Cloning and rational mutagenesis of kexstatin I, a potent proteinaceous inhibitor of Kex2 proteinase

Kohei ODA^{*1}, Hiroshi OYAMA^{*}, Satoshi ITO^{*}, Momoko FUKIHARU^{*}, Yu MIYAGAWA^{*}, Saori TAKAHASHI^{*}, Mika HIROSE[†], Norihisa KIKUCHI^{‡2}, Toru NAKAYAMA^{‡3} and Yuji SHIBANO^{‡4}

^{*}Department of Applied Biology, Faculty of Textile Science, Kyoto Institute of Technology, Matsugasaki, Sakyo-ku, Kyoto 606-8585, Japan, and [†]Department of Molecular Biology, Biomolecular Engineering Research Institute (BERI), 6-2-3 Furuedai, Suita, Osaka 565-0874, Japan

Kexstatin I is a potent proteinaceous inhibitor of Kex2 proteinase (EC 3.4.21.61). In the present study we show the molecular cloning, primary structure determination and expression of the gene encoding kexstatin I. We also demonstrate its enhanced activity and specificity for Kex2 proteinase inhibition by rational mutagenesis. The cloned kexstatin I gene encoded a protein of 145 amino acid residues, including the 35-residue signal sequence for secretion. The amino acid sequence showed 52% identity with those of the *Streptomyces* subtilisin inhibitors (SSIs). Thus kexstatin I is the first SSI-family member that can inhibit Kex2 proteinase. The reactive site of the inhibitor was determined to be -Thr⁶⁹-Lys⁷⁰↓Glu⁷¹-, where ↓ indicates the reactive site. Because Kex2 proteinase generally shows the highest affinity for substrates with basic amino acid residues at the P₁ and P₂ sites, conversion of the Thr⁶⁹-Lys⁷⁰ segment of the inhibitor into

dibasic motifs was expected to result in enhanced inhibitory activities. Thus we constructed kexstatin I mutants, in which the Thr⁶⁹-Lys⁷⁰ sequence was replaced by the Thr⁶⁹-Arg⁷⁰, Lys⁶⁹-Lys⁷⁰ and Lys⁶⁹-Arg⁷⁰ sequences using PCR-based mutagenesis, and analysed them kinetically. Among these mutants, the Lys⁶⁹-Arg⁷⁰ mutant was the most potent inhibitor. The K_i for Kex2 proteinase was 3.2 × 10⁻¹⁰ M, which was 140-fold lower than that of the inhibitor with the Thr⁶⁹-Lys⁷⁰ sequence. Although kexstatin I could also inhibit subtilisin, the enhancement of inhibitory activity upon such mutations was specific for Kex2 proteinase inhibition.

Key words: furin, inhibition, subtilisin, *Streptomyces* subtilisin inhibitors.

INTRODUCTION

Many eukaryotic proteins and peptides are expressed as pro-protein and propeptide precursors. These precursors are then cleaved at specific sites to generate mature species with biological activities. Such post-translational processing, namely limited endoproteolysis, frequently occurs at sites with paired basic amino acid residues. An unequivocal example of an enzyme capable of cleaving at such sites is the product of the *KEX2* gene of the yeast *Saccharomyces cerevisiae*. Kex2 proteinase (also known as kexin; EC 3.4.21.61) is a serine-type endoproteinase catalysing hydrolytic cleavage at sites -Xaa-Xaa-Xaa-Arg↓, where the P₂ site is lysine or arginine for optimal activity [1–4]. The enzyme is known to function in yeast cells as a convertase, producing a mature α-mating factor and killer toxin from their precursors [5]. Recently, the genes of the mammalian counterparts of Kex2, furin and other prohormone convertases have been identified and shown to be involved in the processing of a wide variety of precursor proteins yielding hormones, plasma proteins, receptors, proteinases, viral proteins and bacterial exotoxins [6]. Kex2 and its mammalian analogues have a subtilisin-like catalytic domain and thus belong to the subtilisin superfamily. However, their specificity and biological functions are distinct from those of the previously known degradative members of this superfamily and are categorized into an alterna-

tive class that is known as the subtilisin/Kex2 endoproteinase family (Kex2 family) [6]. The mechanism and specificity of these Kex2-family enzymes have been most intensively studied for the yeast Kex2 enzyme, which has therefore become a paradigm for understanding the molecular basis of processing into bioactive peptides and proteins [6].

The reaction mechanisms of many proteinases have been proposed on the basis of the crystal structures of enzymes, as well as those of enzyme–inhibitor complexes. Therefore there has been considerable interest in the specific proteinase inhibitors, especially the proteinaceous inhibitors. For proprotein-processing proteinases, however, neither the crystal structure nor the specific proteinaceous inhibitor has been identified so far. In 1996, we found a novel proteinaceous proteinase inhibitor, kexstatin I, in the culture filtrate of an actinomycete *Streptomyces platensis* Q268 [7]. Kexstatin I strongly inhibits Kex2 proteinase, subtilisin and, to lesser extent, trypsin (K_i = 5.8 × 10⁻⁷ M; K. Oda, unpublished work), but does not inhibit thermolysin, pepsin, chymotrypsin or cathepsin B. Because kexstatin I is currently the only known example of a proteinaceous inhibitor that potently inhibits Kex2 proteinase, this inhibitor should be a good tool for the identification of the tertiary structure of the Kex2 proteinase–substrate complex, providing important information concerning the specificity and catalytic mechanism of these important processing proteinases.

Abbreviations used: Boc, t-butoxycarbonyl; MCA, methylcoumarinamide; LB, Luria–Bertani; ORF, open reading frame; r-kexstatin I, recombinant kexstatin I; SSI, *Streptomyces* subtilisin inhibitor; Suc, succinyl.

¹ To whom correspondence should be addressed (e-mail bika@ipc.kit.ac.jp).

² Present address: Shionogi Research Laboratories, Shionogi & Co., Ltd., 5-12-4 Sagisu, Fukushima-ku, Osaka, 553-0002, Japan.

³ Present address: Department of Biomolecular Engineering, Graduate School of Engineering, Tohoku University, 07 Aoba Aza Aramaki, Aoba-ku, Sendai, Miyagi 980-8579, Japan.

⁴ Present address: Suntory Ltd., Mori Bldg 31F, 5-7-2 Kojimachi, Chiyoda-ku, Tokyo 102-8530, Japan.

The nucleotide and amino acid sequence data reported (kexstatin I) appear in the GenBank® Database under the accession numbers AB005264 and BAA21112.

In the present study, we cloned and sequenced the kexstatin I gene and overexpressed it in a heterologous system. We also determined the reactive site of the inhibitor, whose structural characteristics did not satisfy the structural requirements for optimal reactivity to the Kex2 enzyme. This observation prompted us to alter the reactive-site sequence into other sequences, which fulfilled such requirements. Thereby, we could successfully obtain more potent and more specific mutant inhibitors for Kex2 proteinase.

The amino acid residues around the cleavage site of the inhibitor are designated as $-P_4-P_3-P_2-P_1\downarrow P_1'-P_2'-$, where the residues on the acyl side of the scissile peptide bond are numbered P_1, P_2, P_3 etc., and those towards the C-terminus are labelled P_1', P_2' etc., with the scissile bond (indicated by $\downarrow$) between P_1 and P_1' . The positional numbering of the amino acid residues in recombinant kexstatin I (r-kexstatin I) and its mutants is the same as that in the mature form of native kexstatin I.

EXPERIMENTAL

Materials

Kex2 proteinase was overexpressed with *Pichia pastoris* GS115 as a soluble protein as previously described [8]. Subtilisin (subtilisin BPN'), which was a product of Nagase Co. (Osaka, Japan) was further purified by gel-filtration chromatography on Superdex 200 HR (Amersham Pharmacia Biotech, Tokyo, Japan) equilibrated with 50 mM sodium acetate buffer (pH 7.0) containing 1 mM CaCl_2 . Synthetic peptide substrates, t-butoxycarbonyl (Boc)-Gln-Arg-Arg-methylcoumarinamide (MCA) and succinyl (Suc)-Ala-Ala-Pro-Phe-MCA, were purchased from Peptide Institute, Inc. (Osaka, Japan). Native kexstatin I was purified from the culture filtrate of *S. platensis* Q268, which is the stock culture of the Department of Applied Biology, Faculty of Textile Science, Kyoto Institute of Technology, as described previously [7]. Plasmid pIN-III-OmpA2 [9] was a gift from Dr S. Taguchi (Institute for Physical and Chemical Research, Riken, Wako-shi, Saitama, Japan).

Amino acid sequence determination and PCR amplification

The purified native kexstatin I [0.1 mg in 100 μl of 0.1 M Tris/HCl (pH 8.0)] was digested with 1 μg of tosylphenylalanyl-chloromethane ('TPCK')-treated trypsin (Sigma) at 37 °C for 3 h. Peptides were purified with an HPLC system equipped with an Aquapore Butyl column (2.1 mm $\times$ 220 mm; Applied Biosystems, Foster City, CA, U.S.A.). A linear gradient elution from 0.1% (v/v) trifluoroacetic acid in water to 64% (v/v) acetonitrile in 0.1% trifluoroacetic acid was carried out at a flow rate of 0.3 ml/min. Peptides were detected by monitoring the changes in the absorbance at 220 nm. The peptide peaks were collected and re-chromatographed until they were homogeneous. Amino acid sequences were subsequently analysed by automated Edman degradation (see below).

The SKI-1 and SKI-3C primers, 5'-ATTACGC(G/C)CC-(C/G)AC(C/G)GA(A/C)CT-3' and 5'-ATGCCCTGCCA(C/G)AC(C/G)CC(C/G)T-3' respectively, were designed on the basis of the partial amino acid sequences of the tryptic peptides, Tyr-Ala-Pro-Thr-Glu-Leu and Asp-Gly-Val-Trp-Gln-Gly, and were used as upstream and downstream primers respectively. The bases, shown in parentheses, indicate the degenerate sites. PCR (25 cycles) was completed using the *S. platensis* genomic DNA as a template and the above PCR primers. The cycling conditions were 94 °C (60 s) for denaturation, 50 °C (90 s) for annealing and 72 °C (90 s) for elongation. The PCR product, estimated to be a 250 bp fragment, was gel-purified and sequenced

to confirm that it was indeed an amplified kexstatin I gene. The PCR product was radiolabelled with [α - ^{32}P]dCTP using an Oligolabeling kit (Amersham Pharmacia Biotech) according to the manufacturer's instructions. This probe was used to screen the *S. platensis* genomic library by colony hybridization (see below).

Cloning of the kexstatin I gene

The chromosomal DNA was isolated from the *S. platensis* Q268 cells by the procedure described by Obata et al. [10]. The chromosomal DNA was digested with the restriction enzymes *Bam*HI and *Sal*I, and the digests were separated by 0.7% agarose gel electrophoresis. DNA fragments with a size of 2000 bp were recovered from the agarose gels by electrophoresis using DNACELL (Daichi Pure Chemicals, Tokyo, Japan) and ligated with *Bam*HI-*Sal*I-digested pUC119. The *Escherichia coli* JM109 cells were transformed with the ligation mixture, and the transformants were grown on Luria-Bertani (LB) agar plates containing 50 $\mu\text{g}/\text{ml}$ ampicillin at 37 °C overnight. The genomic library thus obtained was used for colony hybridization using the radiolabelled 250 bp probe mentioned above. The other DNA manipulation procedures were performed using standard techniques [11]. The DNA sequences of both strands were determined using an ABI PRISM 310 DNA sequencer (Applied Biosystems) by the dideoxy-chain termination method, which were then analysed and aligned using the GeneWorks computer program (Intelligenetics, Inc., Mountain View, CA, U.S.A.).

Expression

The gene encoding a mature form of kexstatin I was heterologously expressed in the *E. coli* cells as follows. A DNA fragment of kexstatin I encoding amino acid residues 36–145 (see Figure 1) was amplified by PCR using the pUCKX1 (see the Results and Discussion sections) as a template and primers a (5'-AAAGAATTTCGGCCTCTACG CCCCAGCG-3') and b (5'-GGATCCTCAGAAAGTTGAAGACCG-3'), where the underlining indicates the *Eco*RI and *Bam*HI sites respectively. The PCR product was digested with *Eco*RI and *Bam*HI and ligated with the *Eco*RI-*Bam*HI-digested pIN-III-OmpA2 to obtain the plasmid pIN-III-OmpA2/KXTN, which was then transformed with the cells of *E. coli* strain JM109. The transformant cells were grown in LB medium containing 50 mg/l ampicillin at 32 °C for 12 h with shaking. The culture (24 ml) was then inoculated into 2400 ml of M9 medium containing 1 mM isopropyl β -D-thiogalactoside. The cells were grown at 32 °C for 24 h with shaking and were subsequently collected by centrifugation.

Purification of r-kexstatin I

The collected cells were disrupted at 4 °C by seven cycles of ultrasonication (where one cycle corresponds to 10 kHz for 2 min followed by an interval of 1 min). The cell debris was removed by centrifugation. To the supernatant solution, solid ammonium sulphate was slowly added to 80% saturation. After the mixture had been left at 4 °C for 12 h, the precipitate was collected by centrifugation and dissolved in 40 ml of 0.02 M Tris/HCl buffer, pH 7.5 (buffer B). A 14 ml portion of the inhibitor solution was applied to a Sephadex G-75 column (25 mm $\times$ 900 mm) equilibrated with buffer B, and eluted. Second and third portions of the inhibitor solution were chromatographed in the same way. The active fractions (180 ml) were combined and applied to a CM-Sephadex CL-6B column (25 mm $\times$ 150 mm) equilibrated with buffer B. The column was washed with three column volumes of this buffer. The inhibitor

Figure 1 Nucleotide and deduced amino acid sequences of kexstatin I

The underlined amino acid sequence is the suggested signal sequence for secretion. The N-terminal Gly¹ of the mature kexstatin I is indicated by an asterisk. A vertical arrow indicates the identified reactive site. The partial amino acid sequences used for designing the SKI-1 and SKI-3C PCR primers are double-underlined. Horizontal arrows under the nucleotide sequence indicate possible transcription terminator sequences with an inverted repeat motif.

was eluted with a linear gradient of NaCl (0–0.6 M) in the same buffer (300 ml). Active fractions were combined, dialysed against distilled water and freeze-dried.

Determination of the reactive site

Native kexstatin I (4.42 nmol), which was purified from the culture filtrate of *S. platensis* Q268, was reacted at 25 °C for 5

days with 0.43 nmol of Kex2 proteinase in a mixture (1.0 ml) containing 50 µmol of Tris/HCl (pH 7.0), 10 µmol of CaCl₂, 0.2 mg of NaN₃ and 0.1 mg of Triton X-100. The mixture was then analysed for proteolytic cleavage of kexstatin I by SDS/PAGE. Protein bands in the gel were transferred on to a PVDF membrane by electroblotting, and the membrane was stained with Coomassie Brilliant Blue R250. Stained portions corresponding to fragments 1 and 2 of kexstatin I were excised using dissecting scissors and subjected to automated Edman degradation in order to determine the N-terminal amino acid sequences of the fragments. Protein bands in the gel were also electroblotted on to low-density Teflon tape (Hewlett Packard) and subjected to automated C-terminal amino acid sequencing (see under 'Analytical methods').

Mutagenesis studies

The r-kexstatin I genes with altered reactive-site sequences (see the Results section and Table 1 for nomenclature) were obtained by the following two-step PCR processes. Primers used in the PCR amplifications are listed in Table 1. To obtain the gene of the TR species of r-kexstatin I, for example, a DNA fragment (termed L) of the r-kexstatin I gene was amplified using the template plasmid pIN-III-OmpA2/KXTN and the PCR primers a and TR-L. PCR amplification was completed by incubating the PCR mixture at 94 °C for 1 min followed by thermal cycling (25 cycles) at 94 °C for 1 min for denaturation, 55 °C for 1 min for annealing, and 72 °C for 2 min for elongation. Similarly, a DNA fragment (termed R) of the r-kexstatin I gene was also amplified using the template plasmid and the PCR primers b and TR-U. L and R, the DNA fragments amplified by these first-step PCR processes, share a common DNA sequence encoding the mutated reactive site. These fragments were gel-purified, combined and subjected to the second-step PCR using the a and b primers under the same conditions as above. The 300 bp PCR product thus obtained was digested with *Xho*I and *Bam*HI and ligated with *Xho*I–*Bam*HI-digested pIN-III-OmpA2 to obtain a vector pIN-III-OmpA2/KXTN-TR. It was then sequenced using the a and b primers to confirm that only the expected mutation was introduced into the r-kexstatin I gene. The KK and KR species were obtained in the same way as above, except that the primers used during the first-PCR process were, instead of the TR-L and TR-U primers, the KK-L and KK-U primers for the KK species and the KR-L and KR-U primers for the KR species. The mutated kexstatins were expressed in the *E. coli* JM109 cells and purified to homogeneity as described above.

Table 1 Nomenclature of mutant kexstatin I and synthetic oligonucleotides used in the mutagenesis experiments

Construction of TR, KK and KR species is described in the Experimental section. During the first-step PCR processes, the TR-L, KK-L and KR-L primers were used with primer a, and the TR-U, KK-U, and KR-U primers were used with primer b, where the primers a and b were 5'-AAAGAATTCGGCTCTACGCCCGACG-3' and 5'-GGATCCTCAGAAGTTGAAGACCG-3', and were sense and antisense primers for amplification of the r-kexstatin I gene. For further details see the text.

Species name	Reactive-site sequence*	Names and sequences of mutagenesis primers
TR	Thr ⁶⁹ -Arg ⁷⁰	TR-L: 5'-TCCACTCCCTGGTGACAGCCGGTCC-3' TR-U: 5'-GTGCACAGGGAGTGAACCCCTCG-3'
KK	Lys ⁶⁹ -Lys ⁷⁰	KK-L: 5'-TCCACTCCTTCTTGACAGCCGGTCCGAGG-3' KK-U: 5'-GTGCAAGAAGGAGTGAACCCCTC-3'
KR	Lys ⁶⁹ -Arg ⁷⁰	KR-L: 5'-CCCTCTTGACAGCCGGTCCGAGG-3' KR-U: 5'-CGGACCGGCTGTGCAAGAGGGAGTGAACCCCTCG-3'

* The reactive-site sequence of r-kexstatin I is Thr⁶⁹-Lys⁷⁰. Numbering of amino acid residues is based on that of native kexstatin I.

Inhibition kinetics

Kex2 proteinase [0.16 pmol in 960 μ l of 0.2 M Tris/HCl (pH 7.0) containing 1 mM CaCl_2 and 0.01 % Triton X-100 (buffer A)] was mixed with varying amounts of an inhibitor (in 20 μ l of buffer A) and incubated at 25 °C for 1 h to complete the binding equilibrium. Following incubation, the synthetic substrate (Boc-Gln-Arg-Arg-MCA, 2 nmol in 20 μ l of buffer A) was added to the mixture. Final concentrations of the inhibitors in the reaction mixture were as follows: 0–160 nM native kexstatin I; 0–240 nM r-kexstatin; 0–200 nM TR species; 0–500 nM KK species; 0–2 nM KR species. Residual enzyme activity was estimated from the initial linear part of fluorescence intensity change (excitation, 380 nm; emission, 460 nm), which was monitored at 25 °C with a Hitachi Fluorescence Spectrophotometer F-2000. Because preliminary studies suggested that kexstatin I was a tight-binding inhibitor whose K_i values were of the same or lower order as the enzyme concentration, we analysed the inhibition according to the procedure described by Tashiro et al. [12] and Laskowski and Kato [13], which has widely been used for the analysis of tight-binding inhibition. Briefly, the slopes were converted into free enzyme concentrations, which were used, together with the initial concentrations of the inhibitor, to calculate the association equilibrium constants, K_a ($= 1/K_i$), using software devised by Professor Tashiro [12] (Department of Food Science and Nutrition, School of Human Environmental Science, Mukogawa Women's University, Nishinomiya, Hyogo, Japan) run on an NEC 9801 computer; each titration curve was fitted to the theoretical equation of Bieth and Frechin [14] using a non-linear least-squares analysis program; the best-fit line through the data points yielded K_a .

For subtilisin assays, the substrate used was Suc-Ala-Ala-Pro-Phe-MCA, and the amount of enzyme was 0.064 pmol. Final concentrations of inhibitors in the reaction mixture were 0–0.3 nM for native kexstatin I, r-kexstatin I and the TR, KK and KR species. Other assay conditions were the same as above.

Inhibitor concentration was determined by the method of Smith et al. [15] with BSA as the standard. The concentration of the Kex2 proteinase stock solution was determined by amino acid analysis; the enzyme was hydrolysed in constant-boiling HCl (5.7 M) under reduced pressure and analysed for amino acid composition using a Hitachi high-performance amino acid analyser (model 8500). The enzyme concentration was calculated on the basis of the submitted amino acid sequence of the enzyme (GenBank® accession number AAA34719).

Analytical methods

The molecular masses of subunits were estimated by SDS/PAGE with 15 % (w/v) polyacrylamide gels according to the procedure of Laemmli [16], using standard proteins (Amersham Pharmacia Biotech), electrophoresed simultaneously, as molecular-mass markers. Proteins were stained with Coomassie Brilliant Blue R250 and destained in destaining solution (methanol/acetic acid/water, 2:1:7, by vol.).

The native molecular mass of the inhibitor was estimated by gel-filtration chromatography on Sephadex G75 (25 mm $\times$ 900 mm) equilibrated with buffer B (see above). For calibration, standard proteins (Amersham Pharmacia Biotech) were chromatographed under the same conditions.

N-terminal amino acid sequences were determined by automated Edman degradation with a gas-phase sequencer, Perkin Elmer/Applied Biosystems model 492. C-terminal amino acid sequences were determined as described by Burkhart et al. [17], using an automated C-terminal protein sequencer G1009A

(Hewlett Packard) equipped with an on-line thiohydantoin analysis system.

CD spectra were measured on a JASCO J-720A spectropolarimeter interfaced with a data processor. A cell pathlength of 0.1 cm and protein concentrations of 82 μ M (wild-type) and 46 μ M (KR species) were used. All spectra were obtained by averaging 32 scans and by subtraction of the base-line spectrum of the buffer. No filtering or smoothing techniques were used to reduce noise levels.

RESULTS

Cloning and structure of the kexstatin I gene

Tryptic peptides of kexstatin I were isolated, and their amino acid sequences were determined. On the basis of the determined amino acid sequences, PCR primers (SKI-1 and SKI-3C, see the Experimental section) were designed. An internal stretch of the kexstatin I gene was then amplified by PCR using the *S. platensis* genomic DNA as a template and these primers, and a 246 bp DNA fragment was obtained (Figure 1). Using the ^{32}P -labelled 246 bp DNA fragment as a probe, the genomic library of *S. platensis* was screened for the kexstatin I gene by colony hybridization. One positive clone was obtained from approximately 100 ampicillin-resistant transformants. This clone harboured a plasmid, termed pUCKX1, with a 2200 bp insert containing an open reading frame (ORF) of 435 bp (Figure 1), which encoded a protein of 145 amino acid residues. The amino acid sequence deduced from the nucleotide sequence of the ORF contained sequences that completely coincided with those determined with tryptic peptides and automated Edman degradation (Figure 1). Therefore we concluded that this ORF coded for kexstatin I. A comparison of the N-terminal amino acid sequence [7] (Figure 1) with the deduced amino acid sequence, taken together with the fact that kexstatin I is an extracellular protein, led us to conclude that the amino acid sequence from positions –1 to –35 should be a signal for secretion [18]; thus, the purified kexstatin I from *S. platensis*, which is the mature form, consists of 110 amino acids, and its molecular mass is 11 788 Da. The molecular mass of native kexstatin I was estimated to be 23 kDa by gel-filtration chromatography on a Sephadex G-75 column; therefore kexstatin I was homodimeric. Kexstatin I showed sequence similarity with proteinase inhibitors belonging to the *Streptomyces* subtilisin inhibitor (SSI) family [19] (Figure 2). The highest identity, 52 %, was found with SSI [20], which is also a homodimeric inhibitor. The SSI family of inhibitors are known to have two intramolecular disulphide bridges; the observed high sequence identity between kexstatin I and SSI, along with the fact that CD spectra of SSI (results not shown) and those of the native and KR species of kexstatin I (Figure 3) essentially coincided with each other, strongly suggest that kexstatin I should also have two intramolecular disulphide bridges as in SSI (Figure 2).

Expression and purification of r-kexstatin I

We constructed an expression vector, pIN-III-OmpA2/KXTN, in which the gene encoding the mature form of kexstatin I was expressed with the ompA signal peptide in the *E. coli* JM109 cells used as a host. The heterologous expression system produced an active inhibitor (approx. 30 mg/l broth), and the recombinant inhibitor (termed r-kexstatin I) could be purified to homogeneity, as judged by SDS/PAGE (results not shown), by a combination of ammonium sulphate fractionation, as well as gel-filtration

Figure 2 Sequence alignment and dendrogram of kexstatin I and other SSI-family inhibitors

(A) Alignment of amino acid sequences of the mature forms of native kexstatin I (shown in bold) and other SSI-family inhibitors: SSI [20,21]; SCI, *Streptomyces* chymotrypsin inhibitor; PSN, plasminostrepin [28]; STIs 1 and 2, *Streptomyces* trypsin inhibitors 1 and 2 [29]; API, alkaline proteinase inhibitor 2c [30]; and SILs 1, 2, 3, 4, 8, 10, 13 and 14, *Streptomyces* subtilisin inhibitor-like proteins 1, 2, 3, 4, 8, 10, 13 and 14 [31]. Identical amino acid residues among all of the inhibitors listed are shadowed. A vertical arrow indicates the reactive site. The alignment suggests the presence of disulphide bridges between Cys³¹ and Cys⁴⁶ (open circles) and between Cys⁶⁸ and Cys⁹⁸ (solid circles) in the mature form of native kexstatin I. (B) A dendrogram of the SSI-family inhibitors. The phylogenetic tree was constructed by the neighbour-joining method [32], using the Genetyx Mac computer program (version 9.0).

Figure 3 CD spectra of the native (solid line) and KR species (broken line) of kexstatin I

For details see the text.

chromatography and cation-exchange chromatography. Automated Edman degradation of the purified inhibitor showed that it possessed an extra tripeptide sequence, Ala-Glu-Phe-, which

was attached to the N-terminal Gly¹ of the mature form of native kexstatin I, as expected (see the Experimental section), indicating that the ompA signal peptide was precisely removed by the action of a signal peptidase in the *E. coli* cells.

Determination of the reactive site

To determine the reactive site of kexstatin I, the native inhibitor was reacted with a relatively large amount of Kex2 proteinase for a prolonged period of time (5 days). SDS/PAGE analysis showed that this incubation resulted in the formation of two degradation products with molecular masses of 7 and 4 kDa, termed fragments 1 and 2 respectively. After electroblotting of the protein bands of the fragments onto PVDF membranes, N- and C-terminal amino acid sequences of the fragments were analysed by automated protein sequencing; the determined N- and C-terminal sequences were Gly-Leu-Tyr-Ala-Pro-Thr- and -Cys-Thr-Lys for fragment 1 and Glu-Trp-Asn-Pro- and -Phe-Asn-Phe for fragment 2. A comparison of these sequences with the deduced amino acid sequence of kexstatin I (Figure 1) showed conclusively that the reactive site of kexstatin I was -Thr⁶⁹-Lys⁷⁰-Glu⁷¹-.

Rational mutagenesis studies

We constructed the TR, KK and KR species of r-kexstatin I, where the Thr⁶⁹-Lys⁷⁰ sequence in the reactive site of the r-kexstatin I was replaced by Thr-Arg, Lys-Lys and Lys-Arg sequences, respectively, by PCR-based mutagenesis (see Table 1). The r-kexstatin I is, therefore, synonymous with the TK species. The mutated kexstatin I genes were expressed with the *E. coli* JM109/pIN-III-OmpA2 system and the mutant proteins

Figure 4 Titration curve of Kex2 proteinase with the KR species of r-kexstatin I

Kex 2 proteinase (final concentration of 0.157 nM) was incubated with various concentrations of KR species (0–2 nM) at 25 °C for 1 h and subsequently assayed as described in the Experimental section.

Table 2 K_i values for inhibition of Kex2 proteinase and subtilisin

Name	Reactive-site sequence*	K_i (M)	
		Kex2 proteinase	Subtilisin
Native kexstatin I	Thr ⁶⁹ -Lys ⁷⁰	4.4×10^{-8}	4.9×10^{-11}
r-Kexstatin I (TK species)	Thr ⁶⁹ -Lys ⁷⁰	4.2×10^{-8}	1.3×10^{-10}
TR species	Thr ⁶⁹ -Arg ⁷⁰	1.2×10^{-8}	4.0×10^{-11}
KK species	Lys ⁶⁹ -Lys ⁷⁰	3.7×10^{-9}	6.9×10^{-10}
KR species	Lys ⁶⁹ -Arg ⁷⁰	3.2×10^{-10}	2.6×10^{-9}

* Numbering of amino acid residues is based on the that of native kexstatin I.

were purified to homogeneity as described above. Automated Edman degradation of the purified mutants showed that N-terminal amino acid sequences of these mutants were the same as that of r-kexstatin I. Also, the C-terminal amino acid sequences of the r-kexstatin I and its mutants were the same as that of the native inhibitor. The molecular masses of these recombinant proteins were estimated to be 23 kDa by gel filtration, whereas SDS/PAGE of these proteins gave single protein bands corresponding to molecular masses of 12 kDa, indicating that these recombinant proteins were homodimeric, as was the native inhibitor. Moreover, CD spectra of these recombinant proteins were essentially identical with that of the native inhibitor (see Figure 3 for KR species), indicating that there was no major alteration in the main-chain conformation from that of the native inhibitor.

Like the native inhibitor [7], r-kexstatin I and all of the mutants inhibited Kex2 proteinase and subtilisin, but not thermolysin, pepsin, chymotrypsin or cathepsin B. Titration curves for the inhibition of Kex2 proteinase and subtilisin by the native and recombinant inhibitors were best fitted to those described in terms of tight-binding competitive inhibition [12,13] (Figure 4). This is consistent with the fact that the SSI, a prototype inhibitor of the SSI family, competitively inhibits subtilisin [21]. Also, slow-binding inhibitors generally act in a competitive manner

[22]. The K_i values for inhibition of Kex2 proteinase and subtilisin are summarized in Table 2. The KR species inhibited the Kex2 enzyme the most strongly and showed a K_i value that was 140-fold smaller than those of native kexstatin I and r-kexstatin I (i.e. the TK species). The KK and TR species could also inhibit the Kex2 enzyme more strongly than native kexstatin I and r-kexstatin I, although they inhibited the Kex2 enzyme more weakly than the KR species; the K_i value of the TR species was comparable with those of native kexstatin I and r-kexstatin I. In contrast, TK and TR species, having threonine residues at the P₂ sites, inhibited subtilisin much more strongly than the KK and KR species.

DISCUSSION

The present study describes the cloning, sequencing, expression and characterization of kexstatin I, which is the first example of a potent proteinaceous inhibitor for Kex2 proteinase. Primary structure analysis of kexstatin I revealed that it is phylogenetically related to the SSI family (see Figure 2B). This is consistent with the fact that Kex2 proteinase is a member of the subtilisin superfamily, with processing functions [6]. Although kexstatin I is the first example of an inhibitor from this family that can strongly inhibit Kex2 proteinase, it still inhibited subtilisin almost 900-fold more strongly than the Kex2 enzyme. Thus another important objective of the present study was to alter this inhibitor, by rational approaches, to enhance its specificity towards Kex2 proteinase inhibition.

It has been proposed that the proteinaceous proteinase inhibitor is essentially a substrate that cannot actually be turned over because of its very tight binding to the enzyme, leading to the notion of the 'reactive site', defined as a peptide bond that is potentially and specifically cleaved by the enzyme in the enzyme-inhibitor complex [19]. Importantly, it is suggested that the specificity of a proteinaceous proteinase inhibitor is, at least in part, governed by the amino acid residue at or near the reactive site of the inhibitor [23–25]. Hence, strongest inhibition would be expected when the nature of the reactive-site amino acids of the inhibitor satisfy the substrate specificity of the proteinase.

Kex2 proteinase exhibits a high selectivity for positively charged residues (i.e. lysine and arginine) at P₂ and a stricter requirement for Arg at P₁ [8], and also recognises the P₄ site with preference for aliphatic and basic residues; interplays between the P₁ and P₄ sites are shown to be important for the specificity of the enzyme [26]. In the native kexstatin I, however, lysine and threonine residues occupy the P₁ and P₂ sites of the inhibitor, respectively, and such structural characteristics did not satisfy the requirements for optimal reactivity to the Kex2 enzyme, although the leucine residue occupies the P₄ site of the inhibitor, providing favourable interactions between the enzyme and the inhibitor. Thus conversion of the reactive-site sequence at P₁–P₂ sites into dibasic sequences was expected to enhance the ability of the inhibitor to inhibit Kex2 proteinase. To address this interesting question, we constructed the TR, KK and KR species of r-kexstatin I, where the Thr⁶⁹-Lys⁷⁰ sequence in the reactive site of r-kexstatin I was replaced by Thr-Arg, Lys-Lys and Lys-Arg sequences, respectively, using PCR-based mutagenesis (see Table 1). The recombinant proteins were then heterologously expressed in *E. coli* cells. The expressed proteins were most probably accumulated in the periplasmic space of the cells and could be easily purified after total lysis of the cells rather than through isolation of the periplasmic fraction. We then analysed some physicochemical and biochemical properties of the recombinant inhibitors to examine the validity of their use for kinetic

assessment. The N- and C-terminal sequences of r-kexstatin I and its mutants were those which were expected (see the Results section). Subunit structures and CD spectra (see Figure 3) of these recombinant proteins were indistinguishable from those of the native inhibitor. These results, along with the fact that r-kexstatin I inhibited Kex2 proteinase as potently as the native inhibitor (see Table 2), corroborated that these recombinant proteins were fully active as inhibitors.

Kinetic analyses showed that KR species inhibited the Kex2 enzyme the most strongly, followed by the KK, TR and TK species, in decreasing order, indicating that dual substitutions at P₁ and P₂ sites with basic amino acid residues, especially with arginine at P₁, substantially enhance the inhibitory activity for Kex2 proteinase. This is in accordance with a preference of Kex2 proteinase for basic residues at P₁ and P₂ sites and an even more stringent requirement for an arginine residue at the P₁ site. Importantly, the enhancement of inhibitory activity by such mutations was specific for Kex2 proteinase inhibition; replacements of P₂ sites with basic amino acids resulted in a diminution of affinity for subtilisin, and KK and KR species, respectively, showed 5- and 20-fold larger K_i values for subtilisin than that of r-kexstatin I. As a net result of these effects, the specificity of kexstatin I could be improved 2600-fold for Kex2 proteinase inhibition by the dual replacement of P₁ and P₂ sites with lysine and arginine residues respectively. The KR species showed a specificity that was reversed from those of the native and recombinant inhibitors. The KR species could inhibit Kex2 proteinase 8-fold more strongly than subtilisin, in contrast with the native and r-kexstatins, which inhibited subtilisin 300–900-fold more strongly than the Kex2 enzyme.

It should be pointed out that some amino acid residues conserved throughout the SSI family are different in kexstatin I (e.g. Thr⁶, Glu⁷, Ile⁵⁹ and Leu⁷⁵), providing a possibility that these differences are responsible for the unusual specificity of kexstatin I. However, this would be less likely, because Thr⁶, Ile⁵⁹ and Leu⁷⁵ in kexstatin I are replaced by similar amino acids (serine, leucine and valine respectively) in the other SSI-family inhibitors; also, Glu⁷ lies distant from the reactive site of this inhibitor and should only minimally affect the specificity of the inhibitor. Nonetheless, it would be possible that, besides the contact between the reactive site and the atoms of the catalytic residues, other contacts between the enzyme and the inhibitor are important for their tight binding. The mutated r-kexstatin I obtained in the present study could not inhibit furin, indicating that a proteinaceous inhibitor can act on one of the Kex2-family enzymes without acting on the other analogues of the same enzyme family. This may be explained in terms of differences in the amino acid sequences other than the reactive site being responsible for the differential protein–protein interactions. This further implies that it might be possible to find, from Nature, an inhibitor that can exclusively act on one analogue, other than the Kex2 proteinase, of this family, and it might also be possible to alter the specificity of kexstatin I through protein engineering approaches. Because some of the Kex2 family analogues have been shown to be differentially expressed with distinct biological functions in mammalian tissues [6,27], specific inhibition by an inhibitor that has a limited specificity would further enhance our knowledge of the biological roles of the respective analogue of this family and should be of interest from a medical point of view [6].

In conclusion, the specificity of kexstatin I could be rationally improved for Kex2 proteinase inhibition by designing the reactive-site sequences. Attempts are currently underway to crystallize the Kex2–mutated kexstatin I complexes to elucidate the stereostructure of the complex by X-ray crystallography.

We are grateful to Dr S. Taguchi for providing us with the plasmid pIN-III-OmpA2.

REFERENCES

- Mizuno, K., Nakamura, T., Ohshima, T., Tanaka, S. and Matsuo, H. (1989) Characterization of KEX2-encoded endopeptidase from yeast *Saccharomyces cerevisiae*. *Biochem. Biophys. Res. Commun.* **159**, 305–211
- Mizuno, K., Nakamura, T., Ohshima, T., Tanaka, S. and Matsuo, H. (1988) Yeast KEX2 genes encodes an endopeptidase homologous to subtilisin-like serine proteases. *Biochem. Biophys. Res. Commun.* **156**, 246–254
- Fuller, R. S., Brake, A. J. and Thorner, J. (1989) Intracellular targeting and structural conservation of a prohormone-processing endoprotease. *Science (Washington, D.C.)* **246**, 482–486
- Bevan, A., Brenner, C. and Fuller, R. S. (1998) Quantitative assessment of enzyme specificity *in vivo*: P2 recognition by Kex2 protease defined in a genetic system. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 10384–10389
- Fuller, R. S., Sterne, R. E. and Thorner, J. (1988) Enzymes required for yeast prohormone processing. *Annu. Rev. Physiol.* **50**, 345–362
- Nakayama, K. (1997) Furin: a mammalian subtilisin/Kex2p-like endoprotease involved in processing of a wide variety of precursor proteins. *Biochem. J.* **327**, 625–635
- Oda, K., Takahashi, S., Kikuchi, N. and Shibano, Y. (1996) A novel proteinaceous Kex2 proteinase inhibitor, kexstatin, from *Streptomyces platensis* Q268. *Biosci. Biotechnol. Biochem.* **60**, 1388–1389
- Brenner, C. and Fuller, R. S. (1992) Structural and enzymatic characterization of a purified prohormone-processing enzyme: secreted, soluble Kex2 protease. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 922–926
- Taguchi, S., Yoshida, Y., Matsumoto, K. and Momose, H. (1993) Improved leader and putative terminator sequences for high-level production of *Streptomyces* subtilisin inhibitor in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **39**, 732–737
- Obata, S., Furukubo, S., Kumagai, I. and Miura, K. (1989) High-level expression in *Streptomyces lividans* 66 of a gene encoding *Streptomyces* subtilisin inhibitor from *Streptomyces albobriseolus*. *J. Biochem. (Tokyo)* **105**, 372–376
- Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Tashiro, M., Kihara, Y., Katayama, Y. and Maki, Z. (1989) Purification and characterization of a major trypsin inhibitor, FMT-II, from foxtail millet grain. *Agric. Biol. Chem.* **53**, 443–451
- Laskowski, Jr., M. and Kato, I. (1980) Protein inhibitors of proteinases. *Annu. Rev. Biochem.* **49**, 593–626
- Blieth, J. and Frechin, J.-C. (1974) Esterase inhibitors as impurities in commercial preparations of soybean trypsin inhibitor (Kunitz). In *Bayer Symposium V, Protease Inhibitors* (Fritz, H., Taschesche, H., Greene, L. J. and Truscheit, E., eds.), pp. 291–304, Springer-Verlag, Berlin
- Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J. and Klenk, D. C. (1985) Measurement of protein using bicinchoninic acid. *Anal. Biochem.* **150**, 76–85
- Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**, 680–685
- Burkhart, W. A., Moyer, M. B., Bailey, J. M. and Miller, C. G. (1996) Electrophoretic transfer of proteins to Teflon tape and membranes for N- and C-terminal sequence analysis. *Anal. Biochem.* **236**, 364–367
- Boyd, D. and Beckwith, J. (1990) The role of charged amino acids in the localization of secreted and membrane proteins. *Cell (Cambridge, Mass.)* **62**, 1031–1033
- Hiroimi, K., Akasaka, K., Mitsui, Y., Tonomura, B. and Murao, S. (1985) Protein Protease Inhibitor – The case of *Streptomyces* Subtilisin Inhibitor (SSI), Elsevier, Amsterdam
- Ikenaka, T., Odani, S., Sakai, M., Nabeshima, S., Sato, A. and Murao, S. (1974) Amino acid sequence of an alkaline proteinase inhibitor (*Streptomyces* subtilisin inhibitor) from *Streptomyces albobriseolus* S-3253. *J. Biochem. (Tokyo)* **76**, 1191–1209
- Inouye, K., Tonomura, B., Hiroimi, K., Sato, S. and Murao, S. (1977) The stoichiometry of inhibition and binding of a protein proteinase inhibitor from *Streptomyces* (*Streptomyces* subtilisin inhibitor) against subtilisin BPN'. *J. Biochem. (Tokyo)* **82**, 961–967
- Morrison, J. F. (1982) The slow-binding and slow, tight-binding inhibition of enzyme-catalyzed reactions. *Trends Biochem. Sci.*, **7**, 102–105
- Omichi, K., Nagura, N. and Ikenaka, T. (1980) The reactive site of *Streptomyces* subtilisin inhibitor. *J. Biochem. (Tokyo)* **87**, 483–489
- Nonaka, T., Suzuki, T., Tanaka, N., Saito, S., Senda, T., Miura, K. and Mitsui, Y. (1996) Structure and function of subtilisin BPN' as studied through crystallographic studies on a series of its complexes with genetically engineered proteinaceous inhibitor SSI. *Adv. Exp. Med. Biol.* **379**, 21–27

- 25 Kojima, S., Obata, S., Kumagai, I. and Miura, K. (1990) Alteration of the specificity of the *Streptomyces* subtilisin inhibitor by gene engineering. *Biotechnology (N.Y.)* **8**, 449–452
- 26 Rockwell, N. C. and Fuller, R. S. (1997) Internally consistent libraries of fluorogenic substrates demonstrate that Kex2 protease specificity is generated by multiple mechanisms. *Biochemistry* **37**, 3386–3391
- 27 Day, R., Schafer, M. K.-H., Cullinan, W. E., Watson, S. J., Chrétien, M. and Seidah, N. G. (1993) Region specific expression of furin mRNA in the rat brain. *Neurosci. Lett.* **149**, 27–30
- 28 Taguchi, S., Kumagai, I., Nakayama, J., Suzuki, A. and Miura, K. (1989) Efficient extracellular expression of a foreign protein in *Streptomyces* using secretory protease inhibitor (SSI) gene fusions. *Bio/Technology* **7**, 1063–1066
- 29 Strickler, J. E., Barka, T. R., Gorniak, J., Fornwald, J., Keys, R., Rowland, J. J., Rosenberg, M. and Taylor, D. P. (1992) Two novel *Streptomyces* protein protease inhibitors. Purification, activity, cloning, and expression. *J. Biol. Chem.* **267**, 3236–3242
- 30 Suzuki, K., Uyeda, M. and Shibata, M. (1981) Partial amino acid sequences of an alkaline protease inhibitor, API-2 (b and c). *Agric. Biol. Chem.* **45**, 629–634
- 31 Terabe, M., Kojima, S., Taguchi, S., Momose, H. and Miura, K. (1995) A subtilisin inhibitor produced by *Streptomyces bikiniensis* possesses a glutamine residue at reactive site P1. *J. Biochem. (Tokyo)* **117**, 609–613
- 32 Saito, N. and Nei, M. (1987) A neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425

Received 13 November 2000/2 January 2001; accepted 30 January 2001