Available online at www.sciencedirect.com

SCIENCE DIRECT

Biochimica et Biophysica Acta 1748 (2005) 35–42

BIOCHIMICA ET BIOPHYSICA ACTA BBA

http://www.elsevier.com/locate/bba

Cloning and expression of ostrich trypsinogen: an avian trypsin with a highly sensitive autolysis site

Borbála Szenthe[a], Carminita Frost[b], László Szilágyi[a], András Patthy[c], Ryno Naudé[b], László Gráf[a,c,*]

[a]Department of Biochemistry Eötvös Lóránd University, Pázmány Péter sétány 1/C, H-1117 Budapest, Hungary
[b]Department of Biochemistry and Microbiology, University of Port Elizabeth, PO Box 1600, Port Elizabeth 6000, South Africa
[c]Biotechnology Research Group of the Hungarian Academy of Sciences, Pázmány Péter sétány 1/C, H-1117 Budapest, Hungary

Received 7 May 2004; received in revised form 7 December 2004; accepted 7 December 2004
Available online 29 December 2004

Abstract

One of ostrich (*Struthio camelus*) trypsinogen genes was cloned from pancreatic cDNA. Its amino acid sequence compared to known trypsin sequences from other species shows high identity and suggests that it is a member of the phylogenetically anionic trypsinogen I subfamily. After cytoplasmic over expression in *Escherichia coli* and renaturation, the activation properties of ostrich trypsinogen were studied and compared to those of human trypsinogen 1 (also called as human cationic trypsinogen). Ostrich trypsinogen undergoes bovine enterokinase activation and autoactivation much faster than human trypsinogen 1 and exhibits on a synthetic substrate a somewhat higher enzymatic activity than the latter one. The most interesting property of ostrich trypsin is its relatively fast autolysis that can be explained via a mechanism different from the common mechanism for rat and human 1 trypsins. The latter proteases have a site, Arg117–Val118, where the autolysis starts and then goes on in a zipper-like fashion. This is absent from ostrich trypsin. Instead it has a couple of cleavage sites within regions 67–98, including two unusual ones, Arg76–Glu77 and Arg83–Ser84. These appear to be hydrolysed fast in a non-consecutive manner. Such an autolysis mechanism could not be inhibited by a single-site mutation which in humans is proposed to lead to pancreatitis.

Keywords: Ostrich; Trypsinogen; Autolysis

1. Introduction

Kühne [1] described first, more than 130 years ago, that cattle pancreatic juice possessed the property to digest proteins. He proposed that this property was due to an 'enzyme' that he named 'trypsin'. It was also reported by Kühne [1] that the extract from freshly removed cattle pancreas exhibited only a low proteolytic activity, but it was increasing when the juice was allowed to stand. Upon longer standing, however, the proteolytic activity started to decrease. This was the first description of autoactivation of the inactive form and autolytic inactivation of the active form of trypsin, respectively [1]. Since then trypsinogen (the inactive, proenzyme form of trypsin) and trypsin have become prototypes of investigating many aspects of protease action, including the molecular mechanisms of proenzyme autoactivation and the autolysis of the active

Abbreviations: *N*-CBZ-GPR-pNA, *N*-benzyloxycarbonyl-Gly-Pro-Arg-*para*-nitroanilide; oligodT, oligonucleotide containing only thymine specific for the poly-adenine tail of the mRNA; LB, Lauria-Bertani broth; MUB, 4-methylumbelliferone; MUGB, 4-methylumbelliferyl-4-guanidinobenzoate; HuTg-1, human trypsinogen 1; HuTr-1, human trypsin 1; IPTG, isopropyl-thio-galactose; GdnHCl, guanidine-hydro-chloride; EDTA, ethylene diamine tetraacetic acid; PCR, polymerase chain reaction; STI-agarose, soybean trypsin inhibitor linked to agarose; Tris–HCl, α,α,α-Tris–(hydroxymethyl)-methylamine hydrochloride; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; ctg, chymotrypsinogen; sp, serin protease; tg, trypsinogen; tr, trypsin; $NH_4(HCO_3)$, ammonium–hydrogen–carbonate

* Corresponding author. Department of Biochemistry Eötvös Lóránd University, Pázmány Péter sétány 1/C, H-1117 Budapest, Hungary. Tel.: +36 1 209 0555/8781; fax: +36 1 381 2172.
E-mail address: graf@ludens.elte.hu (L. Gráf).

1570-9639/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbapap.2004.12.006

protease [2–5]. Much of the fundamental knowledge about the structural basis of substrate specific serine protease action has also been derived from studies on mutants and variants of trypsin provided by either protein engineering [6,7] or natural evolution. As to the latter alternative, recent examples are human pancreatic mesotrypsin that is highly resistant to canonical trypsin inhibitors [5,8], and crayfish trypsin that, on contrary, shows extremely high affinity towards inhibitors that does not inhibit mammalian trypsins [9]. Thus, it's still a rewarding strategy to explore structure–activity relationships in trypsins from evolutionary significant new species [10]. Ostrich belongs to one of these species representing the connecting taxon (Ratite) between birds and reptiles [11]. We thought to clone the cDNA of ostrich trypsinogen, determine its sequence and express the zymogen in *Escherichia coli* in order to compare its biochemical properties to those of some already well-characterized trypsins.

These studies of ostrich trypsinogen provided us with new insight into the mechanism of trypsin autolysis. Our results show that ostrich trypsin has a unique, multi-cleavage site autolysis region that assures fast autolysis of trypsin in both the intestine and pancreas of ostrich. Such an autolysis mechanism cannot simply be blocked by a single-site mutation of trypsin causing the accumulation of the protease in the pancreas.

2. Materials and methods

2.1. Cloning

Total RNA was isolated from ostrich pancreatic tissue (10 mg), stored in RNeasy solution (Qiagen) by using TRI Reagent (Sigma) according to the manufacturers instruction. cDNA was polymerized with oligodT from the isolated RNA using the Fermantas Revert Aid H Minus First Strand cDNA synthesis kit. The ostrich trypsinogen gene was amplified from the cDNA pool with odT and 5′ end [5′GTN CCN GGN GAY GCN GAY GAY GAY AAR ATH 3′] degenerated primer designed on the basis of the N-terminal amino acid sequence of ostrich trypsinogen [12]. The PCR product was ligated into a pKS+ (Stratagen) vector using TA cloning, and sequenced from both directions with T7 and T3 plasmid specific primers, by automated dideoxy sequencing (ABI Prism) using a Big Dye Terminator kit. After obtaining the DNA sequence the gene was amplified with specific primers. For the 5′ end, 5′C GAA GCT TTG CCC GGT GAT GCC GAT GAC GAC AAG ATC G3′ was used. This contains a *Hin*dIII restriction site downstream from the start codon that extends the known propeptide amino acid sequence [12] with an extra MSTQAL sextapeptide at the N-terminus. Because of the latter extension the expressed protein is although differing from the native pretrypsinogen in the signal peptide, should activate similarly to the native trypsinogen, and will produce a native trypsin. For the 3′end, 5′CGG GAG CTC ATC GAG GCA TCA GTA GGC 3′, including a TAG stop codon and an *Sac*I restriction site was applied. The PCR product was ligated into a pET17c cytoplasmic expression vector (Novagen), for use in *E. coli*, and sequenced.

2.2. Expression and refolding

Ostrich trypsinogen was expressed in BL21 DE3 pLysS (Stratagen) cells into the cytoplasm as an inclusion body. Cells were grown in 500 ml LB broth containing 100 μg/ml ampicillin (Sigma) at 37 °C in a shaking incubator to an absorbance of 0.8 at 600 nm and induced with IPTG (Sigma) (final concentration of 100 μg/ml) for 4 h. Cells were harvested by centrifugation at 14300×*g* for 30 min at 4 °C (Beckman J2-MC), redissolved in 50 ml buffer [50 mM Tris–HCl (Sigma) and 20 mM EDTA (Sigma), pH 8.0], frozen at −20 °C and lysed by thawing. The inclusion bodies were isolated from the cell lysate after adding DNAse (Sigma) (5 μg/ml final concentration), centrifuged (20000×*g*, Beckman J2-MC) and washed twice with the abovementioned buffer. Fifty milligrams of the isolated inclusion body was dissolved in 5 ml 6 M GdnHCl containing 0.1 M Tris–HCl, 2 mM EDTA and 48 mM DTT, pH 8.5, flushed with nitrogen. The dissolved inclusion bodies were incubated at 37 °C with shaking for 30 min. The solution was added to a 45 ml buffer [6 M GdnHCl, 1 mM cystine, 1 mM cysteine in 0.1 M Tris–HCl and 2 mM EDTA, pH 8.5] at 4 °C. The whole refolding mixture (50 ml) was placed in a dialysis tube, and 6 M GdnHCl was dialysed out against 250 ml [1 mM cystine, 1 mM cysteine, 0.1 M Tris–HCl and 2 mM EDTA, pH 8.5] buffer so the final concentration of GdnHCl became 1 M. Thereafter the sample was dialysed into 2.5 mM HCl and stored at −20 °C. HuTg-1 was expressed and refolded as described previously [3].

2.3. Purification

Ecotin column chromatography was employed for zymogen purification [13]. Purified zymogen, activated by trypsin-free bovine enterokinase (Sigma) (purified on an ecotin column) with a final concentration of 0.0125 mg/ml for 60 min at 37 °C in buffer [50 mM Tris–HCl, 10 mM $CaCl_2$ and 0.5 M NaCl, pH 8] was loaded onto a 1 ml STI-agarose column (Sigma). The column was washed with distilled water and eluted with 20 mM HCl. The purities of ostrich trypsinogen and trypsin are illustrated by gel electrophoresis in the presence and absence of a reducing agent (Fig. 1). The protein concentration was determined by measuring the absorbance at 280 nm using theoretical molar extinction coefficients for HuTg-1 and ostrich trypsinogen (ε=36160 and ε=38840, respectively).

2.4. Kinetic measurements

The concentration of the purified zymogens after activation with enterokinase (HuTr-1 and ostrich trypsin)

was determined by active site titration using a 50 mM Tris–HCl buffer containing 10 mM $CaCl_2$ pH 8.0, MUB (Sigma) for calibration and MUGB (Sigma) as covalently binding fluorescent substrate on a Spex Fluoro Max fluorometer with excitation at 380 nm and emission at 460 nm. K_M and k_{cat} values (Table 1) were obtained by fitting 3–5 independent progress curves to the Michaelis-Menten equation using the DYNAFIT software [14] on the *N*-CBZ-Gly-Pro-Arg-pNA chromogenic substrate (Sigma, 100 μM final concentration) with three different (12.5, 18.75 and 25 nM for HuTr-1 and 11.5, 17.25, and 23 nM for ostrich trypsin) final active enzyme concentrations. Measurements were carried out on a Shimadzu spectrophotometer at 405 nm (37 °C).

2.5. SDS-PAGE

Laemmli sample buffer was added in 1:1 volume with (reducing conditions) and without β-mercaptoethanol (non-reducing conditions) to ostrich trypsinogen and trypsin, boiled for 3 min and 20 μl was loaded onto a 15% SDS-polyacrylamide minigel and subjected to electrophoresis at 200 V for 45 min (Fig. 1).

2.6. Protein sequencing

N-terminal amino acid sequencing was performed with an Applied Biosystem 471 A pulsed liquid-phase sequencer. Sequenced autolysis samples contained 4 μM enzyme in 50 mM $NH_4(HCO_3)$. 10 μl of the autolysis mixtures was sequenced.

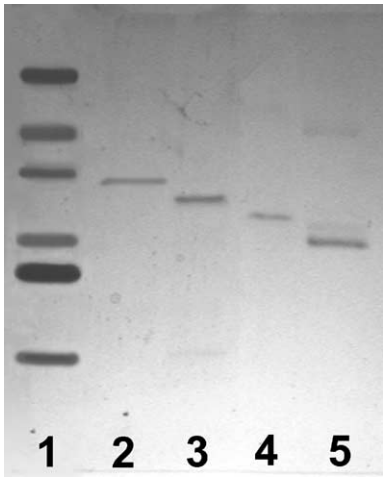

Fig. 1. SDS-PAGE of ostrich trypsinogen (lanes 2 and 4) and trypsin (lanes 3 and 5) in the presence (lanes 2 and 3) and absence of a reducing agent (lanes 4 and 5). Trypsin (lanes 3 and 5) runs farther then the zymogen (lanes 2 and 4) because of the absence of the activation peptide. Both proteins run faster in the absence of a reducing agent (lanes 4 and 5) since the remaining disulfide bridges cause more compact structure and smaller relative volume of the proteins. Lane 1 is Dalton Mark VII-L for SDS-PAGE (Sigma) (Bands from above: 66 kDa; 45 kDa; 36 kDa; 29 kDa; 24 kDa; 20.1 kDa and 14.2 kDa.).

Table 1
Kinetic parameters of ostrich trypsin and HuTr-1 as determined with the *N*-CBZ-GPR-pNA substrate

Enzyme	k_{+1} [$\mu M^{-1}s^{-1}$]	k_{-1} [s^{-1}]	k_{cat} [s^{-1}]	K_m [μM]	k_{cat}/K_m [$\mu M^{-1}s^{-1}$]
Ostrich	11.3±0.34	46.8±2.1	162.3±0.63	18.5±0.8	8.77±0.33
HuTr-1	8.4±0.1	63.2±3.8	182.6±4.3	25.35±3.22	7.4±1.17

2.7. Bioinformatics

Sequences for multiple alignment were collected from Genbank and/or Swiss-Prot/TrEMBL with homology search by Blast [15] using ostrich trypsinogen sequence (without the N-terminal extension, MSTQAL) as a query and the databases (http://www.ncbi.nlm.nih.gov/blast; June, 2004). Sequence alignment and phylogenetic analysis was made with program ClustalW [16] (http://www.ebi.ac.uk/clustlaw, June, 2004). The accession numbers of the protein sequences are the following (Genbank/Swiss-Prot/TrEMBL): Atlantic cod trypsinogen I. X76886/P16049; Atlantic salmon tg I. X70075/P35031; bovine ctg A / P00766; bovine tg II. anionic X54703/Q29463; chicken tg II. P29 U15157/Q90629; dog tg II. Anionic M11589/P06872; human tg I.cationic M22612/P07477; mouse tg 10Like BN000136/Q7M754; ostrich tg AY601749/; pig tg / P00761; rat tg II. anionic V01274/P00763; rat tg III. cationic M16624/P08426; *Xenopus leavis* tg II. U72330/P70059. The physico–chemical parameters of ostrich trypsinogen were calculated using ProtParam at http://www.expasy.org/tools/protparam.html.

3. Results

3.1. Cloning and sequence comparison

The ostrich trypsinogen gene was cloned from pancreatic cDNA. 5′ end primer was designed on the basis of the known N-terminal amino acid sequence of ostrich trypsinogen [12]. Although there are several trypsinogens in ostrich pancreas [17], we picked just one clone for further studies. It was sequenced by automated dideoxy sequencing. The amino acid sequence deduced from the DNA sequence is shown and aligned with sequences of trypsins from some other species in Fig. 2. The expressed ostrich trypsinogen is composed of 237 amino acid residues. Its calculated amino acid composition is practically the same as that reported for the isolated natural ostrich trypsinogen [12]. Due to the sub-cloning from vector pTRAP into expression vector pET17b, the zymogen has got an N-terminal extension, MSTQAL. This modification did not seem to affect its activation. The overall sequence similarity to other known trypsins is about 80%. The deduced molecular mass is 25 191.5 Da. Its theoretical p*I* value is 5.3. Except for human trypsinogen 1, the 12 Cys residues are found at the same positions in the species homologues

Multiple sequence alignment of the propeptides:

```
                          1       10
Bovine ctg.A  -------------CGVPAIQPVLSGLSR 15
Ostrich tg.   -------------MSTQALPGDADDDK 10
Human tg.I    ----MNPLLILTFVAAALAAPFDDDD-K 23
Rat tg.II     ----MRALLFLALVGAAVAFPVDDDD-K 23
Chicken tg.II ---MKFLFLILSCLGAAVAFPGGADDDK 25
Bovine tg.II  ----MHPLLILAFVGAAVAFPSDDDD-K 23
Pig tg.       -------------------FPTDDDD-K  8
Dog tg.II     ----MNPLLILAFLGAAVATPTDDDD-K 23
Rat tg.III    ----MKALIFLAFLGAAVALPLDDDDDK 24
Mouse tg. 10L ----MSTLLFLALVGAAVAFPVDDDD-K 23
XenopusII.tg  IISTMKFLVILVLLGAAVAFE--DDD-K 25
Salamon tg.I  ----MISLVFVLLIGAA----FATEDDK 20
Cod tg.I      -----KSLIFVLLLGAV----FAEED-K 18
                                        :
```

Multiple sequence alignment of the enzymes:

```
                  20        30        40        50        60        70        80        90
Bovine ctr.A  IVNGEEAVPGSWPWQVSLQDKTGFHFCGGSLINENWVVTAAHCGVTTSDVVVAGEFDQGSSSEKIQKLKIAKVFKNS 92
Ostrich tr.   IVGGYNCPAHSVPYQVSLN--AGYHFCGGSLINSQWVLSAAHCYKSSI-QVRLGEYNIDVREDSEVVRSSAAVIRHP 89
Human tr.I    IVGGYNCEENSVPYQVSLN--SGYHFCGGSLINEQWVVSAGHCYKSRI-QVRLGEHNIEVLEGNEQFINAAKIIRHP 97
Rat tr.II     IVGGYTCQENSVPYQVSLN--SGYHFCGGSLINDQWVVSAAHCYKSRI-QVRLGEHNINVLEGNEQFVNAAKIIKHP 97
Chicken tr.II IVGGYTCPEHSVPYQVSLN--SGYHFCGGSLINSQWVLSAAHCYKSRI-QVRLGEYNIDVQEDSEVVRSSSVIIRHP 99
Bovine tr.II  IVGGYTCAENSVPYQVSLN--AGYHFCGGSLINDQWVVSAAHCYQYHI-QVRLGEYNIDVLEGGEQFIDASKIIRHP 97
Pig tr.       IVGGYTCAANSIPYQVSLN--SGSHFCGGSLINSQWVVSAAHCYKSRI-QVRLGEHNIDVLEGNEQFINAAKIITHP 82
Dog tr.II     IVGGYTCEENSVPYQVSLN--AGYHFCGGSLISDQWVVSAAHCYKSRI-QVRLGEYNIDVLEGNEQFINSAKVIRHP 97
Rat tr.III    IVGGYTCQKNSLPYQVSLN--AGYHFCGGSLINSQWVVSAAHCYKSRI-QVRLGEHNIDVVEGGEQFIDAAKIIRHP 98
Mous tr. 10L  IVGGYTCRENSVPYQVSLN--SGYHFCGGSLINDQWVVSAAHCYKSRI-QVRLGEHNINVLEGNEQFIDAANIIKHP 97
Xenopus tr.II IVGGFTCAKNAVPYQVSLN--AGYHFCGGSLINSQWVVSAAHCYKSRI-QVRLGEHNIALNEGTEQFIDSQKVIKHP 99
Salamon tr.I  IVGGYECKAYSQTHQVSLN--SGYHFCGGSLVNENWVVSAAHCYKSRV-EVRLGEHNIKVTEGSEQFISSSRVIRHP 94
Cod tr.I      IVGGYECTKHSQAHQVSLN--SGYHFCGGSLVSKDWVVSAAHCYKSVL-RVRLGEHHIRVNEGTEQYISSSSVIRHP 92
              **.*  .   : . ****:  : * *******:..:**::*.**      * :. *..    .      .  :: :.

                     100       110       120       130       140       150       160
Bovine ctr.A  KYNSLTINNDITLLKLSTAASFSQTVSAVCLPSASDDFAAGTTCVTTGWGLTRYTNANTPDRLQQASLPLLSNTNCK 169
Ostrich tr.   KYSSRSLDNDIMLIKLASPVAYSADVQPIALPSSCV--KAGTKCLISGWGNTLSSGSSFPEILQCLQAPVLSDRECR 164
Human tr.I    QYDRKTLNNDIMLIKLSSRAVINARVSTISLPTAPP--ATGTKCLISGWGNTASSGADYPDELQCLDAPVLSQAKCE 172
Rat tr.II     NFDRKTLNNDIMLIKLSSPVKLNARVATVALPSSCA--PAGTQCLISGWGNTLSSGVNEPDLLQCLDAPLLPQADCE 172
Chicken tr.II KYSSITLNNDIMLIKLASAVEYSADIQPIALPSSCA--KAGTECLISGWGNTLSNGYNYPELLQCLNAPILSDQECQ 174
Bovine tr.II  KYSSWTLDNDILLIKLSTPAVINARVSTLALPSACA--SGSTECLISGWGNTLSSGVNYPDLLQCLEAPLLSHADCE 172
Pig tr.       NFNGNTLDNDIMLIKLSSPATLNSRVATVSLPRSCA--AAGTECLISGWGNTKSSGSSYPSLLQCLKAPVLSDSSCK 167
Dog tr.II     NYNSWILDNDIMLIKLSSPAVLNARVATISLPRACA--APGTQCLISGWGNTLSSGTNYPELLQCLDAPILTQAQCE 172
Rat tr.III    SYNANTFDNDIMLIKLNSPATLNSRVSTVSLPRSCG--SSGTKCLVSGWGNTLSSGTNYPSLLQCLDAPVLSDSSCK 173
Mouse tr.10L  KFKKKTLDNDIMLIKLSSPVTLNARVATVALPSSCA--AAGTQCLISGWGNTLSSGVNNPDLLQCLDAPLLPQADCE 172
Xenopus tr.II NYNSRNLDNDIMLIKLSTTARLSANIQSVPLPSACA--SAGTNCLISGWGNTLSSGTNYPDLLQCLNAPILTDSQCS 174
Salamon tr.I  NYSSYNIDNDIMLIKLSKPATLNTYVQPVALPTSCA--PAGTMCTVSGWGNTMSSTAD-SNKLQCLNIPILSYSDCN 168
Cod tr.I      NYSSYNINNDIMLIKLTKPATLNQYVHAVALPTECA--ADATMCTVSGWGNTMSSVAD-GDKLQCLSLPILSHADCA 166
              .:.   ::*** *:** . .  .  : .: **        .* * :*** *  .  .  . **  . *:*.  .*

             170       180       190       200       210       220       230       240
bov.chytrA   KYWGTKIKDAMICAG--ASGVSSCMGDSGGPLVCKKNGAWTLVGIVSWGSSTCSTSTPGVYARVTALVNWVQQTLAAN- 245
ostrich tr   NAYPGEISSNMICVGFLEGGKDSCQGDSGGPVVCDG----TLQGIVSWGIGCAQKGYPGVYTKVCNYVSWIQETIAAY- 237
huI.tr       ASYPGKITSNMFCVGFLEGGKDSCQGDSGGPVVCNG----QLQGVVSWGDGCAQKNKPGVYTKVYNYVKWIKNTIAANS 247
ratII.tr     ASYPGKITDNMVCVGFLEGGKDSCQGDSGGPVVCNG----ELQGIVSWGYGCALPDNPGVYTKVCNYVDWIQDTIAAN- 246
chickenII.tr EAYPGDITSNMICVGFLEGGKDSCQGDSGGPVVCNG----ELQGIVSWGIGCALKGYPGVYTKVCNYVDWIQETIAAY- 248
bovineII.tr  ASYPGEITNNMICAGFLEGGKDSCQGDSGGPVACNG----QLQGIVSWGYGCAQKGKPGVYTKVCNYVDWIQETIAANS 247
pig tr       SSYPGQITGNMICVGFLEGGKDSCQGDSGGPVVCNG----QLQGIVSWGYGCAQKNKPGVYTKVCNYVNWIQQTIAAN- 231
dogII.tr     ASYPGQITENMICAGFLEGGKDSCQGDSGGPVVCNG----ELQGIVSWGYGCAQKNKPGVYTKVCNFVDWIQSTIAANS 247
ratIII.tr    SSYPGKITSNMFCLGFLEGGKDSCQGDSGGPVVCNG----QLQGVVSWGYGCAQKGKPGVYTKVCNYVNWIQQTVAAN- 247
mous10Ltr    ASYPGKITKNMICVGFLEGGKDSCQGDSGGPVVCNG----QLQGIVSWGYGCAQKDNPGVYTKVCNYVDWIQNTIAAN- 246
xenopusII.tr NSYPGEITKNMFCAGFLAGGKDSCQGDSGGPVVCNG----QLQGVVSWGYGCAQRNYPGVYTKVCNFVTWIQSTISSN- 248
salamonI.tr  NSYPGMITNAMFCAGYLEGGKDSCQGDSGGPVVCNG----ELQGVVSWGYGCAEPGNPGVYAKVCIFNDWLTSTMASY- 242
codI.tr      NSYPGMITQSMFCAGYLEGGKDSCQGDSGGPVVCNG----VLQGVVSWGYGCAERDHPGVYAKVCVLSGWVRDTMANY- 240
               :   *.  *.* *   .* .** ******:.*.      * *:**** . .  . ****::*     *: .*::
```

Fig. 2. Comparison of the newly sequenced ostrich trypsinogen amino acid sequence with some known vertebrate trypsinogen sequences and bovine chymotrypsinogen A for the classical chymotrypsin numbering [18,19] by multiple alignment. Data were collected from Genbank with homology search at http://www.ncbi.nlm.nih.gov/blast using ostrich trypsinogen sequence as query. Sequences were analysed with the alignment tool available at http://www.ebi.ac.uk/tools/clutalw. Accession numbers of the protein sequences are the following (Genbank/Swiss-Prot/TrEMBL): Atlantic cod trypsinogen I. X76886/P16049; Atlantic salmon tg I. X70075/P35031; bovine ctg A /P00766; bovine tg II. anionic X54703/Q29463; chicken tg II. P29 U15157/Q90629; dog tg II. Anionic M11589/P06872; human tg I.cationic M22612/P07477; mouse tg 10Like BN000136/Q7M754; ostrich tg AY601749/; pig tg /P00761; rat tg II. anionic V01274/P00763; rat tg III. cationic M16624/P08426; *X. leavis* tg II. U72330/P70059. Autolysis sites are shown bold with black background. Amino acids that participate in Ca^{2+} binding are shown with gray background. Cleavage site of enterokinase in the propetide sequence is underlined. The six amino acid extension in ostrich propeptide is a by-product of cloning and was not aligned and is shown in gray. Numbers in the first row show the chymotrypsin numbering. Numbers beside the sequences indicate the exact amino acid positions.

compared (Fig. 2). The amino acids involved in calcium binding [18] (Glu70, Asn72, Val75, Glu77 and Glu80, conventional chymotrypsinogen numbering [19,20]) are also conserved for the ostrich sequence (Fig. 2). Since phylogenetic comparisons of vertebrate trypsinogens showed (Fig. 3) that the nearest relative of ostrich trypsinogen is chicken trypsinogen P29 [21], the former one also belongs to the anionic trypsinogen subfamily I [10]. Because there is no reptile trypsinogen sequence available in the protein databanks we could not confirm by a trypsinogen protein sequence based phylogenetic tree that the ostrich (representing the Ratite superorder) is a transitional taxon between the reptiles and the birds.

3.2. Catalytic properties

The pH optimum of 8.0 found for ostrich trypsin (results not shown) is similar to those from other known trypsins. As measured on a synthetic substrate (see Materials and methods) ostrich trypsin was found to have a 20% better catalytic efficiency than HuTr-1 (Table 1).

3.3. Activation by bovine enterokinase and autoactivation

Ostrich trypsinogen displayed a faster activation mechanism when compared to HuTg-1 (Fig. 4). The enterokinase recognition sequence in ostrich trypsinogen has an additional Ala residue preceding the three Asp residues (DADDDK) at position 10 in the ostrich sequence (Fig. 2). This is not an unusual enterokinase cleavage site in non-mammalian species (Fig. 2) [22]. This difference does not decrease, but may rather increase the efficiency of zymogen activation by enterokinase in the non-linear region of the activation curve (Fig. 4). Perhaps for the same structural reason, the autoactivation of ostrich trypsinogen was also faster than that of HuTg-1. HuTg-1 displayed a two fold slower autoactivation (Fig. 5).

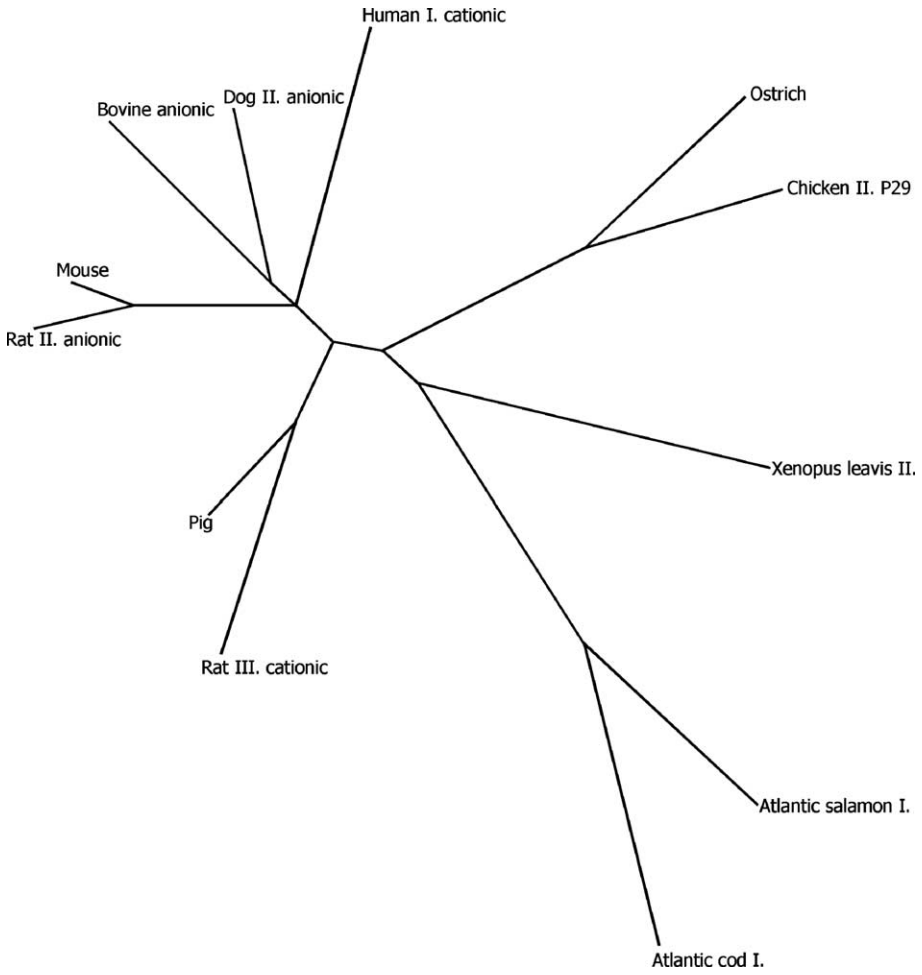

Fig. 3. Dendrogramm aligned with http://www.ebi.ac.uk/clustlaw. Protein sequence data were collected from http://www.ncbi.nlm.nih.gov/blast; June, 2004. The GenBank accession numbers for the sequences are: Accession numbers of the protein sequences are the following (Genbank/Swiss-Prot/TrEMBL): Atlantic cod trypsinogen I. X76886/P16049; Atlantic salmon tg I.; bovine tg II. anionic X54703/Q29463; chicken tg II. P29 U15157/Q90629; dog tg II. Anionic M11589/P06872; human tg I.cationic M22612/P07477; mouse tg 10Like BN000136/Q7M754; ostrich tg AY601749/; pig tg /P00761; rat tg II. anionic V01274/P00763; rat tg III. cationic M16624/P08426; *X. leavis* tg II. U72330/P70059.

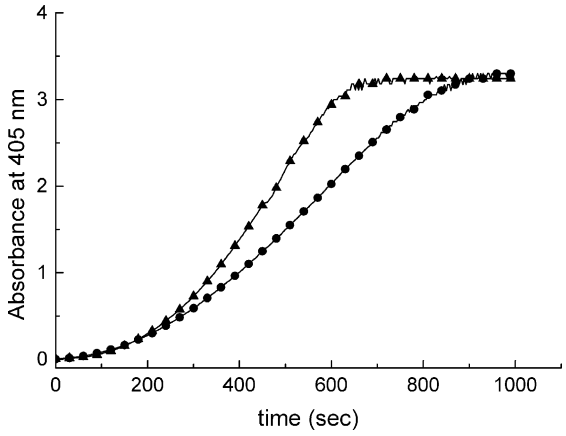

Fig. 4. Activation of ostrich and Hu-1 trypsinogen by bovine enterokinase. Kinetic measurements were done in 1 ml buffer [50 mM Tris–HCl buffer, pH 8, containing 10 mM $CaCl_2$ and 0.1% Triton X-100] with 1.87 μM enzyme and 2.5 μg of bovine enterokinase in the presence of 250 μM *N*-CBZ-GPR-pNA [7]. Ostrich trypsinogen (▲); Hu-Tg-1 (●).

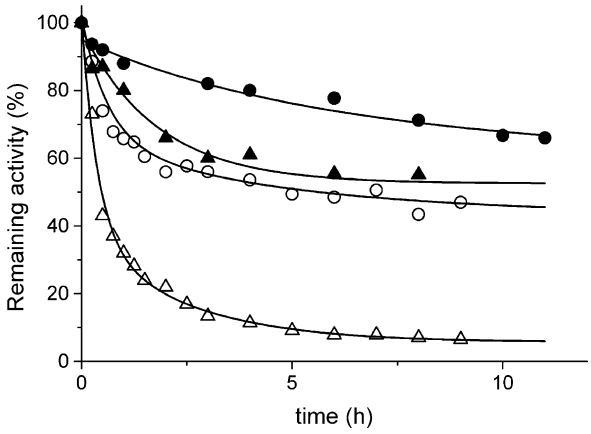

Fig. 6. A time study of the loss of activity for ostrich and Hu-1 trypsin. Activated enzymes were purified on an STI-agarose column. Active enzyme with a final concentration of 4 μM was incubated at 37 °C in 50 mM Tris–HCl buffer (pH 8) with 10 mM $CaCl_2$ or 1 mM EDTA. Aliquots (3 μl) were taken out for activity measurements. HuTr-1+EDTA (○); ostrich trypsin+EDTA (△), HuTr-1+Ca^{2+} (●), ostrich trypsin+Ca^{2+} (▲).

3.4. Autolysis

Autolysis studies were carried out in the presence and absence of Ca^{2+} and followed by enzyme activity measurements. Under both conditions ostrich trypsin degraded much faster than HuTr-1 (Fig. 6). Such a fast self-degradation is not a common attribute of serine proteases [23]. The N-terminal amino acid sequences of fragments from the lysis mixtures for ostrich and Hu-1 trypsins were determined and compared in Table 2 to those previously identified for rat trypsin II [4]. The homologues cleavage sites in the sequences of these three species are also marked in Fig. 2 (bold with black background). It appears from the comparison that autolysis sites Arg67–Leu68, Lys/Arg90–His91 and Lys/Arg97–Thr/Ser98 are common to rat, human and ostrich trypsins. However, cleavage site Arg117–Val118, that was shown to be the abundant and autolysis rate determining site for both rat [4] and Hu-1 trypsins [24], is absent from the structure of ostrich trypsin. Instead, two unusual autolysis sites, Arg76–Glu77 and Arg83–Ser84 (absent from the known structures of other trypsins), are present in the Ca^{2+} binding pocket of ostrich trypsin (Table 2, Fig. 2). Consequently, the cleavages of these two sites are likely to be responsible for the relatively fast autolysis of ostrich trypsin.

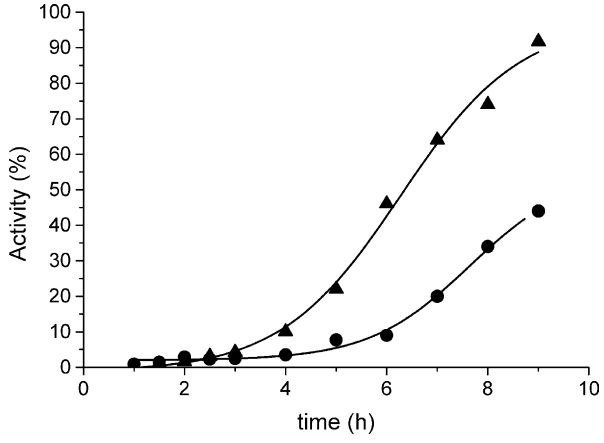

Fig. 5. Autoactivation of ostrich (▲) and Hu-1 trypsinogen (●). Autoactivation experiments were conducted with 1.87 μM final zymogen concentration in 50 mM Tris–HCl buffer, pH 8, containing 10 mM $CaCl_2$ at 37 °C. Aliquots (10 μl) were removed and activity was measured in 1 ml buffer with *N*-CBZ-GPR-pNA (100 μM final concentration) at 405 nm. Activity is shown as a percentage of the maximum activity.

Table 2
Comparison of autolysis sites of ostrich trypsin with rat trypsin II and HuTr-1

Enzyme	Autolysis site	N-terminal sequence of fragments
Rat trypsin II	N-terminus	IVGGY…
	Arg67–Leu68	LGEHN…
	Lys90–His91	HPNFD…
	Arg96–Lys97	KTLNN…
	Lys107–Leu108	LSSPV…
	Arg117–Val118	VATVA…
Human trypsin 1	*N-terminus*	IVGGY…
	Arg67–Leu68	LGEHN…
	Arg90–His91	HPQYD
	Lys97–Thr98	TLNND…
	Arg117–Val118	VSTIS…
	Lys175–Ile176	ITDNM…
Ostrich trypsin	*N-terminus*	IVGGY…
	Arg67–Leu68	LGEYN…
	Arg76–Glu77	***EDSEV***…
	Arg83–Ser84	***SSAAV***…
	Arg90– His91	HPKYS…
	Arg97–Ser98	SLNND…

Sequences of fragments from the autolysis mixtures, determined by protein sequencing, are shown in italics. Unique autolysis sites in ostrich trypsin are shown in bold and italics. Autolysis sites of rat trypsin II are cited from [4].

4. Discussion

The amino acid sequence of ostrich trypsinogen was deduced from the cloned DNA sequence determined in this study and aligned with those of some other vertebrate trypsinogens (Fig. 2). The phylogenetic comparison shows that the cloned ostrich enzyme's nearest relative is P29 chicken trypsinogen (Fig. 3). Since this chicken zymogen belongs to the phylogenetically anionic trypsinogen I subfamily [10,21], ostrich trypsinogen is also a member of the same subfamily. The calculated isoelectric point of ostrich trypsinogen is 5.3. Ostrich trypsin has some specific amino acid mutations at otherwise conserved sites, e.g. Val instead of Gln at position 81 and Val instead of Phe at position 82.

The N-terminal sequences of trypsins are highly conserved across the species compared (Fig. 2). It may be noted that both ostrich trypsin and HuTr-1 contains Asn at position 21, a position suggested to affect the rate of autoactivation [3,8,23]. Thus, the difference in the autoactivation rates of these two zymogens must have another structural cause. It is likely that the faster autoactivation of ostrich trypsinogen (Fig. 5), just like its more efficient activation by enterokinase (Fig. 4), may be due to the ADDDK structure of its activation peptide (Fig. 2). It contains the crucial DDDK recognition pattern for enterokinase cleavage [25], but the Ala in ostrich trypsinogen moderates the negative charge of this cleavage site. Such an effect has recently been reported to accelerate the autoactivation of some human trypsinogen mutants [22].

Perhaps the most unusual property of ostrich trypsin is its relatively fast autolysis (Fig. 6). Ostrich trypsin does not contain the autolysis site, Arg117–Val118, where the autolysis of rat, human and bovine trypsins starts [4,24]. Instead, there are two new autolysis sites, Arg76–Glu77 and Arg83–Ser84, in the structure of ostrich trypsin that together with the other sites also present in HuTr-1 undergo extensive autolysis (Table 2). Our preliminary studies show that these cleavages may occur in a parallel rather than in a consecutive fashion. The latter kind of mechanism was established for rat trypsin autolysis [4]. To provide more adequate experimental evidence for this new mechanism, the elimination of some of these autolytic sites by site-directed mutagenesis and autolysis studies of these mutant trypsins have to be performed as was done with rat trypsin [4].

Acknowledgements

The authors gratefully acknowledge the Hungarian Science and Technology Foundation (HSTF) in Budapest, the Hungarian Scientific Research Fund (OTKA T037568 to L.Sz.), the University of Port Elizabeth and the South African National Research Foundation (NRF) for their financial support, and the Grahamstown abattoir for the supply of ostrich pancreas. They thank Mrs. Énekes Emma for her invaluable technical assistance.

References

[1] W. Kühne, Über die Verdauung der Eiweisstoffe durch den Pancreassaft, Virchows Arch. 39 (1867) 130–185.
[2] S. Halfon, C.S. Craik, Trypsin, in: A.J. Barrett, N.D. Rawlings, J.F. Woessner (Eds.), Handbook of Proteolytic Enzymes, Academic Press, New York, 1998, pp. 12–21.
[3] L. Szilágyi, E. Kénesi, G. Katona, Gy. Kaslik, G. Juhász, L. Gráf, Comparative in vitro studies on native and recombinant human cationic trypsins, J. Biol. Chem. 276 (2001) 24574–24580.
[4] E. Várallyay, G. Pál, A. Patthy, L. Szilágyi, L. Gráf, Two mutations in rat trypsin confer resistance against autolysis, Biochem. Biophys. Res. Commun. 243 (1998) 56–60.
[5] J.-M. Chen, C. Ferec, Human trypsins, in: A.J. Barrett, N.D. Rawlings, J.F. Woessner (Eds.), Handbook of Proteolytic Enzymes, 2 edn., Elsevier, London, 2004, pp. 1489–1493.
[6] L. Hedstrom, L. Szilágyi, W. Rutter, Converting trypsin to chyomotrypsin: the role of surface loops, Science 255 (1992) 1249–1253.
[7] B. Jelinek, J. Antal, I. Venekei, L. Graf, Ala226 to Gly and Ser189 to Asp mutations convert rat chymotrypsin B to a trypsin-like protease, Protein Eng. Des. Sel. 2 (2004) 127–131.
[8] G. Katona, G.I. Berglund, J. Hajdu, L. Gráf, L. Szilágyi, Crystal structure reveals basis for the inhibitor resistance of human brain trypsin, J. Mol. Biol. 315 (2002) 1209–1218.
[9] A. Patthy, S. Amir, Z. Malik, A. Bódi, J. Kardos, B. Asbóth, L. Gráf, Remarkeable phylum selectivity of *Schistocerca gregaria* trypsin inhibitor: the possible role of enzyme-inhibitor flexibility, Arch. Biochem. Biophys. 398 (2002) 179–187.
[10] J.C. Roach, K. Wang, L. Gan, L. Hood, The molecular evolution of the vertebrate trypsinogens, J. Mol. Evol. 45 (1997) 640–652.
[11] J. Ausio, J.T. Soley, W. Burger, J.D. Lewis, D. Barreda, K.M. Cheng, The histidine-rich protamine from ostrich and tinamou sperm. A link between reptile and bird protamines, Biochemistry 38 (1999) 180–184.
[12] M.D. Bodley, R.J. Naude, W. Oelofsen, A. Patthy, Ostrich trypsinogen: purification, kinetic properties and characterization of the pancreatic enzyme, Int. J. Biochem. Cell Biol. 27 (1995) 719–728.
[13] Z. Lengyel, G. Pál, M. Sahin-Tóth, Affinity purification of recombinant trypsinogen using immobilized ecotin, Protein Expr. Purif. 12 (1998) 291–294.
[14] P. Kuzmic, Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase, Anal. Biochem. 237 (1996) 260–273.
[15] S.F. Altschul, T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lippman, Gapped BLAST PSI-Blast: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.
[16] J.D. Thompson, D.G. Higgins, T.J. Gibbson, CLUSTALW: improving the sensitivity of multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrixchoice, Nucleic Acids Res. 22 (1994) 4673–4680.
[17] N. Smith, R.J. Naude, W. Oelofsen, C. Lazure, A. Patthy, The isolation and partial characterization of trypsinogen, pancreatic secretory trypsin inhibitor and multiple forms of chymotrypsinogen and trypsin from the pancreas of the ostrich (*Struthio camelus*), Int. J. Biochem. 24 (1992) 877–885.
[18] W. Bode, P. Schwager, The refined crystal structure of bovine beta-trypsin at 1,8A resolution: II. Crystallographic refinement, calcium binding site, benzamidine binding site and active site at pH 7,0, J. Mol. Biol. 98 (1975) 693–717.

[19] R. Zwilling, H. Neurath, Invertebrate proteases, Methods Enzymol. 80 (1981) 633–664.
[20] D.M. Shotton, B.S. Hartley, Amino-acid sequence of porcine pancreatic elastase and its homologies with other serine proteases, Nature 225 (1970) 802–806.
[21] K. Wang, L. Gan, I. Lee, L. Leroy, Isolation and characterization of chicken trypsinogen gene family, Biochem. J. 307 (1995) 471–479.
[22] J.-M. Chen, Z. Kukor, C. Marechal, M. Toth, L. Tsakiris, O. Raguenes, C. Ferec, M. Sahin-Tóth, Evolution of trypsinogen peptides, Mol. Biol. Evol. 20 (2003) 1767–1777.
[23] M. Sahin-Tóth, Human cationic trypsinogen. Role of Asn-21 in zymogen activation and implications in hereditary pancreatitis, J. Biol. Chem. 275 (2000) 22750–22755.
[24] Z. Kukor, M. Tóth, G. Pál, M. Sahin-Tóth, Arg-117 is the reactive site of an inhibitory surface loop that controls spontaneous zymogen activation, J. Biol. Chem. 277 (2002) 6111–6117.
[25] D. Lu, K. Futterer, S. Korolev, X. Zheng, K. Tan, G. Waksman, J.E. Sandler, Crystal structure of enteropeptidase light chain complexed with an analog of the trypsinogen activation peptide, J. Mol. Biol. 292 (1999) 361–373.