An artificial split enzyme for small molecule recognition

In this chapter an artificial allosteric enzyme is described based on the split protein murine dihydrofolate reductase (mDHFR). In this split enzyme mDHFR fragments were first conjugated to short oligonucleotides and cyclodextrin moieties were attached to oligonucleotides complementary to those appended to the protein fragments. Then, DNA hybridization and supramolecular interactions between a guest molecule and the two cyclodextrins were combined to create an allosteric split enzyme for small molecule recognition. The synthesis of cyclodextrin-DNA conjugates is described and two strategies for the preparation of mDHFR fragments-DNA conjugates were explored, involving maleimide coupling and Cu(I)-catalyzed 1,3-dipolar cycloaddition. Preliminary studies on the reassembly of the split-enzyme are also described.
2.1 Introduction

2.1.1 Split proteins
Split proteins biosensor assays (SBAs), also referred to as protein complementation assays (PCAs), have emerged as powerful tools to study protein-protein interactions. Split proteins are polypeptides that can be dissected into two fragments at the genetic level and whose reassembly into a functional protein is mediated by secondary interactions.1–3 Protein fragments that make a successful split protein reporter have two critical features: (1) they do not spontaneously reassemble into an active conformation in absence of a binding partner and (2) they do not show any activity independently of each other. Furthermore, the reassembly into an active species has to be driven by additional interacting moieties and an easily detectable readout (fluorescence or catalytic activity) should be available upon reassembly.2

The first example of a conditional split protein reassembly was reported by Johnsson and Varshavsky in which two fragments of ubiquitin were attached to interacting proteins pairs (leucine zippers). The secondary interactions provided by the appended polypeptides mediated the reassembly of the ubiquitin fragments into an active conformation (Figure 1a).4

![Figure 1](image_url)

**Figure 1**: (a) Reassembly of split proteins mediated by leucine zippers.4–10 (b) An autoinhibited strategy for the detection of small molecules using supramolecular building blocks: a GFP fragment is conjugated with a receptor moiety (β-cyclodextrin) and a small molecule (coumarin) to form an intramolecular inactive inclusion complex. Addition of an exogenous guest molecule (1-adamantanol) displaces the coumarin, unfolds the peptide and allows the reassembly of the two fragments.11 (c) Split protein fragments fused to FKBP12 (FK506 binding protein) and FRB (FKBP12-rapamycin-binding domain of FRAP). Addition of rapamycin induces protein protein dimerization resulting in the reconstitution of the active protein.12,7,13–15 (d) Q8-activated split-luciferase: split-luciferase fragments containing an N-terminal FGG sequence reassemble upon binding to cucurbit[8]uril (Q8).16
Since then several examples of split proteins have been reported including GFP variants, luciferases, dihydrofolate reductase, β-lactamase, Tobacco Etch Virus (TEV) protease, inteins and chorismate mutase. The reassembly of two inactive fragments can be mediated by protein-protein interactions, protein-oligonucleotide interactions, small molecules and supramolecular interactions or DNA hybridization (Figure 1).

Besides being used to study protein-protein interactions, split protein reporters have proven to be a valuable tool for the study interactions of proteins with non-proteinogenic targets such as small molecules, DNA or RNA.

2.1.2 DNA-templated reassembly of split enzymes
Nucleic acids have proven to be a versatile molecular building block in nanotechnology and to have important biomedical applications, thanks to the programmability of their assemblies and their natural bio-compatibility. DNA, for example, has been used to create sophisticated two- and three-dimensional structures and as a structural building block for modular assembly. Oligonucleotides can, indeed, function as templates to build well-defined structures and several methods have been developed to control protein activity with oligonucleotides. DNA can template the co-localization of protein domains or fragments in order to promote the assembly of protein complexes or increase the efficiency of catalytic reactions. To tether two proteins (protein fragments or domains) on a DNA template, the synthesis of protein-oligonucleotide conjugates with a complementary strand can be used. Another possibility is the fusion of the proteins of interest with DNA-binding domains.

The group of Ghosh was the first to introduce DNA-templated reassembly of split enzymes with a methodology called Sequence Enabled Enzyme Reassembly (SEER). In this approach protein fragments are expressed as fusion proteins with a DNA binding domain, such as zinc finger proteins (Figure 2a). Upon addition of a DNA duplex containing recognition sequences for both DNA binding proteins, the split protein fragments co-localize on the DNA strand in close proximity to each other, allowing the functional reassembly of the enzyme. This approach has been applied to create split protein versions of GFP, β-lactamase and luciferases. Exchanging one zinc finger for a binding domain that recognizes methylated CpG sites, SEER was also employed for the detection of methylated DNA as a potential cancer biosensor. The system was further developed to increase its modularity utilizing DNA hairpin and single-stranded overhangs for the reassembly of a split luciferase. A slight modification of the SEER methodology was also applied by Takeda et al. (Figure 2b). In this work, only one fragment of a split luciferase was fused with a zinc finger protein, while the other was coupled to a
single-stranded oligonucleotide. Upon hybridization with the complementary strand, followed by incubation of the two fragments, reassembly of the luciferase was obtained.

**Figure 2**: Examples of DNA-mediated split protein reassembly. (a) and (b) SEER based on interactions between a DNA binding domain fused to protein fragments.\(^{19,20,35–37,27}\) (c) Reassembly of split proteins mediated by DNA-hybridization.\(^{21,28}\) (d) Modular design of a DNA-hybridization mediated reassembly of a split enzyme.\(^{34}\)

One advantage of using sequence-specific DNA-binding domains, such as zinc fingers, is that these fusion proteins can be obtained by recombinant expression, thereby obviating the need to prepare protein-DNA conjugates. However, this methodology is only applicable to DNA targets for which a peptide-binding domain can be engineered. In contrast, mediating protein reassembly directly through DNA hybridization could, in principle, be applied to any nucleic acid sequence.

The DNA-triggered reassembly of a split green fluorescent protein (EGFP) by Demidov et al. (Figure 2c) was the first example for which DNA hybridization was employed to reassemble a split protein.\(^{21}\) In this work each fragment of EGFP was expressed with a terminal cysteine, biotinylated and incubated with streptavidin. Two complementary oligonucleotides were biotinylated as well, on the 5’ and 3’ respectively. Upon incubation of the split EGFP-streptavidin conjugates with the biotinylated oligonucleotides, non-covalent EGFP fragment-DNA conjugates were obtained and hybridization of the two complementary strands mediated the reassembly of the split EGFP fragments, restoring fluorescence of the protein. A similar approach was also applied by Cissell et al. in which the reassembly of split luciferase fragments was facilitated by covalently linking complementary oligonucleotides to the protein fragments.\(^{28}\)

In order to increase the modularity of the DNA-templated split proteins and allow these constructs to respond to external DNA strands, our group has
previously developed an allosteric split murine dihydrofolate reductase (split mDHFR, Figure 2d). Toward this end, full length mDHFR was first split into two fragments (corresponding to positions 1-105 and 106-186) which were subsequently conjugated to oligonucleotides by cysteine-maleimide coupling. Upon addition of a template DNA strand that is complementary to both oligonucleotidic handles, the protein fragments co-localized on the DNA and reassembled into a functional enzyme. As a result, the catalytic activity of this split mDHFR could be modulated by varying both the concentration and the sequence of the template DNA.

2.2 Aim

In this chapter, we aimed to create an artificial split enzyme that responds to small molecules. Based on the design of the DNA-directed split mDHFR enzyme described above, a synthetic receptor that allows for sensing of an analyte was appended onto two complementary oligonucleotidic sequences. DNA hybridization and supramolecular interactions were combined to create an allosteric split enzyme for small molecule recognition. The incorporation of the synthetic receptors in this design is variable and this strategy would allow for a modular assembly of one split enzyme that could respond to different small guest molecules by only changing the receptors of interest.

2.3 Design

The enzyme murine dihydrofolate reductase (mDHFR) was selected for the design of the envisioned allosteric split enzyme. mDHFR is a monomeric (21 KDa) protein that catalyzes the reduction of 7,8-dihydrofolate to 5,6,7,8-tetrahydrofolate using NADPH as a cofactor. As described above, mDHFR classifies as a split protein: when dissected at specific positions two inactive fragments are formed that can reassemble into an active conformation only in presence of secondary interactions. The structure-activity relationship of DHFR is well-studied and enzymatic activity of mDHFR can be monitored spectrophotometrically or in vivo by cell survival assay.

The design of the split allosteric enzyme for small molecule recognition is depicted in Figure 3 and is based on the DNA-directed split mDHFR enzyme described above (Figure 2d). As before, the polypeptide chain of mDHFR is cut between residues 105 and 106 and conjugated to short oligonucleotides. However, in the present design the template DNA used to reassemble the split mDHFR in the previous example is replaced with two shorter oligonucleotides that are complementary to the mDHFR-linked sequences. Each of these oligonucleotides is further functionalized with a receptor moiety. Hybridization of the
receptor-functionalized oligonucleotides and the ones covalently attached to the protein fragments links a receptor moiety with each mDHFR fragment. Ultimately, simultaneous binding of a guest molecule to both receptors would facilitate the reassembly of the fragments, yielding a catalytically active enzyme.

![Diagram](image)

**Figure 3:** Schematic representation of the mDHFR split enzyme for small molecule recognition using cyclodextrins as receptor moieties.

The most notable feature of this design is its modularity: employing the same protein-DNA conjugates it is possible to create a variety of artificial enzymes by simply changing the receptors linked to the oligonucleotides. No synthesis of a number protein-linked receptor is required, thus obviating the need to optimize reaction conditions and purification for each different variant. Once a library of oligonucleotide-linked receptors is available, one module is readily exchanged for another, giving rise to various artificial enzymes that each respond to the binding of a different molecule.

To explore the feasibility of such a combinatorial split enzyme system, cyclodextrins will be used as model receptor moieties. Cyclodextrins are cyclic oligosaccharides containing a hydrophobic core that enables binding of small, apolar molecules in aqueous environments through hydrophobic interactions. Their host/guest chemistry is well-established and these molecules are known to bind small ligands, such as bis(2-(1-adamantyl)ethyl phosphate, with high affinity forming a 2:1 host:guest complex, which is ideal for this design. Moreover, the synthetic methodology for mono-functionalization of cyclodextrins, including the conjugation to oligonucleotides, is well-developed.

**2.4 Results and discussion**

**2.4.1 Synthesis of cyclodextrin-DNA conjugates**

The synthesis of cyclodextrin-DNA conjugates has been reported in literature using different strategies: coupling of an amino-functionalized cyclodextrin to an activated acid on a DNA strand, coupling of thiolated cyclodextrins to amino-modified oligonucleotides or by employing Cu(I)-catalyzed 1,3-dipolar
azide-alkyne cycloaddition (CuAAC). The latter was selected as the strategy of choice due to the synthetic ease by which azide-functionalized cyclodextrins and alkyne-functionalized oligonucleotides are accessible.

The alkyne moiety was introduced by reaction of commercially available amino-functionalized oligonucleotides with an excess of the N-hydroxysuccinimide ester (NHS) of 5-pentynoic acid in 200 mM NaH₂PO₄ pH 7.2 overnight at 25 °C (Figure 4). Size exclusion chromatography was used to remove the excess of alkyne linker. This procedure gave rise to quantitative formation of alkyne-functionalized oligonucleotides as judged by rp-HPLC (reverse-phase HPLC) and MALDI-TOF (Table 1).

Figure 4: Synthesis of alkyne-functionalized oligonucleotides

Mono-6-deoxy-6-azido-β-cyclodextrin was synthesized in two steps according to literature procedures (Figure 5) and cyclodextrin-DNA conjugates were prepared from the 3’- and 5’-alkyne-modified oligonucleotides.

Figure 5: Synthesis of mono-6-deoxy-6-azido-β-cyclodextrin

To prepare the cyclodextrin-DNA conjugates (Figure 6a), mono-6-deoxy-6-azido-β-cyclodextrin was dissolved in DMSO and added to a solution of alkyne modified oligonucleotide in 200 mM NaH₂PO₄ pH 7.5 containing CuSO₄·5H₂O, THPTA (3,3’3’’-(4,4’,4’’-(nitrilotris(methylene))tris (1H-1,2,3-triazole-4,1-diyl))tris (propan-1-ol)) and sodium ascorbate. Denaturing PAGE (Figure 6b) and rp-HPLC showed full conversion towards the cyclodextrin-DNA conjugates after incubation at 40 °C for 4 h. Purification of the conjugates from the excess of cyclodextrin was achieved by strong anion exchange chromatography from 100 mM Tris pH 8.0 by elution with a high ionic strength buffer, followed by desalting of the resulting solution by size exclusion chromatography (Figure 7).
a)

\[
\text{Oligo11} \quad \text{GTCAGACATGTC - (CH}_2\text{)}_6\text{-NH-CO-(CH}_2\text{)}_2\text{-C}_3\text{N}_2\text{-βCD}
\]

\[
\text{Oligo12} \quad \text{βCD-C}_3\text{N}_2\text{-CO-NH-(CH}_2\text{)}_3\text{-AGCGTTCTCACC}
\]

\[
\text{Oligo13} \quad \text{CAAGGTCAGACATGTC - (CH}_2\text{)}_6\text{-NH-CO-(CH}_2\text{)}_2\text{-C}_3\text{N}_2\text{-βCD}
\]

\[
\text{Oligo14} \quad \text{βCD-C}_3\text{N}_2\text{-CO-NH-(CH}_2\text{)}_3\text{-AGCGTTCTCACCAGTC}
\]

b)

**Figure 6:** a) Synthesis of cyclodextrin-DNA conjugates. b) Representative denaturing PAGE of cyclodextrin-DNA conjugates. Lanes: (1) Reference: bromophenol blue (2) CAAGGTCAGACATGTC-(CH}_2\text{)}_6\text{-NH}_2 \text{(3) Oligo3 (4) Oligo13 (5) H}_2\text{N-(CH}_2\text{)}_6\text{-AGCGTTCTCACCAGTC (6) Oligo4 (7) Oligo14.}

Analysis by MALDI-TOF (Table 4) showed that the first product eluting from the anion exchange column was the unreacted cyclodextrin, which does not interact with the positively charged resin, while the second peak corresponded to the pure cyclodextrin-DNA conjugate.

a)

**Figure 7:** a) UV traces of anion exchange chromatography of oligonucleotide-cyclodextrin conjugates: Oligo13 (left), Oligo14 (right). b) rp-HPLC traces of purified Oligo13 and Oligo14.
2.4.2 Split mDHFR fragments expression and purification
The same design of the mDHFR protein fragments as reported in literature was used (Figure 8). The N-terminal fragment of mDHFR (N\textsubscript{term}-mDHFR), corresponding to residues 1-105, features a C-terminal cysteine and an N-terminal His-tag that facilitates purification. For the C-terminal fragment of mDHFR (C\textsubscript{term}-mDHFR), encompassing residues 106-186, a cysteine was introduced close to the N-terminus followed by a short linker to the mDHFR sequence in order to reduce steric hindrance at the bioconjugation site and improve coupling efficiency. Similarly to the N\textsubscript{term}-mDHFR, a His-tag was introduced at the C-terminus of the sequence.

Expression plasmids encoding for full length mDHFR, N\textsubscript{term}-mDHFR and C\textsubscript{term}-mDHFR fragments were transformed and proteins were produced independently in E. coli BL21 (DE3). Due to insolubility of mDHFR and its fragments, cells were harvested under denaturing conditions and purification via affinity chromatography (Ni-NTA) afforded the full length mDHFR and the N- and C-terminal mDHFR fragments with good purity and in satisfactory yields (30-34 mg/L, 10-14 mg/L and 6-13 mg/L, respectively). High resolution mass spectrometry and Tricine-SDS-PAGE independently confirmed the identity and purity of all the variants. Dimer formation was observed in the Tricine-SDS-PAGE with bands corresponding to 29 and 22 KDa for the N- and C-terminal fragment, resulting from oxidative cystine formation.

2.4.3 Synthesis of mDHFR fragment-DNA conjugates via cysteine-maleimide coupling
The synthesis of mDHFR-DNA conjugates was performed following a previously described procedure. This strategy involves the reaction of maleimide-functionalized oligonucleotides with unique cysteines on the mDHFR fragments. Cysteine-maleimide coupling can be applied to mDHFR since no native cysteine is present in its sequence.

Maleimide-containing oligonucleotides were prepared by conjugation of commercially available amino-functionalized oligonucleotides with an excess of the NHS ester of 3-maleimidopropionic acid in 200 mM Na\textsubscript{2}HPO\textsubscript{4} buffer at pH 7.2 (Figure 9). The reactions were performed at 25 °C for 2 hours and the resulting maleimide-functionalized oligonucleotides were purified by size exclusion chromatography, analyzed by rp-HPLC and MALDI-TOF and lyophilized (Table 2). Isolation of the maleimide-functionalized oligonucleotides by ethanol precipitation was also investigated as an alternative purification strategy to reduce
possible hydrolysis of the maleimide moiety. No significant differences were observed in the mass spectra of the oligonucleotides nor in the efficiency of the following coupling with the protein fragments.

Figure 9: Synthesis of maleimide-functionalized oligonucleotides

Coupling of the mDHFR fragments with 5’- and 3’-maleimide-modified oligonucleotides was performed as reported previously.\textsuperscript{34} Protein solutions in 100 mM NaH\textsubscript{2}PO\textsubscript{4} pH 4.0, 100 mM Tris-HCl, 8 M urea were buffered to pH 8.0 by addition of concentrated NaOH and incubated for 1 h with dithiothreitol (DTT) to reduce disulfides. Proteins were then precipitated by addition of 25% TFA in water and protein pellets were obtained after centrifugation. The pellets were thoroughly washed with ethanol and air dried before redissolving them in the reaction buffer (100 mM CHES buffer pH 9.2 containing 8 M urea). The resulting solutions were then immediately added to the freeze-dried maleimide-modified oligonucleotide. The reactions were incubated at 40 °C overnight and the crude mixtures were then analyzed by Tricine-SDS-PAGE to estimate the conversion. In the Tricine-SDS-PAGE bands corresponding to the unreacted proteins, the protein-DNA conjugates and dimer of the protein were observed (Figure 10).

Figure 10: a) Synthesis of protein-DNA conjugates via cysteine-maleimide coupling b) 12% Tricine-SDS-PAGE of reaction mixtures: 100 µM of protein (40 nmol), 300 µM oligonucleotide (120 nmol) in 100 mM CHES buffer pH 9.2, 8 M urea. Left: coupling reactions of N\textsubscript{term}-mDHFR with (1) Oligo7, 2h (2) Oligo7, 16 h, (3) Oligo8, 2 h, (4), Oligo8, 16 h. Right: coupling reactions of C\textsubscript{term}-mDHFR with (1) Oligo7, 2h (2) Oligo7, 16 h, (3) Oligo8, 2 h, (4) Oligo8, 16 h.
Reaction conditions were optimized to improve conversions and reduce the amount of dimer formed during the coupling reactions. Important factors proved to be a complete removal of DTT and a strict control over the pH of the reaction mixture. Traces of DTT or a pH lower than 9 (probably due to residual TFA in the reaction tube) resulted in low to no conversion. In order to completely remove DTT from the protein solution, two TFA precipitations were performed, followed by several washing steps of the resulting pellet with ethanol to remove traces of both DTT and TFA. Reduction of the disulfide bonds with TCEP, a reducing reagent compatible with cysteine-maleimide coupling, was also tested but did not lead to any significant improvement. Fast re-dissolving of the protein pellet and addition to the freeze-dried oligonucleotide also turned out to be important: slightly longer incubation time in CHES buffer at pH 9.2 increased the amount of dimer formation in the reaction mixture and lowered conversions.

Different protein:oligonucleotide ratios were also tested; an excess of oligonucleotide improved conversions, yet a large excess of oligonucleotide is not recommended to avoid unspecific coupling of the maleimide-modified oligonucleotides to residues other than the desired cysteine. Although, the nucleophilicity of the thiol group is higher when compared to the amino groups of lysines and of the N-terminus, under the denaturing conditions of the coupling reaction, all amino acids are solvent exposed and more likely to react. 3 equivalents of oligonucleotide were found to be ideal. Coupling with oligonucleotides of different length (12-mer, 16-mer) was also tested, as well as 5'- or 3'-modification, but did not result in significant differences.

Under optimized conditions, conversions of around 50-60% were observed (estimated from the Tricine-SDS-PAGE) (Figure 10) although the overall reproducibility of the procedure was poor. Moreover, in several coupling reactions the appearance of bands with an apparent higher molecular weight was observed, indicative of coupling more than one oligonucleotide to the protein fragments. A 1:1 protein:oligonucleotide ratio in the conjugate is important, as proteins with more than one oligonucleotide attached would result in a significant error when determining the concentration of the conjugates (measured via absorption at 260 nm). Ultimately though, the inherent lack of reproducibility of the cysteine-maleimide conjugation forced us to explore different coupling strategies in order to obtain protein-DNA conjugates in higher yields and purities.

2.4.4 Synthesis of protein-DNA conjugates via copper catalyzed azide alkyne cycloaddition (CuAAC)
Cu(I)-catalyzed 1,3-dipolar azide-alkyne cycloaddition (CuAAC) was selected as an alternative to cysteine-maleimide coupling to prepare protein-DNA conjugates. This strategy has already been applied to the ligation of protein and peptides to oligonucleotides. In order to utilize CuAAC for the synthesis of
protein-DNA conjugates an azide and an alkyne functionality needed to be introduced in each of the reactant. Since mDHFR fragments containing terminal cysteines were available from the previous approach, an alkyne functionality was introduced into the protein fragments by alkylation of the thiol group with N-propynyl iodoacetamide as reported in literature.\(^6\) Consequently, the azido moiety was introduced on the oligonucleotides.

![Figure 11: Synthesis of azide-functionalized oligonucleotides.](image)

3'- and 5'-azide-functionalized oligonucleotides were prepared starting from commercially available amino-modified oligonucleotides by conjugation with excess of the NHS ester of 3-azidopropionic acid (Figure 11). Reaction mixtures were incubated overnight at 25 °C and the resulting azide-functionalized oligonucleotides were purified by size exclusion chromatography and analyzed by rp-HPLC and MALDI-TOF and lyophilized (Table 3).

To introduce the alkyne moiety on the mDHFR fragments containing cysteines, a reaction with 30 eq. propynyl iodoacetamide in 100 mM Tris pH 8.5, 8 M urea was performed (Figure 12). The protein fragments were incubated with DTT, precipitated with 25% TFA and thoroughly washed with ethanol as described previously in order to reduce disulfide bonds. Alkylation with the iodoacetamide moiety was performed using degassed buffer and under inert atmosphere at 25 °C overnight under continuous shaking.

![Figure 12: Synthesis of alkyne-functionalized protein fragments](image)

MALDI-TOF analysis showed formation of the desired product, but also traces of di- and tri-functionalized proteins. To prevent the alkylation of other nucleophilic amino acid side chains (i.e. lysines), milder reaction conditions were tested. Propynylbromoacetamide was used instead of the iodo derivative and the reaction was evaluated at different pH (7.8, 8.0, 8.2). 10 equivalents of alkylating reagent in 100 mM Tris buffer pH 8.0 containing 8 M urea provided the best results. MALDI-TOF analysis showed the exclusive formation of the
monoalkylated product, albeit not quantitatively. However, for this reaction lower conversions were considered to be a better option than multi-functionalization of the protein. Unlabeled protein will, indeed, not interfere in the CuAAC to the oligonucleotide and could later be separated from the protein-DNA conjugate. On the other hand, the introduction of more than one alkyne moiety on the protein might result in coupling of more than one oligonucleotidic sequence. The excess of propynyl-bromoacetamide was removed from the reaction mixtures by repeated precipitation of the protein with 25% TFA in water.

Coupling of alkyne-mDHFR fragments with azide-functionalized oligonucleotides was performed under inert atmosphere by dissolving protein pellets in degassed 100 mM NaH$_2$PO$_4$ pH 8.0, 8 M urea containing the azide-modified oligonucleotide (1.2 eq) as well as CuSO$_4·$5H$_2$O, THPTA and sodium ascorbate (Figure 13a). Reactions were incubated overnight at 25 °C and analyzed by Tricine-SDS-PAGE (Figure 13b).

Tricine-SDS-PAGE showed formation of mDHFR-DNA conjugates for both fragments, corresponding to the bands around 20 to 24 KDa for C- and N-terminal fragment, respectively. Conversions cannot be estimated from these gels, since the lower bands presumably contain both the alkyne-modified and free cysteine proteins.

Figure 13: a) Synthesis of mDHFR fragments-DNA conjugates via CuAAC b) Tricine-SDS-PAGE of reaction mixture of CuAAC coupling. 100 μM of protein (100 nmol), 120 μM oligonucleotide in 100 mM NaH$_2$PO$_4$ pH 8.0, 8 M urea at 25 °C, 16 h. Left: triplicates of the reaction of Nterm-mDHFR-alkyne with Oligo9. Right: triplicates of the reaction of Cterm-mDHFR-alkyne with Oligo10.

With the exception of a slight variation in the preceding alkylation reaction, the CuAAC bioconjugation proved much more reproducible than the previously employed cysteine-maleimide strategy. Protein-DNA conjugates were incubated with EDTA to remove Cu$^{2+}$ ions and purified, after removal of EDTA with repeated ultracentrifugation, in a two-step protocol. First, strong anion exchange chromatography was used to separate the protein-DNA conjugates from the alkyne-functionalized protein, the unreacted cysteine proteins and possible dimers. Elution from the anion exchange column with high ionic strength buffer afforded a
mixture that contained protein-DNA conjugates and unreacted oligonucleotide (Figure 14a). These elution fractions were then concentrated and loaded onto a Ni-NTA column, on which only the protein-DNA conjugates were retained. Following several wash steps to remove the excess of unbound oligonucleotides, elution under acidic conditions (pH 4.0) afforded pure protein-DNA conjugates as judged by Tricine-SDS-PAGE (Figure 14b).

![Diagram showing UV traces of anion exchange chromatography of protein-DNA conjugates](image1)

**Figure 14:** a) UV traces of anion exchange chromatography of protein-DNA conjugates. b) Tricine-SDS-PAGE after Ni-NTA affinity chromatography of protein-DNA conjugates. (A) Indicates the first peak eluted from the anion exchange chromatography. c) Tricine-SDS-PAGE of purified protein-DNA conjugates: (1) Nterm mDHFR_Oligo9 and (2) Cterm mDHFR_Oligo10.

### 2.4.5 Split mDHFR reassembly assay

The reassembly of the mDHFR split enzyme was verified first by DNA hybridization as previously reported,\(^{34}\) then the small molecule mediated reassembly was tested. To monitor the reassembly of mDHFR into an active conformation, the catalytic activity of the enzyme was assayed in the NADPH-dependent reduction of dihydrofolate to tetrahydrofolate. This reaction is conveniently monitored by UV-visible spectroscopy following changes in the NADPH absorption at 340 nm.\(^ {46,47}\) A rapid dilution protocol that has previously been reported was used to assess the reassembly of mDHFR: unfolded mDHFR is known to refold and recover catalytic activity when diluted in buffer without denaturing agent in presence of its natural substrates.\(^ {47}\)

Equimolar amounts of Nterm-mDHFR_Oligo9, Cterm-mDHFR_Oligo10 and the template DNA were pre-mixed under denaturing conditions and subsequently diluted into the reaction buffer containing dihydrofolate (100 µM) to a final concentration of 0.1 µM. NADPH was added last (100 µM) to initiate the reaction.
and the progress was monitored following the absorbance at 340 nm over time (Figure 15).

As previously reported, in absence of the template DNA, the two fragments of mDHFR do not reassemble into an active conformation and no catalytic activity was observed (curve 2 – pink). Conversely, in presence of the fully complementary DNA template, a rapid consumption of NADPH was observed (curve 3 – green), demonstrating that the secondary interactions provided by DNA hybridization mediate the reassembly of the split mDHFR enzyme.

**Figure 15:** Kinetic curves for the consumption of NADPH in the reduction of dihydrofolate to tetrahydrofolate. Uncatalyzed NADPH degradation (1 – black, dashed), \( \text{N}_{\text{term}} \text{-mDHFR_Oligo9} + \text{C}_{\text{term}} \text{-mDHFR_Oligo10} \) (2 – pink), \( \text{N}_{\text{term}} \text{-mDHFR_Oligo9} + \text{C}_{\text{term}} \text{-mDHFR_Oligo10} + \text{template DNA} \) (3 – green), full length mDHFR (4 – Blue).

Compared to the previously reported DNA-mediated split mDHFR, this split enzyme proved to catalyze the reduction of dihydrofolate to tetrahydrofolate 3-times slower in presence of the fully complementary template DNA, with an initial rate of 0.1±0.01 sec\(^{-1}\). This difference might either be a result of different purities of the protein-DNA conjugates or of the different linkers between the protein and the oligonucleotides. Indeed due to the different coupling strategies used, different linkers are present between the protein fragments and the oligonucleotides that might result in different conformations of the reassembled enzymes. Nevertheless, reassembly of the enzyme was observed, providing a valuable comparison for the small molecule-mediated reassembly approach.

The reassembly of the split mDHFR mediated by small molecules with the cyclodextrin-DNA conjugates was tested in a similar fashion. Equimolar amounts of protein-DNA conjugates (\( \text{N}_{\text{term}} \text{-mDHFR_Oligo9} \) and \( \text{C}_{\text{term}} \text{-mDHFR_Oligo10} \)) and of cyclodextrin-DNA conjugates (Oligo13 and Oligo14) were pre-mixed under denaturing conditions at a final concentration of 2 µM. The effector molecule bis(2-((1R,3R,5S)-adamantan-1-yl)ethyl) phosphate was added at a final concentration of 4 µM. Then, the pre-mixed solution was diluted 20 times in the reaction buffer containing dihydrofolate. As before, NADPH was added to the mixture and the changes in absorbance at 340 nm were monitored over time (Figure 16).
As expected, no reassembly was observed in presence of the two fragments of mDHFR and the two cyclodextrin-DNA conjugates (line 2, blue), but also no reassembly was detected when the effector molecule was present in the mixture (line 3, pink). This lack of activity could be attributed to the fact that supramolecular interactions between the cyclodextrins and the bisadamantyl phosphate are not strong enough to reassemble the two split mDHFR fragments into an active conformation, or that, under the reaction conditions, the 2:1 complex between the cyclodextrins and the guest molecule is not formed.

![Figure 16: Kinetic curves for the consumption of NADPH in the reduction of dihydrofolate to tetrahydrofolate.](image)

In the assumption that the designed system resembles the structure of cyclodextrin dimers, the binding affinity of the guest molecule to the cyclodextrin is expected to be around $10^7 \text{ M}^{-1}$.\textsuperscript{49,50} In the reaction conditions (0.1 µM cyclodextrin-DNA conjugates, 0.2 µM bisadamantyl phosphate) considering a binding affinity of $10^7 \text{ M}^{-1}$, around 30% of the guest molecules should be bound to the cyclodextrins (considered as a dimer, distribution diagram simulated by HysS2009 software). Addition of an excess of guest molecule would promote the formation of the complex, but it might also favor the formation of 1:1 complexes instead of the 2:1 complex that would be needed to mediate the reassembly of the split enzyme. Also, an excess of guest molecule could lead to unspecific interactions with the protein structure that might alter the activity of the enzyme. Determination of the binding affinity of the guest molecule for the cyclodextrins in the designed system would allow to adjust the concentration of all components and perform the reassembly assay in conditions in which the complex is formed. Unfortunately, attempts to measure the binding affinity via Isothermal Titration Calorimetry (ITC) were not successful yet.
2.5 Conclusions

In conclusion, this chapter described the design and development of an artificial allosteric split enzyme for small molecule recognition. This split enzyme is characterized by a modular assembly of the components. Cyclodextrin-oligonucleotide conjugates were successfully obtained via CuAAC approach from 6-azido-β-cyclodextrins and alkyne-functionalized oligonucleotides. The strategy for the synthesis of mDHFR fragment-oligonucleotide conjugates was optimized from a previously reported method\textsuperscript{34} applying CuAAC. The use of CuAAC, instead of cysteine-maleimide coupling, involves an extra step of protein functionalization to introduce the alkyne moiety on cysteine functionalized mDHFR fragments, but proved to give more reproducible results in terms of conversion towards the protein-DNA conjugate. Optimization of the reaction conditions for the alkyne functionalization of the protein gave rise exclusively to the mono-alkylated protein. The design of protein fragments involving \textit{in vivo} incorporation of non-natural amino acids containing either an alkyne\textsuperscript{55} or an azide\textsuperscript{54} moiety could also be explored in future studies.

As proof of principle, this system enabled the reassembly of the split mDHFR when a fully complementary template strand was used. Unfortunately, when the cyclodextrin-DNA conjugates and bisadamantyl phosphate as an effector molecule were employed, no catalytic activity was observed. Further studies should focus on elucidating the binding affinity of the effector molecule toward the cyclodextrin-DNA conjugates, in order to tune the reactions conditions and ensure the formation of the 2:1 complex between cyclodextrin-DNA conjugates and the effector molecule under the assay conditions. Possibly the use of guest molecules such as 6-(4-tert-butylanilino)naphthalene-2-sulfonate (BNS) that result in an increase in fluorescence signal upon complexation to the cyclodextrins would lead to easier determination of the binding affinity.\textsuperscript{50}

2.6 Experimental section

2.6.1 General remarks

Chemicals were purchased from Sigma Aldrich, Acros, Alfa Aesar (3-maleimidopropionic acid N-hydroxysuccinimide ester) or TCI Europe (6-maleimidohexanoic acid N-hydroxysuccinimide ester) and used without further purification. \textsuperscript{1}H-NMR and \textsuperscript{13}C-NMR spectra were recorded on a Varian 400 MHz in CDCl\textsubscript{3}, DMSO-d\textsubscript{6} or methanol-d\textsubscript{4}. Chemical shifts (δ) are denoted in ppm using residual solvent peaks as internal standard. rp-HPLC analysis were performed on a Shimadzu LC-10AD VP, Waters Xterra MS C18 column (3.0 x 150 mm, particle size 3.5 μm) using a gradient of CH\textsubscript{3}CN/triethylammonium acetate (TEAA) buffer 50 mM at pH 7; gradient: 5/95 0 to 10 min, to 35/65 at 60 min, to 70/30 at 65 min. Flow: 0.5 mL/min. MALDI-TOF measurements were carried out on a Voyager-DE Pro apparatus. Matrices used for measurements: THAP (oligonucleotides): 50% 2, 4, 6- trihydroxyacetophenone in ethanol (10 mg/mL), 50% ammonium citrate dibasic in milliQ water (20 mg/mL), 0.1 % TFA. HPA (oligonucleotides): 90% saturated 3-hydroxypticolinic acid in milliQ water:CH\textsubscript{3}CN 1:1, 10%
ammonium citrate dibasic in milliQ water (100 mg/mL), 0.1 % TFA. **SA** (proteins): 10 mg/mL of sinapic acid in milliQ water:CH₃CN 10:75, 0.1 % TFA. **DHB** (cyclodextrins) 10 mg/mL of 2,5-dihydroxybenzoic acid in milliQ water:CH₃CN 1:1, 0.1 % TFA. UV-visible spectra were recorded on Jasco V-660 Spectrophotometer at 25 °C. mDHFR assays were conducted in quartz cuvettes with a 1 cm path length. Oligonucleotides were purchased from BioTez Berlin-Buch GmbH. Preparative size exclusion chromatography on oligonucleotides was performed with Sephadex TM G-25 DNA Grade. DNA denaturing gels were performed in minigel BioRad apparatus and staining was obtained with Stains-All (Sigma-Aldrich). The concentration of the protein and oligonucleotides was measured with Nanodrop 2000 (Thermo Scientific). Extinction coefficients of proteins (ε 280) and of oligonucleotides (ε 260) were calculated using the Protparam tool on the Expsay server and the oligoanalyzer tool on the Integrated DNA Technology website, respectively. *E. coli* BL21 (DE3) were used for protein production. Centrifugation was performed using a Beckman Coulter Avanti J-E centrifuge. Äkta Purifier 900 (GE Healthcare) was used for Fast Protein Liquid Chromatography (FPLC). FPLC columns (HiTrap QFF, HisTrap HP and Superdex 75 10/300 GL) were purchased from GE Healthcare. Ni-NTA agarose resin was purchased from Qiagen. Tricine-SDS-PAGE were performed in minigel BioRad apparatus and Coumassie staining was obtained with InstantBlue, (Expedeon). The concentration of proteins was achieved with ultracentrifugation filters Amicon-ultra 15mL and 0.5 mL. Expression plasmids pTWINXa-mDHFR, pTWINXa-Νterm-mDHFR and pTWINXa-GOSGG_Cterm-mDHFR encoding for full length mDHFR and 1-105 fragment and 106-186 fragment of mDHFR respectively, were available from previous work.³⁴

### 2.6.2 Synthesis

**Mono-6-deoxy-6-(p-tolylsulfonyl)-β-cyclodextrin** (1) was synthesized according to a literature procedure⁵⁵ starting from 1.5 g (1.32 mmol) of β-cyclodextrin with a 20% yield (335 mg, 260 mmol). The product was analyzed by MALDI-TOF (DHB): m/z meas.1311 Da for [M+Na]+ (m/z calc. C₃₀H₇₆O₇SNa = 1312.16 Da)

**Mono-6-deoxy-6-azido-β-cyclodextrin** (2) was synthesized according to a literature procedure⁵⁵ starting from 177 mg (0.137 mmol) of 1 and it was obtained with a 39% yield (62 mg, 0.053 mmol). The product was analyzed by MALDI-TOF (DHB): m/z meas. 1161 Da for [M+H]+ and 1183 Da for [M+Na]+ (m/z calc. C₂₅H₂₆N₆O₃ = 1160.00 Da).

**4-Pentynoic acid N-hydroxysuccinimide ester** (3) was synthesized adapting a literature procedure.⁴¹ 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC-HCl, 1.1 mg, 5.6 mmol) was added to a solution containing 4-pentynoic acid (500 mg, 5 mmol) and N-hydroxysuccinimide (644 mg, 5.6 mmol) in 50 mL dichloromethane. The resulting mixture was stirred at room temperature overnight and then washed with saturated NaHCO₃ and saturated NaCl. The organic layer was separated, dried over anhydrous Na₂SO₄ and evaporated in vacuo. Recrystallization from cold ethanol afforded the pure product as a white solid in 44% yield (428 mg, 2.2 mmol). Analytical data were in accordance with those previously published. ¹H-NMR (CDCl₃, 400 MHz) δ (ppm): 2.05 (t, J=2.6 Hz, 1H), 2.62 (td J₁=2.6 Hz, J₂=7.0 Hz, 2H), 2.84 (s, 4H), 2.88 (t, J=7.0 Hz, 2H).

**3-azido-1-propanol** (4) was synthesized according to a literature procedure⁴¹ starting from 1.4 g of 3-bromo-1-propanol with a 70% yield (709 mg, 7 mmol). Analytical data were in accordance with those previously published. ¹H-NMR (CDCl₃, 400 MHz) δ (ppm): 1.81 (quin, J=6.0 Hz, 2H), 2.01 (s, 1H), 3.43 (t, J=8.0 Hz, 2H), 3.73 (t, J=8.0 Hz, 2H).
3,3',3''- (4,4',4''-(nitrilotris(methylene))tris(1H-1,2,3-triazole-4-1 diyl))tris(propan-1-ol) (THPTA) (5) was synthesized adapting a literature procedure. Tripropargylamine (130 mg, 1.0 mmol) in an acetonitrile: methanol 1:1 solution (3 mL) was cooled to 0 °C in an ice bath and then 20 mL of 2.49 Hz (1H), 4.10 (s, 2H), 4.11 (2H) 

Organic layer was evaporated. The crude mixture was dissolved in methanol (2 mL) and the product was precipitated in acetonitrile (10 mL) overnight at 4 °C. The product was obtained as a white solid with a 62% yield (942 mg, 2.2 mmol). Analytical data were in accordance with those previously published. 

bis(2-((1R,3R,5S)-adamantan-1-yl)ethy) phosphate (6) was synthesized according to a literature procedure starting from 2 g (11 mmol) of 1-adamantane ethanol with a 9% yield (400 mg, 0.1 mmol). Analytical data were in accordance with those previously published. 

3-(Maleimido)propionic acid N-hydroxysuccinimide ester (7) was synthesized according to a literature procedure starting from 2 g (22 mmol) of β-alanine with a 55% yield (3.2 g, 12 mmol). Analytical data were in accordance with those previously published. 

3-azidopropionic acid (8) was synthesized according to a literature procedure starting from 3 g of bromopropionic (20 mmol) acid with a 47% yield (1.1 g, 9.4 mmol). Analytical data were in accordance with those previously published. 

3-azidopropionic acid N-hydroxysuccinimide ester (9) was synthesized according to a literature procedure starting from 1.1 g of 8 (9.4 mmol) with a 90% yield (1.8 g, 8.4 mmol). Analytical data were in accordance with those previously published. 

N-propynyl iodoacetamide (10) and N-propynyl bromoacetamide (11) were synthesized adapting a literature procedure. Iodoacetic acid (1 g, 5.4 mmol) or bromoacetic acid (0.75 g, 5.4 mmol), propargylamine (326 mg, 5.9 mmol), 1-ethyl-3-(dimethylaminopropyl)carbodiimide hydrochloride (EDC·HCl, 1.03 g, 5.4 mmol) and dimethylaminopyridine (66 mg, 0.54 mmol) were dissolved in 20 mL of CH₂Cl₂. The mixture was stirred overnight at room temperature, then 20 mL of a saturated NaHCO₃ solution were added. The organic layer was separated and the mixture extracted once with 10 mL saturated NaCl. The combined organic layers were dried over anhydrous Na₂SO₄ and the solvent was evaporated in vacuo. N-propynyl iodoacetamide and N-propynyl bromoacetamide were obtained as yellowish solid in 40% (481 mg, 2.16 mmol) and 26% (253 mg, 1.43 mmol) yield, respectively. Analytical data for N-propynyl iodoacetamide were in accordance with those previously published. N-propynyl iodoacetamide \(^1\)H-NMR (CDCl₃, 400 MHz) δ (ppm): 2.28 (t, J=2.49 Hz, 1H), 4.10 (s, 2H), 4.11-4.12 (m, 2H). N-propynyl bromoacetamide \(^1\)H-NMR (CDCl₃, 400 MHz) δ (ppm): 2.28-2.29 (m, 1H), 4.09 (s, 2H), 4.10-4.11 (m, 2H).
2.6.3 Synthesis of alkyne-functionalized oligonucleotides
0.1 μmol of a stock solution of amino-modified-oligonucleotide in H2O were diluted in NaH2PO4 200 mM, pH 7.2 to a total volume of 0.9 mL. 100 μL of a stock solution 4-pentyonic acid succinimidyl ester (200 mM in N,N-dimethylformamide (DMF), 200 equivalents) were added to this solution. The reaction was incubated at 25 °C overnight under continuous shaking (800 rpm). The product was purified from the excess of linker by size exclusion chromatography (SephadexTM G-25 DNA Grade, 50 mM TEAA buffer pH 7.0) and lyophilized. Characterization was performed by rp-HPLC and MALDI TOF (THAP).

Table 1 rp-HPLC retention times and MALDI-TOF of alkyne-functionalized oligonucleotides

<table>
<thead>
<tr>
<th>Oligo</th>
<th>Sequence (5'-3')</th>
<th>r.t. (min)</th>
<th>MW meas</th>
<th>MW calc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oligo1</td>
<td>GTCAGACATGTC-(CH2)6-NH-CO-(CH2)6-C≡C</td>
<td>35.94</td>
<td>3908</td>
<td>3904</td>
</tr>
<tr>
<td>Oligo2</td>
<td>C≡C-(CH2)6-CO-NH-(CH2)6-AGCGTTCTCACCC</td>
<td>35.79</td>
<td>3843</td>
<td>3840</td>
</tr>
<tr>
<td>Oligo3</td>
<td>CAAGTCAGACATGTC-(CH2)6-NH-CO-(CH2)6-C≡C</td>
<td>34.72</td>
<td>5148</td>
<td>5147</td>
</tr>
<tr>
<td>Oligo4</td>
<td>C≡C-(CH2)6-CO-NH-(CH2)6-AGCGTTCTACCCGTC</td>
<td>25.64</td>
<td>5083</td>
<td>5077</td>
</tr>
</tbody>
</table>

2.6.4 Synthesis of maleimide-functionalized oligonucleotides
0.1 μmol of a stock solution of amino-modified-oligonucleotide in H2O were diluted in NaH2PO4 200 mM, pH 7.2 to a total volume of 0.9 mL. 100 μL of a stock solution of 3-maleimidopropionic acid N-hydroxysuccinimide ester (200 mM in DMF, 200 equivalents) were added to this solution. The reaction was incubated at 25 °C for 2 h under continuous shaking (800 rpm). The product was purified from the excess of linker by size exclusion chromatography (SephadexTM G-25 DNA Grade, 50 mM TEAA buffer pH 7.0) and lyophilized. The purification of the maleimide-functionalized oligonucleotides by ethanol precipitation was performed by addition of 100 μL of sodium acetate buffer at pH 5.2 to a final concentration of 0.3 M and incubation of the solution at -20 °C for 1-2 h. The precipitate was centrifuged at 13000 rpm for 30 min and washed 3 times with 70% ethanol before being air dried or lyophilized. Characterization was performed by rp-HPLC and MALDI-TOF (THAP).

Table 2 rp-HPLC retention times and MALDI-TOF of maleimide-functionalized oligonucleotides

| Oligo5 | Maleimide-CO-NH-(CH2)6-GACATGTCACCTTG | 35.12 | 3978 | 3977 |
| Oligo6 | GGTGAGAGACCTG-(CH2)6-NH-CO-Maleimide | 35.01 | 4039 | 4040 |
| Oligo7 | Maleimide-CO-NH-(CH2)6-GACATGTCACCTTG | 36.24 | 5194 | 5200 |
| Oligo8 | GACTGGTGAGAGACCTG-(CH2)6-NH-CO-Maleimide | 35.54 | 5284 | 5275 |

2.6.5 Synthesis of azide-functionalized oligonucleotides
0.1 μmol of a stock solution of amino-modified-oligonucleotide in H2O were diluted in NaH2PO4 200 mM pH 7.2 to a total volume of 0.9 mL. 100 μL of a stock solution 3-azidopropionic acid succinimidyl ester (200 mM in DMF, 200 equivalents) were added to this solution. The reaction was incubated at 25 °C overnight under continuous shaking (800 rpm). The product was purified from the excess of linker by size exclusion chromatography (SephadexTM G-25 DNA Grade, 50 mM TEAA buffer pH 7.0) and lyophilized. Characterization was performed by rp-HPLC and MALDI TOF (THAP/HPA).

Table 3 rp-HPLC retention times and MALDI-TOF of azide-functionalized oligonucleotides

| Oligo9 | N3-(CH2)6-CO-NH-(CH2)6-GACATGTCACCTTG | 29.21 | 5139 | 5147 |
| Oligo10 | GACTGGTGAGAGACCTG-(CH2)6-NH-CO-(CH2)6-N3 | 28.76 | 5210 | 5220 |
2.6.6 Synthesis of cyclodextrin-DNA conjugates via CuAAC
A solution of mono-6-deoxy-6-azido-β-cyclodextrin in DMSO (100 μL, stock solution 20 mM) was added to a solution of alkynyl-modified oligonucleotide in 200 mM NaH₂PO₄ pH 7.5 (250 μL 10 nmol). Subsequently, a pre-mixed solution of CuSO₄·5H₂O, THPTA and sodium ascorbate (150 μL of stock CuSO₄·5H₂O 1.7 mM, THPTA 5.5 mM and sodium ascorbate 17 mM in milliQ water) was added. Final concentrations: mono-6-deoxy-6-azido-β-cyclodextrin 3.5 mM, oligonucleotide 20 μM, CuSO₄ 0.5 mM, THPTA 1.6 mM and sodium ascorbate 5 mM in 500 μL. The reaction was incubated at 40 °C for 4 hours under continuous shaking (800 rpm). The reaction mixture was analyzed by denaturing PAGE (TBE buffer, 250 V, 60 min), injected on an anion exchange column (QFF column) and then flushed with 20 mM Tris, pH 8.0. After elution of the excess cyclodextrin, a gradient of 5 minutes with 20 mM Tris pH 8.0, allowed elution of the conjugates. The elution fractions were pooled, desalted by size exclusion chromatography (Sephadex G-25 DNA-grade, CH₃CN/TEAA 50 mM pH 7.0) and the resulting product was freeze-dried. Characterization was performed by rp-HPLC and MALDI-TOF (THAP/HPA).

Table 4 rp-HPLC retention times (r.t.) and MALDI-TOF of cyclodextrin-DNA conjugates

<table>
<thead>
<tr>
<th>Sequence (5’-3’ )</th>
<th>r.t. (min)</th>
<th>MW meas</th>
<th>MW calc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oligo11 GTCAGACATGTC - (CH₂)₃-βCD-C₂-N₆-βCD</td>
<td>33.52</td>
<td>5063</td>
<td>5064</td>
</tr>
<tr>
<td>Oligo12 βCD-C₂-N₆-βCD-C₂-N₆-βCD</td>
<td>33.25</td>
<td>5001</td>
<td>5000</td>
</tr>
<tr>
<td>Oligo13 CAAGGTACAGATGTC- (CH₂)₃-βCD-C₂-N₆-βCD</td>
<td>26.73</td>
<td>6311</td>
<td>6308</td>
</tr>
<tr>
<td>Oligo14 βCD-C₂-N₆-βCD-C₂-N₆-βCD</td>
<td>26.93</td>
<td>6235</td>
<td>6237</td>
</tr>
</tbody>
</table>

2.6.7 Expression and purification of mDHFR fragments
Plasmids pTWINXα-mDHFR, pTWINXα-N_term-mDHFR and pTWINXα-GGSGG_C_term-mDHFR were transformed into E. coli BL21 (DE3) cells, which were spread onto an agar plate containing 100 μg/mL of ampicillin. Single colonies were selected after overnight growth and used to inoculate 5 mL LB medium containing the same antibiotic. This starter culture was grown at 37 °C overnight and used to inoculate 500 mL fresh LB medium with the same antibiotic. The cultures were grown at 37 °C and when OD₆₀₀ reached 0.6 (0.9 for C_term-mDHFR) isopropyl β-D-1-thiogalactopyranoside (IPTG, 1 mM) was added to induce the expression of the target proteins. Expression was performed at 30 °C overnight (37 °C for N_term-mDHFR). Cells were harvested by centrifugation (6000 rpm, JA-10, 20 min, 4 °C), and the pellet was resuspended in 100 mM NaH₂PO₄, pH 8.0, 100 mM Tris-HCl, 8 M urea (10 mL). The cell lysate was incubated at room temperature for 30 min under continuous shaking and centrifuged (10000 rpm, JA-17, 30 min, 4 °C.). The supernatant was loaded onto a pre-packed Ni Sepharose High Performance column (HisTrap HP, 1 mL) with a flow of 0.5 mL/min. Once all the cell-free extract was injected and no more protein was eluting (no absorbance at 280 nm), the column was washed with 10-15 column volumes (CVs) of 100 mM NaH₂PO₄ pH 6.3, 100 mM Tris-HCl, 8 M urea and finally the protein was eluted with 10-15 CVs of 100 mM NaH₂PO₄ pH 4.0, 100 mM Tris-HCl, 8 M urea (purification of full length mDHFR was performed manually on Ni-NTA column, Qiagen). The elution fractions were analyzed on a 12% polyacrylamide Tricine-SDS gel followed by Coomassie staining. The fractions containing protein were pooled and concentrated using Amicon Ultra-15 centrifugation filters. The concentration of the proteins was measured with Nanodrop 2000, using the calculated extinction coefficients: full length mDHFR ε₂₈₀ = 74600 M⁻¹cm⁻¹, N_term-mDHFR ε₂₈₀ = 12490 M⁻¹cm⁻¹, C_term-mDHFR ε₂₈₀ = 12950 M⁻¹cm⁻¹. Typical expression yields for full length mDHFR, N_term-mDHFR and C_term-mDHFR were typically 30-34 mg/L, 10-14 mg/L and 6-13 mg/L respectively. Exact mass was measured for each protein, (ESI+, TOF).

**Full length mDHFR:** MW_meas=23659 Da (MW_calc=23658 Da)
**N_term-mDHFR:** MW_meas=14341 Da (MW_calc=14341 Da)
**C_term-mDHFR:** MW_meas=11008 Da (MW_calc=11140 Da, -Met 11008 Da)
2.6.8 Alkyne functionalization of mDHFR fragments
A protein solution (100 nmol) in 100 mM NaH₂PO₄ pH 4.0, 100 mM Tris-HCl, 8 M urea was brought to pH 8 with NaOH 2.5 M solution and DTT to a final concentration of 0.1 M was added. The solution was incubated for 1 h at 25 °C under continuous shaking. Then, 1 mL of 25% trifluoroacetic acid (TFA) in water was added and the mixture was incubated on ice for 1 h. The mixture was centrifuged at maximum speed (13000 rpm) for 10 min and the pellet was resuspended in 1 mL of TFA to remove traces of DTT. The mixture was centrifuged at maximum speed (13000 rpm) for 10 min and the pellet was washed with ethanol several times, air dried and then stored under nitrogen atmosphere. The protein pellet was re-dissolved under nitrogen atmosphere in degassed 100 mM Tris-HCl pH 8.0, 2 mM EDTA, 8 M urea (900 μL, final concentration protein 100 μM) and a 10 times molar excess of N-propynyl iodoacetamide or N-propynyl bromoacetamide from a 100 mM stock solution in DMF was added (10 μL, final concentrations 100 μM protein, 1 mM alkyne linker). The reaction was incubated at 25 °C overnight under continuous shaking (800 rpm). After this time 4 mL of 25% TFA in water were added and the mixture was incubated on ice for 1 h. The mixture was centrifuged at maximum speed (13000 rpm) for 10 min and the pellet was washed with ethanol several times, air dried and then stored dry until further use. Before TFA precipitation, a sample of the reaction mixture was taken and the reaction mixture was analyzed by 12% Tricine SDS-PAGE and MALDI-TOF.

Nterm-mDHFR_alkyne: MW_{meas}=14447 Da (MW_{calc}=14436 Da), 14351 Da (starting material, MW_{calc}=14341 Da)

Cterm-mDHFR_alkyne: MW_{meas}=11116 Da (-Met MW_{calc}=11008 Da), 11248 (unknown +142 Da from measured Cterm-mDHFR_alkyne)

2.6.9 Synthesis of protein-DNA conjugates via cysteine-maleimide coupling
Protein solutions eluted from Ni-NTA columns and concentrated in 100 mM NaH₂PO₄ pH 4.0, 100 mM Tris-HCl, 8 M urea, were buffered to pH 8.0 by addition of NaOH (2.5 M) and incubated with 100 mM DTT for 1 h at 25 °C under continuous shaking. After this time 1 mL of 25% TFA in water was added and the mixture was incubated on ice for 1 h. The mixture was centrifuged at maximum speed for 10 min and the pellet was resuspended in 1 mL of TFA to remove traces of DTT. The mixture was centrifuged at maximum speed for 10 min and the pellet was washed with ethanol several times, air dried and then redisolved in 100 mM CHES pH 9.2, 8 M urea to a final concentration of 100 μM (vortex and sonication used to resuspend the pellet, pH 8-9). This protein solution was immediately transferred to the freeze-dried maleimide-functionalized oligonucleotide (3 equivalents). The reaction was incubated at 40 °C overnight under continuous shaking (800 rpm). Reaction mixtures were analyzed by 12% Tricine-SDS-PAGE.

2.6.10 Synthesis of protein-DNA conjugates via CuAAC
Under a nitrogen atmosphere, the freeze dried azide-functionalized oligonucleotide was dissolved in 925 μL of degassed 100 mM NaH₂PO₄ pH 8.0, 8 M urea (120 nmol). This solution was added to the protein pellet of mDHFR fragment (100 nmol). Subsequently, 75 μL of a pre-mixed degassed solution of CuSO₄·5H₂O, THPTA and sodium ascorbate (stock of CuSO₄·5H₂O 13.5 mM, THPTA 2.7 mM and sodium ascorbate 40 mM) in milliQ water were added. Final concentrations were: protein 100 μM, oligonucleotide 120 μM, CuSO₄ 1 mM, THPTA 200 μM and sodium ascorbate 3 mM. The reaction was incubated at 25 °C overnight under continuous shaking (800 rpm). The reaction mixture was analyzed by 12% Tricine-SDS-PAGE. The crude mixture was incubated with EDTA (50 mM) at 25 °C for 1 h under continuous shaking (800 rpm) and concentrated to 200 μL with an Amicon-ultra 0.5 mL centrifugal filter. The concentrate was reconstituted with 400 μL of 100 mM NaH₂PO₄ pH 8.0, 8 M urea and this step was repeated five times. The resulting mixture was injected on a pre-packed strong anion exchange column (HiTrap QFF) and the column was washed with 100 mM Tris-HCl pH 7.5, 8 M urea, to remove all unreacted protein (5 column volumes, 1 mL/min). The conjugates were eluted from the anion exchange in 5 minutes with a high ionic strength buffer.
(100 mM Tris-HCl pH 7.5, 8 M urea, 1 M NaCl). Elution fractions were pooled and loaded onto a pre-equilibrated slurry of Ni-NTA (50% Ni-NTA in 20% ethanol, 100 µL) and incubated for 30 min at room temperature (mixed at 200 rpm on a rotary shaker). The column was washed five times with 1 CV of 100 mM NaH₂PO₄ pH 8.0, 100 mM Tris-HCl, 8 M urea (500 µL) until all unreacted oligonucleotide was removed and the conjugate was eluted with five fractions of 0.5 CV of 100 mM NaH₂PO₄ pH 4.0, 100 mM Tris-HCl, 8 M urea (5x50 µL). Elution fractions were analyzed on a 12% polyacrylamide Tricine-SDS gel followed by Coumassie staining. Fractions containing the protein-DNA conjugate were pooled and concentrated using Amicon-ultra 0.5 mL centrifugal filters. The concentration of the conjugate was measured with Nanodrop 2000, using the calculated extinction coefficients for the oligonucleotide at 260 nm. The contribution of the protein to the absorbance at 260 nm of the conjugate was considered negligible compared to the oligonucleotide. The conjugates were analyzed by analytical size exclusion chromatography using a Superdex 75 HR 10/30 column, pre-equilibrated with 100 mM Tris-HCl pH 7.5, 8 M urea buffer.

2.6.11 Split mDHFR reassembly

Equimolar amounts of N₅₅₅₅-mDHFR_Oligo7, C₅₅₅₅₅₅-mDHFR_Oligo8, Oligo13 and Oligo14 or template DNA were pre-mixed under denaturing conditions (100 mM NaH₂PO₄ pH 4.0, 100 mM Tris-HCl, 8 M urea) to a final concentration of 2 µM (50 µL). When required, the guest molecule, bis(2-(1R,3R,5S)-adamantan-1-yl)ethyl) hydrogen phosphate, was added at a final concentration of 4 µM. This pre-mixed solution was diluted 20 times in the reaction buffer 50 mM Tris pH 7.7, 5 mM MgCl₂, 3.3 mM KCl, 10 mM DTT containing dihydrofolate (100 µM) directly in the cuvette, to a final concentration of 0.1 µM for each component and 0.2 µM for the guest molecule. NADPH was added at a final concentration of 100 µM (from a stock solution of 50 mM in milliQ water). The change in absorbance at 340 nm was monitored over time. The sequence of template DNA used for the reassembly (5’-3’) was CAAGGTCAAGCATGCACGTTTCACCAGTC. The initial rate for the reduction of dihydrofolate to tetrahydrofolate catalyzed by the split enzyme in presence of the fully complementary oligonucleotide was calculated as average of three independent experiments and the error reported corresponds to the standard deviation. In calculating the initial rate the first 10 min of reaction time were discarded due to the presence of a lag phase, as previously reported.³⁴

2.7 References

An artificial split enzyme for small molecule recognition
