Conformational Isomerism of Endothelin in Acidic Aqueous Media: A Quantitative NOESY Analysis[†,‡]

Niels H. Andersen,[*,§] Chinpan Chen,[§] Thomas M. Marschner,[§] Stanley R. Krystek, Jr.,[∥] and Donna A. Bassolino[∥]

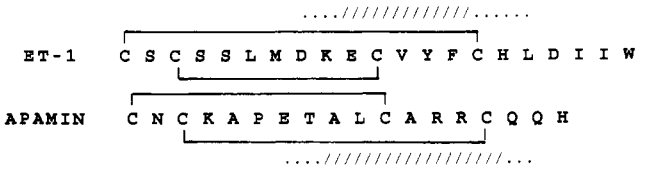
Department of Chemistry, University of Washington, Seattle, Washington 98195, and Department of Macromolecular Modeling, Bristol-Myers Squibb Research Institute, P.O. Box 4000, Princeton, New Jersey 08543

Received May 20, 1991; Revised Manuscript Received September 12, 1991

ABSTRACT: The conformational features of endothelin-1 (ET-1) in mixed water/ethylene glycol media have been studied by two-dimensional ^{1}H NMR experiments throughout the pH range 3.2–7.2. At pH <5 all backbone NH signals can be observed, and NOESY experiments provided a large set of dipolar cross-peaks. Cross-peak intensities from each experiment (different mixing times and H_2O versus D_2O) were converted to distance constraints using a novel algorithm (program DISCON) for removing spin diffusion effects and thus obtain cross-rates rather than cross-peak intensities. A set of 168 nonstereospecific distance bounds (average experimental precision, ±0.38 Å) was used in dynamics simulated annealing refinements. Two consensus structural features were found—a reverse turn at Ser^5 → Asp^8 and an α-helical stretch from Lys^9 to Cys^{15}; however, after constraint-free minimization, structures generated using XPLOR-1.5, CONGEN, and DISCOVER all violated at least 32% of the bounds by more than 0.2 Å, which we ascribe to conformational isomerism. When the constraints were modified to reflect subsequent experimental data and to eliminate constraints that could not be obeyed by any single conformer structure, the relaxed structures still violated at least 15% of this more limited and looser set of constraints. Therefore, a modified procedure for constrained dynamics refinement (using XPLOR-2.1), which allows for conformational isomerism outside of the central helical core region, was developed. This "conformer search procedure" produced structures which fell into five tightly defined conformational clusters. The two most populated clusters correspond to a rotation of the 8,9-amide unit. The conformer which we propose as the major contributor at pH 3.2–5.8 was defined to a backbone rmsd of 0.51 Å over residues 1 → 15. An alternative description of the motional averaging in segments of the endothelin structure as extensive randomization rather than rapid interconversion between a small number of discreet conformers was ruled out by an analysis of NH shift–temperature gradients and exchange rates. This analysis suggests that small $\Delta\delta/\Delta T$ values need not correlate with H-bonding for conformational mixtures. In ET-1 the greatest motional averaging occurs from Ser^2 through Ser^5 (not in the C-terminus) and may be so extensive as to approximate a flexible random coil population as high as 30%. The C-terminus shows less rapid and less extensive conformational averaging, but no definitive structures for individual conformers could be derived in the absence of stereospecific constraints. The pharmacological implications of the consensus structural features are discussed.

In this article we report a NOESY-based conformational study of endothelin-1 (ET-1),[1] originally isolated from porcine aortic endothelial cells (Yanagisawa et al., 1988) but now known to be a major endogenous vasoconstrictor in humans (Yanagisawa & Masaki, 1989) and a member of a new class of peptidic agents that includes additional mammalian ETs and the sarafotoxin snake venom peptides (Kloog et al., 1988). Increasing evidence for the involvement of endothelins in human disease has prompted a major effort in pharmacological evaluation, structure elucidation, and drug design based on the endothelins.

The endothelins bear some resemblance to bee venom apamin (Freeman et al., 1986; Pease & Wemmer, 1988):



Perkins et al. (1990) used this presumed analogy and the obeservation of some degree of helicity by circular dichroism (CD) to create a structural model of ET-1. The regions of apamin and ET-1 that are considered helical are shown in the illustration with some fraying indicated at each end of a central structured core.

To date, no crystal structure has been reported for any member of the endothelin family of peptides. Even if one were available, there is little basis for assuming that either the solution structure or the receptor-associated conformation would bear a close resemblance to that observed in the crystal. In distinct contrast to the situation for a protein, small peptides

[†] The research at the University of Washington was supported initially by an NSF grant and then by ONR Grant N00014-88-K-0202; NMR instrumentation upgrades were partially funded by a grant from E. R. Squibb and Sons.

[‡] The coordinates of residues 1–17 of the major conformer have been deposited in the Brookhaven Protein Data Bank (code number 1EDP).

[*] Address correspondence to this author.

[§] University of Washington.

[∥] Bristol-Myers Squibb Research Institute.

[1] Abbreviations: CD, circular dichroism; DG, distance geometry; DMSO, dimethyl sulfoxide; MM and MD, molecular mechanics and dynamics; ET, endothelin; FR, SA structures that have been relaxed (without NOE distance constraints) into the nearest conformational minimum; HPLC, high-performance liquid chromatography (reverse-phase, hydrophobic supports); SA, dynamics simulated annealing procedure or structure generated from it; TFA, trifluoroacetic acid, TFE, trifluoroethanol. Glycol in all cases refers to ethylene glycol. The common abbreviations of 2D NMR are employed: NOESY, COSY, TPPI, TMS, FID, NOE, d_{NN}, $d_{\alpha N}$, etc. without further comment. Specific NOE interactions and the corresponding derived distance constraints are identified by amino acid (residue number or one-letter symbol and residue number) and hydrogen position—HN (or N), α, β, etc.—for each proton involved; e.g., 9α/12HN represents a cross-peak or constraints for the α-proton of residue 9 and the backbone NH of residue 12. All other references to individual residues are by symbol *and* residue number—L17 corresponds to the leucine at position 17—or use standard format—Leu^{17}.

0006-2960/92/0431-1280$03.00/0

can be expected to display numerous interactions between molecules in the lattice; these and packing forces could be dominant determinants of conformer stability. Molecular modeling suggests that the disulfide constraints in endothelins (1–15, 3–11) are insufficient to restrict the conformational possibilities to a single family of structures. A myriad of quite dissimilar conformers which do not differ greatly in energy, based on standard potential force fields such as that in DISCOVER (Biosym Technologies, Inc.) or CHARMm (Brooks et al., 1983), can be generated. The conformational versatility suggested by such dynamics simulations can be taken as a stern warning that segmental motion and conformational equilibria cannot be ruled out for ET-1 and that these possibilities must be considered in a NOESY-based conformation elucidation.

Protein structure elucidation by NMR has become a well-accepted method, and increasingly better resolution is reported in these studies. In most studies, NOE intensities are converted into rather coarse distance constraints with wide bounds, typically (Clore et al., 1990) large, 1.8–2.7 Å; medium, 2.0–3.3 Å; small, 2.4–5.0 Å; and very small, 2.8–6.0 Å. (In a medium-sized peptide, the reach of the NOE is shorter, and the upper bounds for the last two categories are reduced.) The success of coarse distance constraints[2] in protein structure refinement very likely arises from the occurrence of stretches of regular secondary structure motifs and long-range NOEs reflecting sequence remote contacts associated with tertiary structure (Borgias & James, 1988). Smaller cyclic peptides are not blessed with either of these and are also more likely to display backbone conformations from "high-energy regions" of a Ramachandran plot.

The remarkable increase in precision in some recent protein solution structure elucidations (Moore et al., 1988; Haruyama & Wüthrich, 1989; Kraulis et al., 1989; Clore et al., 1990) arises not from the NOE distance constraints but rather from the use of hydrogen-bonding constraints and the use of scalar coupling information in conjunction with NOEs to derive local sets of $\phi/\psi/\chi$ constraints and βCH_2 *pro-R/S* assignments, using the program STEREOSEARCH (Nilges et al., 1990) or HABAS (Güntert et al., 1989). The specter of conformational equilibria in a peptide structure argues against the use of all scalar coupling data except for maximal values that implicate antiperiplanar proton pairs.

Three notably dissimilar NMR structures have been proposed for ET-1 in DMSO solution (Saudek et al., 1989, 1991; Endo et al., 1989; Munro et al., 1991); none of them display a regular helical region, and with the exception of the structures of Endo et al., the degree of convergence within the bicyclic ring was poor. Munro et al. suggested a somewhat helical structure from Leu[6] through Cys[11]; the conformation from Cys[11] through Cys[15] was also described as well-defined, but apparently nonhelical. The Saudek structures typically include one or two helical turns, including the region from Cys[11] through Cys[15]. We have also examined ET-1 in DMSO (unpublished studies) and reached the conclusion, since noted by Brown et al. (1990), that motional averaging in this denaturing solvent is so extensive as to preclude "structure" definition by NOESY-based methods.

Our initial work (Krystek et al., 1991) and other preliminary accounts [Dalgarno et al., 1990; Tamaoki et al., 1991; Donlan et al., 1991; Reily & Dunbar, 1991; for [Nle[7]]ET-1, see Aumelas et al. (1991)] of NMR studies of endothelins in predominantly aqueous media have appeared recently, largely since this paper was first submitted. Without exception, they present evidence for a helical region (variously reported as residues 9–15 → 9–17), but the remaining structure is ill-defined; however, a reverse turn at Ser[5]–Asp[8] was noted in two cases (Krystek et al., 1991; Tamaoki et al., 1991). The structure within the C-terminus and its orientation to the bicyclic core were completely indeterminant in all these studies.[3]

In the present study we present further details from our initial study (Krystek et al., 1991) and evidence that indicates that ET-1 displays conformational averaging in some portions of the bicyclic core. By the judicious use of tight distance constraints, derived by the DISCON algorithm (Andersen et al., 1990b) which largely removes spin diffusion effects, we have been able to search for major conformers of ET-1 and define the loci of motional averaging and the nature of the averaging. A novel analysis of NH shift–temperature gradients provides a confirmation of the results from the NOESY-based conformers search.

Experimental and Computational Methods

Sample Preparation. Synthetic ET-1 was obtained from Peninsula Labs (Belmont, CA) and was found by HPLC to be consistently pure (Kumajaya et al., 1988); the commercial material (70–80% peptide by weight) was used as obtained for NMR and CD spectroscopy. Weighed samples (2.0–4.0 mg) were placed in small vials, and 400 μL of aqueous glycol (50% or 60% glycol-d_6 by volume) was added. The sample was then transferred to a 5-mm NMR tube and treated with either 12 μL of 79 mM aqueous CF_3CO_2H (TFA) or 8 μL of 800 mM KH_2PO_4/K_2HPO_4 buffer (pH ~7.6). Further pH adjustments of buffered solution were made with aqueous KOH or TFA to uncorrected pH meter readings as reported. Peptide concentrations are estimated as 1.2–2.8 mM, a range over which no significant concentration dependence was observed. Samples in D_2O media were obtained by repeated addition of 99.99+% D_2O and lyophilization. The media composition (percent glycol) of reconstituted samples was estimated (±10%) from the observed selective relaxation rates or geminal pair cross-relaxation rates on the basis of the correlation time dependence on media viscosity (Andersen et al., 1984; Lane et al., 1986). Very minor changes in chemical shift values were observed over the full range of glycol concentrations used in this study (40–75 vol % glycol).

CD Spectroscopy. All spectra were recorded on a JASCO J-720 spectropolarimeter at ambient temperature (~23 °C) over the ranges indicated using short path-length cells—178–205 nm (0.05 mm), 185–250 nm (0.1 or 0.5 mm), and 200–270 nm (2.0 mm). Each spectrum was the average of 16 scans smoothed using the reverse Fourier transform noise reduction software from JASCO. The CD spectra appearing in the supplemental figure were obtained by diluting equal aliquots of an NMR sample (1.8 mM, 55% aqueous ethylene

[2] In the present study we explore the suitability of tight distance bounds in conformation elucidation. More narrowly defined experimental distances have been used in a few other reported studies of peptides. These distances came either from an eigenvalue solution (Olejniczak et al., 1986; Bremer et al., 1984; Boelens et al., 1988, 1989) for the relaxation rates (Fesik et al., 1986; Brown et al., 1989) or from iterative bound adjustment using BKCALC (Summers et al., 1990) to simulate NOESY spectral intensities. The program DISCON (Andersen et al., 1990b), which incorporates a novel algorithm for extracting experimental distances and their precision from each NOESY signal matrix (vide infra), was used for the present study.

[3] However, Saudek and Pelton favor, even in aqueous media, more compact structures for endothelins in which the C-terminus bends back on the core making numerous contacts; see Bortmann et al. (1991) and Pelton (1991).

glycol, pH ~5.8) with set volumes of buffered water or water/alcohol mixtures to final concentrations corresponding to 13.2–1360 μM (based on the expectation value of the difference in ϵ_{242} due to tyrosine that is produced upon addition of aqueous KOH).

NMR Spectroscopy. All NMR spectra were recorded at 500.135 MHz using Bruker instruments. All spectra in protic media has a spectral width of 6024 Hz; in D_2O, the spectral width was 4505 Hz. In media with greater than 0.5% HDO content, water suppression was obtained by saturation during the entire preparatory delay (2 s, 1D experiments; 1.05–1.4 s, 2D experiments) and during the mixing time for NOESY experiments. The temperature gradients for resonances were determined primarily from 1D spectra (8K points, resolution enhanced) recorded at 5-deg intervals from 290 to 315 K. For a few severely overlapped signals the gradients were confirmed by NOESY and/or COSY spectra at 295 and 310 K. For resolved NH resonances these 1D spectra also provided $J_{N\alpha}$ estimates (±1 Hz) and exchange rate estimates. Amide–NH exchange rates were obtained as $t_{1/2}$ values from serial spectra recorded at 5 min to 36 h after ET-1 was dissolved in D_2O media (75% glycol, pH = 3.2) and were confirmed by NH reappearance rates upon addition of H_2O (to a final concentration corresponding to 58% glycol) to the fully exchanged sample. All data, including 2D serial files, were accumulated on Aspect 3000 computers and processed on a Silicon Graphics 4D/20 workstation.

NOESY spectra were generated by the phase-sensitive, States–Haberkorn method (States et al., 1982), consisting of two serial files (284–400 t_1 values) of 2K complex FIDs, collected and processed as previously described (Andersen et al., 1987) with the following exceptions: all samples were spinning, the first point in t_1 was multiplied by 0.6 (Otting et al., 1986), and a skewed 60° shifted sine bell (qsb 60° 0.7) apodization was applied in both dimensions. The latter has been shown (Andersen et al., 1989) to produce no measurable distortion in cross-/auto-peak intensity ratios. After zero-filling all NOESY spectra were 1K × 1K data matrices. The contour plots were examined without base-plane correction.[4] Two protic media NOESY spectra were used for computer-aided constraint extraction (percent glycol, pH, T, τ_m, scans per t_1 per file): experiment 1, 60, 3.2, 309 K, 100 ms, 192; experiment 2, 60, 3.2, 295 K, 160 ms, 64. A later experiment (55, 5.8, 310 K, 200 ms, 192) was used only for qualitative constraints for connectivities involving 16HN and 17HN which were resolved under these conditions. The experiments in D_2O media were experiment 3, ~40, 7.0, 295 K, 200 ms, 64; and experiment 4, ~75, 3.5, 295 K, 120 ms, 128.

Phase-sensitive TPPI-COSY spectra (Marion & Wüthrich, 1983; Drobny et al., 1979) were collected for each pH and temperature in Table I, typically 800–1024 t_1 values, 4K complex points in t_2, and 32 scans/t_1. A high-resolution TPPI-COSY, 8K complex points in t_2, 4K × 4K after zero-filling, was collected at 305 K in 50% aqueous glycol (pH ~7). The latter provided $J_{N\alpha}$ and $J_{\alpha\beta}$ estimates for signals from crowded spectral regions.

Distance Constraint Extraction. For each spectrum the peak contour levels of all signals were recorded (or estimated, for overlapping peaks). A selection of cross-peaks and auto-peaks (including at least two NH, two αH, one member of a prochiral βCH_2, a Me peak, and an aromatic H in each case) were also quantitated as volume integrals. Fortunately, essentially all specific protons displayed at least one cross-peak which was of reasonable size and fully resolved. Peak contour levels were converted to volume estimates assuming that all peaks for a specific proton resonance position (on ω_2) have the same height/volume ratio.

The data from experiments 2 and 4 were used as such. The largest single data set (642 observations) was constructed on the basis of experiment 1 supplemented with scaled intensities from experiments 2 and 3. The intensities from experiment 2 were weak cross-peaks that did not appear at τ_m = 100 ms (309 K) or were better resolved at the lower temperature. Those from experiment 3 were for regions partially masked by the water suppression used in experiment 1. Estimated cross-relaxation rates were extracted from each data set using the program DISCON.

The DISCON *algorithm* is based on eq 1, which is fully supported by experiment simulations (Andersen et al., 1989; Eaton & Andersen, 1987). The left half of the equation

$$\frac{1}{2\tau_m}\left(\frac{S_{ij}}{S_{ii}}+\frac{S_{ji}}{S_{jj}}\right)=\sigma_{ij}^{eff}\cong\sigma_{ij}+\frac{\tau_m}{2}\sum_{k\neq i,j}\sigma_{ik}\sigma_{kj} \qquad (1)$$

defines an effective cross-rate as a cross-diagonal average of driver-normalized NOEs. It is assumed that this is well approximated as a phenomenological summation of direct and indirect transfer rates. The relationship between eq 1 and a two-term Taylor expansion of $\mathbf{A} = \exp(\mathbf{R}\tau_m)$ has been detailed (Andersen et al., 1990a).

The DISCON algorithm is a reorganization of eq 1 to produce an iterative solution for the σ-value of each dipolar-coupled spin pair:

$$\text{first estimate:}\quad \sigma_{ij}^{(1)}=\sigma_{ij}^{eff}-\frac{\tau_m}{2}\sum_{k\neq i,j}\sigma_{ik}^{(eff)}\sigma_{kj}^{(eff)}$$
$$\sigma_{ij}^{(a+1)}=\sigma_{ij}^{eff}-\frac{\tau_m}{2}\sum_{k\neq i,j}\sigma_{ik}^{(a)}\sigma_{kj}^{(a)},\ \ldots,\Rightarrow\sigma_{ij}^{est} \qquad (2)$$

In order to apply this algorithm systematically, missing points in the 2D data matrix are filled with surrogate values. We assume that a measured S_{ij} value is a good approximation of S_{ji} if the latter is unmeasurable due to streaks or other spectral artifacts. The remaining surrogate entries are class averages or zeros. The auto-peak classes are backbone NH, α-methines, methylene protons, β- and γ-methines, methyls, aromatic methines, and all others. Four classes of cross-peaks are recognized: $d_{\alpha N}(i,i+1)$, intraresidue vicinal, all geminals, and all other intraresidue interactions. The remaining unquantified cross-peaks become surrogate zeros.

The program (Andersen et al., 1990b) carries out the iteration in eq 2 and reports σ_{ij}^{est} for any pair with an observed S_{ij} value and a maximum σ_{ij} for any observed zero or instance of $\sigma_{ij}^{est} < 0.2\ \sigma_{ij}^{eff}$. These are then converted into distance constraints: $d_{ij} = k(\sigma_{ij}^{est})^{-1/6}$. The best (and extreme values) of k are determined as those which produce d_{ij} values that do not violate the ranges of values for intraresidue and interresidue connectivities $(i,i+1)$ $d_{\alpha N}$ and d_{NN} which are defined on the basis of holonomic constraints. Each data set thus provides two types of constraints: those with both low and high bounds (type a) and those with only a low bound (type b). The final set of constraints was obtained by combining d_{ij} estimates (and ranges) from independent determinations. The values from experiment 1 were given twice the statistical weight on the basis of the greater signal to noise and the lesser contribution

[4] The final matrices for spectra in protic media were baseline-corrected separately on each side of the residual solvent signal in ω_2 using a third-order polynomial (cubic spline method in FTNMR). Several trial sets of points in ω_2 (each corresponding to baseline values in the control spectra) were examined for each side of the residual solvent signal.

of spin diffusion under the experimental conditions employed. The format written by the DISCON programs and required by XPLOR and PFS (vide infra) is illustrated:

	d	d_-	d_+
10Hα/13HN	3.50	0.22	0.25
6Hα/9HN	3.95	0.45	3.10

The latter illustrates a type b constraint, for which d_+ is >3.0, assigned on the basis of the sequence separation involved, as 3.1, 6.1, or 9.1 Å. In the case of the 6α/9HN connectivity a very weak cross-peak was observed in experiment 2, but the DISCON analysis indicated that σ^{est} fell to below the cuttoff for inclusion. The (d, d_-, d_+) format for constraints is used throughout.

Structure Generation and Refinement Procedures. XPLOR (Polygen Corp.), DISCOVER (Biosym Technologies Inc.), and CONGEN (Bruccoleri & Karplus, 1987) were used for distance-constrained dynamics. In all three protocols, disulfide closure was effected by using distance constraints, with S–S bond introduction at the minimization stage following the first annealing procedure. Hydrogen-bond potentials were excluded at all stages of the XPLOR procedures. Two acyclic structures (α, ϕ/ψ −65/−50; loose coil, ϕ/ψ −80/+90) were used in all protocols. A β-structure (ϕ/ψ −150/+150) was used in CONGEN and DISCOVER. A fully extended linear structure was used for XPLOR and DISCOVER. Two of the published DMSO NMR structures were also used as initial structures for XPLOR refinements. Distances involving prochiral βCH_2 protons, prochiral methyl groups, and aromatic δ protons were $\langle r^{-6} \rangle$ averaged (Brunger et al., 1986; Clore et al., 1986; Bassolino et al., 1988) when possible.[5]

The CONGEN and DISCOVER procedures were described in our initial report (Krystek et al., 1991). In initial runs using version 1.5 of XPLOR, no repulsive van der Waals terms were used, and only the improper torsions had inflated force constants (500 kcal·mol^{-1}·rad^{-2}). The NOE constraints were divided into six groups, named and defined as follows: key, all well-determined $d_{\alpha N}(i+1)$, $d_{\beta N}(i+1)$, and $d_{NN}(i+1)$ connectivities and the shorter $d_{\alpha N}(i)$, $d_{\alpha N}(i+3)$, and $d_{\alpha\beta}(i+3)$ connectivities; short, intraresidue connectivities and less well determined sequential $d_{\alpha N}$, d_{NN}, and $d_{\beta N}$ connectivities (i,$i+n$; $n < 4$); clos, interresidue connectivities across the disulfide linkages; Me, constraints involving methyl groups; long, all other intraresidue constraints; and disu, a set of distance constraints to predispose noncyclic structures for closure and to provide a penalty for disulfide torsion angles deviating widely from ±90°. The same classification was retained in conformer search procedures using XPLOR-2.1.

In XPLOR-1.5 runs, dynamics were carried out at 800° (with a 2000° upper limit for revectorization). During their course, scale factors (initially set at 1.0) were uniformly increased by a factor of 1.15–1.60 every 30 fs until they reached preset maximum values (60 kcal·mol^{-1}·Å^{-2} for all categories except key and short set at 100). Subsequently, at either 0.6- or 1.2-ps intervals the scale factors for the distinct constraint classes were reduced by a randomly varied factor (8–250) and then increased as previously to newly set maxima. The random rescaling procedure greatly reduced the tendency of dynamic courses to lock into nonoptimal local minima. Each run generates two or three structures during the final stages of 7–9-ps dynamics courses with the scale factors at high values. These are cooled and subjected to a 200-step Powell minimization during which the k_{NOE} values are reduced to 20–40 kcal·mol^{-1} Å^{-2}, affording the SA structures. In XPLOR-1.5 runs commencing with noncyclic structures, the clos, disu, and Me scale factors were initially set at 0.05 in order to avoid excessive force contributions that can distort the structures and counteract the effects of the other constraint classes. The latter and the random rescaling procedure were not required in the conformer search protocol which used XPLOR-2.1. Sample protocols and annotated comfiles are included in the supplementary material.

"SA structures" generated with CONGEN, DISCOVER, and version of 1.5 of XPLOR were relaxed in the CHARMm force field of CONGEN. The heavier NOE weighting in DISCOVER and XPLOR-1.5 produces significant distortion of normal bond angles and introduces severe van der Waals repulsions. However, 100 steps of steepest descent minimization without NOE constraints do not alter conformations; the maximum backbone atom rmsd observed upon relaxation was less than 0.25 Å. The increase in NOE constraint violations associated with structure relaxation has been reported (Krystek et al., 1991). The relaxed structures are designated as FR structures.

Structure Evaluation and Comparison Methods. Since some structures were generated in different force fields with E_{NOE} terms of contrasting weight, an independent method of judging fit to NOE constraints was used. Both SA and FR structures were examined. The functions which proved particularly useful have been incorporated in the program PFS (*P*enalty *F*unction *S*cans; Andersen et al., 1990b). The four measures used in the present work are the rmsd, the fraction of bounds violated by ≥0.2 Å, the sum of absolute violations, and an rms weighted violation measure defined as

rmswv = rms weighted violations =

$$\left[\frac{1}{n-1}\sum_k\left(\frac{\text{violation}}{\text{range}}\right)_k^2\right]^{1/2}$$

A violation is defined as any excursion of r_{ij} outside of the constraint range, $d - d_- \rightarrow d + d_+$, and range, in the equation for rmswv, is given by $d_- + d_+$. For constraints with no upper bound (type b), only the violation of the lower bound was used both for the calculation of the rmsd and for all fit measures based on violations (the meaningless range value is replaced by the average range for all type a constraints). Any statistically valid structure would be expected to display an rms deviation less than $\langle \overline{\text{range}} \rangle/\sqrt{2}$. Well-fitting structures display an rms deviation on the order of $\langle \overline{\text{range}} \rangle/2$. Penalty functions that are based on "weighted violations" provide a superior measure of fit since well-determined data (reflected by smaller ranges) cannot be violated without incurring significant penalty accumulation. In practice we find that all well-fitting structures display rmswv values that are smaller than the rmsd value.

Conformer Search Protocol. In our search for potential conformers, XPLOR-2.1 was employed and included inflated bond-stretching and -bending force constants in analogy to the reported procedure of Kraulis et al. (1989). During the initial stages of each dynamics course, van der Waals repulsion was turned off and soft square-well potentials were employed. The repulsive force was ramped on and the NOE penalty was converted to a standard biharmonic well during the last 11 ps of the 1000° dynamics. A representative comfile appears in the supplementary material. All initial runs employed ϕ_2

[5] In the then current versions, DISCOVER did not allow $\langle r^{-6} \rangle$ averaging and CONGEN allowed it only for protons attached to the same carbon atom. The βCH_2 and methyl constraints used in DISCOVER were expanded (Wuthrich et al., 1983) for use as center averages.

$= \phi_3 = \phi_{20} = -125 \pm 30°$ as torsion constraints (with a modest forcing energy). The starting structures were seven relaxed XPLOR-1.5 structures and the initial noncyclic linear and α-helical conformations. For the noncyclic structures parallel runs with (and without) torsion constraints ($\psi_9 \rightarrow \psi_{15}$, ±30°, as listed in column 2 of Table II) were examined. Two constraint sets were used for each structure—the "major conformer set" with and without the "unfolded helices" extensions. The manner in which these new constraint tables were generated is outlined in the Results and Discussion section. Only 30% of the runs included an extensive set of constraints involving residues 18–21. The annotated constraint table for the conformer search procedure is included in the supplementary material. In excess of 120 structures were generated, of which ca. half could be eliminated on the basis of either multiple significant (>0.2 Å) violations or a single unacceptable violation (>0.6 Å) of the key long-range constraints. The remaining structures were subjected to a conformation cluster analysis.

The C-terminal tail was ignored in the cluster analysis. A cluster of bicyclic ring conformations was defined as any set of structures with a complete run of backbone torsions for the bicyclic core with all torsion rmsd values less than 40°. The validity of each cluster was judged by repeating the constrained dynamics course using the cluster mean torsions (±20°) as constraints. We eliminated any cluster for which a torsion-constrained dynamics increases the sum of violations by greater than 10% or introduced any unacceptable violations of the key long-range constraints.

At this point some 40-odd structures, 90% of which fell into five conformational families, remained. NOESY simulations were carried out for at least two structures in each family in order to identify close contacts which are inconsistent with the observed NOESY spectra. A structure (or family of structures) was eliminated if rerefinement, including the full set of unobserved NOEs as constraints with $d_{ij} \geq 3.4$ Å, failed to correct the contact or introduced significant violations of the key long-range constraints. During this final refinement stage parallel runs were performed with and without torsion constraints. In order to allow the NOE constraints to control the dynamics course, the new set of torsion constraints[6] was given a variable scale factor. The penalty for violation of torsion constraints was periodically reduced to 5% or 15% of its maximal value. In numerous instances the torsion angles at ϕ_4, ϕ_6, and ϕ_9 changed sign during this refinement stage. However, new structures always correspond to one of the already recognized conformation clusters. In addition, structures with $\phi_2 = +60°$ (which also predicts a large value of $J_{N\alpha}$) were generated by these refinements. Each acceptable structure was relaxed in the default CHARMm force field of QUANTA-3.0 (Polygen Corp.), excluding the H-bond term. Thus any hydrogen bonds that are observed follow from the NOESY data rather than the minimization procedures.

NOESY Simulation. Expectation spectra for experiment 1 were calculated for each conformer model using the program NOESYSIM (Andersen et al., 1990a). These were scanned for predicted cross-peaks (corresponding to $r_{ij} < 3.4$ Å) which are entirely absent in the experimental data. For prochiral βCH_2 units, all possible locations in the experimental matrix were examined. All trial conformers were relaxed (100 steps, steepest descent) in the default CHARMm force field of QUANTA-3.0 (Polygen Corp.) prior to use in NOESYSIM.

Results and Discussion

ET-1 displays well-dispersed, high-resolution, proton NMR spectra in aqueous CD_3CN and glycol. In both cases a structured state is indicated by the wide range of amide-NH shift–temperature gradients ($\Delta\delta/\Delta T = +1.9 \rightarrow -11.4$ ppb/°C at pH ~3.2 in 60% glycol). Under these conditions the majority of the amide-NH peaks show somewhat slowed exchange ($t_{1/2}$ 20 min → 2 h at 305 K); however, several (residues 1, 2, and 4 in particular) exchange rapidly, $t_{1/2} < 2$ min, and these same resonances are completely absent at pH ≥6 due to cross-relaxation during a 2-s presaturation. We therefore chose the lower pH for our initial studies. The glycol-containing media was selected since the lower viscosity aqueous acetonitrile gave much smaller NOEs and could not be used at low temperatures. CD spectroscopy (supplementary figures) was used to ascertain whether glycol had a major effect on secondary structure preference. Throughout the pH range 3–7 the spectra in 4% and 54% (v/v) glycol media were comparable. The same pH-induced changes in spectra occurred whether or not glycol was present. In contrast, and in conflict with a literature report (Perkins et al., 1990), we find that TFE produces a significant increase in helicity. Given the recent report (Bennes et al., 1990) suggesting that ET-1 aggregates at concentrations higher than 22 μM, we also determined the CD under the NMR conditions throughout the concentration range 13.2 μM–1.4 mM; no changes were observed.

Sequential Resonance Assignments. All resonances of endothelin spectra could be identified by a relatively straightforward application of the COSY/NOESY strategy used for protein structures (Wüthrich, 1986). The use of two temperatures for the NOESYs (see Figure 1) was essential for the resolution of the otherwise extensive set of shift coincidences in the NH region. Other annotated spectra (COSY fingerprint in H_2O, α/β-region COSY in D_2O, d_{NN} connectivities at 295 K) are available upon request. The assignment procedure is outlined here.

ET-1 contains three Ser, two Leu, a Met and Glu, and a host of Cys, Asp, and aryl residues that need to be distinguished. Each aromatic Hδ gave an NOE connectivity to the corresponding Hα and/or Hβ resonances which were thus distinguished unambiguously. The serines were readily sequenced in the NOESY; only S5-HN displays NOEs to two sets of serine α and β resonances. This also identified S4 and, by default, S2. A clear unambiguous string $(i,i+1)$ $d_{\alpha N}$ connectivities than spanned C1 through L6. Another series spanned D18 → W21. With D18 identified, the leucine sequential assignment could be confirmed since 18HN showed an NOE to the uniquely upfield Leu β (its spin system was thus assigned to L17). Unique residues K9 and V12 were linked by a dense web of $(i,i+3)$ connectivities, and the 9α/10N peak (lower panel, Figure 1) served to distinguish E10 and M7. The remaining Cys residues were clear from the downfield NOESY, which displayed an unambiguous string of $(i,i+1)$ d_{NN} connectivities from K9 → H16. The string was broken by the near NN shift coincidence of L17 and H16. This and a few other near shift coincidences (11β, 14β; 11α, 13α; 6α, 10α; and 15α, 18α) result in a few remaining ambiguities in the chart of sequential NOEs (Figure 2); however, the sequential assignment itself is secure. The complete resonance assignments appear in Table I. A qualitative analysis

[6] The set of constraints used at this point was as follows: with ±30° bounds, $\phi_2 = \phi_3 = \phi_{20} = -120°$, $\psi_9 = -10°$; with ±35° bounds, $\phi_7 = -90°$, $\phi_{15} = -95°$; with ±20° bounds, $\phi_{10} = -70°$, $\psi_{10} = -55°$, $\phi_{11} = -55°$, $\psi_{11} = -25°$, $\phi_{12} = -75°$, $\psi_{12} = -40°$, $\phi_{13} = -40°$, $\psi_{13} = -30°$, $\phi_{14} = -75°$, $\psi_{14} = -40°$. The force constant was adjusted so as to produce an ~1.5 kcal penalty for a further 25° deviation from the specified range during the final Powell minimization. In contrast, the NOE distances were more heavily weighted; a typical total E_{NOE} value at this stage was 40 kcal/mol.

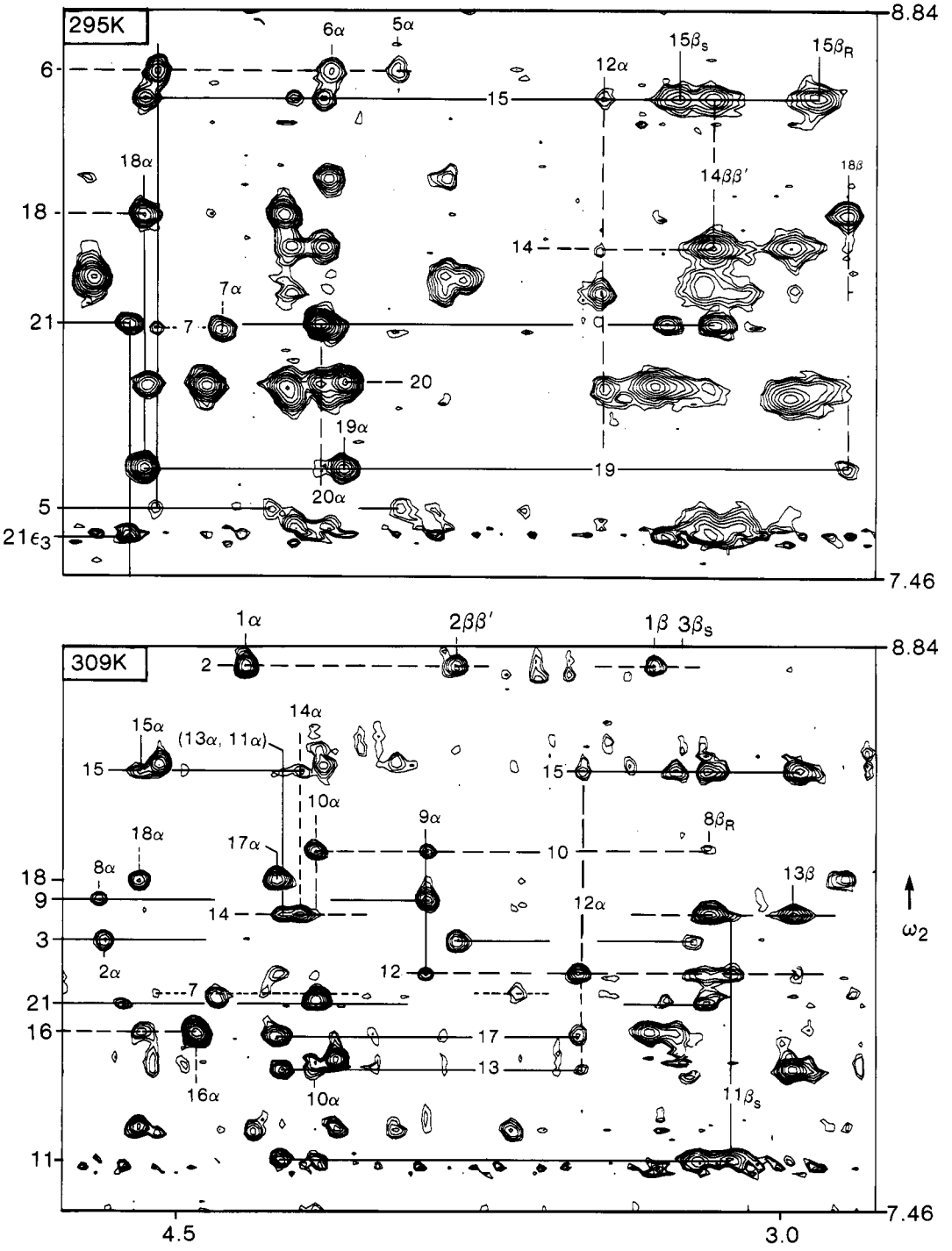


FIGURE 1: Partially annotated NH/α,β region of NOESY spectra of ET-1 at 295 and 309 K. The segments correspond to δ_2 = 7.46–8.84 ppm by δ_1 = 2.77–4.78 ppm in both cases. Throughout, resonance positions for even-numbered residues are shown by dashed lines and odd-numbered residues by solid lines except for M7, which is dotted for clarity. The R/S assignments shown for βCH_2 units are tentative and were not used in the refinement presented in this account; in all cases Cα is given a higher priority than the other Cβ substituents. The downfield portion of the 295 K spectrum showing d_{NN} connectivities appears as a supplementary figure.

of Figure 2 suggests a central helical domain, residues 10–15, which may extend in some "frayed" version to residues 6 and 17.

Stereospecific Assignments. Only 7 of the 18 βCH_2 groups of ET-1 were sufficiently nonequivalent and displayed differential NOEs upon which a *pro-R/S* assignment could be based. Although five of these could be assigned securely on the basis of the standard criteria (Nilges et al., 1990; Güntert et al., 1989; Hyberts et al., 1987), we chose not to include these assignments in our constraint table. Our simulation studies of medium-sized peptide conformation refinement indicate that even a single incorrectly attributed NOE can be disastrous and we thus err on the side of certainty. Also we specifically wished to ascertain whether higher precision of NOE bounds can reduce the need for "risky" stereospecific assignments and torsion constraints in an NOE-based conformation elucidation. We did, however, assign (Zuiderweg et al., 1985; Hyberts et al., 1987) the V12 methyls so as to include the rich web of interresidue connectivities associated with them in the constraint set.

High-Precision Distance Constraints. The combined data from four NOESY experiments were used to extract a set of distance constraints (corrected for secondary NOEs) as described in the Experimental and Computational Methods

Table I: Chemical Shifts[a] for Endothelin at pH 3.2 (and pH ~7)[b]

	NH [$\Delta\delta/\Delta T$][c]	α	β, β'	other
C1	ex	4.33	3.31, 3.31	
		(3.91)	(3.16, 3.03)	
S2	8.79 [−5.7]	4.67	3.82, 3.81	
	(ex)			
C3	8.125 [−4.0]	5.055	3.215, 2.56	
	(8.18)	(~4.95)		
S4	8.945 [−11.4]	4.31	3.95, 3.895	
	(ex)			
S5	7.65 [+1.9]	4.54	3.955, 3.66	
	(ex?)			
L6	8.555 [−10.5]	4.14	1.65, 1.61	γ, 1.68; δ, 0.91, 0.82
	(8.525 [−11.7])			
M7	7.995 [−4.7]	4.40	2.175, 1.91	γ, 2.61, 2.45
	(8.05 [~−4.3])			ϵ, 2.08
D8	7.375 [+1.1]	4.685	3.18, 2.78	
	(7.405 [+3])		(3.09, 2.63)	
K9	8.22 [+1.4]	3.885	1.85, 1.82	γ, 1.53, 1.43; δ, 1.68, 1.64
	(8.18 [0])			ϵ, 2.98, 2.95
E10	8.34 [−7.1]	4.15	2.145	γ, 2.49 (2.32)
	(8.44 [−1.9)			
C11	7.59 [+0.1]	4.25	3.20, 3.125	
	(7.72 [+0.9])	(4.32)		
V12	8.045 [−7.8]	3.50	2.04	γ, 0.975, 0.835
	(8.13 [−10.2])	(3.62)		
Y13	7.805 [−6.1]	4.235	2.97, 2.95	δ, 6.84; ϵ, 6.66
	(7.83 [−7])			
F14	8.19 [−5.9]	4.195	3.20, 3.17	δ, 7.325; ϵ, 7.355
	(8.01 [−6.5])		(3.24, 3.15)	ζ, 7.38
C15	8.54 [−7.2]	4.585	3.26, 2.95	
	(8.35 [−10.6])	4.585	(3.18, 2.99)	
H16	7.895 [−2]	4.45	3.33, 3.30	δ2, 7.25; ϵ1, 8.58
	(7.96 [−1.2])	(4.525)	(3.21, 3.18)	(δ2, 7.00); (ϵ1, 7.95)
L17	7.885 [−2.9]	4.25	1.63, 1.54	γ, 1.52; δ, 0.82, 0.815
	(7.84 [−4.8])			
D18	8.27 [−5.9]	4.59	2.84, 2.72	
	(8.27 [−6])		(2.635, 2.56)	
I19	7.665 [−4.1]	4.105	1.705	γ1, 1.335, 1.005
	(7.655 [−1.9]			γ2, 0.605; δ, 0.78
I20	7.83 [−7.7]	4.155	1.76	γ1, 1.37, 1.08
	(7.75 [−7.0])			γ2, 0.795; δ, 0.79
W21	7.97 [−7.8]	4.63	3.29, 3.18	δ1, 7.17; ϵ1, 9.98
	(7.54 [−6.5])	(4.49)	(3.28, 3.105)	ϵ3, 7.565; ζ3, 7.05
				ζ2, 7.35; η2, 7.11

[a] Chemical shifts are reported to the nearest 0.005 ppm; the reference is internal TMS at 307 ± 2 °C in 60% (v/v) glycol-d_6/H_2O. [b] For CH resonances, chemical shifts at pH 7 are reported only when they differ from the values at pH 3.2 by more than 0.07 ppm. [c] Temperature gradients are in ppb/°C, and the usual negative value indicates an upfield shift with increasing temperature. Complete or partial cross-relaxation during a 2-s presaturation is indicated by ex.

section. Since multiple determinations of d_{ij} values were generally available, we had an independent check of the bounds precision assigned by the computational method. In excess of 220 conformationally dependent constraints with upper and lower bounds were identified together with >40 constraints with only a lower bound. The inclusion, as constraints, of observed zero NOEs as minimum distances cannot be overemphasized. They frequently provide the only distinction between models which would otherwise fit those constraints which have both a lower and an upper bound.[7] We deleted all constraints involving 19γ1, 20γ1, and 21β protons, in part due to our lack of prochirality assignments but also due to literature-based expectation that the C-terminus might be unstructured. With the deletion of constraints based on possibly ambiguous assignments and the conservative conversion of those involving βCH_2 units to wild-card entries, the data set reduced to 137 type a entries (average range 0.76 Å) and 31 type b entries. A complete listing of the initial distance bounds is available in the supplementary material.

Conformation Elucidation Based on Distance Constraints. We used two commercial dynamics packages (DISCOVER and XPLOR) and CONGEN with six highly divergent starting structures in our studies. In the XPLOR runs,[5] the full set of constraints could be used with $\langle r^{-6} \rangle$ averaging for methyls, prochiral methylenes, and symmetry-related aromatic ring hydrogens. No hydrogen-bonding constraints were used, and the H-bond potentials were excluded in all XPLOR runs and CHARMm minimizations.

The CONGEN and DISCOVER protocols were used to generate respectively 9 and 18 SA structures; the XPLOR protocol (version 1.5, Polygen Corp.) was examined more exhaustively. We found that the incidence of runs producing low-violation structures was significantly increased when the constraint set was divided into six categories whose force constants could be randomly rescaled (and then reincremented) during dynamics annealing. Apparently this procedure reduces the severity of the local minimum problem (Metzler et al., 1989, and discussion and references therein). It was used with all starting structures, and 68 SA structures were generated in this manner.

[7] The work of Summers et al. (1990) on a Zn-finger domain is an excellent demonstration of this. In that study, models obtained through distance-geometry methods routinely predicted strong NOEs which were not observed. These were identified and added as constraints with no upper bound. Eventually these new bounds outnumbered the original constraints by 2 to 1, and a very high resolution structure resulted.

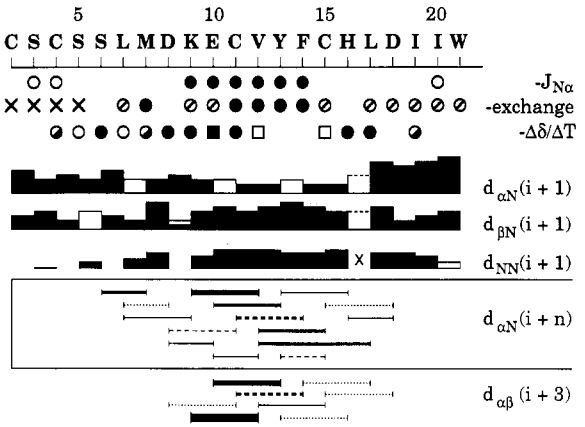


FIGURE 2: Summary of sequential connectivities and amide-NH characteristics for endothelin-1 in acidic aqueous glycol. Symbols: for $J_{N\alpha}$, (●) <6 Hz, (○) ≥9 Hz; for backbone NH exchange rates, (●) $t_{1/2}$ >30 min, (⊘) $t_{1/2}$ 7–25 min, (×) $t_{1/2}$ <3 min; for NH temperature gradients ($\Delta\delta/\Delta T$, ppm/°C), (●) >−3.1 at pH ≤3.5, (○) −4 → −4.7 at pH ≤3.5, (■) −7 → −8 but becomes >−2.5 at pH 7–8, (□) −7 → −8 but becomes <−10 at pH 7–8, (○) <−10 at pH ≤3.5. Some exchange rates could not be measured due to extensive peak overlap. For the (i,i+1) connectivities, bar thickness reflects NOE size, open bars indicate a potentially ambiguous assignment, and a × indicates a shift coincidence prevents detection. For (i,i+n; n = 2–5) connectivities, a dashed line or bar indicates a probable assignment (potential ambiguity) and a dotted line indicates that a specific connectivity could be obscured by other peaks, but is assumed absent; all others are unambiguously absent.

Table II: Comparisons of ϕ/ψ Averages (and rmsds) for NOE-Refined Structures

	XPLOR-1.5 structures[a]			
	rmswv < 0.30	$J_{N\alpha}$ filter	relaxed structures[b]	major conformer[c]
ϕ_6	−18 (55)	−9 (69)	−36 (67)	−56 (9)
ψ_6	5 (37)	31 (31)	9 (38)	−25 (9)
ϕ_7	−131 (30)	−137 (22)	−119 (36)	−59 (9)
ψ_7	−45 (41)	−60 (29)	−39 (29)	−52 (3)
ϕ_8	−70 (55)	−61 (42)	−83 (27)	−64 (5)
ψ_8	75 (114)	41 (93)	78 (100)	+93 (5)
ϕ_9	10 (102)	12 (88)	−11 (93)	−48 (5)
ψ_9	−4 (25)	−14 (21)	4 (35)	−24 (10)
ψ_{10}	−70 (27)	−59 (21)	−87 (31)	−58 (3)
ψ_{10}	−59 (14)	−55 (12)	−57 (13)	−53 (4)
ϕ_{11}	−71 (12)	−75 (14)	−68 (8)	−59 (7)
ψ_{11}	−20 (7)	−22 (7)	−24 (12)	−30 (5)
ϕ_{12}	−66 (10)	−68 (12)	−68 (18)	−71 (6)
ψ_{12}	−53 (5)	−53 (4)	−43 (10)	−54 (5)
ϕ_{13}	−48 (9)	−46 (9)	−53 (13)	−44 (8)
ψ_{13}	−41 (13)	−46 (13)	−27 (18)	−36 (6)
ϕ_{14}	−76 (18)	−75 (20)	−99 (27)	−56 (9)
ψ_{14}	−11 (13)	−11 (25)	−6 (29)	−66 (6)
ϕ_{15}	−128 (28)	−120 (22)	−127 (29)	−69 (7)
ψ_{15}	−62 (17)	−54 (7)	−70 (17)	−23 (7)
ϕ_{16}	−93 (37)	−99 (28)	−85 (49)	+64 (13)
ψ_{16}	−31 (15)	−35 (13)	−12 (51)	+37 (71)
no. of structures	39	21	16	5

[a] rmswv is the weighted rms violation measure defined in the Experimental and Computational Methods section. [b] From Krystek et al. (1991). [c] Obtained from the conformer search protocol (vide infra).

Our first comparison of the NMR structures was based on the backbone dihedral values. The complete mean (and rmsd) dihedral comparison is given elsewhere;[8] only the stretch of dihedrals which include relatively well-defined angles (rmsd <40°) will be considered here (Table II). For the XPLOR structures only those displaying an rms weighted violation penalty measure of less than 0.3 Å are included in this table. The only region that could be said to be well-defined is a helical domain from ψ_9 through ψ_{14}. It was bracketed by an extremely variable 8(9)-peptide unit (ψ_8/ϕ_9 displays rms >80° by all protocols) and a nonhelical, but well-determined value for ϕ_{15}. As can be seen in Figure 2, ET-1 does display some scalar coupling values ($J_{N\alpha}$ < 6 Hz and $J_{N\alpha}$ ≥ 9 Hz) that are unlikely to represent conformational averages. Rather than use these as torsion constraints, we chose to apply them as a filter of structural hypotheses that displayed low NOE penalty values. Of the 39 XPLOR structures displaying rmswv <0.30 Å, 21 were consistent with the $J_{N\alpha}$ data. The $J_{N\alpha}$ filtering did not, however, remove the large variance of ψ_8/ϕ_9. The observed coupling in K9 can be attributed to either positive or negative values of ϕ_9.

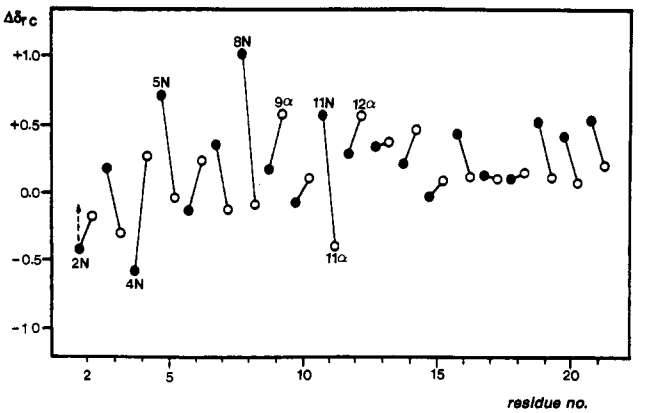


FIGURE 3: Sequential plot of shift deviations from random coil values (Wüthrich, 1986). For each residue the HN deviation is shown first (filled symbols) and then the Hα deviation (open symbols). The deviation ($\Delta\delta_{rc}$) is defined as $\delta_{rc} - \delta_{obs}$ such that observed upfield deviations correspond to positive values. The dashed arrow at N2 reflects our expectation that HN of residue 2 would display a downfield shift [we are currently determining the statistical significance of this observation, the 2-effect, first reported by Hammen (1990)].

All further comparisons between structures from different dynamics protocols used structures relaxed by a steepest descent minimization in the CHARMm force field. These FR structures (independent of the dynamics protocol used to generate the related SA structure) violated 32–49% of the initial set of NOE bounds by greater than 0.2 Å. The FR structures (four from CONGEN, five from DISCOVER, and seven from XPLOR) are summarized in third column of Table II. As previously noted and illustrated,[8] two consensus structural features could be identified in these structures: a "helical core" from residues 9 → 15 and a reverse turn at residues 6 and 7. The poor fit outside of the helical region suggests that some form of conformational averaging or structural randomization applies. Under either circumstance, no single structure would be found that fits all the constraints. *It should be noted that in the excess of 30 2D spectra which have recorded at pH = 5.8–7.8 and 3.2–4.5, not a single peak which might indicate a slowly interconverting minor conformer has been observed.* Multiple individual conformers, if present, must have lifetimes of less than 10 ms.

We chose to adopt, for ET-1 under these experimental conditions, a multiple populated conformer model. We assumed that the region from $\psi_9 \rightarrow \psi_{15}$ has essentially the same structure in all contributing conformers. For the remaining regions, we rejected an unstructured "random coil" model on the basis of two observations.

First, we do not observe the standard random coil shifts for any significant portion of the endothelin sequence. Figure 3

[8] For further details see the partial preliminary account (Krystek et al., 1991).

plots the deviation from random coil values for the HN and Hα resonances of ET-1. Only for residues 10, 15, 17, and 18 were both backbone resonances within ±0.2 ppm of the standard values. Second, intermediate and long-range NOEs are observed within (2HA/6HD%%, 3.6 Å; 3HB% (and/or 11HB%)/5HN, 3.25 Å; 3HN/6HD%%, 3.40 Å; 5HA/7HN, 3.8 Å; 5HB%/8HN, 3.15 Å; and 5HB%/7HN, 3.3 Å) and between motionally averaged regions and the persistently structured core (6HA/12HG2%, 2.7 Å; 7HA/9HE%, 3.7 Å; 8HB%/10HN, 3.35 Å; 12HA/17HN, 3.2 Å; 12HG1%/16HD2, 3.3 Å; and 12HG1%/16HE1, 3.6 Å), where percent indicates a single-character wild card. These interactions must reflect (a) discreet conformer(s). Finally, the slow exchange of M7-HN and the nearly zero $\Delta\delta/\Delta T$ values for K9, D8, and S5 could also be used to argue for persistent structure in a well-populated conformer.

Since it was apparent that another round of constrained dynamics runs would be required to search for specific conformers, rather than a single "structure" that should fit all constraints, we took this opportunity to confirm all NOE constraints. Additional experiments had suggested that three of the previous constraints represented misassigned or ambiguous peaks. From Figure 1 (and the chemical shifts in Table I) the peak ascribed to 12α/17HN, by careful examination of peak alignments, could contain a significant component of 12α/16HN intensity. In fact, the only conditions we have found that separate the backbone NH resonances of H16 and L17 are pH 5.8 at 310 K. This is shown in Figure 4. This experiment provided a number of new constraints as well. Figure 4 also shows β,γ,δ/HN and β,γ,α/α interactions under other experimental conditions, which provide unambiguous assignment for the key connectivities. In every case, the spectra at pH 3.2 are concordant with the assignments, even if they are not fully definitive by themselves.

Defining Conformers of ET-1 That Contribute in Acidic Media. Any structure that can be legitimately postulated as a major populated conformer under the experimental conditions would have to satisfy the following criteria. (1) The structure should retain any region of persistent structure required to rationalize backbone amide-NH exchange rate and/or temperature gradient data that suggests H-bonding or other sequestration from bulk water, and within that region, *all* NOE constraints should be valid. (2) In the other regions an acceptable conformer should display, for all DISCON-derived distances that are near either their minimum or maximum values (based on holonomic constraints), no significant violations. (3) The structure should not predict any strong- or medium-intensity NOEs ($r_{ij} < 3.2$ Å) which are clearly absent in the experimental data.

Turning to the first criterion, what constitutes the region of persistent structure that is presumably shared by the major conformers? Y13 and F14 display the slowest exchange ($t_{1/2} > 80$ min) and are in a helical domain; assuming the usual (i–4)C=O···H-N(i) H-bonding pattern, we take this as evidence for structuring from $\psi_9 \rightarrow \psi_{13}$ consistent with our previous conclusions based on NMR structure statistics.[9] A string of $i(i+3)$ connectivities extends the persistently structured region to at least ψ_{15}: in ca. one-third of the conformer search dynamics runs, the values (±30°) for $\psi_9 \rightarrow \psi_{15}$ appearing in the second column of Table II were used as torsion constraints (with a modest forcing energy).

For conformer searches, we have retained *all* $i(i+3)$ and sequential connectivities from the helical region. A number of $d_{\alpha N}(i+4)$ connectivities appear in this region, as is often observed in persistent helices in proteins, but these were either deleted or loosened since it has been suggested (Bruch et al., 1989) that $d_{NN}(i+2)$ and $d_{\alpha N}(i+4)$ NOEs for such regions are largely due to secondary NOE contributions. We retain, however, other long-range connectivities for which no obvious secondary relay point exists, most notably 12α → 17HN (12α → 16HN has zero intensity, but *we do observe a moderate-intensity NOE due to a side-chain interaction, 12α → 17β,γ, which was not included due to the potential ambiguity in the 17β,β′,γ assignment*).

Criterion 2 suggests $\phi = -125 \pm 30°$ (for residues 2, 3, and 20) as an added constraint necessary in major conformers in order to keep $J_{N\alpha} \geq 9$ Hz for the population-weighted average. Criterion 2, however, forces major pruning or loosening of NOE constraints outside of the helical region. All type b constraints (lower bound only) can be retained, which incorporates part of criterion 3 into the constraints. The complete lack of d_{NN} connectivities S5/L6 and D8/K9 fall in this category. The extremely weak $d_{\beta N}$ connectivities (3/4, 6/7, 8/9, 18/19) were also retained as low bound constraints. The only other $i(i+1)$ connectivities retained as high penalty constraints were based on the very strong α_i/HN_{i+1} NOEs observed for 1/2, 3/4, 5/6, 17/18, 19/20, and 20/21.

The occasional very weak $d_{\alpha N}(i+3)$, and more numerous $d_{\alpha N}(i+2)$, observations for the S5 → C11 region present a dilemma for a model that assumes conformational dynamics. This behavior has been noted previously for "nascent helices" (Dyson et al., 1988b). The $i(i+3)$ connectivities could be attributed to the partial population of helical conformations and were segregated into a separate constraint table which, when included, would direct the search toward frayed helices. However, the $d_{\alpha N}(i+2)$ connectivities have two possible origins for this type of system: (1) a specific set of linked turns characterized by $\psi_{i+1} \rightarrow 0°$, which gives the minimum value for the α_i/HN_{i+2} vector (3.05 Å), or (2) a "dynamic secondary NOE". The first explanation tolerates a much wider range of values for ψ_i and ϕ_{i+1}. The dynamic option is $\alpha_i \rightarrow HN_{i+1}$ (in one conformer, $\psi_i = +120°$ for optimum efficiency) followed by $NH_{i+1} \rightarrow HN_{i+2}$ (possibly in another conformer). The second step of the dynamics process requires $\phi_{i+1} = 0 \pm 70°$ but allows a larger variation for ψ_{i+1}. This argues against the use of $d_{\alpha N}(i+2)$ distance bounds for systems that are suspected to be dynamically unfolding. A number of $d_{\beta N}(i+2)$ distance bounds, however, were retained in unaltered form since the foregoing argument could not apply on the basis of intensities observed for the separate stages of potential two-step pathways. Since we were primarily concerned with defining a set of acceptable conformers for the bicyclic core, we excluded all but the shortest distance constraints for the C-terminus from the majority of the runs. The resulting annotated "conformer search" constraints are given in full in the supplementary material; the series of XPLOR-2.1 refinements and NOESY simulations used is outlined in the Experimental and Computational Methods section. Criterion 3 was reapplied several times during the conformer search, ultimately by performing a NOESY simulation for each conformer family being considered—predicted NOEs that were not observed in

[9] The $\Delta\delta/\Delta T$ data in Figure 2 present a contrary view if one accepts the general premise that there is a correlation between sequestration from water exposure (typically as an H-bond) and $\Delta\delta/\Delta T$ values that are significantly less negative than −6 ppb/°C [e.g., Dyson et al. (1988a) and Rose et al., (1985)]. Y13 and F14 display gradients which fall in the range (−6 → −8 ppb/°C) typical for flexible random coil peptides (Ni et al., 1988; Price et al., 1988). The near zero $\Delta\delta/\Delta T$ values appear at D8, K9, and C11 (and at pH ~7 also at E10). We are forced to conclude that there are other factors that influence $\Delta\delta/\Delta T$ than those commonly cited and will return to this point.

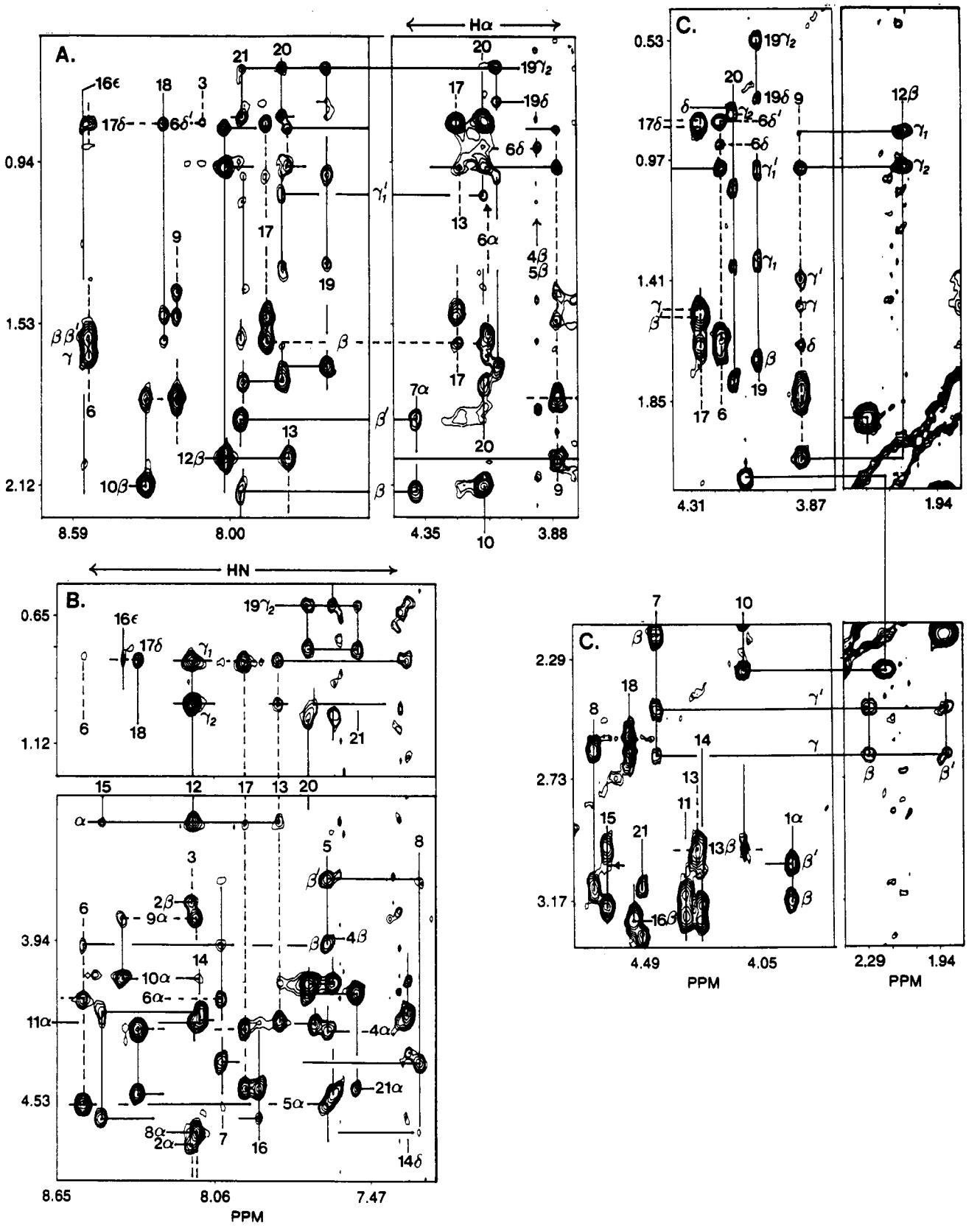


FIGURE 4: Definition of key NOE assignments through temperature and pH studies. Segments of three NOESY experiments are shown: panel A, the upfield portion of the lower panel in Figure 1 (pH 3.2, 60% glycol/H_2O, 309 K, τ_m = 100 ms); panel B, NH to α and upfield resonances at pH 5.8 (55% glycol/H_2O, 310 K, τ_m = 200 ms); and panel C, $\alpha H/\beta H$, $\alpha H/CH_3$, and $\beta H/CH_3$ cross-peaks at pH 7 (40% glycol/D_2O, 295 K, τ_m = 200 ms). In all panels, resonance positions corresponding to protons in residues 3, 6, 9, 13, and 17 are shown as dashed lines.

any experiment were converted to constraints, typically $d_{ij} \geq$ 3.4 Å, for the subsequent refinement stages.

Of 120 runs initiated, 22 conformer structures, which retained an excellent fit to the NOE distances upon minimization without distance constraints, were retained for further consideration and for further studies of ET conformational equilibrium changes with media. A comparison of these structures, and the initial set (Krystek et al., 1991), appears in Table III (and Table II). The structures resulting from the conformer search protocol all displayed acceptably low de-

Table III: Structure NOE Constraint and Geometry Violation Statistics[a]

		conformer search results			
			relaxed structures[b]		
	16 selected structures from Krystek et al. (1991)	22 SA structures	all (22)	major conformer (5)	others (5)
NOE penalty functions[c]					
original constraints[d]					
(136 ± 32 type b)					
⟨|viol|⟩	0.22, 0.30	0.21, 0.31	0.22, 0.32	0.24, 0.31	0.32, 0.34
% viol >0.2 Å	32, 36	26, 30	26, 30	27, 29	27, 29
search constraints[e]					
(89 + 81 type b)					
⟨|viol|⟩	0.11, 0.19	0.043, 0.059	0.042, 0.060	0.046, 0.051	0.053, 0.056
rmswv	0.39, 0.65	0.17, 0.24	0.16, 0.25	0.18, 0.21	0.18, 0.21
% viol >0.2 Å	15, 21	6, 9.7	6, 9.5	6, 7.8	8, 10.1
108 key constraints					
⟨|viol|⟩	0.10, 0.19	0.029, 0.045	0.028, 0.042	0.028, 0.033	0.037, 0.040
rmswv	0.39, 1.25	0.12, 0.21	0.12, 0.21	0.12, 0.14	0.12, 0.16
% viol >0.2 Å	14, 20	3, 7.3	3, 7.0	3, 5.2	6, 6.8
XPLOR statistics[f]					
E_{NOE} (kcal/mol)	nc	68.5 ± 35.9	77.8 ± 35.7	64.6 ± 15.4	59.3 ± 9.6
bond deviations (Å)	0.017 ± 0.007	0.007 ± 0.0004	0.020 ± 0.007	0.020 ± 0.003	0.021 ± 0.009
angle deviations (deg)	3.07 ± 0.19	2.42 ± 0.03	3.28 ± 0.06	3.29 ± 0.05	3.30 ± 0.06
improper torsions (deg)	2.78 ± 1.02	1.08 ± 0.05	4.40 ± 0.54	4.46 ± 0.46	4.43 ± 0.42
E_{repel} (kcal/mol)	2.5 ± 2.0	14.5 ± 6.4	1.7 ± 1.3	2.3 ± 1.9	2.0 ± 1.5
CHARMm energies[g]					
E_{L-J} (kcal/mol)	nc	153 ± 33	19.7 ± 12.0	14.6 ± 4.4	16.1 ± 11.6
E_{total} (kcal/mol)	nc	−83 ± 54	−297 ± 36	−296 ± 13	−286 ± 55

[a] nc indicates that no valid comparison is available between the structures from the preliminary study and the present conformer search structures. The units for all other measures are given in the left-hand column. [b] The major conformer group includes all structures in Figure 5; the other columns include the other structures appearing in Figure 6. [c] NOE constraint deviation measures are from the program PFS and include the average absolute deviation, the percentage of constraints violated by 0.2 Å or more, and the root-mean-square weighted violation (rmswv) as defined in the Experimental and Computational Methods section. The best and mean values for each set of structures are given. [d] The original set of constraints as modified only to reflect new data and previously unrecognized ambiguities. The average precision was ±0.37 Å. [e] The constraints used in the conformer search protocol (see supplementary table). The average precision was ±0.43 Å.

viations from standard geometry and an excellent fit to the constraints. In the case of conformers generated with XPLOR-2.1, constraint-free minimization produces a ca. 210 kcal/mol reduction in the CHARMm energy (of which 135 kcal/mol is in the van der Waals term) without a significant increase in any of the constraint violation measures. The FT conformers also display a superior fit to the initial set of constraints. When the fit is restricted to only those constraints (170) which should apply to all acceptable conformers, the ten best conformer models display eight or less violations greater than 0.2 Å. Five of these represented the same conformer and are illustrated in Figure 5. We consider this to be the best model for the major conformer of the core of endothelin in acidic aqueous media. The members of this cluster display no violations exceeding 0.38 Å, and all violations over 0.2 Å involve methyl distances.

Figure 6 illustrates the best refined structures from those conformation clusters which occurred with high frequency in the initial structure generation phase of the search. The best-fitting "major conformer", shown boldly in red-orange, was also obtained with highest frequency. A set of atom coordinates ($Cys^1 \rightarrow Leu^{17}$, only) for the major conformer has been deposited in the Brookhaven Protein Data Bank; the other structure coordinates are available upon request.

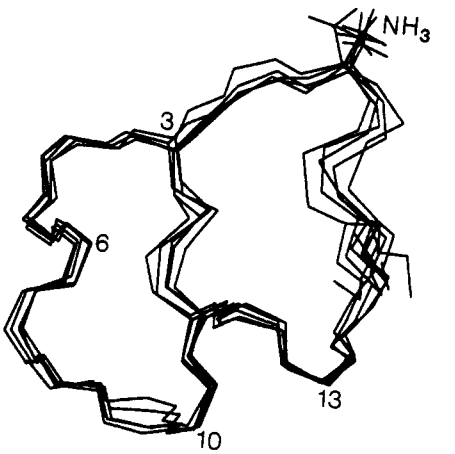


FIGURE 5: Major conformer cluster for endothelin in acidic aqueous media. The backbone atoms (Cys^1 through His^{16}) and disulfide linkages of five structures are shown, least-squares superpositioned from residue 1 through residue 15 (rmsd = 0.51 Å). The terminal NH_3 group and four Cα atoms are labeled.

CONCLUSIONS

In the case of endothelin we believe our search of conformational space is sufficiently exhaustive to allow us to conclude that no single conformer can rationalize the set of NOEs observed in acidic aqueous media. With the use of DISCON, the NOEs could be converted to a set of experimental distances and experimental precision estimates. No conformer that we have found comes close to a fit within the established experimental bounds of the constraints for the entire sequence. However, certain regions appear to have well-defined persistent structure (vide infra). Parallel constrained dynamics runs using the wider bounds typical of protein NMR studies indicate that this degree of structural definition would not have been reached without the measures of constraint precision available from the program DISCON. These precision estimates were also essential for distinguishing between the interconverting conformer model which we have adopted and a less informative "disordered" model. When we attempted to fit all constraints to a single structure, the family of NMR structures had wide variance in the C3 → D8 loop region. In contrast, the conformer search protocol allowed us to obtain the single best-fitting (and presumably highest population)

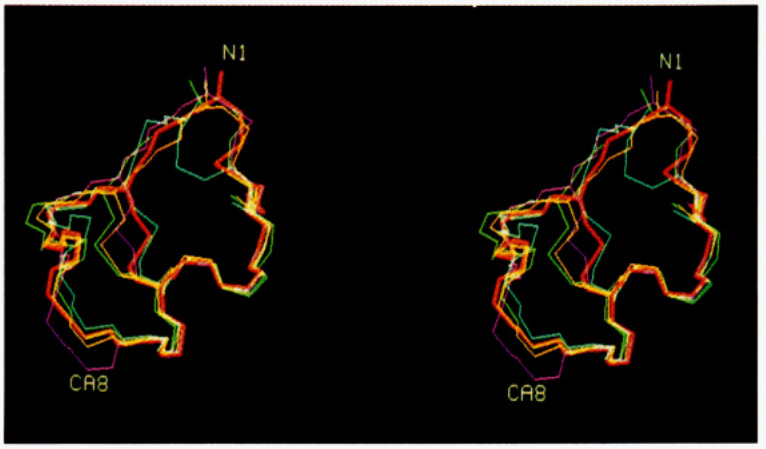


FIGURE 6: Stereoview of endothelin-1 conformer models. In each case the disulfides and backbone atoms of residues 1 through 16 are displayed, least-squares fitted over the backbone atoms of residues 9 through 15. The conformer shown boldly (red-orange) is viewed as the best model for the major contributor and has negative torsion values for ϕ_2, ϕ_4, ϕ_6, and ϕ_9 and both disulfides. The 8(9)-amide rotamer of the major conformer is shown in magenta. The minor conformers are blue (1,15-disulfide isomer), green (ϕ_6 positive), and yellow ($\phi_2 = +70°$).

conformer as a cluster with excellent definition throughout the bicyclic core (backbone rmsd = 0.51 Å over residues 1 → 15); see Figure 5.

Finally it should be noted that this high level of structural definition was obtained without prochirality assignments for βCH_2 units. If secure stereochemical assignments become available, they may serve to further refine the details of side-chain conformations and eliminate some of the potential conformer models from further consideration. For example, fully 30% of the otherwise acceptable structures could be rejected if our tentative assignment of the chirality at the Cys^3-βCH_2 was assumed to be definitive.

Turning to the structural and pharmacological implications of our studies, structure–activity data (Kimura et al., 1988; Nakajima et al., 1989; Kitazumi et al., 1990; Hunt et al., 1991) clearly reveal that most of the residues in the C-terminal tail are required for both receptor binding and activity; the other key residues are D8, E10, Y13, and F14. The structure and location of the C-terminus has been a major point of contention. Our data do not fully resolve the question, but they do provide firm evidence that the C-terminus is not the most mobile portion of the structure. Motional averaging is in fact more evident in an N-terminal loop region, Ser^2 → Ser^5. Within the C-terminal tail (L17 → W21), we observe large cross-peak intensities for $d_{\alpha N}$ and many other inter- and intraresidue connectivities. This would *not* be expected for a rapidly varying "unstructured" state. In contrast to this, a number of NOE intensities involving residues 2 → 5 are significantly smaller than expected (these include the intraresidue $d_{N\alpha}$'s for S2, C3, and S5, which yield distance constraints longer than allowed for vicinal protons). The NOESY data for apamin (Pease & Wemmer, 1988), which has an increasingly mobile C-terminus, show only $d_{\alpha N}$ connectivities near the C-terminus, and these decrease in intensity even as the structure becomes less helical: the expected result only in the case of increasing high-frequency motional averaging. A critical review of all reported NMR data for endothelins reveals that certain unusual chemical shifts are always observed; for example, the Ile^{19}-βMe is always far upfield, suggesting a preserved structural feature in all media, in analogues [ET-3 (Bortmann et al., 1991) and sarafotoxin 6b, in which one of the Val^{19}-Me signals corresponds to the Ile^{19}-βMe (Mills et al., 1991)], and in C-terminal fragments (unpublished studies). This site also shows an unusual NOE connectivity, 21HN/19γ_2; however, this and several other medium-range distance constraints cannot be fitted without inducing unacceptable violations of sequential constraints. We therefore propose that the C-terminus of endothelin is another instance of interconversion between at least two discreet conformers with lifetimes in the microsecond domain, not a case of rapid motional averaging. In the absence of stereospecific constraints, this multiconformer problem is underdetermined; we have been unable to define the contributing conformers.

Regarding the bicyclic core, all acceptable conformer models—independent of the inclusion of torsion constraints during the refinement or the use of the additional NOE constraints (that could dictate a frayed helix extending from K9 toward the N-terminus)—have the same backbone structure from K9-Cα through C15-C (rmsd over 9 → 15: N, Cα, C′, O; av 0.42 Å, maximum 0.57 Å). We were unable to generate any acceptable conformers with a frayed helix extending to Leu^6. However, a hydrophobic cluster that includes this side chain was evident (Figure 7, top panel) and may be a stabilizing feature which dictates the formation and specific conformation of the helical core. All ϕ and ψ torsion rmsds were less than ±18° from ψ_7 through ϕ_{16} (except at ψ_8/ϕ_9, as discussed below). As a result, the side chains of the central-region residues previously indicated as key (by pharmacological studies) are all well determined. The bottom panel of Figue 7 shows the consensus positions of D8, E10, Y13, and F14. In the case of D8, obviously stereospecific assignment will be required to refine the position. While it remains to be determined whether these side-chain orientations are retained near neutral pH or in membrane-mimicking media, the model should serve for modeling-based antagonist design.

Two primary loci of structural variance were found in the loop connecting the N-termini of the disulfide linkages and the persistent helix: ψ_3/ϕ_4 (which takes on values of +165/−65° or +20/+60°) and ψ_8/ϕ_9 (+85/−48° or −15/+55°). Additional points of variance were found at ψ_5/ϕ_6 (+100/−70° or +20/+50°) and in the disulfide dihedrals. In each case, the first set of values is found with greater frequency by our conformer search protocol, and we propose that this represents the major conformer of ET-1 under our experimental conditions. We can provide no firm evidence to support a correlation between the frequency of obtaining a particular conformation cluster in NOE-based refinement and conformer population but use it here as a working hypothesis.

The helical core is not a perfectly regular one; in most structures the Lys^9-CO is situated between 12HN and 13HN, generally better situated for H-bonding 13 → 9. The 15 →

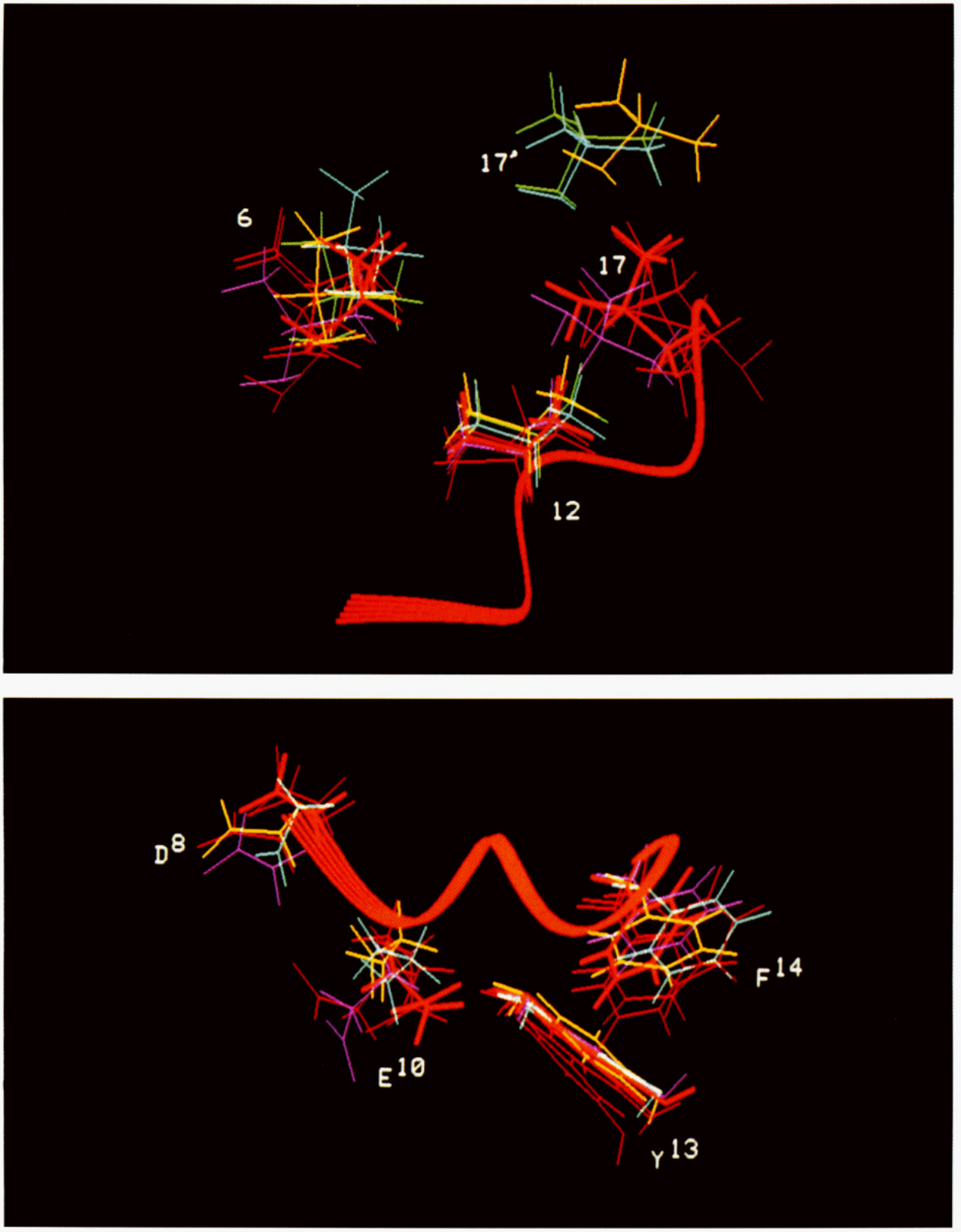


FIGURE 7: Side-chain dispositions in ET-1 conformers. In both panels the ribbon representation (D8 → H16) is for the major isomer of Figure 6, and the same coloring scheme applies. Four additional structures from the major cluster (Figure 5) are included to show the typical structural variance between NMR-derived structures. The perspective for the top panel corresponds to that of Figure 6; that of the bottom panel is related to the top panel by the rotation matrix: 0.693, −0.665, 0.276/0.621, 0.747, 0.238/−0.365, 0.006, 0.931. The top panel, the hydrophobic cluster (Leu[6], Val[12], and Leu[17]), shows two alternative positions for the Leu[17] methyl groups; that for the less populated conformer is shown as 17′. The bottom panel shows the location of the key binding and activity-imparting residues.

11 H-bond in our models has excellent geometry while that for 14 → 10 is rather long. The very slow exchange of 12HN may also reflect sequestration from water by the 3,11-disulfide and the hydrophobic cluster consisting of the side chains L6, V12, and L17 (and in some conformers also I19). The slow exchange of 11HN can be attributed to an H-bond with the Asp[8]-CO (present in most conformer models). A bent 10HN to 8CO hydrogen bond appears in a number of conformers.

The remaining instance of somewhat retarded amide-HN exchange is Met[7]. We can find no ready rationale for this observation. The previously noted reverse turn sequesters 8HN (bonded to 5CO), for which no exchange rate could be measured due to signal overlap. This somewhat ill-defined turn generally resembles a type I β-turn at residues 5 → 8.

The present study raises an issue of general interest: the interpretation of amide-NH $\Delta\delta/\Delta T$ data and its correlation with solvent exposure and/or persistence of H-bonding. For endothelin, we observe an unusually wide range of values, −11.7 → +3 ppb/°C, unlike that found in proteins or in smaller linear peptides. Changes occur upon neutralization, suggesting a conformational equilibrium that depends on the ionization state at C1/H16. Throughout, the enhanced negative values of $\Delta\delta/\Delta T$ bracket regions with very small or positive temperature gradients.

For a rapidly interconverting mixture of conformers, $\Delta\delta/\Delta T$ values will also reflect population changes when the interconverting conformers display widely divergent NH shifts. The plot of δ_{HN} deviations from standard values (Figure 3) shows a partial correlation with the $\Delta\delta/\Delta T$ values in Table I. The three most divergent δ_{HN} values (upfield for S5 and D8, downfield for S4) correspond to three of the most extreme $\Delta\delta/\Delta T$ values: D8, +3 ppb/°C (at pH ~7); S5, +1.9 ppb/°C; and S4, 11.4 ppb/°C. If the conformational state equilibrium can be satisfactorily modeled as

$$\text{structured} \rightleftarrows \text{random coil}$$

the increase in random coil population with increasing temperature would predict a linear correlation of $\Delta\delta/\Delta T$ and $\Delta\delta_{rc}$. Panel A of Figure 8 shows this correlation for all observations recorded in this study. A trend is observed; however, some of the extreme values of $\Delta\delta/\Delta T$, ca. −10 for C15 and V12 at pH 7 and near zero for K9 and E10 (at pH 7 only, for E10), cannot be rationalized in this manner. When the plot is restricted (panel B) to residues for which evidence of motional averaging has already been presented, the correlation becomes clear. It would appear that *the unusual gradient values for S4 and S5 reflect their highly divergent $\Delta\delta_{HN}$ values in a structured form* which is decreasingly populated at higher temperatures. The state (or states) which become(s) populated at higher temperatures apparently have NH shifts that are, on average, well approximated by the standard random coil values. We expect this type of analysis will be generally useful for recognizing and characterizing conformational averaging in medium-sized peptides. In the present case it confirms that all segments of the endothelin structure have definable structural motifs at 290–295 K. At the high-temperature limit (315 K) of this study, the averaging near Ser[4,5] may be so extensive as to approximate a flexible random coil model in the less populated (~30%) state.[10] However, this level of disorder does not preclude structural analysis. The flexibility in this region is coupled with changes in the disulfide dihedral (certainly for the 1,15-linkage, possibly also for the 3,11-linkage). Concerning this point, it can be noted that the

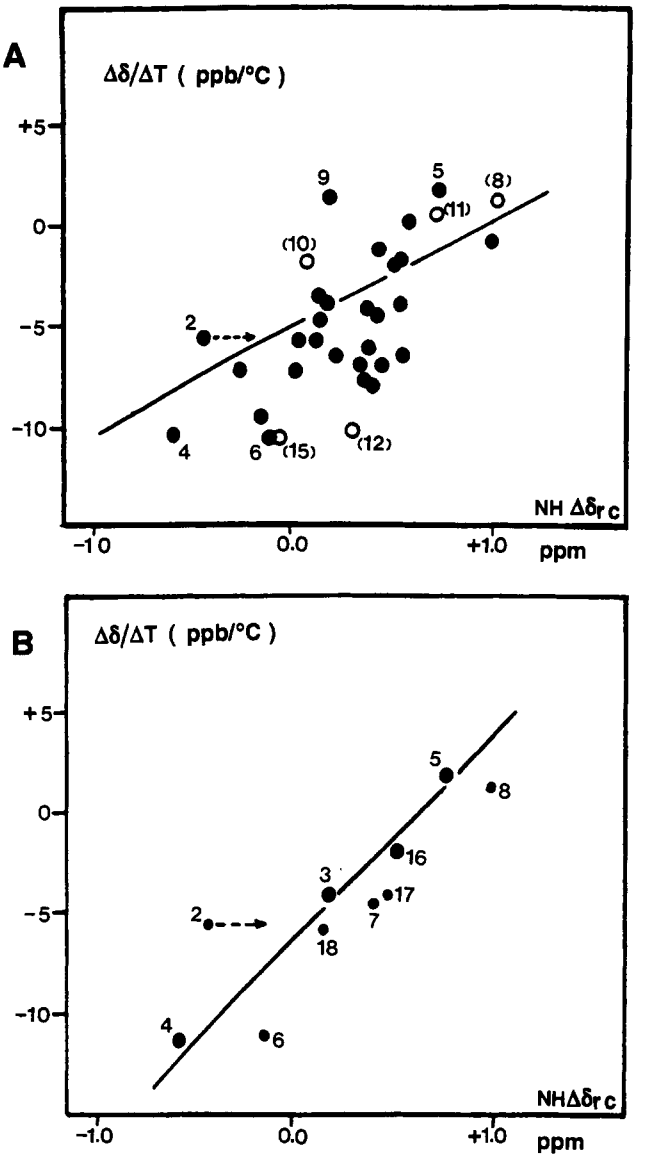


FIGURE 8: Correlation of $\Delta\delta/\Delta T$ and $\Delta\delta_{rc}$: panel A, all observations; panel B, restricted to residues for which motional averaging is suggested by either or both NOE analysis and HN cross-saturation. The statistical fit in panel A (r = 0.58) is all inclusive; in panel B the line is for residues 3, 4, 5, and 16 (r = 0.99). For all residues in panel B, r = 0.85. Residue labels in parentheses and open symbols are for pH 7 values.

C1-βCH$_2$ group is shift coincident at pH 3.2 but becomes nonequivalent at pH >5.8. This suggests that the degree of disorder is reduced upon changing the ionization state at H16 and/or C1 and that the region of residual disorder may shift with pH. Throughout, the C-terminus shows qualitative features suggesting a nonrandom state which has eluded definition to date. Studies of endothelin conformational equilibria at pH 5.1–7.6 are now in progress for both the free solution and micelle-associated states. More rigid analogues retaining full potency are also being examined. These results will be presented in due course.

ACKNOWLEDGMENTS

We thank David Floyd and J. T. Hunt (Bristol-Myers Squibb) for stimulating discussions concerning endothelin pharmacology and structure–activity studies, additional samples of commercial ET-1, and HPLC assay confirmations of peptide purity. We also thank Scott M. Harris for extensive help in recording CD spectra.

[10] The population estimate for the random coil state follows the slope of the correlation between $\Delta\delta_{rc}$ and the temperature gradient. The NOE data for this region can also be rationalized in this manner. Alternatively, the motions of the 8,9-amide unit and in the vicinity of Ser[4,5] may be sufficiently rapid so as to invalidate any kind of isotropic model including one based on discreet interconverting conformers. This would be the case for large amplitude motions on a time scale of less than $5\tau_c$ (ca. 10 ns in the present case). Cross-relaxation with a rapidly equilibrating hydration shell provides an alternative rationale for the diminished NOE intensities.

Supplementary Material Available

Three figures showing the HN/HN segment corresponding to the 295 K panel of Figure 1, CD spectra of ET-1 in aqueous buffer with and without added glycol and TFE, and CD spectra in the NMR medium spanning the concentration range from 13 to 1360 μM and three tables giving an outline of the XPLOR-1.5 protocol using NOE scale factor variation to avoid local minima traps, fully annotated comfiles for the XPLOR-2.1 conformer generation procedure, and a listing of the original and conformer search constraints (19 pages). Ordering information is given on any current masthead page.

Registry No. ET-1, 123626-67-5.

References

Andersen, N. H., Lin, B. S., & Nguyen, K. T. (1984) *Biochem. Biophys. Res. Commun. 121*, 702–709.

Andersen, N. H., Eaton, H. L., & Nguyen, K. T. (1987) *Magn. Reson. Chem. 25*, 1025–1034.

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

Andersen, N. H., Lai, X., Hammen, P. K., & Marschner, T. M. (1990a) in *NMR Applications in Biopolymers* (Finley, J. W., Schmidt, S. J., & Seranni, A. S., Eds.) pp 95–134, Plenum Publishing Co., New York.

Andersen, N. H., Lai, X., & Marschner, T. (1990b) *NOESYSIM/DISCON Documentation*, University of Washington, Seattle, WA.

Aumelas, A., Chiche, L., Mahe, E., Le-Nguyen, D., Sizun, P., Berthault, P., & Perly, B. (1991) *Int. J. Protein Res. 37*, 315–324.

Bassolino, D. S., Hirata, F., Kitchen, D., Kominos, D., Pardi, A., & Levy, R. M. (1988) *Int. J. Supercomput. Appl. 2*, 41–61.

Bennes, R., Calas, B., Chabrier, P. E., Demaille, J., & Heitz, F. (1990) *FEBS Lett. 257*, 21–24.

Boelens, R., Koning, T. M. G., & Kaptein, R. (1988) *J. Mol. Struct. 173*, 299–311.

Boelens, R., Koning, T. M. G., van der Marel, G. A., van Boom, J. H., & Kaptein, R. (1989) *J. Magn. Reson. 82*, 290–308.

Borgias, B. A., & James, T. L. (1988) *J. Magn. Reson. 79*, 493–512.

Bortmann, P., Hoflack, J., Pelton, J. T., & Saudek, V. (1991) *Neurochem. Int. 18*, 491–496.

Bremer, J., Mendz, G. L., & Moore, W. J. (1984) *J. Am. Chem. Soc. 106*, 4691–4696.

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

Brown, F. K., Hempel, J. C., Dixon, J. S., Amato, S., Mueller, L., & Jeffs, P. W. (1989) *J. Am. Chem. Soc. 111*, 7328–7333.

Brown, S. C., Donlan, M. E., & Jeffs, P. W. (1990) in *Proceedings of the 11th American Peptide Symposium* (Rivier, J. E., & Marshall, G. R., Eds.) pp 552–556, Pierce Chemical Co., Rockford, IL.

Bruccoleri, R. E., & Karplus, M. (1987) *Biopolymers 26*, 137–165.

Bruch, M. D., McKnight, C. J., & Gierasch, L. M. (1989) *Biochemistry 28*, 8554–8561.

Brunger, A. T., Clore, G. M., Groenenborn, A. M., & Karplus, M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3801–3805.

Clore, G. M., Brunger, A. T., Karplus, M., & Gronenborn, A. M. (1986) *J. Mol. Biol. 191*, 523–551.

Clore, G. M., Appella, E., Yamada, M., Matsushima, K., & Groenenborn, A. M. (1990) *Biochemistry 29*, 1689–1700.

Dalgarno, D., Senior, M., Slater, L., & Chackalamannil, S. (1990) Abstract, Eastern Analytical Symposium, Inc., Somerset, NJ.

Donlan, M. E., Brown, S. C., & Jeffs, P. W. (1991) *J. Cell. Biochem., Suppl. 15G*, 85.

Drobny, G., Pines, A., Sinton, S., Weitekamp, D. P., & Wemmer, D. (1979) *Faraday Symp. Chem. Soc. 13*, 49–55.

Dyson, H. J., Rance, M., Houghten, R. A., Lerner, R. A., & Wright, P. E. (1988a) *J. Mol. Biol. 201*, 161–200.

Dyson, H. J., Rance, M., Houghten, R. A., Wright, P. E., & Lerner, R. A. (1988b) *J. Mol. Biol. 201*, 201–217.

Eaton, H. L., & Andersen, N. H. (1987) *J. Magn. Reson. 74*, 212–225.

Endo, S., Inooka, H., Ishibashi, Y., Kitada, C., Mizuta, E., & Fujino, M. (1989) *FEBS Lett. 257*, 149–154.

Fesik, S. W., O'Donnell, T. J., Gampe, R. T., Jr., & Olejniczak, E. T. (1986) *J. Am. Chem. Soc. 108*, 3163–3170.

Freeman, C. M., Catlow, C. R. A., Hemmings, A. M., & Hider, R. C. (1986) *FEBS Lett. 197*, 289–296.

Güntert, P., Braum, W., Billeter, M., & Wüthrich, K. (1989) *J. Am. Chem. Soc. 111*, 3997–4004.

Hammen, P. K. (1990) Ph.D. Dissertation, University of Washington, Seattle, WA.

Haruyama, H., & Wüthrich, K. (1989) *Biochemistry 28*, 4301–4312.

Hunt, J. T., Lee, V. G., Stein, P. D., Hedberg, A., Liu, E. C.-K., McMullen, D., & Moreland, S. (1991) *Bio-Org. Med. Chem. Lett. 1*, 33–38.

Hyberts, S. G., Märki, W., & Wagner, G. (1987) *Eur. J. Biochem. 164*, 625–635.

Kimura, S., Kasuya, Y., Sawamura, T., Shinmi, O., Sugita, Y., Yanagisawa, M., Goto, K., & Masaki, T. (1988) *Biochem. Biophys. Res. Commun. 156*, 1182–1186.

Kitazumi, K., Shiba, T., Nishiki, K., Furukawa, Y., Takasaki, C., & Tasaka, K. (1990) *FEBS Lett. 260*, 269–272.

Kloog, Y., Ambar, I., Sakolvsky, M., Kochva, E., Wollberg, Z., & Bdolah, A. (1988) *Science 242*, 268–270.

Kraulis, P. J., Clore, G. M., Nilges, M., Jones, T. A., Petterson, G., Knowles, J., & Gronenborn, A. M. (1989) *Biochemistry 28*, 7241–7257.

Krystek, S. R., Jr., Bassolino, D. A., Novotny, J., Chen, C., Marschner, T. M., & Andersen, N. H. (1991) *FEBS Lett. 281*, 212–218.

Kumagaya, S.-I., Kuroda, H., Nakajima, K., Watanabe, T. X., Kimura, T., Masaki, T., & Sakakibara, S. (1988) *Int. J. Pept. Protein Res. 32*, 519–526.

Lane, A. N., Lefevre, J.-F., & Jardetzky, O. (1986) *J. Magn. Reson. 66*, 201–218.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.

Metzler, W. J., Hare, D. R., & Pardi, A. (1989) *Biochemistry 28*, 7045–7052.

Mills, R. G., Atkins, A. R., Harvey, T., Junius, F. K., Smith, R., & King, G. F. (1991) *FEBS Lett. 282*, 247–252.

Moore, J. A., Case, D. A., Chazin, W. J., Gippert, G. P., Havel, T. G., Powls, R., & Wright, P. E. (1988) *Science 240*, 314–317.

Munro, S., Craik, D., McConville, C., Hall, J., Searle, M., Bicknell, W., Scanlon, D., & Chandler, W. (1991) *FEBS Lett. 278*, 9–13.

Nakajima, K., Kubo, S., Kumagaye, S.-I., Nishio, H., Tsunemi, M., Inui, T., Kuroda, H., Chino, H., Watanabe, T. X., Kimura, T., & Sakakibara, S. (1989) *Biochem. Biophys.*

Res. Commun. 163, 424-429.
Ni, F., Scheraga, H. A., & Lord, S. T. (1988) *Biochemistry 27*, 4481-4491.
Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988) *FEBS Lett. 229*, 317-324.
Nilges, M., Clore, G. M., & Gronenborn, A. M. (1990) *Biopolymers 29*, 813-822.
Olejniczak, E. T., Gampe, R. T., Jr., & Fesik, S. W. (1986) *J. Magn. Reson. 78*, 28-41.
Otting, G., Widmer, H., Wagner, G., & Wüthrich, K. (1986) *J. Magn. Reson. 66*, 187-193.
Pease, J. H., & Wemmer, D. E. (1988) *Biochemistry 27*, 8491-8498.
Pelton, J. T. (1991) *Neurochem. Int. 18*, 485-489.
Perkins, T. D. T., Hider, R. C., & Barlow, D. J. (1990) *Int. J. Pept. Protein Res. 36*, 128-133.
Price, W. S., Medz, G. L., & Martenson, T. E. (1988) *Biochemistry 27*, 8990-8999.
Reily, M. D., & Dunbar, J. B., Jr. (1991) *Biochem. Biophys. Res. Commun. 178*, 570-577.
Rose, G. D., Gierasch, L. M., & Smith, J. A. (1985) *Adv. Protein Chem. 37*, 1-106.
Saudek, V., Hoflack, J., & Pelton, J. T. (1989) *FEBS Lett. 257*, 145-148.
Saudek, V., Hoflack, J., & Pelton, J. T. (1991) *Int. J. Pept. Protein Res. 37*, 174-179.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286-292.
Tamaoki, H., Kobayashi, Y., Nishimura, S., Ohkubo, T., Kyogoku, Y., Nakajima, K., Kumagaye, S., Kimura, T., & Sakakibara, S. (1991) *Protein Eng. 4*, 509-518.
Summers, M. F., South, T. L., Kim, B., & Hare, D. R. (1990) *Biochemistry 29*, 329-340.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Wüthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol. 169*, 949-961.
Yanagisawa, M., & Masaki, T. (1989) *Trends Pharmacol. Sci. 10*, 374-378.
Yanagisawa, M., Kurihara, H., Kimura, S., Tomobe, Y., Kobayashi, M., Mitsui, Y., Yazaki, Y., Goto, K., & Makaki, T. (1988) *Nature 332*, 411-415.
Zuiderweg, E. R. P., Boelens, R., & Kaptein, R. (1985) *Biopolymers 24*, 601-611.

Spectroscopic Evidence for an Intermediate in the T_6 to R_6 Allosteric Transition of the Co(II)-Substituted Insulin Hexamer†

Larry Gross‡ and Michael F. Dunn*
Department of Biochemistry, University of California at Riverside, Riverside, California 92521-0129
Received August 5, 1991; Revised Manuscript Received October 31, 1991

ABSTRACT: The phenol-induced conformational transition in the insulin hexamer is known to involve a large change in structure wherein residues 1-8 of the insulin B-chain are transformed from an extended coil (T-state) to a helix (R-state). This change in protein conformation both exposes a cryptic protein pocket on each subunit to which phenol binds and forces the HisB10 zinc sites to undergo a change in coordination geometry from octahedral to tetrahedral [Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Reynolds, C. D., Smith, G. D., Sparks, C., & Swensen, D. (1989) *Nature 338*, 593-596]. Substitution of Co(II) for Zn(II) at the HisB10 sites introduces a sensitive chromophoric probe of the structural and chemical events that occur during this allosteric transition [Roy, M., Brader, M. L., Lee, R. W.-K., Kaarsholm, N. C., Hansen, J. F., & Dunn, M. F. (1989) *J. Biol. Chem. 264*, 19081-19085]. In this study, using rapid-scanning stopped-flow (RSSF) UV-visible spectroscopic studies, we demonstrate that a transient chemical intermediate is formed during the phenol-induced conversion of Co(II)-substituted hexamer from the T-state to the R-state. Decomposition of the RSSF spectra gave a spectrum for the intermediate with *d-d* transitions consistent with the assignment of the intermediate as either a distorted tetrahedral or a 5-coordinate Co(II) species. Possible structures for the intermediate and the implications of these findings to the allosteric mechanism are considered.

Crystallographic structure determinations (Baker et al., 1988; Derewenda et al., 1989) have shown that the T_6 to R_6 conformational change involves large alterations both of the conformation of residues 1-8 of the insulin B-chain and of the geometry of the HisB10 metal sites.[1] The conformation of residues B1-B8 is extended in the T-state and coiled to a helix in the R-state. The T_6-R_6 transition is driven by ligand binding; the change to a helical conformation opens six binding pockets for phenol which are occluded in the T-state by the LeuB6 side chains. The HisB10 metal sites undergo a change in geometry from octahedral coordination in the T-state (with a ligand field consisting of three HisB10 imidazolyl moieties and three water molecules), to tetrahedral coordination in the R-state (wherein the three water molecules are replaced by

† Supported by NIH Grant 1-RO1-DK 42124-02.
* To whom correspondence should be addressed: Department of Biochemistry 015, University of California at Riverside, Riverside, CA 92521-0129.
‡ Present address: Upjohn Co., Kalamazoo, MI 49001.

[1] Abbreviations: T, the conformation of an insulin subunit in the two-zinc hexamer; R, the conformation of an insulin subunit in the phenol-induced hexamer; T_6, T_3R_3, and R_6, the three crystallographically identified allosteric forms of the insulin hexamer; M(II)-T_6 or M(II)-R_6 (where M = Zn or Co), the metal ion substituted in the HisB10 sites of the insulin hexamer; RSSF; rapid-scanning stopped flow; SWSF, single-wavelength stopped flow.

0006-2960/92/0431-1295$03.00/0