FEBS Letters 580 (2006) 4195–4199

Mouse stefins A1 and A2 (*Stfa1* and *Stfa2*) differentiate between papain-like endo- and exopeptidases

Marko Mihelič[a], Cory Teuscher[b], Vito Turk[a], Dušan Turk[a,*]

[a] *Department of Biochemistry and Molecular Biology, J. Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia*
[b] *Departments of Medicine and Pathology, C317 Given Medical Building, University of Vermont, Burlington, VT 05405, USA*

Received 23 April 2006; revised 13 June 2006; accepted 28 June 2006

Available online 5 July 2006

Edited by Miguel De la Rosa

Abstract **Stefin A (*Stfa*) acts as a competitive inhibitor of intracellular papain-like cysteine proteases which play important roles in normal cellular functions such as general protein turnover, antigen processing and ovarian follicular growth and maturation. In the mouse there are at least three different variants of *Stfa* (*Stfa1*, *Stfa2* and *Stfa3*). Recent genetic studies identified structural polymorphisms in *Stfa1* and *Stfa2* as candidates for *Aod1b*, a locus controlling susceptibility to day three thymectomy (D3Tx)-induced autoimmune ovarian disease (AOD). To evaluate the functional significance of these polymorphisms, recombinant allelic proteins were expressed in *Escherichia coli*, purified and characterized. The polymorphisms do not markedly alter the folding characteristics of the two proteins. *Stfa1* and *Stfa2* both act as fast and tight binding inhibitors of endopeptidases papain and cathepsins L and S, however their interaction with exopeptidases cathepsins B, C and H was several orders of magnitude weaker compared to human, porcine and bovine *Stfa*. Notwithstanding, the K_i values for the interactions of *Stfa1-b* from AOD resistant C57BL/6J mice was 10-fold higher than that of the *Stfa1-a* allele from susceptible A/J mice for papain, cathepsins B, C and H but not L and S. In contrast, the inhibitory activities of *Stfa2-a* and *Stfa2-b* were found to be roughly equivalent for all targets peptidases.**

Keywords: Stefin A; Cystatins; Cathepsin; Autoimmunity; Antigen processing; Ovarian dysgenesis

1. Introduction

The thymus is the primary site of T-cell development [1] and its removal from 3-day-old mice (day 3 thymectomy – D3Tx) leads to a loss of peripheral tolerance and the development of organ-specific autoimmune diseases during adulthood. D3Tx does not lead to autoimmune disease in all mouse strains, indicating that the development of these diseases is genetically controlled [2]. For example, A/J mice are highly susceptible to autoimmune ovarian disease (AOD) whereas C57BL/6J mice are not. Previously, using quantitative trait locus (QTL) linkage analysis we identified five regions of the mouse genome harboring loci (*Aod1–Aod5*) controlling susceptibility to AOD and its intermediate phenotypes, anti-ovarian autoantibody production, oophoritis, and atrophy [3]. Subsequently, congenic mapping revealed that *Aod1* encompassed two linked loci with opposing allelic effects on disease susceptibility (*Aod1a* and *Aod1b*) and lead to the identification of structural polymorphisms in Stefin A1 (*Stfa1*) and A2 (*Stfa2*) as candidates for *Aod1b*.

Stefins A (*Stfa*) are non-glycosylated intracellular inhibitors of cysteine cathepsins [4] including cathepsin S and L. Cathepsins S and L plays an important role in antigen processing and presentation [5] and cathepsin S has been shown to play a role in the development of D3Tx-induced autoimmune disease [6]. Additionally, cathepsins are involved in ovarian follicle growth and maturation [7]. They exhibit complex temporal and spacial expression patterns at different stages of the estrous cycle and follicular development consistent with divergent functions for specific cathepsins in follicular development, growth and rupture [8]. In contrast to human, at least three different *Stfa* variants are encoded within the mouse genome (*Stfa1*, *Stfa2*, and *Stfa3*) [9]. They share 60–65% identity in amino acid sequences and 55–60% identity when compared to human *Stfa*.

The structural polymorphisms that distinguish the C57BL/6J (*Stfa1-b* and *Stfa2-b*) and A/J (*Stfa1-a* and *Stfa2-a*) alleles may influence the folding properties and/or their affinity of binding to their target proteases. In order to evaluate the functional significance of these polymorphisms all four variants of mouse *Stfa* were expressed and their interaction constants with papain and a variety of cysteine cathepsins were determined. The alleles of *Stfa1* and *Stfa2* were essentially equivalent as functional inhibitors of several papain-like cysteine proteases; however, the *Stfa1-a* and *Stfa2-a* alleles have a detectable impact on their interaction with these proteases compared to *Stfa1-b* and *Stfa2-b*.

2. Materials and methods

2.1. Materials

Plasmids and expression strains for bacterial and yeast expression were from Novagen and Invitrogen, respectively. Restriction endonucleases and other molecular biology enzymes were from New England Biolabs. Ion-exchange and gel-filtration carriers were from Amersham Pharmacia Biotech. Fluorogenic substrates were from Bachem, other chemicals and salts used in experiments were from Serva. cDNA clones for mouse cathepsin L (clone ID: IRAKp961H028Q2) and cathepsin S (clone ID: IMAGp998M103751Q3) were obtained from RZPD (www.rzpd.de).

2.2. Cloning, expression and purification of mouse stefin A variants

The cDNA clones for mouse *Stfa1* and *Stfa2* were amplified by polymerase chain reaction (PCR). The PCR products were digested with *Nde*I and *Bam*HI and subcloned into pET-3a or pET-11a expression vector.

*Corresponding author. Fax: +386 1 477 3984.
E-mail address: dusan.turk@ijs.si (D. Turk).

0014-5793/$32.00
doi:10.1016/j.febslet.2006.06.076

Escherichia coli strain BL21DE3 pLysS was used for expression. Cells transformed with recombinant plasmids were grown at 37 °C in Luria-Bertani (LB) medium, containing 100 μg/ml ampicillin and 50 μg/ml chloramphenicol to OD600 0.6. Expression was induced by addition of isopropyl β-D-thiogalactoside (IPTG) to a final concentration of 0.4 mM. After 1–4 h of induction, the cells were harvested by centrifugation and resuspended in the ice-cold lysis buffer (50 mM Tris–HCl, 0.1 M NaCl, 1 mM EDTA, pH 8). Cell disruption was achieved by freezing and thawing of the cell suspension and by sonification. Nucleic acids were precipitated by the addition of 5% solution of polyethyleneimine to a final concentration of 0.1% and removed together with the insoluble fractions of bacterial cells by centrifugation. Clear cytosolic fraction was concentrated and loaded onto a Superdex G-75 gel filtration column equilibrated in 30 mM Tris–HCl, 0.3 M NaCl, pH 8.0. Eluted fractions were analyzed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE). Fractions containing recombinant protein were collected, dialyzed against 30 mM Tris–HCl, pH 8, and applied onto a SP-Sepharose Fast Flow column equilibrated with the same buffer. The elution was carried out with a linear gradient of NaCl (0–0.5 M) in the starting buffer. Fractions containing pure protein, as judged by SDS–PAGE analysis, were pooled and concentrated.

2.3. Cloning, expression and purification of mouse cathepsins L and S

The cDNA region encoding mouse procathepsins L and S were amplified by PCR. Amplified DNA products were digested with *Xho*I and *Not*I and ligated into pPIC9 expression plasmid. The recombinant plasmids were subsequently linearized with *Sal*I and electroporated into *Pichia pastoris* strain GS115 using Gene Pulser (Bio-Rad). The recombinant clones of *P. pastoris* were selected on MD agar plates. The integration of recombinant plasmid into the genome of *P. pastoris* was additionally examined by the yeast colony PCR as described in [10]. 40 recombinant clones were tested for protein production using small 50 ml bioreactors (TPP). The clone with the highest level of protein production, as judged by the SDS–PAGE analysis, was selected for large-scale expression.

Mature mouse cathepsin S was purified from the expression media using ion-exchange chromatography step on SP-Sepharose Fast Flow equilibrated in 30 mM sodium acetate, 1 mM EDTA, pH 5.5. Proteins were eluted using linear gradient of NaCl (0–0.5 M) in the starting buffer. Fractions containing cathepsin S were additionally purified on Superdex G-75 gel filtration column.

Crude expression media containing mouse procathepsin L was dialyzed against 30 mM Tris–HCl, 1 mM EDTA, pH 7.5 buffer and loaded onto a Q-Sepharose Fast Flow column. Procathepsin L containing fractions were eluted with equilibration buffer containing 0.5 M NaCl, dialyzed against 30 mM sodium acetate, 1 mM EDTA, 2 mM DTT, pH 5, and applied onto SP-Sepharose Fast Flow ion exchange column equilibrated with the same buffer. Pure fractions of mature cathepsin L were eluted using equilibration buffer containing 0.3 M NaCl.

2.4. Circular dichroism spectroscopy

The protein was dissolved in the 10 mM phosphate buffer, pH 5.5, at concentration of 0.5 mg/ml. The far-UV CD spectra were measured at room temperature (22 °C) on the Aviv 60DS spectrometer (Piscataway, NJ, USA). Data points were recorded from 250 to 190 nm using 1 nm interval and a cell path length of 0.1 cm. The bandwidth was set to 1 nm.

2.5. Kinetics of inhibition

In all kinetic experiments, papain, cathepsin S, cathepsin H and cathepsin B were assayed using 0.1 M phosphate buffer, pH 6.0, containing 3 mM DTT and 1 mM EDTA, whereas for cathepsin C, 100 mM NaCl was added into assay buffer. Cathepsins L was assayed in 0.1 M acetate buffer, pH 5.5, containing 3 mM DTT and 1 mM EDTA. The active concentration of enzymes was determined by active site titration using cysteine protease inhibitor E-64 or human recombinant *Stfa* pretitrated with papain.

All inhibition experiments were performed under the pseudo first-order conditions with inhibitor concentration at least 10-fold higher than the enzyme concentrations and less than 10% of substrate was hydrolyzed during these experiments. *Stfa* variants were in various concentrations mixed with substrate solution dissolved in the appropriate buffer in the fluorometric cuvette. The reaction was initiated by the addition of enzyme in a negligible volume. Constant concentrations of papain (50 pM), cathepsin L (30 pM), cathepsin S (60 pM), cathepsin C (100 pM), cathepsin H (300 pM) and cathepsin B (100 pM) and substrates (10 μM Z-Phe-Arg-MCA for papain, cathepsin L, cathepsin B and cathepsin S and 10 μM Z-Arg-MCA for cathepsin H) were used throughout the measurements. The progress curves were monitored at excitation and emission wavelengths of 370 and 460 nm, respectively, using a C-61 fluorimeter (Photon Technology International).

In all cases, typical biphasic curves were observed and were analyzed by the least-square fitting to the appropriate equation [11]. Linear dependence of pseudo first-order rate constant against the inhibitor concentration was observed for all enzyme-inhibitor pairs. Apparent association rate constants, k_{ass}, were obtained from these plots and were additionally corrected for substrate competition using adequate K_m values. K_{diss} were obtained from the intercepts of slopes with the ordinate and K_i values were calculated from the obtained k_{ass} and k_{diss} values. If the intercept was too close to the origin of the axis, precluding an accurate determination of k_{diss}, K_i values were determined by the equilibrium method [12]. In these cases, k_{diss} values were calculated from $k_{diss} = K_i/k_{ass}$.

3. Results and discussion

3.1. Location of polymorphisms within Stfa1 and Stfa2 alleles

The sequences of the C57BL/6J and A/J *Stfa1* and *Stfa2* alleles are homologous to human *Stfa*. They posses all secondary structural elements constituting the fold of stefins, the N-terminal helix and four stranded beta sheets, as well as the structural features crucial for effective binding to the active sites of their target proteases (Fig. 1, using human *Stfa* nomenclature), the N-terminal trunk, the first hairpin loop with the characteristics of QVVAG motif (where the first V47 can be replaced by A and the A49 by Q) and the second hairpin loop [13,14].

The N-terminal trunk of *Stfa1-a* is one amino acid shorter than *Stfa1-b*, whereas the *Stfa2-a* and *Stfa2-b* alleles both possess a seven residue long N-terminal extension when compared to human *Stfa*. As can be seen from the theoretical models of *Stfa1* and *Stfa2*, most variable positions are located in the regions distant to the active site cleft (Fig. 2). Among the variable positions only positions 2 (S2 *Stfa1-b* – M2 *Stfa1-a* substitution) and 3 (L3 *Stfa1-b* – Y3 *Stfa1-a*, I3 *Stfa2-b* – M3 *Stfa2-a* substitution) at the N-terminal trunk and position 47 within the first hairpin loop, where V47 has been changed to A in *Stfa1-a*, are contributing to the direct interactions with the active site of the target enzymes. This short analysis suggests that there are likely no significant differences in folding properties between the predominant and variable forms of mouse *Stfa*, whereas a possibility exists that there are differences in their binding affinities primarily due to variability of the N-terminal trunk region.

3.2. Folding of mouse Stfa1 and Stfa2

Stfa1 and *Stfa2* from C57BL/6J and A/J mice were expressed in *E. coli* using a vector for cytosolic expression. Formation of appropriately folded structures was verified by assessing their far UV-CD spectra (Fig. 3). The spectra of *Stfa2* markedly differ from the spectra of *Stfa1* in that they exhibit an increased negative molecular ellipticity between 210 and 220 nm and decreased molecular ellipticity at lower wavelengths. This difference is indicative of an increase in disordered structure which is presumably due to the elongated N-terminal trunk of *Stfa2*. Aside from this difference, the spectra for *Stfa1* and *Stfa2* are similar to human *Stfa* spectra. Furthermore, the spectra