Journal of Computer-Aided Molecular Design, 10 (1996) 11–22
ESCOM

J-CAMD 327

Structure of a cyclic peptide with a catalytic triad, determined by computer simulation and NMR spectroscopy

Björn Walse[a], Magnus Ullner[a], Christer Lindbladh[b], Leif Bülow[b], Torbjörn Drakenberg[a] and Olle Teleman[c,*]

Departments of [a]Physical Chemistry 2 and [b]Pure and Applied Biochemistry, Centre for Chemistry and Chemical Engineering, Lund University, P.O. Box 124, S-221 00 Lund, Sweden
[c]Biotechnology and Food Research, Technical Research Institute of Finland, P.O. Box 1500, FIN-02044 VTT Espoo, Finland

Received 22 September 1995
Accepted 7 December 1995

Keywords: Peptide conformation; Peptide design; Serine protease mimic; Molecular dynamics simulation; NMR spectroscopy; Structure determination; Cyclic peptide

Summary

We report the design of a cyclic, eight-residue peptide that possesses the catalytic triad residues of the serine proteases. A manually built model has been relaxed by 0.3 ns of molecular dynamics simulation at room temperature, during which no major changes occurred in the peptide. The molecule has been synthesised and purified. Two-dimensional NMR spectroscopy provided 35 distance and 7 torsion angle constraints, which were used to determine the three-dimensional structure. The experimental conformation agrees with the predicted one at the β-turn, but deviates in the arrangement of the disulphide bridge that closes the backbone to a ring. A 1.2 ns simulation at 600 K provided extended sampling of conformation space. The disulphide bridge reoriented into the experimental arrangement, producing a minimum backbone rmsd from the experimental conformation of 0.8 Å. At a later stage in the simulation, a transition at Ser^3 produced more pronounced high-temperature behaviour. The peptide hydrolyses *p*-nitrophenyl acetate about nine times faster than free histidine.

Introduction

The design of a protein or a peptide into a predetermined three-dimensional (3D) structure has for a long time posed a challenge to scientists. Although we do not understand how the amino acid sequence of a protein determines its specific 3D structure, various rules and prediction methods have more recently become available to the protein engineer [1,2]. These have been created from statistical analysis of sequences and conformations of known protein and peptide structures and from the results of previous protein design work. Recent advances in computational chemistry have allowed the extension of these methods [3], sometimes by reducing the complexity of the interaction model [4]. In the present work, a molecular dynamics (MD) simulation was carried out without experimental restraints, as a primary step in the design of a peptide mimicking the catalytic triad of a serine protease, but also to explore the correlation with experiment and with a view to using simulation as a predictive technique in protein design. A similar approach was recently used to study the helix I/loop I fragment of barnase as an early folding unit [5].

The structures and catalytic mechanism of the serine proteases are among the most thoroughly studied [6–14]. These proteases are characterised by the presence of a serine, a histidine and an aspartic acid residue, the catalytic triad, in the active centre. The first step in the hydrolysis of a peptide or ester bond involves a nucleophilic attack of the serine hydroxyl oxygen on the substrate carbonyl carbon, which then forms an acylated enzyme intermediate. The nucleophilicity of the serine hydroxyl oxygen is potentiated by the other two amino acids in the triad, histidine and aspartic acid. Hydrolysis of the acyl serine intermediate completes the reaction.

Serine proteases have been mimicked in various artificial enzyme models, where the catalytically active residues Ser, His and Asp have been mounted on various carriers

*To whom correspondence should be addressed.

Ace Ala Cys Ser Pro Gly His Cys Glu O^-

Fig. 1. The primary structure of CTP.

in positions predicted to form a charge-relay system. Examples of such designed peptides are an undecapeptide adopting an α-helical conformation [15], a 73-residue peptide, chymohelizyme-1, with four covalently linked chemically synthesised amphiphilic α-helical peptides [16], and a peptide consisting of seven 'esterase' units fused to the N-terminal end of galactokinase by genetic engineering methods [17]. These mimics display activities far below those of the proper enzymes. The problem lies in the design of a molecule that adopts a certain 3D structure essential for the desired function. Another attempt to achieve this is a 29-residue peptide with the N- and C-termini connected by a disulphide bridge [18], although its catalytic properties have not yet been verified [19].

Here we have investigated the design of a very small and stable peptide, onto which functionally important amino acid residues can be grafted to obtain an enzyme-like catalytic triad. In this context, we temporarily choose to disregard all contributions from the oxyanion hole as well as from the large enzyme framework in general. Serine, histidine and glutamic acid were placed at appropriate positions in an octameric peptide designed to fold with high probability into a stable β-turn and to be locked by a disulphide bridge near the termini. Glu replaces Asp in this mimic for geometric reasons. The desired conformation was model built manually. A 300 ps MD simulation indicated short-term stability of the conformation; consequently, the peptide was synthesised and its conformation determined by a combination of NMR spectroscopy and simulated folding. The conformational space was probed further by a 1.2 ns simulation at 600 K, and we demonstrate that this simulation visits the experimentally determined conformation. We describe the structure and dynamics of the molecule and compare the predicted and physical conformations, including a control simulation based on the experimental structure. Hydrolytic activity data are also given and we conclude by discussing improvements in the design.

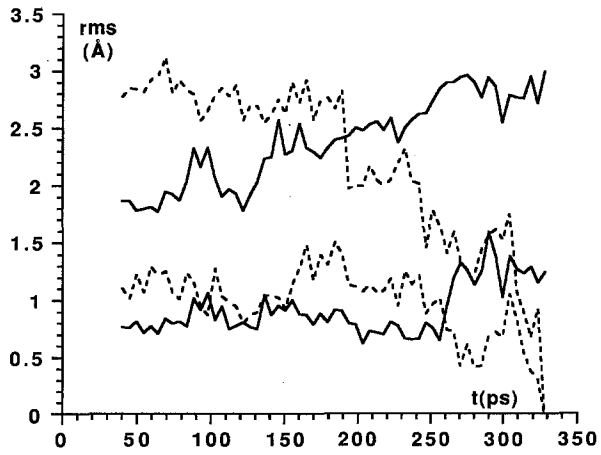

Fig. 2. Rms deviations during the original model simulation. The two bottom curves are the rmsd for the simulated backbone against the simulated backbone initially (full line) and after 328 ps (dashed). The two top curves give, similarly, the rmsd for all atoms in the peptide. The trends in the all-atom rmsd indicate that equilibrium has not been reached for the side chains. The term backbone indicates the C, C^{α}, N and O atoms of residues 2–7 plus C^{β} and S^{γ} in Cys^2 and Cys^7, i.e., the 'ring'.

Methods

Initial model

A physical stick model was built manually from the catalytic triad conformation of chymotrypsin [6,20]. The triad was incorporated into a very small peptide, acetate-Ala-Cys-Ser-Pro-Gly-His-Cys-Glu-O^-, Fig. 1. We will refer to this peptide as CTP, for *c*atalytic *t*riad *p*eptide. The molecule backbone forms a ring closed by a disulphide bridge between residues 2 and 7. To achieve maximum probability for a type II β-turn in residues 3–6, residues 4–5 are Pro-Gly [21].

TABLE 1
SIMULATION PARAMETERS

Parameter	Original model	Control	600 K	Annealing
Box size ($Å^3$)	27.9	identical	identical	identical
Number of atoms in CTP	101	identical	identical	identical
Number of sodium ions	2	identical	identical	identical
Number of water molecules	669	673	669	669
Total number of atoms	2110	2122	2110	2110
Potential truncation (Å)	10	identical	identical	identical
Time step (fs)	0.2/1.2	identical	identical	identical
Equilibration[a] (ps)	2/19/19	38	154	–
Analysed simulation (ps)	288	154	1017	15×19
IBM RS-6000 340 CPU time (days)	7	4	25	7
Temperature scaling interval (ps)	0.12	0.12	0.024	0.012
Coordinate sampling interval (ps)	0.24	0.24	0.24	0.24

[a] In vacuum but with ε = 80, constrained/in water, constrained/in water, unconstrained. The total equilibration time was 40 ps.

TABLE 2
SIMULATION RESULTS

	Original model	Control	600 K
Temperature (K)	302 ± 2	301 ± 1	601 ± 5
of translation	306 ± 6	305 ± 5	608 ± 9
of rotation	299 ± 4	298 ± 4	592 ± 10
of vibration	300 ± 3	300 ± 2	601 ± 4
Pressure (bar)	−650 ± 80	−600 ± 100	5260 ± 40
Radius of gyration (Å)	5.1 ± 0.1	4.6 ± 0.1	5.2 ± 0.3
Principal moments of inertia (au Å^2)	12.6×10^3	11.0×10^3	–[a]
	15.3×10^3	12.2×10^3	–[a]
	15.8×10^3	12.2×10^3	–[a]
Backbone rmsd (Å)	0.9	0.5	0.8–2.2
Diffusion coefficient (m^2 s^{-1})	0.32×10^{-9}	0.67×10^{-9}	1.5×10^{-9}

[a] As the peptide undergoes several conformational changes during the high-temperature simulation, the average principal moments of inertia lack physical significance.

Interaction potential

The interaction potential described by Ahlström et al. [22,23] was used for the MD simulations. All atoms, including nonpolar hydrogens, were treated explicitly. A potential with explicit hydrogens is able to reproduce local dynamics and certain relaxation phenomena better than a potential using united atoms [24].

Nonbonded interactions were described by a Lennard-Jones (LJ) and a Coulomb potential. The relative dielectric permittivity was 1, since the surrounding water was treated explicitly. The parameter values were based on literature data [25–27]. For water the simple point charge (SPC) model was used [28], but with incorporation of intramolecular flexibility [29]. No explicit hydrogen bond potential was used; instead, hydrogen bonds were adequately described by Coulomb and LJ terms.

The internal bond and bond angle vibrations were treated explicitly and were represented by a harmonic potential function. The force constants and the equilibrium values for bond lengths and angles were taken from the literature [26,28,30,31]. The interactions due to internal rotation were handled by means of periodic dihedral (torsion) potentials up to an order of three. No improper torsions were used; instead, their contribution was realised via the normal bond angles.

TABLE 3
HYDROGEN BONDS

Hydrogen bond	Acceptor	Proton	Relative frequency (%)
Simulation[a]	O Ser^3	H$^{\gamma}$ Ser^3	23
(original model)	O Ser^3	H His^6	93
	O Pro^4	H His^6	17
	N$^{\varepsilon 2}$ His^6	H Ser^3	8
	N$^{\varepsilon 2}$ His^6	H$^{\gamma}$ Ser^3	21
	O His^6	H Ser^3	46
	O His^6	H$^{\gamma}$ His^6	10
	N Glu^8	H$^{\delta 1}$ His^6	24
	O$^{\varepsilon 1}$ Glu^8	H Ala^1	24
	O$^{\varepsilon 1}$ Glu^8	H Cys^2	31
	O$^{\varepsilon 1}$ Glu^8	H$^{\delta 1}$ His^6	13
	O$^{\varepsilon 1}$ Glu^8	H Cys^7	9
	O$^{\varepsilon 2}$ Glu^8	H Ala^1	33
	O$^{\varepsilon 2}$ Glu^8	H Cys^2	38
	O$^{\varepsilon 2}$ Glu^8	H Glu^8	37
	O Glu^8	H$^{\delta 1}$ His^6	34
	(O^{xt} Glu^8	H$^{\delta 1}$ His^6	5)
Experimental structure[b]	O Ser^3	H His^6	90
	O Pro^4	H His^6	80
	N Gly^5	H Cys^7	30

[a] The hydrogen bond criterion according to Linse et al. [42]. The hydrogen–acceptor formation and breaking distances were 2.5 and 3.6 Å and the donor–hydrogen–acceptor formation and breaking angles were 135° and 100°.

[b] The hydrogen–acceptor distance criterion was 3.0 Å. The donor–hydrogen–acceptor angle criterion was 120°.

Simulation technique

All simulations were done using the molecular dynamics package MUMOD [32,33]. Newton's equation of motion was solved numerically by means of a double-

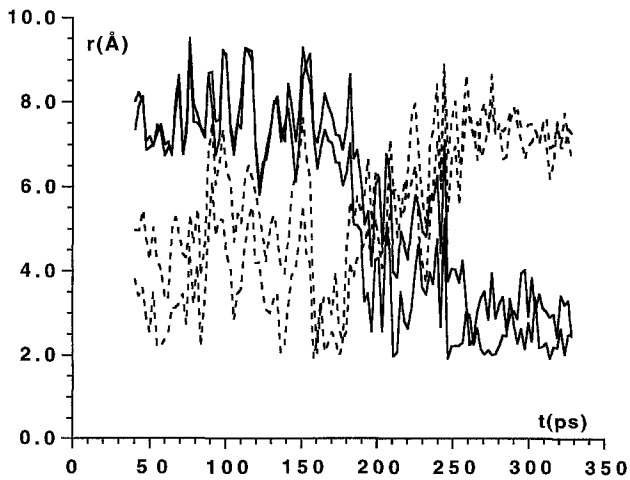

Fig. 3. The exchange of histidine imidazole ligands at about 200 ps in the original model MD simulation. Full lines show the distances as a function of time from H$^{\delta 1}$ of His^6 to the C-terminal carboxyl oxygens; dashed lines show the distances to the two Glu^8 side-chain carboxyl oxygens.

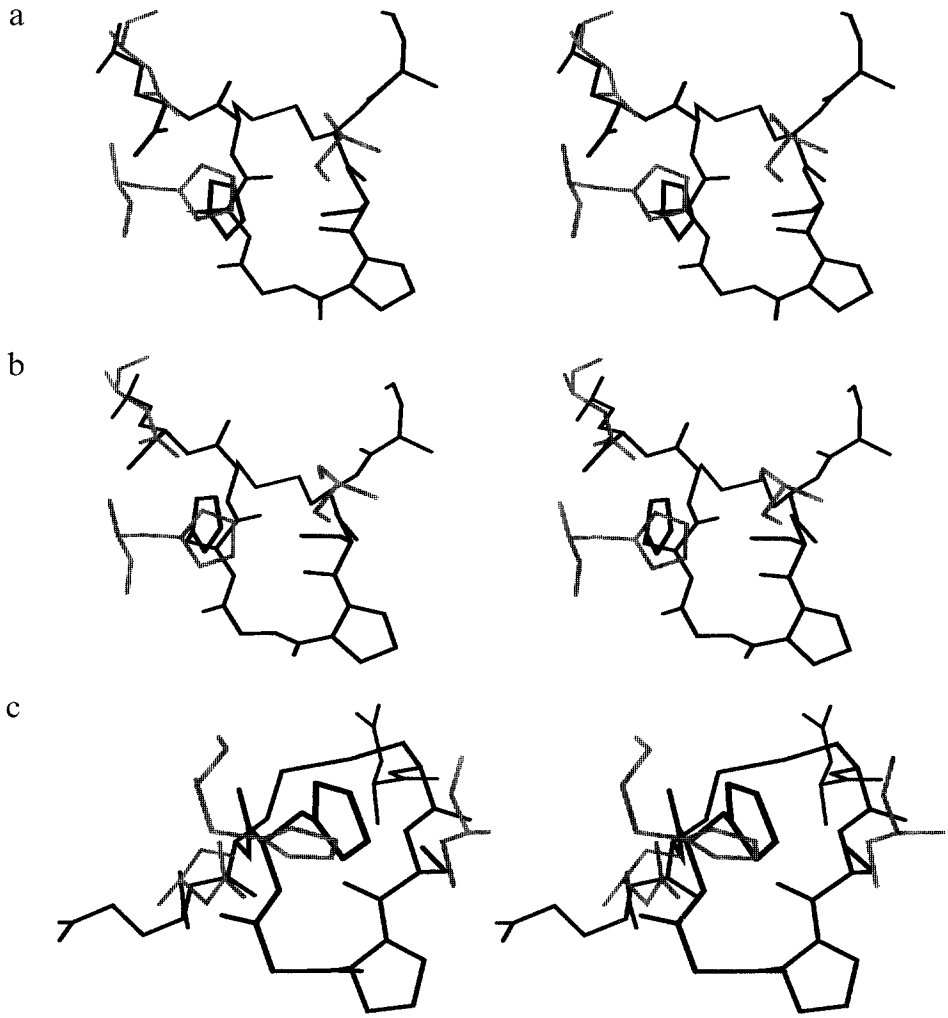

Fig. 4. Stereoviews, showing a comparison of CTP (black) and the catalytic triad (Ser^{195}, His^{57} and Asp^{102}) in α-chymotrypsin (grey). Coordinates of α-chymotrypsin were taken from Birktoft and Blow [8]. The figures were generated using the UCSF software MidasPlus. (a) The catalytic triad superimposed on the side chains of Ser^3, His^6 and Glu^8 in the original simulated conformation of CTP; rmsd = 2.36 Å. (b) The catalytic triad superimposed on the side chains of Ser^3, His^6 and the carboxy terminus in the original simulated conformation of CTP; rmsd = 1.34 Å. (c) The catalytic triad superimposed on the side chains of Ser^3, His^6 and the carboxy terminus in one of the NMR-derived conformations of CTP; rmsd = 3.26 Å.

time-step, fourth order predictor–corrector Gear algorithm. Covalent bonds and bond angles were integrated with a time step of 0.2 fs, and all other degrees of freedom with a time step of 1.2 fs. The two-time-step algorithm gives radial distribution functions and translational diffusion coefficients identical to those obtained from a one-step algorithm with a time step of 0.2 fs [29].

The simulation was started from the model-built conformation of CTP. Initially, 729 water molecules were placed on the basis of a primitive cubic lattice of 9 × 9 × 9 points with a repeat distance of 3.1 Å. Water molecules overlapping with the peptide were removed and all water orientations were randomised. Water initialised in this fashion has been shown to equilibrate within 10 ps [29]. Of the remaining 671 water molecules, two were randomly selected and exchanged for sodium ions to achieve electroneutrality. The peptide and solvent filled a cubic cell with a side length of 27.9 Å, which amounts to a concentration of 76 mM, and full periodic boundary conditions were used.

The initial model-built conformation had a very high potential energy and was equilibrated for 0.025 ps at low temperature, after which the temperature was increased to 300 K. At first, the catalytic triad was maintained by harmonic constraints between H^{γ} Ser^3 ↔ $N^{\varepsilon 2}$ His^6 and $H^{\delta 1}$ His^6 ↔ $O^{\varepsilon 1}$ Glu^8, both with an equilibrium distance of 2.0 Å. CTP was simulated with these constraints for 2 ps in vacuum, but with a dielectric permittivity of 80, and thereafter in explicit water for 19 ps. The constraints were then released and the simulation was continued for 19 ps to give a total of 40 ps of equilibration. An ensuing 288 ps of simulation was analysed.

As a control, an MD simulation was performed with one of the NMR-derived conformations as the starting

TABLE 4
NMR CHEMICAL SHIFTS

Residue	NH	$C^{\alpha}H$	$C^{\beta}H$	$C^{\gamma}H$	Other
Ala[1]	8.42	4.31	1.36		
Cys[2]	8.42	4.77	3.56		
			3.01		
Ser[3]	8.74	4.77	3.88, d		
Pro[4]	–	4.30	2.32	2.18	3.98 (H^{δ})
			1.91	2.04	3.68
Gly[5]	8.95	4.14			
		3.67			
His[6]	8.69	4.91	3.46		8.51 (H^{ε})
			3.13		7.14 (H^{δ})
Cys[7]	8.61	4.83	3.31		
			2.99		
Glu[8]	8.30	4.12	2.06	2.19, d	
			1.90		

^{1}H NMR chemical shifts are given in ppm at 5 °C. Chemical shifts are relative to the H_2O signal at 4.98 ppm and are accurate to ±0.02 ppm. Cases where degeneracy is assumed are indicated with 'd'.

point. This conformation was more robust in the early equilibration, 38 ps, after which a trajectory of 154 ps was generated and analysed.

To search conformation space more efficiently, a longer MD simulation was performed at 600 K. The equilibration for this simulation was started as for the original model. After an extra 19 ps at 300 K, the temperature was raised in steps of 50 K every 10 ps, followed by 1084 ps of simulation at 600 K. Conformations were taken from the high-temperature simulation every 60 ps for the last 840 ps and were annealed by linearly decreasing the temperature to 0 K during 19 ps. The three trajectories will be referred to as 'original model', 'control' and '600 K'. The simulation parameters used are summarised in Table 1.

Peptide synthesis

The peptide CTP was assembled by stepwise solid-phase synthesis on Pam resin in an ABI430A synthesiser at the molecular biology core facility at Lund University, and according to the ABI standard program with capping after each coupling. Cleavage from the resin was performed in HF:DMS:*p*-cresol:thiocresol, 32:4:3:1, for 3 h at −5 °C. After washing and extraction in ether and 20% acetic acid, the peptide was lyophilised and redissolved in water, the pH was adjusted to pH 8 with NH_3 and the peptide was oxidised overnight. The oxidised peptide was finally purified by high-performance liquid chromatography (HPLC) on a YMC reversed phase C_{18} column (20× 250 mm) using an acetonitrile gradient (12–28%) in 0.1% TFA and 0.05% NH_3.

NMR spectroscopy

Two samples were used for NMR measurements. In the first, CTP was dissolved at an approximate concentration of 7 mM in a mixture of H_2O and D_2O (10%), and the concentration was 4 mM in pure D_2O in the second sample. The pH was adjusted to 6.3 with small amounts of 0.4% DCl and 0.4% NaOD.

All ^{1}H NMR spectra were collected on a General Electric Omega 500 spectrometer operating at 500.13 MHz, and the acquired data were processed with the GE Omega software on SUN workstations. Pure absorption, phase-sensitive COSY [34], TOCSY [35] and ROESY [36] spectra were recorded at 5 °C using the hypercomplex method of data collection [37] with the carrier frequency placed on the solvent resonance. A spectral width of 5 kHz was employed in all cases, with a relaxation delay of 1.5 s incorporated into each sequence. TOCSY spectra were recorded by the MLEV17 sequence [38] and the spin-lock field was applied for 120 ms. ROESY spectra were re-

TABLE 5
COUPLING CONSTANTS AND STEREOCHEMISTRY

Parameter	Ala[1]	Cys[2]	Ser[3]	Pro[4]	Gly[5]	His[6]	Cys[7]	Glu[8]
$^3J_{NH^{\alpha}}$ (Hz)	5.6	6.9	6.9	–	z.q.	9.7	8.6	7.4
$^3J_{\alpha\beta1}$ (Hz)	–	5.3	d	n.d.	–	10.2	9.7	3.6
$^3J_{\alpha\beta2}$ (Hz)	–	5.7	d	n.d.	–	5.1	5.4	3.5
$NOE_{\alpha\beta1}$	–	strong	d	weak	–	weak	medium	strong
$NOE_{\alpha\beta2}$	–	strong	d	strong	–	strong	strong	strong
$NOE_{NH^{\beta1}}$	–	weak	d	–	–	strong	strong	weak
$NOE_{NH^{\beta2}}$	–	medium	d	–	–	weak	weak	strong
Rotamer	n.d.	g^2g^3	n.d.	n.d.	n.d.	t^2g^3	t^2g^3	g^2g^3
χ^1 (°)	n.d.	60	n.d.	n.d.	n.d.	−60	−60	60
Φ_{lower} (°)	−90	–	–	–	–	−160	−160	–
Φ_{upper} (°)	−40	–	–	–	–	−80	−80	–
χ^1_{lower} (°)	–	20	–	–	–	−100	−100	20
χ^1_{upper} (°)	–	100	–	–	–	−20	−20	100

Coupling constants were extracted from a COSY spectrum collected at 5 °C after resolution enhancement; the NOE cross peaks were evaluated in a 300 ms ROESY spectrum collected at 5 °C. 'z.q.' indicates that the coupling constant could not be determined due to zero-quantum contribution, 'n.d.' that the stereochemical assignment has not been determined and 'd' that the ^{1}H β chemical shifts were degenerated.

TABLE 6
DISTANCE CONSTRAINTS

Atom i	Atom j	Upper bound (Å)	Lower bound (Å)	Atom i	Atom j	Upper bound (Å)	Lower bound (Å)
Ace^0 M	Ala^1 H	3.750	1.925	Gly^5 H	His^6 H	4.110	2.877
Ala^1 H	Ala^1 M^{β}	3.710	1.897	Gly^5 $H^{\alpha 2}$	His^6 H	3.840	2.688
Ala^1 H^{α}	Ala^1 M^{β}	3.410	1.800	His^6 H	His^6 $H^{\beta 1}$	3.240	2.268
Cys^2 H	Cys^2 H^{α}	3.410	2.387		His^6 $H^{\delta 2}$	3.610	2.527
	Cys^2 $H^{\beta 1}$	3.590	2.513	Cys^7 H	Cys^7 $H^{\beta 1}$	3.070	2.149
	Cys^2 $H^{\beta 2}$	3.430	2.401		Glu^8 H	3.520	2.464
Cys^2 H^{α}	Cys^2 $H^{\beta 1}$	2.670	1.869	Cys^7 H^{α}	Cys^7 $H^{\beta 1}$	2.950	2.065
	Cys^2 $H^{\beta 2}$	2.710	1.897		Cys^7 $H^{\beta 2}$	2.620	1.834
Cys^2 $H^{\beta 1}$	Ser^3 H	3.490	2.443		Glu^8 H	3.040	2.128
Cys^2 S^{γ}	Cys^7 S^{γ}	2.500	1.800	Cys^7 $H^{\beta 2}$	Glu^8 H	3.720	2.604
Ser^3 H	Ser^3 L^{β}	4.200	2.310	Glu^8 H	Glu^8 H^{α}	3.000	2.100
	His^6 H	4.350	3.045		Glu^8 $H^{\beta 1}$	3.690	2.583
	His^6 $H^{\beta 1}$	3.690	2.583		Glu^8 $H^{\beta 2}$	3.070	2.149
Ser^3 H^{α}	Ser^3 L^{β}	3.590	1.883		Glu^8 L^{γ}	4.070	2.219
	Pro^4 $H^{\delta 1}$	2.360	1.800	Glu^8 H^{α}	Glu^8 $H^{\beta 1}$	2.640	1.848
	Pro^4 $H^{\delta 2}$	3.790	2.653		Glu^8 $H^{\beta 2}$	2.870	2.009
Pro^4 H^{α}	Gly^5 H	2.670	1.869		Glu^8 L^{γ}	3.580	1.876
	Cys^7 H	3.510	2.457				

L denotes a pseudoatom that represents both protons in a methylene group and M denotes one that represents the three protons of a methyl group.

corded with a mixing time of 300 ms. For each FID, 64 transients were acquired. The data size was 512×2048 complex points for all 2D experiments. The raw data sets were multiplied by an unshifted sine-bell window function in both dimensions for COSY spectra and by a 45° phase-shifted sine squared-bell window function for TOCSY and ROESY spectra. Zero-filling in the evolution dimension provided final 2D spectra of 1024×2048 complex points.

Structure determination

The NOE cross-peak volumes were integrated from the ROESY spectra using the in-house software MAGNE, and interpreted in terms of interproton distances under the isolated two-spin approximation (ISPA) [39]. The geminal H^{β} protons in Cys^2 and Cys^7, d = 1.77 Å, were used for calibration.

HN-H^{α} coupling constants were extracted from a COSY spectrum (sample dissolved in H_2O/10% D_2O), zero-filled to a final data size of 1024×8192 points. The Φ dihedral angle constraints were classified into ranges $-90° < \Phi < -40°$ for ${}^3J_{NH^{\alpha}} \leq 5.6$ Hz and $-160° < \Phi < -80°$ for ${}^3J_{NH^{\alpha}} \geq 8.5$ Hz. The side-chain rotamers were determined via the H^{β}-H^{α} coupling constants and quantification of the intraresidue NOEs for the residues that show NMR patterns consistent with one of the three classical side-chain rotamers g^2g^3, g^2t^3 and t^2g^3 [40]. The H^{β}-H^{α} coupling constants were determined from a COSY spectrum (sample dissolved in D_2O) recorded with a data size of 2048×2048 points.

The amide proton exchange was measured by collecting 1D spectra on a sample with the peptide freshly dissolved in D_2O.

Simulated folding (SF), which is essentially restrained MD with a simplified description of nonbonded interactions, was used to calculate the tertiary structure from NMR constraints. For a detailed description, see Ref. 41. No constraints representing plausible hydrogen bonds were included. The disulphide bridge was represented by a constraint rather than a covalent bond. The SF runs were performed by means of an adapted version of MUMOD.

Hydrolysis of p-nitrophenylacetate

The reaction rates were determined in acetonitrile:ethanol:70 mM phosphate buffer, pH 7.6 (1:4:2.5; v/v/v) (buffer A) containing 0.2 mM catalyst and 0.1 mM *p*-nitrophenyl acetate. The formation of *p*-nitrophenol was followed spectrophotometrically at 405 nm. Relative hydrol-

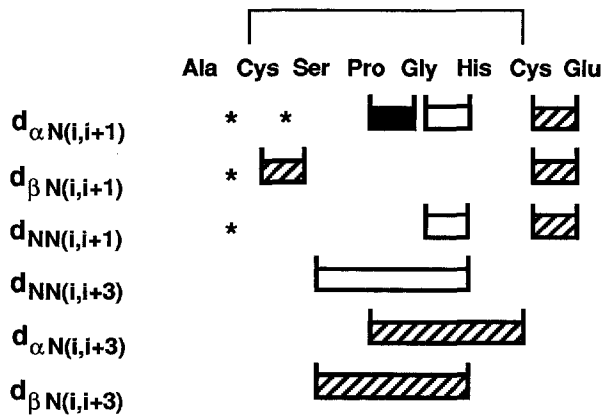

Fig. 5. Schematic representation of interresidue NOE connectivities with d < 2.5 Å (filled bar), 2.5 Å < d < 3.5 Å (hatched bar) or d > 3.5 Å (open bar). Connectivities that could not be identified due to degeneracy of resonances or saturation are marked with an asterisk.

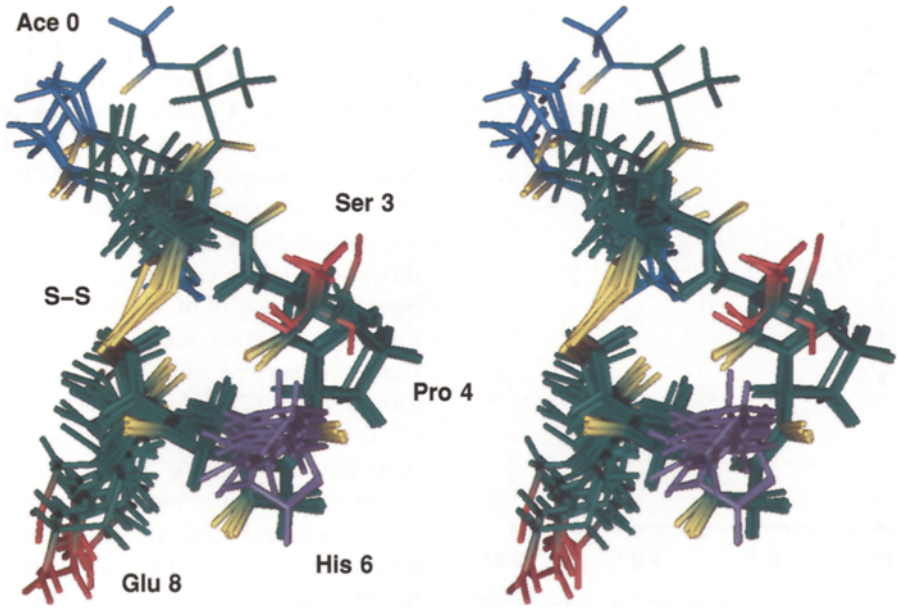

Fig. 6. Stereoview, showing 10 conformations based on the NMR constraints. The amino terminus is at the top and the carboxy terminus at the bottom of the picture. All backbone oxygens are yellow, as is the disulphide bridge closing the ring. The hydroxyl group of Ser^3 and the carboxyl group of Glu^8 are drawn in red, while the imidazole of His^6 is violet. The picture was drawn using the SCARECROW software [45].

ysis rates were calculated after correction for spontaneous hydrolysis of *p*-nitrophenyl acetate in buffer A. The rate for histidine was used as the reference. The reference peptides, Ser-His and Ser-His-Asp, were obtained from Serva (Heidelberg).

Results and Discussion

Original model and its refinement by MD simulation

An initial model was built manually by grafting a serine, a histidine and a glutamic acid side chain onto a peptide using the active site of chymotrypsin as a template. The linear peptide contained eight residues, and its backbone was made cyclic by means of a disulphide bridge connecting Cys^2 and Cys^7.

This model was relaxed and conformationally refined by means of MD simulation. Table 2 lists thermodynamic averages and other properties obtained from the simulation. The internal rms deviation from the initial model, Fig. 2, fluctuates around 0.8 Å for about 210 ps and then increases to about 1.3 Å.

Hydrogen bonds were evaluated using an asymmetric criterion, as discussed by Linse et al. [42], and are given in Table 3. The charge-relay system is not intact, in that the hydrogen bonds between H^{γ} in Ser^3 and $N^{\varepsilon 2}$ in His^6 and between $H^{\delta 1}$ in His^6 and $O^{\varepsilon 1}$ in Glu^8 are broken relatively soon. After about 200 ps, the carboxy terminus turns around and forms, some 50 ps later, a hydrogen bond with the side chain of His^6, replacing its hydrogen bond with the Glu^8 carboxyl group, see Fig. 3.

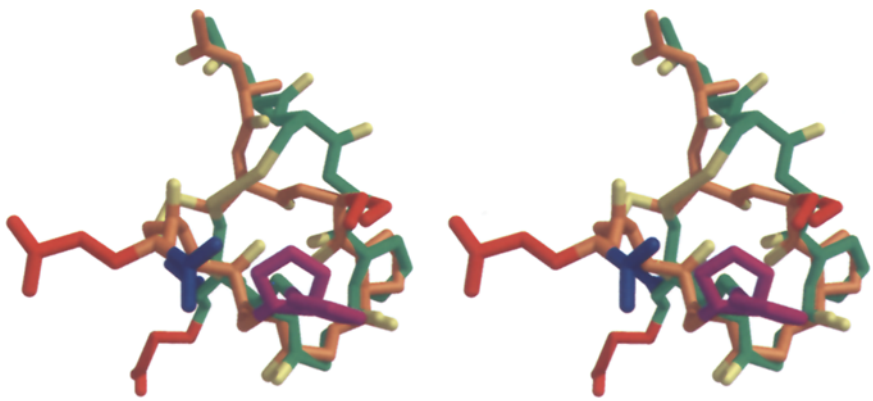

Fig. 7. Comparison of NMR-based conformation (backbone green) and simulated original model (backbone orange). Backbone oxygens and the disulphide bridge are yellow. Ser^3 and Glu^8 side chains are red, the His^6 imidazole is violet and the carboxy terminus is blue. The most notable discrepancy in the backbone is the placement of the disulphide bridge, owing to differences in the Cys^2 Ψ and His^6 Ψ backbone torsion angles. The figure was generated using the UCSF software MidasPlus.

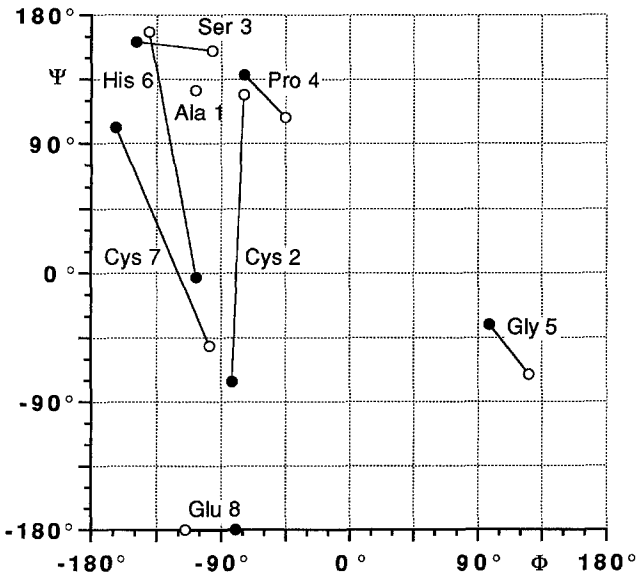

Fig. 8. Ramachandran diagram for backbone torsion angles. Filled circles are average angles for the experimental conformations, open circles are based on the original model simulation. The Ψ angle for Ala[1] could not be determined from the NMR data and the filled circle is thus not drawn. Glu[8] does not have a Ψ angle and is drawn at the bottom.

Superimposition of CTP residues Ser[3], His[6] and Glu[8] (end conformation of the original simulation) onto the catalytic triad residues in α-chymotrypsin (Ser[195], His[57] and Asp[102]) gave an rmsd of 2.36 Å (based on all side-chain atoms). Figure 4a shows that the side chain of Glu[8] points in the opposite direction compared to Asp[102]. When the superimposition is altered such that Asp[102] is matched against the CTP carboxy terminus, the rmsd decreases to 1.34 Å. The C-terminal carboxyl group of CTP has a similar orientation as the Asp[102] side chain.

Assignment of NMR resonances

The synthesised peptide was subjected to 2D COSY, TOCSY and ROESY NMR experiments. Complete assignments for all proton signals were provided from COSY and TOCSY experiments (data not shown). The amino acid residues were identified according to their characteristic pattern of chemical shifts, based on connectivities from scalar homonuclear couplings. Additional cross peaks in the spectra arose from the *cis*-Pro[4] form of the peptide at a relative population of 17%, but only the trans form was used in the structure determination.

Sequential assignments [43] of the amino acid residues were carried out from the analysis of a ROESY spectrum with a mixing time of 300 ms. Sequential NOEs between Ala[1] and Cys[2] could not be seen, due to an overlap of the HN chemical shifts. A possible sequential NOE between HN Ser[3] and H^{α} Cys[2] was also hidden, due to overlap with H^{α} Ser[3]. All chemical shifts are given in Table 4 and the coupling constants are listed in Table 5.

Structure determination

The 3D structure was calculated from proton–proton distances, derived from NOE cross peaks and coupling constant data. All amide protons in the CTP peptide showed fast exchange with the solvent. The final list of NMR constraints contained 35 NOE distances (Table 6) and seven torsion angle constraints (Table 5). The NOEs are represented schematically in Fig. 5. Starting from a different random conformation each time, 20 SF runs were performed, 10 of which converged to the same conformational family, see Fig. 6. Among these conformations, the pairwise rmsd at optimal rigid-body superposition is 0.44±0.25 Å for the backbone of residues 2–7 (N, C^{α}, C and O atoms). The rmsd against the mean conformation is 0.29±0.18 Å. Ala[1] and Glu[8], which are not part of the ring, were not included in this average.

The obtained structure possesses the intended type II β-turn, but the catalytic triad is not present in the obtained structure. Instead, Ser[3], His[6] and Glu[8] appear to form hydrogen bonds with the surrounding water. No specific interaction involves the three side chains and it appears likely that the hydrogen bond between Ser[3] and His[6] is formed occasionally. While Glu[8] is sterically unrestricted, the ring backbone is bent so as to make it impossible for the carboxyl group to reach the His[6] side chain. If the connection occurs, it is mediated by at least one water molecule. The position of H^{γ} Ser[3] is not well determined, since it exchanges for deuterium. The closest connectivities involve the β protons, leaving an uncertainty in the H^{γ} position of about 2 Å in any direction. All three

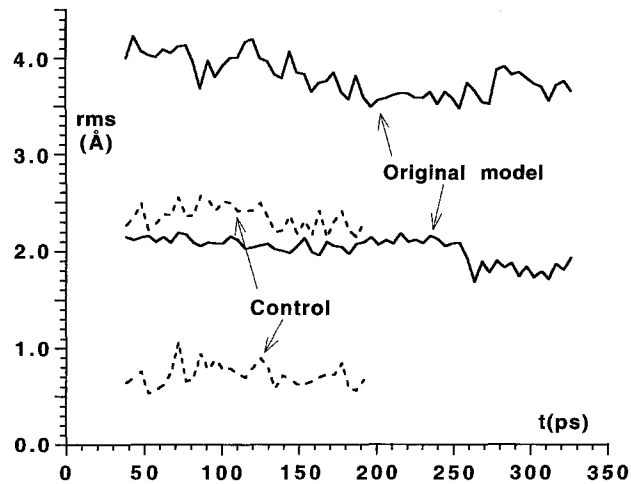

Fig. 9. Comparison of simulated and experimental conformations by means of rms deviations. The full lines give the rmsd as a function of time between the original simulated conformation and the NMR-derived conformations. The rmsd was calculated separately for each experimental conformation and then averaged. The upper curve is for all peptide atoms, the lower for the ring backbone (see the caption of Fig. 2). The dashed lines give, similarly, the rmsd for the control simulation. Note that for the backbone in the control run the rmsd was about 0.5 Å against its own starting conformation and about 0.8 Å when averaged over 10 NMR conformations.

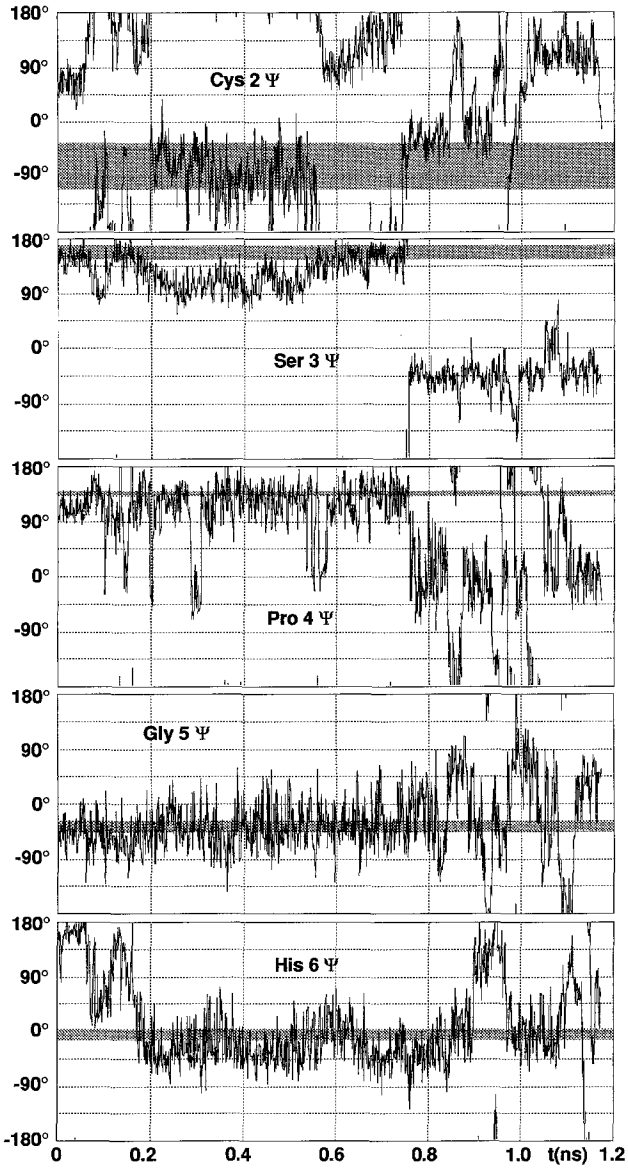

Fig. 10. The five most important backbone torsion angles as a function of time during the 600 K simulation. The grey bands indicate the experimental value for each torsion angle, where the width of the band represents the experimental uncertainty.

rotamers around the C^{α}-C^{β} bond in Ser^3 are represented in the NMR structures.

In Fig. 4c, the catalytic triad residues of α-chymotrypsin are superimposed on the experimental structure of CTP. Rmsd values of 3.62 and 3.26 Å are obtained when Asp^{102} is matched to the side chain of Glu^8 or to the carboxy terminus, respectively. Glu^8 is too far away and has the wrong orientation to establish the necessary hydrogen bonds with His^6.

Comparison of original model to experimental conformation

The experimental and simulated conformations are superimposed in Fig. 7. The backbone ring is quite restricted and agrees accurately with the simulated conformation for residues 3–6. The Ψ backbone torsions of Cys^2 and His^6 correspond to an α-helix conformation rather than a β-sheet conformation. This gives the disulphide a substantially different position in the experimental conformation and prevents the Glu^8 side chain from reaching His^6. The original model simulation was performed at room temperature and is too short for a backbone transition to occur in the ring. The secondary structure was evaluated by means of Ramachandran diagrams. Figure 8 gives such a Φ,Ψ-space comparison of the experimental conformation with the end point of the original model simulation. In the original trajectory, Ala^1 ends in a β conformation but is rather flexible and undergoes backbone transitions in the simulation. Ala^1 Ψ is undetermined in the experimental conformation; it is α in three conformations and β in seven.

The rmsd for the backbone of residues 2–7 plus the disulphide bridge, between the 10 NMR conformations and the simulated conformation, is 2.1 ± 0.1 Å for about 200 ps, but decreases, some time after the rearrangement of the carboxy terminus, to 1.8 ± 0.1 Å, see Fig. 9. The rmsd based on all atoms decreases slowly.

High-temperature simulation

As the intuitive model building in combination with room temperature simulation failed to predict the correct arrangement of the disulphide bridge, the conformational search was extended by MD simulation at 600 K for a longer period of time (Table 1). The higher temperature helps to overcome potential energy barriers and thus extends the conformational space traversed. Already after 0.2 ns, the disulphide moiety reorients and the backbone conformation of both His^6 and Cys^2 changes from β to α,

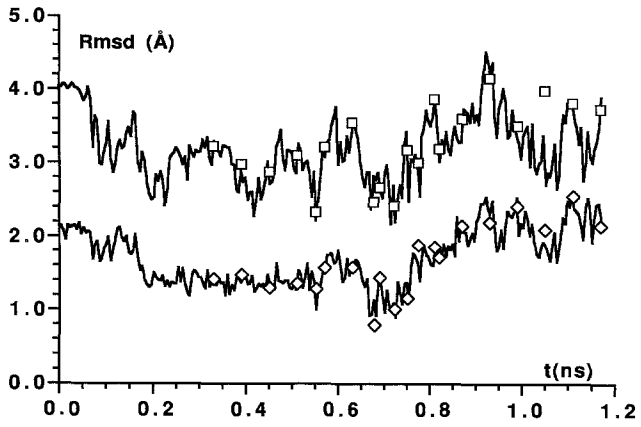

Fig. 11. Rmsd values between CTP conformations during the 600 K simulation and the 10 experimental conformations. The rmsd was calculated separately for each experimental conformation and then averaged. The rmsd is given for the 'ring' (defined in the caption of Fig. 2) and for all atoms (upper curve). The diamonds ('ring') and squares (all atoms) represent similar rmsd values, but based on annealed conformations.

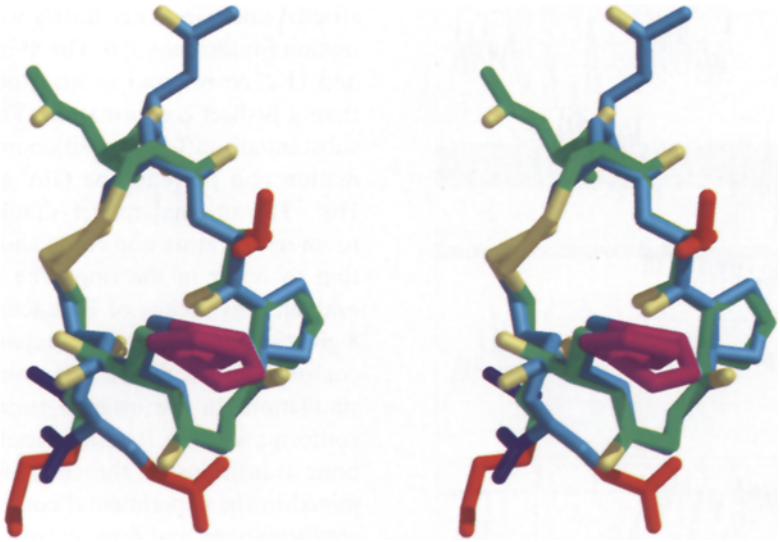

Fig. 12. Superimposition of an experimental conformation (green backbone) and the conformation annealed from the one after 678.0 ps of MD simulation at 600 K (blue backbone). The other colours are as in Fig. 7. The figure was generated using UCSF MidasPlus.

in agreement with the experimental data (Fig. 10). This situation is maintained for about 0.5 ns, when Ser^3 makes a transition from β to α. This conformational change appears to entail increased flexibility and a relatively large number of conformational transitions are observed during the remaining 0.4 ns. The Ser^3 α conformation is thus likely to be a high-temperature state. The ring conformation is in principle determined by the Ψ torsion angles for residues 2–7 plus the Φ angle of Gly^5. The number of independent states can therefore be estimated to be 2^7 or, considering the ring closure, $2^6 = 64$. With about 10 torsional transitions observed during the simulation, we can conclude that conformation space would be adequately sampled in 10–50 ns.

The rmsd was calculated against the 10 experimental conformations. The averages of the 10 values are given in Fig. 11, both for the ring backbone and for all atoms. A total of 15 snapshots were taken at regular intervals from the high-temperature simulation and cooled linearly to 0 K during 20 ps. Another five snapshots were taken based on similarity to the experimental conformation (3) or on a low total electrostatic energy (2), and these were likewise annealed (Fig. 11). It turns out that the annealing does not affect the conformations very much and produces rmsd values close to the high-temperature values.

The annealed conformation closest to the experimental ones has a ring rmsd of 0.79 Å (rmsd = 2.2 Å for all atoms). This rmsd is equal to that obtained between the 'control' simulation and the experimental conformations, which were not the starting point for the control simulation. In other words, some of the annealed conformations do not deviate more from the experimental structures than what is produced by room temperature thermal fluctuations plus the experimental uncertainty in the structure determination.

An annealed and an experimental conformation are superimposed in Fig. 12; it can be seen that they overlap well. Their secondary structure is compared in Fig. 13 and we note the increased similarity as compared to Fig. 8. The high-temperature simulation thus samples phase space well enough to visit the correct conformation after less than a nanosecond (678 ps), i.e., computer simulations can be made sufficiently long to cover the conformational space of a peptide of such a small size. We were unable to detect any simple criterion that would have identified the correct conformation without recurrence to the experimental structure. Variations in the total potential energy were obscured by thermal fluctuations.

Hydrolytic activity

We investigated CTP hydrolytic activity using *p*-nitrophenyl acetate as the substrate. The hydrolytic activity was compared to that obtained for the peptides Ser-His, Ser-His-Asp and the free amino acids serine, histidine and aspartic acid, using the hydrolysis rate of free histidine as the reference. All amino acids and peptides tested were used at equimolar concentrations. The hydrolytic activity of CTP was monitored in the pH range 6.5–8.0. CTP had

TABLE 7
HYDROLYSIS OF *p*-NITROPHENYL ACETATE

Reactant	Relative activity
His	1.0
Glu	0.05
Ser	0.05
Ser + His + Asp	1.6
Ser-His	1.7
Ser-His-Asp	2.0
CTP	9.0

The activity of histidine was used as the standard.

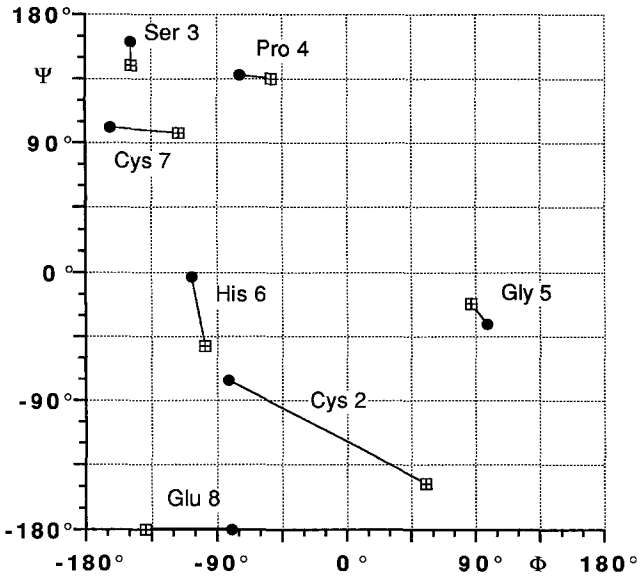

Fig. 13. Comparison of secondary structure for the annealed conformation of Fig. 12 to that of the experimental conformation. The symbols are as in Fig. 8, but with crossed squares for the annealed conformation. Note the improved agreement.

a pH optimum of 7.6–7.8, which is close to the pK_a of histidine. At lower pH values the activity was drastically reduced. Table 7 gives the hydrolytic activities at pH 7.6, and it is obvious that there is a cooperative effect between the functional amino acids, since Ser-His, Ser-His-Asp and CTP also exhibit faster hydrolysis than the individual amino acids or a mixture of amino acids. CTP gives 4–5 times faster hydrolysis than Ser-His-Asp. A likely reason for this is that the disulphide bridge restricts the peptide backbone motion and thereby increases the local concentration of the functionally important side chains. Since cysteine residues are potent nucleophiles, the activity was measured only for the oxidised form of the peptide. The flexibility of the involved side chains is substantial; this is particularly true for the side chain of Glu^8, which is prevented from direct hydrogen bonding of His^6. The high polarizability of water means that a water-mediated charge-relay system may work, but we expect that a more constrained triad would be more efficient.

Conclusions

The combination of theoretical tools with synthesis and structure determination enabled, as a first step, the realisation of a cyclic peptide structure that is about 10 times as efficient as free histidine in the hydrolysis of *p*-nitrophenyl acetate. A probable reason for this enhanced reactivity is an increased local concentration of the relevant side chains. Similar hydrolytic activities have been obtained for the substantially larger peptide chymohelizyme-1 [16]. The obtained activity is still far from those of real enzymes. For instance, chymotrypsin and papain are 250- and 25 000-fold more active, respectively, on *p*-nitrophenyl acetate than CTP.

The original, manually built model turned out to be incorrect for the Cys^2 Ψ and His^6 Ψ angles. This could not be rectified through relaxation by MD simulation at room temperature, since backbone rotations are much slower than the simulation time scale. High-temperature simulation produced the desired rearrangement of the disulphide bridge in some 200 ps and an rmsd relative to the experimental conformation of at best somewhat below 0.8 Å. A slower β to α transition for Ser^3 entailed increased flexibility and true high-temperature behaviour.

Based on the conformational data, the design could be extended in order to improve activity by replacing Glu^8 with D-Glu. Substrate binding to the reactant is another important factor for obtaining high reactivity [44] and we intend to explore this by coupling suitable affinity tails to the characterised peptide structure.

Acknowledgements

B.W., M.U. and T.D. acknowledge grants by the Swedish Board of Technology Development, NUTEK. L.B. was supported by the National Swedish Research Foundation and C.L. wishes to thank Klaus Mosbach for his support. We are also grateful to Dr. Leif Laaksonen for use of the graphics software SCARECROW [45] and to Peter Drakenberg for the MAGNE software.

References

1 DeGrado, W.F., Adv. Protein Chem., 39 (1988) 51.
2 Tuchscherer, G. and Mutter, M., J. Pept. Sci., 1 (1995) 3.
3 Rost, B. and Sander, C., J. Mol. Biol., 232 (1993) 584.
4 Wallqvist, A. and Ullner, M., Protein Struct. Funct. Genet., 18 (1994) 267.
5 Braxenthaler, M., Avbelj, F. and Moult, J., J. Mol. Biol., 250 (1995) 239.
6 Blow, D.M., Birktoft, J.J. and Hartley, B.B., Nature, 221 (1969) 337.
7 Blow, D.M. and Steitz, T.A., Annu. Rev. Biochem., 39 (1970) 63.
8 Birktoft, J.J. and Blow, D.M., J. Mol. Biol., 68 (1972) 187.
9 Kraut, J., Annu. Rev. Biochem., 46 (1977) 331.
10 Kraut, J., Science, 242 (1988) 533.
11 Carter, P. and Wells, J.A., Nature, 332 (1988) 564.
12 Tsukada, H. and Blow, D.M., J. Mol. Biol., 184 (1985) 703.
13 Craik, C.S., Roczniak, S., Langman, C. and Rutter, W.J., Science, 237 (1987) 909.
14 Warshel, A., Naray-Szabo, G. and Hwang, J.-K., Biochemistry, 28 (1989) 3629.
15 Metha, R.V., Mathur, K.B. and Dahr, M.M., Indian J. Chem., 15B (1977) 458.
16 Corey, M.J., Hallakova, E., Pugh, K. and Stewart, J.M., Appl. Biosci. Biotechnol., 47 (1994) 199; in this paper previous results were discussed: Hahn, K.W., Klis, W.A. and Stewart, J.M, Science, 248 (1990) 1544.
17 Bülow, L. and Mosbach, K., FEBS Lett., 210 (1987) 147.
18 Atassi, M.Z. and Manshouri, T., Proc. Natl. Acad. Sci. USA, 90 (1993) 8282.

19 a. Matthews, B.W., Craik, C.S. and Neurath, H., Proc. Natl. Acad. Sci. USA, 91 (1994) 4103.
b. Corey, D.R. and Phillips, M.A., Proc. Natl. Acad. Sci. USA, 91 (1994) 4106.
c. Wells, J.A., Fairbrother, W.J., Otlewski, J., Laskowski, M. and Burnier, J., Proc. Natl. Acad. Sci. USA, 91 (1994) 4110.
20 Matthews, B.W., Siegler, P.B., Henderson, R. and Blow, D.M., Nature, 214 (1967) 652.
21 Zimmermann, R. and Scheraga, H.A., Biopolymers, 16 (1977) 811.
22 Ahlström, P., Teleman, O., Jönsson, B. and Forsén, S., J. Am. Chem. Soc., 109 (1987) 1541.
23 Ahlström, P., Teleman, O., Kördel, C.-J., Forsén, S. and Jönsson, B., Biochemistry, 28 (1989) 3205.
24 Teleman, O., Jönsson, B. and Svensson, B., Comput. Phys. Commun., 62 (1991) 307.
25 Hermans, J., Berendsen, H.J.C., Van Gunsteren, W.F. and Postma, J.P.M., Biopolymers, 23 (1984) 1513.
26 Van Gunsteren, W.F. and Karplus, M., Macromolecules, 15 (1982) 1528.
27 Margenau, M. and Kestner, N.R., Theory of Intermolecular Forces, Pergamon Press, New York, NY, 1969.
28 Berendsen, H.J.C., Postma, J.P.M., Van Gunsteren, W.F. and Hermans, J., In Pullman, B. (Ed.) Intermolecular Forces, Reidel, Dordrecht, 1981, pp. 331–342.
29 Wallqvist, A. and Teleman, O., Mol. Phys., 74 (1991) 515.
30 Dolphin, D. and Wick, A.E., Tabulation of Infrared Data, Wiley, New York, NY, 1977.
31 Herzberg, G., Molecular Spectra and Molecular Structure: Infrared and Raman Spectra of Polyatomic Molecules, van Nostrand, Princeton, NJ, 1945.
32 Teleman, O. and Jönsson, B., J. Comput. Chem., 7 (1986) 58.
33 Teleman, O., Ahlström, P. and Jönsson, B., Mol. Simul., 7 (1991) 181.
34 Aue, W.P., Bartholdi, E. and Ernst, R.R., J. Chem. Phys., 64 (1976) 2229.
35 Braunschweiler, L. and Ernst, R.R., J. Magn. Reson., 53 (1983) 521.
36 Bax, A. and Davis, D.G., J. Magn. Reson., 63 (1985) 207.
37 States, D.J., Haberkorn, R.A. and Ruben, D.J., J. Magn. Reson., 48 (1982) 286.
38 Bax, A. and Davis, D.G., J. Magn. Reson., 65 (1985) 355.
39 Neuhaus, D. and Williamson, M., The Nuclear Overhauser Effect in Structural and Conformational Analysis, VCH, New York, NY, 1989.
40 Wagner, G., Braun, W., Havel, T.F., Schaumann, T., Gō, N. and Wüthrich, K., J. Mol. Biol., 196 (1987) 611.
41 Ullner, M., Selander, M., Persson, E., Stenflo, J., Drakenberg, T. and Teleman, O., Biochemistry, 31 (1992) 5974.
42 Linse, S., Teleman, O. and Drakenberg, T., Biochemistry, 29 (1990) 5925.
43 Wüthrich, K., NMR of Proteins and Nucleic Acids, 2nd ed., Wiley, New York, NY, 1986.
44 Fersht, A., Enzyme Structure and Mechanism, 2nd ed., Freeman, New York, NY, 1985.
45 Laaksonen, L., J. Mol. Graphics, 10 (1992) 33.