New, Sensitive Fluorogenic Substrates for Human Cathepsin G
Based on the Sequence of Serpin-reactive Site Loops*

Sophie Réhault‡, Michèle Brillard-Bourdet‡, Maria A. Juliano§, Luiz Juliano§,
Francis Gauthier‡, and Thierry Moreau‡¶

*This work was supported by the Association pour la Recherche sur le Cancer and, in Brazil, by the Fundação de Amparo a Pesquisa do Estado de São Paulo and the Conselho Nacional de Desenvolvimento Científico e Tecnológico. 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.

‡To whom correspondence should be addressed. Tel.: 33-2-4736-6177; Fax: 33-2-4736-6046; E-mail: moreaut@univ-tours.fr.

¶Laboratory of Enzymology and Protein Chemistry, University François Rabelais, 2bis Boulevard Tonnellé, 37032 Tours Cedex, France and the ‡Laboratorio de Biologia Molecular, Universidade Federal de São Paulo, 100 Rua Três de Maio, São Paulo 04044-020, Brazil

The New, Sensitive Fluorogenic Substrates for Human Cathepsin G were developed by the authors. Cathepsin G, a protease found in human neutrophils and monocytes, has both trypsin- and chymotrypsin-like activities. The development of sensitive fluorogenic substrates for Cathepsin G was necessary to study its enzymatic properties. The authors synthesized a series of intramolecularly quenched fluorogenic peptides based on the sequence of various serpin loops. These peptides were assayed as substrates for Cathepsin G and other chymotrypsin-like enzymes including chymotrypsin and chymase. The most sensitive substrate for Cathepsin G was a peptide derived from the S1 and S2 regions of the enzyme. This substrate was able to be cleaved with a specificity constant of 5-20 mM s⁻¹. The fluorogenic activity of Cathepsin G was compared with that of such chymotrypsin-like enzymes as trypsin and chymotrypsin. The authors used this substrate to study the interaction of Cathepsin G with its inhibitors and to develop new inhibitors for the enzyme.

Human cathepsin G, neutrophil elastase, and proteinase 3 are the major proteolytic enzymes in the azurophilic granules of polymorphonuclear leukocytes (neutrophils) (1-2). They are also present in mast cells (3) and monocytes (4) and are involved in the degradation of foreign organisms or dead tissues within the phagolysosome (5) during inflammatory reactions. They are released into the extracellular space, by leakage and/or cell death, where they are controlled by plasma-derived inhibitors such as α₁-antitrypsin or α₁-proteinase inhibitor. Any imbalance between these proteases and their inhibitors because of genetic or acquired deficiencies of these inhibitors, or chronic inflammation, can result in the uncontrolled digestion of most proteins of the extracellular matrix.

The proteolytic activity of cathepsin G has been shown to be involved in the proteolysis of elastin, collagen type I and type II, cartilage proteoglycans, fibronectin, laminin, and immunoglobulins G and M (see Ref. 6 for a review). Cathepsin G is able to interact with the V3 loop of human immunodeficiency virus type 1 gp120 protein (11) and is involved in the proteolytic processing/activation of interleukin-8 (12), complement C3 (13), and factor V (14).

Investigation of the substrate specificity of cathepsin G using oxidized insulin B chain (15), peptide 4-nitroanilide substrates (16, 17), and peptide thioesters (18) has shown that cathepsin G has both chymotrypsin- and trypsin-like specificities but that it is much less active against all of the synthetic substrates assayed so far than are homologous serine proteinases. As no sensitive substrate suitable for enzymatic studies or the search for specific inhibitors of cathepsin G was available, we have developed a series of fluorogenic peptide substrates derived from the inhibitory loop of α₁-antitrypsin, the main physiological inhibitor of cathepsin G (19). Some other serpins, including α₁-proteinase inhibitor (19), have also been reported to be cathepsin G inhibitors. Interaction of these two serpins with their cognate enzymes involves the cleavage of their reactive inhibitory loop with the production of a SDS-stable complex as a result of the formation of a covalent bond. Therefore, the sequence of serpin reactive loops may be used to develop peptide substrates of the corresponding target enzyme(s), as has been shown recently for human kallikreins hK1 and hK2 (20). We therefore prepared a series of intramolecularly quenched fluorogenic substrates of cathepsin G based on the sequence of various serpin-reactive site loops and bearing an ortho-aminobenzoyl (Abz) and quenching N-(2,4-dinitrophenyl) ethylenediamine (EDDnp) as donor-acceptor pairs at the N and C termini of the peptides, respectively. The hydrolytic activity of cathepsin G was compared with that of such chymotrypsin-like enzymes as trypsin, chymotrypsin, and elastase.

1 The abbreviations used are: hK1/hK2/hK3, human kallikreins 1, 2, and 3, respectively; Abz, ortho-aminobenzoyl acid; AC, antichymotrypsin-deprived substrate; ACT, α₁-antichymotrypsin; EDDnp, N-(2,4-dinitrophenyl) ethylenediamine; HPLC, high performance liquid chromatography; pNA, para-nitroanilide; SCCA, squamous cell carcinoma antigen; RSL, reactive site loop.
zymes as chymotrypsin, chymase, and human Kallikrein hK3 (also called prostate-specific antigen). The effects of residues at the P1 and P2 positions on substrate binding were assessed using the crystal structure of human cathepsin G (21). The nomenclature used for the individual amino acid residues (P1, P2, etc.) of the substrate and corresponding residues of the enzyme subites (S1, S2, etc.) is that of Schecter and Berger (22).

**EXPERIMENTAL PROCEDURES**

**Materials**—Human cathepsin G (EC 3.4.21.20) was from ICN Pharmaceuticals, bovine chymotrypsin (EC 3.4.21.1) was from Roche Molecular Biochemicals, and human recombinant chymase (EC 3.4.21.39) was kindly provided by Prof. Norman Schecter (Dept. of Dermatology, University of Pennsylvania, Philadelphia, PA). hK3 was purified as described previously (23). N,N-Dimethylformamide and acetonitrile were from Merck, and C18 cartridges for reverse-phase chromatography were from Brownlee Laboratories. All other reagents were of analytical grade.

**Design and Synthesis of Quenched Fluorescent Substrates**—All quenched fluorogenic substrates were synthesized by solid-phase procedures with the Fmoc (N-(9-fluorenylmethoxycarbonyl) methodology using a multiple automated peptide synthesizer (PSSM-8, ShimaKyo Co.) according to previously described procedures (24–26). Glutamine was the C-terminal residue in all peptides because of a requirement of the synthesis strategy (24). Substrate purity was checked by matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF; Micromass, ToFSpec-E) and by reverse-phase chromatography on a C18 column eluted at 2 ml/min with a 10-min linear gradient of acetonitrile (0–60%) in 0.075% trifluoroacetic acid. Substrate stock solutions (2–5 mM) were prepared in N,N-dimethylformamide and diluted with activity buffer.

**Enzyme Assays**—Assays were carried out at 37 °C in 50 mM Hepes buffer, pH 7.4, 50 mM NaCl for cathepsin G and chymotrypsin, in 0.1 M Tris/HCl, pH 8.0, 1.0 mM NaCl for human chymase, and in 50 mM Tris/HCl, pH 8.3, 1 mM EDTA for hK3.

The hydrolysis of Abz-peptidyl-EDDnp substrates was followed by measuring the fluorescence at λem = 320 nm and λex = 420 nm in a Hitachi F-2000 spectrophotometer. The system was standardized using Abz-FR-OH prepared from the total tryptic hydrolysis of an Abz-FR-NA methodology (MALDI-TOF; Micromass, ToFSpec-E) and by reverse-phase chromatography on a C18 column eluted at 2 ml/min with a 10-min linear gradient of acetonitrile (0–60%) in 0.075% trifluoroacetic acid. Substrate stock solutions (2–5 mM) were prepared in N,N-dimethylformamide and diluted with activity buffer.

Specificity constants (kobs/Ke) were determined under first-order conditions using a substrate concentration far below the Km (0.1–6 μM, depending on the enzyme). Final enzyme concentrations were 40–100 nM for cathepsin G, 15–70 nM for chymotrypsin, and 10–50 nM for chymase. Under these conditions, the Michaelis-Menten equation is reduced to the following: v = kobs[S], where kobs = Vmax/Km. Integrating this equation over time gives ln [S] = -kobs·t + ln[S]0 with [S] being the substrate concentrations at time 0 and time t, respectively. Because Vmax = kcat[E], where [E] is the final enzyme concentration, dividing kobs by [E] gave the kobs/Km ratio. The kcat for the first-order substrate hydrolysis was calculated by fitting experimental data to the first-order law using Enzfitter software (Elsevier Science Publishers, Amsterdam).

Cathespin G and chymotrypsin were titrated with α1-antichymotrypsin as described previously (27), whereas the concentration of active chymase was determined under standard assay conditions using 1 mM succinyl-Ala-Ala-Pro-Phe-pNA as a substrate and assuming a specific activity of 2.7 μmol of p-nitroaniline released min⁻¹/mmol of chymase (28).

The Km and Vmax for the hydrolysis of Abz-peptidyl-EDDnp substrates by cathepsin G were determined using 8–10 substrate concentrations (1–15 μM). The final concentration of cathepsin G was 20–100 nM depending on the substrate used. Experimental data were fitted to the hyperbolic Michaelis-Menten rate equation using Enzfitter software. Values for kcat were obtained from Vmax/[E] = kcat.

**Chromatographic Procedures and Analysis of Peptide Products**—Fluorogenic substrates (4–80 μM final concentration) were incubated with cathepsin G (200–400 nM), chymotrypsin (7–75 nM), or chymase (20–120 nM) at 37 °C in their respective buffers. The reaction was blocked by adding 3 μl of trifluoroacetic acid, and aliquots (200 μl) were removed at times ranging from 2 min to 1 h depending on the substrate and enzyme used. The substrate fragments generated were purified by reverse-phase chromatography on a C18 column (2.1 × 30 mm, Brownlee), a P200 pump (Thermo Separation Products), and a SpectraSystem UV3000 detector ((Thermo Separation Products), at a flow rate of 1 ml/min, with a linear (0–85%, v/v) gradient of acetonitrile in 0.075% trifluoroacetic acid over 12 min. Eluted peaks were simultaneously monitored at three wavelengths (220, 320, and 360 nm), which allowed the direct identification of EDDnp-containing peptides prior to sequencing or amino acid analysis. Cleavage sites were identified by N-terminal sequencing with an Applied Biosystems 477A pulsed liquid sequencer with the chemicals and program recommended by the manufacturer. Phenylthiocarbamide-amino acids were detected at 254 nm and identified and quantified by reference to phenylthiocarbamide-amino acid standards.

**Model Building**—The model of a Pz-P1' hexapeptide (TPFSSM) complexed to cathepsin G was constructed using the Insight II and Discover molecular modeling and energy minimization programs (Biosym Technologies Inc.). The coordinates for cathepsin G (Protein Data Bank code 1cgH) were from (21), where the complexed succinyl-Val-Pro-Phe'-(Oph) inhibitor (Pz = diphenyl ether) was first respectively superimposed onto the ecotin-crab collagen complex (29). The Pz-P1' side chains in bound ecotin were changed to those for the model substrate by keeping the same conformation, except for Phe at the P1' position, in which the side chain was manually moved to adopt a conformation similar to that of the Phe occupying the S1' pocket of cathepsin G. Flanking residues were deleted and the substrate-cathepsin G complex was minimized using the steepest descent algorithm until the derivative was <0.1 kcal/molÅ⁻² with the cathepsin G atoms fixed to their crystallographic positions and the atoms of the substrate backbone tethered with a force constant of 500 kcal/molÅ².

**RESULTS AND DISCUSSION**

The activity of proteolytic enzymes is generally measured by the amidoletic release of a chromophore or a fluorophore covalently attached to a peptide moiety that confers the specificity to the enzyme under study. Cathepsin G, like some other serine proteases including proteinase 3 and hK3 (also called prostate-specific antigen), has little activity toward these substrates because of a low intrinsic catalytic activity of the proteinase or a low affinity with the peptide moiety. We have therefore used peptide sequences derived from proteolytically sensitive loops in natural serpin inhibitors to synthesize substrates with intramolecularly quenched fluorescence (Table I) that have residues on both sides of the cleavage site. The interaction of serpins with their cognate enzymes results in cleavage of the reactive loop, leading to either the formation of a covalent, enzymatically inactive, complex (inhibitory pathway) or the release of an inactive serpin of lower intramolecular energy that has undergone a large conformational change (substrate pathway). This conformational change results in the insertion of the cleaved loop in a β-sheet of the protein. Partitioning of serpin between the substrate and the inhibitor pathway depends on the serpin-proteinase pair and explains why there can be stoichiometries of inhibition >1 (30). Because serpin specificity toward a proteinase is governed mainly by the inhibitory loop sequence, synthetic peptides derived from the variable loop structure of individual serpin inhibitors should be good substrates for the corresponding cognate serine proteinase(s). This approach was recently used to develop substrates of hK2 and hK1.

**Hydrolysis of α1-Antichymotrypsin-derived Fluorogenic Substrates by Cathepsin G and Other Chymotrypsin-like Enzymes**—α1-Antichymotrypsin (ACT) is the major physiological inhibitor of cathepsin G and is also an inhibitor of chymase and chymotrypsin (31). Peptide substrates of increasing length (Table I), reproducing the sequence of human ACT loop from P1 to P15' (nomenclature based on the Leu-Ser cleavage site in the
Table I

Amino acid sequences of intramolecularly quenched Abz-peptidyl-EDDnp substrates derived from serpin-reactive site loops

The sequence alignment is based on the scissile bond at the P1-P1' site in the corresponding native serpins.

| No. | Serpin   | Amino acid sequence (P5 4 3 2 1 1' 2' 3' 4' 5') |
|-----|----------|-----------------------------------------------|
| AC-2| ACT      | Abz - L L S A L Q - EDDnp                     |
| AC-3| ACT      | Abz - T L L S A L Q - EDDnp                   |
| AC-4| ACT      | Abz - I T L L S A L Q - EDDnp                 |
| AC-5| ACT      | Abz - I T L L S A L V E T R Q - EDDnp         |
| AC-6| ACT      | Abz - K I T L L S A L V E T R T I V R P N R Q - EDDnp |
| AP-1| a1-PIb   | Abz - E A I P M S I P P E V K F N K Q - EDDnp |
| SC-1| SCCA1    | Abz - G P G S S P T O - EDDnp                 |
| SC-2| SCCA2    | Abz - V E L S S P Q - EDDnp                   |
| MA-1| Maspin   | Abz - V P G A R I L Q - EDDnp                 |

* Sequence positions around the cleavage site in serpin-reactive site loops are designated according to the nomenclature of Schechter and Berger (22).

a PI, proteinase inhibitor.

**ACT reactive loop denoted as P1-P1'-1 (22),** were synthesized and assayed as substrates of cathepsin G and related enzymes having chymotrypsin-like specificity, such as human chymase, bovine chymotrypsin, and human kallikrein hK3. The specificity constants $k_{cat}/K_m$ for the hydrolysis of substrates were determined under first-order conditions, i.e. using substrate concentrations far below the supposed $K_m$, thus avoiding the intermolecular quenching that occurs when the substrate concentration and/or length increases. As the specificity constant is not influenced by the nature of the rate-limiting step or by nonproductive substrate binding (32), it can be compared for different substrates and enzymes.

All ACT-derived substrates were hydrolyzed by cathepsin G, chymotrypsin, and chymase (Table II), although at different rates, in keeping with the inhibition spectrum of ACT. The three enzymes cleaved the shortest substrate of the series (AC-2) with the same $k_{cat}/K_m$, but the specificity constant increased dramatically with the peptide length using chymotrypsin and chymase, except for the longest substrate (AC-6), whereas it remains essentially constant for cathepsin G. This shows the importance of the residues from P13 to P5, and to a lesser extent of residues at P4-P5', for the interaction between the substrate and chymotrypsin-like proteinases. Under the experimental conditions used, we found no significant hydrolysis of these substrates by hK3, despite the fact that this chymotrypsin-like enzyme is inhibited by a1-antichymotrypsin in vivo (33). This result could be because of the low reactivity of hK3 against synthetic substrates (33) or because the inhibition mechanism of this proteinase by ACT is different and results in a different site of cleavage or no cleavage of the reactive loop.

The $K_m$ values for the hydrolysis of the ACT-derived substrates AC-2–AC-6 (Table III) are far lower than those reported for the peptidyl 4-nitroanilide substrates currently used for cathepsin G (Suc-VPF-pNA, $K_m = 1.4$ mM (17) or succinyl-AAPF-pNA $K_m = 1.7$ mM (27)). However, there was a limited improvement in the specificity constants because the $k_{cat}$ values were lower than those for the peptidyl 4-nitroanilide substrates (Table III), in agreement with the fact that the para-nitroanilide group is a very good leaving group for chromogenic substrates.

Cleavage sites in ACT-derived substrates were identified by reverse-phase HPLC fractionation of hydrolysis products and sequencing of the EDDnp-containing fragment(s) and/or by amino acid analysis of the Abz-containing fragment. Cathepsin G, chymotrypsin, and chymase all cleaved the same Leu-Ser bond in all substrates tested, whatever the substrate length (Fig. 1). This cleavage site is identical to that in a1-antichymotrypsin-proteinase complexes. This result emphasizes the importance of this sequence as a whole fitting into the enzyme active site, because there was no further cleavage after the other two Leu residues.

**Other Serpin-derived Substrates for Cathepsin G—** Substrates derived from the sequence of a1-proteinase inhibitor and squamous cell carcinoma antigen-2 (SCCA2) (34), which are two other serpin inhibitors of cathepsin G, were synthesized (Table I) and assayed with chymotrypsin-like proteinases including cathepsin G. a1-Proteinase inhibitor-derived substrate AP-1 was hydrolyzed by cathepsin G and chymase, although less efficiently than ACT-derived substrates, but was a very good substrate for chymotrypsin (Table IV). All enzymes cleaved at the same Met-Ser bond as that reported in a1-proteinase inhibitor upon interaction with most of target proteinases. A different cleavage site, at the Phe1-1-proteinase inhibitor-derived substrate AP-1, which does not contain the Phe-Leu pair. The SCCA2-derived substrate SC-2 was poorly hydrolyzed by cathepsin G but was very efficiently cleaved by chymase at the same Leu-Ser bond. Chymotrypsin did not hydrolyze the SCCA2-derived substrate SC-2, which agrees with the fact that it is not inhibited by native SCCA2, unlike cathepsin G or chymase (34).

We also tested a substrate that reproduces part of the sequence of SCCA1, a serpin that is 92% identical to SCCA2, with no serine proteinase target identified so far. It has been reported as a cysteine proteinase inhibitor (35). The SCCA1-derived substrate was not hydrolyzed by cathepsin G, chymotrypsin, or chymase, although it contained a Phe residue. The reactive site loop of maspin, a serpin that possesses a tumor-suppressing activity, could also be a good candidate for designing a substrate for cathepsin G. The inhibitory activity of this serpin is controversial, although recent studies have identified tissue-type plasminogen activator as its potential target (36). Other studies indicate that the maspin RSL is very sensitive to proteolysis, which leads to maspin inactivation (37) and would explain its lack of inhibitory activity. The substrate derived from the maspin RSL was a very poor substrate of cathepsin G and chymase, but it was rapidly hydrolyzed by chymotrypsin with a specificity constant of 41.7 ± 0.8 mM$^{-1}$s$^{-1}$ (Table IV).

**Substitution at Positions P2 and P1 of the Lead ACT-derived Substrate, Abz-TLLSALQ-EDDnp—** The nature of the P1 amino acid is critical for defining the substrate specificity of many serine proteinases. Cathepsin G has a preference for Phe, Lys, Arg, Leu, and Met at the P2 position of substrates (16–17, 38). We therefore replaced the P2 Leu residue of the scissile Leu-Ser bond in Abz-TLLSALQ-EDDnp with an aromatic (Phe) or a basic (Arg or Lys) residue to obtain a substrate of improved
Synthetic Substrates for Cathepsin G

| Substrate | No. | k_{cat}/K_{m} | k_{cat}/K_{m} | k_{cat}/K_{m} |
|-----------|-----|---------------|---------------|---------------|
| Abz-LLSALQ-EDDnp | AC-2 | 7.7 ± 0.2 | 7.6 ± 0.2 | 7.0 ± 0.1 |
| Abz-LLSALQ-EDDnp | AC-3 | 17.4 ± 0.1 | 52.0 ± 0.3 | 360.3 ± 5.5 |
| Abz-LLSALQ-EDDnp | AC-4 | 3.9 ± 0.1 | 180.2 ± 1.6 | 292.3 ± 1.8 |
| Abz-LLSALQ-EDDnp | AC-5 | 14.2 ± 0.2 | 245.2 ± 3.6 | 660.0 ± 3.5 |
| Abz-LLSALQ-EDDnp | AC-6 | 7.5 ± 0.2 | 236.4 ± 7.4 | 72.8 ± 3.0 |

*Values are means ± S.D. (n = 3).

Kinetic parameters for the hydrolysis of fluorogenic substrates derived from the α1-antichymotrypsin-reactive site loop by cathepsin G

| Substrate | No. | K_{m} | k_{cat} | k_{cat}/K_{m} |
|-----------|-----|-------|--------|---------------|
| Abz-LLSALQ-EDDnp | AC-3 | 30.4 ± 8.0 | 0.30 ± 0.06 | 9.8 |
| Abz-LLSALQ-EDDnp | AC-4 | 21.7 ± 5.8 | 0.33 ± 0.06 | 17.4 ± 0.1 |
| Abz-LLSALQ-EDDnp | AC-5 | 8.2 ± 0.2 | 0.40 ± 0.03 | 48.8 |
| Abz-LLSALQ-EDDnp | AC-6 | 15.5 ± 4.3 | 0.08 ± 0.02 | 5.6 |
| Abz-LLSALQ-EDDnp | AC-9p | 10.7 ± 1.8 | 1.00 ± 0.09 | 95.8 |
| Succ-VPF-pNA | | 1400 | 9.6 | 6.9 |
| Succ-AAPF-pNA | | 1700 | 3.1 | 1.8 |

*Values are means ± S.D. (n = 3).

The important contribution of a Phe at P1 or a Pro-Phe pair at P2-P1 in cathepsin G substrates is further shown by the substantially lower free energy (Table V) required to reach the transition state complex for hydrolysis of AC-9 and AC-9p (ΔG = −4.5 kJ mol^{-1} for P1 Leu → Phe and −5.62 kJ mol^{-1} for P2-P1 Leu-Leu → Pro-Phe). However, flanking residues at the P and P’ positions also contribute to recognition and cleavage efficiency, because substrates retaining the Pro-Phe moiety in their sequences are not cleaved with the same efficiency by cathepsin G. This is emphasized by the low cleavage susceptibility of Abz-AAPFSQ-EDDnp (k_{cat}/K_{m} = 3.8 m{s}^{-1} s^{-1} ± 0.28) by cathepsin G. Other substitutions at P1 resulted in a slight increase in the relative transition state binding energy (∆ΔG), the least favorable residue being Arg at P1.

Taken together, the results show that cathepsin G has the following preference for P residues: Phe > Leu > Lys > Arg. This order of preference is similar but not identical to that recently described by Polanowska et al. (38), who used a series of p-nitroanilide substrates and found the following order: Phe > Lys > Arg > Met. These differences may be attributed to synergistic interactions between cathepsin G substrates, as shown for human chymase (40), and/or to the presence of P’ residues in the intramolecularly quenched substrates.

Earlier studies were done using peptide sequences derived from the RSL of α1-antichymotrypsin (17) or α1-proteinase inhibitor (41) to design cathepsin G substrates, but the kinetic parameters indicated that they were poorer substrates than those developed in this study, probably because of their short length, or lack of residues on the prime side.

ACT variants mutated at the P1 residue of their RSL have specificity toward cathepsin G. In keeping with the preference of chymotrypsin-like proteinases for an aromatic residue at P1, replacement of the P1 Leu by Phe in AC-9 resulted in a 2–6-fold increase in the specificity constants (Table V) for the three proteinases tested. Introduction of an Arg (AC-7) or a Lys (AC-8) at P1 resulted in a slight decrease in k_{cat}/K_{m} for cathepsin G and chymotrypsin, whereas chymase hardly cleaved these two substrates.

Combining a Phe residue at P1 with a Pro at P2, which is often a preferred residue at position P2 of synthetic substrates for many serine proteinases (39) including cathepsin G (16), resulted in an approximately 2-fold increase in the k_{cat}/K_{m} values. This was thus the best substrate (AC-9p) for cathepsin G and chymase in this series. With Leu, Arg, and Lys at P1, the presence of a P2 Pro did not significantly modify the catalytic efficiency of the enzymes, except for the substrate containing Pro-Arc, AC-7p, which gave a slightly lower k_{cat}/K_{m} value when tested with cathepsin G and chymase.

The substrates containing Phe (AC-9 and AC-9p) and Leu (AC-3 and AC-3p) at P1 were cleaved by all three enzymes at the Phe-Ser and Leu-Ser bonds, but those containing Arg or Lys residues were further cleaved at the Leu-Gln bond by chymotrypsin and chymase but not by cathepsin G. Therefore the experimentally determined k_{cat}/K_{m} value for the hydrolysis of AC-7, AC-7p, AC-8, and AC-8p by chymotrypsin and chymase is an average value for the parallel hydrolysis of two peptide bonds. They are probably cleaved with similar efficiencies, i.e. with the same k_{cat}/K_{m} because experimental data for first-order substrate hydrolysis fit a simple exponential (data not shown). The presence of a basic residue at P1 therefore directs the cleavage by chymotrypsin and chymase, but not by cathepsin G, toward the Leu-Gln bond. Thus, chymotrypsin and chymase have a more restricted specificity for aromatic/hydrophobic residues at P1 than does cathepsin G, indicating that cathepsin G has almost equal trypsin- and chymotrypsin-like specificities.
has been produced and assayed as inhibitors of chymotrypsin-like proteinases (28). The stoichiometry of inhibition (SI) varies significantly, depending on the nature of the P1 residue. For example, a SI of 7.0 was obtained for a Leu358-Phe variant (rACT L358R) interacting with chymase, compared with a SI of 4.5 with native ACT, indicating the predominance of the substrate pathway for the Phe P1 ACT variant. A variant of ACT with an arginyl residue at P1 (rACT L358R) did not inhibit human chymase. Our results agree with these observations indicating an absence of the substrate pathway in the overall mechanism of inhibition. Instead, cleavage of the inhibitory loop must occur to form an irreversible complex with the enzyme trapped as an acyl complex. This is reflected in the rather low turn-over number and high affinity of cathepsin G for ACT-derived substrates.

The conformation of the inhibitory RSL of native ACT is probably more restricted than that of peptides derived from this loop, which should make the loop to bind more tightly and be cleaved even more slowly than the model substrates used here. Thus, the predominance of the substrate pathway over the inhibitory pathway during the interaction of a serpin-proteinase pair seems to be directly correlated with the susceptibility of the serpin inhibitory loop to proteolysis.

Structural Determinants of Substrate Specificity in Cathepsin G—Putative interactions between substrates and the extended binding site of cathepsin G were examined using molecular modeling to gain insight into the structural basis for cathepsin G substrate specificity. A model of the P2-P′ region of the substrate (TPFSAL) of the substrate Abz-TPFSALQ-EDDnp within the active site was constructed based on the P2-P′ region of the ecotin inhibitor bound to crab collagenase, a serine protease in which the structure is very similar to that of cathepsin G. The conformation of the inhibitory loop of ecotin has been suggested to mimic that of a polypeptide substrate (29) and is thus a good template for modeling a substrate complex of cathepsin G. Optimal superimposition of the cathepsin G structure complexed to a transition state analogue inhibitor (Protein Data Bank code 1cgh) (21) over the ecotin-collagenase complex (Protein Data Bank code 1azz) (29) resulted in 210 topologically equivalent α-carbon atoms with a root mean square deviation of 0.86 Å. The inhibitor bound to cathepsin G was removed, and the P2-P′ residues in the ecotin binding loop (Ser-Thr-Met-Ala-Cys) were replaced with the corresponding residues of the hexapeptide portion (Thr-Pro-Phe-Ser-Ala-Leu) of the AC-9p substrate, Abz-TPFSALQ-EDDnp. This procedure yielded an initial cathepsin G-substrate complex, which was energy-minimized as indicated in “Experimental Procedures.”

Interactions that could mediate binding of the hexapeptide model TPFSAL to cathepsin G are shown in the complex model (Fig. 2). Phe at P1 is bound in the S1 pocket defined by the 189–191 and 214–217 segments of cathepsin G. Because Glu226 is present in the S1 pocket, which is hydrogen-bonded to Ala190, the benzyl group of Phe P1 does not completely enter the pocket, but the edge of its phenyl ring makes favorable electrostatic interactions with Glu226, a residue that is thought to be a major structural determinant for the preference of Phe at P1 (21).

The Phe P1 of the modeled substrate was replaced by Leu, Arg, or Lys (Fig. 2) to examine how the S1 site accommodates aromatic and/or hydrophobic as well as positively charged side chains. Glu226 prevents P1 Leu from entering the very far pocket so that it lies at the mouth of the pocket establishing hydrophobic contacts with Phe191, Lys192, and Tyr215. The side chains of Lys and Arg may extend almost fully into the pocket to contact Glu226 through hydrogen bond(s) and electrostatic interactions. The unique location of Glu226 in the S1 subsite of cathepsin G makes it a key residue of the S1 specificity for both Phe and basic residues. In agreement with the experimental results presented here, the modeling studies indicate that the S1 subsite of cathepsin G can accommodate several amino acid side chains without requiring any enzyme conformational change. Therefore, the ability of cathepsin G to accommodate various P1 residues can be explained by the shape and nature of the primary specificity pocket rather than by any structural plasticity of the active site, as in some serine proteases such as α-lytic protease (44). Some flexibility may be possible, however, to optimally bind the substrate because of the absence of the disulfide bond at Cys191-Cys220.

Cathepsin G has a preference for a Pro at P2 like many other chymotryptsin-like enzymes. Although lacking hydrogen bond donor functions, this residue allows a change in the direction of the substrate chain as it threads through the active site, thus
residues in cathepsin G substrates tend to be small amino acids. Lys192 should favor the presence of acidic residues and disfavor basic ones in agreement with data from subsite mapping experiments using a series of natural inhibitory serpins, a structural basis for the preference of Ser at P1 toward the active site. This restricts the residues that can be accommodated in the S2’ pocket to small amino acids such as Ala or Val. The binding energy is provided by residues 39–42 of the substrate and form antiparallel main chain hydrogen bonds with the P2 residue. It is also possible that the positively charged side chain of Arg41 is flexible and that an acidic residue may be accommodated at P2. The proximity of Arg41 may also provide some discrimination at P3.

These observations thus suggest that the 30s loop, because of its unique conformation in cathepsin G, acts as a major discriminating structural determinant of the P1′-P3′ specificity of cathepsin G. For this reason, P3′-elongated substrates, such as the intramolecularly quenched substrates used in this study, are far better substrates than conventional chromogenic or fluorogenic substrates of cathepsin G that have residues only in the P3′ site.

The modeling studies also suggest that P1′-Ser of the substrate may form a hydrogen bond with Ser40 of cathepsin G (distance Oś Ser P1′=Oś Ser40 = 3.4 Å), thus providing a structural basis for the preference of Ser at P1′, as in the natural inhibitory serpins, α1-antichymotrypsin, α1-proteinase inhibitor, and SCFA2. The segment 39–41, belonging to a surface loop (loop 30s), has an unusual conformation that is also found in crab collagenase (29) and chymase (45) but not in most serine proteinases. This conformation results in a narrowing of the S1′ pocket, suggesting why the preferred P1′ residues in cathepsin G substrates tend to be small amino acids.

The binding clefts in the S2′ and S3′ sites appear to be less well defined than for other serine proteinases because of the peculiar backbone path of segment 39–41. Consequently, cathepsin G shows no significant preference for the P2′ and P3′ residues. However, the side chain of Arg41, unlike the equivalent residues in collagenase (29) and chymase (45), does not project toward the solvent but is packed on the protein body toward the active site. This restricts the residues that can be accommodated in the S3′ pocket to small amino acids such as Ala or Val. The binding energy is provided by residues 39–42 of cathepsin G (Fig. 2), which run antiparallel to the modeled substrate and form antiparallel main chain hydrogen bonds with the P2′ residue. It is also possible that the positively charged side chain of Arg41 is flexible and that an acidic residue may be accommodated at P2′. The proximity of Arg41 may also provide some discrimination at P3′.

avoiding steric hindrance with His57 and leading to an optimal positioning of the scissile bond (Fig. 2). The conformation of the P2 Pro in the model substrate obtained after replacement of the Ser residue by Pro in ecotin was very similar to that of the Pro belonging to the inhibitor complexed to cathepsin G. This observation emphasizes the necessity for the substrate or inhibitor backbone of the P2 residue to adopt a conformation allowing an optimal adaptation of the polypeptide chain in the enzyme active site. Examination of the S3′ pocket reveals no obvious binding determinants for the P3′ side chain, although Lys192 should favor the presence of acidic residues and disfavor basic ones in agreement with data from subsite mapping experiments using a series of p-nitroanilide substrates (16).

The modeling studies also suggest that P1′-Ser of the substrate may form a hydrogen bond with Ser40 of cathepsin G (distance Oś Ser P1′=Oś Ser40 = 3.4 Å), thus providing a structural basis for the preference of Ser at P1′, as in the natural inhibitory serpins, α1-antichymotrypsin, α1-proteinase inhibitor, and SCFA2. The segment 39–41, belonging to a surface loop (loop 30s), has an unusual conformation that is also found in crab collagenase (29) and chymase (45) but not in most serine proteinases. This conformation results in a narrowing of the S1′ pocket, suggesting why the preferred P1′ residues in cathepsin G substrates tend to be small amino acids.

The binding clefts in the S2′ and S3′ sites appear to be less well defined than for other serine proteinases because of the peculiar backbone path of segment 39–41. Consequently, cathepsin G shows no significant preference for the P2′ and P3′ residues. However, the side chain of Arg41, unlike the equivalent residues in collagenase (29) and chymase (45), does not project toward the solvent but is packed on the protein body toward the active site. This restricts the residues that can be accommodated in the S3′ pocket to small amino acids such as Ala or Val. The binding energy is provided by residues 39–42 of cathepsin G (Fig. 2), which run antiparallel to the modeled substrate and form antiparallel main chain hydrogen bonds with the P2′ residue. It is also possible that the positively charged side chain of Arg41 is flexible and that an acidic residue may be accommodated at P2′. The proximity of Arg41 may also provide some discrimination at P3′.

These observations thus suggest that the 30s loop, because of its unique conformation in cathepsin G, acts as a major discriminating structural determinant of the P1′-P3′ specificity of cathepsin G. For this reason, P3′-elongated substrates, such as the intramolecularly quenched substrates used in this study, are far better substrates than conventional chromogenic or fluorogenic substrates of cathepsin G that have residues only on the P side.

Another structural determinant that could be important in defining the dual substrate specificity of cathepsin G is Gly216. Its conformation is determined by distal surface loops and is correlated with the P1′ preference in serine proteinases of the chymotrypsin family (44). The backbone conformation of Gly216 in cathepsin G (Φ = −173.66°, Ψ = −160.89°), like that in crab collagenase (Φ = −152.28°, Ψ = −166.56°), a protease having an even broader specificity than cathepsin G (46), is intermediate between that of trypsin (Φ = 177.30°, Ψ = 174.87°) and chymotrypsin (Φ = −112.28°, Ψ = −161.89°).
chymotrypsin ($
\Phi = -165.56^\circ$, $
\Psi = -160.82^\circ$). It has also been demonstrated that variant trypsins (46) lacking a chymotrypsin-like backbone at Gly216 do not accelerate hydrolysis of P \textsubscript{1} Phe-containing substrates. Because the Gly 216 of cathepsin G has a conformation close to that in chymotrypsin, the conformation of Gly216 probably plays an important part in governing the dual trypsin- and chymotrypsin-like specificity of cathepsin G.

We have shown here that sequences derived from a given serpin-reactive site loop may be useful for designing substrates of the corresponding target enzyme. Most of the intramolecularly quenched fluorogenic substrates described here for cathepsin G have greater $k_{cat}/K_m$ ratios than the best substrates currently used to measure cathepsin G activity. They have low $K_m$ values and rather moderate $k_{cat}$ values, which make them suitable for several applications. The specificity constant values determined for the various substrates used in this study indicate that cathepsin G has about 40–60% chymotrypsin activity, 10–15% chymase activity, and is much more active than hK3.

Acknowledgments—We thank Prof. Norman Schechter of the Department of Dermatology at the University of Pennsylvania (Philadelphia) for the gift of human recombinant chymase. We thank Catherine Girardin for technical assistance and Dr. Owen Parkes for editing the English text.

REFERENCES

1. Baggiolini, M., Bretz, U., Dewald, B., and Feigenson, M. E. (1978) Agents Actions 8, 3–10
2. Barrett, A. J. (1981) Methods Enzymol. 80, 561–565
3. Harvima, I. T., Horsmanheimo, L., Naukkarinen, A., and Horsmanheimo, M. (1994) Arch. Dermatol. Res. 287, 61–67
4. Senior, R. M., and Campbell, E. J. (1984) J. Immunol. 132, 2547–2551
5. Watorek, W., Farley, D., Salvesen, G., and Travis, J. (1988) Adv. Exp. Med. Biol. 240, 23–31
6. Bietli, J. G. (1986) in Biology of the Extracellular Matrix (Mecham, R. P., ed) Vol. 1, pp. 217–220, Academic Press, New York
7. LaRosa, C. A., Rohrer, M. J., Benoit, S. E., Rodino, L. J., Barnard, M. R., and Michelon, A. D. (1994) J. Vasc. Surg. 19, 306–318
8. Chertov, O., Ueda, H., Xu, L. L., Tani, K., Murphy, W. J., Wang, J. M., Howard, O. M., Sayers, T. J., and Oppenheim, J. J. (1997) J. Exp. Med. 186, 739–747
9. Yamazaki, T., and Aoki, Y. (1998) Immunology 93, 115–121
10. Shafer, W. M., Pohl, J., Onunka, V. C., Bangalore, N., and Travis, J. (1991) J. Biol. Chem. 266, 112–116
11. Avril, L. E., Di Martino-Ferrer, M., Brillard-Bourdet, M., and Gauthier, F. (1995) FEBS Lett. 367, 251–256
12. Padrines, M., Wolf, M., Walz, A., and Baggiolini, M. (1994) FEBS Lett. 352, 231–235
13. Maison, C. M., Villiers, C. L., and Colomb, M. G. (1991) J. Immunol. 147, 921–926
14. Allen, D. H., and Tracy, P. B. (1995) J. Biol. Chem. 270, 1408–1415
15. Levy, H., and Peinastein, G. (1979) Biochim. Biophys. Acta 567, 35–42
16. Tanaka, T., Minematsu, Y., Reilly, C. F., Travis, J., and Powers, J. C. (1985) Biochemistry 24, 2540–2547
17. Powers, J. C., Tanaka, T., Harper, J. W., Minematsu, Y., Barker, L., Lincoln, D., Crumley, K. V., Fraki, J. E., Schechter, N. M., Lazarus, G. G., Nakajima, K., Nakashino, K., Neurath, H., and Woodburry, R. G. (1985) Biochemistry 24, 2048–2058

FIG. 2. Models of cathepsin G complexed to a hexapeptide substrate. A, the P\textsubscript{3}–P\textsubscript{3} TPFSAL residues of the AC-9p substrate were docked in the active site of cathepsin G as described under “Experimental Procedures.” The residues of cathepsin G forming the S\textsubscript{1} site (189–191 and 214–217) and the loop 30s (residues 39–41) interacting with P\textsubscript{3} residues of the substrates are shown as a thin line. Glu\textsuperscript{226} located at the bottom of the S\textsubscript{1} site and thought to be the key residue defining the P\textsubscript{1} specificity of cathepsin G, is also shown. Residues P\textsubscript{3}–P\textsubscript{3} of the substrate are shown as a thick line. The hatched line indicates the modeled hydrogen bond interaction between the Ser\textsuperscript{40} and the P\textsubscript{1} Ser of the substrate. The P\textsubscript{1} Phe residue of the modeled AC-9p substrate was replaced by P\textsubscript{1} Leu (B), P\textsubscript{1} Arg (C), and P\textsubscript{1} Lys (D). The S\textsubscript{1} site may thus accommodate all of these residues without any conformational changes in the enzyme.
Synthetic Substrates for Cathepsin G

18. Harper, J. W., Ramirez, G., and Powers, J. C. (1981) Anal. Biochem. 118, 382–387
19. Beatty, K., Bieth, J., and Travis, J. (1980) J. Biol. Chem. 255, 3931–3934
20. Bourgeois, L., Brillard-Bourdet, M., Duggerthe, D., Juliano, M. A., Juliano, L., Tremblay, R. R., Dube, J. Y., and Gauthier, F. (1997) J. Biol. Chem. 272, 29590–29595
21. Hof, P., Mayr, I., Huber, R., Kozus, E., Potempa, J., Travis, J., Powers, J. C., and Bode, W. (1996) EMBO J. 15, 5481–5491
22. Schechter, I., and Berger, A. (1967) Biochem. Biophys. Res. Commun. 27, 157–162
23. Wang, M. C., Valenzuela, L. A., Murphy, G. P., and Chu, T. M. (1979) Invest. Urol. 17, 159–163
24. Hirata, J. Y., Cezari, M. H. S., Nakano, C. R., Bosheev, P., Ito, A. S., Juliano, M., and Juliano, L. (1994) Lett. Pept. Sci. 1, 299–308
25. Oliveira, M. C., Hirata, I. Y., Chagas, J. R., Bosheev, P., Gomes, R. A., Figueredo, A. F., and Juliano, L. (1992) Anal. Biochem. 203, 39–46
26. Chagas, J. R., Juliano, L., and Prade, E. S. (1991) Anal. Biochem. 192, 419–425
27. Ermolli, J., Boudier, C., Laine, A., Meyer, B., and Bieth, J. G. (1994) J. Biol. Chem. 269, 29502–29508
28. Schechter, N. M., Jordan, L. M., James, A. M., Cooperman, B. S., Wang, Z. M., and Rubin, H. (1993) J. Biol. Chem. 268, 23626–23633
29. Perona, J. J., Tsu, C. A., Craik, C. S., and Fletterick, R. J. (1997) Biochemistry 36, 5381–5392
30. Schechter, N. M., Sprows, J. L., Schoenberger, O. L., Lazarus, G. S., Cooperman, B. S., and Rubin, H. (1989) J. Biol. Chem. 264, 21308–21315
31. Hayem, A., Marko, D., Laine, A., and Davrl, M. (1988) Adv. Exp. Med. Biol. 240, 97–99
32. Fersht, A. (1985) Enzyme Structure and Mechanism., pp. 105–111, W. H. Freeman and Co., New York
33. Christensson, A., Laurell, C. B., and Lilja, H. (1990) Eur. J. Biochem. 194, 755–763
34. Schick, C., Kamachi, Y., Bartuski, A. J., Cataltepe, S., Schechter, N. M., Pemberton, P. A., and Silverman, G. A. (1997) J. Biol. Chem. 272, 1849–1855
35. Schick, C., Pemberton, P. A., Shi, G. P., Kamachi, Y., Bartuski, A. J., Gornstein, E. R., Brunme, D., Chapman, H. A., and Silverman, G. A. (1998) Biochemistry 37, 5258–5266
36. Sheng, S., Truong, B., Fredricksen, D., Wu, R., Pardee, A. B., and Sager, R. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 499–504
37. Fitzpatrick, P. A., Wong, D. T., Bier, P. J., and Pemberton, P. A. (1996) Protein Eng. 9, 585–589
38. Polanowska, J., Krokoszynska, I., Czapisinska, H., Watarek, W., Dadejz, M., and Otleswski, J. (1996) Biochim. Biophys. Acta 1386, 169–198
39. Bauer, C. A., Thompson, R. C., and Blout, E. R. (1970) Biochemistry 15, 1291–1295
40. Sanker, S., Chandrashekharan, U. M., Wilk, D., Glynias, M., Karnik, S. S., and Husson, A. (1997) J. Biol. Chem. 272, 2963–2968
41. McRae, B., Nakajima, K., Travis, J., and Powers, J. C. (1980) Biochemistry 19, 3973–3978
42. Lomas, D. A., Stone, S. R., Llewellyn-Jones, C., Keogan, M. T., Wang, Z., Rubin, H., Carrell, R. W., and Stockley, R. A. (1995) J. Biol. Chem. 270, 23437–23443
43. Duranton, J., Adam, C., and Bieth, J. G. (1998) Biochemistry 37, 11239–11245
44. Perona, J. J., and Craik, C. S. (1995) Protein Sci. 4, 337–369
45. McGrath, M. E., Mirzadegan, T., and Schmid, B. F. (1997) Biochemistry 36, 14318–14324
46. Tsu, C. A., Perona, J. J., Fletterick, R. J., and Craik, C. S. (1997) Biochemistry 36, 5393–5401