

# Cleavage Efficiency of the Novel Aspartic Protease Yapsin 1 (Yap3p) Enhanced for Substrates with Arginine Residues Flanking the P1 Site: Correlation with Electronegative Active-Site Pockets Predicted by Molecular Modeling<sup>†,‡</sup>

Vicki Olsen,<sup>§</sup> Kunchur Guruprasad,<sup>||,⊥</sup> Niamh X. Cawley,<sup>§</sup> Hao-Chia Chen,<sup>#</sup> Tom L. Blundell,<sup>||</sup> and Y. Peng Loh<sup>\*,§</sup>

Section on Cellular Neurobiology, Laboratory of Developmental Neurobiology, the Endocrinology and Reproduction Research Branch, National Institute of Child Health and Human Development, National Institutes of Health, Maryland 20892, and the Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1QW, England

Received October 7, 1997

**ABSTRACT:** Yapsin 1, a novel aspartic protease with unique specificity for basic residues, was shown to cleave CCK<sub>13–33</sub> at Lys<sub>23</sub>. Molecular modeling of yapsin 1 identified the active-site cleft to have negative residues close to or within the S6, S3, S2, S1, S1', S2', and S3' pockets and is more electronegative than rhizopuspepsin or endothiapepsin. In particular, the S2' subsite has three negative charges in and close to this pocket that can provide strong electrostatic interactions with a basic residue. The model, therefore, predicts that substrates with a basic residue in the P1 position would be favored with additional basic residues binding to the other electronegative pockets. A deletion of six residues close to the S1 pocket in yapsin 1, relative to rhizopuspepsin and other aspartic proteases of known 3D structure, is likely to affect its specificity. The model was tested using CCK<sub>13–33</sub> analogues. We report that yapsin 1 preferentially cleaves a CCK<sub>13–33</sub> substrate with a basic residue in the P1 position since the substrates with Ala in P1 were not cleaved. Furthermore, the cleavage efficiency of yapsin 1 was enhanced for CCK<sub>13–33</sub> analogues with arginine residues flanking the P1 position. An alanine residue, substituting for the arginine residue in the P6 position in CCK<sub>13–33</sub>, resulted in a 50% reduction in the cleavage efficiency. Substitution with arginine residues downstream of the cleavage site at the P2', P3', or P6' position increased the cleavage efficiency by 21-, 3- and 7-fold, respectively. Substitution of Lys<sub>23</sub> in CCK<sub>13–33</sub> with arginine resulted not only in cleavage after the substituted arginine residue, but also forced a cleavage after Met<sub>25</sub>, suggesting that an arginine residue in the S2' pocket is so favorable that it can affect the primary specificity of yapsin 1. These results are consistent with the predictions from the molecular model of yapsin 1.

Peptide hormones and neuropeptides are synthesized as larger polypeptides that undergo a series of posttranslational modifications to produce bioactive peptides (1–4). Important among these modifications is the limited proteolytic cleavage performed by endoproteases at specific basic residue sites of the precursors. This type of processing is performed by endoproteases belonging to the subtilisin-like serine protease family, the proprotein convertases (5–14), as well as thiol- (15, 16), metallo- (17), and aspartic proteases (18–23), most notably the yapsins. The yapsins form a novel class of aspartic proteases with specificity for basic residues

and are related to the yeast yapsin 1 (Yap3p).<sup>1</sup> There are two members from yeast, yapsin 1 (18) and yapsin 2 (Mkc7) (19). In mammals, there is yapsin 3 (Pro-opiomelanocortin-Converting-Enzyme (PCE) (20), which has been well characterized and shown to be related to yapsin 1 (21). In addition, an aspartic protease from bovine chromaffin granules (CG) with properties similar to yapsin 3 has been reported (22).

The molecular modeling studies presented here provide the first insight into the nature of this unusual family of aspartic proteases. They reveal that the yapsin 1 structure contains several electronegative active-site pockets, including the S1 pocket, thereby supporting early results suggesting that additional basic residues between the P2–P6 and P2'–P6', flanking the cleavage site, enhance the catalytic efficiency of yapsin 1 (24, 25). A thorough kinetic analysis is reported here using CCK<sub>13–33</sub><sup>2</sup> analogues as model

<sup>†</sup> This research was supported by the Danish Natural Science Council, Grant 9400095 to V.O., and by the BBSRC Linked-Research Group in Protein Engineering of Aspartic Proteinases to K.G.

<sup>‡</sup> The coordinates of the yapsin 1 model will be deposited in the Brookhaven Protein Data Bank.

\* Correspondence should be addressed to this author at LDN, NIH, Bldg. 49, Rm. 5A38, Bethesda, MD 20892. Phone: (301) 496 3239. Fax: (301) 496 9938. E-mail: ypl@codon.nih.gov.

<sup>§</sup> Section on Cellular Neurobiology, Laboratory of Developmental Neurobiology.

<sup>||</sup> Department of Biochemistry, University of Cambridge, England.

<sup>⊥</sup> Present address: Bioinformatics Centre, University of Pune, India.

<sup>#</sup> The Endocrinology and Reproduction Research Branch.

<sup>1</sup> Abbreviations: CCK<sub>13–33</sub>, cholecystokinin<sub>13–33</sub>; Yap3p, Yeast aspartic protease 3; CG, chromaffin granules; PCE, pro-opiomelanocortin converting enzyme; ACTH<sub>1–39</sub>, adrenocorticotrophic hormone; PTH, phenylthiohydantoin; SCR, structurally conserved regions; SVR, structurally variable regions.

<sup>2</sup> Numbering of the residues in the synthetic CCK<sub>13–33</sub> peptides are done from the C-terminal end of the peptides.

Table 1: Selection of Fragments for Structurally Conserved Regions (SCRs) and Structurally Variable Regions (SVRs) in the Model of Yapsin 1<sup>a</sup>

fragment	type	PDB code <sup>b</sup>	fragment	type	PDB code <sup>b</sup>
Glu2-Thr7	SCR	6APR	Gln189-Gly198	SCR	6APR
Asn8-Gln11	SVR	4APE:Phe257	Ile199-Lys203	SVR	* 2YPR:A/Tyr15
Ser12-Asp44	SCR	4APE	Thr204-Leu235	SCR	4APE
Asn44a-Tyr53	SVR	not modeled	Gly236	SVR	6APR:E/Arg236
Gly54-Ser66	SCR	6APR	Ala237-Ser241	SCR	4APE
Asn67-Asn68	SVR	6APR:E/Tyr66	Arg242-Ile242a	SVR	4APE:Ser239
Thr69-Glu106	SCR	6APR	Gly243-Cys249	SCR	4APE
Thr107-Met111	SVR	1GOF:Phe194	Pro250-Glu257	SVR	6APR:E/Ile250
Gly119-Thr131	SCR	4APE	Ile258-Ser277	SCR	6APR
Tyr131a-Ser131i	SVR	4MDH:B/Asn195	Thr278-Gly279	SVR	2LDX:Val24
Gly132-Pro141	SCR	6APR	Thr280-Thr289	SCR	4APE
Ile141a-Gly144	SVR	not modeled	Ser290-Gly294	SVR	4APE:Gln288
Ala145-Ser158	SCR	6APR	Thr295-Leu315	SCR	6APR
Asp159-Ala160	SVR	* 2YPR <sup>c</sup> :A/Phe291	Glu316-Leu318	SVR	6APR:E/Phe311
Met161-Ile184	SCR	6APR	Glu319-Ala325	SCR	6APR
Val185-Ile188	SVR	not modeled	Arg326	SVR	5PEP:Lys320

<sup>a</sup> Residues are numbered according to porcine pepsin (PDB code, 5PEP) as in Figure 1. <sup>b</sup> The PDB codes for modeling the SCRs, and the PDB code, chain identifier, starting residue and residue number for modeling SVRs are shown. <sup>c</sup> \*2YPR: coordinates of yeast protease A complexed with an inhibitor (personal communication, Dr. Nora Cronin, Birkbeck College, U.K.).

substrates to test the model. CCK<sub>13-33</sub> was chosen as the model substrate, since recent studies showing the colocalization of CCK mRNA with yapsin-1-like immunoreactivity in rat cortex and hippocampus (21), and the ability of yapsin 1 to cleave CCK<sub>33</sub> to CCK<sub>22</sub> and CCK<sub>8</sub> (25), suggests that yapsin 1-related aspartic proteases are likely to play a role in the processing of pro-CCK and other prohormones in endocrine/neuroendocrine cells in vivo.

## EXPERIMENTAL PROCEDURES

### Materials

**Peptide Substrates.** CCK<sub>13-33</sub> and its analogues were custom synthesized by Peptide Technologies Inc. (Gaithersburg, MD). ACTH<sub>1-39</sub> was purchased from Bachem California (Torrence, CA).

**Yapsin 1.** Secreted yapsin 1 from *Saccharomyces cerevisiae* was purified as described (47). The concentration of the enzyme was determined by N-terminal sequence analysis. The specificity of secreted yapsin 1 has been compared to the membrane bound form and no evident difference in specificity was observed (26).

### Methods

**Modeling Yapsin 1.** The sequence of yapsin 1 was aligned with the sequences of other aspartic proteases of known three-dimensional structure obtained from the SWISSPROT protein sequence databank (27) using the program ALIGN (28). A phylogenetic tree was constructed on the basis of evolutionary distances derived from pairwise sequence identity comparisons using the method of Fitch and Margoliash (29). The three-dimensional model was constructed using the rule-based comparative modeling approach encoded in the COMPOSER suite of computer programs (30, 31) that are available in SYBYL software (Tripos Inc.). The model was based on the three-dimensional structures of the homologous rhizopuspepsin (29.7% sequence identity) (PDB code, 6APR at 2.5 Å resolution) (32) and endothiapepsin (28.2% sequence identity) (PDB Code, 4APE at 2.1 Å resolution) (33). As the sequence identity of the model to the known three-dimensional structures is less than 30%, template-based methods were used as a guide to align the

sequence of yapsin 1 to the sequences of the known three-dimensional structures. A structural template was derived by alignment of aspartic proteases of known three-dimensional structure using the structure comparison program COMPARE (34). The sequence of yapsin 1 was then aligned with this template using the programs SPROF/QSLAVE (28). The 3D structures of rhizopuspepsin and endothiapepsin were superimposed using MNYFIT (30) and the alignment of yapsin 1 derived from the template alignment was used to designate the position of loops in yapsin 1. The sequence alignment used for modeling is shown in Figure 1. The structurally conserved regions (SCRs) and the structurally variable regions (SVRs) used to model yapsin 1 are shown in Table 1. The model was refined using energy minimization techniques in the SYBYL software (Tripos Inc.).

**Modeling Peptide Substrates/Inhibitor in the Binding Sites.** To try to understand the basis for the substrate specificity of yapsin 1, we used the three-dimensional structure of mouse renin complexed with the decapeptide inhibitor CH-66 (PDB Code, 1SMR at 2.0 Å) (35) to model the CH-66 inhibitor in yapsin 1. The structure of CH-66 was chosen because it is the longest inhibitor defined by X-ray analysis (as a complex with mouse renin) and was valuable in defining S6 to S3'. The inhibitor modeling was achieved by superimposing the yapsin 1 model onto the crystal structure of the mouse renin-inhibitor complex and transferring the coordinates of the inhibitor onto the model. Residues in positions P6-P3' of CH-66 were replaced to correspond to equivalent residues in a modified CCK<sub>20-28</sub> peptide with arginine residues at the P2' and P3' positions in the place of Leu<sub>21</sub><sup>3</sup> and Gln<sub>20</sub>. The geometry around the scissile peptide bond P1-P1' was fixed as in the crystal structure complex of mouse renin in order to mimic a transition-state analogue. Replacements and small adjustments that were required to optimize the fit were carried out using FRODO (36) and further refinement was achieved

<sup>3</sup> For clarification, the numbering of yapsin 1, corresponding to porcine pepsin numbering, is displayed as normal sized characters whereas the numbering of the amino acid residues of the substrates are shown as subscript.

		ab10	20	30	40
PORC PEPSN	----	IGDEPLENY--	LDTEYFGTIG	IGTPAQDFTVIF	DTGSSNLWVPSV
CARP YEAST	----	GGHDVPLTNY--	LNAQYYTDTIT	LGTTPQNFKVIL	DTGSSNLWVPSN
CARP CRYPA	----	STGSATTTTPI	DSLDDAYITPV	QIGTPAQTLNLD	FDGSSDLWVFS
CARP RHINI	----	ASGSVPMVDYE-	NDVEYYGEVTV	GTGPIKLLKLD	FDGSSDMWFAST
Yapsin 1	-----	EIIITN----	QSFYSVDLEV	GTTPQNVTVLV	DTGSSDLWIMGS
		10	20	30	40
	abc45				
PORC PEPSN	Y----	CSS-----			
CARP YEAST	E----	CGS-----			
CARP CRYPA	E----	TT-----			
CARP RHINI	L----	CSS-----			
Yapsin 1	DNPYCSSNS	MGSSRRRVID	KRDDSSSGG	SLINDINPFG	WLTGTGSAIGPT
		50	60	70	80
			50	60	a 70
PORC PEPSN	-----	LACSDHNQFNP	DDSSSTFEAT-	SQELSSITY	
CARP YEAST	-----	LACFLHSKYD	HEASSSYKAN-	GTEFAIQY	
CARP CRYPA	-----	ASEVDGQTIY	TPSKSTTAKL	LSGATWSISY	
CARP RHINI	-----	CSNSHTKYD	PKKSSTYAAD-	GRTWSISY	
Yapsin 1	ATGLGGGSG	TATQSVPA	SEATMDCQQY	GTFTSTSGS	STFRSN-NTYFSISY
		100	110	120	130
		80	90	100	110
PORC PEPSN	GTG-SMTG	ILGYDTVQ	VGGISDTNQ	IFGLSETEP	GSFLYYAPFD
CARP YEAST	GTG-SLEG	YISQDTLS	IGDLTIPK	QDFAEATSE	PGLTFAFGK
CARP CRYPA	GDGSSSG	GDVYTD	TVSVDGLT	VTGQAVES	AKKVSSTF
CARP RHINI	GDGSSAS	GILATD	NVNLGGL	LKIKQTIEL	AKRESSAF
Yapsin 1	GDGTFA	SGTDFG	TDVLDL	SDLNVTGL	SFAVANET
		150	160	170	180
		129	abcde	fghi	139
PORC PEPSN	YPSISAS-	-----	GATPVFDNL	W----	DQGLVSDLF
CARP YEAST	YDTISVD-	-----	KVVPPFFYN	AI----	QDLDLDEK
CARP CRYPA	FSTLNTV	SP-----	TQKKTFFD	NAKA----	SLDSPVFT
CARP RHINI	FNTITTV-	-----	RGVKTPVD	NLI----	SQGLISRPI
Yapsin 1	LPELEV	TYSGST	ASHSGKAY	KYDNFPI	VLKNSGAI
		190	200	210	220
		d 160	170	180	abcde
PORC PEPSN	-DDSGSV	VLLGGID	SSYYTG	SLNWNVP	VSVE-----
CARP YEAST	DTENGGE	ATFGGID	ESKFKG	DITWLPV	RRK-----
CARP CRYPA	-APGTYN	FFGFID	TTAYTG	SITYTA	VSTKQ-----
CARP RHINI	SNGGGGE	YIFGGY	DSSKFK	GSLLTVP	IDNSE-----
Yapsin 1	-DAMHGT	ILFGAV	DHSKYTG	TLTYTIP	IVNTLSA
		240	250	260	270
		abcd200	210	220	230
PORC PEPSN	M----	DGETIAC	SGGCQA	IVDTGTSL	LTGPTSAI
CARP YEAST	L----	GDEYAE	LESHGAA	IDTGTSL	LITLPSGL
CARP CRYPA	V----	GSSTFK	STSIDG	IADFTGT	TLTYLP
CARP RHINI	I----	GGTTVS	SASFDA	ILDGTTL	LLLPDD
Yapsin 1	ISDSGS	SNKTLT	TTTKIP	ALVDSG	TTLTYLP
		280	290	300	310
			250	260	270
PORC PEPSN	GEMVISC	SSIDSL	PDI	IVFTING	VQYPLSP
CARP YEAST	GQYTLDC	NRDNL	PDLIF	NFNNGY	NFTIGP
CARP CRYPA	GGYVFPC	SA-T-L	PSFTF	GVSAR	IVIPGDY
CARP RHINI	GTYSITC	DT-SK	LQPLV	FTLGS	STFEVPS
Yapsin 1	GYVLD	DCPS-	DDSM	EIVFDF	GGFHIN
		330	340	350	360
		290	300	310	a 320
PORC PEPSN	MDVPTSS	GELWIL	GDFVFI	RQYYTV	FDR-ANNK
CARP YEAST	MDFPEP	VGP	LAIVG	DAFLR	KYYYSI
CARP CRYPA	S----	AGIGIN	IFGDV	ALKAA	FVVFNG
CARP RHINI	----	AAGGD	LAIFG	DVFLK	NNYVVF
Yapsin 1	TSDDT	TGT---	ILGDS	FLTNA	YVVYDL
		380	390	400	

FIGURE 1: Alignment of the yapsin 1 with other aspartic proteases. The yapsin 1 sequence was aligned with the sequences of porcine pepsin (SWISSPROT Code: PORC PEPSN), rhizopuspepsin (CARP RHINI), endothiapepsin (CARP CRYPA), and yeast proteinase A (CARP YEAST) using the program MALIGN (28). Only the mature sequences were used in the alignment. The numbering above PORC PEPSN corresponds to the numbering in porcine pepsin as in the Brookhaven PDB code 5PEP and is corrected after residue number 230. The lowercase letters indicate insertions in the yapsin sequence compared with other sequences. The numbering below yapsin 1 corresponds to the sequential numbering in yapsin 1. The first 71 amino acid residues corresponding to the signal peptide and the propeptide, and the 91 C-terminal residues are not shown and these residues were not modeled, since appropriate fragments to model these regions were not available in the protein structure databank.

by using options in SYBYL (Tripos Inc.) and the energy minimization procedures therein. Interactions with residues in the peptide binding subsites were identified by defining a 4.0 Å cutoff value between yapsin 1 and the peptide substrate atoms and these are listed in Table 2 (37, 38).

#### Assay of Yapsin 1 and Identification of Products

**CCK<sub>13-33</sub> and Its Analogues.** Yapsin 1 (0.2–6 nM) was incubated with an appropriate amount of peptide (20–500 μM) in a total volume of 100 μL (0.1 M sodium citrate, pH

Table 2: Binding Subsite Interactions in the Model of Yapsin 1 with Modified CCK<sub>20–28</sub><sup>a</sup>

binding subsites <sup>b</sup>	yapsin1 residues
S6	Gln11, <sup>c</sup> Ser12, Ser158, Ser277
S5	Ser12
S4	Thr219, Leu220
S3	Phe13, Asp77, Gly217, Thr218, Thr219
S2	Tyr75, Gly76, Asp77, Gly217, Thr218, Tyr222
S1	Phe13, Leu30, Asp32, Gly34, Ser35, Tyr75, Val120, Asp215, Gly217, Thr218
S1'	Gly34, Gln189, Leu213, Asp215, Thr218
S2'	Gly34, Ser35, Ser36, Asp37, Ile73, Tyr75, Val120, Leu128, Gln189
S3'	Leu128, Gln189, Leu213, Asp292

<sup>a</sup> Modified CCK<sub>20–28</sub>: Arg(P6)-Val(P5)-Ser(P4)-Met(P3)-Ile(P2)-Lys(P1)-Asn(P1')-Arg(P2')-Arg(P3'). <sup>b</sup> A distance cutoff (4.0 Å) was used to define the binding subsites. <sup>c</sup> Residues are numbered according to porcine pepsin (Protein Data Bank code 5PEP) (37, 38 ).

4.0) at 37 °C for 5–60 min. The reaction was stopped by the addition of 10  $\mu$ L glacial acetic acid. Products were separated from substrate on an LKB 2150 HPLC system using a Bio-Rad HiPore RP-318 column (5  $\times$  250 mm). Buffer A was 0.1% trifluoroacetic acid, and buffer B was 80% acetonitrile in 0.1% trifluoroacetic acid. For the synthetic CCK<sub>13–33</sub> peptides, a linear gradient of 10–40% buffer B in 30 min at 1 mL/min was used, except for CCK-(P2' Arg) and CCK(P2' Arg, P6' Arg), where a two-step gradient was used: 10 min at 1 mL/min of 5% buffer B followed by 5–40% buffer B in 30 min at 1 mL/min. Products were monitored by absorbance at 214 nm, and individual peaks were collected for identification by direct N-terminal amino acid sequencing. The N-terminal sequence analysis was carried out by Edman degradation using a Perkin-Elmer/Applied Biosystems Model 494A Procise Protein Sequencer.  $\beta$ -Lactoglobulin was used to determine the sequencing efficiency.

**ACTH<sub>1–39</sub>.** For a standard assay, 10  $\mu$ g (22  $\mu$ M) of ACTH<sub>1–39</sub> was incubated with enzyme (0.2 nM) in 100  $\mu$ L, 0.1 M sodium citrate, pH 4.0, at 37 °C for 30 min. The reaction was stopped with 10  $\mu$ L glacial acetic acid. Cleavage products were identified by HPLC as described previously (25).

#### Kinetic Analysis of Products Generated

**Determination of  $K_m$  and  $V_{max}$ .** The method of Lineweaver–Burk (39) was used to calculate  $K_m$  and  $V_{max}$  values for the generation of products from each substrate. The amount of enzyme to be used for each substrate was determined individually to ensure that >90% of the substrate remained at the end of the reaction, to minimize product inhibition. Substrate concentrations spanning approximate  $K_m$  values were from our previous report (25). Products from each substrate were measured in millimeters of peak height of absorbance at 214 nm and converted to nanomoles of product using standard curves generated under identical gradient conditions. The assays were done in triplicate, with an accompanying ACTH<sub>1–39</sub> standard assay to demonstrate that the specific activity of the enzyme was constant throughout this study. The kinetic parameters,  $K_m$  and  $V_{max}$ , were determined directly from Lineweaver–Burk plots created in SigmaPlot.

**Determination of  $k_{cat}$ .** The absolute concentration of yapsin 1 was determined by N-terminal sequence analysis of a purified preparation of the enzyme.  $K_{cat}$  values were obtained from  $V_{max}/[E]_T = k_{cat}$ , where  $[E]_T$  = total enzyme. Since the protein concentration of yapsin 1 and not the active

enzyme concentration was used to calculate  $k_{cat}$ , these values may be underestimates. However, this will not affect the comparison of the relative cleavage efficiencies ( $k_{cat}/K_m$ ) for the different substrates in this study, since the same preparation of enzyme was used for all assays and was shown to maintain its initial specific activity throughout the study.

## RESULTS

**Yapsin 1 Model.** Yapsin 1 was modeled to identify and predict the characteristics of the substrate binding subsites of this enzyme. A superposition of the  $\alpha$ -carbon atoms in the main chain of the yapsin 1 model and the crystal structure of rhizopuspepsin (Figure 2) identifies the sites of insertions and a deletion in yapsin 1 relative to rhizopuspepsin. The model contains 311 of the 407 aligned amino acid residues of the yapsin 1 sequence. The breaks in the polypeptide chain of the model of yapsin 1 correspond, in sequential numbering, to Asn43-Tyr118 (76 amino acid residues), Ile210-Gly216 (seven residues) and Val257-Ile269 (13 residues). The equivalent porcine pepsin numbers (PDB code, 5PEP), which will be used in the rest of this paper (as shown in Figure 1), are Asn44a-Tyr53, Ile141a-Gly144, and Val185-Ile188, respectively. These insertions in yapsin 1 could not be modeled, since appropriate fragments were not available from database searches. The long insertion (Asn44a-Tyr53) may correspond to a domain insertion that is probably stabilized by the disulfide bridge Cys45-Cys50. In the related yeast Bar1p (40), this is a much smaller insertion (26 residues). Such large insertions are also observed in the plant aspartic proteases and range between 99 and 104 amino acid residues, but are located elsewhere in the three-dimensional structure (41). It may be possible that such domains have a functional role. In yapsin 1, this domain is predicted to be distant from the active-site cleft and therefore not likely to be directly involved in the enzyme specificity.

**Binding Subsite Interactions.** The yapsin 1 model was superimposed onto the crystal structure complex of mouse renin with the CH-66 decapeptide inhibitor (PDB code, 1SMR). This positions the CH-66 decapeptide in the active-site cleft of the yapsin 1 model. The residues corresponding to P6–P3' of CH-66 were replaced with modified CCK<sub>20–28</sub> residues. The interactions between atoms in the peptide and yapsin 1 model were optimized using energy minimization techniques in SYBYL. Residues in each of the binding subsites (S6–S3') were indicated by a 4.0 Å cutoff values between atoms in the peptide and yapsin 1 (Table 2.) The S6 pocket is exposed and likely to accommodate a residue with a big side chain. Near the S6 pocket, but outside the

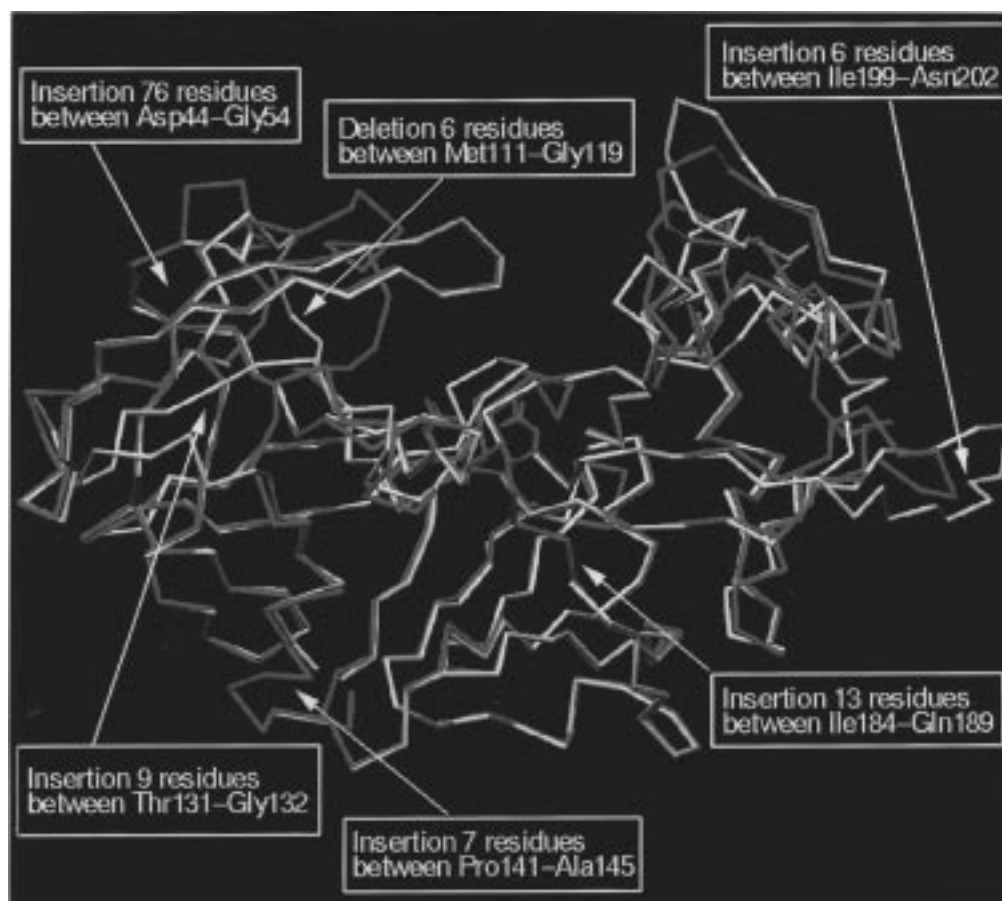


FIGURE 2: Structural superposition of the C $\alpha$  trace of yapsin 1 and the crystal structure of rhizopuspepsin (PDB code, 6APR) showing spatial location of the sites of insertions and deletions in the model. Color code: C $\alpha$  trace of yapsin 1 model (brown), C $\alpha$  trace of rhizopuspepsin (blue), side chains of catalytic aspartates (red), and disulfide bridges (yellow).

4.0 Å cutoff distance, two negatively charged residues (Asp157 and Asp159) may influence the preference for basic residues at the P6 position. The S5 subsite is exposed to the solvent.

The S3 pocket is large and would possibly prefer a large hydrophobic group such as methionine, whereas a phenylalanine may not be accommodated due to steric hindrance with Phe13. Asp77, also present in rhizopuspepsin and endotheiapepsin (Figure 1) is located on a "flap" that influences residues accommodated by the S3, S2, and S1 pockets.

The deletion of six residues (forming a helix) relative to rhizopuspepsin, between Met111 and Gly119 in yapsin 1, is predicted to affect yapsin 1 specificity. In other aspartic proteases, this region contributes to the S1 subsite interactions. In yapsin 1, the absence of residues close to the S1 pocket is likely to result in a more open and less hydrophobic pocket.

The side chain of an arginine residue in the P2' position of a peptide substrate is likely to interact with the negatively charged Asp37, also present in rhizopuspepsin and endotheiapepsin (Figure 1). On the basis of the CH-66 inhibitor, two negatively charged residues, Glu127 and Glu129, could also provide electrostatic interactions with the P2' residue. The insertion of nine residues between Leu188 and Gln189, not modeled, is likely to affect the S2' pocket further, possibly by providing a more hydrophobic environment. The presence of Asp292 in the S3' pocket is likely to favor the presence of a basic residue in the P3' position.

A view of the charge distribution calculated by GRASP (42) is shown in Figure 3. Overall the active-site cleft of yapsin 1 appears to be electronegative, with negatively charged residues either close to or within the S6, S3, S2, S1, S1', S2' and S3' pockets (see Table 2). This is likely to lead to a general preference for positively charged residues close to the scissile bond.

*Effect of Lysine versus Alanine at the P1 Site of CCK<sub>13-33</sub> on Cleavage by Yapsin 1.* Kinetic studies were performed using synthetic CCK<sub>13-33</sub> analogues as substrates to test the predicted properties of the substrate binding subsites determined from the yapsin 1 model. Incubation of CCK<sub>13-33</sub> with yapsin 1 generated two products in a time dependent manner (Figure 4A). N-terminal sequence analysis of the products identified two peptides (KAPSGRVSMIK<sub>23</sub> and NLQSLDPSHR<sub>13</sub>), indicating cleavage on the carboxyl side of Lys<sub>23</sub>. The kinetic studies determined the parameters:  $K_m$ ,  $k_{cat}$ , and  $k_{cat}/K_m$  as shown in Table 3. This result is consistent with the model predicting the ability of the S1 subsite to accommodate a residue with a long basic side chain, such as Lys<sub>23</sub>. Substitution of Lys<sub>23</sub> with an alanine residue resulted in minimal cleavage of the peptide (CCK P1 Ala) by yapsin 1. The two generated products were identified as KAPSGRVSM<sub>25</sub> and IANLQSLDPSHR<sub>13</sub> (Figure 5A). The poor cleavage of this substrate was demonstrated by cleavage of 50  $\mu$ M CCK<sub>13-33</sub> and CCK(P1 Ala) for 1 h with 2.1 pmol of yapsin 1 resulting in 71% cleavage of CCK<sub>13-33</sub>, and <1% cleavage of CCK(P1 Ala). (Data not shown).

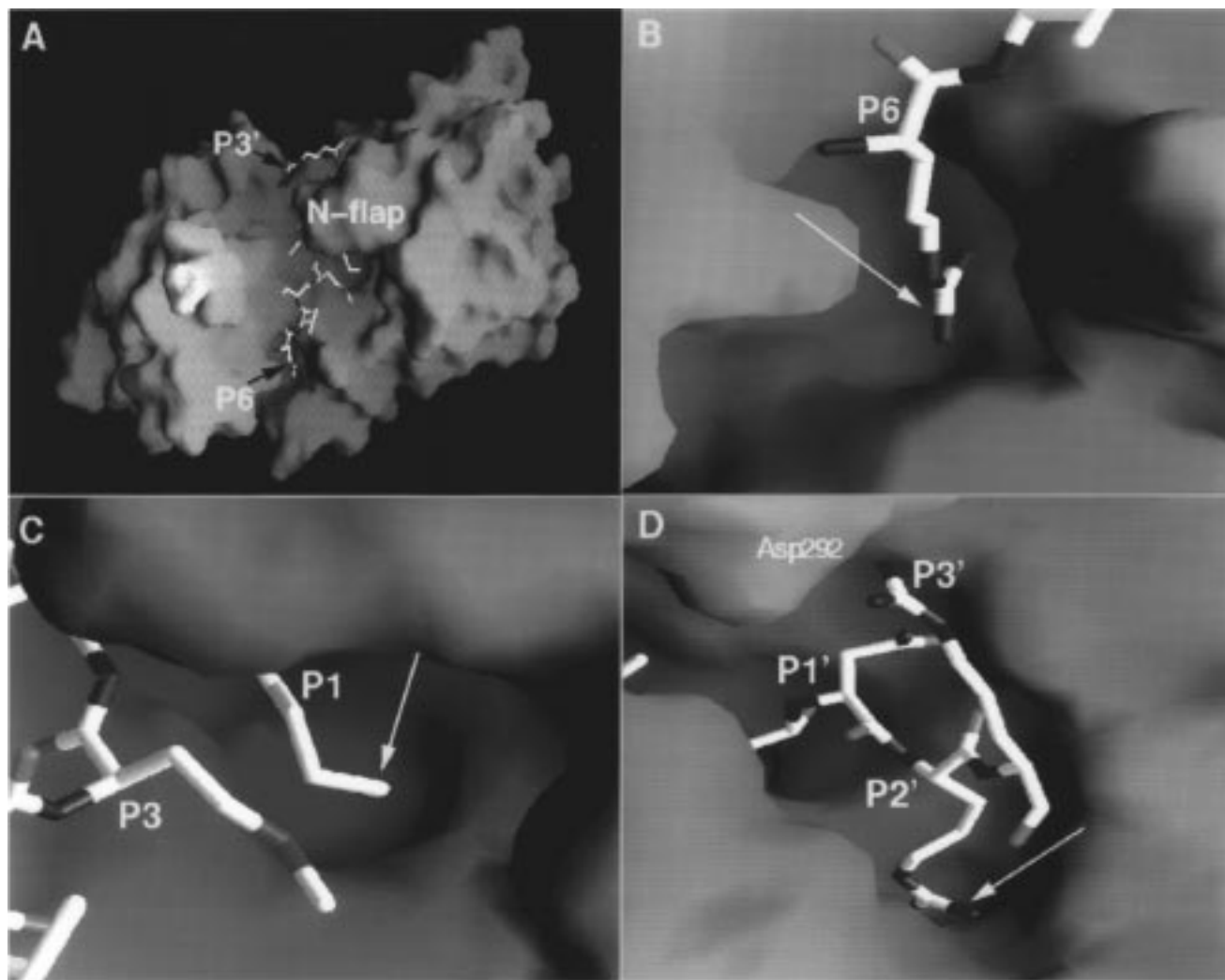


FIGURE 3: Modified CCK<sub>20–28</sub> peptide substrate in the active-site cleft of yapsin 1. (A) The molecular surface drawn by using GRASP (42) showing the relative charge distribution; (B) view of the S6 pocket showing accommodation of Arg (P6); (A) view of the S1 pocket showing accommodation of Lys (P1); (D) view of the S2' pocket showing accommodation of Arg (P2') and the P3' close to Asp292. Color code: negative charges (red), positive charges (blue).

**Effect of Substituting P6 Arg with Ala in CCK<sub>13–33</sub> on Cleavage by Yapsin 1.** The yapsin 1 model predicts that the S6 pocket is likely to show a preference for interaction with a residue having a long side chain. A CCK<sub>13–33</sub> analogue was synthesized with an alanine in the P6 position, CCK(P6 Ala), to examine the effect of a residue with a short side chain in this subsite as opposed to the arginine residue found in the original substrate. Incubation of this peptide with yapsin 1 resulted in the generation of two products: KAPSGAVSMIK<sub>23</sub> and NLQSLDPHSR<sub>13</sub> (Figure 4B). Kinetic analysis of the cleavage of this substrate showed a 50% decrease in  $k_{\text{cat}}/K_{\text{m}}$  compared to CCK<sub>13–33</sub> (Table 3).

**Effect of a Charged Residue in the P2' Position of CCK<sub>13–33</sub> on Cleavage by Yapsin 1.** According to the model, the negatively charged residues Asp37, Glu127, and Glu129 in the S2' subsite are likely to interact favorably with a positive charged residue in the P2' position. This hypothesis was tested by introduction of an arginine in the P2' position. This analogue, CCK(P2' Arg), was cleaved after Lys<sub>23</sub> by yapsin 1 as determined by N-terminal sequence analysis of the two generated products (KAPSGRVSMIK<sub>23</sub> and NRQSLDPHSR<sub>13</sub>) (Figure 4C). Kinetic analysis of the cleavage of this CCK<sub>13–33</sub> analogue by yapsin 1 showed a

significant 20.8-fold increase in  $k_{\text{cat}}/K_{\text{m}}$  compared to CCK<sub>13–33</sub>. Since the presence of an arginine in the P2' position so dramatically enhanced the catalytic efficiency, we also investigated the effect of introducing a negative charge in the P2' position. The analogue CCK(P2' Asp) was cleaved by yapsin 1 after Lys<sub>23</sub> to yield two products: KAPSGRVSMIK<sub>23</sub> and NDQSLDPHSR<sub>13</sub> (Figure 4D). Kinetic analysis of the cleavage efficiency of peptide CCK(P2' Asp) revealed a decrease of 60% compared to CCK<sub>13–33</sub> (Table 3). To investigate whether a lysine in the P2' position enhances the catalytic efficiency of yapsin 1 to the same degree as arginine, the peptide CCK(P2' Lys) was synthesized. Cleavage by yapsin 1 resulted in the generation of two peaks (Figure 4E) identified as KAPSGRVSMIK<sub>23</sub> and NKQSLDPHSR<sub>13</sub> by N-terminal amino acid analysis, indicating that yapsin 1 cleaved after Lys<sub>23</sub>, but not after the introduced lysine residue in the P2' position. Kinetic analysis of the cleavage efficiency did, surprisingly reveal a 40% reduction compared to CCK<sub>13–33</sub> (Table 3). To investigate further the degree of influence of the P2' arginine on cleavage, the Lys<sub>23</sub> was substituted for an alanine while having Arg<sub>21</sub> located in the P2' position, relative to the original cleavage site. Incubation of CCK(P1 Ala, P2' Arg)

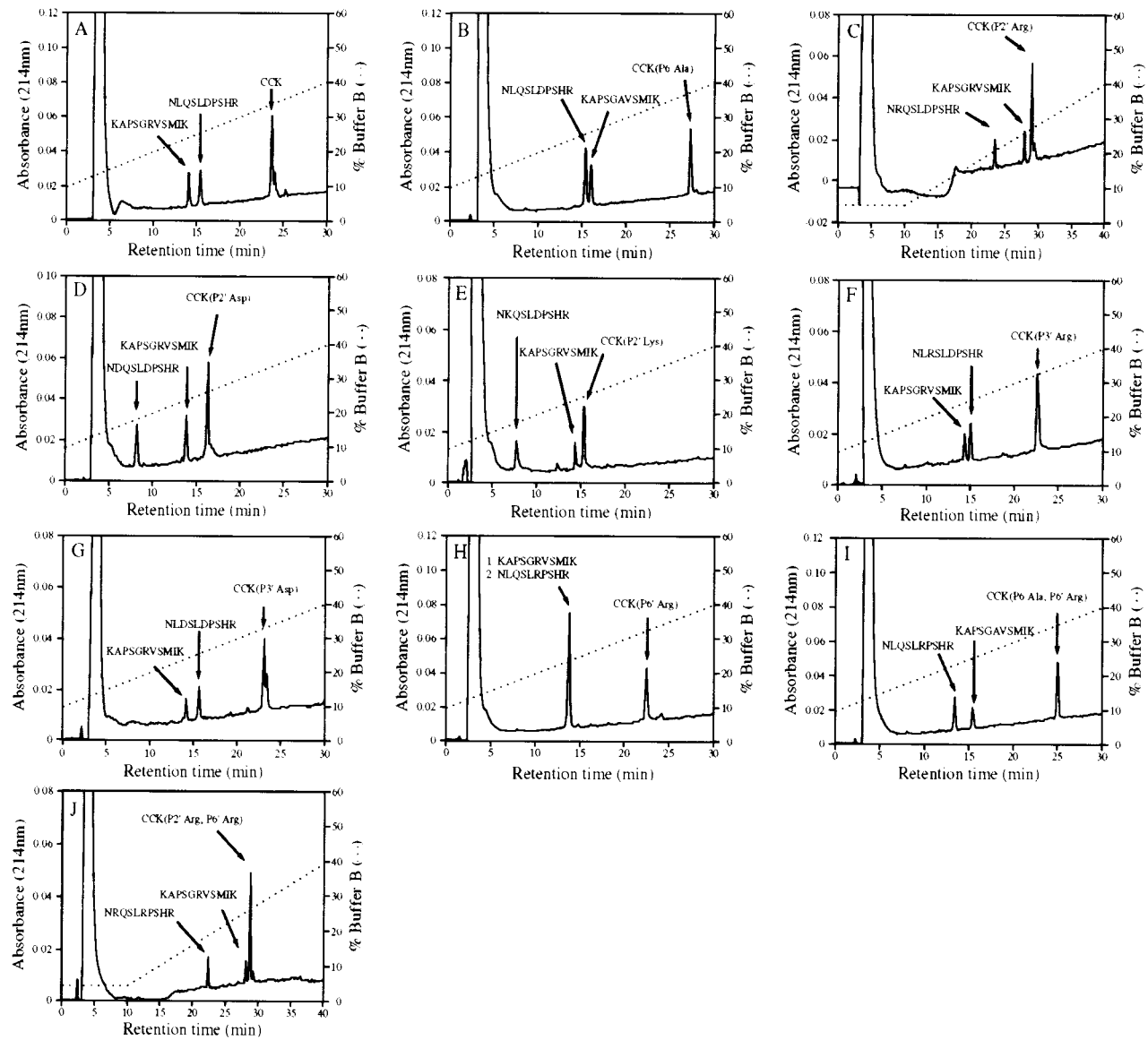


FIGURE 4: Proteolytic processing of synthetic CCK<sub>13–33</sub> peptides. HPLC profiles of products generated from synthetic CCK<sub>13–33</sub> peptides incubated with yapsin 1. (A) 6  $\mu$ g CCK incubated with 0.42 pmol of yapsin 1 for 1 h; (B) 6  $\mu$ g of CCK(P6 Ala) incubated with 0.84 pmol of yapsin 1 for 40 min; (C) 6  $\mu$ g of CCK(P2' Arg) incubated with 0.84 pmol of yapsin 1 for 40 min; (D) 12  $\mu$ g of CCK(P2' Asp) incubated with 2.1 pmol of yapsin 1 for 1 h; (E) 6  $\mu$ g of CCK(P2' Lys) incubated with 2.1 pmol of yapsin 1 for 1 h; (F) 8  $\mu$ g of CCK(P3' Arg) incubated with 0.42 pmol yapsin 1 for 1 h; (G) 8  $\mu$ g of CCK(P3' Asp) incubated with 2.1 pmol of yapsin 1 for 1 h; (H) 6  $\mu$ g of CCK(P6 Arg) incubated with 0.21 pmol yapsin 1 for 40 min; (I) 6  $\mu$ g of CCK(P6 Ala, P6' Arg) incubated with 0.21 pmol yapsin 1 for 40 min; (J) 4  $\mu$ g of CCK(P2' Arg, P6' Arg) incubated with 0.21 pmol yapsin 1 for 20 min.

Table 3: Kinetic Constants of Yapsin 1 for the Hydrolysis of Synthetic CCK<sub>13–33</sub> Peptides

	P6 P5 P4 P3 P2 P1 P1' P2' P3' P4' P5' P6'															
	K <sub>31</sub> A <sub>32</sub> P <sub>31</sub> S <sub>30</sub> G <sub>29</sub> R <sub>28</sub> V <sub>27</sub> S <sub>26</sub> M <sub>25</sub> I <sub>24</sub> K <sub>23</sub> N <sub>22</sub> L <sub>21</sub> Q <sub>20</sub> S <sub>19</sub> L <sub>18</sub> D <sub>17</sub> P <sub>16</sub> S <sub>15</sub> H <sub>14</sub> P <sub>13</sub>															
	$k_{\text{cat}}$ s <sup>-1</sup>				$K_{\text{m}}$ mol/L				$k_{\text{cat}}/K_{\text{m}}$ (mol/L) <sup>-1</sup> s <sup>-1</sup>				relative $k_{\text{cat}}/K_{\text{m}}$			
CCK	5.4 $\pm$ 1.2				(2.1 $\times$ 10 <sup>-4</sup> ) $\pm$ 0.6				(2.5 $\times$ 10 <sup>4</sup> ) $\pm$ 0.1				1.0			
CCK(P6 Ala)	1.7 $\pm$ 0.6				(1.5 $\times$ 10 <sup>-4</sup> ) $\pm$ 0.4				(1.3 $\times$ 10 <sup>4</sup> ) $\pm$ 0.4				0.5			
CCK(P2' Arg)	22.0 $\pm$ 3.9				(4.3 $\times$ 10 <sup>-5</sup> ) $\pm$ 1.4				(5.2 $\times$ 10 <sup>5</sup> ) $\pm$ 0.8				20.8			
CCK(P2' Asp)	3.0 $\pm$ 0.7				(2.7 $\times$ 10 <sup>-5</sup> ) $\pm$ 0.8				(1.1 $\times$ 10 <sup>4</sup> ) $\pm$ 0.1				0.4			
CCK(P2' Lys)	4.0 $\pm$ 1.7				(3.0 $\times$ 10 <sup>-4</sup> ) $\pm$ 1.6				(1.4 $\times$ 10 <sup>4</sup> ) $\pm$ 0.2				0.6			
CCK(P3' Arg)	8.5 $\pm$ 0.7				(1.1 $\times$ 10 <sup>-4</sup> ) $\pm$ 0.2				(8.1 $\times$ 10 <sup>4</sup> ) $\pm$ 0.5				3.2			
CCK(P3' Asp)	4.5 $\pm$ 0.6				(2.5 $\times$ 10 <sup>-5</sup> ) $\pm$ 0.5				(1.8 $\times$ 10 <sup>4</sup> ) $\pm$ 0.1				0.7			
CCK(P6' Arg)	8.0 $\pm$ 1.5				(4.7 $\times$ 10 <sup>-5</sup> ) $\pm$ 1.6				(1.7 $\times$ 10 <sup>5</sup> ) $\pm$ 0.2				6.8			
CCK(P6 Ala, P6' Arg)	6.0 $\pm$ 0.3				(5.5 $\times$ 10 <sup>-5</sup> ) $\pm$ 0.5				(1.1 $\times$ 10 <sup>5</sup> ) $\pm$ 0.1				4.4			
CCK(P2' Ala, P6' Arg)	25.4 $\pm$ 3.8				(4.4 $\times$ 10 <sup>-5</sup> ) $\pm$ 1.3				(5.8 $\times$ 10 <sup>5</sup> ) $\pm$ 0.9				23.2			

with yapsin 1 resulted in two product-peaks: KAPSGRVSMIANR<sub>21</sub> and QSLDPSHR<sub>13</sub> (Figure 5B), i.e., cleav-

age after Arg<sub>21</sub>. The cleavage rate of this substrate was, however, so slow that kinetic analysis was not performed.

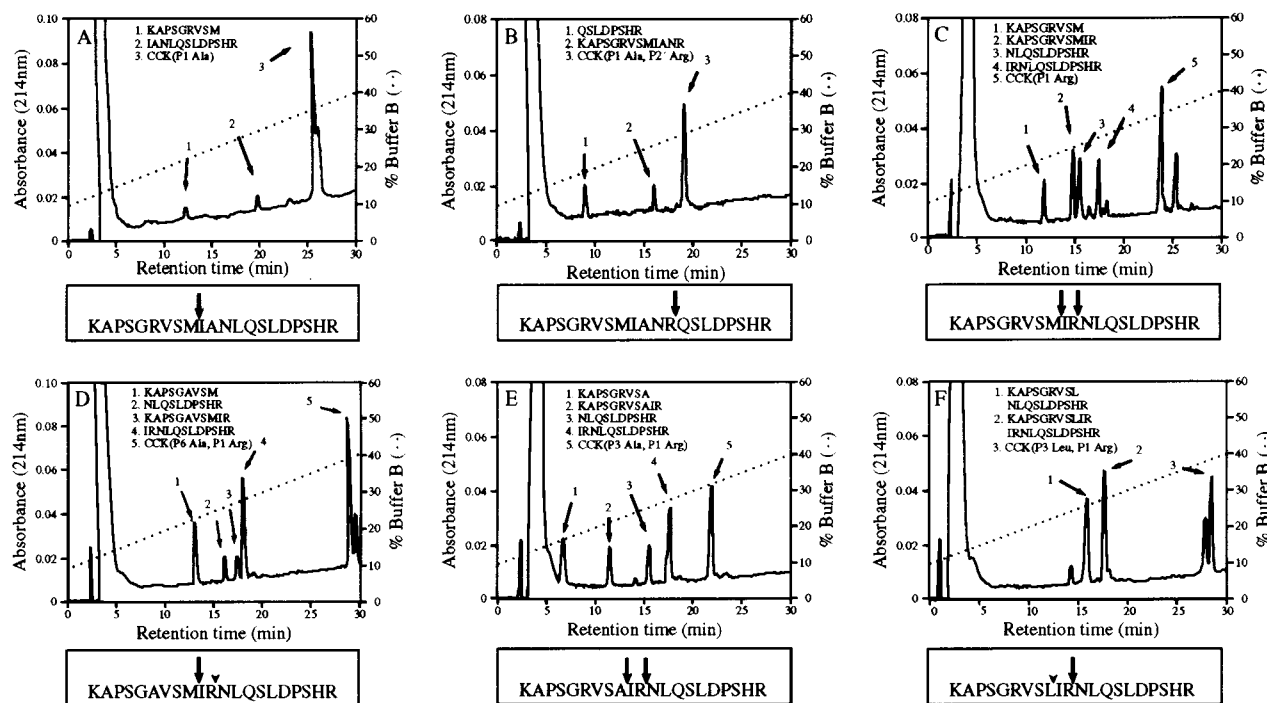


FIGURE 5: Proteolytic processing of synthetic CCK<sub>13-33</sub> peptides. HPLC profiles of products generated from synthetic CCK<sub>13-33</sub> peptides incubated with yapsin 1. The cleavage site of the individual peptides are indicated by arrows. An arrowhead indicated a minor cleavage relative to a second (major) in the same substrate. (A) 15  $\mu$ g of CCK(P1 Ala) incubated with 2.1 pmol of yapsin 1 for 4 h; (B) 6  $\mu$ g of CCK(P1 Ala, P2' Arg) incubated with 21 pmol of yapsin 1 for 30 min; (C) 15  $\mu$ g of CCK(P1 Arg) incubated with 0.21 pmol of yapsin 1 for 1 h; (D) 15  $\mu$ g of CCK(P6 Ala, P1 Arg) incubated with 0.21 pmol of yapsin 1 for 30 min; (E) 15  $\mu$ g of CCK(P3 Ala, P1 Arg) incubated with 2.1 pmol of yapsin 1 for 1 h; (F) 15  $\mu$ g of CCK(P3 Leu, P1 Arg) incubated with 0.21 pmol of yapsin 1 for 50 min.

Incubation of 50  $\mu$ M CCK<sub>13-33</sub> and CCK(P1 Ala, P2' Arg) for 1 h with 2.1 pmol of yapsin 1 resulted in 71% cleavage of the CCK<sub>13-33</sub>, whereas only 11% of CCK(P1 Ala, P2' Arg) was cleaved (data not shown).

**Effect of a Charged Residue in the P3' Position of CCK<sub>13-33</sub> on Cleavage by Yapsin 1.** The model predicts that Asp292 in the S3' subsite may significantly influence substrate interactions thereby favoring a basic residue in the P3' position. Arginine in the P3' position, CCK(P3' Arg), produced two products, KAPSGRVSMIK<sub>23</sub> and NLRSLDP-SHR<sub>13</sub>, after cleavage by yapsin 1 (Figure 4F). A 3-fold increase in  $k_{cat}/K_m$  was determined by kinetic analysis (Table 3). The effect of a negative charged residue in the P3' position was determined by examining the cleavage of CCK-(P3' Asp) by yapsin 1. After incubation with yapsin 1 the two generated products (Figure 4G) were identified as KAPSGRVSMIK<sub>23</sub> and NLDSLDP-SHR<sub>13</sub>. As shown in Table 3 the catalytic efficiency was decreased 30% compared to CCK<sub>13-33</sub>.

**Effect of an Arg in the P6' Position of CCK<sub>13-33</sub> on Cleavage by Yapsin 1.** Although no predictions could be made from the modeling studies with regard to the S6' subsite of yapsin 1, which is exposed to the solvent, previous studies have suggested that basic residues downstream of the cleavage site (P2'–P6') enhance the cleavage by yapsin 1 (25). To investigate this, CCK(P6' Arg) was incubated with yapsin 1, generating one product peak (Figure 4H). N-terminal amino acid analysis revealed two different species, KAPSGRVSMIK<sub>23</sub> and NLQSLRPSHR<sub>13</sub>. Kinetic analysis of the cleavage of this peptide showed an overall 6.8-fold increase in  $k_{cat}/K_m$  compared to CCK<sub>13-33</sub> (Table 3). The peptide CCK(P6 Ala, P6' Arg), was incubated with yapsin

1, generating two products, KAPSGAVSMIK<sub>23</sub> and NLQSLRPSHR<sub>13</sub> (Figure 4I). Kinetic studies showed a 4.4-fold increase in  $k_{cat}/K_m$  for this substrate compared to CCK<sub>13-33</sub> (Table 3). To determine if several arginine residues surrounding the cleavage site will increase the cleavage efficiency of yapsin 1 further, the peptide CCK(P2' Arg, P6' Arg) was synthesized. Incubation with yapsin 1 generated two products, KAPSGRVSMIK<sub>23</sub> and NRQSLRPSHR<sub>13</sub> (Figure 4J). Kinetic studies showed a 23.2-fold increase in the  $k_{cat}/K_m$  of CCK(P2' Arg, P6' Arg) compared to CCK<sub>13-33</sub> (Table 3).

**Effect of Substituting Arg for Lys at the P1 Position in CCK<sub>13-33</sub> on Cleavage by Yapsin 1.** To investigate whether an arginine residue in the P1 position of CCK<sub>13-33</sub> will be recognized as a cleavage site, yapsin 1 was incubated with CCK(P1 Arg). Surprisingly, four product peaks were generated from this substrate, suggesting cleavage at two different sites (Figure 5C). The presence of pepstatin A completely inhibited the formation of these products (data not shown), verifying that all products are a result of yapsin 1 cleavage. N-terminal sequence analysis of the generated products determined the cleavages to occur after Arg<sub>23</sub> and after Met<sub>25</sub>. On the basis of the relative amount of the four products generated (data not shown), it appears that yapsin 1 cleaves at the two positions without any preference of one over the other. Substitution of the P6 arginine with an alanine also resulted in the generation of four products after incubation of CCK(P6 Ala, P1 Arg) with yapsin 1 (Figure 5D). The cleavage sites were determined to be similar to those found for CCK(P1 Arg). A preference for cleavage after Met<sub>25</sub> was, however, found. To determine whether Met<sub>25</sub> is accommodated well in the S1 pocket, substitution



of this residue with alanine and leucine was performed. The peptides CCK(P3 Ala, P1 Arg) and CCK(P3 Leu P1 Arg) were cleaved by yapsin 1 after Arg<sub>23</sub> and Ala<sub>25</sub>/Leu<sub>25</sub> (Figure 5E and 5F). Like CCK(P1 Arg), CCK(P3 Ala, P1 Arg) was cleaved without a preference of one site over the other, whereas yapsin 1 cleaved CCK(P3 Leu P1 Arg) preferentially after Arg<sub>23</sub>. Incubation of 25  $\mu$ M CCK<sub>13–33</sub>, CCK(P1 Arg), CCK(P6 Ala, P1 Arg), CCK(P3 Ala, P1 Arg), and CCK(P3 Leu, P1 Arg) for 1 h with 0.21 pmol of yapsin 1 resulted in 29, 79, 83, 6, and 70% cleavage, respectively of the substrates (data not shown).

## DISCUSSION

Yapsin 1 is the model enzyme of a new subclass of aspartic proteases which differs from the traditional aspartic proteases by having specificity toward basic residues as opposed to the usual hydrophobic preference. To gain insight into the unique specificity of yapsin 1 for basic residues, modeling studies were carried out. The model revealed that yapsin 1 has a more open S1 pocket than other aspartic proteases due to a deletion of six residues in the yapsin 1 sequence between Met111 and Gly119, relative to rhizopuspepsin, thereby facilitating the accommodation of a large residue in this pocket (Figures 1 and 3C). The nature of the S1 pocket therefore suggests that a small neutral amino acid residue such as alanine would not fully occupy this pocket. Moreover, Asp77, which has been demonstrated to be very important for the cleavage of peptides with a lysine residue in the P1 position by rhizopuspepsin (43), is located on the end of a flap pointing down into the active-site cleft. Similarly, Asp77 is also likely to affect the binding at P1 in yapsin 1, which is close to the S1 pocket, although, slightly more than 4.0 Å. Indeed, in this study, we show that yapsin 1 cleaved CCK<sub>13–33</sub> and CCK(P1 Arg) preferentially at the basic residue cleavage site, and substitution of Lys<sub>23</sub> with Ala resulted in a substrate that was very poorly cleaved (Figure 5A).

Furthermore the model revealed that most of the other pockets within the active-site cleft of yapsin 1 were also highly electronegative (Table 2 and Figure 3), indicating that the enzyme not only prefers substrates with a basic residue in the P1 position but the catalytic efficiency will likely be enhanced for substrates with additional basic residues in the P6, P2', or P3' positions. Since S3, S1, and S2' are contiguous on one side, whereas S2, S1', and S3' are contiguous on the other side, interactions between residues forming one pocket and the substrate can influence binding at adjacent pockets. To test this prediction, the effects of deleting the arginine residue in the P6 position and introduction of an arginine in the P2', P3', and P6' position of CCK<sub>13–33</sub> were analyzed. Substituting the P6 Arg with an alanine resulted in a decrease of the cleavage efficiency by 50% (Table 3), similar to that observed for proalbumin peptides described by Ledgerwood et al. (44). In light of the yapsin 1 model, an alanine residue can be accommodated in the S6 pocket, but probably does not interact favorably with residues forming the S6 pocket. In contrast, an arginine residue can interact favorably with residues in the electronegative S6 pocket resulting in more efficient cleavage (Figure 3B).

Compared to CCK<sub>13–33</sub>, the cleavage efficiency of CCK-(P2' Arg) was almost 21-fold higher, indicating a strong preference for an arginine residue in the P2' position (Figure 3D). Similar results were also observed for a pro-albumin

peptide analogue, when an arginine was substituted for an Ala in the P2' position (44). Interestingly, introduction of a lysine residue in the P2' position resulted in 40% reduction of the cleavage efficiency compared to CCK<sub>13–33</sub>. The clear distinction of yapsin 1 between the binding of lysine versus arginine to this pocket could be due to the different properties of the two amino acids; arginine is less flexible and hydrophobic due to the presence of the guanadinium group, and it has the capability to form more hydrogen bonds than lysine. As shown in Table 3, the binding constant of CCK(P2' Arg) was 10-fold lower than for CCK(P2' Lys), supporting this interpretation. Introduction of a negatively charged residue in the P2' position, peptide CCK(P2' Asp), resulted in a 60% decrease in the catalytic efficiency of yapsin 1 for this substrate compared to CCK<sub>13–33</sub>, probably due to electrostatic repulsion of the aspartic acid residue from Asp37 in this subsite. Like yapsin 1, rhizopuspepsin has an aspartic acid in position 37 (Figure 1); however, a preference for an arginine residue over alanine in the P2' position is not observed (45). The high affinity for an arginine residue in the S2' subsite of yapsin 1 could therefore in addition to Asp37 be due to Glu127 and Glu129, not found in rhizopuspepsin.

As predicted by the model, an arginine residue in the P3' position was found to be favorable, possibly due to interaction with Asp292 in the S3' subsite. However, the effect of an arginine in the P3' position is not as favorable as in the P2' position; likewise, an aspartic acid is not as unfavorable in the P3' position as in the P2' position, indicating that an arginine residue in the P2' position has more influence on enhancing the catalytic efficiency of yapsin 1 than in the P3' position. Introduction of an arginine in the P6' position resulted in an almost 7-fold increase in the catalytic efficiency. Since this pocket is exposed to the solvent and there is no suitable crystal structure complex available, no attempt was made to model the S6' subsite.

The catalytic efficiency of yapsin 1 appears to be enhanced even more when the substrate has more than one arginine flanking the basic residue as predicted by the model. This is exemplified by the finding that the cleavage of CCK(P2' Arg, P6' Arg), which has a total of three arginine residues (in the P6, P2', and P6' positions) flanking the cleavage site, has an overall catalytic efficiency that was 23-fold better than CCK<sub>13–33</sub> due to higher binding affinity (lower  $K_m$ ; Table 3). This also explains why the cleavage of ACTH<sub>1–39</sub> by yapsin 1 occurs after the first of four basic residues (KKRR) (46) in high efficiency (25), since there is one arginine residue interacting with the S2' pocket and another with the S3' pocket.

Our present study also showed that an arginine at the P2' position not only had a strong effect in enhancing the catalytic efficiency but would also influence the primary specificity of yapsin 1. This was exemplified by the cleavage of CCK(P1 Arg) by yapsin 1. CCK(P1 Arg) was equally cleaved at two positions: Arg<sub>23</sub>, the usual cleavage site, and after Met<sub>25</sub> (Figure 5C). Furthermore, CCK(P6 Ala, P1 Arg) was preferentially cleaved after Met<sub>25</sub>, resulting in Arg<sub>23</sub> being in the P2' position (Figure 5D). Thus, the lack of the upstream P6 arginine residue which causes a loss of favorable interaction in the S6 pocket appears to have rendered yapsin 1 to prefer binding the substrate with the arginine residue in the P2' position. Substitution of Met<sub>25</sub> with leucine, in CCK-(P3 Leu, P1 Arg), also resulted in cleavage at Arg<sub>23</sub> and at

Leu<sub>25</sub> (Figure 5F), showing that the cleavage at the Met<sub>25</sub> or Leu<sub>25</sub> was due to the strong affinity of the S2' subsite for an arginine residue in competition with the S1 subsite for the same residue, rather than to a direct specificity for methionine in the P1 position per se. This is further supported by the result showing that CCK(P3 Ala, P1 Arg), which has the Met<sub>25</sub> substituted for an alanine, was cleaved after Ala<sub>25</sub>, albeit inefficiently (Figure 5E), due to the small size of this residue which interacts poorly with S1 pocket.

In conclusion, the kinetic and molecular modeling studies presented here have provided the first understanding of how the novel aspartic protease, yapsin 1, interacts with its substrate at the active-site cleft. The model presented shows a more open S1 pocket and an overall electronegative charge of several pockets in the active-site cleft, thus accounting for the unique accommodation of a basic residue at the P1 site and the enhancement of the catalytic efficiency for substrates having additional basic residues surrounding the P1. This information will be important for future studies on the search for natural substrates for this enzyme, and hence the biological function of yapsin 1, as well as facilitating the design of specific inhibitors for this class of aspartic proteases.

## ACKNOWLEDGMENT

We thank Dr. Alla Gustchina, NCI- Frederick Cancer Research and Development Center, for helpful comments and discussions on the yapsin 1 model.

## REFERENCES

- Loh, Y. P., Beinfeld, M. C., and Birch, N. P. (1993) in *Mechanisms of Intracellular Trafficking and Processing of Proproteins* (Loh, Y. P., Ed.) pp 179–224, CRC Press: Boca Raton, FL.
- Douglass, J., Civelli, O., and Herbert, E. (1984) *Annu. Rev. Biochem.* 53, 665–715.
- Loh, Y. P. (1984) *Annu. Rev. Neurosci.* 7, 189–222.
- Mains, R. E., Dickerson, I. M., May, V., Stoffers, D. A., Perkins, S. N., Ouafik, L., Husten, E. J., and Eipper, B. A. (1990) *Front. Neuroend.* 11, 52–89.
- van den Ouweland, A. M. W., van Duijnhoven, H. L. P., Keizer, G. D., Dorssers, L. C. J., and Van de Ven, W. J. M. (1990) *Nucleic Acids Res.* 18, 664.
- Smekens, S. P., Avruch, A. S., LaMendola, J., Chan, S. J., and Steiner, D. F. (1991) *Proc. Natl. Acad. Sci. U.S.A.* 88, 340–344.
- Smekens, S. P., and Steiner, D. F. (1990) *J. Biol. Chem.* 265, 2997–3000.
- Kiefer, M. C., Tucker, J. E., Joh, R., Landsberg, K. E., Saltman, D., and Barr, P. J. (1991) *DNA Cell Biol.* 10, 757–769.
- Seidah, N. G., Day, R., Hamelin, J., Gaspar, A., Collard, M. W., and Chrétien, M. (1992) *Mol. Endocrinol.* 6, 1559–1570.
- Nakayama, K., Kim, W.-S., Torii, S., Hosaka, M., Nakagawa, T., Ikemizu, J., Baba, T., and Murakami, K. (1992) *J. Biol. Chem.* 267, 5897–5900.
- Nakagawa, T., Hosaka, M., Torii, S., Watanabe, T., Murakami, K., and Nakayama, K. (1993) *J. Biochem.* 113, 132–135.
- Lusson, J., Vieau, D., Hamelin, J., Day, R., Chrétien, M., and Seidah, N. G. (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90, 6691–6695.
- Nakagawa, T., Murakami, K., and Nakayama, K. (1993) *FEBS Lett.* 327, 165–171.
- Seidah, N. G., Hamelin, J., Mamarbachi, M., Dong, W., Tadros, H., Mbikay, M., Chretien, M., and Day, R. (1996) *Proc. Natl. Acad. Sci. U.S.A.* 93, 3388–3393.
- Krieger, T. J., and Hook, V. Y. H. (1991) *J. Biol. Chem.* 266, 8376–8383.
- Devi, L., and Goldstein, A. (1984) *Proc. Natl. Acad. Sci. U.S.A.* 81, 1892–1896.
- Pierotti, A. R., Prat, A., Chesneau, V., Gaudoux, F., Leseney, A.-M., Foulon, T., and Cohen, P. (1994) *Proc. Natl. Acad. Sci. U.S.A.* 91, 6078–6082.
- Egel-Mitani, M., Flygenring, H. P., and Trier Hansen, M. (1990) *Yeast* 6, 127–137.
- Komano, H., and Fuller, R. S. (1995) *Proc. Natl. Acad. Sci. U.S.A.* 92, 10752–10756.
- Loh, Y. P., Parish, D. C., and Tuteja, R. (1985) *J. Biol. Chem.* 260, 7194–7205.
- Cawley, N. X., Pu, L.-P., and Loh, Y. P. (1996) *Endocrinology* 137, 5135–5143.
- Azaryan, A. V., Schiller, M., Mende-Mueller, L., and Hook, V. Y. H. (1995) *J. Neurochem.* 65, 1771–1779.
- Mackin, R. B., Noe, B. D., and Spiess, J. (1991) *Endocrinology* 129, 1951–1957.
- Cawley, N. X., Noe, B. D., and Loh, Y. P. (1993) *FEBS Lett.* 332, 273–276.
- Cawley, N. X., Chen, H.-C., Beinfeld, M. C., and Loh, Y. P. (1996) *J. Biol. Chem.* 271, 4168–4176.
- Cawley, N. X., Wong, M., Pu, L.-P., Tam, W., and Loh, Y. P. (1995) *Biochemistry* 34, 7430–7437.
- Bairoch, A., and Apweiler, R. (1996) *Nucleic Acids Res.* 24, 21–25.
- Johnson, M. S., Overington, J. P., and Blundell, T. L. (1993) *J. Mol. Biol.* 231, 735–752.
- Fitsch, W. M., and Margoliash, E. (1967) *Science* 155, 279–284.
- Sutcliffe, M. J., Haneef, I., Carney, D., and Blundell, T. L. (1987) *Protein Eng.* 1, 377–384.
- Sutcliffe, M. J., Hayes, F. R. F., and Blundell, T. L. (1987) *Protein Eng.* 1, 385–392.
- Suguna, K., Padlan, E. A., Bott, R., Boger, J., Parris, K. D., and Davies, D. R. (1992) *Proteins* 13, 195–205.
- Blundell, T. L., Jenkins, J. A., Sewell, B. T., Pearl, L. H., Cooper, J. B., Tickle, I. J., Veerapandian, B., and Wood, S. P. (1990) *J. Mol. Biol.* 211, 919–941.
- Sali, A., and Blundell, T. L. (1990) *J. Mol. Biol.* 212, 403–428.
- DeAlwis, C. G., Frazzo, C., Badasso, M., Cooper, J. B., Tickle, I. J., Driessen, H., Blundell, T. L., Murakami, K., Miyazaki, H., Sueirasdiaz, J., Jones, D. M., and Szelke, M. (1994) *J. Mol. Biol.* 236, 342–360.
- Jones, T. A., and Thirup, S. (1986) *EMBO J.* 5, 819–822.
- Abola, E. E., Bernstein, F. C., Bryant, S. H., Koetzle, T. F., and Weng, J. (1987) in *Crystallographic Databases—Information Content, Software Systems, Scientific Applications* (Allen, F. H., Bergerhoff, G., and Sievers, R., Eds.) pp 107–132, Data Commission of the International Union of Crystallography, Bonn, Germany.
- Bernstein, P. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr, Brice, M. D., Rogers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1977) *J. Mol. Biol.* 112, 535–542.
- Lineweaver, H., and Burk, D. (1934) *J. Am. Chem. Soc.* 56, 568–566.
- MacKay, V. L., Welch, S. K., Insley, M. Y., Manney, T. R., Holly, J., Saari, G. C., and Parker, M. L. (1988) *Proc. Natl. Acad. Sci. U.S.A.* 85, 55–59.
- Guruprasad, K., Tormakangas, K., Kervinen, J., and Blundell, T. L. (1994) *FEBS Lett.* 352, 131–136.
- Nichols, A., and Honig, B. (1991) *J. Comput. Chem.* 12, 435–445.
- Lowther, W. T., Majer, P., and Dunn, B. M. (1995) *Protein Sci.* 4, 689–702.
- Ledgerwood, E. C., Brennan, S. O., Cawley, N. X., Loh, Y. P., and George, P. M. (1996) *FEBS Lett.* 383, 67–71.
- Balboa, M., Cunningham, A., and Hofmann, T. (1993) *Arch. Biochem. Biophys.* 306, 297–303.
- Azaryan, A., Wong, M., Friedman, T. C., Cawley, N. X., Estivariz, F. E., Chen, H.-C., and Loh, Y. P. (1993) *J. Biol. Chem.* 268, 11968–11975.
- Cawley, N. X., Olsen, Y., Zhang, C.-F., Chen, H.-C., Tan, M., and Loh, Y. P. (1998) *J. Biol. Chem.* 273, 584–591.