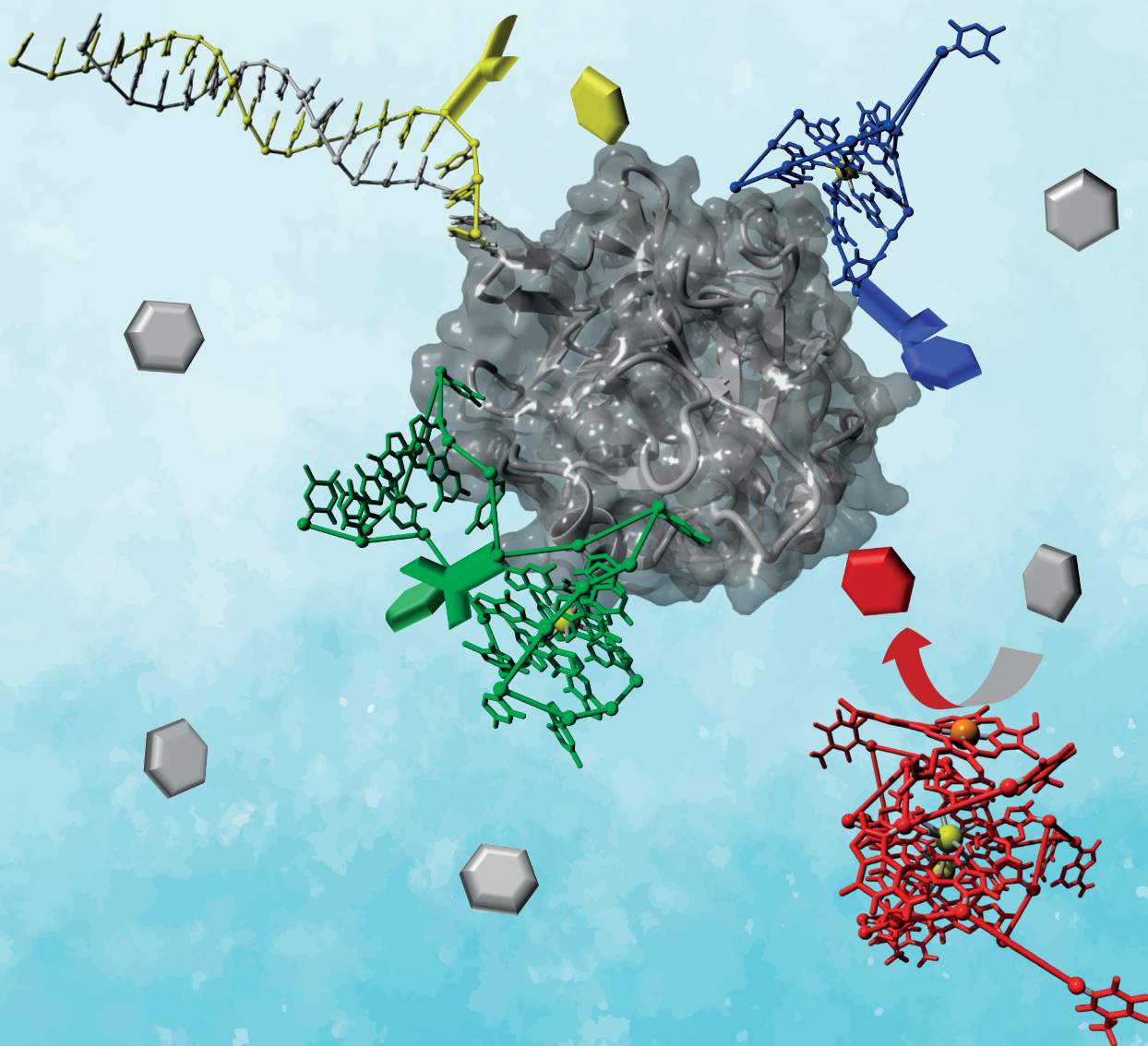


Controlled Modification of Wild-type Proteins using DNA Nanotools



Jordi Frederik Keijzer

Propositions

1. DNA is an efficient tool for pre- and posttranslational protein modification. (*this thesis*)
2. A DNA system will be the first artificial protein modification catalyst that has all features of a natural enzyme. (*this thesis*)
3. The claim that RNAzymes selectively modify RNA-binding proteins in cell lysate is an overstatement. (*Chem Commun.* **2020**, 56, 11641-11644)
4. More advanced medicine causes more advanced genetic diseases.
5. Reading a publication of four pages requires about the same amount of energy as making a 15 minute walk.
6. Nowadays education empowers the student more than the teacher.
7. Research, fishing and dating require the same mindset.

Propositions belonging to the thesis, entitled

“ Controlled Modification of Wild-type Proteins using DNA nanotools ”

Jordi Frederik Keijzer

Wageningen, 2nd of September 2022

Controlled Modification of Wild-type Proteins using DNA nanotools

Jordi Frederik Keijzer

Thesis Committee

Promotor

Prof. Dr H. Zuilhof
Professor of Organic Chemistry
Wageningen University and Research

Co-promotor

Dr. H. B. Albada
Assistant professor, Laboratory of Organic Chemistry
Wageningen University and Research

Other members

Prof. Dr W. J. H. van Berkel, Wageningen Universiteit & Research
Prof. Dr. J.G. Roelfes, Stratingh Inst for Chemistry, University of Groningen
Prof. Dr. M. Merkx, Eindhoven University of Technology
dr. W.A. Velema, Radboud University, Nijmegen

This research was conducted under the auspices of the Graduate School Food Technology
Agrobiotechnology Nutrition and Health Science.

Controlled Modification of Wild-type Proteins using DNA nanotools

Jordi Frederik Keijzer

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University,
by the authority of the Rector Magnificus,
Prof Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Friday 2 September 2022
at 4 in the Omnia Auditorium.

Jordi F. Keijzer
Controlled Modification of Wild-type Proteins using DNA nanotools
160 pages

PhD thesis, Wageningen University, Wageningen, the Netherlands (2022)
With references, with summary in English

ISBN: 978-94-6447-274-5
DOI: <https://doi.org/10.18174/571669>

Voor mijn grootouders, ouders, broer en partner

Table of contents

Abbreviations		9
Preface	Samenvatting voor de niet-wetenschapper	13
	Summary for the non-scientist	15
Chapter 1	General introduction	17
Chapter 2	Site-Specific and Trigger-Activated Modification of Proteins by Means of Catalytic Hemin/G-quadruplex (hGQ) DNAzyme Nanostructures	39
Chapter 3	Aptamer-Assisted Bioconjugation of Tyrosine Derivatives using hemin/G-quadruplex (hGQ) DNAzyme Nucleoapzyme Nanostructures	63
Chapter 4	Site-Selective and Inducible Acylation of Thrombin using DNA Aptamer-Catalyst Conjugates	77
Chapter 5	Calibrating site-selective protein modification by catalytic DNA nanostructures	105
Chapter 6	Discussion	129
Chapter 7	Summary	143
Acknowledgements		149
About the author		155

Abbreviations

5Hexynyl	= hexynyl chain at the 5' end of a DNA sequence
ABAL	= aptamer based affinity labelling
ACN	= acetonitrile
ANANS	= alkylated <i>N</i> -acyl- <i>N</i> -sulfonamide
Anh.	= anhydrous
α S	= alfa-synuclein
BA	= benzoic acid
BaTAm	= scrambled tyrosinamide binding aptamer
BCN	= (1R,8S,9s)-Bicyclo[6.1.0]non-4-yn
BME	= 2-mercaptoethanol / β -mercaptoethanol
brine	= saturated NaCl solution in water
BSA	= Bovine Serum Albumin
CDCl ₃	= deuterated chloroform
CHCl ₃	= chloroform
CuAAC	= copper-assisted alkyne-azide cycloaddition
CHY	= chymotrypsin
d	= doublet
dd	= doublet of doublets
ddH ₂ O	= double-distilled water
DCM	= dichloromethane
DIPEA	= <i>N,N</i> -diisopropylethylamine
DMAP	= dimethylaminopyridine
DMF	= dimethylformamide
DMSO	= dimethyl sulfoxide
DMSO-d ₆	= deuterated dimethyl sulfoxide
DNA _{temp}	= template DNA strand
DNA _{react}	= reacting DNA strand
dsDNA	= double-stranded DNA
dt	= doublet of triplets
DTT	= dithiothreitol
eq.	= equivalents
Et	= ethyl
EtOAc	= ethyl acetate
EtOH	= ethanol
ESI	= electron spray ionisation

Abbreviations

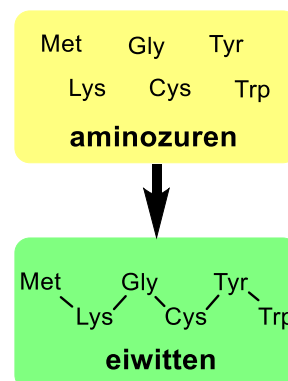
FA	= formic acid
h	= hour
(h)GQ	= (hemin/) <i>G</i> -Quadruplex
HEPES	= 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	= histidine
HPLC	= high-performance liquid chromatography
HRP	= horseradish peroxidase
Hz	= hertz
IAA	= iodoacetic acid
iAzideN	= non-terminal azide-functionalized thymine in a DNA sequence
IgG	= immunoglobulin G
i5OctdU	= non-terminal octynyl-functionalized thymine in a DNA sequence
J	= coupling constant
Kolla	= tsuur
LBA	= lysozyme binding aptamer
... ^{LC}	= on the light chain
LC-MS	= liquid chromatography–mass spectrometry
Lys	= lysine
m	= multiplet
Me	= methyl
MeCN	= acetonitrile
MeOD	= methanol- <i>d</i> ₄
MeOH	= methanol
min	= minute(s)
MS	= mass spectrometry
N ₃	= azide
NAD(H)	= nicotinamide
NEt ₃	= triethylamine
NF-κB	= nuclear factor-κ-β
NHS	= <i>N</i> -hydroxysuccinimide
NML	= <i>N</i> -methylluminol
NMR	= nuclear magnetic resonance
NTA	= tris(nitrilotriacetic acid)
Phe	= phenylalanine
PTAD	= 4-Phenyl-1,2,4-triazole-3,5-dione
PyOx	= pyridinecarbaldehyde oxime
q	= quartet
RT	= room temperature (20-25°C)

s	= singlet
SAA	= surface accessible area
SDS-PAGE	= sodium dodecyl sulfate–polyacrylamide gel electrophoresis
SELEX	= Systematic evolution of ligands by exponential enrichment
SPAAC	= strain-promoted alkyne-azide cycloaddition
ssDNA	= single-stranded DNA
t	= triplet
TamBA	= tyrosinamide binding aptamer
TAMN	= (0.5% TFA + 30% ACN + 70% ddH ₂ O + 400 mM NaCl)
TBA	= thrombin binding aptamer, a.k.a. HD01
TBA2	= thrombin binding aptamer 2, a.k.a. HD22
tBu	= <i>tert</i> -butyl
Temp	= template
THF	= tetrahydrofuran
Thrombin	= human alpha thrombin
THPTA	= tris(3-hydroxypropyltriazolylmethyl)amine
TLC	= thin layer chromatography
Tras	= trastuzumab
Tris	= tris(hydroxymethyl)aminomethane
TRM	= thrombin
Trp	= tryptophan
Tsuur	= kolla
Tyr	= tyrosine
TyrAm	= L-tyrosinamide
UV	= ultraviolet
UV/VIS	= ultraviolet/visible light

Preface

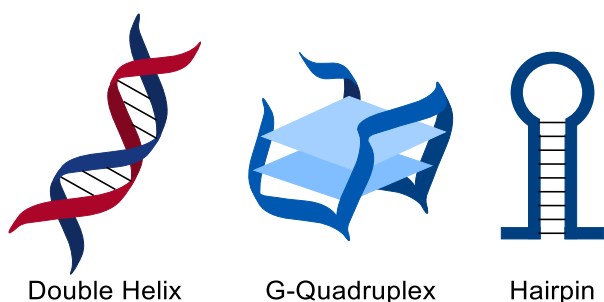
Samenvatting voor de niet-wetenschapper

Eiwitten zijn grote moleculen die in de natuur voorkomen en die elk bestaan uit 20 verschillende soorten bouwstenen (*aminozuren*). Omdat eiwitten soms uit wel meer dan 500 van deze bouwstenen bestaan, zijn er enorm veel combinaties mogelijk. Elke combinatie van bouwsteentjes kan leiden tot een nieuwe functie. Drie voorbeelden zijn: (i) *lysozyme*, een eiwit in je speeksel dat helpt met het afbreken van suikers, (ii) *trombine*, een eiwit dat ervoor zorgt dat je bloed stolt als je een wond hebt en (iii) *antilichamen*, eiwitten die je lichaam helpen om ziekteverwekkers, zoals het coronavirus, te herkennen.



Figuur S1. Eiwitten bestaan uit kettingen van aminozuren.

Eiwit modificatie is een techniek waarbij een verandering gemaakt wordt aan een eiwit, veelal door er een ander molecuul aan vast te maken. Hoofdstuk 1 behandelt reeds bestaande technieken voor eiwit modificatie en de vele toepassingen ervan. In microbiologie worden bijvoorbeeld lichtgevende moleculen aan eiwitten vastgemaakt, zodat deze kunnen worden bestudeerd in cellen onder een microscoop. Een ander voorbeeld is in nieuwe genezingswijzen, waar antibiotica aan een antilichaam worden vastgemaakt, zodat deze ze naar een ziekteverwekker kan brengen. Eiwit modificatie kan op twee manieren worden gedaan. Enerzijds kan het *genetisch*, door aanpassingen te maken in de blauwdruk van het eiwit en te veranderen hoe het wordt gemaakt. Dit is een precieze methode, maar het duurt lang en werkt niet altijd. Anderzijds kan het *chemisch*, door moleculen aan het eiwit te koppelen nadat het eiwit al is gemaakt. Dit gaat sneller en werkt vaker, maar kan minder precies zijn als het niet voorzichtig wordt gedaan. In dit proefschrift kiezen we voor de chemische manier.



Figuur S2. Drie mogelijke vormen die DNA kan hebben.

Het innovatieve aan dit werk is het gebruik van kunstmatig DNA om de precisie van eiwit modificatie met chemicaliën te verbeteren. De reden dat we DNA gebruiken is, omdat we DNA kunnen ontwerpen zodat het een bepaalde vorm of functie aanneemt. De meest bekende vorm is wanneer twee DNA strengen om elkaar heen krullen (*Double Helix*), maar er zijn ook andere vormen mogelijk, zoals een kubus (*G-Quadruplex*) of een lus (*Hairpin*). De vorm van

het DNA kan een bepaald effect teweeg brengen, zoals het versnellen van een chemische reactie (*katalyse*) of het vasthouden van een molecuul (*aptameer*). Het zijn deze twee functies die in dit proefschrift worden gebruikt om eiwit modificatie te verbeteren.

In hoofdstuk 2 maken we gebruik van een DNA kubus die samen met een ijzerhoudend molecuul een katalysator vormt (*hGQ*). Deze hGQ gebruikt waterstof peroxide om een ander molecuul (*NML*) aan een eiwit te koppelen. Het type DNA kubus dat gekozen wordt is hierbij

van belang, omdat het van invloed is op hoe snel de reactie is en waar het molecuul op het eiwit wordt geplakt. Dit principe hebben we bewezen op verschillende eiwitten, waaronder de drie voorbeelden uit de eerste alinea. Daarnaast hebben we het DNA zo ontworpen dat we de reactie kunnen controleren door de kubus open te vouwen (stoppen) en terug dicht te vouwen (starten).

In hoofdstuk 3 gebruiken we dezelfde katalysator als in hoofdstuk 2, maar nu met in het verlengde daarvan een aptameer (*TamBA*) die een stukje eiwit (een peptide) kan vasthouden. In dit ontwerp houdt de aptameer het peptide op zijn plaats (zoals een hand een spijker) waarna de katalysator de reactie erop kan uitvoeren (zoals een hamer). Hierdoor verloopt de reactie in zijn geheel sneller en efficiënter dan in hoofdstuk 2.

In hoofdstuk 4 gebruiken we een DNA kubus die een aptameer (*TBA*) is en het eiwit trombine kan vasthouden. Aan TBA maken we een katalysator (*DMAp* of *PyOx*) vast, die alleen op korte afstand werkt. Alleen wanneer TBA zich vasthoudt aan trombine, is de katalysator dichtbij genoeg om moleculen aan trombine te koppelen. Een andere gevolg van dat de katalysator alleen op korte afstand werkt, is dat modificatie alleen plaats vindt in de buurt van waar TBA zich vasthoudt. Daardoor kunnen we met een tweede aptameer (*TBA2*) die zich ergens anders vasthoudt, trombine op een andere plaats modificeren. Als derde konden we deze DNA systemen zo ontwerpen dat we de aptameren kunnen vast- of losmaken op commando en op deze manier de reactie starten of stoppen wanneer wij dat willen.

In hoofdstuk 5 analyseren we de uiterste en optimale afstand waarop de katalysatoren uit hoofdstuk 2 en 4 werken. Hiervoor maken we een DNA streng vast aan een eiwit wat daarna een houvast is voor een tweede DNA streng (die dan samen een stabiele dubbele helix vormen). Door hiervoor zes dezelfde DNA strengen te gebruiken die elk hun katalysator een stukje verder hadden, konden we deze katalysatoren op een toenemende afstand plaatsen. Door daarna de efficiëntie van deze systemen te testen, konden we de uiterste en optimale afstand bepalen van de (drie eerder genoemde) katalysatoren. Deze resultaten hebben we daarna kunnen verklaren met behulp van een computer-berekend model van onze DNA systemen.

Hoofdstuk 6 bevat vooral extra opmerkingen en discussies over het werk uit de hoofdstukken 2 tot en met 5, zoals factoren die van invloed zijn op de beschreven reacties of voor- en nadelen van het gebruik van aptameren. Aan de hand hiervan worden daarna ook aanbevelingen voor toekomstig onderzoek gedaan, zoals diepgaander onderzoek naar het effect van de DNA kubus of het gebruik van alternatieve vormen van katalytisch DNA.

In hoofdstuk 7 wordt een wetenschappelijke samenvatting van dit proefschrift gegeven.

Summary for the non-scientist

Proteins are large molecules in nature that consist of 20 different types of building blocks (*amino acids*). Because proteins sometimes consist of more than 500 of these building blocks, many combinations are possible. Every combination of building blocks can lead to a new function. Three examples are: (i) *lysozyme*, a protein in your saliva that helps to degrade sugars, (ii) *thrombin*, a protein which causes blood to clot when you have an open wound and (iii) *antibodies*, proteins which help your body identify pathogens, such as the coronavirus.

Protein modification is a technique where a change is made to a protein, mostly by attaching another molecule to it. Chapter 1 discusses already existing techniques for protein modification and its many applications. In microbiology for example fluorescent molecules are attached to the proteins so they can be studied in cells under a microscope. Another application is in novel therapeutics, where antibiotics are attached to an antibody so it can bring them to a pathogen. Protein modification can be done in two ways. On one hand, it can be done *genetically*, by making alterations to the blue print of a protein and change the way it is created. Although this is a precise method, it takes time and does not always work. On the other hand it can be done *chemically*, by attaching molecules to the protein after it has been created. This is faster and works more often, but can also be less precise if it is done cautiously. In this thesis we choose the latter option.

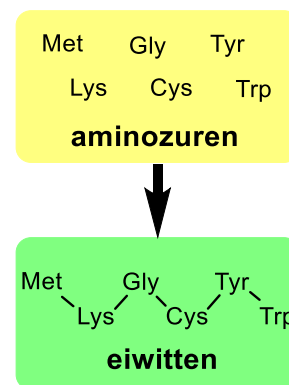


Figure S1. Proteins consist of chains of amino acids.

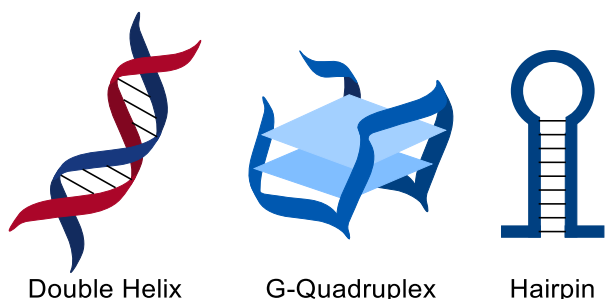


Figure S2. Three examples of shapes that DNA can adopt.

The innovative part of this work is the use of artificial DNA to improve the precision of the modification of proteins with chemicals. The reason we use DNA, is because we can design the DNA in such a way that it takes a particular shape or function. The most familiar shape is when two DNA strands curl around one another (*Double Helix*), but also other shapes are possible, such as a cube (*G-Quadruplex*) or a loop (*Hairpin*). The shape of the DNA can provide a certain effect, such as speeding up a reaction (*catalysis*) or to grab hold of another molecule (*aptamer*). It is these two functions that are used in this thesis to improve protein modification.

In chapter 2 we make use of a DNA cube that together with an iron-containing molecule forms a catalyst (*hGQ*). This hGQ uses hydrogen peroxide to attach another molecule (*NML*) to a protein. The type of DNA cube that is chosen is hereby of importance, because it influences the speed of the reaction and where the molecule is attached to the protein. We proved this principle on various proteins, among others the three examples from the first paragraph. Additionally, we designed the DNA in such a way that we could control the reaction by unfolding (stopping) and refolding (starting) the cube.

In chapter 3 we use the same catalyst as in chapter 2, but now extended with an aptamer (*TamBA*) which can hold on to a small piece of a protein (a peptide). In this design, the aptamer holds the peptide into position (like a hand holds a nail) after which that the catalyst can perform the reaction on it (like a hammer). As a result, the reaction proceeds faster and more efficient than in chapter 2.

In chapter 4 we use a DNA cube that is an aptamer (*TBA*) and can hold onto the protein thrombin. To TBA, we attach a catalyst (*DMAP* or *PyOx*) that only works at short range. Only when the aptamer is holding on to thrombin is the catalyst close enough to attach molecules to thrombin. Another consequence of the catalyst only working at short range, is the modification takes place solely in a circle around the position where the aptamer is holding on. Because of this, we could use a second aptamer (*TBA2*) that holds onto somewhere else, to modify thrombin in a different position. Thirdly, we could design these DNA systems in such a way that we could command the aptamers to hold on or let go and thus start or stop the reaction at will.

In chapter 5 we analyse the maximal and optimal distance at which the catalysts of chapter 2 and 4 work. To this end, we tether a DNA strand to a protein, which could provide a toehold for a second DNA strand (which together form a stable double helix). By using six times the same DNA strand of which each had their catalyst a wee bit further, we could fixate these catalysts at an increasing distance. By then testing the efficiency of these systems, we could determine the utmost and optimal distance of the (three earlier mentioned) catalysts. Afterwards, we could explain these results by using a computer-calculated model of our DNA systems.

Chapter 6 features mainly additional remarks and discussions regarding the work from chapters 2 to 5, such as factors that are of influence on the described reactions or the pros and cons of the usage of aptamers. Following these discussions, we then give recommendations for future research, such as more in depth research into the effect of the DNA cube or the use alternative forms of catalytic DNA.

In chapter 7 a scientific summary of this thesis is provided.

Chapter 1



General introduction

Parts are adapted from:

Chem. Rev. **2021**, 121 (12), 7032–7058

and

Biopolymers **2021**, 113 (3), 1–8

1.1 Abstract

Protein modification is an important asset for many research fields, most notably for chemical biology, medicinal chemistry, and biochemistry. However, optimal and reliable results are only obtained if the modification is performed efficiently and with high precision. Therefore, many different methods have been established to achieve this, ranging from reagents optimized to modify a singular amino acid residue, incorporation of genetically encoded unique reaction handles, or usage of highly specific catalysts such as enzymes, metal complexes or organic moieties. This thesis describes how the current protein modification toolbox is expanded by developing methods that use the versatility and adaptability of synthetic DNA to label native proteins in a site-selective and trigger-responsive fashion.

1.2 Protein modification

1.2.1 Proteins

Biological processes in organisms are often regulated by proteins. Blueprints for these proteins are stored in the DNA of an organism, which is temporarily copied in the form of mRNA that is used for protein synthesis by assembling amino acid building blocks.¹ These amino acids can be divided into classes based on several criteria; their division here is based on the reactive functionality of their side chain (Figure 1.1).^{1,2} On top of this basic level of functionalities, proteins are often decorated with other functional groups by a process known as post-translational modification (PTM). These PTMs can be as small as hydroxylation of proline, acetylation of lysine, or phosphorylation of serine or threonine, but can also be large, for example, the glycosylation of the Fc-domain of immunoglobulins or the ubiquitylation of proteins that have approached the end of their life-cycle. At the moment, >200 PTMs are known, all of which are performed by dedicated enzymes and substrates, and which are designed to tailor the ribosomally synthesized biopolymer for its biological role.^{1,3}

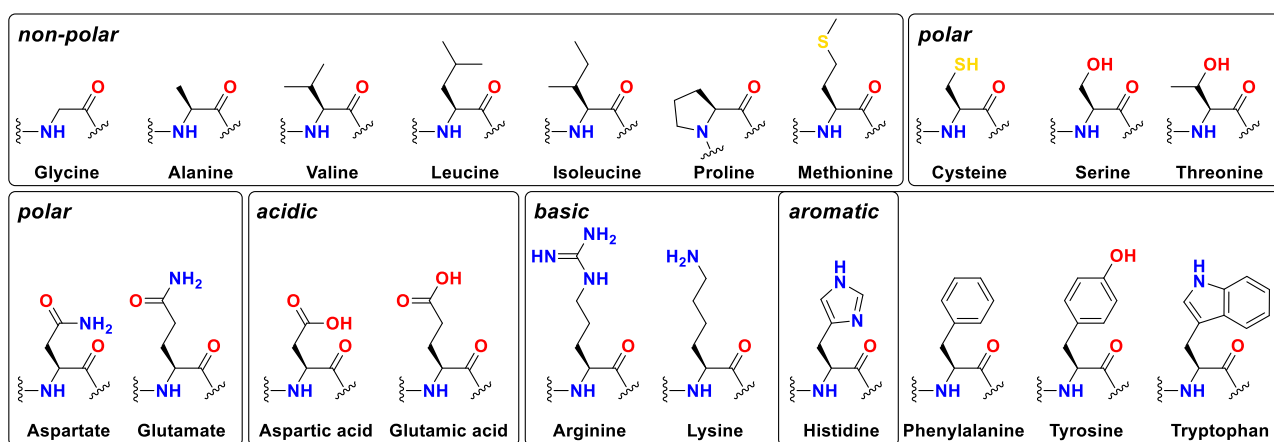


Figure 1.1. The structures of the 20 proteinogenic amino acid residues divided according to their side-chain functionality.^{1,2}

Whereas proteins are often optimally evolved to fit their designated biological purpose, this does not necessarily imply that those functions enable optimal use in our intended applications. In most cases, efficient use of the vast pool of functionalities that is accessible via proteins is only possible after modification of these biopolymers with artificial and often unnatural moieties.

1.2.2 Artificial protein modification

Artificial protein modification is a subcategory in the field of ‘bioconjugate chemistry’, which entails the covalent attachment of a foreign moiety to a biomolecule, for example, a nucleic acid, protein, glycan or lipid.⁴ When restricted to proteins, two pathways to obtain chemically modified functional proteins are available: (i) genetic alteration of the DNA blueprint of a protein so that it contains a uniquely reactive residue, and (ii) chemical alteration of a proteinogenic amino acid in a native protein.^{5–8}

Synthetic routes for the chemical modification of proteins help various biochemical research fields to visualize, track, guide or further functionalize proteins for their intended purposes (Figure 1.2). For example, fluorescent labels are attached to proteins to allow their visualisation within cells, which helps to understand intracellular protein functions.^{9–11} Similarly, protein modification is performed to anchor proteins such as antibodies to surfaces to create sensors^{12,13}, *e.g.*, for the SARS-CoV-2 viral protein.¹³ For therapeutic applications, proteins can be modified with drug-like entities, as is apparent from antibody-drug conjugates that are designed to increase the selectivity of toxins by guiding them to specific malignant cells and tissues.^{14–17}

To help these disciplines achieve their goals, protein modification needs to be as accurate as possible. For that, existing methods for modification of wild type protein are not sufficient.^{5–8,18} As proteins are typically composed of more than one copy of a specific amino acid, a lack of distinction between functional groups that have similar reactivities can lead to heterogeneously modified proteins. In these mixtures, different fractions carry the modification on different residues, whereas other fractions can contain different combinations of multiple modifications. For many of these members, the original structure or/and function of the protein

is/are compromised. Moreover, inseparable conjugates often display different biological activities, which not only makes it difficult to ascribe a measured effect to a specific conjugate, but also contains the risk that certain conjugates have undesired activities. Therefore, a modification is preferably installed on just one or a few residues, where the impact of modification on the biological activity of the construct is more reliably determined.^{5,6,8}

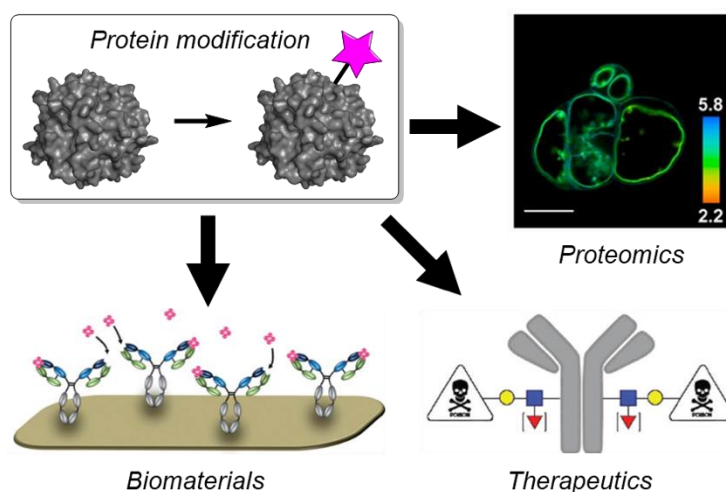


Figure 1.2. Schematic depiction of protein modification and how the products are applied to visualize biological structures in cells (top right),⁹ to attach antibodies on surfaces of sensors (bottom left),¹² or to generate therapeutics such as antibody-drug conjugates (bottom right).¹⁴

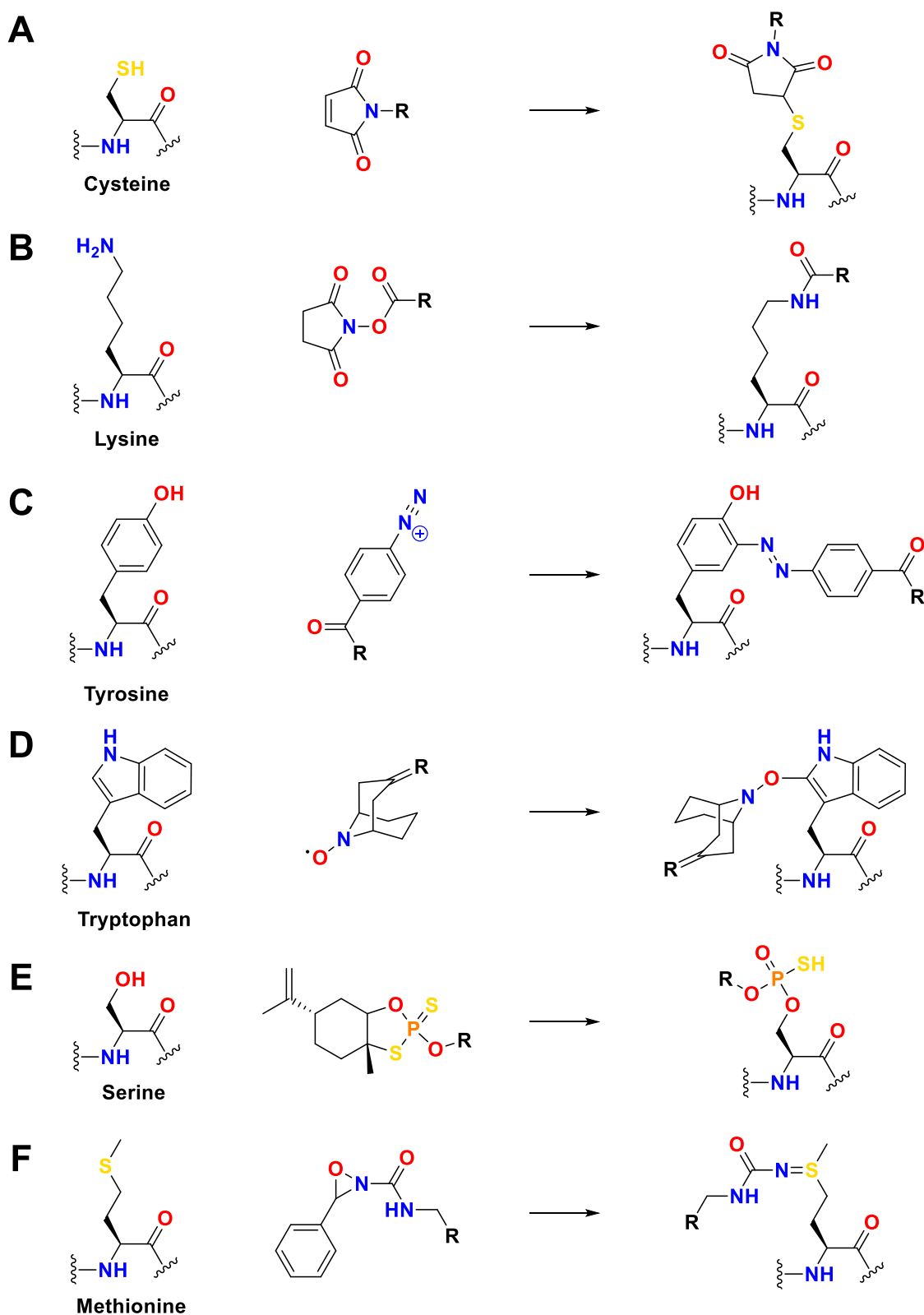
1.3 Selective artificial protein modification

Methods for selective protein modification have different levels of selectivity and are most commonly divided into three levels. The first level is to target one type of functional group in an amino acid. Such chemo-selectivity can be achieved by choosing a type of chemistry that preferably or exclusively reacts with the functional group of choice. The second level is to target one specific amino acid residue among multiple of the same type. This is called site-selective modification, and it is achieved with optimized reagents or catalysts, but is much harder as these methods rarely work equally effective when used for different protein targets. The third level of control is achieved by genetic incorporation of an unnatural amino acid residue into the protein sequence. This residue contains a unique functional group that is the only target of the used chemical reagent, achieving the highest level of selectivity.^{5–8,18,19}

1.3.1 Chemo-selective reagents

To acquire selectivity in protein modification, the first step is to use chemistry to target one type of amino acid. This is called chemo-selectivity and reagents have been developed to target most amino acids that have a functional group (Scheme 1.1).^{4–7,18} Perhaps the most common example of a chemical reagent for the modification of a proteinogenic amino acid residue is the maleimide group, which was found to react solely with the thiol functionality of cysteine (Cys) residues to form thioether bonds (Scheme 1.1A). Even though the resulting thioether is not very stable and suffers from a retro-Michael reaction, it is still one of the most-applied approaches to make protein conjugates.^{20–22} The *N*-hydroxy-succinimide ester on the other hand reacts with the primary amines of lysine (Lys) residues and the *N*-terminus, forming highly stable peptide bonds (Scheme 1.1B). For two aromatic amino acids, diazonium-aryl salts were discovered to modify the phenol of tyrosine (Tyr) residues (Scheme 1.1C),²³

which was recently joined by 9-azabicyclo[3.3.1]nonane-3-one-*N*-oxyl that selectively modifies the indole ring of tryptophan (Trp) residues (Scheme 1.1D).²⁴ Other recent designs are the limonene-derived oxathiophospholane to derivatize serine residues (Ser, Scheme 1.1E)²⁵ and the 3-phenyl-1,2-oxaziridine-2-carboximide to exclusively functionalize methionine residues (Met, Scheme 1.1F).²⁶



*Scheme 1.1. Selective sample of chemo-selective reagents that modify different amino acid residues: (A) maleimide reacts with sulfhydryl of Cys;²⁰⁻²² (B) *N*-hydroxy-succinimide ester reacts with amines of Lys and *N*-terminus; (C) diazonium-aryl salt reacts with phenol of Tyr;²³ (D) 9-azabicyclo[3.3.1]nonane-3-one-*N*-oxyl reacts with indole ring of Trp;²⁴ (E) oxathiophospholane reacts with primary alcohol of Ser;²⁵ (F) 3-phenyl-1,2-oxaziridine-2-carboximide reacts with thioether of Met.²⁶*

1.3.2 Site-selective protein modification with optimized reagents and conditions

Taking selectivity one step further is to design a reagent for site-selective protein modification. For this, various chemo-selective reagents have been optimized to target singular amino acids on a protein by making use of their so-called micro-environment. The micro-environment refers to local influences that enhance the reactivity of a nearby functional group, for example through steric or electronic effects. Sulfonyl acrylate (Figure 1.3A, *left*) was computer-engineered and optimized to target single Lys residues in several proteins.²⁷ Similarly, methylsulfone phenyloxadiazole (Figure 1.3A, *right*) was specifically designed for targeting Lys99 in humanized catalytic antibody h38C2, leading to homogeneous site-specifically modified proteins.²⁸ Although the results of this methodology can thus be outstanding, the major drawback lies in the laborious optimization required for each protein. Another example is the transamination reagent pyridoxal phosphate, which selectively activates *N*-terminal residues of proteins by forming an imine intermediate. If the *N*-terminal residue is a glutamine residue, decarboxylation can occur and a hydrogen is transferred to the pyridinium ring. Should this occur, transamination to attach an amine-functionalized modification and the pyridoxal phosphate is regenerated.²⁹

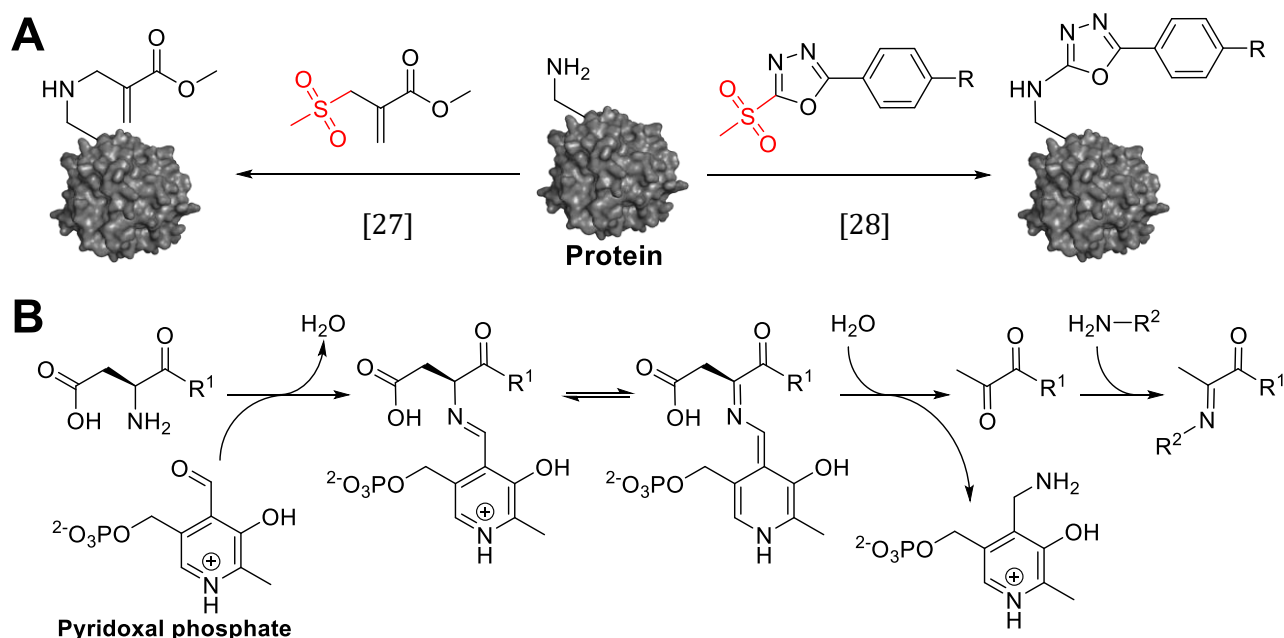
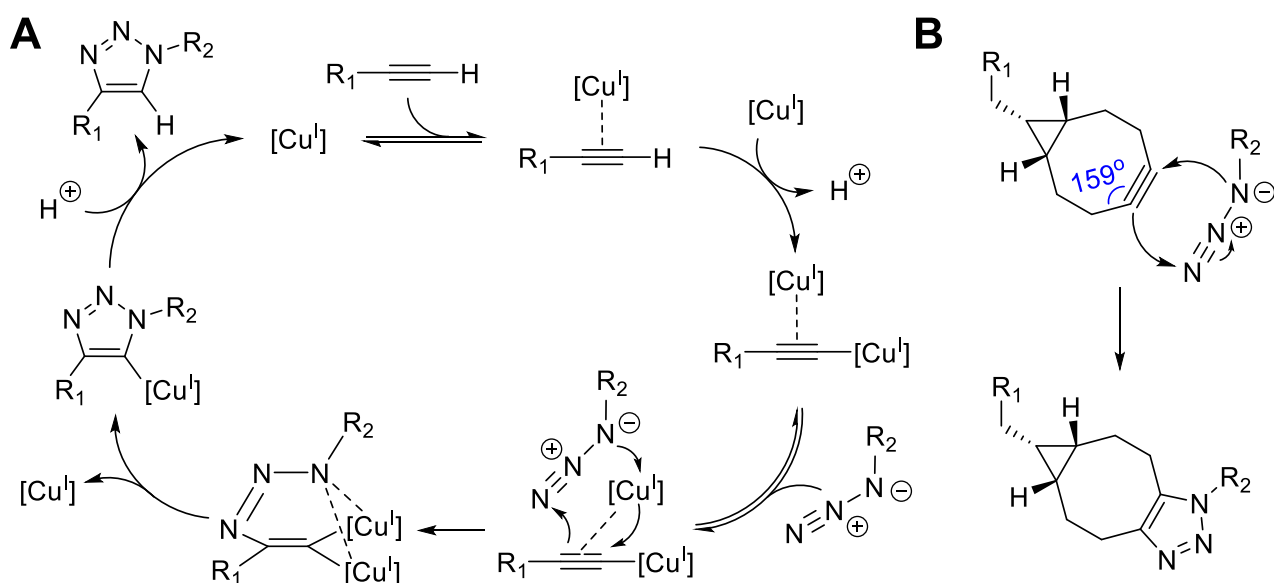


Figure 1.3. (A) Optimized reagents that target single Lys residues. The red part acts as leaving group upon nucleophilic attack by the ε-amine of the Lys residue.^{27,28} (B) Pyridoxal phosphate site-selectively transaminates *N*-terminal Gln residues on proteins.²⁹

Alternatively, unnatural reaction couples can be used in order to avoid interference with biological functionalities. Perhaps the best known examples of these so-called ‘bio-orthogonal’ reactions are the copper-catalysed alkyne–azide cycloaddition (CuAAC)³⁰ and the strain-promoted alkyne–azide cycloaddition (SPAAC) (Scheme 1.2).^{31,32} By introducing either of these reaction partners in a protein, clean and site-selective modification with the partner is almost guaranteed. Incorporation of such an unnatural amino acid can be done by the earlier mentioned genetic alteration with which an unnatural amino acid carrying the bio-orthogonal reaction handle is included at the desired site.³³ Although effective, this strategy suffers from (i) the laborious work to get a mutated protein expressed, (ii) the tolerance of the producing organism for the genetic alteration, as well as (iii) the unpredictability of the impact of the inserted residues on the protein’s structure and function.³³



Scheme 1.2. Reaction mechanisms of two widely applied click reactions between an azide and alkyne. **(A)** Copper-catalysed alkyne-azide cycloaddition (CuAAC), where $Cu(I)$ catalyses the coupling of R_1 with R_2 .³⁰ **(B)** Strain-promoted alkyne-azide cycloaddition (SPAAC), where the strain in BCN is the driving force resulting in the coupling of R_1 with R_2 .^{31,32}

1.3.3 Enzymatic approaches

As was mentioned, nature usually modifies proteins using enzymes. Some of these evolved to have a binding site that recognises one specific substrate and modify it in a chemo-selectively and site-specific manner. Naturally, methods have been set up to harness the high-efficiency enzymes for protein modification. For examples, sortase A is a Ca^{2+} -dependent transpeptidase enzyme that recognizes the amino acid motif Leu-Pro-(X)-Thr-Gly and cleaves the peptide bond between Thr and Gly (Figure 1.4A).³⁴ The subsequently formed acyl-sortase intermediate can be attacked by an amine to yield a new peptide bond. When this amine contains a synthetic molecule,³⁵ *e.g.*, a peptide³⁶ or polymer,³⁷ it is attached to the protein. Alternatively, mushroom tyrosinase can oxidize exposed Tyr residues to their corresponding *ortho*-quinone,³⁸ which can subsequently be coupled to (1*R*,8*S*,9*S*)-bicyclo[6.1.0]non-4-yn (BCN) by means of inverse electron-demand Diels-Alder cycloaddition (Figure 1.4B).¹⁴ Although initial studies relied on pre-engineered exposed Tyr residues, in the case of antibodies a naturally hidden Tyr residue could be exposed by removal of the glycan on the Fc domain with a second enzyme Glycanase.^{14,39} Alternatively to these examples where the enzyme activates the protein, inverse approaches are also described where the enzyme activates the substrate. Specifically, the enzyme horseradish peroxidase (HRP; oxidation potential 1.1 V)⁴⁰ has been used to activate *N*-methyl luminol (NML) derivatives using hydrogen peroxide and the activated NML derivatives couple to exposed Tyr residues on the protein surface.⁴¹ This method was improved by changing to the enzyme laccase (oxidation potential: 0.8–1.0 V¹⁰⁹), which generated higher conversions and used oxygen instead of hydrogen peroxide, thus presenting far less side-oxidation (Figure 1.4C).⁴²

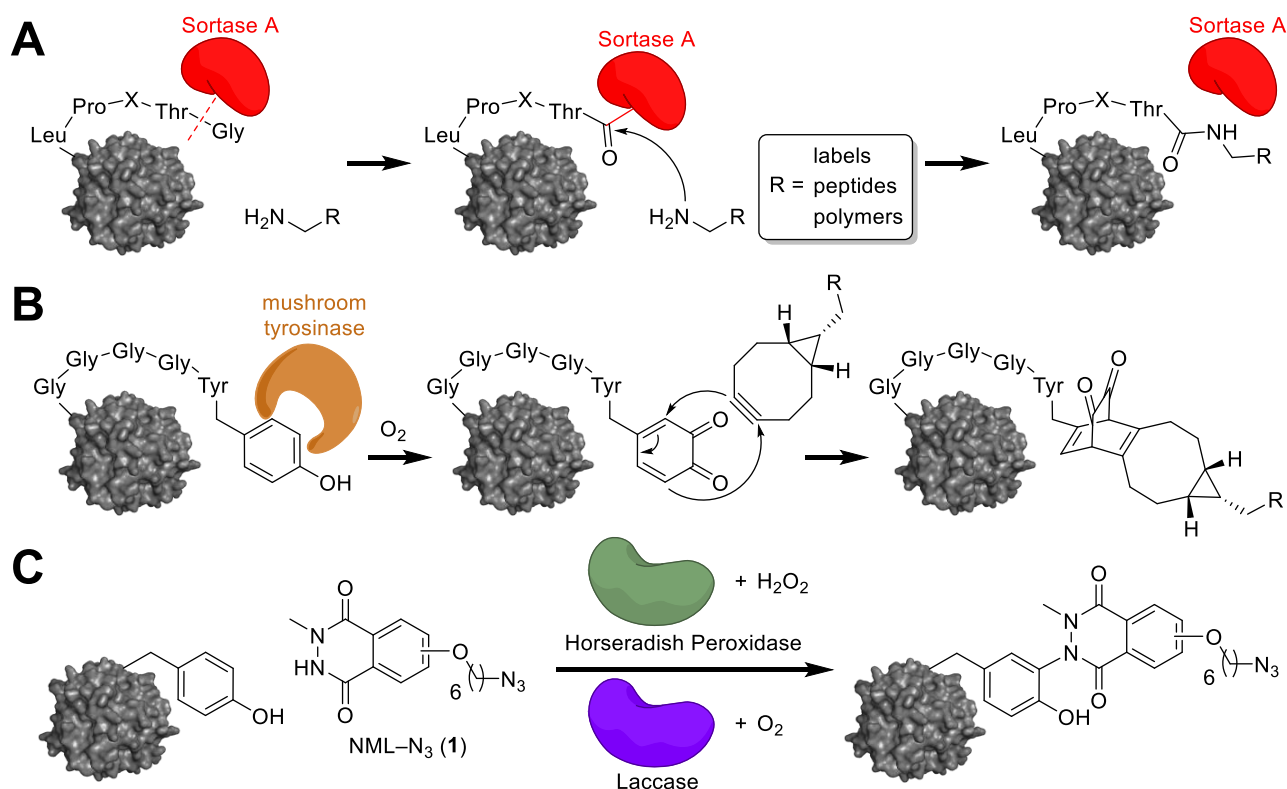


Figure 1.4. Examples of enzymatic protein modification. (A) Sortase A severs the peptide bond between Tyr and Gly in the Leu-Pro-(X)-Thr-Gly motif and attaches a new amine group.^{34–37} (B) Mushroom tyrosinase oxidizes exposed Tyr residues to quinones which can react further with the BCN moiety.¹⁴ (C) HRP or laccase can couple NML to Tyr residues by using hydrogen peroxide or oxygen, respectively.⁴²

1.3.4 Artificial inorganic catalytic approaches

The tremendous advantage of the high specificity of enzymatic protein modification is simultaneously their major drawback, as it poses a restriction on the organic substrates that can be used. In addition, enzymes are notoriously intolerant towards non-physiological conditions, which limits the conditions that can be used. As a result, methods based on artificial catalysis have been developed.

The well-known photocatalyst Ru(bipy)₃ complex can oxidize exposed Tyr residues to tyrosyl radicals when irradiated with light. The tyrosyl radical subsequently is scavenged by a tyrosyl radical trapping agents (RTA), such as *N*-(4-dimethylaniline)amide (Figure 1.5A). Importantly, control over the site of modification was obtained when the photocatalyst was bound to benzenesulfonamide; a ligand for carbonic anhydrase. As such, ligand-directed protein modification that activates a protein-bound functional group (*i.e.*, Tyr) that subsequently reacts with a RTA offered specific advantages over methods that activate a soluble label that has to diffuse to a reactive site on the protein in order to achieve oxidative cross-coupling.^{43,44}

Iron(III)-containing protoporphyrin IX, more commonly referred to as hemin, is the metal complex responsible for oxygen binding in the active sites of haemoglobin.⁴⁵ In the absence of haemoglobin and the presence of hydrogen peroxide, hemin acts as an oxidation catalyst (oxidation potential: 0.35 V)⁴⁶ that oxidizes a variety of substrates, including phenols and luminol derivatives.^{46–49} Using the *N*-methylated luminol derivative NML and hemin, proteins were modified (Figure 1.5B) in a similar fashion as was observed for HRP (Figure 1.4C). Whereas hemin was more tolerant towards changes in the conditions than HRP, substantial higher amounts of the catalyst were needed to achieve similar conversions.⁴⁶

Where these catalysts target Tyr residues, the peptide-bound dirhodium catalyst can alkylate Trp, His⁵⁰ and Asn⁵¹ residues via a carbene intermediate that is formed upon reacting with diazo compounds (Figure 1.5C). The catalyst is simply bound to the peptide through complexation of $\text{Rh}_2(\text{OAc})_3\text{TFA}$ with deprotonated Asp or Glu residues. By binding the catalyst to protein-affinity peptides, it was successfully used to generate alkylated Fyn kinase,⁵⁰ human IgG and trastuzumab.⁵¹ Although the dirhodium catalyst is not always chemo- and/or site-selective, different peptides appear able to target different residues,^{50,51} thus it is likely that such selectivities are attainable through design of the peptide.

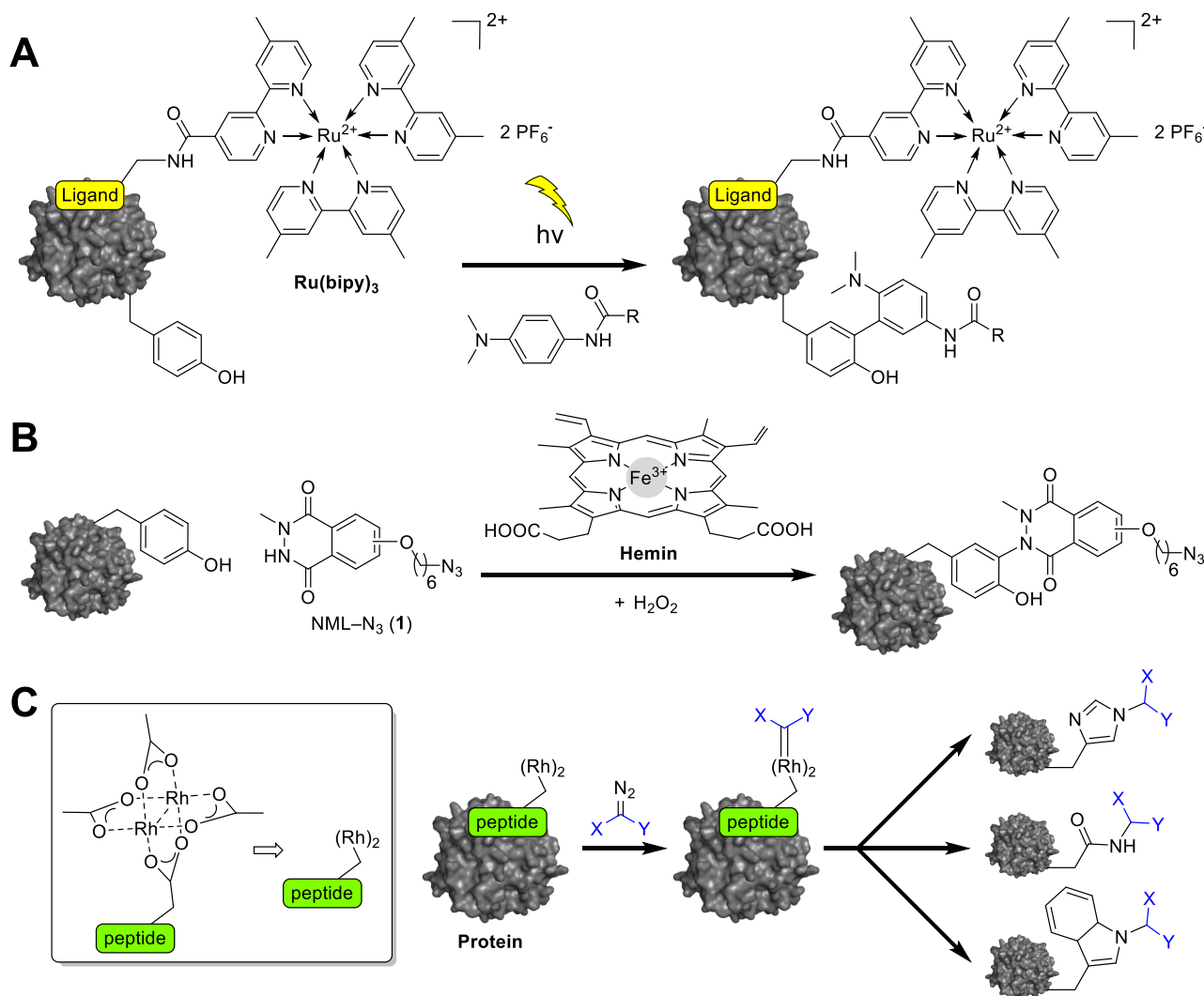


Figure 1.5. Two examples of catalytic modification of tyrosine (Tyr) residues. (A) Photocatalyst $\text{Ru}(\text{bipy})_3$ spawns radicals on Tyr residues, which can then be modified with tyrosyl trapping agents such as dimethylaniline.^{43,44} (B) Protoporphyrin IX can radically conjugate NML derivatives to Tyr residues.⁴⁶ (C) Rh_2 paddlewheel metalloptides were used to alkylate Trp, His or Asn residues via a carbene intermediate. The peptide was designed to have affinity for different proteins and to target different residues.^{50,51}

1.3.5 Artificial organocatalytic approaches

Apart from a few metal-based catalysts, organocatalytic approaches have also been developed.⁵² In an elegant approach, the group of Rai⁵³ used so-called linchpin catalysis in which a bifunctional probe reversibly binds Lys residues on one end and irreversibly modifies His residues on the other end of the tether (Figure 1.6A). Afterwards, unreacted probe is washed away and only probes that had the correct dimensions to facilitate binding to His residues remained on the protein.⁵³ Our last example is taken from the work of the group of Hamachi, which applied ligand-directed catalysis using the acyl transfer catalyst dimethylaminopyridine (DMAP) connected to lectin-binding saccharides. After binding of the DMAP-saccharide to the relevant lectin, the catalyst resides in close proximity to the protein surface, thereby enabling selectively acylation of proximal Lys residues using acyl donor molecules from solution (Figure 1.6B).⁵⁴

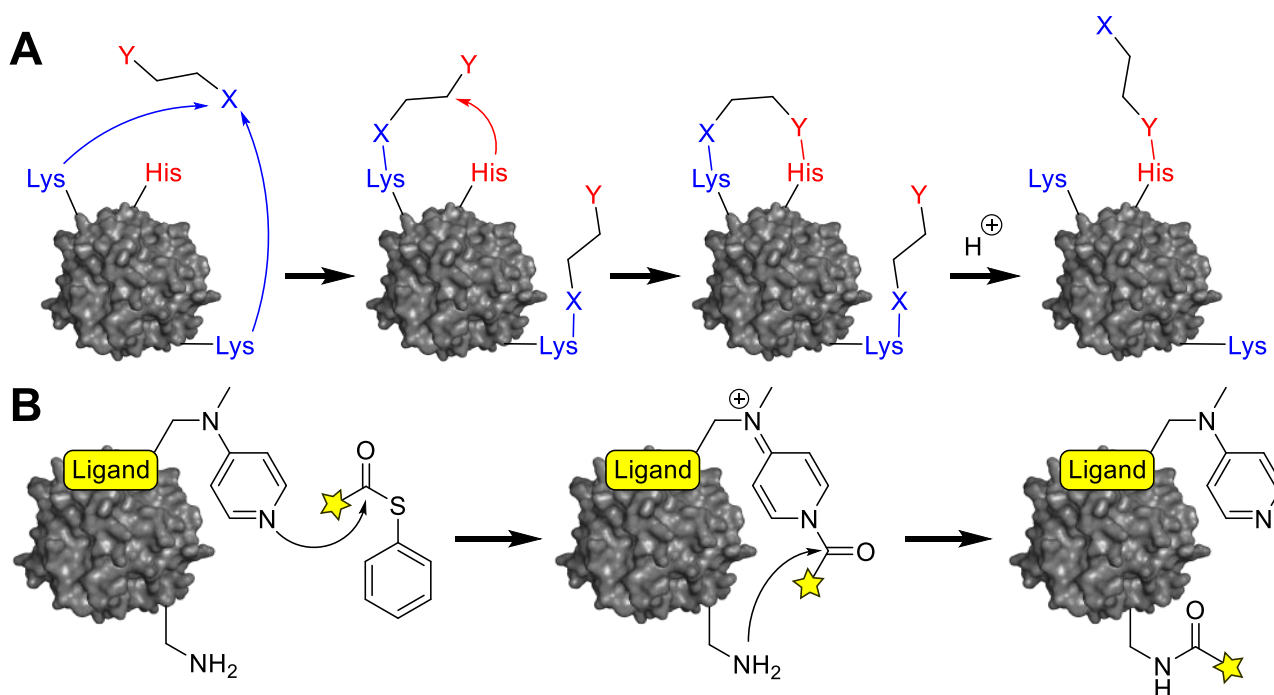


Figure 1.6. Two examples for the chemical derivatization of amino groups of proteins. (A) Linchpin catalysis, where Lys is bound reversibly and acts as a handle to conjugate nearby His residues after which acidic treatment reverses the Lys bonds.⁵³ (B) Ligand-directed DMAP catalyses the acylation of nearby Lys residues.⁵⁴

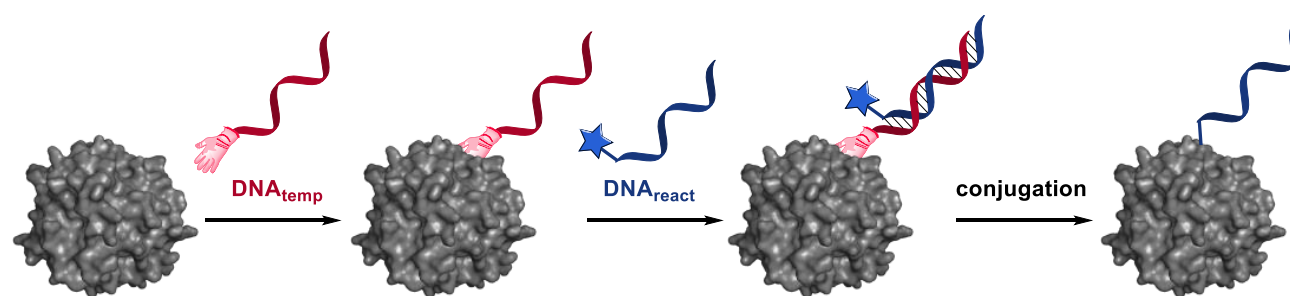
1.4 DNA-assisted site-selective protein modification techniques

1.4.1 DNA templated approach

In the past few years, different approaches where DNA has been used for protein modification appeared. The first reported use of DNA for protein modification was by the group of Li⁵⁵ and involved 'DNA templation'. In this strategy, various small molecules were used as recognition elements to guide DNA template strands (DNA_{temp}) to the surface of their protein binding partners (Figure 1.7). Afterwards, a second reacting DNA strand (DNA_{react}), carrying a photo-activatable diazirine moiety, was hybridized with the DNA_{temp}, positioning the reactive moiety in close proximity to the protein surface. Upon irradiation, crosslinking of DNA_{react} to the protein occurred, leading to a covalent protein-DNA conjugate. It was demonstrated that this approach was compatible with multiple different small molecules, even in the presence of competitive assay conditions, such as HeLa cell lysate. Later on, this method was used as a tool to identify protein targets of a DNA-encoded small molecule library.⁵⁶ After incubating the library with cell lysate conjugation was induced using UV light. Residual unbound DNA strands were digested by Exonuclease I, leaving only the undigested protein-dsDNA conjugates. Analysis of the remaining strands enabled identification of both the ligand and the target protein.

A comparable DNA-templated protein modification strategy was adopted by the group of Gothelf,^{57,58} which used metal-affinity probes to selectively modify His₆-tagged and metalloproteins (Figure 1.7). A DNA_{temp} strand conjugated with the known chelating agent tris(nitrilotriacetic acid) (NTA) could form a complex with the His₆-tag of various proteins in the presence of nickel(II) or copper(II) ions. A DNA_{react} strand functionalized with an *N*-hydroxysuccinimide (NHS) was then hybridized to the NTA-DNA_{temp}, enabling subsequent covalent coupling of the complementary strand to the protein target.⁵⁷ This method was not only applicable to His₆-tagged protein, but also on metalloproteins and even on an IgG antibody. Further analyses of the modified proteins revealed that conjugation occurred mainly in the vicinity of metal-binding sites, making this a site-selective conjugation method. In later work, the DNA-protein conjugate was used as an intermediate that could be oxidatively cleaved to leave an aldehyde on the protein surface. As such, the strategy uses DNA to site-selectively install aldehyde groups on proteins which could subsequently be used for oxime ligation.⁵⁸

A third strategy in the category of DNA-templated protein modification uses peptides to guide a hybridized reactive DNA strand to a specific protein (Figure 1.7).^{59,60} Using a DNA_{temp} that contained a trimethylated histone H3 peptide as a guiding moiety, proteins that detect histone modification could be selectively bound within complex protein mixtures.⁵⁹ Once bound a diazirine-bearing DNA_{react} was hybridized to the DNA_{temp} strand, after which irradiation yielded the desired conjugate that could be fished out and identified. The group of Gothelf⁶⁰ performed a similar strategy by attaching a DNA_{temp} to Fc-III, a cyclic peptide that is known to bind the Fc region of human IgG.⁶¹ After hybridization, nucleophilic attack from a lysine residue on the aldehyde of DNA_{react} led to the formation of an imine that could be reductively aminated. For the antibody rituximab, this resulted in 75% of DNA_{react} being selectively conjugated to its Fc region. Additionally, the obtained conjugate was assembled into pentameric IgG superstructures, using a star shape DNA nanostructure as a core, to synthetically mimic IgM antibodies.^{60,62}






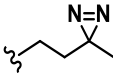
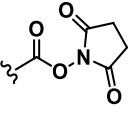

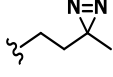

Guiding moiety 	Reactive group 	Protein 	Reference(s)
LPCBS		carbonic anhydrase II	[55,56]
AP1497		FKBP12	
desthiobiotin		Avidin	
antipain		trypsin	
tris(nitrilotriacetic acid) with Ni ²⁺ / Cu ²⁺ ions		Metalloproteins	[57]
		His ₆ -tagged proteins	[58]
trimethylated histone H3 peptide		histone modification reading proteins	[59]
Fc-III peptide		Fc region of IgG	[60]

Figure 1.7. DNA-templated protein modification, where the template strand (DNA_{temp}) is guided by different moieties (red hand). The reacting strand (DNA_{react}) hybridizes to the template and follow-up proximal conjugation results in site-selective attachment of DNA_{react} to the protein. Various guiding moieties have been adopted to target specific proteins or protein groups, using a set of reactive groups (blue star).

1.4.2 Using DNA as a ligand

Biological interactions between DNA and proteins was also exploited to modify proteins that have DNA as a substrate. A self-conjugating dsDNA probe was used by the group of Khodyreva⁶³ to label active DNA polymerase β (Figure 1.8A). A photosensitizer that was incorporated into the DNA polymerase β was triggered by irradiation. This in turn activated the 4-azido-2,3,5,6-tetrafluorophenyl group on DNA_{react} of the dsDNA probe, resulting in conjugation of the proximal DNA strand and trapping the probe in the DNA-polymerase active site. Even though the reported 50% conversions included unwanted products, the desired conjugates could be attained proving this conjugation approach. Similarly, the group of Li⁶⁴ used a half-dsDNA half-ssDNA probe that was designed for the targeting of dsDNA-binding proteins, including Nuclear Factor- κ - β (NF- κ B) (Figure 1.8B). The probe was then used for DNA-templated protein modification (described above) and a DNA_{react} bearing a diazirine moiety was hybridized to the probe which was, after photo-activation, selectively conjugated to the protein target. The system enabled conjugation of DNA to various transcription factors and showed selectivity for dsDNA binding proteins when used in cell lysate.

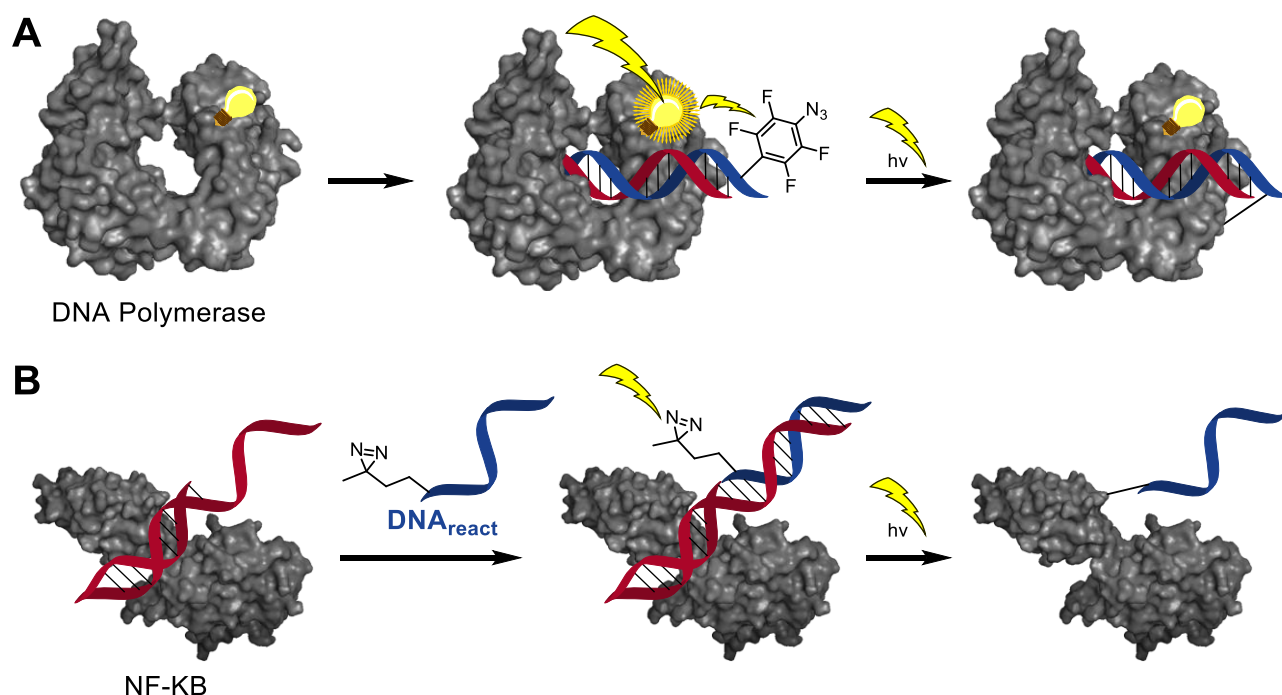


Figure 1.8. DNA-guided protein modification. (A) A photosensitizer was built into a DNA polymerase to trigger a photo-activatable group on a dsDNA probe. This results in the covalent trapping of the DNA probe in the active site.⁶³ (B) A dsDNA probe with a ssDNA extension is used to bind the dsDNA-binding protein NF-KB. After hybridization with a diazirine-bearing DNA_{react} strand, photo-activation results in covalent attachment of DNA_{react} to NF-KB.⁶⁴

1.4.3 Direct conjugation with DNA aptamers

Although the previous works already present DNA-guided protein modification, they are limited to DNA-binding proteins and are not always protein-specific. Alternatively, DNA-guided modification that relied on the affinity of certain aptamers for proteins was employed. Aptamers are oligonucleotide sequences – DNA, RNA, synthetic or hybrid – that bind non-covalently with high affinity and selectivity to a variety of targets, ranging from small molecules and metal ions, to large proteins and even cells.⁶⁵ This versatility enabled aptamers to become widely used and serve as selective tools to benefit research varying from proteomics studies to therapeutic applications.^{66,67} Aptamers are discovered in a high-throughput methodology called SELEX, where a library of DNA sequences is incubated with a target (*i.e.*, the ligand), washing away non-binding sequences and cloning the binding ones by means of PCR. After multiple cycles, isolation of the remaining sequences can lead to the discovery of one or more aptamers.^{67,68}

Even though aptamers can have high affinity for proteins, their binding mode not always favours efficient protein modification. To counter this, the group of Koch⁶⁹ incorporated the unnatural nucleic acid 5-bromo-2'-deoxyuridine into aptameric sequences for various proteins, among others human α -thrombin (Figure 1.9A). These reactive aptamers could bind their protein target and then be covalently linked upon irradiation. This approach particularly increased the capture yield of sequences with a low affinity. The group of Famulok⁷⁰ developed a strategy called Aptamer-Based Affinity Labelling (ABAL), in which DNA/RNA aptamers bearing reactive moieties were used to trap the protein to which the aptamers bind (Figure 1.9A). This approach worked for aptamers for hepatocyte growth factor receptor, IgE, and cytohesin-2, although conversions were moderate with 30%. Interestingly, however, the strategy was not only effective in complex protein mixtures, but also *in vivo* on non-small-cell lung carcinoma cells. Along similar lines, the group of Mayer⁷¹ employed the ABAL approach to modify thrombin using its aptamer TBA, which is known to bind exosite I of human alpha thrombin. TBA was

functionalized with a sulfo-*N*-hydroxysuccinimide on the 3'-end via a photocleavable tether and as such, TBA could almost fully inhibit the activity of the enzyme. Enzyme activity could be restored by irradiation of the formed conjugate, which allowed TBA to dissociate from thrombin and regain its activity.⁷¹ Importantly, when the aptamer was functionalized with an α,α -difluoromethyl carboxyl group to attach to amine functionalities, efficient conversion was obtained when the carboxyl group was attached on a T₈-linker as shorter tethers hampered modification.⁷²

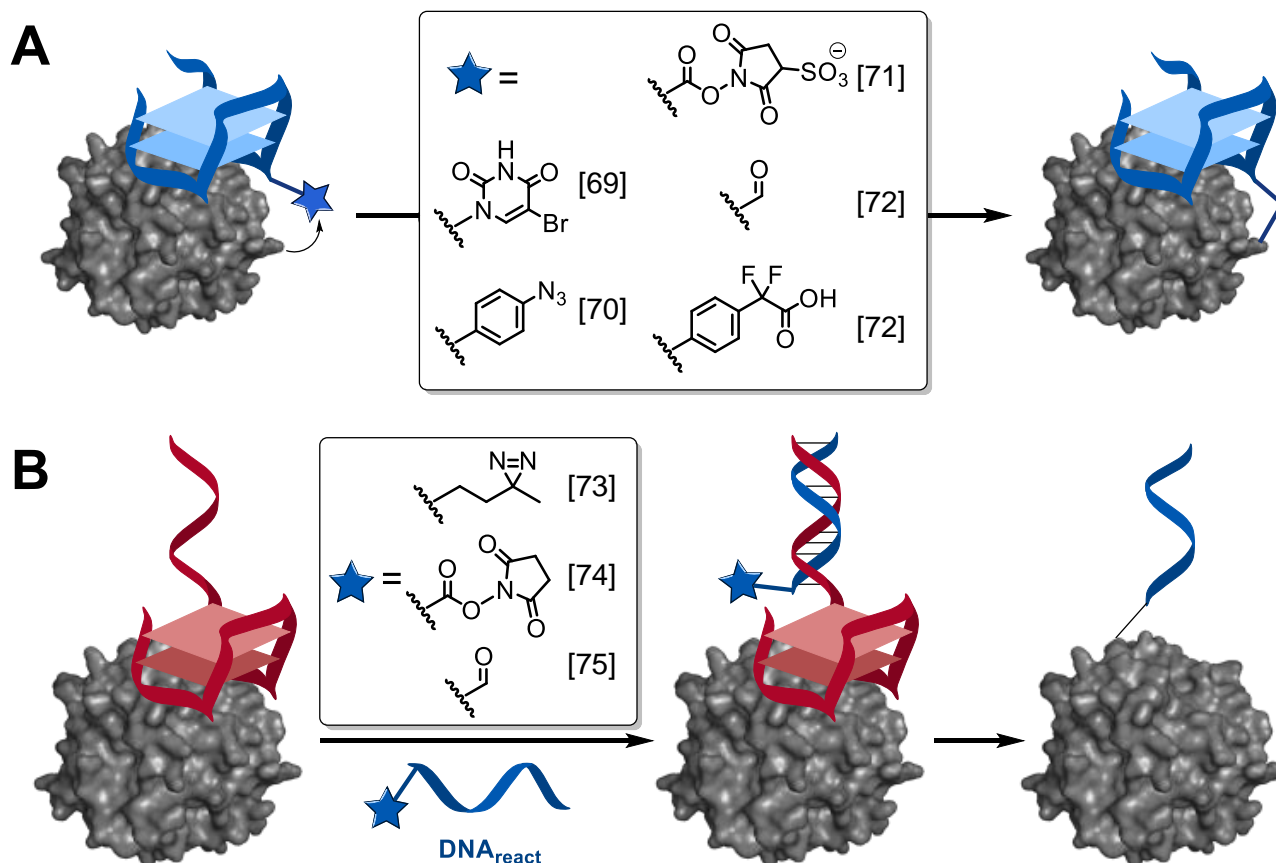









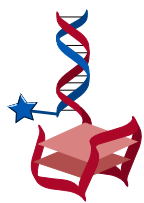
Figure 1.9. (A) Aptamer Based Affinity Labelling (ABAL) uses DNA aptamers that contain reactive groups to self-conjugate after finding their target. Panel shows the reactive moieties used for ABAL and their references. (B) DNA aptamers used for DNA-templated protein modification in which $\text{DNA}_{\text{react}}$ contains a photo-activatable diazirine⁷³, electrophilic NHS ester⁷⁴ or an aldehyde.⁷⁵

1.4.4 DNA templated approach with DNA aptamers

In order to reach beyond the aptameric sequence, aptamers have been used in a DNA-templated format as well. Work by the group of Zhang described how an aptamer for lysozyme C could be extended with a template DNA sequence that was functionalized with a diazirine moiety (Figure 1.9B).⁷³ This DNA-templated ABAL strategy proved efficient in a competitive assay against BSA, in the presence of HeLa cell lysate, and even in raw chicken egg white with hardly any off-target conjugation. Similarly, the group of Tan⁷⁴ placed a template DNA strand on TBA to which a $\text{DNA}_{\text{react}}$ bearing an electrophilic NHS ester was hybridized, resulting in thrombin-DNA conjugates with 85% conversion and 56% isolated yields. Tryptic digestion analyses showed that the site modification was limited to only two lysine residues, indicating high site-selectivity of their method. Comparable results were obtained with aptamer TBA2 (a.k.a. HD22), another aptamer for thrombin that binds to exosite II. This approach was also applied on the platelet-derived growth factor, streptavidin, and human IgG.⁷⁴ This last class of proteins was also the target of the group of Gothelf.⁷⁵ Optimized strategies for the application of ABAL and DNA-templated ABAL on IgG antibodies were designed, resulting in conjugates of the therapeutic antibodies

trastuzumab, rituximab, and cetuximab, with yields around 60% for the DNA-templated ABAL. When performing ABAL, it was revealed that this approach was more efficient, producing up to 90% conversion with just one equivalent of aptamer. Unfortunately, full site-selectivity could not be achieved as conjugation to the light chain of the Fab domain (which is not part of the Fc domain) was observed in both the direct and DNA-templated ABAL.⁷⁵

Table 1.1. Overview of the DNA-assisted protein modification strategies mentioned in section 1.4 along with their conversions, advantages, disadvantages and appropriate references.

Method	Details	Conv	Advantages	Disadvantages	Ref.
DNA- templation 	 small molecule	0.1–2*	<ul style="list-style-type: none"> • Small guiding unit • Highly specific • Many ligands known 	<ul style="list-style-type: none"> • Limited to available ligands • Increase of K_D by DNA attachment 	[55,56]
	 metal- affinity	25–60%	<ul style="list-style-type: none"> • Multi-applicable on any proteins with His₆-tag 	<ul style="list-style-type: none"> • Metal binding site required • Metal ion required 	[57]
	 peptide	50–100%	<ul style="list-style-type: none"> • Many protein-binding peptides • Variation in attachment point 	<ul style="list-style-type: none"> • Bulky peptide groups • Poorly defined peptide-protein interaction 	[58–60]
DNA substrate 	dsDNA probe	±50%	<ul style="list-style-type: none"> • Selective for dsDNA binding proteins • Simple binding probe 	<ul style="list-style-type: none"> • Requires intricate engineering of functionalized dsDNA • Case-specific strand length optimization • Limited to dsDNA binding proteins 	[63]
	templated dsDNA probe	0.1–2*			[64]
DNA aptamer as guiding unit 		20–100%	<ul style="list-style-type: none"> • Protein specific • No additional binding unit • Enables weak-binding aptamers 	<ul style="list-style-type: none"> • Only self-conjugate • Not many aptamers known • Poorly defined aptamer-protein interaction 	[69–72,75]
		45–85%	<ul style="list-style-type: none"> • Protein specific • Various DNA strands possible 	<ul style="list-style-type: none"> • Only conjugate DNA • Not many aptamers known • Poorly defined peptide-protein interaction 	[73–75]

Note: * these yields are normalized with respect a positive control.

1.5 Controlling Catalytic Transformations using DNA

The aforementioned DNA-based protein modification techniques are all designed to generate protein-DNA conjugates, whereas the attachment of small molecules would be more suited for wider application. In this thesis we propose that this is not only possible, but by making use of recent developments in DNA nanotechnology, can do so in a controlled and selective fashion. If successful, this would provide tools for the precision modification of wild-type proteins with functionalities from fluorescent probes to drug-like molecules.

The foundation of DNA nanotechnology lies in the predictability of DNA. DNA consists of nucleobase-carrying ribose units that are connected via phosphate linkages, where the four types of nucleobases form the two binding pairs: adenine (A) pairs with thymine (T) and cytosine (C) pairs with guanine (G).¹ The reliable formation of these pairs enables the design of complex structures that are composed often of multiple DNA strands.⁷⁶ Over the years, many types of nanostructures have been demonstrated, ranging from 2D figures to nanosized boxes (Figure 1.7A,B).^{77,78} On top of this, DNA has also been programmed to be able to switch between two or more different conformations (Figure 1.7C). These so-called DNA switches are induced using external triggers, which can be stabilizing cations (K^+ , Pb^{2+} , Mg^{2+}), (de)protonation as a result of pH, built-in molecular motors or other DNA strands.⁷⁹ As a result, the well-known and predictable dimensions of DNA nanostructures have also been utilized to increase our understanding of biological phenomena. For example, the group of Seitz used DNA as a molecular ruler to determine the spatial parameters of the tandem SH2 domain of Syk kinase.⁸⁰

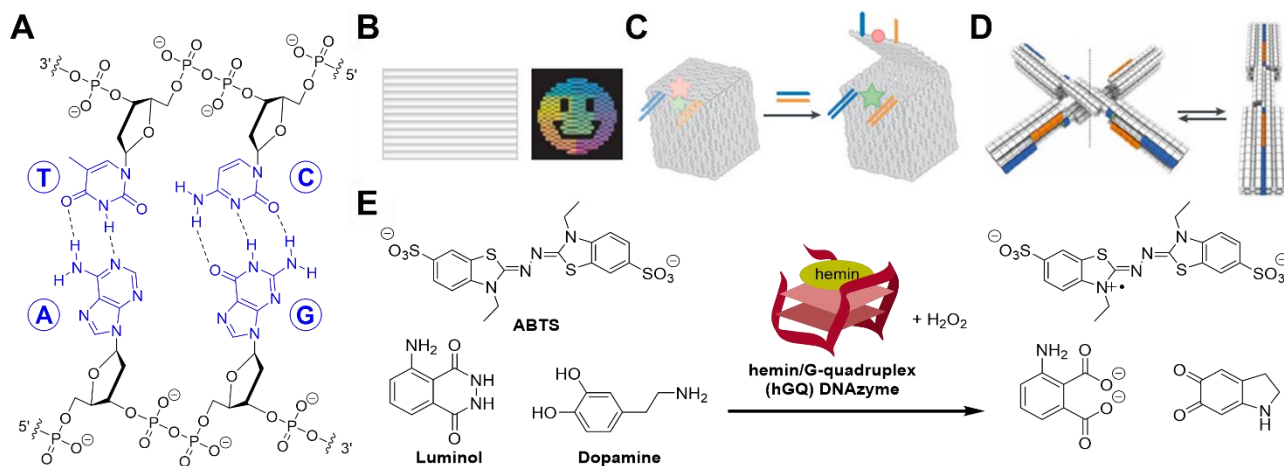


Figure 1.10. Synthetic DNA has been used to construct a variety of shapes, responsive structures, and catalytic entities. (A) Molecular structure of dsDNA with the nucleobase pairs in blue. (B-D) Images taken from Dey et al.⁷⁷ showing examples of DNA origami, such as (B) the 2D planar smiley face, (C) the box with dynamic lid and (D) the nanoswitch that can change between two conformations; (E) hGQ DNAzyme and its aromatic substrates that after oxidation have unique absorptions, useful for sensors.^{81,82}

Additional advantages of DNA are its relatively stability, even at higher temperatures^{83,84} and the commercial availability of custom-made oligonucleotides containing additional functionalities such as fluorophores or click handles (alkynes, azides).⁸⁵ These features make DNA a readily-accessible and stable biopolymer that enables the design of complex supramolecular structures through its programmability.^{77,78} It is not surprising that DNA structures have also been reported to exert catalytic properties in the presence of a suitable 'cofactor', such as RNA/DNA hydrolysis and Friedel-Craft alkylation with metal ions.⁸⁶⁻⁸⁸ Perhaps the best known example is the horseradish peroxidase mimicking hemin/G-Quadruplex DNAzyme, which can oxidize aromatic substrates, such as ABTS, luminol and dopamine, which after oxidation have unique absorbances (Figure 1.7D).^{81,82} It is therefore not surprising that DNA architectures have been used to develop catalytic nanostructures of which the activities arise from the interplay between the active units. For example, the group of Willner developed DNA-based nucleozymes, in which the catalytic features of a hGQ DNAzyme were substantially

enhance by conjugation to an aptamer for the substrate dopamine.^{81,82} Other chemical conversions have also been assisted by DNA. Specifically, the group of Roelfes developed achiral catalytic metal complexes that formed chiral Friedel-Crafts and Diels-Alder products in the presence of DNA.⁸⁹ In a different approach, the group of Silverman developed aptamer-constructs that were able to phosphorylate tyrosine residues in a complementary DNA-peptide construct.^{90,91} Clearly, the accessibility, programmability and programmability of DNA-associated catalytic events could help the design catalytic constructs that can raise the level of our current abilities to modify proteins.

1.6 Outline of this thesis

DNA has already shown great potential as guiding agent for site-selective and efficient protein modification, although aforementioned techniques (section 1.5 and Table 1.1) only conjugated the DNA strands themselves. In this thesis, we explored the use of DNA for nanostructured catalysts for site-selective modification of native proteins with small molecules by merging various aspects of enzymatic protein modification in one artificial DNA-based approach.

In [chapter 2](#), we harnessed the catalytic DNA system hGQ that uses hydrogen peroxide to oxidatively activate *N*-methyl luminol (NML) molecules and conjugates them to the tyrosine residues on proteins. By changing the structure of the DNA system, we altered the activity and selectivity of this catalyst and even designed it to include an activity-control switch.

In [chapter 3](#), we continued with the catalytic system described in chapter 2 and included an aptameric sequence for tyrosinamide in the hGQ system. This study showed how the aptameric sequence could enhance the rate of catalytic modification. The aptamer also helped to increase the modification efficiency of larger tyrosinamide-containing biomolecules, *i.e.*, peptides.

In [chapter 4](#), we took protein-binding aptamers and coupled them to acylation catalysts DMAP or PyOx in order to selectively acylate their target protein, also in the presence of other proteins. We focused on thrombin, and used two aptamers that bind thrombin on opposite sides. Additionally, design of the DNA system allowed again the incorporation of an activity-control switch.

In [chapter 5](#), we eliminated the dynamic binding factor of the aptamers by synthesizing covalently bound protein-DNA conjugates to which we hybridized a second catalyst-functionalized DNA strand. With this, we assessed the optimal and maximum range at which DMAP, PyOx and hGQ can perform protein modification.

[Chapter 6](#) covers an integrated discussion regarding the work of the aforementioned chapters, followed by recommendations for future research.

Lastly, [chapter 7](#) gives a brief summary of chapters 1 to 6 of this thesis.

Publication online:

Parts of the text of this chapter are published in:



Chem. Rev. **2021**, 121 (12), 7032–7058



Biopolymers **2021**, 113 (3), 1–8

References

- 1 B. Alberts, A. Johnson, J. Lewis, D. Morgan, M. Raff, K. Roberts and P. Walter, *The Cell*, **2014**, vol. 6.
- 2 J. J. Clayden, N. Greeves, S. Warren and P. Wothers, *Organic Chemistry*, **2001**, vol. 40.
- 3 A. L. Lehninger, D. L. Nelson and M. M. Cox, *Lehninger Principles of Biochemistry*, W.H. Freeman & Co Ltd, fifth edit., **2008**.
- 4 J. Kalia and R. T. Raines, *Curr Org Chem.* **2010**, 14 (2), 138–147.
- 5 O. Boutureira and G. J. L. Bernardes, *Chem. Rev.* **2015**, 115 (5), 2174–2195.
- 6 N. C. Reddy, M. Kumar, R. Molla and V. Rai, *Org. Biomol. Chem.* **2020**, 18 (25), 4669–4691.
- 7 N. Stephanopoulos and M. B. Francis, *Nat. Chem. Biol.* **2011**, 7 (12), 876–884.
- 8 E. A. Hoyt, P. M. S. D. Cal, B. L. Oliveira and G. J. L. Bernardes, *Nat. Rev. Chem.* **2019**, 3 (3), 147–171.
- 9 L. Michels, V. Gorelova, Y. Harnvanichvech, J. W. Borst, B. Albada, D. Weijers and J. Sprakel, *Proc. Natl. Acad. Sci. U. S. A.* **2020**, 117 (30), 18110–18118.
- 10 M. A. Kasper, M. Glanz, A. Stengl, M. Penkert, S. Klenk, T. Sauer, D. Schumacher, J. Helma, E. Krause, M. C. Cardoso, H. Leonhardt and C. P. R. Hackenberger, *Angew. Chem. Int. Ed.* **2019**, 58 (34), 11625–11630.
- 11 J. V Jun, C. M. Haney, R. J. Karpowicz, S. Giannakoulis, V. M. Y. Lee, E. J. Petersson and D. M. Chenoweth, *J. Am. Chem. Soc.* **2019**, 141 (5), 1893–1897.
- 12 A. K. Trilling, J. Beekwilder and H. Zuilhof, *Analyst* **2013**, 138 (6), 1619–1627.
- 13 R. L. Pinals, F. Ledesma, D. Yang, N. Navarro, S. Jeong, J. E. Pak, L. Kuo, Y. C. Chuang, Y. W. Cheng, H. Y. Sun and M. P. Landry, *Nano Lett.* **2021**, 21 (5), 2272–2280.
- 14 J. J. Bruins, D. Blanco-Ania, V. Van Der Doef, F. L. Van Delft and B. Albada, *Chem. Commun.* **2018**, 54 (53), 7338–7341.
- 15 T. Wu, M. Liu, H. Huang, Y. Sheng, H. Xiao and Y. Liu, *Chem. Commun.* **2020**, 56 (65), 9344–9347.

- 16 K. L. Wu, C. Yu, C. Lee, C. Zuo, Z. T. Ball and H. Xiao, *Bioconjugate Chem.* **2021**, 32 (9), 1947–1959.
- 17 A. Flemming, *Nat. Rev.* **2014**, 13, 178.
- 18 D. M. Valcourt, J. Harris, R. S. Riley, M. Dang, J. Wang and E. S. Day, *Nano Res.* **2018**, 11 (10), 4999–5016.
- 19 J. F. Keijzer and B. Albada, *Biopolymers* **2021**, 113 (3), 1–8.
- 20 R. Ohri, S. Bhakta, A. Fourie-O'Donohue, J. Dela Cruz-Chuh, S. P. Tsai, R. Cook, B. Wei, C. Ng, A. W. Wong, A. B. Bos, F. Farahi, J. Bhakta, T. H. Pillow, H. Raab, R. Vandlen, P. Polakis, Y. Liu, H. Erickson, J. R. Junutula and K. R. Kozak, *Bioconjugate Chem.* **2018**, 29 (2), 473–485.
- 21 D. M. Chigoho, Q. Lecocq, R. M. Awad, K. Breckpot, N. Devoogdt, M. Keyaerts, V. Caveliers, C. Xavier and J. Bridoux, *Pharmaceuticals* **2021**, 14 (6), 550.
- 22 E. W. McConnell, A. L. Smythers and L. M. Hicks, *J. Am. Soc. Mass Spectrom.* **2020**, 31 (8), 1697–1705.
- 23 J. M. Hooker, A. P. Esser-Kahn and M. B. Francis, *J. Am. Chem. Soc.* **2006**, 128 (49), 15558–15559.
- 24 Y. Seki, T. Ishiyama, D. Sasaki, J. Abe, Y. Sohma, K. Oisaki and M. Kanai, *J. Am. Chem. Soc.* **2016**, 138 (34), 10798–10801.
- 25 J. C. Vantourout, S. R. Adusumalli, K. W. Knouse, D. T. Flood, A. Ramirez, N. M. Padial, A. Istrate, K. Maziarz, J. N. Degruyter, R. R. Merchant, J. X. Qiao, M. A. Schmidt, M. J. Deery, M. D. Eastgate, P. E. Dawson, G. J. L. Bernardes and P. S. Baran, *J. Am. Chem. Soc.* **2020**, 142 (41), 17236–17242.
- 26 S. Lin, X. Yang, S. Jia, A. M. Weeks, M. Hornsby, P. S. Lee, R. V. Nichiporuk, A. T. Iavarone, J. A. Wells, F. D. Toste and C. J. Chang, *Science* **2017**, 355 (6325), 597–602.
- 27 M. J. Matos, B. L. Oliveira, N. Martínez-Sáez, A. Guerreiro, P. M. S. D. Cal, J. Bertoldo, M. Maneiro, E. Perkins, J. Howard, M. J. Deery, J. M. Chalker, F. Corzana, G. Jiménez-Osés and G. J. L. Bernardes, *J. Am. Chem. Soc.* **2018**, 140 (11), 4004–4017.
- 28 D. Hwang, K. Tsuji, H. Park, T. R. Burke and C. Rader, *Bioconjugate Chem.* **2019**, 30 (11), 2889–2896.
- 29 J. M. Gilmore, R. A. Scheck, A. P. Esser-Kahn, N. S. Joshi and M. B. Francis, *Angew. Chem. Int. Ed.* **2006**, 45 (32), 5307–5311.
- 30 V. Hong, S. Presolski, C. Ma and M. G. Finn, *Angew. Chem. Int. Ed.* **2009**, 48 (452), 9879–9883.
- 31 J. Dommerholt, S. Schmidt, R. Temming, L. J. A. Hendriks, F. P. J. T. Rutjes, J. C. M. Van Hest, D. J. Lefeber, P. Friedl and F. L. Van Delft, *Angew. Chem. Int. Ed.* **2010**, 49 (49), 9422–9425.
- 32 C. G. Gordon, J. L. MacKey, J. C. Jewett, E. M. Sletten, K. N. Houk and C. R. Bertozzi, *J. Am. Chem. Soc.* **2012**, 134 (22), 9199–9208.
- 33 K. Lang and J. W. Chin, *Chem. Rev.* **2014**, 114 (9), 4764–4806.
- 34 H. Mao, S. A. Hart, A. Schink and B. A. Pollok, *J. Am. Chem. Soc.* **2004**, 126 (9), 2670–2671.
- 35 M. Rashidian, L. Wang, J. G. Edens, J. T. Jacobsen, I. Hossain, Q. Wang, G. D. Vitoria, N. Vasdev, H. Ploegh and S. H. Liang, *Angew. Chem. Int. Ed.* **2016**, 128 (2), 538–543.
- 36 M. D. Witte, J. J. Cragnolini, S. K. Dougan, N. C. Yoder, M. W. Popp and H. L. Ploegh, *Proc. Natl. Acad. Sci. U. S. A.* **2012**, 109 (30), 11993–11998.
- 37 J. Hu, G. Wang, W. Zhao, X. Liu, L. Zhang and W. Gao, *Biomaterials* **2016**, 96, 84–92.
- 38 S. Osaki, *Arch. Biochem. Biophys.* **1963**, 100, 378–384.
- 39 J. J. Bruins, A. H. Westphal, B. Albada, K. Wagner, L. Bartels, H. Spits, W. J. H. Van Berkel and F. L. Van Delft, *Bioconjugate Chem.* **2017**, 28 (4), 1189–1193.

- 40 P. J. Kersten, B. Kalyanaraman, K. E. Hammel, B. Reinhammar and T. K. Kirk, *Biochem. J.* **1990**, 268 (2), 475–480.
- 41 S. Sato, K. Nakamura and H. Nakamura, *ChemBioChem* **2017**, 18 (5), 475–478.
- 42 S. Sato, K. Nakane and H. Nakamura, *Org. Biomol. Chem.* **2020**, 18 (19), 3664–3668.
- 43 S. Sato and H. Nakamura, *Angew. Chem. Int. Ed.* **2013**, 125 (33), 8843–8846.
- 44 S. Sato, K. Hatano, M. Tsushima and H. Nakamura, *Chem. Commun.* **2018**, 54 (46), 5871–5874.
- 45 A. Arnone, *Annu. Rev. Med.* **1974**, 25, 123–130.
- 46 S. Sato, K. Nakamura and H. Nakamura, *ACS Chem. Biol.* **2015**, 10 (11), 2633–2640.
- 47 S. Baj and T. Krawczyk, *J. Photochem. Photobiol. A Chem.* **2006**, 183 (1–2), 111–120.
- 48 S. Usha Devi and T. Ramasarma, *Mol. Cell. Biochem.* **1987**, 77 (2), 111–120.
- 49 H. Shalit, A. Libman and D. Pappo, *J. Am. Chem. Soc.* **2017**, 139 (38), 13404–13413.
- 50 F. Vohidov, J. M. Coughlin and Z. T. Ball, *Angew. Chem. Int. Ed.* **2015**, 127 (15), 4670–4674.
- 51 J. Ohata and Z. T. Ball, *J. Am. Chem. Soc.* **2017**, 139 (36), 12617–12622.
- 52 K. S. Egorova and V. P. Ananikov, *Organometallics* **2017**, 36 (21), 4071–4090.
- 53 S. R. Adusumalli, D. G. Rawale, U. Singh, P. Tripathi, R. Paul, N. Kalra, R. K. Mishra, S. Shukla and V. Rai, *J. Am. Chem. Soc.* **2018**, 140 (44), 15114–15123.
- 54 Y. Koshi, E. Nakata, M. Miyagawa, S. Tsukiji, T. Ogawa and I. Hamachi, *J. Am. Chem. Soc.* **2008**, 130 (1), 245–251.
- 55 G. Li, Y. Liu, Y. Liu, L. Chen, S. Wu, Y. Liu and X. Li, *Angew. Chem. Int. Ed.* **2013**, 125 (36), 9723–9728.
- 56 P. Zhao, Z. Chen, Y. Li, D. Sun, Y. Gao, Y. Huang and X. Li, *Angew. Chem. Int. Ed.* **2014**, 53 (38), 10056–10059.
- 57 C. B. Rosen, A. L. B. Kodal, J. S. Nielsen, D. H. Schaffert, C. Scavenius, A. H. Okholm, N. V. Voigt, J. J. Enghild, J. Kjems, T. Tørring and K. V. Gothelf, *Nat. Chem.* **2014**, 6 (9), 804–809.
- 58 A. L. B. Kodal, C. B. Rosen, M. R. Mortensen, T. Tørring and K. V. Gothelf, *ChemBioChem* **2016**, 1338–1342.
- 59 X. Bai, C. Lu, J. Jin, S. Tian, Z. Guo, P. Chen, G. Zhai, S. Zheng, X. He, E. Fan, Y. Zhang and K. Zhang, *Angew. Chem. Int. Ed.* **2016**, 128 (28), 8125–8129.
- 60 T. B. Nielsen, R. P. Thomsen, M. R. Mortensen, J. Kjems, P. F. Nielsen, T. E. Nielsen, A. L. B. Kodal, E. Cló and K. V. Gothelf, *Angew. Chem. Int. Ed.* **2019**, 58 (27), 9068–9072.
- 61 K. Sasaki, Y. Miyashita, D. Asai, D. Funamoto, K. Sato, Y. Yamaguchi, Y. Mishima, T. Iino, S. Takaishi, J. Nagano, A. Kishimura, T. Mori and Y. Katayama, *MedChemComm* **2018**, 9 (5), 783–788.
- 62 C. Zhang, M. Su, Y. He, X. Zhao, P. A. Fang, A. E. Ribbe, W. Jiang and C. Mao, *Proc. Natl. Acad. Sci. U. S. A.* **2008**, 105 (31), 10665–10669.
- 63 S. V. Dezhurov, I. R. Grin, I. V. Safronov, G. V. Shishkin, O. I. Lavrik and S. N. Khodyreva, *Russ. Chem. Bull.* **2005**, 54 (5), 1311–1321.
- 64 Y. Liu, W. Zheng, W. Zhang, N. Chen, Y. Liu, L. Chen, X. Zhou, X. Chen, H. Zheng and X. Li, *Chem. Sci.* **2015**, 6 (1), 745–751.
- 65 S. Klusmann, *The Aptamer Handbook: Functional Oligonucleotides and Their Applications*, **2006**.
- 66 P. R. Bouchard, R. M. Hutabarat and K. M. Thompson, *Annu. Rev. Pharmacol. Toxicol.* **2010**, 50,

237–257.

- 67 P. Röthlisberger and M. Hollenstein, *Adv. Drug Deliv. Rev.* **2018**, *134*, 3–21.
- 68 R. Stoltenburg, C. Reinemann and B. Strehlitz, *Biomol. Eng.* **2007**, *24* (4), 381–403.
- 69 D. Smith, B. D. Collins, J. Heil and T. H. Koch, *Mol. Cell. Proteomics* **2003**, *2* (1), 11–18.
- 70 J. L. Vinkenburg, G. Mayer and M. Famulok, *Angew. Chem. Int. Ed.* **2012**, *51* (36), 9176–9180.
- 71 F. Rohrbach, F. Schäfer, M. A. H. Fichte, F. Pfeiffer, J. Müller, B. Pötzsch, A. Heckel and G. Mayer, *Angew. Chem. Int. Ed.* **2013**, *52* (45), 11912–11915.
- 72 R. Wang, D. Lu, H. Bai, C. Jin, G. Yan, M. Ye, L. Qiu, R. Chang, C. Cui, H. Liang and W. Tan, *Chem. Sci.* **2016**, *7* (3), 2157–2161.
- 73 W. Bi, X. Bai, F. Gao, C. Lu, Y. Wang, G. Zhai, S. Tian, E. Fan, Y. Zhang and K. Zhang, *Anal. Chem.* **2017**, *89* (7), 4071–4076.
- 74 C. Cui, H. Zhang, R. Wang, S. Cansiz, X. Pan, S. Wan, W. Hou, L. Li, M. Chen, Y. Liu, X. Chen, Q. Liu and W. Tan, *Angew. Chem. Int. Ed.* **2017**, *129* (39), 12116–12119.
- 75 M. B. Skovsgaard, M. R. Mortensen, J. Palmfeldt and K. V. Gothelf, *Bioconjugate Chem.* **2019**, *30* (8), 2127–2135.
- 76 Y. Yin and X. S. Zhao, *Acc. Chem. Res.* **2011**, *44* (11), 1172–1181.
- 77 S. Dey, C. Fan, K. V. Gothelf, J. Li, C. Lin, L. Liu, N. Liu, M. A. D. Nijenhuis, B. Saccà, F. C. Simmel, H. Yan and P. Zhan, *Nat. Rev. Methods Prim.* **2021**, *1* (1), 1–24.
- 78 Y. J. Chen, B. Groves, R. A. Muscat and G. Seelig, *Nat. Nanotechnol.* **2015**, *10* (9), 748–760.
- 79 F. Wang, X. Liu and I. Willner, *Angew. Chem. Int. Ed.*, **2015**, *54*, 1098–1129.
- 80 H. Eberhard, F. Diezmann and O. Seitz, *Angew. Chem. Int. Ed.* **2011**, *50* (18), 4146–4150.
- 81 H. B. Albada, E. Golub and I. Willner, *Chem. Sci.* **2016**, *7* (5), 3092–3101.
- 82 E. Golub, H. B. Albada, W. C. Liao, Y. Biniuri and I. Willner, *J. Am. Chem. Soc.* **2016**, *138* (1), 164–172.
- 83 H. Clausen-Schaumann, M. Rief, C. Tolksdorf and H. E. Gaub, *Biophys. J.* **2000**, *78* (4), 1997–2007.
- 84 R. D. Blake and S. G. Delcourt, *Nucleic Acids Res.* **1998**, *26* (14), 3323–3332.
- 85 F. X. Montserrat, A. Grandas, R. Eritja and E. Pedroso, *Tetrahedron* **1994**, *50* (8), 2617–2622.
- 86 M. Hollenstein, *Molecules*, **2015**, *20*, 20777–20804.
- 87 S. K. Silverman, *Acc. Chem. Res.* **2015**, *48* (5), 1369–1379.
- 88 L. Stefan, F. Denat and D. Monchaud, *Nucleic Acids Res.* **2012**, *40* (17), 8759–8772.
- 89 A. J. Boersma, R. P. Megens, B. L. Feringa and G. Roelfes, *Chem. Soc. Rev.* **2010**, *39* (6), 2083–2092.
- 90 S. M. Walsh, A. Sachdeva and S. K. Silverman, *J. Am. Chem. Soc.* **2013**, *135* (40), 14928–14931.
- 91 V. Dokukin and S. K. Silverman, *Chem. Commun.* **2014**, *50* (66), 9317–9320.

Chapter 2



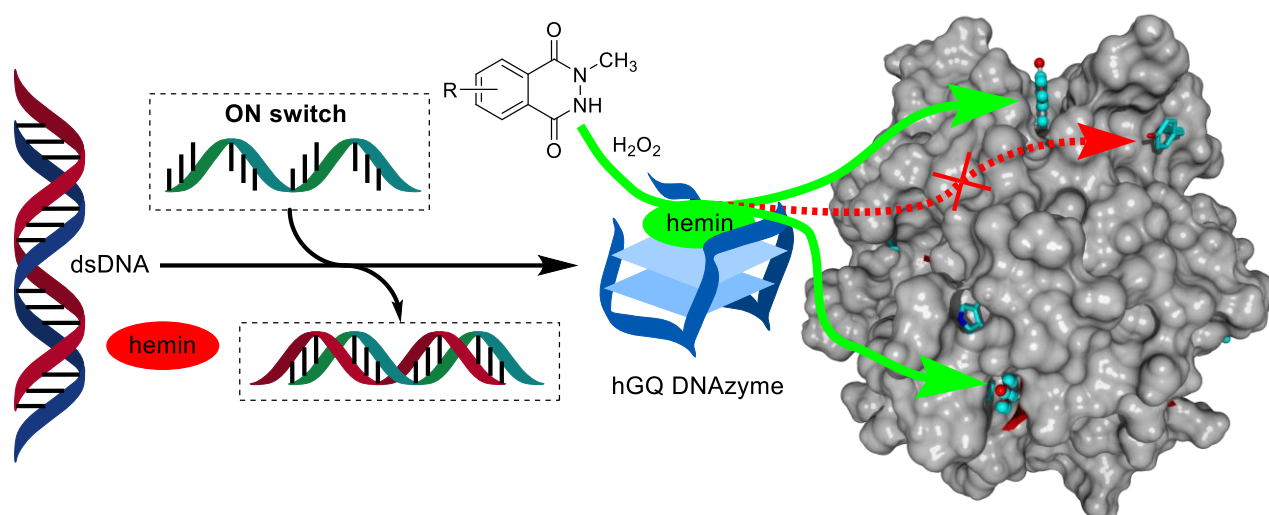
Site-Specific and Trigger-Activated Modification of Proteins by Means of Catalytic Hemin/G- quadruplex (hGQ) DNAzyme Nanostructures

Work by:

Jordi F. Keijzer and Bauke Albada

Adapted from:

Bioconjugate Chem. **2020**, 31 (10), 2283–2287



2.1 Abstract

Catalytic nanostructures offer a level of control over the positioning of involved functionalities that is required for the mimicry of enzymatic features. In this chapter, we show how the non-covalent complex between hemin and G-quadruplex DNA efficiently catalyses the modification of proteins with *N*-methyl luminol derivatives in the presence of hydrogen peroxide. Final conversions are reached within 15–30 min, and LC-MS analysis of tryptic digests of the proteins shows that the reaction proceeds with high chemo-selectivity for electron-rich aromatic residues (Tyr \gg Trp), and that the efficiency and site-selectivity of the modification depends on the sequence and secondary structure of the G-quadruplex nano-structure. Furthermore, the modification can be applied on proteins with different biomedical functions, and the nanostructure can be designed to contain a regulatory element in order to control protein modification by means of an external stimulus.

2.2 Introduction

The hemin/G-quadruplex (hGQ) complex is a DNA-based catalyst that can mimic oxidation reactions performed by peroxidase enzymes.¹⁻³ The hGQ complex forms a compact nanostructure by a G-rich DNA strand that forms one or more intermolecular guanine tetrads in the presence of a stabilizing cation (e.g., Na⁺, K⁺, NH₄⁺, Pb²⁺). When two of these G-tetrads are stacked, hemin can bind and form a catalytically active complex through π - π interactions. The catalytic activity of the resulting hGQ DNAzyme is significantly enhanced when compared to hemin alone,^{1,4} and depends on the sequence and secondary structure formed by the layers of guanine tetrads.^{5,6} Further enhancement of the catalytic activity (k_{cat}) can be achieved by nucleotide supplements⁷ or by conjugation of the hGQ DNAzyme to an aptamer sequence that binds to the substrate.^{8,9} Such so-called nucleozymes can be subjected to rational design¹⁰ or incorporation into supramolecular assemblies in order to increase their activity or tailor their properties.⁹ The predictable formation of the catalytically active hGQ DNAzyme nanostructure has led to its incorporation into complex oligonucleotide assemblies of which the activity depended on the application of an external trigger.¹¹ Apart from the oxidation of chemical substrates in sensor-type setups,^{2,3} some of the chemical conversions mimic biological processes, such as the oxidation of dopamine to aminochrome or *N*-hydroxy-L-arginine to nitric oxide and L-citrulline.⁸ The application of this interesting catalytic nanostructure for more demanding chemical transformations, *i.e.*, on larger and more complex (bio)molecules such as proteins, was hardly explored. In view of the variety of the hGQ DNAzyme nanostructures that can be formed depending on the sequence, and the ability to control the catalytic activity by neighbouring functionalities, we assessed whether hGQ DNAzymes could be used as potent catalysts for protein modification.

Protein modification is ideally performed in a rapid, efficient and site-specific manner.^{12,13} The latest methods for this apply bio-orthogonal click-reactions, which are superior in rate and bio-orthogonality when compared to classic reactions between a nucleophile and electrophile.^{14,15} Alternatively, synthetic catalysts¹⁶ or enzymes^{17,18} have been applied for the modification of native proteins or the conversion of (a) genetically encoded handle(s).¹⁹⁻²² When it comes to the application of biomimetic catalysts however, only a few methods have been reported, examples of which are pyridoxal-5-phosphate for the modification of protein N-termini,²³ or Ru(bipy)₃-induced modification of Tyr residues.²⁴

We set out to test if the hGQ DNAzyme nanostructure would transfer the differences in topology to variations in protein modification ability and selectivity and if the hGQ DNAzyme could be embedded in a nanostructure that responds to external triggers. For this, we used multiple hGQ DNAzymes based on various topologies²⁵ as catalysts for the oxidative modification of proteins using *N*-methyl luminol (NML)^{17,26} derivatives and hydrogen peroxide (H₂O₂) (Figure 2.1A-B). Having the ability to design the local environment of a protein-modifying catalyst would allow mimicry of features that give enzymes the ability to perform highly selective protein modification.²³ To test this, we used lysozyme, thrombin, bovine serum albumin (BSA), and the therapeutically relevant immunoglobulin trastuzumab as a representative set that covers a large range of protein sizes, from 14 kDa lysozyme to 150 kDa trastuzumab.

2.3 Results & Discussion

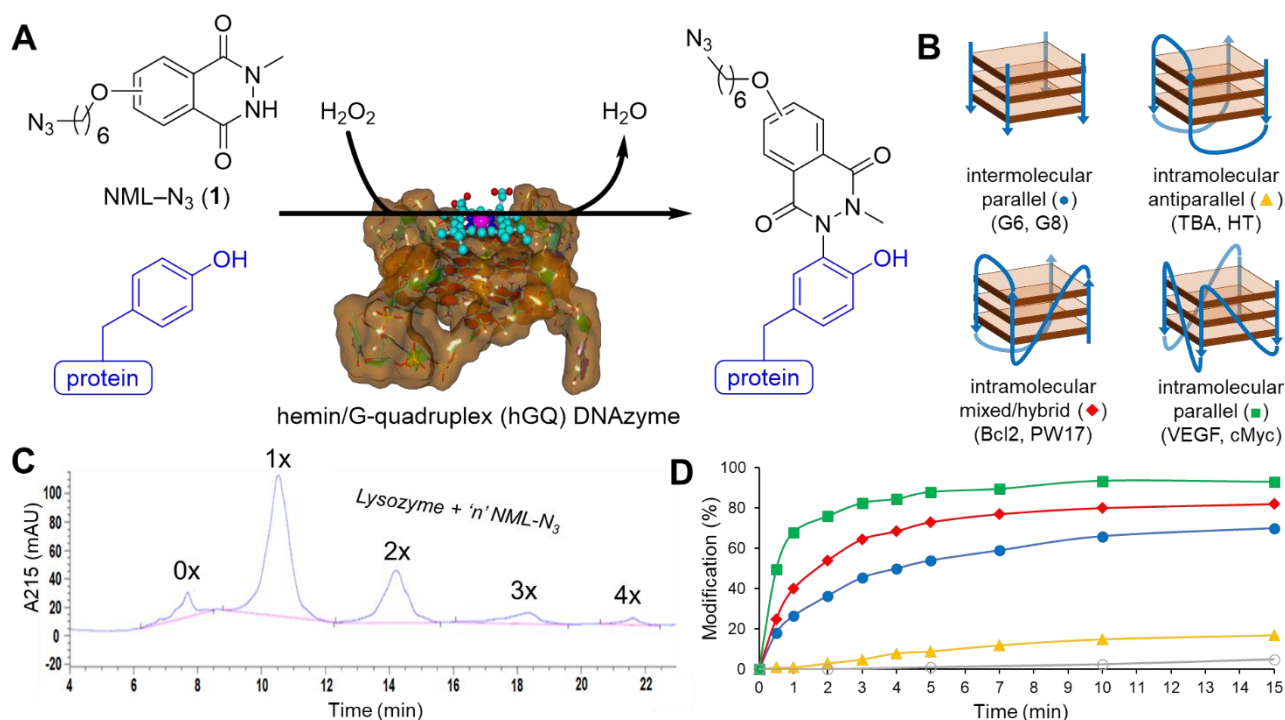


Figure 2.1. (A) hemin/G-quadruplex (hGQ) DNAzyme catalysed modification of a protein-based tyrosine residue with *N*-methyl luminol derivative (NML) **1** in the presence of H_2O_2 ; (B) G-quadruplex topologies used to construct the DNAzymes in this study (a general depiction of three G tetrads is shown for all systems, but the amount can differ per sequence); (C) HPLC-UV(A280) trace of lysozyme modified with NML **1**, each peak representing 'n' NML **1** molecules conjugated to the protein; (D) Graph of lysozyme conjugation with NML **1** over time by the different hemin/G-quadruplex systems. The colour codes are for ● intermolecular, ▲ for anti-parallel, ◆ for mixed type, ■ for antiparallel and ○ for non GQ-forming. Conditions: 10 μM hemin, 10 μM DNA, 140 μM lysozyme, 700 μM NML **1**, and 2800 μM H_2O_2 , pH: 7.0, at 25 °C for 30 min.

Lysozyme (14.3 kDa) contains three tyrosine (Tyr) residues with different solvent accessible area (SAA): Tyr23 (SAA: 32%), Tyr20 (SAA: 28%), Tyr53 (SAA: 15%).²⁷ The presence of potentially competing aromatic amino acid residues tryptophan (Trp28 (SAA: 0%), Trp62 (SAA: 50%), Trp63 (SAA: 28%), Trp108 (SAA: 2%), Trp111 (SAA: 7%), Trp123 (SAA: 27%), phenylalanine and histidine allowed assessment of the chemo-selectivity of the reaction. Regular protein modification experiments started with the incubation of a DNA strand, hemin, NML **1** and protein in phosphate buffer (400 mM NaCl, 10 mM KCl, pH 7.0) for 10 min at 25 °C. The modification reaction was initiated by addition of H_2O_2 and performed at 25 °C for 30 minutes, after which it was terminated by quenching with catalase. The percentage of modified lysozyme could subsequently be calculated by separation with HPLC(-MS) and integration of the product peaks as detected by UV (at 215 nm) (Figure 2.1C and Figure A2.1).

Much to our delight, LC-MS analysis of the reaction mixtures revealed substantial levels (32–96%) of lysozyme modification by various hGQ DNAzyme nanostructures (Table 2.1). Whereas only 2% of lysozyme was modified by hemin in the absence of DNA or in the presence of unstructured ssDNA or dsDNA, the presence of G-quadruplex forming sequences led to substantially higher amounts of modified protein. Interestingly, notable differences were observed for different G-quadruplex topologies: intramolecular parallel GQs formed the most active complexes, followed by the intramolecular mixed type GQs and intermolecular parallel GQs, and with the intramolecular anti-parallel GQs generating the least active hGQ DNAzymes. Time-resolved HPLC-analysis of the reaction mixtures using a representative hGQ catalyst from each topology revealed that modification was nearly complete after 15 minutes (Figure 2.1D).

Table 2.1. The names of the DNA sequences of the different DNazymes, the reported folding conformation of their G-quadruplex, and the percentages of un-, mono-, di-, tri-, and tetra-modified lysozyme produced after 30 min reaction in the dark at 20 °C. Sequences in bold were subjected to more detailed studies.

DNA sequence*	GQ folding topology	total % mod.	% 1× mod.	% 2× mod.	% 3-4× mod.	% 0× mod.
-	-	2	2			98
ssDNA	no GQ structure	2	2			98
dsDNA	no GQ structure	1	1			99
G5	intermolecular	20	18	2		80
G6	intermolecular	36	36			64
G7	intermolecular	51	48	3		49
G8	intermolecular	78	69	8	1	22
G9	intermolecular	86	67	17	2	14
G10	intermolecular	82	70	12		18
HT	antiparallel	32	32			68
TBA	antiparallel	35	35			75
TBA2	antiparallel	33	33			77
TA	antiparallel	40	40			50
TA2	antiparallel	46	46			54
Oxy28	antiparallel	44	44			56
PW17	mixed type	74	65	8	1	26
rPS2M	mixed type	62	59	3		38
Bcl2	mixed type	72	62	10		28
EA2	parallel	91	70	18	3	9
cMyc	parallel	96	63	27	6	4
EAD	parallel	93	66	22	5	7
EAD2	parallel	77	68	8	1	23
EAD3	parallel	94	68	22	4	6
EAD4	parallel	93	69	20	4	7
EAD6	parallel	96	61	26	9	4
VEGF	parallel	54	49	5		46
cKit21	parallel	81	70	10	1	19
HIF-1α	parallel	82	67	13	2	18
RET	parallel	69	62	6	1	31
LBA1-HT	antiparallel	17	17	-	-	83
LBA1-PW17	mixed type	91	73	16	2	9
LBA1-cMyc	parallel	96	66	26	4	4

*Conditions: 10 μM hemin, 10 μM DNA, 140 μM lysozyme, 700 μM NML (**1**), and 2800 μM H₂O₂.

*The DNA sequences can be found in Table A2. 1. The sequences in bold were used for the site-selectivity studies.

We also analysed the influence of the concentration(s) of the different components. As could be expected, increasing the concentration of the DNzyme led to a sharp increase in the modification (Figure 2.2A). Similarly, increasing the concentration of NML **1** led to a similar result, although the increment was less steep than for the DNzyme concentration (Figure 2.2B). As can be seen from the data, increasing H₂O₂ concentrations is beneficial up to 1000 μM, after which this effect disappears (Figure 2.2C). This effect even inverts above 3000 μM, possibly due to rapid catalyst degradation, which has been shown for hGQ DNzymes.^{1,4} Interestingly, closer examination of this data revealed that the concentration of H₂O₂ is of greater consequence when the concentration of NML **1** is lower (Figure 2.2D). When comparing different concentrations of NML with identical equivalents of H₂O₂, we can conclude that lower concentrations of NML can be compensated by higher concentrations of H₂O₂ to obtain similar modification levels, an effect also observed the other way around.

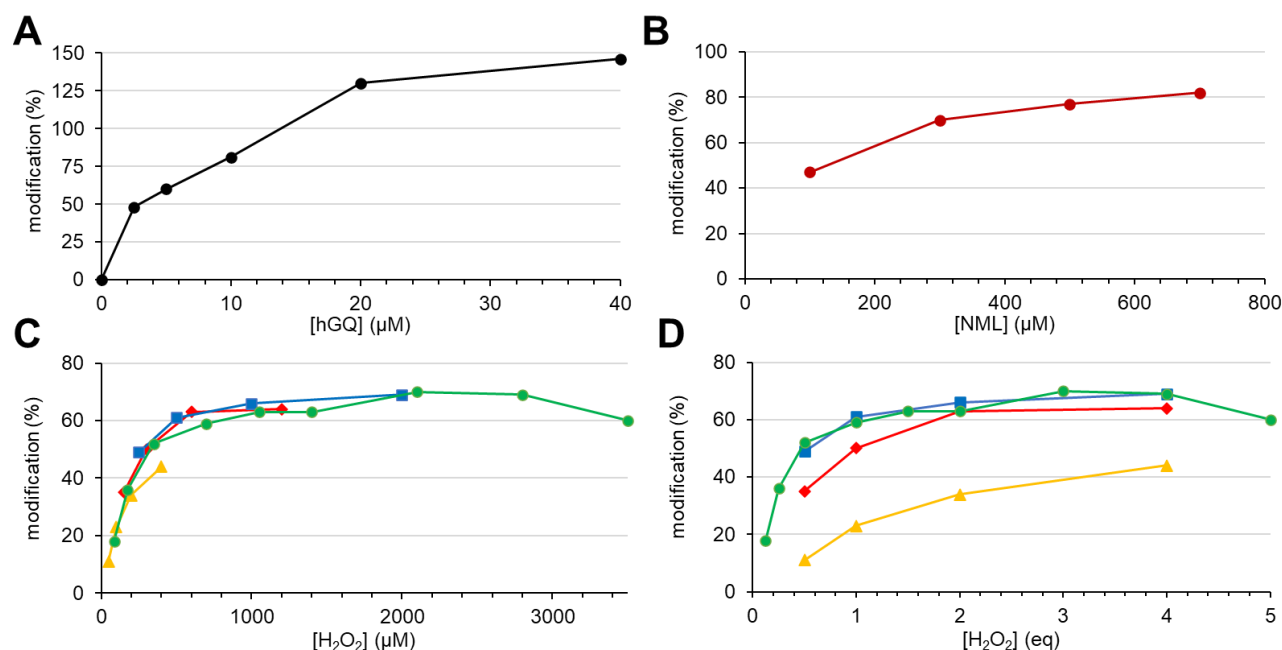


Figure 2.2. (A) Graph showing the effect of DNAzyme concentration on lysozyme modification. Conditions: 140 μM lysozyme, 700 μM NML 1, 2800 μM H_2O_2 , pH: 7.0, at 25 $^{\circ}C$ for 30 min. (B) Graph showing the effect of NML concentration on lysozyme modification. Conditions: 10 μM hGQ DNAzyme, 140 μM lysozyme, 4x[NML] μM H_2O_2 , pH: 7.0, at 25 $^{\circ}C$ for 30 min. (C) Graphs showing the effect of H_2O_2 concentration at different NML concentrations and (D) changing the ratio between NML and H_2O_2 . Conditions: 140 μM lysozyme, pH: 7.0, at 25 $^{\circ}C$ for 30 min. The colour codes in (C–D) are \blacktriangle for 100 μM , \blacklozenge for 300 μM , \blacksquare for 500 μM and \bullet for 700 μM NML 1. Percentages exceed 100%, because multiple modifications are counted separately.

Given the effect of G-quadruplex topology on modification quantity, we performed tryptic digestion in combination with LC-MS/MS analysis to determine if different hGQ DNAzymes also targeted different Tyr residues. This analysis revealed that of the aromatic residues on lysozyme, not only Tyr residues were modified, but also Trp residues; Phe and His residues were unmodified. As expected however, the obtained fragments revealed that modification of Tyr was preferred over the modification of Trp, indicating high chemo-selectivity for Tyr. Singly modified lysozyme occurred on either Tyr23 or Tyr20 (Figure 2.3) which are the most exposed residues with SAA of 32% and 28%, respectively. That site-selectivity was not merely dictated by the quantity of modification, is apparent from the results obtained for G8 and PW17: these systems show very similar modification ability, yet the G8-based hGQ globally modifies lysozyme, whereas PW17 restricts its modification to one side of the protein, *i.e.* Tyr20 and Tyr23 (Figure 2.3A). Interestingly, when the PW17 sequence is conjugated to a lysozyme-binding aptamer (LBA), an additional modification of Tyr53 is observed (Table 2.2), which shows the potential influence of an aptamer on the modification ability of hGQ DNAzymes.²⁸ Importantly, LBA itself did not enhance the background activity of hemin and only marginal levels of modification of lysozyme were observed. As expected, activity tests of lysozyme after modification by the hGQ DNAzyme revealed that the NML modifications cause a decrease in the glycanhydrolase activity (Figure A2.2).²⁹

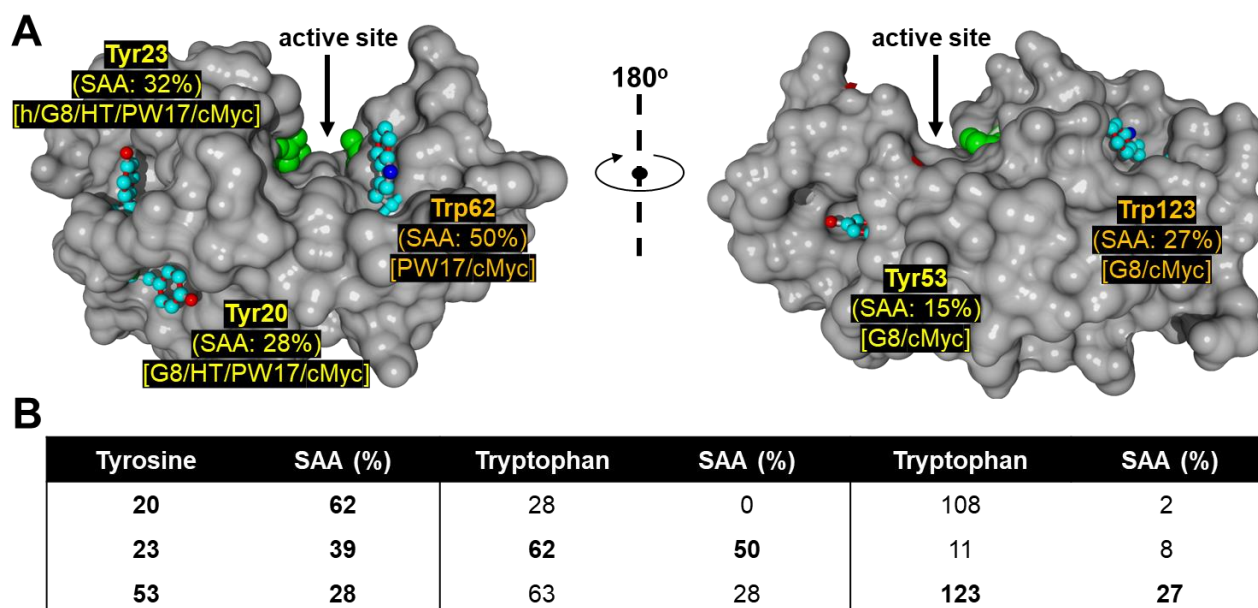


Figure 2.3. (A) Relevant sides of lysozyme C with the position and solvent accessible area (SAA) of the residues that were modified with the respective hGQ DNazymes (given between the square brackets; an 'h' indicates that this residue is also modified by hemin alone), or the Tyr/Trp residues in thrombin that have a high SAA but that are not modified (in red) [based on PDB-code: 3JIV (lysozyme)]. Modified residues are displayed in ball-and-stick, unmodified residues as sticks; active site residues are shown in green ball display. (B) List of all Tyr and Trp residues in thrombin in decreasing SAA percentage, with the residues that are modified in bold.

Following these encouraging results, we subjected the much larger protein human alpha-thrombin (36 kDa)³⁰ to modification conditions using the same hGQ DNzyme sequences. As expected, hGQ DNazymes effectively modified thrombin and displayed differences in their activity and site-specificity with comparable topology-related trends as were found for lysozyme (Table A2.3). We also found that for both lysozyme and thrombin protein crosslinking did not occur in the presence, but also absence, of NML **1**.

Upon studying the influence of the reactant concentrations, similar results as for lysozyme were obtained for thrombin. Higher DNzyme and NML **1** concentrations led again to higher percentages of modification (Figure 2.4A–B) and after 5 μ M of DNzyme, its effect decreases. However, in contrast to lysozyme, increasing H₂O₂ concentration appears always beneficial, even at higher concentrations of NML (Figure 2.4C–D). Possibly this is the result of a larger number of residues being available on the thrombin surface, when compared to lysozyme (Figure 2.2D).

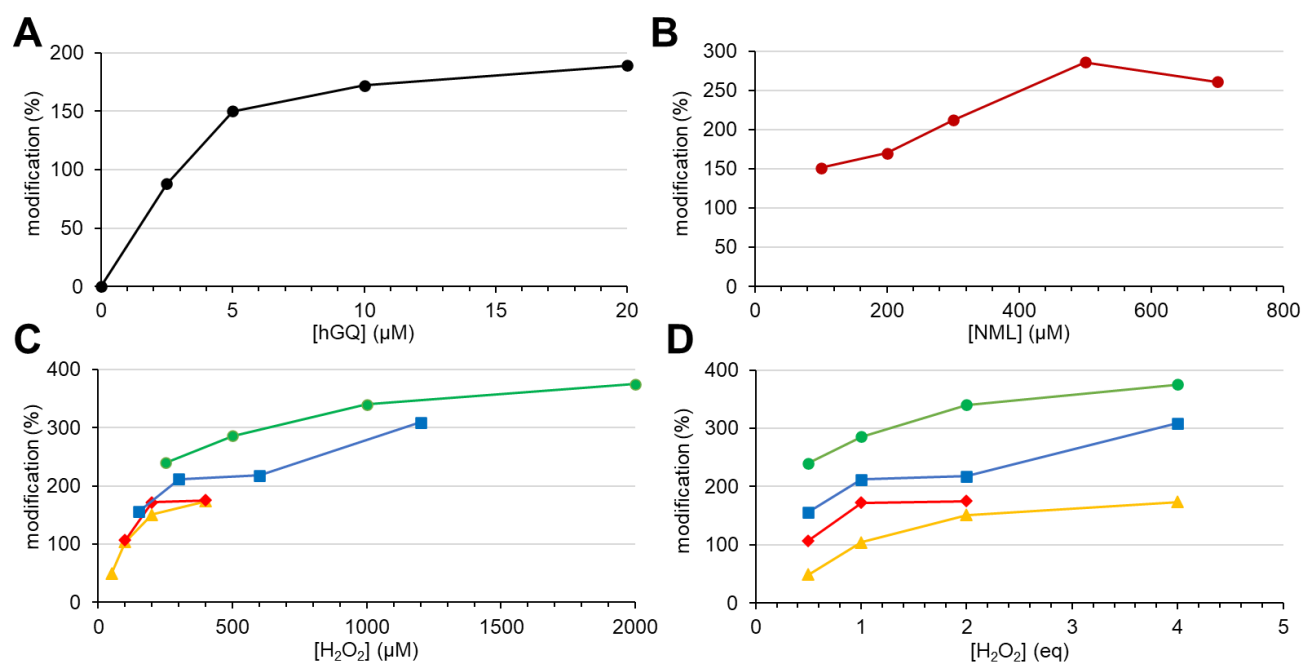


Figure 2.4. (A) Graph showing the effect of DNAzyme concentration on thrombin modification. Conditions: 42 μM thrombin, 200 μM NML 1, 200 μM H₂O₂, pH: 7.0, at 25 °C for 30 min. (B) Graph showing the effect of NML concentration on thrombin modification. Conditions: 10 μM hGQ DNAzyme, 42 μM thrombin, 1x[NML] μM H₂O₂, pH: 7.0, at 25 °C for 30 min. (C) Graphs showing the effect of H₂O₂ concentration at different NML concentrations and (D) changing the ratio between NML and H₂O₂. Conditions: 42 μM thrombin, pH: 7.0, at 25 °C for 30 min. The colour codes in (C–D) are ▲ for 100 μM, ◆ for 300 μM, ■ for 500 μM and ● for 700 μM NML 1. Percentages exceed 100%, because multiple modifications are counted separately.

Thrombin contains various Tyr and Trp residues that would be available for modification judging from their SAA (Figure 2.5B–C). LC-MS/MS analysis of tryptic digests of the reaction mixtures with thrombin revealed that in the presence of any of the hGQ DNAzymes, NML 1 and H₂O₂ modification took place on Tyr85 (SAA: 24%) and Trp148 (SAA: 79%) (Table 2.2). Depending on the G-quadruplex structure, additional modifications were detected. Interestingly, two exposed Tyr residues were not modified: Tyr71 (SAA: 62%) and Tyr47 (SAA: 28%) (Figure 2.5C). Tyr71 is located at the anion-binding exosite I of thrombin, which is also the known binding site of the G-quadruplex thrombin binding aptamer (TBA).³¹ Similarly, Tyr47 is located at the periphery of cationic exosite II, which is the binding site for thrombin binding aptamer HD22. Based on this we propose that the hGQ DNAzymes interact at those sites, thereby blocking modification of these specific residues.

Table 2.2. Modified residues on lysozyme or thrombin by hGQ DNAzyme catalysed modification with NML 1.

GQ type	Residues on lysozyme*	Residues on thrombin [#]
Hemin alone	Tyr: 23	Tyr: 85
Intermolecular parallel (G8)	Tyr: 20 / 23 / 53 Trp: 123	Tyr: 85 / 114 / 134 / 190 Trp: 148 / 190
Intramolecular antiparallel (HT)	Tyr: 20 / 23	Tyr: 85 / 134 Trp: 148
Intramolecular mixed type (PW17)	Tyr: 20 / 23 Trp: 62	Tyr: 85 / 114 / 190 Trp: 148
Intramolecular parallel (cMyc)	Tyr: 20 / 23 / 53 Trp: 62 / 123	Tyr: 85 / 114 / 134 / 190 Trp: 148 / 190 / 27(LC)
Intramolecular mixed type with Lysozyme-binding aptamer (PW17-LBA)	Tyr: 20 / 23 / 53 Trp: 62	n/a

*Conditions: 10 μM hemin, 10 μM DNA, 140 μM lysozyme, 700 μM NML (1), and 2800 μM H₂O₂ (reaction time: 30 min).

[#]Conditions: 10 μM hemin, 10 μM DNA, 42 μM thrombin, 300 μM NML (1), and 300 μM H₂O₂ (reaction time: 30 min).

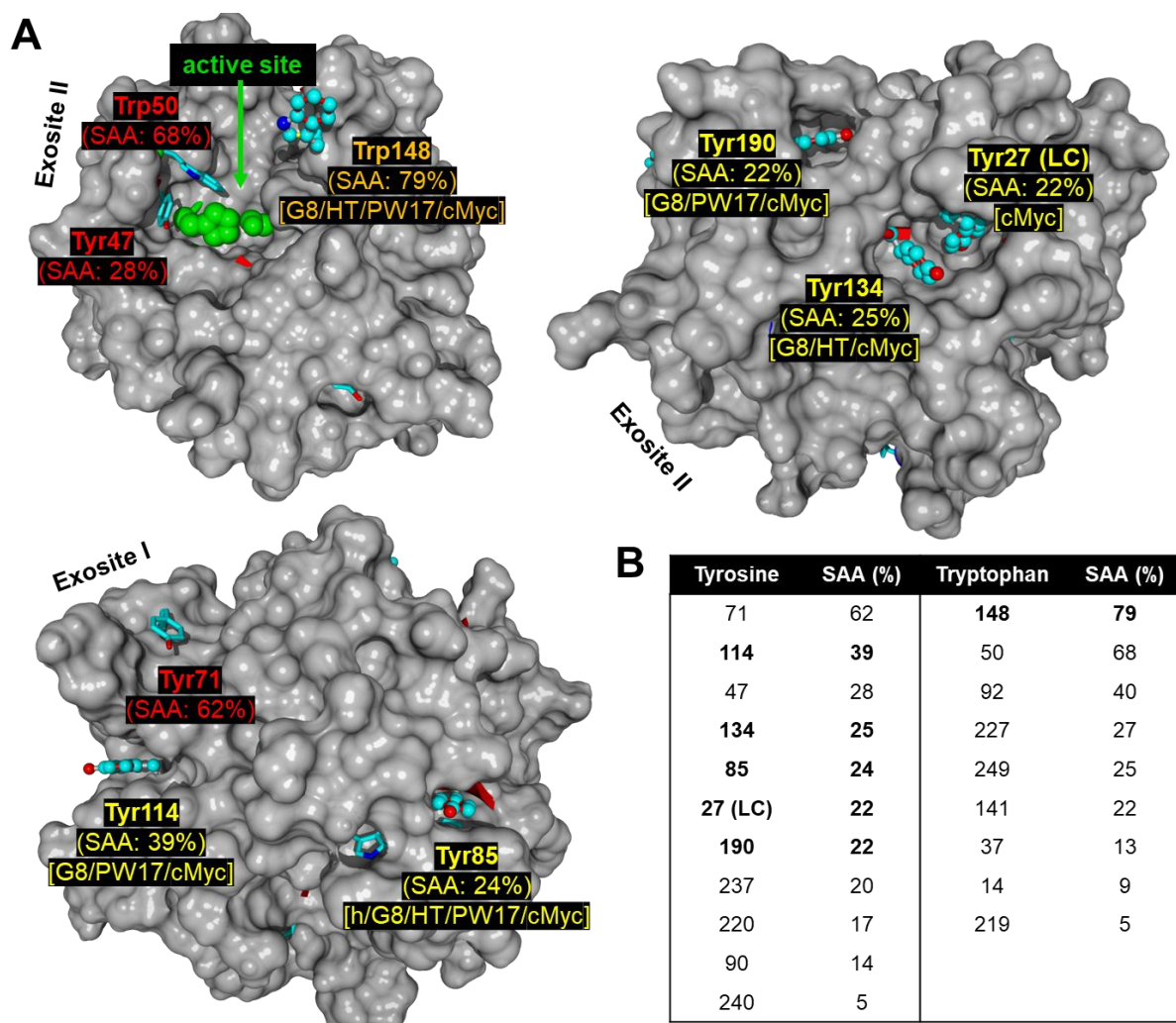


Figure 2.5. (A) Relevant sides of human alpha-thrombin with the position and solvent accessible area (SAA) of the residues that were modified with the respective hGQ DNazymes (given between the square brackets; an 'h' indicates that this residue is also modified by hemin alone), or the Tyr/Trp residues in thrombin that have a high SAA but that are not modified (in red) [based on PDB-code: 5EW2 (thrombin)]. Modified residues are displayed in ball-and-stick, unmodified residues as sticks; active site residues are shown in green ball display. (B) List of all Tyr and Trp residues in thrombin in decreasing SAA percentage, with the residues that are modified in bold (LC refers to the thrombin light chain, the other residues are on its heavy chain).

Apparently, modification with NML **1** by hGQ DNazymes in the presence of H₂O₂ can be limited to only a few sufficiently exposed residues. To examine if this also applied to larger proteins that potentially contain many more exposed reactive residues, we investigated modification of bovine serum albumin (BSA, 66 kDa) and the monoclonal therapeutic antibody trastuzumab (150 kDa). For these proteins we used SDS-PAGE analysis and visualized the modification with a two-step labelling approach in which protein was first modified with NML **1** and then derivatized by means of a strain-promoted alkyne-azide cycloaddition (SPAAC) reaction to a 4 kDa BCN-functionalized PEG unit. As expected, both proteins were modified in the presence of NML **1**, H₂O₂ and hGQ DNazymes (Figure A2.4 and Figure A2.5). DNzyme activities appear to be similarly related to the different topologies as was the case for lysozyme and thrombin. Specifically, whereas BSA was modified once by hemin alone, in the presence of GQ sequences higher numbers of modification were observed. Interestingly, the therapeutically relevant antibody trastuzumab was primarily modified on the heavy chain with up to three modifications for the most active hGQ DNazymes (*i.e.* PW17 and cMyc) and higher concentrations of reagents (Figure A2.5). As was observed for thrombin, the number of modifications decreased when the amount of DNzyme was reduced 2- or 4-fold (from 0.9 eq. with respect to the protein, to 0.45 and 0.225). The light chain was never found to be modified more than once, of which the quantity increased with more active DNazymes. It appears that conditions might require optimization for each protein.

Now that we established that our protein modifying catalysts display features normally only associated with enzymes (i.e., high rate, high chemo-selective, and high levels of site-specificity), we designed a system that allowed regulation of the hGQ catalysed protein modification reaction by means of a switchable element (Figure 2.6A).^{32,33} Upon addition of a deactivation ssDNA sequence that is complementary to the PW17 sequence (functioning as an OFF switch), a dsDNA duplex is formed that does not have the ability to bind and activate hemin. Indeed, we were able to switch the activity of the DNAzyme between its active ("ON") and inactive ("OFF") state by means of an external stimulus. Specifically, in the ON state, the DNAzyme modifies approximately 80% of lysozyme with NML **1**, whereas in the OFF state the modification conversion drops to ~5%, which is similar to hemin alone (HPLC traces can be found in the SI of the published work³⁴). Importantly, the hGQ DNAzyme was reformed after addition of an activation ssDNA strand that was complementary to the entire deactivation ssDNA strand. Since the activating strand contains a high number of guanine bases, we designed a strand that in itself does not form an active hGQ DNAzyme. Indeed, the reformed hGQ DNAzyme complex regained its original protein modifying ability. This was also the case when a DNAzyme-aptamer conjugate was applied (see online SI³⁴). Using the larger and fluorescent NML-lissamine **2** (Figure 2.6B) the switchable character of the hGQ DNAzyme could be visualized by SDS-PAGE analysis for both lysozyme and thrombin (Figure 2.6C). We note that for NML-lissamine **2** the modification efficiency was reduced compared to the smaller NML **1**, which attained conversions of 80%.

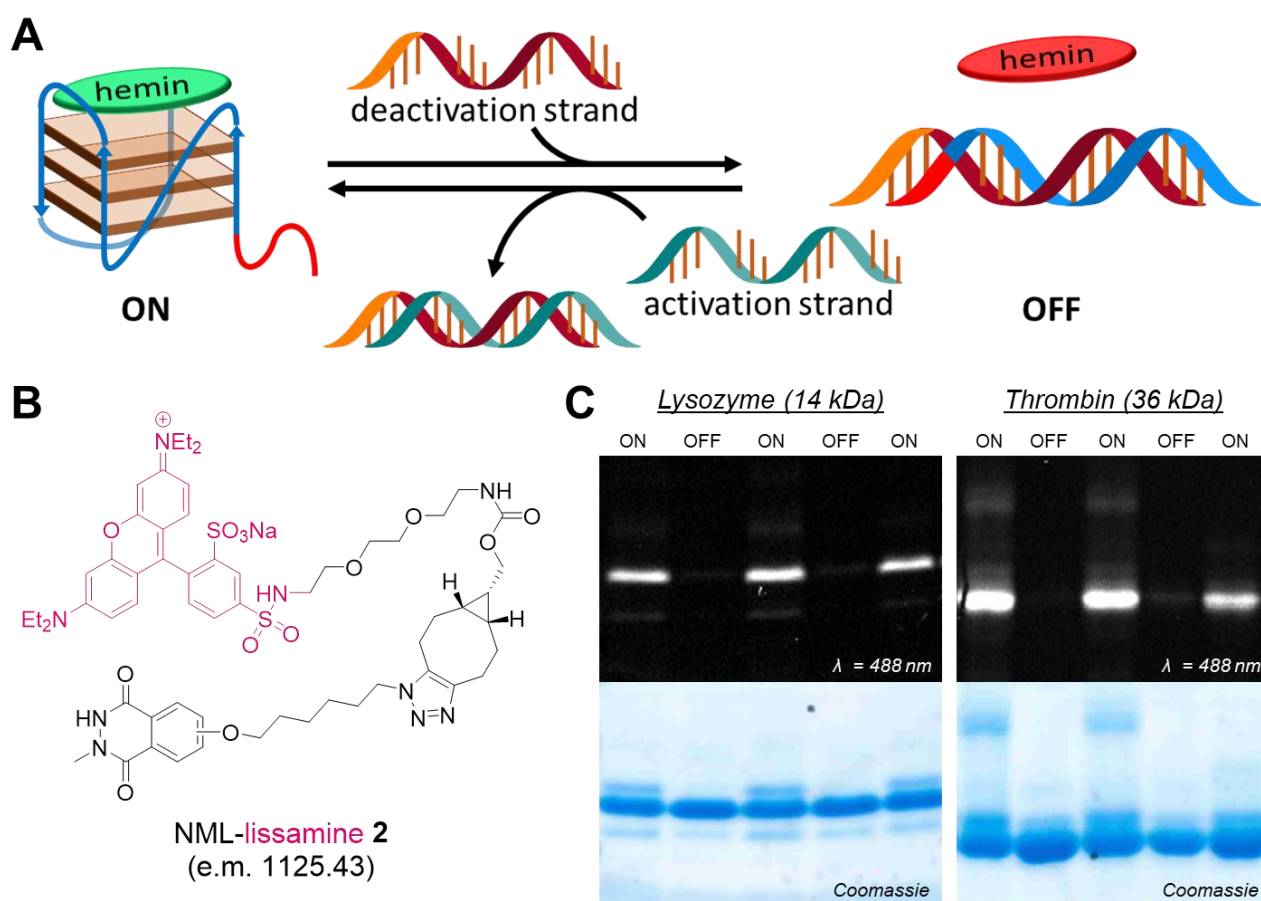


Figure 2.6. (A) Design of the switchable hGQ DNAzyme system. (B) Structure of NML-lissamine **2**. (C) SDS-PAGE analysis of the trigger-regulated modification of lysozyme (left) and thrombin (right) by means of fluorescent NML-lissamine **2**. Top image is a photo of the gel under UV, visualizing fluorescent lissamine moieties. The bottom image show the gel after being stained with Coomassie blue, visualizing all proteins. The combination of these two gel photos shows that thrombin is only conjugated when the DNAzyme is in the 'ON' mode.

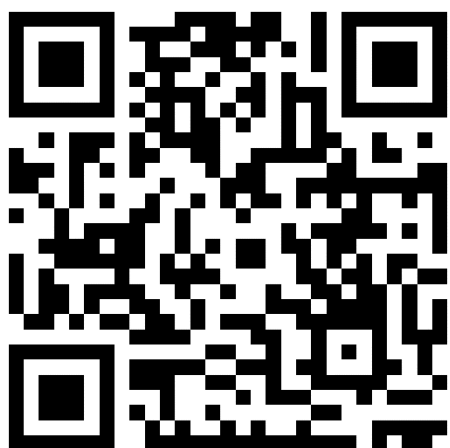
2.4 Conclusion

In this chapter, we described how the hemin/G-quadruplex (hGQ) DNAzyme nanostructure can be used for the oxidative modification of Tyr residues, and to a lesser extent Trp residues, with *N*-methyl luminol derivatives in the presence of H₂O₂. As expected, we found that higher concentrations of DNAzyme, NML and H₂O₂ led to increased quantity of modification, but also that the effect of NML and H₂O₂ concentrations can be influenced by one another, *i.e.* the decrease in modification caused by lower concentrations of one component can be compensated by increasing the concentration of the other. We observed a correlation of the differences in protein modification and the GQ folding conformation, where the parallel GQ sequences are more active than antiparallel GQs. Furthermore, we found preferences for specific residues that are modified by the different GQ topologies. This suggests differences in interaction between the various hGQ DNAzyme with the different target proteins, a process that can further be affected by the application of protein-binding aptamers. Lastly, we show that the catalytic nanostructure can be inactivated by the application of an external trigger, thereby lowering protein modification to the background activity of <5% that we observed for hemin alone.

The observed hGQ DNAzyme-catalysed modification of Tyr residues in proteins is novel, adding a yet unknown C–N bond forming reaction to the hGQ DNAzyme catalytic repertoire. Furthermore, the rapid rate of modification (full conversions are achieved within 30 mins), its high chemo-selectivity (Tyr>>Trp), site-specificity (which is potentially influenced by the presence of a protein-binding aptamer), and ability to respond to an external trigger, make this biomimetic protein modification process not that dissimilar from biological (enzymatic) protein modification processes. In view of the many GQ structures that can bind to proteins,³⁵ the growing applications for DNA nanotechnology^{36,37} and the importance of modified proteins for many lines of research in many scientific disciplines, we expect that our approach will uncover novel catalysts with specific protein modification abilities. Lastly, this study shows that DNA-based catalysts have the remarkable ability to modify proteins, and it is expected that DNA represents not only a biological genetic (indirect) entry to modified proteins, but also an artificial (direct) access to protein modification by means of their exquisite programmable catalytic functions.³⁸

Publication online:

This work is also published³⁴ and can be found here:



Bioconjugate Chem. **2020**, *31* (10), 2283–2287

References

- 1 P. Travascio, Y. Li and D. Sen, *Chem. Biol.* **1998**, *5* (9), 505–517.
- 2 E. Golub, C. H. Lu and I. Willner, *J. Porphyr. Phthalocyanines* **2015**, *19* (1–3), 65–91.
- 3 L. Stefan and D. Monchaud, *Nat. Rev. Chem.* **2019**, *3* (11), 650–668.
- 4 D. M. Kong, W. Yang, J. Wu, C. X. Li and H. X. Shen, *Analyst* **2010**, *135* (2), 321–326.
- 5 W. Li, Y. Li, Z. Liu, B. Lin, H. Yi, F. Xu, Z. Nie and S. Yao, *Nucleic Acids Res.* **2016**, *44* (15), 7373–7384.
- 6 X. Cheng, X. Liu, T. Bing, Z. Cao and D. Shangguan, *Biochemistry* **2009**, *48* (33), 7817–7823.
- 7 L. Stefan, F. Denat and D. Monchaud, *Nucleic Acids Res.* **2012**, *40* (17), 8759–8772.
- 8 E. Golub, H. B. Albada, W. C. Liao, Y. Biniuri and I. Willner, *J. Am. Chem. Soc.* **2016**, *138* (1), 164–172.
- 9 H. B. Albada, J. W. De Vries, Q. Liu, E. Golub, N. Klement, A. Herrmann and I. Willner, *Chem. Commun.* **2016**, *52* (32), 5561–5564.
- 10 H. B. Albada, E. Golub and I. Willner, *Chem. Sci.* **2016**, *7* (5), 3092–3101.
- 11 F. Wang, C. H. Lu and I. Willner, *Chem. Rev.* **2014**, *114* (5), 2881–2941.
- 12 E. A. Hoyt, P. M. S. D. Cal, B. L. Oliveira and G. J. L. Bernardes, *Nat. Rev. Chem.* **2019**, *3* (3), 147–171.
- 13 O. Koniev and A. Wagner, *Chem. Soc. Rev.* **2015**, *44* (15), 5495–5551.
- 14 B. L. Oliveira, Z. Guo and G. J. L. Bernardes, *Chem. Soc. Rev.* **2017**, *46* (16), 4895–4950.
- 15 O. Boutureira and G. J. L. Bernardes, *Chem. Rev.* **2015**, *115* (5), 2174–2195.
- 16 P. G. Isenegger and B. G. Davis, *J. Am. Chem. Soc.* **2020**, *141* (20), 8005–8013.
- 17 S. Sato, K. Nakamura and H. Nakamura, *ChemBioChem* **2017**, *18* (5), 475–478.
- 18 Y. Zhang, K. Y. Park, K. F. Suazo and M. D. Distefano, *Chem. Soc. Rev.* **2018**, *47* (24), 9106–9136.
- 19 K. Minamihata, M. Goto and N. Kamiya, *Bioconjugate Chem.* **2011**, *22* (1), 74–81.
- 20 J. J. Bruins, D. Blanco-Ania, V. Van Der Doef, F. L. Van Delft and B. Albada, *Chem. Commun.* **2018**, *54* (53), 7338–7341.
- 21 J. J. Bruins, C. Van De Wouw, K. Wagner, L. Bartels, B. Albada and F. L. Van Delft, *ACS Omega* **2019**, *4* (7), 11801–11807.
- 22 J. J. Bruins, A. H. Westphal, B. Albada, K. Wagner, L. Bartels, H. Spits, W. J. H. Van Berkel and F. L. Van Delft, *Bioconjugate Chem.* **2017**, *28* (4), 1189–1193.
- 23 J. M. Gilmore, R. A. Scheck, A. P. Esser-Kahn, N. S. Joshi and M. B. Francis, *Angew. Chemie - Int. Ed.* **2006**, *45* (32), 5307–5311.
- 24 S. Sato and H. Nakamura, *Molecules* **2019**, *24* (21), 1–17.
- 25 S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd and S. Neidle, *Nucleic Acids Res.* **2006**, *34* (19), 5402–5415.
- 26 S. Sato, K. Nakamura and H. Nakamura, *ACS Chem. Biol.* **2015**, *10* (11), 2633–2640.
- 27 J. Esque, C. Oguey and A. G. De Brevern, *J. Chem. Inf. Model.* **2010**, *50* (5), 947–960.
- 28 A. A. Bastian, A. Marcozzi and A. Herrmann, *Nat. Chem.* **2012**, *4* (10), 789–793.
- 29 D. Shugar, *Biochim. Biophys. Acta* **1952**, *8*, 302–309.
- 30 G. B. Villaneuva, V. Perret and J. W. Fenton II, *Thromb. Res.* **1984**, *36*, 377–387.
- 31 A. Pica, I. R. Krauss, V. Parente, H. Tateishi-Karimata, S. Nagatoishi, K. Tsumoto, N. Sugimoto and F. Sica, *Nucleic Acids Res.* **2017**, *45* (1), 461–469.

- 32 F. Wang, X. Liu and I. Willner, *Angew. Chemie - Int. Ed.*, **2015**, 54, 1098–1129.
- 33 X. Liu, C. H. Lu and I. Willner, *Acc. Chem. Res.* **2014**, 47 (6), 1673–1680.
- 34 J. F. Keijzer and B. Albada, *Bioconjugate Chem.* **2020**, 31 (10), 2283–2287.
- 35 S. Ray, D. Tillo, R. E. Boer, N. Assad, M. Barshai, G. Wu, Y. Orenstein, D. Yang, J. S. Schneekloth and C. Vinson, *ACS Chem. Biol.* **2020**, 15 (4), 925–935.
- 36 P. Chidchob and H. F. Sleiman, *Curr. Opin. Chem. Biol.* **2018**, 46, 63–70.
- 37 K. E. Bujold, A. Lacroix and H. F. Sleiman, *Chem* **2018**, 4 (3), 495–521.
- 38 S. K. Silverman, *Acc. Chem. Res.* **2015**, 48 (5), 1369–1379.

Appendix 2

A2.1 DNA codes with their respective topologies and nucleotide sequences

Table A2.1. DNA sequence codes, folding topologies, and sequences for each strand that was used in this study.

Code	Topology	DNA Sequence (5' to 3')
Non-GQ	Unknown/unstructured	GAC ACG CCC TGG TTC CGC GCC ATG CCT TTC GCA TTA AGT CGT AGC
G5	Intermolecular	TTT TGG GGG TTT T
G6	Intermolecular	TTT TGG GGG GTT TT
G7	Intermolecular	TTT TGG GGG GGT TTT
G8	Intermolecular	TTT TGG GGG GGG TTT T
G9	Intermolecular	TTT TGG GGG GGG GTT TT
G10	Intermolecular	TTT TGG GGG GGG GGT TTT
HT	Antiparallel	AGG GTT AGG GTT AGG GTT AGG G
TBA	Antiparallel	GGT TGG TGT GGT TGG
TBA2	Antiparallel	TCC GTG GTA GGG CAG GTT GGG GTG AC
TA	Antiparallel	GGT TGG TGT GGT TGG
TA2	Antiparallel	GGG GTT GGG GTG TGG GGT TGG GG
Oxy28	Antiparallel	GGG GTT TTG GGG TTT TGG GGT TTT GGG G
PW17	mixed type	GGG TAG GGC GGG TTG GG
rPS2M	mixed type	GTG GGT AGG GCG GGT TGG
Bcl2	mixed type	GGG GCG CGG GAG GAA GGG GGC GGG
EA2	Parallel	CGA GGT GGG TGG GTG GGA
cMyc	Parallel	TGA GGG TGG GGA GGG TGG GGA A
EAD	Parallel	CTG GGT GGG TGG GTG GGA
EAD2	Parallel	CTG GGA GGG AGG GAG GGA
EAD3	Parallel	CTG GGC GGG CGG GCG GGA
EAD4	Parallel	CTG GGT TGG GTT GGG TTG GGA
EAD6	Parallel	CTG GGG TGG GGT GGG GTG GGG A
VEGF	Parallel	GGG CGG GCC GGG GGC GGG
cKit21	Parallel	CGG GCG GGC GCG AGG GAG GGG
HIF-1α	Parallel	GGG AGG GAG AGG GGG CGG G
RET	Parallel	GGG CGG GCG CGG GCG GG
SwitchGQ*	mixed type	TTACGATTTGCTTT GGGTAGGGCGGGTTGGG
OFF switch	Complementary strand for SwitchGQ	CTGTGCCCCGACCAACCCGCCCTACCCAAAGCCTAAT TCAGCATCG
ON switch	Complementary strand for OFF switch	CGATGCTGAATTAGGCTTTGCGTAGCGCGGGTTGGT CGCGCACAG
LBA-HT	Antiparallel	ATCAGGGCTAAAGAGTGCAGAGTTACTTAGT AGGGTTAGGGTTAGGGTTAGGG
LBA-PW17	mixed type	ATCAGGGCTAAAGAGTGCAGAGTTACTTAGT GGGTAGGGCGGGTTGGG
LBA-cMyc	Parallel	ATCAGGGCTAAAGAGTGCAGAGTTACTTAGT TGAGGGTGGGGAGGGTGGGGAA
LBA-SwitchGQ*	mixed type	ATCAGGGCTAAAGAGTGCAGAGTTACTTAG TTACGATTTGCTTT GGGTAGGGCGGGTTGGG

* additional bases required for the activity switching are indicated in red.

A2.2 Protocol for modification of lysozyme with hGQ DNAzyme

DNA stock solutions were annealed at 95 °C for 5 min prior to use in a reaction. A mixture was typically prepared containing 10 µM GQ DNA (from 100 µM stock in ddH₂O), 10 µM hemin (from 100 µM stock in DMSO), 140 µM lysozyme (from a 1400 µM stock solution in ddH₂O) (*Sigma Aldrich*) and 700 µM of compound **1** (taken from a 7 mM stock solution in DMSO) in PO₄ buffer [50 mM, pH=7.0, with 400 mM NaCl and 5 mM KCl]. This mixture was allowed to stand for 30 min after which H₂O₂ (from 28 mM stock in ddH₂O) was added so that a final concentration of 2.8 mM was obtained. The reaction mixture was then kept in the dark at 25 °C for the indicated time period (*i.e.* 30 min for the data in Tables S2–S7). The reaction was quenched by adding catalase to a final concentration of 0.01 mg/mL (from 0.2 mg/mL stock in (NH₄)₂SO₄ buffer).

For the switchable system, the hGQ DNAzyme was switched OFF by adding 1.2 eq. of an anti-GQ DNA strand after which the mixture was incubated at 40 °C for 20 min. Similarly, the OFF-switched DNAzyme was switched back ON by adding 1.5 eq. (with respect to the original GQ concentration, *i.e.* 1.25 eq. with respect to the OFF strand) of anti-anti-GQ DNA strand after which the mixture was incubated at 40 °C for 20 min. The activities of the switched systems were measured.

The kinetics of lysozyme modification was studied using five representative hGQ DNAzymes, *i.e.* intermolecular species G8, antiparallel species HT, mixed type species PW17, and the parallel species cMyc. From a reaction mixture, aliquots were taken and quenched at the following time points: 0.5, 1, 2, 3, 4, 5, 7, 10, and 15 minutes.

A2.3 Protocol for the analysis of protein modification on HPLC(-MS)

The reaction mixture was aspirated three times with a pipette, after which 10 µL was added to an HPLC vial insert that already contained 10 µL of the TAMN solution. The resulting mixture was also aspirated three times. This sample was then run over a Thermo Fischer MAbPAC RP column 3.0 × 100 mm, at 80 °C, the gradient varying per protein. For lysozyme, the gradient started with 19% (ACN + 0.1% FA) ending with 31% (ACN + 0.1% FA) in (H₂O + 5% ACN + 0.1% FA) (flow rate: 0.5 mL/min) over 20 min. The system used was an Agilent 1220 Infinity LC system with DAD detector.

For mass spectrometry analysis, reaction mixtures were diluted to a final protein concentration of 0.25 mg/mL. Protein samples were then analysed on a Thermo Scientific™ Q Exactive Focus Orbitrap using the same gradient as was used for the HPLC analyses.

A2.4 Tryptic digestion of protein and subsequent analysis to determine site-specificity of the modification

To a sample that contained approximately 200 µg protein, 30 µg of DNase I [taken from a 3 mg/mL DNase I in 100 mM (Tris (pH 7.9) with 25 mM MgCl₂ and 5 mM CaCl₂) stock solution] was added and subsequently kept at RT for 16 h. The mixture was then transferred to an Amicon ultra centrifugal filter unit (MWCO depending on protein size) to wash away the oligonucleotide fragments and to change the buffer to 8 M urea in 100 mM Tris (pH 7.9). The samples were then treated with 10 mM DTT at 37 °C for 1 h and with 20 mM IAA in the dark at RT for 30 min. After removal of excess DTT and IAA by centrifugal filtration, the sample buffer was changed to 1 M urea in 100 mM NH₄HCO₃ (pH 8.0), and treated with 4 µg trypsin gold for 16 h. The peptide fragments were then collected by spinning down and the samples, were desalted and concentrated using Pierce® C18 Tips pipette tips.

Peptide digests were analysed on an EASY nanoLC connected to Thermo Scientific™ Q Exactive PLUS. Peptides were trapped onto a PepSep trap column (2 cm × 100 µm ID, 5 µm C18 ReproSil) and subsequently separated on a PepSep analytical column (8 cm × 75 µm ID, 3 µm C18 ReproSil, PepSep). Elution was achieved using a gradient that started with 5% (ACN + 0.1% FA) ending with 40% (ACN + 0.1% FA) in (H₂O + 0.1% FA), washing the column with 80% (ACN + 0.1% FA) afterwards.

The eluted peaks were analysed using MaxQuant software, searching for peptides with mass modification corresponding to H(17)O(3)C(15)N(5) (*i.e.* the substitution of a proton on the protein by NML derivative **1**) and limiting criteria of 1% PSM FDR and a minimal peptide score of 50. As protein database the entire 'chicken hen egg genome' was used, obtained from www.unitprot.org.

A2.5 Protocol for modification of alpha-thrombin with hGQ DNzyme

DNA stocks were annealed at 95 °C for 5 min prior to usage in a reaction. A mixture was prepared containing 10 µM GQ DNA (from 100 µM stock in ddH₂O), 10 µM hemin (from 100 µM stock in DMSO), 42 µM human α-thrombin (from 419 µM stock in ddH₂O) (*Haematologic Technologies*) and 200 µM of compound **1** (taken from a 2 mM stock in DMSO) in PO₄ buffer [50 mM, pH=7.0, 400 mM NaCl, 5 mM KCl]. This mixture was allowed to stand for 30 min after which H₂O₂ was added to a final concentration of 200 µM (from 2 mM stock in ddH₂O). The reaction mixture was then kept in the dark at 25 °C for 30 min. The reaction was quenched by adding Catalase to a final concentration of 0.01 mg/mL (from 0.2 mg/mL stock in (NH₄)₂SO₄ buffer).

For the switchable system, the hGQ DNzyme was switched OFF by adding 1.2 eq. of an anti-GQ DNA strand after which the mixture was incubated at 40 °C for 20 min. Similarly, the OFF-switched DNzyme was switched back ON by adding 1.5 eq. (with respect to the original GQ concentration, *i.e.* 1.25 eq. with respect to the OFF strand) of anti-anti-GQ DNA strand after which the mixture was incubated at 40 °C for 20 min. The activities of the switched systems were measured.

For HPLC(-MS) analysis of modified thrombin, the gradient started with 20% (ACN + 0.1% FA) ending with 36% (ACN + 0.1% FA) in (H₂O + 5% ACN + 0.1% FA) (flow rate 0.5 mL/min over 40 min).

A2.6 Varying reagent concentrations for lysozyme modification

Table A2.2. Different concentrations of reagents tested for the modification of lysozyme with DNA strand PW17 and the following percentages of un-, mono-, di-modified lysozyme that were produced. Conditions: 140 μ M lysozyme, pH: 7.0, at 25°C for 30 min. Concentration of hemin was always equal to the concentration of PW17.

Conc. PW17 (μ M)	Conc. NML (μ M)	Conc. H ₂ O ₂ (μ M)	total % mod.	% 1 \times mod.	% 2-3 \times mod.	% 0 \times mod.
10	100	50	11	11	-	89
		100	23	23	-	77
		200	34	32	2	66
		400	44	41	3	56
	300	150	35	34	1	65
		300	50	47	3	50
		600	63	57	6	37
		1,200	64	58	6	36
	500	250	49	47	2	51
		500	61	57	4	38
		1,000	66	61	5	34
		2,000	69	61	8	31
	700	88	18	18	-	82
		175	36	34	2	64
		350	52	45	7	48
		700	59	51	8	41
		1,050	63	52	11	37
		1,400	63	52	11	37
		2,100	70	57	13	30
		2,800	69	57	12	31
		3,500	60	50	10	40
		5,600	57	50	7	43
		10,000	48	44	4	52
2.5	700	2,800	46	44	2	54
5			55	50	5	45
10			69	57	12	31
20			95	60	35	5
40			98	51	48	2

A2.7 MS data of modified lysozyme

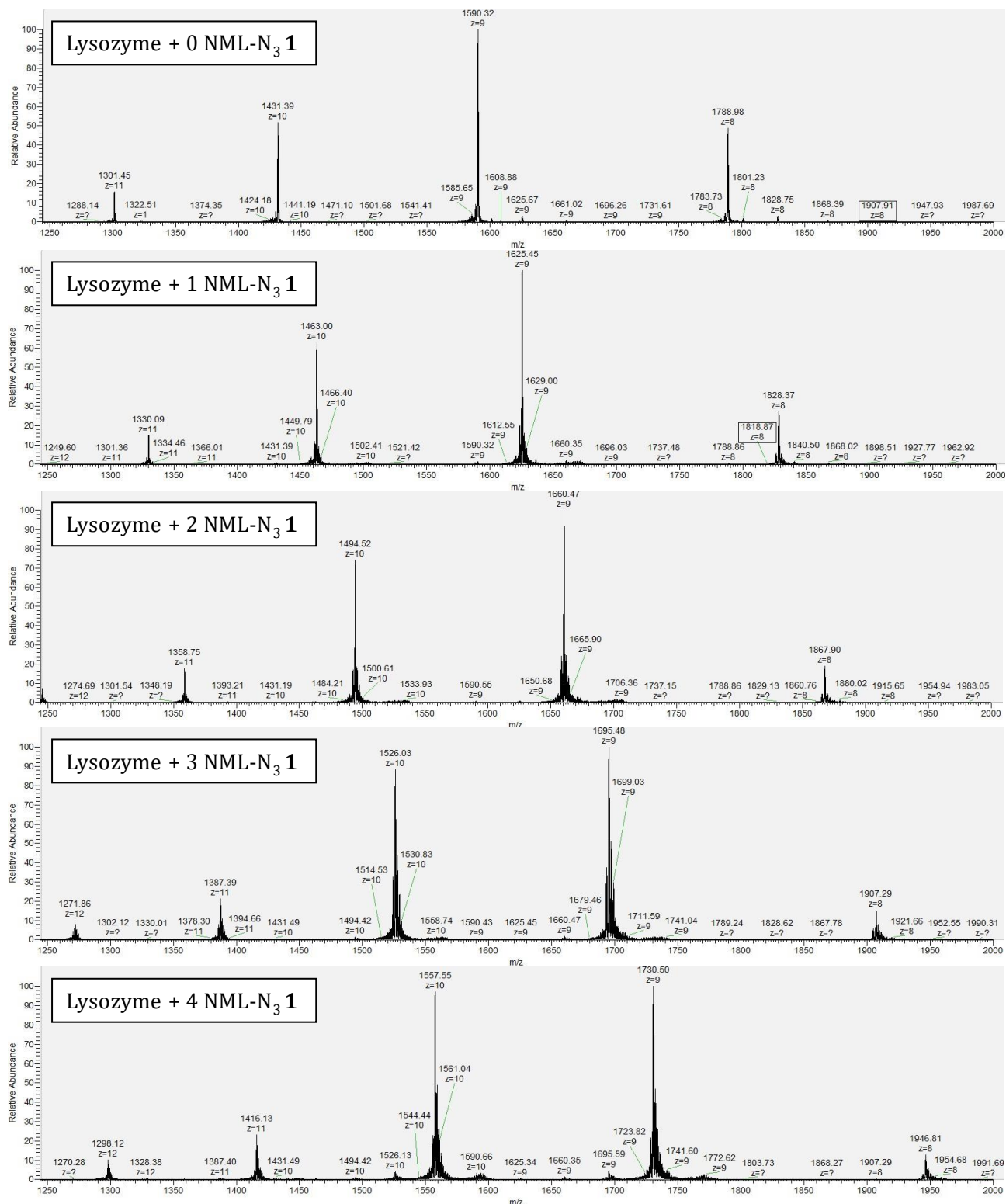


Figure A2.1. MS data associated to Figure 2. 1C, where each separated peak corresponds to lysozyme that has 0, 1, 2, 3 or 4 NML 1 tags attached.

A2.8 Protocol for lysozyme activity assay

The protocol described by Shugar was followed.²⁹ In short, 50 mM PO₄ buffer (pH 6.24) was prepared using monobasic and dibasic potassium phosphate salts, and its pH was tuned using concentrated KOH and HCl. For the substrate, a suspension of *Micrococcus lysodeikticus* (13 mg) in PO₄ buffer (85 mL) was prepared. For the enzyme, a dilution of (un)treated lysozyme in PO₄ buffer (final concentration of 0.01 mg/mL) was used. A disposable cuvette was filled with 2.5 mL of the substrate solution and 0.1 mL of enzyme solution. The two were mixed by inverting the cuvette that was covered with parafilm, after which the absorbance at 450 nm was measured for 5 min. All tests were performed in triplicate (Figure A2. 2) and absorbance values were used to calculate the activity with the formula:

$$\text{Units/mL enzyme} = \frac{(\Delta A_{450}/\text{min } t(x) - \Delta A_{450}/\text{min } t(x-1))(df)}{(0.001) * (0.1)}$$

in which df = dilution factor, ΔA_{450} = change in absorbance as Unit per volume (in mL) of enzyme solution.

GQ topology	code	cell wall cleavage	
		Activity ^o	Reduction
-	(unmodified)	202	
-	hemin	180	-11%
Intermolecular	G8	125	-38%
Antiparallel	HT	140	-31%
Mixed/hybrid	PW17	113	-44%
Parallel	cMyc	89	-56%

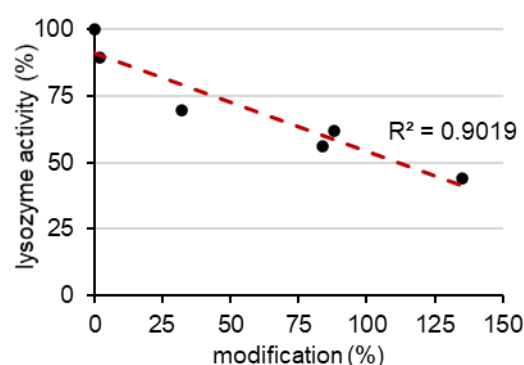


Figure A2.2. The enzymatic activity of lysozyme after modification with different hGQ DNAs with enzymatic activity displayed in Units • (0.01 mg lysozyme)⁻¹ • mL⁻¹. The graph displays the correlation between the modification percentage and the resulting activity decrease, where modification percentage exceeds 100%, because multiple modifications are counted separately.

A2.10 HPLC analysis of modified thrombin

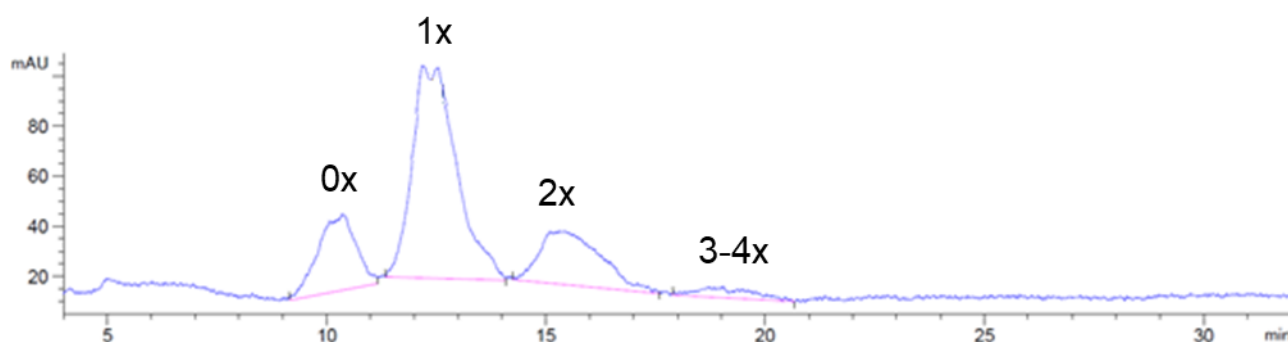


Figure A2.3. HPLC trace of the separation of the products formed during thrombin modification by PW17-based hGQ DNAs with NML 1. The numbers indicate that the mass of the peak corresponds to thrombin + “x” NML modifications.

A2.9 Thrombin modification with hGQ DNAzymes

Table A2.3. Different DNAzymes and the percentages of un-, mono-, di-, tri- and tetra-modified human α -thrombin produced. Conditions: 'x' μ M hGQ DNAzyme, 42 μ M thrombin, 140 μ M lysozyme, 200 μ M NML **1** and 200 μ M H₂O₂, pH: 7.0, at 25°C for 30 min.

Conc. DNA (μ M)	DNA sequence	GQ folding topology	total % mod. throm.	% 1 \times mod. throm.	% 2 \times mod. throm.	% 3-4 \times mod. throm.	% 0 \times mod. throm.
10	-	-	5	5			95
	G8	intermolecular	82	58	20	4	18
	HT	antiparallel	39	35	4		66
	PW17	mixed type	97	41	39	17	3
	EA2	parallel	96	41	40	15	4
	cMyc	parallel	94	59	27	8	6
5	-	-	5	5			95
	G8	intermolecular	70	56	14		30
	HT	antiparallel	44	38	6		54
	PW17	mixed type	96	50	38	8	4
	EA2	parallel	93	56	33	4	7
	cMyc	parallel	93	59	31	3	7
2.5	-	-	2	2			98
	G8	intermolecular	39	35	4		61
	HT	antiparallel	13	13			87
	PW17	mixed type	73	57	14	2	27
	EA2	parallel	71	57	14		29
	cMyc	parallel	65	53	12		35

A2.11 BSA modification with hGQ DNAzymes

	1	2	3	4	5	6	7	8	9	10	11
<i>hemin</i>	-	+	+	+	+	+	+	+	+	+	-
<i>interm</i>	-	-	+	+	-	-	-	-	-	-	-
<i>antipar</i>	-	-	-	-	+	+	-	-	-	-	-
<i>hyrbid</i>	-	-	-	-	-	-	+	+	-	-	-
<i>parallel</i>	-	-	-	-	-	-	-	-	+	+	-
conv(%)		29	53	68	48	51	73	74	77	81	

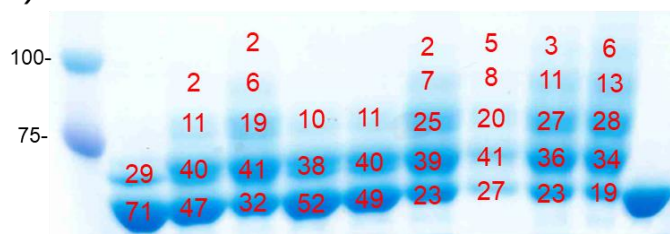


Figure A2.4. SDS-PAGE results of BSA modified with NML **1** and a BCN-functionalized masstag by different DNAzymes. The red numbers display conversion percentages that were calculated with ImageJ. Conditions: 10 μ M hGQ DNAzyme, 30 μ M BSA, 300 μ M NML **1** and 300 μ M H₂O₂, pH: 7.0, at 25°C for 30 min.

A2.12 Trastuzumab modification with hGQ DNazymes

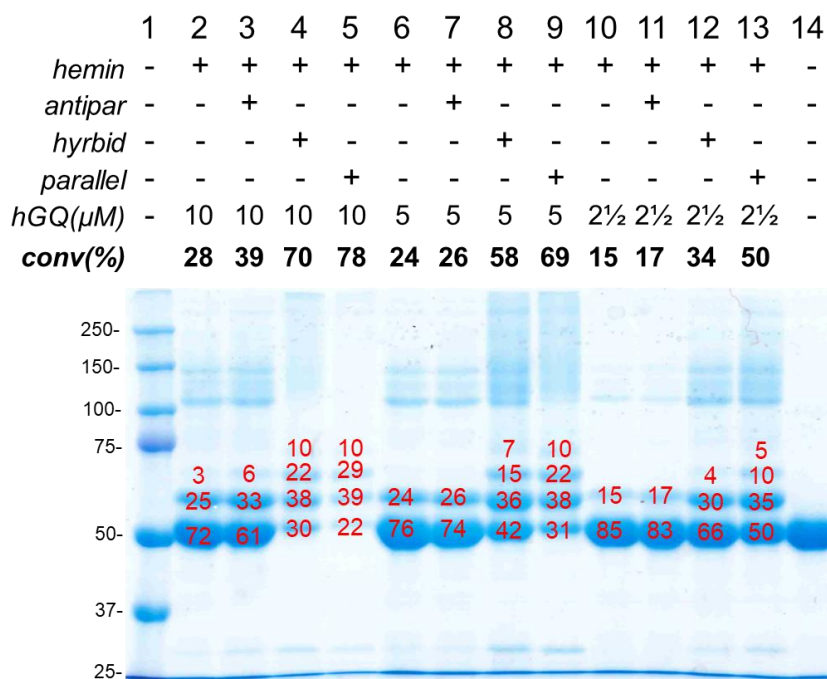
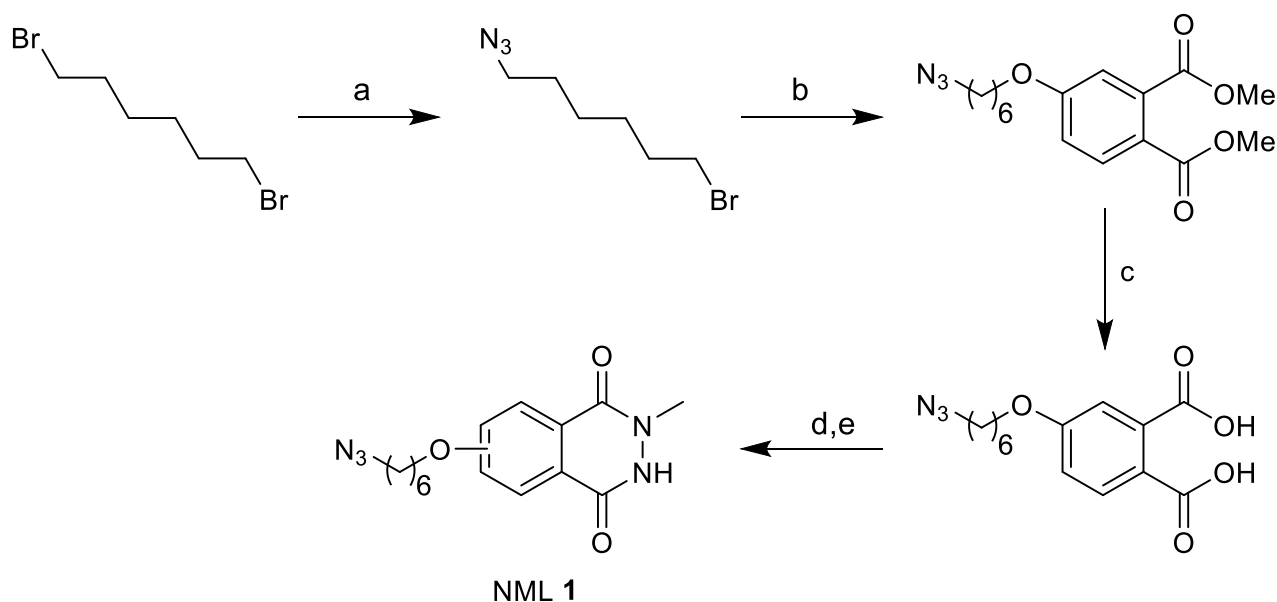


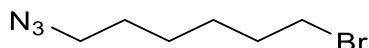
Figure A2.5. SDS-PAGE analysis of the modification of trastuzumab with NML **1** and a BCN-functionalized masstag by means of different DNazymes. The red numbers display conversion percentages calculated with ImageJ. Conditions: 10 μ M hGQ DNzyme, 11 μ M Trastuzumab, 300 μ M NML **1** and 300 μ M H₂O₂, pH: 7.0, at 25°C for 30 min. Trastuzumab is detected on SDS-PAGE by a band at 25 kDa (originating from the two light chains) and a band at 50 kDa (corresponding to the two heavy chains).

A2.13 Synthesis of organic compounds



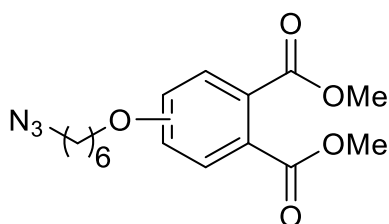
Scheme A2.1. Synthesis of NML **1**. (a) sodium azide, DMF, 55°C–60 °C, 16 h, **23%**; (b) dimethyl 4-hydroxyphthalate, K₂CO₃, DMF, 60°C, 8 h, **85%**; (c) NaOH, MeOH, THF, H₂O, rt, 18 h, **quant.**; (d) acetic anhydride, anh. THF, 70°C, 3 h; (e) methylhydrazine, anh. THF, 82°C, 2 h, **67%**.

1-azido-6-bromohexane:



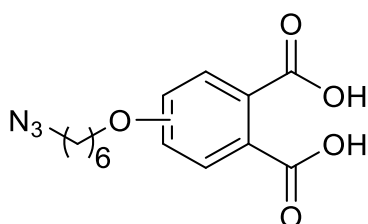
1,6-dibromohexane (9.0 g, 36.9 mmol, 1.5 eq) was dissolved in DMF and whilst stirring, heated to 55 °C. Sodium azide (1.6 g, 24.6 mmol, 1 eq) was gradually added as four smaller batches over the course of 30 minutes and the resulting mixture was stirred at 60 °C for 16 h. The solvent was removed under *vacuo* and DCM (100 mL) and H₂O (100 mL) were added after which the product was extracted with DCM (3 × 70 mL). The organic layer was dried over MgSO₄ and concentrated under reduced pressure. The residue was purified via flash column chromatography (SiO₂, eluent: 0–1% Et₂O in PE40–60), yielding a colourless oil (1.1 g, 23 %). ¹H NMR (400 MHz, CDCl₃), δ 3.41 (td, *J* = 6.9, 1.7 Hz, 2H), 3.28 (t, *J* = 6.9 Hz, 2H), 1.87 (p, *J* = 6.9 Hz, 2H), 1.61 (dq, *J* = 15.9, 8.8, 7.6 Hz, 2H), 1.53–1.35 (m, 4H) ppm. ¹³C NMR (101 MHz, CDCl₃), δ 51.5, 33.8, 32.7, 28.8, 27.8, 26.0 ppm.

Dimethyl 4-((6-azidoheptyl)oxy)phthalate:



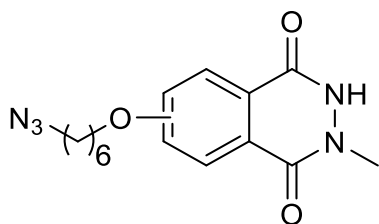
Dimethyl 4-hydroxyphthalate (210.2 mg, 1.0 mmol, 1.0 eq) and 1-azido-6-bromohexane (247.3 mg, 1.2 mmol, 1.2 eq) were dissolved in DMF (4 mL) and K₂CO₃ (221.1 mg, 1.6 mmol) was added and the resulting mixture was stirred at 60 °C for 8 h. The reaction was quenched with brine (20 mL) and the product was extracted with EtOAc (3 × 15 mL). The organic layer was dried over MgSO₄ and the solvent removed under reduced pressure. The residue was purified via flash column chromatography (SiO₂, eluent: 20% EtOAc in heptane), yielding a colourless oil (284.7 mg, 85%). ¹H NMR (400 MHz, CDCl₃), δ 7.72 (dd, *J* = 8.7, 1.1 Hz, 1H), 6.98 (dd, *J* = 2.6, 0.9 Hz, 1H), 6.89 (ddd, *J* = 8.7, 2.6, 1.1 Hz, 1H), 3.92 (td, *J* = 6.3, 1.4 Hz, 2H), 3.83 (d, *J* = 1.4 Hz, 3H), 3.78 (d, *J* = 1.4 Hz, 3H), 3.19 (td, *J* = 6.8, 1.3 Hz, 2H), 1.80–1.64 (m, 2H), 1.53 (q, *J* = 6.7 Hz, 2H), 1.47–1.30 (m, 4H) ppm. ¹³C NMR (101 MHz, CDCl₃), δ 168.6, 166.6, 161.5, 135.6, 131.4, 121.8, 115.9, 113.9, 68.1, 52.5, 52.1, 51.2, 28.7, 28.6, 26.3, 25.4 ppm.

4-((6-azidoheptyl)oxy)phthalic acid:



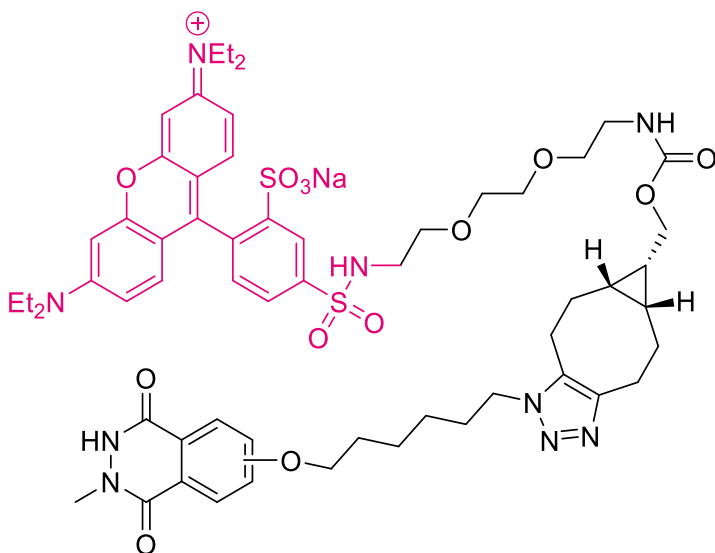
Dimethyl 4-((6-azidoheptyl)oxy)phthalate (200 mg, 596 μmol, 1 eq) and NaOH (143 mg, 3.58 mmol, 6 eq) were dissolved in MeOH:THF:H₂O (2.25 mL) and stirred vigorously at r.t. for 18 h. After this, HCl (1M) was added until pH was 3 (this was reached with approx. 4.5 mL) and the product was extracted with EtOAc (3 × 10 mL). The organic layer was dried over Na₂SO₄, and concentrated under reduced pressure, yielding a white solid (176 mg, 96%). ¹H NMR (400 MHz, MeOD), δ 7.82 (d, *J* = 8.6 Hz, 1H), 7.09 (d, *J* = 2.6 Hz, 1H), 7.03 (dd, *J* = 8.7, 2.6 Hz, 1H), 5.19 (s, 2H), 4.02 (t, *J* = 6.4 Hz, 2H), 3.34–3.18 (m, 2H), 1.77 (dt, *J* = 8.1, 6.3 Hz, 2H), 1.59 (p, *J* = 7.0 Hz, 2H), 1.53–1.35 (m, 4H) ppm. ¹³C NMR (101 MHz, MeOD), δ 172.1, 169.8, 163.0, 138.0, 132.9, 123.5, 116.5, 115.2, 69.4, 52.3, 30.0, 29.8, 27.5, 26.6 ppm.

N-methyl luminol derivative (1):



4-((6-azidohexyl)oxy)phthalic acid (120 mg, 488 μmol , 1 eq) and acetic anhydride (185 μL , 1.95 mmol, 5 eq) were dissolved in anhydrous THF (5 mL) and stirred at 70 $^{\circ}\text{C}$ for 3 h. The solvent was removed under reduced pressure and the residue was dissolved in ethanol (20 mL). Methylhydrazine (77 μL , 1.46 mmol, 3.75 eq) was added and the mixture was stirred at 82 $^{\circ}\text{C}$ for 2 h. The reaction was quenched with saturated NH_4Cl solution (35 mL) and the product was extracted with EtOAc (3×20 mL). The organic layer was then dried over Na_2SO_4 and concentrated under reduced pressure. The target compound was purified via flash column chromatography (SiO_2 , eluent: 50–60% EtOAc in heptane), yielding a white solid (104 mg, 67% over two steps). HRMS (ESI): calculated for $[\text{M}+\text{H}]^+$ 318.1566, found: 318.1566 ^1H NMR (400 MHz, CDCl_3), δ 8.33 (d, $J = 8.8$ Hz, 1H), 8.04 (d, $J = 8.8$ Hz, 1H), 7.77 (d, $J = 2.6$ Hz, 1H), 7.43 (d, $J = 2.5$ Hz, 1H), 7.35 (dd, $J = 8.8, 2.6$ Hz, 1H), 4.14 (t, $J = 6.4$ Hz, 2H), 3.72 (d, $J = 9.8$ Hz, 3H), 3.30 (td, $J = 6.8, 2.7$ Hz, 2H), 1.91–1.81 (m, 2H), 1.65 (p, $J = 7.0$ Hz, 2H), 1.58–1.42 (m, 2H) ppm. ^{13}C NMR (101 MHz, CDCl_3) δ 163.2, 162.9, 157.9, 153.3, 152.6, 132.0, 129.5, 127.2, 122.9, 122.5, 118.5, 108.6, 106.6, 68.7, 51.5, 37.0, 36.8, 29.1, 29.0, 28.9, 26.6, 25.8, 25.8 ppm.

N-methyl-luminol lissamine (2):



BCN-PEG₂-lissamine B (1.7 mg, 2.0 μmol , 1 eq) and azide-functionalized NML **1** (3.1 mg, 9.8 μmol , 5 eq) were dissolved in DMF (0.75 mL) and stirred for 16 h. The target compound was purified via flash column chromatography (SiO_2 , eluent: 10–15% (10% NH_3 in MeOH) in DCM), yielding a deep-pink solid (1.5 mg, 65%). HRMS (ESI) calculated for $[\text{M}+\text{H}]^+$ 1182.5004; found: 1182.5001.

Chapter 3



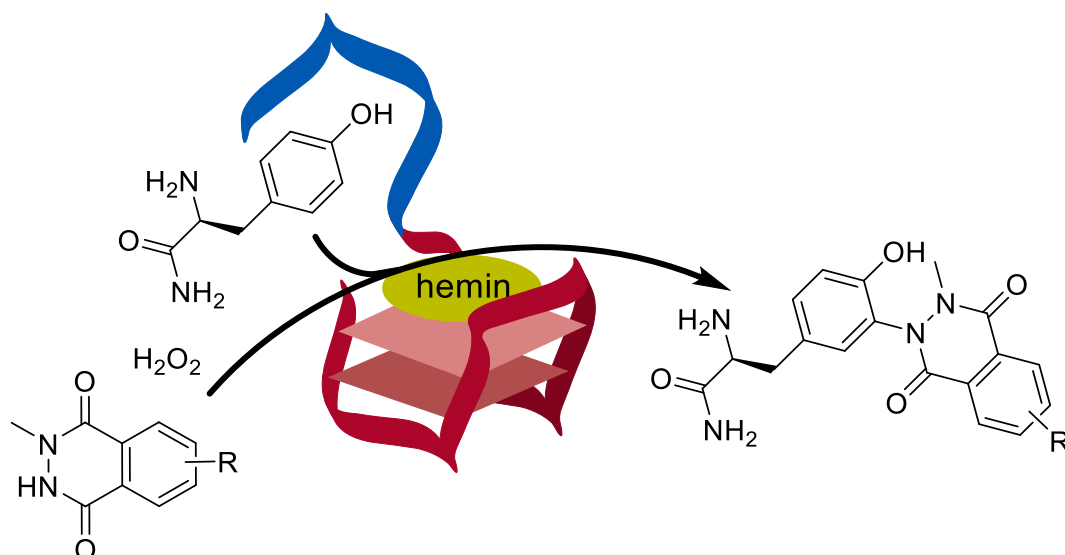
Aptamer-Assisted Bioconjugation of Tyrosine Derivatives using hemin/G-quadruplex (hGQ) DNAzyme Nucleoapzyme Nanostructures

Work by:

Jordi F. Keijzer, Sophie Wintermans, Marte Dros, Han Zuilhof, Bauke Albada

Partially adapted from:

ChemCatChem **2021**, 13, 4618–4624



3.1 Abstract

Hemin/G-quadruplex (hGQ) DNAzymes are horseradish peroxidase-mimicking catalysts that can oxidize various organic substrates, including tyrosine. In this chapter, we implement aptamer-functionalized hGQ DNAzymes, also known as nucleoapzymes, to achieve increased bioconjugation of *N*-methyl luminol to tyrosine-containing residues and peptides. We found that the presence of a tyrosinamide-binding aptamer led to a 12-fold increase in the catalytic rate constant (k_{cat}) and the saturation kinetics curves that were obtained provide evidence for the involvement of the substrate binding site in the reaction. Application of the best performing nucleoapzymes for the modification of Tyr-containing peptides reveals that (i) the aptamer also recognizes the ligand structure when this is embedded in a larger peptide structure, and (ii) distant residues in a peptide substrate can influence the conversion. As such, we show that nucleoapzymes display enzyme-like features and provide an additional instrument in the toolbox of bioconjugation chemistry.

3.2 Introduction

The enzymatic modification of bioactive molecules with chemical entities is a popular strategy for the preparation of molecular constructs with desired biomedical properties.¹ Although enzymes are unmatched when it comes to catalytic activity and substrate selectivity, their use imposes constraints on the experimental conditions that can be applied (*e.g.*, limited range in temperature and pH, limited tolerance for the presence of cosolvents) and on the chemical transformations that can be performed (*e.g.*, limited set of reactions, limited substrate scope). Therefore, it is of interest to develop strategies that mimic the attractive features of enzyme-catalysed bioconjugation reactions² but now for unnatural chemical transformations.

Recently, we³ and others⁴ have shown that the horseradish peroxidase-mimicking hemin/G-quadruplex (hGQ) DNAzyme catalyst⁵ has the ability to efficiently modify tyrosine (Tyr) residues in a variety of proteins with *N*-methyl luminol derivatives.⁶ Applications of such hGQ DNAzymes as protein-modifying catalyst in a switchable nanostructure showed that the modification could be controlled by an externally added trigger.³ Although the hGQ nanostructure itself provided some substrate selectivity, conjugation to an element that specifically recognizes the substrate could, in principle, enable a higher level of selectivity.

One method to achieve specific recognition in DNA-based nanostructures could be the involvement of an aptamer, which (usually) is a nucleic acid sequences that binds a non-nucleic acid molecule. This basic concept of aptamer-assisted catalysis was recently uncovered for the oxidation of dopamine.^{7,8} For these so-called nucleopzymes, it was shown that the conjugation of an aptamer to a substrate-converting catalyst led to increased conversions. Even rational design of catalysts with improved properties⁹, incorporation in supramolecular micelle structures,¹⁰ or the attachment of an artificial catalyst was shown.¹¹ However, in all cases the reaction entailed the oxidation or hydrolysis of one organic substrate, such as dopamine or ATP,¹¹ in the presence of H₂O₂ or H₂O, respectively. Whether aptamer-assisted catalysis would lead to more efficient conjugation of two organic substrates remained unknown, let alone the influence of mutations in the periphery of the reactive moiety of the substrates.

To address this, we explored the potential use of nucleopzymes for bioconjugation reactions. To this aim, we performed a detailed study of the catalysed conjugation of an *N*-methyl-luminol (NML) derivative **1** to tyrosine residues and several tyrosine-containing bioactive peptides^{12–14} by L-tyrosinamide aptamer-functionalized hGQ DNAzyme nucleopzyme nanostructures. We investigate to which extent the reaction depends on the type of tyrosine or otherwise aromatic amino acid residue and if it is influenced (i) by the presence of an aptamer, (ii) by the relative positioning of the aptamer with respect to the hGQ DNAzyme, and (iii) by the influence of proximal amino acid residues in the peptide chain. Finally, we briefly discuss the potential of these novel enzyme-mimicking catalysts for bioconjugation reactions.

3.3 Results & Discussion

First, we determined the ability of hGQ DNAzymes to conjugate NML **1** to the phenol-ring of the tyrosine side-chain, using L-tyrosinamide (TyrAm) as a model substrate (Figure 3.1A). For this, we tested 24 different G-quadruplex forming sequences (Table A2.1, page 40), originating from both biological and artificial sources, that together with hemin could form so-called hemin/G-quadruplex (hGQ) DNAzymes. They consisted of at least four fundamentally different conformation types: intramolecular parallel, anti-parallel, hybrid/mixed, and intermolecular parallel (Figure 3. 1B). For the quantification of the conversion, we used reversed phase HPLC-MS with benzoic acid as internal standard. Although hemin alone also catalysed the conjugation reaction, nearly all G-quadruplex-forming sequences enhanced the

activity of hemin (Figure 3.1C); only the hGQ DNAzyme based on the human telomere (HT) repeat sequence did not display enhanced activity. In accordance with earlier findings from our group³ and others,^{15,16} the anti-parallel hGQ DNAzyme is the least active, followed by the hybrid/mixed and intermolecular parallel structures, and with the hGQ based on the parallel conformation being the most active. Differences in activity among DNAzymes of the same type can be attributed to subtle differences in the interaction between the DNA G-tetrads and hemin, caused by the other nucleobases in the DNA sequence, and to subtle differences in the interaction between hGQ and substrate.¹⁷ Whereas formation of TyrAm-TyrAm dimers and trimers was observed in the absence of NML **1** – phenol dimerization is known to occur with tyrosine residues^{17,18} – this was not seen in the presence of NML **1**.

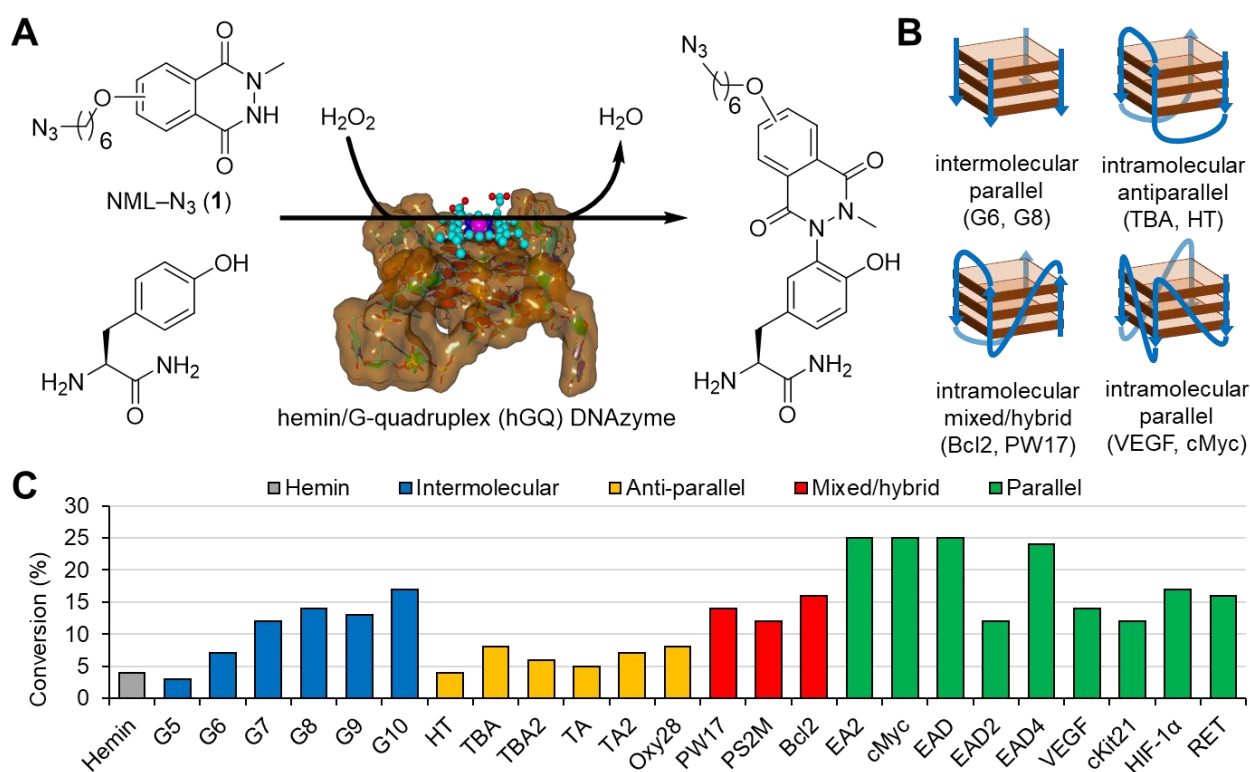


Figure 3.1. (A) Conjugation reaction of TyrAm to NML **1**, catalysed by a hemin/G-quadruplex (hGQ) DNAzyme and H₂O₂. (B) G-quadruplex topologies used to construct the DNAzymes in this study (a general depiction of three G tetrads is shown for all systems, but the amount can differ per sequence). (C) Conversions for hemin and the various hGQ DNAzymes; no conversion was obtained in the absence of hemin. The different G-quadruplex topologies are indicated by the colour code. Conditions: 5 μM hGQ DNAzyme, 100 μM TyrAm, 100 μM NML **1**, and 100 μM H₂O₂, pH: 7.0, at 25 °C for 30 min.

Next, we explored whether fusion of a substrate-binding aptamer to the hGQ DNAzyme could generate a nucleoapzyme for enhanced tyrosine modification with NML **1**. Assuming that the TyrAm-NML product would not display affinity for a tyrosine-binding aptamer, we anticipated that minimal product inhibition would enable multiple turn-overs for each construct. To detect the increase in activity caused by the aptamer, we selected the hybrid/mixed type hGQ DNAzyme PW17, as this by itself showed intermediate peroxidase activity and thus enabled us to detect clear changes in reactivity as result of substrate binding. Since spatial positioning of the substrate binding site relative to the hGQ catalytic site was shown to be important for optimal catalysis,^{7,9} a total of 18 nucleoapzyme constructs was tested. We designed the nucleoapzymes by extending the sequence of PW17 with that of the full or split sequence of L-tyrosinamide-binding aptamer (TamBA, 49 nucleotides, *K_D* = 45 nM; Table A3.1)¹⁹. More specifically, TamBA was placed either as a whole on the 3' or the 5'-end of PW17 or in a split-aptamer approach,²⁰ where the two (varying) parts TamBA's sequence were placed on either side of the PW17 sequence (Figure 3.2A).

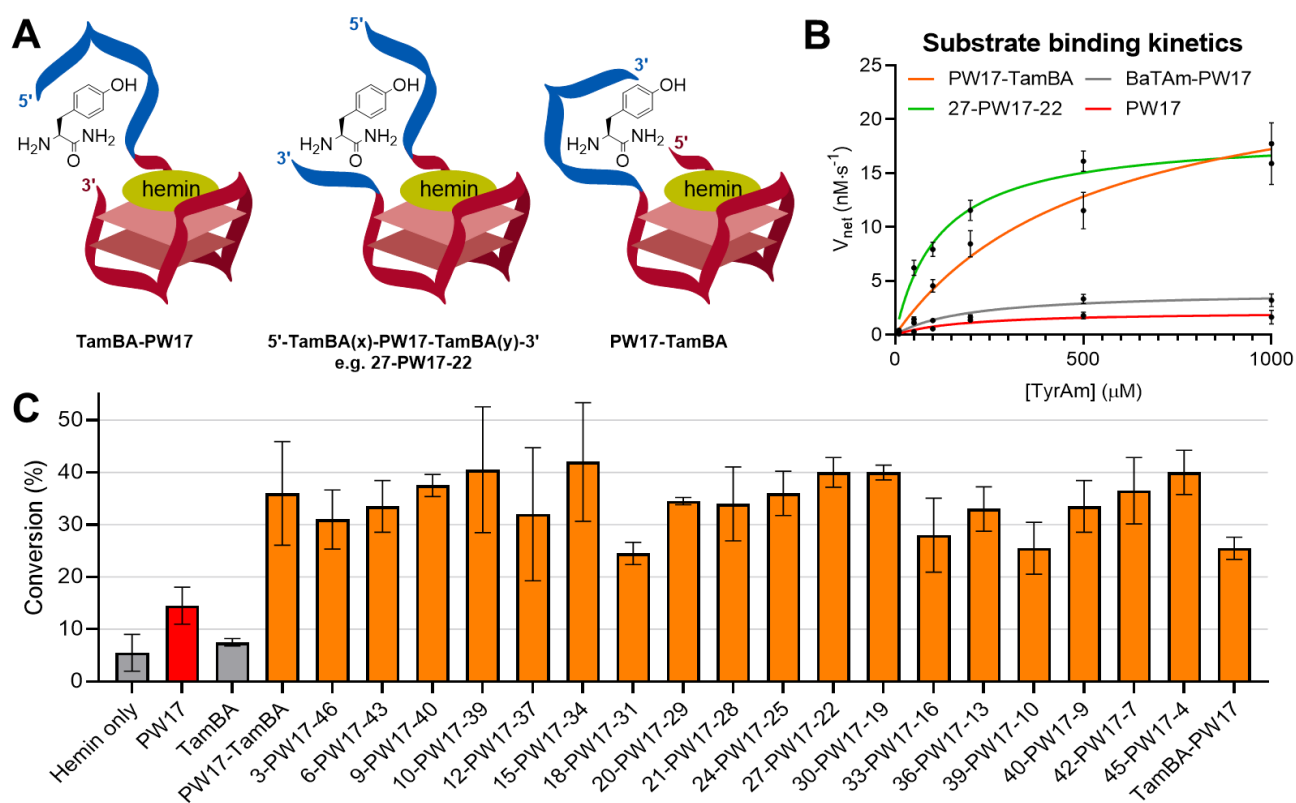


Figure 3.2. **(A)** Schematic depiction of the PW17-based nucleozymes. The (x) and (y) indicate the number of nucleobases from the aptamer on either side of the PW17 sequence. **(B)** Michaelis-Menten kinetics of the isolated PW17 DNAzyme, the DNAzyme-aptamer conjugates PW17-TamBA and 27-PW17-22, and BaTAm-PW17 (5 μM catalyst). $V_{net} = V_{obs} - V_{hemin}$. **(C)** Conversions for the various nucleozymes. Conditions: 5 μM nucleozyme, 100 μM TyrAm, 100 μM NML **1**, and 100 μM H₂O₂, pH: 7.0, at 25 °C for 30 min.

These different nucleozyme constructs were tested for their ability to catalyse the conjugation of TyrAm and NML **1** in the presence of H₂O₂ (Figure 3.2C). Importantly, the aptamer TamBA itself did not increase the activity of hemin (third bar from the left). The TamBA-PW17-based nucleozymes, however, showed 2–3 times higher conversion compared to the PW17 DNAzyme. The nucleozyme in which the full TamBA sequence was joined to the 3'-end of PW17 (PW17-TamBA) was slightly more active than the nucleozyme in which TamBA was joined to the 5'-end (TamBA-PW17). Higher conversions were also observed for the constructs that contained a split-aptamer, although no construct with a clear maximal activity was found. Nevertheless, our observation that the highest conversions were observed for nucleozymes 10-PW17-39, 15-PW17-34, 27-PW17-22, 30-PW17-19 and 45-PW17-4, suggests better substrate binding or alignment with respect to the catalytic hemin moiety in these constructs.

To prove that the TyrAm binding site of the aptamer indeed contributed to the reaction rate, we determined the modification rates at various substrate concentrations (Figure 3.2B). Inclusion of a substrate binding site in the hGQ DNAzyme-forming sequence should lead to saturation kinetics, showing relatively high activities at lower substrate concentrations. Analysis of the reaction curves indeed revealed saturation kinetics for nucleozymes PW17-TamBA and 27-PW17-22, whereas the PW17 DNAzyme and a nucleozyme consisting of the PW17 DNAzyme joined to a scrambled TamBA sequence (*i.e.*, BaTAm) displayed almost linear kinetics (Table 3.1).²¹ The small 1.8-fold higher activity of PW17-BaTAm compared to PW17, is attributed to flanking nucleobases that enhance the activity of the hGQ DNAzyme,²² potentially assisted by electrostatic attraction between the cationic TyrAm substrate and the higher number of negative charges on the extended oligonucleotide. Importantly, the full-aptamer nucleozyme PW17-TamBA and split-aptamer nucleozyme 27-PW17-22 caused a substantially greater enhancement, with 12-fold and 8.5-fold higher V_{max} values, respectively.

Interestingly, the K_M value of the split-aptamer construct 27-PW17-22 was lower than that of PW17-TamBA, which indicates a closer interaction between the substrate binding site and the active centre.²² Additionally, removal of either the 5'- or 3'-end extension of the split-aptamer nucleoapzymes revealed that the 3'-end was more important for the rate enhancement (Figure A3.1). Nevertheless, the conjugate that contained the intact aptamer sequence (*i.e.*, in PW17-TamBA and in TamBA-PW17) ultimately displayed a higher activity as the integrity of both the DNAzyme and the aptamer are likely better maintained than in the split-aptamer approach. Therefore, it can be concluded that incorporation of a substrate binding site in the PW17 DNAzyme leads to substantially higher reaction rates, especially at lower substrate concentrations.

Table 3.1. Kinetic parameters of catalysed reaction between TyrAm and NML 1 in the presence of various hGQ DNAzyme catalysts. Conditions: 5 μ M hemin, 5 μ M DNA, 10–1000 μ M TyrAm / NML 1 / H₂O₂, pH: 7.0, at 25 °C for 30 min.

DNA construct	k_{cat} (10^{-3} s^{-1})	K_M (μM)	k_{cat}/K_M ($10^{-3} \text{ s}^{-1} \cdot \mu\text{M}^{-1}$)	V_{max} ($\text{nM} \cdot \text{s}^{-1}$)	Enhancement (PW17 = 1)
PW17	0.44 ± 0.07	-	-	2.2 ± 0.1	-
PW17-TamBA	5.25 ± 0.75	522 ± 155	0.010	26.2 ± 1.3	12
27-PW17-22	3.72 ± 0.23	119 ± 24	0.031	18.6 ± 0.9	8.5
BaTA-m-PW17	0.81 ± 0.11	-	-	4.0 ± 0.2	1.8

After this, we examined whether the modification of other tyrosine derivatives and aromatic proteinogenic amino acids with NML 1 could be catalysed by our systems (Figure 3.3). For this, we chose hemin, the PW17 DNAzyme, and the two best-performing nucleoapzymes: PW17-TamBA and 27-PW17-22. We found that although native tyrosine (column 2) performed slightly better than TyrAm (column 1), all ester functionalities (column 3-5) displayed slightly lower conversions. The complete absence of the carboxyl group in tyramine (column 6) does not negatively influence the conversion. Moreover, an amide on the N-terminus (column 7) boosts the conversion for Tyram, but reduces that of tyrosine (column 8). Attachment of bulky groups on the N-terminus (column 9 and 10 vs column 8) led to lower conversions, particularly for the nucleoapzymes. As expected, *tert*-butyl protection of the phenol group (column 11) prevents modification almost entirely. Of the other aromatic proteinogenic amino acids, phenylalaninamide and histidine were not modified (column 12 and 13), only tryptophan was conjugated to NML 1 (column 14). Interestingly, the modification of L-tryptophan with NML 1 was also enhanced by the presence of aptamers. In summary, we found enhanced conversions by either of the two nucleoapzymes for all tyrosine derivatives, as well as for tryptophan.

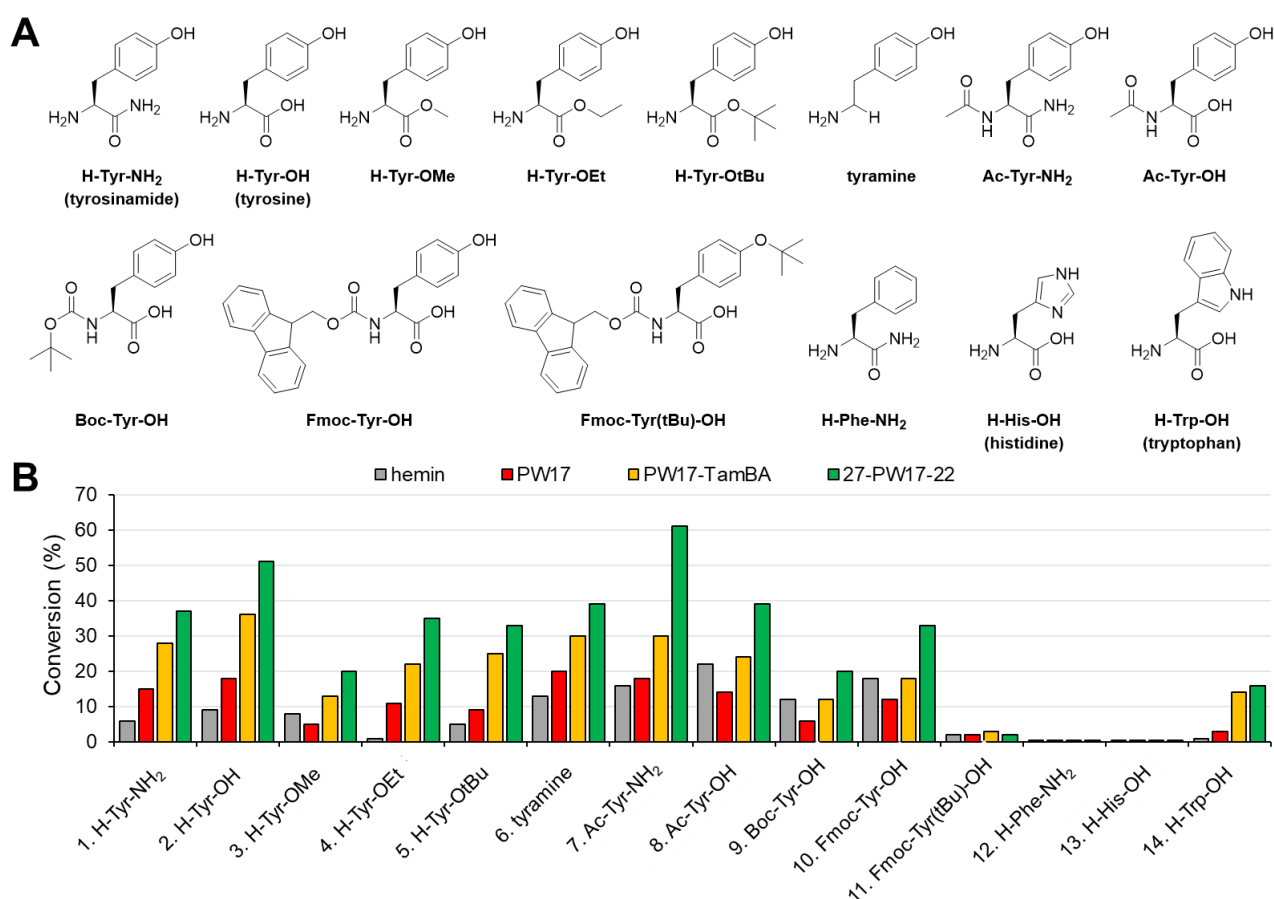


Figure 3.3. (A) The substrates tested in the aptamer-enhanced conjugation reaction. (B) Conversions of the various TyrAm analogues by hemin, PW17, PW17-TamBA and 27-PW17-22. Conditions: 5 μ M nucleopzyme, 100 μ M of tyrosine analogue, 100 μ M NML **1**, and 100 μ M H₂O₂, pH: 7.0, at 25 °C for 30 min.

Following these encouraging results, we tested our systems on biologically relevant tyrosine-containing peptides. For these studies, we chose the opioid neuroactive peptides Leu-enkephalin and Met-enkephalin (Figure 3.4A), which contain a tyrosine residue on their N-terminus (Figure 3.3, column 8). We first modified conjugated NML **1** to these residues using 20 different hGQ DNAzymes. We observed that for these peptides not all hGQ DNAzymes were able to reach higher conversions than hemin (Figure 3.4D, E). Specifically, the intermolecular parallel and intramolecular anti-parallel structures did not show substantially enhanced activity, with a notable exception for the G8 DNAzyme, which doubled the conversion of Met-enkephalin compared to hemin. For the hybrid and intramolecular parallel sequences, conversions all reached over 30% for both substrates. As these peptides only differ at the 5th position with respect to the Tyr residue (Figure 4A), we expected limited influence of the Leu-to-Met mutation on the conversions. Indeed, two peptides showed conversions in the same range with similar differences between the different DNAzyme topologies as seen for TyrAm. The small differences that were observed when comparing the two peptides, are likely the result of interactions between the DNAzyme nanostructure and the microenvironment (caused by distant residues in the peptide sequence), influencing the reactivity of the Tyr residue. As such, the hGQ DNAzyme family displays the ability to discriminate between substrates using very subtle interactions, an essential feature also found in enzymes.

After mapping the abilities of the hGQ DNAzymes, we subjected Leu-enkephalin and Met-enkephalin to NML **1**-conjugation using our nucleopzymes and studied the effect of the aptamer sequence on the conversions (Figure 3.4F, G). Since the N-terminal tyrosine residue matches the most active tyrosine derivative (Figure 3.3, column 8), we predicted the nucleopzymes to enhance conjugation to the peptides as well. Indeed, most of the nucleopzyme constructs achieved substantially higher

conversions compared to the PW17 DNAzyme alone. Whereas PW17 reached a conversion of $22 \pm 6\%$ for Leu-enkephalin and $25 \pm 10\%$ for Met-enkephalin, PW17-TamBA increased the conversions to $30 \pm 3\%$ for Leu-enkephalin and $40 \pm 15\%$ for Met-enkephalin. The most active nucleoapzymes include 15-PW17-34, 27-PW17-22, 30-PW17-19 and 45-PW17-4, a finding that matches our data from the modification of TyrAm. Importantly, analysis of Met-enkephalin by mass spectrometry did not reveal any oxidized Met, proving that the oxidative nature of the hGQ DNAzyme does not cause side-reactions on this residue. Although a limited 2.5-fold enhancement of the conversion of tyrosine residues in peptides was observed, this study demonstrated that TamBA can assist in the conjugation of NML **1** to Tyr-containing peptides. Proof that this process is indeed assisted by substrate-binding, can be derived from the saturation curves obtained for Leu-enkephalin in the presence of PW17-TamBA and 27-PW17-22 (Figure 3.4B). Whereas PW17 and BaTam-PW17 displayed linear curves, PW17-TamBA and 27-PW17-22 displayed typical saturation kinetics curves, supporting the notion that a substrate binding site is involved in the reaction. Furthermore, a distinct difference in the rate of conjugation of Met-enkephalin to NML **1** was observed between 27-PW17-22 and PW17-TamBA (Figure 3.4C), indicating that the split-aptamer approach can be used to enhance the interaction between peptide substrate and aptamer. As this effect was not seen for Leu-enkephalin, we conclude that this is highly sequence specific and shows that our bioconjugation nucleoapzymes can indeed interact with substrates via subtle interactions.⁸

Table 3.2. Kinetic parameters of catalysed reaction between Leu-Enkephalin and NML **1** in the presence of PW17-based hGQ DNAzyme, two nucleoapzyme constructs and PW17 conjugated with a scrambled TamBA sequence (BaTam). Conditions: $5 \mu\text{M}$ hemin, $5 \mu\text{M}$ DNA, $10\text{--}1000 \mu\text{M}$ Leu-enkephalin / NML **1** / H_2O_2 , pH: 7.0, at 25°C for 30 min.

DNA construct	k_{cat} (10^{-3} s^{-1})	K_{M} (μM)	$k_{\text{cat}}/K_{\text{M}}$ ($10^{-3} \text{ s}^{-1} \cdot \mu\text{M}^{-1}$)	V_{max} ($\text{nM} \cdot \text{s}^{-1}$)	enhancement (PW17 = 1)
PW17	0.47 ± 0.1	-	-	2.3 ± 0.1	-
PW17-TamBA	3.6 ± 0.5	5027 ± 1247	0.0007	18.1 ± 0.9	7.8
27-PW17-22	2.1 ± 0.4	1533 ± 659	0.0014	10.3 ± 0.5	4.4
BaTam-PW17	1.2 ± 0.4	-	-	6.0 ± 0.3	2.6

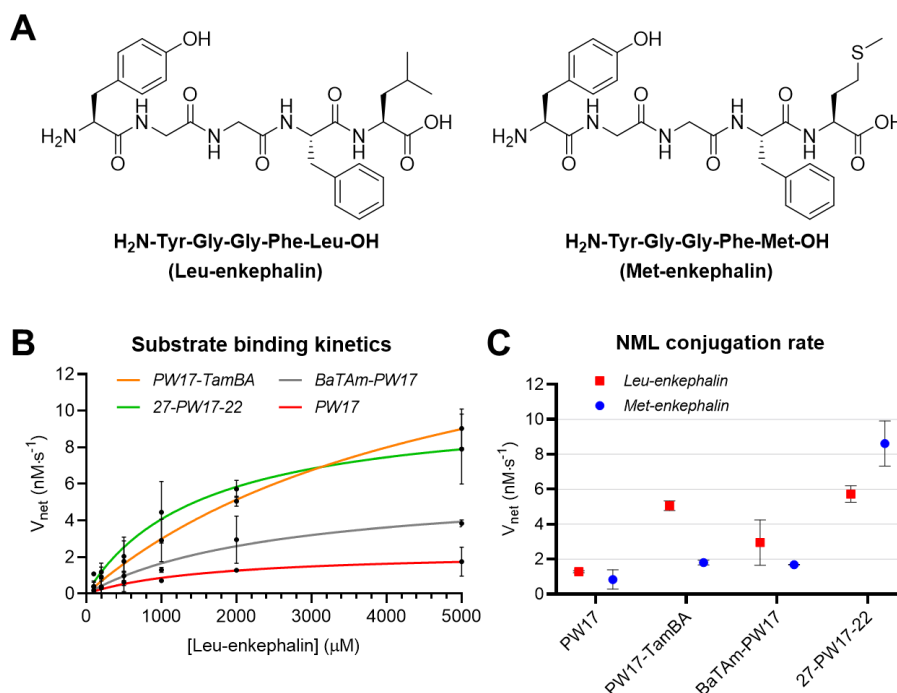


Figure 3.4. (A) Structures of Leu-enkephalin and Met-enkephalin. (B) Michaelis-Menten kinetics of PW17, PW17-TamBA and 27-PW17-22, and PW17-BaTam, where $V_{\text{net}} = V_{\text{obs}} - V_{\text{hemin}}$. (C) Rates of conjugation of Leu-enkephalin or Met-enkephalin in the presence of PW17, PW17-TamBA, BaTam-PW17 and 27-PW17-22. Conditions: $5 \mu\text{M}$ nucleoapzyme, $2000 \mu\text{M}$ peptide, $2000 \mu\text{M}$ NML **1**, and $2000 \mu\text{M}$ H_2O_2 .

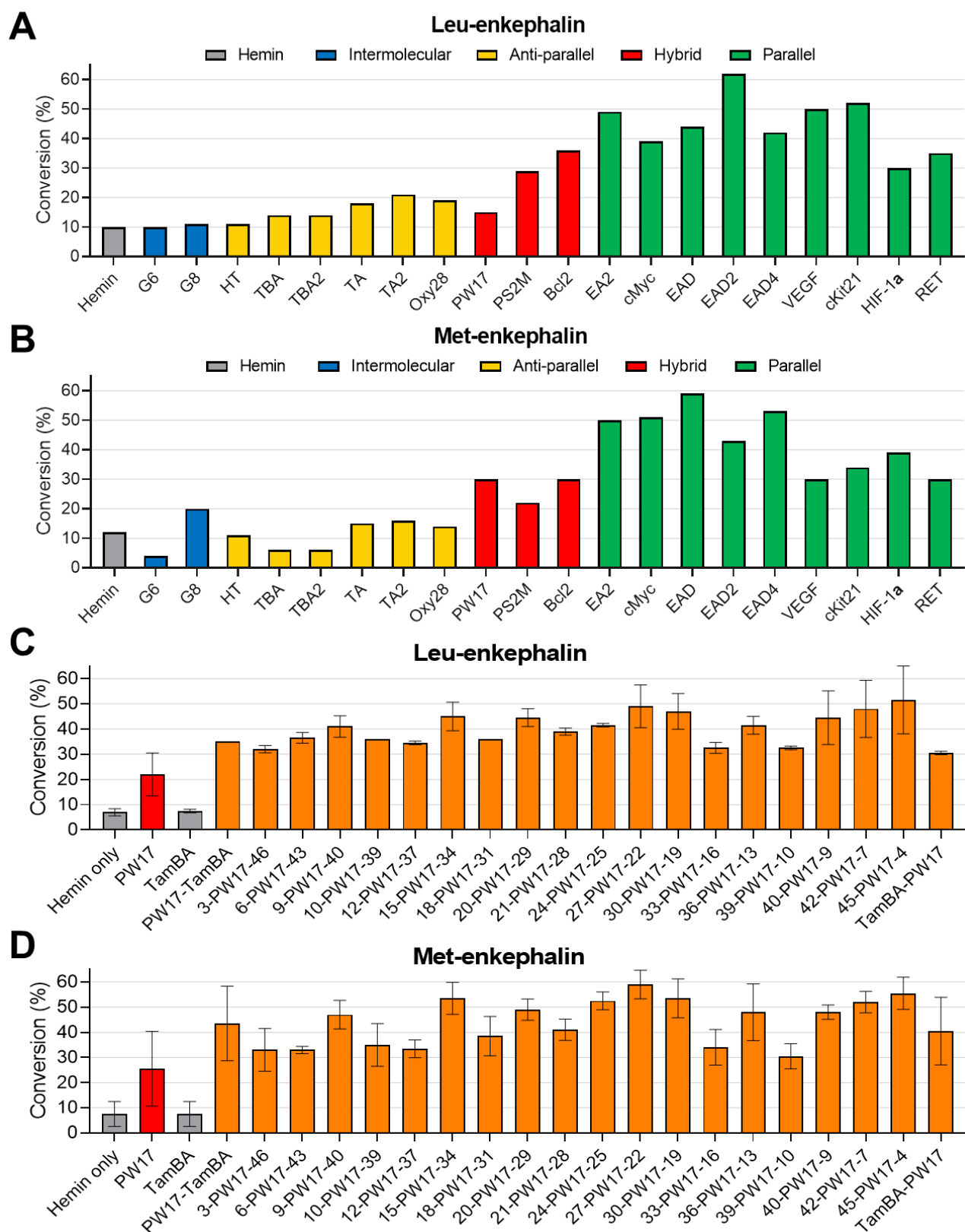


Figure 3.5. Conversions for hemin and the various hGQ DNAzymes (A,B) and nucleozymes (C,D) with Leu-Enkephalin (A,C) or Met-Enkephalin (B,D). Conditions: 5 μ M hGQ/nucleozyme, 100 μ M peptidez, 100 μ M NML 1, and 100 μ M H₂O₂, pH: 7.0 at 25 $^{\circ}$ C for 30 min.

3.4 Conclusions

In this chapter, we described the bioconjugation of tyrosine derivatives and *N*-methyl luminol by means of peroxidase-mimicking hemin/G-quadruplex (hGQ) DNAzymes. First, we showed that the phenol side-chain as found in tyrosine is the preferred functional group of all proteinogenic amino acids when it comes to conjugation to NML. Furthermore, the conversion of this reaction performed on tyrosinamide was enhanced up to 12-fold by extending the DNAzyme with an aptamer sequence that binds tyrosinamide-like substrates. The observed saturation kinetics for the DNAzyme-aptamer constructs revealed that the binding site of this aptamer assists in the bioconjugation reaction, confirming enzyme-mimicry for nucleoapzymes in bioconjugation chemistry. The conjugation of *N*-methyl luminol and the opioid peptides Leu- and Met-enkephalin could also be increased by our hGQ nucleoapzymes, with respect to the DNAzymes. Upon comparison, small differences in conversion were observed for Leu- and Met-enkephalin, which could be caused by residues distally positioned from the site of conjugation can potentially influence the reaction.

Where previous reports of nucleoapzymes concerned only the oxidation or hydrolysis of singular organic substrates,¹¹ our current study proved that conjugation of two substrates can also be enhanced by the use of nucleoapzymes. Therefore, we expect that the current results encourage applications of nucleoapzymes in selective bioconjugation of not only hGQ-driven tyrosine modification, but perhaps other catalysts and residues as well.^{6,23,24}

Publication online:

Part of this work is also published²⁵ and can be found here:



ChemCatChem **2021**, *13*, 4618–4624

References

- 1 E. M. Milczek, *Chem. Rev.* **2018**, *118* (1), 119–141.
- 2 E. Montanari, A. Gennari, M. Pelliccia, L. Manzi, R. Donno, N. J. Oldham, A. Macdonald and N. Tirelli, *Bioconjug. Chem.* **2018**, *29* (8), 2550–2560.
- 3 J. F. Keijzer and B. Albada, *Bioconjug. Chem.* **2020**, *31* (10), 2283–2287.
- 4 T. Masuzawa, S. Sato, T. Niwa, H. Taguchi, H. Nakamura and T. Oyoshi, *Chem. Commun.* **2020**, 56 (78), 11641–11644.

- 5 P. Travascio, Y. Li and D. Sen, *Chem. Biol.* **1998**, *5* (9), 505–517.
- 6 B. Albada, J. F. Keijzer, H. Zuilhof and F. van Delft, *Chem. Rev.* **2021**, *121* (12), 7032–7058.
- 7 E. Golub, H. B. Albada, W. C. Liao, Y. Biniuri and I. Willner, *J. Am. Chem. Soc.* **2016**, *138* (1), 164–172.
- 8 Y. Biniuri, B. Albada and I. Willner, *J. Phys. Chem. B* **2018**, *122* (39), 9102–9109.
- 9 H. B. Albada, E. Golub and I. Willner, *Chem. Sci.* **2016**, *7* (5), 3092–3101.
- 10 H. B. Albada, J. W. De Vries, Q. Liu, E. Golub, N. Klement, A. Herrmann and I. Willner, *Chem. Commun.* **2016**, *52* (32), 5561–5564.
- 11 Y. Biniuri, Z. Shpilt, B. Albada, M. Vázquez-González, M. Wolff, C. Hazan, E. Golub, D. Gelman and I. Willner, *ChemBioChem* **2020**, *21* (1–2), 53–58.
- 12 S. Sato, K. Nakane and H. Nakamura, *Org. Biomol. Chem.* **2020**, *18* (19), 3664–3668.
- 13 S. Sato, K. Nakamura and H. Nakamura, *ACS Chem. Biol.* **2015**, *10* (11), 2633–2640.
- 14 S. Sato, M. Matsumura, T. Kadonosono, S. Abe, T. Ueno, H. Ueda, H. Nakamura, S. Sato and H. Nakamura, *Bioconjug. Chem.* **2020**, *31* (5), 1417–1424.
- 15 X. Cheng, X. Liu, T. Bing, Z. Cao and D. Shanguan, *Biochemistry* **2009**, *48* (33), 7817–7823.
- 16 D. M. Kong, W. Yang, J. Wu, C. X. Li and H. X. Shen, *Analyst* **2010**, *135* (2), 321–326.
- 17 O. J. Einarson and D. Sen, *Nucleic Acids Res.* **2017**, *45* (17), 9813–9822.
- 18 T. Köhler, P. A. Patsis, D. Hahn, A. Ruland, C. Naas, M. Müller and J. Thiele, *ACS Omega* **2020**, *5* (13), 7059–7064.
- 19 E. Vianini, M. Palumbo and B. Gatto, *Bioorganic Med. Chem.* **2001**, *9* (10), 2543–2548.
- 20 G. F. Luo, Y. Biniuri, M. Vázquez-González, V. Wulf, M. Fadeev, R. Lavi and I. Willner, *Adv. Funct. Mater.* **2019**, *29* (33), 1–10.
- 21 A. Cornish-Bowden, *Fundamentals of Enzyme Kinetics*, John Wiley & Sons, **2013**.
- 22 W. Li, Y. Li, Z. Liu, B. Lin, H. Yi, F. Xu, Z. Nie and S. Yao, *Nucleic Acids Res.* **2016**, *44* (15), 7373–7384.
- 23 D. Alvarez Dorta, D. Deniaud, M. Mével and S. G. Gouin, *Chem. - A Eur. J.* **2020**, *26* (63), 14257–14269.
- 24 T. Tamura and I. Hamachi, *J. Am. Chem. Soc.* **2019**, *141*, 2782–2799.
- 25 S. Wintermans, J. F. Keijzer, M. Dros, H. Zuilhof and B. Albada, *ChemCatChem* **2021**, *13*, 4618–4624.

Appendix 3

A3.1 DNA codes with their respective topologies and nucleotide sequences

Table A3.1. The sequence codes, folding topologies, and sequences for each DNA strand that was used in this study.

Code	Topology	DNA Sequence (5' to 3')
PW17	Mixed / hybrid GQ	GGGTAGGGCGGGTTGGG
TamBA	L-Tyrosinamide aptamer ³	AATTCGCTAGCTGGAGCTTGGATTGATGTGGTGTGTGAGTGC GGTGCCC
PW17-TamBA	Full-aptamer nucleoapzyme	GGGTAGGGCGGGTTGGGTAATTCGCTAGCTGGAGCTTGGATT GATGTGGTGTGTGAGTGCGGTGCCC
TamBA-PW17	Full-aptamer nucleoapzyme	AATTCGCTAGCTGGAGCTTGGATTGATGTGGTGTGTGAGTGC GGTGCCCTGGGTAGGGCGGGTTGGG
BaTAm-PW17	Scrambled TamBA nucleoapzyme	AAGTATCGTGAGGTGGCTTCTAGTGATTTGCTGTGTAGGCGC GGGCTGCTGGGTAGGGCGGGTTGGG
3-PW17-46	Split-aptamer nucleoapzyme	AATTGGGTAGGGCGGGTTGGGTTTCGCTAGCTGGAGCTTGGAT TGATGTGGTGTGTGAGTGCGGTGCCC
6-PW17-43		AATTCGTGGGTAGGGCGGGTTGGGTCTAGCTGGAGCTTGGAT TGATGTGGTGTGTGAGTGCGGTGCCC
9-PW17-40		AATTCGCTATGGGTAGGGCGGGTTGGGTGCTGGAGCTTGGAT TGATGTGGTGTGTGAGTGCGGTGCCC
10-PW17-39		AATTCGCTAGTGGGTAGGGCGGGTTGGGTCTGGAGCTTGGAT TGATGTGGTGTGTGAGTGCGGTGCCC
12-PW17-37		AATTCGCTAGCTTGGGTAGGGCGGGTTGGGTGGAGCTTGGAT TGATGTGGTGTGTGAGTGCGGTGCCC
15-PW17-34		AATTCGCTAGCTGGATGGGTAGGGCGGGTTGGGTGCTTGGAT TGATGTGGTGTGTGAGTGCGGTGCCC
18-PW17-31		AATTCGCTAGCTGGAGCTTGGGTAGGGCGGGTTGGGTGGAT TGATGTGGTGTGTGAGTGCGGTGCCC
20-PW17-29		AATTCGCTAGCTGGAGCTTGTGGGTAGGGCGGGTTGGGTGAT TGATGTGGTGTGTGAGTGCGGTGCCC
21-PW17-28		AATTCGCTAGCTGGAGCTTGGTGGGTAGGGCGGGTTGGGTAT TGATGTGGTGTGTGAGTGCGGTGCCC
24-PW17-25		AATTCGCTAGCTGGAGCTTGGATTGGGTAGGGCGGGTTGGG TGATGTGGTGTGTGAGTGCGGTGCCC
27-PW17-22		AATTCGCTAGCTGGAGCTTGGATTGATTGGGTAGGGCGGGTT GGGTGTGGTGTGTGAGTGCGGTGCCC
30-PW17-19		AATTCGCTAGCTGGAGCTTGGATTGATGTGTGGGTAGGGCGG GTTGGGTGTGTGTGAGTGCGGTGCCC
33-PW17-16		AATTCGCTAGCTGGAGCTTGGATTGATGTGGTGTGGGTAGGG CGGGTTGGGTGTGTGAGTGCGGTGCCC
36-PW17-13		AATTCGCTAGCTGGAGCTTGGATTGATGTGGTGTGTTGGGTA GGGCGGGTTGGGTGAGTGCGGTGCCC
39-PW17-10		AATTCGCTAGCTGGAGCTTGGATTGATGTGGTGTGTGAGTGG GTAGGGCGGGTTGGGTGCGGTGCCC
40-PW17-9		AATTCGCTAGCTGGAGCTTGGATTGATGTGGTGTGTGAGTTG GGTAGGGCGGGTTGGGTGCGGTGCCC
42-PW17-7		AATTCGCTAGCTGGAGCTTGGATTGATGTGGTGTGTGAGTGC TGGTAGGGCGGGTTGGGTGGTGCCC
45-PW17-4		AATTCGCTAGCTGGAGCTTGGATTGATGTGGTGTGTGAGTGC GGTTGGGTAGGGCGGGTTGGGTGCCC

15-PW17	Half split-aptamer nucleopzyme, conjugated at 5'	AATTCGCTAGCTGGATGGGTAGGGCGGGTTGGG
PW17-34	Half split-aptamer nucleopzyme, conjugated at 3'	GGGTAGGGCGGGTTGGGTGCTTGGATTGATGTGGTGTGTGAGTGC GGTTGCC
27-PW17	Half split-aptamer nucleopzyme, conjugated at 5'	AATTCGCTAGCTGGAGCTTGGATTGATTGGGTAGGGCGGGTTGGG
PW17-22	Half split-aptamer nucleopzyme, conjugated at 3'	GGGTAGGGCGGGTTGGGTGTGGTGTGTGAGTGC GGTTGCC
30-PW17	Half split-aptamer nucleopzyme, conjugated at 5'	AATTCGCTAGCTGGAGCTTGGATTGATGTGTGGGTAGGGCGGGTTGGG
PW17-19	Half split-aptamer nucleopzyme, conjugated at 3'	GGGTAGGGCGGGTTGGGTGTGTGTGAGTGC GGTTGCC
45-PW17	Half split-aptamer nucleopzyme, conjugated at 5'	AATTCGCTAGCTGGAGCTTGGATTGATGTGGTGTGTGAGTGC GGTTGGGTAGGGCGGGTTGGG
TamBA-HT	Full-aptamer nucleopzyme	AATTCGCTAGCTGGAGCTTGGATTGATGTGGTGTGTGAGTGC GGTTGCCCTAGGGTTAGGGTTAGGGTTAGGG
HT-TamBA	Full-aptamer nucleopzyme	AGGGTTAGGGTTAGGGTTAGGGTAATTCGCTAGCTGGAGCTTGGATTGATGTGGTGTGTGAGTGC GGTTGCC
27-HT-22	Split-aptamer nucleopzyme	AATTCGCTAGCTGGAGCTTGGATTGATTAGGGTTAGGGTTAGGGTTAGGGTGTGGTGTGTGAGTGC GGTTGCC
TamBA-EAD	Full-aptamer nucleopzyme	AATTCGCTAGCTGGAGCTTGGATTGATGTGGTGTGTGAGTGC GGTTGCCCTCTGGGTGGGTGGGTGGGA
EAD-TamBA	Full-aptamer nucleopzyme	CTGGGTGGGTGGGTGGGATAATTCGCTAGCTGGAGCTTGGATTGATGTGGTGTGTGAGTGC GGTTGCC
27-EAD-22	Split-aptamer nucleopzyme	AATTCGCTAGCTGGAGCTTGGATTGATTCTGGGTGGGTGGGTGGGTGGGTGGGTGTGTGAGTGC GGTTGCC

A3.2 Protocol for TyrAm-NNL coupling experiments

Prior to the application of our DNA-based catalysts, DNA stock solutions were annealed at 95 °C for 5 minutes, after which they were allowed to cool to room temperature. In a typical experiment, a mixture containing 5 µM hemin (from 50 µM stock in DMSO), 5 µM DNA (from 50 µM stock in ddH₂O), 100 µM tyrosine-containing substrate (from 1 mM stock in DMSO, containing 2 mM benzoic acid (BA) as internal standard), 100 µM NML **1** (from 1 mM stock in DMSO) in Na₂HPO₄ buffer (50 mM, pH 7.0, with 400 mM NaCl and 5 mM KCl) was prepared in a PCR tube (0.5 mL volume, clear, certified DNase free). This mixture was mixed gently and allowed to stand for 10 minutes at 25 °C, after which H₂O₂ (from a freshly made 1 mM stock in ddH₂O) was added to obtain a final concentration of 100 µM H₂O₂ in the reaction mixture. This mixture was incubated at 25 °C for 30 minutes and subsequently quenched by adding catalase (from 0.1 mg/mL stock in (NH₄)₂SO₄ buffer) to a final concentration of 0.01 mg/mL.

For kinetic experiments, variable concentrations of L-tyrosine-containing substrate, containing the internal standard BA, and NML **1** were added (final concentrations of 10, 50, 100, 200, 500 and 1000 µM). After this, equimolar amounts of H₂O₂ were added and the mixture was incubated at 20 °C. From each reaction mixture, aliquots were taken at various time points after H₂O₂ addition. Each sample was quenched with catalase (from 0.1 mg/mL stock in (NH₄)₂SO₄ buffer, final concentration of 0.01 mg/mL).

A3.3 Protocol for HPLC(-MS) analysis and subsequent calculations

HPLC measurements were performed with an Agilent 1290 Infinity LC System with DAD detector using a ReproSil Gold 300 C₄ (3 µm 250x4 mm) column. LC-MS measurements were performed with an Agilent 1220 Infinity LC system with DAD detector connected to a Thermo Scientific™ Q-Exactive Focus Mass Spectrometer. Reaction mixtures were homogenized by aspirating and dispensing with a pipette several times. Then, 10 µL was added to an HPLC vial insert that contained 10 µL of the TAMN solution, and the resulting mixture was homogenized with a pipette. The injected samples were eluted with a gradient of buffer A (ddH₂O + 5% ACN + 0.1% FA) to buffer B (ACN + 0.1% FA) (see Supporting Table S2 for details of each gradient and specifics of the column).

From the HPLC traces at 269 nm, substrate conversions were quantified using benzoic acid (BA) as internal standard. For every Tyr derivative and peptide, a response factor *F* was calculated *in triplicate* using formula 1:

$$F = \frac{Area_{Tyr^*}[BA]}{[Tyr^*]Area_{BA}} \quad (1)$$

in which $Area_{Tyr^*}$ and $Area_{BA}$ is measured in units of mAU·s⁻¹, and the concentrations of Tyr and BA are measured in mM. Calculated *F* values were averaged and used to determine the concentrations of modified Tyr (*i.e.*, Tyr*) by means of Formula 2:

$$[Tyr^*] = \frac{Area_{Tyr^*}[BA]}{F \cdot Area_{BA}} \quad (2)$$

Using the concentrations of [Tyr*] at the specific concentrations of substrates, saturation curves were made in GraphPad Prism 8. Data points were fitted with a non-linear regression (the '*k_{cat}*' analysis). From the nonlinear fit, the reaction kinetics parameters (*k_{cat}*, *K_M* and *V_{max}*) were automatically determined.

A3.4 Conversions of nucleoapzymes with partial TamBA sequences

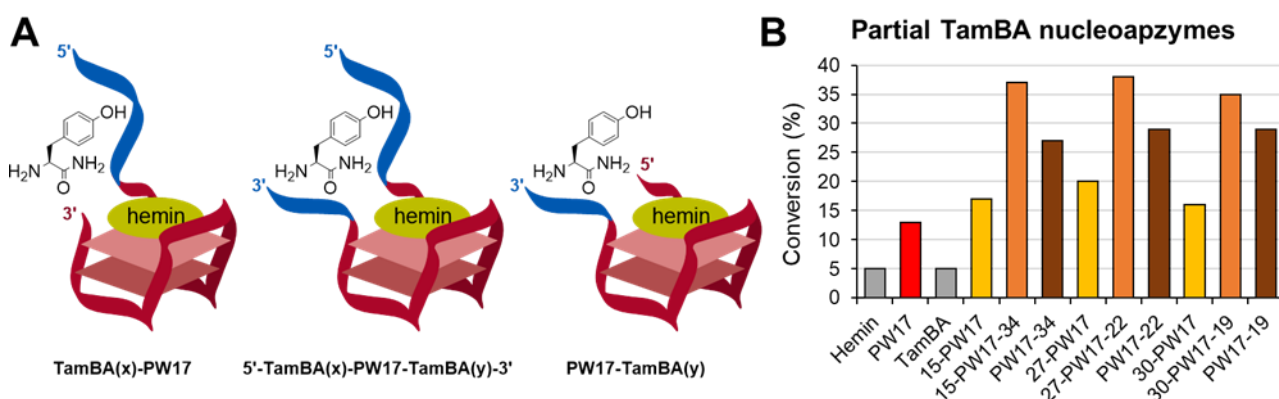


Figure A3.1. (A) Schematic depiction of the PW17-based nucleoapzymes. The (x) and (y) indicate the number of nucleobases from the aptamer on either side of the PW17 sequence. (B) Conversions obtained for the coupling of TyrAm and NML 1 as catalysed by single halves of the split-aptamer nucleoapzymes. Conditions: 5 µM hGQ DNAzyme, 100 µM TyrAm, 100 µM NML 1 and 100 µM H₂O₂, pH: 7.0 at 25 °C for 30 min.

Chapter 4



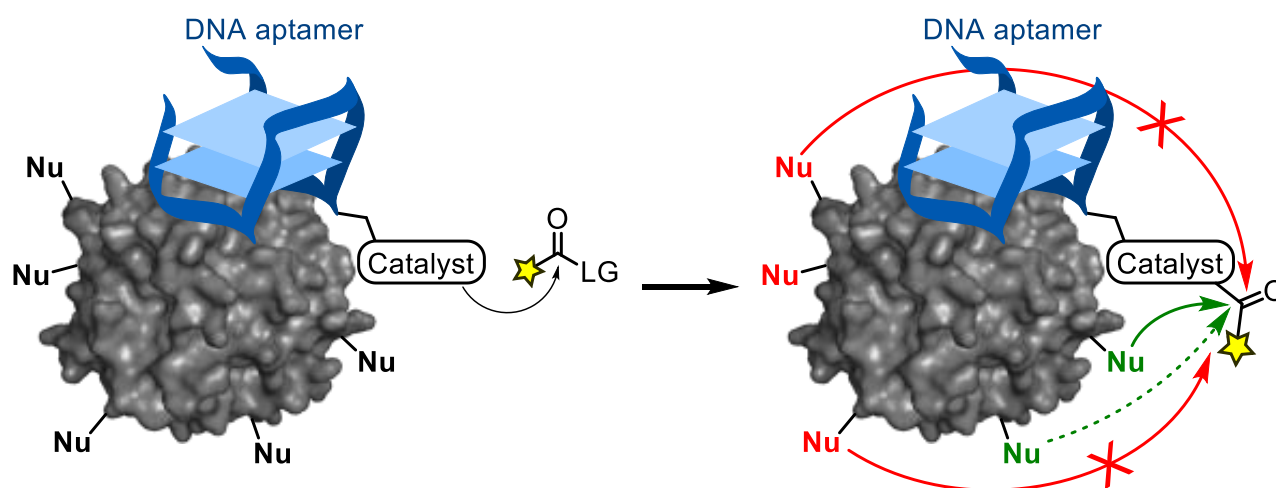
Site-Selective and Inducible Acylation of Thrombin using DNA Aptamer-Catalyst Conjugates

Work by:

Jordi F. Keijzer, Judith Firet, Bauke Albada

Adapted from:

Chem. Commun. **2021**, 57, 12960–12963



4.1 Abstract

Enzymes are nature's most efficient catalysts for protein modification as they possess not only efficiency, but also site selectivity, substrate specificity and controllable activity. Mimicking all these features by one artificial system is challenging and most current methodologies lack one or more of the aforementioned enzymatic features. In our work, we conjugated two acyl-transfer catalysts (DMAP and PyOx) to thrombin-binding DNA aptamers to acylate thrombin. For both catalysts, modification occurred site-selectively on Lys (>>Ser) residues proximal to the respective aptamer-thrombin interface and was selective for thrombin in the presence of other proteins. For DMAP, changing the position of the catalyst on the aptamer led to acylation of different residues on thrombin, which for PyOx could be done by changing to another thrombin aptamer. Additionally, the activity of both DNA-catalysts could be controlled through an external trigger. As such, our artificial approach to protein modification integrates an unparallel set of the afore-mentioned features that, so far, have only been found in protein-modifying enzymes.

4.2 Introduction

The artificial chemical modification of proteins has led to new insights in protein function and novel applications of this important class of biomolecules. Examples of the latter include therapeutics such as antibody-drug conjugates that carry toxins to specific tissues,¹⁻³ and of the former include chemical biology approaches where fluorescently labelled proteins are used to pinpoint their localisation and/or track their movement within cells.^{4,5}

Currently, a remaining challenge is, however, the site-selective modification of native proteins. More than a few conditions have been established to derivatize specific residues in some proteins, such as genetic incorporation of non-canonical amino acids,⁶⁻⁸ reagents that target single residues^{9,10} or a linchpin strategy, where a cleavable bond is formed to confer site-selectivity.¹¹ However, tedious optimization is often required when these methods are applied to different proteins. Alternatively, catalytic modification of reactive residues^{12,13} has resulted in various degrees of selective protein modification, especially when guided by a protein-binding ligand.^{14,15}

Theoretically, this principle can also be applied using DNA aptamers,¹⁶⁻¹⁹ which are nucleotide-based receptor molecules that bind a specific target such as proteins.²⁰⁻²² Protein-binding aptamers have been applied for protein sensors,²³ as well as protein-DNA bioconjugation,^{18,19} yet their application for steering catalytic protein modification remains unexplored. This is unfortunate as DNA aptamers are not only conveniently synthesized²⁴ and functionalized,²⁵ but their affinity for a specific site on their target protein would enable site-selective modification.²¹ A fourth advantage of DNA aptamers is that they are easily included in more complex DNA systems²⁶⁻²⁸ that, for example, facilitate ON/OFF switchable characteristics.^{29,30} As such, aptamers offer a potentially rich platform for trigger-responsive site-selective modification of native proteins (Figure 4.1).

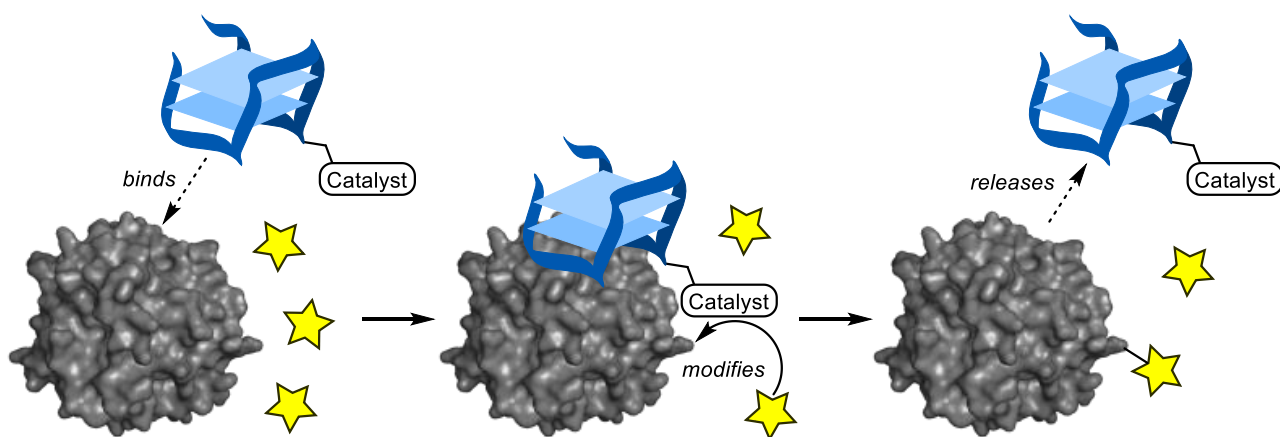


Figure 4.1. Schematic image of an aptamer-catalyst conjugate that binds to its protein target, site-selectively modifies the bound protein with soluble donor molecules, and afterwards dissociates as a result of the dynamic nature of the protein-aptamer interaction, or as result of a conformational change caused by the modification. This cycle can repeat itself with other protein molecules or the same one as long as soluble donor molecules are present.

In this work, we show how catalyst-functionalized aptamers can be used for the selective modification of a native protein, *i.e.*, human α -thrombin. Specifically, one of two acyl-transfer catalysts, DMAP^{31,32} or PyOx,¹⁵ were tethered at various positions to two thrombin-binding aptamers (TBA and TBA2, 15 and 26 nucleotides, respectively) and we assessed the acylation of thrombin, a 36 kDa serine protease that is associated with blood clotting. Additionally, the system could be programmed to include an ON/OFF switchable activity to allow control by means of an external trigger.

4.3 Results & Discussion

Our first set of catalytic DNA constructs consisted of a DMAP catalyst attached to various sites on the smallest aptamer, TBA. For this, a BCN-functionalized dimethylaminopyridine derivative (BCN-DMAP) was synthesized and covalently attached to the appropriate azide-functionalized thymine at one of all seven thymine residues in the TBA sequence (Figure 4.2B)³³ by means of strain-promoted azide-alkyne cycloaddition (SPAAC) chemistry. Subsequent protein modification experiments were performed by incubating thrombin with one of the various TBA-DMAP constructs for 1 h at 37 °C, after which azido-thioester **1** (Figure 4.3A) was added (final concentration: 150 μ M). After modification, an excess of BCN-PEG2000 was added and the mixture was analysed on SDS-PAGE. The various degrees of thrombin modification that we observed showed that different TBA-DMAP constructs led to different enhanced levels of single modification, with TBA³-DMAP and TBA¹²-DMAP as the best performing constructs showing conversions of 27% and up to 49%, respectively (Table A4.2). Inspection of the single crystal X-ray structure of the TBA-thrombin complex (PDB-code: 5EW1)³⁴ confirmed that these two positions of the TBA are indeed closest to the protein surface, which explains their higher activity (Figure 4.4A&B). Kinetics analysis of TBA¹²-DMAP indicated that after 2 h the acylation reaction was mostly complete (Figure 4.2E).

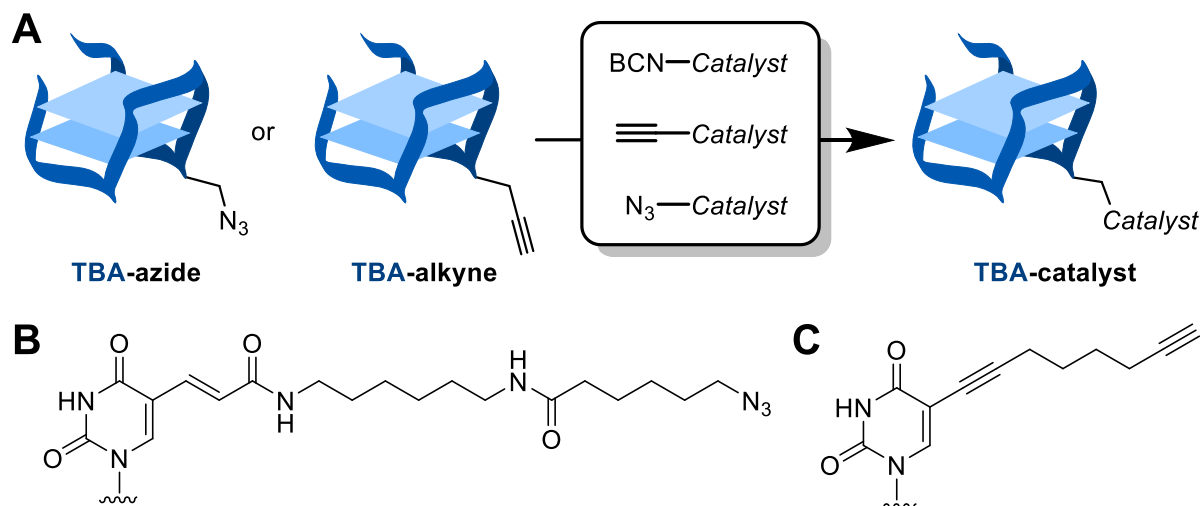


Figure 4.2. (A) Catalysts are attached to the DNA aptamer(s) through strain promoted alkyne-azide cycloaddition (SPAAC; used for BCN) or copper-catalysed alkyne-azide cycloaddition (CuAAC; used for alkyne); (B-C) Structures of (B) azide-functionalized thymine and (C) alkyne functionalized thymine.

Subsequent tryptic digestion and analysis by LC-MS/MS of azide-modified thrombin revealed that the aptamer-assisted single modification of thrombin was performed in a site-selective manner on Lys residues proximal to the respective DMAP-containing thymine bases (Figure 4.4A&B). Our observation that not all residues in proximity to the modified base are acylated, *e.g.*, K77 (Figure 4.4A) and K17 of the light chain (LC) (Figure 4.4B) while apparently more remote residues are acylated, *e.g.*, K154 (Figure 4.4A) and K106 and K107 (Figure 4.4B), confirms that the protein-aptamer interaction is quite dynamic and that subtle interactions influence the site-selectivity.¹⁶

In order to benefit from the affinity of an aptamer for its target, we incubated a mixture of proteins (including thrombin) with TBA³-DMAP or TBA¹²-DMAP and exposed it to alkyne-thioester **2** (Figure 4.3B). Modified proteins were visualized by click conjugation to azido-lissamine by means of the post-modification copper(I)-catalyzed azide-alkyne modify its target protein. In an attempt to counter the high reactivity of thioesters themselves led to high background labelling (up to 20% with 300 μ M **1**), we synthesized divalent DMAP catalyst

constructs (Figure 4.3C) that could more efficiently transfer the acyl group of the substrate to the protein by doubling the amount of activated acyl groups near the protein surface.³⁵ Regrettably, even a divalent DMAP catalyst provided only marginally improved conversions, from 36–49% for DMAP to 48–61% for diDMAP (Table A4.3) cycloaddition (CuAAC). As only fluorescently labelled thrombin was observed and none of the other proteins was modified (Figure 4.3B), we conclude that, indeed, DMAP-aptamer conjugates specifically modify its target protein.

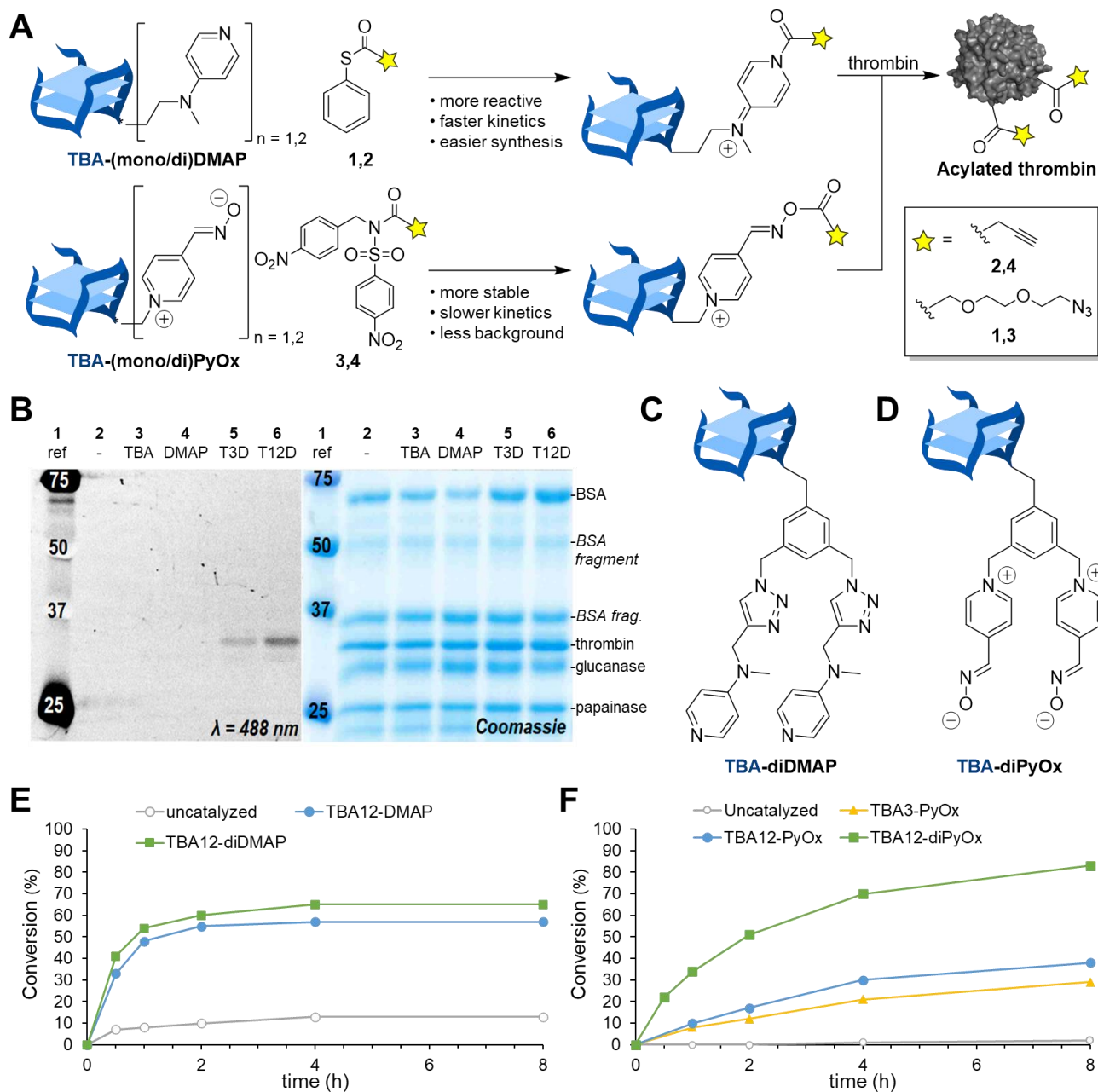


Figure 4.3. (A) DNA-tethered DMAP accepts thioesters **1** and **2** as substrates, whereas PyOx accepts alkylated N-acyl-N-sulfonamides **3** and **4**. Both catalysts form a reactive intermediate from which the acyl is transferred to a nearby nucleophile on the protein. (B) SDS-PAGE analysis and visualization of the selectivity of TBA³-DMAP ("T3D") and TBA¹²-DMAP ("T12D") for thrombin in a selectivity assay. Proteins modified with alkyne-thioester **2** are coupled with azido-lissamine by means of CuAAC, and visualized by fluorescence detection on the gel (left grey-scale image). (C–D) Structure of the divalent catalyst species for (C) diDMAP and (D) diPyOx. (E) Acylation over time by the TBA-DMAP and TBA-diDMAP constructs, compared to the uncatalyzed thioester **1** modification. (F) Acylation over time by TBA³-PyOx, TBA¹²-PyOx, and TBA¹²-diPyOx constructs compared to uncatalyzed ANANS **3** modification.

As the high background activity of the thioesters prevented us from increasing conversions to modified thrombin, we switched to an acyl transfer catalyst that is more nucleophilic than DMAP and can activate less reactive alkylated *N*-acyl-*N*-sulfonamide (ANANS) substrates, *i.e.*, 4-pyridinecarbaldehyde oxime (abbreviated as PyOx, Figure 4.3A).¹⁵ For this study, we focused on TBA positions that gave the best results in the TBA-DMAP constructs, *i.e.*, positions 3 and 12. Thus, after synthesizing alkyne-functionalized PyOx, CuAAC with TBA³-N₃ or TBA¹²-N₃ generated constructs TBA³-PyOx and TBA¹²-PyOx. In addition, azido or alkyne-functionalized ANANS derivatives **3** and **4** were also synthesized (see online SI³⁶). Although exposure of thrombin to TBA³-PyOx or TBA¹²-PyOx and ANANS **3** under the same conditions as for TBA-DMAP revealed even lower conversions, the absence of detectable background at pH 7.2 enabled us to double the substrate concentration and triple the reaction time (Figure 4.3F). This resulted in 18% and 28% conversion to singly-modified thrombin for TBA³-PyOx and TBA¹²-PyOx (at pH 7.2), and 29% and 38%, respectively, at pH 8.0. LC-MS/MS analysis of tryptic digests of thrombin modified with azido-functionalized ANANS **3** revealed, again, site-selective acylation around the TBA-bound PyOx catalyst (Table A4.4). Although overlap was observed with corresponding TBA-DMAP constructs (*i.e.*, acylation of K21, K77, K106, K107, K154 and K18^{LC}), modification of additional Lys residues was also observed (*i.e.*, K83, K174 and K23^{LC}) (Figure 4.4D). To our surprise, two proximal serine residues were also modified, *e.g.*, Ser22 and Ser158 by TBA¹²-PyOx.

Encouraged by this cleaner conversion to modified thrombin, we also designed a divalent PyOx catalyst by mono-azidation of 1,3,5-tri(bromomethyl)benzene and subsequent attachment of PyOx to the resulting 1-(azidomethyl)-3,5-di(bromomethyl)benzene moiety, to yield azido-diPyOx (Figure 4.3D). After CuAAC conjugation of azido-diPyOx to position 3 or 12 on alkyne-functionalized TBA (Figure 4.2A–B), TBA³-diPyOx and TBA¹²-diPyOx were obtained. Upon testing with 300 μ M of azido-ANANS **3** we found that these diPyOx constructs were much more efficient in the modification of thrombin, reaching >90% conversion to acylated thrombin (at pH 7.2; Table A4.4). That this higher percentage also entailed larger quantities of multi-modified protein (specifically: 28% single, 30% double, 22% triple and 11% quadruple modified thrombin) provides evidence that a higher concentration of activated PyOx catalyst is located at the surface of the protein. Interestingly, TBA^{3,12}-bis(diPyOx), which contains two diPyOx catalysts at different positions, was not more active (Table A4.4). Tryptic digestion analysis of thrombin modified at pH 8.0 showed that, when compared to monovalent TBA¹²-PyOx, divalent TBA¹²-diPyOx modified additional residues K196 and K23^{LC}, while at the same time omitting modification of residues S22 and K83 (Figure 4.4D). Larger differences between PyOx and diPyOx functionalized aptamers were observed at pH 7.2, with 3.9–5.2-fold enhancement when placed on TBA³ and 2.7–3.3-fold enhancement when placed on TBA¹². Apparently, the diPyOx catalyst generates a higher concentration of reactive acyl-catalyst complex in proximity to the protein surface than the DMAP catalyst (Table A4.3 and Table A4.4). As was the case for the DMAP conjugates, TBA¹²-diPyOx also only modified thrombin in competitive assays (Figure A4.2). LC-MS analysis of TBA¹²-diPyOx in the presence of ANANS **3** revealed mono- and di-acylated constructs as well as dehydrated catalyst in the absence of thrombin (see online SI³⁶). Apparently, an acylated PyOx catalyst is formed that dehydrates in the absence of a substrate; self-acylation of the DNA was not observed by this analysis.

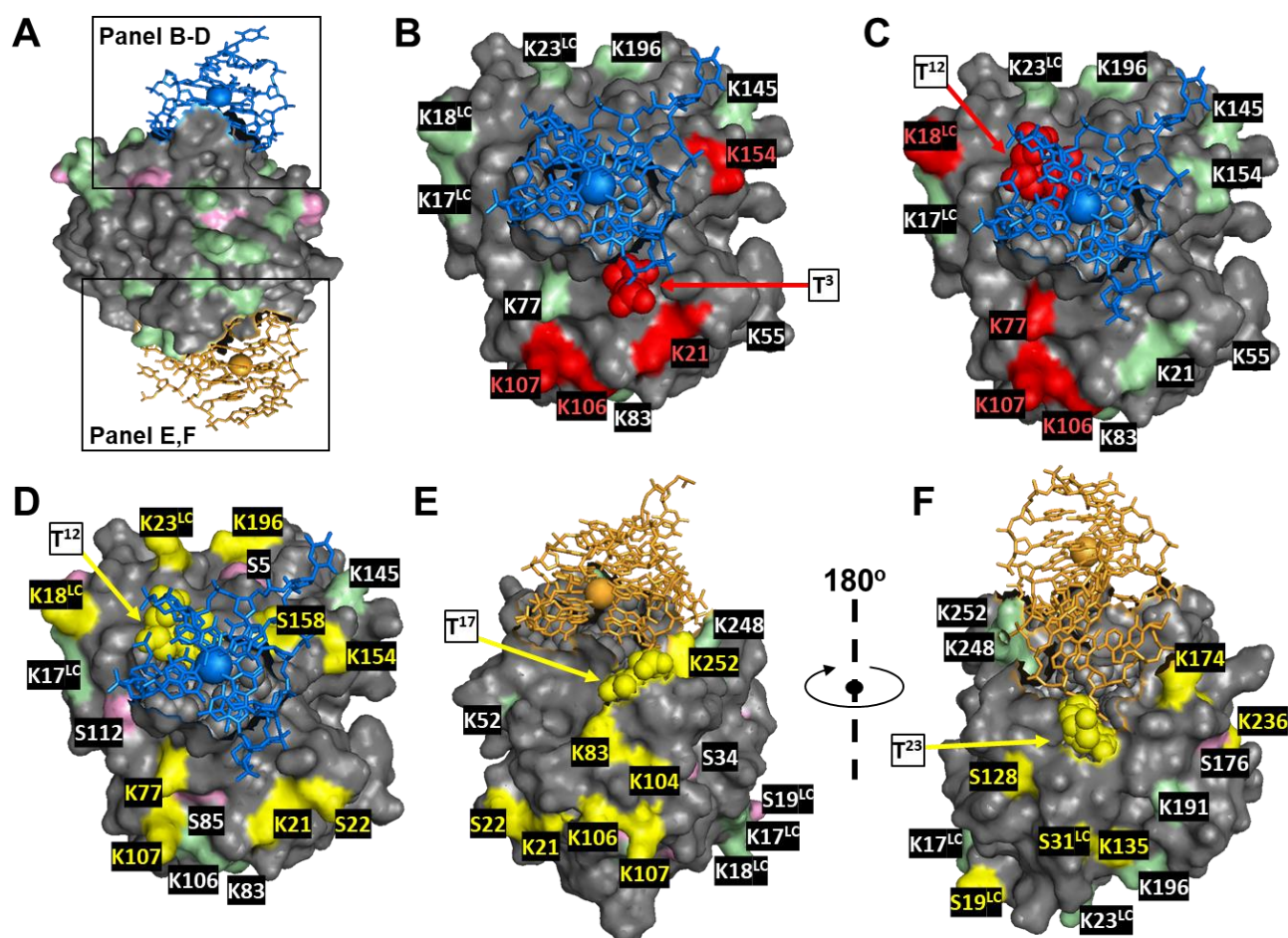


Figure 4.4. Site-selective modification of thrombin by catalyst-aptamer conjugates. The colour of the residues that are modified matches the colour that marks the catalyst position; the aptamer is shown in stick presentation (blue for TBA in panels A, B, D and E, and orange for TBA2 in panel F), the atoms of the base that is modified with the catalyst are shown as spheres. (A) Single crystal X-ray structure of thrombin with both TBA and TBA2 as well as exposed Lys (green) and Ser (pink) residues (PDB-code: 5EW1³⁴). (B–C) Thrombin residues modified with azido-thioester **1** by (A) TBA³-DMAP or (B) TBA¹²-DMAP in red. (D–E) Thrombin residues modified with azido-ANANS **3** by (D) TBA¹²-diPyOx, (E) TBA²¹⁷-diPyOx or (F) TBA²²³-diPyOx in yellow. Images were generated using PyMol.

Now that we identified a catalyst system that yields high conversions, we implemented another thrombin-binding aptamer (TBA2, which binds to the opposite side of thrombin), in order to pursue the modification of the other side of thrombin. Indeed, the TBA2-diPyOx catalysts modified thrombin at residues positioned in close proximity to the respective positions of the catalyst (Figure 4.4E), even though the conversions of 20–27% were substantially lower than for the other aptamer (Table A4.5). Nonetheless, our results show that different sides of the protein can be subjected to modification using different protein-binding aptamers. That direct conjugation of the catalyst to the protein binding aptamer is required for modification is shown by the absence of modification when the catalyst was tethered to a template strand that could hybridize with either the TBA or TBA2 aptamer that was extended with its complementary sequence (Figure A4.4). Conversions of TBA¹²-diPyOx and TBA²¹⁷-diPyOx that were studied in the presence of unmodified TBA and TBA2, revealed that aptamer affinity after catalyst functionalization is similar to their native counterparts (Figure A4.5). Additionally, TBA¹²-diPyOx conversion was unaffected by the presence of TBA2, neither was the activity of TBA²¹⁷-diPyOx hampered by TBA.

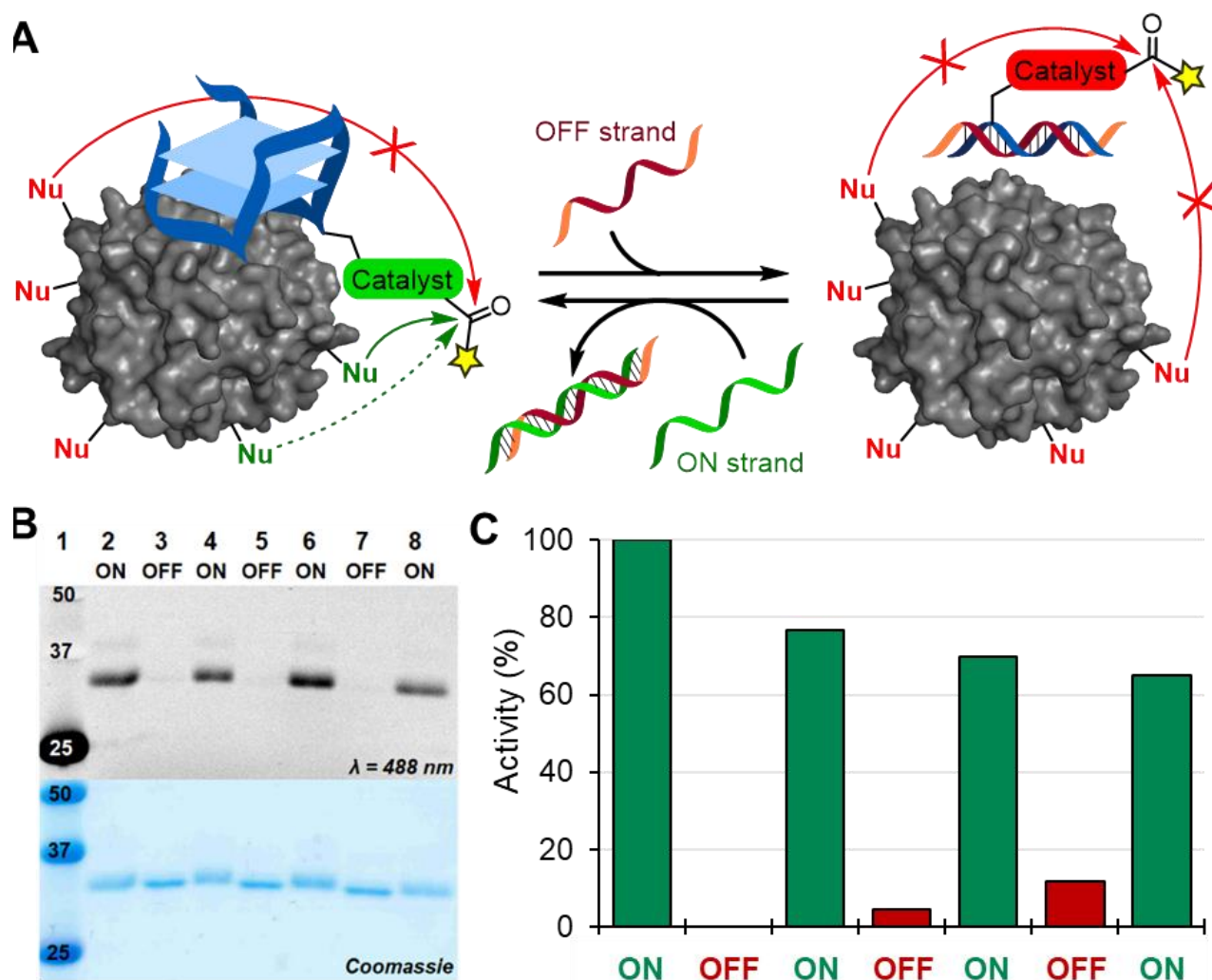


Figure 4.5. Activity control of TBA-catalyst constructs. (A) Schematic depiction of the activity switch where ssDNA OFF or ON strands suppress or regenerate, respectively, the protein-binding ability of the aptamer. In the OFF-state, the construct should not modify the protein. (B) SDS-PAGE analysis and visualization of the activity switch of TBA¹²-diPyOx after CuAAC modification of thrombin-bound alkyne-ANANS 4 with azido-lissamine. (C) The activity of TBA¹²-diPyOx could be switched OFF and ON for three cycles (quantities were determined from SDS-PAGE after modification with azido-ANANS 3 and BCN-PEG2000 using ImageJ software).

In our ambition to imitate the ability of enzymes to not only site-selectively modify proteins, but also do this in a trigger-dependent manner, we incorporate an activity-control switch in our catalytic constructs. For this, we designed an ssDNA OFF-strand that could hybridize with our TBA-catalyst constructs to form a dsDNA duplex that is unable to form the G-quadruplex structure that is essential to thrombin binding (Figure 4.5A). Indeed, upon addition of this complementary ssDNA OFF-strand, the system turned to its OFF state as no protein modification was detected (Figure 4.5B and A4.3). After removing the OFF-strand by addition of a second DNA strand that is fully complementary to the OFF strand (including its toeholds), the original TBA is reformed and the modification of thrombin is again efficient. We could use this switch to successfully control the activity of TBA¹²-DMAP, TBA¹²-PyOx and TBA¹²-diPyOx, and the activity of the latter could be regulated up to at least three full cycles (Figure 4.5C). The switch also worked *in situ*, where a 4 h reaction generated the same conversion as a system that was in OFF modus for 4 of the 8 h, even with different ON-OFF patterns, *i.e.*, 2–4–2 hrs ON-OFF-ON or 2–2–2–2 hrs ON-OFF-ON-OFF (Figure A4. 6).

4.4 Conclusions

We present the first DNA-based catalyst for the site-selective modification of a specific protein using its affinity for an aptamer. We show that both DMAP and PyOx can be used as acyl transfer catalysts, although the latter leads to significantly higher conversions and uses a less reactive substrate that suppresses uncatalyzed labelling. Furthermore, we show that nucleobases 3 and 12 of the TBA are the best sites for catalyst conjugation, and that nucleobase 17 is the best for TBA2. Both TBA and TBA2 catalytic constructs modified Lys and, to a lesser extent, Ser residues that are in proximity to the catalyst. Interestingly, we found that residue K145, which was modified by isolated DMAP, was not acylated by the DNA¹²-DMAP conjugate, but also found that acylation of other residues was achieved which were not modified by the isolated DMAP. This shows that the nanocatalyst constructs can overrule the inherent tendency of the isolated catalyst. Importantly, the activity of our DNA-based catalysts could be repeatedly regulated by an external stimulus. As protein modification efficiency, substrate specificity, site-selectivity, and external control over the activity are four essential features that define enzymatic protein modification,¹ our DNA-catalyst constructs are biomimetics of naturally occurring acylating enzymes. In view of the existence of aptamers for other proteins,¹³ we expect that our methodology will become applicable for other proteins in the future, and will be a valuable addition to the protein acylation toolbox.

Publication online:

This work is also published³⁶ and can be found here:



Chem. Commun. **2021**, 57, 12960–12963

References

- 1 J. J. Bruins, D. Blanco-Ania, V. Van Der Doef, F. L. Van Delft and B. Albada, *Chem. Commun.* **2018**, 54 (53), 7338–7341.
- 2 T. Wu, M. Liu, H. Huang, Y. Sheng, H. Xiao and Y. Liu, *Chem. Commun.* **2020**, 56 (65), 9344–9347.
- 3 J. J. Bruins, C. Van De Wouw, K. Wagner, L. Bartels, B. Albada and F. L. Van Delft, *ACS Omega* **2019**, 4 (7), 11801–11807.
- 4 M. A. Kasper, M. Glanz, A. Stengl, M. Penkert, S. Klenk, T. Sauer, D. Schumacher, J. Helma, E. Krause,

- M. C. Cardoso, H. Leonhardt and C. P. R. Hackenberger, *Angew. Chemie - Int. Ed.* **2019**, *58* (34), 11625–11630.
- 5 Y. Song, F. Xiong, J. Peng, Y. M. E. Fung, Y. Huang and X. Li, *Chem. Commun.* **2020**, *56* (45), 6134–6137.
- 6 O. Boutureira and G. J. L. Bernardes, *Chem. Rev.* **2015**, *115* (5), 2174–2195.
- 7 E. A. Hoyt, P. M. S. D. Cal, B. L. Oliveira and G. J. L. Bernardes, *Nat. Rev. Chem.* **2019**, *3* (3), 147–171.
- 8 N. C. Reddy, M. Kumar, R. Molla and V. Rai, *Org. Biomol. Chem.* **2020**, *18* (25), 4669–4691.
- 9 M. J. Matos, B. L. Oliveira, N. Martínez-Sáez, A. Guerreiro, P. M. S. D. Cal, J. Bertoldo, M. Maneiro, E. Perkins, J. Howard, M. J. Deery, J. M. Chalker, F. Corzana, G. Jiménez-Osés and G. J. L. Bernardes, *J. Am. Chem. Soc.* **2018**, *140* (11), 4004–4017.
- 10 V. Laserna, A. Istrate, K. Kafuta, T. A. Hakala, T. P. J. Knowles, M. Alcarazo and G. J. L. Bernardes, *Bioconjugate Chem.* **2021**, *32* (8), 1570–1575.
- 11 S. R. Adusumalli, D. G. Rawale, U. Singh, P. Tripathi, R. Paul, N. Kalra, R. K. Mishra, S. Shukla and V. Rai, *J. Am. Chem. Soc.* **2018**, *140* (44), 15114–15123.
- 12 C. A. Pérez Ríquez, O. Abian and J. M. Palomo, *Chem. Commun.* **2019**, *55*, 12928–12931.
- 13 J. F. Keijzer and B. Albada, *Bioconjugate Chem.* **2020**, *31* (10), 2283–2287.
- 14 S. Tsukiji, M. Miyagawa, T. Ogawa, Y. Koshi, I. Hamachi and E. Nakata, *J. Am. Chem. Soc.* **2007**, *130* (1), 245–251.
- 15 T. Tamura, Z. Song, K. Amaike, S. Lee, S. Yin, S. Kiyonaka and I. Hamachi, *J. Am. Chem. Soc.* **2017**, *139* (40), 14181–14191.
- 16 M. Kohlberger, S. Wildner, C. Regl, C. G. Huber and G. Gadermaier, *PLoS One* **2020**, *15* (11 November), 1–19.
- 17 L. Zhang, X. Fang, X. Liu, H. Ou, H. Zhang, J. Wang, Q. Li, H. Cheng, W. Zhang and Z. Luo, *Chem. Commun.* **2020**, *56* (70), 10235–10238.
- 18 J. L. Vinkenborg, G. Mayer and M. Famulok, *Angew. Chemie - Int. Ed.* **2012**, *51* (36), 9176–9180.
- 19 F. Rohrbach, F. Schäfer, M. A. H. Fichte, F. Pfeiffer, J. Müller, B. Pötzsch, A. Heckel and G. Mayer, *Angew. Chemie - Int. Ed.* **2013**, *52* (45), 11912–11915.
- 20 R. Stoltenburg, C. Reinemann and B. Strehlitz, *Biomol. Eng.* **2007**, *24* (4), 381–403.
- 21 S. Klusmann, *The Aptamer Handbook: Functional Oligonucleotides and Their Applications*, **2006**.
- 22 I. Russo Krauss, A. Merlino, A. Randazzo, E. Novellino, L. Mazzarella and F. Sica, *Nucleic Acids Res.* **2012**, *40* (16), 8119–8128.
- 23 A. Vasilescu, Q. Wang, M. Li, R. Boukherroub and S. Szunerits, *Chemosensors* **2016**, *4* (2), 1–20.
- 24 F. X. Montserrat, A. Grandas, R. Eritja and E. Pedroso, *Tetrahedron* **1994**, *50* (8), 2617–2622.
- 25 M. S. Kupryushkin, M. D. Nekrasov, D. A. Stetsenko and D. V. Pyshnyi, *Org. Lett.* **2014**, *16* (11), 2842–2845.
- 26 N. C. Seeman, *Annu. Rev. Biophys. Biomol. Struct.* **2002**, *27* (1), 225–248.
- 27 S. Wintermans, J. F. Keijzer, M. Dros, H. Zuilhof and B. Albada, *ChemCatChem* **2021**, *13*, 4618–4624.
- 28 T. A. Ngo, H. Dinh, T. M. Nguyen, F. F. Liew, E. Nakata and T. Morii, *Chem. Commun.* **2019**, *55* (83), 12428–12446.

- 29 T. Li, S. Dong and E. Wang, *J. Am. Chem. Soc.* **2010**, *132* (38), 13156–13157.
- 30 H. B. Albada, E. Golub and I. Willner, *Chem. Sci.* **2016**, *7* (5), 3092–3101.
- 31 G. Höfle, W. Steglich and H. Vorbrüggen, *Angew. Chemie Int. Ed. English* **1978**, *17* (8), 569–583.
- 32 M. S. Xie, B. Huang, N. Li, Y. Tian, X. X. Wu, Y. Deng, G. R. Qu and H. M. Guo, *J. Am. Chem. Soc.* **2020**, *142* (45), 19226–19238.
- 33 D. M. Tasset, M. F. Kubik and W. Steiner, *J. Mol. Biol.* **1997**, *272* (5), 688–698.
- 34 A. Pica, I. R. Krauss, V. Parente, H. Tateishi-Karimata, S. Nagatoishi, K. Tsumoto, N. Sugimoto and F. Sica, *Nucleic Acids Res.* **2017**, *45* (1), 461–469.
- 35 K. Shiraiwa, R. Cheng, H. Nonaka, T. Tamura and I. Hamachi, *Cell Chem. Biol.* **2020**, *27* (8), 970–985.
- 36 J. F. Keijzer, J. Firet and B. Albada, *Chem. Commun.* **2021**, *57*, 12960–12963.

Appendix 4

A4.1 DNA codes with their respective topologies and nucleotide sequences

Table A4.1. DNA sequence code and the respective sequences for each strand that was used in this study.

Code	DNA Sequence (5' to 3')
TBA	GGT TGG TGT GGT TGG
TBA-3-azide	GG/iAzideN/ TGG TGT GGT TGG
TBA-4-azide	GGT /iAzideN/GG TGT GGT TGG
TBA-7-azide	GGT TGG /iAzideN/GT GGT TGG
TBA-9-azide	GGT TGG TG/iAzideN/ GGT TGG
TBA-12-azide	GGT TGG TGT GG/iAzideN/ TGG
TBA-13-azide	GGT TGG TGT GGT /iAzideN/GG
TBA-3-alkyne	GG/i5OctdU/ TGG TGT GGT TGG
TBA-12-alkyne	GGT TGG TGT GG/i5OctdU/ TGG
TBA-3,12-dialkyne	GG/i5OctdU/ TGG TGT GG/i5OctdU/ TGG
TBA2 (HD22)	TCC GTG GTA GGG CAG GTT GGG GTG AC
TBA2-1-azide	/5AzideN/CC GTG GTA GGG CAG GTT GGG GTG AC
TBA2-5-azide	TCC G/iAzideN/G GTA GGG CAG GTT GGG GTG AC
TBA2-8-azide	TCC GTG G/iAzideN/A GGG CAG GTT GGG GTG AC
TBA2-17-azide	TCC GTG GTA GGG CAG G/iAzideN/T GGG GTG AC
TBA2-18-azide	TCC GTG GTA GGG CAG GT/iAzideN/ GGG GTG AC
TBA2-23-azide	TCC GTG GTA GGG CAG GTT GGG G/iAzideN/G AC
TBA2-26-azide	TCC GTG GTA GGG CAG GTT GGG GTG A/3AzideN/
TBA2-5'-end-alkyne	/5Hexynyl/TCC GTG GTA GGG CAG GTT GGG GTG AC
TBA2-5-alkyne	TCC G/i5OctdU/G GTA GGG CAG GTT GGG GTG AC
TBA2-17-alkyne	TCC GTG GTA GGG CAG G/i5OctdU/T GGG GTG AC
TBA2-23-alkyne	TCC GTG GTA GGG CAG GTT GGG G/i5OctdU/G AC
TBA2-26-alkyne	TCC GTG GTA GGG CAG GTT GGG GTG A/35OctdU/
hexynyl-TEMP1	/5Hexynyl/A AAA TAT ATA TAT ATA AAA
TEMP2-TBA	TTT TAT ATA TAT ATA TTT T GGT TGG TGT GGT TGG
TEMP2-TBA2	TTT TAT ATA TAT ATA TTT T TCC GTG GTA GGG CAG GTT GGG GTG AC
TBA-OFF	ATG CCC AAC CAC ACC AAC CAT GC
TBA-ON	GCA TGG TTT GTG TGT TTG GGC AT

* Note: 'iAzideN' indicates internal azide-functionalized thymine (see Figure 4. 2B for details); 'i5OctdU' indicates internal alkyne-functionalized thymine (see Figure 4. 2C for details); '5Hexynyl' indicates alkyne modification on the 5'-end; 3AzideN and 5AzideN indicate a 3'-end or 5'-end positioned 6-azido-hexanoic amide.

A4.2 Protocol for modification of thrombin with DNA-bound acyl transfer catalyst

A mixture was typically prepared containing 10 μ M thrombin (from a 200 μ M stock solution in 50% glycerol in ddH₂O), 30 μ M DNA construct (from varying stock concentrations in ddH₂O) and in HEPES buffer [50 mM, pH=8.0, with 350 mM NaCl and 50 mM KCl]. This mixture was incubated in the dark for 60 min at 37 °C, after which acyl donor (from varying stock concentrations in DMSO) was added. The reaction mixture was again incubated in the dark at 37 °C, shaking the tubes at 500 rpm (reaction times: 2 hrs for DMAP catalyst, 5 hrs for PyOx).

Prior to SDS-PAGE analysis, additional functionalization is required to visualize the modifications. Two approaches were used: band shifting or fluorescent staining.

- **Band shifting:** thrombin modified with an azide-carrying acyl donor was treated with 6 equivalents of BCN-PEG2000 (purchased from Synaffix B.V.) with respect to the concentration of acyl donor and incubated at 12 °C overnight.
- **Fluorescent staining:** Thrombin modified with an alkyne-carrying acyl donor was treated with 6 equivalents of azido-sulphorhodamine B (purchased from Tenova Chemicals B.V.) with respect to the acyl donor, 100 μ M [Cu•THPTA] (complex of CuSO₄ and tris(3-hydroxypropyl)triazolylmethyl)amine mixed in a ratio of 1:5 in ddH₂O) and 1 mM sodium ascorbate (from a freshly made stock of 10 mM in ddH₂O) and incubated at 12 °C overnight.

A4.3 Protocol for switchable activity of DNA-bound acyl transfer catalyst

A mixture was typically prepared containing 10 μ M thrombin (from a 200 μ M stock solution in 50% glycerol in ddH₂O), 30 μ M DNA construct (from varying stock concentrations in ddH₂O) and in HEPES buffer [50 mM, pH=8.0, with 350 mM NaCl and 50 mM KCl]. This mixture was incubated in the dark for 60 min at 37 °C. To switch OFF the activity of TBA-catalyst construct, 1.2 equivalents of TBA-OFF DNA was added after which the mixture was incubated at 37 °C for 15 min. Similarly, the TBA-catalyst construct was switched back ON by adding 1.5 equivalents (with respect to the original GQ concentration, *i.e.*, 1.25 equivalents with respect to the OFF strand) of TBA-ON DNA after which the mixture was incubated at 37 °C for 15 min. The acyl donor (from varying stock concentrations in DMSO) was added after each switching event, and the reaction mixture was incubated in the dark at 37 °C, shaking the tubes at 500 rpm. Reaction times were 2 h when using the DMAP catalyst and 5 h when using the PyOx catalyst.

A4.4 Protocol for SDS-PAGE analysis

Acrylamide gels (12%) were prepared according to Bio-Rad bulletin 6201 protocol. Specifically, reaction mixtures containing 2–5 μ g of protein were diluted with one volume equivalent of SDS-PAGE sample buffer (2 \times) containing 10% BME and incubated for 10 minutes at 95 °C. The denatured sample was then used for SDS-PAGE analysis (12% acrylamide gel). Precision Plus Protein™ Dual Color Standards was used as a reference protein ladder. After running, if one of the proteins was modified with a fluorophore, a UV-photo of the gel was taken. Gels were then stained using Coomassie brilliant blue (0.1% Coomassie Blue R250 in 10% acetic acid, 50% methanol and 40% demineralized water) by shaking gently for 0.5 hours, and destained with destaining solution (10% acetic acid, 50% methanol, and 40% demineralized water) by shaking gently for 1 hour. Afterwards, the destaining solution was replaced with H₂O and shaken gently overnight at room temperature. When the BCN-PEG2000 mass-tag was used, quantification was performed by integrating the intensity of the Coomassie stained bands of the SDS-PAGE gel using ImageJ software.

A4.5 Protocol for the analysis of protein modification on HPLC(-MS)

The reaction mixture was aspirated three times with a pipette, after which 10 μ L was added to an HPLC vial insert that already contained 10 μ L of buffer (200 mM citrate and 400 mM NaCl; pH: 5.5). The resulting mixture was also aspirated three times. This sample was then run over a Thermo Fischer MAbPAC RP column 3.0 \times 100 mm, at 80 $^{\circ}$ C, with the gradient starting with 23% (ACN + 0.1% FA) and ending with 33% (ACN + 0.1% FA) in (95% H₂O + 5% ACN + 0.1% FA) (flow rate 0.5 mL/min) over 25 min. The system used was an Agilent 1220 Infinity LC system with DAD detector.

For mass spectrometry analysis, reaction mixtures were diluted to a final protein concentration of 0.25 mg/mL. Protein samples were then analysed on a Thermo Scientific™ Q Exactive Focus Orbitrap using the same gradient as was used for the HPLC analyses.

A4.6 Tryptic digestion of protein and subsequent analysis to determine site-specificity of the modification

Modified protein samples were subjected to SDS-PAGE separation and the desired protein bands excised from the gel and cut up to small pieces. The pieces were washed by incubating three times with 50 mM NH₄HCO₃ (pH: 8.0) in 50% ACN in ddH₂O and subsequently dried in a Speedyvac vacuum centrifuge. The dry pieces were swollen in 50 μ L DTT [10 mM in 100 mM NH₄HCO₃ (pH: 8.0)] and incubated for 45 minutes at 56 $^{\circ}$ C. The supernatant was removed and 50 μ L of IAA (55 mM in 100 mM NH₄HCO₃ (pH: 8.0)) was added and the pieces were incubated in the dark at rt for 30 min. The supernatant was removed and the pieces were washed by incubating once with 50 mM NH₄HCO₃ (pH: 8.0) in 50% ACN in ddH₂O and subsequently dried in a vacuum centrifuge. The gel pieces were swollen in 40 μ L trypsin gold (125 ng/ μ L) and incubated at 37 $^{\circ}$ C for 16–18 h. The initial supernatant was collected and the gel pieces were washed by incubating 15 min at 37 $^{\circ}$ C with 20 μ L NH₄HCO₃ (100 mM, pH: 8.0) and 15 min at 37 $^{\circ}$ C when diluted with 20 μ L. The collected supernatants were combined and dried in a vacuum centrifuge and the dry peptide digest dissolved in 20 μ L 0.1% FA.

Peptide digests were analysed on an EASY nanoLC connected to Thermo Scientific™ Q Exactive PLUS. Peptides were trapped onto a PepSep trap column (2 cm \times 100 μ m ID, 5 μ m C18 ReproSil) and subsequently separated on a PepSep analytical column (8 cm \times 75 μ m ID, 3 μ m C18 ReproSil, PepSep). Elution was achieved using a gradient that started with 5% (ACN + 0.1% FA) ending with 40% (ACN + 0.1% FA) in (H₂O + 0.1% FA), washing the column with 80% (ACN + 0.1% FA) afterwards.

The eluted peaks were analysed using MaxQuant software, searching for peptides with mass modification corresponding to H(11)O(3)C(7)N(3) (*i.e.*, the substitution of a proton on the protein by the acyl group of thioester **1** or ANANS **3**) and limiting criteria of 1% PSM FDR and a minimal peptide score of 80. As protein database human proteome was used, obtained from www.unitprot.org (code: UP000005640).

A4.7 TBA-DMAP selectivity data

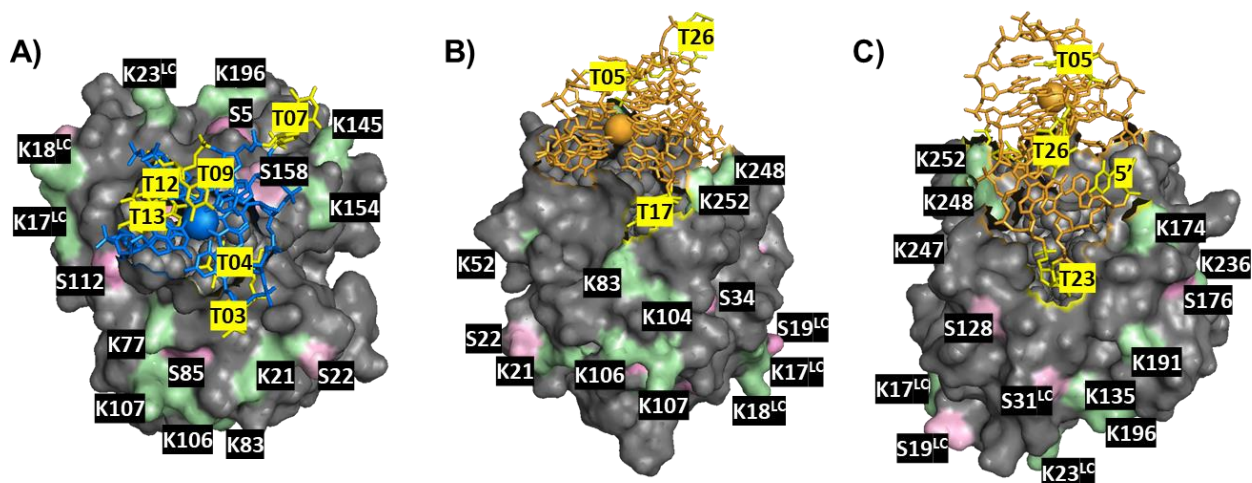


Figure A4.1. Crystal structures of thrombin complexed with its two aptamers TBA and TBA2 (a.k.a. HD22). (A) TBA and its proximal residues and (B/C) TBA2 and proximal residues on either side of the aptamer. Lysine residues are depicted in light green, serine residues are depicted in pink, TBA is shown blue, TBA2 is shown orange and thymine nucleobases to which catalysts were attached are shown in yellow. (PDB-code: 5EW1)³⁴

Table A4.2. Conversions (calculated from SDS-PAGE data) and targeted residues for the different TBA-DMAP constructs, in comparison to the absence of catalyst or with free DMAP. The red **X** highlights modifications that are also found in the absence of catalyst or with free DMAP and are background acylation. Conditions: 10 μ M thrombin, 30 μ M catalyst, 150 μ M azide-thioester **1**, pH: 8.0, 37 $^{\circ}$ C, 2 h. Numerical data regarding these tryptic digestion results can be found in the Supporting Information of the published work.³⁶

Catalyst	Conv. (%)	K21	K77	K83	K106	K107	K145	K154	K196	K236	K252	K18 ^{LC}	K23 ^{LC}
-	5–8			X			X		X	X	X		X
free DMAP	5–10			X			X		X	X	X		X
TBA ³ -DMAP	27	X		X	X	X	X	X	X	X	X		X
TBA ⁴ -DMAP	12	X		X	X				X	X	X		X
TBA ⁷ -DMAP	16			X			X	X	X	X	X		X
TBA ⁹ -DMAP	8			X					X	X	X		X
TBA ¹² -DMAP	36–49		X	X	X	X			X	X	X	X	X
TBA ¹³ -DMAP	17		X	X	X	X			X	X	X	X	X

A4.8 TBA-DMAP conversion data

Table A4.3. Conversions (calculated from SDS-PAGE data) of all DMAP and diDMAP constructs as calculated from SDS-PAGE results using ImageJ. Conditions: 10 μ M thrombin, 30 μ M catalyst, azido-thioester 1, pH: 8.0, 37 $^{\circ}$ C, 2 h.

Catalyst	1 (μ M)	Conv (%)	Catalyst	1 (μ M)	Conv (%)
-	150	5–8	-	300	12–20
DMAP	150	5–10	DMAP	300	18–20
TBA ³ -DMAP	150	27	TBA ¹² -DMAP	300	53–61
TBA ⁴ -DMAP	150	12	TBA ¹² -diDMAP	300	59–65
TBA ⁷ -DMAP	150	16	TBA2 ¹ -DMAP	300	22
TBA ⁹ -DMAP	150	8	TBA2 ⁵ -DMAP	300	12
TBA ¹² -DMAP	150	36–49	TBA2 ⁸ -DMAP	300	21
TBA ¹³ -DMAP	150	17	TBA2 ¹⁷ -DMAP	300	14–18
TBA ¹² -diDMAP	150	48–61	TBA2 ¹⁸ -DMAP	300	19
TBA2 ¹⁷ -DMAP	150	5	TBA2 ²³ -DMAP	300	20
TBA2 ¹⁷ -diDMAP	150	9	TBA2 ²⁶ -DMAP	300	10
			TBA2 ¹⁷ -diDMAP	300	15

A4.9 DNA-(di)PyOx selectivity data

Table A4.4. Details of the modification by the various catalyst-aptamer constructs, including conversion and the position(s) of the modification(s). The red **X** highlights modifications that are also found in the absence of catalyst or with free PyOx and are background acylation. Conditions: 10 μ M thrombin, 30 μ M catalyst, 300 μ M azido-ANANS 3, 37 $^{\circ}$ C, 6 h. Numerical data regarding these tryptic digestion results can be found in the Supporting Information of the published work.³⁶

pH	Catalyst		Lysine										Serine		
	Code	Conv (%)	21	77	83	106	107	145	154	196	18 ^{LC}	23 ^{LC}	5	22	158
7.2	free PyOx	0					X	X							
	TBA ³ -PyOx	18	X	X	X	X	X	X	X					X	X
	TBA ¹² -PyOx	28	X	X	X	X	X	X	X					X	X
	TBA ³ -diPyOx	71–93	X			X	X	X		X		X	X	X	X
	TBA ¹² -diPyOx	75–91	X	X			X	X	X	X	X	X		X	X
	TBA ^{3,12} -bis(diPyOx)	73	not analyzed												
8.0	free PyOx	2		X			X	X							
	TBA ³ -PyOx	29	X	X	X	X	X	X	X		X			X	X
	TBA ¹² -PyOx	38	X	X	X	X	X	X	X		X			X	X
	TBA ¹² -diPyOx	83	X	X		X	X	X	X	X	X	X			X
	TBA ^{3,12} -bis(diPyOx)	84	not analyzed												

A4.10 TBA2-diPyOx selectivity data

Table A4.5. Conversions (calculated from SDS-PAGE data) and modified residues for different TBA2-diPyOx constructs, in comparison to free diPyOx. The red **X** highlights modifications that are also found in the absence of catalyst or with free diPyOx and are background acylation. Conditions: 10 μ M thrombin, 30 μ M catalyst, 300 μ M azido-ANANS **3**, pH: 7.2, 37 $^{\circ}$ C, 6 h. Numerical data regarding these tryptic digestion results can be found in the Supporting Information of the published work.³⁶

Catalyst		Lysine													Serine			
Code	Conv (%)	21	77	83	104	106	107	135	145	174	236	247	248	252	22	128	19 ^{LC}	31 ^{LC}
free diPyOx	0		X						X									
TBA2 ^{5end} -diPyOx	15		X					X	X	X	X							X
TBA2 ⁵ -diPyOx	6	X	X	X			X		X	X		X	X	X	X			
TBA2 ¹⁷ -diPyOx	27	X	X	X	X	X	X		X					X	X			
TBA2 ²³ -diPyOx	20		X					X	X	X	X					X	X	X
TBA2 ²⁶ -diPyOx	20		X	X		X	X		X	X	X	X	X	X			X	

A4.11 TBA-diPyOx substrate specificity

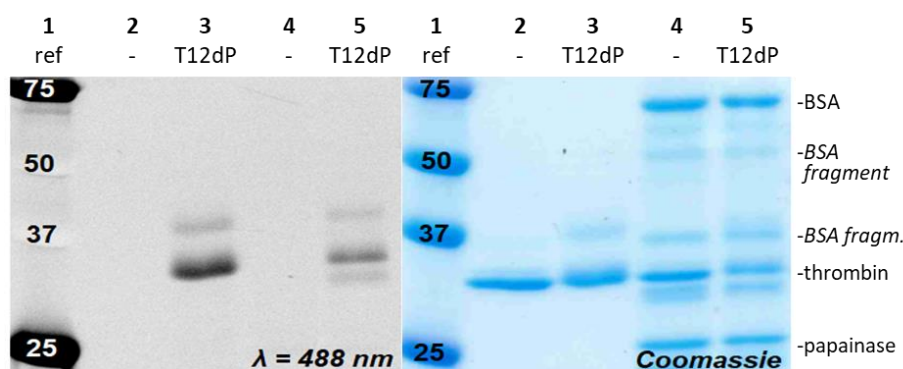


Figure A4.2. SDS-PAGE results of thrombin modified by TBA¹²-diPyOx with alkyne-ANANS **4** and subsequent CuAAC with an azido-PEG-lissamine. The only band that becomes fluorescent originates from thrombin, indicating specificity of TBA¹²-diPyOx over the proteins BSA and papainase.

A4.12 TBA-DMAP activity switch

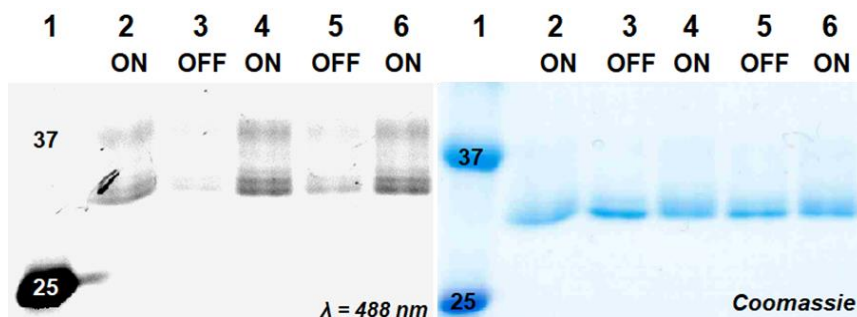


Figure A4.3. SDS-PAGE results of thrombin modified with alkyne-thioester **2** and subsequent CuAAC with an azido-PEG-lissamine. Lane 2: TBA¹²-DMAP; Lane 3: TBA¹²-DMAP + TBA-OFF; Lane 4: TBA¹²-DMAP + TBA-OFF + TBA-ON; Lane 5: TBA¹²-DMAP + TBA-OFF + TBA-ON + TBA-OFF; Lane 6: TBA¹²-DMAP + TBA-OFF + TBA-ON + TBA-OFF + TBA-ON. Conditions: 10 μ M thrombin, 30 μ M TBA¹²-DMAP, 150 μ M thioester substrate **2**, pH: 8.0, 37 $^{\circ}$ C, 2 h.

A4.13 Templated catalyst approach

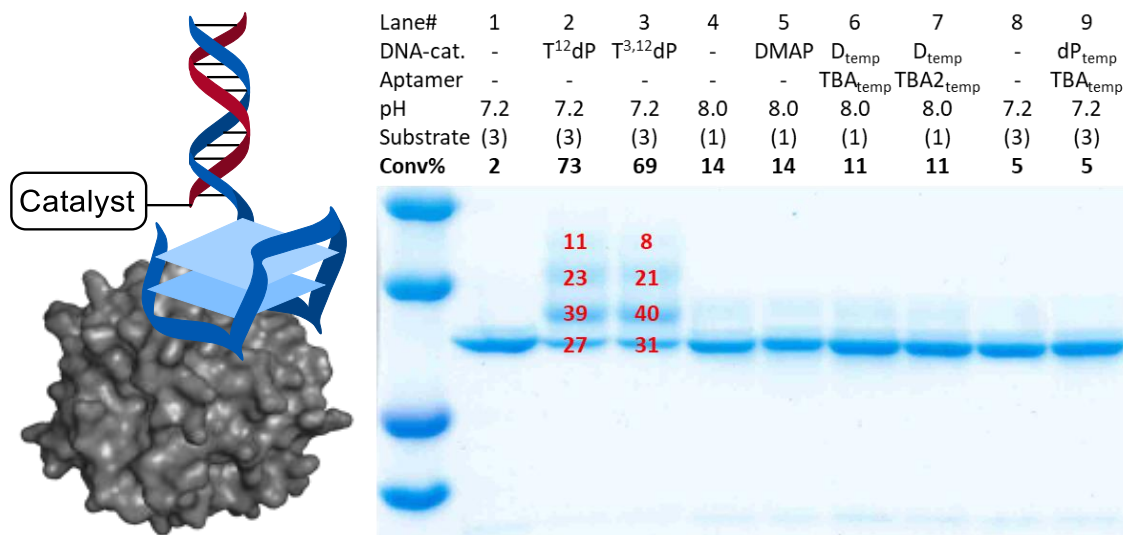


Figure A4.4. SDS-PAGE results of thrombin modified with azido-thioester **1** or azido-ANANS **3** and subsequently SPAACed with a BCN-functionalized 2 kDa masstag. Lane 1–3: comparison of diPyOx and bis(diPyOx). Lanes 4–7: DNA-templated DMAP acylation of thrombin. Lanes 8–9: DNA-templated diPyOx acylation of thrombin. Conditions: 10 μ M thrombin, 30 μ M DNA, 300 μ M acyl donor, 37 $^{\circ}$ C in the dark for 6 h. The bottom row displays the modification percentages of each lane as calculated with ImageJ, when possible, this is indicated for individual bands in red. Abbreviations: T¹²dP = TBA¹²-diPyOx; T^{3,12}dP = TBA^{3,12}-bis(diPyOx); D_{temp} = DMAP-template; TBA_{temp} = TBA-template; TBA2_{temp} = TBA2-template; dP_{temp} = diPyOx-template.

A4.14 TBA(2)-catalyst vs TBA(2) competition experiment

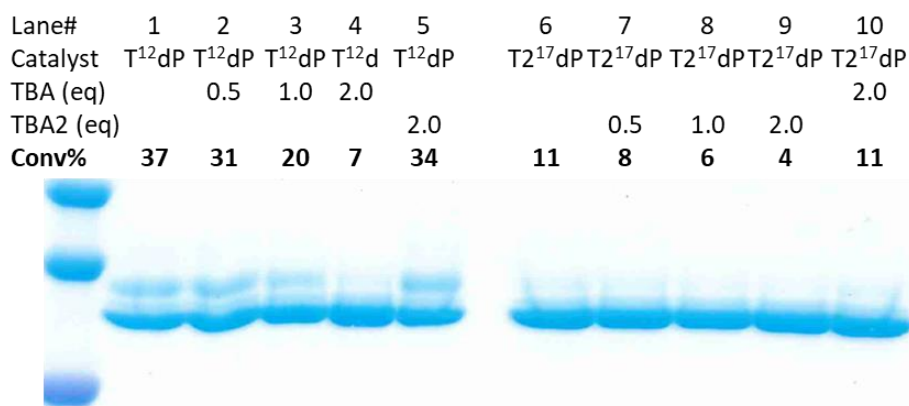


Figure A4.5. SDS-PAGE results of thrombin modified with azido-thioester **1** or azido-ANANS **3** and subsequently SPAACed with a BCN-functionalized 2 kDa masstag. Varying equivalents of native aptamer were added to observe the effect on the yield and estimate whether the affinity of the modified aptamer was affected by the catalyst. Conditions: 20 μ M thrombin, 30 μ M DNA, 300 μ M acyl donor, 37 $^{\circ}$ C in the dark for 8 h. The changes in modification percentages indicate that the affinity of the catalyst-aptamer construct is comparable to that of the native aptamer. Reported conversions were calculated with ImageJ.

A4.15 TBA-diPyOx activity switch *in situ*

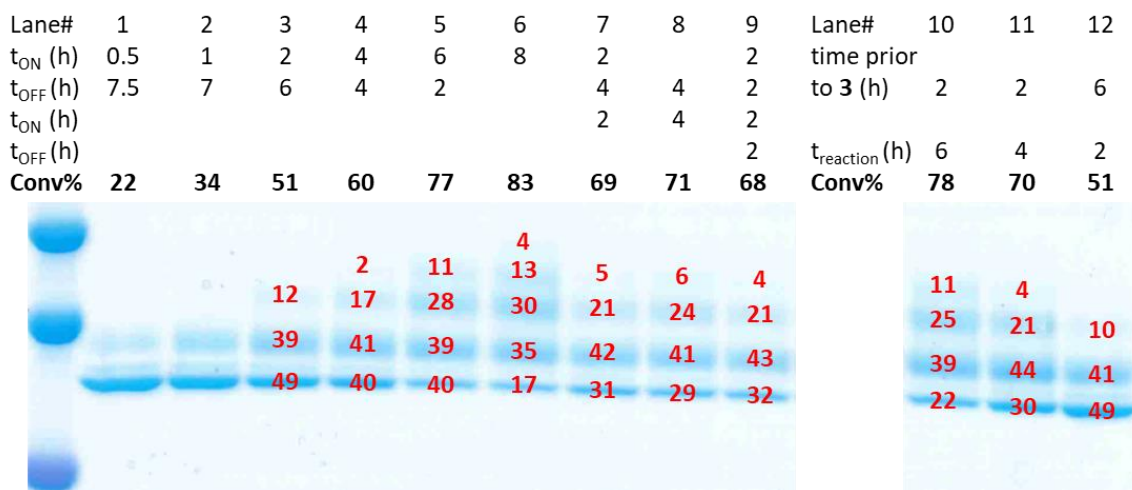


Figure A4.6. SDS-PAGE results of thrombin modified with azido-ANANS **3** and subsequently SPAACed with a BCN-functionalized 2 kDa masstag. Lane 1–6: reactions run with TBA¹²-diPyOx for ‘x’ hours in ON and stopped with TBA-OFF after ‘y’ hours. Lanes 7–9: similar to 1–6, but with more *in situ* switch steps. Lane 10–12: azido-ANANS **3** was added after waiting for ‘x’ hours to compare ‘x’ hour reaction time to ‘x’ hour OFF time. Conversions match, thus the switch works. Conditions: 10 μ M thrombin, 30 μ M DNA, 1.1 eq TBA ON/OFF, 300 μ M acyl donor **3**, 37 °C in the dark for 8 h. The bottom row displays the modification percentages of each lane as calculated with ImageJ. Red numbers in the gel indicate the various percentages of labelled products.

A4.16 Synthesis of TBA-DMAP constructs

DNA sequences containing azido-thymine modification were purchased as HPLC-purified lyophilized powders from Integrated DNA Technologies (IDT). The DNA was treated with 50 equivalents of compound **1** with respect to the DNA concentration and the reaction mixture was incubated in the dark at 4 °C for 16–20 hours. The synthesized DNA construct was purified by spin-filtration over 3 kDa MWCO Amicon® Ultra-15 Centrifugal Filter Units, washing 5 times with 400 mM NaCl solution in ddH₂O. Purity and concentration were determined by HPLC-MS and UV-Vis. (Schematic in Figure 4. 2A–B)

A4.17 Synthesis of DNA-PyOX, DNA-diPyOx and DNA-diDMAP

DNA sequences with alkyne-thymine modification were purchased as HPLC-purified lyophilized powders from Integrated DNA technologies. The powders were dissolved in oxygen-poor ddH₂O. The DNA was treated with 10 equivalents of alkyne-PyOx, azido-diDMAP or azido-diPyOx (from 100 mM stock in DMSO) with respect to the DNA concentration, 100 μ M [Cu•THPTA] (complex of CuSO₄ and THPTA mixed in a ratio of 1:5 in ddH₂O) and 10 mM sodium ascorbate (from a freshly made stock of 100 mM in ddH₂O) and incubated in the dark at 12 °C for 16–20 hours. The synthesized DNA construct was purified by spin-filtration over 3 kDa MWCO Amicon® Ultra-15 Centrifugal Filter Units, washing 3 times with 400 mM NaCl solution in ddH₂O. Purity and concentration were determined by HPLC-MS and UV-Vis. (Schematic in Figure 4. 2A–C)

A4.18 ImageJ software

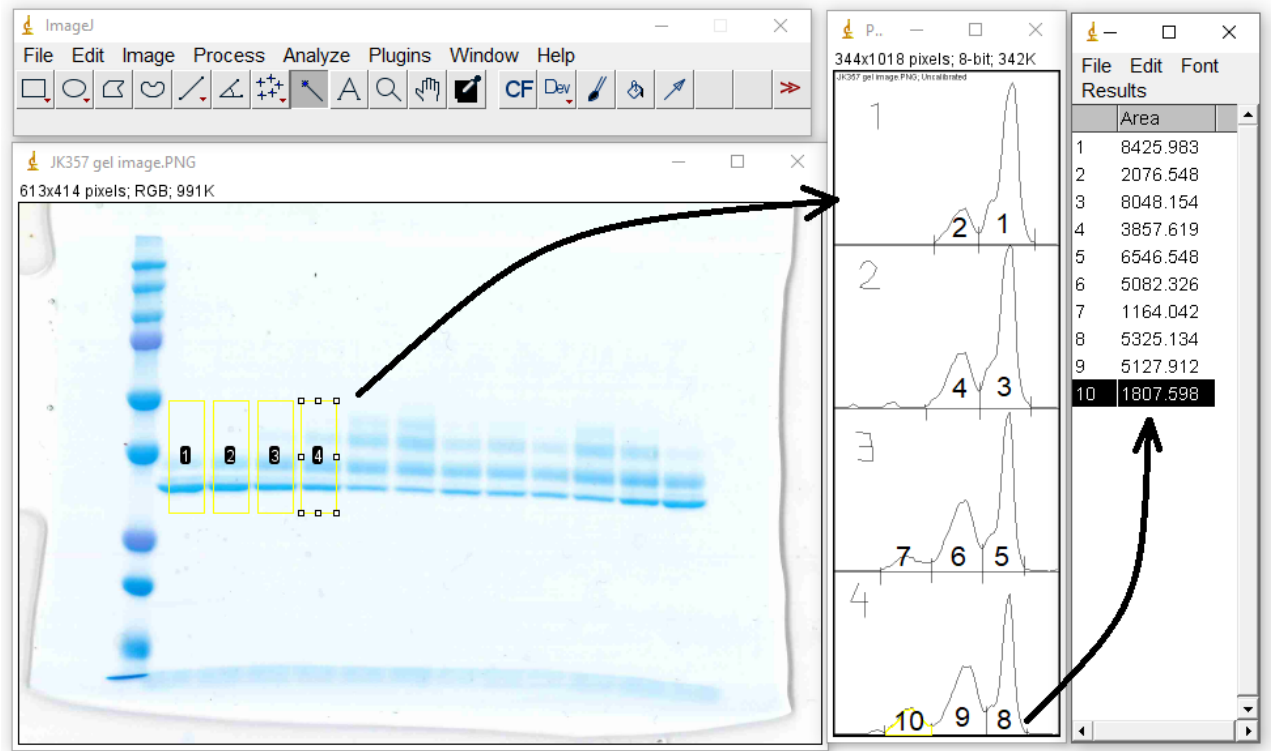
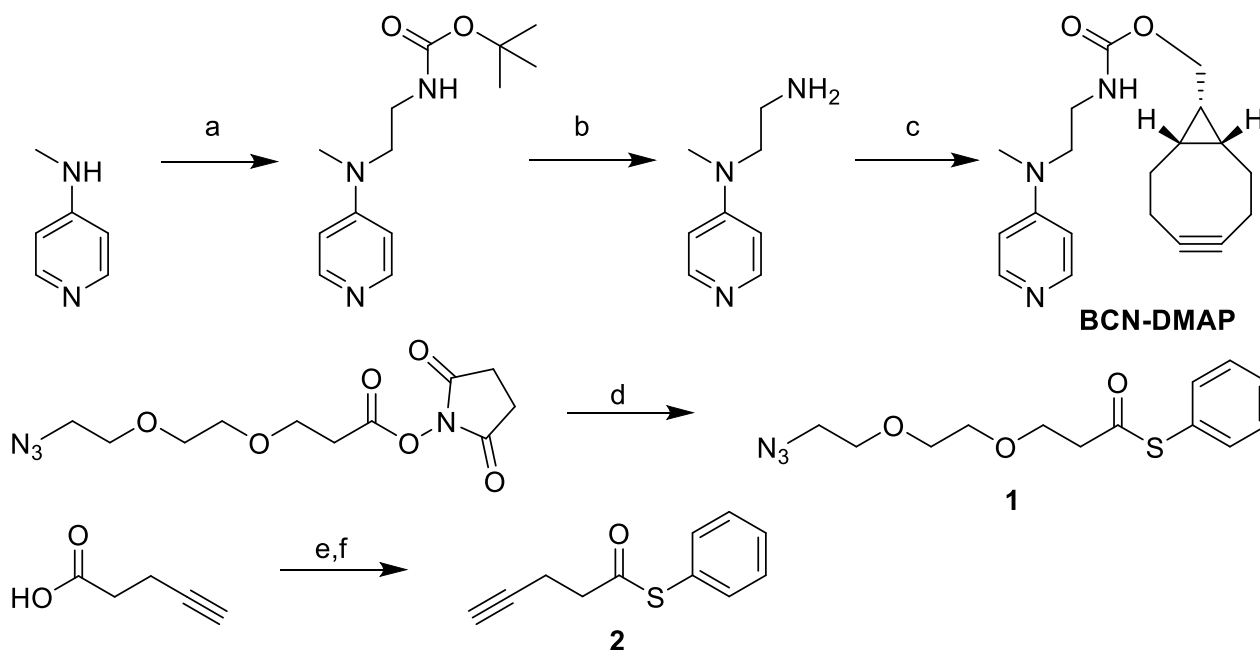


Figure A4.7. Example of integration with ImageJ, where pixel density is integrated to obtain values for calculation of relative percentages of protein concentration on the gels. The analysis starts by uploading a scan image of a gel into the program. Square boxes are then drawn to determine the areas in which the pixel density needs to be integrated, resulting in graphs for each of the boxes. In these graphs, the peaks match with the bands from bottom to top and after separating the peaks into separate areas, the program can calculate the areas of these peaks. These values can be used to determine the ratio of modification.

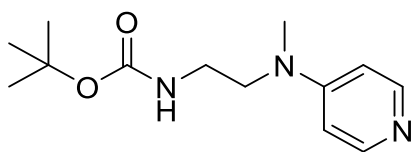
A4.19 Synthesis of organic compounds

Synthesis of BCN-DMAP and its alkyne-/azide-functionalized thioester substrates



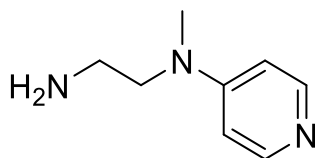
Scheme A4.1. Synthesis of BCN-DMAP and thioesters **1** and **2**. Reaction conditions: (a) NaH, *tert*-butyl (2-bromoethyl)carbamate, anhyd. THF, rt, 24 h, **9%**; (b) TFA (60 vol%), DCM, rt, 36 h, **quant**; (c) BCN-succinimidyl carbonate, NEt₃, anhyd. THF, rt, 18 h, **69%**; (d) thiophenol, toluene, rt, 16 h, **85%**; (e) isobutyl chloroformate, NEt₃, DCM, rt, 10 min; (f) thiophenol, NEt₃, DCM, rt, 16 h, **54%**.

Boc-ethylamine-DMAP:



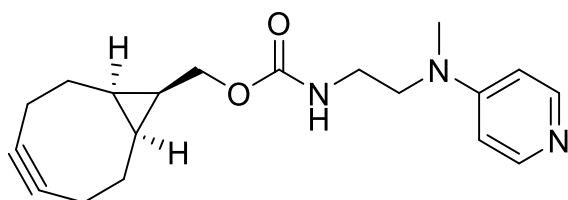
4-(methylamino)-pyridine (200 mg, 1.85 mmol, 1.0 eq) and NaH (592 mg, 14.8 mmol, 8.0 eq) were dissolved in 2 mL of anhyd. THF and the resulting mixture was stirred at rt for 90 min until H₂ evolution was no longer evident. Then, a solution of *tert*-butyl (2-bromoethyl)carbamate (622 mg, 2.77 mmol, 1.5 eq) in 1 mL anhyd. THF was added, and the resulting mixture was stirred at rt for 24 h. The reaction was quenched with saturated NH₄Cl solution (5 mL) and the product was extracted with DCM (6 × 10 mL). The combined organic layer was dried over Na₂SO₄ and concentrated under reduced pressure. The target compound was purified via flash column chromatography (SiO₂, eluent: 5–9% [10% (25% NH₃ in H₂O) in MeOH] in DCM), yielding a white solid (41 mg, 9%). HRMS (ESI): calculated for [M+H]⁺ 252.1712; found: 252.1709. ¹H NMR (400 MHz, CDCl₃) δ 8.17 (d, *J* = 5.8 Hz, 2H), 6.51 (d, *J* = 5.8 Hz, 2H), 4.86 (s, 1H), 3.48 (dd, *J* = 12.2, 5.8 Hz, 2H), 3.28 (q, *J* = 6.4 Hz, 2H), 2.97 (s, 3H), 1.41 (s, 9H) ppm. ¹³C NMR (101 MHz, CDCl₃) δ 156.1, 153.7, 149.7, 106.5, 79.5, 50.7, 37.9, 37.6, 28.4 ppm.

Ethylamine-DMAP:



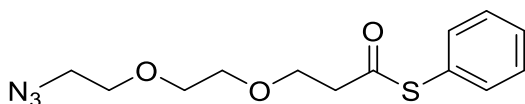
Tert-butyl (2-(methyl(pyridin-4-yl)amino)ethyl)carbamate (41 mg, 163 μ mol, 1.0 eq) was dissolved in 1.2 mL CH_2Cl_2 and cooled to 0 $^\circ\text{C}$. Then, trifluoroacetic acid (1.8 mL 60 vol%) was added dropwise and the resulting mixture was stirred at rt for 36 h. Afterwards, the mixture was concentrated under reduced pressure and the target compound was purified via flash column chromatography [SiO_2 , eluent: 6–14% (10% NH_3 in MeOH) in DCM], yielding a light-brown sticky oil (20 mg, 81%). ^1H NMR (400 MHz, D_2O) δ 8.04–7.97 (m, 2H), 6.61 (d, J = 6.0 Hz, 2H), 3.41 (t, J = 6.7 Hz, 2H), 2.94 (s, 3H), 2.79 (t, J = 6.2 Hz, 2H) ppm. ^{13}C NMR (101 MHz, D_2O) δ 156.3, 147.8, 106.1, 53.0, 37.7, 35.8 ppm.

BCN-DMAP:



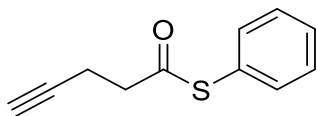
*N*¹-methyl-*N*¹-(pyridin-4-yl)ethane-1,2-diamine (20 mg, 132 μ mol, 1.0 eq) and triethylamine (148 μ L, 1.06 mmol, 8.0 eq) were dissolved in 0.7 mL anhydrous THF and the resulting mixture was stirred at rt for 10 min. To this, a solution of BCN-succinimidyl carbonate (47.8 mg, 164 μ mol, 1.24 eq) in 0.7 mL anhydrous THF was added dropwise, after which the mixture was allowed to warm up to rt and stirred at rt for an additional 18 h. The mixture was concentrated under reduced pressure and the target compound was purified via flash column chromatography (SiO_2 , eluent: 5–7% [10% (25% NH_3 in H_2O) in MeOH] in DCM), yielding a colourless sticky oil (30 mg, 69%). ^1H NMR (400 MHz, CDCl_3) δ 8.19–8.11 (m, 2H), 6.54–6.48 (m, 2H), 5.27 (t, J = 6.0 Hz, 1H), 4.11 (d, J = 8.1 Hz, 2H), 3.48 (t, J = 6.7 Hz, 2H), 3.33 (q, J = 6.5 Hz, 2H), 2.97 (s, 3H), 2.32–2.03 (m, 6H), 1.53 (q, J = 11.2, 10.4 Hz, 2H), 1.30 (p, J = 8.7 Hz, 1H), 0.90 (t, J = 10.0 Hz, 2H) ppm. ^{13}C NMR (101 MHz, CDCl_3) δ 157.0, 153.7, 149.7, 106.6, 98.9, 63.1, 50.7, 38.2, 37.7, 29.1, 21.5, 20.2, 17.8 ppm.

Azido-thioester (1):



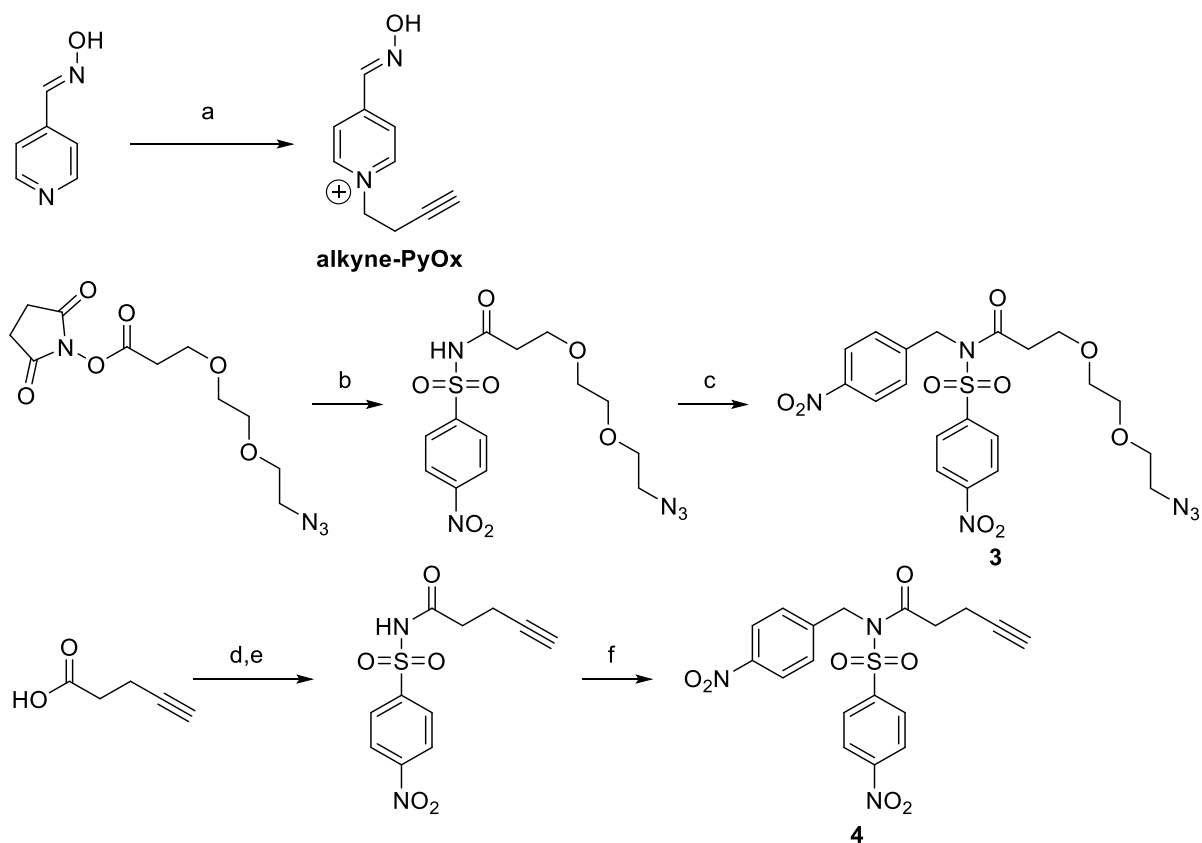
2,5-dioxopyrrolidin-1-yl 3-(2-(2-azidoethoxy)ethoxy)propanoate (150 mg, 500 μ mol, 1.03 eq) and thiophenol (49 μ L, 480 μ mol, 1.0 eq) were dissolved in 2 mL toluene, followed by dropwise addition of triethylamine (81 μ L, 580 μ mol, 1.2 eq). The resulting mixture was stirred at rt for 16 h, after which the reaction was quenched with brine (8 mL) and the product was extracted with EtOAc (3 \times 6 mL). The organic layer was then dried over Na_2SO_4 and concentrated under reduced pressure. The target compound was purified via flash column chromatography (SiO_2 , 20% EtOAc in petroleum ether (40/60)), yielding a colourless oil (121 mg, 85%). HRMS (ESI) calculated for $[\text{M}+\text{Na}]^+$: 318.0888; found $[\text{M}+\text{Na}]^+$: 318.0877. ^1H NMR (400 MHz, CDCl_3) δ 7.45–7.35 (m, 5H), 3.81 (t, J = 6.3 Hz, 2H), 3.71–3.58 (m, 6H), 3.35 (t, J = 5.1 Hz, 2H), 2.92 (t, J = 6.3 Hz, 2H) ppm. ^{13}C NMR (101 MHz, CDCl_3) δ 195.3, 134.4, 129.4, 129.1, 127.6, 70.5, 70.5, 70.0, 66.6, 50.6, 43.9 ppm.

Alkyne-thioester (**2**):



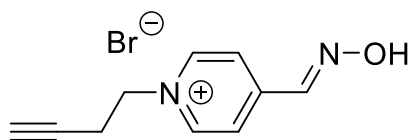
4-pentynoic acid (1.0 g, 10 mmol, 1.0 eq) was dissolved in 40 mL DCM, after which isobutyl chloroformate (1.5 mL, 11 mmol, 1.1 eq) and triethylamine (1.4 mL, 10 mmol, 1.0 eq) were added and the resulting mixture was stirred at rt for 10 min under argon atmosphere. Thiophenol (2.1 mL, 20 mmol, 2.0 eq) and triethylamine (1.4 mL, 10 mmol, 1.0 eq) were added subsequently, and the mixture was stirred at rt for 16 h. After this, the mixture was filtered, the solids washed with DCM, and the filtrate was washed with 1M KHSO₄, water and brine. The obtained organic phase was dried over Na₂SO₄ and concentrated under reduced pressure. The target compound was purified via automated flash column chromatography (SiO₂, 5% EtOAc in petroleum ether (40/60)), yielding a yellow oil (1.0 g, 54%). ¹H NMR (400 MHz, CDCl₃) δ 7.42 (s, 5H), 2.90 (dd, *J* = 7.9, 6.8 Hz, 2H), 2.58 (td, *J* = 7.4, 2.6 Hz, 2H), 2.02 (t, *J* = 2.7 Hz, 1H) ppm. ¹³C NMR (101 MHz, CDCl₃) δ 197.4, 135.8, 129.7, 129.4, 127.4, 82.0, 69.6, 42.2, 14.7 ppm.

Synthesis of alkyne-PyOx and its azide-/alkyne-functionalized ANANS substrates



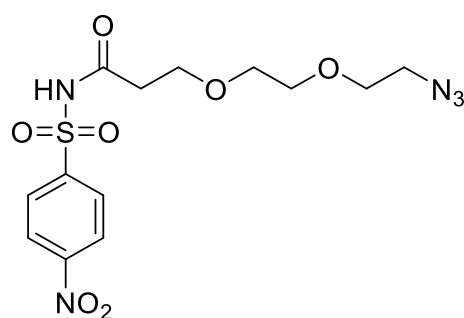
Scheme A4.2. Synthesis of alkyne-PyOx and ANANS **3** and **4**. Reaction conditions: (a) 4-bromobut-1-yne, anh. ACN, 82 °C, 25 h, **92%**; (b) 4-nitrobenzenesulfonamide, DIPEA, DCM, rt, 20 h, **90%**; (c) DIPEA, 1-(bromomethyl)-4-nitrobenzene, anh. THF, 50 °C, 17 h, **72%**; (d) isobutyl chloroformate, NEt₃, DCM, rt, 10 min; (e) 4-nitrobenzenesulfonamide, NEt₃, DCM, rt, 18 h, **11%**; (f) DIPEA, 1-(bromomethyl)-4-nitrobenzene, anh. THF, 50 °C, 18 h, **72%**.

Alkyne-PyOx:



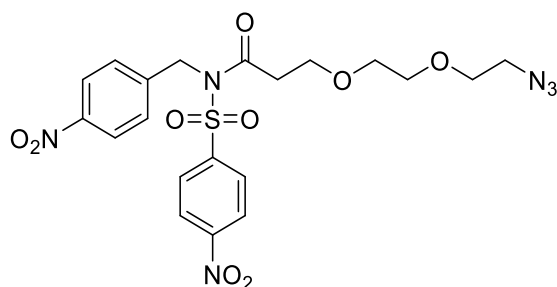
Pyridine-4-aldoxime (370 mg, 3.03 mmol, 1.0 eq) and 4-bromobut-1-yne (600 μ L, 6.67 mmol, 2.2 eq) were dissolved in 6 mL anhydrous ACN and the mixture was refluxed for 25 h. After this, the reaction mixture was filtered, the residues were washed with ACN and were left to dry overnight, yielding a light-brown solid (488 mg, 92%). HRMS (ESI) calculated for $[M]^+$: 175.0871; found $[M]^+$: 175.0863. ^1H NMR (400 MHz, D_2O) δ 8.96–8.89 (m, 2H), 8.42 (s, 1H), 8.29–8.22 (m, 2H), 4.77 (s, 2H), 3.01 (td, J = 6.3, 2.6 Hz, 2H), 2.54 (q, J = 2.6, 2.2 Hz, 1H) ppm. ^{13}C NMR (101 MHz, D_2O) δ 150.0, 146.3, 144.0, 125.7, 78.6, 74.7, 59.3, 20.0 ppm.

Azido-ANANS precursor:



4-nitrobenzenesulfonamide (162 mg, 799 μ mol, 1.2 eq) was dissolved in 1 mL DCM, after which *N,N*-diisopropylethylamine (232 μ L, 1.33 mmol, 2.0 eq) was added and the resulting mixture was stirred for 5 min. Then, 2,5-dioxopyrrolidin-1-yl 3-(2-(2-azidoethoxy)ethoxy)propanoate (200 mg, 666 μ mol, 1.0 eq) was added and the mixture was stirred at rt for 20 h. After this, the reaction was washed with 1 M HCl (8 mL) and the product extracted with DCM (3 \times 8 mL). The organic layer was washed with brine (25 mL) and the product extracted with DCM (2 \times 15 mL). The organic layer was dried over Na_2SO_4 and concentrated under reduced pressure. The target compound was purified via flash column chromatography (SiO_2 , starting with 3% [10% (25% NH_3 in H_2O) in MeOH] in DCM, then 5% MeOH in DCM), yielding a yellow oil (233 mg, 90%). HRMS (ESI) calculated for $[M+H]^+$: 410.0746; found $[M+H]^+$: 410.0730. ^1H NMR (400 MHz, CDCl_3) δ 8.23 (d, J = 8.5 Hz, 2H), 8.16 (d, J = 8.4 Hz, 2H), 3.60 (dt, J = 12.1, 5.4 Hz, 8H), 3.45 (t, J = 4.9 Hz, 2H), 2.49 (t, J = 5.7 Hz, 2H) ppm. ^{13}C NMR (101 MHz, CDCl_3) δ 177.3, 149.8, 147.9, 128.7, 123.9, 70.3, 70.0, 69.7, 67.7, 50.7, 39.0 ppm.

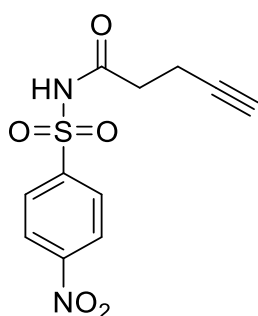
Azido-ANANS (3):



3-(2-(2-azidoethoxy)ethoxy)-*N*-((4-nitrophenyl)sulfonyl)propanamide (120 mg, 307 μ mol, 1.0 eq) was dissolved in 1 mL anhydrous THF, after which *N,N*-diisopropylethylamine (268 μ L, 1.54 mmol, 5.0 eq) was added and the resulting mixture was stirred at rt for 5 min. Then, 1-(bromomethyl)-4-nitrobenzene

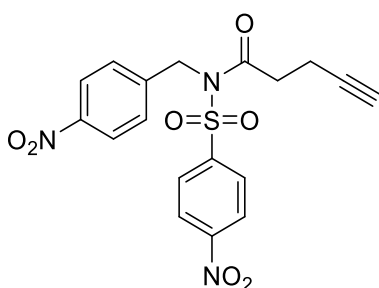
(265 mg, 1.23 mmol, 4.0 eq) was dissolved in 0.7 mL anh. THF and added to the solution and the resulting mixture was stirred at 50 °C for 17 h. After this, the reaction mixture was washed with brine (5 mL) and extracted with EtOAc (3 × 5 mL). The organic layer was dried over Na₂SO₄ and concentrated under reduced pressure. The target compound was purified via flash column chromatography (SiO₂, 30–40% EtOAc in petroleum ether(40-60)) yielding a yellow oil (121 mg, 76%). HRMS (ESI) calculated for [M+H]⁺: 545.1067; found [M+H]⁺: 545.1049. ¹H NMR (400 MHz, CDCl₃) δ 8.38 (d, *J* = 8.5 Hz, 2H), 8.23 (d, *J* = 8.4 Hz, 2H), 8.11 (d, *J* = 8.6 Hz, 2H), 7.55 (d, *J* = 8.3 Hz, 2H), 5.17 (s, 2H), 3.70 (t, *J* = 6.0 Hz, 2H), 3.62 (t, *J* = 4.9 Hz, 2H), 3.54 (q, *J* = 5.3, 4.2 Hz, 4H), 3.35 (t, *J* = 4.9 Hz, 2H), 2.84 (t, *J* = 6.0 Hz, 2H) ppm. ¹³C NMR (101 MHz, CDCl₃) δ 171.4, 150.9, 147.8, 144.7, 143.3, 129.5, 128.4, 124.6, 124.2, 70.6, 70.1, 66.4, 50.8, 49.5, 37.0 ppm.

Alkyne-ANANS precursor:



4-pentynoic acid (0.5 g, 5 mmol, 1 eq) was dissolved in 20 mL DCM, after which isobutyl chloroformate (0.75 mL, 5.5 mmol, 1.1 eq) and triethylamine (0.7 mL, 5 mmol, 1 eq) were added and the resulting mixture was stirred at rt for 10 min under argon atmosphere. Then, 4-nitrobenzenesulfonamide (1.31 g, 6.50 mmol, 1.3 eq) and triethylamine (0.7 mL, 5 mmol, 1 eq) were added and the mixture was stirred at rt for 18 h. After this, the resulting suspension was filtered and the filtrate was concentrated under reduced pressure. The target compound was purified with automated flash column chromatography (SiO₂, 0–5% [10% (25% NH₃ in H₂O) in MeOH] 5% MeOH in DCM), yielding a white solid (157 mg, 11%). ¹H NMR (400 MHz, MeOD) δ 8.42–8.37 (m, 2H), 8.25–8.18 (m, 2H), 2.44 (td, *J* = 6.8, 1.5 Hz, 2H), 2.39–2.32 (m, 2H), 2.19 (t, *J* = 2.6 Hz, 1H) ppm. ¹³C NMR (101 MHz, MeOD) δ 174.4, 151.7, 147.7, 130.5, 124.9, 83.1, 69.8, 36.3, 13.8 ppm.

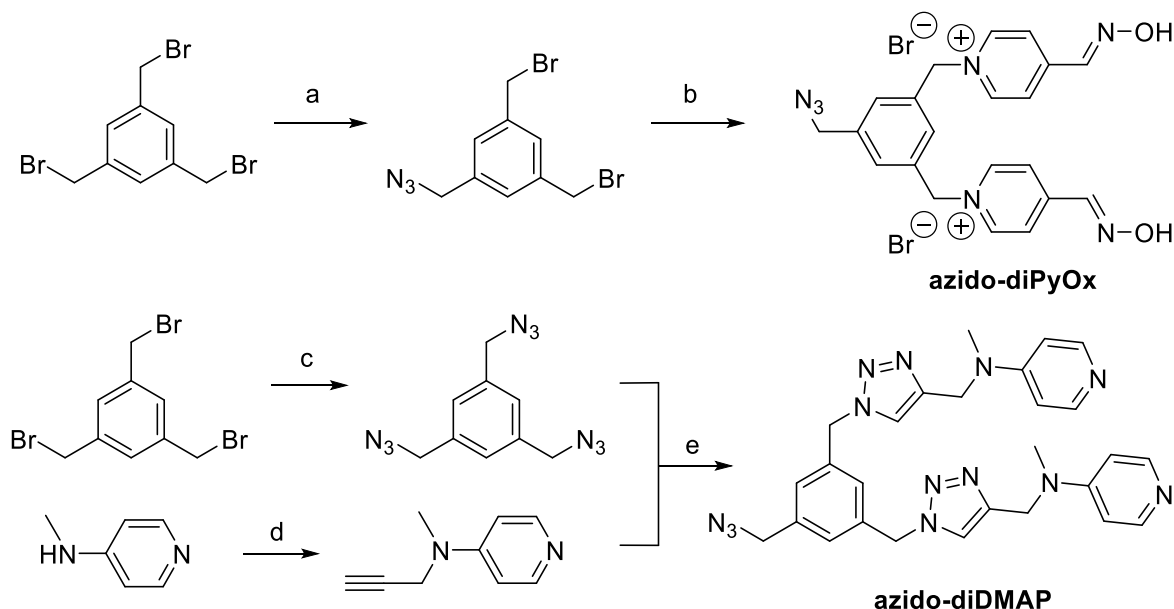
Alkyne-ANANS (4):



N-((4-nitrophenyl)sulfonyl)pent-4-ynamide (157 mg, 556 μmol, 1.0 eq) and was dissolved in 4 mL anh. THF. To this, *N,N*-diisopropylethylamine (0.5 mL, 3 mmol, 5 eq) was added and the resulting mixture was stirred at rt for 10 min. Then, 1-(bromomethyl)-4-nitrobenzene (481 mg, 2.23 mmol, 4.0 eq) was added and the mixture was stirred at 50 °C for 18 h. After this, brine (10 mL) was added to the mixture and the organic compounds were extracted with EtOAc (3 × 8 mL). The organic layer was dried over Na₂SO₄ and concentrated under reduced pressure. The target compound was purified via automated flash column chromatography (SiO₂, 20–40% EtOAc in petroleum ether(40-60)) yielding a pale-yellow solid (168 mg, 72%). HRMS (ESI) calculated for [M+Na]⁺: 440.0528; found [M+Na]⁺: 440.0511. ¹H NMR

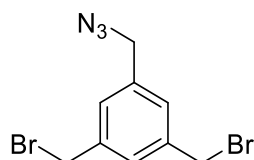
(400 MHz, CDCl_3) δ 8.48–8.32 (m, 2H), 8.23 (dt, J = 6.9, 3.0 Hz, 2H), 8.14–8.03 (m, 2H), 7.56 (dt, J = 6.6, 3.1 Hz, 2H), 5.18–5.12 (m, 2H), 2.82 (td, J = 7.0, 3.0 Hz, 2H), 2.45 (tt, J = 6.4, 3.1 Hz, 2H), 1.89 (q, J = 3.0 Hz, 1H) ppm. ^{13}C NMR (101 MHz, CDCl_3) δ 171.1, 150.9, 147.9, 144.5, 143.0, 129.4, 128.5, 124.8, 124.3, 81.8, 69.7, 49.5, 35.8, 14.1 ppm.

Synthesis of Azido-diDMAP and Azido-diPyOx



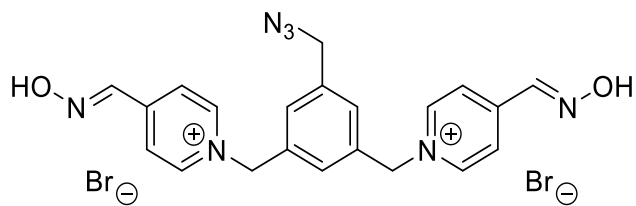
*Scheme A4.3. Synthesis of the divalent catalysts. Reaction conditions: (a) NaN_3 , DMF, rt, 16 h, **57%**; (b) pyridine-4-aldoxime, ACN, 65°C , 32 h, **64%**; (c) NaN_3 , DMF, rt, 16 h, **98%**; (d) $n\text{-BuLi}$, propargylbromide, anh. THF, -80°C – 0°C , 30 min, **57%**; (e) $\text{Cu(I)(ACN)}_4\text{PF}_6$, O_2 -poor THF, rt, 18 h, **11%**.*

1-(azidomethyl)-3,5-bis(bromomethyl)benzene:



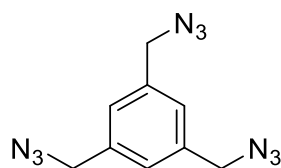
1,3,5-tris(bromomethyl)benzene (30 mg, 84 μmol , 1.0 eq) and sodium azide (5.5 mg, 84 μmol , 1.0 eq) were dissolved in 100 μL DMF and stirred at rt for 16 h. The volatile compounds were removed using a stream of air, after which the residue was dissolved in 400 μL DCM. The products were separated by means of preparative TLC (5% diethyl ether in petroleum ether(40–60)) and the desired product recovered with diethyl ether, filtered and concentrated under reduced pressure, yielding a white solid (15.1 mg, 57%). HRMS (ESI) calculated for $[\text{M}+\text{H}]^+$: 317.9241 / 319.9240; found $[\text{M}+\text{H}]^+$: 317.9318 / 319.9297. ^1H NMR (400 MHz, CDCl_3) δ 7.39 (s, 1H), 7.27 (d, J = 8.3 Hz, 2H), 4.46 (d, J = 7.5 Hz, 4H), 4.37 (s, 2H) ppm. ^{13}C NMR (101 MHz, CDCl_3) δ 139.3, 137.1, 129.6, 128.7, 54.3, 32.4 ppm.

Azido-diPyOx:



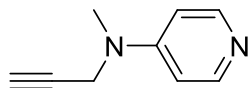
1-(azidomethyl)-3,5-bis(bromomethyl)benzene (8.0 mg, 25 μ mol, 1.0 eq) and pyridine-4-aldoxime (15 mg, 125 μ mol, 5 eq.) were dissolved in 500 μ L ACN and the resulting mixture was stirred at 65 $^{\circ}$ C for 32 h. The mixture was transferred to an Eppendorf tube, the reaction vessel was washed with ACN and the resulting combined ACN solutions were centrifuged with a tabletop centrifuge (1 min, 6000 rpm). The ACN was carefully removed from the precipitate, after which the solid was washed with clean ACN (3 \times 1 mL) using the same centrifugation procedure. The residue was allowed to dry overnight under a flow of air, yielding a brown solid (6.5 mg, 64%). HRMS (ESI) calculated for $[M-H]^+$: 402.1678; found $[M-H]^+$: 402.1678. ^1H NMR (400 MHz, D_2O) δ 8.86 (dt, J = 9.9, 4.8 Hz, 4H), 8.39 (d, J = 2.3 Hz, 2H), 8.25–8.19 (m, 4H), 7.55 (d, J = 2.8 Hz, 2H), 7.51 (s, 1H), 5.86 (d, J = 2.4 Hz, 4H), 4.49 (d, J = 2.3 Hz, 2H) ppm. ^{13}C NMR (101 MHz, D_2O) δ 149.3, 146.2, 144.8, 144.7, 138.7, 134.8, 129.9, 129.0, 125.0, 63.4, 53.3 ppm.

1,3,5-tris(azidomethyl)benzene:



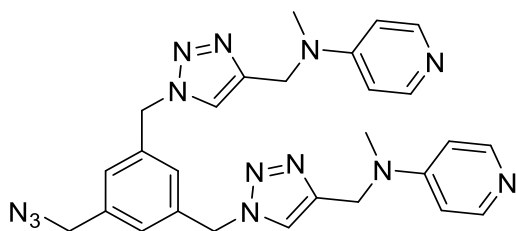
1,3,5-tris(bromomethyl)benzene (200 mg, 560 μ mol, 1.0 eq) and sodium azide (219 mg, 3.36 mmol, 6.0 eq) were dissolved in 0.5 mL DMF and the resulting mixture was stirred at rt for 16 h. The volatile compounds were removed using a stream of air, after which the residue was dissolved in 1 mL DCM. The mixture was washed with H_2O and the obtained aqueous phase was extracted with DCM (3 \times 5 mL). The combined organic phase was dried over Na_2SO_4 and concentrated under reduced pressure, yielding a clear oil (134 mg, 98%). ^1H NMR (400 MHz, CDCl_3) δ 7.24 (s, 3H), 4.38 (s, 6H) ppm. ^{13}C NMR (101 MHz, CDCl_3) δ 137.0, 127.4, 54.3 ppm.

Alkyne-DMAP:



4-(methylamino)-pyridine (200 mg, 1.9 mmol, 1.0 eq) was placed in a flame-dried 25 mL flask under argon and dissolved in anhydrous THF (1.5 mL). The solution was cooled to -90 $^{\circ}$ C and *n*-butyllithium (0.8 mL, 2.0 mmol, 1.1 eq) was added; the resulting mixture was stirred for 15 min at -90 $^{\circ}$ C. Then, propargyl bromide (200 μ L, 2.77 mmol, 1.5 eq) was added and the reaction mixture was stirred for 15 min at -90 $^{\circ}$ C and 15 min at 0 $^{\circ}$ C. After this, the reaction was quenched with sat. NH_4Cl (5 mL) and the organic product was extracted with THF (3 \times 6 mL). The combined organic phase was dried over Na_2SO_4 and resulting filtrate was concentrated under reduced pressure. The target compound was purified by flash column chromatography (SiO_2 , 5% (10% [25% NH_3 in H_2O] in MeOH) in EtOAc) yielding a brown oil (152.8 mg, 57%). HRMS (ESI) calculated for $[M+H]^+$: 147.0922; found $[M+H]^+$: 147.0917. ^1H NMR (400 MHz, CDCl_3) δ 8.13 (s, 2H), 6.44 (d, J = 5.8 Hz, 2H), 3.90 (d, J = 2.6 Hz, 2H), 2.85 (s, 3H), 2.14 (t, J = 2.5 Hz, 1H) ppm. ^{13}C NMR (101 MHz, CDCl_3) δ 153.0, 149.7, 107.4, 78.0, 72.4, 40.5, 37.4 ppm.

Azido-diDMAP:



1,3,5-tris(azidomethyl)benzene (30 mg, 123 μmol , 1.0 eq), *N*-methyl-*N*-(prop-2-yn-1-yl)pyridin-4-amine (36 mg, 247 μmol , 2.0 eq) and diisopropylethylamine (107 μL , 617 μmol , 5.0 eq) were mixed in acetonitrile (3 mL) and the resulting mixture was bubbled with argon for 30 minutes. Tetrakis(acetonitrile)copper(I) hexafluorophosphate (230 mg, 617 μmol , 5.0 eq) was added and the reaction was stirred at rt overnight. The mixture was washed with 10% 3M NaOH in brine (5 mL) and the product was extracted from the aqueous phase using EtOAc (3×10 mL). The combined organic phase was dried over Na_2SO_4 and after filtration the filtrate was concentrated under reduced pressure. The target compound was purified via flash column chromatography (SiO_2 , 8-12% (10% [25% NH_3 in H_2O] in MeOH) in DCM)) yielding a pale yellow solid (7 mg, 11 %). HRMS (ESI) calculated for $[\text{M}+\text{H}]^+$: 536.2747; found $[\text{M}+\text{H}]^+$: 536.2740. ^1H NMR (400 MHz, CD_3CN) δ 8.14 (s, 4H), 7.70 (s, 2H), 7.19 (d, J = 1.7 Hz, 2H), 7.13 (d, J = 1.8 Hz, 1H), 6.75 (s, 4H), 5.49 (s, 4H), 4.66 (s, 4H), 4.35 (s, 2H), 3.08 (s, 6H) ppm. ^{13}C NMR (101 MHz, CD_3CN) δ 155.0, 148.6, 144.7, 138.5, 138.2, 128.5, 128.0, 123.7, 108.3, 54.5, 53.8, 47.4, 38.2 ppm.

Chapter 5



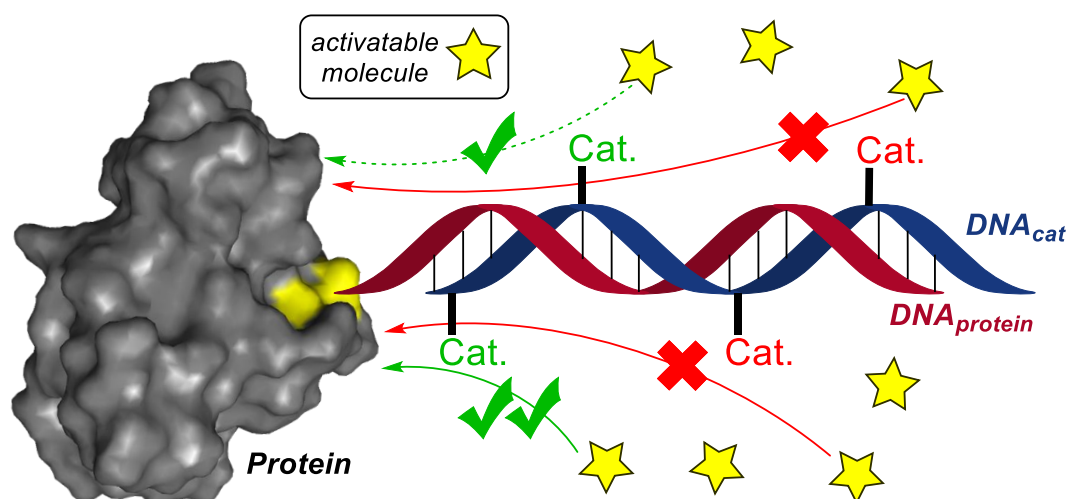
Calibrating Site-Selective Protein Modification by Catalytic DNA Nanostructures

Work by:

Jordi F. Keijzer, Han Zuilhof, Bauke Albada

Adapted from:

ChemEur. **2022**, <https://doi.org/10.1002/chem.202200895>



5.1 Abstract

Modified proteins are crucial for a wide range of biochemical and biological fields. To this aim, modification of proteins is often obtained by attaching reactive or catalytic moieties to protein-affinity tags. However, optimization of the position of the moiety with respect to the protein-affinity tag is typically out of scope and control. In this chapter, we synthesized three different covalently bound protein-DNA conjugates by modification with an azide and subsequent copper-catalysed click with alkyne-functionalized DNA. The protein-DNA conjugates were used as template strands for catalyst-functionalized DNA strand to assess the optimal and maximum range at which (three) different catalyst can perform protein modification. We found a strong correlation between the catalyst-to-protein distance and the efficiency of protein modification for acyl transfer catalysts, which operate via a covalently bound reactant intermediate. Additionally, we found that the catalyst's distance and orientation with respect to the protein surface, also influences its site-selectivity. The catalyst operating with unbound reactant intermediates showed only enhanced efficiency. Our results are rationalized using computational simulations, showing that one-point anchoring of the DNA construct leads to notable differences in the site of modification.

5.2 Introduction

Modified proteins are important to a large variety of scientific and commercial applications, including biomaterials,¹ therapeutics² and proteomics.³ Generally applicable approaches for artificial chemical modification of proteins are thus of great importance. But despite recent progress,^{4–7} widely applicable methods for site-specific protein modification remain elusive, because the outcome of many strategies changes when they are applied in alternate settings or on different proteins with other micro-environments. Methods have been established to derivatize specific residues in some proteins, *e.g.*, via genetic incorporation of orthogonal groups, such as azides or alkynes,^{5,6} recombinant proteins with unique micro-environments,⁸ or by optimized reagents designed to target single residues.^{9–11} Even the total synthesis of small proteins is a possibility.¹²

site-specific protein modification remains a major challenge, and no widely applicable approach is currently available. So far various methods have been established to derivatize specific residues in some proteins, *e.g.*, via reactive residues in unique micro-environments, genetic incorporation of uniquely reactive or maybe even orthogonal groups, such as azides or alkynes, or by optimized reagents designed to target single residues.^{9–11} However, none of these have been proven to be generally applicable to a variety of wild-type proteins, and most depend for their specificity on specific details of the protein involved.

As alternative to the approaches mentioned above, catalytic protein modification applies a molecular unit that activates an inert moiety, which in turn reacts with amino acid residues on the protein surface.^{5,13} Such protein modification catalysts can be organometallic,^{14,15} organic,^{16–18} enzymatic¹⁹ or even based on DNA.^{20,21} In order to achieve site selectivity in the modification, catalysts rely on a protein-binding element that brings them to a specific site of the target protein.^{5,13} These elements can bind proteins covalently, such as linchpins^{18,22} or non-covalently, such as ligands^{15–17} and DNA aptamers.²³ Although effective, the issue that arises in this strategy, is finding the ideal position of the catalyst with respect to the interface between protein and protein-binding moiety.^{23–25}

Previous work from our group showed that the position of acyl transfer catalysts with respect to the protein affected the conversion to modified protein and site-selectivity of the modification.²³ In order to develop nanometre-sized protein-binding catalysts for precision modification of proteins, we needed to gain better insight in the interplay between dimensions of the nanostructure and positioning of the

catalyst. Therefore, we covalently linked catalytic constructs and protein via a dsDNA structure in order to extract design principles that would allow us to develop next-generation protein modification tools.

Specifically, we covalently conjugated the 5' end of a template DNA strand (DNA_{temp}) to a protein and using it to hybridize its complementary strand that carries a protein-modifying catalyst at various distances from the 3' end (DNA_{catalyst}, Figure 5.1). For this, we synthesized protein-DNA_{temp} constructs using three proteins of various size and analysed the effective catalyst distance of three different catalysts: the acylation catalysts dimethylaminopyridine (DMAP) and pyridinecarbaldehyde oxime (PyOx), and the

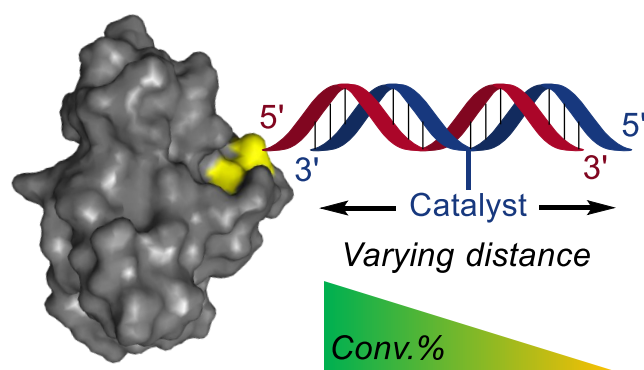


Figure 5.1. Schematic depiction of the work described in this chapter. We determine the distance at which different catalysts still efficiently modify proteins, hypothesizing a reduction in both efficiency and precision with increasing distance between catalyst and protein. For this, a template DNA strand is tethered to a protein and hybridized to a complementary DNA strand that contains one of these catalysts attached at varying distances.

oxidative cross-coupling catalyst hemin/G-Quadruplex. Whereas the first two catalysts acylate the nucleophilic residues Lys and Ser by means of a catalyst-bound reactive intermediate,²³ the latter catalyst generates a soluble reactive species that reacts with the electron-rich aromatic rings of Tyr and Trp.²⁰

5.3 Results & Discussion

Initial attempts entailed the site-specific attachment of an alkyne-DNA_{temp} strand to papain, a 26 kDa cysteine protease. Unfortunately, these resulted in incomplete conversion to DNA-papain, while conjugation was also insufficiently controlled for our goals. Therefore, we switched to glutaredoxin 1 (GRX), a small 9.5 kDa protein containing a single disulfide-bridge as its active site (Figure 5.2A). Using a slight excess (2 eq.) of dithiolthreitol (DTT), the bridge was reduced and 1-azidomethyl-3,5-bis(bromomethyl)benzene could be conjugated to the active site by reacting with both Cys residues, resulting in the installation of a single azide group on GRX. After purification by spin-filtration, mass spectrometry confirmed the identity of the product and that no crosslinking of GRX had occurred. Then, DNA_{temp} was attached to the azide using CuAAC, and the formed GRX-DNA_{temp} conjugate could be purified using ion-exchange FPLC (Figure 5.2B). The identity and purity of the GRX-DNA_{temp} conjugate was confirmed with HPLC-MS (Figure 5.2C) and SDS-PAGE (Figure 5.2D).

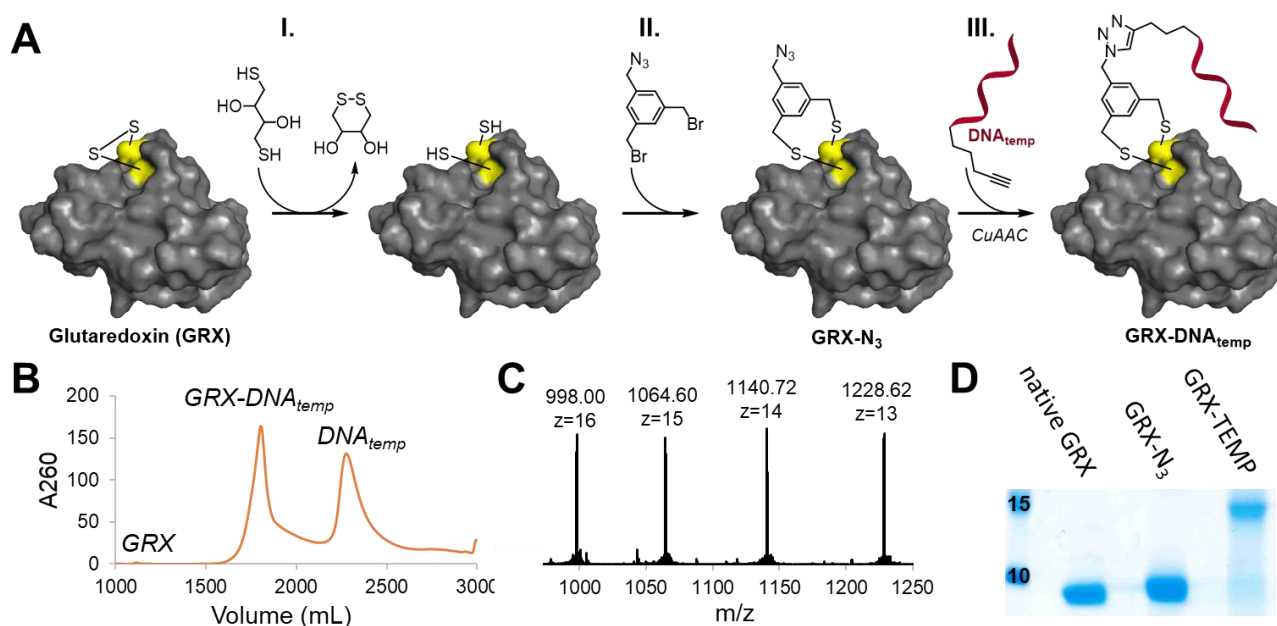
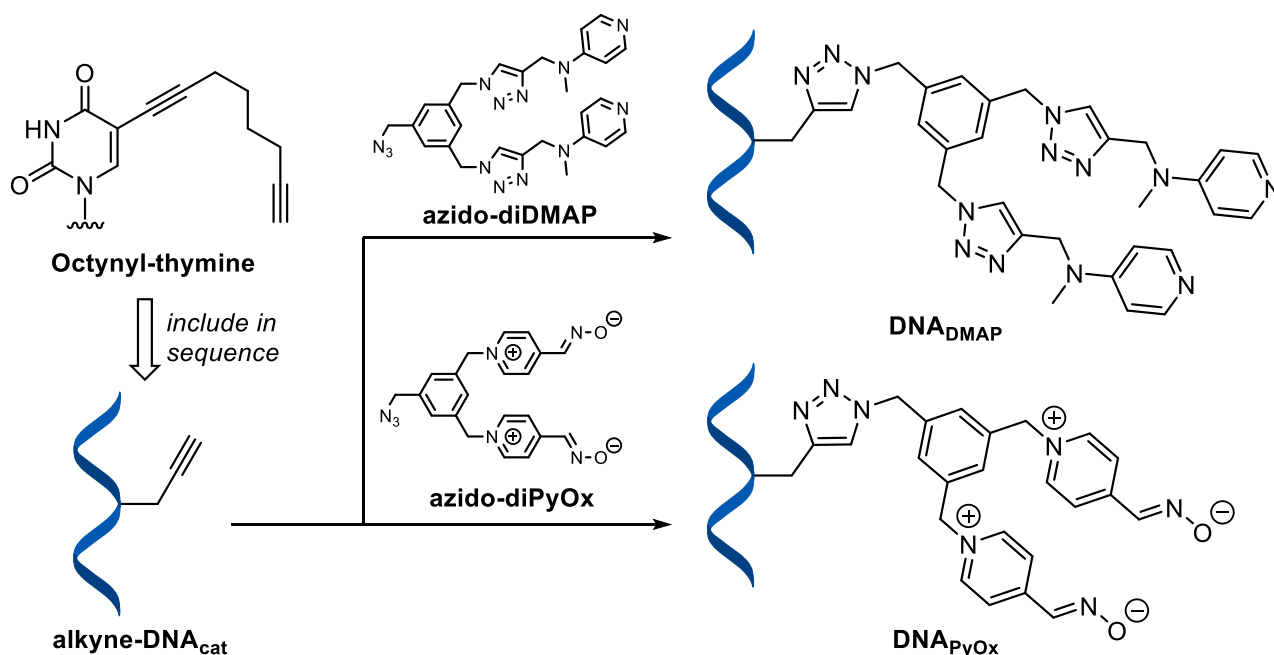


Figure 5.2. Preparation of the building blocks used to calibrate our catalytic DNA systems. (A) Synthesis of GRX-DNA_{temp} by means of (I) DTT-mediated opening of the disulfide bridge, (II) insertion of 1-azidomethyl-3,5-bis(bromomethyl)benzene to install a uniquely reactive azide at the active site, and (III) attachment of alkyne-DNA_{temp} by means of copper-catalysed click. (B) FPLC trace of the separation of GRX, GRX-DNA_{temp} and DNA_{temp} by ion-exchange (detection at 260 nm). (C) Mass spectrometry data of GRX-DNA_{temp} (calculated peaks: mass(z) = 998.0(16), 1064.6(15), 1140.7(14), 1228.3(13)). (D) SDS-PAGE analysis of native GRX, azide-functionalized GRX and the GRX-DNA_{temp} conjugate, which has an upward shift of ~6 kDa.

The complementary DNA-catalyst strand (*i.e.*, DNA_{DMAP} or DNA_{PyOx}) was prepared by CuAAC ligation of azido-diDMAP or azido-diPyOx to different complementary DNA strands. These strands contained one octynyl-modified thymine at different positions in the oligomer, namely at position 1, 2, 3, 4, 6 or 8 with respect to the 3' end of the oligomer (Scheme 5.1). After purification by spin filtration, HPLC-MS confirmed that the correct constructs were formed and obtained with >95% purity.



Scheme 5.1. Attachment of the different catalysts to commercially available DNA strands containing one alkyne-modified thymine using CuAAC in combination with the respective azido-functionalized catalyst, resulting in the catalyst-functionalized DNA pairing strands. Formation of these products was confirmed with HPLC-MS and indicated quantitative conversion to the DNA-catalysts.

After the different components required for the protein modification studies were obtained, we assessed how the efficiency of modification was affected by the details of the protein-catalyst interaction. For this, we first incubated the DNA_{DMAP} strands in a 1:1 ratio with GRX-DNA_{temp} in HEPES buffer (pH: 8.0) for 30 min. Then, thioester **1** (Figure 5.3A) was added and after 2 h the reaction was quenched by addition of an excess ethanolamine. Subsequently, the acylated protein was clicked using SPAAC with a bicyclononyne (BCN)-functionalized 2kDa PEG unit. The different protein derivatives were separated from GRX-DNA_{temp} using SDS-PAGE, and the individual bands were integrated and quantified with ImageJ. We found that positioning of the diDMAP further away from the protein surface did not markedly vary with positioning in the first three sites, but resulted in a gradual decrease in acylation of GRX-DNA_{temp} when the catalyst was moved further out (Figure 5.3B). This demonstrates a clear correlation between catalyst-to-protein distance and modification efficiency.

We also incubated GRX-DNA_{temp} with various DNA-diPyOx strands in HEPES buffer (pH: 7.2) for 30 min, after which its dedicated substrate ANANS **3** (Figure 5.3A) was added. In this case, the reaction was quenched by the addition of citrate buffer (pH: 5) and worked up as described above for the diDMAP-catalyzed modifications. For diPyOx, we found a similar decrease in acylation by distancing the catalyst from the protein surface and again T1 being the most efficient position (Figure 5.3B). Surprisingly, the total conversion performed by the diDMAP catalyst far exceeded that of the diPyOx system, as the former generated similar conversions using only a third of the concentrations of acyl donor (Figure 5.3B). This finding is opposed to what was observed in aptamer-based experiments,²³ in which diPyOx generated higher yields (Figure 4.3E-F, page 69).

After this, we analysed the site-selectivity of the modification by tryptic digestion and follow-up MS/MS for both of these catalysts. GRX contains a total of six Lys residues: K18 (SAA: 269 Å²), K23 (SAA: 276 Å²), K45 (SAA: 280 Å²), K51 (SAA: 275 Å²), K54 (SAA: 299 Å²), K80 (SAA: 274 Å²); and two Ser residues: S9 (SAA: 207 Å²) and S25 (SAA: 206 Å²). All positions of diDMAP (T1-T8) acylate five of the six Lys residues, where K80 is situated furthest away from the catalysts and is the only unmodified residue (Figure 5.3C). Similarly, diDMAP on T8 also cannot reach K54, which is likely again caused by distance. Interestingly, diPyOx resulted in a more varied modification landscape: K80 was modified when diPyOx was positioned on T1 and T8, but not on T2-T6. Similarly, K18 remained unmodified by diPyOx on T3

and T8, and K54 was unmodified by diPyOx on T1 and T3. We hypothesize that this results not only from increasing distances, but also the helical shape of the dsDNA, which yields with this variation in distance also a distance in spatial orientation. Apparently, this is not critical from diDMAP, but is critical for diPyOx, suggesting a subtle interplay between catalyst efficiency and specificity on the one hand and substrate on the other hand.

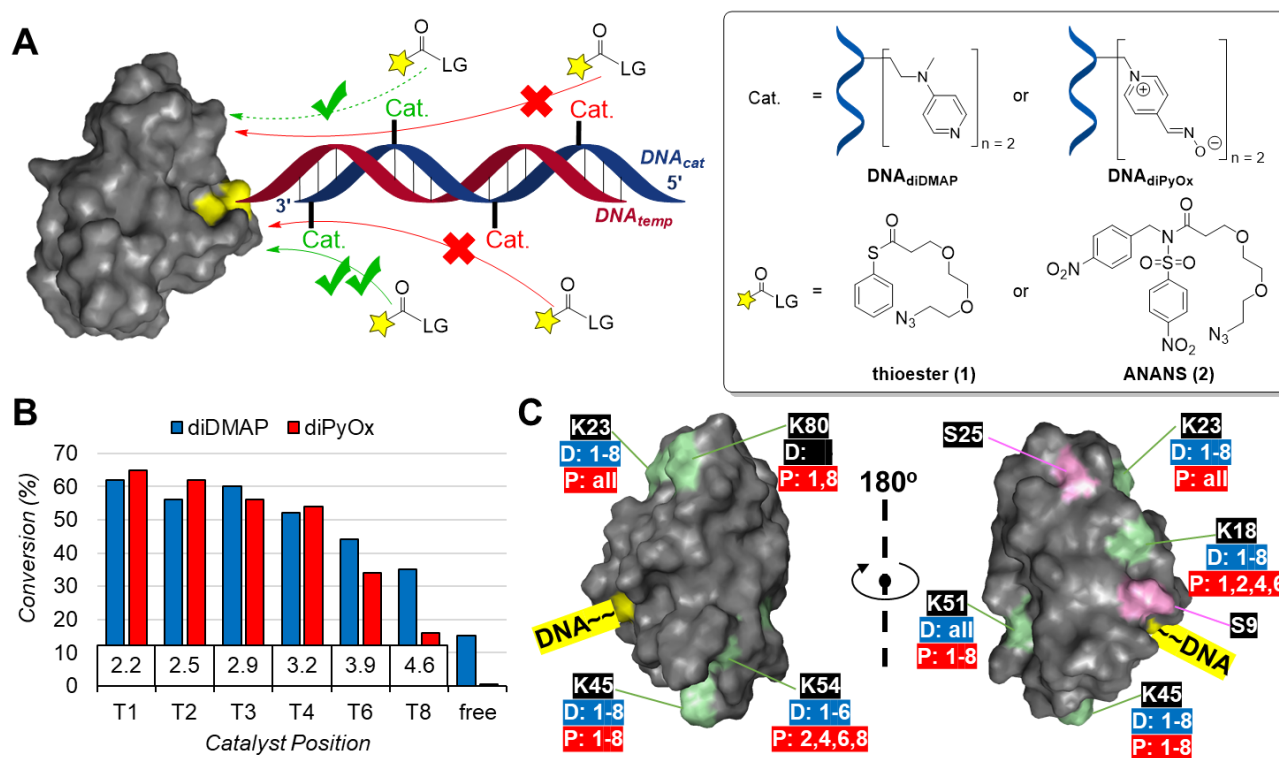


Figure 5.3. (A) Principle of the reaction setup where the efficiency of catalysts at increasing distance is analysed. The side panel shows diDMAP and diPyOx that use acyl donors **1** and **2**, respectively; (B) Diagram showing the decline in conversion percentages of GRX-DNA_{temp} by diDMAP or diPyOx when positioned further away from the protein surface with T1 being the closest. The numbers in the boxes show the distance between nucleobase and protein surface in nm. (C) Crystal structure(s) of GRX (PDB code: 1EGO) showing Lys residues (green) with the numbers of catalyst positions that modify them (D = diDMAP (blue), P = diPyOx (red), all = including free catalyst) as well as the attachment site on the DNA strand. Conditions: 20 μ M GRX-DNA_{temp} with (a) 22 μ M DNA_{diDMAP} and 100 μ M thioester **1**, pH: 8.0, at 37 $^{\circ}$ C for 2 h or (b) 22 μ M DNA_{diPyOx} and 300 μ M ANANS **3**, pH: 7.2, at 37 $^{\circ}$ C for 6 h.

Next to control over the selectivity of the modification by these catalysts that operate via a covalently-bound substrate, we hypothesized that the distance between a hemin/G-quadruplex (hGQ) catalyst and the protein would provide similar control over its modification efficiency. As the hGQ DNAzyme generates a soluble unbound radical reactant that has to diffuse to a proximal reactive site, we anticipated to see an effect of the distance on the modification efficiency. Therefore, the hybridizing DNA strand was designed to include a G-Quadruplex folding sequence, i.e., PW17, so that upon hybridization and subsequent addition of hemin, a protein-bound hGQ DNAzyme was formed. When the PW17 DNAzyme forms a hybrid GQ structure (Table A2.1, page 40), the 3'- and 5'-end are positioned in close proximity. As a result, we anticipate grafting on dsDNA strands will not interfere with the hGQ DNAzyme function.²⁶ The PW17 sequence was included at different sites from the 3'-end to the 5'-end (Figure 5.4). Afterwards, N-methyl-luminol (NML) **1** and H₂O₂ were added to initiate the conjugation and after 30 min catalase was added to stop the reaction by removing the H₂O₂. The excess of NML **1** was removed by spin filtration, and a SPAAC reaction between azide-functionalized protein and BCN-PEG₂₀₀₀ was performed. The components in the mixture were separated using SDS-PAGE, and integration of the Coomassie-stained bands revealed the modified proteins. Even though bound hGQ resulted in higher conversions than unbound hGQ, different positions of the catalyst on the DNA oligomer did not generate much variation, not even when the catalyst was positioned at T20, which is \sim 9 nm away from the protein

surface.²⁷ Apparently, the produced NML radical survives sufficiently long to diffuse to over at least 9 nm to reactive sites on the protein. In view of the similarities of the level of modification at different positions, the site-selectivity of the modification was not further examined. Clearly, use of a catalyst that generates a soluble reactive species requires additional levels of control when compared to an approach that relies on an activated catalyst-bound intermediate.

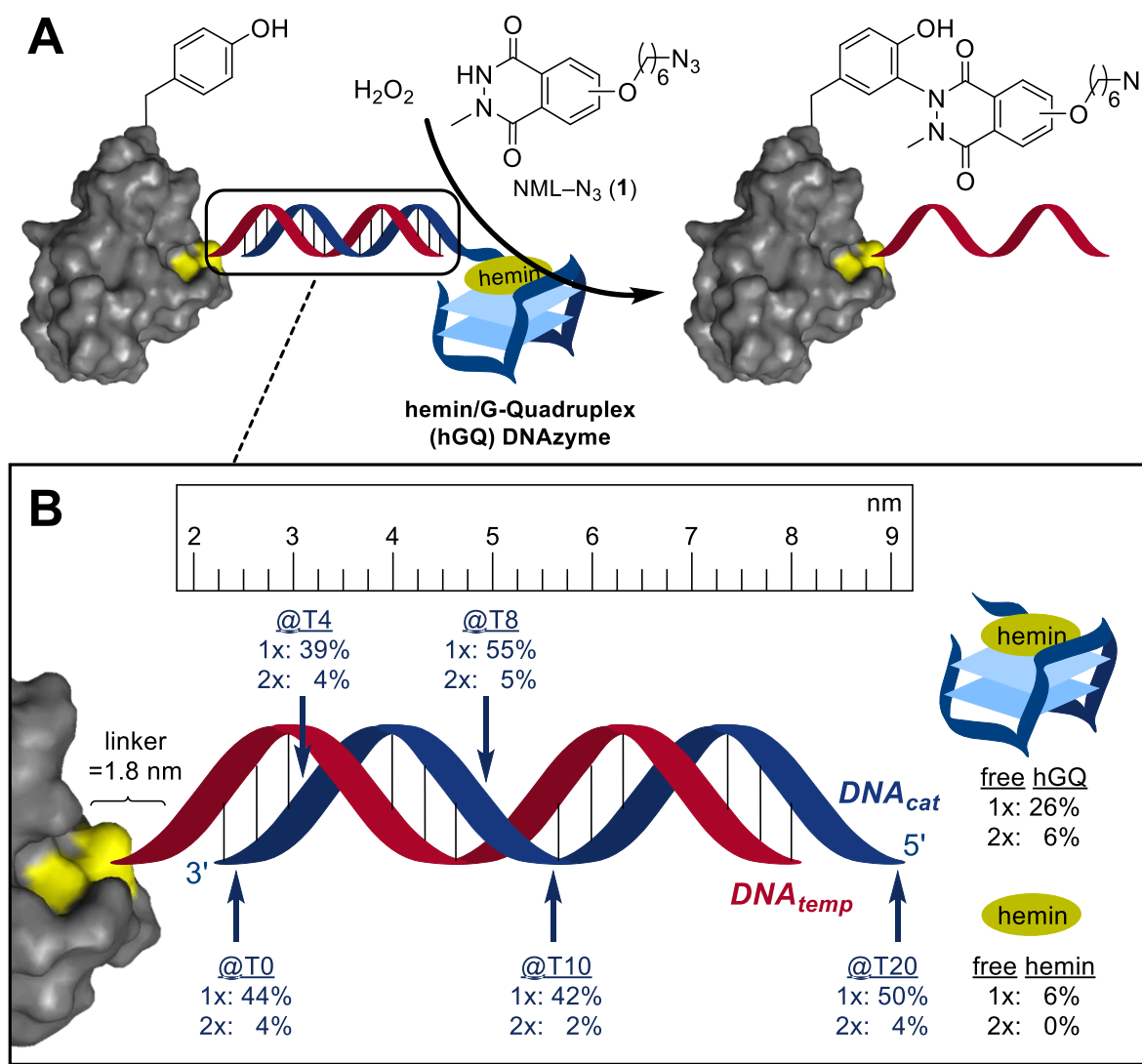
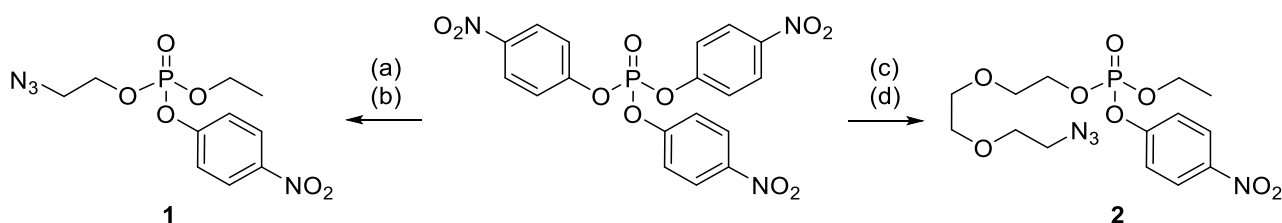


Figure 5.4. (A) The G-Quadruplex-forming sequence PW17 is included in the hybridizing strand and by addition of hemin, a protein-bound hGQ DNAzyme is formed. When H₂O₂ is present, the DNAzyme conjugates N-methyl-luminol (NML) **1** to tyrosine residues on the protein, which can be visualized after removal of the hGQ-containing DNA strand. (B) The different positions where PW17 was included with the percentages of single and double modification that the DNAzyme generated. Conditions: 20 μ M GRX-DNA_{temp}, 22 μ M DNA-hGQ, 30 μ M NML **1** and 100 μ M H₂O₂, pH: 7.0, at 25 $^{\circ}$ C for 30 min.

After these encouraging results for the modification of the small protein GRX using DMAP and PyOx catalysts, we applied larger proteins in order to determine how the distance between catalyst and protein would affect the region of acylation. We thus devised a way to synthesize one or more protein-DNA conjugates, large enough to observe a possible difference between sites of modification. Paraoxon is a selective inhibitor for serine proteases and binds the serine residue in their active site.²⁸ Indeed, incubation of serine proteases chymotrypsin (25 kDa) and human α -thrombin (36 kDa) with paraoxon generated nearly quantitative single modification. We thus synthesized an azido-functionalized paraoxon, which could easily be made from commercially available tris(*p*-nitrophenol)phosphate in a single reaction with azido-ethanol (Scheme 5.2, **1**).



Scheme 5.2. One-pot syntheses of azido-functionalized paraoxon derivatives. Reaction conditions: (a) 2-azidoethanol, DBU, DCM, 0 °C, 1 h, (b) ethanol, DBU, DCM, 0 °C, 1 h, **35%** (over 2 steps). (c) 2-azido-EG₂-ethanol, DBU, DCM, 0 °C, 1 h, (d) ethanol, DBU, DCM, 0 °C, 1 h, **25%** (over 2 steps).

Incubation chymotrypsin (CHY) and thrombin (TRM) with this azido-paraoxon and subsequent purification by spin filtration (Figure 5.5A), resulted in quantitative formation of singly modified proteases as determined by LC-MS (Figure 5.5B,D). DNA_{temp} was then attached by means of CuAAC, generating yields of 67% and 50% for chymotrypsin (CHY-Et-DNA_{temp}) and thrombin (TRM-Et-DNA_{temp}), respectively. As we anticipated that the potentially hidden azide in this first approach might hamper attachment of the DNA anchor strand, we also synthesized an azido-EG₂-paraoxon species (Scheme 5.2, **2**) to make the protein-bound azide moiety more accessible. Indeed, conjugation yields of the DNA to the proteins increased to 72% for chymotrypsin (CHY-EG₂-DNA_{temp}) and 65% for thrombin (TRM-EG₂-DNA_{temp}), respectively. These two different tethers enabled us to correlate their effect on the modification performance of the different catalysts.

Purification of these protein-DNA_{temp} conjugates was again performed by ion-exchange FPLC. Unfortunately, whereas product formation was confirmed with SDS-PAGE (Figure 5.5C,E), the products suffered from degradation during or after the ion-exchange FPLC purification. By carefully monitoring the product in each synthetic step, we discovered that the protein-DNA_{temp} conjugates collapse in the absence of glycerol (Figure A5.4). As such, we choose to purify these larger protein-DNA_{temp} conjugates only by spin filtration using a molecular weight cut off value of 10 kDa. Fortunately, this was sufficient to remove the bulk of the remaining DNA_{temp} (~6 kDa), while it still allowed the use of glycerol in the solutions to retain the stability of the constructs.

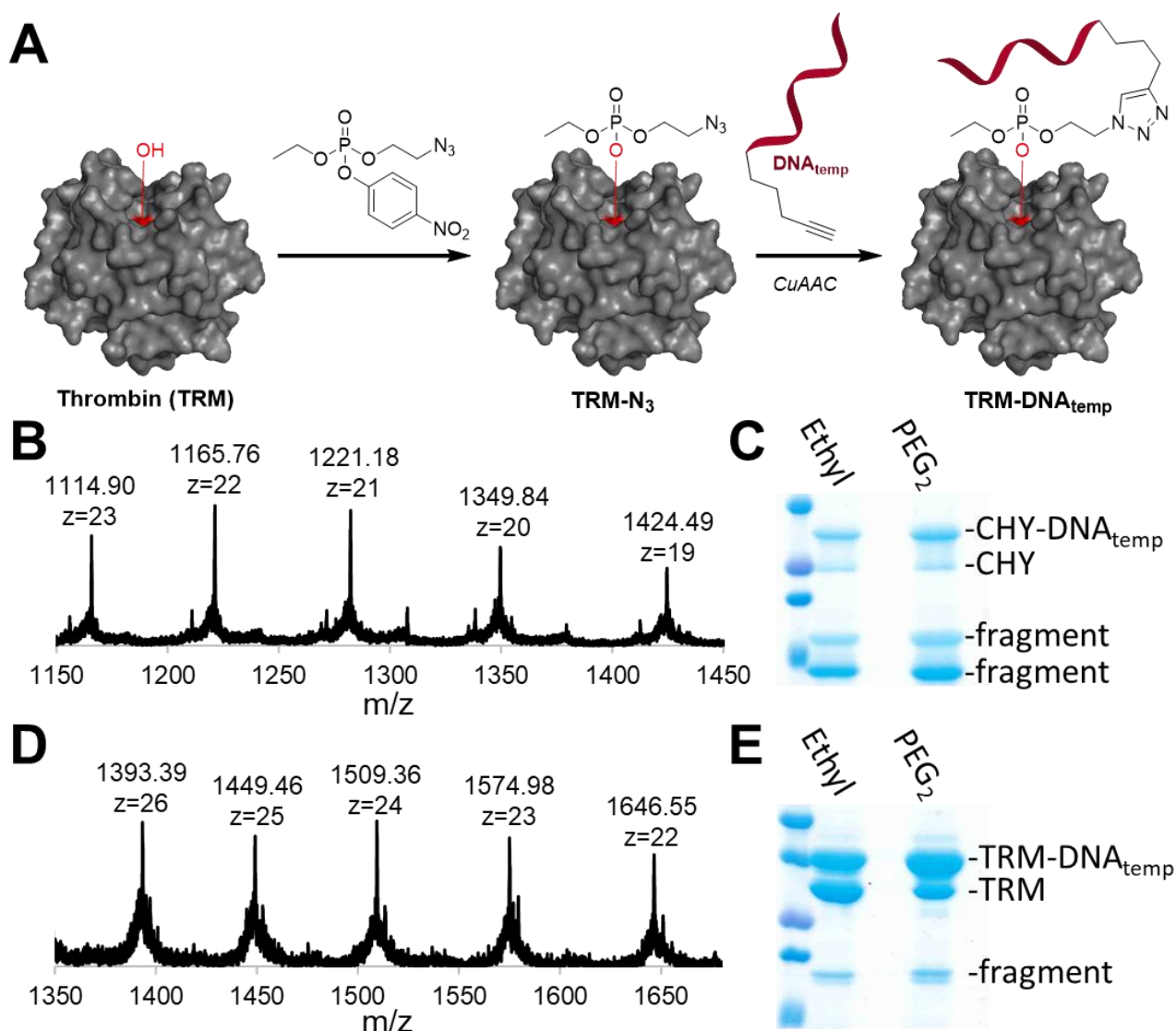


Figure 5.5. (A) Synthesis of thrombin-DNA (TRM-DNA_{temp}) and chymotrypsin-DNA (CHY-DNA_{temp}) by using paraoxon derivative **1** (or **2**); (B) Mass spectrometry data of CHY-Et-N₃ (calculated peaks: mass(z) = 1115.0(23), 1165.7(22), 1221.2(21), 1282.2(20), 1349.8(19), 1424.5(18)); (C) CHY-DNA_{temp} synthesized with the ethyl or EG₂ linker; (D) Mass spectrometry data of TRM-Et-N₃ (calculated peaks: mass(z) = 1393.4(26), 1449.1(25), 1509.3(24), 1575.0(23), 1645.5(22)); (E) SDS-PAGE result of synthesized TRM-DNA_{temp} with ethyl or EG₂ linker.

When performing the diDMP and diPyOx acylation with CHY-DNA_{temp}, the results were difficult to interpret due to protein degradation both during the reaction and after acylation and clicking with BCN-PEG₂₀₀₀ or BCN-lissamine. Only normalized values for the conversion could eventually be obtained (Figure 5.6A). Thrombin-DNA, however, did not degrade during the reaction and after PEGylation in the second step, we could integrate product concentrations from the SDS-PAGE with ImageJ. For both diDMP and diPyOx we found the same inverse correlation between acylation percentage and catalyst distance as we observed for GRX-DNA_{temp} (Figure 5.6B), and found that the best conversions are obtained when the catalyst is connected at the first few nucleotides of the protein-bound dsDNA unit. Remarkably, for both PEG-linked conjugates, the ideal position of diDMP seems to be T3. This does however, not seem the case for the ethyl-linked conjugates where T4 appears to be slightly higher. It appears that diDMP is optimally positioned at roughly 4 nucleobases distance when the dsDNA is connected to the protein by means of an ethyl unit, whereas for the slightly longer EG₂ linker the optimal position is at nucleotide T3. Notably, tests of TRM-DNA_{temp} with the hGQ DNAzyme revealed little to no variation in modification percentage (Figure A5.9 and A5.10), which confirms our finding for GRX-DNA_{temp} that this protein-modifying catalyst does not benefit from close proximity to the protein.

Analysis of the site selectivity of the modifications by these dsDNA-bound catalysts was performed by tryptic digestion and follow-up MS/MS. Firstly, the covalently-bound catalytic DNA constructs were able to acylate residues that were not acylated by the free catalyst. This is most clearly seen for DMAP, which by itself only acylates K83, but when connected to the protein by means of the dsDNA construct many additional residues were acylated. To our delight, we found that both catalysts primarily acylated residues in close proximity to the active site where the catalytic dsDNA construct was anchored (Figure 5.6C–D). In fact, distant residues are only modified when the catalyst is positioned close enough to the anchor point of the dsDNA. As such, K226 is only modified by diDMAP at a distance of <8 nucleobases and K106 only at <4 nucleobases. Similarly, K226 and S22 are only modified by diPyOx at <8 nucleobases, K83 only at <6 nucleobases and K06 and K232 only at 1 nucleobase distance. This implies that by positioning the catalyst further away, a trade-off between conversion and selectivity can be made where conversion could be reduced to attain a higher level of site selectivity, and *vice versa*. Interestingly, although diPyOx can also acylate Ser residues (in contrast to diDMAP), of all nearby residues only Ser22 was modified. Because the SAA of Ser22 is in the same range as the others (Ser22 = 206 Å², Ser67 = 206 Å², Ser158 = 203 Å², Ser176 = 205 Å² and Ser216 = 201 Å²) this is likely the only residue within reach. Once more, however, the spatial orientation caused by the helicity of the dsDNA influences site selectivity, as modification of some residues (K77 and K83 for diDMAP and K77 for diPyOx) does not directly correlate to distance.

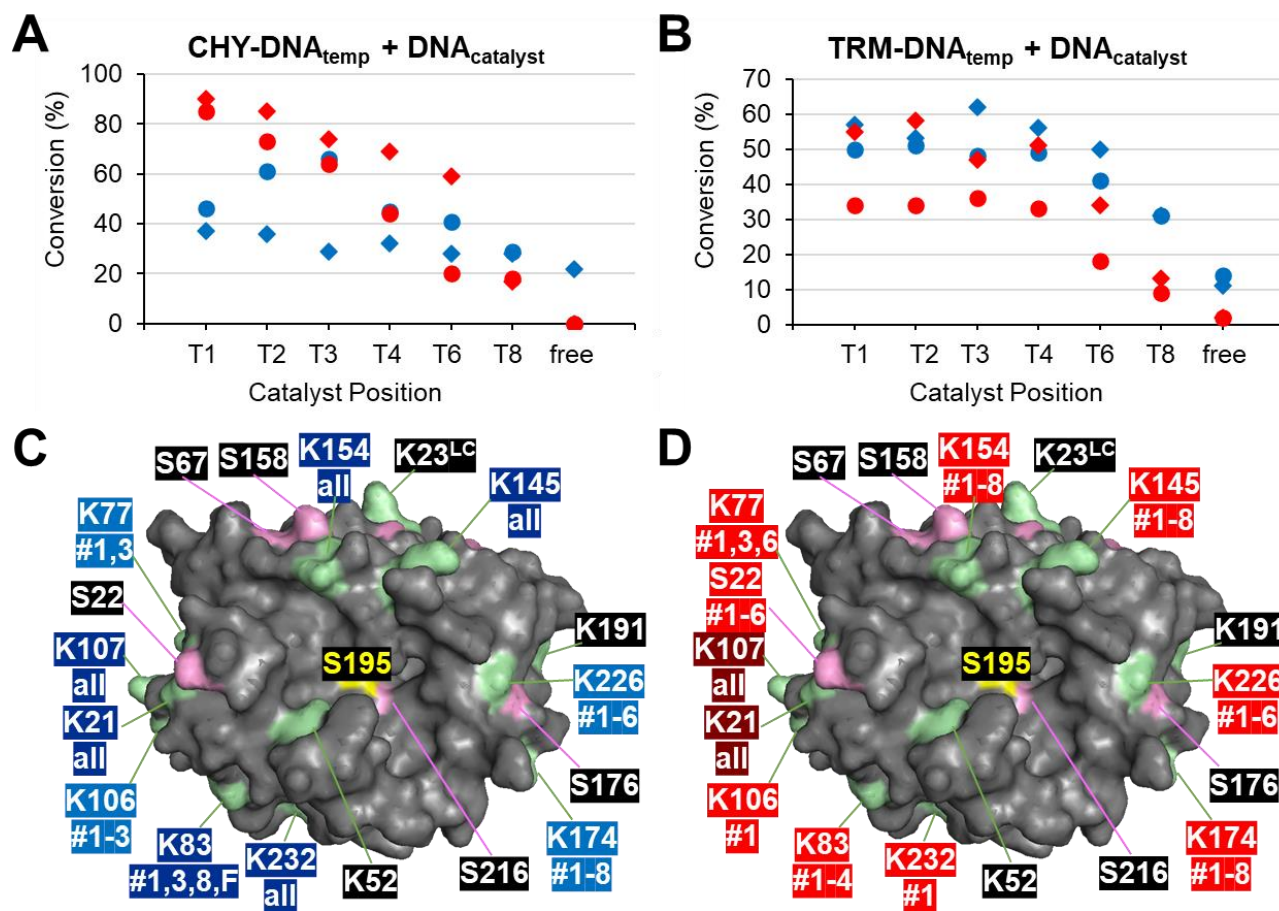


Figure 5.6. (A–B) Graphs showing the decline in conversion percentages of (A) CHY-DNA_{temp} and (B) TRM-DNA_{temp} by DNADMAP or DNAPyOx when positioned further away from the protein surface (T1 is 3' end). The shapes indicate ♦ for ethyl and • for EG₂ linkers, where the colours indicate blue for diDMAP and red for diPyOx. Conversions in (A) are normalized values. Conditions: 20–26 μM protein-DNA_{temp} with (i) 23–28 μM DNADMAP and 100 μM thioester **1**, pH: 8.0, at 37 °C for 2 h or (ii) 23–28 μM DNAPyOx and 300 μM ANANS **3**, pH: 7.2, at 37 °C for 6 h; (C–D) Crystal structure of thrombin showing its Lys (green) and Ser (pink) residues with respect to the active site where DNATemp is attached (S195) (PDB-code: 5EW1²⁹). #^{LC} = Light Chain. (C) Modification sites by DNADMAP. (D) Modification sites by DNAPyOx. Light coloured numbers indicate residues modified with only bound catalysts and dark coloured numbers indicate residues also modified with unbound catalyst.

Finally, in order to obtain a better understanding of how the positioning of the catalyst on the dsDNA affects the site of modification, we performed computational simulations using thrombin as our protein of interest. First, we constructed the dsDNA-thrombin conjugate using the single crystal X-ray structure of the protein (PDB-code 5EW1),²⁹ and a model of the appropriate dsDNA equipped with a diPyOx catalyst on position T1. As the difference in residues that were modified by ethyl and EG₂-linked dsDNA constructs was minimal, we focused this analysis on the EG₂-linked constructs (Figure 5.7A). We sampled the conformational space of the construct using internal coordinates of the model, with focus on the linker between Ser195 and dsDNA and on the diPyOx-functionalized T1. Furthermore, dihedral angles were set to be limited to energetically favour the staggered conformation, and bumps between structures were minimized. As a result, an umbrella of 10 different structures was generated (Figure 5.7B-D), revealing an impression of the reach of the catalyst when attached at T1. Closer inspection of the positions of the catalysts confirms that all the modified residues lie within reach of the catalyst, and that unmodified residues are either outside of the range that can interact with the catalyst, or are via its spatial orientation hidden from the catalyst by the dsDNA unit, which is the case for K52 that is next to the active site (Figure 5.6C,D). As the distances between some of the more remotely positioned residues are large, this analysis reveals the presence of considerable freedom in this system. This could be expected based on the degrees of freedom in the linkers, which translates to the rotational freedom of the dsDNA part. Therefore, more control over the precision of the modification can likely be obtained by additional anchoring of the catalytic DNA constructs.

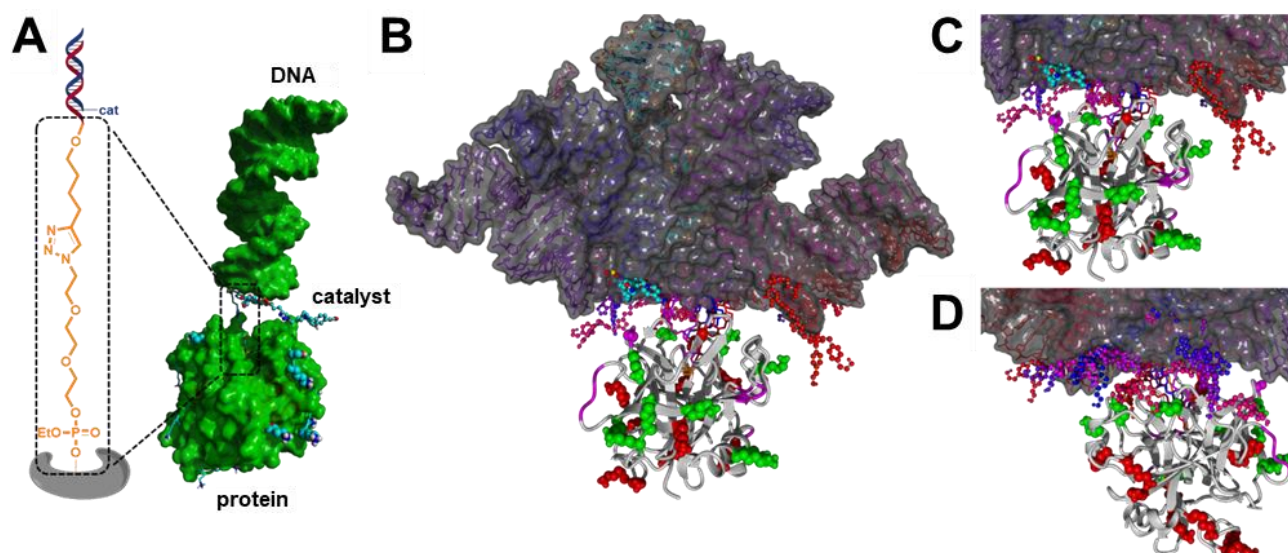


Figure 5.7. Molecular model of the thrombin-dsDNA-diPyOx construct to rationalize the observed modifications. (A) Details of the model, with emphasis on the linker between protein and dsDNA; the linker is shown in sticks, the diPyOx catalyst in ball-and-stick, the modified Lys residues and one Ser residue in balls, and the unmodified Lys residues in sticks; surfaces of the protein and DNA are depicted in green, except for the previously described details. In the model, the dsDNA unit is shown on top and the protein at the bottom. (B) Resulting 10 structures of sampling the dihedral angles of the spacer between protein and dsDNA, and of diPyOx-functionalized T1. Modified Lys residues are shown as green balls, modified Ser residue as magenta balls, unmodified Lys residues as red balls. The position of the protein is fixed, the 10 differently positioned dsDNA-diPyOx units are coloured from blue to red, their surfaces are shown in grey. (C-D) Zoomed parts of the interface between dsDNA-diPyOx and thrombin.

5.4 Conclusions

The catalytic activity and site-specificity of DNA-bound protein-modifying moieties was tested as a function of distance to the reactive site at three proteins, including two serine proteases. Firstly, whereas the isolated catalysts resulted in poor modification of the proteins and only of one exposed lysine residue (K83), the covalently attached constructs resulted in much higher levels of modification.

For the covalently linked catalysts, two specific relations were found: (i) If the catalyst has to reach at the site of the protein to perform its action, then a strong tapering down of activity was observed with increasing distance. This situation takes place, for example, with DMAP and PyOx modification of Lys and Ser residues on the protein surface. Basically, no distance dependence was observed for a distance of 1-2 base pairs, after which a gradual decline to far lower reactivity was observed for catalysts at an 8 base pairs distance from the reactive site. (ii) If, in contrast, the catalyst produces a reactive, soluble intermediate that itself reacts at the protein surface, then there is no clear distance dependence if the intermediate is sufficiently long-lived to allow diffusion-mediated reactivity. This situation occurs with the hGQ DNAzyme, which uses H_2O_2 to produce a long-lived, yet sufficiently reactive intermediate radical, then no decline in reactivity is observed even with up to 20 nucleobases distance. Lastly, we noticed that in specific cases, the effect of even one additional nucleobase could be quite significant, which we attribute to a combination of distance and spatial orientation, as confirmed by molecular modelling. Following these conclusions, we anticipate that these results will help the design of future probes with either of the two types of catalysts (activation via a bound reactant intermediate or an unbound radical reactant) investigated in this work by taking into account the described effects of distance and spatial orientation for. We also note that our work adds another strategy to create single DNA-protein conjugates, which are in high demand for applications in the fields of diagnostics, medicine and nanotechnology. More specifically, for bioassays, caged-enzymes or carriers for enzymes or nucleic acids for gene editing and disease treatment.³⁰⁻³²

Publication online

This work is also published and can be found here:



References

- 1 J. A. Shadish and C. A. DeForest, *Matter* **2020**, 2 (1), 50–77.
- 2 M. J. Birrer, K. N. Moore, I. Betella and R. C. Bates, *J. Natl. Cancer Inst.* **2019**, 111 (6), 538–549.
- 3 Y. Zhao and O. N. Jensen, *Proteomics* **2009**, 9 (20), 4632–4641.
- 4 N. C. Reddy, M. Kumar, R. Molla and V. Rai, *Org. Biomol. Chem.* **2020**, 18 (25), 4669–4691.
- 5 E. A. Hoyt, P. M. S. D. Cal, B. L. Oliveira and G. J. L. Bernardes, *Nat. Rev. Chem.* **2019**, 3 (3), 147–171.
- 6 B. Albada, J. F. Keijzer, H. Zuilhof and F. van Delft, *Chem. Rev.* **2021**, 121 (12), 7032–7058.
- 7 J. F. Keijzer and B. Albada, *Biopolymers* **2021**, 113 (3), 1–8.

- 8 A. C. Conibear, E. E. Watson, R. J. Payne and C. F. W. Becker, *Chem. Soc. Rev.* **2018**, 47 (24), 9046–9068.
- 9 M. J. Matos, B. L. Oliveira, N. Martínez-Sáez, A. Guerreiro, P. M. S. D. Cal, J. Bertoldo, M. Maneiro, E. Perkins, J. Howard, M. J. Deery, J. M. Chalker, F. Corzana, G. Jiménez-Osés and G. J. L. Bernardes, *J. Am. Chem. Soc.* **2018**, 140 (11), 4004–4017.
- 10 D. Hwang, K. Tsuji, H. Park, T. R. Burke and C. Rader, *Bioconjugate Chem.* **2019**, 30 (11), 2889–2896.
- 11 V. Laserna, A. Istrate, K. Kafuta, T. A. Hakala, T. P. J. Knowles, M. Alcarazo and G. J. L. Bernardes, *Bioconjugate Chem.* **2021**, 32 (8), 1570–1575.
- 12 N. Hartrampf, A. Saebi, M. Poskus, Z. P. Gates, A. J. Callahan, A. E. Cowfer, S. Hanna, S. Antilla, C. K. Schissel, A. J. Quartararo, X. Ye, A. J. Mijalis, M. D. Simon, A. Loas, S. Liu, C. Jessen, T. E. Nielsen and B. L. Pentelute, *Science* **2020**, 368 (6494), 980–987.
- 13 P. G. Isenegger and B. G. Davis, *J. Am. Chem. Soc.* **2020**, 141 (20), 8005–8013.
- 14 C. A. Pérez Ríquez, O. Abian and J. M. Palomo, *Chem. Commun.* **2019**, 55, 12928–12931.
- 15 H. A. Beard, J. R. Hauser, M. Walko, R. M. George, A. J. Wilson and R. S. Bon, *Commun. Chem.* **2019**, 2 (1), 1–9.
- 16 S. Tsukiji, M. Miyagawa, T. Ogawa, Y. Koshi, I. Hamachi and E. Nakata, *J. Am. Chem. Soc.* **2007**, 130 (1), 245–251.
- 17 T. Tamura, Z. Song, K. Amaike, S. Lee, S. Yin, S. Kiyonaka and I. Hamachi, *J. Am. Chem. Soc.* **2017**, 139 (40), 14181–14191.
- 18 S. R. Adusumalli, D. G. Rawale, U. Singh, P. Tripathi, R. Paul, N. Kalra, R. K. Mishra, S. Shukla and V. Rai, *J. Am. Chem. Soc.* **2018**, 140 (44), 15114–15123.
- 19 J. J. Bruins, A. H. Westphal, B. Albada, K. Wagner, L. Bartels, H. Spits, W. J. H. Van Berkel and F. L. Van Delft, *Bioconjugate Chem.* **2017**, 28 (4), 1189–1193.
- 20 J. F. Keijzer and B. Albada, *Bioconjugate Chem.* **2020**, 31 (10), 2283–2287.
- 21 T. Masuzawa, S. Sato, T. Niwa, H. Taguchi, H. Nakamura and T. Oyoshi, *Chem. Commun.* **2020**, 56 (78), 11641–11644.
- 22 D. G. Rawale, K. Thakur, P. Sreekumar, S. T. K., R. A., S. R. Adusumalli, R. K. Mishra and V. Rai, *Chem. Sci.* **2021**, 12 (19), 6732–6736.
- 23 J. F. Keijzer, J. Firet and B. Albada, *Chem. Commun.* **2021**, 57, 12960–12963.
- 24 S. E. Lee, A. Sidorov, T. Gourelain, N. Mignet, S. J. Thorpe, J. A. Brazier, M. J. Dickman, D. P. Hornby, J. A. Grasby and D. M. Williams, *Nucleic Acids Res.* **2001**, 29 (7), 1565–1573.
- 25 D. F. Sauer, M. Bocola, C. Broglia, M. Arlt, L. L. Zhu, M. Brocker, U. Schwaneberg and J. Okuda, *Chem. - An Asian J.* **2015**, 10 (1), 177–182.
- 26 F. Wang, C. Lu and I. Willner, *Chem. Rev.* **2014**, 114, 2881–2941.
- 27 K. I. Mortensen, J. Sung, H. Flyvbjerg and J. A. Spudich, *Nat. Commun.* **2015**, 6 (8621), 1–9.
- 28 N. R. Ileperuma, S. D. G. Marshall, C. J. Squire, H. M. Baker, J. G. Oakeshott, R. J. Russell, K. M. Plummer, R. D. Newcomb and E. N. Baker, *Biochemistry* **2007**, 46 (7), 1851–1859.
- 29 A. Pica, I. R. Krauss, V. Parente, H. Tateishi-Karimata, S. Nagatoishi, K. Tsumoto, N. Sugimoto and F. Sica, *Nucleic Acids Res.* **2017**, 45 (1), 461–469.
- 30 C. Wang, L. Yue and I. Willner, *Nat. Catal.* **2020**, 3 (11), 941–950.
- 31 J. B. Trads, T. Tørring and K. V. Gothelf, *Acc. Chem. Res.* **2017**, 50 (6), 1367–1374.
- 32 D. Zhao, Y. Kong, S. Zhao and H. Xing, in *DNA Nanotechnology*, eds. C. Fan and Y. Ke, Springer, Cham, 1st edn., **2020**, pp. 83–124.

Appendix 5

A5.1 DNA codes with their respective topologies and nucleotide sequences

Table A5.1. The codes and nucleobase sequences for each DNA strand used in this study. “5Hexynyl” indicates a hexynyl attached at the 5’ end; “35OctdU” indicates a 3’ end thymine nucleobase with an octynyl attached; “i5OctdU” indicates an internal thymine nucleobase with an octynyl attached.

Code	DNA Sequence (5’ to 3’)
DNA_{temp}	/5Hexynyl/ AAA ATA TAT ATA TAT AAA A
DNA_{catalyst}	TTT TAT ATA TAT ATA TTT T
DNA_{cat}-T1-alkyne	TTT TAT ATA TAT ATA TTT /35OctdU/
DNA_{cat}-T2-alkyne	TTT TAT ATA TAT ATA TT/i5OctdU/ T
DNA_{cat}-T3-alkyne	TTT TAT ATA TAT ATA T/i5OctdU/T T
DNA_{cat}-T4-alkyne	TTT TAT ATA TAT ATA /i5OctdU/TT T
DNA_{cat}-T6-alkyne	TTT TAT ATA TAT A/i5OctdU/A TTT T
DNA_{cat}-T8-alkyne	TTT TAT ATA TA/i5OctdU/ ATA TTT T
PW17	GGG TAG GGC GGG TTG GG
DNA_{cat}-T0-PW17	TTT TAT ATA TAT ATA TTT TGG GTA GGG CGG GTT GGG
DNA_{cat}-T4-PW17	TTT TAT ATA TAT ATA GGG TAG GGC GGG TTG GG
DNA_{cat}-T8-PW17	TTT TAT ATA TAG GGT AGG GCG GGT TGG G
DNA_{cat}-T10-PW17	TTT TAT ATA TGG GTA GGG CGG GTT GGG ATA TAT TTT
DNA_{cat}-T20-PW17	GGG TAG GGC GGG TTG GGT TTT ATA TAT ATA TAT TTT

A5.2 Protocol for modification of protein-DNA_{temp}

A mixture was typically prepared containing (a) 20 μ M GRX-DNA_{temp} (from 100 μ M stock in 20 mM NH₄HCO₃ pH: 8.0) (b) 20–25 μ M CHY-DNA_{temp} (from 150–185 μ M stock in 20 mM NH₄HCO₃ pH: 8.0) (c) 20–26 μ M TRM-DNA_{temp} (from a 150–200 μ M stock solution in 20% glycerol in 20 mM NH₄HCO₃ pH: 8.0), 30 μ M DNA_{catalyst} (from varying stock concentrations in ddH₂O) in HEPES buffer [50 mM, pH: 8.0, with 350 mM NaCl and 50 mM KCl]. This mixture was incubated in the dark for 20–30 min at 37 °C, after which acyl donor (from varying stock concentrations in DMSO) was added. The reaction mixture was again incubated in the dark at 37 °C, shaking the tubes at 500 rpm (reaction times: 2 hrs for DMAP catalyst, 6 hrs for PyOx).

Prior to SDS-PAGE analysis, additional functionalization is required to visualize the modifications. Two approaches were used: band shifting or fluorescent staining.

- **Band shifting:** protein modified with an azide-carrying acyl donor was treated with at least 6 equivalents of BCN-PEG₂₀₀₀ (purchased from Synaffix B.V.) with respect to the concentration of acyl donor and incubated at 12 °C overnight.
- **Fluorescent staining:** protein modified with an alkyne-carrying acyl donor was treated with 6 equivalents of BCN-sulphorhodamine B with respect to the acyl donor, 1 and incubated at 12 °C overnight.

A5.3 Protocol for SDS-PAGE analysis

Acrylamide gels (12%) were prepared according to Bio-Rad bulletin 6201 protocol. Specifically, reaction mixtures containing 2–5 µg of protein were diluted with one volume equivalent of SDS-PAGE sample buffer (2×) containing 10% BME and incubated for 10 minutes at 95 °C. The denatured sample was then used for SDS-PAGE analysis (12% acrylamide gel). Precision Plus Protein™ Dual Color Standards was used as a reference protein ladder. After running, if one of the proteins was modified with a fluorophore, a UV-photo of the gel was taken. Gels were then stained using Coomassie brilliant blue (0.1% Coomassie Blue R250 in 10% acetic acid, 50% methanol and 40% demineralized water) by shaking gently for 0.5 hours, and destained with destaining solution (10% acetic acid, 50% methanol, and 40% demineralized water) by shaking gently for 1 hour. Afterwards, the destaining solution was replaced with H₂O and shaken gently overnight at room temperature. When the BCN-PEG₂₀₀₀ mass-tag was used, quantification was performed by integrating the intensity of the Coomassie stained bands of the SDS-PAGE gel using ImageJ software.

A5.4 Protocol for the analysis of protein modification on HPLC(-MS)

The reaction mixture was aspirated three times with a pipette, after which 2–10 µL was added to an HPLC vial insert that already contained 18–10 µL of buffer (200 mM Citrate and 400 mM NaCl; pH: 5.5). The resulting mixture was also aspirated three times. This sample was then run over a Thermo Fischer MAbPAC RP column 3.0 × 100 mm, at 80 °C, the gradient varying per protein. For thrombin, the gradient started with 23% (ACN + 0.1% FA) ending with 33% (ACN + 0.1% FA) in (95% H₂O + 5% ACN + 0.1% FA) (flow rate 0.5 mL/min) over 25 min. The system used was an Agilent 1220 Infinity LC system with DAD detector.

For mass spectrometry analysis, reaction mixtures were diluted to a final protein concentration of 0.25 mg/mL. Protein samples were then analysed on a Thermo Scientific™ Q Exactive Focus Orbitrap using the same gradient as was used for the HPLC analyses.

A5.5 Tryptic digestion of protein and subsequent analysis to determine site-specificity of the modification

Modified protein samples were subjected to SDS-PAGE separation and the desired protein bands cut from the gel and cut up to small pieces. The pieces were washed by incubating three times with 50 mM NH₄HCO₃ (pH: 8.0) in 50% ACN in ddH₂O and subsequently dried in a Speedyvac vacuum centrifuge. The dry pieces were swollen in 50 µL DTT [10 mM in 100 mM NH₄HCO₃ (pH: 8.0)] and incubated for 45 minutes at 56 °C. The supernatant was removed and 50 µL of IAA (55 mM in 100 mM NH₄HCO₃ (pH: 8.0)) was added and the pieces were incubated in the dark at rt for 30 min. The supernatant was removed and the pieces were washed by incubating once with 50 mM NH₄HCO₃ (pH: 8.0) in 50% ACN in ddH₂O and subsequently dried in a vacuum centrifuge. The gel pieces were swollen in 40 µL trypsin gold (125 ng/µL) and incubated at 37 °C for 16–18 h. The initial supernatant was collected and the gel pieces were washed by incubating 15 min at 37 °C with 20 µL NH₄HCO₃ (100 mM, pH: 8.0) and 15 min at 37 °C when diluted with 20 µL. The collected supernatants were combined and dried in a vacuum centrifuge and the dry peptide digest dissolved in 20 µL 0.1% FA.

Peptide digests were analysed on an EASY nanoLC connected to Thermo Scientific™ Q Exactive PLUS. Peptides were trapped onto a PepSep trap column (2 cm × 100 µm ID, 5 µm C18 ReproSil) and subsequently separated on a PepSep analytical column (8 cm × 75 µm ID, 3 µm C18 ReproSil, PepSep).

Elution was achieved using a gradient that started with 5% (ACN + 0.1% FA) ending with 40% (ACN + 0.1% FA) in (H₂O + 0.1% FA), washing the column with 80% (ACN + 0.1% FA) afterwards.

The eluted peaks were analysed using MaxQuant software, searching for peptides with mass modification corresponding to H(11)O(3)C(7)N(3) (*i.e.*, the substitution of a proton on the protein by the acyl group of thioester **1** or ANANS **3**) and limiting criteria of 1% PSM FDR and a minimal peptide score of 80. As protein database human proteome was used, obtained from www.uniprot.org (code: UP000005640).

A5.6 Synthesis of GRX-DNA_{temp}

Glutaredoxin 1 was incubated with DTT (1.5 eq.) in 15 mM NH₄HCO₃ (pH: 8.0) at 37 °C for 30 min, followed by incubation with 1-(azidomethyl)-3,5-bis(bromomethyl)benzene (2.5 eq.) at 37°C for 2.5 h and subsequently purified by spin filtration over 3 kDa MWCO Amicon® Ultra-15 Centrifugal Filter Units, washing 3 times with 50 mM NaCl solution in 25 mM Tris (pH: 8.0). The synthesized azido-protein was treated with 2 equivalents of DNA_{temp} (from 500 µM stock in oxygen-poor ddH₂O) with respect to the protein concentration, 100 µM [Cu•THPTA] (complex of CuSO₄ and THPTA mixed in a ratio of 1:5 in ddH₂O) and 10 mM sodium ascorbate (from a freshly made stock of 100 mM in oxygen-poor ddH₂O) and incubated in the dark at 12 °C for 16–20 hours. The formed GRX-DNA_{temp} construct was purified by FPLC, using an ion-exchange MonoQ column (Vol: 1 mL) using a gradient from 0-1 M NaCl in 20 mM Tris (pH: 8.0). The collected fractions were and concentrated by spin filtration over 3 kDa MWCO Amicon® Ultra-15 Centrifugal Filter Units, washing 3 times with 50 mM NaCl solution in 25 mM Tris (pH: 8.0). The concentration of GRX-DNA_{temp} was quantified from absorption values determined with a Scientific™ Nanodrop 2000, using a 1:1 mixture of native GRX with DNA_{temp} as a reference.

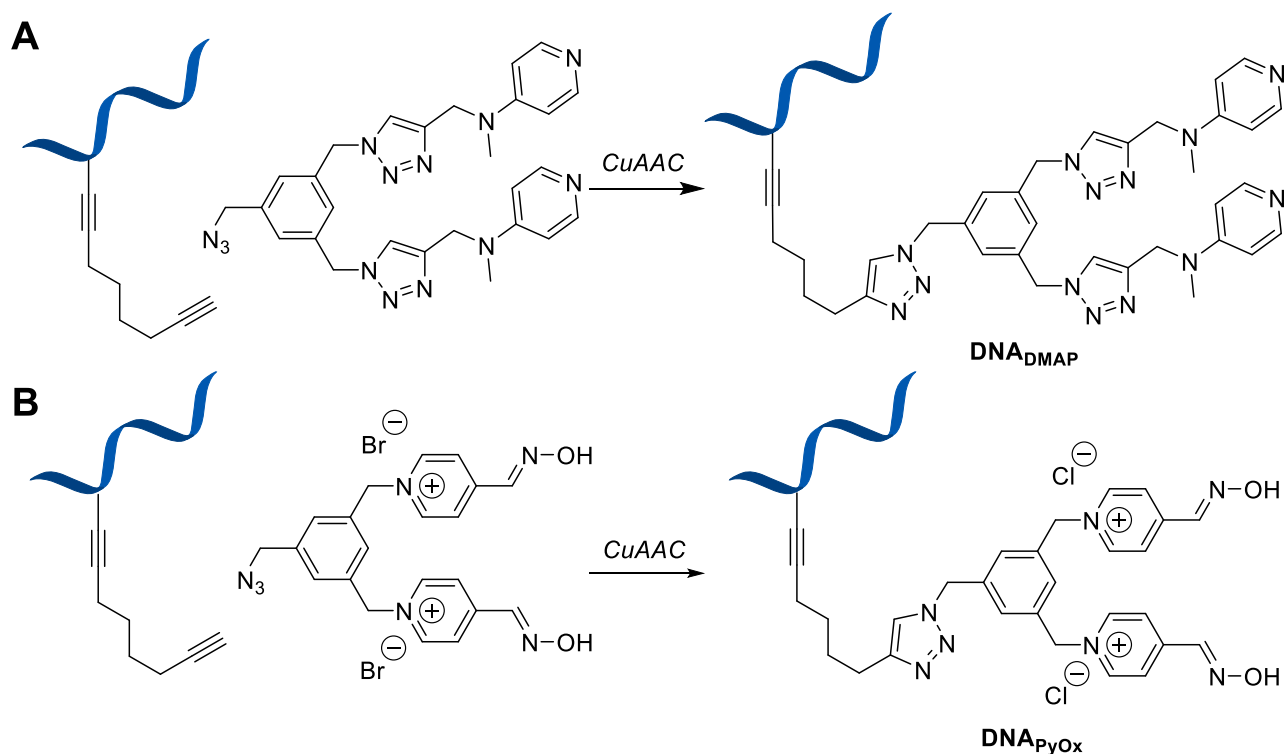
A5.7 Synthesis of CHY-DNA_{temp} and TRM-DNA_{temp}

Thrombin or Chymotrypsin was incubated with azido-paraoxon (20 eq.) in 20% glycerol in 50 mM HEPES (pH: 7.2) at 37 °C for 3 h and subsequently purified by spin filtration over 10 kDa MWCO Amicon® Ultra-15 Centrifugal Filter Units, washing 3 times with 20% glycerol in 50 mM NaCl solution in 25 mM Tris (pH: 8.0) for thrombin and 50 mM NaCl solution in 25 mM Tris (pH: 8.0) for chymotrypsin. The respective azido-protein was treated with 2-3 equivalents of DNA_{temp} (from 500 µM stock in oxygen-poor ddH₂O) with respect to the protein concentration, 100–150 µM [Cu•THPTA] (complex of CuSO₄ and THPTA mixed in a ratio of 1:5 in ddH₂O) and 7–10 mM sodium ascorbate (from a freshly made stock of 100 mM in oxygen-poor ddH₂O) and incubated in the dark at 12 °C for 16–20 hours. The synthesized protein-DNA_{temp} constructs were purified by spin filtration over 10 kDa MWCO Amicon® Ultra-15 Centrifugal Filter Units, washing 3 times with 20% glycerol in 50 mM NaCl solution in 25 mM Tris (pH: 8.0) for thrombin and 50 mM NaCl solution in 25 mM Tris (pH: 8.0) for chymotrypsin. The concentrations were quantified from absorption values determined with a Scientific™ Nanodrop 2000, using a 1:1 mixture of native protein with DNA_{temp} as a reference.

A5.8 Synthesis of DNA_{DMAP} and DNA_{PyOx}

DNA-alkyne sequences with alkyne-thymine modification were purchased as HPLC-purified lyophilized powders from Integrated DNA technologies. The powders were dissolved in oxygen-poor ddH₂O. The DNA was treated with 10 equivalents of compound azido-diDMAP or azido-diPyOx (from 100 mM stock in DMSO) with respect to the DNA concentration, 100 µM [Cu•THPTA] (complex of CuSO₄ and THPTA mixed in a ratio of 1:5 in ddH₂O) and 10 mM sodium ascorbate (from a freshly made stock of 100 mM in

ddH₂O) and incubated in the dark at 12 °C for 16–20 hours. The synthesized DNA_{DMAP} or DNA_{PyOx} construct was purified by spin filtration over 3 kDa MWCO Amicon® Ultra-15 Centrifugal Filter Units, washing three times with 400 mM NaCl solution in ddH₂O. Purity and concentration were determined by HPLC-MS and UV-Vis.



Scheme A5.1. Syntheses of DNA_{catalyst} constructs. (A) DNA-alkyne and azido-diPyOx or (B) azido-diDMAP are coupled via copper-catalysed alkyne-azide click chemistry and purified via spin filtration.

A5.9 GRX-DNA_{temp} with DNA_{DMAP} & 100 μ M thioester

Lane#	1	2	3	4	5	6	7	8
Cat pos	T1	T2	T3	T4	T6	T8	-	free
Conv%	62	56	60	52	44	36	-	15

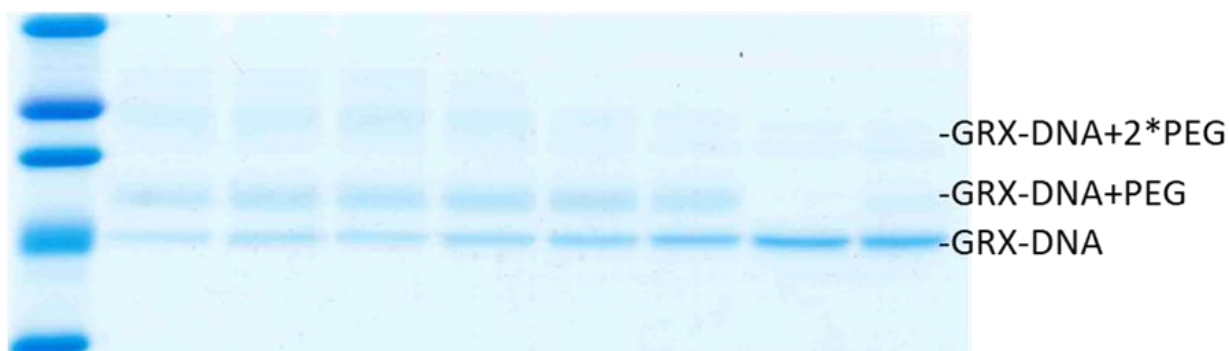


Figure A5.1. GRX-DNA_{temp} modified by DNA_{DMAP} with thioester 1 and afterwards PEGylated with BCN-PEG₂₀₀₀. Conversions were calculated using ImageJ. Conditions: 20 μ M GRX-DNA_{temp}, 22 μ M DNA_{DMAP} and 100 μ M thioester 1, pH: 8.0, at 37 °C for 2 h.

A5.10 GRX-DNA_{temp} with DNA_{PyOx}

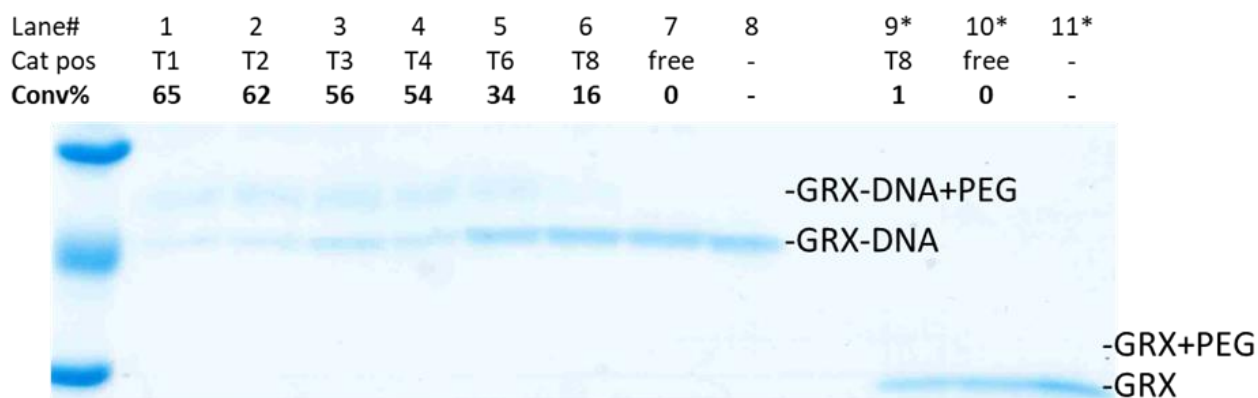


Figure A5.2. GRX-DNA_{temp} modified by DNA_{PyOx} with ANANS 3 and afterwards PEGylated with BCN-PEG₂₀₀₀. Conversions were calculated using ImageJ. Slots 9-11 contain native GRX. Conditions: 20 μ M GRX-DNA_{temp}, 22 μ M DNA_{PyOx} and 300 μ M ANANS 3, pH: 7.2, at 37 °C for 6 h. * = reactions with native GRX.

A5.11 GRX-DNA_{temp} with DNA_{cat}-hGQ DNAzyme

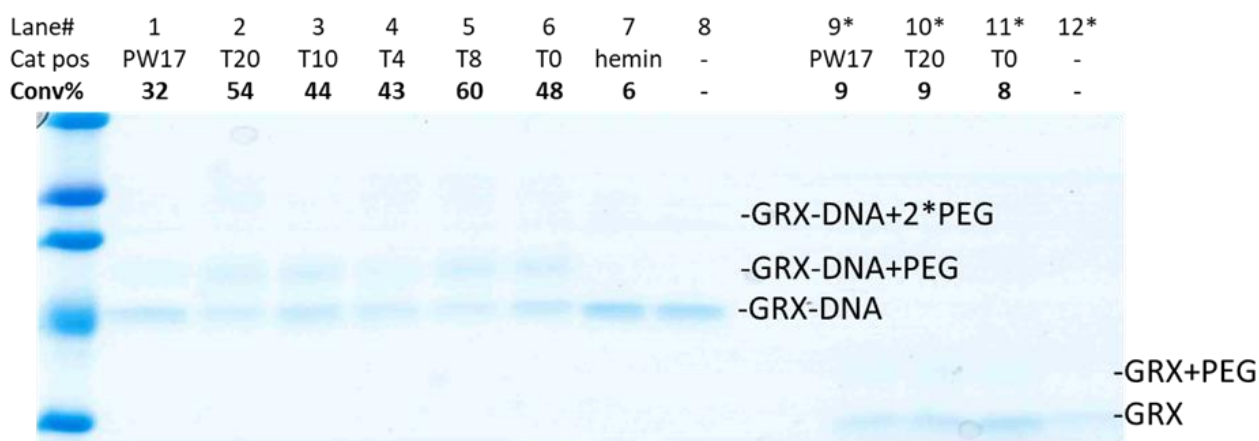


Figure A5.3. GRX-DNA_{temp} modified by DNA_{cat}-hGQ with NML 1 and afterwards PEGylated with BCN-PEG₂₀₀₀. Conversions were calculated using ImageJ. Slots 9-12 contain native GRX. Conditions: 20 μ M GRX-DNA_{temp}, 22 μ M DNA_{cat}-hGQ, 30 μ M NML 1 and 100 μ M H₂O₂, pH: 7.0, at 25 °C for 30 min. * = reactions with native GRX.

A5.12 Degradation of TRM-DNA_{temp}

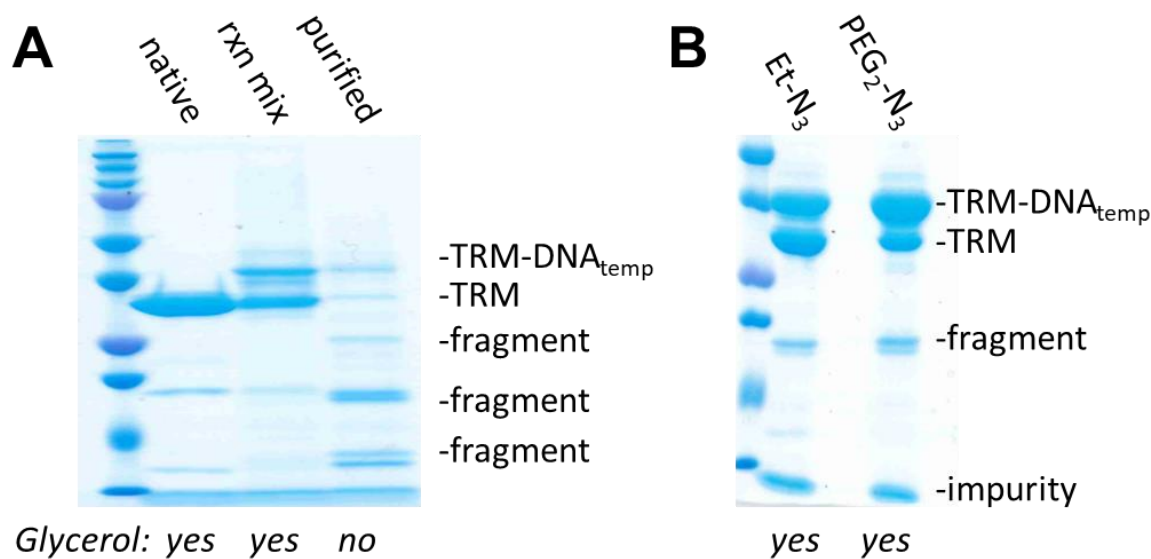


Figure A5.4. (A) TRM-DNA_{temp} degrades after purification, which did not occur when it was still in the reaction mixture, which contained 5% glycerol. As such, (B) purification with spin filtration and 10% glycerol was performed and there TRM-DNA_{temp} did not degrade.

A5.13 Example of CHY-Et-DNA_{temp} on gel

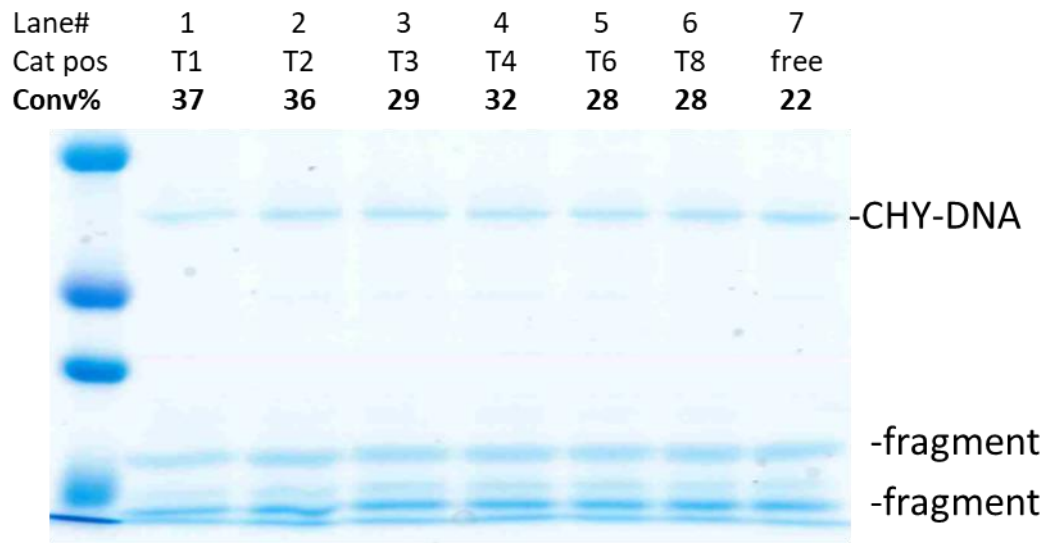


Figure A5.5. CHY-Et-DNA_{temp} modified by DNA_{DMAP} with thioester **1** and afterwards PEGylated with BCN-PEG₂₀₀₀. Conversions are normalized and were calculated using ImageJ. Conditions: 20 μ M CHY-Et-DNA_{temp} with 23 μ M DNA_{DMAP} and 100 μ M thioester **1**, pH: 8.0, at 37 $^{\circ}$ C for 2 h.

A5.14 Example of CHY-EG₂-DNA_{temp} on gel

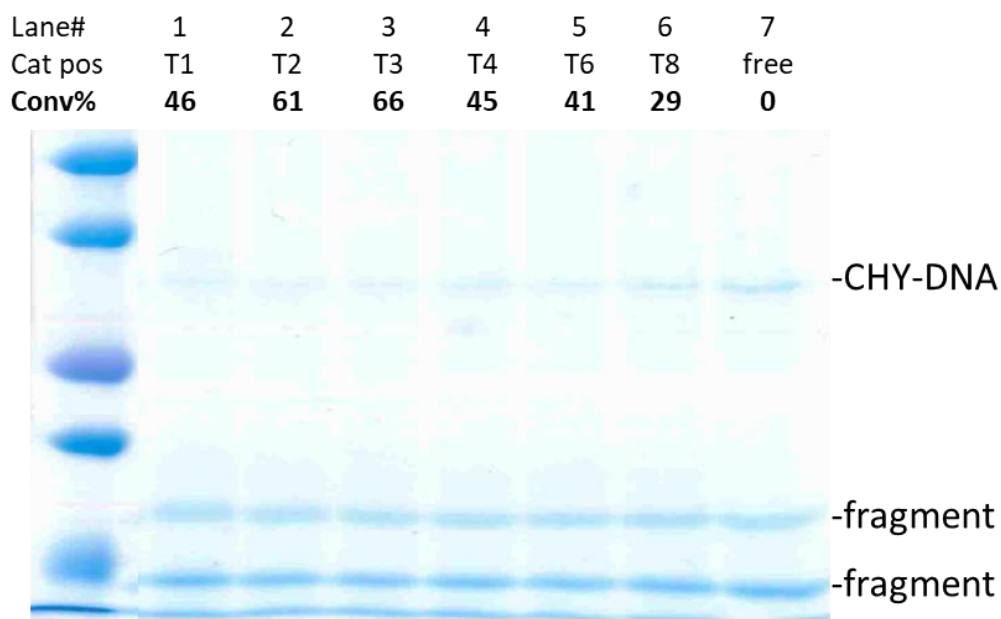


Figure A5.6. CHY-EG₂-DNA_{temp} modified by DNA_{DMAP} with thioester **1** and afterwards PEGylated with BCN-PEG₂₀₀₀. Conversions are normalized and were calculated using ImageJ. Conditions: 25 μ M CHY-EG₂-DNA_{temp} with 28 μ M DNA_{DMAP} and 100 μ M thioester **1**, pH: 8.0, at 37 °C for 2 h.

A5.15 Example of TRM-Et-DNA_{temp} on gel

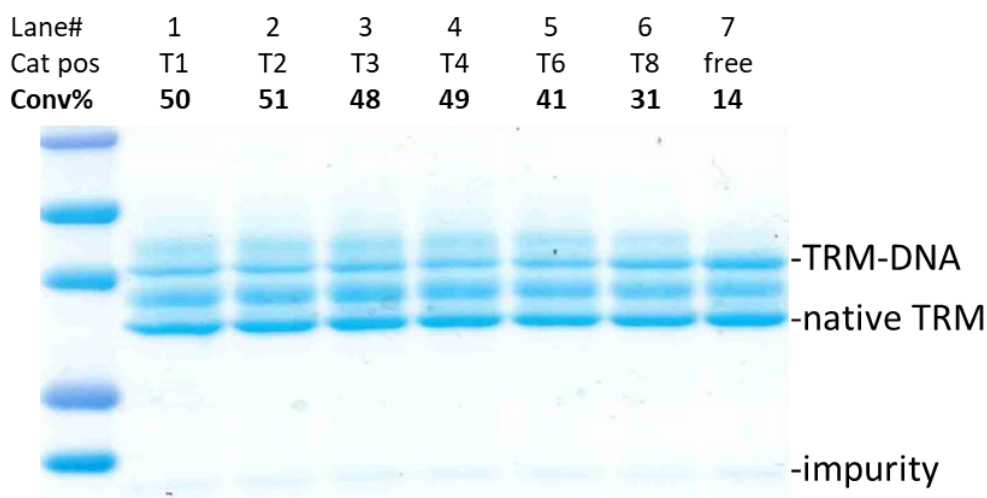


Figure A5.7. TRM-Et-DNA_{temp} modified by DNA_{DMAP} with thioester **1** and afterwards PEGylated with BCN-PEG₂₀₀₀. Conversions are normalized and were calculated using ImageJ. Conditions: 20 μ M TRM-Et-DNA_{temp} with 23 μ M DNA_{DMAP} and 100 μ M thioester **1**, pH: 8.0, at 37 °C for 2 h.

A5.16 Example of TRM-EG₂-DNA_{temp} on gel

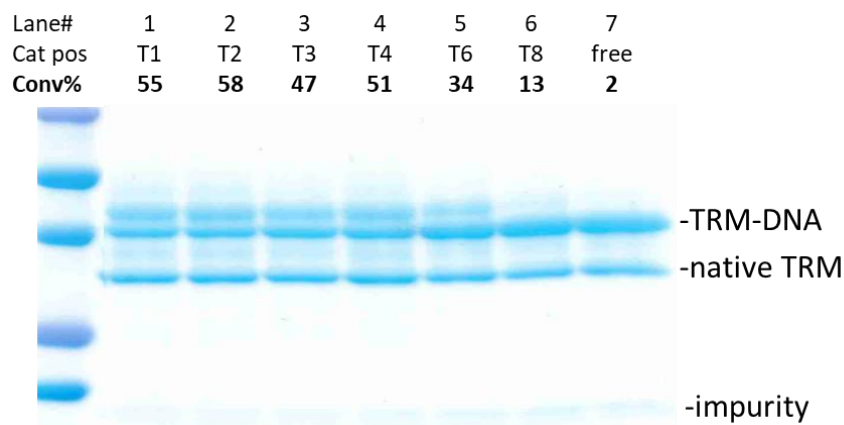


Figure A5.8. TRM-EG₂-DNA_{temp} modified by DNA_{pyOx} with ANANS **3** and afterwards PEGylated with BCN-PEG₂₀₀₀. Conversions are normalized and were calculated using ImageJ. Conditions: 26 μ M TRM-EG₂-DNA_{temp} with 28 μ M DNA_{pyOx} and 300 μ M ANANS **3**, pH: 7.2, at 37 °C for 2 h.

A5.17 TRM-Et-DNA_{temp} with DNA_{cat}-hGQ DNAzyme

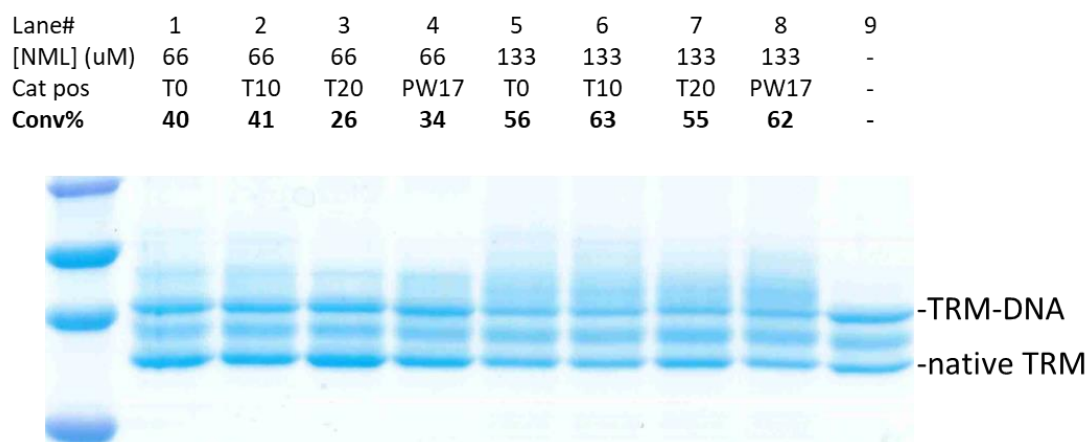


Figure A5.9. TRM-Et-DNA_{temp} modified by DNA_{cat}-hGQ DNAzyme with NML **1** and afterwards PEGylated with BCN-PEG₂₀₀₀. Conversions were calculated using ImageJ. Conditions: 20 μ M TRM-Et-DNA_{temp}, 30 μ M DNA_{cat}-hGQ and [NML **1**] = [H₂O₂] = 66 or 133 μ M, pH: 7.0, at 25 °C for 30 min.

A5.18 TRM-EG₂-DNA_{temp} with DNA_{cat}-hGQ DNzyme

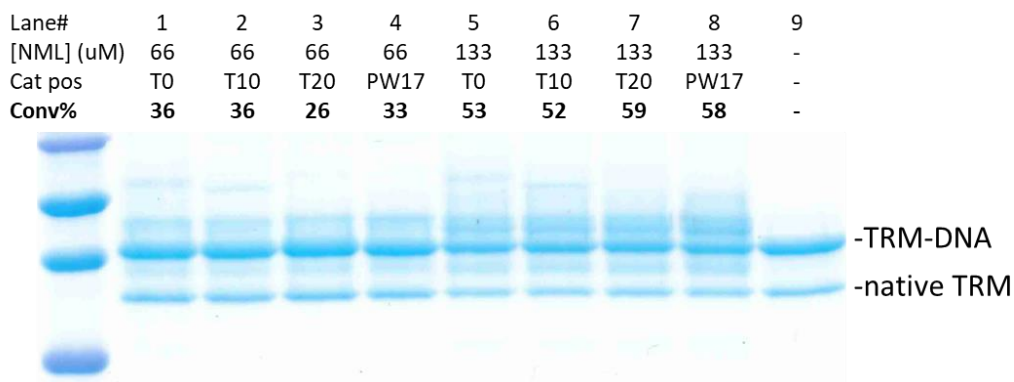


Figure A5.10 TRM-EG₂-DNA_{temp} modified by DNA_{cat}-hGQ DNzyme with NML **1** and afterwards PEGylated with BCN-PEG₂₀₀₀. Conversions were calculated using ImageJ. Conditions: 26 μ M TRM-EG₂-DNA_{temp}, 30 μ M DNA_{cat}-hGQ and [NML **1**] = [H₂O₂] = 66 or 133 μ M, pH: 7.0, at 25 °C for 30 min.

A5.19 MS data of CHY-EG₂-N₃ and TRM-EG₂-N₃

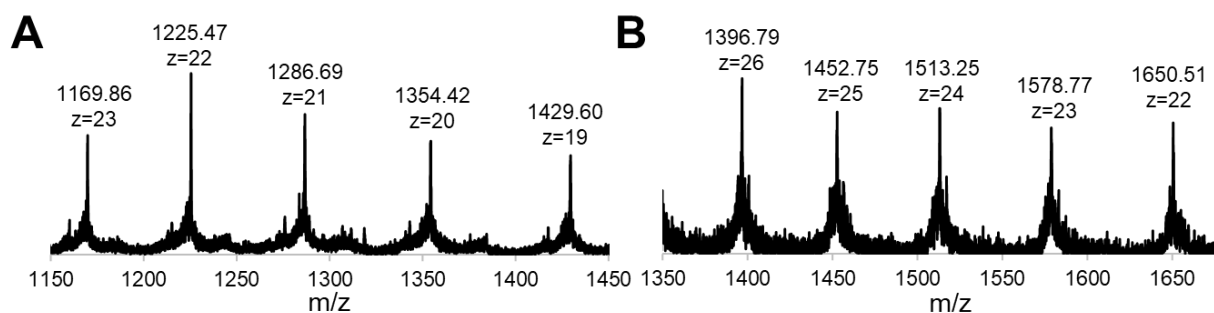


Figure A5.11 (A) Mass spectrometry data of CHY-EG₂-DNA_{temp}. Calculated peaks: mass(*z*) = 1169.8(22), 1225.4(21), 1286.7(20), 1354.3(19), 1429.5(18); (B) Mass spectrometry data of TRM-EG₂-DNA_{temp}. Calculated peaks: mass(*z*) = 1396.8(26), 1452.7(25), 1513.1(24), 1578.9(23), 1650.6(22).

A5.20 Estimated lengths of components

Table A5.2. Lengths of all components in this work. Values were estimated using YASARA with each molecule in the most linear state. The independent linker lengths include also the hexynyl chain of the DNA, thus indicating the distance between protein surface and the 5' end of DNA_{temp}.

Code	Length (nm)	Code	Length (nm)
GRX-DNA(T1)	2.2	CHY/TRM-EG ₂ -DNA(T1)	2.5
GRX-DNA(T2)	2.5	CHY/TRM-EG ₂ -DNA(T2)	2.8
GRX-DNA(T3)	2.9	CHY/TRM-EG ₂ -DNA(T3)	3.2
GRX-DNA(T4)	3.2	CHY/TRM-EG ₂ -DNA(T4)	3.5
GRX-DNA(T6)	3.9	CHY/TRM-EG ₂ -DNA(T6)	4.2
GRX-DNA(T8)	4.6	CHY/TRM-EG ₂ -DNA(T8)	4.9
CHY/TRM-Et-DNA(T1)	2.0	GRX linker	1.8
CHY/TRM-Et-DNA(T2)	2.4	Et linker	1.7
CHY/TRM-Et-DNA(T3)	2.7	EG ₂ linker	2.1
CHY/TRM-Et-DNA(T4)	3.1	Octynyl (thymine linker)	1.0
CHY/TRM-Et-DNA(T6)	3.8	diDMAP (N ₃ to acyl group)	1.4
CHY/TRM-Et-DNA(T8)	4.5	diPyOx (N ₃ to acyl group)	1.3

A5.21 HPLC and MS data of DNA_{DMAP}

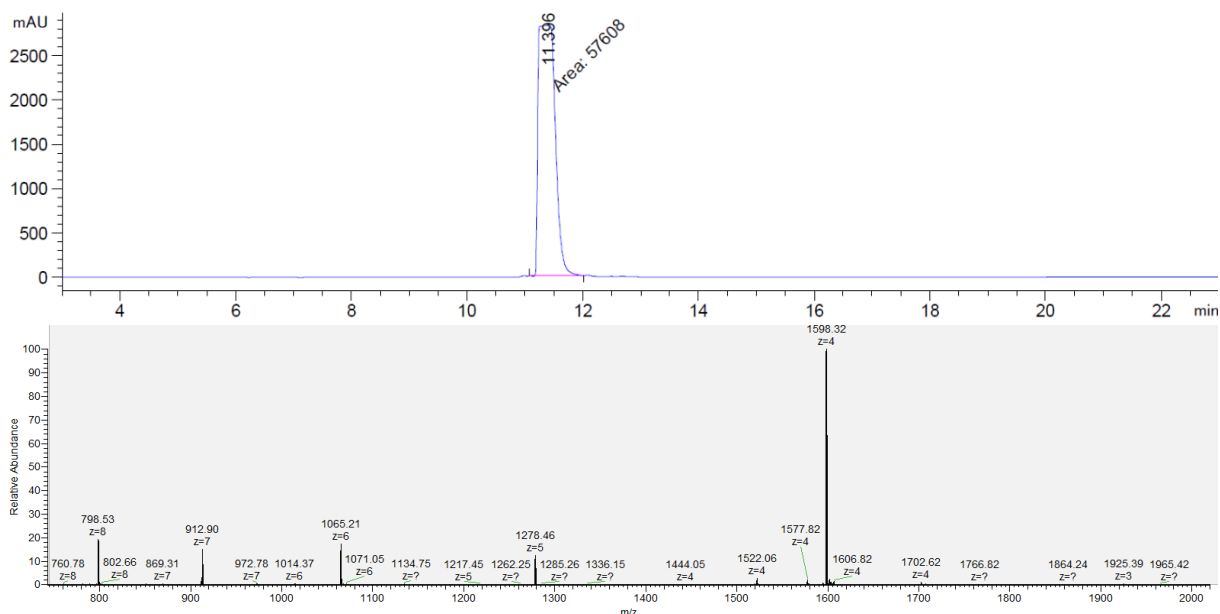


Figure A5.12 DNA_{DMAP}¹ HPLC trace (Abs 260 nm) and mass spectrometry data. Calculated peaks: mass(z) = 798.5(8), 912.9(7), 1065.2(6), 1278.4(5), 1598.3(4). Spectra of the other diDMAP constructs looked the same.

A5.22 HPLC and MS data of DNA_{PyOx}

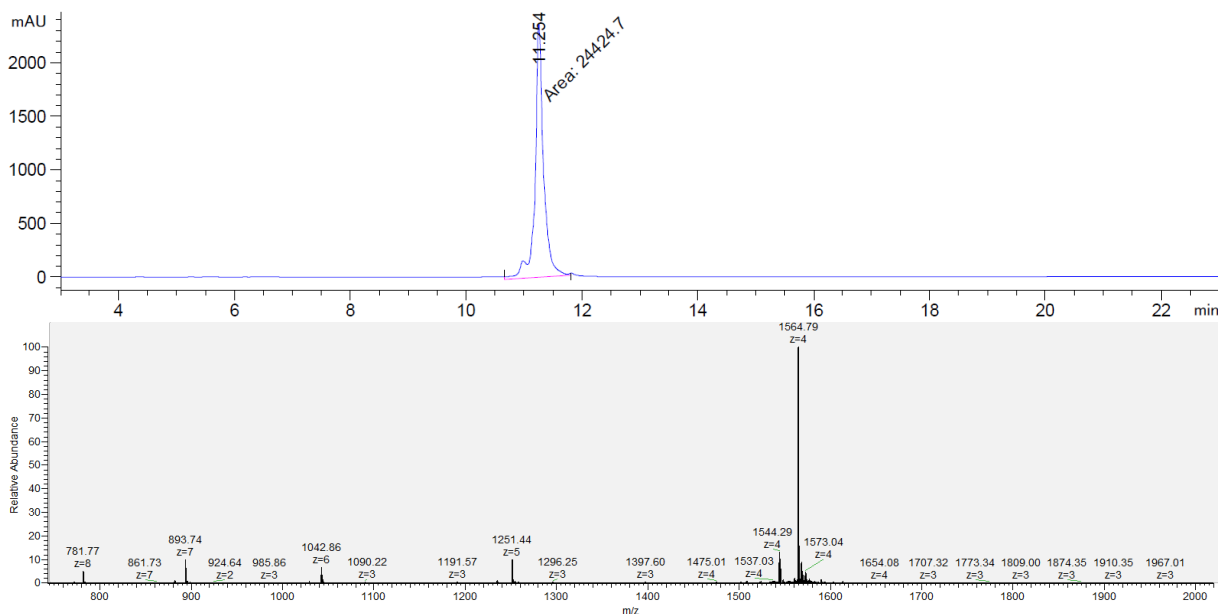


Figure A5.13 DNA_{PyOx}¹ HPLC trace (Abs 260 nm) and mass spectrometry data. Calculated peaks: mass(z) = 781.9(8), 893.7(7), 1042.7(6), 1251.6(5), 1564.8(4). Spectra of the other diPyOx constructs looked the same.

Chapter 6



General Discussion

6.1 General discussion of DNA for protein modification

As the work in this thesis has demonstrated, catalytically activated DNA systems can be effective tools for native protein modification. Even though the modification is not conferred by systems consisting of solely DNA, the DNA scaffold is vital for the catalytic protein modification reaction, as its removal reduced the modification quantity to near insignificant. In fact, we exploited this by incorporating a conformation switch that controls the activity of the catalytic DNA constructs. By locking away the DNA in a dsDNA duplex, the modification percentage dropped to background levels for all systems, and this could be reversed by releasing the DNA to its original ssDNA form. As such the described protein modification methodologies are indeed regulated by DNA.

In order to use DNA for protein modification, a few hurdles needed to be overcome to make the systems work in a consistent manner and to accurately analyse the reaction mixtures. The first issue that arises when mixing DNA and proteins (in concentrations above $\sim 1\ \mu\text{M}$ and higher) is the risk of precipitation. We hypothesised that this is caused by the phosphodiester backbone that forms ion pairs with positively charged protein residues (*i.e.*, most notably Lys and Arg). Indeed, the formed precipitate could be dissolved by raising the pH above the pI value of the protein or by using a buffer with high ionic strength, *i.e.*, salt concentrations $\geq 400\ \text{mM}$ (NaCl or KCl). Although we witnessed no influence by this ionic strength on protein activity (activity of lysozyme with and without $\geq 400\ \text{mM}$ NaCl was the same), the effect of salt concentration on the stability of proteins is still not fully understood and could possibly be a hurdle in future applications.¹⁻³

The differences in charges of DNA and protein also mean that optimization for simultaneous analysis on the same technique is required. Concentrated LC-MS samples will precipitate during elution as the salt concentration of the samples is diluted during the run. The eluent for protein-based LC-MS is often kept at acidic pH to have a more even charge distribution, which gives higher ion counts and more slender peaks. Acidic pH will however, cause the reverse effect for the ionisation of negatively charged DNA, which is better detected at more neutral pH (6-8). On the other hand, neutral pH has the advantage that the eluent contains a (volatile) salt buffer (*e.g.*, NH_4OAc or Et_3NHOAc), enabling the injection of concentrated samples, although this will also result in a decreased detection in MS due to ion suppression and variable masses due to varying numbers of salt ion adducts.

To circumvent this, we digested soluble DNA using DNase I, followed by spin filtration to remove the small nucleotide fragments. This was even required as the relatively high DNA concentrations interfered with protein digestion by trypsin (chapter 2). Should it be preferred to avoid the use of DNase I, SDS-PAGE separation of the mixtures is the alternative (chapters 4 and 5). In this, proteins and DNA can be separated although interpretation is only possible with suitable controls for protein and DNA. In addition, DNA requires different gel staining reagents for visualisation than those used for proteins. Fortunately, after DNA visualisation with quercetin stain,⁴ protein staining with Coomassie Blue was virtually unaffected. For more accurate DNA analyses, agarose gel electrophoresis could be an alternative.⁵

6.2 Additional insights into the hGQ DNzyme

The work of chapter 2 demonstrates that the horseradish peroxidase-mimicking hGQ DNzyme is an effective catalyst for protein modification and it adds protein modification to the DNzyme toolbox. The DNzyme mimics enzymes not only in oxidative activity, but also in tuneable selectivity and controllable output. The influence of the reactant concentrations were already described in the chapter, but later experiments and literature suggested other conditions could potentially further benefit the hGQ DNzyme and the NML conjugation reaction it induces.

Temperature usually has great impact on reaction kinetics,⁶ although for the DNzyme-induced conjugation, not much difference was observed when considering temperatures between 16 °C and 40 °C. Concerning the pH, however, the effect on the hGQ DNzyme was significantly notable, as it operates optimally at pH between 5.5 and 7.5. Interestingly, the type of buffering agent was even more crucial. Citrate and phosphate-buffered solutions worked well, as did cacodylate ($C_2H_7AsO_2$). In stark contrast, the so-called Good's buffers HEPES and TAPS performed very poorly, almost completely suppressing the NML conjugation when compared to phosphate buffer at pH 8. Similarly, in Tris buffer the DNzyme was only able to generate half of the conversions it could attain in phosphate buffer. As these three buffers share few structural similarities, it is difficult to predict how other buffering agents would affect this particular DNzyme reaction.

The additional components of the buffer are also important. In line with other works,⁷ we found that K^+ is an important ion for DNzyme activity. Without K^+ , modification generated by the DNzymes drops by roughly 20%, which is accordance with literature data that state the ion stabilises the GQ structure.⁸ We did not yet test other cations, such as $Hg(II)$ or $Pb(II)$, even though these were also reported to enhance stability and activity.^{8,9} In addition, work from Stefan *et al.*¹⁰ demonstrates that having the nucleotide supplements CTP or ATP in solution with the hGQ DNzyme results in a higher oxidation rates of ABTS and 3,3',5,5'-tetramethylbenzidine,¹¹ thus their inclusion could also be considered. Similarly, our work did not use any surfactants, but work by Kosman *et al.*¹² demonstrated that detergents Triton X-100 and Tween 20 can increase ABTS oxidation rate of the PS2.M DNzyme, whereas the use of the detergent Brij 58 improved the oxidation by the HT DNzyme. This shows that different surfactants can stabilize different GQ topologies and/or the resulting DNzymes (Table A2.1, page 40).

Other factors to be considered are the concentration of target protein and the type of catalytic co-factor. Experiments with twice the concentration of lysozyme (280 μM against the 140 μM that we normally used in our work) resulted in trace amounts of dimerization. On the other hand, lysozyme concentrations below 50 μM affect the conversion percentages negatively. This critical concentration is likely dependant on the size of the protein, but was not further investigated. Alternative catalytic co-factors are also possible as various metalloporphyrins exists, some of which show oxidative properties as well.¹³⁻¹⁷ In unpublished work from our group, we attempted to use $Mn(II)$ porphyrin, which shows peroxidase-like features when bound to dsDNA.¹⁵ Although the system showed promising results when oxidizing dopamine at seemingly equal rates as the hGQ DNzyme, it regrettably failed to perform the Tyr-NML conjugation in the presence of either dsDNA or GQ-forming ssDNA.

In chapter 2 we confirmed for proteins what others have found for small substrates,⁷ namely that the DNA sequence plays a major role in the activity and selectivity of the DNAzyme-induced conjugation. However, the underlying cause for the existing correlation between the type of GQ folding and the activity is not precisely known. The GQs used in our experiments all adopt one of the four folding types^{7,18} mentioned in chapters 2 and 3, yet within these categories different subvarieties exist as well.^{18,19} These varieties differ in the type of loop that connects the guanine stacks, which can be a lateral, diagonal or propeller loop (Figure 6.1A). As such, the hybrid GQ can have form-1, 2 or 3 (Figure 6.1B) and the antiparallel GQ can have a chair or a basket conformation (Figure 6.1C),¹⁸ not to mention that loops can also vary in length and nucleotide composition.^{18,20} All of these small distinctions alter properties, including volume of the hGQ complex, electron density and interaction surface of the GQ, all of which ultimately affects the binding and amplification of the hemin co-factor in the hGQ DNAzyme complex. Upon comparison of the intramolecular structures (Figure 6.1) to their respective activities (Table 2.1, page 31 and Figure 3.1C, page 54)) it can be concluded that lateral or diagonal loops decrease DNAzyme activity. Most likely, such loops hinder binding of hemin²¹ or subsequent activation of small molecules by the metalloporphyrin. A direct comparison can be made with the hybrid sequence PS2.M, which has the exact same sequence as its derivative PW17, but has one G on another terminus, giving it one sterically hampering loop more than PW17 and consequently, PS2.M has a lower activity. This steric clash could perhaps also be the underlying cause for the variations seen in the targeted residues on proteins by different GQ topologies (Table 2.2, page 34). It would thus be useful to study the effect of the type and length of the loop on selectivity to assess whether loop design could be used to steer the modification site of hGQ DNAzymes.

Furthermore, DNAzyme activity can be increased by extension with a cytidine-rich⁸ or an adenine-rich sequence.²² It has been shown that such extensions boost DNAzyme activity and it was proposed that the nucleobases in these chains occupy a similar role as the distal His residue in Microperoxidase 11.¹¹ In our observations (chapters 2 and 3) we found that extension of the GQ sequence in some cases indeed resulted in higher conversions, which is likely the result of the described flanking contribution. These contributions could also be a cause for the observed activity variations in GQs with similar topologies.

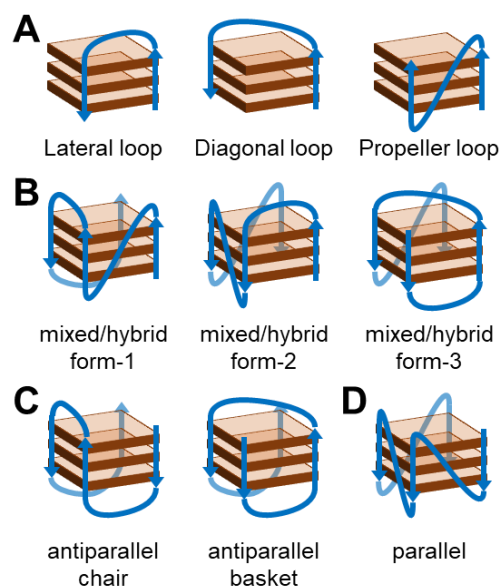


Figure 6.1 Different intramolecular GQ topologies. (A) Nomenclature of GQ loops; (B) subtypes of mixed/hybrid GQ's; (C) subtypes of antiparallel GQ's; (D) parallel.^{18,19}

6.3 Application of TamBA-hGQ for protein modification

Chapter 3 describes how the tyrosinamide-binding aptamer (TamBA) was used to enhance NML conjugation of tyrosinamide (TyrAm). The work distinguishes itself from earlier nucleozyme work²³ by adding a third component to aptamer-DNAzyme-catalysed reactions: the molecule NML that is oxidatively attached to the substrate. Because of this, the methodology shows potential for enhancing the catalytic protein modification described in chapter 2.

As such, we tested the best-performing PW17-TamBA systems for NML conjugation on lysozyme and trastuzumab. Indeed, higher conversions were attained for lysozyme modified with PW17-TamBA and 22-PW17-27 in comparison to PW17 alone and PW17-BaTAm, however trastuzumab modification was unaffected by the presence of TamBA. We hypothesized that this was due to TamBA not being able to recognize the Tyr residues as the features it relies on for binding are not exposed well enough in trastuzumab. To test this, we performed NML conjugation with an in-house mutant Laminarinase A, which had a genetically encoded (Gly)₄Tyr sequence at its C-terminus. This tag was sufficiently exposed to achieve selectivity with mushroom tyrosinase enzyme²⁴ and could be recognized by TamBA as well. Regrettably, the results of NML conjugation with mutant Laminarinase A pointed mainly towards dimerization of the protein, which is interesting but beyond our goals.

Since we focussed on the detection of conversion enhancement, tryptic digestion studies would be recommended to observe possible influence by TamBA on selectivity, as was caused by LBA in chapter 2. Additional tests in substrate selectivity, however, gave tentative results as it was shown that TamBA enhanced the modification of not only various tyrosinamide derivatives, but also of tryptophan. Naturally, in the original selection process of TamBA, it was selected for its strong binding to TyrAm and not for being uniquely selective for just TyrAm.²⁵ As such, it is possible that TamBA also binds phenylalaninamide and histidine, but these appear chemically inert for reaction with NML. If TamBA indeed binds other amino acids residues than Tyr, achieving chemo-selectivity on proteins will be impossible with this aptamer.

6.4 DNA-catalyst nanostructures

Whereas DNA aptamers have already been used for the conjugation of DNA to proteins,²⁶ the work of chapter 4 uses aptamers in a more widely applicable approach by guiding catalytic moieties that attach small molecules. This gives the option of protein modification with all sorts of chemical groups rather than just DNA and controllable activity through the design of the catalytic DNA nanostructure. In addition, the organic catalysts DMAP and PyOx that are used in chapters 4 and 5 only work in close proximity and thus the modification is substrate-specific and even has a degree of site selectivity. Although DMAP performs faster and has a more confined area of modification than PyOx, PyOx presents much lower background modification, higher conversions and works in constructs of both TBA and TBA2. It is thus a trade-off and the properties of either determine which of the two catalysts is preferred. Nevertheless, the background activity of DMAP enabled us to gain insight in how the inherent reactivity of amino acid residues could be overruled by the presence of the catalytic nanostructure. Specifically, we noted that a lysine residue that was modified by free DMAP was not modified by DMAP-functionalized aptamers.

A remarkable difference between the two acyl transfer catalysts is their varying efficiency in the aptamer-based and templated approach. In the aptamer approach of chapter 4, diPyOx acylates more substrate than diDMAP at 300 μ M acyl donor. In the templated approach of chapter 5, however, diDMAP acylates more and can achieve similar acylation levels with just 100 μ M donor as diPyOx with 300 μ M. This could be the result of optimal positioning, since diDMAP is slightly longer and has more rotational freedom. Alternatively, it could be that the dynamic binding of the aptamer presents less interaction time between activated DMAP and targeted residues, whereas in the templated approach, DMAP is always in proximity.

6.5 Points of attention for quantification of protein-DNA conjugates

Going deeper into the importance of catalyst distance and orientation, the work of chapter 5 describes the use of covalent protein-DNA constructs for the calibration of catalysts DMAP, PyOx and hGQ. We correlated the distance between catalyst and protein to efficiency of modification of protein for all catalysts, providing insight for the development of future ligand or aptamer-directed catalytic nanostructures.

It should be noted, however, that quantification of protein-DNA conjugates and their reaction products was difficult using LC-MS. As such, our quantification was performed by integration of Coomassie-stained SDS-PAGE gel results with ImageJ software. This software provides a number from the pixel density of the protein bands, which allows comparison of bands on the same gel to find relative percentages. This method is, however, not very robust and two parameters are highly influential of the results that are obtained: (i) the quality of the photo, and (ii) the staining procedure. To optimize photo quality lighting was kept constant for the gel by scanning gels that were sandwiched between two plastic transparent sheets. For every gel image, new sheets were taken to avoid smudges or salts that might cause pixel variations. The exact same protocol was used for staining and destaining of the gels, and shaking the gel during (de)staining was applied to evenly colour the gel.

When calculating conversions of the protein to protein-DNA conjugates, it should be taking into account that Coomassie Blue staining is less efficient on protein-DNA conjugates. In fact, attachment of a DNA strand to a protein can decrease Coomassie Blue staining of the protein-DNA conjugate by approximately 30%, thus lessening the intensity of the bands on SDS-PAGE. This effect has been described in literature,^{27,28} and results in a decrease in conversion when calculating conversions with our methodology, because we directly correlate intensity to the concentration of product. As such, our quantification of protein-DNA conjugates is likely suboptimal due to a decreased band intensity, although this also suggests that our conjugation strategy is actually more efficient than we reported.

6.6 Future prospects

6.6.1 Protein modification with catalytic DNA

Studying the influence of the DNA sequence has in our search been limited to just four types of GQ forming constructs. However, as mentioned in section 6.3, subvarieties and alternative loops are possible. To truly assess the scope of the influence, hGQ DNAzymes based on said subvarieties should be tested, as well as artificially designed GQ's with variations in the length and nucleobase content of the loops. Boosting the activity by flanking residues can be included in these sequences to design the most efficient and perhaps most selective DNAzymes for NML conjugation. Perhaps peptide nucleic acids (PNAs) (Figure 6.2) could be considered to create hGQ hybrid PNA-DNAzymes^{29,30} with peptide extensions that subtly interact with the protein of interest. Alternatively, the selectivity of the DNAzymes could be tailored using protein-affinity groups. Just as simple sequence extension was used to include a lysozyme binding aptamer (LBA) in the hGQ DNAzyme, in PNA-based DNAzymes the inclusion of protein-binding peptides can be done by mere sequence extension. Such peptide sequences could potentially increase selectivity of the PNA-peptide hybrids.³¹

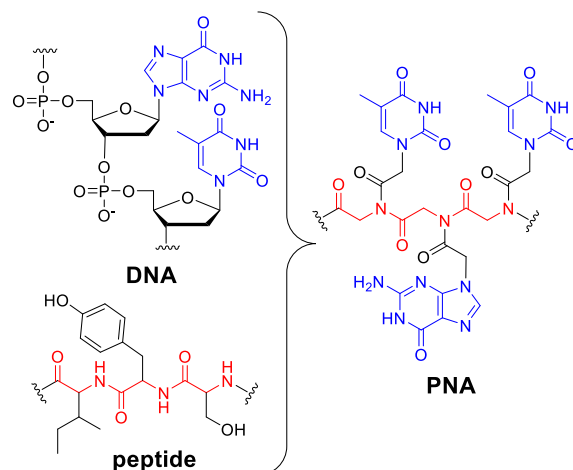


Figure 6.2. PNA is a combination of DNA and peptides where DNA nucleobases (blue) are attached to the amide groups of a peptide backbone (red).

In addition to Tyr-NML conjugation, the hGQ DNAzyme can also perform other reactions. In one example a carbene reaction between styrene and ethyldiazoacetate^{32,33} was effective on peptides and small proteins. Specifically, after installation of a styrene group on Leu-enkephalin or GRX with 4-chloromethyl-styrene (Figure 6.3A), the hGQ DNAzyme could in the presence of sodium dithionite ($\text{Na}_2\text{S}_2\text{O}_4$), conjugate ethyldiazoacetate (EDA) to styrene more efficiently than hemin alone. Formation of the products was confirmed using LC-MS (Figure 6.3B–D). Although this reaction is a novel DNA-catalysed bioconjugation that is orthogonal to native amino acid residues, it needs a strong reducing agent and an extra step to introduce the styrene group, both of which make this reaction inferior to the NML conjugation.

As mentioned in section 6.2, the use of other catalytic co-factors is also an option. Metalloporphyrins with other transition metal cores than Fe can also bind GQ, possibly generating different properties. For example the Mn- and Ru-based porphyrins have different oxidation potentials than hemin and the Zn-based porphyrin is activated by light rather than hydrogen peroxide.^{13–17} Preliminary data from our lab already indicated that a Zn-porphyrin DNAzyme is able to perform the Tyr-NML conjugation with peptides and proteins when activated by 400 nm light. Following this data, together with the known capacity of GQ DNA to bind and stabilize metalloporphyrins, we anticipate more protein modifying DNAzymes to emerge in the near future.

Next to the porphyrin-based DNAzymes, other catalytic DNA structures have been developed to modify protein-like substrates. Among these, there are DNA sequences that hydrolyse peptide bonds,³⁴ modify the ϵ -amine of Lys residues,³⁵ perform reductive amination,³⁶ even dephosphorylate³⁷ and phosphorylate Tyr^{38,39} or Ser⁴⁰ residues. With such a varied toolbox of potential reactions, we feel that our work on protein modification with catalytic DNA, is only the beginning of DNA-assisted protein modification.

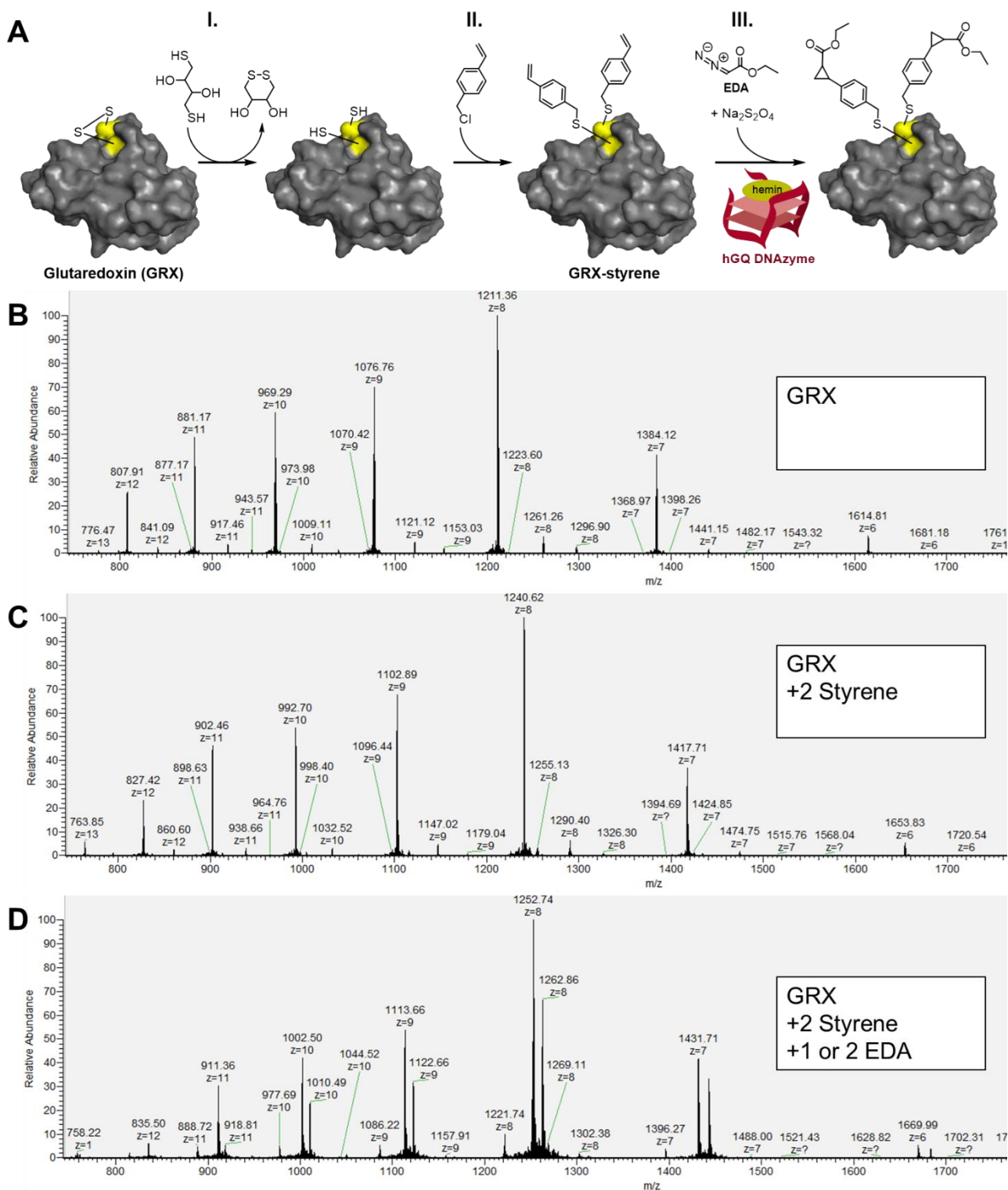


Figure 6.3. (A) Bioconjugation of GRX with styrene and EDA in three steps: (I) DTT-mediated opening of the disulfide bridge, (II) insertion of 4-chloromethylstyrene to install two styrene groups, and (III) carbene reaction of GRX-styrene with EDA performed by the hGQ DNAzyme. (B) Mass spectrometry data corresponding to native GRX, calculated peaks: mass(*z*) = 807.9(12), 881.3(11), 969.3(10), 1076.9(9), 1211.4(8), 1384.3(7), 1614.8(6). (C) Mass spectrometry data corresponding to GRX with two styrene groups on either, calculated peaks: mass(*z*) = 827.3(12), 902.4(11), 992.5(10), 1102.7(9), 1240.4(8), 1417.4(7), 1653.5(6). (D) Mass spectrometry data corresponding to GRX with styrene coupled with one or two ethyl diazoacetate (EDA) molecules, calculated values (1xEDA): mass(*z*) = 835.5(12), 911.4(11), 1002.4(10), 1113.7(9), 1252.8(8), 1431.6(7), 1670.0(6) and calculated values (2xEDA): mass(*z*) = 841.8(12), 918.2 (11), 1010.1(10), 1122.0(9), 1262.2(8), 1442.3(7), 1682.5(6).

6.6.2 DNA aptamers for protein modification

Aptamers can be effective protein-binding elements that display high affinity and specificity for their target. By harnessing these features, we demonstrated that they could be applied for catalytic protein modification in a substrate-specific, site-selective and controllable fashion. In chapter 2, inclusion of a protein-binding aptamer could be used to change the selectivity of the conjugation performed by the hGQ DNAzyme, in chapter 3 we used it to enhance the rate of conjugation with small peptides and in chapter 4 both of these concepts were translated to protein modification with the acyl transfer catalysts. Research with SELEX still leads to regular discovery of new protein binding aptamers,⁴¹⁻⁴⁴ which can be used in these aptamer-based approaches. Aptamers have already been demonstrated for clinical use,⁴⁵ *in vivo* targeting,^{44,46} and even the binding of live cells.⁴⁷ Therefore it is not unrealistic to assume that research regarding aptamer-based *in vivo* protein labelling and cell surface labelling will lead to many more protein modification strategies.

In our work, we used information from the crystal structures of TBA and TBA2 to design the catalytic DNA nanostructures. Unfortunately, only a few aptamer-protein complexes have resolved crystal structures. To circumvent the necessity for crystal structures of protein-binding aptamers, a catalytic DNA construct could be hybridized to template strands carrying peptides,^{48,49} metals^{28,50} and small molecules^{51,52} as anchors that guide the catalytic construct to the protein. Positioning of the catalyst with respect to these protein anchors is crucial,^{28,53-55} which was also visible in our findings in chapter 4. Considering the impact of spatial orientation, the effect of linkers should be studied further. Ideally, a series of linker lengths is tested with the same catalytic moiety to analyse the linker's influence. As a limited set is commercially available,⁵⁶ self-made oligonucleotides might be required. Insight and availability of different anchors and linkers is a prerequisite for future widespread application of DNA constructs.

6.6.3 The protein-DNA conjugate system

The protein-DNA conjugates synthesized in chapter 5 have the potential for many future applications. One example is to perform research into alternate catalysts as we only examined three different catalysts in chapter 5, two of which are acyl transfer catalysts. Given the sharp increase of the activity of DMAP in the templated DNA approach, it would be interesting to investigate other proximity catalysts as well, for example the Ru(bipy)₃ photocatalyst^{57,58} or the Rh₂(OAc)₄ metallocatalyst^{59,60} (section 1.3.3 page 11-12). In addition to distance calibration, such constructs could lead to DNA-based modification strategies of other types of amino acid residues.

Another example is the design and development of more complex DNA nanostructures. A design could be envisioned where a hairpin loop is hybridized to the template strand, pointing out horizontally over the protein surface. Catalysts can covalently be tethered to the loop region of the hairpin or be hybridized to it by matching strands. Another possibility is the use of a hairpin template strand with both ends close together in ssDNA form. If a proximity catalyst is placed on the opposite end of the protein-bound DNA strand, the system would be catalytically active in its hairpin form.⁶¹ Addition of a complementary strand results in a dsDNA duplex that will inactivate the system by moving the catalyst away from the protein surface. In line with these examples, many more designs are possible, underscoring the potential of DNA nanotechnology for protein modification.

6.7 DNA as tool for artificial enzymes

Enzymes are perhaps the most efficient catalysts for site-selective protein modifying known as they combine several crucial features. (i) Catalytic modification, which allows an enzyme to convert large quantities of substrate molecules while itself being in low quantity.^{62–64} The efficiency differs per catalyst, but most enzymes have far greater protein modifying efficiency than their artificial counterparts.^{65–67} They often possess (ii) substrate selectivity, which is the result of a binding site that has a unique affinity for a particular molecular motive, giving them the ability to distinguish their target amongst a large numbers of other molecules. This binding site also often results in (iii) site selectivity, because the substrate can often fit the binding site in just one way, whereby only one or a few residues are positioned near the reactive centre.^{62–64} (iv) The activity of enzymes can be regulated to increase or decrease their output in accordance with the needs of the cell, which often happens through post-translational modifications of proteins.^{62,63}

As such, mimicking enzymes is a hot research field^{65,68} and examples of enzyme mimicry include: the transamination catalyst pyridoxalphosphate⁶⁹ and the ligand-directed Ru(bipy)₃⁵⁷ and DMAP.⁷⁰ Regrettably, all these examples lack one of the four features mentioned above, in contrast to the DNA-based systems described in this thesis, which meet all four characteristics. In chapter 4, the aptamer part of the DNA-catalyst constructs gives both site-selectivity and substrate-specificity with the attached DMAP or PyOx performing the catalysis and the single strand nature of the aptamer allowing for activity control with a complementary strand. In chapter 2, the same activity control switch is used, but the catalysis and selectivity are now performed by the hGQ DNAzyme and although we did not test for substrate specificity, signs of this feature have been shown by Masuzawa *et al.*⁷¹ and could otherwise be achieved by using DNA aptamers.

Synthetic DNA is a great material for enzyme mimicry, as many of the functions associated to enzyme activity can be programmed into DNA nanostructures. In addition, many other DNA-catalysed reactions have already been described^{72,73} as well as new functions and geometries that can be programmed into DNA nanostructures.^{74–78} On top of this, artificial nucleobase pairs are being designed to allow for wider set of programmable DNA nucleobase pairs.^{79–81} This expanding toolbox of DNA-based functions and chemistries will pave the way for more efficient and elaborate protein modification strategies.

In conclusion, this thesis describes the development of new protein modification strategies using DNA as nanometre-sized constructs in a catalytic, protein-binding and/or templated fashion. The DNA systems were designed to site-selectively modify wild-type proteins in a substrate-specific manner, generating high conversions when doing so and to contain a control switch for their activity. This proves that DNA is not only a functional tool for catalytic protein modification, but also has the potential for enzyme mimicry.

References

- 1 H. Weingärtner, C. Cabrele and C. Herrmann, *Phys. Chem. Chem. Phys.* **2012**, 14 (2), 415–426.
- 2 H. I. Okur, J. Hladílková, K. B. Rembert, Y. Cho, J. Heyda, J. Dzubiella, P. S. Cremer and P. Jungwirth, *J. Phys. Chem. B* **2017**, 121 (9), 1997–2014.
- 3 D. Usoltsev, V. Sitnikova, A. Kajava and M. Uspenskaya, *Biomolecules* **2020**, 10 (606), 1–17.
- 4 X. Wang, M. Ni, C. Niu, X. Zhu, T. Zhao, Z. Zhu, Y. Xuan and W. Cong, *EuPA Open Proteomics* **2014**, 4, 156–164.
- 5 M. R. Green and J. Sambrook, *Cold Spring Harb. Protoc.* **2019**, 2019 (1), 6–15.
- 6 J. J. Clayden, N. Greeves, S. Warren and P. Wothers, *Organic Chemistry*, **2001**, vol. 40.
- 7 X. Cheng, X. Liu, T. Bing, Z. Cao and D. Shangguan, *Biochemistry* **2009**, 48 (33), 7817–7823.
- 8 T. Chang, H. Gong, P. Ding, X. Liu, W. Li, T. Bing, Z. Cao and D. Shangguan, *Chem. - A Eur. J.* **2016**, 22 (12), 4015–4021.
- 9 T. Li, S. Dong and E. Wang, *J. Am. Chem. Soc.* **2010**, 132 (38), 13156–13157.
- 10 L. Stefan, F. Denat and D. Monchaud, *Nucleic Acids Res.* **2012**, 40 (17), 8759–8772.
- 11 A. Spector, W. Zhou, W. Ma, C. F. Chignell and K. J. Reszka, *Exp. Eye Res.* **2000**, 71 (2), 183–194.
- 12 J. Kosman and B. Juskowiak, *Cent. Eur. J. Chem.* **2012**, 10 (2), 368–372.
- 13 M. M. Pereira, L. D. Dias and M. J. F. Calvete, *ACS Catal.* **2018**, 10784–10808.
- 14 M. Costas, *Coord. Chem. Rev.* **2011**, 255 (23–24), 2912–2932.
- 15 J. Xu, J. Wu, C. Zong, H. Ju and F. Yan, *Anal. Chem.* **2013**, 85 (6), 3374–3379.
- 16 Z. Zhang, E. Sharon, R. Freeman, X. Liu and I. Willner, *Anal. Chem.* **2012**, 84 (11), 4789–4797.
- 17 T. S. Metzger, R. Tel-Vered, H. B. Albada and I. Willner, *Adv. Funct. Mater.* **2015**, 25 (41), 6470–6477.
- 18 Y. Ma, K. Iida and K. Nagasawa, *Biochem. Biophys. Res. Commun.* **2020**, 531 (1), 3–17.
- 19 M. A. Keniry, *Biopolymers* **2000**, 56 (3), 123–146.
- 20 A. Guédin, J. Gros, P. Alberti and J. L. Mergny, *Nucleic Acids Res.* **2010**, 38 (21), 7858–7868.
- 21 J. M. Nicoludis, S. T. Miller, P. D. Jeffrey, S. P. Barrett, P. R. Rablen, T. J. Lawton and L. A. Yatsunyk, *J. Am. Chem. Soc.* **2012**, 134 (50), 20446–20456.
- 22 W. Li, Y. Li, Z. Liu, B. Lin, H. Yi, F. Xu, Z. Nie and S. Yao, *Nucleic Acids Res.* **2016**, 44 (15), 7373–7384.
- 23 H. B. Albada, E. Golub and I. Willner, *Chem. Sci.* **2016**, 7 (5), 3092–3101.
- 24 J. J. Bruins, A. H. Westphal, B. Albada, K. Wagner, L. Bartels, H. Spits, W. J. H. Van Berkel and F. L. Van Delft, *Bioconjugate Chem.* **2017**, 28 (4), 1189–1193.
- 25 E. Vianini, M. Palumbo and B. Gatto, *Bioorganic Med. Chem.* **2001**, 9 (10), 2543–2548.
- 26 J. F. Keijzer and B. Albada, *Biopolymers* **2021**, 113 (3), 1–8.
- 27 M. Synakewicz, D. Bauer, M. Rief and L. S. Itzhaki, *Sci. Rep.* **2019**, 9 (1), 1–10.
- 28 C. B. Rosen, A. L. B. Kodal, J. S. Nielsen, D. H. Schaffert, C. Scavenius, A. H. Okholm, N. V. Voigt, J. J. Enghild, J. Kjems, T. Tørring and K. V. Gothelf, *Nat. Chem.* **2014**, 6 (9), 804–809.

- 29 B. Datta, C. Schmitt and B. A. Armitage, *J. Am. Chem. Soc.* **2003**, *125* (14), 4111–4118.
- 30 Y. Krishnan-ghosh, E. Stephens and S. Balasubramanian, *J. Am. Chem. Soc.* **2004**, *126* (19), 5944–5945.
- 31 N. Stephanopoulos, *Bioconjugate Chem.* **2019**, *30* (7), 1915–1922.
- 32 H. Ibrahim, P. Mulyk and D. Sen, *ACS Omega* **2019**, *4* (12), 15280–15288.
- 33 A. Rioz-Martínez, J. Oelerich, N. Ségaud and G. Roelfes, *Angew. Chem. Int. Ed.* **2016**, *55* (45), 14136–14140.
- 34 B. M. Brandsen, A. R. Hesser, M. A. Castner, M. Chandra and S. K. Silverman, *J. Am. Chem. Soc.* **2013**, *135* (43), 16014–16017.
- 35 B. M. Brandsen, T. E. Velez, A. Sachdeva, N. A. Ibrahim and S. K. Silverman, *Angew. Chem. Int. Ed.* **2014**, *53* (34), 9045–9050.
- 36 O. Y. Wong, A. E. Mulcrone and S. K. Silverman, *Angew. Chem. Int. Ed.* **2011**, *123* (49), 11883–11888.
- 37 J. Chandrasekar and S. K. Silverman, *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (14), 5315–5320.
- 38 V. Dokukin and S. K. Silverman, *Chem. Commun.* **2014**, *50* (66), 9317–9320.
- 39 S. M. Walsh, A. Sachdeva and S. K. Silverman, *J. Am. Chem. Soc.* **2013**, *135* (40), 14928–14931.
- 40 A. Sachdeva and S. K. Silverman, *Chem. Commun.* **2010**, *46* (13), 2215–2217.
- 41 A. Schmitz, A. Weber, M. Bayin, S. Breuers, V. Fieberg, M. Famulok and G. Mayer, *Angew. Chem. Int. Ed.* **2021**, *60* (18), 10279–10285.
- 42 M. Liu, J. Wang, Y. Chang, Q. Zhang, D. Chang, C. Y. Hui, J. D. Brennan and Y. Li, *Angew. Chem. Int. Ed.* **2020**, *59* (20), 7706–7710.
- 43 N. Kuntip, D. Japrun and P. Pongprayoon, *Biopolymers* **2021**, *112* (3), 1–9.
- 44 X. Huang, J. Zhong, J. Ren, D. Wen, W. Zhao and Y. Huan, *Oncol. Lett.* **2019**, *18* (1), 265–274.
- 45 E. W. M. Ng, D. T. Shima, P. Calias, E. T. Cunningham, D. R. Guyer and A. P. Adamis, *Nat. Rev. Drug Discov.* **2006**, *5*, 123–132.
- 46 E. M. McConnell, K. Ventura, Z. Dwyer, V. Hunt, A. Koudrina, M. R. Holahan and M. C. Derosa, *ACS Chem. Neurosci.* **2019**, *10* (1), 371–383.
- 47 X. Xiong, H. Liu, Z. Zhao, M. B. Altman, D. Lopez-Colon, C. J. Yang, L.-J. Chang, C. Liu and W. Tan, *Angew. Chem. Int. Ed.* **2013**, *125* (5), 1512–1516.
- 48 X. Bai, C. Lu, J. Jin, S. Tian, Z. Guo, P. Chen, G. Zhai, S. Zheng, X. He, E. Fan, Y. Zhang and K. Zhang, *Angew. Chem. Int. Ed.* **2016**, *128* (28), 8125–8129.
- 49 T. B. Nielsen, R. P. Thomsen, M. R. Mortensen, J. Kjems, P. F. Nielsen, T. E. Nielsen, A. L. B. Kodal, E. Cló and K. V. Gothelf, *Angew. Chem. Int. Ed.* **2019**, *58* (27), 9068–9072.
- 50 A. L. B. Kodal, C. B. Rosen, M. R. Mortensen, T. Tørring and K. V. Gothelf, *ChemBioChem* **2016**, 1338–1342.
- 51 G. Li, Y. Liu, Y. Liu, L. Chen, S. Wu, Y. Liu and X. Li, *Angew. Chem. Int. Ed.* **2013**, *125* (36), 9723–9728.
- 52 P. Zhao, Z. Chen, Y. Li, D. Sun, Y. Gao, Y. Huang and X. Li, *Angew. Chem. Int. Ed.* **2014**, *53* (38), 10056–10059.
- 53 S. E. Lee, A. Sidorov, T. Gourlain, N. Mignet, S. J. Thorpe, J. A. Brazier, M. J. Dickman, D. P. Hornby, J. A. Grasby and D. M. Williams, *Nucleic Acids Res.* **2001**, *29* (7), 1565–1573.

- 54 D. F. Sauer, M. Bocola, C. Broglia, M. Arlt, L. L. Zhu, M. Brocker, U. Schwaneberg and J. Okuda, *Chem. - An Asian J.* **2015**, *10* (1), 177–182.
- 55 R. Wang, D. Lu, H. Bai, C. Jin, G. Yan, M. Ye, L. Qiu, R. Chang, C. Cui, H. Liang and W. Tan, *Chem. Sci.* **2016**, *7* (3), 2157–2161.
- 56 N. Z. Fantoni, A. H. El-Sagheer and T. Brown, *Chem. Rev.* **2021**, *121* (12), 7122–7154.
- 57 S. Sato and H. Nakamura, *Angew. Chem. Int. Ed.* **2013**, *125* (33), 8843–8846.
- 58 S. Sato, K. Hatano, M. Tsushima and H. Nakamura, *Chem. Commun.* **2018**, *54* (46), 5871–5874.
- 59 F. Vohidov, J. M. Coughlin and Z. T. Ball, *Angew. Chem. Int. Ed.* **2015**, *127* (15), 4670–4674.
- 60 J. Ohata and Z. T. Ball, *J. Am. Chem. Soc.* **2017**, *139* (36), 12617–12622.
- 61 E. B. Pimentel, T. M. Peters-Clarke, J. J. Coon and J. D. Martell, *J. Am. Chem. Soc.* **2021**, *143* (50), 21402–21409.
- 62 B. Alberts, A. Johnson, J. Lewis, D. Morgan, M. Raff, K. Roberts and P. Walter, *The Cell*, **2014**, vol. 6.
- 63 A. L. Lehninger, D. L. Nelson and M. M. Cox, *Lehninger Principles of Biochemistry*, W.H. Freeman & Co Ltd, fifth edit., **2008**.
- 64 P. Intasian, K. Prakinee, A. Phintha, D. Trisrivirat, N. Weeranoppanant, T. Wongnate and P. Chaiyen, *Chem. Rev.* **2021**, *121* (17), 10367–10451.
- 65 N. C. Reddy, M. Kumar, R. Molla and V. Rai, *Org. Biomol. Chem.* **2020**, *18* (25), 4669–4691.
- 66 B. Albada, J. F. Keijzer, H. Zuilhof and F. van Delft, *Chem. Rev.* **2021**, *121* (12), 7032–7058.
- 67 Y. Zhang, K. Y. Park, K. F. Suazo and M. D. Distefano, *Chem. Soc. Rev.* **2018**, *47* (24), 9106–9136.
- 68 S. Sato and H. Nakamura, *Molecules* **2019**, *24* (21), 1–17.
- 69 J. M. Gilmore, R. A. Scheck, A. P. Esser-Kahn, N. S. Joshi and M. B. Francis, *Angew. Chem. Int. Ed.* **2006**, *45* (32), 5307–5311.
- 70 Y. Koshi, E. Nakata, M. Miyagawa, S. Tsukiji, T. Ogawa and I. Hamachi, *J. Am. Chem. Soc.* **2008**, *130* (1), 245–251.
- 71 T. Masuzawa, S. Sato, T. Niwa, H. Taguchi, H. Nakamura and T. Oyoshi, *Chem. Commun.* **2020**, *56* (78), 11641–11644.
- 72 A. J. Boersma, R. P. Megens, B. L. Feringa and G. Roelfes, *Chem. Soc. Rev.* **2010**, *39* (6), 2083–2092.
- 73 S. K. Silverman, *Acc. Chem. Res.* **2015**, *48* (5), 1369–1379.
- 74 O. I. Wilner and I. Willner, *Chem. Rev.* **2012**, *112* (4), 2528–2556.
- 75 M. Madsen and K. V. Gothelf, *Chem. Rev.* **2019**, *119* (10), 6384–6458.
- 76 S. J. W. Henry and N. Stephanopoulos, *Wiley Interdiscip. Rev. Nanomedicine Nanobiotechnology* **2021**, *13* (6), 1–21.
- 77 Y. J. Chen, B. Groves, R. A. Muscat and G. Seelig, *Nat. Nanotechnol.* **2015**, *10* (9), 748–760.
- 78 P. Chidchob and H. F. Sleiman, *Curr. Opin. Chem. Biol.* **2018**, *46*, 63–70.
- 79 A. Marx and K. Betz, *Chem. - A Eur. J.* **2020**, *26* (16), 3446–3463.
- 80 E. T. Kool, *Acc. Chem. Res.* **2002**, *35* (11), 936–943.
- 81 I. Hirao, M. Kimoto and R. Yamashige, *Acc. Chem. Res.* **2012**, *45* (12), 2055–2065.

Chapter 7



Summary

Summary

Proteins are biological macromolecules that consist of 20 different types of building blocks, called amino acids. Some of these amino acids have different reactive groups on their side chains, such as amines, alcohols or thiols, each with different reactivities. The field of artificial protein modification aims at the (bio)chemical alteration of proteins to functionalize them, for example for analytical or drug delivery purposes. If this is done in a chemo-selective manner, one specific type of amino acid is modified. Some existing reagents for chemo-selective modification are the maleimide group for thiol modification on Cys residues and the *N*-hydroxy-succinimide ester for amine modification on Lys residues. Naturally, multiple amino acids of the same type can exist within a single protein. The term site selectivity is used for strategies that target only one or a few residues among various others of the same type. This can be obtained by genetic incorporation of unnatural amino acids that have unique side chains that are the only ones to react with the used chemical reagent, for example an azide that reacts specifically with an alkyne. However, genetic alteration of proteins is time-consuming and positive outcomes are not always guaranteed. As such, many strategies for site-selective modification of native proteins have been devised over the years.

Apart from the application of chemical reagents for protein modification, catalysts can also be used to confer selectivity. This can be done with naturally occurring enzymes, such as Sortase A that substitutes peptides for amine-containing moieties, or mushroom tyrosinase that oxidatively activates Tyr residues. Alternatively, artificial catalysts such as hemin or $\text{Rh}_2(\text{OAc})_4$ can be used for the oxidative conjugation of NML and Tyr residues or the alkylation of Trp and His residues, respectively. To confer selectivity, artificial catalysts are often armed with protein-binding ligands. Examples are sugar-tethered DMAP for the modification of lectins, and benzenesulfonamide $\text{Ru}(\text{bipy})_3$ for the modification of the periphery of carbonic anhydrase.

Protein modification has also been made selective by using DNA in various ways. In the templated approach, a DNA strand was functionalized with a protein-binding moiety and, after incubation with a protein, used as a template to hybridize a second reactive DNA strand. Via the template strand, this second strand was brought in close proximity to the protein and can site-selective react with the protein. In the ligand-directed approach, DNA-binding proteins were modified using DNA strands carrying photo-reactive warheads that upon irradiation, covalently bind the interacting protein. A third approach is the use of DNA aptamers, which are oligonucleotide sequences that bind a specific target such as a protein. These aptamers have been armed with reactive warheads for self-conjugation to their protein target or by using them in a DNA-templated approach, to conjugate other DNA strands.

Although these techniques are effective, they have only been used to prepare protein-DNA conjugates. However, DNA has the potential to modify proteins with other functionalities, as catalytic DNA strands exist that can modify chemical groups that also occur in proteins. In addition, the field of DNA nanotechnology has demonstrated that the programmable nature of DNA allows for additional functions and forms through the design of the oligonucleotide nanostructures, ranging from conformational switches to nano-sized boxes that can open and close. In this thesis, we aim to use DNA for the modification of native proteins with small molecules, as well as to design the used DNA systems to include extra functions, such as substrate-binding sites and activity switches.

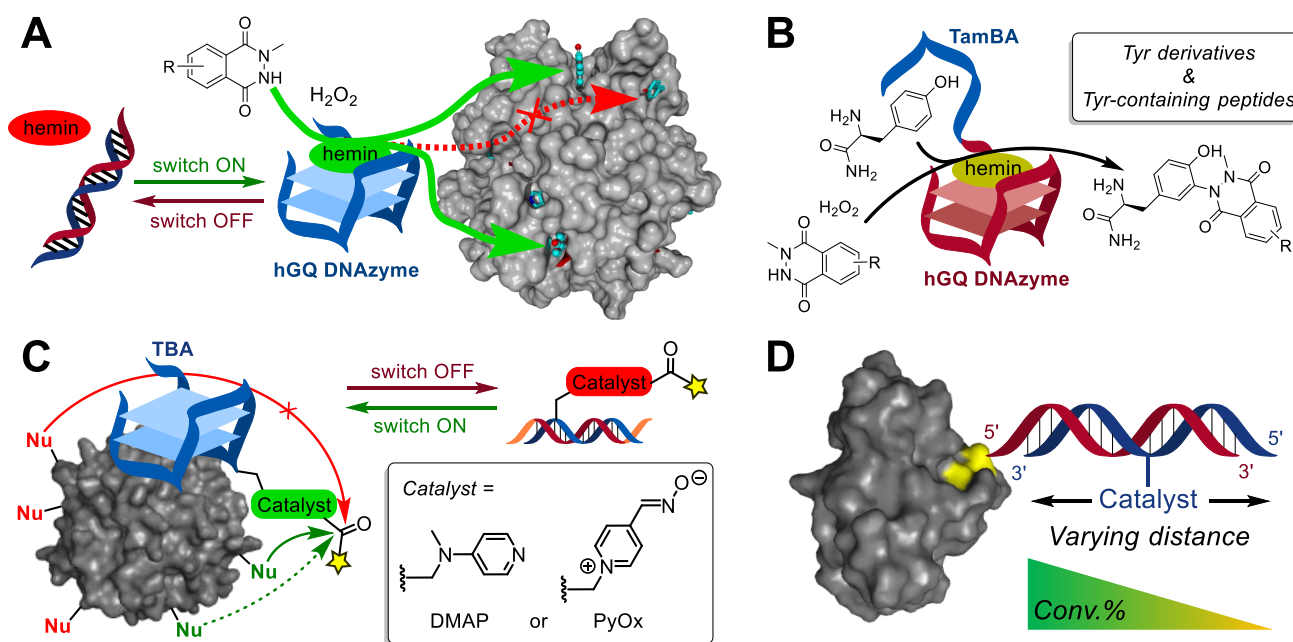


Figure 7.1. Overview of DNA-assisted modification strategies from this thesis. **(A)** The hGQ DNAzymes conjugate NML to Tyr residues on protein surfaces. They could also be programmed with an activity control switch; **(B)** Extension of an hGQ DNAzyme with a tyrosine-binding aptamer (TamBA) enhanced its conversion levels on Tyr derivatives and small peptides; **(C)** Acyl transfer catalysts DMAP and PyOx were covalently attached to TBA (and TBA2) and could catalyze the acylation of thrombin. They could also be programmed with an activity control switch; **(D)** The correlation between spatial orientation and conversion of DMAP, PyOx and hGQ were analysed by tethering them to a template strand and hybridizing these strands to protein-DNA conjugates.

Chapter 2 describes protein modification using DNA in a catalytic fashion (Figure 7.1A). The horseradish peroxidase-mimicking hemin/G-quadruplex (hGQ) DNAzyme conjugates *N*-methyl luminol (NML) derivatives to Tyr and Trp residues in the presence of H_2O_2 . By testing on the model protein lysozyme C (14.3 kDa), we examined the activities of 22 different hGQ DNAzymes that adopt different GQ topologies. We found that the DNAzymes presented varying activities, even among DNAzymes of the same topological group, where the intramolecular parallel GQs were generally the most active (conversions of 54–96%), followed by the intramolecular hybrid (60–72%), then the intermolecular (20–86%) and lastly the intramolecular antiparallel GQs (32–44%). As the reaction requires H_2O_2 , it could be abruptly stopped by adding a trace amount of the enzyme catalase (<0.1 mg/mL) and time-resolved kinetics could be studied. These studies showed that the NML conjugation proceeds very fast, achieving 80% conversion within 5 minutes for the most active DNAzymes and reaching completion after just 15–30 minutes. Investigation of the influence of the reactants led to the conclusion that higher concentrations of hGQ generally led to sharp increases in conversion and that lower concentrations of NML could be compensated by higher concentrations of hydrogen peroxide to obtain similar conversions, and *vice versa*. With this knowledge we performed DNAzyme-induced NML conjugation on three larger proteins, *i.e.*, thrombin (36 kDa), bovine serum albumin (66 kDa) and the antibody trastuzumab (150 kDa). For all proteins, conjugation was successful and the DNAzyme topologies showed a similar activity hierarchy. To analyse the selectivity of the reaction, tryptic digestion and follow-up MS/MS of modified lysozyme and thrombin was performed, revealing that different GQ topologies target different Tyr (and to a lesser extent Trp) residues. Interestingly, by attaching a lysozyme-binding aptamer (LBA) to the hybrid hGQ, the selectivity on the lysozyme could be shifted to target an additional Tyr residue. Lastly, we designed one of the hybrid GQ DNAzymes to contain a regulatory element with which the conformation of the DNAzyme could be switched between GQ (ON) and dsDNA duplex (OFF), thereby controlling the activity of the DNAzyme.

In chapter 3, we further investigated the potential of combining the hGQ DNAzyme with an aptamer (Figure 7.1B). For this, we extended the hybrid-type GQ sequence PW17 with tyrosinamide-binding

aptamer (TamBA) to form so-called nucleopzymes. We analysed the catalytic efficiency of these nucleopzymes in the conjugation reaction of NML with tyrosinamide and found that PW17-TamBA gave a 12-fold increase in the catalytic rate constant (k_{cat}) compared to regular PW17. Tests with a scrambled TamBA sequence and studies into the saturation kinetics provided evidence for the involvement of the substrate binding site in this reaction, demonstrating the increased activity was the result of TamBA and not mere sequence extension. We then applied the best-performing nucleopzymes for the NML conjugation of two Tyr-containing peptides and found again a TamBA-induced increase in NML conjugation, leading us to conclude that the aptamer also recognizes Tyr as part of a larger peptide structure. Additionally, we found different enhancement rates for both peptides, leading us to believe that distant residues in a peptide substrate can be of influence on the generated conversion.

In chapter 4, we set out to modify proteins with DNA aptamers as protein-affinity elements to guide covalently attached catalysts (Figure 7.1C). Specifically, we conjugated the acyl-transfer catalyst dimethylaminopyridine (DMAP) to different positions of one of two thrombin-binding aptamers (TBA and TBA2) and analysed the catalytic efficiency of these DNA-DMAP constructs on the acylation of Lys residues on thrombin with thioesters. We found that the position of attachment for DMAP is crucial for its activity as only the two constructs with DMAP closest to the protein-aptamer interaction surface (TBA³-DMAP and TBA¹²-DMAP) showed catalytic activity and none of the TBA2-DMAP constructs. Final conversions were obtained after 2 h and reached values of around 40% and 60% for DMAP and divalent DMAP (diDMAP), respectively. Remarkably, tryptic digestion and follow-up MS/MS of acylated thrombin revealed that TBA³-DMAP and TBA¹²-DMAP site-selectively target different Lys residues. To reduce the high background activity of the thioesters (8–12%), we implemented the acyl-transfer catalyst pyridinecarbaldehyde oxime (PyOx) that is more nucleophilic and accepts less reactive alkyl-*N*-acyl-*N*-sulfonamides (ANANS). Even though final conversions were only obtained after 8 h, PyOx showed much less background (<2%) whilst producing similar or higher amounts of modified protein (around 30% and 80% conversion for PyOx and divalent PyOx (diPyOx), respectively). Interestingly, the TBA2-diPyOx constructs were also catalytically active (20–27% conversion) and follow-up tryptic digestion and MS/MS revealed three important findings: (i) TBA-diPyOx and TBA2-diPyOx acylate on opposite sides of thrombin in accordance with the binding sites of the aptamers; (ii) the acylation site of TBA2-diPyOx was also dependant on the position of the catalyst, the acylation site of TBA-diPyOx was not; (iii) PyOx acylates not only Lys but also Ser residues. Despite these differences, both DMAP and PyOx performed modification exclusively on thrombin in the presence of other proteins and the conformation of both TBA-catalyst constructs could be switched between ssDNA (ON) and dsDNA duplex (OFF), thereby controlling their activity.

In chapter 5, we sought to analyse the effective distance of the three catalysts used in the previous chapters (Figure 7.1D). For this, we synthesized three different covalently bound protein-DNA conjugates using two approaches. In the first approach we modified the single thiol bridge of glutaredoxin 1 (9.5 kDa) by reduction, followed by modification with 1-azidomethyl-3,5-bis(bromomethyl)benzene and finally CuAAC with alkyne-DNA. The second approach was the modification of serine proteases chymotrypsin (26 kDa) and α -thrombin (36 kDa) with an azide-functionalized paraoxon derivative that reacts exclusively with the serine residue in their active site and subsequent CuAAC with alkyne-DNA. The strand of the different protein-DNA conjugates was used as a template strand to which we could hybridize a library of complementary DNA strands that were functionalized with catalysts at different positions. As such, we could assess the optimal and maximum range at which each catalyst (diDMAP, diPyOx, hGQ) performs protein modification. The catalysts diDMAP and diPyOx that both have a covalently bound reactant intermediates, demonstrated a strong correlation between their distance to the protein surface and their efficiency. Follow-up tryptic digestion and MS/MS revealed that the distance between these catalysts and the protein surface also affects their site-selectivity, which suggested that conversion could be reduced to attain a higher level

of site selectivity. We attempted to rationalize the selectivity data using computational models and were able to explain that the modification pattern found on thrombin matches the reach of the dsDNA bound catalysts. The third catalyst we tested, the hGQ DNAzyme, uses unbound radical reactants with a limited life-span that dissociate to their target. Its conversion was unaffected at ranges up to 7 nm, which was the maximal length of our linker and therefore no further digestion analyses were performed.

Chapter 6 features additional remarks and discussions regarding the work of this thesis as well as recommendations for future research. The analysis of protein and DNA mixtures can be cumbersome due to the opposite charges of the overall positively charged protein and negatively charged DNA. To tackle this issue, high ionic strength is required to avoid precipitation and reaction analyses require DNA digestion or separation of protein and DNA. We also pose additional insights for the DNAzyme-catalysed NML-Tyr conjugation, including the importance of pH, buffer agents and cations, as well as the effect of the nucleobases and length of the loops in the four GQ topologies. Preliminary research into the application of nucleozymes for protein modification regrettably gave inconclusive results and we note that TamBA could lack the selectivity needed for site-selective protein modification. We also compared the other catalysts in this thesis, *i.e.* DMAP and PyOx, concluding that both have properties useful for different objectives and that the spatial orientation is more crucial for DMAP, since it had a much higher activity in the templated approach (chapter 5) with respect to the aptameric approach (chapter 4). Following this, we note that protein quantification with ImageJ (and of protein-DNA conjugates) is greatly influenced by Coomassie stain efficiency and proper lighting of the gel image.

For the future, we recommend the use of (a combination of) small ligands and of tailor-made synthetic catalysts that provide more control over the positioning of the catalyst construct and the modification chemistry. In addition, new hGQ DNAzyme-catalysed reactions or alternative metalloporphyrin-DNA complexes or other catalysts that use a covalently-bound substrate for novel catalysed reaction. We also advise the development of a catalytic DNA-templated aptamer approach to enable a more widespread applicability and research into linker length, orientation and rigidity to ensure optimal positioning of catalysts or reactive moieties. The protein-DNA_{temp} conjugates have potential for future studies into other proximity-based catalysts, such as Rh₂(OAc)₄, and the development of more complex DNA nanostructures for protein modification, including hairpin structures that protrude sideways or that open and close, moving catalysts away or towards the protein. We finalize the discussion by underscoring the success of DNA for catalytic protein modification, and that DNA systems can be designed to include catalytic sites, substrate binding sites and a regulatory domain. As such, our nanostructured catalytic constructs mimic crucial features hitherto only found in protein-modifying enzymes.

Acknowledgements



“Neem maar van mij aan, het was unaniem een belachelijk groot succes”

– Ivo Niehe

Acknowledgements/Dankwoord

It was a strange turn of events that led me to my PhD in Wageningen, but looking back, I would not change it even if I could. Despite the stressful times, I can say my PhD was a valuable and educative time for me and now I feel arrogant enough to say that also the project was a success. But naturally, I could never have achieved this on my own. So in this chapter I will thank everyone that both helped my PhD project succeed and made me enjoy my time here in Wageningen. I will do my best to thank everyone accordingly and without spelling errors, because if we are all honest, this is the chapter that is read the most.

To my supervisor *Bauke Albada*, I want to express my sincere gratitude for the many work-related and social things that you taught me during my PhD. Your door was always open for me and luckily so. DNA and proteins are stubborn molecules and the plenty of issues that surrounded them could always be tackled together; a success visibly underlined by all the publications we achieved over the years. What I also enjoyed is that besides work, we could easily talk about our personal lives and our enthusiasm for birds, such as our eagle-spotting trip to the Blaauwe Kamer at 6:30 in the morning. Looking back at when I started, your guidance has truly helped me to become an independent and dependable researcher as well as a better person in general. For this I will be forever grateful and I thank you again for helping me reach this milestone in my life.

I would like to thank my promotor *Han Zuilhof* for giving me the opportunity to do my PhD at ORC. Although we didn't always talk about my research, you provided me with advice in many other occasions, such as the organisation of the Israel trip, the science of philosophy and handling my reviewers. I also owe you a big thank you for supporting me in doing a postdoc internship in the group and for your efforts to keep the lab open during Corona. Thank you!

I wish to thank my thesis committee *Willem van Berkel*, *Gerard Roelfes*, *Maarten Merckx* and *Willem Velema* for accepting the invitation to my committee and taking the time to assess my thesis.

I thank *Floris van Delft* for the discussions in the beginning of my PhD. When I was still sorting out how to improve my workflow, you were there to give me advice on how to do so. Particularly, your emphasis on 'not to re-invent the wheel' was something I needed to hear whenever I embarked on new chemistry.

To my paranymps, thank you for standing with me during my defence. This is of course not the first time you helped me out. During my work, I could always rely on your advice in both lab and office. *Ian*, although our research topics were not entirely the same, we had enough overlap to help each other out. In addition we shared great times when organizing the Israel trip, going to CHAINS and attending the AMOCC summer school. You are a great person and I am glad that I could share most of my PhD time with you. You are a great friend and I am happy that we have met! *Jorick*, when I started my research you were my example. From you I learned all the basics of protein modification and more than once I find myself still copying your behaviour. This includes not just working meticulously and only trusting my own chemical stocks, but also cleaning up after everyone and loudly playing Foo Fighters over the stereo when working alone in the lab. Thank you for all your lessons and for being a great friend.

Sjoerd. When I joined ORC, you were the third member in our office. You were like my big brother in the lab to which I could turn when stuff was unknown. Even after you left ORC, I could still reach out to you for questions, social talks and board games. For this I thank you!

Yuri, ringbearer, Gnoerd the Gnurrier, gossip queen, my soup brother. It can easily be said that when it is just the two of us in the office, we produce enough sound for a full one. Thanks for all the meaningful and meaningless talks we had and will have, ranging from synthesis and proteins to döner and movie analogies for ORC!

Dieuwertje. Thank you for the lunchbreaks outside in which we could brainstorm about synthesis and peptides as well as our private lives. It was sad that you had to leave our office and lab, but luckily you still visit(ed) us regularly so we can get a 'broodje kroket'!

Alice, thank you for being you! What a cliché, but who cares! I love (and miss) your cheerful spirit in the lab! I still gladly make use of your phrase "italian minutes" even though people always argue that I am not Italian. Also thank you for answering my many questions about the final stages of the PhD.

Michel, thanks for taking the time to discuss our work issues and when those were unsolvable, we could always play some online games on Steam to forget them. For some of these issues however, there were only two possible choices...

Judith thank you for your help with my lab work when I was ill and for being so extremely organized. It really made my (work)life a lot better!

I thank *Frank* and *Hans* for all your help with the LC-MS. If you hadn't been around, none of my work would have been possible. When the HPLC-MS had concocted a new error, you guys were there to help me out, for which I and my thesis are extremely grateful.

Thanks go to *Barend* for helping with my NMR analyses. So far you have been able to avoid my opening doors and luckily so, because it would be an unfair repay of your kindness!

I thank *Adrie Westphal*, *Willy van den Berg* and *Willem van Berkel* for your knowledge on protein purification and FPLC and the biochemistry department for allowing me to use their equipment.

Thanks to *Twan America* for teaching me how to use MaxQuant and all the peptide analyses you have done for me on the nano-LC.

Also a big thank you to *Aleida Ruisch*, *Elly Geurtsen*, *Meta Bakker*, *Esther Kloppenberg-Fakkert* and *Erik van Rozendaal*[†]. By courtesy of your administration my project and teaching was possible and we were able to plan our Israel trip. *Aleida*, thank you for fighting with me to get my defence date. *Elly*, thank you for helping me understand my administrative work.

Thank you *Maurice* for being such a cheerful person. I could always share a laugh with you and could rely on your experience in biochemistry and with the WUR and its (many) policies. Also thank you for telling me to stay calm during the weeks leading up to my defence.

Thanks to *Hendra* for being my mentor in Orion and *Tjerk*, *Judith*, *Anne-Marie*, *Sandra* and *Eric* for with helping out with the teaching in general.

Henny, thank you for your assistance with purchases and the many social talks. You have to carry on the WE-day without me now.

Ellen, Esther, Julian, Lucas, Sybren, thank you for the genuine interest in my life and PhD project and for reminding me that those two are (surprisingly) not the same. *Kaustub*, bro, thank you for complaining with me about work and tinder alike. *Alexandre, Jacob, Satesh* our tennis games were a great distraction from my work stress. And naturally also thanks to the other PhD/PD colleagues from ORC that helped further my research with scientific (and social) discussions: *Alyssa, Andriy, Annemieke, Ariadni, Daniele, Esther van Andel, Fred, Fridolin, Hamit, Irene, Jasper, Jay, Kushal, Milou, Muthusamy, Pepijn, Rafael, Sevil* (good luck on the 29th!), *Simon, Zhen*.

I would also like to thank *Maarten Smulders, Fedor Miloserdov, Louis de Smet, Caroline Paul, Gert Salentijn, Sidhu Pujari* and *Michel Nielen* for their intellectual contributions to my project(s) as well as all others within ORC.

To my students: *Kenan Özbasi, Michael Hoekstra, Ayleen Lascaris, Sizèd van Enk, Jeanne Muizelaar, Sophie Wintermans* and *Marte Dros*. Thank you for the extra pairs of hands in the lab. Hopefully I was able to raise your interest in Organic Chemistry and to teach you all that an LC-MS is not user-friendly.

I would like to thank *Gijs van der Marel, Mark Overhand* and *Mathieu Noteborn[†]* for helping me start my PhD. Your letters of recommendation gave me the chance to start my PhD, thus making all of this possible. *Matthieu[†]*, you are missed.

Een kort bedankje voor *Ouwehands Dierenpark* en de mooie paraplu die ik van hen kreeg.

Robert, Branco, Mark, Jerwin en *Sander*. Bedankt voor jullie goede vriendschap. Ik mis de dagelijkse potjes kaarten die we deden op de universiteit en ik ben blij dat we elkaar nog regelmatig spreken! *Robert*, kom gerust nog een keer fietsen hier. Er zijn nog genoeg kastelen! *Branco*, samen gamen is altijd gezellig en met *Sander* erbij is het leuk om met z'n drieën te werken aan onze culturele opvoeding. Laten we snel weer een museum bezoeken! *Mark*, hopelijk hebben we de nieuwe Jurassic World gezien wanneer je dit leest! *Jerwin*, "een beetje geluk hebben" zijn nog altijd wijze woorden die ik tijdens mijn PhD ook meer dan eens heb herhaald.

Aan mijn middelbare school vrienden. *Ben, Ellen, Ferry, Jan Willem, Jillis, Mark, Menno, Talitha, Tim, Sander* en (sinds kort ook) *Lisette*. De hechtheid van onze vriendengroep wordt benadrukt door het feit dat we elkaar na meer dan tien jaar nog altijd spreken. Bedankt voor alle gezelligheid! Ik heb jullie gezelschap gemist tijdens de corona crisis en ik ben blij dat we inmiddels weer gezellig met de groep dingen kunnen doen! *Mark* bedankt voor al onze leuke golf tripjes! *dr. Jan Willem*, bedankt voor al onze gesprekken over PhD zijn!

Aan Jaap en mijn vrienden van het gamen. Bedankt voor de gezelligheid en zo af en toe ook wat input in mijn project. Hopelijk zien we elkaar binnenkort weer eens wat vaker in het echt!

Peter, Wilma, Pim en *Kim* bedankt voor jullie gastvrijheid en vriendelijkheid. Ik heb geboft met jullie als schoonfamilie, want jullie zijn geweldige mensen! *Peter*, jouw generositeit kent geen grenzen en daarvoor wil ik je bedanken. *Wilma*, bedankt voor je kookkunsten en dat we nu beiden onze scheikunde verhalen kwijt kunnen. *Pim*, bedankt voor je loyaliteit en natuurlijk die

lekkere broodjes in Ede. *Kim*, je levensadviezen gebruik ik regelmatig, dankjewel! *Guust*[†], bedankt voor je gezelschap tijdens het schrijven van mijn proefschrift.

Mijn grootouders[†], *opa en oma Vliegenthart* en *opa en oma Keijzer*. Ondanks dat jullie dit niet meer lezen wil ik jullie toch nog even noemen. Ik wil jullie bedanken voor alle liefde en zeggen dat ik jullie mis op een moment als dit, maar ik geloof er oprecht in dat jullie nog altijd over mijn schouder meekijken en daarin vind ik vaak steun. Daarom wil ik jullie mijn proefschrift deels aan jullie opdragen.

Lieve *Mam* en *Pap*. Het is onmogelijk om jullie genoeg te bedanken in slechts een stukje tekst, mede omdat jullie bijdrage aan mijn proefschrift niet alleen financieel en intellectueel was, maar ook emotioneel. Van kinds af aan heb ik me bij jullie veilig en geliefd gevoeld en dankzij jullie steun en opleiding ben ik de persoon geworden die dit proefschrift heeft kunnen maken. Jullie hebben altijd in mij geloofd en mij aangemoedigd om voor het hoogste te gaan, waardoor ik de kracht en motivatie had om niet alleen mijn bachelor en master, maar nu ook mijn PhD te behalen. Mijn levensdoel zal dan ook altijd zijn om jullie trots te maken en ik ben blij dat ik mijn proefschrift deels aan jullie mag opdragen. Ik wil jullie beiden bedanken voor dit alles en zeggen hoe blij ik ben dat jullie er op 2 september bij zijn om me ook dan weer te steunen. Ik hou immens veel van jullie!

Maikel, je bent niet alleen mijn broer, maar je bent ook mijn allerbeste vriend. Al sinds ik klein was, ben jij mijn rolmodel en ik heb mijn proefschrift ook aan jou opgedragen, omdat ik nooit zo ver was gekomen zonder jou. Ik ben superblij dat we zoveel dingen samen kunnen doen en hebben gedaan. Golfen, vakanties, videogames en nog veel meer. Bedankt voor het luisteren en meedenken over een onderwerp wat je eigenlijk niet helemaal lag. Ook jij bent er op 2 september weer bij en dat geeft mij moed. Ik hou van jou, broer!

Mijn allerliefste *Janka*. Een PhD titel behalen is prachtig, maar jou ontmoeten was misschien wel mijn grootste overwinning van de afgelopen jaren. Het feit dat het op een haar na allemaal nooit gebeurd had kunnen zijn, doet zelfs een scepticus als ik geloven in het lot. Naar mijn mening kan ik je niet genoeg bedanken voor wat je me waard bent. Jouw steun en liefde zijn alles wat ik wilde en nodig had in mijn leven en zo heb jij (onbewust) een enorme bijdrage aan mijn proefschrift geleverd. We vangen elkaar op na slechte dagen en vieren samen elkaars successen. Je bent de persoon op wie ik altijd kan rekenen, die me altijd steunt en die me altijd kan laten lachen. Ik ben trots en zielsgelukkig dat jij op 2 september bij me bent, want met niemand wilde ik dit moment liever delen dan met jou. Mijn lebenslangerschicksalsschatz, dankjewel voor alles. Ik hou van jou.

About the author



Curriculum vitae

Jordi Frederik Keijzer was born on the 30th of March 1993 in Zwijndrecht, the Netherlands. After primary school, he graduated from Develstein College in Zwijndrecht in 2011. He continued his education by studying Molecular Science and Technology at Leiden University and TU Delft, earning a bachelor's degree in 2014. Two years thereafter, in 2016, he obtained his master's degree in Chemistry at Leiden University, specializing in Organic Synthesis and Chemical Biology. This specialization led him to perform his promotion research at the Organic Chemistry Group of Wageningen University & Research Centre, where he worked under Bauke Albada and Han Zuilhof to harness the benefits of DNA for site-selective and trigger-responsive modification of native proteins. Currently, he continues to work in the same department as a postdoctoral researcher, developing complex DNA nanostructures for the selective and trigger-responsive modification of native proteins.



List of publications

D. Streefkerk, J. F. Keijzer, J. Lusunga, B. Albada, H. Zuilhof “SuFEx-based inhibitors for chymotrypsin” *manuscript in preparation*.

J. F. Keijzer, H. Zuilhof and B. Albada, “Calibrating catalytic DNA nanostructures for site-selective protein modification” *Chem. Eur.* **2022**, <https://doi.org/10.1002/chem.202200895>

J. F. Keijzer and B. Albada, “DNA-assisted Site-selective Protein Modification” *Biopolymers* **2021**, *113* (3), 1-8.

J. F. Keijzer, J. Firet and B. Albada, “Site-selective and inducible acylation of thrombin using aptamer-catalyst conjugates” *Chem. Commun.* **2021**, *57*, 12960-12963

J. F. Keijzer, S. Wintermans, M. Dros, H. Zuilhof and B. Albada, “Aptamer-Assisted Bioconjugation of Tyrosine Derivatives with hemin/G-quadruplex (hGQ) DNAzyme Nucleoapzyme Nanostructures” *ChemCatChem* **2021**, *13*, 4618–4624.

B. Albada, J. F. Keijzer, H. Zuilhof and F. L. van Delft, “Oxidation-Induced “One-Pot” Click Chemistry” *Chem. Rev.* **2021**, *121* (12), 7032–7058.

J. F. Keijzer and B. Albada, “Site-Specific and Trigger-Activated Modification of Proteins by Means of Catalytic Hemin/G-quadruplex DNAzyme Nanostructures” *Bioconjug. Chem.* **2020**, *31* (10), 2283–2287.

J. J. Bruins, C. van de Wouw, J. F. Keijzer, B. Albada and F. L. van Delft, “Inducible, Selective Labeling of Proteins via Enzymatic Oxidation of Tyrosine” *Methods in Molecular Biology* **vol 2012**. Humana, New York, NY., ISBN: 978-1-4939-9546-2

S. H. Grimm, B. Gagestein, J. F. Keijzer, N. Liu, R. H. Wijdeven, E. B. Lenselink, A. W. Tuin, A. M. C. H. van den Nieuwendijk, G. J. P. van Westen, C. A. A. van Boeckel, H. S. Overkleeft, J. Neefjes and M. van der Stelt, “Comprehensive structure-activity-relationship of azaindoles as highly potent FLT3 inhibitors” *Bioorganic Med. Chem.* **2019**, *27* (5), 692–699.

Overview of completed training activities

Discipline-specific courses	Organizing institute	Year
Summer Course Glycosciences	VLAG / GBB	2018
Advanced Organic Chemistry	ORC	2018-2021
CHAINS symposium 2017	KNCV	2017
Sym. Medicinal Chemistry of Antimicrobials to Antiparasitics	KNCV	2018
Wageningen Symposium of Organic Chemistry	KNCV	2018
AMOCC / NIOK	HRSMC	2018
Seminar Series - Molecular Life Sciences 2017-2018	WUR	2017-2018
CHAINS symposium 2018	KNCV	2018
Seminar Series - Molecular Life Sciences 2018-2019	WUR	2018-2019
CHAINS symposium 2019 – poster presentation	KNCV	2019
CHAINS symposium 2020 – oral presentation	KNCV	2020
CHAINS symposium 2021 – oral presentation	KNCV	2021
General courses	Organizing institute	Year
VLAG PhD week	VLAG	2018
Brain training	WUR Library	2018
Scientific Artwork - vector graphics and images	WUR Library	2018
Scientific Publishing	WUR Library	2018
Scientific Writing	VLAG	2019
Career Perspectives	VLAG	2021
Last Stretch of the PhD Program	VLAG	2022
Additional activities	Organizing institute	Year
Research proposal	ORC	2017
Organizing committee of PhD Study Trip to Israel	ORC	2017-2019
PhD Study Trip to Israel	ORC	2019
Labhead	ORC	2018-2021
Tri-weekly group meetings	ORC	2017-2022
Colloquia	ORC	2017-2022
Teaching activities	Organizing institute	Year
ORC-12803: Organic chemistry 1	ORC	2017-2019
ORC-20306: Bio-organic chemistry	ORC	2017-2021
ORC-31303: Research Methods in Organic Chemistry	ORC	2019-2020
Proctor at student exams	ORC/WUR	2018-2019
Thesis (BSc/MSc) supervision	ORC/WUR	2017-2021

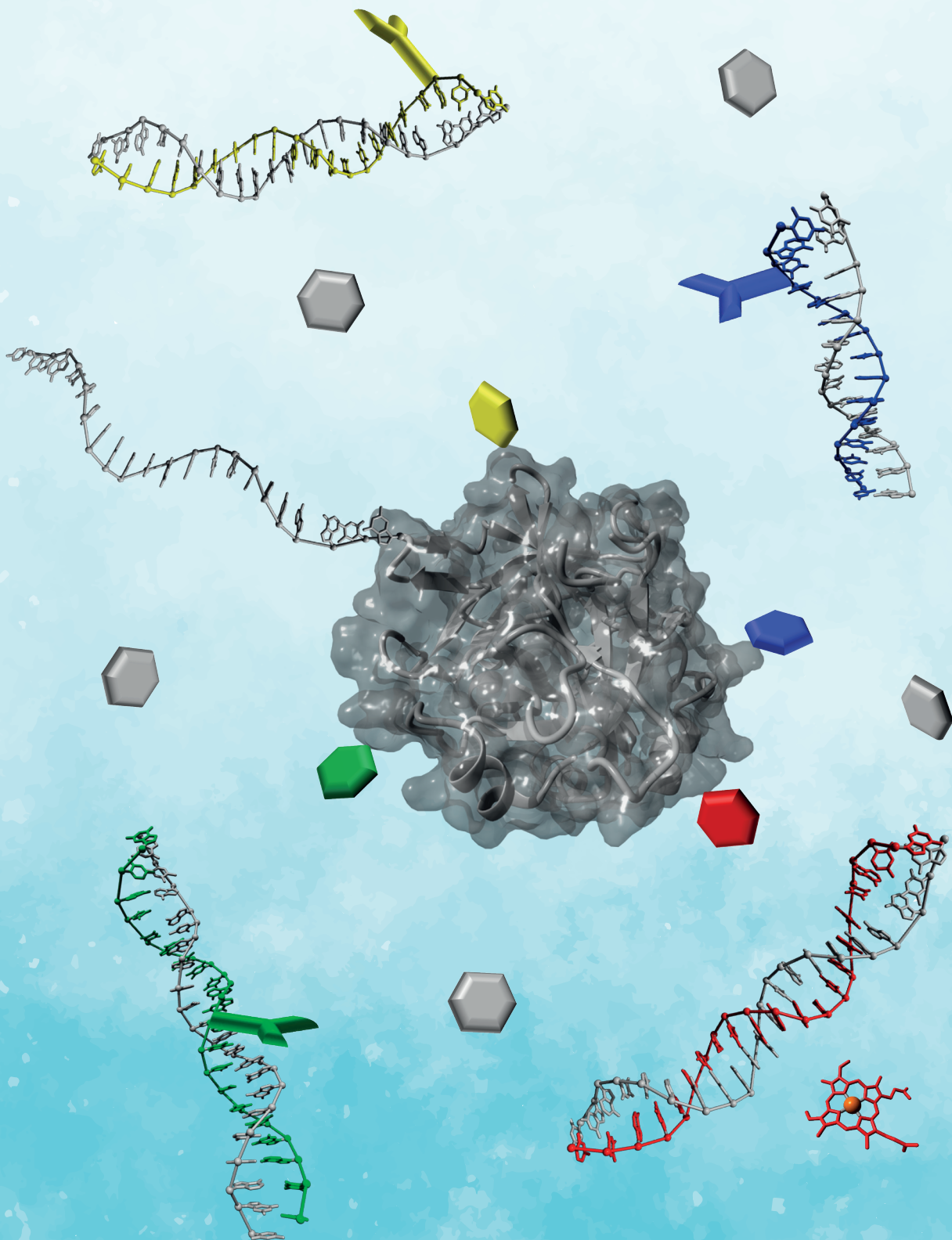
The research described in this thesis was financially supported by NWO, de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (project number: 711.017.004).

Financial support from Wageningen University for printing this thesis is gratefully acknowledged.

Cover design by author

Crystal structures on cover: 6EDB (DNA duplex), 5EW1 (protein, blue and green aptamer), 1XAV (red aptamer).

Printed by Proefschriften.nl, Deventer, The Netherlands



Invitation

to the public defence
of my Ph.D. thesis
entitled:

Controlled Modification of Wild-type Proteins using DNA Nanotools

2nd of September 2022
at 16:00 in Omnia
6708 PH, Wageningen

Jordi F. Keijzer

jordi.keijzer@wur.nl
jordikeijzer@hotmail.com

Paranymphs:

Ian de Bus
ian.debus@hotmail.com

Jorick Bruins
jorickbruins@hotmail.com