Sturtevant, J. M. (1987) *Annu. Rev. Phys. Chem. 38*, 463–488.
Surewicz, W. K., & Mantsch, H. H. (1988) *Biochim. Biophys. Acta 952*, 115–130.
Surewicz, W. K., & Mantsch, H. H. (1990) in *Protein Engineering—Approaches to the Manipulation of Protein Folding* (Narang, S., Ed.) pp 131–157, Butterworth, Stoneham, MA.
Surewicz, W. K., Szabo, A. G., & Mantsch, H. H. (1987) *Eur. J. Biochem. 167*, 519–523.
Surewicz, W. K., Stepanik, T. M., Szabo, A. G., & Mantsch, H. H. (1988) *J. Biol. Chem. 263*, 786–790.
Susi, H. (1969) in *Structure and Stability of Biological Macromolecules* (Timasheff, S. N., & Fasman, G., Eds.) pp 575–663, Marcel Dekker, New York.
Tomasi, M., Battistini, A., Cardelli, M., Sonnino, S., & D'Agnolo, G. (1984) *Biochemistry 23*, 2520–2526.
van Heyningen, W. E., Carpenter, C. C. J., Pierce, N. F., & Greenough, W. (1971) *J. Infect. Dis. 124*, 415–418.

^{1}H NMR Assignment and Secondary Structure of the Ca^{2+}-Free Form of the Amino-Terminal Epidermal Growth Factor like Domain in Coagulation Factor X†

Maria Selander,‡ Egon Persson,§ Johan Stenflo,§ and Torbjörn Drakenberg*,‡

Department of Physical Chemistry 2, Chemical Centre, University of Lund, Box 124, S-221 00 Lund, Sweden, and Department of Clinical Chemistry, Malmö General Hospital, University of Lund, S-214 01 Malmö, Sweden

Received February 15, 1990; Revised Manuscript Received May 22, 1990

ABSTRACT: Blood coagulation factor X is composed of discrete domains, two of which are homologous to the epidermal growth factor (EGF). The N-terminal EGF like domain in factor X (fX-EGF_N), residues 45–86 of the intact protein, contains a β-hydroxylated aspartic acid and has one Ca^{2+}-binding site. Using 2D NMR techniques, we have made a full assignment of the 500-MHz ^{1}H NMR spectrum of Ca^{2+}-free fX-EGF_N. On the basis of this assignment and complementary NOESY experiments, we have also determined the secondary structure of Ca^{2+}-free fX-EGF_N in water solution. Residues 45–49 are comparatively mobile, whereas residues 50–56 are constrained by two disulfide bonds to one side of an antiparallel β-sheet involving residues 59–64 and 67–72. Another antiparallel β-sheet involves residues 76–77 and 83–84. A small, parallel β-sheet connects residues 80–81 and 55–56 and thereby orients the two antiparallel β-sheets relative to each other. Four β-turns are identified, involving residues 50–53, 56–59, 64–67, and 73–76. Residues 78–82 adopt an extended bend structure. On the basis of secondary structure and the location of the three disulfide bonds, we find that Asp 46, Asp 48, and Hya 63 are sufficiently close to each other to form a Ca^{2+}-binding site. However, the amino terminus of the Ca^{2+}-free form of fX-EGF_N is not part of a triple-stranded β-sheet as in other EGF like peptides. Differences and similarities between fX-EFG_N and murine EGF with respect to secondary structure and conformational shifts are discussed.

During the past years, a large number of extracellular and membrane proteins have been found to contain domains homologous to the epidermal growth factor (EGF)[1] (Appella et al., 1988). Among these are proteins involved in blood coagulation, i.e., factors VII, IX, X, and XII and protein C (Furie & Furie, 1988), and in fibrinolysis, such as urokinase and the tissue-type plasminogen activator (Patthy, 1985). In particular, factors IX and X and protein C have two tandemly arranged EGF like domains, where the N-terminal one contains a hydroxylated aspartic acid residue (*erythro*-β-hydroxyaspartic acid, Hya) formed by postribosomal hydroxylation of aspartic acid (Drakenberg et al., 1983; McMullen et al., 1983). The Hya residue is always found in the consensus sequence Cys-X-Hya-X-X-X-X-Tyr/Phe-X-Cys-X-Cys (Stenflo et al., 1987), and the EGF like, Hya-containing domains usually also have two Asp/Glu N-terminal to the first Cys residue (Figure 1).

It has recently been demonstrated that the Hya-containing EGF like domains in factor IX (Huang et al., 1989; Handford et al., 1990), factor X (Persson et al., 1989), and protein C (Öhlin et al., 1988) bind Ca^{2+}. A Ca^{2+}-dependent interaction between a monoclonal antibody and the isolated EGF like region of protein C suggests that binding of Ca^{2+} imposes structural changes in the domain (Öhlin & Stenflo, 1987). This local change also seems to cause a global structural change in the protein, since binding of Ca^{2+} to the EGF like domain has a dramatic effect on the rate of activation of protein C by thrombin–thrombomodulin (Esmon et al., 1983).

Coagulation factor X, the subject of the study, is a vitamin K dependent zymogen of a serine protease. In its active form

† This work was supported by the Swedish Natural Science Research Council (K-KU-2545-300), the Swedish Medical Research Council (B87X-04487-13B), and the Ekedahl-Lundberg Foundation. The NMR spectrometer was purchased with generous grants from the Knut and Alice Wallenberg Foundation and the Swedish Council for Planning and Coordination of Research (FRN).

* To whom correspondence should be addressed.
‡ Department of Physical Chemistry 2.
§ Department of Clinical Chemistry.

[1] Abbreviations: NMR, nuclear magnetic resonance; COSY, *J*-correlated spectroscopy; NOE, nuclear Overhauser enhancement; NOESY, two-dimensional nuclear Overhauser enhancement spectroscopy; TOCSY, total correlation spectroscopy; R-COSY, relayed COSY; DR-COSY, double-relayed COSY; 2QF-COSY, double-quantum-filtered COSY; 2Q, double-quantum; 2D, two-dimensional; EGF, epidermal growth factor; TGFα, transforming growth factor α; Hya, *erythro*-β-hydroxyaspartic acid; Hyn, *erythro*-β-hydroxyasparagine; Gla, γ-carboxyglutamic acid.

0006-2960/90/0429-8111$02.50/0

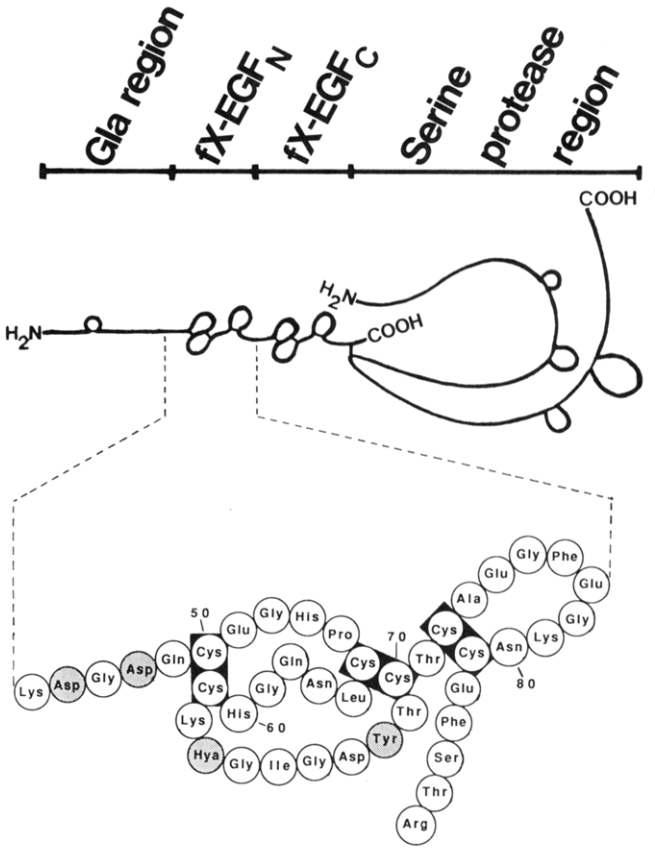

FIGURE 1: Schematic representation of bovine factor X with a blowup of the isolated N-terminal EGF like domain fX-EGF$_N$ corresponding to residues 45–86 of the intact protein. Residues Asp 46 and 48 and Hya[1] 63, which are thought to be involved in Ca^{2+} binding, are shaded, as well as Tyr 68 required by the aspartyl β-hydroxylase.

it constitutes the enzymatically active part of the prothrombinase complex that activates prothrombin into the serine protease thrombin (Mann et al., 1988). Factor X, analogous to factor IX and protein C, is composed of four distinct domains (Figure 1). The N-terminal domain contains 12 γ-carboxyglutamic acids (Gal[1]), binds Ca^{2+}, and is required for the membrane interaction that is a prerequisite for biological activity (Stenflo & Suttie, 1977; Suttie et al., 1985). The Gla-containing domain is followed by two EGF like domains where the N-terminal one contains a Hya residue (Drakenberg et al., 1983; McMullen et al., 1983). The isolated EGF like Hya-containing domain binds Ca^{2+} with an association constant of 1.3×10^3 M^{-1} (Persson et al., 1989), comparable to the Ca^{2+} affinity of factor X without the Gla-containing region (Sugo et al., 1984). The C-terminal domain contains the serine protease part of the protein.

Attempts at crystallizing factor X or any of the proteins with the same domain structure have been fraught with difficulties. Therefore, little is known about the structure–function relationships in these molecules, including the role of the EGF-like domains. In contrast, the tertiary structures of murine EGF (Montelione et al., 1987; Kohda et al., 1988) and human EGF (Cooke et al., 1987) have been successfully determined by 2D NMR[1] techniques and found to be very similar. The structure of TGFα,[1] an EGF homologue, has also been studied by 2D NMR (Kohda et al., 1989; Montelione et al., 1989; Tappin et al., 1989; Brown et al., 1989) and shows similar features of secondary structure. There is no evidence that human EGF, murine EGF, or TGFα bind Ca^{2+} and none of them contain Hya or the Asp residues postulated to take part in Ca^{2+} binding. However, if the structures of the Ca^{2+}-binding EGF like domains found in factors IX and X and protein C are similar to that of human EGF, residues Asp 46, Asp 48, and Hya 63 (Figure 1) would lie on the same side of a triple-stranded β-sheet and thus be able to form a Ca^{2+}-binding site (Cooke et al., 1987). On the other hand, the sequence similarity between the Ca^{2+}-binding EGF like domain of factor X and mouse EGF is sparse except for the identical arrangement of the disulfide bonds (Höjrup et al., 1987). High-resolution structures of the Ca^{2+}-binding EGF like domains with and without Ca^{2+} are thus needed as a first step to investigate the structural role of these domains in blood coagulation.

In the present work, a complete sequence-specific resonance assignment of the ^1H NMR spectrum at pH 5.8 of the Ca^{2+}-free form of the amino-terminal EGF like domain of bovine factor X (fX-EGF$_N$) is presented together with the secondary structure as obtained from NMR data. Secondary structure and conformational shifts of fX-EGF$_N$ are compared with those of murine EGF as determined by Montelione et al. (1986, 1987, 1988).

Materials and Methods

The amino-terminal EGF like domain of bovine factor X (fX-EGF$_N$), corresponding to residues 45–86 of the intact protein (Figure 1), was isolated from enzymatic digests of factor X as described by Persson et al. (1989). The lyophilized protein samples were dissolved in 0.45 mL of H_2O, containing 5% (v/v) ^2H_2O for the lock, or in 100% ^2H_2O, to give a protein concentration of 4–5.5 mM. To adjust the pH, microliter amounts of 0.05 M NaOH, HCl, NaO^{2}H, or ^{2}HCl were added. The pH value was not corrected for isotopic effects. There was no evidence of aggregation as judged by the narrow line widths in the NMR spectra. The Ca^{2+} concentration was determined by atomic absorption spectroscopy and was less than 8% of the protein concentration for each sample. As the Ca^{2+} binding constant of isolated fX-EGF$_N$ is 1.3×10^3 M^{-1} at pH 7.5 (Persson et al., 1989) and probably weaker at pH 5.8, differences in shift between the Ca^{2+}-loaded and Ca^{2+}-free forms must be larger than 0.4 ppm in order to affect the shifts by more than 0.02 ppm. Such large shift changes in all likelihood involve displacement of an aromatic group, for which there is no evidence in previous titrations (Persson et al., 1989). Sequence analysis demonstrated that there were no internal cleavages in the fragment and that any contaminating sequence was below the 2% level. No additional resonances due to impurities or heterogenity in the protein were observed during the assignment procedure.

A General Electric Omega 500 spectrometer working at 500.13 MHz for ^1H was used for the NMR experiments. The probe was thermostated to 28 °C unless otherwise stated. Standard pulse sequences and phase cycling were used to obtain hypercomplex phase-sensitive COSY[1] (Marion & Wüthrich, 1983), R-COSY,[1] and DR-COSY[1] (Wagner, 1983) in ^1H_2O and 2QF-COSY[1] (Rance et al., 1983) in ^2H_2O. NOESY[1] experiments were performed as described by Bax (1985) with a mixing time of 150 ms and TOCSY[1] experiments as described by Bax & Davis (1985). 2Q[1] experiments were performed by using the pulse sequence and phase cycling described by Braunschweiler et al. (1983) with a composite 180° pulse (Levitt & Freeman, 1979).

In each experiment a total of 512 spectra were collected, each with 2048 complex points. In the COSY experiments, 64 scans were acquired per spectrum, whereas in all other experiments 96 scans were acquired. The carrier was set on the water resonance for all experiments. The water resonance was normally suppressed by selective irradiation by using the

decoupler during a preparation period of 1.3 s.

In processing the data, sine-bell window functions were used in both dimensions for sine-modulated datasets, whereas phase-shifted (typically 60°) sine-bell functions were applied to cosine-modulated data sets. Zero filling in the evolution dimension was used so that the final 2D spectra contained 1024 × 2048 complex data points. This resulted in a resolution of 6 and 3 Hz in ω_1 and ω_2, respectively.

The strategy used in the spin system assignment procedure was according to Chazin et al. (1988). In this strategy, relayed scalar connectivities from the NH protons to the α-protons and side-chain protons are observed in COSY, R-COSY, DR-COSY, and TOCSY experiments in ^1H_2O to determine the shifts for most of the side-chain resonances. For long-side-chain spin systems like Lys and Leu, it is often necessary to perform the same experiments in ^2H_2O and observe the relayed connectivities from $C^\alpha H$ and terminal protons to the side-chain protons. The residual assignment is then completed by observing the coincidence of relayed connectivities. The aromatic residues have two spin systems, which are not joined by indirect spin–spin couplings. In the NOESY experiment, however, dipolar coupling between hydrogens in the aromatic ring and the C^β protons gives rise to crosspeaks that connect the side chain with the aromatic ring so that the assignment can be completed. The same procedure is used in the assignment of the side-chain amide protons of Asn and Gln. Finally, a sequential assignment is made by observing the sequential NOE's[1] between adjacent backbone hydrogens as described by Billeter et al. (1982).

Slowly exchanging amide protons were identified by an exchange experiment in which a sample lyophilized from ^1H_2O was dissolved in ^2H_2O and the remaining resonances were observed in 1D and 2D spectra. Two classes of slow exchange were distinguished: one containing the amide protons that are still visible in a 1D spectrum half an hour after dissolution in ^2H_2O (intermediate exchange) and the other containing those that can also be observed in a 2.5-h COSY started 45 min after dissolution (slow exchange).

Backbone NH/$C^\alpha H$ coupling constants were measured at pH 5.0 from a 1D spectrum with a resolution of 1.5 Hz/point. The pH was chosen to minimize overlap among the NH resonances, which were identified from a COSY experiment at the same pH. Only minor shift changes were observed when the pH was lowered from 5.8 to 5.0, indicating no significant structural changes.

Results

In this presentation of the assignment and secondary structure of fX-EGF_N, the amino acid residues are numbered as in the intact protein (Figure 1).

Spin System Assignment. Spectra of fX-EGF_N, which has 42 amino acid residues, are expected to give 40 crosspeaks in the NH–$C^\alpha H$ region of a COSY spectrum, as neither the N-terminal residue nor Pro 54 gives rise to NH–$C^\alpha H$ crosspeaks. In a COSY spectrum at 28 °C, 38 crosspeaks were found, one of which was later identified as the $N^\delta H$–$C^\gamma H$ crosspeak of Arg 86. Two of the remaining three crosspeaks were not observable at 28 °C due to interference with the water resonance (His 60, Lys 62). However, relayed COSY and TOCSY experiments provided NH–$C\beta H$ and other side-chain crosspeaks, which made it possible to assign the NH and side-chain shifts, whereas the $C^\alpha H$ shifts were obtained from ^2H_2O experiments. No relayed scalar connectivities to the amide proton could be detected for Asp 46 in the COSY experiments, probably due to fast exchange of this amide proton. Thus, for three residues, Lys 45, Asp 46, and Pro 54, the assignment of the side-chain shifts had to be based on the $C^\alpha H$ resonances.

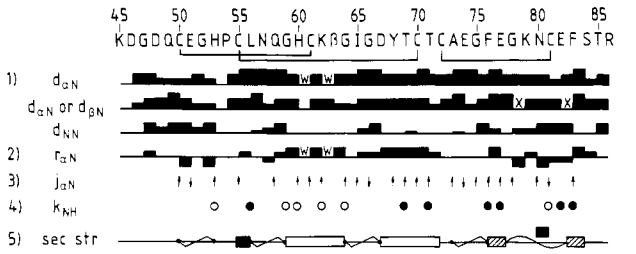

FIGURE 2: Survey of NMR data used for sequential assignment and identification of secondary structure elements in fX-EGF_N. The Hya residue is marked β. From top to bottom: (1) Sequential $d_{N\alpha}$, $d_{N\alpha'}$, $d_{N\beta}$, and d_{NN} NOE's where the height of the bar represents the size of the NOE. Interference with the water resonance is marked W, whereas lack of information due to overlap is marked X. (2) The parameter $r_{\alpha N}$ illustrates the strength of sequential C^α_{i-1}–NH_i NOE's relative to internal $C_{\alpha i}$–NH_i NOE's. Positive bars indicate that the sequential NOE was significantly larger than the internal NOE, whereas for negative bars the internal NOE was the larger one. When no significant difference between the two NOE's is observed, no bar is drawn. (3) $J_{NH-\alpha}$ coupling constants larger than 8 Hz are labeled ↑, while those smaller than 6 Hz are labeled ↓. (4) Amide proton exchange rates, k_{NH}, classified as intermediate are marked with empty circles; slowly exchanging amide protons are marked with filled circles. (5) A schematic diagram of the secondary structure of fX-EGF_N. β-Turns are drawn as jagged lines. Residues involved in β-sheets are enclosed in a box with markings according to which strands lie opposite each other in the β-sheet. The extended bend structure identified in fX-EGF_N is shown as wavy lines.

From COSY, R-COSY, and DR-COSY experiments in ^1H_2O, NH, $C^\alpha H$, and $C^\beta H$, shifts were obtained for all residues except for Lys 45, Asp 46, and Pro 54. TOCSY experiments in ^1H_2O with mixing times of 40 and 70 ms provided sufficient information to assign the C^γ protons. In order to complete the assignments of the six residues with C^ϵ and/or C^δ protons, a TOCSY experiment in ^2H_2O with a mixing time of 60 ms was performed and the scalar connectivities to the α-protons were analyzed. This experiment was also necessary to assign the proline spin system and the exact shifts of α-protons close to the water resonance. The spin system assignments were confirmed from an analysis of a 2QF-COSY spectrum in ^2H_2O. Four residues, Gln 49, Asn 57, Gln 58, and Asn 80, had side-chain amide groups that were assigned from a NOESY spectrum in ^1H_2O. From this spectrum, the ring systems of Tyr 68, Phe 76, and Phe 83 were also connected to their respective side chains. Within the aromatic rings, the δ-proton resonance assignments were based on this NOE's to the C^β protons. The ring shifts of the two histidines were the same at this pH. For the spin systems where only one $C^\beta H$ shift was found, possible degeneracy of the C^β protons was examined in a 2Q experiment ^2H_2O. Five of these amino acid residues had degenerate β-protons, whereas for the others the lacking β-shift was found from the $\alpha/(\beta 1 + \beta 2)$ crosspeak.

Sequence-Specific Assignments. The sequence-specific assignment of fX-EGF_N, given in Table I, was based on sequential NOE's as collected from three NOESY spectra in ^1H_2O at 25 °C, 28 and 33 °C. The use of the different temperatures was very helpful as ambiguities due to overlap could often be resolved because of the strong and different temperature dependences of the amide protons. A summary of the sequential connectivities as given in Figure 2 shows that only a few short stretches could be connected by means of NH–NH NOE's. In contrast, almost all sequential $C^\alpha H_i$–NH_{i+1} and $C^\beta H_i$–NH_{i+1} NOE's were observed and, therefore, the sequential assignment relied primarily on these observations (Figure 3).

Slowly Exchanging Amide Protons. At pH 5.8, all amide

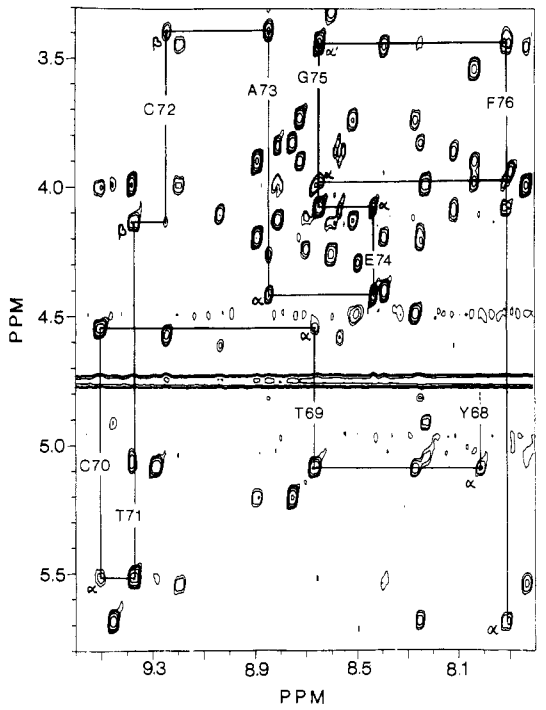

FIGURE 3: Region of a NOESY spectrum at 28 °C showing NH–C$^{\alpha}$H and NH–C$^{\beta}$H crosspeaks. The sequential assignment procedure is demonstrated for residues Tyr 68–Phe 76. Intraresidual crosspeaks are marked α or β. The residue number corresponding to the NH shift is shown on the vertical line. The C$^{\alpha}$H resonances of residues Thr 71 and Cys 72 are too close to the water resonance at 28 °C to observe any sequential NH–C$^{\alpha}$H crosspeaks. However, sequential NH–C$^{\beta}$H crosspeaks are observed, and the sequential NH–C$^{\alpha}$H crosspeaks were observed at other temperatures.

protons exchanged completely with ^2H_2O within half an hour in an exchange-out experiment as described above. This is too short a time to enable identification of slowly exchanging amide protons in a 2D experiment. The pH was therefore lowered to 3.8 in order to reduce the rate of exchange (Englander et al., 1972) and the experiment was repeated. Shift changes due to the change in pH from 5.8 to 3.8 were found to be less than 0.05 ppm for the NH resonances observed in the COSY spectrum. Identification of the amide protons whose resonances could only be seen in the 1D spectrum was therefore made on the basis of the assignments at pH 5.8 together with an analysis of the coupling constants. Between pH 3.8 and 5.8 the four carboxylic side chains probably ionize, two of which are situated close to the N-terminus and, as will be shown, do not take part in any well-defined secondary structure. As the shifts of the NH resonances are similar at pH 3.8 and 5.8, we expect that the change in pH does not affect the secondary structure to any large extent and that the amide protons that exchange slowly at pH 5.8 also do so at pH 3.8 but on a longer time scale. No broadening of the amide proton resonances was observed on lowering the pH as found for TGFα, where there are differences in the structure at pH 3.8 and 6.5 (Tappin et al., 1989). The amide protons that exchange slowly at pH 3.8 are given in Figure 2.

Determination of Secondary Structure. The secondary structure determination for fX-EGF$_N$ was primarily based on nonsequential NH–NH, NH–C$^{\alpha}$H, NH–C$^{\beta}$H, and C$^{\alpha}$H–C$^{\alpha}$H NOE connectivities. No α-helical structure was found as judged by the absence of C$^{\alpha}$H$_i$–NH$_{i+3}$ NOE's (Wüthrich, 1986). Eighteen out of 42 residues take part in three separate β-sheets, identified by their characteristic interstrand NOE's (Wüthrich et al., 1984). As a slowly exchanging amide proton is possibly, but not necessarily, involved in hydrogen bonding (Englander et al., 1972), no hydrogen bonds were postulated without supportative NOE data.

The secondary structure of fX-EGF$_N$ is shown in Figure 4 together with the NOE connectivities used for the structure determination. Information concerning coupling constants, slow exchange of amide protons, and relative strength of sequential NH–C$^{\alpha}$H NOE's as compared to intraresidual NH–

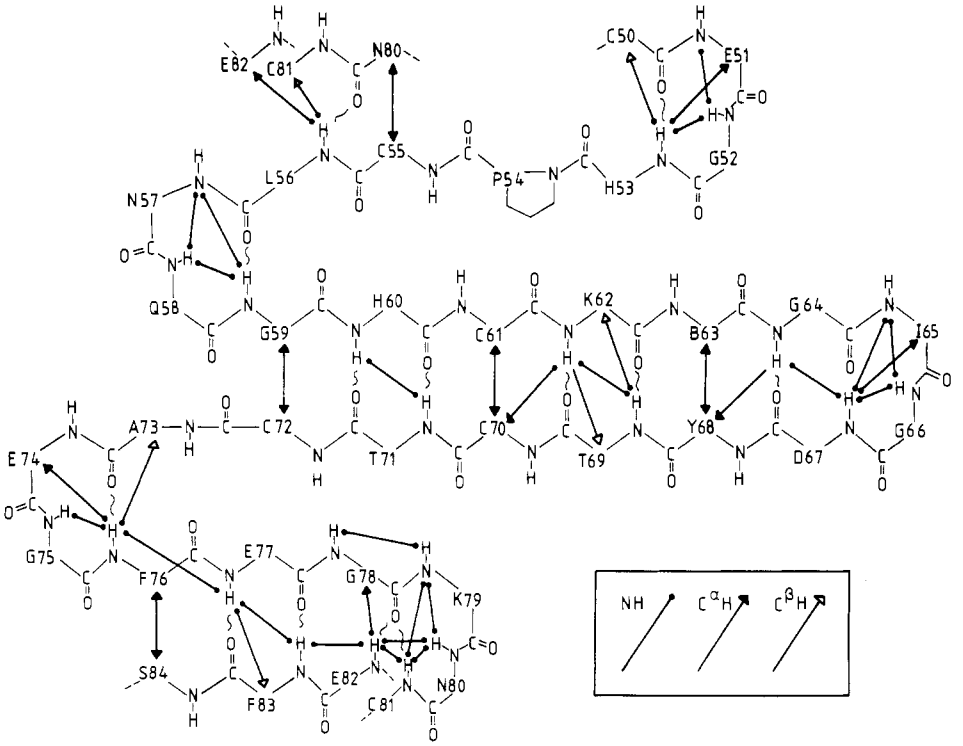

FIGURE 4: Display of the secondary structure of fX-EGF$_N$ including the NOE's on which the analysis is based. NOE's are shown with thick lines, where the markings at the ends of the lines indicate what types of hydrogens are involved (see inset). If NOE's from one proton to both α- and β-protons of another residue are observed, only the α-NOE is included in the picture. Hydrogen bonds, indicated with wavy lines, are postulated when supported by both NOE and exchange data.

Table I: ^{1}H NMR Chemical Shifts (ppm) of Ca^{2+}-Free fX-EGF_N at 301 K, pH 5.8[a]

residue	NH	$C^\alpha H$	$C^\beta H$	$C^\gamma H$	$C^\delta H$	other
K45		4.08	1.97, 1.97	1.39, 1.48	1.74, 1.74	3.04, d ($C^\epsilon H_2$)
D46		4.58	2.56, 2.69			
G47	8.58	3.88, 4.13				
D48	8.10	4.66	2.77, 2.81			
Q49	9.04	4.11	2.16, 2.25	2.54, d		6.76, 7.51 ($N^\epsilon H_2$)
C50	8.71	4.27	3.11, 3.11			
E51	7.75	3.92	1.95, 2.03	2.26, d		
G52	8.02	3.54, 4.00				
H53	7.72	4.37	3.17, 3.33		7.20	8.50 ($C^\epsilon H$)
P54		4.18	1.62, 1.75	0.78, 1.43	3.20, 3.38	
C55	7.59	4.42	2.28, 2.62			
L56	8.39	4.21	1.28, 1.54	1.74		0.79, 0.84 ($C^\delta H_3$)
N57	8.90	3.91	1.06, 1.78			7.16, 7.89 ($N^\epsilon H_2$)
Q58	8.73	3.73	2.27, 2.27	3.11, d		6.79, 7.48 ($N^\epsilon H_2$)
G59	7.16	3.31, 4.27				
H60	8.53	4.92	3.12, 3.19		7.20	8.50 ($C^\epsilon H$)
C61	8.98	5.15	3.02, 3.02			
K62	9.29	4.78	1.79, 1.91	1.82, d	1.54, d	3.08, d ($C^\epsilon H_2$)
β63	8.48	4.48	4.36			
G64	8.29	3.75, 4.52				
I65	8.50	4.13	1.87	1.17, 1.48		0.95 ($C^\gamma H_3$)
						0.89 ($C^\delta H_3$)
G66	8.80	3.84, 3.98				
D67	7.42	4.73	2.46, 2.68			
Y68	7.99	5.16	2.71, 3.10			6.84 ($C^\delta H$)
						6.82 ($C^\epsilon H$)
T69	8.66	4.59	4.05	1.21		
C70	9.51	5.54	2.76, 3.07			
T71	9.36	4.61	4.15	1.34		
C72	9.22	4.88	2.77, 3.42			
A73	8.85	4.43	1.56			
E74	8.44	4.11	1.98, 2.07	2.30, d		
G75	8.65	3.47, 4.01				
F76	7.91	5.73	2.97, 3.13			6.98, ($C^\delta H$)
						7.14, d ($C^\epsilon H$, $C^\zeta H$)
E77	9.45	4.95	2.10, 2.14	2.19, d		
G78	8.23	4.00, 5.10				
K79	9.38	3.98	1.79, 1.99	1.37, d	1.63, 1.63	2.66, 2.77 ($C^\epsilon H_2$)
N80	9.19	5.58	2.29, 3.47			6.96, 7.52 ($N^\epsilon H_2$)
C81	7.83	4.02	3.18, 3.47			
E82	10.63	3.96	1.56, 1.73	1.10, d		
F83	8.89	5.24	2.82, 3.24			7.26 ($C^\delta H$)
						7.45 ($C^\epsilon H$)
						7.53 ($C^\zeta H$)
S84	8.75	4.85	3.85, d			
T85	8.25	4.24	4.21	1.02		
R86	7.74	4.16	1.66, 1.82	1.55, 1.55	3.14, d	7.10 ($N^\epsilon H$)

[a] Chemical shifts are referenced to the H_2O signal at 4.75 ppm and are accurate to ±0.02 ppm. Where degeneracy of shifts is confirmed from $2Q^1$ experiments, the shifts are written separately, whereas cases where degeneracy is assumed are indicated with d. Hya^1 is marked β.

$C^\alpha H$ NOE's is included in Figure 2. The proline residue 54 is found to be trans, as NOE connectivities are found between His 53–$C^\alpha H$ and Pro 54–$C^\delta H$ but not between the $C^\alpha H$'s of these residues (Wüthrich et al., 1984). Furthermore, NOE's are found between β-protons in the bridge-forming cysteines, indicating that the disulfide bonds are intact.

Discussion

Secondary Structure of fX-EGF_N. The main feature of the secondary structure of fX-EGF_N is a six-residue-long antiparallel β-sheet connecting residues 59–64 and 67–72. Most of the interstrand NOE's characteristic for an antiparallel β-sheet are observed and the hydrogens involved in hydrogen bonding exchange slowly with the solvent. For the residues in the sheet, the sequential NH–$C^\alpha H$ connectivities are much stronger than the intra NH–$C^\alpha H$ connectivities. In addition, only very weak, if any, sequential NH–NH NOE's are found for these residues and most of the backbone NH/$C^\alpha H$ coupling constants are significantly larger than 8 Hz. All this indicates that the residues adapt an extended conformation as is expected for a β-sheet (Wagner et al., 1986; Wüthrich, 1986). Another antiparallel β-sheet is found in the C-terminal end of the peptide involving residues 76–77 and 83–84. As for the β-sheet involving residues 59–64 and 67–72, its existence is confirmed by NOE's and exchange data.

Several NOE's connect residues 55–56 and 80–82 in such a way that a parallel arrangement of the peptide chains is well defined. As the amide proton of residue 56 is subject to slow exchange, it is plausible that there is a hydrogen bond connecting these two strands. The structure would then be characterized as a very small parallel β-sheet. This sheet has a kink due to the S–S bridge connecting cysteines 72 and 81.

Four typical β-turns are identified in fX-EGF_N. The turns at residues 50–53, 56–59, and 73–76 all satisfy the NOE criteria for β-turns (Wüthrich et al., 1984) and have slowly exchanging amide protons to confirm the hydrogen bond. These turns are of type I, I, and II, respectively, as judged by the presence or absence of NOE connectivities between the amide protons of residues 2 and 3 in the turn (Wagner et al., 1986). The β-turn at residues 64–67, type I as judged by the NOE's, has no slowly exchanging amide proton, which may be due to solvent exposure of this part of the peptide. It should be noted that hydrogen bonding is not a necessary condition for a β-turn (Richardson, 1981).

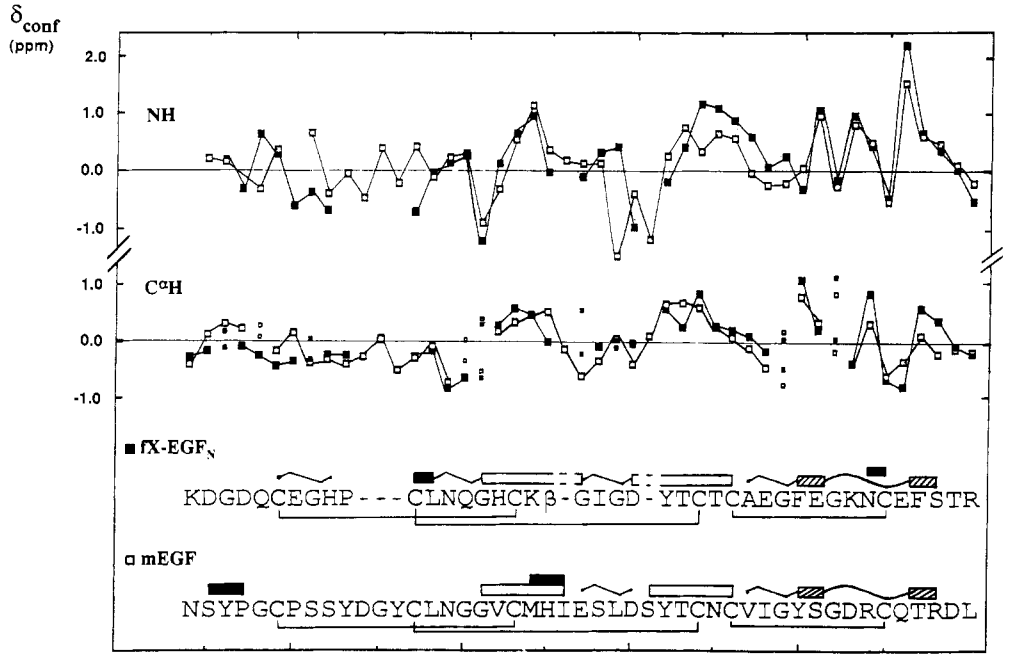

FIGURE 5: Conformational shifts for fX-EGF$_N$ (■) and murine EGF (mEGF) (□) together with the secondary structure of the two peptides represented as in Figure 2. The triple-stranded β-sheet in murine EGF is shown with two boxes on top of each other indicating which strands are joined. For the glycines, both $C^{\alpha}H$ shifts are shown with smaller square markings. The alignment of the sequences is based on the location of the S–S bridges. Gaps have been introduced in the fX-EGF$_N$ sequence for optimal alignment.

Residues 78–81 have the NOE characteristics of a type II β-turn. However, the peptide chain has a kink at residue 81 due to the S–S bridge, which makes it possible for the amide protons of both residues 81 and 82 to form hydrogen bonds to the carbonyl oxygen of residue 78, as is confirmed by the NH exchange data. Therefore, the turn involving residues 78–82 could be characterized as an extended bend. The kink at residue 81 also makes it possible for residues 76–77 and 83–84 to form the antiparallel β-sheet as described above.

The three N-terminal residues Lys 45–Gly 47 give weak crosspeaks for the sequential NOE connectivities; the line width of the amide proton resonance of Gly 47 is very narrow compared to those of the other amide resonances in the peptide, and the amide proton of Asp 46 exchanges so fast with the solvent that its resonance is saturated by the decoupling of the water resonance. All this is in agreement with the amino terminus being mobile as compared to the rest of the peptide. Furthermore, for residues Lys 45–Gln 49, no $J_{NH-\alpha}$ coupling constants are larger than 8 Hz or smaller than 6 Hz and there are no slowly exchanging amide protons, no exceptionally large sequential or internal NH-α NOE's, and no backbone NOE's to the rest of the peptide. This indicates that the first five N-terminal amino acid residues of fX-EGF$_N$ do not take part in any well-defined secondary structure.

Indications on Tertiary Fold. From the secondary structure shown in Figure 4, the information on sequential NOE's given in Figure 2, and the location of the cysteine bridges, a model of fX-EGF$_N$ is easily constructed. The protein fragment is found to have two domains, one (residues Ala 73–Arg 86) where the main feature is the C-terminal β-sheet and one that consists of the major antiparallel β-sheet and the 14 N-terminal amino acids (Lys 45–Cys 72). The orientation of the C-terminal domain relative to the large β-sheet is defined from the direction of the β-sheet between residues 55–56 and 80–82. Residues Lys 45–Pro 54 are not structurally confined to the rest of fX-EGF$_N$ by any NOE's involving the backbone protons. As Cys 61 and Cys 70 are located opposite to each other in the main antiparallel β-sheet, their side chains protrude on the same side of the β-sheet. Therefore, residues 51–58, constrained by the S–S bridges 51–61 and 55–70, are then fixed on this side of the sheet. A detailed discussion of the tertiary structure has to await a structure determination based on the full set of observed NOE's and refined by means of restrained molecular dynamics simulation, which is now in progress in our laboratory.

Comparison with Other EGF Homologous Structures. It is interesting to compare the assignment and secondary structure of fX-EGF$_N$ presented in this work with those of other EGF homologous structures. Here, we will compare the secondary structure fX-EGF$_N$ with that of murine EGF (Montelione et al., 1986, 1988), which is very similar to that of human EGF (Cooke et al., 1987) and rat EGF (Mayo et al., 1989) and has therefore been chosen to represent the EGF structures. As the secondary structure of TGFα has been determined with different results by different groups (Brown et al., 1989; Kohda et al., 1989; Montelione et al., 1989; Tappin et al., 1989), no comparison with TGFα will be made.

The secondary structures of fX-EGF$_N$ and murine EGF are remarkably similar, even though, after having aligned the two peptides according to the location of the disulfide bonds, only eight residues except for the cysteines are identical (Figure 5). A common structure in the two peptides is the antiparallel β-sheet corresponding to residues 59–64 and 67–72 in fX-EGF$_N$. The structure of the C-terminal domain also seems to be conserved in the two peptides with its β-turn, its short β-sheet, and the extended bend.

In fX-EGF$_N$, as opposed to murine and human EGF, there is no evidence of a triple-stranded β-sheet connecting the three N-terminal residues to the rest of the structure. Instead, the N-terminus seems to be a mobile part of the peptide. This may be a genuine structural difference with implications on Ca^{2+} binding, which will be discussed below. It should be emphasized that in fX-EGF$_N$ there are two aspartic acid residues in the amino terminus that are probably negatively charged at pH 5.8. As no negatively charged residues are present in the amino terminus of murine EGF, the difference in secondary structure may be due to the difference in charge distribution.

The orientation of the C-terminal domain relative to the N-terminal domain in EGF is a matter of discussion in the litterature. In fX-EGF_N, however, the orientation is defined by the short parallel β-sheet. Montelione et al. (1986) find a slowly exchanging amide proton in the position in murine EGF corresponding to Leu 56 in fX-EGF_N but do not give a structural explanation of this slow exchange (Montelione et al., 1987). In both human EGF (Carver et al., 1986) and murine EGF (Kohda et al., 1988) interactions between residues corresponding to residues 55–56 and 80–82 in fX-EGF_N are found but no β-sheet is identified. Either this small β-sheet is unique for fX-EGF_N or it may also be identified in other EGF like structures at a closer look. It should also be observed that the pHs at which the two structures were determined were different (pH 3 for murine EGF, pH 5.8 for fX-EGF_N). This may very well affect the secondary structure as there are residues with carboxylic side chains in this region in both murine EGF and fX-EGF_N that normally titrate between pH 3 and 5.

By comparing the shifts of the backbone protons of fX-EGF_N and murine EGF (Montelione et al., 1988), structural similarities can be investigated from another point of view. The shifts of the backbone hydrogens are affected by primary, secondary, and tertiary structures (Wüthrich, 1986). To be able to focus on conformational shift contributions and remove most of the effects of the different primary structures in fX-EGF_N and murine EGF, conformation-dependent shifts were calculated by subtraction of the corresponding random-coil chemical shifts from the observed shifts (Pardi et al., 1983). The conformation-dependent NH and $C^{\alpha}H$ shifts in fX-EGF_N and murine EGF are compared in Figure 5 together with the secondary structure of the two peptides. In the C-terminal part, both NH and $C^{\alpha}H$ shifts show a very similar pattern in the two peptides. The shift pattern for the main β-sheet also displays the same features: the C^{α} protons generally show positive conformation-dependent shifts in the sheet. This is expected in an antiparallel β-sheet as the C^{α}-protons get close to the carbonyl oxygens (Wagner et al., 1983). Residues in regular β-sheets are shifted by on average +0.4 ppm with respect to the random-coil value (Szilagyi & Jardetzky, 1989), which agrees well with the size of the conformational shifts in the β-sheets of fX-EGF_N and murine EGF. In the N-terminal part, however, the shift pattern differs. The positive $C^{\alpha}H$ conformation-dependent shifts in murine EGF corresponding to the triple-stranded β-sheet in this peptide have no counterpart in fX-EGF_N. This is consistent with the fact that no such β-sheet is observed in fX-EGF_N. Similarly, the shift pattern in the region of the first β-turn in fX-EGF_N, a turn that is not found in murine EGF, is different in the two peptides. Differences are also obvious in the region of the β-turn in the main antiparallel β-sheet, but this is probably due to the gaps that have been introduced in the fX-EGF_N sequence for optimal alignment as murine EGF has a larger number of amino acids than fX-EGF_N. An interesting feature is that in the region of the postulated small parallel β-sheet in fX-EGF_N the shift pattern is very similar in the two peptides. As other regions with the same secondary structure in the two peptides have similar conformation-dependent shifts, this indicates that the short parallel β-sheet orienting the two domains in fX-EGF_N relative to each other may also exist in murine EGF.

Structural Requirements for Hydroxylation of Asp 63. The consensus sequence for EGF like domains that have Hya or β-hydroxyasparagine (Hyn),[1] Cys-X-Hya/Hyn-X-X-X-X-Tyr/Phe-X-Cys-X-Cys (Stenflo et al., 1987), seems to be an absolute but not a sufficient requirement for hydroxylation. For example, human factor VII has the sequence but is not hydroxylated (Thim et al., 1988). In factor X, our present results show that the four residues between Hya 63 and Tyr 68 form a β-turn, which reverses the chain so that all residues in the consensus sequence become part of an antiparallel β-sheet, where Hya 63 and Tyr 68 lie opposite to each other. As linear synthetic peptides that have the consensus sequence are not hydroxylated (Stenflo et al., 1989), it is probable that these peptides do not adopt the secondary structure required for hydroxylation. However, the fact that factor VII is not hydroxylated implies that there may be further tertiary structural requirements for aspartyl β-hydroxylase in order to hydroxylate Asp/Asn.

Ca^{2+} Binding of fX-EGF_N. The secondary structure of fX-EGF_N allows Asp 46, Asp 48, and Hya 63 to get close enough in space to coordinate a Ca^{2+} ion. This location of the Ca^{2+}-binding site is supported by the fact that the shifts of the aromatic ring protons in Tyr 68 (Figure 3) are affected by Ca^{2+} binding (Persson et al., 1989). However, no triple-stranded β-sheet is found in the Ca^{2+}-free form of fX-EGF_N, a structural element found in murine EGF and human EGF. By analogy, this structure was proposed to form a Ca^{2+}-binding site in factor IX, as the three conserved acidic residues corresponding to Asp 46, Asp 48, and Hya 63 in factor X would then be on the same side of this sheet (Cooke et al., 1987). In fX-EGF_N, the four amino-terminal residues of fX-EGF_N are mobile as compared to the rest of the protein fragment. It remains to be seen if a triple-stranded β-sheet is stabilized in the Ca^{2+}-loaded form of fX-EGF_N.

ACKNOWLEDGMENTS

We thank Sara Linse and Ann-Kristin Öhlin for helpful discussions and continuous encouragement. We also extend special thanks to Mats Lundell for drawing Figures 2 and 4.

Registry No. EGF, 62229-50-9; factor X, 9001-29-0.

REFERENCES

Appella, E., Weber, I. T., & Blasi, F. (1988) *FEBS Lett. 231*, 1–4.

Bax, A. (1985) *J. Magn. Reson. 65*, 142–145.

Bax, A., & Davis, P. (1985) *J. Am. Chem. Soc. 107*, 2820.

Billeter, M., Braun, W., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 321–346.

Braunschweiler, L., Bodenhausen, G., & Ernst, R. R. (1983) *Mol. Phys. 48*, 535–560.

Brown, S. C., Mueller, L., & Jeffs, P. W. (1989) *Biochemistry 28*, 593–599.

Carver, J. A., Cooke, R. M., Esposito, G., Campbell, I. D., Gregory, H., & Sheard, B. (1986) *FEBS Lett. 205*, 77–81.

Chazin, W. J., Rance, M., & Wright, P. E. (1988) *J. Mol. Biol. 202*, 603–622.

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

Drakenberg, T., Fernlund, P., Roepstorff, P., & Stenflo, J. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 1802–1806.

Englander, S. W., Downer, N. W., & Teitelbaum, H. (1972) *Annu. Rev. Biochem. 41*, 903–1024.

Esmon, N. L., DeBault, L. E., Esmon, C. T. (1983) *J. Biol. Chem. 258*, 5548–5553.

Furie, B., & Furie, B. C. (1988) *Cell 53*, 505–518.

Handford, P. A., Baron, M., Mayhew, M., Willis, A., Beesley, T., Brownlee, G. G. & Campbell, I. D. (1990) *EMBO J. 9*, 475–480.

Höjrup, P., and Magnusson, S. (1987) *Biochem. J. 245*, 887–892.
Huang, L. H., Ke, X.-H., Sweeney, W., & Tam, J. P. (1989) *Biochem. Biophys. Res. Commun. 160*, 133–139.
Kohda, D., Go, N., Hayashi, K., & Inagaki, F. (1988) *J. Biochem. 103*, 741–743.
Kohda, D., Shimada, I., Miyake, T., Fuwa, T., & Inagaki, F. (1989) *Biochemistry 28*, 953–958.
Levitt, M., & Freeman, R. (1979) *J. Magn. Reson. 33*, 473–476.
Mann, K. G., Jenny, R. J., & Krishnaswamy, S. (1988) *Annu. Rev. Biochem. 57*, 915–956.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.
Mayo, K. H., Cavalli, R. C., Peters, A. R., Boelens, R., & Kaptein, R. (1989) *Biochem. J. 257*, 197–205.
McMullen, B. A., Fujikawa, K., Kisiel, W., Sasagawa, T., Howald, W. N., Kwa, E. Y., & Weinstein, B. (1983) *Biochemistry 22*, 2875–2884.
Montelione, G. T., Wüthrich, K., Nice, E. C., Burgess, A. W., & Scheraga, H. A. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 8594–8598.
Montelione, G., Wüthrich, K., Nice, E. D., Burgess, A. W., & Sheraga, H. A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 5226–5230.
Montelione, G. T., Wüthrich, K., & Scheraga, H. A. (1988) *Biochemistry 27*, 2235–2243.
Montelione, G. T., Winkler, M. E., Burton, L. E., Rinderknecht, E., Sporn, M. B., & Wagner, G. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 1519–1523.
Öhlin, A.-K., & Stenflo, J. (1987) *J. Biol. Chem. 262*, 13798–13804.
Öhlin, A.-K., Linse, S., & Stenflo, J. (1988) *J. Biol. Chem. 263*, 7411–7417.
Pardi, A., Wagner, G., & Wüthrich, K. (1983) *Eur. J. Biochem. 137*, 445–454.
Patthy, L. (1985) *Cell 41*, 657–663.
Persson, E., Selander, M., Linse, S., Drakenberg, T., Öhlin, A.-K., & Stenflo, J. (1989) *J. Biol. Chem. 264*, 16897–16904.
Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 131*, 1094–1102.
Rees, D. J. G., Jones, I. M., Handford, P. A., Walter, S. J., Esnouf, M. P., Smith, K. J., & Brownlee, G. G. (1988) *EMBO J. 7*, 2053–2061.
Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167–339.
Stenflo, J., & Suttie, J. W. (1977) *Annu. Rev. Biochem. 46*, 157–172.
Stenflo, J., Lundvall, Å., & Dahlbäck, B. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 368–372.
Stenflo, J., Holme, E., Lindstedt, S., Chandramouli, N., Tsai Huang, L. H., Tam, J. P., & Merrifield, R. B. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 444–447.
Sugo, T., Björk, I., Holmgren, A., & Stenflo, J. (1984) *J. Biol. Chem. 259*, 5705–5710.
Suttie, J. W. (1985) *Annu. Rev. Biochem. 54*, 459–477.
Szilagyi, L., & Jardetzky, O. (1989) *J. Magn. Reson. 83*, 441–449.
Tappin, M. J., Cooke, R. M., Fitton, J. E., & Campbell, I. d. (1989) *Eur. J. Biochem. 179*, 629–637.
Thim, L., Bjoern, S., Christensen, M., Nicolaisen, E., Lund-Hansen, T., Pedersen, A. H., & Hedner, V. (1988) *Biochemistry 27*, 7785–7793.
Wagner, G. (1983) *J. Magn. Reson. 55*, 151–156.
Wagner, G., Pardi, A., & Wüthrich, K. (1983) *J. Am. Chem. Soc. 105*, 5948–5949.
Wagner, G., Neuhaus, D., Wörgötter, E., Vasak, M., Kagi, J. R. H., & Wüthrich, K. (1986) *J. Mol. Biol. 187*, 131–135.
Wüthrich, K., (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Wüthrich, K., Billeter, M., & Braun, W. (1984) *J. Mol. Biol. 180*, 715–740.