E* conformers and the less intense signals corresponding to the *Z* conformers. In all cases, the percentage of *Z* conformer was less than 30%. Fig. 2 presents the ^1H NMR spectrum of inhibitor **2** where

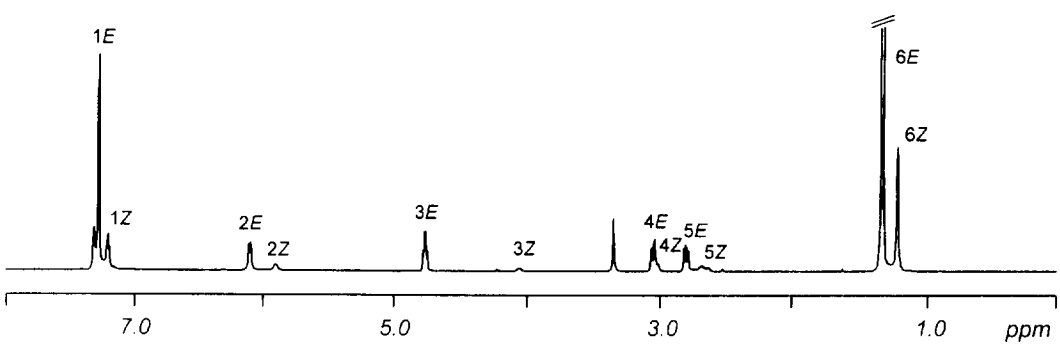

Fig. 2. ^1H NMR spectrum of enol inhibitor **2** showing the presence of both the *E* and *Z* forms. Peak assignments are: 1, aromatic protons of Phe; 2, NH of Phe; 3, CHα of Phe; 4, CHβ' of Phe; 5, CHβ of Phe; 6, CH_3's of Boc.

Table 1. ^{1}H Chemical Shifts of Enol Inhibitors

Type of proton[b]	Chemical shift (ppm)[a]				
	1	**2**	**3**	**4**	**5**
R-CHo			7.97	7.84	
R-CHm			7.99	7.47	
R-CHp				7.51	
R-CH_2					3.86
					3.94
R-CH_3					1.21
Phe-CHα	4.43	4.77	4.78	4.38	4.62
		4.06	4.09	4.31	4.01
Phe-CH_2β	2.68	2.80	2.73	2.56	2.87
		2.68	2.91		2.71
Phe-CH_2β'	2.92	3.05	3.04	2.84	3.07
		3.02	3.14		2.96
Phe-CHδ	7.29	7.24	7.27	7.17	7.17
		7.28	7.19		7.23
Phe-CHϵ	7.30	7.24	7.20	7.17	7.24
		7.28			
Phe-CHζ	7.22	7.17	7.13	7.13	7.16
		7.24			
Phe-NH	6.55	6.10	6.33	6.13	6.12-6.19
		5.91	5.98	5.91	5.84-5.91
Boc-CH_3	1.34	1.32	1.31	1.26	1.25-1.31
		1.20	1.16	1.05	1.19-1.25

[a]Values on the first line of each type of proton correspond to the *E* form and values on the second line correspond to the minor *Z* form. All spectra were recorded in DMSO-d_6 at 300 K. [b]Symbols o, m and p correspond to the ortho, meta and para protons on aromatic rings.

the percentage of *Z* conformer was estimated at 23% from the integration of the signals of the Phe-NH and Boc-CH_3 protons for both conformers. The percentages of *Z* conformer were 19, 23 and 29% for inhibitors **3**, **4** and **5**, respectively.

Chemical shifts are relatively similar for the enol inhibitors with different R groups (Tables 1 and 2). Significant differences are observed for the ^{13}C signal of enol C2 (Table 2). The lowest field signal is observed when the R group is a pentafluorophenyl (enol **2**), followed by the p-nitrophenyl (enol **3**), the cyano (enol **1**) and the phenylsulfoxy (enol **4**) and diethylphosphono (enol **5**) inhibitors. This order generally corresponds to the electron withdrawing potential of the R groups, the pentafluorophenyl being expected to be the strongest electroattractor due to the presence of five fluorine substituents. Other significant chemical shift differences are noticed for the NH, α, β and β' protons of the phenylalanyl portion of the inhibitors (Table 1) as well as for the enol C3 (Table 2). The nature of the substituents and/or their spatial orientation produce effects that can be felt at such a distance. Variations of the $^3J_{\alpha\beta}$ values are additional evidence that the structure of the phenylalanyl portion of the molecule is modified (Table 3).

The *E* and *Z* isomers of a same inhibitor display ^{1}H chemical shifts and coupling constants that differ

Table 2. ^{13}C Chemical Shifts of Enol Inhibitors

Type of carbon-13	Chemical shift (ppm)[a]				
	1	**2**	**3**	**4**	**5**
Enol C2	67.1	77.8	74.4	61.7	61.1
CN	122.1	113.8	121.2	121.2	122.5
	120.6[b]				
R-Cq[c]		137.6	137.1	145.7	
R-Co		136.8[d]	123.7	126.1	
R-Cm		137.3[d]	121.1	128.2	
R-Cp		143.8	148.8	131.2	
R-CH_2					61.0
R-CH_3					16.0-16.1
Enol C3	191.1	182.8	187.0	185.7	192.2
Phe-Cα (C3)	56.5	55.5	57.7	56.8	57.0
		53.4	59.7		56.0
Phe-Cβ (C4)	38.0	39.1	38.8	38.0	37.5
		40.2			39.1
Phe-Cγ	138.7	138.6	138.8	138.2	139.0
					138.3
Phe-Cδ	129.1	129.3	129.2	129.1	129.2-129.5
		129.4	129.5		
Phe-Cϵ	127.9	127.6	127.7	127.9	127.4-128.0
		127.9	127.8		
Phe-Cζ	125.6	125.7	125.8	125.6	125.5-126.0
		125.7	125.7		
Boc-CO	154.9	154.4	154.8	154.7	154.5-155.4
Boc-Cq[c]	77.5	77.5	77.4	77.6	77.4
Boc-CH_3	28.2	28.2	28.2	28.1	28.5
		27.7			28.1

[a]Values on the first line of each type of carbon-13 correspond to the *E* form and values on the second line correspond to the minor *Z* form. [b]Value for the second CN group on C2. [c]Cq is the quaternary carbon of the aromatic ring. [d]These assignments could be inverted.

significantly (Tables 1-3; Fig. 2). Differences are especially large for the Phe-CHα, Phe-NH and Boc-CH_3 which are the groups in closest proximity to the R group in the *Z* form. Measurable $^3J_{NH\alpha}$ and $^3J_{\alpha\beta}$ also display significant differences arising from a different orientation of the chemical groups. To better characterize the dihedral angle variation between the *E* and *Z* forms, Karplus-type equations were used to calculate the dihedral angles formed by protons around the NH-α bond (dihedral angle $\theta_{NH\alpha}$) and the α-β bond (dihedral angles $\theta_{\alpha\beta}$ and $\theta_{\alpha\beta'}$) of the Phe group from the three-bond coupling constants for inhibitors **2** and **3**.[7,8] Such equations yield two or four possible dihedral angle values. By comparison to the dihedral angles measured in the molecular model generated by energy minimization, unique sets of $\theta_{NH\alpha}$, $\theta_{\alpha\beta}$ and $\theta_{\alpha\beta'}$ values could be identified. For inhibitor **2**, the calculated values of $\theta_{NH\alpha}$, $\theta_{\alpha\beta}$ and $\theta_{\alpha\beta'}$ were 160, 140 and -116°, respectively, for the *E* form and 135, 145 and -116°, respectively, for the *Z* form. A difference of 25° is therefore calculated for $\theta_{NH\alpha}$ between the two forms whereas the $\theta_{\alpha\beta}$ and $\theta_{\alpha\beta'}$ angles are approximately the same. Fig. 3 illustrates the molecular model of inhibitor **2** constructed with dihedral

Table 3. ^{1}H-^{1}H Coupling Constants of Enol Inhibitors

Type of coupling	Coupling constant (Hz)[a]				
	1	**2**	**3**	**4**	**5**
$^3J_{o,m}$ (R)			8.8	7.2	
$^3J_{m,p}$ (R)				7.6	
$^3J_{NH\alpha}$ (Phe)	8.6	8.6	8.5	8.5	
		5.8	5.4		
$^3J_{\alpha\beta}$ (Phe)	10.3	8.5	9.9	9.3	7.4
		9.5	7.2		8.9
$^3J_{\alpha\beta'}$ (Phe)	3.3	4.3	3.1	3.2	3.4
		4.3	3.1		3.4
$^2J_{\beta\beta'}$ (Phe)	13.5	13.3	13.6	13.3	13.4
		13.1	13.0		12.5
$^3J_{\delta\varepsilon}$ (Phe)		7.1	7.7		
$^3J_{\varepsilon\zeta}$ (Phe)	5.8		7.0		
$^4J_{\delta\zeta}$ (Phe)	2.2				

[a]Values on the first line of each type of coupling correspond to the *E* form and values on the second line correspond to the minor *Z* form.

angles in agreement with those calculated from the coupling constants. In the *E* form, the measured dihedral angles are compatible with the formation of a hydrogen bond between the carbonyl oxygen of the Boc group and the hydroxyl hydrogen on C3 (Fig. 3A), the two atoms being distanced by 2.4 Å. Such is not the case for the *Z* form (Fig. 3B) where the two atoms are separated by more than 4.1 Å as a result of the repulsion between the bulky cis substituents on C2 and C3. The formation of the hydrogen bond in the *E* form, in addition to the reduced

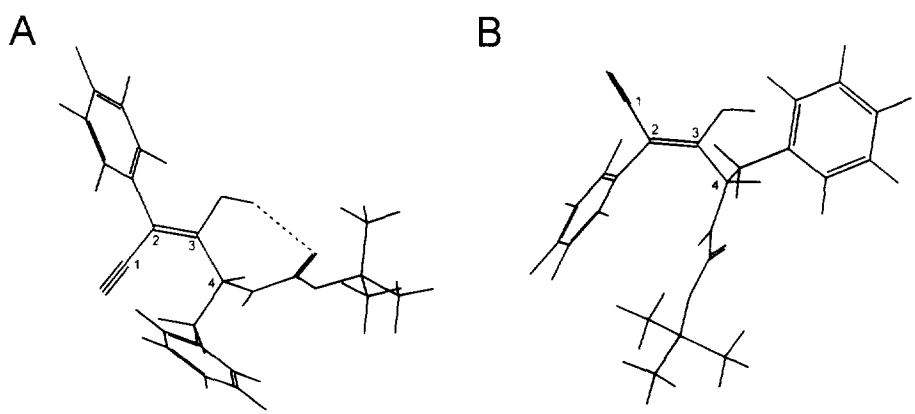

Fig. 3. Molecular model of (A) *E* form and (B) *Z* form of enol inhibitor **2** calculate by conjugated gradients energy minimization using dihedral angles constraints determined by NMR: $\theta_{NH\alpha} = 160°$, $\theta_{\alpha\beta} = 136°$ and $\theta_{\alpha\beta'} = -111°$ for the *E* form, and $\theta_{NH\alpha} = 135°$, $\theta_{\alpha\beta} = 139°$ and $\theta_{\alpha\beta'} = -108°$ for the *Z* form. Hydrogen bonding is formed between the hydroxyl hydrogen on C3 and the carbonyl oxygen on the Boc, distanced by 2.4 Å in the *E* form.

substituent repulsion, is a stabilizing factor explaining the predominance of the *E* form. Very similar results were obtained with inhibitor **3**. Indeed, the values calculated for the $\theta_{NH\alpha}$, $\theta_{\alpha\beta}$ and $\theta_{\alpha\beta'}$ dihedral angles were 158, 148 and -107°, respectively, for the *E* form, and 132, 132 and -107°, respectively, for the *Z* form. The molecular model of inhibitor **3** is very similar to that of inhibitor **2** (Fig. 3) with formation of the same hydrogen bonding in the *E* form.

The temperature coefficients of the NH chemical shifts were lower than -7 x 10^{-3} ppm/K for all enol inhibitors in both the *E* and *Z* forms. Such values are indicative of free exchange of the protons and are inconsistent with the formation of hydrogen bonding involving the NH group. The OH group was too broad to be observed and therefore, it was not possible to assess its involvement in hydrogen bonding by this method.

Binding to HIV-1 protease

Binding of the enol inhibitors to HIV-1 protease was assessed both experimentally, by measuring the inhibition constants of the protease and, theoretically, by conducting docking experiments with molecular models. Table 4 presents a comparison of the inhibition constants and of the minimal energies of the docked conformer for enol inhibitors in the *E* form. The results display a good correlation between both values, the best inhibition being obtained with inhibitor **2** and the worst with inhibitor **1**. A review of the crystal structures of inhibitor-protease complexes[9] indicates that the major factor for action is the location of a hydroxyl group of the inhibitor in the vicinity of the catalytic site, especially the side chains of Asp^{25} and $Asp^{25'}$. In our series, it can be observed that this criterion is best respected for inhibitor **2** (Fig. 4A) which equally displays the best experimental and theoretical binding parameters (Table 4).

For inhibitors **2-5**, the *Z* form of each inhibitor could not present adequate binding to the protease active site. In all cases, the best minimal energy was superior to 10^5 kcal/mol and the positioning of the inhibitor was different from that of the *E* form, especially regarding the orientation of the OH group on C3. The example of inhibitor **2** is displayed in Fig. 4. It can clearly be seen that the *Z* form (Fig. 4B) is oriented differently from the *E* form (Fig. 4A) and that the location of the OH group is different, pointing away from Asp^{25} and $Asp^{25'}$.

In summary, this study has allowed to determine both the ^1H and ^{13}C NMR parameters of a series of enol

Table 4. Correlation between inhibition constants (K_i) and minimal intermolecular energies (E_{min}) of the docked conformers of the enol inhibitors with HIV-1 protease

Enol inhibitor	K_i (μM)	E_{min} (kcal/mol)[a]
1: Boc-Phe-CN-CN	531	764
2: Boc-Phe-CN-ΦF$_5$	0.48	173
3: Boc-Phe-CN-ΦNO$_2$	2.3	223
4: Boc-Phe-CN-SO$_3$Φ	29	344
5: Boc-Phe-CN-PO$_3$Et$_2$	41	374

[a] Minimal intermolecular energies were measured within a 10 Å sphere around the inhibitor.

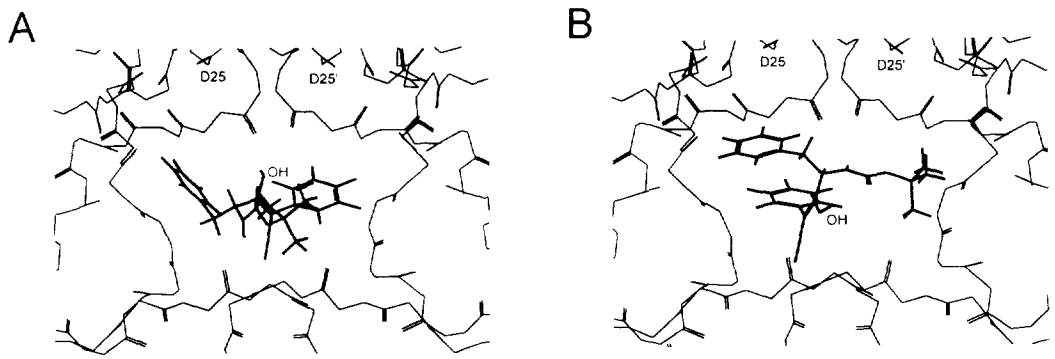

Fig. 4. Molecular models of the (A) *E* form and (B) *Z* form of the enol inhibitor **2** docked into the binding site of HIV-1 protease. Biological activity is associated with the location of the hydroxyl group of the inhibitor in proximity of the Asp^{25} and Asp^{25} as is the case in (A). The minimal intermolecular energies were 173 kcal/mol for the *E* form and 1.6 x 10^5 kcal /mol for the *Z* form.

inhibitors of the HIV protease, to characterize their structures, to determine their relative binding affinities and to demonstrate that the *E* form is predominant and biologically active. These parameters should help to design inhibitors of this nature with even stronger biological action.

ACKNOWLEDGMENTS

This work was supported by the Medical Research Council of Canada and Health and Welfare Canada.

REFERENCES

1. Huff, J.R. *J. Med. Chem.* **1991**, 34, 2305-2314.
2. Wlodawer, A.; Erickson, J.W. *Annu. Rev. Biochem.* **1993**, 62, 543-585.
3. Vaillancourt, M.; Vanasse, B.; Cohen, E.; Sauvé, G. *Bioorg. Med. Chem. Lett.* **1993**, 3, 1169-1174.
4. Vaillancourt, M.; Vanasse, B.; Le Berre, N.; Cohen, E.; Sauvé, G. *Bioorg. Med. Chem.* **1994**, 2, 343-355.
5. Vaillancourt, M.; Cohen, E.; Sauvé, G. *J. Enzyme Inhib.* **1995**, 9, 217-233.
6. Dreyer, G.B.; Lambert, D.M.; Meek, T.D.; Carr, T.J.; Tomaszek, Jr., T.A.; Fernandez, A.V.; Bartus, H.; Cacciavillani, E.; Hassell, A.M.; Minnich, M.; Petteway, Jr., S.R.; Metcalf, B.W.; Lewis, M. *Biochemistry* **1992**, 31, 6646-6659.
7. Ludvigsen, S.; Andersen, K.V.; Poulsen, F.M. *J. Mol. Biol.* **1991**, 217, 731-736.
8. DeMarco, A.; Llinas, M.; Wüthrich, K. *Biopolymers* **1978**, 17, 637-650.
9. Appelt, K. *Persp. Drug Disc. Design* **1993**, 1, 23-48.

(*Received in USA* 9 *December* 1996; *accepted* 31 *January* 1997)