# Solution Structure of the Amino-Terminal Fragment of Urokinase-Type Plasminogen Activator‡

Andrew P. Hansen, Andrew M. Petros, Robert P. Meadows, David G. Nettesheim, Andrew P. Mazar, Edward T. Olejniczak, Robert X. Xu, Terry M. Pederson, Jack Henkin, and Stephen W. Fesik*

*Pharmaceutical Discovery Division, Abbott Laboratories, Abbott Park, Illinois 60064*

*Received December 23, 1993; Revised Manuscript Received February 22, 1994*®

ABSTRACT: The amino-terminal fragment (ATF) of urokinase-type plasminogen activator is a two domain protein which consists of a growth factor and a kringle domain. The $^{1}H$, $^{13}C$, and $^{15}N$ chemical shifts of this protein have been assigned using heteronuclear two- and three-dimensional NMR experiments on selective and uniformly $^{15}N$- and $^{15}N/^{13}C$-labeled protein isolated from mammalian cells that overexpress the protein. The chemical shift assignments were used to interpret the NOE data which resulted in a total of 1299 NOE restraints. The NOE restraints were used along with 27 $\phi$ angle restraints and 21 hydrogen-bonding restraints to produce 15 low energy structures. The individual domains in the structures are highly converged, but the two domains are structurally independent. The root mean square deviations (rmsd) between residues 11–46 in the growth factor domain and the mean atomic coordinates were 0.99 ± 0.2 for backbone heavy atoms and 1.65 ± 0.2 for all non-hydrogen atoms. For residues 55–130 in the kringle domain, the rmsd was 0.84 ± 0.2 for backbone heavy atoms and 1.42 ± 0.2 for all non-hydrogen atoms. The overall structures of the individual domains are very similar to the structures of homologous proteins. However, important structural differences between the growth factor and other homologous proteins were observed in the region which has been implicated in binding the urokinase receptor which may explain, in part, why other growth factors show no appreciable affinity for the urokinase receptor.

Urokinase-type plasminogen activator (u-PA)[1] is a 54-kDa glycoprotein that catalyzes the conversion of plasminogen to plasmin, a broad specificity protease responsible for the degradation of fibrin clots and extracellular matrix components such as fibrin, fibronectin, and laminin (de Munk & Rijken, 1990). u-PA is used clinically for its ability to speed the dissolution of blood clots in the treatment of pulmonary embolisms, myocardial infarction, and peripheral arterial occlusion (Haber et al., 1989). In addition, u-PA plays an important role in tissue remodeling and metastatic tissue invasion (Ossowski, 1988; Crowley et al., 1993; Jankun et al., 1993). Indeed, high levels of u-PA found at tumor sites have been shown to correlate with the metastatic potential of tumors and decreased survival in patients with certain forms of cancer [for a review, see Duffy (1993)].

The structure of u-PA consists of three individual modules: a growth factor domain (GFD), a kringle domain, and a serine protease domain. The N-terminal GFD is responsible for localizing u-PA on the cell surface by binding to the u-PA receptor (Appella et al., 1987). This interaction is critical for tumor cell metastasis and is specific for the growth factor domain of u-PA since other homologous growth factors (e.g., epidermal growth factor) do not show appreciable binding affinity for the urokinase receptor (Vassalli et al., 1985). The function of the kringle domain is less clear. Unlike the homologous kringles of plasminogen and tissue plasminogen activator (t-PA), the u-PA kringle does not bind to fibrin. Instead, the u-PA kringle has been shown to bind to heparin, suggesting a possible role in binding polyanionic molecules such as the proteoglycans which could localize u-PA at the site of tissue invasion (Stephens et al., 1992). The C-terminal serine protease domain functions as the catalytic portion of u-PA that activates plasminogen to plasmin. It is homologous to other serine proteases such as trypsin and chymotrypsin but exhibits a different substrate and inhibitor specificity (Towle et al., 1993).

Structural information on u-PA is important to further our understanding of the specific structural features that define its biological functions. Knowledge of the three-dimensional structure of u-PA is of particular importance due to its critical role in metastatic tissue invasion. Structural information on u-PA and its molecular complexes could aid in the design of antimetastatic agents that inhibit the enzymatic activity of u-PA or block the interaction of u-PA with the urokinase receptor. However, no three-dimensional structure of intact u-PA or any of its fragments have thus far been reported. Although three-dimensional structures of growth factors (Montelione et al., 1987; Cooke et al., 1987; Kline et al., 1990; Moy et al., 1993), kringles (de Vos et al., 1992; Mulichak et al., 1991), and serine proteases (Birktoft & Blow, 1972; Tulinsky et al., 1973; Huber et al., 1974; Stroud et al., 1974; Shotten & Watson, 1970; Sawyer et al., 1978) that are homologous to the individual domains of u-PA have been previously determined, the three-dimensional structure of u-PA may differ in key areas, since its biological functions and binding specificities are vastly different from the homologous proteins whose structures have been solved.

* Author to whom correspondence should be addressed.

‡ The coordinates of the NMR-derived solution structure of the amino-terminal fragment of urokinase-type plasminogen activator have been deposited with the Brookhaven Protein Data Bank with the identity code 1URK.

® Abstract published in *Advance ACS Abstracts,* April 1, 1994.

[1] Abbreviations: u-PA, urokinase-type plasminogen activator; GFD, growth factor domain; EGF, epidermal growth factor; t-PA, tissue-type plasminogen activator; NMR, nuclear magnetic resonance; 2D, two dimensional; ATF, amino-terminal fragment of u-PA; 3D, three dimensional; NOE, nuclear Overhauser enhancement; NOESY, nuclear Overhauser enhancement spectroscopy; HSQC, heteronuclear single-quantum coherence; TOCSY, total correlation spectroscopy; HMQC, heteronuclear multiple-quantum coherence; CT, constant time; DG, distance geometry; SA, simulated annealing; rmsd, root mean square deviation.

0006-2960/94/0433-4847$04.50/0 

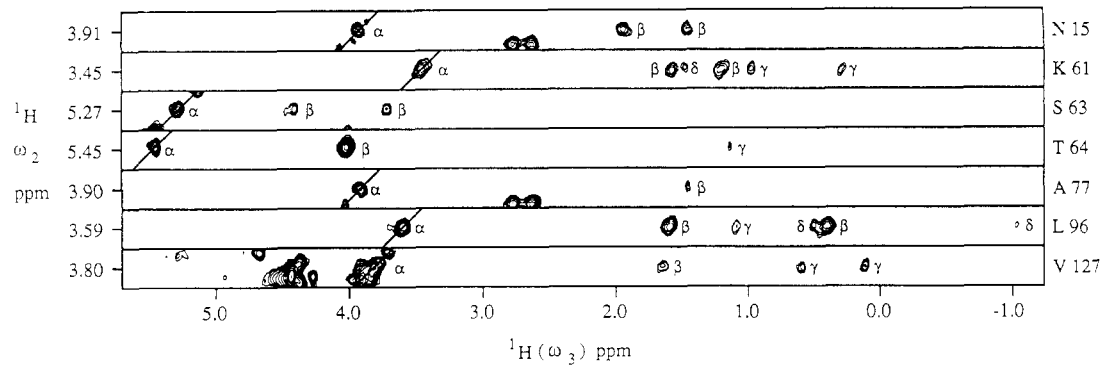


FIGURE 1: $^1$H($\omega_2$), $^1$H($\omega_3$) planes extracted from a 3D HCCH-TOCSY spectrum at the $^{13}$C chemical shifts of selected $\alpha$ carbons.

In earlier NMR studies using 2D homonuclear NMR methods, Dobson and co-workers assigned the $^1$H NMR resonances and determined the secondary structure of the u-PA kringle (Li et al., 1992). In another study, Bokman et al. (1993) reported on the $^1$H and $^{13}$C NMR assignments of the aromatic side chains of the u-PA kringle. In addition, they have analyzed the changes in chemical shift of the aromatic resonances of the u-PA kringle upon the addition of heparin in an effort to probe the heparin binding site. Other investigations have focused on interdomain interactions within u-PA as studied by differential scanning calorimetry (Novokhatny et al., 1992, 1993) and $^1$H NMR (Bogusky et al., 1989; Nowak et al., 1993).

In this paper we present the $^1$H, $^{13}$C, and $^{15}$N assignments and three-dimensional structure of the amino-terminal fragment (ATF) of u-PA (residues 6–135), which consists of the growth factor and kringle domains. In addition to characterizing the structures of the two individual domains, we have examined the possibility of interdomain interactions. The assignments and structure of the amino-terminal fragment of u-PA were obtained from an analysis of several heteronuclear multidimensional NMR data sets recorded on uniformly $^{15}$N- and $^{15}$N-/$^{13}$C-labeled ATF isolated from mammalian cells that overexpress the protein (Hansen et al., 1992). Although several examples have appeared in the literature on the three-dimensional structure determination of isotopically labeled proteins from bacteria (Clore & Gronenborn, 1991; Ikura et al., 1992; Clubb et al., 1991; Meadows et al., 1993; Theriault et al., 1993), to our knowledge, this represents the first example of a structure determination of a protein that has been expressed and isotopically labeled in mammalian cells.

## MATERIALS AND METHODS

*Sample Preparation.* Recombinant u-PA was cloned into Sp2/0 cells and expressed in mammalian cell culture (Lo & Gillies, 1991). Uniformly $^{15}$N-labeled and [U-$^{13}$C,$^{15}$N]u-PA was obtained by expressing u-PA in cell culture media containing labeled amino acids from bacteria or algae that have been supplemented with isotopically labeled glutamine and cysteine (Hansen et al., 1992). [$^{15}$N]Gln and [U-$^{13}$C,$^{15}$N]-Gln were enzymatically synthesized from [$^{15}$N]Glu or [U-$^{13}$C,$^{15}$N]Glu and $^{15}NH_4Cl$ (Hansen et al., 1992). [$^{15}$N]-Cys was purchased from Cambridge Isotopes and was added to both [U-$^{15}$N] and [U-$^{13}$C,$^{15}$N] media since doubly labeled cysteine was not available. u-PA was purified from the harvested media and subjected to limited autoproteolysis to generate the amino-terminal fragment (residues 6–135) as previously described (Mazar et al., 1992; Hansen et al., 1992). The [U-$^{15}$N]ATF and [U-$^{13}$C,$^{15}$N]ATF NMR samples ($\sim$1 mM) were prepared by dissolving the purified protein in 400 $\mu$L of sodium acetate-$d_3$ buffer (pH 4.5, 90% $H_2O$/10% $D_2O$). For the 3D HCCH-TOCSY and $^{13}$C-resolved NOESY experiments, a 2 mM sample of [U-$^{13}$C,$^{15}$N]ATF sample was prepared in a microcell (Shigemi) in 150 $\mu$L of $D_2O$ solution containing sodium acetate-$d_3$ (pD 4.5).

Four additional ATF samples were prepared that were selectively $^{15}$N-labeled by amino acid type. These samples were produced by incorporating an $^{15}$N-labeled amino acid (Lys, Leu, Val, and Phe) into the cell culture media containing all of the other amino acids in their unlabeled form. The relative concentration of each amino acid was the same as found in commercial media (Hansen et al., 1992). The selectively labeled samples were purified and prepared in the same manner as the uniformly labeled samples. A fifth ATF sample was prepared by incorporating both (*R*)-[2-$^2$H,$^{15}$N]-Gly and $^{15}$N-labeled arginine. (*R*)-[2-$^2$H,$^{15}$N]Gly was synthesized as described by Curley et al. (1994).

*NMR.* All NMR spectra were collected at 30 °C on Bruker AMX 500 and Bruker AMX 600 NMR spectrometers. NMR spectra were processed and analyzed using in-house written software on Silicon Graphics computers.

The $^1$H–$^{15}$N heteronuclear single-quantum correlation (HSQC) spectra (Bodenhausen & Ruben, 1980) were acquired on the AMX 500 using 256 complex points in $t_1$ and 2048 complex points in $t_2$, with 16 scans per $t_1$ experiment. The sweep width was 8333 Hz in $\omega_2$ ($^1$H) and 1773 Hz in $\omega_1$ ($^{15}$N).

For the $^{15}$N-resolved 3D NOESY-HSQC experiments (Fesik & Zuiderweg, 1988; Marion et al., 1989a), 48($t_1$) $\times$ 128($t_2$) $\times$ 1024($t_3$) complex points were acquired using spectral widths of 2100 Hz ($\omega_1$, $^{15}$N), 7692 Hz ($\omega_2$, $^1$H), and 10 000 Hz ($\omega_3$, $^1$H) and 24 scans per increment. Three 3D NOESY-HSQC experiments were recorded with mixing times of 50, 100, and 150 ms. Water suppression was accomplished using a homospoil pulse during the mixing time and spin-lock pulses according to the method of Messerle et al. (1989).

Two $^{15}$N-resolved 3D TOCSY-HSQC (Marion et al., 1989b) spectra were acquired with essentially the same parameters as the $^{15}$N-resolved NOESY-HSQC experiment. TOCSY mixing times of 22 and 44 ms (clean MLEV) (Griesinger et al., 1988) were used in the two experiments.

The HNHB experiment (Archer et al., 1991) was acquired with spectral widths of 1773, 5208, and 8333 Hz and 48 $\times$ 96 $\times$ 2048 complex points in $t_1$($^{15}$N), $t_2$($^1$H), and $t_3$($^1$H), respectively. The data were acquired on the AMX 500 with 16 scans per increment for a total experimental time of 82 h.

The $^3J_{H^N,H^\alpha}$ coupling constants were measured from 2D HMQC-J (Kay & Bax, 1990) and 3D HNHA (Vuister & Bax, 1993) spectra. The 2D HMQC-J experiment was collected using sweep widths of 1748 and 8333 Hz and 896 $\times$ 2048 complex points in $\omega_1$($^{15}$N) and $\omega_2$($^1$H). The 3D HNHA experiment was collected with sweep widths of 1748, 6410, and 8333 Hz in $\omega_1$($^{15}$N), $\omega_2$($^1$H), and $\omega_3$($^1$H), respec-

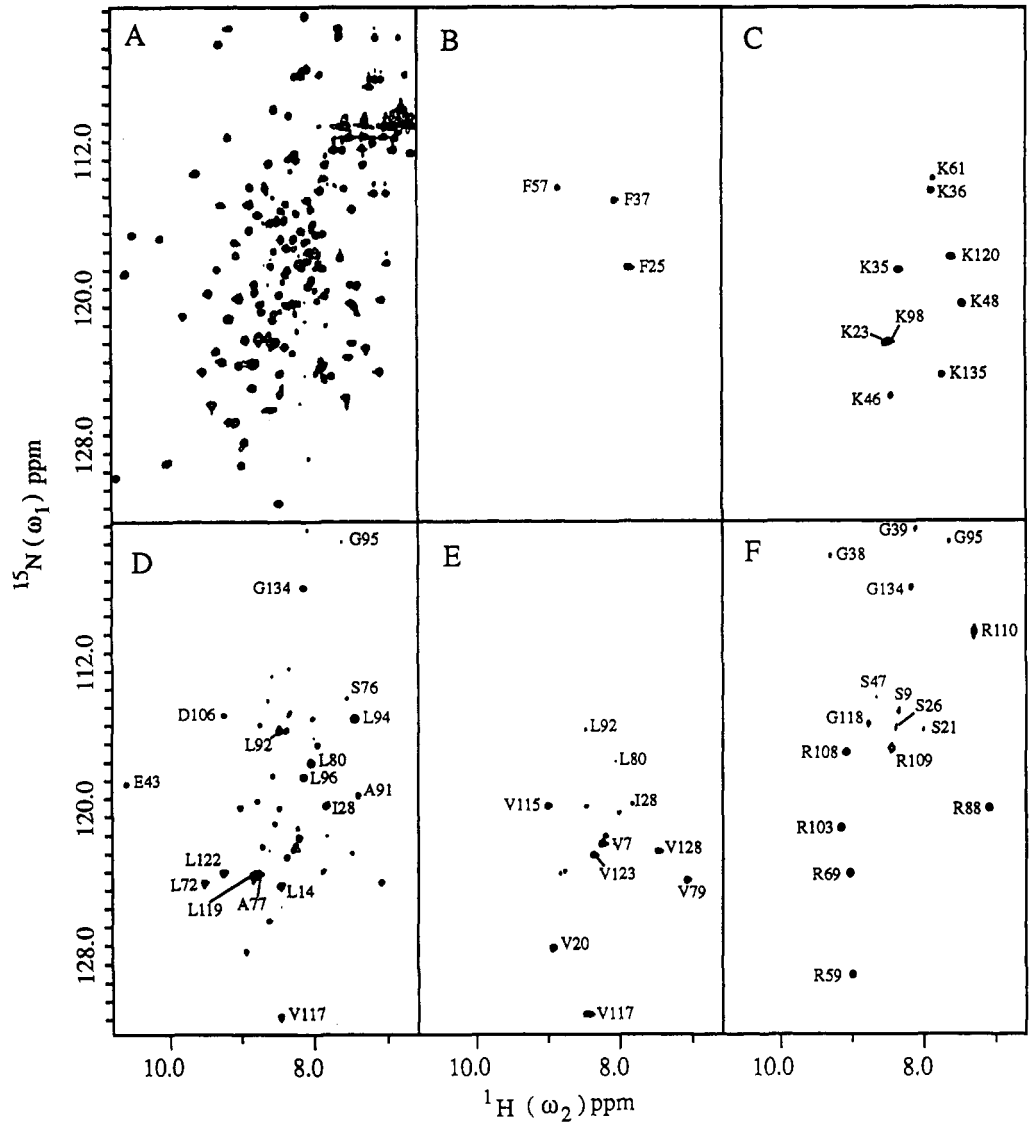


FIGURE 2: $^{1}H$-$^{15}N$ HSQC spectra of (A) [U-$^{15}N$], (B) [$^{15}N$-Phe], (C) [$^{15}N$-Lys], (D) [$^{15}N$-Leu], (E) [$^{15}N$-Val], and (F) [$^{15}N$-Arg; (*R*)-2-$^{2}H$, $^{15}N$-Gly]ATF showing the amide assignments of the selectively labeled amino acids and representative examples of mixing to other amino acid types.

tively. The delay during which the $^{3}J_{H^N,H^\alpha}$ coupling is active was set to 26.0 ms.

The 3D HNCA experiment (Kay et al., 1990) and 3D HN(CO)CA experiment (Bax & Ikura, 1991) were acquired with oversampling in $\omega_3$ using identical sweep widths of 1773 Hz ($\omega_1$, $^{15}N$), 3019 Hz ($\omega_2$, $^{13}C$), and 33 333 Hz ($\omega_3$, $^{1}H$). The HNCA was acquired in 80 scans per increment with 39 × 36 × 4096 complex points, and the HN(CO)CA experiment was acquired with 112 scans per increment and 36 × 32 × 4096 complex points.

The 3D CBCA(CO)NH experiment (Grzesiek & Bax, 1992) was collected using sweep widths of 1773, 6792, and 33333 Hz and 48 × 24 × 4096 complex points with 32 scans per increment.

The 3D HCCH-TOCSY experiments (Bax et al., 1990) were acquired using the AMX 600. For two of the experiments, the carbon carrier was placed at 41.25 ppm, and the data were collected using sweep widths of 5055, 6250, and 10 000 Hz over 40 × 120 × 1024 complex points in $\omega_1$($^{13}C$), $\omega_2$($^{1}H$), and $\omega_3$($^{1}H$), respectively. Mixing times of 10 and 16 ms using a DIPSI-2 pulse sequence were used in the two experiments.

In order to identify the *J*-coupled signals in the aromatic region of the spectrum, a third 3D HCCH-TOCSY experiment was recorded with the $^{13}C$ carrier set at 125 ppm. In this experiment the data were collected using a mixing time of 9 ms and sweep widths of 5055, 7246, and 10 000 Hz over 40 × 96 × 1024 complex points.

Three 3D $^{13}C$-resolved NOESY-HSQC experiments (Ikura et al., 1990; Zuiderweg et al., 1990) were collected using the AMX 600 with mixing times of 80, 100, and 140 ms. Gradients were used to suppress artifacts and the residual water signal (Bax & Pochapsky, 1992). The carbon carrier was placed at 58.0 ppm, and sweep widths of 5055, 6250, and 10 000 Hz over 38 × 128 × 1024 complex points were used.

The $^{3}J_{C',C\gamma}$ coupling constants were measured using a 2D $^{13}C$-{$^{13}CO$} spin-echo difference CT-HSQC experiment (Grzesiek et al., 1993). A total of 1024 × 8333 complex points were collected with sweep widths of 12 500 Hz ($\omega_1$, $^{13}C$) and 8333 Hz ($\omega_2$, $^{1}H$).

*NOE-Derived Distance Restraints.* NOE cross peaks from the 3D $^{15}N$- and $^{13}C$-resolved NOESY spectra were categorized as strong, medium, or weak according to their integrated volumes and given lower bounds of 1.8 Å and upper bounds of 3.0, 4.0, and 5.0 Å, respectively. The intensity of NOEs between protons separated by known distances in rigid structural elements were different for the kringle and growth factor domains, presumably due to differences in molecular

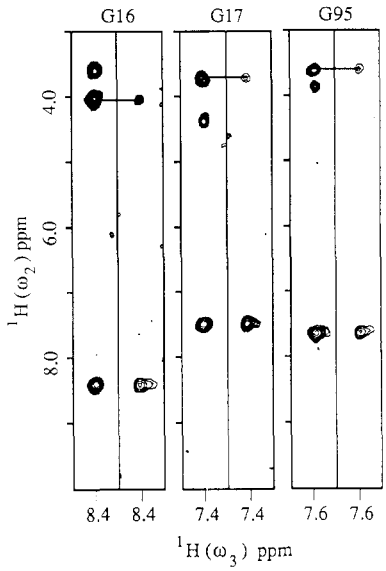


FIGURE 3: Selected $^1H(\omega_2)$, $^1H(\omega_3)$ planes extracted at the $^{15}N(\omega_1)$ chemical shifts of HNHA spectra of [U-$^{15}$N]ATF (left) and (*R*)-[2-$^2$H-$^{15}$N]Gly (right). Assignment of the (*S*)-H$^\alpha$-Gly is shown with a horizontal line.

reorientation times. Therefore, the NOE cross peak intensities for the two domains were categorized separately. NOE-derived distance restraints from the $^{13}$C-resolved NOESY involving alanine, serine, and glycine were derived from the NOE cross peak originating on the fully labeled residue. Standard corrections for center averaging (Wüthrich et al., 1983) were applied.

*Torsional Angle Restraints.* Torsional angle restrains for the $\phi$ angles were obtained from an analysis of the $^3J_{H^N,H^\alpha}$ coupling constants measured from 2D HMQC-J and 3D HNHA spectra. Restraints were included for those angles that exhibited $^3J_{H^N,H^\alpha}$ coupling constants of >9.0 Hz (−120 ± 30°) or <6.0 Hz (−60 ± 30°) in both NMR experiments. Since $^3J_{H^N,H^\alpha}$ values of <6.0 Hz may correspond to more than one $\phi$ angle, $\phi$ angles that were determined from small coupling constants were used as restraints only if they appeared in regions of $\alpha$-helical secondary structure as determined from the NOE data and amide exchange patterns.

*Hydrogen-Bonding Restraints.* Amide proton exchange was monitored with 12 $^1$H–$^{15}$N HSQC spectra collected at 4-h intervals after a dried sample of [U-$^{15}$N]ATF was redissolved in 100% $D_2O$. A total of 50 slowly exchanging amide protons were identified in the spectra. Analysis of the three-dimensional structures of ATF produced during the early rounds of refinement allowed unambiguous assignment of hydrogen bond acceptors for many of these slowly exchanging amides. Hydrogen bonds were included in subsequent calculations with upper and lower bounds of 1.8–2.5 Å ($H^N \rightarrow O$) and 2.5–3.3 Å ($N \rightarrow O$).

*Structure Calculations.* Structure calculations were performed with the X-PLOR 3.1 program (Brünger, 1992) using a distance geometry/simulated protocol (Nilges et al., 1988) as described in the X-PLOR manual (Brünger, 1992). A linear representation of the polypeptide chain was used as the initial structure for all calculations. In addition to the standard amino acids, a fucosylated threonine was constructed and included into the protein at residue 18. Six disulfide bonds were treated as covalent bonds in the final round of calculations. The pairing of cysteines was chosen on the basis of the conserved pairing in homologous proteins which was found to be consistent with mass spectral data of proteolytic fragments of u-PA (Alex Buko, personal communication).

Since initial attempts to calculate structures of the intact protein using a single stage DG/SA protocol on the two domain protein generally resulted in poor convergence, we adopted a two-stage strategy similar to that previously described for the NMR-based structure determination of human complement factor H (Barlow et al., 1993). The protocol involves calculating structures for the two domains separately, linking them together, and using the concatenated coordinates as starting positions for a final round of refinement. For ATF, the growth factor (residues 6–49) and kringle (residues 50–135) domains were calculated independently using the standard DG/SA protocol. This first round of calculations produced 40 low energy structures for each of the two structural domains. Next, each growth factor structure was arbitrarily paired with a kringle domain and loosely docked to form a two domain protein using QUANTA (Molecular Simulations Inc., Waltham, MA). After the appropriate nomenclature modifications, the concatenated, multidomain Cartesian coordinates were used as the starting structures for the final round of high temperature molecular dynamics and simulated annealing.

## RESULTS AND DISCUSSION

*Labeling in Mammalian Cells.* In order to prepare the large quantities of u-PA necessary for the structural studies, u-PA was expressed in mammalian cells. Unlike bacterially expressed u-PA that requires extensive refolding (Winkler & Blaber, 1986; Orsini et al., 1991), u-PA expressed in mammalian cells was readily isolated in its properly folded form and contains a fucose moiety attached to threonine 18 (Buko et al., 1991). However, since uniformly $^{15}$N- and $^{15}$N/$^{13}$C-labeled recombinant proteins had not been previously prepared from mammalian cells, we had to develop a method for obtaining the uniform $^{15}$N- and $^{15}$N/$^{13}$C-labeled proteins (Hansen et al., 1992). As previously described (Hansen et al., 1992), [U-$^{15}$N]ATF prepared by this method showed no dilution of the $^{15}$N-labeling upon incorporation into the protein. However, some dilution of the label was observed in the preparation of the uniformly $^{15}$N/$^{13}$C-labeled sample of ATF. In our first attempts to label the protein, unlabeled glucose was used as part of the medium. Under these conditions some dilution of the $^{13}$C-labeling might be expected, since the carbon skeleton of alanine can be made from pyruvate, and serine and glycine can be synthesized from 3-phosphoglycerate—intermediates in the glycolysis pathway. Similarly, glutamate and proline can be made from $\alpha$-ketoglutarate, and aspartate and asparagine can be obtained from oxaloacetate in the citric acid cycle. Nevertheless, minimal dilution of the carbon labeling of amino acids synthesized from the citric acid cycle was observed. In contrast, from a comparison of the cross peak intensities on different sides of the diagonal in a $^{13}$C-resolved NOESY spectrum, we estimated that the $^{13}$C-labeling of alanine, serine, and glycine was 60–75%. However, as shown in Figure 1 this amount of labeling was sufficient for analysis of the $^{13}$C-resolved data.

In addition to uniformly $^{15}$N- and $^{15}$N/$^{13}$C-labeled samples of protein, NMR samples of ATF were prepared selectively labeled by amino acid type. Figure 2B–F depict $^1$H–$^{15}$N HSQC spectra of selectively $^{15}$N-labeled ATF. The HSQC spectra of the [$^{15}$N-Phe]ATF (Figure 2B) and [$^{15}$N-Lys]ATF (Figure 2C) indicate that the label was not mixed during the mammalian cell growth. However, the HSQC spectrum of the sample which contained [$^{15}$N]Arg and [$^{15}$N]Gly (Figure 2F) showed a small amount of mixing to serine from the glycine. As shown in Figure 2E, mixing of the label was also

Table 1: Chemical Shifts for ATF[a]

| residue | HN $^{15}$N | Hα Cα | Hβ Cβ | Hγ Cγ | others |
|---|---|---|---|---|---|
| Q6 | | 4.40 | 2.51, 2.03 | 2.45, 2.45 | |
| | | 59.0 | 27.4 | | |
| V7 | 8.20 | 4.46 | 2.10 | 0.99, 0.95 | |
| | 120.9 | 59.5 | 32.3 | 20.8, 20.0 | |
| P8 | | 4.46 | 2.30, 1.91 | 2.03, 1.97 | Hδ 3.89, 3.70 |
| | | 62.9 | 32.1 | 27.1 | Cδ 50.7 |
| S9 | 8.36 | 4.43 | 3.89, 3.83 | | |
| | 115.7 | 57.5 | 64.0 | | |
| N10 | 8.39 | 4.70 | 2.79, 2.76 | | |
| | 119.3 | 53.0 | 38.4 | | |
| C11 | 8.24 | 4.35 | 3.09, 3.09 | | |
| | 117.1 | | | | |
| D12 | 8.31 | 4.47 | 2.70, 2.64 | | |
| | 120.1 | 53.5 | 39.4 | | |
| C13 | 8.17 | 4.57 | 2.60, 2.60 | | |
| | 117.6 | | | | |
| L14 | 8.44 | 4.39 | 1.63, 1.39 | 1.78 | Hδ 0.87, 0.79 |
| | 124.7 | 54.1 | 43.3 | 27.4 | Cδ 25.3, 22.1 |
| N15 | 8.52 | 3.93 | 1.94, 1.46 | | |
| | 110.1 | 54.8 | 36.8 | | |
| G16 | 8.38 | 4.06, 3.63 | | | |
| | 101.9 | 45.0 | | | |
| G17 | 7.46 | 4.44, 3.78 | | | |
| | 102.6 | 45.7 | | | |
| T18 | 8.88 | 4.65 | 3.87 | 1.16 | |
| | 116.8 | 60.1 | 75.8 | 17.9 | |
| C19 | 9.08 | 4.61 | 3.12, 2.89 | | |
| | 127.5 | | | | |
| V20 | 8.98 | 4.06 | 0.83 | 0.69, 0.41 | |
| | 128.7 | 61.1 | 33.1 | 21.8, 20.0 | |
| S21 | 7.98 | 4.68 | 3.71, 3.71 | | |
| | 116.4 | 57.2 | 65.0 | | |
| N22 | 8.54 | 4.76 | 3.31, 2.84 | | Hδ2 7.83, 7.29 |
| | 121.7 | 53.4 | 39.9 | | Nδ2 113.2 |
| K23 | 8.54 | 3.97 | 1.58, 1.39 | 1.01, 1.01 | Hδ 1.49, 1.49; Hε 2.86, 2.86 |
| | 123.1 | 57.9 | 32.3 | 25.3 | Cδ 28.9; Cε 41.7 |
| Y24 | 8.15 | 4.36 | 2.73, 2.64 | | Hδ 7.03; Hε 6.79 |
| | 116.9 | 58.5 | 38.4 | | Cδ 133.0; Cε 118.2 |
| F25 | 7.86 | 4.87 | 3.31, 2.89 | | Hδ 7.30; Hε 7.37; Hζ 7.28 |
| | 119.0 | 56.4 | 39.7 | | Cδ 132.1; Cε 131.5; Cζ 132.6 |
| S26 | 8.39 | 4.38 | 3.98, 3.92 | | |
| | 116.4 | 60.1 | 63.5 | | |
| N27 | 8.58 | 4.72 | 2.90, 2.90 | | |
| | 116.4 | 53.5 | 37.7 | | |
| I28 | 7.83 | 4.37 | 2.05 | 1.02m 1.53, 1.19 | Hδ 0.85 |
| | 120.4 | 60.6 | 38.4 | 16.9 26.8 | Cδ 12.5 |
| H29 | 8.33 | 5.48 | 3.21, 3.09 | | Hδ 7.05 |
| | 121.0 | 53.8 | 31.6 | | Cδ 119.7 |
| W30 | 8.79 | 5.08 | 3.44, 3.34 | | Hδ 6.97; Hε1 10.03; Hε3 7.29 |
| | 119.7 | 56.4 | 30.5 | | Cδ 127.4; Nε1 129.4; Cε3 120.8; Hζ2 7.39; Hζ3 6.95; Hη 7.08; Cζ2 115.0; Cζ3 121.3; Cη 123.7 |
| C31 | 8.84 | 5.23 | 2.91, 2.67 | | |
| | 115.3 | | | | |
| N32 | 9.33 | 5.18 | 2.89, 2.54 | | |
| | 123.4 | 51.8 | 39.1 | | |
| C33 | 8.93 | 5.29 | 3.19, 2.75 | | |
| | 122.8 | | | | |
| P34 | | 4.67 | 2.69, 2.21 | 2.55, 2.06 | Hδ 3.66, 3.53 |
| | | 62.7 | 32.1 | 27.9 | Cδ 50.6 |
| K35 | 8.35 | 4.09 | 1.88, 1.88 | 1.55, 1.55 | Hδ 1.75, 1.75; Hε 3.05, 3.05 |
| | 119.1 | 59.3 | 32.3 | 24.7 | Cδ 29.2; Cε 42.0 |
| K36 | 7.89 | 3.94 | 1.58, 1.42 | 1.09, 1.01 | Hδ 1.52, 1.52; Hε 2.85, 2.85 |
| | 114.6 | 57.4 | 32.3 | 25.3 | Cδ 29.2; Cε 41.7 |
| F37 | 8.06 | 5.63 | 3.26, 2.59 | | Hδ 6.96; Hε 7.25; Hζ 7.21 |
| | 115.1 | 56.7 | 43.1 | | Cδ 132.4; Cε 131.1; Cζ 130.0 |
| G38 | 9.28 | 4.76, 3.79 | | | |
| | 106.7 | 43.3 | | | |
| G39 | 8.10 | 4.85, 3.92 | | | |
| | 105.5 | 43.6 | | | |
| Q40 | 9.43 | 4.02 | 2.06, 1.87 | 2.43, 2.21 | |
| | 120.2 | 58.5 | 28.9 | 34.2 | |
| H41 | 9.61 | 5.29 | 3.95, 2.98 | | |
| | 113.7 | 54.3 | 28.7 | | |
| C42 | 7.57 | 4.29 | 3.76, 3.04 | | |
| | 111.2 | | | | |

Table 1 (Continued)

| residue | HN $^{15}$N | Hα Cα | Hβ Cβ | Hγ Cγ | others |
|---|---|---|---|---|---|
| E43 | 10.55 | 4.22 | 2.21, 1.96 | 2.61, 2.28 | |
| | 119.1 | 56.9 | 29.4 | 35.5 | |
| I44 | 8.79 | 4.26 | 1.52 | 0.733m 1.54, 1.12 | Hδ 0.85 |
| | 124.2 | 60.1 | 38.6 | 16.6 27.1 | Cδ 11.9 |
| D45 | 8.50 | 4.51 | 2.84, 2.54 | | |
| | 126.5 | 53.0 | 40.7 | | |
| K46 | 8.42 | 4.01 | 1.83, 1.53 | 1.32, 1.27 | Hδ 1.52, 1.52; Hε 2.78, 2.78 |
| | 125.1 | 57.7 | 31.8 | 23.9 | Cδ 28.9; Cε 41.7 |
| S47 | 8.62 | 4.37 | 3.90, 3.77 | | |
| | 114.3 | 59.8 | 64.0 | | |
| K48 | 7.40 | 4.45 | 1.94, 1.82 | 1.57, 1.50 | Hδ 1.78, 1.78; Hε 3.09, 3.09 |
| | 121.2 | 56.2 | 33.1 | 24.8 | Cδ 28.9; Cε 42.0 |
| T49 | | 4.58 | 4.43 | 1.24 | |
| | | 61.6 | 70.0 | 21.4 | |
| C50 | | | 1.95, 1.95 | | |
| Y51 | | 5.33 | 3.12, 2.46 | | Hδ 6.51; Hε 6.83 |
| | | 54.8 | 40.9 | | Cδ 133.6; Cε 118.2 |
| E52 | 9.29 | 4.76 | 2.11, 2.02 | 2.34, 2.27 | |
| | 122.14 | 54.3 | | 36.2 | |
| G53 | 9.16 | 4.18, 3.87 | | | |
| | | 46.7 | | | |
| N54 | 8.62 | 4.93 | 3.47, 2.76 | | |
| | 122.9 | 51.7 | 37.6 | | |
| G55 | 9.21 | 4.18, 3.64 | | | |
| | 106.3 | 46.7 | | | |
| H56 | 8.74 | 4.47 | 3.53, 2.88 | | Hδ 6.43; Hε 8.37 |
| | 120.8 | 57.7 | 28.4 | | Cδ 120.3; Cε 135.9 |
| F57 | 8.92 | 4.56 | 3.52, 2.94 | | Hδ 7.31; Hε 7.45; Hζ 7.42 |
| | 114.7 | 57.2 | 37.8 | | Cδ 131.8; Cε 131.8; Cζ 130.0 |
| Y58 | 7.46 | 4.54 | 3.08, 3.29 | | Hδ 7.22; Hε 6.75 |
| | 120.5 | 59.6 | 38.4 | | Cδ 133.4; Cε 118.2 |
| R59 | 8.97 | 4.34 | 2.34, 1.95 | | Hδ 3.32, 3.32 |
| | 129.6 | 55.1 | 28.3 | | Cδ 45.2 |
| G60 | 4.45 | 4.31, 3.41 | | | |
| | 104.1 | 44.9 | | | |
| K61 | 7.81 | 3.45 | 1.58, 1.21 | 0.97, 0.28 | Hδ 1.47, 1.47; Hε 2.88, 2.88 |
| | 113.9 | 55.4 | 31.8 | 25.0 | Cδ 29.2; Cε 41.7 |
| A62 | 7.89 | 4.34 | 1.55 | | |
| | 123.9 | 54.3 | 19.0 | | |
| S63 | 8.56 | 5.29 | 4.41, 3.71 | | |
| | 113.3 | 59.0 | 65.1 | | |
| T64 | 7.47 | 5.46 | 4.02 | 1.14 | |
| | 111.7 | 59.0 | 72.9 | 21.1 | |
| D65 | 8.81 | 5.02 | 3.48, 2.42 | | |
| | 120.1 | 52.7 | 41.7 | | |
| T66 | 7.62 | 4.07 | 4.42 | 1.19 | |
| | 106.1 | 64.5 | 68.7 | 21.8 | |
| M67 | 8.42 | 4.70 | 2.33, 2.13 | 2.57, 2.52 | Hε 2.10 |
| | 120.2 | 54.3 | 32.6 | 32.6 | Cε 16.9 |
| G68 | 8.24 | 4.25, 3.67 | | | |
| | 108.5 | 45.2 | | | |
| R69 | 8.99 | 4.53 | 2.06, 2.06 | 1.75, 1.62 | Hδ 3.40, 3.00 |
| | 124.0 | 54.3 | | 27.6 | Cδ 42.8 |
| P70 | | 4.67 | 2.28, 1.93 | 2.24, 2.13 | Hδ 3.89, 3.71 |
| | | 61.9 | 32.0 | 27.6 | Cδ 49.8 |
| C71 | 8.07 | 4.51 | 3.15, 2.54 | | |
| | 119.8 | | | | |
| L72 | 9.52 | 4.58 | 1.46, 1.08 | 1.74 | Hδ 0.98, 0.61 |
| | 124.6 | 52.5 | 42.5 | 26.8 | Cδ 23.2, 26.0 |
| P73 | | 4.44 | 2.26, 1.88 | 2.19, 2.04 | Hδ 4.11, 3.52 |
| | | 61.6 | 31.8 | 27.9 | Cδ 50.1 |
| W74 | 8.69 | 4.65 | 3.89, 3.27 | | Hδ 7.33; Hε1 11.67; Hε3 7.81 |
| | 121.3 | 59.0 | 27.6 | | Cδ 129.7; Nε1 131.3; Cε3 120.0; Hζ2 7.44; Hζ3 7.15; Hη3 5.68; Cζ2 114.8; Cζ3 123.4; Cη3 123.9 |
| N75 | 8.23 | 4.65 | 3.14, 2.05 | | |
| | 112.6 | 51.2 | 38.1 | | |
| S76 | 7.54 | 4.26 | 4.27, 3.97 | | |
| | 114.5 | 58.3 | 65.1 | | |
| A77 | 8.84 | 3.91 | 1.46 | | |
| | 124.2 | 55.6 | 17.9 | | |
| T78 | 8.08 | 4.01 | 3.95 | 1.14 | |
| | 108.3 | 65.1 | 68.5 | 22.1 | |
| V79 | 7.09 | 3.28 | 2.06 | 1.48, 0.69 | |
| | 124.4 | 66.4 | 31.3 | 23.9, 26.3 | |
| L80 | 8.06 | 3.55 | 1.64, 1.07 | 1.51 | Hδ 0.56, 0.29 |
| | 118.2 | 56.7 | 41.2 | 26.1 | Cδ 25.5, 22.6 |

Table 1 (Continued)

| residue | HN $^{15}$N | Hα Cα | Hβ Cβ | Hγ Cγ | others |
|---|---|---|---|---|---|
| Q81 | 6.65 | 4.31 | 2.32, 1.95 | 2.44, 2.32 | |
| | 112.6 | 55.1 | 28.7 | 33.6 | |
| Q82 | 8.06 | 4.52 | 2.58, 2.18 | 2.45, 1.14 | Hε2 7.15, 6.99 |
| | 118.3 | 52.7 | 28.7 | 32.8 | Nε2 114.6 |
| T83 | 8.22 | 3.89 | 3.87 | 0.67 | |
| | 112.9 | 66.4 | 68.9 | 21.9 | |
| Y84 | 7.86 | 4.36 | 3.58, 2.88 | | Hδ 7.10; Hε 6.85 |
| | 117.1 | 59.6 | 39.9 | | Cδ 132.0; Cε 118.7 |
| H85 | 5.82 | 5.25 | 3.77, 3.21 | | Hδ2 7.40; Hε1 7.42 |
| | 110.2 | 54.3 | 33.4 | | Cδ2 121.3; Cε1 138.1 |
| A86 | 8.02 | 4.01 | 1.20 | | |
| | 115.6 | 53.0 | 19.5 | | |
| H87 | 10.47 | 5.23 | 3.71, 3.33 | | Hδ2 6.94; Hε1 8.65 |
| | 116.8 | 53.0 | 28.7 | | Cδ2 117.1; Cε1 136.3 |
| R88 | 7.07 | 4.30 | 2.51, 1.45 | 1.43, 0.65 | Hδ 3.85, 2.87 |
| | 120.5 | 56.4 | 33.6 | 28.1 | Cδ 43.8 |
| S89 | 9.30 | 4.28 | 4.03, 4.03 | | |
| | 118.8 | 61.1 | 62.4 | | |
| D90 | 8.05 | 4.88 | 3.18, 2.51 | | |
| | 118.2 | 51.9 | 39.7 | | |
| A91 | 7.39 | 3.64 | 1.63 | | |
| | 119.8 | 56.7 | 19.3 | | |
| L92 | 8.49 | 4.18 | 1.73, 1.64 | 1.65 | Hδ 0.93, 0.86 |
| | 116.5 | 58.0 | 40.9 | 26.8 | Cδ 24.2, 24.2 |
| Q93 | 7.98 | 4.07 | 2.23, 2.16 | 2.56, 2.44 | |
| | 119.3 | 58.8 | 27.9 | 34.2 | |
| L94 | 7.47 | 4.24 | 1.57, 1.26 | 1.81 | Hδ 0.88, 0.86 |
| | 115.7 | 54.1 | 43.6 | 26.1 | Cδ 21.6, 24.7 |
| G95 | 7.65 | 3.88, 3.66 | | | |
| | 106.1 | 45.9 | | | |
| L96 | 8.14 | 3.60 | 1.59, 0.41 | 1.10 | Hδ 0.51, −1.03 |
| | 118.8 | 52.7 | 40.2 | 25.3 | Cδ 26.1, 18.5 |
| G97 | 5.36 | 4.37, 3.65 | | | |
| | 107.5 | 44.6 | | | |
| K98 | 8.51 | 4.20 | 1.74, 1.74 | 1.54, 1.26 | Hδ 1.78, 1.78; Hε 3.06, 3.06 |
| | 122.6 | 55.1 | 29.7 | 24.5 | Cδ 29.7; Cε 41.8 |
| H99 | 8.19 | 4.59 | 2.97, 1.77 | | Hδ2 6.87; Hε1 7.73 |
| | 118.6 | 54.3 | 31.6 | | Cε1 137.3 |
| N100 | 7.47 | 4.77 | 2.43, 1.81 | | |
| | 119.8 | 50.7 | 38.4 | | |
| Y101 | 10.11 | 5.38 | 3.46, 2.60 | | Hδ 6.28; Hε 6.73 |
| | 117.4 | 53.9 | 38.1 | | Cδ 133.1; Cε 117.1 |
| C102 | 9.08 | 4.51 | 3.18, 2.62 | | |
| | 118.3 | | | | |
| R103 | 9.13 | 5.00 | 1.45, 1.01 | 1.96, 1.16 | Hδ 3.77, 3.05 |
| | 121.6 | 52.5 | 35.7 | 26.8 | Cδ 43.3 |
| N104 | 8.30 | 5.34 | 2.49, 1.95 | | |
| | 110.5 | 49.6 | 37.8 | | |
| P105 | | 4.39 | 1.76, 1.63 | 1.03, 0.99 | Hδ 3.36, 2.55 |
| | | 64.3 | 33.9 | 26.8 | Cδ 49.1 |
| D106 | 9.24 | 4.51 | 3.03, 2.54 | | |
| | 115.4 | 52.4 | 38.9 | | |
| N107 | 7.16 | 4.02 | 2.78, 2.64 | | |
| | 112.0 | 55.6 | 36.3 | | |
| R108 | 9.07 | 4.39 | 2.15, 1.81 | 1.84, 1.54 | Hδ 2.68, 2.45 |
| | 117.5 | 57.2 | 30.2 | 27.9 | Cδ 42.5 |
| R109 | 8.43 | 4.09 | 1.96, 1.91 | 1.74, 1.67 | Hδ 3.24, 3.24 |
| | 117.2 | 59.0 | 30.5 | 26.8 | Cδ 43.3 |
| R110 | 7.26 | 4.51 | 1.89, 1.47 | 1.58, 1.50 | Hδ 3.23, 3.23 |
| | 110.8 | 53.5 | 30.5 | 25.2 | Cδ 43.8 |
| P111 | | 4.03 | 1.58, 1.58 | | |
| | | 63.7 | | | |
| W112 | 8.99 | 5.42 | 3.36, 2.93 | | Hδ 7.28; Hε1 10.03; Hε3 7.00 |
| | 120.5 | 55.5 | 33.6 | | Cδ 129.2; Nε1 130.2; Cε3 121.6; Hζ2 6.46; Hζ3 5.43; Hη3 7.22; Cζ2 113.5; Cζ3 122.1; Cη3 122.4 |
| C113 | 9.15 | 4.63 | 3.48, 2.90 | | |
| | 111.7 | | | | |
| Y114 | 8.53 | 5.14 | 2.71, 2.53 | | Hδ 6.90; Hε 6.88 |
| | 118.0 | 58.3 | 39.4 | | Cε 118.7 |
| V115 | 9.02 | 4.30 | 1.68 | 0.97, 0.81 | |
| | 120.5 | 60.3 | 36.5 | 21.8, 21.0 | |
| Q116 | 8.82 | 4.16 | 2.10, 1.94 | 2.28, 1.95 | Hε2 6.61, 6.99 |
| | 125.4 | 55.9 | 28.2 | 33.1 | Nε2 111.1 |
| V117 | 8.47 | 4.18 | 2.09 | 0.91, 0.85 | |
| | 131.7 | 62.2 | 32.1 | 21.1, 20.6 | |

Table 1 (Continued)

| residue | HN $^{15}$N | Hα Cα | Hβ Cβ | Hγ Cγ | others |
|---|---|---|---|---|---|
| G118 | 8.76 | 4.05, 3.69 | | | |
| | 115.9 | 46.5 | | | |
| L119 | 8.85 | 4.21 | 1.78, 1.78 | 1.67 | Hδ 0.95, 0.86 |
| | 124.2 | 55.4 | 42.3 | 26.8 | Cδ 25.0, 22.5 |
| K120 | 7.62 | 4.91 | 1.76, 1.76 | 1.41, 1.41 | Hδ 1.75, 1.75; Hε 3.04, 3.04 |
| | 118.4 | 53.0 | 34.9 | 24.2 | Cδ 28.7; Cε 42.0 |
| P121 | | 4.90 | 1.66, 1.48 | 1.98, 1.75 | Hδ 3.71, 3.64 |
| | | 61.1 | 31.0 | 27.1 | Cδ 49.6 |
| L122 | 9.25 | 4.90 | 1.67, 1.63 | 1.75 | Hδ 0.97, 0.95 |
| | 123.9 | 53.3 | 46.1 | 26.8 | Cδ 24.8, 23.9 |
| V123 | 8.38 | 3.80 | 1.66 | 0.59, 0.11 | |
| | 123.2 | 64.5 | 31.3 | 21.8, 19.3 | |
| Q124 | 9.15 | 4.69 | 2.38, 1.97 | 2.53, 2.22 | |
| | 127.0 | 53.5 | 33.9 | 33.9 | |
| E125 | 8.57 | 4.27 | 2.04, 1.92 | 2.48, 2.29 | |
| | 118.7 | 56.7 | 29.9 | 37.3 | |
| C126 | 8.36 | 4.75 | 2.86, 2.86 | | |
| | 118.2 | | | | |
| M127 | 8.60 | 4.40 | 2.51, 2.39 | 2.08, 1.74 | Hε 2.10 |
| | 122.9 | 54.8 | 32.1 | 31.8 | Cε 16.9 |
| V128 | 7.41 | 4.22 | 1.98 | 0.79, 0.60 | |
| | 122.8 | 60.6 | 32.8 | 20.5, 19.8 | |
| H129 | 8.72 | 4.94 | 3.26, 3.15 | | Hδ2 7.30; Hε1 8.60 |
| | 122.8 | 54.3 | 30.3 | | Cδ2 120.0; Cε1 136.8 |
| D130 | 8.54 | 4.32 | 3.20, 2.55 | | |
| | 120.9 | 52.5 | 42.3 | | |
| C131 | | | | | |
| A132 | 8.71 | 4.34 | 1.40 | | |
| | 122.5 | 52.5 | 19.0 | | |
| D133 | 7.96 | 4.79 | 3.24, 2.86 | | |
| | 116.8 | 51.4 | 37.1 | | |
| G134 | 8.17 | 3.93, 3.93 | | | |
| | 108.6 | 45.3 | | | |
| K135 | 7.75 | 4.12 | 1.85, 1.74 | 1.40, 1.40 | Hδ 1.69, 1.69; Hε 3.02, 3.02 |
| | 124.6 | 57.2 | 33.4 | 24.7 | Cδ 28.9; Cε 42.0 |

[a] $^{1}$H chemical shifts are referenced to $H_2O$ (4.74 ppm), $^{13}$C relative to external TSP (0.0 ppm), and $^{15}$N relative to $NH_4NO_3$ (119.4 ppm).

observed when [$^{15}$N]valine was incorporated in the mammalian cell culture media. Although HSQC cross peaks corresponding to leucine and isoleucine were observed, the valine peaks were the largest in the spectrum and could easily be distinguished from the other peaks. The most scrambling was observed for the $^{15}$N-labeling of leucine. In addition to the HSQC cross peaks corresponding to the leucine amides, peaks were also observed for isoleucine, valine, aspartate, glutamate, alanine, serine, and glycine. Despite the mixing of the labels, the identification of the amide signals by amino acid type could be readily made from careful analysis of the relative peak intensities. In some cases, the additional peaks observed in the spectra aided in the assignment process. Our results with selective labeling of proteins in Sp2/0 cells are consistent with that previously observed by Archer et al. (1993) who employed 4–6 $^{15}$N-labeled amino acids in each preparation.

*NMR Assignments.* The $^{1}$H, $^{13}$C, and $^{15}$N chemical shifts were assigned from analysis of several 2D and 3D NMR spectra. In the first step of the assignment procedure, the $^{1}$H and $^{15}$N amide frequencies of each of the amino acid spin systems were correlated to the $C^{\alpha}$ frequencies of both the *i* and *i*–1 amino acids in a 3D HNCA experiment. Additional $C^{\alpha}_{i-1}$ correlations were observed in a 3D HN(CO)CA experiment which also allowed the correlations between the amides and $C^{\alpha}_{i}$ or $C^{\alpha}_{i-1}$ to be distinguished. Next, the $^{1}$H and $^{15}$N amide chemical shifts were correlated to the $H^{\alpha}$ frequencies of both the *i* and *i*–1 amino acids. Amide/$H^{\alpha}_{i}$ correlations were obtained from either a 3D $^{15}$N-resolved TOCSY-HSQC or a 3D HNHA experiment while the amide/$H^{\alpha}_{i-1}$ correlations were observed in a $^{15}$N-resolved NOESY-HSQC spectrum. From analysis of these data and other spectral information (e.g., NOEs between amide protons), many of the backbone resonances of the adjacent amino acid spin systems were identified. In order to sequentially assign the backbone signals, the amino acid spin systems had to be identified by amino acid type. This was accomplished by recording 2D HSQC spectra of ATF that was selectively $^{15}$N-labeled by amino acid type (Figure 2) or by analyzing the side chain $^{1}$H and $^{13}$C chemical shifts obtained from 3D CBCA(CO)NH, HCCH-TOCSY, and HNHB spectra. Once the backbone signals were assigned, the $^{1}$H and $^{13}$C resonances of the aliphatic side chains were assigned by correlating the $C^{\alpha}$ and $H^{\alpha}$ signals with the $^{1}$H and $^{13}$C signals of the side chains in a 3D HCCH-TOCSY experiment (Figure 1). The $^{1}$H and $^{13}$C signals of the aromatic side chains were assigned by correlating the $H^{\delta}/C^{\delta}$ chemical shifts of the aromatic rings with their corresponding $H^{\beta}/C^{\beta}$ chemical shifts in a $^{13}$C-resolved NOESY experiment and then correlating the $H^{\delta}$ and $C^{\delta}$ frequencies to the remaining signals of the aromatic rings in a 3D HCCH-TOCSY experiment.

Stereospecific assignments of the α-protons of the glycine residues were obtained using a sample of ATF in which (*R*)-[2-$^{2}$H,$^{15}$N]Gly (Curley et al., 1994) was incorporated. As shown in Figure 3, the stereospecific assignments were made by comparing the individual $^{1}$H–$^{1}$H planes of the 3D HNHA spectrum of uniformly $^{15}$N-labeled ATF with an ATF sample containing stereospecifically deuterated and $^{15}$N-labeled Gly. Stereospecific assignments of valine methyl groups were obtained from $^{3}J_{C',C\gamma}$ coupling constants measured from a 2D $^{13}$C–{$^{13}$CO} spin-echo difference CT-HSQC experiment (Grzesiek et al., 1993) and the observed pattern of NOEs (Zuiderweg et al., 1985). The stereospecific assignments of the leucine methyl groups were obtained from $^{3}J_{N,H^{\beta}}$ and $^{3}J_{H^{\alpha},H^{\beta}}$ coupling constants estimated from HNHB and $^{15}$N TOCSY-

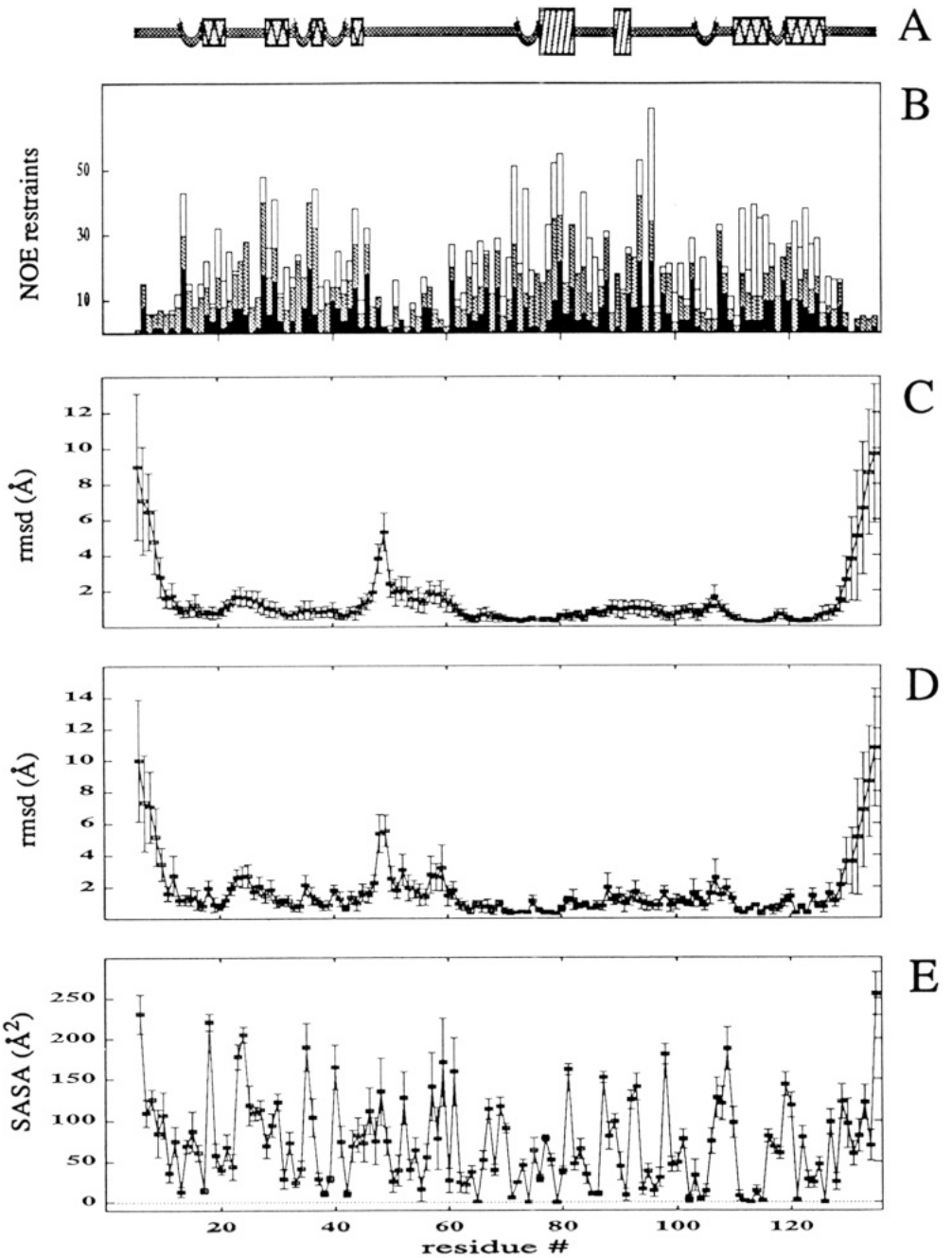


FIGURE 4: (A) Schematic views of the secondary structure of the amino-terminal fragment of u-PA in which lines denote nonregular secondary structure, semicircles show $\beta$-turns, thin sawtooth boxes depict $\beta$-strands, and large striped boxes denote $\alpha$-helical residues. (B) Distributions of NOEs that were used in the structure calculations. The height reflects the total number of NOE-derived restraints for each residue, while the black, gray, and white portions denote the intraresidue, short-range, and long-range restraints, respectively. (C) Residue-based root mean square deviation to the mean coordinates for backbone heavy atoms (C$\alpha$, N, C′). (D) Residue-based root mean square deviation to the mean coordinates for all heavy atoms. Note that in panels C and D, the rmsd for each domain was calculated separately, and the results were concatenated. (E) Solvent-accessible surface area based upon the X-PLOR implementation of the Lee and Richards algorithm. Probe radius was 1.6 Å.

HSQC experiments, respectively, and intraresidue NOE data (Powers et al., 1993). Attempts to stereospecifically assign the $\beta$-methylene protons were unsuccessful due to the relatively large line widths of ATF signals which made $^3J_{H^\alpha,H^\beta}$ coupling constants difficult to measure.

The $^1$H, $^{13}$C, and $^{15}$N assignments for the amino-terminal fragment of u-PA are given in Table 1. Except for cysteine 131, most of the NMR signals were assigned. The failure to unambiguously identify any resonances of C131 may be due, in part, to the unfavorable relaxation properties for this portion of ATF. In fact, preliminary analysis of $^{15}$N relaxation data shows that the amides in the region around C131 exhibit very short $^{15}$N spin–spin relaxation rates.

*NMR-Derived Restraints.* A total of 1299 NOE-derived distance restraints were obtained from analysis of the 3D $^{15}$N- and 3D $^{13}$C-resolved NOESY spectra. For the growth factor domain (residues 6–49), 415 NOE-derived distance restraints were obtained, including 121 nontrivial intraresidue, 144 sequential, 55 short-range ($|i-j| < 5$, where $i$ and $j$ are residue numbers), and 95 long-range ($|i-j| \geq 5$) NOEs. For the kringle domain (residues 50–135), 884 NOE-derived distance restraints were employed in the structure calculations, including 246 nontrivial intraresidue, 227 sequential, 113 short-range, and 298 long-range NOEs. A histogram of the NOE restraints for each residue is shown in Figure 4. In addition to the NOE-derived distance restraints, torsional angle

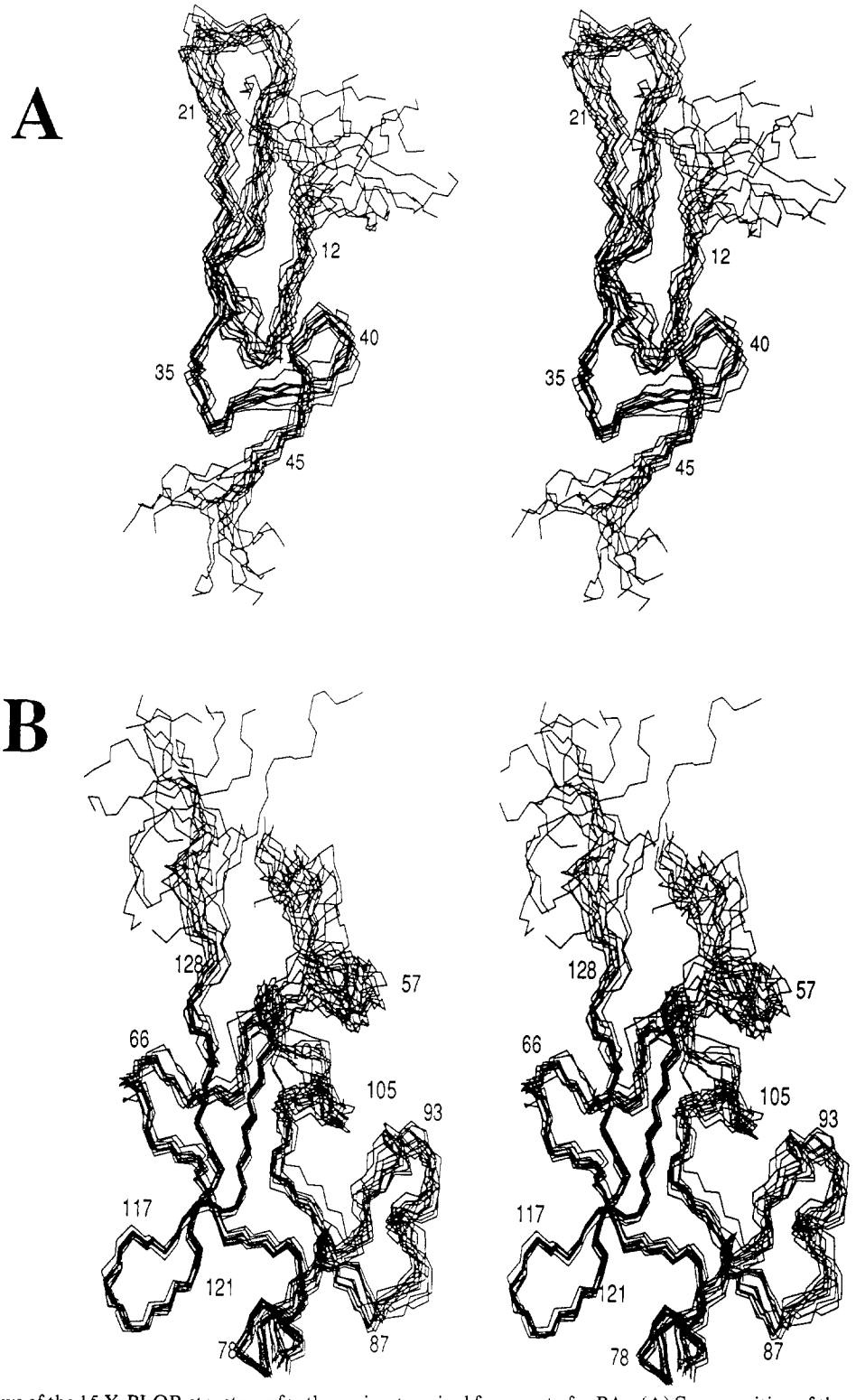


FIGURE 5: Stereoviews of the 15 X-PLOR structures for the amino-terminal fragment of u-PA. (A) Superposition of the growth factor domain (residues 6–49) onto the mean atomic coordinates of residues 11–46. The rmsd for the backbone heavy atoms of residues 11–46 was 0.99 ± 0.2 Å. (B) Superposition of the kringle domain (residues 50–135) onto the mean atomic coordinates of residues 55–130. Backbone rmsd for these residues was 0.85 ± 0.2 Å.

restraints for 27 $\phi$ angles were obtained from analysis of the 2D HMQC-J and 3D HNHA spectra. Also, by inspection of initial ATF structures produced after several rounds of refinement without hydrogen-bonding restraints, hydrogen-bonding partners could be unambiguously identified for 21 out of the 50 slowly (>15 min) exchanging amides. The total number of experimentally derived restraints was thus 1368 (10.5/residue).

*Quality of the ATF Structure.* From the 40 structures obtained from the final round of high temperature molecular dynamics and simulated annealing, the 15 structures with the lowest total energy were retained for further analysis. A superposition of these structures to the average structure is shown in Figure 5, structural statistics are given in Table 2, and residue based root mean square deviations (rmsd) and solvent exposure are presented in Figure 4. As shown in the

Table 2: Structural Statistics and RMSD for 15 ATF Structures[a]

| structural statistics | ⟨SA⟩ | $\overline{\langle SA \rangle}_{gfr}$ | $\overline{\langle SA \rangle}_{kr}$ |
|---|---|---|---|
| X-PLOR energies (kcal·mol$^{-1}$) | | | |
| $E_{tot}$ | 271 ± 10 | 85 | 203 |
| $E_{bond}$ | 30 ± 1 | 8 | 22 |
| $E_{angle}$ | 157 ± 5 | 50 | 123 |
| $E_{impr}$ | 20 ± 1 | 8 | 16 |
| $E_{vdw}$[b] | 24 ± 4 | 11 | 22 |
| $E_{cdih}$[c] | 0.9 ± 0.2 | 0.2 | 0.6 |
| $E_{noe}$[d] | 26 ± 4 | 8 | 20 |
| RMSD from idealized values | | | |
| bonds (Å) | 0.04 ± 0.00 | 0.04 | 0.04 |
| angles (deg) | 0.54 ± 0.54 | 0.54 | 0.58 |
| impropers (deg) | 0.50 ± 0.02 | 0.50 | 0.58 |

| Cartesian coordinate rmsd (Å) | N, Cα, C′ | all non-H |
|---|---|---|
| ⟨SA⟩ vs $\overline{\langle SA \rangle}_{gf}$[e] | 0.99 ± 0.2 | 1.65 ± 0.2 |
| ⟨SA⟩ vs $\overline{\langle SA \rangle}_{k}$[f] | 0.84 ± 0.2 | 1.42 ± 0.2 |

[a] Where ⟨SA⟩ is the ensemble of 15 final X-PLOR structures, $\overline{\langle SA \rangle}_{gf}$ is the Cartesian coordinates for residues 6–49 obtained by averaging the ensemble of structures following a least-squares superposition of the backbone heavy atoms (N, Cα, C′) of residues 11–46, $\overline{\langle SA \rangle}_{gfr}$ is the energy minimized averaged growth factor coordinates, $\overline{\langle SA \rangle}_{k}$ is the average Cartesian coordinates of residues 50–135 obtained after a least-squares superposition of the backbone heavy atoms of residues 55–130, and $\overline{\langle SA \rangle}_{kr}$ is the energy-minimized averaged kringle coordinates. [b] The X-PLOR $F_{repel}$ function was used to simulate the van der Waals potential with atomic radii set to 0.8 times their CHARMM (Brooks et al., 1983) values. [c] Dihedral angles were given force constants of 200 kcal·mol$^{-1}$·rad$^{-2}$ and were applied at the beginning of the annealing refinement stage. No experimental dihedral was violated by >5.0° from the applied bounds. [d] NOE-derived distance restraints were applied with a square-well potential with lower bounds set to 1.8 Å and upper bounds of 3.0, 4.0, or 5.0 Å. The force constant was 50 kcal·mol$^{-1}$·Å$^{-2}$. Hydrogen bonds are included in $E_{noe}$ and were given bounds of 1.8–2.3 and 2.5–3.3 Å with force constants of 50 kcal·mol$^{-1}$·Å$^{-2}$. No distance restraint was violated by >0.35 Å in any of the final structures. [e] RMSD for residues 11–46. [f] RMSD for residues 55–130.

figures, most of the backbone in the kringle and growth factor domains exhibits good structural convergence. In addition, many of the side chains in the kringle domain are also well-defined (Figure 6). The root mean square deviation for residues 55–130 of the kringle domain to the averaged coordinates was 0.84 ± 0.2 and 1.4 ± 0.2 Å for the backbone and all heavy atoms, respectively, while residues 11–46 of the growth factor exhibited an rmsd of 0.99 ± 0.2 and 1.65 ± 0.2 Å. An additional measure of the structural convergence was obtained using the angular order parameter, $S^{angle}$ (Hyberts et al., 1992), which is a measure of the local torsional variation. A value of 1 indicates perfect convergence, while a value of 0 results from dihedral heterogeneity. As shown in Figure 7, the $\phi$ and $\psi$ $S^{angle}$ values follow the trends of the residue based rmsd (Figure 4). High angular convergence ($S^{angles} \geq 0.9$) was observed for 63 of the 130 residues. These included residues 14–24, 32–42, 63–83, 87, 90–95, 99, 111–117, 120–124, and 129. As shown in Figure 8B, these residues are all within the favorable region of the Ramachandran plot. Also as noted in Figure 8B, positive $\phi$ angles for non-glycine residues were observed for structurally converged residues N15, N27, Q40, C42, C71, and H87.

Although the 15 low energy structures are well-defined in many regions, there remain several residues that exhibit high positional heterogeneity. Specifically, five-residue segments at the amino (6–10) and carboxy (131–135) termini are not defined by the NMR data. Also, residues 47–55, which serve to link the growth factor and kringle domains, are conformationally diverse. As a result of the heterogeneity in the linker region, the spatial relationship between the two domains is ill-defined. Depicted in Figure 9 is the superposition of the heavy atoms for the kringle onto the mean atomic coordinates for that domain. The rmsd for the backbone heavy atoms in the growth factor domain for the superpositions shown in Figure 9 was 23.3 ± 6 Å, indicating that the position of the growth factor domain is not influenced by the structure of the kringle. As discussed below, difference in the relative line widths observed for the two domains also suggest independent motion for the two domains.

*Description of the Structure.* The growth factor domain of ATF consists of two regions of antiparallel β-sheet and three β-turns (Figure 10A). One two-stranded sheet is formed by residues T18–S21 and H29–N32 while the other is a very small sheet formed by residues F37–G38 and I44–D45. From the $\phi$,$\psi$ angles in the 15 final structures, a type I′ turn was identified for L14–G17, a type I turn for P34–F37, while residues G39–C42 adopt a type II′ β-turn. Although the growth factor domain does not have a hydrophobic core, residues T18, V20, W30, and N32 cluster on one side of the β-sheet as shown in Figure 10A. Another key contact was found between Q43 and L14. Q43 forms a bulge between the last reverse turn and the second strand of the small β-sheet. Its interaction with L14 helps to fix the orientation of the C-terminal, double turn region relative to the major β-sheet. One of the most interesting structural elements observed in the growth factor domain is the seven-residue omega loop (Leszczynski & Rose, 1986) which connects the two strands of the major β-sheet. In this loop the side chains of N22, F25, and I28 point inward and are clustered toward the center of the loop (Figure 10A), while the side chains of K23, Y24, S26, and N27 point outward.

The kringle domain is composed of one antiparallel β-sheet, two small regions of β-sheet-like structure, two short helices, and several reverse turns. A two-stranded β-sheet is formed by residues P111–V117 and K120–C126 which are connected by a reverse turn, which could not be clearly classified, from V117–K120. One region of β-sheet-like structure was found between T64–D65 and R69–P70. While the pattern of NOEs, slowly exchanging amides, and Cα chemical shifts observed in this region are indicative of a β-sheet, main chain hydrogen bonds indicative of β-sheet structure could not be definitively identified due to the close proximity of several potential hydrogen bond acceptor side chains. The second region of β-sheet-like structure was found between P73–W74 and H99–N100. Although the observed cross-strand NOEs between these residues are indicative of a β-sheet, P73 and W74 are also the first two residues of a type I turn, suggesting that this is not a regular β-sheet.

A six-residue α-helix was found from A77 to Q82 in the kringle domain. In this helix, A77, T78, Q81, and Q82 are oriented toward the solvent while V79 and L80 are oriented inward and pack against the hydrophobic core of the kringle domain. This hydrophobic core is composed of a number of aromatic and aliphatic side chains which include W74, Y84, L94, L96, W112, and Y114 (Figure 6). A second, helical region was identified from D90–L94. Other regions of secondary structure in the kringle include type I turns from residues P73–S76 and N104–N107 and an irregular turn involving a pentapeptide from D65–R69. The two central disulfide bonds C71–C113 and C102–C126 are stacked orthogonal to one another and lie on top of the hydrophobic core as shown in the orientation depicted in Figure 10B.

In earlier studies based on homonuclear NMR data, Dobson and co-workers reported the $^1$H chemical shift assignments and secondary structure of the u-PA kringle (Li et al., 1992).

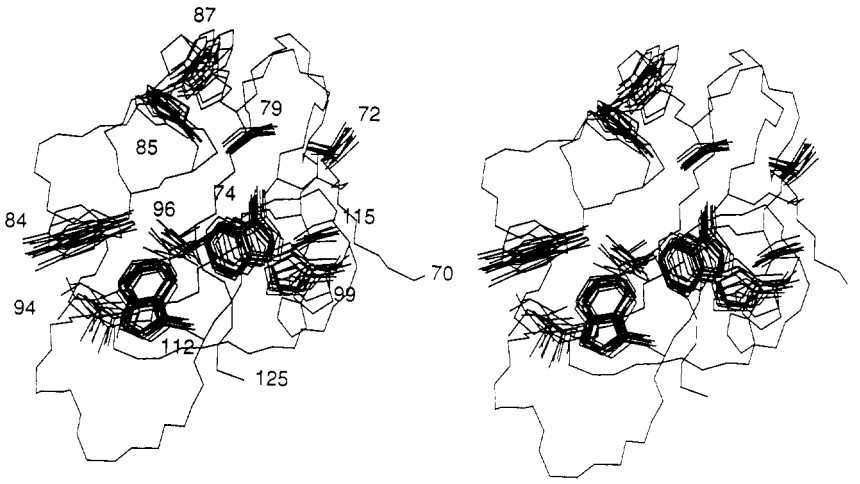


FIGURE 6: Stereoview of the superposition of selected side chains in the kringle domain.

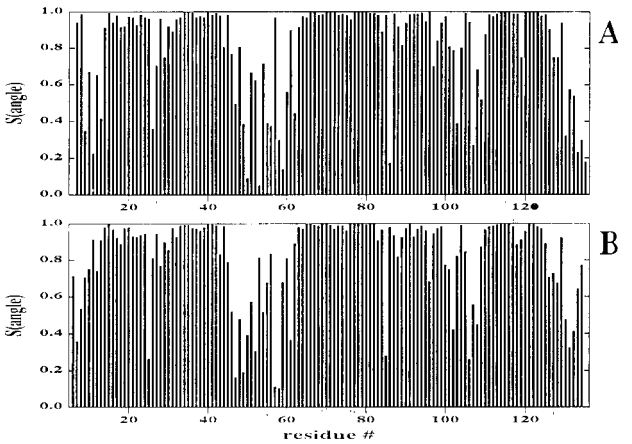


FIGURE 7: Angular order parameters for the $\phi$ (A) and $\psi$ (B) angles for the 15 low energy structures of the amino-terminal fragment of u-PA.

In general, the previously reported amide chemical shifts of the u-PA kringle are similar to those presented here for ATF (Table 1). However, significant proton chemical shift differences (>0.5 ppm) of the amides were observed for residues R59, N107, R108, and R109. In addition, several of the proton chemical shifts of the side chains of ATF differed from those reported for the u-PA kringle (Li et al., 1992). Nevertheless, in general, the secondary structure of the u-PA kringle in ATF was found to be similar to that previously described for the u-PA kringle by itself (Li et al., 1992), except for some minor deviations in the locations (±1 residue) of the helices.

*Domain–Domain Interactions.* Many proteins consist of multiple domains or modules with individual binding or catalytic functions that are linked together to form a large protein. Many of these modules have been studied in the absence of the rest of the protein, and only a relatively small amount of structural information is available on the interaction between modules in multidomain proteins. In a previous study of a protein containing two fibronectin type 1 modules, Williams et al. (1993) found 58 NOEs between the domains linked together by a regular turn. The relative orientation between the modules was found to be fixed with little or no reorientation of the intermodule linker. In a second study on the solution structure of a pair of complement modules, Barlow et al. (1993) observed numerous NOEs between both domains and the three amino acid linker region and between the two domains. The structure calculated with these NOEs showed a poorly ordered linker region and a limited range of intermodular orientations (Barlow et al., 1993).

The structural independence of the individual domains of u-PA has been studied by differential scanning calorimetry (Novokhatny et al., 1992, 1993) and 1D NMR (Bogusky et al., 1989). It was shown that the domains of u-PA unfold independently. In another study, Dobson and co-workers (Oswald et al., 1989; Nowak et al., 1993) concluded, on the basis of the line widths of tryptophan residues, that there is a high degree of motional freedom between the kringle and protease domains. In addition, it was suggested that the GFD and kringle domains are probably not highly correlated, but strong interactions could not be ruled out (Nowak et al., 1993).

One of the goals in the structure determination of the amino-terminal fragment of u-PA was to investigate the possibility of interdomain interactions between the growth factor and the kringle domains. Only three relatively weak NOEs were observed between the two domains (I44 $H^{\gamma 2}$–Y101 $H^{\epsilon}$, I44 $H^{\delta 1}$–Y101 $H^{\epsilon}$, I44 $H^{\delta 1}$–Y101 $H^{\delta}$). Initial structure calculations that included these NOEs are restraints converged to low energy structures. However, if these structures existed more than transiently, other short-range interdomain proton–proton interactions that were predicted from the structure are not observed in the NOE spectra. Further support of the structural independence of the two domains was found in the marked difference in line width and intensity of the peaks observed in many of the NMR spectra for the kringle and GFD domains. In general, the line width was greater for the NMR signals of the kringle compared to the growth factor, suggesting different correlation times for the two domains consistent with their different molecular weight. The lack of interdomain NOEs, the apparent difference in signal intensity and line widths of the peaks in the two domains, and the very broad signals observed in the linker region have lead us to conclude that the growth factor and kringle domains are structurally independent with very little interaction between them. The structural independence of the two domains is much greater than that previously observed in NMR studies of other multidomain proteins.

*Comparison of the u-PA Growth Factor to Homologous Proteins and Biological Implications of the Structure.* High-resolution solution structures have been obtained for both murine epidermal growth factor (EGF) at pH 3.1 (Montelione et al., 1992) and transforming growth factor $\alpha$ (TGF-$\alpha$) at pH 6.3 (Kline et al., 1990). Figure 11 compares the average, minimized NMR structure of the growth factor domain from u-PA (Figure 11A) to the NMR structures of TGF-$\alpha$ (Figure 11B) and murine EGF (Figure 11C). Although the overall sequence homology is weak between the u-PA growth factor

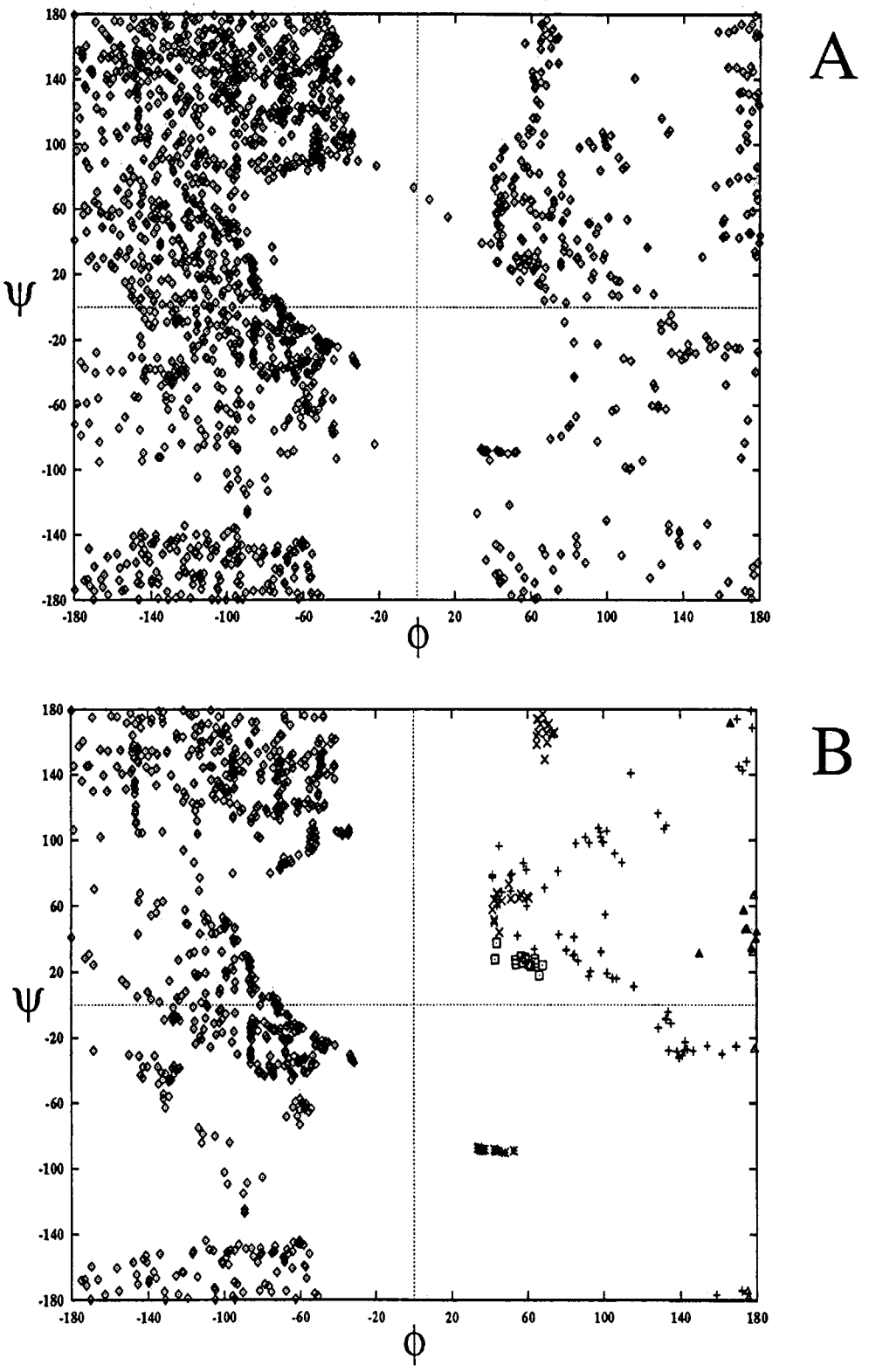


FIGURE 8: Ramachandran plot for the 15 NMR-derived structures of the amino-terminal fragment of u-PA for (A) all $\phi/\psi$ angles and (B) for those residues with $\phi$ and $\psi$ $S^{angle}$ values $\geq 0.9$ (63 residues). In panel B, positive $\phi$ angles are noted for glycine residues (+), N15 and N27 (□), Q40 (*), C42 and C71 (×), and H87 (△).

domain and the other two growth factors, structurally they are similar. All three can be divided into two subdomains (Moy et al., 1993), an amino-terminal subdomain composed of an antiparallel $\beta$-sheet (residues 6–33 in u-PA) and a carboxy-terminal subdomain composed of a double turn structure (residues 34–46). Although the amino-terminal $\beta$-sheet of the u-PA growth factor is two stranded, both TGF-$\alpha$ and EGF have a loosely attached third strand (residues 5–6 in TGF-$\alpha$ and 2–5 in EGF). For TGF-$\alpha$, this third strand was found to dissociate at low pH (Kline et al., 1990). Therefore, although we found no evidence for a third strand in the u-PA growth factor domain at pH 4.5, we cannot rule out its existence at higher pH. The other difference between these domains, which is apparent in Figure 11, is the relative orientation of the two subdomains. In EGF (Figure 11C), the carboxy-terminal subdomain more closely contacts the

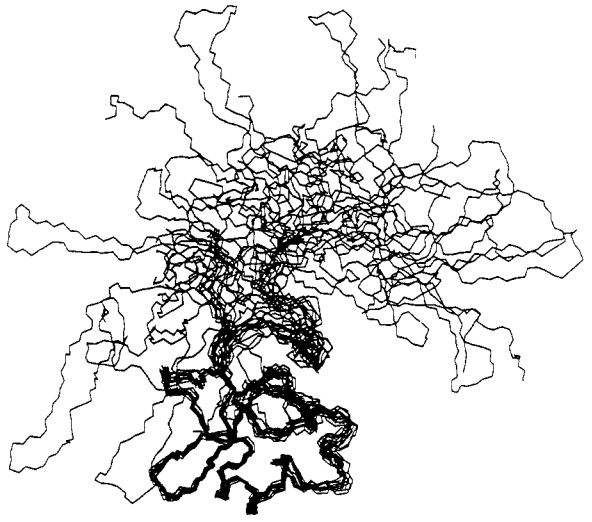

FIGURE 9: Superposition of the 15 NMR-derived structures of the amino-terminal fragment of u-PA where the Cartesian coordinates of the kringle domain have been superimposed onto the mean atomic coordinates for residues 55–130.

amino-terminal subdomain than in either TGF-$\alpha$ or the u-PA growth factor domain.

The most important difference between the u-PA growth factor domain and the other two growth factors is in the turn region of the major $\beta$-sheet (residues 22–28). From an analysis of the u-PA receptor affinity measured for a series of peptides and other growth factors (e.g., mouse ATF), Appella et al. (1987) showed that residues 20–30 confer the u-PA receptor binding specificity of the u-PA growth factor domain. Figure 12 depicts an overlay of the NMR structures of EGF, TGF-$\alpha$, and the u-PA growth factor for residues 19–31. For both EGF and TGF-$\alpha$, a type I $\beta$-bend was identified for S25–L26 and Q26–E27, respectively. In contrast, as shown in Figure 12, the two $\beta$-strands of the u-PA kringle are joined by a seven-residue omega loop. This difference in structure is most likely the consequence of a one amino acid insertion in the $\beta$-sheet region of the u-PA growth factor as compared to EGF and TGF-$\alpha$. It is possible that the high binding affinity for the u-PA receptor displayed by the u-PA growth factor compared to the negligible binding observed for EGF or TGF-$\alpha$ may be explained, in part, by this loop versus turn motif.

*Comparison of the u-PA Kringle to Homologous Proteins and Biological Implications of the Structure.* Kringle domains have been found in a number of proteins including tissue plasminogen activator (t-PA), plasminogen, prothrombin, factor XII, and apolipoprotein(a) (Patthy, 1985; McLean et al., 1987). In plasminogen and t-PA, these domains mediate interactions with fibrin that are critical for fibrinolysis, whereas the function of the kringle domains in the other proteins is less clear. High-resolution crystal structures have been obtained for the second kringle domain from t-PA (de Vos et al., 1992), for the fourth kringle domain from plasminogen (Mulichak et al., 1991), and for the kringle domain from prothrombin (Seshadri et al., 1991). In addition, the solution structure of the second kringle domain from t-PA has been reported (Byeon & Llinás, 1991). Of these kringles whose structures have been solved, the u-PA kringle domain shows the most homology ($\sim$46%) with the second kringle domain from t-PA and consists of the same number of residues. The homology between the u-PA kringle and the other kringle domains is less: $\sim$32% with the prothrombin kringle and $\sim$34% with the plasminogen kringle. Both of the other kringles contain amino acid insertions and deletions compared to the u-PA and t-PA kringles. Despite the relatively low sequence homology among the different kringles, the overall structures of these domains are very similar. In the u-PA kringle as well as the other kringles, a two-stranded antiparallel $\beta$-sheet is conserved. Also, reverse turns are located in similar positions in all of the proteins, although their exact classification may vary. In addition, many of the residues which form the hydrophobic core of the kringles are either identical or conservatively substituted. As might be expected on the basis of the sequence homology, the three-dimensional structure of the u-PA kringle most closely resembles the t-PA kringle. Despite the similarity, however, there are significant differences between the structures in the backbone conformation for residues A77–Q82 and D90–L94. Figure 13A shows an overlay of these regions for the u-PA kringle, the t-PA kringle, and plasminogen kringle 4. For the u-PA kringle domain (Figure 13A), residues A77–Q82 form an $\alpha$-helix, contrary to the other kringle domains which adopt a nonregular extended loop in this region. For residues D90–L94, both the u-PA kringle and t-PA kringle are composed of a distorted helix, whereas, in plasminogen kringle 4, this helix is absent due to a three residue deletion.

Despite the high structural similarity of the u-PA kringle domain to the t-PA kringle domain, these two kringles exhibit very different binding specificities. The t-PA kringle binds to positively charged lysine residues of fibrin, whereas the u-PA kringle binds to polyanionic molecules such as heparin. As determined by both NMR and X-ray crystallography, the lysine binding site is on the surface of the protein and consists of an anionic center and a cationic center with a hydrophobic cleft in between. As shown in Figure 13B, superposition of the lysine binding site of the t-PA kringle with the homologous residues of the u-PA kringle reveals that the hydrophobic and ionic character of this region is very different corroborating an earlier observation by Llinás and co-workers on the basis of modeling studies (Bokman et al., 1993). Two of the four aromatic side chains which form the lysine binding site in the t-PA kringle are substituted by a valine (residue 123 in ATF) and a glutamate (residue 125 in ATF). Also, in the catioinic center, a lysine residue in the t-PA kringle is replaced by a glutamine in the u-PA kringle. The aspartic acid residue which forms part of the anionic center (residue 108 in ATF) in the t-PA kringle is replaced by an arginine in u-PA.

These differences of amino acids in this region serve to mediate different interactions involving the kringle. It has been suggested (Stephens et al., 1992) on the basis of peptide competition assays and NMR titration experiments that three arginine residues (108, 109, 110) and three histidine residues (85, 87, 99) of the u-PA kringle are involved in binding to heparin. All six of these residues, whose side chains are depicted in Figure 10B, are located on the surface of the protein. Interestingly, the three arginine residues are in the vicinity of the "nonfunctional" lysine binding site of the u-PA kringle (Figure 13B) correspond to D57, A58, and K60 of the t-PA kringle. This may explain, in part, the affinity of the u-PA kringle for a highly anionic molecule such as heparin.

## CONCLUSIONS

The three-dimensional solution structure of the amino-terminal fragment of u-PA was determined by heteronuclear, multi-dimensional NMR methods. Both the assignments and structure determination were greatly aided by the use of uniformly $^{15}$N- and $^{13}$C/$^{15}$N-labeled ATF prepared from mammalian cells (Hansen et al., 1992). Although more difficult than the preparation of isotopically labeled proteins

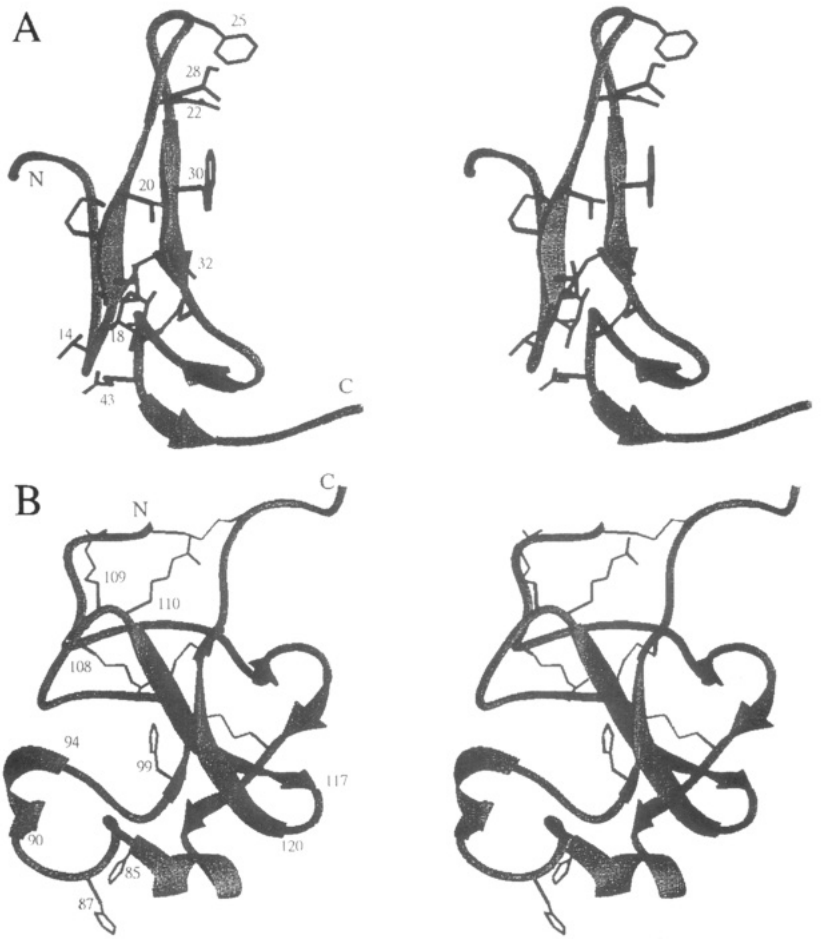


FIGURE 10: Stereoview ribbon diagram (Carson, 1987) of (A) growth factor and (B) kringle domain of ATF. Selected side chains and disulfide bonds are shown in black.

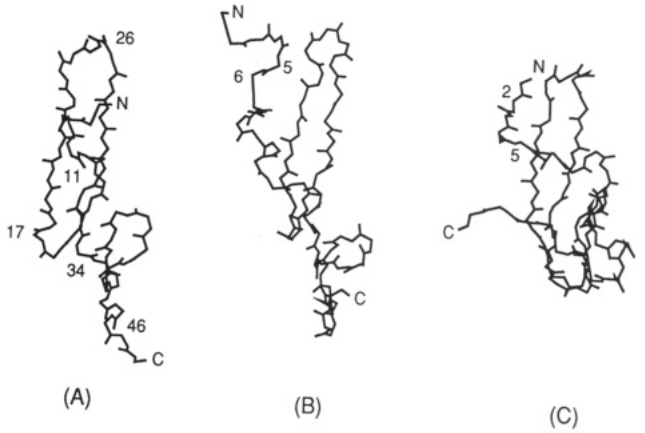


FIGURE 11: Comparison of the solution structures of the (A) u-PA growth factor domain to (B) TGF-$\alpha$ and (C) murine EGF.

from bacteria, isotope-labeling can be achieved for proteins expressed in mammalian cells and, as demonstrated here, can be very useful for determining the structures of larger proteins by NMR.

The three-dimensional structure of the amino-terminal fragment of u-PA is of considerable interest due to its important biological activities. Indeed, u-PA plays a critical role in metastatic tissue invasion which requires the binding of the growth factor domain to the urokinase receptor for localizing u-PA on the cell surface (Crowley et al., 1993; Duffy, 1993). Although the overall three-dimensional structure of the u-PA growth factor was found to be similar to homologous growth factors whose structures have been previously determined (Montelione et al., 1987; Cooke et al., 1987; Kline et al., 1990; Moy et al., 1993), key structural differences were observed in the turn region of the major $\beta$-sheet. It has been suggested (Appella et al., 1987) that this part of the u-PA growth factor domain is important for binding to the urokinase receptor. The structural differences observed between u-PA and the other growth factors could explain, in part, why other growth factors show no appreciable affinity for the u-PA receptor. Furthermore, the three-dimensional structure of the u-PA growth factor domain could aid in the design antimetastatic agents that block the interaction of u-PA with its receptor. However, care must be taken not to overinterpret the structure of the growth factor domain in the absence of the u-PA receptor as its conformation could be altered upon binding to the receptor. The most relevant structural information for designing potential antimetastatic agents will come from the 3D structure of the u-PA growth factor/receptor complex (work in progress).

As observed for the u-PA growth factor and homologous growth factor domains, the overall three-dimensional structure of the u-PA kringle was found to be similar to the X-ray and NMR structures of homologous kringles that have been previously determined (de Vos et al., 1992; Mulichak et al., 1991; Seshadri et al., 1991; Byeon & Llinás, 1991). Yet the biological activities of the different kringles are markedly different. For example, the plasminogen and t-PA kringles bind to positively charged lysine residues of fibrin, whereas the u-PA kringle binds to polyanionic molecules such as heparin (Stephens et al., 1992). As previously suggested (Bokman et al., 1993), these differences in binding can be explained by the presence of different amino acid residues located on the

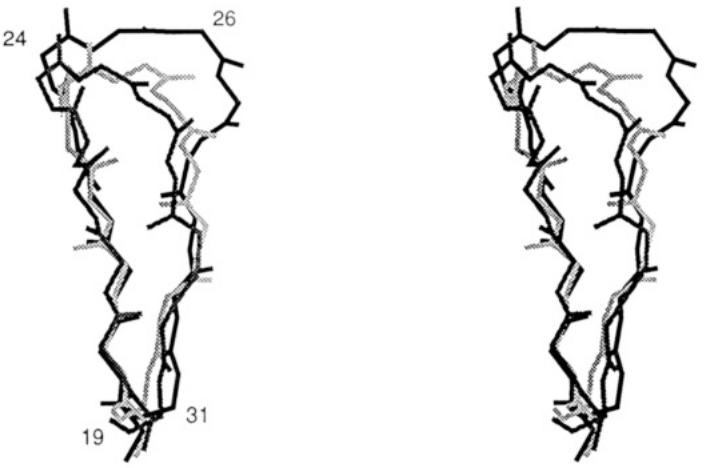


FIGURE 12: Overlay of the backbone of EGF (light gray) and TGF-$\alpha$ (dark gray) to residues 19–31 of u-PA growth factor domain (black). The amino acid residues of the u-PA growth factor domain are numbered.

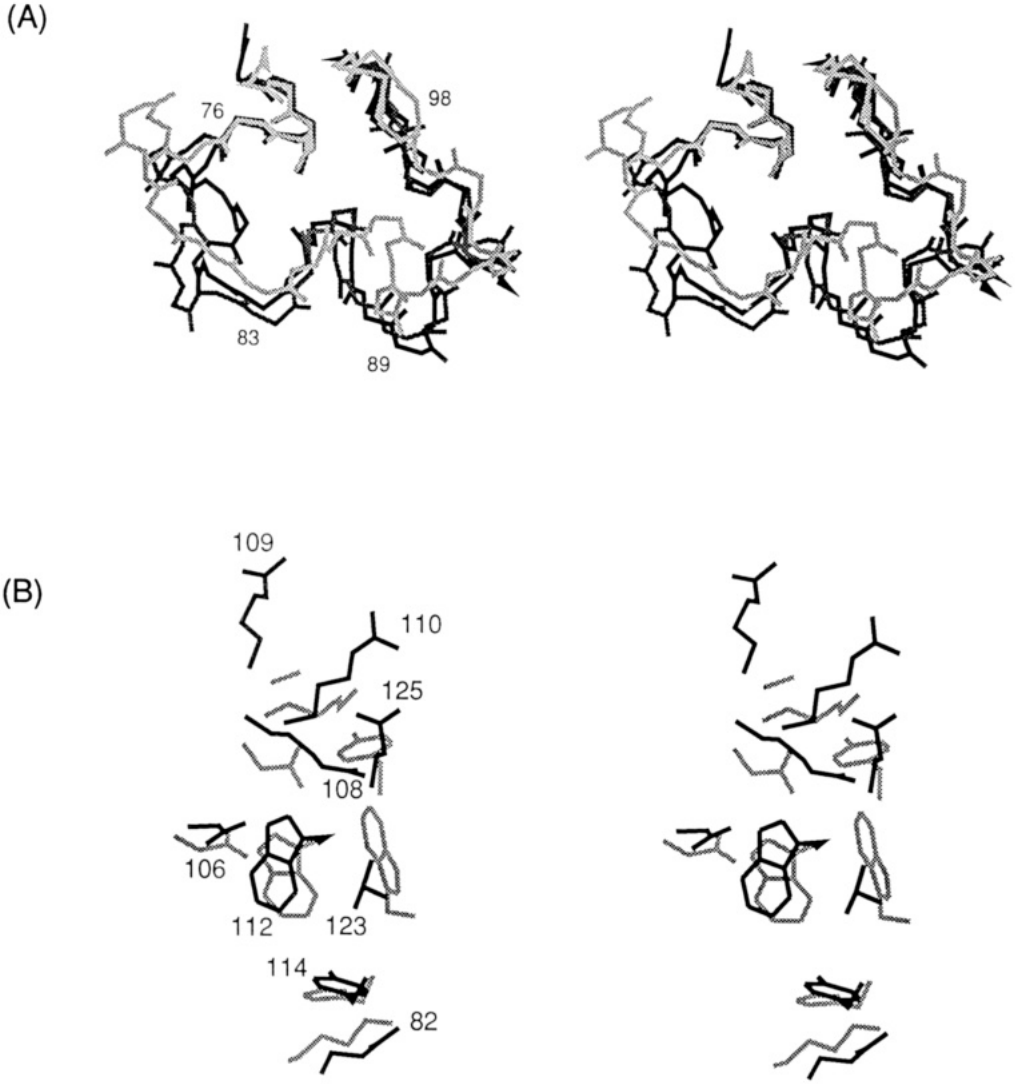


FIGURE 13: (A) Overlay of the backbone of plasminogen kringle 4 domain (light gray) and t-PA kringle 2 domain (dark gray) to residues 74–100 of the u-PA kringle domain (black). (B) Superposition of the side chains of the lysine-binding site of t-PA kringle 2 (gray) on the homologous region of the u-PA kringle domain.

surface of the proteins. The differences observed in the biological activities and binding specificities for proteins within the same structural class appears to be a general phenomenon and is consistent with the idea that the growth factor, kringle, and other structural motifs (e.g., SH2, SH3 domains) serve as scaffolds for proper positioning the side chains of key amino acids that dictate their biological functions (Baron et al., 1991). For some modular proteins (e.g., a protein containing two fibronectin type 1 modules), the orientation of the individual domains is fixed relative to one another (Williams et al., 1993). In contrast, the kringle and growth factor domains of u-PA appear to be structurally independent with very little inter-

domain interaction. By stringing together individual modules that have been tailored through mutagenesis for a particular function, proteins with remarkable biological activities can be efficiently designed. Experimentally determined three-dimensional structures of individual protein modules and proteins containing more than one module may help explain their exquisite binding specificities and unique biological functions.

## NOTE ADDED IN PROOF

An NMR-derived solution structure of the kringle domain of urokinase-type plasminogen activator has recently appeared (Li et al., 1994).

## REFERENCES


Appella, E., Robinson, E., Ulrich, S., Stoppelli, M., Corti, A., Cassani, G., & Blasi, F. (1987) *J. Biol. Chem. 262*, 4437–4440.

Archer, S. J., Ikura, M., Torchia, D. A., & Bax, A. (1991) *J. Magn. Reson. 95*, 636–641.

Archer, S. J., Bax, A., Roberts, A. B., Sporn, M. B., Ogawa, Y., Piez, K. A., Weatherbee, J. A., Tsang, M. L.-S., Lucas, R., Zheng, B.-L., Wenker, J., & Torchia, D. A. (1993) *Biochemistry 32*, 1152–1163.

Barlow, P. N., Steinkasserer, A., Norman, D. G., Kieffer, B., Wiles, A. P., Sim, R. B., & Campbell, I. D. (1993) *J. Mol. Biol. 232*, 268–284.

Baron, M., Norman, D. G., & Campbell, I. D. (1991) *Trends Biochem. Sci. 16*, 13–17.

Bax, A., & Ikura, M. (1991) *J. Biomol. NMR 1*, 99–104.

Bax, A., & Pochapsky, S. (1992) *J. Magn. Reson. 99*, 638–643.

Bax, A., Clore, G. M., & Gronenborn, A. M. (1990) *J. Magn. Reson. 88*, 425–431.

Birktoft, J. J., & Blow, D. M. (1972) *J. Mol. Biol. 68*, 187–240.

Bodenhausen, G., & Ruben, D. J. (1980) *Chem. Phys. Lett. 69*, 185–189.

Bogusky, M. J., Dobson, C. M., & Smith, R. A. G. (1989) *Biochemistry 28*, 6728–6735.

Bokman, A. M., Jimenez-Barbero, J., & Llinás, M. (1993) *J. Biol. Chem. 268*, 13858–13868.

Brünger, A. T. (1992) *X-PLOR 3.1 Manual*, Yale University Press, Yale University, New Haven, CT.

Buko, A. M., Kentzer, E. J., Petros, A. M., Menon, G., Zuiderweg, E. R. P., & Sarin, V. K. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 3992–3996.

Byeon, I.-J. L., & Llinás, M. (1991) *J. Mol. Biol. 222*, 1035–1051.

Campbell, I. D., & Bork, P. (1993) *Curr. Opin. Struct. Biol. 3*, 385–392.

Carson, M. (1987) *J. Mol. Graphics 5*, 103.

Clore, G. M., & Gronenborn, A. M. (1991) *Science 252*, 1390–1399.

Clubb, R. T., Thanabal, V., Osborne, C., & Wagner, G. (1991) *Biochemistry 30*, 7718–7730.

Cooke, R. M., Wilkinson, A. J., Baron, M., Pastore, A., Tappin, M. J., Cambell, I. D., Gregory, H., & Sheard, B. (1987) *Nature 327*, 339–341.

Crowley, C., Cohen, R., Lucas, B., Liu, G., Shuman, M., & Levinson, A. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 5021–5025.

Curley, R. W., Panigot, M. J., Hansen, A. P., & Fesik, S. W. (1994) *J. Biomol. NMR* (in press).

de Munk, G. A. W., & Rijken, D. C. (1990) *Fibrinolysis 4*, 1–9.

de Vos, A. M., Ultsch, M. H., Kelley, R. F., Padmanabhan, K., Tulinsky, A., Westbrook, M. L., & Kossiakoff, A. A. (1992) *Biochemistry 31*, 270–279.

Duffy, M. J. (1993) *Fibrinolysis 7*, 295–302.

Fesik, S. W., & Zuiderweg, E. R. P. (1988) *J. Magn. Reson. 87*, 588–593.

Grzesiek, S., & Bax, A. (1992) *J. Am. Chem. Soc. 114*, 6291–6293.

Grzesiek, S., Vuister, G. W., & Bax, A. (1993) *J. Biomol. NMR 3*, 487–493.

Haber, E., Quertermous, T., Matsueda, G. R., & Runge, M. S. (1989) *Science 243*, 51–56.

Hansen, A. P., Petros, A. M., Mazar, A. P., Pederson, T. M., Rueter, A., & Fesik, S. W. (1992) *Biochemistry 31*, 12713–12718.

Huber, R., Kukla, D., Bode, W., Schwager, P., Bartel, K., Diesenhofer, J., & Steigemann, W. (1974) *J. Mol. Biol. 89*, 73–101.

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

Ikura, M., Kay, L. E., Tschudin, R., & Bax, A. (1990) *J. Magn. Reson. 86*, 204–209.

Jankun, J., Merrick, H. W., & Goldblatt, P. J. (1993) *J. Cell. Biochem. 53*, 135–144.

Kay, L. E., & Bax, A. (1990) *J. Magn. Reson. 86*, 110–126.

Kay, L. E., Ikura, M., Tschudin, R., & Bax, A. (1990) *J. Magn. Reson. 89*, 496–514.

Kline, T. P., Brown, F. K., Brown, S. C., Jeffs, P. W., Kopple, K. D., & Mueller, L. (1990) *Biochemistry 29*, 7805–7813.

Leszczynski, J. F., & Rose, G. D. (1986) *Science 234*, 849–855.

Li, X., Smith, R. A. G., & Dobson, C. M. (1992) *Biochemistry 31*, 9562–9571.

Li, X., Bokman, A. M., Llinás, M., Smith, R. A. G., & Dobson, C. M. (1994) *J. Mol. Biol. 235*, 1548–1559.

Lo, K.-M., & Gillies, S. D. (1991) *Biochim. Biophys. Acta 1088*, 217–224.

Marion, D., Kay, L. E., Sparks, S. E., Torchia, D. A., & Bax, A. (1989a) *J. Am. Chem. Soc. 111*, 1515–1517.

Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., & Clore, G. M. (1989b) *Biochemistry 28*, 6150–6156.

Massague, J. (1983) *J. Biol. Chem. 258*, 13614–13620.

Masucci, M., & Blasi, F. (1990) *Thromb. Res. (Suppl. XI)*, 49–60.

Mazar, A. P., Buko, A., Petros, A. M., Barnathan, E. S., & Henkin, J. (1992) *Fibrinolysis 6 (Suppl. 1)*, 49–55.

McLean, J. W., Tomlinson, J. E., Kuang, W.-J., Eaton, D. L., Chen, E. Y., Fless, G. M., Scanu, A. M., & Lawn, R. M. (1987) *Nature 230*, 132–137.

Meadows, R. P., Nettesheim, D. G., Xu, R. X., Olejniczak, E. T., Petros, A. M., Holzman, T. F., Severin, J., Gubbins, E., Smith, H., & Fesik, S. W. (1993) *Biochemistry 32*, 754–765.

Messerle, B. A., Wider, G., Otting, G., Weber, C., & Wüthrich, K. (1989) *J. Magn. Reson. 85*, 608–613.

Montelione, G. T., Wüthrich, K., Nice, E. C., Burgess, A. W., & Sheraga, H. A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 5226–5230.

Montelione, G. T., Wüthrich, K., Burgess, A. W., Nice, E. C., Wagner, G., Gibson, K. D., & Sheraga, H. A. (1992) *Biochemistry 31*, 326–249.

Moy, F. J., Li, Y.-C., Rauenbuehler, P., Winkler, M. E., Scheraga, H. A., & Montelione, G. T. (1993) *Biochemistry 32*, 7334–7353.

Mulichak, A. M., Tulinsky, A. & Ravichandran, K. G. (1991) *Biochemistry 30*, 10576–10588.

Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988) *FEBS Lett. 229*, 317–324.

Novokhatny, V., Medved, L., Mazar, A., Marcotte, P., Henkin, J., & Ingham, K. (1992) *J. Biol. Chem. 267*, 3878–3885.

Novokhatny, V. V., Medved, L. V., Mazar, A., & Ingham, K. C. (1993) *J. Biol. Chem. 268*, 17211–17218.

Nowak, U. K., Li, X., Teuten, A. J., Smith, R. A. G., & Dobson, C. M. (1993) *Biochemistry 32*, 298–309.

Orsini, G., Brandazza, A., Sarmientos, P., Molinari, A., Lansen, J., & Cauet, G. (1991) *Eur. J. Biochem. 195*, 691–697.

Ossowski, L. (1988) *J. Cell. Biol. 107*, 2437–2445.

Oswald, R. E., Bogusky, M. J., Bamberger, M., Smith, R. A. G., & Dobson, C. M. (1989) *Nature 337*, 579–582.

Patthy, L. (1985) *Cell 41*, 657–663.
Powers, R., Garrett, D. S., March, C. J., Frieden, E. A., Gronenborn, A. M., & Clore, G. M. (1993) *Biochemistry 32*, 6744–6762.
Sawyer, L., Shotton, D. M., Campbell, J. W., Wendel, P. L., Muirhead, H., Watson, H. C., Diamond, R., & Ladner, R. C. (1978) *J. Mol. Biol. 118*, 137–208.
Seshadri, T. P., Tulinsky, A., Skrzypczak-Jankun, E., & Park, C. H. (1991) *J. Mol. Biol. 220*, 481–494.
Shotton, D. M., & Watson, H. C. (1970) *Nature 225*, 811–816.
Spera, S., & Bax, A. (1991) *J. Am. Chem. Soc. 113*, 5490–5492.
Stephens, R. W., Bokman, A. M., Myöhänen, H. T., Reisberg, T., Tapiovaara, H., Pedersen, N., Grøndahl-Hansen, J., Llinás, M., & Vaheri, A. (1992) *Biochemistry 31*, 7572–7579.
Stroud, R. M., Kay, L. M., & Dickerson, R. E. (1974) *J. Mol. Biol. 83*, 185–208.
Theriault, Y., Logan, T. M., Meadows, R., Yu, L., Olejniczak, E. T., Holzman, T. F., Simmer, R. L., & Fesik, S. W. (1993) *Nature 361*, 88–91.
Towle, M. J., Lee, A., Moduakor, E. C., Schwartz, C. E., Bridges, A. J., & Littlefield, B. A. (1993) *Cancer Res. 53*, 2553–2559.
Tulinsky, A., Mani, N. V., Morimoto, C. N., & Vandlen, R. L. (1973) *Acta Crystallogr. Sect. B 29*, 1309–1322.
Vassalli, J.-D., Baccino, D., Belin, D. (1985) *J. Cell. Biol. 100*, 86–92.
Vuister, G. W., & Bax, A. (1993) *J. Am. Chem. Soc. 115*, 7772–7777.
Williams, M. J., Phan, I., Baron, M., Driscoll, P. C., & Campbell, I. D. (1993) *Biochemistry 32*, 7388–7395.
Winkler, M. E., & Blaber, M. (1986) *Biochemistry 25*, 4041–4045.
Wüthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol. 169*, 949–961.
Zuiderweg, E. R. P., Boelens, R., & Kaptein, R. (1985) *Biopolymers 24*, 601–611.
Zuiderweg, E. R. P., McIntosh, L. P., Dahlquist, F. W., & Fesik, S. W. (1990) *J. Magn. Reson. 86*, 210–216.