Structural and Conformational Studies of [Ile^7] and [Leu^7]Surfactins from *Bacillus subtilis natto*

Hideji ITOKAWA,* Toshihiko MIYASHITA, Hiroshi MORITA, Koichi TAKEYA, Toshihiko HIRANO, Masato HOMMA, and Kitaro OKA

Tokyo College of Pharmacy, Horinouchi 1432–1, Hachioji, Tokyo 192–03, Japan.
Received July 7, 1993; accepted November 9, 1993

A novel [Ile^7]surfactin (1), which showed anti-human immunodeficiency virus activity, has been isolated from *Bacillus subtilis natto*. Structural and conformational analysis of the peptide backbone of [Ile^7]surfactin was conducted by a combination of various two-dimensional (2D) nuclear magnetic resonance (NMR), circular dichroism (CD) spectroscopy and simulated annealing calculations, compared with a known [Leu^7]surfactin (2). Both surfactins were shown to exist in different conformational states in both polar and apolar solvents.

Keywords surfactin; *Bacillus subtilis natto*; conformation; NMR; simulated annealing calculation; anti-HIV activity

Surfactins[1,2] are cyclic depsipeptides produced by *Bacillus subtilis* and *B. subtilis natto*, which contain a β-hydroxy fatty acid and seven amino acids. They show potent antifungal activities,[3] antitumor activities against Ehrlich ascites carcinoma cells[4] and inhibit fibrin clot formation,[1] and also function as an inhibitor of cyclic adenosine 3′,5′-monophosphate phosphodiesterase.[5]

In the course of our studies on antitumorous natural cyclic peptidic compounds,[6,7] we re-examined a series of surfactins produced by *Bacillus subtilis natto*, and reported the satisfactory separation and MS-MS spectrometric analysis of crude surfactins in a preliminary form.[7] In mycosubtilin[8] and iturin,[9] known as peptidic lipids, their congeners with various peptide sequences showed different conformations and different biological activities. Structural and conformational analyses of the peptide backbones of a novel [Ile^7]surfactin (**1**) and [Leu^7]surfactin (**2**) isolated from the viscous phase of "natto", which is a popular Japanese food made of soybeans fermented with *Bacillus subtilis natto*, by means of various two dimensional (2D) NMR, circular dichroism (CD) spectra and simulated annealing calculations, showed that **1** contains isoleucine at residue 7 and exhibits a different conformation from that of **2**, with leucine at residue 7 in solution. The present paper characterizes a new variant of surfactin, [Ile^7]surfactin (**1**), with a peptide sequence and conformation different from the known [Leu^7]surfactin (**2**). Anti-human immunodeficiency virus (anti-HIV) activities of isolated surfactins are also described.

Results and Discussion

Isolation and Structure Elucidation of [Ile^7] and [Leu^7]-Surfactins The viscous phase of natto was extracted with hot water and, on the addition of methanol, the soluble part was concentrated. Further, this extract was subjected to Celite column chromatography using a 10% ethanol–methylene chloride solvent system to give a crude surfactin. Final purification was carried out by silica gel HPLC using a mixture of distilled water, acetic acid, ethanol, and *n*-hexane (0.2 : 0.2 : 7.0 : 92.6) as a solvent system[7] to give two cyclic depsipeptides (**1** and **2**).

Compound **1** was given as colorless needles, mp 131—133 °C, and the molecular formula, $C_{53}H_{93}N_7O_{13}$, was established by high resolution (HR)-fast atom bombardment mass spectrum (FAB-MS). The IR spectrum showed absorption bands due to amide (3309 cm^{-1}) and amide carbonyl (1651 cm^{-1}) groups, which are characteristic of peptide bonds. Hydrolysis of **1** with 6 N HCl, followed by derivation with Marfey's reagent and HPLC analysis,[10] confirmed the presence of L-Glu, L-Leu, D-Leu × 2, L-Val, L-Asp and L-Ile.

The amino acid sequence was elucidated by a combination of 2D NMR techniques, such as correlated spectroscopy (COSY), homonuclear Hartmann–Hahn (HOHAHA)[11] and phase sensitive nuclear Overhauser and exchange spectroscopy (NOESYPH)[12] spectra. Pyridine-d_5 was selected as a measuring solvent because the peaks were broadened in other solvents, except for pyridine-d_5 and methanol-d_4. In the ^{1}H-NMR spectrum, seven signals due to amide protons around δ 8.5—9.4 were observed. In ^{1}H–^{1}H COSY and HOHAHA spectra, the presence of individual amino acids of Glu, Leu × 3, Val, Asp and Ile was revealed by the analysis of correlated peaks from these amide protons. Then, the ^{13}C signals which corresponded to the above amino acids could be assignable by a heteronuclear multiple quantum coherence (HMQC) spectrum.[13] The amino acid sequence was determined to be Glu–Leu–Leu–Val–Asp–Leu–Ile by the NOE relationship between the Hα proton and the neighboring amide proton (Fig. 2). In the NOESYPH spectrum, a strong cross peak between the amide proton of L-Glu and the α and β-protons in a fatty acid was observed. Furthermore, in the heteronuclear multiple bond correlation spectroscopy (HMBC) spectrum,[14] a cross peak between the carbonyl carbon in a fatty acid and the amide proton of L-Glu was observed, together with a cross peak between this carbon and the α and β-protons in a fatty acid, corroborating the presence of a fatty acid between L-Glu and L-Ile. From these 2D NMR measurements, complete ^{1}H and ^{13}C assignments of **1** were conducted as shown in Tables I and II.

Compound **2** was given as colorless needles, mp 129—131 °C, and the molecular formula, $C_{52}H_{91}N_7O_{13}$, was established by HR-FAB-MS spectrum. Amino acid analysis of Marfey's derivatives after acid hydrolysis suggested the same amino acid composition as the known

surfactin. Complete ^1H and ^{13}C assignments of **2** can be made by the same procedure described in the case of **1**, and they confirm the same amino acid sequence as the known surfactin.[1,2] From the molecular formulae of **1** and **2**, the length of fatty acid composed in **2** was considered to be less than that in **1** by one methylene, and these fatty acids in **1** and **2** were deduced to be as shown in Fig. 1. In this study, however, the identification of these fatty acids has not been examined. From the above results, the structures of compounds **1** and **2** were established to be [Ile^7] and [Leu^7]surfactins, respectively, as shown in Fig. 1.

Conformational Properties of the Peptide Backbones of 1 and 2 The flexibility of peptides in solution, in general, is somewhat difficult to determine their conformation. NMR provides the most reliable information about the structure in solution, provided that an unambiguous assignment of the molecular constitution is possible. NOE relationships observed in **1** and **2**, in pyridine-d_5, are listed in Fig. 2. As can be seen in Fig. 2, different relationships were observed between **1** and **2**, suggesting that **1** and **2**

TABLE I. ^1H-NMR Chemical Shifts of [Ile^7] and [Leu^7]Surfactins in Pyridine-d_5 at 303 K (500 MHz)

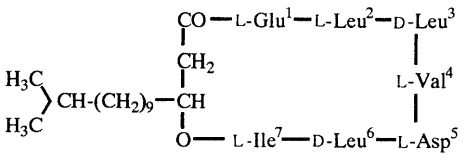
Amino acid		[Ile^7]Surfactin			[Leu^7]Surfactin	
L-Glu^1	Hα	5.03	$J_{\alpha N}=6.5$		4.84	
	Hβ1	2.60	$J_{\alpha\beta 1}=6.8$		2.59	
	Hβ2	2.69	$J_{\alpha\beta 2}=6.8$		2.68	
	Hγ1	2.95			2.97	
	Hγ2	2.95			2.97	
	HN	8.89			8.78	
L-Leu^2	Hα	4.92			5.08	
	Hβ1	1.99			2.19	
	Hβ2	2.15			2.19	
	Hγ	1.99			1.97	
	Hδ1	a)			0.97	
	Hδ2	a)			1.11	
	HN	9.12			8.70	
D-Leu^3	Hα	4.94	$J_{\alpha N}=5.7$		4.79	
	Hβ1	1.90			2.01	
	Hβ2	2.05			2.01	
	Hγ	1.90			2.00	
	Hδ1	a)			0.81	
	Hδ2	a)			0.98	
	HN	9.03			8.59	
L-Val^4	Hα	4.77	$J_{\alpha N}=7.1$		4.68	$J_{\alpha N}=5.9$
	Hβ	2.63	$J_{\alpha\beta}=7.1$		2.59	$J_{\alpha\beta}=5.9$
	Hγ1	1.19	$J_{\beta\gamma 1}=6.9$		1.13	$J_{\beta\gamma 1}=6.9$
	Hγ2	1.21	$J_{\beta\gamma 2}=6.9$		1.16	$J_{\beta\gamma 2}=6.8$
	HN	9.08			9.69	
L-Asp^5	Hα	5.63	$J_{\alpha N}=5.9$		5.72	
	Hβ1	3.40	$J_{\alpha\beta 1}=5.7$		3.54	$J_{\alpha\beta 1}=9.4$
	Hβ2	3.59	$J_{\alpha\beta 2}=8.4$		3.73	$J_{\alpha\beta 2}=3.4$
			$J_{\beta 1\beta 2}=16.3$			$J_{\beta 1\beta 2}=16.3$
	HN	9.31			8.95	
D-Leu^6	Hα	5.13	$J_{\alpha N}=7.7$		5.02	
	Hβ1	1.97			1.97	
	Hβ2	2.13			2.27	
	Hγ	1.97			2.18	
	Hδ1	a)			0.85	
	Hδ2	a)			1.02	
	HN	8.63			8.30	
L-Ile^7 (L-Leu^7)	Hα	4.71	$J_{\alpha N}=7.4$	Hα	4.60	
	Hβ	2.31	$J_{\alpha\beta}=7.0$	Hβ1	2.18	
	Hγ1	1.46		Hβ2	2.28	
	Hδ2	1.78		Hγ	2.01	
	HγME	1.13	$J_{\beta\gamma b}=6.8$	Hδ1	0.95	
	Hδ	0.93		Hδ2	0.99	
	HN	9.20		HN	9.64	

a) It could not be determined in the present study.

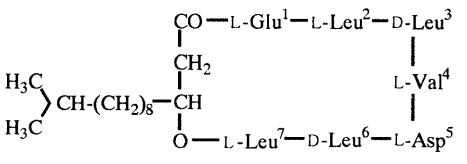
TABLE II. ^{13}C-NMR Chemical Shifts of [Ile^7] and [Leu^7]Surfactins in Pyridine-d_5 at 303 K (125 MHz)

Amino acid		[Ile^7]Surfactin		[Leu^7]Surfactin	
L-Glu^1	C_α	54.22		55.49	
	C_β	28.47		28.11	
	C_γ	31.43		31.36	
	C_δ	175.69		175.47	
	$C_{C=O}$	172.90		172.36	
L-Leu^2	C_α	53.07		52.04	
	C_β	39.99		39.26	
	C_γ	25.16		25.30	
	$C_{\delta 1}$	21.58[a)]		23.56	
	$C_{\delta 2}$	21.73[a)]		21.10	
	$C_{C=O}$	174.48		175.09	
D-Leu^3	C_α	52.94		53.72	
	C_β	39.99		39.82	
	C_γ	24.90		25.34	
	$C_{\delta 1}$	21.92[a)]		21.95	
	$C_{\delta 2}$	23.02[a)]		22.70	
	$C_{C=O}$	174.22		172.21	
L-Val^4	C_α	60.81		61.33	
	C_β	30.50		29.91	
	$C_{\gamma 1}$	19.00		18.56	
	$C_{\gamma 2}$	19.60		19.45	
	$C_{C=O}$	172.26		172.01	
L-Asp^5	C_α	51.48		51.58	
	C_β	37.46		37.22	
	C_γ	173.85		173.89	
	$C_{C=O}$	172.11		173.77	
D-Leu^6	C_α	52.43		52.82	
	C_β	41.61		42.45	
	C_γ	34.52		24.85	
	$C_{\delta 1}$	23.11[a)]		21.46	
	$C_{\delta 2}$	23.27[a)]		23.27	
	$C_{C=O}$	173.41		175.98	
L-Ile^7 (L-Leu^7)	C_α	58.14	C_α	52.50	
	C_β	36.70	C_β	39.53	
	C_γ	25.63	C_γ	25.10	
	$C_{\gamma ME}$	16.22	$C_{\delta 1}$	21.36	
	C_δ	22.70	$C_{\delta 2}$	22.85	
	$C_{C=O}$	172.20	$C_{C=O}$	173.35	

a) Assignment may be interchanged.



1



2

Fig. 1. Structures of [Ile^7]Surfactin (**1**) and [Leu^7]Surfactin (**2**)

The position of L- and D-leucines was arranged according to those in the known surfactin cited in the previous paper.[1,2]

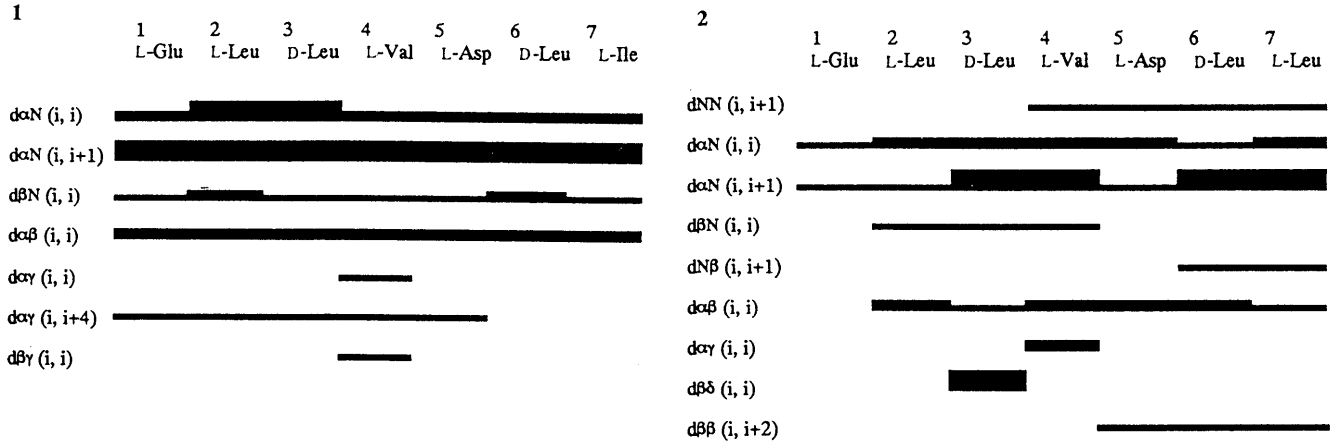
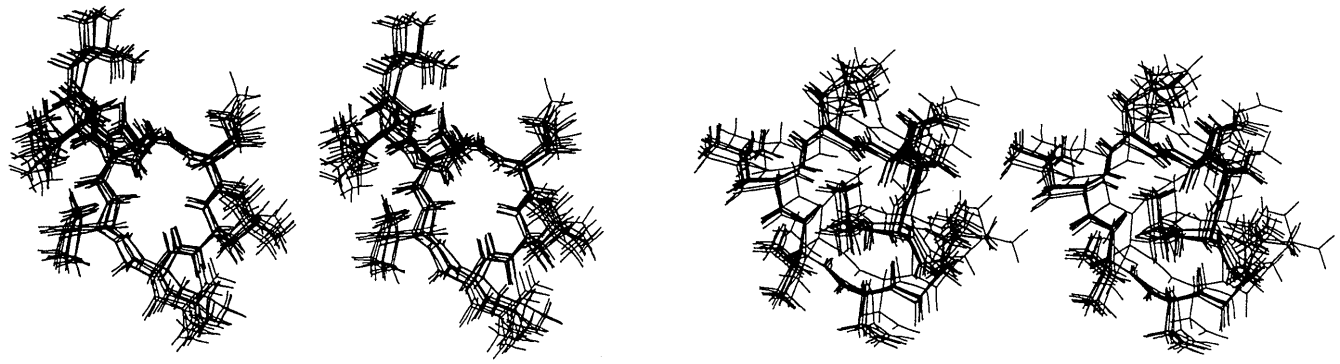


Fig. 2. The Sequence of [Ile^7]Surfactin (**1**) and [Leu^7]Surfactin (**2**) Together with a Summary of the Observed Short-Range Intraresidue NOEs

The intensities were obtained by reading heights of cross peaks in NOESYPH spectra. The height of the bars show an intensity classification of cross peaks into strong, medium and weak.



1 **2**

Fig. 3. Stereoviews Are Shown of a Superposition of Each of the 10 Best Energy-Minimized Structures of [Ile^7]Surfactin (**1**) and [Leu^7]Surfactin (**2**) after Simulated Annealing

TABLE III. Effect of Temperature on the NH Chemical Shifts of [Ile^7] and [Leu^7]Surfactins in Pyridine-d_5, $-\Delta\delta/\Delta T$ (10^{-3} ppm/K)

Compounds	Glu^1	Leu^2	Leu^3	Val^4	Asp^5	Leu^6	Leu(Ile)^7
[Ile^7]Surfactin	7.6	15.0	14.0	11.0	12.6	11.0	10.6
[Leu^7]Surfactin	3.0	6.6	8.0	13.6	4.3	6.0	15.3

exist in different conformational states in pyridine-d_5.

The involvement of NH groups in the intramolecular hydrogen bond was confirmed on the basis of the effect of temperature on the NH chemical shifts. The temperature effect on NH chemical shifts is often used to identify external or internal NH orientations.[15] The temperature coefficients ($d\delta/dT$) of **1** and **2**, in pyridine-d_5, are listed in Table III. As can be seen in Table III, the effect of temperature on the NH chemical shifts of **1** was larger than on those of **2**. The NH proton of Glu-1 in **2** was involved in a weak intramolecular hydrogen bond. However, most of the NH protons in **1** and **2** were not involved in intramolecular hydrogen bonds. From this temperature effect, different conformations in pyridine-d_5 between **1** and **2** were again deduced.

For the purpose of predicting or analyzing complicated conformational features of this series of compounds, it is necessary to use a computational method which can give us a result which doesn't depend on the starting structure. As it is very difficult to cover all the possible ring structures, including conformational freedom in substituent groups on the ring, molecular mechanics calculation is not efficient in this molecular system. The molecular dynamics (MD) calculation is much better in overcoming the local minima problem. The NOE constraint MD techniques as a tool for simulated annealing[16] was tested in the case of the molecules of these surfactins.

The starting geometries of **1** and **2** for the simulation were modeled by NOE constraint energy minimization. A simulation was performed using a time step of 1 fs, and the structures were sampled every 90 fs. Each system was equilibrated for 5400 fs with a thermal bath at 500 K, and thereafter, successively, for 900 fs with a thermal bath 10 K lower in temperature, until a final temperature of 50 K was obtained. Ten cycles are performed, giving a total simulation time of 63 ps, and each freezed conformation was sampled from the minimum temperature at 50 K. The minimized structures were ranked in order of increasing energy, and every ten matching low energy conformations are depicted in Fig. 3. Simulated annealing supports the conclusion that the molecules exhibit different conformations from each other.

The CD spectra of **1** and **2** in $CHCl_3$ and MeOH are shown in Fig. 4. Both **1** and **2** show a positive Cotton

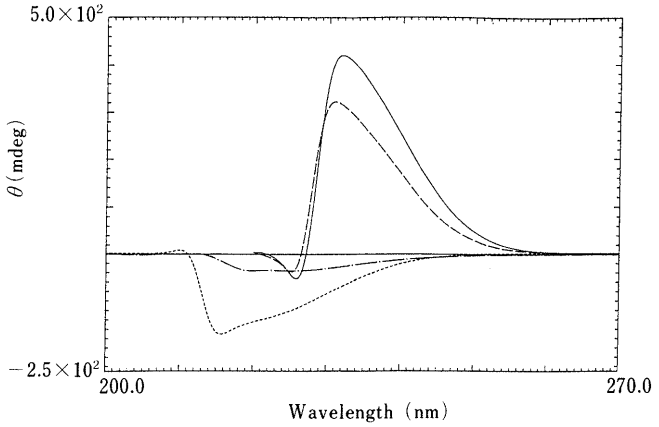


Fig. 4. CD Spectra of [Ile⁷] and [Leu⁷]Surfactins in $CHCl_3$ and MeOH

——, [Ile⁷]surfactin in $CHCl_3$; – – – –, [Leu⁷]surfactin in $CHCl_3$; —·—, [Ile⁷]surfactin in MeOH; ------, [Leu⁷]surfactin in MeOH.

TABLE IV. Optical Rotations of [Ile⁷] and [Leu⁷]Surfactins

Compounds	Solvents	$[\alpha]_D$ (°)	c
[Ile⁷]Surfactin	$CHCl_3$	69.1	0.11
	MeOH	−12.7	0.11
[Leu⁷]Surfactin	$CHCl_3$	44.0	0.10
	MeOH	−38.0	0.10

curve in $CHCl_3$; however, in a polar solvent such as MeOH, the Cotton curve is negative. From the above CD curves, **1** and **2** were considered to take on different conformations in $CHCl_3$ and MeOH. Then, the values of optical rotations of **1** and **2** in each solvent shown in Table IV are related to the CD curves. Though the signs almost corresponded, different values were observed by comparisons between **1** and **2**, suggesting different conformations in each solvent. In general, these surfactins were deduced to take different conformations under hydrophilic and lipophilic conditions.

The surfactins were well known to show various pharmacological activities.[1-5] We examined the other biological activities of isolated surfactins **1** and **2** and found out that they showed moderate anti-HIV activities in XTT formazan assay for HIV-1 cytopathic effects (**1**: EC_{50} 2.00×10^{-5} M; **2**: 1.40×10^{-5} M).[17] No difference in anti-HIV activities between **1** and **2** was observed.

Experimental

All melting points were recorded on a Yanagimoto MP-3 micromelting point apparatus and are uncorrected. The spectral data were obtained on the following instruments: IR spectrum on a JASCO A-302, CD on a JASCO J-700, NMR on a Bruker AM500 and MS on a VG AutoSpec. Preparative HPLC was carried out on a Senshu Pak Silica-5251-N column (20 i.d. × 250 mm, Senshu Scientific Co., Ltd.) packed with 5 μm silica gel.

Extraction and Isolation Commercial natto (3.5 kg) was extracted with hot water and, on the addition of methanol, the methanol soluble phase was concentrated to give a crude extract (5.45 g). This extract was passed through a Celite column eluted with 10% ethanol/methylene chloride to give crude surfactins (2.25 g). Chromatographic purification by HPLC was carried out using a preparative column to give compounds **1** (30 mg) and **2** (150 mg). The mobile phase solvent on HPLC was a mixture of distilled water, acetic acid, ethanol, and *n*-hexane (0.2 : 0.2 : 7.0 : 92.6) at a flow rate of 5.0 ml/min.

[Ile⁷]Surfactin (**1**): Colorless needles, mp 131.0—133.0 °C, $[\alpha]_D$ −12.7° (c=0.11, MeOH). HR-FAB-MS: Calcd for $C_{53}H_{94}N_7O_{13}$ 1036.6909 $[M+1]^+$. Found 1036.6847. IR (KBr) cm^{-1}: 3309 (NH), 1651 (amide C=O). CD: $[\theta]_{232\,nm}$: $+4.30\times10^5$ ($c=0.97\times10^{-3}$, $CHCl_3$), $[\theta]_{219\,nm}$: -4.45×10^4 ($c=0.97\times10^{-3}$, MeOH).

[Leu⁷]Surfactin (**2**): Colorless needles, mp 129.0—131.0 °C, $[\alpha]_D$ −38.0° (c=0.10, MeOH). HR-MS: Calcd for $C_{52}H_{92}N_7O_{13}$ 1022.6753 $[M+1]^+$. Found 1022.6688. IR (KBr) cm^{-1}: 3317 (NH), 1651 (amide C=O). CD $[\theta]_{231\,nm}=+3.20\times10^5$ ($c=0.98\times10^{-3}$, $CHCl_3$). $[\theta]_{215\,nm}$: -2.00×10^5 ($c=0.98\times10^{-3}$, MeOH).

Amino Acid Analysis of the Hydrolysate of 1 and 2 Solutions of **1** and **2** (each containing 1 mg of peptides) in 6 N HCl were heated at 100 °C for 13 h. After being cooled, each solution was concentrated to dryness. The residue was soluble in water and treated with 1-fluoro-2,4-dinitrophenyl-5-L-alanine amide (Marfey's reagent) and 1 M $NaHCO_3$ at 40 °C for 1 h. After being cooled, 2 M HCl was added and then concentrated to dryness. This residue was subjected to HPLC (Lichrospher 100, RP-18 (10 μm), Merck), flow rate 2 ml/min, detection 340 nm, solvent A: 10—50% MeOH/50 mM triethylamine phosphate (TEAP) buffer. The t_R values were L-Asp 15.7, L-Glu 16.2, L-Val 26.9, L-Leu and L-Ile 32.9, D-Leu 39.6. Solvent B: 23% CH_3CN/50 mM–TEAP buffer. The t_R values of L-Leu and L-Ile were 32.7 and 29.8, respectively.

Bioassay of HIV-1 Cytopathic Effects See previous paper.[17]

Simulated Annealing Calculation Computer modeling and all calculations were performed using the molecular-modeling software SYBYL (Tripos Associates, St. Louis, MO) on an IRIS 4-D work station. Molecular dynamics calculation was employed by the standard SYBYL force field.[18] The dielectric constant (ε) was assumed to be proportional to interatomic distances (r) as $\varepsilon=r$. Solvent molecules were not included in the calculations. The structures obtained by simulated annealing were further minimized with the standard SYBYL force field.

References

1) K. Arima, A. Kakinuma, G. Tamura, *Biochem. Biophys. Res. Commun.*, **31**, 488 (1968).
2) A. Kakinuma, M. Hori, M. Isona, G. Tamura, K. Arima, *Agr. Biol. Chem.*, **33**, 971 (1969); A. Kakinuma, H. Sugino, M. Isona, G. Tamura, K. Arima, *ibid.*, **33**, 973 (1969); A. Kakinuma, M. Hori, H. Sugino, J. Yoshida, M. Isono, G. Tamura, K. Arima, *ibid.*, **33**, 1523 (1969); A. Kakinuma, A. Ouchida, T. Shima, H. Sugino, M. Isono, G. Tamura, K. Arima, *ibid.*, **33**, 1669 (1969); K. Hosono, H. Suzuki, *J. Antibiot.*, **36**, 667, 674 (1983).
3) J. Vater, *Progr. Cooloid Polymer Sci.*, **72**, 12 (1986).
4) Y. Kameda, K. Matsui, H. Kato, T. Yamada, H. Sagai, *Chem. Pharm. Bull.*, **20**, 1551 (1972); Y. Kameda, S. Ouhira, K. Matsui, S. Kamatomo, T. Hase, T. Atsusaka, *ibid.*, **22**, 938 (1974).
5) K. Hosono, H. Suzuki, *J. Antibiot.*, **36**, 679 (1983).
6) H. Morita, T. Yamamiya, K. Takeya, H. Itokawa, C. Sakuma, J. Yamada, T. Suga, *Chem. Pharm. Bull.*, **41**, 781 (1993); H. Morita, S. Nagashima, K. Takeya, H. Itokawa, *ibid.*, **41**, 992 (1993).
7) K. Oka, T. Hirano, M. Homma, H. Ishii, K. Murakami, S. Mogami, A. Motizuki, H. Morita, K. Takeya, H. Itokawa, *Chem. Pharm. Bull.*, **41**, 1000 (1993).
8) F. Peypoux, M.-T. Pommier, D. Marion, M. Ptak, B. C. Das, G. Michel, *J. Antibiot.*, **39**, 636 (1986); M. Genest, D. Arion, A. Caille, M. Ptak, *Eur. J. Biochem.*, **169**, 389 (1987).
9) F. Peypoux, M. Guinand, G. Michel, L. Delcambe, B. C. Das, E. Lederer, *Biochemistry*, **17**, 3992 (1978); F. Peypoux, D. Marion, R. Maget-D, M. Ptak, B. C. Das, G. Michel, *Eur. J. Biochem.*, **153**, 335 (1985).
10) P. Marfey, *Carlsberg Res. Commun.*, **49**, 591 (1984).
11) A. Bax, D. G. Davis, *J. Magn. Reson.*, **65**, 355 (1985).
12) G. Bodenhauser, H. Koger, R. R. Ernst, *J. Magn. Reson.*, **58**, 370 (1984).
13) A. Bax, S. Subramanian, *J. Magn. Reson.*, **67**, 565 (1986).
14) A. Bax, M. F. Summers, *J. Am. Chem. Soc.*, **108**, 2094 (1986).
15) H. Kessler, *Angew. Chem.*, **94**, 509 (1982).
16) S. R. Wilson, W. Cui, J. Moskowits, K. E. Schmidt, *Tetrahedron Lett.*, **29**, 4373 (1988).
17) O. S. Weislow, R. Kiser, D. L. Fine, J. Bader, R. H. Shoemaker, M. R. Boyd, *J. Natl. Cancer Inst.*, **81**, 577 (1989).
18) J. G. Vinter, A. Davis, M. R. Saunders, *J. Comput.-Aided Mol. Design*, **1**, 31 (1987); M. Clark, R. D. Cramer III, N. V. Opdembosch, *J. Comput. Chem.*, **10**, 982 (1989).