

Synergistic DNA- and Protein-Based Recognition Promote an RNA-Templated Bio-orthogonal Reaction

Niall M. McLoughlin,^[a, b] Arne Kuepper,^[c] Saskia Neubacher,^[a, b] and Tom N. Grossmann^{*[a, b, c]}

Abstract: Biomolecular assemblies composed of proteins and oligonucleotides play a central role in biological processes. While in nature, oligonucleotides and proteins usually assemble via non-covalent interactions, synthetic conjugates have been developed which covalently link both modalities. The resulting peptide-oligonucleotide conjugates have facilitated novel biological applications as well as the design of functional supramolecular systems and materials. However, despite the importance of concerted protein/oligonucleotide recognition in nature, conjugation approaches have barely utilized the synergistic recognition abilities of such com-

plexes. Herein, the structure-based design of peptide-DNA conjugates that bind RNA through Watson-Crick base pairing combined with peptide-mediated major groove recognition is reported. Two distinct conjugate families with tunable binding characteristics have been designed to adjacently bind a particular RNA sequence. In the resulting ternary complex, their peptide elements are located in proximity, a feature that was used to enable an RNA-templated click reaction. The introduced structure-based design approach opens the door to novel functional biomolecular assemblies.

Introduction

DNA, RNA and proteins are central biomacromolecules that possess distinct molecular recognition properties. For many biological processes, the supramolecular interplay between these biomacromolecules is essential, with complexes composed of proteins and nucleic acids playing diverse roles in transcription, translation and the modulation of gene expression. In these complexes, interaction partners usually adopt particular three-dimensional structures which define the shape of involved binding interfaces.^[1] For example, the minor and major groove of DNA and RNA duplexes are recognized by certain protein folds including the zinc-finger domain and the helix-loop-helix motif.^[2,3] While in nature, protein/nucleic acid complexes are usually assembled through non-covalent interactions, artificial hybrid-architectures have been created


through the covalent linkage of oligonucleotides and proteins or peptides.^[4–6] For example, synthetic conjugates have been generated to promote the cellular uptake of nucleic acids with cell-penetrating peptides,^[7,8] or to barcode peptide libraries with unique oligonucleotide sequences.^[9,10] The programmable Watson-Crick base-pairing abilities of oligonucleotides have also been used to control the conformation^[11,12] or the multivalent presentation of conjugated peptides,^[13–17] and to facilitate the DNA- and RNA-templated formation of peptide ligands.^[18,19]


In most synthetic conjugates, the peptide and oligonucleotide components are used orthogonally where each component contributes with a particular function.^[6] This contrasts with the situation in natural complexes where protein and nucleic acid-based recognition and function often work in concert. Only rarely, synthetic conjugates have utilized the synergistic binding of oligonucleotides and peptides. For example, conjugates have been devised whereby the oligonucleotide component forms a duplex with target DNA which then acts as a template for peptide binding.^[20,21] However, the assembly of larger and more complex conjugate structures with tunable properties has not been realized so far. Herein, we report biomolecular assemblies in which RNA serves as the scaffold to recruit conjugates composed of peptide and DNA components. These peptide-DNA conjugates bind the RNA through a synergy of Watson-Crick base-pairing and peptide-mediated major groove recognition. Two distinct types of conjugates were developed that adjacently bind the RNA target sequence to align their peptide components in close proximity. These conjugates were subsequently equipped with biorthogonal handles, to facilitate an RNA-templated reaction.

[a] N. M. McLoughlin, Dr. S. Neubacher, Prof. Dr. T. N. Grossmann
Department of Chemistry and Pharmaceutical Sciences
Vrije Universiteit Amsterdam
Amsterdam, 1081 HZ (The Netherlands)
E-mail: t.n.grossmann@vu.nl

[b] N. M. McLoughlin, Dr. S. Neubacher, Prof. Dr. T. N. Grossmann
Amsterdam Institute of Molecular and Life Sciences (AIMMS)
Vrije Universiteit Amsterdam
Amsterdam, 1081 HZ (The Netherlands)

[c] Dr. A. Kuepper, Prof. Dr. T. N. Grossmann
Chemical Genomics Centre of the Max Planck Society
Dortmund, 44227 (Germany)

 Supporting information for this article is available on the WWW under <https://doi.org/10.1002/chem.202101103>

 © 2021 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Results and Discussion

A fragment of TAV2b recognizes RNA/DNA duplexes

The protein TAV2b is a viral suppressor of RNA silencing that recognizes the major groove of double-stranded RNA and thereby binds duplexes in a sequence-independent manner.^[22] We previously identified a minimized fragment of TAV2b (TAV2b₂₁₋₅₃, hereby referred to as **sTAV2b**) capable of binding duplex RNA but not single-stranded RNA or duplex DNA.^[23] We showed that two **sTAV2b** peptides bind to a 19-mer RNA duplex while adopting a mainly α -helical conformation.^[23] Herein, we use the complex of two **sTAV2b** peptides with double-stranded RNA (Figure 1a) as the starting point for the design of peptide-DNA conjugates that bind single-stranded RNA.^[22,23] For initial experiments, we chose two complementary 22-nucleotide RNA sequences that form a duplex comprising 21-base pairs (RNA ^{α} /RNA ^{β} , Figure 1b). In addition, analogous duplexes composed of an RNA and a DNA strand (RNA ^{α} /DNA ^{β}) and of two DNA strands (DNA ^{α} /DNA ^{β}) were included (Figure 1b). To assess the binding of **sTAV2b** to these three duplex structures, isothermal titration calorimetry (ITC) was performed revealing the expected high affinity of **sTAV2b** for RNA ^{α} /RNA ^{β} ($K_d = 89$ nM, Supporting Information Table S4 and Figure S1). As anticipated, we did not observe binding of **sTAV2b** to double-

stranded DNA (DNA ^{α} /DNA ^{β} , Supporting Information Figure S2). The mixed duplex (RNA ^{α} /DNA ^{β}), however, was bound by **sTAV2b** ($K_d = 599$ nM, Figure 1c) though with ca. 10-fold lower affinity than duplex RNA.

Double-stranded RNA usually assembles into an A-form duplex while DNA tends to adopt the B-form.^[24] Duplexes composed of an RNA and a DNA strand, however, often fold into an A-form duplex.^[24] For the three nucleic acid duplexes in this study (Figure 1b), we performed circular dichroism (CD) spectroscopy indeed revealing characteristic A-form spectra ($\lambda_{\min} = 210$ nm, $\lambda_{\max} = 267$ nm,) for double-stranded RNA (RNA ^{α} /RNA ^{β} , Supporting Information Figures S4) and the mixed duplex (RNA ^{α} /DNA ^{β} , blue line, Figure 1d).^[24] The DNA duplex (DNA ^{α} /DNA ^{β}) showed the expected B-form pattern (Supporting Information Figures S4). We also used CD to investigate **sTAV2b** as well as its complex with RNA ^{α} /DNA ^{β} . While the spectrum of unbound **sTAV2b** reveals a typical random coil profile (orange line, Figure 1d), its complex with the mixed duplex, led to a distinct increase in ellipticity in the region between $\lambda = 208$ and 222 nm (black line, Figure 1d). This effect is not accounted for by the simple addition of the individual **sTAV2b** and RNA ^{α} /DNA ^{β} spectra (dashed line, Figure 1d) but rather suggests that **sTAV2b** adopts an α -helical conformation upon binding. Notably, the analogous behaviour was observed for **sTAV2b** binding to duplex RNA and is in line with previous results (Supporting Information Figure S5).^[23] Taken together, double-stranded RNA (RNA ^{α} /RNA ^{β}) and the mixed duplex (RNA ^{α} /DNA ^{β}) both adopt an A-form structure, and induce increased α -helicity in **sTAV2b** upon binding. This can explain why we observe binding of the mixed duplex but not of double-stranded DNA and suggests a similar complex structure of **sTAV2b** bound to RNA ^{α} /RNA ^{β} as well as RNA ^{α} /DNA ^{β} .

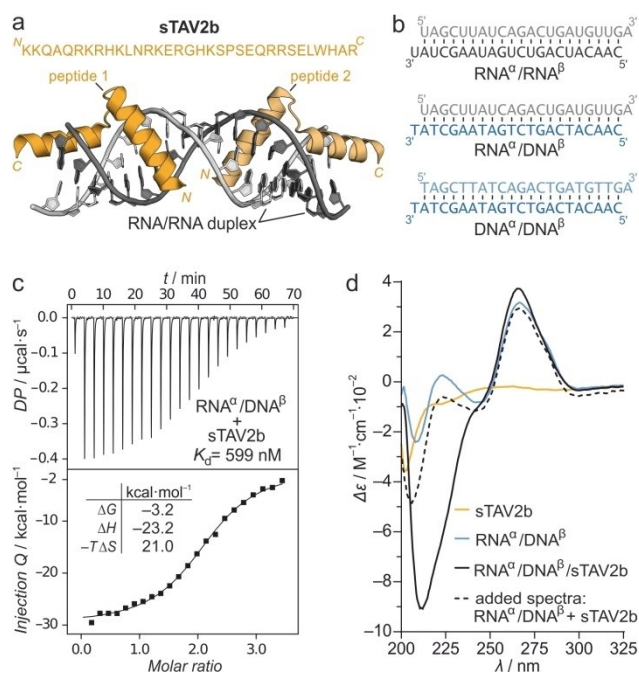


Figure 1. (a) Model of two identical TAV2b fragments (aa 21–53, **sTAV2b**, orange) bound to duplex RNA (grey). Model was derived from crystal structure PDB ID: 2Z10. (b) Oligonucleotide sequences used to test duplex binding (RNA: gray, DNA: blue). (c) Representative ITC plots of complex formation between **sTAV2b** ($c = 108$ μ M) and RNA ^{α} /DNA ^{β} ($c(\text{duplex}) = 6$ μ M). Measurements were performed in triplicate, errors = 1 σ ; for full data see Supporting Information Table S4 and Figure S4. Buffer: Phosphate-buffered saline (pH 7.4) (d) CD spectra of **sTAV2b** ($c = 4$ μ M; orange line), RNA ^{α} /DNA ^{β} ($c(\text{duplex}) = 2$ μ M; blue line), spectra of RNA ^{α} /DNA ^{β} ($c(\text{duplex}) = 2$ μ M) with **sTAV2b** ($c = 4$ μ M) (black line) and the sum of the two individual spectra (dotted black line). Buffer: 10 mM sodium phosphate (pH 7.4), 100 mM NaCl.

Peptide-DNA conjugation promotes RNA binding

Binding of peptide **sTAV2b** to the mixed duplex RNA ^{α} /DNA ^{β} provides a unique supramolecular assembly. We envisaged to combine the sequence-specific binding capability of DNA with the major groove recognition conveyed by **sTAV2b**, for the design of peptide-DNA conjugates capable of binding single-stranded RNA. For conjugate design, we considered the length of the DNA and the peptide component, as well as the position and nature of their linkage. We decided to use 12-mer DNA as the starting point since we expected this length to form a defined duplex with RNA. The TAV2b/duplex structure (Figure 1a) suggests two potential peptide-DNA arrangements (**A** and **B**), where DNA **A12** targets the 5'- and DNA **B12** the 3'-region of RNA ^{α} . We aimed to minimize the length of the peptides used in these two arrangements while maintaining the number of peptide contacts to the target RNA (for details see Supporting Methods). Based on these considerations, we chose a 28-mer TAV2b sequence (1, TAV2b₂₆₋₅₃) for arrangement **A** (left, Figure 2a, Supporting Information Figure S6), and a 21-mer sequence (2, TAV2b₂₆₋₄₇) for arrangement **B** (right, Figure 2a, Supporting Information Figure S7).

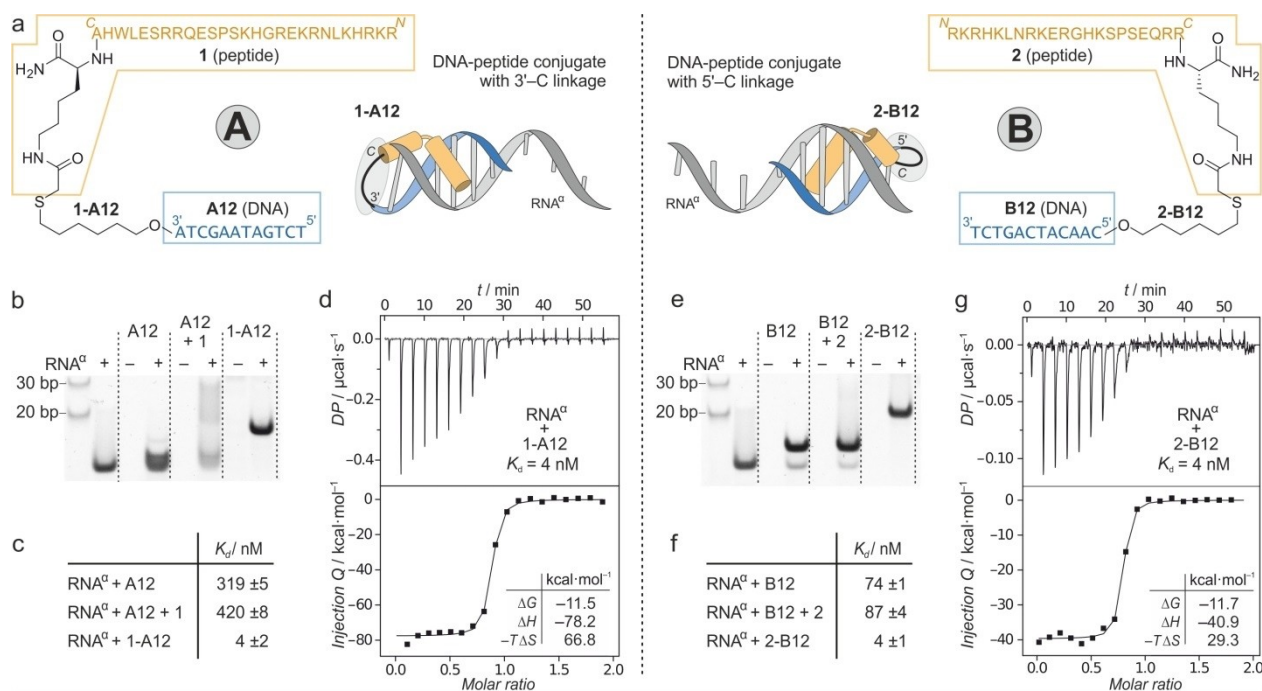


Figure 2. (a) Schematic overview of the design of conjugate type **A** and **B**: **sTAV2b** (orange) bound to duplex RNA/DNA (grey/blue) serves as starting point for the design of peptide-DNA conjugates (**1-A12** and **2-B12**) targeting single stranded RNA^α (grey). Chemical structure of linkages between peptide and DNA are shown. (b) EMSA of RNA^α in the presence and absence of DNA **A12**, a mixture of DNA and peptide (**A12 + 1**) and conjugate **1-A12**. Experiments employed 15% native polyacrylamide gel electrophoresis (PAGE) ($c(\text{RNA}) = 3 \mu\text{M}$, $c(\text{ligand}) = 4.5 \mu\text{M}$. Running buffer: 1x TAE, stain: SYBR gold (full gel in Supporting Information Figure S8). (c) K_d -values determined by ITC for RNA^α-ligand interactions (triplicate measurements, errors = 1 σ ; for full data see Supporting Information Table S4 and Figures S9–11). (d) Representative ITC of **1-A12** with the RNA^α (for full data see Supporting Information Figure S11). (e) EMSA of RNA^α in the presence and absence of DNA **B12**, a mixture of DNA and peptide (**B12 + 2**) and conjugate **2-B12** (for details see caption Figure 2b, full gel in Supporting Information Figure S12). (f) K_d -values determined by ITC for RNA^α/ligand interactions (triplicate measurements, errors = 1 σ ; for full data see Supporting Information Table S4 and Figures S13–15). (g) Representative ITC of **2-B12** with the RNA^α (for all data see Supporting Information Figure S15).

In arrangement **A** (left, Figure 2a), we expect the C-terminus of peptide **1** to be in proximity to the 3'-end of DNA **A12**. We envisioned a flexible linker structure to facilitate a 3'-to-C-terminus linkage. Preferably, this linker would be assembled via a biocompatible reaction using appropriately modified peptide and DNA molecules. For that purpose, peptide **1** was synthesized with a C-terminal lysine bearing a bromoacetamide-functionalized side chain (Supporting Information Figure S16). DNA **A12**, on the other hand, was obtained with a 3'-thiol modification facilitating the formation of the thioether linkage in conjugate **1-A12** (left, Figure 2a, Supporting Information Figure S17). In arrangement **B** (right, Figure 2a), we expect the C-terminus of peptide **2** to be in proximity to the 5'-end of DNA **B12**. Aiming for a similar linker geometry as for arrangement **A**, an analogous synthesis was applied however using a 5' thiol-modified DNA **B12** analogue (right, Figure 2a).

Initially, binding of conjugate **1-A12** to RNA^α was analysed using electrophoretic mobility shift assays (EMSA) applying non-denaturing conditions that allow the detection of supramolecular complexes. In addition, the binding of DNA **A12** in the presence and absence of peptide **1** was investigated (Figure 2b). Incubation of RNA^α with **A12** led to the appearance of a new slightly elevated band, indicative of RNA^α/**A12** duplex formation. Adding peptide **1** to this solution (**A12 + 1**) caused the disappearance of discrete bands (Figure 2b), suggesting

non-specific binding of **1** to the RNA^α/**A12** duplex. In contrast, incubation of RNA^α with conjugate **1-A12** resulted in the formation of a defined high molecular weight band indicating the desired complex formation. Most notably, this goes in hand with a complete disappearance of the band for free RNA^α (Figure 2b) suggesting that conjugate **1-A12** forms a higher affinity complex with RNA^α than DNA **A12** alone. We next performed ITC experiments to quantify binding, revealing moderate affinities for DNA **A12** both in the presence and absence of peptide **1** ($K_d = 319 \text{ nM}$ and $K_d = 420 \text{ nM}$, respectively; Figure 2c). Conjugate **1-A12**, on the other hand shows high affinity for RNA^α ($K_d = 4 \text{ nM}$) corresponding to an 80-fold increase when compared to DNA **A12** (Figure 2c).

Analogous binding experiments were performed for conjugate **2-B12**, DNA **B12** and peptide **2**. EMSA experiments revealed a behaviour similar to the **A**-family. Though, a notable difference is the appearance of a discrete band for the RNA^α/**B12** complex in the presence of peptide **2** (Figure 2e). This indicates a lower tendency of the unconjugated peptide to interact with the RNA/DNA duplex. ITC experiments revealed an overall higher affinity of DNA **B12** to RNA^α ($K_d = 74 \text{ nM}$, Figure 2f) than of DNA **A12** ($K_d = 319 \text{ nM}$) which can be explained by the higher GC-content in the resulting **B12**/RNA^α duplex. Nevertheless both conjugates show equal affinity for RNA^α (for both, $K_d = 4 \text{ nM}$, Figures 2d and 2g). This suggests a more

pronounced synergistic effect from peptide-DNA conjugation in 1-A12 than in 2-B12 (80- vs. 19-fold increased affinity compared to DNA only). The large increase in affinity for conjugation of 1 and A12 is accompanied by a 2.2-fold larger enthalpic contribution to binding (A12 vs. 1-A12, $\Delta H = -34$ vs. -78 kcal mol⁻¹, Supporting Information Table S4). Notably, this is not observed for the conjugation of 2 and B12 (B12 vs. 2-B12, $\Delta H = -37$ vs. -41 kcal mol⁻¹). The large enthalpic contribution to binding for 1-A12, however, is partially compensated by an increased entropic penalty (A12 vs. 1-A12, $-T\Delta S = 26$ vs. 67 kcal mol⁻¹, Supporting Table S4). Comparing both conjugates, the greater enthalpy of binding observed for 1-A12 can arise from the C-terminally elongated sequence of peptide 1 which can be expected to form additional contacts with the DNA-RNA duplex.

A- and B-series conjugates can bind simultaneously

Nucleic acid duplexes are typically analyzed regarding their melting temperature (T_m), characterizing the transition from the folded duplex to the corresponding single strands.^[25] We therefore performed thermal denaturation experiments using the CD signal at $\lambda = 267$ nm as the readout. This wavelength is indicative of the presence of A-form duplexes and thereby facilitates the measurement of T_m -values.^[24] For the duplexes ($c = 2$ μ M) composed of RNA^q and DNA A12 or B12, we determined T_m -values in the anticipated range. In analogy to the ITC measurements, RNA^q/A12 shows a lower thermal stability ($T_m = 36$ °C, Figure 3a) than RNA^q/B12 ($T_m = 44$ °C, Figure 3b). For both mixed duplexes, the addition of the corresponding peptide did not result in meaningful changes. However when comparing the RNA-bound conjugates (1-A12 and 2-B12) to the corresponding RNA^q/DNA duplexes, we observed considerably increased T_m -values ($\Delta T_m = 11$ and 9 °C, A- and B-family, Figure 3a and 3b). When compared to melting temperatures calculated by nearest-neighbor methods,^[26] such increases are equivalent to the addition of 3–6 nt respectively (Supporting Information Table S5).

Next, we were interested to assess how the length of the DNA sequence affects conjugate binding to RNA^q. For this reason, 10 and 11 nucleotide DNA strands were conjugated to peptides 1 and 2, respectively, providing an A- and B-series of conjugates (1-AX and 2-BX, X = number of nucleotides in DNA, Figure 3c and 3d). Subsequently, thermal denaturation experiments with RNA^q were performed for the new conjugates and the corresponding DNA sequences (AX and BX, X = number of nucleotides) in the presence and absence of peptide. Overall, decreasing DNA length was associated with reduced T_m -values both for the unmodified DNA and the conjugates (Figure 3c and 3d). In particular, the removal of the cytosine (C) and the corresponding GC base pair affected the thermal stability of the DNA and the conjugates in both series. Nevertheless in all cases, peptide conjugation improved stability with 1-A11 exhibiting the highest and 2-B11 lowest stabilizing effect upon conjugation ($\Delta T_m = 12$ °C and 5 °C, respectively).

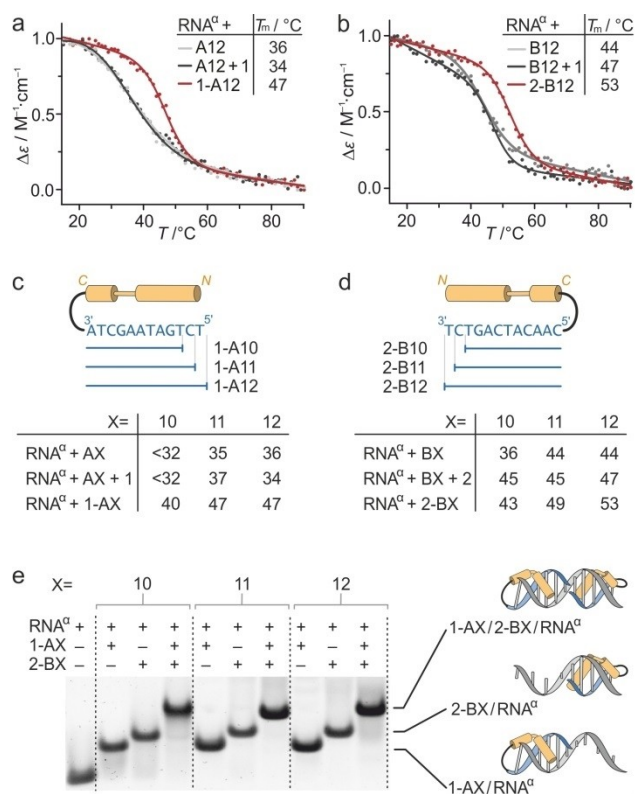


Figure 3. (a) Melting temperature profiles of RNA^q in the presence of A12, an equimolar mixture of A12 + 1, and 1-A12 respectively. (b) Melting temperature profiles of RNA^q in the presence of B12, an equimolar mixture of B12 + 2, and 2-B12 respectively ($\lambda = 267$ nm, $c(\text{RNA}^q) = 2$ μ M, $c(\text{binding partners}) = 2$ μ M. Buffer: 10 mM sodium phosphate, 100 mM NaCl (pH 7.4)). (c) Sequences of DNA-truncated A-family conjugates. Table of T_m -values of conjugates with RNA^q (for melting curves see Supporting Information Figure S18). (d) Sequences of DNA-truncated B-family conjugates. Table of T_m -values of conjugates with RNA^q (for melting curves see Supporting Information Figure S18). (e) EMSA of RNA^q in the presence of different conjugates (Supporting Information Figure S19) including co-incubation of A- and B-series members with DNA sequences of equal lengths. Experiments employed 15 % native PAGE ($c(\text{RNA}^q) = 3$ μ M, $c(\text{conjugate}) = 4.5$ μ M. Running buffer: TAE, stain: SYBR gold. Cartoon representations of the proposed RNA/conjugate complexes corresponding to band species are presented on the right hand side.

We then used EMSA experiments to verify the presence of complexes between RNA^q and conjugates. Indeed, incubation of RNA^q with any of the six conjugates resulted in elevated bands which indicated the desired complex formation (Figure 3e). Knowing that the conjugates of the A-series target the 5'-region of RNA^q while the B-series targets the 3'-region, we were interested to assess the possibility for their simultaneous binding. For that reason, conjugates with equal DNA length were co-incubated with RNA^q and included in the EMSA analysis. For all combinations, we observe a band further elevated than incubation with the individual conjugates (Figure 3e). This clearly supports simultaneous binding in those three combinations. In this respect, it is noteworthy that the complementary regions of the DNA in 1-A11 and 2-B11 overlap by one nucleotide and in 1-A12 and 2-B12 by three nucleotides (Supporting Information Figure S20). In these cases, we expect

the corresponding DNA nucleotides to be to some extent unpaired.

Simultaneous conjugate binding accelerates crosslinking

Having confirmed the simultaneous binding capacities of both conjugate families, we next sought to utilize the unique structural features of this supramolecular assembly. In particular, the relative orientation of the peptide *N*-termini caught our attention. While the peptides are connected to the outwards facing DNA termini, we expect that major groove binding would orient their *N*-termini inwards. We considered to use the resulting proximity of both termini to enable an RNA^α-templated linkage of both conjugates (Figure 4a). For that purpose, we chose the bio-orthogonal strain-promoted azide-alkyne cycloaddition.^[27] Inspired by the two-component click reaction used for peptide stapling,^[28] each of the two peptide *N*-termini was equipped with an azide group. For crosslinking, bis-alkyne 5 (Supporting Information Figure S21) was used in which two dibenzocyclooctyne moieties are connected via a polyethylene glycol spacer. To implement the templated reaction, the peptides in conjugates 1-A12 and 2-B12 were *N*-terminally modified with an azide group, yielding conjugates 3 and 4, respectively (Supporting Information Figure S22). The desired reaction of 3 and 4 with bis-alkyne 5 would then result in heterodimeric product 3*4 (Figure 4a and 4b).

To assess the strain-promoted click reaction, product formation was determined using high-performance liquid chromatography (HPLC) as the readout. At the beginning of the reaction ($t=0$, Figure 4c), we observed two peaks corresponding to conjugates 3 and 4. When incubating the two conjugates (each $c=5\ \mu\text{M}$) for 6 h with RNA^α ($c=5\ \mu\text{M}$) and bis-alkyne 5 ($c=5.75\ \mu\text{M}$), we observe decreasing signals for 3 and 4 and the appearance of a new peak (red line, Figure 4c). Isolation of this new species and its characterization via electrospray ionization mass spectrometry (ESI-MS) provided a spectrum that after deconvolution indicates the presence of a molecule with a molecular weight of $15187\ \text{g mol}^{-1}$ (Figure 4d). This is in agreement with the molecular weight expected for the desired heterodimeric product 3*4 ($M=15186\ \text{g mol}^{-1}$, Figure 4b). Based on the peak areas in the chromatogram, 67% product 3*4 was formed (Figure 4c).

To evaluate the actual impact of RNA^α-mediated assembly on the click reaction, conjugates 3 and 4 were also reacted with bis-alkyne 5 in absence of the RNA^α (no RNA, Figure 4c). After 6 h, a number of new peaks were observed which could be assigned to the previously obtained product 3*4 (10%) as well as the homo-dimer 3*3 (5%, Supporting Information Figure S24). We also detected additional signals which we associated with homodimer 4*4 (5%) as well as the mono-reacted conjugates 3* (12%) and 4* (18%) originating from the reaction of one of the two alkyne moieties in 5 with the azide in 3 or 4 (Supporting Information Figure S25). As an additional control experiment, we reacted 3, 4 and bis-alkyne 5 in the presence of non-complementary RNA^β (grey line, Figure 4c), which resulted in a chromatogram very similar to the reaction

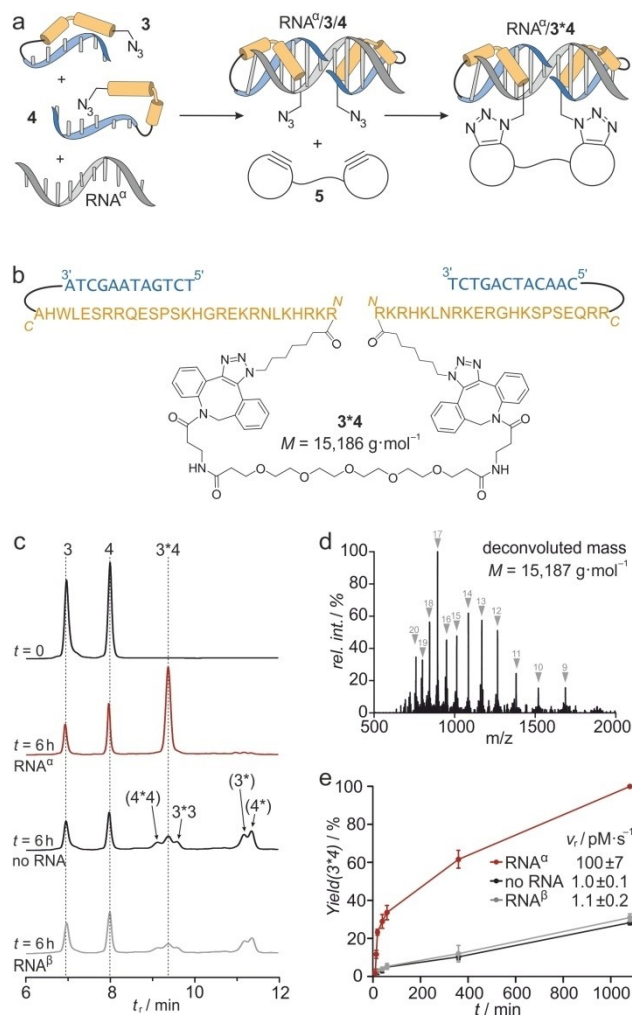


Figure 4. (a) Schematic overview of RNA^α-templated reaction. Azide-modified conjugates 3 and 4 bind RNA^α to form a ternary complex which reacts with a bis-alkyne crosslinker to yield a complex between ligated conjugate 3*4 and RNA^α. (b) Structure of heterodimeric product 3*4 including the chemical structure of the reacted crosslink. (c) HPLC traces of unreacted 3 and 4 ($t=0$) and of reactions between 3, 4 and 5 after 6 h in the presence of RNA^α (red line), in the absence of RNA (no RNA), and in the presence of RNA^β (gray). (d) Electrospray ionization (ESI) mass spectrum of heterodimeric product 3*4 (full characterization see Supporting Information Figure S23). (e) Time-dependent yields of 3*4 as determined by HPLC including initial rates (v_i). Conjugates ($c=5\ \mu\text{M}$) were incubated for 18 h at room temperature in the presence or absence of an RNA template ($c=5\ \mu\text{M}$) and bis-alkyne crosslinker 5 ($c=5.75\ \mu\text{M}$). Reaction buffer: 100 mM NH_4HCO_3 , 20% acetonitrile, 1% DMSO (pH=8.0). Templated reactions were performed in triplicate, errors = 1σ; for full data see Supporting Information Table S6 and Figure S27.

in absence of RNA. This can be explained by the absence of conjugate binding to RNA^β which was also confirmed using EMSA experiments (Supporting Information Figure S26). Next, we determined the time-dependent formation of product 3*4 in the three discussed cases (RNA^α, no RNA, RNA^β). Again using HPLC as the readout, time courses were obtained in triplicates (Figure 4e). Notably in the presence of RNA^α and after 18 h, the formation of 96% click product 3*4 was observed, while under the same conditions only about 30% product is detected without RNA or in the presence of RNA^β (Figure 4e). These time-

dependent measurements allowed the determination of initial rates (v_r) revealing a two orders of magnitude faster reaction in the presence ($v_r = 100 \text{ pMs}^{-1}$, Figure 4e) than in the absence of RNA $^\alpha$ ($v_r = 1.0 \text{ pMs}^{-1}$). Incubation with control RNA $^\beta$ provides an initial rate ($v_r = 1.1 \text{ pMs}^{-1}$) comparable to reactions in the absence of RNA. Overall, the observed proximity-induced reaction clearly supports the anticipated structural preorganization in ternary complex RNA $^\alpha$ /3*4 (Figure 4a).

Conclusions

In this study, we use a protein/oligonucleotide complex as the basis for the design of synthetic peptide-DNA conjugates with tunable assembly characteristics. Initially, we observed that the RNA-binding motif of TAV2b recognizes double-stranded RNA as well as RNA/DNA hybrid duplexes. Pairing this insight with a previously reported TAV2b/RNA crystal structure,^[22] we designed two distinct families of peptide-DNA conjugates, which form a complex with a single-stranded RNA target. RNA binding is enabled via a combination of Watson-Crick base-pairing and peptide-mediated duplex recognition, resulting in conjugates with increased RNA affinity when compared to the isolated DNA and peptide components. The RNA affinity of these conjugates can be tuned by the variation of the DNA length. We also showed that two conjugates (one from each family) can bind simultaneously to the target RNA.

Given the structural basis of our design, we rationalized that simultaneous RNA binding would bring the peptide N-termini of both conjugates into proximity and could therefore allow a templated reaction of both termini. Employing a strain-promoted double click reaction, we found that reactions carried out in the absence of RNA or in the presence of a non-complementary sequence generated mixtures of different low-yielding click products. Incubation with template RNA $^\alpha$, leads to the exclusive formation of the hetero-dimeric product 3*4. This proximity-induced reaction highlights the potential of structure-based design in the development of such modular assemblies, particularly given the growing number of available protein/oligonucleotide complex structures.^[29,30] Moreover, this expands both the complexity and functional spectrum of recent efforts towards chimeric molecules utilizing peptide and oligonucleotide features.^[31–36] Taken together, our work serves as a proof-of-concept for the structure-based design of peptide-DNA conjugates possessing the ability to assemble into structurally-defined complexes. Such systems can open the door to novel functional biomolecular assemblies.

Acknowledgements

This work was supported by the European Research Council (ERC starting grant number 678623), by AstraZeneca, Bayer CropScience, Bayer HealthCare, Boehringer Ingelheim, Merck KGaA, and the Max Planck Society. We thank the group of Prof. Iwan de Esch (VU Amsterdam) for access and assistance with the isothermal titration calorimeter, Prof. Govert Somsen, Ben

Bruyneel and Dr. Erika Amstalden van Hove (VU Amsterdam) for access and assistance with the MALDI-TOF spectrometer, and we thank the group of Prof. Martine Smit (VU Amsterdam) for access to gel imaging. We also thank Dr. Carolin Mueller for useful discussions and Dr. Alan Gerber for critical reading of this manuscript.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: bio-conjugation · peptides · RNA recognition · self-assembly · strain-promoted click reaction

- [1] J. Wilce, J. Vivian, M. Wilce, in *Protein Dimerization and Oligomerization in Biology* (Ed. J. M. Matthews), Springer, New York, 2012, pp. 91–104.
- [2] M. Cassandri, A. Smirnov, F. Novelli, C. Pitolli, M. Agostini, M. Malewicz, G. Melino, G. Raschella, *Cell Death Dis.* 2017, 3, 17071.
- [3] S. Jones, *Genome Biol.* 2004, 5, 226.
- [4] D. Zhao, Y. Kong, S. Zhao, H. Xing, *Top. Curr. Chem.* 2020, 378, 41.
- [5] I. Dovgan, O. Koniev, S. Kolodych, A. Wagner, *Bioconjugate Chem.* 2019, 30, 2483–2501.
- [6] T. MacCulloch, A. Buchberger, N. Stephanopoulos, *Org. Biomol. Chem.* 2019, 17, 1668–1682.
- [7] C. M. Troy, D. Derossi, A. Prochiantz, L. A. Greene, M. L. Shelanski, *J. Neurosci.* 1996, 16, 253–261.
- [8] Y.-L. Chiu, A. Ali, C. Chu, H. Cao, T. M. Rana, *Chem. Biol.* 2004, 11, 1165–1175.
- [9] M. C. Needels, D. G. Jones, E. H. Tate, G. L. Heinkel, L. M. Kochersperger, W. J. Dower, R. W. Barrett, M. A. Gallop, *Proc. Natl. Acad. Sci. USA* 1993, 90, 10700–10704.
- [10] Y. Li, R. De Luca, S. Cazzamalli, F. Pretto, D. Bajic, J. Scheuermann, D. Neri, *Nat. Chem.* 2018, 10, 441–448.
- [11] L. Röglin, F. Altenbrunn, O. Seitz, *ChemBioChem* 2009, 10, 758–765.
- [12] S. Thurley, L. Röglin, O. Seitz, *J. Am. Chem. Soc.* 2007, 129, 12693–12695.
- [13] M. Marczyneke, K. Gröger, O. Seitz, *Bioconjugate Chem.* 2017, 28, 2384–2392.
- [14] B. M. G. Janssen, E. H. M. Lempens, L. L. C. Olijve, I. K. Voets, J. L. J. Van Dongen, T. F. A. De Greef, M. Merckx, *Chem. Sci.* 2013, 4, 1442–1450.
- [15] C. Lou, M. C. Martos-Maldonado, C. S. Madsen, R. P. Thomsen, S. R. Midtgaard, N. J. Christensen, J. Kjems, P. W. Thulstrup, J. Wengel, K. J. Jensen, *Nat. Commun.* 2016, 7, 1–9.
- [16] E. Spruijt, S. E. Tusk, H. Bayley, *Nat. Nanotechnol.* 2018, 13, 739–745.
- [17] J. Jin, E. G. Baker, C. W. Wood, J. Bath, D. N. Woolfson, A. J. Turberfield, *ACS Nano* 2019, 13, 9927–9935.
- [18] K. Gorska, N. Winssinger, *Angew. Chem. Int. Ed.* 2013, 52, 6820–6843; *Angew. Chem.* 2013, 125, 6956–6980.
- [19] O. Seitz, *J. Pept. Sci.* 2019, 25, 1–19.
- [20] M. Ejlersen, N. J. Christensen, K. K. Sørensen, K. J. Jensen, J. Wengel, C. Lou, *Bioconjugate Chem.* 2018, 29, 1025–1029.
- [21] E. Pazos, C. Portela, C. Penas, M. E. Vázquez, J. L. Mascareñas, *Org. Biomol. Chem.* 2015, 13, 5385–5390.
- [22] H.-Y. Chen, J. Yang, C. Lin, Y. A. Yuan, *EMBO Rep.* 2008, 9, 754–60.
- [23] A. Kuepper, N. McLoughlin, S. Neubacher, E. Collado-Camps, N. Chandran, S. Mukherjee, L. Bethge, J. Bujnicki, R. Brock, S. Heinrichs, T. Grossmann, *ChemRxiv*. 2020, <https://doi.org/10.26434/chemrxiv.13325264>.
- [24] J. Kypr, I. Kejnovska, D. Renciuik, M. Vorlickova, *Nucleic Acids Res.* 2009, 37, 1713–1725.
- [25] G. R. Bishop, J. B. Chaires, *Curr. Protoc. Nucleic Acid Chem.*, John Wiley & Sons, Inc., Hoboken, 2002, pp. 7.11.1–7.11.8.
- [26] W. A. Kibbe, *Nucl. Acids Res.* 2007, 35, W43–W46.
- [27] J. C. Jewett, E. M. Sletten, C. R. Bertozzi, *J. Am. Chem. Soc.* 2010, 132, 3688–3690.
- [28] Y. H. Lau, Y. Wu, M. Rossmann, B. X. Tan, P. de Andrade, Y. S. Tan, C. Verma, G. J. McKenzie, A. R. Venkitaraman, M. Hyvönen, D. R. Spring, *Angew. Chem. Int. Ed.* 2015, 54, 15410–15413; *Angew. Chem.* 2015, 127, 15630–15633.

- [29] D. M. Krüger, S. Neubacher, T. N. Grossmann, *RNA* **2018**, *24*, 1457–1465.
- [30] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucl. Acids Res.* **2000**, *28*, 235–242.
- [31] Z. Chen, P. A. Lichtor, A. P. Berliner, J. C. Chen, D. R. Liu, *Nat. Chem.* **2018**, *10*, 420–427.
- [32] A. Ghidini, A. Cléry, F. Halloy, F. H. T. Allain, J. Hall, *Angew. Chem. Int. Ed.* **2021**, *60*, 3163–3169; *Angew. Chem.* **2021**, *133*, 3200–3206.
- [33] S. Pomplun, Z. P. Gates, G. Zhang, A. J. Quartararo, B. L. Pentelute, *J. Am. Chem. Soc.* **2020**, *142*, 19642–19651.
- [34] C. S. Swenson, A. Velusamy, H. S. Argueta-Gonzalez, J. M. Heemstra, *J. Am. Chem. Soc.* **2019**, *141*, 19038–19047.
- [35] W. Engelen, K. Zhu, N. Subedi, A. Idili, F. Ricci, J. Tel, M. Merkx, *ACS Cent. Sci.* **2020**, *6*, 22–31.
- [36] S. Angerani, N. Winssinger, *J. Am. Chem. Soc.* **2020**, *142*, 12333–12340.

Manuscript received: March 26, 2021

Accepted manuscript online: April 29, 2021

Version of record online: June 1, 2021