JOC The Journal of Organic Chemistry

Article

Synthesis and RNA binding properties of extended nucleobases for triplex-forming peptide nucleic acids

Ilze Kumpina, Nikita Brodyagin, James Arthur MacKay, Scott D Kennedy, Martins Katkevics, and Eriks Rozners

J. Org. Chem., **Just Accepted Manuscript** • DOI: 10.1021/acs.joc.9b01133 • Publication Date (Web): 20 Sep 2019

Just Accepted

is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Synthesis and RNA binding properties of extended nucleobases for triplex-forming peptide nucleic acids

Ilze Kumpina,[1] Nikita Brodyagin,[2] James A. MacKay,[3] Scott D. Kennedy,[4] Martins Katkevics,[1] Eriks Rozners[2,*]

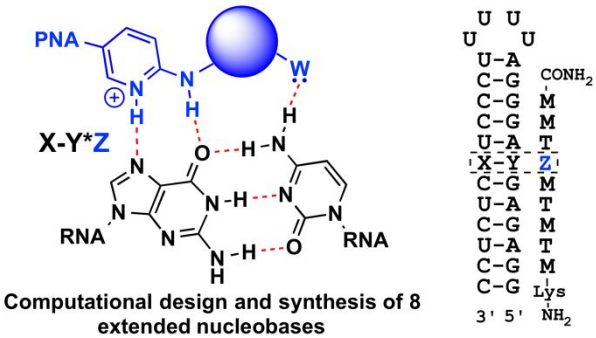
[1]Latvian Institute of Organic Synthesis, Aizkraukles 21, Riga, LV-1006, Latvia

[2]Department of Chemistry, Binghamton University, State University of New York, Binghamton, NY 13902, USA

[3]Department of Chemistry and Biochemistry, Elizabethtown College, Elizabethtown, PA 17022, USA

[4]Department of Biochemistry and Biophysics, University of Rochester School of Medicine and Dentistry, Rochester, NY 14642, USA

* e-mail erozners@binghamton.edu



Computational design and synthesis of 8 extended nucleobases

Abstract: Triple helix formation, using Hoogsteen hydrogen bonding of triplex-forming oligonucleotides, represents an attractive method for sequence-specific recognition of double-stranded nucleic acids. However, practical applications using triple helix-forming oligonucleotides and their analogues are limited to long homopurine sequences. The key problem for recognition of pyrimidines is that they present only one hydrogen bond acceptor or donor group in the major groove. Herein, we report our first attempt to overcome this problem by using peptide nucleic acids (PNAs) modified with extended nucleobases that form three hydrogen bonds along the entire Hoogsteen edge of the Watson-Crick base pair. Five new nucleobase triples were designed and their hydrogen bonding feasibility was confirmed by *ab initio* calculations. PNA monomers carrying the modified nucleobases were synthesized and incorporated in short model PNA sequences. Isothermal titration calorimetry (ITC) showed that these nucleobases had modest binding affinity for their double-stranded RNA (dsRNA) targets. Finally, molecular modeling of the modified triples in PNA-dsRNA helix suggested that the modest binding affinity was caused by subtle structural deviations from ideal hydrogen bonding arrangements or disrupted pi-stacking of the extended nucleobase scaffolds.

INTRODUCTION

Triple helices are naturally occurring structures in RNA.[1] Additionally, triple helical binding of an oligonucleotide in the major groove of double-stranded RNA (dsRNA) may be used for sequence-selective molecular recognition of folded RNAs.[1, 2] However, triple helices formed by negatively charged oligonucleotides are destabilized by electrostatic repulsion. The Rozners group[3] were the first to use peptide nucleic acid (PNA),[4, 5] a DNA analogue built on neutral amide backbone (Figure 1), to overcome the unfavorable electrostatic repulsion in triple helical recognition of dsRNA. Other groups have also explored thermodynamics and kinetics of PNA-dsRNA triplex formation[6] and developed chemical modifications to improve PNA binding to dsRNA.[7] Recently, the Rozners group discovered that PNAs with cytosines (C) replaced by 2-aminipyridine (M) nucleobases formed exceptionally stable and sequence-selective triple helices with dsRNA under physiologically relevant conditions.[8-10] The key innovative feature was the increased basicity of M (pK_a of M^+ ~6.5) compared to C (pK_a of C^+ ~4.5) that enabled formation of protonated M*G-C triple (Figure 1) at physiological pH. This discovery opened the door for using PNA-dsRNA triplexes for developing novel fluorescent RNA detection methods[11] and modulating RNA function in biological systems.[12] Taken together, these studies identified M-modified PNA as a promising ligand for molecular recognition of complex dsRNAs.

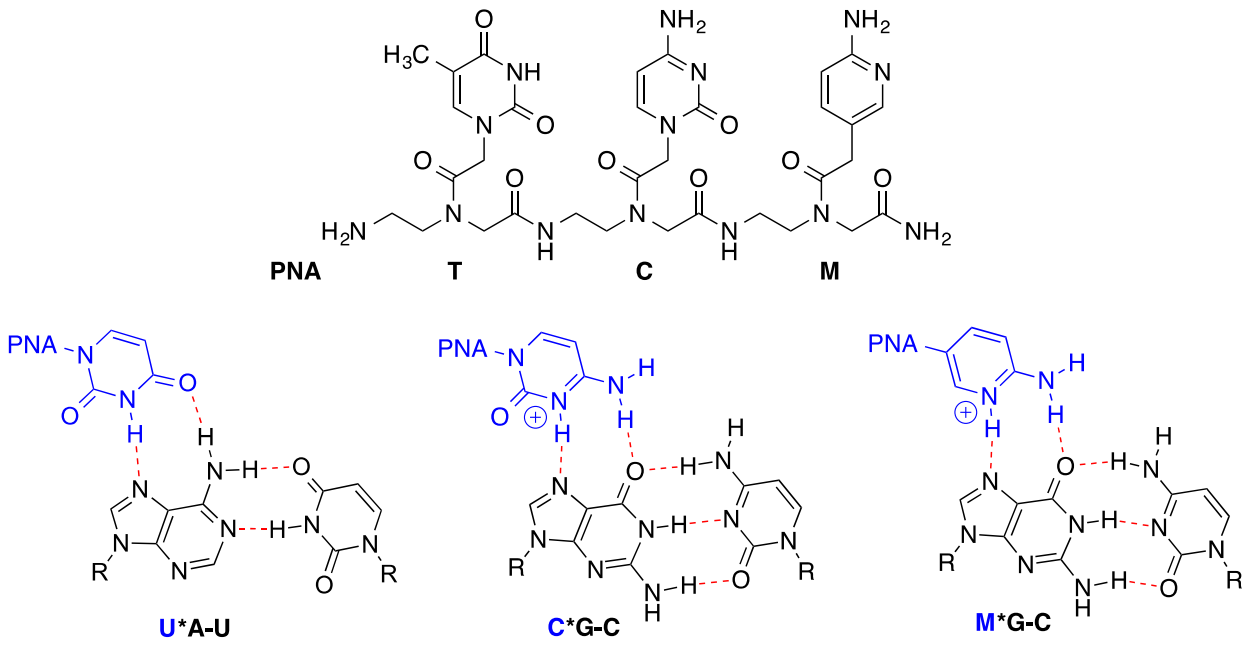


Figure 1. Structure of PNA and Hoogsteen base triples, U*A-U, C*G-C and M*G-C.

A remaining general problem for triple helical recognition of nucleic acids is that the common Hoogsteen hydrogen-bonded triples (e.g., U*A-U and C*G-C or M*G-C, Figure 1) limit the sequences that can be targeted to motifs where one of the strands of double helix is formed by purine nucleosides only. Despite extensive effort in developing modified nucleobases that would recognize pyrimidines in U-A (or T-A in DNA) and C-G base pairs, triple helix formation is still limited to relatively long homopurine sequences that may be interrupted by only a few pyrimidines.[13-17] The problem of designing high stability triples for recognition of U (or T) and C is that they present only one hydrogen bond acceptor or donor group, respectively, in the major groove. To overcome this problem, we envisioned designing extended nucleobases that would form three hydrogen bonds along the entire Hoogsteen edge of the Watson-Crick base pair.

The idea of extending the Hoogsteen hydrogen bonding over both nucleosides of a Watson-Crick base pair has been entertained since the early days of triple-helical

DNA recognition. Dervan and co-workers[18] pioneered the first design, where a nucleobase analogue (**D_3**, Figure 2) spanned the major grove to form hydrogen bonds with both C and G of the inverted C-G base pair. Other groups followed up with similar designs.[19-22] These early efforts shared common problems of reduced affinity (compared to the native triples) and lack of discrimination between the pyrimidines of C-G and T-A base pairs.[14] NMR structures later showed that **D_3** did not form the expected hydrogen bonds, but instead intercalated in the double helix and stacked over the pyrimidine-purine base pairs.[22-24] Other studies[25-27] used NMR in organic solvent to evaluate hydrogen bonding potential and design novel extended nucleobases, e.g., ureido isoindolin-1-one (**Ind**, Figure 2). However, strong hydrogen bonding in organic solvent did not necessarily translate into formation of strong and selective DNA triple helices suggesting that other factors, such as, constrains on backbone conformation, base triple isomorphism and nucleobase stacking, played subtle but important roles in overall efficiency of triple helical binding.[27, 28]

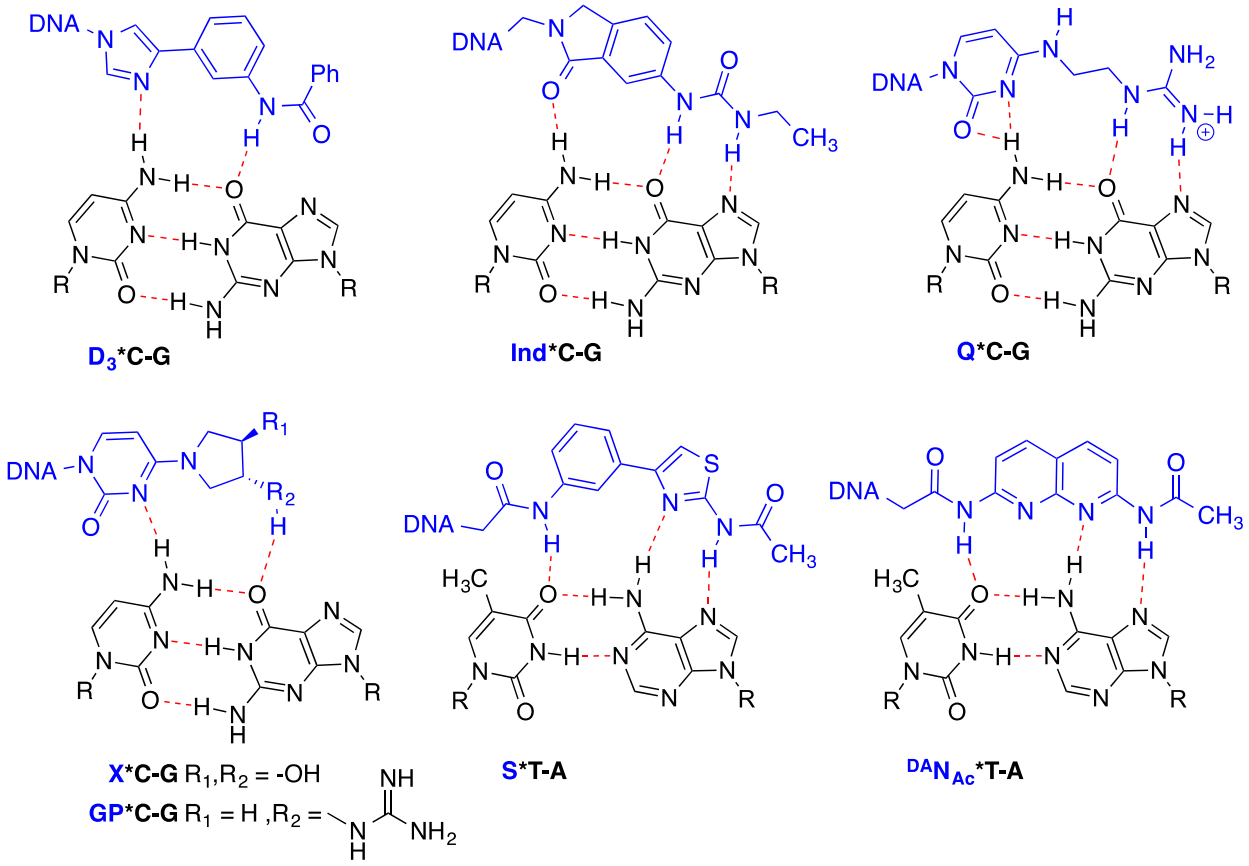


Figure 2. Selected examples of nucleobases that hydrogen bond to entire Hoogsteen edge of Watson-Crick base pairs. In specific studies, the nucleobases were attached to either native DNA, or various backbone modified analogues (e.g., LNA, PNA, etc).

In more recent studies, cytosine derivatives *N*[4]-(2-guanidoethyl)-5-methylcytosine **Q**,[29] 4-[(3*R*,4*R*)-dihydroxypyrrolidino]-pyrimidin-2-one **X**[30] and the related guanidine derivative **GP**[31] (Figure 2), and *N*-(4-(3-acetamidophenyl)thiazol-2-yl)acetamide[32] (**S**, Figure 2) showed promising binding to C-G and T-A base pairs, respectively. Overall, while several nucleobases have given encouraging preliminary results, current designs are far from optimal and typically suffer from either modest binding affinity (e.g., **Q**) or modest sequence selectivity (e.g., **S**). In attempted triple helix formation involving all four base pairs, Fox and co-workers[33, 34] used **S** for T-A recognition, while Ohkubo et

al.[35] used N-acetyl-2,7-diamino-1,8-naphtyridine($^{DA}\mathbf{N}_{Ac}$, Figure 2) and **Q** for T-A and C-G recognition, respectively. Chen and co-workers[36] used **Q**-modified PNAs to form a triple helix on a seven purine long stretch of dsRNA interrupted by one C, and showed that the triplex formation stimulated ribosomal frameshifting in a cell-free translation assay.[37] Despite the promising results of these studies, a general solution to recognition of any sequence of dsDNA or dsRNA is still elusive and design of novel nucleobases remains an active area of research.[17, 38]

The present manuscript describes our first efforts in designing extended heterocyclic systems that would recognize the entire Hoogsteen edge of Watson-Crick base pairs in the major groove of dsRNA. Electronic and geometric considerations discussed above influenced our strategy for designing extended nucleobases. While we used *ab initio* calculations of nucleobase triples as simple models for initial assessment of the hydrogen bonding feasibility, the ultimate test was the experimentally determined affinity of the modified PNAs for their matched target dsRNA hairpins.

We report syntheses of five new scaffolds of extended nucleobases (**M1** through **M4** and **T1**, Figure 3). Isothermal titration calorimetry (ITC) showed that these nucleobases had modest binding affinity for their dsRNA targets. Finally, molecular modeling of the modified triples in a PNA-dsRNA helix provided additional insights into structural reasons behind the observed affinities suggesting that the loss of binding affinity compared to the canonical Hoogsteen triples was due to subtle structural deviations from ideal hydrogen bonding arrangements or disrupted pi-stacking of the extended nucleobase scaffolds.

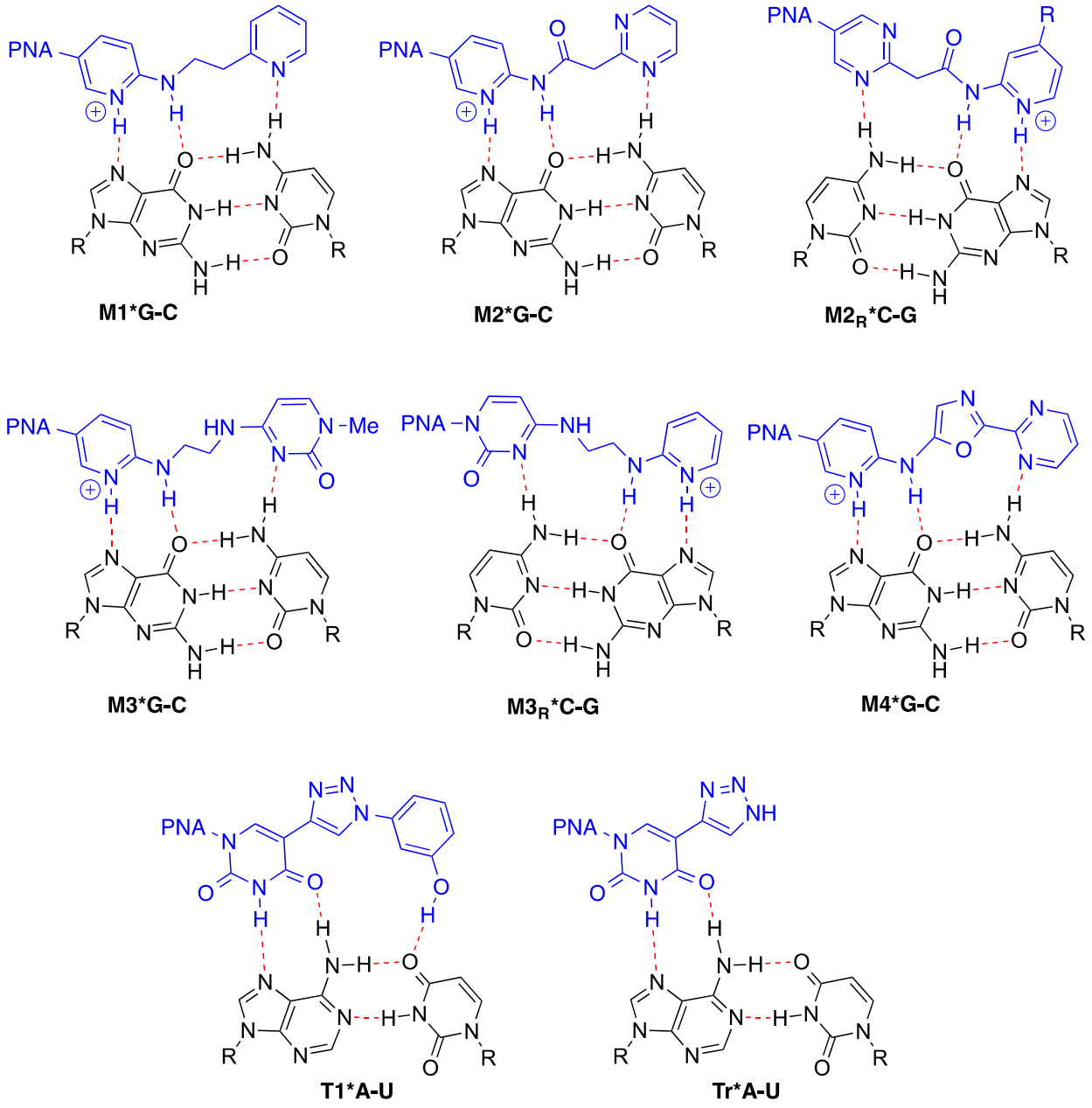


Figure 3. Extended nucleobases and proposed hydrogen bonding schemes for recognition of the entire Hoogsteen edge of Watson-Crick base pairs.

RESULTS AND DISCUSSION

Overall design. Our design of extended nucleobases (Figure 3) was to extend M and T using an appropriate linker that would position a single hydrogen bond acceptor or donor across the major grove for additional interaction with C and U, respectively. We anticipated the identification of an ideal scaffold that could ultimately be used for either G-C or C-G recognition depending on the attachment point to PNA backbone (c.f., **M2***G-C and **M2R***C-G in Figure 3). The idea was that a successful extended nucleobase would retain and improve the excellent binding properties of M and T observed in our previous studies.[8-10]

Toward this end, we first envisioned nucleobase **M1**, linking the M scaffold to a pyridine through a simple ethylene linker (Figure 3). To test the geometrical viability of our design, we used docking of the extended nucleobases on the G-C or A-U base pairs followed by *ab initio* methods (HF-631G(d)) to confirm the intended H-bonding geometries and conformational preferences of the triples. Models of the **M1** nucleobase containing a methyl group in the location of attachment to the PNA backbone were built in silico. We performed a geometry optimization on the G-C base pair, docked it with a geometry optimized **M1** nucleobase, and finally re-ran the geometry optimization on the base triple. We were pleased to find that the structure of the energy minimized triple showed H-bonding distances in the typical range and agreement with our expectations (Figure 4 **M1**). Furthermore, the minimized geometry of the **M1** triple suggested that planarity across the entire extended system was attainable and the triple energy was lower than the sum of the **M1** nucleobase and G-C energies. These considerations together suggested that this strategy may prove useful in identifying privileged scaffolds.

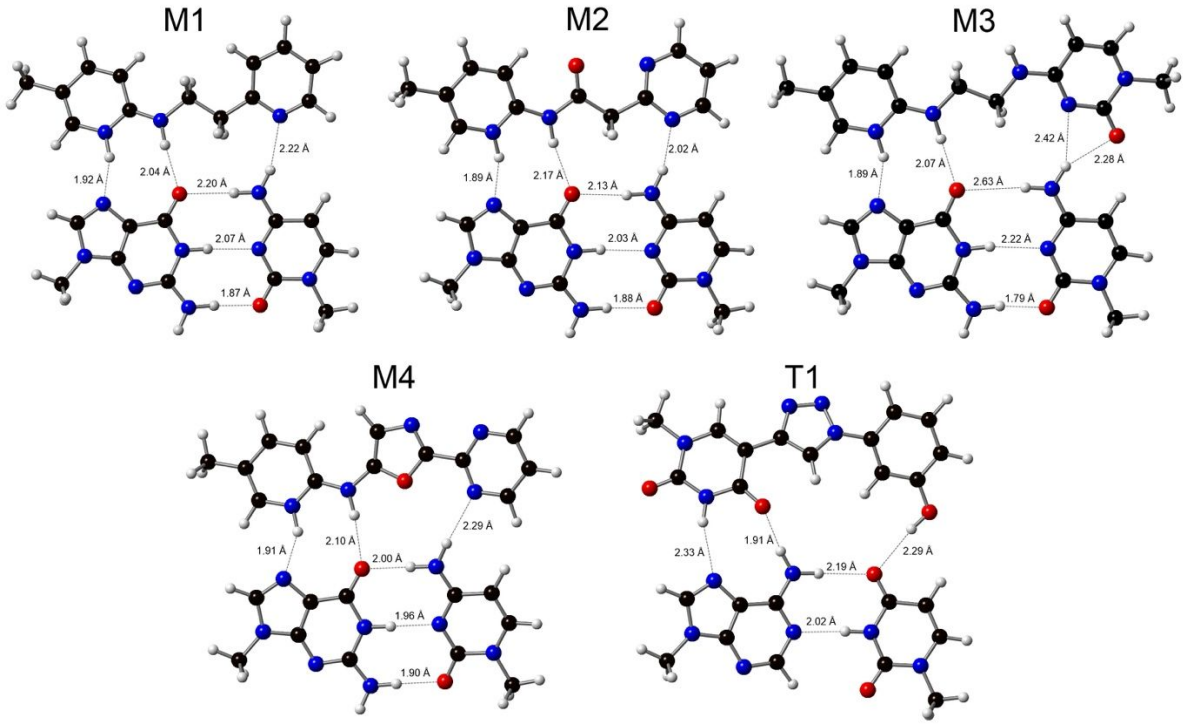


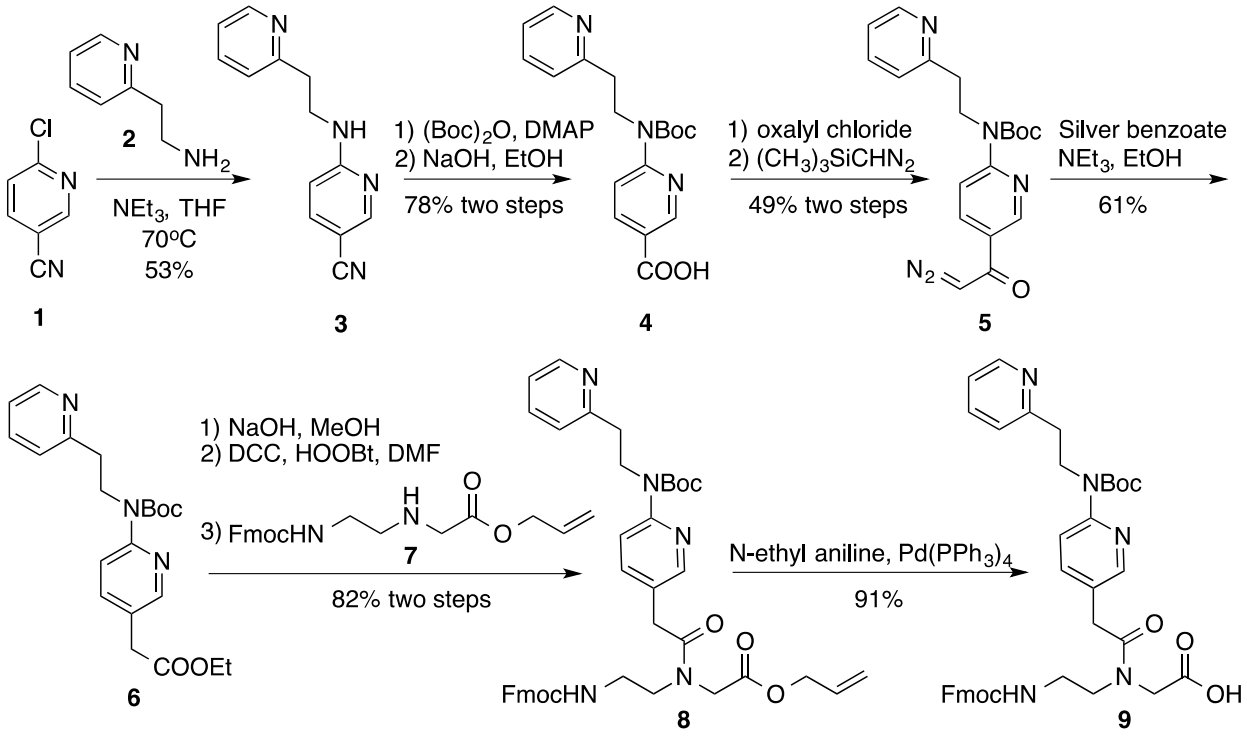
Figure 4. Molecular modeling of triples (M1) **M1*G-C** geometry optimized using HF-631g(d) and (M2-T1) views of initial docking of extended nucleobases on G-C and U-A base pairs.

A similar strategy was used to identify scaffolds **M2**-**M4** and **T1**. In each of these examples the docking of the nucleobase with its respective Watson-Crick base pair resulted in H-bonding geometries within expected distances of 2.5 Å or better (Figure 4). However, for each of these scaffolds, geometry optimized base triples led to unexpected deviations of planarity of either the nucleobase itself (**M2**-**M4**) or the nucleobase coming out of the plane of the Watson-Crick base pair (**T1**) (see the Supporting Information, Figure SL for details). Returning to the docking geometries in Figure 4, molecular energies for these arrangements predicted favorable interactions in

all cases. We reasoned that the ultimate test would be in the experimental binding affinity of the PNA for dsRNA given difficult to predict electronic effects such as H-bonding in aqueous solution, the added complexity of the PNA triplex, and pi-stacking interactions. Thus, we set out to prepare these extended nucleobases with the aim of evaluating their binding properties in PNA-dsRNA triple helix.

Synthesis of M1 PNA monomer. The synthesis of the **M1** PNA monomer started with nucleophilic aromatic substitution of commercially available 6-chloronicotinonitrile **1** with 2-(pyridin-2-yl)ethan-1-amine **2** (Scheme 1). Protection of the secondary amine with di-*tert*-butyl dicarbonate followed by saponification of the nitrile gave the carboxylic acid **4**. Arndt-Eistert homologation of **4** using thionyl chloride to make the requisite acyl chloride gave a complex mixture that did not contain the desired product; however, oxalyl chloride in the presence of catalytic amount of DMF yielded the acyl chloride that was *in situ* converted to diazoketone **5** using TMS-diazomethane. Completion of Arndt-Eistert homologation was achieved with concomitant esterification with ethanol. The esterification was preferred over a more direct route to carboxylic acid, because the ester was much easier to isolate and purify. Saponification of the ethyl ester **6** gave the carboxylic acid, which was coupled to the PNA backbone **7** to give after deallylation the final **M1** PNA monomer **9**.[39]

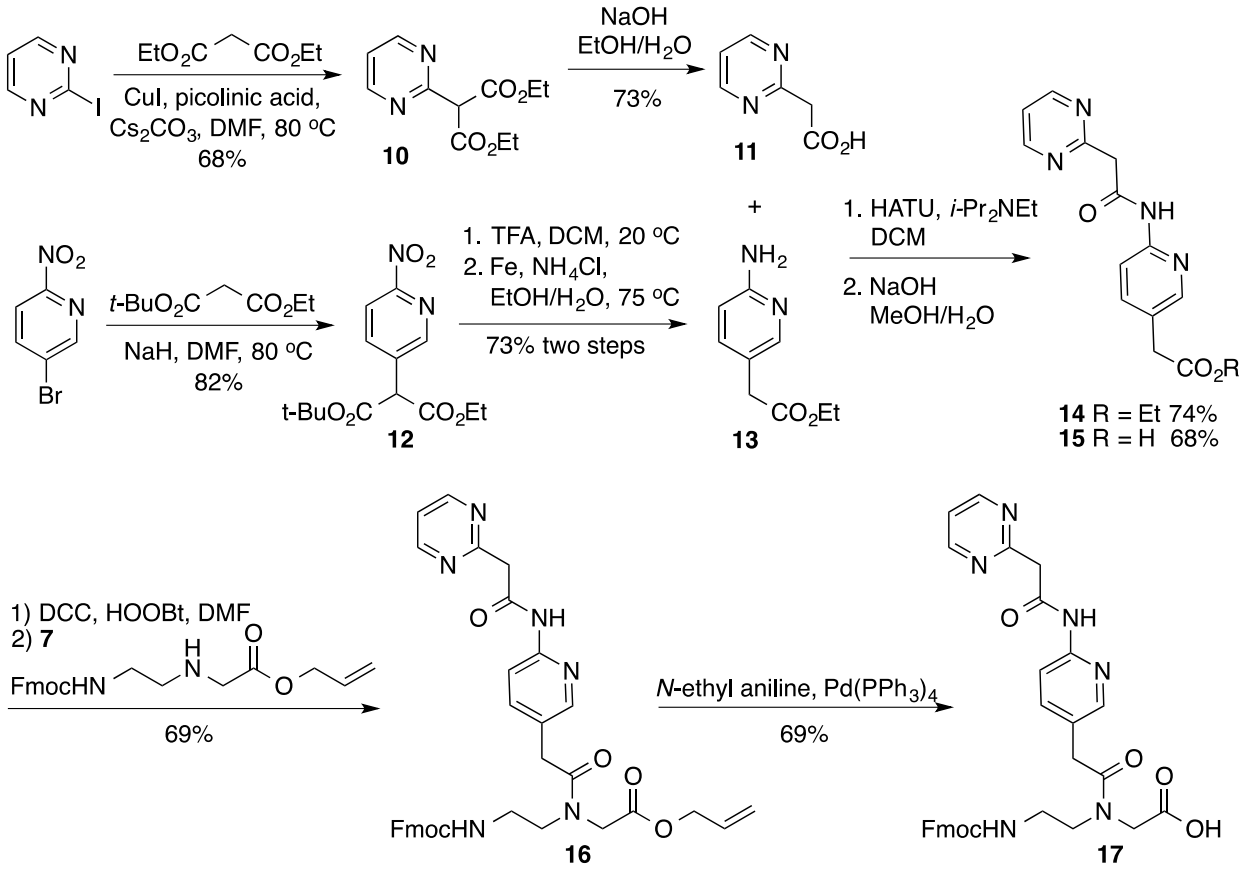
Scheme 1. Synthesis of M1 PNA monomer



Synthesis of M2 PNA monomer. Nucleobases **M2**, **M2R** and **^{MeO}M2R** were variations of **M1** where the one of the methylene groups was replaced by a carbonyl group. The key consideration in their design was the synthetic accessibility of amide linkers that would allow easier optimization of extended nucleobases. A convergent synthesis (Scheme 2) started with coupling of halogenated heterocycles with malonate derivatives. The copper-catalyzed coupling[40] of 2-iodopyrimidine with diethylmalonate gave a better yield of **10** than nucleophilic aromatic substitution used for the synthesis of **12**. Basic hydrolysis of **10** afforded upon acidification 2-(pyrimidin-2-yl) acetic acid **11**. Using a similar approach, ethyl aminopyridyl acetate **13** was made by a nucleophilic aromatic substitution reaction of 4-bromo-2-nitropyridine and *tert*-butyl ethyl malonate, with subsequent cleavage of *tert*-butyl group, decarboxylation and reduction of the nitro function. HATU mediated coupling of **11** with aminopyridyl derivative **13**, followed by

hydrolysis furnished the acetic acid derivative **15**, which was converted to **M2** PNA monomer **17** using the same chemistry as in Scheme 1.

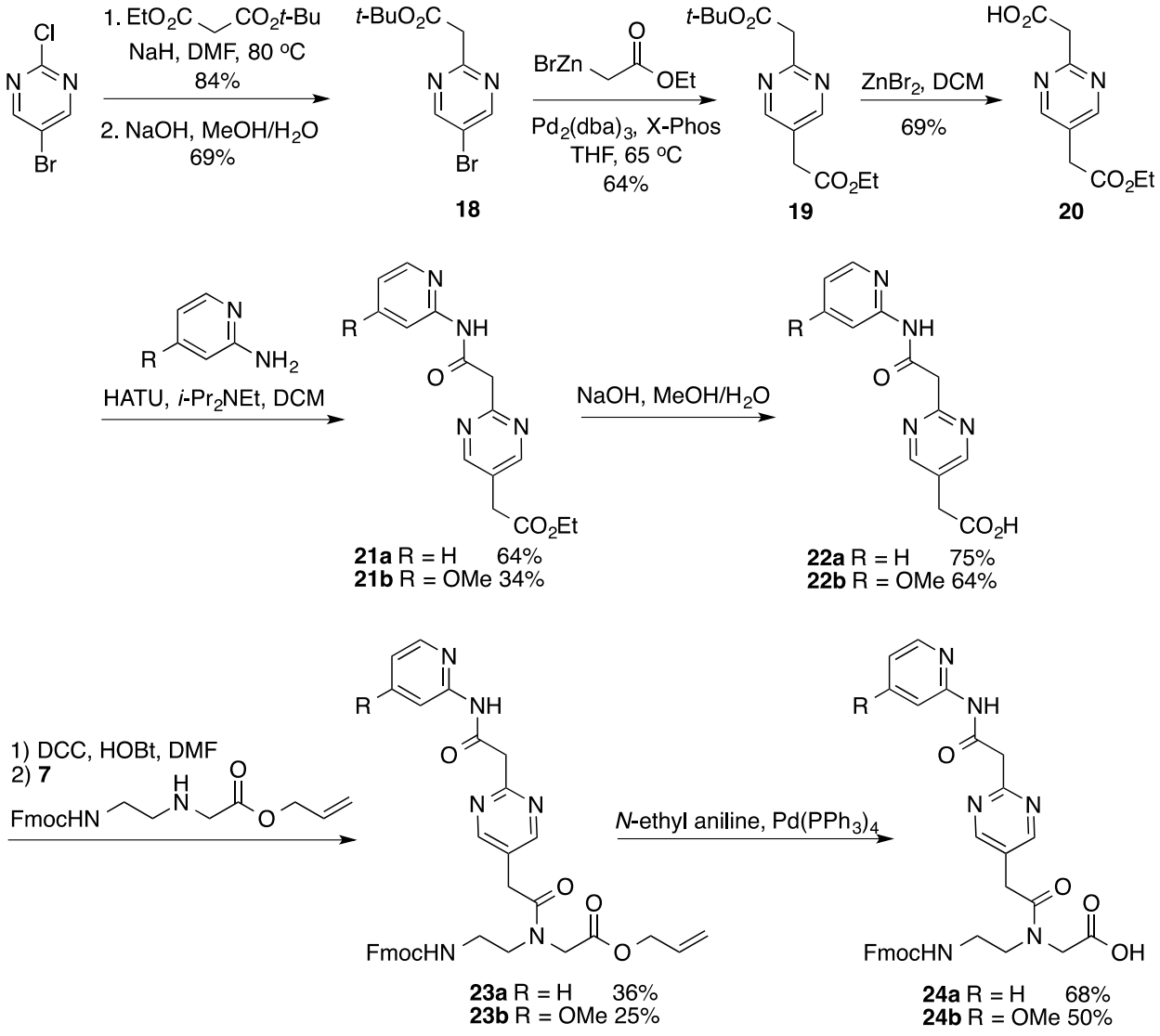
Scheme 2. Synthesis of M2 PNA monomer



Synthesis of M2R and ^{MeO}M2R. The synthesis of orthogonally protected pyrimidyl diester **19**, the key intermediate for acids **22a** and **22b**, started from 5-bromo-3-chloropyrimidine (Scheme 3). The first acetate moiety was introduced following an exact literature precedent,[41] while for the second we adopted a variant of Reformatsky-Negishi reaction.[42] Attempts to use TFA in CH_2Cl_2 for removal of the *tert*-butyl group in **19** resulted in decomposition of **20** after evaporation of the reaction mixture even at room temperature. On the other hand, $ZnBr_2$ mediated cleavage of the *tert*-butyl group[43]

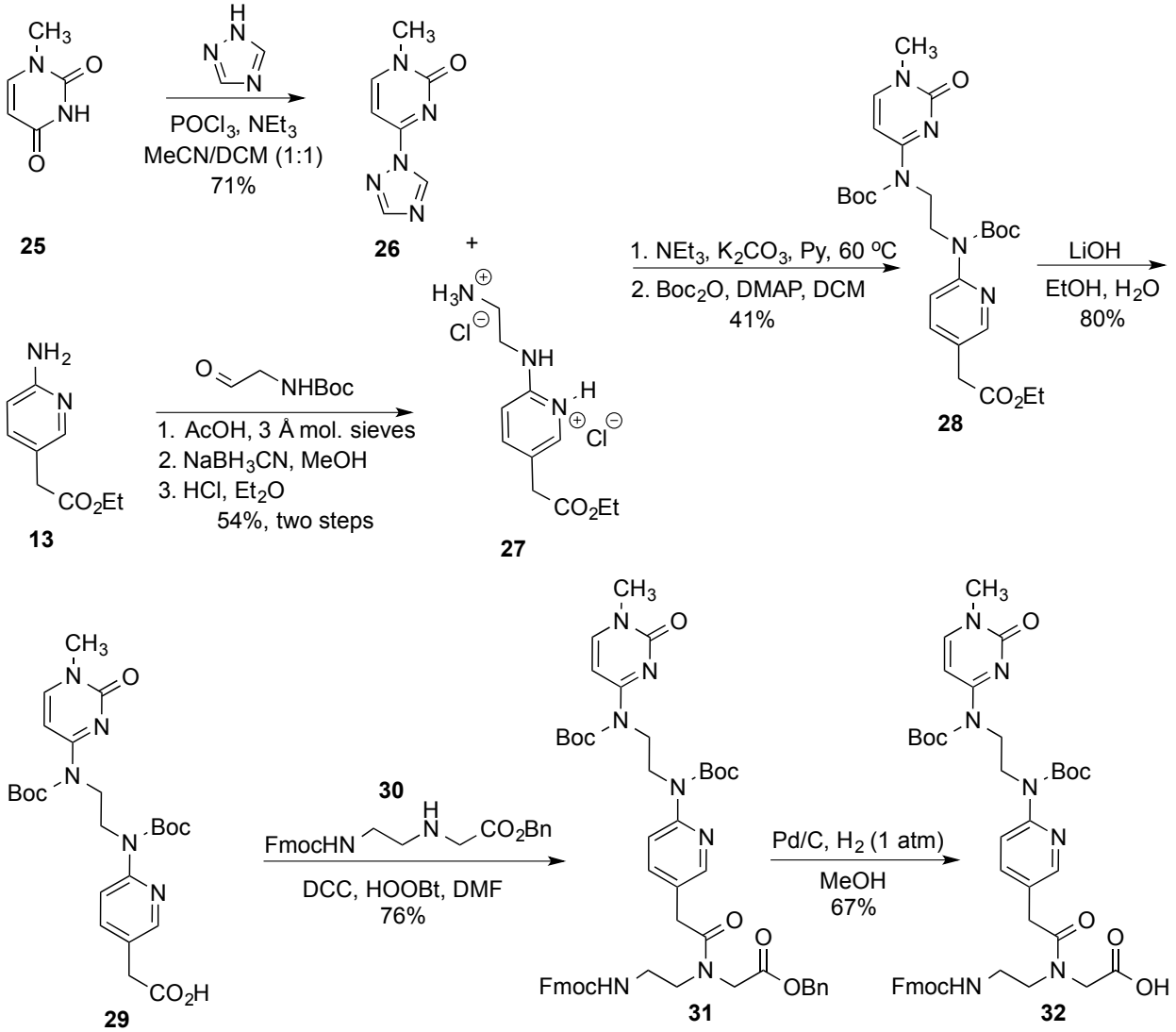
afforded pyrimidylacetic acid **20** in good yield. HATU mediated coupling of **20** with 2-aminopyridine or 2-amino-4-methoxypyridine, followed by basic hydrolysis of ethyl esters provided the acetic acid derivatives **21a** and **21b**, which were converted to **M2R** and ^{MeO}**M2R** PNA monomers **24a** and **24b**, respectively, using the same chemistry as in Schemes 1 and 2.

Scheme 3. Synthesis of M2R and ^{MeO}M2R PNA monomers



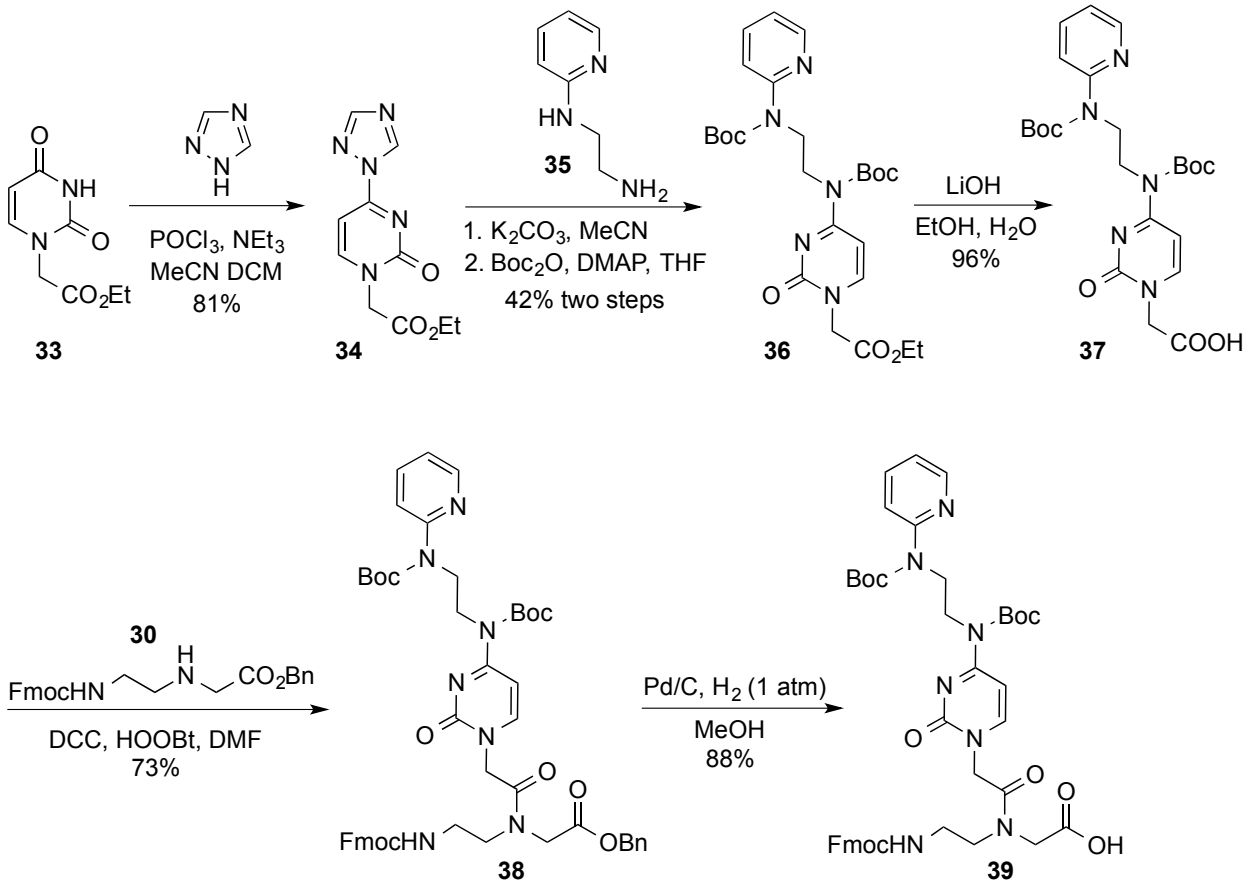
Synthesis of M3. Inspired by the promising binding properties reported for **Q**[29, 36, 44] (Figure 2) and our own experience with **M** (Figure 1), we designed **M3** and **M3R** where the guanidine substituent of **Q** was replaced by the structurally isosteric **M** (Figure 3). Synthesis of **M3** (Scheme 4) started with reductive alkylation of aminopyridine **13** followed by coupling with triazolyl uracil **26**.[36] After standard protection and deprotection steps the carboxylic acid precursor **29** was coupled to the benzyl protected PNA backbone **30**.[45] Hydrogenation gave the final **M3** PNA monomer **32**.

Scheme 4. Synthesis of M3 PNA monomer



Synthesis of M3R. Synthesis of **M3R** (Scheme 5) followed a similar route as for **M3** (Scheme 4) starting from uracil ethyl acetate **33**.[46] Coupling of the triazolyl derivative **34** with amine **35**[47] gave, after protection group manipulations, the acetic acid derivative **37**, which was converted to **M3R** PNA monomer **39**, using the same chemistry as in Scheme 4.

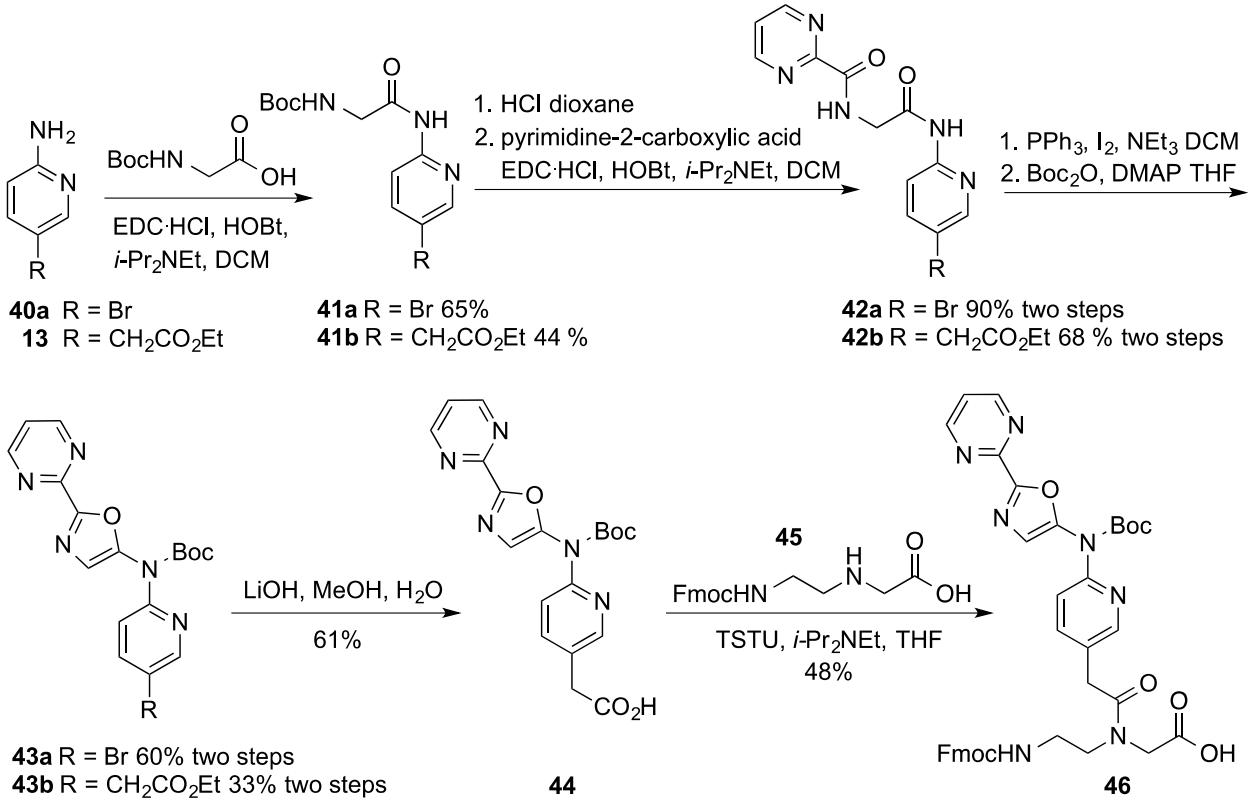
Scheme 5. Synthesis of M3R PNA monomer



Synthesis of M4. In parallel to nucleobases connected with flexible alkyl linkers, we pursued designs using heterocyclic linkers, such as oxazole in **M4** (Figure 3). The initial attempt at **M4** used, as a precursor of oxazole formation, amide **42a**, which was prepared by successive couplings of 4-bromo-2-aminopyridine (**40a**) with Boc-glycine

and, after deprotection, pyrimidine-2-carboxylic acid (Scheme 6). While trifluoroacetic acid anhydride[48] did not produce the desired ring closure reaction, treatment of **42a** with triphenylphosphinium iodide[49] gave the target oxazole that precipitated from the reaction mixture in a reasonably pure form. However, neither the previously used Cu catalyzed arylation of malonates (Scheme 2), nor the Negishi reaction (Scheme 3), nor the Heck reaction with *tert*-butyl((1-methoxyvinyl)oxy)dimethylsilane[50] gave the desired ester **43b**. An alternative approach starting from **13** that had the acetic acid residue in place was lower yielding but successful in providing the key precursor **44**. In this case, oxazole did not precipitate from the reaction mixture and could be purified only as **43b** after installation of the Boc group. Basic hydrolysis gave the carboxylic acid **44**, which was coupled with the unprotected version of PNA backbone **45**[51] to give the **M4** PNA monomer **46**. This one step procedure is preferable for compounds with limited solubility that are difficult to separate from residual Ph_3PO after removal of the allyl protecting group. However, the yields tend to be moderate due to an incomplete coupling step.

Scheme 6. Synthesis of M4 PNA monomer



40a R = Br
13 R = CH_2CO_2Et

41a R = Br 65%
41b R = CH_2CO_2Et 44 %

42a R = Br 90% two steps
42b R = CH_2CO_2Et 68 % two steps

43a R = Br 60% two steps
43b R = CH_2CO_2Et 33% two steps

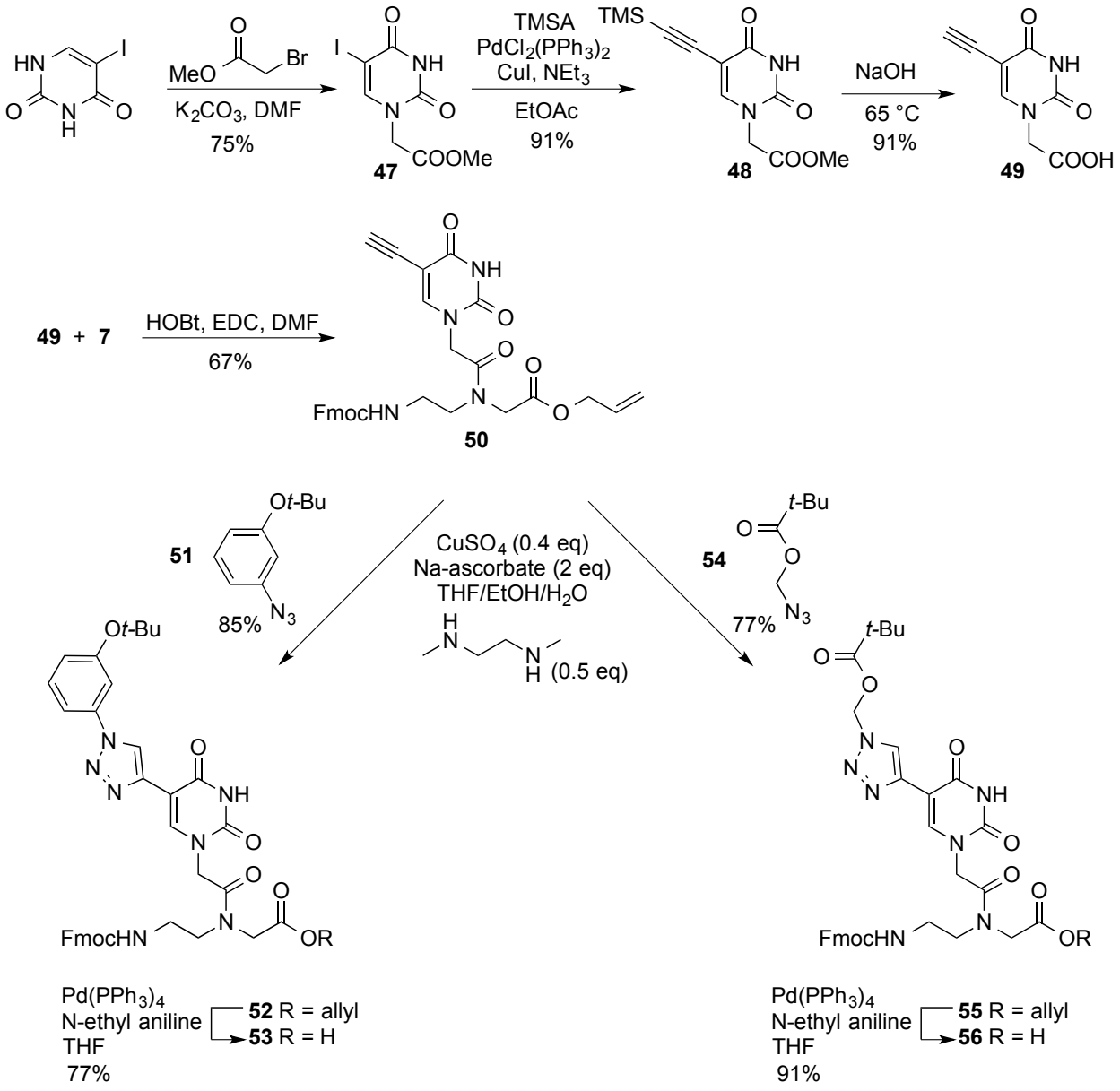
44

46

Synthesis of T1. For recognition of A-U base pairs, we explored linking additional functionality to U using the well-established substitution chemistry at the 5-position. The **T1** nucleobase (Figure 3) was envisioned as the product of a copper-catalyzed click reaction between 3-azidophenol and a 5-alkynl uracil derivative. In our initial synthetic studies, 5-ethynyl uracil was alkylated at N-1 with a bromoacetate ester. Unfortunately, the desired product was accompanied by significant inseparable dialkylation (N-1 and N-3) products. The copper catalyzed cycloaddition with 3-azido phenol afforded the desired nucleobase, however, it was found that the phenolic residue was problematic in the coupling to the PNA backbone. In a revised route (Scheme 7), we chose to perform amide coupling to the PNA backbone prior to the key cycloaddition step. This would in turn also allow for a late stage cycloaddition with various azide

coupling partners. To this end, 5-iodouracil was treated with methylbromoacetate to afford **47** in 75% yield. Notably, dialkylation was still observed (~5-10%), but a single recrystallization removed the undesired impurity. Sonigashira coupling afforded alkynyl uracil **48** in excellent yield, and treatment of **48** with aqueous sodium hydroxide accomplished both hydrolysis of the methyl ester and concomitant removal of the TMS group to afford **49**. Notably, **49** could be easily prepared on a gram scale from readily available starting materials. Finally coupling of **49** to the PNA backbone **7** occurred uneventfully using HOBt and EDC.[52]

Scheme 7. Synthesis of T1 and Tr PNA monomers



To avoid the previously encountered problems with the phenolic functional group at the stage of PNA synthesis, protection as a *t*-butyl ether was proposed. While other protecting groups may be more common on phenols, we chose *t*-butyl due to its general stability and because it would be easily removed using TFA upon cleavage of the final PNA from the solid support resin. To this end, attachment of the *t*-butyl group was

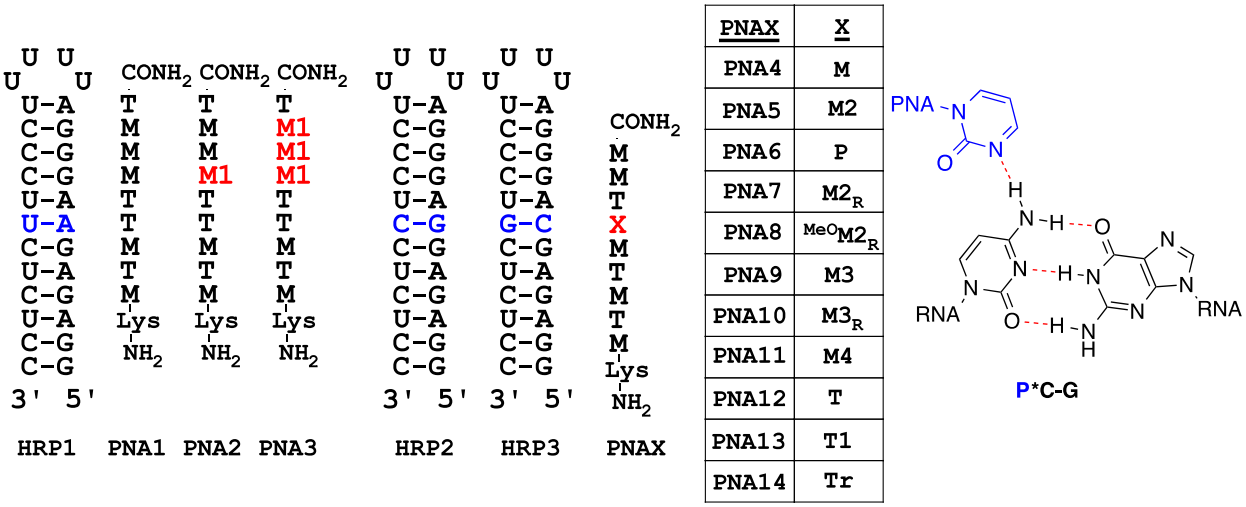
accomplished using the method of Bartoli involving either $Sc(OTf)_3$ or $Mg(OCl_4)_2$ as a catalyst and di-*tert*-butyl dicarbonate as the *t*-butyl source.[53] Butylation of 3-azido phenol afforded similar yields of **51** using either catalyst, so the less expensive $Mg(ClO_4)_2$ catalyst was chosen for larger scale reactions.

With both the azide (**51**) and alkyne (**50**) cycloaddition partners available, the [3+2] cycloaddition step was explored. Initial studies utilized copper sulfate/sodium ascorbate as the catalyst system in a 2:1 THF/water mixed solvent. To our delight, a 44% yield of the desired triazole **52** was obtained. Using *N*,*N*'-dimethylethylene diamine as an additive increased the yields further, and in an effort to better dissolve the reactants a 1:1:0.4 mixture of THF/EtOH/H_2O was employed. Under these optimized conditions, 85% yield of triazole **52** resulted, which was converted to the **T1** monomer **53** in 77% yield under standard deallylation conditions.[39]

An additional monomer, **Tr**, was made as a control to determine if the triazole affects binding of the uracil derivative to A-U. The pivaloylmethyl (POM) group was chosen to protect the triazole as it is known to be stable to the Fmoc chemistry used in PNA synthesis and could be readily removed with aqueous ammonia from PNA to form the free triazole.[54] **Tr** monomer (**56**) was prepared by the [3+2] cycloaddtion of azidomethyl pivalate **54**[54] with **50** following the same conditions for cyclization as described for **52**. Removal of the allyl protecting group on the PNA backbone afforded the POM-protected 5-triazolyl monomer **56**.

Synthesis and RNA binding studies of PNAs modified with extended nucleobases. We used the PNA monomers carrying the modified nucleobases **M1-M4** and **T1** along with the commercially available T and previously described M monomers

to synthesize modified PNAs (Figure 5) on an Expedite 8909 DNA/RNA/PNA synthesizer following our previously developed protocols.[8, 10, 39] The binding of PNAs to the complementary model hairpins was studied using isothermal titration calorimetry (ITC), as in our previous studies that characterized M-modified PNAs.[8, 10, 39] The hairpins **HRP1-HRP3** contain a variable base pair in the middle of the stem (colored blue in Figure 5) that allows evaluation of modified nucleobases targeting U-A, C-G, and G-C base pairs, respectively.



PNAX	X
PNA4	M
PNA5	M2
PNA6	P
PNA7	$M2_R$
PNA8	$^{MeO}M2_R$
PNA9	M3
PNA10	$M3_R$
PNA11	M4
PNA12	T
PNA13	T1
PNA14	Tr

Figure 5. Sequences of dsRNA hairpins and modified PNAs. The modified nucleobase is highlighted in red; the variable base pair is in blue.

To test the extended nucleobase **M1** we used a slightly modified PNA sequence (**PNA1**, Figure 5) that differs from our standard model 9-mer[10] by being shifted two base pairs up closer to the U_4 loop that closes the hairpin stem. The reason for initially choosing **PNA1** was that it has three consecutive M nucleobases, which allows checking for possible cooperative stabilization due to stacking of the extended nucleobases. Substitution of one M with **M1** resulted in significant loss of binding affinity

(~11 fold) of **PNA2** (Table 1). Three consecutive M residues further decreased the affinity of **PNA3** ~2.5 fold compared to the singly modified **PNA2**. While the latter result suggested that there might be some favorable contribution from stacking of the extended nucleobases, clearly, the additional pyridine ring of **M1** did not produce the expected stabilizing effect. Comparison of ΔH and ΔS values for **PNA1**, **PNA2** and **PNA3** (Table 1, columns 3 and 4) suggested that the decrease in binding affinity was caused by unfavorable enthalpy, which was compensated by increase in entropy.

Table 1. Binding affinities of modified PNAs for dsRNA hairpins.[a]

PNA (modified base)	Target dsRNA (variable base pair)	$K_a \times 10^6$ M^{-1}	ΔH (kcal/mol)	-TΔS (kcal/mol)
PNA1 (M)	**HRP1** (U-A)	170 ± 30	-71 ± 3	60 ± 3
PNA2 (M1)	**HRP1** (U-A)	15 ± 1	-44 ± 1	35 ± 1
PNA3 (3 × M1)	**HRP1** (U-A)	6 ± 1	-18 ± 1	9 ± 1
PNA4 (M)	**HRP2** (C-G)	110[b]	-103	92
PNA5 (M2)	**HRP2** (C-G)	8 ± 1	-48 ± 3	38 ± 3
PNA6 (P)	**HRP3** (G-C)	3[c]	NR[d]	NR[d]
PNA7 ($M2_R$)	**HRP3** (G-C)	6.5 ± 0.5	-40 ± 1	31 ± 1
PNA8 ($^{MeO}M2_R$)	**HRP3** (G-C)	4 ± 0.5	-38 ± 0	29 ± 0
PNA9 (M3)	**HRP2** (C-G)	21 ± 3	-68 ± 4	58 ± 4

PNA10 ($M3_R$)	**HRP3** (G-C)	7 ± 1	-41 ± 3	32 ± 3
PNA11 (M4)	**HRP2** (C-G)	4 ± 0.5	-36 ± 1	27 ± 1
PNA12 (T)	**HRP3** (U-A)	94[b]	-82[b]	72[b]
PNA13 (T1)	**HRP3** (U-A)	16 ± 3	-42 ± 6	32 ± 6
PNA14 (Tr)	**HRP3** (U-A)	72 ± 7	-79 ± 3	67 ± 3

[a] Association constant $K_a \times 10^6$ M^{-1}, measured by ITC using a Malvern MicroCal iTC200 in 2 mM $MgCl_2$, 90 mM KCl, 10 mM NaCl, 50 mM potassium phosphate buffer pH 7.4 at 25 °C. [b] Data from reference.[10] [c] Data from reference.[55] [d] Not reported in reference.[55]

Because **M1** did not show the expected improvement in binding affinity or significant cooperative stabilization, we decided to test other scaffolds using our standard model system **PNAX** where we had previous data to benchmark the new extended nucleobases.[10, 55] Substitution of M with **M2** resulted in significant loss of binding affinity for **PNA5,** ~14 fold compared to **PNA4**. **PNA7** and **PNA8** having the "reversed" nucleobases, **M2R** and **^{MeO}M2R** that target G-C interruptions in polypurine tracts, showed similarly low binding affinity. The *N*-acyl group in **M2** and **M2R** was expected to decrease the pK_a of aminopyridine, which lowers the binding affinity. We hypothesized that the electron donating methoxy group in **^{MeO}M2R** would counterbalance the unfavorable effect of the electron withdrawing *N*-acyl group, however, we did not observe an increase in binding affinity of **^{MeO}M2R** compared to **M2R**. A concurrent study in our group supports the hypothesis, but also explains why binding affinity is not restored upon the addition of 4-methoxy substituent on 2-aminopyridine nucleobase.[56]

A promising result was obtained with **M3**-modified **PNA9**, where the ~5 lower affinity compared to **PNA4** suggested that the one atom longer linker in **M3** was better tolerated than the shorter linker in **M1**. Interestingly, the "reversed" nucleobase, **M3R** showed notably lower affinity, consistent with our previous observation that cytosine is not a good match for recognition of G-C interruptions in polypurine tracts.[10] Taken together these results suggested that the **M3** scaffold could be further improved by replacing of cytosine with a better ligand for H-bonding to G-C.

Scaffolds built of more rigid, five membered linkers did not improve binding. Affinity of **M4** was weak and while **T1** showed encouraging binding, the affinity was ~6 lower than that for the control PNA having unmodified T (cf., **PNA12** and **PNA13** in Table 1). To understand the individual contributions of linker and second "nucleobase" in **T1**, we synthesized and tested nucleobase **Tr**, which contained only the triazole linker residue. Binding affinity of **Tr**-modified **PNA14** was comparable with that of the control **PNA12** carrying unmodified T. This result suggested that the heterocyclic linker alone did not interfere with triplex formation and that the decreased affinity of **T1** was most likely due to lack of favorable interactions by the phenol ring or steric clashes due to non-planar conformations about the tricyclic system.

Taken together our results demonstrate the challenges in triple-helical molecular recognition of pyrimidine interruptions in polypurine tracts of helical nucleic acids. Interestingly, the binding of PNAs having our "reversed" nucleobases, **M2R**, **^{MeO}M2R** and **M3R** was comparable or better than that of **PNA6** containing **P**, our current best nucleobase for recognition of G-C interruptions in polypurine tracts.[13] This result emphasizes the most significant unsolved problem in triple-helical recognition – while it

is relatively easy to form the standard Hoogsteen triples, triples involving binding to pyrimidine Hoogsteen edge are inherently difficult to attain. However, an encouraging result was obtained with **M3**, which may be improved by future optimization.

Molecular modeling of PNA-dsRNA triple helices containing extended nucleobases. To gain more insight into reasons behind the weak affinity of extended nucleobases, we modeled **M1** and **T1** in a PNA-dsRNA triplex based on our recent structural study.[56] A potential reason for the weak RNA binding was that the hydrogen-bonding conformations predicted by *ab initio* calculations were not favorable for attachment of extended nucleobases to the PNA backbone. Bases **M1** and **T1**, energy minimized in the absence of RNA, were fit to M* and T* of PNA residues of M*G-C and T*A-U triples, respectively, in the PNA-dsRNA structure. These models were subjected to simulated-annealing calculations in which distance restraints derived from our recent NMR study[56] were applied to the modified residues while atom positions were fixed for all other residues except the PNA residue preceding the modified residue. The base triples resulting from these calculations are shown in Figures 5A and 5B (green carbons). **M1** in the M*-aligned **M1***G-C triple (Figure 6A) is apparently well-positioned to form a hydrogen-bond with the C amino group (2.1 Å for N-H distance), consistent with the *ab initio* structure. Hydrogen-bonding angles and distances did not reveal any distortions that could explain the low binding affinity of **PNA2** and **PNA3** observed in ITC experiments (Table 1).

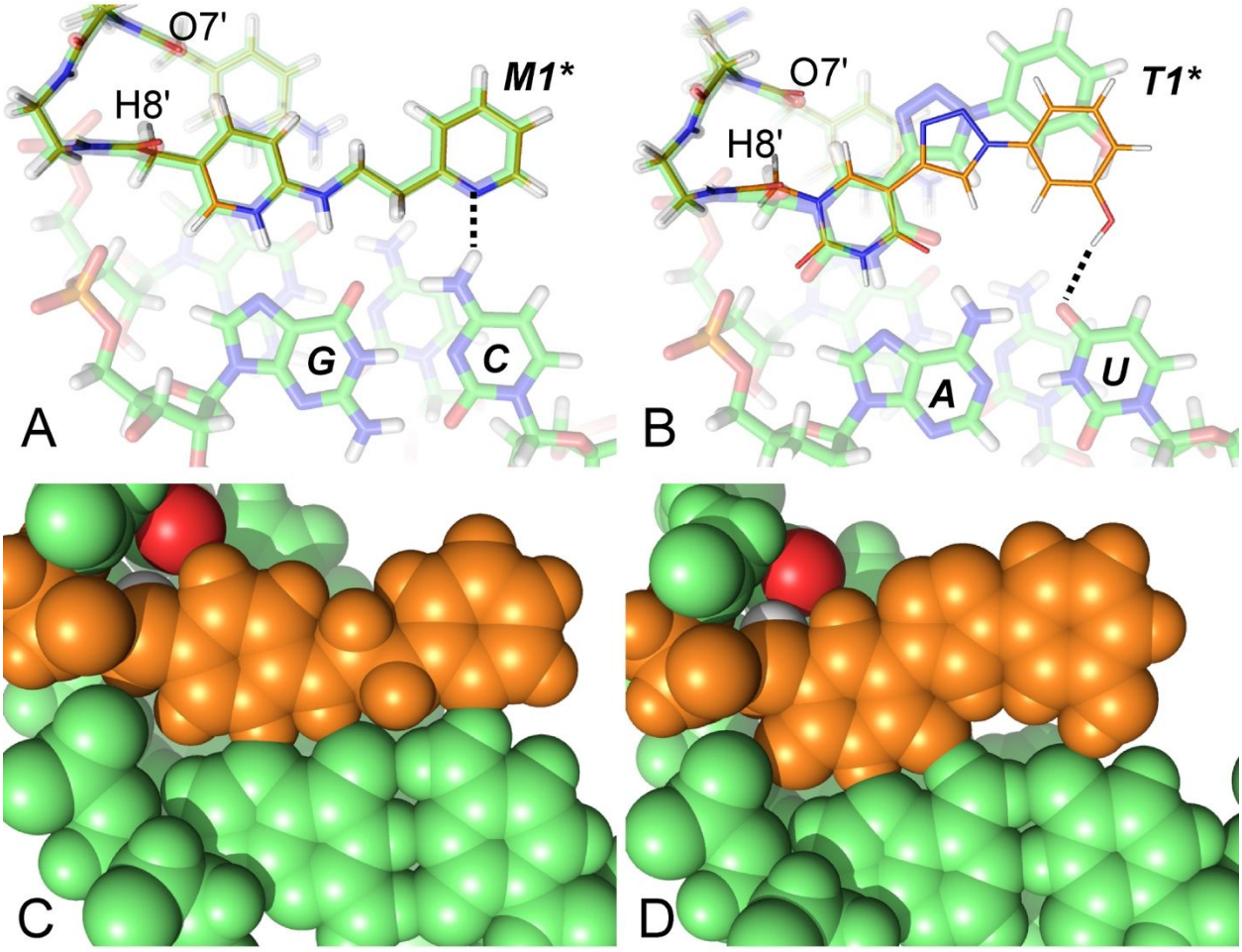


Figure 6. Simulated-annealing calculations of **M1***G-C and **T1***A-U triples. A and B, green carbons show **M1***G-C and **T1***A-U calculated using NMR restraints[56] applied in structure calculations of M*G-C and T*A-U triples, that were used as starting coordinates; A and B, thin and brown lines show **M1** and **T1** in the same calculations with addition of a distance restraint (maximum 2.4 Å) for the expected third hydrogen-bond (black, dashed line). C and D, space-filling view of models in A and B, respectively, that include the third hydrogen-bond restraint. Note contact between O7' (red) and H8' (white) in D (T1).

In contrast, **T1** in the T*-aligned **T1***A-U triple (Figure 6B) extends out away from the A-U pair and is not well-positioned to form a hydrogen-bond with the U carbonyl oxygen (4.8 Å for H-O distance). This contrasts with the *ab initio* structure. A second simulated-annealing calculation was performed on each model with the addition of a distance restraint for the expected third hydrogen-bond to C or U nucleobases in G-C and A-U base pairs (dashed lines in Figures 5A and 5B) The resulting conformations are shown as thin, brown lines in Figures 5A and 5B. In the **M1***G-C triple, this restraint had no effect as the N-H distance is already within the applied restraint (maximum 2.4 Å). In the **T1***A-U triple, **T1** is pulled back into the major groove (2.45 Å for H-O distance) and apparently can form a triple much as observed in *ab initio* calculations. However, there is a steric cost for this conformation. Introduction of the “hydrogen-bond” distance restraint in **T1***-AU forces the proximal end (H8’) of the nucleobase to swing out towards the major groove resulting in contact with oxygen of the carbonyl group (O7’) of the previous PNA residue (Figure 6D, white and red). This contact, as well as increased **T1**-A contact, is accompanied by an increase of van der Waals energy by 4.6±0.6 kcal (average of 10 simulations). As suggested by Figure 6A, introduction of the hydrogen-bond distance in **M1***G-C results in no change in van der Waals energy (-0.2±0.5 kcal). The close tolerances surrounding the nucleobase-to-PNA linker is demonstrated in Figure 6C. While steric factors near the PNA backbone demonstrate a possible consideration in the design of extended nucleobases, they do not explain the similarly poor affinities of **M1** and **T1**. Similar docking of **M2**, **M3**, and **M4** also do not reveal obvious explanations for their affinities. Higher level modeling may be required to explain the binding results.

CONCLUSIONS

Our present study illustrates the challenges in designing extended heterocyclic systems that hydrogen-bond the entire Hoogsteen edge of Watson-Crick base pairs. Initial modeling and ab initio calculations suggested the feasibility of several hydrogen-bonding scaffolds. The first challenge was to develop synthetic routes that link the heterocyclic nucleobases and form the proposed extended scaffolds. To that end, we successfully prepared five new extended nucleobases, **M1**-**M4** and **T1**, incorporated these at selected positions in PNA 9-mers, and measured their RNA binding affinity using ITC. Somewhat surprisingly, the extended nucleobases displayed reduced binding affinity compared to the original M and T nucleobases. Simulated-annealing calculations of **T1** identified another challenge of designing extended heterocyclic systems. While **T1** could achieve reasonable three hydrogen bond binding, it was accompanied by significant distortion in PNA backbone leading to unfavorable stereoelectronic interactions. In contrast, similar calculations on **M1** did not reveal any reasons for the decreased binding affinity, suggesting that other unidentified factors, for example, subtle conformational preferences and distortions of hydration network, may play significant role in overall binding energetics. Taken together, these results provide important insights for development of next generation nucleobase analogues that may address the long-standing limitation of triple helical recognition of nucleic acids – the requirement for long polypurine tracts.

EXPERIMENTAL SECTION

General synthetic procedures. All chemicals were obtained from commercial suppliers and were used without further purification unless stated otherwise. THF and toluene were dried by passing over activated alumnia. Methylene chloride, acetonitrile, and pyridine were obtained by refluxing over CaH_2 for six hours followed by distillation. All the dry reactions were carried out under an atmosphere of nitrogen using a Schlenk line or argon from a balloon. Silacycle 0.25 mm 60 Å silica gel F254 plates were used for TLC analysis and visualization was aided by UV light, iodine, or $KMnO_4$ stain. Column chromatography was performed using P60 230-400 mesh silica gel. Melting points were measured on a MEL-TEMP apparatus and were uncorrected. NMR spectra were obtained using a Bruker AM 600, 400 or 300 spectrometers with the chemical shift (δ) reported in parts per million (ppm) relative to TMS or to the solvent peak (DMSO-d6 or $CDCl_3$) as a reference. High Resolution Mass Spectrometry (HRMS) analyses using positive electrospray ionization (ESI+) were recorded on a Micromass Q-Tof micro™ instrument. Elemental analyses were performed on Carlo-Erba EA-1108 instrument.

6-((2-(Pyridin-2-yl)ethyl)amino)nicotinonitrile (3). 6-chloronicotinonitrile **1** (4.0 g, 28.9 mmol, 1 equiv.) and 2-(pyridin-2-yl)ethan-1-amine **2** (3.8 mL, 31.8 mmol, 1.1 equiv.) were dissolved in anhydrous THF (60 mL) and stirred at room temperature for 5 minutes. Then anhydrous triethylamine (8.1 mL, 57.8 mmol, 2 equiv.) was added and the reaction was purged with nitrogen for 15 minutes. The reaction mixture was brought to 80 °C (oil bath) and refluxed for 24 h. The precipitate (triethylammonium chloride) was filtered off, the filtrate was concentrated under reduced pressure, and the crude product was re-dissolved in 50 mL of EtOAc. The organic phase was washed with 30

mL of water and 30 mL of brine. The organic layer was dried over anhydrous Na_2SO_4, filtered and concentrated under reduced pressure. Silica gel chromatography using a linear gradient of 1-3% MeOH in CH_2Cl_2 gave product **3** (3.4 g, 53%) as a colorless oil. R_f = 0.37 (5% MeOH in CH_2Cl_2). HRMS (ESI/TOF) m/z: [M + H]+ calcd for $C_{13}H_{13}N_4$ 225.1140; found 225.1140. IR (neat, cm^{-1}) 3261.9; 2214.2. ^1H NMR (600 MHz, $CDCl_3$, ppm) δ: 8.45 (1H, d, J = 4.6 Hz), 8.24 (1H, d, J = 1.8 Hz), 7.54 (1H, td, J = 7.7, 1.7 Hz), 7.42 (1H, d, J = 7.9 Hz), 7.14 – 7.05 (1H, m), 6.31 (1H, d, J = 8.8 Hz), 6.09 (1H, s), 3.71 (2H, q, J = 5.9 Hz), 3.01 (2H, t, J = 6.5 Hz). $^{13}C\{^1H\}$ NMR (151 MHz, $CDCl_3$, ppm) δ: 159.8, 159.1, 153.2, 149.2, 139.2, 136.8, 123.6, 121.8, 118.8, 107.2, 96.4, 41.0, 36.8.

***tert*-Butyl (5-cyanopyridin-2-yl)(2-(pyridin-2-yl)ethyl)carbamate.** 6-((2-(Pyridin-2-yl)ethyl)amino)nicotinonitrile **3** (369 mg, 1.65 mmol, 1 equiv.) was dissolved in anhydrous DMF (10 mL) and the mixture was stirred at room temperature until homogenous. Then di-*tert*-butyl dicarbonate (0.46 mL, 1.97 mmol, 1.2 equiv.) was added, followed by the addition of 4-dimethylaminopyridine (40 mg, 0.33 mmol, 0.2 equiv.). The reaction mixture was stirred for 2 h at room temperature under nitrogen. The crude mixture was concentrated under reduced pressure and re-dissolved in 30 mL of CH_2Cl_2. The organic phase was washed with 30 mL of water and 30 mL of brine. The organic layer was dried over anhydrous Na_2SO_4, filtered and concentrated under reduced pressure. Silica gel chromatography using a linear gradient of 1-2% MeOH in CH_2Cl_2 gave the title compound (0.53 g, 99%) as a colorless oil. R_f = 0.43 (hexanes/EtOAc = 1:1). HRMS (ESI/TOF) m/z: $[M + H]^+$ calcd for $C_{18}H_{21}N_4O_2$ 325.1665; found 325.1666. IR (neat, cm^{-1}) 2980; 2224; 1699. ^1H NMR (600 MHz, $CDCl_3$, ppm) δ: 8.55 (1H, d, *J* = 2.2 Hz), 8.45 (1H, d, *J* = 4.2 Hz), 7.95 (1H, d, *J* = 8.9

Hz), 7.74 (1H, dd, *J* = 8.9, 2.3 Hz), 7.51 (1H, td, *J* = 7.6, 1.7 Hz), 7.18 – 7.02 (2H, m), 4.50 – 4.27 (2H, m), 3.20 – 3.01 (2H, m), 1.44 (9H, s). $^{13}C\{^{1}H\}$ NMR (151 MHz, $CDCl_3$, ppm) δ: 159.3, 156.8, 153.3, 151.0, 149.4, 139.5, 136.2, 123.31, 121.34, 117.6, 117.2, 103.5, 82.5, 46.3, 37.4, 28.1.

6-((*tert*-Butoxycarbonyl)(2-(pyridin-2-yl)ethyl)amino)nicotinic acid (4). *tert*-Butyl (5-cyanopyridin-2-yl)(2-(pyridin-2-yl)ethyl)carbamate (1.33 g, 4.10 mmol, 1 equiv.) was dissolved in EtOH (34 mL) and the mixture was stirred at room temperature until homogenous. Then 2 M aqueous NaOH solution (15 mL) was added to the reaction. The reaction mixture was brought to 80 ^{o}C (oil bath) and was refluxed for 2 h. The reaction was cooled to room temperature and EtOH was removed under reduced pressure. The crude mixture was slowly stirred in an ice bath while 20% aqueous citric acid was added until pH became 6 and product precipitated. The precipitate was filtered off, washed with cold deionized water, and dried on high vacuum to give **4** (1.137 g, 80%) as a white solid. R_f = 0.50 (20% MeOH in CH_2Cl_2). HRMS (ESI/TOF) m/z: $[M + H]^+$ calcd for $C_{18}H_{22}N_3O_4$, 344.1610; found 344.1610. IR (neat, cm^{-1}) 2974.2, 1714.7. ^{1}H NMR (600 MHz, DMSO-d6, ppm) δ: 13.19 (1H, s), 8.86 (1H, s), 8.45 (1H, s), 8.18 (1H, s), 7.72 (2H, d, *J* = 61.2 Hz), 7.19 (2H, s), 4.31 (2H, s), 3.04 (2H, s), 1.42 (9H, s). $^{13}C\{^{1}H\}$ NMR (151 MHz, DMSO-d6, ppm) δ: 165.8, 158.7, 156.6, 152.7, 148.8, 137.9, 136.2, 123.1, 121.6, 121.3, 117.7, 81.1, 45.9, 36.6, 27.5.

***tert*-Butyl (5-(2-diazoacetyl)pyridin-2-yl)(2-(pyridin-2-yl)ethyl)carbamate (5).** Dry 6-((*tert*-butoxycarbonyl)(2-(pyridin-2-yl)ethyl)amino)nicotinic acid **4** (1.865 g, 5.41 mmol, 1 equiv.) was dissolved in anhydrous CH_2Cl_2 (35 mL) and the mixture was stirred under nitrogen in an ice bath for 5 minutes. Oxalyl chloride (1.16 mL, 13.51 mmol, 2.5

equiv.) was added to the reaction followed by 1 drop of anhydrous DMF. The reaction was stirred in an ice bath for 1 h, then the ice bath was removed and the mixture was stirred for an additional 2.5 h at room temperature (slow flow of nitrogen was kept throughout the reaction). The reaction was put in a warm water bath (~40-50 ^{o}C) and concentrated under a strong flow of nitrogen. The residue was re-dissolved in anhydrous 1:1 mixture of acetonitrile and THF (23 mL) and stirred in an ice bath until solution became homogenous. Then 2.0M (trimethylsilyl)diazomethane solution in diethyl ether (13.4 mL, 5 equiv.) was added to the reaction dropwise and the mixture was stirred in an ice bath for 2 h under a slow flow of nitrogen. The ice bath was removed and the reaction mixture was left stirring for 12 h at room temperature under a slow flow of nitrogen. The reaction was cooled in an ice bath and quenched with 5% aqueous Na_2SO_3 solution. The aqueous layer was extracted with EtOAc (3 × 50 mL). The organic layer was dried over anhydrous Na_2SO_4, filtered and concentrated under reduced pressure. Silica gel chromatography using a linear gradient of 1-4% MeOH in CH_2Cl_2 gave product **5** (0.98 g, 49%) as a yellow oil. R_f = 0.56 (5% MeOH in CH_2Cl_2). HRMS (ESI/TOF) m/z: $[M + H]^+$ calcd for $C_{19}H_{22}N_5O_3$, 368.1723; found 368.1719. IR (neat, cm^{-1}) 2976; 2102; 1707. ^{1}H NMR (600 MHz, $CDCl_3$, ppm) δ: 8.69 (1H, d, *J* = 2.1 Hz), 8.49 (1H, d, *J* = 4.3 Hz), 7.97 (1H, dd, *J* = 8.8, 2.4 Hz), 7.83 (1H, d, *J* = 8.8 Hz), 7.54 (1H, td, *J* = 7.6, 1.6 Hz), 7.12 (1H, d, *J* = 7.7 Hz), 7.08 (1H, dd, *J* = 6.9, 5.3 Hz), 5.87 (1H, s), 4.45 – 4.35 (2H, m), 3.19 – 3.08 (2H, m), 1.47 (9H, s). ^{13}C{^{1}H} NMR (151 MHz, $CDCl_3$, ppm) δ: 183.9, 159.6, 157.5, 153.7, 149.5, 146.5, 136.3, 135.4, 127.4, 123.5, 121.4, 118.0, 82.1, 54.4, 46.7, 37.7, 28.3.

Ethyl 2-(6-((*tert*-butoxycarbonyl)(2-(pyridin-2-yl)ethyl)amino)pyridin-3-yl)acetate (6). Dry *tert*-butyl (5-(2-diazoacetyl)pyridin-2-yl)(2-(pyridin-2-yl)ethyl)carbamate **5** (946 mg, 2.57 mmol, 1 equiv.) was suspended in absolute EtOH (56 mL) followed by addition of anhydrous triethylamine (0.54 mL, 3.86 mmol, 1.5 equiv.) and silver benzoate (176 mg, 0.77 mmol, 0.3 equiv.). The reaction was brought to 80 °C (oil bath) and stirred under reflux for 2 h. The reaction was cooled to room temperature and concentrated under reduced pressure. The crude product was re-dissolved in EtOAc (50 mL) and washed with 30 mL of 5% aqueous $NaHCO_3$, water, and brine. The organic layer was dried over anhydrous Na_2SO_4, filtered, and concentrated under reduced pressure. Silica gel chromatography using a linear gradient of 1-4% MeOH in CH_2Cl_2 gave **6** (609 mg, 61%) as a colorless oil. R_f = 0.39 (5% MeOH in CH_2Cl_2). HRMS (ESI/TOF) m/z: $[M + H]^+$ calcd for $C_{21}H_{28}N_3O_4$, 386.2080; found 386.2077. IR (neat, cm^{-1}) 2978, 1701. ^1H NMR (600 MHz, $CDCl_3$, ppm) δ: 8.50 – 8.42 (1H, m), 8.23 (1H, s), 7.57 – 7.44 (3H, m), 7.10 (1H, d, *J* = 7.8 Hz), 7.07 – 7.00 (1H, m), 4.33 – 4.23 (2H, m), 4.11 (2H, q, *J* = 7.1 Hz), 3.52 (2H, s), 3.15 – 3.03 (2H, m), 1.40 (9H, s), 1.21 (3H, t, *J* = 7.1 Hz). $^{13}C\{^1H\}$ NMR (151 MHz, $CDCl_3$, ppm) δ: 170.8, 159.7, 154.0, 153.5, 149.2, 147.9, 137.7, 136.1, 125.4, 123.4, 121.2, 119.5, 81.0, 61.1, 46.8, 37.9, 37.6, 28.2, 14.1

2-(6-((*tert*-Butoxycarbonyl)(2-(pyridin-2-yl)ethyl)amino)pyridin-3-yl)acetic acid. Ethyl 2-(6-((*tert*-butoxycarbonyl)(2-(pyridin-2-yl)ethyl)amino)pyridin-3-yl)acetate (1.286 g, 3.34 mmol, 1 equiv.) was dissolved in MeOH (10 mL) followed by addition of 1.6 M aqueous NaOH solution (10 mL) and the mixture was stirred until homogenous. Then reaction was brought to 75 °C (oil bath) and stirred under reflux for 2 h. Then

reaction mixture was cooled to room temperature and concentrated under reduced pressure. The crude title compound was used in the next step without any further purification (assumed 100% conversion).

Allyl ester of M1 PNA monomer (8). Dry 2-(6-((*tert*-butoxycarbonyl)(2-(pyridin-2-yl)ethyl)amino)pyridin-3-yl)acetic acid (178 mg, 0.50 mmol, 1 equiv.) and 3-hydroxy-1,2,3-benzotriazin-4(3*H*)-one hydrate (HOOBt) (89 mg, 0.55 mmol, 1.1 equiv.) were dissolved in 3.5 mL of anhydrous DMF and cooled to 0 °C. *N,N'*-dicyclohexylcarbodiimide (DCC) (124 mg, 0.60 mmol, 1.2 equiv.) was added and the resulting mixture was stirred at 0 °C for 80 minutes. PNA backbone **7** (194 mg, 0.51 mmol, 1 equiv.) was added as a solid in one portion, the ice bath was removed, and the reaction was stirred at room temperature for 24 h. The resulting orange-brown solution was concentrated by rotatory evaporation, re-dissolved in CH_2Cl_2 (40 mL), and washed with 5% aqueous $NaHCO_3$ (2 × 20 mL). The aqueous layer was back-extracted with DCM (2 × 20 mL), then all the organic layers were combined and washed with brine (20 mL). The combined organic layers were dried over Na_2SO_4 and concentrated to a crude brown oil. Silica gel chromatography using a linear gradient of 1-3% MeOH in CH_2Cl_2 gave product **8** (295 mg, 82%) as a white foam. R_f = 0.73 (5% MeOH in CH_2Cl_2). HRMS (ESI/TOF) m/z: $[M + H]^+$ calcd for $C_{41}H_{46}N_5O_7$, 720.3397; found 720.3405. ^1H NMR (600 MHz, DMSO-d6, ppm, mixture of rotamers) δ: 8.45 (1H, t, *J* = 4.3 Hz), 8.23 (1H, s), 8.19 (1H, d, *J* = 1.5 Hz), 7.88 (2H, d, *J* = 7.5 Hz), 7.74 – 7.60 (2H, m), 7.54 (1H, dd, *J* = 13.3, 5.1 Hz), 7.43 (3H, dt, *J* = 14.7, 8.0 Hz), 7.36 – 7.24 (2h, m), 7.18 (2H, q, *J* = 4.7 Hz), 5.92 (1H, dddd, *J* = 27.9, 22.4, 10.7, 5.4 Hz), 5.40 – 5.14 (2H, m), 4.64 (1H, d, *J* = 5.4 Hz), 4.58 (1H, d, *J* = 5.3 Hz), 4.41 (1H, s), 4.37 (1H, d, *J* = 6.7 Hz), 4.31 (1H,

d, *J* = 6.9 Hz), 4.23 (1H, t, *J* = 6.7 Hz), 4.18 (2H, dd, *J* = 14.2, 6.7 Hz), 4.10 (1H, s), 3.76 (1H, s), 3.63 (1H, s), 3.50 (1H, t, *J* = 6.1 Hz), 3.39 (1H, t, *J* = 6.3 Hz), 3.35 (1H, s), 3.25 (1H, dd, *J* = 11.9, 5.8 Hz), 3.17 (1H, dd, *J* = 12.2, 6.1 Hz), 3.00 (2H, t, *J* = 7.3 Hz), 1.39 (9H, s). $^{13}C\{^{1}H\}$ NMR (151 MHz, DMSO-d6, ppm, mixture of rotamers) δ: 170.8, 170.5, 169.4, 169.0, 159.0, 156.4, 156.1, 153.1, 152.42, 152.36, 149.0, 147.9, 147.8, 143.9, 143.8, 140.74, 140.71, 138.3, 138.0, 136.29, 136.27, 132.3, 132.2, 127.6, 127.2, 127.0, 125.1, 125.0, 123.1, 121.3, 120.1, 119.2, 119.1, 118.2, 117.8, 80.3, 65.4, 65.4, 64.8, 50.0, 48.0, 47.7, 46.7, 46.7, 46.4, 38.9, 38.2, 36.9, 35.5, 35.1, 27.7.

M1 PNA monomer (9). The allyl ester of **M1** PNA monomer **8** (295 mg, 0.41 mmol, 1 equiv.) was dissolved in anhydrous THF (15 mL). $Pd(PPh_3)_4$ (183 mg, 0.16 mmol, 0.4 equiv.) was added followed by *N*-ethylaniline (0.98 mL, 7.75 mmol, 19 equiv.). The yellow solution was stirred under nitrogen for 2 h at room temperature. After the reaction was complete by TLC, the solvent was removed under reduced pressure and the yellow residue was re-dissolved in CH_2Cl_2 (40 mL). The resulting solution was washed with 10% aqueous $KHSO_4$ solution (3 × 20 mL) and the combined aqueous layers were back-extracted with CH_2Cl_2 (30 mL). The combined organic extracts were washed with brine (20 mL), dried over Na_2SO_4, and concentrated under reduced pressure to afford a crude yellow solid containing the product contaminated with allyl anilines. The solid was re-dissolved in CH_2Cl_2, adsorbed onto ~600 mg of celite, and purified using flash chromatography using a linear gradient of 0-10% MeOH in CH_2Cl_2 on a 20 g Silica-CM column (Agela Technologies) to afford PNA monomer **9** (241 mg, 91%) as a white foam. R_f = 0.40 (20% MeOH in CH_2Cl_2). HRMS (ESI/TOF) m/z: $[M + H]^+$ calcd for $C_{38}H_{42}N_5O_7$, 680.3084; found 680.3076. ^{1}H NMR (600 MHz,

DMSO-d6, ppm): (mixture of rotamers) δ 12.69 (1H, s), 8.44 (1H, dd, *J* = 6.7, 2.9 Hz), 8.21 (1H, d, *J* = 1.7 Hz), 8.17 (1H, d, *J* = 2.1 Hz), 7.88 (1H, d, *J* = 7.5 Hz), 7.72 – 7.60 (2H, m), 7.58 – 7.50 (1H, m), 7.42 (3H, dt, *J* = 14.3, 6.8 Hz), 7.31 (2H, tt, *J* = 10.7, 4.9 Hz), 7.18 (2H, td, *J* = 7.3, 3.8 Hz), 4.34 (1H, d, *J* = 6.8 Hz), 4.28 (1H, d, *J* = 7.0 Hz), 4.25 – 4.19 (1H, m), 4.16 (2H, dd, *J* = 14.0, 6.3 Hz), 3.97 (1H, s), 3.73 (1H, s), 3.60 (1H, d, *J* = 7.4 Hz), 3.51 (1H, s), 3.46 (1H, t, *J* = 6.4 Hz), 3.39 – 3.34 (1H, m), 3.23 (1H, dd, *J* = 12.0, 5.9 Hz), 3.14 (1H, dd, *J* = 12.4, 6.2 Hz), 2.98 (2H, t, *J* = 7.3 Hz), 1.38 (9H, s). $^{13}C\{^{1}H\}$ NMR (151 MHz, DMSO-d6, ppm): (mixture of rotamers) δ 170.7, 170.3, 158.9, 156.3, 153.1, 152.4, 148.99, 148.97, 148.0, 147.8, 143.9, 143.8, 140.7, 140.7, 138.3, 138.1, 136.4, 136.3, 127.6, 127.4, 127.0, 125.1, 125.0, 123.1, 121.4, 120.1, 119.2, 80.3, 69.8, 65.4, 47.9, 47.5, 46.7, 46.4, 38.8, 36.8, 35.0, 27.8.

Diethyl 2-(pyrimidin-2-yl)malonate (10) was prepared following a literature procedure.[57] 2-Iodopyridine (7.60 g, 36.9 mmol), Cs_2CO_3 (20.04 g, 73.8 mmol, 2 equiv.), CuI (1.40 g, 7.38 mmol, 0.2 equiv.) and 2-picolinic acid (1.81 g, 14.8 mmol, 0.2 equiv.) were placed into a dry flask under an argon atmosphere. DMF (50 mL) was added followed by diethyl malonate (11.2 mL, 73.8 mmol, 2 equiv.) and the reaction mixture was stirred at 80 °C (oil bath) for 21 h. The brown mixture was cooled to room temperature and filtered through a pad of celite. The pad was washed with EtOAc (4 × 30 mL), the combined organic solvents were extracted with saturated aqueous NH_4Cl (70 mL), dried with Na_2SO_4, and evaporated under reduced pressure. The crude product was purified by silica gel column chromatography using a linear gradient (0-70%) of EtOAc in petroleum ether (40-60 °C) to afford the title compound **10** (1.65 g, 68% yield) as a yellow oil. R_f = 0.63 (hexanes/EtOAc, 1:1). HRMS (ESI/Q-TOF) m/z: [M

+ H]$^+$ calcd for $C_{11}H_{15}N_2O_4$, 239.1032; found 239.1040. IR (neat, cm^{-1}) 3473, 2984, 2941, 2907, 1756, 1568, 1423, 1308, 1255, 1179, 1153. ^{1}H NMR (300 MHz, $CDCl_3$, ppm) δ: 8.75 (2H, d, *J* = 4.9 Hz), 7.25 (1H, t, *J* = 4.9 Hz), 5.10 (1H, s), 4.28 (4H, t, *J* = 7.1 Hz), 1.29 (6H, t, *J* = 7.1 Hz). ^{13}C{^{1}H} NMR (100.6 MHz, $CDCl_3$, ppm) δ: 166.7, 163.7, 157.6, 120.0, 62.2, 62.0, 14.1.

2-(Pyrimidin-2-yl)acetic acid (11). A 3 M solution of NaOH (1.91 g, 47.9 mmol, 3 equiv.) in water (16 mL) was added to a solution of diethyl 2-(pyrimidin-2-yl)malonate **10** (3.80 g, 16.0 mmol) in EtOH (20 mL). The reaction mixture was stirred at room temperature for 22 h, partly evaporated, and acidified with 1M HCl to pH 3-4. The aqueous mixture was extracted with EtOAc (8 × 25 mL). The organic phases were combined, dried with Na_2SO_4, and evaporated under reduced pressure to afford the title compound (1.60 g, 73% yield) as an off-white solid that was used in next reaction without further purification.

1-(*tert*-Butyl)-3-ethyl-2-(6-nitropyridin-3-yl)malonate (12). *tert*-Buthyl ethyl malonate (28.0 mL, 148 mmol, 2 equiv.) was added dropwise to a suspension of NaH (60% dispersion in mineral oil, 5.91 g, 148 mmol, 2 equiv.) in dry DMF (75 mL) at 0 °C (ice bath) and under an argon atmosphere. The suspension was stirred at room temperature till the mixture become a clear solution, then 5-bromo-2-nitropyridine (15.0 g, 73.9 mmol) was added. The reaction mixture was stirred at 80 °C (oil bath) for 19 h, cooled to room temperature and saturated aqueous NH_4Cl (70 mL) was added. The pH was adjusted to 4 by addition of aqueous 1 N HCl solution and the aqueous phase was extracted with EtOAc (2 × 100 mL). The combined organic layers were washed with brine (70 mL), dried with Na_2SO_4 and concentrated under reduced pressure. The crude

product was purified by silica gel column chromatography using a linear gradient of EtOAc (5-20%) in petrol ether to afford the title compound **12** (18.9 g, 82% yield) as a light-yellow oil. R_f = 0.52 (hexanes/EtOAc = 3:1). HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_{14}H_{19}N_2O_6$, 311.1243; found 311.1247. IR (neat, cm^{-1}) 3456, 3110, 2981, 2938, 2908, 1743, 1466, 1370, 1313, 1238, 1144. ^1H NMR (300 MHz, $CDCl_3$, ppm) δ: 8.60 (1H, d, *J* = 2.2 Hz), 8.27 (1H, d, *J* = 8.4 Hz), 8.20 (1H, dd, *J* = 8.4, 2.2 Hz), 4.70 (1H, s), 4.32 – 4.18 (2H, m), 1.46 (9H, s), 1.29 (3H, t, *J* = 7.1 Hz). $^{13}C\{^1H\}$ NMR (100.6 MHz, $CDCl_3$, ppm) δ: 166.8, 165.4, 156.4, 149.6, 141.1, 135.6, 117.9, 84.1, 62.7, 56.2, 27.9, 14.2. The ^1H data were in agreement with the previously published data.[58]

Ethyl 2-(6-nitropyridin-3-yl)acetate. TFA (46.8 mL, 609 mmol, 10 equiv.) was added to a solution of 1-ethyl 3-isopropyl 2-(6-nitropyridin-3-yl)malonate **12** (18.9 g, 60.9 mmol) in CH_2Cl_2 (100 mL) at 0 °C (ice bath). After 25 h at room temperature, the solvent was evaporated. The residue was co-evaporated under reduced pressure with toluene (2 × 50 mL) to afford the title compound (12.7 g, 74% yield) as a brown solid. R_f = 0.63 (hexanes/EtOAc = 1:1). HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_9H_{11}N_2O_4$, 211.0719; found 211.0718. IR (neat, cm^{-1}) 3447, 3110, 3060, 2984, 2938, 1729, 1537, 1534, 1465, 1354, 1233, 1181. ^1H NMR (300 MHz, $CDCl_3$, ppm) δ: 8.57 (1H, d, *J* = 2.2 Hz), 8.26 (1H, d, *J* = 8.4 Hz), 8.03 (1H, dd, *J* = 8.4, 2.2 Hz), 4.20 (2H, q, *J* = 7.1 Hz), 3.80 (2H, s), 1.27 (3H, t, *J* = 7.1 Hz). $^{13}C\{^1H\}$ NMR (100.6 MHz, $CDCl_3$, ppm) δ: 169.4, 155.9, 149.6, 140.8, 136.5, 117.9, 61.9, 38.1, 14.2. The ^1H data were in agreement with the previously published data.[58]

Ethyl 2-(6-aminopyridin-3-yl)acetate (13). Aqueous saturated NH_4Cl (18 mL) and Fe (1.43 g, 25.7 mmol, 2 equiv.) were added to a solution of ethyl 2-(6-nitropyridin-

3-yl)acetate (2.70 g, 12.8 mmol) in EtOH (20 mL). The mixture was heated for 2 h at 75 °C (oil bath) and filtered through pad of celite. The pad was washed with water (15 mL) and EtOH (3 × 15 mL). The combined filtrates and washes were partly evaporated and obtained oily residue that was extracted with CH_2Cl_2 (4 × 30 mL). The organic phases were combined, dried with Na_2SO_4 and evaporated under reduced pressure to afford the title compound **13** (2.1 g, 91% yield) as a black solid. R_f = 0.25 (5% MeOH in CH_2Cl_2). HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_9H_{13}N_2O_2$, 181.0977; found 181.0975. IR (neat, cm^{-1}) 3319, 3144, 2981, 2706, 1736, 1671, 1629, 1559, 1477, 1376, 1336, 1235, 1182. ^{1}H-NMR (300 MHz, $CDCl_3$, ppm) δ: 7.89 – 7.80 (1H, m), 7.49 (1H, dd, *J* = 8.6, 2.3 Hz), 6.57 (1H, dd, *J* = 8.6, 0.8 Hz), 6.0 – 5.0 (2H, br s), 4.15 (2H, q, *J* = 7.1 Hz), 3.47 (2H, s), 1.25 (3H, t, *J* = 7.1 Hz). $^{13}C\{^1H\}$ NMR (100.6 MHz, $CDCl_3$, ppm) δ: 171.2, 156.7, 144.2, 140.8, 119.5, 110.2, 61.3, 37.5, 14.3. The ^{1}H data were in agreement with the previously published data.[59]

Ethyl 2-(6-(2-(pyrimidin-2-yl)acetamido)pyridin-3-yl)acetate (14). HATU (825 mg, 2.17 mmol) was added to a solution of 2-(pyrimidin-2-yl)acetic acid **11** (300 mg, 2.17 mmol), ethyl 2-(6-aminopyridin-3-yl)acetate **13** (391 mg, 2.17 mmol) and DIPEA (0.75 mL, 4.34 mmol, 2 equiv.) in dry CH_2Cl_2 (15 mL) under an argon atmosphere. After 2 h at room temperature, the solvent was evaporated under reduced pressure and the residue was purified by silica gel column chromatography using a linear gradient (50-80%) of EtOAc in petrol ether to afford the title compound **14** (483 mg, 74% yield) as a yellow oil. R_f = 0.36 (5% MeOH in CH_2Cl_2). HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_{15}H_{17}N_4O_3$, 301.1301; found 301.1298. IR (neat, cm^{-1}) 3283, 2983, 1733, 1695, 1564, 1424, 1400, 1306, 1231, 1174. ^{1}H NMR (300 MHz, $CDCl_3$, ppm) δ: 8.78 (2H, d, *J* = 4.9

Hz), 8.23 – 8.17 (2H, m), 7.64 (1H, dd, *J* = 8.5, 2.4 Hz), 7.27 (1H, t, *J* = 5.0 Hz), 4.15 (2H, q, *J* = 7.1 Hz), 4.14 (2H, s), 3.56 (2H, s), 1.24 (3H, t, *J* = 7.1 Hz). $^{13}C\{^{1}H\}$ NMR (100.6 MHz, $CDCl_3$, ppm) δ: 170.9, 166.5, 164.9, 157.5, 150.5, 148.3, 139.3, 125.9, 119.6, 114.0, 61.3, 47.7, 38.14, 14.3.

2-(6-(2-(Pyrimidin-2-yl)acetamido)pyridin-3-yl)acetic acid (15). A 3 M solution of NaOH (531 mg, 13.29 mmol, 3 equiv.) in water (4.4 mL) was added to a solution of ethyl 2-(6-(2-(pyrimidin-2-yl)acetamido)pyridin-3-yl)acetate **14** (1.65 g, 4.43 mmol) in EtOH (10 mL). The reaction mixture was stirred at room temperature for 2 h. The solvent was evaporated under reduced pressure. The residual oil was acidified with 4 M HCl (pH~3-4) and extracted with EtOAc (4 × 30 mL). The combined organic phases were extracted with saturated aqueous NaCl (23 mL) and dried with Na_2SO_4. The solvent was evaporated under reduced pressure to afford the title compound **15** (820 mg, 68% yield) as an off-white solid. HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_{13}H_{13}N_4O_3$, 273.0988; found 273.1000. IR (neat, cm^{-1}) 3118, 2436, 1947, 1710, 1599, 1564, 1423, 1398, 1385, 1308, 1155, 1137. ^{1}H NMR (400 MHz, DMSO-d6, ppm) δ: 12.45 (1H, s), 10.74 (1H, s), 8.76 (2H, d, *J* = 4.9 Hz), 8.20 (1H, dd, *J* = 2.4, 0.8 Hz), 8.02 (1H, d, *J* = 8.5 Hz), 7.67 (1H, dd, *J* = 8.5, 2.4 Hz), 7.40 (1H, t, *J* = 4.9 Hz), 4.09 (2H, s), 3.59 (2H, s). $^{13}C\{^{1}H\}$ NMR (100.6 MHz, DMSO-d6, ppm) δ: 172.5, 168.0, 165.5, 157.4, 150.6, 148.4, 139.1, 126.3, 119.6, 112.8, 47.1, 37.0.

Allyl ester of M2 PNA monomer (16). 2-(6-(2-(pyrimidin-2-yl)acetamido)pyridin-3-yl)acetic acid **15** (600 mg, 2.2 mmol, 1.1 equiv.), PNA backbone **7** (760 mg, 2.0 mmol) and HOOBt (423 mg, 2.6 mmol, 1.3 equiv.) were dissolved in anhydrous DMF (10 mL) under an argon atmosphere. The solution was cooled in ice bath and DCC (577 mg, 2.8

mmol, 1.3 equiv.) was added. After 10 min, the ice bath was removed and the solution was stirred overnight (18 h) at room temperature. The reaction mixture was evaporated under reduced pressure, the residue was dissolved in EtOAc (120 mL) and extracted with saturated aqueous $NaHCO_3$ (25 mL). The organic layer was dried with Na_2SO_4 and evaporated under reduced pressure. The crude product was purified by silica gel column chromatography using a linear gradient of EtOH (0-5%) in CH_2Cl_2 to afford compound **16** (875 mg, 69% yield) as an off-white foam. R_f = 0.79 (10% MeOH in CH_2Cl_2). HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_{35}H_{35}N_6O_6$, 635.2618; found 635.2620. IR (neat, cm^{-1}) 3299, 3012, 2946, 1959, 1746, 1704, 1645, 1564, 1531, 1520, 1307, 1248, 1189. ^1H NMR (400 MHz, $CDCl_3$, ppm) (mixture of rotamers) δ: 9.93 (1H, s), 8.78 – 8.72 (2H, m), 8.22 – 8.10 (2H, m), 7.75 (2H, d, *J* = 7.6 Hz), 7.58 (3H, d, *J* = 7.5 Hz), 7.39 (2H, t, *J* = 7.6 Hz), 7.30 (2H, t, *J* = 7.5 Hz), 7.25 – 7.21 (1H, m), 5.97 – 5.83 (1H, m), 5.69 (0.7 H, t, *J* = 6.0 Hz), 5.40 (0.3H, t, *J* = 5.2 Hz) 5.34 (1H, d, *J* = 17.1 Hz), 5.26 (1H, d, *J* = 11.2 Hz), 4.70 – 4.57 (2H, m), 4.44 – 4.32 (2H, m), 4.20 (1H, t, *J* = 6.8 Hz), 4.14 – 4.00 (4H, m), 3.66 – 3.50 (4H, m), 3.43 – 3.29 (2H, m). $^{13}C\{^1H\}$ NMR (100.6 MHz, $CDCl_3$, ppm) (mixture of rotamers) δ: 171.3, 170.0, 166.5, 164.9, 157.5, 156.7, 150.4, 148.1, 144.0, 143.9, 141.4, 139.3, 139.2, 131.5, 131.2, 127.9, 127.8, 127.2, 126.3, 125.1, 120.1, 120.1, 119.6, 119.2, 114.1, 67.0, 66.7, 66.3, 49.7, 49.3, 47.6, 47.3, 39.6, 37.1, 36.5.

M2 PNA monomer (17). $Pd(PPh_3)_4$ (140 mg, 0.126 mmol, 0.1 equiv.) and N-methylaniline (275 µl, 2.52 mmol, 2 equiv.) were added to solution of allyl ester of **M2** PNA monomer **16** (800 mg, 1.26 mmol) in dry THF (20 mL) under an argon atmosphere and the solution was stirred overnight (18 h) at room temperature. The solvent was

evaporated under reduced pressure and obtained residue was purified by reverse phase chromatography using a linear gradient (0-30%) of MeCN in water to afford compound **17** (518 mg, 69% yield) as a yellow foam. HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_{32}H_{31}N_6O_6$, 595.1469; found 595.1460. ^{1}H NMR (300 MHz, DMSO-d6, ppm) (mixture of rotamers) δ: 10.70 (1H, s), 8.76 (2H, d, *J* = 4.9 Hz), 8.17 – 8.07 (1H, m), 8.04 – 7.80 (4H, m), 7.70 – 7.53 (3H, m), 7.44 – 7.34 (3H, m), 7.34 – 7.25 (2H, m), 4.37 – 4.22 (1H, m), 4.20 (2H, s), 4.07 (2H, s), 3.80 – 3.50 (4H, m), 3.47 – 3.30 (3H, m), 3.22 – 3.10 (2H, m). ^{13}C{^{1}H} NMR (100.6 MHz, DMSO-d6, ppm) (mixture of rotamers) δ: 170.9, 170.3, 167.8, 165.5, 157.4, 156.1, 150.3, 150.2, 148.4, 143.9, 143.8, 142.6, 140.7, 140.6, 139.4, 139.2, 137.4, 128.9, 127.7, 127.6, 127.3, 127.1, 125.3, 125.1, 121.4, 120.0, 119.5, 112.7, 109.7, 65.5, 59.7, 53.6, 47.4, 47.1, 46.7, 38.3, 35.8, 20.8, 14.1.

1-(*tert*-Butyl)-3-ethyl-2-(5-bromopyrimidin-2-yl)malonate. *tert*-Butyl ethyl malonate (5.19 mL, 27.4 mmol, 2 equiv.) was added dropwise to a suspension of NaH, 60% dispersion in mineral oil (1.09 g, 27.4 mmol, 2 equiv.) in dry DMF (25 mL) at 0 °C (ice bath) under an argon atmosphere. The suspension was stirred at room temperature till the mixture became clear (~15 min), then 5-bromo-2-chloropyrimidine (2.65 g, 13.7 mmol) was added. The reaction mixture was heated at 70 °C (oil bath) for 21 h, then cooled to room temperature and saturated aqueous NH_4Cl (40 mL) was added. The pH was adjusted to 3 by addition of a 1 M HCl and the aqueous phase was extracted with EtOAc (3 × 50 mL). The combined organic layers were extracted with saturated aqueous NaCl (50 mL), dried with Na_2SO_4, and concentrated under reduced pressure. The residue was purified by silica gel column chromatography using a linear gradient

(10-20%) of EtOAc in petrol ether to afford the title compound as a light-yellow oil (4.0 g, 84% yield). R_f = 0.40 (hexanes/EtOAc = 9:1). HRMS (ESI/Q-TOF) m/z: $[M + Na]^+$ calcd for $C_{13}H_{17}BrN_2O_4Na$, 367.0269; found 367.0268. Found 367.0268, error = -0.3 ppm. IR (neat, cm^{-1}) 3467, 2981, 2936, 1753, 1545, 1458, 1425, 1369, 1302, 1252, 1140. ^1H NMR (300 MHz, $CDCl_3$, ppm) δ: 8.79 (2H, s), 4.97 (1H, s), 4.28 (2H, q, *J* = 7.0 Hz), 1.48 (9H, s), 1.29 (3H, t, *J* = 7.0 Hz). $^{13}C\{^1H\}$ NMR (100.6 MHz, $CDCl_3$, ppm) δ: 166.6, 165.4, 162.0, 158.2, 119.5, 83.2, 62.2, 62.1, 28.0, 14.2.

***tert*-Butyl 2-(5-bromopyrimidin-2-yl)acetate (18).** A solution of NaOH (1.35 g, 33.8 mmol, 2 equiv.) in water (30 mL) was added to solution of 1-(*tert*-butyl)-3-ethyl-2-(5-bromopyrimidin-2-yl)malonate (5.6 g, 16.9 mmol) in MeOH (90 mL). The mixture was stirred at room temperature for 5.5 h. An aqueous 1 M HCl solution was added till pH 5 and the mixture was extracted with Et_2O (3 × 50 mL). The combined organic phases were extracted with brine (20 mL), dried with Na_2SO_4 and concentrated under reduced pressure to afford the title compound **18** (3.18 g, 69% yield) as an off-white solid. R_f = 0.31 (hexanes/EtOAc = 9:1). Anal. calcd for $C_{10}H_{13}BrN_2O_4$: C, 43.98; H, 4.80; N, 10.26. Found: C, 43.92; H, 4.79; N, 10.14. IR (neat, cm^{-1}) 3441, 3017, 2973, 2943, 1902, 1733, 1729, 1544, 1436, 1368, 1281, 1252, 1196, 1153. ^1H NMR (300 MHz, $CDCl_3$, ppm) δ: 8.76 (2H, s), 3.91 (2H, s), 1.46 (9H, s). $^{13}C\{^1H\}$ NMR (100.6 MHz, $CDCl_3$, ppm) δ: 168.7, 163.3, 158.1, 118.8, 81.9, 46.0, 28.2. The ^1H data were in agreement with the previously published data.[41]

***tert*-Butyl 2-(5-(2-ethoxy-2-oxoethyl)pyrimidin-2-yl)acetate (19).** Preparation of Zn reagent for reaction: Zn dust (7.19 g, 110.0 mmol, 2 equiv.) was placed in flask under an argon atmosphere and heated for 10 min at 90 °C (oil bath). THF (40 ml) and

TMSCl (0.7 mL, 5.5 mmol, 0.1 equiv.) were added sequentially. The mixture was activated by heating for 10 min (oil bath, 40-45 °C). Ethyl bromoacetate (6.1 mL, 55.0 mmol) was dropwise added over 40 min at 20-25 °C. The grey suspension (Zn dust in yellow solution) was cooled to room temperature and THF (70 ml) was added. The mixture was stirred at room temperature 30 min, then stored at +4 °C as a ~0.5 M stock solution in THF. The solution of Zn reagent (0.47 M, 34 mL, 2.5 equiv.) was added to a solution of *tert*-butyl 2-(5-bromopyrimidin-2-yl)acetate **18** (1.76 g, 6.44 mmol), $Pd_2(dba)_3$ (177 mg, 0.19 mmol, 3 mol%), XPhos (184 mg, 0.38 mmol, 6 mol%) in dry THF (20 mL). The reaction mixture was heated at 65 °C (oil bath) for 4 h, cooled to room temperature, diluted with EtOAc (60 mL), and quenched with saturated aqueous NH_4Cl (10 mL). The organic phase was separated and extracted with saturated aqueous NaCl (15 mL), dried with Na_2SO_4, and evaporated under reduced pressure. The crude product was purified by silica gel column chromatography using a linear gradient (0-50%) of EtOAc in petrol ether to afford the title compound **19** (1.51 g, 64% yield) as a yellow oil that solidified overnight. R_f = 0.46 (hexanes/EtOAc = 1:1). HRMS (ESI/Q-TOF) m/z: $[M + Na]^+$ calcd for $C_{14}H_{20}N_2O_4Na$, 303.1321; found 303.1329. IR (neat, cm^{-1}) 3452, 2981, 2934, 1741, 1594, 1557, 1449, 1369, 1287, 1258, 1152. ^1H NMR (400 MHz, $CDCl_3$, ppm) δ: 8.65 (2H, s), 4.18 (2H, q, *J* = 7.2 Hz), 3.94 (2H, s), 3.60 (2H, s), 1.46 (9H, s), 1.27 (3H, t, *J* = 7.1 Hz). $^{13}C\{^1H\}$ NMR (100.6 MHz, $CDCl_3$, ppm) δ: 169.9, 169.1, 163.8, 157.8, 125.5, 81.6, 61.7, 46.3, 35.9, 28.2, 14.3.

2-(5-(2-Ethoxy-2-oxoethyl)pyrimidin-2-yl)acetic acid (20). $ZnBr_2$ (11.0 g, 48.8 mmol, 7.2 equiv.) was added to solution of *tert*-butyl 2-(5-(2-ethoxy-2-oxoethyl)pyrimidin-2-yl)acetate **19** (1.9 g, 6.8 mmol) in dry CH_2Cl_2 (30 mL). After 17 h at

room temperature, water (10 mL) was added and the mixture was stirred for 10 min. Then saturated aqueous NaCl (10 mL) was added and the water phase was extracted with CH_2Cl_2 (3 × 50 mL). The organic phases were combined, dried with Na_2SO_4, and evaporated under reduced pressure to afford the title compound **20** (1.21 g, 69% yield) as a light brown solid. HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_{10}H_{13}N_2O_4$, 225.0875; found 225.0881. IR (neat, cm^{-1}) 3420, 3140, 2922, 2852, 2261, 2127, 1734, 1630, 1399, 1179. ^1H NMR (300 MHz, DMSO-d_6, ppm) δ: 8.69 (2H, s), 4.12 (2H, q, *J* = 7.1 Hz), 3.81 (2H, s) 3.80 (2H, s), 1.21 (3H, t, *J* = 7.1 Hz). $^{13}C\{^1H\}$ NMR (100.6 MHz, DMSO-d_6, ppm) δ: 170.3, 158.0, 126.0, 60.8, 34.3, 14.0.

Ethyl 2-(2-(2-oxo-2-(pyridin-2-ylamino)ethyl)pyrimidin-5-yl)acetate (21a). 2-(5-(2-Ethoxy-2-oxoethyl)pyrimidin-2-yl)acetic acid **20** (1.28 g, 5.7 mmol), 2-aminopyridine (536 mg, 5.7 mmol), HATU (2.16 g, 5.7 mmol) were dissolved in dry CH_2Cl_2 (25 mL) under an argon atmosphere and *i*-Pr_2NEt (1.97 mL, 11.4 mmol, 2 equiv.) was added. After 3 h at room temperature mixture was diluted with CH_2Cl_2 (100 mL), extracted with 10% aqueous citric acid (2 × 15 mL) and saturated aqueous $NaHCO_3$ (15 mL), dried with Na_2SO_4, and evaporated under reduced pressure. The crude product was purified by silica gel column chromatography using a linear gradient of EtOAc (30-80%) in petrol ether to afford the title compound **21a** (650 mg, 64% yield) as a light brown oil. R_f = 0.25 (EtOAc). HRMS (ESI/Q-TOF) m/z: $[M+H]^+$ calcd for $C_{15}H_{17}N_4O_3$, 301.1301; found 301.1314. IR (neat, cm^{-1}) 3242, 3194, 3004, 2981, 1717, 1578, 1559, 1527, 1338, 1302, 1202. ^1H NMR (400 MHz, $CDCl_3$, ppm) δ: 9.98 (1H, s), 8.71 (2H, s), 8.28 (1H, ddd, *J* = 4.9, 1.9, 1.0 Hz), 8.22 (1H, d, *J* = 8.5 Hz), 7.68 (1H, ddd, *J* = 7.4, 1.9, 0.5 Hz), 7.01 (1H, ddd, *J* = 7.4, 4.9, 1.0 Hz), 4.19 (2H, q, *J* = 7.1 Hz), 4.13

(2H, s), 3.62 (2H, s), 1.27 (3H, t, *J* = 7.1 Hz). $^{13}C\{^{1}H\}$ NMR (100.6 MHz, $CDCl_3$, ppm) δ: 169.7, 166.6, 163.6, 158.0, 151.5, 148.0, 138.4, 126.1, 119.9, 114.4, 61.8, 47.4, 35.8, 14.3.

Ethyl 2-(2-(2-((4-methoxypyridin-2-yl)amino)-2-oxoethyl)pyrimidin-5-yl)acetate (21b). 2-(5-(2-Ethoxy-2-oxoethyl)pyrimidin-2-yl)acetic acid **20** (1.40 g, 6.24 mol), 2-amino-4-methoxypyridine (775 mg, 6.24 mmol) and HATU (2.37 g, 6.24 mmol) were dissolved in dry CH_2Cl_2 (60 mL) under an argon atmosphere, then *i*-Pr_2NEt (2.15 mL, 12.48 mmol, 2 equiv.) was added. After 1 h at room temperature, the mixture was diluted with CH_2Cl_2 (50 mL), extracted with 10% citric acid (2 × 15 mL), and saturated aqueous $NaHCO_3$ (20 mL). The organic layer was dried with Na_2SO_4 and evaporated under reduced pressure. The product was purified by silica gel column chromatography using a linear gradient (50-100%) of EtOAc in petrol ether to afford the title compound **21b** (695 mg, 34% yield) as a brown oil. R_f = 0.28 (EtOAc). HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_{16}H_{19}N_4O_4$, 331.1406; found 331.1406. IR (neat, cm^{-1}) 2992, 2656, 1714, 1637, 1572, 1402, 1219, 1193, 1137. ^{1}H NMR (400 MHz, $CDCl_3$, ppm) δ: 9.89 (1H, s), 8.71 (2H, s), 8.08 (1H, d, *J* = 5.8 Hz), 7.86 (1H, d, *J* = 2.4 Hz), 6.58 (1H, dd, *J* = 5.8, 2.4 Hz), 4.19 (2H, q, *J* = 7.1 Hz), 4.12 (2H, s), 3.86 (3H, s), 3.62 (2H, s), 1.28 (3H, t, *J* = 7.1 Hz). $^{13}C\{^{1}H\}$ NMR (100.5 MHz, $CDCl_3$, ppm) δ: 169.7, 167.6, 166.8, 163.5, 158.0, 153.1, 148.6, 126.1, 107.9, 99.0, 61.8, 55.9, 55.5, 47.4, 35.8, 14.3.

2-(2-(2-Oxo-2-(pyridin-2-ylamino)ethyl)pyrimidin-5-yl)acetic acid (22a). A 3 M solution of NaOH (138 mg, 3.46 mmol, 2 equiv.) in water was added to a solution of ethyl 2-(2-(2-oxo-2-(pyridin-2-ylamino)ethyl)pyrimidin-5-yl)acetate **21a** (520 mg, 1.73 mmol) in EtOH (3 mL). The reaction mixture was stirred at room temperature for 2 h.

Amberlite was added till pH ~3-4, the mixture was filtered, the solids were washed with a mixture of water and EtOH (1:1, 3 × 15 mL), and the combined filtrates were evaporated under reduced pressure to afford the title compound **22a** (350 mg, 75% yield) as a yellow foam. HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_{13}H_{13}N_4O_3$, 273.0988; found 273.0994. IR (neat, cm^{-1}) 3130, 2412, 1942, 1702, 1586, 1545, 1437, 1399, 1312, 1188, 1164. ^1H NMR (400 MHz, DMSO-d6, ppm) δ: 13.1 – 12.2 (1H, br s), 10.74 (1H, s), 8.65 (2H, s), 8.33 (1H, ddd, *J* = 4.9, 2.0, 1.0 Hz), 8.07 (1H, d, *J* = 8.4 Hz), 7.81 – 7.73 (1H, m), 7.11 (1H, ddd, *J* = 7.3, 4.9, 1.0 Hz), 4.07 (2H, s), 3.67 (2H, s). $^{13}C\{^1H\}$ NMR (100.6 MHz, DMSO-d6, ppm) δ: 171.8, 168.1, 163.6, 157.9, 151.9, 148.0, 138.2, 126.7, 119.5, 113.4, 46.7, 34.7.

2-(2-(2-((4-Methoxypyridin-2-yl)amino)-2-oxoethyl)pyrimidin-5-yl)acetic acid (22b). A 3 M solution of NaOH (168 mg, 4.20 mmol, 2 equiv.) in water (1.4 mL) was added to the solution of ethyl 2-(2-(2-((4-methoxypyridin-2-yl)amino)-2-oxoethyl)pyrimidin-5-yl)acetate **21b** (695 mg, 2.10 mmol) in EtOH (10 mL). The reaction mixture was stirred at room temperature for 2 h. Amberlite was added till pH ~5, the mixture was filtered, the solids were washed with EtOH (3 × 10 mL), and the combined filtrates were evaporated under reduced pressure. The crude product was dissolved in CH_2Cl_2/EtOH (9:1, 2 mL) and filtered through a pad of silica gel eluting with mixture of CH_2Cl_2/EtOH (9:1, 100 mL) to afford the title compound **22b** (430 mg, 64% yield) as a light yellow solid. HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_{14}H_{15}N_4O_4$, 303.1093; found 303.1101. IR (neat, cm^{-1}) 3396, 3030, 2407, 1703, 1630, 1583, 1446, 1313, 1212, 1210. ^1H NMR (400 MHz, DMSO-d6, ppm) δ: 12.8 – 12.5 (1H, s), 10.71 (1H, s), 8.65 (2H, s), 8.14 (1H, d, *J* = 5.8 Hz), 7.71 (1H, d, *J* = 2.4 Hz), 6.72 (1H, dd, *J* = 5.8, 2.4 Hz),

4.06 (2H, s), 3.68 (2H, s). $^{13}C\{^{1}H\}$ NMR (100.6 MHz, DMSO-d6, ppm) δ: 171.9, 168.3, 166.6, 163.6, 157.9, 153.5, 149.0, 126.6, 106.5, 98.6, 55.3, 46.7, 34.6.

Allyl ester of M2R PNA monomer (23a). 2-(2-(2-Oxo-2-(pyridin-2-ylamino)ethyl)pyrimidin-5-yl)acetic acid **22a** (350 mg, 1.30 mmol, 1.1 equiv.), PNA backbone **7** (450 mg, 1.18 mmol) and HOOBt (250 mg, 1.54 mmol, 1.3 equiv.) were dissolved in anhydrous DMF (8 mL) under an argon atmosphere. The solution was cooled in ice bath and DCC (341 mg, 1.65 mmol, 1.4 equiv.) was added. After 10 min, the ice bath was removed and the solution was stirred overnight (18 h) at room temperature. The reaction mixture was evaporated, the residue was dissolved in EtOAc (80 mL) and extracted with saturated aqueous $NaHCO_3$ (20 mL). The organic layer was dried with Na_2SO_4 and evaporated under reduced pressure. The crude product was purified by reverse phase column chromatography using a linear gradient (0-50%) of MeCN in water to afford the title compound **23a** (271 mg, 36% yield) as an off-white foam. HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_{35}H_{35}N_6O_6$, 635.2618; found 635.2630. IR (neat, cm^{-1}) 3316, 3018, 2948m 1713, 1645, 1529, 1434, 1247, 1188. ^{1}H NMR (300 MHz, $CDCl_3$, ppm) (mixture of rotamers) δ: 10.02 (1H, s), 8.62 (2H, d, *J* = 4.1 Hz), 8.31 – 8.15 (2H, m), 7.75 (2H, d, *J* = 7.5 Hz), 7.71 – 7.63 (1H, m), 7.57 (2H, d, *J* = 7.4 Hz), 7.39 (2H, t, *J* = 7.4 Hz), 7.30 (2H, t, *J* = 7.6 Hz), 7.04 – 6.96 (1H, m), 6.00 – 5.81 (1H, m), 5.72 (0.6H, t, *J* = 5.8 Hz), 5.44 – 5.18 (2.4H, m), 4.73 – 4.62 (2H, m), 4.51-4.34 (2H, m), 4.26 – 3.99 (3H, m), 4.10 (2H, s), 3.70 – 3.30 (6H, m). $^{13}C\{^{1}H\}$ NMR (100.6 MHz, $CDCl_3$, ppm) (mixture of rotamers) δ: 170.0, 169.8, 166.6, 163.3, 158.3, 158.2, 156.8, 151.5, 147.9, 144.0, 143.8, 141.5, 138.4, 131.4, 128.0, 127.9, 127.3, 126.6, 125.0, 120.2, 119.9, 119.4, 114.4, 67.1, 66.9, 66.4, 49.7, 49.4, 47.3, 39.6, 33.9.

M2R PNA monomer (24a). $Pd(PPh_3)_4$ (43 mg, 0.04 mmol, 0.1 equiv.) and *N*-methylaniline (82 µL, 0.76 mmol 2 equiv.) were added to a solution of allyl ester of **M2R** PNA monomer **23a** (240 mg, 0.38 mmol) in anhydrous THF (10mL) under an argon atmosphere. The solution was stirred overnight (18 h) at room temperature. The solvent was evaporated under reduced pressure and the residue was purified by reverse phase column chromatography using a linear gradient (0-40%) of MeCN in water to afford the title compound **24a** (154 mg, 68% yield) as an off-white foam. HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_{32}H_{31}N_6O_6$, 595.2305; found 595.2320. IR (neat, cm^{-1}) 3116, 1700, 1654, 1441, 1399, 1190, 1148. ^1H NMR (400 MHz, DMSO-d6, ppm) (mixture of rotamers) δ: 13.0 – 12.5 (1H, br s), 10.74 (1H, s), 8.55 (2H, d, *J* = 5.6 Hz), 8.32 (1H, dd, *J* = 4.9, 2.0 Hz), 8.07 (1H, d, *J* = 8.4 Hz), 7.87 (2H, d, *J* = 7.6 Hz), 7.81 – 7.55 (5H, m), 7.49 – 7.23 (5H, m), 7.10 (1H, dd, *J* = 7.4, 4.9 Hz), 4.38 – 4.10 (4H, m), 4.07 (1H, s), 4.06 (1H, s), 3.96 (1H, s), 3.80 (1.3H, s), 3.64 (0.7H, s), 3.53 – 3.33 (2H, m), 3.27 – 3.07 (2H, m). $^{13}C\{^1H\}$ NMR (100.6 MHz, DMSO-d6, ppm) (mixture of rotamers) δ: 171.2, 170.6, 170.3, 169.7, 168.4, 163.3, 163.2, 158.5, 158.2, 158.0, 157.8, 156.4, 151.5, 147.1, 143.9, 143.8, 140.8, 140.7, 139.0, 127.6, 127.4, 127.1, 125.2, 125.0, 120.1, 119.6, 113.7, 65.4, 47.9, 47.7, 46.7, 33.1, 32.8, 30.7.

Allyl ester of ^{MeO}M2R PNA monomer (23b). 2-(2-(2-((4-methoxypyridin-2-yl)amino)-2-oxoethyl)pyrimidin-5-yl)acetic acid **22b** (400 mg, 1.32 mmol), PNA backbone **7** (503 mg, 1.32 mmol) and HOBt (280 mg, 1.72 mmol, 1.3 equiv.) were dissolved in anhydrous DMF (10 mL) under an argon atmosphere. The solution was cooled in ice bath and DCC (382 mg, 1.85 mmol, 1.4 equiv.) was added. After 10 min, the ice bath was removed and the solution was stirred overnight (18h) at room

temperature. The reaction mixture was evaporated under reduced pressure. The residue was dissolved in EtOAc (80 mL) and extracted with saturated aqueous $NaHCO_3$ (20 mL). The organic layer was dried with Na_2SO_4 and evaporated under reduced pressure. The product was purified by silica gel column chromatography using a linear gradient (0-5%) of EtOH in EtOAc to afford the title compound **23b** (220 mg, 25% yield) as an off-white foam. R_f = 0.21 (5% MeOH in CH_2Cl_2). HRMS (ESI/Q-TOF) m/z: [M + H]$^+$ calcd for $C_{36}H_{37}N_6O_7$, 665.2724; found 665.2719. IR (neat, cm^{-1}) 3279, 3012, 2947, 1706, 1652, 1575, 1538, 1532, 1447, 1444, 1436, 1320, 1247, 1202. ^{1}H NMR (300 MHz, $CDCl_3$, ppm) (mixture of rotamers) δ: 9.95 (1H, s), 8.62 (2H, d, *J* = 3.2 Hz), 8.05 (1H, dd, *J* = 5.8, 4.0 Hz), 7.86 (1H, t, *J* = 2.3 Hz), 7.75 (2H, d, *J* = 7.5 Hz), 7.57 (2H, d, *J* = 7.4 Hz), 7.39 (2H, t, *J* = 7.5 Hz), 7.30 (2H, t, *J* = 7.4 Hz), 6.58 – 6.53 (1H, m), 6.00 – 5.83 (1H, m), 5.72 (0.6H, t, *J* = 5.8 Hz), 5.41 – 5.20 (2.4H, m), 4.70 – 4.61 (2H, m), 4.50 – 4.34 (2H, m), 4.25 – 4.02 (3H, m), 4.09 (2H, s), 3.85 (3H, s), 3.71 – 3.28 (6H, m).

^{MeO}M2R PNA monomer (24b). $Pd(PPh_3)_4$ (38 mg, 0.03 mmol, 0.1 equiv.) and *N*-methylaniline (72 µL, 0.66 mmol, 2 equiv.) were added to a solution of allyl ester of **^{MeO}M2R** PNA monomer **23b** (220 mg, 0.33 mmol) in anhydrous THF (10mL) and the solution was stirred overnight (18 h) at room temperature. The solvent was evaporated under reduced pressure and the residue was purified using reverse phase chromatography using a linear gradient (0-45%) of MeCN in water to afford the title compound **24b** (103 mg, 50% yield) as an off-white foam. HRMS (ESI/Q-TOF) m/z: [M + H]$^+$ calcd for $C_{33}H_{33}N_6O_7$, 625.1469; found 625.1460. ^{1}H NMR (400 MHz, DMSO-d6, ppm) (mixture of rotamers) δ: 13.6 – 12.2 (1H, br s), 10.71 (1H, s), 8.55 (2H, d, *J* = 6.0 Hz), 8.13 (1H, d, *J* = 5.8 Hz), 7.91 – 7.84 (2H, m), 7.75 – 7.62 (3H, m), 7.54 – 7.15 (5H,

m), 6.71 (1H, dd, *J* = 5.8, 2.4 Hz), 4.40 – 4.10 (4H, m), 4.09 – 3.94 (3H, m), 3.80 (1H, s), 3.79 (3H, s), 3.64 (s, 1H), 3.52 – 3.32 (2H, m), 3.30-3.09 (2H, m). $^{13}C\{^{1}H\}$ NMR (100.6 MHz, DMSO-d6, ppm) (mixture of rotamers) δ: 170.2, 169.6, 168.3, 166.6, 163.4, 157.9, 157.8, 156.3, 153.5, 149.0, 143.9, 143.8, 140.72, 140.68, 133.9, 133.8, 127.6, 127.4, 127.3, 127.0, 125.1, 125.0, 120.09, 120.06, 106.4, 98.6, 65.4, 55.2, 47.9, 46.7, 32.8, 30.7.

1-Methyluracil (25) TMSCl (5.7 mL, 45.0 mmol, 0.56 equiv.) was added to a suspension of uracil (9.0 g, 80.3 mmol) in hexamethyldisilazide (HMDS, 59.0 mL, 281.0 mmol, 3.5 equiv.) under an argon atmosphere. The white suspension was stirred under reflux (oil bath 130 ºC) until it became a clear solution (3 h). Excess HMDS was evaporated under reduced pressure, the residue was suspended in DCE (20 mL), and MeI (21.7 mL, 348.5 mmol, 4.34 equiv.) was added. The reaction was refluxed (oil bath 60 ºC) for 19 h. The solvent was evaporated under reduced pressure and the residue was suspended in *i*-PrOH (50 mL). The precipitate was filtered, washed with *i*-PrOH (15 mL) and crystallized from water (30 mL) to afford the title compound **25** (6.0 g, 59% yield) as off-white crystals. R_f = 0.16 (EtOAc). ^{1}H NMR (300 MHz, DMSO-d6, ppm) δ: 11.20 (1H, s), 7.61 (1H, d, J = 7.8 Hz), 5.51 (1H, d, J = 7.8 Hz), 3.22 (3H, s). The ^{1}H NMR was in agreement with previously published data.[60]

1-Methyl-4-(1,2,4-triazolyl)uracil (26). $POCl_3$ (2.22 mL, 23.8 mmol, 2 equiv.) and NEt_3 (16.6 mL, 118.9 mmol, 10 equiv.) were added to an ice-cold stirring solution of 1,2,4-triazole (6.57 g, 95.2 mmol, 8 equiv.) in anhydrous MeCN (30 mL) under an argon atmosphere. A solution of 1-methyluracil[60] (1.5 g, 11.9 mmol) in a mixture of dry MeCN/CH_2Cl_2 (1:1, 30 mL) was added dropwise to the reaction mixture. The reaction

mixture was stirred at room temperature for 23 h. Precipitates were filtered and washed with a mixture of MeCN/CH_2Cl_2 (1:1, 2 × 30 mL) to afford the title compound **26** (1.54 g, 71% yield) as an off-white solid. HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_7H_8N_5O$, 178.0729; found 178.0733. IR (neat, cm^{-1}) 3113, 1680, 1554, 1476, 1424, 1399, 1341, 1287, 1156, 1127. ^1H NMR (400 MHz, DMSO-d6, ppm) δ: 9.41 (1H, s), 8.46 (1H, d, *J* = 7.0 Hz), 8.39 (1H, s), 6.92 (1H, d, *J* = 7.0 Hz), 3.50 (3H, s); (400 MHz, D_2O, ppm) δ 9.56 (1H, s), 8.58 (1H, d, *J* = 7.0 Hz), 8.53 (1H, s), 7.37 (1H, d, *J* = 7.0 Hz), 3.87 (3H, s).

Ethyl 2-(6-((2-((*tert*-butoxycarbonyl)amino)ethyl)amino)pyridin-3-yl)acetate. Molecular sieves (3 Å) and AcOH (795 µL, 13.87 mmol, 2.5 equiv.) were added to a solution of the ethyl 2-(6-aminopyridin-3-yl)acetate **13** (1.0 g, 5.55 mmol) and *N*-Boc-2-aminoacetaldehyde (1.81 g, 11.43 mmol, 2.06 equiv.) in dry MeOH (20 mL). After 3 h, $NaBH_3CN$ (691 mg, 11.1 mmol, 2 equiv.) was added and the reaction mixture was stirred overnight at room temperature. The mixture was diluted with EtOAc (20 mL), filtered, the solids were washed with EtOAc (2 ×10 mL), and the combined filtrates were evaporated under reduced pressure. Purification by silica gel column chromatography using a linear gradient of EtOAc (50-70%) in petrol ether gave the title compound (1.05 g, 58% yield) as a yellow oil. R_f = 0.57 (hexanes/EtOAc = 3:7). ^1H NMR (300 MHz, $CDCl_3$, ppm) δ: 7.80 (1H, s), 7.44 – 7.34 (1H, m), 6.48 (1H, d, *J* = 8.8 Hz), 5.67 (1H, s), 5.03 (1H, s), 4.10 (2H, t, *J* = 7.2 Hz), 3.43 – 3.28 (6H, m), 1.38 (9H, s), 1.19 (3H, t, *J* = 7.2 Hz). $^{13}C\{^1H\}$ NMR (100.6 MHz, $CDCl_3$, ppm) δ: 171.7, 157.7, 156.5, 147.1, 139.1, 118.6, 107.5, 79.5, 61.1, 42.6, 40.6, 37.7, 28.5, 14.3.

Ethyl 2-(6-((*tert*-butoxycarbonyl)(2-((*tert*-butoxycarbonyl)(1-methyl-2-oxo-1,2-dihydropyrimidin-4-yl)amino)ethyl)amino)pyridin-3-yl)acetate (28). HCl in dry

Et_2O (30.8 mL, 61.84 mmol, 20 equiv.) was added to a cold (ice bath) solution of ethyl 2-(6-((2-((tert-butoxycarbonyl)amino)ethyl)amino)pyridin-3-yl)acetate (1.0 g, 3.09 mmol) in CH_2Cl_2 (40 mL) and the reaction mixture was stirred at room temperature for 3 h 30 min. Solvent was evaporated under reduced pressure and the residue was co-evaporated with dry toluene (3 × 15 mL) to afford compound **27** (1.3 g, 93% yield) as a yellow oil, which was used without further purification in the next step.

1-Methyl-4-(1,2,4-triazolyl)uracil **26** (748 mg, 4.22 mmol), K_2CO_3 (1.75 g, 12.66 mmol, 3 equiv.) and NEt_3 (2.94 mL, 21.10 mmol, 5 equiv.) were added to solution of **27** (1.25 g, 4.22 mmol) in dry pyridine (20 mL). The reaction mixture was stirred at 60 ^oC (oil bath) for 88 h. The reaction mixture was filtered, the solids were washed with MeCN (20 mL), and the combined filtrates were evaporated under reduced pressure to afford crude product (1.05 g) as a brown oil that was used directly in the next step. A small sample for analytical purpose was purified by reverse phase column chromatography using a linear gradient of MeCN (10-35%) in water/MeCN. ^1H NMR (400 MHz, DMSO-d6, ppm) δ: 7.84 (1H, d, *J* = 2.4 Hz), 7.65 (1H, t, *J* = 5.3 Hz), 7.50 (1H, d, *J* = 7.2 Hz), 7.28 (1H, dd, *J* = 8.6, 2.4 Hz), 6.59 (1H, t, *J* = 5.3 Hz), 6.45 (1H, d, *J* = 8.6 Hz), 5.63 (d, *J* = 7.2 Hz, 1H), 4.06 (2H, q, *J* = 7.1 Hz), 3.45 (2H, s), 3.41 – 3.33 (4H, m), 3.20 (3H, s), 1.17 (3H, t, *J* = 7.1 Hz). $^{13}C\{^1H\}$ NMR (100.6 MHz, DMSO-d6, ppm) δ: 171.5, 164.1, 157.8, 156.3, 147.6, 145.6, 137.9, 137.8, 117.4, 107.9, 93.7, 60.2, 36.7, 36.6, 14.1.

DMAP (74 mg, 0.60 mmol, 0.2 equiv.) and Boc_2O (2.96 g, 13.58 mmol, 4.5 equiv.) were added to a solution of crude ethyl 2-(6-((2-((1-methyl-2-oxo-1,2-dihydropyrimidin-4-yl)amino)ethyl)amino)pyridin-3-yl)acetate (1.0 g, 3.02 mmol) in dry CH_2Cl_2 (40 mL). The reaction mixture was stirred at room temperature for 25 h. The

solvent was partly evaporated under reduced pressure, silica gel was added and the residual solvent was evaporated. The silica gel with absorbed crude product was added to a silica gel column and eluted with a linear gradient of EtOAc (50-100%) in petrol ether to afford the title compound **28** (653 mg, 41% yield over two steps) as a brown oil. R_f = 0.44 (CH_2Cl_2/EtOH = 19:1). HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_{26}H_{38}N_5O_7$, 532.2771; found 532.2792. IR (neat, cm^{-1}) 2979, 2933, 1733, 1668, 1629, 1482, 1389, 1342, 1226, 1153. ^1H NMR (400 MHz, $CDCl_3$, ppm) δ: 8.12 (1H, dd, *J* = 2.5, 0.8 Hz), 7.76 (1H, d, *J* = 8.6 Hz), 7.51 (1H, dd, *J* = 8.6, 2.5 Hz), 7.37 (1H, d, *J* = 7.5 Hz), 7.07 (1H, d, *J* = 7.5 Hz), 4.45 – 4.38 (2H, m), 4.33 – 4.27 (2H, m), 4.13 (2H, q, *J* = 7.1 Hz), 3.52 (2H, s), 3.48 (3H, s), 1.44 (9H, s), 1.32 (9H, s), 1.24 (3H, t, *J* = 7.1 Hz). $^{13}C\{^1H\}$ NMR (100.6 MHz, $CDCl_3$, ppm) δ: 171.1, 164.8, 156.5, 154.2, 153.7, 153.5, 147.5, 146.5, 137.8, 124.9, 118.3, 99.1, 83.0, 81.5, 61.2, 45.4, 45.2, 38.1, 38.0, 28.3, 27.9, 14.3.

2-(6-((*tert*-Butoxycarbonyl)(2-((*tert*-butoxycarbonyl)(1-methyl-2-oxo-1,2-dihydropyrimidin-4-yl)amino)ethyl)amino)pyridin-3-yl)acetic acid (29). A ~2 M solution of LiOH (95 mg, 3.95 mmol, 3 equiv.) in water (2 mL) was added to a solution of ethyl 2-(6-((*tert*-butoxycarbonyl)(2-((*tert*-butoxycarbonyl)(1-methyl-2-oxo-1,2-dihydropyrimidin-4-yl)amino)ethyl)amino)pyridin-3-yl)acetate **28** (700 mg, 1.32 mmol) in EtOH. The reaction mixture was stirred at room temperature for 2.5 h. The solvent was partially evaporated under reduced pressure and the residual mixture was acidified to pH~ 4-5 with 1 M HCl and extracted with EtOAc (4 × 15 mL). The organic layers were combined, extracted with saturated aqueous NaCl (15 mL), dried with Na_2SO_4, and evaporated under reduced pressure to afford the title compound **29** (530 mg, 80% yield)

as a brown oil of sufficient purity for use in next step. HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_{24}H_{34}N_5O_7$, 504.2458; found 504.2470. IR (neat, cm^{-1}) 3148, 2979, 1395, 1145. ^1H NMR (400 MHz, DMSO-d6, ppm) δ: 12.5 – 12,3 (1H, br s), 8.14 (1H, t, *J* = 1.6 Hz), 7.90 (1H, d, *J* = 7.4 Hz), 7.60 (2H, d, *J* = 1.6 Hz), 6.78 (1H, d, *J* = 7.4 Hz), 4.23 – 4,17 (4H, m), 3.54 (2H, s), 3.35 (3H, s), 1.37 (9H, s), 1.30 (9H, s). $^{13}C\{^1H\}$ NMR (100.6 MHz, DMSO-d6, ppm) δ: 172.3, 163.8, 155.2, 153.2, 152.7, 152.6, 148.5, 147.5, 138.2, 126.1, 118.0, 97.5, 82.3, 80.5, 44.6, 44.3, 37.1, 36.9, 27.7, 27.4.

Benzyl ester of M3 PNA monomer (31). Benzyl protected PNA backbone **30**[45] (470 mg, 1.09 mmol, 1.1 equiv.), 2-(6-((*tert*-butoxycarbonyl)(2-((*tert*-butoxycarbonyl)(1-methyl-2-oxo-1,2-dihydropyrimidin-4-yl)amino)ethyl)amino)pyridin-3-yl)acetic acid **29** (500 mg, 0.99 mmol), and HOOBt (194 mg, 1.19 mmol, 1.2 equiv.) were dissolved in dry DMF (15 mL) and DCC (225 mg, 1.09 mmol, 1.1 equiv.) was added under an argon atmosphere. The reaction mixture was stirred at room temperature for 20 h, diluted with EtOAc (50 mL), extracted with saturated aqueous $NaHCO_3$ (20 mL), and saturated aqueous NaCl (20 mL). The organic layers were dried with Na_2SO_4 and evaporated under reduced pressure. The crude residue was purified by silica gel column chromatography using a linear gradient of EtOAc (75-100%) in petrol ether followed by 4% EtOH in EtOAc to afford the title compound **31** (690 mg, 76% yield) as a yellow foam. R_f = 0.44 (EtOAc). HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_{50}H_{58}N_7O_{10}$, 916.4245; found 916.4252. IR (neat, cm^{-1}) 2979, 1717, 1652, 1526, 1471, 1389, 1224, 1151. ^1H NMR (400 MHz, DMSO-d6, ppm) (mixture of rotamers) δ: 8.10 (0.7H, d, *J* = 2.3 Hz), 8.03 (0.3H, d, *J* = 2.3 Hz), 7.90 – 7.84 (3H, m), 7.70 – 7.64 (2H, m), 7.58 – 7.26 (12H, m), 6.79 (1H, d, *J* = 7.4 Hz), 5.18 (0.5H, s), 5.13 (1.5H, s), 4.41 – 4.09 (9H, m),

3.71 (1.4H, s), 3.71 (0.6H, s), 3.50 – 3.34 (2H, m), 3.32 (3H, s), 3.25 – 3.12 (2H, m), 1.36 (9H, s), 1.30 (2.6H, s), 1.27 (6.4H, s). $^{13}C\{^{1}H\}$ NMR (100.6 MHz, DMSO-d6, ppm) (mixture of rotamers) δ: 170.8, 170.5, 169.2, 163.8, 156.4, 155.2, 153.2, 152.6, 152.4, 148.5, 147.5, 147.2, 143.9, 143.8, 140.74, 140.71, 138.2, 137.9, 135.8, 135.6, 128.5, 128.4, 128.2, 128.1, 128.0, 127.8, 127.6, 127.0, 126.6, 125.1, 125.0, 120.1, 117.9, 97.5, 82.2, 82.2, 80.5, 80.5, 66.5, 65.8, 65.4, 47.9, 47.8, 46.7, 44.6, 44.3, 37.1, 35.1, 27.7, 27.4, 27.3.

M3 PNA monomer (32). Pd/C (146 mg) was added to a solution of benzyl ester of **M3** PNA monomer **31** (630 mg, 0.69 mmol) in MeOH (40 mL). Hydrogen gas (1 atm) was bubbled through the reaction mixture for 2 h. The mixture was filtered through a pad of Celite and the pad was washed with MeOH (3 × 15 mL). The combined filtrates were evaporated under reduced pressure to afford the title compound **32** (381 mg, 67% yield) as a yellow foam. HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_{43}H_{52}N_7O_{10}$, 826.3776; found 826.3766. IR (neat, cm^{-1}) 3118, 2980, 1399, 1155. ^{1}H NMR (400 MHz, DMSO-d6, ppm) (mixture of rotamers) δ: 13.0 – 12.5 (1H, br s), 8.10 (0.6H, d, *J* = 2.2 Hz), 8.05 (0.4H, d, *J* = 2.2 Hz), 7.90 – 7.84 (3H, m), 7.67 (2H, d, *J* = 7.6 Hz), 7.58 – 7.28 (7H, m), 6.77 (1H, d, *J* = 7.4 Hz), 4.50 – 4.13 (7.7H, m), 3.96 (1.3H, s), 3.69 (1.2H, s), 3.55 (0.8H, s), 3.46 – 3.35 (2H, m), 3.33 (3H, s), 3.25 – 3.10 (2H, m), 1.36 (9H, s), 1.30 (3H, s), 1.28 (6H, s). $^{13}C\{^{1}H\}$ NMR (100.6 MHz, DMSO-d6, ppm) (mixture of rotamers) δ: 170.8, 170.3, 163.8, 156.3, 155.2, 153.2, 152.6, 152.4, 152.3, 148.5, 147.5, 147.3, 143.9, 143.8, 140.8, 140.7, 138.3, 138.0, 127.6, 127.0, 126.7, 125.1, 125.0, 120.1, 117.9, 97.6, 82.2, 82.2, 80.5, 80.5, 65.4, 54.9, 47.9, 47.6, 46.7, 44.5, 44.3, 37.1, 35.1, 27.7, 27.37, 27.35.

1-(Ethoxycarbonylmethyl)uracil (33). Ethyl chloroacetate (10.0 mL, 94.6 mmol, 1.06 equiv.) was added portion-wise to a suspension of uracil (10.0 g, 89.2 mmol) and K_2CO_3 (16.0 g, 116 mmol, 1.3 equiv.) in anhydrous DMF (50 mL) under an argon atmosphere. After addition, the resulting reaction mixture was vigorously stirred at 60 ^oC (oil bath) for 20 h. The mixture was cooled to room temperature, diluted with EtOAc (200 mL), and filtered. The solids were washed with EtOAc (2 × 30 mL) and the combined filtrates were evaporated under reduced pressure. The residue was purified by silica gel column chromatography using a linear gradient (50-100%) of EtOAc in petrol ether to afford the title compound **33** (4.49 g, 54% yield) as a white solid. R_f = 0.31 (hexanes/EtOAc = 3:7). ^{1}H NMR (300 MHz, DMSO-d6, ppm) δ: 11.39 (1H, s), 7.62 (1H, d, *J* = 7.9 Hz), 5.61 (1H, d, *J* = 7.9 Hz), 4.50 (2H, s), 4.15 (2H, q, *J* = 7.1 Hz), 1.21 (3H, t, *J* = 7.1 Hz). ^{13}C{^{1}H} NMR (100.6 MHz, DMSO-d6, ppm) δ: 168.2, 163.7, 151.0, 145.9, 101.1, 61.2, 48.6, 14.0. The ^{1}H and ^{13}C NMR data were in agreement with the previously published data.[46]

1-(Ethoxycarbonylmethyl)-4-(1,2,4-triazolyl)uracil (34). Under an argon atmosphere, $POCl_3$ (2.35 mL, 25.23 mmol, 2 equiv.) and NEt_3 (17.6 mL, 126.15 mmol, 10 equiv.) were added into an ice-cold stirring suspension of 1,2,4-triazole (6.97 g, 100.92 mmol, 8 equiv.) in anhydrous MeCN (50 mL). A solution of 1-(ethoxycarbonylmethyl)uracil (**33**) (2.5 g, 12.6 mmol) in a mixture of dry MeCN/CH_2Cl_2 (1:1, 20 mL) was added dropwise and the reaction mixture was stirred at room temperature for 2 h. The mixture was partly evaporated, diluted with EtOAc (70 mL) and washed with saturated aqueous $NaHCO_3$ (40 mL). The water phase was extracted with EtOAc (5 × 40 mL). The combined organic phases were extracted with saturated

aqueous NaCl (30 mL), dried over Na_2SO_4, and evaporated under reduced pressure. The residue was purified by filtration through a small pad of silica gel. Elution with EtOAc gave the title compound **34** (2.54 g, 81% yield) as a white solid. R_f = 0.34 (EtOAc). HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_{10}H_{12}N_5O_3$, 250.0940; found 250.0943. IR (neat, cm^{-1}) 3125, 2925, 2867, 1735, 1684, 1550, 1477. ^1H NMR (300 MHz, $CDCl_3$, ppm) δ: 9.29 (1H, s), 7.80 (1H, d, *J* = 7.1 Hz), 7.10 (1H, d, *J* = 7.1 Hz), 4.71 (2H, s), 4.29 (2H, q, *J* = 7.1 Hz), 1.32 (3H, t, *J* = 7.1 Hz). $^{13}C\{^1H\}$ NMR (100.6 MHz, $CDCl_3$, ppm) δ: 166.8, 160.0, 155.4, 154.2, 151.3, 143.6, 95.6, 62.7, 51.6, 14.2.

***N*^{1}-(pyridin-2-yl)ethane-1,2-diamine (35).** A solution of 2-chloropyridine (9.4 mL, 0.10 mol) and 1,2-diaminoethane (66.8 mL, 1.0 mol, 10 equiv.) was stirred at 130 °C (oil bath) for 23 h. The excess of 1,2-diaminoethane was evaporated under reduced pressure and the residue was dried in vacuum (40 °C, 0.1 mbar). The dried residue was dissolved in CH_2Cl_2 (150 mL), water (50 mL) was added, and the pH was adjusted to 11 with 40% aqueous NaOH. The water phase was separated and extracted with CH_2Cl_2 (2 × 100 mL). Combined organic phases were evaporated under reduced pressure to afford the title compound **35** (9.65 g, 70% yield) as a white solid. ^1H NMR (300 MHz, DMSO-d6, ppm) δ: 8.00 – 7.88 (1H, m), 7.39 – 7.27 (1H, m), 6.52 – 6.34 (3H, m), 3.19 (2H, q, *J* = 6.4 Hz), 2.67 (2H, t, *J* = 6.4 Hz), 1.7 – 1.2 (2H, br s). The ^1H NMR data were in agreement with the previously published data.[47]

Ethyl 2-(4-((tert-butoxycarbonyl)(2-((tert-butoxycarbonyl)(pyridin-2-yl)amino)ethyl)amino)-2-oxopyrimidin-1(2H)-yl)acetate (36). 1-(Ethoxycarbonylmethyl)-4-(1,2,4-triazolyl)uracil **34** (1.50 g, 6.02 mmol), *N*^{1}-(pyridin-2-yl)ethane-1,2-diamine **35** (990 mg, 7.22 mmol, 1.2 equiv.) and K_2CO_3 (2.50 g, 18.06

mmol, 3 equiv.) were mixed in MeCN (40 mL). The reaction mixture was stirred at room temperature for 6 h and filtered through pad of Celite. The pad was washed with MeCN (2 × 15 mL) and the combined filtrates were evaporated under reduced pressure. The residual yellow oil was dissolved in dry THF (40 mL), and DMAP (73 mg, 0.6 mmol 0.1 equiv.) and Boc_2O (6.55 g, 30 mmol 5 equiv.) were added. After 40 h at room temperature, the solvent was partly evaporated under reduced pressure, silica gel was added and the residual solvent was evaporated. The silica gel with absorbed crude product was added to a silica gel column and eluted with a linear gradient of EtOAc (40-70%) in petrol ether to afford the title compound **36** (1.32 g, 42% yield) as an off-white foam. R_f = 0.32 (hexanes/EtOAc = 3:7). HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_{25}H_{36}N_5O_7$, 518.2615; found 518.2627. IR (neat, cm^{-1}) 2980, 2950, 1726, 1668, 1630, 1471, 1369, 1224, 1150. ^1H NMR (400 MHz, $CDCl_3$, ppm) δ: 8.26 (1H, ddd, *J* = 4.9, 2.0, 0.9 Hz), 7.77 (1H, d, *J* = 8.5 Hz), 7.56 (1H, ddd, *J* = 8.5, 7.2, 2.0 Hz), 7.31 (1H, d, *J* = 7.6 Hz), 7.22 (1H, d, *J* = 7.6 Hz), 6.92 (1H, ddd, *J* = 7.2, 4.9, 0.9 Hz), 4.57 (2H, s), 4.46 – 4.40 (2H, m), 4.35 (2H, dt, *J* = 6.6, 2.3 Hz), 4.24 (2H, q, *J* = 7.1 Hz), 1.45 (9H, s), 1.34 (9H, s), 1.29 (3H, t, *J* = 7.1 Hz). $^{13}C\{^1H\}$ NMR (100.6 MHz, $CDCl_3$, ppm) δ: 167.6, 165.2, 155.7, 154.6, 154.2, 153.4, 147.3, 146.0, 136.9, 119.0, 118.8, 99.2, 83.3, 81.4, 62.1, 50.6, 45.3, 28.4, 27.9, 14.3.

2-(4-((tert-Butoxycarbonyl)(2-((tert-butoxycarbonyl)(pyridin-2-yl)amino)ethyl)amino)-2-oxopyrimidin-1(2H)-yl)acetic acid (37). ~2 M LiOH (116 mg, 4.83 mmol, 2 equiv.) in water (2.4 mL) was added to a cold solution of ethyl 2-(4-((tert-butoxycarbonyl)(2-((tert-butoxycarbonyl)(pyridin-2-yl)amino)ethyl)amino)-2-oxopyrimidin-1(2H)-yl)acetate **36** (1.25 g, 2.41 mmol) in EtOH (25 mL). After 1 h at 0 °C

(ice bath) Amberlite was added till pH~5. The mixture was filtered, the solids were washed with EtOH (3 × 10 mL), and the combined filtrates were evaporated under reduced pressure to afford the title compound **37** (1.16 g, 96% yield) as an off-white foam. HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_{23}H_{32}N_5O_7$, 490.2302; found 490.2300. IR (neat, cm^{-1}) 3111, 2980, 1405, 1147. ^1H NMR (300 MHz, DMSO-d6, ppm) δ: 14.0 – 12.5 (1H, br s), 8.35 – 8.24 (1H, m), 7.89 (1H, dd, *J* = 7.4, 2.5 Hz), 7.78 – 7.57 (2H, m), 7.06 (1H, ddd, *J* = 6.7, 4.8, 1.9 Hz), 6.89 (1H, dd, *J* = 7.4, 2.5 Hz), 4.50 (2H, s), 4.27 – 4.20 (4H, m), 1.37 (9H, s), 1.30 (9H, s). $^{13}C\{^1H\}$ NMR (100.6 MHz, DMSO-d6, ppm) δ: 169.3, 164.2, 154.7, 153.8, 153.3, 152.6, 148.5, 147.1, 137.1, 119.2, 118.6, 97.4, 82.4, 80.6, 50.4, 44.5, 31.3, 27.7, 27.3.

Benzyl ester of M3R PNA monomer (38). 2-(4-((*tert*-Butoxycarbonyl)(2-((*tert*-butoxycarbonyl)(pyridin-2-yl)amino)ethyl)amino)-2-oxopyrimidin-1(2*H*)-yl)acetic acid **37** (1.06 g, 2.17 mmol, 1.1 equiv.), benzyl protected PNA backbone **30** (0.85 g, 1.97 mmol) and HOOBt (354 mg, 2.17 mmol, 1.1 equiv.) were dissolved in anhydrous DMF (20 mL). The solution was cooled on ice and *N*,*N*'-dicyclohexylcarbodiimide (489 mg, 2.37 mmol, 1.2 equiv.) was added. The ice bath was removed and the solution was stirred overnight at room temperature. The reaction mixture was evaporated under reduced pressure, the residue was dissolved in EtOAc (70 mL) and extracted with saturated aqueous $NaHCO_3$ (2 × 20 mL). The organic layer was dried over Na_2SO_4 and evaporated under reduced pressure. The product was purified by silica gel column chromatography using a linear gradient of EtOAc (50-100%) in petrol ether to afford the title compound **38** (1.3 g, 73% yield) as a white foam. R_f = 0.34 (EtOAc). HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_{49}H_{56}N_7O_{10}$, 902.4089; found 902.4084. IR (neat, cm^{-1}) 3333, 2978, 2933, 1723, 1662,

1517, 1457, 1368, 1250, 1225, 1150. ^{1}H NMR (300 MHz, DMSO-d6, ppm) (mixture of rotamer) δ: 8.30 – 8.26 (1H, m), 7.88 (2H, d, *J* = 7.6 Hz), 7.77 – 7.61 (5H, m), 7.47 – 7.25 (10H, m), 7.08 – 7.01 (1H, m), 6.88 (1H, d, *J* = 7.5 Hz), 5.22 (0.6H, s), 5.13 (1.4H, s), 4.85 (1.4H, s), 4.76 (0.6H, s), 4.48 – 4.11 (9H, m), 3.56 – 3.34 (2H, m), 3.32 – 3.06 (2H, m), 1.36 (9H, s), 1.30 (9H, s). ^{13}C{^{1}H} NMR (100.6 MHz, DMSO-d6, ppm) (mixture of rotamer) δ: 169.4, 168.9, 167.6, 167.3, 164.1, 156.3, 154.7, 153.8, 153.3, 152.6, 148.8, 147.1, 143.9, 140.7, 140.7, 139.4, 137.4, 137.1, 135.8, 135.6, 128.9, 128.5, 128.4, 128.2, 128.1, 128.0, 127.9, 127.6, 127.3, 127.0, 125.1, 121.4, 120.1, 120.0, 119.2, 118.6, 109.7, 97.4, 82.4, 80.6, 66.6, 65.9, 65.5, 49.2, 48.0, 46.7, 44.5, 44.4, 27.7, 27.3.

M3R PNA monomer (39). Pd/C (297 mg) was added to solution of benzyl ester of **M3R** PNA monomer **38** (1.26 g, 1.40 mmol) in MeOH (25 mL). Hydrogen gas (1 atm) was bubbled through the reaction mixture for 2 h. The mixture was filtered through a pad of Celite and the pad was washed with MeOH (3 × 15 mL). The combined filtrates were evaporated under reduced pressure to afford the title compound **39** (1.0 g, 88% yield) as a white foam. HRMS (ESI/Q-TOF) m/z: $[M + H]^{+}$ calcd for $C_{42}H_{50}N_7O_{10}$ 812.3619; found 812.3615. IR (neat, cm^{-1}) 3139, 1399, 1385, 1149. ^{1}H NMR (300 MHz, DMSO-d6, ppm) (mixture of rotamer) δ: 13.3 – 12.6 (1H, br s), 8.30 – 8.26 (1H, m), 7.89 (2H, d, *J* = 7.4 Hz), 7.78 – 7.63 (5H, m), 7.45 – 7.27 (5H, m), 7.08 – 7.01 (1H, m), 6.87 (1H, t, *J* = 7.7 Hz), 4.82 (1H, s), 4.63 (1H, s), 4.40 – 4.07 (8H, m), 4.00 (1H, s), 3.54 – 3.31 (3H, m), 3.16 – 3.04 (1H, m), 1.36 (9H, s), 1.30 (9H, s). ^{13}C{^{1}H} NMR (100.6 MHz, DMSO-d6, ppm) (mixture of rotamers) δ: 170.5, 167.5, 167.0, 164.1, 164.0, 156.3, 154.7, 153.8, 153.3, 152.6, 148.9, 147.1, 143.9, 140.7, 140.7, 137.2, 127.6, 127.1,

125.1, 120.1, 119.2, 118.6, 97.4, 82.3, 82.3, 80.6, 65.5, 49.2, 47.8, 46.7, 44.5, 44.4, 27.7, 27.3.

***tert*-Butyl (2-((5-bromopyridin-2-yl)amino)-2-oxoethyl)carbamate (41a).** *i*-Pr_2NEt (4.0 mL, 23.1 mmol, 2 equiv.) was added to a mixture of Boc-protected glycine (2.13 g, 12.1 mmol, 1.05 equiv.), 2-amino-5-bromopyridine **40a** (2.0 g, 11.6 mmol) and HATU (4.84 g, 12.7 mmol, 1.1 equiv.) in DCM (20 mL). After 21 h at room temperature, the mixture was diluted with DCM (70 mL), extracted with 10% citric acid (50 mL), saturated aqueous $NaHCO_3$ (50 mL), and saturated aqueous NaCl (50 mL). The organic layers were dried with Na_2SO_4 and evaporated under reduced pressure. The crude product was purified by silica gel column chromatography using a linear gradient of EtOAc (0-50%) in petrol ether. The product obtained after column chromatography was contaminated with tetramethylurea and was further crystallized from DCM/MeOH to afford the title compound **41a** (2.50 g, 65% yield) as white crystals. mp = 156.3-157.1 °C. R_f = 0.68 (hexanes/EtOAc = 1:1). HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_{12}H_{17}BrN_3O_3$ 330.0453; found 330.0450. IR (neat, cm-1) 3246, 3055, 2976, 1707, 1681, 1570, 1531, 1456, 1369, 1250, 1166. ^1H NMR (300 MHz, $CDCl_3$, ppm) δ: 8.51 (1H, s), 8.33 (1H, d, *J* = 2.4 Hz), 8.14 (1H, d, *J* = 8.8 Hz), 7.80 (1H, dd, *J* = 8.8, 2.4 Hz), 5.15 (1H, s), 3.98 (2H, d, *J* = 6.0 Hz), 1.48 (9H, s). $^{13}C\{^1H\}$ NMR (100.6 MHz, $CDCl_3$, ppm) δ: 168.5, 156.2, 149.7, 148.9, 141.0, 115.4, 115.1, 81.0, 45.4, 28.4.

2-Amino-*N*-(5-bromopyridin-2-yl)acetamide hydrochloride. A solution of HCl in dry dioxane (15.0 mL, 60 mmol, 8 equiv.) was slowly added to a solution of *tert*-butyl (2-((5-bromopyridin-2-yl)amino)-2-oxoethyl)carbamate **41a** (2.45 g, 7.4 mmol) in dry dioxane (15 mL) at 0 °C (ice bath) under an argon atmosphere. After 20 h at room

temperature, the solvent was evaporated under reduced pressure to afford the title compound (1.97 g, quantitative yield) as a white solid. HRMS (ESI/Q-TOF) m/z: [M+H]+ Calcd. for $C_7H_9BrN_3O$ 229.9929; Found 229.9925. IR (neat, cm-1) 3132, 3015, 2654, 1717, 1636, 1572, 1570, 1481, 1399, 1395, 1385, 1219, 1193, 1137. ^{1}H NMR (300 MHz, DMSO-d6, ppm) δ 11.12 (1H, s), 8.52 – 8.47 (m, 1H), 8.28 (3H, s), 8.07 (1H, td, *J* = 8.8, 2.5 Hz) 8.07 – 7.97 (1H, m), 3.83 (2H, q, *J* = 5.7 Hz). ^{13}C{^{1}H} NMR (100.6 MHz, DMSO-d6, ppm) δ: 166.0, 150.2, 148.8, 141.0, 115.2, 114.1, 41.1.

***N*-(2-((5-bromopyridin-2-yl)amino)-2-oxoethyl)pyrimidine-2-carboxamide (42a).** HOBt (1.22 g, 9.0 mmol, 1.3 equiv.), pyrimidine-2-carboxylic acid (947 mg, 7.63 mmol, 1.1 equiv.) and DCC (1.98 g, 9.0 mmol, 1.3 equiv.) were added to a mixture of 2-amino-*N*-(5-bromopyridin-2-yl)acetamide hydrochloride (1.85 g, 6.94 mmol) and *i*-Pr_2NEt (3.6 mL, 20.8 mmol, 3 equiv.) in dry DCM (50 mL) under an argon atmosphere. After 18 h at room temperature, the mixture was diluted with DCM (50 mL), undissolved solid was filtered off, the filtrate was evaporated under reduced pressure, and the residue was purified by silica gel column chromatography using a linear gradient of EtOH (0-5%) in DCM to afford the title compound **42a** (2.10 g, 90% yield) as a white solid. HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_{12}H_{11}BrN_5O_2$ 336.0096; found 336.0092. ^{1}H NMR (300 MHz, DMSO-d6, ppm) δ: 10.81 (1H, s), 9.12 (1H, t, *J* = 6.0 Hz), 9.00 (2H, d, *J* = 4.9 Hz), 8.45 (1H, t, *J* = 1.6 Hz), 8.05 – 8.00 (2H, m), 7.72 (1H, t, *J* = 4.9 Hz), 4.19 (2H, d, *J* = 6.0 Hz). ^{13}C{^{1}H} NMR (100.6 MHz, DMSO-d6, ppm) δ: 168.3, 162.5, 157.8, 157.5, 150.7, 148.6, 140.8, 123.2, 115.0, 113.5, 43.3.

N-(5-bromopyridin-2-yl)-2-(pyrimidin-2-yl)oxazol-5-amine. NEt_3 (3.8 mL, 27.4 mmol, 4 equiv.) and suspension of *N*-(2-((5-bromopyridin-2-yl)amino)-2-

oxoethyl)pyrimidine-2-carboxamide (**42a**) (3.38 g, 6.84 mmol) in dry DCM (30 mL) were added to a solution of PPh_3 (3.59 g, 13.7 mmol, 2 equiv.) and I2 (3.47 g, 13.7 mmol, 2 equiv.) in dry DCM (60 mL). The reaction mixture was stirred at room temperature for 2 h. The mixture became dark and thick. Solid precipitates were filtered off and washed with DCM. Evaporation of solvents gave the title compound (1.38 g, 64% yield) as a light yellow solid. ^{1}H NMR (300 MHz, DMSO-d6, ppm) δ: 10.97 (1H, s), 8.90 (2H, d, *J* = 4.9 Hz), 8.37 (1H, d, *J* = 2.5 Hz), 7.87 (1H, dd, *J* = 8.8, 2.5 Hz), 7.51 (1H, t, *J* = 4.9 Hz), 7.26 (1H, s), 6.91 (1H, d, *J* = 8.8 Hz).

***tert*-Butyl (5-bromopyridin-2-yl)(2-(pyrimidin-2-yl)oxazol-5-yl)carbamate (43a).** *N*-(5-Bromopyridin-2-yl)-2-(pyrimidin-2-yl)oxazol-5-amine (317 mg, 1.0 mmol), Boc2O (326 mg, 1.49 mmol, 1.5 equiv.) and DMAP (12 mg, 0.1 mmol, 0.1 equiv.) were dissolved/suspended in dry THF (6 mL) under an argon atmosphere. After 14 h at room temperature, the solvent was partly evaporated under reduced pressure. The residue was absorbed on silica gel and purified by column chromatography using a linear gradient of EtOAc in petrol ether to afford the title compound **43a** (390 mg, 94% yield) as an off-white foam. HRMS (ESI/Q-TOF) m/z: $[M + Na]^+$ calcd for $C_{17}H_{16}BrN_5O_3Na$ 440.0334; found 440.0341. IR (neat, cm-1) 3457, 2130, 3049, 2981, 2933, 1738, 1617, 1564, 1543, 1460, 1416, 1370. ^{1}H NMR (300 MHz, CDCl3, ppm) δ: 8.89 (2H, d, *J* = 4.9 Hz), 8.35 (1H, dd, *J* = 2.5, 0.7 Hz), 7.83 (1H, dd, *J* = 8.8, 2.5 Hz), 7.69 (1H, dd, *J* = 8.8, 0.7 Hz), 7.36 (1H, t, *J* = 4.9 Hz), 7.26 (1H, s), 1.47 (9H, s). ^{13}C{^{1}H} NMR (100.6 MHz, CDCl3, ppm) δ: 157.9, 156.2, 154.7, 151.6, 151.4, 149.1, 146.7, 140.3, 124.7, 121.3, 120.0, 117.1, 84.1, 28.0.

Ethyl 2-(6-(2-((*tert*-butoxycarbonyl)amino)acetamido)pyridin-3-yl)acetate (41b). Boc-protected glycine (2.24 g, 12.82 mmol, 1.1 equiv), ethyl 2-(6-aminopyridin-3-yl)acetate **40b** (2.10 g, 11.65 mmol), HOBt (1.89 g, 13.98 mmol, 1.2 equiv.), EDC·HCl (2.90 g, 15.15 mmol, 1.3 equiv.) were dissolved in dry CH_2Cl_2 (40 mL) and *i*-Pr_2NEt (6.0 mL, 34.96 mmol, 3 equiv.) was added under an argon atmosphere. The reaction mixture was stirred at room temperature for 20 h, then diluted with CH_2Cl_2 (70 mL), extracted with 10% aqueous citric acid (40 mL), saturated aqueous $NaHCO_3$ (40 mL), and saturated aqueous NaCl (40 mL). The organic layers were dried with Na_2SO_4 and evaporated under reduced pressure. The crude product was purified by silica gel column chromatography using a linear gradient (10-50%) of EtOAc in petrol ether to afford the title compound **41b** (2.51 g, 44% yield) a as a yellow oil. R_f = 0.44 (hexanes/EtOAc = 1:1). HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_{16}H_{24}N_3O_5$ 338.1716; found 338.1707. IR (neat, cm^{-1}) 3275, 2979, 1699, 1593, 1528, 1398, 1303, 1172. ^1H NMR (300 MHz, $CDCl_3$, ppm) δ: 8.57 (1H, s), 8.22 – 8.13 (2H, m), 7.66 (1H, dd, *J* = 8.6, 2.3 Hz), 5.18 (1H, s), 4.16 (2H, q, *J* = 7.1 Hz), 3.98 (2H, d, *J* = 6.0 Hz), 3.58 (2H, s), 1.25 (3H, t, *J* = 7.1 Hz). $^{13}C\{^1H\}$ NMR (100.6 MHz, $CDCl_3$, ppm) δ: 170.9, 168.2, 150.0, 148.3, 139.5, 126.2, 113.9, 80.8, 61.3, 45.3, 38.1, 28.4, 14.3.

Ethyl 2-(6-(2-aminoacetamido)pyridin-3-yl)acetate hydrochloride. A solution of HCl in dry dioxane (13.0 mL, 51.87 mmol, 7 equiv.) was slowly added to a solution of ethyl 2-(6-(2-((*tert*-butoxycarbonyl)amino)acetamido)pyridin-3-yl)acetate **41b** (2.5 g, 7.41 mmol) in dry dioxane (15 mL) at 0 °C (ice bath). After 20 h at room temperature, the solvent was evaporated under reduced pressure to afford the title compound (2.0 g, quantitative yield) as a white solid. HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for

$C_{11}H_{16}N_3O_3$ 238.1192; found 238.1189. IR (neat, cm^{-1}) 3436, 2946, 1730,1356, 1598, 1490, 1426, 1339, 1235, 1183. ^1H NMR (400 MHz, DMSO-d6, ppm) δ: 11.35 (1H, s), 8.48 (3H, s), 8.26 (1H, dd, *J* = 2.2, 0.8 Hz), 8,00 – 7.92 (1H, m), 7.82 (1H, dd, *J* = 8.6, 2.4 Hz), 4.08 (2H, q, *J* = 7.1 Hz), 3.84 (2H, q, *J* = 5.8 Hz), 3,72 (2H, s), 1.18 (3H, t, *J* = 7.1 Hz). $^{13}C\{^1H\}$ NMR (100.6 MHz, DMSO-d6, ppm) δ: 170.8, 166.0, 149.4, 147.3, 140.7, 126.5, 113.5, 60.6, 41.0, 36.6, 14.1.

Ethyl 2-(6-(2-(pyrimidine-2-carboxamido)acetamido)pyridin-3-yl)acetate (42b). Pyrimidine-2-carboxylic acid (997 mg, 8.03 mmol, 1.1 equiv.), ethyl 2-(6-(2-aminoacetamido)pyridin-3-yl)acetate hydrochloride (2.0 g, 7.31 mmol), HOBt (1.18 g, 8.77mmol, 1.2 equiv.) and EDC · HCl (1.82 g, 9.5 mmol, 1.3 equiv.) were dissolved in dry CH_2Cl_2 (40 mL) and *i*-Pr_2NEt (3.8 mL, 21.9 mmol, 3 equiv.) was added under an argon atmosphere. After 21 h at room temperature, the mixture was diluted with CH_2Cl_2 (50 mL) and extracted with 10% aqueous citric acid (15 mL), saturated aqueous $NaHCO_3$ (15 mL), and saturated aqueous NaCl (20 mL). The organic layer was dried with Na_2SO_4 and evaporated under reduced pressure. The crude product was purified by silica gel column chromatography using a linear gradient (2-8%) of EtOH in CH_2Cl_2 to afford the title compound **42b** (1.7 g, 68% yield) as a yellow solid. R_f = 0.81 (10% MeOH in CH_2Cl_2). HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_{16}H_{18}N_5O_4$ 344.1359; found 344.1367. IR (neat, cm^{-1}) 3384, 3184, 31223, 3012, 2986, 1734, 1680, 1550, 1517, 1512, 1506, 1405, 1307, 1203, 1154. ^1H NMR (300 MHz, DMSO-d6, ppm) δ: 10.62 (1H, s), 9.11 (1H, t, *J* = 6.0 Hz), 9.00 (2H, d, *J* = 4.9 Hz), 8.22 (1H, d, *J* = 2.3 Hz), 8.00 (1H, d, *J* = 8.5 Hz), 7.74 – 7.67 (2H, m), 4.19 (2H, d, *J* = 6.0 Hz), 4.09 (2H, q, *J* = 7.1 Hz), 1.19 (3H, t, *J* = 7.1 Hz). $^{13}C\{^1H\}$ NMR (100.6 MHz, DMSO-d6, ppm) δ:

170.9, 167.9, 162.5, 157.8, 157.5, 150.6, 148.4, 139.2, 125.7, 123.2, 112.9, 60.4, 43.2, 36.7, 14.0.

Ethyl 2-(6-((2-(pyrimidin-2-yl)oxazol-5-yl)amino)pyridin-3-yl)acetate. NEt_3 (900 µL, 6.11 mmol, 3 equiv.) and ethyl 2-(6-(2-(pyrimidine-2-carboxamido)acetamido)pyridin-3-yl)acetate **42b** (700 mg, 2.04 mmol) were added to a solution of PPh_3 (1.12 g, 4.28 mmol, 2.1 equiv.) and I_2 (1.09 g, 4.28 mmol, 2.1 equiv.) in dry CH_2Cl_2 (30 mL) under an argon atmosphere. After 6 h at room temperature, the solvent was evaporated under reduced pressure and the residue was purified by silica gel column chromatography using a linear gradient (0-4%) of EtOH in EtOAc to afford the title compound (285 mg, 43% yield) as a yellow solid. R_f = 0.32 (5% MeOH in CH_2Cl_2). HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_{16}H_{16}N_5O_3$ 326.1253; found 326.1253. ^{1}H NMR (400 MHz, $CDCl_3$, ppm) δ: 8.85 (2H, d, *J* = 4.9 Hz), 8.63 (1H, s), 8.37 (1H, dd, *J* = 2.4, 0.8 Hz), 7.60 (1H, dd, *J* = 8.5, 2.4 Hz), 7.28 (1H, t, *J* = 4.9 Hz), 7.21 (1H, s), 6.85 (1H, dd, *J* = 8.5, 0.8 Hz), 4.15 (2H, q, *J* = 7.1 Hz), 3.58 (2H, s), 1.24 (3H, t, *J* = 7.1 Hz). ^{13}C{^{1}H} NMR (100.6 MHz, $CDCl_3$, ppm) δ: 171.2, 157.9, 154.8, 151.9, 151.4, 149.3, 148.8, 139.5, 122.7, 120.5, 111.8, 109.0, 61.3, 37.9, 14.3.

Ethyl 2-(6-((*tert*-butoxycarbonyl)(2-(pyrimidin-2-yl)oxazol-5-yl)amino)pyridin-3-yl)acetate (43b). Ethyl 2-(6-((2-(pyrimidin-2-yl)oxazol-5-yl)amino)pyridin-3-yl)acetate (664 mg, 2.04 mmol), Boc_2O (485 mg, 4.08 mmol, 2 equiv.) and DMAP (13 mg, 0.1 mmol, 0.05 equiv.) were dissolved/suspended in dry THF (12 mL) under an argon atmosphere. After 16 h at room temperature, the solvent was partly evaporated under reduced pressure and the residue was absorbed on silica gel and purified by column chromatography using a linear gradient (50-100%) of EtOAc in

petrol ether to afford the title compound **43b** (287 mg, 76% yield) as a yellow oil. R_f = 0.39 (EtOAc). HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_{21}H_{24}N_5O_5$, 426.1777; found 426.1772. ^{1}H NMR (300 MHz, $CDCl_3$, ppm) δ: 8.88 (2H, d, *J* = 4.9 Hz), 8.23 (1H, t, *J* = 1.6 Hz), 7.69 (2H, d, *J* = 1.6 Hz), 7.34 (1H, t, *J* = 4.9 Hz), 7.26 (1H, s), 4.15 (2H, q, *J* = 7.1 Hz), 3.58 (2H, s), 1.47 (9H, s), 1.24 (3H, t, *J* = 7.1 Hz). ^{13}C{^{1}H} NMR (100.6 MHz, $CDCl_3$, ppm) δ: 170.6, 157.9, 156.2, 154.9, 152.0, 151.8, 148.7, 147.3, 138.8, 127.6, 124.6, 121.2, 118.9, 83.8, 61.3, 38.0, 28.1, 14.3.

2-(6-((tert-butoxycarbonyl)(2-(pyrimidin-2-yl)oxazol-5-yl)amino)pyridin-3-yl)acetic acid (44). A solution of LiOH (59 mg, 2.47 mmol, 3 equiv.) in water (2 mL) was added to a solution of ethyl 2-(6-((*tert*-butoxycarbonyl)(2-(pyrimidin-2-yl)oxazol-5-yl)amino)pyridin-3-yl)acetate **43b** (350 mg, 0.82 mmol) in MeOH (5 mL). After 4 h at room temperature, Amberlite was added till pH~ 4-5, the mixture was filtered, the solids were washed with a mixture of water and MeOH (1:1, 4 × 10 mL) and evaporated under reduced pressure to afford the title compound **44** (200 mg, 61% yield) as a yellow solid. ^{1}H NMR (300 MHz, DMSO-d6, ppm) δ: 12.53 (1H, s), 8.95 (2H, d, *J* = 4.9 Hz), 8.23 (1H, dd, *J* = 2.4, 0.8 Hz), 7.82 (1H, dd, *J* = 8.4, 2.4 Hz), 7.64 (1H, dd, *J* = 8.4, 0.8 Hz), 7.61 (1H, t, *J* = 4.9 Hz), 7.42 (1H, s), 1.44 (9H, s).

M4 PNA monomer (46). TSTU (167 mg, 0.553 mmol, 1.1 equiv.) was added to a solution of 2-(6-((tert-butoxycarbonyl)(2-(pyrimidin-2-yl)oxazol-5-yl)amino)pyridin-3-yl)acetic acid **44** (200 mg, 0.503 mmol) and *i*-Pr_2NEt (113 µL, 0.654 mmol, 1.3 equiv.) in dry THF (5 mL). After 1 h at room temperature, a solution of PNA backbone **45**[51] (171 mg, 0.503 mmol) and *i*-Pr_2NEt (113 µL, 0.654 mmol, 1.3 equiv.) in a mixture of H_2O/MeCN (1:1, 15 mL) were added. The resulting solution was stirred for 5 h at room

temperature, then acidified to pH ~2-3 using a 20% aqueous solution of citric acid. The reaction mixture was extracted with CH_2Cl_2 (3 × 15 mL). The organic layers were combined, extracted with saturated aqueous NaCl (15 mL), dried with Na_2SO_4, and evaporated under reduced pressure. The crude product was purified by reverse phase column chromatography using a linear gradient (20-60%) of MeCN in water to afford the title compound **46** (174 mg, 48% yield) as a white foam. HRMS (ESI/Q-TOF) m/z: $[M + H]^+$ calcd for $C_{38}H_{38}N_7O_8$, 720.2782; found 720.2783. IR (neat, cm^{-1}) 3429, 3265, 3050, 2979, 2935, 2521, 2255, 2126, 1736, 1654, 1545, 1254, 1156. ^1H NMR (300 MHz, DMSO-d6, ppm) (mixture of rotamers) δ: 12.61 (1H, br s), 8.95 (2H, d, *J* = 4.9 Hz), 8.15 (1H, dd, *J* = 9.4, 2.3 Hz), 7.90 – 7.84 (2H, m), 7.75 – 7.58 (5H, m), 7.43 – 7.26 (6H, m), 4.38 – 4.16 (4H, m), 3.96 (1H, s), 3.78 (1.3H, s), 3.63 (0.7H, s), 3.50 – 3.33 (2H, m), 3.24 – 3.11 (2H, m), 1.44 (9H, s). $^{13}C\{^1H\}$ NMR (100.6 MHz, DMSO-d6, ppm) (mixture of rotamers) δ: 170.7, 170.0, 158.1, 156.3, 155.4, 153.8, 151.4, 150.6, 150.6, 148.8, 148.7, 147.1, 143.9, 143.8, 140.7, 139.5, 127.6, 127.0, 125.1, 125.0, 123.1, 122.0, 120.1, 119.0, 83.2, 65.4, 46.7, 27.5.

Methyl 2-(5-iodo-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)acetate (47) 5-Iodouracil (8.0 g, 33.6 mmol) and anhydrous K_2CO_3 (3.95 g, 28.6 mmol) were added to a round bottom flask that was flushed with nitrogen. After the addition of anhydrous DMF (75 mL), the heterogenous mixture was cooled to 0 °C (ice bath). Methyl bromoacetate (3.18 mL, 33.6 mmol) was added dropwise over ca. 15 minutes and the reaction was stirred for an additional 15 minutes at 0 °C (ice bath). The cooling bath was removed and the reaction stirred at room temperature. After 5.5 h, the resulting mixture was concentrated under reduced pressure. To the crude oil was added 25 mL

of 10% aqueous HCl and 20 mL of water. The precipitated solid was filtered on a Buchner funnel and washed with cold water (10 mL) and cold EtOAc (10 mL). The solid was dried under reduced pressure to afford 10.9 g of product containing ca. 5% of a dialkylated impurity. The solid was recrystallized from ca. 140 mL of hot ethanol to obtain 6.97 g of a white solid (mp= 199.5-200.0 °C, lit = 197-199[61]). The mother liquor was concentrated and recrystallized to afford 0.813 g of additional pure product (total 7.78 g, 75% yield); R_f= 0.22 (1:1 hexanes/EtOAc), dialkylated product R_f= 0.39. HRMS (ESI/ion trap) m/z: $[M + H]^+$ calcd for $C_7H_8N_2O_4I$, 310.9529; found 310.9525. IR (neat, cm^{-1}) 3120; 3515; 2953; 1712; 1660; 1606. ^1H NMR (400 MHz, DMSO-d6, ppm) δ: 11.80 (1H, br s) 8.20 (1H, s) 4.52 (2H, s) 3.69 (3H, s). $^{13}C\{^1H\}$ NMR (100.6 MHz, DMSO-d6, ppm) δ: 168.9, 161.5, 151.1, 150.5, 68.8, 52.9, 48.9.

Methyl 2-(2,4-dioxo-5-((trimethylsilyl)ethynyl)-3,4-dihydropyrimidin-1(2*H*)-yl)acetate (48). In a glovebox, methyl 2-(5-iodo-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)acetate **47** (3.065 g, 9.89 mmol), $PdCl_2(PPh_3)_2$ (208 mg, 0.30 mmol), and CuI (75 mg, 0.40 mmol) were added to a flask that was stoppered with a septum and removed from the glovebox. The solids were dissolved in freshly distilled ethyl acetate (26 mL) followed by the dropwise addition of (trimethylsilyl)acetylene (1.72 mL, 12.16 mmol) and freshly distilled triethylamine (7.0 mL, 50.6 mmol). The reaction mixture was stirred 11 h then quenched by the addition of 100 mL of brine. The biphasic solution was filtered through a 2 cm pad of celite directly into a separatory funnel and the layers were separated. The aqueous phase was extracted with Et_2O (3 × 50 mL) with each portion first passing through the celite pad to wash it of remaining product. The combined organics were dried (Na_2SO_4) and concentrated to a crude residue. Purification by flash

chromatography, (5.5 × 14 cm) eluting with 1:1 hexanes/EtOAc afforded pure **48** (2.531 g, 91% yield). Analytical TLC, 1:1 hexanes/EtOAc eluent, R_f= 0.30. An analytical sample was obtained by dual chamber crystallization with EtOAc/hexanes to afford pale off-white plates, Melting point: 134.5-135.0 °C. HRMS (ESI/ion trap) m/z: $[M + H]^+$ calcd for $C_{12}H_{17}N_2O_4Si$, 281.0958; found 281.0961. IR (neat, cm^{-1}) 3165, 3038, 2952, 2845, 2162, 1745, 1698, 1619. ^1H NMR (600 MHz, $CDCl_3$, ppm) δ 8.68 (1H, br s) 7.39 (1H, s) 4.45 (2H, s) 3.79 (3H, s) 0.20 (9H, s). $^{13}C\{^1H\}$ NMR (150.9 MHz, $CDCl_3$) δ 167.2, 161.1, 149.5, 147.3, 100.9, 100.3, 100.3, 94.5, 53.1, 48.9, -0.2.

2-(5-Ethynyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)acetic acid (49). Methyl 2-(2,4-dioxo-5-((trimethylsilyl)ethynyl)-3,4-dihydropyrimidin-1(2*H*)-yl)acetate **48** (2.53 g, 9 mmol) was suspended in aqueous 1 M NaOH (22 mL, 22 mmol) and the mixture was heated and stirred at 65 °C on an aluminum heating block for 1 h during which the reaction became homogeneous. After cooling to 0 °C, conc. HCl (2 mL) was added to precipitate the product. The resulting suspension was filtered, washed with 10 mL of cold water, 5 mL of cold EtOH, then dried to afford the tittle compound **49** (1.604 g, 91% yield) as a tan powder. Melting point: 215 °C (dec). HRMS (ESI/ion trap) m/z: $[M + H]^+$ calcd for $C_8H_7N_2O_4$, 195.0406; found 195.0406. IR (neat, cm^{-1}) 3343, 3295, 3032, 1756, 1682, 1627. ^1H NMR (600 MHz, DMSO-d6, ppm) δ 13.26 (1H, br s) 11.72 (1H, s) 8.11 (1H, s) 4.43 (2H, s) 4.10 (1H, s). $^{13}C\{^1H\}$ NMR (150.9 MHz, DMSO-D6) δ 169.6, 162.6, 150.5, 150.4, 97.2, 84.1, 76.3, 49.3.

Allyl *N*-(2-((((9*H*-fluoren-9-yl)methoxy) carbonyl)amino)ethyl)-*N*-(2-(5-ethynyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)acetyl)glycinate (50). Carboxylic acid **49** (325 mg, 1.67 mmol) and HOBt (295 mg, 2.18 mmol) were taken up in 5 mL of

dry DMF and concentrated to dryness under high vac. The resulting mixture was re-dissolved in 20 mL of DMF and cooled to 0 °C. *N*-(3-Dimethylaminopropyl)-*N*′-ethylcarbodiimide (EDC) (0.36 mL, 2.0 mmol) was added dropwise via syringe and the resulting mixture was stirred at 0 °C for 80 minutes. PNA backbone **7** (700 mg, 1.69 mmol) was added as a solid in one portion, the cooling bath was removed, and the reaction was stirred at room temperature for 24 h. The resulting orange-brown solution was concentrated by rotatory evaporation, re-dissolved in CH_2Cl_2 (40 mL), and washed with 5% aqueous $NaHCO_3$ (2 × 40 mL). The aqueous layer was back-extracted with CH_2Cl_2 (2 × 10 mL) then all the organic layers were combined and washed with brine (20 mL). The solution was dried (Na_2SO_4) and concentrated to a crude brown oil. Purification by flash chromatography, (2.5 x 15 cm) eluting with 1:9 hexanes/EtOAc (200 mL) followed by pure EtOAc afforded pure **50** (629 mg, 67% yield) as a white foam. Analytical TLC, EtOAc eluent, R_f= 0.37. An analytical sample was obtained by recrystallization from hot MeOH to afford fluffy white plates, Melting point: 170.0-171.0 °C. HRMS (ESI/ion trap) m/z: $[M + H]^+$ calcd for $C_{30}H_{29}O_7N_4$, 557.2036; found 557.2022. IR (neat, cm^{-1}) 3340, 3287,1654, 1204. ^1H NMR (600 MHz, DMSO-d6, ppm): (3:1 mixture of rotamers) δ 11.66 (1H, br s) 7.94 (0.75H, s) 7.92-7.84 (2.25H, m) 7.72-7.62 (2H, m) 7.41 (2H, t, *J* = 7.4 Hz) 7.40-7.37 (0.75H, br m) 7.33 (2H, t, *J* = 7.2 Hz) 7.24 (0.25H, br t) 5.96 (0.25H, ddt, *J* = 16.5, 11.0, 5.5 Hz) 5.90 (0.75H ddt, *J* = 16.0, 10.6, 5.3) 5.36 (0.25H, d, *J* = 17.3 Hz) 5.31 (0.75H, d, *J* = 17.3 Hz) 5.25 (0.25H, d, *J* = 10.7 Hz) 5.21 (0.75H, d, *J* = 10.7 Hz) 4.74 (1.5H, s) 4.66 (0.5H, d, *J* = 5.5 Hz) 4.58 (1.5H, d, *J* = 5.3 Hz) 4.57 (0.5H, s) 4.41-4.27 (2.5H, m) 4.23 (1H, t, *J* = 6.8 Hz) 4.16-4.03 (2.5H, m) 3.47-3.01 (4H, m). $^{13}C\{^1H\}$ NMR (150.9 MHz, DMSO-d6): (mixture of rotamers) δ

168.9, 168.6, 167.3, 167.0, 162.1, 156.3, 156.1, 150.32, 150.27, 149.9, 143.8, 140.7, 132.2, 132.1, 127.6, 127.0, 125.1, 120.1, 118.3, 117.9, 96.62, 96.60, 83.5, 76.0, 65.5, 65.4, 65.3, 64.9, 48.9, 48.2, 48.0, 47.9, 46.9, 46.7, 40.1, 38.7, 37.9.

1-Azido-3-(*tert*-butoxy)benzene (51). Following a literature procedure,[62] 3-aminophenol (2.0 g, 18.3 mmol) was dissolved in 40 mL of 2 M HCl. The solution was cooled to 0 °C and a solution of $NaNO_2$ (1.57 g, 23 mmol) in 8 mL of H_2O was added dropwise by pipet. The solution turned dark orange along with gas evolution. After stirring for 30 min, NaN_3 (1.56 g, 23 mmol) solution in 10 mL of H_2O was carefully added dropwise by pipet. *Warning: Vigorous gas evolution and frothing*. The solution was gradually warmed to room temperature and stirred for 3 h. The resulting brown solution was extracted with EtOAc (3 × 50 mL), washed with brine (20 mL), dried (Na_2SO_4), and concentrated to give 3-azido phenol (2.45 g, 98% yield) as a pure red orange oil that solidified upon storage at -20 °C. The product had spectral properties (^1H, ^{13}C, and IR) matching the literature[62] and was used without further purification. Based on a literature precedent,[53] 3-azidophenol (1.0 g, 7.4 mmol) was dissolved in 8 mL of CH_2Cl_2. To the solution was added $Mg(ClO_4)_2$ followed by a solution of di-*tert*-butyldicarbonate (4.0 g, 18.5 mmol) in 2 mL of CH_2Cl_2. Bubbling was immediately observed and the reaction was heated at 40 °C for 6 h. Water (50 mL) was added to the reaction and the mixture was extracted with CH_2Cl_2 (4 × 20 mL), dried (Na_2SO_4), and carefully concentrated by rotatory evaporation. The crude product containing unreacted phenol was purified by flash chromatography (2.5 × 15 cm) eluting with a gradient of 1 to 10% Et_2O in pentane. Fractions containing product were concentrated to afford **51** (856 mg, 61% yield) as a pale yellow oil. Analytical TLC, 15:1 pentane/Et_2O, R_f= 0.61. HRMS (EIMS) m/z: $[M]^+$

calcd for $C_{10}H_{13}N_3O$, 191.1059; found 191.1057. IR (neat, cm^{-1}) 2978, 2108, 1587, 1478, 1160. ^1H NMR (400 MHz, $CDCl_3$, ppm) δ: 7.23 (1H, t, *J* = 8.0 Hz) 6.77 (2H, m) 6.65 (1H, t, *J* = 2.2 Hz) 1.36 (9H, s). $^{13}C\{^1H\}$ NMR (100.6 MHz, $CDCl_3$) δ: 156.9, 140.6, 129.8, 120.5, 114.7, 113.9, 79.1, 28.9.

Allyl ester of T1 PNA monomer (52). Alkyne **50** (265 mg, 0.476 mmol) was taken up in 12 mL of degassed 1:1 THF/EtOH and heated to 40 °C on an aluminum heating block until the solution was homogeneous. Sodium ascorbate (189 mg, 0.952 mmol), azide **51** (137 mg, 0.715 mmol), and degassed water (3 mL) were added sequentially and the stirred suspension was evacuated and purged with nitrogen. Addition of *N*,*N*'-dimethylethylenediamine (26 µL, 0.238 mmol) via syringe was followed by dropwise addition of aqueous $CuSO_4$ solution (0.47 M, 0.4 mL, 0.190 mmol) during which the solution darkened and then color dissipated. The resulting yellow/orange solution was stirred at 40 °C for 90 minutes after which TLC showed complete conversion. The reaction mixture was diluted with water (40 mL), extracted with CH_2Cl_2 (3 × 25 mL), dried (Na_2SO_4), and concentrated to a crude oil. The product was purified by flash chromatography (2.5 × 15 cm), 1:1:0.1 toluene/CH_2Cl_2/MeOH eluent. After the elution of ca 150 mL of eluent, a yellow band eluted over 60 mL which was collected and concentrated to afford pure **52** (300 mg, 85% yield) as a pale yellow solid. Analytical TLC, 1:1:0.1 toluene/CH_2Cl_2/MeOH eluent, R_f= 0.23. An analytical sample was obtained by recrystallization in EtOAc/Et_2O to afford pale yellow amorphous crystals, melting point: 181.0-182.8 °C. HRMS (ESI/ion trap) m/z: $[M + H]^+$ calcd for $C_{40}H_{42}O_8N_7$, 748.3095; found 748.3077. IR (neat, cm^{-1}) 3361, 3178, 3067, 2978, 1672, 1602, 1473, 1367, 1248, 1208, 1193, 1149. ^1H NMR (400 MHz, $CDCl_3$, ppm, 7:3

mixture of rotamers) δ: 8.77 (0.3H, br s) 8.71 (0.7H, br s) 8.60 (0.7H, s) 8.58 (0.3H, s) 8.29 (0.3H, s) 8.25 (0.7H, s) 7.79-7.71 (2H, m) 7.62 (1.4H, d, *J* = 7.6 Hz) 7.58 (0.6H, d, *J* = 7.6 Hz) 7.47-7.34 (5H, m) 7.30 (2H, dt, *J* = 7.6, 0.9 Hz) 7.08-7.03 (1H, m) 5.99-5.80 (1.7H, m) 5.42-5.22 (2.3H, m) 4.70 (0.6H, d, *J* = 6.2 Hz) 4.68-4.60 (2.8H, m) 4.57 (0.6H, s) 4.51 (1.4H, d, *J* = 6.2 Hz) 4.38 (0.6H, d, *J* = 7.0 Hz) 4.25 (0.7H, t, *J* = 6.2 Hz) 4.23-4.16 (0.9H, m) 4.08 (1.4H, s) 3.72-3.28 (4H, m) 1.41 (9H, s). $^{13}C\{^{1}H\}$ NMR (100.6 MHz, $CDCl_3$, mixture of rotamers) δ: 169.1, 168.7, 166.8, 161.2, 156.8, 156.7, 149.9, 143.9, 143.8, 141.9, 141.7, 141.3, 139.5, 137.5, 131.3, 130.9, 129.9, 127.71, 127.65, 127.1, 125.1, 123.9, 120.08, 120.05, 119.95, 119.2, 116.0, 115.0, 105.7, 79.6, 66.9, 66.8, 66.3, 49.2, 48.9, 48.7, 47.3, 39.4, 39.0, 28.9.

T1 PNA monomer (53). Allyl ester of **T1** PNA monomer **52** (192 mg, 0.26 mmol) was dissolved in anhydrous THF (5.1 mL). $Pd(PPh_3)_4$ (12 mg, 0.01 mmol) was added followed by *N*-ethylaniline (59 μL, 0.47 mmol). The yellow solution was stirred under nitrogen for 2 h at room temperature. After the reaction was complete by TLC, the solvent was removed under reduced pressure and the yellow residue was re-dissolved in CH_2Cl_2 (20 mL). The resulting solution was washed with 10% aqueous $KHSO_4$ solution (3 × 10 mL) and the combined aqueous layers were back-extracted with CH_2Cl_2 (15 mL). The combined organic extracts were washed with brine (10 mL), dried with Na_2SO_4, and concentrated under reduced pressure to afford a crude yellow solid containing the product contaminated with allyl anilines. The solid was re-dissolved in CH_2Cl_2, adsorbed onto ~600 mg of celite, and purified using automated chromatography using a linear gradient (0-10%) of MeOH in CH_2Cl_2 on a 20 g Silica-CM column (Agela Technologies) to afford **T1** monomer **53** (140 mg, 77% yield) as an off-white solid;

analytical TLC, 3:1 CH_2Cl_2/MeOH, R_f= 0.17. HRMS (ESI/ion trap) m/z: $[M + H]^+$ calcd for $C_{37}H_{38}N_7O_8$, 708.2782; found 708.2771. IR (neat, cm^{-1}) 3313, 3164, 3055, 1720, 1697, 1686, 1665. ^1H NMR (400 MHz, DMSO-d6, ppm, 7:3 mixture of rotamers) δ: 12.8 (1H, br s) 11.77 (0.7H, s) 11.76 (0.3H, s) 8.89 (0.7H, s) 8.87 (0.3H, s) 8.39 (0.7H, br s) 8.29 (0.3H, br s) 7.87 (2H, d, *J* = 7.5 Hz) 7.72-7.61 (3H, m) 7.54-7.26 (7H, m) 7.11 (1H, d, *J* = 7.9 Hz) 4.93 (1.4H, s) 4.74 (0.9H, m) 4.39-4.15 (3.3H, m) 4.01 (1.4H, s) 3.52-3.00 (4H, m) 1.36 (9H, s). $^{13}C\{^1H\}$ NMR (100.6 MHz, DMSO-d6, mixture of rotamers) δ: 168.0, 167.6, 162.1, 162.0, 156.8, 156.7, 150.7, 144.4, 143.3, 143.1, 141.20, 141.17, 140.5, 137.7, 130.8, 128.1, 127.5, 125.6, 124.0, 120.6, 120.5, 115.7, 115.3, 104.24, 104.21, 79.6, 66.0, 65.9, 48.7, 47.4, 47.2, 29.0.

Allyl ester of Tr PNA monomer (55). Alkyne **50** (297 mg, 0.53 mmol) was dissolved in 14 mL of degassed 1:1 THF/EtOH. Sodium ascorbate (212 mg, 1.07 mmol), azidomethyl pivalate **54**[54] (126 mg, 0.80 mmol), and degassed water (3.5 mL) were added sequentially and the stirred suspension was evacuated and purged with nitrogen. Addition of *N*,*N*'-dimethylethylenediamine (28 μL, 0.27 mmol) via syringe was followed by dropwise addition of aqueous $CuSO_4$ solution (0.47 M, 0.45 mL, 0.21 mmol) during which the solution darkened and then color dissipated. The resulting yellow solution was stirred at 40 °C on an aluminum heating block for 2 h after which TLC showed complete conversion. The reaction mixture was diluted with water (30 mL) and extracted with CH_2Cl_2 (2 × 30 mL). The combined organics were washed with bring (20 mL), dried (Na_2SO_4), and concentrated to a crude oil. The product was purified by flash chromatography (3 × 16 cm) using a linear gradient (1-4%) of MeOH in CH_2Cl_2. The fractions containing product were combined to afford **55** (291 mg, 77% yield) as a pale

yellow solid. Analytical TLC, 20:1 CH_2Cl_2/MeOH eluent, R_f= 0.19. HRMS (ESI/ion trap) m/z: $[M + H]^+$ calcd for $C_{36}H_{40}O_9N_7$, 714.2888; found 714.2869. IR (neat, cm^{-1}) 1709,1674,1523, 1449, 1187, 1121. ^1H NMR (400 MHz, $CDCl_3$, ppm, 6:4 mixture of rotamers) δ: 8.84 (0.4H s) 8.79 (0.6H, s) 8.40 (1H, s) 8.23 (0.6H s) 8.19 (0.4H s) 7.74 (2H, d, *J* = 7.8 Hz) 7.64-7.54 (2H, m) 7.38 (2H, t, *J* = 7.8 Hz) 7.29 (2H, t, *J* = 7.3 Hz) 6.23 (1.2H, s) 6.22 (0.8H, s) 4.69 (0.8H, d, *J* = 4.6 Hz) 4.67-4.58 (2.4H, m) 4.55 (0.8H, s) 4.50 (2.4H, d, *J* = 6.7 Hz) 4.38 (0.8H, d, *J* = 6.7 Hz) 4.24 (0.6H, t, *J* = 6.7 Hz) 4.24-4.16 (1.4H, m) 4.07 (1.2H, s) 3.63-3.34 (4H, m) 1.19 (9H, s). $^{13}C\{^1H\}$ NMR (100.6 MHz, $CDCl_3$, mixture of rotamers) δ: 177.3, 169.1, 168.7, 166.9, 161.0, 156.8, 149.9, 143.9, 143.8, 141.9, 141.7, 141.3, 139.5, 131.2, 130.9, 127.7, 127.1, 125.1, 123.2, 120.1, 119.9, 119.2, 105.6, 69.8, 66.9, 66.8, 66.3, 49.2, 48.9, 48.6, 47.3, 39.4, 39.0, 38.8, 26.8.

Tr PNA monomer (56). Allyl ester on **Tr** PNA monomer **55** (286 mg, 0.40 mmol) was dissolved in anhydrous THF (8 ml). $Pd(PPh_3)_4$ (19 mg, 0.02 mmol) was added followed by *N*-ethylaniline (91 μL, 0.72 mmol). The yellow solution was stirred under nitrogen for 2 h at room temperature. After the reaction was complete by TLC, the solvent was removed under reduced pressure and the yellow residue was re-dissolved in CH_2Cl_2 (20 mL). The resulting solution was adsorbed onto ~600 mg of celite, and purified using automated chromatography on a linear gradient of 0-10% MeOH in CH_2Cl_2 on a 20 g Silica-CM column (Agela Technologies) to afford **Tr** monomer **56** (241 mg, 91%) as an off-white solid; analytical TLC, 3:1 CH_2Cl_2/MeOH, R_f= 0.17. HRMS (ESI/ion trap) m/z: $[M + H]^+$ calcd for $C_{33}H_{36}N_7O_8$, 708.2782; found 708.2771. IR (neat, cm^{-1}) 3173, 3069, 1701, 1686, 1671, 1120. ^1H NMR (600 MHz, DMSO-d6, ppm, 6:4 mixture of rotamers) δ: 13.1 (0.4H, br s) 12.7 (0.6H, br s) 11.74 (0.6H, s) 11.72 (0.4H, s) 8.49 (0.6H, s) 8.48

(0.4H, s) 8.37 (0.6H, s) 8.27 (0.4H, s) 7.87 (2H, d, *J* = 7.5 Hz) 7.70-7.64 (2H, m) 7.44 (0.6H, t, *J* = 5.5 Hz) 7.39 (2H, t, *J* = 7.5 Hz) 7.34-7.30 (2H, m) 7.27 (0.4H, t, *J* = 5.5 Hz) 6.36 (1.2H, s) 6.35 (0.8H, s) 4.90 (1.2H, s) 4.71 (0.8H, s) 4.35 (1.2H, d, *J* = 6.6 Hz) 4.29 (0.8H, d, *J* = 6.9 Hz) 4.24 (0.6H, t, *J* = 6.6 Hz) 4.23-4.17 (0.8H, m) 4.00 (1.2H, t, *J* = 6.4 Hz) 3.38-3.15 (2H, m) 3.15-3.09 (0.8H, m). $^{13}C\{^{1}H\}$ NMR (150.9 MHz, DMSO-d6, mixture of rotamers) δ: 176.5, 170.7, 170.3, 167.5, 167.1, 161.61, 161.57, 156.3, 156.1, 150.1, 143.8, 142.6, 142.4, 140.70, 140.68, 139.2, 139.1, 127.6, 127.0, 125.1, 123.1, 120.1, 103.61, 103.58, 70.0, 65.5, 65.4, 49.1, 48.6, 48.3, 48.1, 47.7, 47.0, 46.9, 46.7, 38.6, 38.2, 37.9, 26.4.

PNA sequences were synthesized using standard 2-μmol scale Fmoc protocol on an Expedite 8909 DNA synthesizer using NovaSyn TG Sieber resin support and Fmoc chemistry as previously reported.[39]

Isothermal titration calorimetry experiments were performed on a MicroCal iTC200 instrument at 25 °C in phosphate buffer containing 2 mM $MgCl_2$, 90 mM KCl, 10 mM NaCl, 50 mM potassium phosphate. 2.45 μL aliquots of 80 μM PNA solution were sequentially injected from a 40 μL rotating syringe (750 rmp) into 200 μL of 8 μM RNA hairpin solution. For ITC titration traces, see Supporting Information, Figures SA-SK.

Acknowledgements

This work was supported by the Latvian Institute of Organic Synthesis, internal research fund IG-2018-04 and the US National Science Foundation grants CHE-1708761 to E.R. and CHE-1708699 to J.A.M. The Regional NMR Facility (600 MHz instrument) at Binghamton University is supported by NSF (CHE-0922815).

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI:

Spectral data for new compounds; LC-MS characterization of PNA monomers and oligiomers; PNA synthesis; preparation of RNA hairpins; representative ITC images; computational data.

REFERENCES

1. Devi, G.; Zhou, Y.; Zhong, Z.; Toh, D.-F. K.; Chen, G. RNA triplexes: from structural principles to biological and biotech applications. *Wiley Interdisciplinary Reviews: RNA* **2015,** *6*, 111-128.
2. Zhou, Y.; Kierzek, E.; Loo, Z. P.; Antonio, M.; Yau, Y. H.; Chuah, Y. W.; Geifman-Shochat, S.; Kierzek, R.; Chen, G. Recognition of RNA duplexes by chemically modified triplex-forming oligonucleotides. *Nucleic Acids Res.* **2013,** *41*, 6664-6673.
3. Li, M.; Zengeya, T.; Rozners, E. Short Peptide Nucleic Acids Bind Strongly to Homopurine Tract of Double Helical RNA at pH 5.5. *J. Am. Chem. Soc.* **2010,** *132*, 8676-8681.
4. Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P. E. PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen-bonding rules. *Nature* **1993,** *365*, 566-568.
5. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science* **1991,** *254*, 1497-1500.
6. Sato, T.; Sakamoto, N.; Nishizawa, S. Kinetic and thermodynamic analysis of triplex formation between peptide nucleic acid and double-stranded RNA. *Org. Biomol. Chem.* **2018,** *16*, 1178-1187.
7. Devi, G.; Yuan, Z.; Lu, Y.; Zhao, Y.; Chen, G. Incorporation of thio-pseudoisocytosine into triplex-forming peptide nucleic acids for enhanced recognition of RNA duplexes. *Nucleic Acids Res.* **2014,** *42*, 4008-4018.
8. Zengeya, T.; Gupta, P.; Rozners, E. Triple Helical Recognition of RNA Using 2-Aminopyridine-Modified PNA at Physiologically Relevant Conditions. *Angew. Chem., Int. Ed.* **2012,** *51*, 12593-12596.
9. Muse, O.; Zengeya, T.; Mwaura, J.; Hnedzko, D.; McGee, D. W.; Grewer, C. T.; Rozners, E. Sequence Selective Recognition of Double-Stranded RNA at

Physiologically Relevant Conditions Using PNA-Peptide Conjugates. *ACS Chem. Biol.* **2013,** *8*, 1683-1686.

10. Hnedzko, D.; McGee, D. W.; Karamitas, Y. A.; Rozners, E. Sequence-selective recognition of double-stranded RNA and enhanced cellular uptake of cationic nucleobase and backbone-modified peptide nucleic acids. *RNA* **2017,** *23*, 58-69.
11. Kim, K. T.; Chang, D.; Winssinger, N. Double-Stranded RNA-Specific Templated Reaction with Triplex Forming PNA. *Helv. Chim. Acta* **2018,** *101*, e1700295.
12. Endoh, T.; Hnedzko, D.; Rozners, E.; Sugimoto, N. Nucleobase-Modified PNA Suppresses Translation by Forming a Triple Helix with a Hairpin Structure in mRNA In Vitro and in Cells. *Angew. Chem., Int. Ed.* **2016,** *55*, 899-903.
13. Gupta, P.; Zengeya, T.; Rozners, E. Triple helical recognition of pyrimidine inversions in polypurine tracts of RNA by nucleobase-modified PNA. *Chem. Commun.* **2011,** *47*, 11125-11127.
14. Purwanto, M. G. M.; Weisz, K. Non-natural nucleosides for the specific recognition of Watson-Crick base pairs. *Curr. Org. Chem.* **2003,** *7*, 427-446.
15. Hari, Y.; Obika, S.; Imanishi, T. Towards the sequence-selective recognition of double-stranded DNA containing pyrimidine-purine interruptions by triplex-forming oligonucleotides. *Eur. J. Org. Chem.* **2012,** *2012*, 2875-2887.
16. Gowers, D. M.; Fox, K. R. Towards mixed sequence recognition by triple helix formation. *Nucleic Acids Res.* **1999,** *27*, 1569-1577.
17. Fox, K. R.; Brown, T.; Rusling, D. A. DNA Recognition by Parallel Triplex Formation. In *Chemical Biology No. 7: DNA-targeting Molecules as Therapeutic Agents*, Waring, M. J., Ed. Royal Society of Chemistry, 2018; Vol. 7, pp 1-32.
18. Griffin, L. C.; Kiessling, L. L.; Beal, P. A.; Gillespie, P.; Dervan, P. B. Recognition of all four base pairs of double-helical DNA by triple-helix formation: design of nonnatural deoxyribonucleosides for pyrimidine purine base pair binding. *J. Am. Chem. Soc.* **1992,** *114*, 7976-82.
19. Huang, C. Y.; Miller, P. S. Triplex formation by an oligodeoxyribonucleotide containing N4-(6-aminopyridinyl)-2'-deoxycytidine. *J. Am. Chem. Soc.* **1993,** *115*, 10456-7.
20. Zimmerman, S. C.; Schmitt, P. Model Studies Directed toward a General Triplex DNA Recognition Scheme: A Novel DNA Base That Binds a CG Base-Pair in an Organic Solvent. *J. Am. Chem. Soc.* **1995,** *117*, 10769-70.
21. Huang, C.-Y.; Bi, G.; Miller, P. S. Triplex formation by oligonucleotides containing novel deoxycytidine derivatives. *Nucleic Acids Res.* **1996,** *24*, 2606-2613.
22. Lehmann, T. E.; Greenberg, W. A.; Liberles, D. A.; Wada, C. K.; Dervan, P. B. Triple-helix formation by pyrimidine oligonucleotides containing non-natural nucleosides with extended aromatic nucleobases. Intercalation from the major groove as a method for recognizing C.G and T.A base pairs. *Helv. Chim. Acta* **1997,** *80*, 2002-2022.

23. Wang, E.; Koshlap, K. M.; Gillespie, P.; Dervan, P. B.; Feigon, J. Solution structure of a pyrimidine·purine·pyrimidine triplex containing the sequence-specific intercalating non-natural base D3. *J. Mol. Biol.* **1996,** *257*, 1052-69.

24. Koshlap, K. M.; Gillespie, P.; Dervan, P. B.; Feigon, J. Nonnatural deoxyribonucleoside D3 incorporated in an intramolecular DNA triplex binds sequence-specifically by intercalation. *J. Am. Chem. Soc.* **1993,** *115*, 7908-9.

25. Mertz, E.; Mattei, S.; Zimmerman, S. C. Synthetic Receptors for CG Base Pairs. *Org. Lett.* **2000,** *2*, 2931-2934.

26. Lengeler, D.; Weisz, K. Novel phthalimide nucleosides for the specific recognition of a CG Watson-Crick base pair. *Tetrahedron Lett.* **2001,** *42*, 1479-1481.

27. Wang, W.; Purwanto, M. G. M.; Weisz, K. CG base pair recognition by substituted phenylimidazole nucleosides. *Org. Biomol. Chem.* **2004,** *2*, 1194-1198.

28. Mertz, E.; Mattei, S.; Zimmerman, S. C. Synthesis and duplex DNA recognition studies of oligonucleotides containing a ureido isoindolin-1-one homo-N-nucleoside. A comparison of host-guest and DNA recognition studies. *Bioorg. Med. Chem.* **2004,** *12*, 1517-1526.

29. Semenyuk, A.; Darian, E.; Liu, J.; Majumdar, A.; Cuenoud, B.; Miller, P. S.; MacKerell, A. D., Jr.; Seidman, M. M. Targeting of an Interrupted Polypurine:Polypyrimidine Sequence in Mammalian Cells by a Triplex-Forming Oligonucleotide Containing a Novel Base Analogue. *Biochemistry* **2010,** *49*, 7867-7878.

30. Hari, Y.; Akabane, M.; Hatanaka, Y.; Nakahara, M.; Obika, S. A 4-[(3R,4R)-dihydroxypyrrolidino]pyrimidin-2-one nucleobase for a CG base pair in triplex DNA. *Chem. Commun.* **2011,** *47*, 4424-4426.

31. Hari, Y.; Akabane, M.; Obika, S. 2',4'-BNA bearing a chiral guanidinopyrrolidine-containing nucleobase with potent ability to recognize the CG base pair in a parallel-motif DNA triplex. *Chem. Commun.* **2013,** *49*, 7421-7423.

32. Guianvarc'h, D.; Benhida, R.; Fourrey, J.-L.; Maurisse, R.; Sun, J.-S. Incorporation of a novel nucleobase allows stable oligonucleotide-directed triple helix formation at the target sequence containing a purine·pyrimidine interruption. *Chem. Commun.* **2001**, 1814-1815.

33. Rusling, D. A.; Powers, V. E. C.; Ranasinghe, R. T.; Wang, Y.; Osborne, S. D.; Brown, T.; Fox, K. R. Four base recognition by triplex-forming oligonucleotides at physiological pH. *Nucleic Acids Res.* **2005,** *33*, 3025-3032.

34. Wang, Y.; Rusling, D. A.; Powers, V. E. C.; Lack, O.; Osborne, S. D.; Fox, K. R.; Brown, T. Stable Recognition of TA Interruptions by Triplex Forming Oligonucleotides Containing a Novel Nucleoside. *Biochemistry* **2005,** *44*, 5884-5892.

35. Ohkubo, A.; Yamada, K.; Ito, Y.; Yoshimura, K.; Miyauchi, K.; Kanamori, T.; Masaki, Y.; Seio, K.; Yuasa, H.; Sekine, M. Synthesis and triplex-forming properties of

oligonucleotides capable of recognizing corresponding DNA duplexes containing four base pairs. *Nucleic Acids Res.* **2015,** *43*, 5675-5686.

36. Toh, D.-F. K.; Devi, G.; Patil, K. M.; Qu, Q.; Maraswami, M.; Xiao, Y.; Loh, T. P.; Zhao, Y.; Chen, G. Incorporating a guanidine-modified cytosine base into triplex-forming PNAs for the recognition of a C-G pyrimidine-purine inversion site of an RNA duplex. *Nucleic Acids Res.* **2016,** *44*, 9071-9082.
37. Patil, K. M.; Toh, D.-F. K.; Yuan, Z.; Meng, Z.; Shu, Z.; Ong, A. A. L.; Krishna, M. S.; Lu, Y.; Chen, G.; Zhang, H.; Lu, L. Incorporating uracil and 5-halouracils into short peptide nucleic acids for enhanced recognition of A-U pairs in dsRNAs. *Nucleic Acids Res.* **2018,** *46*, 7506-7521.
38. Gadre, S.; Peters, M. S.; Serrano, A.; Schrader, T. A Synthetic Methodology Toward Pyrrolo[2,3-b]pyridones for GC Base Pair Recognition. *Org. Lett.* **2018,** *20*, 6961-6964.
39. Hnedzko, D.; Cheruiyot, S. K.; Rozners, E. Using triple-helix-forming Peptide nucleic acids for sequence-selective recognition of double-stranded RNA. *Curr. Protoc. Nucleic Acid Chem.* **2014,** *58*, 4.60.1-4.60.23.
40. Yip, S. F.; Cheung, H. Y.; Zhou, Z.; Kwong, F. Y. Room-Temperature Copper-Catalyzed α-Arylation of Malonates. *Org. Lett.* **2007,** *9*, 3469-3472.
41. Morgentin, R.; Jung, F.; Lamorlette, M.; Maudet, M.; Menard, M.; Ple, P.; Pasquet, G.; Renaud, F. An efficient large-scale synthesis of alkyl 5-hydroxy-pyridin- and pyrimidin-2-yl acetate. *Tetrahedron* **2009,** *65*, 757-764.
42. Wong, B.; Linghu, X.; Crawford, J. J.; Drobnick, J.; Lee, W.; Zhang, H. A chemoselective Reformatsky-Negishi approach to α-haloaryl esters. *Tetrahedron* **2014,** *70*, 1508-1515.
43. Kaul, R.; Brouillette, Y.; Sajjadi, Z.; Hansford Karl, A.; Lubell William, D. Selective tert-butyl ester deprotection in the presence of acid labile protecting groups with use of $ZnBr_2$. *J. Org. Chem.* **2004,** *69*, 6131-3.
44. Puah, R. Y.; Jia, H.; Maraswami, M.; Kaixin Toh, D.-F.; Ero, R.; Yang, L.; Patil, K. M.; Lerk Ong, A. A.; Krishna, M. S.; Sun, R.; Tong, C.; Huang, M.; Chen, X.; Loh, T. P.; Gao, Y.-G.; Liu, D. X.; Chen, G. Selective Binding to mRNA Duplex Regions by Chemically Modified Peptide Nucleic Acids Stimulates Ribosomal Frameshifting. *Biochemistry* **2018,** *57*, 149-159.
45. Feagin, T. A.; Shah, N. I.; Heemstra, J. M. Convenient and scalable synthesis of Fmoc-protected peptide nucleic acid backbone. *J. Nucleic Acids* **2012**, 354549, 5 pp.
46. Babkov, D. A.; Valuev-Elliston, V. T.; Paramonova, M. P.; Ozerov, A. A.; Ivanov, A. V.; Chizhov, A. O.; Khandazhinskaya, A. L.; Kochetkov, S. N.; Balzarini, J.; Daelemans, D.; Pannecouque, C.; Seley-Radtke, K. L.; Novikov, M. S. Scaffold hopping: Exploration of acetanilide-containing uracil analogues as potential NNRTIs. *Bioorg. Med. Chem.* **2015,** *23*, 1069-1081.

47. Robinson, A.; Thomas, G. L.; Spandl, R. J.; Welch, M.; Spring, D. R. Gemmacin B: bringing diversity back into focus. *Org. Biomol. Chem.* **2008,** *6*, 2978-2981.

48. Lipshutz, B. H.; Hungate, R. W.; McCarthy, K. E. Heterocycles as masked diamide/dipeptide equivalents. Formation and reactions of substituted 5-(acylamino)oxazoles as intermediates en route to the cyclopeptide alkaloids. *J. Am. Chem. Soc.* **1983,** *105*, 7703-13.

49. Falorni, M.; Giacomelli, G.; Porcheddu, A.; Dettori, G. New oxazole-based conformationally restricted peptidomimetics: design and synthesis of pseudopeptides. *Eur. J. Org. Chem.* **2000**, 3217-3222.

50. Agnelli, F.; Sulikowski, G. A. Synthesis of arylacetates by the palladium-catalyzed cross-coupling of aryl bromides and copper(II) enolates. *Tetrahedron Lett.* **1998,** *39*, 8807-8810.

51. Debaene, F.; Da Silva, J. A.; Pianowski, Z.; Duran, F. J.; Winssinger, N. Expanding the scope of PNA-encoded libraries: divergent synthesis of libraries targeting cysteine, serine and metallo-proteases as well as tyrosine phosphatases. *Tetrahedron* **2007,** *63*, 6577-6586.

52. Novosjolova, I.; Kennedy, S. D.; Rozners, E. 2-Methoxypyridine as a Thymidine Mimic in Watson-Crick Base Pairs of DNA and PNA: Synthesis, Thermal Stability, and NMR Structural Studies. *ChemBioChem* **2017,** *18*, 2165-2170.

53. Bartoli, G.; Bosco, M.; Carlone, A.; Dalpozzo, R.; Locatelli, M.; Melchiorre, P.; Sambri, L. Alcohols and Di-tert-butyl Dicarbonate: How the Nature of the Lewis Acid Catalyst May Address the Reaction to the Synthesis of tert-Butyl Ethers. *J. Org. Chem.* **2006,** *71*, 9580-9588.

54. Loren, J. C.; Krasinski, A.; Fokin, V. V.; Sharpless, K. B. NH-1,2,3-triazoles from azidomethyl pivalate and carbamates: Base-labile N-protecting groups. *Synlett* **2005**, 2847-2850.

55. Hnedzko, D.; Rozners, E. Sequence-specific recognition of structured RNA by triplex-forming peptide nucleic acids. In *Methods in Enzymology: RNA Recognition*, Hargrove Amanda, E., Ed. Elsevier, 2019; Vol. 623, pp 401-416.

56. Kotikam, V.; Kennedy, S. D.; MacKay, J. A.; Rozners, E. Synthetic, Structural, and RNA Binding Studies on 2-Aminopyridine-Modified Triplex-Forming Peptide Nucleic Acids. *Chem. Eur. J.* **2019,** *25*, 4367-4372.

57. Rani, V. J.; Aminedi, R.; Polireddy, K.; Jagadeeswarareddy, K. Synthesis and Spectral Characterization of New Bis(2-(pyrimidin-2-yl)ethoxy)alkanes and Their Pharmacological Activity. *Arch. Pharm.* **2012,** *345*, 663-669.

58. Blomgren, P. A. C., K.S.; Kropf, J.E.; Lee, S.H.; Mitchell, S.A.; Schmitt, A.C.; Xu, J; Zhao, Z. Imidazopyrydines sky inhibitors. US pat. 0148430A1, 2014.

59. Pasternak, A. D., R. K.; Zhu, Y.; Yang, L.; Walsh, S.; Tang, H.; Kim, D.; Pio, B.; Shahripour, A.; Belyk, K. Inhibitors of the Renal Outer Medullary Potassium Channel. US pat. US20130217680, 2013.

60. Holmes, B.; Lee, J.; Landon, K. A.; Benavides-Serrato, A.; Bashir, T.; Jung, M. E.; Lichtenstein, A.; Gera, J. Mechanistic Target of Rapamycin (mTOR) Inhibition Synergizes with Reduced Internal Ribosome Entry Site (IRES)-mediated Translation of Cyclin D1 and c-MYC mRNAs to Treat Glioblastoma. *J. Biol. Chem.* **2016,** *291*, 14146-14159.

61. Qu, G.; Zhang, Z.; Guo, H.; Geng, M.; Xia, R. Microwave-promoted facile and efficient preparation of N-(alkoxycarbonylmethyl)nucleobases. Building blocks for peptide nucleic acids. *Molecules* **2007,** *12*, 543-551.

62. Demeter, O.; Fodor, E. A.; Kallay, M.; Mezo, G.; Nemeth, K.; Szabo, P. T.; Kele, P. A Double-Clicking Bis-Azide Fluorogenic Dye for Bioorthogonal Self-Labeling Peptide Tags. *Chem. Eur. J.* **2016,** *22*, 6382-6388.