Available online at www.sciencedirect.com
ScienceDirect
Tetrahedron 63 (2007) 6577–6586
Tetrahedron

Expanding the scope of PNA-encoded libraries: divergent synthesis of libraries targeting cysteine, serine and metallo-proteases as well as tyrosine phosphatases

François Debaene, Julien A. Da Silva, Zbigniew Pianowski, Fernando J. Duran and Nicolas Winssinger*

Institut de Science et d'Ingénierie Supramoleculaires, Université Louis Pasteur, CNRS, 8 allée Gaspard Monge, 67000 Strasbourg, France

Received 17 January 2007; revised 28 February 2007; accepted 2 March 2007
Available online 12 March 2007

Abstract—Seven PNA-encoded combinatorial libraries targeting proteases and phosphatases with covalent reversible and irreversible mechanism-based inhibitors were prepared. The libraries were synthesized using modified PNA monomers, which dramatically increase the water solubility of the libraries in biologically relevant buffers. The libraries were shown to selectively inhibit targeted enzymes.

1. Introduction

We have previously demonstrated that libraries encoded with PNA tags can be synthesized in a split and mix format, cleaved from the resin to obtain the library as a mixture in solution, and reformatted by hybridization to an oligonucleotide microarray (Fig. 1).[1–6] We have further shown that such libraries can be used to measure multiple enzymatic activities in complex biological samples such as crude cell lysates or used to discover new inhibitors. This technology has recently been embraced by other research groups.[7–9] For small molecule screening, an interesting asset of this method is that library members that bind to an enzyme in a mixture can be separated from unbound compounds by size exclusion separation or SDS gel.[3,6] A limiting factor in these separations is the poor solubility of PNA in biologically relevant buffers and PNA's inherent propensity to aggregate due to the neutral peptidic backbone. This fact has thus far required the use of denaturing conditions for the size exclusion separation and is limiting the concentration or size of the libraries that can be used.

2. Results and discussion

2.1. Modified PNA monomers

PNA monomers containing a primary amino group charged at biologically relevant pHs were evaluated for their propensity to disrupt aggregation and improve water solubility. To

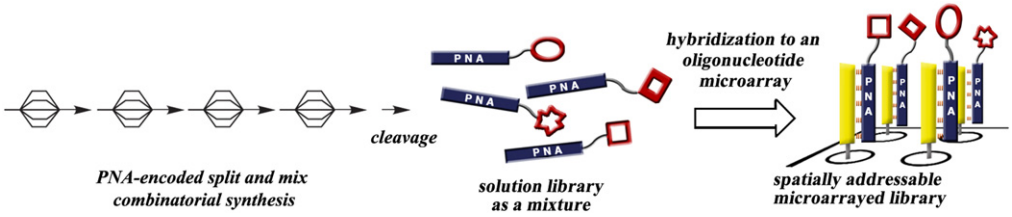


Figure 1. PNA-encoded split and mix synthesis and conversion of crude libraries to a microarray format by sequence-specific self-assembly.

Keywords: PNA; Protease inhibitors; Phosphatase inhibitors; Combinatorial chemistry; Microarrays.
Abbreviations: CDI, 1,1′-carbodiimidazole; DBU, 1,8-diazabicyclo[5.4.0]undec-7-ene; DIC, diisopropylcarbodiimide; 4-DMAP, 4-dimethylaminopyridine; DMP, Dess–Martin periodinane; DSO, di(*N*-succinimidyl)oxylate; EDC, 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride; EDTA, ethylenediamine tetraacetic acid; HOBt, 1-*N*-hydroxybenzotriazole; HBTU, *O*-benzotriazole-*N*,*N*,*N*′,*N*′-tetramethyl-uronium-hexafluoro-phosphate; kDa, kilodaltons; KHMDS, potassium hexamethyldisilazide; NMM, *N*-methyl morpholines; NMP, *N*-methylpyrrolidone; PBS, phosphate buffered saline; SucOH, *N*-hydroxy-succinimide; TFA, trifluoroacetic acid; TSTU, *O*-(*N*-succinimidyl)-*N*,*N*,*N*′,*N*′-tetramethyluronium tetrafluoroborate.
* Corresponding author. Tel.: +33 (0)3 90 24 51 13; fax: +33 (0)3 90 24 51 12; e-mail: winssinger@isis.u-strasbg.fr

0040–4020/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tet.2007.03.033

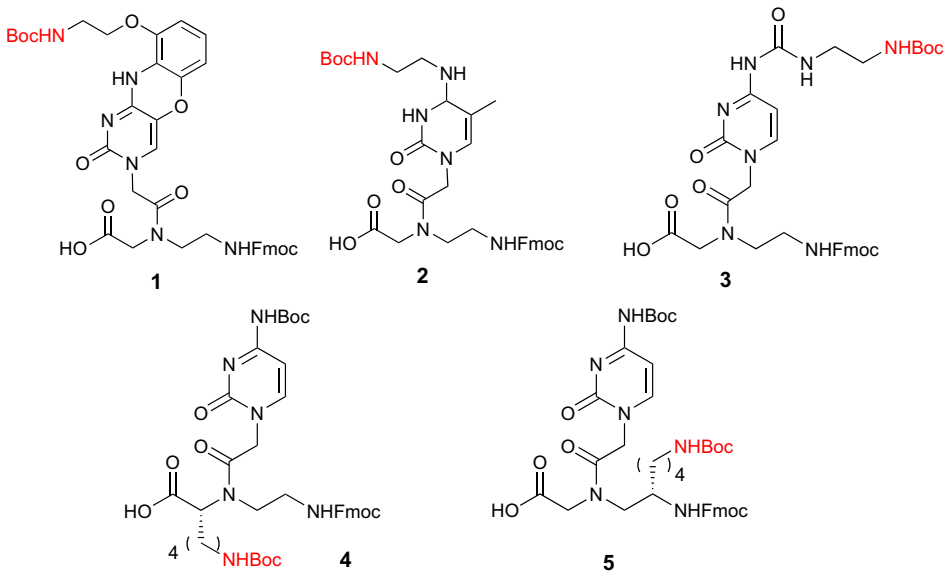
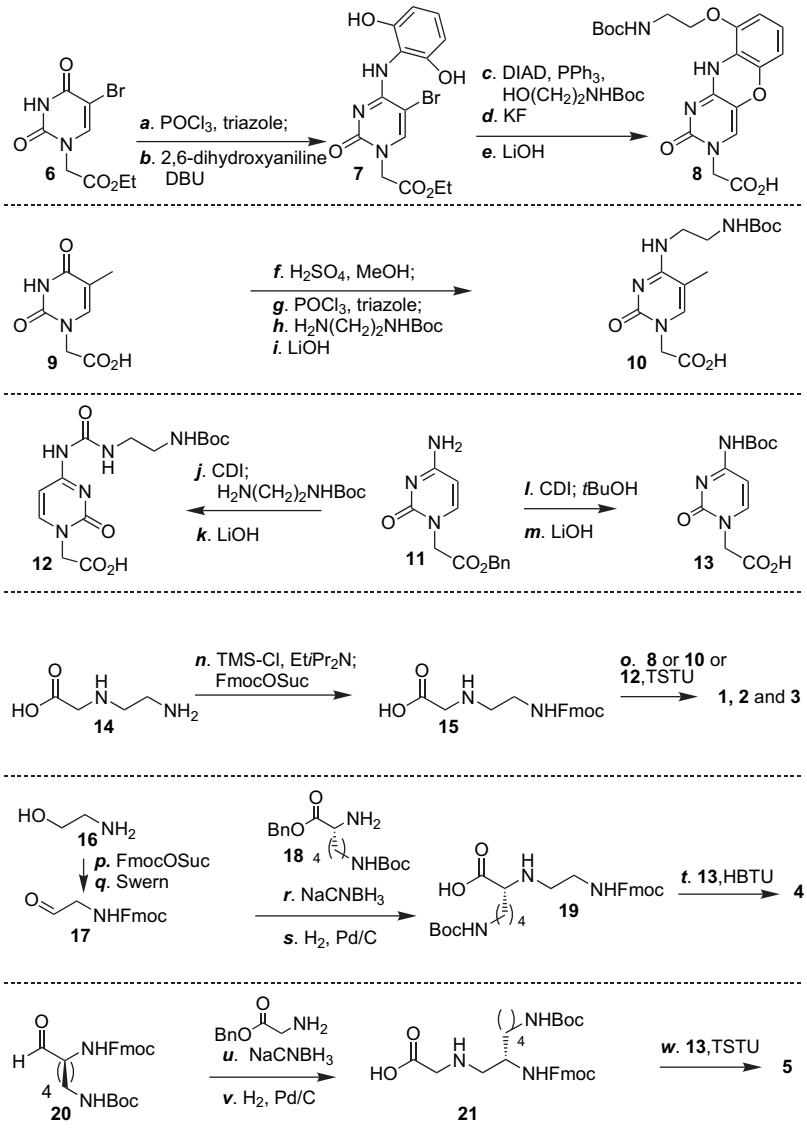


Figure 2. Modified PNA monomers designed to improve water solubility and disrupt aggregation.

this end, a series of modified PNA monomers bearing a Boc-protected primary amino group were envisioned (Fig. 2). Upon cleavage of the library from Rink resin, these groups would be deprotected and provide a solubilizing charge. We have reported a validated codon system with optimized hybridization conditions[6] and based on the prominence of cytosine in the codons, we opted to focus on modified cytosines. The so-called G-clamps (**1**, Fig. 2) have been described in the literature[10,11] and are known to dramatically improve hybridization affinity, however, the impact of the additional amino group on the overall solubility of PNA oligomers was not reported. A simplified version deprived of the tricyclic moiety but preserving the protected amino group was also considered (**2**, Fig. 2). While this later modification was not expected to improve binding affinity, it should improve water solubility. *N*-carbamoyl derivatives of cytosine have also been described in the context of natural phosphoribose-based oligonucleotides,[12,13] which led us to consider PNA monomer **3**. Last but not least, two types of backbone modifications were also considered, modification of the glycine portion[14,15] and of the ethyl-amino portion[16,17] with the side chain of a lysine (**4** and **5**, respectively, Fig. 2). The substitution of the glycine moiety for a lysine is known to improve water solubility and leads to a slight increase in hybridization affinity for the *R* stereoisomer.[14,15,18] The monomers bearing an Fmoc at the N-terminus and Boc for the exocyclic nitrogen and side chain were prepared as shown in Scheme 1. The tricyclic moiety of the G-clamp was prepared following previously reported chemistry, however, with a Boc-protected side chain. Thus, bromo-thymine **6** (Scheme 1) was converted to compound **7** and one of the hydroxyl groups was derivatized via Mitsunobu reaction followed by a cyclization of the other hydroxyl to obtain the tricyclic system. Hydrolysis of the methyl ester with LiOH afforded the modified cytosine **8** in 60% yield from **6**. Monomer **10** was prepared in four steps from thymine acetic acid **9** using the same activation chemistry ($POCl_3$, triazole) as for **6** followed by displacement with mono-Boc-protected ethylenediamine and hydrolysis of the methyl ester. The *N*-carbamoyl cytosine derivative **12** was prepared in two steps from benzyl acetate **11**[19] via formation of the isocyanate and reaction with mono-Boc-protected ethylenediamine followed by hydrolysis of the benzyl ester. We had previously used a Bhoc to protect the exocyclic nitrogens of the nucleotides, however, we have found that the corresponding Boc-protected monomers are more soluble in the coupling solutions (NMP) and afford better results in addition to being more acid stable. Boc-protected cytosine acetic acid **13** was obtained from the cytosine benzyl acetate **11**[19] in two steps.

The *N*-(2-aminoethyl)glycine **14** was selectively protected with Fmoc via in situ silylation[20] and treatment with Fmoc-OSuc to afford **15**. Surprisingly, it was found that coupling yield of the different nucleotide acetic acids to this backbone **15** was better if the acid functionality of **15** remained unprotected. The modified oligonucleotides **8**, **10**, and **12** were coupled to **15** using TSTU activation to obtain products **1**–**3**, respectively, in moderate to good yield. The aminoethyl lysine backbone **19** was prepared in four steps from ethanolamine starting with an Fmoc protection of the amino group and Swern oxidation of the alcohol to obtain aldehyde **17**, which was engaged in a reductive amination with H-Lys(Boc)OBn followed by a hydrogenolysis. Compound **21** was prepared via a similar sequence involving a reductive amination of glycine benzyl ester and Fmoc-protected lysinal **20** followed by a reductive debenzylation. Coupling of cytosine acetic acid **13** to the modified backbone **19** required the more reactive HBTU activation affording compound **4** in 79% while coupling of **13** to **21** was effective with TSTU (58%).

Oligomers containing two to four modifications capped with a Cy3 fluorophore were prepared on an automated synthesizer (Applied Biosystem Expedite) or by semi-automated synthesis (Argonaut Quest) using standard protocols (HBTU activation in NMP and 20% piperidine in DMF for Fmoc deprotection). It is interesting to note that monomers **1** and **2** did not require a protecting group on the secondary aniline. We found that reaction with acetic anhydride in the presence of excess DMAP did not yield any acetylated product. These modified monomers were incorporated in PNA oligomers with comparable efficiency to the regular ones



Scheme 1. Synthesis of modified PNA monomers **1**–**5**. Reagents and conditions: (a) $POCl_3$ (2.0 equiv), 1,2,4-triazole (9.0 equiv), Et_3N (15 equiv), MeCN, 0 °C, 30 min, then **6**, 23 °C, 4 h; (b) 2,6-dihydroxyaniline (1.5 equiv), DBU (1.6 equiv), CH_2Cl_2, 23 °C, 5 h, 70% (two steps); (c) PPh_3 (1.5 equiv), $HO(CH_2)_2NHBoc$ (1.2 equiv), DIAD (1.2 equiv), THF, 23 °C, 4 h; (d) KF (10 equiv), EtOH, reflux, 36 h; (e) LiOH (4.0 equiv), MeOH/H_2O (2:1), 23 °C, 30 min, 86% (three steps); (f) H_2SO_4 concd (0.7 equiv), MeOH, reflux, 1 h, 98%; (g) $POCl_3$ (2.0 equiv), 1,2,4-triazole (9.0 equiv), Et_3N (15 equiv), MeCN, 0 °C, 30 min, add nucleobase, 23 °C, 4 h; (h) $H_2N(CH_2)_2NHBoc$ (1.5 equiv), DBU (2.0 equiv), CH_2Cl_2, 23 °C, 14 h, 55% (two steps); (i) LiOH (2.0 equiv), MeOH/H_2O (1:1), 23 °C, 20 min, 98%; (j) CDI (1.6 equiv), $H_2N(CH_2)_2NHBoc$ (1.5 equiv), DMF, 6 h, 92%; (k) LiOH (4 equiv), MeOH, 0 °C, 86%; (l) CDI (1.6 equiv), *t*-BuOH (1.6 equiv), DMF, 60 °C, 12 h, 93%; (m) LiOH (4.0 equiv), MeCN, H_2O, 23 °C, 10 min, 70%; (n) TMS-Cl (2.0 equiv), Fmoc-OSuc (0.95 equiv), NMM (4.0 equiv), DMF, CH_2Cl_2, 0 °C, 2.5 h, 61%; (o) **8**, **10** or **12** (1.0 equiv), DIPEA (1.3 equiv), TSTU (1.1 equiv), DMF, 0 °C, 20 min, then **15** (1.0 equiv), Et*i*Pr_2N (1.3 equiv), H_2O, MeCN, 23 °C, 1 h, 27% for **1**, 56% for **2**, 57% for **3**; (p) Fmoc-OSu (0.9 equiv), Et*i*Pr_2N (1.2 equiv), CH_2Cl_2, 23 °C, 2.5 h, 91%; (q) $(COCl)_2$ (1.5 equiv), Me_2SO (3.0 equiv), Et_3N (4.0 equiv), CH_2Cl_2/THF (2:1), −78 °C→0 °C, 30 min; (r) **18** (1.0 equiv), MeOH, 0 °C, 30 min, then CH_3CO_2H concd (1.2 equiv), $NaBH_3CN$ (0.5 equiv), 0 °C, 1 h, 59% (two steps); (s) Pd/C 5%, MeOH, H_2, 23 °C, 4 h, 94%; (t) **13** (2.2 equiv), HBTU (2.0 equiv), Et*i*Pr_2N (2.5 equiv), 2,6-lutidine (2.0 equiv), DMF, 23 °C, 20 min, then **19** (1.0 equiv), Et*i*Pr_2N (1.3 equiv), 23 °C, 2 h, 74%; (u) **20** (1.0 equiv), benzyl glycinate (1.0 equiv), $NaBH_3CN$ (0.55 equiv), AcOH (1.3 equiv), MeOH, 0 °C, 30 min; 23 °C, 1 h, 46%; (v) Pd/C 5%, MeOH, H_2, 23 °C, 2.5 h, 92%; (w) **13** (1.0 equiv), TSTU (1.1 equiv), Et*i*Pr_2N (2.6 equiv), DMF, MeCN, H_2O, 23 °C, 80 min, 58%.

(96–99% average yield). For monomer **3**, while the couplings were efficient, mass spectrometric analysis (MALDI) of the intermediates and the final product were complicated by fragmentation of the product making qualitative assessment of the success of a synthesis difficult. Monomer **4** coupled with high efficiency, however, monomer **5** gave poor yield after the second or third incorporation and its oligomer could not be evaluated for binding affinity. The oligomers containing modified monomers **1**–**4** were then individually hybridized to a microarray containing all permutations of the codons (625 sequences) using the condition optimized[6] for unmodified PNA (see Supplementary information for hybridization data). The oligomers containing the G-clamp **1** showed reduced affinity compared to unmodified PNA while oligomers containing modified monomer **2** and **3** showed no affinity under these conditions. The reduced affinity observed with the oligomer containing the G-clamp may appear contradictory to previous reports. However, the use of formamide in the hybridization may reduce the hydrophobic stacking contribution of this tricyclic system. Oligomers containing the modified monomer **4** were found to bind with comparable affinity and specificity in these conditions.

Table 1. Solubility of unmodified (Lys(Alloc)-Lys- GCCG-CGA-CGA-GACG-Lys(Cy3)NH_2) and modified (Lys(Alloc)-Lys- GCC*G-C*GA-C*GA-GAC*G-Lys(Cy3)NH_2) PNAs at pH 5.5 (LCPB buffer: 100 mM NaAc, 100 mM NaCl, 1 mM EDTA, and 0.01% Brij-35); pH 7 (PBS buffer 200 mM) and in pure H_2O

	pH 5.5 (LCPB)	pH 7.0 (PBS)	H_2O
Unmodified PNA	23%	19%	21%
PNA containing monomer **1**	60%	51%	47%
PNA containing monomer **4**	100%	100%	100%

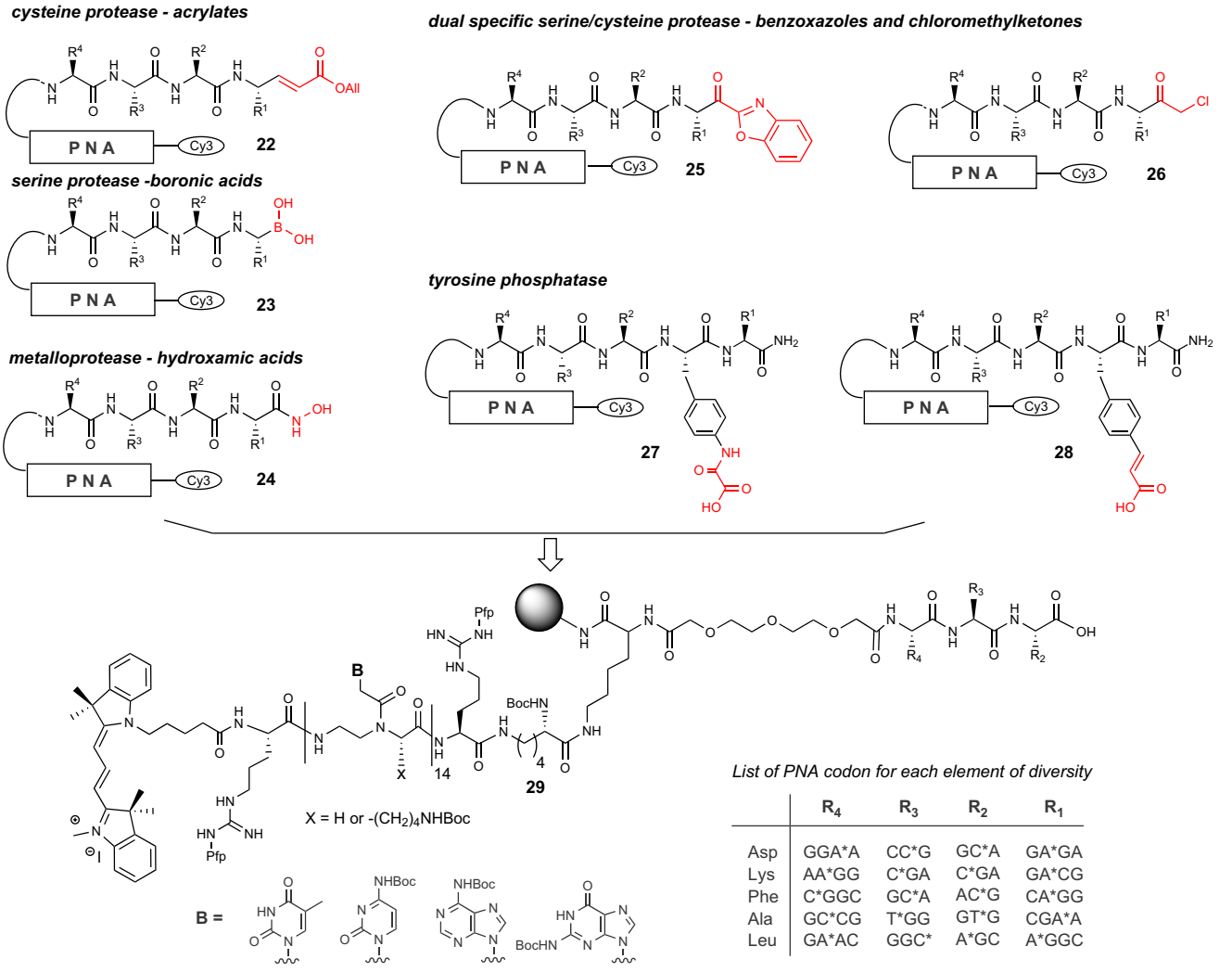
The values represent the percentage of PNA remaining in solution after sonication of a 10 μM PNA solution and centrifugation (concentration measured by fluorescence).

Table 2. Size exclusion filtration of a 10 μM solution of unmodified (Lys(Alloc)-Lys- GCCG-CGA-CGA-GACG-Lys(Cy3)NH_2) and modified PNAs(Lys(Alloc)-Lys- GCC*G-C*GA-C*GA-GAC*G-Lys(Cy3)NH_2) PNAs at pH 5.5 in LCPB buffer (100 mM NaAc, 100 mM NaCl, 1 mM EDTA, and 0.01% Brij-35)

Unmodified PNA	20%
PNA containing monomer **1**	36%
PNA containing monomer **4**	70%

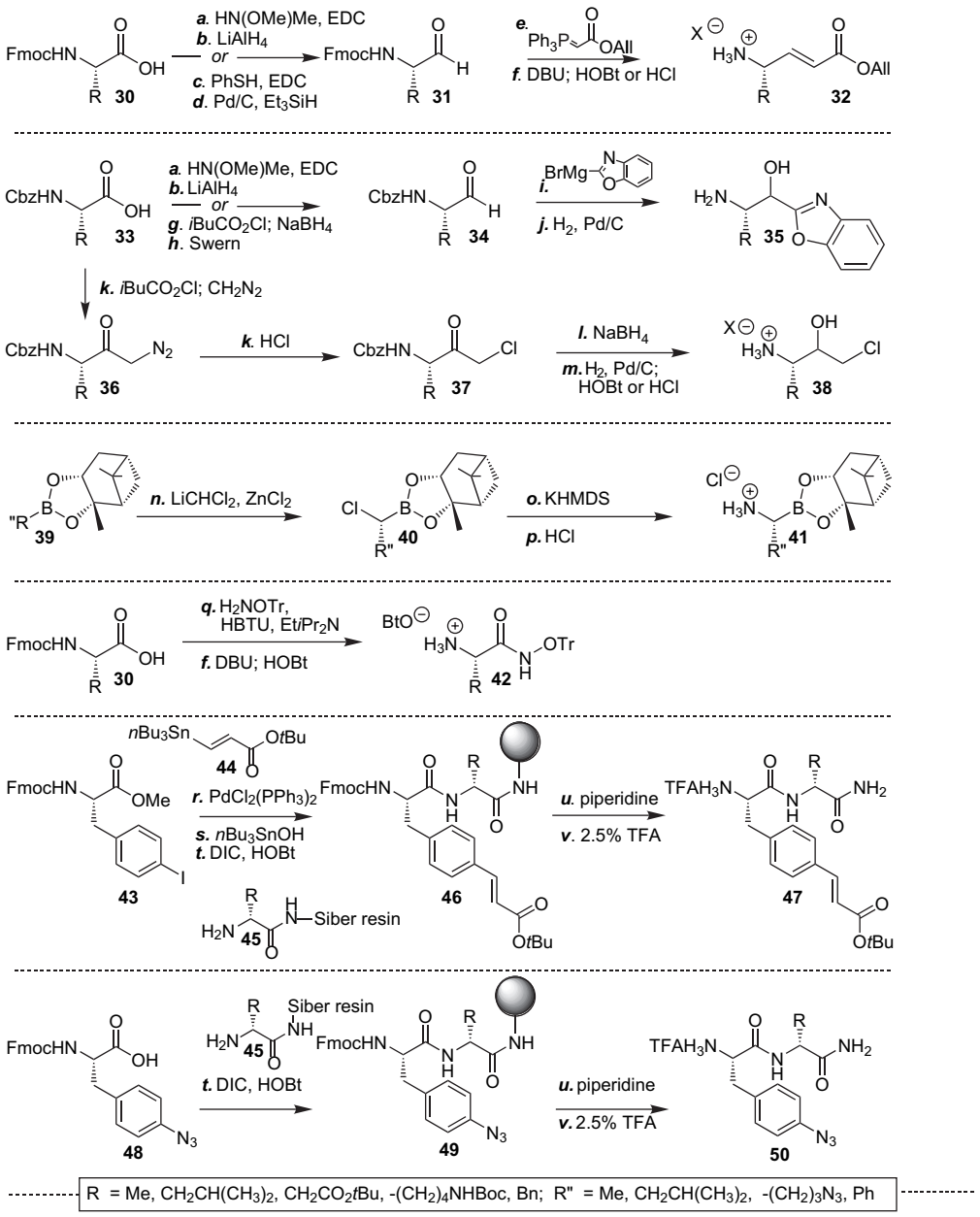
The values represent the percentage of the PNA that passed through a 50 kDa size exclusion filter (VivaScience, Sartorius) in the first filtration.

It should be noted that formamide was previously found to be important in reducing non-specific interaction during hybridization and the addition of up to six cationic charges in a 14-mer PNA did not lead to unspecific binding to the microarray due to electrostatic interactions with the negatively charged DNA probes. Based on subsequent finding regarding the solubility of the different oligomers (vide infra), hybridization of G-clamp oligomers with different buffers was not investigated. Oligomers containing modified monomers **1** and **4** were evaluated for their solubility in pure water and in buffers containing 200 mM salt at pH 5.5 and 7.0. The oligomers were sonicated at 50 °C for 10 min and then centrifuged to pellet all colloidal suspensions. As shown in Table 1, the unmodified oligomer is poorly soluble at 10 μM concentration with only 19–23% of the compound remaining in solution after centrifugation. The incorporation of G-clamp monomer **1** did improve solubility of PNA, however, precipitates still remained. Gratifyingly, the oligomer with modified monomer **4** was completely soluble under these conditions. We then tested the efficiency of size exclusion filtration using different oligomers by comparing the fluorescence intensity of the filtrates using a 10 μM solution. As shown in Table 2, while 70% of the fluorescence was



	R_4	R_3	R_2	R_1
Asp	GGA*A	CC*G	GC*A	GA*GA
Lys	AA*GG	C*GA	C*GA	GA*CG
Phe	C*GGC	GC*A	AC*G	CA*GG
Ala	GC*CG	T*GG	GT*G	CGA*A
Leu	GA*AC	GGC*	A*GC	A*GGC

Figure 3. General strategy for the divergent synthesis of libraries **22–28** targeting diverse enzyme families. All libraries are prepared from common intermediate **29** (asterisk in the codons at positions of modified PNA).



Scheme 2. Synthesis of modified amino acids **32**, **35**, **38**, **41**, **42**, **47**, and **50**. Reagents and conditions: (a) HN(OMe)Me (1.1 equiv), EDC (1.1 equiv), HOBt (1.1 equiv), Et*i*Pr$_2$N (1.1 equiv), DMF, 12 h; (b) LiAlH$_4$ (1.0 equiv), THF, −20 °C, 1 h, 66–80% (two steps); (c) PhSH (1.1 equiv), DIC (1.6 equiv), 4-DMAP (0.05 equiv), CH_2Cl_2, 23 °C, 30 min, 94% for Asp, 94% for Lys; (d) Et_3SiH (4.0 equiv), Pd/C (0.2 mol %), CH_2Cl_2, 23 °C, 90 min, 80–87%; (e) phosphonium iodide (1.6 equiv), Et*i*Pr$_2$N (1.4 equiv), toluene, 80 °C, 74–84%; (f) DBU (1.0 equiv), CH_2Cl_2, 10 min; HOBt (2.0 equiv); (g) *i*-BuCO$_2$Cl (1.0 equiv), Et_3N (1.2 equiv), THF, 30 min, −15 °C→23 °C, $NaBH_4$ (1.2 equiv) in H_2O, 30 min, 70% for Asp; (h) oxalyl chloride (1.5 equiv), DMSO (3.0 equiv), Et_3N (4.0 equiv), CH_2Cl_2, −78 °C→23 °C, 30 min, 65% for Asp; (i) benzoxazole (3.0 equiv), *i*-PrMgCl (2.0 equiv), THF, −10 °C, 40 min; **34** (1.0 equiv), −10 °C, 2 h→23 °C, 2 h, 65–75%; (j) Pd/C (0.1 mol %), EtOH, H_2, 23 °C, 12 h, 85–96%; (k) NMM (1.4 equiv), *i*-BuCO$_2$Cl (1.3 equiv), THF, −25 °C, 1 h; CH_2N_2 (2.0 equiv), −25 °C→0 °C, 1 h; HCl concd/CH_3CO_2H (1:1); (l) $NaBH_4$ (1.05 equiv), THF/EtOH (10:1), 23 °C, 30 min, 34–65% (two steps); (m) H_2, Pd/C (0.1 mol %), HCl (1.0 equiv) or HOBt (1.1 equiv), MeOH, 23 °C, >98%; (n) CH_2Cl_2 (1.2 equiv), LDA (1.1 equiv), $ZnCl_2$ in Et_2O (1.7 equiv), −20 °C→23 °C overnight, 66–78%; (o) KHMDS (1.0 equiv), THF, from −20 °C to 23 °C, 12 h; (p) HCl in dioxane (3.3 equiv), THF, −10 °C→23 °C, 12 h, 43–56% (two steps); (q) HBTU (1.1 equiv), Et*i*Pr$_2$N (1.5 equiv), DMF, 23 °C, 15 min; H_2NOTr (1.0 equiv), Et*i*Pr$_2$N (1.3 equiv), 23 °C, 1 h, 30–67%; (r) $PdCl_2(PPh_3)_2$ (0.1 equiv), *n*-Bu$_3$SnCH═COO*t*-Bu (2.0 equiv), DMF, 90 °C, 4 h, 60%; (s) *n*-Bu$_3$SnOH (10 equiv), Cl(CH$_2$)$_2$Cl, reflux, 71%; (t) HOBt (2.5 equiv), DIC (25 equiv), DMF, 23 °C, 10 min; **45** (0.5 equiv), 23 °C, 2 h; (u) 20% piperidine in DMF, 23 °C, 5 min; (v) 2.5% TFA in CH_2Cl_2, 23 °C, 3×20 min.

recovered in the filtrates after the first concentration of oligomer with modified PNA **4**, only 20% and 36% were recovered for the unmodified PNA and oligomer containing the G-clamp **1**, respectively. These differences can be attributed to the level of aggregation of PNA in the buffer. Hence, the most suitable PNA modification for our purpose is monomer **4**.

2.2. General design of inhibitors

We then turned our attention to the synthesis of libraries targeting different families of enzymes implicated in cellular regulation through post-translational modification of proteins. We have previously shown that the acrylamide functionality could be used to target cysteine proteases.[2,3] The

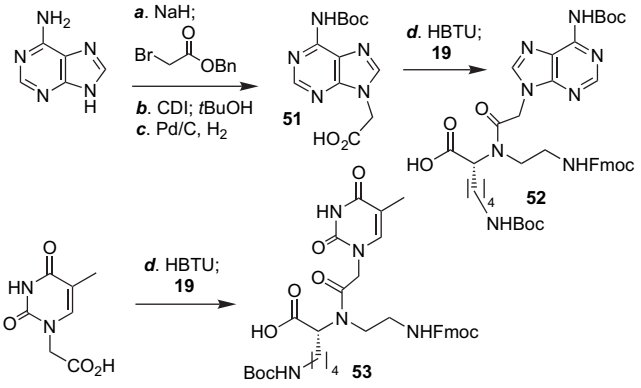
synthetic strategy previously used precluded the incorporation of other types of warheads. We reasoned that carrying out the synthesis of the peptide fragment from N- to C-terminus would allow us to incorporate the different warheads required for targeting specific enzymatic mechanisms. As shown in Figure 3, intermediate **29** could be used to access libraries containing acrylates (**22**), which target specifically cysteine protease,[21,22] boronic acids (**23**), which target specifically serine protease,[23,24] hydroxamic acids (**24**), which target metalloprotease,[25] and benzoxazole ketones (**25**)[26,27] and chloromethyl ketones (**26**),[22,28] which target both serine and cysteine proteases. Beyond proteases, phosphatases are known to be important post-translational regulators.[29] Two types of phosphotyrosine surrogate targeting tyrosine phosphatase were envisioned (**27** and **28**). Library **28** is based on a Michael acceptor, which can react with the nucleophilic cysteine residue involved in the phosphate hydrolysis. While Michael acceptors are known for phosphatase inhibition in the context of natural products[30] or peptidomimetics,[31,32] the tyrosine acrylate analog has not been reported. Conversely, the oxamic acid (**27**) functionality, which is an unhydrolizable mimic of a phosphophenol has been reported to lead to sub micromolar inhibitors.[33,34] The library was envisioned to have four points of diversity with five different amino acids at each position (Ala, Leu, Phe, Lys, and Asp) representing the diversity of amino acid functionalities, namely an aliphatic, aromatic, basic, and acidic amino acid.

2.3. Synthesis of modified amino acids

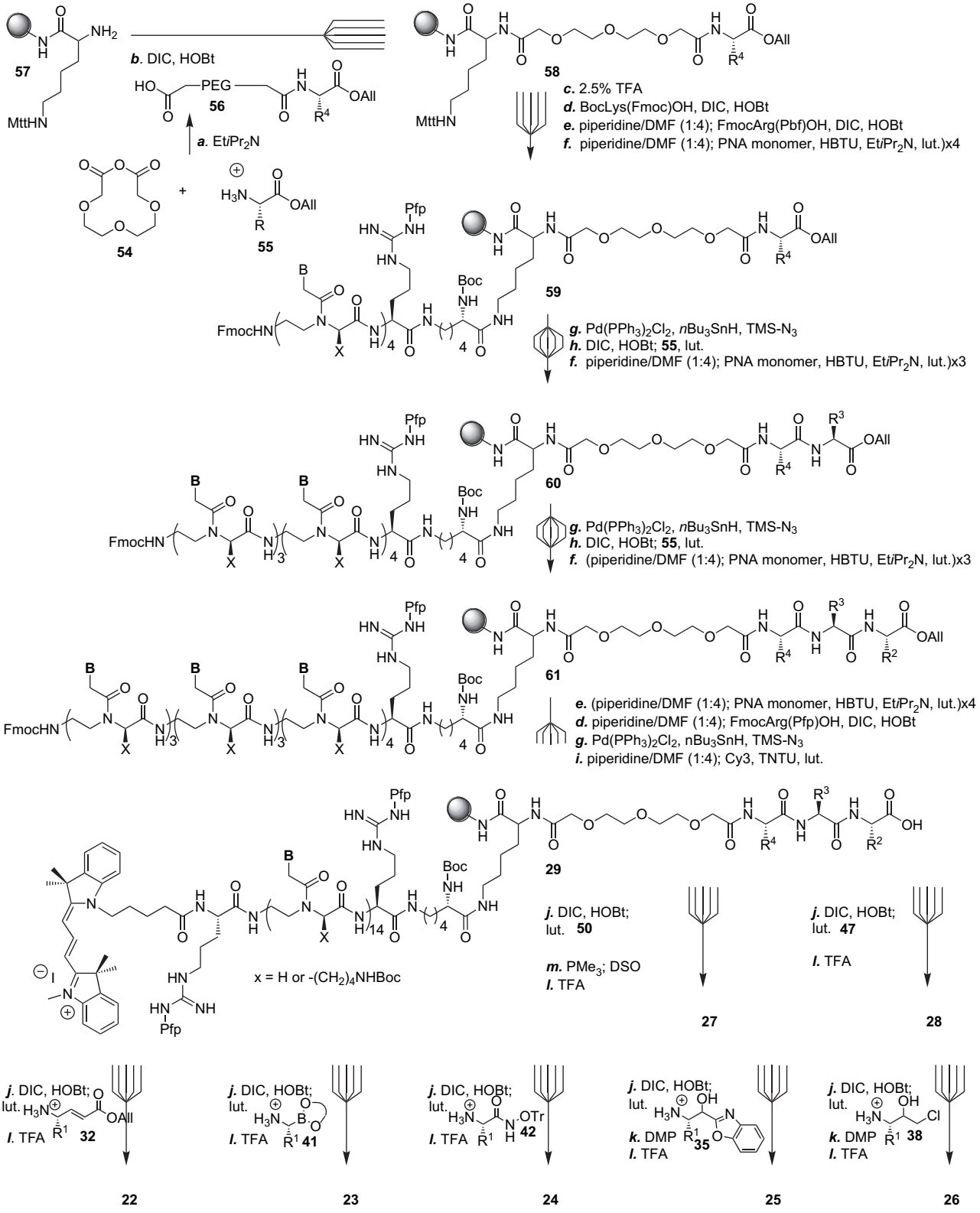
The required amino acid warheads were prepared as shown in Scheme 2. The acrylates **32** were obtained in four steps using similar procedures as previously reported for the acrylamide library.[2] Thus Fmoc-protected amino acids **30** were reduced to the aldehyde **31** either via a thioester (Asp and Lys) or a Weinreb amide (Ala, Phe, Leu) and were condensed with commercially available allyl phosphoniumacetate to obtain the Fmoc-protected acrylates. It should be noted that Fmoc deprotection of these acrylates with piperidine led to conjugate addition of the piperidine. Nevertheless, the compounds could be smoothly deprotected with DBU. Quenching of the reactions with HOBt allowed for the crude reaction products to be used directly in subsequent couplings (the DBU–HOBt salts do not interfere with the reaction). For the purpose of benzoxazole derivatives **35** and halohydrin derivatives **38**, the synthesis was carried out starting with Cbz-protected amino acid **33** thus allowing a facile final deprotection with hydrogenolysis. The amino acids **33** were converted to the corresponding aldehydes via the Weinreb amide except for Asp, which was reduced to the corresponding alcohol and reoxidized to the aldehyde by Swern oxidation. The aldehydes **34** were then reacted with the metalated benzoxazole prepared in situ with *i*-PrMgCl and subsequently hydrogenated to obtain the desired products **35** as diastereomeric mixtures. Deprotonation of the benzoxazole with *i*-PrMgCl afforded better results than with *n*-BuLi as previously reported for similar reactions.[35,36] Conversely, amino acids **33** were activated with *i*-$BuCO_2Cl$ and treated with diazomethane to obtain the diazonium compounds **36**, which were treated in situ with HCl thus affording chloromethyl ketones **37**.[37] Compounds **37** could not be deprotected under hydrogenolysis due to dechlorination and/or polymerization, however, reduction of the ketone ($NaBH_4$) followed by hydrogenation in the presence of 1 equiv of acid (HCl or HOBt) afforded the desired compound **38** without the aforementioned side reactions. All the benzoxazole and halohydrin monomers **35** and **38** were coupled to a resin loaded with phenylalanine, oxidized with DMP, cleaved from the resin, and analyzed by LC-MS. All the 10 reactions showed a single peak corresponding to the desired product suggesting that the preparation of **35** and **38** does not engender any epimerization. The boronic acids warheads were prepared according to the optimized procedure reported by Wityak and co-workers.[38] (+)-Pinanediol alkylboronates **39** were treated with the anion of dichloromethane generated in situ with LDA followed by $ZnCl_2$ to afford compounds **40**, which were engaged in S_N2 reactions with KHMDS followed by acid treatment to remove the silyl groups thus affording boronates **41**. The trityl-protected hydroxamic acids[39] **42** were prepared from the corresponding Fmoc-protected amino acids **30** using commercially available O-protected hydroxylamine. As for the acrylate, it was found that the crude product from the Fmoc deprotection could be used directly in the subsequent coupling if carried out with DBU and quenched with HOBt. The modified tyrosine warhead was obtained by a Stille coupling between 4-iodophenylalanine **43** and stannane **44**. Hydrolysis of the methyl ester using tributyl tin hydroxide[40] (no Fmoc deprotection was observed under these conditions) followed by coupling to Sieber resin **45** afforded polymer-bound dipeptide **46**. Deprotection of the Fmoc and cleavage from the resin with low TFA concentrations afforded the dipeptides **47** without cleavage of the side chain protecting groups. Dipeptide **50** was obtained in a similar fashion from 4-azidophenylalanine via intermediate **49**.

2.4. Divergent synthesis of libraries 22–28

Based on the improved properties of oligomers containing modified monomer **4**, we opted to include a modified PNA monomer in each codon of the library yielding a final library bearing a total of four homogeneously distributed modified PNA (Fig. 3). This required the preparation of modified A and T monomers (Scheme 3). Boc-protected adenine acetic



Scheme 3. Synthesis of modified A and T monomers **52** and **53**. Reagents and conditions: (a) NaH (1.2 equiv), DMF, 23 °C, 4 h; benzyl bromoacetate (1.1 equiv), 0 °C→23 °C overnight, 65%; (b) CDI (1.5 equiv), DMF, 105 °C, 2 h, *t*-BuOH (1.5 equiv), 23 °C, 12 h, 60%; (c) Pd/C, H_2, EtOH, 23 °C, 15 min, >98%; (d) HBTU (0.9 equiv), Et*i*Pr$_2$N (1.1 equiv), 2,6-lutidine (0.9 equiv), DMF, 23 °C, 15 min; **19** (0.5 equiv), Et*i*Pr$_2$N (0.6 equiv), 23 °C, 2 h, 61% for **52** and 55% for **53**.



Scheme 4. Synthesis of libraries **22**–**28**. Reagents and conditions: (a) **55** (1.0 equiv), Et*i*Pr$_2$N (2.0 equiv), 4-DMAP (0.1 equiv), CH_2Cl_2, 2 h, 23 °C, 34–52%; (b) **56** (5.0 equiv), HOBt (5.0 equiv), DIC (5.0 equiv), Et*i*Pr$_2$N (5.0 equiv), DMF, 2×3 h, 23 °C; (c) TFA 2.5%, Et_3SiH 5%, CH_2Cl_2, 3×15 min; Et*i*Pr$_2$N 0.1 M, 2×10 min, 23 °C; (d) BocLys(Fmoc)OH (4.0 equiv), HOBt (4.0 equiv), DIC (4.0 equiv), 2,6-lutidine (6.0 equiv), DMF, 3 h, 23 °C; (e) piperidine/DMF (1:4), 5 min; FmocArg(Pfp)OH (4.0 equiv), HOBt (4.0 equiv), DIC (4.0 equiv), 2,6-lutidine (6.0 equiv), DMF, 3 h, 23 °C; (f) piperidine/DMF (1:4), 5 min; PNA monomer (5.0 equiv), HBTU (4.4 equiv), 2,6-lutidine (7.5 equiv), Et*i*Pr$_2$N (5.0 equiv), NMP, 1 h; Ac_2O (10 equiv), 2,6-lutidine (10 equiv), DMF, 5 min, 23 °C; (g) Pd(PPh$_3$)$_2$Cl$_2$ (0.2 equiv), Ph_3P (0.8 equiv), TMS-N$_3$ (10 equiv), *t*-Bu$_3$SnH (5.0 equiv), CH_2Cl_2, 3×15 min, 23 °C; (h) **55** (5.0 equiv), HOBt (4.5 equiv), DIC (4.5 equiv), 2,6-lutidine (10 equiv), DMF, 10 h, 23 °C; (i) piperidine/DMF (1:4), 5 min; Cy3 (5.0 equiv), TNTU (4.5 equiv), 2,6-lutidine (10 equiv), premixed 1 h, NMP, 23 °C, 2×3 h; (j) warhead (5.0 equiv), HOBt (4.5 equiv), DIC (4.5 equiv), 2,6-lutidine (10 equiv), DMF, 2×3 h, 23 °C; (k) DMP (10 equiv), CH_2Cl_2, 1 h, 23 °C; (l) TFA/*m*-Cresol (4:1), 2 h, 23 °C; (m) PMe_3 (10 equiv), THF, 10 min, 23 °C; di(*N*-succinimidyl)oxylate (DSO) (10 equiv), DMF, 2 h, 23 °C.

Table 3. Inhibition of representative enzymes by the different libraries

	22	23	24	25	26	27	28
Cathepsin B	1–10 μM	>100 μM	>100 μM	1–10 nM	1–10 nM	>100 μM	>100 μM
Trypsin	>100 μM	10–100 nM	>100 μM	1–10 nM	10–100 nM	>100 μM	>100 μM
MMP2	>100 μM	>100 μM	1–10 μM	>100 μM	>100 μM	>100 μM	>100 μM
PTP 1B	>100 μM	>100 μM	>100 μM	>100 μM	>100 μM	2–10 μM	>100 μM

acid **51** was prepared in three steps from adenine and coupled to **19** using the same procedure as for cytosine to afford the required monomer **52**. The thymine monomer **53** was prepared using the same coupling procedure with commercially available thymine acetic acid.

The library was prepared on a Rink resin loaded at 0.2 mmol/g with a bifunctional lysine[2] (**57**, Scheme 4), which was split into five pools. In order to introduce a spacer between the inhibitor and the PNA tag, the first amino acid residue **55**, protected at the C-terminus as an allyl ester[41] was reacted with PEG-anhydride **54**.[42] A simple work up was sufficient to remove all unreacted amine and PEG and the crude product **56** could be coupled directly to the resin **57** to obtain five pools of polymer bound intermediates **58**. For each pool, the trityl group was removed with 2.5% TFA, and the respective PNA codon was coupled after the introduction of a lysine spacer and arginine residue affording five pools of resin **59**. The completion of each reaction sequence was verified by LC-MS showing a single peak for the desired product. The five pools were then combined and the allyl ester was removed under π-allyl transfer catalysis. While this deprotection is very fast with *n*-Bu_3SnH, it sometimes fails to go to completion due to the competing decomposition of tin hydride to ditin and molecular hydrogen. The addition of a second nucleophile (TMS-N_3) in the reaction was found to increase the reliability of the deprotection (carrying out the deprotection with TMS-N_3 alone is sluggish). Two reiterations of resin splits and mixes with an amino acid coupling and PNA encoding afforded intermediate **61** via **60**. The completions of these reactions were monitored by MALDI analysis of the crude cleavage product at every cycle. The resin was then split into five pools and the respective PNA codons were introduced followed by an arginine residue and a Cy3[6,43] to afford the key intermediate **29**. With the five pools of polymer bound intermediate **29** at hand, there remained to couple the warhead and cleave the libraries from the solid support. Thus, a fraction of each pool was coupled, respectively, to the acrylate **32**, boronate **41**, hydroxamate **42**, amino alcohol **35** and **38**, and dipeptide **47** and **50** each containing a modified tyrosine. Libraries **22**–**24** and **28** were obtained after the coupling by TFA treatment of the resin. For the purpose of libraries **25** and **26**, each pool was subjected to a DMP oxidation prior to TFA cleavage. Library **27** was obtained by reduction of the azide using PMe_3 and direct coupling of the iminophosphorane to di(*N*-succinimidyl)oxylate prior to TFA cleavage. For the coupling of the boronate **41**, it was found to be essential to wash the resin with KCN solution in DMSO to remove residual palladium in order to avoid deboronation during the coupling/cleavage.[44] While the TFA cleavage only leads to partial hydrolysis of the boronate, preliminary screens showed that the boronate was as active as the boronic acids. All the libraries were precipitated in Et_2O, resuspended in H_2O, and lyophilized.

2.5. Screening of libraries

The integrity of the warhead from each library was then validated with a screen of each library against a panel of enzymes representative of the families targeted by the libraries: cathepsins B, a cysteine protease; trypsin, a serine protease; MMP2, a metalloprotease; and PTP 1B, a phosphatase. As shown in Table 3, library **22** with the acrylate warhead show low micromolar inhibition of cathepsin B in agreement with our previous observations,[6] but did not have notable inhibition for the other enzymes. Library **23** with the boronic acid warhead, which can form a reversible covalent bond to the nucleophilic serine residue of a serine protease inhibits trypsin at nanomolar concentration but does not inhibit the other proteases. Library **24** with the hydroxamic functionality, which can coordinate to the metallo-center of the protease inhibits selectively MMP2 at low micromolar. Libraries **25** and **26** are both ketone library, which can react with cysteine or serine protease but library **25** does so reversibly whereas **26** does so irreversibly. Both of these libraries showed potent inhibition of cathepsins B and trypsin. Libraries **27** and **28** did not inhibit any protease despite the acrylate function present in the tyrosine surrogate of library **28**. On the other hand, library **27** showed selective inhibition of PTP 1B. Disappointingly, library **28** did not show significant inhibition in our assay conditions.

Together, these results add further evidence that PNA and the modified PNA used in these libraries do not interfere with biological assays and that the warhead was successfully incorporated in the libraries.

3. Conclusion

The incorporation of modified PNA monomers dramatically increases the solubility of PNA-encoded libraries in biologically relevant buffer without compromising the specificity of hybridization and should widen the scope of such libraries. The divergent strategy developed for the library synthesis allows the incorporation of a wide variety of warhead at the last step as was demonstrated with the synthesis of seven libraries extending the scope of enzyme classes that can be targeted with PNA-encoded libraries. The libraries of inhibitors were shown to be selective for their target enzyme families.

4. Experimental

4.1. General

All solid phase reactions were carried out on an Argonaut Quest 210 or an Applied Biosystems Expedite PNA synthesizer under a nitrogen atmosphere at room temperature with

dry solvents obtained by passing them through commercially available activated alumina columns (PureSolv, Innovative Technology, MA). Substituted polystyrene resins (100–200 mesh, 1% DVB) and amino acids were purchased from Advanced Chemtech or Novabiochem. MALDIs were performed on a Bruker Autoflex TOF. LC-MS were recorded on HP1100/Surveyor MSQ spectrometers.

4.1.1. Library synthesis. Rink amide resin (0.8 mmol/g) was loaded with a substoichiometric amount of *N*-α-Fmoc-*N*-ε-4-methyltrityl-lysine (0.2 mmol/g), DIC (1.0 equiv), and HOBt (1.0 equiv) in DMF (10 mL/g) for 12 h and acetic anhydride (5.0 equiv) with 2,6-lutidine (5.0 equiv) was added to cap the resin (2 h). The resin was distributed into five pools of 20 μmol and, after 30 min swelling in CH_2Cl_2, the resins were deprotected with 20% piperidine in DMF for 5 min. An automated washing procedure was used after each step involving four DMF washes (2 mL/g of resin) with a 30 s agitation (no contracting solvents such as methanol or diethyl ether were used). The resins were then treated with a solution of compound **56** (5.0 equiv) preactivated (5 min) with HOBt (5.0 equiv) and DIC (5.0 equiv) in DMF (6 μL/mg of resin) for 3 h to obtain intermediates **58**. The reaction was repeated to ensure its completion. The pools were then treated with a 2.5% TFA solution containing 5% Et_3SiH three times for 10 min and washed with a 0.1 M solution of Et*i*Pr$_2$N to free base the polymer-bound amine. Each pool was then coupled to Boc-Lys(Fmoc)OH (4.0 equiv) by preactivating the acid with DIC (4.0 equiv) in the presence of HOBt (4.0 equiv) in DMF (0.14 M) for 5 min prior to its addition to the resin. 2,6-Lutidine (6 equiv) was then added and the reaction was allowed to proceed for 3 h. The Fmoc group was removed (20% piperidine in DMF for 5 min) followed by coupling to FmocArg(Pbf)OH using the same condition. Each pool were then analyzed by cleavage of an analytical sample (0.1–0.3 mg) using 4:1 TFA/*m*-cresol for 1 h followed by pelleting from Et_2O. The product from each pool showed a single peak by LC-MS with the expected mass (see Supplementary data for spectra). Each pool was encoded according to the general procedure below to obtain polymer bound intermediate **59**. Cleavage of an analytical sample from each pool showed the desired compound as a single peak without truncated sequences. The pools were mixed and the allyl group was removed according to the general procedure below. The resin was redistributed into five new pools for a second cycle involving an amino acid coupling and encoding according to the general procedures below afforded resins **60**. Cleavage of an analytical sample from each pool showed the desired compounds as the five expected product without truncated sequences. The third cycle of mix and split afforded five pools of **61**. Cleavage of an analytical sample from each pool showed the desired compounds (25 per pool) as the expected range of masses without significant truncated sequences due to incomplete reactions. The library was mixed and redistributed in five pools to introduce the last PNA codon, remove the allyl, and to introduce the fluorophore (Cy3 (5.0 equiv) was preactivated with TNTU (4.5 equiv) in the presence of 2,6-lutidine (10 equiv) to form the NHS-ester for 1 h in NMP and the mixture was added to the resin and agitated for 2×3 h) thus yielding **29**. Cleavage of an analytical sample from each pool showed the desired range of molecular weights (125 compounds per pool (it should be noted that the presence of the Cy3-fluorophore greatly reduce the ionization of the library). A tenth of each pool (2 μmol) was then coupled the respective warheads (**32**, **41**, **42**, **35**, **38**, **50**, and **47**) according to the general procedure for amino acid coupling. For **35** and **38**, the coupling was followed by a treatment with DMP (10 equiv) in CH_2Cl_2 for 1 h. For library **27**, the coupling was followed by treatment with PMe_3 (10 equiv) for 10 min, filtration and addition of di(*N*-succinimidyl)oxalate (10 equiv) in wet DMF (1% water) for 2 h. Cleavage of the library using a 4:1 mixture of TFA with *m*-cresol for 2 h (500 μL) followed by precipitation in Et_2O (5 mL) and centrifugation afforded a pink pellet characteristic of the Cy3 color. The pellets were washed with Et_2O, resuspended in water and lyophilized to afford 5–6 mg for each pool (expected ca. 12 mg). MALDI analysis of each pool showed a shift in the molecular weight range corresponding to the weight of the warhead. While this shift in molecular weight can not be used to asses the productivity of the DMP oxidation, the selective inhibition of cathepsins B and trypsin with both these libraries can only come from the ketone product.

4.2. General procedure for PNA encoding

The resin was treated with 20% piperidine in DMF for 5 min, washed according to the aforementioned procedure. During the course of the deprotection, the PNA monomer **11** (5.0 equiv) was preactivated for 5 min using Et*i*Pr$_2$N (5.0 equiv), HBTU (4.4 equiv), and 2,6-lutidine (7.5 equiv) in NMP (0.17 M in PNA monomer) for 1 h. This solution was added to the deported resin and the reaction mixture was agitated for 1 h. The resin was filtered without washing and acetic anhydride (5.0 equiv) followed by 2,6-lutidine (5.0 equiv) in DMF (0.17 M) was added. The resin was agitated for 5 min and then washed according to the aforementioned washing protocol.

4.3. General procedure for allyl deprotection

The resin was treated with a solution of $Pd(PPh_3)_2Cl_2$ (0.2 equiv), Ph_3P (0.8 equiv), and TMS-N_3 (10 equiv) as a CH_2Cl_2 solution (0.1 M) followed by a solution of *n*-Bu_3SnH (5.0 equiv) in CH_2Cl_2 (1 M) for 15 min. The reaction was repeated twice and then washed five times with CH_2Cl_2 prior to the aforementioned washing procedure. Prior to the introduction of the boronates, it was found to be important to wash the resin with a saturated solution of KCN in DMSO (3×5 min) to remove any precipitated palladium.

4.4. General procedure for amino acid coupling

The polymer bound acid was preactivated with DIC (4.5 equiv) and HOBt (4.5 equiv) in DMF (0.15 M) for 5 min. The amine (5.0 equiv) was then added with 2,6-lutidine (10 equiv) as a DMF solution (1 M for the amine) and the reaction was continued for 10 h.

Acknowledgements

We thank Human Frontier Science Program for their support of this work (HFSP 0080/2003). The Région d'Alsace and

Ministry of Science are gratefully acknowledged for fellowships to F.D., Z.P. and J.D. The authors thank Laurence Oswald for her technical assistance.

Supplementary data

Procedures and analytical data for the preparation of modified PNA **1–21** and **51–52** as well as amino acids **3–50**. LC-MS and MALDI analysis of libraries **22–29** and their synthetic precursors **56–61**. Hybridization data and procedures for the biological assays. Supplementary data associated with this article can be found in the online version, at doi:10.1016/j.tet.2007.03.033.

References and notes

1. Winssinger, N.; Ficarro, S.; Schultz, P. G.; Harris, J. L. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11139–11144.
2. Debaene, F.; Mejias, L.; Harris, J. L.; Winssinger, N. *Tetrahedron* **2004**, *60*, 8677–8690.
3. Harris, J.; Mason, D. E.; Li, J.; Burdick, K. W.; Backes, B. J.; Chen, T.; Shipway, A.; Van Heeke, G.; Gough, L.; Ghaemmaghami, A.; Shakib, F.; Debaene, F.; Winssinger, N. *Chem. Biol.* **2004**, *11*, 1361–1372.
4. Winssinger, N.; Damoiseaux, R.; Tully, D. C.; Geierstanger, B. H.; Burdick, K.; Harris, J. L. *Chem. Biol.* **2004**, *11*, 1351–1360.
5. Harris, J. L.; Winssinger, N. *Chemistry* **2005**, *11*, 6792–6801.
6. Urbina, H. D.; Debaene, F.; Jost, B.; Bole-Feysot, C.; Mason, D. E.; Kuzmic, P.; Harris, J. L.; Winssinger, N. *Chembiochem* **2006**, *7*, 1790–1797.
7. Sano, S.; Tomizaki, K. Y.; Usui, K.; Mihara, H. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 503–506.
8. Diaz-Mochon, J. J.; Bialy, L.; Bradley, M. *Chem. Commun.* **2006**, 3984–3986.
9. Diaz-Mochon, J. J.; Bialy, L.; Keinicke, L.; Bradley, M. *Chem. Commun.* **2005**, 1384–1386.
10. Rajeev, K. G.; Maier, M. A.; Lesnik, E. A.; Manoharan, M. *Org. Lett.* **2002**, *4*, 4395–4398.
11. Ausin, C.; Ortega, J.-A.; Robles, J.; Grandas, A.; Pedroso, E. *Org. Lett.* **2002**, *4*, 4073–4075.
12. Miyata, K.; Tamamushi, R.; Ohkubo, A.; Taguchi, H.; Seio, K.; Santa, T.; Sekine, M. *Org. Lett.* **2006**, *8*, 1545–1548.
13. Dutta, S. P.; Chheda, G. B. *J. Carbohydr. Nucleos. Nucleot.* **1980**, *7*, 217–240.
14. Sforza, S.; Corradini, R.; Ghirardi, S.; Dossena, A.; Marchelli, R. *Eur. J. Org. Chem.* **2000**, 2905–2913.
15. Haaima, G.; Lohse, A.; Buchardt, O.; Nielsen, P. E. *Angew. Chem., Int. Ed.* **1996**, *35*, 1939–1941.
16. Englund, E. A.; Appella, D. H. *Org. Lett.* **2005**, *7*, 3465–3467.
17. Wu, Y.; Xu, J. C. *Tetrahedron* **2001**, *57*, 8107–8113.
18. Menchise, V.; De Simone, G.; Tedeschi, T.; Corradini, R.; Sforza, S.; Marchelli, R.; Capasso, D.; Saviano, M.; Pedone, C. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12021–12026.
19. Debaene, F.; Winssinger, N. *Org. Lett.* **2003**, *5*, 4445–4447.
20. Bolin, D. R.; Sytwu, I. I.; Humiec, F.; Meienhofer, J. *Int. J. Pept. Protein Res.* **1989**, *33*, 353–359.
21. Hernandez, A. A.; Roush, W. R. *Curr. Opin. Chem. Biol.* **2002**, *6*, 459–465.
22. Powers, J. C.; Asgian, J. L.; Ekici, O. D.; James, K. E. *Chem. Rev.* **2002**, *102*, 4639–4750.
23. Kettner, C. A.; Shenvi, A. B. *J. Biol. Chem.* **1984**, *259*, 15106–15114.
24. Adams, J. *Curr. Opin. Chem. Biol.* **2002**, *6*, 493–500.
25. Whittaker, M.; Floyd, C. D.; Brown, P.; Gearing, A. J. *Chem. Rev.* **1999**, *99*, 2735–2776.
26. Edwards, P. D.; Meyer, E. F., Jr.; Vijayalakshmi, J.; Tuthill, P. A.; Andisik, D. A.; Gomes, B.; Strimpler, A. *J. Am. Chem. Soc.* **1992**, *114*, 1854–1863.
27. Maryanoff, B. E. *J. Med. Chem.* **2004**, *47*, 769–787.
28. Otto, H.-H.; Schirmeister, T. *Chem. Rev.* **1997**, *97*, 133–171.
29. Bialy, L.; Waldmann, H. *Angew. Chem., Int. Ed.* **2005**, *44*, 3814–3839.
30. Craig, M.; Luu, H. A.; McCready, T. L.; Williams, D.; Andersen, R. J.; Holmes, C. F. B. *Biochem. Cell Biol.* **1996**, *74*, 569–578.
31. Cao, X.; Moran, E. J.; Siev, D.; Lio, A.; Ohashi, C.; Majalli, A. M. M. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 2953–2958.
32. Moran, E. J.; Sarshar, S.; Cargill, J. F.; Shahbaz, M. M.; Lio, A.; Mjalli, A. M. M.; Armstrong, R. W. *J. Am. Chem. Soc.* **1995**, *117*, 10787–10788.
33. Xin, Z.; Liu, G.; Abad-Zapatero, C.; Pei, Z.; Szczepankiewicz, B. G.; Li, X.; Zhang, T.; Hutchins, C. W.; Hajduk, P. J.; Ballaron, S. J.; Stashko, M. A.; Lubben, T. H.; Trevillyan, J. M.; Jirousek, M. R. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 3947–3950.
34. Beaulieu, P. L.; Cameron, D. R.; Ferland, J.-M.; Gauthier, J.; Ghiro, E.; Gillard, J.; Gorys, V.; Poirier, M.; Rancourt, J.; Wernic, D.; Llinas-Brunet, M.; Betageri, R.; Cardozo, M.; Hickey, E. R.; Ingraham, R.; Jakes, S.; Kabcenell, A.; Kirrane, T.; Lukas, S.; Patel, U.; Proudfoot, J.; Sharma, R.; Tong, L.; Moss, N. *J. Med. Chem.* **1999**, *42*, 1757–1766.
35. Edwards, P. D.; Wolanin, D. J.; Andisik, D. W.; Davis, M. W. *J. Med. Chem.* **1995**, *38*, 76–85.
36. Costanzo, M. J.; Almond, H. R., Jr.; Hecker, L. R.; Schott, M. R.; Yabut, S. C.; Zhang, H.-C.; Andrade-Gordon, P.; Corcoran, T. W.; Giardino, E. C.; Kauffman, J. A.; Lewis, J. M.; de Garavilla, L.; Haertlein, B. J.; Maryanoff, B. E. *J. Med. Chem.* **2005**, *48*, 1984–2008.
37. Wood Warren, J. L.; Huang, L.; Ellman Jonathan, A. *J. Comb. Chem.* **2003**, *5*, 869–880.
38. Wityak, J.; Earl, R. A.; Abelman, M. M.; Bethel, Y. B.; Fisher, B. N.; Kauffman, G. S.; Kettner, C. A.; Ma, P.; McMillan, J. L., et al. *J. Org. Chem.* **1995**, *60*, 3717–3722.
39. Hanessian, S.; Moitessier, N.; Gauchet, C.; Viau, M. *J. Med. Chem.* **2001**, *44*, 3066–3073.
40. Nicolaou, K. C.; Estrada, A. A.; Zak, M.; Lee, S. H.; Safina, B. S. *Angew. Chem., Int. Ed.* **2005**, *44*, 1378–1382.
41. Waldmann, H.; Kunz, H. *Liebigs Ann. Chem.* **1983**, 1712–1725.
42. Page, P.; Bradley, M.; Walters, I.; Teague, S. *J. Org. Chem.* **1999**, *64*, 794–799.
43. Korbel, G. A.; Lalic, G.; Shair, M. D. *J. Am. Chem. Soc.* **2001**, *123*, 361–362.
44. Tian, Z.-Q.; Brown, B. B.; Mack, D. P.; Hutton, C. A.; Bartlett, P. A. *J. Org. Chem.* **1997**, *62*, 514–522.