Journal of Biomolecular NMR, 9 (1997) 11–24
ESCOM

J-Bio NMR 406

Assignments and structure determination of the catalytic domain of human fibroblast collagenase using 3D double and triple resonance NMR spectroscopy

Mark A. McCoy[a,*,**], Martin J. Dellwo[a,**], Diane M. Schneider[a], Tracey M. Banks[a], Joseph Falvo[b], Karen J. Vavra[b], Alan M. Mathiowetz[a], M. Walid Qoronfleh[a], Richard Ciccarelli[a], Ewell R. Cook[b], Tricia A. Pulvino[a], Robert C. Wahl[a] and Hsin Wang[a,***]

[a]*Sterling Winthrop Pharmaceuticals Research Division, 1250 South Collegeville Road, Collegeville, PA 19426, U.S.A.*
[b]*Eastman Kodak Company, Rochester, NY 14650, U.S.A.*

Received 17 July 1996
Accepted 8 October 1996

Keywords: Collagenase; Assignments; Structure; MMP; 3D NMR; Matrix metalloproteinase

Summary

We report here the backbone ^{1}HN, ^{15}N, $^{13}C^{\alpha}$, ^{13}CO, and $^{1}H^{\alpha}$ NMR assignments for the catalytic domain of human fibroblast collagenase (HFC). Three independent assignment pathways (matching ^{1}H, $^{13}C^{\alpha}$, and ^{13}CO resonances) were used to establish sequential connections. The connections using $^{13}C^{\alpha}$ resonances were obtained from HNCOCA and HNCA experiments; ^{13}CO connections were obtained from HNCO and HNCACO experiments. The sequential proton assignment pathway was established from a 3D ($^{1}H/^{15}N$) NOESY-HSQC experiment. Amino acid typing was accomplished using ^{13}C and ^{15}N chemical shifts, specific labeling of ^{15}N-Leu, and spin pattern recognition from DQF-COSY. The secondary structure was determined by analyzing the 3D ($^{1}H/^{15}N$) NOESY-HSQC. A preliminary NMR structure calculation of HFC was found to be in agreement with recent X-ray structures of human fibroblast collagenase and human neutrophil collagenase as well as similar to recent NMR structures of a highly homologous protein, stromelysin. All three helices were located; a five-stranded β-sheet (four parallel strands, one antiparallel strand) was also determined. β-Sheet regions were identified by cross-strand $d_{\alpha N}$ and d_{NN} connections and by strong intraresidue $d_{\alpha N}$ correlations, and were corroborated by observing slow amide proton exchange. Chemical shift changes in a selectively ^{15}N-labeled sample suggest that substantial structural changes occur in the active site cleft on the binding of an inhibitor.

Introduction

The matrix metalloproteinases (MMPs) are a family of proteinases that includes collagenase, stromelysin, and gelatinase. (For a review on MMPs, see Docherty and Murphy, 1990.) They are synthesized in connective tissues and are of great importance in both physiological and pathological tissue degradation. Collagenase is responsible for the specific cleavage of triple helical collagen (Fields et al., 1987), which is a major structural component of articular cartilage. The inhibition of this enzyme may therefore result in the control of osteoarthritis and rheumatoid arthritis.

Members of the MMP family share a high degree of sequence homology, including an N-terminal catalytic domain and a C-terminal hemopexin-like domain that may be involved in matrix binding. The only exception is the punctuated metalloproteinase (PUMP), which lacks the C-terminal domain (Busiek et al., 1992). The catalytic domain of each MMP contains two zinc ions, one catalyt-

*To whom correspondence should be addressed.
**Present address: Department of Chemistry, University of Pennsylvania, 231 South 34th street, Philadelphia, PA 19104, U.S.A.
***Present address: Department of Chemistry, College of Staten Island, Staten Island, NY 10314, U.S.A.

Abbreviations: MMPs, matrix metalloproteinases; HFC, human fibroblast collagenase; HFS, human fibroblast stromelysin; HNC, human neutrophil collagenase; HSQC, heteronuclear single quantum correlation; NOESY, nuclear Overhauser enhancement spectroscopy; TOCSY, total correlation spectroscopy; COSY, correlated spectroscopy; NOE, nuclear Overhauser enhancement; TSP, sodium 3-(trimethylsilyl)propionate-2,2,3,3-d_4.

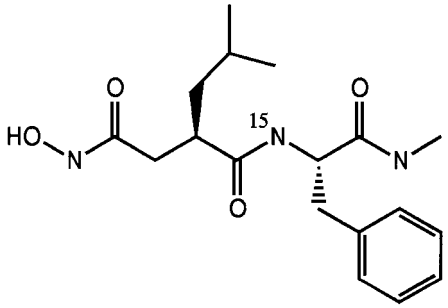

Fig. 1. The structure of the hydroxamate inhibitor used in this study. Note that the P'_2 site phenylalanine is labeled with ^{15}N.

ic and one structural, and a number of calcium ions. With the exception of PUMP, MMPs have a molecular weight greater than 50 kDa, and are sometimes glycosylated. Glycosylation and large molecular weights are clearly unfavorable for structural studies using techniques such as NMR spectroscopy and X-ray crystallography. We (Vavra et al., 1993) as well as others have constructed truncated forms of MMPs that correspond to the catalytic domains and have expressed them in *E. coli* for biophysical studies. Several structural analyses of truncated catalytic MMP domains have been reported in the literature. X-ray structures have been reported for truncated human fibroblast collagenase (Borkakoti et al., 1994; Lovejoy et al., 1994a,b; Spurlino et al., 1994) and truncated human neutrophil collagenase (Reinemer et al., 1994; Stams et al., 1994). NMR assignments (Gooley et al., 1993; Van Doren et al., 1993) of the catalytic domain of stromelysin have been reported and both X-ray (Becker et al., 1995) and NMR (Gooley et al., 1994,1996; Van Doren et al., 1995) structures have been solved. In this work we report the NMR backbone assignments, and the secondary and tertiary structure of the 19.2 kDa catalytic domain of human fibroblast collagenase. We rely on triple resonance NMR techniques for the sequential backbone assignment of all non-proline amino acid residues in the protein and on NOE connectivity patterns and distance restraints for determining the secondary structure.

Materials and Methods

Expression and purification of trHFC

A recombinant C-terminally truncated procollagenase with a termination codon after Pro^{250} was cloned into a pET11a plasmid, pKV150, and grown in a BL21 (DE3) *E. coli* expression host. The purification procedure, which includes the recovery of protein from inclusion bodies and the activation of the latent enzyme by (aminophenyl)-mercuric acetate (APMA), has been described previously (Vavra et al., 1993). Although the cDNA was obtained from Dr. Gregory Goldberg (Washington University, St. Louis, MO, U.S.A.), three locations (Thr^{116}, Asp^{201}, and Arg^{210}) in our sequence differ from published sequences (Witham et al. (1986) reported Thr^{116}, His^{201}, Thr^{210} and Goldberg et al. (1986) reported Arg^{116}, Asp^{201}, Arg^{210}). These discrepancies have been verified by DNA sequencing, mass spectrometry, and X-ray diffraction (Spurlino et al., 1994) studies.

Preparation of isotope-labeled trHFC for NMR studies

Uniformly ^{15}N-labeled protein was produced in minimal media containing 1 g/l $^{15}NH_4Cl$ (Martek, Columbia, MD, U.S.A.; Cambridge Isotope, Woburn, MA, U.S.A.). Uniformly ^{13}C,^{15}N dual labeled protein was produced in the same media containing 4 g/l [$^{13}C_6$]-D-glucose (Martek). The ^{15}N-Leu labeled protein was produced in a modified LeMaster medium (LeMaster and Richards, 1988) using the same prototrophic strain. All the supplemental amino acids and nucleic acid bases were from Sigma (St. Louis, MO, U.S.A.) and ^{15}N-Leu was from Cambridge Isotope. All the protein samples were inhibited by a hydroxamate inhibitor (Fig. 1) that was labeled with ^{15}N at the P'_2 site (Ghose et al., 1995) in a slight excess of a 1:1 molar ratio of inhibitor to protein, to ensure complete inhibition. The K_i of the inhibitor was 6 nM. All except one of the NMR samples were prepared in an H_2O buffer at pH 5.5 containing 50 mM sodium acetate-d_3, 10 mM $CaCl_2$, 0.2 M NaCl, and 0.02% NaN_3. The ^{15}N-labeled sample was prepared in D_2O at pD 5.5. All samples contained 4% DMSO-d_6, which was used to introduce the inhibitor and served as a deuterium lock in all the experiments except for those carried out on the D_2O sample. Protein concentration was measured according to Mach et al. (1992). The protein concentrations of the NMR samples were: ^{15}N-labeled trHFC in H_2O, 0.7 mM; ^{15}N-labeled trHFC in D_2O, 0.85 mM; ^{13}C,^{15}N-trHFC, 0.9 mM; and ^{15}N-Leu labeled trHFC, 0.5 mM.

NMR spectroscopy

All NMR experiments were carried out at 30 °C. Proton shifts were referenced against 0.08 M TSP (sodium 3-(trimethylsilyl)propionate-2,2,3,3-d_4, from Cambridge Isotope) in the same buffer solution. Carbon shifts were also referenced against 0.08 M TSP ($\delta = 1.7$ ppm, methyl carbon). Nitrogen shifts were referenced against 2.9 M NH_4Cl in 1 M HCl ($\delta = 24.93$ ppm; Levy and Lichter, 1979). All the referencing spectra were measured with 4% DMSO-d_6 as the deuterium lock.

The HNCACO (Clubb et al., 1992) data set was collected on a Varian Unity PLUS 600 spectrometer with three rf channels and a (^1H/^{15}N/^{13}C) triple resonance probe with z-gradients. All the other data sets were collected on a three-channel Bruker AMX 500 with a (^1H/^{15}N/^{13}C) triple resonance probe. The indirect dimensions of all double and triple resonance experiments were collected using States-TPPI phase incrementation (Marion et al., 1989a). A weak (~25 Hz) off-resonance rf field was ap-

plied for 900 ms to suppress the solvent resonance in those experiments with the carrier at 8.36 ppm (^{1}H). Otherwise, a weak (~25 Hz) cw presaturation field was applied on-resonance (4.712 ppm) for 900 ms.

Two (^{1}H/^{15}N) HSQC experiments (Bodenhausen and Ruben, 1980; Marion et al., 1989a) were performed on the ^{15}N-trHFC sample in H_2O. The first HSQC has a sweep width of 2000 Hz in the ^{15}N dimension and allows us to label each ^{1}H/^{15}N correlation peak with good resolution in the ^{15}N dimension. The second experiment has a sweep width of 10 000 in the indirect ^{15}N dimension and allows us to determine the nitrogen chemical shift of arginine and histidine side chains, which are aliased in the first experiment. The 3D NOESY-HSQC and 3D TOCSY-HSQC (Fesik and Zuiderweg, 1988; Marion et al., 1989b) were acquired with 1024 complex points in the acquisition dimension (f3), over a sweep width of 8000 Hz. A total of 32 transients were averaged for each indirect point; 128 complex points were sampled in the indirect ^{1}H dimension over a sweep width of 8000 Hz and 32 complex points were sampled in the ^{15}N dimension over a sweep width of 2000 Hz. The NOESY mix time was set to 100 ms. An MLEV-16 sequence (Levitt and Freeman, 1981) with a peak B_1 field of 6 kHz and a duration of 37 ms was used to achieve isotropic mixing of scalar coupled proton spin networks. DQF-COSY and 2D TOCSY experiments were performed on unlabeled trHFC in H_2O and D_2O and were useful in confirming some spin systems.

The HNCA (Kay et al., 1990) and HNCOCA (Ikura and Bax, 1991) triple (^{1}H/^{15}N/^{13}C) resonance experiments were acquired with $512 \times 32 \times 22$ complex points in the ^{1}H, ^{15}N, and $^{13}C^{\alpha}$ dimensions, respectively. The HNCO (Kay et al., 1990) was collected with $512 \times 32 \times 22$ complex points in the ^{1}H, ^{15}N, and ^{13}CO dimensions, respectively. The HNCACO (Clubb et al., 1992) was collected with $1024 \times 32 \times 28$ complex points in the ^{1}H, ^{15}N, and ^{13}CO dimensions, respectively. The CBCACONH (Grzesiek and Bax, 1992a) was acquired with $1024 \times 32 \times 27$ complex points in the ^{1}H, ^{15}N, and $^{13}C^{\alpha/\beta}$ dimensions, respectively. The center frequencies for HNCA, HNCOCA, and HNCO were 8.36 ppm (^{1}H) , 121.66 ppm (^{15}N), 57.1 ppm ($^{13}C^{\alpha}$), and 179.84 ppm (^{13}CO); the proton carrier for

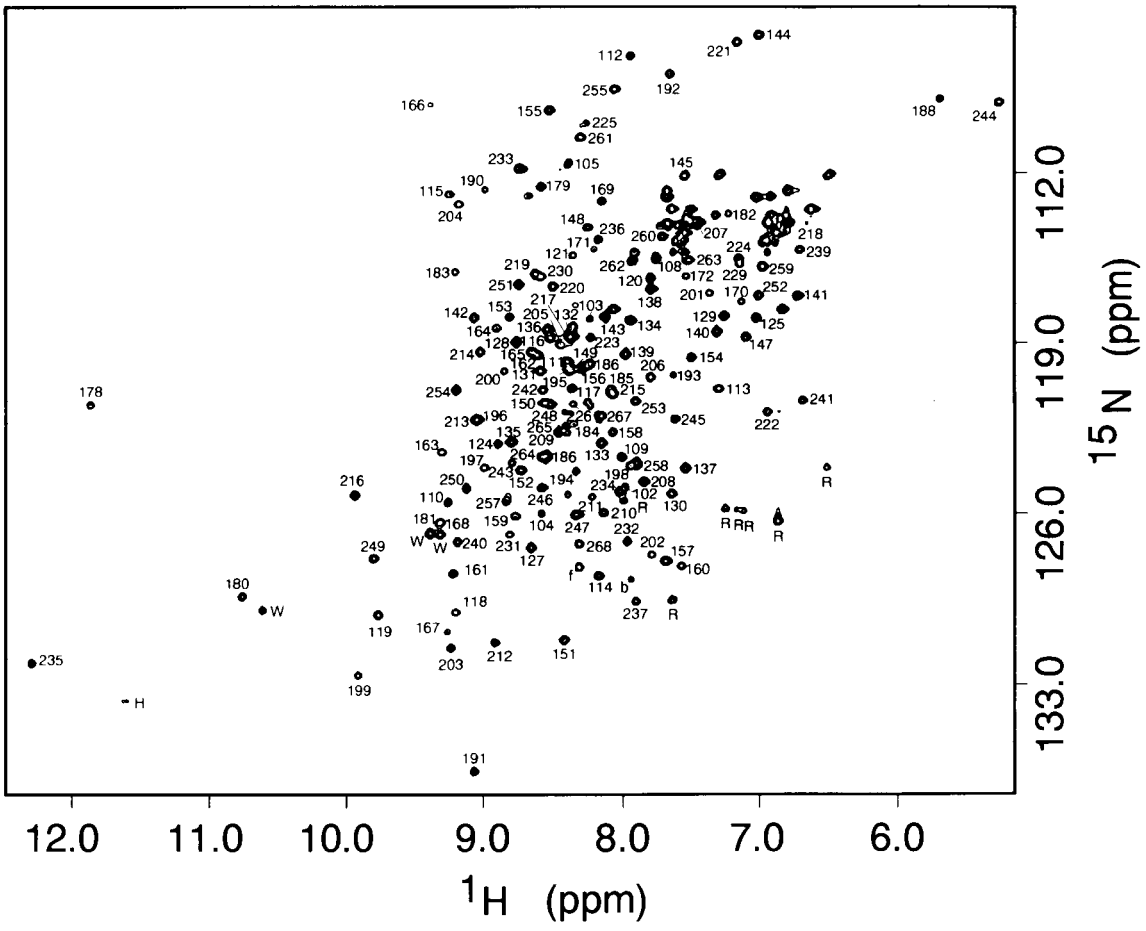

Fig. 2. ^{15}N-HSQC spectrum of trHFC in H_2O at 30 °C. The sample conditions are described in the text. The correlation peaks are indicated by residue numbers, except that f denotes the labeled amide of the free inhibitor, b that of the bound inhibitor, W the HNε1 of tryptophan, H the side chain of histidine (aliased), and R the HNε of arginine (aliased). Cross peaks due to the side chains of asparagine and glutamine are not labeled.

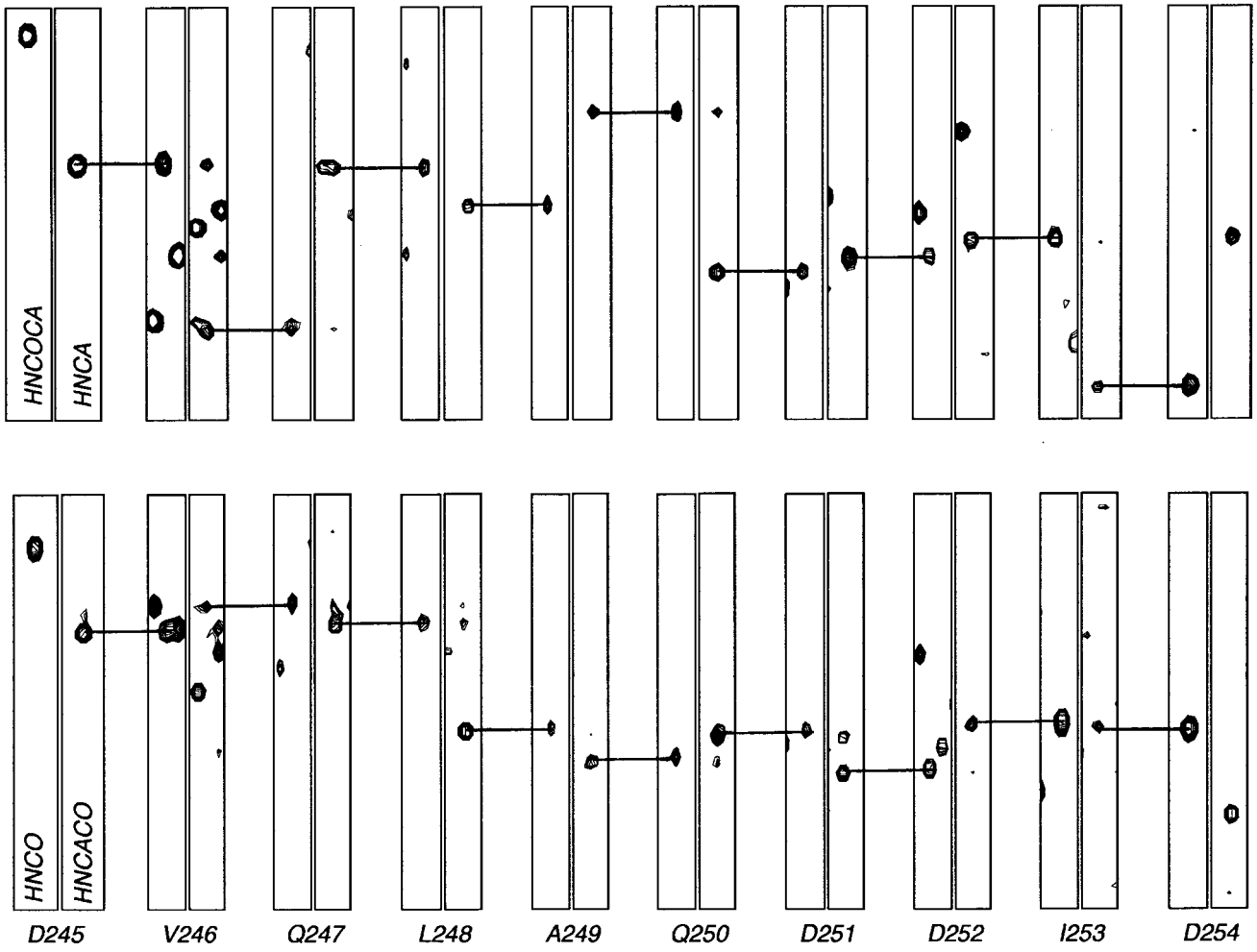

Fig. 3. Correlation of residues Asp^{245}–Asp^{254} of trHFC using four 3D triple resonance NMR data sets. Four ^{15}N planes, one for each of the HNCOCA, HNCA, HNCO, and HNCACO experiments, are shown for each residue. A sequential link was established between Asp^{245}–Val^{246} using two independent ($^{13}C^{\alpha}$ and ^{13}CO) pathways. The HNCA plane for Asp^{245} contains a peak with the same $^{13}C^{\alpha}$ chemical shift as the HNCOCA plane for residue Val^{246}. Similarly, the HNCACO plane for residue Asp^{245} contains a peak with the same ^{13}CO chemical shift as the HNCO plane for residue Val^{246}. Residues Val^{246}–Gln^{247}, etc. are similarly linked. Breaks in this assignment strategy occur at proline residues (trHFC has 11 proline residues) when no HNCACO correlation could be made (17 missing peaks for non-proline residues) or when there are degeneracies for ^{1}H, ^{15}N, $^{13}C^{\alpha}$, and ^{13}CO resonances.

the HNCACO experiment was positioned at 4.712 ppm. Carrier frequencies of 4.712 ppm (^{1}H), 121.66 ppm (^{15}N), and 45.0 ppm ($^{13}C^{\alpha/\beta}$) were used for the CBCACONH experiment. All double and triple resonance experiments used a WALTZ-16 decoupling field (Shaka et al., 1983) of 1.5 kHz to decouple ^{15}N during acquisition.

The ^{1}H, ^{15}N, and $^{13}C^{\alpha}$ dimensions of the HNCA and HNCOCA experiments were sampled over spectral ranges of 4000, 2000, and 3000 Hz, respectively. The ^{1}H, ^{15}N, and ^{13}CO dimensions of the HNCO experiment were sampled over spectral ranges of 4000, 2000, and 2000 Hz, respectively. The ^{1}H, ^{15}N, and ^{13}CO dimensions of the HNCACO experiment were sampled over spectral ranges of 10 000, 2400, and 2300 Hz, respectively. The ^{1}H, ^{15}N, and $^{13}C^{\alpha/\beta}$ dimensions of the CBCACONH experiments were sampled over spectral ranges of 8000, 2000, and 6667 Hz, respectively. Typical rf field strengths of 25, 4.2, 4.2, and 7.8 kHz were used for the excitation of ^{1}H, $^{13}C^{\alpha}$, ^{13}CO, and ^{15}N coherences, respectively. Proton decoupling during the evolution of carbon coherence in the HNCACO experiment was achieved using the DIPSI-2 sequence (Shaka et al., 1988) with a peak field strength of 13.2 kHz. The peak proton decoupling field for the CBCACONH experiment, also using DIPSI-2, was 5.5 kHz. Experiment times ranged from 2.5 days for HNCO (64 transients for each t_3 point) to 7 days for HNCACO (160 transients for each t_3 point). The HNCAHA (Kay et al., 1992), CBCANH (Grzesiek and Bax, 1992b), and HNCACB (Wittekind and Mueller, 1993) experiments were also attempted, but the low sample concentration and the relatively poor sensitivity of these experiments made them unsuitable for this study.

All data processing was done using the FELIX software (Biosym Technologies Inc., San Diego, CA, U.S.A.).

Low-frequency deconvolution was used to remove the residual water signal. After Fourier transformation, the F3 (detection) trace was baseline corrected, typically by averaging the last 10% of the spectrum. The first point in each indirect FID was multiplied by 0.5 to reduce errors arising from the discrete Fourier transform. The carbon dimension of all 3D triple resonance experiments was extended by complex linear prediction. A 90° shifted sine-bell apodizing function was used in all dimensions. After zero-filling and subsequent Fourier transformation, the data matrix dimensions were: 512×64×64 for the HNCA, HNCO, HNCACO, and HNCOCA experiments; 512×256×64 for NOESY-HSQC and TOCSY-HSQC; and 512×64×128 for CBCACONH. The NOESY-HSQC, TOSCY-HSQC, HNCACO, and CBCACONH experiments were acquired with the carrier centered on the H_2O resonance; the right half of the F3 dimension of each of the resulting data sets was removed during processing. The assignment process was aided by custom macros written for FELIX which simplify the tasks of identifying and correlating chemical shifts across multiple spectra (M.J. Dellwo, unpublished results).

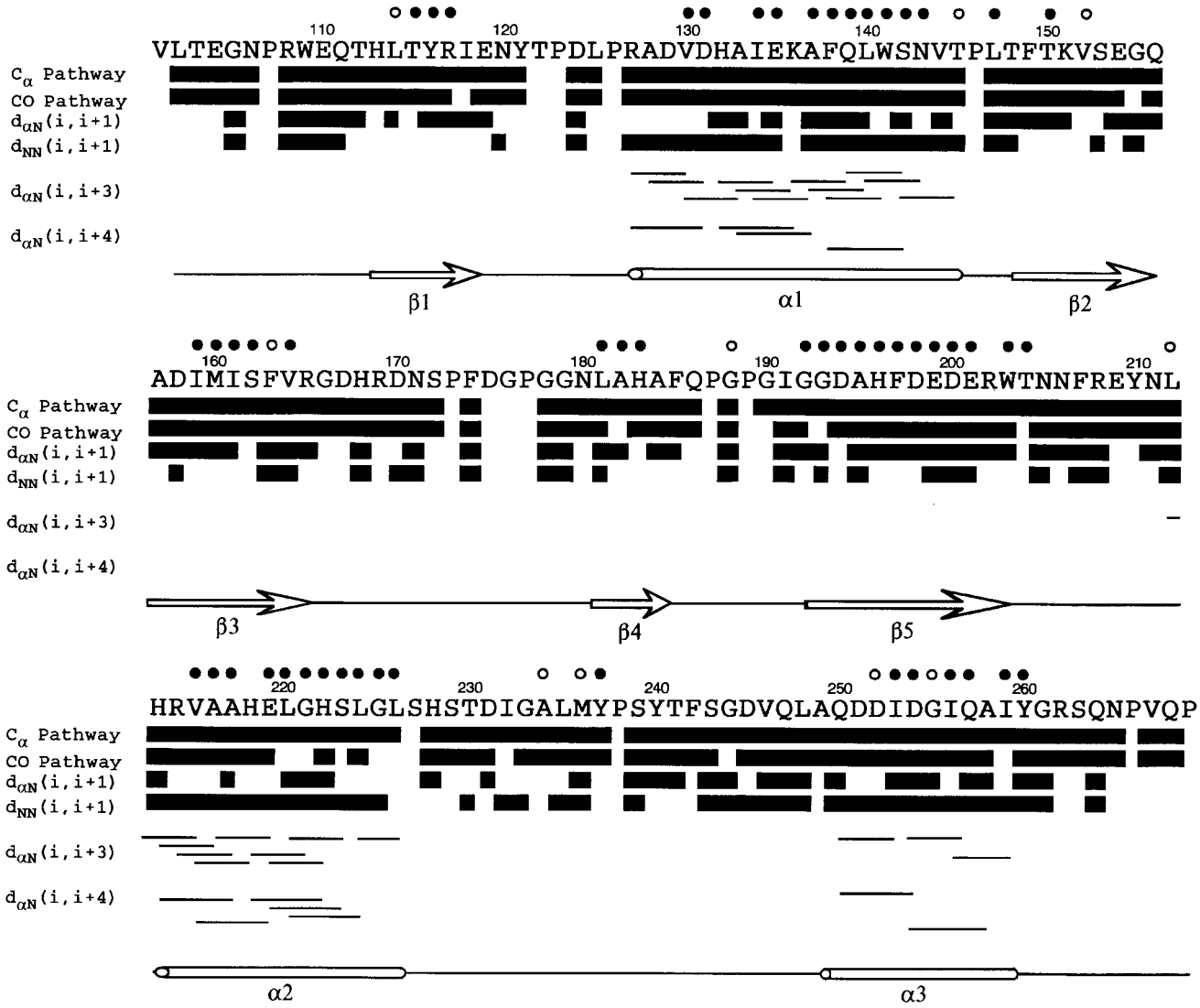

Fig. 4. Summary of sequential connections and NOE distances. The top four lines are a diagrammatic representation of the sequential connections from the $^{13}C^{\alpha}$, ^{13}CO, and proton assignment pathways. The $^{13}C^{\alpha}$ assignment pathway is established through the HNCA and HNCOCA data sets, the ^{13}CO assignment pathway is established with the HNCO and HNCACO data sets, and the proton pathway is established using the 3D ($^{1}H/^{15}N$) NOESY-HSQC and 3D ($^{1}H/^{15}N$) TOCSY-HSQC experiments. A solid bar represents a sequential assignment that has been made using that specific pathway; an open space, primarily at proline residues, represents a break in the assignment pathway. For non-proline breaks in the assignment pathway, a low signal intensity from the HNCACO data set coupled with $^{13}C^{\alpha}$ chemical shift degeneracy leads to the break. Below the sequential assignment representation is a summary of $d_{NN}(i,i\pm1)$, $d_{NN}(i,i\pm2)$, $d_{\alpha N}(i,i+1)$, $d_{\alpha N}(i,i+3)$, and $d_{\alpha N}(i,i+4)$ NOE connections that are taken from the 3D ($^{1}H/^{15}N$) NOESY-HSQC spectrum. Secondary structure is indicated by the schematic helices and sheet symbols. Proton exchange is indicated, above the sequence, by open circles (NH correlations that persist after several days, but exchange after 5 months) and closed circles (NH correlations that persist in D_2O after 5 months).

TABLE 1
BACKBONE RESONANCE ASSIGNMENTS FOR THE CATALYTIC DOMAIN OF HUMAN FIBROBLAST COLLAGENASE

Residue	^{1}HN	^{15}N	^{13}CO	$^{13}C^{\alpha}$	$^{1}H^{\alpha}$	Residue	^{1}HN	^{15}N	^{13}CO	$^{13}C^{\alpha}$	$^{1}H^{\alpha}$
Val[101]			176.82	64.06		Ile[161]	9.22	128.47	176.15	62.02	5.57
Leu[102]	7.98	124.90	178.72	54.65	4.56	Ser[162]	8.61	119.49	173.86	58.44	5.08
Thr[103]	8.24	118.04	175.75	63.39	4.31	Phe[163]	9.30	123.49	178.29	58.46	5.28
Glu[104]	8.58	126.01	178.32	58.32	4.31	Val[164]	8.90	118.40	175.29	60.54	4.83
Gly[105]	8.39	111.63	175.23	46.71	3.93	Arg[165]	8.65	119.39	177.67	55.85	5.21
Asn[106]	8.24	121.58		53.05	4.88	Gly[166]	9.38	109.21	177.76	48.55	
Pro[107]			176.59	64.87		Asp[167]	9.26	130.83	176.89	56.71	4.53
Arg[108]	7.76	115.53	177.01	55.83	4.53	His[168]	9.31	126.38	175.35	55.82	5.07
Trp[109]	8.01	123.69	179.03	58.65	3.22	Arg[169]	8.15	113.18	177.27	59.74	3.93
Glu[110]	9.25	125.55	178.28	58.90	4.38	Asp[170]	7.13	117.30	177.82	53.83	4.68
Gln[111]	8.40	119.77	176.22	55.86	4.66	Asn[171]	8.21	115.18	176.35	55.91	4.48
Thr[112]	7.94	107.21	174.93	62.95	4.25	Ser[172]	7.54	116.26		56.24	4.69
His[113]	7.30	120.89	175.21	54.98	5.11	Pro[173]					
Leu[114]	8.17	128.54	178.34	55.16	4.57	Phe[174]					
Thr[115]	9.24	112.90	176.19	59.86	5.55	Asp[175]					
Tyr[116]	8.52	118.82		56.75	5.63	Gly[176]					
Arg[117]	8.37	120.86	175.75	53.29	4.62	Pro[177]			179.87	65.14	
Ile[118]	9.20	130.05	177.03	62.85	4.37	Gly[178]	11.86	121.54	175.77	69.02	4.05
Glu[119]	9.76	130.16	177.36	61.24	3.88	Gly[179]	8.59	112.59	176.08	48.15	3.91
Asn[120]	7.80	116.34	184.87	53.45	4.67	Asn[180]	10.76	129.39	179.62	55.81	4.67
Tyr[121]	8.36	115.42		60.10	4.17	Leu[181]	9.32	126.36	177.50	57.77	4.41
Thr[122]						Ala[182]	7.23	113.68	177.85	53.25	4.66
Pro[123]			178.96	64.40		His[183]	9.20	116.10	173.48	56.17	4.76
Asp[124]	8.89	123.16	177.25	57.60	4.12	Ala[184]	8.35	122.34	176.96	52.64	5.02
Leu[125]	7.03	117.96		52.83	4.89	Phe[185]	8.09	120.91	176.91	58.95	4.37
Pro[126]			179.45	63.86		Gln[186]	8.23	119.87		55.92	
Arg[127]	8.66	127.40	180.02	61.79		Pro[187]			175.62	65.80	
Ala[128]	8.76	118.98	181.97	56.45	4.68	Gly[188]	5.69	108.97		46.26	3.69
Asp[129]	7.26	117.91	181.39	58.81	4.56	Pro[189]			180.18	64.15	
Val[130]	7.65	125.20	178.96	68.27	3.52	Gly[190]	8.99	112.71	176.79	49.04	3.99
Asp[131]	8.59	120.17	181.55	59.62		Ile[191]	9.06	136.49	176.01	63.40	4.25
His[132]	8.37	118.77	177.84	58.16	4.65	Gly[192]	7.66	107.94	176.86	48.52	3.81
Ala[133]	8.16	123.12	182.23	56.79	4.04	Gly[193]	7.64	120.34	173.75	48.16	
Ile[134]	7.94	118.11	179.03	64.38	3.63	Asp[194]	8.34	124.28	175.66	57.49	4.69
Glu[135]	8.79	123.08	181.60	61.94	4.13	Ala[195]	8.40	120.11	176.46	52.46	4.94
Lys[136]	8.54	118.46	180.71	59.81	4.07	His[196]	9.04	122.17	174.88	52.42	5.59
Ala[137]	7.55	124.16	179.55	57.11	3.99	Phe[197]	8.99	124.13	175.89	58.14	4.26
Phe[138]	7.80	116.80	180.35	65.47	3.78	Asp[198]	7.95	124.05	181.12	55.90	4.34
Gln[139]	7.98	119.49	179.23	59.78	4.22	Glu[199]	9.91	132.60	178.54	58.09	4.50
Leu[140]	7.31	118.54	180.59	59.15	3.92	His[200]	8.85	120.16	178.56	57.95	4.64
Trp[141]	6.73	117.07	181.92	60.05	4.68	Glu[201]	7.37	116.96	177.99	54.96	4.66
Ser[142]	9.06	117.98	177.98	62.67	4.34	Arg[202]	7.79	127.68	175.40	56.22	4.37
Asn[143]	8.13	117.95	178.72	56.74	4.71	Trp[203]	9.23	131.48	179.07	57.60	5.40
Val[144]	7.00	106.36	176.65	61.87	4.92	Thr[204]	9.18	113.33	176.88	61.39	4.81
Thr[145]	7.55	112.14		61.55	5.38	Asn[205]	8.45	119.11	176.61	52.67	5.05
Pro[146]			178.10	63.84		Asn[206]	7.80	120.43	175.60	53.89	4.93
Leu[147]	7.10	118.76	178.94	56.63	4.52	Phe[207]	7.45	114.06	178.96	59.76	4.38
Thr[148]	8.25	114.26	173.58	61.53	4.38	Thr[208]	7.85	124.70	177.14	57.71	3.95
Phe[149]	8.37	120.11	177.62	57.61	5.43	Glu[209]	8.46	122.68	180.12	54.04	
Thr[150]	8.56	121.46	173.21	63.17	4.49	Tyr[210]	8.14	125.97		63.99	3.82
Lys[151]	8.42	131.17	178.04	57.20	3.36	Asn[211]	8.22	125.27	177.75	56.62	4.93
Val[152]	8.73	124.24	177.73	61.22	4.63	Leu[212]	8.92	131.29	179.07	60.04	
Ser[153]	8.81	117.96	175.14	60.56	4.53	His[213]	9.05	122.17	177.73	61.68	
Glu[154]	7.50	119.61	176.42	56.26	4.48	Arg[214]	9.02	119.37	180.43	59.45	
Gly[155]	8.53	109.44	174.96	46.52	4.07, 3.79	Val[215]	8.07	121.09	179.96	52.65	3.87
Gln[156]	8.29	120.05	176.54	56.89	4.31	Ala[216]	9.93	125.27	179.92	57.38	3.97
Ala[157]	7.69	127.94	177.69	51.29	4.53	Ala[217]	8.38	118.76	180.15	60.12	3.83
Asp[158]	8.07	122.68	179.09	60.16	4.43	His[218]	6.84	114.29	177.87	60.53	4.39
Ile[159]	8.77	126.13	175.68	62.90	4.22	Glu[219]	8.63	116.19	179.97	59.67	
Met[160]	7.57	128.16	178.74	54.77	5.28	Leu[220]	8.50	116.69	179.62	57.83	3.99

TABLE 1
(continued)

Residue	^{1}HN	^{15}N	^{13}CO	$^{13}C^{\alpha}$	$^{1}H^{\alpha}$	Residue	^{1}HN	^{15}N	^{13}CO	$^{13}C^{\alpha}$	$^{1}H^{\alpha}$
Gly^{221}	7.16	106.66	179.58	49.63	3.94, 2.39	Asp^{245}	7.62	122.13	176.48	54.63	4.75
His^{222}	6.94	121.82	181.41	59.74	4.86	Val^{246}	8.58	124.97	175.64	64.08	4.10
Ser^{223}	8.23	118.78	176.64	64.46		Gln^{247}	8.34	126.05	176.17	54.75	4.53
Leu^{224}	7.15	115.53	179.69	56.68	4.59	Leu^{248}	8.52	121.54	179.50	56.97	4.03
Gly^{225}	8.27	109.98	175.90	46.28		Ala^{249}	9.79	127.85	180.46	51.71	4.65
Leu^{226}	8.36	121.53		56.66	4.47	Gln^{250}	9.12	125.00	179.61	60.77	4.72
Ser^{227}						Asp^{251}	8.75	116.61	180.76	59.96	4.70
His^{228}				57.91		Asp^{252}	7.01	117.06	179.36	58.98	4.78
Ser^{229}	7.15	115.75	177.65	56.78	4.89	Ile^{253}	7.91	121.40	179.56	67.45	3.59
Thr^{230}	8.59	116.28	176.07	63.75	4.59	Asp^{254}	9.19	120.95	182.05	58.81	4.43
Asp^{231}	8.81	126.86	177.97	55.51	4.60	Gly^{255}	8.06	108.58	178.07	49.23	3.91
Ile^{232}	7.97	127.15	178.42	64.38	2.59	Ile^{256}	8.55	123.64	180.18	62.71	4.64
Gly^{233}	8.73	111.86	175.74	46.33	4.11, 3.55	Gln^{257}	8.83	125.50	181.54	59.83	4.40
Ala^{234}	8.02	125.13	181.44	53.01	4.59	Ala^{258}	7.90	123.94	180.12	56.04	4.16
Leu^{235}	12.30	132.10	183.93	59.07	4.31	Ile^{259}	6.98	115.85	179.16	65.64	3.81
Met^{236}	8.18	114.76	178.41	55.40	4.88	Tyr^{260}	7.71	114.63	177.56	61.57	4.57
Tyr^{237}	7.91	129.59		59.37	5.09	Gly^{261}	8.31	110.56	173.79	45.82	4.41, 3.92
Pro^{238}			177.26	66.68		Arg^{262}	7.93	115.64	178.66	57.06	4.73
Ser^{239}	6.70	115.17	175.48	58.70	5.32	Ser^{263}	7.52	115.60	176.41	59.99	4.09
Tyr^{240}	9.18	127.16	176.35	62.76	3.91	Gln^{264}	8.58	123.70	177.15	57.16	4.37
Thr^{241}	6.69	121.37	173.24	61.87	4.33	Asn^{265}	8.41	122.40		52.97	4.97
Phe^{242}	8.57	120.95	176.42	59.61	4.93	Pro^{266}			178.48	64.63	
Ser^{243}	8.79	123.95	175.77	58.83	4.52	Val^{267}	8.16	122.01	177.57	64.00	4.08
Gly^{244}	5.25	109.09	173.61	47.46	4.29, 2.93	Gln^{268}	8.31	127.25		54.64	4.68

Results and Discussion

Protein stability

Our recombinant trHFC (Val^{101}–Pro^{269})* was found to have slightly higher specific activity than native full-length HFC. One factor that may have contributed to the higher activity is that trHFC does not suffer from autoproteolyzation, as does native HFC. In the presence of the inhibitor and at pH 5.5, all our labeled NMR samples, except for one, showed no detectable degradation or changes in HSQC (^{1}H/^{15}N) spectra (data not shown) over a period of a year and a half. In one sample proteolytic fragments appeared after 6 months. Filtration with a Centricon-10 (Amicon) removed the fragments and the sample remained stable for the next year. It is therefore concluded that trHFC may have proteolyzed a co-purified contaminant in that sample. Even in the absence of the inhibitor, we have evidence that the ^{15}N-Leu trHFC sample (see the discussion below) was stable for at least 2 days.

Sequential assignment strategy

Figure 2 shows the (^{1}H/^{15}N) HSQC spectrum of trHFC. Each backbone ^{1}H/^{15}N correlation is labeled according to its assignment to a specific amino acid in the protein sequence. Peaks from the side chains of tryptophan, histidine, arginine, and the inhibitor are also labeled in Fig. 2. Peaks due to asparagine and glutamine side chains account for nearly all the unlabeled peaks in this spectrum. To achieve sequential assignments with experiments that only use ^{1}H/^{1}H and ^{1}H/^{15}N correlations, we need to be able to correlate H^{α} and H^{β} resonances to backbone ^{1}HN and ^{15}N resonances. Out of the 158 non-proline residues in the 3D TOCSY-HSQC, 141 showed H^{α} resonances and 51 showed H^{β} resonances. The H^{α} resonances were, however, severely degenerate and it was impractical to achieve the sequential assignments via the d_{NN} and $d_{\alpha N}$ NOEs (Wüthrich, 1986). Sequential backbone assignments were made using 3D triple resonance methods on a single, uniformly ^{13}C,^{15}N-labeled sample in H_2O. Due to considerable $^{13}C^{\alpha}$ degeneracy, only ~20% of the sequential connections can be made unambiguously from the HNCA–HNCOCA data alone. Assignment of the backbone residues was therefore accomplished using several independent connectivity pathways. Sequential pairwise connections were established only when the $^{13}C^{\alpha}$ chemical shift of the ^{1}H(i)-^{15}N(i)-$^{13}C^{\alpha}$(i) correlation peak in the HNCA data matched the $^{13}C^{\alpha}$ chemical shift of the ^{1}H(i)-^{15}N(i)-$^{13}C^{\alpha}$(i–1) peak in the HNCOCA data, and the ^{13}CO chemical shift of the ^{1}H(i)-^{15}N(i)-^{13}CO(i) correlation peak in the HNCACO data matched the carbonyl chemical shift of the ^{1}H(i)-^{15}N(i)-^{13}CO(i–1) peak in the HNCO data. Figure 3 shows the utilization of the $^{13}C^{\alpha}$/^{13}CO pathways to make sequential assignments for residues Asp^{245}–Asp^{254}. The sequential connections made via heteronuclear scalar

*The numbering is according to Witham et al. (1986), which includes the proenzyme and 19 residues of signal peptides.

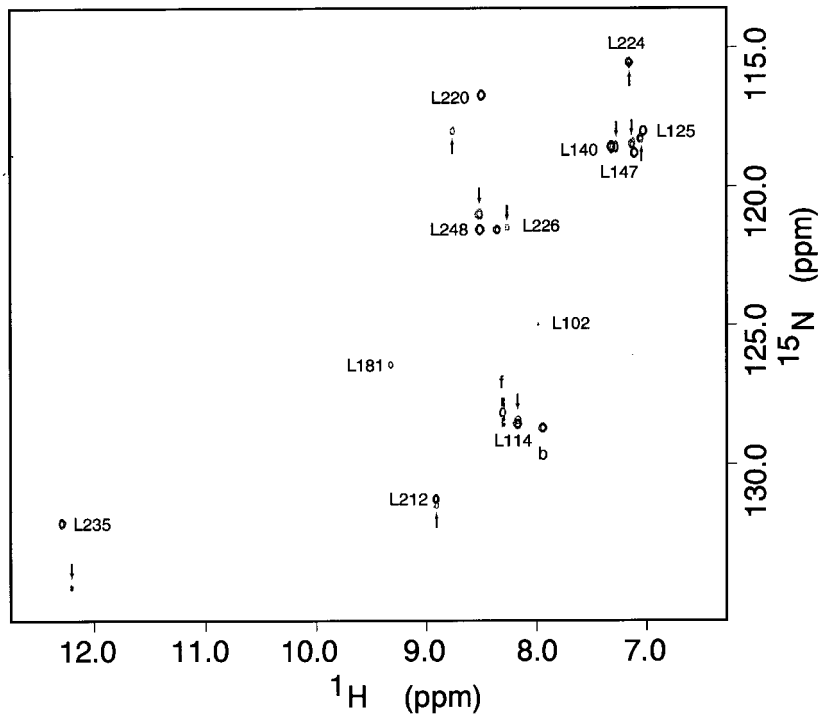

Fig. 5. ^{15}N-HSQC spectra of trHFC that is selectively labeled with ^{15}N-Leu. The spectrum of the inhibited protein is shown in solid lines and that of the uninhibited form is in dashed lines (highlighted with arrows for clarity). Assignments are indicated in the figure; f denotes the labeled amide (at the P'_2 site) of the free inhibitor and b that of the bound inhibitor.

correlations in the triple resonance experiments were verified and many remaining ambiguities were resolved when sequential d_{NN} and $d_{\alpha N}$ connections could be made via ^{1}H-^{1}H dipolar correlations in the 3D (^{1}H/^{15}N) NOESY-HSQC spectrum. In Fig. 4, a diagrammatic representation is given of the sequential connections from the $^{13}C^{\alpha}$, ^{13}CO, and proton assignment pathways. A listing of the ^{1}HN, ^{15}N, $^{13}C^{\alpha}$, ^{13}CO, and $^1H^{\alpha}$ assignments is contained in Table 1. There remain six residues where a clear assignment is not possible: Thr^{122}, Phe^{174}–Gly^{176} and Ser^{227}–His^{228}. Missing assignments arise when there is severe $^{13}C^{\alpha}$ chemical shift degeneracy and a missing correlation peak in the HNCACO data set. The lack of peaks is related to the poor sensitivity of this method relative to, for example, the HNCO experiment. The connections provided by the HNCACO experiment are, however, crucial in resolving many of the ambiguous assignments that arise from the $^{13}C^{\alpha}$ assignment pathway. Of the 17 non-proline residues that do not have an HNCACO peak, 13 could be resolved based on $^{13}C^{\alpha}$ or proton connections.

Residue typing by chemical shifts

Due to the severe proton chemical shift overlap and large (~20 Hz) proton line widths, only four residues were typed using spin pattern recognition from the 2D COSY and 2D TOCSY spectra in H_2O and D_2O. For residue typing, we relied heavily on carbon chemical shift information and the identification of leucine residues by specific labeling. $^{13}C^{\alpha}$ chemical shifts were useful in suggesting several residue types: glycine (43–49 ppm), alanine (49–57 ppm), threonine (62–65 ppm), and valine and isoleucine (> 62 ppm). All have characteristic, though not unique, chemical shift ranges. Additionally, glycine residues often have a nitrogen chemical shift of 105–110 ppm. Even more useful are the C^{β} chemical shifts from the CBCA-CONH experiment. The $^{13}C^{\beta}$ chemical shift ranges of threonine (66–72 ppm), serine (60–66 ppm), and alanine (17–23 ppm) are essentially exclusive; no other residue type falls within these chemical shift ranges. The remaining residues are clustered into two groups. Valine, tryptophan, methionine, histidine, glutamic acid, glutamine, proline, arginine and lysine are found in the range of 28–34 ppm, while leucine, isoleucine, cysteine, phenylalanine, tyrosine, asparagine and aspartic acid are found in the range of 37–45 ppm. Knowing both the $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts then becomes a powerful assignment tool as many residue types can be determined or, in the worst case, be reduced to a choice of 4–5 residue types. When this information is compared to the primary sequence of the protein, many clusters of 4–5 amino acids can be uniquely placed. For example, residues Gly^{166}–Asp^{170} (GDHRD) can be denoted as GYXXY, where X implies a $^{13}C^{\beta}$ shift of 28–33 ppm and Y implies a $^{13}C^{\beta}$ shift of 37–45 ppm: Gly^{166} is known from its $^{13}C^{\alpha}$ shift. The GYXXY pattern is unique in trHFC. When the GYXXY connections are found, they can easily be placed in the protein without any further knowledge of the side-chain assignments or residue type.

Specific labeling of leucine residues

Figure 5 shows an HSQC ($^1H/^{15}N$) spectrum of inhibited trHFC in which only leucine residues are ^{15}N labeled. All the 12 leucine residues in trHFC can be readily identified. Like Van Doren et al. (1993), we did not detect any scrambling even though aminotransferase-deficient strains were not used in the preparation of ^{15}N-Leu trHFC. Leu^{102} has much diminished intensity, which is readily explained by N-terminal heterogeneity. Grant et al. (1987) showed that activation of the proHFC resulted in a mixture of N-termini starting at Val^{101} or Leu^{102}. N-terminus heterogeneity was also indicated from protein sequencing and mass spectrometry.

Figure 5 also shows an HSQC spectrum of the uninhibited ^{15}N-Leu trHFC. Significant changes in both ^1H and ^{15}N resonances were observed in several leucine residues between the inhibited and uninhibited forms of the protein. It is interesting to note that the sample had been in the uninhibited state at 30 °C and at pH 5.5 for 2 days before the inhibitor was added to the sample. The HSQC spectrum of ^{15}N-Leu trHFC measured after the inhibitor was added indicates that the enzyme assumes the same conformation as observed in the spectra of our inhibited samples, without any sign of degradation.

Secondary and tertiary structure

Three helices and five β-strands were identified in trHFC from NOE and amide proton exchange data, as summarized in Fig. 4. Helical regions were found by locating strong $d_{NN}(i,i\pm1)$ and $d_{NN}(i,i\pm2)$ connections and characteristic $d_{\alpha N}(i,i+3)$ and $d_{\alpha N}(i,i+4)$ connections. Figure 6 displays a series of contiguous 2D slices from the 3D ($^1H/^{15}N$) NOESY-HSQC spectrum that shows the NOE network in helix I (Arg^{127}–Val^{144}). When arranged in daisy wheel plots with a pitch of 3.6 residues per turn (not shown), helix I is strongly amphipathic throughout its entire length with the exception of Ala^{128}, whereas helices II (Leu^{212}–Ser^{233}) and III (Ala^{249}–Tyr^{260}) are much less so. The presence of $d_{\alpha N}(i,i+3)$ and $d_{\alpha N}(i,i+4)$ connections and no observable $d_{\alpha N}(i,i+2)$ NOEs suggest that all three helices are essentially α-helical. From our data, helices I and II are well restrained. In addition to NOE restraints, helix II contains His^{218} and His^{222} which are coordinated (and restrained) to the catalytic zinc and thus form part of the active site. We did observe substantially fewer $d_{\alpha N}(i,i+3)$ and $d_{\alpha N}(i,i+4)$ NOE restraints in helix III compared to helices I and II, which might lead one to suggest that it is substantially less structured. Aside from the lack of numerous NOEs, other data strongly support the existence of a well-formed helix III. First, we note that most

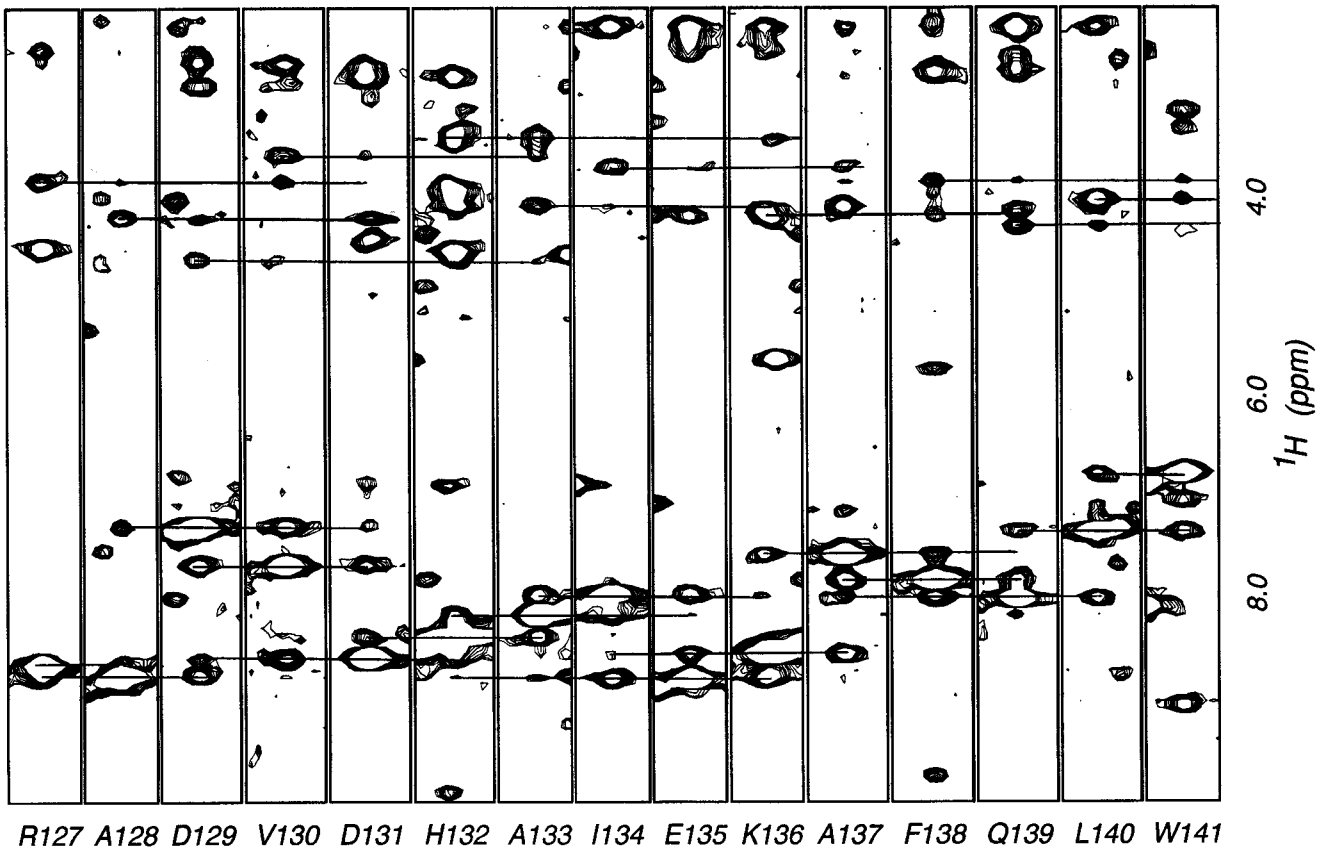

Fig. 6. Proton–proton planes from the 3D ($^1H/^{15}N$) NOESY-HSQC spectrum at the corresponding ^{15}N frequency for sequential residues Leu^{127}–Trp^{141} of trHFC. Characteristic $d_{NN}(i,i\pm1,i\pm2)$ and $d_{\alpha N}(i,i+1,i+3,i+4)$ connections are noted with horizontal lines connecting the appropriate planes.

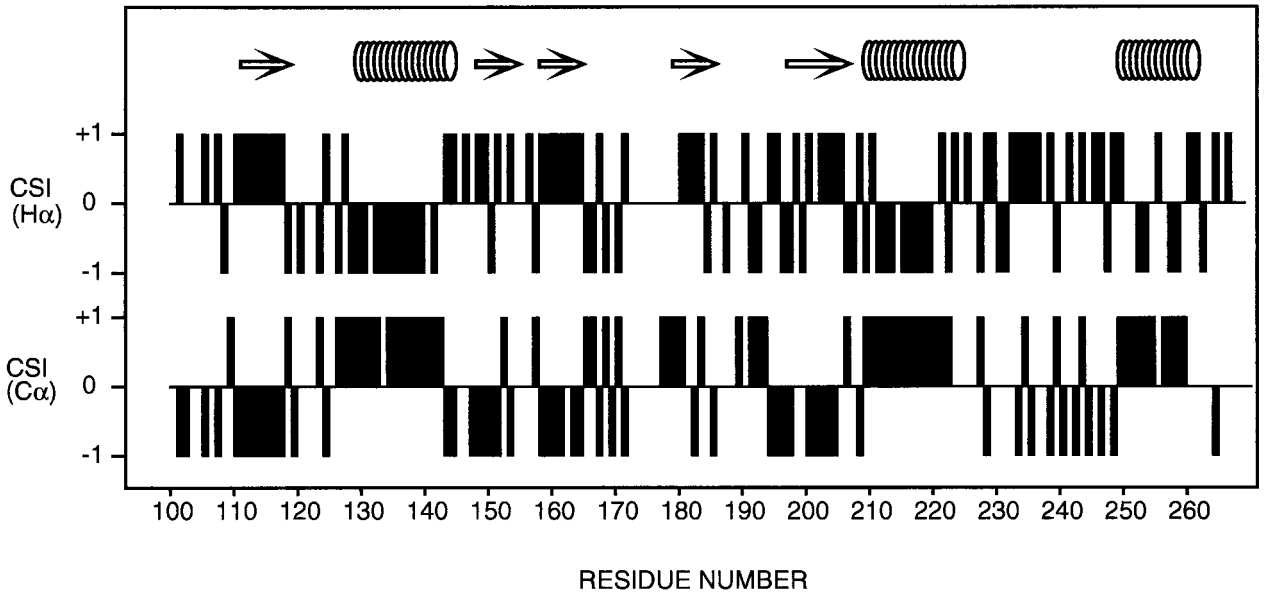

Fig. 7. Chemical shift index (Wishart and Sykes, 1994) as a function of residue number for H^{α} and C^{α} chemical shifts, correlated with the elements of secondary structure.

TABLE 2
SUMMARY OF STRUCTURAL STATISTICS

Rmsd from experimental distance restraints (Å)		
Total distance restraints (965)		
Intraresidue (310)	0.028 ± 0.004	
Sequential (346)	0.053 ± 0.006	
Short range (96)[a]	0.073 ± 0.008	
Long range (104)[a]	0.074 ± 0.021	
Hydrogen bonds (102)[b]	0.062 ± 0.013	
Metal, zinc (7)[c]	0.018 ± 0.010	
Number of NOE violations > 0.5 Å	0.5 ± 0.60	
Average violation	0.6 ± 0.11	
Rmsd from experimental angle restraints (°)		
Dihedral angles (87)	1.32 ± 0.38	
Rmsd from idealized geometry		
Bonds (Å)	0.004 ± 0.0003	
Angles (°)	0.74 ± 0.027	
Impropers (°)	0.46 ± 0.027	
Cartesian rmsd (Å)[d]	<SA> from (SA)	(SA) from crystal structure[e]
Sheet 1 (Leu^{114}–Glu^{119})	1.16	0.79
Helix 1 (Arg^{127}–Asn^{143})	1.09	0.90
Sheet 2 (Thr^{148}–Gly^{155})	2.26	1.26
Sheet 3 (Asp^{158}–Arg^{165})	0.84	0.54
Sheet 4 (Leu^{181}–Ala^{184})	0.95	0.84
Sheet 5 (Gly^{193}–His^{200})	1.03	1.45
Helix 2 (Leu^{212}–Leu^{224})	1.03	0.89
Helix 3 (Ala^{249}–Tyr^{260})	1.49	1.02
Backbone (Arg^{108}–Tyr^{260})	1.85	3.02
All heavy atoms	2.56	–

[a] Short- and long-range NOEs are j < i+5 and j > i+5, respectively.
[b] Hydrogen bond restraints were incorporated as two NOEs per hydrogen bond and were implemented by restraining the O-NH bond to 1.7–2.3 Å and the N-O bond to 2.8–3.3 Å. A total of 102 restraints were used for 51 hydrogen bonds.
[c] Nε2 and Nε1 for His^{168}, His^{183}, His^{196}, His^{218}, His^{222}, and His^{228} to Zn distances were restrained to 2.27 ± 0.2 Å and OD1 from Asp^{170} to Zn was restrained to 2.07 ± 0.2 Å.
[d] <SA> is the ensemble of 26 simulated annealing structures and (SA) is the average structure.
[e] Coordinates for the crystal structure are from John Spurlino, 3D Pharmaceuticals (Exton, PA, U.S.A.).

of the backbone amide protons in helix III persist in D_2O (Fig. 4), which is expected for a well-formed α-helix. Only Asp^{252} and Gly^{255} exchange completely in 5 months time. In fact, Asp^{252}, Gly^{255}, and Ala^{258} are on the side of the helix that may be more accessible to water. Second, in Fig. 7, we show the chemical shift index (CSI; Wishart and Sykes, 1994) for H^{α} and C^{α} chemical shifts which confirms our secondary structure assignments and strongly supports the presence of helix III. Finally, in terms of evolution, helices II and III are very well conserved. Among HFC, HNC, and HFS the identities for helices I, II, and III are 21%, 73%, and 73%, respectively. The lack of NOEs for helix III is perhaps a consequence of the truncation, which might lead to heterogeneity at the C-terminus of the molecule.

The residues that make up the β-sheet region of trHFC and a summary of the characteristic d_{NN} and $d_{\alpha N}$ NOE connections are shown in Fig. 8. With the NOEs and deuterium exchange data, we were able to locate unambiguously the hydrogen bonding network between strands. Like other members of the MMP family, the β-sheet of trHFC is comprised of four parallel strands (in the order of β2, β1, β3, β5) and one antiparallel strand (β4). From Fig. 8, it can be deduced that both the imidazole side chains of His^{183} (β4) and His^{196} (β5) lie on the same side of the sheet. They are well positioned to coordinate the structural zinc, in agreement with the zinc coordination picture depicted for HFS by Gooley et al. (1993,1994) in their NMR studies and for HFC by Spurlino et al. (1994) in their X-ray structure.

In addition to the backbone assignments listed in Table 1, we have made side-chain assignments using 3D NOESY-HSQC, 2D DQF-COSY, and 2D TOCSY spectra. All the methyl resonances were identified, except for Thr^{122} (not assigned in Table 1) and the δ-methyl groups of Ile^{159} and Ile^{256}. Several aromatic spin systems were also identified, including the complete assignment of the three tryptophans. With these additional assignments we were able to calculate a preliminary NMR structure. Figure 9 shows the basic fold of trHFC using MOLSCRIPT (Kraulis, 1991). This structure is the average of 26 lowest energy simulated annealing structures generated by X-PLOR (Brünger, 1992). Starting structures were calculated using a distance geometry algorithm. A total of 965 distance restraints (including 51 hydrogen bonds) and 87 dihedral angle restraints were used in the structure calculation. NOE intensities were classified as strong, medium, and weak corresponding to distance restraints of 1.8–3.2 Å,

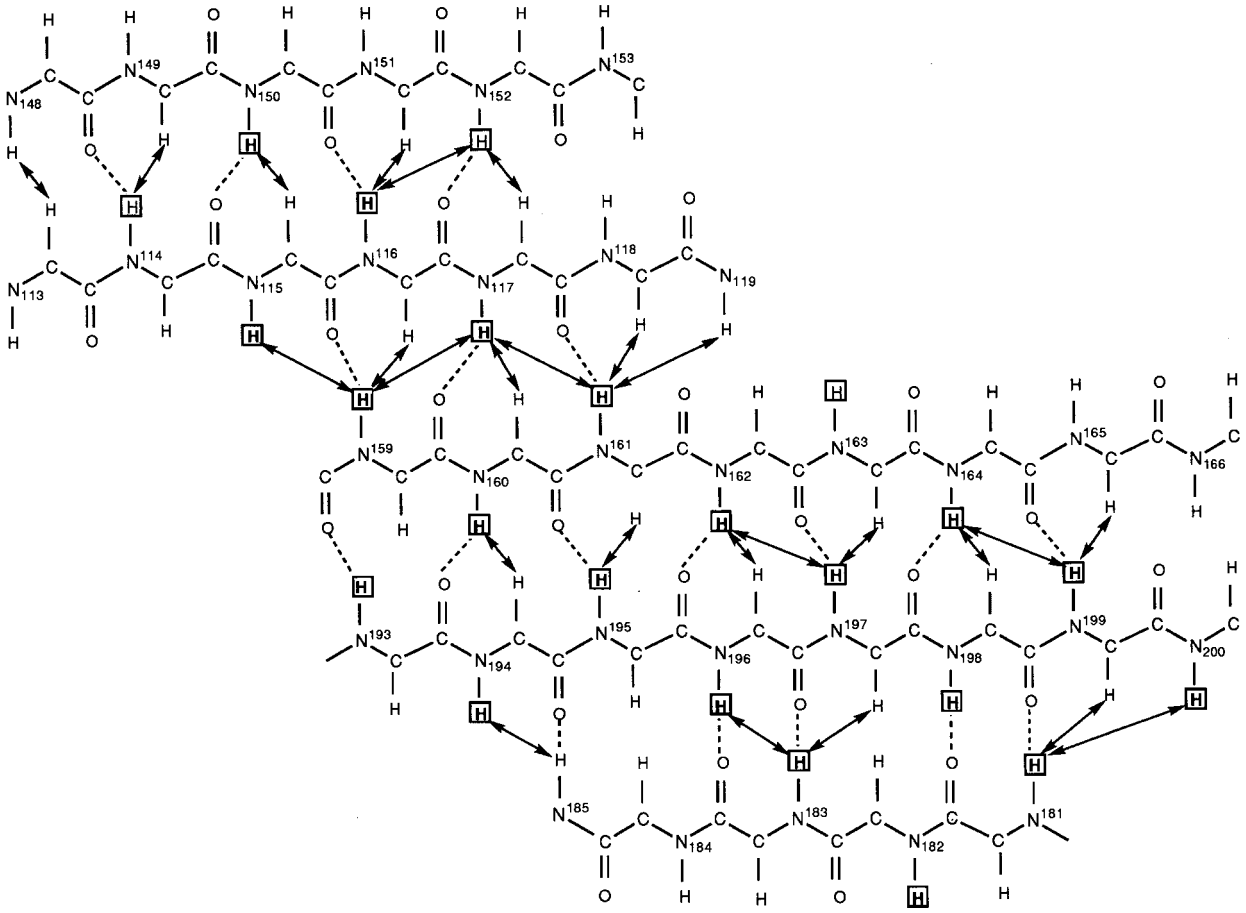

Fig. 8. Summary of NOEs (solid arrows) and hydrogen bonds (dotted lines) across the five-stranded β-sheet of trHFC. Amide protons that are slowly exchanged in D_2O are highlighted in boxes; amide protons that persist after 5 months are boldfaced within boxes.

1.8–4.0 Å, and 1.8–5.0 Å, respectively. Dihedral angle restraints were obtained from the HMQC-J (Kay and Bax, 1990) and HNHA (Vuister and Bax, 1993) experiments. Dihedral angles were restrained to $-140 \pm 60°$ or to $60 \pm 55°$ for $^{3}J_{H^{N}H^{\alpha}}$ measurements of > 8.5 Hz or < 5 Hz, respectively. Hydrogen bond restraints were deduced from hydrogen/deuterium exchange monitored as a function of time with a $^{1}H/^{15}N$-HSQC experiment. Pseudoatom restraints were used where no stereospecific assignment could be made. The average structure was calculated using restrained energy minimization and a repulsive term to simulate the correct van der Waals potential. A summary of relevant structural statistics is presented in Table 2.

In general, most secondary structure elements show an rmsd of 1.0–1.5 Å for the ensemble of structures <SA> relative to the average structure (SA). The precision of the ensemble is related directly to the number of NOEs (particularly long-range) used in the calculation. A more extensive use of carbon-edited NOE data, inclusion of NOEs from trHFC to the inhibitor, inclusion of restraints from the protein to calcium metal and refinement against chemical shift will all help improve the accuracy of the current structure. Nonetheless, the current NMR structure of HFC accurately depicts the fold of the X-ray structure. An overlay of the backbone traces of the X-ray and average NMR structures (Fig. 10) shows, with the exception of the loops, good correspondence. Statistics detailing the quality of the fit between these structures are given in Table 2. As a check on the quality of the stereochemistry of the resulting structures, PROCHECK (Laskowski et al., 1993) was used. The results of the Ramachandran plot are quite similar to those obtained for stromelysin (Van Doren et al., 1995) in that $90 \pm 3\%$ of the backbone dihedral angles are in the allowed region of ϕ,ψ space while only $2 \pm 2\%$ are in the disallowed region. Many of the violations occur at the N- and C-termini and in the less well restrained loops.

The uninhibited trHFC

In Fig. 5, the leucine residues that experience the largest $^{1}H/^{15}N$ shift upon inhibition of the enzyme are located near the active site of the enzyme. Leu^{220} sits between His^{218} and His^{222}, two coordinating histidines to the catalytic zinc; Leu^{235} is at the opposite side of His^{218}. At a concentration of 0.5 mM, the already weak Leu^{181} peak was beyond detection in the uninhibited enzyme. The X-ray structure of trHFC inhibited with the same

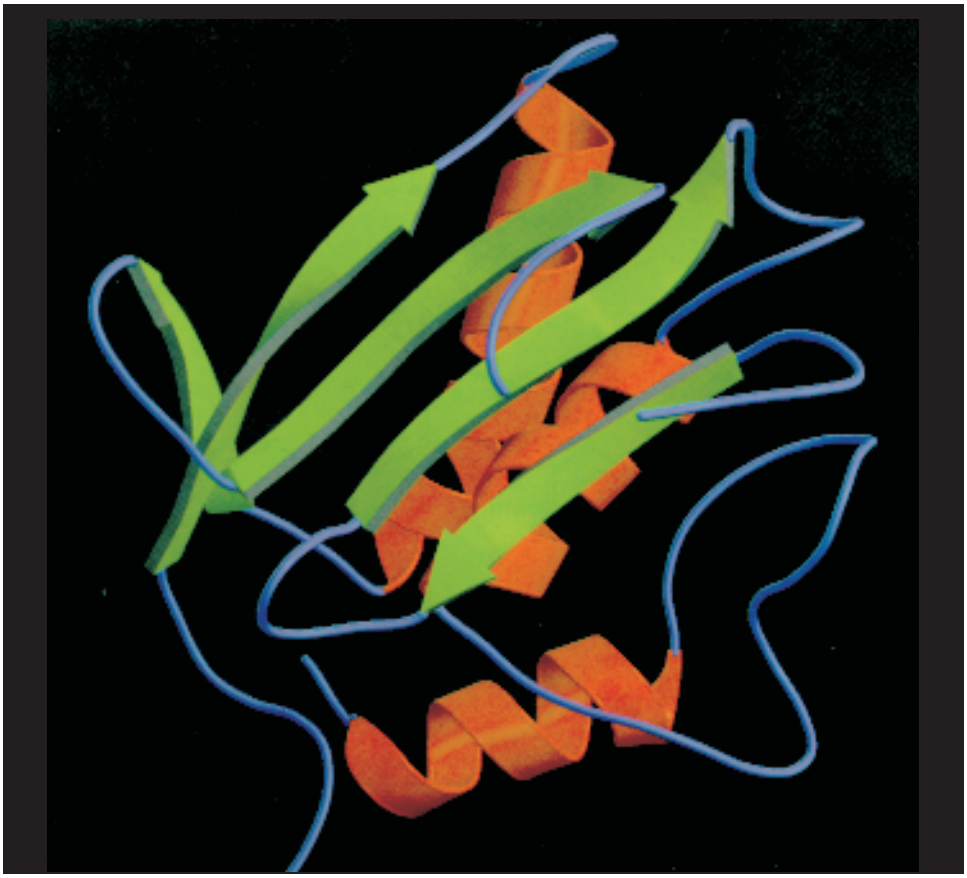

Fig. 9. MOLSCRIPT (Kraulis, 1991) diagram of the average NMR structure (SA) for residues 103–263 of trHFC. Key structural features include a five-stranded β-sheet (four parallel strands, one antiparallel strand) and three α-helices. Statistics for the ensemble of SA structures, <SA>, relative to (SA) are given in Table 2.

Fig. 10. An overlay of the C^{α} trace of the average NMR structure (red) with the crystal structure (in blue, from Spurlino et al., 1994) using GRASP (Nicholls et al., 1991). Statistics for the fit of the two structures are given in Table 2. The overlay is reasonably good for the elements of secondary structure and consistent tertiary packing is observed. Large deviations in the loops are due to a low number of NOE restraints in these regions.

inhibitor (Spurlino et al., 1994) clearly shows a hydrogen bond between the Leu^{181} amide nitrogen and the P_1' carbonyl. Without a hydrogen bond anchor, Leu^{181} may be too mobile to be detected. The HSQC spectra of ^{15}N-Leu trHFC (Fig. 5) reflect the structural perturbation along the active site cleft in the presence and absence of the inhibitor. From the drug-design point of view, it is important to study the uninhibited structure as well as a series of inhibited ones to understand how the enzyme accommodates inhibitors in an 'induced' lock and key fashion. To date, there has been only one X-ray report on an MMP without an inhibitor. In that report, however, Lovejoy et al. (1994b) found trHFC complexed to itself. The N-terminus residues (Leu^{102}–Gly^{105}) occupy the active site cleft of a symmetry-related molecule, which leads to an infinite lattice of trHFC molecules. The two HSQC spectra in Fig. 5 show that the line widths of the NMR peaks are similar, which suggests that both the uninhibited and inhibited forms of the enzyme are monomers of comparable size in solution. With a preliminary demonstration of the stability of the enzyme at a presumably unfavorable pH (5.5) (Van Doren et al., 1993), an NMR investigation of the uninhibited trHFC should be both feasible and useful.

Conclusions

We have reported nearly complete NMR backbone assignments of trHFC by using a combination of double and triple resonance techniques. The sequential assignments were made by matching $^{13}C^{\alpha}$, ^{13}CO, and, when possible, $^{1}H^{\alpha}$ pathways. Residue typing was achieved by employing characteristic chemical shift ranges and specific ^{15}N labeling of leucine residues. NOE restraints derived from these assignments clearly delineate the major secondary structure elements. Three α-helices and one five-stranded β-sheet were identified. These findings are supported by deuterium exchange data and are in full agreement with published structures of the MMP family (Borkakoti et al., 1994; Gooley et al., 1994,1996; Lovejoy et al., 1994a,b; Reinemer et al., 1994; Spurlino et al., 1994; Stams et al., 1994; Van Doren et al., 1995). A preliminary structure calculation reveals that the fold of trHFC is very similar to that determined in the X-ray crystal struc-

ture. In addition, we were able to observe HSQC spectra of ^{15}N-Leu trHFC with and without an inhibitor. The results indicate that active site structural perturbation can be clearly monitored by NMR. Our preliminary data suggest that trHFC is stable in the absence of inhibitors. The NMR investigation of the uninhibited trHFC is therefore feasible and could provide insight on how the enzyme accommodates substrates and inhibitors.

Acknowledgements

We wish to thank Patricia Brake for protein sequencing, and Jamshied Eshraghi and Swapan Chowdhury for mass spectrometric analysis.

References

Becker, J.W., Marcy, A.I., Rokosz, L.L., Axel, M.G., Burbaum, J.J., Fitzgerald, P.M.D., Cameron, P.M., Esser, C.K., Hagmann, W.K., Hermes, J.D. and Springer, J.P. (1995) *Protein Sci.*, **4**, 1966–1976.

Bodenhausen, G. and Ruben, D.J. (1980) *Chem. Phys. Lett.*, **69**, 185–188.

Borkakoti, N., Winkler, F.K., Williams, D.H., D'Arcy, A., Broadhurst, M.J., Brown, P.A., Johnson, W.H. and Murray, E.J. (1994) *Nat. Struct. Biol.*, **1**, 106–110.

Brünger, A.T. (1992) *X-PLOR, A System for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT, U.S.A.

Busiek, D.F., Ross, F.P., McDonnell, S., Murphy, G., Matrisian, L.M. and Welgus, H.G. (1992) *J. Biol. Chem.*, **267**, 9087–9092.

Clubb, R.T., Thanabal, V. and Wagner, G. (1992) *J. Magn. Reson.*, **97**, 213–217.

Docherty, A.J.P. and Murphy, G. (1990) *Ann. Rheum. Dis.*, **49**, 469–479.

Fesik, S.W. and Zuiderweg, E.R.P. (1988) *J. Magn. Reson.*, **78**, 588–593.

Fields, G.B., Van Hart, H.E. and Birkedal-Hansen, H. (1987) *J. Biol. Chem.*, **262**, 6221–6226.

Ghose, A.K., Logan, M.E., Treasurywala, A.M., Wang, H., Wahl, R.C., Tomczuk, B., Gowravaram, M., Jaeger, E.P. and Wendoloski, J.J. (1995) *J. Am. Chem. Soc.*, **117**, 4671–4682.

Goldberg, G.I., Wilhelm, S.M., Kronberger, A., Bauer, E.A., Grant, G.A. and Eisen, A.Z. (1986) *J. Biol. Chem.*, **261**, 6600–6605.

Gooley, P.R., Johnson, B.A., Marcy, A.I., Cuca, G.C., Salowe, S.P., Hagmann, W.K., Esser, C.K. and Springer, J.P. (1993) *Biochemistry*, **32**, 13098–13108.

Gooley, P.R., O'Connell, J.F., Marcy, A.I., Cuca, G.C., Salowe, S.P., Bush, B.L., Hermes, J.D., Esser, C.K., Hagmann, W.K., Springer, J.P. and Johnson, B.A. (1994) *Nat. Struct. Biol.*, **1**, 111–118.

Gooley, P.R., O'Connell, J.F., Marcy, A.I., Cuca, G.C., Axel, M.G., Caldwell, C.G., Hagmann, W.K. and Becker, J.W. (1996) *J. Biomol. NMR*, **7**, 8–28.

Grant, G.A., Eisen, A.Z., Marmer, B.L., Roswit, W.T. and Goldberg, G.I. (1987) *J. Biol. Chem.*, **161**, 5886–5887.

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

Grzesiek, S. and Bax, A. (1992b) *J. Magn. Reson.*, **99**, 201–207.

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

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

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

Kay, L.E., Wittekind, M., McCoy, M.A., Friedrichs, M.S. and Mueller, L. (1992) *J. Magn. Reson.*, **98**, 443–450.

Kraulis, P. (1991) *J. Appl. Crystallogr.*, **24**, 946.

Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993) *J. Appl. Crystallogr.*, **26**, 283–291.

LeMaster, D.M. and Richards, F.M. (1988) *Biochemistry*, **27**, 142–150.

Levitt, M.H. and Freeman, R. (1981) *J. Magn. Reson.*, **43**, 502–507.

Levy, G.C. and Lichter, R.L. (1979) *Nitrogen-15 Nuclear Magnetic Resonance Spectroscopy*, Wiley, New York, NY, U.S.A.

Lovejoy, B., Cleasby, A., Hassell, A.M., Longly, K., Luther, M.A., Weigl, D., McGeehan, G., McElroy, A.B., Drewry, D., Lambert, M.H. and Jordan, S.R. (1994a) *Science*, **263**, 375–377.

Lovejoy, B., Hassell, A.M., Luther, M.A., Weigl, D. and Jordan, S.R. (1994b) *Biochemistry*, **33**, 8207–8217.

Mach, H., Middaugh, C.R. and Lewis, R.V. (1992) *Anal. Biochem.*, **200**, 74–80.

Marion, D., Ikura, M., Tschudin, R. and Bax, A. (1989a) *J. Magn. Reson.*, **85**, 393–399.

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

Nicholls, A., Sharp, K. and Honig, B. (1991) *Protein Struct. Funct. Genet.*, **11**, 281–296.

Reinemer, P., Grams, F., Huber, R., Kleine, T., Schnierer, S., Piper, M., Tschesche, H. and Bode, W. (1994) *FEBS Lett.*, **338**, 227–233.

Shaka., A.J., Keeler, J., Frenkiel, T. and Freeman, R. (1983) *J. Magn. Reson.*, **52**, 335–338.

Shaka, A.J., Lee, C.J. and Pines, A. (1988) *J. Magn. Reson.*, **77**, 274–293..

Spurlino, J.C., Smallwood, A.M., Carlton, D.D., Banks, T.M., Vavra, K.J., Johnson, J.J., Cook, E.R., Falvo, J., Wahl, R.C., Pulvino, T.A., Wendoloski, J.J. and Smith, D.L. (1994) *Proteins Struct. Funct. Genet.*, **19**, 98–109.

Stams, T., Spurlino, J.C., Smith, D.L., Wahl, R.C., Ho, T.F., Qoronfleh, M.W., Banks, T.M. and Rubin, B. (1994) *Nat. Struct. Biol.*, **1**, 119–123.

Van Doren, S.R., Kurochkin, A.V., Ye, Q.-Z., Johnson, L.J., Hupe, D.J. and Zuiderweg, E.R.P. (1993) *Biochemistry*, **32**, 13109–13120.

Van Doren, S.R., Kurochkin, A.V., Hu, W., Ye, Q.-Z., Johnson, L.L., Hupe, D.J. and Zuiderweg, E.R.P. (1995) *Protein Sci.*, **4**, 2487–2498.

Vavra, K., Falvo, J., Banks, T., Schneider, E., Ryerson, C., Sochor, A., Roth, E., Carlton, D., Lirette, R., Dixon, D., Brake, P., Koehn, J., Pulvino, T., Ciccarelli, R. and Wahl, R. (1993) *Protein Sci.*, **2**, . 1), 350-M.

Vuister, G.W. and Bax, A. (1993) *J. Am. Chem. Soc.*, **115**, 7772–7777.

Wang, H., Dellwo, M., Mathiowetz, A.M., Banks, T.M., Pulvino, T.A., Wahl, R.C., Schneider, D.M. and McCoy, M.A. (1994) *Protein Sci.*, **3**, . 1), 398-T.

Wishart, D.S. and Sykes, B.D. (1994) *J. Biomol. NMR*, **4**, 171–180.

Witham, S.E., Murphy, G., Angel, P., Rahmsdorf, H.-J., Smith, B.J., Lyons, A., Harris, T.J.R., Reynolds, T.J., Herrlich, P. and Docherty, A.J.P. (1986) *Biochem. J.*, **240**, 913–916.

Wittekind, M. and Mueller, L. (1993) *J. Magn. Reson.*, **B101**, 201–205.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York, NY, U.S.A.