Stereostructure of (−)-Chloropeptin I, a Novel Inhibitor of gp120−CD4 Binding, via High-Temperature Molecular Dynamics, Monte Carlo Conformational Searching, and NMR Spectroscopy

Hiroaki Gouda, Keiichi Matsuzaki, Haruo Tanaka, and Shuichi Hirono*

School of Pharmaceutical Sciences Kitasato University, Minato-ku, Tokyo 108, Japan

Satoshi Ōmura*

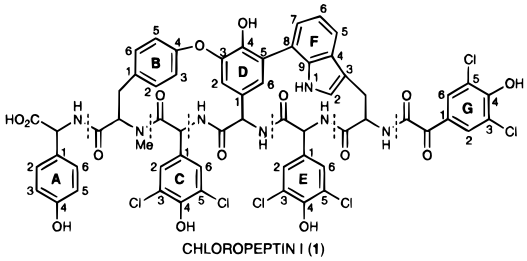
Research Center for Biological Function, The Kitasato Institute, and School of Pharmaceutical Sciences Kitasato University, Minato-ku, Tokyo 108, Japan

John A. McCauley, Paul A. Sprengeler, George T. Furst, and Amos B. Smith, III*

Department of Chemistry, University of Pennsylvania Philadelphia, Pennsylvania 19104

Received August 26, 1996

Recently, we reported the isolation and planar structure of (−)-chloropeptin I (**1**), an unusual chlorinated hexapeptide produced by the soil actinomycete *Streptomyces* sp. WK-3419.[1] Chloropeptin I strongly inhibits in vitro binding of the HIV-1 gp120 envelope glycoprotein to the CD4 protein (IC_{50} 2.0 μM); in addition, **1** selectively inhibits HIV replication in peripheral human lymphocytes.[1] These findings suggest that **1** might be used in combination with reverse transcriptase (e.g., AZT, ddC, ddI, and d4T) and/or protease inhibitors to improve management of the disease. Herein, we describe the elucidation of the relative and absolute stereochemistry of chloropeptin I.



Acidic hydrolysis of (−)-**1** (6 N HCl) furnished (*R*)-(*p*-hydroxyphenyl)glycine (residue A) and (*R*)-(3,5-dichloro-4-hydroxyphenyl)glycine (residues C and E), the absolute configurations of which were readily determined by optical rotation and chiral TLC.

For assignment of the C_α configurations of tyrosine (residue B), (*p*-hydroxyphenyl)glycine (residue D), and tryptophan (residue F), we employed two independent approaches, each combining computer modeling with NMR analysis.[2] Among the numerous computational methods for determining solution structures, molecular dynamics and Monte Carlo conformational searching have emerged as the most powerful. Molecular dynamics (MD) is particularly useful for molecules with many degrees of freedom. One drawback, however, is the failure to consistently generate molecule conformations separated by large energy barriers.[2e] Monte Carlo techniques do not suffer from this shortcoming but instead become less efficient as the number of degrees of freedom increases. Chloropeptin I falls within the scope of both methods, and it was not clear a priori which would be more effective.

The high-temperature molecular dynamics (MD) calculations were performed with the CAChe package (CAChe Scientific, Inc.) and CAMDAS program (conformational analyzer with molecular dynamics and sampling). The latter, developed in our laboratory,[3] generates the energetically accessible conformers of a target molecule by performing an MD calculation and sampling conformers along the trajectory. CAMDAS then clusters similar conformers based upon root-mean-square (rms) deviations of the atomic positions. In a test analysis of cyclodecane, CAMDAS reproduced the results of the systematic search reported previously.[3]

As our point of departure, we employed the CAChe package to prepare structure files for the eight diastereomers differing in the C_α configuration of residues B, D, and F. Each diastereomer required two initial structures with the positive and negative values for the side chain χ_2 dihedral angle (C_α−C_β−C3−C2) of residue F, because the high energy barrier prevented computational interconversion of the corresponding families of rotamers. Each of the 16 MD calculations was carried out for 1000 ps with an integral time step of 1 fs. The temperature of the system was maintained at 1200 K, and the covalent bond lengths were fixed with the SHAKE algorithm throughout the MD simulation.[4] We used the Allinger MM2 force field and its extended version to evaluate the potential energy surface of the molecule.[5,6] To mimic the shield effects of solvent molecules on electrostatic interactions, the electrostatic potential term was neglected. The potential-scaled method was adopted to accelerate the conformational search.[7,8] Conformers were sampled at 50-step (i.e., 0.05 ps) intervals to produce 20 000 conformers for each initial structure. The sets of conformers were preclustered with a root-mean-square deviation (rmsd) threshold of 0.5 Å for all non-hydrogen atoms, reducing the number of conformers per structure to about 2000. Each conformer was then minimized until the rms of the potential energy gradient fell below 0.05 kcal mol^{-1} $Å^{-1}$. The minimized conformers were reclustered with an rmsd threshold of 0.3 Å, furnishing a final conformer set for each initial structure. Table 1 lists the number of distinct conformers generated by CAMDAS for each diastereomer.

NOESY, double-quantum-filtered (DQF)-COSY, and exclusive E.COSY spectra of **1** provided requisite NOE data and coupling constants.[9−12] The nonequivalence of the C_β proton resonances for residue F enabled us to establish the range of the χ_1 angle via Karplus analysis of the $^3J_{\alpha\beta}$ values, in conjunction with the intraresidual NH−C_βH and C_αH−C_βH NOEs. The χ_1 angle proved to be 180 ± 60°, independent of configuration. The $^3J_{HN\alpha}$ coupling constants for residues D and

(1) Matsuzaki, K.; Ikeda, H.; Ogino, T.; Matsumoto, A.; Woodruff, H. B.; Tanaka, H.; Omura, S. *J. Antibiot.* **1994**, *47*, 1173.

(2) For MD methods, see: (a) Bruccoleri, R. E.; Karplus, M. *Biopolymers* **1990**, *29*, 1847. (b) van Schaik, R. C.; van Gunsteren, W. F.; Berendsen, H. J. C. *J. Comput.-Aided Mol. Des.* **1992**, *6*, 97. (c) Hooft, R. W. W.; van Eijck, B. P.; Kroon, J. *J. Chem. Phys.* **1992**, *97*, 6690. (d) Tsujishita, H.; Moriguchi, I.; Hirono, S. *J. Phys. Chem.* **1993**, *97*, 4416. For Monte Carlo methods, see: (e) Saunders, M.; Houk, K. N.; Wu, Y.-D.; Still, W. C.; Lipton, M.; Chang, G.; Guida, W. C. *J. Am. Chem. Soc.* **1990**, *112*, 1419. (f) Byrne, D.; Li, J.; Platt, E.; Robson, B.; Weiner, P. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 67.

(3) Tsujishita, H.; Hirono, S. *J. Comput.-Aided Mol. Des.* In press.

(4) Rychaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.

(5) Allinger, N. L. *J. Am. Chem. Soc.* **1977**, *99*, 8127.

(6) *CAChe Reference version 3.6*. CAChe Scientific, Inc., 1994.

(7) Tsujishita, H.; Moriguchi, I.; Hirono, S. *J. Phys. Chem.* **1993**, *97*, 4416−4420.

(8) Tsujishita, H.; Moriguchi, I.; Hirono, S. *Biophys. J.* **1994**, *66*, 1815.

(9) Jeener, J.; Meier, B. H.; Bachmann, P.; Ernst, R. R. *J. Chem. Phys.* **1979**, *71*, 4546.

(10) Rance, M.; Sorensen, O. W.; Bodenhausen, G.; Wagner, G.; Ernst, R. R.; Wüthrich, K. *Biochem. Biophys. Res. Commun.* **1983**, *117*, 479.

(11) Griesinger, C.; Sorensen, O. W.; Ernst, R. R. *J. Magn. Reson.* **1987**, *75*, 474.

(12) The spectra were recorded on a Varian UNITY-400 spectrometer at a ^1H frequency of 400 MHz. The sample was prepared by dissolving purified and lyophilized chloropeptin I in DMSO-d_6. The probe temperature was maintained at 25 °C. The NOESY mixing times were either 100 or 200 ms.

S0002-7863(96)03003-X CCC: $12.00

Table 1. Root-Mean-Square Deviations of CAMDAS-Generated Conformers of **1** from Experimentally Determined Values

Configuration[a]	Smallest RMS deviations from expti distance constraints (Å)	RMS deviation from expti dihedral constraints (°)[b]	Number of conformers obtained by CAMDAS
RRR$^-$	0.24	0.0	411
RRR$^+$	0.24	83.0	392
RRS$^-$	0.37	0.0	428
RRS$^+$	0.24	99.3	379
RSR$^-$	0.14	60.2	392
RSR$^+$	0.38	46.9	393
RSS$^-$	0.38	48.4	421
RSS$^+$	0.25	64.7	376
SRR$^-$	0.00	0.0	375
SRR$^+$	0.17	14.1	479
SRS$^-$	0.33	71.5	428
SRS$^+$	0.29	2.3	398
SSR$^-$	0.17	43.3	448
SSR$^+$	0.32	50.2	419
SSS$^-$	0.29	43.3	404
SSS$^+$	0.05	54.4	359

[a] The + and − refer to positive and negative χ_2 dihedral angles of residue F, respectively. [b] Calculated for the conformer having the smallest rms deviations from the distance constraints.

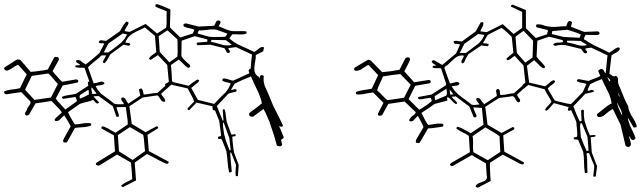


Figure 1. Stereoview of the solution conformation of (−)-chloropeptin I (**1**).

E, greater than 8 Hz, dictated ϕ angles of 120 ± 40° for D and −120 ± 40° and 120 ± 40° for the *S* and *R* diastereomers of E, respectively. Using these three torsion angle constraints and a total of 24 distance constrants, the latter derived from NOE data, we calculated rms deviations for all conformers generated by CAMDAS in order to determine which configurations can adopt conformers accommodating the constraints. Table 1 lists the smallest rms distance deviations obtained for each initial structure. The single family of satisfactory conformers embodied the *S*, *R* and *R* configurations at C_α of residues B, D, and F, respectively. The strong NOE between the C_α protons of residues B and C suggests that the peptide bond linking these residues is cis. The NOE between H(2) of the indole unit and the C_α proton in residue F conclusively established the spatial orientation of the side chain. Figure 1 provides a stereoview of the resultant solution conformation.

Independently, we also pursued a stepwise alternative analysis of the stereochemistry and solution conformation of **1**, employing Monte Carlo conformational searches[13] with the AMBER* force field[14] in conjunction with NOSEY and DQF-COSY NMR techniques (500 MHz; mix times of 50, 100, 200, and 600 ms). For molecular modeling, we initially assigned the *R* configuration to the central amino acid D. We then performed Monte Carlo conformational searches on the core of chloropeptin I for the four diastereomers created by varying the unknown C_α configurations in residues B and F. Residues A and D, which do not reside within the core ring, were excluded. In each case, MacroModel generated 10 000 conformers which in turn were minimized. NOE correlations served to illuminate four remaining structural questions: (1) the spatial orientation of the indole ring (residue F χ_2 angle); (2) the stereochemistry of the tryptophan residue F; (3) the configuration of the tyrosine residue B; (4) the geometry (cis or trans) of the *N*-methyl amide bond. Table 2 illustrates the predicted interatomic distances

Table 2. Comparison of Key NOEs and Computer Generated Distances of **1**

χ_2 Angle of Residue F				Residue C Amide			
Protons	Distance (Å)[a]			Protons	Distance (Å)		
	$+\chi_2$	$-\chi_2$	NOE		trans	cis	NOE
H1(F)-C_αH(E)	3.71,3.68	4.81,4.82	W	C_αH(C)-C_αH(B)	4.52	2.14	S
H1(F)-NH(E)	4.45,4.13	3.45,3.28	M	C_αH(C)-NCH_3(B)	1.7-3.2	4.5-4.8	None
H2(F)-C_αH(E)	3.19,3.11	4.86,4.85	W	Residue B Configuration			
H2(F)-NH(E)	4.01,3.71	2.28,2.29	S	Protons	Distance (Å)		
Residue F Configuration					R	S	NOE
Protons	Distance (Å)			C_αH(B)-NCH_3(B)	2.8	3.2	W
	S	R	NOE	C_αH(B)-H2(B)	3.9	4.32	W
C_αH(F)-NH(E)	3.52	2.40	S	C_αH(B)-H6(B)	4.5	2.78	S
C_αH(F)-H2(F)	4.10	2.44	S	C_αH(B)-C_αH(C)	3.92	2.14	S
C_αH(F)-NH(F)	2-3	4.5	W	C_αH(C)-H6(B)	3.1	2.85	S
				H6(B)-NH(A)	1.9	3.8-4.5	M

[a] The two distances are for the *S* and *R* configurations of residue F, respectively.

for the relevant protons in the lowest-energy minimized conformations for the various diastereomers determined by the Monte Carlo conformational search and the strong (s) or medium strong (ms) NOEs, indicating relatively short interatomic distances and the weak (w) NOEs corresponding to longer separations. Comparison of the NOE data with the calculated distances indicated that (A) the indole nitrogen is pointed back, as drawn, (B and C) the configurations of the tryptophan residue F and the tyrosine residue B are D and L, respectively, permitting the equatorial disposition of both side chains, and (D) the methylated amide bond is cis. Additional calculations employing the opposite C_α stereochemistry for residue D could not be reconciled with the NOE data. Incorporating the B and F configurations, we next performed a Monte Carlo conformational search for the entire chloropeptin I structure, generating more than 5000 conformers. Energy minimization via a volume-based continuum model for water reduced the conformer population to 340. Finally, NOE distance constraints, derived from the volume integrals of the NOESY crosspeaks, were added to the model; these constraints excluded all but a single family of low-energy conformers, comprising side chain rotamers appended to a common central core. The resultant stereostructure proved to be essentially identical to that determined via the CAMDAS method.

The elucidation of the stereostructure of chloropeptin I could lead to an improved understanding of the mechanism of inhibition of gp120−CD4 binding. For example, the extended main chain of **1** could serve as either an acceptor or donor in hydrogen bonding with β-sheet target proteins; CD4 contains β-sheets, although their role in gp120 binding remains to be established. Interestingly, the shape of the rings and central peptide strand in **1** strongly resembles the crystal structure of a vancomycin derivative;[15] the major difference is that the sugar moiety and extra ring of vancomycin form a well-defined binding pocket which is not observed for **1**. The β-strand of -L-Lys-D-Ala-D-Ala, a bacterial cell wall peptide, is known to participate in binding with vancomycin.[15,16] Further studies of chloropeptin I should facilitate the rational design of more effective gp120 binding inhibitors, which may prove useful in the treatment of AIDS.

Acknowledgment. Support in the U.S. was provided by the National Institutes of Health (Institutes of General Medical Sciences) through Grant GM-29028. The authors wish to thank Dr. Christopher S. Shiner for helpful suggestions and critical comments.

Supporting Information Available: Tables of ^{1}H and ^{13}C chemical shifts, the experimental constraints used in the calculation of rms deviations, and figures illustrating the data in Table 2 (5 pages). See any current masthead page for ordering and Internet access instructions.

JA963003D

(13) The MacroModel Interactive Molecular Modeling System version 5.0 was employed, see: Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Compt. Chem.* **1990**, *11*, 440.

(14) McDonald, D. Q.; Still, W. C. *Tetrahdron Lett.* **1992**, *33*, 7743. Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Chio, C.; Alagona, G.; Profeta, S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765. Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230.

(15) Sheldrick, G. M.; Jones, P. G.; Kennard, O.; Williams, D. H.; Smith, G. A. *Nature* **1978**, *271*, 223.

(16) Knox, J. R.; Pratt, R. F. *Antimicrob. Agents Chemother.* **1990**, *34*, 1342. See also: Williams, D. H. *Acc. Chem. Res.* **1984**, *17*, 364.