# Interplay between $S_1$ and $S_4$ Subsites in Kex2 Protease: Kex2 Exhibits Dual Specificity for the $P_4$ Side Chain†

Nathan C. Rockwell‡,§,|| and Robert S. Fuller*,§

*Department of Biological Chemistry, University of Michigan Medical Center, Room 5413 Medical Science I, 1301 East Catherine, Ann Arbor, Michigan 48109, and Department of Biochemistry, B400 Beckman Center, Stanford University, Stanford, California 94305*

*Received October 13, 1997*

ABSTRACT: The yeast Kex2 protease is the prototype of a family of eukaryotic proprotein processing proteases that includes PC1, PC2, and furin. The catalytic domains of these enzymes are homologous to the degradative serine proteases of the subtilisin family. Kex2 exhibits optimal activity toward substrates with Lys or Arg at $P_2$ and Arg at $P_1$ (Lys-Arg or Arg-Arg cleavage sites). However, mammalian proprotein processing proteases such as furin exhibit more stringent requirements for basic residues at $P_4$ than at $P_2$. Here we demonstrate that Kex2 protease also recognizes $P_4$, with dual specificity for aliphatic and basic residues. Recognition of $P_4$ is even more readily apparent in substrates having a poor $P_1$ residue (Lys). Kinetic analysis of a series of otherwise identical fluorogenic substrates with Lys at $P_1$ and different residues at $P_4$ indicates that large, aliphatic $P_4$ residues increase $k_{cat}/K_M$ by 100-fold. However, smaller residues or acidic residues at $P_4$ do not. $P_4$ Arg also confers efficient cleavage on such a substrate, but the uncharged isostere of Arg, citrulline, does not. Kex2 may thus possess distinct subsites that recognize aliphatic or basic $P_4$ side chains. Although a favorable $P_4$ residue can partially compensate for the defects in $k_{cat}$ and $k_{cat}/K_M$ seen with Lys in place of Arg at $P_1$, this substitution resulted in a change in rate-determining step for all substrates examined. As previously seen in the case of subtilisin, effects of substitutions at the $P_1$ and $P_4$ positions were not independent, suggesting that interplay between these two positions is a common feature of substrate specificity for both processing proteases and degradative enzymes of the subtilisin superfamily.

Many eukaryotic proteins and peptides are expressed as proproteins or propeptide precursors that are posttranslationally cleaved at specific sites to generate the mature species. In the secretory pathway, such processing frequently occurs at basic motifs, and at least two families of enzymes capable of cleaving at such sites are now known (*1*−*3*). Of these two types of enzymes, a family of serine proteases homologous to the degradative proteases of the subtilisin family is the better understood, and the *Saccharomyces cerevisiae* member of this family, Kex2 protease (kexin, EC 3.4.21.61), has become a paradigm for understanding how these enzymes achieve the rapid catalysis and exquisite specificity that are key to their function as processing enzymes (*4*).

This family recognizes two principal types of sites, dibasic sites and sites of the form Arg-Xaa-Xaa-Arg↓,[1] although it is thought that Arg at the $P_6$ position may be able to substitute for the $P_4$ Arg in some contexts (*5*). Thus, these enzymes primarily possess specificity at the $P_4$, $P_2$, and $P_1$ positions, which are the principal sites of substrate specificity in the degradative subtilisins as well (*6*). It is not yet clear whether residues C-terminal to the scissile bond play a major role in substrate recognition, although the mammalian Kex2 homologue furin has been reported to cleave a substrate containing C-terminal residues more efficiently than substrates containing a C-terminal reporter group, aminomethylcoumarin (AMC, *7*). However, a recent examination of Kex2 specificity at the $P_3$, $P_2$, and $P_1$ positions showed no significant difference between cleavage of substrates with C-terminal amino acids and cleavage of AMC substrates (*8*).

Kex2 protease has been characterized in more detail than any other member of this family (*8, 9*). The enzyme exhibits high selectivity for positively charged residues at $P_2$ and an even more stringent requirement for Arg at $P_1$ (*8, 9*). Analysis of known cleavage sites suggested a preference for aliphatic residues at $P_4$ (*10*), but the activity of the purified enzyme with a set of commercially available peptidyl-MCA

† This work was supported by NIH Training Grant 2T32GM07599 (N.C.R.) and NIH Grant GM39697 (to R.S.F.).

* To whom correspondence should be addressed.

‡ Stanford University.

§ University of Michigan Medical Center.

|| Present address: Department of Biological Chemistry, University of Michigan Medical Center, Room 5413 Medical Science I, 1301 E. Catherine, Ann Arbor, MI 48109.

[1] Abbreviations: Abu or $\epsilon$, α-aminobutyric acid; Ac, acetyl; AMC, 7-amino-4-methylcoumarin; Bistris, [bis(2-hydroxyethyl)amino]tris-(hydroxymethyl)methane; boc, *tert*-butoxycarbonyl; Cha or $\chi$, ($\beta$-cyclohexyl)alanine; Cit or Ç, citrulline; DMSO, dimethyl sulfoxide; Lys(Z), ($\epsilon$-amino)benzyloxycarbonyllysine; MCA, C-terminal methylcoumarinamide; Nle or $\beta$, norleucine; Nvl or $\pi$, norvaline; Xaa, an arbitrary amino acid; Z, benzyloxycarbonyl; ↓, the scissile bond. Throughout we designate 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. Recognition sites for individual substrate side chains are designated $S_3$, $S_2$, $S_1$, etc. (*35*).

S0006-2960(97)02534-8 CCC: $15.00 
Published on Web 02/20/1998

substrates suggested $P_4$ recognition was of minor importance at most for efficient catalysis (*9*). Cleavage of the substrate Ac-Pro-Met-Tyr-Lys-Arg↓MCA by Kex2 exhibited an initial burst of product formation, demonstrating the existence of a rate-limiting step after acyl-enzyme formation (*9*).

Comparison of a tripeptidyl-MCA Lys-Lys substrate with a pentapeptidyl-MCA Lys-Arg substrate based on a physiological cleavage site suggested that substitution of Lys for Arg at $P_1$ could result in a defect in $k_{cat}/K_M$ of 3000-fold and a defect in $k_{cat}$ of >100-fold (*9*). However, a recent examination of Kex2 specificity using a set of extended peptide substrates with Lys, Orn, or Arg at $P_1$ showed that Lys-Lys or Lys-Orn sequences were cleaved with $k_{cat}/K_M$ values only 70−110-fold below the value observed with comparable Lys-Arg substrates (*8*). These substrates were cleaved with $k_{cat}/K_M$ values approximately 100-fold higher than would be expected based on comparison of previously reported AMC substrates (*9*). The extended substrates contained residues C-terminal to the scissile bond not present in AMC substrates and also contained additional residues N-terminal to the cleaved bond which were not present in the tripeptide Lys-Lys substrate examined previously (*9*). The improved cleavage of Lys-Lys sequences in the extended context raised the possibility that processing proteases might have specific contacts with the residues C-terminal to the cleavage site that are not present in the AMC substrates frequently used to examine the specificity of these proteases (*9*, *11*). However, it was also possible that interactions with the additional amino acids N-terminal to the cleavage site were responsible for this stimulation in cleavage. In particular, these extended substrates contained a $P_4$ residue (Nle), unlike the Lys-Lys tripeptidyl-MCA substrate previously examined; if the apparent preference for aliphatic residues at this position *in vivo* (*10*) were the result of Kex2 specificity at this position, then the improved cleavage of the extended Lys-Lys substrate might arise from the presence of a favorable $P_4$ side chain.

In this work, we synthesized a set of AMC substrates in which residues at $P_1$ and $P_4$ were systematically varied. Kinetic analysis of the cleavage of these substrates demonstrates that a favorable N-terminal contact between Kex2 and the side chain of the $P_4$ residue accounts for the entire increase in $k_{cat}/K_M$ observed with the extended Lys-Lys substrate. Moreover, Kex2 exhibited dual specificity at $P_4$, as aliphatic residues or basic residues at this position resulted in efficient cleavage but aromatic residues, small residues, uncharged polar residues, and acidic residues did not. A favorable $P_4$ interaction increased $k_{cat}$ for such substrates, but all Lys-Lys substrates examined exhibited a change in rate-determining step relative to Lys-Arg substrates, because the initial burst seen in cleavage of Lys-Arg substrates (*9*) was not seen in cleavage of even the best Lys-Lys substrate. This indicated that substitution of Lys for Arg at $P_1$ resulted in a defect in acylation regardless of the nature of the $P_4$ residue. Furthermore, the effects of substitutions at $P_4$ and $P_1$ were not independent, a phenomenon also seen with degradative subtilisins (*12*). The interaction between $P_1$ and $P_4$ recognition may therefore be a conserved feature of substrate recognition by both processing and degradative proteases of the subtilisin superfamily.

## MATERIALS AND METHODS

*Reagents, Substrates, and Enzyme*. All reagents were from Sigma, Aldrich, or Fisher. Preparation of Kex2 protease has been described elsewhere (*8*). Ac-Pro-Met-Tyr-Lys-Arg↓MCA was the generous gift of J. Thorner. Boc-Glu-Lys-Lys↓MCA and boc-Gln-Gly-Arg↓MCA were from Peninsula Laboratories, and other AMC substrates were prepared as described previously (*8*) except for Ac-Asp-Tyr-Lys-Lys↓MCA, which was prepared as follows. Ac-Asp-Tyr-Lys(Z)-Lys(Z) (43 mg, University of Michigan Core Facility) was reacted with 13 μL of isobutyl chloroformate and 3 μL of *N*-methylmorpholine in 400 μL of *N*-methylpyrrolidinone on ice for 30 min. AMC (88 mg) was then added in an additional 700 μL of *N*-methylpyrrolidinone, and the reaction was shaken overnight at room temperature to yield Ac-Asp-Tyr-Lys(Z)-Lys(Z)↓MCA in good yield (estimated at approximately 40% of starting peptide by HPLC). The mixture was then precipitated with ether and centrifuged. The precipitate was resuspended in 1 mL of methanol and added to 50 mg of palladium on activated charcoal (Degussa type E101 NO/W, Aldrich). This suspension was flushed with argon and then placed under hydrogen for 5 h. Palladium/carbon catalyst was then removed by filtration, and Ac-Asp-Tyr-Lys-Lys↓MCA was purified by HPLC as described (*8*). A small amount of product (<1% of starting peptide) was recovered after hydrogenation. A similar loss was not observed in preparation of Ac-Nle-Tyr-Lys-Arg↓MCA from Ac-Nle-Tyr-Lys(Z)-Arg↓MCA (95% yield at the hydrogenation step), indicating that this loss was specific to Ac-Asp-Tyr-Lys-Lys↓MCA. The identity of this substrate was confirmed by its UV and mass spectra. Although this substrate was cleaved well by trypsin (data not shown), cleavage of this substrate by Kex2 protease was so poor that it was possibly due to background cleavage by a trace contaminating protease. The value we report for $k_{cat}/K_M$ for this substrate is therefore only an estimated upper bound.

*Steady-State Kinetics*. Pseudo-first-order kinetics and reaction conditions were as previously described (*8*). For saturation kinetics, 100 μL reactions were started by addition of enzyme and quenched with 600 μL of 0.5 M $ZnSO_4$. All substrates were examined at substrate concentrations at least 4-fold above and below $K_M$ except for Ac-Ala-Tyr-Lys-Lys↓MCA, which was examined at substrate concentrations up to only 2.5-fold above $K_M$ because of limited solubility. Reactions were calibrated with substrate-only blanks with or without small amounts of free AMC as a standard, fluorescence was measured as described (*9*), and initial reaction velocities were fit directly to the Michaelis−Menten equation using nonlinear regression. Only reactions with <6% conversion of substrate to product were used in fitting. All values reported are the mean values of three independent trials, and errors are reported as standard deviation.

*Rapid-Quenched-Flow Measurements*. Reactions were carried out under standard conditions except that reactions were carried out at 21 °C. Enzyme was mixed with substrate in a Kintek RQF-3 rapid-chemical-quench mixer for times between 2 ms and 32 s depending on the substrate. Reactions were quenched by addition of 70 μL of 0.75 M sodium citrate, pH 2.95, followed by an additional 150 μL of quench 20 ms later upon exit of the reaction from the rapid-quench mixer. Each reaction was diluted with a further 400 μL of

Table 1: Kex2 Protease Possesses $P_4$ Specificity[a]

| sequence | $k_{cat}/K_M$ ($M^{-1}$ $s^{-1}$) | relative $k_{cat}/K_M$ | error (%) |
|---|---|---|---|
| Ac$\beta$YKK↓MCA | $9.2 \times 10^4$ | 1 | 7 |
| Ac$\pi$YKK↓MCA | $7.8 \times 10^4$ | 0.85 | 6 |
| Ac$\epsilon$YKK↓MCA | $5.7 \times 10^3$ | 0.062 | 6 |
| AcAYKK↓MCA | $1.2 \times 10^3$ | 0.013 | 18 |
| AcVYKK↓MCA | $1.4 \times 10^4$ | 0.15 | 6 |
| AcFYKK↓MCA | $3.7 \times 10^4$ | 0.40 | 21 |
| Ac$\chi$YKK↓MCA | $1.3 \times 10^5$ | 1.4 | 4 |
| AcRYKK↓MCA | $1.2 \times 10^5$ | 1.3 | 28 |
| AcÇYKK↓MCA | $4.9 \times 10^3$ | 0.053 | 11 |
| AcDYKK↓MCA | <250 | <0.0027 | N/D |
| Ac$\beta$YKR↓MCA | $4.3 \times 10^7$ | 1 | 23 |
| AcAYKR↓MCA | $4.3 \times 10^6$ | 0.10 | 2.3 |

[a] All data are reported as the mean of at least three independent trials, and errors are reported as standard deviations. Cleavage sites are indicated by arrows. $\epsilon$ = α-aminobutyric acid, $\pi$ = norvaline, $\beta$ = norleucine, $\chi$ = ($\beta$-cyclohexyl)alanine, and Ç = citrulline. Relative $k_{cat}/K_M$ values are normalized to the substrates with norleucine at $P_4$ for each $P_1$ residue.

quench prior to reading samples so that the final volume was suitable for use with standard fluorescence microcuvettes (Starna). No reactivity of the enzyme has been observed in the initial quenching volume. For the substrate boc-Gln-Gly-Arg↓MCA, reactions were quenched with 1 M acetic acid. No difference was observed between quench solutions.

## RESULTS

*Kex2 Recognizes the $P_4$ Side Chain.* Previously, we demonstrated that Kex2 cleaves an extended, internally quenched fluorogenic peptide substrate with Lys at $P_1$ and $P_2$ 100-fold more efficiently than a tripeptidyl-MCA substrate with Lys at these positions (*8*, *9*). Comparison of sequences of physiological Kex2 cleavage sites suggested a preference for aliphatic residues at the $P_4$ position (*10*). We hypothesized that recognition of the $P_4$ residue might be the specific interaction conferring reactivity on the extended substrates. We therefore synthesized a set of substrates of the form Ac-Xaa-Tyr-Lys-Lys↓MCA and examined their cleavage by Kex2 (Table 1). The results show clearly that Kex2 possesses specificity for the $P_4$ side chain and that $P_4$ recognition accounts for the observed discrepancy between the extended Lys-Lys substrate and the previously reported tripeptidyl-MCA substrate.

A long aliphatic side chain at $P_4$ was sufficient to confer improved cleavage upon a Lys-Lys substrate, but shorter side chains did not cause this enhancement (Table 1). The values of $k_{cat}/K_M$ for the substrates Ac-Nle-Tyr-Lys-Lys↓MCA and Ac-Ala-Tyr-Lys-Lys↓MCA differed by approximately 70-fold despite the fact these substrates differed only by three methylene groups in the $P_4$ side chain. Examination of equivalent substrates with Nvl or Abu at $P_4$ showed that the three-carbon alkyl chain of Nvl was also sufficient to raise $k_{cat}/K_M$ to a similar value ($9.2 \times 10^4$ $M^{-1}$ $s^{-1}$ with Nle at $P_4$, $7.8 \times 10^4$ $M^{-1}$ $s^{-1}$ with Nvl at $P_4$). However, the ethyl side chain of Abu did not confer such stimulation and resulted in an intermediate value of $k_{cat}/K_M$ (5700 $M^{-1}$ $s^{-1}$ with Abu at $P_4$, versus 1200 $M^{-1}$ $s^{-1}$ with Ala at $P_4$).

Interestingly, substituting Val at $P_4$ resulted in a $k_{cat}/K_M$ approximately 2-fold above the value seen with $P_4$ Abu, or approximately 6-fold below the value observed with $P_4$ Nvl. That is, the substrate containing an isopropyl side chain (Val) was significantly worse than the substrate containing an *n*-propyl side chain (Nvl) and only slightly better than the substrate containing an ethyl side chain (Abu). Therefore, the $P_4$ pocket of Kex2 can accommodate $\beta$-branched residues, but only a portion of such a residue's side chain can actually interact with the enzyme. However, the enzyme can accommodate quite large residues at this position; in fact, an equivalent substrate containing the bulky Cha (cyclohexyl-Ala) at $P_4$ had a $k_{cat}/K_M$ of $1.3 \times 10^5$ $M^{-1}$ $s^{-1}$, making it one of the best substrates in this library. In contrast, substitution of Phe for Nle at $P_4$ introduces a modest defect in cleavage (Table 1). Because a cyclohexyl moiety is accommodated at this position, the phenyl ring of Phe must be disfavored because of its rigid planar character.

Table 2: Interaction with the $P_4$ Side Chain Increases $k_{cat}$[a]

| sequence | $k_{cat}$ ($s^{-1}$) | $K_M$ ($\mu$M) |
|---|---|---|
| bocEKK↓MCA | 0.19 | 55 |
| bocQGR↓MCA | 6.1 | 320 |
| Ac$\beta$YKR↓MCA | 50 | 1 |
| Ac$\beta$YKK↓MCA | 5.0 | 38 |
| AcAYKK↓MCA | 0.36 | 350 |

[a] Data for substrates other than bocEKK-MCA and bocQGR-MCA are reported as the mean of three independent trials with standard deviations for all values below 35%. Cleavage sites are indicated by arrows. $\beta$ = norleucine. Data for bocEKK↓MCA and bocQGR↓MCA are from ref 9, but little difference in $k_{cat}$ and $K_M$ between Kex2 preparations was observed with these substrates (data not shown).

Kex2 does not exclusively recognize hydrophobic residues at $P_4$, since the substrate Ac-Arg-Tyr-Lys-Lys↓MCA was cleaved with an efficiency similar to substrates containing Nle or Cha at $P_4$ (Table 1). The favorable cleavage of a substrate with Arg at $P_4$ suggested that Kex2 either recognized the aliphatic portion of the Arg side chain, which is similar to the Nvl side chain, or possessed dual specificity at $P_4$, interacting favorably with either aliphatic or basic side chains. To distinguish between these possibilities, we examined cleavage of Ac-Cit-Tyr-Lys-Lys↓MCA, in which the $P_4$ side chain of Cit is isosteric to that of Arg but lacks positive charge. This substrate was cleaved poorly (approximately 20-fold lower $k_{cat}/K_M$ than with Nle or Arg at $P_4$; Table 1). This indicates that recognition of Arg at $P_4$ relies on the positive charge of the guanidinium group rather than the aliphatic portion of the Arg side chain. An equivalent substrate with Asp at $P_4$ was cleaved extremely poorly, as might be expected if the $P_4$ pocket recognizes hydrophobic or positively charged residues.

*Favorable $P_4$ Residues Increase $k_{cat}$.* Although a favorable $P_4$ side chain is clearly able to improve $k_{cat}/K_M$ for Lys-Lys substrates, such enhancement might conceivably arise from a decrease in $K_M$ rather than direct compensation for the severe $k_{cat}$ defect (>100-fold) seen with the tripeptidyl Lys-Lys-MCA substrate (Table 2; *9*). To distinguish between these possibilities, we measured $k_{cat}$ and $K_M$ for Ac-Nle-Tyr-Lys-Lys↓MCA and Ac-Ala-Tyr-Lys-Lys↓MCA (Table 2). The results show that the enhanced reactivity of Lys-Lys substrates with favorable $P_4$ residues was due largely to an improvement in $k_{cat}$. $K_M$ for Ac-Nle-Tyr-Lys-Lys↓MCA was decreased less than 2-fold relative to the $K_M$ for the tripeptidyl Lys-Lys substrate, but $k_{cat}$ was increased approximately 25-fold. In fact, $k_{cat}$ for the Lys-Lys substrate with Nle at $P_4$ was comparable to that for a Gly-Arg substrate (Table 2;

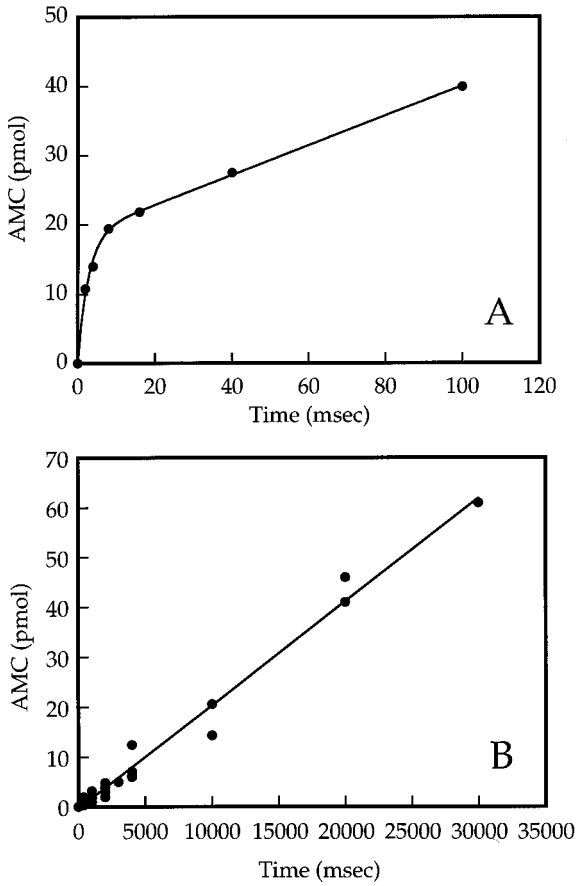


FIGURE 1: (A) Kex2 cleavage of Ac-Nle-Tyr-Lys-Arg↓MCA displays an initial burst of product formation (18.5 pmol of Kex2 per time point). (B). No such burst is seen with the substrate boc-Glu-Lys-Lys↓MCA (24 pmol of Kex2 per time point).

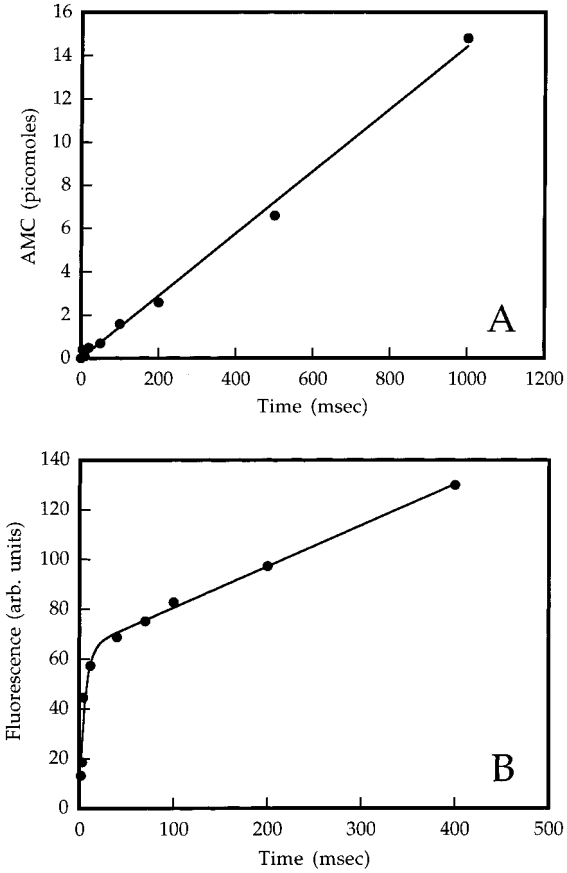


FIGURE 2: Although the substrate Ac-NleTyrLysLys↓MCA has a very similar $k_{cat}$ to the substrate boc-GlnGlyArg↓MCA (Table 2), these substrates have different rate-limiting steps. (A) No initial burst is seen with Ac-NleTyrLysLys↓MCA (15 pmol of Kex2 per time point). (B) A burst is seen with boc-GlnGlyArg↓MCA. Values are reported as raw fluorescence.

this substrate has the lowest $k_{cat}$ measured thus far with Arg at $P_1$). $k_{cat}$ for the Lys-Lys substrate with Ala at $P_4$ was only slightly higher than that for the tripeptide substrate boc-Glu-Lys-Lys↓MCA, while $K_M$ was actually about 6-fold higher. These results show that the nature of the $P_4$ side chain, rather than its mere presence, is required to compensate for the $k_{cat}$ defect observed with Lys-Lys substrates.

*Substitution of Lys for Arg at $P_1$ Changes the Rate-Limiting Step.* Unlike other serine proteases such as subtilisin BPN′ and chymotrypsin (*13*), Kex2 is unusual in that an initial burst of product formation is seen in its reaction with a Lys-Arg amide substrate (Figure 1A; *9*). Such a burst was observed in cleavage of an Arg-Arg substrate by the mammalian Kex2 homologue furin (*14*). However, the tripeptidyl Lys-Lys substrate boc-Glu-Lys-Lys↓MCA, which has a $k_{cat}$ defect of >100-fold (*9*), did not show such a burst of product formation (Figure 1B), indicating that the rate-determining step in cleavage of this substrate is different from that for the Lys-Arg substrate. However, as shown above, the presence of a favorable $P_4$ residue can restore $k_{cat}$ for a Lys-Lys substrate to a level comparable to that of certain substrates with Arg at $P_1$ and only an order of magnitude below the $k_{cat}$ seen with an optimal substrate (Table 2). This raised the possibility that the rapid acylation seen with substrates containing Arg at $P_1$ might also be restored in a Lys-Lys substrate by including a favorable $P_4$ residue.

However, cleavage of the substrate Ac-Nle-Tyr-Lys-Lys↓MCA by Kex2 did not exhibit an initial burst (Figure 2A), indicating that the partial compensation conferred by the $P_4$ side chain did not restore rapid acylation. The substrate boc-Gln-Gly-Arg↓MCA, whose $k_{cat}$ is comparable to that of Ac-Nle-Tyr-Lys-Lys↓MCA, did show an initial burst (Figure 2B), indicating that the rate-limiting step in cleavage of these two substrates is different despite their similar $k_{cat}$ values.

*Enzyme−Substrate Interactions at $P_4$ and $P_1$ Are Not Independent.* Although the $P_4$ side chain can compensate for Lys at $P_1$, previous work on Kex2 had shown little difference between tripeptidyl and pentapeptidyl Lys-Arg substrates (*9*). However, a mutant form of Kex2 with a perturbed oxyanion hole exhibited much more severe defects in cleavage of a tripeptidyl Lys-Arg substrate than in cleavage of an analogous pentapeptidyl substrate (*15*), suggesting that contacts beyond $P_3$ may compensate for a lesion in the active site. The work cited above relied on substrates differing in N-terminal protecting groups and $P_3$ residues. To examine $P_4$ effects on Lys-Arg substrates within consistent sequence contexts, we compared cleavage of Ac-Nle-Tyr-Lys-Arg↓MCA and Ac-Ala-Tyr-Lys-Arg↓MCA by wild-type Kex2 protease. The $k_{cat}/K_M$ values for these substrates differed by only 10-fold, as opposed to >70-fold for the equivalent Lys-Lys substrates (Table 1). Thus, the effect of an optimal $P_4$ interaction upon $k_{cat}/K_M$ is nearly 10-fold greater in the context of a poor $P_1$ residue (Lys) than with the optimal $P_1$ residue (Arg). Therefore, we conclude

Table 3: Physiological Kex2 Cleavage Sites[a]

| cleavage site | sequence | reference |
|---|---|---|
| k1 killer toxin (44↓45) | -LLPR↓EA- | 30 |
| k1 killer toxin (149↓150) | -VARR↓DI- | 30 |
| k1 killer toxin (188↓189) | -YVKR↓SD- | 31 |
| k1 killer toxin (233↓234) | -VAKR↓YV- | 30 |
| MFα1 pro-α-factor (85↓86) | -LDKR↓EA- | 32 |
| MFα1 pro-α-factor (104↓105)[b] | -MYKR↓EA- | 32 |
| KEX2 prodomain (109↓110) | -LFKR↓LP- | 9 |
| exo−1,3-$\beta$-glucanase (40↓41) | -NKKR↓YY- | 33 |
| MFα2 pro-α-factor (80↓81) | -LAKR↓EA- | 34 |
| MFα2 pro-α-factor (101↓102) | -MYKR↓EA- | 34 |
| *S. cerevisiae* HSP150 (71↓72) | -KAKR↓AA- | 17 |

[a] Eleven known Kex2 cleavage sites from *S. cerevisiae* are listed. Residues surrounding the cleavage site are shown from $P_4$ to $P_2'$. The residue numbers of $P_1$ and $P_1'$ are given, and cleavage sites are indicated by arrows. [b] The same cleavage site is repeated at 125↓126 and 145↓146.

that recognition of $P_1$ and $P_4$ by Kex2 protease are interdependent.

## DISCUSSION

Here, we have shown that Kex2 protease exhibits specificity at $P_4$ in addition to its previously recognized specificity at $P_2$ and $P_1$. Favorable interactions between the enzyme and the $P_4$ side chain substantially improve values of $k_{cat}/K_M$ and $k_{cat}$ for Lys-Lys substrates. However, this partial compensation for Lys at $P_1$ is insufficient to restore the burst kinetics seen with Lys-Arg substrates; that is, even Lys-Lys substrates with favorable $P_4$ residues form the acylenzyme intermediate more slowly than Lys-Arg substrates. Furthermore, the effects of substitutions at $P_1$ and $P_4$ on catalysis are not independent of one another.

These data clearly demonstrate a preference for either an aliphatic or a basic residue at this position (Table 1). Although a relatively small number of *in vivo* Kex2 cleavage sites are known, 8 out of a list of 11 known or likely Kex2 cleavage sites in *S. cerevisiae* proproteins possess aliphatic $P_4$ residues (Table 3). However, expression of mammalian proproteins in yeast has shown that Kex2 can efficiently cleave substrates *in vivo* with Arg at $P_4$ (*16*). Moreover, one probable physiological cleavage site does contain a basic residue at $P_4$ (*17*). Therefore, *in vivo* observations are in keeping with the data presented here.

There is no crystal structure yet available for a proprotein processing protease, but the structures of a number of bacterial subtilisins have been solved (*18*−*20*), allowing identification of residues which may be involved in $P_4$ recognition (*21*). Furthermore, the role of several of these residues has been examined by mutagenesis in degradative subtilisins (*22*−*26*), and quantitative effects on $P_4$ specificity are indeed seen. An attempt to modulate furin specificity using site-directed mutagenesis of such residues has also been reported (*27*), and qualitative effects on specificity are seen in these mutants as well. It is thus possible to tentatively identify candidate residues for $P_4$ specificity in both Kex2 and metazoan homologs, such as furin, which have specificity for Arg at $P_4$ (Table 4). Kex2 possesses both hydrophobic and acidic residues at these locations, which may explain the observed dual specificity for aliphatic and basic $P_4$ residues. This is also in keeping with the proposal that the $P_4$ pocket has two subsites with potentially different characteristics (*28*).

Table 4: Residues Implicated in $P_4$ Specificity[a]

| enzyme | residue 104 | residue 107 | residue 126 | residue 130 | residue 135 |
|---|---|---|---|---|---|
| Kex2 | Thr | Glu | Trp | Asp | Val |
| furin | Asp | Glu | Trp | Asp | Ala |
| subtilisin | Tyr | Ile | Leu | Ser | Leu |

[a] Residues implicated in $P_4$ specificity for Kex2, subtilisin BPN′, and human furin were compared using a previously published alignment of the catalytic domains (*21*). A more recent alignment (*28*) agrees well with the earlier alignment. The numbering of subtilisin BPN′ is used throughout. The presence of both aliphatic and acidic residues in the Kex2 $P_4$ pocket may explain its recognition of both aliphatic and basic $P_4$ side chains.

No Lys-Lys substrate examined exhibited the initial burst of product formation seen with Lys-Arg substrates. This suggests that Arg at $P_1$ may be necessary for rapid acylation of the enzyme. The initial burst indicates that initial acylation and AMC release are faster than the subsequent steady-state turnovers, implying that a step after acylenzyme formation and AMC release must be rate-limiting in the steady state. The lack of such a pre-steady-state burst with Lys-Lys substrates suggests that acylation may be the rate-determining step in cleavage of such substrates. Favorable $P_4$ contacts do not confer burst kinetics upon Lys-Lys substrates. However, a favorable interaction with the $P_4$ side chain does increase $k_{cat}$ for a Lys-Lys substrate, so it must accelerate the rate-limiting step for such substrates. Thus, the $P_4$ contact may specifically increase the rate of acylation of Kex2 protease by Lys-Lys substrates, though not to the rates achieved with the more physiological substrates having Arg at $P_1$.

The nonadditive effects of substitutions at $P_1$ and $P_4$ indicate that the subsites that bind the $P_1$ and $P_4$ side chains do not do so independently. In crystal structures of degradative subtilisins with bound inhibitors, the $P_4$ and $P_1$ pockets are adjacent to each other. It is thus quite feasible that the nature of the $P_4$ residue affects the alignment of the $P_1$ residue with the $S_1$ subsite and vice versa. This is also in keeping with the observed interdependence of the $S_1$ and $S_4$ binding sites in degradative subtilisins (*12*).

An extensive survey of substrate specificity has been carried out for two degradative subtilisins (*6*). By comparing otherwise identical substrates with Phe or Asp at the $P_4$ position, $P_4$ specificity was estimated to contribute approximately 5 kcal/mol toward catalysis. However, comparison of substrates with Phe or Ala at this position gave a much smaller value of 1.1−1.9 kcal/mol.

The similarity in our $k_{cat}/K_M$ values for Kex2 cleavage of Lys-Lys substrates with Cha, Nle, or Arg at $P_4$ (Table 1) suggests that $P_4$ has been optimized in this context. If $k_{cat}/K_M$ for Ac-Nle-Tyr-Lys-Lys-MCA or Ac-Arg-Tyr-Lys-Lys-MCA is indeed an upper bound for Kex2 cleavage of these Lys-Lys substrates, then the difference between these values and the lowest value measured provides a reasonable estimate of the energetic contribution of the $S_4$−$P_4$ interaction to Kex2 catalysis. This corresponds to a free energy of 2.6−2.8 kcal/mol with a substrate containing $P_4$ Ala as a lower bound or $\geq$4 kcal/mol with a substrate containing $P_4$ Asp. These values are quite comparable to the energetics of discrimination for similar residue substitutions in degradative subtilisins discussed above. However, in the case of Kex2 protease,

the energies of interaction with substrate $P_1$ and $P_2$ side chains are higher still (*8*). $P_4$ specificity would thus play an accessory role in proprotein processing by Kex2, though for enzymes such as furin it is clearly a primary determinant of cleavage site selection (*29*).

Kex2 had long been considered to have specificity for dibasic sites, but little apparent $P_4$ specificity, because substrates containing either aliphatic or basic residues at $P_4$ were cleaved with comparable efficiency (*9*). However, we have shown that the reason for this is that Kex2 exhibits dual recognition for aliphatic and basic $P_4$ residues. Furthermore, this conclusion would suggest that characterization of $P_4$ specificity in related processing proteases should include examination of residues such as Ala or Asp rather than simple comparison of the effects of a large aliphatic residue and a basic residue. The interaction between the $S_1$ and $S_4$ subsites in determining cleavage site selection makes this an important consideration. The molecular basis for this communication between $S_1$ and $S_4$ subsites now apparent in both processing and degradative members of the subtilisin superfamily is yet to be elucidated.

## ACKNOWLEDGMENT

We thank the University of Michigan Protein and Carbohydrate Core Facility for peptide synthesis and mass spectroscopy. We also thank Israel Schechter and the members of the Fuller lab for helpful discussions and comments and gratefully acknowledge Damian Krysan for advice on substrate synthesis.

## REFERENCES

1. Seidah, N. G., Chretien, M., and Day, R. (1994) *Biochimie 76*, 197−209.
2. Rouille, Y., Duguay, S. J., Lund, K., Furuta, M., Gong, Q., Lipkind, G., Oliva, A. A., Jr., Chan, S. J., and Steiner, D. F. (1995) *Front. Neuroendocrinol. 16*, 322−361.
3. Komano, H., and Fuller, R. S. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 10752−10756.
4. Brenner, C., Bevan, A., and Fuller, R. S. (1994) *Methods Enzymol. 244*, 152−167.
5. Watanabe, T., Nakagawa, T., Ikemizu, J., Nagahama, M., Murakami, K., and Nakayama, K. (1992) *J. Biol. Chem. 267*, 8270−8274.
6. Grøn, H., Meldal, M., and Breddam, K. (1992) *Biochemistry 31*, 6011−6018.
7. Angliker, H., Neumann, U., Molloy, S. S., and Thomas, G. (1995) *Anal. Biochem. 224*, 409−412.
8. Rockwell, N. C., Wang, G., Krafft, G., and Fuller, R. S. (1997) *Biochemistry 36*, 1912−1917.
9. Brenner, C., and Fuller, R. S. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 922−926.
10. Park, C.-M., Bruenn, J. A., Ganesa, C., Flurkey, W. F., Bozarth, R. F., and Koltin, Y. (1994) *Mol. Microbiol. 11*, 155−164.
11. Jean, F., Boudreault, A., Basak, A., Seidah, N. G., and Lazure, C. (1995) *J. Biol. Chem. 270*, 19225−19231.
12. Grøn, H., and Breddam, K. (1992) *Biochemistry 31*, 8967−8971.
13. Perona, J. J., and Craik, C. S. (1995) *Protein Sci. 4*, 337−360.
14. Bravo, D. A., Gleason, J. B., Sanchez, R. I., Roth, R. A., and Fuller, R. S. (1994) *J. Biol. Chem. 269*, 25830−25837.
15. Brenner, C., Bevan, A., and Fuller, R. S. (1993) *Curr. Biol. 3*, 498−506.
16. Franzusoff, A., Volpe, A. M., Josse, D., Pichuantes, S., and Wolf, J. R. (1995) *J. Biol. Chem. 270*, 3154−3159.
17. Russo, P., Kalkinnen, N., Sareneva, H., Paakkola, J., and Makarow, M. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 3671−3675.
18. Kraut, J. (1977) *Annu. Rev. Biochem. 46*, 331−358.
19. McPhalen, C. A., and James, M. N. G. (1988) *Biochemistry 27*, 6582−6598.
20. Takeuchi, Y., Satow, Y., Nakamura, K. T., and Mitsui, Y. (1991) *J. Mol. Biol. 221*, 309−325.
21. Siezen, R. J., de Vos, W. M., Leunissen, J. A. M., and Dijkstra, B. W. (1991) *Protein Eng. 4*, 719−737.
22. Bech, L. M., Sørensen, S. B., and Breddam, K. (1992) *Eur. J. Biochem. 209*, 869−874.
23. Sørensen, S. B., Bech, L. M., Meldal, M., and Breddam, K. (1993) *Biochemistry 32*, 8994−8999.
24. Rheinnecker, M., Baker, G., Eder, J., and Fersht, A. R. (1993) *Biochemistry 32*, 1199−1203.
25. Rheinnecker, M., Eder, J., Pandey, P. S., and Fersht, A. R. (1994) *Biochemistry 33*, 221−225.
26. Ballinger, M. D., Tom, J., and Wells, J. A. (1996) *Biochemistry 35*, 13579−13585.
27. Creemers, J. W. M., Siezen, R. J., Roebroek, A. J. M., Ayoubi, T. A. Y., Huylebroeck, D., and Van de Ven, W. J. M. (1993) *J. Biol. Chem. 268*, 21826−21834.
28. Siezen, R. J., and Leunissen, J. A. M. (1997) *Protein Sci. 6*, 501−523.
29. Molloy, S. S., Bresnahan, P. A., Leppla, S. H., Klimpel, K. R., and Thomas, G. (1992) *J. Biol. Chem. 267*, 16396−16402.
30. Bostian, K. A., Elliott, Q., Bussey, H., Burn, V., Smith, A., and Tipper, D. J. (1984) *Cell 36*, 741−751.
31. Zhu, Y.-S., Zhang, X.-Y., Cartwright, C. P., and Tipper, D. J. (1992) *Mol. Microbiol. 6*, 511−520.
32. Kurjan, J., and Herskowitz, I. (1982) *Cell 30*, 933−943.
33. Basco, R. D., Giménez-Gallego, G., and Larriba, G. (1990) *FEBS Lett. 268*, 99−102.
34. Singh, A., Chen, E. Y., Lugovoy, J. M., Chang, C. N., Hitzeman, R. A., and Seeburg, P. H. (1983) *Nucleic Acids Res. 11*, 4049−4063.
35. Schechter, I., and Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157−162.

BI972534R