intermediate" by NMR measurements in the presence of enzyme is highly implausible.

Barlow et al. (1989) also report a nonspecific phosphatase, which must be a contaminant in their preparations because 800 μM S3P and 1 mM glyphosate only inhibit the rate by 80%; these concentrations are 40 times higher than should be required to see 98% inhibition of EPSP synthase (Anderson et al., 1988a).

Other side reactions were also shown to occur, involving the hydrolysis of EPSP at the active site to produce pyruvate and breakdown of the intermediate to form the ketal. These reactions are also slow; the rate of EPSP hydrolysis in the absence of phosphate is 4.7×10^{-4} s^{-1} (Anderson et al., 1988b); with 5% of the sites occupied by EPSP under internal equilibrium conditions, we predict a rate of pyruvate formation of 2.4×10^{-5} s^{-1}, which is consistent with the observed rate. These slow side reactions add little new mechanistic information to illuminate the events at the active site. For example, the hydrolysis of EPSP is probably a function of the protonation of the substrate at the active site followed by attack by water rather than phosphate. The formation of the EPSP ketal might have implications for the geometry of the intermediate bound to the active site, but the slowness of the reaction and uncertainty of the pathway of formation of the ketal could lead one to question the validity of such conclusions. Rate measurements are necessary to extract mechanistic conclusions from structural data, and, certainly, the converse is equally true.

Current data provide overwhelming support to conclude that the EPSP synthase reaction proceeds by a simple addition–elimination mechanism via the formation of a tetrahedral intermediate, as originally proposed in the pioneering work from Sprinson's laboratory (Bondinell et al., 1971).

References

Anderson, K. S., & Johnson, K. A. (1990) *J. Biol. Chem.* (in press).

Anderson, K. S., Sikorski, J. A., & Johnson, K. A. (1988a) *Biochemistry 27*, 1604–1610.

Anderson, K. S., Sikorski, J. A., & Johnson, K. A. (1988b) *Biochemistry 27*, 7395–7406.

Anderson, K. S., Sikorski, J. A., Benesi, A. J., & Johnson, K. A. (1988c) *J. Am. Chem. Soc. 110*, 6577–6579.

Anton, D. L., Hedstrom, L., Fish, S. M., & Abeles, R. H. (1983) *Biochemistry 22*, 5903–5908.

Barlow, P. N., Appleyard, R. J., Wilson, B. J. O., & Evans, N. S. (1989) *Biochemistry 28*, 7985–7991.

Bondinell, W. E., Vnek, J., Knowles, P. F., Sprecher, M., & Sprinson, D. B. (1971) *J. Biol. Chem. 246*, 6191–6196.

Franz, J. E. (1985) in *The Herbicide Glyphosate* (Grossbard, E., & Atkinson, D., Eds.) pp 1–17, Butterworth, Boston, MA.

Leo, G. C., Sikorski, J. A., & Sammons, R. D. (1990) *J. Am. Chem. Soc.* (in press).

Millar, G., Lewendon, A., Hunter, M. G., & Coggins, J. R. (1986) *Biochem. J. 237*, 427–437.

Rogers, S. G., Brand, L. A., Holder, F. B., Sharps, E. S., & Brackin, M. J. (1983) *Appl. Environ. Microb. 46*, 37–43.

Steinrucken, H. C., & Amrhein, N. (1980) *Biochem. Biophys. Res. Commun. 94*, 1207–1212.

Sequence-Specific ^{1}H NMR Assignments and Secondary Structure of Eglin c†

Sven G. Hyberts and Gerhard Wagner*

Biophysics Research Division, Institute of Science and Technology, University of Michigan, 2200 Bonisteel Boulevard, Ann Arbor, Michigan 48109

Received June 26, 1989; Revised Manuscript Received October 16, 1989

ABSTRACT: Sequence-specific nuclear magnetic resonance assignments were obtained for eglin c, a polypeptide inhibitor of the granulocytic proteinases elastase and cathepsin G and some other proteinases. The protein consists of a single polypeptide chain of 70 residues. All proton resonances were assigned except for some labile protons of arginine side chains. The patterns of nuclear Overhauser enhancements and coupling constants and the observation of slow hydrogen exchange were used to characterize the secondary structure of the protein. The results indicate that the solution structure of the free inhibitor is very similar to the crystal structure reported for the same protein in the complex with subtilisin Carlsberg. However, a part of the binding loop seems to have a significantly different conformation in the free protein.

Eglin c is a proteinase inhibitor from the leech *Hirudo medicinalis* (Seemüller et al., 1977). It inhibits the granulocytic proteinases elastase and cathepsin G and some other proteinases such as chymotrypsin and subtilisin. Its inhibitory properties make the protein an interesting molecule for pharmaceutical application as antiinflammatory agent. Eglin c is a member of the potato inhibitor 1 family. It consists of a single polypeptide chain with 70 residues. When eglin c was first detected, it was isolated together with hirudin and eglin b, which were identified as protein fractions a and b, respectively, while eglin c was fraction c. Eglins b and c differ by a single amino acid exchange at position 35 (histidine in eglin b, tyrosine in eglin c). Although eglin is lacking disulfide bonds, it is very stable against denaturation by heat (Seemüller et al., 1977, 1980). The crystal structure of the protein has been determined in the complex with subtilisin Carlsberg (McPhalen et al., 1985, 1987; Bode et al., 1986, 1987). The structure of the homologous serine proteinase inhibitor CI-2 from barley seeds has been determined in the complex with proteinases by X-ray crystallography (McPhalen & James,

† This work was supported by the Schweizerischer Nationalfonds (Project 3.197-0.85), by the National Science Foundation (Grant DMB-8616059), and by the National Institutes of Health (Grant 1GM38608).

* Address correspondence to this author.

0006-2960/90/0429-1465$02.50/0

1988) and in the free form by X-ray crystallography (McPhalen & James, 1987) and by NMR[1] spectroscopy (Clore et al., 1987a,b). In this paper we present sequence-specific assignments of the ^1H NMR spectrum of eglin c which were obtained with standard procedures described earlier (Dubs et al., 1979; Wagner et al., 1981, 1982; Wüthrich et al., 1982). In addition, we have obtained stereospecific assignments of a part of the β-methylene protons and the valine γ-methyl protons with methods developed in the work with this protein (Hyberts et al., 1987). For the assignments of the proton spectrum, in particular the aromatic resonances, we have also employed heteronuclear ^1H–^{13}C correlation spectroscopy as described previously (Brühwiler & Wagner, 1986; Wagner & Brühwiler, 1986). Furthermore, we have examined the patterns of nuclear Overhauser enhancements (NOE's) and spin–spin coupling constants, combined with analysis of hydrogen–deuterium exchange rates, to elucidate the secondary structure of the protein.

The assignments presented here are the basis for a structure determination in solution of this protein and several mutants which were designed to alter the structure of the proteinase binding loop. Part of these mutants that we are investigating have lost their inhibitory function, and in one case, the NMR data provide an indication for the structural basis of the loss of inhibitory function (Wagner et al., 1989).

Materials and Methods

The protein eglin c was produced by Ciba-Geigy, Basel, Switzerland, from a synthetic gene cloned and expressed in high yield in *Escherichia coli*. Structural analysis showed that the expressed eglin c is modified to N^{α}-acetyleglin c in a posttranslational event (Rink et al., 1984; Märki et al., 1985; Richter et al., 1985). For the NMR measurements 0.005 M solutions in H_2O or ^2H_2O were used. For the measurements in ^2H_2O all labile protons were exchanged with deuterons prior to the measurements. All experiments were carried out at pH 3.0. The NMR spectra were measured on a Bruker WM-500 and a General Electric GN-500 spectrometer. Double quantum filtered COSY (DQF-COSY) spectra were recorded as described (Piantini et al., 1982; Shaka & Freeman, 1983; Rance et al., 1984). 2D RELAY (Eich et al., 1982) and double RELAY spectra were recorded as described previously (Wagner, 1983). The 2D total correlation spectra (TOCSY; Braunschweiler & Ernst, 1983; Müller & Ernst, 1979) were recorded in H_2O and ^2H_2O. The isotropic mixing was performed with an MLEV-17 sequence as described (Bax & Davis, 1985; Davis & Bax, 1985). Heteronuclear ^1H–^{13}C COSY spectra were recorded with ^1H detection as described (Brühwiler & Wagner, 1986; Wagner & Brühwiler, 1986) following previously developed methods (Müller, 1979; Vidusek et al., 1982; Bax et al., 1983a,b; Bendall et al., 1983; Frey et al., 1985). NOESY spectra (Jeener et al., 1979; Kumar et al., 1980) were recorded with variation of the mixing time as outlined (Macura et al., 1981; Rance et al., 1985) to minimize zero-quantum contributions in NOESY cross peaks of coupled protons. NOESY spectra with mixing times of 50, 75, 100, 150, and 200 ms were recorded. All 2D NMR spectra were recorded in a way suitable for phase-sensitive Fourier transforms. The sign discrimination along ω_1 was achieved by time-proportional phase incrementation (TPPI) as described by Marion and Wüthrich (1983) or by quadrature detection along ω_1 (States et al., 1982).

[1] Abbreviations: NMR, nuclear magnetic resonance; 1D, one dimensional; 2D, two dimensional; ppm, parts per million; COSY, 2D correlated spectroscopy; DQF-COSY, double quantum filtered COSY; NOE, nuclear Overhauser effect; NOESY, 2D NOE spectroscopy.

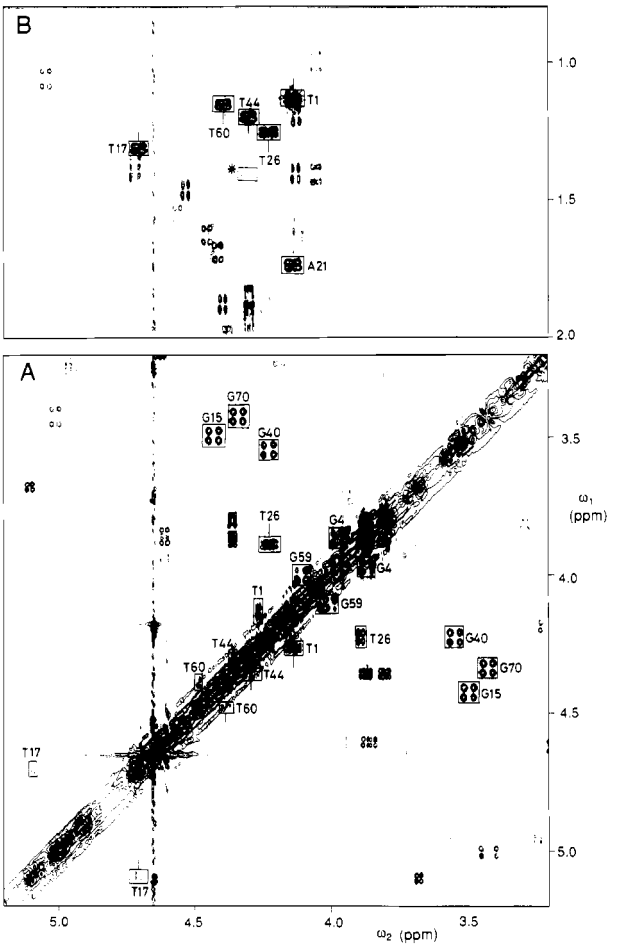

FIGURE 1: Assignments of glycine, threonine, and alanine spin systems in a DQF-COSY spectrum of a 5 mM solution of eglin c in ^2H_2O at 36 °C, p^{2}H 3.0. (A) Glycine α–α cross peaks and threonine α–β cross peaks. (B) Threonine β–γ cross peaks and an alanine α–β cross peak. The asterisk marks an extra Ala/Thr spin system of low intensity, which indicates the presence of a chemical or conformational heterogeneity.

Results and Discussion

The assignment of the proton resonances of eglin c was obtained with the strategies first described by Dubs et al. (1978), Nagayama and Wüthrich (1981), Wagner et al. (1981, 1982), Wüthrich et al. (1982), and Wider et al. (1982). Following these procedures, (i) we describe the assignment to residue type of the spin systems of the nonlabile protons as far as they involve α-protons and connectivities to labile protons of amino acid side chains; (ii) we present the identification of the aromatic side-chain resonances; (iii) we describe how we have obtained the connectivities between α-protons and backbone amide protons; (iv) we document the sequence-specific assignments; (v) we show spectral parameters that characterize the secondary and tertiary structure of eglin c. The chemical shifts for all spin systems assigned are listed in Table I. In the figures the peaks are identified with the sequence positions as obtained after the final sequential assignment procedure.

Assignments of Nonlabile Side-Chain Spin Systems. Figure 1 shows regions of a double quantum filtered COSY spectrum of eglin c recorded in ^2H_2O containing the cross peaks of the glycine, threonine, and alanine spin systems. Figure 1A contains all α–β cross peaks of the five glycines and the α–β cross peaks of the five threonines. The β–γ cross peaks of the threonines and the α–β cross peaks of the threonines and the

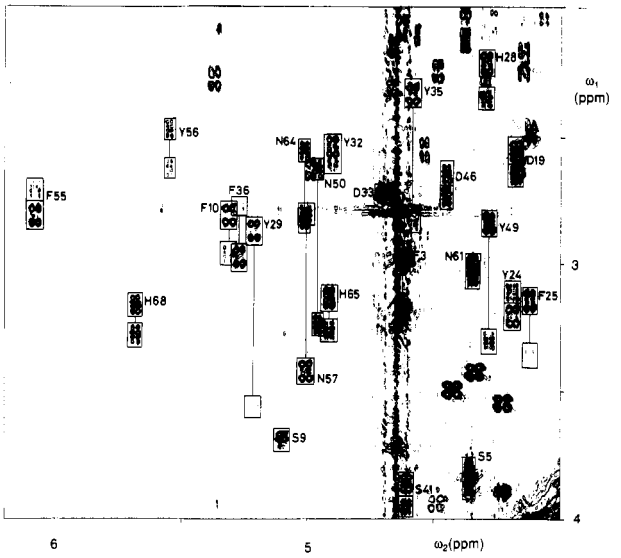

FIGURE 2: Assignments of the AMX spin systems in another expansion of the same DQF-COSY spectrum as shown in Figure 1.

α–β cross peaks of the single alanine (Ala 21) are documented in Figure 1B. The assignments of all the threonine spin systems were confirmed with RELAY and TOCSY experiments where cross peaks between α-protons and γ-protons were observed. The peak labeled with an asterisk in Figure 1B has the same multiplet structure as an alanine or a threonine cross peak, but it has much lower intensity. Since eglin c contains only five threonines and one alanine, this peak indicates a chemical or conformational heterogeneity. Probably, there is a minor fraction of molecules where the N-terminus is not acetylated, and the extra peak belongs to Thr 1 of this protein fraction.

Figure 2 is another expansion of the same DQF-COSY spectrum as shown in Figure 1 containing all α–β cross peaks of the AMX spin systems. Eglin c has six tyrosines, five phenylalanines, three histidines, three serines, four asparagines, and three aspartic acids. The identification of the serine spin systems was obtained on the basis of the small β–β coupling constant (Nagayama & Wüthrich, 1981) and on the basis of the low-field shift of the β-protons. This identification of serine spin systems was confirmed at a later stage in the process of sequential assignments. For Ser 9 the β–β coupling constant could not be measured because of the near degeneracy of the two β-protons. Therefore, for the assignment of this serine as such we had to rely only on chemical shifts and finally on the sequential assignments. The identification of tyrosine, phenylalanine, and histidine spin systems was achieved by observation of NOE's between the β-protons and the aromatic δ-protons (Wagner & Wüthrich, 1982; Wider et al., 1982) as described below. Similarly, the identification of the asparagine spin systems was obtained by observation of NOESY cross peaks between the N^{δ} protons and the β-protons in a NOESY spectrum in H_2O. The β-protons of most AMX spin systems are nondegenerate. Only Asp 33 has completely degenerate β-protons. This was found from a double quantum spectrum in ^2H_2O, which shows a cross peak at the ω_2-position of the α-proton (4.70 ppm) and at the ω_1-position of the sum of the β-proton positions (Wagner & Zuiderweg, 1983).

Valine and leucine spin systems are analyzed in Figure 3. Eglin c contains 11 valine spin systems. The α–β cross peaks and the β–γ cross peaks are identified in panels A and B of Figure 3, respectively. For Val 52 the two γ-methyl resonances are almost entirely degenerate. This becomes evident when one inspects the line shape of the cross peak in Figure 3B, which is slightly broader than for the other nondegenerate cross peaks. All connectivities between the γ-methyl groups and the α-protons were obtained from the COSY spectrum of Figure 3, from a RELAY spectrum, and from α–γ cross peaks in a TOCSY spectrum in ^2H_2O.

Eglin c contains five leucine spin systems. The five pairs of cross peaks between the δ-methyl groups and the γ-methine protons can readily be recognized in Figure 3B. The connectivities to the α- and β-protons were obtained by RELAY and TOCSY experiments. Figure 4 shows the region of a

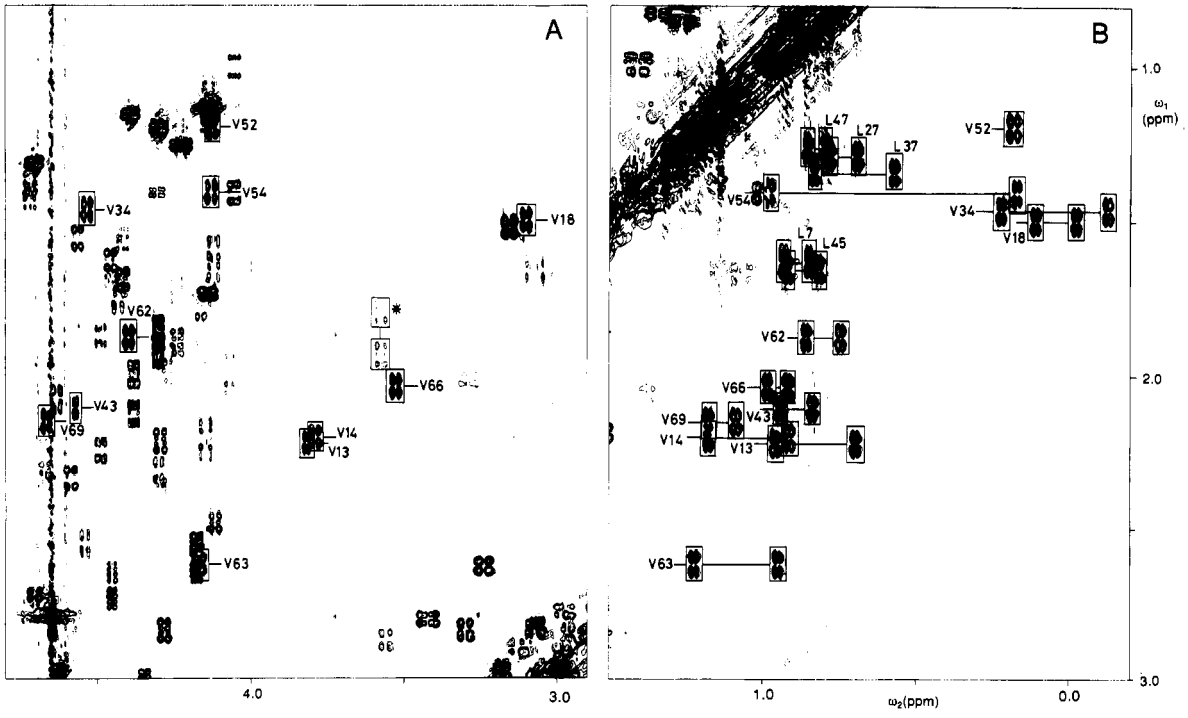

FIGURE 3: Assignment of valine and leucine spin systems in other expansions of the same DQF-COSY spectrum as shown in Figure 1. (A) The 11 valine α–β cross peaks and (B) the valine β–γ and the leucine γ–δ cross peaks.

Table I: Chemical Shifts of the Assigned Proton NMR Resonances of Eglin c at pH 3.0, 36 °C[a]

	H^N	H^α	H^β	H^γ	others		H^N	H^α	H^β	H^γ	others
Thr 1	8.10	4.27	4.14	1.14	*N*-Ac, 2.09	Tyr 35	8.23	4.59	2.34, 2.84		H^δ, 6.93
Glu 2	8.32	4.30	*1.93*, *1.86*	2.18, 2.29							H^ϵ, 6.64
Phe 3	8.17	4.62	*3.19*, *2.99*		H^δ, 7.24	Phe 36	8.84	5.27	*2.78*, *2.98*		H^δ, 7.05
					H^ϵ, 7.31						H^ϵ, 7.15
					H^ζ, 7.27						H^ζ, 7.10
Gly 4	8.22	3.87, 3.98				Leu 37	9.25	5.08	1.15, 1.15	1.35	H^δ, 0.58, 0.82
Ser 5	8.10	4.36	*3.88*, *3.81*			Pro 38		3.58	1.81, 1.94	1.98, 2.11	H^δ, 3.68, 3.92
Glu 6	8.27	4.38	2.00, 2.13	2.40, 2.45		Glu 39	8.82	4.07	1.95, 2.02	2.33, 2.47	
Leu 7	8.07	4.42	1.57, 1.70	1.64	H^δ, 0.85, 0.93	Gly 40	8.67	3.55, 4.23			
Lys 8	7.89	4.12	1.60, 1.68	1.08, 1.18	H^δ, 1.68	Ser 41	7.72	4.61	3.86, 3.96		
					H^ϵ, 2.83, 3.07	Pro 42		4.49	1.88, 2.25	2.02, 2.10	H^δ, 3.92, 3.72
					H^ζ, 7.50	Val 43	8.07	4.57	2.11	*0.84*, *0.95*	
Ser 9	7.89	5.10	3.69, 3.71			Thr 44	7.66	4.36	4.30	1.21	
Phe 10	7.66	5.31	2.82, 2.97		H^δ, 6.98	Leu 45	8.60	4.46	1.63, 1.71	1.66	H^δ, 0.82, 0.92
					H^ϵ, 6.35	Asp 46	8.16	4.45	2.65, 2.74		
					H^ζ, 6.36	Leu 47	8.13	4.72	1.38, 1.45	1.30	H^δ, 0.80, 0.86
Pro 11		4.30	2.10, 2.32	2.02	H^δ, 3.29, 3.82	Arg 48	8.82	4.56	1.56, 1.82	1.53	H^δ, 3.15
Glu 12	9.53	4.16	*1.94*, *1.78*	2.05, 2.35							$N^\epsilon H$, 7.30
Val 13	7.09	3.82	2.23	*0.96*, *0.70*		Tyr 49	8.46	4.28	*3.31*, *2.85*		H^δ, 7.15
Val 14	7.18	3.52	2.04	*0.92*, *0.99*							H^ϵ, 6.83
Gly 15	9.10	3.50, 4.43				Asn 50	7.94	4.96	2.63, 3.25		H^δ, 7.63, 6.89
Lys 16	7.84	4.63	2.07, 2.09	1.36	H^δ, 1.36, 1.48	Arg 51	7.29	5.37	1.45, 2.29	1.79	H^δ, 2.85, 2.93
					H^ϵ, 2.76, 2.88						$N^\epsilon H$, 9.12
					H^ζ, 7.55						$N^\eta H$, 6.58, 6.64
Thr 17	8.43	5.09	4.70	1.32		Val 52	8.37	4.13	1.20	0.19, 0.19	
Val 18	8.07	3.10	1.50	*−0.02*, *0.11*		Arg 53	8.62	4.43	1.49, 1.76	0.89, 1.21	H^δ, 2.90, 3.14
Asp 19	8.39	4.18	2.56, 2.66								$N^\epsilon H$, 7.32
Gln 20	7.70	4.12	2.49, 2.49	1.99, 2.05	H^ϵ, 7.49, 6.87	Val 54	8.69	4.13	1.41	0.18, 0.98	
Ala 21	9.39	4.14	1.74			Phe 55	9.21	6.08	2.72, 2.83		H^δ, 7.13
Arg 22	9.06	4.24	1.73, 1.90	1.50, 1.59	H^δ, 3.09, 3.19						H^ϵ, 6.94
					H^ϵ, 7.01						H^ζ, 6.94
Glu 23	7.59	4.15	2.22	2.55, 2.19		Tyr 56	8.13	5.55	*2.64*, *2.48*		H^δ, 6.84
Tyr 24	8.11	4.19	*3.11*, *3.21*		H^δ, 6.50						H^ϵ, 6.60
					H^ϵ, 6.41	Asn 57	8.60	5.01	2.81, 3.43		H^δ, 7.62, 7.24
Phe 25	8.62	4.12	3.15, 3.36		H^δ, 7.39	Pro 58		4.67	2.31	2.10	H^δ, 4.12, 4.12
					H^ϵ, 6.18	Gly 59	8.54	4.00, 4.11			
					H^ζ, 6.35	Thr 60	7.28	4.48	4.40	1.16	
Thr 26	8.30	3.89	4.23	1.26		Asn 61	8.89	4.35	3.01, 3.06		H^δ, 7.32, 6.68
Leu 27	7.74	4.06	*1.01*, *1.41*	1.30	H^δ, 0.70, 0.79	Val 62	7.41	4.40	1.88	*0.75*, *0.86*	
His 28	7.80	4.29	*2.36*, *2.22*		H^δ, 6.62	Val 63	8.87	4.17	2.62	*1.23*, *0.95*	
					H^ϵ, 8.64	Asn 64	8.91	5.01	2.56, 2.83		H^δ, 7.55, 6.84
Thr 29	8.38	5.21	*3.57*, *2.88*		H^δ, 6.99	His 65	7.77	4.92	3.14, 3.27		H^δ, 7.26
					H^ϵ, 7.04						H^ϵ, 8.62
Pro 30		4.54	2.05, 2.56	1.85, 1.99	H^δ, 3.16, 3.56	Val 66	9.17	3.79	2.21	0.92, 1.18	
Gln 31	9.53	4.27	1.88, 1.98	1.53, 2.18	H^ϵ, 7.51, 7.10	Pro 67		4.99	1.79, 2.63	2.25, 2.13	H^δ, 3.94, 4.49
Tyr 32	7.35	4.90	2.54, 2.61		H^δ, 6.97	His 68	8.82	5.68	*3.17*, *3.29*		H^δ, 7.41
					H^ϵ, 6.97						H^ϵ, 8.71
Asp 33	9.04	4.70	2.73, 2.73			Val 69	8.53	4.67	2.15	*1.18*, *1.09*	
Val 34	8.10	4.54	1.47	*−0.13*, *0.23*		Gly 70	8.84	3.42, 4.34			

[a] Stereospecific assignments were obtained for those β-methylene protons and for those γ-methyl protons of valines that are italic (Hyberts et al., 1987). In these cases, the numbers listed first and second correspond to $H^{\beta 2}$ and $H^{\beta 3}$, respectively (for β-methylene protons), and to $H^{\gamma 1}$ and $H^{\gamma 2}$, respectively (for γ-methyl protons of valines).

DQF-COSY spectrum in ^2H_2O that contains the α–β cross peaks of all residues with long side chains. Besides the five leucines, eglin c contains four arginines, two lysines, five glutamic acids, two glutamines, and six prolines. The chemical shifts are listed in Table I. The α–β cross peaks of Pro 38 are outside the spectral range displayed in Figure 4. They can be seen in Figure 3A where they are labeled with an asterisk. The γ-protons of the glutamines and the glutamic acids were obtained from a TOCSY spectrum. The other resonances of the long side chains that are listed in Table I were obtained from TOCSY spectra in H_2O and ^2H_2O. In particular, we have assigned the lysine side chains completely. All arginine side chains were assigned out to the $N^\epsilon H$ protons, and for Arg 51 the side-chain $N^\eta H$ guanidinium protons could also be assigned, showing strong NOE's to the $N^\epsilon H$ proton. This was possible because the $N^\epsilon H$ proton of this residue is unusually far shifted to lower field (9.12 ppm; see Table I). Therefore, the resonance is well resolved from the signals of the other labile protons of this side chain, and NOE's are resolved, whereas in most other cases $N^\epsilon H$ protons and $N^\eta H$ guanidinium protons are nearly degenerate and cross peaks merge with the diagonal.

Aromatic Side Chains. The elucidation of the spin systems of the aromatic side chains was achieved with a combination of double quantum spectra (Figure 5A), double quantum filtered COSY, and heteronuclear ^1H–^{13}C COSY spectra. The connectivities between the aromatic spin systems and the respective β-protons were obtained from a NOESY spectrum in ^2H_2O, which is documented in Figure 5B. Five of the six tyrosines have well-separated resonances (Tyr 24, Tyr 29, Tyr 35, Tyr 49, and Tyr 56) so that the cross peaks can readily be recognized in COSY and double quantum spectra. Tyr 32 was more difficult to assign. No cross peak could be found for this sixth tyrosine in the DQ spectrum (Figure 5A) and in a DQF-COSY. It could only be identified with a heteronuclear COSY spectrum. At the ω_1-position of the C^δ carbon resonances there are cross peaks for six tyrosines. The sixth tyrosine at a proton chemical shift of 6.97 ppm must be Tyr

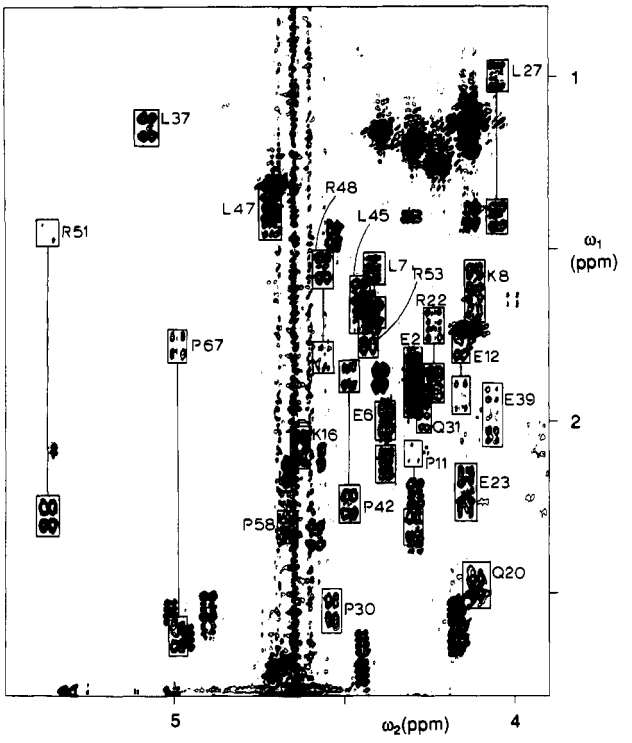

FIGURE 4: Assignment of the α-β cross peaks of the residues with long side chains in another expansion of the same DQF-COSY spectrum as shown in Figure 1.

32. The fact that it is not visible in the DQ spectrum and in the DQF-COSY spectrum can only be explained by a complete degeneracy of the δ- and ϵ-protons. The resonance position at 6.97 ppm coincides with those of the ϵ- and the ζ-proton resonances of Phe 55. Thus, this tyrosine could not have been identified without the heteronuclear experiment where the proton degeneracy is resolved by the resolution along the carbon axis. The spin system of Phe 25 was elucidated from, the DQ spectrum (Figure 5A). All resonances are well resolved, and the signal at ω_2 = 7.39 ppm can be assigned to the δ-protons on the basis of the small multiplet splitting (doublet) whereas the other two signals have the characteristic COSY triplet structure where the center component is canceled. The identification of the δ-proton resonances is confirmed by observation of a NOESY cross peak to the β-protons. For Phe 10 and Phe 55 the δ-proton resonances are well resolved, but the ϵ- and the ζ-protons are degenerate. The δ-protons can be identified as such from the NOESY cross peaks to the β-protons (Figure 5). The 4-protons can only be resolved from the 3,5-protons in the heteronuclear spectrum due to the separation of the 4-carbon resonance from the ϵ-carbon resonance along the ω_1-axis. For Phe 3 and Phe 36 all aromatic resonances have very similar chemical shifts (Figure 5A). The δ-proton resonances can be identified from the NOESY cross peaks to the β-protons in Figure 5B. In the heteronuclear ^1H–^{13}C spectrum the resonances are better resolved and can readily be identified since the ζ-carbon resonance is always shifted to higher field relative to the δ- and ϵ-carbon resonances. The histidine δ- and ϵ-proton resonances have well-resolved COSY cross peaks, and the connectivities to the β-protons are obtained from the NOESY spectrum of Figure 5B (see Table I).

NH Connectivities. Figure 6 shows the fingerprint region of a COSY spectrum of eglin c in H_2O at 36 °C. This spectrum, a DQF-COSY at 20 °C, a RELAY, and a TOCSY spectrum provided unambiguous connectivities between backbone amide protons and the spin systems of the nonlabile

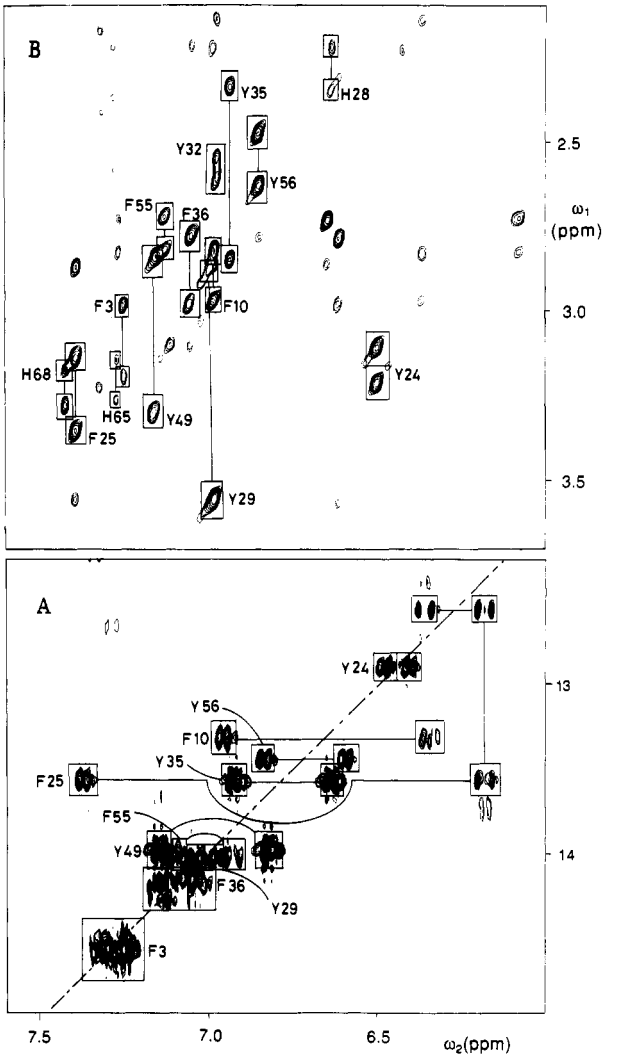

FIGURE 5: Connectivities between aromatic spin systems and β-protons. The assignments of the aromatic spin systems are shown in a DQ spectrum (A). Connectivities to the β-protons of the same residues were obtained in a NOESY spectrum (B).

protons. In the TOCSY spectrum we obtained, in addition, connectivities between the side-chain amino groups of the lysines and the rest of the side chains, as well as between the $N^{\epsilon}H$ of arginines and the γ-methylene protons. The H^{ϵ} of Arg 51 has an unusual low-field shift and is resolved from the proton resonances of the guanidinium group. Therefore, we can observe NOE's to the latter resonances. These signals are located at 6.58 and 6.64 ppm. Both signals are broad. The assignments of these signals are important since, according to the crystal structure, the side chain of Arg 51 is involved in several hydrogen bonds stabilizing the proteinase binding loop.

Sequential Assignments. The connectivities obtained in the process of sequential assignments are summarized in Figure 7. We follow the standard notations for $d_{\alpha N}$, d_{NN}, and $d_{\beta N}$ connectivities as introduced and used previously (Wagner & Wüthrich, 1982; Wüthrich et al., 1982, 1984). For prolines, connectivities to the δ-protons stand in place of connectivities to the amide protons. Filled circles in the first line of Figure 7 indicate slow amide proton exchange at pH 3.0, 20 °C. Amide proton exchange rates were measured quantitatively in NOESY spectra of proteins dissolved in ^2H_2O. The sequential NOESY connectivities were obtained from spectra at different mixing times (50, 75, 100, 150, and 200 ms), at 36 °C. A few additional spectra were recorded at 20 °C to

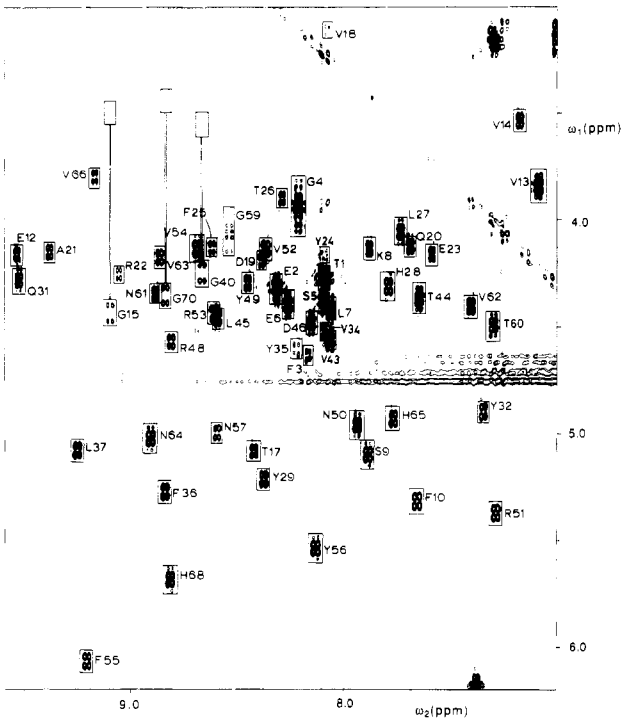

FIGURE 6: Region of the H^N–H^α cross peaks in a DQF-COSY spectrum in H_2O (fingerprint region).

resolve problems of overlap with the water resonance. The documentation is given in spectra with a mixing time of 200 ms, 36 °C. Figure 8 shows d_{NN} connectivities; Figure 9 provides $d_{\alpha N}$ connectivities. In Figure 9 we have used the same representation of $d_{\alpha N}$ connectivities as first described in the assignments of metallothionein 2 (Wagner et al., 1986). The rectangular boxes indicate the positions of the intraresidue NH–H^α cross peaks, for which the positions are known from the COSY of Figure 6. From these boxes horizontal lines lead to sequential $d_{\alpha N}$ cross peaks connecting with the following residue; vertical lines lead to the preceding residue. $d_{\alpha N}$ connectivities listed in Figure 7 that are not shown in Figure 9 were observed at the lower temperature (20 °C). At least one sequential connectivity was found for each peptide link, except for the sequence Asp[46]-Leu[47]. This is because the amide protons of these two residues coincide (Table I) so that sequential connectivities cannot be distinguished from intraresidue cross peaks.

Figure 9A shows $d_{\alpha N}$ connectivities from Thr 1 to Gly 4. Since the N-terminal region of eglin seems to be mobile, residual J contributions are present in the intraresidue cross peaks for Thr 1 and Glu 2. Figure 9B shows $d_{\alpha N}$ connectivities from Gly 4 to Ser 5, from Leu 7 to Phe 10, from Pro 11 to Val 14, from Gly 15 to Lys 16, and from Thr 17 to Gln 20. The connectivity between Ser 5 and Glu 6 cannot be identified unambiguously because it coincides with the intraresidue peak of Glu 6. The connectivity between Glu 6 and Leu 7 is partially overlapped between the intraresidue cross peaks of Ser 5 and Leu 7 (see Figure 9B). The $d_{\alpha\delta}$ connectivity between Phe 10 and Pro 11 is in another part of the spectrum and very strong. The cross peak between Val 14 and Gly 15 coincides with the intraresidue peak of Gly 15 and cannot be identified unambiguously. Figure 9C shows all the $d_{\alpha N}$ connectivities between Ala 21 and Tyr 29. A $d_{\alpha N}$ connectivity between Gln 20 and Ala 21 cannot be established unambiguously because of the near degeneracy of the α-protons of both residues. A $d_{\alpha\delta}$ connectivity is observed between Tyr 29 and Pro 30 (not shown). Figure 9D shows the $d_{\alpha N}$ connectivities from Pro 30 to Leu 37, from Pro 38 to Ser 41, from Pro 42 to Asp 46, and from Arg 48 to Tyr 49. The $d_{\alpha\delta}$ connectivities between Leu 37 and Pro 38 and between Ser 41 and Pro 42 are observed in a ^2H_2O NOESY spectrum. The $d_{\alpha N}$ connectivity between Asp 46 and Leu 47 cannot be obtained because the amide protons of both residues are nearly degenerate. For the same reason neither a d_{NN} nor a $d_{\beta N}$ connectivity can be identified. This is the only part of the sequence where there is a break in the sequential connectivities. The $d_{\alpha N}$ connectivity between Leu 47 and Arg 48 is bleached out by solvent irradiation. Figure 9E shows the $d_{\alpha N}$ connectivities from Tyr 49 to Asn 57, from Gly 59 to Val 66, and from Pro 67 to Val 69. The $d_{\delta N}$ connectivities between Pro 58 and Gly 59 overlap with the intraresidue cross peaks of Gly 59. Therefore, the sequential connectivities between Pro 58 and Gly 59 rely only on the $d_{\beta N}$ connectivities (Figure 8). The $d_{\alpha N}$ connectivity between Pro 58 and Gly 59 is absent because the H^α of Pro 58 overlaps with the water signal (Table I). The connectivity is present at lower temperature, but it overlaps with the intraresidue cross

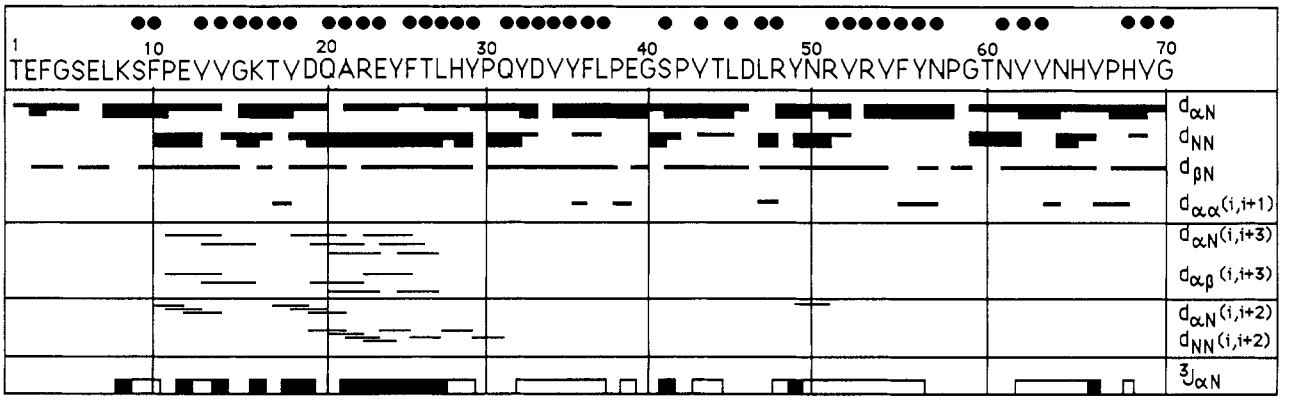

FIGURE 7: Summary of sequential and medium-range NOE's. The sequence is given with one-letter symbols of the amino acids. Filled circles above the sequence indicate residues with slowly exchanging amide protons. The first four lines below the sequence give (i,i+1) connectivities. Only the $d_{\alpha N}$ and the d_{NN} connectivities were analyzed semiquantitatively. For the $d_{\alpha N}$ connectivities we have drawn thin lines when the connectivities were only observed in the 200-ms NOESY spectrum. Medium lines were drawn when the connectivities were seen in a 50-ms NOESY spectrum but had weak intensities. Thick lines indicate intense NOE's in the 50-ms NOESY spectrum. The d_{NN} connectivities were quantified a little differently. Thin, medium, and thick lines indicate the presence of NOESY cross peaks in spectra with mixing times of 200, 75, and 50 ms, respectively. $d_{\beta N}$ connectivities are listed whenever they were observed in either a 50-, 75-, 100-, or 200-ms NOESY spectrum. A few $d_{\alpha\alpha}$ connectivities were observed in a 200-ms NOESY spectrum. All of them are very weak. Medium-range NOE's typical for helical structures are listed in the lower four lines. At the bottom is a summary of $^3J(H^\alpha,H^N)$ coupling constants. Open boxes, coupling constants $\geq$8 Hz; filled boxes, coupling constants $\leq$6 Hz.

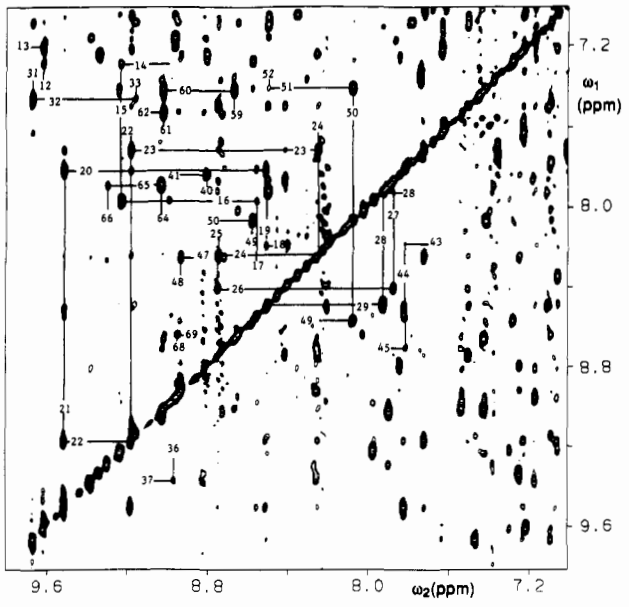

FIGURE 8: d_{NN} connectivities as observed in a NOESY spectrum of eglin in H_2O, recorded with a mixing time of 200 ms.

peak of Val 69. We observe a strong $d_{\alpha\delta}$ connectivity between Val 66 and Pro 67 in NOESY spectra in ^2H_2O. The $d_{\alpha N}$ connectivity between Gly 70 and Val 69 is absent since the H^{α} of Val 69 (see Table I) is bleached out by the water irradiation at this temperature. The connectivity is seen at 20 °C. We have recorded a NOESY spectrum in ^2H_2O with a high signal-to-noise ratio, with a mixing time of 200 ms. In this spectrum we observed a few $d_{\alpha\alpha}(i,i+1)$ connectivities. These are summarized in Figure 7. All of these cross peaks are very weak, and there is no evidence for cis peptide bonds.

Secondary Structure. We have analyzed the NOE's for typical patterns of regular secondary structures. These are presented in Figure 7. There is evidence for two helical structures in the N-terminal region of the protein. The first extends from Phe 10 to Val 13 and the second from Val 18 to Tyr 29. For both helices we observe strong d_{NN} connectivities in a 50-ms NOESY spectrum (not shown). We also observe a number of characteristic $d_{\alpha N}(i,i+3)$ and $d_{\alpha\beta}(i,i+3)$ NOE's (Figure 7). For the first helix we see three $d_{\alpha N}(i,i+2)$ NOE's similarly as for the beginning of the second helix. Only for the second helix do we observe a larger number of d_{NN}-$(i,i+2)$ NOE's. At this stage the experimental data do not distinguish whether these helices are α-helices or 3_{10}-helices since we have no resolved $d_{\alpha N}(i,i+4)$ NOE cross peaks which would identify a helix as an α-helix (Wüthrich et al., 1984).

There is experimental evidence for an irregular four-stranded β-sheet. The summary of the NOE's observed and the patters of slowly exchanging amide protons are given in Figure 10. Lys 8 to Phe 10 form a short antiparallel sheet with His 68 and Val 69. Lys 16 and Thr 17 are in an antiparallel sheet with Val 62 and Val 63. The strands Lys 8 to Val 18 and Val 62 to Gly 70 can be considered a longer β-sheet, interrupted by a short helix on the one strand and a β-bulge on the other strand (Figure 10). The strand from Val 62 to Val 63 forms an antiparallel β-sheet with Tyr 56 to Asn 57, and Val 69 shows NOE's to Val 52 (Figure 10). The strands Arg 51 to Asn 57 and Asp 33 to Glu 39 form a regular parallel β-sheet.

The NOE's summarized in Figure 7 indicate that there are turns centered at Gly^{15}-Lys^{16}, Pro^{30}-Gln^{31}, Gly^{40}-Ser^{41}, Leu^{47}-Arg^{48}, Tyr^{49}-Arg^{51}, Gly^{59}-Val^{62}, and Asn^{64}-His^{65}. However, only for residues Tyr^{49}-Arg^{51} do we observe a $d_{\alpha N}$-$(i,i+2)$ NOE. No attempts have been made so far to characterize the types of turns.

A summary of all interresidue NOE's is given in the NOE distance plot of Figure 11. The data represent 820 interresidue NOE's. The upper triangle of Figure 11 shows all NOE's; the lower triangle lists only the backbone–backbone NOE's. The antiparallel arrangement of the strands Lys 8 to Val 18 and Val 62 to Val 69 can readily be recognized. The interruption of the antiparallel β-sheet by the helix on one strand and the β-bulge on the other strand is manifested in a lack of interstrand contacts (Figures 10 and 11). The second antiparallel β-sheet and the parallel β-sheet are also clearly manifested in the distance plot of Figure 11.

The structure of the binding loop (Pro 42 to Asp 46) is very poorly characterized by NOE constraints. Besides intraresidue and sequential NOE's, there is only a single medium-range NOE between the NH of Asp 46 and the methyl group of Thr 44. In the crystal structure of the complex of eglin c with subtilisin Carlsberg, the conformation of the binding loop is stabilized by H-bonds between the side chains of Arg 51 and Arg 53 and the C-terminal carboxyl group of Gly 70 on the one side and the side-chain hydroxyl group of Thr 44, the main-chain carbonyl of Thr 44, the NH of Asp 46, and the side chain of Asp 46 on the other side. In addition, there are a number of H-bonds between an internal water molecule and the residues Thr 44, Asp 46, and Arg 51 (Bode et al., 1986, 1987; McPhalen & James, 1987). H-bonds are difficult to characterize by NMR. Involvement of a NH in an intramolecular H-bond can be recognized by observation of slow hydrogen exchange. However, identification of the H-bond acceptor has been possible so far only within regular secondary structures where also NOE's are observed which are characteristic for this type of secondary structure. If the H-bond acceptor is a titrable group, and the pK_a of this titrable group is known, the interaction with the H-bond donor may be recognized from a large pH-dependent chemical shift change of the proton of the H-bond acceptor. This is not the case for the H-bonds to side chains that are supposedly stabilizing the binding loop of eglin c.

In the crystal structure of the complex between eglin c and subtilisin Carlsberg, the eglin strand Pro 42 to Leu 45 is the central strand of a triple-stranded β-sheet where the two outer strands belong to the proteinase (Bode et al., 1986). It can therefore be expected that there are differences in the structures of free and complexed eglin c. Indeed, the H-exchange rates measured for free eglin c (Figure 7) are inconsistent with a conformation of the binding loop identical with the conformation of the complex. The amide protons of the residues Ser 41, Val 43, Leu 45, and Leu 47 exchange slowly although they are not involved in intramolecular H-bonds in the crystal structure of eglin. This indicates that this part of the molecule has a different structure in the free inhibitor. There is other evidence that the binding loop has a slightly different conformation in the free inhibitor. For example, in the crystal structure of the complex, the side chain of Arg 53 forms hydrogen bonds with side-chain and main-chain oxygens of Thr 44, and the δ- and ϵ-protons of Arg 53 are in close contact with the β-proton and the γ-methyl groups of Val 43 so that strong NOE's would be expected. We also have simulated the NOESY spectrum of eglin c (Beeson, Hyberts, and Wagner, unpublished results) with the coordinates of the crystal structure (Bode et al., 1987), using the CORMA algorithm (Keepers & James, 1984). This simulation yields strong NOESY cross peaks between the two residues Arg 53 and Val 43. However, in the experiment the expected NOE's are

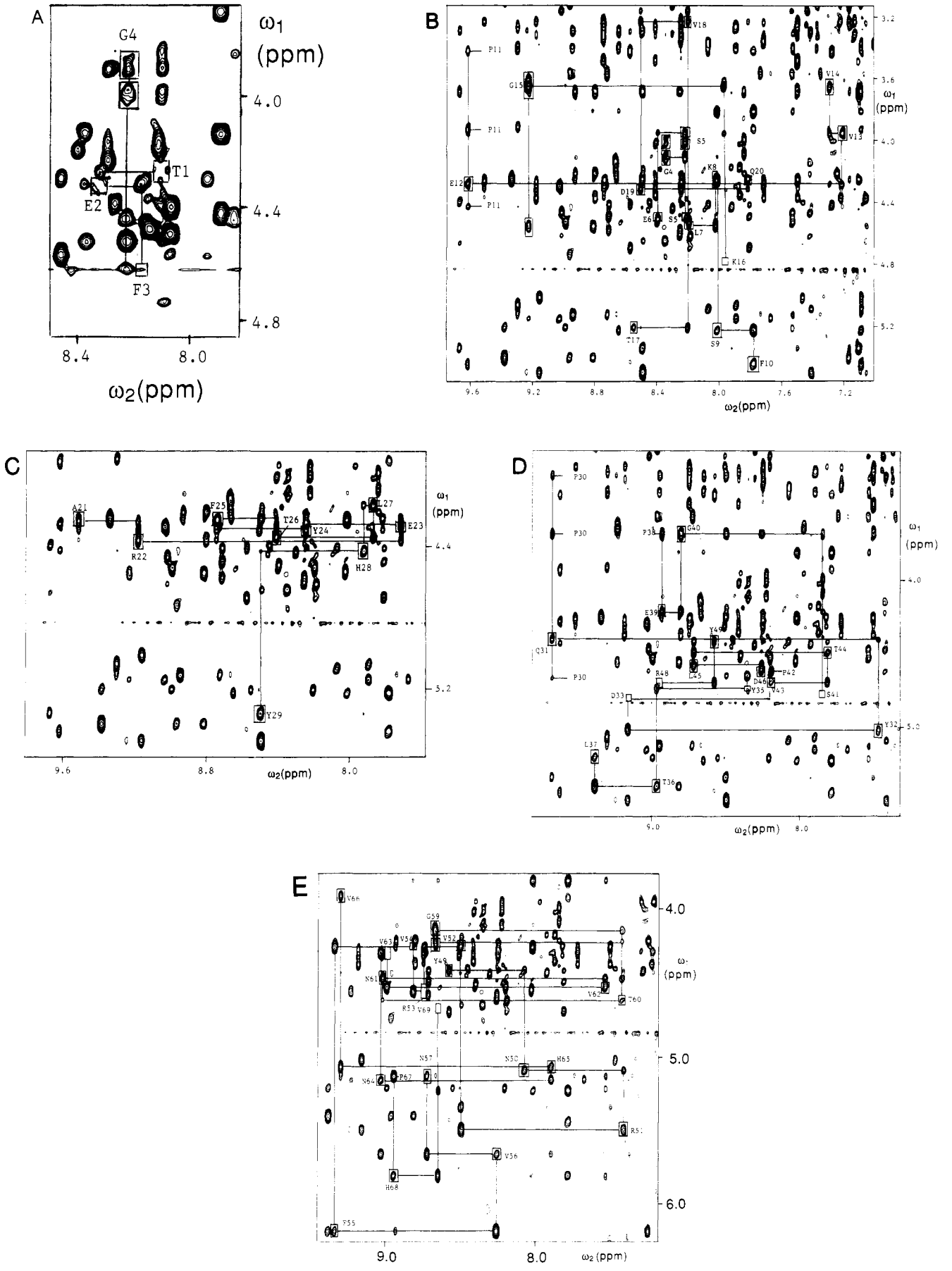

FIGURE 9: $d_{\alpha N}$ connectivities as observed in a NOESY spectrum of eglin in H_2O, recorded with a mixing time of 200 ms: (A) Thr 1 to Gly 4; (B) Gly 4 to Ser 5, Leu 7 to Phe 10, Pro 11 to Val 14, Gly 15 to Lys 16, and Thr 17 to Gln 20; (C) Ala 21 to Tyr 29; (D) Pro 30 to Leu 37, Pro 38 to Ser 41, Pro 42 to Asp 46, and Arg 48 to Tyr 49; (E) Tyr 49 to Asn 57, Gly 59 to Val 66, and Pro 67 to Val 69.

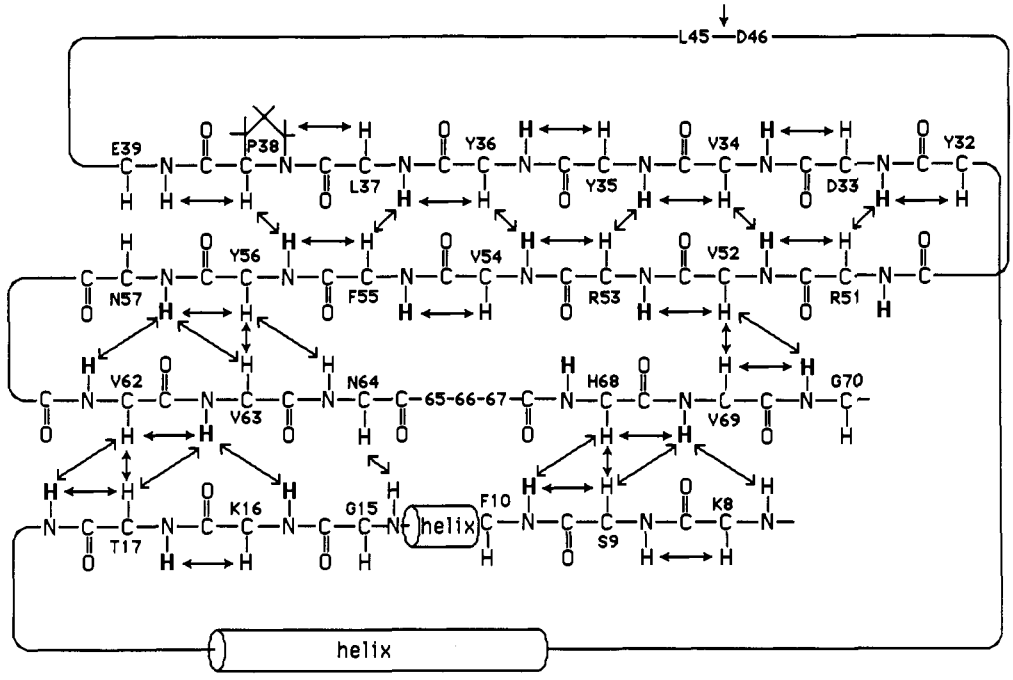

FIGURE 10: Schematic drawing of the regular secondary structure elements of eglin c. The helical segments are indicated with horizontal cylinders; the β-sheet structures are drawn explicitly. The NOE's observed in the β-sheets are indicated with arrows. Slowly exchanging amide protons are in bold letters.

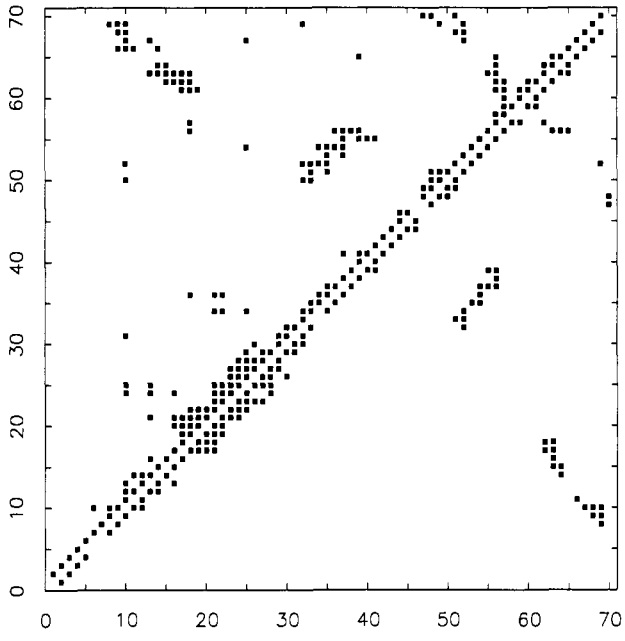

FIGURE 11: Diagonal plot of interresidue NOE distance constraints. The axes are calibrated with the amino acid sequence numbers. Each square indicates that at least one NOE was observed between the two residues in NOESY spectra of 50, 75, 100, or 200 ms. The data of this figure represent a total of 820 interresidue NOE's. The upper triangle contains all NOE constraints observed; the lower triangle contains only the backbone–backbone constraints.

completely absent. That the contact between Arg 53 and the residues Val 43 and Thr 44 is not important for the function of eglin c has further been corroborated by the fact that mutation of Arg 53 to a lysine and also mutation of Thr 44 to a proline do not affect the function of the inhibitor (Wagner et al., 1989; Grütter, Heinz, Hyberts, and Wagner, unpublished results). In the crystal structure of the complex with the proteinase, there is also a contact between the H^{α} of Arg 51 and the peptide amide proton of Arg 48, which would, according to the CORMA calculations, result in a strong NOE cross peak. However, in the experimental spectra it is completely absent at mixing times between 50 and 150 ms. At 200 ms its presence cannot be ruled out, but the cross peak would be only a little above the noise level, whereas in the same spectrum more than 800 interresidue NOE's are clearly observable. There are several other minor differences between the crystal structure of the inhibitor in the complex with the proteinase. These include, in particular, orientations of side chains such as histidines and valines. These may in part be due to the different pH at which the NMR and the X-ray studies were carried out. We have performed most of the spectroscopy at pH 3.0 to keep amide proton exchange slow and to increase the solubility of the protein. However, we have recorded 1D spectra at pH values between 2 and 8 using dilute samples. Qualitatively, the main features of the spectra do not change up to pH 8. We also have recorded 2D spectra up to pH 5 with the purpose to measure individual pK_a's; however, we did not exhaustively analyze these spectra. We did not find significant changes of the spectra. However, details, such as the orientation of individual side chains, could not be followed in this pH-variation study. At present, we are carrying out structure calculations. They result in conformations similar to those of the crystal structure of the complex, as far as the core of the protein is concerned, whereas the binding loop is little constrained and many different conformations of this strand are possible. Details of the structure calculation will be described in a subsequent publication.

ACKNOWLEDGMENTS

We thank Dr. Walter Märki, Ciba-Geigy, Basel, Switzerland, for the gift of the protein eglin c, and we are grateful to Dr. Marc Adler for preparation of Figure 11.

Registry No. Eglin c, 96380-69-7.

REFERENCES

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

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

Bax, A., Griffey, R. H., & Hawkins, B. L. (1983a) *J. Am. Chem. Soc. 105*, 7188–7190.
Bax, A., Griffey, R. H., & Hawkins, B. L. (1983b) *J. Magn. Reson. 55*, 301–315.
Bendall, M. R., Pegg, D. T., & Doddrell, D. M. (1983) *J. Magn. Reson. 52*, 81–117.
Bode, W., Papamokos, E., & Musil, D. (1987) *Eur. J. Biochem. 166*, 673–692.
Bode, W., Papamokos, E., Musil, D., Seemüller, U., & Fritz, H. (1986) *EMBO J. 5*, 813–818.
Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521–528.
Brühwiler, D., & Wagner, G. (1986) *J. Magn. Reson. 69*, 546–551.
Clore, G. M., Gronenborn, A. M., Kjaer, M., & Poulsen, F M. (1987a) *Protein Eng. 1*, 305–311.
Clore, G. M., Gronenborn, A. M., James, M. N. G., Kjaer, M., McPhalen, C. A., & Poulsen, F. M. (1987b) *Protein Eng. 1*, 313–318.
Dubs, A., Wagner, G., & Wüthrich, K. (1979) *Biochim. Biophys. Acta 577*, 177–194.
Frey, M. H., Wagner, G., Vašak, M., Sørensen, O. W., Neuhaus, D., Wörgötter, E., Kägi, J. H. R., & Wüthrich, K. (1985) *J. Am. Chem. Soc. 107*, 6847–6851.
Hyberts, S. G., Märki, W., & Wagner, G. (1987) *Eur. J. Biochem. 164*, 625–635.
Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4553.
Keepers, J. W., & James, T. L. (1984) *J. Magn. Reson. 57*, 404–426.
Macura, S., Huang, Y., Suter, D., & Ernst, R. R. (1981) *J. Magn. Reson. 43*, 259–281.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.
Märki, W., Rink, H., Schnebli, H. P., Liersch, M., Raschdorf, F., & Richter, W. (1985) in *Peptides: Structure and Function, Proceedings of the 9th American Peptide Symposium* (Deber, C. M., Hruby, V. J., & Kopple, K. D., Eds.) pp 385–388, Pierce Chemical Co., Rockford, IL.
McPhalen, C. A., & James, M. N. G. (1987) *Biochemistry 26*, 261–269.
McPhalen, C. A., & James, M. N. G. (1988) *Biochemistry 27*, 6582–6598.
McPhalen, C. A., Schnebli, H. P., & James, M. N. G. (1985) *FEBS Lett. 188*, 55–58.
Müller, L. (1979) *J. Am. Chem. Soc. 101*, 4481–4484.
Müller, L., & Ernst, R. R. (1979) *Mol. Phys. 38*, 963–992.
Nagayama, K., & Wüthrich, K. (1981) *Eur. J. Biochem. 115*, 653–657.
Piantini, U., Sørensen, O. W., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6800–6801.
Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1984) *Biochem. Biophys. Res. Commun. 117*, 479–485.
Rance, M., Bodenhausen, G., Wagner, G., Wüthrich, K., & Ernst, R. R. (1985) *J. Magn. Reson. 62*, 497–510.
Seemüller, U., Meier, M., Ohlson, K., Müller, H.-P., & Fritz, H. (1977) *Hoppe-Seyler's Z. Physiol. Chem. 358*, 1105–1117.
Seemüller, U., Eulitz, M., Fritz, H., & Strobl, A. (1980) *Hoppe-Seyler's Z. Physiol. Chem. 361*, 1841.
Shaka, A. J., & Freeman, R. (1983) *J. Magn. Reson. 51*, 169–173.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.
Wagner, G. (1983) *J. Magn. Reson. 55*, 151–156.
Wagner, G., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 347–384.
Wagner, G., & Zuiderweg, E. R. P. (1983) *Biochem. Biophys. Res. Commun. 113*, 854–860.
Wagner, G., & Brühwiler, D. (1986) *Biochemistry 25*, 5839–5843.
Wagner, G., Hyberts, S. G., Heinz, D. W., & Grütter, M. G. (1989) *Proceedings of the Sixth Conversation in the Discipline Biomolecular Stereodynamics* (Sarma, R. H., Ed.) Adenine Press, Albany, NY (in press).
Wider, G., Lee, K. H., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 367–388.
Wüthrich, K., Wider, G., Wagner, G., & Braun, W. (1982) *J. Mol. Biol. 155*, 311–319.
Wüthrich, K., Billeter, M., & Braun, W. (1984) *J. Mol. Biol. 180*, 715–740.