Molecular Cloning and Primary Structure of Rat α_1-Antitrypsin†

Steven Chao,‡ Karl X. Chai,§ Lee Chao,§ and Julie Chao*,‡

Departments of Pharmacology and Biochemistry, Medical University of South Carolina, 171 Ashley Avenue, Charleston, South Carolina 29425

Received June 8, 1989; Revised Manuscript Received August 29, 1989

ABSTRACT: A cDNA clone encoding rat α_1-antitrypsin has been isolated from a λgt-11 rat liver cDNA library using an antigen-overlay immunoscreening method. The nucleotide sequence of this cDNA clone is 1306 base pairs in length and has a coding region of 1224 base pairs which can be translated into an α_1-antitrypsin precursor protein consisting of 408 amino acid residues. The cDNA sequence contains a termination codon, TAA, at position 1162 and a polyadenylation signal sequence, AATAAT, at position 1212. The calculated molecular weight of the translated mature protein is 43 700 with 387 amino acid residues; this differs from purified rat α_1-antitrypsin's apparent molecular weight of 54 000 because of glycosylation. Five potential glycosylation sites were identified on the basis of the cDNA sequence. The translated mature protein sequence from the cDNA clone matches completely with the N-terminal 33 amino acids of purified rat α_1-antitrypsin, which has an N-terminal Glu. The cDNA encoding rat α_1-antitrypsin shares 70% and 80% sequence identity with its human and mouse counterparts, respectively. The reactive center sequence of rat α_1-antitrypsin is highly conserved with respect to human α_1-antitrypsin, both having Met-Ser at the P1 and P1′ residues. Genomic Southern blot analysis yielded a simple banding pattern, suggesting that the rat α_1-antitrypsin gene is single-copy. Northern blot analysis using the cDNA probe showed that rat α_1-antitrypsin is expressed at high levels in the liver and at low levels in the submandibular gland and the lung. Dot blot analysis showed that rat α_1-antitrypsin RNA levels in the liver increased by 2-fold after induction by acute-phase inflammation and that it is sexually dimorphoric, with 5-fold higher levels in male than in female rats.

Human α_1-antitrypsin, a major plasma serine proteinase inhibitor, has been extensively studied and characterized (Travis & Salvesen, 1983; Heidtmann & Travis, 1986). Neutrophil elastase is the target enzyme of α_1-antitrypsin, which also exerts an inhibitory effect on a broad spectrum of serine proteinases, including protein C (Heeb & Griffin, 1988), thrombin, cathepsin G, trypsin, chymotrypsin, and kallikrein (Heidtmann & Travis, 1986). Deficiency of α_1-antitrypsin is known to be involved in the pathogenesis of human emphysema and other diseases (Heidtmann & Travis, 1986). The nucleotide sequences of both the cDNA and the gene of human α_1-antitrypsin have been revealed (Long et al., 1984). In addition, the protein structure, as well as the regulated, tissue-specific expression of the α_1-antitrypsin gene, has been well studied (Latimer et al., 1987; Monaci et al., 1988). Crystallographic data of postcomplex human α_1-antitrypsin reveal the reaction mechanism of the inhibition of serine proteinases by the serpin superfamily (Loebermann et al., 1984).

We have been studying the structure and function of tissue kallikrein in both human and rat systems (Chen et al., 1988; Lu et al., 1989; Murray et al., 1989; Shai et al., 1989). It has been reported that human α_1-antitrypsin is a very slow and progressive inhibitor of human tissue kallikrein, with a $t_{1/2}$ of around 10 h (Geiger et al., 1981a,b; Hirano et al., 1984; Horl et al., 1981, 1982). There has been a great deal of interest concerning the control of both kallikrein activity and bioavailability by specific inhibitors at the posttranslational level. We have previously identified a novel tissue kallikrein-binding protein and the endogenous complexes of kallikrein with the binding protein in both humans and rats (Chao et al., 1986a,b; Chao & Chao, 1988). This new protein may be the major modulator of tissue kallikrein. However, it is not clear whether a high concentration of α_1-antitrypsin in the circulation is relevant to the physiological function of tissue kallikrein. In addition, relatively little is known about rat α_1-antitrypsin and its interaction with tissue kallikrein. In order to understand the structure, expression, and the bioregulatory action of α_1-antitrypsin, we have purified the protein from rat serum to apparent homogeneity and isolated cDNA clones encoding α_1-antitrypsin from a rat liver expression cDNA library using a highly sensitive and specific immunoscreening method recently developed in our laboratory (Chao et al., 1989a). In this paper, we describe the biochemical and immunological properties of rat α_1-antitrypsin, along with the complete sequence and characterization of its cDNA and its tissue-specific expression.

Experimental Procedures

Purification of Rat α_1-Antitrypsin and Preparation of Antiserum. Rat α_1-antitrypsin was purified from pooled rat sera by using procedures similar to those reported for the purification of human α_1-antitrypsin (Pannel et al., 1974). Briefly, 40 mL of rat sera was applied to an Affigel-blue column and then to two DEAE-Sepharose CL-6B columns at pH 8.8 and 6.5, respectively. Fractions containing α_1-antitrypsin were combined and lyophilized and then further separated on a non sodium dodecyl sulfate (SDS) preparative polyacrylamide gel. Purified rat α_1-antitrypsin was resolved in an analytical 7.5–15% gradient SDS–polyacrylamide gel (PAGE).

New Zealand White rabbits were injected with 50–100 μg of the purified α_1-antitrypsin mixed with complete Freund's adjuvant (DIFCO). Booster injections of 50 μg of the purified α_1-antitrypsin were administered until a high titer of antiserum was detected by a radioimmunoassay developed as described previously (Chao et al., 1989b). α_1-Antitrypsin (5 μg) was

† This work was supported by NIH Grant HL29397 and American Cancer Society Grant BC556.

* Address correspondence to this author.

‡ Department of Pharmacology.

§ Department of Biochemistry.

0006-2960/90/0429-323$02.50/0

labeled with ^{125}I by the lactoperoxidase method (specific activity 50 μCi/μg of protein) (Shimamoto et al., 1980).

Lectin Blot Analysis. After analysis on 7.5–15% SDS–PAGE, the proteins were transferred onto nitrocellulose using a Mini Trans-Blot system (Bio-Rad). The nitrocellulose was blotted for 1 h in a buffer containing 3% bovine serum albumin (BSA) pretreated with periodic acid at 40 °C. The nitrocellulose was then incubated sequentially at room temperature for 1 h in a buffer containing 5 μg/mL concanavalin A (Con A) (Serva) or 20 μg/mL wheat germ agglutinin (WGA) (Sigma), for 1 h in rabbit antisera against either Con A or WGA (1:800 dilution) (Serva), for 30 min in goat antisera against rabbit IgG (1:100 dilution) (Cappel), and for 30 min in the peroxidase–antiperoxidase complex of rabbit origin (1:400 dilution) (Calbiochem). The nitrocellulose was washed 5 times for 5 min each between every incubation and then was stained with 0.03% 3,3′-diaminobenzidine and 0.005% hydrogen peroxide in 50 mM Tris-HCl, pH 7.6.

Amino Acid Sequence Analysis. Sequential Edman degradation was performed with an ABI Model 477 gas-phase sequencer equipped with an on-line narrow-bore PTH-amino acid analyzer (Applied Biosystems Inc.).

Western Blot Analysis. Samples (30 μL) containing either purified rat α_1-antitrypsin (0.010–1.0 μg) or rat serum (0.015–1.2 μL) in 0.01 M sodium phosphate, pH 7.0, were subjected to SDS–PAGE on 7.5–15% gradient slab gels and then electroblotted onto nitrocellulose (Burnett, 1981). The immunoblotting procedures using the antigen-overlay method were as described previously (Lammle et al., 1986). Briefly, the nitrocellulose was blocked with BLOTTO [5% w/v dry milk in 0.01 M sodium phosphate, pH 7.4, 0.14 M NaCl, 1 μM (*p*-aminophenyl)methanesulfonyl fluoride, 1 mg/L thimerosal, 200 mg/L NaN_3, and 0.01% antifoam] (Johnson et al., 1984) for 1 h at 30 °C and then incubated with rabbit anti-rat α_1-antitrypsin antiserum (1:500 in BLOTTO). After 3-h incubation at 30 °C with gentle shaking, the nitrocellulose membranes were washed 3 times with BLOTTO and then incubated with ^{125}I-α_1-antitrypsin (~250 000 cpm/mL). The nitrocellulose was washed 3 times with BLOTTO and once with phosphate-buffered saline (0.01 M sodium phosphate/0.14 M NaCl, pH 7.4), dried, and exposed to Kodak X-Omat film. The quantity of the immunoreactive α_1-antitrypsin was measured by cutting the autoradiographic positive spots from the nitrocellulose and counting in a γ counter.

Immunological Screening of the Rat Liver cDNA Library. A rat liver λgt-11 cDNA library of approximately 2.4×10^7 PFU/mL (plaque forming units per milliliter) was plated in soft agarose with *Escherichia coli* strain Y1090 as the host cell. Nitrocellulose filters soaked in 10 mM IPTG (isopropyl β-D-thiogalactoside) were used for plaque lifting according to the procedures of Young and Davis (1983). Filters were washed in TBST (50 mM Tris, pH 8.0, 150 mM NaCl, and 0.05% Tween-20) 3 times for 10 min each and then blocked for 1 h at 30 °C with BLOTTO. After being blocked, the filters were incubated for 1 h at 30 °C with gentle shaking in rabbit anti-rat α_1-antitrypsin antiserum (1:500 in BLOTTO). The filters were then washed 3 times in BLOTTO for 10 min each and incubated for 90 min with ^{125}I-α_1-antitrypsin (~250 000 cpm/mL) in BLOTTO. After being washed 3 times with BLOTTO and once with phosphate-buffered saline (0.01 M sodium phosphate/0.14 M NaCl, pH 7.4), the filters were air-dried and exposed to Kodak X-Omat film. Purified α_1-antitrypsin at concentrations of 1.5, 0.15, and 0.015 μg was spotted onto nitrocellulose and processed as above as positive controls.

Nucleic Acid Sequencing. The positive clone which contained the largest cDNA insert was subcloned into the M13 mp19 sequencing vector according to the method described by Messing (1983). Nucleic acid sequencing was performed by using Sanger's dideoxy chain termination method (Sanger et al., 1977). Sequence reactions were carried out in both directions.

RNA Preparation and Northern Blot Hybridization. Sprague-Dawley rats (Charles River Laboratories) weighing 250 g were anesthetized with pentabarbital (50 mg/kg body weight) and perfused with normal saline via cardiac puncture (Chao et al., 1988). The organs of interest were removed and homogenized separately in a solution of 4 M guanidine isothiocyanate, 0.5% sodium laurylsarcosine, 25 mM sodium citrate, and 0.1 M 2-mercaptoethanol. Total RNAs were then extracted by centrifuging the tissue homogenates through a 5.7 M CsCl gradient. The RNA was denatured in formaldehyde, separated by electrophoresis in a 1.5% agarose–formaldehyde gel, and then transferred to Zeta probe (Bio-Rad) by capillary blotting. The filter was air-dried and baked under vacuum at 80 °C for 2 h. Prehybridization was performed for 4 h at 55 °C in a solution of 5 × SSC (1 × SSC = 0.15 M NaCl/0.015 M sodium citrate), 50 mM sodium phosphate, pH 6.8, 1 mM sodium pyrophosphate, 100 μg/mL sonicated herring sperm DNA, 0.02% BSA, and 0.02% poly(vinylpyrrolidone). Hybridization was carried out in the same solution at 60 °C overnight with the nick-translated α_1-antitrypsin cDNA probe [$(1–2) \times 10^8$ cpm/μg]. The final washing of the Zeta probe was in 3 × SSC containing 0.5% Sarkosyl and 0.05% SDS at 50 °C for 30 min. Autoradiography was carried out at −70 °C with Kodak X-Omat film.

Southern Blot Hybridization of Rat Genomic DNA. High molecular weight rat genomic DNA was purified according to the procedures described by Maniatis et al. (1982). DNA samples of 10 μg each were digested with *Bam*HI, *Bgl*II, *Eco*RI, *Hin*dIII, *Sst*I, and *Stu*I (Bethesda Research Laboratories). DNA fragments were resolved in a 0.7% agarose gel in 1 × TEA buffer (0.05 M Tris-HCl, 0.02 M sodium acetate, 0.002 M Na_2EDTA, and 0.018 M NaCl, pH 8.2), followed by capillary transfer to nitrocellulose using a 10 × SSC solution. Hybridization was carried out at 42 °C in 50% formamide, using the nick-translated α_1-antitrypsin cDNA as the probe [$(1–2) \times 10^8$ cpm/μg]. The final wash of the blot was in a solution of 1 × SSC at 55 °C for 20 min and autoradiographed with Kodak X-Omat film.

Results

Biochemical and Immunological Characterization of Rat α_1-Antitrypsin. Rat α_1-antitrypsin was purified from serum to apparent homogeneity. Figure 1 shows the SDS–PAGE profiles of purified rat α_1-antitrypsin stained with Coomassie Blue (lane 2) and lectin immunoblots using antibodies to concanavalin A (lane 3) and wheat germ agglutinin (lane 4). Purified rat α_1-antitrypsin migrates as a 54-kDa protein in a 7.5–15% SDS–PAGE gradient gel under reducing conditions. The results indicate that α_1-antitrypsin is a glycoprotein containing methyl α-D-mannoside and *N*-acetyl-α-glucosamine groups, since it has binding activities to both concanavalin A and wheat germ agglutinin (lanes 3 and 4). Purified rat α_1-antitrypsin is homogeneous since N-terminal amino acid sequence analysis reveals a single N-terminal Glu. In the immunoblot assay of total rat serum (0.015–1.2 μL) with the antigen-overlay method using antiserum against purified α_1-antitrypsin, only one species of protein with the same molecular weight as the purified α_1-antitrypsin (54 000) was recognized by the antiserum (Figure 2, left panel). Detection

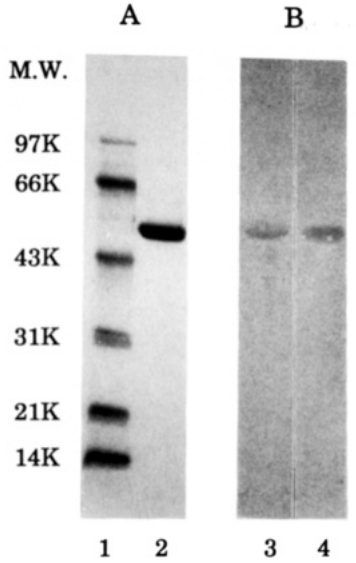

FIGURE 1: SDS–polyacrylamide gel electrophoresis (PAGE) of rat α_1-antitrypsin. Purified α_1-antitrypsin (5 μg) was resolved in a 7.5–15% gradient SDS–PAGE. The left (A) panel is protein staining with Coomassie blue. Protein molecular weight markers (lane 1); α_1-antitrypsin (lane 2). The right (B) panel is lectin immunoblot. α_1-Antitrypsin immunoblotted with antiserum to concanavalin A (lane 3); wheat germ agglutinin (lane 4).

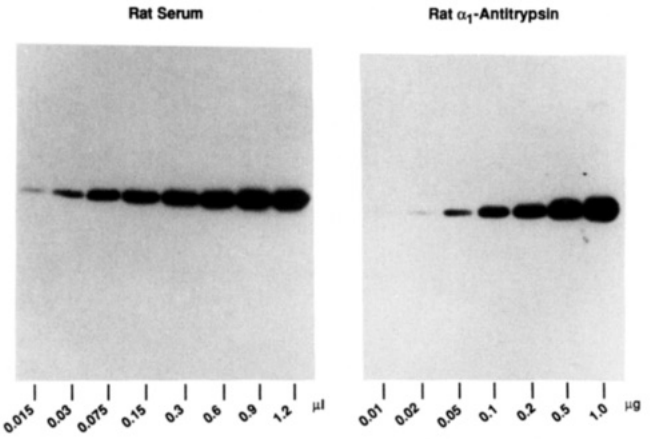

FIGURE 2: Western blot analysis of α_1-antitrypsin using an antigen-overlay method. Rat serum (0.015–1.2 μL as indicated in the left panel) and purified α_1-antitrypsin (0.01–1 μg as indicated in the right panel) were subjected to a 7.5–15% linear gradient SDS–PAGE under reducing conditions and then electrotransferred to nitrocellulose. Blots were incubated with antiserum against α_1-antitrypsin followed by ^{125}I-α_1-antitrypsin and exposed for 16 h.

of immunoreactive α_1-antitrypsin in rat serum is dose-dependent, with a minimum limit of 0.015 μL of serum and 10 ng of the purified α_1-antitrypsin, as shown in Figure 2, right panel. The results demonstrate both the specificity and the sensitivity of the immunoblot method for the detection of rat α_1-antitrypsin in complex biological fluids such as plasma. The calculated concentration of α_1-antitrypsin in rat serum, approximately 3 mg/mL, is similar to the reported level of human α_1-antitrypsin, 2–4 mg/mL serum (Daniels, 1975).

Isolation of cDNA Clones Encoding Rat α_1-Antitrypsin. A rat liver λgt-11 cDNA expression library was immunoscreened using polyclonal antiserum made against the purified rat α_1-antitrypsin as described previously (Chao et al., 1989a). In the primary screening, approximately 0.03% of the clones screened appeared positive. In the secondary screening, the 16 clones selected randomly from the primary screening were all positive. These positive clones were further plaque-purified and used for nucleotide sequence analysis.

Nucleotide Sequence Analysis and Characterization of Rat α_1-Antitrypsin cDNA. The positive clones obtained from the secondary screening were first examined for their cDNA insert size. The clone containing the longest insert [1300 base pairs (bp), as determined by agarose gel electrophoresis] was subcloned into the M13 mp19 sequencing vector and sequenced. The nucleotide and the translated protein sequences of rat α_1-antitrypsin are shown in Figure 3. The total length of the rat α_1-antitrypsin cDNA is 1306 bp, with a coding region of 1224 bp. When translated into the corresponding protein sequence, the cDNA encodes a protein with 408 amino acid residues. Both the nucleotide and the translated protein sequences were searched against the National Biomedical Research Foundation (NBRF) sequence data base. The algorithms used in the computer analysis were the FASTN (for nucleic acids) and the FASTP (for proteins) programs by Lipman and Pearson (1985). The k-tup values selected were 6 for nucleic acids and 2 for proteins. The computer analysis showed that the rat α_1-antitrypsin sequences share 70% identity with human α_1-antitrypsin and 80% identity with mouse α_1-antitrypsin at the nucleic acid and protein levels. The N-terminal 33 amino acid residues of the purified rat α_1-antitrypsin match completely with a segment of the translated protein sequence from the cDNA starting from residue 22, which is a Glu (as shown in the underlined region in Figure 3). The peptide bond between the Glu residue and the Ala residue N-terminal to Glu is the site of cleavage during processing to the mature protein. Although the N-terminal amino acid sequence of purified α_1-antitrypsin differs from a partial N-terminal 26 amino acid sequence (Gross et al., 1983) at Arg/Ser-18 and Ser/Pro-22, our protein sequence is confirmed by the cDNA sequence (Figure 3). As compared by aligning the rat and the human α_1-antitrypsin precursor sequences, the rat α_1-antitrypsin sequence is three amino acid residues (or seven base pairs at the nucleic acid level) short of the initiator Met (Figure 4). The cDNA clone thus covers the complete mature protein and the majority of the signal peptide. The rat α_1-antitrypsin cDNA contains a polyadenylation signal sequence of AATAAT at position 1212, downstream from the termination codon TAA at position 1162. The 3′-untranslated region is 77 bp in length with no poly(A) tail. The base composition of the cDNA is 26% A, 23% G, 23% T, and 28% C. The calculated molecular weight of the mature rat α_1-antitrypsin derived from the cDNA sequence is 43 700, which is smaller than the 54 000-dalton glycosylated α_1-antitrypsin existing in serum (Figures 1 and 2). In the translated protein sequence, five amino acid residues are identified as potential glycosylation sites. As indicated in Figure 3, they are Asn-40, Asn-77, Thr-79, Asn-241, and Thr-391. These glycosylation sites were identified by comparison with two consensus sequences, i.e., Asn-X-Ser/Thr and Ser/Thr-X-X-Pro (Marshall, 1974).

Sequence Comparison of Rat α_1-Antitrypsin with Human and Mouse α_1-Antitrypsin. In Figure 4, the translated protein sequence of rat α_1-antitrypsin is aligned with that of human α_1-antitrypsin (Long et al., 1984) and the partial protein sequence derived from mouse α_1-antitrypsin cDNA (Krauter et al., 1986). As compared with the human α_1-antitrypsin sequence, rat α_1-antitrypsin has two deletions, one of five residues and the other of one residue, in the N-terminal region. The mouse α_1-antitrypsin sequence has 1 insertion (Asn), 34 residues N-terminal from the reactive center P1 residue as compared with rat and human sequences. All three of the α_1-antitrypsins use a common Met–Ser peptide bond as the cleavage site by the target serine proteinases. However, when the amino acid residues of P1–P5′ are compared (boxed, Figure

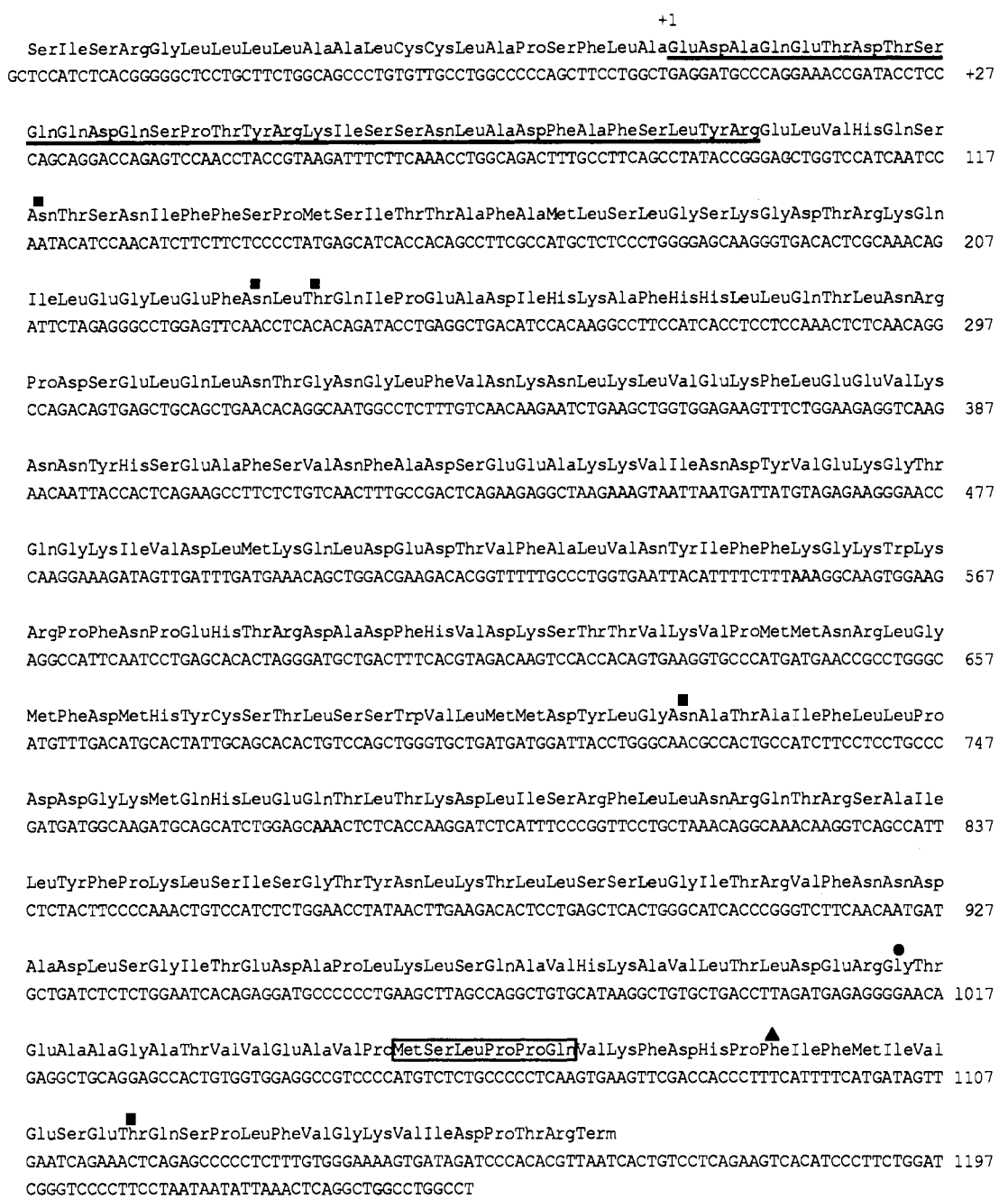

FIGURE 3: Translation of rat α_1-antitrypsin cDNA. The nucleotide sequence is bounded by *Eco*RI linkers. The numbering of the nucleotide sequence starts from the first nucleotide of the codon coding for the mature peptide N-terminal Glu residue. The underlined region is the sequence verified by the amino acid sequence analysis of purified rat α_1-antitrypsin. The boxed region is the reactive center including residues P1–P5′. Solid squares indicate potentially glycosylated residues. The solid circle indicates the "hinge" Gly residue, and the solid triangle indicates the conserved Phe residue in serpins C-terminal to the reactive center.

4), rat α_1-antitrypsin is found to be closer to the human α_1-antitrypsin because the replacements in the rat sequence, P2′ Leu-Ile and P5′ Gln-Glu, are replacements between similar amino acid residues while in the mouse sequence the replacements are less conservative: P2′ Met-Ile and P5′ Ile-Glu. Fourteen residues N-terminal from the P1 Met, in all three of the sequences, there exists a Gly residue surrounded by the sequence Glu-X-Gly-X-Glu. This Gly residue serves as the hinge residue for the stressed loop structure on which the reactive-site peptide bond is exposed (Carrell & Boswell, 1986). Seven residues C-terminal from the P5′ residue, a Phe residue is found in all three sequences. This Phe residue has been found in all of the known serpins and is believed to be functionally important (Carrell et al., 1987). The overall sequence identity between rat and human α_1-antitrypsin is close to 70%, and that between rat and mouse (with respect to the available sequence) is close to 80%.

Expression and Regulation of Rat α_1-Antitrypsin by Northern Blot Analysis. Tissue-specific expression of α_1-antitrypsin was analyzed by Northern blot analysis using a ^{32}P-labeled α_1-antitrypsin cDNA probe as shown in Figure 5. α_1-Antitrypsin mRNA can be detected at high levels in the liver (lanes 1–4), and elevated levels during acute-phase inflammation (lanes 1 vs 2) and a major difference in levels between male and female rats (lanes 3 vs 4) are evident. We detected low levels of rat α_1-antitrypsin mRNA in both the lung and the submandibular gland (lanes 5 and 6), indicating that the α_1-antitrypsin mRNA also is synthesized outside of the liver. In quantitative dot blot analysis of the rat liver total mRNA, the α_1-antitrypsin mRNA levels are found to be in-

```
                                  +1
RAT       SISRGLLLLAALCCLAPSFLAEDAQ-----ETDTSQQDQS-PTYRKISSNLADFAFSLYRELVHQSNTSNIFFSPMS  80
HUMAN  MPSSVSWGILLLAGLCCLVPVSLAEDPQGDAAQKTDTSHHDQDHPTFNKITPNLAEFAFSLYRQLAHQSNSTNIFFSPVS

RAT    ITTAFAMLSLGSKGDTRKQILEGLEFNLTQIPEADIHKAFHHLLQTLNRPDSELQLNTGNGLFVNKNLKLVEKFLEEVKN 130
HUMAN  IATAFAMLSLGTKADTHDEILEGLNFNLTEIPEAQIHEGFQELLRTLNQPDSQLQLTTDGGLFLSEGLKLVDKFLEDVKK

RAT    NYHSEAFSVNFADSEEAKKVINDYVEKGTQGKIVDLMKQLDEDTVFALVNYIFFKGKWKRPFNPEHTRDADFHVDKSTTV 210
HUMAN  LYHSEAFTVNFGDTEEAKKQINDYVEKGTQGKIVDLVKELDRDTVFALVNYIFFKGKWERPFEVKDTEEEDFHVDQVTTV
MOUSE                                                          KWKKPFDPENTEEAEFHVDESTTV

RAT    KVPMMNRLGMFDMHYCSTLSSWVLMMDYLGNATAIFLLPDDGKMQHLEQTLTKDLISRFLLNRQTRSAILYFPKLSISGT 290
HUMAN  KVPMMKRLGMFNIQHCKKLSSWVLLMKYLGNANAIFFLPDEGKLQHLENELTHDIITKFLENEDRRSASLHLPKLSITGT
MOUSE  KVPMMTLSGMLDVHHCSTLSSWVLLMDYAGNATAVFLLPDDGKMQHLEQTLSKELISKFLLNRRRRLAQIHFPRLSISGE

RAT    YNLKTLLSSLGITRVFNNDADLSGITEDA-PLKLSQAVHKAVLTLDEKGTEAAGATVVEAVPMSLPPQVKFDHPFIFMIV 369
HUMAN  YDLKSVLGQLGITKVFSNGADLSGVTEEA-PLKLSKAVHKAVLTIDEKGTEAAGAMFLEAIPMSIPPEVKFNKPFVFLMI
MOUSE  YNLKTLMSLLGITRIFNNGADLSGITEEANPLKLSQAVHKAVLTIDETGTEAAAVTVLQMVPMSMPPIILRFDHPFLFIIF

RAT    ESETQSPLFVGKVIDPTR 387
HUMAN  EQNTKSPLFMGKVVNPTQK
MOUSE  EEHTQSPIFVGKVVDPTHK
```

FIGURE 4: Comparison between the translated rat α_1-antitrypsin sequence and its human and mouse counterparts. The numbering of the rat sequence starts from the mature peptide N-terminal Glu residue. The "hinge" sequence, the reactive-center sequence, and the conserved Phe (F) residue C-terminal to the reactive center are boxed. Dashes indicate deletions. The arrow indicates the cleavage site for the signal peptide.

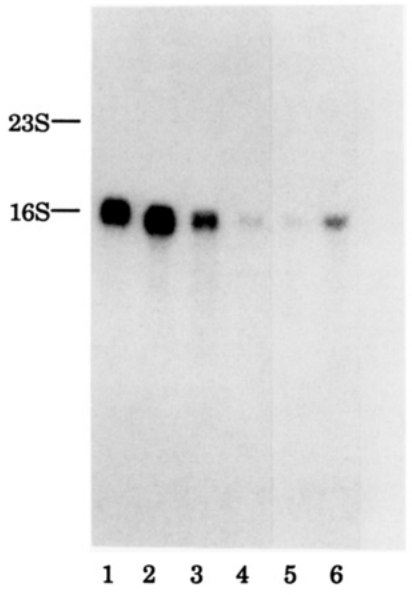

FIGURE 5: Northern blot analysis of α_1-antitrypsin in male and female Sprague–Dawley rat tissues. Total male rat liver RNA (10 μg of each) from control (lane 1) and after induction by acute-phase inflammation caused by turpentine as described previously (Chao et al., 1989b) (lane 2). Total liver RNA from male (lane 3) and female (lane 4) rats. Note that the first two lanes and the middle two lanes are experimentally separate. The rats represented in lanes 1 and 2 are age-matched, and the RNA from these rats was prepared separately from the RNA from the rats represented in lanes 3 and 4, which were also separately age-matched. Lung RNA (115 μg, lane 5) and submandibular RNA (82 μg, lane 6). Lanes 1–4 were developed after 6 h, and lanes 5 and 6 were developed after 3 days.

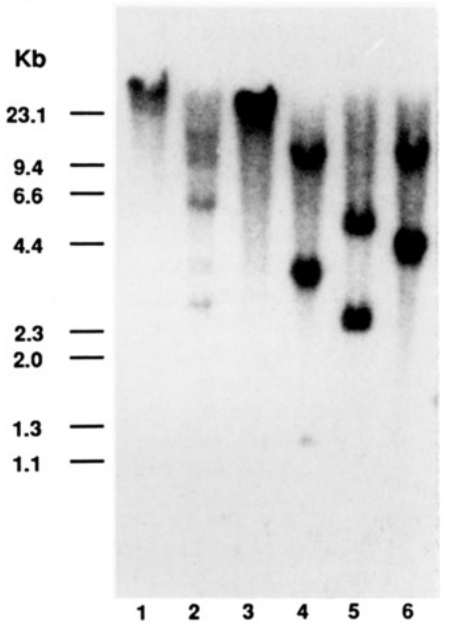

FIGURE 6: Southern blot hybridization of rat genomic DNA. Approximately 10 μg of rat DNA was digested with *Bam*HI (lane 1), *Bgl*II (lane 2), *Eco*RI (lane 3), *Hin*dIII (lane 4), *Sst*I (lane 5), and *Stu*I (lane 6).

duced up to 2-fold after acute-phase inflammation and to be 5-fold higher in male than in female rats (data not shown).

Southern Blot Hybridization of Rat Genomic DNA. As shown in Figure 6, Southern blot hybridization of rat genomic DNA with the rat α_1-antitrypsin cDNA probe showed one band (30 kb) when the DNA was digested with *Bam*HI, four bands (total size 35 kb) with *Bgl*II, one band (30 kb) with *Eco*RI, three bands (total size 19 kb) with *Hin*dIII, two bands (total size 8 kb) with *Sst*I, and two bands (total size 15 kb) with *Stu*I.

Discussion

We have isolated rat α_1-antitrypsin from serum and have generated antiserum against the purified protein. The antibodies against α_1-antitrypsin were used to identify cDNA clones encoding α_1-antitrypsin from a rat liver expression cDNA library with the antigen-overlay method (Chao et al., 1989a). A cDNA clone of 1306 bp encoding the α_1-antitrypsin

precursor was isolated following primary and secondary screening. The identity of the cDNA clone was established by matching its translated sequence with the N-terminal 33 amino acid sequence of the purified protein and by sequence comparisons with human and mouse α_1-antitrypsin. Nucleotide sequence analysis showed that the cDNA encodes the complete, mature α_1-antitrypsin peptide and a part of the signal peptide. An initiation codon cannot be identified at the 5′ end, however, as compared with the protein sequence of the human α_1-antitrypsin precursor; the cDNA clone is seven base pairs short of encoding a full-length signal peptide.

In the reactive-center region of rat α_1-antitrypsin, a Met residue could serve as the residue which is attacked by the active-site Ser residue of the target serine proteinase. This feature, which is the same in human and mouse α_1-antitrypsin, renders the inhibitor susceptible to oxidation at the reactive site, hence causing the loss of its ability to inhibit the target enzyme. The fact that both human and rat α_1-antitrypsin have a Met–Ser bond as their cleavage site in the reactive center suggests that rat α_1-antitrypsin may react with tissue kallikrein as well as its human counterpart does. However, in our protein assays following prolonged incubation for 24 h at 37 °C, only human α_1-antitrypsin was found to be able to form SDS-stable complexes with human tissue kallikrein (urinary kallikrein). No interaction between rat α_1-antitrypsin and rat tissue kallikrein was detectable under the same conditions (data not shown). The lack of interaction between rat α_1-antitrypsin and rat tissue kallikrein is an unexpected result, the cause of which is an open question. The protein sequence of rat α_1-antitrypsin is significantly (70%) similar to that of human α_1-antitrypsin, and both contain only one Cys residue (residue 256), indicating that no intramolecular disulfide bond exists; therefore, we would expect the folding of rat α_1-antitrypsin to be similar to that of human α_1-antitrypsin. However, even with similar folding, replacement of amino acid residues in the reactive-center region (P2′ Leu-Ile, P5′ Gln-Glu) could be responsible for the lack of binding between rat α_1-antitrypsin and rat tissue kallikrein.

It has been well established that α_1-antitrypsin is synthesized primarily in the liver and in limited amounts in human macrophages (Rogers et al., 1983). Our Northern and dot blot analysis using an α_1-antitrypsin cDNA probe indicates that rat α_1-antitrypsin mRNAs can be detected in two new sites, the submandibular gland and the lung, and the levels in these tissues are estimated to be more then 100-fold lower than those in the liver. We have shown by Western blot analysis that both the lung and the submandibular gland have immunoreactive α_1-antitrypsin of the same molecular weight as that of the protein purified from rat sera (data not shown). Particularly interesting is the finding of the in situ synthesis of α_1-antitrypsin in the lung, which is the site of emphysema development when α_1-antitrypsin is inactivated by oxidation (Heidtmann & Travis, 1986). The synthesis of α_1-antitrypsin in the lung could then serve as a backup source of α_1-antitrypsin. However, the role of the extrahepatic synthesis of α_1-antitrypsin is unclear at the present time.

Human α_1-antitrypsin is known to be one of the acute-phase proteins in plasma (Chandra et al., 1981). Both Northern and dot blot analyses show that rat α_1-antitrypsin mRNA levels increase during acute-phase inflammation (Figure 5). Using a newly developed radioimmunoassay for rat α_1-antitrypsin, we also found that α_1-antitrypsin concentrations in the serum increase 2-fold after acute-phase inflammation (data not shown). Furthermore, Northern blot analysis showed that α_1-antitrypsin mRNA levels in the liver of male rats are severalfold higher than those in female rats (Figure 5, lanes 3 and 4). The implications of both the sexual dimorphism and inflammatory induction of α_1-antitrypsin gene expression in the liver remain to be investigated. When rat genomic DNA was digested with the six restriction endonucleases indicated in Figure 6 and probed with rat α_1-antitrypsin cDNA, a rather simple banding pattern was observed, suggesting that the rat α_1-antitrypsin gene is single-copy.

Registry No. α_1-Antitrypsin (rat protein moiety), 124041-98-1; α_1-antitrypsin, 9041-92-3.

References

Burnett, W. N. (1981) *Anal. Biochem. 112*, 195–203.

Carrell, R. W., & Boswell, D. R. (1986) *Proteinase Inhibitors*, pp 403–419, Elsevier Science Publishers BV, Amsterdam, The Netherlands.

Carrell, R. W., Pemberton, P. A., & Boswell, D. R. (1987) *Cold Spring Harbor Symp. Quant. Biol. 52*, 527–535.

Chandra, T., Kurachi, K., Davie, E. W., & Woo, S. L. C. (1981) *Biochem. Biophys. Res. Commun. 2*, 751–758.

Chao, J., & Chao, L. (1988) *J. Hypertension 6*, 551–557.

Chao, J., Tillman, D. M., Wang, M., Margolius, H. S., & Chao, L. (1986a) *Biochem. J. 239*, 325–331.

Chao, J., Chao, L., Woodley, C. M., Gerald, W. L., & Margolius, H. S. (1986b) *Adv. Exp. Med. Biol. 198A*, 181–187.

Chao, J., Swain, C., Chao, S., Xiong, W., & Chao, L. (1988) *Biochim. Biophys. Acta 964*, 329–339.

Chao, S., Chao, L., & Chao, J. (1989a) *BioTechniques 7*, 68–71.

Chao, S., Chao, L., & Chao, J. (1989b) *Biochim. Biophys. Acta 991*, 477–483.

Chen, Y., Chao, J., & Chao, L. (1988) *Biochemistry 27*, 7189–7196.

Daniels, J. C. (1975) in *Serum Protein Abnormalities: Diagnostic and Clinical Aspects* (Ritzmann, S. E., & Daniels, J. C., Eds.) pp 243–263, Little, Brown and Company, Boston.

Geiger, R., Stuckstedte, U., Clausnitzer, B., & Fritz, H. (1981a) *Hoppe Seyler's Z. Physiol. Chem. 362*, 317–325.

Geiger, R., Konig, G., & Fruhmann, G. (1981b) *Hoppe-Seyler's Z. Physiol. Chem. 362*, 1013–1015.

Gross, V., Kaiser, C., Tran-Thi, T., Schmelzer, E., Witt, I., Plummer, T. H., Jr., & Heinrich, P. C. (1983) *FEBS Lett. 2*, 201–205.

Heeb, M. J., & Griffin, J. H. (1988) *J. Biol. Chem. 263*, 11613–11616.

Heidtmann, H., & Travis, J. (1986) in *Proteinase Inhibitors*, pp 441–456, Elsevier Science Publishers BV, Amsterdam, The Netherlands.

Hirano, K., Okumura, Y., Hayakawa, S., Adachi, T., & Sugiura, M. (1984) *Hoppe-Seyler's Z. Physiol. Chem. 365*, 27–32.

Horl, W. H., & Heidland, A. (1981) *Klin. Wochenschr. 59*, 761–763.

Horl, W. H., Schafer, R. M., & Heidland, A. (1982) *Eur. J. Clin. Pharmacol. 22*, 541–544.

Johnson, D. A., Gautxch, J. W., Spoasman, J. D., & Elder, J. H. (1984) *Gene Anal. Tech. 1*, 3–8.

Krauter, K. S., Citron, B. A., Hsu, M. T., Powell, D., & Darnell, J. E., Jr. (1986) *DNA 5*, 29–36.

Lammle, B., Berrettini, M., & Griffin, J. H. (1986) *Anal. Biochem. 156*, 118–125.

Latimer, J. J., Berger, F. G., & Baumann, H. (1987) *J. Biol. Chem. 26*, 12641–12646.

Lipman, D. J., & Pearson, W. R. (1985) *Science 227*, 1435.

Loebermann, H., Tokuoka, R., Deisenhofer, J., & Huber, R. (1984) *J. Mol. Biol. 177*, 531-556.
Long, G. L., Chandra, T., Woo, S. L. C., Davie, E. W., & Kurachi, K. (1984) *Biochemistry 23*, 4828-4837.
Lu, H., Lin, F., Chao, L., & Chao, J. (1989) *Int. J. Pept. Protein Res. 33*, 237-249.
Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning—A Laboratory Manual*, pp 280-281, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Marshall, R. D. (1974) *Biochem. Soc. Symp. 40*, 17-26.
Messing, J. (1983) *Methods Enzymol. 101*, 20-77.
Monaci, P., Nicosia, A., & Cortese, R. (1988) *EMBO J. 7*, 2075-2087.
Murray, S., Chao, J., Lin, F., & Chao, L. (1989) *J. Cardiovasc. Pharmacol.* (in press).
Pannell, R., Johnson, D., & Travis, J. (1974) *Biochemistry 13*, 5349-5445.
Rogers, J., Kalsheker, N., Wallis, S., Speer, A., Coutelle, C. H., Woods, D., & Humphries, S. E. (1983) *Biochem. Biophys. Res. Commun. 116*, 375-382.
Sanger, F., Nicklen, S., & Coulson, A. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463-5467.
Shai, S., Woodley-Miller, C., Chao, J., & Chao, L. (1989) *Biochemistry 28*, 5334-5343.
Shimamoto, K. Chao, J., & Margolius, H. S. (1980) *J. Clin. Endocrinol. Metab. 51*, 840-848.
Travis, J., & Salvesen, G. S. (1983) *Annu. Rev. Biochem. 52*, 655-709.
Young, R. A., & Davis, R. W. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 1194-1198.

High-Resolution Structure of an HIV Zinc Fingerlike Domain via a New NMR-Based Distance Geometry Approach†

Michael F. Summers,*,‡ Terri L. South,‡ Bo Kim,‡ and Dennis R. Hare*,§
Department of Chemistry and Biochemistry, University of Maryland Baltimore County, Baltimore, Maryland 21228, and Hare Research, Inc., 14810 216th Avenue N.E., Woodinville, Washington 98072
Received May 11, 1989; Revised Manuscript Received August 26, 1989

ABSTRACT: A new method is described for determining molecular structures from NMR data. The approach utilizes 2D NOESY back-calculations to generate simulated spectra for structures obtained from distance geometry (DG) computations. Comparison of experimental and back-calculated spectra, including analysis of cross-peak buildup and auto-peak decay with increasing mixing time, provides a quantitative measure of the consistence between the experimental data and generated structures and allows for use of tighter interproton distance constraints. For the first time, the "goodness" of the generated structures is evaluated on the basis of their consistence with the actual experimental data rather than on the basis of consistence with other generated structures. This method is applied to the structure determination of an 18-residue peptide with an amino acid sequence comprising the first zinc fingerlike domain from the *gag* protein p55 of HIV. This is the first structure determination to atomic resolution for a retroviral zinc fingerlike complex. The peptide [Zn(p55F1)] exhibits a novel folding pattern that includes type I and type II NH-S tight turns and is stabilized both by coordination of the three Cys and one His residues to zinc and by extensive internal hydrogen bonding. The backbone folding is significantly different from that of a "classical" DNA-binding zinc finger. Residues C(1)-F(2)-N(3)-C(4)-G(5)-K(6) fold in a manner virtually identical with the folding observed by X-ray crystallography for related residues in the iron domain of rubredoxin; superposition of all main-chain and Cys side-chain atoms of residues C(1)-K(6) of Zn(p55F1) onto residues C(6)-Y(11) and C(39)-V(44) of rubredoxin gives RMSDs of 0.46 and 0.35 Å, respectively. The side chains of conservatively substituted Phe and Ile residues implicated in genomic RNA recognition form a hydrophobic patch on the peptide surface.

During assembly and budding stages of the retroviral life cycle, a *gag* polyprotein identifies and forms a complex with viral RNA (Dickson et al., 1985; Bolognesi et al., 1978). Subsequent to transporting the RNA to the cell wall for budding, the *gag* polyprotein is cleaved to give, among other products, a low molecular weight nucleic acid binding protein (NABP).[1] Without exception, retroviral *gag* polyproteins and their NABP products contain either one or two conserved sequences of the type C-X_2-C-X_4-H-X_4-C (Henderson et al., 1981; Copeland et al., 1984; Karpel et al., 1987). Berg has

†This work was supported by ACS Institutional Research Grant IN-147F, by the MD Cancer Program/University of Maryland (M.F.S.), and by NIH Grants GM42561-01 (M.F.S.) and GM35620-02 (D.R.H.).
*Author to whom correspondence should be addressed. Reprint requests should be addressed to M.F.S.
‡University of Maryland Baltimore County.
§Hare Research, Inc.

[1] Abbreviations: 2D NMR, two-dimensional nuclear magnetic resonance; 2QF-COSY, double quantum filtered correlated spectroscopy; BKCALC, NOESY back-calculation program; DG, distance geometry; DSPACE, distance geometry program; HIV, human immunodeficiency virus; HMQC, heteronuclear multiple quantum coherence spectroscopy; HOHAHA, homonuclear Hartmann-Hahn spectroscopy; HS, homospoil pulse; HTLV, human T-cell leukemia virus; k_{CR}, parameter used by BKCALC which governs the rate of cross relaxation; k_{ZL}, parameter used by BKCALC to control z leakage rates; MuLV, murine leukemia virus; NABP, nucleic acid binding protein; NOESY, nuclear Overhauser effect spectroscopy; RMSD, root-mean-square deviation; ROESY, rotating-frame Overhauser effect spectroscopy.

0006-2960/90/0429-329$02.50/0