Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 10384–10389, September 1998
Biochemistry

Quantitative assessment of enzyme specificity *in vivo*: P_2 recognition by Kex2 protease defined in a genetic system

ALISON BEVAN*, CHARLES BRENNER*†, AND ROBERT S. FULLER*‡§

*Department of Biochemistry, Stanford University School of Medicine, Stanford, CA, 94305; and ‡Department of Biological Chemistry, University of Michigan Medical School, Ann Arbor, MI 48109

Communicated by Gregory A. Petsko, Brandeis University, Waltham, MA, June 25, 1998 (received for review May 6, 1997)

ABSTRACT **The specificity of the yeast proprotein-processing Kex2 protease was examined *in vivo* by using a sensitive, quantitative assay. A truncated prepro-α-factor gene encoding an α-factor precursor with a single α-factor repeat was constructed with restriction sites for cassette mutagenesis flanking the single Kex2 cleavage site (-SLDKR ↓ EAEA-). All of the 19 substitutions for the Lys (P_2) residue in the cleavage site were made. The wild-type and mutant precursors were expressed in a yeast strain lacking the chromosomal genes encoding Kex2 and prepro-α-factor. Cleavage of the 20 sites by Kex2, expressed at the wild-type level, was assessed by using a quantitative-mating assay with an effective range greater than six orders of magnitude. All substitutions for Lys at P_2 decreased mating, from 2-fold for Arg to $>10^6$-fold for Trp. Eviction of the Kex2-encoding plasmid indicated that cleavage of mutant sites by other cellular proteases was not a complicating factor. Mating efficiencies of strains expressing the mutant precursors correlated well with the specificity (k_{cat}/K_M) of purified Kex2 for comparable model peptide substrates, validating the *in vivo* approach as a quantitative method. The results support the conclusion that K_M, which is heavily influenced by the nature of the P_2 residue, is a major determinant of cleavage efficiency *in vivo*. P_2 preference followed the rank order: Lys > Arg > Thr > Pro > Glu > Ile > Ser > Ala > Asn > Val > Cys > As*P* > Gln > Gly > His > Met > Leu > Tyr > Phe > Trp.**

Substrate discrimination is central to the physiological function of enzymes (1). Much insight into the energetics of enzyme-substrate recognition has come from studies of serine proteases that mediate key pathways, including protein degradation, coagulation, and proprotein processing (2, 3). The substrate specificity of Kex2 protease ("kexin," EC No. 3.4.21.61) from the yeast *Saccharomyces cerevisiae* is of particular interest because it is the prototype of the eukaryotic subtilisin-related proprotein processing enzymes that cleave sites consisting of pairs or clusters of basic residues (ref. 4 and see refs. 5 and 6 for reviews).

Although three-dimensional structures of Kex2 family enzymes are not available, kinetic studies of cleavage of model substrates (7–11), mutational analysis of substrate cleavage sites (12, 13), and modeling studies (14) suggest that like degradative subtilisins, Kex2 family enzymes interact with substrate P_1, P_2, and P_4 side chains¶ but are vastly more discriminating at each position. Studies of secreted, soluble Kex2 protease using peptidyl methylcoumarinamide (peptidyl-MCA) substrates revealed a highly efficient enzyme, with k_{cat}/K_M values for cleaving the best Lys-Arg substrates in the range of 1–5 $\times$ 10^7 s^{-1} M^{-1} (7, 9, 10). These same studies suggested rules for recognition of P_1 and P_2 residues (7, 9). Kex2 prefers Arg over Lys at P_1 by up to a factor of 100- to 10,000-fold in k_{cat}/K_M (7, 9). A 130-fold decrease in k_{cat} with Lys at P_1 suggested that Arg is required at P_1 for achieving a high catalytic rate (7). In contrast, substitutions for Lys at P_2 resulted in increases in K_M without significantly altering k_{cat} (7). More recent kinetic studies have extended our understanding of recognition of the P_1 and P_2 residues and demonstrated the existence of an S_4 subsite in Kex2 (9, 10).

Until now, comparable detailed information about specificity *in vivo* has been lacking. Here, we introduce a method for analyzing Kex2 specificity *in vivo* that relies on the essential role of the enzyme in yeast mating. Localized to a late Golgi compartment (16, 17), Kex2 cleaves the α-mating pheromone precursor, pro-α-factor, carboxyl to Lys-Arg sites (Fig. 1*A*) (19, 22). Kex2 cleavage is essential for production of mature α-factor and therefore is essential for mating of α haploid cells with haploid cells of the a-mating type. A sterile α strain was constructed that lacks both the chromosomal genes encoding the α-factor precursors (23, 24) and the *KEX2* gene. Mating competence can be restored by introducing plasmids encoding Kex2 and prepro-α-factor. The wild-type *KEX2* gene is reintroduced on a single-copy plasmid under the control of its own promoter to preserve a normal level of expression. A second single-copy plasmid contains a modified *MFα1* gene, *MFα1–100*, that encodes a truncated α-factor precursor (T-pαf) with a single α-factor repeat unit (Fig. 1*B*). Mating of the strain depends on cleavage of T-pαf by Kex2 at a single site that can be modified by cassette mutagenesis (Fig. 1*C*). The effects of alterations of the cleavage site on the efficiency of cleavage by Kex2 protease are determined by using a quantitative-mating assay that has an effective range of greater than six orders of magnitude (25). This permits a sensitive, quantitative assessment of the interaction of Kex2 protease with wild-type and mutant cleavage sites *in vivo*.

MATERIALS AND METHODS

Plasmid Constructions. Routine molecular biology techniques were performed as described (26). The backbone of pAL5-*KEX2* was constructed from the plasmid pRS314 (27). Silent mutations were introduced into the unique *Xba*I and *Hin*dIII sites in the *TRP1* gene by one-step, site-directed mutagenesis (8) by using a 4.3-kbp PCR fragment generated by mutagenic primers 5′-acagaggccgcagaatgtgcCctagattccg-3′ and 5′-gctaacataaaatgtaagctCtcggggctct-3′. A ≈3.9-kbp *Bam*HI fragment of plasmid mYCpN3-*KEX2* (17) containing the *KEX2* promoter, ORF and 3′-untranslated sequences was inserted anticlockwise into a *Bam*HI linker (New England

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: Peptidyl-MCA, peptidyl methylcoumarin amide; T-pαf, truncated pro-α-factor.

†Present address: Kimmel Cancer Institute, Thomas Jefferson University, Philadelphia, PA 19107.
§To whom reprint requests should be addressed. e-mail: bfuller@umich.edu.

¶Substrate residues are designated P_i. . . -P_2-P_1- ↓ -P_1'-P_2'-. . . P_j' with the cleaved bond between residues P_1 and P_1' (15).

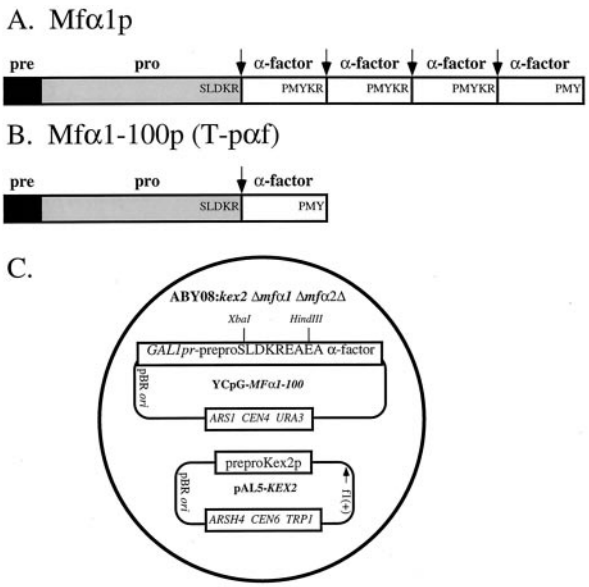

FIG. 1. Schematic structures of α-factor precursors and overview of the genetic system. (*A*) Signal peptidase cleavage of the major form of prepro-α-factor (165 residues), which is encoded by the *MFα1* gene, results in production of 146 residue pro-α-factor (18). Kex2 protease cleaves pro-α-factor between the glycosylated "pro" domain and the first α-factor repeat carboxyl to the sequence Ser-Leu-Asp-Lys-Arg ↓ and at three sites between the α-factor repeats carboxyl to the sequence Pro-Met-Tyr-Lys-Arg ↓ (19). Production of mature α-factor requires exoproteolysis by Kex1 carboxypeptidase to remove C-terminal Lys and Arg residues from the internal repeats (20) and by Ste13 dipeptidyl aminopeptidase to remove N-terminal GluAla and AspAla dipeptides from each repeat (21). (*B*) The product of the *MFα1–100* gene contains a single α-factor repeat. Production of mature α-factor requires cleavage by Kex2 at a single site. (*C*) Schematic of yeast strain ABY08 carrying plasmids YCpG-*MFα1–100* and pAL5-*KEX2*. The *Xba* I and *Hin*dIII sites in YCpG-*MFα1–100* were used for cassette mutagenesis of the P_2 Lys (Lys^{84}). An important feature of this system is that cleavage of both wild-type and mutant sites must occur in the correct "phase," i.e., immediately following the P_1 Arg. This is because production of mature α-factor requires removal of two Glu-Ala dipeptides by the Ste13 peptidase.

Biolabs), replacing the existing polylinker. The *KEX2*-coding sequence contained a unique *Sac*II site introduced at the $Pro^{102}Arg^{103}$ codons of the *KEX2*-coding sequence with the primer 5′-ccatgatttattcccgcgGaacgaccta-3′ by the method of Kunkel (28).

KEX2 knockout plasmid pAL2-*kex2Δ2*::URA3 was constructed by inserting the 3,843-bp *Bgl*II–*Bam*HI fragment of pNKY51, containing the *URA3* gene flanked by direct repeats of *Salmonella typhimurium hisG*-sequences (29), into the *Bgl*II site of pKX11 (R.S.F., unpublished observations), replacing the entire *KEX2*-coding sequence.

Plasmid pBM258 is a derivative of YCp50 (30) that contains a 685-bp fragment with the bidirectional *GAL1,10* promoter (31). A 530-bp *Bam*HI to *Sal*I fragment of pAB111 (gift of A. Brake, Chiron Corp.) containing the *MFα1* ORF (23) was inserted into pBM258 generating plasmid pBM258-*MFα1* in which the *MFα1* structural gene is placed under *GAL1*-promoter control. Cleavage with *Hin*dIII and religation generated plasmid pBM258-*MFα1-ΔHIII* in which the first cleavage site was fused to the fourth mature α-factor repeat. A silent *Xba*I restriction site 5′ of the DNA encoding the Kex2 cleavage site was introduced (as above) with the oligonucleotide 5′-ctctcttttatctagaGaTacccttc-3′ to generate plasmid YCpG-*MFα1–100*.

Yeast Strain Construction. Yeast media and methods were essentially as described (32). SD is synthetic minimal medium containing 2% (wt/vol) glucose, SDC is synthetic complete medium with 2% (wt/vol) glucose, and SGC is synthetic complete medium containing 2% (wt/vol) galactose. Omissions from complete media formulations to create "drop-out" media are indicated as "-Ura," etc. Strains are described in Table 1. Strains SEY6210 and SEY6211 (33) were mated and sporulated and tetrads dissected to generate ABY02. Strain JE104 containing plasmid pBM258-*MFα1* was crossed with ABY02 and the resulting diploid was cured of pBM258-*MFα1* and transformed with a 5,124-bp *Bam*HI fragment of pAL2-*kex2Δ2*::*URA3* to delete the chromosomal *KEX2* gene. The diploid was sporulated, and sterile *kex2Δ2*::*URA3 his3 HIS4* segregants were identified by complementation and cosegregation of Ura^+ and cold sensitivity (34, 35). A *MATα* candidate was plated on 5-fluoroorotic acid medium (36) to select to generate Ura^- strain ABY06. Strain ABY08 was derived from ABY06 by selection on α-aminoadipic acid (37) to isolate a Lys^- auxotroph shown subsequently to be *lys2 LYS5* by complementation.

Mutagenesis of MFα1–100 and Mating Assays. Plasmid YCpG-*MFα1–100* was cleaved with *Xba*I and *Hin*dIII, and the digest was ligated to a 100-fold excess of 20 mer oligonucleotides 5′-ctagatNNNagagaggctga-3′and 5′-agcttcagcctctct-NNNat-3′. Mutations were confirmed by sequencing alkali-denatured, double-stranded plasmid DNA. *MFα1–100* alleles encoding substitutions for the P_2 Lys residue at position 84 were designated as shown in Table 2. To perform quantitative-mating assays, mutant plasmids were introduced into ABY08 with and without pAL5-*KEX2*. In the qualitative assay, patches of ABY08 transformants were grown on YPAGal plates for 12 hr at 30°C and then replica-plated onto lawns of DC14 cells on SD plates. In the quantitative assay (25), a fixed number of DC14 cells ($1–1.5 \times 10^7$) were mixed with exponentially growing cultures of α-cells ($2–5 \times 10^6$), and each mixture was collected on a nitrocellulose filter (HATF, Millipore). Filters were incubated on YPAGal plates at 30°C for 12 hr to allow mating. Cells were resuspended in distilled water and dilutions of the mating reaction were plated on SD plates to select prototrophic diploids. Mating frequency was expressed as the number of prototrophic colonies on minimal (SD) plates per input α cell. Typically, the mating frequency of ABY08 [pAL5-*KEX2*, YCpG-*MFα1–100*] was ≈0.2. Input α cells were measured by titering serial dilutions on permissive medium.

Table 1. Yeast strains used in this study

Strain	Genotype	Source
SEY6210	MATα *leu2-3,112 ura3-52 his3-Δ200 trpl-Δ900 lys2-810 suc2-Δ9* Gal^+	S. D. Emr
SEY6211	MAT**a** *leu2-3,112 ura3-52 his3-Δ200 trp1-Δ900 ade2-101 suc2-Δ9* Gal^+	S. D. Emr
ABY02	MAT**a** *leu2-3,112 ura3-52 his3-Δ200 trp1-Δ900 ade2-101 lys2-810 suc2-Δ9* Gal^+	This work
JE104	*MATα mfα1Δ::LEU2 mfα2Δ::LYS2 ura3-52 his4 lys2 suc2-Δ9* Gal^+	J. Thorner
DC14	MAT**a** *his1*	J. Thorner
DC17	MATα *his1*	J. Thorner
CRY2	MATα *his3-11,15 leu2-3,112 ura3-1 trp1-1 ade2-1 can1-100*	Fuller lab
ABY06	MATα *mfa1Δ::LEU2 mfa2Δ::LYS2 kex2Δ::hisg ura3-52 his3-Δ200 trp1-Δ900 ade2-101 suc2-Δ9* Gal^+	This work
ABY08	MATα *mfα1Δ::LEU2 mfα2Δ::lys2 kex2Δ2::hisg ura3-52 his3-Δ200 trp1-Δ900 ade2-101 suc2-Δ9* Gal^+	This work

S. D. Emr, University of California, San Diego; J. Thorner, University of California, Berkeley.

Table 2. Effects of P_2 substitutions on processing of T-pαf *in vivo* measured by quantitative mating*

Genotype†			Plasmids‡					Relative mating	
KEX2	*MFα1*	*MFα2*	*KEX2*	MFα1	(-allele)	P2	P1 ↓	efficiency	± % Error§
+	+	+	—	—		LYS	-Arg	6.4	4
Δ	Δ	Δ	+	+	-wt	LYS	-Arg	4.5	10
Δ	Δ	Δ	+	+	-100	LYS	-Arg	1.0	13
Δ	Δ	Δ	+	+	-114	ARG	-Arg	0.51	11
Δ	Δ	Δ	+	+	-116	THR	-Arg	0.13	26
Δ	Δ	Δ	+	+	-112	PRO	-Arg	0.095	8
Δ	Δ	Δ	+	+	-104	GLU	-Arg	0.088	10
Δ	Δ	Δ	+	+	-108	ILE	-Arg	0.054	7
Δ	Δ	Δ	+	+	-115	SER	-Arg	0.043	12
Δ	Δ	Δ	+	+	-101	ALA	-Arg	0.032	16
Δ	Δ	Δ	+	+	-111	ASN	-Arg	0.029	2
Δ	Δ	Δ	+	+	-117	VAL	-Arg	0.014	15
Δ	Δ	Δ	+	+	-102	CYS	-Arg	8.5×10^{-3}	7
Δ	Δ	Δ	+	+	-103	ASP	-Arg	2.7×10^{-3}	5
Δ	Δ	Δ	+	+	-113	GLN	-Arg	2.4×10^{-3}	14
Δ	Δ	Δ	+	+	-106	GLY	-Arg	2.3×10^{-3}	4
Δ	Δ	Δ	+	+	-107	HIS	-Arg	7.2×10^{-4}	1
Δ	Δ	Δ	+	+	-110	MET	-Arg	6.8×10^{-4}	14
Δ	Δ	Δ	+	+	-109	LEU	-Arg	4.3×10^{-4}	17
Δ	Δ	Δ	+	+	-119	TYR	-Arg	1.7×10^{-4}	4
Δ	Δ	Δ	+	+	-105	PHE	-Arg	6.3×10^{-6}	7
Δ	Δ	Δ	+	+	-118	TRP	-Arg	$<1.0 \times 10^{-6}$	NA
Δ	Δ	Δ	+	+	-120	amb	-Arg	$<1.0 \times 10^{-6}$	NA
Δ	Δ	Δ	+	−		- - - - - - - - - -		$<1.0 \times 10^{-6}$	NA
Δ	Δ	Δ	−	+	-100	LYS	-Arg	$<1.0 \times 10^{-6}$	NA

*Assays were performed three to eight times for each amino acid substitution at P_2 but only twice for nonmating controls (e.g., amber codon at P_2 or complete absence of KEX2 or MFα1 gene).
†"+ + +" strain was CRY2; "Δ Δ Δ" was ABY08.
‡The KEX2 plasmid was pAL5-*KEX2*, the *MFα1*-WT plasmid was pBM258-*MFα1*, and the rest of the plasmids were YCpG-*MFα1-100* or derivatives with the indicated P_2 substitution.
§"% Error" is SEM expressed as a percentage.

Mating efficiency was the mating frequency of experimental samples divided by that for control strain ABY08 [pAL5-*KEX2*, YCpG-*MFα1–100*]. The mating efficiency of strain ABY08 [pAL5-*KEX2*, YCpG-*MFα1–100*] is defined as 1.0. Mating with Glu at P_2 was unexpectedly high. To establish definitively that this was due to cleavage after Glu-Arg, the plasmid encoding the putative *MFα1–104* was recovered, sequenced, and reintroduced into ABY08. The mating behavior of this strain was identical to that previously observed.

Detection of Pro-α-Factor by Immunoblotting. Anti-(pro-α-factor) antibodies were obtained by immunizing two virgin New Zealand white rabbits with a purified, *E. coli*-expressed β-galactosidase/pro-α-factor fusion protein (38) followed by four additional boosts with purified *E. coli*-expressed pro-α-factor (A.B. and R.S.F., unpublished experiments). Antibodies were affinity purified (16). Yeast transformants expressing T-pαf were grown in SGC-Ura medium to an OD of 2.0 at 600 nm, cells were removed by centrifugation, BSA (Pierce) was added to culture medium (3 μg/ml), and protein was precipitated by adding 0.1 volume of 0.15% (wt/vol) sodium deoxycholate and 90% (wt/vol) trichloroacetic acid at 0°C. After centrifugation, pellets were washed three times with 1 ml of ice-cold acetone and resuspended in 20 μl of 50 mM sodium phosphate, pH 5.5, 0.02% (wt/vol) sodium dodecylsulfate, 0.5 mM phenylmethylsulfonyl fluoride, and 10 mM NaN_3 and heated to 97°C for 3 min. Protein samples were deglycosylated by digestion with 0.5 munit of Endoglycosidase H (Boeringer Mannheim) for 16 hr at 37°C. Samples were subjected to SDS/PAGE (15% acrylamide), electroblotted to nitrocellulose (Millipore), and probed with a 1:5000 dilution of anti-(pro-α-factor) antibody. Immunoblots were probed with goat anti-rabbit horseradish peroxidase IgG and developed with enhanced chemiluminescence (Amersham) according to the manufacturer's recommendations.

RESULTS

Sensitivity and Range of the Assay System. Several controls were used to evaluate the genetic system. *In vivo*, wild-type pro-α-factor is processed to completion by Kex2, resulting in high levels of α-factor secretion. To determine whether the quantitative-mating assay was sensitive to slight reductions in α-factor production from the wild-type levels, the mating ability of strains expressing plasmid-encoded α-factor precursors with one or four α-factor repeats was examined. The mating efficiency of strain ABY08 containing the wild-type *KEX2* gene on plasmid pAL5-*KEX2* and expressing the wild-type *MFα1* gene under *GAL1* promoter control was quite similar to that of the wild-type *MATα* strain, CRY2 (Table 2). In contrast, the mating efficiency of ABY08 containing pAL5-*KEX2* and expressing the truncated pro-α-factor (T-pαf) encoded by the *Mfα1–100* gene under *GAL1*-promoter control was 4.5-fold lower (Table 2). Thus the mating assay responded to differences in levels of α-factor production even under conditions of efficient processing. Mating was undetectable in transformants lacking either the *KEX2* or prepro-α-factor gene. Mating of ABY08 with *MFα1–100* and pAL5-*KEX2* was at least 10^6-fold greater (Table 2). Thus, the quantitative-mating assay provided a sensitivity range of six orders of magnitude with which to assess the effects of substitutions at the P_2 position in T-pαf on the efficiency of cleavage by Kex2 protease *in vivo*. The precision of the assay was such that 2-fold differences in mating efficiency were reproducible.

Effects of Substitutions at P_2. As shown in Table 2, all substitutions for Lys at P_2 reduced mating efficiency, with the effects spread over the entire sensitivity range of the assay. As expected, the best cleavage sites contained Lys or Arg at P_2, although substitution of Arg at P_2 led reproducibly to a 2-fold lower mating efficiency. Analysis of the data identified several

general features as well as certain side-chain specific effects. With three clear exceptions, there was an approximate, inverse correlation between the total solvent accessible surface area of the P_2 side chain and the log of the mating efficiency (Fig. 2*A*). The exceptions were Lys, Arg, and Gly. In the case of Lys and Arg, efficient cleavage is likely due to favorable electrostatic interactions with the enzyme contacts. Studies of purified Kex2 appear to rule out contributions to substrate binding or catalysis from interactions between the methylene groups of P_2 side chains and the enzyme (9) Moreover, the Met side chain is nearly isosteric with the aliphatic portion of the Lys side chain with similar packing volume and hydrophobic surface area. Substitution of Met for Lys at P_2 resulted in a decrease of ≈1,500-fold in mating efficiency, whereas substitution of Ala for Lys resulted in only a ≈30-fold decrease. In the case of Gly at P_2, the most likely explanation for poor mating is poor cleavage due to high conformational flexibility. Substitution of large, aromatic side chains at the P_2 position had the most drastic effects. Introduction of Phe at P_2 resulted in barely detectable mating, with the efficiency reduced 160,000-fold. Introduction of Trp at P_2 resulted in apparent sterility. This was equivalent to the reduction of mating efficiency seen with strain ABY08 lacking either *KEX2* or an *MFα* gene, a level at least 10^6-fold below wild-type mating, the same as the level of mating observed with insertion of an amber codon at the P_2 codon position.

Correlation between mating efficiency and solvent-accessible nonpolar surface area of P_2 side chains also was examined (Fig. 2*B*). With the notable exception of Ile (see below), increasing solvent-accessible nonpolar surface area of hydrophobic residues (Pro, Val, Met, Leu, Tyr, Phe, and Trp) exhibited a better inverse correlation with mating efficiency than did total solvent accessible area, consistent with the general conclusion that the P_2-binding site is relatively hydrophilic. However, the mating efficiency in the case of residues with small and/or hydrophilic side chains (Asp, Gln, Asn, Ser, Glu, Ala, Thr) exhibited no obvious correlation in this plot.

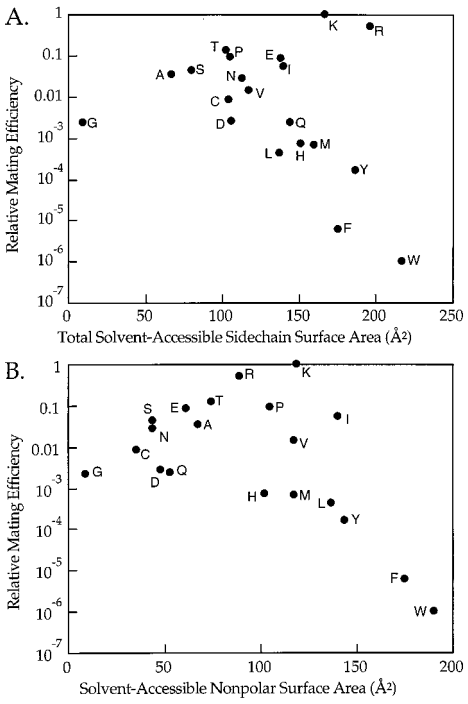

FIG. 2. (*A*) Plot of relative mating efficiency vs. surface area of the P_2 side chain. Total solvent-accessible side chain surface areas for all residues except Gly were as described by Creighton (39). For Gly, total solvent-accessible side chain surface was estimated as half the surface area of a hydrogen atom. (*B*) Plot of relative mating efficiency vs. accessible nonpolar surface area. Total solvent-accessible nonpolar side chain surface areas for all residues except Gly were as described by Creighton (39). For Gly, total solvent accessible nonpolar side chain surface was estimated as one-half the surface area of a hydrogen atom.

The effects of certain substitutions were unexpected and will require detailed information about the three-dimensional structure of the active site to be explained. Introduction of Glu at P_2 caused only an 11-fold reduction in mating, despite the introduction of negative charge. Introduction of Asp resulted in a much greater decrease in mating (370-fold). Interestingly, mating with Gln was reduced ≈420-fold, whereas mating with Asn was only reduced 34-fold. The P_2-binding site appears to favor β-branched residues. Ile at P_2 resulted in a 19-fold reduction in mating whereas with Leu, mating was reduced 2,300-fold. Thr was the third best residue at P_2, reducing mating only 7-fold relative to Lys whereas Ser reduced mating efficiency 23-fold.

Observed Cleavage Is Due to Kex2 Protease. One potentially complicating factor in the assay was the possibility that some of the P_2 mutant precursors were being cleaved by a protease other than Kex2. To examine this possibility, plasmid pAL5-*KEX2* was evicted from each of the 20 strains expressing a form of T-pαf with either the wild-type or one of the mutant cleavage sites. In each case, loss of the *KEX2* plasmid resulted in negligible mating, indicating that mating depended on cleavage by Kex2 protease (data not shown).

Mutations Do Not Affect T-pαf Stability or Secretion. Reduced mating associated with any or all of the mutant precursors might have been caused by instability or improper localization rather than by reduced efficiency of cleavage by Kex2 protease. Culture media of ABY08 transformants expressing *MFα1–100* or one of the P_2 substitution alleles, but lacking the *KEX2* gene, were concentrated and digested with Endo-glucosaminidase H to remove N-linked oligosaccharides. Affinity-purified antibody raised against pro-α-factor was used to characterize the deglycosylated, secreted α-factor precursors in these media fractions by immunoblotting (Fig. 3). Polypeptides similar in mobility and abundance to the *MFα1–100* protein were detected in the culture media of all but three transformants. As expected, the strain containing the vector, pBM258, without an insert produced no cross-reacting material. Likewise, the medium from ABY08 expression *MFα1–120*, which contained a nonsense codon in the P_2 position, showed no detectable antigen. The strain expressing *MFα1–*

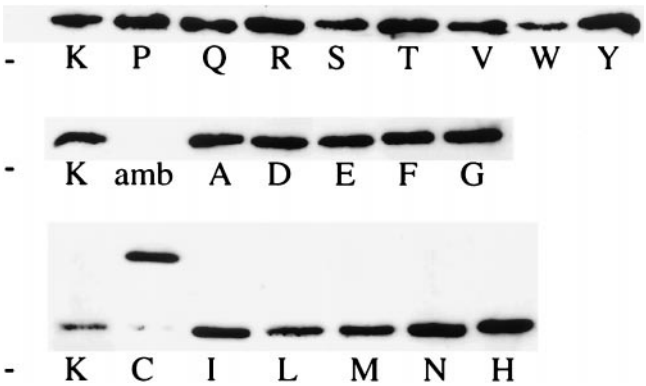

FIG. 3. Expression of mutant α-factor precursors. Yeast strain ABY08 was transformed with pBM258 (control plasmid or "−"), YCpG-*MFα1–100* (K), and with plasmids carrying P_2-substituted mutant alleles of *MFα1–100* indicated using the single letter amino acid code. The amber termination codon in place of Lys84 is labeled "amb." Enzymatically deglycosylated proteins from culture media were separated by SDS/PAGE, blotted, and probed with anti-(pro-α-factor) antiserum as described in *Materials and Methods*. Mfα1–120p with Cys at P_2 formed an intermolecular disulfide and was only partially digested (shown) with Endoglycosidase H unless reduced with β-mercaptoethanol before deglycosylation (data not shown).

102, which encodes a precursor with a substitution of Cys for the P_2 Lys^{84}, produced a precursor-related antigen that migrated much more slowly than the others. This precursor appeared to dimerize through an interchain disulfide because (*i*) Cys^{84} was the only Cys residue in the polypeptide and (*ii*) this precursor could only be deglycosylated if first treated with β-mercaptoethanol. Thus with the exception of Mfα1–102p, all mutant precursors containing an amino acid substitution at the P_2 position were produced in normal amounts and transported normally through the secretory pathway.

DISCUSSION

Most studies of the specificity of Kex2 family proteases using natural proprotein substrates, although informative, have been qualitative in nature. Studies with purified furin established requirements for P_1 and P_4 Arg residues in cleavage of the anthrax protective antigen and suggested that the nature of the P_2 and P_3 residues was less important (13). One *in vivo* study used a cotransfection assay to assess the ability of wild-type furin and various furin mutants to process the wild-type cleavage site (Arg-Ser-Lys-Arg $\downarrow$) and three mutant cleavage sites containing single substitutions at P_1 (Gly), P_2 (Ala), or P_4 (Ala) in pro Von Willebrand factor (12). Bravo and coworkers (40) used immunoblotting to follow cleavage of wild-type and mutant forms of the human insulin proreceptor by purified secreted, soluble furin.

We have detailed a genetic system in which enzyme-substrate interactions can be quantified under physiological conditions. These studies provide a comprehensive profile for recognition of substrate P_2 residues by Kex2 *in vivo*, raising the question of how well these results compare with the kinetic analysis of cleavage of synthetic peptide substrates by purified Kex2. Overall, the relative k_{cat}/K_M values for cleavage of small peptidyl-MCA substrates (Fig. 4*A*) and internally quenched fluorogenic peptide substrates (Fig. 4*B*) by purified Kex2 protease correlate well with the relative mating efficiencies observed with comparable T-pαf variants. The general agreement between these sets of data leads to several conclusions. First, studies with model peptides have shown that with Arg at P_1, most substitutions for Lys at P_2 affected K_M without significantly altering k_{cat}. Thus, the correlation between k_{cat}/K_M values and mating efficiencies indicates that K_M is a major determinant of specificity *in vivo.* Second, pro-α-factor cleavage *in vivo* must be regulated by K_M even for very good, low K_M substrates. Even the relatively conservative substitution of Arg for Lys at P_2 in T-pαf resulted in a measurable and reproducible change in the mating assay. Studies with model peptides indicated that substitution of Arg for Lys at P_2 results in only a 4-fold increase in K_M, along with a small increase in k_{cat} (7). Third, only three of the six most efficiently cleaved sites reflect sequences found in natural substrates of Kex2 protease. Lys-Arg is present in the four cleavage sites of α-factor precursor (23) and is the site of intramolecular cleavage at of the N-terminal "pro" domain of Kex2 (7, 43). Kex2 protease also cleaves at a Lys-Arg, an Arg-Arg, and probably a Pro-Arg site in the M1 killer protoxin *in vivo* (44, 45). Hsp150p (46) and the *EXG1*-endcoded exoglucanase (47) also are cleaved at Lys-Arg sites. It is interesting, therefore, that the current study reveals Thr-Arg to be almost as good as Arg-Arg and slightly better than Pro-Arg as a cleavage site for Kex2 protease, suggesting Thr-Arg as a candidate physiological cleavage site for the enzyme although no such sites are known in natural substrates. This prediction is supported by analysis of furin specificity by the substrate phage display method, which suggested that in addition to Lys and Arg, Thr and Pro are acceptable P_2 residues for furin (48). Finally, the correlations observed indicate that cleavage of the T-pαf P_2 sequence variants *in vivo* can be accounted for by recognition of substrate sequences by Kex2 in small, unstructured peptides, suggesting that higher order structure may be relatively unimportant in utilization of pro-α-factor as a substrate by the enzyme.

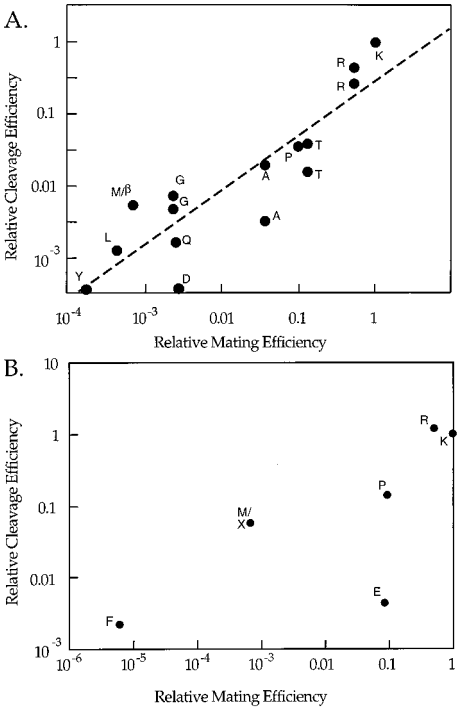

FIG. 4. *In vitro* vs. *in vivo* cleavage efficiency. (*A*) Peptidyl-MCA substrates. Relative efficiency of the peptidyl-MCA substrates is correlated with the *in vivo*-mating efficiency. Relative cleavage efficiency is defined as the ratio of k_{cat}/K_M for a peptidyl-MCA substrate with the indicated residue at P_2 to the k_{cat}/K_M for a peptidyl-MCA substrate with Lys at P_2. Data for k_{cat}/K_M of purified secreted, soluble Kex2 protease with peptidyl-MCA substrates are from Brenner and Fuller (7), Rockwell, *et al.* (9), Rockwell and Fuller (10), and N. Rockwell and R.S.F. (unpublished results). P_2 residues of substrates compared are indicated in the single letter code (β represents norleucine). A substrate with nor-leucine at P_2 was compared with mating efficiency with Met at P_2 *in vivo* (YCpG-MFα1–110). In certain cases (Arg, Thr, Ala, and Gly), two synthetic substrates were available for comparison. The dashed line represents a line fit for direct proportionality between log (k_{cat}/K_M) and the log of the mating efficiency (correlation coefficient, r = 0.886; slope = 0.75). (*B*) Internally quenched fluorogenic peptide substrates. A correlation plot of relative efficiency of cleavage of the internally quenched fluorogenic peptide substrates (k_{cat}/K_M) v. relative mating efficiency for corresponding P_2 substitutions. Relative cleavage efficiency is defined as the ratio of k_{cat}/K_M for an internally quenched fluorogenic peptide substrate with the indicated residue at P_2 to the k_{cat}/K_M for an internally quenched fluorogenic peptide substrate with Lys at P_2. Data for k_{cat}/K_M of purified secreted, soluble Kex2 protease with internally quenched fluorogenic peptide substrates are from ref. 9. P_2 residues of substrates compared are indicated in the single letter code. The internally quenched fluorogenic peptides were of the sequence Arg-GlE-norLeu-Tyr-Xaa-Arg-$\downarrow$-Glu-Ala-Glu-Ala-LyD-Arg, in which "GlE" represents glutamyl-EDANS, [EDANS, 5-(2-aminoethyl-amino)-naphthalene-1-sulfonic acid], "Xaa" represents the particular P_2 substitution, and "LyD" represents lysyl-DABSYL (DABCYL, 4-(4-dimethylaminophenyl)-azobenzoyl) (41, 42). This sequence was based on the Kex2 cleavage site between the α-factor repeats in pro-α-factor. A substrate with nor-leucine (X) at P_2 was compared with mating efficiency with Met at P_2 *in vivo* (YCpG-*MF*α1–110).

The most deleterious substitutions at P_2 involved residues with bulky, nonpolar side chains, suggesting that steric exclusion poses the largest barrier to efficient cleavage, revealed most dramatically in the case of Trp at P_2. In the case of His, although the imidazole ring is polar, its bulk presumably prevents a good fit in the P_2 pocket. Comparison of Tyr and Phe reveal that although these residues have similar bulky

aromatic side chains, cells expressing T-pαf with Tyr at P_2 mated 27-fold better than cells expressing T-pαf with Phe at P_2. Presumably, this results from the greater polarity of the phenolic side chain of Tyr. The poor mating with His or Tyr at P_2 indicates that neither polarity nor the possibility of positive charge can fully compensate for the unfavorable steric interactions of the large cyclic side chains.

Two exceptions to the general agreement between the *in vivo* and *in vitro* data emerged with acidic residues at P_2. As shown in Fig. 4, cells expressing T-pαf with Glu or Asp at P_2 mated significantly better than would be expected from the relative k_{cat}/K_M values. A possible explanation for this observation may lie in differences in sequence context. In T-pαf, the sequence spanning the cleavage site was Ser-Leu-Asp-Lys-Arg-↓-Glu-Ala-Glu-Ala. Substituting Asp or Glu at P_2 therefore introduces an acidic residue adjacent to Asp at P_3. In contrast, the peptidyl-MCA substrate with Asp at P_2 had Asn at P_3 (N. Rockwell and R.S.F., unpublished experiments). Likewise, the internally quenched substrate with Glu at P_2 was based on the sequence Met-Tyr-Lys-Arg-↓-Glu-Ala-Glu-Ala, with Tyr at P_3 (8). Repulsion between adjacent acidic side chains at P_2 and P_3 in T-pαf may result in a substrate conformation that promotes cleavage.

This work demonstrates the utility of the mating assay as a quantitative method for analyzing substrate specificity of Kex2 protease *in vivo* and establishes methods that form part of a genetic and biochemical to approach to understanding and engineering protease specificity (49). The combined information obtained from the physiological and biochemical studies provides a comprehensive selectivity profile of the wild-type protease that defines the range and nature of binding contacts used by the enzyme. This genetic assay also can be used in screens to identify mutant enzymes having altered substrate specificity (A.B. and R.S.F., unpublished experiments).

The authors would like to thank J. Rothblatt for providing the pro-α-factor fusion plasmid and N. Rockwell for kinetic data on purified Kex2 protease and assistance in analyzing correlations with quantitative mating data. This work was supported by National Institutes of Health training Grant 5T32 GM07365 (A.B.), National Cancer Institute training Grant 5T32CA09302 (C.B.), and National Institutes of Health Grant GM39697 (to R.S.F.).

1. Jencks, W. P. (1975) *Adv. Enzymol. Relat. Areas Mol. Biol.* **43,** 219–410.
2. Perona, J. J. & Craik, C. S. (1995) *Protein Sci.* **4,** 337–360.
3. Neurath, H. (1991) *Enzyme* **45,** 239–243.
4. Fuller, R. S., Brake, A. J. & Thorner, J. (1989) *Science* **246,** 482–486.
5. Seidah, N. G., Chretien, M. & Day, R. (1994) *Biochimie (Paris)* **76,** 197–209.
6. Rouille, Y., Duguay, S. J., Lund, K., Furuta, M., Gong, Q., Lipkind, G., Oliva, A. J., Chan, S. J. & Steiner, D. F. (1995) *Front. Neuroendocrinol.* **16,** 322–361.
7. Brenner, C. & Fuller, R. S. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 922–926.
8. Brenner, C., Bevan, A. & Fuller, R. S. (1993) *Curr. Biol.* **3,** 498–506.
9. Rockwell, N. C., Wang, G. T., Krafft, G. A. & Fuller, R. S. (1997) *Biochemistry* **36,** 1912–1917.
10. Rockwell, N. C. & Fuller, R. S. (1998) *Biochemistry* **37,** 3386–3391.
11. Jean, F., Boudreault, A., Basak, A., Seidah, N. G. & Lazure, C. (1995) *J. Biol. Chem.* **270,** 19225–19231.
12. Creemers, J. W., Siezen, R. J., Roebroek, A. J., Ayoubi, T. A., Huylebroeck, D., Van de Ven, W. J. (1993) *J. Biol. Chem.* **268,** 21826–21834.
13. Molloy, S. S., Bresnahan, P. A., Leppla, S. H., Klimpel, K. R. & Thomas, G. (1992) *J. Biol. Chem.* **267,** 16396–16402.
14. Siezen, R. J., Creemers, J. W. & Van de Ven, W. J. (1994) *Eur. J. Biochem.* **222,** 255–266.
15. Schechter, I. & Berger, A. (1967) *Biochem. Biophys. Res. Commun.* **27,** 157–162.
16. Redding, K., Holcomb, C. & Fuller, R. S. (1991) *J. Cell Biol.* **113,** 527–538.
17. Wilcox, C. A., Redding, K., Wright, R. & Fuller, R. S. (1992) *Mol. Biol. Cell* **3,** 1353–1371.
18. Waters, M. G., Evans, E. A. & Blobel, G. (1988) *J. Biol. Chem.* **263,** 6209–6214.
19. Julius, D., Brake, A., Blair, L., Kunisawa, R. & Thorner, J. (1984) *Cell* **37,** 1075–1089.
20. Dmochowska, A., Dignard, D., Henning, D., Thomas, D. Y. & Bussey, H. (1987) *Cell* **50,** 573–584.
21. Julius, D., Blair, L., Brake, A., Sprague, G. & Thorner, J. (1983) *Cell* **32,** 839–852.
22. Fuller, R. S., Sterne, R. E. & Thorner, J. (1988) *Annu. Rev. Physiol.* **50,** 345–362.
23. Kurjan, J. & Herskowitz, I. (1982) *Cell* **30,** 933–943.
24. Singh, A., Chen, E. Y., Lugovoy, J. M., Chang, C. N., Hitzeman, R. A. & Seeburg, P. H. (1983) *Nucleic Acids Res.* **11,** 4049–4063.
25. Hartwell, L. H. (1980) *J. Cell Biol.* **85,** 811–822.
26. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
27. Sikorski, R. S. & Hieter, P. (1989) *Genetics* **122,** 19–27.
28. Kunkel, T. A., Roberts, J. D. & Zakour, R. A. (1987) *Methods Enzymol.* **154,** 367–382.
29. Alani, E., Cao, L. & Kleckner, N. (1987) *Genetics* **116,** 541–545.
30. Rose, M. D., Novick, P., Thomas, J. H., Botstein, D. & Fink, G. R. (1987) *Gene* **60,** 237–243.
31. Johnston, M. & Davis, R. W. (1984) *Mol. Cell. Biol.* **4,** 1440–1448.
32. Rose, M. D., Winston, F. & Hieter, P. (1990) *Methods in Yeast Genetics, a Laboratory Course Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York).
33. Robinson, J. S., Klionsky, D. J., Banta, L. M. & Emr, S. D. (1988) *Mol. Cell. Biol.* **8,** 4936–4948.
34. Martin, C. & Young, R. A. (1989) *Mol. Cell. Biol.* **9,** 2341–2349.
35. Komano, H. & Fuller, R. S. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 10752–10756.
36. Boeke, J. D., LaCroute, F. & Fink, G. R. (1984) *Mol. Gen. Genet.* **197,** 345–346.
37. Chattoo, B. B., Sherman, F., Fjellstedt, T. A., Mehnert, D. & Ogur, M. (1979) *Genetics* **93,** 51–65.
38. Rothblatt, J. A. & Meyer, D. I. (1986) *Cell* **44,** 619–628.
39. Creighton, T. E. (1993) *Proteins, Structures, and Molecular Properties* (Freeman, New York).
40. Bravo, D. A., Gleason, J. B., Sanchez, R. I., Roth, R. A. & Fuller, R. S. (1994) *J. Biol. Chem.* **269,** 25830–25837.
41. Matayoshi, E. D., Wang, G. T., Krafft, G. A. & Erickson, J. (1990) *Science* **247,** 954–958.
42. Wang, G. T. & Krafft, G. A. (1992) *Bioorg. Med. Chem. Lett.* **2,** 1665–1668.
43. Gluschankof, P. & Fuller, R. S. (1994) *EMBO J.* **13,** 2280–2288.
44. Zhu, Y. S., Zhang, X. Y., Cartwright, C. P. & Tipper, D. J. (1992) *Mol. Microbiol.* **6,** 511–520.
45. Zhu, H., Bussey, H., Thomas, D. Y., Gagnon, J. & Bell, A. W. (1987) *J. Biol. Chem.* **262,** 10728–10732.
46. Russo, P., Kalkkinen, N., Sareneva, H., Paakkola, J. & Makarow, M. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 3671–3675.
47. Basco, R. D., Gimenez, G. G. & Larriba, G. (1990) *FEBS Lett.* **268,** 99–102.
48. Matthews, D. J., Goodman, L. J., Gorman, C. M. & Wells, J. A. (1994) *Protein Sci.* **3,** 1197–1205.
49. Brenner, C., Bevan, A. & Fuller, R. S. (1994) *Methods Enzymol.* **244,** 152–167.