Synthetic life
(continuation of „The molecular origins of life” SoSe 2021)

WiSe 2021/22
Zbigniew Pianowski
Overview of the course

**artificial ribozymes and aptamers** for efficient catalysis and recognition (SELEX, DNAzymes, foldamers);

**unnatural base pairing** – expansion of the genetic alphabet;

**Artificial genetic polymers** and oligonucleotide analogues (XNA);

biosynthetic incorporation of **unnatural aminoacids (UAAs)** into proteins;

**enzyme engineering** – production of enzymes with unknown or unnatural properties, *ab initio* protein design, directed evolution, theozymes;

**Artificial lipid vesicles** as models for protocell multiplication;

**design of artificial organisms**
CHAPTER 1

OLIGONUCLEOTIDES

Part 3 – noncanonical backbone
Artificial genetic polymers

DNA

RNA (natural)

RNA (unnatural)

2'-modified RNA

Phosphorothioate

Boranophosphate

Hexitol Nucleic Acid (HNA)

Threose Nucleic Acid (TNA)

Peptide Nucleic Acid (PNA)
**Native chemical ligation**

*Native chemical ligation* or *NCL* is an important extension of the chemical ligation field, a concept for constructing a large polypeptide formed by the assembling of two or more unprotected peptides segments. Especially, NCL is the most powerful ligation method for synthesizing proteins (native or modified) of moderate size (i.e., small proteins < 200 AA).

*Originated from Intein splicing*
Spiegelmers: L-RNA

Aptamers (from the Latin aptus – fit, and Greek meros – part) are oligonucleotide or peptide molecules that bind to a specific target molecule. Aptamers are usually created by selecting them from a large random sequence pool, but natural aptamers also exist in riboswitches.

An L-ribonucleic acid aptamer (L-RNA aptamer, trade name Spiegelmer – from German Spiegel "mirror" – by Noxxon Pharma) is an RNA-like molecule built from L-ribose units. It is an artificial oligonucleotide named for being a mirror image of natural oligonucleotides.

L-RNA aptamers are a form of aptamers. Due to their L-nucleotides, they are highly resistant to degradation by nucleases. Spiegelmers are considered potential drugs and are currently being tested in clinical trials.

A. Vater, S. Klussmann, Drug Discovery Today 2015, 20, 147-155
Spiegelmers: L-RNA

Spiegelmers: L-RNA

Aptamer selection against RNA with natural D-chirality

Spiegelmers/“AptamiRs” - a new class of miR inhibitors.

Spiegelmers with modified nucleotides


Synthesis of a mirrored 832-residue Taq polymerase or the 604-residue Klenow Fragment from D-aminoacids is still beyond reach with the current methods. However, progress in long peptide synthesis enabled solid-support synthesis followed by fragment ligation (NCL) of the mirror image configuration of polymerase X from African swine fever virus (ASFV), the shortest known polymerase (174 amino acids).

PDB: 1JQR

**L-DNA polymerase**

**D-ASFV Pol X** elongated an L-DNA primer with L-dNTPs - a functional 56-mer L-DNAzyme was made within 36 hours.

This poses an important proof of concept, however, polymerase X is a thermo-labile repair enzyme and its catalytic activity does not meet the requirements for a standard PCR.

---

**A thermostable L-DNA-polymerase**

Cell of *Sulfolobus* infected by virus STSV1 observed under microscopy. Two spindle-shaped viruses were being released from the host cell.

A thermostable mirror-image polymerase **D-Dpo4-3C (352 aa)** is able to amplify L-DNA in a classical PCR reaction and can be used to assemble an L-DNA gene from L-DNA oligonucleotides.

Tyr12Ala or Tyr12Ser \(\rightarrow\) DNA-dependent RNA polymerase

---

**Assembly of a mirror-image gene**

A thermostable mirror-image polymerase **D-Dpo4-3C (352 aa)** is able to amplify L-DNA in a classical PCR reaction and can be used to assemble an L-DNA gene from L-DNA oligonucleotides.

Tyr12Ala or Tyr12Ser \(\rightarrow\) DNA-dependent RNA polymerase

---

A thermostable L-DNA-polymerase

A. Pech, S. Klussmann et al. Nucl. Acid Res. 2017, 45, 3997-4005
A functional DNA-ligase in the D-enantiomeric conformation has been synthesized. It exhibited DNA ligation activity on chirally inverted nucleic acids in L-conformation, but not acting on natural substrates and with natural co-factors. The ligase was based on the known structure of the *Paramecium bursaria* chlorella virus DNA-ligase and the homologous but shorter DNA-ligase of *Haemophilus influenza*. The structure and the activity of the mirror-image ligase were characterized, documenting its enantiospecific functionality.

J. Weidmann *et al.*
*Cell Chemical Biology* 2019 26(5), 645-651.e3
D-Protein LigA: (A) The product of the final chemical ligation after 12 h. The two peptides 5 and 10 and the final, full-length protein of about 30 kDa can be seen. (B) MALDI-TOF mass spectrum of LigA. Measured: 29,974.5 Da theoretical: 29,966.1 Da.

J. Weidmann et al.
Cell Chemical Biology 2019 26(5), 645-651.e3
The transcription of a mirror-image gene into L-RNA, as well as reverse transcription of L-RNA into L-DNA by synthetic D-polymerases, based on designed mutants of Dpo4, have been demonstrated. The efficient mirror-image transcription system may enable enzymatic preparation of L-RNA molecules to further enable clinical applications of nuclease-resistant aptamer biosensors and drugs or studies on mirror-image or cross-chiral ribozymes and aptamers. The enzymatically transcribed L-5S rRNA shown in this study could be used as one component in a future effort to assemble a mirror-image ribosome, a step toward the realization of mirror-image life.
**Mirror-image transcription, reverse transcription and amplification**

*Sulfolobus solfataricus* P2 DNA polymerase IV (Dpo4)

Currently, the biggest challenge in establishing a mirror-image version of the central dogma is to build a mirror-image ribosome-based translation machine.

The natural and mirror-image versions of three ribosomal proteins (L5, L18, and L25) in the large subunit of the *Escherichia coli* ribosome have been chemically synthesized. The synthetic mirror-image proteins can fold *in vitro* and assemble with enzymatically transcribed mirror-image 5S ribosomal RNA into ribonucleoprotein complexes.
In addition, the RNA–protein interactions are chiral-specific in that the mirror-image ribosomal proteins do not bind with natural 5S ribosomal RNA and vice versa. The synthesis and assembly of mirror-image 5S ribonucleoprotein complexes are important steps towards building a functional mirror-image ribosome.

However, the large subunit of bacterial ribosome alone is composed of a 5S RNA subunit (120 nucleotides), a 23S RNA subunit (2900 nucleotides) and 31 proteins...

J.J. Ling et al., Angew. Chem. Int. Ed. 2020, 59 (9), 3724-3731
**Phage display**

Phage display is a laboratory technique for the study of protein–protein, protein–peptide, and protein–DNA interactions that uses bacteriophages (viruses that infect bacteria) to connect proteins with the genetic information that encodes them. A gene encoding a protein of interest is inserted into a phage coat protein gene, causing the phage to "display" the protein on its outside while containing the gene for the protein on its inside, resulting in a connection between genotype and phenotype. These displaying phages can then be screened against other proteins, peptides or DNA sequences, in order to detect interaction between the displayed protein and those other molecules. In this way, large libraries of proteins can be screened and amplified in a process called in vitro selection, which is analogous to natural selection.

The most common bacteriophages used in phage display are M13 and fd filamentous phage, T4, T7, and λ phage.


**Nobel prize in chemistry 2018**

1) fusion proteins for a viral coat protein + the gene to be evolved (typically an antibody fragment) are expressed in bacteriophage.
2) the library of phage are washed over an immobilised target.
3) the remaining high-affinity binders are used to infect bacteria.
4) the genes encoding the high-affinity binders are isolated.
5) those genes may have random mutations introduced and used to perform another round of evolution. The selection and amplification steps can be performed multiple times at greater stringency to isolate higher-affinity binders.
Polypeptides composed entirely of D-amino acids (D-proteins) inherently have in vivo properties that are proposed to be near-optimal for a large molecule therapeutic agent. Specifically, D-proteins are resistant to degradation by proteases and are anticipated to be nonimmunogenic. Furthermore, D-proteins are manufactured chemically and can be engineered to have other desirable properties, such as improved stability, affinity, and pharmacokinetics.

The 204-residue covalent-dimer vascular endothelial growth factor (VEGF) with full mitogenic activity was prepared from three unprotected peptide segments by one-pot native chemical ligations. The covalent structure of the synthetic VEGF was confirmed by precise mass measurement, and the three-dimensional structure of the synthetic protein was determined by high-resolution X-ray crystallography.

VEGF is a signal protein produced by cells that stimulates the formation of blood vessels. Cancers that can express VEGF are able to grow and metastasize. Overexpression of VEGF can cause vascular disease in the retina of the eye and other parts of the body.

L-VGEF: K. Mandal, S. Kent Angew. Chem., Int. Ed. 2011, 50(35), 8029-8033
D-VGEF: K. Mandal, S. Kent et al. PNAS 2012, 109 (37), 14779-14784
RFX037.D is a \textit{D}-protein antagonist of natural vascular endothelial growth factor A (VEGF-A) that inhibited binding to its receptor, with extreme thermal stability ($T_m > 95 \degree C$) and high affinity for VEGF-A ($K_d = 6 \text{ nM}$). It was identified using a combination of total chemical protein synthesis and mirror image phage display of proteins. Comparison of the two enantiomeric forms of RFX037 revealed that the \textit{D}-protein is more stable in mouse, monkey, and human plasma and has a longer half-life \textit{in vivo} in mice. Significantly, RFX037.D was nonimmunogenic in mice, whereas the \textit{L}-enantiomer generated a strong immune response. These results confirm the potential utility of synthetic \textit{D}-proteins as alternatives to therapeutic antibodies.

\textit{D-proteins: almost ideal therapeutic agents}


X-ray crystal structure of RFX037:VEGF-A heterochiral protein complex. (A) Two RFX037.L molecules (yellow) bound to one d-VEGF-A homodimer (green) and two RFX037.D molecules (blue) bound to one l-VEGF-A homodimer (cyan). (B) Superposition of RFX037.D (blue) and RFX001.D (gray, rcsb accession 4GLS). (C) RFX037.D side chains (shown as sticks) that contact l-VEGF-A. (D) The contact surfaces of l-VEGF-A to RFX037.D (blue), VEGFR1 (salmon), or both (red). (E) Hydrogen bond networks formed by intramolecular polar contacts originated from additional N- and C-terminal residues in RFX037.D.

An immune checkpoint TIGIT is a novel target in cancer immunotherapy of anti-PD-1 resistant tumors. Here, mirror-image phage display bio-panning was performed using the $d$-enantiomer of TIGIT synthesized by hydrazide-based native chemical ligation.

The 3D structure of $^{d}$IgV$^{TIGIT}$ represented as a mirror image of the reported crystal structure (PDB entry: 3UDW).

X. Zhou et al., Angew. Chem. Int. Ed. 2020, 59 (35), 15114-15118
Using the mirror-image phage display technique, a D-peptide DTBP-3 was identified, which could occupy the binding interface and effectively block the interaction of TIGIT with its ligand PVR. DTBP-3 showed proteolytic resistance, tumor tissue penetrating ability, and could inhibit tumor growth and metastasis in anti-PD-1 resistant tumor model.

The negative signaling pathway mediated by immune checkpoints (such as PD-1/PD-L1) leads to the exhaustion of immune cells and immune escape of cancer. However, the therapeutic effects of PD-1/PD-L1 blockade are limited (<30% response rates), and adaptive resistance is often observed. TIGIT is a novel immune checkpoint molecule expressed on Nk and T cells, which competes with costimulatory receptor CD226 for the shared ligand PVR to deliver immunosuppressing signals and a potential therapeutic target.

A biotin labeled D-enantiomer of the immunoglobulin variable domain of TIGIT (D\text{IgV}^{\text{TIGIT}}\text{-biotin}) was synthesized by hydrazide-based native chemical ligation (NCL) and removable backbone modification (RBM) strategy.

X. Zhou et al., Angew. Chem. Int. Ed. 2020, 59 (35), 15114-15118
D-peptide for cancer immunotherapy

**DTBP-3 is the first D-peptide targeting TIGIT, which could serve as a potential candidate for cancer immunotherapy.**

C) Near-infrared fluorescence imaging of CT26 tumor-bearing mice injected (i.v.) with Cy5.5 and DTBP-3-Cy5.5. D) Representative imaging and E) average fluorescent intensity (n=3) of tissues 24 h post injection. F) Representative fluorescence microscopy images of sectioned tumors. Scale bar=200 mm.

The effects of DTBP-3 on anti-PD-1 responsive and resistant tumor models. Tumor growth curves of B16-OVA (A) or 4T1 (B). C) Representative images and statistics of lung metastases loci in 4T1 tumor model (*P<0.05).

CHAPTER 1

OLIGONUCLEOTIDES

Part 3 – noncanonical backbone – Xeno Nucleic Acids
XNA – Xeno Nucleic Acids

Hexitol Nucleic Acid (HNA)
Threose Nucleic Acid (TNA)
Peptide Nucleic Acid (PNA)

(Fluoro) Arabino Nucleic Acid
Locked Nucleic Acid
Apio Nucleic Acid
Cyclohexene Nucleic Acid
Altritol Nucleic Acid
Glycerol Nucleic Acid

A
DNA:DNA
DNA:RNA
RNA:RNA

C
ANA:RNA
CeNA:RNA
FANA:RNA
HNA:RNA
LNA:RNA
PNA:DNA
PNA:RNA

N1
CeN:CeNA
hDNA:hDNA
FRNA:FRNA
GNA:GNA
HNA:HNA
LNA:LNA
PNA:PNA

2′
TNA:TNA
dXyNA:dXyNA
XyNA:XyNA

3′
FAF:FAF
FRNA:FRNA (alt)
FRNA:FANA (chim)
Overview of XNA

The 2'→4' linked pentopyranosyl family

- Homo-DNA: \( \beta-D, 4' \rightarrow 6' \)
- Altropyranosyl-NA: \( \beta-D, 4' \rightarrow 6' \)
- \( \beta-D-Ribopyranosyl-NA \) (pRNA)
- \( \beta-D-Xylopyranosyl-NA \)
- \( \beta-D-Ribopyranosyl-NA \) \( 3' \rightarrow 4' \)
- \( \alpha-L-Lyxopyranosyl-NA \) \( 3' \rightarrow 4' \)
- Glucopyranosyl-NA: \( \beta-D, 4' \rightarrow 6' \)
- Allopyranosyl-NA: \( \beta-D, 4' \rightarrow 6' \)
- \( \alpha-L-Lyxopyranosyl-NA \)
- \( \alpha-L-Arabinoxyranosyl-NA \)
- \( \alpha-L-Threofuranosyl-NA \) TNA

Review on the oligonucleotide modifications:
A. Eschenmoser Angew. Chem., Int. Ed. 2011, 50, 12412-12472
Overview of XNA

**XNA – Xeno Nucleic Acids**

**XNA** - synthetic alternative to DNA and RNA as information-storing biopolymers that differs in the sugar backbone.
- at least 6 XNAs can store and retrieve genetic information
- Ongoing research to create synthetic polymerases to transform XNA

**Xenobiology**
- (XNA) as information carriers, expanded genetic code and, incorporation of non-proteinogenic amino acids into proteins
- the **origin of life**: Primordial soup → (XNA) → RNA → RNA(+DNA)+Proteins
- development of industrial production systems with novel capabilities (pathogen resistance, biopolymer engineering)
- „genetic firewall“ – excludes the risk of contaminating currently existing organisms (horizontal gene transfer)

The **long-term goal** - a cell that stores its genetic information on XNA, with different base pairs, using non-canonical amino acids and an altered genetic code. So far cells have been constructed that incorporate only one or two of these features
**XNA – Xeno Nucleic Acids**

*XNA are not recognized by natural polymerases.*

One of the major challenges is to find or create novel types of polymerases that will be able to replicate these new-to-nature constructs. The method of polymerase evolution and design successfully led to the storage and recovery of genetic information (of less than 100bp length) from six alternative genetic polymers based on simple nucleic acid architectures not found in nature.

*XNA aptamers*, which bind their targets with high affinity and specificity, were also selected, demonstrating that beyond heredity, specific XNAs have the capacity for Darwinian evolution and folding into defined structures.

Thus, heredity and evolution, two hallmarks of life, are not limited to DNA and RNA but are likely to be emergent properties of polymers capable of information storage.

**Engineering XNA polymerases**

TgoT, a variant of the replicative polymerase of *Thermococcus gorgonarius*

(A) Sequence alignments showing mutations from wtTgo in polymerases Pol6G12 (red), PolC7 (green), and PolD4K (blue).

(B) Mutations are mapped on the structure of Pfu (PDB: 4AIL).

Yellow - template; dark blue - primer; orange - mutations present in the parent polymerase TgoT

HNA synthesis

Pol6G12 extends the primer (p) incorporating 72 hNTPs against template T1 to generate a full-length hybrid molecule with a 37,215-dalton expected molecular mass.

HNA reverse transcription (DNA synthesis from an HNA template). Polymerase-synthesized HNA (from template YtHNA4) is used as template by RT521 for HNA-RT

**XNA genetic polymers.**

Structures and PAGE of synthesis (+72 xnt), and reverse transcription (+93 nt)

(E) PAGE of LNA synthesis [primer (41 nt) + 72 Int] and LNA-RT (red). LNA synthesis (green) migrates at its expected size (113 nt) and comigrates with reverse transcribed DNA (red) synthesized from primer PRT2 (20 nt).

XNART–polymerase chain reaction. Amplification products of expected size (133 base pairs) are obtained only with both XNA forward synthesis and RT (RT521 or RT521K)

Characterization of HNA aptamers. Anti-TAR aptamer T5-S8-7 and anti-HEL aptamer LYS-S8-19.

(A and B) Aptamer binding specificity against TAR variants (red, sequence randomized but with base-pairing patterns maintained) and different protein antigens (human lysozyme, HuL; cytochrome C, CytC; streptavidin, sAV; biotinylated-HEL bound to streptavidin, sAV-bHEL). OD, optical density.

(C) Affinity measurements of aptamer binding by SPR. RU, response units.

(D) FACS analysis of fluorescein isothiocyanate (FITC)–labeled aptamers binding to plasmacytoma line J558L with and without expression of membrane-bound HEL (mHEL). wt, wild type.
**XNA – Xeno Nucleic Acids**

XNA – complementarity to DNA, also used as genetic catalysts.

FANA, HNA, CeNA and ANA - cleave RNA (*XNAzymes*).

FANA *XNAzymes* can also ligate DNA, RNA and XNA substrates.

**Chemical synthesis yields an active RNA endonuclease XNAzyme**

Secondary FANAzyme structure

1. 1 and 3 - matching RNA
2. 2 and 4 - scrambled RNA

Multiple turnovers

**An RNA ligase XNAzyme (FANA)**

FANA XNAzymes can also ligate DNA, RNA and XNA substrates.

---

*Secondary FANAzyme structure*

---

Pre-steady state trimolecular reaction rate ($k_{obs}$) at 25 °C ($n = 3$; error bars, s.d.).

---

XNA–XNA ligase XNAzyme (FANA): catalysis without natural nucleic acids

XNA–XNA ligase XNAzyme (FANA): catalysis without natural nucleic acids

**XNA–XNA ligase XNAzyme (FANA): catalysis without natural nucleic acids**

FpImR4_2-catalysed oligomerization of XNA (FANA) substrates

XNAzyme-catalysed assembly of an active XNAzyme. A variant XNA ligase (FpImR4_2mut) catalyses ligation (lane 2) of FANA substrates LigS1F NUC and LigS2F NUC. The product (LigPF NUC) is a variant of XNAzyme FR17_6 min (Fig. 2), which cleaves RNA substrate NucSVR (lanes 5 and 6), but not scrambled RNA (NucSR SCRAM2)(lanes 3 and 4).

### Table 1: Polymerase-mediated synthesis of XNAs

<table>
<thead>
<tr>
<th>Pol Family</th>
<th>Polymerase</th>
<th>Novel Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pol A</td>
<td>Taq, Tth, Pol B</td>
<td>2F RNA, 2′OMe RNA, 2′-azido RNA</td>
</tr>
<tr>
<td>Pol B</td>
<td>Tgo, KOD, 9N, Pfu, phi29</td>
<td>CeNA, LNA, phNA, HNA, FANA, CyDNA, 2F RNA, ANA, TNA, 2′azido RNA, tPhoNA</td>
</tr>
<tr>
<td>Pol Y</td>
<td>(D-aa) Dpo4</td>
<td>L-DNA</td>
</tr>
<tr>
<td>Pol X</td>
<td>(D-aa) ASFV pol</td>
<td>L-DNA, L-RNA</td>
</tr>
<tr>
<td>Pol X</td>
<td>T7 RNAP, SynS</td>
<td>2F RNA, 2′OMe RNA, Ds-Pa UBP</td>
</tr>
<tr>
<td>Pol X</td>
<td>HIV-RT</td>
<td>pyDAD-puADA UBP</td>
</tr>
</tbody>
</table>

K. Duffy, S. Arangundy-Franklin, P. Holliger *BMC Biology, 2020, 18, Art.# 112*
<table>
<thead>
<tr>
<th>Drug name (trade name)</th>
<th>Target</th>
<th>Modifications</th>
<th>Mechanism</th>
<th>Indication</th>
<th>Approval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pegaptanib (Macugen)</td>
<td>Vascular endothelial growth factor (VEGF)</td>
<td>2′F, 2′OMe, PEG conjugate</td>
<td>Aptamer</td>
<td>Neovascular (wet) age-related macular degeneration</td>
<td>FDA approved (2004)</td>
</tr>
<tr>
<td>Mipomersen (Kynamro)</td>
<td>Apolipoprotein B-100 mRNA</td>
<td>2′MOE, PS, 5mC</td>
<td>ASO (RNase H)</td>
<td>Homozygous familial hypercholesterolemia</td>
<td>FDA approved (2013)</td>
</tr>
<tr>
<td>Eteplirsen (Exondys 51)</td>
<td>Exon 51 in dystrophin mRNA</td>
<td>PMO</td>
<td>ASO (splicing modulation)</td>
<td>Duchenne muscular dystrophy</td>
<td>FDA approved (2016)</td>
</tr>
<tr>
<td>Nusinersen (Spinraza)</td>
<td>Survival of motor neuron 2 (SMN2) pre-mRNA</td>
<td>2′MOE, PS, 5mC</td>
<td>ASO (splicing modulation)</td>
<td>Spinal muscular atrophy</td>
<td>FDA (2016) and EMA (2017) approved</td>
</tr>
<tr>
<td>Patisiran (Onpattro)</td>
<td>Transthyretin (TTR) mRNA</td>
<td>2′OMe</td>
<td>siRNA</td>
<td>Hereditary transthyretin-mediated amyloidosis</td>
<td>FDA and EMA approved (2018)</td>
</tr>
<tr>
<td>Inotersen (Tegsedi)</td>
<td>Transthyretin (TTR) mRNA</td>
<td>2′MOE, PS, 5mC</td>
<td>ASO (RNase H)</td>
<td>Hereditary transthyretin-mediated amyloidosis</td>
<td>FDA and EMA approved (2018)</td>
</tr>
<tr>
<td>Volanesorsen (Waylivra)</td>
<td>Apolipoprotein C3 (apo-CIII) mRNA</td>
<td>2′MOE, PS, 5mC</td>
<td>ASO (RNase H)</td>
<td>Familial chylomicronemia syndrome</td>
<td>EMA approved (2019)</td>
</tr>
<tr>
<td>Givosiran (Givlaari)</td>
<td>Aminolevulinate synthase 1 (ALAS1) mRNA</td>
<td>PS, 2′F, 2′OMe, GalNAc conjugate</td>
<td>siRNA</td>
<td>Acute hepatic porphyria</td>
<td>FDA approved (2019)</td>
</tr>
<tr>
<td>Golodirsen (Vyondys 53)</td>
<td>Exon 53 in dystrophin mRNA</td>
<td>PMO</td>
<td>ASO (splicing modulation)</td>
<td>Duchenne muscular dystrophy</td>
<td>FDA approved (2019)</td>
</tr>
</tbody>
</table>
Peptidonucleic acids – functional DNA analogues

PNA – stable ex vivo, the backbone detected in cyanobacteria

Applications: antigene, antisense agents; fluorescent DNA probes (FISH), anticancer, antiviral, antibacterial, antiparasitic agents; diagnostics, mol. biology
Structural modifications of the PNA - \( \alpha \text{GPNA, } \gamma \text{GPNA} \)

- \( \alpha \text{GPNA} \): Alkylguanidinium residues (Arg side chains)
  - enhanced water solubility
  - cell permeability (analogous to oligoarginine CPPs)
- \( \gamma \text{GPNA} \)

\[ \begin{align*}
\alpha \text{GPNA} & : \quad X \quad \alpha \quad D \text{-arginine} \\
\gamma \text{GPNA} & : \quad X \quad \gamma \quad L \text{-arginine} \\
\end{align*} \]

\[ R = \text{Alkylguanidinium residues} \]

\[ X = A; \ C; \ G; \ T \]
Cell-penetrating αGPNA

HeLa cells incubated with 1 μM GPNA (FITC-DCCDACCCTDCTDGCDCACDGDCGTT-NH₂) for 16 h, fixed, stained with DAPI. Nuclei (blue), GPNA (green).

• GPNA 16-mers targeting the epidermal growth factor receptor (EGFR) in preclinical models as therapeutic modality for head and neck squamous cell carcinoma (HNSCC) and nonsmall cell lung cancer (NSCLC)
• Elicited potent antisense effects in NSCLC and HNSCC preclinical models
• When administered intraperitoneally in mice, EGFRAS-GPNA was taken-up by several tissues including the xenograft tumor
• Systemic administration of EGFRAS-GPNA induced antitumor effects in HNSCC xenografts, with similar efficacies as the FDA-approved EGFR inhibitors: cetuximab and erlotinib.

D. Ly et al. ACS Chem. Biol. 2013, 8, 345-352
Cell-penetrating αGPNA for in vivo catalytic oligonucleotide sensing

Z. Pianowski, N. Winssinger Chem. Comm. 2007, 37, 3820-3822
Cell-penetrating αGPNA for in vivo catalytic oligonucleotide sensing

Inside living cells

A, B – controls (+/-) C – matching PNA D – mismatched PNA

(A) For nucleic acids replicators, templating is based on base pairing, so the formation of a mutant template is rare. Once formed, the mutant replicator forms a competing replication cycle. (B) For a peptide replicator, templating is less exact, so the formation of a mutant template is common. The mutant template can catalyze formation of mutant progeny or parental progeny, and the two species form a mutualistic network.

Nonenzymatic templated nucleic acid synthesis – monomer/short oligomer

Problems:
- very slow reactions
- limited range of templates (mostly C-rich)
- poor regiospecificity (2’-5’ linkages, predominant in some cases)
- 3’-aminonucleotides perform better, but undergo intramolecular cyclizations as side reaction

**Templated nucleic acid synthesis – short oligomer coupling**

Limitations:
- slightly distorted backbone (amine instead of amide backbone every 5 bases)
- only carefully designed pentamers work – limiting the diversity for functional selection

Templated nucleic acid synthesis – base filling

Advantages:
- no cross-reactivity
- selectivity increased by proximity of the reaction to the hybridization site

Limitations:
- single or double abasic sites (most efficient inside of the chain)
- Aldehydes give better yields and accuracy, but worse hybridization of the product

A polyamide responsive to selection pressure

A polyamide responsive to selection pressure

Dynamic polymer responsive to template changes with high fidelity

The DNA primer region affords a high level of control over the location and register of the tPNA backbone in relation to the template strand.

Templated Self-Assembly of Dynamic Peptide Nucleic Acids

Dynamic and Responsive DNA-like Polymers

A. Formation of tel-DNA by ROP

B. Comparison of DNA with tel-DNA

- 6-membered periodicity
- uncharged backbone
- possible anti-sense polymer
- dynamic and responsive polymers

C. Concept towards template directed synthesis (TDS)

C: use of a thiolate initiator to form a dynamic library of tel-DNA oligomers in situ. Addition of a defined DNA template will equilibrate the dynamic mixture to a thermodynamically favorable complementary sequence defined polymer.

Sequence-Controlled Synthesis of Clickable Synthetic Oligonucleotides

\[
\begin{align*}
\text{1} \quad \xrightarrow{\text{NaOH, DCM}} \quad \text{2} & \quad \xrightarrow{\text{DBU, DCM, 66\%}} \quad \text{3} \\
\text{3} \quad \xrightarrow{\text{N(Boc)_2, DMAP, DCM, then MeOH, reflux, 94\%}} \quad \text{4} & \quad \xrightarrow{\text{DBU, DCM, 77\%}} \quad \text{5} \\
\text{5} \quad \xrightarrow{\text{O(Boc)_2, DMAP, DCM}} \quad \text{6} & \quad \xrightarrow{\text{DBU, DCM, 61\%}} \quad \text{7} \\
\text{7} \quad \xrightarrow{\text{NaOH, MeOH, THF, Bu_NI, 60 \degree C}} \quad \text{8} & \quad \xrightarrow{\text{KSAc, NaI, DMF, 50 \degree C, 80\%}} \quad \text{9} \\
\text{11} \quad \xrightarrow{\text{O(Boc)_2, DMAP, DCM}} \quad \text{12} & \quad \xrightarrow{\text{DBU, DCM, 81\%}} \quad \text{13} \\
\text{12} \quad \xrightarrow{\text{NaOH, MeOH, THF, Bu_NI, 60 \degree C, 31\%}} \quad \text{14} & \quad \xrightarrow{\text{KSAc, NaI, DMF, 50 \degree C, 83\%}} \quad \text{15} \\
\text{17} \quad \xrightarrow{\text{NaOH, MeOH, DMPA, 365 nm, DCM}} \quad \text{18}
\end{align*}
\]

DCM = dichloromethane, TEA = triethylamine, DBU = 1,8-diazabicyclo[5.4.0]undec-7-ene, Boc = tert-butyloxycarbonyl, Ac = acetyl, DMF = dimethyl formamide

X. Han et al. Macromol. Rapid Commun. 2020, 41, 2000327
Sequence-Controlled Synthesis of Clickable Synthetic Oligonucleotides

X. Han et al. Macromol. Rapid Commun. 2020, 41, 2000327