DNA-Binding Ligands from Peptide Libraries Containing Unnatural Amino Acids

Theo Lescrinier, Chris Hendrix, Luc Kerremans, Jef Rozenski, Andreas Link, Bart Samyn, Arthur Van Aerschot, Eveline Lescrinier, Ramon Eritja, Jozef Van Beeumen, and Piet Herdewijn*

Abstract: An unnatural peptide-based library, bound on a solid support, was screened for double-stranded-DNA (dsDNA)-binding ligands. For this purpose, fluorescein and rhodamine were used to label the single-stranded oligodeoxynucleotides. Beads containing products with affinity to dsDNA turned red in visible light and fluoresced yellow in UV light. A similar technique can be used for the selection of ligands which bind to a hairpin RNA, using a monolabelled oligoribonucleotide. The screening process revealed a high structure–affinity relationship in the successful products. Only six out of the twelve unnatural amino acids were selected, with the repeated appearance of AlaU, Sar and the secondary amino acids (Hyp, Inp). The affinity and selectivity for the target was determined using a DNase I protection assay.

Keywords: combinatorial chemistry · DNA recognition · peptides · structure–activity relationships

Introduction

Creating compounds that recognise specific DNA sequences is a central goal in the development of DNA-targeted drugs. These compounds may be able to manipulate the transcription of individual genes and, hence, are potentially useful for the treatment of cancers and genetic diseases. The most successful approaches so far are the pyrrole–imidazole polyamide motif, developed by P. Dervan,[1] and the peptide nucleic acid, described by P. Nielsen.[2] As an alternative to the stepwise approach of compound selection by iterative synthesis, we propose the generation of dsDNA recognition libraries.

[*] Prof. Dr. P. Herdewijn, Dr. T. Lescrinier, Dr. C. Hendrix, L. Kerremans, Dr. J. Rozenski, Dr. A. Link, Dr. A. Van Aerschot, E. Lescrinier
Laboratory of Medicinal Chemistry
Rega Institute for Medical Research
Katholieke Universiteit Leuven, Minderbroedersstraat 10, B-3000 Leuven (Belgium)
Fax: (+ 32) 16-337-387
E-mail: piet.herdewijn@rega.kuleuven.ac.be

B. Samyn, Prof. Dr. Van Beeumen
Laboratory of Protein Biochemistry and Protein Engineering
Universiteit Gent, K. L. Ledeganckstraat 35, B-9000 Gent (Belgium)

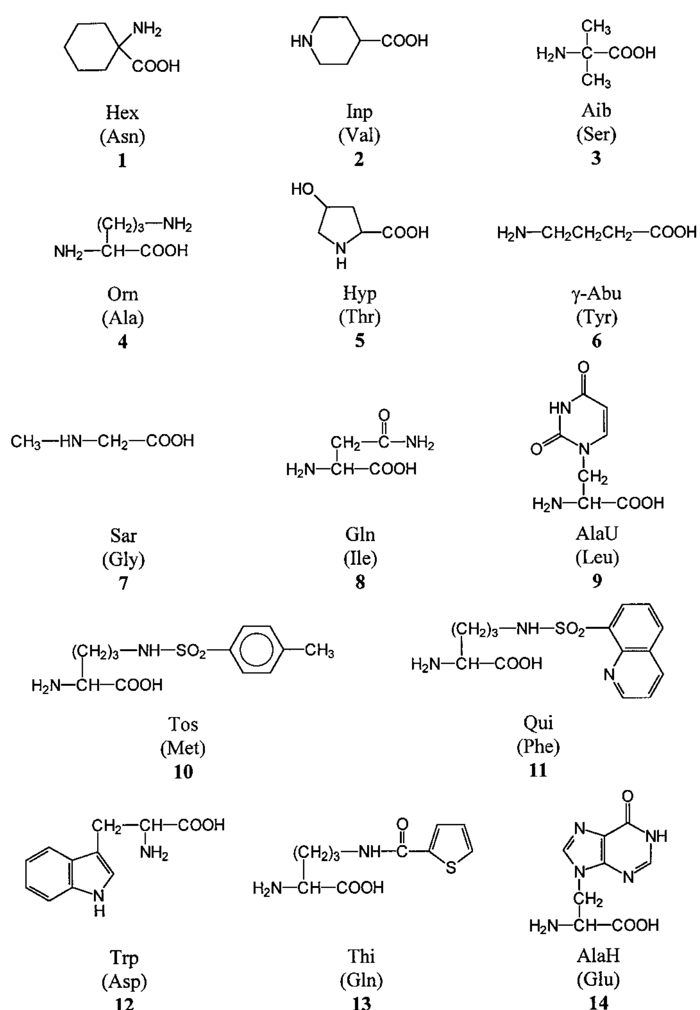
Dr. R. Eritja
European Molecular Biology Laboratory
Meyerhofstrasse 1, D-69012 Heidelberg (Germany)

Results and Discussion

The specificity of DNA–protein interactions served as model to select the monomers which were used to synthesise the libraries (Scheme 1). Synthetic amino acids were selected that are able to form complementary hydrogen-bonding donor–acceptor interactions with the nucleobases in the grooves of the double helix. As the stabilisation energy of hydrogen-bonding interaction is not very large, synthetic amino acids were introduced in the library able to strengthen the binding to dsDNA by stacking, intercalation and electrostatic interactions. Glutamine was incorporated in the library as it is known to be employed fairly often in protein–DNA recognition by bidentate hydrogen bonding of its side chain to an adenine base in the major groove. Several amino acids with a hydrophobic side chain were included to allow hydrophobic and Van der Waals contacts with the C-6 methyl group of thymine and the 2′-deoxy position of the furanoses. To allow optimal interaction it may be expected that upon complex formation one or both partners have to distort within the limit of allowed conformational changes. The deformability of the DNA double helix is sequence-dependent and this can contribute to the specificity of the recognition process. Amino acids were selected with appropriate conformational rigidity to allow both optimal fit and a minimisation of unfavourable entropy changes during complex formation.

The selection of an amide library is based on already existing examples[1, 2] of dsDNA recognition structure, the

 © WILEY-VCH Verlag GmbH, D-69451 Weinheim, 1998 0947-6539/98/0403-0425 $ 17.50+.25/0



Scheme 1. Building blocks used for the generation of the first library. Encoding amino acids are given in parentheses.

presence of this backbone in natural proteins, the straightforward synthetic chemistry and the possible involvement of main-chain NH groups in recognition by dsDNA. The selection of the limited length of the sequence of recognition elements (see below) is based on synthetic considerations combined with concerns about cellular uptake. Moreover, in protein–DNA recognition only a limited number of amino acid residues are important in the direct readout of the nucleotide sequence, and DNA-binding small molecules such as antitumor agents likewise bind preferentially to certain sequences. As the library contains 14 amino acids (Scheme 1), yielding a theoretical mixture of 537824 compounds, we expect our library not only to contain a universe of structures with different electronic and functional group distributions but, likewise, a universe of structural motifs.

Before the construction of the library on a solid phase was started, coupling efficiency as well as the occurrence of possible side reactions was evaluated by the synthesis of more test sequences. These were selected based on expected worst-case scenarios (Table 1). Monophenylglycine (Mfg) was originally a library member, but was removed due to its high susceptibility to racemisation.[3] The library was synthesised by mix-and-split synthesis[4] giving 14^5 different individual compounds and approximately 200–300 pmol of compound per bead. The chemical structure of the peptides was determined with encoding sequences[5] of natural amino acids and Edman degradations. The encoding acid-labile Ddz [2-(3,5-dimethoxyphenyl)-propyl-2-oxycarbonyl] amino acids were introduced before the unnatural Fmoc (9-fluorenylmethyloxycarbonyl) amino acids. The coding system is represented in Scheme 1.

Table 1. Selected oligopeptides used to investigate coupling efficiency between the modified amino acids.

	Sequence of oligopeptides	LSIMS $[M+H]^+$
1	HOOC-β-Ala-Sar-Gln-Hex-Inp-Hex-NH_2	650
2	HOOC-Sar-Hyp-Inp-AlaU-Trp-NH_2	886
3	HOOC-AlaH-Inp-Aib-Mfg-NH_2	553
4	HOOC-Aib-Aib-Inp-Inp-NH_2	411
5	HOOC-β-Ala-Tos-Inp-Thi-Qui-β-Ala-NH_2	1069

The library thus obtained was utilised in a screening procedure using a 14 base-pair dsDNA sequence (5′-ACATTGCACAATCT-3′ with its complement) as target. This represents the binding site, within the enhancer region of the human IL-6 gene, for NF-IL-6.[6] To ensure targeting of the dsDNA, one single strand was labelled with fluorescein and the other with rhodamine. In UV light, beads containing rhodamine-labelled oligos are coloured red, while fluorescein-labelled oligos turn the beads green. A yellow bead must therefore be labelled with both fluorescein and rhodamine and, thus, contains double-stranded DNA (Figure 1A). In visible light, all rhodamine-labelled oligo beads are coloured red. This implies that beads containing products with an affinity to dsDNA can be easily isolated using double-labelled DNA with a screening procedure using UV as well as visible light. A strategy using two different dyes to address selectivity problems has been described before by Boyce et al., although for different purposes.[7] The T_m of the unlabelled double-stranded DNA (4 μM) is 55 °C in 0.1 M NaCl (phosphate buffer pH 7.4), while the T_m of the 5′-labelled duplex under the same conditions is 54 °C. This ensures formation of double-stranded structures at room temperature and the absence of major influences of the dyes on the hybridising properties.

The screening procedure itself was carried out with the labelled dsDNA in a solution of 0.1 M NaCl in phosphate buffer (pH 7.4) for 4 h. Beads were collected, washed six times with buffered 0.1 M NaCl and visualised under a fluorescent

Abstract in Dutch: *Door het merken van enkelstrengige oligodeoxynucleotiden met fluoresceine en rhodamine werd een methode ontwikkeld voor de identificatie van liganden met bindingsaffiniteit voor dsDNA. De screening van een bank op basis van onnatuurlijke peptiden resulteerde in een hoge structuur–affiniteitsrelatie. Deze methode kan nuttig zijn voor de identificatie van sequentiespecifieke dsDNA-bindende producten.*

 © WILEY-VCH Verlag GmbH, D-69451 Weinheim, 1998 0947-6539/98/0403-0426 $ 17.50+.50/0

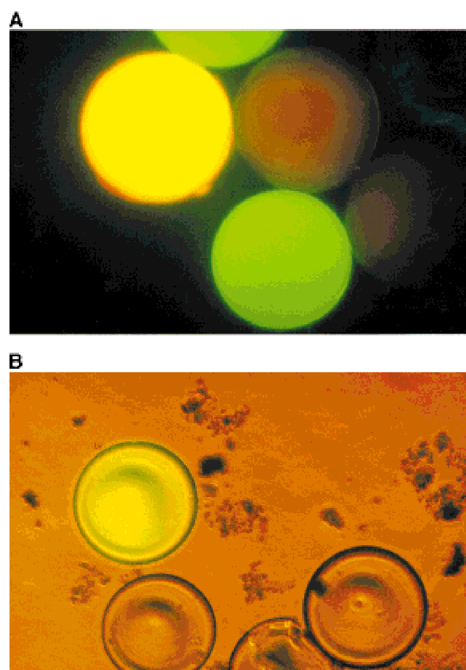


Figure 1. A) Differentiation between labelling with rhodamine, fluorescein and both dyes. B) Contrast between a bead bearing a peptide with affinity and beads without any interaction with the TAR-loop (TAR = *trans*-activation response element) target in UV and visible light at low intensity. The hairpin RNA was monolabelled with fluorescein and protected against 3′-exonuclease degradation with propanediol.

microscope. The yellow fluorescent beads were isolated under UV light and evaluated in a second selection cycle in visible light allowing isolation of the intensely red beads (Figure 2). In the next round, the hydroxybutyrate linker[8] connecting the noncoding sequence was hydrolysed using 0.1N NaOH/ CH_3CN and the screening procedure using double-labelled dsDNA was repeated. The absence of both fluorescence and colour emphasises the importance of the nonnatural peptide sequence for interaction with the target.

To test the versatility of the approach, we likewise screened the second library for peptides interacting with TAR-RNA. As interaction between Tat protein (Tat = *trans*-acting transcriptional activation) and TAR-RNA is critical for virus replication, molecules which could prevent or reduce the affinity of Tat for TAR might be useful as antiviral agents. Due to the high lability of the RNA target, supplementary precautions for screening were taken. The RNA hairpin[9] target was monolabelled at the 5′-end with fluorescein and protected at the 3′-end against 3′-exonuclease degradation using propanediol.[10] The screening buffer consisted of 0.1M NaCl, 20 mM KH_2PO_4 and 0.1 mM EDTA (pH 7.4), and was treated with 7 OD units of labelled target in 3 mL of screening buffer for 16 h. The beads were washed and were divided over 4 petri dishes and evaluated under UV light. In the presence of visible light at low intensity the difference between binders and nonbinders could be detected (Figure 1B).

Decoding of the peptides, selected during dsDNA affinity screening, was based on Edman degradation and resulted in the identification of a series of 16 peptides from the initial screening and a series of 9 peptides resulting from a second screening round at higher oligo concentrations and longer exposure time. They may be classified as high- and low-affinity sequences based on the screening circumstances and the intensity of the fluorescence (colour) of the beads. Unfortunately, almost all isolated compounds (list not shown) proved to be polyornithines (three or four ornithines in a 6-mer), reflecting the importance of ion–ion interactions between the peptides and the polyphosphate backbone of the dsDNA for affinity binding. As it was not our aim to select new polycationic DNA-binding substances but more specific molecules, the assembly of a second library was undertaken.

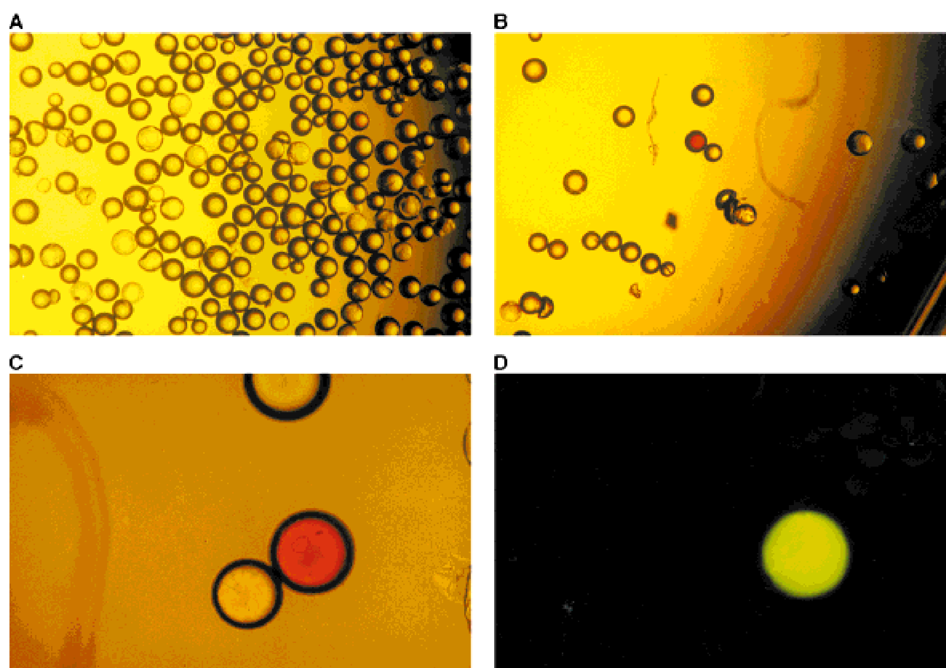


Figure 2. Selection process: panel A gives an overview of part of the library; part B gives a view of part of the library containing one positive bead in the middle, which is enlarged in panel C (visible light) and panel D (UV light).

Four important points distinguish the second library from the first one: 1) The positively charged ornithine was excluded from the randomisation steps and only one positively charged amino acid (arginine) was incorporated at a fixed (end) position. 2) γ-Aminobutyric acid was removed, as it induces high flexibility in the peptide backbone, unfavourable for affinity binding, and as it was hardly present in the peptides isolated from the first library. 3) Upon analysis of the sequencing results of the first library, it was noticed that no negatively charged coding amino acid was detected. In parallel, the selected sequences did not contain the hypo-

 © WILEY-VCH Verlag GmbH, D-69451 Weinheim, 1998 0947-6539/98/0403-0427 $ 17.50+.25/0

xanthine derivative (AlaH) or tryptophan (Trp) as they were coded for by aspartic acid and glutamic acid, respectively. This was most probably due to electrostatic repulsion between the negatively charged polyphosphate backbone and the amino acid side chains of the coding sequence. Therefore both encoding amino acids were eliminated from the library. 4) To avoid bias, a higher randomisation was introduced by varying the encoding–coded amino acid pairs in each cycle. In the first library the encoding system consisted of the same coding pairs of natural/unnatural amino acids throughout the whole peptide. The amino acids used in the second library and the encoding system are given in Table 2. All coupling reactions were verified by spectrophotometrical Fmoc determination and averaged between 90% and 100%.

Table 2. The coding system applied for the second library.

	Cycle 1	Cycle 2	Cycle 3	Cycle 4	Cycle 5
Thi	Gln	Ser	Val	Thr	Val
Hex	Asn	Val	Thr	Gln	Tyr
Hyp	Thr	Met	Ala	Ile	Gly
Tos	Met	Gly	Leu	Ser	Ile
AlaH	Tyr	Phe	Ser	Leu	Met
AlaU	Leu	Tyr	Met	Asn	Ala
Trp	Ala	Gln	Gln	Ala	Gln
Sar	Gly	Thr	Phe	Tyr	Thr
Gln	Ile	Asn	Ile	Val	Phe
Aib	Ser	Leu	Asn	Met	Ser
Qui	Phe	Ala	Gly	Gly	Asn
Inp	Val	Ile	Tyr	Phe	Leu

The screening technology was identical to the one for the first library; 11 UV-positive beads were selected (Table 3). The importance of the unnatural part of the bifurcations in the binding process was confirmed the same way as described before. The bound structure was determined by Edman sequencing.

Table 3. Sequencing results of positive hits of the second library.

Peptide sequence
Ac-Arg-AlaU-Sar-AlaU-Sar-AlaU
Ac-Arg-AlaU-Sar-Hyp-AlaU-Inp
Ac-Arg-AlaU-Sar-Inp-Sar-Sar
Ac-Arg-AlaU-AlaU-Gln-AlaU-Sar
Ac-Arg-AlaU-Sar-AlaU-Hyp-Gln
Ac-Arg-AlaU-AlaU-Inp-AlaU-Sar
Ac-Arg-AlaU-Hyp-Inp-Thi-AlaH
Ac-Arg-AlaU-Sar-Sar-Sar-Sar
Ac-Arg-AlaU-Sar-Hyp-Thi-Hyp *or*
Ac-Arg-AlaU-Sar-Hyp-Thi-Tos
Ac-Arg-AlaU-AlaU-Inp-AlaU-Hyp
Ac-Arg-AlaU-AlaU-Sar-Sar-Sar

Given the fact that we started with a library of about 250000 compounds and 12 different amino acids, the results of the screening procedure revealed a high structure–affinity relationship. Next to the fixed arginine, the β-uracilyl alanine analogue (AlaU) was invariably found in the encoded sequence. At the third position the preference is limited to practically two amino acids: sarcosine (Sar, 55%) and AlaU (36%). Except for the fourth position (where isonipecotic acid (Inp) is the most common amino acid) the diversity is rather limited. Only six out of the twelve unnatural amino acids are used and there is the repetitive appearance of AlaU, Sar and secondary amino acids (Hyp, Inp). The most frequent occurring combinations are: Ac-Arg-AlaU-Sar-Inp-AlaU-Sar-β-Ala-Gly and Ac-Arg-AlaU-Sar-Inp-Sar-Sar-β-Ala-Gly.

All unnatural peptides as well as the bifurcated compounds containing both the natural and the unnatural peptide part were synthesised. The latter synthesis is necessary as the screening procedure does not exclude the participation of the coding sequence in the interaction with the dsDNA fragment in combination with the encoded sequence. The synthesis of the bifurcations might also be interesting since the two arms (coded and encoded part) might engirdle the dsDNA. This is already known for the natural leucine zipper for which a scissor grip model is proposed.[11] Five bifurcations were prepared, four peptides selected from the library and the most frequently occurring combination (Table 4).

Table 4. The structure of the bifurcations synthesised for solution screening.

$H_2NC(=O)$—Lys(R_1)—GABA—Gly—βAla—R_2

Bif1-5

	R_1	R_2	
Bif 1	-Leu-Thr-Met-Tyr-Ala-NH_2	-AlaU-Sar-AlaU-Sar-AlaU-Arg-NHAc	**OP 1**
Bif 2	-Gly-Thr-Tyr-Tyr-Ala-NH_2	-Sar-Sar-Inp-Sar-AlaU-Arg-NHAc	**OP 2**
Bif 3	-Gly-Tyr-Tyr-Asn-Ala-NH_2	-Sar-AlaU-Inp-AlaU-AlaU-Arg-NHAc	**OP 3**
Bif 4	-Thr-Ser-Ala-Tyr-Ala-NH_2	-Hyp-Thi-Hyp-Sar-AlaU-Arg-NHAc	**OP 4**
Bif 5	-Gly-Tyr-Tyr-Tyr-Ala-NH_2	-Sar-AlaU-Inp-Sar-AlaU-Arg-NHAc	**OP 5**

The affinity of the synthetic peptides for the dsDNA target, however, was too low to be detected by gel shift experiments. Therefore we used a more sensitive functional assay, DNase I activity, for evaluating binding of peptides to dsDNA in solution (which is also used in DNase I footprint titration experiments). It is expected that a DNA-bound peptide would protect the phosphodiester backbone of DNA from DNase I-catalysed hydrolysis, which is used here in a DNase I protection assay. For this purpose, one DNA strand was radiolabelled, hybridised with the complementary strand, mixed with an excess of the respective peptides and subjected to DNase I-catalysed degradation. The degraded samples were analysed by denaturing polyacrylamide gel electrophoresis and visualised by autoradiography. The percentage of the remaining full-length oligonucleotide after different incubation times was quantified by scanning laser densitometry. The unnatural octapeptide (**OP**), the bifurcated peptide (**Bif**) and random peptides (**RP**) were all evaluated.

All the bifurcations evaluated (1000 equiv each, Table 4, compounds **Bif 2–5**), were able to slow down the degradation rate of dsDNA by DNase I in a comparable manner. After 2 h of incubation in the presence of a bifurcation, an average amount of 34% of intact dsDNA was still present compared to 18% in the case of the control experiment. The unnatural octapeptides (**OP 2** to **OP 5**) corresponding to the encoded part of the above mentioned bifurcations, were able to inhibit the degradation more efficiently than the corresponding bifurcations. The autoradiographs of the control experiment (A) and the experiment containing 1000 equiv of **OP 2** (B) are depicted in Figure 3. The results obtained with the four

 © WILEY-VCH Verlag GmbH, D-69451 Weinheim, 1998

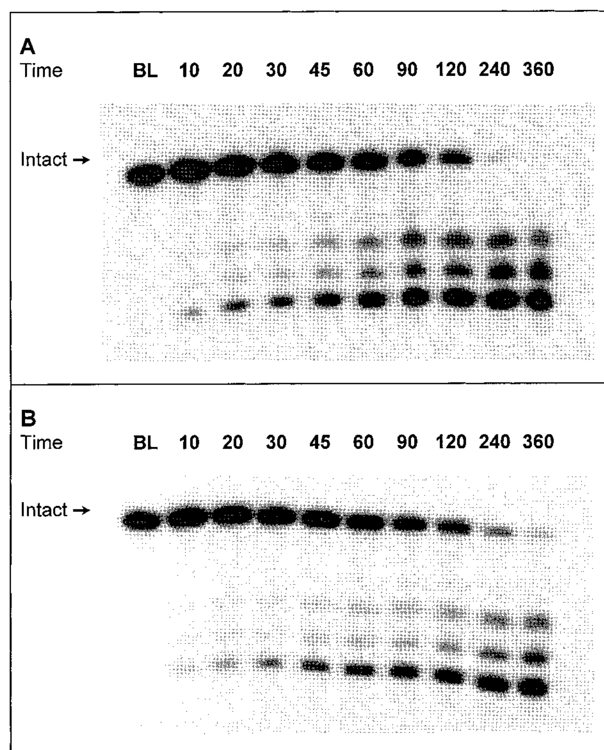


Figure 3. Autoradiographs of the degradation of the dsDNA target by DNase I. A) Control: no peptide added; B) 10^3 equiv of compound **OP2** added. Aliquots were taken at the time (in min) indicated on top. BL = blank, no enzyme added.

octapeptides are summarised in Figure 4. The presence of the coding strand seems to play an unfavourable role in the

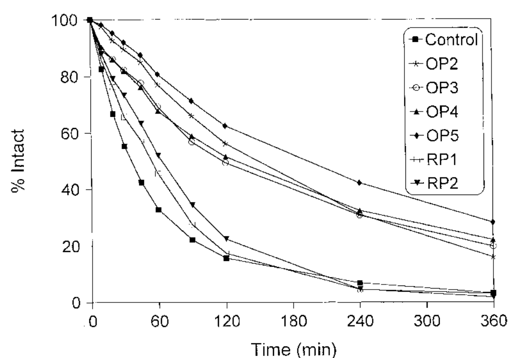


Figure 4. DNase I degradation of the dsDNA target in the presence of 1000 equiv of the peptides listed. The dsDNA target was incubated for the indicated times, and full-length (intact) and digested oligonucleotides were resolved on 20% denaturing polyacrylamide gels and quantified, as described in the Experimental Section.

binding of the peptides to dsDNA. Two random pentapeptides, one containing three amino acids with a positive charge in their side chain [Lys-Lys-Cha-Arg-Phe (**RP1**)] and one with a net charge of zero [Phe-Tyr-Glu-Lys-Pro (**RP2**)], were unable to bind to the dsDNA fragment and had degradation rates comparable to those in control experiments. The most efficient compound (Ac-Arg-AlaU-Sar-Inp-AlaU-Sar-β-Ala-Gly) corresponds to one of the two most frequently occurring combinations mentioned before. After 2 h of incubation with DNase I, 64% of the dsDNA was still intact, compared with 18% in the control experiment, and after 6 h more than 35% of intact dsDNA was still present while no intact dsDNA could be seen in the control experiment, with the bifurcations or with random peptides. In order to further evaluate the specificity of the compounds and to exclude binding of the peptide to DNase I, the degradation assay was repeated in the presence of 1000 equiv of **OP5** using $(dA)_{17} \cdot (T)_{17}$ as target. However, this duplex was degraded at the same rate as if no **OP5** were present (Figure 5). The unnatural oligopeptide

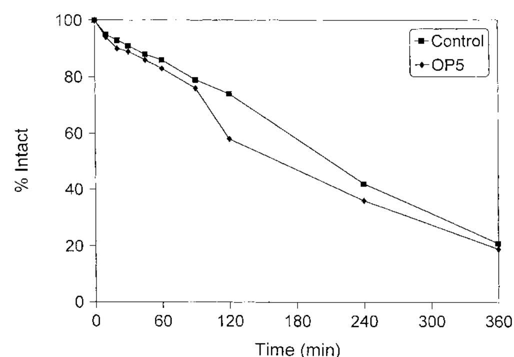


Figure 5. DNase I degradation of $dA_{17} \cdot T_{17}$ in the presence of 1000 equiv of **OP5**. $dA_{17} \cdot T_{17}$ was incubated for the indicated times, and full-length (intact) and digested oligonucleotides were resolved on 20% denaturing polyacrylamide gels and quantified as described in the Experimental Section.

shows selective protection of the sequence originally used during the screening process. This library approach for selecting dsDNA binding unnatural oligopeptides will now be followed by structure optimisation using solution-phase screening to increase affinity and selectivity for the target and to analyse its precise binding site. This technique may be useful for identifying new sequence-specific dsDNA- and RNA-binding ligands able to interfere with gene expression.

Experimental Section

Melting points were determined in capillary tubes with a Büchi–Totolli apparatus and are uncorrected. Ultraviolet spectra were recorded with a Philips PU 8740 UV/Vis spectrophotometer. The ^{13}C NMR spectra were recorded with a Varian Gemini 200. Exact mass measurements by liquid secondary ion mass spectrometry (LSIMS) were obtained using a Kratos Concept ^1H mass spectrometer. Precoated Machery–Nagel Alugram® Sil G/UV254 plates were used for TLC and the spots were examined with UV light, sulfuric acid/anisaldehyde spray or ninhydrin spray. Column chromatography was performed on Acros silica gel (0.06–200 nm). Anhydrous solvents were obtained as follows: THF was refluxed on $LiAlH_4$ overnight and was distilled; dichloromethane (DCM) was stored on calcium hydride, refluxed and distilled. Pyridine, triethylamine (TEA) and *N,N*-diisopropylethylamine (DIEA) were refluxed overnight on potassium hydroxide and distilled. *N,N*-dimethylformamide (DMF) was stored on activated molecular sieves for 3 days and was always tested for the absence of dimethylamine by the bromophenol test prior to use. CH_3CN for HPLC was purchased from Rathburn (grade S) and water for HPLC purification was distilled twice. Ddz-protected amino acids and benzotriazol-1-yloxytris-pyrrolidinophosphonium hexafluorophosphate (PyBoP) were purchased from Advanced ChemTech (Louisville, Kentucky). Fmoc–Arg(Pmc) (Pmc = 2,2,5,7,8-pentamethylchroman-6-sulfonyl) and Fmoc–Lys(Dde) [Dde = 1-(4,4-dimethyl-2,6-dioxocyclohex-1-ylidene)ethyl] were obtained

from Novabiochem (Switzerland). Tentagel-NH_2 was obtained from RAPP-Polymere (Tübingen, Germany). Dichloromethane, *N,N*-dimethylformamide, acetic anhydride (Ac_2O) and pyridine were obtained from BDH (Poole, England). Piperidine, trifluoroacetic acid (TFA), diisopropylcarbodiimide (DIC), 1-hydroxybenzotriazole (HOBt), 1-methylimidazole (NMI), dimethylaminopyridine (DMAP) and 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (DEC) were supplied by ACROS (Geel, Belgium).

Amino acids used in the libraries: For the assembly of the first library the following building blocks were used: *N*-Fmoc-1-amino-1-cyclohexanoic acid,[12] *N*-Fmoc-isonipecotic acid,[12] *N*-Fmoc-α-aminoisobutyric acid,[13] *N*-Fmoc-sarcosine,[14] *N*-Fmoc-γ-aminobutyric acid,[15,16] *N*-Fmoc-4-*O-tert*-butylproline,[17] *N*-Fmoc-β-(uracil-1-yl)-α-D-alanine, *N*-Fmoc-β-(hypoxanthin-9-yl)-α-D-alanine, N^{α}-Fmoc-N^{δ}-(2-thiophenecarbonyl)ornithine,[18] N^{α}-Fmoc-N^{δ}-(*p*-toluenesulphonyl)ornithine,[18] N^{α}-Fmoc-N^{δ}-8-(quinolinesulfonyl)ornithine,[18] N^{in}-*tert*-butyloxycarbonyl-N^{α}-Fmoc-tryptophan, N^{δ}-*tert*-butyloxycarbonyl-N^{α}-Fmoc-ornithine and *N*-trityl-N^{α}-Fmoc-glutamine. The latter three were obtained from Advanced ChemTech, while synthesis of the other building blocks was carried out as described. Analytical data lacking in the literature are provided. During the course of our work the building blocks Fmoc–Sar, Fmoc–γ-Abu and Fmoc–Hyp (*t*Bu) have become available from Advanced Chemtech or Novabiochem.

General procedure for fluorenylmethyloxycarbonylations: Protection of the α-amino moiety with the fluorenylmethyloxycarbonyl group was carried out as described here for the synthesis of Fmoc–Hex.

***N*-(9-fluorenylmethyloxycarbonyl)-1-aminocyclohexanecarboxylic acid (Fmoc–Hex)**:[12] Na_2CO_3 (25 g, 5 equiv), 1,4-dioxane (200 mL) and 9-fluorenylmethyloxycarbonyl chloride (14.2 g, 55 mmol, 1.1 equiv) were added to a solution of 1-aminocyclohexanecarboxylic acid hydrochloride (8.9 g, 50 mmol) in H_2O (250 mL). After stirring for 16 h, the reaction mixture was poured into 300 mL of H_2O and the solution was extracted three times with diethyl ether. The aqueous layer was acidified with a 2N solution of hydrochloric acid in water to pH 2 and the colourless suspension was extracted three times with ethyl acetate. The combined organic layer was dried on anhydrous $MgSO_4$, filtered and evaporated. The obtained white solid was crystallised from nitromethane to yield 11.68 g of Fmoc–Hex (32 mmol, 64 %). M.p. 178 °C; ^{13}C NMR ([D_6]DMSO): $\delta = 21.30$ (C-3, C-5), 25.28 (C-4), 32.09 (C-2, C-6), 47.00 (Fmoc, C-9′), 58.44 (C-1), 65.60 (Fmoc, CH_2O), 120.32 (Fmoc, C-4′), 125.63 (Fmoc, C-1′), 127.31 (Fmoc, C-2′), 127.87 (Fmoc, C-3′), 140.96 (Fmoc, C-11′), 144.12 (Fmoc, C-10′), 155.45 (OCONH), 176.22 (COOH); exact mass (LSIMS, thioglycerol) calcd for $C_{22}H_{24}NO_4$ [M+H]$^+$ 366.1705, found 366.17160; elem. anal. for $C_{22}H_{23}NO_4$ calcd C 72.32, H 6.34, N 3.83; found C 72.37, H 6.34, N 3.77.

***N*-(9-fluorenylmethyloxycarbonyl)isonipecotic acid (Fmoc–Inp)**:[12] The title compound was prepared according to the general procedure as described for Fmoc–Hex and was crystallised from nitromethane (80 %). M.p. 182 °C; ^{13}C NMR ($CDCl_3$): $\delta = 28.06$ (C-3), 41.06 (C-2), 43.64 (2 × C-4), 47.85 (Fmoc, C-9′), 67.84 (Fmoc, CH_2O), 120.48 (Fmoc, C-4′), 125.43 (Fmoc, C-1′), 127.54 (Fmoc, C-2′), 128. 19 (Fmoc, C-3′), 141.84 (Fmoc, C-11′), 144.48 (Fmoc, C-10′), 155.80 (OCONH), 180.44 (COOH); exact mass (LSIMS, thioglycerol) calcd for $C_{21}H_{22}NO_4$ [M+H]$^+$ 352.1548, found 352.1536; elem. anal. for $C_{21}H_{21}NO_4$ calcd C 71.78, H 6.02, N 3.99; found C 71.78, H 6.09, N 3.93.

***N*-(9-fluorenylmethyloxycarbonyl)-α-aminoisobutyric acid (Fmoc–Aib)**:[13] Following the general procedure, the title compound was synthesised and crystallised from nitromethane in 83 % overall yield. M.p. 178 °C (ref. [13] 179–180 °C); ^{13}C NMR ([D_6]DMSO): $\delta = 25.28$ (2 × CH_3), 46.79 (Fmoc, C-9′), 55.29 (α-C), 65.43 (Fmoc, CH_2O), 120.21 (Fmoc, C-4′), 125.42 (Fmoc, C-1′), 127.20 (Fmoc, C-2′), 127.75 (Fmoc, C-3′), 140.82 (Fmoc, C-11′), 143.96 (Fmoc, C-10′), 155.09 (OCONH), 175.98 (COOH); exact mass (LSIMS, glycerol) calcd for $C_{19}H_{20}NO_4$ [M+H]$^+$ 326.1392, found 326.1389; elem. anal. for $C_{19}H_{19}NO_4$ calcd C 70.14, H 5.89, N 4.30; found C 69.89, H 5.92, N 4.48.

***N*-(9-fluorenylmethyloxycarbonyl)sarcosine (Fmoc–Sar)**:[14] The title compound was crystallised from hexane–EtOAc in 90 % yield. M.p. 113 °C (ref. [14] 43–45 °C); ^{13}C NMR ([D_6]DMSO): $\delta = 35.13$, 35.78 (NCH_3), 50.02, 50.40 (α-C), 67.18, 67.35 (Fmoc, CH_2O), 120.39 (Fmoc, C-4′), 125.31 (Fmoc, C-1′), 127.39 (Fmoc, C-2′), 127.94 (Fmoc, C-3′), 141.00 (Fmoc, C-11′), 127.94 (Fmoc, C-3′), 141.00 (Fmoc, C-12′), 144.04 (Fmoc, C-10′), 155.79, 156.17 (OCONH), 171.13, 171.28 (COOH); exact mass (LSIMS, NBA–NaOAc) calcd for $C_{18}H_{17}NO_4Na$ [M+Na]$^+$ 334.1055, found 334.1054; elem. anal. for $C_{18}H_{17}NO_4$ calcd C 69.44, H 5.50, N 4.50; found C 69.42, H 5.58, N 4.70.

***N*-(9-fluorenylmethyloxycarbonyl)γ-aminobutyric acid (Fmoc–γ-Abu)**:[15, 16] Sodium carbonate (10 g), 1,4-dioxane (50 mL) and 9-fluorenylmethyloxycarbonyl chloride (5.72 g, 22 mmol, 1.1 equiv) were added to a solution of γ-amino butyric acid (2.1 g, 20 mmol) in H_2O (100 mL). The reaction mixture was stirred for 2.5 h and poured into H_2O (200 mL). The solution was extracted three times with diethyl ether; the aqueous layer was acidified under vigorous stirring with a 2N solution of hydrochloric acid in water. The white precipitate was isolated by filtration, washed with diethyl ether and dried in vacuo. The name product was crystallised from ethyl acetate to yield 6.3 g (19.2 mmol, 96 %). Analytical data are in agreement with ref. [16]. M.p. 170 °C (ref. [16] 166–168 °C); elem. anal. for $C_{19}H_{19}NO_4$ calcd C 70.14, H 5.89, N 4.30; found C 70.25, H 5.77, N 4.28.

***N*-(9-fluorenylmethyloxycarbonyl)-4-*O-tert*-butyl proline [Fmoc–Hyp(*t*-Bu)]**: The title compound was synthesised in 3 steps from hydroxyproline as described.[17] ^{13}C NMR ([D_6]DMSO): $\delta = 28.15$ (*t*Bu, CH_3), 37.74 (C-3), 46.48 (Fmoc, C-9′), 53.00, 53.99 (C-5), 58.62 (C-2), 66.51, 66.99 (Fmoc, CH_2O), 68.40, 69.20 (C-4), 73.44 (*t*Bu, C–O), 120.15 (Fmoc, C-4′), 125.14 (Fmoc, C-1′), 127.16 (Fmoc, C-2′), 127.75 (Fmoc, C-3′), 140.85 (Fmoc, C-11′), 144.00 (Fmoc, C-10′), 154.07 (OCONH), 175.23 (COOH); exact mass (LSIMS, thioglycerol) calcd for $C_{25}H_{27}NO_4$ [M+H]$^+$ 410.19673; found 410.19673.

N^{α}-*tert*-butyloxycarbonyl-β-(uracil-1-yl)-α-D-alanine: This compound was synthesised before as a racemic mixture using a Michael addition reaction[19] and as an enantiomerically pure compound by opening of Boc-serine β-lactone.[20] To a suspension of uracil (5.04 g, 45 mmol, 1.1 equiv) in anhydrous DMF (170 mL), NaH (80 % dispersion in oil, 1.8 g, 40 mmol, 1 equiv) was added. The suspension was stirred at room temperature for 1 h and cooled to −78 °C. A solution of *N-tert*-Boc-D-serine-β-lactone[21,22] (7.4 g 40 mmol) in anhydrous DMF (70 mL) was added dropwise for 30 min. After stirring for 16 h at room temperature, the solvent was evaporated to dryness. The obtained oil was dissolved in water and acidified with a 2N solution of hydrochloric acid in water. The colourless suspension was extracted three times with ethyl acetate. The organic layer was dried on anhydrous $MgSO_4$, filtered and evaporated to dryness. The residual oil was purified by flash column chromatography on silica gel (elution 93 % DCM/6 % MeOH/1 % HOAc or TEA) yielding 8.03 g of the title compound (26.8 mmol, 67 %). ^{13}C NMR ([D_6]DMSO): $\delta = 8.71$ (TEA, CH_3), 28.21 (Boc, CH_3), 45.01 (TEA, CH_2), 50.87 (β-CH_2), 53.19 (α-CH), 77.96 (Boc, C–O), 100.04 (U, C-5), 146.33 (U, C-6), 151.08 (U, C-2), 155.21 (OCONH), 164.05 (U, C-4), 172.14 (COO^-); exact mass (LSIMS, thioglycerol) calcd for $C_{12}H_{18}N_3O_6$ [M+H]$^+$ 300.1195; found 300.1198.

N^{α}-(9-fluorenylmethyloxycarbonyl)-β-(uracil-1-yl)-α-D-alanine (Fmoc–AlaU): A solution of N^{α}-*tert*-Boc-β-uracil-1-yl)-α-alanine (8.00 g) in a mixture of TFA/DCM 1:1 (100 mL) was stirred for 16 h. The solvent was evaporated and the foam obtained was dissolved in a 10 % aqueous solution of sodium carbonate (200 mL). To this clear yellow solution was added 1,4-dioxane (100 mL) and a solution of FmocCl (7.58 g, 29.4 mmol, 1.1 equiv). The reaction mixture was stirred for 16 h at room temperature, poured into water (200 mL) and extracted three times with diethyl ether. The aqueous phase was acidified by a 2N aqueous solution of hydrochloric acid to pH 2 and extracted three times with ethyl acetate. The combined organic layer was dried on anhydrous $MgSO_4$, filtered and evaporated. The title compound was crystallised from DCM yielding 8.42 g (20 mmol, 77 %). M.p. 221 °C; ^{13}C NMR ([D_6]DMSO): $\delta = 46.68$ (Fmoc, C-9′), 48.65 (β-CH_2), 52.24 (α-CH), 65.93 (Fmoc, CH_2O), 100.65 (U, C-5), 120.67 (Fmoc, C-4′), 125.30 (Fmoc, C-1′), 127.20 (Fmoc, C-2′), 127.79 (Fmoc, C-3′), 140.83 (Fmoc, C-11′), 143.80 (Fmoc, C-10′), 146.16 (U, C-6), 150.99 (U, C-2), 156.15 (OCONH), 163.83 (U, C-4), 171.23 (COOH); exact mass (LSIMS, thioglycerol) calcd for $C_{22}H_{20}N_3O_6$ [M+H]$^+$ 422.1352; found 422.1359; elem. anal. for $C_{22}H_{19}N_3O_6$ calcd C 62.70, H 4.54, N 9.97; found C 62.63, H 4.29, N 10.03.

N^{α}-*tert*-Butyloxycarbonyl-β-(6-chloropurin-9-yl)-α-D-alanine: Likewise, this compound was synthesised as a racemic mixture by means of a Michael addition reaction[19] and as an enantiomeric pure compound by opening of Boc-serine β-lactone.[20] To a solution of 6-chloropurine (5.41 g, 35 mmol, 1.2 equiv) in anhydrous DMF (150 mL) was added NaH (80 %

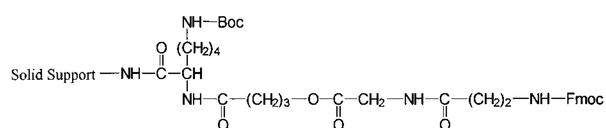
© WILEY-VCH Verlag GmbH, D-69451 Weinheim, 1998
0947-6539/98/0403-0430 $ 17.50+.50/0 *Chem. Eur. J.* **1998**, *4*, No. 3

dispersion in oil, 900 mg, 30 mmol, 1 equiv). After 1 h, the solution was added dropwise over 1 h to a chilled solution of *N-tert*-Boc-D-serine-*β*-lactone (5.4 g, 30 mmol, 1 equiv) in anhydrous DMF (90 mL). After stirring for 1 h at $-78\,^{\circ}C$, the reaction mixture was kept at room temperature for 16 h. The solvent was evaporated to dryness and the residual oil was taken up in a saturated solution of sodium bicarbonate and extracted three times with diethyl ether to remove most of the excess of 6-chloropurine. The aqueous layer was acidified with a 2 N aqueous solution of hydrochloric acid to pH 2 and extracted three times with ethyl acetate. The combined organic layer was dried on anhydrous $MgSO_4$, filtered and evaporated. Purification of the title compound was realised by medium-pressure chromatography on silica gel (elution DCM 93 %/MeOH 6 %/HOAc 1 %) to afford 5.8 g of the title compound (17 mmol, 56 %). ^{13}C NMR ([D$_6$]DMSO): $\delta = 28.00$ (Boc, CH_3), 45.21 (β-CH_2), 53.67 (α-CH), 78.34 (Boc, C–O), 130.80 (Pu, C-5), 147.97 (Pu, C-8), 148.88 (Pu, C-4), 151.51 (Pu, C-2), 152.47 (Pu, C-6), 155.01 (OCONH), 171.19 (COOH); exact mass (LSIMS, thioglycerol) calcd for $C_{13}H_{17}N_5O_4Cl$ $[M+H]^+$ 342.0969; found 342.0962.

N^{α}-9-fluorenylmethyloxycarbonyl-*β*-(hypoxanthin-9-yl)-*α*-D-alanine (Fmoc–AlaH): A solution of N^{α}-*tert*-Boc-*β*-(6-chloropurin-9-yl)-*α*-D-alanine (5.46 g, 16 mmol) in a mixture TFA/water 3:1 (100 mL) was stirred for 48 h. The solvents were evaporated and the residue was coevaporated twice with toluene. The oil thus obtained was dissolved in H_2O (200 mL) while sodium carbonate (20 g) was added carefully in fractions. Thereafter 1,4-dioxane (100 mL) and a solution of FmocCl (4.6 g, 18 mmol, 1.13 equiv) in 1,4-dioxane (75 mL) were added. The reaction mixture was stirred for 16 h and poured into water and extracted three times with diethyl ether. The aqueous phase was acidified with a 2 N solution of hydrochloric acid in water to pH 2 and was extracted three times with ethyl acetate. The extracts were pooled, dried over anhydrous $MgSO_4$ and evaporated in vacuo. The title compound was crystallised from EtOH to afford 6.0 g (13.6 mmol, 85 %). M.p. $196\,^{\circ}C$; ^{13}C NMR ([D$_6$]DMSO): $\delta = 45.35$ (β-CH_2), 46.64 (Fmoc, C-9′), 55.96 (α-CH), 65.66 (Fmoc, CH_2O), 120.15 (Fmoc, C-4′), 123.69 (Pu, C-5), 125.15 (Fmoc, C-1′), 127.15 (Fmoc, C-2′), 127.66 (Fmoc, C-3′), 140.70 (Pu, C-8), 140.93 (Fmoc, C-11′), 143.82 (Fmoc, C-10′), 145.34 (Pu, C-2), 148.95 (Pu, C-4), 155.63 (OCONH), 156.82 (Pu, C-6), 172.33 (COOH); exact mass (LSIMS, thioglycerol) calcd for $C_{23}H_{20}N_5O_5$ $[M+H]^+$ 446.1464; found 446.1463; elem. anal. for $C_{23}H_{19}N_5O_5$ calcd C 62.02, H 4.30, N 15.72; found C 61.81, H 4.57, N 15.49.

Evaluation of coupling characteristics of selected building blocks: Oligopeptides selected for building block evaluation: HOOC-*β*-Ala-Sar-Gln-Hex-Inp-Hex-NH_2; HOOC-Sar-Hyp-Inp-AlaU-Trp-NH_2, HOOC-AlaH-Inp-Aib-Mfg-NH_2, HOOC-Aib-Aib-Inp-Inp-NH_2, HOOC-*β*-Ala-Tos-Inp-Thi-Qui-*β*-Ala-NH_2. Peptide bonds were formed by means of DIC, HOBt, DIEA and 2 equiv PyBoP on a 4-hydroxybutyrate-derivatised Tentagel S-NH_2.[8] No PyBoP was added in the case of less sterically hindered amino acids (Qui, Thi, Tos). DMAP was used as catalyst for the incorporation of the first amino acid. The standard procedure used 4 equiv of Fmoc amino acid. In the case of AlaU and AlaH, the substitution of HOBt by HOAt as well as the addition of PyBoP was necessary to get satisfactory coupling yields. For the coupling of Fmoc–Tos and Fmoc–Qui building blocks, DIEA had to be removed from the condensation mixture as it induced an intramolecular cyclisation reaction.[10] These coupling conditions gave over 95 % yield for every individual reaction. Peptides were cleaved from the solid support using 0.2 N NaOH/CH_3CN (1:1). The purity was verified by PLRP-S^R chromatography and the identity was determined by LSIMS analysis. All profiles showed one main compound peak.

Synthesis of the encoded library: For the encoded sequence, Ddz-protected natural amino acids were assembled by means of HOBt, DIC and DIEA with DCM as solvent (except for Ddz–Asn, for which DMF was used). Fmoc–Lys(Boc) was used to functionalise the amino resin, using the α-amino function to attach the unnatural building blocks and the ε-amino function to assemble the encoding natural amino acids. The derivatised support is shown in Scheme 2. A hydroxybutyric acid linker,[8] followed by glycine and a β-alanine spacer, was used between the α-amino function of lysine and the first unnatural amino acid. 4-*O*-MMTr-4-hydroxybutyric acid was coupled by TEA/DEC/DMAP in pyridine. After removal of the MMTr group with 2 % trichloroacetic acid in DCM, Fmoc–Gly was coupled by means of HOBt, DIC, NMI. β-Alanine was introduced to avoid diketopiperazine formation using HOBt, DIEA, DIC in DCM/DMF. Completion of acylation was monitored by the ninhydrin[23] test, and residual amines were acetylated with pyridine/Ac_2O/NMI in order to avoid formation of contaminant hybrid structures. Coupling reactions using Fmoc amino acids were evaluated by means of bromophenol blue[24] and coupling yields were determined by UV spectrometry. Following each cycle, the same capping procedure was executed as used for the Ddz amino acid coupling. Before starting the screening procedure, the acid-labile side-chain protecting groups were removed using TFA in the presence of thioanisole, *m*-cresol and ethanedithiol.



Scheme 2. Structure of the derivatised solid support used for the synthesis of the encoded library.

Introduction of the hydroxybutyric acid linker: DEC (2.29 g) and DMAP (800 mg) were added to a solution of 4-*O*-monomethoxytrityl-4-hydroxybutyric acid[8] (4.27 g, 2 mmol) in anhydrous pyridine (12 mL), TEA (712 mL). After 20 min of preactivation, Tentagel S-NH_2 ($\phi = 130$ µm, 300 µmol/g, 10.5 g) was added and the suspension was shaken at room temperature for 24 h. The beads were filtered from the solution, washed with anhydrous pyridine and DCM and dried in vacuo. By measuring the absorption of a solution of an accurate aliquot of resin (≈ 5 mg) in 50.0 mL of a mixture of $HClO_4$/EtOH (3:2) at 475 nm, the loading was calculated ($\varepsilon = 55\,000$) and was found to be complete.

Coupling of Ddz-amino acids: 1,4-Dioxane (30 mL), ice, and a cooled 0.5 N solution of $KHSO_4$ in water to acidify the mixture were added to a solution of the dicyclohexylammonium or cyclohexylammonium salt of the Ddz-amino acid (1 mmol) in water (50 mL). The mixture was extracted three times with ethyl acetate. The organic layer was dried on $MgSO_4$, filtered and evaporated to dryness, yielding the Ddz-amino acid as a carboxylic acid. The obtained oily product was dissolved in DCM (2.5 mL) and HOBt (135 mg, 1 mmol), DIC (160 µL, 1 mmol), and DIEA (180 µL, 1 mmol) were added. The reaction mixture was kept at room temperature for 15 min and transferred into the reaction vessel, which was shaken at 200 rpm for 4 h. The amino acid solution was removed and the resin was washed three times with DMF and three times with DCM.

Coupling of Fmoc amino acids: Deprotection was carried out with 20 % piperidine in DMF for 15 min and the support was washed five times each with DMF and with DCM. A solution of Fmoc amino acid (1 mmol), HOBt/HOAt (1 mmol), DIC (1 mmol), PyBoP (1 mmol), DIEA (1 mmol), in DMF (3 mL) was stored at room temperature for 15 min and was added to the amino-functionalised resin. The reaction vessels were shaken for 16 h, washed three times with DMF, three times with DCM and once with diethyl ether. The beads were dried in vacuo and the Fmoc substitution level was determined by dissolving an accurately measured quantity of resin (10 mg) in 20 % piperidine in DMF (exactly 25.0 mL). After 10 min, the absorbance was measured at 300 nm versus a 20 % solution of piperidine in DMF as a blank allowing calculation of the substitution level ($\varepsilon = 7500$). All the coupling yields for each amino acid are represented in Tables 5 and 6.

Recombining and mixing: All the beads were pooled in a 500 mL Omnifit bottle. The flask was half-filled with DMF/DCM 1:1 and shaken in three dimensions for 30 min, after which the suspension was transferred back to a large sintered-glass filter. The solvent was then removed by suction filtration and the resin was washed twice with DCM and once with diethyl ether, dried and redistributed.

Preparation of the library for screening: After the final assembly cycle, the resin (1 g) was collected and Fmoc-deprotected, and the liberated amino groups were acetylated. Finally, the acid-labile side-chain protecting groups were cleaved by a 50 % TFA treatment for 3 h in the presence of the traditional carbocation scavengers thioanisole and *m*-cresol. Additionally, 1,2-ethanedithiol was added to the deprotection cocktail as it is an extremely good scavenger for *tert*-butyl cations and is particularly effective in preventing the acid-catalysed oxidation of tryptophan.[25] In this way, the *tert*-butyl, Boc and Ddz groups were removed quantitatively. Extensive

Table 5. The quantitative Fmoc determinations of each amino acid in each cycle during the synthesis of the first library.

Fmoc AA	Cycle 1	Cycle 2	Cycle 3	Cycle 4	Cycle 5
AlaU	96%	105%	112%	89%	104%
Hex	88%	87%	97%	89%	87%
Inp	93%	94%	97%	100%	93%
Hyp	109%	88%	101%	101%	88%
Sar	101%	89%	105%	104%	100%
Trp	96%	89%	107%	102%	101%
Orn	86%	93%	103%	101%	96%
Qui	98%	99%	108%	109%	100%
Gln	81%	87%	103%	99%	96%
γ-Abu	93%	97%	107%	108%	96%
Thi	94%	93%	103%	98%	104%
Aib	87%	91%	103%	99%	92%
Tos	89%	95%	106%	102%	98%
AlaH	100%	90%	100%	100%	104%

Table 6. The quantitative yields of Fmoc for the individual amino acids in each coupling step of the second library.

	Cycle 1	Cycle 2	Cycle 3	Cycle 4	Cycle 5
Thi	84%	94%	90%	107%	98%
Hex	83%	88%	90%	88%	93%
Hyp	83%	96%	87%	107%	104%
Tos	90%	100%	97%	102%	98%
AlaH	102%	105%	100%	106%	105%
AlaU	92%	102%	95%	106%	106%
Trp	95%	103%	87%	98%	94%
Sar	95%	102%	96%	103%	105%
Gln	86%	95%	86%	95%	109%
Aib	86%	87%	86%	86%	90%
Qui	90%	103%	93%	106%	92%
Inp	91%	89%	87%	94%	96%

washing with DCM, 5% DIEA in DCM, and DCM, and drying yielded the library ready to be screened against any soluble target.

Synthesis of the TAR hairpin: On a 1-*O*-dimethoxytrityl-1,3-propanediol-functionalised LCAA-CPG, prepared analogously to a reported synthesis,[26] the oligoribonucleotide TAR loop sequence 5′-fluorescein-GGCCAGAUCUGAGCCUGGGAGCUCUCUGGCC-1,3-propanediol-3′ was assembled from phosphoramidites purchased from Biogenex (San Ramon, CA). The standard 1 μmol scale RNA synthesis cycle from an Applied Biosystems 394 DNA synthesiser was used. The addition of the 5′-fluorescein was performed with Fluoro-Prime (Pharmacia) phosphoramidite. The oligoribonucleotide was cleaved from the support by concentrated NH_3/EtOH (3:1) at 55 °C for 5 h, filtered, evaporated and treated with a triethylamine trihydrofluoride solution in methylpyrrolidone[27] (0.5 mL) for 1.5 h at 60 °C. After precipitation with 3 M sodium acetate and 1-butanol, the precipitate was purified by reversed-phase HPLC. Two main fluorescent products were obtained in a 2:1 ratio and were assigned to the different isomers present on the fluorescein molecule. The different fractions containing the desired product were pooled and desiccated by extraction of water by 1-butanol.

Screening of the encoded library: The labelled oligonucleotides were obtained from Eurogentec, Belgium. The T_m values were measured in a buffer containing 0.1 M NaCl, 0.02 M K_2HPO_4 (pH 7.4), 0.1 mM EDTA at a 4 μM concentration. The duplex was prepared by addition of equimolar amounts of each single-stranded oligonucleotide to a solution of 0.1 M sodium chloride in phosphate buffer, pH 7.4. After 30 min of equilibration, the resin (1.1 g) was added and incubated for 3 h at room temperature with the doubly labelled target. The beads were collected on a filter and washed 6 times with the same 0.1 M NaCl solution, pH 7.4 (20 mL). The library was divided between 8 petri dishes and was visualised under a fluorescence microscope. In the UV light positive yellow fluorescent beads were isolated with the help of a micropipette and evaluated in a second selection cycle. During the second cycle the beads were evaluated in visible light to compare the intensity of the red colour of the beads.

In the second part of the screening procedure, the hydroxybutyrate linker connecting the noncoding sequence was hydrolysed using 0.1 N NaOH:CH_3CN (1:1) for 4 h. The beads were neutralised by transfer to a petri dish filled with a buffer containing 0.1 M NaCl, 0.02 M K_2HPO_4 (pH 7.4) and 0.1 mM EDTA (4 mL). Repetition of the screening procedure on these beads using affinity interaction with the doubly labelled dsDNA fragment did not generate any colour or fluorescence in any of these beads. Resynthesis of only the natural coding peptide sequence proved redundant as no interaction could be seen with the target.

Finally the beads were washed, picked up, fixed individually on filters and subjected to Edman degradation.

Synthesis and purification of peptides selected by the screening process: Peptide syntheses were carried out starting from 270 μmol of benzhydrylamine-modified Tentagel[28] useful for the synthesis of C-terminal amides (since screening was carried out on a solid support). Each coupling reaction was monitored by Fmoc determination. Fmoc amino acids were coupled using 4 equiv of building blocks, HOBt, DIC, DIEA (and PyBoP, HOAt for AlaU and AlaH). Capping was performed with Ac_2O/pyr/NMI 1:4:1 for 5 min. After synthesis of oligopeptides, side-chain protecting groups were removed and products were cleaved from the solid support using TFA, thianisol, *m*-cresol, H_2O. The peptide was precipitated by means of diisopropylether, purified by PLRP-S^R (gradient 5% CH_3CN/H_2O to 50% CH_3CN/H_2O) and identified by LSIMS. The total isolated yield ranged from 50 to 86% (Table 7).

For the synthesis of the bifurcated compound, likewise, the benzhydrylamine linker was selected. The hydroxybutyrate linker was substituted by its amino analogue (γ-aminobutyric acid) in order to increase the stability of compounds. The protected (Fmoc–Lys(Dde)) lysine was bound to the solid support using HOBt, DIC, DIEA in DMF. The Fmoc group was removed with 20% piperidine in DMF. The molecules were obtained by first assembling the unnatural peptide sequence on the α-amino group of the lysine moiety followed by preparation of the natural peptide on the δ-amino group (using Fmoc chemistry in both cases). The Dde group at the ε-amino group of lysine was removed using 2% hydrazine in DMF. Cleavage and purification conditions are identical to those for the octapeptides. The

Table 7. Coupling yields, isolated weight, calculated and observed molecular weight of the individual peptides as determined by LSIMS.

Peptide sequence		Weight (mg)	Yield (%)	M_r calcd	M_r found
Ac-Arg-AlaU-Sar-AlaU-Sar-AlaU-β-Ala-Gly-$CONH_2$	**OP1**	63	63	1028	1029
Ac-Arg-AlaU-Sar-Hyp-AlaU-Inp-β-Ala-Gly-$CONH_2$		51	54	1000	1001
Ac-Arg-AlaU-Sar-Inp-Sar-Sar-β-Ala-Gly-$CONH_2$	**OP2**	53	70	848	849
Ac-Arg-AlaU-AlaU-Gln-AlaU-Sar-β-Ala-Gly-$CONH_2$		64	50	1085	1086
Ac-Arg-AlaU-Sar-AlaU-Hyp-Gln-β-Ala-Gly-$CONH_2$		46	48	1017	1018
Ac-Arg-AlaU-AlaU-Inp-AlaU-Sar-β-Ala-Gly-$CONH_2$	**OP3**	90	72	1068	1069
Ac-Arg-AlaU-Hyp-Inp-Thi-AlaH-β-Ala-Gly-$CONH_2$		85	70	1177	1178
Ac-Arg-AlaU-Sar-Sar-Sar-Sar-β-Ala-Gly-$CONH_2$		63	67	808	809
Ac-Arg-AlaU-Sar-Hyp-Thi-Hyp-β-Ala-Gly-$CONH_2$	**OP4**	84	72	1045	1046
Ac-Arg-AlaU-AlaU-Inp-AlaU-Hyp-β-Ala-Gly-$CONH_2$		71	57	1110	1111
Ac-Arg-AlaU-Sar-Inp-AlaU-Sar-β-Ala-Gly-$CONH_2$	**OP5**	99	86	958	959
Ac-Arg-AlaU-AlaU-Sar-Sar-Sar-β-Ala-Gly-$CONH_2$		68	66	918	919

Table 8. Bifurcated peptides synthesised on the α and ε amino group of lysine.

Bifurcation number	Weight (mg)	Yield (%)	M_r calcd	M_r found
1	62	34	1820.8	1850.3
2	65	40	1616.8	1617.0
3	102	55	1849.8	1850.1
4	110	63	1751.8	1752.0
5	40	25	1788.8	1789.2

identity of the bifurcations was determined by electrospray MS. The total yield of isolated compound ranged from 25% to 63% (Table 8).

The DNase I experiments: The oligonucleotide 3′-TCTAACACGTTACA-5′ was radiolabelled (^{32}P) at the 5′-end by means of T4 polynucleotide kinase (Gibco BRL) and [γ-^{32}P]ATP (4500 Ci/mmol, ICN) by standard procedures[29] and purified on a NAP-5® column (Pharmacia). The radiolabelled oligonucleotide and its cold complement 5′-AGATTGTG-CAATGT-3′ were dissolved in water, both in 1 µM concentration. The peptides were dissolved in 1 mM concentration in PBS. In a total reaction volume of 30 µL, 3 µL of both oligonucleotides were mixed with 3 µL of 10 × DNase I reaction buffer (200 mM Tris–HCl pH 8.4, 20 mM $MgCl_2$ and 500 mM KCl). The mixture was stored at room temperature for 15 min to allow hybridisation of the DNA strands. Subsequently, 3 µL of the peptide solution was added, resulting in a dsDNA–peptide ratio of 1:1000. The mixtures were further stored at room temperature for 40 h. The control tube received no peptide. The degradation reaction was started upon addition of 0.0125 U DNase I (Amersham) and was performed at 25 °C. At appropriate time intervals aliquots were withdrawn, mixed with twice the volume of formamide stop mix (xylene cyanol FF 0.05%, bromophenol blue 0.05%, 50 mM EDTA in 90% formamide), frozen on dry ice and stored in a freezer until analysis. Full-length (intact) and digested oligonucleotides were resolved on a 20% denaturing polyacrylamide gel containing urea (8.3 M) with TBE buffer[29] at 1000 V over 1.25 h, followed by autoradiography. Scanning laser densitometry was performed with a DeskTop Densitometer (pdi, NY, USA) equipped with Discovery Series® (Diversity One®) software. The resulting figures were corrected for the amount of degradation already present in the nondegraded blank. The results are the average of two independent experiments.

Acknowledgments: The authors are grateful to the Onderzoeksfonds K. U. Leuven for an OT grant. This work was supported by VIS 95/11 and by a grant from the FWO-Vlaanderen (Project 3.0105.94). Dr. T. Lescrinier is a research assistant and Dr. A. Van Aerschot is a research associate of the Belgian National Fund for Scientific Research. Prof. J. Van Beeumen is indebted to the Flemish Government for Concerted Research Action 12052293. We are indebted to M. Wiersma for the excellent synthesis of the TAR hairpin. We thank M. Vandekinderen for editorial help.

Received: June 30, 1997 [F744]

[1] J. W. Trauger, E. E. Baird, P. B. Dervan, *Nature* **1996**, *382*, 559–561.
[2] P. E. Nielsen, M. Egholm, R. H. Berg, O. Buchardt, *Science* **1991**, *254*, 1497–1500.
[3] L. A. Carpino, *J. Org. Chem.* **1988**, *53*, 875–878.
[4] K. S. Lam, S. E. Salmon, E. M. Hersh, V. J. Hruby, W. M. Kazmierski, R. J. Knapp, *Nature* **1991**, *354*, 82–84.
[5] a) J. M. Kerr, S. C. Banville, R. N. Zuckermann, *J. Am. Chem. Soc.* **1993**, *115*, 2529–2531; b) V. Nikolaiev, A. Stierandova, V. Krchnak, B. Seligmann, K. S. Lam, S. E. Salmon, M. Lebl, *Peptide Res.* **1993**, *6*, 161–170.
[6] H. Isshiki, S. Akira, O. Tanabe, T. Nakajima, T. Shimamoto, T. Hirano, T. Kishimoto, *Mol. Cell. Biol.* **1990**, *10*, 2757–2764.
[7] R. Boyce, G. Li, H. P. Nestler, T. Suenaga, W. C. Still, *J. Am. Chem. Soc.* **1994**, *116*, 7955–7956.
[8] a) J. Haralambidis, L. Duncan, G. W. Tregear, *Tetrahedron Lett.* **1987**, *28*, 5199–5202; b) J. Haralambidis, L. Duncan, K. Angus, G. W. Tregear, *Nucleic Acids Res.* **1990**, *18*, 493–499.
[9] M. Gait, J. Karn, *Trends Biochem. Sci.* **1993**, *18*, 255–259.
[10] A. Van Aerschot, T. Saison-Behmoaras, J. Rozenski, C. Hendrix, G. Schepers, G. Verhoeven, P. Herdewijn, *Bull. Soc. Chim. Belges* **1995**, *104*, 717–720.
[11] C. R. Vinson, P. B. Sigler, S. L. McKnight, *Science* **1989**, *246*, 911–916.
[12] A. P. Combs, T. M. Kapoor, S. Feng, J. K. Chen, L. G. Daude-Snow, S. L. Schreiber, *J. Am. Chem. Soc.* **1996**, *118*, 287–288.
[13] G. Valle, G. M. Bonora, C. Toniolo, *Can. J. Chem.* **1984**, *62*, 2661–2666.
[14] B. E. B. Sandberg, C.-M. Lee, M. R. Hanley, L. L. Iversen, *Eur. J. Biochem.* **1981**, *114*, 329–337.
[15] A. Van Aerschot, P. Herdewijn, H. Vanderhaeghe, *Nucleosides Nucleotides* **1988**, *7*, 75–90.
[16] J. W. Drijfhout, W. Bloemhoff, *Int. J. Peptide Protein Res.* **1991**, *37*, 27–32.
[17] C. G. Fields, V. L. VanDrisse, G. B. Fields, *Pept. Res.* **1993**, *6*, 39–47.
[18] T. Lescrinier, R. Busson, H. De Winter, C. Hendrix, G. Janssen, C. Pannecouque, J. Rozenski, A. Van Aerschot, P. Herdewijn, *Peptide Res.* **1997**, *49*, 183–189.
[19] A. J. H. Nollet, U. K. Pandit, *Tetrahedron* **1969**, *25*, 5983–5987.
[20] a) I. Lewis, *Tetrahedron Lett.* **1993**, *34*, 5697–5700; b) I. Lewis, C. Bruns, *Tetrahedron* **1994**, *50*, 7485–7494.
[21] L. D. Arnold, T. H. Kalantar, J. C. Vederas, *J. Am. Chem. Soc.* **1985**, *107*, 7105–7109.
[22] L. D. Arnold, J. C. G. Drover, J. C. Vederas, *J. Am. Chem. Soc.* **1987**, *109*, 4649–4659.
[23] E. Kaiser, R. L. Colescott, C. D. Bossinger, P. J. Cook, *Anal. Biochem.* **1970**, *34*, 595–598.
[24] V. Krchnak, J. Vagner, P. Safar, M. Lebl, *Coll. Czech. Chem. Commun.* **1988**, *53*, 2542–2548.
[25] D. King, C. Fields, G. Fields, *Int. J. Peptide Protein Res.* **1990**, *36*, 255–266.
[26] J. Y. Tang, Q. Guo, A. Roskey, S. Agrawal, *J. Cell. Biochem.* **1993**, *17*, 214 (S417).
[27] F. Wincott, A. DiRenzo, C. Shaffer, S. Grimm, D. Tracz, C. Workman, D. Sweedler, C. Gonzalez, S. Scaringe, N. Usman, *Nucleic Acids Res.* **1995**, *23*, 2677–2684.
[28] M. S. Bernatowicz, S. B. Daniels, H. Koster, *Tetrahedron Lett.* **1989**, *30*, 4645–4648.
[29] T. Maniatis, E. Fritsch, J. Sambrook, *Molecular cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, New York, **1989**, p. 12.1–12.13.