Determination of the Solution Structure of the Peptide Hormone Guanylin: Observation of a Novel Form of Topological Stereoisomerism

Nicholas J. Skelton,*,‡ K. Christopher Garcia,‡,§ David V. Goeddel,§ Clifford Quan,|| and John P. Burnier||

Departments of Protein Engineering, Molecular Biology, and Bio-Organic Chemistry, Genentech, Inc., South San Francisco, California 94080

Received June 20, 1994; Revised Manuscript Received September 2, 1994⊗

ABSTRACT: Guanylin is a 15 amino acid mammalian hormone containing two disulfide bonds. Guanylin shares sequence similarity with the bacterial heat-stable enterotoxin (STa) and is capable of binding to and stimulating the STa guanylyl cyclase receptor. Biologically active peptides have been prepared by two methods: (1) enzymatic treatment of a 99 residue proprotein (denoted proguanylin) expressed in *Escherichia coli* and (2) solid-phase chemical synthesis. Although both sources yield material that is pure by high-performance liquid chromatography and mass spectrometry, analysis by nuclear magnetic resonance (NMR) indicates that peptides from both sources contain two conformationally distinct species present in a 1:1 ratio. The chemical shift differences between the two species are large, allowing unambiguous sequential NMR assignments to be made for both sets of resonances. Exchange between the two forms was not observed even at 70 °C. Structural restraints have been generated from nuclear Overhauser effects and scalar coupling constants and used to calculate structures for both forms using distance geometry and restrained energy minimization. The resulting structures for the first isoform are well defined (root-mean-square deviation from the average structure for backbone atoms of 0.47 Å) and adopt a right-handed spiral conformation, similar to that observed for heat stable enterotoxin. The second isoform is less well defined (root-mean-square deviation from the average structure for backbone atoms of 1.07 Å) but clearly adopts a very different fold consisting of a left-hand spiral. The differences in structure suggest that the two forms may have very different affinities toward the STa receptor. The observation of such isomerism has important implications for the common practice of introducing multiple disulfide bonds into small peptides to limit conformational flexibility and enhance bioactivity.

The heat stable enterotoxins, ST,[1] have been known for some time to activate an intestinal guanylyl cyclase denoted STaR (Schulz et al., 1990; de Sauvage et al., 1991). The initial activation and overproduction of cGMP in intestinal endothelial cells disturbs the balance of fluid uptake and ion secretion in the intestines, resulting in severe diarrhea (Field et al., 1978; Giannella & Drake, 1979; Rao et al., 1981). Infection by enterotoxigenic strains of *Escherichia coli* producing these toxins are the cause of substantial infant mortality in many developing countries and also of traveler's diarrhea (Gordon, 1971; Levine et al., 1977; Sack, 1980). Numerous studies have revealed that the heat stable enterotoxins are small peptides, 18 or 19 residues in length, containing three disulfide bonds (Chan & Giannella, 1981; Yoshimura et al., 1985). Chemical studies (Yoshimura et al., 1985; Shimonishi et al., 1987), NMR (Gariépy et al., 1986; Ohkubo et al., 1986), and X-ray crystallography (Ozaki et al., 1991) have all been employed to determine the structure of active portions of the toxins.

Recently, an endogenous ligand to intestinal guanylyl cyclase has been isolated and cloned from a number of mammalian sources (Currie et al., 1992; de Sauvage et al., 1992; Schulz et al., 1992; Wiegand et al., 1992). The 15 residue peptide responsible for this activity, given the name guanylin, binds to STaR, competes with STa for binding to STaR, and also stimulates cGMP production (Currie et al., 1992) and chloride ion secretion (Forte et al., 1993) in intestinal T_{84} cells. In the cGMP production assay, a higher concentration of guanylin than STa was needed to elicit a response, and the maximal yield of cGMP was approximately half of that produced by STa (de Sauvage et al., 1992). In humans, guanylin is found predominantly in the small intestine (Currie et al., 1992; de Sauvage et al., 1992) and thus has been postulated to act in the regulation or modulation of intestinal fluid and electrolyte absorption (de Sauvage et al., 1992).

Isolation of mRNA indicates that guanylin is expressed *in vivo* as a 99 amino acid proprotein, with the active peptide residing at the C-terminus (de Sauvage et al., 1992). The probable processing sites of the proprotein suggest that the

* To whom correspondence should be addressed.
‡ Department of Protein Engineering.
§ Department of Molecular Biology.
|| Department of Bio-Organic Chemistry.
⊗ Abstract published in *Advance ACS Abstracts,* October 15, 1994.
[1] Abbreviations: ACM, acetamidomethyl; BOC, *tert*-butyloxycarbonyl; COSY, correlation spectroscopy; DIEA, diisopropylethylamine; DG, distance geometry; FID, free induction decay; Guan13 (Guan22), C-terminal 13 (22) residue peptide of proguanylin; HBTU, *O*-benzotriazol-1-yl-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate; JR-NOESY, 2D NOESY spectrum acquired with a jump−return observe pulse; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, two-dimensional NOE spectroscopy; rEM, restrained energy minimization; RMSD, root-mean-square deviation; TOCSY, total correlation spectroscopy; STa, heat-stable enterotoxin originally obtained from *E. coli* strains isolated from humans (also known as STh and ST1b); STaR, STa receptor, also known as the guanylyl cyclase C receptor (GC-C) or cytoskeletal-associated GC receptor; STp, heat-stable enterotoxin originally obtained from bacterial isolates of porcine origin (see Figure 7A for sequence details); TPPI, time-proportional phase incrementation.

0006-2960/94/0433-13581$04.50/0

physiologically important fragment may be larger than the species initially found to possess activity; the 15 residue peptide may have been the result of cleavage of an acid labile aspartate−proline peptide bond during purification (Garcia et al., 1993). The proprotein, termed proguanylin, has been cloned and expressed in *E. coli*, allowing the efficient production of material in quantities sufficient for biophysical characterization (Garcia et al., 1993). Proguanylin itself has no binding affinity for STaR and cannot induce production of cGMP in cells transfected with STaR. Removal of different N-terminal portions from proguanylin by proteolysis revealed that the C-terminal 32 or 22 residues do possess biological activity and can displace ^{125}I-labeled STa from STaR (Garcia et al., 1993).

Although guanylin shares considerable sequence identity with the STa toxins, nothing is known about its three-dimensional structure. In this paper, we report the solution structure of human guanylin determined by NMR methods. Experiments have been performed on the 22 residue C-terminal fragment of proguanylin, prepared by tryptic cleavage of the proprotein expressed in *E. coli*. In addition, a smaller 13 residue peptide prepared by chemical synthesis has been studied in order to answer questions about the cysteine pairings of disulfide bonds in the recombinant material. These structures allow a comparison to previous structural studies on STa peptides, providing insight into the structural epitope required by both peptides for activity. Further, this provides a step toward understanding the interaction of peptides with the intestinal guanylyl cyclase receptor and thus the development of therapeutically useful antagonists.

MATERIALS AND METHODS

Sample Preparation. The expression of proguanylin in *E. coli*, the purification of the proprotein, and generation of the tryptic C-terminal fragments have been described previously (Garcia et al., 1993). The recombinant peptide, named Guan22, was biologically active, pure by analytical HPLC analysis, and had the expected mass of 2258 Da (Garcia et al., 1993).

Peptide Synthesis. The peptide corresponding to the 13 C-terminal residues of proguanylin (Guan13) was synthesized by the solid-phase method (Merrifield, 1963) manually on BOC-Cys(4-methylbenzyl)phenylacetamidomethyl resin. Coupling was effected with HBTU and diisopropylethylamine in dimethylacetamide. During the synthesis, the cysteine thiol groups were protected with 4-methylbenzyl (residues 99 and 91) or ACM groups (residues 96 and 88). The peptide was cleaved from the resin with hydrogen fluoride in a mixture of anisole and ethyl methyl sulfide at 0 °C for 60 min. The linear peptide was purified by reverse-phase chromatography, and mass spectrometry indicated that ACM groups were still present on residues 88 and 96. The partially protected peptide was air-oxidized to form the first disulfide bond between Cys91 and Cys99. The ACM protecting groups were removed and the second disulfide bond formed between Cys88 and Cys96 using iodine and acetic acid. The resulting peptide was purified by reverse-phase HPLC on a 45 × 300 mm Vydac C_{18} column, using a gradient of 10−25% acetonitrile in 0.1% TFA, and had the expected mass of 1304.37.

NMR Spectroscopy. For NMR spectroscopy, peptide was dissolved in 440 μL of 95% H_2O/5% D_2O and the pH adjusted to 5.0 by microliter additions of 1 M NaOH, yielding total peptide concentrations of 4.0 or 3.0 mM for Guan22 and Guan13, respectively. Samples in D_2O solution were prepared by repeated lyophilization from D_2O, with final dissolution in 99.996% D_2O (Cambridge Isotopes); after the initial lyophilizations, the pH was adjusted to 5.0 (uncorrected meter reading) by the addition of 0.1 M NaOD or DCl.

All spectra were acquired at 275 or 280 K on a Bruker AMX-500 spectrometer. NOESY (Kumar et al., 1980; Bodenhausen et al., 1984) and TOCSY (Braunschweiler & Ernst, 1983; Bax & Davis, 1985) spectra were acquired from H_2O and D_2O solution. In addition, COSY (Aue et al., 1975), 2Q (Braunschweiler et al., 1984; Rance & Wright, 1986), and JR-NOESY (Plateau & Guéron, 1982) spectra were acquired from H_2O solution, and COSY-35 (Bax & Freeman, 1981) and 2QF-COSY (Rance et al., 1983) spectra were acquired from D_2O solution. All data were complex in ω_2, with phase discrimination in ω_1 achieved with TPPI (Marion & Wüthrich, 1983). In all spectra (except the JR-NOESY), water suppression was achieved by phase-coherent low power irradiation of the water resonance prior to the pulse sequence for 1.5 to 2.0 s and during the mixing time of NOESY data acquired from H_2O solution. NOESY, JR-NOESY, and TOCSY spectra were acquired as described previously (Akke et al., 1991); in addition, first-order phase corrections were avoided by acquisition in a sine-modulated fashion in ω_1. TOCSY mixing was achieved with the clean DIPSI-2rc sequence applied for 65 or 95 ms (Cavanagh & Rance, 1992). NOESY spectra were collected with mixing times of 200, 300, and 400 ms for Guan22 and 200, 400, and 500 ms for Guan13.

Slow amide proton exchange with solvent was determined by lyophilizing the peptide from H_2O solution. The sample was then dissolved in ice-cold D_2O, immediately placed in a pretuned and calibrated spectrometer, and short TOCSY spectra acquired consecutively for the next 90 min. The first TOCSY was started 4 min after addition of the solvent, and each TOCSY required a total acquisition time of 12 min (four scans for each of 128 t_1 increments). Intrinsic amide proton exchange rates were calculated as described previously (Englander et al., 1979).

Data Processing and Analysis. The spectra were processed and analyzed using Felix 2.1 (Biosym Technologies Inc., San Diego, CA). Prior to Fourier transformation in ω_2, each FID was corrected for DC offset using the final 5% of the FID, the first point was linear predicted and halved in intensity (Otting et al., 1986), and the FID multiplied by a weak Lorentzian-to-Gaussian window function. In ω_1, the first point of cosine modulated data was multiplied by 0.5, and the interferogram multiplied by a cosine bell followed by the application of a weak Lorentzian-to-Gaussian function; the interferograms were zero-filled to yield 1024 real points in each of the final columns of the matrix.

Distance Restraints. NOE distance restraints were generated from spectra acquired at 275 or 280 K and with mixing times of 300 ms for Guan22 and 500 ms for Guan13. Cross-peaks in the H_2O NOESY spectra were categorized as strong, medium, or weak on the basis of the integrated volume and assigned upper bounds of 3.3, 4.0, or 5.0 Å, respectively. The restraint distance used for overlapped cross-peaks was increased by one category or set to the maximum distance of 5.0 Å depending on the severity of the overlap. Cross-

peak volumes derived from all spectra involving methyl groups, degenerate methylene resonances, and protons of rapidly flipping tyrosine side chains, were divided by the number of protons contributing to the resonance before categorization (Yip, 1990). Restraints involving nonstereospecifically assigned methylene protons and the ring protons of tyrosine were applied to pseudoatoms, and the upper bounds increased by 1.0 and 2.0 Å, respectively (Wüthrich et al., 1983). Pseudo atom corrections of 0.5 Å were applied to restraints involving methyl groups (Koning et al., 1990).

Dihedral Angle Restraints and Stereospecific Assignments. $^3J_{HN-H\alpha}$ scalar coupling constants were determined from high digital resolution COSY spectra, with ω_2 cross sections through H^N-H^α (Kim & Prestegard, 1990). ϕ backbone dihedral angles were restrained to $-90°$ to $-40°$ if $^3J_{HN-H\alpha}$ was less than 6.0 Hz and $-140°$ to $-90°$ if $^3J_{HN-H\alpha}$ was greater than 8.0 Hz. For non-glycine residues in which $^3J_{HN-H\alpha}$ was in the range 6.0 → 8.0 Hz, ϕ was restrained to be negative if the intraresidue H^N-H^α NOE was less intense than the sequential $H^\alpha-H^N$ NOE (Clubb et al., 1994). For two residues (Cys91 in the A-form and Cys96 in the B-form), the intraresidue NOE was more intense, and initial structure calculations with no restraint imposed on these angles consistently produced structures with a positive value of ϕ. Subsequent structures were calculated with these two ϕ angles restrained to 30° → 90° (Ludvigsen & Poulsen, 1992).

$^3J_{H\alpha-H\beta}$ coupling constants were obtained from COSY-35 spectra. Stereospecific assignments of $C^\beta H_2$ groups could not be obtained directly from analysis of the intraresidue and sequential NOEs; in cases where one large (>9.0 Hz) and one small (<5.0 Hz) $^3J_{H\alpha-H\beta}$ coupling was observed, initial structures were calculated with χ_1 restrained to the two classical rotamer wells $-60 \pm 60°$ and $180° \pm 60°$. Due to the influence of the restraints imposed by the disulfide bonds and the medium and long-range NOEs, only a single classic rotamer was populated in the resulting structures; stereospecific assignments could then be made and χ_1 restraints tightened to include only a single rotamer well on the basis of observed $^3J_{H\alpha-H\beta}$.

Calculation of Structures. Structures were calculated using the DG-II program (Havel, 1991) within the INSIGHT package (Biosym Technologies Inc., San Diego, CA). The bounds matrix was subject to triangle and tetrangle smoothing prior to prospective embedding of all atoms in four dimensions. The embedded structures were optimized by annealing for 10000 steps in four-dimensional space while cooling from 200 K with all atom masses set to 1000. A subset of the optimized structures having the lowest penalty functions was further subjected to (100 steepest descents + 1000 conjugate gradient) × 2 cycles of energy minimization in the all atom AMBER force field (Weiner et al., 1986) using the DISCOVER package (Biosym Technologies Inc., San Diego, CA) with a nonbonded cutoff of 12 Å. In order to minimize artifacts from the absence of explicit solvent molecules in these calculations, partial atomic charges on charged residues were reduced to give a total charge of 0.2, and the dielectric constant was set to $1/r$, where r is the distance between the interacting atoms. Experimental restraints were enforced by flat-bottomed square-well potentials with force constants of 25 kcal·mol^{-1}·Å^{-2} and 25 kcal·mol^{-1}·rad^{-2} for distances and dihedral angles, respectively.

Families of 30 to 50 structures were calculated for the initial rounds of calculation. In the final round of structure calculations, 160 structures were embedded. The resulting structures exhibited a range of penalty values after the optimization stage of DG-II (the penalty value reflects both deviations from perfect covalent geometry and violations of the input restraints). The 80 structures with lowest penalty function (<0.013 for the A-form and <0.010 for the B-form) were further refined by rEM. The 20 structures having the lowest violation energy term after rEM were used in further analysis. Restraints to enforce hydrogen bonds were not used owing to the difficulty of identifying hydrogen bond acceptors in nonregular regions of secondary structure.

Analysis of the Solution Structures. For RMSD comparison of structures, the structure with the lowest restraint energy was selected as a reference, all others were superimposed on it, and the geometric mean structure calculated. In some cases, the structure with the lowest backbone RMSD from the geometric mean was used as a single representative structure (Sutcliffe, 1993). The overlaid structures were visually inspected using INSIGHT-II (Biosym Technologies Inc., San Diego, CA) or MIDAS (Ferrin et al., 1988) running on Silicon Graphics Indigo-2 workstations. Angular order parameters (S^{ang}) were calculated for ϕ, ψ, and χ_1 dihedral angles (Hyberts et al., 1992); an order parameter of unity indicates that the angle is the same in all structures, where as an order parameter of zero indicates that there is no preferred orientation about a particular torsion angle.

RESULTS

Resonance Assignments. Initial analysis of the COSY and TOCSY spectra of the Guan22 peptide indicated that twice the expected number of spin systems were present. Peak intensities in the 1D and 2D spectra indicated that the two forms were present in approximately equal amounts. Extensive reverse-phase HPLC analysis failed to separate the two species, and the addition of cosolvents (up to 35% v/v d_3-acetonitrile or d_6-DMSO as cosolvent) or changing temperature (from 275 K up to 340 K) failed to induce broadening of the resonances indicative of exchange between the species. However, the spectrum was well resolved, and resonance assignments could be made for both species, which are designated "A" and "B" in the remaining discussion. Complete spin systems were identified using the COSY, 2Q, and TOCSY spectra (Figure 1), and the spin systems were assigned sequence specifically using sequential H^N-H^N, $H^\alpha-H^N$ and $H^\beta-H^N$ NOEs in a standard manner (Wüthrich, 1986). These assignments are listed in Table 1, and the sequential $H^\alpha-H^N$ NOEs used for assignment of the A-form are shown in Figure 2. Note that the residue numbering scheme of proguanylin is used throughout; hence the residues in Guan22 are numbered from 78 to 99. An inspection of Table 1 indicates that the chemical shifts of the A and B form are very different for the 13 C-terminal residues, with 11 H^N resonances differing by more than 0.2 ppm (5 differ by more than 0.5 ppm) and 8 H^α protons differing by more than 0.1 ppm. This observation suggests that the backbone fold of the two forms is quite different.

The sequential NOEs used in the assignments of the two forms along with $^3J_{HN-H\alpha}$ coupling constants and the rates of amide proton exchange with solvent are summarized in Figure 3. The magnitude of $^3J_{HN-H\alpha}$ coupling constants,

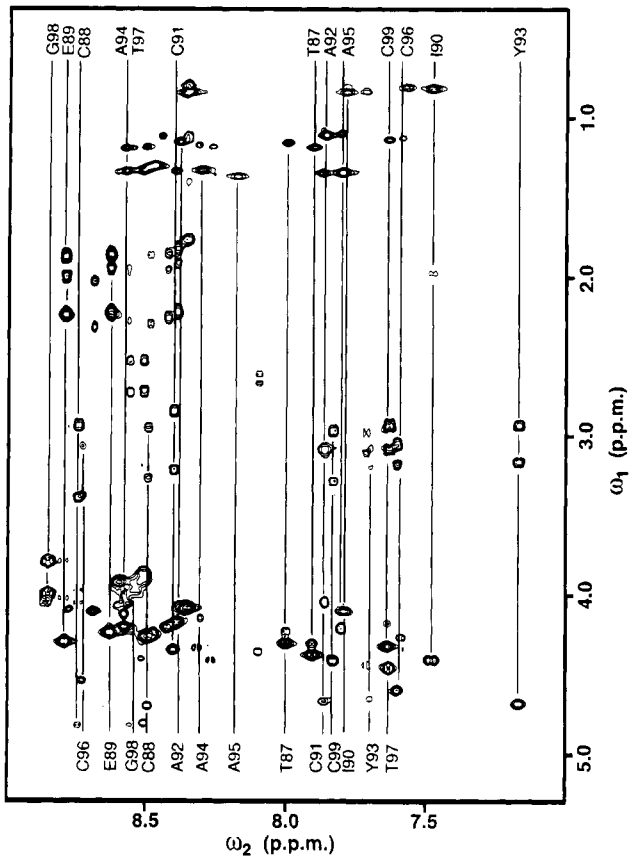

FIGURE 1: A section of the TOCSY spectrum (τ_m = 90 ms) of Guan22 acquired from H_2O solution at pH 5.0 and 280 K. Cross-peaks between the amide and side-chain protons are shown. Labels indicate assignments for the structured part of the peptide (Thr87 to Cys99) for the A- (top) and B- (bottom) forms of the peptide. Vertical lines connect correlations at each of the amide proton chemical shifts.

weak sequential H^N-H^N NOEs, and the absence of any medium or long-range NOEs involving residues Leu78-Gly86 indicate that the N-terminus of the peptide is unstructured or in an extended conformation. However, within the C-terminal region (Thr87-Cys99) there are several large (>8.0 Hz) and small (<6.0 Hz) $^3J_{HN-H\alpha}$, intense sequential H^N-H^N NOEs and numerous medium- and long-range NOEs. In addition, the rates of exchange with solvent for several amide protons in this region are reduced by up to 3 orders of magnitude relative to the rates expected for small unstructured peptides. These observations indicate that the C-terminus of Guan22 does have a definable conformation, although the data are not consistent with any elements of regular secondary structure. Also worthy of note is the resonance broadening observed for certain protons in the B-form (notably Tyr93 H^N and Ala95 H^N) at temperatures below 290 K. This phenomenon is indicative of intermediate exchange, probably involving motion of the Tyr93 side chain, and suggests that although medium- and long-range NOEs are observed for the B-form of Guan22, a unique (i.e., static) model of the structure may not satisfy all of the observed data. In addition, the broadening of these resonances also reduces the number of NOEs which can be observed for the B-form of Guan22.

Comparison of Guan22 and Guan13. Earlier studies have shown definitively that the active form of guanylin contains disulfide bonds from Cys88 to Cys96 and Cys91 to Cys99 (Wiegand et al., 1992; Cuthbert et al., 1994), thus being analogous to the STa toxins (Shimonishi et al., 1987). In order to confirm that the correct disulfide bonds were being formed during the *E. coli* expression of proguanylin (from

Table 1: Chemical Shifts Obtained for Guan22 in H_2O Solution at 280 K and pH 5.0[a]

residue	H^N	H^α [b]	H^β [c]	H^γ	other
A-form					
Leu78		3.95	1.65, 1.59	1.30	0.91 ($C^\delta H_3$)
Glu79	8.80	4.29	1.86, 1.70	2.25	
Glu80	8.65	4.22	1.95, 1.86	2.25	
Ile81	8.38	4.05	1.75	1.39, 1.09	0.73 ($C^\delta H_3$), 0.82 ($C^{\gamma 2} H_3$)
Ala82	8.49	4.21	1.28		
Glu83	8.39	4.17	1.92, 1.82	2.25	
Asp84	8.53	4.80	2.71, 2.54		
Pro85		4.37	2.19, 1.96	1.93	3.75 (H^δ)
Gly86	8.52	3.86			
Thr87	7.91	4.34 (7.1)	4.29 (4.3)	1.17	
Cys88	8.74	4.79 (7.0)	**3.37, 2.92** (3.4, 10.0)		
Glu89	8.66	4.10 (6.1)	2.04	2.36	
Ile90	7.48	4.39 (10.0)	1.96 (5.6)	1.24, 1.06	0.80 ($C^\delta H_3$), 0.79 ($C^{\gamma 2} H_3$)
Cys91	8.40	4.33 (6.5)	**3.19, 2.82** (4.0, 9.5)		
Ala92	7.88	4.02 (5.5)	1.09		
Tyr93	7.16	4.66 (9.1)	3.14, 2.92 (6.1, 8.6)		7.00 (H^δ), 6.75 (H^ϵ)
Ala94	8.31	3.99 (4.3)	1.31		
Ala95	7.82	4.19 (5.6)	1.32		
Cys96	7.60	4.58 (5.9)	**3.15, 3.04** (3.3, 10.7)		
Thr97	8.58	4.19 (5.5)	4.10 (5.5)	1.18	
Gly98	8.87	3.98, 3.76			
Cys99	7.65	4.44 (8.5)	**3.05, 2.92** (3.2, 10.3)		
B-form					
Leu78		3.96	1.65, 1.59	1.30	0.91 ($C^\delta H_3$)
Glu79	8.80	4.29	1.86, 1.70	2.25	
Glu80	8.50	4.24	1.87	2.25	
Ile81	8.39	4.06	1.76	1.39, 1.10	0.78 ($C^\delta H_3$), 0.83 ($C^{\gamma 2} H_3$)
Ala82	8.52	4.24	1.29		
Glu83	8.43	4.20	1.95, 1.85	2.31	
Asp84	8.59	4.82	2.71, 2.53		
Pro85		4.39	2.22, 1.97	1.95	3.81 (H^δ)
Gly86	8.60	3.91			
Thr87	7.99	4.27	4.21 (5.0)	1.14	
Cys88	8.49	4.68 (6.8)	3.25, 2.93 (5.6, 7.9)		
Glu89	8.54	4.07 (6.0)	1.95, 1.91 (4.3, 8.9)	2.31	
Ile90	7.83	4.09 (8.1)	1.89 (7.3)	1.30, 1.10	0.78 ($C^\delta H_3$), 0.82 ($C^{\gamma 2} H_3$)
Cys91	7.89	4.66 (7.4)	3.06		
Ala92	8.41	4.05 (5.2)	1.13		
Tyr93	7.67	4.41	**2.97, 3.13** (10.2, 4.9)		7.03 (H^δ), 6.78 (H^ϵ)
Ala94	8.32	4.13 (5.2)	1.31		
Ala95	8.17	4.32 (6.8)	1.36		
Cys96	8.73	4.53 (6.2)	**3.03, 3.35** (9.0, 5.5)		
Thr97	7.63	4.31 (6.7)	4.16 (4.6)	1.11	
Gly98	8.55	3.94			
Cys99	7.84	4.41 (8.4)	**2.94, 3.26** (10.5, 3.5)		

[a] Chemical shifts are listed in ppm, referenced to the internal H_2O resonance at 4.96 ppm, and are accurate to ±0.02 ppm. [b] Values of $^3J_{HN-H\alpha}$ and $^3J_{H\alpha-H\beta}$ (Hz) are shown in parentheses in the H^α and H^β column, respectively. [c] Methylene groups for which stereospecific assignments were made (see Materials and Methods) are listed in bold, with the *pro-R* chemical shift listed first.

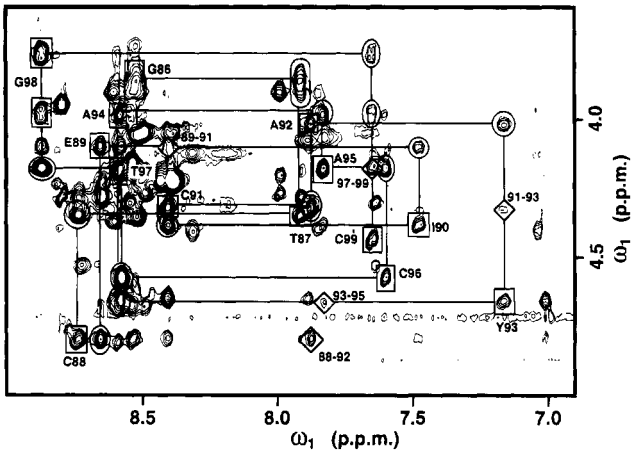

FIGURE 2: A section of the JR-NOESY spectrum (τ_m=300 ms) of Guan22 acquired from H_2O solution at pH 5.0 and 275 K. Assignments are shown for residues Gly86 to Cys99 for the A-form of the peptide. Intraresidue cross-peaks are surrounded by rectangles and labeled with the residue type and residue number. Cross-peaks arising from sequential connectivities are surrounded by ellipses. The horizontal and vertical lines indicate the pathway by which sequential assignments were made. Several cross-peaks arising from medium-range $H^\alpha-H^N$ connectivities are indicated by the diamond-shaped boxes.

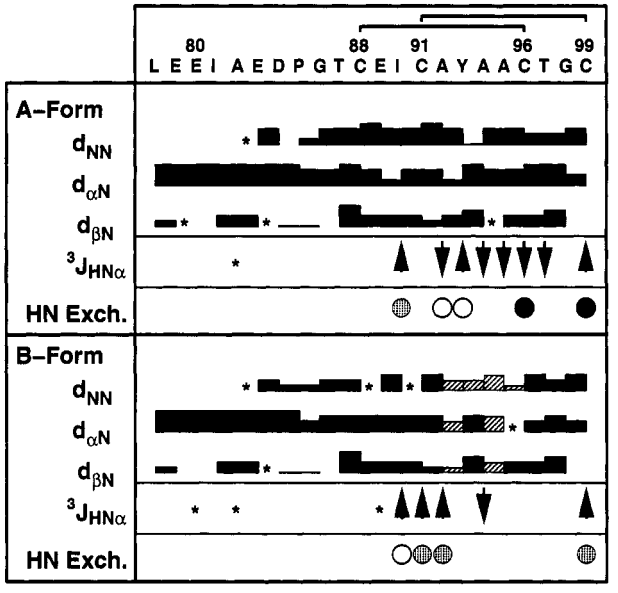

FIGURE 3: Summary of sequential NOE, amide proton exchange, $^3J_{HN-H\alpha}$, and medium-range NOE data for the A- and B-forms of Guan22. For the $d_{\alpha N}$, d_{NN}, and $d_{\beta N}$ rows, bars connect residues where the sequential NOE was observed. The thickness of the bar indicates the relative intensity of each sequential NOE, and asterisks indicate data missing because of chemical shift degeneracy. Shaded bars indicate data obtained from the Guan13 peptides because of resonance broadening in Guan22 (see text). NOEs to H^δ were used in place of H^N for Pro87. $^3J_{H\alpha-HN}$ less than 6.0 Hz or greater than 8.0 Hz are indicated by the downward or upward pointing arrows, respectively. Circles indicate backbone amide resonances exhibiting cross-peaks in a TOCSY started 4 min after dissolution of the peptide in D_2O at 280 K and pH 4.5. Protection factors greater than 10 or 100 are indicated by the shaded and filled circles, respectively.

which Guan22 was derived), synthetic peptide was also prepared. From the studies of Guan22, residues Leu78-Gly86 were known to be largely unstructured in solution; hence the chemical synthesis was simplified by producing the shorter, structured portion of the peptide from Thr87 to Cys99 (Guan13). During the synthesis, the thiol groups of Cys88 and Cys96 were blocked with the ACM protecting group, and Cys91 and Cys99 were protected with the 4-methylbenzyl protecting group. After synthesis, the peptide was cleaved from the resin and purified. The conditions used for cleavage were sufficient to remove the 4-methylbenzyl protecting groups, but not the ACM protecting groups; hence Cys91 and Cys99 could be air-oxidized to form the first disulfide bond specifically between these residues. Cys88 and Cys96 were then deprotected and the 88−96 disulfide bond formed. In this way, the synthetic peptide unequivocally contained the desired disulfide bond pattern. Such an approach has successfully been used for the study of the disulfide cross-linking in conotoxin G1 (Nishiuchi & Sakakibara, 1982). These authors found that peptides with specific disulfide bonds could be formed and that the disulfide bonds were stable at neutral or acidic pH.

NMR spectra of the synthetic peptide also contained twice the expected number of resonances, with the ratio of species again being approximately 1:1. Thus, the possibility of disulfide bond scrambling can be eliminated as the cause of the resonance doubling. Resonances of Guan13 were assigned by standard methods and the chemical shifts compared to those of the equivalent residues of Guan22. The chemical shifts of one of the Guan13 components were very similar to those of the B-form of Guan22 ($\Delta\delta$ < 0.1 ppm for H^N and H^α resonances of all residues except the H^N of Cys88 and Glu89 at the N-terminus of Guan13). The lack of chemical shift differences indicate that these resonances probably correspond to peptides having very similar conformations and that this conformation can be investigated by studying the Guan22 or Guan13 peptide. However, the low-temperature resonance broadening described above for the B-form of Guan22 was less pronounced for Guan13; since the B-form resonances of Guan13 were also less subject to problems of chemical shift degeneracy, data from this peptide were used to calculate structures for the B-form of guanylin. The more pronounced line broadening in Guan22 at a given temperature probably arises from a slower motional process relative to the frequency differences between the species involved in the exchange.

A similar comparison of Guan22 A-form and the other set of Guan13 resonances revealed differences in chemical shift ($\Delta\delta$ > 0.2 ppm) for H^N of Cys88, Glu89, Cys91 and Tyr93 and less significant differences ($\Delta\delta$ > 0.1 ppm) for six other H^N and six H^α resonances. These chemical shift differences indicate that the conformation of the A-form of Guan22 and Guan13 are subtly different, although very slight changes in backbone carbonyl or aromatic ring orientation could give rise to effects of this magnitude. Indeed, several NOEs observed between the side chain of Tyr93 and Thr87 in Guan22 were not present in the Guan13 spectra, suggesting that the time-averaged position of the aromatic ring is affected by removal of residues Leu78-Gly86. This could arise from interactions of Tyr93 with residues at the N-terminus of Guan22 that are sufficiently weak or transient that they fail to produce observable NOEs. Since differences in the two A-form conformations are probably slight, the Guan22 data were chosen as a basis for structure calculations since it contained more numerous and more intense medium- and long-range NOEs (presumably because of its longer rotational correlation time), and fewer instances of chemical shift degeneracy than the Guan13 A-form data.

Table 2: Summary of Input Restraints for the Structure Calculations and the Violations, RMSDs, and Energies of the Final 20 Structures for the A- and B-Forms of Guanylin

	A-form	B-form
input to structure calculations		
$^1H-^1H$ distance restraints[a]		
total	78	53
intraresidue	16	18
sequential	27	18
medium range	15	6
long range	20	11
dihedral angle restraints		
ϕ[b]	9, 1, 2	4, 1, 7
χ_1	5	4
stereospecific assignments	C88, C91, C96, C99	C88, Y93, C96, C99
total restraints/residue	7.3	5.3
results of structure calculations[c]		
energy (kcal·mol^{-1})		
total	-88.3 ± 2.1	-106.0 ± 4.8
nonbonded	-19.0 ± 1.8	-16.6 ± 0.7
restraint violation	0.12 ± 0.05	0.19 ± 0.09
NOE violations		
number > 0.0 Å	3.75 ± 0.94	2.10 ± 1.1
number > 0.1 Å	0.0	0.0
maximum (Å)	0.05 ± 0.01	0.06 ± 0.03
sum (Å)	0.11 ± 0.03	0.09 ± 0.05
dihedral angle violations		
number > 0.5°	0.95 ± 0.50	1.15 ± 0.35
sum (deg)	1.00 ± 0.76	2.2 ± 1.6
RMSD from mean structure		
backbone, 88–99	0.47 ± 0.11	1.07 ± 0.27
heavy atoms, 88–99	0.87 ± 0.14	1.39 ± 0.30

[a] NOE restraints are only included if they are deemed structurally useful, i.e., if the corresponding upper bound is shorter than that allowed by covalent geometry; *all* intraresidue NOEs and sequential NOEs to H^N were checked to see if they met this criterion. [b] The first value is the number of restraints derived from 8.0 Hz < $^3J_{HN-H\alpha}$ < 6.0 Hz; the second value is the number of positive ϕ restraints; the final value is the number of negative ϕ restraints (see Materials and Methods). [c] Values are means and standard deviations for 20 structures.

Structure Calculations. Owing to the lack of structural restraints for the N-terminus of the Guan22 peptide, structures were calculated for residue Thr87 to Cys99 only. In addition to the NMR-derived restraints (Table 2), the disulfide bonds were explicitly included from Cys88 to Cys96 and from Cys91 to Cys99; specific disulfide bond angle restraints were not employed. In the final calculations of the A-form, input data consisted of 78 structurally useful NOE-derived distance restraints, 45% of which were between protons in nonadjacent residues. In the final calculations of the B-form, input data consisted of 53 NOE-derived distance restraints, 32% of which were between protons in nonadjacent residues.

The resulting structures for both the A and B form of guanylin satisfied the input data very well, with no distance restraint violations greater than 0.1 Å and no dihedral angle restraint violations greater than 2° (Table 2). As expected from the number and distribution of restraints used as input to the calculations, the A-form structures are globally better defined (mean backbone atom RMSD from the average structure 0.47 Å) than the B-form structures (mean RMSD from the average structure 1.07 Å). $S^{ang}(\phi,\psi)$ also indicate that the local geometry is better defined in the A-form (Figure 4), with mean $S^{ang}(\phi,\psi)$ values of 0.94 and 0.86 for the A- and B-forms, respectively. The ensemble of 20 structures each for the A- and B-forms of guanylin have been deposited with the Brookhaven Protein Data Bank, Accession Numbers 1GNA and 1GNB, respectively.

Structural Comparison of the A- and B-Form of Guanylin. Analysis of the final 20 structures calculated for the A-form of guanylin indicates the presence of two reverse turns from Cys88 to Cys91 and from Tyr93 to Cys96 (Figure 5A,B), both of which are of type I (Wilmot & Thornton, 1988). At the C-terminus of the peptide, there is a third reversal of the chain direction encompassing residues Cys96 to Cys99 (Figure 5C). In 2 of the 20 structures, the backbone dihedral angles and close separation of Cys99 H^N and Cys96 O indicate a type II reverse turn. However, in the other 18 structures, the ψ angle of Thr97 is outside the range expected for a type II turn (actual angle = $-170° \pm 7°$; ideal angle

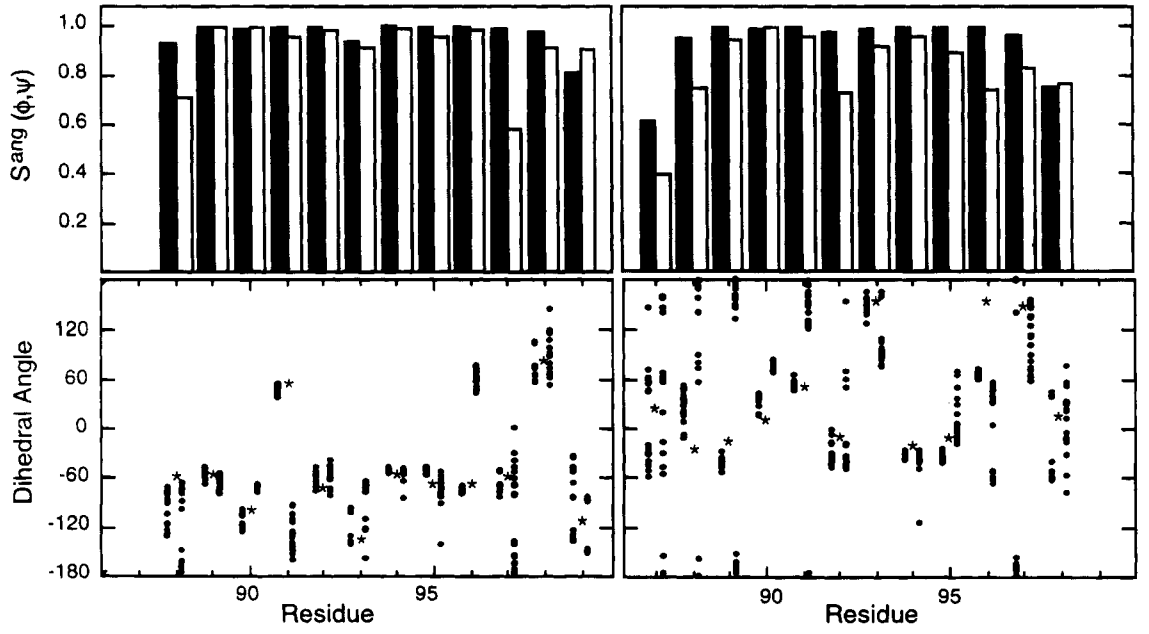

FIGURE 4: Angular order parameters and backbone dihedral angles for guanylin and STp. The bars in the upper panels indicate $S^{ang}(\phi)$ and $S^{ang}(\psi)$ on the left and right, respectively, for the final 20 structures of the A-form (solid bars) and B-form (open bars) of guanylin. The lower panels indicate the ϕ (left) and ψ (right) dihedral angles in the families of solution structures and the crystal structure of STp. For each column, data for STp are indicated by stars, while for clarity the data points for the A and B structures are shifted slightly to the left and right, respectively.

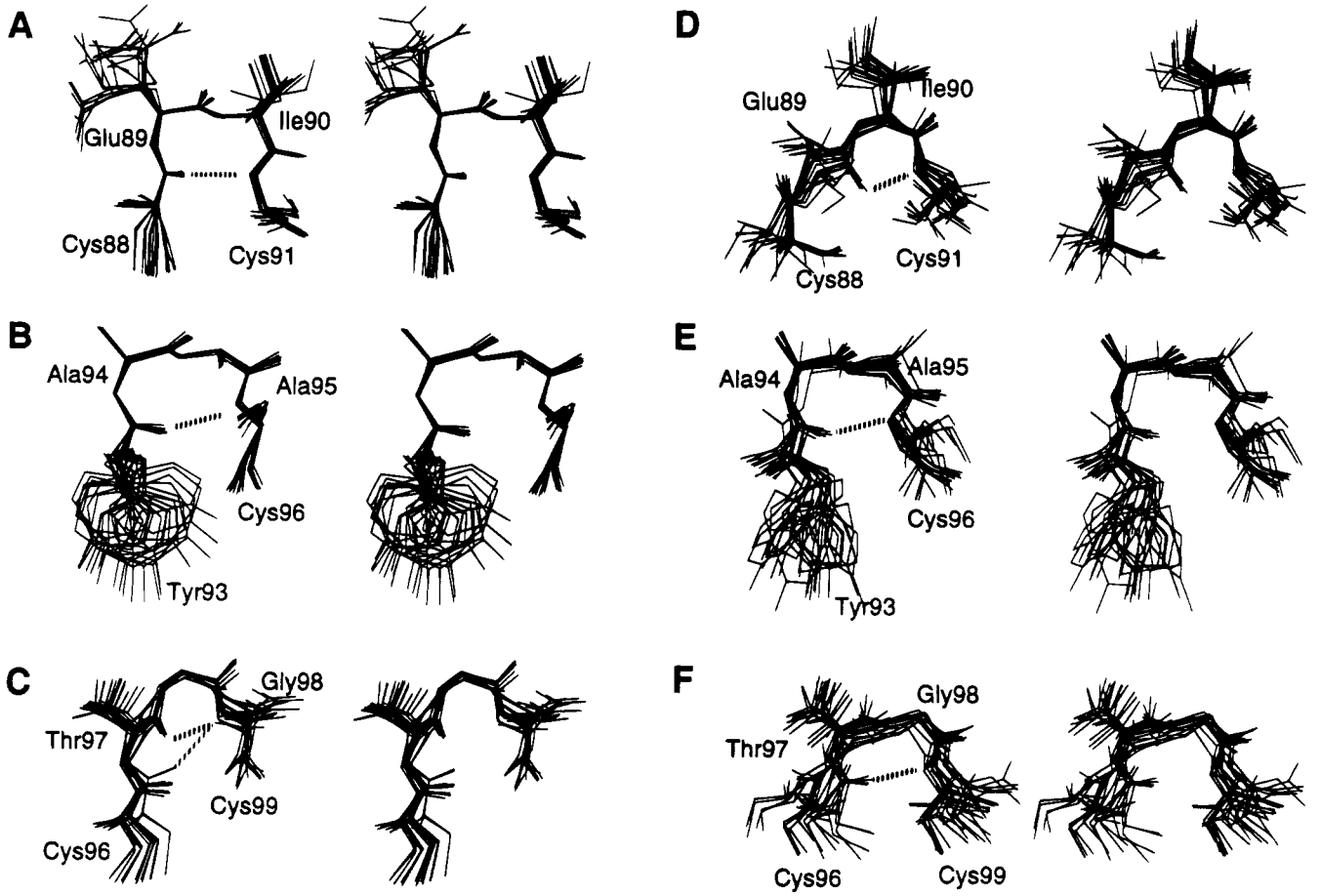

FIGURE 5: Stereoviews of the reverse turns present in the A-form (A−C) and B-form (D−F) structures of guanylin. Solid lines connect heavy backbone atoms of 20 structures; the dashed lines connect H^N and backbone O atoms close enough to be involved in a hydrogen bond interaction in at least some of the structures.

= 120°) and the separation of Cys96 O and Cys99 H^N is 3.75 ± 0.48 Å. The observation of a short separation of Thr97 O and Cys99 HN (1.90 ± 0.02 Å) indicates that a γ-turn is present in 13 structures. The remaining 5 structures have Cys99 H^N pointing away from the rest of the peptide. The lack of definition of the structures in this region most likely results from a scarcity of restraints, due either to fast motional processes or to the particular geometric arrangement of the protons. The calculation protocol may also be contributing to the observed heterogeneity of the structures: the lack of explicit water molecules in the final minimization may lead to an unrealistic number of intramolecular hydrogen bonds. Although the calculations were conducted in such a way as to minimize these effects (i.e., by using a distance dependent dielectric constant and reducing the total charge on the termini and charged side chains), hydrogen bonds not actually present in solution may be energetically favorable in the calculations if there are not enough experimental restraints to define a unique orientation.

The complete fold of the A-form of guanylin is best described as three turns arranged in a right-handed spiral (Figure 6A,B). The core of the structure is composed of the cysteine side chains, with the backbone wrapped around these atoms, and the other side chains pointing out into solution. The lack of a hydrophobic core suggests that the disulfide bonds are essential to maintain the fold of the peptide. Although all four cysteine χ_1 angles are well defined ($S^{ang}(\chi_1) > 0.99$), variations in χ_2 and χ_3 cause both disulfides to adopt two conformations; for both disulfides, approximately half of the structures have disulfide dihedral angles similar to those of a left handed spiral, the remaining structures have nonstandard geometries. Thus, the restraints are unable to define unique disulfide bond geometries, and fast exchange between the calculated conformations (and possibly other conformations) cannot be ruled out. Aside from the cysteines, the side chains of Thr87, Ile90, and Thr97 are reasonably well defined ($S^{ang}(\chi_1) > 0.92$), as expected from the nonaveraged nature of $^3J_{H\alpha-H\beta}$ (Table I). Tyr93 and Glu89 have the least well defined side chains ($S^{ang}(\chi_1)$ 0.86 and 0.37, respectively).

In the structures of the B-form of guanylin, the backbone is found to contain turns in the same general regions as the A-form. Due to the less precise nature of these structures, the turns are not easily categorized. The first is present between residues 89 and 91; Cys91 H^N and Glu89 O are within hydrogen-bonding distance (1.91 ± 0.02 Å) in 19 of the structures (Figure 5D) although the ϕ,ψ angles are not consistent with a normal γ-turn. A type I reverse turn is observed in 16 of the structures between Tyr93 and Cys96 (Figure 5E). Finally, six of the structures contain a type II reverse turn at the C-terminus, three others contain a γ-turn (Figure 5F), and the remainder do not adopt conformations consistent with regular reverse turns. As with the A-form structures, the core of the B-form is composed of the cysteine side chains, with the other side chains pointing into the solvent. The side-chain orientations of Tyr93, Cys96, Thr97, and Cys99 are all well defined ($S^{ang}(\chi_1) > 0.95$), in keeping with the observed values of $^3J_{H\alpha-H\beta}$. Other side chains are

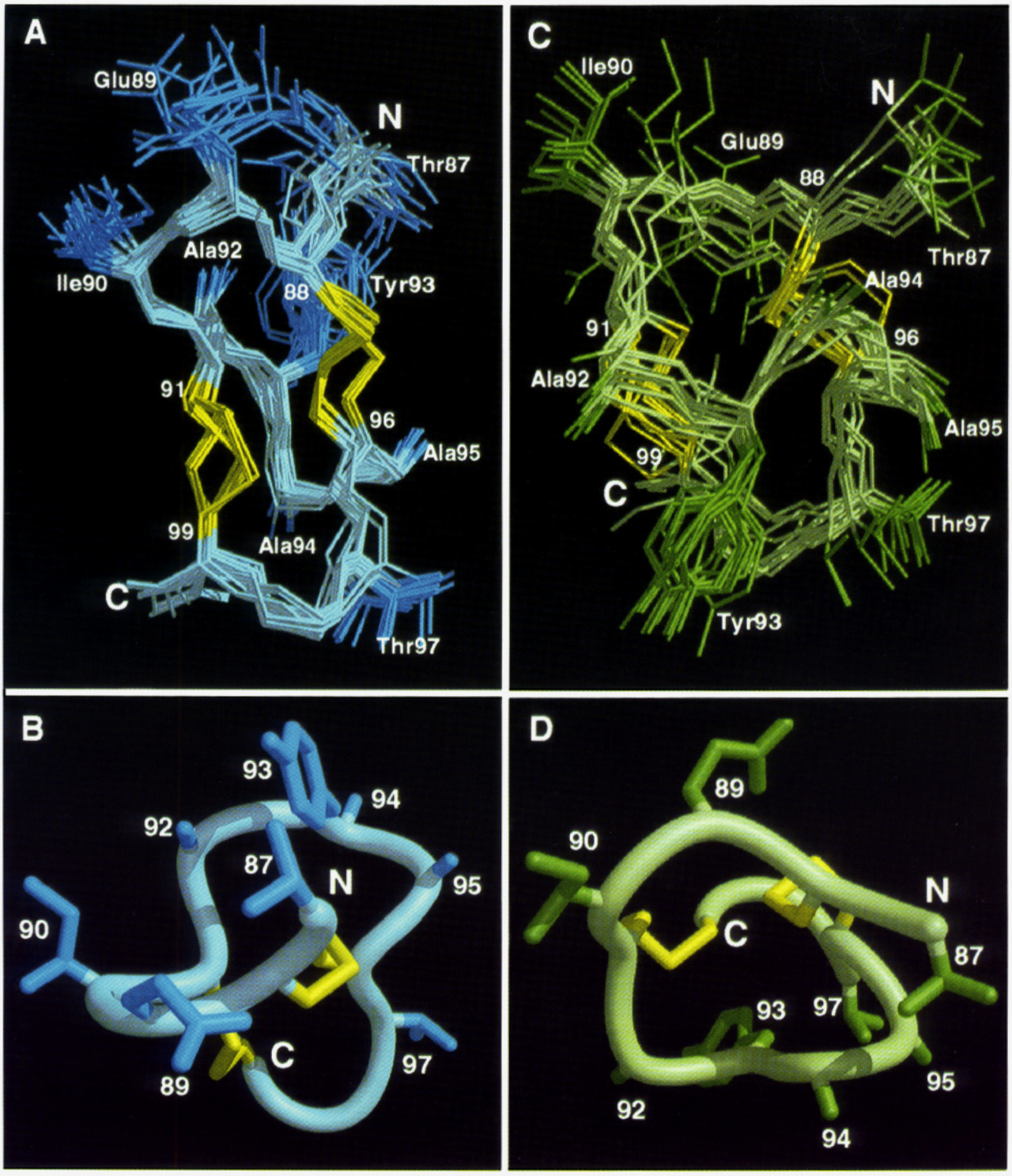

FIGURE 6: Solution structures of the A-form (A, B), and the B-form (C, D) of guanylin determined from NMR data. Panels A and C include all heavy atoms (except carbonyl oxygen) for residues 87 to 99 of the final structures of each form. For clarity, 4 of the 20 structures are omitted from C. The C^{β} and S^{γ} atoms of cysteine residues are colored yellow. In panels B and D a single representative structure of each form is depicted (see Materials and Methods) with the backbone and side chains represented as a tube and sticks, respectively. These views were obtained by a 90° rotation about a horizontal axis relative to A and C. In all panels, side-chain atoms are a slightly darker shade of blue (A-form) or green (B-form). Labels indicate side-chain positions.

not well defined and probably sample a number of different conformations in solution.

During the analysis of the B-form resonances, several low-intensity NOEs were observed between the aromatic ring protons of Tyr93 and the methyl groups of Ala94 and Ala95. Inclusion of these NOEs in the structure calculations increased the residual violations of the restraints and caused Tyr93 to adopt an eclipsed conformation with $\chi_1 \approx -120°$. This conformation is both energetically unlikely and also in conflict with $^3J_{H\alpha-H\beta}$ (Table 1), the latter indicating a χ_1 of 180° or −60°. The structures presented here for the B-form of guanylin were calculated without these NOEs. Tyr93 adopts a noneclipsed conformation ($\chi_1 = -64° \pm 9°$) and the restraint violations are reduced. The side chain of Tyr93 points away from Ala94 and Ala95 in these structures, and the long distances between the ring and methyl protons (>5 Å) would not be expected to yield the observed NOEs. These observations suggest that there is some dynamic

process occurring in this region of the B-form of guanylin. Such a conclusion was also drawn from the temperature dependent line-broadening effect observed for Tyr93 H^N and Ala95 H^N. Attempts to interpret the time-averaged NOEs in terms of a single static conformation results in strained and distorted structures with violations of the input restraint. Only by ignoring three of the NOEs can a family of converged structures be obtained. The dynamic process appears to be localized to the region around Tyr93 as there are no other inconsistencies between the NMR data and the calculated structures. However, it should be pointed out that the structures described here only represent one of the areas of conformational space accessible to this form of the peptide.

Although the structure of the B-form of guanylin is not as well defined as the A-form, the available NMR data are sufficient to define a global fold that is quite distinct from that of the A-form. The RMSD of backbone atoms in residues 88 to 99 between the mean A-form and mean B-form structure is 4.48 Å, much higher than the uncertainty in the pairwise RMSDs within each family of structures. The B-form structure is best described as three turns connected in a left-hand spiral (Figure 6C,D). Thus, although the A and B forms of the peptide have exactly the same covalent bonds, they differ in the handedness of the global fold. Panels B and D of Figure 6 emphasize the difference in spiral handedness of the two forms, whilst panels A and C indicate how the Cys91-Cys96 loop is below the plane formed by the disulfide bonds in the A-form but above the plane in the B-form. As a result of the different folds, the side chain of Tyr93 extends toward the N-terminus in the A-form (in the A-form NOEs are observed between Tyr93 and Thr87) but toward the C-terminus in the B-form (in the B-form NOEs are observed between Tyr93 and Cys99). The different locations of the aromatic ring in the two peptides no doubt contributes to the marked differences in backbone chemical shifts of the two species.

Examination of the lower panels in Figure 4 indicate that only three backbone dihedral angles vary markedly between the A- and B-forms of guanylin. Two of the large changes occur at residues between the reverse turns (Cys91 ϕ(A) = 47° ± 5°, ϕ(B) = −132° ± 15°; Cys96 ϕ(A) = −74° ± 2°, ϕ(B) = 62° ± 8°). In this way, the overall fold of the A- and B-forms of guanylin are quite different, even though the three turns are found at similar positions in the primary sequence. Thus, in order to fold up guanylin and form both disulfide bonds, the ϕ angle of either Cys91 *or* Cys96 must adopt a relatively high energy positive value, with the two possibilities leading to the two observed backbone folds. The third large change in backbone dihedral angle (Glu89 ψ(A) = −42° ± 5°, ψ(B) = 174° ± 4°) results in the change from a type I to a γ turn between Cys88 and Cys91 in the A- and B-forms, respectively (see above).

There is no evidence of exchange between the two guanylin species present in solution. Heating the sample up to 340 K failed to induce any line broadening, as shown by signals from the two overlapped Tyr93 H^δ doublets, whose four lines were still clearly observable at this temperature ($\Delta\delta$ approximately 4.0 Hz). These observations suggest that if the A- and B-forms of guanylin do interconvert, the exchange must be very slow with a half-life of seconds or longer even at 340 K. From transition state theory, exchange this slow indicates an activation barrier in excess of 20 kcal·mol^{-1}. Figure 6A,C suggests why the energetic penalty of such an interconversion is high: effectively, the loop containing Ala92-Ala95 (including the bulky tyrosine side chain) must be pulled through the larger loop in the plane of these figures. Restrained molecular dynamics and template forcing calculations lend support to the high energetic barrier to this process. Starting with representative A- or B-form structures, numerous attempts were made to force an interconversion between the two forms; transition from one state to the other could not be induced with a realistic force field (R. McDowell, personal communication).

DISCUSSION

Comparison with the STp Enterotoxin Structure. STa peptides have been studied previously by NMR (Gariépy et al., 1986; Ohkubo et al., 1986). Model structures were proposed based on the observed NOEs, although the definition of the structures were limited by uncertainty over the cysteine pairing within the disulfide bonds. Interestingly, in aqueous solution many of the resonances were reported to exhibit temperature-dependent broadening that was interpreted in terms of conformational heterogeneity with intermediate to fast interconversion between conformers (Gariépy et al., 1986) . However, there was no evidence of two conformations in very slow exchange, as found in the present work on guanylin.

The best defined structural information has been obtained from crystallographic analysis of a derivative of STp (Ozaki et al., 1991). The pattern of disulfides determined by chemical means (Gariépy et al., 1987; Shimonishi et al., 1987) was confirmed by the crystal structure analysis. The structure was composed of three reverse turns; using the alignment of Figure 7A and the numbering scheme of guanylin, the turns occur between residues 88−91, 93−96 and 96−99, i.e., in the same places that turns were observed in guanylin (Figure 5). Moreover, the turns were arranged in a right-handed spiral comparable to that found in the A-form of guanylin. The similarity in structure between STp and the A-form of guanylin is borne out by the 1.44 Å backbone atom RMSD from the mean guanylin A structure; in contrast, the equivalent RMSD for the B-form of guanylin is 4.67 Å.

The structures of STp and guanylin are most similar in the region from residue 90 to 96 (Figure 7B), with the atomic displacements of backbone atoms (using the mean A-form structure of guanylin) all less than 1.3 Å for these residues. Previous mutational studies have implicated residues in the 92−95 loop, especially residue 93, as being important for the receptor binding and enterotoxic activity of STa (Okamoto et al., 1988; Carpick & Gariépy, 1991; de Sauvage et al., 1991). Thus, the backbone similarity of STp and the A-form of guanylin in this region may be considered as confirmatory evidence of their similar mode of action. However, interactions of guanylin with intestinal cells or with cell lines transfected with STaR have IC_{50} values at least 100 times higher than STa peptides (Carpick & Gariépy, 1993; Garcia et al., 1993). Even a guanylin mutant with residues Tyr93 and Ala94 replaced with the amino acids found in STa (asparagine and proline, respectively) does not have IC_{50} values comparable to those of STa (Carpick & Gariépy, 1993). Thus, similarity in the backbone conformation of residues 90 to 96, or the correct side chains at positions 93 and 94 are not sufficient for optimal receptor

A

```
                  80                  88    91        96    99
Guan22    L E E I A E D P G T C E I C A Y A A C T G C
Guan13                      T C E I C A Y A A C T G C
STa             N S S N Y C C E L C C N P A C T G C Y
STp                         C C E L C C N P A C A G C
```

B

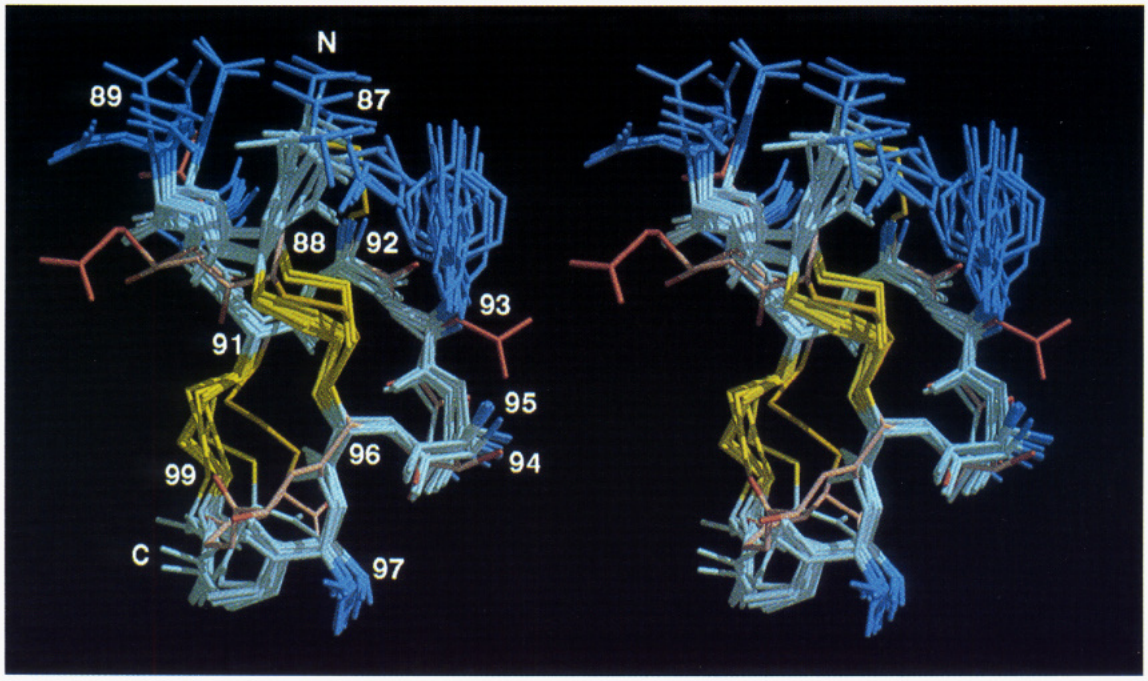

FIGURE 7: (A) Primary sequences of guanylin and ST peptides discussed in the text. (B) Stereoviews of overlaid structures of STp (pink and red) and the A-form of guanylin (cyan) based on the heavy backbone atoms of residues 88 to 99. In both cases, the C^{β} and S^{γ} atoms of cysteine are colored yellow. Coordinates for STp were taken from Ozaki et al. (1991). Residue numbers of non-cysteine side chains are included.

interaction. Interestingly, the loop from residue 91 to 96 of both forms of guanylin is similar to that observed in STp: all three structures contain a type I reverse turn and backbone atom RMSDs from STp to the mean A and B-form structures of guanylin are 0.68 and 1.11 Å, respectively. The possible significance of the relatively close fit to the B-form loop is discussed in more detail below.

The differences in activity of guanylin and STa may be the result of the large difference in backbone conformation of residues 89 and 97 to 99: the displacements of the overlaid structures are as large as 3.0 Å for some atoms. The differences in the orientation of the C-terminal turn primarily result from changes in the backbone ψ angle of Cys96 (Figure 4; guanylin A-form = 69° ± 4°, STp = 149°). However, aside from the changes in amino acid sequence, the apparent differences in the conformations of guanylin (A-form) and STp may also arise because of crystal packing influences in the solid state or because of the poor definition of the C-terminus in solution (see above). In addition, the relationship of the free peptides to the receptor-bound state is unknown at this time. Thus, more structure−function analysis of mutant peptides and structural studies of bound peptides will be necessary for a complete understanding of the functionally important features of this interaction.

Topological Stereoisomerism in Disulfide-Bonded Peptides. "Topological stereoisomerism" has been defined as the existence of multiple possibilities for the molecular topology of a nonplanar polypeptide chain (Mao, 1989). In this respect, "planar" means that the structure may be drawn schematically on a flat surface with no bonds intersecting, and from a theoretical point of view, all peptides containing three or fewer disulfide bonds are planar. Such isomerism results in molecules that contain exactly the same chemical connectivities but differ in the relative spatial arrangement of the loops and disulfide bonds.

Guanylin contains only two disulfide bonds and can be drawn schematically in planar form (Figure 8A). However, this is not a realistic depiction as the loops between the cysteine residues are of unequal length, and the longer loop between residue 91 and 96 must bulge above or below the plane of Figure 8A (Figure 8B,C). Thus, in real peptides where steric bulk has a marked impact on conformation, the planarity may be compromised and a form of topological stereoisomerism may be observed even when there are only two disulfide bonds present. The structures calculated above indicate that isomerism of this kind is occurring for guanylin, giving rise to the two distinct sets of NMR resonances. Since there is no evidence of interconversion between the forms,

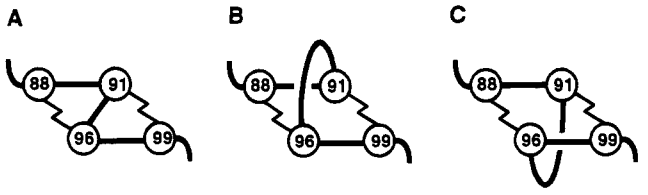

FIGURE 8: Schematic views of the backbone fold and cross-linking of guanylin in planar format (A) and in more realistic views in which the central loop is above (B) or below (C) the plane of the four cysteine residues. The numbered circles represent the cysteine residues, thick lines the backbone chain tracing, and thin jagged lines the disulfide bridges. Panel B actually corresponds to the A-form of guanylin, while panel C corresponds to the B-form.

the observed 1:1 ratio cannot necessarily be related to the thermodynamic stability of each state but may be the result of kinetic trapping during the formation of the disulfide bonds.

To our knowledge, this type of isomerism has not been encountered in other small disulfide linked peptides. Simple air oxidation of chemically synthesized peptides related to apamin (4 cysteines) and STa (6 cysteines) yield predominantly a single covalent and geometric isomer (Gariépy et al., 1986; Kubota et al., 1989; Pease et al., 1990). Presumably, the specific spacing of the cysteines, or the nascent structure prior to cross-linking, is such that extensive covalent or geometric isomerization is not encountered in these cases. STa and STp contain a third disulfide bond from residue 87 to 92. The 87 C^{α}–92 C^{α} distance in the crystal structure of STp is 5.4 Å and in the A-form of guanylin is 6.9 ± 0.6 Å; hence this disulfide bond could be accommodated in this form of guanylin without significant rearrangement. However, the equivalent distance in the B-form of guanylin is 11.3 ± 1.2 Å, with the side chain of Ala92 pointing away from that of Thr87 (Figure 7C,D). Thus, in all probability, the STa peptides cannot attain a "B-form" conformation without disruption of the disulfide bonds. This may account for the lack of topological stereoisomerism observed for the STa peptides. However, others have noted that, during distance geometry calculations of STa, some low-energy "mirror image" conformations are observed (Glunt et al., 1994; J. G. Shelling and D. J. Ward, personal communication) and that conformational equilibria involving these states may be occurring in solutions of this peptide.

In larger protein systems containing an equivalent pattern of disulfide bonds, larger loops between cysteines may conceivably allow a fast interconversion between the two possible geometric isomers. In this case, the tertiary contacts stabilizing the protein are likely to favor one isomer over the other, leading to the observation of a single conformational state. The topological stereoisomerization observed for guanylin is also distinct from the conformational heterogeneity observed in a mutant of BPTI (Otting et al., 1993). In the BPTI case, the conformational change involves a difference in χ_1 of one cysteine, resulting in a change from a left- to a right-handed chirality for the disulfide bond. This interconversion is fast, and exchange peaks are observed in short mixing time ROESY and NOESY spectra acquired at room temperature.

Implications of the Present Results. These studies performed on guanylin have shown that in addition to the three chemically distinct isomers of a peptide containing four cysteines and two disulfide bonds, one of the geometric isomers may also exhibit a form of topological stereoisomerism. Peptides prepared by chemical synthesis and bacterial expression exhibit this isomerism and contain equivalent ratios of isomers. The bacterial expression directed secretion into the periplasmic space, hence the polypeptide is folded in a manner analogous to the mammalian molecule produced *in vivo* (provided that no other chaperonin type proteins are involved with the *in vivo* folding). The observation of two species in material produced in this way suggests that either the two forms are present in the proprotein or that exchange (albeit very slow) can occur once the mature hormone is cleaved enzymatically from the remainder of the proprotein. Structural studies on proguanylin are underway to shed light on this matter.

The structure of the A- and B-forms of guanylin are sufficiently different that they undoubtedly interact differently with STaR. Assuming that the conformation of STp observed in the crystal structure represents an active conformation, the similarity of the A-form to the crystal structure suggests that this form of guanylin is capable of binding to STaR, thereby producing an increase in cGMP levels. This hypothesis is strengthened by the inability of STa to obtain a conformation similar to the B-form (see above), thereby discounting a radical conformational change to an active B-form like conformation in order to bind to the receptor. Thus, the possible biological role of the B-form of guanylin is more speculative but may include interactions with different receptors. Alternatively, if the A- and B-forms are in very slow exchange, binding to the STa receptor may have a biphasic binding profile, with a fast initial phase from the hormone already in the A-form and a slow second phase resulting from the isomerization to the A-form prior to binding. One caveat to these hypotheses is that the 91–96 loop of the B-form of guanylin is similar to that of the STp crystal structure; hence this form of the peptide may have intrinsic binding to STaR in its own right.

The observation of this type of isomerization has important consequences for the rational design of peptides capable of specific interactions with larger proteins. Disulfide bonds are often used to limit the conformations accessible to the backbone of the peptide in the hope that the resulting structure will have a higher affinity for a target receptor or enzyme. The design process is dependent on the introduction of cross-links to affect the structure in a specific and predictable manner. Formation of a topological isomer at the cross-linking stage raises the possibility of forcing the peptide into an unexpected and unwanted area of conformational space. The absence of activity of such an isomer may lead the researcher to erroneously assume that the designed area of conformational space has been explored and is not important for activity when in actuality, only the isomeric conformation has been probed. Such a conclusion will not make the difficult job of rational design any easier. Clearly, the possibility of forming topological isomers must be considered when more than one disulfide bond is introduced. Moreover, structural studies are warranted throughout the design and synthesis process to ensure that the correct peptide topology is being obtained.

ACKNOWLEDGMENT

We thank Dr. Jim Bourell for performing mass spectrometric analyses and attempting to chromatographically

separate the two isoforms, Dr. Frederic de Sauvage for performing the tryptic cleavage of proguanylin, Dr. Bob McDowell for attempting to interconvert the A and B isoforms during molecular dynamics simulations, and Dr. Wayne Fairbrother for numerous helpful discussions during the calculation of the structures. We also thank Drs. David Ward and Judith Shelling for providing details of their STa calculations prior to publication.

REFERENCES

Akke, M., Skelton, N. J., Kördel, J., & Chazin, W. J. (1991) in *Techniques in Protein Chemistry II* (Villafranca, Ed.) pp 401−408, Academic Press, Inc., Boca Raton, FL.

Aue, W. P., Bartholdi, E., & Ernst, R. R. (1975) *J. Chem. Phys. 64*, 2229−2246.

Bax, A., & Freeman, R. (1981) *J. Magn. Reson. 44*, 542 561.

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

Bodenhausen, G., Kogler, H., & Ernst, R. R. (1984) *J. Magn. Reson. 58*, 370−388.

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

Braunschweiler, L., Bodenhausen, G., & Ernst, R. R. (1984) *Mol. Phys. 48*, 535−560.

Carpick, B. W., & Gariépy, J. (1991) *Biochemistry 30*, 4803−4809.

Carpick, B. W., & Gariépy, J. (1993) *Infect. Immun. 61*, 4710−4715.

Cavanagh, J., & Rance, M. (1992) *J. Magn. Reson. 96*, 670−678.

Chan, S. K., & Giannella, R. A. (1981) *J. Biol. Chem. 256*, 7744−7746.

Clubb, R. T., Fergusen, S. B., Walsh, C. T., & Wagner, G. (1994) *Biochemistry 33*, 2761−2772.

Currie, M. G., Fok, K. F., Kato, J., Moore, R. J., Hamra, F. K., Duffin, K. L., & Smith, C. E. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 947−951.

Cuthbert, A. W., Hickman, M. E., MacVinish, L. J., Evans, M. J., Colledge, W. H., Ratcliff, R., Seale, P. W., & Humphrey, P. P. A. (1994) *Br. J. Pharmacol. 112*, 31−36.

de Sauvage, F. J., Camereto, T. R., & Goeddel, D. V. (1991) *J. Biol. Chem. 266*, 17912−17918.

de Sauvage, F. J., Keshav, S., Kuang, W. J., Gillett, N., Henzel, W., & Goeddel, D. V. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 9089−9093.

Englander, J. J., Calhoun, D. B., & Englander, S. W. (1979) *Anal. Biochem. 92*, 517−524.

Ferrin, T. E., Huang, C. C., Jarvis, L. E., & Langridge, R. (1988) *J. Mol. Graphics 6*, 13−27.

Field, M., Graf, L. H., Laird, W. J., & Smith, P. L. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 2800−2804.

Forte, L. R., Eber, S. L., Turner, J. T., Freeman, R. H., Fok, K. F., & Currie, M. G. (1993) *J. Clin. Invest. 91*, 2423−2428.

Garcia, K. C., de Sauvage, F. J., Struble, M., Henzel, W., Reilly, D., & Goeddel, D. V. (1993) *J. Biol. Chem. 268*, 22397−22401.

Gariépy, J., Lane, A., Frayman, F., Wilbur, D., Robien, W., Schoolnik, G. K., & Jardetzky, O. (1986) *Biochemistry 25*, 7854−7866.

Gariépy, J., Judd, A. K., & Schoolnik, G. K. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8907−8911.

Giannella, R. A., & Drake, K. W. (1979) *Infect. Immun. 24*, 19−23.

Glunt, W., Hayden, T. L., Wells, C., Shelling, J. G., & Ward, D. J. (1994) *J. Math. Chem. 15*, 353−366.

Gordon, J. E. (1971) *Ann. N.Y. Acad. Sci. 176*, 9−15.

Havel, T. F. (1991) *Prog. Biophys. Mol. Biol. 56*, 43−78.

Hyberts, S., Goldberg, M. S., Havel, T. F., & Wagner, G. (1992) *Protein Sci. 1*, 736−751.

Kim, Y., & Prestegard, J. H. (1990) *J. Magn. Reson. 84*, 9−13.

Koning, T. M. G., Boelens, R., & Kaptein, R. (1990) *J. Magn. Reson. 90*, 111−123.

Kubota, H., Hidaka, Y., Ozaki, H., Ito, H., Hirayama, T., Takeda, Y., & Shimonishi, Y. (1989) *Biochem. Biophys. Res. Commun. 161*, 229−235.

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

Levine, M. M., Caplan, E. S., Waterman, D., Cash, R. A., Hornick, R. B., & Snyder, M. J. (1977) *Infect. Immun. 17*, 78−82.

Ludvigsen, S., & Poulsen, F. M. (1992) *J. Biomol. NMR 2*, 227−233.

Mao, B. (1989) *J. Am. Chem. Soc. 111*, 6132−6136.

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

Merrifield, R. B. (1963) *J. Am. Chem. Soc. 85*, 2149−2154.

Nishiuchi, Y., & Sakakibara, S. (1982) *FEBS Lett. 128*, 260−262.

Ohkubo, T., Kobayashi, Y., Shimonishi, Y., Kyogoku, Y., Braun, W., & Gō, N. (1986) *Biopolymers 25*, S123−S134.

Okamoto, K., Okamoto, K., Yukitake, J., & Miyama, A. (1988) *Infect. Immun. 56*, 2144−2148.

Otting, G., Widmer, H., Wagner, G., & Wüthrich, K. (1986) *J. Magn. Reson. 66*, 187−193.

Otting, G., Liepinsh, E., & Wüthrich, K. (1993) *Biochemistry 32*, 3571−3582.

Ozaki, H., Sato, T., Kubota, H., Hata, Y., Katsube, Y., & Shimonishi, Y. (1991) *J. Biol. Chem. 266*, 5934−5941.

Pease, J. H., Storrs, R. W., & Wemmer, D. E. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 5643−5647.

Plateau, P., & Guéron, M. (1982) *J. Am. Chem. Soc. 104*, 7310−7311.

Rance, M., & Wright, P. E. (1986) *J. Magn. Reson. 66*, 372−378.

Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479−485.

Rao, M. C., Orellana, S. A., Field, M., Robertson, D. C., & Giannella, R. A. (1981) *Infect. Immun. 33*, 165−170.

Sack, R. B. (1980) *J. Infect. Dis. 142*, 279−286.

Schulz, S., Green, C. K., Yuen, P. S., & Garbers, D. L. (1990) *Cell 63*, 941−948.

Schulz, S., Chrisman, T. D., & Garbers, D. L. (1992) *J. Biol. Chem. 267*, 16019−16021.

Shimonishi, Y., Hidaka, Y., Koizumi, M., Hane, M., Aimoto, S., Takeda, T., Miwatani, T., & Takeda, Y. (1987) *FEBS Lett. 215*, 165−170.

Sutcliffe, M. J. (1993) *Protein Sci. 2*, 936−944.

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

Wiegand, R. C., Kato, J., Huang, M. D., Fok, K. F., Kachur, J. F., & Currie, M. G. (1992) *FEBS Lett. 311*, 150−154.

Wilmot, C. M., & Thornton, J. M. (1988) *J. Mol. Biol. 203*, 221−232.

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.

Yip, P. (1990) *J. Magn. Reson. 90*, 382−383.

Yoshimura, S., Ikemura, H., Watanabe, H., Aimoto, S., Shimonishi, Y., Hara, S., Takeda, T., Miwatani, T., & Takeda, Y. (1985) *FEBS Lett. 181*, 138−143.