THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 277, No. 30, Issue of July 26, pp. 27273–27281, 2002

Printed in U.S.A.

The Structure of Human Prokallikrein 6 Reveals a Novel Activation Mechanism for the Kallikrein Family*

Received for publication, February 14, 2002, and in revised form, May 16, 2002
Published, JBC Papers in Press, May 16, 2002, DOI 10.1074/jbc.M201534200

F. Xavier Gomis-Rüth‡§, Álex Bayés¶, Georgia Sotiropoulou‖, Georgios Pampalakis‖, Theodoros Tsetsenis‖, Virtudes Villegas¶, Francesc X. Avilés¶, and Miquel Coll‡**

From the ‡Institut de Biologia Molecular de Barcelona, Consejo Superior de Investigaciones Científicas, c/Jordi Girona, 18-26, Barcelona 08034, Spain, ¶Institut de Biotecnologia i de Biomedicina, Departament de Bioquímica i Biologia Molecular, Unitat de Ciéncies, Universitat Autònoma de Barcelona, Bellaterra, Barcelona 08193, Spain, and the ‖Department of Pharmacy, School of Health Sciences, University of Patras, Rion, Patras 265 00, Greece

Zyme/protease M/neurosin/human kallikrein 6 (hK6) is a member of the human kallikrein family of trypsin-like serine proteinases and was originally identified as being down-regulated in metastatic breast and ovarian tumors when compared with corresponding primary tumors. Recent evidence suggests that hK6 may serve as a circulating tumor marker in ovarian cancers. In addition, it was described in the brain of Parkinson's disease and Alzheimer's disease patients, where it is implicated in amyloid precursor protein processing. It is thus a biomarker for these diseases. To examine the mechanism of activation of hK6, we have solved the structure of its proform, the first of a human kallikrein family member. The proenzyme displays a fold that exhibits chimeric features between those of trypsinogen and other family members. It lacks the characteristic "kallikrein loop" and forms the six disulfide bridges of trypsin. Pro-hK6 displays a completely closed specificity pocket and a unique conformation of the regions involved in structural rearrangements upon proteolytic cleavage activation. This points to a novel activation mechanism, which could be extrapolated to other human kallikreins.

The kallikreins constitute a subfamily of mammalian serine proteases, originally defined as enzymes cleaving vasoactive peptides (kinins) from plasma α_2-globulin substrates called kininogens (1). Their name was coined by Werle and co-workers (2) in the early 1930s while analyzing pancreatic fluids (from Greek "kallikreas," pancreas). Initially, these enzymes were subdivided into two subgroups. One was formed by plasma kallikreins, high molecular mass glycoproteins present in blood plasma, and the other comprised glandular, organ, and tissue kallikreins, medium-sized (~30–50-kDa) glycoproteins mainly occurring in pancreas, salivary glands, saliva, kidney, and urine (1, 3, 4). Recently, kallikrein gene expression has been detected in an increasing number of organs, fluids, and tissues (*e.g.* central nervous system, glands (thyroid, thymus), prostate, seminal fluid, testis, ovary, uterus, colon, skin, heart, breast, lung, and trachea), many of them subjected to hormone regulation (5). The enzymes described to date form a large family in rodents (6, 7), among them neuropsin (KLK8[1]; Ref. 8), tonin (KLK2; Ref. 9), kallikrein 13 (KLK13; Ref. 10), and 7 S γ-nerve growth factor (KLK3; Ref. 11). In humans they currently encompass at least 15 genes (5), ranging from the classic enzymes, tissue kallikrein (KLK1; Ref. 12), glandular kallikrein (KLK2; Ref. 13), and prostate-specific antigen (PSA; KLK3; Ref. 14) to the newly discovered gene products, like normal epithelium-specific 1 gene (KLK10; Ref. 15), zyme/protease M/neurosin/human kallikrein 6 (KLK6; Ref. 16), and prostase (KLK4; Ref. 17).

Kallikreins are secreted endopeptidases of ~230 amino acid residues (the classic chymotrypsin numbering (18) will be used throughout; see Fig. 1), initially synthesized as preproenzymes. These are proteolytically processed to render proenzymes by removal of the signal peptide, followed by activation to the mature forms. This proceeds equally by limited proteolysis, normally targeted to an arginine/lysine-valine/isoleucine peptide bond, in the manner characteristic for serine proteases (19). Kallikreins display mostly trypsin-like specificity, thus hydrolyzing peptide bonds preferentially after lysine and arginine residues of oligopeptide substrates (20). Unlike trypsin, however, they do not display significant activity toward larger proteins. This specificity makes them candidate autoactivators or activating enzymes for other prokallikreins, putatively in cascade reactions similar to those described for blood coagulation, digestion, fibrinolysis, and apoptosis.

* This work was supported in part by Ministerio de Educación y Cultura and Ministerio de Ciencia y Tecnología, Spain, Grants PB98-1631, BIO2000-1659, BIO98-0362, BIO2001-2046, and 2FD97-0518. Grants 1999SGR188 and 2001SGR346 and the Centre de Referència en Biotecnologia (all from the Generalitat de Catalunya) further funded this work. Diffraction data collection as performed at EMBL and BM-14-S protein crystallography beamlines (ESRF, Grenoble) was supported financially by the ESRF and by European Union Grant HPRI-CT-1999-00022 (awarded to the EMBL). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

This work is dedicated to Prof. Wolfram Bode on the occasion of his 60th birthday to acknowledge his hallmark contribution to the structural biology of serine proteinases over the past 3 decades.

The atomic coordinates and structure factors (code 1gvl) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

§ To whom correspondence should be addressed. Tel.: 3493-400-61-44; Fax: 3493-204-59-04; E-mail: xgrcri@ibmb.csic.es.

** Supported by a NATO collaborative research grant (contract 973152) and by the "K. Karatheodoris" grant awarded by the Research Committee of the University of Patras (contract 1950).

[1] The abbreviations used are: KLK*x*, kallikrein-family gene ascription number *x* according to Ref. 5 and the human gene nomenclature committee (www.gene.ucl.ac.uk/nomenclature/); hK6, human kallikrein 6; PDB, Protein Data Bank; PSA, prostate-specific antigen; S_1, S_2, etc. and P_1, P_2, etc., protease active site cleft and substrate subsites, respectively, N-terminal of the scissile peptide bond, S_1', S_2', etc. and P_1', P_2', etc., protease active site cleft and substrate subsites, respectively, at the C terminus of the scission, in accordance with Ref. 64; SP, Swiss-Prot/TrEMBL sequence database accession number (www.expasy.org/sprot); RF, rotation function; CC, correlation coefficient in amplitudes.

In humans, tissue kallikrein gene family members consist of five coding exons and map in tandem to the same chromosomal locus (19q13.3–13.4) (5, 21), which shows genetic linkage to Alzheimer's disease (22) and which is nonrandomly rearranged in a variety of human solid tumors, like pancreatic carcinomas, astrocytomas, ovarian cancers, and thyroid tumors, among others (23–25). Distinct members of the kallikrein family have been associated with cancer, and two members, PSA and glandular kallikrein, are used as biomarkers for the diagnosis and monitoring of prostate and breast cancer, whereas kallikrein 11 has been assessed for ovarian carcinoma (26).

Human kallikrein 6 (hK6) is a member of the kallikrein gene family, and it is also known as protease M, neurosin, or zyme (16, 27, 28), a 223-amino acid serine proteinase with trypsin-like activity and restricted substrate specificity. It derives from a 244-amino acid precursor containing a 16-residue presignal peptide and a 5-amino acid activation propeptide and is glycosylated at Asn^{132}. This protein may have orthologues in mouse, rat, and monkey (28–31). It is distinct from neuropsin (KLK8), a hippocampal protease involved in kindling epileptogenesis in mice (32). Like the other members of the human kallikrein family, the KLK6 gene lies at chromosome locus 19q13.4, telomeric to the PSA (KLK3) and KLK2 genes (5, 21). (Pro)hK6 or its mRNA was found in human brain, spinal cord, cerebellum, kidney, uterus, mammary gland, pancreatic tissue, cerebrospinal fluid, prostate and salivary glands and, at a lower level, in spleen (16, 27, 28, 33, 34). Immunohistochemical localization of hK6 in various tissues, showing epithelial cell secretion, has been published (35). It has also been detected in pathological tissues, as human colon adenocarcinoma and ovarian and breast cancer. In the metastatic breast tumor cell line BT474, estrogens and progestins up-regulate the expression of hK6 in a dose-dependent manner (36). The connection between breast and prostate cancer and steroid hormones has been established in many epidemiological studies (37). hK6 was shown to be down-regulated in the metastatic cells but strongly expressed, although less so than in normal tissues, in primary breast cancer cells, ovarian cancer tissue, and tumor cell lines, whereas no expression was detected in normal ovary (16). Plasminogen was suggested as a putative target in the search for physiological substrates of hK6. Proteolytic cleavage of plasminogen *in vitro* resulted in the generation of angiostatin and angiostatin-like fragments. Angiostatin is an endogenous inhibitor of tumor angiogenesis and metastatic growth. Therefore, this expression pattern suggests that hK6 contributes to several cancers in establishing breast and ovarian primary tumors and that it may function later during tumor progression as a potential metastasis and angiogenesis inhibitor (16, 38).

The protein has also been found in adult but not in fetal brain. It is decreased in Parkinson's disease and Alzheimer's disease patients (33). Cells transfected with the enzyme and amyloid precursor protein accumulate amyloidogenic fragments (28). Accordingly, hK6 has been implicated in the progression of Alzheimer's disease. This activity of hK6 in brain is consistent with the finding that its mRNA is induced in the perilesioned region upon injury of the central nervous system (39). These results and the fact that hK6 has been found in significantly increased concentrations in the serum of ovarian cancer patients and in the cerebrospinal fluid, plasma, and blood of Alzheimer's disease patients have led to the development of an immunofluorometric assay for the measurement of this new biomarker for breast and ovarian cancer and for Alzheimer's disease (40–42).

hK6 is a serine proteinase expressed as an inactive zymogen to prevent proteolytic activity in the wrong environment. Upon adequate stimulation, limited proteolysis is induced, leading to proenzyme activation. For the first time, we have analyzed the proform of a member of the (human) kallikrein family by x-ray crystallography, since no three-dimensional structure of a prokallikrein was available. We present the essential features of the structural determinants for activation of human pro-hK6, which can be extrapolated to the other 14 members of the family.

TABLE I
Data collection and refinement statistics

Parameter	Value
Space group	$P2_12_12_1$
Cell constants (Å)[a]	46.0, 66.4, 71.2
Wavelength (λ in Å)	0.934
No. of measurements	64,722
No. of unique reflections	19,445
Whole resolution range (Å)	38.6 to 1.80
Completeness (%)/R_{merge} (%)[b]	94.0/9.5
Average intensity ($\langle I\rangle/\langle\sigma(I)\rangle$)/average multiplicity	4.6/3.3
Last resolution shell (Å)	1.90 to 1.80
Completeness (%)/R_{merge} (%)[b]	90.1/43.6
Average intensity ($\langle I\rangle/\langle\sigma(I)\rangle$)/average multiplicity	1.7/3.0
B factor (Wilson plot) (Å^2)	20.0
Resolution range for refinement (Å)	33.2 to 1.80
No. of reflections (working set/test set)	18,024/1,418
Crystallographic R_{factor} (free R_{factor})[c]	18.5 (22.5)
No. of protein atoms (active/inactive)	1,591/125
No. of solvent molecules	158
Root mean square deviation from target values	
Bonds (Å)	0.013
Angles (degrees)	1.39
Bonded *B*-factors (Å^2)	3.63
Average *B*-factors for protein/solvent atoms (Å^2)	21.1/30.0

[a] Values for *a*, *b*, and *c*.

[b] $R_{merge} = \{\Sigma_{hkl}\Sigma_i|I_i(hkl) - \langle I(hkl)\rangle|/\Sigma_{hkl}\Sigma_i I_i(hkl)\} \times 100$, where $I_i(hkl)$ is the *i*th measurement of reflection *hkl*, including symmetry-equivalent ones, and $\langle I(hkl)\rangle$ is its mean intensity.

[c] $R_{factor} = \{\Sigma_{hkl}||F_o| - k|F_c||/\Sigma_{hkl}|F_o|\} \times 100$, with F_o and F_c as the observed and calculated structure factor amplitudes, calculated for the working set reflections; free R_{factor}, same for a test set of 7% reflections (>500) not used during refinement.

EXPERIMENTAL PROCEDURES

Protein Preparation—Initial trials to produce homogenous protein including the residues of the chemical sequence of the hK6 precursor Ala^{10}–Lys^{245} (chymotrypsin numbering (18); corresponds to Ala^{16}–Lys^{244} of the sequential protein numbering (see Fig. 1)) failed due to autolysis after Arg^{76}.[2] Accordingly, this residue was mutated to glutamine. The *N*-glycosylation site at Asn^{132} prevented the protein from crystallizing, so it was also replaced by glutamine. Finally, a further mutation occurred in the activation peptide region during PCR cloning (Q13R). This construct, harboring three mutations, was used to express the recombinant proteins used in the present study. The protein was heterologously overexpressed in yeast following an approach recently reported for other proproteases (43). Briefly, the vector pPIC9 containing the mutant gene was transformed into *Pichia pastoris* strain KM71 by the spheroblast method. After expression at 30 °C, the culture supernatant containing pro-hK6 was separated by centrifugation and subjected to two steps of purification (hydrophobic and anion-exchange fast protein liquid chromatography). A final size exclusion chromatography step resulted in highly purified protein, finally concentrated to 10 mg/ml (in 10 mM Tris-HCl, pH 8.5, 10 mM benzamidine). The presence of a single continuous protein chain was assessed by mass spectrometry and automated Edman degradation analysis of the N-terminal sequence. Pro-hK6 was activated to render fully active mature hK6 by trypsin and lysyl endopeptidase cleavage; the active enzyme was tested for activity against the synthetic substrate N^{α}-benzoyl-L-arginine ethyl ester (Sigma). This shows that our pro-hK6 form is not a nonactivable truncated species, as described in the form of subunit III for the serine protease proenzyme proprotease E (44).

Crystallization and Data Collection—The protein was crystallized using the sitting drop vapor diffusion method from equivolumetric drops of protein and precipitant solution (30% polyethylene glycol 4000, 0.1 M Tris-HCl, pH 8.5, 0.2 M $MgCl_2$). Well shaped orthorhombic crystals

[2] T. Tsetsenis, unpublished results.

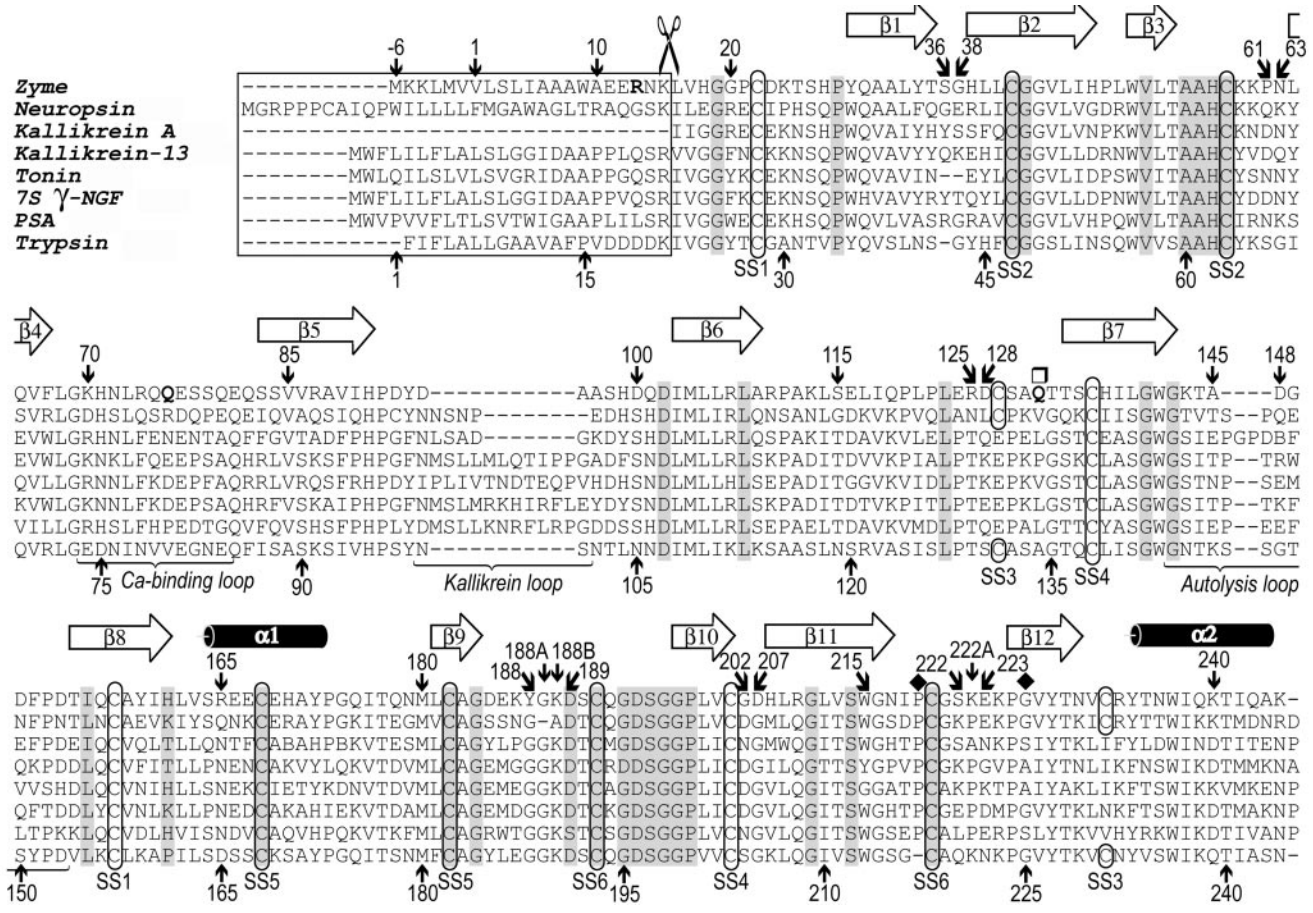

FIG. 1. **Primary and secondary structures of pro-hK6.** Shown is a structure-based sequence alignment of members of the kallikrein family of known three-dimensional structure, together with PSA and trypsinogen. Displayed are the zymogen forms (pre and/or pro sequences are *boxed*) of hK6 (this work; SwissProt/TrEMBL (SP) sequence access code Q92876; kallikrein family ascription number KLK6 (3, 5)), mouse-brain neuropsin (PDB accession code 1npm; SP Q61955; KLK8 (32)), mouse glandular kallikrein 13 (PDB 1ao5; SP P36368; KLK13 (10)), rat glandular tonin (PDB 1ton; SP P00759; KLK2 (9)), mouse neuronal 7S γ-nerve growth factor (PDB 1sgf; SP P00756; KLK3 (77)), human PSA (SP P07288; KLK3), and bovine pancreatic trypsin (PDB 1tgb; SP P00760 (78)). The sequence of mature porcine pancreatic kallikrein A (PDB 2pka; SP P00752; KLK1 (67)), for which the sequence upstream of the activation cleavage site has not been reported, is also shown. This protein has been chemically sequenced, so that *B* in the sequence stands for aspartate/asparagine. The sequential pro-hK6 numbering is shown at the *bottom*, and the chymotrypsinogen-based numbering (18), used throughout, is on *top*. The three introduced amino acid substitutions (Q13R/R76Q/N132Q) in this sequence are shown in *boldface type*. Regular secondary structure elements correspond to pro-hK6; the *arrows* represent β-strands, and *rods* represent α-helices. The cysteine residues engaged in the formation of up to six disulfide bonds are *boxed* and marked with labels (*SS1–SS6*). The 29 "invariant" residues (70) in serine proteinases and the residue in the bottom of the S_1 specificity pocket are shown with a *gray background*, characteristic residues for kallikreins (67) appear with a *rhombus*, the *N*-glycosylation site of hK6 is shown with a *square*, and the limited proteolysis activation site is indicated by *vertical scissors*. According to the chymotrypsinogen numbering, this cleavage occurs between residues Arg/Lys[15] and Ile/Leu/Val[16].

belonging to space group $P2_12_12_1$ (a = 46.0 Å, b = 66.4 Å, c = 71.2 Å) were obtained. These crystals diffract to better than 1.8-Å resolution and harbor one monomer in the asymmetric unit and 43% solvent (V_m = 2.2 Å^3/Da (45)). A cryo-protecting solution consisted of mother liquor with 20% glycerol. Crystals briefly soaked in this solution were immediately flash cryo-cooled in liquid nitrogen and stored for diffraction analysis. Diffraction data collection was performed at European Synchrotron Radiation Facility beamlines (Grenoble, France) with 165-mm marCCD and ADSC Quantum4 CCD detectors. Data were processed with MOSFLM version 6.11b (46) and scaled, merged, and reduced with SCALA (47). Table I provides a summary for the data collection and processing.

Structure Solution and Refinement—A preliminary search for sequence similarity of pro-hK6 within the Protein Data Bank revealed the closest relationship with murine neuropsin (PDB 1npm (32)). Accordingly, these coordinates were used as a searching model to solve the structure by molecular replacement, using the stand alone version of AMoRe (48). These calculations, performed in the 15 to 4.0-Å resolution range, confirmed $P2_12_12_1$ as the correct space group. The rotation function (RF) showed a clear peak at Eulerian angles α = 26.7°, β = 87.5°, and γ = 74.0°. The correlation coefficient in amplitudes (CC (48)) for this solution equals 3.7%, the crystallographic R_{factor} (see Table I for a definition; computed after data expansion to P_1) equals 56.9%, and RF = 9.5 (second highest solution: CC = 1.5%, R_{factor} = 57.2%, RF = 8.0). The translation function also rendered a clear solution (26.7, 87.5, 74.0, 0.3396, 0.0569, 0.2789; α, β, γ in degrees, and x, y, z in fractional cell coordinates; CC = 17.1%; R_{factor} = 53.8%; second highest peak CC = 11.2%, R_{factor} = 56.2%). A final rigid body refinement with FITING (48) led to better figures, CC = 20.0% and R_{factor} = 52.9% (28.3, 87.7, 72.0, 0.3386, 0.0553, and 0.2773). The input coordinates were rotated and translated according to this solution and subjected to a simulated annealing refinement step and positional/temperature factor refinement using maximum likelihood as a criterion after performing bulk solvent correction and anisotropic temperature factor refinement (R_{factor} = 41.3%; free R_{factor} = 44.2%). Program CNS version 1.0 (49) was used for this purpose. The subsequently computed σ_A-weighted $2mF_o - DF_c$ and $mF_o - DF_c$ electron density maps, inspected together with the initially refined model on a SGI Graphic Work station with the program TURBO-FRODO (50), showed the correctness of the solution and positive difference density for the regions diverging from the phasing model. At this stage, an electron density/model improvement step was undertaken with ARP/wARP (51) using all data up to 1.8-Å resolution and the warpNtrace-mode with protein sequence information. This calculation rendered excellent electron density maps (see Fig. 2, *a* and *b*) and a model consisting of five chains and 185 residues identified within the sequence. In successive cycles, this model was improved by alternating manual model building and crystallographic refinement with CNS and refmac5 (52) under translations and liberations of pseudo-rigid bodies

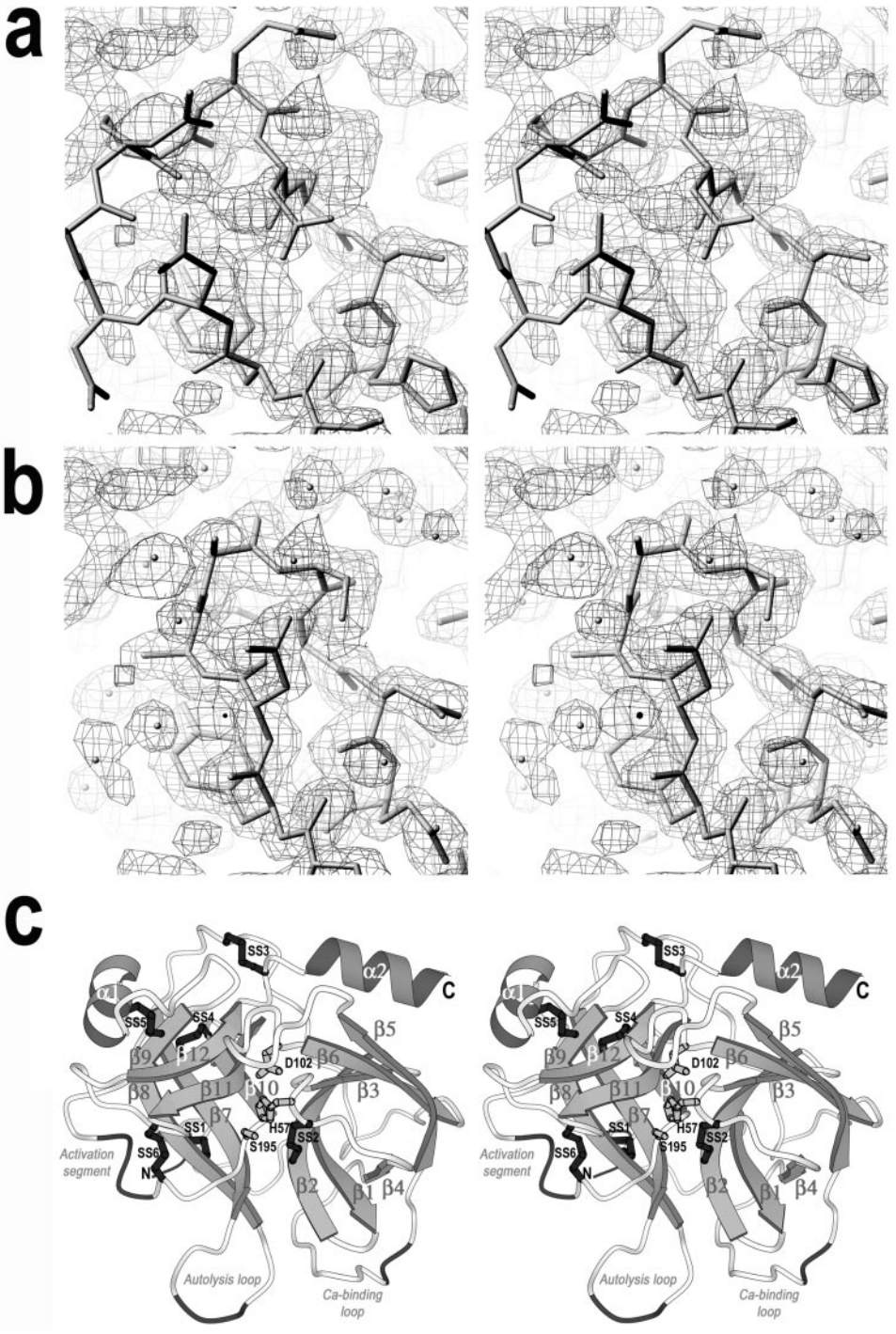

FIG. 2. **Structure solution and pro-hK6 topology.** *a*, stereo schematic diagram displaying the σ_A-weighted $2mF_o - DF_c$ electron density map obtained after ARP/wARP refinement, superimposed with the model used for initial phasing, contoured at 1 σ above average. *b*, same as *a*, with the final refined model superimposed. *c*, stereo ribbon plot of pro-hK6 shown in the traditional serine proteinase standard orientation (79) (*i.e.* looking into the active site cleft) with a bound substrate running from the *left* (nonprimed subsites prior to the scissile peptide bond; Ref. 64) to *right* (primed subsites). The regular secondary structure elements are displayed as *arrows* (β-strands) and *ribbons* (α-helices) and labeled (β_1–β_{12} and α_1–α_2). The side chains of the residues of the catalytic triad (*light gray*) and the six disulfide bonds (*dark gray*; *SS1–SS6*) are also shown as *stick models* and labeled. The N and C termini and the positions of characteristic structural loops are also indicated, and poorly defined and undefined main-chain stretches (see "Experimental Procedures") are shown as *dark gray coils*.

(TLS) refinement (53). In the final stages, solvent molecules were introduced where appropriate, if present in $2mF_o - DF_c$ and $mF_o - DF_c$ maps contoured at 1 and 2.5 σ, respectively. The final model comprises the residues ranging from Lys^{15} to Ala^{244} of the chemical sequence of pro-hK6 and 158 solvent molecules (labeled W 501 to W 658). As observed in other serine proteinase zymogens, chain segments corresponding to parts of the calcium-binding loop (in particular, Ser^{78}–Gln^{80}), autolysis loop (Asp^{148}–Phe^{151}), and base-forming segment (Tyr^{188}–Ser^{190}) are flexible. Together with the first 3 residues (Lys^{15}–Val^{17}), they have been tentatively traced based on very weak electron density and set to zero occupancy. All active residues lie within the allowed regions of a Ramachandran plot as determined with PROCHECK (54). The peptide bond Ile^{218}-Pro^{219} is in the *cis*-conformation. Table I provides a summary of the final refinement statistics.

Miscellaneous—Structure superimpositions were calculated with TURBO-FRODO (50), LSQKAB of the CCP4 suite (55), and LSQMAN of the RAVE package (Uppsala Software Factory; on the World Wide Web at alpha2.bmc.uu.se/~gerard/manuals). Structure and sequence simi-

larity searches were performed with the DALI server at www.ebi.ac.uk/dali (56) and the CE (Combinatorial Extension) server at cl.sdsc.edu/ce.html (57). Structural classifications are according to SCOP at scop.mrc-lmb.cam.ac.uk/scop. Figures were prepared with TURBO-FRODO and MOLSCRIPT (58). Model quality was assessed with PROCHECK (54). The final coordinates of pro-hK6 have been deposited with the Protein Data Bank (accession code 1gvl).

RESULTS AND DISCUSSION

Overall Structure—Pro-hK6 displays a frustum-like shape with approximate overall dimensions of 50 × 45 × 30 Å with a larger flat frontal surface (according to the standard orientation; see Fig. 2*c*) and a smaller parallel surface at the back, where the N and C termini are located. It is composed of two hydrophobically interacting domains of about 120 residues, the N-terminal one ranging approximately until Leu^{121} and further including Tyr^{234}–Ala^{244} at the C terminus, and a C-terminal domain from Pro^{122} to Arg^{233}, both characteristic for the serine proteinase fold (59, 60). Each domain features a six-stranded antiparallel β-barrel, consisting of two intercalating Greek key motifs sharing the two central strands (see Fig. 2*c*), and an α-helix (α_1 belongs to the C-terminal domain and α_2 to the N-terminal one). Strands β_1–β_6 make up the N-terminal β-barrel (β-barrel 1), and β_7–β_{12} make up the C-terminal one (β-barrel 2). The longitudinal axis of β-barrel 1 is approximately perpendicular to the frontal frustum back surface, whereas β-barrel 2 is almost parallel. At the interface between the barrels on the frontal surface, a shallow depression marks the active site cleft (see below).

The polypeptide chain is covalently cross-linked by six disulfide bridges (SS1–SS6; see Figs. 1 and 2*c*). SS1 anchors the N-terminal segment to the C-terminal domain (at β_8), and SS2 connects strand β_2 with loop $\beta_3\beta_4$, harboring one of the catalytic residues, His^{57} (see below), and encompassing a single helical turn between Ala^{55} and Cys^{58}. SS3–SS6 are located in the C-terminal domain. In particular, SS3 anchors loop $\beta_{12}\alpha_2$ to the segment connecting both domains on the back surface, SS5 fixes α_1 to the flank of the β-barrel moiety and SS6 links the activation segment undergoing major conformational changes upon activation (see below) with the top of the C-terminal barrel.

The N-terminal segment on the surface protrudes from the molecular body (see Fig. 2*c*) at the back frustum surface and is only visible in electron density due to inter-main-chain interactions between Lys^{15}–Val^{17} and His^{91}-Pro^{92} of a symmetry-related molecule, as observed in the structure of proprotease E (61). The main chain enters the molecular body at His^{27}, leading to the first strand of β-barrel 1. Embedded between β_4 and β_5 of this barrel and similarly to other known serine proteinases, a surface-exposed cation-lacking calcium-binding loop (62) Lys^{70}–Gln^{80} is placed, folded back on top of its flanking strands β_4 and β_5 and approaching loop $\beta_1\beta_2$. The tip of Lys^{70} occupies the approximate position of a calcium ion in trypsin, which displays metal-coordinating glutamate residues at positions 70 and 80 (see Fig. 1). This loop is fully ordered, except for the central residues Ser^{78}–Gln^{80}, which are flexible. In this region, at the *lower right* part of the molecule (see Fig. 2*c*), an autolysis-sensitive point is located after Arg^{76}, which when processed leads to an inactive species. This cleavage was observed during the purification of the protein and may have (auto)regulatory functions as observed in chymotrypsinogen, which is cut during activation at its autolysis loop beside the primary activation point (63). In order to obtain stable, proteolytically unprocessed protein, this residue was mutated to glutamine in the present study (see "Experimental Procedures" and Fig. 1).

After leaving β-barrel 1, a 26-residue linker (Ala^{109}–Thr^{134}) running across the smaller back frustum surface leads to β-barrel 2. Before entering its first strand, however, the main chain describes a wide loop, passing over β-hairpin $\beta_{10}\beta_{11}$ and harboring at position Asn^{132} (mutated to glutamine in the present study) an *N*-glycosylation site that accounts for additional 9 kDa in the molecular mass of the native protein (16). After the first strand of β-barrel 2 and spatially adjacent to the calcium-binding loop, in the *middle lower part* of the molecule (Fig. 2*c*), the surface-located autolysis loop, so called again by similarity with other known serine proteinases, is found, running from Gly^{142} to Asp^{153}. This loop is also fully defined, except for the central three residues Asp^{148}–Phe^{151} (note that the experimentally determined autolysis-sensitive point is not in the autolysis loop in pro-hK6). The main chain then enters strand β_8, and at its exit site, again on the molecular surface, it features helix α_1 (Ser^{164}–Tyr^{172}), which, after passing through β_9, enters the active site and substrate-binding site base-forming segment Asp^{189}-Ser^{195}. This segment harbors the connecting disulfide bridge SS6 and is flexible from Tyr^{188} to Ser^{190}. After hairpin $\beta_{10}\beta_{11}$ of the barrel, the main chain enters section Gly^{216}–Tyr^{228}, including the entrance frame-forming segment Val^{213}–Cys^{220} and the back-forming region Gly^{226}–Tyr^{228}. It is located on the front surface and anchored via SS6 to the previously mentioned base-forming segment. This region is involved in shaping the S_1 pocket (see below). Following this and after passing through strand β_{12}, the main chain runs over β-barrel 1, reaching the opposite surface and the C terminus in the form of helix α_2, which packs against the strands of the second Greek key motif of β-barrel 1.

Active Site Cleft and Specificity Pocket of Pro-hK6—The essential residues shaping the active site and the specificity pocket (S_1 site (64)) are provided by the loops connecting the third and fourth and the fifth and sixth β-strands of each barrel on the frustum front. The active site cleft is centered on the catalytic residue Ser^{195}, which together with residues His^{57} and Asp^{102} establishes the charge-relay system characteristic of serine proteinases. His^{57} and Asp^{102} are provided by loops $\beta_3\beta_4$ and $\beta_5\beta_6$, respectively, of β-barrel 1, whereas Ser^{195} is presented to the active site cleft by the end of the activation segment prior to strand β_{10} of β-barrel 2 (see Fig. 2*c*).

Substrate binding is achieved in mature serine proteinases by the extended segment Ser^{214}–Gly^{216} (see Fig. 4*b*), which binds the peptide main chain via its own main chain, generating an antiparallel β-ribbon. The first two residues are in the "active" extended conformation, but the chain deviates at Gly^{216}, turning back to cover the S_1 pocket. The oxyanion hole, which stabilizes the negatively charged transition state reaction intermediate, is provided by atoms Gly^{193} N and Ser^{195} N, although only the latter is in the position probably required in the active enzyme.

The active site and S_1 pocket base-forming segment, running from Asp^{189} to Ser^{195}, is folded outward, occupying the space of the autolysis loop and the enzyme moiety-penetrating N terminus in mature enzymes (see below). This drags region Asn^{217}–Ser^{222} inward by means of SS6, thus blocking the S_1 pocket via the side chain of Ile^{218}, linked to Pro^{219} by a *cis*-peptide bond. The S_1 site of pro-hK6 is mainly determined in its depth by Gly^{216} and Gly^{226} that form a narrow slot-like pocket that can receive large substrate side chains in P_1 in the mature enzyme. It is shaped at its bottom by the side chains of Val^{213} and Tyr^{228} and by the main-chain segments Trp^{215}–Cys^{220} and Asp^{189}–Gln^{192}. Asp^{189}, residing at the bottom of the pocket in mature enzymes (see below), is flexible in our proenzyme structure but will account for the trypsin-like preference for lysine and arginine residues in P_1. Asp^{194}, which upon activation will become engaged with the newly formed N terminus in an internal salt bridge, is folded outward and interacts with Trp^{141} N and is

TABLE II
Structural similarity between human pro-hK6 and other serine (pro)proteinases

Protein	PDB code	Root mean square deviation	Sequence identity	Aligned residues/ total size	*Z* score[a]
		Å	%		
Mouse hippocampal neuropsin	1npm	2.7	49	209/225	7.0
Bovine pancreatic trypsinogen	1tgn	2.8	44	211/229	7.0
Bovine pancreatic chymotrypsinogen A	2cga	2.5	38	204/245	7.0
Porcine pancreatic kallikrein	2pka	2.6	44	207/232	6.9
Mouse glandular kallikrein 13	1ao5	2.2	37	196/237	6.5
Rat submaxillary gland tonin	1ton	3.5	39	200/235	5.9
Mouse 7 S γ-nerve growth factor	1sgf	3.7	37	205/233	5.2

[a] As defined in Ref. 57. Ordered following the *Z* score and then sequence identity.

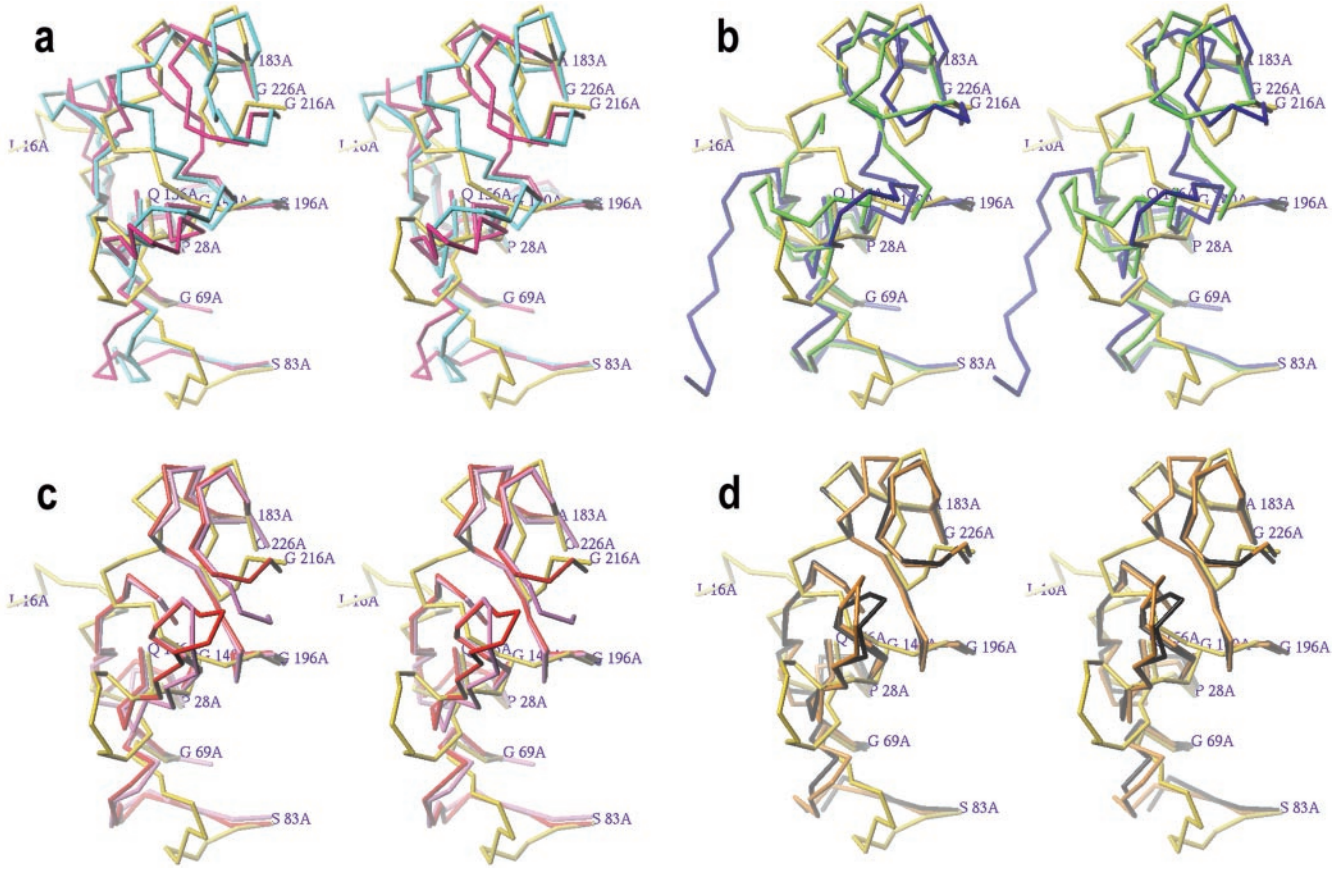

FIG. 3. **Unique structural features in pro-hK6.** Shown are superimpositions of the C-α traces of the protein segments corresponding to the N-terminal region, the calcium-binding loop, the autolysis loop, the activation segment around Asp^{189}–Gly^{196}, and segment Gly^{216}–Gly^{226} of pro-hK6 (*yellow sticks*) with distinct serine proproteases chymotrypsinogen A (PDB 2cga; *magenta sticks*) and chymotrypsinogen C (PDB 1pyt; *cyan*) (*a*) and trypsinogen (PDB 1tgn; *green*) and proprotease E (PDB 1pyt; *dark blue*) (*b*) and with closely related active serine proteinases kallikrein A (PDB 2pka; *red*) and tonin (PDB 1ton; *pink*) (*c*) and neuropsin (PDB 1npm; *black*) and trypsin (PDB 2ptn; *orange*) (*d*). In all cases, the orientation corresponds to that of Fig. 2*c*, further rotated clockwise ~90°. Selected pro-hK6 residues are labeled.

trapped by the end of loop $\beta_1\beta_2$ and the beginning of strand β_2 in a similar manner as other zymogens like chymotrypsinogen A and C and proproteinase E (61, 65). Unlike other zymogens like proproteinase E and trypsinogen, however, it is not engaged in a "second triad" (66), as the positions corresponding to the second histidine and serine residues are Leu^{40} and Ala^{32}, respectively, in pro-hK6.

Comparison with Related Serine (Pro)Proteinases—Pro-hK6 has the archetypal structure of trypsinogen/chymotrypsinogen-like serine (pro)proteinases, consisting of the two interacting β-barrels and the two α-helices. These regions, including between 196 and 211 structurally aligned residues, superimpose well in pro-hK6 and neuropsin (49% sequence identity; see Fig. 1 and Table II), bovine trypsin/trypsinogen (44%), and porcine pancreatic kallikrein (44%), with higher sequence identity than to other kallikreins (40–42%). The sequence identity with bovine chymotrypsinogen A is significantly lower (38%). Pro-hK6 has chimeric features between the other kallikrein family members and trypsinogen. Unlike most of the former, but similarly to the latter, pro-hK6 lacks the kallikrein loop (see below and Fig. 1), important for kallikrein specificity and a point at which the polypeptide is cleaved and where an *N*-glycosylation site is located in kallikrein A, kallikrein-13, and other family members although not in pro-hK6 (67). However, only 3 of the 15 human kallikreins have the complete loop, believed to confer specificity for kininogenase activity (5). The classic kallikreins display five disulfide bonds, a subset of the six found in bovine trypsin, whereas pro-hK6 and most of the recently sequenced members also display the sixth (see Fig. 1 and Ref. 5). Our human proenzyme further displays an intermediate helix α_1 conformationally equivalent to that of trypsin and distinct from kallikrein A. This feature constitutes one of the most evident

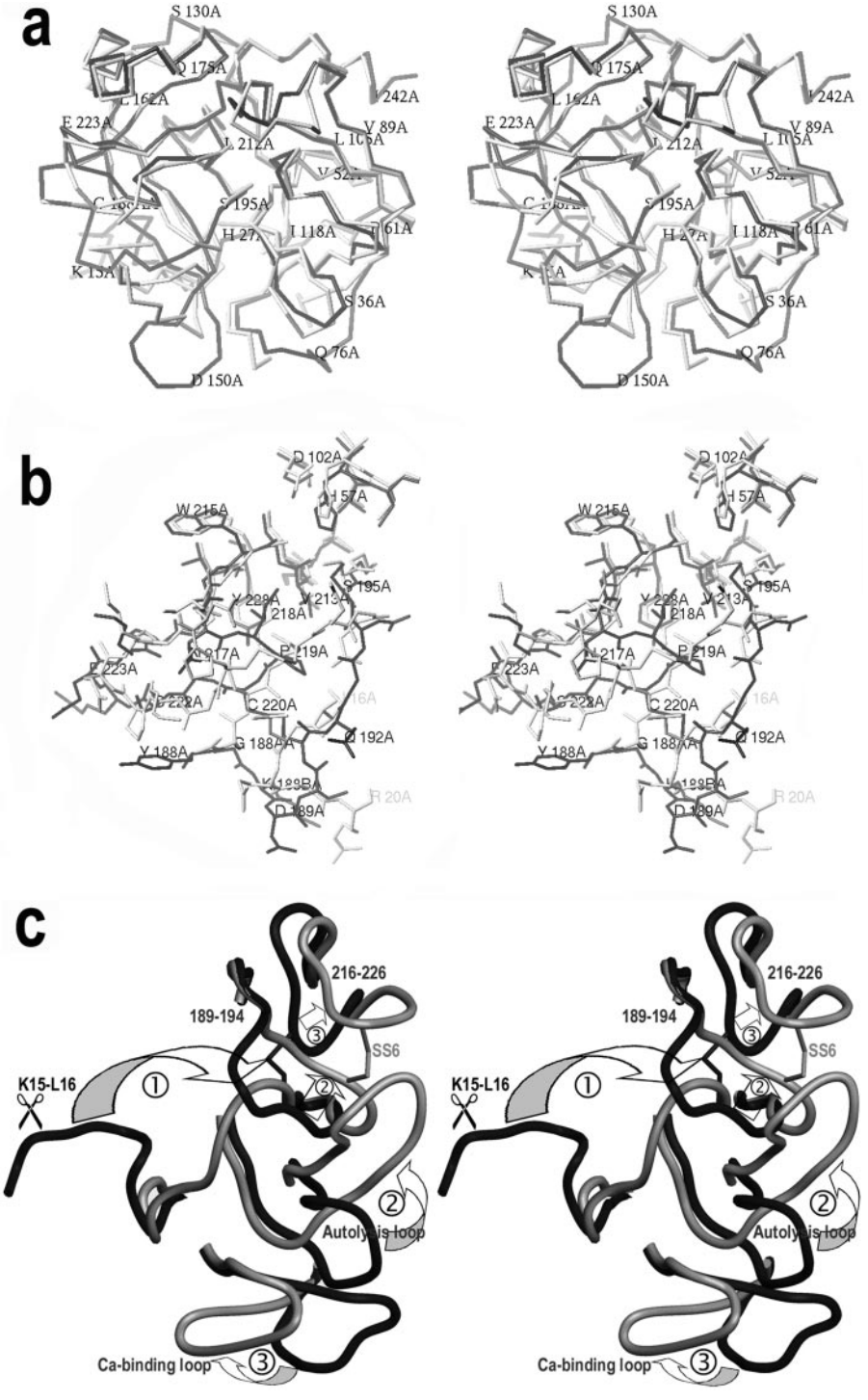

FIG. 4. **Proposed activation of pro-hK6.** *a*, superimposition of the C-α traces of pro-hK6 (*dark sticks*) and neuropsin (*light sticks*) shown in standard orientation. *b*, close up view of the regions of pro-hK6 putatively involved in major structural rearrangements upon activation (*dark sticks* and *labels*) superimposed with the equivalent zones of neuropsin (*light trace* and *labels*), considered as a structural homologue of active hK6. *c*, scheme highlighting the proposed structural rearrangements observed on going from pro-hK6 (*dark coil*) to neuropsin (*light coil*).

changes between the latter two (67). Moreover, like kallikrein A, pro-hK6 has a deletion at position 37 belonging to loop $\beta_1\beta_2$, which adopts an equivalent conformation, distinct from trypsin. It also has the characteristic *cis*-Pro^{219}, part of an extended loop forming an enlarged bump from His^{217} to Cys^{220}, and a specific orientation of the carbonyl-oxygen of *cis*-Pro^{219}, which points out of the pocket in contrast to other serine proteinases (67). This feature is present in all kallikreins but missing in trypsin (Fig. 1). When compared with the latter, kallikrein A shows large deviations in the external loops, mostly surrounding the substrate binding site and forming a more compact rampart around it (67). This contributes to the generally restricted accessibility of the kallikrein substrate-binding site compared with bovine trypsin, possibly to achieve distinct specificity or the prevention of unwanted inhibition by proteinaceous inhibitors targeting trypsin. The specificity determining aspartate at position 189 is common to most kallikreins except PSA, which, like chymotrypsin, has a serine at this position. This explains its dual trypsin- and chymotrypsin-like specificity (68). Further exceptions are human KLK7 (asparagine), KLK9 (glycine), and KLK15 (glutamate) (5). 27 of the reported "invariant" amino acids (29–31, depending on the source (59, 69, 70)) surrounding the active site of trypsin-like serine proteinases and involved in catalysis or substrate anchoring are totally conserved in pro-hK6. Of the other two residues, Leu^{155} in trypsin is conservatively exchanged for an isoleucine, as in kallikrein A, and trypsin Pro^{161} is not conserved in most of the other kallikrein sequences either (see Fig. 1).

Despite the similarity mentioned above, the loops and segments connecting the regular secondary structure elements form a completely novel structure (see Fig. 3, *a–d*), not attributable to crystallographic contacts in the crystal and accounting for a unique serine proproteinase fold. These zones include the entrance frame (Val^{213}–Cys^{220}) with the connecting disulfide bridge SS6 (Cys^{191}–Cys^{220}) and loop Gly^{216}–Gly^{226}, containing the back-forming segment (Gly^{226}–Tyr^{228}), and the base-forming region Asp^{189}–Ser^{195}, flattened on top of the C-terminal β-barrel on its front. The unique orientation of these regions is accompanied by further singularities in the regions corresponding to the so-called autolysis and calcium-binding loops (Fig. 3, *a–d*). Although partially undefined, the latter is positioned ~10 Å further away (as calculated for Gln^{75} C-α, at the center of the loop) when compared with trypsin(ogen) and with other (also cation-depleted) kallikrein structures. The latter both display remarkably similar main-chain conformations despite the lack of sequence similarity. This distinct arrangement induces the shift of the autolysis loop in pro-hK6, which invades the space released by the calcium-binding loop. The former is unique in being 2 residues shorter than in trypsin and other related serine proteinases (see Fig. 1) and even 4 residues shorter than in kallikrein A. This zone shows high conformational variability among serine (pro)proteinases (see Fig. 3), although none is as close to the calcium-binding loop as in pro-hK6. On the other side of this loop, these changes lead the end of the base-forming segment (distinct from Ala^{183}–Asp^{194}; see Fig. 3, *a–d*) to occupy the free space. Upon further transmission of these changes, loop Gly^{216}–Lys^{224}, together with disulfide bond SS6, collapses into the molecule, with a maximum distance of 6–7 Å for Cys^{220}.

The surface-located segment Ser^{130}–Thr^{134} also adopts a distinct conformation (not being involved in any intra- or intermolecular contacts), with a maximum deviation of ~7 Å. It is folded toward β-hairpin $\beta_{10}\beta_{11}$ at Thr^{133} and Ala^{131}. This region, including the only glycosylation site at position 132, may play a role in a regulatory autoactivation/autoinactivation mechanism or in promoting sugar-mediated adhesion to cells or other molecules, where its activity is required.

Finally, the final refined structure has a unique rotamer of the side chain of active-site Ser^{195}. Two solvent molecules and His^{57} Nϵ2 coordinate its Oγ atom in a tetrahedral manner. The unique feature resides in the *gauche*$^-$ conformation of its χ_1 angle (+67°), rendering a high steric hindrance to the interaction with a substrate scissile peptide bond. The side chain can only approach the substrate if the angle becomes negative (*gauche*$^+$), as observed in all other zymogen and active protease structures analyzed (see Fig. 3, *a–d*; values oscillate between −43 and −96°).

Proposed Activation Mechanism—Pro-hK6 is activated *in vivo* probably by limited proteolysis of peptide bond Lys^{15}–Leu^{16}, either autolytically or by another yet unknown proteinase with a compatible expression pattern. As previously mentioned, pro-hK6 displays the closest structural and sequential similarity with neuropsin (see Fig. 4), although the base-forming segment Ala^{183}–Ser^{195} is one residue longer in neuropsin. This causes significant structural changes in this loop and in that of the disulfide-connected segment Gly^{216}–Gly^{226}, endowing hK6 with a slightly different substrate specificity upon maturation. Both proteins have been detected in mammalian brain (32) and, like other homologues such as myelencephalon-specific protease (71), KLK14 (72), reelin (73), spinesin (74), and neurotrypsin (75), show higher expression levels in the central nervous system than in most peripheral tissues. A recent report has even shown that high levels of plasma kallikrein are present in the brain, where it could play a role in the nervous system (76). Both pro-hK6 and neuropsin share, besides a sequence identity of 49%, the absence of the kallikrein loop (see Fig. 1) and the presence of the sixth SS bond found in bovine trypsin, although absent in other kallikrein structures. Accordingly, neuropsin, the only brain-associated kallikrein of known structure (32), can be considered as a valid model for active hK6 and thus suggests which structural rearrangements would possibly occur upon pro-hK6 activation (see above and Fig. 4).

Thus, upon proteolytic cleavage of the pro-hK6 target bond (Fig. 4*c*, ①), the new N-terminal segment Leu^{16}–Pro^{28} would fold toward the molecular surface, mainly due to two concerted ~90° rotations around the peptide bond between Lys^{24} and Thr^{25} and around Cys^{157} C-α–C-β, the position of Cys^{157} C-α in neuropsin and pro-hK6 being just 0.8 Å away. This would displace the connected Cys^{22} (4 Å for its C-α atom) and probably cause the establishment of a main-chain hydrogen bond between Ile^{155} O and Cys^{22} N in the mature enzyme. The preceding segment Leu^{16}–Pro^{21} could be rotated toward the molecular body, forming a further double main-chain hydrogen bond between Gly^{20} and Cys^{157}. In this movement, Leu^{16} penetrates the protein moiety originating a displacement of segment Gly^{188A}–Asp^{194} (Fig. 4*c*, ①), which would be folded inward with a maximum displacement of ~10 Å for Asp^{189} C-α. This is mainly permitted by a ~150° rotation around Gly^{188A} N-Cα, facilitated by the absence of a side chain in this residue, and a further rotation of ~130° around Asp^{194} C-α–C. In this way, the acidic side chain of the latter residue would rotate to establish the classical salt bridge with the new N terminus, Leu^{16} N, and the oxyanion hole-shaping main-chain nitrogen atom of Gly^{193} would be properly positioned. Also, these changes might trigger a ~180° rotation of the side chain of active site Ser^{195} around its χ_1 angle to reach the negative value observed in neuropsin and other active serine proteinases. The residue following the N terminus, Val^{17}, would be trapped by Asp^{189} via a main-chain double hydrogen bond. The further intrusion of the new N terminus would also rearrange the autolysis loop (Fig. 4*c*, ②), induced by rotations around the C-α–C bonds of Asp^{153} (~130°) and Trp^{141} (~90°), to occupy the space left by Gly^{188A}-Asp^{194}.

These movements also affect the connecting disulfide bridge SS6 (Cys^{191}–Cys^{220}), since Cys^{191} C-α would move for about ~7–8 Å inward in the mature structure, inducing the extrusion of loop Gly^{216}–Glu^{223} (Fig. 4*c*, ③) by a ~90° rotation around Gly^{216} N–C-α. Furthermore, the rearrangement of the autolysis loop could push out the vicinal calcium-binding loop to adopt the conformation observed in neuropsin (partially undefined in this structure however), but also in kallikrein A, tonin, and kallikrein 13 (see Fig. 3, *a–d*). This structural element would be folded out with a maximum displacement of ~13 Å for Gln^{76}, explaining the susceptibility of this region to autolytic cleavage upon proteolytic activation. In any case, this movement, mainly enabled by a ~60° rotation around Leu^{73} N–C-α and a second of ~90° around Glu^{81} C-α–C, would imply the rupture of the strong double main chain hydrogen bond between Leu^{73} and Asp^{153} observed in pro-hK6.

Conclusion—Despite close overall structural similarity with other members of the trypsin/chymotrypsin-like family of serine (pro)proteinases, the structure of pro-hK6 has unique features that mainly affect those regions involved in rearrangement upon activation. The close sequential and structural similarity and tissue distribution of pro-hK6 with mouse neuropsin makes the latter a valid structural model for active human hK6 in the absence of an experimental structure and allows the formulation of a novel activation mechanism. The proposed activation may be achieved either by extrinsic trypsin(-like) activity involving other kallikreins or intrinsically

due to a regulatory autoactivation mechanism. In addition to this activation pathway, the autolysis-sensitive position after Arg^{76} leads to an inactive species and may provide a second regulatory mechanism; once mature protein has been generated from pro-hK6 in the presence of adequate stimuli and in the proper environment, the capacity for autolysis at this position may provide a complementary autoinactivation strategy, acting when the local concentration of hK6 becomes too high or its activity is no longer required.

Acknowledgments—We are indebted to Rosa Pérez-Luque and Tània González for technical assistance in crystallization experiments and Dimitrios Vahliotis for DNA sequencing. We especially thank the ESRF and EMBL local contacts and the team of the Spanish CRG Beamline BM-14-S at ESRF for providing assistance and support for measurements.

REFERENCES

1. Schachter, M. (1979) *Pharmacol. Rev.* **31,** 1–17
2. Kraut, H., Frey, E. K., and Werle, E. (1930) *Hoppe-Seylers Z. Physiol. Chem.* **189,** 97–106
3. Nustad, K., Orstavik, T. B., Gautvik, K. M., and Pierce, J. V. (1978) *Gen. Pharmacol.* **9,** 1–9
4. Fiedler, F. (1979) in *Bradykinin, Kallidin, and Kallikrein* (Erdös, E. G., ed), pp. 103–161, Springer, Berlin
5. Yousef, G. M., and Diamandis, E. P. (2001) *Endocr. Rev.* **22,** 184–204
6. Clements, J. A. (1997) in *The Kinin System* (Farmer, S. G., ed) pp. 71–97, Academic Press, Inc., San Diego, CA
7. Gerald, W. L., Chao, J., and Chao, L. (1986) *Biochim. Biophys. Acta* **866,** 1–14
8. Chen, Z. L., Momota, Y., Kato, K., Taniguchi, M., Inoue, N., Shiosaka, S., and Yoshida, S. (1998) *J. Histochem. Cytochem.* **46,** 313–320
9. Fujinaga, M., and James, M. N. G. (1987) *J. Mol. Biol.* **195,** 373–396
10. Timm, D. E. (1997) *Protein Sci.* **6,** 1418–1425
11. Pérez-Polo, J. R., De Jong, W. W., Straus, D., and Shooter, E. M. (1972) *Adv. Exp. Med. Biol.* **32,** 91–97
12. Evans, B. A., Yun, Z. X., Close, J. A., Tregear, G. W., Kitamura, N., Nakanishi, S., Callen, D. F., Baker, E., Hyland, V. J., and Sutherland, G. R. (1988) *Biochemistry* **27,** 3124–3129
13. Schedlich, L. J., Bennetts, B. H., and Morris, B. J. (1987) *DNA* **6,** 429–437
14. Riegman, P. H., Vlietstra, R. J., van der Korput, J. A., Romijn, J. C., and Trapman, J. (1989) *Biochem. Biophys. Res. Commun.* **159,** 95–102
15. Liu, X. L., Wazer, D. E., Watanabe, K., and Band, V. (1996) *Cancer Res.* **56,** 3371–3379
16. Anisowicz, A., Sotiropoulou, G., Stenman, G., Mok, S. C., and Sager, R. (1996) *Mol. Med.* **2,** 624–636
17. Nelson, P. S., Gan, L., Ferguson, C., Moss, P., Gelinas, R., Hood, L., and Wang, K. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 3114–3119
18. Hartley, B. S. (1964) *Nature* **201,** 1284–1287
19. Neurath, H. (1975) in *Proteases and Biological Control* (Reich, E., Rifkins, D. B., and Shaw, E., eds) Vol. 2, pp. 51–64, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
20. Clements, J. A. (1994) *Mol. Cell. Endocrinol.* **99,** C1–C6
21. Harvey, T. J., Hooper, J. D., Myers, S. A., Stephenson, S. A., Ashworth, L. K., and Clements, J. A. (2000) *J. Biol. Chem.* **275,** 37397–37406
22. Pericak-Vance, M. A., Bebout, J. L., Gaskell Jr., P. C., Yamaoka, L. H., Hung, W. Y., Alberts, M. J., Walker, A. P., Bartlett, R. J., Haynes, C. A., and Welsh, K. A. (1991) *Am. J. Hum. Genet.* **48,** 1034–1050
23. Mitelman, F. (1994) *Catalog of Chromosome Aberrations in Cancer*, 5th Ed., John Wiley & Sons, Inc., New York
24. Amfo, K., Neyns, B., Teugels, E., Lissens, W., Bourgain, C., De Sutter, P., Vandamme, B., Vamos, E., and De Greve, J. (1995) *Oncogene* **11,** 351–358
25. Ozen, M., Hopwood, V. L., Balbay, M. D., Johnston, D. A., Babaian, R. J., Logothetis, C. J., von Eschenbach, A. C., and Pathak, S. (2000) *Int. J. Oncol.* **17,** 113–117
26. Diamandis, E. P., Okui, A., Mitsui, S., Luo, L. Y., Soosaipillai, A., Grass, L., Nakamura, T., Howarth, D. J., and Yamaguchi, N. (2002) *Cancer Res.* **62,** 295–300
27. Yamashiro, K., Tsuruoka, N., Kodama, S., Tsujimoto, M., Yamamura, Y., Tanaka, T., Nakazato, H., and Yamaguchi, N. (1997) *Biochim. Biophys. Acta* **1350,** 11–14
28. Little, S. P., Dixon, E. P., Norris, F., Buckley, W., Becker, G. W., Johnson, M., Dobbins, J. R., Wyrick, T., Miller, J. R., MacKellar, W., Hepburn, D., Corvalan, J., McClure, D., Liu, X., Stephenson, D., Clemens, J., and Johnstone, E. M. (1997) *J. Biol. Chem.* **272,** 25135–25142
29. Mason, A. J., Evans, B. A., Cox, D. R., Shine, J., and Richards, R. I. (1983) *Nature* **303,** 300–307
30. Evans, B. A., Drinkwater, C. C., and Richards, R. I. (1987) *J. Biol. Chem.* **262,** 8027–8034
31. MacDonald, R. J., Southard-Smith, E. M., and Kroon, E. (1996) **271,** 13684–13690
32. Kishi, T., Kato, M., Shimizu, T., Kato, K., Matsumoto, K., Yoshida, S., Shiosaka, S., and Hakoshima, T. (1999) *J. Biol. Chem.* **274,** 4220–4224
33. Ogawa, K., Yamada, T., Tsujioka, Y., Taguchi, J., Takahashi, M., Tsuboi, Y., Fujino, Y., Nakajima, M., Yamamoto, T., Akatsu, H., Mitsui, S., and Yamaguchi, N. (2000) *Psych. Clin. Neurosci.* **54,** 419–426
34. Okui, A., Kominami, K., Uemura, H., Mitsui, S., and Yamaguchi, N. (2001) *Neuroreport* **12,** 1345–1350
35. Petraki, C. D., Karavana, V. N., Skoufogiannis, P. T., Little, S. P., Howarth, D. J., Yousef, G. M., and Diamandis, E. P. (2001) *J. Histochem. Cytochem.* **49,** 1431–1441
36. Yousef, G. M., Luo, L. Y., Scherer, S. W., Sotiropoulou, G., and Diamandis, E. P. (1999) *Genomics* **62,** 251–259
37. López-Otín, C., and Diamandis, E. P. (1998) *Endocr. Rev.* **19,** 365–396
38. Tanimoto, H., Underwood, L. J., Shigemasa, K., Parmley, T. H., and O'Brien, T. J. (2001) *Tumour Biol.* **22,** 11–18
39. Tomizawa, K., He, X., Yamanaka, H., Shiosaka, S., and Yoshida, S. (1999) *Brain Res.* **824,** 308–311
40. Diamandis, E. P., Yousef, G. M., Soosaipillai, A. R., and Bunting, P. (2000) *Clin. Biochem.* **33,** 579–583
41. Diamandis, E. P., Yousef, G. M., Petraki, C., and Soosaipillai, A. R. (2000) *Biochemistry* **33,** 663–667
42. Diamandis, E. P., Yousef, G. M., Soosaipillai, A. R., Grass, L., Porter, A., Little, S., and Sotiropoulou, G. (2000) *Clin. Biochem.* **33,** 369–375
43. Ventura, S., Villegas, V., Sterner, J., Larson, J., Vendrell, J., Hershberger, C. L., and Avilés, F. X. (1999) *J. Biol. Chem.* **28,** 19925–19933
44. Gomis-Rüth, F. X., Gómez-Ortiz, M., Vendrell, J., Ventura, S., Bode, W., Huber, R., and Avilés, F. X. (1998) *Eur. J. Biochem.* **251,** 839–844
45. Matthews, B. W. (1968) *J. Mol. Biol.* **33,** 491–497
46. Leslie, A. G. (1999) *Acta Crystallogr. Sect. D* **55,** 1696–1702
47. Evans, P. (1993) in *Data collection and Processing: Proceedings of the CCP4 Study Weekend 29–30 January 1993* (Sawyer, L., Isaacs, N., and Bailey, S., eds) pp. 114–122, SERC Daresbury Laboratory, Warrington, UK
48. Navaza, J. (1994) *Acta Crystallogr. Sect. A* **50,** 157–163
49. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D* **54,** 905–921
50. Roussel, A., and Cambilleau, C. (1989) in *Silicon Graphics Geometry Partners Directory*, pp. 77–79, Silicon Graphics, Mountain View, CA
51. Perrakis, A., Sixma, T. K., Wilson, K. S., and Lamzin, V. S. (1997) *Acta Crystallogr. Sect. D* **53,** 448–455
52. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. Sect. D* **53,** 240–255
53. Winn, M. D., Isupov, M. N., and Murshudov, G. N. (2001) *Acta Crystallogr. Sect. D* **57,** 122–133
54. Laskowski, R. A., Macarthur, M. W., and Thornton, J. M. (1998) *Curr. Opin. Struct. Biol.* **8,** 631–639
55. Collaborative Computing Project 4 (1994) *Acta Crystallogr. Sect. D* **50,** 760–763
56. Holm, L., and Sander, C. (1998) *Nucleic Acids Res.* **26,** 316–319
57. Shindyalov, I. N., and Bourne, P. E. (1998) *Prot. Eng.* **11,** 739–747
58. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950
59. Lesk, A. M., and Fordham, W. D. (1996) *J. Mol. Biol.* **258,** 501–537
60. Rawlings, N. D., and Barrett, A. J. (1994) *Methods Enzymol.* **244,** 19–60
61. Gomis-Rüth, F. X., Gómez, M., Bode, W., Huber, R., and Avilés, F. X. (1995) *EMBO J.* **14,** 4387–4394
62. Bode, W., and Schwager, P. (1975) *FEBS Lett.* **56,** 139–143
63. Harel, M., Su, C. T., Frolow, F., Silman, I., and Sussman, J. L. (1991) *Biochemistry* **30,** 5217–5225
64. Schechter, I., and Berger, A. (1967) *Biochem. Biophys. Res. Commun.* **27,** 157–162
65. Wang, D., Bode, W., and Huber, R. (1985) *J. Mol. Biol.* **185,** 595–624
66. Madison, E. L., Kobe, A., Gething, M.-J., Sambrook, J. F., and Goldsmith, E. J. (1993) *Science* **262,** 419–421
67. Bode, W., Chen, Z., Bartels, K., Kutzbach, C., Schmidt-Kastner, G., and Bartunik, H. (1983) *J. Mol. Biol.* **164,** 237–282
68. Watt, K. W. K., Lee, P.-J., M'Timkulu, T., Chan, W.-P., and Loor, R. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83,** 3166–3170
69. James, M. N. G., Delbaere, L. T. J., and Brayer, G. D. (1978) *Can. J. Biochem.* **56,** 396–402
70. Young, C. L., Barker, W. C., Tomaselli, C. M., and Dayhoff, M. O. (1978) in *Atlas of Protein Sequence and Structure* (Dayhoff, M. O., ed) Vol. 5, Suppl. 3, pp. 73–94, National Biomedical Research Foundation, Silver Spring, MD
71. Scarisbrick, I. A., Isackson, P. J., Ciric, B., Windebank, A. J., and Rodriguez, M. (2001) *J. Comp. Neurol.* **431,** 347–361
72. Yousef, G. M., Magklara, A., Chang, A., Jung, K., Katsaros, D., and Diamandis, E. P. (2001) *Cancer Res.* **61,** 3425–3431
73. Quattrocchi, C. C., Wannenes, F., Persico, A. M., Ciafre, S. A., D'Arcangelo, G., Farace, M. G., and Keller, F. (2002) *J. Biol. Chem.* **277,** 303–309
74. Yamaguchi, N., Okui, A., Yamada, T., Nakazato, H., and Mitsui, S. (2002) *J. Biol. Chem.* **277,** 6806–6812
75. Gschwend, T. P., Krueger, S. R., Kozlov, S. V., Wolfer, D. P., and Sonderegger, P. (1997) *Mol. Cell. Neurosci.* **9,** 207–219
76. Cerf, M. E., and Raidoo, D. M. (2000) *Metab. Brain Dis.* **15,** 315–323
77. Bax, B., Blundell, T. L., Murray-Rust, J., and McDonald, N. Q. (1997) *Structure* **5,** 1275–1285
78. Fehlhammer, H., Bode, W., and Huber, R. (1977) *J. Mol. Biol.* **111,** 415–438
79. Huber, R., Kukla, D., Bode, W., Schwager, P., Bartels, K., Deisenhofer, J., and Steigemann, W. (1974) *J. Mol. Biol.* **89,** 73–101