

# Amino acid sequence elucidation of human acrosin-trypsin inhibitor (HUSI-II) reveals that Kazal-type proteinase inhibitors are structurally related to $\beta$ -subunits of glycoprotein hormones

Edwin Fink<sup>1</sup>, Christa Hehlein-Fink<sup>1</sup> and Manfred Eulitz<sup>2</sup>

<sup>1</sup>Department of Clinical Chemistry and Clinical Biochemistry, University of Munich, Nussbaumstrasse 20, D-8000 Munich 2, FRG and <sup>2</sup>Institute for Clinical Molecular Biology, Gesellschaft für Strahlen- und Umweltforschung, Marchioninistrasse 25, D-8000 Munich 70, FRG

Received 30 July 1990

The amino acid sequence of the acrosin-trypsin inhibitor HUSI-II from human seminal plasma is presented which unequivocally identifies HUSI-II as being of Kazal-type. In addition, the HUSI-II sequence shows a striking similarity to the middle part of glycoprotein hormone  $\beta$ -subunits thus revealing a hitherto unknown structural and evolutionary relationship between Kazal-type inhibitors and glycoprotein hormones.

Kazal-type inhibitor; Glycoprotein hormone  $\beta$ -subunit; Acrosin-trypsin inhibitor (human seminal plasma); Chorionic gonadotropin; Luteinizing hormone

## 1. INTRODUCTION

Kazal-type inhibitors are proteinaceous inhibitors of serine proteinases with amino acid sequences similar to the sequence of bovine pancreatic secretory trypsin inhibitor first described by Kazal et al. [1]. Inhibitors of Kazal-type have been detected in all vertebrates investigated. They are present in eggs of birds and amphibians and in many tissues and body fluids such as blood plasma, seminal plasma, seminal vesicles, pancreas, pancreatic juice, submandibular gland and saliva [2,3].

The only inhibitor of human origin which was proven by sequence comparison as being of Kazal-type is the pancreatic secretory trypsin inhibitor [4]. Another trypsin inhibitor, the trypsin-acrosin inhibitor HUSI-II from human seminal plasma [5-8] has been suspected to be of Kazal-type like the inhibitors found in seminal fluids of boar [9] and bull [10,11]. The final proof, however, the elucidation of the amino acid sequence of HUSI-II, has never been brought forth. We now report the amino acid sequence of HUSI-II which unequivocally identifies this polypeptide as a Kazal-type inhibitor.

*Correspondence address:* E. Fink, Department of Clinical Chemistry and Clinical Biochemistry, University of Munich, Nussbaumstrasse 20, D-8000 Munich 2, FRG

*Abbreviations:* PSTI, pancreatic secretory trypsin inhibitor; hCG, human chorionic gonadotropin; hLH, human luteinizing hormone; hFSH, human follicle stimulating hormone; hTSH, human thyroid stimulating hormone; suffix  $\beta$ ,  $\beta$ -subunit

## 2. EXPERIMENTAL

The acrosin-trypsin inhibitor HUSI-II was isolated from a pool of human seminal plasma basically as described [6-8]. In brief: after centrifugation the cell-free human seminal plasma was incubated with SP-Sephadex C-50 at low ionic strength at pH 5.4. The ion exchanger was washed and eluted in a column with a linear gradient of 0-1.0 M NaCl in 0.05 M ammonium acetate buffer, pH 5.4. Trypsin inhibitors were isolated from the eluate by affinity chromatography on immobilized trypsin and subsequently separated by gel filtration into two fractions, HUSI-I (now designated MPI [12]) and HUSI-II. The HUSI-II preparation was separated by ion exchange chromatography on SP-Sephadex C-25 (sodium phosphate of ionic strength 0.01 and pH 6.1-7.3 in 0.19 mol/l NaCl) into at least 7 peaks of very similar, though not completely identical amino acid compositions. The two prevalent of these chromatographic forms, HUSI-II-5 and -6, were used for sequencing. Further purification of these forms and isolation of peptides obtained after reduction and carboxymethylation or treatment with pyroglutamyl peptidase [13] was achieved by reversed phase HPLC (butyl or octadecyl column, gradient of acetonitrile in 0.1% trifluoroacetic acid). N-terminal pyroglutamic acid was removed from reduced, carboxymethylated inhibitor by treatment with pyroglutamyl peptidase (Boehringer, Mannheim) [13]. The amino acid sequences of purified peptides were determined employing an Applied Biosystems model 470A gas-phase sequencer with an on-line HPLC 120A system for identification of the PTH derivatives. The NBRF sequence data bank (release 15) was searched by using the program MICROGENIE, version 5 (Beckman).

## 3. RESULTS AND DISCUSSION

In HUSI-II-6 the reactive site peptide bond P<sub>1</sub>-P'<sub>1</sub> [14] (positions 23-24 in Fig. 1) is cleaved, therefore, two peptides were separated after reduction and alkylation. Their sequences correspond to positions 6-23 and

	10	20	↓	30	40	50	60
HUSI-II	QFGLFSKYRTPNC	SQYRLPGCPRHF	NPVC	GS	DMSTYANECT	LCMKIREGGH	NIKIIRNGPC
bPSTI	NILGREAKCTNEV	NGCPRIYNP	VC	GT	DGVTYSNE	CLCNENKER	QTPVLIQKSGPC
hPSTI	DSLREAKCYNE	NGCTKIYD	PV	CGT	DGNTPNE	CVLGFENK	KRQTSILLQKSGPC
pPSTI	TSPQREATCTSE	VSGCPKIYN	PV	CGT	DGITYSNE	CVLCS	ENKKRQTPVLIQKSGPC
BUSI-I	EIYFEPDFGF	PPDCKVYTE	ACTREYN	PI	CDSAAK	YSNE	CTFCNEKMNDADLHFQHFGECEY
BUSI-IIA	QGAQVDCAEFK	DPKVYCTRE	SNPH	CGS	NGEYGN	KCAF	CKAVMKSGGKINLKHRRCK
POSI	TRKQPNCNVYR	SHLFFCT	RQMD	PL	CGT	NGKSYAN	PCIFCSEKGLRNQKDFDGHWHCREYTSARS

Fig. 1. Amino acid sequence of HUSI-II and comparison with the sequences of pancreatic secretory trypsin inhibitors and of seminal plasma acrosin-trypsin inhibitors. b-, h-, pPSTI: bovine [22], human [4,23], porcine PSTI [24]; BUSI-I, BUSI-IIA: bull semen inhibitor types I [10] and IIA [11]; POSI: porcine semen inhibitor [9]. The amino acids which are identical to those of the HUSI-II sequence are shaded. The position of the P<sub>1</sub> residue [14] is marked by an arrow.

24–61 in Fig. 1. Amino acid analysis of the inhibitor fraction HUSI-II-5 revealed that it contained 5 additional amino acids. The N-terminal amino acid of this fraction cannot be detected by automatic Edman degradation because it is blocked. After incubation with pyroglutamyl peptidase the sequence of the 11 N-terminal residues (2–12 in Fig. 1) was identified. The blocked N-terminus was obviously a pyroglutamyl residue most probably derived from an original glutaminyll.

Sequence comparisons clearly demonstrated that HUSI-II is a Kazal-type serine proteinase inhibitor (Fig. 1). The sequence of HUSI-II is more similar to the sequences of PSTIs than to those of inhibitors from seminal plasma of different species (Fig. 1 and Table I). An exception is an inhibitor isolated from guinea pig seminal vesicles (paper in preparation; sequence data not shown here), the sequence similarity of 50.8% to this inhibitor is significantly higher than to PSTIs.

A search for proteins with sequence similarities to HUSI-II disclosed a high degree of similarity between HUSI-II and the glycoprotein hormone  $\beta$ -subunits (Fig. 2, Table II). The family of glycoprotein hormones [15] is composed of chorionic gonadotropin, luteinizing hormone, follicle-stimulating hormone, and thyroid-stimulating hormone. Within one species, all glycoprotein hormones share a common  $\alpha$ -subunit, whereas the  $\beta$ -subunits are similar but not identical and determine the biological specificity [15].

The sequence similarity of HUSI-II to glycoprotein hormone  $\beta$ -subunits, especially to hCG $\beta$  and hLH $\beta$  is strikingly high. The other known human Kazal-type inhibitor, hPSTI, shows a lower, but nevertheless clearly recognizable similarity to the  $\beta$ -subunits which is of about the same order of magnitude as that of Kazal-type inhibitors of different species like the 3 domains of chicken ovomucoid (Table II). Presumably due to this lower degree of similarity the structural relationship between Kazal-type inhibitors and glycoprotein hormones had not been detected before.

One sequence segment extending over 7 amino acids (Fig. 2, numbers 17–23 in the top line) is absolutely identical in HUSI-II, hCG $\beta$  and hLH $\beta$ . To our

knowledge an identical segment of this length has never been found before in apparently unrelated proteins.

The conservation of this 17–23 segment during evolution might indicate that this part of the molecule plays a functionally important role both in HUSI-II and in glycoprotein hormones. In fact, this can be taken for granted for HUSI-II since in Kazal-type inhibitors the corresponding segment includes the major part of the region of primary contact between inhibitor and proteinase in the complex (discussed in detail by Laskowski et al. [16]). In addition, the C-terminal Arg of the 17–23 segment (position 23 in Fig. 2) represents the P<sub>1</sub> residue [14] of the reactive inhibitor site of HUSI-II, i. e. the residue primarily responsible for the specificity of the inhibitor. It has been shown that the side chain of the P<sub>1</sub> residue is inserted into the specificity pocket of the proteinases when complexes with the cognate enzymes are formed [17,18]. Though no special function has been assigned to the 17–23 segment of the glycoprotein hormone  $\beta$ -subunits it is tempting to speculate that this part of the molecule might also be involved in a strong protein–protein interaction such as complex formation with either the  $\alpha$ -subunit or the receptor. In this context it is especially interesting that proteinaceous proteinase inhibitors can prevent the binding of hCG to receptor [19,20]. This finding might suggest that the 17–23 segment has a tertiary structure similar to that of proteinase inhibitors and is reacting with a binding site similar to that of proteinases, or it might directly combine with a proteinase necessary for the binding of the hormone to the receptor. The involvement of such a

Table I

Comparison of the sequences of HUSI-II and various Kazal-type inhibitors

		bPSTI	hPSTI	pPSTI	BUSI-I	BUSI-IIA	POSI
HUSI-II	AA	23	21	20	17	16	15
6–61	%	41.1	37.5	35.7	30.4	28.6	26.8

AA and %: number and percentage of amino acids identical in HUSI-II and the respective inhibitor. For the calculation of percentage only the sequence section common to all inhibitors (6 to 61 of the numbering of the top line of Fig. 1) was taken into account

```

      10      20      30      40      50      60
HUSI-II  QFGLFSKYrTPNCSQYrlpgcprhFnpvCGsDMSTYANECtlcmKIREGGHNkiIRnGPC 1- 61
hCGβ    ...LPQVVCNIRDVRFESIRlPGCPRGVNPV-VSYAVALSCQCALCRRSTTDCGGPKDHPLTCD... 52-111
hLHβ    ...LPQ-VCTIRDVRFESIRlPGCPRGVDPV-VSFPVALSCRGPCRRSTSDCGGPKDHPLTCD... 52-110
hTSHβ   ...LSQDVCTIRDFIYRTVEIPCCLHVAPY-FSYPVALSCKGGKNTDYSDCIHEAEKTNYCT... 47-106
hFSHβ   ...KIQKTCTFKELVYETVAVPCAHADSL-YTYPVATQCHGKCDSDSTDCTVRGLGPSYCS... 46-105

```

Fig. 2. Comparison of the HUSI-II sequence to parts of the sequences of the human glycoprotein hormone  $\beta$ -subunits [15]. Lowercase bold letters in the HUSI-II sequence indicate that the same amino acid is found in at least one of the hormone sequences. The amino acids of the  $\beta$ -subunit sequences which are identical to those of the HUSI-II sequence are shaded. The numbering in the top line represents the numbering of HUSI-II, the right column gives the residue numbers of the respective sequence. The position of the P<sub>1</sub> residue [14] is marked by an arrow.

Table II

Comparison of sequences of several Kazal-type inhibitors and  $\beta$ -subunits of glycoprotein hormones.

		hCG $\beta$ 52-111	hLH $\beta$ 52-110	hTSH $\beta$ 47-106	hFSH $\beta$ 46-105
HUSI-II 6-61	AA	17	15	13	7
	%	30.4	26.8	23.2	12.5
hPSTI 1-56	AA	8	8	7	8
	%	14.3	14.3	12.5	14.3
OMCHI1 7-62	AA	8	7	5	8
	%	14.3	12.5	8.9	14.3
OMCHI2 5-60	AA	9	7	5	5
	%	16.1	12.5	8.9	8.9
OMCHI3 1-51	AA	8	8	8	6
	%	14.3	14.3	14.3	10.7

OMCHI1, -2, -3: first, second and third domain of chicken ovomucoid [21]; AA, %: number and percentage of identical amino acids (100% = 56). The numbers below the names indicate which sections of the protein sequences were compared

proteinase has been proposed [19]. On the other hand, it can be deduced that  $\beta$ -subunits could act as proteinase inhibitors. However, in our preliminary experiments with hCG $\beta$  no inhibitory effect on bovine trypsin could be detected (data not shown). Currently we are starting experiments to investigate whether HUSI-II can antagonize the binding of hCG to its receptor.

**Acknowledgements:** We thank H. Fritz for his stimulating and helpful discussions. This work was supported by the Sonderforschungsbereich 207 of the Deutsche Forschungsgemeinschaft.

## REFERENCES

- [1] Kazal, L., Spicer, D.S. and Brahinsky, R.A. (1948) *J. Am. Chem. Soc.* 70, 3034-3040.
- [2] Laskowski, M., Jr. and Kato, I. (1980) *Annu. Rev. Biochem.* 49, 593-626.
- [3] Laskowski, M., Jr. (1986) in: *Nutritional and Toxicological Significance of Enzyme Inhibitors in Foods* (Friedman, M. ed.) pp. 1-17, Plenum, New York.
- [4] Bartelt, D.C., Shapanka, R. and Greene, L.J. (1977) *Arch. Biochem. Biophys.* 179, 189-199.
- [5] Haendle, H., Fritz, H., Trautschold, I. and Werle, E. (1965) *Hoppe Seyler's Z. Physiol. Chem.* 343, 185-188.
- [6] Fink, E., Jaumann, E., Fritz, H., Ingrisich, H. and Werle, E. (1971) *Z. Physiol. Chem.* 352, 1591-1594.
- [7] Fink, E., Fritz, H., Jaumann, E., Schiessler, H., Förg-Brey, B. and Werle, E. (1973) in: *Protides of the Biological Fluids*, Vol. 20 (Peeters, H. ed.) pp. 425-431, Pergamon, Oxford.
- [8] Schiessler, H., Fink, E. and Fritz, H. (1976) *Methods Enzymol.* 65, 847-859.
- [9] Tschesche, H., Kupfer, S., Klauser, R., Fink, E. and Fritz, H. (1976) in: *Protides of the Biological Fluids*, Vol. 23 (Peeters, H. ed.) pp. 255-266, Pergamon, Oxford.
- [10] Meloun, B., Jonáková, V. and Cechová, D. (1983) *Coll. Czech. Chem. Commun.* 48, 2558-2568.
- [11] Meloun, B., Cechová, D. and Jonáková, V. (1984) *Hoppe-Seyler's Z. Physiol. Chem.* 364, 1665-1670.
- [12] Fritz, H. (1988) *Biol. Chem. Hoppe-Seyler* 369, Suppl., 79-82.
- [13] Podell, D.N. and Abraham, G.N. (1978) *Biochem. Biophys. Res. Commun.* 81, 176-185.
- [14] Schechter, I. and Berger, A. (1967) *Biochem. Biophys. Res. Commun.* 27, 157-162.
- [15] Pierce, J.G. and Parson, T.F. (1981) *Annu. Rev. Biochem.* 50, 465-495.
- [16] Laskowski, M., Jr., Kato, I., Ardelt, W., Cook, J., Denton, A., Empie, M.W., Kohr, W.J., Park, S.J., Parks, K., Schatzley, B.L., Schoenberger, O.L., Tashiro, M., Vichot, G., Whatley, H.E., Wiczorek, A. and Wiczorek, M. (1987) *Biochemistry* 25, 202-221.
- [17] Papamokos, E., Weber, E., Bode, W., Huber, R., Empie, M.W., Kato, I. and Laskowski, M., Jr. (1982) *J. Mol. Biol.* 158, 515-537.
- [18] Bolognesi, M., Gatti, G., Menegatti, E., Guarneri, M., Marquart, M., Papamokos, E. and Huber, R.J. (1982) *Mol. Biol.* 162, 839-868.
- [19] Wilks, J.W. and Hui, O.J. (1987) *Endocrinology* 120, 946-952.
- [20] Willey, K.P. and Leidenberger, F.J. (1989) *J. Biol. Chem.* 264, 19716-19729.
- [21] Kato, I., Schrode, J., Kohr, W.J. and Laskowski, M., Jr. (1987) *Biochemistry* 26, 193-201.
- [22] Greene, L.J. and Bartelt, D.C. (1969) *J. Biol. Chem.* 244, 2646-2657.
- [23] Yamamoto, T., Nakamura, Y., Nishide, J., Emi, M., Ogawa, M., Mori, T. and Matsubara, K. (1985) *Biochem. Biophys. Res. Commun.* 132, 605-612.
- [24] Bartelt, D.C. and Greene, L.J. (1971) *J. Biol. Chem.* 246, 2218-2229.