Vol. 137, No. 1, 1986 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
May 29, 1986 Pages 431-436

ISOLATION AND CHARACTERIZATION OF A PARTIAL cDNA CLONE FOR HEPARIN COFACTOR II[1]

Roger C. Inhorn and Douglas M. Tollefsen[2]

Division of Hematology-Oncology, Departments of Internal Medicine and Biological Chemistry, Washington University School of Medicine, St. Louis, Missouri 63110

Received April 12, 1986

Summary: A human fetal liver cDNA library constructed in λgt11 was screened with affinity-purified rabbit antibodies raised against heparin cofactor II. One positive clone was plaque purified and the cDNA insert was completely sequenced. The clone encodes the C-terminal 167 amino acid residues of heparin cofactor II as well as the entire 3'-untranslated region of the message. Proline and leucine were identified in the P_2 and P_1 positions of the protease cleavage site, providing a possible exlanation for the ability of heparin cofactor II to inhibit both thrombin and chymotrypsin-like proteases. The coding sequence is identical to that of the recently published human leuserpin 2 (Ragg (1986) Nucl. Acids Res. 14, 1073). © 1986 Academic Press, Inc.

Heparin cofactor II (HCII)[3] is a 65,600 dalton glycoprotein that is one of several protease inhibitors found in human plasma (1-5). HCII, like antithrombin III (ATIII), inhibits thrombin by formation of an equimolar complex that is stable to sodium dodecyl sulfate polyacrylamide gel electrophoresis. Heparin accelerates the rate of thrombin inactivation by both HCII and ATIII ~1000-fold. In contrast, dermatan sulfate specifically increases the rate of thrombin-HCII complex formation but has no significant effect on ATIII (6).

The plasma protease inhibitors are thought to function as suicide substrates by mimicking the cleavage sites of physiological substrates and trapping the appropriate protease in a stable complex. The amino acids surrounding the cleavage site, particularly the residue in the P_1

[1]A preliminary report of this work was presented at the annual meeting of the American Society of Hematology and was published in abstract form in Blood 66 (Suppl 1), 336a (1985).
[2]To whom correspondence should be addressed.
[3]Abbreviations: HCII, heparin cofactor II; ATIII, antithrombin III; hLS2, human leuserpin 2; dATP[α^{35}S], deoxyadenosine 5'-[[α^{35}S]thio]triphosphate.

0006-291X/86 $1.50
Copyright © 1986 by Academic Press, Inc.
All rights of reproduction in any form reserved.

position, are thought to be critical in conferring protease specificity to the inhibitor (7). ATIII, which contains an arginine in the P_1 position (8), inhibits all of the serine proteases in the intrinsic clotting cascade, as well as plasmin and factor VIIa (9, 10). These proteases preferentially cleave proteins following basic amino acid residues. Thrombin is the only one of the above proteases known to be inhibited by HCII (11); however, HCII also inhibits leukocyte cathepsin G (11) and chymotrypsin (12), which differ from trypsin-like proteases in that they generally cleave following hydrophobic amino acid residues. The ability of HCII to inhibit chymotrypsin-like proteases may be explained in part by the identification of a leucine in the P_1 position of the inhibitor (13).

Based upon limited amino acid sequence information, HCII has been classified as a member of the gene family that includes ATIII, α-1-antitrypsin, and α-1-antichymotrypsin (14-16). Recently, Ragg isolated and sequenced a 2.1 kb cCNA clone termed human leuserpin 2 (hLS2) that represents a new member of this family (17). Ragg isolated the clone using an oligonucleotide probe against a conserved region of the inhibitor family. In the current study we have characterized a 1.2 kb cDNA clone encoding the C-terminus of HCII as well as the entire 3'-untranslated region of the message. The clone was isolated from a λgt11 cDNA library prepared from human fetal liver using affinity-purified polyclonal antibodies against HCII. We report here that HCII and hLS2 are in fact identical, firmly establishing HCII as a member of this family of protease inhibitors. Furthermore, we propose that the P_1 leucine and P_2 proline of HCII together determine the unusual bifunctional protease specificity exhibited by the inhibitor.

Materials and Methods

Materials

Restriction endonucleases, T4 DNA ligase, and the Klenow fragment of *Escherichia coli* DNA polymerase were from Bethesda Research Laboratories. Calf intestine alkaline phosphatase was from Boehringer

Mannheim. Deoxyadenosine 5'-[[α^{35}S]thio]triphosphate (dATP[α-^{35}]S) and Na^{125}I were from Amersham. Deoxynucleotides and dideoxynucleotides were from Pharmacia P-L Biochemicals. Nitrocellulose BA-85 was from Schleicher and Schuell.

Isolation of clone

The human fetal liver λgt11 library was kindly provided by Dr. Savio Woo, Baylor University. Rabbit antiserum against HCII was prepared as described by Jaffe et al. (18) and was affinity purified on an HCII-Sepharose column. The library was plated at a density of 50,000 recombinants per 150-mm plate, and nitrocellulose lifts were hybridized sequentially with affinity-purified antibody and ^{125}I-labeled (~5 x 10^6 cpm/μg) Staphylococcus aureus protein A as described by Young and Davis (19). One positive clone, λHCII.1, was plaque purified and DNA was prepared by a plate lysis method (20,21).

DNA sequencing

The cDNA insert from λHCII.1 was subcloned directly into M13 mp18, and sequencing was performed by the dideoxynucleotide chain termination method (22) using dATP[α^{35}S] and buffer gradient gels (23). 20-base oligonucleotides were synthesized with a DNA synthesizer (Applied Biosystems, Foster City, CA) and used as sequencing primers. The identity of each nucleotide was confirmed by sequencing both strands in their entirety. DNA sequences were analyzed with the computer programs of CompuGene (St. Louis, MO) and the Protein Identification Resources (24).

Results and Discussion

A human liver cDNA library prepared in λgt11 was screened with affinity-purifed rabbit antibodies raised against HCII, and antibody-positive plaques were detected using ^{125}I-labeled protein A. Among approximately 100,000 independent recombinants screened, one positive phage, λHCII.1, was identified and plaque purified. The 1.2 kb cDNA insert was subcloned into M13 mp18 and was completely sequenced by the dideoxy method. λHCII.1 contains a 501-nucleotide open reading frame at its 5'-end followed by a stop codon (TAG) and 659 nucleotides of 3'-noncoding sequence (Fig. 1). The predicted amino acid sequence at the 3'-end of the open reading frame matches the published carboxy terminal sequence of HCII (14) in 34 of 37 residues. A polyadenylation signal of AATAAA was identified 20 nucleotides prior to the 34-nucleotide poly (A) tail.

λHCII.1 is identical in sequence to nucleotides 942-2081 of hLS2 (17) with the exception of nucleotide 1525 in the 3'-untranslated region (nucleotide 584 in Fig. 1), where our analysis reveals a T rather than a

```
1     CTC GCA GCA AAT GAC CAG GAG CTG GAC TGC GAC ATC CTC CAG CTG GAA TAC GTG
1     Leu Ala Ala Asn Asp Gln Glu Leu Asp Cys Asp Ile Leu Gln Leu Glu Tyr Val

55    GGG GGC ATC AGC ATG CTA ATT GTG GTC CCA CAC AAG ATG TCT GGG ATG AAG ACC
19    Gly Gly Ile Ser Met Leu Ile Val Val Pro His Lys Met Ser Gly Met Lys Thr

109   CTC GAA GCG CAA CTG ACA CCC CGG GTG GTG GAG AGA TGG CAA AAA AGC ATG ACA
37    Leu Glu Ala Gln Leu Thr Pro Arg Val Val Glu Arg Trp Gln Lys Ser Met Thr

163   AAC AGA ACT CGA GAA GTG CTT CTG CCG AAA TTC AAG CTG GAG AAG AAC TAC AAT
55    Asn Arg Thr Arg Glu Val Leu Leu Pro Lys Phe Lys Leu Glu Lys Asn Tyr Asn

217   CTA GTG GAG TCC CTG AAG TTG ATG GGG ATC AGG ATG CTG TTT GAC AAA AAT GGC
73    Leu Val Glu Ser Leu Lys Leu Met Gly Ile Arg Met Leu Phe Asp Lys Asn Gly

271   AAC ATG GCA GGC ATC TCA GAC CAA AGG ATC GCC ATC GAC CTG TTC AAG CAC CAA
91    Asn Met Ala Gly Ile Ser Asp Asn Arg Ile Ala Ile Asp Leu Phe Lys His Gln

325   GGC ACG ATC ACA GTG AAC GAG GAA GGC ACC CAA GCC ACC ACT GTG ACC ACG GTG
109   Gly Thr Ile Thr Val Asn Glu Glu Gly Thr Gln Ala Thr Thr Val Thr Thr Val

379   GGG TTC ATG CCG CTG TCC ACC CAA GTC CGC TTC ACT GTC GAC CGC CCC TTT CTT
127   Gly Phe Met Pro Leu Ser Ile Gln Val Arg Phe Thr Val Asp Arg Pro Phe Leu

433   TTC CTC ATC TAC GAG CAC CGC ACC AGC TGC CTG CTC TTC ATG GGA AGA GTG GCC
145   Phe Leu Ile Tyr Glu His Arg Thr Ser Cys Leu Leu Phe Met Gly Arg Val Ala

487   AAC CCC AGC AGG TCC TAG  AGGTGGAGGTCTAGGTGTCTGAAGTGCCTTGGGGGCACCCTCATTT
163   Asn Pro Ser Arg Ser Stop

551   TGTTTCCATTCCAACAACGAGAACAGAGATGTTTTGGTCATCATTTTACGTAGTTTACGCTACCAATCTGA
622   ATTCGAGGCCCATATGAGAGGAGCTTAGAAACGACCAAGAAGAGAGGCTTGTTGGAATCAATTCTGCACAA
693   TAGCCCATGCTGTAAGCTCATAGAAGTCACTGTAACTGTAGTGTGTCTGCTGTTACCTAGAGGGTCTCACC
764   TCCCCACTCTTCACAGCAAACCTGAGCAGCGCGTCCTAAGCACCTCCCGCTCCGGTGACCCCATCCTTGCA
835   CACCTGACTCTGTCACTCAAGCCTTTCTCCACCAGGCCCCTCATCTGAATACCAAGCACAGAAATGAGTGG
906   TGTGACTAATTCCTTACCTCTCCCAAGGAGGGTACACAACTAGCACCATTCTTGATGTCCAGGGAAGAAGC
977   CACCTCAAGACATATGAGGGGTGCCCTGGGCTAATGTTAGGGCTTAATTTTCTCAAAGCCTGACCTTTCAA
1048  ATCCATGATGAATGCCATCAGTCCCTCCTGCTGTTGCCTCCCTGTGACCTGGAGGACAGTGTGTGCCATGT
1119  CTCCCATACTAGAGATAAATAAATGTAGCCACATTTACTGTG (A)34
```

Figure 1 Nucleotide sequence for the cDNA insert of λHCII.1. Nucleotide 1 corresponds to nucleotide 942 in the sequence reported by Ragg (17). The numbering of amino acids is arbitrary, with amino acid 1 representing residue 314 of the mature protein. The arrow represents the site of thrombin cleavage, and the polyadenylation signal of AATAAA is enclosed by a box. Amino acid residues downstream of the thrombin cleavage site have been determined for heparin cofactor II (14); the sequence predicted by λHCII.1 is identical except for three discrepancies whose positions are indicated by underlining.

C. Since our clone was isolated using affinity-purified antibodies that recognize only one protein from either plasma or hepatoma post-culture medium (18), we argue that hLS2 is in fact HCII. This is supported by the perfect match between the predicted amino terminal sequence of hLS2 and the published 24 amino terminal residues of HCII. The three discrepancies between the published HCII protein sequence and that predicted by the cDNA clones probably result from errors in the amino acid sequence determination.

Our sequence analysis provides a possible explanation for the ability of HCII to inhibit both thrombin and chymotrypsin-like proteases. The

Table I. Cleavage sites of physiological thrombin substrates

Substrate	P_2	P_1	P_1'	P_2'	Reference
heparin cofactor II	Pro	Leu	Ser	Thr	this paper
α-1-antitrypsin	Pro	Met	Ser	Ile	25
antithrombin III	Gly	Arg	Ser	Leu	8
prothrombin	Pro	Arg	Ser	Glu	26
"	Pro	Arg	Thr	Phe	26
protein C	Pro	Arg	Leu	Ile	27
factor VIII	Pro	Arg	Ser	Phe	28
"	Pro	Arg	Ser	Phe	28
"	Gln	Arg	Glu	Ile	28
fibrinogen, Aα	Pro	Arg	Val	Val	29
"	Val	Arg	Gly	Pro	29
fibrinogen, Bβ	Ala	Arg	Gly	His	30
factor XIII	Pro	Arg	Gly	Val	31

The frequent occurrence of proline in the P_2 position is emphasized by underlining.

presence of a P_1 leucine (13), which is also found in α-1-antichymotrypsin (16), is consistent with the ability of HCII to inhibit leukocyte cathepsin G and chymotrypsin (11,12). We propose that the P_2 proline confers anti-thrombin activity upon HCII. Proline is frequently found in the P2 position of known physiological thrombin substrates (Table I). In addition, the k_{cat}/K_m values measured for thrombin cleavage of peptide p-nitroanilides containing proline or proline homologs in the P_2 positions are consistently higher than those found with peptides containing other P_2 residues (32). Our prediction that the proline and leucine residues together determine the bifunctional protease specificity of HCII provides a model that can be directly tested using site-specific mutagenesis.

Acknowledgments

We thank Dr. Evan Sadler for invaluable discussions, and John Bell for synthesis of oligonucleotides. This investigation was supported in part by NIH Research Service Award GM 07200, Medical Scientist, NIH grants HL 27589 and HL 01079, and a grant from the Monsanto Company.

References

1. Briginshaw, G.F., and Shanberge, J.N. (1974) Arch. Biochem. Biophys. 161, 683-690.
2. Briginshaw, G.F., and Shanberge, J.N. (1974) Thromb. Res. 4, 463-467.
3. Tollefsen, D.M., and Blank, M.K. (1981) J. Clin. Invest. 68, 589-596.
4. Tollefsen, D.M., Majerus, D.W., and Blank, M.K. (1982) J. Biol. Chem. 257, 2162-2169.
5. Griffith, M.J., Carraway, T., White, G.C., and Dombrose, F.A. (1983) Blood 61, 111-118.
6. Tollefsen, D.M., Pestka, C.A., and Monafo, W.J. (1983) J. Biol. Chem. 258, 6713-6716.
7. Carrell, R., and Travis, J. (1985) Trends Biochem. Sci. 10, 20-24.
8. Chandra, T., Stackhouse, R., Kidd, V.J., and Woo, S.L.C. (1983) Proc. Natl. Acad. Sci. USA 80, 1845-1848.
9. Rosenberg, R.D. (1977) Semin. Hematol. 14, 427-439.
10. Broze, G.J., and Majerus, P.W. (1980) J. Biol. Chem. 255, 1242-1247.
11. Parker, K.A., and Tollefen, D.M. (1985) J. Biol. Chem. 260, 3501-3505.
12. Church, F.C., Noyes, C.M., and Griffith, M.J. (1985) Proc. Natl. Acad. Sci. USA 82, 6431-6434.
13. Griffith, M.J., Noyes, C.M., Tyndall, J.A., and Church, F.C. (1985) Biochemistry 24 6777-6782.
14. Griffith, M.J., Noyes, C.M., and Church, F.C. (1985) J. Biol. Chem. 260, 2218-2225.
15. Hunt, L.T., and Dayhoff, M.O. (1980) Biochem. Biophys. Res. Commun. 95, 864-871.
16. Chandra, T., Stackhouse, R., Kidd, V.J., Robson, J.H., and Woo, S.L.C. (1983) Biochemistry 22, 5055-5061.
17. Ragg, H. (1986) Nucl. Acids Res. 14, 1073-1088.
18. Jaffe, E.A., Armellino, D., and Tollefsen, D.M. (1985) Biochem. Biophys. Res. Commun. 132, 368-374.
19. Young, R.A., and Davis, R.W. (1983) Proc. Natl. Acad. Sci. USA 80, 1194-1198.
20. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.), pp. 371-372.
21. Garber, R.L., Kuroiwa, A., and Gehrung, W.J. (1983) EMBO J. 2, 2027-2036.
22. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
23. Biggin, M.D., Bigson, T.J., and Hong, G.F. (1983) Proc. Natl. Acad. Sci. USA 80, 3963-3965.
24. Dayhoff, M.O., Barker, W.C., and Hunt, L.T. (1983) Methods Enzymol. 65, 499-560.
25. Kurachi, K., Chandra, T., Degen, S.J.F., White, T.T., Marchioro, T.L., Woo, S.L.C., and Davie, E.W. (1981) Proc. Natl. Acad. Sci USA 78, 6826-6830.
26. Degen, S.J.F., MacGillivray, R.T.A., and Davie, E.W. (1983) Biochemistry 22, 2087-2097.
27. Foster, D., and Davie, E.W. (1984) Proc. Natl. Acad. Sci. USA 81, 4766-4770.
28. Vehar, G.A., Keyt, B., Eaton, D. Rodriguez, H., O'Brien,D.P., Rotblat, F., Oppermann, H., Keck, R., Wood, W.I., Harkins, R.N., Tuddenham, E.G.D., Lawn, R.M., and Capon, D.J. (1984) Nature 312, 337-342.
29. Rixon, M.W., Chan, W., Davie, E.W., and Chung, D.W. (1983) Biochemistry 22, 3244-3250.
30. Chung, D.W., Que, B.G., Rixon, M.W., Mace, M., and Davie, E.W. (1983) Biochemistry 22, 3244-3250.
31. Takagi, T., and Doolittle, R.F. (1974) Biochemistry 13, 750-756.
32. Lottenberg, R., Hall, J.A., Blinder, M., Binder, E.P., and Jackson, C.M. (1983) Biochim. Biophys. Acta 742, 539-557.