Volume 236, number 1, 14–18 FEB 06165 August 1988

Efficient production of native, biologically active human cystatin C by *Escherichia coli*

Magnus Abrahamson, Henrik Dalbøge[+], Isleifur Olafsson, Søren Carlsen [+] and Anders Grubb

Department of Clinical Chemistry, University of Lund, Malmö General Hospital, S-214 01 Malmö, Sweden and [+]*Nordisk Gentofte A/S, DK-2820 Gentofte, Denmark*

Received 13 June 1988

A cDNA encoding the mature human cysteine proteinase inhibitor cystatin C was fused to the coding sequence for the *Escherichia coli* outer membrane protein A signal peptide, and the recombinant gene was expressed in *E. coli* under the control of the λ P_R promoter, an optimized Shine-Dalgarno sequence and the λ cI 857 repressor. When induced at 42°C, such cells expressed large amounts of recombinant cystatin C. The recombinant protein was isolated in high yield and characterized. All physicochemical properties investigated, including the positions of disulfide bonds, indicated that the *E. coli* derived cystatin C was identical to cystatin C isolated from human biological fluids, except that the proline residue in position three was not hydroxylated. The recombinant protein displayed full biological activity against papain, cathepsin B and dipeptidyl peptidase I.

Cystatin C; Cysteine proteinase inhibitor; Gene expression; Amyloidosis; Cerebral hemorrhage

1. INTRODUCTION

Cystatin C belongs to a recently defined superfamily of proteins called the cystatin superfamily comprising inhibitors of cysteine proteinases [1,2] and is present in all human biological fluids in concentrations implying a physiological importance of cystatin C as a controlling inhibitor of extracellular cysteine proteinases [3]. Cystatin C is the dominating cysteine proteinase inhibitor in cerebrospinal fluid and accidental chymopapain injection into the cerebrospinal fluid at cysteine proteinase treatment of sciatica results in massive cerebral hemorrhage when the inhibiting capacity of cystatin C is exceeded [4]. In hereditary cystatin C amyloid angiopathy young adults are affected by massive, often fatal, cerebral hemorrhage. This disease is associated with an abnormally low concentration of cystatin C in cerebrospinal fluid [5] and with deposition of a cystatin C variant as amyloid [6,7]. Access to large amounts of native cystatin C is required for further elucidation of the biological activities of the protein as well as for preventive and therapeutic trials at cerebral hemorrhages associated with low cerebrospinal fluid levels of cystatin C. The present work was undertaken to establish an *E. coli* expression system allowing not only the production of large amounts of native cystatin C but also the future production of cystatin C variants enabling an assessment of the biological and clinical importance of various amino acid substitutions in the molecule.

Correspondence address: M. Abrahamson, Department of Clinical Chemistry, University of Lund, Malmö General Hospital, S-214 01 Malmö, Sweden

Abbreviations: E-64, L-3-carboxy-*trans*-2,3-epoxypropionyl-leucylamido-(4-guanidino)butane; Z, carboxybenzoyl; Bz, benzoyl; pNA, *p*-nitroaniline; NHMec, aminomethylcoumarin; OmpA, outer membrane protein A

2. MATERIALS AND METHODS

2.1. *Materials*

Restriction enzymes and other DNA modifying enzymes were from Amersham. Carboxypeptidase Y was purchased from Boehringer. Cathepsin B (EC 3.4.22.1) and dipeptidyl peptidase I (type X) (EC 3.4.14.1), both from bovine spleen, were obtained from Sigma. Papain (EC 3.4.22.2) isolated from crude

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/88/$3.50

Papaya latex by affinity chromatography on Gly-Gly-Tyr(Bzl)-Arg-Sepharose [8], activatable to 90%, was a kind gift from Drs I. Björk and P. Lindahl, Uppsala, Sweden. Enzyme substrates were from Bachem. Q Sepharose and Biogel P-60 were from Pharmacia and BioRad, respectively.

2.2. *Construction of expression vector and conditions for expression*

Plasmid pUC18/C6a [9] which contains a full-length cystatin C cDNA inserted at the *Eco*RI site in the polylinker of pUC18, was cut with *Nco*I/*Hin*dIII. The N-terminal part of the gene was substituted with a synthetic linker reconstituting the coding sequence for mature cystatin C enabling in frame fusion to a DNA fragment encoding the signal peptide from the *E. coli* outer membrane protein A (OmpA). From the resulting plasmid an approx. 720 bp *Cla*I/*Eco*RI DNA fragment containing the OmpA signal peptide fused in frame to the cystatin C gene was isolated. The fragment was introduced into plasmid pHD234 which contains the λ cI 857 temperature sensitive repressor gene, the λ P_R promoter, an optimized Shine-Delgarno sequence [10] and the polylinker from pUC18 resulting in the expression plasmid pHD313.

The plasmid was introduced into *E. coli* MC1061 and propagated in 1 l TB medium containing 100 μg/ml ampicillin. The expression of cystatin C was induced by increasing the temperature from 30°C to 42°C for 3 h at an initial A_{600} of approx. 5. The cells were harvested by centrifugation.

2.3. *Isolation and physicochemical characterization of cystatin C*

Cystatin C from bacterial suspensions or human urine was isolated by anion-exchange and gel chromatographies using columns of Q Sepharose and Biogel P-60 after concentration of the starting material by ultrafiltration, under the conditions described in [3]. Individual references and equipments used for agarose and SDS-polyacrylamide gel electrophoreses, amino acid analysis and automated N-terminal amino acid sequence analysis are given in [3]. Direct sequencing from crude bacterial extracts was accomplished by blotting proteins separated by agarose gel electrophoresis onto a PVDF membrane, staining by Coomassie blue and putting the membrane segment carrying the relevant protein directly into the gas-phase sequenator [11]. Analysis of disulfide-linked peptides by diagonal high-voltage electrophoresis was performed as described earlier [12]. Concentrations of isolated cystatin C were determined spectrophotometrically ($A_{280,1\%}$ = 9.1 [13]) and by amino acid analysis and of cystatin C in crude extracts by single radial immunodiffusion using a polyclonal antiserum and a standard previously described [3]. Total protein was measured by quantitative amino acid analysis.

2.4. *Analysis of biological activity*

Papain and cathepsin B activity was assayed at 37°C in 0.1 M phosphate buffer, pH 6.5, containing 1 mM DTT and 1 mM EDTA. The concentration of active cysteine proteinase in papain and cathepsin B solutions was determined by titration with compound E-64 [14]. Dipeptidyl peptidase activity was also assayed at 37°C and pH 6.5, but in 0.1 M phosphate buffer containing 50 mM NaCl, 2 mM DTT and 1 mM EDTA. Cysteine proteinase inhibitory activity in crude *E. coli* extracts was determined by incubation with papain and assessment of residual activity using Z-Phe-Arg-NHMec. The activities in solutions of isolated expression product or human cystatin C were compared by papain titration using the substrate Bz-Arg-pNA and relating the results to total cystatin C content as measured by amino acid analysis. Apparent equilibrium and dissociation rate constants for interactions between the isolated expression product and papain, cathepsin B and dipeptidyl peptidase I were calculated from continuous rate assays [15] using 10 μM Z-Phe-Arg-NHMec, Z-Phe-Arg-NHMec, and Gly-Arg-NHMec, respectively.

3. RESULTS

In order to produce mature cystatin C in the periplasm of *E. coli*, the 5′-end of a cystatin C cDNA insert, including the part coding for the signal peptide as well as the 13 N-terminal amino acids of mature cystatin C, was excised and replaced by a synthetic linker reconstituting the coding sequence for mature cystatin C and enabling in frame fusion to a DNA encoding the *E. coli* OmpA signal peptide (fig.1). The modified cDNA insert was ligated into an expression vector containing the λ P_R promoter, an optimized Shine-Dalgarno sequence and the λ cI 857 repressor (fig.2). After expression had been induced in *E. coli* transformed with this plasmid, cell extracts obtained by osmotic shock contained large amounts of cystatin C immunoreactive material, in contrast to cell extracts from non-induced bacteria or bacteria not harbouring the expression plasmid. The extracts also displayed papain inhibitory activity. After optimizing the conditions for growth and expression, expression levels corresponding to 10% of the total protein content was obtained. Sequencing of the expression product gave the N-terminal sequence SSPGKPPRLVGGPMDASVE-EEGV--ALDFAVGEYNKASNDMY-S, showing that the OmpA signal peptide-cystatin C fusion protein was correctly processed at the Ala^{-1}-Ser^{1} bond by the *E. coli* signal peptidase (fig.1).

The expression product was purified to homogeneity from concentrated cell extracts by a simple procedure comprising chromatography on Q Sepharose at pH 9.5 and gel chromatography on Biogel P-60 of the non-adsorbed fraction. No significant proteolytic degradation was obtained at the high pH used in the initial phase of the isolation, thereby obviating the need for addition of protease inhibitors. The yield of the isolation procedure was 50–55%.

The physicochemical properties of the isolated

```
          -21
          MetLysLysThrAlaIleAlaIleAlaVal
  5'- ATCGATGAAAAAAACTGCTATCGCTATCGCTGTT
  3'- TAGCTACTTTTTTTGACGATAGCGATAGCGACAA
       ClaI
-11                          -1 ↓ 1
AlaLeuAlaGlyPheAlaThrValAlaGlnAla SerSer
GCTCTGGCTGGTTTCGCTACTGTTGCTCAGGCG TCTTCT
CGAGACCGACCAAAGCGATGACAACGAGTCCGC AGAGAG

                          11
ProGlyLysProProArgLeuValGlyGlyProMet...
CCGGGTAAACCGCCGCGTCTGGTTGGTGGTCCCATGG -
GGCCCATTTGGCGGCGCAGACCAACCACCAGGGTACC -
                                  NcoI
```

Fig.1. DNA and corresponding amino acid sequences of the modified region of the cystatin C cDNA. The linker sequence reconstituting the sequence encoding mature cystatin C was assembled from two oligonucleotides, and ligated to *Nco*I digested cDNA. This construct was then fused in frame to a synthetic sequence encoding the OmpA signal peptide. The site of cleavage by the *E. coli* signal peptidase is shown by an arrow.

expression product was compared to those of cystatin C isolated from human urine by the same procedure. Agarose and SDS-polyacrylamide gel electrophoreses showed that the proteins had the same charge and molecular mass (fig.3). Amino acid analyses gave identical results for the two proteins, within the limits of experimental errors (not shown). These data in conjunction with the 42 identified N-terminal amino acid residues of the expression product strongly indicated that the primary structure of the *E. coli* produced cystatin C was essentially identical to that of cystatin C of

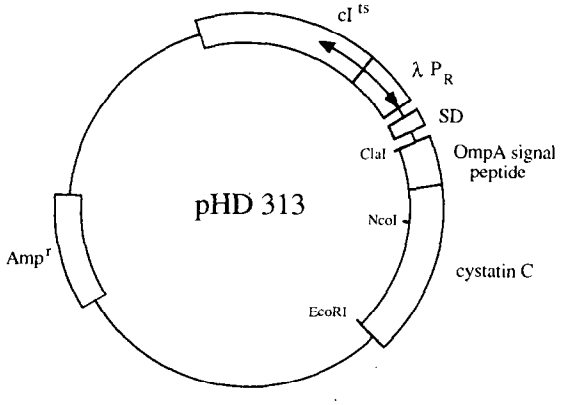

Fig.2. The expression vector pHD313 carrying DNA sequences for an OmpA signal peptide-cystatin C fusion preprotein. Only relevant restriction sites are shown.

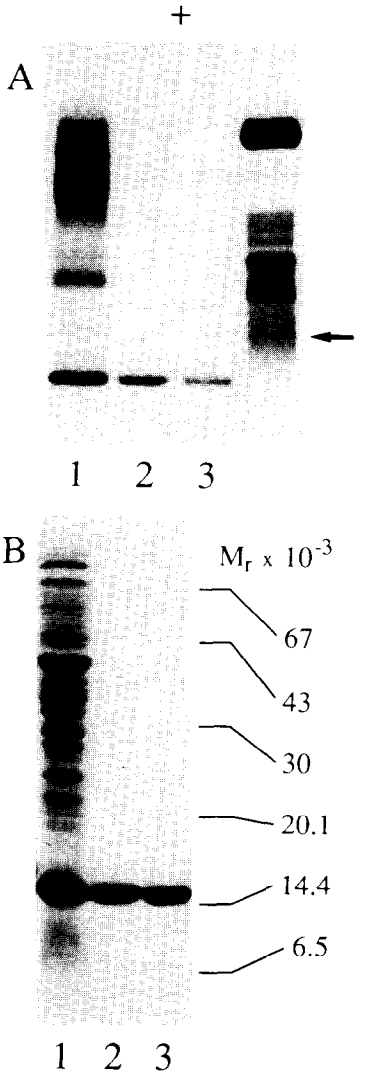

Fig.3. Electrophoretic analyses of (lane 1) extract of pHD313 transformed *E. coli* MC1061 after induction of expression; (2) isolated *E. coli* produced cystatin C; and (3) cystatin C isolated from human urine. (A) Agarose gel electrophoresis at pH 8.6. The anode and point of sample application is shown by a plus sign and an arrow, respectively. (B) SDS-polyacrylamide electrophoresis at reducing conditions. (Far right) The positions of molecular mass markers.

human origin. In addition, amino acid analysis of the disulfide bonded peptides showed that the *E. coli* produced cystatin C had two disulfide bridges spanning the same cysteine residues as the disulfide bridges in human cystatin C, and also gave strong evidence for a correct C-terminal sequence [12].

Table 1

Comparison between inhibitory properties of *E. coli* produced cystatin C and cystatin C from human urine

	$K_{i(app)}$ (nM)	
	E. coli produced	Human origin
Papain	<0.005	<0.005
Cathepsin B	0.50	0.27
DPP I	3.1	6.3

The molar concentration of active cysteine proteinase inhibitor in solutions of isolated expression product and cystatin C from human urine was measured by titration of a papain solution of known active enzyme concentration (see section 2) and compared to the total cystatin C concentration as measured by quantitative amino acid analysis. As estimated this way, the expression product displayed a molar inhibitory activity of 83%, whereas that of cystatin C from human urine only displayed an inhibitory activity of 55%. The apparent equilibrium constants, $K_{i(app)}$, for interactions between *E. coli* produced cystatin C and papain, cathepsin B and dipeptidyl peptidase I at 37°C and pH 6.5 were determined and compared to the corresponding values for cystatin C isolated from human urine (table 1). Also, the rate constant for dissociation of the enzyme-inhibitor complex, k_{-1}, was calculated for the two cystatin C preparations interacting with cathepsin B. The $k_{-1(app)}$ value for *E. coli* produced cystatin C was 0.9×10^{-3} s^{-1}, in close agreement with the value of 0.8×10^{-3} s^{-1} obtained for cystatin C of human origin.

4. DISCUSSION

The presently described *E. coli* expression system allows production of cystatin C indistinguishable from cystatin C of human origin with respect to all tested physicochemical properties, the only exception being that the proline in position three of cystatin C isolated from human biological fluids typically is hydroxylated to 50% [16] whereas the *E. coli* produced material is not. The biological activity of *E. coli* produced cystatin C was very similar to that of the human material, since it displayed tight binding to the model enzyme for cysteine proteinases, papain, and to the well-characterized mammalian lysosomal cysteine proteinase, cathepsin B; it also inhibited the lysosomal mammalian exopeptidase dipeptidyl peptidase I. The K_i and k_{-1} values calculated for those interactions were within experimental errors identical to the corresponding values for cystatin C from human urine reported here and earlier [3,17].

The highest molar inhibitory concentration to date measured in preparations of cystatin C of human origin (measured as inhibition of papain activity against low M_r substrates), amounts to 50–60% of the total protein concentration [3]. The reason for the occurrence of a large proportion of apparently biologically inactive cystatin C in such preparations is unknown and it is therefore of considerable interest that the preparations of recombinant cystatin C described in this work have biological activities amounting to more than 80% of their protein concentrations.

The *E. coli* expression system described in this work for the production of large amounts of recombinant cystatin C, apparently identical to human cystatin C, will probably allow preventive and therapeutic trials aiming at increasing the cysteine proteinase inhibitory capacity of biological fluids in which low levels of cystatins are associated with pathophysiological processes. The system might also be used for site-directed mutagenesis to produce variants of cystatin C for studies on structure-function relationships, with respect to inhibitory properties and, more specifically, to investigate if the replacement of a leucine residue at position 68 of cystatin C with a glutamine residue could be the cause of amyloid formation in hereditary cystatin C amyloid angiopathy [7].

Acknowledgements: The work in Malmö was supported by A. Österlunds Stiftelse, Direktör A. Pålssons Stiftelse, G. och J. Kocks Stiftelse, the Medical Faculty, University of Lund and the Swedish Medical Research Council (project no.05196).

REFERENCES

[1] Barrett, A.J., Fritz, H., Grubb, A., Isemura, S., Järvinen, M., Katunuma, N., Machleidt, W., Müller-Esterl, W., Sasaki, M. and Turk, V. (1986) Biochem. J. 236, 312.

[2] Barrett, A.J. (1987) Trends Biochem. Sci. 12, 193–196.

[3] Abrahamson, M., Barrett, A.J., Salvesen, G. and Grubb, A. (1986) J. Biol. Chem. 261, 11282–11289.
[4] Grubb, A. (1988) Acta Orthop. Scand. 59, 63–65.
[5] Grubb, A., Jensson, O., Gudmundsson, G., Arnason, A., Löfberg, H. and Malm, J. (1984) New Engl. J. Med. 311, 1547–1549.
[6] Cohen, D.H., Feiner, H., Jensson, O. and Frangione, B. (1983) J. Exp. Med. 158, 623–628.
[7] Ghiso, J., Jensson, O. and Frangione, B. (1986) Proc. Natl. Acad. Sci. USA 83, 2974–2978.
[8] Blumberg, S., Schechter, I. and Berger, A. (1970) Eur. J. Biochem. 15, 97–102.
[9] Abrahamson, M., Grubb, A., Olafsson, I. and Lundwall, Å. (1987) FEBS Lett. 216, 229–233.
[10] Dalbøge, H., Carlsen, S., Jensen, E.B., Christensen, T. and Dahl, H.M.H. (1988) DNA, in press.
[11] Matsudaira, P. (1987) J. Biol. Chem. 262, 10035–10038.
[12] Grubb, A., Löfberg, H. and Barrett, A.J. (1984) FEBS Lett. 170, 370–374.
[13] Cejka, J. and Fleischmann, L.E. (1973) Arch. Biochem. Biophys. 157, 168–176.
[14] Barrett, A.J., Kembhavi, A.A., Brown, M.A., Kirschke, H., Knight, C.G., Tamai, M. and Hanada, K. (1982) Biochem. J. 201, 189–198.
[15] Nicklin, M.J.H. and Barrett, A.J. (1984) Biochem. J. 223, 245–253.
[16] Grubb, A. and Löfberg, H. (1982) Proc. Natl. Acad. Sci. USA 79, 3024–3027.
[17] Barrett, A.J., Davies, M.E. and Grubb, A. (1984) Biochem. Biophys. Res. Commun. 120, 631–636.