Dysfunctional C $\bar{\text{I}}$ -inhibitor(At), isolated from a type II hereditary-angio-oedema plasma, contains a P1 'reactive centre' (Arg⁴⁴⁴ → His) mutation

K. S. AULAK,* P. A. PEMBERTON,† F. S. ROSEN,‡ R. W. CARRELL,† P. J. LACHMANN* and R. A. HARRISON*§

*MIP Unit, MRC Centre, Hills Road, Cambridge CB2 2QH, U.K., †Department of Haematological Medicine, University Medical Schools, Hills Road, Cambridge CB2 2QH, U.K., and ‡The Center for Blood Research, 800 Huntington Avenue, Boston, MA 02115, U.S.A.

Simple rapid procedures for identification and analysis of dysfunctional C $\bar{\text{I}}$ -inhibitor proteins mutated at the reactive-centre P1 residue have been developed and used to define structurally a C $\bar{\text{I}}$ -inhibitor protein, C $\bar{\text{I}}$ -inhibitor(At), isolated from an individual with hereditary angio-oedema. The observed mutation, Arg⁴⁴⁴ → His, is compatible with a single base change in the codon used for Arg⁴⁴⁴ in the native protein.

INTRODUCTION

C $\bar{\text{I}}$ -inhibitor (C $\bar{\text{I}}$ -inh) is a plasma serine proteinase inhibitor (serpin) active against proteinases of the complement, coagulation, kinin and fibrinolytic systems. It is the sole plasma inhibitor of C $\bar{\text{I}}$ r and C $\bar{\text{I}}$ s [1], and a significant physiological inhibitor of kallikrein [2,3] and factor XIIa [4]. Genetic deficiency of C $\bar{\text{I}}$ -inh is the biochemical defect that gives rise to the disease hereditary angio-oedema (HAE) [5,6]. Two forms of the disease, types I and II, both inherited in an autosomal dominant fashion, have been described [7,8]. Type I HAE is characterized by both low functional and low antigenic plasma levels of C $\bar{\text{I}}$ -inh, whereas individuals with the latter form of the disease synthesize and secrete into the plasma both normal (functional) and mutant (dysfunctional) inhibitors. In these individuals, the observed lower (approx. 15%) than-expected (50%) level of functional inhibitor, but close to normal antigenic level of inhibitor protein, results from increased consumption of the normal inhibitor and decreased consumption of the mutant antigenically indistinguishable protein [9]. Analysis of mutant proteins has indicated variability in their degree of dysfunction and in the nature of their structural abnormality [10–15].

The crystal structure of one plasma serpin, α_1 -antitrypsin, cleaved at the reactive-centre P1 residue, has been solved at 0.3 nm (3Å) [16]. In this, the P1 and P'1 residues (those immediately N- and C-terminal to the bond split during inhibition) are separated by 6.9 nm (69Å). For these residues to be contiguous in the native protein, an exposed stressed peptide loop must exist. Supporting evidence for this hypothesis is seen in the sensitivity to proteolytic attack at residues surrounding the P1–P'1 site. Cleavage of α_1 -antitrypsin in this region generates an inactive inhibitor with a distinctive increase in antigenic heat-stability [17], and proteolytic sensitivity around the P1–P'1 site combined with an increase in antigenic heat-stability of the cleaved inactive protein is now recognized as a common feature of many serpins,

including C $\bar{\text{I}}$ -inh [17; P. A. Pemberton, R. A. Harrison, P. J. Lachmann & R. W. Carrell, unpublished work]. This suggests that they too possess an exposed stressed peptide loop (the 'bait' region) containing the reactive-centre residues. A major determinant of serpin specificity is the reactive-centre P1 residue [19], and mutations at this site result in severe functional impairment [20–22]. In C $\bar{\text{I}}$ -inh the P1 residue is arginine, and there are no other basic residues within the bait region [23]. We have exploited these properties of C $\bar{\text{I}}$ -inh in devising simple procedures whereby P1 residue mutations can be readily identified, and have used them in the first structural definition of a mutant C $\bar{\text{I}}$ -inh protein.

MATERIALS AND METHODS

Normal (functional) C $\bar{\text{I}}$ -inh was purified from pooled human plasma as described by Harrison [24]. C $\bar{\text{I}}$ -inh(At) was purified, by using the same procedure, from 500 ml of plasma drawn from an individual with type II HAE {plasma donor C.El.; see [25]; before withdrawal of plasma the patient had been receiving Danazol (100 mg/day) therapy, but this was interrupted for the 10 days preceding plasma donation}. During isolation, behaviour of the mutant protein was unaltered from that of the normal, and thus the purified protein will be contaminated with small amounts of normal inhibitor.

Trypsin (bovine pancreas; Tos-Phe-CH₂Cl-treated) was purchased from Sigma, and *Pseudomonas aeruginosa* elastase was kindly given by Dr. K. Morihara, Kyoto Research Laboratories, Tokyo, Japan. Trypsin digestion was performed as follows: 10 μ l of trypsin (0.43 μ M in 50 mM-Tris/HCl/100 mM-NaCl, pH 8.0) was added to 10 μ l of C $\bar{\text{I}}$ -inh (10 μ M in 100 mM-KCl/20 mM-sodium/potassium phosphate, pH 7.0) and digestion continued for 50 min at 37 °C. Digestion was terminated either by injection of the sample on to the h.p.l.c. column or, for samples to be analysed by SDS/polyacrylamide-gel electrophoresis, by addition of SDS sample-preparation buffer [26] and heating to 100 °C for 2 min. The

Abbreviations used: C $\bar{\text{I}}$ -inh, C $\bar{\text{I}}$ inhibitor; C $\bar{\text{I}}$ -inh(At), C $\bar{\text{I}}$ -inhibitor isolated from a type II hereditary-angio-oedema (HAE) plasma; Tos-Phe-CH₂Cl, tosylphenylalanylchloromethane ('TPCK'), TFA, trifluoroacetic acid.

§ To whom correspondence and reprint requests should be sent.

conditions for *Pseudomonas* elastase digestion were as follows: C $\bar{\text{I}}$ -inh (10 μM in 100 mM-KCl/20 mM-phosphate, pH 7.0) and *Pseudomonas aeruginosa* elastase (0.3 μM in 50 mM-Tris/HCl/200 mM-NaCl/5 mM-CaCl $_2$ /50 μM -ZnCl $_2$, pH 7.0) were mixed at an enzyme/inhibitor molar ratio of 1:400 and digestion continued for 2.5 h at 37 °C. The reaction was terminated either by addition of EDTA to a final concentration of 10 mM and loading on to the h.p.l.c. column or by addition of SDS sample preparation buffer and heating to 100 °C for 2 min.

SDS/polyacrylamide-gel electrophoresis was performed by using the buffer and sample-preparation systems of Laemmli [26] and a 10–20% (w/v) polyacrylamide gradient gel. Proteins (approx. 10 μg) were prepared for electrophoresis under non-reducing conditions, and the gel stained with Coomassie Blue. M_r marker proteins for SDS/polyacrylamide-gel electrophoresis were purchased from BDH.

C $\bar{\text{I}}$ -inh and its digestion products were chromatographed on a PLRP-S 300A column (250 mm $\times$ 4.6 mm; Polymer laboratories, Church Stretton, Shropshire, U.K.) fitted to Waters h.p.l.c. instrumentation, using a linear gradient from 5% to 80% (v/v) acetonitrile in 0.1% TFA and a flow rate of 1 ml/min. After injection of the sample, the column was washed with 5% acetonitrile/0.1% TFA for 10 min before initiation of the gradient. The gradient was then developed over 40 min, followed by a further 10 min wash in 80% (v/v) acetonitrile and a return to equilibration conditions.

For *N*-terminal sequence analysis of the *C*-terminal reactive-centre-containing peptides, 5 nmol of C $\bar{\text{I}}$ -inh was digested with *Pseudomonas* elastase at an enzyme/inhibitor ratio of 1:400 and the digest resolved by h.p.l.c. as described above. The *C*-terminal peptide peak was collected, evaporated to dryness, then redissolved in TFA. *N*-Terminal sequence analysis of 600 pmol of peptide was performed by Dr. L. Packman at the Protein Sequencing Facility of the Department of Biochemistry, University of Cambridge, Cambridge, U.K., on an Applied Biosystems gas-phase sequencer.

RESULTS AND DISCUSSION

The critical nature of the P1 residue in determining serpin function and specificity makes it a prime candidate for alteration in at least some of the structurally distinct mutant C $\bar{\text{I}}$ -inh proteins that have been described in individuals with type II HAE. We have therefore explored ways in which the reactive-centre region of these proteins might specifically be analysed. The reactive-centre residues are highly sensitive to proteolytic attack by a number of proteinases [P. A. Pemberton, R. A. Harrison, P. J. Lachmann & R. W. Carrell, unpublished work; 23] and, as the only basic residue in this region is the P1 arginine, we investigated the possibility that tryptic digestion could be limited to the P1–P'1 site. The results of trypsin digestion can be revealed either by SDS/polyacrylamide-gel electrophoresis (Fig. 1), or by h.p.l.c. (Fig. 2c). Under non-reducing conditions the native protein has an apparent molecular mass of 100 kDa (Fig. 1, track 1), and limited tryptic digestion generates a major fragment of apparent size 85 kDa (Fig. 2c, peak 4) and a minor fragment of apparent size 2.5 kDa (Fig. 1, track 3; Fig. 2c, peak 3). *N*-Terminal sequence analysis of the 2.5 kDa peptide confirmed that it was derived from cleavage between the P1 arginine and P'1 serine

Fig. 1. SDS/polyacrylamide-gel analysis of limited tryptic cleavage of C $\bar{\text{I}}$ -inh

Proteins (approx. 10 μg) were prepared and run as described in the Materials and methods section. Track 1, C $\bar{\text{I}}$ -inh (normal); 2, C $\bar{\text{I}}$ -inh(At); 3, C $\bar{\text{I}}$ -inh (normal) treated with trypsin; 4, C $\bar{\text{I}}$ -inh(At) treated with trypsin; 5, C $\bar{\text{I}}$ -inh (normal) treated with *Pseudomonas aeruginosa* elastase; 6, C $\bar{\text{I}}$ -inh(At) treated with *Pseudomonas aeruginosa* elastase. Track M, marker proteins. Markers used were: ovotransferrin (78 kDa), albumin (67 kDa), ovalbumin (43 kDa), carbonic anhydrase (30 kDa), myoglobin (17 kDa) and myoglobin CNBr-cleavage products I + II (14 kDa) I and II [8/6 kDa (unresolved)] and III (2.5 kDa).

residues. A limited number of small peptides derived from the *N*-terminal section of the protein are also generated. These were eluted from the h.p.l.c. column at about 25 min and are not detected at 280 nm.

All point mutations of the codon used for the P1 residue (Arg 444 ; codon CGC) in the normal protein [27–29] would result in loss of trypsin-sensitivity. This is demonstrated for C $\bar{\text{I}}$ -inh(At) in Fig. 1 (track 4) and Fig. 2(d), where only limited digestion (of the co-purifying normal protein) is observed. The slight increase in mobility of the major fragment in track 4 relative to track 2 is due to the loss of small *N*-terminal peptides. It is also evident from track 2 that the C $\bar{\text{I}}$ -inh(At) preparation used contained a small amount of protein of slightly higher mobility than the native protein. This probably reflects heterogeneity in glycosylation, as the doublet is retained in the trypsin-treated protein (track 4) (and cannot therefore be due to *N*-terminal peptide loss).

Pseudomonas aeruginosa elastase cleaves between the

Fig. 2. H.p.l.c. of C $\bar{\text{I}}$ -inh and its digestion products

C $\bar{\text{I}}$ -inh and its digestion products were chromatographed as described in the Materials and methods section. (a), (c) and (e) show the normal inhibitor and (b), (d) and (f) show C $\bar{\text{I}}$ -inh(At). (a, b) Undigested protein; (c, d) protein digested with trypsin; (e, f) proteins digested with *Pseudomonas aeruginosa* elastase. Peak 1, native C $\bar{\text{I}}$ -inh; peaks 2 and 7, C $\bar{\text{I}}$ -inh(At); peaks 3 and 5, C-terminal peptide derived by tryptic digestion of the normal inhibitor; peaks 4 and 6, the major fragment given by tryptic digestion of the normal inhibitor; peaks 8 and 10, C-terminal peptide derived by elastase digestion of both the normal and the At inhibitors; peaks 9 and 11, the major fragment given by *Pseudomonas* elastase digestion of both inhibitors.

P4 and P3 residues of the normal protein [P. A. Pemberton, R. A. Harrison, P. J. Lachmann & R. W. Carrell, unpublished work; 23], again generating major 85 kDa and minor 2.5 kDa fragments (Fig. 1, track 5) as

well as smaller N-terminal-derived fragments (see also Fig. 2e). If loss of trypsin-sensitivity in C $\bar{\text{I}}$ -inh(At) were due solely to a point mutation in the P1 codon, the bait region should remain as a stressed loop with unaltered sensitivity to elastase. This is demonstrated in Fig. 1, track 6, where, as with the normal inhibitor, 85 kDa and 2.5 kDa fragments are observed. H.p.l.c. of the digests also demonstrates cleavage (Fig. 2e and 2f) and allows recovery of the released C-terminal peptide. N-Terminal sequence analysis of this will identify the mutated P1 residue; that determined for the C-terminal peptide released from C $\bar{\text{I}}$ -inh(At) is shown in Fig. 3, the P1 arginine (residue 444) being replaced by histidine. This substitution is compatible with a single base change in the Arg⁴⁴⁴ codon (CGC → CAC). Interestingly, this point mutation does not result in an altered restriction site for any of the currently available enzymes. Loss of trypsin-sensitivity might also be predicted for certain mutations either at the P'1 site (e.g. Thr⁴⁴⁵ → Pro), or indeed, elsewhere in the reactive-centre region. However, such mutations, provided that they did not disrupt folding of the protein and hence specific proteolytic digestion and peptide release, would also be identified by sequence analysis of the *Pseudomonas*-elastase-released C-terminal peptide.

The drastic consequences of a mutation at the P1 site are illustrated by antithrombin Pittsburgh, a natural anti-trypsin mutant which has a substitution of methionine for arginine at the P1 residue [20,30]. Activity against its natural substrate, elastase, was greatly diminished, but that against thrombin, kallikrein and factor XIIa was greatly increased. By analogy, P1 mutations in C $\bar{\text{I}}$ -inh might also generate inhibitors with altered specificity rather than fully dysfunctional proteins. In addition to its clearly defined role within the complement system, C $\bar{\text{I}}$ -inh is an important regulator of kallikrein release. It may also be consumed during excessive plasminogen activation. Possibly because of these diverse activities, there is some controversy as to the pathogenesis of HAE [31,32]. Structural definition and functional analysis of mutant proteins is one method by which the proteinases involved in HAE mediator release can be defined, and the procedures presented here offer rapid and simple methods by which one class of these can be analysed. Using them, we now have preliminary data on six additional dysfunctional C $\bar{\text{I}}$ -inh proteins, isolated from unrelated patients. Of these, three are P1 variants.

In addition, the procedures permit improved interpre-

Proteinase cleavage site	Ps. a. E			Trypsin							
C $\bar{\text{I}}$ -inh (normal) A.A.S.	S	A	I	S	V	A	R	T	L	L	V
cDNA nucleotide sequence	TCC	GCC	ATC	TCT	GTG	GCC	CGC	ACC	CTG	CTG	GTC
I.N.S. of C $\bar{\text{I}}$ -inh(At)	---	---	---	---	---	---	CAC	---	---	---	---
C $\bar{\text{I}}$ -inh(At) A.A.S.				(S)	V	A	H	T	L	L	V
Residue no.	P7	P6	P5	P4	P3	P2	P1	P'1	P'2	P'3	P'4

Fig. 3. Amino acid sequence of the reactive-centre region of C $\bar{\text{I}}$ -inh(At)

Abbreviations used: Ps. a. E, *Pseudomonas aeruginosa* elastase; A.A.S., amino acid (single-letter notation) sequence; I.N.S., inferred nucleotide sequence.

tation of the properties of isolated variant proteins. As patients with the type II form of the disease synthesize both normal and mutant inhibitor proteins, and these co-purify, all purified variant proteins contain a variable (but significant) amount of normal inhibitor, making interpretation of inhibitory profiles complex. The level of this contamination can now be accurately assessed, either by quantification of the amount of protein cleaved by trypsin (and resolved by h.p.l.c.) or by quantification of the relative recoveries of arginine and the mutated P1 residue at the relevant step in sequence analysis. From these we estimate a contamination of C $\bar{\text{I}}$ -inh(At) with the normal protein of 10–15%.

The techniques described here also have potential application, with some modification, in analysis of other dysfunctional serpins. Two requirements must be met. Firstly, the proteinase used for cleavage at the active site must not be inactivated by the inhibitor. Secondly, a proteinase cleaving within the *N*-terminal section of the 'bait' region (i.e. *N*-terminal to the P1 residue), preferably with restricted activity elsewhere in the molecule, must be identified. In some cases (e.g. antithrombin III), reduction of disulphide bonds between the potential *C*-terminal peptide and the rest of the molecule would also be necessary before resolution and sequence analysis.

Identification of a single amino acid substitution in C $\bar{\text{I}}$ -inh(At) is not in itself proof that the altered inhibitory properties arise as a consequence of this change. However, comparative peptide mapping has not indicated any other differences between the At and normal proteins (K. S. Aulak, unpublished work). In addition, neither protein nor cDNA sequence analysis has indicated any residue other than arginine at the P1 position of the normal inhibitor, and the critical nature of this residue strongly supports the Arg $\rightarrow$ His substitution as the cause of dysfunction. This contention can only be confirmed by full sequence analysis of the protein or by functional analysis of an Arg⁴⁴⁴ $\rightarrow$ His construct.

During the course of this work K.S.A. was the recipient of an MRC research studentship, and P.A.P. was supported by funds from the Tobacco Advisory Council and the Wellcome Trust.

REFERENCES

1. Sim, R. B., Reboul, A., Villiers, C. L. & Colomb, M. G. (1979) *FEBS Lett.* **97**, 111–115
2. Schapira, M., Scott, C. F. & Colman, R. W. (1982) *J. Clin. Invest.* **69**, 462–468
3. van der Graff, F., Koedam, J. A. & Bouma, B. N. (1983) *J. Clin. Invest.* **71**, 149–158
4. de Agostini, A., Lijnen, H. R., Pixley, R. A., Colman, R. W. & Schapira, M. (1984) *J. Clin. Invest.* **73**, 1542–1549
5. Landerman, N. S., Webster, M. E., Becker, E. L. & Ratcliffe, H. E. (1962) *J. Allergy* **33**, 330–341
6. Donaldson, V. H. & Evans, R. R. (1963) *Am. J. Med.* **35**, 37–44
7. Osler, W. (1888) *Am. J. Med. Sci.* **95**, 362–367
8. Rosen, F. S., Charache, P., Pensky, J. & Donaldson, V. H. (1965) *Science* **148**, 957–958
9. Lachmann, P. J. & Rosen, F. S. (1984) *Acta. Pathol. Microbiol. Immunol. Scand. Section C Suppl.* **92**, 316–329
10. Rosen, F. S., Alper, C. A., Pensky, J., Klemperer, M. R. & Donaldson, V. H. (1971) *J. Clin. Invest.* **50**, 2143–2149
11. Harpel, P. C., Hugli, T. E. & Cooper, N. R. (1975) *J. Clin. Invest.* **55**, 605–611
12. Laurell, A.-B. & Martensson, U. (1971) *Eur. J. Immunol.* **1**, 146–149
13. Curd, J. G., Yelvington, M., Ziccardi, R. J., Mathison, D. A. & Griffin, J. H. (1981) *Clin. Exp. Immunol.* **45**, 261–270
14. Donaldson, V. H., Harrison, R. A., Rosen, F. S., Bing, D. H., Kindness, G., Canar, J., Wagner, C. J. & Awad, S. (1985) *J. Clin. Invest.* **75**, 124–132
15. Donaldson, V. H., Wagner, C. J., Tsuei, B., Kindness, G., Bing, D. H., Harrison, R. A. & Rosen, F. S. (1987) *Blood* **69**, 1096–1101
16. Loebermann, H., Tokuoka, R., Deisenhofer, J. & Huber, R. (1984) *J. Mol. Biol.* **177**, 531–556
17. Carrell, R. W. & Owen, M. C. (1985) *Nature (London)* **317**, 730–732
18. Reference deleted.
19. Carrell, R. W. & Boswell, D. R. (1986) in *Proteinase Inhibitors* (Barrett, A. & Salvesen, G. S., eds.), chapter 12, Elsevier Scientific Publications, Amsterdam
20. Owen, M. C., Brennan, S. O., Lewis, J. H. & Carrell, R. W. (1983) *N. Engl. J. Med.* **309**, 694–698
21. Erdjument, H., Lane, D. A., Ireland, H., Panico, M., di Marzo, V., Blench, I. & Morris, H. R. (1987) *J. Biol. Chem.* **262**, 13381–13384
22. Owen, M. C. & Beresford, C. H. (1988) *FEBS Lett.* **231**, 311–316
23. Salvesen, G. S., Catanese, J. J., Kress, L. F. & Travis, J. (1985) *J. Biol. Chem.* **260**, 2432–2436
24. Harrison, R. A. (1983) *Biochemistry* **22**, 5001–5007
25. Hadjiyannaki, K. & Lachmann, P. J. (1971) *Clin. Allergy* **1**, 221–223
26. Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685
27. Bock, S. C., Skriver, K., Nielsen, E., Thorgersen, H. C., Wiman, B., Donaldson, V. H., Eddy, R. L., Marrinan, J., Radziejewska, E., Huber, R., Shows, T. B. & Magnusson, S. (1986) *Biochemistry* **25**, 4292–4301
28. Tosi, M., Duponchel, C., Bourgarel, P., Colomb, M. & Meo, T. (1986) *Gene* **42**, 265–272
29. Que, B. G. & Petra, P. H. (1986) *Biochem. Biophys. Res. Commun.* **137**, 620–625
30. Scott, C. F., Carrell, R. W., Glaser, C. B., Kueppers, F., Lewis, J. H. & Colman, R. W. (1986) *J. Clin. Invest.* **77**, 631–634
31. Frank, M. M. (1987) *J. Allergy Clin. Immunol.* **78**, 848–850
32. Cullmann, W. & Opferkuch, W. (1986) *Progr. Allergy* **39**, 311–334