Artificial genetic polymers



Intein splicing

An **intein** is a segment of a protein that is able to excise itself and join the remaining portions (the **exteins**) with a peptide bond in a process termed protein splicing. Inteins have also been called "protein introns".

Intein-mediated protein splicing occurs after the intein-containing mRNA has been translated into a protein. This precursor protein contains three segments—an **N-extein** followed by the intein followed by a **C-extein**. After splicing has taken place, the resulting protein contains the N-extein linked to the C-extein; this splicing product is also termed an extein.



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).



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



A. Vater, S. Klussmann, *Drug Discovery Today* **2015**, *20*, 147-155 B.E. Young, N. Kundu, J.T. Szczepanski, *Chem. Eur.J.* **2019**, *25*, 7981–7990

Mirror-image RNA that binds D-Adenosine





Figure 5. Stability of the 58-mer RNA ligands in human serum. (A) Aptamer D-A42d and (B) spiegelmer L-A42d. Aliquots were taken at the indicated times. L marks the size standard (10 bp DNA ladder). The results were reproduced in an independent experiment.

S. Klussmann, J. Furste, et al. Nature Biotech. 1996, 14, 1112-1115

Selected pharmaceutically relevant targets of spiegelmers

Monocyte chemoattractant peptide MCP-1/CCL2 NOX-E36/ emapticap pegol

Placebo-controlled Phase I, SAD and four-week MAD completed, PDeffect:reduction of CCR2. monocytes in peripheral blood. Double-blind, placebo-controlled Phase IIa trial in diabetic patients with albuminuria (three months treatment, three months follow-up): reductions in albuminuria and improved glycemic control(HbA1c); renoprotective effect maintained in absence of drug (three months)

Stromal-cell-derived factor-1 (SDF-1/CXCL12)

NOX-A12/ olaptesed pegol

Hematopoietic stem cell/WBC mobilization in mice and/or monkeys; Phase I, SAD: hematopoietic stem cell and WBC mobilization Phase IIa for chemosensitization in MM (with VD) and CLL (with BR) ongoing

Hepcidin NOX-H94/ lexaptepid pegol

Phase I:dose-dependent increases in serum iron in healthy Volunteers; PK/PD study: inhibition of LPS-induced serum iron decrease Phase IIa in anemia of cancer ongoing

A. Vater, S. Klussmann, Drug Discovery Today 2015, 20, 147-155

Aptamer selection against RNA with natural D-chirality



L-6-4t-TAR → TAR → An *L*-RNA aptamer was developed that binds the natural dform of the HIV-1 trans-activation responsive (TAR) RNA (K_d = 100 nM). The aptamer initially was obtained as a *D*-aptamer against *L*-TAR RNA through *in vitro* selection.

It binds *D*-TAR exclusively at the six-nucleotide distal loop, but does so through tertiary interactions rather than simple Watson–Crick pairing.

This complex is the first example of two nucleic acids molecules of opposing chirality that interact through a mode of binding other than primary structure.

Binding of the *L*-aptamer to *D*-TAR RNA inhibits formation of the Tat-TAR ribonucleoprotein complex that is essential for TAR function \rightarrow therapeutic interest

J.T. Szczepanski, G.F. Joyce, J. Am. Chem. Soc. 2013, 135, 36, 13290–13293

Spiegelmers/"AptamiRs" - a new class of miR inhibitors.



In vitro selection was used to obtain *L*-RNA aptamers that bind the distal stem-loop of various precursor microRNAs (premiRs). These *L*-aptamers, termed "aptamiRs", bind their corresponding pre-miR target through highly specific tertiary interactions rather than Watson–Crick pairing. Formation of a pre-miR–aptamiR complex inhibits Dicer-mediated processing of the pre-miR, which is required to form the mature functional microRNA. One of the aptamiRs, which was selected to bind oncogenic pre-miR-155, inhibits Dicer processing under simulated physiological conditions, with an IC₅₀ of 87 nM

MicroRNAs (miRs) are small, noncoding RNAs that act as post-transcriptional regulators of gene expression, involved in development, differentiation, and apoptosis. Alterations in their expression patterns can contribute to the pathogenesis of human disease.



J.T. Szczepanski, G.F. Joyce, J. Am. Chem. Soc. 2015, 137, 51, 16032–16037

Spiegelmers with modified nucleotides

D-19a-6



By employing 5-aminoallyl-UTP during mirror image in vitro selection, a modified *L*-RNA aptamer was isolated with exceptional affinity (IC_{50} = 4nM) towards oncogenic precursor microRNA-19a. These studies demonstrate that expanding the chemical functionality of *L*-aptamers enables development of robust RNA-binding reagents.



Processing of the mirrored genetic information

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

Showalter A.K. et al. *Nat Struct Biol* **2001**, *8*, 942-946 Z. Wang, W. Xu, L. Liu, T. F. Zhu *Nature Chem.* **2016**, *8*, 698-704



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.





Z. Wang, W. Xu, L. Liu, T. F. Zhu Nature Chem. 2016, 8, 698-704



L-DNA polymerase

a, Template-directed primer extension by synthetic L-ASFV pol X (natural system) and D-ASFV pol X (mirror-image system) with the corresponding D- and L-DNA primers, templates and dNTPs. **b**, Repeated cycles of polymerization by D-ASFV pol X: **c**,**d**, The nucleotide substrate specificities of synthetic L- and D-ASFV pol X. **e**, Chiral specificity assay with different chiral combinations of polymerases, primer/template pairs and dNTPs.

L-DNA polymerase



a,b, Primer extension by synthetic L- and D-ASFV pol X with the corresponding D-DNA primer (5'-Cy5 labelled) and L-DNA primer (5'-FAM labelled), templates and dNTPs. **c**, The above two polymerization reactions were carried out in a racemic mixture under the same conditions as described above, with the L- and D-ASFV pol X, D- and L-primers, D- and L-templates and D- and L-dNTPs added, incubated for up to 4 h at 37 °C.

L-DNAzyme



a, Sequence and predicted secondary structure of the previously reported Zn²⁺-dependent self-cleaving DNAzyme.
 b, Primer extension on a 66 nt template to produce the Zn²⁺-dependent self-cleaving DNAzyme.
 c, Self-cleavage of the enzymatically polymerized Zn²⁺-dependent D- and L-DNAzymes.

A thermostable L-DNA-polymerase

A thermostable mirror-image polymerase **D-Dpo4-3C** has been produced, that is able to amplify L-DNA in a classical PCR reaction and can even be used to assemble an L-DNA gene from L-DNA oligonucleotides. This artificial enzyme is a mutant of DNA polymerase IV from *Sulfolobus solfataricus*, a Y-family polymerase consisting of 352 amino acids, the longest protein made by chemical synthesis thus far.



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

Furthermore, with an additional single point mutation (Tyr12Ala or Tyr12Ser), this DNA polymerase can be tuned to accept also ribonucleotides as substrates with reasonable efficiency. Thus, this enzyme may be hijacked to act as a DNA-dependent RNA polymerase to prepare longer stretches of *L*-RNA

A. Pech, S. Klussmann et al. Nucl. Acid Res. 2017, 45, 3997-4005

A thermostable L-DNA-polymerase



Synthesis strategy for d-Dpo4-3C. (**A**) five fragments were synthesized and assembled as follows: (i) native chemical ligation (NCL) of fragments 1 and 2. Isolated yield \approx 18%. (ii) Segment condensation of fully protected fragments 4 and 5 followed by deprotection. Isolated yield \approx 15%. (iii) NCL of fragments 3 and 4•5 followed by Z-deprotection. Isolated yield \approx 25%. (iv) Thioester-conversion of fragment 1•2 and NCL with fragment 3•4•5. Isolated yield: 10%. (v) Folding. (**B**) sequence of d-Dpo4-3C; coloring as in panel A. (**C**) folded d-Dpo4-3C (artist impression based on PDB 3PR4 (31)).

A. Pech, S. Klussmann et al. Nucl. Acid Res. 2017, 45, 3997-4005

A thermostable L-DNA-polymerase



Assembly of a mirror-image gene. (**A**) schematic of the oligonucleotide setup. (**B**) lane 1, 3 µl of 10 bp DNA ladder. Lane 2, mirror-image no-enzyme control. Lane 3, mirror-image gene assembly. Lane 4, empty. Lane 5, natural handedness no enzyme control. Lane 6, natural handedness gene assembly.

A. Pech, S. Klussmann et al. Nucl. Acid Res. 2017, 45, 3997-4005

Mirror-image DNA ligase

MANSDLMLLHTYNNQPIEGWVMSEKLDGVRGYWNGKQLLTRQGQRLSPPAYFIKDFPPFAIDGELFSERNHFEEISTITKCFKGDGW EKLKLYVFDVPDAEGNLFERLAKLKAHLLEHPT**CYIEIIEQIPVKDKTHLYQFLAQVENLQGEGVVVRNPNAPYERKRSSQILKLKT** ARGEQCTVIAHHKGKGQFENVMGALTCKNHRGEFKIGSGFNLNERENPPPIGSVITYKYRGITNSGKPRFATYWREKKGGGHHHHHH

Α

В

2 7 Activation Ligation Ligation hz-Cys conversion Desulfurization 138-195 9 3 Ligation Activation Thz-Cys conversion idation 10 5 Activation Ligation 11

A functional DNA-ligase in the Denantiomeric conformation has been synthesized. It exhibited DNA ligation activity on chiraly 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

Mirror-image DNA ligase



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.



Enzymatic Ligation of Gene Fragments Made of Synthetic L-DNA

J. Weidmann *et al. Cell Chemical Biology* **2019** *26(5), 645-651.e3*

Mirror-image transcription, reverse transcription and amplification



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*-**SS rRNA** shown in this study could be used as one component in a future effort to assemble a mirror-image life

Sulfolobus solfataricus P2 DNA polymerase IV (Dpo4)

M.Wang et al. Chem. 2019, 5 (4), 848-857



Mirror-image transcription, reverse transcription and amplification

Sulfolobus solfataricus P2 DNA polymerase IV (Dpo4)

M.Wang et al. Chem. 2019, 5 (4), 848-857

Mirror-image ribonucleoprotein complex



Α

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.



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

Mirror-image ribonucleoprotein complex



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.

G.P. Smith *Science* **1985**, *228* (4705), 1315–1317 J. Scott, G.P. Smith *Science* **1990**, *249* (4967), 386–390 J.W. Kehoe, B.K. Kay *Chem. Rev.* **2005**, *105* (*11*), 4056–72



Phage display cycle.

Thomas Shafee - Own work, CC BY 4.0

 fusion proteins for a viral coat protein + the gene to be evolved (typically an antibody fragment) are expressed in bacteriophage.
 the library of phage are washed over an immobilised target.
 the remaining high-affinity binders are used to infect bacteria.
 the genes encoding the high-affinity binders are isolated.
 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.

D-proteins: almost ideal therapeutic agents

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

D-proteins: almost ideal therapeutic agents



RFX037.D is a *D*-protein antagonist of natural vascular endothelial growth factor A (VEGF-A) that inhibited binding to its receptor, with extreme thermal stability ($T_m > 95$ °C) and high affinity for VEGF-A ($K_d = 6$ 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 *D*-protein is more stable in mouse, monkey, and human plasma and has a longer half-life *in vivo* in mice. Significantly, RFX037.D was nonimmunogenic in mice, whereas the *L*-enantiomer generated a strong immune response. These results confirm the potential utility of synthetic *D*-proteins as alternatives to therapeutic antibodies.

T.N.M. Schumacher, P. Kim *et al. Science* **1996**, *271 (5257)*, 1854-1857
M. Uppalapati, S. Kent *et al.*, ACS Chem. Biol. **2016**, *11*, 1058-1065



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 I-VEGF-A. (D) The contact surfaces of I-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.

M. Uppalapati, S. Kent et al., ACS Chem. Biol. 2016, 11, 1058-1065



D-peptide for cancer immunotherapy

An immune checkpoint TIGIT is a

(A) HHHHHH TG RVAPGPGI GL TLOSI T

D-peptide for cancer immunotherapy

A	Name	Sequence	Frequency (n/34)	Κ _D (μΜ)	Relative Blocking
	DTBP-1	LTPHKHHKHLHA	14	2.79 ± 0.51	1
	DTBP-3	GGYTFHWHRLNP	3	5.60 ± 2.56	16.2
	DTBP-6	AGMHVHYNWNHL	2	N/A	N/A
	DTBP-13	HGVALHIRLHAG	2	N/A	N/A
	DTBP-17	SAIHFHHPRWKP	2	48.3 ± 11.2	4.8
	DTBP-5	GNLTLHMHRSPS	1	6.75 ± 1.07	3.6
	DTBP-3S	NRHPWGLYGHTF		NB	1.6
в			Relative binding (%)	* I I	↓ ↓ ↓ ↓ ↓
c	Givi	Phe5 Phe5 Phe5 His6 Asn70 Asn58	235	Asn11 007 Ser80 Asp63	bro12

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 (^DIgV^{TIGIT}-biotin) was synthesized by hydrazide-based native chemical ligation (NCL) and removable backbone modification (RBM) strategy.

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.

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).

X. Zhou et al., Angew. Chem. Int. Ed. 2020, 59 (35), 15114-15118

Biology of the mirror world



Thermostable L-DNA/L-RNA polymerase and reverse transcriptase

Mirror-image ribonucleoprotein complex (→ mirrored ribosome)



