# Increase of the $P_1$ Lys/Leu Substrate Preference of Carboxypeptidase Y by Rational Design Based on Known Primary and Tertiary Structures of Serine Carboxypeptidases†

Kjeld Olesen and Klaus Breddam*

*Department of Chemistry, Carlsberg Laboratory, Gamle Carlsbergvej 10, DK-2500 Copenhagen Valby, Denmark*

*Received June 12, 1995; Revised Manuscript Received September 26, 1995*⊗

ABSTRACT: The $P_1$ substrate preference of serine carboxypeptidases, as expressed by the Lys/Leu ratio, differs by up to $10^5$-fold. Predictions of the major determinants of this preference are made by correlating primary and tertiary structures to substrate preferences. In carboxypeptidase Y from yeast it is predicted that Trp312 constitutes such a determinant, reducing the $P_1$ Lys/Leu substrate preference of this enzyme. The predictions are tested by the construction and kinetic characterization of ten mutant enzymes of carboxypeptidase Y. All of these enzymes exhibit changes in their $P_1$ substrate preference. Generally, small decreases in activity ($k_{cat}/K_m$) are observed with substrates containing uncharged $P_1$ side chains. With substrates containing acidic $P_1$ side chains, i.e., FA-Glu-Ala-OH, the activity generally increases slightly, 7-fold in the case of W312K. The most dramatic effects of the Trp312 substitutions are observed with substrates containing basic $P_1$ side chains, i.e., $k_{cat}/K_m$ for the hydrolysis of FA-Lys-Ala-OH with W312E has increased 1150-fold, exclusively as a result of increased $k_{cat}$ values. Similar results have previously been obtained by mutational substitution at position 178 of carboxypeptidase Y. The construction and kinetic characterization of position 178+312 double mutants demonstrate that the kinetic effects of substitutions at these two positions are not additive. The $P_1$ Lys/Leu substrate preference of one double mutant, L178D+W312D, has changed 380 000-fold as compared to the wild type enzyme, and the overall $P_1$ substrate preference of this enzyme closely resembles that of carboxypeptidase WII from wheat.

The serine carboxypeptidases is a group of proteolytic enzymes that catalyze the sequential removal of amino acids from the C-terminus of peptide substrates. The primary structure (amino acid sequence) of about 30 enzymes has been determined, but less than one-third of these enzymes have been purified and characterized kinetically. Although all the enzymes characterized so far exhibit rather broad substrate preferences, both with respect to the C-terminal ($P_1'$) and penultimate ($P_1$) amino acid residue of the substrate, they can nevertheless be divided into two distinct groups on the basis of their $P_1$ substrate preferences. One group exhibits a very high preference (about $10^3-10^4$-fold) for substrates containing hydrophobic $P_1$ side chains relative to basic $P_1$ side chains, while the other group of enzymes exhibits a weak preference (about 1−50-fold) for substrates containing basic $P_1$ side chains. Hereafter, these two groups of enzymes will be referred to as having low and high $P_1$ Lys/Leu substrate preferences, respectively.

The tertiary structure has been determined for two serine carboxypeptidases, carboxypeptidase Y (CPD-Y)[1] from yeast, *Saccharomyces cereviciae* (Endrizzi et al., 1994) and carboxypeptidase II (CPD-WII) from wheat, *Triticum aestivum* (Liao & Remington, 1990). These two enzymes are examples of the two groups of serine carboxypeptidases, as they differ more than $10^5$-fold with respect to their $P_1$ Lys/Leu substrate preferences. However, the tertiary structures of the two enzymes are similar and do not provide an explanation for the huge difference in substrate preference.

Thus, in an attempt to gain new insight into the nature of the different $P_1$ substrate preferences of serine carboxypeptidases we have turned our focus to the primary structures. By correlating an alignment of the primary structures of all known serine carboxypeptidases with the $P_1$ substrate preference of characterized enzymes we predict which structural features, identified in the tertiary structures of CPD-Y and CPD-WII, are important for the $P_1$ Lys/Leu substrate preference of such enzymes. These predictions are tested with CPD-Y as a model by the construction and kinetic characterization of mutant enzymes.

## MATERIALS AND METHODS

*Database Search and Alignment.* The EMBL database was analyzed for sequences similar to CPD-Y and/or CPD-WII using the TFASTA program while the Swissprot and PIR databases were analyzed using both the BLAST and FASTA programs on the BIOBASE server at the University of Århus, Denmark. The alignment of sequences displaying

† This work was supported by funds from BioNebraska Inc., Lincoln, NE.

* Corresponding author. Tel.: +45 33 27 52 40. FAX: +45 33 27 47 08. E-mail: ibs@biobase.dk.

⊗ Abstract published in *Advance ACS Abstracts,* November 15, 1995.

[1] Abbreviations: CPD-Y, yeast carboxypeptidase Y; CPD-WII, wheat carboxypeptidase II; dsDNA, double-stranded DNA; FA, 3-(2-furyl-acryloyl); kDa, kilodalton; LB, Luria broth; L178X, mutant protein with leucine 178 replaced with amino acid X; L178X+W312Z, mutant protein with leucine 178 replaced with amino acid X and tryptophan 312 replaced with amino acid Z; ssDNA, single-stranded DNA; Wt, wild type; W312X, mutant protein with tryptophan 312 replaced with amino acid X. The binding site notation is that of Schechter and Berger (1967). Accordingly, the C-terminal amino acid of the substrate is denoted $P_1'$, and those in the amino-terminal direction from the scissile bond are denoted $P_1$, $P_2$, ... , $P_n$. In analogy, binding sites are denoted $S_1'$ and $S_1$, $S_2$, ... , $S_n$.

0006-2960/95/0434-15689$09.00/0 

significant similarity was constructed in Excel 5.0 by manual alignment of sequences.

*Reagents and Strains.* FA-Lys-Ala-OH, FA-Arg-Leu-OH, and FA-Lys-Leu-OH were from Bachem AG, while other FA-dipeptides were synthesized as previously described (Olesen *et al.*, 1994). Oligonucleotides were synthesized on an Applied Biosystems 394 DNA/RNA synthesizer. LB, 2× YT, and SOC media were prepared according to Sambrook *et al.* (1989). SC and YPD media were prepared according to Sherman (1991). *Escherichia coli* BMH71−18mutS {*thi supE* Δ(*lac-proAB*) [mutS::Tn10] F′[*proAB*$^+$ *laqI*$^q$ *lacZ*ΔM15]} (Kramer *et al.*, 1984; Zell & Fritz, 1987), *E. coli* JM109 {*recA1 supE44 endA1 hsdR17 gyrA96 relA1 thi* Δ(*lac-proAB*) F′[*traD36 proAB*$^+$ *laqI*$^q$ *lacZ*ΔM15]} (Yanisch-Perron *et al.*, 1985), and *S. cerevisiae* K2579LLR (Olesen & Kielland-Brandt, 1993), a spontaneous mutant of W2579 (*MAT***a** Δ*prc1 leu2−3 leu2−112 ura3−52 vps1−1*) (Nielsen *et al.* 1990) able to grow on low leucine concentrations, were from in-house stocks.

*Transformation and Isolation of Single-Stranded DNA.* Transformation of *E. coli* was performed by electroporation with a Bio-Rad Gene Pulser set at 25 μF, 200 Ω, and 2.5 kV in 2 mm cuvettes according to Dower *et al.* (1988). Transformation of yeast was performed according to Schiestl and Gietz (1989) and Gietz *et al.* (1992). Single-stranded DNA of pYSP1 (Olesen & Kielland-Brandt, 1993) was prepared from a transformed culture of JM109 as described by Olesen and Kielland-Brandt (1993).

*Mutagenesis and Sequencing.* Mutagenesis was performed by a modification of the procedure by Lewis and Thompson (1990). Single-stranded pYSP1 (0.2 pmol) was mixed with 0.5 pmol of ampicillin-repair oligonucleotide (Promega, Altered Sites Kit) and 2 pmol of the mutagenic oligonucleotide in 80 μL of 1× annealing buffer (20 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 50 mM NaCl), heated to 70 °C for 5 min, and cooled slowly (~1 h) to room temperature. After the annealing reaction 10 μL of 10× synthesis buffer (100 mM Tris-HCl, pH 7.5, 5 mM of each of the four dNTPs, 10 mM ATP, 20 mM DTT), 10 Weiss units of T4 ligase (New England Biolabs), 20 units of T4 DNA polymerase (Promega), and $H_2O$ to 100 μL were added. The polymerase/ligation mixture was incubated at 37 °C for 2 h. Subsequently, it was phenol extracted, ethanol precipitated, resuspended, and used to transform electrocompetent *E. coli* BMH71−18mutS cells. After incubation for 30 min at 37 °C in SOC, an aliquot of the transformed cells was plated on LB with 60 mg of ampicillin/L to determine the number of transformants while the rest of the cells were grown overnight in 50 mL of LB with 60 mg of ampicillin/L. A plasmid preparation from this primary culture contains up to 50% ampicillin-sensitive plasmids originating from the unmutated parent strand. To enrich for mutants, 1 μg of miniprep DNA from this culture was used for a secondary transformation of JM109, followed by a second preparation of plasmid DNA. Sequencing was performed by the Applied Biosystems dsDNA Taq DyeDeoxy terminator procedure for use with the Applied Biosystems model 373A DNA-sequencing system to confirm the introduced mutations and to ensure that no secondary mutations had occurred. Mutagenic oligonucleotides were 33 bases long with the mutated codon situated at the center. Double mutants (L178X+W312X) were constructed by simultaneous mutagenesis with two oligonucleotides.

*Isolation and Kinetic Characterization of Mutant CPD-Y.* Overexpression of mutant enzymes was performed according to Nielsen *et al.* (1990). The mutant *Bam*HI-*Bam*HI *PRC1* fragments of pYSP1 were inserted into the *GAL*-expression vector pRA21 (Olesen & Kielland-Brandt, 1993) and introduced into yeast strain K2579LLR. Mutant forms of CPD-Y were purified from culture supernatants by affinity chromatography as described by Johansen *et al.* (1976). The purity of the mutant enzymes was ascertained to be >98% by SDS−PAGE on 12.5% homogeneous gels using the PhastSystem from Pharmacia. Enzyme concentrations were calculated using $\epsilon_{280}$ = 88 $mM^{-1}$ $cm^{-1}$ for enzymes mutated at position 312, $\epsilon_{280}$ = 95 $mM^{-1}$ $cm^{-1}$ for all other enzymes, and MW = 64 kDa. The rates of hydrolysis of FA-peptide substrates were measured spectrophotometrically at 337−352 nm (depending on substrate concentration) in 50 mM MES, 1 mM EDTA, pH 6.5, 25 °C. The following substrate concentrations were used: FA-Phe-Ala-OH, 0.05−0.5 mM; FA-Leu-Ala-OH and FA-Val-Ala-OH, 0.2−2.0 mM; FA-Ala-Ala-OH, 0.1−2.0 mM; FA-Ser-Ala-OH and FA-Arg-Leu-OH, 0.1−1.0 mM; FA-Glu-Ala-OH, 0.5−5.0; and 0.25−2.5 mM for the remaining substrates. Precautions were taken to avoid exposure to light due to the sensitivity of the FA group (Kanstrup & Buchardt, 1991). Kinetic parameters were determined by fitting data to the Michaelis−Menten equation using the GraFit 3.01 program from Erithacus Software (Sigma).

## RESULTS AND DISCUSSION

*Alignment of Serine Carboxypeptidase Sequences.* Alignments of serine carboxypeptidases previously published have been based on only a few primary structures [e.g., Sørensen et al. (1987)]. Since large portions of the structures differ and many insertions/deletions are necessary to maximize the identity, the reliability of such alignments is very limited. To overcome this problem we have included all primary structures from the EMBL, Swissprot, and PIR databases (Table 1) that are similar to CPD-Y and CPD-WII in our alignment (Figure 1). In the region corresponding to positions 200−320 of CPD-Y the enzymes show only little identity, causing errors in previous alignments, but due to the number of sequences included in our alignment we believe that the identity we find in this region is reliable. The identity among the primary structures is not evenly distributed throughout the sequences. Rather, regions of high identity alternate with regions of low identity which often contain numerous insertions/deletions. In the tertiary structures of CPD-Y and CPD-WII, regions of high identity generally are situated in the interior of the enzymes while regions of low identity reside on the surface of the enzymes. Structurally, two families of serine carboxypeptidases exist; some, like CPD-Y, consist of a single polypeptide chain while others, like CPD-WII, consist of two chains as a result of either a nick or the excision of an internal part of the protein during maturation of the proenzyme. The resulting C- and N-terminal parts of the resulting two peptide chains are cross-linked through a disulfide bridge (corresponding to positions 262 and 268 in CPD-Y). This disulfide bridge is also present in most of the one-chain enzymes. The KEX1 enzyme contains a 235 amino acid C-terminal extension relative to CPD-Y (Dmochowska et al., 1987). This extension, probably preventing peptides other than the mating factor from being substrates, is not included in the alignment.

Table 1: Serine Carboxypeptidases (and Carboxypeptidase-like Hydroxy Nitrile Lyases)

| enzyme | species / database entry / reference |
|---|---|
| CPD-Y | baker's yeast (*S. cerevisiae* / SW:CBPY_YEAST / Valls et al., 1987; Svendsen et al., 1982 |
| YBY9 | baker's yeast (*S. cerevisiae*) / SW:YBY9_YEAST / Becam et al., 1994 |
| KEX1 | baker's yeast (*S. cerevisiae*) / SW:KEX1_YEAST / Dmochowska et al., 1987 |
| SXA2 | fission yeast (*S. pombe*) / SW:SXA2_SCHPO / Imay & Yamamoto, 1992 |
| CPY1 | yeast (*Candida albicans*) / SW:CBPCANAL / Mukhtar et al., 1992 |
| CPD-S1 | *Penicillium janthinellum* (fungus) / SW:CPS1_PENJA / Svendsen et al., 1993 |
| CPD-S3 | *Penicillium janthinellum* (fungus) / / Svendsen & Day, 1995 |
| PEPF/CPD-An2 | *Aspergillus niger* (fungus) / PIR3:S47152 / van den Homberg et al., 1994 |
| CPD-As | *Aspergillus satoi* (fungus) / D25288 / Chiba et al., 1995 |
| CPD-Z | *Absidia zychae* (zygomycete) / EMBL:AZSCPZ / Lee et al., 1995 |
| CPD-M1 | barley (*Hordeum vulgare*) / SW:CBP1_HORVU / Doan & Fincher, 1988; Sørensen et al., 1986 |
| CPD-M2 | barley (*Hordeum vulgare*) / SW:CBP2_HORVU / Sørensen et al., 1987 |
| CPD-M2−1 | barley (*Hordeum vulgare*) / EMBL:HVACXPII1 / Dal Degan et al., 1994 |
| CPD-M2−2 | barley (*Hordeum vulgare*) / EMBL:HVACXPII2 / Dal Degan et al., 1994 |
| CPD-M2−3 | barley (*Hordeum vulgare*) / PIR3:S44191 / Dal Degan et al., 1994 |
| CPD-M3 | barley (*Hordeum vulgare*) / SW:CBP3_HORVU / Sørensen et al., 1989 |
| CPD-W2 | wheat (*T. aestivum*) / SW:CBP2_WHEAT / Breddam et al., 1987 |
| CPD-W3 | wheat (*T. aestivum*) / SW:CBP3_WHEAT / Baulcombe et al., 1987 |
| CPD-R1 | rice (*Oryza sativa*) / SW:CBP1_ORYSA / Washio & Ishikawa, 1994 |
| CPD-R3 | rice (*Oryza sativa*) / SW:CBP3_ORYSA / Washio & Ishikawa, 1992 |
| CPD-R3−1 | rice (*Oryza sativa*) / PX:D17587 / Washio & Ishikawa, 1993 |
| CBP-X | mouse-ear cress (*Arabidopsis thalina*) / SW:CBPX_ARATH / Bradley, 1992 |
| HNL-Sb | sorghum (*Sorghum bicolor*) / EMBL:SBHYDMAN / Wajant et al., 1994 |
| CPD-cef13d12 | *Caenorhabditis elegans* (nematode) / EMBL:CEF13D12 / Wilson et al., 1994 |
| CPD-cef32a5 | *Caenorhabditis elegans* (nematode) / EMBL:CEF32a5 / Wilson et al., 1994 |
| CPD-cef41c3 | *Caenorhabditis elegans* (nematode) / EMBL:CEF41c3 / Wilson et al., 1994 |
| CPD-Nf | *Naegleria fowleri* (amoeba) / PIR3:A43828 / Hu et al., 1992 |
| CPD-Ae | yellow fever mosquito (*Aedes aegypti*) / PIR2:A41612 / Cho et al., 1991 |
| PRTP-M | mouse (*Mus musculus*) / SW:PRTP_MOUSE / Galjart et al., 1990 |
| PRTP-H | human (*Homo sapiens*) / SW:PRTP_HUMAN / Galjart et al., 1988 |

Finally, it is interesting to note that a high similarity exists between serine carboxypeptidases and hydroxy nitrile lyase from *Sorghum bicolor* (HNL-Sb) (Wajant et al., 1994).

*$S_1$ Binding Pockets of CPD-Y and CPD-WII.* So far, only the three-dimensional structures of CPD-Y (Endrizzi et al., 1994) and CPD-WII (Liao & Remington, 1990) have been determined. Both of these enzymes belong to the $\alpha/\beta$ hydrolase fold (Ollis et al., 1992), and generally, the tertiary structures of the two enzymes are very similar. However, due to the insertions/deletions in the region corresponding to residues 180−317 of CPD-Y, which represents an insertion into the basic $\alpha/\beta$ hydrolase fold, the structure of this region is different within these two enzymes (Endrizzi et al., 1994). As a consequence, an α-helix, situated on the rim of the $S_1$ binding pocket of CPD-Y is positioned at a distance from the binding pocket in CPD-WII. Hence, the $S_1$ binding subsite is more secluded from solvent water in CPD-Y than in CPD-WII. Furthermore, some of the amino acids constituting the $S_1$ binding pockets of the two enzymes reside within this region. Thus, as the structures of this region differ, the $S_1$ binding pockets of the two enzymes are not entirely composed of the same portions of the primary structures. In CPD-Y, the $S_1$ binding pocket comprises the residues Tyr147, Leu178, Tyr185, Tyr188, Trp312, Ile340, and Cys341, while in CPD-WII the residues Tyr156 (corresponding to position 147 of CPD-Y), Leu187 (178), Phe224, Ile225, Val334 (340), and Val335 (341) constitute the solvent accessible surface of the $S_1$ binding pocket. If viewed from the point of the catalytically essential Ser146, the $S_1$ pockets can be defined as having a left side, a right side, an end, and a floor. By this definition, in CPD-Y Ile340 and Cys341 constitute the left side, Tyr147 and Trp312 the right side, Tyr185 and Tyr188 the end, and Leu178 the floor of the binding pocket (Figure 2A). In CPD-WII Val334 and Val335 constitute the left side, Tyr156 the right side, Phe224 and Ile225 the end, and Leu187 the floor of the binding pocket (Figure 2B). It should be noted that when the $P_1$ side chain of a substrate is bound to the binding pocket of either enzyme it may be partially accessible to solvent upward as the binding pockets have no "roof".

*Relationship between Primary Structure, Tertiary Structure, and $P_1$ Lys/Leu Substrate Preference.* On the basis of sequence identity and lengths of inserts/deletions in the region corresponding to positions 180−317 in CPD-Y we predict that the 12 uppermost enzymes listed in the alignment share the CPD-Y-like fold (α-helix on the rim of the $S_1$ binding pocket, single peptide chain), while the 14 bottommost enzymes share the CPD-WII like fold (no α-helix on the rim of the $S_1$ binding pocket, consisting of two peptide chains). The fold of the remaining four enzymes cannot be predicted. When correlating $P_1$ substrate preferences (when known) to the primary structures, it is observed that all enzymes predicted to share the CPD-WII-like fold have a high $P_1$ Lys/Leu substrate preference, regardless of the nature of the amino acids within the $S_1$ binding pocket (Table 2). On the other hand, not all enzymes predicted to share the CPD-Y-like fold exhibit a low $P_1$ Lys/Leu substrate preference. For the CPD-Y-like enzymes the residues corresponding to Tyr147, Tyr188, Ile340, and Cys341 in CPD-Y are fully conserved. Residues 178, 185, and 312 vary for enzymes exhibiting a high $P_1$ Lys/Leu substrate preference ($\geq 1$), while for those with a low $P_1$ Lys/Leu substrate preference ($\ll 1$) they are conserved as Leu, Tyr, and Trp, respectively (Table 2). Thus, the determinant for either low or high $P_1$ Lys/Leu substrate preference appears to be residue 178, 185, and/or 312. In CPD-S1 and CPD-Z, position 178 is occupied by Trp and Phe, respectively. It has previously been shown that the $P_1$ Lys/Leu substrate preference of CPD-Y increases 50-fold (to 0.003) by substituting Leu178 with a Trp and 24-fold (to 0.0015) with a Phe (Olesen et

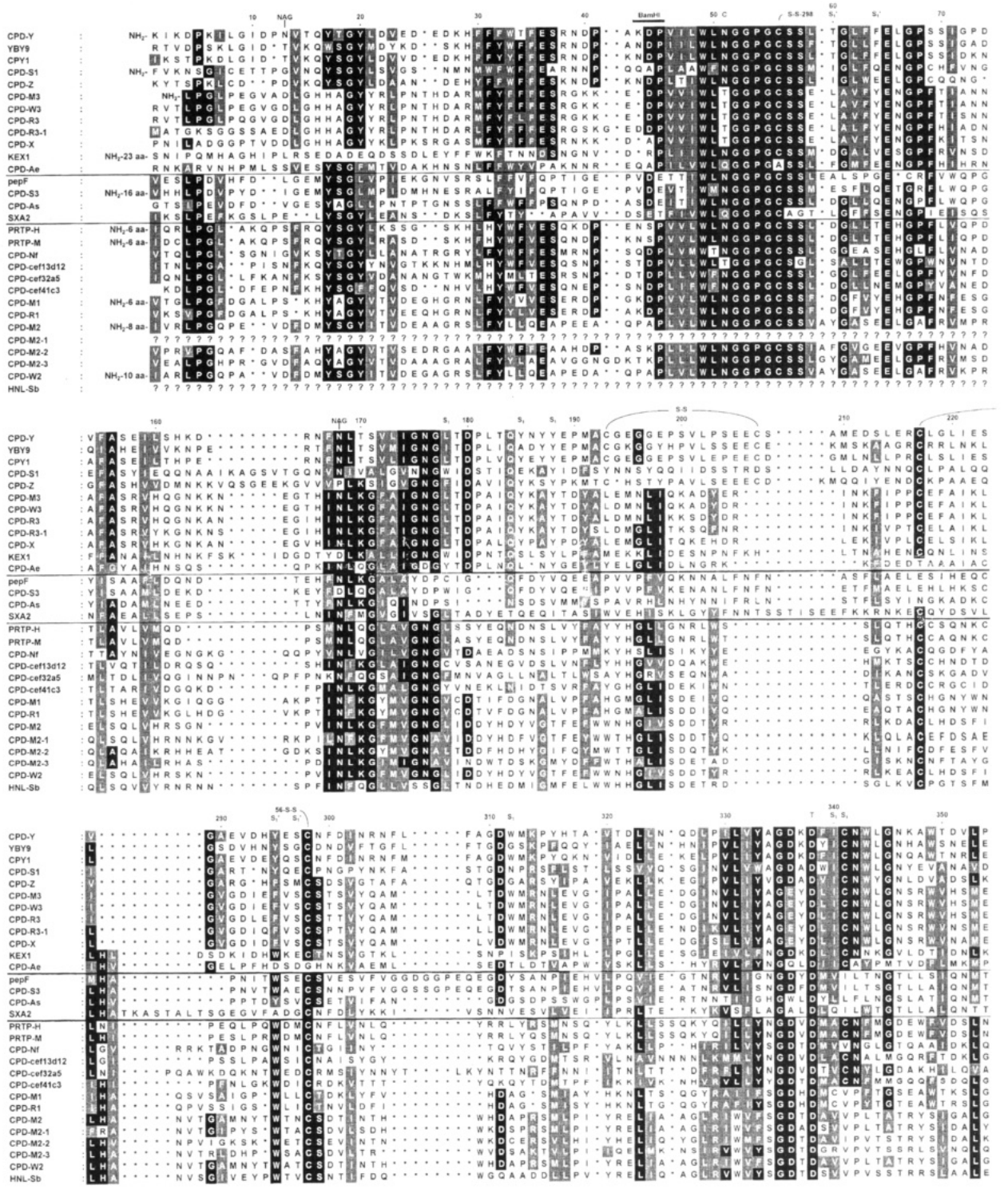
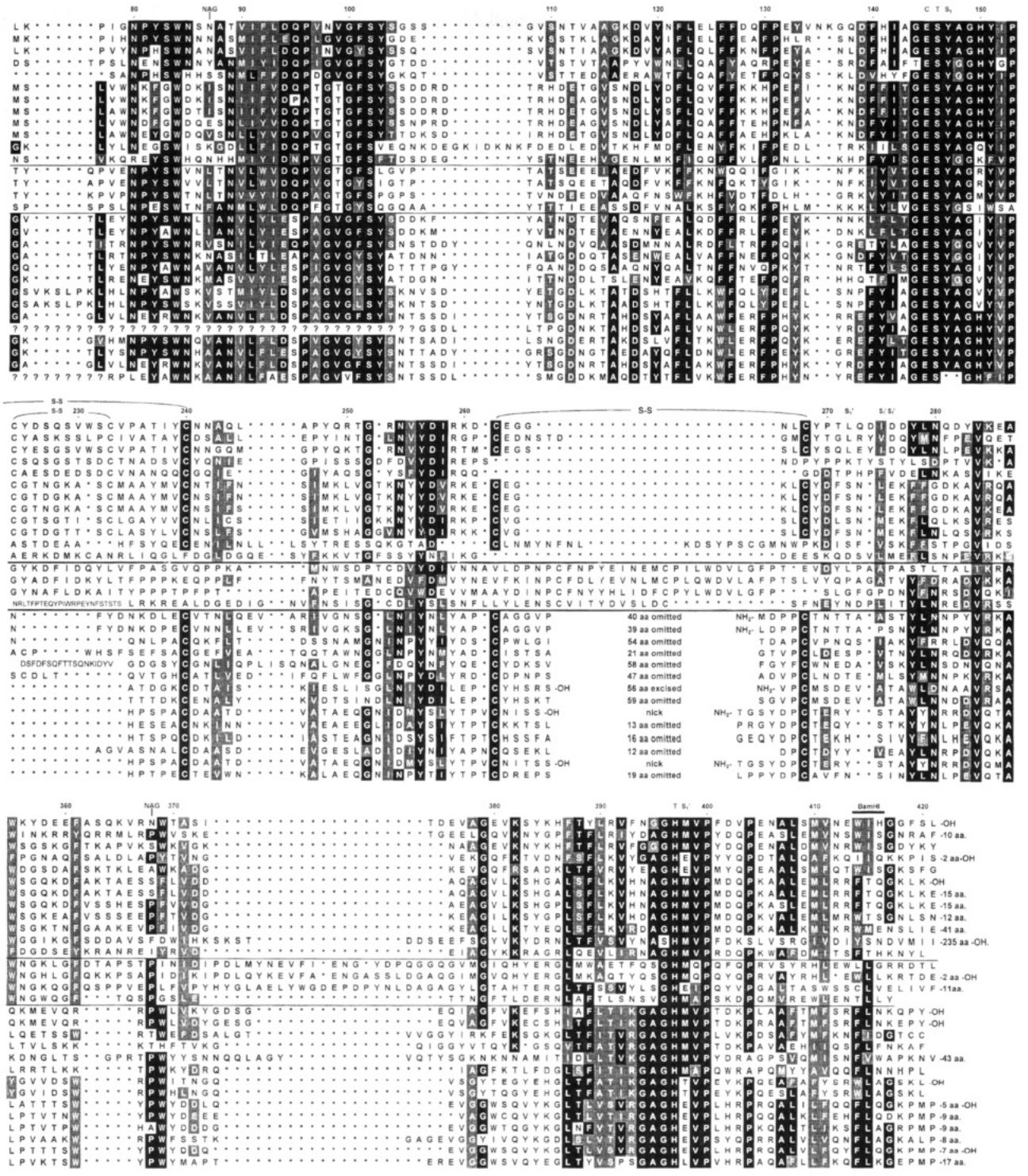


FIGURE 1: Alignment of serine carboxypeptidases and similar sequences. Conserved positions (identical in at least 50% of sequences) are indicated by white letters on black background, while semiconserved positions (similar in at least 50% of sequences) are indicated by white letters on grey background. Numbers correspond to amino acid positions of mature CPD-Y. Disulfide bridges and N-glycosylation sites of CPD-Y are indicated by "-S-S-" and "NAG", respectively. Residues of the catalytic triad are indicated by "T", while residues constituting either $S_1$, $S_1'$ or the binding pocket for the C-terminal carboxyl group of the substrate are indicated by "$S_1$", "$S_1'$", and "C", respectively. Pre- and proregions are not included. Sequences with N- or C-terminal extensions (in the mature enzyme) relative to CPD-Y are truncated

al., 1994), but this is still far lower than the $P_1$ Lys/Leu substrate preference of 64 observed for CPD-S1 (Breddam, 1986) and 2 for CPD-Z (Lee et al., 1993). Hence, the major determinant for high versus low $P_1$ Lys/Leu substrate preference in CPD-Y appears to be position 185 and/or 312, neither of which has previously been the target for a mutational investigation.

*Mutagenesis of CPD-Y.* To investigate the significance of positions 185 and 312 for the $P_1$ Lys/Leu substrate preference of CPD-Y, we have employed a strategy of random mutagenesis combined with functional screens as previously described (Olesen & Kielland-Brandt, 1993). For the sake of completion, position 188 was also included, as this would otherwise be left as the only uninvestigated




to the same N- or C-terminal position as CPD-Y; the length of the extension is indicated. Known N- and C-termini are indicated by "$NH_2$-" and "-OH", respectively. For two chain enzymes the number of amino acids excised are indicated. CPD-M2−3 and HNL-Sb are partial sequences; the unknown C-terminal parts are indicated by "?". Enzymes predicted to share the CPD-Y-like fold (top) or the CPD-WII-like fold (bottom) and enzymes that could not be assigned to either fold (middle) are separated by horizontal lines. Other partial sequences from the databases similar to serine carboxypeptidases but shorter than 200 amino acids, such as "expressed sequence tags" from *Caenorhabditis elegans*, *Arabidopsis thalina*, and *Dictyostelium discoideum* (EMBL:DD20608), rice and human, are not included in the alignment.

residue of the $S_1$ binding pocket in CPD-Y. The strategy involves making a library of random mutants by degenerated oligonucleotide directed mutagenesis, identifying transformants of yeast expressing active CPD-Y by a chromogenic overlay assay and finally, estimating the substrate preference of active mutants by a chromogenic coupled microtiter assay using CBZ-Xaa-Leu-OH substrates where Xaa = Phe, Gly, Pro, Glu, His, or Lys. As expected, no position 188 mutants with pronounced changes in substrate preference were found, consistent with it being fully conserved regardless of the substrate preference of the CPD-Y like enzymes. For positions 185 and 312, enzymes with elevated activity toward CBZ-Lys-Leu-OH were only isolated if Trp312 had been replaced by Asp or Glu, respectively. Hence, this position was chosen for a more detailed investigation, and another seven mutants were produced by site-directed mutagenesis

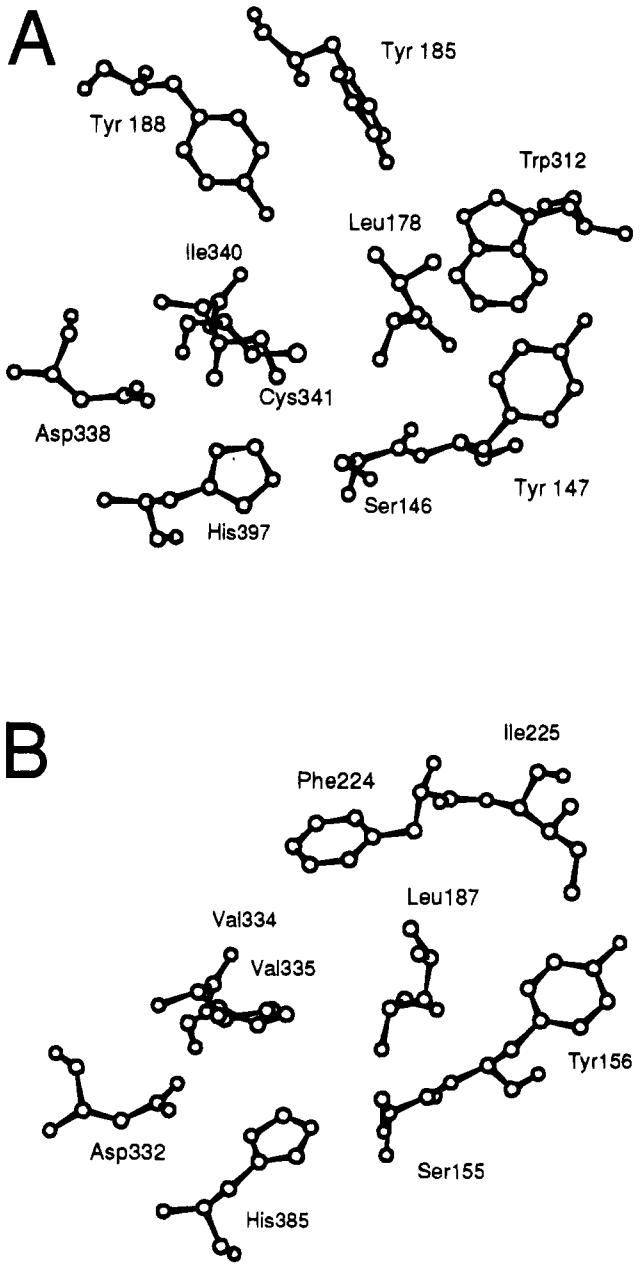



FIGURE 2: Catalytic triads and residues constituting the $S_1$ binding pockets of CPD-Y (A) and CPD-WII (B). In CPD-Y, Ser146, His397, and Asp338 constitute the catalytic triads while in CPD-WII it is constituted by Ser155, His385 and Asp332. In CPD-Y, Leu178 constitutes the floor of the $S_1$ binding pockets, Ile340 and Cys341 constitute the left side, Tyr147 and Trp312 constitute the right side, and Tyr185 and Tyr188 constitute the end. In CPD-WII Leu187 constitutes the floor, Val334 and Val335 constitute the left side, Tyr156 constitutes the right side, and Phe224 and Ile225 constitute the end of the $S_1$ binding pockets. Numbers refer to amino acid positions of the mature enzymes.

in which Trp312 had been replaced by Phe, Leu, Ala, Ser, Gln, Asn, or Lys. Finally, as CPD-S1 has a Glu at position 185 and displays a high $P_1$ Lys/Leu substrate preference, a mutant was produced in which Tyr185 had been replaced by Glu.

*Kinetic Characterization.* The resulting 10 mutants were characterized kinetically with respect to their $P_1$ substrate preferences using a series of FA-Xaa-Ala-OH peptide substrates systematically varied at the $P_1$ position. The $P_1$ substrate preference of CPD-WII has previously been described using Bz-Xaa-OMe substrates (Breddam et al., 1987), but in order to allow for direct comparison with the CPD-Y mutants the preference toward FA-Xaa-Ala-OH substrates was determined (Table 3). For the W312X mutants the substrates have been divided into two groups, one with uncharged amino acid residues (Phe, Leu Val, Ala, or Ser) (Table 4) and one with charged amino acid residues (Glu, Arg, or Lys) (Table 5) in the $P_1$ position.

With substrates containing hydrophobic $P_1$ side chains the general effect of substitutions at position 312 is a slight decrease in activity, up to 22-fold in the case of W312D with FA-Val-Ala-OH, primarily as a result of increased $K_m$ values. With this substrate acylation is probably the rate-limiting step ($k_2 \ll k_3$) in the CPD-Y-catalyzed hydrolysis, as it is normally the case with serine protease-catalyzed peptide hydrolysis (Fersht, 1985), and hence $K_m \approx K_s$. Thus, the increases in $K_m$ reflect weaker binding of the substrate as a result of reduced hydrophobic interactions with the position 312 residue. The substrate FA-Phe-Ala-OH represents an exception from this general trend as it is hydrolyzed with increased $k_{cat}/K_M$ values by most mutant enzymes, primarily as a result of decreased $K_m$ values. Presumably, a Phe side chain is too large to be accommodated in the $S_1$ binding pocket when a bulky Trp resides at position 312.

Table 2: Predicted Folding Type, $S_1$ Binding Pocket Constitution, and $P_1$ Lys/Leu Substrate Preference of Characterized Carboxypeptidases (with Known Primary Structure)

| enzyme | fold | $S_1$ residues[a] | $P_1$ Lys/Leu substrate preference (ref) |
|---|---|---|---|
| CPD-Y | Y | YL YYW IC | 0.00005 |
| CPD-S1 | Y | YW EYN IC | 64 (Breddam, 1988) |
| CPD-Z | Y | YF YYG IC | 2.0 (Lee et al., 1993) |
| CPD-M3 | Y | YL YYW IC | 0.0037 (Breddam & Sørensen, 1987) |
| CPD-S3 | Y | YW SYI IC | 1.3 (Dey & Aasmul-Olsen, 1993) |
| PEPF/CPD-An2 | ? | YC ?? VI | 4.2 (Dal Degan et al., 1992) |
| CPD-M1 | WII | YV WN CV | 15 (Breddam & Ottesen, 1983) |
| CPD-M2 | WII | YL FI VV | 7.1 (Breddam, 1985) |
| CPD-W2 | WII | YL FI VV | 5.2 |

[a] $S_1$ residues of enzymes with CPD-Y-like fold correspond to amino acid positions 147, 178, 185, 188, 312, 34,0 and 341 in the alignment, while $S_1$ residues of enzymes with CPD-WII-like fold correspond to positions 147, 178, 222, 223, 340, and 341 in the alignment.

Table 3: Kinetics of CPD-Y and CPD-WII

| substrate | enzyme | $k_{cat}$ ($min^{-1}$) | $K_m$ (mM) | $k_{cat}/K_m$ ($min^{-1}$·$mM^{-1}$) |
|---|---|---|---|---|
| FA-Phe-Ala-OH | CPD-Y | 6700[a] | 0.18[a] | 38000[a] |
| | CPD-WII | nd[b] | nd | 1800 |
| FA-Leu-Ala-OH | CPD-Y | 3800[a] | 0.11[c] | 35000[c] |
| | CPD-WII | 1500[a] | 0.98[c] | 1600[a] |
| FA-Val-Ala-OH | CPD-Y | 410[a] | 0.17[a] | 2400[a] |
| | CPD-WII | nd | nd | 30[a] |
| FA-Ala-Ala-OH | CPD-Y | nd | nd | 2300[a] |
| | CPD-WII | 430[a] | 1.6[c] | 270[a] |
| FA-Ser-Ala-OH | CPD-Y | 500[a] | 1.0[a] | 490[a] |
| | CPD-WII | 57[d] | 1.7[d] | 34[c] |
| FA-Glu-Ala-OH | CPD-Y | nd | nd | 41[a] |
| | CPD-WII | nd | nd | 11[a] |
| FA-Arg-Ala-OH | CPD-Y | nd | nd | 14[a] |
| | CPD-WII | nd | nd | 7200[a] |
| FA-Lys-Ala-OH | CPD-Y | 9[d] | 5[d] | 2[a] |
| | CPD-WII | 2000[a] | 0.24[a] | 8300[a] |

[a] Standard deviation, ±0%−10%. [b] nd, not determined due to high $K_m$. [c] SD ±10%−20%. [d] SD, ±20%−30%.

Table 4: Kinetics of CPD-Y Mutants toward Peptide Substrates Containing Uncharged $P_1$ Residues

| substrate | enzyme | $k_{cat}$ ($min^{-1}$) | $K_m$ (mM) | $k_{cat}/K_m$ ($min^{-1}\cdot mM^{-1}$) |
|---|---|---|---|---|
| FA-Phe-Ala-OH | wild type | 6700[a] | 0.18[a] | 38000[a] |
| | W312F | 7100[a] | 0.047[a] | 150000[a] |
| | W312L | 8300[a] | 0.076[a] | 110000[a] |
| | W312A | 17000[a] | 0.26[a] | 56000[a] |
| | W312S | 6800[a] | 0.20[b] | 35000[a] |
| | W312N | 10000[a] | 0.26[b] | 41000[a] |
| | W312Q | 8400[a] | 0.066[a] | 130000[a] |
| | W312D | 4700[a] | 0.17[a] | 27000[a] |
| | W312E | 6400[a] | 0.096[b] | 67000[a] |
| | W312K | 6400[a] | 0.86[b] | 7400[a] |
| FA-Leu-Ala-OH | wild type | 3800[a] | 0.11[b] | 35000[b] |
| | W312F | 5000[a] | 0.26[a] | 19000[a] |
| | W312L | 7200[a] | 0.39[b] | 18000[b] |
| | W312A | 5900[a] | 0.39[b] | 15000[a] |
| | W312S | 3600[a] | 0.51[a] | 7000[a] |
| | W312N | 5000[a] | 0.45[a] | 11000[a] |
| | W312Q | 9700[a] | 0.44[a] | 22000[a] |
| | W312D | 3400[a] | 0.43[a] | 8000[a] |
| | W312E | 3900[a] | 0.28[a] | 14000[a] |
| | W312K | 3000[a] | 1.7[b] | 1800[a] |
| FA-Val-Ala-OH | wild type | 410[a] | 0.17[a] | 2400[a] |
| | W312F | 200[a] | 0.43[a] | 450[a] |
| | W312L | 180[a] | 0.34[a] | 510[a] |
| | W312A | 270[a] | 1.1[a] | 250[a] |
| | W312S | 240[a] | 0.72[a] | 340[a] |
| | W312N | 230[a] | 0.71[a] | 320[a] |
| | W312Q | 170[a] | 0.67[a] | 260[a] |
| | W312D | 120[a] | 1.1[b] | 110[a] |
| | W312E | 150[a] | 0.60[a] | 250[a] |
| | W312K | 360[a] | 1.6[a] | 230[a] |
| FA-Ala-Ala-OH | wild type | nd[c] | nd | 2300[a] |
| | W312F | 160[a] | 0.27[a] | 580[a] |
| | W312L | 630[a] | 0.67[a] | 940[a] |
| | W312A | nd | nd | 790[a] |
| | W312S | 1300[b] | 2.3[b] | 580[a] |
| | W312N | 1600[a] | 1.7[a] | 940[a] |
| | W312Q | 510[a] | 0.71[b] | 720[a] |
| | W312D | 710[a] | 1.4[a] | 530[a] |
| | W312E | 1000[a] | 1.1[b] | 880[a] |
| | W312K | nd | nd | 270[a] |
| FA-Ser-Ala-OH | wild type | 500[a] | 1.0[a] | 490[a] |
| | W312F | 470[a] | 0.62[a] | 750[a] |
| | W312L | 340[a] | 0.58[a] | 580[a] |
| | W312A | 350[a] | 0.58[a] | 610[a] |
| | W312S | 320[a] | 0.77[a] | 420[a] |
| | W312N | 450[a] | 0.85[b] | 530[a] |
| | W312Q | 330[a] | 0.43[a] | 760[a] |
| | W312D | 110[a] | 0.43[a] | 250[a] |
| | W312E | 220[a] | 0.44[a] | 520[a] |
| | W312K | 840[a] | 3.4[a] | 250[a] |

[a] SD, ±0%−10%. [b] SD, ±10%−20%. [c] nd, not determined due to high $K_m$.

This idea is supported by the finding that for the wild type enzyme $K_m$ is actually smaller with Leu in $P_1$ than with Phe, even though the latter has the potential to make stronger hydrophobic contacts with the binding pocket. Furthermore, substituting Leu178 with the much larger Trp has no effect on $K_m$ with Phe in $P_1$ while it increases $K_m$ by a factor of 5 with Leu in $P_1$ (Olesen et al., 1994).

The activity with FA-Ser-Ala-OH, with an uncharged hydrophilic $P_1$ side chain is only slightly affected by the Trp312 substitutions. Presumably, Trp312 does not interact with the Ser side chain. Consistently, the $k_{cat}/K_M$ values with this substrate are comparable to those for substrates containing Val or Ala in $P_1$ lacking hydrophobic interactions in the mutant enzymes.

Table 5: Kinetics of CPD-Y Mutants toward Substrates Containing Charged $P_1$ Residues

| substrate | enzyme | $k_{cat}$ ($min^{-1}$) | $K_m$ (mM) | $k_{cat}/K_m$ ($min^{-1}\cdot mM^{-1}$) |
|---|---|---|---|---|
| FA-Glu-Ala-OH | wild type | nd[a] | nd | 41[b] |
| | W312F | nd | nd | 120[b] |
| | W312L | nd | nd | 160[b] |
| | W312A | nd | nd | 110[b] |
| | W312S | nd | nd | 84[b] |
| | W312N | nd | nd | 85[b] |
| | W312Q | nd | nd | 130[b] |
| | W312D | nd | nd | 120[b] |
| | W312E | nd | nd | 170[b] |
| | W312K | 850[b] | 2.9[c] | 290[b] |
| FA-Arg-Ala-OH | wild type | nd | nd | 14[b] |
| | W312F | nd | nd | 190[b] |
| | W312L | nd | nd | 410[b] |
| | W312A | nd | nd | 35[b] |
| | W312S | nd | nd | 49[b] |
| | W312N | nd | nd | 53[b] |
| | W312Q | nd | nd | 180[b] |
| | W312D | nd | nd | 830[b] |
| | W312E | nd | nd | 680[b] |
| | W312K | nd | nd | 3.6[c] |
| FA-Lys-Ala-OH | wild type | 9[d] | 5[d] | 2.0[b] |
| | W312F | nd | nd | 8.4[b] |
| | W312L | nd | nd | 16[b] |
| | W312A | nd | nd | 26[b] |
| | W312S | nd | nd | 64[b] |
| | W312N | nd | nd | 50[b] |
| | W312Q | 340[c] | 3.2[c] | 110[b] |
| | W312D | 5700[c] | 6.6[c] | 860[b] |
| | W312E | 3000[b] | 1.3[b] | 2300[b] |
| | W312K | nd | nd | 0.62[b] |
| FA-Arg-Leu-OH | wild type | 130[b] | 0.34[b] | 380[b] |
| | W312N | 2000[c] | 1.9[d] | 1000[b] |
| | W312Q | 3700[b] | 1.5[b] | 2400[b] |
| | W312D | 8000[b] | 0.71[b] | 11000[b] |
| | W312E | 3700[b] | 0.69[b] | 5400[b] |
| FA-Lys-Leu-OH | wild type | 29[b] | 0.68[c] | 43[b] |
| | W312F | nd | nd | 240[b] |
| | W312L | 490[b] | 2.0[b] | 250[b] |
| | W312A | nd | nd | 750[b] |
| | W312S | nd | nd | 660[b] |
| | W312N | nd | nd | 610[b] |
| | W312Q | nd | nd | 600[b] |
| | W312D | 7800[b] | 1.4[b] | 5600[b] |
| | W312E | 11000[b] | 0.71[b] | 15000[b] |
| | W312K | nd | nd | 9.3[b] |

[a] nd, not determined due to high $K_m$. [b] SD, ±0%−10%. [c] SD, ±10%−20%. [d] SD, ±20%−30%.

With FA-Glu-Ala-OH the activity increases slightly, up to 7-fold with W312K. With this substrate it is interesting to note that the activity also increases with W312D and W312E. Potentially, when both the $S_1$ binding pocket and the $P_1$ side chain are negatively charged, charge repulsion could occur. However, at least at the pH used for kinetic characterization (6.5), the distal ends of the position 312 residue and a $P_1$ Glu side chain appear to be sufficiently separated, or otherwise one or both side chains are uncharged.

The largest effects of the Trp312 substitutions occur with substrates containing basic $P_1$ side chains, where $k_{cat}/K_M$ increases up to 1150 times in the case of W312E with FA-Lys-Ala-OH. Only when Trp312 is replaced by a basic residue does the activity with such substrates decrease. To facilitate the determination of $k_{cat}$ and $K_m$ with Arg and Lys in $P_1$, some mutants have also been characterized with FA-Arg-Leu-OH and FA-Lys-Leu-OH (Table 5). The kinetics with these substrates demonstrate that the increases in activity with substrates containing basic $P_1$ side chains are entirely

Table 6: Kinetics of CPD-Y and Y185E

| substrate | enzyme | $k_{cat}$ ($min^{-1}$) | $K_m$ (mM) | $k_{cat}/K_m$ ($min^{-1}\cdot mM^{-1}$) |
|---|---|---|---|---|
| FA-Phe-Ala-OH | wild type | 6700[a] | 0.18[a] | 38000[a] |
| | Y185E | 8600[a] | 0.22[b] | 39000[a] |
| FA-Leu-Ala-OH | wild type | 3800[a] | 0.11[b] | 35000[b] |
| | Y185E | 3500[a] | 0.27[a] | 13000[a] |
| FA-Val-Ala-OH | wild type | 410[a] | 0.17[a] | 2400[a] |
| | Y185E | 240[a] | 0.24[a] | 1000[a] |
| FA-Ala-Ala-OH | wild type | nd[c] | nd | 2300[a] |
| | Y185E | nd | nd | 1400[a] |
| FA-Ser-Ala-OH | wild type | 500[a] | 1.0[a] | 490[a] |
| | Y185E | 340[a] | 1.6[b] | 210[a] |
| FA-Glu-Ala-OH | wild type | nd | nd | 41[a] |
| | Y185E | nd | nd | 9.0[a] |
| FA-Arg-Ala-OH | wild type | nd | nd | 14[a] |
| | Y185E | 45[b] | 1.7[b] | 26[a] |
| FA-Lys-Ala-OH | wild type | 9[d] | 5[d] | 2[a] |
| | Y185E | 16[a] | 1.3[a] | 12[a] |

[a] SD, ±0%−10%. [b] SD, ±10%−20%. [c] nd, not determined due to high $K_m$. [d] SD, ±20%−30%.

due to increases in $k_{cat}$. Thus, the beneficial effect of the Trp312 substitutions for the catalysis of these substrates is not on the ground state but entirely on the transition state complex.

The effects of the Tyr185 to Glu substitution are generally similar to those observed for Trp312 substitutions, except that the activity with Glu in $P_1$ decreases rather than increases and that the beneficial effects for the hydrolysis with Lys and Arg in $P_1$ are much smaller and are due more to changes in $K_m$ than in $k_{cat}$ (Table 6). Since CPD-Z exhibits a high $P_1$ Lys/Leu substrate preference with a Tyr at position 185, the position corresponding to 312 in CPD-Y is the major determinant for high versus low Lys/Leu ratio of the CPD-Y like carboxypeptidases.

*Interplay of Positions 178 and 312.* The altered substrate preference of some of the W312X mutant enzymes are very similar to those previously observed with some of the L178X enzymes (Olesen et al., 1994). In both cases a decrease is observed in activity toward substrates with uncharged amino acids in the $P_1$ position, primarily as a result of increased $K_m$ values, and an increase in activity toward those with basic $P_1$ side chains, as a result of increased $k_{cat}$ values. Thus, it was of interest to investigate whether these residues interfere leading to nonadditive effects of the mutational replacements, especially with respect to the hydrolysis of substrates with basic $P_1$ side chains.

To investigate this, three double mutants, L178S+W312N, L178D+W312N, and L178D+W312D, were constructed and characterized kinetically (Table 7). Like each of the corresponding single-substituted enzymes, all of these double-substituted enzymes exhibit decreased $k_{cat}/K_m$ values with substrates containing an uncharged $P_1$ side chain and increased $k_{cat}/K_m$ values with those containing basic $P_1$ side chains, the latter again primarily due to increases in $k_{cat}$. With L178D+W312D it is again of interest to note that $k_{cat}/K_m$ for the hydrolysis of FA-Glu-Ala-OH has increased although this enzyme possesses two potentially negatively charged residues in the $S_1$ binding pocket with the capacity of repulsing a negative charge of the Glu side chain of the substrate.

If the kinetic effects of substituting Leu178 and Trp312 of CPD-Y were additive, then the change in activation energy required for the enzymatically catalyzed transition

Table 7: Kinetics of CPD-Y Double Mutants

| substrate | enzyme | $k_{cat}$ ($min^{-1}$) | $K_m$ (mM) | $k_{cat}/K_m$ ($min^{-1}\cdot mM^{-1}$) |
|---|---|---|---|---|
| FA-Phe-Ala-OH | wild type | 6700[a] | 0.18[a] | 38000[a] |
| | L178S + W312N | 1200[a] | 0.24[a] | 4700[a] |
| | L178D + W312N | 260[a] | 0.42[b] | 620[a] |
| | L178D + W312D | 130[a] | 0.27[a] | 500[a] |
| FA-Leu-Ala-OH | wild type | 3800[a] | 0.11[b] | 35000[b] |
| | L178S + W312N | 770[a] | 0.60[a] | 1300[a] |
| | L178D + W312N | 140[a] | 1.1[a] | 130[a] |
| | L178D + W312D | 70[a] | 1.4[a] | 51[a] |
| FA-Val-Ala-OH | wild type | 410[a] | 0.17[a] | 2400[a] |
| | L178S + W312N | 42[a] | 0.68[a] | 62[a] |
| | L178D + W312N | 10[a] | 0.88[a] | 12[a] |
| | L178D + W312D | 2.1[a] | 0.58[a] | 3.6[a] |
| FA-Ala-Ala-OH | wild type | nd[c] | nd | 2300[a] |
| | L178S + W312N | 320[a] | 1.3[b] | 250[a] |
| | L178D + W312N | nd | nd | 21[a] |
| | L178D + W312D | nd | nd | 12[a] |
| FA-Ser-Ala-OH | wild type | 500[a] | 1.0[a] | 490[a] |
| | L178S + W312N | 120[a] | 0.53[a] | 230[a] |
| | L178D + W312N | 23[a] | 0.87[a] | 27[a] |
| | L178D + W312D | 6.5[a] | 0.39[a] | 17[a] |
| FA-Glu-Ala-OH | wild type | nd | nd | 41[a] |
| | L178S + W312N | 74[a] | 1.1[a] | 69[a] |
| | L178D + W312N | nd | nd | 30[a] |
| | L178D + W312D | 6.3[a] | 0.068[a] | 93[a] |
| FA-Arg-Ala-OH | wild type | nd | nd | 14[a] |
| | L178S + W312N | nd | nd | 640[a] |
| | L178D + W312N | nd | nd | 210[a] |
| | L178D + W312D | 280[a] | 0.39[a] | 710[a] |
| FA-Lys-Ala-OH | wild type | 9[d] | 5[d] | 2[a] |
| | L178S + W312N | nd | nd | 200[a] |
| | L178D + W312N | nd | nd | 150[a] |
| | L178D + W312D | 330[a] | 0.32[a] | 1100[a] |

[a] SD, ±0%−10%. [b] SD, ±10%−20%. [c] nd, not determined due to high $K_m$. [d] SD, ±20%−30%.

of a substrate into product(s) with a double mutant [$\Delta\Delta G_T^{\ddagger}$(Wt→L178X+W312X)] should equal the sum of each individual single mutant [$\Delta\Delta G_T^{\ddagger}$(Wt→L178X) + $\Delta\Delta G_T^{\ddagger}$(Wt→W312X)], where $\Delta\Delta G_T^{\ddagger}$(Wt→mutant) = $\Delta\Delta G_T^{\ddagger}$(Wt) − $\Delta\Delta G_T^{\ddagger}$(mutant) = $RT$ ln[$k_{cat}/K_m$(Wt)] − $RT$ ln[$k_{cat}/K_m$(mutant)] (Fersht, 1985). Qualitatively, this relation has been considered met when an observed value does not deviate by more than 4−6 kJ/mol from its corresponding calculated value [e.g., LiCarta and Ackers, (1995)]. Hence, by this definition, the kinetic effects of replacing Leu178 and Trp312 appear to be additive in about fifty percent of the cases (Table 8). However, if the data is viewed as a whole the observed data should relate to the calculated data according to the function $\Delta G_T^{\ddagger}$(observed) = 1 × $\Delta G_T^{\ddagger}$(calculated) in case of additivity. As seen in Figure 3 this is not the case; rather the data relate by the function $\Delta G_T^{\ddagger}$(observed) = 0.64 × $\Delta G_T^{\ddagger}$(calculated) ($r$ = 0.98). If instead the values observed for the double mutants instead are plotted against the numeric maximum values of the corresponding single mutants (rather than their sum), it is observed that this set of data correlate with a slope of 0.9 ($r$ = 0.97) (Figure 3). Hence, rather than inheriting the sum of the effects of the Leu178 and Trp312 substitutions, the double mutant enzymes exhibit characteristics corresponding to the most significant of the single mutations with respect to the activity toward each substrate.

*Salt Dependency.* The activity of W312D, W313E, and L178D+W312D (potentially with a negative charge in the $S_1$ binding pocket) toward FA-Glu-Ala-OH increases relative to the wild type enzyme. This is not due to a cation-mediated binding of the negatively charged Glu side chain of the

Table 8: $\Delta\Delta G_T^{\ddagger}$(Wt→mutant) Values of CPD-Y Double-Mutant and Corresponding Single-Mutant Enzymes

| substrate | enzyme | $^{\ddagger}\Delta\Delta G_T^{\ddagger}$ (kJ/mol) L178→X | W312→X | Σ | obsd |
|---|---|---|---|---|---|
| FA-Phe-Ala-OH | L178S+W312N | 2.3 | −0.2 | 2.1 | 5.2 |
| | L178D+W312N | 12.3 | −0.2 | 12.1 | 10.2 |
| | L178D+W312D | 12.3 | 0.8 | 13.1 | 10.7 |
| FA-Leu-Ala-OH | L178S+W312N | 5.9 | 2.9 | 8.8 | 8.1 |
| | L178D+W312N | 16.9 | 2.9 | 19.8 | 13.8 |
| | L178D+W312D | 16.9 | 3.8 | 20.7 | 16.1 |
| FA-Val-Ala-OH | L178S+W312N | 7.4 | 5.0 | 12.4 | 9.0 |
| | L178D+W312N | 19.1 | 5.0 | 24.1 | 13.1 |
| | L178D+W312D | 19.1 | 7.6 | 26.7 | 16.1 |
| FA-Ala-Ala-OH | L178S+W312N | 4.6 | 2.2 | 6.8 | 5.5 |
| | L178D+W312N | 14.7 | 2.2 | 16.9 | 11.6 |
| | L178D+W312D | 14.7 | 3.6 | 18.3 | 13.0 |
| FA-Ser-Ala-OH | L178S+W312N | 2.1 | −0.2 | 1.9 | 1.9 |
| | L178D+W312N | 3.9 | −0.2 | 3.7 | 7.2 |
| | L178D+W312D | 3.9 | 1.7 | 5.6 | 8.3 |
| FA-Glu-Ala-OH | L178S+W312N | −0.2 | −1.8 | −2.0 | −1.3 |
| | L178D+W312N | 4.1 | −1.8 | 2.3 | 0.8 |
| | L178D+W312D | 4.1 | −2.7 | 1.4 | −2.0 |
| FA-Arg-Ala-OH | L178S+W312N | −9.1 | −3.3 | −12.4 | −9.4 |
| | L178D+W312N | −7.2 | −3.3 | −10.5 | −6.7 |
| | L178D+W312D | −7.2 | −10.1 | −17.3 | −9.7 |
| FA-Lys-Ala-OH | L178S+W312N | −10.5 | −7.9 | −18.4 | −11.4 |
| | L178D+W312N | −11.1 | −7.9 | −19.0 | −10.7 |
| | L178D+W312D | −11.1 | −15.0 | −26.1 | −15.6 |

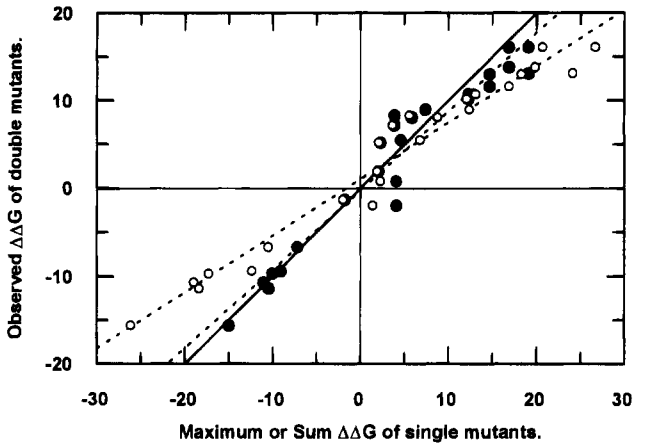



FIGURE 3: $\Delta\Delta G_T^{\ddagger}$(observed) of double-mutant enzymes plotted against $\Delta\Delta G_T^{\ddagger}$(sum) (○) and $\Delta\Delta G_T^{\ddagger}$(max) (●) values of corresponding single-mutant enzymes. Line indicates a one to one relationship.

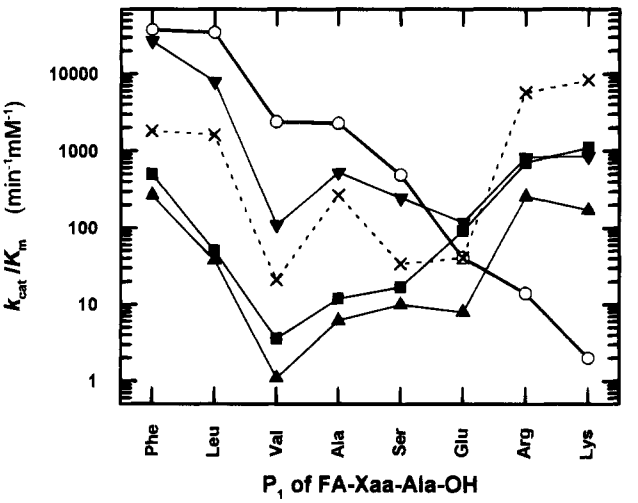



FIGURE 4: Substrate preference profiles of wild type CPD-Y (○), L178D (▲), W312D (▼), L178D+W312D (■), and CPD-WII (×). Substrates are ordered according to decreasing activity with wild type CPD-Y, and connecting lines are entirely pictorial.

substrate since the influence of NaCl, KCl, LiCl, $MgCl_2$, or $CaCl_2$ on the activity toward FA-Glu-Ala-OH is indistinguishable from that with FA-Leu-Ala-OH. The activity of these enzymes toward substrates with a basic $P_1$ side chain also increases. This is probably not due to the formation of a salt bridge between substrate and enzyme since the activity is not affected by addition of 0−1 M NaCl.

*Substrate Preference.* The two model enzymes, CPD-Y and CPD-WII, differ $10^5$-fold with respect to their $P_1$ Lys/Leu substrate preference (Table 1, Figure 4), to some extent due to the higher activity of CPD-Y toward substrates containing hydrophobic $P_1$ residues but in particular due to its much lower activity toward substrates containing basic $P_1$ residues (Table 1, Figure 4).

The properties of the amino acids constituting the $S_1$ binding pockets of the two enzymes offer no obvious explanation for this difference in substrate preference. The main difference between the two binding pockets is that the $S_1$ subsite of CPD-Y is more secluded from solvent water due to the α-helix on the rim. This may account for the higher activity of CPD-Y relative to CPD-WII toward substrates containing hydrophobic $P_1$ side chains.

It is highly unusual that the activity toward substrates containing basic $P_1$ side chains may be increased to such an extent by such a varied array of modifications, not only at one but at different positions within the $S_1$ binding pocket. It is unlikely that all of these modifications could lead to similar beneficial interactions between basic $P_1$ residues and the $S_1$ binding pocket. Likewise, the high activity of CPD-WII with these substrates cannot be accounted for by specific interactions between residues of the $S_1$ binding pocket and the distal positive charge of the $P_1$ side chain. Our interpretation is that binding between CPD-WII and a basic $P_1$ side chain involves hydrophobic interaction between the $S_1$ binding pocket and the hydrophobic part of the side chain while the distal hydrophillic charge remains "dissolved" in

solvent water. In CPD-Y, this may not be possible due to the α-helix on the rim deepening the binding pocket. Thus, reducing the bulk character of the position 312 residue renders the distal charge of the bound $P_1$ side chain accessible to solvent water through the side of the binding pocket. In the case of the Leu178 mutants it is possible that mutational substitution at position 178 causes Trp312 to adopt a tertiary conformation different from that in the wild type enzyme. If this is the case, then some of the catalytic effects of substituting Leu178 may reflect altered substrate interactions with Trp312 and hence, the position 178 mutants are in a sense already "double mutants". This would also explain the lack of additivity of the Leu178 and Trp312 substitutions.

Enzymes like L178D, W312D, and L178D+W312D display substrate preferences very different from CPD-Y but quite similar to CPD-WII (Figure 4). In the case of L178D the change in substrate preference is primarily due to an extreme reduction in activity with substrates containing hydrophobic $P_1$ side chains, while in the case of W312D it is primarily due to a large increase in activity with substrates containing basic $P_1$ side chains. L178D+W312D inherits both the very adverse effect of the L178→D substitution with respect to hydrolysis of substrates with hydrophobic $P_1$ side chains as well as the very beneficial effect of the W312→D substitution with respect to hydrolysis of substrates with basic $P_1$ side chains. As a consequence, this enzyme exhibits a larger change in $P_1$ Lys/Leu substrate preference than any of the corresponding single mutants (Figure 4). While the $P_1$ Lys/Leu substrate preference of the L178D and W312D enzymes has changed 80 000- and 1900-fold respectively, relative to the wild type CPD-Y, it has changed by 380 000-fold with the double mutant.

A number of mutation studies have been performed on other proteases in order to reveal the nature of substrate preference. With respect to producing mutant enzymes with greatly changed substrate preference one of the most successful examples comes from the studies of trypsin (Hedstrom et al., 1992), in which case the $P_1$ substrate preference was changed $3.5 \times 10^8$-fold. This example resembles our work with carboxypeptidases in two ways: (a) it was based on primary structures, tertiary structures, and substrate preferences of enzymes similar to trypsin, and (b) the effects were achieved by altering both the activity toward hydrophobic $P_1$ side chains as well as toward basic $P_1$ side chains. However, in one respect the example of trypsin is very different from that of the carboxypeptidases. With trypsin the level of activity toward substrates containing basic $P_1$ side chains was dependent on specific beneficial interactions between the $S_1$ binding pocket and the charge of the $P_1$ side chain, whereas with the carboxypeptidases it is dependent on the availability of solvent water to the charge of the $P_1$ side chain.

It is curious why, during the course of evolution, Trp at position 312 has been selected in CPD-Y, since with many other amino acids at this position the average catalytic capacity increases without a detectable decrease in the expression level of the enzyme. For example, with Gln or Glu at this position the activity dramatically increases with substrates containing Arg or Lys residues at the $P_1$ position, slightly increases with substrates containing Phe, Ser, or Glu, slightly decreases with Leu or Ala, and only decreases dramatically with Val. Although overexpression of mutant enzymes like W312D and W312E, both in a secreting *vps1* strain and a nonsecreting *VPS1* strain of yeast, does not seem to have any adverse effect on the growth of the cells, these enzymes might degrade peptides required at other stages of the yeast life cycle, e.g., mating, sporulation, or cell dormancy. If so, Trp could have been selected in the course of evolution to reduce activity with substrates containing Arg or Lys at $P_1$ in order to minimize hydrolysis of essential peptides. However, CPD-Y is generally thought to be involved in the breakdown of proteins in the vacuole, and therefore, we think that it is more likely that Trp at position 312 has been selected to optimize activity toward substrates with small nonpolar amino acids in $P_1$. Hence, Trp represents a compromise, which allows reasonable activity with small nonpolar $P_1$ side chains, especially Val, at the expense of optimal activity with Phe in $P_1$ and the almost total sacrifice of activity toward Arg and Lys. Conceivably, other carboxypeptidases are responsible for the breakdown of peptides containing Arg or Lys in the $P_1$ position.

## CONCLUDING REMARKS

We have shown that the substrate preference of one enzyme, CPD-Y, could be engineered to closely resemble that of another enzyme, CPD-WII, by combining information from primary structures, tertiary structures, and substrate preferences of similar carboxypeptidases. In the process it has been demonstrated that with serine carboxypeptidases the level of activity toward substrates containing basic $P_1$ side chains is not dependent on specific electrostatic interactions between the $P_1$ side chain and the $S_1$ binding pocket of the enzyme. Rather, with these enzymes the activity toward such substrates is dependent on the accessibility of solvent water to the distal charge of the $P_1$ side chain.

## REFERENCES


Baulcombe, D. C., Barker, R. F., & Jarvis, M. G. (1987) *J. Biol. Chem. 262*, 13726−13735.

Becam, A.-M., Cullin, C., Grzybowska, E., Lacroute, F., Nasr, F., Ozier-Kalogeropoulos, O., Palucha, A., Slonimski, P. P., Zagulski, M., & Herbert, C. J. (1994) *Yeast 10*, 1−11.

Bone, R., Silen, J. L., & Agard, D. A. (1989) *Nature 339*, 191−195.

Bradley, D. (1992) EMBL:ATCPYLP.

Breddam, K. (1983) *Carlsberg Res. Commun. 48*, 9−19.

Breddam, K., Sørensen, S. B., & Svendsen, I. (1987) *Carlsberg Res. Commun. 52*, 297−311.

Chiba, Y., Midorikawa, T., & Ichishima, E. (1995) *Biochem. J. 308*, 405−409.

Cho, W. L., Deitsch, K. W., & Raikhel, A. S. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 10821−10824.

Dal Degan, F., Ribadeau-Dumas, B., & Breddam, K. (1992) *Appl. Environ. Microbiol. 58*, 2144−2152.

Dal Degan, F., Rocher, A., Cameron-Mills, V., & von Wettstein, D. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 8209−8213.

Dey, E. S., & Aasmul-Olsen, S. (1993) *Enzyme Microb. Technol. 15*, 1042−1049.

Dmochowska, A., Dignard, D., Henning, D., Thomas, D. Y., & Bussey, H. (1987) *Cell 50*, 573−584.

Doan, N. P., & Fincher, G. B. (1988) *J. Biol. Chem 263*, 11106−11110.

Dower, W. J., Miller, J. F., & Ragsdale, C. W. (1988) *Nucleic Acids Res. 16*, 6127−6145.

Endrizzi, J. A., Breddam, K., & Remington, S. J. (1994) *Biochemistry 33*, 11106−11120.

Fersht, A. (1985) *Enzyme Structure and Mechanism*, 2nd ed., W. H. Freeman and Company, New York.

Galjart, N. J., Gillemans, N., Harris, A., van de Horst, G. T. J., Verheijen, F. W., Galjaard, H., & D'Azzo, A. (1988) *Cell 54*, 755−764.

Galjart, N. J., Gillemans, N., Meijer, D., & D'Azzo, A. (1990) *J. Biol. Chem. 265*, 4678−4684.
Gietz, D., St. Jean, A., Woods, R. A., & Schiestl, R. H. (1992) *Nucleic Acids Res. 20*, 1425.
Hedstrom, L., Szilagyi, L., & Rutter, W. J. (1992) *Science 255*, 1249−1253.
Hu, W. N., Kopachik, W., & Band, R. N. (1992) *Infect. Immunol. 60*, 2418−2424.
Huber, R., & Bode, W. (1978) *Acc. Chem. Res. 11*, 114−122.
Imay, Y., & Yamamoto, M. (1992) *Mol. Cell. Biol. 12*, 1827−1834.
Johansen, J. T., Breddam, K., & Ottesen, M. (1976) *Carlsberg Res. Commun. 41*, 1−14.
Kanstrup, A., & Buchardt, O. (1991) *Anal. Biochem. 194*, 41−44.
Kramer, B., Kramer, W., & Fritz, H.-J. (1984) *Cell 38*, 879−887.
Lee, B., Takeuchi, M., & Kobayashi, Y. (1993) *Biosci. Biotech. Biochem. 57*, 618−622.
Lee, B., Takeuchi, M., & Kobayashi, Y. (1995) *Curr. Genet. 27*, 159−165.
Lewis, M. K., & Thompson, D. V. (1990) *Nucleic Acids Res. 18*, 3439−3443.
Liao, D.-I., & Remington, S. J. (1990) *J. Biol. Chem. 265*, 6528−6531.
LiCarta, V. J., & Ackers, G. K. (1995) *Biochemistry 34*, 3133−3139.
Mukhtar, M., Logan, D. A., & Kaeufer, N. F. (1992) *Gene 121*, 173−177.
Nielsen, T. L., Holmberg, S., & Petersen, J. G. L. (1990) *Appl. Microbiol. Biotechnol. 33*, 307−312.
Olesen, K., & Kielland-Brandt, M. C. (1993) *Protein Eng. 6*, 409−415.
Olesen, K., Mortensen, U. H., Aasmul-Olsen, S., Kielland-Brandt, M. C., Remington, S. J., & Breddam, K. (1994) *Biochemistry 33*, 11121−11126.
Ollis, D., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S. M., Harel, M., Remington, S. J., Silman, I., Schrag, J., Sussman, J. L., Verschueren, K. H. G., & Goldman, A. (1992) *Protein Eng. 5*, 197−211.
Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Schechter, I., & Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157−162.
Schiestl, R. H., & Gietz, R. D. (1989) *Curr. Genet. 16*, 339−346.
Sherman, F. (1991) in *Methods in Enzymology*: *Guide to Yeast Genetics and Molecular Biology* (Guthrie, C., & Fink, C. R., Eds.) Vol. 194, pp 3−21, Academic Press, Inc., New York.
Sørensen, S. B., Breddam, K., & Svendsen, I. (1986) *Carlsberg Res. Commun. 51*, 475−485.
Sørensen, S. B., Svendsen, I., & Breddam, K. (1987) *Carlsberg Res. Commun. 52*, 285−295.
Sørensen, S. B., Svendsen, I., & Breddam, K. (1989) *Carlsberg Res. Commun. 54*, 193−202.
Svendsen, I., Day, E. S. (1995) *FEBS Lett. 371*, 1−3.
Svendsen, I., Hofmann, T., Endrizzi, J., Remington, S. J., & Breddam, K. (1993) *FEBS Lett. 333*, 39−43.
Svendsen, I., Martin, M., Viswanatha, T., & Johansen, J. T. (1982) *Carlsberg Res. Commun. 47*, 15−27.
Valls, L. A., Hunter, C. P., Rothman, J. H., & Stevens, T. H. (1987) *Cell 48*, 887−897.
van den Hombergh, J. P. T. W., Jarai, G., Buxton, F. P., & Visser, J. (1994) *Gene 151*, 73−79
Wajant, H., Mundry, K. W., & Pfitzenmaier, K. (1994) *Plant Mol. Biol. 26*, 735−746.
Washio, K., & Ishikawa, K. (1992) *Plant Mol. Biol. 19*, 631−640.
Washio, K., & Ishikawa, K. (1993) EMBL:OSCBP31.
Washio, K., & Ishikawa, K. (1994) *Biochim. Biophys. Acta. 1199*, 311−314.
Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J., Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A., Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M., Johnston, L., Jones, M., Kershaw, J., Kirsten, J., Laister, N., Latreille, P., Lightning, J., Lloyd, C., McMurray, A., Mortimore, B., O'Callaghan, M., Parsons, J., Percy, C., Rifken, L., Roopra, A., Saunders, D., Shownkeen, R., Smaldon, N., Smith, A., Sonnhammer, E., Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M., Vaughan, K., Waterston, R., Watson, A., Weinstock, L., Wilkinson-Sproat, J., & Wohldman, P. (1994) *Nature 368*, 32−38.
Yanisch-Perron, C., Vieira, J., & Messing, J. (1985) *Gene 33*, 103−119.
Zell, R., & Fritz, H.-J. (1987) *EMBO J. 6*, 1809−1815.

BI951309B