E., Dayton, C., & Cardarelli, P. M. (1986) *Am. J. Med. 80*, 229-240.
Sekiguchi, K., Hakomori, S., Funahashi, M., Matsumoto, I., & Seno, N. (1983) *J. Biol. Chem. 258*, 14359-14365.
Skorstengaard, K., Jensen, M. S., Sahl, P., Petersen, T. E., & Magnusson, S. (1986) *Eur. J. Biochem. 161*, 441-453.
Stathakis, N. E., & Mosesson, M. W. (1977) *J. Clin. Invest. 60*, 855-865.
Sturtevant, J. M. (1974) *Annu. Rev. Biophys. Bioeng. 3*, 33-51.
Thomas, L., Smith, R. T., & VonKorff, R. W. (1954) *Proc. Soc. Exp. Biol. Med. 86*, 813-819.
Wallace, D. G., Donovan, J. W., Schneider, P. M., Meunier, A. M., & Lundblad, J. L. (1981) *Arch. Biochem. Biophys. 212*, 515-524.
Yamada, K. M. (1983) *Annu. Rev. Biochem. 52*, 761-799.
Yamada, K. M., & Olden, K. (1978) *Nature (London) 275*, 179-194.
Yamada, K. M., Kennedy, D. W., Kimata, K., & Pratt, R. M. (1980) *J. Biol. Chem. 255*, 6055-6063.

NMR Study of the Solution Conformation of Rat Atrial Natriuretic Factor 7-23 in Sodium Dodecyl Sulfate Micelles

E. T. Olejniczak,* R. T. Gampe, Jr., T. W. Rockway, and S. W. Fesik*
Pharmaceutical Discovery Division, Abbott Laboratories, Abbott Park, Illinois 60064
Received February 23, 1988; Revised Manuscript Received May 16, 1988

ABSTRACT: The conformation of the cyclic portion (7-23) of naturally occurring rat atrial natriuretic factor, ANF(1-28), has been examined in sodium dodecyl sulfate (SDS) micelles using high-resolution NMR techniques. Evidence is presented which shows that ANF(7-23) has several regions of definable structure in SDS micelles which were not observed in earlier studies in bulk solvents. The ^{1}H NMR resonances of ANF(7-23) in SDS micelles were assigned using sequential assignment techniques, and the conformational properties were analyzed primarily from proton-proton distances obtained from the quantitative analysis of two-dimensional nuclear Overhauser effect spectra. Three-dimensional structures consistent with the NMR data were generated by using distance geometry and constrained minimization/dynamics. Several similar but not identical structures were found which adequately satisfied the NMR constraints. Although none of the structures adopted a standard secondary structure, the conformations of three different sections of the peptide, 8-13, 14-17, and 18-21, were nearly identical in all of the predicted structures when individually superimposed.

Atrial natriuretic factor (ANF) is a peptide hormone which exhibits natriuretic, vasorelaxant, and hypotensive activities (de Bold, 1985; Cantin & Genest, 1985; Sonnenberg et al., 1983). Considerable effort has been expended to develop therapeutically useful mimics of this peptide by synthesizing analogues with amino acid substitutions and deletions. These studies have helped define the amino acid residues important for eliciting a biological response. However, in order to fully understand the factors governing the biological activity of a peptide hormone and to mimic its function, it is necessary to know not only the peptide's important functional groups but also their relative spatial orientation when bound to its receptor.

Several attempts have been made to deduce the bioactive conformation of ANF. In one approach, hypothetical models for the biologically active conformation were proposed on the basis of the structure activity data. In one such study, the biological activity of a series of ANF analogues in which each amino acid was systematically replaced with its D enantiomer indicated that the residues between Phe-8 and Ile-15 had the largest effect on activity. Furthermore, point substitutions of D-Ala for Gly suggested that the conformation of the ANF ring consisted of multiple turns (Nutt & Veber, 1987). These data can be compared to predictions based on Chou-Fasman rules which predict that the bioactive conformation consists of helical regions with no β sheet (Flynn, 1985).

Molecular modeling techniques have also been used to probe the accessible conformation space of the peptide to identify conformational features which could be tested by the synthesis of specific ANF analogues. In one study, 11 different families of related low-energy structures of ANF were identified from a short dynamics run (D. H. J. Mackay, personal communication). However, it is clear that an even larger number of families of structures would be found from a much longer run (S. Burt and C. Hutchins, personal communication). Because of the number of possible conformations for this flexible molecule, it is difficult to determine the bioactive conformation by these methods without additional data.

Another approach is to determine the preferred solution conformation of ANF and test whether this is the biologically active conformation by synthesizing appropriate conformationally restricted analogues. However, several studies have found that due to the flexibility of the peptide no preferred conformation could be observed in bulk solvents (Fesik et al., 1985; Theriault et al., 1987). Even for ANF analogues in which proline was substituted for glycine in order to rigidify the structure, no preferred conformation was found (Gampe et al., 1988).

In this paper, the conformational properties of the cyclic portion (7-23)

$$\underbrace{C^{7}\text{-F-G-G-R}^{11}\text{-I-D-R-I}^{15}\text{-G-A-Q-S}^{19}\text{-G-L-G-C}^{23}} \quad (1)$$

0006-2960/88/0427-7124$01.50/0

of naturally occurring rat ANF have been examined by NMR in an aqueous solution of sodium dodecyl sulfate (SDS) micelles. As in other studies (Brown, 1979; Briggs & Gierasch, 1984; Braun et al., 1983; Wakamatsu et al., 1986; Zetta et al., 1986; Weiner et al., 1987), SDS was used to approximate an anisotropic, amphiphilic environment such as that found in biological membranes, hoping that under these conditions ANF will adopt a preferred conformation.

Conformational constraints from the quantitative analysis of nuclear Overhauser effects and coupling constants have been used to elucidate conformational features of ANF(7–23) in SDS micelles. Structures consistent with the NMR data were obtained by a combined use of distance geometry and constrained dynamics/minimization. Preliminary results from a ^{13}C T_1 study of the peptide in SDS micelles are described.

Materials and Methods

NMR. NMR samples were prepared by dissolving ANF-(7–23) (10 mM) into a 200 mM H_2O/D_2O (9/1) solution (pH 5.0) of perdeuteriated SDS-d_{25} (MSD Isotopes). In order to create a more uniform suspension, samples were briefly (5 min) sonicated before use. As evidence for the incorporation or interaction of ANF(7–23) with SDS, the proton NMR resonances of ANF(7–23) in SDS micelles were found to be broader compared to the corresponding signals of the peptide in aqueous solution. This broadening was attributed to the incorporation of the peptide into the micelle and not to viscosity changes in the solution.

Proton NMR spectra were recorded on a General Electric GN 500 spectrometer and processed off-line on a VAX 11/780 with an attached CSPI-minimap array processor using software written by Dr. E. R. P. Zuiderweg. All plotting and other data analysis were done by using a modified version of Dr. Dennis Hare's FTNMR program (Hare Research, Inc.). All of the two-dimensional NMR experiments reported here were recorded at 40 °C. However, additional experiments were also performed at 20 °C (unpublished data) to help eliminate some ambiguities in NOE cross-peak assignments.

Two-dimensional double quantum filtered correlation experiments (DQF-COSY)(Piantini et al., 1982) consisted of 512 pairs (real and imaginary) of t_1 slices of 4096 complex points. The data were processed in the phase-sensitive mode using a shifted sine bell of 0° (t_2) and 45° (t_1) to give a final data set size of 1K by 4K with a digital resolution of 1.22 Hz/point in ω_2.

Relayed coherence-transfer experiments were performed as described previously (Eich et al., 1982; Bax & Drobny, 1985), by collecting 160 t_1 slices of 1024 complex points. A mixing time of 20 ms was used to optimize the coherence transfer for the type of spin systems present (Bax & Drobny, 1985). To process the data, a sine bell window function was used in both dimensions, and the data were zero filled in t_1. After Fourier transformation of the data, the absolute value was calculated.

The two-dimensional nuclear Overhauser experiments (2D NOE) (Kumar et al., 1980; States et al., 1982) were collected using mixing times of 50, 100, 200, 300, and 400 ms. The experiments were performed in both D_2O and H_2O and analyzed separately. The H_2O resonance in the spectra was suppressed either by a jump return sequence (Clore et al., 1983; Plateau & Gueron, 1982) or by coherent irradiation of the H_2O resonance (Zuiderweg et al., 1984). The data obtained with a jump return sequence were used to check for any bleaching effect that presaturation may have on amide to α NOE cross-peak intensities. In each data set, 350 pairs of t_1 slices consisting of 1024 complex points were accumulated. A shifted sine bell of 45° (t_2) and 90° (t_1) was used during processing. The final data set size was 1K by 1K, giving a final digital resolution of 4.88 Hz/point. The NOE data were base line corrected along each column (ω_1) and row (ω_2), using a fifth-order polynomial prior to cross-peak and diagonal peak integrations.

The NOE data were analyzed by using a multispin approach (Olejniczak et al., 1986). In this method, all of the cross and diagonal peak volumes for a set of dipolar-coupled spins are measured. This matrix of volumes, $V(\tau_m)$, is directly related to the relaxation rate matrix, Γ, by the equation (Macura & Ernst, 1980; Perrin & Gipe, 1984):

$$V(\tau_m) = \exp^{-\Gamma\tau_m} V^0 \tag{2}$$

where V^0 is the volume at a mixing time of $\tau_m = 0.0$. In principle, this method takes into account all of the complicated relaxation pathways that result in the observed cross-peak volume. Thus, multispin effects are accounted for even though there is no prior knowledge of the three-dimensional structure of the molecule. The data analysis for each individual mixing time gives a separate measurement of the relaxation rate constants. The average value for each relaxation rate constant was used in all of the calculations. The standard deviation of the measurements was used for estimating errors. A detailed discussion of the methodology used in this type of analysis has been presented earlier (Olejniczak et al., 1986).

Distances were obtained from the measured cross-relaxation rates and a known fixed distance, r_{kl}, using the equation:

$$r_{ij} = \left(r_{kl}^6 \frac{\sigma_{kl}}{\sigma_{ij}} \right)^{-1/6} \tag{3}$$

This approach assumes that the dynamics of both interproton vectors are similar. Our known distance was chosen to be the distance between geminal protons (i.e., $r_{kl} = 1.8$ Å). The standard deviation of the measured σ for five well-resolved geminal protons was small, and the average was used for σ_{kl} in eq 3. An estimate of the error in the distance between protons l,m caused by uncertainties in σ_{lm} is given by

$$\Delta r_{lm} = K \frac{\Delta\sigma_{lm}}{\sigma_{lm}^{5/6}} \tag{4}$$

$\Delta\sigma_{lm}$ can be estimated from the standard deviation in the relaxation constant. The standard deviation of σ is generally less than 20% of the average value. This leads to very small errors in the distance. A greater source of distance uncertainties is due to the effect that motional averaging may have on the observed σ values (Olejniczak et al., 1984). This source of error is very difficult to quantitate.

The ^{13}C experiments were carried out on a Bruker AM-500. The ^{13}C NMR signals were assigned from the proton NMR resonances using a proton-detected proton/carbon chemical shift correlation experiment (Bax & Subramanian, 1986; Summers et al., 1986). The ^{13}C T_1's were measured with an inversion recovery sequence and were obtained by a least-squares fit to

$$I_z(t) = M_0[1.0 - 2.0 \exp(-t/T_1)] \tag{5}$$

Molecular Modeling. Distance geometry calculations were performed by using an algorithm supplied by Professor I. Kuntz (Havel et al., 1983). All distances fixed by the geometry of the peptide were used as input constraints. This included upper and lower bounds for all atoms connected by one or two intervening bonds. The lower bounds between all remaining atoms were always greater than or equal to the sum of their van der Waals radii. Chirality constraints were also included to ensure that only structures with L-amino acids would be predicted. Experimental distance constraints were

Table I: Experimental Chemical Shifts, $^3J_{HN^\alpha}$ (40 °C), and Amide Temperature Coefficients

residue	chemical shifts (ppm from TSP) NH	H^α	H^β	other	$^3J_{HN^\alpha}$	$(\Delta\delta/\Delta T) \times 10^3$
Cys-7		4.35	3.26			no[c]
Phe-8	8.83	4.67	3.11[a]	$H^{\delta,\epsilon}$ 7.26[a] H^ζ 7.16	4.4 ± 1.0	no[c]
Gly-9	8.41	3.97[b] 3.67[b]				4.7
Gly-10	8.08	3.87[b] 3.91[b]				6.0
Arg-11	8.07	4.22	1.85[a]	$H^{\gamma 1}$ 1.71[b] $H^{\gamma 2}$ 1.64[b] $H^{\delta 1,2}$ 3.16[a]	7.6	3.5
Ile-12	7.56	4.18	1.98	$H^{\gamma 1}$ 1.48[b] $H^{\gamma 2}$ 1.16[b] H^δ 0.92 H^γ 0.82	10.0	0.2
Asp-13	8.22	4.50	2.81[a]		7.4	5.7
Arg-14	7.86	4.20	1.83, 1.87[b]	$H^{\gamma 1,2}$ 1.68[a] $H^{\delta 1,2}$ 3.19[a]	6.8	3.7
Ile-15	7.48	4.14	1.96	$H^{\gamma 1}$ 1.51[b] $H^{\gamma 2}$ 1.16[b] H^δ 0.92 H^γ 0.87	9.9	0.2
Gly-16	8.11	3.90[b] 3.95[b]				3.0
Ala-17	7.91	4.25	1.41		6.5	4.2
Gln-18	8.16	4.31	2.05, 2.19[b]		8.8	3.5
Ser-19	7.98	4.40	3.89[a]		8.0	4.0
Gly-20	8.16	3.91[b] 3.99[b]				4.0
Leu-21	7.92	4.36	1.65[a]	H^γ 1.86 $H^{methyl_{1,2}}$ 0.90	7.9	3.7
Gly-22	8.24	3.94[a]				4.7
Cys-23	7.84	4.44	3.04		9.0	5.2

[a] Proton chemical shifts were degenerate. [b] Protons are not stereospecifically assigned. [c] Not observed.

obtained from vicinal three-bond scalar coupling constants between the amide and H^α protons and from the quantitative analysis of 2D NOE data sets. Additional "minimum approach" constraints were included in the distance geometry calculation which reflected the lack of NOEs between protons. If the lack of an NOE was not clear due to overlap, or if other NOEs were not observed to the protons in question, then the additional minimum approach constraint was not included. The minimum approach distance set for these constraints was 3.0 Å, a value that corresponds to NOEs easily observable under the experimental conditions employed. In our experience, it is important to include these lower bound constraints to exclude structures containing protons close together that contradict the lack of an observable NOE (Fesik et al., 1987a).

The constrained energy minimization and dynamics calculations were performed using the DISCOVER program (Biosym Technologies) which treats all of the protons in the system explicitly. The parameters used in the calculation were the default parameters of the program except for the ionic charges. These were scaled to one-tenth their normal values. The structures generated by the distance geometry algorithm were minimized in two runs. The first run was a short, 50-step steepest descent minimization in which all of the heavy atoms were tethered to the original starting structure by a 50 kcal/mol force constant. This short minimization was used to relax any bad van der Waals contacts while retaining the features of the original distance geometry structure. This was followed by a longer conjugate gradient minimization which included the NMR distance constraints. The NMR distance constraints were incorporated into the calculation by using a harmonic restraint between the two atoms of the form $K(r^S - r^E)^2$ where K is the force constant, r^E is the interatom distance obtained experimentally, and r^S is the current distance in the structure being minimized (Kaptein et al., 1985). The force constant, K, was adjusted for each constraint depending on the estimated accuracy of the measured interproton distance. The constrained dynamics were run at a temperature of 400 K. Structures obtained from the constrained dynamics were subjected to constrained minimization until the gradient was less than 0.01 kcal mol^{-1} Å^{-1}.

Results and Discussion

NMR. The proton NMR resonances were assigned by using standard two-dimensional experiments and sequential assignment techniques (Wagner & Wüthrich, 1982; Wüthrich, 1986a,b). Scalar coupling networks were identified for each of the amino acids by DQF-COSY and relayed COSY experiments. Next, adjacent residues were connected by using backbone NOE connectivities. A list of the proton assignments obtained by this sequential assignment procedure is given in Table I.

The conformational properties of ANF(7–23) in SDS micelles were examined from vicinal spin–spin coupling constants, NOEs, and the temperature dependence of the amide chemical shifts. Coupling constants were measured from one-dimensional spectra or in complicated regions of the spectra from high-resolution DQF-COSY experiments. The coupling constants between the amide and α-protons for some of the amino acids are large and are indicative of structure (Table I). However, the interpretation of the remaining coupling constants is ambiguous and could be due either to a defined structure or to an averaging among many different conformations (Bundi & Wüthrich, 1979; Hoch et al., 1985).

The NOE data were particularly helpful in defining the conformation of ANF(7–23). A contour plot of the NOE data observed for the amide protons (ω_2) is shown in Figure 1. A total of 73 observable NOEs were used to obtain distance

Table II: Experimental Distances Compared to Distances in Structures

proton–proton	distance (Å) from NOE	structure I	structure II	structure III	proton–proton	distance (Å) from NOE	structure I	structure II	structure III
Cys-7 H^{α}–Cys-7 H^{β_1}	2.9	3.0	3.0	2.9	Ile-15 H^{N}–Gly-16 H^{N}	2.7	2.7	2.7	2.8
Cys-7 H^{α}–Cys-7 H^{β_2}	2.8	2.7	2.6	2.8	Ile-15 H^{α}–Gly-16 H^{N}	2.5	2.4	2.2	2.2
Cys-7 H^{α}–Phe-8 H^{N}	2.8	2.8	2.7	3.1	Ile-15 H^{α}–Ala-17 H^{N}	3.7	3.8	3.5	3.7
Cys-7 H^{β_1}–Phe-8 H^{N}	3.2	3.5	3.3	3.2	Ile-15 H^{N}–Ala-17 H^{N}	4.3	5.0	5.2	5.4
Cys-7 H^{β_2}–Phe-8 H^{N}	3.2	2.5	2.5	2.7	Ile-15 H^{α}–Gln-18 H^{N}	4.6	4.6	4.7	5.3
Cys-7 H^{α}–Gly-22 H^{N}	4.4	4.6	4.4	4.6	Gly-16 H^{N}–Ala-17 H^{N}	3.1	3.0	2.7	2.8
Phe-8 H^{α}–Gly-9 H^{N}	3.0	3.0	3.0	3.0	Ala-17 H^{N}–Ala-17 H^{α}	2.5	2.4	2.4	2.4
Phe-8 $H^{\delta,\epsilon}$[b]–Ile-12 H^{γ_2}[c]	3.1	3.3	4.7	3.7	Ala-17 H^{N}–Gln-18 H^{N}	3.0	3.0	2.9	3.0
Phe-8 $H^{\delta,\epsilon}$[b]–Ile-12 H^{δ_1}[b]	3.1	4.7	3.6	3.3	Ala-17 H^{α}–Gln-18 H^{N}	2.7	2.6	2.6	2.6
Gly-9 H^{N}–Gly-9 H^{α_1}	2.7	2.8	2.8	2.9	Ala-17 H^{β}[c]–Gln-18 H^{N}	4.1	4.4	4.4	4.4
Gly-9 H^{N}–Gly- H^{α_2}	2.6	2.8	2.5	2.6	Ala-17 H^{α}–Ser-19 H^{N}	3.9	3.6	3.8	3.6
Gly-9 H^{N}–Gly-10 H^{N}	2.8	2.6	2.5	2.6	Ala-17 H^{β}[c]–Ser-19 H^{N}	4.4	4.8	4.4	4.6
Gly-9 H^{α_1}–Gly-10 H^{N}	2.5	2.3	2.5	2.4	Gln-18 H^{N}–Gln-18 H^{α}	3.0	2.9	2.6	3.0
Gly-9 H^{α_2}–Gly-10 H^{N}	3.0	3.2	3.2	3.2	Gln-18 H^{N}–Gln-18 H^{β_1}	3.5	3.6	3.5	3.4
Gly-9 H^{N}–Arg-11 H^{N}	4.6	4.5	5.2	4.9	Gln-18 H^{N}–Gln-18 H^{β_2}	2.8	2.8	2.8	2.6
Gly-10 H^{N}–Ile-12 H^{N}	4.3	4.3	3.8	4.2	Gln-18 H^{N}–Ser-19 H^{N}	2.8	2.7	2.9	2.6
Gly-10 H^{N}–Asp-13 H^{N}	3.6	3.7	4.1	4.2	Gln-18 H^{α}–Ser-19 H^{N}	2.7	2.9	2.8	3.0
Gly-10 H^{α_2}–Ile-12 H^{N}	4.1	4.5	4.8	4.3	Gln-18 H^{β_1}–Ser19 H^{N}	3.5	3.7	3.5	3.6
Arg-11 H^{N}–Arg-11 H^{α}	2.9	3.0	2.7	2.9	Gln-18 H^{β_2}–Ser-19 H^{N}	4.1	4.0	4.2	4.0
Arg-11 H^{N}–Ile-12 H^{N}	2.8	2.5	2.3	2.7	Ser-19 H^{N}–Ser-19 H^{α}	3.1	3.0	3.1	3.0
Arg-11 H^{N}–Ile-12 H^{β}	3.6	3.6	4.0	3.5	Ser-19 H^{N}–Gly-20 H^{N}	3.0	3.1	3.2	3.3
Arg-11 H^{α}–Ile-12 H^{N}	2.9	3.0	2.7	2.8	Ser-19 H^{α}–Gly-20 H^{N}	2.9	2.8	2.8	2.9
Ile-12 H^{N}–Ile-12 H^{α}	3.0	2.9	2.6	3.0	Gly-20 H^{N}–Gly-20 H^{α_1}	2.4	2.5	2.6	2.5
Ile-12 H^{N}–Ile-12 H^{β}	2.5	2.4	2.3	2.5	Gly-20 H^{N}–Gly-20 H^{α_2}	2.2	2.2	2.2	2.2
Ile-12 H^{N}–Asp-13 H^{N}	3.0	3.1	3.2	3.1	Gly-20 H^{N}–Leu-21 H^{N}	3.1	3.1	3.2	3.2
Ile-12 H^{α}–Asp-13 H^{N}	2.4	2.6	2.4	2.3	Leu-21 H^{N}–Leu-21 H^{α}	2.9	2.7	2.6	2.9
Arg-14 H^{N}–Arg-14 H^{α}	2.7	2.6	2.8	2.8	Leu-21 H^{N}–Leu-21 H^{δ_3}[c]	4.0	4.7	4.4	4.4
Arg-14 H^{N}–Arg-14 H^{β_1}	3.3	3.4	3.3	3.2	Leu-21 H^{N}–Gly-22 H^{N}	2.6	2.4	2.3	2.4
Arg-14 H^{N}–Arg-14 H^{β_2}	2.8	2.8	2.6	2.6	Leu-21 H^{α}–Gly-22 H^{N}	2.7	2.8	3.0	2.8
Arg-14 H^{α}–Arg-14 H^{β_1}	2.5	3.0	2.8	2.5	Leu-21 H^{β_1}–Gly-22 H^{N}	3.3	3.3	3.3	3.5
Arg-14 H^{α}–Arg-14 H^{β_2}	2.4	2.5	2.5	2.4	Leu-21 H^{β_2}–Gly-22 H^{N}	3.3	3.5	3.5	3.3
Arg-14 H^{N}–Ile-15 H^{N}	2.7	3.0	2.8	2.4	Gly-22 H^{N}–Cys-23 H^{N}	2.7	2.7	2.8	2.7
Arg-14 H^{α}–Ile-15 H^{N}	2.8	2.5	3.1	3.1	Cys-23 H^{N}–Cys-23 H^{α}	3.0	2.9	3.0	2.9
Arg-14 H^{β_1}–Ile-15 H^{N}	3.2	3.0	3.2	3.6	Cys-23 H^{N}–Cys-23 H^{β_1}	3.4	3.5	3.6	3.4
Arg-14 H^{β_2}–Ile-15 H^{N}	3.9	4.0	4.0	4.0	Cys-23 H^{N}–Cys-23 H^{β_2}	2.9	2.8	2.7	2.8
Ile-15 H^{N}–Ile-15 H^{α}	3.0	2.7	2.9	3.0					

[a] β-Protons are not stereospecifically assigned. The indicated proton assignment is based solely on agreement between the experimental distances and corresponding distances in the structures. [b] The δ and ϵ protons of Phe-8 are degenerate. The reported distance is the r^{-6} average to the closest of either the Phe δ- or the ϵ-protons in the structures. The distance to methyl groups in the structures was calculated with $r_j = [(1/3)\sum_{i=1}^{3}(1/r_{ij}^{6})]^{-1/6}$.

constraints to define the conformation of the peptide. An additional 80 lower bound, minimum approach distances were included as constraints. Other proton–proton distance constraints were obtained from an analysis of isotope-filtered 2D NOE experiments as described elsewhere (Fesik et al., 1987b). A list of distances derived from the quantitative analysis of the 2D NOE spectra is given in Table II.

Several of the NOEs that were observed indicate that ANF(7–23) has regions of defined structure. For example, NOEs were observed between protons of amino acids separated by two or more intervening residues (medium/long-range NOEs). Although a flexible molecule may have many different close approaches between protons on distant residues, the occurrence of any particular close approach will tend to be rare, resulting in an unobservably small NOE. The medium/long-range NOEs observed include NOEs between the aromatic protons of Phe-8 and the methyls of Ile-12, between Gly-10 amide and Asp-13 amide, and between Ile-15 amide and Ala-17 amide as well as other medium/long-range NOEs (see Table II). A summary of the observed interresidue NOEs is given in Table III. The table shows that all of the medium/long-range NOEs are observed in two different sections of the peptide. No medium/long-range NOEs connect these two separate sections.

In addition to the medium/long-range NOEs, the short-range backbone NOEs that were observed are distinctly different from what we have empirically found for random-coil peptides (Gampe et al., 1988). In a random-coil peptide, the $\alpha(i)$ to amide $(i + 1)$ NOEs tend to be larger than the amide(i) to amide($i + 1$) NOEs. This difference usually corresponds to a factor of between 3 and 6 in the cross-peak volumes. This is exactly what was observed in rat ANF analogues in dimethyl sulfoxide and in aqueous solution (Fesik et al., 1985; Gampe et al., 1988). However, in the present data, many of these NOEs are of similar magnitude.

The most noticeable feature of the experimental backbone constraints is the large number of amide(i) to amide($i + 1$) NOEs which correspond to short NH-NH distances. These data indicate that a β or extended structure for the peptide is unlikely. Strong NH–NH NOEs are observed in an α helix. An α helix is however unlikely due to the lack of the characteristic long-range NOEs between $\alpha(i)$ and amide($i + 3$) or $\alpha(i)$ and $\beta(i + 3)$ protons and because the $\alpha(i)$ to amide($i + 1$) NOEs are strong in the data but are weak in α helixes (Wüthrich et al., 1984; Wüthrich, 1986a,b). The fact that the amide hydrogen exchange rates for the molecule are quite fast also indicates that a stable, standard, helical structure is not present.

A recent CD study (Epand & Stahl, 1987) indicated that ANF(5–28), in a variety of amphiphilic environments, primarily adopts a β or extended structure with little or no α helix. The short NH–NH distances observed in this study are inconsistent with an extended structure, but our data also indicate that no standard α helix is present.

Table I lists the amide chemical shifts of ANF(7–23) as a function of temperature. The amides show only a small

Table III: Summary of the Interresidue NOEs[a]

X	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
7	⊕	⊕														⊕	
8	⊕	⊕	⊕			⊕											
9		⊕	⊕	⊕	⊕												
10			⊕	⊕		⊕	⊕										
11			⊕		⊕	⊕											
12		⊕		⊕	⊕	⊕	⊕										
13				⊕		⊕	⊕										
14								⊕	⊕								
15								⊕	⊕	⊕	⊕	⊕					
16									⊕	⊕	⊕						
17									⊕	⊕	⊕	⊕	⊕				
18									⊕		⊕	⊕	⊕				
19											⊕	⊕	⊕	⊕			
20													⊕	⊕	⊕		
21														⊕	⊕	⊕	
22	⊕														⊕	⊕	⊕
23																⊕	⊕

[a] The presence of NOEs connecting two residues is indicated by the symbol ⊕.

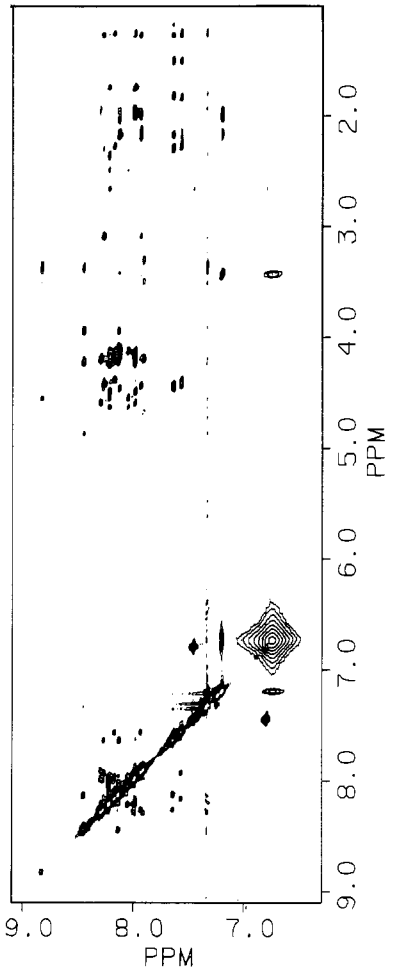

FIGURE 1: Portion of a two-dimensional NOE contour map of ANF(7–23)/SDS in H_2O. The mixing time is 200 ms; all other parameters are as described in the text.

change in their amide resonance frequency as a function of temperature (i.e., $\Delta\delta/\Delta T$ less than 0.004 ppm/°C), while the amides of Ile-12 and -15 are nearly invariant over the measured temperature range. The small amide temperature coefficients are significant since they are often indicative of more conformationally defined regions of peptides (Smith & Pease, 1980; Wüthrich, 1986a,b). However, since the ANF-(7–23) molecule is involved in a micelle complex, intra- as well as intermolecular interactions may be responsible for the small amide temperature coefficients.

Additional information on the peptide was obtained from ^{13}C NMR studies. The $^{13}C^{\alpha}$ resonances were assigned from a proton/carbon correlation experiment using the H^{α} assignments. The measured $^{13}C^{\alpha}$ chemical shifts indicate sequence specific differences. For example, the C^{α} chemical shifts of the two Ile residues and the two Arg residues differ by more than 3 ppm. Furthermore, one of the Gly α-carbons is shifted downfield by more than 10 ppm from the other Gly resonances.

Differences in the motional averaging of the different residues were examined by measuring the T_1's of six well-resolved C^{α} resonances as well as the average T_1 of three overlapping Gly residues. The measured T_1's (scaled by the number of attached protons for the Gly residues) were found to be similar (all around 330 ms). Although more accurate T_1 values would have to be measured for a thorough analysis, the preliminary data indicate that there are no dramatic differences in the flexibility of different parts of the peptide backbone.

Molecular Modeling. Three-dimensional structures of ANF(7–23) were generated that were consistent with the NMR data by the combined use of distance geometry and constrained minimization/dynamics (Clore & Gronenborn, 1987). First, 250 structures were generated by distance geometry as starting structures for further refinement. Distance geometry structures were selected for further refinement if they had no violations of the upper and lower bounds of the NMR distance constraints greater than 0.7 Å and if the structures did not predict any additional short (i.e., less than 3.0 Å) interproton distances between protons with no observable NOE. Thirty-seven structures met these criteria and were chosen for further refinement with constrained energy minimization. Following the constrained minimization, structures were again checked for consistency with the NMR data. Only structures where the average fractional constraint violation [i.e., $(1/N)\sum_{i=1}^{n}|(r_i^E - r_i^S)|/r_i^E$] was less than 0.15 were retained. The remaining structures were then grouped into three different families based on the root mean square (rms) difference of the backbone atoms in the structures. One structure from each of the different families was further refined with a 10-ps constrained dynamics run at 400 K, followed by constrained energy minimization of selected structures generated in the run. One fully refined structure from each of these runs, that best fit the NMR data, was selected. These three structures are representative examples from each of the families and of the full range of structures that satisfy the NMR constraints and are used for further comparisons.

A comparison of the distances in these three structures to those obtained from the NMR data is given in Table II. As

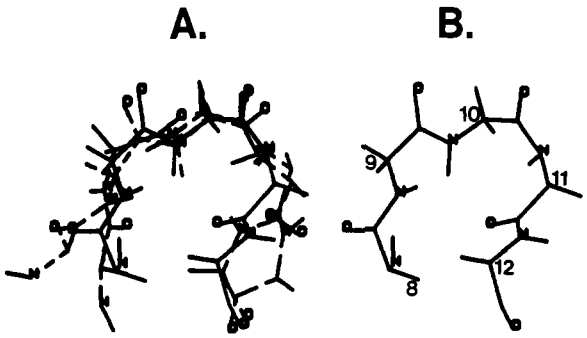

FIGURE 2: Backbone atoms of the region between Phe-8 and Ile-12. (A) Superposition of the backbone atoms of three predicted structures. The largest rms between the backbone atoms of any pair of the three structures was 1.3 Å. When compared to the average coordinates, the average rms is 0.88 Å. (B) Backbone of the structure which best fit the NMR data.

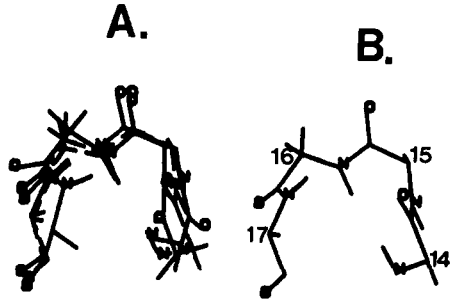

FIGURE 3: Backbone atoms of the region between Arg-14 and Ala-17. (A) Superposition of the backbone atoms of three predicted structures. The largest rms between the backbone atoms of any pair of the three structures was 0.82 Å. When compared to the average coordinates, the average rms is 0.43 Å. (B) Backbone of the structure which best fit the NMR data.

can be seen, all of the observable short-distance NOEs are accounted for by the structures. In addition, all protons closer than 3 Å in the structure can be accounted for by the observed NMR data.

All three structures predicted similar conformational features for particular regions of the peptide. When each of these three regions was superimposed, the rms difference between the backbone atoms in any pair of the three structures was less than 1.3 Å, and the largest rms between any structure and the local average coordinates was always less than 0.88 Å.

The first of the three regions (Phe-8 to Ile-12) is shown in Figure 2. The NOEs between the Phe-8 aromatic protons and the Ile-12 methyl groups constrain the side chains of these two residues to be on the same side of the molecule. In the predicted structures, the amide of Ile-12 is pointing outward. This suggests that internal hydrogen bonding is most likely not the reason for its small amide temperature coefficient. As shown in the plot of the backbone atoms (Figure 2A), there is a reasonably good agreement between all three of the structures. Specific features of the backbone can be seen more clearly in Figure 2B where for clarity only the backbone atoms of one of the three structures are shown.

The second region which is almost identical in all three structures is the region from Arg-14 to Ala-17. Plots of the backbone of this region are shown in Figure 3. The turn found here is similar to but clearly not identical with a β I turn. In particular, the α to amide NOE between Arg-14 and Ile-15 indicates a distance of 2.7 Å versus the 3.6 Å found in a standard turn.

The two regions (Phe-8 to Ile-12 and Arg-14 to Ala-17) form two different loops. Good, simultaneous overlap of all of the residues in the two loops between Phe-8 and Ala-17 does not occur in the three families of structures. In one of the structures, the two loops almost stack on one another, forming a section that is similar to a very loose left-handed helix. In the other two cases, the two loops are adjacent to each other, forming roughly an "S"-like shape.

All three structures had another loop between Gln-18 and Leu-21. The rms between any pair of structures in this region is 1.3 Å. When the structures are compared to the local average coordinates, the largest rms is 0.65 Å. The high reproducibility of this section of the peptide in the predicted structures is surprising, since it is only constrained by short-range NOEs (Table III). A NOE between Ala-17 and Ser-19 provides a constraint between this loop and the section 14–17.

Although short sections of the peptide superimpose very well, the overall all rms difference of the backbone atoms between any pair of predicted structures is quite large. The average rms difference between the backbone atoms of any two structures is 3.0 Å. When compared to the average coordinates of the structures, the average rms difference is 1.8 Å.

To show features of the overall conformation of the peptide, we have derived an average structure from the average coordinates. This was accomplished by tethering the structure to the backbone atoms of the original average coordinates and then minimizing the structure while slowly increasing the contribution of the nonbonded terms in the potential function. The NMR constraints were also included in the potential function. An energy-refined structure was obtained that adequately satisfied the NMR constraints using this procedure.

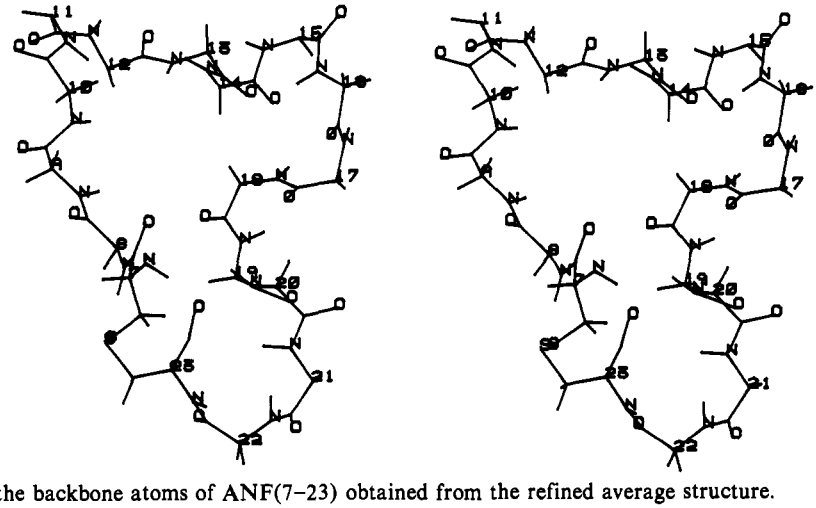

FIGURE 4: Stereo view of the backbone atoms of ANF(7–23) obtained from the refined average structure.

The rms between the final structure and the original average backbone coordinates was 1.5 Å. Since the average coordinates do not represent an energetically feasible molecule, an rms value of this size is reasonable. When the three structures are compared to this refined structure, the average rms of the backbone atoms is 2.2 Å. This is close to the value obtained using the average coordinates.

Even though there are differences in the predicted structures, they all exhibit similar features. Figure 4 shows the backbone atoms of the average structure. In this structure, the three-loop "cloverleaf"-like shape is apparent for the backbone of the molecule. This shape is common to all three families of predicted structures. Although there is considerable variability in the dihedral angles of the structures, variability in the dihedral angles of the residues connecting the loops seems to have the greatest effect on the overall shape of the molecule.

CONCLUSIONS

Peptides are inherently flexible molecules in solution. For many small peptides, this precludes any detailed analysis of their solution structure. In some cases, evidence of a less flexible more defined structure is present. Careful analysis of a variety of NMR parameters can then be used to obtain useful structural information for the less flexible parts of the peptide.

In this study, we have obtained NMR data with which we have characterized the conformation of three sections (i.e., residues 8–12, 14–17, and 18–21) of ANF(7–23). Due to an insufficient number of long-range NOE distance constraints, we were unable to adequately characterize the relative orientation of these short sections of the peptide. Thus, even though the information about local regions of the peptide is very detailed, the overall conformation of the peptide is not as well determined. None of the sections of the peptide that could be analyzed in detail conforms to any standard secondary structure typically observed in the X-ray crystal structures of proteins. The lack of a standard secondary structure in the peptide can be rationalized in several ways. First, it could be indicative of some restricted motional averaging in the molecule. Thus, even if sections of the peptide are predominantly found in some standard secondary structure, this well-defined secondary structure may not be reflected in the motionally averaged NMR parameters. Another possibility is that the lack of any standard secondary structure is a consequence of the amphiphilic environment of the peptide in SDS micelles. The interaction of SDS with the peptide may help stabilize a less regular three-dimensional conformation than found in large proteins.

It is interesting to note that the proposed structure(s) of ANF(7–23) is (are) consistent with some of the features deemed important by structure/activity studies (Nutt & Veber, 1987). For example, our NMR data indicate that Phe-8 and Ile-12 are positioned on the same side of the molecule. Structure/activity data for these two residues have shown that it is important to have hydrophobic residues in these positions. It has therefore been suggested that they interact with a hydrophobic pocket on the receptor.

In this study, we provided evidence and described regions of definable structure in ANF(7–23). In order to judge the biological significance of the NMR structure, it will be necessary to synthesize constrained analogues to test if the conformation determined here is a useful model for the receptor-bound conformation.

ACKNOWLEDGMENTS

We gratefully acknowledge the support and encouragement of Dr. Thomas J. Perun and Dr. William H. Holleman.

Registry No. ANF(7–23), 96235-91-5; SDS, 151-21-3.

REFERENCES

Bax, A., & Drobny, G. (1985) *J. Magn. Reson. 61*, 306.
Bax, A., & Subramanian, S. J. (1986) *J. Magn. Reson. 67*, 565.
Braun, W., Wider, G., Lee, K. H., & Wüthrich, K. (1983) *J. Mol. Biol. 169*, 921.
Briggs, M. S., & Gierasch, L. M. (1984) *Biochemistry 23*, 3111.
Brown, L. (1979) *Biochim. Biophys. Acta 557*, 135.
Bundi, A., & Wüthrich, K. (1979) *Biopolymers 18*, 285.
Cantin, M., & Genest, J. (1985) *Endocr. Rev. 6*, 107.
Clore, G. M., & Gronenborn, A. M. (1987) *Protein Eng. 1*, 275.
Clore, G. M., Kimber, B. J., & Gronenborn, A. M. (1982) *J. Magn. Reson. 54*, 170.
de Bold, A. J. (1985) *Science (Washington, D.C.) 230*, 767.
Eich, G., Bodenhausen, G., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 3731.
Epand, R., & Stahl, G. L. (1987) *Int. J. Pept. Protein Res. 29*, 238.
Fesik, S. W., Holleman, W. H., & Perun, T. J. (1985) *Biochem. Biophys. Res. Commun. 131*, 517.
Fesik, S. W., Bolis, G., Sham, H. L., & Olejniczak, E. T. (1987a) *Biochemistry 26*, 1851.
Fesik, S. W., Gampe, R. T., Jr., & Rockway, T. W. (1987b) *J. Magn. Reson. 74*, 366.
Flynn, T. G. (1985) *Abstracts, 9th American Peptide Symposium*, P-15.
Gampe, R. T., Jr., Connolly, P. J., Rockway, T., & Fesik, S. W. (1988) *Biopolymers 27*, 313.
Havel, T. F., Kuntz, I. D., & Crippen, G. M. (1983) *Bull. Math. Biol. 45*, 665.
Hoch, J. C., Dobson, C. M., & Karplus, M. (1985) *Biochemistry 24*, 3831.
Kaptein, R., Zuiderweg, E. R. P., Scheek, R. M., Boelens, R., & Van Gunsteren, W. F. (1985) *J. Mol. Biol. 182*, 179.
Kumar, A., Ernst, R. R., & Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1.
Macura, S., & Ernst, R. R. (1980) *Mol. Phys. 41*, 95.
Nutt, R. F., & Veber, D. F. (1987) *Endocrinol. Metab. Clinics North America 16*, 19.
Olejniczak, E. T., Dobson, C. M., Karplus, M., & Levy, R. M. (1984) *J. Am. Chem. Soc. 106*, 1923.
Olejniczak, E. T., Gampe, R. T., & Fesik, S. W. (1986) *J. Magn. Reson. 67*, 28.
Perrin, C. L., & Gipe, R. K. (1984) *J. Am. Chem. Soc. 106*, 4036.
Piantini, U., Sorensen, O. W., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6800.
Plateau, P., & Gueron, M. (1982) *J. Am. Chem. Soc. 104*, 7310.
Smith, J. A., & Pease, L. G. (1980) *CRC Crit. Rev. Biochem. 8*, 315.
Sonneberg, H., Milojevic, S., Chong, C. K., & Veress, A. T. (1983) *Hypertension (Dallas) 5*, 672.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286.
Summers, M. F., Marzelli, L. G., & Bax, A. (1986) *J. Am. Chem. Soc. 108*, 4285.
Theriault, Y., Boulanger, Y., Weber, P. L., & Reid, B. R. (1987) *Biopolymers 66*, 1075.
Wagner, G., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 347.
Wakamatsu, K., Okado, A., Higashhyima, T., & Miyazawa, T. (1986) *Biopolymers 25*, s193.

Weiner, J. H., Dettman, H. D., Henry, G. D., O'Neil, J., & Sykes, B. D. (1987) *Biochem. Soc. Trans. 15*, 87.
Wüthrich, K. (1986a) *NMR in Biological Research: Peptides and Proteins*, North-Holland, Amsterdam.
Wüthrich, K. (1986b) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Wüthrich, K., Billeter, M., & Braun, W. (1984) *J. Mol. Biol. 180*, 715.
Zetta, L., De Marco, A., & Zannoni, G. (1986) *Biopolymers 25*, 2315.
Zuiderweg, E. R. P., Hallenga, K., & Olejniczak, E. T. (1986) *J. Magn. Reson. 70*, 207.

Complete Sequence and Organization of the Murine β-Glucuronidase Gene[†,‡]

Melanie A. D'Amore,[§,∥] Patricia M. Gallagher,[§] Thomas R. Korfhagen,[⊥] and Roger E. Ganschow[*,§]
Division of Basic Science Research and Division of Pulmonary Biology, Children's Hospital Research Foundation, Cincinnati, Ohio 45229, and Graduate Program in Developmental Biology, University of Cincinnati, Cincinnati, Ohio 45221
Received March 3, 1988; Revised Manuscript Received May 25, 1988

ABSTRACT: The murine β-glucuronidase structural gene (*Gus-s*) has been isolated from a BALB/cJ sperm DNA bacteriophage library and its nucleotide sequence established. The gene is organized into 12 exons comprising 17.5% of the 14 009 base pair (bp) region spanning the interval between transcription initiation and the putative site of polyadenylation. A TATA box sequence, embedded within a GC-rich region, is found 28 bp upstream from the transcription initiation site. Eleven members of the B1 family and eight members of the B2 family of murine repetitive elements were identified within *Gus-s* and 2440 bp of flanking sequence. Other novel sequences found within *Gus-s*, including a $(AC)_{19}$ homocopolymer tract within intron 3 and a 23 base pair complex direct repeat within intron 9, are presented and discussed.

The murine β-glucuronidase gene complex, designated [*Gus*], provides an excellent model system for examining the structure and function of mammalian regulatory elements which serve to control and modify gene expression. The GUS structural gene, *Gus-s*, and three GUS-specific regulatory elements, identified through characterization of natural variants of GUS expression, define the GUS gene complex on chromosome 5 of the mouse [for a review, see Paigen (1979)]. Three common alleles of *Gus-s* ($-s^a$, $-s^b$, $-s^h$) specify allozymes which differ in electrophoretic mobility, heat stability, or both (Paigen, 1961; Swank et al., 1973; Lalley & Shows, 1974). Specific alleles of each of three GUS regulatory elements (*Gus-r*, *Gus-t*, and *Gus-u*) are associated with specific alleles of *Gus-s*, and these associations define three common haplotypes, $[Gus]^a$, $[Gus]^b$, and $[Gus]^h$, and several rare haplotypes. The effects of each regulatory variant on the expression of GUS have been examined in considerable detail at the cellular and biochemical levels.

Gus-r, a cis-active regulatory element tightly linked to *Gus-s*, controls the androgen responsiveness of kidney GUS mRNA (Palmer et al., 1983). Recently, a rare GUS haplotype, $[Gus]^{or}$, was described in which kidney GUS does not respond to androgen (Lund et al., 1988). $[Gus]^{or}$ represents a cis-active, "null" variant of the androgen response of kidney GUS which may represent either a mutant form of a novel, GUS-specific response element or an allele of *Gus-r*. To date, no androgen-responsive element has been identified in an androgen-responsive gene. Structural and functional comparisons of DNAs within and surrounding the GUS structural gene between this variant and haplotypes which respond normally could provide identification of this element.

A second element, *Gus-u*, is tightly linked to *Gus-s* and controls in cis the levels of GUS synthesis in all tissues at all times (Lusis et al., 1983). In addition, a third regulatory element, designated *Gus-t*, is tightly linked to *Gus-s* and exerts an additional control in trans over GUS synthesis in certain tissues (Meredith & Ganschow, 1978; Lusis et al., 1983). Recent preliminary studies from our laboratory strongly infer that the control of the rates of GUS synthesis by *Gus-u* and *Gus-t* is not exerted through control of GUS mRNA levels (Wawrzyniak and Ganschow, unpublished experiments). Systematic comparisons and genetic manipulations of the DNA within the GUS gene complex could reveal the DNA determinants of these regulatory elements.

Clearly, efforts to identify the DNA determinants of the GUS-specific regulatory elements require the characterization of the GUS structural gene (*Gus-s*). To this end, we have used a GUS cDNA,[1] pGUS-1 (Palmer et al., 1983), to isolate a series of overlapping bacteriophage clones which span $Gus\text{-}s^a$. Determination of the entire nucleotide sequence from one haplotype provides molecular access into the GUS gene complex and allows the design of experiments to identify the differences in DNA structure within [*Gus*] which are causally related to differences in regulation among the GUS haplotypes.

Materials and Methods

Materials. Restriction enzymes, T4 DNA ligase, and M13 sequencing primer (17-mer) were purchased from New England BioLabs, Inc. T4 polynucleotide kinase, RNase T1, and T3 polymerase were obtained from Bethesda Research Laboratories. Sodium 2′,3′-dideoxynucleoside 5′-triphosphates, sodium 2′-deoxynucleoside 5′-triphosphates, and M13mp18

[†] Supported by Grants DK14770 and GM36266 to R.E.G. and Grant GM10799 to T.R.K. from the National Institutes of Health.
[‡] The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number J02836.
[*] To whom correspondence should be addressed at Children's Hospital Research Foundation, IDR Room 720, Elland and Bethesda Avenues, Cincinnati, OH 45229-2899.
[§] Division of Basic Science Research.
[∥] Graduate Program in Developmental Biology.
[⊥] Division of Pulmonary Biology.

[1] Abbreviations: bp, base pair(s); cDNA, complementary DNA; RF, replicative form.

0006-2960/88/0427-7131$01.50/0