FEBS Letters 354 (1994) 267–270

FEBS 14749

Specific amino acid substitutions in human collagenase cause decreased autoproteolysis and reveal a requirement for a second zinc atom for catalytic activity

D.H. Williams, E.J. Murray*

Roche Research Centre, PO Box 8, Welwyn Garden City, Herts AL7 3AY, UK

Received 7 September 1994; revised version received 3 October 1994

Abstract We have previously reported the crystal structure of truncated human collagenase (domain II) complexed with a low molecular weight inhibitor. Attempts to crystallise full-length active collagenase (i.e. domain II + III) have been hindered by autoproteolysis at the domain II/III junction at high protein concentrations. To overcome this problem, we have generated an inactive enzyme via a H149→L,D151→N double substitution which displaces the non-catalytic zinc atom, and shown that the altered collagenase is unable to cleave a synthetic substrate. We have also generated an I251→S substitution at the domain II/III junction and demonstrate an increased resistance to proteolysis compared to wild-type collagenase.

Key words: Recombinant collagenase; Site directed mutagenesis; Autoproteolysis; Zinc; Catalysis

1. Introduction

Metalloproteinases (MMPs) play a pivotal role in the degradation of connective tissue matrices which is the basis of tissue remodelling and cellular trafficking. Under pathological conditions, MMPs are instrumental in tumour invasiveness during metastasis and cartilage destruction during arthritis. They represent members of a gene family which includes fibroblast collagenase [1], neutrophil collagenase [2], gelatinases [3,4], stromelysins I, II, III [5–7], and pump-1 [5]. Sequence comparison and functional analysis of the MMPs has revealed the presence of individual domains within the protein which give the MMPs characteristic properties. Domain I (1–82 aa) includes the signal sequence and pro-peptide which contains the conserved sequence PRCGVPD. This sequence maintains latency of the secreted pro-enzyme via a 'cysteine switch' [8,9]. Domain II (83–250 aa) includes the active site [10] and the catalytic Zn^{2+} binding region (VAAHELGHSLG). The C-terminal domain III, homologous to hemopexin, has been less-well characterised, but may function as part of the binding site for natural inhibitors such as TIMP [5,10]. Furthermore, it has been proposed that domain III imposes substrate specificity upon the active site in domain II [8]. The key to understanding the basis of different substrate specificities of the MMP family members lies in the detailed knowledge of the interaction between domains II and III. We have therefore attempted to crystallise full-length active human collagenase (domain II+III) but have been hindered by autoproteolysis at the domain II/III junction. To avoid this problem, we have altered the likely target peptide bond in this region.

We have also investigated the function of the second non-catalytic zinc atom [11], which has been confirmed in the crystal structure of collagenase domain II [12,13]. Analysis of the structure revealed the chelating groups to be H149, D151, H164 and H177. Thus we have substituted H149 to leucine (H149→L) and D151 to asparagine (D151→N) to generate a double-substituted collagenase designed to be depleted of the second zinc.

*Corresponding author. Fax: (44) (707) 373 504.

2. Materials and methods

2.1. Generation of wild-type and mutated collagenase cDNA expression constructs

A 1.4 kb partial *Xba*I fragment of collagenase cDNA (bp 136–1567, see [1]) was blunted by end-filling using Klenow polymerase and cloned into a *Sma*I site previously incorporated in the expression vector pDS56RBSII-6HIS [14]. This construct is designated pHCOL, and encodes procollagenase. The collagenase domain II was generated via PCR using the oligonucleotides 5'-CTCACAGGGATCCAGCGACTCTAGAAACACAA-3' and 5'-GTTTGTGAAGCTTAGGGCTGGACAGGATTTTGGGA-3' as 5' and 3' primers, respectively, upon pHCOL DNA as a template, and the resulting domain II-encoding DNA was cloned into the *Bam*H1 site of pDS56RBSII-6HIS. This construct was designated pTCOL. PCR conditions using Vent polymerase were as recommended by the supplier (New England Biolabs). Site-directed mutagenesis was performed on M13mp18-pHCOL single-stranded DNA using Muta-Gene kit (Stratagene) using mismatched primers.

2.2. Isolation and purification of N-terminal hexahistidine-tagged proteins

200 ml cultures of pHCOL or pTCOL in JM103 were grown up overnight at 37°C in M9 + 50 μg ampicillin + 1 μg thiamine. Each culture was used to inoculate 800 ml Luria broth + 50 μg ampicillin. IPTG was added at a broth optical density $OD_{580\,nm}$ of 0.6–0.8 to a concentration of 2 mM, and bacteria were harvested 4 h later by centrifugation. The bacteria were lysed and proteins denatured by resuspension and stirring for 1 h in 6 M guanidine hydrochloride, 0.1 M NaH_2PO_4, pH 8.0. The cleared lysate was loaded onto a Ni^{2+}-NTA agarose column. Washing was performed using 6 M guanidine hydrochloride, 0.1 M NaH_2PO_4, pH 8.0, then 8 M urea, 0.1 M NaH_2PO_4, 0.01 M Tris-HCl, pH 7.5, to remove non-specifically bound protein. Renaturation of protein was performed on the column using 2 M urea, 0.1 M NaH_2PO_4, 0.01 M Tris-HCl, pH 7.5, followed by 0.2 M NaCl, 50 mM Tris-HCl, 2 mM $CaCl_2$, pH 7.4. Bound protein was eluted from the column using 0.5 M imidazole, 0.2 M NaCl, 50 mM Tris-HCl, 2 mM $CaCl_2$, pH 7.4. The purified protein was then dialysed extensively against 0.2 M NaCl, 50 mM Tris-HCl, 2 mM $CaCl_2$, pH 7.4.

2.3. Isolation of activated recombinant collagenase

Purified non-activated TCOL (domains I + II: aa 20–270 of collagenase as in [1], plus N-terminal hexahistidine tag) or HCOL (domains I + II + III: aa 20 to C-terminus of collagenase as in [1], plus N-terminal hexahistidine tag), was treated with 40 μg/ml trypsin for 50 min at 37° C. This protein solution was then passed twice down a soya bean trypsin inhibitor agarose column to remove contaminating trypsin. The protein was then precipitated at 4°C at 80% ammonium sulphate and dissolved

SSDI 0014-5793(94)01136-2

in 0.2 M NaCl, 50 mM Tris-HCl, 2 mM $CaCl_2$, pH 7.4, to a concentration of approximately 1 mg/ml. This solution was then dialysed against the same buffer. Further protein concentration was performed using 10 kDa cut-off ultraconcentrators (Amicon).

2.4. Synthetic hexapeptide assay

This is a modified version of the assay described in [15] using the substrate Ac-Pro-Gln-Gly-Leu-Leu-Gly-OEt. Briefly, 0.5 mM substrate was incubated with purified enzyme at 37°C for between 2 and 16 h in 50 mM boric acid (pH 7.4), 1 mM $CaCl_2$. A further 37°C incubation was performed, for 1 h, after the addition of 1% picrylsulphonic acid in 2% w/v sodium bicarbonate dissolved in aqueous 50% methanol. 1 M HCl was added and spectrophotometric readings taken 20 min later at 335 nm to monitor the amount of trinitrophenyl peptide formed.

2.5. Collagen fibril assay

^{14}C-Labelled collagen was dried to a film, in 96-well microtitre plates, from thermally reconstituted fibrils [16]. 200 μl samples of purified enzyme in 0.2 M NaCl, 50 mM Tris-HCl (pH 7.4), 2 mM $CaCl_2$, were incubated in the microtitre wells for 8 h at 37°C. The amount of ^{14}C in the supernatant, representing soluble collagen degradation products, was measured by scintillation counting. An excess of bacterial collagenase was used to determine the maximum amount of degradation under these conditions.

2.6. Mass spectrometry analysis

HCOL (10 mg/ml) and HCOL I251→S (17 mg/ml) were mixed with an equal volume of sinapinic acid and analysed on a FinneganMat MALDI.

3. Results and discussion

3.1. Recombinant procollagenase undergoes autoproteolysis

Human procollagenase cDNA was sub-cloned into the hexahistidine-tagged expression vector pDS56RBSII-6HIS. Subsequent expression and purification of the recombinant collagenase from *E. coli* produced variable amounts of a 52 kDa protein (consistent with the size of procollagenase), and an abundant 46 kDa protein, consistent with the size of collagenase activated by autoproteolysis (Fig. 1A, lane 2). When these purified preparations of human collagenase are stored at 4°C at >0.1 mg/ml for a few weeks, further autoproteolysis occurs. The derived products represent domain III (approx 26 kDa) and domain II (approx 20 kDa) (Fig. 1B, compare lanes 4 and 6). Previous reports using N-terminal amino acid sequencing have revealed that P250–I251 is the likely cleavage site in activated collagenase [10], although this has yet to be confirmed by site-directed mutagenesis. In addition, previous reports have also shown that a collagenase(domain II)–stromelysin 2(domain III) chimera shows limited autoproteolysis [17]. Inspection of the domain II–III junction sequences shows that stromelysin 2 contains a serine at the corresponding position of I251 in collagenase (Fig. 2a). We therefore generated an I251→S substitution in recombinant collagenase to investigate whether this would increase resistance to proteolysis and facilitate the purification of full-length activated enzyme. It was found that, as expected, the substitution does not alter the collagenolytic properties of the enzyme (data not shown). After prolonged storage at 4°C, the I251→S-substituted collagenase was more resistant to proteolysis than the wild-type enzyme (Fig. 1B, compare lanes 4 and 5). We investigated the limited degradation which had taken place in the I251→S variant to check whether a shift to an alternative cleavage site had occurred to generate similar sized products. Mass spectrometry measurements revealed predominant degradation products of

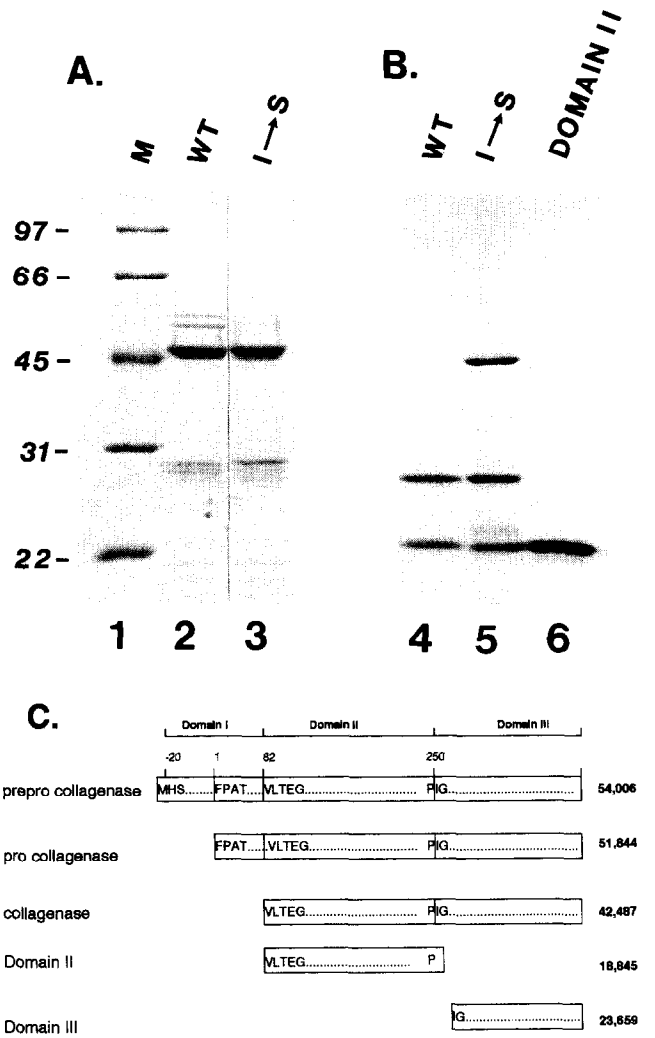

Fig. 1. (A) SDS-PAGE gel stained with Coomassie blue showing freshly purified wild-type (lane 2) and I251→S mutant (lane 3) recombinant collagenase. (B) Similar gel showing wild-type (lane 4) and I251→S mutant (lane 5) collagenase after storage at 4°C compared with truncated (83–250 aa) collagenase which constitutes domain II (lane 6). (C) Scheme showing molecular weights of collagenase domains as calculated on Wisconsin GCG software.

23,690 ± 47 and 18,826 ± 38 Da for wild-type collagenase, and 23,671 ± 47 and 18,796 ± 38 Da for the I251→S variant (data not shown). The similarity in these molecular weights suggests that the I251→S substitution has not induced a shift to an alternative cleavage site, therefore the collagenase active site can accommodate a serine at P1′ instead of isoleucine, albeit with lower efficiency. We note that only human fibroblast collagenase has the target P250–I251 peptide bond at the domain II–III junction (Fig. 2a). It would be interesting to investigate whether such autoproteolysis occurs in vivo as a protective autoregulatory mechanism to avoid excessive collagen fibril damage in pathological conditions. Thus, over-production of fibroblast collagenase would result in a separation of domain II from domain III and therefore loss of collagen specificity. This would sustain collagen fibril levels and change the specificity of the activated truncated collagenases toward denatured collagenous fragments [10]. Clearly, the proposed autoregula-

a

neutro. coll	IQ.....	..PTGPSTP.	.....KP	CDPSLTFDAI TTLRGEILFF
fibro. coll	VQ.....	..P**I**GPQTP.	.....KA	CDSKLTFDAI TTIRGEVMFF
stromelysin 1	PETPLVPTE	PVPPEPGTP.	.....AN	CDPALSFDAV STLRGEILIF
stromelysin 2	TEEPLVPTK	SVPSGSEMP.	.....AK	CDPALSFDAI STLRGEYLFF
rat coll	PN.....	..PKHPKTP.	.....EK	CDPALSLDAI TSLRGETMIF
72kd gelatinase	...LGTGPT.	.PTLGPVTP.	.....EI	CKQDIVFDGI AQIRGEIFFF
95kd gelatinase	TGPPSAGPTG	PPTAGPSTAT	TVPLSPVDDA	CNVNI.FDAI AEIGNQLYLF
Consensus	----------	--P-G---P-	----------	CD--L-FDAI -T-RGE-L-F

b

neutro. coll	TRISQGEADI	NIAFYQRD**H**G	**D**NSPFDGPNG	ILA**H**AFQPGQ GIGGDA**H**FDA
fibro. coll	TKVSEGQADI	MISFVRGD**H**R	**D**NSPFDGPGG	NLA**H**AFQPGP GIGGDA**H**FDE
stromelysin 1	SRLYEGEADI	MISFAVRE**H**G	**D**FYPFDGPGN	VLA**H**AYAPGP GINGDA**H**FDD
stromelysin 2	SRLYEGEADI	MISFAVKE**H**G	**D**FYSFDGPGH	SLA**H**AYPPGP GLYGDI**H**FDD
rat coll	TRIHDGTADI	MISFGTKE**H**G	**D**FYPFDGPSG	LLA**H**AFPPGP NLGGDA**H**FDD
stromelysin 3	TEVHEGRADI	MIDFARYW**D**G	**D**DLPFDGPGG	ILA**H**AFFPKT HREGDV**H**FDY
72kd gelatinase	SRIHDGEADI	MINFGRWE**H**G	**D**GYPFDGKDG	LLA**H**AFAPGT GVGGDS**H**FDD
95kd gelatinase	TRVYSRDADI	VIQFGVAE**H**G	**D**GYPFDGKDG	LLA**H**AFPPGP GIQGDA**H**FDD
Consensus	----EG-ADI	MI-F---E**H**G	**D**--PFDGP--	-LA**H**AF-PG- ---GD-**H**FD-

Fig. 2. Sequence homologies between MMPs using PILEUP on Wisconsin GCG software showing sequence comparison of (a) domain II and III junction, and (b) chelating residues for non-catalytic zinc atom. The relevant amino acids discussed in the text are highlighted in bold-type.

tory mechanism would act in concert with trans-regulatory protease systems, utilising stromelysin and uPA, which also exert fine control over active collagenolytic activity.

3.2. *Effect of H149→L,D151→N site-directed mutagenesis*

Crystallographic studies have revealed a second zinc in the collagenase molecule. The sequence comparison in Fig. 2b shows that the residues involved in the coordination complex are highly conserved throughout the MMP family. We have introduced a H149→L,D151→N double substitution in HCOL and thus have presumably depleted collagenase specifically of the second zinc. Fig. 3 shows that this results in a dramatic loss of proteolytic activity compared to I251→S HCOL. This inactivity is not due to collapse of secondary structure since the CD spectra of the inactive and active variants demonstrates a fairly small difference, with a slight apparent increase in the α-helix content of the inactive form (Fig. 4). Therefore, the inactivity is due to relatively small perturbations in secondary structure, mediated by the displaced H149 and D151 chelating residues. This suggests that the second zinc atom is not fundamentally important for the maintenance of the tertiary structure of collagenase. The four chelating residues are nevertheless strongly conserved across the MMP family, and our observation suggests that they are crucial for catalytic activity. Thus we would hypothesise that the second zinc atom plays a major role in defining the local environment of the active site. We cannot, as yet, define the molecular details of

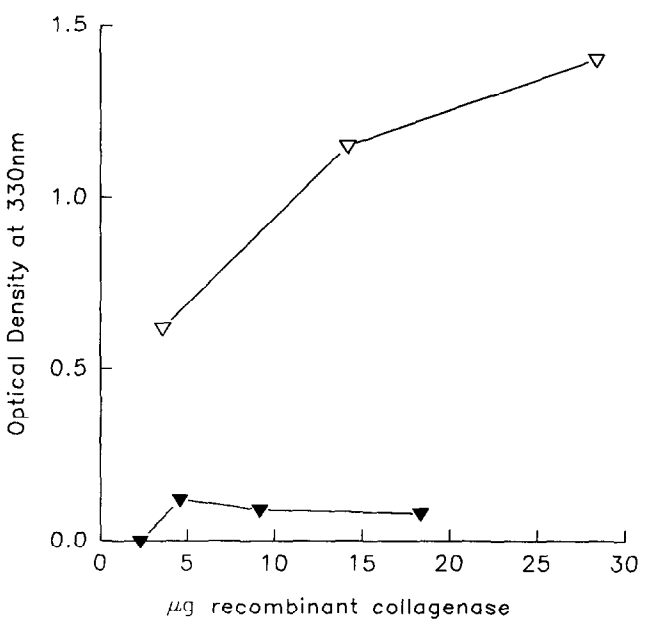

Fig. 3. Comparison of proteolytic activities of I251→S (▽) and H149→L,D151→N mutant (▼) collagenases using a hexapeptide substrate.

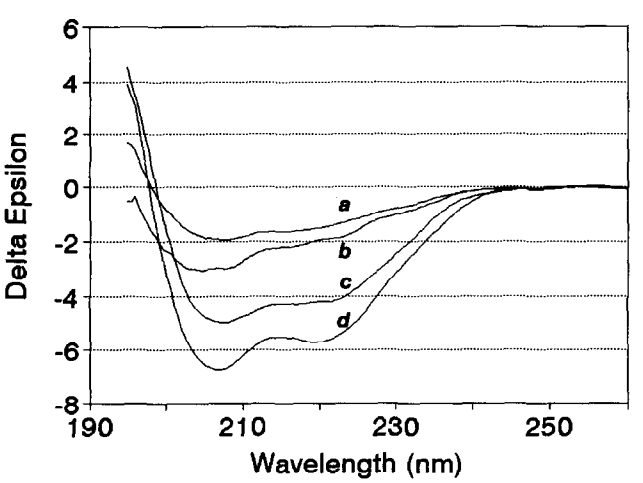

Fig. 4. CD spectra of I251→S collagenase in water (a) and 50% trifluroethanol (c), and H149→L,D151→N collagenase in either water (b) or 50% trifluroethanol (d) using a Jasco J-600 spectrophotometer.

this role, but we would speculate that the second zinc atom is instrumental in fixing the position of an exposed loop (aa 146–162) which may have medium-range effects (approximately 12 Å) upon the position of the catalytic helix.

Acknowledgements: We are grateful to Dr. Steve Martin (NIMR, Mill Hill) for CD expertise, Dr. Raj Handa (Roche) for oligonucleotide synthesis and Kevin Bottomly (Roche) for technical assistance, and Gavin Towerzey for help with the graphics.

References

[1] 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.
[2] Mallaya, S.K., Mookhtiar, K.A., Gao, Y., Brew, K., Diozegi, M., Birkedal-Hansen, H. and Van Wart, H.E. (1990) Biochemistry 29, 10628–10634.
[3] Goldberg, G.I., Marmer, B.L., Grant, G.A., Eisen, A.Z., Wilhelm, S.M. and He, C. (1989) Proc. Natl. Acad. Sci. USA 86, 8207–8211.
[4] Wilhelm, S.M., Collier, I.E., Eisen, A.Z., Grant, G.A. and Goldberg, G.I. (1989) J. Biol. Chem. 264, 17213–17221.
[5] Muller, D., Quantin, B., Gesnel, M.-C., Millon-Collard, R., Abecassis, J. and Breathnach, R. (1988) Biochem. J. 253, 187–192.
[6] Whitham, S.E., Murphy, G., Angel, P., Rhamsdorf, H.-J., Smith, B.J., Lyons, A., Harris, T.J.R., Reynolds, J.J., Herrlich, P. and Docherty, A.J.P (1986) Biochem. J. 240, 913–916.
[7] Basset, P., Bellocq, J.P., Wolf, C., Stoll, I., Hutin, P., Limacher, J.M., Podhajcer, O.L., Chenard, M.P., Rio, M.C. and Chambon, P. (1990) Nature 348, 699–704.
[8] Windsor, L.J., Birkedal-Hansen, H., Birkedal-Hansen, B., and Engler, J.A. (1991) Biochemistry 30, 641–647.
[9] Van Wart, H.E. and Birkedal-Hansen, H. (1990) Proc. Natl. Acad. Sci. USA 87, 5578–5582.
[10] Clark, I.M. and Cawston, T.E. (1989) Biochem. J. 263, 201–206.
[11] Salowe, S.P., Marcy, A.I., Cuca, G.C., Smith, C.K., Kopka, I.E., Hagmann, W.K. and Hermes, J.D. (1992) Biochemistry 31, 4535–4540.
[12] 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) Struct. Biol. 1, 106–110.
[13] Lovejoy, B., Cleasby, A., Hassell, A.M., Longley, K., Luther, M., Wiegle, D., McGeehan, G., McElroy, A.B., Drewery, D., Lambert, M.H. and Jordan, S.P. (1994) Science 263, 375–377.
[14] Stuber, D., Matile, H., and Garotta, G. (1990) Immunological Methods, vol. IV (Lefkovits and Pernis, eds).
[15] Johnson, W.H., Roberts, N.A. and Borkakoti, N. (1987) J. Enzyme Inhibit. 2, 1–22.
[16] Johnson-Wint, B. (1980) Anal. Biochem. 104, 175–181.
[17] Sanchez-Lopez, R., Alexander, C.M., Behrendtsen, O., Breathnach, R. and Werb, Z. (1993) J. Biol. Chem. 268, 7238–7247.