Journal of Biomolecular NMR, 5 (1995) 271–286
ESCOM

J-Bio NMR 245

High-resolution solution structure of siamycin II: Novel amphipathic character of a 21-residue peptide that inhibits HIV fusion

Keith L. Constantine[a,*], Mark S. Friedrichs[a], David Detlefsen[b], Maki Nishio[b], Mitsuaki Tsunakawa[b], Tamotsu Furumai[b], Hiroaki Ohkuma[b], Toshikazu Oki[c], Susan Hill[b], Robert E. Bruccoleri[a], Pin-Fang Lin[b] and Luciano Mueller[a]

[a]*Bristol-Myers Squibb Pharmaceutical Research Institute, P.O. Box 4000, Princeton, NJ 08543, U.S.A.*
[b]*Bristol-Myers Squibb Pharmaceutical Research Institute, P.O. Box 5100, Wallingford, CT 06492, U.S.A.*
[c]*Biotechnology Research Center, Toyama Prefectural University, Kosugi, Toyama 939-03, Japan*

Received 29 August 1994
Accepted 11 October 1994

Keywords: Amphipathic peptide; Anti-HIV peptide; Restrained energy minimization; Variable target function

Summary

The 21-amino acid peptides siamycin II (BMY-29303) and siamycin I (BMY-29304), derived from *Streptomyces* strains AA3891 and AA6532, respectively, have been found to inhibit HIV-1 fusion and viral replication in cell culture. The primary sequence of siamycin II is CLGIGSCNDFAGCGYAIV-CFW. Siamycin I differs by only one amino acid; it has a valine residue at position 4. In both peptides, disulfide bonds link Cys^1 with Cys^{13} and Cys^7 with Cys^{19}, and the side chain of Asp^9 forms an amide bond with the N-terminus. Siamycin II, when dissolved in a 50:50 mixture of DMSO and H_2O, yields NOESY spectra with exceptional numbers of cross peaks for a peptide of this size. We have used 335 NOE distance constraints and 13 dihedral angle constraints to generate an ensemble of 30 siamycin II structures; these have average backbone atom and all heavy atom rmsd values to the mean coordinates of 0.24 and 0.52 Å, respectively. The peptide displays an unusual wedge-shaped structure, with one face being predominantly hydrophobic and the other being predominantly hydrophilic. Chemical shift and NOE data show that the siamycin I structure is essentially identical to siamycin II. These peptides may act by preventing oligomerization of the HIV transmembrane glycoprotein gp41, or by interfering with interactions between gp41 and the envelope glycoprotein gp120, the cell membrane or membrane-bound proteins [Frèchet, D. et al. (1994) *Biochemistry*, **33**, 42–50]. The amphipathic nature of siamycin II and siamycin I suggests that a polar (or apolar) site on the target protein may be masked by the apolar (or polar) face of the peptide upon peptide/protein complexation.

Introduction

The majority of the drugs available for treating HIV infection, as well as most ongoing anti-AIDS drug discovery efforts, target the viral enzymes. Inhibition of these enzymes requires penetration of the cellular membrane by the inhibitor. Therefore, identification of compounds that interfere with the viral entry into cells could potentially provide powerful new therapeutic approaches for the prevention and treatment of AIDS.

Screening assays of *Streptomyces* metabolites isolated at the former Bristol-Myers Squibb Research Institute, Tokyo, revealed that two related polypeptides, siamycin II and siamycin I, inhibit the HIV-induced fusion between

*To whom correspondence should be addressed.

Abbreviations: ABNR, adopted basis Newton Raphson; AIDS, acquired immunodeficiency syndrome; CW, continuous wave; DMSO, dimethylsulfoxide; DQF-COSY, two-dimensional double-quantum-filtered correlation spectroscopy; HIV, human immunodeficiency virus; HSQC, heteronuclear single-quantum coherence; NOE, nuclear Overhauser enhancement; NOESY, two-dimensional nuclear Overhauser enhancement spectroscopy; ppm, parts per million; P.E.-COSY, two-dimensional primitive exclusive correlation spectroscopy; REDAC, redundant dihedral angle constraint; rf, radio frequency; rmsd, root-mean-square difference; SIV, simian immunodeficiency virus; sw, spectral width; τ_m, mixing time; TOCSY, two-dimensional total correlation spectroscopy; TSP, trimethylsilyl-2,2,3,3-$^2H_4$-propionate; 2D, two-dimensional.

gp160 expressing cells and Hela $CD4^+$ cells (Tsunakawa, M. et al., unpublished results). These peptides are also known by their Bristol-Myers Squibb compound numbers: BMY-29303 (siamycin II) and BMY-29304 (siamycin I). Further studies showed that the peptides also display inhibitory activity against the replication of the retroviruses HIV-1, HIV-2 and SIV. Both peptides protect CEM-SS cells against HIV-1 RF-strain acute infection with an ED_{50} of 0.1 μM and a therapeutic index of 500–1500 (Lin, P.-F. et al., unpublished results).

The primary structure of siamycin I was determined by a variety of methods, including amino acid composition analysis, mass spectrometry, chromatography and NMR spectroscopy (Detlefsen, D. et al., unpublished results). The primary structure of siamycin I was found to be CLGVGSCNDFAGCGYAIVCFW (one-letter amino acid code), with disulfide bonds between Cys^1 and Cys^{13} and between Cys^7 and Cys^{19}. The side chain of Asp^9 forms an amide bond with the N-terminus, leading to an unusual tricyclic covalent structure. A very similar anti-HIV peptide, RP-71955, has been isolated and characterized by researchers at Rhône–Poulenc Rorer (Frèchet et al., 1994). RP-71955 differs from siamycin I at two sites: Val^4 of siamycin I becomes Ile^4 in RP-71955 and Ile^{17} in siamycin I becomes Val^{17} in RP-71955.

Amino acid composition analysis revealed that siamycin II differs from siamycin I by only one amino acid: a valine residue in siamycin I is replaced by an isoleucine residue in siamycin II. Using 2D NMR methods, the ^1H NMR resonances of siamycin II have been assigned for the peptide dissolved in 100% DMSO at 25 °C, and dissolved in a 50:50 mixture of DMSO and H_2O at 20 °C and 35 °C. The sequence of siamycin II was established to be CLGIGSCNDFAGCGYAIVCFW. A series of NOESY spectra were recorded using the mixed solvent at 20 °C. An exceptionally well-defined solution structure ensemble has been determined using the variable target function minimization program DIANA (Güntert and Wüthrich, 1991; Güntert et al., 1991; Güntert, 1992) and restrained energy minimization using X-PLOR (Brünger, 1992) and CONGEN (Bruccoleri and Karplus, 1987; Bruccoleri, 1990). The final structures were determined using 13 dihedral angle constraints and 335 NOE distance constraints. A procedure for quantitatively deriving cross-relaxation rates from NOE buildup curves (see Materials and Methods) was applied to help improve the accuracy of the distance constraints used for modeling. Although the overall fold of siamycin II is similar to that reported for RP-71955 (Frèchet et al., 1994), several important differences in local regions of the structures are found. Inspection of the siamycin II structures reveals a novel amphipathic arrangement of side-chain and main-chain groups. Siamycin I is demonstrated to adopt an extremely similar conformation. These results have important implications for the biological activities of these peptides.

Materials and Methods

NMR sample preparation

The initial NMR sample of siamycin II contained 15 mg of the peptide dissolved in 0.55 ml of DMSO-d_6 (Cambridge Isotope Laboratories, Inc., Woburn, MA), resulting in a peptide concentration of ~14 mM. After acquiring NMR data for this sample, it was divided equally into two separate portions. To the first portion, 275 μl of 10 mM acetate buffer (pH = 4.6) in 100% H_2O was added and a corresponding buffer in 99.96% D_2O (Isotec, Inc., Miamisburg, OH) was added to the second portion. The acetate buffer was prepared with sodium acetate-d_3 and acetic acid-d_4 (Cambridge Isotope Laboratories, Inc.). A sample of siamycin I was prepared by dissolving 7.8 mg of the peptide in 0.55 ml of a 55:45 mixture of DMSO-d_6 and 10 mM acetate buffer (pH = 4.6) in H_2O.

NMR data acquisition and processing

All spectra were recorded on a Varian Unity 600 NMR spectrometer operating at 599.91 MHz proton frequency. The temperature was set to 25 °C for all experiments that were performed on the 100% DMSO-d_6 sample. For samples in the DMSO–water mixtures, data sets were collected at 20 °C and 35 °C.

For all 2D experiments, half-dwell sampling was employed in t_1 to improve baseline offset and flatness (Bax et al., 1991). DQF-COSY (Rance et al., 1983) and clean-TOCSY (Braunschweiler and Ernst, 1983; Bax and Davis, 1985; Griesinger et al., 1988) spectra were recorded. For the clean-TOCSY experiments, mixing times of 71 and 17 ms were used with a window period equal to the duration of a 90° pulse. The rf field strengths for hard and TOCSY pulses were set to 38.7 and 11.06 kHz, respectively. For each sample, NOESY spectra (Kumar et al., 1980; Macura and Ernst, 1980; Bodenhausen et al., 1984) with different mixing periods were collected. For the DMSO sample, a P.E.-COSY spectrum (Mueller, 1987) was recorded. Water suppression was achieved by low-power CW irradiation during the relaxation delays and the NOESY mixing periods. In addition, for the H_2O-containing samples, water peak suppression was achieved by very low power (6 Hz) CW irradiation (to avoid radiation damping in t_1) and by the WATERGATE suppression scheme (Piotto et al., 1992). Spectral widths of 8.0 kHz in each proton dimension were used for all 2D spectra. In general, the acquisition time was set to 0.16 s, followed by a 1.8 s recovery delay. For all NOESY experiments, 32 scans were accumulated per t_1 increment. For the TOCSY experiments, 16 or 32 scans were averaged per t_1 increment. However, for the freshly prepared DMSO-d_6 / D_2O sample, TOCSY spectra were recorded rapidly with eight scans per t_1 increment.

Cosine-bell and 60° shifted skewed sine-bell apodization was employed prior to Fourier transformation of the

TOCSY t_1 and t_2 data, respectively, and 45° shifted sine-bell apodization was applied to the DQF-COSY t_1 and t_2 data. For the NOESY spectra, the t_1 and t_2 data were subjected to exponential-to-Gaussian conversion using an exponential broadening of −7 Hz and a Gaussian broadening of 7 Hz. In addition, the t_1 data were multiplied by a cosine function to minimize truncation errors. The P.E.-COSY data were processed as described previously (Mueller, 1987; Marion and Bax, 1988). Zero-filling was applied to yield 2048×2048 spectral matrices for the NOESY and TOCSY spectra. The DQF-COSY and P.E.-COSY data matrices were zero-filled to 4096×4096 points. The ^{1}H chemical shifts were referenced to the residual DMSO resonance, which was assumed to resonate 2.57 ppm downfield relative to an external TSP standard at all temperatures.

Two standard ^{1}H-^{13}C HSQC spectra (Bodenhausen and Ruben, 1980) were recorded in 100% DMSO-d_6 with the ^{13}C carrier frequency set to 41 ppm (^{13}C-aliphatic) and 126 ppm (^{13}C-aromatic), respectively. A ^{13}C spectral width of 5.0 kHz and 64 t_1 increments were used. For both spectra half-dwell t_1 sampling was employed to improve baseline flatness and to identify aliased peaks (Bax et al., 1991), and a total of 112 scans were averaged per t_1 value. In the samples containing the DMSO/D_2O solvent mixture, two gradient-enhanced and sensitivity-enhanced ^{1}H-^{13}C HSQC spectra (Kay et al., 1992) were recorded, again with ^{13}C carrier settings at 41 and 126 ppm. The number of t_1 increments was 80 and the carbon spectral width was 5.0 kHz. A total of 112 scans were averaged to record the aromatic ^{1}H-^{13}C HSQC, and 256 scans were averaged per t_1 increment to record the aliphatic ^{1}H-^{13}C HSQC. All HSQC spectra were processed to yield matrices of 2048×2048 data points.

Constraint derivation

The constraints used for determining the solution structure of siamycin II were derived primarily from data recorded at 20 °C using the samples dissolved in 50% DMSO/50% H_2O. Backbone ϕ dihedral constraints were derived from $^3J_{NH\alpha}$ coupling constants, which were estimated from splittings observed in the DQF-COSY spectrum and from NH-H$^\alpha$ cross-peak intensities measured in a TOCSY spectrum with short mixing time (τ_m = 16 ms). The ϕ dihedral angles were constrained to $-120° \pm 50°$ when the $^3J_{NH\alpha}$ was estimated to be greater than 8.0 Hz and when the intraresidue HN-H$^\alpha$ distance was estimated to be > 2.6 Å. This latter consideration rules out $\phi \sim 60°$. Stereospecific H$^\beta$ and valine γ-methyl assignments, and χ^1 dihedral angle constraints, were obtained from H$^\alpha$-H$^\beta$ cross-peak splittings observed in the P.E.-COSY spectrum, and from analysis of peak intensity patterns in the NOESY and short mixing time (τ_m = 16 ms) TOCSY spectra (Constantine et al., 1994a). In cases where the data indicated one predominant χ^1 rotamer (~60°, ~−60° or ~180°), the dihedral angle was constrained to be within 40° of the indicated staggered conformation. Just one of the three favored χ^1 rotamers was excluded in some cases.

NOESY spectra were recorded for the DMSO/H_2O sample of siamycin II with mixing times of 30, 50, 90, 150 and 300 ms. In order to generate a set of preliminary structures, the volumes of unambiguously assigned cross peaks in the 90 ms NOESY spectrum were integrated using a modified version of the FELIX program (Hare Research, Inc.; Friedrichs, M., unpublished results). These were conservatively classified into strong, medium, weak and very weak NOE bins, which correspond to upper-bound interproton distance constraints of 2.5, 3.0, 4.0 and 6.0 Å, respectively. These calibrations were based on comparison of intraresidue HN-H$^\alpha$ NOE intensities with the allowed distance ranges for these interactions (Wüthrich, 1986). Preliminary models were used to assign additional NOESY cross peaks. In order to help account for spin diffusion effects and to obtain more accurate distance constraints, the NOESY time series data were fit using the cross-diagonal average of diagonal peak normalized NOE volumes (Andersen et al., 1989; Lai et al., 1993), modified to properly treat equivalent or degenerate protons:

$$[1/(n_i+n_j)][{}^gI_{ij}(\tau_m)/{}^gI_i(\tau_m)+{}^gI_{ji}(\tau_m)/{}^gI_j(\tau_m)] \approx -{}^i\sigma_{ij}\tau_m + B\tau_m^2 \quad (1)$$

On the left side of this expression, n_i and n_j are the number of protons in groups i and j (e.g., $n_i = 1$ for an amide proton, 2 for equivalent tyrosine H_2^ε protons, 3 for a resolved methyl group and 6 for a degenerate pair of methyl groups), ${}^gI_{ij}(\tau_m)$ and ${}^gI_{ji}(\tau_m)$ are the measured intensities of the ij and ji cross peaks involving group i and group j, and ${}^gI_i(\tau_m)$ and ${}^gI_j(\tau_m)$ are the measured intensities of the diagonal peaks of groups i and j. For the diagonal peaks, superscript ^{g} denotes that the entire group of protons is represented. The superscript ^{i} is used to represent quantities involving individual protons. For example, the diagonal peak intensity of a single methyl proton, ${}^iI_i(\tau_m)$, is equivalent to ${}^gI_i(\tau_m)/n_i$. On the right-hand side of Eq. 1, B is a collection of terms composed of two-step leakage interactions and spin diffusion involving one intervening group, and ${}^i\sigma_{ij}$ is an 'effective cross rate' (Lai et al., 1993) between an individual proton of group i and an individual proton of group j. The total group i ↔ group j cross rate, ${}^g\sigma_{ij}$, is given by

$${}^g\sigma_{ij} = n_i n_j {}^i\sigma_{ij} \quad (2)$$

Equation 1 is an empirical relationship that is similar to a second-order Taylor expansion of the relaxation matrix. It differs from a true Taylor expansion in that diagonal peak intensities at $\tau_m = 0$, ${}^gI_i(0)$, are replaced by

TABLE 1
NMR ASSIGNMENTS FOR SIAMYCIN II UNDER VARIOUS CONDITIONS

Residue	Chemical shift (ppm)[a]			
	NH	H^{α}	H^{β}	Others
100% DMSO, 25 °C				
Cys^{1}	8.59	4.64 (50.8)[b]	3.57, 2.62 (46.7)	
Leu^{2}	9.38	4.55	1.65, 2.22[c]	H^{γ} 2.30 (27.2); $H_3^{\delta 1}$ 1.11[c]; $H_3^{\delta 2}$ 0.98[c]
Gly^{3}	9.21	3.64, 4.23[c] (45.9)		
Ile^{4}	6.74	4.54	1.59 (42.3)	$H^{\gamma 1}$ 0.92[c]; $H^{\gamma 2}$ 1.46[c]; H_3^{γ} 0.82 (22.5); H_3^{δ} 0.83 (18.2)
Gly^{5}	7.96	4.51, 3.42[c]		
Ser^{6}	7.72	4.44	3.63, 3.79[c] (65.4)	
Cys^{7}	7.93	4.71 (54.9)	3.40, 3.24 (47.2)	
Asn^{8}	8.94	5.22 (53.9)	2.36, 2.27[c] (40.6)	H^{δ} 7.01, 6.87
Asp^{9}	7.99	4.45	2.84 (41.0)	
Phe^{10}	8.07	4.31 (57.2)	2.78 (42.1)	H_2^{δ} 7.31; H_2^{ε} 7.27 (129.6); H^{ζ} 7.31
Ala^{11}	8.80	3.94 (53.7)	1.21 (19.3)	
Gly^{12}	8.86	3.94, 3.47 (45.9)		
Cys^{13}	8.20	4.38	2.89, 3.50[c]	
Gly^{14}	7.34	3.04, 3.58[c] (46.7)		
Tyr^{15}	9.13	5.44 (57.1)	2.00, 2.87[c] (41.0)	H_2^{δ} 6.84 (133.6); H_2^{ε} 6.54 (117.4)
Ala^{16}	8.27	4.65	1.23 (23.6)	
Ile^{17}	8.14	4.50	2.00	$H^{\gamma 1}$ 1.12[c]; $H^{\gamma 2}$ 1.46[c]; H_3^{γ} 0.91; H_3^{δ} 0.87
Val^{18}	7.01	4.20 (61.1)	1.91 (33.4)	$H_3^{\gamma 1}$ 0.84[c]; $H_3^{\gamma 2}$ 0.90[c]
Cys^{19}	8.30	4.52	2.87, 2.61(45.9)	
Phe^{20}	7.68	4.67 (56.3)	2.64, 2.83[c] (41.1)	H_2^{δ} 6.96 (132.5); H_2^{ε} 7.10 (130.9); H^{ζ} 7.16 (129.3)
Trp^{21}	8.52	4.51	3.27, 3.05 (30.5)	$H^{\delta 1}$ 7.23 (127.1); $H^{\varepsilon 1}$ 10.94; $H^{\zeta 2}$ 7.41 (114.6); $H^{\eta 2}$ 7.15 (124.1); $H^{\zeta 3}$ 7.08 (121.6); $H^{\varepsilon 3}$ 7.64 (121.3)
50% DMSO / 50% H_2O, pH 4.6, 20 °C				
Cys^{1}	8.88	4.59	3.48, 2.56	
Leu^{2}	9.69	4.50	1.56, 2.10[c]	H^{γ} 1.99; $H_3^{\delta 1}$ 1.00[c]; $H_3^{\delta 2}$ 0.88[c]
Gly^{3}	9.06	3.67, 4.15[c]		
Ile^{4}	6.90	4.24	1.58	$H^{\gamma 1}$ 0.91[c]; $H^{\gamma 2}$ 1.28[c]; H_3^{γ} 0.73; H_3^{δ} 0.74
Gly^{5}	7.36	4.42, 3.48[c]		
Ser^{6}	7.83	4.53	3.71, 3.89[c]	
Cys^{7}	8.06	4.87	3.40, 3.04	
Asn^{8}	8.90	5.12	2.40, 2.26[c]	H^{δ} 7.22, 6.82
Asp^{9}	7.92	4.37	2.82	
Phe^{10}	8.50	4.27	2.89, 2.55	H_2^{δ} 7.25; H_2^{ε} 7.20; H^{ζ} 7.25
Ala^{11}	8.87	3.88	1.22	
Gly^{12}	8.77	3.98, 3.55		
Cys^{13}	–	4.25	3.06, 3.36[c]	
Gly^{14}	7.38	3.09, 3.70[c]		
Tyr^{15}	9.24	5.40	1.96, 2.70[c]	H_2^{δ} 6.71; H_2^{ε} 6.55
Ala^{16}	8.35	4.65	1.16	
Ile^{17}	8.24	4.22	1.75	$H^{\gamma 1}$ 1.00[c]; $H^{\gamma 2}$ 1.45[c]; H_3^{γ} 0.83; H_3^{δ} 0.77
Val^{18}	7.34	3.89	1.84	$H_3^{\gamma 1}$ 0.66[c]; $H_3^{\gamma 2}$ 0.80[c]
Cys^{19}	7.97	4.40	2.38, 2.25	
Phe^{20}	8.06	4.45	2.54, 2.81[c]	H_2^{δ} 6.84; H_2^{ε} 7.05; H^{ζ} 7.09
Trp^{21}	7.66	4.39	3.24, 3.03	$H^{\delta 1}$ 7.08; $H^{\varepsilon 1}$ 10.32; $H^{\zeta 2}$ 7.34; $H^{\eta 2}$ 7.08; $H^{\zeta 3}$ 7.01; $H^{\varepsilon 3}$ 7.54
50% DMSO / 50% H_2O, pH 4.6, 35 °C				
Cys^{1}	8.73	4.60	3.48, 2.57	
Leu^{2}	9.63	4.53	1.57, 2.10[c]	H^{γ} 1.99; $H_3^{\delta 1}$ 1.01[c]; $H_3^{\delta 2}$ 0.89[c]
Gly^{3}	8.98	3.67, 4.15[c]		
Ile^{4}	6.90	4.24	1.58	$H^{\gamma 1}$ 0.93[c]; $H^{\gamma 2}$ 1.29[c]; H_3^{γ} 0.73; H_3^{δ} 0.74
Gly^{5}	7.27	4.43, 3.48[c]		
Ser^{6}	7.81	4.53	3.72, 3.88[c]	
Cys^{7}	8.04	4.87	3.38, 3.08	
Asn^{8}	8.82	5.13	2.41, 2.31[c]	H^{δ} 7.15, 6.76
Asp^{9}	7.90	4.40	2.83	
Phe^{10}	8.43	4.28	2.89, 2.62	H_2^{δ} 7.25; H_2^{ε} 7.19; H^{ζ} 7.25
Ala^{11}	8.74	3.89	1.22	

TABLE 1
(continued)

Residue	Chemical shift (ppm)[a]			
	NH	H^{α}	H^{β}	Others
Gly^{12}	8.74	3.88, 3.48		
Cys^{13}	–	4.36	3.06, 3.32[c]	
Gly^{14}	7.37	3.12, 3.71[c]		
Tyr^{15}	9.22	5.40	1.96, 2.70[c]	H_2^{δ} 6.71; H_2^{ε} 6.56
Ala^{16}	8.34	4.65	1.18	
Ile^{17}	8.20	4.21	1.76	$H^{\gamma 1}$ 1.02[c]; $H^{\gamma 2}$ 1.45[c]; H_3^{γ} 0.84; H_3^{δ} 0.77
Val^{18}	7.29	3.91	1.85	$H_3^{\gamma 1}$ 0.68[c]; $H_3^{\gamma 2}$ 0.81[c]
Cys^{19}	7.90	4.40	2.42, 2.28	
Phe^{20}	8.00	4.46	2.57, 2.83[c]	H_2^{δ} 6.86; H_2^{ε} 7.07; H^{ζ} 7.10
Trp^{21}	7.56	4.39	3.23, 3.03	$H^{\delta 1}$ 7.08; $H^{\varepsilon 1}$ 10.26; $H^{\zeta 2}$ 7.34; $H^{\eta 2}$ 7.08; $H^{\zeta 3}$ 7.01; $H^{\varepsilon 3}$ 7.54

[a] The ^1H chemical shift values (±0.02 ppm) were referenced relative to the residual DMSO signal, which was assumed to resonate 2.57 ppm downfield of TSP.

[b] The values given in parentheses following ^1H chemical shift values are the ^{13}C chemical shifts of the attached carbon, referenced indirectly to TSP (Fairbrother et al., 1992).

[c] Stereospecific assignment. For the H^{β} resonances, the first chemical shift is for $H^{\beta 2}$ (pro-*R*), and the second is for $H^{\beta 3}$ (pro-*S*). The glycine H^{α} chemical shifts are given in the order $H^{\alpha 1}$, $H^{\alpha 2}$.

${}^{g}I_i(\tau_m)$. This relation has been shown to provide more accurate distance estimates than a second-order Taylor expansion (Andersen et al., 1989; Hyberts and Wagner, 1989; Baleja et al., 1990; Lai et al., 1993). Effective distances were derived from

$$r_{ij} = r_{cal}({}^{g}\sigma_{cal} / {}^{g}\sigma_{ij})^{1/6} \quad (3)$$

where r_{cal} and ${}^{g}\sigma_{cal}$ are a reference distance and cross rate, respectively. For this study, the effective cross rate (–1.3 s^{-1}) between the Trp^{21} $H^{\delta 1}$ and $H^{\varepsilon 1}$ protons (r_{cal} = 2.5 Å) was used. This calibration was found to be consistent with the range of cross rates found for other intraresidue cross peaks. Corrections of (at least) +10% and –20% were applied to the r_{ij} in order to establish upper and

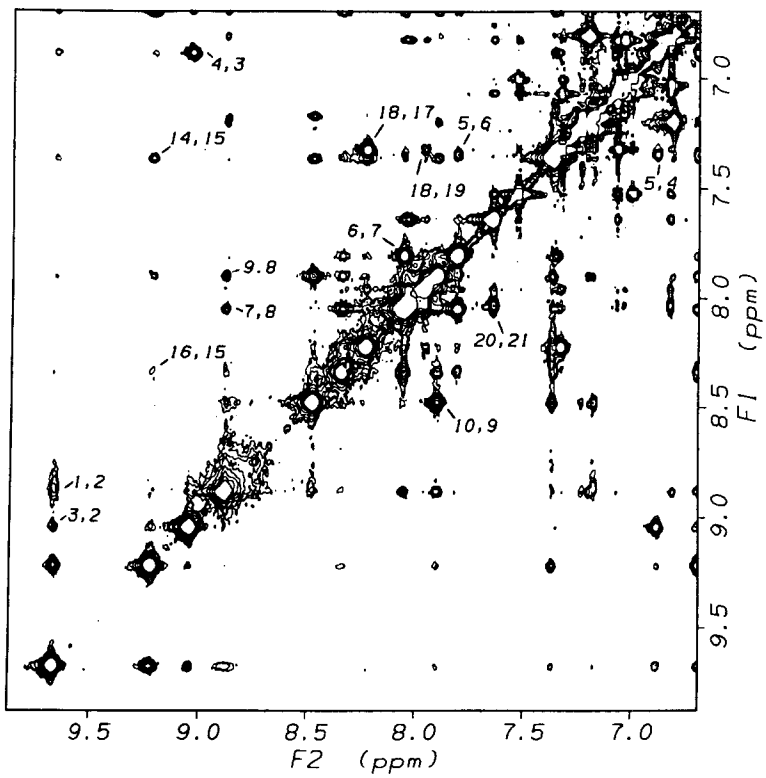


Fig. 1. Amide–amide region of the NOESY spectrum of siamycin II. Sequential NH-NH cross peaks are labeled res1-res2, where the first residue number corresponds to the F1 NH resonance and the second residue number to the F2 NH resonance. The spectrum was recorded with τ_m = 90 ms at 20 °C. The peptide concentration was approximately 7 mM in H_2O/DMSO (1:1, v/v), 10 mM sodium acetate, pH 4.6.

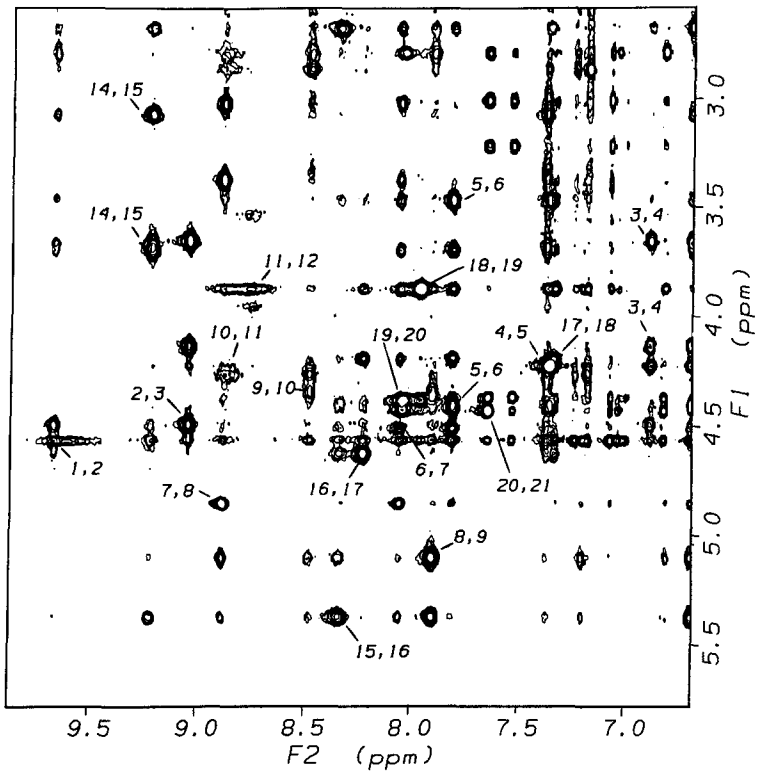


Fig. 2. Fingerprint region of the NOESY spectrum of siamycin II. Sequential H^{α}-NH cross peaks are labeled res1-res2, where the first residue number corresponds to the H^{α} (F1) resonance and the second residue number to the NH (F2) resonance. Conditions are the same as for Fig. 1.

lower bound distance constraints, respectively, with no lower bounds being allowed to exceed 3.5 Å. The $\tau_m = 0$, ${}^{g}I_{ij}(\tau_m) = 0$ point was included in all fits. The NOESY data with $\tau_m = 300$ ms were excluded from all cross-rate fits, since many cross peaks begin to decay before 300 ms. If ${}^{g}I_{ij}(\tau_m)$ was measured but ${}^{g}I_{ji}(\tau_m)$ was not, then ${}^{g}I_{ji}(\tau_m)$ was set equal to ${}^{g}I_{ij}(\tau_m)$. If ${}^{g}I_{i}(\tau_m)$ and/or ${}^{g}I_{j}(\tau_m)$ were not measured, the corresponding class average volume (Lai et al., 1993) was used.

Molecular modeling

Models of the siamycin II solution structure were generated with the variable target function program DIANA (Güntert et al., 1991). NOE distance constraints were used with appropriate pseudoatom distance corrections (Wüthrich et al., 1983) for equivalent and non-stereospecifically assigned protons. Structures were generated using one cycle of the REDAC protocol (Güntert and Wüthrich, 1991) with the standard minimization parameters (Güntert, 1992), with the exception that the local target function acceptance cutoff for generating redundant constraints was increased from 0.4 to 0.6 Å^2. Accepted DIANA structures were subjected to restrained energy minimization using the X-PLOR Powell routine (Brünger, 1992) and the CONGEN (Bruccoleri and Karplus, 1987; Bruccoleri, 1990,1993) adopted basis Newton Raphson (ABNR) algorithm (Brooks et al., 1983). Special patches were constructed to define the amide bond between the Asp[9] side chain and the N-terminus. Instead of using pseudoatoms, distance constraints involving equivalent and non-stereospecifically assigned protons were incorporated as $(\Sigma r^{-6})^{-1/6}$ effective distances (Constantine et al., 1992, 1994b). Constraints involving the degenerate Ile[4] H^{γ}_3 and H^{δ}_3 resonances (Table 1) were also treated in this manner. This approach allows the use of tighter constraints, and

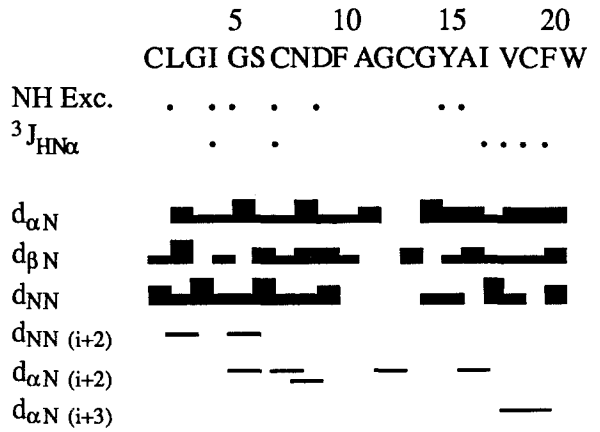


Fig. 3. Summary of the sequential and medium-range backbone–backbone NOEs observed for siamycin II. Strong, medium and weak sequential NOEs are indicated by thick, medium and thin solid horizontal bars between adjacent residues, respectively. Thin horizontal lines denote medium-range NOE connectivities. Residues with slowly exchanging backbone NH protons or with ${}^{3}J_{NH\alpha} > 8$ Hz are indicated by filled circles in the first and second rows below the sequence, respectively.

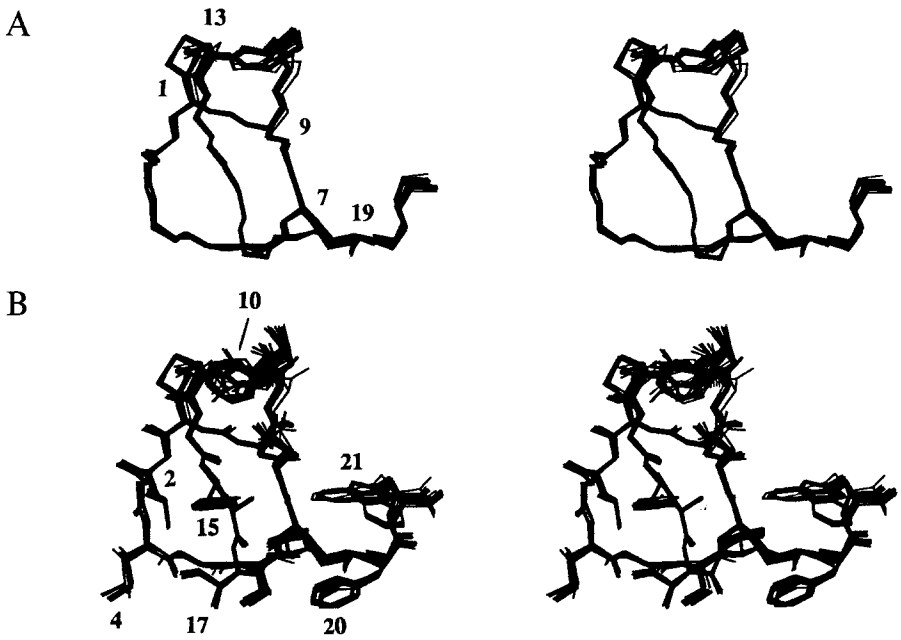


Fig. 4. Stereoviews of the 30 final energy-minimized siamycin II structures. (A) Diagram showing all backbone N, C, and C^{α} atoms, the disulfide bridges and the Asp^{9} C^{β} and C^{γ} atoms of the amide bond between the Asp^{9} side chain and Cys^{1} N. Residues 1, 7, 9, 13 and 19 are labeled. (B) Diagram showing all heavy atoms. Selected side chains are labeled.

it is fully consistent with distances derived directly from the ${}^{g}\sigma_{ij}$ (see Eqs. 1–3). Experimental constraints, electrostatics, van der Waals, harmonic bond length and harmonic bond angle terms were included in all minimization steps. Empirical dihedral angle functions were excluded. The dielectric parameter was set numerically equal to one

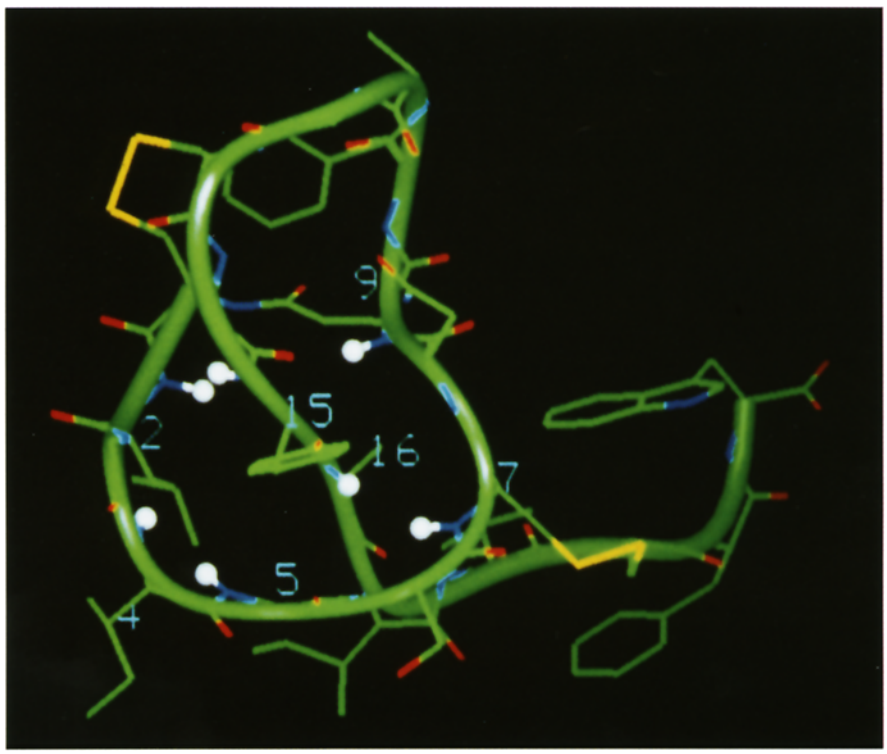


Fig. 5. Schematic illustration of a representative siamycin II structure, showing the locations of the slowly exchanging backbone amide hydrogens. The peptide backbone is traced with a solid green ribbon, and side chains are displayed in a stick representation, with carbons shown in green, oxygens in red, nitrogens in blue and sulfurs in yellow. The only hydrogens shown (white balls) are those that exhibit slow H–D exchange; the corresponding residues are labeled.

times (CONGEN) or two times (X-PLOR) the internuclear separation in Å, and the cutoff distance for nonbonded interactions was set to 8.0 Å. NOE distance constraint and dihedral angle constraint force constants were set to 50 kcal/mol Å^2 and 60 kcal/mol rad^2, respectively. Calculations were performed on Silicon Graphics 4D/440 VGX or Onyx computers or on a CRAY YMP2 supercomputer. Molecular graphics analysis was performed with the Insight II program (Biosym Technologies, San Diego, CA) and with GRASP (Nicholls et al., 1991).

Results and Discussion

Spin system identification and sequential assignment of siamycin II

An initial NMR study of siamycin II was performed at 25 °C using a ~14 mM sample of the peptide dissolved in 100% DMSO. Spin systems were identified by analysis of the TOCSY and DQF-COSY spectra in the usual manner (Wüthrich, 1986). Two natural-abundance ^{13}C-edited 2D HSQC spectra were also recorded to assist spin system identification. One leucine, one valine, two isoleucine, two alanine, four glycine and 11 {NH, H^{α}, H^{β}, $H^{\beta'}$} spin systems were identified. In addition, the aromatic ring spin systems of one tryptophan, one tyrosine and two phenylalanine residues were identified, along with the side-chain NH_2 of the asparagine residue. These five side-chain spin systems were connected to their corresponding backbone {NH, H^{α}, H^{β}, $H^{\beta'}$} spin systems by observation of intraresidue NOE cross peaks in the NOESY spectra. The serine spin system was identified by the downfield locations of its H^{β} (3.63 ppm), $H^{\beta'}$ (3.79 ppm) and C^{β} (65.4 ppm) resonances. Based on the composition analysis (Detlefsen, D. et al., unpublished results), the five remaining unidentified {NH, H^{α}, H^{β}, $H^{\beta'}$} spin systems include four cysteine residues and one aspartate residue.

As a working hypothesis, the primary structure of siamycin II was assumed to be similar to those of siamycin I (Detlefsen, D. et al., unpublished results) and RP-71955 (Frèchet et al., 1994), with one of the valine residues replaced by an isoleucine residue in siamycin II. An analysis of sequential NOE connectivities confirmed that the primary sequence of siamycin II is CLGIGSCNDFAGC-GYAIVCFW. Siamycin II differs from siamycin I at position 4, and it differs from RP-71955 at position 17. In addition, NOE cross peaks were observed between the NH resonance of Cys^1 and the H^{β} and H^{α} resonances of Asp^9, confirming that siamycin II contains an amide bond between the side chain of Asp^9 and the N-terminus. NOEs were also observed between the H^{α} resonance of Cys^1 and the H^{β} resonances of Cys^{13}, confirming the Cys^1-Cys^{13} and (by elimination) the Cys^7-Cys^{19} disulfide pairings.

In 100% DMSO, the resonances of Cys^1, Ala^{11}, Gly^{12} and Cys^{13} are relatively weak and broad. Therefore, we decided to explore different solvent systems. Using half of the original DMSO sample, a ~7 mM sample in 50% DMSO/50% H_2O was prepared. While the resonances of Cys^1, Ala^{11}, Gly^{12} and Cys^{13} remained rather weak, in general the spectra were much improved in the mixed solvent system. In particular, NOESY spectra for the mixed solvent displayed a large number of additional cross peaks relative to the 100% DMSO data. The NH-NH and NH-H^{α} regions of a NOESY spectrum with $\tau_m = 90$ ms, recorded at 20 °C using the mixed solvent system, are shown in Figs. 1 and 2, respectively, with sequential NOE cross peaks labeled. The Cys^1 H^{α} (4.59 ppm) resonance overlaps with the H_2O resonance at 20 °C. Therefore, NOEs involving the Cys^1 H^{α} resonance were confirmed by inspection of NOESY spectra recorded at 35 °C (not shown). Figure 3 summarizes the observed sequential and medium-range backbone–backbone NOEs; residues with $^3J_{NH\alpha}$ estimated to be >8 Hz or with a slowly exchanging NH proton are also indicated. The H^{β} protons of Tyr^{15} and the γ-methyl groups of Val^{18} were stereospecifically assigned by inspection of the NMR data. These were the only stereospecific assignments used for the structure determination (see below). Additional stereoassignments were obtained subsequent to the structure determination on the basis of the structural models and the derived cross-relaxation rates. The resonance assignments for siamycin II are reported in Table 1.

Siamycin II solution structure determination

For the NOESY spectrum with $\tau_m = 90$ ms, recorded at 20 °C using the mixed solvent system, 885 cross peaks and 32 resolved diagonal peaks were picked and integrated. This is an exceptionally large number of cross peaks for a 21-residue peptide, indicating an unusually compact conformation. An initial analysis of this spectrum yielded 251 unambiguous, conformationally relevant NOE distance constraints. These constraints were derived qualitatively by classifying the cross peaks into bins. The initial set of distance constraints included 45 intraresidue, 88 sequential, 25 medium-range and 93 long-range NOEs. Also, six ϕ and seven χ^1 dihedral angle constraints were obtained.

TABLE 2
STRUCTURAL STATISTICS FOR THE 30 FINAL ENERGY-MINIMIZED SIAMYCIN II STRUCTURES

⟨Backbone N, C and C^{α} atom rmsd to mean⟩	0.24 ± 0.07 Å
⟨All heavy atom rmsd to mean⟩	0.52 ± 0.07 Å
⟨Rms distance constraint violation⟩	0.092 ± 0.002 Å
⟨Number of distance constraint violations >0.3 Å⟩	3.8 ± 1.3
⟨Rms dihedral angle constraint violation⟩	0.03° ± 0.01°
⟨Bond length deviation from ideality[a]⟩	0.010 ± 0.001 Å
⟨Bond angle deviation from ideality⟩	2.56° ± 0.02°
⟨Improper torsion deviation from ideality⟩	1.19° ± 0.02°

[a] Deviations from ideal values refer to those in the X-PLOR parallhsa.pro parameter set.

TABLE 3
AVERAGE ϕ AND ψ ANGLES AND LOCAL RMSD VALUES FOR THE FINAL ENERGY-MINIMIZED SIAMYCIN II STRUCTURES

Residue	ϕ (°)	ψ (°)	Backbone rmsd[a] (Å)	Heavy-atom rmsd (Å)
1			0.17	0.68
2	−69 ± 2	−74 ± 3	0.13	0.12
3	129 ± 1	5 ± 1	0.15	0.19
4	−134 ± 4	26 ± 37	0.21	0.54
5	−55 ± 36	−149 ± 2	0.19	0.22
6	−150 ± 2	14 ± 2	0.14	0.31
7	−143 ± 2	76 ± 3	0.11	0.15
8	−13 ± 2	111 ± 1	0.11	0.43
9	−53 ± 2	−59 ± 2	0.13	0.14
10	−161 ± 7	132 ± 48	0.25	0.43
11	62 ± 44	146 ± 72	0.46	1.03
12	−43 ± 71	28 ± 54	0.70	1.40
13	−73 ± 51	−47 ± 10	0.37	0.57
14	−167 ± 12	104 ± 7	0.14	0.14
15	−158 ± 2	−174 ± 2	0.09	0.14
16	−159 ± 1	157 ± 2	0.07	0.08
17	−102 ± 1	−3 ± 1	0.08	0.15
18	−95 ± 3	149 ± 4	0.09	0.14
19	−165 ± 6	34 ± 1	0.12	0.15
20	−110 ± 4	4 ± 3	0.20	0.32
21			0.33	1.45

[a] Local rmsd values were computed after global fitting of the backbone atoms of residues 1–21.

The initial constraint set was employed to generate 80 structures using the DIANA program (Güntert et al., 1991) with one cycle of REDAC (Güntert and Wüthrich, 1991). The amide bond between the Asp^9 side chain and the N-terminus was not explicitly defined; an upper-bound distance constraint (2.7 Å) between the Asp^9 C^γ and Cys^1 N was incorporated to approximately position the Asp^9 side chain. From this DIANA run, 34 structures were selected which had a DIANA target function < 3.00 $Å^2$. These structures were then subjected to 2000 steps of restrained Powell minimization using X-PLOR (Brünger, 1992). Twenty-one structures were successfully minimized, whereas the minimization failed to properly form the Asp^9 side chain–Cys^1 amide bond in 13 structures. Of the 21 structures, 16 that contained no NOE violations > 0.5 Å were chosen as a preliminary ensemble of siamycin II structures. This ensemble has an average backbone N, C and C^α atom rmsd to the mean coordinates of 0.83 Å; the average value for all heavy atoms is 1.49 Å.

The preliminary ensemble was searched in order to assign additional NOE cross peaks. The analysis of the preliminary structures yielded an additional 63 NOE constraints for a total of 314 distance constraints, including 49 intraresidue NOEs, 109 sequential NOEs, 36 medium-range NOEs and 120 long-range NOEs. Instead of using a qualitative classification of constraints into bins, the set of NOESY spectra recorded with mixing times of 30, 50, 90 and 150 ms were analyzed by a NOESY time series fit (see Materials and Methods) in order to obtain more precise distance contraints. In addition to the +10% that was automatically added to the upper bounds, additional corrections of 0.5 to 2.0 Å were added to constraints involving overlapping cross peaks or cross peaks that did not appear in the 30 ms NOESY spectrum.

Using DIANA, 80 new structures were calculated using 13 dihedral angle constraints and the 314 refined distance constraints. Of these, 38 with target functions < 2.00 $Å^2$ were selected for additional refinement. These 38 structures have average backbone atom and all heavy atom rmsd values to the mean structure of 0.31 and 0.69 Å, respectively, demonstrating a large improvement in structural resolution over the preliminary structures. As noted above, the X-PLOR Powell minimization routine often failed to correctly establish the amide bond between the Asp^9 side chain and the N-terminus; therefore, the DIANA structures were first subjected to a 1000-step energy minimization using the more robust ABNR energy minimization routine of CONGEN. This resulted in properly formed Asp^9 side chain–Cys^1 N amide bonds for 37 of the 38 structures. The 37 structures were then reminimized with X-PLOR (1000 Powell steps) and 30 structures with no constraint violations greater than 0.5 Å were chosen as the penultimate set of siamycin II structures. These 30 structures have average backbone atom and all heavy atom rmsd values to the mean structure of 0.26 and 0.54 Å, respectively.

The high resolution of the penultimate structure set afforded several new NOE constraints, including one intraresidue constraint, six sequential constraints, eight medium-range constraints and six long-range constraints, thus increasing the total number of NOE constraints to 335. All constraints are available from the authors upon request. The penultimate structure set was subjected to a final 1000-step X-PLOR restrained energy minimization, resulting in a final set of 30 structures with no constraint violations > 0.5 Å. This set has average backbone atom and all heavy atom rmsd values to the mean structure of 0.24 and 0.52 Å, respectively. The structures changed only slightly during the final energy minimization – the backbone atom and all heavy atom rmsd values between the mean coordinates of the penultimate and final structure sets are 0.09 and 0.12 Å, respectively. Structural statistics for the 30 final structures are given in Table 2.

In order to help judge the accuracy of the final structure set, the entire set was searched for proton–proton distances less than 3.5 Å. For interactions involving equivalent groups, the $(\Sigma r^{-6})^{-1/6}$ effective distance was used in the search. In nearly all cases where distances less than 3.5 Å were found, the proton–proton pair was either associated with an assigned cross peak, or an intensity above the noise floor was observed in the expected loca-

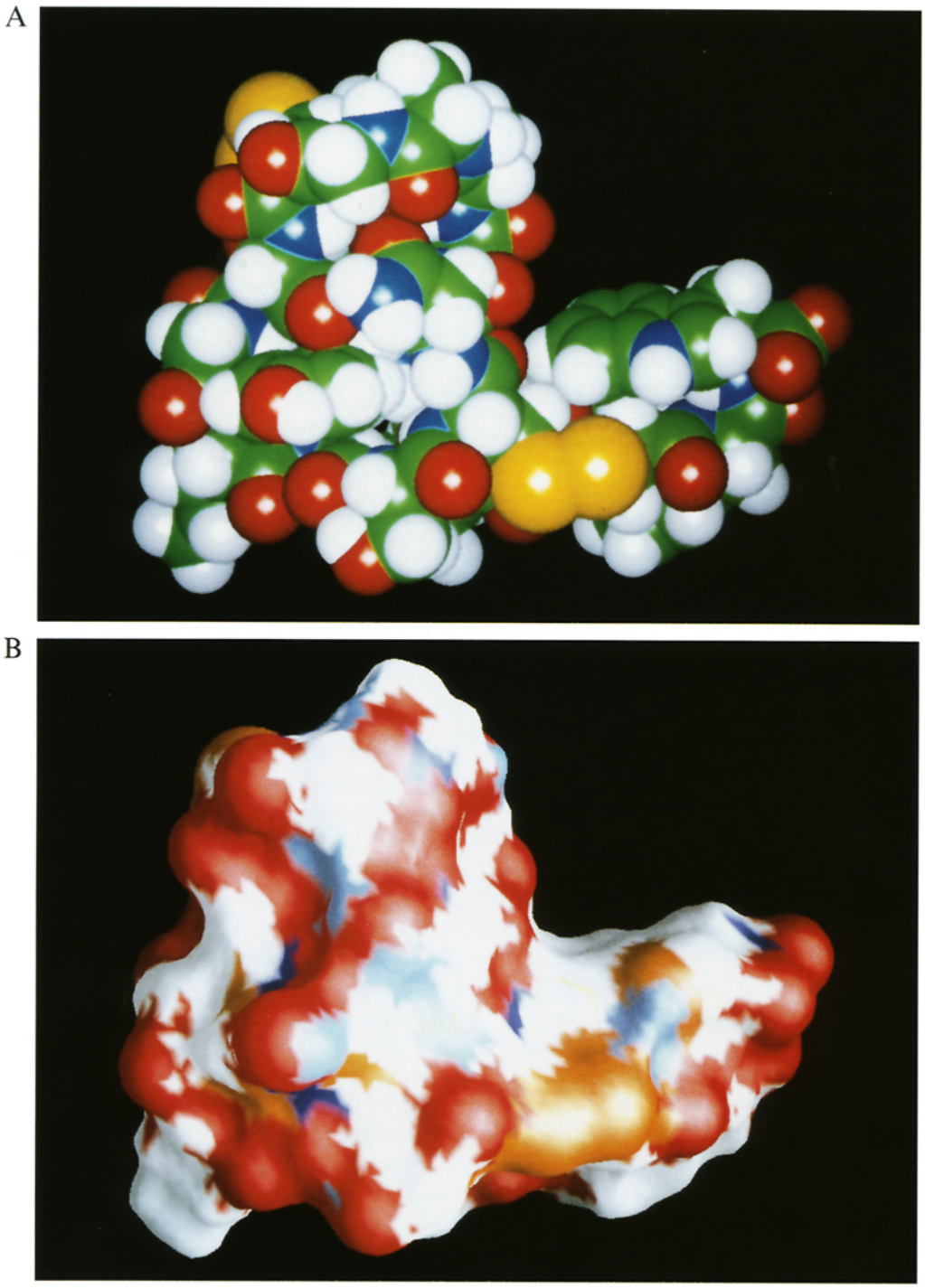
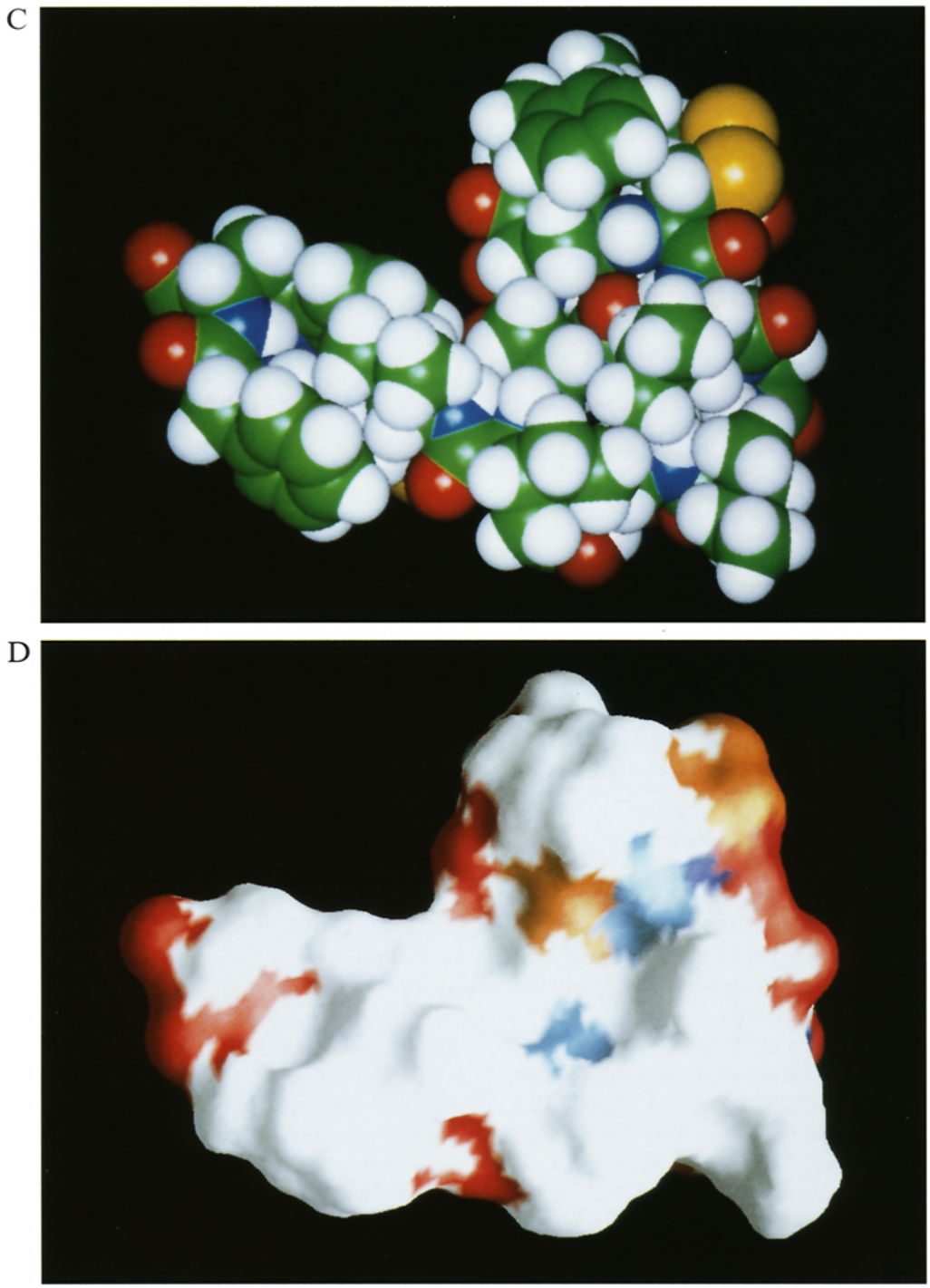


Fig. 6. Diagrams displaying surface features of a representative siamycin II structure. (A) Spacefilling depiction, showing the predominantly polar face, generated using InsightII (Biosym Technologies, San Diego, CA). Carbons are shown in green, hydrogens in white, nitrogens in blue, oxygens in red and sulfurs in yellow. (B) GRASP (Nicholls et al., 1991) depiction, showing the predominantly polar face. The surface is color coded according to atomic partial charges given in the AMBER all-atom potential (Weiner et al., 1986): partial charges >0.4 are shown in dark blue;

tion in the 90 ms NOESY. The only exceptions were proximities present in some structures between the Asn^{8} side-chain protons and the Gly^{12} NH and H^{α} protons, and between Phe^{10} H^{α} and Gly^{12} NH. The Gly^{12} resonances are broadened and weakened by a conformational exchange process. Therefore, transient conformations that do not give rise to detectable NOEs cannot be ruled out for interactions involving Gly^{12}.

Siamycin II solution conformation

Figure 4A shows a stereoview of the superimposed backbone N, C and C^{α} atoms of the 30 final siamycin II



those between 0.2 and 0.4 are shown in light blue; those between –0.2 and 0.2 are shown in white; those between –0.4 and –0.2 are shown in orange; and those <–0.4 are shown in red. (C) Spacefilling depiction, showing the predominantly apolar face. The color coding is the same as for A. (D) GRASP depiction, showing the predominantly apolar face. The color coding is the same as for B.

structures. The disulfide bridges and the amide linkage between the Asp^9 side chain and the N-terminus are also shown. In general, the polypeptide backbone and bridge conformations are very well defined. Table 3 lists the average ϕ and ψ angles, their standard deviations, and local rmsd values for the backbone atoms and all heavy atoms for individual residues. Residues 11 through 13 show higher local backbone atom rmsd values than the rest of the structure. Relatively few constraints were obtained for this region. As noted previously, many resonances in this region are weak and broad due to a conformational exchange process that occurs approximately on the millisecond time scale. It is noteworthy that varied conformations were found for the same region in models

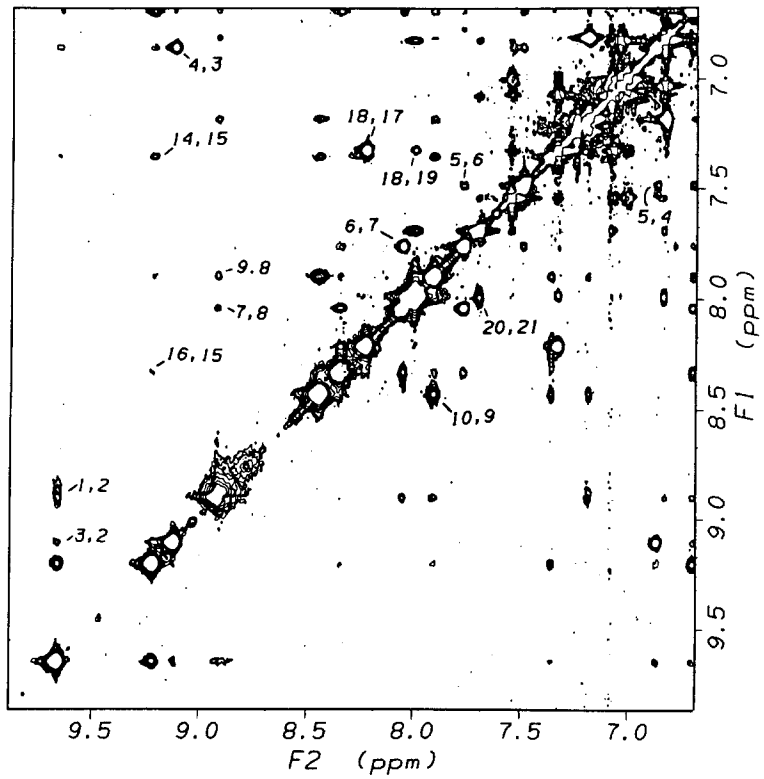


Fig. 7. Amide–amide region of the NOESY spectrum of siamycin I. Sequential NH-NH cross peaks are labeled res1-res2, where the first residue number corresponds to the F1 NH resonance and the second residue number to the F2 NH resonance. The spectrum was recorded with $\tau_m = 90$ ms at 20 °C. The peptide concentration was approximately 7 mM in H_2O/DMSO (45:55, v/v), 10 mM sodium acetate, pH 4.6.

of the RP-71955 peptide studied by the Rhône–Poulenc Rorer group (Frèchet et al., 1994). The orientation of the amide plane between Ile[4] and Gly[5] also displays variability. In 27 structures Ile[4] ψ lies between −33° and −30° and Gly[5] ϕ between −52° and −48°, while the remaining three structures (17, 18 and 25) have Ile[4] ψ between 86° and 91° and Gly[5] ϕ between −170° and −164°.

The polypeptide backbone conformations of residues 19–21 are well defined in the siamycin II structures (Table 3). In contrast, this region was found to be highly disordered in RP-71955 (Frèchet et al., 1994). Long-range NOEs are observed between Trp[21] ring protons and Cys[7] H^{α} and H^{β} protons for siamycin II. These NOE constraints, which greatly restrict the conformational space accessible to the C-terminal tail, apparently were not observed in the case of RP-71955. The degree of conformational disorder in the C-terminus may be sensitive to the different solvent conditions used for siamycin II and RP-71955.

In siamycin II, the disulfide bridge between Cys[7] and Cys[19] is well defined, with the Cys[7] χ^1 angle falling in the range −90° to −76°, and the Cys[19] χ^1 angle in the range −142° to −132°. On the other hand, multiple conformations are observed for the Cys[1]-Cys[13] disulfide bond. Three families of conformers are observed for Cys[1] χ^1, with 15 structures having χ^1 between −108° and −95°, 13 structures with χ^1 between 168° and −175°, and two structures with χ^1 of −51° and −50°. Three families of conformers are also observed for Cys[13] χ^1, with χ^1 between 133° and 169° for 15 structures, χ^1 between −166° and −150° for 13 structures, and a χ^1 of −31° for two structures. Inspection of the various cysteine χ^1 angles reported above reveals that some of them deviate by more than 30° from the staggered conformations. This may be due in part to the fact that empirical dihedral angle functions were excluded from the force field; i.e., the staggered rotamer conformations were not explicitly enforced. In any case, given the highly compact, knot-like conformation obtained (see below), locally distorted dihedral angles cannot be ruled out.

The overall fold of siamycin II is rather unusual. Although most of the backbone is well defined, there are no identifiable 'standard' β-turns (Richardson, 1981; Rose et al., 1985; Robson and Garner, 1986). The backbone conformation is composed of strands and irregular bends and turns. In the view shown in Fig. 4A, Cys[1] is near the upper left corner of the peptide. Cys[1], Leu[2] and Gly[3] form the left edge of the peptide that leads into a loop at the bottom of the structure containing Ile[4], Gly[5] and Ser[6]. Cys[7] through Phe[10] form the right edge of the peptide. The central cyclic ring (residues 1 through 9) is closed by the Asp[9] side chain–Cys[1] N amide bond. Residues Ala[11], Gly[12] and Cys[13] compose a disordered turn at the top of the peptide. This is followed by a central strand (Gly[14],

Tyr[15] and Ala[16]) that passes in front of the Asp[9]-Cys[1] amide bridge and exits behind the lower right corner of the central ring, forming a knot-like threading of the peptide backbone. Ile[17] and Val[18] form a bend that leads to the C-terminal tail (residues Cys[19], Phe[20] and Trp[21]).

Figure 4B shows a stereoview of the 30 final siamycin II structures, displaying all heavy atoms. Most of the free (nonbridging) side-chain conformations are well defined, especially those of Leu[2], Tyr[15], Ile[17] and Val[18] (Table 3). The Tyr[15] side chain forms the core of the central cyclic ring. In the case of the Trp[21] ring, two distinct conformations are found. In 17 structures, the χ^2 angle is between 74° and 77°, while in the 13 remaining structures this angle is between −109° and −110°. Many of the NOE constraints involving the Trp[21] ring were interpreted loosely, since this side chain is at the end of the C-terminal tail and is therefore likely to be flexible. This could result in NOEs that reflect transiently populated conformers. Also, many of the NOE constraints involving the Trp[21] ring are very weak or absent in the 30 ms NOESY spectrum. For example, an NOE between the Trp[21] $H^{\epsilon 3}$ and Phe[20] H^{α} is very weak in this spectrum. Direct fitting of the NOE buildup for this interaction yielded an upper bound of 4.22 Å. This value was increased by 1.5 Å to allow for possible dynamic and spin diffusion effects. If this constraint had not been loosened, only the $\chi^2 \sim -109°$ class of conformers would satisfy the constraint. However, use of the tighter constraint would probably be an overinterpretation of the data. Both the $\chi^2 \sim 75°$ and the $\chi^2 \sim -109°$ conformers are consistent with the loose constraints. Additional dynamic studies will be necessary to determine if the Trp[21] side chain adopts a unique conformation, or if it indeed samples multiple conformational states.

The overall stability of the conformation was examined by performing molecular dynamics simulations in vacuum. A representative NMR model was used as the starting structure. Additional details of the simulations are given in the Supplementary Material (available upon request from the authors). An 80 ps restrained simulation was performed at 293 K with the NOE distance constraint and dihedral angle constraint force constants set to 5 kcal/mol Å^2 and 6 kcal/mol rad^2, respectively. Structures were saved every 4 ps, yielding 20 snapshots (Fig. S1 of the Supplementary Material). The average backbone atom and all heavy atom rmsd values between the starting structure and the snapshots are 0.52 and 0.69 Å, respectively, and they reduce to 0.40 and 0.61 Å, respectively, if residues 19–21 are excluded. Thus, even with mild restraints, relatively high precision is maintained, particularly if the three C-terminal residues are not considered. Two 80 ps unrestrained simulations were performed: one at 293 K and one at 500 K. Twenty snapshots were saved from each run. For residues 1–18, the average backbone

TABLE 4
NMR ASSIGNMENTS FOR SIAMYCIN I AT 20 °C IN 45% H_2O (pH 4.6), 55% DMSO

Residue	Chemical shift (ppm)[a]			
	NH	H^{α}	H^{β}	Others
Cys[1]	8.94	4.58	3.50, 2.57	
Leu[2]	9.67	4.51	1.57, 2.14[b]	H^{γ} 2.01; $H_3^{\delta 1}$ 1.02[b]; $H_3^{\delta 2}$ 0.89[b]
Gly[3]	9.13	3.67, 4.17[b]		
Val[4]	6.88	4.27	1.82	$H_3^{\gamma 1}$ 0.71[b]; $H_3^{\gamma 2}$ 0.76[b]
Gly[5]	7.50	4.45, 3.47[b]		
Ser[6]	7.78	4.52	3.69, 3.88[b]	
Cys[7]	8.06	4.86	3.41, 3.07	
Asn[8]	8.92	5.13	2.39, 2.24[b]	H^{δ} 7.20, 6.83
Asp[9]	7.92	4.37	2.81	
Phe[10]	8.45	4.28	2.88, 2.56	H_2^{δ} 7.26; H_2^{ϵ} 7.20; H^{ζ} 7.25
Ala[11]	8.90	3.88	1.21	
Gly[12]	8.77	3.98, 3.55		
Cys[13]	7.37	4.30	3.06, 3.37[b]	
Gly[14]	7.37	3.07, 3.68[b]		
Tyr[15]	9.23	5.40	1.96, 2.70[b]	H_2^{δ} 6.85; H_2^{ϵ} 6.56
Ala[16]	8.35	4.64	1.17	
Ile[17]	8.24	4.21	1.75	$H^{\gamma 1}$ 1.01[c]; $H^{\gamma 2}$ 1.47[c]; H_3^{γ} 0.84; H_3^{δ} 0.77
Val[18]	7.34	3.91	1.86	$H_3^{\gamma 1}$ 0.68[b]; $H_3^{\gamma 2}$ 0.81[b]
Cys[19]	8.01	4.40	2.43, 2.29	
Phe[20]	8.01	4.45	2.55, 2.80[b]	H_2^{δ} 6.71; H_2^{ϵ} 7.06; H^{ζ} 7.10
Trp[21]	7.70	4.37	3.24, 3.02	$H^{\delta 1}$ 7.09; $H^{\epsilon 1}$ 10.40; $H^{\zeta 2}$ 7.35; $H^{\eta 2}$ 7.10; $H^{\zeta 3}$ 7.00; $H^{\epsilon 3}$ 7.56

[a] The ^1H chemical shift values (±0.02 ppm) were referenced relative to the residual DMSO signal, which was assumed to resonate 2.57 ppm downfield of TSP.

[b] Stereospecific assignment. For the H^{β} resonances, the first chemical shift is for $H^{\beta 2}$, and the second is for $H^{\beta 3}$ (IUPAC-IUB Commission, 1970). The glycine H^{α} chemical shifts are given in the order $H^{\alpha 1}$, $H^{\alpha 2}$.

atom and all heavy atom rmsd values between the starting structure and the snapshots are 1.03 and 1.68 Å for the 293 K run (Fig. S2 of the Supplementary Material), and 1.12 and 1.88 Å for the 500 K run. The overall fold is well maintained in both unrestrained simulations.

Experimental results verify the high stability of siamycin II. After adding D_2O (to 50% v/v) to the sample dissolved in DMSO, seven backbone amide NH protons were still observable after 24 h at 35 °C. The residues with slowly exchanging amide protons are Leu^2, Ile^4, Gly^5, Cys^7, Asp^9, Tyr^{15} and Ala^{16}. Figure 5 shows a schematic diagram of a representative structure that illustrates the locations of these protons. Leu^2, Ile^4, Gly^5, Cys^7 and Asp^9 are in the central ring, and the amide protons of these residues are buried and all point roughly inward toward the center of the ring. Tyr^{15} and Ala^{16} pass through the center of the ring. These results demonstrate that this region of the peptide is highly stable. Also, only two of the slowly exchanging amides are involved in hydrogen bonds. Leu^2 NH forms a hydrogen bond with Tyr^{15} O, and Cys^7 NH forms a hydrogen bond with Ala^{16} O. The remaining slowly exchanging amide hydrogens must be protected by the rigidity of the structure. Furthermore, it is noteworthy that the Rhône–Poulenc Rorer group found that the Leu^2 NH was exposed to solvent (Frèchet et al., 1994) in RP-71955, while the Leu^2 side chain was partially buried. These results were based on a large NH proton temperature coefficient, and on the use of 'anti-distance constraints'. We find the opposite situation for Leu^2 in siamycin II. At 20 °C, the Leu^2 NH resonates at 9.69 ppm (Table 1) in siamycin II, and it resonates at 9.77 ppm in RP-71955 at 10 °C (Frèchet et al., 1994). These downfield shifted positions are more consistent with a buried, hydrogen-bonded NH than with a solvent-exposed NH. Also, an exposed NH is not consistent with slow exchange. This example illustrates the dangers inherent in drawing structural conclusions from amide temperature coefficients. All other factors being equal, an amide resonance that has a large secondary shift (i.e., a large displacement from its 'random coil' value) is likely to have a larger temperature coefficient than one with a small secondary shift.

In addition to having an unusual backbone fold, siamycin II displays interesting surface features. Figures 6A and B show spacefilling CPK and color-coded molecular surface images, respectively, of a representative siamycin II structure in an orientation similar to the one displayed in Figs. 4 and 5. This side of the peptide will be referred to as the 'polar face'. Polar (hydrophylic) groups exposed on this face include the Tyr^{15} ring hydroxyl, the Ser^6 hydroxyl, the Asn^8 side chain and the backbone carbonyl oxygens of Gly^3, Ile^4, Ser^6, Gly^{12}, Cys^{13}, Gly^{14} and Cys^{19}. A 180° rotation about the vertical axis produces the views shown in Figs. 6C and D. This side of the peptide is the 'apolar face'. Apolar (hydrophobic) groups on this face include the side chains of Leu^2, Ile^4, Phe^{10}, Ala^{16}, Ile^{17}, Val^{18} and Phe^{20}. These structural features indicate that siamycin II, as well as siamycin I and RP-71955, possess a novel amphipathic character.

Similarity of the siamycin I, siamycin II and RP-71955 structures

In order to directly compare NMR data obtained for siamycin I with the siamycin II data, a sample of siamycin I was prepared in DMSO/H_2O. In this solvent, siamycin I yields spectra that are very similar to those obtained for siamycin II. An example is given in Fig. 7, which shows the NH-NH region of a NOESY spectrum with $\tau_m = 90$ ms of siamycin I at 20 °C. Note the similarity of this spectral region to that shown in Fig. 1. Highly similar patterns of NOESY cross peaks are observed in all regions of the NOESY spectra. Assignments for siamycin I (Table 4) were readily obtained. Within the error of the measurements, the majority of the chemical shifts are identical between corresponding protons in siamycins II and I. Overall, the data indicate that both compounds adopt very similar conformations. Excluding the H^β resonances of residue 4, the largest chemical shift change occurs for Gly^5 NH, which shifts downfield by 0.14 ppm in siamycin I relative to siamycin II. In siamycin II, the Ile^4 H^α-Gly^5 HN cross peak overlaps with the Ile^{17} H^α-Val^{18} HN cross peak; therefore, loose constraints were used for these interactions. The Val^4 H^α-Gly^5 HN cross peak is well resolved and strong in siamycin I, indicating

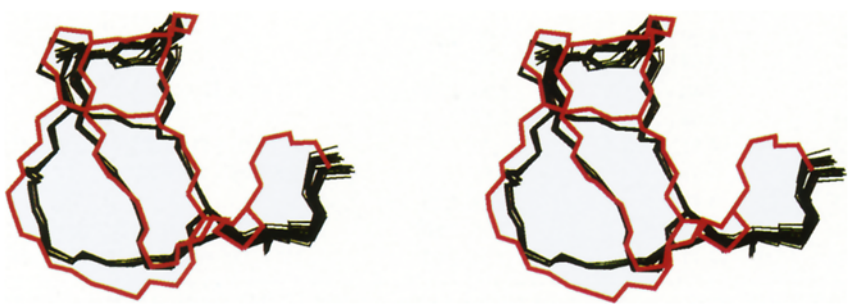


Fig. 8. Stereoviews of the 30 final energy-minimized siamycin II structures (black thin trace) and the restrained minimized RP-71955 structure (red heavy trace). All backbone N, C and C^α atoms, the disulfide bridges and the Asp^9 C^β and C^γ atoms of the amide bond between the Asp^9 side chain and Cys^1 N are shown.

that the minor Ile[4]-Gly[5] amide plane conformation (Ile[4] ψ ~88° and Gly[5] ϕ ~−167°) observed for three siamycin II structures may be the true conformation. For siamycin I, the data indicate that the Val[4] side chain adopts a χ^1 ~60° conformation, which is topologically equivalent to the Ile[4] conformation observed in siamycin II.

Similar chemical shift positions are observed for siamycin II in 100% DMSO, siamycin II in 50% DMSO/50% H_2O and RP-71955 in 50% methanol/50% H_2O (Figs. S3–S6 of the Supplementary Material). As noted previously, the structures of siamycin II and RP-71955 are overall quite similar. In Fig. 8, we compare the restrained minimized mean structure of RP-71955 (red heavy trace) and our ensemble of siamycin II structures. Excluding residues 19–21, the average backbone N, C and C^α atom rmsd between the siamycin II structures and the RP-71955 mean structure is 1.08 ± 0.07 Å. It is apparent that the overall knot-like fold is maintained in the different solvent conditions examined. This result, along with the high stability observed for the central core of siamycin II, suggests that large-scale conformational changes upon binding to the biological target protein(s) are unlikely; i.e., the reported structures are probably very similar to the biologically active conformation.

Biological implications: Possible structure–function relationships

It is likely that the amphipathic character of siamycin II and siamycin I plays a role in their biological activity. These peptides inhibit HIV-induced cell fusion (Lin, P.-F. et al., unpublished results). The polar or apolar face of the peptides may bind to a complementary surface on the HIV envelope glycoprotein gp41, or possibly on the HIV envelope glycoprotein gp120 or on an as yet unidentified cell surface receptor. Such an interaction would sterically block the complementary protein surface, and it would result in the *presentation of a surface of the wrong polar/apolar character*. This latter consequence of binding siamycin II or siamycin I may play a crucial role in blocking interactions between HIV and the T-cell surface, an essential step in the fusion process. For example, the apolar face of the siamycins may bind directly to the hydrophobic gp41 fusion domain. This would 'cap' the domain with a solvent-exposed polar surface, thus preventing the fusion domain from penetrating into the membrane. Alternatively, siamycin binding may trap gp41 in an inactive conformation. Indeed, gp41 appears to be the most likely target, since sequencing of DNA from resistant HIV strains revealed mutations in gp41 (Lin, P.-F. et al., unpublished results). Experiments aimed at localizing the siamycin II/siamycin I binding site are underway.

Conclusions

Siamycin II and siamycin I display an unusual tricyclic structure, with a highly stable central ring (residues 1 through 9) and a knot-like threading of residues 14–16 through the central ring. This highly compact arrangement afforded an exceptionally large number of NOE constraints (~16 per residue) for siamycin II, resulting in the determination of a highly resolved solution structure ensemble. Analysis of NMR data for siamycin I revealed that it adopts a conformation that is essentially identical to the siamycin II structure. Not only are the backbone folds unusual for these peptides, the molecular surfaces also have novel features. Overall, the peptides are wedge-shaped, with one face being predominantly hydrophobic, and the other predominantly hydrophilic. The polar/apolar character may be important for the anti-HIV activity of these peptides. Delineation of the precise target and binding mode of siamycins II and I could potentially yield important new weapons in the fight against AIDS.

Acknowledgements

We thank Dr. Niels Andersen for enlightening discussions regarding the fitting of NOESY time series data, and Dr. Stanley Krystek for assistance with the GRASP program.

References

Andersen, N.H., Eaton, H.L. and Lai, X. (1989) *Magn. Reson. Chem.*, **27**, 515–528.

Baleja, J.D., Moult, J. and Sykes, B.D. (1990) *J. Magn. Reson.*, **87**, 375–384.

Bax, A. and Davis, D.G. (1985) *J. Magn. Reson.*, **65**, 355–360.

Bax, A., Ikura, M., Kay, L.E. and Zhu, G. (1991) *J. Magn. Reson.*, **91**, 174–178.

Bodenhausen, G. and Ruben, D.J. (1980) *Chem. Phys. Lett.*, **69**, 185–188.

Bodenhausen, G., Kolger, H. and Ernst, R.R. (1984) *J. Magn. Reson.*, **58**, 370–388.

Braunschweiler, L. and Ernst, R.R. (1983) *J. Magn. Reson.*, **53**, 521–528.

Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S. and Karplus, M. (1983) *J. Comput. Chem.*, **4**, 187–217.

Bruccoleri, R.E. and Karplus, M. (1987) *Biopolymers*, **26**, 137–168.

Bruccoleri, R.E. (1990) CONGEN Manual (Version 2.0), Bristol-Myers Squibb Pharmaceutical Research Institute, Princeton, NJ.

Bruccoleri, R.E. (1993) *Mol. Simul.*, **10**, 151–174.

Brünger, A.T. (1992) X-PLOR Manual (Version 3.1), Yale University Press, New Haven, CT.

Constantine, K.L., Madrid, M., Banyai, L., Trexler, M., Patthy, L. and Llinás, M. (1992) *J. Mol. Biol.*, **223**, 281–298.

Constantine, K.L., Friedrichs, M.S. and Mueller, L. (1994a) *J. Magn. Reson. Ser. B*, **104**, 62–68.

Constantine, K.L., Friedrichs, M.S., Metzler, W.J., Wittekind, M., Hensley, P. and Mueller, L. (1994b) *J. Mol. Biol.*, **236**, 310–327.

Fairbrother, W.J., Palmer, A.G., Rance, M., Reizer, J., Saier Jr., M.H. and Wright, P.E. (1992) *Biochemistry*, **31**, 4413–4425.

Frèchet, D., Guitton, J.D., Herman, D., Faucher, G., Helynck, G., Monegier du Sorbier, B., Ridoux, J.P., James-Surcouf, E. and Vuilhorgne, M. (1994) *Biochemistry*, **33**, 42–50.

Griesinger, C., Otting, G., Wüthrich, K. and Ernst, R.R. (1988) *J. Am. Chem. Soc.*, **110**, 7870–7872.

Güntert, P., Braun, W. and Wüthrich, K. (1991) *J. Mol. Biol.*, **217**, 517–530.

Güntert, P. and Wüthrich, K. (1991) *J. Biomol. NMR*, **1**, 447–456.

Güntert, P. (1992) DIANA User's Manual (Version 2.0), Eidgenössische Technische Hochschule, Zürich.

Hyberts, S.G. and Wagner, G. (1989) *J. Magn. Reson.*, **81**, 418–422.

IUPAC-IUB Commission on Biochemical Nomenclature (1970) *J. Mol. Biol.*, **52**, 1–17.

Kay, L.E., Keifer, P. and Saarnev, T. (1992) *J. Am. Chem. Soc.*, **114**, 10663–10665.

Kumar, A., Ernst, R.R. and Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun.*, **95**, 1–6.

Lai, X., Chen, C. and Andersen, N.H. (1993) *J. Magn. Reson. Ser. B*, **101**, 271–288.

Macura, S. and Ernst, R.R. (1980) *Mol. Phys.*, **41**, 95–117.

Marion, D. and Bax, A. (1988) *J. Magn. Reson.*, **80**, 528–533.

Mueller, L. (1987) *J. Magn. Reson.*, **72**, 191–197.

Nicholls, A., Sharp, K. and Honig, B. (1991) *Protein Struct. Funct. Genet.*, **11**, 281–296.

Piotto, M., Saudek, V. and Sklenář, V. (1992) *J. Biomol. NMR*, **2**, 661–665.

Rance, M., Sørensen, O.W., Bodenhausen, G., Wagner, G. and Ernst, R.R. (1983) *Biochem. Biophys. Res. Commun.*, **69**, 979–987.

Richardson, J.S. (1981) *Adv. Protein Chem.*, **34**, 167–339.

Robson, B. and Garner, J. (1986) *Introduction to Proteins and Protein Engineering*, Elsevier, New York, NY.

Rose, G.D., Gierasch, L.M. and Smith, J.A. (1985) *Adv. Protein Chem.*, **37**, 1–109.

States, D.J., Haberkorn, R.A. and Ruben, D.J. (1982) *J. Magn. Reson.*, **48**, 286–292.

Weiner, S.J., Kollman, P.A., Nguyen, D.T. and Case, D.A. (1986) *J. Comput. Chem.*, **7**, 230–252.

Wüthrich, K., Billeter, M. and Braun, W. (1983) *J. Mol. Biol.*, **169**, 949–961.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York, NY.