

# The Solution Conformation of the Antibacterial Peptide Cecropin A: A Nuclear Magnetic Resonance and Dynamical Simulated Annealing Study<sup>†</sup>

Tadeusz A. Holak,<sup>†</sup> Åke Engström,<sup>§</sup> Per J. Kraulis,<sup>||</sup> Gunnar Lindeberg,<sup>§</sup> Hans Bennich,<sup>§</sup> T. Alwyn Jones,<sup>||</sup>  
Angela M. Gronenborn,<sup>\*,†,‡</sup> and G. Marius Clore<sup>\*,†,‡</sup>

Max-Planck-Institut für Biochemie, D-8033 Martinsried bei München, FRG, Department of Immunology, Biomedical Centre, University of Uppsala, P.O. Box 582, S-75123 Uppsala, Sweden, and Department of Molecular Biology, Biomedical Centre, University of Uppsala, P.O. Box 591, S-75123 Uppsala, Sweden

Received February 23, 1988; Revised Manuscript Received April 8, 1988

**ABSTRACT:** The solution conformation of the antibacterial polypeptide cecropin A from the *Cecropia* moth is investigated by nuclear magnetic resonance (NMR) spectroscopy under conditions where it adopts a fully ordered structure, as judged by previous circular dichroism studies [Steiner, H. (1982) *FEBS Lett.* 137, 283-287], namely, 15% (v/v) hexafluoroisopropyl alcohol. By use of a combination of two-dimensional NMR techniques the <sup>1</sup>H NMR spectrum of cecropin A is completely assigned. A set of 243 approximate interproton distance restraints is derived from nuclear Overhauser enhancement (NOE) measurements. These, together with 32 distance restraints for the 16 intrahelical hydrogen bonds identified on the basis of the pattern of short-range NOEs, form the basis of a three-dimensional structure determination by dynamical simulated annealing [Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988) *FEBS Lett.* 229, 317-324]. The calculations are carried out starting from three initial structures, an  $\alpha$ -helix, an extended  $\beta$ -strand, and a mixed  $\alpha/\beta$  structure. Seven independent structures are computed from each starting structure by using different random number seeds for the assignments of the initial velocities. All 21 calculated structures satisfy the experimental restraints, display very small deviations from idealized covalent geometry, and possess good nonbonded contacts. Analysis of the 21 converged structure indicates that there are two helical regions extending from residues 5 to 21 and from residues 24 to 37 which are very well defined in terms of both atomic root mean square differences and backbone torsion angles. For the two helical regions individually the average backbone rms difference between all pairs of structures is  $\sim 1$  Å. The long axes of the two helices lie in two planes, which are at an angle of 70-100° to each other. The orientation of the helices within these planes, however, cannot be determined due to the paucity of NOEs between the two helices.

The principal component of the humoral immune system of the giant silk moth *Hyalophora cecropia*, as well as of a number of other insects, resides in a family of small proteins known as cecropins (Faye et al., 1975; Hultmark et al., 1980; Boman & Hultmark, 1987). They are induced in the hemolymph following injection of live bacteria and display strong bacteriolytic activity against a variety of Gram-positive and Gram-negative bacteria (Boman & Hultmark, 1981; Boman & Steiner, 1981). The first two cecropins to be isolated and sequenced were the A and B forms (Steiner et al., 1981; van Hofsten, 1985). They are composed of 37 and 35 residues, respectively, are  $\sim 70\%$  homologous, and consist of a strongly basic amino-terminal half followed by a hydrophobic, essentially neutral, C-terminal region. In contrast to other lytic peptides such as mellitin and apolipoprotein A1, the cecropins only have a lytic effect on the membranes of prokaryotic cells. Studies using synthetic peptides have shown that the 1-33 fragment of cecropin A is antibacterial (Merrifield et al., 1982) and that an aromatic residue in position 2 is important for

activity (Andreu et al., 1983, 1985). Circular dichroism (CD)<sup>1</sup> studies have indicated that although cecropins have a largely random coil conformation in water alone, they adopt an ordered, mainly helical (60-80%), conformation in the presence of only small amounts (10%) of hexafluoroisopropyl alcohol (HFP), a solvent that mimics a hydrophobic membranelike environment by reducing the water activity (Steiner, 1982; Andreu et al., 1985).

In the present study, we determine by NMR spectroscopy the solution conformation of cecropin A under conditions where it adopts an ordered structure (15% v/v HFP). With a variety of two-dimensional NMR techniques (Ernst et al., 1987), the <sup>1</sup>H NMR spectrum is assigned in a sequential manner, and a set of 243 approximate interproton distance restraints is derived from nuclear Overhauser enhancement (NOE) measurements. These form the basis of a three-dimensional structure determination by dynamical simulated annealing (Nilges et al., 1988a,b).

## EXPERIMENTAL PROCEDURES

**Sample Preparation.** Cecropin A was synthesized by automated solid-phase synthesis (Merrifield, 1963) on an Applied Biosystems (Foster City, CA) Model 430A peptide synthesizer by using a protocol similar to that described by Merrifield et al. (1982). Symmetric anhydrides were coupled twice, except

<sup>†</sup> This work was supported by the Max-Planck-Gesellschaft and Grant 321/4003/0318909A from the Bundesministerium für Forschung und Technologie (G.M.C. and A.M.G.), the Swedish Natural Science Research Council (T.A.J. and P.J.K.), Grant 3556 from the Swedish Medical Research Council (H.B.), and Grant 86-3787 from the Swedish National Board for Technological Development (H.B. and Å.E.).

<sup>‡</sup> Max-Planck-Institut für Biochemie.

<sup>§</sup> Department of Immunology, University of Uppsala.

<sup>||</sup> Department of Molecular Biology, University of Uppsala.

<sup>‡</sup> Present address: Laboratory of Chemical Physics, Building 2, National Institute of Diabetes and Digestive and Kidney Disorders, National Institutes of Health, Bethesda, MD 20892.

<sup>1</sup> Abbreviations: CD, circular dichroism; HFP, hexafluoroisopropyl alcohol or 1,1,1,3,3,3-hexafluoropropan-2-ol; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, two-dimensional NOE spectroscopy; HOHAHA, homonuclear Hartmann-Hahn spectroscopy; CD, circular dichroism; SD, standard deviation; rms, root mean square.

for those residues previously shown to give incomplete reactions and which required triple couplings. After cleavage with HF, the crude peptide was subjected to ion-exchange chromatography using a Mono-S column (Pharmacia AB, Uppsala, Sweden) in a sodium phosphate buffer, pH 6.5, containing 10% (v/v) methanol. Further purification was carried out by reversed-phase chromatography using a PEP-RPC column (Pharmacia AB) in 0.1% v/v trifluoroacetic acid and acetonitrile (gradient 20–40%). The purity of the final product was ascertained by  $^{252}\text{Cf}$ -plasma desorption time of flight mass spectroscopic analysis (Sundqvist & MacFarlane, 1985) on a BIOION Model 20 spectrometer (Bio-Ion Nordic AB, Uppsala, Sweden), as well as by Edman sequence analysis (Edman & Henschen, 1975).

The samples for NMR contained ~6 mM cecropin A in 2.5 mM sodium phosphate buffer, pH 5.0, and 15% HFP- $d_2$ ; in addition, one sample contained 85%  $\text{D}_2\text{O}$  and the other 77%  $\text{H}_2\text{O}$ /8%  $\text{D}_2\text{O}$ . These conditions were chosen on the basis of CD studies (Steiner, 1982; Andreu et al., 1982), which indicated that the transition from a random coil to a fully ordered conformation was complete between 10% and 15% (v/v) HFP.

**NMR Spectroscopy.** All NMR measurements were carried out at 500 MHz on a Bruker AM-500 spectrometer. All two-dimensional spectra were recorded in the pure-phase absorption mode by using the time proportional incrementation methods (Redfield & Kuntz, 1975; Bodenhausen et al., 1980; Marion & Wüthrich, 1983). The following spectra were recorded in both  $\text{D}_2\text{O}$  and  $\text{H}_2\text{O}$ : NOESY (Jeener et al., 1979; Macura et al., 1981) with mixing times of 100 and 200 ms and HOHAHA (Braunschweiler et al., 1983; Davis & Bax, 1985) with a MLEV17 mixing sequence (Bax & Davis, 1985) of 20–60-ms duration. All spectra were measured at 25 °C. In addition, a NOESY spectrum was acquired at 14 °C, the lowest temperature for which the resonance lines in the spectrum were not appreciably broadened. For the NOESY spectra in  $\text{H}_2\text{O}$ , the water resonance was suppressed by the use of a semiselective excitation pulse in which the last 90° pulse in the sequence was replaced by the jump-return sequence (90° $_x$ – $\tau$ –90° $_{-x}$ ) with the carrier placed at the position of the solvent (Plateau & Gueron, 1982) and a value of 80  $\mu\text{s}$  for  $t$ . For the HOHAHA spectra in  $\text{H}_2\text{O}$ , the water resonance was suppressed either by presaturating the water resonance or by adding the pulse sequence 90° $_x$ – $H$ – $\Delta$ –90° $_x$ – $\tau$ –90° $_{-x}$  at the end of the MLEV17 $_y$  sequence (Bax et al., 1987), where  $H$  is a homospoil pulse (4 ms),  $D$  a recovery delay (4 ms), and 90° $_x$ – $t$ –90° $_{-x}$  the semiselective jump-return sequence with a value of 80  $\mu\text{s}$  for  $t$ . For the majority of the spectra, 650–700  $t_1$  increments were collected, each with 4K data points, over a spectral width of 6 kHz in both dimensions.

**Calculations.** All minimization and dynamical simulated annealing calculations were carried out as described by Nilges et al. (1988a,b) on either VAX-8550 or CONVEX C1-XP2 computers with the program XPLOR (A. T. Brünger, unpublished data; Brünger et al., 1987a,b), which is a vectorized version of the program CHARMM (Brooks et al., 1983) especially adapted for restrained molecular dynamics. Displaying of structures was carried out with a modified version of the function network of FRODO (Jones, 1978) interfaced with XPLOR on an Evans & Sutherland PS390 color graphics system.

## RESULTS AND DISCUSSION

**Sequential Resonance Assignment and Secondary Structure.** Sequence-specific resonance assignment was carried out in a sequential manner [Wüthrich et al., 1982; Billeter et al.,

1982; for reviews, see Wüthrich (1986) and Clore and Gronenborn (1987)]. HOHAHA spectra (Davis & Bax, 1985), recorded at several mixing times ranging from 20 to 60 ms to detect successively direct, single and relayed through-bond (scalar) connectivities, were used to identify amino acid spin systems. NOESY spectra were used to observe through-space (<5 Å) connectivities. For the purpose of sequential assignment the most useful NOEs involve the NH, C $^{\alpha}\text{H}$ , and C $^{\beta}\text{H}$  protons, as well as the C $^{\delta}\text{H}$  protons of proline, and are of the type NH( $i$ )–NH( $i \pm 1$ ), C $^{\alpha}\text{H}(i)$ –NH( $i + 1, 3, 4$ ), and C $^{\alpha}\text{H}(i)$ –C $^{\beta}\text{H}(i + 3)$ . (Note that the sequential NOEs observed for C $^{\delta}\text{H}$  protons of proline are similar to those of the NH protons of the other amino acids.) Most of the methyl resonances of valines could be assigned stereospecifically by using the method of Zuiderweg et al. (1985). The  $^3J_{\alpha\beta}$  coupling constants for Val-26, Val-28, and Val-29 could be estimated to be larger than 10 Hz from the splitting along the F2 dimension of the C $^{\alpha}\text{H}$ –C $^{\beta}\text{H}$  cross peaks in the HOHAHA and NOESY spectra. Large values for  $^3J_{\alpha\beta}$  (10–12 Hz) corresponding to a trans orientation of the C $^{\alpha}\text{H}$  and C $^{\beta}\text{H}$  protons ( $g^+$  conformation about the C $^{\alpha}$ –C $^{\beta}$  bond) are generally found for valine residues in regular  $\alpha$ -helical segments (Zuiderweg et al., 1985). In this conformation, the C $^{\gamma}2\text{H}_3$  methyl group is closer to its own NH proton than the C $^{\gamma}1\text{H}_3$  group (IUPAC nomenclature). This is easily detected in the NOESY spectra. The  $\beta$ -protons for proline-24 were also assigned stereospecifically by examining relative intensities of C $^{\alpha}\text{H}$ –C $^{\beta}\text{H}$  cross peaks in the NOESY spectra (Clore et al., 1986a). Examples of NOESY and HOHAHA spectra are shown in Figures 1 and 2, and a complete list of assignments is given in Table I.

A summary of the short-range ( $|i - j| \leq 4$ ) NOEs involving the NH, C $^{\alpha}\text{H}$ , and C $^{\beta}\text{H}$  protons is shown in Figure 3 and provides the basis for a qualitative interpretation of the secondary structure (Wüthrich et al., 1984; Wüthrich, 1986; Clore & Gronenborn, 1987). The stretches of NH( $i$ )–NH( $i + 1$ ) NOEs together with the presence of C $^{\alpha}\text{H}(i)$ –NH( $i + 3, 4$ ) and C $^{\alpha}\text{H}(i)$ –C $^{\beta}\text{H}(i + 3)$  NOEs suggest that there are two helices extending from residues 5–21 and 24–37.

**Interproton Distance Restraints.** A set of 243 approximate interproton distance restraints, comprising 181 interresidue and 62 intraresidue distances, was derived from the NOESY spectra recorded at 25 °C with mixing times of 100 and 200 ms. The distances were classified into four distance ranges, 1.8–2.6, 1.8–3.1, 1.8–3.7, and 1.8–5.0 Å, corresponding to very strong, strong, medium, and weak NOEs, respectively (Williamson et al., 1985; Clore et al., 1986b, 1987a). The upper limits for distances involving either methyl or methylene proton were corrected for center averaging (Clore et al., 1987b) in the same manner as described by Wüthrich et al. (1983) for the equivalent pseudo-atom representation. In addition, a further 0.5 Å was added to NOEs involving methyl groups to account for the higher apparent intensity of methyl resonances (Clore et al., 1987c; Wagner et al., 1987).

The interproton distance restraints were supplemented by 32 restraints for the 16 NH( $i + 4$ )–O( $i$ ) intrahelical hydrogen bonds identified on the basis of C $^{\alpha}\text{H}(i)$ –NH( $i + 3, 4$ ) and C $^{\alpha}\text{H}(i)$ –NH( $i + 3$ ) NOEs (Wagner et al., 1987). For each hydrogen bond the NH( $i + 4$ )–O( $i$ ) and N( $i + 4$ )–O( $i$ ) distances were restrained to distance ranges of 1.3–2.3 and 2.3–3.3 Å, respectively.

The complete list of all distance restraints used in the computations is available as supplementary material.

**Structure Computation.** The approach used for the determination of the three-dimensional structure of cecropin involved the application of dynamical simulated annealing

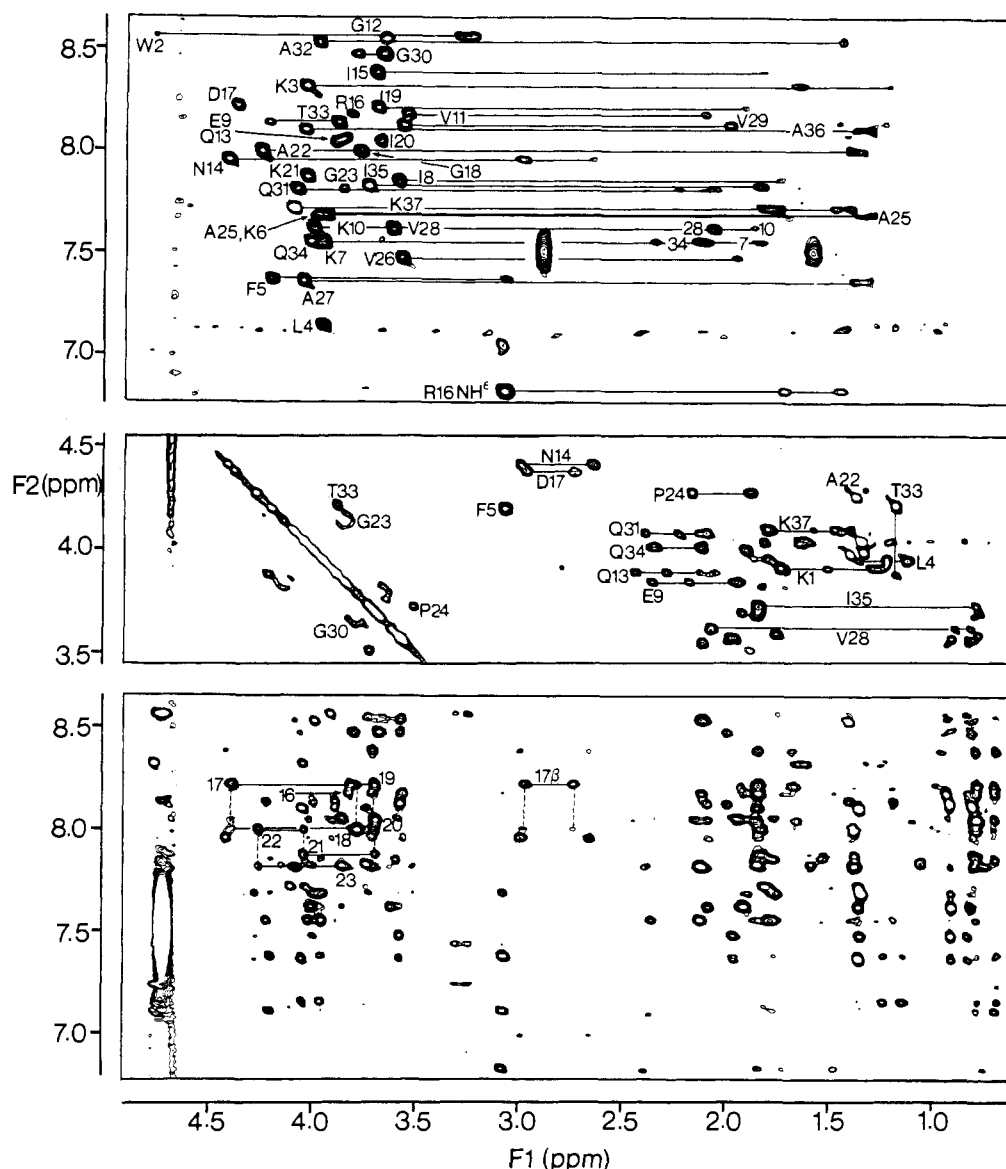


FIGURE 1: (a, top) HOHAHA spectrum (41-ms mixing time) of the NH(F2 axis)-aliphatic(F1 axis) region, (b, middle) HOHAHA spectrum (38-ms mixing time) of the C $\alpha$ H(F2 axis)-aliphatic(F1 axis) region, and (c, bottom) NOESY spectrum (200-ms mixing time) of the NH(F2 axis)-aliphatic(F1 axis) region of cecropin in 15% HFP at 25 °C. The cross peaks in (a) and (b) arise from direct and relayed through-bond connectivities. The cross peaks in (c) arise from through-space (<5 Å) connectivities, and some examples of C $\alpha$ H(*i*)-NH(*i* + 1) and C $\beta$ H(*i*)-NH(*i* + 1) connectivities are shown for residues 16–23 with the labeling at the position of the intraresidue C $\alpha$ H-NH and C $\beta$ H-NH cross peaks. (a) and (c) were recorded on the 77% H $_2$ O/8% D $_2$ O sample while (b) was recorded on the 85% D $_2$ O sample.

(Nilges et al., 1988). The basis of simulated annealing involves raising the temperature of the system followed by slow cooling in order to overcome local minima and locate the global minimum region of the target function (Kirkpatrick et al., 1983). This can be achieved either by using the Metropolis algorithm (Metropolis et al., 1953) or by solving Newton's equations of motion. Dynamical simulated annealing makes use of the latter and is thus similar in spirit to restrained molecular dynamics (Clare et al., 1985, 1986; Brünger et al., 1986, 1987a,b; Kaptein et al., 1985).

Dynamical simulated annealing differs from restrained molecular dynamics, however, insofar as all nonbonded interaction terms of the empirical energy function used in molecular dynamics (i.e., the Lennard-Jones van der Waals, hydrogen-bonding, electrostatic, and dihedral potentials) are replaced by a single hard-sphere repulsion term,  $F_{\text{repel}}$ , to prevent unduly close nonbonded contacts (Nilges et al., 1988):

$$F_{\text{repel}} = 0, \quad \text{if } r \geq sr_{\min} \\ k_{\text{rep}}(s^2r_{\min}^2 - r^2)^2, \quad \text{if } r < sr_{\min} \quad (1)$$

$k_{\text{rep}}$  is a variable force constant,  $r_{\min}$  is the standard van der Waals radii as represented by the Lennard-Jones potential used in the CHARMM empirical energy function (Brooks et al., 1983), and  $s$  is a van der Waals radius scale factor (set to 0.8 in our case).

Thus the total target function for which the global minimum is searched is made up of the terms:

$$F_{\text{tot}} = F_{\text{covalent}} + F_{\text{repel}} + F_{\text{NOE}} + F_{\phi} \quad (2)$$

$F_{\text{covalent}}$  maintains correct bond lengths, angles, planes, and chirality. The force constants for the bond, angle, and improper torsions are set to uniform high values to ensure near perfect stereochemistry of single amino acids throughout the calculation, namely, 600 kcal·mol $^{-1}$ ·Å $^{-2}$ , 500 kcal·mol $^{-1}$ ·rad $^{-2}$ , and 500 kcal·mol $^{-1}$ ·rad $^{-2}$ , respectively. (Note that the improper torsion terms serve to maintain appropriate planarity and chirality; unlike conventional dihedral angle restraints they have only one minimum described by a single harmonic potential. Further, the peptide bond is assumed to be planar and

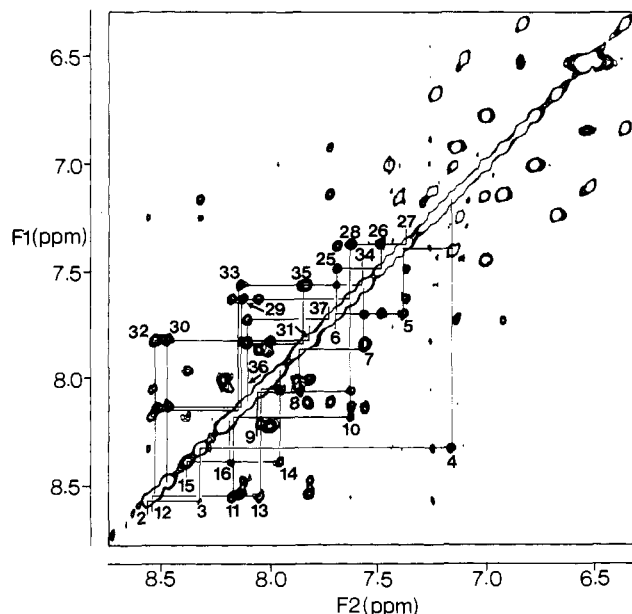


FIGURE 2: NH(F1 axis)–NH(F2 axis) region of the NOESY spectrum (200-ms mixing time) of cecropin in 15% (v/v) HFP, 77% H<sub>2</sub>O, and 8% D<sub>2</sub>O at 25 °C, displaying through-space connectivities. NH-(i)–NH(i + 1) connectivities for residues 2–16 and 25–37 are shown below and above the diagonal, respectively. Cross peaks are labeled according to the F2 chemical shift of the corresponding NH proton.

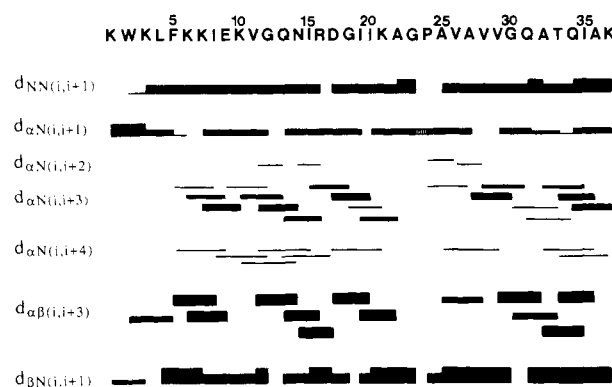


FIGURE 3: Summary of the short-range ( $|i - j| \leq 4$ ) connectivities involving the NH, C $\alpha$ H, and C $\delta$ H protons. The intensities of the NOEs are reflected in the thickness of the lines. The C $\alpha$ H(Gly-24)–C $\delta$ H(Pro-25) NOE is shown as a hatched box along the same line as the C $\alpha$ H(i)–NH(i + 1) connectivities.

trans.)  $F_{\text{NOE}}$  is a square-well potential for the experimental distance restraints (Clare et al., 1986b) and has the form:

$$F_{\text{NOE}} = \begin{cases} k_{\text{NOE}}(r_{ij} - r_{ij}^u)^2, & \text{if } r_{ij} > r_{ij}^u \\ 0, & \text{if } r_{ij}^l \leq r_{ij} \leq r_{ij}^u \\ k_{\text{NOE}}(r_{ij} - r_{ij}^l)^2, & \text{if } r_{ij} < r_{ij}^l \end{cases} \quad (3)$$

where  $k_{\text{NOE}}$  is the NOE restraints force constant,  $r_{ij}$  is the calculated distance, and  $r_{ij}^u$  and  $r_{ij}^l$  are the upper and lower limits of the experimental distance restraints. Finally,  $F_\phi$  is a square-well dihedral potential term (Clare et al., 1986b) with a force constant of 200 kcal·mol<sup>-1</sup>·Å<sup>-2</sup> used to restrict the  $\phi$  backbone torsion angles of all residues, except Gly and Asn, to the allowed region of the  $\phi$ ,  $\psi$  map ( $\phi = -5^\circ$  to  $-180^\circ$ ) (Brünger et al., 1987b). The form of  $F_\phi$  is analogous to that of  $F_{\text{NOE}}$ . The basis for the latter lies in the observation that the  $\phi$  backbone torsion angles of all residues lie within this range in high-resolution crystal structures, except in the case of Gly or Asn residues in type II turns (Richardson, 1981).

Three starting structures (Figure 4a) were used for the calculations: an  $\alpha$ -helix, denoted as Ini $\alpha$  ( $\phi$ ,  $\psi = -57^\circ, -47^\circ$ );

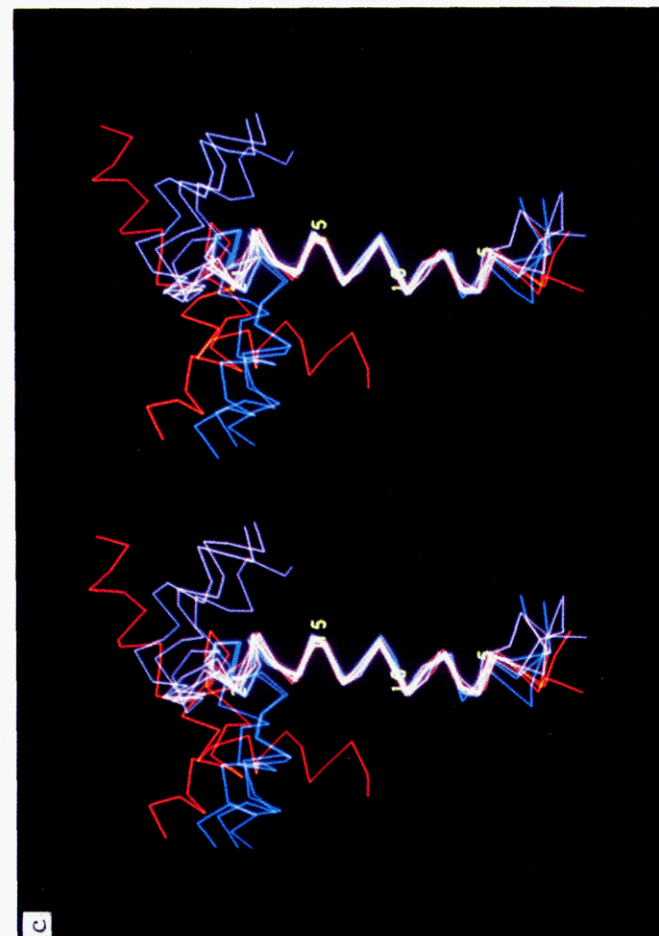
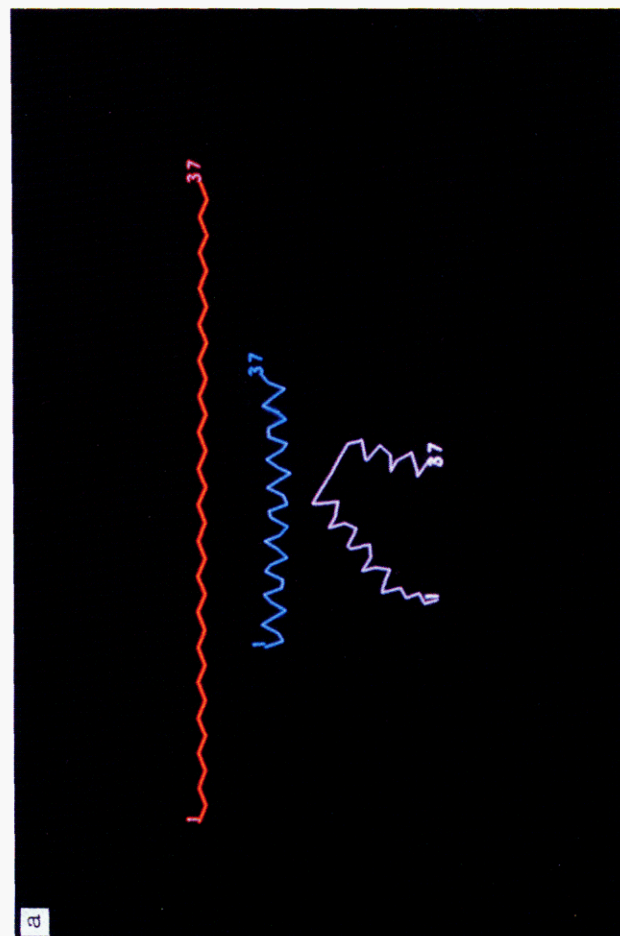
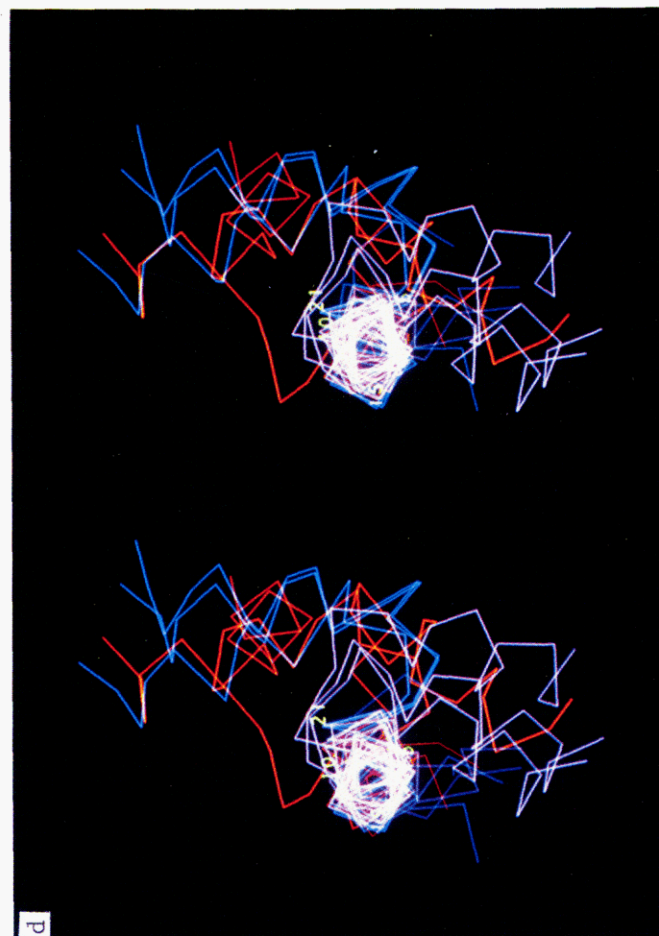
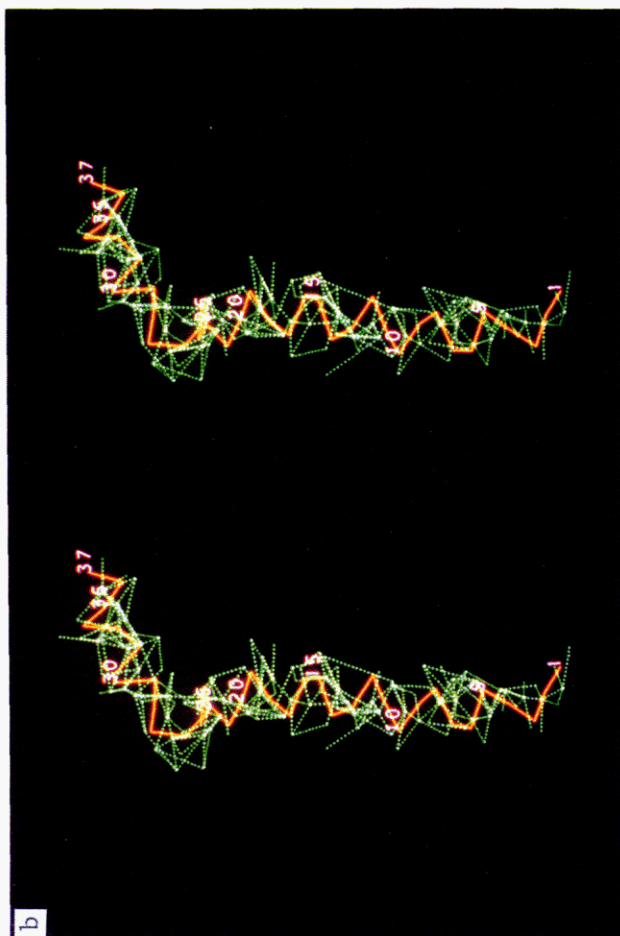
Table I: Proton Resonance Assignment of Cecropin A in 15% HFP, pH 5.0, 25 °C

residue	NH	C $\alpha$ H	C $\delta$ H	others
Lys-1	NO <sup>a</sup>	3.91	1.74	C $\gamma$ H <sub>2</sub> 1.32; C $\delta$ H <sub>2</sub> 1.54; C $\epsilon$ H <sub>2</sub> 2.81
Trp-2	8.56	4.76	3.24, 3.31	C $\delta$ H <sup>7</sup> 2.24; NH <sup>8</sup> 9.88; C $\epsilon$ H <sup>9</sup> 7.43; C $\epsilon$ H <sup>10</sup> 6.98
Lys-3	8.32	4.04	1.64, 1.69	C $\gamma$ H <sup>11</sup> 7.13; C $\delta$ H <sup>12</sup> 7.39
Leu-4	7.17	3.95	1.14, 1.24	C $\gamma$ H <sup>13</sup> 1.27, 1.37; C $\delta$ H <sup>14</sup> 1.64; C $\epsilon$ H <sup>15</sup> 2.90
Phe-5	7.38	4.19	3.07	C $\gamma$ H <sup>16</sup> 1.37; C $\delta$ H <sup>17</sup> 0.68, 0.78
Lys-6	7.68	3.93	1.77	C $\delta$ H <sup>18</sup> 6.99; C $\epsilon$ H <sup>19</sup> 7.23; C $\epsilon$ H <sup>20</sup> 7.23
Lys-7	7.55	3.94	1.83, 1.88	C $\gamma$ H <sup>21</sup> 1.30, 1.44; C $\delta$ H <sup>22</sup> 1.63; C $\epsilon$ H <sup>23</sup> 2.89
Ile-8	7.85	3.60	1.77	C $\gamma$ H <sup>24</sup> 1.24, 1.39; C $\delta$ H <sup>25</sup> 1.56; C $\epsilon$ H <sup>26</sup> 2.78
Glu-9	8.05	3.85	1.95	C $\gamma$ H <sup>27</sup> 1.07, 1.54; C $\delta$ H <sup>28</sup> 0.80; C $\epsilon$ H <sup>29</sup> 0.70
Lys-10	7.62	4.00	1.93	C $\gamma$ H <sup>30</sup> 2.18, 2.37
Val-11	8.16	3.57	2.12	C $\gamma$ H <sup>31</sup> 1.47; C $\delta$ H <sup>32</sup> 1.61; C $\epsilon$ H <sup>33</sup> 2.89
Gly-12	8.52	3.67	3.70 ( $\alpha$ )	C $\gamma$ H <sup>34</sup> 0.84, 0.94
Gln-13	8.03	3.90	2.08, 2.16	C $\gamma$ H <sup>35</sup> 2.30, 2.46; NH <sup>36</sup> 6.66, 7.20
Asn-14	7.95	4.42	2.67, 2.99	NH <sup>37</sup> 6.76, 6.98
Ile-15	8.37	3.70	1.85	C $\gamma$ H <sup>38</sup> 0.98, 1.67; C $\delta$ H <sup>39</sup> 0.84; C $\epsilon$ H <sup>40</sup> 0.70
Arg-16	8.16	3.82	1.85	C $\gamma$ H <sup>41</sup> 1.47, 1.73; C $\delta$ H <sup>42</sup> 3.07; NH <sup>43</sup> 6.84
Asp-17	8.20	4.37	2.74, 2.97	
Gly-18	7.98	3.77		
Ile-19	8.19	3.71	1.94	C $\gamma$ H <sup>44</sup> 0.98, 1.68; C $\delta$ H <sup>45</sup> 0.80; C $\epsilon$ H <sup>46</sup> 0.70
Ile-20	8.03	3.70	1.84	C $\gamma$ H <sup>47</sup> 1.10, 1.66; C $\delta$ H <sup>48</sup> 0.81; C $\epsilon$ H <sup>49</sup> 0.77
Lys-21	7.87	4.04	1.83	C $\gamma$ H <sup>50</sup> 1.52; C $\delta$ H <sup>51</sup> 1.60; C $\epsilon$ H <sup>52</sup> 2.89
Ala-22	7.99	4.26	1.38	
Gly-23	7.81	3.85	4.15 ( $\alpha$ )	
Pro-24	4.27	1.90, 2.19		C $\gamma$ H <sup>53</sup> 1.94; C $\delta$ H <sup>54</sup> 3.52, 3.73
Ala-25	7.68	3.99	1.36	
Val-26	7.48	3.58	1.96	C $\gamma$ H <sup>55</sup> 0.83; C $\delta$ H <sup>56</sup> 0.91
Ala-27	7.37	4.05	1.36	
Val-28	7.62	3.62	2.09	C $\gamma$ H <sup>57</sup> 0.83; C $\delta$ H <sup>58</sup> 0.91
Val-29	8.11	3.57	1.99	C $\gamma$ H <sup>59</sup> 0.81; C $\delta$ H <sup>60</sup> 0.92
Gly-30	8.46	3.67	3.80 ( $\alpha$ )	
Gln-31	7.81	4.08	2.12	C $\gamma$ H <sup>61</sup> 2.26, 2.38; NH <sup>62</sup> 6.35, 6.83
Ala-32	8.52	3.99	1.42	
Thr-33	8.13	3.89	4.22	C $\gamma$ H <sup>63</sup> 1.20
Gln-34	7.56	4.02	2.13	C $\gamma$ H <sup>64</sup> 2.36; NH <sup>65</sup> 6.50, 7.10
Ile-35	7.82	3.74	1.85	C $\gamma$ H <sup>66</sup> 1.08, 1.59; C $\delta$ H <sup>67</sup> 0.81; C $\epsilon$ H <sup>68</sup> 0.74
Ala-36	8.09	4.04	1.35	
Lys-37	7.71	4.10	1.81	C $\gamma$ H <sup>69</sup> 1.43, 1.50; C $\delta$ H <sup>70</sup> 1.61; C $\epsilon$ H <sup>71</sup> 2.89; CONH <sup>72</sup> 6.92, 7.14

<sup>a</sup>NO, not observed.

an extended  $\beta$ -strand denoted as Ini $\beta$  ( $\phi$ ,  $\psi = -139^\circ, 135^\circ$ ); and a mixed  $\alpha\beta$  structure, denoted as Ini $\alpha\beta$ , in which residues 5–21 and 25–37 were  $\alpha$ -helices, as predicted from the qualitative interpretation of the NOE data, and residues 1–4 and 22–24  $\beta$  strands. All the side chains were placed in an extended conformation ( $\chi_i = 180^\circ$ ).

The annealing schedule was carried out in two steps. Step 1 comprised 50 cycles of 75-fs dynamics each. The initial velocities were chosen from a Maxwellian distribution at 1000 K, and after each cycle the velocities were rescaled to this temperature. The value of the force constant for the experimental distance restraints was doubled at the beginning of each cycle from an initial value of 0.2 kcal·mol<sup>-1</sup>·Å<sup>-2</sup> to a maximum value of 50 kcal·mol<sup>-1</sup>·Å<sup>-2</sup>. To make rearrangement easily possible, the force constant for the repulsion term was





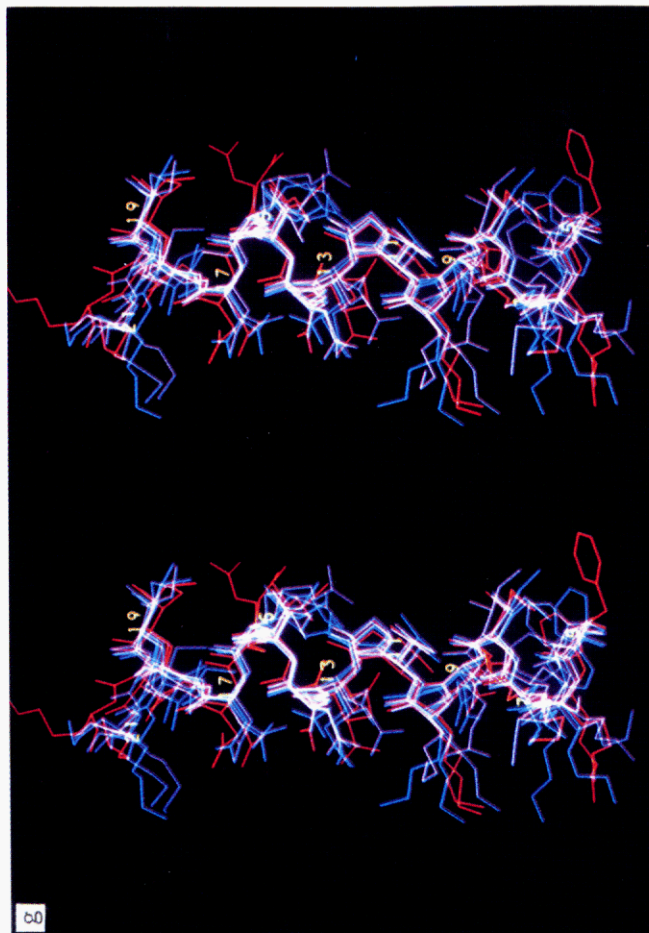
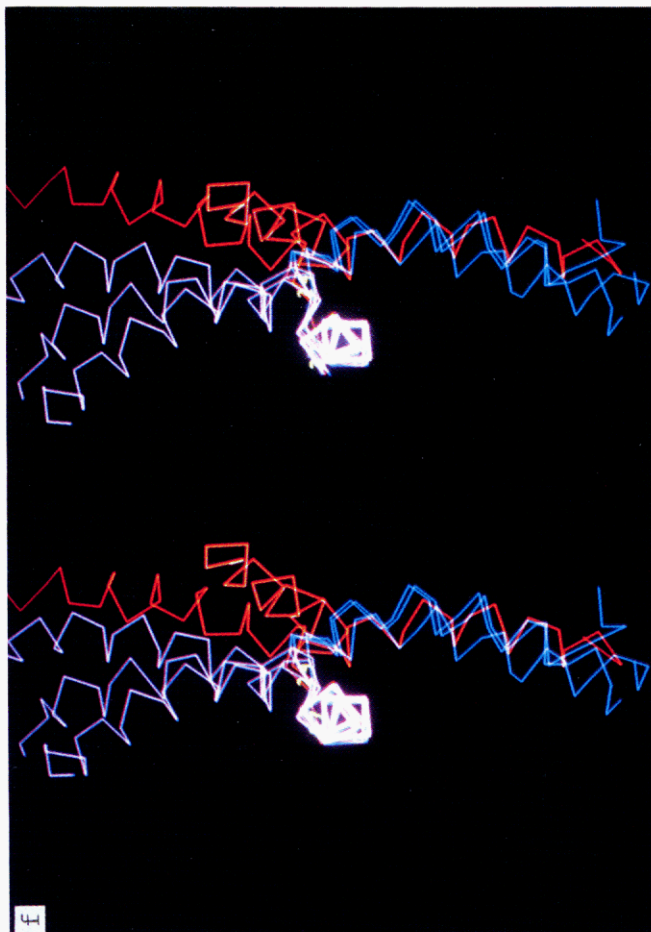
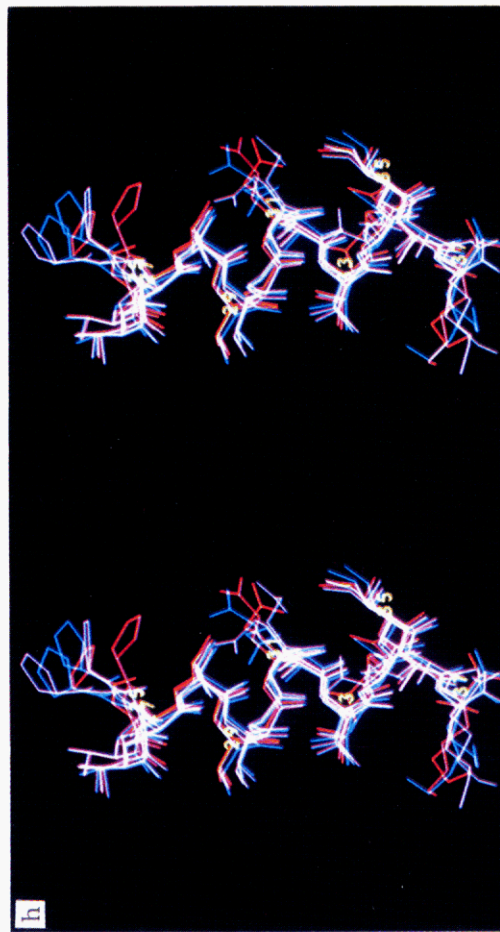


FIGURE 4: Stereoviews of the initial and final structures. The color code used throughout is as follows: the  $\text{Ini}\alpha$  and  $\text{SA}\alpha$  structures are shown in blue, the  $\text{Ini}\alpha\beta$  and  $\text{SA}\alpha\beta$  structures in magenta, and the  $\text{Ini}\beta$  and  $\text{SA}\beta$  structures in red. (a) The three initial structures; (b) superposition of the NOE distance restraints (green) on the  $\text{C}^\alpha$  backbone of one the final  $\text{SA}\beta$  structures; (c and d) superposition of the  $\text{C}^\alpha$  backbones of three  $\text{SA}\alpha$ , three  $\text{SA}\alpha\beta$ , and three  $\text{SA}\beta$  structures best fitted to residues 5-21 viewed along and down the long axis of helix 1, respectively; (e and f) superposition of the



$\text{C}^\alpha$  backbones of three  $\text{SA}\alpha$ , three  $\text{SA}\alpha\beta$ , and three  $\text{SA}\beta$  structures best fitted to residues 24-37 viewed along and down the long axis of helix 2, respectively; (g) all atom (excluding protons) best fit superposition of residues 5-21 of two  $\text{SA}\alpha$ , two  $\text{SA}\alpha\beta$ , and two  $\text{SA}\beta$  structures; (h) all atom (excluding protons) best fit superposition of residues 24-37 of two  $\text{SA}\alpha$ , two  $\text{SA}\alpha\beta$ , and two  $\text{SA}\beta$  structures.

multiplied by a factor of  $400^{1/50}$  prior to each cycle from an initial value of  $0.01 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}^{-2}$  to a final value of  $4 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}^{-2}$ . Step 2 consisted of 1.5-ps dynamics at 300 K with values of the force constants  $k_{\text{NOE}}$  and  $k_{\text{rep}}$  equal to their final values at the end of step 1. The velocities were rescaled to 300 K every 150 fs during this period. This was followed by 200 cycles of restrained Powell minimization. The entire procedure takes  $\sim 45$  min on a VAX-8550 computer. It should be noted that the values of the force constants are somewhat arbitrary. They do, however, maintain nearly perfect covalent geometry and ensure that no unduly close contacts occur and that the experimental restraints are effectively introduced and satisfied.

In the case of the simulations starting from  $\text{Ini}\alpha$  and  $\text{Ini}\beta$ , the dynamical simulated annealing schedule was carried out only once per structure calculation. For the computations starting off from  $\text{Ini}\beta$ , however, we found that the entire procedure had to be repeated twice with different random number seeds in order to obtain convergence with respect to the experimental restraints (i.e., the structure obtained after the first pass through the annealing schedule was used as the input structure for a second pass). Seven independent calculations were carried out from each initial structure by using different random number seeds for the assignment of the initial velocities. The seven structures calculated from  $\text{Ini}\alpha$  are referred to collectively as  $\langle \text{SA}\alpha \rangle$ , those from  $\text{Ini}\beta$  as  $\langle \text{SA}\beta \rangle$ , and those from  $\text{Ini}\alpha\beta$  as  $\langle \text{SA}\alpha\beta \rangle$ . All 21 final structures are referred to collectively as  $\langle \text{SA} \rangle$ . The mean structure obtained by best fitting the coordinates of the 21 final structures to residues 5–21 is referred to as  $\overline{\text{SA}}(5\text{--}21)$  and that obtained by best fitting to residues 24–37 as  $\overline{\text{SA}}(24\text{--}37)$ .

The structural statistics of the initial and final structures are given in Table II. Although the initial structures have good covalent geometry, they all exhibit large deviations with respect to the experimental data and poor nonbonded contacts. All the  $\langle \text{SA} \rangle$  structures, on the other hand, satisfy the experimental restraints within  $0.5 \text{ \AA}$ , have good nonbonded contacts as evidenced by low values of  $F_{\text{repel}}$  and negative Lennard-Jones van der Waals energies, and display very small deviations from idealized covalent geometry.

**Converged Structures.** Superpositions of the final  $\langle \text{SA} \rangle$  structures are shown in Figure 4 and plots of atomic rms distributions and  $\phi$ ,  $\psi$  rms difference in Figures 5 and 6, respectively. Atomic rms difference between the various structures are given in Table III.

The main structural feature to emerge from these data is that there are two helical regions extending from residues 5–21 and residues 24–37 that are very well defined in terms of both atomic rms differences (Figures 4 and 5) and backbone torsion angles (Figure 6). The average backbone atomic rms difference between all pairs of structures for the two helical regions individually is  $<1.0 \text{ \AA}$ , and the average rms difference in the  $\phi$ ,  $\psi$  angles is  $<20^\circ$ . Indeed, the overall average values of  $-61 \pm 17^\circ$  and  $-41 \pm 14^\circ$  for the  $\phi$  and  $\psi$  angles, respectively, in these two regions are very close to those of an idealized  $\alpha$ -helix ( $\phi = -56^\circ$ ,  $\psi = -47^\circ$ ). The atomic rms shifts for the helical regions from the two initial structures  $\text{Ini}\alpha$  and  $\text{Ini}\alpha\beta$  to the final  $\langle \text{SA} \rangle$  structures are clearly small ( $1\text{--}1.5 \text{ \AA}$ ). Confidence in the conformation of these regions in the  $\langle \text{SA} \rangle$  structures is afforded, however, by the excellent convergence from  $\text{Ini}\beta$ , which is accompanied by backbone atomic rms shifts of  $9.7 \pm 0.1 \text{ \AA}$  for helix 1 and  $8.4 \pm 0.3 \text{ \AA}$  for helix 2. In addition, we note that many of the side-chain positions are also well defined within the two helical regions (Figures 4g,h and 5).

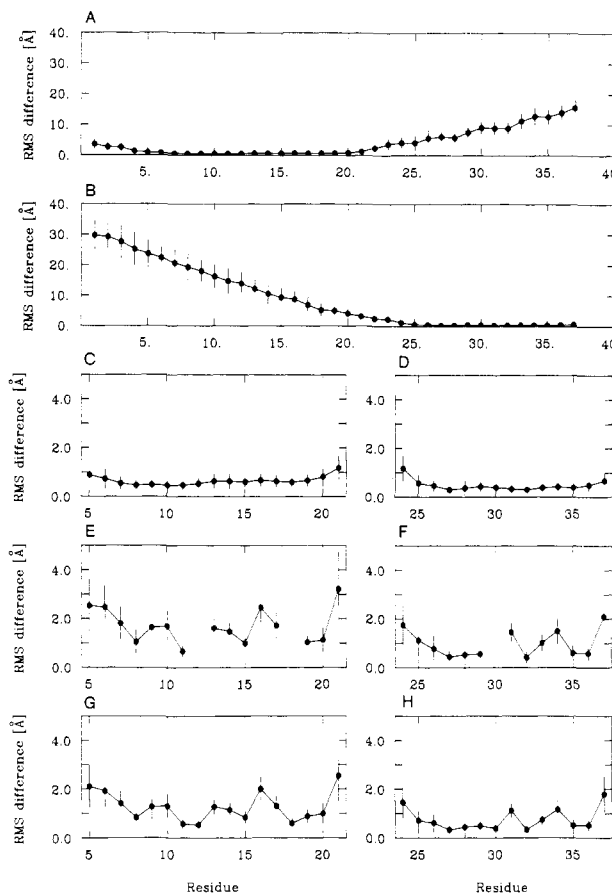


FIGURE 5: Atomic rms distributions of the individual 21  $\langle \text{SA} \rangle$  structures about the mean structures  $\overline{\text{SA}}(5\text{--}21)$  or  $\overline{\text{SA}}(24\text{--}37)$ . (A)  $\langle \text{SA} \rangle$  vs  $\overline{\text{SA}}(5\text{--}21)$  for all residues; (B)  $\langle \text{SA} \rangle$  vs  $\overline{\text{SA}}(24\text{--}37)$  for all residues; (C, E, G) residues 5–21 of  $\langle \text{SA} \rangle$  vs  $\overline{\text{SA}}(5\text{--}21)$  for the backbone atoms, side-chain atoms, and all atoms, respectively; (D, F, H) residues 24–37 of  $\langle \text{SA} \rangle$  vs  $\overline{\text{SA}}(24\text{--}37)$  for the backbone atoms, side-chain atoms, and all atoms, respectively. The backbone atoms comprise the C, C $\alpha$ , N, and O atoms. The filled in circles ( $\bullet$ ) represent the average values and the bars the standard deviations in these values.

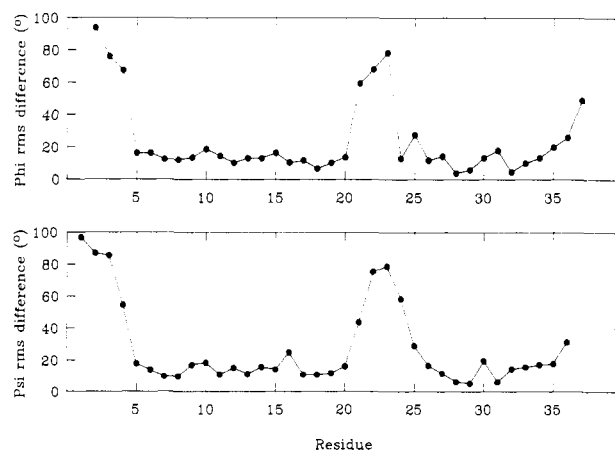


FIGURE 6: Average angular  $\phi$  and  $\psi$  rms differences between all pairs of  $\langle \text{SA} \rangle$  structures.

The exact orientation of the two helices, with respect to each other, cannot be defined. This is clearly evident in the atomic rms distribution plots (Figure 5) and by the large angular rms differences in the  $\phi$ ,  $\psi$  angles of residues 21–24 (Figure 6). This is not very surprising as all the interresidue NOEs are of a short-range nature, and only two NOEs involving residues more than four apart in the sequence could be detected, even in the NOESY spectrum at  $14^\circ\text{C}$ . These are the NOEs

Table II: Structural Statistics<sup>a</sup>

	initial structures			final structures		
	Ini $\alpha$	Ini $\alpha\beta$	Ini $\beta$	$\langle SA\alpha \rangle$	$\langle SA\alpha\beta \rangle$	$\langle SA\beta \rangle$
rms deviations from exptl restraints ( $\text{\AA}$ ) <sup>b</sup> (violations $>0.5 \text{ \AA}$ )						
all (275)	0.549 (15)	0.931 (22)	4.277 (136)	$0.039 \pm 0.002$ (0)	$0.041 \pm 0.008$ (0)	$0.045 \pm 0.00$ (0)
interresidue (181)	0.672 (12)	1.144 (19)	3.819 (99)	$0.045 \pm 0.004$ (0)	$0.047 \pm 0.010$ (0)	$0.051 \pm 0.00$ (0)
intraresidue (62)	0.148 (3)	0.150 (3)	0.778 (15)	$0.015 \pm 0.014$ (0)	$0.022 \pm 0.014$ (0)	$0.020 \pm 0.01$ (0)
H bond (32)	0.000 (0)	0.000 (0)	8.587 (32)	$0.030 \pm 0.007$ (0)	$0.031 \pm 0.009$ (0)	$0.031 \pm 0.01$ (0)
$F_{\text{NOE}}$ (kcal·mol <sup>-1</sup> ) <sup>c</sup>	$4.2 \times 10^3$	$1.2 \times 10^4$	$2.5 \times 10^5$	$20.0 \pm 3.0$	$24.5 \pm 10.0$	$27.6 \pm 3.9$
$F_{\phi}$ (kcal·mol <sup>-1</sup> ) <sup>d</sup>	0	0	0	0	$2.9 \pm 2.1$	0
$F_{\text{repel}}$ (kcal·mol <sup>-1</sup> ) <sup>e</sup>	1633	2272	298	$18.3 \pm 3.0$	$19.2 \pm 7.8$	$17.6 \pm 2.3$
$E_{\text{L-J}}$ (kcal·mol <sup>-1</sup> ) <sup>f</sup>	$4.3 \times 10^3$	$1.5 \times 10^3$	165	$-72.1 \pm 6.3$	$-71.0 \pm 7.5$	$-69.8 \pm 6.2$
deviations from idealized geometry <sup>g</sup>						
bonds ( $\text{\AA}$ ) (604)	0.016	0.042	0.042	$0.007 \pm 0.0003$	$0.007 \pm 0.0005$	$0.007 \pm 0$
angles (deg) (1109)	2.583	3.096	3.096	$1.936 \pm 0.008$	$1.935 \pm 0.018$	$1.932 \pm 0.0$
impropers (deg) (151)	1.151	0.245	0.242	$0.366 \pm 0.019$	$0.396 \pm 0.084$	$0.315 \pm 0.0$

<sup>a</sup> The notation of the structures is as follows: Ini $\alpha$ , Ini $\beta$ , and Ini $\alpha\beta$  are the three initial structures, namely, an  $\alpha$ -helix, an extended  $\beta$ -strand, and a mixed  $\alpha/\beta$  structure, respectively (see text).  $\langle SA\alpha \rangle$ ,  $\langle SA\beta \rangle$ , and  $\langle SA\alpha\beta \rangle$  are the three final categories of converged structures derived from Ini $\alpha$ , Ini $\beta$ , and Ini $\alpha\beta$ , respectively. There are seven individual structures in each final category (see text). <sup>b</sup> The rms deviations from the experimental restraints are calculated with respect to the upper and lower limits of the distance restraints (Clare et al., 1986b). The number of violations  $>0.5 \text{ \AA}$  is given in parentheses below the values for the rms deviations. The number of distances in each category is given in parentheses next to the category name. <sup>c</sup> The values of the square-well NOE potential  $F_{\text{NOE}}$  are calculated with a force constant of  $50 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}^{-2}$ . <sup>d</sup> The values of  $F_{\phi}$  are calculated with a force constant of  $200 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{rad}^{-2}$ . <sup>e</sup>  $F_{\phi}$  is a square-well dihedral potential that is used to restrict the  $\phi$  backbone torsion angles of all residues, except Gly and Asn, to a range between  $-5^\circ$  and  $-180^\circ$  during the entire calculation. <sup>f</sup> The values of the van der Waals repulsion term  $F_{\text{repel}}$  (cf. eq 1) are calculated with a force constant of  $4 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}^{-2}$  with the hard-sphere van der Waals radii set to 0.8 times the standard values used in the CHARMM empirical function (Brooks et al., 1983). <sup>g</sup>  $E_{\text{L-J}}$  is the Lennard-Jones van der Waals energy calculated with the CHARMM empirical energy function (Brooks et al., 1983). <sup>h</sup> The number of bond, angle, and improper terms is given in parentheses. The improper terms serve to maintain planarity and appropriate chirality.

Table III: Atomic rms Differences<sup>a</sup>

	atomic rms difference ( $\text{\AA}$ )					
	all residues		residues 5-21		residues 24-37	
	backbone atoms	all atoms	backbone atoms	all atoms	backbone atoms	all atoms
(A) rms Differences between Initial Structures						
Ini $\alpha$ vs Ini $\alpha\beta$	9.8	10.6	0.3	0.7	0.4	1.0
Ini $\alpha$ vs Ini $\beta$	21.2	22.4	9.9	11.0	8.1	9.1
Ini $\alpha\beta$ vs Ini $\beta$	27.9	29.4	9.7	10.7	8.0	9.1
(B) rms Shifts						
Ini $\alpha$ vs $\langle SA\alpha \rangle$	$8.4 \pm 0.7$	$9.3 \pm 0.7$	$1.2 \pm 0.3$	$2.1 \pm 0.2$	$1.5 \pm 0.1$	$2.1 \pm 0.2$
Ini $\alpha\beta$ vs $\langle SA\alpha\beta \rangle$	$5.3 \pm 0.2$	$6.6 \pm 0.3$	$0.9 \pm 0.3$	$2.0 \pm 0.5$	$1.5 \pm 0.1$	$2.1 \pm 0.2$
Ini $\beta$ vs $\langle SA\beta \rangle$	$27.2 \pm 2.5$	$28.1 \pm 2.4$	$9.7 \pm 0.1$	$10.7 \pm 0.2$	$8.4 \pm 0.3$	$9.4 \pm 0.3$
(C) rms Distributions of Final Structures						
$\langle SA \rangle$ vs $\langle SA \rangle$			$1.0 \pm 0.3$	$2.1 \pm 0.4$	$0.8 \pm 0.2$	$1.4 \pm 0.3$
$\langle SA \rangle$ vs $\overline{SA}(5-21)$			$0.7 \pm 0.2$	$1.5 \pm 0.3$		
$\langle SA \rangle$ vs $\overline{SA}(24-37)$					$0.5 \pm 0.1$	$0.9 \pm 0.2$

<sup>a</sup> The notation of the structures is the same as that in Table II. In addition,  $\langle SA \rangle$  comprise all 21 final structures (i.e., seven  $SA\alpha$ , seven  $SA\alpha\beta$ , and seven  $SA\beta$  structures), and  $\overline{SA}(5-21)$  and  $\overline{SA}(24-37)$  are the mean structures obtained by averaging the coordinates of the 21 final  $\langle SA \rangle$  structures best fitted to residues 5-21 and 24-37, respectively.

between Ile-20 C $\gamma^2$ H<sub>3</sub> and Ala-25 C $\alpha$ H and between Ala-22 C $\alpha$ H and Val-28 C $\gamma^2$ H<sub>3</sub>, which provide the only distance restraints bridging the two helices. In this respect, it is interesting to note that the relative orientation of the two helices in the seven  $SA\alpha$  structures is similar with an overall average backbone atomic rms difference between all pairs of structures of  $2.0 \pm 0.5 \text{ \AA}$ . The same is true of the seven  $SA\alpha\beta$  structures (overall average backbone atomic rms difference of  $1.8 \pm 0.4 \text{ \AA}$ ), although the orientation of the helices in these structures is quite different from that in the  $SA\alpha$  structures (Figure 4c-f), as well as that in Ini $\alpha\beta$ . In the case of the  $SA\beta$  structures, however, the orientation of the two helices is different in all the structures (Figure 4c-f).

Further analysis of the  $\langle SA \rangle$  structures reveals that the two planes in which the helices lie (i.e., the planes parallel to the helix axes) are reasonably well defined. The angle between these two planes ranges from  $70^\circ$  to  $100^\circ$ . What is not defined

is in fact the position of the helices within each plane. This is clearly evident in the superpositions best fitted to one or the other helix. Thus when residues 5-21 are superimposed, it is clear that although the second helix can adopt a number of different orientations, the long axes of helix 2 all lie in approximately the same plane (Figure 4c,d) and similarly for helix 1 when residues 24-37 are superimposed (Figure 4e,f).

To see whether there was a most likely conformation for the hinge region connecting the two helices, we carried out a search of the crystallographic protein data bank using the procedure described by Jones and Thirup (1986). For each structure, we were able to find several eight-residue segments in the protein data bank within  $1 \text{ \AA}$  of the backbone atom positions of residues 20-27. We therefore conclude that all the calculated conformations are equally probable.

**Structure-Function Relationship.** The amphipathic  $\alpha$ -helix has been shown to be an important structural feature of a



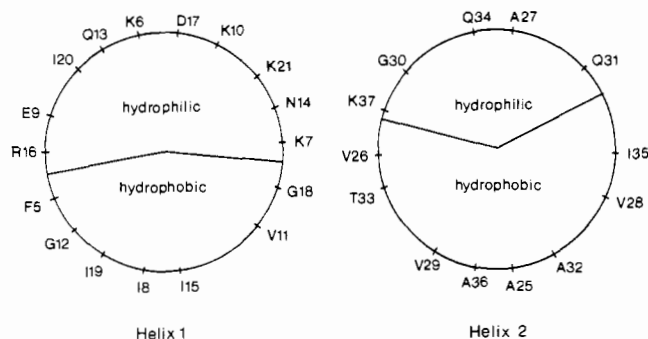


FIGURE 7: Helical wheel projection (Schiffer & Edmundson, 1967) of helices 1 (residues 5–21) and 2 (residues 25–37).

number of lytic peptides, such as apolipoprotein A1 (Fukushima et al., 1979), melittin (DeGrado et al., 1981), and *d*-hemolysin (Lee et al., 1987). Cecropin is no exception in this respect. Examination of the side-chain distribution along the two  $\alpha$ -helices (Figure 4g,h) indicates that they are indeed amphipathic in nature, confirming at part previous predictions (Merrifield et al., 1982; Steiner, 1982). Helix 1 (residues 5–21) is an almost perfect example of an amphipathic  $\alpha$ -helix with a clear and approximately equal separation of the hydrophilic (nine polar residues interspersed by a single nonpolar residue) and hydrophobic (seven nonpolar residues) faces. In the case of helix 2 (residues 25–37), however, the hydrophilic face constitutes only a third of the surface and comprises three polar (Gln-31, Gln-34, Lys-37) as well as two nonpolar (Ala-27, Gly-30) residues. These features are easily appreciated from the helical wheel projections shown in Figure 7.

The present structure in conjunction with the results from synthetic studies suggests a possible model for the interaction of cecropin A with bacterial membranes. It seems likely that the primary anchor site for cecropin A on the membrane is provided by an aromatic residue (Trp or Phe) at position 2, which serves to bring the rest of the peptide into sufficiently close contact with the membrane for helix formation to occur. When this aromatic residue is either removed (Andreu et al., 1983) or replaced by another residue (Andreu et al., 1985), this step can no longer take place and antibacterial activity is severely reduced (by >90%). We further suggest that the mainly hydrophobic portion of helix 2 (residues 25–33) serves to further strengthen the interaction of cecropin A with the membrane, thereby stabilizing the formation of helix 1. This is supported by the observation that although the 1–21 fragment of cecropin A still displays antibacterial activity, it is severely reduced (by ~90%) relative to that of native cecropin A (DeGrado, 1983). The four C-terminal residues, on the other hand, seem to be of little importance as the 1–33 fragment is just as active as intact cecropin A (Merrifield et al., 1982). Finally, we suggest that the main lytic activity of cecropin A resides in the amphipathic helix 1. For some bacteria, such as *Pseudomonas aeruginosa*, *Bacillus megatherium*, and *Micrococcus luteus*, a complete helix 1 seems to be required, as disruption of this helix by the introduction of a helix breaker, such as proline, at position 4 or 8 drastically reduces antibacterial activity (Andreu et al., 1985). In the case of *Escherichia coli*, on the other hand, a reduced helix 1 extending from residue 9 to 21 seems to be sufficient as [Pro<sup>8</sup>]cecropin A is as active as native cecropin A (Andreu et al., 1985).

#### SUPPLEMENTARY MATERIAL AVAILABLE

One table giving the complete list of NOE interproton distance restraints used in the computation of the three-di-

mensional structure of cecropin (6 pages). Ordering information is given on any current masthead page.

**Registry No.** Cecropin A, 80451-04-3.

#### REFERENCES

- Andreu, D., Merrifield, R. B., Steiner, H., & Boman, H. G. (1983) *Proc. Natl. Acad. Sci. U.S.A.* 80, 6475–6479.
- Andreu, D., Merrifield, R. B., Steiner, H., & Boman, H. G. (1985) *Biochemistry* 24, 1683–1688.
- Bax, A., & Davis, D. G. (1985) *J. Magn. Reson.* 65, 355–366.
- Bax, A., Sklenar, V., Clore, G. M., & Gronenborn, A. M. (1987) *J. Am. Chem. Soc.* 109, 6511–6513.
- Billeter, M., Braun, W., & Wüthrich, K. (1982) *J. Mol. Biol.* 153, 321–345.
- Bodenhausen, G., Vold, R. L., & Vold, R. R. (1980) *J. Magn. Reson.* 37, 93–106.
- Boman, H. G., & Hultmark, D. (1981) *Trends Biochem. Sci. (Pers. Ed.)* 6, 306–309.
- Boman, H. G., & Steiner, H. G. (1981) *Curr. Top. Microbiol. Immunol.* 94/95, 75–91.
- Boman, H. G., & Hultmark, D. (1987) *Annu. Rev. Microbiol.* 41, 103–126.
- Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson.* 53, 521–558.
- Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem.* 4, 187–217.
- Brünger, A. T., Clore, G. M., Gronenborn, A. M., & Karplus, M. (1986) *Proc. Natl. Acad. Sci. U.S.A.* 83, 3801–3805.
- Brünger, A. T., Kuryan, J., & Karplus, M. (1987a) *Science (Washington, D.C.)* 235, 458–460.
- Brünger, A. T., Clore, G. M., Gronenborn, A. M., & Karplus, M. (1987b) *Protein Eng.* 1, 399–406.
- Clore, G. M., & Gronenborn, A. M. (1987) *Protein Eng.* 1, 275–288.
- Clore, G. M., Gronenborn, A. M., Brünger, A. T., & Karplus, M. (1985) *J. Mol. Biol.* 185, 435–455.
- Clore, G. M., Gronenborn, A. M., Carlson, G., & Mayer, E. F. (1986a) *J. Mol. Biol.* 190, 259–267.
- Clore, G. M., Nilges, M., Sukumaran, D. K., Brünger, A. T., Karplus, M., & Gronenborn, A. M. (1986b) *EMBO J.* 5, 2729–2735.
- Clore, G. M., Brünger, A. T., Karplus, M., & Gronenborn, A. M. (1986c) *J. Mol. Biol.* 191, 523–551.
- Clore, G. M., Sukumaran, D. K., Nilges, M., & Gronenborn, A. M. (1987a) *Biochemistry* 26, 1732–1745.
- Clore, G. M., Gronenborn, A. M., Nilges, M., Sukumaran, D. K., & Zarbock, J. (1987b) *EMBO J.* 6, 1833–1842.
- Clore, G. M., Gronenborn, A. M., Nilges, M., & Ryan, C. A. (1987c) *Biochemistry* 26, 8012–8023.
- Davis, D. G., & Bax, A. (1985) *J. Am. Chem. Soc.* 107, 2821–2822.
- DeGrado, W. F. (1983) in *Peptides: Structure and Function* (Hruby, V., & Rich, D. H., Eds.) pp 195–198, Pierce Chemical Co., Rockford, IL.
- DeGrado, W. F., Kezdy, F. J., & Kaiser, E. T. (1981) *J. Am. Chem. Soc.* 103, 679–681.
- Edman, P., & Henschen, A. (1975) in *Protein Sequence Analysis* (Needleman, S. B., Ed.) 2nd ed., pp 232–279, Springer-Verlag, Berlin.
- Ernst, R. R., Bodenhausen, G., & Wokaun, A. (1987) *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Clarendon Press, Oxford.
- Faye, I., Pye, A., Rasmuson, T., Boman, H. G., & Boman, I. A. (1975) *Infect. Immun.* 12, 1426–1438.

- Fukushima, D., Kupferberg, J. P., Yokoyama, S., Kroon, D. J., Kaiser, E. T., & Kezdy, F. J. (1979) *J. Am. Chem. Soc.* 101, 3703-3704.
- Hultmark, D., Steiner, H., Rasmuson, T., & Boman, H. G. (1980) *Eur. J. Biochem.* 106, 7-16.
- Jeener, J., Meier, B. H., Bachman, P., & Ernst, R. R. (1979) *J. Chem. Phys.* 71, 4546-4553.
- Jones, T. A., & Thirup, S. (1986) *EMBO J.* 5, 819-822.
- Kaptein, R., Zuiderweg, E. R. P., Scheek, R. M., Boelens, R., & van Gunsteren, W. F. (1985) *J. Mol. Biol.* 182, 179-182.
- Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983) *Science (Washington, D.C.)* 220, 671-680.
- Lee, K. H., Filton, J. E., & Wüthrich, K. (1987) *Biochim. Biophys. Acta* 911, 144-153.
- Macura, S., Huang, Y., Suter, D., & Ernst, R. R. (1981) *J. Magn. Reson.* 43, 259-281.
- Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.* 113, 967-974.
- Merrifield, R. B. (1963) *J. Am. Chem. Soc.* 85, 2149-2154.
- Merrifield, R. B., Vizioli, L. D., & Boman, H. G. (1982) *Biochemistry* 21, 5020-5031.
- Metropolis, N., Rosenbluth, M., Rosenbluth, A., Teller, A., & Teller, E. (1983) *J. Chem. Phys.* 21, 1087-1092.
- Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988a) *FEBS Lett.* 229, 317-324.
- Nilges, M., Gronenborn, A. M., Brünger, A. T., & Clore, G. M. (1988b) *Protein Eng.* 2, 27-38.
- Plateau, P., & Gueron, M. (1982) *J. Am. Chem. Soc.* 104, 7310-7311.
- Redfield, A. G., & Kuntz, S. D. (1975) *J. Magn. Reson.* 19, 250-254.
- Richardson, J. S. (1981) *Adv. Protein Chem.* 34, 167-339.
- Schiffer, M., & Edmundson, A. B. (1967) *Biophys. J.* 7, 121-135.
- Steiner, H. (1982) *FEBS Lett.* 137, 283-287.
- Steiner, H., Hultmark, D., Engström, Å., Bennich, H., Kapur, R., & Boman, H. G. (1981) *Nature (London)* 292, 246-248.
- Sunsqvist, B., & Macfarlane, R. D. (1985) *Mass Spectrom. Rev.* 4, 421-460.
- van Hofsten, P., Faye, I., Kockrum, K., Lee, J.-Y., Xanthopoulos, K. G., Boman, I. A., Boman, H. G., Engström, Å., Andreu, D., & Merrifield, R. B. (1985) *Proc. Natl. Acad. Sci. U.S.A.* 82, 2240-2243.
- Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Go, N., & Wüthrich, K. (1987) *J. Mol. Biol.* 196, 611-640.
- Williamson, M. P., Havel, T. F., & Wüthrich, K. (1985) *J. Mol. Biol.* 182, 295-315.
- Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
- Wüthrich, K., Wider, G., Wagner, G., & Braun, W. (1982) *J. Mol. Biol.* 155, 311-319.
- Wüthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol.* 169, 949-961.
- Wüthrich, K., Billeter, M., & Braun, W. (1984) *J. Mol. Biol.* 180, 715-740.
- Zuiderweg, E. R. P., Boelens, B., & Kaptein, R. (1985) *Biopolymers* 24, 601-611.

## *Pseudomonas aeruginosa* Exotoxin A: Effects of Mutating Tyrosine-470 and Tyrosine-481 to Phenylalanine<sup>†</sup>

Maja Lukac<sup>‡</sup> and R. John Collier\*

Department of Microbiology and Molecular Genetics, Harvard Medical School, and Shipley Institute of Medicine, 200 Longwood Avenue, Boston, Massachusetts 02115

Received April 8, 1988; Revised Manuscript Received June 21, 1988

**ABSTRACT:** Directed mutagenesis was used to probe the functions of Tyr-470 and Tyr-481 of *Pseudomonas aeruginosa* exotoxin A (ETA) with respect to cytotoxicity, ADP-ribosylation of elongation factor 2 (EF-2), and NAD-glycohydrolase activity. Both of these residues lie in the active site cleft, close to Glu-553, a residue believed to play a direct role in catalysis of ADP-ribosylation of EF-2. Substitution of Tyr-470 with Phe caused no change in any of these activities, thus eliminating the possibility that the phenolic hydroxyl group of Tyr-470 might be directly involved in catalysis. Mutation of Tyr-481 to Phe caused an approximately 10-fold reduction in NAD:EF-2 ADP-ribosyltransferase activity and cytotoxicity but no change in NAD-glycohydrolase activity. The latter mutation did not alter the  $K_M$  of NAD in the NAD-glycohydrolase reaction, which suggests that the phenolic hydroxyl of Tyr-481 does not participate in NAD binding. We hypothesize that the phenolic hydroxyl of Tyr-481 may be involved in the interaction of the toxin with substrate EF-2.

**E**xotoxin A (ETA), the most toxic protein produced by *Pseudomonas aeruginosa*, acts by an enzymic mechanism similar to that of diphtheria toxin (DT). Domain III, the

carboxyl-terminal folding domain defined in the crystallographic structure of Allured et al. (1986, 1987), shows sequence homology with the catalytic fragment of DT (fragment A) (Carroll & Collier, 1988; Brandhuber et al., 1988; Zhao & London, 1988), and like fragment A, domain III catalyzes transfer of the ADP-ribosyl moiety of NAD to elongation factor 2 (EF-2). Thereby, EF-2 is inactivated and protein synthesis of sensitive eukaryotic cells is inhibited.

Glu-553 of ETA and Glu-148 of DT were identified by photoaffinity labeling as functionally homologous active site residues (Carroll & Collier, 1987). Results of site-directed

<sup>†</sup>Supported by NIH Grants AI-22021 and AI-22848 from the National Institute of Allergy and Infectious Diseases. Partial support was also received from the Shipley Institute of Medicine. M.L. was supported in part by funds from the Swiss National Science Foundation.

\*Address correspondence to this author at the Department of Microbiology and Molecular Genetics, Harvard Medical School.

<sup>‡</sup>Present address: Swiss Federal Institute for Water Resources and Water Pollution Control, CH-6047 Kastanienbaum, Switzerland.