Pergamon

Bioorganic & Medicinal Chemistry Letters, Vol. 4, No. 22, pp. 2657-2662, 1994

0960-894X/94 $7.00+0.00

0960-894X(94)00395-5

PROTEOLYTIC STUDIES OF HOMOLOGOUS PEPTIDE AND N-SUBSTITUTED GLYCINE PEPTOID OLIGOMERS

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

University of California, Department of Pharmaceutical Chemistry, San Francisco, CA 94143-0446 (S.M.M.)
Chiron Corporation, 4560 Horton Street, Emeryville, CA 94608-2916 (R.J.S., S.N., R.N.Z., J.M.K., W.H.M.)

Abstract. Homologous L-amino acid, D-amino acid, and parallel and anti-parallel (retro) sequence N-substituted glycine peptide and peptoid oligomers were prepared and incubated with proteases from each major class. The L-amino acid containing peptides were readily cleaved by the appropriate enzymes, while equivalent D-amino acid containing and N-substituted glycine containing oligomers were essentially untouched.

A new field of pharmaceutical research, "molecular diversity", has exploded in the last several years.[1] In the preparation of drug candidates, the automated and combinatorial use of genetically encoded building blocks, such as the standard 20 amino acids, as well as non-natural building blocks, like N-substituted glycines (NSGs), now allows the generation and screening of unprecedented numbers of compounds.

However, the discovery aspects of molecular diversity approaches represent only one component of pharmaceutical research and development. Many leads ultimately fail because of deficiencies in one or more important pharmaceutical properties, such as absorption, distribution, metabolism, and excretion (ADME), not to mention safety and efficacy. While small 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 libraries, especially for orally available therapeutics (at least until drug delivery methods achieve greater success).

One new class of diversity is based 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.[2] The NSG peptoids, which can mimic both peptides and non-peptides, demonstrate potent and specific biological activity.[3,4] From an ADME perspective, work by Conradi et al. would predict improved absorption characteristics for NSG peptoids because of N-substitution.[5] 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 NSG peptoids against a series of common proteases.

For the purposes of this study, we selected enzymes representative of several known classes of relevant proteases including serine-proteases: chymotrypsin, elastase, trypsin; cysteine-proteases: papain; acid-proteases: pepsin; metallo-proteases: carboxypeptidase A. Based on known sequence specificities, we designed L-amino acid peptide (all-L) substrates that would allow fair comparisons with D-amino acid containing (all-D) and N-substituted glycine containing oligomers. Both parallel (all-N) and anti-parallel (retro all-N) NSGs were prepared to cover both possible reading frames.

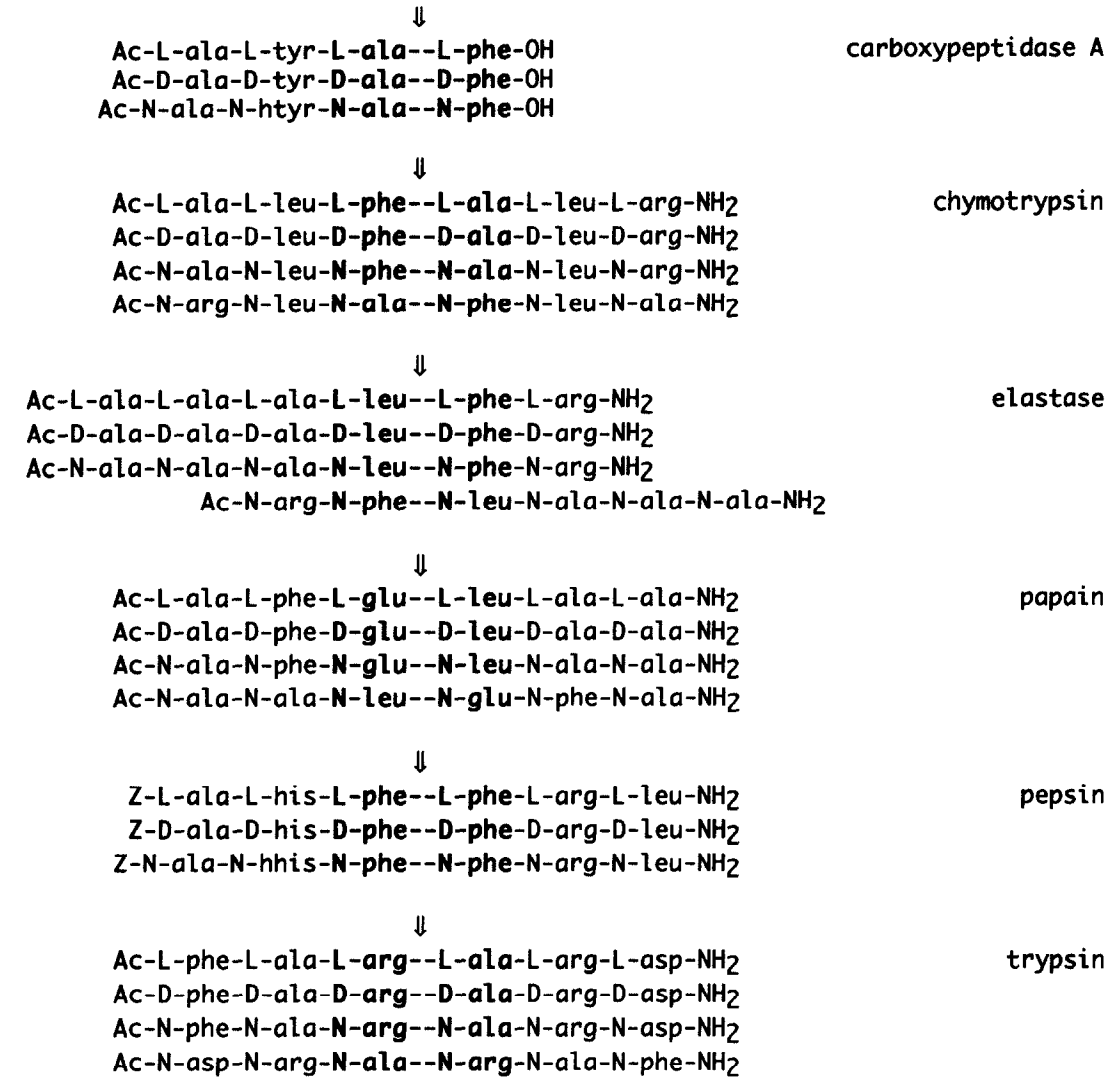

Figure 1. Sequences of peptides and NSG peptoids designed for the indicated proteases. Expected cleavage sites are indicated by boldface residues and arrows.

Design of Peptides and NSG Peptoids. Figure 1 summarizes the peptide and NSG peptoid sequences designed and synthesized for these proteolysis studies. Hexapeptides were chosen for most of the enzymes since previous studies indicate extended recognition sites for most proteases of at least 6 or 7 residues.[6] The sequences are based on literature studies defining good peptide substrates and/or inhibitors. For chymotrypsin, P_3 - P_1 residues are based on the analysis of extended peptides as substrates,[7] while $P_{1'}$ - $P_{3'}$ residues are based

on tight binding inhibitors.[8] 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.[9] Thus, the sequence of P_4 - P_1 is based on studies that show that substitution of L-leu for L-ala at P_1 gives a better k_{cat}/K_m value.[9,10] L-phe at the $P_{1'}$ site of the elastase peptide is based on studies by Bauer et al.;[11] and L-arg at $P_{2'}$ was included to improve solubility of the peptide. For papain, extended binding from P_4 - $P_{3'}$ gives improved activity and the primary specificity is for aromatic residues at P_2.[12,13] Inclusion of L-glu-L-leu at the P_1-$P_{1'}$ sites is based on studies demonstrating the further specificity of the $P_{1'}$ site of papain for large hydrophobic residues in substrates.[14] For pepsin, the P_3 -$P_{1'}$ sequence and the inclusion of a carbobenzyloxy capping group is based on the extensive studies that explore site specificities and demonstrate an extended substrate binding site from P_4 -$P_{3'}$.[15] The L-arg-L-leu sequence in the $P_{2'}$ -$P_{3'}$ position is based on an analysis of cleavage sites reported for a wide variety of protein substrates for pepsin.[16] For trypsin, the P_3 -P_1 sequence is based on substrate studies,[17] while the $P_{1'}$ -$P_{3'}$ fragment is based on the sequence of the pancreatic trypsin inhibitor complexed with trypsin.[18] 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 active site.[19]

Proteolysis Results. Data for the proteolysis of the respective peptides and peptoids are displayed as fractional conversions, by enzyme, in Figure 2. As indicated by the closed circles for each plot, the all-L peptides are indeed substrates for their respective enzymes. The peptide designed for elastase (panel B) clearly undergoes cleavage at more than one site over the course of the experiment. With the other enzymes, a fairly rapid but variable extent of cleavage of the all-L peptides occurs in the first 12 min of the assay followed by a much slower, further cleavage (except for pepsin and carboxypeptidase A, 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 estimated 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 this 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 two all-D peptides (3% and 6% after 20 h with trypsin and carboxypeptidase A), and no convincing cleavage of any of the all-N or retro all-N peptoids. The only NSG peptoid giving any hint of cleavage is the all-N peptoid for papain, which may be cleaved to 0.2%, however, 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. While a lack of cleavage of the NSG peptoids is consistent with a lack of proteolytic susceptibility, there would also be no cleavage if the NSG peptoids were irreversible inhibitors. To test the latter possibility, we followed reactions of the enzymes with mixtures of approximately equimolar concentrations of all-L + all-N (or retro all-N) compounds. Inhibition of the rate of cleavage of the all-L peptides under these conditions appears to be no more than 5% for any of the compounds tested (data not shown), indicating that irreversible reaction of the NSG peptoids with the enzymes has not occurred.

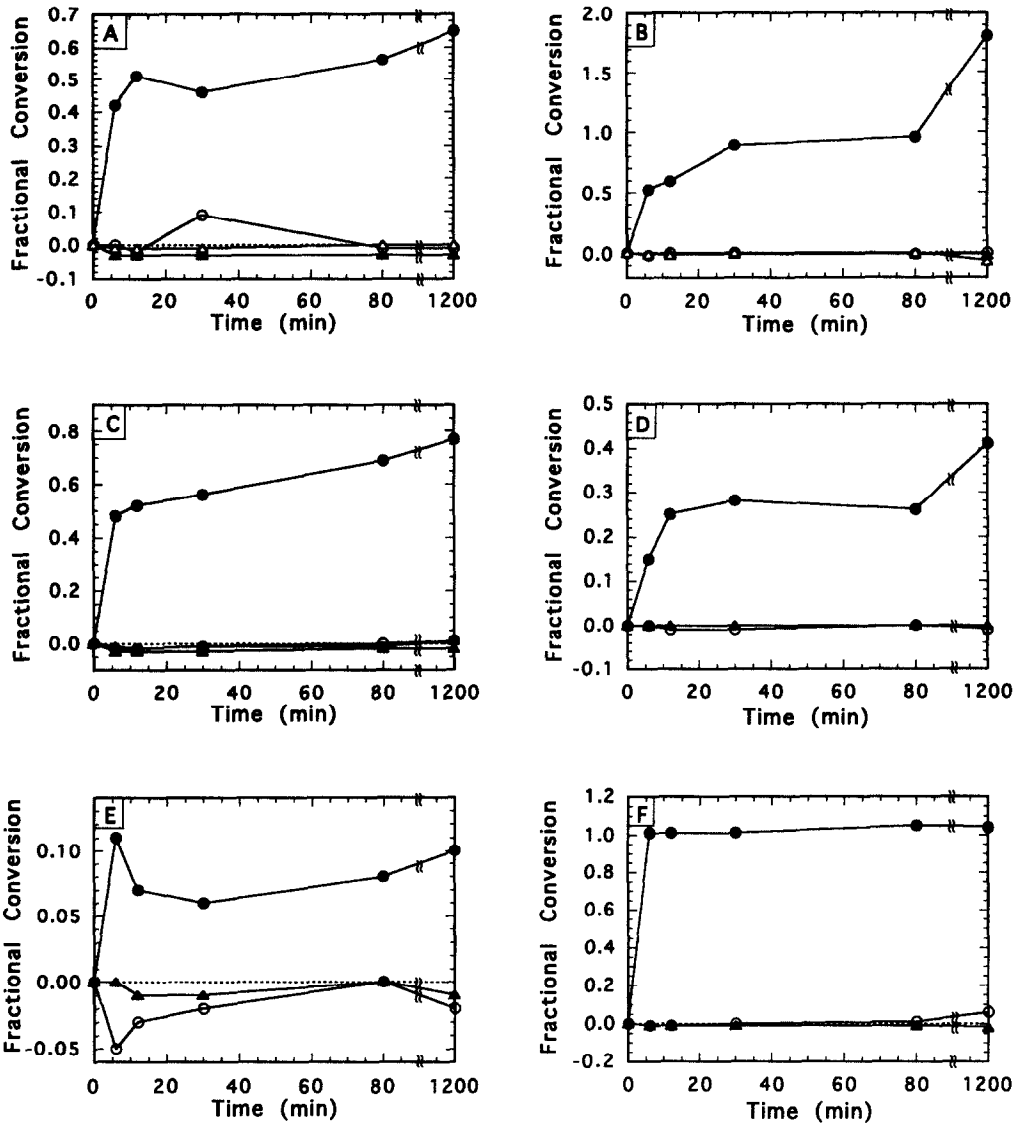

Figure 2. Fractional conversions of peptides and peptoids. Cleavage was monitored by reaction of the free amine products with 4-fluoro-7-nitro-benzofurazan to produce fluorescent products.[20] Concentrations of amino product were calculated using the appropriate standard for the resultant amino terminal residue at the expected cleavage site, as indicated in Figure 1, and are divided by the initial substrate concentration to obtain fractional conversion. For each panel (●) all-L peptide, (○) all-D peptide, (▲) all-N peptoid, (Δ) retro all-N peptoid. Enzymes are (A) chymotrypsin, (B) elastase, (C) trypsin, (D) papain, (E) pepsin, (F) carboxypeptidase A.

Discussion While molecular diversity approaches to lead discovery and lead optimization are receiving increasing attention, few groups have addressed potential development hurdles such as ADME characteristics. 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. Our pursuit of NSG peptoids both as homologs of peptides[3] and as *de novo* leads[21], has led us to begin studies of the proteolytic stabilities of these compounds. In this paper we have begun 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*. Further 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)[21] do indicate that the compound is proteolytically stable upon extended storage in plasma (J. Gibbons & K. Spear, Chiron, unpublished results).

For the present study, we selected enzymes representative of the major classes of relevant proteases: carboxypeptidase A, chymotrypsin, elastase, papain, pepsin, trypsin. Based on known sequence specificities for these proteases, we designed L-amino acid peptide substrates that would allow fair comparisons with D-amino acid containing and N-substituted glycine containing oligomers. Parallel and anti-parallel (retro) NSG peptoids were prepared to cover both possible reading frames along the peptoid (i.e., the "register", N to C, or C to N). 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 carbonyl groups are the same for the two. 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 synthesizing 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.[20] The L-amino acid containing peptides chosen for this study were readily proteolyzed by all of the enzymes studied. In contrast, D-amino acid containing peptides, N-substituted glycine containing peptoids and their anti-parallel or "retro" sequences were essentially untouched by the enzymes studied.

For quantitating the extent of hydrolysis of the compounds, we have assumed that cleavage occurs at the sites indicated in Figure 1. 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.

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.

References and Notes

1. Moos, W.H.; Green, G.D.; Pavia, M.R. *Annu. Rep. Med. Chem.* **1993,** *28,* 315.
2. Zuckermann, R.N.; Kerr, J.M.; Kent, S.B.H.; Moos, W.H. *J. Am. Chem. Soc.* **1992,** *114,* 10646.
3. Simon, R.J.; Kania, R.S.; Zuckermann, R.N.; Huebner, V.D.; Jewell, D.A.; Banville, S.; Ng, S.; Wang, L.; Rosenberg, S.; Marlowe, C.K.; Spellmeyer, D.C.; Tan, R.; Frankel, A.D.; Santi, D.V.; Cohen, F.E.; Bartlett, P.A. *Proc. Natl. Acad. Sci. USA* **1992,** *89,* 9367.
4. Zuckermann, R.N.; Martin, E.J.; Spellmeyer, D.C.; Stauber, G.B.; Shoemaker, K.R.; Kerr, J.M.; Figliozzi, G.M.; Goff, D.A.; Siani, M.A.; Simon, R.J.; Banville, S.C.; Brown, E.G.; Wang, L.; Richter, L.S.; Moos, W.H. *J. Med. Chem.* **1994,** *37,* 2678.
5. Conradi, R.A.; Hilgers, A.R.; Ho, N.F.H.; Burton, P.S. *Pharm. Res.* **1992,** *9,* 435.
6. Polgár, L. *Mechanisms of Protease Action.*; CRC Press, Inc.: Boca Raton, 1989.
7. Fischer, G.; Bang, H.; Berger, E.; Schellenberger, A. *Biochim. Biophys. Acta* **1984,** *791,* 87.
8. Imperiali, B.; Abeles, R.H. *Biochemistry* **1987,** *26,* 4474.
9. Thompson, R.C.; Blout, E.R. *Biochemistry* **1973,** *12,* 57.
10. Harper, J.W.; Cook, R.R.; Roberts, C.J.; McLaughlin, B.J.; Powers, J.C. *Biochemistry* **1984,** *23,* 2995.
11. Bauer, C.A.; Thompson, R.C.; Blout, E.R. *Biochemistry* **1976,** *15,* 1296.
12. Schechter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967,** *27,* 157.
13. Schechter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1968.** *32,* 898.
14. Lowbridge, J.; Fruton, J.S. *J. Biol. Chem.* **1974,** *249,* 6754.
15. Fruton, J.S. *Adv. Enzymol. Rel. Areas Mol. Biol.* **1976,** *44,* 1.
16. Powers, J.C.; Harley, A.D.; Myers, D.V. *Acid Proteases, Structure, Function, and Biology;* Tang, J., Ed.; Plenum: New York, 1977; p 141.
17. Pozsgay, M.; CS. Szabó, G.; Bajusz, S.; Simonsson, R.; Gáspár, R.; Elodi, P. *Eur. J. Biochem.* **1981,** *115,* 497.
18. Huber, R.; Bode, W. *Acc. Chem. Res.* **1978,** *11,* 114.
19. Rees, D.C.; Lipscomb, W.N. *J. Mol. Biol.* **1982,** *160,* 475.
20. Miyano, H.; Toyo'oka, T.; Imai, K. *Anal. Chim. Acta* **1985,** *170,* 81.
21. Zuckermann, R.N.; Martin, E.J.; Spellmeyer, D.C.; Stauber, G.B.; Shoemaker, K.R.; Kerr, J.M.; Figliozzi, G.M.; Goff, D.A.; Siani, M.A.; Simon, R.J.; Banville, S.C.; Brown, E.G.; Wang, L.; Richter, L.S.; Moos, W.H. *J. Med. Chem.* **1994,** *37,* 2678.

(*Received in USA* 8 *September* 1994; *accepted* 5 *October* 1994)