Specificity Studies on Retroviral Proteinase from Myeloblastosis-Associated Virus

Petr Strop,*,‡ Jan Konvalinka,‡ Dalibor Stys,‡ Libuse Pavlickova,‡ Ivo Blaha,‡ Jiri Velek,‡ Miloslav Travnicek,§ Vladimir Kostka,‡ and Juraj Sedlacek§

Institute of Organic Chemistry and Biochemistry, Czechoslovak Academy of Science, 166 10 Prague, Czechoslovakia, and Institute of Molecular Genetics, Czechoslovak Academy of Science, 166 37 Prague, Czechoslovakia

Received April 25, 1990; Revised Manuscript Received November 1, 1990

ABSTRACT: The specificity of the p15 proteinase of myeloblastosis-associated virus (MAV) was tested with nonviral high molecular weight substrates and with synthetic peptides. Peptides with sequences spanning known cleavage sites in viral polyproteins of Rous sarcoma virus (RSV) and avian leukemia viruses, as well as in BSA and HSA, were synthesized, and the rate of their cleavage by the MAV proteinase was compared. Synthetic peptides require for successful cleavage at least 4 residues at the N-terminal side and 3 residues at the C-terminal side. The proteinase shows a preference for hydrophobic residues with bulky side chains (Met, Tyr, Phe) in P3, although Arg and Gln can also be accepted. Small hydrophobic residues are required in P2 and P2′, and large hydrophobic residues (Tyr, Met, Phe/*p*-nitro-Phe) are preferred in both P1 and P1′. The difference between the specificity of the p15 proteinase and that of the HIV-1 proteinase mostly pertains to position P2′ of the substrate, where bulkier side chains are accepted by the HIV-1 proteinase (Richards et al., 1990). A good chromogenic substrate for the MAV and RSV proteinases was developed and used to further characterize the MAV proteinase activity with respect to ionic strength and pH. The activity of the proteinase is strongly dependent on ionic strength and pH. Both the k_{cat} and K_m values contribute to a higher cleavage efficiency at higher salt concentrations and show a bell-shaped pH dependence curve with a sharp maximum at pH 5.5 (k_{cat}) and 6.5 (K_m).

The replication of all retroviruses involves proteolytic processing of the de novo synthesized viral polyproteins into matured proteins, which subsequently form new viral particles (Krausslich & Wimmer, 1988). The maturation step is of vital importance: all proteinase-deficient mutants are noninfectious (Weiss et al., 1982). Much effort has been devoted to studies on these enzymes, the prospective goal being the possibility of affecting the processing of viral proteins and hence virus maturation in infected cells.

Several authors reported recently on the cloning and expression of retroviral proteinases (Mous et al., 1988; Giam & Boros, 1988; Graves et al., 1988; Sedlacek et al., 1988; Seelmeier et al., 1988; Krausslich & von der Helm, 1988; Darke et al., 1989) and on the results of the first inhibition studies (Seelmeier et al., 1988; Billich et al., 1988; Moore et al., 1988; Richards et al., 1989a,b) with these enzymes. Detailed substrate specificity studies are relatively rare so far (Billich et al., 1988; Darke et al., 1988; Copeland & Oroszlan, 1988). Proteinases of retroviruses were regarded as enzymes endowed with an unusual specificity difficult to understand: these enzymes cleave viral polyproteins at very specific yet differing sequences. Data on the crystal structures of the HIV-1,[1] RSV, and MAV proteinases (Miller et al., 1989a,b; Foundling et al., 1989) provide evidence that proteinases encoded by retroviruses are structurally homologous to aspartate proteinases. The active form of retroviral proteinases, however, is the dimer, an oblong ellipsoid, formed almost symmetrically by two monomers. Each monomer contributes one catalytic aspartate to the catalytic site, a wide yet short cleft to which the access is hindered, unlike in other aspartate proteinases, by two long and flexible loops known as "flaps". The flaps are hydrogen bonding with the substrate (Miller et al., 1989b) and also form a part of subsite S2′. The structure of the flap of both the HIV-1 proteinase and also of its complex with the inhibitor is known. The flap of the RSV and MAV proteinases is larger and considerably more disordered in crystals.

In this report we present studies on the substrate specificity of proteinase p15 encoded by the *gag* gene of MAV, using protein substrates and custom-synthesized peptides mimicking sequences around processing sites in MAV and RSV virus polyprotein precursors.

Experimental Procedures

The MAV p15 proteinase was expressed in *Escherichia coli* and purified to homogeneity as described elsewhere (Sedlacek et al., 1988, 1989; Strop et al., 1989a).

Cleavage and Sequencing of Protein Substrates. HSA or BSA (Fluka, BRD) was dissolved to 1% solution in citrate buffer, pH 5, containing 0.5% SDS and 0.3% 2-mercaptoethanol, and the solution was boiled for 2 min. Subsequently, each protein solution was diluted with the buffer to a final 15 μM concentration and to a final 1.3 μM concentration of proteinase. The cleavage was allowed to proceed at 40 °C for 3 h. The products were separated by SDS–PAGE (10% gel) and electroeluted. SDS was removed by methanol precipitation. Samples of products were purified by repeated dilution–concentration with an Amicon B15 membrane concen-

* To whom correspondence should be addressed.
‡ Institute of Organic Chemistry and Biochemistry.
§ Institute of Molecular Genetics.

[1] Abbreviations: MAV, myeloblastosis-associated virus; RSV, Rous sarcoma virus; HSA, human serum albumin; BSA, bovine serum albumin; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; t-Boc, *tert*-butyloxycarbonyl; Z, benzyloxycarbonyl; RP-HPLC, reversed-phase high-performance liquid chromatography; FAB, fast atom bombardment; TFA, trifluoroacetic acid; HIV, human immunodeficiency virus; PheSta, PhenylStatine [3(*R*)-hydroxy-4(*S*)-amino-5-benzylpentanoic acid], *, expected cleavage site. Formal nomenclature of Berger and Schechter (1970) was used to describe proteinase specificity and enzyme–substrate interactions. Enzyme binding subsites (S) for side chains of substrates (P) are numbered consecutively from the cleavable bond. C-Terminal side chains (P′) and subsites (S′) are distinguished by primes.

0006-2960/91/0430-3437$02.50/0

Table I: Comparison of Rates of Cleavage of Synthetic Peptides Mapping Processing Sites of Viral Polyproteins and Cleavage Sites in BSA and HSA[a]

pcp-tide no.	P subsite 9 8 7 6 5 4 3 2 1		P′ subsite 1 2 3 4 5 6 7 8	rel rate of cleavage A[b]	rel rate of cleavage B[c]	origin of cleavage site
1		GlyProValValPhe(NO_2)	* SerValSerThrLysGluLeuPro	100	66	MAV ideal substrate
2		LysThrValGlyThrSerCysTyr	* HisCysGlyThrAlaIleGlyCys	0		P19/P2a 155-156 gag
3		ThrAlaIleGlyCysAsnCysAla	* ThrAlaSerAlaProProProPro	0		P2a/P2b 166-167 gag
4	AlaSerAlaProProProProTyrValGly		* SerGlyLeuTyrProSer	50		P2b/P2c 175-176 gag
5		ThrGlyProProValValAlaMet	* ProValValIleLysThrGluGly	10		P10/P27 239-240 gag
6		AspGlnGlyIleAlaAlaAlaMet	* SerSerAlaIleGlnProLeuIle	0		P27/P3 479-480 gag
7		ThrMetArgGluLysValLeuThr	* SerSerAlaArgGlnArgLeuArg	0		BSA 189-190 (+)
8		SerAlaIleGlnProLeuIleMet	• AlaValValAsnArgGluArgAsp	0		P3/P12 488-489 gag
9		GlyProGluProProAlaValSer	* LeuAlaMetThrMetGluHisLys	80		P12/P15 577-578 gag
10		SerGlnAlaThrPheGlnAlaTyr	* ProLeuArgGluAlaValIleLeu	90	52	P63/P32 572-573 pol
11		AspGluAlaSerProLeuPheAla	* GlyIleSerAspThrIlePro	0		P32/P4 858-859 pol
12		CysGluValThrGlyValArgAla	* AspValHisLeuLeuGluGlnPro	35		P7/P40 64-65 env
13		LeuGluLysSerHisCysIleAla	* GluValGluLysAspAlaIlePro	50		BSA 289-290 (+)
14		ArgThrGlyIleArgArgLysArg	* SerValSerHisLeuAspAspThr	0		P40/P23 405-406 env
15		ArgLeuValArgProGluValAsp	* ValMetCysThrAlaPheHisAsp	0		HSA 121-122 (+)

[a] The HPLC assay conditions of the cleavage are given under Experimental Procedures. Peptides 2, 3, 6–8, 11, 14, and 15 are not cleaved under the conditions given; peptide 15 is cleaved in 2 M NaCl only. (4-Nitrophenyl)alanine was introduced into position P1 of some peptides in order that their cleavage might be directly monitored. [b] 120-min digestion. [c] 30-min digestion.

trator. The sequencing of the fragments and peptides was carried out in an Applied Biosystems AB 410 gas-phase protein sequencer using Polybrene and on-line HPLC detection.

Peptide Synthesis. All peptide substrates were synthesized by the solid-phase method with a cross-linked (chloromethyl)polystyrene resin. The amino groups were protected by t-Boc groups, and the side chains of Arg, Asp, Glu, Cys, His, Lys, Tyr, and Trp were protected by tosyl, benzyl, *p*-methylbenzyl, t-Boc, Z, benzyl, and formyl groups, respectively. The peptides were cleaved from the support in HF by the low–high method (Tam et al., 1983); they were purified by RP-HPLC and characterized by amino acid analysis and FAB mass spectrometry.

HPLC Assays. Solutions (0.370 mM) of synthetic peptides in phosphate–citrate buffer, pH 6.0, were digested by a 3.3×10^{-6} M solution of the proteinase for different periods of time at 40 °C. The reaction was stopped by the addition of an equal volume of 20% acetic acid. The cleavage was followed by RP-HPLC in the system methanol–H_2O. The cleavage products were collected and were identified by amino acid analysis and sequencing. The degree of cleavage was determined by integration of the areas under the corresponding peaks.

Spectrophotometric Assays. The cleavage of substrates with a *p*-nitro-Phe residue in P1′ was carried out in sodium phosphate buffer, pH 6.0, containing 1.9 M NaCl, at 40 °C. The enzyme concentration determined by active site titration as described below was 0.08 μM and the substrate concentration varied between 6 and 120 μM. The reaction rate was determined from the change in absorbance at 305 nm in a Hi Tech SF-51 stopped-flow or Aminco DW2000 spectrophotometer. The K_m and k_{cat} values were determined from the double-reciprocal Lineweaver–Burk plot and by a nonlinear regression program. The values showed standard deviations that were 5–18% of the value measured. Concentration of the active form of the enzyme was measured by active-site titration with the tight-binding inhibitor ($K_i = 5 \times 10^{-9}$ M) ProProCysValPheStaAlaMetThrMet [PheSta = 3(*R*)-hydroxy-4(*S*)-amino-5-benzylpentanoic acid] in 1.9 M NaCl, pH 6. The inhibitor, in the concentration range 0–0.3 μM, was preincubated for 15 min with the proteinase solution, and the absorbance change at 305 nm was measured after the addition of the substrate AlaThrHisGlnValTyr*Phe(NO_2)-ValArgLysAla. Since the enzyme can be active as a dimer only (Miller et al., 1989a), k_{cat} values are calculated by using the concentration of the dimer.

Protein Concentration. The method of Bradford (1978) was used. Alternatively, protein concentration was calculated from amino acid analysis performed in a Durrum D 500 amino acid analyzer.

Results and Discussion

Cleavage of Nonviral Proteins. Information on the sequence specificity of the proteinase was obtained from cleavage experiments with serum albumins (HSA and BSA) and subsequent N-terminal sequencing of the major fragments. The major cleavage of BSA occurs at the Ala(289)–Glu(290) bond in the sequence ...SerHisCysIleAla*GluValGluLysAspAla.... The sequence around this cleavage site is homologous to the P7/P40 (64–65 *env*) processing site of the RSV polyprotein with regard to positions P1, P1′, and P2′ (see Table I). The proteinase also cleaves the bond Thr(189)–Ser(190) in the BSA sequence ...GluLysValLeuThr*SerSerAlaArgGlu... homologous to the 479–480 *gag* polyprotein processing site in positions P1′, P2′, P3′, and P5′. In HSA the proteinase cleaves preferentially the bond Asp(121)–Val(122) in the sequence

Table II: Effect of Peptide Chain Length on Its Cleavage by MAV Proteinase[a]

peptide no.	P subsite	P′ subsite	rel rate of cleavage
	8 7 6 5 4 3 2 1	1 2 3 4 5 6 7 8 9	
16	SerGlnAlaThrPheGlnAlaTyr	• ProLeuArgGluAlaValIleLeuLys	6
17	ThrPheGlnAlaPhe(NO_2)	• ProLeuArgGluAlaVal	33
18	ThrPheGlnAlaPhe(NO_2)	• ProLeuArgGluAlaVal	42
19	PheGlnAlaPhe(NO_2)	• ProLeuArgGluAlaVal	20
20	GlnAlaPhe(NO_2)	• ProLeuArgGlnAlaVal	0
21	AlaPhe(NO_2)	* ProLeuArgGlnAlaVal	0
22	AlaThrPheGlnAlaPhe(NO_2)	• ProLeuArgGlu	14
23	AlaThrPheGlnAlaPhe(NO_2)	• ProLeuArg	6
24	AlaThrPheGlnAlaPhe(NO_2)	• ProLeu	0
	Chromogenic substrate		
25	AlaThrHisGlnValTyr	• Phe(NO_2)ValArgLysAla	100

[a] Solutions (0.25 mM) of peptides in 4-morpholineethanesulfonate buffer, pH 6.0, with 0.5 M NaCl were digested for 30 min at 38 °C by 2.5 μM p15.

ArgProGluValAsp*ValMetCysTyrAla.

Cleavage of Synthetic Peptides Derived from Processing Sites in Viral Polyproteins. The rate of cleavage of synthetic peptides, designed according to known processing sites in viral polyproteins, and of peptides mimicking sequences cleaved in BSA and HSA is compared in Table I.

The sequence of peptide 1 in the table contains the residues most frequently occurring in the individual positions (Strop et al., 1989b). Peptide 1 is cleaved faster than any other peptide mimicking a native viral sequence. The Tyr–Pro bonds in peptide 10 spanning the P63/P32 boundary (572–573 *pol*) and peptide 9 spanning the Ser–Leu bond at the N-terminus of the proteinase (P12/P15 boundary, 577–578 *gag*) are hydrolyzed at almost the same rate as the "optimized" sequence. Approximately at a rate 2 times slower proceeds cleavage of the Gly–Ser bond in peptide 4 spanning the processing site between P19 and P10 in RSV (P2b/P2c, 175–176 *gag*) and also cleavage of the Ala–Glu bond in peptide 13 based on the 289–290 cleavage site in BSA. The lowest was the rate of cleavage of the Met–Pro bond in peptide 5 mimicking the P10/P27 boundary (*gag* 239–240) in RSV. Nine peptides were not cleaved at all in solutions of low ionic strength, on which the proteinase activity is highly dependent (see below). Peptide 14 spanning the P40/P23 (405–406 *env*) cleavage site was split at a high salt concentration only (1.9 M NaCl). The proteinase also cleaves the peptide corresponding to the P7/P40 boundary (64–65 *env*) and a sequence in BSA that is homologous to P7/P40 in positions P7, P1, P2′, and P8′ and with Glu instead of Asp in P1′ (Table I). This finding is not contradictory to the observation that the *env* polyprotein was also cleaved by a proteinase-defective viral mutant since correct processing (i.e., N-terminal sequence) was not confirmed in this work (Crawford & Goff, 1985).

Cleavage of Conformationally Restricted Peptide Substrates. For experiments intended to cast light on the possibility that the proteinase may prefer looplike structures we synthesized linear and cyclic substrates designed according to the two most rapidly cleaved peptides (1 and 10 in Table I).

The only peptides cleaved by the proteinase were linear peptides *Cys*AsnAlaGluAlaThrHisGluAlaPhe(NO_2)*ProValSer*Cys* and *Cys*TrpProAlaGluAlaThrHisGlnAlaPhe(NO_2)*ProValSerAsn*Cys*. When the flexibility was restricted by disulfide bridges between C- and N-terminal Cys residues these peptides were not cleaved. The other peptides not cleaved involving Ser or Thr in the P1′ site were ThrAlaThrAlaThrProValValPhe(NO_2)*SerValThrAlaThrAla and ThrAlaThrAlaThrAlaProGlyValPhe(NO_2)*ThrAlaThrAla. Since it was assumed that the structure of peptides may play a role in the specificity of the proteinase (Kay & Dunn, 1990) we have investigated the conformations of 10 linear peptide substrates of MAV and HIV proteinases in solution by 2D NMR (Strop et al., unpublished results). We have found that good substrates are predominantly in random conformation. When peptides are bound to the proteinase they exist most probably in an extended conformation from P3 to P3′. This assumption can be made on the basis of known X-ray structure of the HIV-1 proteinase complex with various inhibitors (peptide analogues) (Miller et al., 1989b), preliminary NMR data for MAV proteinase (Strop et al., unpublished results), and on molecular modeling of MAV proteinase and substrate.

Minimal Substrate Size. The effect of peptide chain length on its cleavability was examined with peptides spanning the 572–573 *pol* polyprotein P63/P32 cleavage site (peptide 10 in Table I). The number of amino acid residues on one side of the cutoff site was kept constant while the chain length on the other side was stepwise shortened to a minimum of 2 residues. The proteinase apparently requires seven sites (S4–S3′) occupied for sufficient cleavage to proceed, four residues on the N-terminal side and at least three at the C-terminus (Table II). The substrate that has four residues on both sides is cleaved considerably better. Essentially the same results have been reported previously for the HIV-1 and RSV proteinases (Moore et al., 1989; Kotler et al., 1988).

Specificity Studies with Peptides Derived from pol P63/P32 Processing Site of RSV. A series of peptides based on the chromogenic substrate and on the *pol*63/32 cleavage site was synthesized with variations of the individual residues in positions P4–P4′ (Table III) to investigate the subsite specificity and substrate symmetry. The fast cleavage was observed with

Table III: Subsite Specificity of MAV Proteinase: Relative Rates of Cleavage of Peptides Derived from *pol* 63/32 Sequence[a]

peptide no.	P subsite 6 5 4 3 2 1	P′ subsite 1 2 3 4 5	rel rate of cleavage
26[b]	AlaThrPheGlnAlaPhe(NO_2)*ProLeuArgGluAla		6.3
27[c]	Tyr*Phe(NO_2)		0
28[d]	Tyr*	AlaGlnPheThr	88.5
29[e]	Tyr*Phe(NO_2)	AlaGlnPheThr	12.9
30[f]	LysAlaGluArgLeuTyr *		0
25[g]	AlaThrHisGlnValTyr*Phe(NO_2)	ValArgLysAla	51.8
31[h]		GlnHisThr	54.2
32	P1,P1′ Met*	Pro	52.0
33	P1,P1′ Phe(NO_2)*	Pro	0
34	P1′ *	Leu	0
35	P1 Met *		75.3
36	P1 Phe *		48.1
37	P1 Ser *		9.4
38	P1 Arg *		0
39	P1 Glu *		0
40	P2 Ile		36.3
41	P2 Ala		12.2
42	P2 Leu		2.5
43	P2 Phe		0
44	P2 Gly		0
45	P3 Glu		50.0
46	P3 Met		71.5
47	P3 Arg		53.7
48	P3 Asp		32.2
49	P3 Asn		30.3
50	P3 Val		3.0
51	P3 Pro		0
52	P3 Tyr		45.0
53	P4 Pro		100.0

[a] The relative rates of cleavage of palindromic peptides derived from the *pol* 63/32 processing site and of peptides derived from the chromogenic substrate based on the *pol* 63/32 sequence with variation of residues in P1, P2, P3, and P4 positions are listed. Amino acids different from the original peptide are shown only. Original peptide sequence is underlined. Peptide solutions (0.13 mM) in phosphate buffer, pH 6.0, were digested for 30 min by 0.12 μM p15 proteinase at 38 °C and analyzed by HPLC. [b] *pol* P63/32 sequence. [c] P1′ Pro → Phe(NO_2). [d] P2–P5 palindromic sequence. [e] P1–P5 palindromic sequence. [f] P2′–P5′ palindromic sequence. [g] Chromogenic substrate. [h] P1–P5 palindromic sequence.

peptide 28 with Tyr–Pro in P1–P1′ and palindromically arranged residues in positions 2–6. An almost identical peptide with Phe(NO_2) instead of Pro (peptide 29) in P1′, i.e., symmetrical with regard to positions 2–5, is cleaved at a substantially lower rate. In peptide 30 with the C-terminal sequence in P2′–P5′ preserved and with the same residues yet in reversed order in positions P5–P2, the assumed scissile bond Tyr–Pro is not cleaved at all.

The data in Table III show that the combination of the residues in positions P2–P1 and P1′–P2′ has a decisive effect on cleavage. The substrate specificity of the proteinase cannot be explained solely in terms of structural requirements of the individual binding pockets but rather by assuming a concerted

Table IV: Concerted Effect of Side Chain in Positions P2, P1, P1′, and P2′ on the Cleavage of Homologous Peptides by MAV Proteinase[a]

peptide no.	P2	P1	P1′	P2′	cleavage
10	Ala	Tyr	Pro	Leu	+
27	Ala	Tyr	Phe(NO_2)	Leu	−
25	Val	Tyr	Phe(NO_2)	Val	+
33	Val	Phe(NO_2)	Pro	Val	−

[a] Arrows indicate variations in positions P2 through P2′. + marks more than 2% of cleavage of a peptide. For conditions see Table III and Experimental Procedures.

effect of several side chains of the substrate. We may thus postulate that there is a large binding cleft in the structure of the proteinase, which can adopt various orientations.

The results of cleavage experiments with peptides 10 (Table I), with the P2–P2′ sequence –AlaTyr*ProLeu–, and 27, with the sequence –AlaTyr*Phe(NO_2)Leu– (Table III), may serve as an example. Peptide 10 is cleaved efficiently; peptide 27, not at all. On the other hand, peptide 25, with the sequence –ValTyr*Phe(NO_2)Val–, is rapidly cleaved, whereas peptide 33, –ValPhe(NO_2)*ProVal–, is resistant (see Table IV for schematic presentation). If, however, the valines in the P2 positions are replaced by alanines (peptide 28, –AlaTyr*ProAla–, or peptide 32, where Met is introduced into P1 instead of Tyr), a fast cleavage can be observed (Table III).

From the results obtained, the following conclusions can be drawn:

(i) The structural requirements of the S2–S2′ binding subsites of the proteinase (Berger & Schechter, 1970) cannot be considered independently; when S2 and S2′ are occupied by valines, the residues best fitting into S1 are methionine, tyrosine, and (*p*-nitrophenyl)alanine. The specificity of the MAV proteinase with regard to the S1 and S1′ pockets is similar to that of the HIV-1 proteinase, which cleaves best substrates with Met-Met, Phe-Phe(NO_2), or Tyr-Pro in positions P1 and P1′. The difference in the specificity of the HIV-1 and MAV proteinases pertains mostly to subsites S2 and S2′ (Strop et al., 1989b).

(ii) The S2 and S2′ binding pockets play the key role in the specificity of the MAV proteinase. The side chains of amino acids in P2 probably interact with the flap region [residues 60–70 of MAV covering the catalytical cleft (Miller et al., 1989b)] of the molecule. A correct explanation is difficult since this region is in the disordered part of the crystals and is not seen in the X-ray crystal structures (Miller, 1989a; Foundling et al., 1989). Moreover, it may be also partly involved in the crystal packing. The S2 pocket of the MAV proteinase can accommodate confortably small hydrophobic residues like valine. Peptides with Ile and Leu in P2 are still cleaved but at a slower rate. The Phe residue (peptide 43) is probably too bulky, while Gly (peptide 44) is on the contrary too small for the necessary interactions to occur. The S2′ binding subsite can accommodate aliphatic side chains of Ala, Val, or Leu when Pro is in P1′. Leu is not accepted in S2′ when the bulkier Phe(NO_2) is in S1′.

(iii) The S3 pocket of the MAV proteinase is also considerably large. The best results were achieved when Met (peptide 46) was in P3. Still accepted yet less favorably in this position are ARg, Gln, or Tyr, while valine or substrates with proline are almost completely resistant to cleavage.

The inhibitor that matches the specificity requirements for S4, S3, S2, and S1 subsites, ProProCysValPheStaAlaMet-

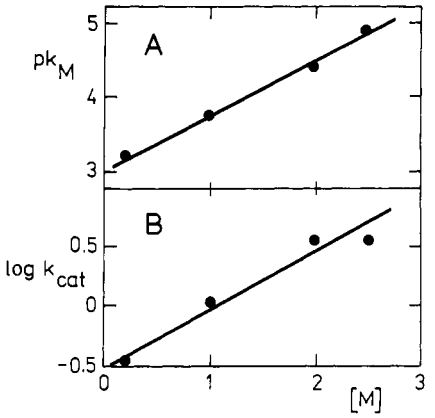

FIGURE 1: Effect of ionic strength on hydrolysis of synthetic peptides by MAV proteinase. Dependence of pK_m (A) and log k_{cat} (B) for peptide 25, AlaThrHisGlnValTyr*Phe(NO_2)ValArgLysAla, on concentration of NaCl. Spectrophotometric assay conditions are given under Experimental Procedures.

ThrMet, was found to be a tight-binding inhibitor ($K_i = 5 \times 10^{-9}$ M) useful for an active-site titration.

Effect of Salt Concentration. We observed a significant dependence of the rate of cleavage of two synthetic peptides, AlaThrHisGlnValTyr*Phe(NO_2)ValArgLysAla (25) and GlyProValValPhe(NO_2)*SerValSerThrLysGluLeuPro (1), on the concentration of NaCl. Salt concentration enhances the rate of cleavage of peptides through a decrease in K_m and through an increase in k_{cat} (Figure 1). The lower K_m values observed at higher salt concentrations can be accounted for by an increase of the hydrophobic binding of the substrate in the catalytic cleft. We have found a preference for hydrophobic side chains in S1, S1′, S2, S2′, and S3 of the MAV protease binding subsites. The binding pocket is largely hydrophobic in both the MAV or the RSV proteinase and in the HIV-1 proteinase as documented by X-ray data (Miller et al., 1989; Foundling et al., 1989).

The increase of the k_{cat} values at higher salt concentrations is not caused by a higher concentration of the active dimer. Its active-site titration carried out with the tight-binding inhibitor ProProCysValPheStaAlaMetThrMet in 1.5, 2, and 3 M NaCl revealed approximately the same concentration of the active dimer in all three NaCl solutions.

pH Optimum. The pH profile of the activity of the MAV proteinase was tested with pig tubulin and with three peptides (1, 36, and 53). For all four substrates the pH optimum of the MAV proteinase was found to be around pH 6. The pH dependence of k_{cat} and K_m for the last two peptides is shown in Figure 2. Both peptides display a very similar pH dependence of K_m: a decrease of the K_m from pH 4 toward neutral pH, a sharp minimum at pH 6.5, and a dramatic drop in binding of peptide substrates above pH 6.5. The k_{cat} values show bell-shaped curves with maxima approximately 1 pH unit lower than the corresponding minima of K_m. A plausible explanation of this pH dependence of k_{cat} frequently observed with classical aspartic proteinases is that the decrease of k_{cat} at higher pH is a result of the ionization of the catalytically active carboxylates. It has been proved that an active aspartic proteinase is formed of two domains or two monomers, each providing one carboxylate for the active site. These two carboxylates share one hydrogen in active form. At high pH, as ionization proceeds, the hydrogen is gradually lost and activity drops. The other possibility, i.e., the dissociation of the MAV proteinase dimer to monomers at higher pH, can be eliminated since we have observed in centrifugation experiments approximately the same concentration of the dimer

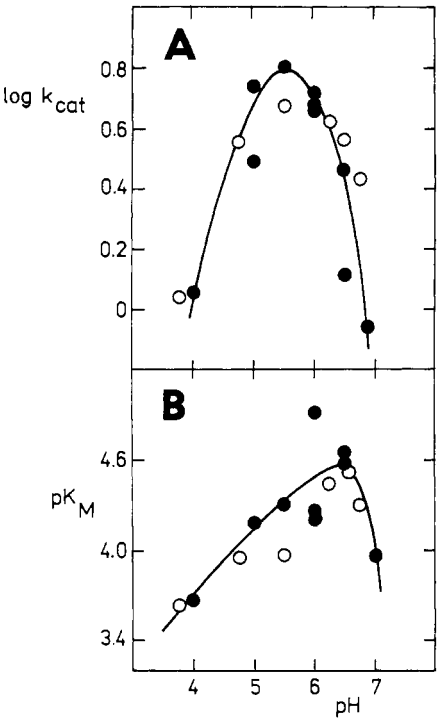

FIGURE 2: pH dependence of kinetic parameters (K_m, k_{cat}) of hydrolysis of two chromogenic peptides by MAV proteinase: peptide 36, AlaThrHisGlnValPhe*Phe(NO_2)ValArgLysAla (open circles), and peptide 53, AlaThrProGlnValTyr*Phe(NO_2)ValArgLysAla (full circles). Spectrophotometric assay conditions are given under Experimental Procedures.

at pH 5 and 7. There is an important difference in the values of k_{cat} for the HIV-1 and the MAV proteinases (Richards, 1989b, 1990). In the case of HIV-1 k_{cat} is constant over the range pH 4–9, while the MAV proteinase shows a distinct maximum at pH 5.5. Likewise, the K_m values for the HIV-1 proteinase increase steadily over the pH range 5.5–9 (Richards et al., 1990). When the structures of the HIV-1 proteinase and the MAV or RSV proteinase are compared (Miller et al., 1990), considering only the residues within a 4.1-Å sphere around the bound substrate analogue, two important differences only are found. In the flap region Gly 48 in HIV-1 is substituted for His (His 65 in MAV) and Pro 81 in the loop between the c′ and the d′ β chain is substituted for Arg (Arg 105 in MAV). The most probable candidates responsible for the differences in the pH profiles of k_{cat} between the HIV-1 proteinase and the MAV proteinase are His in the flap and Arg 105.

Conclusions

The p15^{gag} proteinase of MAV is a nonspecific and also inefficient enzyme. The apparent effectiveness and specificity of cleavage of viral polyproteins is most probably a result of the aggregation and high local concentration of the enzyme and protein substrate in the budding immature viral particles. The minimal substrate length required by the proteinase is 4 residues at the N-terminus and 3 residues at the C-terminus. The proteinase prefers hydrophobic residues with bulkier side chains (Tyr, Met) in P3. Arg and Gln are also accepted in P3. Small hydrophobic residues are required in P2 and P2′ (Val). P1 and P1′ prefer bulkier hydrophobic residues such as Tyr, Met, Phe, or 4-nitro-Phe, but a Pro residue can also fit into P1′ when a larger residue (Leu) occupies P2′. The difference in the specificity between the MAV and the HIV-1 proteinases is probably due to the participation of the flap region of the proteinase molecules in the interactions of the

S2′ pocket. The sequence of the flap region of the MAV proteinase is less homologous than the conserved parts of its structure with the corresponding regions of the HIV-1 proteinase; it is moreover extended by additional residues (Weber, 1989). Glu or Gln or Ile, accepted in P2′ of the substrates of the HIV-1 proteinase (Miller et al., 1989b), is too bulky for the MAV proteinase if a large hydrophobic group occupies P1′. Preference of the MAV proteinase for Pro in P4 or P5 in protein substrates might reflect the probability of location of the scissile bond near to surface turns or loops.

The activity of the MAV proteinase is strongly increased at high salt concentrations. This is reflected in both the k_{cat} and K_m values. The increase in k_{cat} is not caused by a higher concentration of the active dimer as determined by active-site titration. The proteinase shows a sharp pH optimum at pH 5.5–6.0 for both the synthetic peptides and for protein substrates. k_{cat} for peptide substrates reaches a sharp maximum at pH 5.5, while K_m is minimal at pH 6.5.

To understand better the specificity and function of the MAV proteinase a few lines of approach are being followed at present: investigation of new series of substrates with simultaneous (concerted) mutations in P1–P1′, P2–P1′, and P1–P2′, site-directed mutagenesis of MAV proteinase and specificity studies, comparison of the specificity with another retroviral proteinase of known structure (i.e., HIV-1), and finally, attempts to cocrystallize wild-type and mutant MAV proteinase with substrate analogues.

ACKNOWLEDGMENTS

We thank Dr. F. Kapralek and P. Jecmen for the cultivation of biomass, Dr. M. Pavlik for the determination of the N-terminal sequences of fragments of cleaved BSA and HSA, Mrs. Zdenka Senfelderova for careful technical assistance, and Mr. Jaroslav Zbrozek for amino acid analyses. We are thankful to Prof. J. Kay for critical reading through the manuscript and for most valuable comments.

REFERENCES

Alexander, F., Leis, J., Soltis, D. A., Crowl, R. M., Danho, W., Poonian, M. S., Pan, Y.-C., & Skalka, A. M. (1987) *J. Virol. 61*, 534–542.

Berger, A., & Schechter, I. (1970) *Philos. Trans. R. Soc. London, B 257*, 249–264.

Bhown, A. S., Bennet, J. C., & Hunter, E. (1980) *J. Biol. Chem. 255*, 6962–6965.

Billich, S., Knoop, M.-T., Hausen, J., Strop, P., Sedlacek, J., Merz, R., & Moelling, K. (1988) *J. Biol. Chem. 263*, 17905–17909.

Bradford, M. (1978) *Anal. Biochem. 86*, 142–146.

Brown, J. R. (1975) *Fed. Proc. 34*, 591–596.

Copeland, T. D., & Oroszlan, S. (1988) *Gene Anal. Tech. 5*, 109–115.

Copeland, T. D., Grandgenett, D. P., & Oroszlan, S. (1980) *J. Virol. 36*, 115–119.

Crawford, S., & Goff, S. P. (1985) *J. Virol. 28*, 106–118.

Darke, P. L., Nutt, R. F., Brady, S. F., Garsky, V. M., Ciccarone, T. M., Leu, T. M., Lumma, P., Freidinger, R. M., Veber, D. F., & Sigal, I. S. (1988) *Biochem. Biophys. Res. Commun. 156*, 290–296.

Darke, P. L., Leu, T. M., Lumma, P., Freidinger, R. M., Veber, D. F., & Sigal, I. S. (1989) *Biochem. Biophys. Res. Commun. 156*, 290–296.

Foundling, S. I., Salemme, R. R., Korant, B., Wendoloski, J. J., Weber, P. C., Treharne, A. C., Schadt, M. C., Jaskolski, M., Miller, M., Wlodawer, A., Strop, P., Kostka, V., & Ohlendorf, D. H. (1989b) in *Viral Proteinases as Targets for Chemotherapy* (Krausslich, H.-G., Oroszlan, S., & Wimmer, E., Eds.) pp 181–190, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Giam, C. Z., & Boros, I. (1988) *J. Biol. Chem. 263*, 14617–14620.

Grandgenett, Quinn, T., Hippenmeyer, P. J., & Oroszlan, S. (1985) *J. Biol. Chem. 260*, 8243–8249.

Graves, M. C., Lim, J. J., Heimer, E. P., & Kramer, R. A. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 2449–2453.

Hunter, E., Bennet, J. C., Bhown, A., Pepinsky, R. B., & Vogt, V. M. (1986) *Cell 45*, 375–385.

Kay, J., & Dunn, B. M. (1990) *Biochim. Biophys. Acta 1048*, 1–18.

Kotler, M., Katz, R. A., Danho, W., Leis, J., & Skalka, A. M. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 4185–4189.

Krausslich, H.-G., & von der Helm, K. (1988) *Virology 156*, 246–252.

Krausslich, H.-G., & Wimmer, E. (1988) *Annu. Rev. Biochem. 57*, 701–754.

Meloun, B., Moravek, L., & Kostka, V. (1975) *FEBS Lett. 58*, 134–137.

Miller, M., Jaskolski, M., Rao, J. K. M., Leis, J., & Wlodawer, A. (1989a) *Nature 337*, 576–579.

Miller, M., Schneider, J., Sathyanarayna, B. K., Toth, M. V., Marshall, G. R., Clawson, L., Selk, L., Kent, S. B. H., & Wlodawer, A. (1989b) *Science 246*, 1149–1152.

Misono, K. S., Sharieff, F. S., & Leis, J. P. (1980) *Fed. Proc. Exp. Biol. Abs. 60*, 1611–1616.

Moore, M. L., Bryan, W. M., Fakhoury, S. A., Magaard, V. W., Huffman, W. F., Dayton, B. D., Meek, T. D., Hyland, L., Dreyer, G. B., Metcalf, B. W., Strickler, J. E., Gorniak, J. G., & Debouck, C. (1989) *Biochem. Biophys. Res. Commun. 159*, 420–425.

Mouse, J., Heimer, E. P., & Le Grice, S. F. J. (1988) *J. Virol. 62*, 1433–1436.

Pepinsky, R. B., Mattaliano, R. J., & Vogt, V. M. (1986) *J. Virol. 58*, 50–58.

Richards, A. D., Roberts, R., Dunn, B. M., Graves, M. C., & Kay, J. (1989a) *FEBS Lett. 247*, 113–117.

Richards, A. D., Roberts, R., Dunn, B. M., Graves, M. C., & Kay, J. (1989b) *FEBS Lett. 253*, 214–216.

Richards, A. D., Phylip, L. H., Farmerie, W. G., Scarborough, P. E., Alvarez, A., Dunn, B. M., Hirel, Ph.-H., Konvalinka, J., Strop, P., Pavlickova, L., Kostka, V., & Kay, J. (1990) *J. Biol. Chem. 265*, 7733–7736.

Sauer, R. T., Allen, D. W., & Niall, H. D. (1981) *Biochemistry 20*, 3784–3791.

Schwartz, D. E., Tizard, R., & Gilbert, W. (1983) *Cell 32*, 853–869.

Sedlacek, J., Strop, P., Kapralek, F., Pecenka, V., Kostka, V., Travnicek, M., & Riman, J. (1988) *FEBS Lett. 237*, 187–190.

Sedlacek, J., Kapralek, F., Strop, P., Pichova, I., Kostka, V., & Travnicek, M. (1989) in *Proteases of Retroviruses* (Kostka, V., Ed.) pp 49–59, W. de Gruyter, Berlin, New York.

Seelmeier, S., Schmidt, H., Turk, V., & von der Helm, K. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6612–6616.

Strop, P., Pichova, I., Kostka, V., Kapralek, F., & Sedlacek, J. (1989a) in *Proteases of Retroviruses* (Kostka, V., Ed.) pp 62–72, W. de Gruyter, Berlin, New York.

Strop, P., Konvalinka, J., Pavlickova, L., Blaha, I., Soucek, M., Urban, J., Velek, J., Stys, D., Kostka, V., & Sedlacek, J. (1989b) in *Viral Proteinases as Targets for Chemo-*

therapy (Krausslich, H.-G., Oroszlan, S., & Wimmer, E., Eds.) pp 259-268, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Tam, J. P., Heath, W. F., & Merrifield, R. B. (1983) *J. Am. Chem. Soc. 105*, 6442-6455.
Vogt, V. M., Pepinsky, R. B., & Southard, L. E. (1985) *J. Virol. 56*, 31-39.
von der Helm, K., Seelmeier, S., & Schmidt, H. (1989) in *Proteases of Retroviruses* (Kostka, V., Ed.) pp 103-110, W. de Gruyter, Berlin, New York.
Weber, I. T., Miller, M., Jaskolski, M., Leis, J., Skalka, A. M., & Wlodawer, A. (1989) *Science 243*, 928-931.
Weiss, R. A., Teich, N. M., Varmus, H. E., & Coffin, J. M., Eds. (1982) *Molecular Virology of Tumor Viruses: RNA Tumor Viruses*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

β-NGF-endopeptidase: Structure and Activity of a Kallikrein Encoded by the Gene *mGK-22*†

Margaret Fahnestock,‡ James E. Woo,§ George A. Lopez,§ June Snow,∥ Daniel A. Walz,∥ Melih J. Arici,§ and William C. Mobley*,§

Departments of Neurology and Pediatrics and the Neuroscience Program, University of California, San Francisco, California 94143, Department of Physiology, Wayne State University Medical School, Detroit, Michigan 48201, and Department of Molecular Biology, SRI International, 333 Ravenswood Avenue, Menlo Park, California 94025

Received August 20, 1990; Revised Manuscript Received November 26, 1990

ABSTRACT: Mouse nerve growth factor (NGF) is cleaved at a histidine–methionine bond to release an NH_2-terminal octapeptide (NGF^{1-8}). The enzyme responsible, β-NGF-endopeptidase, is structurally and functionally similar to γ-NGF and epidermal growth factor-binding protein (EGF-BP) and cleaves mouse low molecular weight kininogen to produce bradykinin-like activity. These data have suggested that, like γ-NGF and EGF-BP, β-NGF-endopeptidase is a mouse glandular kallikrein. Evidence for a physiological role for NGF^{1-8} encouraged studies to further characterize the structure and function of this enzyme. Purified β-NGF-endopeptidase migrated as a single band on isoelectric focusing and reducing SDS–polyacrylamide gels. As was expected, it removed NGF^{1-8} from NGF. Interestingly, enzymatic activity on an artificial substrate, and on NGF, was inhibited by NGF^{1-8} and by bradykinin. These studies further supported the view that β-NGF-endopeptidase acts on both NGF and kininogen. The first 30 NH_2-terminal amino acids of β-NGF-endopeptidase were sequenced. This analysis demonstrated that the enzyme is encoded by the gene designated *mGK-22* (Evans et al., 1987). The sequence of this gene corresponds to that of EGF-BP type A (Anundi et al., 1982; Drinkwater et al., 1987), and so studies were performed to determine whether or not β-NGF-endopeptidase participates in EGF complex formation. Chromatographic and kinetic data gave no evidence that β-NGF-endopeptidase is an EGF-binding protein. Our studies suggest that contamination of high molecular weight (HMW) EGF preparations with β-NGF-endopeptidase erroneously led to earlier designation of the product of *mGK-22* as an EGF-BP. We conclude that *mGK-22* codes for β-NGF-endopeptidase and that this enzyme is a kallikrein active on both NGF and kininogen. These data suggest that the enzyme may have a biological role through production of NGF^{1-8} and kinins.

Mouse glandular kallikreins comprise a multigene family located on chromosome 7. The glandular kallikreins are serine proteases, structurally similar to trypsin; however, unlike trypsin, they have been reported to exhibit highly defined substrate specificities (Mason et al., 1983). Kallikreins may play an important role in enzymatically processing the precursors of polypeptide hormones. Twenty-four separate mouse kallikrein genes have been sequenced, and as many as 14 different active gene products may be produced (Evans et al., 1987). Members of the kallikrein family produced in the mouse submandibular gland include the α and γ subunits of 7S NGF[1] (Thomas et al., 1981; Isackson & Bradshaw, 1984), glandular kallikrein (Drinkwater & Richards, 1987), renal kallikrein (van Leeuwen et al., 1986), γ-renin (Poe et al., 1983; Drinkwater et al., 1988), and a number of other kallikreins whose functions have not been defined (Fahnestock et al., 1986; Drinkwater & Richards, 1988). In addition, three separate members of the kallikrein family have been identified as mouse EGF-binding proteins (EGF-BP). These proteins are EGF-BP type A, a single-chain protein (Anundi et al., 1982; Ronne et al., 1983), EGF-BP type B, a two-chain protein (Ronne et al., 1983), and EGF-BP type C, a three-chain protein and the only

†Supported by grants from NIH (NS24054), The March of Dimes, The Transamerica Foundation, and Cephalon Inc. to W.C.M., by NIH Grant NS23075 to M.F., and by U.S. Public Health Service Grant HL27073 to D.A.W.

*To whom correspondence should be addressed.
‡SRI International.
§University of California.
∥Wayne State University Medical School.

[1] Abbrevations: NGF, nerve growth factor; NGF^{1-8}, NH_2-terminal octapeptide of NGF; des^{1-8}NGF, NGF chain lacking the first eight NH_2-terminal amino acids; EGF, epidermal growth factor; HMW-EGF, high molecular weight complex containing EGF; EGF-BP, epidermal growth factor binding protein; SDS-PAGE, SDS–polyacrylamide gel electrophoresis; BAPNA, N^{α}-benzoyl-DL-arginine-*p*-nitroanilide; BK, bradykinin; Lys-BK, lysylbradykinin; CM-Cys, (carboxymethyl)cysteine. As used herein, NGF refers to NGF molecules irrespective of the protocol used for purification.

0006-2960/91/0430-3443$02.50/0