New chromogenic substrates of human neutrophil cathepsin G containing non-natural aromatic amino acid residues in position P₁ selected by combinatorial chemistry methods

Magdalena Wysocka · Anna Łęgowska · Elżbieta Bulak · Anna Jaśkiewicz · Hanna Miecznikowska · Adam Lesner · Krzysztof Rolka

Abstract Specificity of human cathepsin G was explored using combinatorial chemistry methods. Deconvolution of a tetrapeptide library, where 5-amino-2-nitrobenzoic acid served as a chromophore attached at the C-terminus, yielded the active sequence Phe-Val-Thr-Tyr-Anb₅,₂-NH₂. This sequence was used for a second-generation library with the general formula Ac-Phe-Val-Thr-X-Anb₅,₂-NH₂, where position X was replaced with several amino acids: L-pyridylalanine (Pal), 4-nitro-L-phenylalanine (Nif), 4-amino-L-phenylalanine (Amf), 4-carboxy-L-phenylalanine (Cbf), 4-guanidine-L-phenylalanine (Gnf), 4-methyloxycarbonyl-L-phenylalanine (Mcf), 4-cyano-L-phenylalanine (Cyf), Phe, Tyr, Arg and Lys. Specificity ligand parameters, \( k_{\text{cat}} \) and \( K_M \), with human cathepsin G were determined for all chromogenic substrates synthesized. The highest value of the specificity constant \( (k_{\text{cat}}/K_M) \) was obtained for a substrate with the Gnf residue in position P₁. This peptide was 10 times more active than the second most active substrate which contained the Amf residue. The following order of potency was established: Gnf > Amf > Tyr = Phe > Arg = Lys > Cyf. Substrate specificity for cathepsin G is greatly enhanced when an aromatic side chain and a strong positive charge are incorporated in residue P₁.

Keywords Cathepsin G · Chromogenic substrates · Peptidomimetics

Human cathepsin G is a member of the serine protease family. Along with leukocyte elastase and protease-3, human cathepsin G is produced, accumulated and secreted by various blood cell types involved in the degradation of dead tissues during inflammatory reaction in the human body [1,2]. They also take part in early response to the invasion of microorganisms [3,4]. Cathepsin G is located mainly in neutrophils but can also be found in mastocytes and other cells [5]. In healthy people its proteolytic activity is controlled by \( \alpha_1 \)-antichymotrypsin, an inhibitor that belongs to the serpin family [6].

Cathepsin G displays a unique substrate specificity, combining both chymotrypsin-like and trypsin-like specificity. There are few enzymes known up to date sharing the same broad spectrum of recognition and hydrolysis of the peptide bond. Duodenase and crab fiddler collagenase are the straight examples of such proteases [7,8]. In the case of cathepsin G this phenomena is a consequence of the arrangement of an S₁ substrate pocket. As described by Hof et al. [9], the structure of the substrate pocket of cathepsin G is similar to that of chymotrypsin. The difference is that side chain Glu226, located at the bottom of the enzyme’s substrate pocket, divides it into subcompartments that can accommodate the positively charged side chains Lys, Arg or the phenyl ring of Phe or Tyr. Since there are no systematic structure–activity relationship studies on cathepsin G substrates (most of cathepsin G substrates were designed for chymotrypsin), the main goal of our research was to synthesize substrates that, after undergoing selective hydrolysis at the C-terminus by this enzyme, would release a quantifiable chromophore. We decided to apply combinatorial chemistry methods. According to our best knowledge, there are no literature reports that utilized this approach to select cathepsin G substrate(s).

In order to find sensitive chromogenic substrates of cathepsin G, a two-step procedure was applied. First we designed and synthesized the tetrapeptide library shown below:
X₄₋X₃₋X₂₋X₁₋Anb₅₋₂-NH₂

where the following amino acid residues were present in position X₁ = Lys, Arg, Ala, Val, Leu, Nle, Ser, Phe, Tyr; X₂ = all proteinogenic amino acids except Cys; X₃ = Ala, Pro, Val, Phe, Ser, Lys, Arg, Asp; X₄ = Phe, Lys, Val, Ile, Abu, Ala, Asp.

Such a library consisted of 9,576 tetrapeptides. In X₃ and X₄ (corresponding to substrate positions P₃ and P₄, respectively) a smaller representative set of amino acids was included containing aliphatic (Ala, Abu, Val, Ile, Pro), aromatic (Phe), basic (Lys, Arg) and acidic (Asp) side chains. The 5-amino-2-nitrobenzoic amide unit (Anb₅₋₂-NH₂) served as a chromophore [10]. The library was synthesized on solid support using mix and split method. Its deconvolution against the experimental enzyme was carried out applying an iterative approach, as described in our previous work [11].

In the second-generation substrates we focused our attention on position P₁, in which the following amino acid residues were introduced: l-pyridylalanine (Pal), 4-nitro-l-phenylalanine (Nif), 4-amino-l-phenylalanine (Amf), 4-carboxy-l-phenylalanine (Cbf), 4-guanidine-l-phenylalanine (Gnf), 4-methylxycarbonyl-l-phenylalanine (Mcf), 4-cyano-l-phenylalanine (Cyf), Phe, Tyr, Arg and Lys (Fig. 1). All these substrates were synthesized by the classical “one-bead–one-compound” approach.

Materials and methods

Solid-phase peptide synthesis

The peptide library and individual peptides were synthesized manually by solid-phase synthesis using Fmoc chemistry as described previously [12]. TentaGel S RAM (substitution 0.25 meq/g) (RAPP Polymere, Germany) was used as solid support. 5-Amino-2-nitrobenzoic acid was coupled by the DIPCDI/HOBt method. The peptides were cleaved from the resin using a TFA/phenol/triisopropylsilane/H₂O mixture (88:5:2:5, v/v/v/v) [13]. Peptide libraries were used without further purification.

Preparation of the peptide library

The peptide library was synthesized by the portioning-mixing method [14,15] as described earlier [16], applying the same reagents as described above. Initially, 10 g of the solid support (TentaGel S RAM) was used. Five-fold molar amino acid was used for the coupling.

Screening of the peptide library for cathepsin G substrate activity

Cathepsin G (Biocentrum Kraków, Poland) was isolated from human neutrophils [17]. Deconvolution of the peptide library synthesized was performed by the iterative method in solution [18]. Stock solution was prepared by dissolving 5 mg of the peptide library in 250 µL of DMSO from which 10 µL was used for analysis. Enzymatic hydrolysis of the peptide library was performed in 0.1 M Tris–HCl (pH 7.5) buffer with 500 mM of NaCl at 25°C. Measurements were carried out at enzyme concentrations of 6 × 10⁻⁷ M. Absorbance of Anb₅₋₂-NH₂ at 405 nm was recorded as a function of time. The calculated initial rates were used as measures of the substrate activity of the peptide library investigated. Each experiment was repeated at least three times. Data differing from the mean value by more than 10% were rejected.

Physicochemical properties of the peptide library

In order to prove that the library synthesized contained the designed peptides, HPLC and MS analyses were performed for randomly chosen sublibraries at each stage of synthesis and deconvolution. A Pro Star system (Varian, Australia) equipped with a Kromasil 100 C₈ column (8 × 250 mm²) (Knauer, Germany) and a UV-VIS detector were used. A linear gradient from 10 to 90% B within 30 min was applied (A: 0.1% TFA; B: 80% acetonitrile in A). The peptides analyzed were monitored at 226 nm. Mass spectra were recorded using a Biflex III MALDI TOF mass spectrometer (Bruker, Germany). Selected sublibraries were additionally characterized using a method described by Furlong et al. [19]. Herein, sublibraries with fixed amino acid residues (e.g. Val and Ala in position P₃) were analyzed by mass spectrometry. For each library (PheValX₂X₁Anb₅₋₂-NH₂ and PheAlaX₂X₁Anb₅₋₂-NH₂), the most intense signal corresponding to the centre of distribution was identified. The presence of substrates with the lowest and highest molecular weights was also confirmed. The molecular weight shift between two centers of distribution was compared with the molecular weight difference between Val and Ala.

Kinetic studies

All measurements were performed using a Cary 3E spectrophotometer (Varian, Australia). The concentration of bovine β-trypsin stock solution was determined by titration with p-nitrophenyl-p’-guanidinobenzoate (NPGB). Such trypsin stock solution was used for titration of SFTI-1, the mutual inhibitor of both bovine β-trypsin and cathepsin G. Solution