J. Mol. Biol. (1996) **259**, 819–827

JMB

Selection of Chymotrypsin Inhibitors from a Conformationally-constrained Combinatorial Peptide Library

Jeffrey D. McBride[1], Neil Freeman[2], Gonzalo J. Domingo[1] and Robin J. Leatherbarrow[1]*

[1]*Department of Chemistry Imperial College of Science Technology and Medicine South Kensington, London SW7 2AY, UK*

[2]*Department of Biomolecular Structure, Glaxo Research and Development Ltd Stevenage, Herts, SG1 2NY UK*

**Corresponding author*

A synthetic library of cyclic peptides was constructed utilizing the anti-tryptic loop region of the Bowman-Birk inhibitor, D4 from *Macrotyloma axillare*, as a template. The loop region of this proteinase inhibitor was reproduced by an 11 residue sequence, conformationally constrained by the presence of a disulfide bridge, to act as a mimetic of the functional reactive site region of this protein. This sequence, plus a pentaglycine spacer arm, was used to create a ''one bead, one peptide'' combinatorial library after on-resin deprotection and cyclization. Randomization at three positions considered to be important for proteinase specificity (P_2, P_1 and P_2') with the genetically coded amino acids (minus cysteine) plus norleucine generated 8000 permutations. Screening this library with biotinylated α-chymotrypsin under appropriate conditions revealed a small number (<0.05%) of beads that selectively bound the labeled proteinase. The sequences present on these active beads were determined, and found to have a well-defined consensus. Analysis of chymotrypsin inhibition in solution using re-synthesized peptides reveals that the sequences identified are potent inhibitors with K_i values in the nanomolar range. These results show that directed randomization of the canonical loop is a powerful way of generating proteinase inhibitors with targeted specificities. Incorporation of selective random changes within a defined structural framework is found to be an effective means of generating variation in large synthetic systems. The functional basis for inhibition by the identified sequences is discussed.

Keywords: combinatorial chemistry; proteinase inhibitor; Bowman-Birk inhibitor; canonical loop; protein engineering

Introduction

Almost all biochemical processes rely upon specific molecular recognition between peptides or proteins and other biologically active molecules. Combinatorial approaches have recently provided new strategies capable of selecting bioactive molecules from large libraries of compounds (for reviews, see Gallop *et al*., 1994; Gordon *et al*., 1994). Such methods reduce the need for repetitive syntheses of related compounds, and closely link the synthetic and screening processes.

Proteinases play important roles in the regulation of many biological processes, by post-translational modification of precursor proteins or by controlling protein turnover (Neurath, 1989). Regulation of proteolytic activity in such systems is vital, and many pathways that deploy proteinase enzymes use specific inhibitor proteins to control the balance between desired and undesired proteolysis. Like their target proteinases, inhibitors are ubiquitous, and can be found throughout the animal and plant

Present address: G. J. Domingo, Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge, UK.

Abbreviations used: BBI, Bowman-Birk inhibitor; PBS-T, phosphate buffered saline containing 0.05% (v/v) Tween-20, pH 7.4; P_n, the *n*th peptide subsite of a substrate or inhibitor on the carboxyl side of the scissile bond; P_n', the *n*th peptide subsite of a substrate or inhibitor on the amino side of the scissile bond; STI, soybean trypsin inhibitor.

0022–2836/96/240819–09 $18.00/0

kingdoms. The pivotal role of proteinases in biological regulation makes them important targets for therapeutic intervention, and there is much interest in identification of new inhibitors. A key question in inhibitor design is how to construct novel inhibitors with directed specificities. Combinatorial libraries of potential inhibitor compounds offer an attractive way to achieve this.

Inhibition by proteinase inhibitor proteins involves the formation of a stoichiometric enzyme-inhibitor complex, which resembles a Michaelis complex. Unlike a good substrate, however, the energy barrier for hydrolysis is large and unfavorable, resulting in low rates of hydrolysis (Laskowski & Kato, 1980; Read & James, 1986; Longstaff *et al*., 1990). In most cases the inhibition is mediated by an exposed loop that is fixed in a characteristic ''canonical'' conformation, and which fits into the active site of the proteinase in a manner thought to be similar to that of a substrate (Laskowski & Kato, 1980; Bode & Huber, 1992). This loop is frequently constrained by the presence of disulfide bridges and/or extensive hydrogen-bonding networks that act to lock the structure into the correct canonical structure. The sequence of this loop determines the specificity of the inhibition, which mirrors the specificity of proteinases for their substrates. For example, most trypsin inhibitors have Arg or Lys as their P_1 residue.

Inhibitors of the Bowman-Birk family have a network of conserved disulfide bridges that help form a symmetrical structure of two tricyclic domains (Chen *et al*., 1992; Werner & Wemmer, 1992; Lin *et al*., 1993), each containing an independent serine proteinase binding site (Figure 1A). The binding loop is contained within a nonapeptide region joined *via* a disulfide between flanking cysteine residues (Figure 1B). The identity of the amino acid residue at the P_1 site on each domain is the main determinant of the serine proteinase inhibited. Native domains possess lysine or arginine for trypsin, leucine or tyrosine for chymotrypsin, and alanine for elastase (Ikenaka & Norioka, 1986). In addition, serine is highly conserved at the P_1' position and proline at the P_3' position.

The individual loop regions of Bowman-Birk inhibitor (BBI) are well suited as a models for proteinase inhibitor design. Nishino *et al*. (1977) first demonstrated that a nine residue monocyclic peptide based on the trypsin binding domain was sufficient to retain the inhibitory properties of the native inhibitor. More recently, Maeder *et al*. (1992) designed an 11 residue monocyclic peptide based on the core nine residue sequence of the anti-tryptic loop region of the Bowman-Birk inhibitor, D4 from *Macrotyloma axillare*. This peptide exhibits an improved K_i value against trypsin, and better resistance to proteolysis (unpublished results). We have used this peptide sequence to produce a defined mixture of P_1 variants in order to study proteinase-inhibitor recognition (Domingo *et al*., 1995). We now report the creation of an immobilized cyclic combinatorial peptide library using this 11 residue sequence as a template to screen for inhibitors active against α-chymotrypsin. Multiplicity has been generated by incorporating the natural amino acids (minus cysteine but plus norleucine) in the P_2, P_1 and P_2' positions using the ''one bead, one peptide'' approach, obtained from split synthesis (Furka *et al*., 1991; Lam *et al*., 1991), to generate 8000 variants. To date, most combinatorial synthetic libraries have been created using complete randomization. Our approach contrasts with this, in that we have elected to conserve a framework of residues in order to maintain a structure-inducing scaffold from our archetypal inhibitor. This will result in a library that, while less diverse than those with totally random sequences, is better targeted at the generation of active proteinase inhibitors.

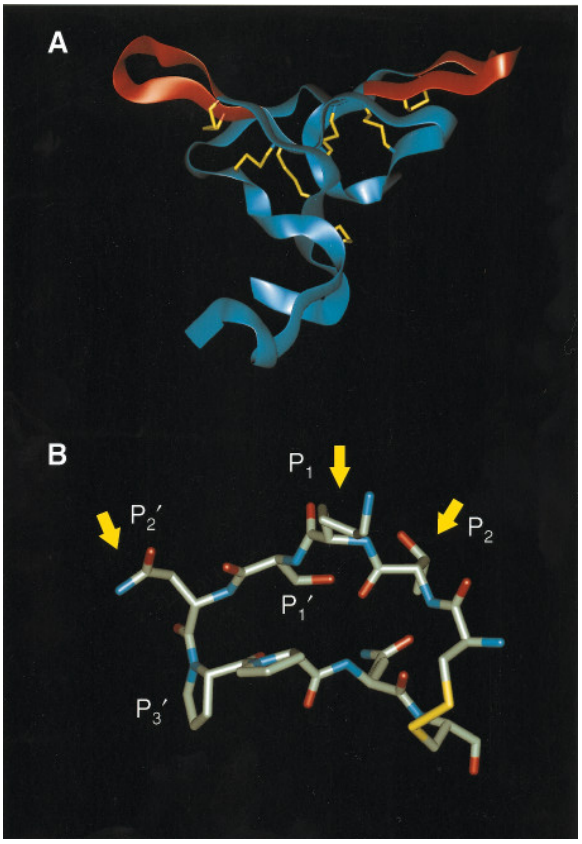

Figure 1. Structure of the Bowman-Birk inhibitor. The coordinates of Bowman-Birk inhibitor protein (Werner & Wemmer, 1992) were obtained from the Brookhaven Protein Data Bank and displayed using the program Quanta. A, The backbone of the Bowman-Birk inhibitor is shown as a ribbon. The two proteinase-binding loops are colored red, and the side-chains of the cystine residues are in yellow. B, Detail of the three-dimensional structure of the binding loop. The residues that were randomized are indicated by arrows.

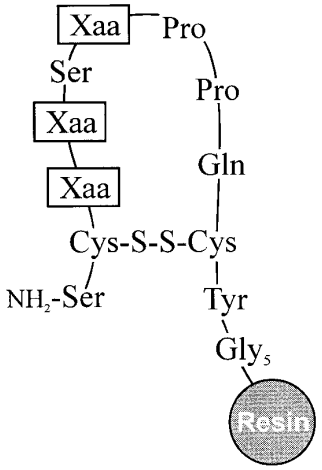

Figure 2. Sequence of the inhibitor library. The resin-bound inhibitor library was synthesized with a pentaglycine spacer between the TentaGel support and an 11 residue inhibitor sequence containing three sites that were subject to randomization. The random sites, marked Xaa, had all the genetically encoded amino acids minus cysteine, plus norleucine to give 20 possible combinations. This results in a library with 8000 permutations.

Results

C-terminal attachment

Disulfide-linked peptide loops based on the sequence of Bowman-Birk inhibitor retain much of the inhibitory capability of the parent inhibitor protein (Nishino *et al*., 1977; Terada *et al*., 1978; Maeder *et al*., 1992; Domingo *et al*., 1995). However, in order to screen for inhibitors using covalently linked peptides, it is first essential to demonstrate that C-terminal attachment is not detrimental to the inhibition properties. The K_i value for inhibition of trypsin by the cyclic peptide SCTKSIPPQCY is 10 nM (Domingo *et al*., 1995). The addition of a pentaglycine spacer resulted in an inhibitor with a K_i value of 4 nM, which shows that addition of a C-terminal tail of glycine residues does not markedly change the inhibition constant. Immobilization of this latter sequence *via* the C terminus to a solid support was found not to prevent trypsin binding (not shown), showing that covalent attachment of this peptide loop is a viable way to screen for proteinase binding.

Synthesis of the inhibitor library

The cyclic peptide library shown in Figure 2 was synthesized by inclusion of three split-synthesis steps (Furka *et al*., 1991; Lam *et al*., 1991). Edman sequencing of several individual resin beads was performed to verify that the correct sequence had been generated. No apparent bias was observed in these sequences, and the randomized positions showed the expected arbitrary distribution of sequences (results not shown).

Screening of the inhibitor library

Screening of the library was performed indirectly by using biotinylated α-chymotrypsin, which was detected by means of an avidin-alkaline phosphatase conjugate to provide an enzyme-linked assay. For this reason, it was first important to establish that biotinylation of chymotrypsin did not perturb the enzyme activity. To ensure that only active species were present, the biotin-chymotrypsin was purified by affinity chromatography on STI-Sepharose. The kinetic parameters of the modified enzyme (biotin to protein 0.9:1; K_m 53 μM, k_{cat} 42 s^{-1}) were found not to differ significantly from the unmodified species (K_m 48 μM, k_{cat} 49 s^{-1}).

The library was screened using a relatively low concentration of biotinylated chymotrypsin in order to select only the most tightly binding species. This procedure was found to result in the selection of a small number of beads (approximately 13 very highly stained beads out of a total of ~32,000). A typical ''positive'' bead is shown in Figure 3, illustrating the great distinction between the positive beads and the remainder. The positive beads were individually isolated, and their sequences determined by Edman degradation.

Sequences discovered

In all cases only a single sequence was obtained from each bead, and repetitive yields from the Edman reactions were 85 to 90% at each position. Of the 13 sequences that were found, ten were confirmed as potent competitive inhibitors of chymotrypsin after subsequent re-synthesis; these sequences are given in Table 1. In addition, we found three sequences that did not display inhibition activity after re-synthesis. These sequences were $SCEWSIPPQCG_5$, $SCTTSTPPQCG_5$ and $SCQQSYPPQCG_5$. Multiple copies were found for most of the active peptides; for example, the sequence $SCTFSIPPQCYG_5$ was found on five independent beads. Of the positions randomized, all the active sequences found had threonine at the P_2 position; tyrosine and phenylalanine were present at P_1; leucine, norleucine and isoleucine were obtained at the P_2' site. Each of the sequences was re-synthesized as soluble 11 residue peptides (the spacer was omitted), and the activity of these molecules as chymotrypsin inhibitors was measured. The inhibition constants are given in Table 1. All the active sequences were found to display low K_i values of ~20 nM.

Discussion

Combinatorial libraries

Synthetic combinatorial libraries represent a significant advance in the study of molecular recognition and compound screening. They allow a large number of distinct species to be synthesized

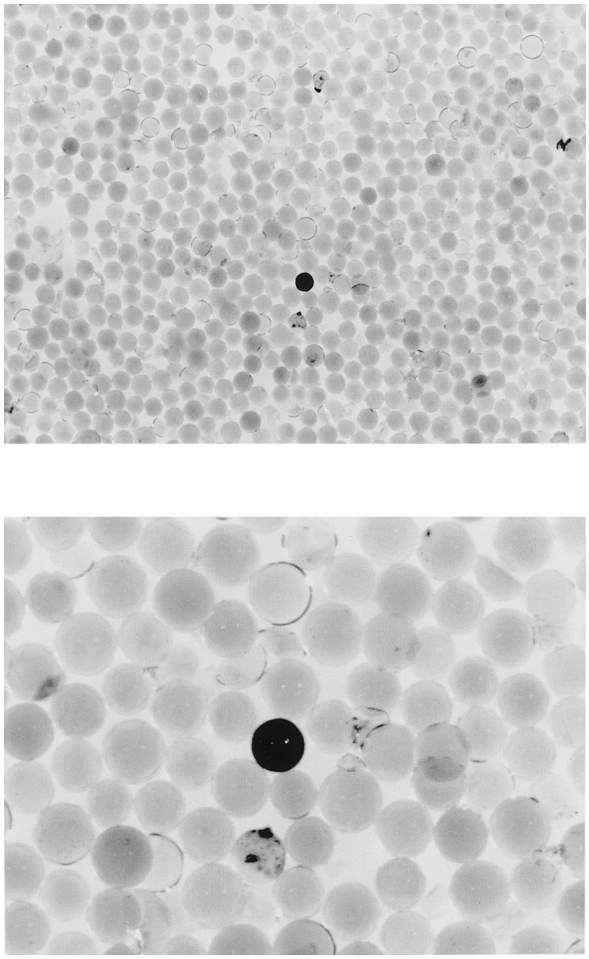

Figure 3. Identification of chymotrypsin-binding beads. The library was incubated with biotinylated α-chymotrypsin. After washing, beads that had retained bound chymotrypsin were visualized by incubation with Extravidin-alkaline phosphatase followed by addition of BCIP/NBT. This treatment results in positive beads becoming stained a deep purple. Such positive beads were removed and subjected to Edman sequencing. The Figure shows a sample of the library containing a single positive bead, at two different levels of magnification. The mean bead diameter is 80 μm.

and tested in a short space of time. In principle, the techniques are applicable to any synthetic methodology (Gordon *et al*., 1994), although most applications reported so far, including the present work, have used libraries built from peptide components (Gallop *et al*., 1994). There are two conceptually distinct classes of synthetic library: those where the library components are present as a solution-phase mixture (Houghten *et al*., 1991; Dooley & Houghten, 1993), and those that involve components covalently linked to a solid support (Geysen *et al*., 1986; Fodor *et al*., 1991; Lam *et al*., 1991). Each has advantages and disadvantages. Solution libraries can be tested under conditions that resemble more closely a typical inhibition or binding assay, but identification of the active components can be ambiguous as such assays always involve complex mixtures. Covalently linked libraries usually allow unambiguous identification of active sequences, but screening covalently coupled compounds may be inappropriate for certain systems and it is essential that the compound retains activity when covalently attached. However, this latter problem can be overcome by a hybrid approach, where a resin-bound library incorporates cleavable linkers that allow release and assay in solution (Salmon *et al*., 1993).

In addition to synthetic libraries, it is possible to generate variation by genetic means. Randomized peptides can be displayed on the surface of filamentous phage (Cwirla *et al*., 1990; Devlin *et al*., 1990; Scott & Smith, 1990). The ''phage-display'' method is a powerful means of selecting active peptide and protein sequences, and has been used to good effect in the engineering of larger protein domains, particularly antibodies (McCafferty *et al*., 1990; Hoogenboom *et al*., 1991). The major disadvantage of phage display is its limitation to only the genetically coded amino acids. There is also some possibility of biological bias in the sequences generated by an expression system (Kay, 1994).

Proteinase inhibition is obviously an important target for combinatorial studies, and there are several examples where libraries have been used to screen for proteinase inhibitors. Multicomponent tetrapeptide mixtures have been used to search for HIV proteinase inhibitors (Owens *et al*., 1991). Solution libraries of linear peptides have been used to identify trypsin inhibitors (Eichler & Houghten, 1993), and the same group has used solution libraries of derivatized cyclic Lys_3Glu to screen for

Table 1. Sequences of active chymotrypsin inhibitors identified from the library

Library template	S C **X** **X** S **X** P P Q C Y G_5	Number of independent sequences found	K_i value for the re-synthesized peptide (nM)
Identified chymotrypsin inhibitors	S C **T** **F** S **I** P P Q C Y G_5	5	19
	S C **T** **F** S **L** P P Q C Y G_5	1	20
	S C **T** **F** S **Nle** P P Q C Y G_5	2	19
	S C **T** **Y** S **I** P P Q C Y G_5	2	17

The positions marked **X** in the library template represent the randomized positions corresponding to the P_2, P_1 and P_2' inhibitor sites, respectively. Twenty different amino acids were introduced at each of these loci, giving a total library size of 8000 sequences. The library template was cyclized by a disulfide between the two cysteine residues, and was attached to TentaGel resin *via* the C-terminal residue. The K_i values were measured in solution assays using re-synthesized 11-mer peptides (the glycine spacer was omitted).

chymotrypsin inhibition (Eichler *et al*., 1994). Meldal & Svendsen (1995) have made use of a novel fluorogenic detection system to identify linear peptide inhibitors in a combinatorial bead library. Using phage display of BPTI variants, Roberts *et al*. (1992) have selected novel neutrophil elastase inhibitors.

Design of our library

For construction of the library reported here, we elected to randomize the P_2, P_1 and P_2' residues on the inhibitor loop. There were two concerns in this selection. First, we wished to target residues that were known to be important for interactions with the proteinase in the natural BBI-trypsin complex. This means that residues adjacent to the scissile bond become important candidates for variation. Second, it is impracticable to vary very many sites, or the library size becomes excessively large. Tsunogae *et al*. (1986) first described how the inhibitory loop of BBI was a two-stranded antiparallel β-sheet comprising two regions separated by the *cis*-Pro at P_3'. The ''back side'' (P_3' to P_8) acts principally to restrain the reactive site region in an inhibitory conformation whilst not directly interacting with the enzyme. All contacts with the proteinase are *via* the ''front side'' of the loop (P_3 to P_2'; Figure 1B). Of these interacting residues, the P_1' serine residue is well conserved, possibly due to its involvement in intramolecular hydrogen bonding within the inhibitor loop (Lin *et al*., 1993), and the P_3 cysteine residue is essential for disulfide formation.

The multiplicity of the combinatorial bead library is limited by the size of the individual resin beads and the amount of resin that it is feasible to screen. Here, we screened approximately 32,000 beads with a theoretical 8000 different sequences present, giving an average of four copies of each sequence present in the assay. This means that the screening regime should, in principle, locate multiple copies of any active components. Various methods have been suggested that allow sequence determination from a single bead. These include direct Edman sequencing of peptide libraries (Lam *et al*., 1991), use of mass spectroscopy (Brummel *et al*., 1994), and indirect tagging by oligonucleotides (Brenner & Lerner, 1992; Nielsen *et al*., 1993) or by chromatographic markers (Ohlmeyer *et al*., 1993). The sequences reported here were determined by direct Edman sequencing. Although this methodology is specific to peptide-based libraries, it is highly appropriate for the system used here and gives unambiguous results.

Sequences identified

All beads that were analyzed showed the presence of a single peptide sequence, and it is important to note that no hydrolysis product was detectable. Testing of peptide inhibitors of proteinases in solution is complicated by the inevitable presence of hydrolysis products (e.g. see Domingo *et al*., 1995). However, the TentaGel resin that we have used here is essentially impermeable to proteins and only the surface of the resin bead is accessible. This means that the overwhelming majority of the peptide, which is located internally within the bead matrix, remains unaffected by the proteinase, even after extended incubation.

The screening procedure resulted in selection of 13 positive beads. Sequencing data were obtained from each of these, and all the different sequences identified were individually synthesized as 11 residue disulfide-cyclized peptides. Competitive inhibition kinetics were measured on each peptide in solution, and K_i values were determined. It was found that ten of the beads (77%) had sequences that produced tight-binding chymotrypsin inhibitors, whereas three sequences did not show any competitive inhibition by the re-synthesized peptide. Non-specific binding is always a potential problem in library screening (Lam & Lebl, 1994), and so it is essential to authenticate the activity of derived sequences. All subsequent discussion relates to the sequences that were confirmed to have inhibitory activity after assays on re-resynthesized peptides.

Implications for inhibition mechanism

The results presented in Table 1 show that a clear consensus sequence exists for chymotrypsin inhibition. It is striking that we found multiple copies for all but one of the inhibitors. This gives us confidence in the selection procedure used, as we screened sufficient resin to include, on average, four copies of each sequence. The peptide inhibitor template used in this study is expected to inhibit proteinases in the same manner as the parent BBI protein. The results of this study can therefore, by extrapolation, shed light on the structural requirements for inhibition by this class of inhibitor protein. Of the three locations varied, the P_2 site was found to contain exclusively threonine; the P_1 site had phenylalanine and tyrosine; and leucine, isoleucine and norleucine were found at P_2'. The preference for aromatic residues at P_1 is consistent with the substrate preference of chymotrypsin (Schellenberger *et al*., 1991). These results are also in accord with previous work where the P_1 residue alone was varied (Domingo *et al*., 1995). The sequences obtained at the other locations were less predictable from the substrate specificity, and it is particularly striking that exclusively threonine was found at the P_2 location. In the crystal structure of mung bean inhibitor complexed to trypsin, the P_2 threonine residue is found to make internal hydrogen bonds within the inhibitor loop *via* the O^{γ}, and to form hydrophobic contacts *via* C^{β} and C^{γ} to His57 within the catalytic triad (Lin *et al*., 1993). It is likely that these interactions would be important with all serine proteinases, which

explains the conservation. Nevertheless, it is surprising that we did not find any structurally homologous amino acid such as serine, which implies that the structural requirements for inhibition are quite restrictive at this site. We are currently investigating this phenomenon further by constructing a series of systematic variations at this locus. The amino acid residues identified at P'_2 are clearly related (they are isomers of each other), and suggest a preference for large aliphatic substituents at this location. Table 1 records the inhibition constants that were determined after each of the peptides identified had been individually synthesized. All the sequences were found to be potent inhibitors of chymotrypsin, with K_i values of around 20 nM. To our knowledge, the K_i values found for these peptides represent the lowest so far recorded for sequences based on the BBI reactive site when inhibiting chymotrypsin.

There are two features of the combinatorial library described here that distinguish it from most previous reports. First, the presence of a disulfide bond to constrain the conformation. In the context of the peptides used here, the disulfide is essential for these sequences to act as proteinase inhibitors; reduction to a linear peptide destroys the inhibitory activity. The advantages of a constrained system in this particular instance are exemplified by comparing the results of the present investigation with those obtained using libraries of linear peptides to search for trypsin inhibitors (Eichler & Houghten, 1994). Those authors used a solution library of hexapeptides, corresponding to the reactive sites of known proteinaceous trypsin inhibitors, and a ''positional scanning'' (Dooley & Houghten, 1993) approach with lysine or arginine as the P_1 residue in fixed positions. Our constrained library, although less diverse, contained inhibitors at least three orders of magnitude more potent than those discovered for the linear peptide library. The second distinguishing feature of our system is that randomization was undertaken at a limited number of sites within a larger functional motif. This strategy has recently been applied to larger peptide domains (Bianchi *et al*., 1995), and is clearly an optimal way of creating interesting functional variants when using larger systems. Indeed, the BBI loop itself can be considered to represent a small polypeptide domain, as it is a distinct structural unit that retains functional activity when isolated from the rest of the protein.

An advantage of synthetic combinatorial libraries over those produced using genetic techniques is the possibility to include residues that are not genetically encoded. Our library was quite limited in this regard, with only a single unnatural amino acid (norleucine). However, it is interesting that this amino acid was found amongst the active peptides, and it is likely that future libraries will make much more use of unnatural residues and non-peptide components.

Implications for protein design

There has been much recent interest in the use of small protein domains as scaffolds onto which novel functionality can be constructed (Pessi *et al*., 1993; Bianchi *et al*., 1995; Vita *et al*., 1995). These molecules provide a means of generating conformationally defined structures, libraries of which are likely to have greater potential as pharmacophores than do simple linear sequences. The specialized scaffold used in our study is even smaller than those described above, yet still retains the constrained conformation required for functional activity.

The mechanism of action of proteinase inhibitors is well established, and changes in activity by single-site changes within the reactive site region either by natural mutation (Carrell *et al*., 1982), site-directed mutagenesis (Rosenberg *et al*., 1984; Courtney *et al*., 1985; Longstaff *et al*., 1990) or peptide synthesis (Terada *et al*., 1978; Favel *et al*., 1989; Domingo *et al*., 1995) have been described. However, it seems likely that efficient tailoring of specificity between a proteinase and an inhibitor sequence would benefit from a more extensive series of changes within the inhibitor canonical loop. It is striking, and perhaps quite fortunate from a design standpoint, that very many proteinase inhibitors act by presenting a section of peptide in this particular canonical structure. This means that judicious variation of these interacting residues gives a way of engineering inhibition specificity. The work reported here is a first step in the construction of a generalized library of constrained canonical loop structures from which specific inhibitors can be selected.

Materials and Methods

Materials

N-α-9-Fluorenylmethoxycarbonyl-protected amino acids and derivatives were purchased from Calbiochem-Novabiochem or Bachem with the following side-chain protecting groups: Ala, Arg(2,2,5,7,8-pentamethylchroman-6-sulfonyl), Asn(trityl), Asp(tBu), Cys(trityl), Gln(trityl), Glu(tBu), Gly, His(trityl), Ile, Leu, Lys(t-butoxycarbonyl), Met, Nle, Phe, Pro, Ser(tBu), Thr(tBu), Trp, Tyr(tBu), Val. Dimethylformamide (DMF) and *N*-methylpyrrolidone (NMP) were peptide synthesis grade from Rathburn Chemicals (Walkerburn, UK) and 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) and *N*-hydroxybenzotriazole (HOBT) were from SNPE (Croydon, UK). Bovine pancreatic α-chymotrypsin (treated with Nα-*p*-tosyl-L-lysine chloromethyl ketone), biotin *N*-hydoxysuccinimide ester, phenol, ethanedithiol, thioanisole, dimethylsulfoxide (DMSO), 5-bromo-4-chloro-3-indolyl phosphate (BCIP), nitro blue tetrazolium (NBT) and trifluoroacetic acid (TFA) were purchased from Sigma. Soybean trypsin inhibitor (STI; Sigma) was coupled to CNBr-activated Sepharose 4B (Pharmacia-LKB) as described by the manufacturer.

Biotin labeling of chymotrypsin

α-Chymotrypsin (37 mg) was dissolved in 350 μl of 50 mM $NaHCO_3$ (pH 9). To this was added 11 mg of biotin *N*-hydroxysuccinimide ester (Sigma) in 70 μl of DMF. Labeling was performed while maintaining gentle shaking for one hour. The reaction was quenched by the addition of an equal volume of 1 M ethanolamine, 1 M NaCl. Chymotrypsin was then purified by affinity chromatography using STI-Sepharose equilibrated in 50 mM Tris-HCl (pH 8), 500 mM NaCl. Active chymotrypsin was eluted with 10 mM HCl and concentrated by centrifugal ultrafiltration using a Centricon (Amicon), with a 10 kDa exclusion limit. The degree of biotinylation was determined using 4-hydroxyazobenzene-2'-carboxylic acid (Sigma) by the method of Green (1970), and the concentration of protein was found using the BCA protein assay (Pierce) with unlabeled α-chymotrypsin as standard, to derive a biotin to protein ratio. Labeled enzyme was stored at –70°C.

Peptide synthesis and purification

Free peptides for kinetic analysis were synthesized by standard solid-phase techniques using Fmoc chemistry (Atherton & Sheppard, 1989) and HBTU / HOBT activation (Knorr *et al*., 1989) as described (Domingo *et al*., 1995) using either an Applied Biosystems (ABI) 431A peptide synthesizer at 0.1 mmol scale with fourfold excess of amino acids or a Shimadzu PSSM-8 peptide synthesizer at 0.025 mmol scale with tenfold excess of amino acids.

Cleavage of the peptides from the resin and deprotection of the side-chain protecting groups was performed using a mixture of phenol (0.75 g), ethanedithiol (0.25 ml), thioanisole (0.5 ml), water (0.5 ml) and TFA (10 ml) with stirring for two hours. Peptide was separated from resin on a sintered glass funnel, precipitated in 50 ml of ice-cold tert-methyl butyl ether (Aldrich) and collected by centrifugation at 2000 ***g***. The precipitation was repeated five times, and the peptide was dried under vacuum over silica gel.

Disulfide bridge formation was performed using DMSO oxidation (Tam *et al*., 1991). The dried peptides were dissolved in 5% (v / v) acetic acid and 20% (v / v) DMSO to give a 0.05 mg ml^{-1} solution. The pH was adjusted to 6 with aqueous ammonia, and stirring was performed at room temperature for 24 hours. Peptide purification was by reverse-phase HPLC using a Gilson system with a Waters C_{18} Radial Pak column (25 mm × 10 cm) using mobile phases of water and acetonitrile, each containing 0.1% (v / v) TFA. All peptides were characterized by FAB-MS and analytical reverse-phase HPLC (Waters C_{18} Radial Pak column 10 mm × 5 mm).

BBI library synthesis

The BBI library (Figure 2) was synthesized by incorporating several ''split synthesis'' steps (Furka *et al*., 1991; Lam *et al*., 1991). Poly(ethylene glycol)-grafted polystyrene resin, TentaGel-S-NH_2, was used as the solid-phase support, having a diameter of 80 μm and a substitution of 0.23 meq g^{-1} (Rapp Polymere, Tubingen, Germany). Five glycine residues were first attached as a spacer. Next, the C-terminal portion of the 11 residue peptide, up to and including the proline P'_3 residue was synthesized on the ABI 431A synthesizer as described earlier, but using the 0.25 mmol scale. The N terminus was deprotected before equal separation into 20 capped polypropylene tubes in order to carry out the split synthesis. These contained individually 0.125 mmol in DMF of all the natural L-amino acids (except for cysteine) and L-norleucine as their pentafluorophenyl (OPfp) esters or 3-hydroxy-4-oxo-3,4-dihydro-1,2,3-benzotriazine(Od-hbt) (serine and threonine) esters. One equivalent of HOBT to amino acid was included as catalyst and the final volume was 2.5 ml in DMF. Coupling was allowed to proceed for two hours with shaking. Reagent was removed by filtration and the resin was washed with NMP. The individual tubes were then re-pooled into the reaction vessel of the ABI 431A synthesizer and N-terminal deprotection conducted on the synthesizer as described above. The next residue, serine at the P'_1 position, was added on the synthesizer. The split synthesis procedure was then repeated for the P_1 and P_2 residues. The remaining two residues were added, and final Fmoc deprotection performed, on the synthesizer. The resin-bound library was washed with methanol and dried. Deprotection was performed as above using 15 ml of the cleavage mixture for 3.5 hours at room temperature with gentle shaking. The library was washed on a sintered glass filter with TFA followed by methanol and again dried overnight. Cyclization was performed as described previously using a one liter solution of 5% acetic acid, 20% DMSO pH 6 with gentle shaking for 48 hours. Control experiments established that these conditions allow intramolecular disulfide formation to occur in almost quantitative yield (McBride *et al*., 1996). The resin was then finally filtered, thoroughly washed with water and methanol, dried and stored at 4°C.

Screening the BBI library for chymotrypsin inhibitors

The library was washed with PBS-T, and sufficient was resuspended in this buffer to give approximately 32,000 beads in a volume of 6.5 ml. Biotin chymotrypsin was added to give a final concentration of 0.36 μM, and the suspension incubated for 15 minutes at room temperature on a shaker. The beads were then thoroughly washed with PBS-T and collected on a Millipore RA 1.2 μm filter. They were then resuspended in 10 ml of PBS-T containing 25 μg ml^{-1} Extravidin-alkaline phosphatase (Sigma Chemical Co.; Extravidin is a chemically modified form of Avidin) in PBS-T for one hour at room temperature with gentle shaking. Again the beads were thoroughly washed with PBS-T followed by 100 mM Tris-HCl (pH 9.5), 0.5 mM $MgCl_2$. The beads were re-suspended in 10 ml of the same buffer with 35 mM BCIP, 37 mM NBT as alkaline phosphatase substrate for approximately five minutes. The reaction was then quenched by immediate re-filtration and washing with 25 mM EDTA (pH 7.4). Positively stained beads were detected by visualization under a low-power microscope and transferred individually to glass-fiber sequencing filters. The filters with their single beads were washed with 50 mM HCl, air-dried and then the bead sequenced by standard Edman degradation on an ABI 467A protein sequencer.

Kinetic measurements

Enzymatic hydrolysis of *N*-succinyl-Ala-Ala-Pro-Phe *p*-nitroanilide (Sigma) was conducted at 25°C in 144 mM

Tris-HCl (pH 7.8). Absorbance increase at 410 nm was measured upon addition of chymotrypsin. The substrate concentration was determined from the final absorbance ($\epsilon = 8800$, Erlanger *et al*., 1961). For K_m and k_{cat} determination of labeled and unlabeled chymotrypsin, a range of substrate concentrations were examined. The initial slopes were converted into molar rates and the kinetic constants obtained by non-linear regression (Leatherbarrow, 1990) using the GraFit computer program (Leatherbarrow, 1992).

Inhibition kinetics

The inhibition kinetics were measured for various active peptides identified from the biotin-chymotrypsin screening. These were determined by competitive binding studies using the chromogenic substrate succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide. Assays were performed at pH 7.8 in 144 mM Tris-HCl. The active chymotrypsin concentration was measured by active site titration with *p*-nitrophenyl acetate (Bender *et al*., 1966) and peptide concentrations by the absorbance at 280 nm (Edelhoch, 1967). Substrate hydrolysis was monitored at 410 nm and substrate concentration determined by the final absorbance at 410 nm. Initial velocity data were fit using the GraFit software package to find K_i. Values of K_i were corrected to account for competition by substrate ($K_i = K_{i(observed)} / (1 + [S] / K_m)$).

Acknowledgements

We thank John Barton from our department for performing the mass spectrometry on the peptides, and Dr Phil Jackson of Perkin Elmer/ABI, Warrington, UK, for advice on peptide sequencing. This work was supported by the BBSRC and Glaxo Group Research.

References

Atherton, E. & Sheppard, R. (1989). *Solid Phase Peptide Synthesis, A Practical Approach*. IRL Press, Oxford.

Bender, M. L., Begué-Canton, M. L., Blakley, R. L., Brobacher, L. J., Feder, J., Gunter, C. R., Kézdy, F. J., Killheffer, J. V., Jr, Marshall, T. H., Miller, C. G., Roeske, R. W. & Stoops, J. K. (1966). The determination of the concentrations of hydrolytic enzyme solutions: α-chymotrypsin, trypsin, papain, elastase, subtilisin and acetylcholinesterase. *J. Am. Chem. Soc.* **88**, 5890–5913.

Bianchi, E., Folgori, A., Wallace, A., Nicotra, M., Acali, S., Phalipon, A., Barbato, G., Bazzo, R., Cortese, R., Felici, F. & Pessi, A. (1995). A conformationally homogenous combinatorial peptide library. *J. Mol. Biol.* **247**, 154–160.

Bode, W. & Huber, R. (1992). Natural protein proteinase inhibitors and their interaction with proteinases. *Eur. J. Biochem.* **204**, 433–451.

Brenner, S. & Lerner, R. A. (1992). Encoded combinatorial chemistry. *Proc. Natl Acad. Sci. USA*, **89**, 5381–5383.

Brummel, C. L., Lee, I. N. W., Zho, Y., Benkovic, S. J. & Winograd, N. (1994). A mass spectrometric solution to address problem of combinatorial libraries. *Science*, **264**, 399–402.

Carrell, R. W., Jeppsson, J. O., Laurell, C.-B., Brennan, S. O., Owen, M. C., Vaughan, L. & Boswell, D. R. (1982). Structure and variation of human α_1-antitrypsin. *Nature*, **298**, 329–334.

Chen, P., Rose, J., Loves, R., Wei, C. H. & Wang, B. (1992). Reactive sites of an anticarcinogenic Bowman-Birk proteinase inhibitor are similar to other trypsin inhibitors. *J. Biol. Chem.* **267**, 1990–1994.

Courtney, M., Jallat, S., Tessier, L.-H., Benavente, A., Crystal, R. G. & Lecocq, J.-P. (1985). Synthesis in E. coli of α_1-antitrypsin variants of therapeutic potential for emphysema and thrombosis. *Nature*, **313**, 149–151.

Cwirla, S. E., Peters, E. A., Barrett, R. W. & Dower, W. J. (1990). Peptides on phage: a vast library of peptides for identifying ligands. *Proc. Natl Acad. Sci. USA*, **87**, 6378–6382.

Devlin, J. J., Panganiban, L. C. & Devlin, P. E. (1990). Random peptide libraries: a source of specific protein binding molecules. *Science*, **249**, 404–406.

Domingo, G. J., Leatherbarrow, R. J., Freeman, N., Patel, S. & Weir, M. (1995). Synthesis of a mixture of cyclic peptides based on the Bowman-Birk reactive site loop to screen for serine protease inhibitors. *Int. J. Pept. Protein Res.* **46**, 79–87.

Dooley, C. T. & Houghten, R. A. (1993). The use of positional scanning synthetic peptide combinatorial libraries for the rapid identification of opioid receptor ligands. *Life Sci.* **52**, 1509–1517.

Edelhoch, H. (1967). Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry*, **6**, 1948–1954.

Eichler, J. & Houghten, R. A. (1993). Identification of substrate-analog trypsin inhibitors through the screening of synthetic peptide combinatorial libraries. *Biochemistry*, **32**, 11035–11041.

Eichler, J., Lucka, A. W. & Houghten, R. A. (1994). Cyclic peptide template combinatorial libraries: synthesis and identification of chymotrypsin inhibitors. *Pept. Res.* **7**, 300–307.

Erlanger, B., Kokowsky, N. & Cohen, W. (1961). The preparation and properties of two new chromogenic substrates of trypsin. *Arch. Biochem. Biophy.* **95**, 271–278.

Favel, A., L-Nguen, D., Coletti-Previero, M. A. & Castro, B. (1989). Active site chemical mutagenesis of *Ecballium elaterium* trypsin inhibitor II: new microproteins inhibiting elastase and chymotrypsin. *Biochem. Biophys. Res. Commun.* **162**, 79–82.

Fodor, S. P. A., Read, J. L., Pirrung, L., Stryer, L., Lu, A. T. & Solas, D. (1991). Light-directed, spatially addressable parallel chemical synthesis. *Science*, **251**, 767–773.

Furka, A., Sebestyen, F., Asgedom, M. & Dibo, G. (1991). General method for rapid synthesis of multicomponent peptide mixtures. *Int. J. Pept. Protein Res.* **37**, 487–493.

Gallop, M. A., Barrett, R. W., Dower, W. J., Fodor, S. P. A. & Gordon, E. M. (1994). Applications of combinatorial technologies to drug discovery. 1. Background and peptide combinatorial libraries. *J. Med. Chem.* **37**, 1233–1251.

Geysen, H. M., Rodda, S. J. & Mason, T. J. (1986). A priori delineation of a peptide which mimics a discontinuous antigenic determinant. *Mol. Immunol.* **23**, 709–715.

Gordon, E. M., Barrett, R. W., Dower, W. J., Fodor, S. P. A. & Gallop, M. A. (1994). Applications of combinatorial technologies to drug discovery. 2. Combinatorial organic synthesis, library screening strategies, and future directions. *J. Med. Chem.* **37**, 1385–1401.

Green, N. M. (1970). Spectrophotometric determination of avidin and biotin. *Methods Enzymol.* **18** (Part A), 418–424.

Hoogenboom, H. R., Griffiths, A. D., Johnson, K. S., Chiswell, D. J., Hudson, P. & Winter, G. (1991). Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody variable domains. *Nucl. Acids Res.* **19**, 4133–4137.

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

Ikenaka, T. & Norioka, S. (1986). Bowman-Birk family serine proteinase inhibitors. In *Proteinase Inhibitors* (Barrett, A. J. & Salvesen, G. S., eds), pp. 23–51, Elsevier Science Publishers, Amsterdam.

Kay, B. K. (1994). Biologically displayed random peptides as reagents in mapping protein-protein interactions. *Perspect. Drug Divers. Design*, **2**, 251–268.

Knorr, R., Trzeciak, A., Bornwarth, W. & Gillessen, D. (1989). New coupling reagents in peptide chemistry. *Tetrahedr. Letters*, **30**, 1927–1930.

Lam, K. S. & Lebl, M. (1994). Selectide technology: bead-binding screening. *Methods*, **6**, 372–380.

Lam, K. S., Salmon, S. E., Hersh, E. M., Hruby, V. J., Kazmierski, W. M. & Knapp, J. J. (1991). A new type of synthetic library for identifying ligand-binding activity. *Nature*, **354**, 82–84.

Laskowski, M., Jr & Kato, I. (1980). Protein inhibitors of proteinases. *Annu. Rev. Biochem.* **49**, 593–626.

Leatherbarrow, R. J. (1990). Using linear and non-linear regression to fit biochemical data. *Trends Biochem. Sci.* **15**, 455–458.

Leatherbarrow, R. J. (1992). *GraFit (Version 3.0)*, Erithacus Software Ltd., Staines, UK.

Lin, G., Bode, W., Huber, R., Chi, C. & Engh, R. A. (1993). The 0.25-nm X-ray structure of the Bowman-Birk-type inhibitor from mung bean in ternary complex with porcine trypsin. *Eur. J. Biochem.* **212**, 549–555.

Longstaff, C., Campbell, A. F. & Fersht, A. R. (1990). Recombinant chymotrypsin inhibitor 2: expression, kinetic analysis of inhibition with α-chymotrypsin and wild-type and mutant subtilisin BPN', and protein engineering to investigate inhibitory specificity and mechanism. *Biochemistry*, **29**, 7339–7347.

Maeder, D. L., Sunde, M. & Botes, D. P. (1992). Design and inhibitory properties of synthetic Bowman-Birk loops. *Int. J. Pept. Protein Res.* **40**, 97–102.

McBride, J. D., Harbeck, B. & Leatherbarrow, R. J. (1996). Resin-coupled cyclic peptides as proteinase inhibitors. *Protein Pept. Letters*, in the press.

McCafferty, J., Griffiths, A. D., Winter, G. & Chiswell, D. J. (1990). Phage antibodies: filamentous phage displaying antibody variable domains. *Nature*, **348**, 552–554.

Meldal, M. & Svendsen, I. (1995). Direct visualisation of enzyme inhibitors using a portion mixing inhibitor library containing a quenched fluorogenic peptide substrate. Part 1. Inhibitors for subtilisin Carlsberg. *J. Chem. Soc. Perkin. Trans.* **1**, 1591–1596.

Neurath, H (1989). Proteolytic processing and physiological regulation. *Trends Biochem. Sci.* **14**, 268–271.

Nielsen, J., Brenner, S. & Janda, K. D. (1993). Synthetic methods for the implementation of encoded combinatorial chemistry. *J. Am. Chem. Soc.* **115**, 9812–9813.

Nishino, N., Aoyagi, H., Kato, T. & Izumiya, N. (1977). Studies on the synthesis of proteinase inhibitors. *J. Biochem. (Tokyo)*, **82**, 901–909.

Ohlmeyer, M. H. J., Swanson, R. N., Dillard, L. W., Reader, J. C., Asouline, G., Kobayashi, R., Wigler, M. & Still, W. C. (1993). Complex synthetic chemical libraries indexed with molecular tags. *Proc. Natl Acad. Sci. USA*, **90**, 10922–10926.

Owens, R. A., Gesellchen, B. J., Houchins, B. J. & DiMarchi, R. D. (1991). The rapid identification of HIV protease inhibitors through the synthesis and screening of defined peptide mixtures. *Biochem. Biophys. Res. Commun.* **181**, 402–408.

Pessi, A., Bianchi, E., Crameri, A., Venturini, S., Tramontano, A. & Solazzo, M. (1993). A designed metal-binding protein with a novel fold. *Nature*, **362**, 367–369.

Read, R. & James, M. N. G. (1986). Introduction to the proteinase inhibitors: X-ray crystallography. In *Proteinase Inhibitors* (Barrett, A. J. & Salvesen, G. S., eds), pp. 301–336, Elsevier Science Publishers, Amsterdam.

Roberts, B. L., Markland, W., Ley, A. C., Kent, R. B., White, D. W., Gutterman, S. K. & Ladner, R. C. (1992). Directed evolution of a protein: selection of potent neutrophil elastase inhibitors displayed on M13 fusion phage. *Proc. Natl Acad. Sci. USA*, **89**, 2429–2433.

Rosenberg, S., Barr, P. J., Najarian, R. C. & Hallewell, R. A. (1984). Synthesis in yeast of a functional oxidation-resistant mutant of human α_1-antitrypsin. *Nature*, **312**, 77–80.

Salmon, S. E., Lam, K. S., Lebl, M., Kandola, A., Khattri, P. S., Wade, S., Pátek, M., Kocis, P., Krchnák, V., Thorpe, D. & Felder, S. (1993). Discovery of biologically active peptides in random libraries: solution-phase testing after staged orthogonal release from resin beads. *Proc. Natl Acad. Sci. USA*, **90**, 11808–11712.

Schellenberger, V., Braune, K., Hofmann, H. J. & Jakubke, H. D. (1991). The specificity of chymotrypsin. *Eur. J. Biochem.* **199**, 623-636.

Scott, J. K. & Smith, G. P. (1990). Searching for peptide ligands with an epitope library. *Science*, **249**, 386–390.

Tam, J. P., Wu, C. R., Liu, W. & Zhang, J. W. (1991). Disulfide bond formation in peptides by dimethyl sulfoxide. Scope and applications. *J. Am. Chem. Soc.* **113**, 6657–6662.

Terada, S., Sato, K., Kato, T. & Izumiza, N. (1978). Inhibitory properties of nonapeptide loop structures related to reactive sites of soybean Bowman-Birk inhibitor. *FEBS Letters*, **90**, 89–92.

Tsunogae, Y., Tanaka, I., Yamane, T., Kikkawa, J., Ashida, T., Ishikawa, C. & Takashahi, K. (1986). Structure of the trypsin-binding domain of Bowman-Birk type protease inhibitor and its interaction with trypsin. *J. Biochem (Tokyo)*, **100**, 1637–1646.

Vita, C., Roumestand, C., Toma, F. & Ménez, A. (1995). Scorpion toxins as natural scaffolds for protein engineering. *Proc. Natl Acad. Sci. USA*, **92**, 6404–6408.

Werner, M. H. & Wemmer, D. E. (1992). Three-dimensional structure of soybean trypsin/chymotrypsin Bowman-Birk inhibitor in solution. *Biochemistry*, **31**, 999–1010.

Edited by A. R. Fersht

(Received 6 March 1996; accepted 26 March 1996)