Biochimie 94 (2012) 318–327

Contents lists available at ScienceDirect

Biochimie

journal homepage: www.elsevier.com/locate/biochi

Research paper

Substrate specificity of *Staphylococcus aureus* cysteine proteases – Staphopains A, B and C

Magdalena Kalińska [a,1], Tomasz Kantyka [a,1], Doron C. Greenbaum [c], Katrine S. Larsen [d], Benedykt Władyka [b], Abeer Jabaiah [e], Matthew Bogyo [f], Patrick S. Daugherty [e], Magdalena Wysocka [g], Marcelina Jaros [g], Adam Lesner [g], Krzysztof Rolka [g], Norbert Schaschke [h], Henning Stennicke [d], Adam Dubin [b], Jan Potempa [a,i], Grzegorz Dubin [a,*]

[a] Department of Microbiology, Faculty of Biochemistry Biophysics and Biotechnology, Jagiellonian University, Gronostajowa 7, 30-387 Krakow, Poland
[b] Department of Analytical Biochemistry, Faculty of Biochemistry Biophysics and Biotechnology, Jagiellonian University, Gronostajowa 7, 30-387 Krakow, Poland
[c] Department of Pharmacology, University of Pennsylvania School of Medicine, Philadelphia, PA 19104-6018, USA
[d] Protein Engineering, Novo Nordisk A/S, 2760 Maaloev, Denmark
[e] Department of Chemical Engineering, University of California at Santa Barbara, Santa Barbara, CA 93106-5080, USA
[f] Department of Pathology, Stanford University School of Medicine, Stanford, CA 94305-5324, USA
[g] Faculty of Chemistry, University of Gdansk, Sobieskiego 18/19, 80-952 Gdansk, Poland
[h] Department of Chemistry, Bielefeld University, 33501 Bielefeld, Germany
[i] University of Louisville School of Dentistry, Oral Health and Systemic Disease, Louisville, KY 40202, USA

ARTICLE INFO

Article history:
Received 1 March 2011
Accepted 13 July 2011
Available online 23 July 2011

Keywords:
Protease
Staphopain
Staphylococcal virulence
Substrates library
Substrate specificity

ABSTRACT

Human strains of *Staphylococcus aureus* secrete two papain-like proteases, staphopain A and B. Avian strains produce another homologous enzyme, staphopain C. Animal studies suggest that staphopains B and C contribute to bacterial virulence, in contrast to staphopain A, which seems to have a virulence unrelated function. Here we present a detailed study of substrate preferences of all three proteases. The specificity of staphopain A, B and C substrate-binding subsites was mapped using different synthetic substrate libraries, inhibitor libraries and a protein substrate combinatorial library. The analysis demonstrated that the most efficiently hydrolyzed sites, using Schechter and Berger nomenclature, comprise a P2–Gly↓Ala(Ser) sequence motif, where P2 distinguishes the specificity of staphopain A (Leu) from that of both staphopains B and C (Phe/Tyr). However, we show that at the same time the overall specificity of staphopains is relaxed, insofar as multiple substrates that diverge from the sequences described above are also efficiently hydrolyzed.

1. Introduction

Staphylococcus aureus is an important human pathogen [1]. Its increasing antibiotic resistance presents a major clinical challenge, driving intense research into the physiology of the bacterium [2,3]. Multiple lines of evidence suggest that secreted proteases play a role in staphylococcal virulence [4,5]. Even so, the exact roles of particular enzymes are only beginning to be clarified [6–8].

An analysis of randomly generated *S. aureus* mutant strains in a murine infection model pointed to the importance of V8 protease in the virulence of *S. aureus* [9]. However, it was soon demonstrated that targeted V8 inactivation does not by itself attenuate virulence [10], but significantly affects the V8-dependent activation of staphopain B protease zymogen. The virulence of a staphopain B knockout strain was attenuated in a mouse model, although the V8 protease level was unaffected [8]. Because staphopain B is the final enzyme in the staphylococcal protease activation cascade [8], the results described above indicate that it contributes directly to the bacterium's pathogenicity. Moreover, it was demonstrated in a mouse model that elevated levels of staphopain B were produced

Abbreviations: Ac, acetyl; ABZ, amino benzoic acid; ANB-NH_2, amide of 5-amino-2-nitrobenzoic acid; ACC, 7-amino-4-carbamoylmethylcoumarin; CL, competition labeling; CLiPS, cellular library of peptide substrates; DTT, dithiothreitol; E-64, trans-epoxysuccinyl-L-leucylamido(4-guanidino)butane; eCPX, circularly permutated outer membrane protein X; FRET, fluorescence resonance energy transfer; GFP, green fluorescent protein; LSTS, library of synthetic tetrapeptide substrates; MBP, maltose binding protein; pNA, p-nitroanilide; PS-IL, positional scanning inhibitor library; PS-SCL, positional scanning synthetic combinational library.

* Corresponding author. Tel.: +48 12 664 63 62; fax: +48 12 664 6902.
E-mail address: grzegorz.dubin@uj.edu.pl (G. Dubin).

[1] These authors contributed equally to this work.

0300-9084/$ – see front matter © 2011 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.biochi.2011.07.020

in vivo during the infection with strains of Community Associated Methicillin Resistant *S. aureus* [11]. In contrast, a staphopain A knockout strain showed no marked difference in virulence when compared with the wild-type strain [8].

Apart from human infections, *S. aureus* is a significant cause of poultry diseases, and is a large economic burden on the broiler chicken industry. Strains isolated from chicken dermatitis lesions ubiquitously express staphopain C, whereas its expression has never been demonstrated in human strains [12]. Comparative studies by Takeuchi and colleagues [13] suggest that this protease is directly involved in the pathogenesis of the chicken disease.

Although staphopains A, B, and C share a high degree of sequence similarity [14], their functional divergence is well established by the data cited above and other evidence. Staphopains B and C play roles in human and avian diseases, respectively, whereas staphopain A seems to be unrelated to virulence, with a possible housekeeping function. We hypothesized that this functional variability may reflect the diverse substrate specificity of staphopains. Here, we report the use of multiple high-throughput profiling methods to draw comprehensive conclusions regarding the overall substrate preferences of staphopains A, B and C.

2. Materials and methods

2.1. Protein purification

Staphopains A and B were purified from the culture supernatant of *S. aureus* strain V8BC10 [15]. The bacteria were grown overnight in TSB medium supplemented with β-glycerophosphate (5 g/L) at 37 °C with shaking. The following purification steps were conducted at 4 °C. The cells were removed by centrifugation (10,000 × g, 20 min), and the proteins in the supernatant were precipitated with ammonium sulphate at 80% saturation (561 g/L) and collected by centrifugation. The resulting pellets were resuspended and dialyzed overnight against buffer A (50 mM sodium acetate, pH 5.5). After dialysis, the proteins were separated chromatographically on Q Sepharose FF (Amersham-Pharmacia) equilibrated with buffer A. Staphopain A was collected as the flow-through fraction. Staphopain B was eluted with a gradient of 0–300 mM NaCl in buffer A. Fractions containing staphopain B (assayed on Azocoll) (Merck) were supplemented with ammonium sulphate (final concentration 2 M) and applied to Phenyl Sepharose (Amersham-Pharmacia) in buffer B (50 mM Tris–HCl, pH 7.5) containing 2 M ammonium sulphate. Staphopain B was recovered using a linear gradient (from 2 M to 0 M) of ammonium sulphate in buffer B. Fractions containing the active enzyme were pooled, dialyzed against buffer B, and stored frozen at −20 °C until analysis. Staphopain A was further purified on CM Sepharose FF (Amersham-Pharmacia) using a linear gradient of 0–300 mM NaCl in buffer A. Fractions containing staphopain A (assayed on azocasein) were pooled, dialyzed against buffer B, and kept frozen at −20 °C until further analysis.

Staphopain C was obtained from the culture supernatant of *S. aureus* strain CH-91 [16]. The initial purification steps were identical to those for staphopains A and B. The pellets obtained after ammonium sulphate precipitation were dissolved and dialyzed against buffer C (20 mM phosphate buffer, pH 8.0). Staphopain C was purified on Q Sepharose FF using a linear gradient of 0–0.5 M NaCl in buffer C. Fractions containing the proteolytic activity (assayed on Suc–GFG–pNa) were further purified by gel filtration on Superdex 75 (Amersham-Pharmacia; 5 mM Tris–HCl, 50 mM NaCl, pH 8). The samples were stored at −20 °C until analysis.

Preparations with at least 95% purity, as assayed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE), were used for the experiments. The protein concentrations were determined with the BCA assay (Sigma). The concentrations of the active enzymes were determined by titration with E-64 (staphopain A, staphopain C) or with staphostatin B (staphopain B) [17]. All the specified enzyme concentrations are expressed as active enzyme concentrations. Only preparations with more than 80% active enzyme were used in the experiments.

Different variants of the MBP–linker–GFP fusion protein (Table 3) were obtained and purified as described previously for MBP–GFP [18]. In brief, the CLiPS consensus sequences were engineered into the linker region by site-directed mutagenesis. The sequence selected in CLiPS by staphopain B (LAFGA) was used as a reference and was modified according to CLiPS data to generate substrates for staphopain A (LVLGA) and staphopain C (LVFGA). Fusion proteins were expressed in *Escherichia coli* and purified by affinity chromatography on Chelating Sepharose FF (Pharmacia). The pure proteins were obtained after the Source Q (Pharmacia) purification step.

2.2. General activity assays

Before the protease activity was assessed, each sample was activated in assay buffer (50 mM Tris–HCl, pH 7.6, supplemented with 2 mM DTT and 5 mM EDTA) for 15 min at 37 °C. Staphopains A and B were routinely assayed on chromogenic substrates. Stock suspensions (15 mg/mL) of Azocoll, hide powder azure, and Congo Red–elastin were prepared in 0.6 M sucrose, 0.05% Triton X-100, 0.02% NaN_3, and azocasein was prepared as a 3% (w/v) solution in deionized water. The substrate solutions were diluted with assay buffer to a final concentration of 5 mg/mL or 1% for azocasein. Staphopain C was assayed using Suc–GFG–pNA at a final concentration of 1 mM. All the enzymes were assayed at mid-nanomolar concentrations. The reactions were developed for 1 h (Azocoll, hide powder azure, azocasein, and Suc–GFG–pNA) or overnight (Congo Red elastin) at 37 °C with shaking. The optical density of the supernatant was determined as a measure of enzyme activity (at 520 nm for Azocoll, 495 nm for Congo Red–elastin, 595 nm for hide powder azure, 360 nm for azocasein, and 405 nm for Suc–GFG–pNA).

2.3. Competition labeling of the S2 subsite

The S2 substrate binding pocket of the staphopains was probed with a library of synthetic, peptidomimetic inhibitors, as described previously [19]. In brief, a small molecule library based on E-64 was tested, which consisted of 60 sub-libraries, in each of which a particular natural or non-natural residue (Table S1, see Supplemental Information) was fixed in position P2 and the remaining subsites contained an equimolar mixture of all the natural residues (Cys was excluded, Met was substituted with norleucine). Staphopains A and B (0.5 μM) were pre-incubated separately with each P2 sub-library (10 μM) in 50 mM Tris–HCl (pH 7.6) containing 2 mM DTT for 30 min at room temperature. The samples were then labeled with ^{125}I-DCG-04, a derivative of E-64 [20] (approximately 10^6 cpm per sample) for 1 h at room temperature. The samples were resolved by SDS–PAGE and analyzed by phosphoimaging. Numerical values for the percentage competition were normalized to an untreated control sample and visualized using the programs TreeView and Cluster [21], as described previously [22].

2.4. P1′ positional scanning inhibitor library (PS-IL) and P2′ preference determination

PS-IL, containing structural derivatives of E-64, was synthesized and assayed as previously described [23]. The P1′ library contained

19 sub-libraries, each containing a specific residue of the 19 natural amino acids (excluding Met and Cys, and including norleucine) for the interaction with the protease S1′ subsite and an equimolar mixture of those residues to interact with the S2′ subsite. The P2′ preference was determined with the P2′ subgroup of 19 individual inhibitors, with an Ala residue fixed in the P1′ position and a specific proteinogenic amino acid in position P2′ (Cys was excluded, Met was substituted with norleucine). The inhibitory activity of either the library compounds or the P2′ subgroup (10 μM) against staphopain A (75 nM), staphopain B (120 nM), or staphopain C (350 nM) was determined by incubating each enzyme with the inhibitor for 30 min at 37 °C, followed by an assessment of the residual activity using the Azocoll assay (staphopain A and B) or the synthetic substrate Suc–GFG–pNa (staphopain C).

2.5. Libraries of synthetic tetrapeptide substrates (LSTS)

The LSTS are fluorescence-quenched substrates with four variable positions (P4–P3–P2–P1), flanked by an N-terminal ABZ (fluorophore) and a C-terminal ANB–NH_2 (quencher). The library preparation and testing have been described previously [24,25]. In brief, 19 sub-libraries were tested for each enzyme, each containing a P4 position fixed with one of the 19 natural amino acid residues (except Cys) and positions P3–P1 containing equimolar mixtures of these residues. Each sub-library (3 mg/mL final concentration) was incubated for 10 min with activated enzyme (1.3 μM). The increase in fluorescence resulting from the excitation of the ABZ moiety (325 nm) was monitored at 400 nm (emission). P4 was then fixed with the residue corresponding to the most efficiently hydrolyzed substrate and 19 sub-libraries were synthesized, in each of which the P3 position contained one of the 19 natural amino acid residues and positions P2–P1 contained equimolar mixtures of these residues. The scheme was iterated until all the positions of the most efficiently hydrolyzed substrate were deconvoluted for a particular enzyme. The best substrate determined for each enzyme was resynthesized and the site of hydrolysis was established by liquid chromatography–mass spectrometry (LC–MS). The kinetics of the hydrolysis of each substrate were assessed as described previously [26,27].

2.6. Positional scanning synthetic combinational libraries (PS-SCL)

The substrate specificity of the staphopains was assessed using PS-SCL, as described previously [28]. In brief, all the libraries contained fluorescent substrates with the general structure Ac–P4–P3–P2–P1–ACC. The P1 substrate preference was assessed using a library containing 18 sub-libraries, in each of which the P1 position contained one of the natural amino acids (except Met and Cys), and an equimolar mixture of those amino acids occupied the remaining subsites. The P2–P4 substrate preference was evaluated using three distinct libraries. Each library contained 18 sub-libraries, each having fixed residues at either the P2, P3, or P4 position, and all having Arg fixed at P1. The remaining positions were randomized with a mixture of natural amino acids, excluding Met and Cys. All the libraries were tested at a substrate concentration of 9 μM in a buffer containing 100 mM Tris–HCl (pH 7.6), 200 mM NaCl, 10 mM cysteine, and 5 mM $CaCl_2$. The active enzyme concentration was 100 nM for staphopains A and B and 9 μM for staphopain C. Enzymatic activity was monitored as the increase in fluorescence emission at 455 nm (380 nm excitation).

2.7. Cellular library of peptide substrates (CLiPS)

The complete consensus sequence recognized and cleaved by each staphopain was determined using the CLiPS methodology, as described previously [29]. In brief, a library based on the eCPX protein containing on the N-terminal the streptavidin-binding peptide ligand, followed by a substrate sequence composed of eight randomized amino acids, and the SH3 domain binding peptide on C-terminal, was displayed on the cell surface of *E. coli* [30]. The initial library contained ~10^8 substrates (different clones; diversity of the library was assessed by sequencing of randomly selected clones). The samples were incubated with the enzymes for 1 h at 37 °C and the concentrations of the enzymes were gradually reduced from 650 nM to 65 nM at each sorting step. The library was enriched with protease-recognized clones in 0.2 M Tris–HCl (pH 7.6) for staphopains A and B, and in 5 mM Tris–HCl (pH 8.0), 50 mM NaCl for staphopain C. The staphopain-induced hydrolysis of the substrates and the display of the eCPX protein was monitored with fluorescence-activated cell sorting (FACS). Clones with intact baits displayed red and green fluorescence after incubation with phycoerythrin-conjugated streptavidin (50 nM) and GFP-SH3 (250 nM), respectively. Clones expressing substrates that were specifically hydrolyzed were isolated by sorting for green cells only. After several rounds of sorting, the individual clones were incubated with 120 nM the appropriate staphopain for 1 h and the clones susceptible to proteolysis were sequenced. The data were analyzed to determine the consensus sequence recognized by each protease tested.

2.8. Cleavage of MBP–linker–GFP fusion proteins

Three variants of the MBP–linker–GFP fusion protein (Table 3) were prepared as described in section 2.1. An appropriate variant was cleaved by a particular staphopain at a molar ratio of 1:200, in 25 mM Tris (pH 7.6) supplemented with 2 mM DTT at 37 °C. After incubation for 1 h, the reaction products were separated by SDS–PAGE, transferred to PVDF membrane (Hybond), and the newly released N-terminal was characterized by Edman degradation sequencing (BioCentrum).

3. Results

3.1. Assessment of the S2 subsite preference by competition labeling (CL)

Based on *in vivo* data that suggest that staphopain B plays a significant role in staphylococcal virulence and that staphopain A probably has a housekeeping function, we hypothesized that these two homologous enzymes should exhibit different substrate specificities. To roughly identify the expected differences, we first assessed the S2 subsite preference of the two enzymes (subsite nomenclature after [31]). The rationale for assessing the S2 subsite first was that staphopains belong to the clan CA (papain-like) proteases [32], in which the S2 subsite is a major determinant of specificity [33]. An inhibitor library based on the scaffold of a general inhibitor of clan CA proteases (E-64) was constructed containing 61 P2-residue-fixed sub-libraries (Fig. 1A). The competition of each sub-library with the general radioactive probe ^{125}I-DCG-04 was assessed. The reduction in labeling upon incubation with the particular sub-library, compared with the labeling of a blank sample (without inhibitor), was used as the measure of activity.

The overall substrate preferences of staphopains A and B at the S2 subsite were not strikingly different (Fig. 1B) and most closely resembled that of human cathepsin H (Fig. 1C). Both staphopains most readily accepted the branched side chains of Leu, Ile, and Val at the S2 subsite. Nonetheless, marked differences were also evident. Staphopain B readily accepted some non-natural bulky side chains (phenyl glycine, 2-thienyl alanine) at S2, which were

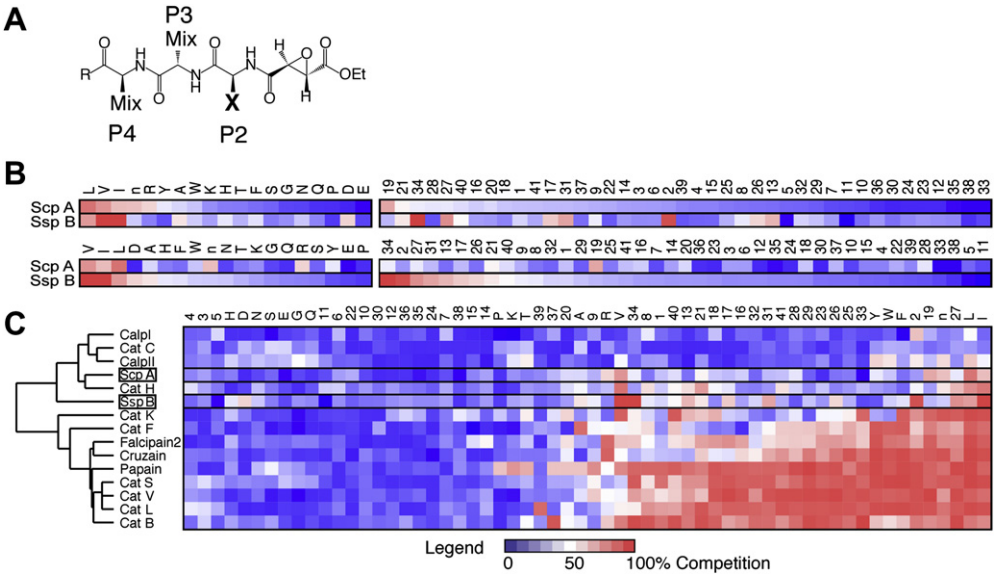

Fig. 1. Comparison of the S2 subsite preferences of staphopains A and B. The competition between different E-64 derivatives and a general radioactive probe to label the active sites of the staphopains was assessed. (A) The general structure of the inhibitor sub-library. Each tested sub-library contained a fixed residue in the P2 position (X) and equimolar mixtures of 18 natural amino acids (Mix) at the P3 and P4 positions. (B) The colour-coded activity of each of the 60 sub-libraries tested. Inhibitors were arranged according to the extent of competition with staphopain A labeling (top) and staphopain B labeling (bottom). (C) Comparative cluster analysis of the S2 subsite preferences of the staphopains and other papain-like proteases. The tree structure illustrates the similarities in the S2 subsite specificities of the CA clan enzymes. Capital letters indicate proteinogenic amino acid residue fixed at the P2 position of a particular sub-library; "n" stands for norleucine. Numbers designate non-natural amino acids (for chemical structures, see Table S1 in the Supplementary Data). The deep red colour indicates 100% competition (high activity of a sub-library) and the dark blue colour indicates 0% competition.

not tolerated by staphopain A. Of possible physiological relevance, staphopain B preferred Ile over Leu, whereas staphopain A had the reverse preference. Moreover, Asp was only accepted by staphopain B, in contrast to staphopain A, which accommodated Arg and Tyr at the analyzed subsite. Altogether, these results were considered as a proof of concept for our hypothesis. The differences in the substrate preferences of the enzymes analyzed were not striking, but were clearly distinguishable. Therefore, we have proceeded to further characterize differences in the specificities of staphopains A, B, and C.

3.2. Mapping the primed site amino acid preference

To gain greater insight into staphopain substrate specificity, we verified their S1′ and S2′ subsite preferences. For this purpose, an extended version of the CL experiment with PS-IL was used. The library was based on E-64 derivatives (Fig. 2A) assembled into two groups of structures. The first group contained 19 sub-libraries, each with a fixed residue in the P1′ position and an equimolar mixture of the 19 natural amino acids at position P2′ (omitting Cys; Met was substituted with norleucine). The second group was synthesized as components of the P1′ sub-library, with an Ala residue fixed in this position and a specific amino acid in position P2′ (Cys was excluded and Met was substituted with norleucine). Each group was incubated with the enzymes, and the residual enzymatic activity was determined.

In this assay, staphopain A exhibited almost no selectivity at the P1′ subsite, whereas the two other staphopains were more selective (Fig. 2B). Staphopain B preferred Ala and Gly, whereas staphopain C preferred Leu, Ile, and Val. Inhibitors containing Gln, Asn, His, or Asp at P1′ were the least effective against all three staphopains. These inhibitors were almost ineffective against staphopain B, whereas they retained more than 50% activity against staphopains A and C under the conditions used.

All three staphopains strongly selected Ala at the P2′ position. This preference was most evident for staphopain B, for which all the tested inhibitors containing residues other than Ala were less than 50% active. The selectivity of staphopain C was less clear because all the inhibitors, except that containing Ser at P2′, were still more than 50% effective. Staphopain A was least selective of all the enzymes analyzed, accepting not only Ala but also Trp and Tyr, and to a lesser extent Phe, Ile, and Leu, at the P2′ position.

Overall, in the assay described, staphopain B exhibited the highest selectivity at both the P1′ and P2′ positions, followed by staphopain C. Staphopain A was the least selective of the three enzymes.

3.3. Specificity profiling with synthetic substrate libraries

The competition experiments roughly confirmed our initial assumption that the substrate preferences of staphopains A, B, and C differ. To extend these findings more comprehensively, we profiled the substrate specificities of the staphopains using an LSTS with the general formula ABZ–P4–P3–P2–P1–ANB–NH_2. The activity of all the tested enzymes was initially determined against a library composed of 19 sub-libraries, each containing a fixed residue at the P4 position and equimolar mixtures of the 19 natural amino acids (excluding cysteine) at the other positions. Then a second library was constructed for each protease, in which P4 contained the residue yielding the highest activity in the previous test, and containing 19 sub-libraries, each with a fixed residue at the P3 position and an equimolar mixture of the 19 natural amino acids at P2 and P1. This scheme was followed until the most efficiently hydrolyzed substrate (residues P4–P1) was determined for each protease (Fig. 3). The cleavage sites in the deconvoluted, most efficiently hydrolyzed substrates were then determined with LC–MS (Fig. S2, see Supplemental Information). Staphopain A (optimal substrate: ABZ–Phe–Gly–Ala–Lys↓ANB–NH_2) and staphopain B (ABZ–Ile–Ala–Ala–Gly↓ANB–NH_2) exhibited similar substrate preferences in this assay, preferably hydrolyzing substrates with Ala residue in the P2 position. The most preferred residues in the other positions distinguish the two enzymes. Staphopain C exhibited a distinct substrate preference, most efficiently hydrolyzing ABZ–Ile–Ala–Lys–Asp↓ANB–NH_2. It is noteworthy

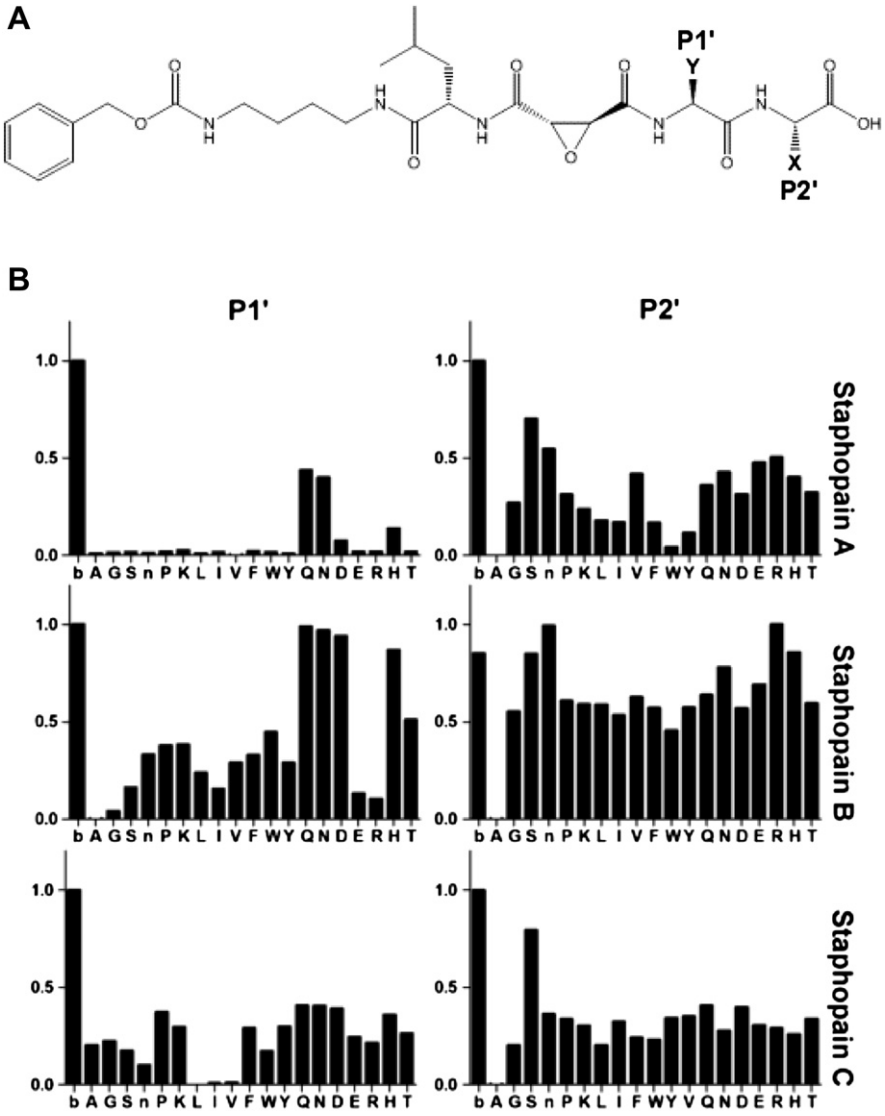

Fig. 2. Substrate preferences of staphopains at the primed subsites. The residual activity of the staphopains was determined in the presence of different E-64 derivatives to assess the P1′ and P2′ residue preferences of the proteases. (A) Markush structure representation of the general structure of a single sub-library of the Positional Scanning Inhibitory Library (PS-IL) used in this study. Each sub-library of the P1′ library contains a fixed residue at position Y and an equimolar mixture of the 19 tested amino acids at position X. Each individual inhibitor of P2′ subgroup contains an Ala residue fixed at position X and one of 19 amino acids at position Y. (B) The activity of the P1′ and P2′ inhibitor sub-libraries against staphopains A, B, and C. The residual proteolytic activity (*y* axis) after pre-incubation with a particular inhibitor sub-library (*x* axis – the fixed residue is represented by the single-letter amino acid code; "n", norleucine). The residual activity was normalized to the control, which lacked inhibitor ("b").

that, when determined in this way, the specificities of staphopains A and B are consistent with the established selectivity of the clan CA proteases at the S2 subsite. The LSTS assay demonstrated the relaxed preference of staphopain C at this position, suggesting a broader specificity of this enzyme.

The optimal substrates established for each enzyme were resynthesized and their kinetic parameters were determined (Table 1). For each substrate, the k_{cat}/K_M value was only significant for the enzyme for which the substrate was initially developed, confirming the differences in the substrate specificities of the proteases analyzed. Their virtual lack of cross reactivity means that the substrates developed should be useful tools for staphopain characterization.

Because the LSTS profiling results were largely inconsistent with the inhibitor-based library profiling reported here, we assessed how the deconvolution scheme and the type of reporter group affected the preferences established at particular subsites. A PS-SCL was constructed with the general structure Ac–P4–P3–P2–P1–ACC. The deconvolution scheme differed from that used for LSTS library in a way that P1 subsite preference was determined first and followed by parallel determination of preference at remaining subsites. For P1 profiling, the library contained 18 sub-libraries, each of which contained a fixed residue in the P1 position and equimolar mixtures of 18 proteinogenic amino acids (excluding Met and Cys) at positions P2–P4. The modified deconvolution scheme yielded results only partially consistent with the LSTS profiling. Both methods demonstrated selectivity of staphopain A for Thr at the P1 subsite, but they varied in other residues likewise efficiently hydrolyzed. The PS-SCL assay also indicated Gln and Ala as amino acids selected in the analyzed position (Fig. 4), whereas this preference was less clear in the LSTS assay, where Lys or Leu were processed at the highest rate (Fig. 3). In the PS-SCL assay, staphopain B exhibited strict specificity for residues with

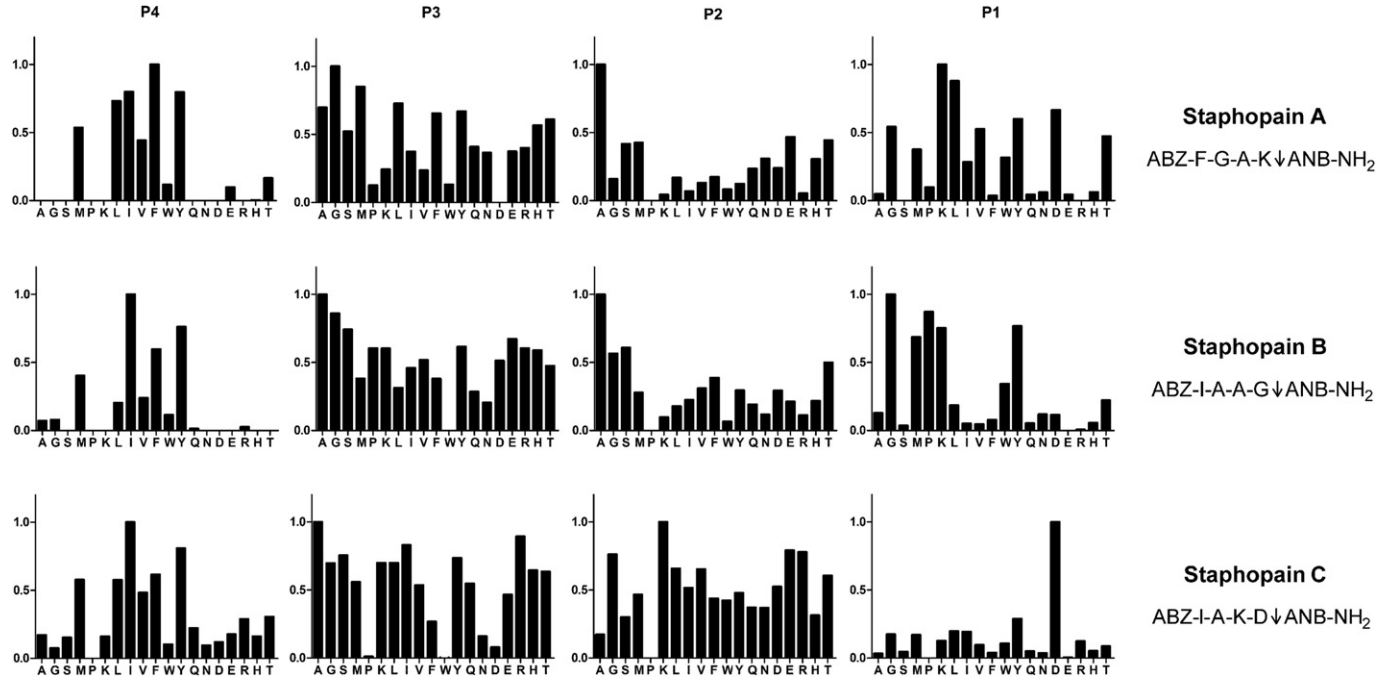

Fig. 3. Comparison of the substrate specificities of staphopains A, B, and C determined using Library of Synthetic Tetrapeptide Substrates (LSTS). The substrate specificities of the staphopains were profiled using substrates with the general structure ABZ–P4–P3–P2–P1–ANB–NH_2, as described in Section 2.5. For each enzyme, the specificity at particular subsite, the optimal substrate sequence, and the cleavage site (indicated with "↓") are shown. Vertical bars indicate the activity of the enzyme against a particular sub-library, normalized to the most active sub-library in each set. Residues fixed at particular subsites (indicated at the top of each panel) are indicated with the single-letter amino acid code.

small side chains (Gly, Ala, Ser) at the P1 subsite. In fact, no other residues were accepted. Even though Gly was also the most preferred residue at P1 in the LSTS assay, many other residues were additionally accepted with a significant yield of substrate hydrolysis products. The results for staphopain C were most inconsistent between the two methods. Whereas LSTS revealed a strong preference for Asp at P1, the PS-SCL assay indicated a relaxed preference at this subsite, with Asp being one of the least preferred residues.

The other subsites were only profiled for staphopains A and C. The library used for profiling the P2 preference contained 18 sub-libraries, each having one of 18 natural amino acids fixed at P2 and Arg at P1. Positions P3 and P4 contained equimolar mixtures of the 18 natural amino acids. Contrary to the broad preference at the S2 subsite observed with LSTS, PS-SCL revealed high specificities for the tested staphopains. Staphopain A strongly preferred Val (also accepting Phe, Pro, and Leu), whereas staphopain C selected Tyr (also accepting Phe and Leu) as a preferred P2 residue. The presence of residues other than those mentioned almost completely abolished the enzymatic activity of the staphopains against the tested substrates.

At the other positions (P3 and P4), staphopain A exhibited a preference for Glu at P3, whereas staphopain C was rather unspecific. Both enzymes showed little preference at the P4 subsite.

Taking these data together, a comparison of the substrate specificities of the tested staphopains determined using LSTS and PS-SCL indicated that the substrate preference is assay specific, rather than protease specific. Nonetheless, a comparison of the results for the different enzymes within each of the two assays demonstrated that the specificities of the staphopains differ, which probably reflects true differences between the enzymes.

3.4. Comprehensive specificity profiling using CLiPS

The synthetic substrate/inhibitor libraries did not yield a consistent picture of staphopain specificity because they were affected by positional deconvolution and the reporter groups. Therefore, we assessed the specificity of the staphopains using a bacterial cell-surface display system based on proteinacious substrates, CLiPS. The library was constructed with a highly variable substrate sequence of eight randomized constitutive amino acids in the context of an outer-membrane bacterial protein. Overall, the

Table 1
Kinetic parameters of the hydrolysis by staphopains of tetrapeptide substrates determined in the LSTS assay.

Substrate sequence	Enzyme	k_{cat} [s^{-1}] × 10^{-1}	K_M [M] × 10^{-6}	k_{cat}/K_M [$s^{-1}\,M^{-1}$] × 10^3
ABZ–Phe–Gly–Ala–Lys–ANB–NH_2	Staphopain A	7.16 ± 0.83	5.6 ± 0.4	127.8 ± 7.4
	Staphopain B	1.12 ± 0.47	154.6 ± 15.5	0.7 ± 0.1
	Staphopain C	0.23 ± 0.03	385.3 ± 20.4	Below 0.1
ABZ–Ile–Ala–Ala–Gly–ANB–NH_2	Staphopain A	0.56 ± 0.03	163.9 ± 12.2	3.4 ± 0.2
	Staphopain B	8.86 ± 6.02	7.6 ± 1.1	118.0 ± 12.2
	Staphopain C	0.02 ± 0.01	271.4 ± 10.2	Below 0.1
ABZ–Ile–Ala–Lys–Asp–ANB–NH_2	Staphopain A	0.32 ± 0.11	587.1 ± 48.1	Below 0.1
	Staphopain B	0.12 ± 0.07	612.9 ± 52.7	Below 0.1
	Staphopain C	0.89 ± 0.58	14.3 ± 1.35	62.4 ± 3.1

Values are the means ± SD of triplicate experiments.

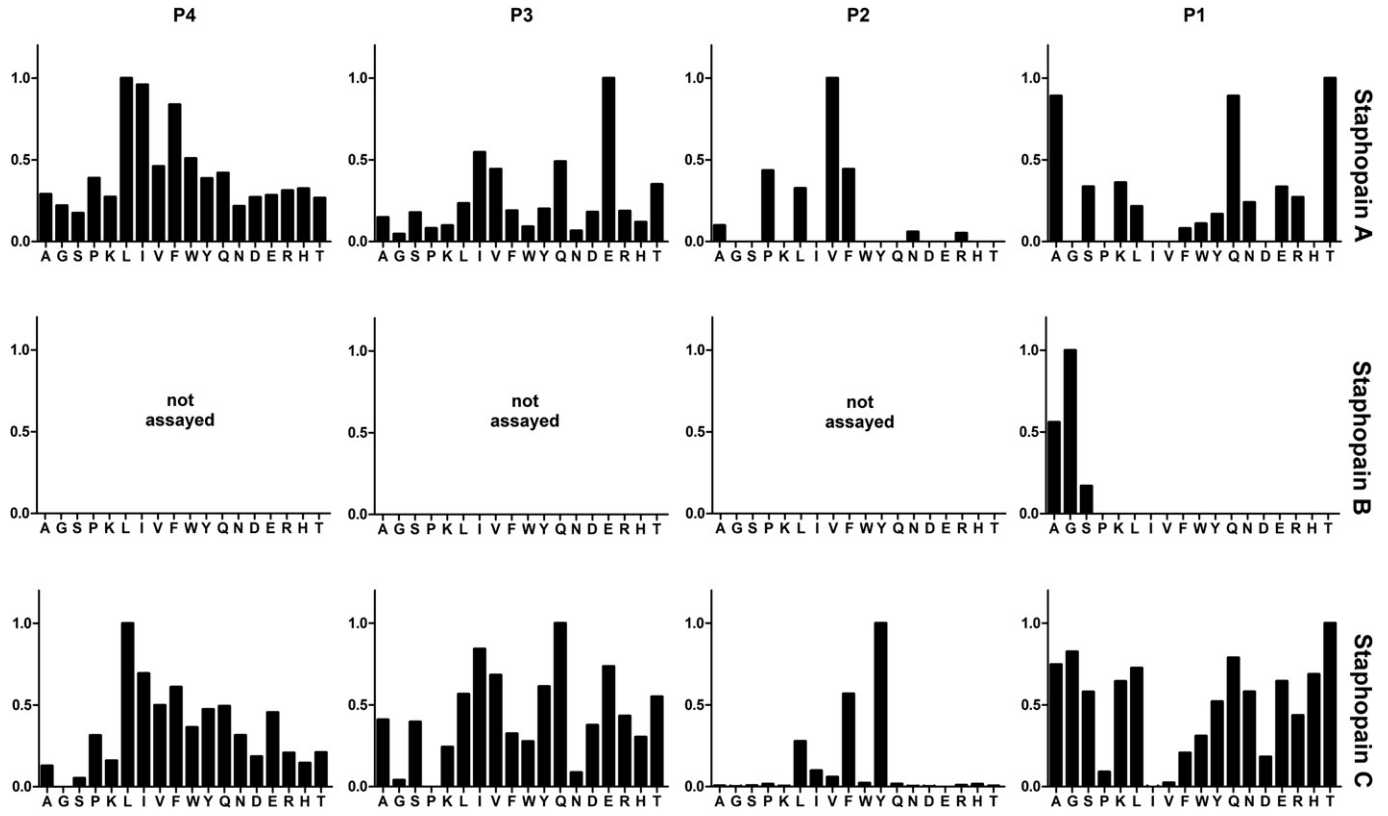

Fig. 4. Substrate specificities of staphopains A, B, and C determined using Positional Scanning Synthetic Combinational Libraries (PS-SCL). The substrate specificities of the staphopains were profiled using a PS-SCL library with the general structure Ac–P4–P3–P2–P1–ACC. Eighteen sub-libraries, each with a fixed P1 residue and equimolar mixtures of residues at all other subsites, were assayed to determine the P1 preference of each enzyme. The P2, P3, and P4 sub-libraries were similar to P1, except that the residues at the respective subsites were fixed and all the substrates contained Arg in the P1 position. Vertical bars indicate the activities of the enzymes against a particular sub-library, normalized to the most active sub-library in each set. Residues fixed at particular subsites (indicated at the top of each column) are identified with the single-letter amino acid code.

tested library contained 10^8 different substrates (out of 3.7×10^{10} possible combinations). Multiple rounds of selection, comprising a protease treatment and the isolation of the hydrolyzed clones, were performed to reveal the consensus sequences recognized and hydrolyzed by staphopains A, B, and C (Table 2). The exact cleavage site within the consensus was determined by analyzing the hydrolysis of purified fusion proteins containing sequences derived from the CLiPS selection (Table 3). Taken together, the general sequence pattern recognized and cleaved by all three staphopains can be roughly presented as: P3–P2–Gly↓Ser(Ala). The P2 preference distinguishes staphopain A from staphopains B and C. Both of the latter enzymes prefer the bulky aromatic side chains of Phe or Tyr at P2, whereas staphopain A preferentially recognizes and cleaves substrates containing Leu in this position. The preference at P3 is less stringent and is similar for all enzymes, with middle-sized residues including Leu, Ile, Val, Ser, Thr, and Ala being selected. The specificity was most stringent for staphopain C, which almost exclusively selected Leu, Ile, or Val. Staphopain B preferred Ser, Thr, Ala, and Pro at this site, and Pro was only specific for this enzyme. Staphopain A was the least selective and all the residues mentioned above (except Pro) were acceptable in the P3 position. The above described preferences are, however, only discriminatory under a high selection pressure (CLiPS). The test fusion proteins (Table 3) containing either one of optimal sequences selected for staphopain B (LAFG↓A), or general pattern most preferred by staphopain A (LVLG↓A) or substrate sequence modified to reflect staphopain C preference (LVFG↓A), are equally well hydrolyzed by all three enzymes, at least to the extent possible to distinguish with a semiquantitative assay used in this study.

4. Discussion

Staphopains B and C contribute to staphylococcal virulence, whereas staphopain A seems to have another unrelated function. We hypothesized that mapping their substrate specificity may help explain the observed differences in physiological functions.

Staphopains A and B have been extensively studied over the years. The degradation of the α_1-inhibitor [34] and few synthetic substrates [35] has been investigated, but because of the small pools of analyzed cleavage sites, no conclusive results regarding the enzymes' substrate preferences have been obtained, apart from the fact that staphopains hydrolyze a broad range of substrates. Structural analysis of the staphostatin B-staphopain B complex cast more light on the substrate preference of the protease [36], but at the same time contradicted the previous fragmentary findings based on the analysis of cleavage sites, so again no overall consistent picture emerged. Staphopain C was first described only recently [13] and no data on its substrate preferences are available. In this report, we have described the results of the first focused study of the substrate preferences of the staphopains, using currently available high-throughput profiling methods: a combinatorial inhibitor library, and synthetic and proteinacious substrate libraries. Each specificity mapping method showed clear differences at the individual substrate-binding subsites in the characterized enzymes. However, the results obtained for each particular enzyme with the different methods were largely inconsistent. This phenomenon might be explained by a combination of the drawbacks inherent in each method and the overall low substrate specificity of analyzed proteases. For each profiling method used in

Table 2
Sequences of the most efficiently hydrolyzed protein substrates determined by CLiPS for staphopains A, B, and C.

	P4	P3	P2	P1	P1′	P2′
Staphopain A						
SORT 5[a]	L	**L**	**L**	**g**	g	s
	L	T	**L**	E	S	A
	F	R	**L**	T	A	g
	R	S	**L**	**g**	g	s
	V	S	**L**	**g**	g	s
	Q	**L**	**L**	T	S	g
	L	**V**	**L**	**g**	g	s
	W	A	**L**	A	**S**	P
	A	**L**	**L**	**g**	g	s
	I	**V**	Y	**g**	g	s
	L	S	F	**G**	g	g
	T	S	**L**	M	M	g
	F	G	**L**	S	g	g
	S	W	**L**	S	T	V
	V	**V**	**L**	**g**	g	s
	L	**L**	**L**	**g**	g	s
	L	**V**	**L**	**g**	g	s
SORT 7	L	T	**L**	**G**	A	S
	L	**V**	**L**	**G**	**S**	S
	L	**V**	**L**	**G**	**S**	S
	R	**L**	**L**	**G**	**S**	S
	R	**L**	**L**	**G**	**S**	S
	M	Q	**L**	**G**	**S**	S
	L	**V**	M	T	**S**	g
	G	**L**	D	**G**	M	I
	R	**V**	R	**G**	H	F
	L	**V**	F	**g**	g	s
		L/V	**L**	**G**	**S**	–
Staphopain B						
SORT 5	**L**	A	**F**	A	**A**	S
	L	S	**F**	**g**	g	s
	V	S	**F**	**g**	g	s
	L	S	**F**	**g**	g	s
	S	T	**Y**	**g**	g	s
	G	P	I	**g**	g	s
	T	P	A	**G**	S	S
	L	P	**V**	**g**	g	s
	I	P	**Y**	**g**	g	s
	G	P	**Y**	**g**	g	s
	Q	A	**Y**	**g**	g	s
	M	S	**Y**	**g**	g	s
	T	A	**Y**	**g**	g	s
	g	A	**Y**	**G**	**A**	Q
	L	S	**Y**	**g**	g	s
	L	L	**Y**	**g**	g	s
	V	V	A	**g**	g	s
	V	V	g	**g**	s	g
	I	V	**F**	**g**	g	s
SORT 7	**L**	A	**F**	**G**	**A**	H
	L	A	**F**	**G**	S	g
	L	I	**F**	**g**	g	s
	H	P	**F**	**G**	S	R
	W	S	**F**	**G**	S	P
	I	A	**F**	**g**	g	s
	V	A	**F**	**G**	g	g
	V	K	**F**	**G**	**A**	Q
	g	S	**F**	**G**	**A**	R
	L	V	**Y**	**G**	**A**	g
	I	T	**Y**	**G**	**A**	S
	I	T	**Y**	**G**	**A**	S
	I	T	**Y**	**G**	**A**	S
	I	T	**Y**	**G**	**A**	S
	I	T	**Y**	**G**	**A**	S
	I/L		**F/Y**	**G**	**A**	
Staphopain C						
SORT 4	g	Y	**Y**	**G**	**S**	A
	L	Q	**Y**	**G**	**S**	S
	E	**V**	**F**	**G**	**S**	S
	V	**L**	**F**	**G**	T	A
	H	**L**	**Y**	**g**	g	s
	T	**V**	**Y**	A	V	P
	W	**L**	K	S	I	S
	V	T	**Y**	**g**	g	s
	A	**L**	**Y**	S	g	g
	P	**I**	**Y**	A	Y	g
	L	**I**	L	**G**	**A**	D
	W	**V**	L	**G**	**A**	V
	A	**V**	**Y**	**g**	g	s
	L	T	**F**	**g**	g	s
	P	K	**Y**	**g**	g	s
SORT 6	g	**I**	**Y**	S	W	F
	M	**V**	**F**	**G**	**A**	g
	g	**V**	**F**	**G**	**A**	P
	A	**V**	**F**	**G**	**A**	P
	A	**V**	**F**	**G**	H	P
	Y	**V**	**F**	**G**	**S**	G
	L	**I**	**F**	**G**	**S**	P
	L	**L**	**F**	**G**	**S**	P
	M	**L**	**F**	**G**	**S**	P
	F	**I**	L	**G**	**S**	P
	I	**I**	**Y**	**G**	**S**	A
	V	**I**	**Y**	**G**	L	S
	V	**I**	**Y**	**G**	**A**	g
	L	**L**	**Y**	**G**	T	S
	L	**V**	**Y**	A	F	g
	N	**V**	**Y**	**G**	W	g
	E	**L**	**Y**	**g**	g	s
		L/V/I	**F/Y**	**G**	**S/A**	

Alignment of sequences selected from a library of randomized muteins of eCPX protein based on their susceptibility to hydrolysis by a particular staphopain. Capital and small letters indicate the variable and constant regions of the library, respectively. The consensus sequence of the most efficiently hydrolyzed substrate is given at the bottom of each sequence alignment. The residues highlighted in bold correspond to the determined consensus sequence. Only fragments of selected sequences corresponding to the consensus region are shown. Full sequences are listed in Supporting Information (Table S3).

[a] The sorting round in which the particular sequence originated is indicated on the left of each column.

this study, multiple examples are available in the literature documenting its suitability for the determination of protease substrate preferences [23,29,37,20,38]. Profiling results obtained with different synthetic substrate methods are usually consistent for highly selective proteases and less consistent for promiscuous enzymes. In the highly selective proteases, enzymatic specificity for particular residues plays a major role in substrate selection. However, in the less specific proteases, the final substrate selection results from interactions between the peptidic and non-peptidic parts of the synthetic substrates. For example, it has been demonstrated that identical synthetic substrates, varying only in the C-terminal chromophore group, display different kinetic parameters [39,40]. Our study with synthetic substrates showed that the substrate specificity of the staphopains is rather relaxed, although subtle differences between particular enzymes still

Table 3
MBP–linker–GFP fusion protein variants used in this study.

Linker sequence	Specificity
sLVLG↓Am	Staphopain A
sLAFG↓Am	Staphopain B
sLVFG↓Am	Staphopain C

Linker sequences were engineered according to the CLiPS consensus determined for each particular staphopain. "↓" indicates the experimentally determined cleavage site. Capital letters indicate CLiPS derived consensus-like region, small letters indicate flanking vector sequences.

exist. We confirmed those differences by profiling the enzymes with proteinaceous substrates. All three staphopains recognize a sequence of three consecutive amino acids. Amino acids with aliphatic side chains, including Leu, Ile, Val, and Thr, are selected at the P3 subsite by all three enzymes. Staphopain B also accepts Pro in this position. All the tested enzymes strongly prefer residues with small side chains in the P1 (Gly) and P1′ positions (Ala or Ser). Only the P2 subsite strongly distinguishes the substrate specificity of the staphopains. Both staphopains B and C prefer the bulky aromatic side chains of Phe and Tyr, whereas staphopain A prefers Leu at this subsite. Although the CLiPS results are largely inconsistent with those obtained with synthetic substrates and are only partially consistent with those obtained with inhibitor profiling, we believe that they best characterize the differences between the staphopains. First, the substrates used (i.e., proteins) are closest to those physiologically encountered by the enzymes. Moreover, no synthetic groups are present. Furthermore, no deconvolution is required, so subsite cooperation is taken into account.

It is interesting to compare the CLiPS specificity profiling data with the analysis of known crystal structures of staphopains A and B (staphopain C structure has not been determined to date). The most pronounced difference in specificity of staphopains A and B is found at S2 subsite where the former prefers a medium sized side chain of Leu whereas the latter bulky side chains of Phe and Tyr. In agreement, the structure of staphopain A in complex with E64 inhibitor (PDB code 1CV8) shows the S2 pocket of hydrophobic nature, where Leu is easily accommodated [41]. The structure of staphopain B − staphostatin B complex (PDB code 1PXV) also demonstrates a hydrophobic but slightly more spacious S2 subsite indicating that residues with larger side chain then Ile are accepted [42]. Several staphopain cleavage sites were characterized in host proteins ([34,43,6], www.sanger.ac.uk) but, only a single one resembles preferred sequences determined in this study. The likely reason for such divergence is that in the current study kinetically most preferred substrates were selected whereas in the cited studies particular purified proteins (not necessarily physiological substrates) were incubated with staphopains until cleavage was observed. Therefore the cleavage of analyzed host proteins may be orders of magnitude less efficient than that of optimal substrates. It is of particular interest that the single cleavage site determined in analyzed host proteins which most resembles the preferred sequences characterized in this study is the one found in chemerin (FAFS↓KA), the only host protein studied for which physiological cleavage by staphopain was convincingly demonstrated [6]. As such, only a proteome wide analysis of staphopain substrates would allow to point correlations (if any) between the sequences determined in this study and sequences hydrolyzed in physiologically relevant substrates.

In summary, our initial hypothesis is only partially supported by our experimental findings. The substrate preferences of staphopains B and C are similar and distinct from that of staphopain A. It is tempting to speculate that these properties reflect the distinct physiological roles of the enzymes. However, we have also demonstrated that the substrate specificity of the staphopains is relaxed and cannot be fully described in terms of classical subsite preferences. Under the stringent selection conditions used in CLiPS, the most kinetically favored substrates repeatedly showed the differences between staphopains B and C, and staphopain A, but more relaxed selection and the use of different methods clearly demonstrated that the range of acceptable substrates is much broader. Therefore, we conclude that staphopains differ in their substrate preferences, but the differences are subtle. Other factors must concurrently determine the differential importance of these enzymes in staphylococcal virulence. For example, Wladyka and colleagues recently argued that a differential susceptibility to plasma protease inhibitors explains why the pathogenic function of staphopain C is restricted to birds and does not extend to humans [44]. Similar factors may influence the different roles of staphopains A and B, and these await further investigation. The comprehensive information on the substrate specificity of the staphopains and the utility of novel tools provided by this study should assist the future characterization of the roles of staphopains in staphylococcal virulence.

Acknowledgments

This work was supported in part by grants K PBP 000322 (to JP), N N302 130734 (to AD), N N204 160036 (to AL) and N N301 032834 (to GD), from the Polish Ministry of Science and Higher Education. Funding for NS by a Heisenberg Fellowship of the DFG is gratefully acknowledged. The research was carried out with the equipment purchased thanks to the financial support of the European Regional Development Fund in the framework of the Polish Innovation Economy Operational Program (contract No. POIG.02.01.00-12-167/08, project Malopolska Centre of Biotechnology).

Appendix. Supplementary data

Table S1 presents non-natural derivatives of E64 used in competition labeling experiment. LC-MS data in Fig. S2 support results from LSTS assays. Alignments of full sequences of randomized region obtained from CLiPS selection are listed in Table S3. Most preferred residues at each position determined with used techniques are summarized in Fig. S4.

Supplementary data associated with this article can be found in the online version, at doi:10.1016/j.biochi.2011.07.020.

References

[1] F.D. Lowy, *Staphylococcus aureus* infections, N. Engl. J. Med. 339 (1998) 520–532.
[2] R.J. Gordon, F.D. Lowy, Pathogenesis of methicillin-resistant *Staphylococcus aureus* infection, Clin. Infect. Dis. 46 (Suppl. 5) (2008) S350–S359.
[3] D. Styers, D.J. Sheehan, P. Hogan, D.F. Sahm, Laboratory-based surveillance of current antimicrobial resistance patterns and trends among *Staphylococcus aureus*: 2005 status in the United States, Ann. Clin. Microbiol. Antimicrob. 5 (2006) 2.
[4] A.L. Cheung, K.J. Eberhardt, E. Chung, M.R. Yeaman, P.M. Sullam, M. Ramos, A.S. Bayer, Diminished virulence of a sar-/agr- mutant of *Staphylococcus aureus* in the rabbit model of endocarditis, J. Clin. Invest. 94 (1994) 1815–1822.
[5] A. Karlsson, S. Arvidson, Variation in extracellular protease production among clinical isolates of *Staphylococcus aureus* due to different levels of expression of the protease repressor sarA, Infect. Immun. 70 (2002) 4239–4246.
[6] P. Kulig, B.A. Zabel, G. Dubin, S.J. Allen, T. Ohyama, J. Potempa, T.M. Handel, E.C. Butcher, J. Cichy, *Staphylococcus aureus*-derived staphopain B, a potent cysteine protease activator of plasma chemerin, J. Immunol. 178 (2007) 3713–3720.
[7] J. Smagur, K. Guzik, M. Bzowska, M. Kuzak, M. Zarebski, T. Kantyka, M. Walski, B. Gajkowska, J. Potempa, Staphylococcal cysteine protease staphopain B (SspB) induces rapid engulfment of human neutrophils and monocytes by macrophages, Biol. Chem. 390 (2009) 361–371.
[8] L. Shaw, E. Golonka, J. Potempa, S.J. Foster, The role and regulation of the extracellular proteases of *Staphylococcus aureus*, Microbiology 150 (2004) 217–228.
[9] S.N. Coulter, W.R. Schwan, E.Y. Ng, M.H. Langhorne, H.D. Ritchie, S. Westbrock-Wadman, W.O. Hufnagle, K.R. Folger, A.S. Bayer, C.K. Stover, *Staphylococcus aureus* genetic loci impacting growth and survival in multiple infection environments, Mol. Microbiol. 30 (1998) 393–404.
[10] K. Rice, R. Peralta, D. Bast, J. de Azavedo, M.J. McGavin, Description of staphylococcus serine protease (ssp) operon in *Staphylococcus aureus* and nonpolar inactivation of sspA-encoded serine protease, Infect. Immun. 69 (2001) 159–169.
[11] C. Burlak, C.H. Hammer, M.A. Robinson, A.R. Whitney, M.J. McGavin, B.N. Kreiswirth, F.R. Deleo, Global analysis of community-associated methicillin-resistant *Staphylococcus aureus* exoproteins reveals molecules produced in vitro and during infection, Cell Microbiol. 9 (2007) 1172–1190.
[12] B.V. Lowder, C.M. Guinane, N.L. Ben Zakour, L.A. Weinert, A. Conway-Morris, R.A. Cartwright, A.J. Simpson, A. Rambaut, U. Nubel, J.R. Fitzgerald, Recent

human-to-poultry host jump, adaptation, and pandemic spread of *Staphylococcus aureus*, Proc. Natl. Acad. Sci. U.S.A. 106 (2009) 19545–19550.
[13] S. Takeuchi, K. Matsunaga, S. Inubushi, H. Higuchi, K. Imaizumi, T. Kaidoh, Structural gene and strain specificity of a novel cysteine protease produced by *Staphylococcus aureus* isolated from a diseased chicken, Vet. Microbiol. 89 (2002) 201–210.
[14] G. Dubin, B. Wladyka, J. Stec-Niemczyk, D. Chmiel, M. Zdzalik, A. Dubin, J. Potempa, The staphostatin family of cysteine protease inhibitors in the genus Staphylococcus as an example of parallel evolution of protease and inhibitor specificity, Biol. Chem. 388 (2007) 227–235.
[15] E. Golonka, R. Filipek, A. Sabat, A. Sinczak, J. Potempa, Genetic characterization of staphopain genes in *Staphylococcus aureus*, Biol. Chem. 385 (2004) 1059–1067.
[16] S. Takeuchi, T. Kinoshita, T. Kaidoh, N. Hashizume, Purification and characterization of protease produced by *Staphylococcus aureus* isolated from a diseased chicken, Vet. Microbiol. 67 (1999) 195–202.
[17] M. Rzychon, A. Sabat, K. Kosowska, J. Potempa, A. Dubin, Staphostatins: an expanding new group of proteinase inhibitors with a unique specificity for the regulation of staphopains, Staphylococcus spp. cysteine proteinases, Mol. Microbiol. 49 (2003) 1051–1066.
[18] A. Dummler, A.M. Lawrence, A. de Marco, Simplified screening for the detection of soluble fusion constructs expressed in *E. coli* using a modular set of vectors, Microb. Cell Fact. 4 (2005) 34.
[19] D.C. Greenbaum, A. Baruch, M. Grainger, Z. Bozdech, K.F. Medzihradszky, J. Engel, J. DeRisi, A.A. Holder, M. Bogyo, A role for the protease falcipain 1 in host cell invasion by the human malaria parasite, Science 298 (2002) 2002–2006.
[20] D. Greenbaum, K.F. Medzihradszky, A. Burlingame, M. Bogyo, Epoxide electrophiles as activity-dependent cysteine protease profiling and discovery tools, Chem. Biol. 7 (2000) 569–581.
[21] M.B. Eisen, P.T. Spellman, P.O. Brown, D. Botstein, Cluster analysis and display of genome-wide expression patterns, Proc. Natl. Acad. Sci. U.S.A. 95 (1998) 14863–14868.
[22] D. Greenbaum, A. Baruch, L. Hayrapetian, Z. Darula, A. Burlingame, K.F. Medzihradszky, M. Bogyo, Chemical approaches for functionally probing the proteome, Mol. Cell Proteomics 1 (2002) 60–68.
[23] J. Pfizer, I. Assfalg-Machleidt, W. Machleidt, L. Moroder, N. Schaschke, Primed-site probing of papain-like cysteine proteases, Int. J. Peptide Res. Therapeut. 13 (2007) 93–104.
[24] R.A. Houghten, C. Pinilla, S.E. Blondelle, J.R. Appel, C.T. Dooley, J.H. Cuervo, Generation and use of synthetic peptide combinatorial libraries for basic research and drug discovery, Nature 354 (1991) 84–86.
[25] M. Wysocka, B. Kwiatkowska, M. Rzadkiewicz, A. Lesner, K. Rolka, Selection of new chromogenic substrates of serine proteinases using combinatorial chemistry methods, Comb. Chem. High Throughput. Screen. 10 (2007) 171–180.
[26] A. Lesner, K. Brzozowski, G. Kupryszewski, K. Rolka, Design, chemical synthesis and kinetic studies of trypsin chromogenic substrates based on the proteinase binding loop of Cucurbita maxima trypsin inhibitor (CMTI-III), Biochem. Biophys. Res. Commun. 269 (2000) 81–84.
[27] A. Lesner, G. Kupryszewski, K. Rolka, Chromogenic substrates of bovine beta-trypsin: the influence of an amino acid residue in P1 position on their interaction with the enzyme, Biochem. Biophys. Res. Commun. 285 (2001) 1350–1353.
[28] K.S. Larsen, H. Ostergaard, J.R. Bjelke, O.H. Olsen, H.B. Rasmussen, L. Christensen, B.B. Kragelund, H.R. Stennicke, Engineering the substrate and inhibitor specificities of human coagulation Factor VIIa, Biochem. J. 405 (2007) 429–438.
[29] K.T. Boulware, P.S. Daugherty, Protease specificity determination by using cellular libraries of peptide substrates (CLiPS), Proc. Natl. Acad. Sci. U.S.A. 103 (2006) 7583–7588.
[30] J.J. Rice, P.S. Daugherty, Directed evolution of a biterminal bacterial display scaffold enhances the display of diverse peptides, Protein Eng. Des. Sel. 21 (2008) 435–442.
[31] I. Schechter, A. Berger, On the size of the active site in proteases. I. Papain, Biochem. Biophys. Res. Commun. 27 (1967) 157–162.
[32] N.D. Rawlings, A.J. Barrett, A. Bateman, MEROPS: the peptidase database, Nucleic Acids Res. 38 (2010) D227–D233.
[33] M.E. McGrath, The lysosomal cysteine proteases, Annu. Rev. Biophys. Biomol. Struct. 28 (1999) 181–204.
[34] J. Potempa, W. Watorek, J. Travis, The inactivation of human plasma alpha 1-proteinase inhibitor by proteinases from *Staphylococcus aureus*, J. Biol. Chem. 261 (1986) 14330–14334.
[35] A. Bjorklind, S. Arvidson, Occurrence of an extracellular serineproteinase among *Staphylococcus aureus* strains, Acta Pathol. Microbiol. Scand. B 85 (1977) 277–280.
[36] R. Filipek, M. Rzychon, A. Oleksy, M. Gruca, A. Dubin, J. Potempa, M. Bochtler, The staphostatin–staphopain complex: a forward binding inhibitor in complex with its target cysteine protease, J. Biol. Chem. 278 (2003) 40959–40966.
[37] J.L. Harris, B.J. Backes, F. Leonetti, S. Mahrus, J.A. Ellman, C.S. Craik, Rapid and general profiling of protease specificity by using combinatorial fluorogenic substrate libraries, Proc. Natl. Acad. Sci. U.S.A. 97 (2000) 7754–7759.
[38] M. Debela, V. Magdolen, N. Schechter, M. Valachova, F. Lottspeich, C.S. Craik, Y. Choe, W. Bode, P. Goettig, Specificity profiling of seven human tissue kallikreins reveals individual subsite preferences, J. Biol. Chem. 281 (2006) 25678–25688.
[39] M. Wysocka, A. Lesner, K. Guzow, L. Mackiewicz, A. Legowska, W. Wiczk, K. Rolka, Design of selective substrates of proteinase 3 using combinatorial chemistry methods, Anal. Biochem. 378 (2008) 208–215.
[40] K. Paschalidou, U. Neumann, B. Gerhartz, C. Tzougraki, Highly sensitive intramolecularly quenched fluorogenic substrates for renin based on the combination of L-2-amino-3-(7-methoxy-4-coumaryl)propionic acid with 2,4-dinitrophenyl groups at various positions, Biochem. J. 382 (2004) 1031–1038.
[41] B. Hoffman, D. Schomburg, H.J. Hecht, Crystal structure of a thiol proteinase from *Staphylococcus aureus* V8 in the E-64 inhibitor complex, Acta Crystallograp. (1993) 102.
[42] R. Filipek, R. Szczepanowski, A. Sabat, J. Potempa, M. Bochtler, Prostaphopain B structure: a comparison of proregion-mediated and staphostatin-mediated protease inhibition, Biochemistry 43 (2004) 14306–14315.
[43] I. Massimi, E. Park, K. Rice, W. Muller-Esterl, D. Sauder, M.J. McGavin, Identification of a novel maturation mechanism and restricted substrate specificity for the SspB cysteine protease of *Staphylococcus aureus*, J. Biol. Chem. 277 (2002) 41770–41777.
[44] B. Wladyka, G. Dubin, A. Dubin, Activation mechanism of thiol protease precursor from broiler chicken specific *Staphylococcus aureus* strain CH-91, Vet. Microbiol. 147 (2011) 195–199.