DRUG DEVELOPMENT RESEARCH 35:20–32 (1995)

Research Article

Comparison of the Proteolytic Susceptibilities of Homologous L-Amino Acid, D-Amino Acid, and N-Substituted Glycine Peptide and Peptoid Oligomers

Susan M. Miller, Reyna J. Simon, Simon Ng, Ronald N. Zuckermann, Janice M. Kerr, and Walter H. Moos

University of California, San Francisco, Department of Pharmaceutical Chemistry, San Francisco (S.M.M.) and Chiron Corporation, Emeryville (R.J.S., S.N., R.N.Z., J.M.K., W.H.M.), California

Strategy, Management and Health Policy				
Venture Capital Enabling Technology	Preclinical Research	Preclinical Development Toxicology, Formulation Drug Delivery, Pharmacokinetics	Clinical Development Phases I-III Regulatory, Quality, Manufacturing	Postmarketing Phase IV

ABSTRACT A series of homologous L-amino acid, D-amino acid, and both parallel and antiparallel (retro) sequence N-substituted glycine peptide and peptoid oligomers were prepared and incubated with a series of enzymes representative of the major classes of proteases. Each respective L-amino acid containing peptide sequence was readily cleaved by the appropriate enzyme, namely Ac-L-ala-L-leu-L-phe-L-ala-L-leu-L-arg-NH_2 by chymotrypsin, Ac-L-ala-L-ala-L-ala-L-leu-L-phe-L-arg-NH_2 by elastase, Ac-L-ala-L-phe-L-glu-L-leu-L-ala-L-ala-NH_2 by papain, Z-L-ala-L-his-L-phe-L-phe-L-arg-L-leu-NH_2 by pepsin, Ac-L-phe-L-ala-L-arg-L-ala-L-arg-L-asp-NH_2 by trypsin, and Ac-L-ala-L-tyr-Lala-L-phe-OH for carboxypeptidase A. In contrast, equivalent D-amino acid containing and N-substituted glycine containing oligomers were cleaved minimally or not at all by the respective enzymes. The N-substituted glycine peptoids represent a new class of combinatorial diversity for lead discovery with improved pharmaceutical characteristics relative to L-amino acid containing peptides.

Key Words: combinatorial chemistry, molecular diversity, peptide, peptoid, protease

INTRODUCTION

The use of molecular diversity approaches to drug discovery has mushroomed in the last several years [Maeji et al., 1991; Houghten et al., 1991; Lam et al., 1991; Furka et al., 1991; Gallop et al., 1994; Gordon et al., 1994; Desai et al., 1994]. With automated techniques and the combinatorial use of various types of building blocks, e.g., both genetically encoded standard L-amino acids, as well as non-natural building blocks like N-substituted glycines (NSGs), libraries of unprecedented numbers of compounds can be generated and screened over very short time periods. However, the discovery aspects of molecular diversity represent only one segment of pharmaceutical research and development. Many leads ultimately fail because of deficiencies in one or more important pharmaceutical property, such as absorption, distribution, metabolism, and excretion (ADME), not to mention safety and efficacy. While small (i.e., non-protein) peptide and modified peptide therapeutics are well known (e.g., calcitonin, captopril, cyclosporin, oxytocin), peptides often exhibit several of these deficiencies. Consequently, peptide libraries per se may hold less immediate promise than non-peptide or peptoid li-

Received December 16, 1994; final version accepted January 26, 1995.

Address reprint requests to Susan M. Miller, Department of Pharmaceutical Chemistry, University of California, 513 Parnassus Ave, Box 0446, San Francisco, CA 94143-0446.

TABLE 1. Properties of Enzymes Used in This Study

Enzyme	Specificity[a]	Class	Substrate example
Carboxypeptidase A 3.4.17.1	P_1, aromatic free carboxyl preferred, extended site to P_3 or P_4	Metallo	Z-gly-gly-L-phe-OH, k_{cat}/K_m5.3 × 10^4 M^{-1} s^{-1}, pH 7.6, 27°C[b]
Papain 3.4.22.2	P_2 aromatic or hydrophobic, P_1, large hydrophobic, extended site P_4–P_3,	Cysteine	Z-gly-L-val-L-glu-L-leu-gly-OH, k_{cat}/K_m 3.2 × 10^3 M^{-1} s^{-1}, pH 6.5, 40°C[c]
Pepsin 3.4.23.1	P_1 and P_1, aromatic, extended site P_4–P_3,	Aspartyl	L-phe-gly-L-his-L-(4NO_2)phe-L-phe-L-val-L-leu-OMe, k_{cat}/K_m 1.5 × 10^6 M^{-1} s^{-1}, pH 3.5–4.0, 37°C[d]
Trypsin 3.4.21.4	P_1 L-lys or L-arg, some extension to P_3	Serine	Bz-L-phe-L-val-L-arg-pNA, k_{cat}/K_m 1.1 × 10^6 M^{-1} s^{-1}, pH 8.1, 37°C[e]
Elastase 3.4.21.36	P_1 hydrophobic preferred, but extended site P_4–P_2,	Serine	Boc-L-ala-L-ala-L-leu-SBzl, k_{cat}/K_m 4.3 × 10^6 M^{-1} s^{-1}, pH 7.5, 25°C[f]
Chymotrypsin 3.4.21.1	P_1 aromatic, extended site P_3–P_3,	Serine	Glt-L-ala-L-ala-L-leu-L-phe-pNA, k_{cat}/K_m 1.4 × 10^6 M^{-1} s^{-1}, pH 7.8, 25°C[g]

[a]Polgár [1989].
[b]Z: carbobenzyloxy. Abramowitz et al. [1967].
[c]Lowbridge and Fruton [1974].
[d]Fruton [1976].
[e]Bz: benzoyl; pNA: p-nitroanilide. Pozsgay et al. [1981].
[f]SBzl: $SCH_2C_6H_5$; Boc: t-butyloxycarbonyl. Harper et al. [1984].
[g]Glt: glutaric acid. Fischer et al. [1984].

braries, especially for orally available therapeutics (at least until drug delivery methods achieve greater success).

A significant effort in our molecular diversity research focuses on the seemingly peptide-like oligo (NSGs). We have developed a "sub-monomer" process for synthesizing NSG peptoids that allows the use of simple acetates and amines to prepare highly diverse libraries quickly and inexpensively [Zuckermann et al., 1992a]. The NSG peptoids, which can mimic both peptides and non-peptides, demonstrate potent and specific biological activity [Simon et al., 1992; Zuckermann et al., 1994]. From an ADME perspective, work by Conradi et al. [1992] would predict improved absorption characteristics for NSG peptoids because of N-substitution. But the gut and the bloodstream, to mention only two biological milieus, provide unfavorable environments for peptides owing to the various classes of proteases present in these physiological compartments. As a further step in characterizing the pharmaceutical potential of new diversities, we report herein a qualitative assessment of the stability of a set of NSG peptoids against a series of common proteases.

For this study, we selected enzymes representative of the four known classes of relevant proteases (Table 1): carboxypeptidase A, papain, pepsin, elastase, trypsin, and chymotrypsin. Based on known sequence specificities, we have designed L-amino acid peptide (all-L) substrates to directly compare with homologous D-amino acid containing (all-D) and N-substituted glycine containing oligomers. For most of the enzymes, both parallel (all-N) and anti-parallel (retro all-N) NSGs were prepared to cover both possible reading frames (Fig. 1). A preliminary report of these data has been presented elsewhere [Miller et al., 1994].

MATERIALS AND METHODS

Enzymes and Reagents

Bovine pancreatic chymotrypsin A_4, trypsin, and carboxypeptidase A, porcine pancreatic elastase, and papain were from Boehringer Mannheim (Indianapolis, IN). Porcine stomach mucosa pepsin was from Calbiochem (La Jolla, CA). 4-Fluoro-7-nitro-benzofurazan was from Aldrich (Milwaukee, WI). Z-L-glu-L-phe-OH, Suc-L-phe-pNA, Boc-L-ala-ONp, Bz-D,L-arg-pNA·HCl, and Z-L-his-L-4-nitro-phe-L-phe-OMe were from Bachem Bioscience (Philadelphia, PA). The carboxypeptidase A all-L and all-D peptides were obtained from Chiron Mimotopes (Clayton, Victoria, Australia). All reagents were of analytical grade.

Peptides and NSG Peptoids Design

Figure 2 summarizes the peptide and NSG peptoid sequences designed and synthesized for these proteolysis studies. Hexapeptides were chosen for all of the enzymes except carboxypeptidase A, since previous studies indicate extended recognition sites for most proteases of at least 6 or 7 residues [Polgár, 1989]. In each case, the sequences are based on literature studies defining good peptide substrates and/or inhibitors. For carboxypeptidase A, a tetramer was designed based on the structure of the potato peptide inhibitor bound to the enzyme, which shows interactions of only 4 amino acids of the inhibitor with the

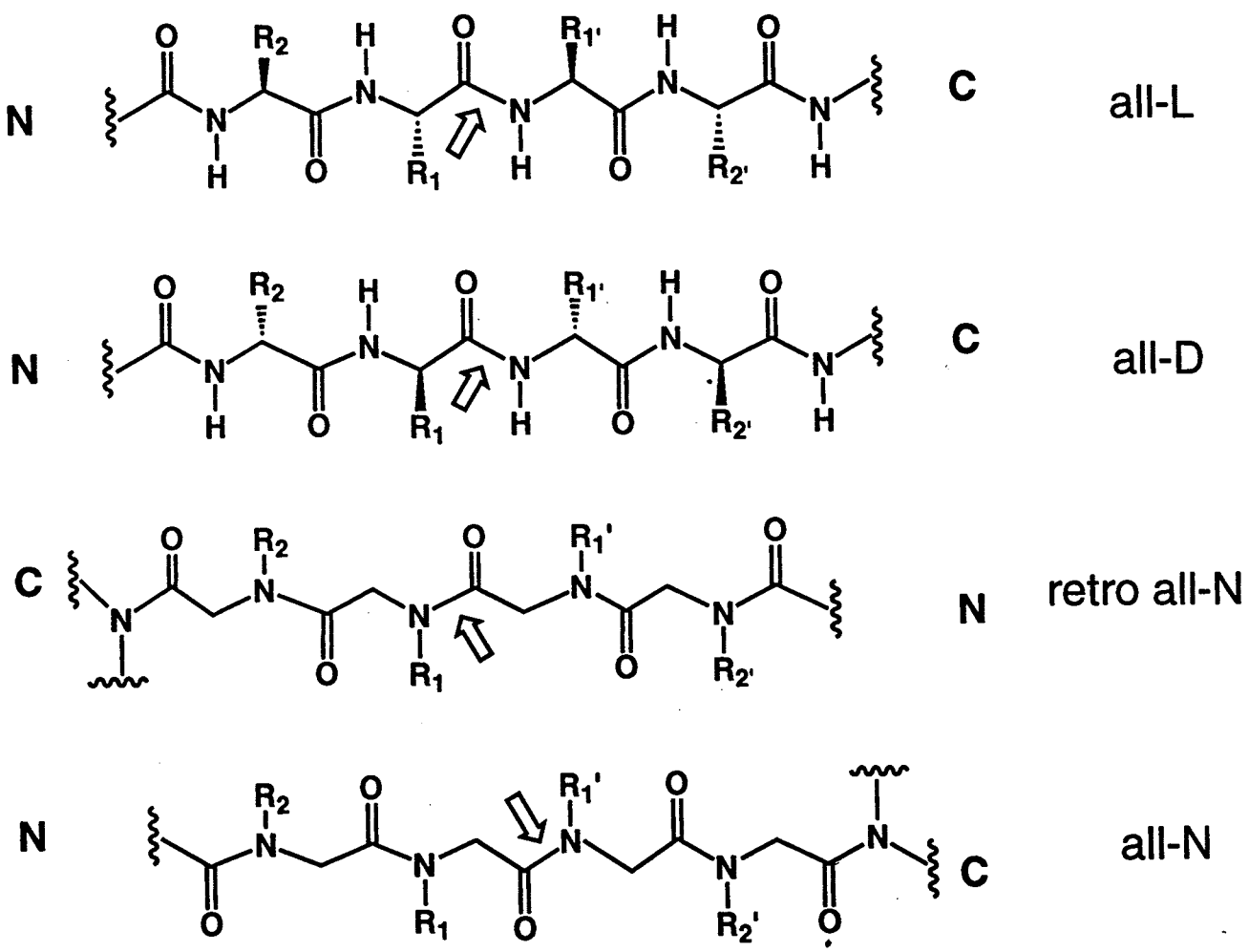


Figure 1. Comparison of general structures of all-X peptides and peptoids. The side chain labels illustrate the standard nomenclature for all-L peptide cleavage by proteases (arrows), which occurs at the carboxyl end of the P_1 residue and the amino end of the $P_{1'}$ residue [Schechter and Berger, 1967].

active site [Rees and Lipscomb, 1982]. Because of the requirement for a free carboxylate group in the carboxypeptidase A substrate, a fully homologous retro NSG peptoid could not be synthesized. For papain, Schechter and Berger [1967] demonstrated that extended binding from P_4–$P_{3'}$ improves activity, and later demonstrated [1968] that the primary specificity is for aromatic residues at P_2. Inclusion of L-glu-L-leu at the P_1–$P_{1'}$ sites is based on the studies of Lowbridge and Fruton [1974], which demonstrate the further specificity of the $P_{1'}$ site of papain for large hydrophobic residues in substrates. For pepsin, the P_3–$P_{1'}$ sequence and the inclusion of a carbobenzyloxy capping group is based on the extensive studies from the lab of Fruton [see Fruton, 1976], which explore site specificities and demonstrate an extended substrate binding site from P_4–$P_{3'}$. The L-arg-L-leu sequence in the $P_{2'}$– $P_{3'}$ position is based on an analysis by Powers et al. [1977] of cleavage sites reported for a wide variety of protein substrates for pepsin. The retro NSG peptoid sequence as a benzamide derivative was designed as a homolog for the pepsin peptide, but was not synthesized for this study. For trypsin, the P_3–P_1 sequence is based on substrate studies by Pozsgay et al. [1981], while the $P_{1'}$–$P_{3'}$ fragment is based on the sequence of the pancreatic trypsin inhibitor complexed with trypsin [Huber and Bode, 1978]. For elastase, extension of the peptide out to P_4 or P_5 appears to be more important for activity than extension on the P′ side [Thompson and Blout, 1973]. Thus, the sequence of P_4–P_1 is based on the studies by Thompson and Blout [1973], as well as those by Harper et al. [1984], which show that substitution of L-leu for L-ala at P_1 gives a better k_{cat}/K_m value. L-phe at the $P_{1'}$ site of the

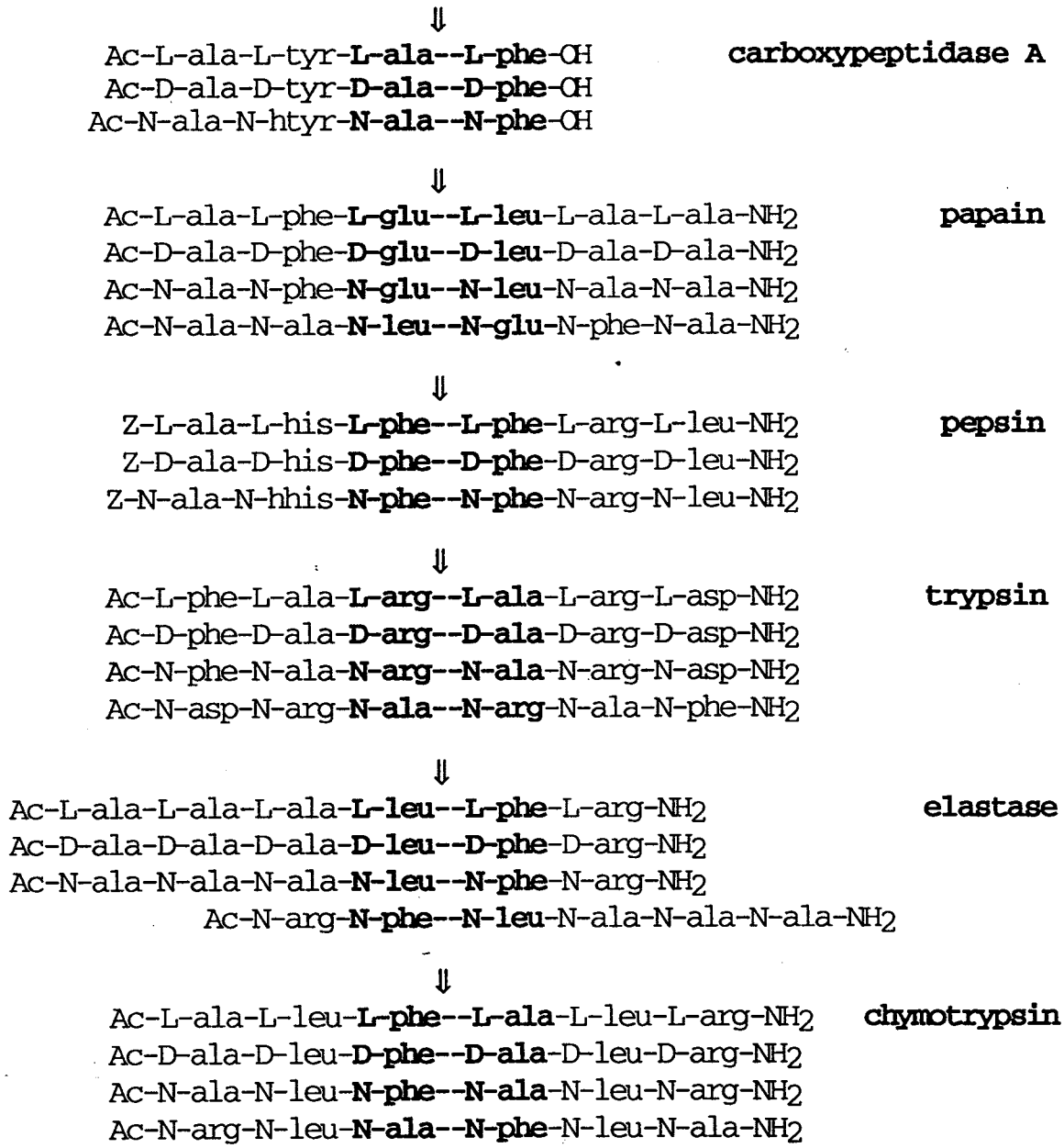


Figure 2. Sequences of peptides and NSG peptoids designed for the indicated proteases. Expected cleavage sites are indicated by boldface residues and arrows. In the NSG oligomers, the R-groups attached to nitrogen are identical to the R-groups of the L- and D-amino acids except for the N-hhis and N-htyr residues, which are extended by an additional methylene group.

elastase peptide is based on studies by Bauer et al. [1976]; and L-arg at $P_{2'}$ was included to improve solubility of the peptide. For chymotrypsin, P_3–P_1 residues are based on the analysis by Fischer et al. [1984] of extended peptides as substrates, while $P_{1'}$–$P_{3'}$ residues are based on tight binding inhibitors described by Imperiali and Abeles [1987].

Synthesis

L- and D-amino acid containing peptides were synthesized by solid-phase 9-fluorenylmethoxycarbonyl (Fmoc) chemistry on an automated synthesizer as described [Zuckermann et al., 1992b]. NSG peptoids, except for the pepsin and carboxypeptidase

all-N oligomers, were prepared using the method of Zuckermann et al. [1992a]. The pepsin all-N hexamer and the carboxypeptidase A all-N tetramer were synthesized from the Fmoc-protected N-substituted glycine monomers [Simon et al., 1992] using solid-phase chemistry on an ABI synthesizer (Foster City, CA).

Purification

All peptides and peptoids were purified by gradient elution from Vydac (Hesperia, CA) C_{18} semiprep or prep HPLC columns using Rainin HPLC chromatographs (Woburn, MA) controlled by software operating on Apple Macintosh (Cupertino, CA) computers. Samples were loaded with the column equilibrated at 0 or 5% acetonitrile in water, with 0.1% trifluoroacetic acid in both phases, and were eluted with a gradient up to 65% acetonitrile, increased at 2%/min. Purified sam ples were lyophilized and stored at 4°C. L- and D-amino acid containing peptides were characterized by amino acid analysis by the ninhydrin method using a Beckman amino acid analyzer (Fullerton, CA). Samples of the purified NSG-peptoids were characterized by mass spectral analysis performed at Mass Search, Inc., 1560 Cummins Dr., Suite C, Modesto, CA 95351.

Proteolysis Controls, Conditions, and Analysis

Solvent controls

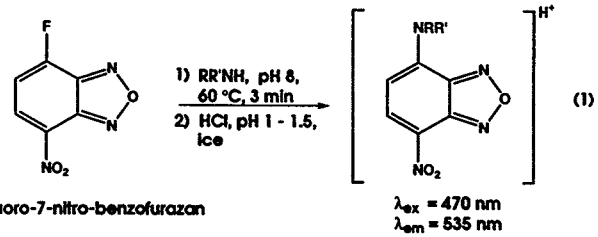
Dissolution of the peptides and peptoids required 10–20% organic solvent, thus the effect of potential solvents (Tween-80 or DMSO) on the activity of the various enzymes was evaluated using colorimetric substrates for each enzyme. Using Suc-L-phe-pNA for chymotrypsin, Bz-D,L-arg-pNA·HCl for trypsin and papain, and Boc-L-ala-ONp for elastase, activity was measured by the increase in absorbance at 405 nm for release of p-nitroaniline or p-nitrophenol. Using Z-L-glu-L-phe-OH for carboxypeptidase A and Z-L-his-L-(4NO_2)phe-L-phe-OMe for pepsin, activity was measured by the decrease in the peptide absorbance band between 220 and 230 nm. Concentrations of 0, 1, 2.5, 5, and 10% of either Tween-80 or DMSO in the buffers indicated below were evaluated for inhibition of activity. Absorbance measurements were made using a Shimadzu UV160U spectrophotometer (Columbia, MD).

Proteolysis conditions

Reaction conditions for each enzyme were set to approximate physiological conditions as shown in Table 2. Preliminary experiments with the all-L peptides were conducted to establish useful concentrations and times of reaction. Stock solutions of enzymes (100 μM) in their respective buffers were made fresh the day of each experiment. Chymotrypsin and trypsin were made as 1.0 mM stocks in 1.0 mM HCl and then diluted to 100 μM final concentration with bicarbonate buffer just before use. Papain was activated by incubation of the 100 μM stock with 2 mM dithiothreitol at 37°C for 5 min before dilution into the reaction mixtures. Concentrations of the peptides and peptoids in the stock solutions were determined by amino acid analysis as described above. For the analysis of the NSG peptoids, the N-phe derivative was found to elute with the same retention time as the L-phe derivative but gives a lower response. By comparing the analysis of several hexapeptides containing five L-amino acids and a single N-phe, the response factor for the N-phe derivative (area under HPLC peak/pmol) was found to be 62.5 ± 1.9% of the L-phe response. Using this percentage and an L-phe standard, the concentration of N-phe, and of each NSG peptoid was quantitated, since each of the compounds of interest contained at least one N-phe residue.

Quantitation of Cleavage

For each reaction, separate enzyme and peptide/peptoid controls at the same dilutions as the complete reaction mixture were incubated simultaneously at 37°C. Reaction mixtures were preequilibrated at 37°C for approximately 3 min before initiation of reaction by addition of enzyme. The extent of peptide cleavage was quantitated in a discontinuous assay by the fluorescence developed upon reaction of the 1° or 2° amine product with 4-fluoro-7-nitro-benzofurazan (NBD-F; see structure below) [Imai and Watanabe, 1981; Miyano et al. 1985].



At appropriate time points, 40 μL aliquots of reaction or control were quenched by dilution into a mixture of 0.2 mL acetonitrile plus 0.2 mL × 0.1 M sodium borate buffer, pH 8.05, in 10 × 75 mm test tubes on ice. NBD-F (2.0 μL × 50 mM in acetonitrile) was added and the samples were immediately incubated for 3.25 min at 60°C in a heating block, after which they were placed on ice, and 20 μL × 3

TABLE 2. Proteolysis Conditions*

Enzyme (MW)[a]	Concentration in assay (μM)	Reaction buffer	Substrate concentration in assay (mM)[b]
Carboxypeptidase A (34,500)	1.0	25 mM sodium bicarbonate, pH 7.8	L: 1.79 D: 1.73 N: 1.03
Papain (23,400)	2.0	50 mM potassium phosphate, pH 6.95	L: 0.92 D: 1.06 N: 2.63 RN: 3.73
Pepsin (34,000)	2.0	8.4 mM HCl/34 mM NaCl, pH 2.05	L: 1.02 D: 1.73 N: 1.62
Trypsin (24,000)	2.0	25 mM sodium bicarbonate, pH 7.8	L: 0.78 D: 0.72 N: 0.83 RN: 0.85
Elastase (22,000)	1.0	25 mM sodium bicarbonate, pH 7.8	L: 0.38 D: 0.48 N: 2.22 RN: 0.73
Chymotrypsin (22,500)	2.0	25 mM sodium bicarbonate, pH 7.8	L: 0.88 D: 0.59 N: 0.85 RN: 1.36

*All reactions were run at 37°C.

[a]Schomburg and Salzmann [1990].

[b]All substrates were dissolved in 20% DMSO/80% respective reaction buffer and were diluted 10-fold into the reactions, except the chymotrypsin all-L peptide, which was dissolved in 13% DMSO/87% buffer and was diluted 5-fold into the assay. Concentrations of the stock solutions were determined by amino acid analysis as described in the text and are accurate to ±4%. Separate reactions were run simultaneously for each compound, where L is the all-L peptide; D is the all-D peptide; N is the all-N peptoid; RN is the retro all-N peptoid.

N HCl was added to each to stabilize the 4-amino-7-nitro-benzofurazan product. The fluorescence was read using a Hitachi fluorimeter (Hitachi Instruments, Danbury, CT) with λ_{ex} = 470 nm and λ_{em} = 535 nm. As standards for approximate quantitation of the peptide/peptoid products, duplicate or triplicate fluorescence curves were generated for 0 to 110 μM concentrations of L-phe, N-benzyl-gly-OEt (N-phe), L-leu, N-leu-N-htyr-NH_2, L-ala, and sar-OEt (N-ala) using the same protocol and measuring a 40 μL aliquot of water for the baseline hydrolysis background. (Because of only limited availability of deprotected N-glu and N-arg monomers and their minimal occurrence as expected amino products, these two standards were not evaluated.) The data were fit to a second-order polynomial equation using the curvefit capability in the graphing program KaleidaGraph (version 3.0.1) (Synergy Software, Reading, PA) for the Macintosh.

RESULTS

Synthesis and Purification

After synthesis, the purity of the peptides and NSG peptoids was evaluated by analytical reverse phase HPLC as described in Materials and Methods. The all-L and all-D peptides showed only very minor impurities by HPLC; however, when tested for background fluorescence, they all showed sufficient signal to warrant purification. Thus, all were purified by preparative HPLC. HPLC analysis of the NSG peptoids showed more than one product for all of the sequences. However, for each forward (all-N) and related retro all-N sequence, a major peak was identified with nearly the same retention time as the corresponding all-L and all-D peptides (which are enantiomeric and have identical retention times). Figure 3 shows a comparison of the all-L, all-N, and retro all-N papain peptides/peptoids before purification. Upon purification and mass spectral analysis, those peaks with the similar retention times were found to have the predicted masses for the desired sequences in all cases.

Solvent Controls

Except for the trypsin and carboxypeptidase A peptides, the designed substrates are fairly hydrophobic and require some organic solvent/buffer mix for dissolution. In an effort to maximize the concentration of peptides in the assays and minimize solvent inhibition of the enzymes, we evaluated the effect of both Tween-80, which is sometimes employed in formulating drugs for animal experiments, and DMSO, which is often a good solvent for peptides but not typically desirable in pharmacological assays. Tween-80, even at only 1%, severely inhibited chymotrypsin, elastase, papain, and trypsin. (Pepsin and carboxypeptidase A were not tested because Tween-80 absorbs in the region used to monitor their activity.) By contrast, up to 2.5% DMSO had very little effect on the activity of any of the enzymes (data not shown). Thus, peptide and NSG peptoid stock solutions were made using 20% DMSO, which upon 10-fold dilution into the assays gave an essentially noninhibitory concentration of 2% DMSO.

Fluorescence Conditions

The conditions for the fluorescence assays are based on the extensive characterization reported by Imai and Watanabe [1981] and Miyano et al. [1985]. From those studies it is clear that: (1) borate buffer at pH 8.0 gives maximal development and stability of fluorescence; (2) the presence of 50% organic solvent increases the yield of amine product and decreases the yield of hydrolysis product; (3) quenching the reaction by addition of hydrochloric acid to give a final

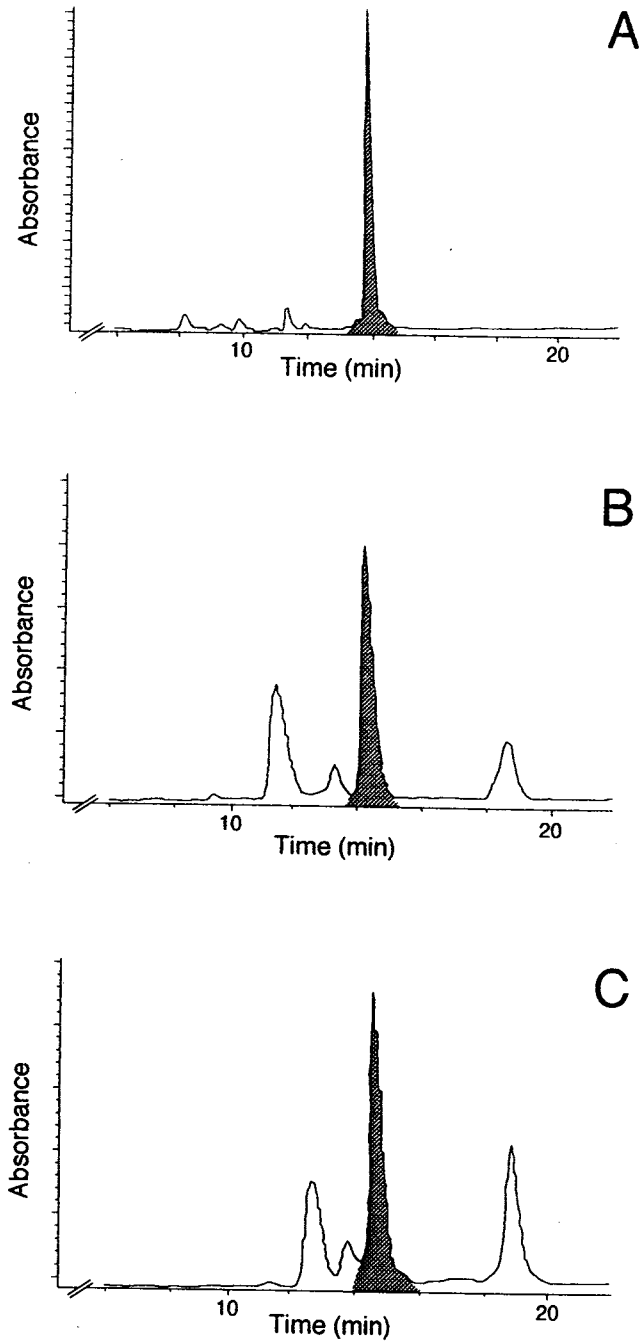


Figure 3. Comparison of analytical reverse phase HPLC elution pattern for papain peptides/peptoids before purification. The peptide gradient was 0 to 60% acetonitrile, 0 to 30 min, while the peptoid gradient was 1 to 61% acetonitrile, 0 to 30 min. Desired compounds are marked. A: All-L peptide elutes at 27.5% acetonitrile. B: All-N peptoid and C: retro all-N peptoid elute at 29.8% acetonitrile.

pH between 1 and 3 further increased and stabilized the net fluorescence of the product. To further optimize the assay for our compounds, we evaluated two

TABLE 3. Optimization of Fluorescence Assay Variables*

	Signal/noise		Net relative fluorescence	
Variable	L-phe	N-phe	L-phe	N-phe
Solvent[a]				
EtOH	8	9	33	31
CH_3CN	19	13	124	78
Temperature and time (min)[b]				
60°C				
1.5	nd	28	nd	57
3.0	nd	20	nd	90
5.0	nd	15	nd	87
10.0	nd	11	nd	83
70°C				
1.0	nd	25	nd	40
2.0	nd	19	nd	69
5.0	nd	10	nd	70
10.0	nd	9	nd	72
Concentration NBD-F (μM)[c]				
215	23	18	99	73
431	19	10	151	80
855	4	2	147	38

*nd = not determined.
[a]Conditions: 50% solvent, 50% 0.1 M sodium borate buffer, pH 8.0, 246 μM 4-fluoro-7-nitro-benzofurazan (NBD-F), 9.9 μM amino acid, heated at 70°C for 4 min and quenched on ice with addition of 20 μL × 3 N HCl.
[b]Conditions as in Solvent with solvent = acetonitrile, but with temperature and time varied as indicated.
[c]Conditions as in Temperature and time with temperature = 60°C and time = 3.0 min, but with varied NBD-F as indicated.

organic solvents, ethanol and acetonitrile, the time and temperature of heating, and the concentration of NBD-F in the assay, focusing primarily on the reaction of N-phe, since the more sterically hindered secondary amines are expected to react more slowly with the reagent and therefore yield lower overall fluorescence. As summarized in Table 3, maximal net fluorescence coupled with nearly optimal signal to noise was obtained for both L-phe and N-phe using acetonitrile as solvent with incubation at 60°C for approximately 3 min in the presence of 215 μM NBD-F. Under these conditions, as little as 0.5 μM L-phe or N-phe could be detected. Higher temperatures, longer incubation times, and higher concentrations of NBD-F led to increased hydrolysis of the reagent and a significant reduction in signal to noise. The final conditions adopted for all of the assays included acetonitrile as solvent, 215 μM NBD-F, and incubation at 60°C for 3.25 min.

Fluorescence Standards

Figure 4 shows average fluorescence standard curves measured for the L- and N-amino acids expected to occur at the amino termini of the peptide or peptoid cleavage products in this study. Standards

were run in each case up to the maximal concentration expected for complete cleavage of the all-L peptides at the expected primary site of cleavage. All of the amino acids evaluated show decreasing fluorescence at high concentrations indicative of self-quenching. As indicated by the dashed lines in Figure 4, the observed fluorescence can be described by the following parabolic dependences on concentration, which result from the inner-filter effect at high concentrations of fluorophore [Schenk, 1973]. Thus,

L-ala-OH:	Fluorescence = $1.63 + 4.49x - 0.014x^2$	$R^2 = 0.981$ (2)
N-ala-OEt:	Fluorescence = $24.38 + 13.53x - 0.116x^2$	$R^2 = 0.927$ (3)
L-leu-OH:	Fluorescence = $5.79 + 9.25x - 0.048x^2$	$R^2 = 0.975$ (4)
N-leu-N-htyr-NH_2:	Fluorescence = $0.26 + 5.49x - 0.031x^2$	$R^2 = 0.999$ (5)
L-phe-OH:	Fluorescence = $2.94 + 11.20x - 0.062x^2$	$R^2 = 0.979$ (6)
N-phe-OEt:	Fluorescence = $-1.28 + 7.17x - 0.033x^2$	$R^2 = 0.999$ (7)

where x equals the concentration of the standard. Concentrations of unknown amine solutions are then quantitated from fluorescence values by solving for the negative root of the appropriately rearranged equation. Standards for N-glu and N-arg, which would only be expected for cleavage of the reverse NSG peptoids for papain and trypsin, respectively, were unavailable for comparison in these studies and were assumed to give at least the minimal fluorescence observed with the N-leu standard in the calculation of fractional conversions below. Since no fluorescence above the background water samples was observed for any time point in the assays with either the trypsin or papain retro all-N compounds, this analysis seems reasonable.

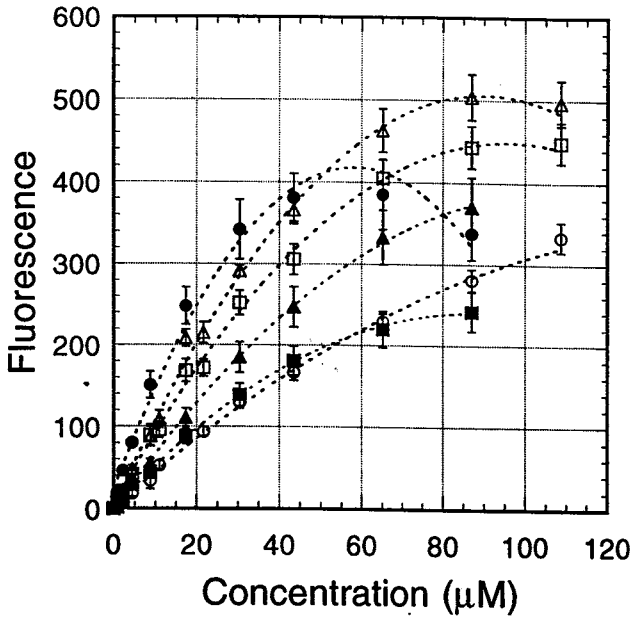


Figure 4. Fluorescence standard curves. Conditions are as indicated in Materials and Methods. Data are for (○) L-ala-OH, (●) N-ala-OEt, (□) L-leu-OH, (■) N-leu-N-htyr-NH2, (△) L-phe-OH, (▲) N-phe-OEt. Points are presented as the mean ± s.e.m. of duplicate (N-substituted) or quadruplicate (L) measurements run on separate days. Dashed lines are parabolic fits to all of the data points for a given compound as described in the text.

Proteolysis

Data for the proteolysis of the respective peptides and peptoids are displayed as fractional conversions, by enzyme, in Figure 5. As indicated by the open circles for each plot, the all-L peptides are indeed substrates for their respective enzymes. The peptide designed for elastase (Fig. 5E) clearly undergoes cleavage at more than one site over the course of the experiment. The concentration of amine product for elastase cleavage was calculated using the equation for fluorescence of an L-phe amino derivative, which is the expected initial cleavage site, as indicated in Figure 2. Further cleavage would likely occur after any of the alanine residues, yielding L-leu and/or L-ala amino termini. Since the fluorescence standard curves for both of these are shallower than for L-phe, the fractional conversions indicated for the elastase peptide represent minimal values.

With the other enzymes, a fairly rapid but variable extent of cleavage of the all-L peptides occurs in the first 6 to 12 min of the assay followed by a much slower, further cleavage (except for pepsin, where cleavage does not appear to proceed any further after 6 min). This behavior could result from instability of enzyme activity upon extended incubation (perhaps due to self-proteolysis) or may be due to inhibition of the enzymes by the cleavage products from the reaction, since several of the peptides were designed for good binding based on sequences and structural information for naturally occurring protein protease inhibitors. Another alternative explanation is that the initial cleavage is nearly complete in the burst, but inaccurately measured by the standards, and the slow phase is due to further cleavage of the primary fragments. Since the issue is not critical to the conclusions of the study, we have not pursued this point further at this time.

In sharp contrast to the significant cleavage seen for the all-L peptides with each enzyme, there is only minimal cleavage of three all-D peptides, and no convincing cleavage of any of the all-N or retro all-N peptoids. With chymotrypsin, there may be 1–2% cleavage of the all-D peptide, but the points are scattered over the course of the incubation. With trypsin and carboxypeptidase A, the all-D peptides show increasing trends in percent cleavage, with approxi-

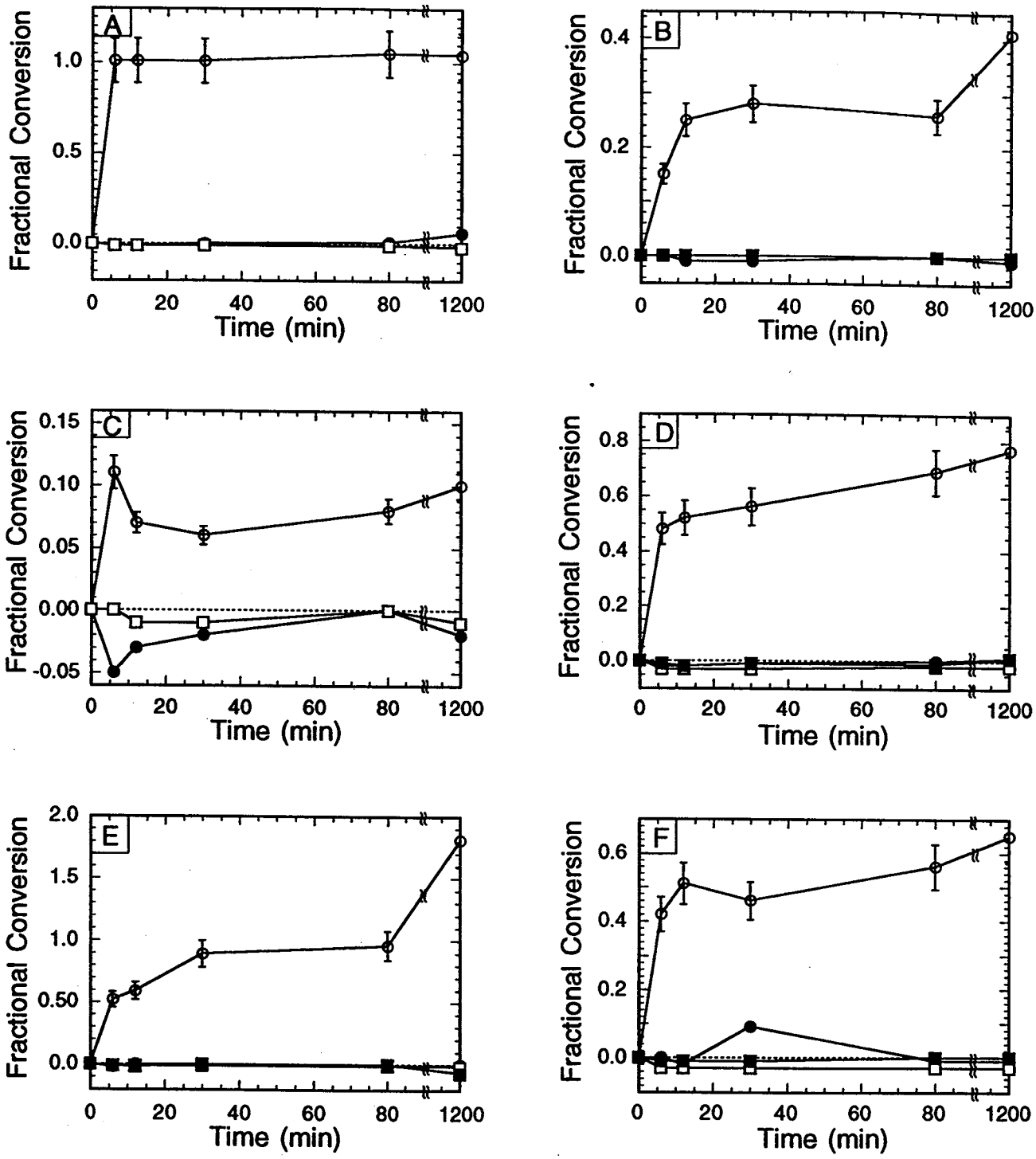


Figure 5. Fractional conversions of peptides and peptoids in simultaneous reactions. Reaction conditions are described in Table 2. Fluorescence values for each time point were corrected for respective enzyme and substrate background fluorescence. Concentrations of amino product were calculated from the fluorescence using the appropriate standard for the resultant amino terminal residue at the expected cleavage site, as indicated in Figure 2, and are divided by the initial substrate concentration to obtain fractional conversion. Error bars indicate the uncertainty of the values propagated from the uncertainties in the fluorescence values and in the peptide and peptoid concentrations. Reactions with all-L substrates run on different days gave fractional conversions within the errors indicated. (No error bars are given for the D or N compounds since the fluorescence values were within experimental error of 0 in all cases.) Concentrations of amino products from the reverse peptoids of trypsin and papain, where the expected amino products are Narg and Nglu, respectively, were estimated using the standard curve for Nleu-Nhtyr-NH_2, the least fluorescent peptoid standard evaluated. For each panel (○) all-L peptide, (●) all-D peptide, (□) all-N peptoid, (■) retro all-N peptoid. Enzymes are (A) carboxypeptidase A, (B) papain, (C) pepsin, (D) trypsin, (E) elastase, (F) chymotrypsin.

mately 1% cleavage for the trypsin all-D and approximately 6% cleavage for the carboxypeptidase A all-D peptides after 20 h. The only NSG peptoid giving any hint of cleavage is the all-N peptoid for papain, which may be cleaved to 0.2%; however, again, the data are scattered over the course of the incubation. No production of detectable amine products was found for any of the other NSG peptoids.

Inhibitory Effects of NSG Peptoids on All-L Cleavage

As described further in the discussion, the absence of cleavage of any of the NSG peptoids is likely to result from a misalignment of the bound peptoids with the catalytic residues at the active sites of the enzymes. However, weak or tight binding properties of the compounds could also contribute to the result. On the one hand, peptoid cleavage may occur at a low rate, but if the K_m values for the peptoids are substantially higher than those for the all-L peptides, reflecting weak binding, the concentrations examined may be too low to observe significant cleavage over the time of the reaction. By contrast, if cleavage occurs, but one or both product fragments bind very tightly with a slow rate of dissociation, the peptoid could act as an inhibitor. As a qualitative test to distinguish between these two possibilities, we evaluated the extent of cleavage of the all-L peptides in the absence and presence of their respective NSG peptoids for four of the enzymes. As shown in Figure 6, the NSG peptoids designed for this study are not potent inhibitors against cleavage of the homologous all-L peptides by these proteases, eliminating the hypothesis that tight binding prohibits extensive cleavage of the NSG peptoids by the proteases. At the single concentrations evaluated, three of the NSG peptoids showed inhibition of cleavage of their homologous all-L peptides: the papain all-N peptoid, present at 2.8-fold greater concentration than the all-L peptide; the elastase all-N peptoid, at 5.8-fold greater concentration than the all-L peptide; and the elastase retro all-N peptoid, at 2-fold greater concentration than the all-L peptide. The inhibition is weak, even with the elevated concentrations, indicating that the NSG peptoids bind rather weakly to the proteases. The papain retro all-N peptoid showed no inhibition at a 4-fold greater concentration over the all-L substrate, suggesting that it binds only very weakly, if at all. The other three NSG peptoids that showed no inhibition were present at only equimolar or lower concentrations than their respective all-L peptides (carboxypeptidase A, N/L = 0.6; trypsin, N/L and RN/L approximately 1). Clearly these three compounds are not tight binders, but weak binding may simply not be detectable at the concentrations evaluated.

DISCUSSION

While molecular diversity approaches to lead discovery and lead optimization are receiving increasing attention, few groups have addressed potential development hurdles, such as the ADME characteristics of the novel classes of oligomeric compounds being generated. Our interest in NSG peptoids both as homologs of peptides [Simon et al., 1992] and as de novo leads [Zuckermann et al., 1994], has led us to begin investigation of the ADME characteristics of these compounds. Notable among the many obstacles encountered in developing orally available peptides and peptidomimetics are the proteases found in the gut, the bloodstream, and other physiological compartments. Since the NSG peptoids are homologous in their backbone structure to peptides, we have begun our investigations with a comparison of the relative stabilities of homologous sequences of all-L, all-D, all-N, and retro all-N peptides/peptoids toward representative classes of proteases that compounds would be expected to encounter in vivo. Additional studies of de novo leads are being pursued and will be reported in detail elsewhere; however, stability studies of one de novo lead (CHIR 2279 [Zuckermann et al. 1994]) do indicate that the compound is proteolytically stable upon extended storage in plasma (Gibbons and Spear, unpublished results).

For the present study, we selected enzymes representative of the major classes of relevant proteases as delineated in Table 1: carboxypeptidase A, papain, pepsin, trypsin, elastase, chymotrypsin. From the extensive data in the literature on sequence specificities for these proteases, we designed L-amino acid peptide substrates, for which we could synthesize fairly homologous D-amino acid containing and N-substituted glycine containing oligomers. Parallel and anti-parallel (retro) NSG peptoids were prepared to cover both the N to C and C to N reading frames along the peptoid. As illustrated by the generic structures in Figure 1, the all-D peptides, being enantiomers of the normal all-L substrates, are not expected to bind properly to the protease active sites, and, hence, should not be subject to cleavage. Likewise, the parallel all-N sequence results in a misalignment of the side chains and the carbonyl groups, causing the susceptible peptide bond to be out of range of the normal nucleophilic catalysts at the active sites, and hence, not cleavable. The retro all-N sequence is more likely to mimic the parent peptide for binding, since the orientation of the side chains and the car-

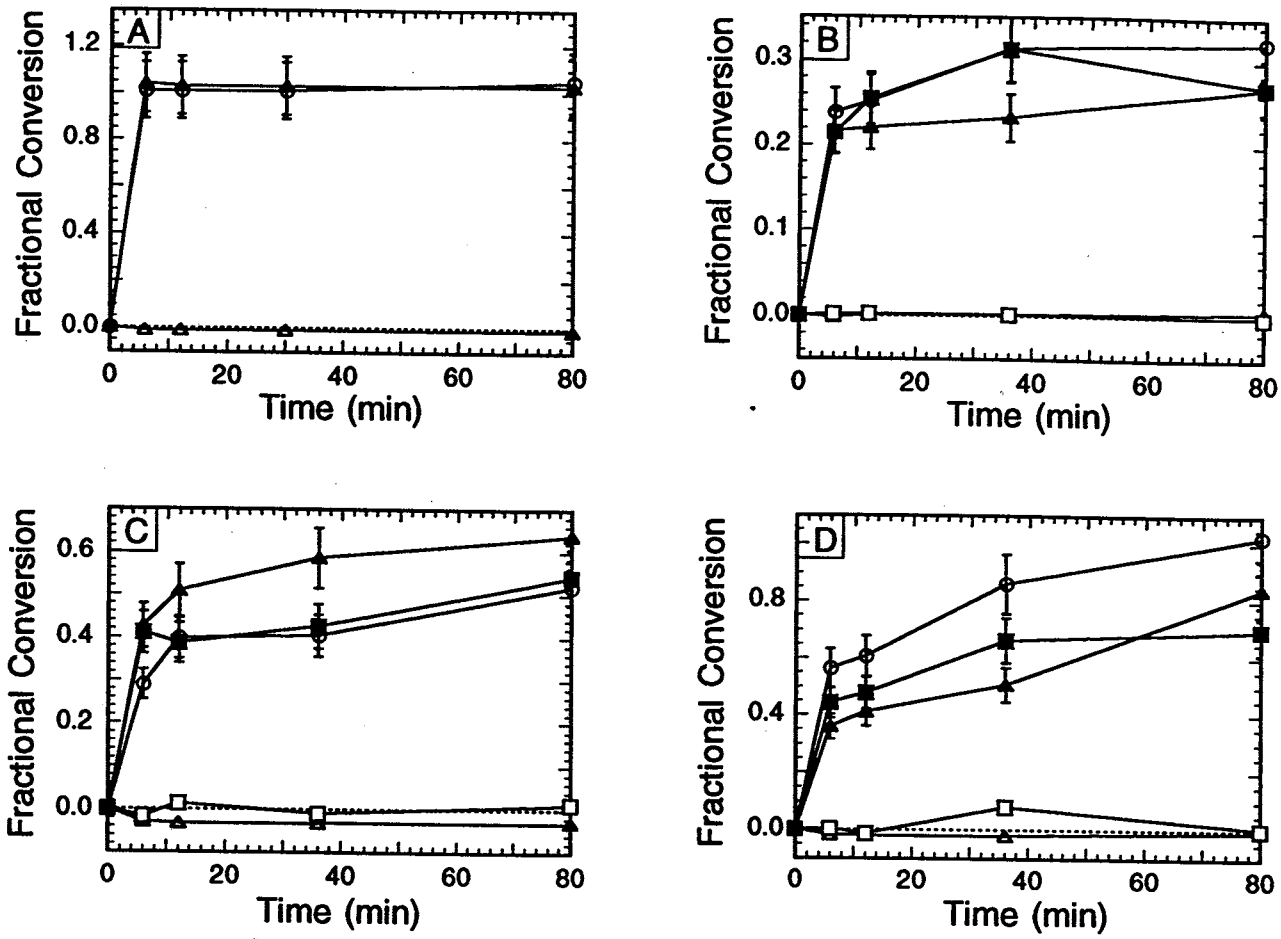


Figure 6. Effect of all-N and retro all-N peptoids on all-L peptide cleavage. Reaction conditions and concentrations of peptides and peptoids for each enzyme are as indicated in Table 2. For the mixtures, equal volumes of the stock peptide and peptoid solutions were diluted to give the same concentrations as in the individual incubations. Fractional conversions and errors are determined as in Figure 5. For each enzyme, reactions are: (○) all-L peptide alone, (△) all-N peptoid alone, (▲) all-L + all-N mixture, (□) retro all-N peptoid alone, (■) all-L + retro all-N mixture. Enzymes shown are: (A) carboxypeptidase A, (B) papain, (C) trypsin, (D) elastase.

bonyl groups are the same for the two (Fig. 1). Note, however, that the polarity of the cleavable C-N bond would be reversed at the active site for the retro peptoid, such that the nitrogen would be out of reach of the normal enzymic groups involved in protonation of the leaving amino group. Thus, binding may be optimized for the retro all-N sequences, but they too might be expected to be stable to proteolysis.

After synthesis and purification of the requisite compounds, we compared their stabilities against the proteases, as measured by fluoresence resulting from conjugation of the cleaved amines with 4-fluoro-7-nitro-benzofurazan, a compound that shows good reactivity with both 1° and 2° amines. The L-amino acid containing peptides designed for this study were readily proteolyzed by all of the enzymes studied. In contrast, but as expected from the above analysis, the D-amino acid containing peptides, N-substituted glycine containing peptoids and their anti-parallel or "retro" sequences, were essentially untouched by the enzymes studied. While this study was designed to compare the proteolytic susceptibility of homologous NSG-peptoid and amino acid peptide sequences to specific proteases, it does not address the more general question of proteolytic stability of random NSG-peptoids to a variety of proteases. However, we have previously examined the susceptibility of several random NSG-peptoid sequences to proteolysis by a panel of different proteases and have not found any evidence for hydrolysis of the peptide bonds [Simon et al., 1992] (Miller et al., unpublished results). Results of the inhibition experiments indicate that some of the

NSG peptoids designed for the study do appear to bind as they weakly inhibit proteolysis of the all-L peptides. It is interesting to note that both forward and reverse peptoids designed for elastase inhibited proteolysis of the homologous all-L peptide, while neither of the peptoids designed for trypsin showed any inhibition. This may reflect the greater importance of extended binding interactions of ligands with elastase as compared with a greater importance of the P_1 residue in binding of ligands to trypsin.

For the quantitation of the extent of hydrolysis of the compounds, we have assumed that cleavage occurs at the sites indicated in Figure 2. With the exception of elastase, the proteases studied have fairly well defined specificities; thus, the assumption is likely to be valid. In the case of elastase, the results clearly indicate multiple cleavage of the all-L peptide, as expected from the low specificity of the enzyme even though the quantitation is not perfect. Another factor that could affect the quantitation is that the relative fluorescence for the expected di- or tripeptide products may differ from that obtained for the simple amino acids used as standards. However, data from the literature indicate only small variations in fluorescence yield (no more than 10%) when reactions of equimolar free amino acid and the same amino acid at the N-termini of several different peptides are compared [Codini et al., 1991]. At this point, the quantitation may not be perfect, but a more complete quantitation seems unnecessary since the reactions with the NSG-peptoids showed no significant fluorescence above the background control reactions, i.e., none of the NSG-peptoids showed any proteolytic cleavage.

In conclusion, under conditions wherein L-amino acid containing peptides are rapidly hydrolyzed, homologous D-amino acid, N-substituted glycine, and retro-sequence N-substituted glycine containing oligomers were essentially untouched by a series of proteases. NSG peptoid diversities thus join the medicinal chemist's repertoire of traditional peptidomimetic and other approaches designed to enhance the absorption, distribution, metabolism, and excretion characteristics of peptides.

ACKNOWLEDGMENTS

The authors thank Cathy Chu for performing the amino acid analyses and Prof. Jack Kirsch for a critical reading of the manuscript.

REFERENCES

Abramowitz N, Schechter I, Berger A (1967): On the size of the active site in proteases. II. Carboxypeptidase-A. Biochem Biophys Res Commun 29:862–867.

Bauer CA, Thompson RC, Blout ER (1976): The active centers of *Streptomyces griseus* protease 3, α-chymotrypsin, and elastase: Enzyme-substrate interactions close to the scissile bond. Biochemistry 15:1296–1299.

Codini M, Palmerini CA, Fini C, Lucarelli C, Floridi A (1991): High-performance liquid chromatographic method for the determination of prolyl peptides in urine. J Chromatogr 536:337–341.

Conradi RA, Hilgers AR, Ho NFH, Burton PS (1992): The influence of peptide structure on transport across Caco-2 cells. II. Peptide bond modification which results in improved permeability. Pharm Res 9:435–438.

Desai MC, Zuckermann RN, Moos WH (1994): Recent advances in the generation of chemical diversity libraries. Drug Dev Res 33:174–188.

Fischer G, Bang H, Berger E, Schellenberger A (1984): Conformational specificity of chymotrypsin toward proline-containing substrates. Biochim Biophys Acta 791:87–97.

Fruton JS (1976): The mechanism of the catalytic action of pepsin and related acid proteinases. Adv Enzymol 44:1–36.

Furka Á, Sebestyen M, Asgedom M, Dibó G (1991): General method for rapid synthesis of multicomponent peptide mixtures. Int J Pept Protein Res 37:487–493.

Gallop MA, Barrett RW, Dower WJ, Fodor SP, Gordon EM (1994): Applications of combinatorial technologies to drug discovery. 1. Background and peptide combinatorial libraries. J Med Chem 37:1233–1251.

Gordon EM, Barrett RW, Dower WJ, Fodor SP, Gallop MA (1994): Applications of combinatorial technologies to drug discovery. 2. Combinatorial organic synthesis, library screening strategies, and future directions. J Med Chem 37:1385–1401.

Harper JW, Cook RR, Roberts CJ, McLaughlin BJ, Powers JC (1984): Active site mapping of the serine proteases human leukocyte elastase, cathepsin G, porcine pancreatic elastase, rat mast cell proteases I and II, bovine chymotrypsin A_α, and *Staphylococcus aureus* protease V-8 using tripeptide thiobenzyl ester substrates. Biochemistry 23:2995–3002.

Houghten R, Pinilla C, Blondelle S, Appel J, Dooley C, Cuervo J (1991): Generation and use of synthetic peptide combinatorial libraries for basic research and drug discovery. Nature 354:84–86.

Huber R, Bode W (1978): Structural basis of the activation and action of trypsin. Accounts Chem Res 11:114–122.

Imai K, Watanabe Y (1981): Fluorimetric determination of secondary amino acids by 7-fluoro-4-nitrobenzo-2-oxa-1,3-diazole. Anal Chim Acta 130:377–383.

Imperiali B, Abeles RH (1987): Extended binding inhibitors of chymotrypsin that interact with leaving group subsites $S_{1'}$–$S_{3'}$. Biochemistry 26:4474–4477.

Lam K, Salmon S, Hersh E, Hruby V, Kazmiersky W, Knapp R (1991): A new type of synthetic peptide library for identifying ligand-binding activity. Nature 354:82–84.

Lowbridge J, Fruton JS (1974): Studies on the extended active site of papain. J Biol Chem 249:6754–6761.

Maeji NJ, Bray AM, Valerio RM, Seldon MA, Wang JX, Geysen HM (1991): Systematic screening of bioactive peptides. Pept Res 4:142–146.

Miller SM, Simon RJ, Ng S, Zuckermann RN, Kerr JM, Moos WH (1994): Proteolytic studies of homologous peptide and N-substi-

tuted glycine peptoid oligomers. Bioorg Med Chem Lett 4:2657–2662.

Miyano H, Toyo'oka T, Imai K (1985): Further studies on the reaction of amines and proteins with 4-fluoro-7-nitrobenzo-2-oxa-1,3-diazole. Anal Chim Acta 170:81–87.

Polgár L (1989): Mechanisms of Protease Action. Boca Raton, FL: CRC Press, Inc.

Powers JC, Harley AD, Myers DV (1977): Subsite specificity of porcine pepsin. In Tang J (ed): Acid Proteases, Structure, Function, and Biology. New York: Plenum Press, pp 141–157.

Pozsgay M, CS. Szabó G, Bajusz S, Simonsson R, Gáspár R, Elodi P (1981): Investigation of the substrate-binding site of trypsin by the aid of tripeptidyl-p-nitroanilide substrates. Eur J Biochem 115:497–502.

Rees DC, Lipscomb WN (1982): Refined crystal structure of the potato inhibitor complex of carboxypeptidase A at 2.5 Å resolution. J Mol Biol 160:475–498.

Schechter I, Berger A (1967): On the size of the active site in proteases. I. Papain. Biochem Biophys Res Commun 27:157–162.

Schechter I, Berger A (1968): On the active site of proteases. III. Mapping the active site of papain; specific peptide inhibitors of papain. Biochem Biophys Res Commun 32:898–902.

Schenk, GH (1973): Absorption of Light and Ultraviolet Radiation: Fluorescence and Phosphorescence Emission. Boston: Allyn and Bacon, Inc.

Schomburg D, Salzmann M (eds) (1990): Enzyme Handbook. Berlin: Springer-Verlag.

Simon RJ, Kania RS, Zuckermann RN, Huebner VD, Jewell DA, Banville S, Ng S, Wang L, Rosenberg S, Marlowe CK, Spellmeyer DC, Tan R, Frankel AD, Santi DV, Cohen FE, Bartlett PA (1992): Peptoids: A modular approach to drug discovery. Proc Natl Acad Sci USA 89:9367–9371.

Thompson RC, Blout ER (1973): Dependence of the kinetic parameters for elastase-catalyzed amide hydrolysis on the length of peptide substrates. Biochemistry 12:57–65.

Zuckermann RN, Kerr JM, Kent SBH, Moos WH (1992a): Efficient method for the preparation of peptoids [oligo (N-substituted glycines)] by sub-monomer solid-phase synthesis. J Am Chem Soc 114:10646–10647.

Zuckermann RN, Kerr JM, Siani MA, Banville SC (1992b): Design, construction and application of a fully automated equimolar peptide mixture synthesizer. Int J Pept Protein Res 40:498–507.

Zuckermann RN, Martin EJ, Spellmeyer DC, Stauber GB, Shoemaker KR, Kerr JM, Figliozzi GM, Goff DA, Siani MA, Simon RJ, Banville SC, Brown EG, Wang L, Richter LS, Moos WH (1994): Discovery of nanomolar ligands for 7-transmembrane G-protein coupled receptors from a diverse (N-substituted) glycine peptoid library. J Med Chem 37:2678–2685.