Internally Consistent Libraries of Fluorogenic Substrates Demonstrate That Kex2 Protease Specificity Is Generated by Multiple Mechanisms†

Nathan C. Rockwell,‡,§ Gary T. Wang,|| Grant A. Krafft,⊥ and Robert S. Fuller*,‡

Department of Biological Chemistry, University of Michigan Medical School, Room 5413 Medical Science I, 1301 Catherine, Ann Arbor, Michigan 48109, Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305, Department of Combinatorial Chemistry, D-4CP AP10, Abbott Laboratories, Abbott Park, Illinois 60064, and Department of Molecular Pharmacology and Biological Chemistry, Northwestern University Medical School, Chicago, Illinois 60611-3008

Received July 19, 1996; Revised Manuscript Received December 2, 1996⊗

ABSTRACT: Kex2 protease from the yeast *Saccharomyces cerevisiae* is the prototype for a family of eukaryotic proprotein processing proteases. To clarify understanding of the interactions responsible for substrate recognition in this family of enzymes, we have carried out a systematic examination of Kex2 substrate specificity using internally consistent sets of substrates having substitutions at only one or two positions. We examined Kex2 sequence recognition for residues at P_3, P_2, and P_1 using two types of fluorogenic peptide substrates, peptidyl-methylcoumarinamides and internally quenched substrates in which cleavage occurs at an actual peptide bond. Kinetic analysis of the two sets of substrates gave comparable data on specificity at these three positions. For the best substrate sequences, high catalytic constants (k_{cat}/K_M) of $(2-5) \times 10^7$ M^{-1} s^{-1} were seen for cleavage of both peptidyl-methylcoumarinamides and peptide bonds. While no evidence for positive interactions with the P_3 residue emerged, Kex2 was found to discriminate against at least one residue, Asp, at this position. Specificity at P_2 was shown to rely primarily on recognition of a positive charge, although steric constraints on the P_2 side chain were also apparent. Kex2 was demonstrated to be exquisitely selective for Arg at P_1. Substitutions with similar charge (Lys, ornithine) or similar hydrogen-bonding capability (citrulline) do not confer efficient catalysis. Comparison of otherwise identical substrates having either Arg or citrulline at P_1 showed that the positive charge of the Arg guanidinium group stabilizes the transition state by approximately 6.8 kcal/mol.

In eukaryotes, many proteins that pass through the secretory pathway are initially synthesized as proproteins that are subjected to both endoproteolysis and exoproteolysis in late secretory compartments to yield their mature forms. Such proteins include both secreted effectors such as insulin and the *Saccharomyces cerevisiae* mating pheromone α-factor and membrane-bound proteins such as HIV gp160 and insulin receptor (Fuller et al., 1988; Rouille et al., 1995; Seidah et al., 1994). Typically, cleavage of these precursors takes place after dibasic sites (Lys-Arg or Arg-Arg) or Arg-Xaa-Xaa-Arg[1] motifs.

The mating pheromone α-factor is synthesized as a precursor, pro-α-factor, containing one Ser-Leu-Asp-Lys-Arg↓Glu-Ala-Glu-Ala cleavage site between a prosequence and the α-factor repeat domain and three Pro-Met-Tyr-Lys-Arg↓Glu-Ala cleavage sites that separate the four α-factor repeat units (Kurjan & Herskowitz, 1982). Kex2 protease ("kexin", E.C. 3.4.21.61) is a transmembrane serine protease in the yeast *trans*-Golgi network that cleaves pro-α-factor carboxyl to these Lys-Arg sites (Julius et al., 1984; Fuller et al., 1988). Kex2 is the prototype of a family of homologous eukaryotic enzymes that appears to account for most basic residue-directed cleavages of proproteins in late secretory compartments (Fuller et al., 1989b; Smeekens & Steiner, 1990; Rouille et al., 1995; Seidah et al., 1994). Kex2 and its homologs share a domain that is approximately 30% identical to the bacterial serine protease subtilisin (Mizuno et al., 1988; Fuller et al., 1989a; Siezen et al., 1991; Perona & Craik, 1995). The Kex2 family of proprotein processing enzymes can be distinguished from a degradative family of subtilisins both by the high sequence specificity found in processing enzymes (Brenner & Fuller, 1992) and by the presence of an additional conserved domain C-terminal to the subtilisin domain which is present in the Kex2 family but not in the degradative subtilisins (Gluschankof & Fuller, 1994).

Catalytically active Kex2 has been purified from yeast as a secreted, soluble enzyme generated by insertion of a stop codon prior to the C-terminal transmembrane domain (Brenner & Fuller, 1992). The substrate specificity of purified Kex2 was characterized using a set of commercially available

† This work funded by NIH Training Grant 2T32GM07599 (N.R.) and NIH GM39697 (R.S.F.).

* Corresponding author: telephone, (313) 936-9764; Fax, (313) 763-7799; email, bfuller@umich.edu.

‡ University of Michigan Medical School.

§ Stanford University School of Medicine.

|| Abbott Laboratories.

⊥ Northwestern University Medical School.

⊗ Abstract published in *Advance ACS Abstracts,* January 15, 1997.

[1] Abbreviations: Ac, acetyl; AMC, 7-amino-4-methylcoumarin; Bistris, bis(2-hydroxyethyl)iminotris(hydroxymethyl)methane; boc, *t*-butoxycarbonyl; Cit or Ç, citrulline; DABCYL, 4-[4-(dimethylamino)-phenyl]azobenzoyl; DMF, dimethylformamide; DMSO, dimethyl sulfoxide; EDANS, 5-[(2-aminoethyl)amino]naphthalene-1-sulfonic acid; Glu(EDANS) or B, 5-[(2-aminoethyl)amino]naphthalene-1-sulfonic acid conjugated to the γ-carboxylate of glutamic acid; Lys(DABCYL) or J, 4-[4-(dimethylamino)phenyl]azobenzoyl conjugated to the ϵ-amino group of lysine; MCA, C-terminal methylcoumarinamide; Nle or β, norleucine; Orn or O, ornithine; TFA, trifluoroacetic acid; Xaa, an arbitrary amino acid; ↓, the scissile bond. We follow the nomenclature of Schechter and Berger (1967) in designating the cleavage site as -P_3-P_2-P_1↓P_1'-P_2'-P_3'-, with the scissile bond between P_1 and P_1' and the C-terminus of the protein on the prime side.

S0006-2960(96)01779-5 CCC: $14.00

peptidyl-methylcoumarinamide substrates (AMC substrates), and purified Kex2 exhibited a marked preference for dibasic sites having Arg at P_1 (Brenner & Fuller, 1992), in keeping with computer models of Kex2 homologs showing specific interactions with Arg at this position (Siezen et al., 1994; Lipkind et al., 1995).

The substrates used in this previous work on Kex2 lacked a consistent sequence context, complicating the interpretation of the data. Moreover, these substrates contained the activated fluorogenic leaving group AMC, leaving open the possibility that activity as measured might not reflect the specificity of the enzyme in cleaving authentic peptide bonds. Here, we measured k_{cat}/K_M for Kex2 cleavage of libraries of internally quenched (IQ) fluorogenic peptide substrates (Matayoshi et al., 1990; Wang & Krafft, 1992) and AMC substrates that allowed specificity to be assessed at P_3, P_2, and P_1 with single amino acid substitutions in otherwise identical substrates based on cleavage sites in the physiological substrate pro-α-factor. Kex2 does not exhibit positive recognition at P_3 but may exclude Asp from the P_3 subsite. In contrast, both P_2 and P_1 are recognized by Kex2. The P_2 pocket discriminates against large side chains and recognizes positively charged residues, while the P_1 pocket is extremely selective for both the charge and structure of Arg.

MATERIALS AND METHODS

Substrates and Reagents. Standard reagents were from Sigma, Aldrich, or Fisher. Ac-Pro-Met-Tyr-Lys-Arg-MCA was the generous gift of J. Thorner. AMC for use as a standard was from Peninsula Laboratories and AMC for use in synthesis was from Aldrich.

Kex2 Purification. Kex2 was purified from culture media of yeast strain CB017 transformed with plasmid pG5-*KEX2Δ613*. Purification was carried out essentially as described (Brenner et al., 1994) with the following modifications. After pressure dialysis, enzyme was further purified by Mono Q chromatography (Brenner & Fuller, 1992). Upon elution, an initial peak of absorbance at 280 nm that did not correlate with activity was followed by two peaks of activity. The first peak had slightly higher specific activity. This fraction was pooled and concentrated by centrifugation in a Centricon 30 (Amicon) to 8800 units/μL. Overall yield was 15%. Final specific activity was 2.6×10^7 units/mg. Active-site titration was done by measuring an initial burst of product formation using a Kintek RQF-3 rapid-quenched flow apparatus as described (Brenner & Fuller, 1992). Standard Kex2 assays were performed at 37 °C in 0.2 M Bistris (pH 7.0 at 37 °C), 1 mM $CaCl_2$, 1% (v/v) DMSO or 0.2−1% (v/v) acetonitrile (depending on the substrate), and 0.1% (v/v) Triton X-100. A unit of Kex2 activity was 1 pmol of AMC produced/min, as described (Fuller et al., 1989a).

IQ Substrates. IQ substrates were prepared as described elsewhere (Wang & Krafft, 1992) and purified by reverse-phase HPLC on a Rainin Dynamax C8 column (21.4 mm × 250 mm). All substrates were >90% pure as judged by HPLC. For quantitation, substrates were cleaved to completion using an excess of either trypsin (Sigma) or Kex2, and a known amount of either peptidyl-EDANS (Pro-Ile-Gln-Val-His-EDANS) or glutamyl-EDANS was used to calibrate the fluorescence signal. Twelve substrates (IQ substrates I-3 through I-14) were of the general form Arg-Lys-(DABCYL)-Nle-Xaa-Xaa-Xaa↓Glu-Ala-Glu-Ala-Glu-(EDANS)-Arg, based on the sequence of the Kex2 cleavage site that occurs between α-factor repeats in pro-α-factor (Met-Tyr-Lys-Arg↓Glu-Ala-Glu-Ala). Nle was chosen as a near isosteric replacement for Met to avoid problems with the oxidation of the thioether. Two additional IQ substrates (I-1 and I-2) were Arg-Glu(EDANS)-Pro-Met-Tyr-Lys-Arg↓Glu-Ala-Glu-Ala-Lys(DABCYL)-Arg and Arg-Glu-(EDANS)-Ser-Leu-Asp-Lys-Arg↓Glu-Ala-Glu-Ala-Lys(D-ABCYL)-Arg, in which the relative order of the DABCYL and EDANS groups was reversed. The first of these substrates contained the Pro-Met-Tyr-Lys-Arg sequence found between α-factor repeats in pro-α-factor. The other contained the Ser-Leu-Asp-Lys-Arg sequence which represents the Kex2 cleavage site that separates the pro-domain from the first α-factor repeat in pro-α-factor.

AMC Substrates. AMC substrates were prepared using a modification of the methodology of Zimmerman et al. (1977). Peptides (University of Michigan core facility) were used as TFA salts. In a typical synthesis, 50 μmol of peptide was added to 400 μL of anhydrous DMF on ice. *N*-Methylmorpholine was added to a stoichiometry of 1:2 with TFA (stoichiometry with the peptide therefore varied depending on the peptide sequence), followed by addition of isobutyl chloroformate to a stoichiometry of 2:1 with peptide. Mixed anhydride formation was allowed to proceed for 30 min on ice, and then AMC was added in 10-fold molar excess over peptide in 1 mL of DMF at room temperature. Reactions were shaken at room temperature for 24 h and then transferred to −80 °C. Substrates were purified by reverse-phase HPLC (Rainin Dynamax C8 or C18 column, 21.4 × 250 mm) using water/acetonitrile gradients containing 0.1% TFA. All substrates eluted between 15% and 30% acetonitrile. Substrates were identified by absorption spectroscopy (scans at 200−450 nm), mass spectroscopy, and digestion with trypsin or subtilisin BPN′ (Sigma). Enzymatic digests with trypsin or subtilisin confirmed not only chemical identity of the substrate but also retention of configuration at the α-carbon of the P_1 residue. Synthetic yields varied with peptide sequence, as would be expected for unprotected peptides with reactive side chains, but were between 15% and 30% of peptide for all substrates. Pseudo-first-order kinetics for all substrates except Ac-Pro-Nle-Tyr-Lys-Cit-MCA (substrate A-11, Table 3) were carried out with HPLC peaks, but cleavage of substrate A-11 could not be detected at such a low substrate concentration. This substrate was therefore concentrated in a Speedvac concentrator (Savant), further purified by gel filtration in 0.1 M acetic acid on a Bio-Gel P-2 column (10 × 1.5 cm), and concentrated in a Speedvac prior to cleavage with Kex2. Concentration of this substrate was determined by cleavage to completion with subtilisin BPN′ (Sigma).

Determination of k_{cat}/K_M Using Pseudo-First-Order Kinetics. Reactions (600 μL) were started by addition of enzyme (final concentration, 52 pM−520 nM). Samples (100 μL) taken at times between 2 and 50 min were quenched with 600 μL of 0.5 M $ZnSO_4$. Enzyme activity did not decay under assay conditions as shown both by control preincubation experiments and by the linearity of time courses with saturating substrate concentrations (data not shown). Fluorescence was measured using a Perkin-Elmer LS5B luminescence spectrometer (IQ substrates $\lambda_{ex} = 338$ nm, $\lambda_{em} = 490$ nm; AMC substrates $\lambda_{ex} = 385$ nm, $\lambda_{em} = 465$ nm).

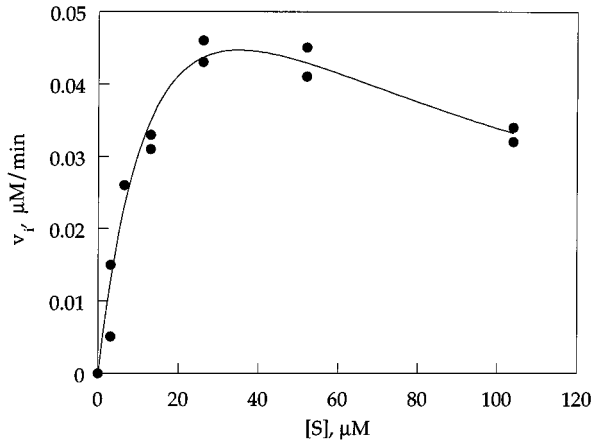

FIGURE 1: IQ substrates do not exhibit normal Michaelis−Menten kinetics. Increasing concentrations of substrate begin to cause inhibition. Data shown are for substrate I-7 and were fit to an equation describing substrate inhibition (eq 2; Fersht, 1985).

Raw fluorescence data were fitted to a first-order curve with floating end points by nonlinear regression using Kaleidagraph v3.0:

$$\text{fluorescence} = F_0 e^{-kt} + F_\infty \quad (1)$$

where F_0 is the amplitude of the change in fluorescence, k is the apparent first-order rate constant, and F_∞ is the end point of the reaction. The resulting apparent first-order rate constants were then divided by enzyme concentration to determine k_{cat}/K_M. The values for k_{cat}/K_M were shown to be independent either of substrate concentration or of enzyme concentration for each substrate, confirming that reactions proceeded under pseudo-first-order conditions. At least two substrate concentrations differing by 10-fold or two enzyme concentrations differing by 4-fold were tested for each substrate. Substrate concentrations ranged between 100 nM and 5 μM. Only progress curves with $r > 0.99$ (Pearson's r-value) were used in determination of k_{cat}/K_M values. Reported kinetic constants are the mean of at least three trials, and errors are reported as standard deviations about the mean.

Saturation Kinetics. Reactions (100 μL) were started by addition of enzyme and quenched by addition of 600 μL of 0.5 M $ZnSO_4$. For IQ substrates, reactions were calibrated with either peptidyl-EDANS or glutamyl-EDANS as a standard. Initial reaction velocities for IQ substrates were fitted to an equation descibing substrate inhibition (Fersht, 1985; a modified form of the Michaelis−Menten equation in which K_s' is the binding constant for the second substrate molecule):

$$v_i = \frac{k_{cat}[E_0][S]}{[S] + K_M + [S]^2/K_s'} \quad (2)$$

RESULTS

Pseudo-First-Order Kinetics with IQ Substrates. To examine Kex2 cleavage of IQ substrates, we attempted to determine k_{cat} and K_M by Michaelis−Menten kinetics as had been done for AMC substrates. However, simple saturation kinetics were not observed. Instead, with each of the five substrates examined, initial velocity decreased at high substrate concentrations, suggesting some form of substrate inhibition (Figure 1). For IQ substrate I-1 (Table 1), inhibition became significant by 10 μM substrate (data not shown). In contrast, cleavage of IQ substrate I-10 (Table 2) by trypsin in otherwise identical conditions did not display detectable substrate inhibition with substrate concentration in excess of 350 μM, showing that the apparent inhibition is true substrate inhibition of Kex2 protease and is not an artifact caused by intermolecular quenching (data not shown). This behavior made it impossible to determine k_{cat} and K_M reliably for IQ substrates. Substrate inhibition was not observed with AMC substrates (Brenner & Fuller, 1992; Rockwell and Fuller, unpublished experiments).

Substrate inhibition requires binding of a second substrate molecule to an enzyme−substrate complex or intermediate (Fersht, 1985). Therefore, at substrate concentrations far below K_M, substrate inhibition should not occur as most of the enzyme will be free. k_{cat}/K_M can be measured under these conditions because cleavage reactions will obey apparent first-order kinetics. Therefore, we used pseudo-first-order kinetics (Figure 2) to measure k_{cat}/K_M for an internally consistent library of IQ substrates and confirmed cleavage sites for five representative substrates by amino acid sequencing (data not shown; sequenced substrates are

Table 1: P_3 Substitutions[a]

substrate	sequence	k_{cat}/K_M (M^{-1} s^{-1})	relative k_{cat}/K_M	error (%)
	(A) IQ Substrates			
I-1	RJβYKR↓EAEABR	2.5×10^7	1	7
I-2	RJβKKR↓EAEABR*	4.1×10^7	1.6	16
I-3	RJβEKR↓EAEABR	1.6×10^7	0.64	11
I-4	RBSLDKR↓EAEAJR	2.6×10^6	0.10	38
I-5	RBPMYKR↓EAEAJR	3.1×10^7	1.2	4
	(B) AMC Substrates			
A-1	AcPMYKR↓MCA	4.5×10^7	1	12
A-2	AcPβYKR↓MCA	3.4×10^7	0.76	5
A-3	AcSLNKR↓MCA	3.9×10^7	0.87	9

[a] Values are the mean of at least three independent trials for each substrate, and errors are reported as the standard deviation of those trials. Cleavage sites are indicated with an arrow (↓), and IQ substrate sequences are therefore in the form -P_3-P_2-P_1↓P_1'-P_2'-P_3'-. IQ substrate cleavage sites indicated with asterisks were determined by amino acid sequencing. AMC substrate sequences are of the form -P_3-P_2-P_1↓MCA, where MCA indicates the C-terminal methylcoumarinamide.

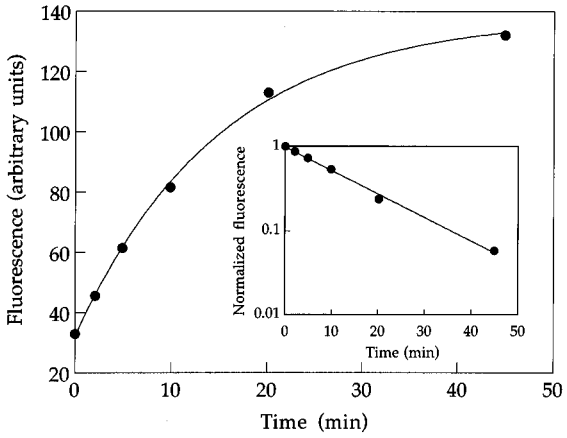

FIGURE 2: IQ substrates follow typical pseudo-first-order kinetics. Raw fluorescence data (arbitrary units) are plotted versus time for a typical reaction. Inset, the data were normalized and plotted on a log scale. Data are for substrate I-8 (Table 2), and they have been fitted to a first-order curve (eq 1) using nonlinear regression. Substrate concentration was 4.2 μM and enzyme concentration was 10.4 nM for this experiment.

indicated with asterisks in the tables). We also determined k_{cat}/K_M for two IQ substrates in which the fluorophore and chromophore were reversed relative to the cleavage site. For comparison, we also examined a number of AMC substrates containing substitutions within similar sequence contexts.

Evidence for Exclusion but Not Recognition at P_3. The IQ substrate library contained substitutions of Lys and Glu for the Tyr found at this position between α-factor repeats (Table 1, substrates I-1 through I-3). Neither of these substitutions produced more than a 2-fold effect on k_{cat}/K_M, suggesting that the nature of the P_3 residue is unimportant for efficient cleavage of Lys-Arg substrates. Although there is no crystal structure yet available for Kex2 or related processing proteases, structures of bacterial subtilisins with bound inhibitors do not show extensive interactions with P_3 (Takeuchi et al., 1991), and kinetic examination of subtilisin specificity also suggests that this substrate residue is relatively unimportant for specificity (Grøn et al., 1992).

However, analysis of other substrates suggested that at least one substitution at P_3 may be unfavorable. Substrates I-4 and I-5, which corresponded to the two different cleavage sites in pro-α-factor, exhibited a 10-fold difference in k_{cat}/K_M. These substrates differed at P_3, P_4, and P_5. Substrate I-4 corresponded to the cleavage site between the prodomain and the first α-factor repeat in pro-α-factor and was cleaved with a k_{cat}/K_M of 2.6×10^6 M^{-1} s^{-1}. Substrate I-5 corresponded to the cleavage site separating the α-factor repeats in pro-α-factor and was cleaved with an almost 10-fold higher k_{cat}/K_M of 3.1×10^7 M^{-1} s^{-1}. k_{cat}/K_M for IQ substrate I-5 was essentially identical to that for IQ substrate I-1 (2.5×10^7 M^{-1} s^{-1}, Table 1), which differed at P_4 and P_5. This suggested that the difference at P_3 between IQ substrates I-4 and I-5 was responsible for their difference in k_{cat}/K_M. Further support for this view came from examination of AMC substrates. AMC substrates A-1 and A-2, comparable to I-1 and I-4, were cleaved with very similar k_{cat}/K_M values. However, substrate A-3, based on the same pro-α-factor cleavage site as substrate I-5 but containing Asn at P_3 rather that Asp, was cleaved with a k_{cat}/K_M essentially identical to those for substrates with Tyr at P_3 (Table 1). Thus, the difference at P_3 between IQ substrate I-5 (P_3 Asp) and AMC substrate A-3 (P_3 Asn) offers the most plausible explanation for the defect in cleavage of substrate I-5. However, substitution of Glu for Tyr at P_3 produced a negligible effect (compare IQ substrates I-1 and I-3, Table 2). Thus, Kex2 may exclude Asp at P_3 despite the apparent lack of specific recognition at this position.

Electrostatic and Steric Contributions to P_2 Recognition. Previous work indicated that Kex2 is specific for basic residues at P_2 (Brenner & Fuller, 1992). This result is borne out by both substrate libraries examined here, but similar P_2 substitutions routinely produced better IQ substrates than AMC substrates (Table 2). Such differences may indicate that amino acids C-terminal to the scissile bond are able to compensate for P_2 substitutions, or it may reflect differences between the sequence contexts of these libraries.

Nonetheless, both libraries gave similar numbers for cleavage of substrates containing either Lys or Arg at P_2 (Table 2, substrates I-1, I-6, and A-3). Substitution of the shorter basic residue Orn produced a modest defect in catalysis (substrate I-7).

As might be expected, replacement of Lys with oppositely charged acidic residues produced a profound defect in cleavage (substrates I-8 and A-4). The uncharged residue Gln was more favorable than Asp at P_2 (substrate A-5); in fact, Gln and Leu (substrate A-6) at P_2 resulted in very similar k_{cat}/K_M values, possibly because both are uncharged γ-branched amino acids.

Table 2: P_2 Substitutions[a]

substrate	sequence	k_{cat}/K_M (M^{-1} s^{-1})	relative k_{cat}/K_m	error (%)
	(A) IQ Substrates			
I-1	RJβYKR↓EAEABR	2.5×10^7	1	7
I-6	RJβYRR↓EAEABR	3.0×10^7	1.2	28
I-7	RJβYOR↓EAEABR	7.5×10^6	0.3	13
I-8	RJβYER↓EAEABR*	1.1×10^5	0.0044	10
I-9	RJβYβR↓EAEABR	1.4×10^6	0.056	43
I-10	RJβYFR↓EAEABR*	5.6×10^4	0.0022	21
I-11	RJβYPR↓EAEABR	3.5×10^6	0.14	37
	(B) AMC Substrates			
A-3	AcSLNKR↓MCA	3.9×10^7	1	9
A-4	AcSLNDR↓MCA	1.4×10^4	3.6×10^{-4}	29
A-5	AcSLNQR↓MCA	6.8×10^4	0.0017	5
A-6	AcSLNLR↓MCA	5.0×10^4	0.0013	7
A-7	AcSLNβR↓MCA	2.2×10^5	0.0056	11
A-8	AcSLNAR↓MCA	1.3×10^5	0.0033	9
A-9	AcSLNYR↓MCA	1.5×10^4	3.8×10^{-4}	31

[a] Values are the mean of at least three independent trials for each substrate, and errors are reported as the standard deviation of those trials. Cleavage sites are indicated with an arrow (↓), and IQ substrate sequences are therefore in the form -P_3-P_2-P_1↓P_1'-P_2'-P_3'-. IQ substrate cleavage sites indicated with asterisks were determined by amno acid sequencing. AMC substrate sequences are of the form -P_3-P_2-P_1↓MCA, where MCA indicates the C-terminal methylcoumarinamide.

These libraries provide clear evidence that recognition of the Lys side chain occurs primarily through the ϵ-amino group. Substitution of Nle for Lys at P_2 resulted in a large defect in catalysis (substrates I-9 and A-7), suggesting that the aliphatic side chain common to these two residues is unimportant for efficient cleavage. Furthermore, k_{cat}/K_M for substrate A-8, which contained Ala at P_2, was within 2-fold of that for substrate A-7, which contained Nle at the same position. We conclude that the energetic contribution to catalysis by the aliphatic portion of the Lys side chain is therefore negligible.

However, substitution of larger hydrophobic residues resulted in more profound defects in cleavage than were seen with Nle or Ala at P_2 (substrates I-10, A-6, and A-9). While Leu was only slightly worse than Ala at P_2, substrates containing aromatic residues at P_2 had k_{cat}/K_M values that were lower than Ala-Arg or Nle-Arg substrates by an order of magnitude. This may reflect steric clash with the S_2 subsite.

In agreement with *in vivo* observations (Zhu et al., 1991), we found that substitution of Pro at P_2 produces only a modest defect in catalysis (substrate I-11). Degradative subtilisins are also able to cleave substrates containing Pro at P_2 efficiently (Grøn et al., 1992), and crystal structures of subtilisin BPN′ with different inhibitors containing Pro or Thr at the P_2 position show that Pro adopts an alternative conformation in which the side chain does not interact with the P_2 pocket (Takeuchi et al., 1991; Heinz et al., 1991), suggesting that this could be the case for Kex2 as well.

In summary, Kex2 specificity at P_2 arises from positive recognition of basic residues coupled with steric exclusion of bulky residues such as Tyr. Comparison of the numbers we present here suggests that the Lys amino group provides

Table 3: P_1 Substitutions[a]

substrate	sequence	k_{cat}/K_M (M^{-1} s^{-1})	relative k_{cat}/K_M	error (%)
	(A) IQ Substrates			
I-1	RJβYKR↓EAEABR	2.5×10^7	1	7
I-12	RJβYKK↓EAEABR*	3.6×10^5	0.014	42
I-13	RJβYKO↓EAEABR	2.2×10^5	0.0088	18
I-14	RJβYRK↓EAEABR*	4.7×10^4	0.0019	28
	(B) AMC Substrates			
A-1	AcPβYKR↓MCA	3.4×10^7	1	5
A-10	AcPβYKK↓MCA	1.5×10^5	0.0044	13
A-11	AcPβYKÇ↓MCA	<500	1.5×10^{-5}	N/D

[a] Values are the mean of at least three independent trials for each substrate, and errors are reported as the standard deviation of those trials. Cleavage sites are indicated with an arrow (↓), and IQ substrate sequences are therefore in the form -P_3-P_2-P_1↓P_1'-P_2'-P_3'-. IQ substrate cleavage sites indicated with asterisks were determined by amino acid sequencing. AMC substrate sequences are of the form -P_3-P_2-P_1↓MCA, where MCA indicates the C-terminal methylcoumarinamide.

about 3 kcal/mol of transition-state stabilization energy, while bulky residues are disfavored by approximately 5 kcal/mol relative to Lys or 2 kcal/mol relative to Ala.

Arginine Is Required at P_1 for Efficient Cleavage. Kex2 protease is clearly specific for Arg at P_1. Substitution of either Lys or Orn for Arg at P_1 (Table 3, substrates I-12 and I-13) resulted in an approximately 100-fold drop in k_{cat}/K_M relative to the canonical substrate sequence (IQ substrate I-1). The defect with Lys at P_1 was exacerbated by replacing Lys at P_2 with Arg (IQ substrate I-14). The resulting substrate was the most poorly cleaved IQ substrate tested in this study. However, k_{cat}/K_M for substrate I-12 was elevated approximately 100-fold relative to that for cleavage of boc-Glu-Lys-Lys-MCA (Brenner & Fuller, 1992) because of the presence of a favorable P_4 residue in these IQ substrates (Rockwell and Fuller, unpublished results).

k_{cat}/K_M for AMC substrate A-10, which contains Lys at P_1, is less than 3-fold below that of the comparable IQ substrate I-12, indicating that similar substitutions have quantitatively similar effects in both types of substrates. The defect in cleavage seen upon substitution of Lys for Arg at P_1 also indicates that recognition of Arg at P_1 relies on more than simple electrostatics, since both the Lys and Orn amino groups are positively charged at the pH employed in our assays.

The fact that recognition of Arg at P_1 is not simply electrostatic suggested that other structural features such as shape or hydrogen-bonding capability could be involved in Kex2 specificity for P_1 Arg. We therefore examined substrate A-11, in which the P_1 residue is citrulline. Although Cit is uncharged, it is essentially isosteric with Arg and contains similar hydrogen-bond donor and acceptor groups. Nonetheless, k_{cat}/K_M for substrate A-11 was below 500 M^{-1} s^{-1}. Loss of a positive charge at this position thus results in a 68 000-fold drop in catalytic efficiency, indicating the importance of a positively charged residue at P_1. The results obtained with substrates A-1 and A-11 suggest that the positive charge on the Arg guanidinium group contributes >6.8 kcal/mol of transition-state stabilization energy.

DISCUSSION

We have examined the substrate specificity of Kex2 protease at three positions: P_3, P_2, and P_1. Kex2 is able to recognize basic residues and exclude aromatic residues at P_2 and selectively cleave C-terminal to P_1 Arg. Though we found no evidence for recognition at P_3, the data suggest that Asp may be preferentially excluded from this position. Kex2 can efficiently cleave optimal peptidyl-methylcoumarinamides and internally quenched peptide substrates with k_{cat}/K_M of $(2-5) \times 10^7$ M^{-1} s^{-1}. The similarity of these values regardless of the nature of the leaving group indicates that the activated coumarin leaving group has little or no effect on catalysis. These k_{cat}/K_M values are similar to those reported for alkaline phosphatase, which has been shown to be diffusion-controlled (Simopoulos & Jencks, 1994). However, the presence of an initial burst of product formation in cleavage of such substrates (Brenner & Fuller, 1992) suggests that Kex2 is not limited by diffusional encounter with substrate under steady-state conditions, since the first turnover must also rely on such an encounter but occurs at a considerably higher rate.

The two types of substrate we have used here give comparable results. Substrates with analogous P_2 substitutions in the two libraries produce qualitatively similar results, but the IQ substrates were consistently cleaved with approximately 5-fold higher k_{cat}/K_M than comparable AMC substrates. However, k_{cat}/K_M values for IQ and AMC substrates containing the same P_1 substitution (Lys for Arg) differed by less than 3-fold, and dibasic substrates with Arg at P_1 were cleaved with almost identical k_{cat}/K_M values regardless of the nature of the leaving group. Therefore, the difference between similar P_2 substitutions in different types of substrates may indicate that residues C-terminal to the scissile bond can partially compensate for poor P_2 substitutions.

The results obtained with both model peptide libraries correlated well with an extensive, quantitative analysis of P_2 specificity of Kex2 protease *in vivo* utilizing a library of P_2 substitutions in pro-α-factor (A. Bevan, C. Brenner, and R. S. Fuller, unpublished experiments). These correlations support the notion that the factors involved in recognition of small, unstructured peptide substrates are paramount in understanding efficient cleavage of structured proprotein substrates in the cell.

Although P_2 substitutions produce large changes in k_{cat}/K_M, the defects in catalysis produced by P_1 substitutions were even more profound despite the much more conservative nature of the P_1 substitutions examined to date. This would suggest that more radical P_1 substitutions would produce still larger defects in catalysis. The Kex2 P_1 pocket would thus seem to be much more specific than the P_2 pocket, as illustrated by comparison of substrates I-1, I-6, and I-12. k_{cat}/K_M values for substrates containing Arg and Lys at P_2 differed by a factor of 1.2, while values for the equivalent pair of P_1 substitutions differed by a factor of approximately 70. Understanding the structural basis of the exquisite specificity of Kex2 for P_1 Arg may well clarify the enzymatic differences between the functionally distinct yet evolutionarily related digestive subtilisins and processing proteases. Indeed, much of the specificity of posttranslational proteolytic processing in eukaryotes may lie in the P_1 specificity of enzymes of the Kex2 family.

ACKNOWLEDGMENT

We thank the University of Michigan Protein and Carbohydrate Core Facility for peptide synthesis, amino acid

sequencing, and mass spectroscopy. We thank Alison Bevan and Michael S. Kay for assistance with Kex2 purification and members of the Fuller lab for helpful discussions.

REFERENCES

Brenner, C., & Fuller, R. S. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 922−926.
Brenner, C., Bevan, A., & Fuller, R. S. (1994) *Methods Enzymol. 244*, 152−167.
Fersht, A. R. (1985) in *Enzyme structure and mechanism*, pp 80−88, W. H. Freeman and Co., New York.
Fuller, R. S., Sterne, R. E., & Thorner, J. (1988) *Annu. Rev. Physiol. 50*, 345−362.
Fuller, R. S., Brake, A., & Thorner, J. (1989a) *Proc. Natl. Acad. Sci. U.S.A. 86*, 1434−1438.
Fuller, R. S., Brake, A., & Thorner, J. (1989b) *Science 246*, 482−486.
Gluschankof, P., & Fuller, R. S. (1994) *EMBO J. 13*, 2280−2288.
Grøn, H., Meldal, M., & Breddam, K. (1992) *Biochemistry 31*, 6011−6018.
Heinz, D. W., Priestle, J. P., Rahuel, J., Wilson, K. S., & Grutter, M. G. (1991) *J. Mol. Biol. 217*, 353−371.
Julius, D., Brake, A., Blair, L., Kunisawa, R., & Thorner, J. (1984) *Cell 37*, 1075−1089.
Kurjan, J., & Herskowitz, I. (1982) *Cell 30*, 933−943.
Lipkind, G., Gong, Q., & Steiner, D. F. (1995) *J. Biol. Chem. 270*, 13277−13284.
Matayoshi, E. D., Wang, G. T., Krafft, G. A., & Erickson, J. (1990) *Science 247*, 954−958.
Mizuno, K., Nakamura, T., Qhshima, T., Tanaka, S., & Matsuo, H. (1988) *Biochem. Biophys. Res. Commun. 156*, 246−254.
Perona, J. J., & Craik, C. S. (1995) *Protein Sci. 4*, 337−360.
Rouille, Y., Duguay, S. J., Lund, K., Furuta, M., Gong, Q., Lipkind, G., Oliva, A. A., Jr., Chan, S. J., & Steiner, D. F. (1995) *Frontiers Neuroendocrin. 16*, 322−361.
Schechter, I., & Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157−162.
Seidah, N. G., Chretien, M., & Day, R. (1994) *Biochimie 76*, 197−209.
Siezen, R. J., de Vos, W. M., Leunissen, J. A. M., & Dijkstra, B. W. (1991) *Protein Eng. 4*, 719−737.
Siezen, R. J., Creemers, J. W. M., & Van de Ven, W. J. M. (1994) *Eur. J. Biochem. 222*, 255−266.
Simopoulos, T. T., & Jencks, W. P. (1994) *Biochemistry 33*, 10375−10380.
Smeekens, S. P., & Steiner, D. F. (1990) *J. Biol. Chem. 265*, 322−361.
Takeuchi, Y., Satow, Y., Nakamura, K. T., & Mitsui, Y. (1991) *J. Mol. Biol. 221*, 309−325.
Wang, G. T., & Krafft, G. A. (1992) *Bioorg. Med. Chem Lett. 2*, 1665−1668.
Zhu, Y.-S., Zhang, X.-Y., Zhang, M., Cartwright, C. P., & Tipper, D. J. (1991) *Mol. Microbiol. 6*, 511−520.
Zimmerman, M., Ashe, B., Yurewicz, E. C., & Patel, G. (1977) *Anal. Biochem. 78*, 47−51.

BI961779L