The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/72127

Please be advised that this information was generated on 2019-04-20 and may be subject to change.
Processive bio-hybrid catalysts based on the T4 clamp protein

Een wetenschappelijke proeve op het gebied van de
Natuurwetenschappen, Wiskunde en Informatica

Proefschrift

Ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de Rector Magnificus prof mr. S.C.J.J. Kortmann,
volgens besluit van het College van Decanen
in het openbaar te verdedigen op donderdag 22 mei 2008
om 10:30 precies

door

Joost Clerx

geboren op 15 januari 1979
te Nijmegen
Promotores:
Prof. dr. R.J.M. Nolte
Prof. dr. A.E. Rowan

Copromotor:
Dr. J.J.L.M. Cornelissen

Manuscriptcommissie:
Prof. G.J.M. Pruijn
Dr. A. Herrmann (Universiteit van Groningen)
Dr. M. Merkx (Technische Universiteit Eindhoven)

Paranimfen:
Marta Comellas Aragonès
Mark Damen

The work described in this thesis was financially supported by the National Research School Combination Catalysis (NRSC-Catalysis)

Realisatie: Ipskamp Printpartners

Omslagontwerp: Hilde Ketelaars

ISBN: 978-90-9022869-3
Research should be the sole focus of your life. I expect you to eat, drink, and sleep research.

I'm allowed to sleep? Only in a lie-awake-at-night, obsessed-over-it kind of way.

Outside interests? Only tell me you're not serious about getting your PhD. I mean, what could be more interesting than our research?

...I can't answer that. Exactly. Nothing.
Preface

The work described in this thesis is part of an on-going research programme aimed at developing so-called processive catalysts, i.e. catalysts can bind and slide over polymeric substrates while catalyzing a reaction. The binding of such catalysts leads to so-called ring-on-a-rod structures termed rotaxanes, in which the catalyst is mechanically linked to its substrate. Since the catalyst cannot dissociate from its substrate, this may provide distinct advantages, such as making the process as a whole more efficient and faster. In billions of years, Nature has evolved enzymes that utilize this principle, and nowadays they provide a source of inspiration to chemists pursuing the development of new and more efficient catalysts.

In this thesis, the possibility of using the T4 DNA polymerase processivity factor as scaffold for the construction of bio-hybrid catalysts is explored. This processivity factor, also called the T4 sliding clamp protein, is a ring-shaped protein, which is able to slide freely over DNA. It is essential for efficient DNA replication, i.e. without this protein the T4 bacteriophage cannot replicate itself. This thesis will start with a literature survey, in which several examples of ring-shaped proteins in Nature are presented. This includes clamp proteins of various organisms, and a detailed discussion of T4 DNA replication and the role of the T4 clamp in this process. Two strategies are developed to construct the bio-hybrid processive oxidation catalysts based on clamp proteins: the catalyst is coupled directly to the protein itself (direct labelling approach), or the catalyst is coupled to a peptide that binds to the clamp via the same interaction that is used by the T4 polymerase (padlock approach). Furthermore, two different chemical oxidation catalysts are selected to be used in this study, one of which is commercially available, while the synthesis of the other is described in this thesis. To explore the feasibility and scope of the proposed methods, it is investigated whether the T4 clamp can bind to synthetic substrates that can potentially be used in combination with the bio-hybrid catalysts. Peptide analogues of the domain of the T4 DNA polymerase that are known to bind to the T4 clamp are synthesized and used to explore its interactions with the clamp. Finally, different biohybrid catalysts are prepared and studied in oxidation experiments according to the different approaches. The oxidation reactions were analyzed with conventional gel electrophoresis methods and a novel streptavidin-labelling approach that allowed the reactions to be studied on the single molecular level by AFM. In order to relate the performance of the biohybrid catalysts and the biological functions of the T4 replication different T4 DNA replication proteins are also studied by AFM, both individually and as complexes with other proteins and DNA.
List of terms and abbreviations

1. General

distributive enzyme  
enzyme that only catalyzes a few turnovers before it dissociates from its template
DNA  
deoxyribonucleic acid
dsDNA  
double stranded DNA
gp  
gene product, number indicates place on genome; integer numbers indicate essential genes, non-integer numbers genes are not or conditionally essential.
processive enzyme  
enzyme that sequentially catalyzes a large amount of turnovers before it dissociates from its template
RB69 phage  
T4-like phage, of which the replication proteins have a high degree of sequence homology with their T4 analogues
RDR  
recombination depended DNA replication
replisome  
complex of proteins dedicated to DNA replication
rotaxane  
mechanically interlocked molecular architecture; structure of molecules that are joined by their shape (topology) rather than direct chemical bonds between them
ssDNA  
single stranded DNA
wt  
wild-type, protein that has amino acid sequence as found in Nature

2. Replication proteins

dda  
distributive and cooperative 5’→3’ helicase
FEN1  
RNase protein that removes the RNA primers after Okazaki fragment synthesis
gp32  
single stranded binding protein
gp33  
transcriptional co-activator, inhibits transcription in absence of gp45
gp41  
processive DNA helicase
gp43  
DNA polymerase (exo-: exonuclease deficient mutant; exo+: wild-type)
gp44/62  
clamp loader protein, 4:1 complex of gp44 and gp62
gp45  
clamp protein, polymerase processivity factor
gp46/47  
recombination exonuclease, generates ssDNA regions for initiation of RDR
gp49  
T4 endonuclease VII, resolves Holliday junctions
gp55  
σ-factor of RNA polymerase, recognizes T4 late promoter sequences
gp59  
helicase loading protein
gp61  
primase
ligase I  
chemically joins two adjacent DNA strands
PCNA  
proliferating cell nuclear antigen, the clamp protein of eukaryotes
RDR  
recombination-dependent replication
RFC  
replication factor C, clamp loader of eukaryotes
RNase H  
nuclease that removes RNA primers of Okazaki fragments, or RNA from RNA-DNA hybrids
uvsW  
DNA/RNA helicase
uvsX  
general recombinase protein, enables the migration of recombination intermediates in one direction (recA analogue)
uvsY  
recombination mediator protein, aids the assembly of uvsX on ssDNA
3. Chemical abbreviations

A  adenosine (referring to DNA), or alanine (referring to amino acid)
AAA DNA sequence of three consecutive adenosine bases
BLAST basic local alignment search tool
BABE p-Bromoacetamidobenzyl-EDTA
Boc tert-butyloxycarbonyl
CI chemical ionization (mass spectroscopy)
DTT DL-dithiothreitol
EM electron microscopy
ESI electrospray ionization (mass spectroscopy)
FeBABE iron(III)-BABE complex
FRET fluorescence resonance energy transfer
GC gas chromatography
IR infrared (spectroscopy)
NMR nucleic magnetic resonance
Maldi-TOF matrix-assisted laser desorption/ionization, time of flight
MS mass spectroscopy
NHS N-hydroxy succinimide
ODN oligodeoxyribonucleotide, a short synthetic DNA strand
PAGE polyacrylamide gel electrophoresis
PDB protein data bank
PEG polyethylene glycol
PIAA polyisocyanoalanine-alanine
PIAAA polyisocyanoalanine-alanine-alanine
ppm parts per million
TLC thin layer chromatography
TMPyP 5-(4-hydroxyphenyl)-10,15,20-tris(N-methylpyridinium-4-yl)porphyrin, trimethyl pyridinium porphyrin
TPyP 5-(4-hydroxyphenyl)-10,15,20-tris(4-pyridyl)porphyrin, tripyridyl porphyrin
Tris tris(hydroxymethyl)aminomethane
TTP 5-(4-hydroxyphenyl)-10,15,20-tris(4-methylphenyl)porphyrin, tritoluyl porphyrin
UV-Vis ultraviolet-visual (spectroscopy)

4. Amino acid codes

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Ala</td>
<td>Alanine</td>
</tr>
<tr>
<td>C</td>
<td>Cys</td>
<td>Cysteine</td>
</tr>
<tr>
<td>D</td>
<td>Asp</td>
<td>Aspartic acid</td>
</tr>
<tr>
<td>E</td>
<td>Glu</td>
<td>Glutamic acid</td>
</tr>
<tr>
<td>F</td>
<td>Phe</td>
<td>Phenylalanine</td>
</tr>
<tr>
<td>G</td>
<td>Gly</td>
<td>Glycine</td>
</tr>
<tr>
<td>H</td>
<td>His</td>
<td>Histidine</td>
</tr>
<tr>
<td>I</td>
<td>Ile</td>
<td>Isoleucine</td>
</tr>
<tr>
<td>K</td>
<td>Lys</td>
<td>Lysine</td>
</tr>
<tr>
<td>L</td>
<td>Leu</td>
<td>Leucine</td>
</tr>
<tr>
<td>M</td>
<td>Met</td>
<td>Methionine</td>
</tr>
<tr>
<td>N</td>
<td>Asn</td>
<td>Asparagine</td>
</tr>
<tr>
<td>P</td>
<td>Pro</td>
<td>Proline</td>
</tr>
<tr>
<td>R</td>
<td>Arg</td>
<td>Arginine</td>
</tr>
<tr>
<td>Q</td>
<td>Gln</td>
<td>Glutamine</td>
</tr>
<tr>
<td>S</td>
<td>Ser</td>
<td>Serine</td>
</tr>
<tr>
<td>T</td>
<td>Thr</td>
<td>Threonine</td>
</tr>
<tr>
<td>V</td>
<td>Val</td>
<td>Valine</td>
</tr>
<tr>
<td>W</td>
<td>Trp</td>
<td>Tryptophan</td>
</tr>
<tr>
<td>Y</td>
<td>Tyr</td>
<td>Tyrosine</td>
</tr>
</tbody>
</table>
# Table of contents

**Preface**  
3

**List of terms and abbreviations**  
4

**Table of contents**  
6

## Chapter 1 – Introduction  
11
1. DNA replication  
12
2. Sliding clamp proteins  
15
3. Clamp proteins as molecular toolbelts  
19
4. Tracking of clamps on DNA  
20
5. T4 bacteriophage DNA replication  
21
   5.1 General  
   21
   5.2 Initiation of replication  
   21
   5.3 Assembly of the T4 holoenzyme  
   22
   5.4 Interactions within the T4 holoenzyme  
   26
   5.5 Primosome assembly  
   27
   5.6 Role of the single strand DNA binding protein (gp32)  
   29
   5.7 The T4 replisome  
   30
   5.8 T4 late transcription  
   32
6. Comparison with other replication complexes  
   34
7. Other processive and toroidal enzymes  
   37
8. Synthetic rotaxane catalysts as models for processive enzymes  
   39
9. References  
   41

## Chapter 2 – Aim and outline of this thesis  
55
1. Aim of this thesis  
56
2. How will the clamp be labelled with catalysts?  
57
3. How can the clamp be loaded onto DNA?  
58
4. Porphyrins as DNA oxidizing catalysts  
59
5. Hydroxyl radical generating species as catalysts  
62
6. Outline of this thesis  
63
7. References  
63
Chapter 3 – Investigations on the interaction of the T4 clamp with synthetic polymers

1. Introduction
2. Results and discussion
   2.1 ATPase stimulation assays with polymeric substrates
   2.2 Pre-steady state fluorescence measurements with polymers
3. Conclusions
4. Experimental section
5. References

Chapter 4 – DNA block-copolymers as T4 clamp substrates

1. Introduction
2. Results and discussion
   2.1 Synthesis of PEG-DNA conjugates
   2.2 Synthesis of poly(butadiene)-DNA conjugates
   2.3 Towards “clicked” DNA-polymer conjugates
   2.4 Synthesis of conjugates via metal-free “click” reactions
3. Conclusions
4. Experimental section
5. References

Chapter 5 – Interaction of the T4 DNA polymerase C-terminus with the T4 clamp

1. Introduction
2. Results and discussion
   2.1 FRET experiments
   2.2 ATPase stimulation studies
3. Conclusions
4. Experimental section
5. References

Chapter 6 – Synthesis and characterization of maleimide functionalized porphyrin derivatives

1. Introduction
2. Results and discussion
   2.1 First generation maleimidotritolyl porphyrins
   2.2 Second generation maleimidotritolylporphyrins
   2.3 Second generation maleimido tripyridylporphyrins
   2.4 Third generation maleimido tripyridylporphyrins
   2.5 Maleimide functionalized Protoporphyrin IX derivatives
3. Conclusions
4. Experimental section
5. References
Chapter 7 – Oxidation of DNA by clamp-based catalysts

1. Introduction
2. Results and discussion
   2.1 Protein expression and labelling
   2.2 Oxidation of DNA plasmids by FeBABE labelled clamps and peptides
   2.3 AFM as an alternative tool for the analysis of porphyrin oxidation experiments
   2.4 Oxidation of DNA plasmids by porphyrin labelled clamps
3. Conclusion
4. Experimental section
5. References

Chapter 8 – Atomic force microscopy analysis of T4 replication proteins and complexes

1. Introduction
2. Results and discussion
   2.1 Imaging of the T4 replication proteins
   2.2 Cross-linked complexes of the T4 replication proteins
   2.3 Complexes of the T4 clamp on DNA
   2.4 Complexes of the T4 clamp loader and DNA polymerase on DNA
   2.5 Imaging of the T4 holoenzyme and T4 replication complexes
3. Conclusion
4. Experimental section
5. References

Appendix

1. Fluorescence resonance energy transfer (FRET)
2. Calculating donor-acceptor distances from stopped-flow fluorescence spectroscopy experiments
3. Analysis of titration experiments
4. Scatchard plot analysis
5. References

Summary

Samenvatting

Dankwoord

Curriculum Vitae
DNA replication is a highly important and intriguing process that lies at the very heart of life. Decades of investigation have revealed that complexes comprising various proteins with dedicated enzymatic functions are responsible for faithfully copying DNA at astonishing speeds. The efficiency of replication is dramatically enhanced by clamp proteins, which are circular proteins that thread onto DNA and tether DNA polymerases to their template. This introduction focuses on these clamp proteins. The structures of clamp proteins of various organisms from all domains of life are discussed, as well as their physical properties when attached to DNA and the interaction of various proteins with these clamps. As a paradigm, the T4 replication process, and the role of the T4 clamp protein herein, is discussed in detail. A brief comparison is made with other well-studied replication systems. This chapter is concluded with a discussion of synthetic rotaxanes that can be regarded as synthetic mimics of toroidal enzymes found throughout Nature.
1. DNA replication

One of the most striking properties of living organisms is their ability to proliferate. In this respect, the process of DNA replication is the most important process that occurs within living species. It is not surprising, therefore, that every organism trusts upon a highly specialized and organized replication machinery for the duplication of its genomic material. The heart of the replication machinery is a DNA polymerase. The DNA polymerases dedicated to replication generally belong to the so-called B family of DNA polymerases and are able to synthesize new DNA starting from the 3' end of DNA with high activity and fidelity. Their structure resembles a half-open right hand, consisting of three subdomains (Figure 1). The active site, built up from a number of highly conserved amino acid sequences, is found in the palm domain at the bottom of the hand, and generally consists of two magnesium ions complexed to aspartates that interact with the incoming nucleotide triphosphate. The finger and thumb domain are less well conserved among DNA polymerases.  

![DNA polymerases](image.png)

**Figure 1** - DNA polymerases. A. Homology model of the crystal structure of RB69 phage DNA polymerase (gp43) bound to DNA. This polymerase has sequence homology to human DNA polymerase α. The three domains (palm, finger and thumb) of the polymerase are coloured differently and indicated with labels. B. Schematic representation of DNA polymerases with the conserved catalytic amino acid motifs A, B, and C. C. Switching of DNA polymerases between the polymerization and exonuclease modes. Figure A is from Steitz et al. Figures B. and C. are adapted from Baker et al.
polymerases, and are involved in interactions with the (single stranded) template strand and the primer-template junction, respectively. All polymerases, except eukaryotic pol α, also have an exonuclease activity, allowing them to excise the terminal nucleotide. This proofreading feature is crucial for maintaining the high fidelity of nucleotide incorporation during replication.

DNA replication in humans is very complex due to the many proteins involved in this process and the structure of the human genome. The yeast *Saccharomyces cerevisiae*, a model for more complex eukaryotes, has an estimated 6,000 genes (one fifth of that of humans), of which 88 are proposed to be involved in DNA replication, and a further 49 in DNA recombination. Because the main features of DNA replication are conserved in all forms of life, the study of replication processes in a number of lower complexity organisms and viruses allows insight into the complex replication process in humans. This is of pivotal importance with respect to diseases related to DNA metabolism and cancer. One of the simplest models is provided by the T7 bacteriophage, which is able to replicate DNA at a rate of 300 nucleotides per second (nt/s) by the action of only four proteins. The more complex T4 bacteriophage encodes a replisome consisting of 8 proteins, which is able to replicate DNA at 500 nt/s. The T4 bacteriophage replicates its 169 kb genome in 8-12 minutes, while it takes *Escherichia coli* 42 minutes to complete the replication of its 4.5 mb genome. The *E. coli* replication proteins assemble to form a 0.9 MDa complex which copies DNA at rates up to 750 nt/s. There are 10-20 copies of replicative DNA polymerases within a single *E. coli* cell. It should be noted that the individual functions within the replisome (i.e. polymerase, primase, helicase activities, etc.) are present in all three systems, irrespective of their wide range of complexity. Replication in humans is much slower, and proceeds at a rate of 100 nt/s. The 3x10^9 base pairs of the human genome are copied in 8 hours; unlike the previous examples, the human genome is divided into replicons, areas that each have one

---

**Figure 2** – Cartoon of DNA replication in 2D. A. Cartoon showing the main enzymatic activities found at the replication fork. B. The polymerases are actually coupled, providing an indication how the replication fork looks like in 3D. This model, in which the single strand DNA binding protein-covered lagging strand elongates until it eventually collapses, is called the trombone model. Figure is adapted from Alberts *et al.*
origin of replication (ori) and are copied simultaneously. There are 10,000-100,000 of these replicons in the human genome. Simian virus 40 (SV40) has been used extensively as a paradigm to study DNA replication in human cells.3,24,25

The investigation of the simpler replication models has provided a wealth of information.26 DNA synthesis occurs in a semi-conservative fashion, i.e. each of the parent strands becomes part of a new duplex.27 The two strands are copied at the same time, and the DNA polymerases are coupled through direct or indirect protein-protein interactions. DNA polymerases can only add nucleotides at the 3’ end of DNA, thus polymerization only occurs in the 3’ → 5’ direction. Because of the polarity of dsDNA, one strand (the leading strand) can be synthesized continuously, while DNA synthesis on the other (lagging strand) occurs in the direction opposite to the movement of the replication fork. As a result, DNA synthesis on the lagging strand is discontinuous, meaning that short stretches of DNA (from 100-200 bases for eukaryotes up to 1-2 kb for prokaryotes), called Okazaki fragments28, are synthesized. Each Okazaki fragment is elongated from short strands of RNA or DNA-RNA hybrids (primers) that are synthesized by primases and provide a starting point for the DNA polymerases. Completion of lagging strand synthesis requires the orchestrated action of a number of enzymes: flap endonucleases remove RNA primers, any remaining gaps in the DNA are filled in by other DNA polymerases, and the strands are finally joined together by DNA ligases.

Besides the primary enzymatic functions that can be identified at a replication fork, a number of other proteins are also involved in the replication of DNA (Figure 2). First, single-stranded binding proteins (SSB) play an important role in stabilizing the single stranded DNA (ssDNA) strands that result from the DNA unwinding activity of the DNA helicase, by preventing any secondary structures (such as hairpins) to be formed that can delay the replication process. They also enhance the activity of the other replication proteins by diminishing unspecific binding of these proteins to ssDNA. Secondly, the replicative DNA polymerases are associated with a circular proteins, called sliding clamps,29 that dramatically enhance the processivity of the DNA polymerase. These proteins are essential for DNA replication since the polymerases themselves are distributive rather than processive enzymes, meaning that they only add a relatively small amount of nucleotides to a growing DNA strand before dissociating from the template DNA. For instance, the main replicative enzyme of E. coli, pol III core, has an intrinsic activity of 20 nt/s and a processivity of not more than 55 nucleotides under optimal conditions.30 In contrast, when associated with its β-subunit, the overall rate increases to 750 nt/s and the processivity is higher than 50,000 nucleotides.20,31 The sliding clamps are assembled on DNA by clamp loader enzymes, called replication factor C (RFC) in eukaryotes. All the individual protein activities need to be coordinated to achieve optimal replication speeds.
2. Sliding clamp proteins

The speed of DNA replication by DNA polymerases is governed by two factors: the intrinsic activity of the DNA polymerase and its processivity. The latter is generally defined as the number of turnovers in a single binding event and depends on the dissociation speed. In the case of DNA replication, this is the number of bases which are added to the growing DNA strand before the DNA polymerase dissociates from its substrate. In most replisomes, the polymerase is associated with a circular protein, called a processivity clamp, to increase its affinity for the DNA substrate and thus its processivity. Together the clamp loader complex, responsible for loading the clamp onto a primer-template site on the DNA, the clamp forms the so-called polymerase accessory proteins. Remarkably, all the clamps share the same topological structure. A number of crystal structures of clamp proteins has been solved in recent years, providing an intriguing insight into this class of proteins. A selection of these crystal structures is represented in Figure 3. The clamps are generally donut-shaped, with a central channel that can accommodate a double stranded DNA molecule. Their structure possesses a six-fold symmetry with α-helices on the inside of the ring containing a number of positively charged amino acids (e.g. lysines, arginines). The helices are held together by β-sheets on the outside of the ring. Despite these common features, the clamp proteins share no sequence similarity, except for closely related organisms and viruses. In fact, a lot of individual clamp proteins have developed their own unique characteristics (Figure 3). For instance, in the clamp proteins of eukaryotes, which are termed proliferating cell nuclear antigens (PCNAs), the six symmetry elements are usually divided into three subunits, while the clamp proteins of bacteria have only two subunits. The homotrimeric clamp protein of the T4 bacteriophage, gene product (gp) 45, is closed in its crystal structure, but one of the three subunit interfaces opens in solution. The clamp protein of the hyperthermophilic archaeon Sulfolobus solfataricus is unique in the sense that it is heterotrimeric, and each of the individual subunits interacts with a specific replication protein (DNA polymerase, flap endonuclease FEN1 or DNA ligase I). The protein trimer is assembled in a specific way; the PCNA3 subunit can only interact with a PCNA1-PCNA2 dimer to form the closed ring trimer. The clamp from cytomegalovirus (UL44) is homodimeric but unlike the prokaryotic clamps its structure resembles the shape of a trimeric clamp. The clamp appears to have lost one subunit, and is C-shaped rather than a closed circle. Finally, the herpes simplex virus clamp (UL42) functions as a single subunit. Although its structure resembles that of other clamp subunits, UL42 did not oligomerize in crosslinking experiments with DNA.
A. Clamps from eukaryotes

PCNA from *Saccharomyces cerevisiae* (1PLQ)  
PCNA from *Homo sapiens* (1VYM)

B. Clamps from prokaryotes

β-subunit of the *Escherichia coli* polymerase (2POL)  
β-subunit of the *Streptococcus pyogenes* polymerase (2AVT)

C. Clamps from archaea and viruses

Gp45 from the T4 Bacteriophage (1CZD)  
PCNA from *Sulfobus solfataricus* (2X2)

UL44 subunit from the cytomegalovirus DNA polymerase (1YYP)  
UL42 subunit of the herpes simplex virus (1DML)
Although the vast majority of organisms and viruses have a clamp protein associated with their replicative DNA polymerases, some use other strategies to tether their polymerase to DNA. Most notably, the T7 DNA polymerase relies on thioredoxin, an ubiquitous redox regulation protein from its host organism *E. coli* rather than a clamp as processivity factor. The processivity factor enables the T7 replisome to copy its genome in one single binding event. The crystal structure of the polymerase (which belongs to the A family) with thioredoxin bound to a primer-template DNA substrate has been solved and is represented in Figure 4. Although not apparent in the crystal structure, the flexible loop of the thioredoxin binding domain was proposed to fold back over the polymerase to form a closed ring around the DNA substrate. Introduction of a thioredoxin binding site in other DNA polymerases (*E. coli* pol I and *Thermus aquaticus* [Taq] polymerase) also led to enhanced processivity in these enzymes. The clamping action of thioredoxin on DNA polymerases is reminiscent of the clamp subunit found in eukaryotic RNA polymerase II, which also acts as a lid for this enzyme.

![Figure 4 - DNA polymerases and their processivity factors. A. Model of the T4 holoenzyme, consisting of the crystal structures of T4 related RB69 DNA polymerase and the T4 clamp protein. B. Crystal structure of the T7 replication complex, showing the T7 DNA polymerase complexed to a DNA substrate and its processivity factor, *E. coli* thioredoxin. Depth cue was added to the latter for clarity reasons. Magnesium ions are represented as small balls.](image)

In addition to their essential role in DNA replication, clamp proteins also play an important role in other processes of DNA metabolism. For instance, the T4 clamp (gp45) is also an activator for ‘late’ transcription (i.e. transcription of the genes encoding the viral protein capsid). The role of eukaryotic PCNAs in processes other than replication has been...
studied intensively, and provided an extensive list of (up to 65) PCNA-interacting proteins (PIPs). For a review the reader is referred to Warbrick,58 Maga et al.,59 Vivona et al.,60 (which also contains a list of proteins interacting with archaea PCNA and the prokaryotic β-subunit), and Tsurimoto et al.61 These proteins are involved in processes as diverse as DNA replication, DNA repair, cell cycle control, chromatin metabolism, gene expression, sister chromatid cohesion (separation of the chromosomes during cell division), and apoptosis.59,62 PCNA is an essential protein for eukaryotes; yeast cells in which the gene encoding PCNA is deleted are not viable.63

The interaction of the vast number of PCNA-interacting proteins with different functions raises an interesting question: how do the individual proteins bind to PCNA and how is the binding of these proteins coordinated for the different processes? The proteins interact with PCNA via defined binding sites (Figure 5A),59 and share a PIP-motif with consensus sequence QXX[L/I/M]XX[F/Y][F/Y] at their N or C-terminus. In this sequence, X can be any amino acid and either one of the amino acids specified within brackets can be part of the sequence at that place. Another interaction motif, the KA box, has also been identified ([A/L/I][A/L/Q]XX[L/V]).64 Thus, the fact that different proteins have to compete for the same binding site provides a way of regulation.65 Although protein-protein interactions also occur at other sites than at the PIP or KA binding sites, these binding motifs are indispensable for interaction with PCNA.

Several clamps from different origins have been crystallized with peptides corresponding to the interaction domains of clamp interacting proteins. These crystal structures include that of PCNA with the C-terminus of p21 (a cell cycle regulatory

![Diagram of PCNA interaction domains](image)

**Figure 5** - A. The overlapping binding sites of PIPs on PCNA, the clamp protein of eukaryotes. Interactions with other proteins take place at the interdomain connecting loop (PIP box: pol δ, p21, MeCTr, DNA ligase I, Fen 1), at the C-terminal tail (pol ε, RF-C, CDK2, Gadd45), or at the inner helices (Cyclin D). The figure is redrawn from Maga et al.,59 using PDB structure 1AXC.66 Colouring based on hydrophobicity (except for the indicated regions). B. Comparison of clamp binding motifs among various organisms and viruses. CAF1 is the histone chaperone chromatin assembly factor 1.67 The herpes simplex virus (HSV1) interaction motif is from Zuccola,44 and the T4 motif is from Berdis and Wong et al.6869 The figure was redrawn from Dalrymple et al.70

---

Chapter 1

- 18 -
protein", FEN1 and δ p66 (polymerase domain of HIV type I reverse transcriptase), and other clamps such as gp45 of the RB69 phage (closely related to the T4 bacteriophage), UL42, UL44, and the β-subunit with the C-terminus of their DNA polymerase. The structures show that the peptides predominantly bind to a hydrophobic patch near the interdomain connecting loop (the largely unstructured loop that connects two of the six symmetry domains of the clamp). Furthermore, comparison of the clamp interaction domains of various organisms, bacteriophages and viruses suggests that this interaction is extremely ancient and has been conserved throughout evolution (Figure 5B). In fact, the clamp from a gram-negative bacterium (E. coli) can still substitute that of a gram-positive bacterium (Staphylococcus aureus), although they diverged over 1 billion years ago. Because peptides analogous to the PCNA binding domain compete with the binding of other proteins to PCNA, they could be used to design inhibitors for PCNA interaction that would lead to novel therapeutics.

3. Clamp proteins as molecular toolbelts

As discussed above, the clamp protein of S. solfataricus has three different subunits that each bind a specific replication protein. Apparently, this clamp functions as a molecular toolbelt, providing all the ‘tools’ needed for the replication of a DNA strand and the maturation of Okazaki fragments. Each subunit of clamp proteins usually has one or two binding sites for clamp interacting proteins, thus the binding of multiple proteins to one clamp is only limited to repulsive steric interactions between the proteins. In the case of gp45, there are secondary binding sites near the interdomain connecting loop (three per protein trimer), while the main interaction takes place at the open subunit interface. Elegant studies of Yang et al. showed that adding an inactive trap polymerase to a processive DNA replication complex rapidly leads to inhibition of DNA synthesis. It was proposed that multiple DNA polymerases bind to one clamp concurrently, and exchange between the secondary and primary binding sites is fast. This study has led to a more refined insight into processive DNA replication complexes; while the complexes are indeed very processive, exchange between different polymerases does occur when allowed by the concentration of DNA polymerase. This dynamic exchange serves an important cellular function and is crucial for the repair of DNA damage. This is illustrated by the binding of different DNA polymerases to the β-subunit. The replicative DNA polymerase of E. coli (pol III), and a so-called lesion by-pass polymerase (pol IV) bind to the β-subunit on the same binding site, and because of the dimeric structure of β, they may do so concurrently. When a replication complex of pol III is stalled because of a lesion in the template strand, pol IV gains control over the primer-template site and polymerizes past the lesion. Although the fidelity of pol IV is significantly lower than that of pol III and this mechanism can result in mutagenesis, these damage-induced DNA polymerases (which besides pol IV also include pols II and V) ensure the continuation of DNA replication and increases the chance of long-term survival.
4. Tracking of clamps on DNA

DNA-processing enzymes that have to search for specific recognition sequences or structural sites (such as damaged DNA, nicks, single strand DNA regions, etc) generally interact weakly with long strands of DNA. As a consequence these enzymes are devoid of any sequence specificity, facilitating the sliding of the enzyme along a DNA strand. The clamp proteins provide a mobile platform on DNA, onto which one or multiple DNA-processing enzymes can bind. Through this binding, the enzymes can ‘scan’ the DNA for recognition sites more efficiently and the activity of these enzymes is normally enhanced by one or two orders of magnitude. In order to slide freely, clamp proteins also should not interact too strongly with DNA. Based on the X-Ray structures of several clamps, it was suggested that one or several layers of water molecules may be present between the negatively charged phosphate backbone of DNA and the protein to facilitate tracking of the proteins along DNA. The diameter of the central hole in clamp proteins is generally larger (34 Å for PCNA, 35 Å for gp45 and 38 Å for β subunit) than the minimal distance required to accommodate double stranded DNA (20-25 Å for the A and B forms of DNA). Nevertheless, direct contact between the (extended) amino acid side chains with the phosphate backbone may not be excluded for PCNA. Tinker et al. have shown experimentally that gp45, E. coli β subunit of polymerase III and PCNA indeed can slide freely over double stranded DNA segments by experiments in which clamps were cross-linked to photoactive nucleotides incorporated into the DNA substrates.93 Furthermore, Stukenberg et al. have investigated the dissociation of the β clamp from circular and linearized plasmids. It was found that the β clamp only dissociates from a DNA plasmid when this plasmid is cut by an endonuclease. The clamp can dissociate by sliding over blunt ends or short overhangs, but not over ends which are blocked by proteins or long single stranded regions. The T4 clamp is also not capable of sliding over 250 nt single stranded DNA regions. In contrast to the stability of the β clamp on DNA, gp45 is not stably bound and can dissociate anywhere from the template. Continuous loading of gp45 by its clamp loader (gp44/62) and ATP is therefore requisite for maintaining the DNA-bound species. This seems to be most effective when ‘roadblocks’ are introduced that prevent gp45 from sliding past a certain position on its template; in which case 80% protection from DNase I digestion was achieved, indicative of a closely packed array of gp45 proteins. The half-time of the β clamp on DNA varies from 5 minutes to 60 minutes, depending on the experimental conditions. The half-life time of the gp45-DNA complex could be extended from 1.2 min to 4.8 min by adding gp45 binding proteins (such as gp55), that track over DNA along with gp45. A similar effect was seen with the DNA polymerase (gp43), rendering the gp45 - gp43 complex on DNA stable enough to pass through a size-exclusion column as a complex. Fu et al. found that the half-life time of gp45 on DNA could be increased five-fold (from 30 s to 2.5 mins) in conditions of macromolecular crowding. The instability of gp45 on DNA can be attributed to its open subunit interface, since PCNA clamps, which are also trimeric, yield stable complexes on DNA (t½ = 24 mins).
5. T4 bacteriophage DNA replication

5.1 General

The process of T4 bacteriophage DNA replication has been a subject of several extensive reviews. Therefore, a comprehensive discussion of the T4 replication literature is beyond the scope of this introduction. This section will give a short overview of the assembly process of the different components of the T4 replisome, and will try to focus on the developments in the past 5 years. The replisome has 7 essential components: gp43 (DNA polymerase), gp45 (sliding clamp), gp44/62 (clamp loader), gp41 (DNA helicase), gp61 (primase) and gp59 (helicase loader). Gp32, the single stranded binding protein, was also proposed to be part of the T4 replisome because it engages in a number of crucial protein-protein interactions.

5.2 Initiation of replication

Replication initiation usually involves the unwinding an origin of replication (ori) region by DNA helicases, followed by the synthesis of a primer by a primase. The primer acts as a recognition site for the assembly of the replisome. Bacteriophages use a wide range of different mechanisms for the initiation of DNA replication (reviewed recently by Weigel and Seitz). For the T4 bacteriophage, two major pathways can be identified. During the early stages of infection, bidirectional DNA synthesis occurs at several of the 5 origins of replication on the T4 genome. This process requires the presence of E. coli RNA polymerase to synthesize a stably bound R-loop that serves as a primer (Figure 6A). The majority of DNA, however, is synthesized by a second mechanism that relies on the conversion of DNA recombination intermediates into replication forks and is promoted by the uvsW protein, a 3’→5’ helicase that removes the R-loop from the origins (Figure 6B). This recombination-dependent-DNA-replication (RDR) process requires the presence of gp32, dda helicase or the combination of gp41 and gp59, the uvsX protein (a recA analogue that promotes unidirectional branch migration), and uvsY, a protein that via several functions stimulates uvsX. Gp46/47, a protein with exonuclease activity is proposed to assist by generating single strands at the ends of the genome. These single stranded DNA region can then invade a homologous region on the T4 genome, after which the branched homologous recombination intermediates are captured by the replication proteins and converted into semi-conservative replication forks. The recombination intermediates are resolved at the end of the late phase by T4 endonuclease VII (gp49) to complete the production of several hundred copies of the linear T4 genome per infected cell.
Figure 6 – Initiation of DNA replication in the T4 bacteriophage. A. Origin-dependent DNA replication starts by the synthesis of a stably bound RNA primer by \textit{E. coli} RNA polymerase at an ori region, to form a R-loop. The primer is then used to assemble the replisome. B. The recombination dependent DNA replication (RDR) system in the T4 bacteriophage phage. In step one, a homologous single stranded DNA strand invades a double strand, via the action of uvsX, uvsY and gp32, to form a so-called D-loop. In step 2 a holoenzyme is assembled at the 3’ end of the invading DNA strand. Helicase action displaces the double strand to allow DNA synthesis (bold line represents newly synthesized DNA), resulting in the movement of the replisome and the melted region of the double stranded template (bubble migration). In step 3, a primosome is assembled on the lagging strand to facilitate the synthesis of Okazaki fragments. Figure B. is redrawn from Bleuit \textit{et al.}^{118}

5.3 Assembly of the T4 holoenzyme

The assembly of the T4 holoenzyme, which is the complex of the clamp (gp45) and DNA polymerase (gp43) bound to DNA and is able to synthesize DNA in a processive manner, has been studied extensively by a variety of experimental methods. These methods include rapid-quench techniques, ATPase activity studies, crosslinking experiments, and standard or stopped-flow fluorescence spectroscopy. The polymerase, gp43, preferably binds to primer-template sites on DNA,\textsuperscript{127} and because of the strong exonuclease activity of gp43,\textsuperscript{128} displays an extraordinary degree of fidelity during polymerization. This results in a frequency of misincorporation of only 1 in 10\textsuperscript{8} nucleotides \textit{in vivo}.\textsuperscript{129-131}

The polymerase accessory proteins gp45 and gp44/62 greatly stimulate the activity of gp43.\textsuperscript{132} The clamp is composed of three identical subunits,\textsuperscript{133,134} and thus has three equivalent subunit interfaces. In contrast to its crystal structure (see Figure 3C), the T4 clamp (gp45) is not a closed ring in solution.\textsuperscript{38} One of the three subunit interfaces is open in solution,\textsuperscript{39,40} with a distance between subunits of 28 Å.\textsuperscript{135} This is a unique characteristic of the clamp proteins of the T4 and T4-related bacteriophages like RB69, as the circular clamps of other organisms remain closed in solution. The clamp loader complex, a 4:1 complex of gp44
and gp61, hydrolyzes ATP to load gp45 onto DNA. The subunits of the clamp loader spontaneously assemble, and their interaction is so tight that they are usually copurified. The complex binds to ssDNA, but gp45 is most efficiently loaded onto primer-template sites, in which the primer can be either DNA or RNA and the template strand is preferably extended with >10 nucleotides at its 5’ position. Gp45 can also be loaded onto DNA at nicked sites, though less efficiently. The clamp loader acts as a catalyst, chaperoning gp45 onto DNA and towards interaction with gp43, and is not part of the final holoenzyme. Its ATPase activity is dependent on the presence of gp45 and a DNA template. There are four identical ATP binding sites, residing in the gp44 subunits. Gp44/62 binds to the ‘rough side’ of gp45 in a 1:1 complex with a K_D of 1-8 nM, although activity studies yielded a higher K_D of 142 nM. The latter studies also demonstrated that the addition of macromolecular crowding agents such as ethylene glycol and dextran polymers results in a strengthened interaction (K_D = 1.2 nM), while the overall activity of the gp44/62—gp45 complex remains unchanged. The binding of gp45 to gp44/62 in the presence of ATP is accompanied with a structural change in gp45 (Figure 7, for a more detailed discussion see Chapter 3). This conformational change is fuelled by the hydrolysis of 2 of the 4 ATP molecules, and results in a further the opening of the open subunit interface of gp45 from 28 to 38 Å, to allow dsDNA to slide through the open subunit interface. ATP hydrolysis is required as the non-hydrolyzable analogue ATPγS fails to promote the open gp45 conformation. When bound to DNA, the open subunit of gp45 closes because of electrostatic interactions and the remaining ATP molecules are hydrolyzed to rearrange gp45 into an out-of-plane conformation, necessary for proper interaction with gp43. Finally, gp44/62 is replaced by gp43, which binds to the same face of the clamp as gp44/62, and the open subunit interface further closes to 11 Å. The kinetic pathways of assembly and disassembly are fully characterized and proceed in 13 steps. The loading of the clamp onto DNA is the limiting step in this process.

![Figure 7](image_url)

**Figure 7** - Conformational changes in gp45 during the holoenzyme assembly process. From left to right: the solution structure of gp45, the conformation of gp45 in the complex with gp44/62, gp45 in the out-of-plane conformation on DNA, and the model of the final T4 holoenzyme. Figure courtesy of Trakselis and Benkovic.

In addition to the mechanism described above, three other routes towards functional T4 holoenzyme complexes have been identified recently (Figure 8). The clamp loader, gp44/62, is able to bind to primer-template sites on DNA before interacting with gp45.
was shown that ATP binding in the gp44 subunits leads to a conformational change that induces DNA binding. A DNA-bound gp44/62 then recruits a gp45 from solution and brings it into the out-of-plane conformation as described in the previous section. In contrast to the previous mechanism, only 1 ATP molecule is hydrolyzed per clamp loading event. Using small forked DNA templates covalently attached to a glass surface, and proteins labelled with different fluorophores, single molecule FRET experiments further identified two more pathways. The fluorophores were chosen such that with different filters, the individual proteins or their complex (through FRET) could be visualized. It was found that gp45 can bind dsDNA without the help of gp44/62. However, this non-specific binding does not lead to the formation of an active holoenzyme after the addition of gp43. Only when gp44/62 is added, the conformation of gp45 is corrected and the holoenzyme can be formed, demonstrating the chaperone function of gp44/62. In a fourth mechanism, gp43 can bind to the primer-template site first, and can then be converted into a functional holoenzyme by the action of a gp44/62—gp45 complex in the presence of ATP.

Figure 8 - T4 holoenzyme assembly pathways. I, the clamp loader (gp44/62) binds to and opens gp45 in solution and transfers it onto DNA. Gp43 replaces the gp44/62 to complete holoenzyme formation. II, gp44/62 first binds to the primer-template site and recruits a clamp from solution. III, gp45 binds to DNA in a non-specific way but is unable to interact with gp43. Only in the presence of gp44/62 functional holoenzyme complexes are formed. IV, gp43 binds first to the primer-template site, and is supplemented with a clamp via the gp44/62-gp45 complex in solution. Figure is from Smiley et al.

Gp45 can also independently interact with gp43 to form a holoenzyme. These complexes are generally short-lived but do lead to active replication complexes, especially in the presence of macromolecular crowding agents. Nevertheless, the action of the clamp
loader complex increases the efficiency of holoenzyme formation, and ensures optimal replication activity at low protein concentrations.

ATPase activity assays show that the clamp loader-clamp complex on DNA is not very stable.\textsuperscript{142} The ATP hydrolysis rate for the complex is high, indicating fast dissociation and association of the complex, in contrast to the more stable holoenzyme complex. This is probably due to the out-of-plane conformation of gp45,\textsuperscript{154} which is prone to dissociation. Moreover, in the absence of gp43 (or any other gp45-interacting protein), gp45 does not stay part of the complex but is released by the clamp loader. Alley \textit{et al.} noted that during stopped-flow experiments, the double stranded region of a small forked DNA substrate, blocked at the end by streptavidin, was fully covered with 4 gp45 proteins.\textsuperscript{152} Likewise, Fu \textit{et al.} found a high protection against DNase I degradation with a similar substrate.\textsuperscript{96} Probably only arrays of gp45 proteins are stable enough to be observed with conventional techniques. Gogol \textit{et al.} managed to visualize these arrays, as well as individual gp45 complexes, with cryo electron microscopy (Figure 9).\textsuperscript{162} Rapid freezing of the sample and low temperatures were needed to stabilize the complexes (originally termed ‘hash’ marks) on DNA. Different DNA substrates were incubated with gp44/62, gp45 and ATP to produce the arrays. No proteins could be visualized on a supercoiled DNA plasmid due to the absence of a clamp loading site. The presence of gp32 (the ssDNA binding protein) greatly stimulated the appearance of the structures, but after removing ATP or magnesium from the solution the structures disappeared within minutes. Remarkably, the addition of gp43 to these complexes resulted in the aggregation of DNA and prevented the observation of the holoenzyme complex.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure9.png}
\caption{Electron microscopy images of gp45 structures on DNA. A. Gp45 arrays and isolate gp45 proteins on plasmid DNA with a single stranded region of ~300 bases. B. Top, close-up from of a gp45-array (top) on RF-M13 DNA. Bottom, isolated gp45 proteins on RF-M13 DNA. Black arrowheads in A. and B. indicate isolated gp45 proteins. Images are from Gogol \textit{et al.}.\textsuperscript{162}}
\end{figure}

The T4 holoenzyme can synthesize DNA on primed ssDNA substrates, provided there are no secondary structures that block its path. Hacker \textit{et al.} showed that while the holoenzyme is stalled for several minutes when nucleotides are omitted from the reaction mixture,\textsuperscript{163} it rapidly dissociates when it encounters a DNA hairpin.\textsuperscript{164} Thus, the holoenzyme
is very processive, but senses the presence of an Okazaki fragment and rapidly dissociates to be reused for the next Okazaki fragment. Carver et al. confirmed these results, but also demonstrated that a forked DNA substrate (rather than a primer) can be displaced to allow the DNA synthesis reaction to proceed. The recycling of the holoenzyme at the lagging strand should be very efficient, since Okazaki fragments are synthesized in a few seconds. The polymerase dissociates from the lagging strand and possibly leaves the clamp behind. There is no evidence that the clamp remains associated with the DNA. Possibly, the clamp is destabilized by the lack of protein binding to its open subunit interface. Soumillion et al. demonstrated that the dissociation of gp45 trimers proceeds via subunit disassembly process, rather than the sliding of DNA through the open subunit interface.

5.4 Interactions within the T4 holoenzyme

The T4 DNA polymerase gp43 is a distributive enzyme (k\text{off} = 6-8 s^{-1}) and binds to primer-template DNA substrates with K_D = 40-70 nM. The T4 holoenzyme in contrast is very processive, which is reflected in the high binding constant of gp43 for ssDNA when associated with gp45 (K_D = 20-30 pM). Furthermore, the k\text{off} rate has been reduced several orders of magnitude to 0.01 s^{-1}. With a half life of 7 minutes, the processivity of the T4 holoenzyme is over 100,000 nucleotides. With in vitro replication speeds of 500 nt/s, a leading strand holoenzyme can in principle replicate the whole T4 genome without dissociating. Like with PCNA interacting proteins, the binding of gp43 to gp45 is mediated via an extruding tail found at the C-terminus of the polymerase, and no holoenzyme can be formed with polymerases lacking the six C-terminal amino acids. Alley et al. demonstrated that a peptide analogue of the gp43 C-terminus cross-linked to the open subunit interface of gp45 at high concentrations, and the binding constant of this peptide to the clamp is K_D = 7.7 \mu M. From these interactions a replisome model was derived, which refined a previous model from the Steitz group based on the crystal structure of the peptide with the clamp. In the latter model the C-terminus was bound to the interdomain connecting loop rather than the open subunit interface of the clamp. The binding constant of the peptide probably reflects the binding to this interdomain connecting loop because the fluorescence titrations were done at low gp45 concentration, in which most of the gp45 trimers are disassembled into dimers and monomers (the dissociation was found to be best described by a cooperative model with the product of dissociation constants k_1k_2 = 0.08-0.21 \mu M^2, see Chapter 5 for more details). An increase from K_D = 480 nM to low nanomolar K_D when the two proteins interact in solution or on DNA, respectively. It was suggested that the rearrangement of gp43 from the interdomain connecting loop (micromolar binding) to the open subunit interface (nanomolar
binding) would allow gp43 to eject gp44/62 from its complex with gp45 to complete the formation of the holoenzyme complex.\textsuperscript{53} Interactions with DNA would then further stabilize the complex to picomolar $K_D$s. Since the exact binding constant of the peptide to the open subunit interface of gp45 has not yet been determined, it remains unclear whether this rearrangement actually takes place.

5.5 Primosome assembly

The DNA helicase gp41 is essential for optimal replisome activity,\textsuperscript{173} as the absence of gp41 diminishes the rate of leading strand DNA synthesis (from $\sim 400 \text{ nt/s}$ to 10 nt/s on a rolling circle template).\textsuperscript{174,175} In the presence of ATP or GTP, gp41 forms a hexameric toroid structure around the lagging strand and translocates in a 5' $\rightarrow$ 3' direction while processively unwinding the double stranded DNA ahead of the replication fork, driven by the hydrolysis of the nucleotides.\textsuperscript{176} Dilution experiments indicated that the half life of gp41 in the replication fork is 11 minutes, comparable with the half life time of the holoenzyme.\textsuperscript{175} Furthermore, gp59 stimulates the assembly of gp41 on the lagging strand, acting as a helicase loader protein.\textsuperscript{177-179} Gp59 requires single stranded template regions covered with gp32. The third component of the primosome is gp61, the primase that synthesized short RNA strands at regular intervals on the lagging strand for the initiation of Okazaki fragment synthesis. Gp41 can interact with gp43 in the presence of a macromolecular crowding agent as these two proteins formed a complex able to synthesize DNA at physiological rates,\textsuperscript{180} but their interaction is too weak to be observed by analytical ultracentrifugation, and is therefore likely not to play an important role \textit{in vivo}.\textsuperscript{181} Rather, the coupling of the holoenzyme and the primosome is mediated by gp59.\textsuperscript{182}

The exact mechanism of helicase loading has been elucidated recently and involves a four-step mechanism (Figure 10). First, gp59 interacts with gp32-covered ssDNA regions mediated via an interaction with the C-terminal domain of gp32. Alternatively, gp59 may also bind to other DNA structures such as forks and cruciforms.\textsuperscript{183,184} In the second step, a putative helicase loading complex is formed, consisting of a condensed complex of six gp59 and six gp32 proteins around which ssDNA is spooled. This complex translocates along the ssDNA strand to a suitable helicase loading position.\textsuperscript{185} Then, gp41 is recruited to this complex. Maximal activity is achieved with a 1:1 ratio of gp59 to gp41, indicating that gp59 indeed functions as a hexamer.\textsuperscript{186} However, the formation of a helicase loading complex, as proposed by Ma \textit{et al.}\textsuperscript{185}, remains elusive because efficient gp41 loading is achieved with a relatively short 45 nt ssDNA region, just enough to accommodate six gp32 proteins (one gp32 binds to 7 bases).\textsuperscript{187,188} Moreover, cross-linking experiments showed that gp59 can lower the affinity of gp32 for ssDNA, and may displace gp32 before forming a hexamer.\textsuperscript{189,190} With the help of single molecule FRET experiments, similar to those which identified multiple pathways of clamp loading, Zhang \textit{et al.} demonstrated that gp59 dissociates from the complex after the formation and loading of the gp41 hexamer.\textsuperscript{187} The primase gp61 is recruited to the gp41 hexamer to complete primosome assembly. The gp41:gp61
stochiometry is 1:1, indicating that also gp61 forms a hexamer in the final primosome, as is does alone on ssDNA. This is in contrast to earlier studies that found a gp41:gp61 ratio of 6:1. The structures of gp41 and gp61 in the presence of ssDNA have also been visualized by electron microscopy, and indeed showed asymmetric hexagonal ring-shaped structures for both proteins. In addition, the gp41 hexamer seems to exist in an open and closed conformation. Another function of gp59 is to inhibit DNA synthesis in the absence of gp41 by forming a complex with gp43 and displacing gp43 from the replication fork. Thus, the premature elongation of leading DNA strands is prevented before a fully functional replisome is formed. If gp41 is added to an inhibited complex, DNA synthesis is restored. The gp59 dependent inhibition of DNA synthesis enables the coupling between leading and lagging strand synthesis.

A.

B.

Figure 10 – Primosome assembly pathways. A. Mechanism proposed by Ma et al. In step 1, a complex is formed on the lagging strand of gp32 and gp59. This complex forms a helicase loading complex (step 2) and translocates to a suitable helicase loading position (step 3). Then the helicase (gp41) hexamer is assembled, which in turn recruits the primase (gp61). B. Mechanism proposed by Zhang et al. The helicase can also assemble without gp59, but not so efficiently. In gp59-mediated helicase assembly, gp59 first interacts with gp32 and forms a hexamer on the lagging strand. This complex loads gp41 and dissociates. The gp41 hexamer then recruits the primase. Images are from Ma and Zhang, respectively.
Besides gp41 and uvsW helicases, the T4 bacteriophage also encodes another the dda helicase that catalyzes the unwinding of dsDNA in the 5’ → 3’ direction.202-205 This non-essential helicase (e.g. phages lacking the dda gene are viable) works via a distributive mechanism, in contrast to the highly processive gp41 hexamer. The dda helicase has specialized but non-essential functions; it was proposed to facilitate the bypass of RNA transcription complexes by the replisome,177,206 and is involved in the opening of DNA replication origins.207 Dda may also play a role in displacing proteins from DNA ahead of the replisome, as discussed below.

The assembly of dda on ssDNA was investigated recently by Ma et al.185 While dda binds to ssDNA autonomously, the binding of dda to DNA seems to be enhanced by gp32. In the dda-gp32 complex, gp32 does not bind to DNA but supposedly induces a more active conformation of dda. Byrd and Raney demonstrated that while a dda monomer translocates along ssDNA, it exerts a force high enough to displace streptavidin from biotinylated oligonucleotides.208 The force is dependent on the number of dda proteins bound to the oligonucleotides, but the individual dda proteins do not bind cooperatively. Thus, the individual helicase activity of several dda proteins in tandem results in this enhancement, like a freight train pulled by several locomotives.209 The dda protein probably does not play an important role in the DNA replication process, given that it is non-essential to the T4 phage. Moreover, dda apparently does not interact with the primase to substitute for gp41, and therefore its action could only be beneficial for leading strand synthesis.

5.6 Role of the single strand DNA binding protein (gp32)

As is apparent from the previous sections, gp32 surpasses its role as a simple single stranded binding protein, and in many ways is a crucial protein within the T4 replisome.32,210 Gp32 is composed of three functional domains: the A and B domains, and the core domain. The A domain interacts with a variety of replication and recombination proteins, including gp59, dda helicase, gp43, gp61.211-214 The B domain is important for gp32-gp32 interactions and cooperative ssDNA binding, and the DNA binding ability resides in the core domain. The interactions of gp32 with ssDNA and dsDNA has been studied recently with biophysical techniques such as DNA unwinding experiments with optical tweezers, and single molecule AFM spectroscopy experiments,215-217 from which a detailed model could be derived.218 The C-terminal domain is complexed to the core domain via electrostatic interactions. This provides a kinetic block that prevents the destabilization of dsDNA at the replication fork, which is perquisite to the gp41 helicase. At high salt conditions this interaction is broken. In vivo, this would represent the state in which the C-terminal domain interacts with other proteins. The ‘unlocking’ of the DNA binding domain results in a very strong cooperative binding of gp32 to ssDNA. In addition, it was shown that gp32 can bind weakly to dsDNA, which would allow it to easily find ssDNA sections via a translocation mechanism. The interaction of replication proteins with the C-terminus of gp32 allows fine tuning of the DNA melting behaviour.
The binding of gp32 to ssDNA sections lowers the non-specific binding of replication proteins, effectively directing these proteins to the replication fork. Furthermore, it prevents the formation of secondary DNA structures, which would have a detrimental effect on replication speeds, and protects DNA from nucleases. Direct interactions with replication proteins, such as gp44/62, increase the binding of these proteins to the replication fork. The presence of gp32 stabilizes the lagging strand holoenzyme, possibly via interactions with gp43. The fidelity of gp43 is also improved, a result of improved substrate specificity rather than changes in exonuclease activity. Recent evidence suggests, however, that sometimes the presence of gp45 and gp32 lower the fidelity of the RB69 polymerase in a sequence specific manner. Finally, Gangisetty et al. demonstrated that the interaction of RNase H (the FEN1 analogue of T4) depends more on the interaction of RNase H with gp32 than with gp45, although RNase H has a terminus domain able to interact with the latter.

5.7 The T4 replisome

Physical cross-linking experiments and footprinting reactions have provided detailed information about the relative position of the individual proteins within the replisome. Ishmael et al. demonstrated that the two holoenzymes are linked to each other via their finger domains, and that interactions between the primosome and lagging strand holoenzyme are mediated by the gp59 protein. This would be in contrast to the FRET experiments described above, which showed that gp59 does not remain bound to the primosome.

Using cleverly designed small circular DNA (minicircle) substrates that allow the rates of DNA synthesis on the leading and lagging strands to be determined independently, Salinas et al. showed that the formation of a gp32-dependent gp43 homodimer results in the coupling of leading and lagging strand synthesis. Via dilution experiments on primed M13 substrates, Kadyrov et al. found that the coupling of leading and lagging strand synthesis also depends on gp44/62 and gp45, indicating that the holoenzyme has to be reassembled for every Okazaki fragment. Low concentrations of these proteins also results in the formation of longer Okazaki fragments, since the reassembly of the holoenzyme at the lagging strand is not efficient enough to keep up with leading strand synthesis. When lagging strand synthesis on minicircle substrates is blocked, synthesis on the leading strand is unaffected. In contrast, the rate of lagging strand synthesis does depend on the rate of leading strand synthesis. Efficient DNA synthesis only takes place when the concentration of the individual replisome components is in a certain range that matches the concentrations the proteins in vivo. Finally, Trakselis et al. demonstrated that the addition of inactive protein mutants to a functional replisome lead to the dissociation of the primase (gp61) from the replisome and thus gp61 needs to be recruited at the start of every Okazaki fragment, although a mechanism in which a number of gp61 proteins remain bound to the replisome would provide a much more efficient reassembly pathway. The addition of inactive clamp loader complexes stops synthesis at the leading strand after 8 minutes (correlating with the
half life of the holoenzyme), and synthesis at the lagging strand after 3 minutes. Furthermore, few replisomes were active throughout the time course of the experiments, implying that under the experimental conditions used, a large fraction of replisomes is stalled. More detailed analysis of the events at the lagging strand showed that the lagging strand polymerase dissociates after completing an Okazaki fragment. The dissociation may be initiated by a collision with a primer, but also by the loading of a clamp at a newly synthesized primer while replication still takes place. This signalling mechanism could account for the large stretches of ssDNA regions found at the lagging strand by electron microscopy studies.

More than 15 years of research into the structure of the T4 replisome has resulted in the model which is depicted in Figure 11. The model represents a static picture of the replisome; as discussed above gp45 and gp43 dissociate after the synthesis of an Okazaki fragment, and gp61 probably dissociates after the synthesis of a primer. Gp44/62 assists the formation of a holoenzyme by loading gp45 and guiding it towards interaction with gp43. The helicase loader can remain bound within the replisome, where it mediates interactions between the leading strand holoenzyme and the primosome. Finally, other proteins may be bound to the clamp, such as additional polymerases that can replace the synthesizing gp43 in the dynamic processivity process as demonstrated by Yang. Recently, T4 α-glucosyltransferase, an enzyme that labels DNA with sugars providing protection from nucleases and to distinguish T4 DNA from host DNA, was also found to be associated with gp45. Nevertheless, its amino acid sequence lack a classical gp45 interaction sequence. This is also the case with the T4 ligase that also is associated with the replisome.

The DNA in the replisome is also highly dynamic, as the replication fork moves with speeds of 250 nt/s. DNA synthesis on the lagging strand occurs in the opposite direction. This implies that the lagging strand between the lagging strand holoenzyme and the replisome will form a loop that is expanding with a speed of ~500 nt/s. If this loop becomes too large, it collapses, most likely after finishing an Okazaki fragment during the switching of a polymerase, since at that stage the coupling between holoenzymes is lost. This process of growing and collapsing lagging strand loops is referred to as the trombone model. Chastain et al. have recently examined replication intermediates of a M13-based rolling-cycle system with electron microscopy. Their results implied that in more than 50 percent of the investigated structures indeed one or multiple loops were present (Figure 11). The intermediates without loops may represent complexes that just finished Okazaki fragment synthesis. The single stranded regions, covered with gp32 and possibly also gp59, were not extended but appeared as compact, ‘bobbin-like’ structures. The protein complexes found at the replication fork were subjected to further investigation by using biotinylated replication proteins, which allowed quantification of those proteins through the addition of streptavidin-DNA complexes. Two polymerases were generally found at the replication fork, but structures with two loops (6% of total) may also contain three polymerases (31%). Furthermore, stalled polymerases were not associated with gp41 helicases, and gp59 was
found at the replication fork after gp41 has been loaded, consistent with the results of Ishmael et al.182

The T4 replisome (Figure 11A) has offered a lot of information about the process of DNA replication. Despite its own unique characteristics, many aspects of its assembly pathway and action can be generalized to eukaryotic DNA replication systems. The T4 system is able to synthesize DNA at high speeds through coordination of all the enzymatic functions at the replication fork.

5.8 T4 late transcription

The process of transcribing DNA to RNA usually is reviewed separately from that of DNA replication. For the T4 phage, however, these processes are more intertwined as they take place simultaneously during the ‘late’ phase of the T4 multiplication cycle.237 In the late phase the components of the T4 viral package are transcribed, and DNA is packaged in the viral capsid.238 In fact, late transcription depends on the polymerase accessory proteins gp45 and gp44/62.140 Also in this case, macromolecular crowding can replace gp44/62 and ATP hydrolysis to load gp45 onto DNA.239 Experiments by the Geiduschek group on transcriptional activation by gp45 have yielded valuable information on the tracking of gp45 on DNA (as discussed in Section 4 of this chapter).56,93,95,98 DNA transcription depends on the E. coli RNA polymerase, and starts from simple TATA-boxes (TATAAATA, of which ~40 exists in the T4 genome). It was shown that gp45 was not loaded at the promoter site itself,
implicating that gp45 has to track from a distal loading site to the promoter site.\textsuperscript{50} If a nick is used as such a site, the nick has to be in the non-transcribed strand to load gp45 in the correct orientation onto DNA.\textsuperscript{95} Two other T4 proteins, gp33 and gp55, are involved in transcriptional regulation as well.\textsuperscript{240} The \(\sigma^{70}\) subunit of the RNA polymerase complex is replaced by gp55 in order to recognize the correct promoter sequence.\textsuperscript{241,242} Gp33 is a transcriptional co-activator, and suppresses transcription by blocking DNA binding by the RNA polymerase until gp45 is part of the transcription complex. In the final transcription complex, gp45 was found at the upstream end of the complex in the vicinity of gp33.\textsuperscript{93} Both gp33 as gp55 have a C-terminus that can interact with gp45\textsuperscript{69,243}, from which it was proposed that both proteins would bind to gp45 simultaneously.\textsuperscript{243-245} Whether gp55 and gp33 both co-track on DNA while bound to gp45 remains unclear; the interaction of gp55 with gp45 would decrease the time needed for the former protein to find a promoter sequence.\textsuperscript{98} Kolesky et al. recently investigated the kinetics of the formation of transcription complexes and found that gp45 activates the opening of a promoter 320 times.\textsuperscript{244} Furthermore, Nechaev et al. located the interaction site of gp33 with RNA polymerase.\textsuperscript{246} A model was proposed in which gp33, gp45 and gp55 collectively mimic the structure of the \(\sigma^{70}\) subunit of a functional \textit{E. coli} RNA polymerase complex (Figure 12). Gp45 bridges the two proteins, so both gp33 and gp55 are bound to gp45 in the RNA polymerase complex, and gp45 would probably orient these proteins in a specific way. This would explain why all three proteins are required for efficient T4 late transcription.

![Figure 12](image)  

\textbf{Figure 12} – Transcription activation in \textit{E. coli} and T4-infected \textit{E. coli}. \(\sigma^{70}\) subunit on an opened promoter. The kink in dsDNA facilitates the opening of the promoter site. The transcription start site (+1) and promoter recognition sites (-10, -10E, -35) are shown on the DNA. A. Cartoon of a functional RNA polymerase complex, redirected by the three T4-encoded proteins gp55, gp45 and gp33 to transcribe T4 late genes. The T4 proteins collectively mimic the \(\sigma^{70}\) subunit found in wt RNA polymerase (A). The red triangles represent the gp45-interacting termini. Images are from Nechaev et al.\textsuperscript{246}

Concurrent DNA replication and transcription also implies that the replisome and transcription complex may collide during the late phase. Indeed, the speed of a replisome is 10-15 times higher than that of a transcription complex\textsuperscript{247}, and if the first has to tail behind the latter, the efficiency of DNA replication would be compromised. In a series of experiments, Liu and Alberts assessed the fate of the transcription complex after a collision...
with a replication complex. If both complexes move in the same direction, the replication fork can pass a transcription complex, while the latter remains bound to the DNA template and is fully able to continue RNA synthesis.\textsuperscript{248} In the original experiments a stalled RNA transcription complex was used, but the findings were later confirmed for transcribing RNA transcription complexes as well.\textsuperscript{249} In contrast, a head-on collision halts the replication fork for many minutes in the absence of gp41. With gp41 the two complexes can pass each other with a delay of only a few seconds. Intriguingly, the RNA polymerase switched from its original template strand to the newly synthesized strand.\textsuperscript{250} The exact mechanism of these processes remain unclear; possibly the dissociation of a lagging strand holoenzyme could facilitate the bypass.\textsuperscript{228}

6. Comparison with other replication complexes

The similarities and differences in clamp loading mechanisms among the most studied DNA replication systems (T4, \textit{E. coli} and yeast) has been reviewed several times.\textsuperscript{251-255} The clamp loading mechanism of the T4 phage is probably the best described in the literature. It is easy to envisage the clamp to be loaded onto DNA, since it has an open subunit interface. In contrast, the $\beta$-subunit of \textit{E. coli} polymerase III forms a closed circle, composed of two identical subunits. It is loaded onto DNA by the $\gamma$-complex, consisting of 7 subunits ($\gamma\tau_2\delta\delta'\psi$), in an ATP dependent fashion. The ATPase activity resides in the $\tau$ and $\gamma$ subunits. All the ATP binding sites need to be filled for the $\gamma$-complex to interact with $\beta$, and to load it onto DNA.\textsuperscript{257} Hydrolysis of ATP is necessary to close $\beta$ around DNA and for the clamp loader to dissociate, in contrast to the T4 clamp loader where ATP hydrolysis is used to open the clamp interface.\textsuperscript{258} The $\delta$ subunit binds to a hydrophobic cleft on $\beta$, which results in the opening of the clamp due to spring tension in the clamp.\textsuperscript{259} The $\delta$ subunit can also unload the clamp from DNA.\textsuperscript{260} Three ATP molecules are hydrolyzed during one clamp loading cycle.\textsuperscript{261} The $\tau$ subunit plays an important role as a processivity switch. When the holoenzyme encounters a duplex DNA molecule, $\tau$ binds to the C-terminal residues of the polymerase and sequesters it from the $\beta$-subunit.\textsuperscript{262} This interaction coordinates the recycling of polymerases at the lagging strand. The holoenzyme dissociates and binds another clamp on a new primer, while leaving the old $\beta$ behind.\textsuperscript{263} The replisome itself consists of an asymmetric dimer of holoenzymes with distinct leading and lagging strand polymerases.\textsuperscript{264} A model of the \textit{E. coli} replisome is shown in Figure 13.
Figure 13 – Schematic model of the *E. coli* replication fork. See text for details. The whole polymerase complex consists of 10 different subunits (including $\beta$ and the $\gamma$-complex), while the core polymerase consists of the subunits $\alpha$ (polymerase activity), $\epsilon$ (proofreading) and $\theta$. Image is from Lodish *et al.*

Eukaryotic PCNAs resemble gp45 in the number of subunits, but are closed trimers in solution. Recently, Zhuang *et al.* showed by FRET experiments that PCNA, like gp45, undergoes an out-of-plane conformation during clamp loading. Moreover, two important studies provided new information about the structure and action of clamp loaders. A crystal structure of the clamp loader (RFC) of *S. cerevisiae* with PCNA showed the interaction of RFC with PCNA, although the clamp was not opened in this structure (Figure 13A). The 5 subunits of RFC were arranged in a spiral conformation on top of PCNA (resembling a horseshoe on a donut). The spiral’s pitch matches that of the DNA double helix, providing important clues about the clamp loading process. Furthermore, if PCNA would interact with the spiralled RFC, it would be brought into the out-of-plane conformation as identified with FRET. A second study reconstituted the shape of an archaeal *Pyrococcus furiosus* RFC complex via electron microscopy. The complex was first assigned to be a structure in which the subunits of the clamp loader lie flat on a closed PCNA trimer. Later experiments, which resolved the structure of the complex to a 12 Å resolution, showed that the first structure was not correctly assigned, and that it could be fitted better by a complex in which the PCNA is open. The PCNA, again in out-of-plane conformation, interacts strongly with the subunits of RFC, which are in a ‘spring washer-shaped ring.’ Although initially thought otherwise, these two studies possibly represent a similar intermediate in the clamp loader process. The crystal structure of the *E. coli* $\gamma$-complex (a minimal analogue $\gamma_3\delta\delta'$ was used) has also been solved, albeit without the $\beta$-subunit, in the absence and presence of ATP$_{\gamma}$S. Both structures look similar, but maybe due to the absence of the $\beta$-subunit and a DNA template, correspond to a stable inactive form of the complex. Nevertheless, the complexes still show structural similarities to RFC. Indeed, a spiral arrangement of subunits that can accommodate a DNA strand in the inner channel was demonstrated by fluorescence assays of mutated $\gamma$-complexes.
Figure 14 – A. Crystal structure of yeast clamp loader RFC in complex with the PCNA clamp (1AXC). The 5 RFC subunits are coloured in turns and spirals out of the paper towards the reader. B. Real and computerized images of the Pfu clamp—clamp loader in different orientations. Scale bar is 100 Å. C. Surface representations from the Pfu clamp—clamp loader complex, as seen side-on and head-on with respect to PCNA. Images are from Miyata et al. The T7 replication system (Figure 15) consists of four proteins; gp2.5, the ssDNA binding protein, gp4, a protein with combined helicase and primase activities, gp5, the DNA polymerase, and the processivity factor thioredoxin which is encoded by E. coli. Despite the lack of a clamp protein (and helicase loader), many aspects of T7 replication are similar to that of T4, such as coordinated leading and lagging strand synthesis on synthetic DNA substrates, and the formation of replication loops that can be visualized with electron microscopy.

Figure 15 – Model of the T7 replication fork. See text for details. The primase is made up from a zinc binding domain (ZBD) and the RNA polymerase domain (RPD). Trx denotes E. coli thioredoxin. Image is from Lee et al.

By virtue of the small number of proteins involved the T7 replication system has been used in a number of biophysical experiments during the last years. A large ssDNA molecule was appended with a magnetic bead, and was replicated by T7 polymerase. During the experiment ssDNA was converted to a dsDNA by the polymerase, a process that
was monitored by the difference in elasticity of single and double stranded DNA reflected in the forces that were exerted on the magnetic bead in a strong magnet field. Under low stretching forces, the activity of the DNA polymerase was similar to those found in bulk experiments (210 vs 300 nt/s). Interestingly, the polymerase frequently paused at different sites on the DNA molecule, possibly being blocked by local hairpin formation. The efficiency of replication was dependent on the stretching force applied to the ssDNA molecule. At forces higher than 20 pN no replication takes place, indicating that multiple bases on the template strand are involved in the rate-limiting step of each turnover. This was explained by an induced fit mechanism that allows error-detection by the enzyme during replication.

The helicase of T7, gp4,281,282 also forms a hexameric ring in the presence of nucleotides, which is able to unidirectionally translocate along DNA at a speed of 130 bases per second.283 Its activity was investigated with a fluorescence-based stopped flow experiment.284 Remarkably, gp4 had poor helicase activity in these experiments (15 nt/s while replication speed is 300 nt/s). The processivity of gp4 is low due to ring opening followed by dissociation. The same group later found that both polymerase and helicase are required for efficient replication.285 The polymerase by itself is unable to perform strand displacement synthesis, while the activity of the helicase is stimulated ~8-fold by the presence of the polymerase. Thus, the trapping of the unwound ssDNA by the polymerase pushes the helicase forward through dsDNA. In another single molecule experiment, the primase activity of gp4 was monitored during replication.278 A long DNA strand, appended with a bead for imaging purposes, was stretched by flow forces. Again, the difference in ssDNA and dsDNA coiling is correlated with the progress of DNA replication. The rate of primer synthesis and recycling of polymerases on the lagging strand should transiently pause the rate of leading strand synthesis. Indeed, such pauses were observed with these experiments, and hence the primase was proposed to act as a molecular brake, facilitating the assembly of the polymerase on the lagging strand without losing coordination between leading and lagging strand synthesis.

7. **Other processive and toroidal enzymes**

The clamp proteins provide a clear example of the advantages of processivity in catalytic processes. They are unique in the sense that they do not have an enzymatic activity of their own, but they are by no means the only ring-shaped proteins involved in processive catalysis. Indeed, the hexameric helicases, discussed above, are generally also very processive.286 Other toroidal proteins include topoisomerases, transcription termination proteins, certain proteases and some structural proteins in phages. The toroidal λ-exonuclease is an exceptional well example of processive toroidal catalysis (Figure 16); it is funnel shaped, and able to efficiently hydrolyze one of two strands of dsDNA templates in the 5′ → 3′ direction in an ATP-independent process.287 One end is large enough to accommodate dsDNA, while only ssDNA fits through the other end. The enzyme binds
preferable to 5’-recessed DNA, but degrades other DNA substrates with bunt or 5’ overhang ends equally well. Interaction with the 5’ phosphate of the DNA strand is important for processive activity, indicating that the subsequent hydrolytic steps pull the enzyme in the right direction over its substrate. The enzymatic action of λ-exonuclease has also been investigated at the single-molecule level, which showed an activity of 1000 hydrolyzed nucleotides per second and a high processivity. Similar experiments, which measured the position of flow-stretched DNA-appended beads as described in the previous section, demonstrated that the activity of λ-exonuclease is dependent on the base content, consistent with a mechanism in which the melting of a base pair is the rate-limited step. Moreover, the enzyme pauses at certain sequences, which may have biological relevance in the inhibition of exonuclease activity at the ends of λ-DNA.

A. Figure 16 – λ-Exonuclease as processive enzyme. A. Crystal structure of λ exonuclease (1AVQ). B. Representation of a λ-exonuclease on a DNA strand, while digesting one of the two strands in the 5’ → 3’ direction.

It is striking that a whole range of DNA and RNA processing enzymes have a toroidal topology. It is not clear whether the shape of DNA caused the appearance of toroidal enzymes in the course of evolution, or whether it is the toroidal shape itself that has an intrinsic high activity. In any case, the very existence of a myriad of toroidal enzymes proves their use for Nature.
8. Synthetic rotaxane catalysts as models for processive enzymes

Despite their ubiquitous appearance in Nature, the potential of processive catalysts in synthetic systems has not been recognized widely. The only synthetic catalysts known to date that can form a rotaxane structure with a linear polymer and can mimic the action of processive enzymes found in Nature have been developed in the Nolte and Rowan groups (Radboud University Nijmegen). By combining a glycouril clip and a porphyrin catalyst, the processive catalyst shown in Figure 17A was obtained.\textsuperscript{296-298} Several small molecules, such as viologen (quarternized 4,4′-dipyridines) or pyridine derivatives can bind inside the cavity with high association constants ($K_A = 10^5$ - $10^6$ M$^{-1}$). Although the cavity is quite constrained in size, the ligands can bind with different binding geometries. Dimethyl viologen or small pyridine ligands bind perpendicular to the porphyrin plane (Figure 16B), with favourable interactions between its $\pi$-surface and the side walls of the clip. Ethanol viologen, however, binds with the $\pi$-surface of one of its pyridine functions parallel to the porphyrin plane, and has extra interactions such as a hydrogen bond with the glycouril’s carbonyl functions, as well as electrostatic interactions with the crown ether oxygens. A crystal structure of the latter complex was obtained (Figure 17C). In addition to its ability to serve as a host for small molecules, the porphyrin clip can also perform catalytic reactions when appropriate metals are bound within the porphyrin. Manganese porphyrins are well known for their catalytic oxidation behaviour.\textsuperscript{299} By using pyridine ligands that can bind to the manganese centre either in or outside the cavity, the catalytic activity of the porphyrin clip can be enhanced, and the catalytic reaction can be directed either on the inside or outside of the cavity.\textsuperscript{300,301} Outside binding is achieved by using ligands that are too large to fit into the cavity; this prevents the formation of $\mu$-oxo-dimers, which significantly elongates the life time of the catalyst.

![Figure 17](image_url)

**Figure 17** – The porphyrin clip. A. Structure of the porphyrin clip. B. Crystal structure of a zinc porphyrin clip complexed with pyridine in its cavity. C. Crystal structure of an methyl ester functionalized free base porphyrin clip complexed with ethanol viologen.
In order to mimic Nature’s processive enzymes, the manganese porphyrin clip was applied for the epoxidation of double bond containing polymers (such as polybutadiene). The polymers were appended with a viologen ‘station’ to facilitate the binding of the catalyst onto the polymer. In the presence of an oxygen donor (PhIO), the double bonds of the polymer were converted to epoxides (Figure 18).\textsuperscript{302} Evidence that this conversion takes place within the cavity comes from the trans/cis ratio of the resulting epoxide, which is significantly lower in the case of a reference porphyrin catalyst (trans/cis ratio is 4 vs. 0.3). Later, olefin metathesis was explored as an easy method to synthesize porphyrin-clip containing rotaxanes and catenanes of polybutadiene polymers.\textsuperscript{303} More recently, ruthenium-porphyrin clips have further expanded the scope of catalytic reactions that can be carried out by the porphyrin clip. The ruthenium catalysts can dimerize and polymerize diazo-esters, which can be used for the auto-catalyzed formation of rotaxanes,\textsuperscript{304} much like the catalytically self-threading cucurbituril rotaxanes reported by Tuncel and Steinke,\textsuperscript{305-307} or the catalytic “click” rotaxanes reported by Leigh \textit{et al.}\textsuperscript{308}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure18}
\caption{Schematic representation of the porphyrin clip-mediated epoxidation reaction of linear polybutadiene. As the clip slides over the polymer, double bonds in the polymer are converted into epoxides.}
\end{figure}

In a complementary approach, the Nolte group has pursued the synthesis of porphyrin boxes as toroidal catalysts for the oxidation of polymer and DNA substrates (Figure 19). These boxes are made from 4 porphyins, each of which occupying one side of a square shaped molecule (Figure 19). Molecular modelling studies showed that these boxes can bind dsDNA in their cavities. The boxes could be synthesized by metal directed self-assembly (MDSA), or were made by sealing alkenes by olefin metathesis in the presence of a template (dynamic combinatorial chemistry approach, DCC).\textsuperscript{309}
A. Synthesis of porphyrin boxes via MDSA or DCC, resulting in inert metal (Pt, Pd) complexes or olefins as linkers, respectively. B. Modelled macrocycle on a DNA template, as seen along and perpendicular to the helix axis, respectively.

Figure 19 - A. Synthesis of porphyrin boxes via MDSA or DCC, resulting in inert metal (Pt, Pd) complexes or olefins as linkers, respectively. B. Modelled macrocycle on a DNA template, as seen along and perpendicular to the helix axis, respectively.

9. References

20 M. Mok and K.J. Marians, The Escherichia coli preprimosome and DNA B helicase can form replication forks that move at the same rate, J. Biol. Chem. 262 (34), 16644-16654 (1987).


Introduction


Introduction


E. Delagoutte and P.H. von Hippel, Molecular mechanisms of the functional coupling of the helicase (gp41) and polymerase (gp43) of bacteriophage T4 within the DNA replication fork, Biochemistry 40 (14), 4459-4477 (2001).


C.E. Jones, T.C. Mueser, and N.G. Nossal, Interaction of the bacteriophage T4 gene 59 helicase loading protein and gene 41 helicase with each other and with fork, flap, and cruciform DNA, J. Biol. Chem. 275 (35), 27145-27154 (2000).


D.H. Jing, F. Dong, G.J. Latham, and P.H. von Hippel, Interactions of bacteriophage T4-coded primase (gp61) with the T4 replication helicase (gp41) and DNA in primosome formation, J. Biol. Chem. 274 (38), 27287-27298 (1999).


C.E. Jones, T.C. Mueser, and N.G. Nossal, Bacteriophage T4 gene 32 protein is required for helicase-dependent leading strand synthesis when the helicase is loaded by the T4 59 helicase-loading protein, J. Biol. Chem. 279 (13), 12067-12075 (2004).


Chapter 1


Introduction


Introduction


Aim and outline of this thesis

This thesis aims to develop processive oxidation catalysts based on the T4 clamp protein. In this chapter the different strategies to obtain the proposed bio-hybrid oxidation system will be explored. First, different procedures to link the chemical catalyst to the protein are introduced, i.e. the direct labelling and the “padlock” approach, respectively. Second, mechanisms by which the catalyst can be loaded onto DNA are identified and discussed. Furthermore, two types of catalysts are introduced that will be used to construct the T4 clamp-based artificial enzyme: cationic manganese porphyrins and footprinting reagents. Whereas the former type is able to specifically cut AAA sequences of DNA, the latter type causes random oxidative damage. Finally, the outline of this thesis is presented.
1. **Aim of this thesis**

The previous chapter shows that clamp proteins are central components in a large variety of DNA metabolic processes. They encircle DNA and tether proteins to DNA via well-understood binding domains in order to enhance their association and confer processivity to their catalytic actions on DNA. Clamp proteins track along DNA, either carrying a second protein or as a single entity, and are loaded and unloaded at specific sites by clamp loader proteins.

The work described in this thesis aims to investigate the use of clamp proteins as scaffolds for chemical catalysts. The so-called porphyrin clip (see Chapter 1, Section 8) represents a well-studied synthetic catalyst threaded onto a macromolecular substrate. Reminiscent of the processive enzymes found in Nature, the rotaxane topology of the porphyrin clip on its substrate will induce processive properties to the catalytic process. For the clamp-based catalysts this property should be very pronounced as their natural function is to serve as processivity enhancers. Furthermore, these conjugates allow us to study chemical catalytic actions on biological templates in aqueous solutions. In the long run, post-synthetic modification of synthetic and natural polymers with the clamp-based catalysts in aqueous solvents can provide routes towards modified macromolecules with tailored properties.

![Figure 1 - Schematic representation of the gp45 catalyst conjugate. The protein is bound to long linear DNA substrates that can be oxidized by the catalysts.](image)

In addition, since clamp proteins are crucial components for DNA replication, one can envisage its use as a target for drug development. Tumour cells rely heavily on their DNA replication apparatuses to enable rapid multiplication. This property of tumour cells is already in clinical use to treat cancers with chemotherapeutic agents that preferentially target fast-growing cells. Most likely, high concentrations of proliferating cell nuclear antigen (PCNA, the clamp protein of eukaryotes) are present in tumour cells to facilitate the
replication of DNA. If one could bring drugs that are conjugated to PCNA-interacting motifs into tumour cells, these motifs would (1) inhibit the rate of DNA replication by distorting the interaction of DNA polymerases with the clamp, and (2) bring chemical drugs into close contact with DNA. The use of a clamp interacting peptide to distort DNA synthesis was already explored for the human herpes simplex virus. The peptide-appended drugs could be chemical nucleases, such as discussed below, or photodynamic therapy agents. The latter are porphyrin-based materials that are gaining popularity in the clinical treatment of cancers. Despite the vast amount of literature on DNA-binding porphyrins, these compounds are uncharged and quite hydrophobic, so it is not expected that they could bind to DNA autonomously. Therefore, the conjugation of these drugs to clamp interacting motifs could also provide a strategy to further enhance the selectivity of photodynamic therapy.

The work described in this thesis will employ the T4 clamp (gp45) as a scaffold for catalyst attachment. A wealth of information is available on the T4 bacteriophage clamp and its role in the T4 replication system. Moreover, a large amount of clamp mutants have been constructed that have cysteine residues on various specific locations on the clamp, which could be used for the construction of the clamp-catalyst hybrids (see below). Gp45 is unique among the clamp proteins, as it has an open subunit interface which is the major site for interaction with clamp binding proteins. This feature may be explored to load the clamp onto DNA. Furthermore, we aim to use the proposed bio-hybrid catalysts to oxidize synthetic polymer substrates, such as poly(isocyanopeptides) or poly(butadiene). Currently, there is no information available whether clamp proteins could be used to fulfil this goal; in particular it is unknown whether the clamp could bind to these polymers. This will therefore be investigated in this thesis. Since the association processes of clamp proteins with DNA is better described, this chapter will focus mainly to use DNA as the substrate for the catalytic oxidation reaction.

2. How will the clamp be labelled with catalysts?

Two strategies will be explored to label clamp proteins with catalysts. First, taking advantage of the large amount of clamp mutants available, catalysts can be conjugated to cysteines at specific sites on the clamp. The catalysts will have to be functionalized with either a maleimide or iodoacetic acid to target the cysteine thiols. Because the clamp is a homotrimer, up to three catalysts per protein ring may be expected. This strategy will be referred to as the direct labelling approach.

Second, the catalyst can be conjugated to a clamp interaction motif, such as the C-terminus region of the T4 DNA polymerase. The peptide will bind to the clamp in the open subunit interface or the interdomain connecting loop. In the first case, the peptide may stabilize the clamp trimer by closing the subunit interface, which is proposed to lead to more stable clamp-DNA complexes. The catalyst will act as a padlock to close the clamp around its substrate. In contrast to the direct labelling approach, the catalysts are now bound to the
clamp in a non-covalent fashion, allowing for dynamic interactions. A peptide corresponding to the C-terminus of gp43 appended with a cysteine is already described in the literature.\textsuperscript{10,12,13} By using the abovementioned thiol-based labelling approach, the same maleimide or iodoacetamide functionalized catalysts can be used as in the direct labelling approach. Moreover, the unlabelled peptide can also be added to the covalent clamp-catalyst hybrids to achieve the same stabilization effect.

**Figure 2** – Different strategies for the construction of processive biohybrid catalysts; A. Direct labelling approach in which chemical catalysts are coupled directly to the T4 clamp protein. The clamp is subsequently stabilized on a DNA template by the addition of a polymerase peptide mimic (small square). B. The padlock approach uses a labelled polymerase peptide mimic to introduce the catalytic functionality onto the protein carrier. Stabilization and introduction of catalyst is achieved in one step.

### 3. How can the clamp be loaded onto DNA?

The T4 clamp is loaded onto DNA by the T4 clamp loader (gp44/62).\textsuperscript{14} This process requires the hydrolysis of ATP and a suitable DNA template, such as a nick or a primer-template site.\textsuperscript{15,16} Provided that the presence of a catalyst label does not inhibit the interaction of gp45 with gp44/62, the modified clamp proteins can be loaded onto DNA by the clamp loader complex. The clamp loading site on the DNA substrate may complicate analysis of the oxidation reactions, e.g. due to loss of sensitivity when analyzed by agarose gel electrophoresis. Therefore it may be worthwhile to look into alternative ways of loading the clamp onto DNA.

In a recent report, it was demonstrated that gp45 binds to a small forked DNA substrate on a surface without the help of the clamp loader.\textsuperscript{17} The clamp was not easily removed by wash steps, indicating that the clamp binds quite strongly to the DNA. The clamp was not able to interact with the DNA polymerase (gp43), however, so it is possible
that the C-terminus peptide analogue cannot bind to this non-specifically bound gp45. It is also uncertain whether the clamp is in the right conformation to track along DNA. On the other hand, the non-specific binding of gp45 onto DNA would provide an easy way of assembling the catalyst on DNA.

Macromolecular crowding\(^{18,19}\) has also been explored as a driving force for the assembly of the T4 holoenzyme. The crowded solutions contain high amounts of hydrophilic polymers and mimic the cellular interior, in which \(\sim 30\%\) of the physical space is occupied by biomolecules. Due to the excluded volume effect, the addition of these polymers effectively promote protein-protein and protein-DNA interactions because the polymers need to be hydrated in solution. Indeed, Jarvis \textit{et al.} showed a 2 fold increase in the binding constant of the complex between gp45 and gp44/62 in the presence of 7.5\% PEG (\(M_N = 12,000\)).\(^{20}\) Reddy \textit{et al.} showed that the presence of high amounts of PEG (\(\sim 10\%\) by weight) promoted the formation of the holoenzyme without the help of the clamp loader or ATP hydrolysis.\(^{21}\) Thus, the use of these crowded solutions could also effectively force the clamp onto DNA.

4. Porphyrins as DNA oxidizing catalysts

Porphyrins are promising drugs because of their inhibitory effect on eukarytic cell growth, antiviral, and antibacterial effects.\(^{22}\) Photosensitation of porphyrin derivatives such as hematoporphyrin (protoporphyrin IX derivatives in which the vinyl groups are replaced with \(\alpha\)-hydroxyethyl moieties) generates reactive oxygen species such as singlet oxygen and hydroxyl radicals\(^{23,24}\) that damage the cellular DNA.\(^{25}\) This is the underlying basis for photodynamic therapy which can be used to treat cancer.\(^{26}\) In bacterial cells, supercoiled plasmids were found to be converted to the circular form due to DNA scission, which can inhibit their replication and transcription. It was found that bacterial clones that survived porphyrin treatment lost their plasmid-encoded antibiotic resistance.\(^{25}\) Porphyrins are also selectively taken up by tumour cells,\(^{27}\) providing some selectivity to the treatment and the ability to image tumour tissues.\(^{28,29}\)

The interaction of DNA with cationic porphyrins, typically those with methylated pyridyl side groups on the \textit{meso} positions, has been studied extensively and has been reviewed a number of times.\(^{30-32}\) In short, there are three different binding modes that depend on the porphyrin substituents, inserted metal, and DNA sequence. Generally, intercalation is favoured at regions rich in GC base pairs, while external binding occurs primarily in the minor groove of AT rich regions.\(^{33-36}\) The latter binding mode causes some disruption of the hydrogen bonding between the base pairs. Porphyrins with metals that have axial ligands (Co, Fe, Mn, Sn, Zn, etc) are too bulky to intercalate and bind on the outside, while free base (2H) and metal containing porphyrins without axial ligands (Au, Cu, Ni, Pd) can intercalate. A third mode of interaction, in which the porphyrin binds to the outside of DNA while stacking, has also been identified.\(^{37,38}\) Preference of a porphyrin for either binding mode may vary with salt concentration\(^{39}\) or location of the pyridyl nitrogen.
atom on the periphery, as well as the DNA conformation (e.g. single or double stranded), as well as the DNA conformation (e.g. single or double stranded, \textsuperscript{43} Z, \textsuperscript{44} triple helix \textsuperscript{45}). Porphyrins or porphyrin-like compounds have been conjugated to nucleic acids, \textsuperscript{46-48} peptides \textsuperscript{49} or amino acids \textsuperscript{50-52} to study either their binding to DNA or to improve their aimed action.

Particularly interesting are the manganese derivatives of meso-tetrakis(N-methylpyridinium-4-yl)porphyrins (Mn-TMPyP, Figure 3). Besides being active catalysts for various organic substrates, they can also promote DNA damage \textsuperscript{53,54} when activated with oxygen donors \textsuperscript{55} such as potassium monopersulfate \textsuperscript{56} or magnesium monoperoxophtalate \textsuperscript{57}, promoting the formation of a high-valent manganese-oxo (Mn\textsuperscript{V}=O) species. \textsuperscript{58} A large amount of methylated tripyridyl porphyrins derivatives has been synthesized \textsuperscript{59,60}, mostly in the group of Meunier, and their DNA cleaving ability has been studied. The porphyrins were also functionalized with intercalating agents such as ellipticine \textsuperscript{61}, which showed increased cytotoxicity against leukaemia cells \textsuperscript{62}, and aminoquinoline, in order to target the G-quadruplex telomere region. \textsuperscript{63,64} Conjugation of oligonucleotides to porphyrins increases its activity and specificity when a proper linker length was used. \textsuperscript{65,66} Mn-TMPyP binds to calf thymus (CT) DNA with an association constant of \( K_A = 10^4 \), but showed significantly higher affinity for poly(A-T) than for poly(G-C) regions (\( K_A = 10^5 \) vs. \( 10^3 \), respectively). Thus, Mn-TMPyP preferably binds to the minor groove of AT rich regions. \textsuperscript{67} The mechanism of DNA oxidation by Mn-TMPyP has been studied in great detail, and was shown to involve multiple pathways, including the oxidation of guanine bases at the 8-position due to electron transfer, \textsuperscript{69} oxygen transfer to C1’ and C4’ positions in deoxyribose units at the 5’ end of oligonucleotides. The main pathway, however, is the single strand cleavage at sequences of three consecutive adenosine-thymine base pairs, by an oxygen transfer reaction to the C5’ position of the deoxyribose. This leaves the 5’ end opposite to the three adenosines of the cleavage site with an aldehyde, which could be reduced and ligated back to the opposite strand to yield the original oligonucleotide duplex. \textsuperscript{70} The reactivity of this aldehyde was further probed by HPLC/MS methods, which showed that it could readily react with nucleophiles (water, amines, hydroxylamines). \textsuperscript{71} Recently, Mourges \textit{et al.} completely characterized the Mn-TMPyP mediated oxidation of short oligonucleotides with HPLC/MS methods, which resulted in the mechanism outlined in Figure 4. \textsuperscript{72}
The specificity of the manganese methylated pyridine porphyrins for AAA sites in the minor groove of dsDNA was suggested to arise from the larger electron potential at AT-rich sequences as compared to GC-rich sequences of double stranded DNA in general, combined with the smaller width of the minor groove at AT-rich sequences, which matches the distance between the positively charged groups in the porphyrin in particular (9 Å vs. 12 Å in normal B-DNA). Furthermore, AT-rich sequences are more flexible than GC-rich sequences, which would allow optimization of the electrostatic interactions. Although Dabrowiak et al. demonstrated by using DNase I footprinting experiments that the porphyrin indeed binds in the minor groove, oxidation experiments do not suggest that the site has to be located in the minor groove to be cleaved, because generally all the AAA sites within the studied oligonucleotides are cleaved. It could also be possible that the cleavage induces a change in the structure of nearby sites to facilitate an oxidation reaction to occur in the latter.

It is evident that the Mn-TMPyP-type porphyrins are interesting DNA oxidation agents because of their specificity towards AAA sites. Moreover, the preparation of derivatives based on TMPyP has been described in detail in the literature. We decided

Figure 4 – Oxidation of DNA by Mn-TMPyP / KHSO₃. Top: C1' hydroxylation; Bottom: C5' hydroxylation, which is the main pathway. Both mechanisms lead to a single strand break in the DNA strand.

Figure 5 – Maleimide derivatized porphyrin, suitable for functionalization of the T4 polymerase C-terminus peptide analogue (padlock approach) and the clamp itself (direct labelling approach).
therefore to functionalize the T4 clamp with these porphyrins. We will need a porphyrin derivative that can react with the cysteine sulfhydryls on the protein, e.g. a porphyrin functionalized with maleimide or iodoacetamide groups. To allow the porphyrin to bind to the minor groove of DNA in the correct conformation, a spacer must be introduced to provide some flexibility to the system.

5. Hydroxyl radical generating species as catalysts

In addition to the sequence-specific Mn-TMPyP catalyst, we decided to also use a non-specific oxidation catalyst. There is an extensive class of compounds available known as artificial metallonucleases or chemical nucleases, involving the well-known members 1,10-phenanthroline copper complexes and the iron-containing catalyst bleomycin. Although the development of nucleases based on Fe(II) and Cu(I) is still an active field of research, other transition metals as diverse as Zn(II), Ce(IV), Ni(II), W(V), and lanthanides have also been employed. Metal-free chemical nucleases have been prepared as well, such as the lysine-enediyne conjugates and tris(2-aminobenzimidazoles), which hydrolyzes RNA. The majority of artificial nucleases need activation (reductant and/or oxygen donor), but some complexes are able to cleave DNA without activation. Several reviews on this topic have been published in the last decades.

Perhaps the best known and most widely used chemical nuclease is the EDTA complex of Fe(III). When incubated with a reductant (such as ascorbic acid or DTT) and an oxygen donor (hydrogen peroxide or dioxygen), Fe(III) is reduced to Fe(II) that subsequently reacts with the oxygen donor to generate hydroxyl radicals, a reaction which is called the Fenton reaction (Figure 6). This reaction is widely used in DNA footprinting experiments, in which DNA binding proteins provide protection against the hydrolytic cleavage of the hydroxyl radicals, and hence generate a ‘footprint’ when the DNA is denatured and analyzed by gel electrophoresis. Analogously, solvent accessibility of complex DNA structures can also be investigated in this way. The small size and high reactivity of the hydroxyl radical provides superior resolution over other, larger footprinting agents such as the DNase I enzyme. It could even be used to measure the helical twist of DNA molecules absorbed on crystal surfaces. The Fe-EDTA system differs from other chemical nucleases in that it is negatively charged and non-intercalating, in contrast to the intercalating 1,10-phenanthroline copper complexes which can also be used for footprinting. Thus, the Fe-EDTA complex has no affinity for DNA. To enhance the efficiency of the Fe-EDTA footprinting reaction, it has therefore been conjugated to intercalating moieties such as methidium bromide (the methyl derivative of ethidium bromide). Detailed information on the interaction pattern in protein-DNA or RNA complexes can be obtained by coupling the Fe-EDTA complex at different sites of a protein and analyzing the resulting footprint. For example, this method was used to determine the contact sites of RNA polymerase with DNA and with its cofactors. Activated Fe-EDTA complexes can also act as artificial peptidases, involving the attack of peroxy-Fe species on the peptide carbonyl bond. This
concept has been expanded to the mapping of protein-DNA interactions via Fe-EDTA conjugated DNA strands by analysis of the different protein fragments rather than DNA fragments.\textsuperscript{113} A commercially available Fe-EDTA compound that can be conjugated to cysteines, \((S)-1-(p\text{-bromoacetamidobenzyl})\text{ethylenediaminetetraacetate (FeBABE, also called Meares' reagent)} \text{ is available.}^{114} \text{ The reach of hydroxyl radical damage with this reagent, measured from the cysteine thiol, is } 22 \text{ Å for DNA damage and 12 Å for proteolytic cleavage (e.g. cleavage of amide bonds of proteins or peptides).}^{109}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure6.png}
\caption{Fe-EDTA complexes and the Fenton reaction. A. The Fenton reaction. The presence of an excess of reducing agent is needed for the catalytic generation of hydroxyl radicals. B. The commercially available FeBABE reagent that can be conjugated to sulfhydryls.}
\end{figure}

In the studies presented in this thesis, the T4 clamp, as well the gp43 C-terminus peptide analogue, will be labelled with the FeBABE compound, in order to study the effect of the T4 clamp on the efficiency of DNA cleavage. In the most optimal case, labelling of the clamp by the direct labelling or the padlock approach provides a mobile footprinting agent that will completely degrade DNA while it is sliding over it.

6. **Outline of this thesis**

In this thesis various aspects of the development of clamp-based catalytic systems will be treated in different chapters. In order to improve our understanding of the behaviour of the clamp in the presence of synthetic polymers, the interaction of the T4 clamp with synthetic polymers is studied (Chapter 3). Subsequently, strategies towards clamp binding onto hydrophobic polymers will be presented in Chapter 4. The interaction of the C-terminus analogue of gp43 with the clamp in order to establish whether the peptide can stabilize the clamp-DNA complex is the topic of Chapter 5. Chapter 6 involves the synthesis of porphyrin derivatives that can be used to create the catalyst-clamp hybrids for the direct labelling approach. Next, a suitable mutant of the clamp is expressed and labelled with the respective catalysts, followed by oxidation studies with the resulting conjugates (Chapter 7). Finally, AFM experiments that were performed to visualize the clamp complex on DNA are presented in Chapter 8.
7. References


Chapter 2


Aim and outline of this thesis


72 S. Mourguès, A. Kupan, G. Pratviel, and B. Meunier, Use of Short Duplexes for the Analysis of the Sequence-Dependent Cleavage of DNA by a Chemical Nuclease, a Manganese Porphyrin, Chembiochem 6 (12), 2326-2335 (2005).


78 J.C. Dabrowiak, B. Ward, and J. Goodisman, Quantitative footprinting analysis using a DNA-cleaving metalloporphyrin complex, Biochemistry 28 (8), 3314-3322 (1989).


Investigations on the interaction of the T4 clamp with synthetic polymers

The interactions of the T4 clamp with synthetic polymers and a DNA plasmid lacking any clamp loader sites are investigated via ATPase assays and stopped-flow fluorescence spectroscopy measurements. Two poly(isocyanopeptides), PIAA and PIAAA, with negatively charged carboxylates on the periphery of the polymer, are proposed to mimic the physical properties of DNA and could therefore be suitable templates for the T4 clamp. Under the conditions used, i.e. in the presence of the clamp loader complex or macromolecular crowding agents, it was concluded that the clamp neither binds to the polymer templates nor to the DNA plasmid with high affinity. FRET studies however suggest a weak interaction with the poly(isocyanopeptide) PIAA in macromolecular crowding conditions.
1. Introduction

One of the long-term goals for the development of T4 clamp based catalytic systems is their application in the oxidation of synthetic polymers. The clamp’s unique protein structure, in particular the inner ring with positive electrostatic charge (Figure 1), may have an effect on the oxidation process which could result in unique behaviour of the catalyst appended clamp proteins on synthetic polymers and the synthesis of polymers with special functional characteristics. First, however, the interaction of the T4 clamp protein with synthetic polymers has to be investigated; it may very well be that the protein does not readily bind to these synthetic entities. Taking into account the charge distribution of the T4 clamp, polymers with negatively charged backbones or functionalities may be the best candidates. Furthermore, the polymer structure is probably of crucial importance. Based on the X-Ray structures of several clamps, it was suggested that one or several layers of water molecules can be present between the DNA and the protein to facilitate tracking of the proteins along DNA.\(^1\) The diameter of the hole in the clamp structures is generally larger (34 Å for PCNA, 35 Å for gp45 and 38 Å for β subunit) than the minimal distance required to accommodate double stranded DNA (18 - 21 Å for the A and B forms of DNA).\(^2\) These distances are calculated by measuring points across the internal hole at which the atomic

![Figure 1](image-url) - Electrostatic potential maps of clamp proteins as calculated by GRASP\(^3\) demonstrate that the central channel of clamp proteins is positively charged. A. Comparison of the molecular surface of gp45 (top) with the electrostatic potential map (bottom) clearly shows the positively charged inner cavity (darker region) of the clamp protein, whereas the protein is overall more negatively charged. B. Molecular surface and electrostatic potential map for yeast PCNA.
density is half of the maximal value. Direct contact between the amino acid side chains with the phosphate backbone however cannot be excluded for PCNA. All the above highlight that efficient translocation requires finely tuned interactions between the protein and its substrate. Tinker et al. have shown with the help of crosslinking experiments employing photoactive nucleotides on DNA substrates that gp45, E. coli β subunit of polymerase III, and PCNA can slide freely over double stranded DNA segments. Stukenberg et al. have investigated the dissociation of the β clamp from circular and linearized plasmids, and found that the β clamp only dissociates from a DNA plasmid when the plasmid is cut by an endonuclease. The β clamp can slide over blunt ends or short overhangs; ends that are blocked by proteins or long single stranded regions, however, prevent the protein from sliding off. The T4 clamp is also not capable of sliding over 250 nt single stranded DNA regions. These combined observations may be explained by the assumption that it is energetically unfavourable for clamp proteins to pass a DNA single strand region because of the presence of a lower amount of negatively charged phosphates. Alternatively, the structure of ssDNA regions may prevent the clamp from sliding off, as ssDNA regions do not have a long persistence length, which results in a more globular structure. It is important to note, however, that these studies do not represent the situation in vivo, where ssDNA regions are covered with the single strand binding protein (gp32), which provides a protein barrier to block dissociation. Nevertheless, the results obtained by the studies above provide guidelines for polymer substrates with regard to being successful for T4 clamp binding. Thus, for the clamp to efficiently slide over synthetic polymers, these polymers have to have a long persistence length, possess negative charges and have an appropriate diameter.

The poly(isocyanopeptides), synthesized and extensively studied in the group of Nolte, might be ideal candidates to serve as templates for clamp binding. Treatment of isocyanides with Ni(II) orients the isocyanides in a square planar geometry and forces either clockwise or counter clockwise polymerization around the Ni(II) centre. This mechanism, referred to as the merry-go-round mechanism, results in polymers with helical 41 polyimide backbone (Figure 2). The polymer can be further rigidified by additional peptide bonds along the helix axis, as in the case of isocyanopeptides. The peptide bonding arrays,

![Figure 2](image-url) - Structure of poly(isocyanopeptides). A. Calculated structure of a 20-mer LLI-PIAAA polymer. Hydrogen bonds are indicated with dashed lines. B. Schematic structure of the polymers, showing the helical backbone and the hydrogen bonding side chains, viewed from the side and top, respectively.
analogous to peptide bonding motifs found in β-sheets, results in rigid polymers with persistence lengths higher than that of DNA. Several polymers or polystyrene block copolymers synthesized from isocyanopeptides with different peptide, thiophene, porphyrin, and perylene functionalities, have been characterized. The block copolymers have been shown to form aggregates with unique morphologies, whereas the polymers with chromophore side chains are studied for their charge transfer properties and their application in organic solar cells.

Furthermore, negatively charged poly(isocyanides) have been prepared and studied as crystallization templates for calcium minerals. These polymers can be synthesized from isocyanide methyl esters followed by base hydrolysis of the esters after the polymer synthesis, or from the corresponding hydrolyzed monomers. The latter procedure yielded polymers with poorly defined backbones, whereas the hydrolyzed polymers prepared by the former procedure had a well defined helical backbone, which was not disturbed upon dissolution of the polymer in aqueous media. Two negatively charged poly(isocyanides) with two (PIAA) and three (PIAAA) alanines in the side chain of the repeating unit of the polymer were thought to efficiently mimic DNA templates in aqueous solutions, as their diameters approach the diameter of the B-form of DNA: 16 Å and 19 Å for PIAA and PIAAA, respectively, compared to 18 Å for common B-DNA (Figure 3).

Two methods were used to investigate the ability of the T4 clamp to interact with various synthetic polymers, including two poly(isocyanides). As an initial study, the interaction of gp45 with these polymers, as well as a supercoiled DNA plasmid was
investigated using an ATPase stimulation assay (Figure 4).\textsuperscript{18} To conveniently monitor the rate of ATP hydrolysis, an enzyme cocktail of pyruvate kinase and lactate dehydrogenase was added, which allowed monitoring of ATP hydrolysis by a decrease in NADH concentration (Figure 4). The rate of NADH oxidation can be monitored with UV spectroscopy at 340 nm. This enzyme coupled assay can be used to measure the ability of the clamp loader to load the clamp onto the substrate.\textsuperscript{19,21} Normally, small DNA substrates are used to mimic primer/template junctions \textit{in vitro} onto which the T4 holoenzyme can be assembled.\textsuperscript{22} The double stranded side of the junction is functionalized with a biotin-streptavidin complex, while at the other end an oligonucleotide is partially annealed to the template strand. Thus the clamp is prevented from sliding off in either direction, effectively stabilizing the complexes that are formed on the DNA.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4}
\caption{The ATPase assay. A. The rate of ATP hydrolysis is coupled to the oxidation of NADH by pyruvate kinase and lactate dehydrogenase. B. A forked DNA substrate (Bio62/34/36) that is routinely used to provide a primer/template junction on which the holoenzyme can be assembled.}
\end{figure}

Each of the ATPase domains of the clamp loader (gp44 subunits) can independently bind one ATP molecule,\textsuperscript{23,24} which are hydrolyzed pairwise in two steps in the presence of the respective cofactors.\textsuperscript{25} The first step involves the formation of a gp45-gp44/62 complex in the presence of the T4 clamp, which results in the hydrolysis of 2 ATP molecules. This leads to an conformational change in gp44/62 and a subsequent opening of the subunit interface of the clamp.\textsuperscript{26-28} Consequently, the ATP hydrolysis rate is typically enhanced by >30 fold. Second, upon interaction with a template with primer-template junctions, such as the forked DNA substrate, gp44/62 hydrolyzes the remaining two ATP molecules. The rate of ATP hydrolysis is increased a further ~10 fold due to the formation of a gp44/62-gp45-DNA complex in which gp45 is twisted.\textsuperscript{29} The resulting complex is unstable and thus prone to dissociation, until gp45 is locked in place by the T4 polymerase gp43. After dissociation, the
complex is reformed, thus explaining the continuous rates of ATP hydrolysis. The addition of gp43 leads to a more stable complex and consequently lowers the rate of ATP hydrolysis.

In addition to the ATPase stimulation assays, fluorescence experiments with labelled gp45 mutants have been shown to yield valuable insight in the formation of the T4 holoenzyme (Figure 5). Von Hippel and co-workers used a S19C mutant labelled with a dansyl group, resulting in the fluorophore being on the rough side of gp45.30 Likewise, a T7C mutant labelled with an ANDB fluorophore on the smooth side of gp45 was used by the Benkovic group.31 Upon formation the gp45 - gp44/62 complex, large changes in fluorescence were observed, because of changes in the local environment of the fluorophores (hydrophobicity, hydrodynamic properties). With the latter mutant a third state of fluorescence was also identified when the clamp was loaded onto the DNA template, either as a complex with gp44/62 or with gp43.31,32

A.  

B.  

Figure 5 – Fluorescence experiments with gp45 labelled mutants. A. Changes in the conformation of gp45 due to protein binding result in increased fluorescence intensity of the fluorophores (dansyl and ANDB). Mutants were S19C and T7C, located near the subunit interface on the rough and smooth side of gp45, respectively. B. FRET across the subunit interface. Tryptophan (W92) is excited and its emission is quenched by a CPM fluorophore positioned at the other side of the subunit interface in a S158C mutant. Changes in fluorescence are brought about by relative opening and closing of the subunit interface. Only one fluorophore per gp45 trimer is shown for reasons of clarity.

The experiments described above were initially carried out in a steady-state fluorescence setup using a standard fluorescence spectrometer, but later repeated with a stopped-flow fluorescence spectrometer to determine the kinetic parameters for the fluorescence transitions. In stopped-flow fluorescence spectroscopy, two solutions are rapidly injected into a mixing chamber while the fluorescence of the mixture is monitored at a specific wavelength. This allows pre-steady-state kinetic analysis of protein binding events. In the case of the ANDB-T7C mutant, these experiments have yielded the rate constants of a 6-step holoenzyme assembly pathway, including the dissociation rate constants of several intermediates.25 The clamp loading step was identified to be the rate-limiting step in this process. These experiments were later refined by measuring fluorescence energy resonance transfer (FRET, see Appendix 1) across the subunit interface of the clamp (Figure 5B). An endogenous tryptophan (W92), located at the subunit surface, together with a coumarin labelled cysteine of the V169C clamp mutant, served as the FRET pair to probe the changes
Investigation on the interaction of the T4 clamp with synthetic polymers

in the distance across the open subunit interface. During the holoenzyme assembly process the clamp protein undergoes different conformational changes as it interacts with the other protein components and DNA templates, which can be monitored via variations of efficiency of the FRET process. The exact experimental details of this method are described in Appendix 1. Using stopped-flow fluorescence spectroscopy, the formation of the holoenzyme could be further unravelled by the Benkovic group to a 10-step process (Figure 6), for which all the kinetic steps were determined by simulation of the fluorescence traces. Interaction of gp45 with gp44/62 in the presence of ATP leads to an opening of the subunit interface from 40 Å to >45 Å. Upon interaction of the transient

![Figure 6](image-url)  

Figure 6 - A. The 10-step holoenzyme assembly pathway as determined by stopped-flow fluorescence experiments. Each conformation of gp45 is labelled with a letter (in superscript). In short: gp45\textsuperscript{A} is gp45 in solution, gp45\textsuperscript{D} is opened gp45 in complex with gp44/62, gp45\textsuperscript{H} is the twisted conformation of gp45 in the gp45-gp44/62-DNA complex, gp45\textsuperscript{K} represents the gp45 conformation in the final holoenzyme, interacting with gp43. B. Schematic representation of the holoenzyme assembly pathway. Small circles on gp44/62 represent ATP molecules. C. Changes in gp45 conformation throughout the formation of the holoenzyme, determined by triangulation experiments. Figures A and C are reproduced from Trakselis et al.29
gp45—gp44/62—DNA complex with gp43, the subunit interface of the clamp closes to 30 Å. This closing was found not to require ATP, but was attributed to interactions of the clamp with DNA and the gp43 C-terminus. As a further clarification, three different gp45 mutants (S158C, V163C and T168C) were constructed, labelled with CPM, and used for the same experiments. The introduced cysteines were positioned such that they would report three dimensional changes in the clamp, rather than the one dimensional change in the previous experiments. Two cysteines (S158C and T168C) were on opposite sides of the ring but close to the subunit interface, while the other (V163C) was positioned at the middle of the subunit interface. This led to the discovery that the clamp in the gp45-gp44/62-DNA complex is actually twisted opposite to the pitch of the DNA template. It is thought that this brings the clamp into the proper conformation to interact with gp43, which is the next step in the formation of the holoenzyme.

2. Results and discussion

2.1 ATPase stimulation assays with polymeric substrates

If gp44/62 is able to load gp45 onto a polymer structure, differences in the ATP hydrolysis pattern can be expected when compared to the assay performed with the standard forked substrate, specifically the step in which the substrate is added to the gp45-gp44/62 complex. ATPase rates during the assembly of the wt holoenzyme were determined and found to be consistent with previous experiments (Table 1).

Titration of the polymer into a solution of the preformed clamp loader – clamp complex did only result in a very modest but significant ATP hydrolysis rate increase for PIAAAA (26%), whereas the forked DNA substrate typically is able to stimulate ATP hydrolysis rate 5-10 fold. The smaller PIAA polymer was not able to significantly stimulate ATP hydrolysis (3%). The lower rate found for the PIAAA polymer compared to a DNA substrate with primer-template junction can obviously be explained by the lack of these sites on the polymer. Hence, gp44/62 is not able to interact with the synthetic polymer as strongly as the DNA substrate and form the twisted gp45 conformation that leads to high ATPase rates. Nevertheless, the clamp loader opens the subunit interface of gp45 which could facilitate the polymer sliding through. Subsequently, the clamp could close around the negatively charged PIAA and PIAAAA substrates. The increase in ATPase activity found after the addition of PIAAA does suggest that the clamp might be complexed to the polymer, as this would lead to the dissociation of the clamp from the clamp loader – clamp complex, and the subsequent recycling of this complex. This process is expected to result in higher ATP hydrolysis rates. The increase in ATPase rate after the addition of PIAAAA polymer is promising, although this result was difficult to reproduce in an experiment that was carried out 5 years later (see Table 1). It is quite well possible that the polymer had denatured in the period between the experiments. Thus, additional experiments are needed to investigate this interaction.
Other polymers were also tested in this assay. Poly(ethyleneimine), which is partly positively charged at the pH used in the assay, resulted in an ATP hydrolysis rate decrease (21%) rather than an increase. This may be explained by the occurrence of interactions of this positively charged polymer with the negatively charged periphery of the clamp protein, inhibiting the formation of the clamp loader – clamp complex, or promoting aggregation of the DNA substrate due to electrostatic interactions with the polymer. Poly(vinylpyrrolidone) was also investigated as a hydrogen bonding scaffold to which the clamp could bind, however this polymer was found to be insoluble under the conditions used.

Finally, experiments with a supercoiled pUC18 plasmid were performed. This plasmid represents the natural DNA substrate for the clamp but does not have any primer-template junctions which serve as recognition sites for the clamp loader. This is a crucial control because the plasmid should be able to accommodate the clamp, but mimics the polymers in their absence of primer-template junctions. Addition of the plasmid to the gp44/62-gp45 complex did not have any influence on the ATPase rate. Apparently, the binding of the clamp loader to DNA is requisite for clamp loading when DNA substrates are used. Obviously, clamp loading onto double stranded regions away from ssDNA regions is desired in vivo. Indeed, Jarvis et al. demonstrated that gp44/62 only binds to ssDNA regions, with a strong preference for primer-template junctions.20 As such sites are not present on synthetic polymers, no loading of the clamp onto these polymers should be expected. Indeed, no significant changes in ATPase activity were measured with PIAA as substrate, thus the binding of clamps to this polymer either does not bring about a change in ATP hydrolysis rates, or it does not take place. The increase found for PIAAA, whilst promising, remains unexplained. When compared to the binding of the clamp to B-DNA, the bigger diameter of PIAAA compared could result in a different interaction mode of the clamp with PIAAA. This in turn could result in the binding of the opened clamp onto the polymer, whereas this process does not readily take place on DNA. Clearly, further studies need to be carried out to understand the delicate process of clamp binding to the polymer substrates.

Table 1 – ATPase stimulation of DNA and polymer templates

<table>
<thead>
<tr>
<th></th>
<th>+44/62</th>
<th>+45</th>
<th>+template</th>
<th>change</th>
<th>+43</th>
<th>change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forked DNA</td>
<td>6.5</td>
<td>91</td>
<td>503</td>
<td>450%</td>
<td>130</td>
<td>-74%</td>
</tr>
<tr>
<td>pUC18</td>
<td>5.0</td>
<td>67</td>
<td>69</td>
<td>3%</td>
<td>N.D.</td>
<td></td>
</tr>
<tr>
<td>PIAA</td>
<td>3.3</td>
<td>40</td>
<td>41</td>
<td>3%</td>
<td>N.D.</td>
<td></td>
</tr>
<tr>
<td>PIAAA</td>
<td>3.8</td>
<td>39</td>
<td>49</td>
<td>26%</td>
<td>N.D.</td>
<td></td>
</tr>
<tr>
<td>PEI</td>
<td>3.2</td>
<td>39</td>
<td>31</td>
<td>-21%</td>
<td>N.D.</td>
<td></td>
</tr>
</tbody>
</table>

*a Rates are in nM/s hydrolyzed ATP.

b No increase in ATP hydrolysis rate could not be observed in a recent experiment. This experiment was carried out 5 years after the initial experiment, and the lack of reproducibility was attributed due to denaturation of the PIAAA polymer in the time between these experiments. Due to lack of time and the fact that no recently synthesized PIAAA polymer was available, this experiment could not be repeated to verify the results.
2.2 Pre-steady state fluorescence measurements with polymers

The putative assembly of the T4 clamp onto synthetic polymers was further examined with the help of stopped-flow fluorescence spectroscopy. Two CPM labelled double mutants (V163C W199F and S158C W199F) were first screened for interaction with gp44/62 in the presence of ATP, of which the latter provided the highest fluorescence amplitude. Thus, in the following experiments this mutant was used to test whether the clamp could interact with the poly(isocyanopeptides). Full examination of the interaction of this mutant was done with the PIAA polymer, because it was rationalized that gp45 would more readily bind to the polymer with the smaller diameter. To correct for interprotein FRET and changes in CPM environment, interaction of the triple mutant (W91F, S158C and W199F) with PIAA polymer was also investigated. The results are listed in Table 2, and show clearly that the clamp does not close after addition of the polymer. States A-D represent the previously determined interaction states of the clamp with the clamp loader (see Figure 5). E is the final conformation of the clamp before addition of the polymer (similar to D). Upon addition of the polymer, two interaction states of gp45 with the polymer (F and G) were detected. The fluorescence changes of these states did not show significant closing of the clamp.

### Table 2 – Fluorescence parameters, transfer efficiencies and distances for W199F, S158C-CPM without PEG present.

<table>
<thead>
<tr>
<th>Holoenzyme State</th>
<th>Assembly State</th>
<th>F&lt;sub&gt;AD&lt;/sub&gt; λ = 290</th>
<th>F&lt;sub&gt;A&lt;/sub&gt; λ = 290</th>
<th>F&lt;sub&gt;AD&lt;/sub&gt; λ = 390</th>
<th>I&lt;sub&gt;AD&lt;/sub&gt;* + X</th>
<th>I&lt;sub&gt;A&lt;/sub&gt;* + X</th>
<th>I&lt;sub&gt;A&lt;/sub&gt;</th>
<th>X</th>
<th>I&lt;sub&gt;AD&lt;/sub&gt;</th>
<th>I&lt;sub&gt;AD&lt;/sub&gt;/I&lt;sub&gt;A&lt;/sub&gt;</th>
<th>E&lt;sub&gt;T&lt;/sub&gt;</th>
<th>R</th>
<th>R&lt;sub&gt;1&lt;/sub&gt;open</th>
</tr>
</thead>
<tbody>
<tr>
<td>45&lt;sup&gt;a&lt;/sup&gt;</td>
<td>A</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.502</td>
<td>1.000</td>
<td>0.000</td>
<td>1.502</td>
<td>1.502</td>
<td>0.409</td>
<td>30.942</td>
<td>48.082</td>
<td></td>
</tr>
<tr>
<td>45&lt;sup&gt;b&lt;/sup&gt;-44/62(4ATP)</td>
<td>B</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.502</td>
<td>1.000</td>
<td>0.000</td>
<td>1.502</td>
<td>1.502</td>
<td>0.409</td>
<td>30.942</td>
<td>48.082</td>
<td></td>
</tr>
<tr>
<td>45&lt;sup&gt;c&lt;/sup&gt;-44/62(4ATP)</td>
<td>C</td>
<td>1.210</td>
<td>1.281</td>
<td>1.239</td>
<td>1.817</td>
<td>1.281</td>
<td>1.239</td>
<td>0.042</td>
<td>1.775</td>
<td>1.433</td>
<td>0.353</td>
<td>32.199</td>
<td></td>
</tr>
<tr>
<td>45&lt;sup&gt;d&lt;/sup&gt;-44/62(2ATP) +PIAA</td>
<td>E</td>
<td>1.333</td>
<td>1.428</td>
<td>1.317</td>
<td>2.423</td>
<td>1.829</td>
<td>1.632</td>
<td>0.198</td>
<td>2.225</td>
<td>1.364</td>
<td>0.296</td>
<td>33.615</td>
<td></td>
</tr>
<tr>
<td>1st</td>
<td>F</td>
<td>1.009</td>
<td>1.006</td>
<td>1.003</td>
<td>2.444</td>
<td>1.840</td>
<td>1.637</td>
<td>0.204</td>
<td>2.241</td>
<td>1.369</td>
<td>0.301</td>
<td>33.494</td>
<td></td>
</tr>
<tr>
<td>2nd</td>
<td>G</td>
<td>1.011</td>
<td>1.006</td>
<td>1.000</td>
<td>2.471</td>
<td>1.851</td>
<td>1.637</td>
<td>0.215</td>
<td>2.257</td>
<td>1.379</td>
<td>0.309</td>
<td>33.287</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Columns are labelled as follows; Holoenzyme State indicates the components of the complexes, Assembly State correlates with the gp45 conformations listed in Figure 5A, F<sub>AD</sub> is the relative fluorescence intensity in presence of acceptor and donor, F<sub>A</sub> the relative fluorescence intensity in the presence of acceptor only, λ indicates the wavelength at which was excited. I<sub>AD</sub> and I<sub>A</sub> are the fluorescence intensities normalized by steady state fluorescence experiments in the presence and absence of donor, respectively. X is the amount of interprotein FRET. E<sub>T</sub> is the total average energy transfer efficiency, and R the average distance between all subunit interfaces, and R<sub>1</sub>open is the distance of the open subunit interface. Distances are given in Å.

The experiments were subsequently repeated in macromolecular crowded conditions (10% PEG 8,000), which yielded the parameters listed in Table 3. The data could be fitted to a two-step process (F and G), which revealed only a very minor closing of the clamp around the polymer of ~1 Å. This lies within the error values and may also have resulted from the changed binding affinity of the clamp towards the clamp loader rather than an increase in affinity of the clamp towards PIAA. From these experiments it cannot be concluded that the T4 clamp is able to bind to the PIAA polymer.
### Table 3 – Fluorescence parameters, transfer efficiencies and distances for W199F, SI58C-CPM with PEG\(^a,^b\)

<table>
<thead>
<tr>
<th>Holoenzyme State</th>
<th>Assembly State</th>
<th>F(_{400}) (\lambda = 290)</th>
<th>F(_{400}) (\lambda = 290)</th>
<th>F(_{400}) (\lambda = 390)</th>
<th>I(_{400}) + X</th>
<th>X</th>
<th>I(_{400})</th>
<th>I(<em>{400})/I(</em>{400})</th>
<th>E(_T)</th>
<th>R</th>
<th>R(_{open})</th>
</tr>
</thead>
<tbody>
<tr>
<td>45(^b)</td>
<td>A</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.502</td>
<td>1.000</td>
<td>1.000</td>
<td>1.502</td>
<td>1.502</td>
<td>0.409</td>
<td>30.942</td>
</tr>
<tr>
<td>45-PIAA 1st</td>
<td>F</td>
<td>1.099</td>
<td>1.004</td>
<td>1.016</td>
<td>1.516</td>
<td>1.004</td>
<td>1.016</td>
<td>-0.012</td>
<td>1.528</td>
<td>1.503</td>
<td>0.410</td>
</tr>
<tr>
<td>2nd</td>
<td>G</td>
<td>1.013</td>
<td>0.998</td>
<td>1.012</td>
<td>1.535</td>
<td>1.002</td>
<td>1.028</td>
<td>-0.026</td>
<td>1.561</td>
<td>1.519</td>
<td>0.422</td>
</tr>
</tbody>
</table>

\(^a\) Column labels are described in Table 2.
\(^b\) No clamp loader was included in this experiment.

### 3. Conclusions

The binding process of the T4 clamp onto several synthetic polymers and a DNA plasmid was investigated with ATPase assays as a steady state method and stopped-flow fluorescence spectroscopy as a pre-steady state method. In both cases, gp45 was opened with the help of gp44/62 in the presence of ATP to facilitate the polymer sliding through. Most notably, ATPase measurements with negatively charged poly(isocyanopeptide) polymers with two (PIAA) or three (PIAAA) alanine side chains, showed slightly different behaviour for the two polymers. While the PIAA was not able to promote any increase in the rate of ATP hydrolysis, there was a small but significant increase in the case of the PIAAA polymer. This could indicate an interaction of some sort, however additional data is needed to confirm this hypothesis. The DNA plasmid, presented as the natural substrate for the clamp without clamp loading sites, failed to stimulate ATP hydrolysis. Thus, the nature of the interaction of the clamp with PIAAA polymer might not be the same as that with DNA.

Furthermore, stopped-flow fluorescence spectroscopy was used to probe changes in the distance across the open subunit interface of the clamp. Binding of the clamp around the substrate is expected to decrease the subunit distance because of electrostatic interactions of the positively charged inner ring of the clamp with the negatively charged polymers. Experiments focussed on the PIAA polymer, because the clamp was supposed to bind the polymer with smaller diameter more easily. The experiments without molecular crowding agents showed no closing of the subunit interface, while in the presence of 10% PEG, a small 1 Å decrease was observed, although this decrease is well within the error value of the experiments.

We have to conclude therefore, that based on the experimental evidence presented here, the binding of gp45 onto synthetic polymers is not completely achieved, at least not by a clamp loader assisted mechanism or via macromolecular crowding agents. Nevertheless, the ATPase increase found for PIAAA and the small closing of the subunit interface of PIAA under macromolecular crowding conditions justify further investigation into the binding of the T4 clamp onto these two polymers.
4. Experimental section

General

ATPase assays were performed using a Cary 100 UV spectrometer. Stopped-flow fluorescence measurements were done using an Applied Photophysics (Surrey, UK) stopped-flow instrument. Pyruvate Kinase/Lactic Dehydrogenase enzymes from rabbit muscle was obtained from Sigma-Aldrich. LDL-PIAAA and LDL-PIAAAA were synthesized by Dr. Gerald Metselaar. All other materials were from standard commercial sources.

ATPase assays
A typical assay was performed in a ~300 μl total volume as follows: 250 μl of a cocktail solution (1.2 mM ATP, 4 mM phosphoenopyruvate, 0.24 mM NADH, 0.6 μM streptavidin, and 3.6 μl pyruvate kinase and lactate dehydrogenase enzyme mix [resulting in ~3.6 units pyruvate kinase and ~2.5 units lactate dehydrogenase per ml cocktail] in 30 mM Tris.OAc pH 7.5, 180 mM KOAc, 12 mM Mg(OAc)₂), was supplemented with 23.9 μl MilliQ and 1.7 μl gp44/62 solution (9.7 μM) and incubated for 3 mins, while the absorption of NADH was monitored at 340 nm. Then, 1.5 μl gp45 solution (49 μM) was added and the solution was incubated for at least three minutes before 1.9 μl forked DNA template (Bio62/34/36, 40 μM) was added. Finally, 15 μl gp43 solution (9 μM) was added. ATPase rates were calculated by fitting linear parts of the measured decrease of the absorbance at 340 nm as a function of time by using the least squares method. Rates were converted to nM/s by multiplying the rate in Abs/min with 2679.5 (= 10⁹ / 60 / εNADH), in which εNADH = 6220 M⁻¹ cm⁻¹.

Experiments with pUC18 were performed as described above, with the following modifications: 25.8 μl MilliQ was added first, and instead of the forked substrate, 0.6 μl of pUC18 solution (763 ng / μl) was added. No gp43 was added.

Experiments with polymers were performed in a similar way with the following modifications: 25.4 μl MilliQ was added first, and instead of the forked substrate, 1 μl of polymer solution (20 μg / μl) was added. No gp43 was added.

Stopped-flow fluorescence
Experiments were performed at 25 °C. Slits were at 2.2 nm, and a cut-off filter of 420 nm was used to filter out all fluorescence not coming from the coumarin dye. Excitation was at 290 nm, and 1000 data points were collected during each measurement. The parameters were calculated as described in Appendix 1.

Protein labelling
One millilitre aliquots of gp45 double mutants V163C/W199F or S167C/W199F (50 μM concentrations) in storage buffer (20 mM HEPES pH 7.5, 50 mM NaCl, 1 mM EDTA, 10 mM β-mercaptoethanol, 10% glycerol) were dialyzed into labelling buffer (20 mM HEPES pH 7.3, 50 mM KOAc, 1 mM EDTA, 10% glycerol, 2 times 1 L changes). Over a time course of 1 hr, 4 x 3.5 μl 10 mM CPM solution in DMF was added, followed by incubation of the protein for 4-5 hrs at room temperature. The reaction mixture was passed through a cellulose spin filter and dialyzed into storage buffer (2 x 1 L changes). Protein concentrations were determined with the Bradford assay, using wt gp45 as the standard. The labelled protein was divided into 100 μl aliquots and frozen in liquid nitrogen.
5. References


23 T.C. Jarvis, L.S. Paul, and P.H. Von Hippel, Structural and Enzymatic Studies of the T4 DNA-Repli-
24 M.C. Young, S.E. Weitzel, and P.H. von Hippel, The kinetic mechanism of formation of the
25 D.J. Sexton, B.F. Kaboord, A.J. Berdis, T.E. Carver, and S.J. Benkovic, Dissecting the order of
26 P. Pietroni, M.C. Young, G.J. Latham, and P.H. von Hippel, Structural analyses of gp45 sliding clamp
interactions during assembly of the bacteriophage T4 DNA polymerase holoenzyme. I.
Conformational changes within the gp44/62-gp45-ATP complex during clamp loading, *J. Biol. Chem.*
272 (50), 31666-31676 (1997).
27 S.C. Alley, E. Abel-Santos, and S.J. Benkovic, Tracking sliding clamp opening and closing during
28 M.A. Trakselis, A.J. Berdis, and S.J. Benkovic, Examination of the role of the clamp-loader and ATP
hydrolysis in the formation of the bacteriophage T4 polymerase holoenzyme, *J. Mol. Biol.* 326 (2), 435-
29 M.A. Trakselis, S.C. Alley, E. Abel-Santos, and S.J. Benkovic, Creating a dynamic picture of the sliding
clamp during T4 DNA polymerase holoenzyme assembly by using fluorescence resonance energy
30 G.J. Latham, P. Pietroni, F. Dong, M.C. Young, and P.H. von Hippel, Fluorescence monitoring of T4
polymerase holoenzyme accessory protein interactions during loading of the sliding clamp onto the
at the bacteriophage T4 DNA replication fork - Characterization of a fluorescently labeled DNA
32 G.J. Latham, D.J. Bacheller, P. Pietroni, and P.H. von Hippel, Structural analyses of gp45 sliding clamp
interactions during assembly of the bacteriophage T4 DNA polymerase holoenzyme. III. The Gp43
dNA polymerase binds to the same face of the sliding clamp as the clamp loader, *J. Biol. Chem.* 272
(50), 31685-31692 (1997).
DNA block-copolymers as T4 clamp templates

DNA block-copolymers, consisting of a DNA and a synthetic polymer block, are introduced as templates for the T4 clamp. The clamp can bind to the DNA block, and may subsequently translocate to facilitate oxidation reactions to take place on the polymer block. Initial studies focussed on the synthesis of DNA-PEG conjugates via the Michael addition of thiol-functionalized oligonucleotides and a commercially available PEG-maleimide polymer. Quantitative yields were obtained with 5 equivalents of polymer and any residual impurities could be removed by size exclusion chromatography. A maleimide appended poly(butadiene) polymer was also synthesized, but did not give rise to conjugates under the conditions tested. Other linking procedures were also tested but failed to give poly(butadiene)-DNA conjugates as well.
1. Introduction

The experiments described in Chapter 3 showed that the binding of the T4 clamp onto synthetic polymers is not readily achieved by clamp loader mediated loading or via autonomous binding driven by electrostatic interactions. It is necessary, therefore, to build primer-template junctions into the polymer to serve as clamp loading sites. When DNA block-copolymers, consisting of a DNA block with a clamp loading site and a synthetic polymer, are used as substrates, the clamp may first be loaded onto the DNA block followed by translocation to the polymer block (Figure 1). One challenge is the prevention of oxidation of the DNA block before the catalyst has moved to the polymer block. This may be achieved by excluding any AAA sites when the porphyrin appended clamp is used, as well as exclusion of guanines on the ends of the DNA block because these are also readily oxidized.

Figure 1 – Schematic representation of an oxidation process facilitated by catalyst appended clamps on DNA block copolymers. The DNA block provides a template for clamp binding, while the clamp itself may translocate to the synthetic polymer block to perform oxidation reactions.

In the last decades, the study of DNA-polymer hybrids has become an increasingly intense field of research, and a comprehensive review is beyond the scope of this thesis. Therefore, this introduction will mainly focus on PEG and hydrophobic polymer conjugates. In the early 1990s, first efforts were made to copolymerize DNA oligonucleotides with pyrrole into a conductive film that could be deposited onto electrodes, which allowed the detection of mismatches in a synthetic ras gene, as well as Hepatitis C virus genotyping. Most efforts, however, have been directed towards the conjugation of ODNs with poly(ethyleneglycol) (PEG) polymers. With positively charged polymers such as poly(ethylenimine), ODNs can be complexed into poly ion complexes (PICs) that can be used for non-viral therapeutic gene transfer in order to inhibit replication, transcription, mRNA translation, or RNA splicing of specific genes. The PEG block provides protection against nucleases and secretion from the body, contributing to the stability of the transfected genes. The conjugates are readily available by solid-phase synthesis techniques via phosphoamidite PEG derivatives, or PEG-modified solid phase supports. A number of solution phase synthesis methods has also been developed. Conjugates of DNA and hydroxyl functionalized PEG were successfully used by the group of Bonora as gene delivery vehicles, and an ODN-PEG-peptide conjugate was also reported. A number of other reactions has also been used to conjugate the ODNs to polymers such as Michael addition and nucleophilic substitution of iodoacetamides, by thiol functionalized ODNs with maleimides, and peptide bond synthesis via treatment of amine appended ODNs with...
activated PEGs. Several degradable linkages that are of interest with respect to the therapeutic release of ODNs within cells were also investigated. Jeong and Park, for instance, made use of the phosphoramidate linkage, which is degraded in acidic environments such as the endosomal compartment of cells, while Oishi et al. used the acid labile β-thiopropionate linkage to make functional PEG-DNA and PEG-RNA hybrids, and also investigated the disulfide linkage as a conjugation tool. Another noteworthy ODN release strategy from PEG conjugates involves a 1,6-benzyl elimination, which liberates the ODN after an enzymatic or hydrolytic cleavage in the linker.

The synthesis of DNA block copolymers containing a hydrophobic polymer block poses a considerable synthetic challenge, which is often overcome by either synthesizing the conjugate on the solid support or by performing the reaction in a common solvent such as DMSO. A limited number of polymers onto which the ODN is grafted, i.e. attached as side chains to the polymer backbone have been synthesized via free radical polymerization, RAFT, and attachment of the ODN to the polymer backbone via standard solid-phase synthesis methods. These DNA-polymer hybrids were successfully used for gene delivery, or single base pair mismatch detection purposes. The first example of a linear ODN - hydrophobic polymer block copolymer was prepared by Jeong and Park by treating an amine functionalized ODN with the NHS activated ester of a copolymer of lactic and glycolic acid in DMSO (Table 1). The amphiphilic nature of the conjugate resulted in the formation of micelles in aqueous solutions with a hydrophobic synthetic polymer core, which could be degraded at pH 8.0 to release the ODN. Polystyrene conjugates were also reported but were synthesized via solid-phase methods. An easy method is to functionalize the ODN strand while it is still on the solid phase, followed by deprotection, cleavage, and isolation of the conjugate from the CPG beads. Upon dispersion of the polymers in aqueous solvents, spherical micelles were formed that could be cross-linked by the addition of complementary ODN functionalized gold particles. Alternatively, Matsushita and co-workers demonstrated that a nucleotide-capped polystyrene polymer can serve as a support for solid-phase synthesis by sequentially building the ODN part from the polymer. Probably due to its short DNA block, the PS-T₅ polymer did not assemble into micelles, rather a cylindrical microphase-separated structure was obtained. The group of Herrmann investigated the application of DNA-block co-polymers in DNA-templated synthesis. In this approach, two different organic groups are brought into close proximity by hybridization of end-labelled ODNs. This can result in a significant rate enhancement when otherwise unreactive groups are used, or a dramatic selectivity when multiple reactants with different ODNs are used in one pot. DNA templated synthesis also allows the construction of libraries of protease inhibiting macrocycles, or, combined with PCR amplification, rapid catalyst screening and discovery of new reactions. The block-copolymers of poly(propyleneoxide) and DNA synthesized in the group of Herrmann were investigated as an unique scaffold for DNA templated synthesis reactions, because the reactants can selectively be placed at the hydrophilic (outside) or hydrophobic interface (at
Table 1 – Overview of linear DNA conjugates of hydrophobic polymers.

<table>
<thead>
<tr>
<th>Conjugate</th>
<th>Synthesis</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Conjugate Diagram" /></td>
<td>Activation of the polymer carboxylic end group with NHS, followed by amide bond formation with aminohexyl appended ODN in DMSO. Polymer is degraded at pH 8.0.</td>
<td>28</td>
</tr>
<tr>
<td><img src="image2" alt="Conjugate Diagram" /></td>
<td>Addition of a PS phosphoamadite in one DNA synthesis cycle, followed by deprotection. $M_N$ of the polymer block is 5.6 kD.</td>
<td>29</td>
</tr>
<tr>
<td><img src="image3" alt="Conjugate Diagram" /></td>
<td>Coupling of a protected thymine to hydroxyl terminated PS, followed by elongation of the DNA block to 5 thymine residues by solid phase synthesis methods. $M_N$ of the polymer block is 11 kD.</td>
<td>30</td>
</tr>
<tr>
<td><img src="image4" alt="Conjugate Diagram" /></td>
<td>Coupling of the phosphoamadite polymer in one DNA synthesis cycle. $M_N$ of the polymer is 1.0 and 6.8 kD.</td>
<td>36</td>
</tr>
</tbody>
</table>

Notably, isoindole formation studied in the core of the micelle proceeded in high yields, possibly because of accumulation of the hydrophobic reactants within the core, and allowed easy detection because of the fluorescent product that is formed.

2. Results and discussion

2.1 Synthesis of PEG-DNA conjugates

Before preparing the synthetically challenging DNA-poly(butadiene) block copolymers, it was decided to perform test reactions with a commercially available maleimide functionalized polyethylene glycol (PEG, $M_N = 5,000$) (I). This polymer was conjugated to various DNA oligonucleotides containing a thiol functionality (I-III). These oligonucleotides are available via solid phase synthesis techniques by appending the oligonucleotide with a hexane thiol moiety, protected with a trityl group. Thus, before conjugation a deprotection step is necessary, which is achieved by treating the ODN with a reducing agent. The excess of reductant has to be removed from the reaction solution as it will also react with the
DNA Block-copolymers as T4 clamp templates

maleimide. Optimization of the reaction conditions for these steps resulted in a 30 min incubation in 100 mM DL-dithiothreitol (DTT) followed by spin column isolation. Deprotection was also successful with other reductants such as tris(2-carboxyethyl)phosphine hydrochloride (TCEP) or treatment with silver nitrate. An alternative to remove the excess reagents from the reaction solution was ethanol.

A.  

\[
\begin{align*}
5' & \quad \text{I, II, III} \\
\overset{\text{O}}{\text{P}} & \quad \overset{\text{O}}{\text{O}} & \quad \overset{\text{O}}{\text{O}} & \quad \overset{\text{O}}{\text{O}} & \quad \overset{\text{O}}{\text{O}} \\
\text{(CH}_2\text{)_6O} & \quad \text{S} & \quad \overset{\text{O}}{\text{P}} & \quad \overset{\text{O}}{\text{O}} & \quad \overset{\text{O}}{\text{O}} & \quad \overset{\text{O}}{\text{O}} \\
\text{5'} & \\
\end{align*}
\]

I-I, II-I, III-I

B.  

\[
\begin{array}{c|cccc}
\text{Mw} & 0 & 1 & 5 & 0 & 1 & 5 \\
\text{unreduced} & & & & \text{reduced} & & \\
\end{array}
\]

\[
\begin{array}{c|cccc}
\text{Mw} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
\end{array}
\]

C.  

\[
\begin{array}{c|cccc}
\text{Mw} & \text{II-I} & \text{III-I} & \text{II-I} \cdot \text{III-I} \\
\end{array}
\]

D.  

\[
\begin{array}{c|cccc}
\text{Mw} & \text{II-I} & \text{III-I} & \text{II-I} \cdot \text{III-I} \\
\end{array}
\]

Figure 2 – Preparation and purification of PEG-DNA conjugates. A. Sulphydryl modified oligonucleotides (ODNs) are reacted with polyethylene glycol maleimide to yield the PEG-ODN conjugates. B. Deprotection of the thiol is necessary before the conjugation reaction, as trityl protected ODNs do not react (left lanes). When deprotected, the ODNs react with the maleimide functionalized polymer, resulting in a band shift of the oligonucleotide. Conjugation is complete with a five fold excess of polymer (right lanes). Samples are analyzed on a 15% non-denaturing TBE gel stained with Stains-All. C. The resulting conjugates are purified with a Superdex 75 size exclusion column. Fraction numbers are indicated. D. The conjugates can be dimerized by heating two complementary conjugates (A-PEG and B-PEG) and allowing the reaction mixture to cool down.
precipitation of the oligonucleotides. The first mentioned procedure was found to be the fastest, and proceeded without significant loss or dilution of the oligonucleotide. Incubation of the deprotected oligonucleotides with a five fold excess of maleimide polymer resulted in the near quantitative formation of the corresponding DNA block copolymers (Figure 2). Purification of the conjugate from excess polymer, unreacted and oxidized oligonucleotides could be successfully achieved by Superdex 75 size exclusion chromatography. The purified PEG-DNA conjugates could be readily dimerized when functionalized with two complementary oligonucleotides (II and III), showing that the availability of the DNA blocks is not impaired by the presence of the synthetic polymer.

2.2 Synthesis of poly(butadiene)-DNA conjugates

In a subsequent series of experiments, a maleimide functionalized 1,4 poly(butadiene) polymer was synthesized and used to construct a DNA conjugate. The rubber-like poly(butadiene)s have carbon-carbon double bonds in their backbones (in the case of the 1,4 polymer) or appended to the backbone (in the case of the 1,2 polymer) that can be epoxydized, and have a low glass temperature of 180 K, providing high flexibility. This makes them interesting targets for studying the oxidation of the double bonds with the catalyst appended clamp proteins. The commercially available poly(butadiene) polymer 2, which is synthesized by anionic polymerization ($M_N = 4100$, $M_W/M_N = 1.05$) and has an amine end functionality was used as starting compound. NMR and Maldi-TOF analyses showed that the commercial sample still contained protected amine functions, *i.e.* silyl groups, so for optimal reactivity the polymer was treated with tetrabutylamine fluoride (TBAF) in THF and precipitated in methanol before use. Furthermore, mass spectroscopy analysis indicated that the actual $M_N$ was 3200 rather than 4100 Da, as reported by the supplier. Reaction with the NHS ester of maleimide caproic acid 3 resulted in the maleimide functionalized polymer 4, which was purified by column chromatography on silica followed by a precipitation step (Scheme 1). This resulted in a poor yield of 31%, which could be attributed to the loss of several impurities present in the starting material. End group analysis of the product showed complete functionalization of the polymer with the maleimide group. The polymer was not stable over long periods of time due to cross-linking of the double bonds.

Scheme 1 – Reagents and conditions (i) a. DSC, DIPEA, 2 hrs, b. DIPEA, 2, CH$_2$Cl$_2$, 16 hrs, 31%. $n \approx 60$. 

[Diagram of the synthesis scheme]
The polymer was subsequently incubated in a Tris buffer (25 mM Tris-HCl pH 7.6, 50 mM NaCl, 1 mM EDTA) with a thiol functionalized oligonucleotide. In contrast to the PEG conjugates that were readily formed, no reaction between the poly(butadiene) polymer and the oligonucleotide occurred. Since the poly(butadiene) polymers are quite hydrophobic, dispersion of these polymers in aqueous solution could result in precipitation or aggregation. Therefore, THF was mixed into the reaction solution to concentrations up to 60% v/v and the excess of polymer was increased to 15-fold. THF was selected from a range of DNA compatible co-solvents (DMF, dioxane, DMSO, acetonitrile) and was found not to promote precipitation of the polymer in a blank experiment. Unfortunately, the addition of THF to the reaction mixture did not stimulate the formation of a conjugate. After an incubation period of 16 hours, most of the ODN was found to be in its oxidized form, probably due to the long incubation times.

A.

Figure 3 – Attempted preparation of poly(butadiene)-DNA conjugates. A. Reaction scheme. B. Incubation of the thiol functionalized oligonucleotides did not result in the formation of a conjugate, even not after mixing in 60% THF as co-solvent to dissolve the polymer; \( n \approx 60 \).

Although the reaction mixtures did not appear cloudy, it is possible that the polymer still precipitated into phase separated structures, even at the high THF concentrations used. This would result in limited availability of the maleimide groups since a large amount of maleimide groups would be located inside the aggregate and is therefore unavailable for reaction with the oligonucleotide. If the head group of the polymer could be made more hydrophilic, a polymer amphiphile would be obtained that will aggregate into well-defined structures with the maleimide groups exposed on the surface of the aggregate. This would allow for a better reaction between the maleimide and the oligonucleotide. With this in mind, the synthesis of a poly(butadiene) polymer in which the maleimide group is separated from the polymer part by a short ethylene glycol linker was undertaken. The acid functionalized Boc-protected spacer 6 was synthesized via reaction of the mono Boc-protected ethylene glycol amine 5 with an excess of succinic anhydride. The carboxylic acid of spacer 6 was subsequently activated with NHS to yield 7 and reacted with the polymer 2. The Boc-protecting group was removed in situ by TFA, yielding amine functionalized polymer 8. Finally, 8 was treated with an excess of the NHS ester of maleimide caproic acid 3 in order to
obtain the resulting maleimide poly(butadiene) polymer 9. The first steps of this synthesis yielded the intermediates in poor yields, because of difficult purification procedures of these hydrophilic compounds. Unfortunately, based on MalDI-TOF mass analysis the last reaction did not yield poly(butadiene) polymer 9.

\[ \text{Scheme 2} - \text{Reagents and conditions (i) 5 eq succinic anhydride, THF / CHCl}_3, 1 \text{ hr, } 62\%. \text{ (ii) 1.1 eq EDC, 1.1 eq NHS, CHCl}_3, 16 \text{ hrs, } 25\%. \text{ (iii) a. 1.1 eq EDC, 1.1 eq NHS, 16 hrs, b. 2 c. TFA, 63\%. (iv) a. 25 eq maleimide caproic acid, 25 eq DSC, 25 eq DIPEA, CHCl}_3, 2 \text{ hrs, b. 1 eq 7. } n \approx 60. \]

2.3 Towards “clicked” DNA-polymer conjugates

The results in the previous section showed that it is difficult to couple an ODN to a poly(butadiene) polymer in the solution phase using conventional conjugation methods. Alternatively, one can synthesize the conjugate on the solid phase, provided that the polymer is stable against the oxidation (I$_2$ or t-butylperoxide) and deprotection (concentrated ammonium hydroxide) methods used. We decided to focus, however, on the synthesis of the conjugates via the Cu(I) catalyzed Huisgen [3+2] dipolar cycloaddition reaction, also called the “click” reaction. Results in our lab had previously demonstrated that conjugates of BSA and polystyrene could successfully be synthesized in aqueous media in heterogeneous conditions using this reaction. Thus, we expected that the reaction of ODN with an azide or alkyn functionality would be feasible with the complementary azide or alkyn substituted polymer.

There is little known about the application of the “click” reaction with ODNs. Although this reaction may appear to be straightforward, success is limited by DNA damage due to radical formation of Cu(I) in solution in the presence of dioxygen. In an early report, Ju and co-workers reported the coupling of an azide ODN to an alkyn modified fluorescein label in solution. An aminohexyl appended ODN had first to be reacted with an azido acid via amide bond formation to provide the azide functionality. The thermally promoted variant of the reaction was used, which leads to a reaction without control over
regioselectivity and results in 1,4 and 1,5 triazole linkages, as opposed to the regioselective formation of a 1,4 triazole obtained by the Cu(I) catalyzed reaction at room temperature. Later, the same group reported the use of the Cu(I) catalyzed reaction to link ODN hairpins to glass chips functionalized with alkyne groups. These hairpins could be extended with the help of Thermo Sequenase DNA polymerase in the presence of nucleotides that are modified with photocleavable fluorophores. The partial sequencing of a 7 base fragment adjacent to the initial hairpin was possible because two nucleotides were linked to specific fluorophores. The procedure involved successive extension and fluorescence read-out steps, followed by cleavage of the fluorophore. This procedure was further improved to sequence a 14 base fragment, which was identified with labelled and photocleavable derivatives of all four nucleotides. The functionalization of azide modified self assembled monolayers on gold with alkyne functionalized ODNs was also demonstrated. In this paper, the author stressed the need for strong Cu(I) co-ordinating ligands such as TBTA to prevent DNA damage. Finally, Burley et al. took advantage of the commercial availability of alkyne functionalized nucleotide triphosphates which could be incorporated into DNA strands and amplified by PCR. Subsequently, an azide aldehyde was clicked to the DNA strands, followed by selective silver deposition at these sites with Tollens’ reagent. The click reaction and metallization procedure was carried out on a PAGE gel. Thus, all the examples known to date that employ the “click” reaction with ODNs are carried out on the solid phase.

Oligonucleotides which are appended with either azide or alkyne groups at the end are not yet commercially available through solid-phase synthesis techniques. Therefore, we looked at the functionalization of ODNs with either an alkyne or azide group. First, we employed the maleimide alkyne spacer 10, which was also used for the synthesis of the maleimide porphyrin labels (discussed in Chapter 6), to react with thiol functionalized ODNs. The Michael addition of thiols to maleimides has been shown to be very effective in the synthesis of DNA-PEG conjugates discussed in this chapter. This procedure was therefore also explored with alkyne maleimide 10 instead of maleimide polymer 1. Unfortunately, only half of the ODN was functionalized, as estimated by the intensity of the shifted band on a PAGE gel (Figure 4A, lanes 1 and 2). Subsequently, a “click” reaction was performed with azide coumarin, using a Cu(I) pentamethyldiethylenetriamine (PMDETA) complex as catalyst. No fluorescent bands could be visualized under UV light (330 nm) after the reaction, indicating that the “click” reaction had not been successful. Moreover, the applied reaction conditions caused a considerable loss of DNA material during the reaction (4A, lanes 3 and 4), most likely due to degradation of DNA by the copper catalyst. The experiment was repeated with an excess of different low molecular weight maleimide azides and alkynes (3-azido N-maleimido propane amine and N-maleimido propyn amine) and clicking of the corresponding ODN azides and alkynes to dansyl alkyne and coumarin azide, respectively, with similar negative results (data not shown). We then looked into different synthetic routes to modify ODNs with azides. An aminohexyl appended ODN was treated with an excess (50 eq) of NHS activated azido acid 11, and analyzed on a PAGE gel (Figure 4B).
A.

\[ (\text{CH}_2)_5\text{N}O\text{O} + \text{NH}_2 \text{O} \]

\[ (\text{CH}_2)_5\text{N}O\text{O} \]

\[ (\text{CH}_2)_6\text{O} \text{SP} \text{O} \text{O}^- \text{O} \]

\[ (\text{CH}_2)_6\text{O} \text{SP} \text{O} \text{O}^- \text{O} \]

B.

\[ \text{maleimide click} \]

\[ \text{Mw} \quad 1 \quad 2 \quad 3 \quad 4 \]

C.

\[ \text{one pot two step} \]

\[ \text{Mw} \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \]
of this reaction could not be determined, because the treated and untreated samples did not show a difference in mobility on the gel (4B, lanes 1 and 2). Alternatively, an amino azide spacer 12 was synthesized according to a literature procedure (see Chapter 6).46 Conjugation of this spacer to the 5' phosphate of an ODN via the procedure of Jeong and Park18 was expected to retard the ODN on a PAGE gel. This, however, was not found to be the case (Figure 4C). Moreover, a test reaction was performed with one third equivalent of tripropargylamine in the presence of a Cu(I) tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA) complex. This did not result in the formation of higher molecular weight DNA species (not shown).

These experiments point out the necessity of confirming the formation of the intermediate DNA alkyne or azide species, for instance via mass spectroscopy methods (Maldi-TOF or ESI). The first method was explored, but no spectra of the ODNs could be obtained. It is known that Maldi-TOF of ODNs is quite difficult and requires specific salt-free conditions. ESI mass spectroscopy was not available in our department at the time of these experiments.

2.4 Synthesis of conjugates via metal-free “click” reactions

The term “click” reaction was first coined for all reactions that are high yielding and proceed under ambient conditions without side products, but quickly has become a synonym for the Cu(I) catalyzed 1,3 dipolar cycloaddition reaction between an azide and alkyne.47 The success of this reaction lies in the high energies that are enclosed in the azide and alkyne functional groups, while the energy barrier for activation remains high. This barrier is lowered by the Cu(I) catalyst, however, the addition of this metal is generally not desirable, when applied to living cells, DNA, or proteins. This is due to the high toxicity of Cu(I) in living systems, caused by radical damage to DNA and proteins. The development of metal-free “click” reactions based on the Huisgen cycloaddition reaction will overcome these problems (Figure 5). Alkynes activated by electron-withdrawing substituents were found to readily give 1,4 triazole adduct with azides without the activation of a catalyst (Figure 5B).48 Good yields (67%) were obtained when a large excess of the small activated alkyne was

Figure 4, opposing page – Routes towards functionalization of ODNs for use in the “click” reaction. Gels are 15% non-denaturing PAGE gels stained with Stains-All. Mw is molecular weight marker. A. ODN functionalization via Michael addition of a sulphhydryl groups to a maleimide. Lanes 1 and 2 are unpurified and purified samples, respectively, from the reaction of the maleimide and the ODN. Lanes 3 and 4 are unpurified and purified samples, respectively, from the “click reaction” of the reaction mixture from lane 2 with a dansyl azide fluorophore. B. ODN functionalization via amide bond formation by treatment of amine functionalized ODN with active esters. Lane 1: ODN IV; Lane 2: ODN IV reacted with NHS azide 11. Lane 3: ODN V; Lane 4: ODN V reacted with NHS azide 11. C. ODN functionalization via functionalization of the 5’ phosphate. Lanes 1 and 2: 5’ phosphate functionalized with aminoazide 12 synthesized by the author and from commercial source, respectively, in the one-pot procedure. Lanes 3 and 4: same as for lanes 1 and 2, but with a purification step of the activated intermediate (see Materials and Methods for details). Lane 5: untreated ODN VI.
“Click” reaction types

<table>
<thead>
<tr>
<th></th>
<th>Reaction Equation</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.</td>
<td>( R_1^{\equiv} + N_3N-N_3^{=} \rightarrow R_2^4R_2^4 )</td>
<td>Sharpless et al.(^{38})</td>
</tr>
<tr>
<td>B.</td>
<td>( R_1^{=}O + N_3N-N_3^{=} \rightarrow O_NNN )</td>
<td>Ju et al.(^{48})</td>
</tr>
<tr>
<td>C.</td>
<td>( R_1^{=}O )</td>
<td>Bertozzi et al.(^{49})</td>
</tr>
<tr>
<td>D.</td>
<td>( R_1^{=}O )</td>
<td>Rutjes, Nolte et al.</td>
</tr>
</tbody>
</table>

**Figure 5** - Click reactions. A. Cu(I) catalyzed Huisgen 1,3 dipolar cycloaddition, leading to regiospecific formation of 1,4 triazoles. B. Metal free click reaction with electron deficient alkynes. C. Strain promoted click reaction of cyclooctene. D. Novel metal free click reaction of oxonorbadiene followed by a retro-Diels-Alder reaction. Reactions B-D are not regioselective. Only one region-isomer is shown for clarity.

incubated with 5’ azide labelled DNA. Nevertheless, the use of corresponding functionalized alkynes with electron withdrawing substituents has not yet been reported, which may point to difficulties in the synthesis of these compounds. Alternatively, strained alkynes also display reactivity towards azides (Figure 5C). A biotin functionalized cyclooctyne was used to label glycoproteins that were expressed in the presence of azide mannosamine.\(^{49}\) The strain-promoted click reaction was not as efficient as the corresponding Cu(I) catalyzed reaction, but the latter was found to diminish immunoreactivity of the protein in the Western blot analysis, pointing to epitope damage caused by Cu(I).

Recently, a novel metal-free click procedure was developed in our laboratory (Figure 5D). Complementary to the work of Ju and co-workers, who developed the procedure with electron withdrawing alkynes, it was found that the double bond in oxonorbordiene derivatives readily react with azides to give a triazoline ring, which subsequently undergoes a retro-Diels-Alder reaction. This procedure, which gives good yields when incubated with moderate excesses of azides, has already been applied to the functionalization of enzyme polymer hybrids. It was of interest to see whether this procedure could also be applied the more demanding click reaction with ODNs. To this end, two ODNs with aminohexyl functionalities were reacted with an acid-functionalized oxonorbordiene derivative. Two different procedures for the EDC mediated peptide bond formation were tested. First, the ODN was incubated with EDC and acid-oxonorbordiene directly. Alternatively, the acid-oxonorbordiene was first activated with EDC and NHS in the absence of ODN to give the NHS active ester, which was then added to the ODN. Subsequently, after purification of the reaction mixtures to remove excess reagents, azide functionalized PEG was added and the
DNA Block-copolymers as T4 clamp templates

Figure 6 – Functionalization of 5’ hexylamine functionalized ODNs IV and V with oxonorbornadiene 13. A. Reaction scheme. B. PAGE analysis. Lanes 1, 2 and 3 are samples of ODN IV, and lanes 4, 5 and 6 are samples of ODN V. Lanes 1 and 4: unfunctionalized ODNs. Lanes 2 and 5: sample of the reaction mixture. Lane 3 and 6: sample of the “click” reaction of the oxonorbornadiene functionalized ODNs with azide functionalized PEG (Mₙ 2,000).

reaction mixtures incubated for 20 hrs. PAGE analysis of the reaction products failed to identify any PEG-DNA conjugates (Figure 5). Additional (mass spectroscopy) analysis of the oxonorbornadiene derivative still has to be done to pinpoint the reason for this lack of reactivity: was the ODN not sufficiently modified with oxonorbornadiene or did the oxonorbornadiene derivative not react with the azide? Unfortunately, these experiments could not be performed due to lack of time.

3. Conclusions

Polymer-DNA conjugates were proposed to provide a template onto which the T4 clamp can be loaded. In the case of catalyst appended clamps, the clamp could slide over the polymer and perform oxidation reactions on the synthetic polymer block.

A growing number of DNA-polymer conjugates are being reported in the literature. Although these conjugates were initially synthesized on a solid phase support, procedures have now been developed to prepare conjugates in the liquid phase. The actual chemistry involved to link the ODN to polymers is strikingly diverse. Conjugates from a number of polymers, ranging from water soluble PEG to hydrophobic polystyrene, and using a variety of polymerization techniques, have been reported. Although the synthesis of the hydrophilic block-copolymers is easily achieved in the liquid phase, preparation of amphiphilic block-copolymers generally require solid phase protocols.

We started by synthesizing PEG conjugates of ODNs as a model system for the synthetically more demanding poly(butadiene)-ODN conjugates. The PEG conjugates were easily prepared via a Michael addition of thiol appended ODNs to PEG maleimides. Purification of the conjugates was achieved via size exclusion chromatography.
This methodology was subsequently applied to the preparation of 1,4 poly(butadiene)-DNA conjugates. First, a maleimide 1,4 poly(butadiene) polymer was synthesized in reasonable yields. Incubation of this polymer with thiol appended ODNs, however, did not result in the formation of any conjugate. Addition of THF to the reaction mixture to dissolve the polymer did not have any effect. It was reasoned that maleimide poly(butadiene) polymers in which the maleimide is separated from the polymer by a hydrophilic spacer could give rise to well-defined aggregates of the polymer in aqueous solution, facilitating the conjugation reaction. Unfortunately, such a polymer could not be synthesized.

In a separate line of research, we looked into the synthesis of DNA-polymer hybrids via “click” chemistry. This efficient cycloaddition reaction between azides and alkynes has already been used in the literature for the modification of surfaces with ODNs. Several procedures were tested for the introduction of azides or alkynes into ODNs. Modification of thiol appended ODNs with an alkyne maleimide derivative was partially successful and could be conveniently monitored because of the different mobilities of the starting material and alkyne functionalized ODN on a PAGE gel. A click reaction with a purified reaction mixture was not successful, however. Modification of amine or 5’ phosphate appended ODNs with different reactions were not successful either, as no product of a model “click” reaction with a fluorescent probe was detected. Finally, the application of a novel metal-free click reaction that was developed in our group did not yield DNA conjugates when reacted with azide functionalized PEG.

Thus, it may be concluded that the protocols for the synthesis of DNA-poly(butadiene) conjugates have to be improved, either by synthesizing the conjugate on the solid support, or by “click” chemistry. In the latter case, the modification of ODNs with azides or alkynes seems to be the limiting factor. It should be noted, however, that azide and alkyne derivatized ODNs are probably readily available directly when synthesized on a solid support using the corresponding azide and alkyne phosphoamidites.

4. Experimental section

General.
NMR spectra were taken on a Varian Inova 400 (400 MHz, 1H and 2D spectra) or on a Bruker DMX300 (75 MHz, 13C spectra) and calibrated to an internal standard of tetramethylsilane. Methoxypolyethylene glycol 5,000 maleimide was from Fluka. Amine functionalized 1,4 poly(butadiene) (MN = 4200, MW/MN = 1.05) was from Polymer Source (Toronto, Canada) and treated with an excess of TBAF in THF to achieve full deprotection of the amine, followed by precipitation in methanol. Maldi-TOF analysis showed that MN ≈ 3400 for the fully deprotected polymer, and contained ~18% of 1,2 polymerized butadienes (determined by NMR). Oligonucleotides were purchased from Biologio (Malden) or Isogen Life Science (IJsselstein) in the case of ODNs IV and V. ODNs were synthesized with high purity methods and purified by extraction of the main products from a PAGE gel by the supplier, and used without further purification. PAGE analysis was done with 15% non-denaturing PAGE gels in 1x TBE buffer (89 mM Tris base, 89 mM boric acid, 2 mM EDTA) and stained with Stains-All (Sigma) according to the directions of the supplier. Molecular weight standards were 25 bp step DNA ladders from New England Biolabs (Ipswich, MA, USA). Microspin G25
spin columns were from Amersham Biosciences (Little Chalfont, UK). Superdex G75 was obtained from Sigma. 3-azido N-maleimido propane amine, N-maleimido propyn amine, and the NHS ester of 5-azido caproic acid were gifts from Dr. Nikos Hatzakis. Coumarin azide was synthesized by Dr. Joe Sly. Acid functionalized oxanorbornadiene was a gift from Mr. Ton Dirks. All other materials were purchased from standard commercial sources.

### Oligonucleotides

<table>
<thead>
<tr>
<th>No.</th>
<th>Sequence (in 5’ to 3’ direction)</th>
<th>Modification</th>
<th>M_W</th>
<th>Melting temperature (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.</td>
<td>CCC CCC CCC CCC CCC</td>
<td>3’ C₆SH</td>
<td>4356</td>
<td>60</td>
</tr>
<tr>
<td>II.</td>
<td>GCG CTC TCC TGT TCC GAC CC</td>
<td>3’ C₆SH</td>
<td>6061</td>
<td>68</td>
</tr>
<tr>
<td>III.</td>
<td>GGG TCG GAA CAG GAG AGC GC</td>
<td>3’ C₆SH</td>
<td>6337</td>
<td>68</td>
</tr>
<tr>
<td>IV.</td>
<td>GCT GCA GGA ATT CGA TAT CA</td>
<td>5’ C₆NH₂</td>
<td>6204</td>
<td>55</td>
</tr>
<tr>
<td>V.</td>
<td>TGA TAT CGA ATT CCT GCA GC</td>
<td>5’ C₆NH₂ 3’ dabcy1</td>
<td>6218</td>
<td>55</td>
</tr>
<tr>
<td>VI.</td>
<td>GCT GCA GGA ATT CGA TAT CA</td>
<td>5’ PO₄²⁻</td>
<td>6221</td>
<td>58</td>
</tr>
</tbody>
</table>

### Synthesis and purification of DNA-PEG conjugates (I-1, II-1, III-1)

Oligonucleotide I (10 μl, 50 μg) was reduced with 0.1 M TCEP for 30 mins. Excess of the reducing reagent was removed via ethanol precipitation (see below), followed by incubation with 1 or 5 equivalents of maleimide functionalized PEG in 10mM Tris pH 8.0, 1 mM EDTA, for 16 hrs. DNA concentration was 0.2 μg/μl in 8.5 μl total volume. Conjugates (250 μl total volume of several samples) were purified on a Superdex 75 column (7 ml volume, Ø = 0.8 cm, l = 14 cm, in 25 mM Tris pH 7.6, 50 mM NaCl and 2 mM EDTA) connected to a peristaltic pump. Eluted fractions (0.5 ml volume) were analyzed on PAGE gel. This procedure was also applied to ODNs II and III, but in these cases the reduction protocol was improved as follows: The ODN was treated with DTT to a final concentration of 0.1 M for 30 mins, followed by purification with a G25 spin column. Conjugation and purification was performed as described above. The purified conjugates were annealed by heating an equimolar solution to 100 °C for 5 mins followed by cooling down to room temperature in 5 mins.

### General ethanol precipitation procedure for oligonucleotides

To the DNA solution, 0.1 volume 3 M NaOAc pH 5.3 and 2.5 volumes ethanol were added, and the solution was stored at -20 °C for longer than 30 mins. The solution was then centrifuged at 13,000 rpm for 15 mins and carefully decanted. The pellet was washed with 70% ice-cold ethanol followed by centrifugation at 13,000 rpm for 10 mins. The solution was decanted and the pellet is dried on the air and dissolved in a buffer of choice.

### Synthesis of maleimide functionalized 1,4 poly(butadiene) (4)

N-maleimido caproate (165 mg, 0.78 mmol) was dissolved in 5 ml distilled dichloromethane, followed by the addition of DIPEA (0.17 ml, 0.86 mmol) and disuccinimyl carbonate (220 mg, 0.86 mmol). This mixture was stirred in the dark for 2 hrs, after which a solution of amine functionalized 1,4 poly(butadiene) 2 was added, and the solution was stirred for 16 hrs in the dark. The reaction mixture was precipitated in methanol, centrifuged, decanted and dried under vacuum. The residue was subjected to column chromatography (methanol in dichloromethane, 1% v/v). Fractions containing product were isolated and
precipitated in methanol, followed by extensive drying of the pellet under vacuum. Yield was 35 mg (31%) of a brownish oil:

**Maldi-TOF MS** (dithranol / AgOTFA)

<table>
<thead>
<tr>
<th>m/z</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>3400</td>
<td>50 MN</td>
</tr>
</tbody>
</table>

**IR**

<table>
<thead>
<tr>
<th>ν</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1709 cm⁻¹</td>
<td></td>
</tr>
</tbody>
</table>

**1H-NMR**

<table>
<thead>
<tr>
<th>δ</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.68</td>
<td>s, 2H, CH₂</td>
</tr>
<tr>
<td>5.56</td>
<td>br m, integrates as 185 H, CH₂ from 1,2 polymer</td>
</tr>
<tr>
<td>5.40</td>
<td>br m, CH₂</td>
</tr>
<tr>
<td>4.96</td>
<td>br m, CH₂ from 1,2 polymer</td>
</tr>
<tr>
<td>3.22</td>
<td>br m, CH₂</td>
</tr>
<tr>
<td>2.05</td>
<td>br m, integrates as 418 H, CH₂</td>
</tr>
<tr>
<td>1.60-1.16</td>
<td>br m, 14H, CH₂</td>
</tr>
</tbody>
</table>

**Attempted synthesis of DNA-poly(butadiene) conjugates (III-4)**

To a solution of ODN III in 25 mM Tris.HCl pH 7.6, 50 mM NaCl and 1 mM EDTA, final concentration 1.0 μg/μl, 15 equivalents of maleimide functionalized 1, 4 poly(butadiene) were added in THF. When the experiments were done at 45% or 60% THF concentrations, more THF was added to the appropriate concentration. After one hr, the solutions were vortexed in an attempt to dissolve any aggregates, and incubated further for 16 hrs. Samples from each solution were analyzed on PAGE gel.

**Synthesis Boc-protected acid functionalized spacer (6)**

Mono Boc-protected diamine spacer (500 mg, 2.01 mmol) 5, of which the synthesis will be presented in Chapter 6, was dissolved in 10 ml distilled CHCl₃, followed by addition of succinic anhydride (1.00 g, 10.0 mmol). Distilled THF was added until a clear solution was obtained. After 1 hr, the solvent was evaporated. The residue was dissolved in CHCl₃ and the organic layer was extracted with 3 x 25 ml saturated NaHCO₃ solution. The water layer was acidified with aqueous KHSO₄ solution to pH ≈ 3 and extracted with 3 x 20 ml CHCl₃. Yield after drying of a brownish oil: 440 mg (62%), containing a impurity of the double Boc-protected diamine spacer:

<table>
<thead>
<tr>
<th>δ</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.45 and 6.97</td>
<td>br s, 2H, NH</td>
</tr>
<tr>
<td>3.62</td>
<td>s, 4H, OC₂H₂C₂H₂O</td>
</tr>
<tr>
<td>3.55</td>
<td>m, 4H, C₂H₂O</td>
</tr>
<tr>
<td>3.46</td>
<td>m, 2H, NHC₂H₂</td>
</tr>
<tr>
<td>3.33</td>
<td>m, 2H, NHC₂H₂</td>
</tr>
<tr>
<td>2.67</td>
<td>t, J=5.7Hz, 2H, CH₂</td>
</tr>
<tr>
<td>2.52</td>
<td>t, J=5.7Hz, 2H, CH₂</td>
</tr>
<tr>
<td>1.46</td>
<td>s, 9H, CH₃</td>
</tr>
</tbody>
</table>

**13C-NMR**

<table>
<thead>
<tr>
<th>δ</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>176.77</td>
<td></td>
</tr>
<tr>
<td>172.22</td>
<td></td>
</tr>
<tr>
<td>157.31</td>
<td></td>
</tr>
<tr>
<td>80.74</td>
<td></td>
</tr>
<tr>
<td>69.95</td>
<td></td>
</tr>
<tr>
<td>69.88</td>
<td></td>
</tr>
<tr>
<td>69.24</td>
<td></td>
</tr>
<tr>
<td>53.14</td>
<td></td>
</tr>
<tr>
<td>51.63</td>
<td></td>
</tr>
<tr>
<td>41.36</td>
<td></td>
</tr>
<tr>
<td>40.03</td>
<td></td>
</tr>
<tr>
<td>39.10</td>
<td></td>
</tr>
<tr>
<td>31.31</td>
<td></td>
</tr>
<tr>
<td>30.14</td>
<td></td>
</tr>
<tr>
<td>28.62</td>
<td></td>
</tr>
<tr>
<td>28.19</td>
<td></td>
</tr>
</tbody>
</table>

**Synthesis of NHS activated Boc-protected acid functionalized spacer (7)**

Acid functionalized Boc-protected spacer 6 (200 mg, 0.57 mmol) was dissolved in 20 ml CHCl₃, followed by addition of EDC (98 mg, 0.63 mmol) and N-hydroxy succinimide (73 mg, 0.63 mmol). The reaction was stirred for 16 hrs under nitrogen atmosphere. Then, the reaction mixture was extracted with 2 x 20 ml water and 1 x 10 ml brine, and the organic layer was dried with NaSO₄, concentrated and dried under vacuum. The residue was purified with silica column chromatography (eluent MeCN / CH₂Cl₂, 1:1 v/v). Yield after drying: 63 mg (25%) of a brownish oil with a minor impurity of N-hydroxyl succinimide:

<table>
<thead>
<tr>
<th>δ</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.48 and 5.08</td>
<td>br s, 2H, NH</td>
</tr>
<tr>
<td>3.61</td>
<td>s, 4H, OCH₂CH₂O</td>
</tr>
<tr>
<td>3.56</td>
<td>m, 4H, CH₂O</td>
</tr>
<tr>
<td>3.47</td>
<td>m, 2H, NHCH₂</td>
</tr>
<tr>
<td>3.31</td>
<td>m, 2H, NHCH₂</td>
</tr>
<tr>
<td>2.98</td>
<td>t, J=7.1Hz, 2H, CH₂</td>
</tr>
<tr>
<td>2.83</td>
<td>s, 4H, CH₂</td>
</tr>
<tr>
<td>2.64</td>
<td>t, J=7.1Hz, 2H, CH₂</td>
</tr>
<tr>
<td>1.44</td>
<td>s, 9H, CH₃</td>
</tr>
</tbody>
</table>

**13C-NMR**

<table>
<thead>
<tr>
<th>δ</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>171.46</td>
<td></td>
</tr>
<tr>
<td>169.83</td>
<td></td>
</tr>
<tr>
<td>168.42</td>
<td></td>
</tr>
<tr>
<td>167.51 (CO)</td>
<td></td>
</tr>
<tr>
<td>79.14 (CCH₃)</td>
<td></td>
</tr>
<tr>
<td>69.84</td>
<td></td>
</tr>
<tr>
<td>69.38</td>
<td></td>
</tr>
<tr>
<td>66.80</td>
<td></td>
</tr>
<tr>
<td>40.07</td>
<td></td>
</tr>
<tr>
<td>39.26</td>
<td></td>
</tr>
<tr>
<td>37.59</td>
<td></td>
</tr>
<tr>
<td>30.26</td>
<td></td>
</tr>
<tr>
<td>28.21</td>
<td></td>
</tr>
<tr>
<td>27.98</td>
<td></td>
</tr>
<tr>
<td>26.57</td>
<td></td>
</tr>
<tr>
<td>25.40</td>
<td></td>
</tr>
<tr>
<td>25.24</td>
<td></td>
</tr>
</tbody>
</table>

**Synthesis of spacer functionalized poly(butadiene) (8)**

Amine terminated 1,4 poly(butadiene) (75 mg, 0.02 mmol) was dissolved in 10 ml CHCl₃ after which NHS activated spacer-Boc 7 (48 mg, 0.11 mmol) was added. The reaction was stirred for 20 hrs under a nitrogen atmosphere in the dark. Then, 1 ml trifluoroacetic acid was added to the reaction mixture. After one hr the reaction mixture was extracted with 3 x 15 ml saturated NaHCO₃ solution and 1 x 15 ml brine. Yield after drying: 50 mg (63%) of a colourless oil, which was immediately used in the next reaction:

<table>
<thead>
<tr>
<th>δ</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.68 and 6.52</td>
<td>br s, NH</td>
</tr>
<tr>
<td>6.02</td>
<td>br s, 2H, NH₂</td>
</tr>
<tr>
<td>5.58</td>
<td>br m, integrates as 16H, CH from 1,2 polymer</td>
</tr>
<tr>
<td>5.42</td>
<td>br m, integrates as 23H, CH</td>
</tr>
<tr>
<td>4.96</td>
<td>br m, integrates as 41, CH₂ from 1,2 polymer</td>
</tr>
</tbody>
</table>
3.45 (m, 4H, CH₂), 3.22 (m, 2H, CH₂), 2.67 (t, J=6.8Hz, 2H, CH₂), 2.51 (t, J=6.8Hz, 2H, CH₂), 2.05 (br m, integrates as 47H, CH₂), 1.4-1.1 (m, integrates as 35H, CH₂, CH₃).

Synthesis of 1,4 poly(butadiene) spacer maleimide (9) Maleimide caproic acid (147 mg, 0.70 mmol) and DIPEA (0.13 ml, 0.80 mmol) were dissolved in 5ml CHCl₃, followed by the addition of disuccinimylcarbonate (180 mg, 0.7 mmol). The reaction was stirred under a nitrogen atmosphere for 2 hrs. Amine spacer functionalized 1,4 poly(butadiene) (95 mg, 0.028 mmol) was then added, and the reaction was allowed to proceed for 16 hrs under a nitrogen atmosphere and in the dark. The reaction mixture was precipitated into methanol, and the precipitate purified by silica column chromatography (eluent: CH₂Cl₂), followed by chromatography on a biobeads column (eluent: distilled THF). Yield: 39 mg (39%) of a colourless oil, which could not be identified as the product:

MALDI-TOF MS (dithranol / AgOTFA)⁵⁰ Mn m/z 3470; end group analysis m/z 496, expected m/z 481;
¹H-NMR (400MHz, CDCl₃) δ 5.58 (br m, integrates as 0.28H, C from 1,2 polymer), (br m, integrates as 2H, CH₂), 4.96 (br m, integrates as 0.32, CH₂ from 1,2 polymer), 2.05 (br m, integrates as 8H, CH₂), 1.60-1.16 (br m, integrates as 0.71H, CH₂, CH₃), small resonances are found at: 7.05, 6.75, 6.68, 4.15, 3.75, 3.66, 3.51, 3.27, 2.83, 2.76, 2.60.

Functionalization of thiol appended ODNs with maleimide alkyne spacer 10 and subsequent click reaction
Conjugation procedure: Thiol functionalized ODN III was reduced with DTT and purified using a spin column as described previously. The reduced ODN was incubated for 16 hrs with 20 eq of maleimide 10 in 25 mM Tris.HCl pH 7.6, 50 mM NaCl and 1 mM EDTA at a final concentration of 2.5 μg/μl.
Click procedure: To the maleimide functionalized ODN was added 50 eq of coumarin azide (10 μl of a 2.18 mg/ml solution), and 10 μl of a Cu(I) PMDETA solution (1.44 mg/ml CuBr, 2.1 μl/ml PMDETA). The reaction was performed in 25 mM carbonate buffer, pH 9.0, 50 mM NaCl, at a final concentration of 0.25 μg/μl ODN in a 50 μl total volume (total amount of DNA ≈ 2 nmol). The reaction mixture was purified after 16 hrs with a G25 spin column, and analyzed by PAGE.

Functionalization of amine appended ODNs with NHS activated azide 11.
A solution of ODN IV or V (0.3 μg/μl) in 25 mM sodium phosphate buffer pH 7.6, 5 mM EDTA, was treated with 50 eq of the NHS ester of 5-azido caproic acid in THF (total THF concentration in the reaction mixture was 20%). The reaction was incubated for 16 hrs in the dark, and analyzed by PAGE.

Functionalization of ODN 5’ phosphate with amino azide spacer 12 and subsequent click reaction.
Protocols were adapted from Jeong and Park.¹⁸
One-pot procedure: ODN VI was treated with EDC (1.25 mg, 6.5 μmol) and 5 μl of a 0.25 M solution of 1-amino 11-azido 3,6,9 trioxaundecane 12 from commercial source or synthesized according to the procedure described in Chapter 6, in 100 mM imidazole.HCl buffer pH 6.0. Total DNA concentration was 37.5 μg (6.1 nmol) in 32.5 μl. The solution was incubated in the dark for 16 hrs, and analyzed on PAGE gel.
Two step procedure: ODN VI treated with EDC (1.25 mg, 6.5 μmol) in 62.5 mM imidazole.HCl buffer pH 6.0 for 1 hr at room temperature, followed by purification with a G25 spin column. The purified solution was incubated with the amino azide spacer in 10 mM sodium phosphate buffer pH 7.5, 50 mM NaCl, 1 mM EDTA, for 16 hrs in the dark, and analyzed on PAGE gel.
Click procedure: Azide functionalized ODNs (1.13 nm) were dissolved in 75 mM sodium phosphate buffer, pH 8.0, and incubated with tripropargyl amine (375 pmol, 2.14 μl/ml) and various amounts of catalyst solution (5.1 mg CuSO₄.5H₂O, 11.75 mg TCEP and 21.76 mg TBTA in 1 ml H₂O) to give 0.1, 1.0 and 10 eq of Cu(I) to azide in the solution. Furthermore, each reaction was also performed with 12.5% t-butanol. All solutions (except ODN solution) were degassed before the reaction, and the centrifuge tubes flushed with Ar. The reactions were incubated for 16 hrs and analyzed using PAGE.
Synthesis of oxanorbornadiene 13 appended ODNs (IV-13) and (V-13) and subsequent click reaction

Two different protocols were investigated for the functionalization of the amine with oxanorbornadiene.

One pot procedure: ODN IV (24 nmol in 30 μl) was subjected to a spin column procedure and 6 μl of a 300 mM HEPES pH 7.6 was added, effectively changing the amine-based Tris buffer to a HEPES buffer. Then, 12 μl acid functionalized oxanorbornadiene 13 solution (16.5 mg/ml in 50% THF/MilliQ, 20 eq) and 12 μl of an EDC solution (7.6 mg/ml, 20 eq) were added. The reaction mixture (60 μl total) was incubated for 3 hrs at room temperature in the dark. The reaction mixture was then purified by using a spin column, supplemented with 10 μl 300 mM HEPES pH 7.6 buffer, and 50 μl of a solution of mono-azide functionalized PEG (MN = 2,000, 30 eq) was added. This reaction mixture, in which the DNA concentration was 0.22 nmol/μl, was incubated for 20 hrs in the dark, and analyzed on a non denaturing PAGE gel.

With pre-activation of the acid: Equal amounts of an acid functionalized oxanorbornadiene 13 solution (16.5 mg/ml in 50% THF/MilliQ, 20 eq) and an EDC solution (7.6 mg/ml, 20 eq) were reacted with each other in the presence of a few crystals of N-hydroxysuccinimide. After one hr, 24 μl of this solution was added to a solution of ODN V in 30mM HEPES pH 7.6 (prepared as described above). After two hrs, this solution was purified and reacted with azide functionalized PEG as described above.

5. References

DNA Block-copolymers as T4 clamp templates
Chapter 4

Interaction of the T4 DNA polymerase C-terminus with the T4 clamp

The C-terminal domain of the T4 DNA polymerase is essential for interaction of the polymerase with the T4 clamp, as is the case for many other clamp-interacting proteins. Two different binding sites for these interaction sequences on the clamp have been identified; peptide analogues of the polymerase C-terminus have been found to bind to the open subunit interface and the interdomain connecting loop of the clamp. The binding constant of the peptide for the first has been studied and determined. In particular, this chapter focuses on the influence of the presence of DNA or macromolecular crowding reagents on the binding process, monitored via FRET based titration experiments. The number of binding sites for the peptide on the clamp has also been investigated and will be discussed.
1. Introduction

If the concept of processive clamp-based oxidation catalysts for the conversion of DNA or polymer substrates is to be realized, the catalysts need to remain bound to its substrate as long as possible. Obviously, if such a catalyst would dissociate rapidly from its substrate, no processive oxidation reaction can occur and the system provides no advantage over simple homologous catalysts that act distributively. Several studies have indicated, however, that the T4 clamp-DNA interaction is not stable;1-4 in fact, dsDNA regions can only be populated by the clamp if the clamp is continuously reloaded from primer-template sites. This binding only seems to be extended to short stretches of DNA and requires that both ends of the region are blocked by proteins, such as the association of streptavidin to biotinylated oligonucleotides,5,6 or the inactivated EcoRI endonuclease mutant (Gln111)3 and the lac promoter binding2 with their respective recognition sites. DNase I footprinting experiments showed that under these conditions ~80% of the DNA is covered.3 Stopped-flow fluorescence spectroscopy on the forked DNA substrate indicated that the double stranded region of the substrate was readily populated and fully covered with 4 clamp proteins.6 Possibly, the interaction between neighbouring proteins within the array stabilizes the clamp proteins in the array, and hence also its presence on the substrate. It should be noted that these densely packed gp45 arrays are not favourable with respect to the proposed processive oxidation catalysts. Preferably, only a small number of oxidation catalysts should bind to each substrate and perform a series of isolated oxidation reactions.

In contrast to the labile binding of isolated gp45 proteins to DNA, interaction and co-tracking of clamp interacting proteins on DNA can lead to stable gp45-DNA complexes. This is indeed the case for a processive holoenzyme, in which gp43 is bound to the clamp via its C-terminus.7 The C-terminus is required for the formation of the polymerase holoenzyme,8,9 and the half-life of this complex on DNA is 7 – 8 mins (k_{off} = 0.001 s^{-1}).10 Interaction between gp45 and gp43 already takes place in the absence of a DNA substrate (K_D = 48 – 480 nM),6,11,12 and is shown to be accompanied with a decrease in the subunit interface of the clamp (see Chapter 3).6 In the presence of a DNA template and the clamp loader protein, a similar closing is observed during the process of holoenzyme assembly.6 The interactions between the T4 clamp and DNA polymerase however are considerably increased in the presence of a DNA substrate (K_D = 20 – 30 pM vs 48 – 480 nM, see Figure 1).5,12 The binding of gp43 to primer-template DNA substrates (K_D = 40 nM for gp43 exo−13 and K_D = 70 nM for gp45 exo+14) is increased by the presence of the clamp.14 The stability of the clamp on DNA is also increased 100 fold as a result of its association with gp43 (from k_{off} = 1 s^{-1} to k_{off} < 0.01 s^{-1}).4 This increase can be attributed to a cooperative interaction of the proteins with DNA. Gp45 is also an activator for T4 late transcription (DNA transcription associated with the late lifecycle stage of the T4 bacteriophage in infected E. coli cells, see Chapter 1).15 The T4 phage depends on the E. coli RNA polymerase for DNA transcription, but the T4 phage does express a transcriptional activator (the σ-family gp55 which replaces E. coli RNA polymerase
Interaction of the T4 DNA polymerase C-terminus with the T4 clamp

Figure 1 - Dissociation constants and kinetic parameters for the various complexes of the T4 clamp (gp45) and the T4 DNA polymerase (gp43) with each other and DNA substrates. For references to the experimental data, see the text.

σ-subunit) and a transcriptional co-activator (gp33, which recognizes transcription promoter sequences). In contrast to DNA replication, DNA transcription does not start from primer-template junctions but from transcription promoter sequences (TATA boxes), and the RNA polymerase holoenzyme has to be assembled at these sites. It was found that gp45 has to be loaded onto DNA upstream of the promoter sequence, followed by migration of gp45 to this sequence, a distance that can span hundreds of base pairs. In this process, gp55 binds with its terminus to gp45 and the gp55-gp45 complex migrates along DNA until it finds a RNA polymerase bound to a promoter site. The half-life time of gp45 on DNA increases four fold when it interacts with gp55 (t½ = 4.8 min vs t½ = 1.2 min). Generally, all T4 proteins that interact with gp45 in similar ways share a common [S/T]LDFL sequence at their C-terminus (Figure 2). Surprisingly, some proteins that were found to bind to gp45 do not appear to have such a sequence (such as α-glycosyltransferase or gp44/62), or even have a recognition sequence on their N-terminus (RNaseH). The reasons for this are not yet fully understood. Interestingly, a BLAST database search also identified the interaction sequence in the middle of T4 DNA ligase, but since this protein has not been crystallized it is impossible to say whether this sequence is available for interaction with the clamp and whether it has any relevance.

Based on the stabilizing effect of gp45-interacting proteins on the gp45 – DNA complex, we reasoned that a peptide analogue of the C-terminus interaction domain should also enhance the life-time of gp45 when on DNA (see Chapter 2). Therefore, C-terminus
A. Peptide
   T4 DNA polymerase (gp43)
   T4 transcriptional co-activator (gp33)
   T4 σ-family protein (gp55)
   T4 RNase H

   RB69 DNA polymerase


   CSDFLFG
   EKASLFDFLFG
   EKTNTLDFLL
   DDSFSDLFLYFEND
   MDLMLMLDEK note: N terminus

B. C.

   Figure 2 - A. Sequence similarities between the C-termini of several gp45-interacting proteins. Recognition sequence is in bold. RNaseH can bind gp45 via its N-terminus.20 B. Model of the peptide (black) binding in the open subunit interface of gp45 (each subunit is coloured differently). Structure derived from the holoenzyme model.22 C. Crystal structure of the RB69 clamp complexed with the peptide in the interdomain connecting loop. For reasons of clarity, the protein and peptide are depicted as a tube and ribbon model, respectively. Sequence of the peptide used was KKA SLFDMFDF, which are the 11 C-terminal amino acids from the RB69 DNA polymerase. The RB69 clamp shares 77% sequence identity with T4 gp45, while RB69 polymerase shares 69% sequence identity with T4 gp43. PDB code of this crystal structure is 1B8H.23

peptide analogues were included in the assembly strategy of the clamp-based catalysts as discussed in Chapter 2. It should be noted, however, that the stabilizing effect of the peptide may not be as pronounced as in the case of the T4 holoenzyme, because the peptide has no affinity for DNA. Furthermore, the peptide may not have the ideal conformation to interact with the subunit interface. Indeed, circular dichroism experiments showed that the peptide is in a random-coil conformation (data not shown), however, the C-terminus in the crystal structure of the polymerase from RB69, which is closely related to that of T4, is largely unstructured as well.24

The interaction of a small peptide that corresponds to the C-terminus of gp43 with gp45 has been studied in the groups of Benkovic and Steitz, and led to the development of two slightly different models for the interactions between gp45 and gp43. The peptide was first shown by Alley et al. to interact with gp45, having a K_D of 7.1 μM.7 Cross-linking experiments pinpointed this interaction to the subunit interface of the clamp (Figure 2B). In contrast, a C-terminus peptide analogue of the T4-related RB69 DNA polymerase was co-crystallized with the RB69 clamp, and the resulting crystal structure showed that the peptide was bound to a hydrophobic patch at the interdomain connecting loop of the clamp located
on the “rough” side of the clamp (Figure 2C). The reason for this alternate binding could be that the crystal packing forces expel the peptide from the subunit interface to a secondary binding site, as the clamp was found to be a closed ring in the crystal structure, similar to the clamp crystallized in the absence of the peptide. On the other hand, there is strong evidence that multiple proteins can bind to the T4 clamp simultaneously. Since only one of the three subunit interfaces of gp45 is open in solution, only one peptide can bind to gp45 in that site, and interactions of additional proteins with gp45 may therefore very well take place at the secondary binding site of the clamp. The interdomain connecting loop is the primary binding site for other clamp proteins such as PCNA. For the T4 system, it was suggested that gp43 first binds to the interdomain connecting loop, and that additional interactions with the rough side of the clamp subsequently redirect the polymerase to the subunit interface. Such a process has been shown to be important in E. coli, where lesion bypass polymerases bind to a secondary binding site on the clamp but displace the replicating DNA polymerase from the primary binding site when the polymerase complex is stalled due to DNA damage. Intriguingly, the crystal structure of the T4 clamp showed that only one peptide was bound to the clamp trimer, although the complex was crystallized in the presence of an excess of peptide. This may indicate that an excess is necessary to saturate all binding sites for the peptide on the clamp, or that only one interdomain connection loop in a gp45 trimer is able to accommodate a peptide. This observation could be the result of the slightly asymmetric conformation of gp45 found in its crystal structure. In contrast to the above, a crystal structure of the p21 C-terminus bound to the more symmetrically human PCNA did show full saturation of all binding sites.

Although the recent studies described above have considerably improved our understanding of the T4 holoenzyme structure, the complex interplay between the clamp and its binding partners is not understood in its fullest details. In this chapter, we focus therefore on the putative stabilizing effect of the peptide on the clamp-DNA complex, and try to determine the number of binding sites for the peptide on the clamp. Different experimental techniques will be used to address this question.

2. Results and discussion

2.1 FRET experiments

The binding constant of the C-terminus peptide analogue of gp43 (hereafter called the peptide) with gp45 was first determined by Alley et al. from a FRET experiment in which the tryptophan (W92) near the subunit interface was quenched after the addition of small aliquots of a dansyl labelled peptide. The titration was performed at a gp45 trimer concentration of 100 nM. We thought it to be of interest to investigate whether the binding constant would change if the titration was performed under different experimental conditions. For instance, would the addition of DNA to the solution affect the binding constant? In principle, such a difference can be expected because interaction of the clamp with DNA leads to a contraction of the clamp, caused by electrostatic interactions between
the positively charged central cavity of the clamp and the negatively charged phosphate backbone of DNA. As a result, the distance across the open subunit interface of the clamp becomes smaller, which could make the binding of the peptide to the subunit interface stronger. On the other hand, it has been shown that the twisted conformation of the clamp, imposed by interaction with the clamp loader in the presence of ATP, is necessary for the clamp to interact with gp43. This may imply that the C-terminus cannot readily interact with the clamp, or that this conformation is necessary for gp43 to interact correctly with the rough side of gp45. Furthermore, would the use of crowding agents have an effect of the binding constant? As will be discussed below, the addition of high concentrations of hydrophilic polymers have a considerable effect on protein–protein and protein–DNA interactions. Pronounced differences in the measured binding constants in these different scenarios may provide insight into the binding of the clamp to DNA and what the role is of the peptide in this process.

The gp43 C-terminus peptide analogue (CSLDFLFG) was synthesized using standard solid phase peptide synthesis techniques. Initially, the cysteine was protected with an S-tert-butyl group, which is not removed during the cleavage of the peptide from the resin. The disulfide bond however could not be easily removed after isolation from the peptide resin.

Figure 3 - Labelling of the gp43 C-terminus peptide analogue with a dansyl fluorophore. A. The free thiol of the peptide is reacted with an iodoacetamide dansyl compound (IAEDANS). B. HPLC chromatogram of the unlabelled and labelled peptide reaction mixture (black and grey, respectively). C. Maldi-TOF spectrum of the main peak in the HPLC chromatogram, corresponding with the labelled peptide.
and the product turned out to be rather insoluble in aqueous solvents. The peptides were therefore re-synthesized using a trityl protected cysteine, which is removed during the cleavage from the resin. The peptides were labelled according to a protocol described by Alley et al. in a HEPES buffer, followed by purification using reversed phase HPLC (Figure 3). Injection of the peptide in HEPES buffer onto the analytical HPLC column resulted in a multitude of peaks due to overloading of the column. This problem was circumvented by using a phosphate buffer, and labelling with iodoacetamide dansyl (IAEDANS) in this buffer was found to proceed to completion within 10 mins. Purification required a preparative HPLC column with a large diameter, typically used for the purification of 50 fold more material. Nevertheless, injection of more than a few milligrams of peptide onto this big column resulted in overloading and splitting of peaks. Small amounts of the labelled peptide could be purified, and Maldi-TOF analysis of the purified peptide showed quantitative and unique labelling of the cysteine in peptide.

The dansyl labelled peptide was used to determine the binding constant of the peptide to gp45 (Figure 4). Unfortunately, initial experiments were frustrated by a decrease in tryptophan fluorescence caused by mixing gp45 solutions in the cuvette by pipetting. This phenomenon was attributed to non-specific binding of gp45 to the walls of the cuvette and the pipette tips. BSA was therefore added to prevent the absorption of gp45, but this turned out to be unsuitable because of the fluorescence of tryptophans in BSA and possible binding of the peptide in one of the hydrophobic patches of BSA. Finally, three different countermeasures were taken to diminish the non-specific binding of gp45. First, the cuvette was silanized with dichlorodimethylsilane to render the walls hydrophobic. Second, Tween 80, a non-ionic sugar based surfactant without UV absorption at 280 nm, was added to the solution to a concentration of 0.002 wt% (the critical micellar concentration of Tween 80 in water). Third, the titration was performed with siliconized (low protein absorption) pipette tips, and the tip with which the solution was mixed after the titration was not changed.

**Figure 4** – Fluorescence titrations. A. Schematic representation of the titration. Tryptophan (W92) is excited and quenched if a dansyl (box) labelled peptide (rectangle) binds to the subunit interface. B. Typical titration curve obtained by performing the experiment described under A.; the tryptophan fluorescence is monitored at 335 nm.
**Table 1** - Calculated dissociation constants of complex between the peptide and gp45 at different concentrations

<table>
<thead>
<tr>
<th>[gp45]</th>
<th>118 nM</th>
<th>235 nM</th>
<th>353 nM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_d$ (± SD) M</td>
<td>$7.73 \times 10^{-6}$ (0.61)</td>
<td>$8.26 \times 10^{-6}$ (0.47)</td>
<td>$8.60 \times 10^{-6}$ (0.29)</td>
</tr>
</tbody>
</table>

during the experiment. These precautionary measures all together were found to inhibit the non-specific binding of gp45. Prior to the experiment, the solution was mixed with the help of a pipette until a stable tryptophan emission intensity was obtained (typically around 80% of the initial fluorescence).

The results of the titration experiment with the labelled peptide and gp45 are listed in Table 1. Under identical conditions as reported by Alley *et al.*, the dissociation constant of the complex between peptide and the clamp was 7.7 μM when fitted to a 1:1 binding model, in good agreement with the value reported in the literature ($K_D = 7.1$ μM). The concentration of gp45 used in these experiments is well below the estimated dissociation constant for gp45 itself (250 nM), although clamp trimer association from monomers could be better described by a cooperative model with the product of equilibrium constants being 0.08-0.21 μM$^2$ (see Figure 5). The question which arises is whether these experiments actually measure the binding of a peptide to a single gp45 subunit, or to the subunit interface (which requires the assembly of a gp45 trimer). If bound to a single gp45 subunit, the peptide could bind to the interdomain connecting loop, or to the sides of the protein (Figure 6). The latter interaction would most likely be weaker than the interaction with the open subunit interface because only one side of the peptide can interact with the protein. On the other hand, the formation of the trimer seems to be templated by the addition of clamp binding proteins, because a quantitative 1:1 binding of gp45 trimer to gp44/62 was observed with concentrations as low as 60 nM for gp45. Possibly, the peptide would have the same effect and induce the trimeric form of the clamp.

**Figure 5** – Assembly of gp45 subunits into gp45 trimers. The curves involving cooperative models are plotted with products of binding constants of 0.08 μM$^2$ and 0.21 μM$^2$; the non-cooperative model is plotted with $K_D = 250$ nM (concentration in gp45 trimers, thus $K_D = 750$ nM in monomers). The concentrations at which 50% of the total amount of gp45 is a gp45 trimer are 285 nM, 460 nM and 750 nM, respectively.
To address these issues, higher concentrations of gp45 were also used in the titration assay. Performing the titration experiments at higher gp45 concentrations (235 nM and 353 nM) resulted only in a slight increase in binding affinity when fitted to a 1:1 binding model (Table 1), which makes it hard to draw any firm conclusions and differentiate between the two binding mode scenarios (Figure 6). In addition, because of the uncertainty in the constants of the cooperative gp45 association models, it is hard to determine the exact ratio of gp45 trimers to gp45 monomers or dimers. Nevertheless, the models were used to determine a window of ratios, which were estimated to be 49 ± 17% trimers at 118 nM, 79 ± 11% trimers at 235 nM and 89 ± 6% trimers at 353 nM (concentrations expressed as one third of gp45 monomer concentration).

Figure 6 – Simplified model for the binding of the gp43 C-terminus analogue to gp45 monomers, dimers and trimers. At low concentration, about half of gp45 exists as a monomer and can bind a peptide in the subunit interface and at the interdomain connecting loop. If peptide binding would invoke the formation of a gp45 trimer (templating mechanism), it would probably only be able to do so via the binding at the interdomain connecting loop. At intermediary concentration, the state of gp45 is represented as a dimer, but it should be noted that the concentration of dimers is probably very low because the cooperative process of gp45. Nevertheless, three binding modes can be imagined; two discussed above and the putative binding of a peptide to the closed subunit interface, made possible because of the absence of strain in the gp45 ring. At high gp45 concentrations it is present as a trimer, which could bind a peptide in the open subunit interface or at the interdomain connecting loop. Multiple peptides can be bound, theoretically to a maximum of 4 per gp45 trimer.

The binding of the peptide to gp45 was also investigated using macromolecular crowding conditions. The addition of high concentrations of hydrophilic polymers to the solution promotes the formation of protein-protein and protein-DNA interactions because of excluded volume effects. Experiments reported in the literature demonstrated that the addition of 7.5% PEG (M_N = 12,000) increased the binding affinity of gp45 for gp44/62 100 fold, furthermore, macromolecular crowding enabled the formation of a functional
hologenzyme without the help of gp44/62 or ATP hydrolysis. This makes this approach particularly interesting for the binding of the catalyst appended clamps onto DNA or polymer substrates.

Surprisingly, the results in the solutions containing PEG (MN = 8,000) indicate that the affinity of the peptide for the clamp is lower in the presence of the polymer. It is known that the crowding effect is more pronounced for larger molecules (such as proteins), because the excluded volume upon binding is higher in these cases. Thus, a high increase in the binding affinity of the peptide for the clamp may not be expected. The decrease in binding affinity suggests that either the polymer is blocking the binding sites on the protein, or that the increased viscosity has a detrimental effect on the binding process.

The addition of DNA (supercoiled pBR322 plasmid) to the solution did not seem to influence the binding of the peptide to the clamp. Similar dissociation constants are obtained either with or without PEG concentrations of 10%, the concentration at which the maximal crowding effect is expected. Thus, it can be concluded that either binding of the clamp to the DNA substrate does not affect the affinity of the peptide for the clamp, or, effective loading of the clamp onto the DNA substrate is not achieved under macromolecular crowding conditions and hence the binding affinity is not affected. Possibly, PEG does not promote gp45 - DNA binding because there is a water layer between the inner ring of gp45 and DNA. This water layer is probably also present in unbound gp45, or water molecules are restricted in movement because of electrostatic interactions. Thus, there is no entropic gain when gp45 binds to DNA because these water molecules are not liberated upon binding.

The binding of the peptide to the clamp was further analyzed with the help of Scatchard plots, in order to determine the number of binding sites. Although the 1:1 binding model was found to describe the fluorescence change very well, and other simple models failed to produce good fits, experimental evidence indicates that there are multiple binding sites for the proteins on one single clamp trimer. As a consequence, multiple peptides may also bind to a clamp trimer. In the crystal structure of the peptide complexed to the interdomain connecting loop of the clamp, the distance between the β carbons of tryptophan (W199) and the alanine in the peptide (which is replaced by the cysteine in our case), is 13 Å, well within the Förster radius reported for the tryptophan-AEDANS donor-

---

**Table 2** - Calculated dissociation constants for the binding of the peptide to the clamp under macromolecular crowding conditions and in the presence or absence of DNA.

<table>
<thead>
<tr>
<th>[gp45]</th>
<th>No DNA $K_d$ (M ± SD)</th>
<th>With DNA $K_d$ (M ± SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0% PEG$^a$</td>
<td>7.73·10$^{-6}$ (0.61)</td>
<td>5.57·10$^{-6}$ (0.74)</td>
</tr>
<tr>
<td>5% PEG</td>
<td>8.57·10$^{-6}$ (1.00)</td>
<td>-</td>
</tr>
<tr>
<td>10% PEG</td>
<td>1.52·10$^{-5}$ (0.32)</td>
<td>1.56·10$^{-5}$ (0.31)</td>
</tr>
</tbody>
</table>

$^a$ PEG MN = 8,000
acceptor pair (22 Å). In comparison, the distance between W92 and the alanine in the C-terminus of gp43 is 15 Å in the holoenzyme model. Thus, when the binding affinities of the peptide for the subunit interface and the secondary binding site on the clamp are similar, tryptophan quenching by peptides binding in the interdomain connecting loop will be significant and the 1:1 binding model may not be the best model to describe the binding of the peptide to the clamp. The analysis presented here can also contribute to the discussion whether multiple proteins can bind to the clamp, and elucidate the stoichiometry of the resulting complexes.

Figure 7 - Scatchard plot analysis of the fluorescence titrations. A. In a Scatchard plot the binding ratio \( \nu \) (which equals \([P]_{bound} / [E]_{total}\)) divided by \([P]_{free}\) is plotted against \( \nu \). An ideal plot gives a straight line with slope \(-1/K_A\) and the number of binding sites (n) as the intercept of the plot with the x-axis. B. Four different traces are shown for the experiment performed at 118 nM gp45, without PEG and DNA. Fitting of the linear parts of the curves led to the estimation of 6.54 binding sites of the peptide in the protein (dotted line). Deviation from the linear part, seen in two of the curves, may suggest binding sites with different affinities.

Scatchard plots were made for all the binding curves described above, and the results are listed in Table 3. The analysis is complicated by the fact that gp45 is a dynamic system at the concentrations used (see Figure 3, and discussion above), and the outcome of the Scatchard plots is extremely sensitive to changes in protein concentration. To cope with the uncertainty in gp45 trimer concentration and to be able to compare the experiments, the number of binding sites is given per gp45 monomer. The number of binding sites obtained in all experiments is significantly higher than expected. The reason for this is unclear at the moment, but it is known that linearization of non-linear data may result in large errors. Another possible explanation lies in the fact that Scatchard plots require the determination of the change in fluorescence intensity upon binding of the peptide to the clamp. This can be determined by titration of very small amounts of peptide to the protein, since the protein is present in large excess to the peptide, all the peptide will be bound to the protein. In the case of gp45, these experiments were frustrated by small changes in fluorescence intensity due to non-specific absorption of the protein. Since the change in fluorescence by peptide binding is very small, non-specific absorption could lead to significant errors in the determination of the stoichiometry of binding. If loss of fluorescence due to non-specific absorption at these small concentrations plays an important role, a
fraction of the quenching could be incorrectly attributed to fluorescence quenching by peptide binding, and consequently, the concentrations of bound peptide and the binding affinity would be overestimated. The number of binding sites is calculated by dividing the concentrations of bound peptide by total protein, hence the above complication may lead to an overestimation of the amount of binding sites. Furthermore, the intersection of the fitted line with the y-axis provides an estimation of the association constant. It is apparent that this constant is too high when compared to the dissociation constants determined by the fluorescence titrations. These experiments could not be corrected for the non-specific absorption process, as this process was not found to be very reproducible.

Table 3 - Scatchard plot results for the binding of the peptide to the clamp.\textsuperscript{a,b}

<table>
<thead>
<tr>
<th>[gp45]</th>
<th>118 nM</th>
<th>Without DNA</th>
<th>235 nM</th>
<th>353 nM</th>
<th>With DNA</th>
<th>118 nM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>K\textsubscript{D}</td>
<td>n</td>
<td>K\textsubscript{D}</td>
<td>n</td>
<td>K\textsubscript{D}</td>
</tr>
<tr>
<td>0% PEG</td>
<td>6.54 (0.54)</td>
<td>0.36 (0.12)</td>
<td>2.74 (0.1)</td>
<td>0.28 (0.05)</td>
<td>2.39 (0.56)</td>
<td>0.21 (0.08)</td>
</tr>
<tr>
<td>5% PEG</td>
<td>5.10</td>
<td>0.50</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10% PEG</td>
<td>7.6 (2.2)</td>
<td>0.17 (0.11)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Standard deviations are given within parenthesis.

\textsuperscript{b} Dissociation constants are given in \textmu M.

Although the exact number of binding sites may not be correct, the trends these numbers provide could still provide useful information. First, under conditions without macromolecular crowding, the number of binding sites is significantly lowered by increasing the protein concentration from 118 nM to 235 nM. This effect cannot be attributed to the protein concentration itself, because a further increase to 353 nM does not lead to a further decrease of the number of binding sites. A possible explanation is that at lower concentration, gp45 exists as monomers and dimers in solution, which leaves more subunit interfaces exposed for binding. Additionally, in the presence of DNA the number of binding sites will also be lowered. These observations suggest that the clamp interacts with DNA, and that this interaction leads to loss of some binding sites. The use of crowding agents surprisingly does not seem to affect the number of binding sites. As the Scatchard plot experiments were plagued by experimental difficulties, caution should be taken with any of the conclusions derived from these experiments. Nevertheless, the results presented here suggest that multiple binding sites do play a role in the interaction of the peptide with gp45. With increasing gp45 concentration, there seems to be a switch between a state in which gp45 can bind more peptide molecules to a state in which the binding is lower. This could be correlated with the increase in the fraction of gp45 trimers in that case. Furthermore, the Scatchard plot show a deviation from the linear part at higher binding site numbers, which also could indicate the occurrence of multiple binding sites.
2.2 ATPase stimulation studies

The interaction between the peptide and gp45 was also studied by the ATPase assay using a small forked DNA substrate, which mimics clamp loading sites in vitro (Figure 8). Addition of gp43 inhibits the synergistic ATPase activity of gp44/62 in the presence of gp45 and a DNA substrate, indicative for the formation of the more stable holoenzyme complex on DNA. If the peptide is able to stabilize the gp45 – DNA complex, the ATP hydrolysis rate should decrease when the peptide is titrated into the solution. This was indeed found to be the case (Figure 8B), and ratio between the initial ATPase activity and measured activity (the

![Diagram of ATPase assay and binding constant determination.](image)

**Figure 8** - Determination of the binding constants of the peptide – clamp complex. A. Schematic representation of the ATPase assay. The clamp (gp45, ellipse) is loaded onto the forked substrate by the clamp loader (gp44/62, rectangle) by ATP hydrolysis. Addition of the peptides to the gp45 - gp44/62 - DNA complex results in a more stable gp45-DNA complex accompanied by lower ATP hydrolysis rates. B. The relative decrease of ATP hydrolysis plotted against peptide concentration was fitted to a 1:1 binding model to give binding constants (K_D) for the clamp-peptide interaction listed in C.

<table>
<thead>
<tr>
<th>Label on peptide</th>
<th>K_D of clamp-peptide complex (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>none</td>
<td>1.59 ± 0.04</td>
</tr>
<tr>
<td>dansyl</td>
<td>2.07 ± 0.37</td>
</tr>
<tr>
<td>fluorescein</td>
<td>3.87 ± 0.26</td>
</tr>
</tbody>
</table>
shutdown ratio) was plotted against the concentration of the peptides. These curves could be fitted to a 1:1 binding model, although the fit was not as good as for the fluorescence titrations. This can be attributed to a decrease in gp45 concentration, which results in less gp45 trimers and lower ATPase rates than expected. The resulting dissociation constants were comparable to the ones obtained from the FRET experiment. The $K_D$ of the unlabelled peptide was slightly lower than the $K_D$’s of the labelled peptides, which may be attributed to steric hindrance of the label and the protein resulting in lower binding strengths for the latter peptides. This assumption is confirmed by the observation that the larger fluorescein label resulted in a higher $K_D$ than the smaller dansyl label.

These experiments do not unambiguously demonstrate a stabilizing effect of the peptide on the clamp - DNA complex. In fact, Berdis et al. attributed the loss of ATPase activity upon titration of the peptide to a gp44/62-gp45 solution in the presence of an unblocked double stranded DNA template to an inhibition of the gp44/62 - gp45 complex. To verify this result, the peptide was titrated into a solution of the gp44/62 - gp45 complex without DNA template present and it was indeed found that it inhibited ATPase rates (data not shown). Possibly, the peptide binds to a site, presumably the interdomain connecting loop that is important for the interaction of the clamp loader with the clamp. This is also observed for the $\delta$-subunit of the *E. coli* clamp loader ($\gamma$-complex), which binds to a site on the $\beta$-subunit that is near a boundary of different domains, and opens one of the two subunit interfaces of $\beta$. Inhibition by peptide binding in the subunit interface seems much more unlikely, but cannot be ruled out. The low ATPase rates (~70 - 80 nM/s) and the error therein prevented an exact number (inhibition constant) to be obtained, but the decrease in the ATPase rates are comparable to the decrease observed in the presence of a DNA template. Surprisingly, the addition of gp43 to the pre-formed gp45 – gp44/62 complex in the absence of a DNA substrate did not lead to a decrease in ATPase rates. Thus, the gp45–interaction domain present in gp43 does not interfere with the binding of gp44/62 to gp45. We postulate that conformational changes associated with gp44/62 binding to DNA templates results in the possibility of gp43 to interact with gp45.

3. Conclusions

The binding of a peptide analogue of the T4 DNA polymerase C-terminus and the T4 clamp was investigated with the help of fluorescence spectroscopy and ATPase assay titrations. The polymerase C-terminus mediates the interactions between the polymerase and the clamp, which take place in the open subunit interface of the clamp. The complex of the clamp on DNA is dynamic, however, its life-time can significantly be enhanced by interaction with clamp binding proteins such as the polymerase. It was proposed therefore, that the addition of the peptide to the clamp would also increase the stability of the clamp on DNA. This is essential for successful processive catalysis to be carried out by artificial clamp based
catalysts that are the topic of this thesis. The dependence of the peptide-clamp interaction on the presence of DNA and macromolecular crowding agents was therefore investigated.

It has been suggested that multiple proteins can interact with the clamp at the same time, and other clamp-interacting proteins are also known to possess homologous C-termini. The clamp can, however, have only one open subunit interface available for interaction at a given time, hence the additional interactions have to take place at secondary binding sites. There is strong and accumulating evidence that these secondary binding sites indeed exist, although their exact nature has not yet been clarified. The experiments described in this chapter were designed to also shine light on these secondary interaction sites.

Fluorescence titrations, in which a dansyl-labelled peptide is titrated to the clamp, were carried out using a literature procedure, and the resulting tryptophan fluorescence quenching was fitted to a 1:1 binding model. It was found that the dissociation constants did not change considerably over a concentration range of gp45, suggesting that the peptide can template the formation of a gp45 trimer making that in each case the same interaction is measured, or that the affinities of the interdomain connecting loop and the open subunit interface for the peptide are similar. Furthermore, the addition of PEG as macromolecular crowding agent resulted in weaker interactions, contrary to the beneficial effect of these agents on protein-protein and protein-DNA interactions. Lastly, the presence of a dsDNA template in the titration solution did not have any effect on the dissociation constant. Either the clamp does not bind to DNA under these conditions, or the interaction of the clamp with DNA does not significantly change the binding affinity of the peptide for the clamp.

To validate the 1:1 binding model and to identify any secondary interaction sites, the number of binding sites was determined with the help of a Scatchard plot. The total number of binding sites was found to be in the range of 2-9 depending on the experimental conditions used. Most likely, these numbers are too high to have any biological significance, but the data do suggest the presence of more than one binding site for the peptide on the clamp. The Scatchard plots were able to sense a decrease in the number of binding sites at higher gp45 concentrations, which could reflect the presence of more gp45 trimers.

Finally, the involvement of the peptide in the formation of the T4 holoenzyme was investigated with the help of the ATPase assay. It was found that the peptide is able to lower the rate of ATP hydrolysis when added to the gp45-gp44/62-DNA complex, much like the T4 polymerase but with lower efficiency. The resulting traces could again be fitted to a 1:1 binding model, and yielded binding constants of similar values as those determined by the fluorescence spectroscopy experiments. The presence of a fluorescent label on the peptide was found to inhibit the peptide-clamp interaction, presumably due to steric interactions between the label and the protein. Thus, the dissociation constants determined by the fluorescence spectroscopy experiments may be somewhat underestimated compared to the unlabelled peptide. The peptide was also found to inhibit the ATPase rates of gp45-gp44/62 complexes with similar efficiency, indicating that the binding sites of the peptide on the clamp play an important role in the interaction between the clamp and the clamp loader.

In conclusion, the binding of the peptide to the clamp follows very complex mechanistic patterns (Figure 6), in which there are binding sites of different nature on the
clamp, and the stoichiometry of binding sites is concentration-dependent. Unfortunately, the experiments presented here have failed to elucidate this complex picture completely.

With respect to the aimed clamp-based oxidation catalysts, the experiments described here cannot categorically demonstrate that the clamp-DNA complex is stabilized by addition of the peptide. Although a situation in which the addition of the peptide partially inhibits the loading of the clamp onto DNA but also stabilizes the clamp on DNA is imaginable, the peptide probably lacks the additional interactions of clamp interacting proteins in order to be able to have a pronounced effect on the life-time of the clamp on DNA. This is also demonstrated by the oxidation experiments described in Chapter 7 (vide infra).

4. Experimental section

General
Acetonitrile was HPLC grade (Biosolve, Valkenswaard). HPLC columns were from Alltech (Breda), analytical: 250 mm x 4.6 mm, preparative: 250 mm x 22 mm. HPLC chromatograms were recorded on a Shimadzu HPLC with LC20AT pumps and a SPD20A detector set at 214 nm and 280 nm, or the specific wavelength of the fluorophore. The column was thermostatted at 30 °C. ATPase measurements were performed on a Cary 100 or Cary 50 Conc (Varian, Middelburg) UV-Vis spectrometer. Fluorescence spectroscopy titrations were performed on a Perkin Elmer Luminescence Spectrometer LS50 (Monza, Italy), in a 3x3 mm fluorescence cuvette (Hellma, Rijswijk). Maldi-TOF ms was performed on a Bruker Biflex III spectrometer, with 2,5-dihydroxybenzoic acid (DHB) or α-cyano-4-hydroxy-3-methoxy-cinnamic acid (αCCE) as matrix. NMR spectra were taken on a Varian Inova 400 (400 MHz, 1H and 2D spectra) or on a Bruker DMX300 (75 MHz, 13C spectra) and calibrated to an internal standard of tetramethylsilane. Abbreviations used are s, singlet; d, doublet; t, triplet; dd, double doublet; m, multiplet.

Calculations related to Figure 5
The cooperative model for gp45 trimer assembly is given by:

\[
[A_1] \leftrightarrow k_1 [A_2] + [A] \quad , \quad [A_2] \leftrightarrow k_2 2[A]
\]

The subscripts denote where the clamp is in a trimer, dimer or monomer conformation. The equilibrium constants for these processes are:

\[
k_1 = \frac{[A][A_2]}{[A_1]} \\
k_2 = \frac{[A]^2}{[A_2]}
\]

Solving the equation of \(k_1\) for \([A_1]\) and \(k_2\) for \([A_2]\), and substituting the latter in the first yields:

\[
[A_1] = \frac{[A]^3}{k_1 k_2}
\]

which was used to construct the curves. The non-cooperative model is given by:

\[
\nu = \frac{[A]}{[A] + K_D}
\]

in which \(\nu\) is the binding density (see Appendix).
Synthesis of the peptide CSLDFLFG

The peptide was synthesized on 2.8 gr Wang resin (0.67 mmol/gr loading) using standard solid phase peptide techniques. The amino acid N-terminals were Fmoc protected, while the side chains were protected by a Boc (t-butyloxycarbonyl) group (lysine), t-butyloxycarbonyl or ester groups (in the case of serine, aspartic acid), or a trityl group (cysteine). The peptide was cleaved from the resin using 95% TFA, 2.5% ethanedithiol, and 2.5% H2O, and subsequently precipitated in ether and lyophilized from acetic acid.

**Maldi-TOF MS (DHB, m/z)** 923 (M+Na+), 1820 (2M+Na+);

**Anal** calcd for C_{44}H_{64}N_{8}O_{14}S (HOAc salt): C, 54.99; H, 6.71; N, 11.66; found: C, 50.93; H, 5.84; N, 10.77, consistent with 93% purity;

\(^{1}H\)-NMR (400MHz, DMSO) δ 8.58 (d, \(J=7.2\)Hz, 1H), 8.23 (t, \(J=5.7\)Hz, 1H), 8.14 (d, \(J=7.5\)Hz, 2H), 7.92 (d, \(J=8.0\)Hz, 1H), 7.77 (d, \(J=7.8\)Hz, 1H), 7.24 (s, 2H), 7.23 (s, 2H), 7.20 (s 1H), 7.18 (t, \(J=4.6\)Hz, 6H), 4.58-4.20 (m, 7H, \(H_{\alpha}\)), 3.98 (t, \(J=5.0\)Hz, 1H), 3.75 (br s, 2H), 3.62 (t, \(J=6.2\)Hz, 2H), 3.05 (dd, \(J=13.8, 4.1\)Hz, 2H), 2.97 (m, 2H), 2.87-2.73 (m, 4H), 2.62 (dd, \(J=16.5, 5.0\)Hz, 2H), 2.40 (dd, \(J=16.6, 8.3\)Hz, 2H), 1.60 (m, 1H), 1.50-1.36 (m, 6H), 1.11 (s, 2H);

\(^{13}C\)-NMR (75MHz, DMSO) δ 171.95, 171.75, 171.65, 171.54, 171.16, 171.01, 170.34, 169.56, 167.29, 137.64, 137.53, 129.16, 127.97, 126.15, 61.74, 61.45, 55.25, 53.85, 53.45, 51.25, 51.01, 49.63, 40.70, 37.54, 37.11, 36.03, 35.92, 27.15, 25.71, 24.03, 23.97, 23.19, 22.96, 21.63, 21.45, 21.32.

**Dansyl labelling of the peptide**

The peptide was dissolved in degassed labelling buffer (50mM phosphate buffer pH 7.4, 1 mM EDTA) to a concentration of 1.5-1.7 mg/ml, in eppendorf tubes that were flushed with argon. Then 4 mg of IAEDANS was added, and the solution was incubated for 3 hrs. The reaction mixture was analysed on an analytical HPLC column, with a gradient of MilliQ supplemented with 5% acetonitrile and 0.1% TFA (solution A) and acetonitrile supplemented with 0.1% TFA (solution B). The gradient was applied as follows: the concentration of B was 15% for 5 mins, and subsequently the peptide was eluted with a gradient to 100% B in 25 mins (flow speed 1 ml/min). The labelled peptide was purified on a preparative HPLC column. First a gradient was applied of 25-50% B in 15 mins, followed by a gradient to 100% B in 5 mins. Isolated fractions were analyzed with Maldi-TOF ms. Pure fractions were pooled, concentrated and dissolved in labelling buffer. Concentrations were determined using the UV-Vis absorption of the label (\(\varepsilon = 5,700 \text{ cm}^{-1}\text{M}^{-1}\) at 336nm).

**Maldi-TOF MS (\(\alpha\)CCA, m/z)** 1228 (M+Na+), 1251 (M-H++2Na+), 1273 (M-2H++3Na+), 1295 (M-3H++4Na+).

**Fluorescein labelling of the peptide**

Labelling and purification were performed as described for dansyl iodoacetamide, using 4 mg 6-iodoacetamidofluorescein (6-IAF) as labelling reagent. Concentrations of the purified samples were determined by UV-Vis using the absorption of the label (\(\varepsilon = 82,000 \text{ cm}^{-1}\text{M}^{-1}\) at 491nm).

**Maldi-TOF MS (\(\alpha\)CCA, m/z)** 1288 (M+Na+), 1310 (M-H++2Na+), 1273 (M-2H++3Na+), 1295 (M-3H++4Na+).

**Fluorescence titrations with dansyl labelled peptides**

Experiments were performed in 25 mM Tris.OAc pH 7.5, 175 mM KOAc, 10 mM Mg(OAc)_2, supplemented with 0.002 wt% Tween 80. The cuvette was first cleaned overnight with piranha acid (concentrated sulphuric acid with a few drops 30% hydrogen peroxide solution), thoroughly washed with MilliQ and ethanol, and dried with air. Then, the cuvette was filled with a 5 vol% solution of dichlorodimethylsilane in distilled chloroform and the solvent was allowed to evaporate. The cuvette was thoroughly washed with chloroform, ethanol, and MilliQ. This treatment led to an increase in the contact angle of aqueous droplets placed in the cuvette, and did not alter the UV-Vis characteristics of the quartz (data not shown).

A solution of dansyl labelled peptide containing the appropriate concentration of the clamp was then titrated into a solution of clamp protein at the concentrations indicated in the tables. If indicated, the solutions also contained the appropriate amount of PEG Mn=8,000 or pBR322 plasmid. Excitation of
tryptophan was performed at 280 nm for experiments without DNA, and at 290 nm for experiments with DNA. The resulting traces were fitted according to a 1:1 binding model using the least squares method in Microsoft Excel:

\[
F_{\text{calc}} = 1 - \frac{1 + K_d [E]_{\text{total}} + K_a [I]_{\text{total}} \pm \sqrt{\left(4 K_d [E]_{\text{total}} [I]_{\text{total}} + (K_d [E]_{\text{total}} + K_a [I]_{\text{total}} + 1)^2\right)}}{2 K_d [E]_{\text{total}}} \times \Delta F_{\text{max}}
\]

in which \(E_{\text{total}}\) is the total protein concentration and \(P_{\text{total}}\) is the total peptide concentration. \(K_d\) is the association constant and is converted to the dissociation constant by taking the reciprocal value. \(F_{\text{obs}}\) is the observed fluorescence, and \(\Delta F_{\text{max}}\) the total final fluorescence value. \(K_a\) and \(\Delta F_{\text{max}}\) were fitted in this model.

**Scatchard plot analysis**

Scatchard plot analysis was performed according to a literature procedure\(^4\), describing the determination of the number of binding sites of a fluorophore on BSA. To determine the change in tryptophan fluorescence due to quenching by the labelled peptide, small quantities of peptide were titrated into a solution of gp45 as described above. Because \([P]<<<[E]\), all peptides are assumed to be bound to the protein. The resulting traces were fitted in Microsoft Excel to a linear formula \((y = ax + b)\). Only fits with \(R^2 > 0.7\) were used to calculate the average change in fluorescence due to peptide quenching. The average was calculated by at least three traces for each experiment. The plots were displayed using the Scatchard equation (see appendix A2):

\[
\frac{\nu}{[P]_{\text{free}}} = nK_a - K_a \nu
\]

where \(\nu\) is the concentration of bound peptide divided by the total concentration of the protein, and \(n\) is the number of binding sites, assuming identical and equivalent binding sites.

**ATPase stimulation assays**

ATPase assays were performed in a 2x10 mm cuvette at 25 °C. The procedure was slightly different from the one described in Chapter 3. First, 60 μl of the ATP hydrolysis mix (1.2 mM ATP, 4 mM phosphoenopyruvate, 0.35 mM NADH, 0.6 μM streptavidin, and 3.6 μl pyruvate kinase and lactate dehydrogenase enzyme mix [resulting in ~1.8 units pyruvate kinase and ~1.3 units lactate dehydrogenase per ml cocktail] in 30 mM Tris.OAc pH 7.5, 175 mM KOAc, 12 mM Mg(OAc)\(_2\)), was added to the cuvette and incubated for two mins. Then, sequentially all the components were added to the cuvette as described in Chapter 3. Final concentrations were 250 nM gp44/62, gp45 and DNA template, and various amounts of the peptides. Dilution of the ATPase mixture up to the final concentrations obtained in the experiments with the peptide did not significantly change the ATP hydrolysis rate.

5. **References**

Interaction of the T4 DNA polymerase C-terminus with the T4 clamp


Synthesis and characterization of maleimide functionalized porphyrin derivatives

This chapter describes the synthesis and characterization of maleimide functionalized porphyrin derivatives that are to be conjugated to the T4 clamp protein. A synthetic procedure is developed, in which the porphyrin is coupled to a linker with a maleimide group. To this end, different synthetic routes are explored, of which the Cu(I) catalyzed [3+2] cycloaddition of azides and alkynes proved to be an efficient method to link porphyrins to the spacers. The scope of this method is explored and a small library of bioconjugatable porphyrins is synthesized.
1. Introduction

This chapter describes the synthesis of a series of maleimide functionalized porphyrin compounds, which can be used to target thiol functionalities of biomolecules such as proteins, cysteine containing peptides, and thiol appended nucleic acids. The maleimide group was chosen over haloacetic acid groups because of its high specificity for thiol functionalities via the Michael-type addition, although haloacetic acid groups react faster with thiols via a nucleophilic substitution reaction. Therefore, normally smaller excesses of the latter compound with respect to the targeted thiol are needed to obtain quantitative functionalization, in comparison to the maleimide route.

The main goal of the research described in this thesis is to perform processive oxidation reactions on a DNA template with a porphyrin functionalized T4 clamp. This sets a number of requirements with regard to the design of the maleimido porphyrins. i) Because the size of the porphyrin is much smaller than that of the T4 clamp, a linker between the maleimide functionality and the porphyrin is needed for the porphyrin to be able to reach the DNA. Because the porphyrin will preferably be attached to the rim of the protein, a linker length of around 30 Å was calculated to be sufficient to allow the porphyrin to complex with the DNA minor groove, which is required for successful oxidation reactions to take place. Furthermore, the oxidation reactions will be performed in aqueous buffered solvents, so the linker has to be water soluble. An ethylene glycol segment was therefore chosen as linker, because it meets the combined requirements of water solubility and high flexibility. Ethylene glycols generally do not suffer from non-specific absorption on protein surfaces, because 2-3 water molecules bind to every ethylene glycol segment. ii) The use of aqueous buffers implies that the porphyrin should preferably be water soluble as well. Besides the corresponding manganese derivatives, other metallation states of the porphyrin (free base, Zn) would be useful for synthetic optimization, analysis and fluorescence studies.

Figure 1 - Retrosynthetic route towards biomolecule – porphyrin conjugates. X and Y are complementary functional groups, R are the porphyrin side groups, which can either have hydrophobic or hydrophilic properties, M is representing the metallation state of the porphyrin (2H, Zn, MnX).
Moreover, the peripheral functionalities of the porphyrin may have a large effect on the efficiency of the oxidation reaction itself. Therefore, a synthesis route that allows one to obtain maleimidoporphyrins in a facile way would be advantageous. A route in which the porphyrin is added to the maleimide functionality in the last step satisfies this requirement. Figure 1 shows a retrosynthetic route that complies with all the abovementioned criteria. It involves the synthesis of a maleimide linker molecule with a synthetic handle (X) that is complementary to that of the porphyrin (Y). The next sections describe the efforts of synthesizing maleimidoporphyrins via the use of different synthetic strategies (X and Y groups).

2. Results and discussion

2.1 First generation maleimidotritolyl porphyrins

Nucleophilic substitution reactions of porphyrin phenolates on alkylhalides are often used for porphyrin functionalization, because of the synthetic availability of the two types of reactants and the fact that these reactions usually proceed in high yields. This approach was therefore tested in the linking of maleimidimide halide spacers to porphyrins. First, a maleimide linker containing an alkyl bromide functionality was synthesized from ethylene glycol diamine 1 by selective protection of one of the two amines with a t-butyloxy carbonyl (Boc) group, followed by reaction with 6-bromohexanoic acid chloride, yielding the Boc-protected bromide 3. The Boc group of 3 was removed by treatment with HCl in ethyl acetate and the amine was subsequently reacted with the N-hydroxysuccinimide (NHS) activated ester of N-maleimido caproic acid (5). The use of the NHS ester of maleimide caproic acid proved to be the most convenient and high yielding route to perform the amide coupling when compared to other methods such as carbodiimide peptide couplings or using the maleimide acid chloride. The activated ester is commercially available, however, because of its high cost, the preferred method was to react the acid with disuccinimylcarbonate (DSC), according to the method of Trester-Zedlitz. The use of DSC to obtain an activated

![Scheme 1](image)

**Scheme 1** - Reagents and conditions (i) 0.2 eq t-butoxydicarbonate, CH₂Cl₂, 99% versus t-butoxydicarbonate. (ii) 1.2 eq 6-bromohexanoylchloride, 1.2 eq DIPEA, CH₂Cl₂, 0 °C to room temperature, 72%. (iii) (a) 2 M HCl/EtOAc (b) 5, CH₂Cl₂, 18%. (iv) 1.0 eq disuccinimylcarbonate, 1.1 eq DIPEA, CH₂Cl₂, 2 hrs.
maleimide analogue was the most successful procedure, compared to other investigated standard amide coupling procedures. Typically, the commercially available 5-maleimide caproic acid 4 was reacted with an equimolar amount of DSC under basic conditions for two hours, after which the crude reaction mixture was slowly added to the amine. No urea-based by-products originating from the reaction of DSC with the amine were found after purification. Alternatively, the activated ester can be purified before reaction with the amine with the help of silica gel chromatography. Unfortunately, this did not lead to significantly improved yields, nor did the application of commercial NHS-activated maleimide caproate. Unfortunately, the excess of NHS in the reaction mixture when using the crude activated maleimide caproic acid leads to a substantial impurity of NHS in the product, even after silica gel chromatography. This impurity can be adequately removed by Biobeads size exclusion chromatography. The reaction of the bromo-functionalized amine spacer 3 and the maleimide NHS ester 5, however, gave low yields (18%), and the proton resonances of the product in the 1H-NMR spectrum displayed broadening, indicative of aggregation. Indeed, after a small amount of DMSO was added to the NMR sample to dissolve these aggregates, the resonances were sharper. This aggregation behaviour could have led to product loss during column chromatography, explaining the poor yields obtained.

Tritolylphenoxyporphyrin (TTP) was investigated as a model porphyrin, as it is easily prepared and purified, and can therefore be used to optimize the proposed nucleophilic substitution reactions. The tripyridylporphyrin (TPyP) derivatives are of more interest, because they are more polar and potentially water soluble, but also synthetically more demanding. TTP 7 was synthesized according to the Adler method, and purified by precipitation and subsequent column chromatography. It was then reacted with the bromide maleimide derivative 6 in DMF with an excess of potassium carbonate as a base (Scheme 2). Even after prolonged reaction times and the addition of a larger excess of potassium carbonate, this did not lead to formation of any product as judged by TLC analysis.

To investigate why the reaction of TTP 7 with spacer 6 did not yield any product, 7 was reacted with bromide 3. The only difference between 3 and 6 is that the former has a Boc-protected amine, while the latter a maleimide functionality. Under similar conditions as for the reaction with the maleimide, the reaction of 3 with 7 yielded 75% of the Boc-protected porphyrin 9. After deprotection of the Boc-group and reaction of the amine with the NHS activated ester of maleimide caproic acid, maleimidoporphyrin 8 was obtained in a disappointingly low yield of 15%. In the last size exclusion purification step a side product was collected that displayed broad NMR signals, which corresponded to 31% of the total yield. The non-aggregated product was determined to be stable for longer than a week in a deuterated chloroform solution when stored in the dark, since no changes in its NMR spectrum were detected.
Synthesis and characterization of maleimide functionalized porphyrin derivatives

Scheme 2 - Reagents and conditions (i) propionic acid, 135 °C, 1.5 hrs, 5% (ii) 1.5 eq 6, 4 eq K₂CO₃, DMF. (iii) 1.2 eq 3, 4 eq K₂CO₃, DMF, 60 hrs, 75%. (iv) (a) 2M HCl/EtOAc (b) 1.1 eq 5, CH₂Cl₂, 15%.

2.2 Second generation maleimidotritoluylporphyrins

The low yields and aggregation problems encountered during the synthesis of maleimidotritoluylporphyrin 8 following nucleophilic substitution reaction paths prompted us to reassess this synthesis route. Therefore, the synthesis of maleimidoporphyrin derivatives via the Cu(I) catalyzed [2+3] dipolar cycloaddition of an alkyne and an azide was investigated. This method, also called the Huisgen’s cycloaddition or the ‘click’ reaction, has emerged during recent years as a powerful method for linking azides to alkynes via a covalent and stable triazole linker under ambient conditions. Originally investigated by Huisgen in the 1960s, the spontaneous reaction between these compounds only proceeds at high temperatures without regioselectivity, i.e. both the 1,5 and 1,3 triazole derivatives are formed in equal amounts. The finding of the Sharpless and Meldal groups that this reaction can be catalyzed by Cu(I) yielding only the 1,5-isomer in near or quantitative yields, and can be performed in the presence of a wide range of functional groups and solvents, sparked its resurgence. The click reaction has shown to be very useful for the synthesis of small organic molecules, of which some are of pharmaceutical interest. Furthermore, because of its bio-orthogonal character, the click reaction has been employed for demanding bioconjugations such as the synthesis of peptide or protein-polymer conjugates.
heterodiprotein conjugates, functionalization of solid phase supports with DNA, and decoration of viruses or viral capsids with fluorophores.

A clickable maleimide linker was synthesized by functionalising the mono-protected amine spacer with 5-pentynoic acid with the help of EDC peptide coupling methods, followed by the deprotection of the Boc protecting group via treatment with hydrochloric acid in ethyl acetate. The reaction with the NHS ester of maleimide carproic acid under the aforementioned reaction conditions (see Scheme 1) subsequently yielded the alkyne maleimide spacer as a white crystalline solid material in a 47% total yield (Scheme 3).

Scheme 3 – Reagents and conditions (i) 1.1 eq 5-pentynoic acid, 1.1 eq EDC, 1.1 eq DIPEA, 1.1 eq HOBt, DMF, 16 hrs, 80% (ii) (a) 2M HCl/EtOAc (b) 1.5 eq 5, CH2Cl2, 16 hrs, 62%.

To obtain a porphyrin that can be used in the click reaction, TTP was reacted with 1,3 dibromopropane, followed by substitution of the bromide with sodium azide in DMF. This yielded the azide functionalized TTP in an overall 44% isolated yield. Alternatively, 3-azido-1-mesylpropane was synthesized from 3-chloropropanol by substitution of the chlorine with sodium azido followed by mesylation of the hydroxyl group. The reaction of with TTP afforded the azide functionalized TTP in one step, in higher yields compared to the previous method, and preventing the laborious separation of the bromide and azide derivatives in the previous route. The azide functionalized TTP could be metallated with zinc acetate tetrahydrate or manganese chloride in DMF at 135 °C in high yields without impairing the azide functionality.

The azide porphyrins, Zn, and Mn were reacted with alkyne functionalized maleimide spacer in click reactions in THF under Schlenk conditions, using a porphyrin concentration of 5 mg/ml and CuBr as catalyst. Reactions were analyzed after 16 hours, usually showing quantitative or near-quantitative consumption of the porphyrin azide. For Zn and Zn, the reaction proceeded smoothly without substantial formation of by-products, as judged from TLC analysis. No copper was inserted in the porphyrin, according to UV-Vis spectroscopy. The reaction of Mn, however, showed a number of by-products, which were found to be yet unidentified higher molecular weight compounds (analyzed by MALDI-TOF ms). The reaction mixtures were subjected to silica column chromatography and the products were precipitated to afford pure maleimide functionalized porphyrins.
Synthesis and characterization of maleimide functionalized porphyrin derivatives

Scheme 4 – Reagents and conditions (i) 20 eq 1,3 dibromopropane, 4 eq K$_2$CO$_3$, DMF, 16 hrs, 63% (ii) 30 eq NaN$_3$, DMF, 70 °C, 16 hrs, 70% (iii) M=Zn, 10 eq Zn(OAc)$_2$.2H$_2$O, DMF, 150 °C, 3 hrs, 77%. M=MnBr, (a) 10 eq MnCl$_2$ DMF, 150 °C, 3 hrs, (b) saturated NaBr(aq) in CHCl$_3$, 85%. (iv) 3 eq NaN$_3$, DMF, 60 °C, 24 hrs, 74%. (v) 1.2eq mesylchloride, 1.2 eq DIPEA, THF, 0 °C to room temperature, 85%. (vi) 1.5 eq 15, 4 eq K$_2$CO$_3$, DMF, 61%.

Scheme 5 – Reagents and conditions (i) 0.2 eq CuBr, 50 eq PMDETA, THF, Schlenk conditions.
and Mn16. NMR experiments on Zn16 revealed a concentration dependence of the resonance signal of the triazole proton, accompanied by a splitting in the porphyrin region, suggesting a complexation of the triazole unit to the zinc metal of the porphyrin. Indeed, Zn16 was diluted to give a tentative binding curve, which however could not be fitted to a 1:1 binding model, providing evidence for an intermolecular rather than an intramolecular binding process.

2.3 Second generation maleimido tripyridylporphyrins

Although the maleimide tritoluylporphyrins 16 were synthesized successfully and can be used for conjugation reactions, the syntheses of the corresponding tripyridyl derivatives were also investigated because these compounds are more polar and more water soluble counterparts of 16. Tripyridylporphyrin (TPyP) 17 was synthesized according to the classical Adler-Longo method.\textsuperscript{21} The phenolic oxygen of \( p \)-hydroxy benzaldehyde was protected by esterification with propionic acid before adding the other reactants to facilitate separation of the resulting mixture of porphyrins by silica column chromatography.\textsuperscript{22} After purification, the ester was removed by treatment with Tesser’s base, and the product was precipitated by acidification, washed and dried thoroughly under high vacuum, to give 18 in 95\% yield. The free phenolic hydroxyl functionality of 18 was reacted with mesyl propyazide 15 to yield the azide appended TPyP 19 in 77\% yield. Metallation to obtain the manganese(III) azide TPyP Mn19 could be carried out in a 76\% yield. In this case collidine was used as base and purification was performed by extraction and subsequent alumina column chromatography.

\begin{align*}
\text{Scheme 6} - \text{Reagents and conditions (i) acetic anhydride, propionic acid, 130 °C, 1.5 hrs, 6\%. (ii) p-dioxane, MeOH, NaOH(aq), 30 min, 95\%. (iii) 20 eq NaOH, 1.5 eq 15, DMF, 16 hrs, 77\%. (iv) 10 eq Mn(OAc)\textsubscript{2}.4H\textsubscript{2}O, 10 eq collidine, DMF 150 °C, 76%}
\end{align*}
Collidine aids the metallation by a twofold mechanism; i) it deprotonates the pyrrolic protons of the porphyrin and, ii) it prevents the formation of large TPyP aggregates by interfering with pyridyl side group - manganese interactions.\textsuperscript{22}

The click reactions of azide TPyP \textbf{19} and \textbf{11} were performed using the same reaction conditions as for the TTP derivatives. In contrast to the click reactions with the tritoluylporphyrins, UV-Vis spectroscopy studies of the reaction mixtures with free base TPyPs pointed to copper insertion in the porphyrin. It is possible that the pyridyl side groups oxidize Cu(I) to Cu(II) to promote insertion of Cu(II) in the porphyrin, however it is more likely that the pyridyl groups deprotonate the porphyrin’s pyrrolic protons, making the porphyrin more susceptible for metallation. In this respect, it is noteworthy that the click reactions with TTP are carried out with an excess of PMDETA, which also acts as a base but apparently is not able to promote copper insertion in the TTP porphyrins. As a result of the copper sequestering by the porphyrin, reactions were not quantitative when 0.2 eq CuBr was used as catalyst. Higher amounts of CuBr (1.0 eq, 5.0 eq) resulted in quantitative consumption of \textbf{19} but also relatively more copper insertion. The copper and free base porphyrins could not be separated by silica chromatography.

\textbf{Scheme 7 –} Reagents and conditions. (i) 1.0 eq CuBr, 50 eq PMDETA, THF, Schlenk conditions.

The click reaction of \textbf{Mn19} and \textbf{11} again showed quantitative formation of the clicked product, because the manganese in the porphyrin excludes copper insertion, as determined by Maldi-TOF ms analysis of the reaction mixture (M\textsuperscript{+} -OAc peak 1190 m/z). Unfortunately, silica and alumina column chromatography purification always led to quick and quantitative conversion to a peak with m/z 1222. This appeared to be a methanol (+32) adduct of the product, a hypothesis that was confirmed by using ethanol instead of methanol in the chromatography steps, in which case a mass peak of 1235 (+45) was found for the product. Moreover, minor signals with m/z value of 1208 were found and attributed to a water adduct (+18). The adducts were determined to be unexchangeable, as upon dissolving the methanol adduct in ethanol and \textit{vice versa} the mass spectrum was found to be essentially the
same. It is unlikely, therefore, that the adducts originated from co-ordination of the alcohols to the manganese center. Thus, it has to be concluded that the alcohols must be covalently linked. \(^1\)H-NMR experiments were performed in deuterated chloroform supplemented with deuterated methanol in order to suppress the amide proton resonance. Although the spectrum displayed significant broadened peaks because of the paramagnetic manganese(III) center, it did not show any resonance that could be attributed to the maleimide double bond. The absence of maleimide functionality was supported by a TLC alkylation assay,\(^{23}\) in which different maleimide porphyrins were spotted on a thiolated TLC plate and the TLC was developed in an appropriate solvent mixture (Figure 2B). When maleimide groups are present, the spotted compound does not migrate from the baseline because it becomes covalently attached to the plate. The efficiency of this process on a TLC plate was shown to be around 85% for compounds with maleimide or haloacetic acid moieties.\(^{23}\) A purified sample of maleimide TPyP (Figure 2B, pur) was used in this assay and failed to produce a baseline spot, in contrast to a sample of the reaction mixture (Figure 2B, rx), of which around 75% did not migrate from the baseline. Although such a case was unprecedented for the TTP derivatives, it was concluded that the alcohol had reacted with the maleimide double bond via a Michael type addition reaction. This reaction can indeed occur under strong basic conditions, accompanied with hydrolysis of the maleimide bond.\(^{24}\) Nevertheless, the latter hydrolysis reaction was ruled out because the imide bond of both maleimide porphyrin

---

**Figure 2** – A. Tentative mechanism for maleimide hydrolysis, porphyrin is depicted from the side as a black bar. \(R^1\) is the linker and porphyrin part of either the same or a different maleimide porphyrin; \(R^2\) can be H, CH\(_3\), or CH\(_2\)CH\(_3\) for water, methanol or ethanol, respectively. B. TLC alkylation assay. Left: samples on thiolated TLC plate, right: control on untreated TLC plate; rf values are lower due to difference in polarity of the treated and untreated TLC plate. Rx: reaction mixture of Mn19, pur: Mn19 after workup with alcohol. C. Maldi-TOF ms spectrum of Mn19 reaction mixture (rx). D. Maldi-TOF ms spectrum of purified Mn19 (pur).
samples (rx and pur) was still present (determined by IR spectroscopy, 1600-1610 cm\(^{-1}\)). The maleimide group, however, is usually very selective towards Michael addition reactions with thiols, and thus has to be activated for reaction with poorer nucleophiles such as alcohols. The co-ordination of one of the maleimide carbonyls to the electron deficient manganese centre in the porphyrin could bring about such an activation, \textit{i.e.} by making the maleimide double bond more electron poor (Figure 2A).

To avoid the above side reactions, the purification of \textbf{Mn19} was performed by extracting the reaction mixture with mild acid to remove copper salts and base, followed by precipitation of the product in hexane. Besides avoiding any column chromatography steps, extraction also improved the yield because any aggregates that arose from co-ordination of the pyridyl side groups with manganese or possibly copper ions were dissolved by the acid treatment. The fact that the product had the desired composition was confirmed by Maldi-TOF ms, which showed only a peak with m/z value of 1222 and no addition products.

To make a truly water soluble catalyst, the pyridyl side groups of \textbf{Mn20} were reacted with methyl iodide to yield \textbf{Mn21}. Because the product could not be analyzed by NMR spectroscopy, a control experiment with free base maleimide TTP porphyrin 16 was performed using identical reaction conditions. This experiment showed that the maleimide double bond protons did not react, although \textsuperscript{13}C-NMR showed that some methylation occurred at other sites, presumably the amide nitrogens, when extended reaction times were used. This might explain why the yield was found to be 120\%, even though the product was precipitated several times.

\textbf{Scheme 8} – Reagents and conditions. (i) 150 eq MeI, DMF, 40 °C, 1 hr, 120\%.

![Scheme 8](image-url)
2.4 Third generation maleimido tripyridylporphyrins

The results in the previous section have shown that the click reaction is quite successful in coupling azido porphyrins with alkyne linkers. To further investigate the scope of the click reaction, the reactions of reactants with reversed functionalities were studied as well. To this end, alkyne porphyrins were synthesized and tested in reactions with azide linkers. The main advantage of the use of alkyne functionalized porphyrins is that, in contrast to their azide counterparts, they can be prepared in one step from the porphyrin reaction mixture. Alkyne porphyrins have been synthesized by Lindsay et al. and used to for the purpose of linking porphyrins or other chromophoric compounds together via Sonogashira or Glaser reactions. The dimer and oligomer porphyrins prepared in these studies had hydrophobic mesityl substituents and their properties were studied in organic solvents. The resulting dyads were found to display interesting energy transfer properties and were studied as synthetic light-harvesting systems or investigated as facilitators in the field of data storage.

Scheme 9 shows the synthesis of an ethynyl porphyrin with polar pyridyl side groups. First, 4-bromo benzaldehyde was functionalized with an ethynyl trimethylsilyl group via a CuI co-catalyzed Sonogashira coupling reaction, and the product purified by column chromatography and subsequent crystallization to give 22 in 61% yield. It was subsequently used to synthesize porphyrin 23, which was purified by tedious silica and alumina column chromatography up to 90-95% purity based on 1H-NMR spectroscopy. More chromatography steps did not improve the purity of 23, and although the MALDI-TOF ms spectrum only showed the molecular ion of 23, 1H-NMR spectroscopy revealed minor additional peaks in the porphyrin region and additional sets of trimethylsilane resonances. In contrast to the more harsh conditions described in the literature, involving either treatment with strong base or TBAF, the trimethylsilane protecting group could be removed in 88% yield under mild conditions with K2CO3. After deprotection, ethynyl porphyrin 23 was metallated with manganese(III) to Mn24 in reasonable yield (65%).
The synthesis of maleimide azide linkers proved to be a challenge (Scheme 10). First, reaction of maleimide bromide 6 was attempted with sodium azide in order to obtain 25, an azide analogue of maleimide alkyne linker 11. Immediately after adding sodium azide to 6, a red brown colour appeared, accompanied with the appearance of a number of side products on TLC, showing that this reaction was not successful. In an alternative route, Boc-protected bromo spacer 3 was converted to Boc-protected azide spacer 26 in 81% yield. The amine was deprotected and reacted with in situ prepared 5, to yield maleimide spacer 25 in a disappointingly poor yield (<10%) after column chromatography. Finally, amine azide spacer 29 was synthesized using a literature procedure via the selective reduction of one of the azides of diazide 28, which was prepared from tetraethylene glycol 27. However, reaction of 27 with in situ prepared maleimide activated ester 5 failed to produce maleimide azide spacer 30 in high yields. The low yields may be explained by the formation of triazolines as a side reaction (see insert Scheme 10). Besides being a Michael acceptor, a maleimide group is also an excellent dienophile and can react with dienes or dipolar dienes such as azides. Normally, triazoline synthesis is performed with large excesses of methyl azide, high temperatures and relatively long reaction times to obtain quantitative conversions. It has also been shown, however, that this cycloaddition occurs when azides and maleimides are brought together in close proximity via hydrogen bonding. Although the synthesis of spacers 25 and 29 was not performed at high temperatures and, in the case of 25, sodium azide instead of methyl azide was used, the stability of the resulting maleimide

Scheme 10 – Reagents and conditions. (i) 5.0 eq NaN₃, DMF, 50 °C. (ii) 5.0 eq NaN₃, TBAI (cat), DMF, 50 °C, 16 hrs, 72%. (iii) (a) 2M HCl/EtOAc (b) 1.1 eq disuccinimylcarbonate, 1.1 eq 4, 1.2 eq DIPEA, CH₂Cl₂, 2 hrs in separate flask, then added to (a). (iv) 2.2 eq methanesulfonyl chloride, 2.2 eq triethylamine, 2.1 eq NaN₃, THF, 0 °C to room temperature, 4.5 hrs, 65%. (v) 0.9 eq triphenylphosphine in diethyl ether, 0.65 M H₃PO₄, 24 hrs, 74%. (vi) 0.9 eq 5, CH₂Cl₂, no product isolated.
azide spacers could have been impaired by triazoline formation via an intra- or intermolecular cycloaddition.

To prove that the ethynyl porphyrins can be useful synthons in the click reaction, the reaction was performed with Boc-protected azide spacer 26 instead of an azide maleimide linker. Indeed, the reaction yielded the resulting Boc-protected porphyrins 30, but suffered from copper insertion like the azide functionalized TPyPs. However, when 1.0 or 5.0 equivalent of catalyst were used, the reaction was quantitative, with the product being partially or fully metallated, respectively. Deprotection of the Boc group and reaction of the amine with the activated ester of the maleimide may still lead to the synthesis of maleimide porphyrins, however, this line was not pursued because the route starting from azide functionalized TPyPs yields the desired maleimide TPyPs in a less laborious procedure.

Scheme 11 – Reagents and conditions. (i) 1.0 or 5.0 eq CuBr, 50 eq PMDETA, THF, Schlenk conditions.

2.5 Maleimide functionalized Protoporphyrin IX derivatives

Protoporphyrin IX is found as prosthetic group in heme proteins, a class of proteins that perform critical tasks in processes like aerobic respiration (haemoglobin) and detoxification (cytochromes, horseradish peroxidase). In these proteins, protoporphyrin IX usually has an Fe(II) center. The porphyrin has two free carboxylic groups, which can be functionalized without impairing the reconstitution of the modified protoporphyrin in the protein. Thus, the porphyrin can be extracted from the protein, modified either at the one of the carboxylic acid moieties or by insertion of a metal, and reconstituted in the protein. This procedure has been exploited by our lab to reconstitute giant amphiphiles made from myoglobin and protoporphyrin modified synthetic polymers. In this work, we aimed at modifying the protoporphyrin with a maleimide group using the procedures developed in the previous sections. Via reconstitution of the maleimide functionalized protoporphyrin with proteins
like myoglobin, cytochrome C, or horseradish peroxidase, enzymes are obtained that are monofunctionalized with a maleimide group. This maleimide group can then be used to attach the enzyme to a surface, to other biomolecules, or beads, etc. The maleimide appended protoporphyrin can in principle also be coupled to the T4 clamp to yield bio-hybrid oxidation catalysts. These catalysts may display toxic properties when eventually used as a medicine in vivo, because protoporphyrins are naturally occurring compounds in the body. Lastly, the protoporphyrin, or derivatives thereof, might find useful application in photodynamic therapy. Via conjugation to the clamp protein it may be targeted towards DNA, while other photodynamic therapy agents usually bind to membranes.39

In our group, a route towards monofunctionalized azide porphyrin 2H31 and the metallated derivatives Zn31 and Mn31 thereof, has been developed.40 Because these protoporphyrin IX derivatives already have an ethylene glycol spacer, it was decided to use a short maleimide alkyne synthon to investigate the click reaction with 2H31. The short maleimide alkyne 32 was synthesized from maleimide caproic acid 4 in a yield of 43%, applying disuccinimyl carbonate (DSC) and propargylamine as nucleophile. Maleimide 32 was subsequently reacted with 2H31 under Schlenk conditions, which resulted in the formation of maleimide functionalized protoporphyrin 2H33. Unfortunately, Cu was found to be inserted into the porphyrin, like in the case of the click reactions with the free base pyridyl porphyrins. In the latter cases, the insertion was attributed to deprotonation by the side groups of the porphyrin, which clearly cannot be the case in porphyrin 31. It could be that the pyrollic protons of protoporphyrins have a higher pKₐ and are more readily

Scheme 12 – Reagents and conditions. (i) a. 1.2 eq DSC, 1.2 eq DIPEA, CH₂Cl₂, 2 hrs, b. 1.4 eq propargylamine hydrochloride, 1.4 eq DIPEA, 15 hrs, 43%. (ii) 2 eq CuBr, 50 eq PMDETA, THF, Schlenk conditions, 33 °C.
deprotonated by PMDETA than the tripyridylporphyrins, which would allow Cu to be inserted into the porphyrin. Cu insertion has also been reported to occur in polymer synthesis procedure employing a free base porphyrin initiation. In this case, ATRP conditions employing elevated temperatures were used. The click reaction described here was also investigated for metallated protoporphyrins Mn and Zn. The reaction with Mn was found to proceed in low yields and gave several by-products (data not shown). Zn was converted to the product Zn in a moderate yield (54%), and could be purified by a tedious silica chromatography procedure.

3. Conclusions

A small library of porphyrin compounds having a maleimide functionality was synthesized by linking maleimide spacers with porphyrins, amongst others, by the ‘click’ reaction. The resulting compounds could not be synthesized via other synthetic strategies, such as nucleophilic substitution of alkyl halides by hydroxyl functionalized porphyrins. The ‘click’ method allowed for modular synthesis of the library, hence the metallation state of the porphyrin and the linker can easily be varied. The syntheses described in this chapter are the first studies toward maleimide functionalized porphyrin compounds and Cu(II) catalyzed ‘click’ chemistry applied to porphyrins. The reactions of linkers containing an alkyne functionality with azide functionalized tritoluyl porphyrins (free base, Zn and MnBr) went smoothly, except for the reaction with the manganese derivative, which was accompanied by the formation of unexplained side products. The alkyne linker was also reacted with azide functionalized tripyridyl porphyrins (free base and MnOAc). Although the ‘click’ reaction again proceeded with complete consumption of the azides, the reaction with the free base porphyrin suffered from copper insertion, which was not observed in the case of the tripyridyl derivatives. It turned out that the maleimide of the manganese derivative was not stable to nucleophiles, which was demonstrated with the help of a TLC alkylation assay. The maleimide tripyridylporphyrin could by purified by extraction. The pyridine functionalities of the manganese maleimide tripyridine porphyrin were subsequently methylated to obtain a water soluble maleimide porphyrin derivative.

In order to investigate the opposite polarity sequence of the reaction, an ethynyl tripyridyl porphyrin was synthesized and clicked to an azide linker. Again, the click reaction proceeded in high yields, but copper was found to be inserted in the porphyrin. The reaction of protoporphyrin IX azides with a small maleimide alkyne was also studied. In the case of the free base porphyrin, the reaction again proceeded smoothly, but was accompanied with copper insertion, while in the case of the manganese porphyrin a large amount of by-products were formed. By applying a zinc azide protoporphyrin IX instead of a free base derivative the corresponding maleimide could be synthesized in a moderate yield (i.e. 54%).
In short, the ‘click’ reaction can be applied successfully to porphyrin azides and alkynes. Sometimes, the reaction is hampered by the formation of by-products, of which the reasons of formation are unknown at this moment. Some free base porphyrins are prone to copper insertion during the reaction, presumably due to deprotonation of the pyrrolic hydrogens of the porphyrin core. Although not demonstrated here, undesired copper insertion may be avoided by first synthesizing the zinc porphyrin derivative and removing the metal after the reaction.

4. Experimental section

General.
NMR spectra were recorded on a Varian Inova 400 (400MHz, 1H and 2D spectra) or on a Bruker DMX300 (75MHz, 13C spectra) spectrometer at 298 K and calibrated using an internal standard of tetramethylsilane or the solvent. Chemical shifts are reported in parts per million (ppm), abbreviations used are: s = singlet, d = doublet, dd = double doublet, t = triplet, dt= double triplet, m = multiplet, br = broad, obs = obscured. Maldi-TOF spectra were recorded on a Bruker Biflex III instrument in reflector mode with dithranol as matrix and were uncorrected. EI and CI mass spectra were recorded on a Fisons VG7070E double-focussing mass spectrometer or on a Finnigan Mat 900s mass spectrometer. ESI spectra were recorded on a Finnigan LCQ Advantage MAX or JEOL Accutof CS JMS-T100CS (accurate mass). UV-Vis spectra were recorded on a Cary 50 Conc spectrometer and IR spectra were recorded on an Anadis Thermo Mattson IR300 spectrometer. Elemental analysis was carried out on a Carbo Erba Instruments CHNSO EA 1108 element analyzer. Melting points were determined on a Büchi B-545 melting point apparatus. Solvents were dried with suitable drying agents and distilled under a nitrogen atmosphere before use when indicated. DMF was Biotech grade (≥ 99.9%, Aldrich), and silica was from Acros (0.035-0.070 mm, pore diameter ca 6 nm) was used for column chromatography. For size exclusion chromatography, Bio-Beads SX1, 200-400 mesh (Bio-Rad) were used. The THF used as eluent for Bio-Beads chromatography was not dried prior to distillation.

TLC alkylation assay
A TLC alkylation assay was performed according to a modified literature procedure: a 20 x 20 cm TLC plate was incubated with a solution of 3-aminopropyltriethoxysilane (APTES) in ethanol (96%) by gentle shaking for 30 mins. The TLC plate was then dried in vacuum (~1 mbar), and the presence of amine functionalities was confirmed by spotting a ninhydrin TLC stain solution on a small section of the plate. 4,4′-dithiobutyric acid (357mg, 1.5 mmol), HOBt hydrate (405 mg, 3.0 mmol) and EDC (316 mg, 1.7 mmol) in CH2Cl2, were dissolved by vigorous stirring and reacted for 1 hr, after which the organic solvent was extracted twice with 0.5 M NaHSO4 and one time with brine solution. The organic layer was diluted with dichloromethane, supplemented with 0.5 ml triethylamine and incubated with the APTES treated TLC plate. After one hr, the plate was washed with dichloromethane and methanol, followed by reduction of the disulfide bridges by washing with a solution of β- mercaptoethanol in methanol for 30 mins. The plate was washed two times with copious amounts of methanol, dried under high vacuum, and ran 5 times with methanol with drying steps in between runs. The presence of thiol groups on the TLC plate was confirmed by spotting a solution of 5,5′-dithiobis(2-nitrobenzoic acid) (Ellman’s reagent) in 20 mM Tris.HCl buffer (pH 7.5) with 1 mM EDTA. The TLC plates were cut to appropriate size and used immediately.
General procedure for the porphyrin-maleimide click reaction.

THF (undistilled) was deoxygenated by bubbling anhydrous and oxygen-free nitrogen through the solvent for a minimal time of 30 mins. The azide and the alkyne derivatives (1.1 eq excess of molecule clicked to the porphyrin) were put in a Schlenk reaction vessel, which was deaerated by at least 5 consecutive pump-N₂ cycles. The porphyrin was dissolved in deaerated THF to a concentration of 5 mg/ml, which was rendered oxygen free by a flow of nitrogen for at least 10 mins. The catalyst solution was prepared by dissolving the desired amount of CuBr in deaerated THF, which was subsequently deaerated more by a nitrogen flow for at least 10 mins, followed by addition of an excess of PMDETA to dissolve the CuBr, keeping a constant flow of nitrogen.

After all the CuBr was dissolved, the desired amount of the resulting yellow-green catalyst solution was transferred to the reaction vessel with an argon-flushed syringe.

2-Amino-3'[(tert-butoxycarbonyl)amino]ethylene glycol diethyl ether (2)

This compound was synthesized according to a slightly modified literature procedure. To a solution of 2,2'-(ethylenedioxy)-bis(ethylamine) (42 g, 0.28 mol) in 200 ml dioxane, a solution of di-tert-butyl dicarbonate (10.2 g, 0.046 mol, 0.2 eq) in 50 ml dioxane was added dropwise. The solution was stirred overnight under nitrogen and concentrated in vacuo. The residue was dissolved in 200 ml water, and extracted with 4 x 100 ml dichloromethane. The combined organic layers were extracted with 3 x 100 ml brine solution, dried over Na₂SO₄ and concentrated to give 6.24 g of a pale brown oil in an overall yield of 54%, containing a minor impurity of the product in which both amines are protected (~10% determined from ¹H-NMR):

CI-MS (m/z) 249 (MH⁺), 219, 203 (-C(CH₃)₂), 193, 187 (-H₂NCH₂CH₂O), 175 (-OC(CH₃)).

¹H-NMR (400 MHz, CDCl₃) δ 5.17 (br s, 1H, NHBOC), 3.63 (s, 4H, OC₃H₂C₃H₂O), 3.57-3.52 (m, 4H, C₃H₂O), 3.32 (br m, 2H, C₃H₂NHBOC), 2.89 (t, J=5.1Hz, 2H, CH₂NH₂), 1.45 (s, 9H, C₃H₃).

¹³C-NMR (75 MHz, CDCl₃) δ 155.8 (NHCO), 77.8 (C₆H₃), 77.4 and 70.4 (CH₃), 42.0 (CH₂NH₂), 40.6 (CH₂NH), 28.7 (C₃H₃).

2-[6''-Bromohexano amide]-3'[(tert-butoxycarbonyl)amino]ethylene glycol diethyl ether (3)

A solution of 6-bromohexanoyl chloride (757 mg, 3.5 mmol) in 10 ml distilled chloroform was added dropwise to a cooled (0 °C) solution of 2 (880 mg, 3.5 mmol) and DIPEA (0.91 ml, 5.3 mmol) in 20 ml distilled chloroform and allowed to stir overnight under a nitrogen atmosphere. The solvent was evaporated under reduced pressure, after which the residue was dissolved in a small amount of ethyl acetate and filtered to remove the hydrochloric acid salt of DIPEA. The filtrate was subjected to silica column chromatography (MeOH in ethyl acetate: 0-6% v/v), pure fractions were collected, dried over Na₂SO₄ and concentrated in vacuo to yield 1.05 g (72%) of a light brown oil:

CI-MS calcld. For C₁₇H₃₃N₂O₅: 424.1573, found: 425.1657 (+H⁺).

¹H-NMR (400 MHz, CDCl₃) δ 6.01 and 4.98 (br s, 1H, NHBOC), 3.61 (s, 4H, OCH₂CH₂O), 3.57-3.52 (m, 4H, CH₂O), 3.47 (m, 2H, CH₂NH₂), 3.41 (t, J=6.7Hz, 2H, CH₂Br), 3.33 (br m, 2H, CH₂NHBOC), 2.21 (t, J=7.4Hz, 2H, CH₂CO), 1.88 (m, 2H, CH₂), 1.68 (m, 2H, CH₂), 1.49 (m, 2H, CH₂), 1.45 (s, 9H, CH₃).

¹³C-NMR (75 MHz, CDCl₃) δ 76.6 (C(CH₃)₃), 70-69 (C₃H₂O), 40.1 and 39.0 (NHCH₃), 36.1, 33.4, and 32.3 (CH₃), 28.2 (CH₃), 27.6 (CH₃), 24.6 (CH₂Br).

5-Maleimido caproic acid N-hydroxysuccinimide ester (5)

This compound was prepared according to a literature procedure. 5-Maleimido caproic acid 4 (125 mg, 0.59 mmol) and distilled triethylamine (0.12 ml, 0.89 mmol) were dissolved in 10 ml distilled dichloromethane, after which disuccinimidylcarbonate (DSC, 228 mg, 0.89 mmol) was added and the solution was stirred for 3 hrs in the dark under a nitrogen flow. The solvent was concentrated in vacuo and subjected to silica column chromatography using ethyl acetate in acetonitrile 50% v/v as eluent. Fractions
containing pure product were isolated, dried with MgSO₄, filtered and concentrated in vacuo to yield 140 mg (77%) of a colourless oil.

Alternatively, an appropriate amount of the ester was prepared using 1.0 eq DIPEA and 1.0 eq DSC and used in the next step without purification.

**1H-NMR (400 MHz, CDCl₃)** δ 6.69 (s, 2H, CH₂), 3.35 (t, 2H, 3J=7.2Hz, NC₃H₂), 2.83 (s, 4H, CH₂), 2.60 (t, 2H, 3J=7.4Hz, CH₂CO), 1.78 (m, 2H, CH₂), 1.64 (m, 2H, CH₂), 1.41 (m, 2H, CH₂).

**13C-NMR (75 MHz, CDCl₃)** δ 170.6, 168.9, 168.2 (CO), 134.0 (C₄H), 37.8 (NCH₂), 31.1, 28.4, 26.2, 26.0, 24.4 (CH₂).

**Maleimide bromo spacer (6)**

(a) To a solution of Boc-protected bromo spacer 3 (132 mg, 0.32 mmol) in a minimal amount of ethyl acetate a solution of HCl in ethyl acetate (~2 mmol/ml, 1 ml) was added, while stirring under nitrogen. After 10 mins, a solid precipitated from the reaction mixture and more HCl solution (0.5 ml) was added. The reaction mixture was stirred for 1 hr, after which the solvent was evaporated under reduced pressure and the excess of HCl was removed by addition and evaporation of EtOAc (1x), and dissolving the product in t-butylalcohol followed by evaporation under reduced pressure (3x). The resulting light brown crude oil was used immediately in the next step.

(b) The solid prepared as described under (a) was dissolved in 5 ml distilled dichloromethane, and adjusted to pH = 10 by addition of trietylamine according to pH indicator paper wetted with water, before adding a solution of 5 prepared using the DSC method on a 0.38 mmol scale (1.2 eq). The resulting solution was stirred overnight under nitrogen in the dark. The solvent was evaporated under reduced pressure and the residue was dissolved in ethyl acetate and filtrated to remove triethylamine salts. The filtrate was subjected to silica column chromatography (MeOH in EtOAc, 0-8% v/v). Fractions containing the product were isolated, evaporated under reduced pressure and purified with a Biobeads column (eluent: distilled dichloromethane) to remove traces of N-hydroxy succinimide. The product was then dissolved in diisopropyl alcohol and precipitated in diisopropyl ether, to yield 30 mg (18%) of a light brown solid.

**ESI-MS** cald. for C₂₂H₃₆BrN₃O₆: 517.1874, found: 518.1866 (+H+).

**1H-NMR (400 MHz, CDCl₃)** δ 6.68 (s, 2H, CH₂), 6.00 and 5.94 (br s, 2H, NH), 3.62 (s, 4H, OCH₂CH₂O), 3.57 (m, 4H, CH₂O), 3.51-3.41 (m, 12H, NHC₃H₂, NC₃H₂, OCH₂, CH₂Br), 2.21 (t, 3J=7.6Hz, 2H COCH₂), 2.18 (t, 3J=7.7Hz, 2H COCH₂), 1.86 (m, 4H, CH₂), 1.67 (m, 4H, CH₂), 1.49 (m, 2H, CH₂), 1.34 (m, 2H, CH₂).

**13C-NMR (75 MHz, CDCl₃)** δ 172.3, 170.3, 133.6, 69.8, 69.7, 69.5, 38.7, 37.2, 35.9, 35.9, 33.2, 32.0, 27.8, 27.3, 25.9, 24.6, 24.3.

**Tritoluylphenoxyporphyrine (7)**

This compound was synthesized according to the Adler method: 4-hydroxybenzaldehyde (4.6 g, 37.5 mmol) and p-tolualdehyde (13.5 g, 113 mmol) were dissolved in 500 ml propionic acid and heated to reflux, followed by addition of freshly distilled pyrrole (10.1 g, 150 mmol). The solution was refluxed for 1.5 hrs, cooled and allowed to crystallize overnight at 4 °C. The solution was filtered and the residue was washed with ethanol until the filtrate was colourless, and dried. The purple solid was purified in 2 g batches with silica column chromatography (MeOH in dichloromethane, 0-3% v/v) to yield a total of 1.2 g (5%) of a purple solid;

**Maldi-TOF MS** (dithranol) m/z calcd 672.8, found 672.3;

**IR** (neat) ν 3321, 2920, 1558, 1507, 1472, 1221, 966, 801 cm⁻¹;

**UV-Vis** (CHCl₃) λ 420, 518, 554, 593, 649 nm;

**1H-NMR (400 MHz, CDCl₃)** δ 8.85 (d, 3J=1.8Hz, β-pyrrole), 8.10 (d, 3J=7.4, 6H, CH toluyl), 8.07 (obs d, 2H, CH phenox), 7.55 (d, 3J=7.6Hz, 6H, CH phenox), 7.21 (d, 3J=8.2Hz, 2H, CH toluyl), 2.70 (s, 9H, CH₃), -2.77 (br s, 2H, NH);
Maleimide functionalized tritolyl porphyrin (8)

(a) Boc protected tritolyl porphyrin 9 (130 mg, 0.13 mmol) was dissolved in 15 ml ethyl acetate by sonication, followed by addition of a solution of HCl in ethyl acetate (~2.5 mmol/ml, 1 ml). After stirring the solution under nitrogen for 10 mins, a solid precipitated from the reaction mixture and the reaction was continued for another hr. The solvent was evaporated under reduced pressure and the residue was washed by addition and evaporation of EtOAc (3x), and dissolving it in t-butylalcohol followed by evaporation under reduced pressure (3x). The resulting green solid was used immediately in the next step.

(b) The residue obtained as described under (a) was dissolved in 10 ml distilled dichloromethane and a sufficient amount of triethyl amine (0.06 ml, 0.43 mmol) was added to make the solution basic. A solution of N-hydroxy succinimide activated ester of 5-maleimidocaproic acid 5, prepared using the DSC method at a 0.14 mmol scale, was added slowly, and the resulting reaction mixture was stirred overnight under a nitrogen atmosphere in the dark.

Maldi-TOF MS (dithranol) m/z calcd: 1109.5, found: 1108.6;

1H-NMR (400 MHz, CDCl3) δ 8.85 (d, J=2.9Hz, 8H, β-pyrrole), 8.11 (d, J=8.4Hz, 2H, CH phenoxy), 8.09 (d, J=7.8Hz, 6H, CH toluyl), 7.54 (d, J=7.8Hz, 6H, CH toluyl), 7.26 (obs d, J=8.4Hz, 2H, CH phenoxy), 6.07 and 5.00 (br s, 2H, NH), 4.24 (t, J=6.3 2H, CH2O), 3.63 (s, 4H, OCH2CH2O), 3.60-3.50 (m, 10H, NHCH2, NCH2, OCH2), 3.35 (br m, 2H, CH2NH), 2.70 (s, 9H, CH3), 2.32 (t, J=7.4Hz, 2H, CH2CO), 2.15 (t, J=7.6Hz, 2H, CH2CO), 2.00 (m, 2H, CH2), 1.85 (m, 4H, CH2), 1.65 (m, 2H, CH2), 1.45 (s, 9H, CH3).

13C-NMR (75 MHz, CDCl3) δ 158.7, 139.2, 137.2, 135.5, 134.4, 133.9, 130.9 (br), 127.4, 120.1, 119.8, 112.69, 70.5, 70.4, 70.2, 70.1, 68.2, 39.5, 39.4, 37.9, 37.0, 36.7, 29.6, 26.4, 25.9, 25.4, 21.9.

Boc protected ethylene glycol functionalized tritolyl porphyrin (9)

Boc-protected bromo spacer 3 (103 mg, 0.25 mmol) and dried K2CO3 (93 mg, 0.67 mmol) were dissolved in 5ml distilled and degassed DMF, after which tritoluylphenoxyporphyrine (114 mg, 0.17 mmol) was added. The reaction was allowed to stir for 60 hrs under a nitrogen atmosphere, after which the reaction was filtered, and the filtrate concentrated in vacuo. The residue was purified using silica column chromatography (MeOH in dichloromethane: 0-3% v/v) and Biobeads size exclusion chromatography (eluent: toluene), to yield 130 mg (75%) of a purple solid:

Maldi-TOF MS (dithranol) m/z calcd: 1016.5, found: 1015.7;

1H-NMR (400 MHz, CDCl3) δ 8.85 (d, J=2.9Hz, 8H, β-pyrrole), 8.11 (d, J=8.4Hz, 2H, CH phenoxy), 8.09 (d, J=7.8Hz, 6H, CH toluyl), 7.54 (d, J=7.8Hz, 6H, CH toluyl), 7.26 (obs d, J=8.4Hz, 2H, CH phenoxy), 6.07 and 5.00 (br s, 2H, NH), 4.24 (t, J=6.3 2H, CH2O), 3.63 (s, 4H, OCH2CH2O), 3.60-3.50 (m, 10H, NHCH2, NCH2, OCH2), 3.35 (br m, 2H, CH2NH), 2.70 (s, 9H, CH3), 2.32 (t, J=7.4Hz, 2H, CH2CO), 2.15 (t, J=7.6Hz, 2H, CH2CO), 2.00 (m, 2H, CH2), 1.85 (m, 4H, CH2), 1.65 (m, 2H, CH2), 1.45 (s, 9H, CH3).

2-[5''-Pentynoic amide]-3'[(tert-butoxycarbonyl)amino]ethylene glycol diethyl ether (10)

5-Pentynoic acid (317mg, 3.23 mmol) and DIPEA (0.55 ml, 3.23 mmol) were dissolved in 5 ml DMF. To this, a solution consisting of 495 mg HOBt monohydrate and 620 mg EDC in 4.3 ml DMF was added. The reaction mixture was stirred overnight under a nitrogen atmosphere, and concentrated in vacuo. The resulting oil was dissolved in 65 ml ethyl acetate and extracted with 2 x 50 ml 5% NaHSO4 solution, 2 x 50 ml saturated NaHCO3 solution, dried over Na2SO4 and concentrated in vacuo to yield 550 mg (76%) of a pale brown oil:

ESI-MS calcd for C8H3Na2O3Na: 351.1896, found: 351.1916;
Synthesis and characterization of maleimide functionalized porphyrin derivatives

IR (neat) v 3200 (br, NH), 1704 (CO), 1648 (CO) cm⁻¹;
¹H-NMR (400 MHz, CDCl₃) δ 6.18 and 5.00 (br s, 2H, CONH), 3.62 (s, 4H, OCH₂CH₂O), 3.58-3.55 (m, 4H, CH₂O), 3.48 (br m, 2H, CH₂NHBOc), 2.54 (m, 2H, CH₂), 2.42 (t, ³J=7.1Hz, 2H, COCH₂), 2.00 (s, 1H, CH), 1.45 (s, 9H, CH₃).
¹³C-NMR (75 MHz, CDCl₃) δ 79.1 (C(CH₃)), 70.0 and 69.9 (C(CH₂)O), 69.6 and 68.9 (C(CH₂)N), 28.3 (CH₃), 14.8 (C(CH₂)C).

2-[5'''-Maleimide caproic amide]-2''-[5'''-pentynoic amide]-ethylene glycol diethyl ether (11)
(a) To a solution of Boc protected alkyne spacer 10 (380mg, 1.22 mmol) in a minimal amount of ethyl acetate a solution of HCl in ethyl acetate (~2.5 mmol/ml, 5 ml) was added, and stirred under nitrogen. After 10 mins, a solid precipitated from the reaction mixture and more of the HCl solution (1.5 ml) was added. The reaction mixture was stirred for 1 hr, after which the solvent was evaporated under reduced pressure and the residue was washed by addition and evaporation of EtOAc (1x), and dissolving it in t-butylalcohol (3x) followed by evaporation under reduced pressure. In the third t-BuOH washing step, some dichloromethane was added. The resulting light brown sticky oil was used immediately in the next step.
(b) The solid prepared as described under (a) was dissolved in 10 ml distilled dichloromethane and made basic by adding of DIPEA to pH = 10, before addition of the solution of 5 prepared using the DSC method on a 1.22 mmol scale. The resulting solution was stirred overnight under nitrogen in the dark. The solvent was evaporated under reduced pressure and the product was subjected to silica column chromatography (MeOH in EtOAc, 0-8% v/v). Fractions containing the product were isolated, evaporated under reduced pressure and purified with a Biobeads column (eluent: distilled THF) to remove traces of N-hydroxy succinimide. Pure fractions were concentrated under reduced pressure to yield 320 mg of a light brown oil (62%), which crystallized into a white crystalline:
EI-MS m/z 421 (M⁺), 299, 237, 211, 193, 186, 124;
Anal cald for C₂₁H₃₁N₃O₆: C, 59.84; H, 7.41; N, 9.97; found: C, 60.52; H, 7.34; N, 9.96;
¹H-NMR (400 MHz, CDCl₃) δ 6.69 (s, 2H, CH), 6.15 and 5.91 (br s, 2H, NH), 3.62 (s, 4H, OCH₂CH₂O), 3.57 (m, 4H, CH₂O), 3.53-3.44 (m, 6H, NHCH₂ and NCH₂), 2.54 (m, 2H, CH₂CCH), 2.42 (t, ³J=7.1Hz, 2H, COCH₂), 2.18 (t, ³J=5.7Hz, 2H, COCH₂), 2.01 (t, ³J=2.5Hz, 1H, CHCH), 1.70-1.57 (m, 4H, CH₂), 1.36-1.24 (m, 2H, CH₂);
¹³C-NMR (75 MHz, CDCl₃) δ 133.6 (C(CH₃)C(CH₃)), 69.8, 69.8, 69.5, 69.4 and 68.8 (OC(CH₂), C(CH₂)CH and C(CH₃)C), 38.8, 38.7 and 37.2 (CH₂N), 36.0 and 34.9 (CH₃CO), 27.8, 25.9, 24.6 and 14.4 (CH₂).

Bromopropyl functionalized tritoluylphenoxyporphyrine (12)
Hydroxyl porphyrin 7 (70 mg, 0.10 mmol) was dissolved in 7 ml DMF, followed by addition of dibromopropane (420 mg, 2.08 mmol) and K₂CO₃ (58 mg, 0.42 mmol). The solution was stirred overnight at 40 °C under a nitrogen atmosphere in the dark, after which the solvent was evaporated under reduced pressure. The residue was dissolved in 25 ml chloroform, extracted three times with 25 ml 5% aqueous KHSO₄, two times 25 ml aqueous saturated NaHCO₃ and 25 ml saturated NaCl solution. The organic layer was dried with Na₂SO₄, concentrated in vacuo, and subjected to silica column chromatography (dichloromethane) to yield 52 mg (63%) of a purple solid:
Maldi-TOF MS (dithranol) m/z calcd: 792.3, found: 794.2;
IR (neat) ν 2099, 1747, 1244 (CO), 668 (C=Br) cm⁻¹;
UV-Vis (CHCl₃) λ, 420, 518, 555, 594, 650 nm;
¹H-NMR (400 MHz, CDCl₃) δ 8.85 (s, 8H, β-pyrrole), 8.08 (d, ³J=7.3Hz, 8H, CH), 7.52 (d, ³J=7.6Hz, 2H, CH toluyl), 7.22 (d, ³J=6.8Hz, 2H, CH phenoxy), 4.32 (d, ³J=5.6Hz, 2H, CH₂O), 3.71 (d, ³J=6.5Hz, 2H, CH₂N), 2.68 (s, 9H, CH₃), 2.46 (m, 2H, CH₂), -2.77 (br s, 2H, NH);
¹³C-NMR (75 MHz, CDCl₃) δ 139.2, 137.2, 135.5, 134.9, 134.4, 131.0 (br), 127.4, 120.1, 119.6, 112.7, 65.8, 32.9 (CH₂Br), 30.5, 21.9.
Azidopropyl functionalized tritoluylphenoxyphorphyrine (13)

Bromopropyl functionalized porphyrin 12 (50 mg, 0.063 mmol) was dissolved in 4 ml DMF, sodium azide (123 mg, 1.89 mmol) was added and the reaction mixture was stirred at 65 °C overnight. The solvent was evaporated in vacuo and the residue was purified using silica column chromatography (hexane in dichloromethane: 80%-40%). Fractions containing product were collected, concentrated under reduced pressure, dissolved in a minimal amount of dichloromethane and precipitated in methanol, to yield 33 mg (70%) of a purple crystalline compound:

**Maldi-TOF MS** (dithranol) m/z calcd: 755.3, found: 755.0, 726.9 (M+-N2);
**IR** (neat) ν 2097 (N3) cm⁻¹;
**UV-Vis** (CH₂Cl₂) λ 420, 518, 553, 591 nm.

**1H-NMR** (400 MHz, CDCl₃) δ 8.85 (s, 8H, β-pyrrole), 8.12 (obs d, 2H, CH phenoxy), 8.09 (d, J=7.8Hz, 2H, CH toluyl), 7.55 (d, J=7.6Hz, 2H, CH toluyl), 7.28 (obs d, 2H, CH phenoxy), 4.36 (d, J=5.8Hz, 2H, CH₂O), 3.71 (d, J=6.6Hz, 2H, CH₂N₃), 2.71 (s, 9H, CH₃), 2.26 (m, 2H, CH₂);

**13C-NMR** (75 MHz, CDCl₃) δ 138.8, 136.8, 135.1, 134.5, 134.5, 130.5 (br), 126.9, 119.6, 112.2, 64.8, 48.0 (CH₂N₃), 28.6, 21.1.

Zinc Azidopropyl functionalized tritoluylphenoxyphorphyrine (Zn13)

Azide porphyrin 13 (25 mg, 0.033 mmol) was dissolved in 10 ml dichloromethane, followed by addition of zinc acetate dihydrate (73 mg, 0.33 mmol) in 5 ml methanol. The solution was refluxed for three hrs in the dark, and the completion of metal insertion was verified with UV-Vis spectroscopy (only 2 Q-bands left) and TLC. The solvent was evaporated in vacuo and the residue was subjected to silica column chromatography (40% hexane in dichloromethane v/v), pure fractions were collected, concentrated under reduced pressure, dissolved in a minimal amount of dichloromethane and precipitated in hexane to yield 21 mg (77%) of a purple solid:

**Maldi-TOF MS** (dithranol) m/z calcd: 817.3, found: 816.8;
**IR** (neat) ν 2100 (N3), 1493, 998, 796 nm;
**UV-Vis** (CH₂Cl₂) λ 421, 549, 591 nm.

**1H-NMR** (400 MHz, CDCl₃) δ 8.96 (s, 8H, β-pyrrole), 8.12 (obs d, 2H, CH phenoxy), 8.09 (d, J=8.0Hz, 6H, CH toluyl), 7.55 (d, J=7.6Hz, 6H, CH toluyl), 7.28 (d, J=8.0Hz, 6H, CH phenoxy), 4.36 (d, J=5.8Hz, 2H, CH₂O), 3.71 (d, J=6.6Hz, 2H, CH₂N₃), 2.71 (s, 9H, CH₃), 2.26 (m, 2H, CH₂);

**13C-NMR** (75 MHz, CDCl₃) δ 158.4, 150.6, 150.5, 140.1, 137.2, 135.7, 135.6, 134.5, 134.5, 130.5 (br), 126.9, 119.6, 112.2, 64.8, 48.4, 29.0, 21.7.

Manganese Azidopropyl functionalized tritoluylphenoxyphorphyrine bromide (Mn13)

Finely crushed manganese dichloride (150 mg, 1.2 mmols) was added to a solution of azide tritoluylporphyrin 13 (90 mg, 0.12 mmols) in 5 ml of DMF. The reaction mixture was heated at 130 °C in the dark during 3 hrs. The solvent was evaporated and the crude residue was dissolved in CHCl₃, and the organic layer was extracted 60 hrs with saturated sodium bromide solution in the dark. The organic layer was concentrated under reduced pressure and purified by column chromatography (5% methanol in dichloromethane) and by precipitation in hexane from dichloromethane, to yield 91 mg (85%) of a green solid:

**Maldi-TOF MS** (dithranol) m/z calcd: 887.2, found: 807.8 (-Br);
**IR** (neat) ν 2927, 2096 (N₃), 1699, 1006, 803 cm⁻¹;
**UV-Vis** (CHCl₃) λ 485, 535, 589, 626 nm;
3-Azidopropanol (14)
This compound was synthesized according to a literature procedure from 3-chloropropanol\textsuperscript{15}, and was distilled under reduced pressure (bp \(47^\circ\text{C}\) at 4 mbar) to yield 74\% of a colourless oil:

**IR** (neat) \(v\) 2096 (N\(_3\)) cm\(^{-1}\);

**\(^{1}\text{H-NMR}\)** (400 MHz, CDCl\(_3\)) \(\delta\) 3.75 (br m, 2H, C\(_2\)H\(_2\)OH), 3.46 (t, \(\delta=6.6\text{Hz}\), 2H, C\(_2\)H\(_2\)N\(_3\)), 2.02 (b s, 1H, OH), 1.84 (m, 2H, CH\(_2\));

**\(^{13}\text{C-NMR}\)** (75 MHz, CDCl\(_3\)) \(\delta\) 59.4 (C\(_2\)H\(_2\)OH), 48.0 (C\(_2\)H\(_2\)N\(_3\)), 31.0 (CH\(_2\)C\(_2\)H\(_2\)).

1-Mesityl-3-azidopropane (15)
A solution of mesyl chloride (4.7 g, 41 mmol) in 40 ml distilled THF was added dropwise to a cooled solution of 3-azidopropanol \(14\) (4.0 g, 40 mmol) and DIPEA (7.0 ml, 41 mmol) in 40ml distilled THF, and stirred at 0 \(^\circ\text{C}\) for 30 mins and stirred at room temperature for another 2 hrs. The solution was then diluted with water and extracted with dichloromethane (3 x 150 ml), dried over MgSO\(_4\) and concentrated in vacuo. The resulting oil was purified with silica column chromatography (ethyl acetate : heptane = 1 : 1 v/v) to yield 5.81 g (85\%) of a colourless oil:

**CI-MS** calcd for C\(_4\)H\(_9\)N\(_3\)O\(_3\)S: 180.0443, found: 180.0441;

**IR** (neat) \(v\) 2097 (N\(_3\)), 1351, 1172 (SO) cm\(^{-1}\);

**\(^{1}\text{H-NMR}\)** (400 MHz, CDCl\(_3\)) \(\delta\) 4.33 (t, \(\delta=6.0\text{Hz}\), 2H, C\(_2\)H\(_2\)S), 3.49 (t, \(\delta=6.4\text{Hz}\), 2H, C\(_2\)H\(_2\)N\(_3\)), 3.04 (s, 3H, C\(_3\)H\(_3\)), 2.01 (m, 2H, C\(_2\)H\(_2\));

**\(^{13}\text{C-NMR}\)** (75 MHz, CDCl\(_3\)) \(\delta\) 66.0 (C\(_3\)H\(_3\)S), 46.9 (C\(_2\)H\(_2\)N\(_3\)), 36.9 (C\(_3\)H\(_3\)), 28.3 (CH\(_2\)C\(_2\)H\(_2\)).

Clicked maleimide functionalized tritoluylporphyrin (16)
This compound was synthesized according to the general procedure starting from 20 mg (0.026 mmol) porphyrin \(13\). The reaction mixture was concentrated in vacuo and subjected to column chromatography (0-10\% MeOH in dichloromethane v/v), the pure fractions were concentrated, dissolved in a small amount of dichloromethane and precipitated in hexane to yield 30.2 mg (99\%) of a purple solid:

**Maldi-TOF MS** (dithranol) \(m/z\) calcd: 1176.6, found: 1177.6;

**IR** (neat) \(v\) 3426, 2923, 1700, 1652, 801;

**UV-Vis** (CHCl\(_3\)) \(\lambda\) 422, 517, 554, 593, 650 nm;

**\(^{1}\text{H-NMR}\)** (400 MHz, CDCl\(_3\)) \(\delta\) 8.85 (s, 8H, \(\beta\)-pyrrole), 8.12 (d, \(\delta=8.5\text{Hz}\), 2H, CH phenoxy), 8.09 (d, \(\delta=7.7\text{Hz}\), 2H, CH toluyl), 7.55 (d, \(\delta=7.9\text{Hz}\), 2H, CH toluyl), 7.24 (obs d, 2H, CH phenoxy), 6.68 (s, 1H, triazole CH), 6.60 (s, 2H, CH\(_3\)), 6.34 and 6.28 (s, 2H, NH\(_2\)), 4.72 (t, \(\delta=6.8\text{Hz}\), 2H, CH\(_2\)O-porph), 3.56 (s, 4H, OCH\(_2\)CH\(_2\)O), 3.54-3.41 (m, 10H, CH\(_2\)O, NHC\(_2\)H\(_2\) and NC\(_2\)H\(_2\)), 2.70 (s, 9H, CH\(_3\)), 2.68 (t, \(\delta=7.2\text{Hz}\), 2H, COCH\(_2\)), 2.57 (m, 2H, CH\(_2\)CH\(_2\)CH\(_2\)), 2.16 (t, \(\delta=7.5\text{Hz}\), 2H, COCH\(_2\)), 1.67-1.54 (m, 6H, CH\(_3\)), -2.77 (br s, NH);

**\(^{13}\text{C-NMR}\)** (75 MHz, CDCl\(_3\)) \(\delta\) 172.5, 171.5, 170.3, 157.8, 146.1, 138.8, 136.9, 135.2, 134.7, 134.0, 133.5, 130.5 (br), 126.9, 121.5, 119.6, 119.0, 112.2, 76.7, 69.8, 69.7, 69.5, 69.3, 64.0, 46.7, 38.7, 38.7, 37.2, 35.8, 35.3, 29.8, 27.8, 25.9, 24.6, 21.1.

Clicked Zinc maleimide functionalized tritoluylporphyrin (Zn16)
This compound was synthesized according to the general procedure starting from 15 mg (0.018 mmol) porphyrin Zn\(13\). The reaction mixture was concentrated in vacuo and subjected to column chromatography (methanol in dichloromethane: 0-4\% v/v) and precipitated multiple times in hexane from dichloromethane to yield 20 mg (91\%) of a purple solid:

**Maldi-TOF MS** (dithranol) \(m/z\) calcd: 1238.5, found: 1237.8;

**IR** (neat) \(v\) 3300, 2923, 2869, 1705, 1649, 995, 797;

**UV-Vis** (CHCl\(_3\)) \(\lambda\) 422, 551, 591 nm;
\[ ^1H\text{-NMR} \ (400 \text{ MHz, CDCl}_3, \text{ supplemented with small amount of pyridine-}d_5) \delta \ 8.88-8.84 \ (m, 8H, \beta\text{-pyrrole}), \ 8.09 \ (\text{obs d, 2H, CH phenoxy}), \ 8.07 \ (d, 3J=7.9Hz, 6H, \text{C}H\text{toluyl}), \ 7.61 \ (\text{br s, triazole C}H), \ 7.51 \ (d, 3J=8.0Hz, 6H, \text{C}H\text{toluyl}), \ 7.21 \ (d, 3J=8.3Hz, 2H, \text{CH phenoxy}), \ 6.60 \ (s, 2H, \text{CH}), \ 6.45 \text{ and } 6.35 \ (\text{br s, 2H, NH}), \ 4.70 \ (t, 3J=6.8Hz, 2H, \text{C}H_2\text{OPorf}), \ 4.24 \ (t, 3J=5.5Hz, 2H, \text{C}H_2\text{-triazole}), \ 3.54 \ (s, 4H, \text{OC}H_2\text{C}H_2\text{O}), \ 3.52-3.39 \ (m, 10H, \text{C}H_2\text{O, NHC}H_2\text{ and NC}H_2), \ 3.11 \ (t, 3J=7.1Hz, 2H, \text{C}H_2\text{-triazole}), \ 2.69 \ (\text{obs t, 3J=7.2Hz, COC}H_2), \ 2.55 \ (m, 2H, \text{C}H_2), \ 2.13 \ (t, 3J=7.5Hz, 2H, \text{COC}H_2), \ 1.66-1.52 \ (m, 6H, \text{C}H_2); \]

\[ ^{13}C\text{-NMR} \ (75 \text{ MHz, CDCl}_3, \text{ supplemented with small amount of pyridine-}d_5) \delta \ 172.5, \ 171.6, \ 170.3, \ 157.3, \ 149.7, \ 148.6, \ 148.3, \ 146.1, \ 140.1, \ 136.1, \ 135.1, \ 135.0, \ 134.7, \ 134.6, \ 134.0, \ 133.5, \ 131.0, \ 130.8, \ 126.5, \ 123.0, \ 122.6, \ 121.5, \ 120.0, \ 119.3, \ 111.8, \ 69.7, \ 69.7, \ 69.5, \ 69.3, \ 63.9, \ 46.7, \ 38.7, \ 38.6, \ 37.2, \ 35.8, \ 35.3, \ 29.8, \ 29.2, \ 27.8, \ 25.9, \ 24.6, \ 21.0. \]

Clicked manganese maleimide functionalized tritoluylporphyrin bromide (Mn16)

This compound was synthesized according to the general procedure starting from 31 mg (0.035 mmol) porphyrin Mn13. The solvent was evaporated under reduced pressure and the crude precipitate was purified with silica column chromatography (methanol in dichloromethane: 0-7% v/v) multiple times, followed by precipitation in hexane from dichloromethane to yield 21 mg (45%) of a green solid.

Maldi-TOF MS (dithranol) m/z calcd: 1308.4, found: 1228.3 (-Br);

\[ \text{IR (neat)} \ \nu \ 2918, \ 1701, \ 1009, \ 802; \]

\[ \text{UV-Vis (CHCl}_3) \ \lambda \ 485, \ 534, \ 589, \ 626 \text{ nm.} \]

Propionic acid ester of tripyridylphenoxyporphyrine (17)

This compound was synthesized according to the Adler method\textsuperscript{21}: 4-hydroxybenzaldehyde (2.75 g, 22.5 mmol) and acetic acid anhydride (12ml, 128 mmol) were dissolved in 180 ml propionic acid and heated to reflux (130 °C). Then 4-pyridinecarboxaldehyde (7.24 g, 67.6 mmol) and freshly distilled pyrrole (6.04 g, 90.1 mmol) were added and the resulting solution was refluxed for 1.5 hrs. The solvent was evaporated in vacuo and the residue was dissolved in dichloromethane and extracted with saturated NaHCO\textsubscript{3} solution until all the acid was neutralized. The organic layer was concentrated and the residue was subjected to silica chromatography (EtOH in dichloromethane, 0-8% v/v) two times, and to basic alumina chromatography (MeOH in dichloromethane, 0-1% v/v) two times, and to basic alumina chromatography (MeOH in dichloromethane, 0-1% v/v) to yield 0.72 g (5%) of a purple solid:

Maldi-TOF MS (dithranol) m/z calcd: 689.3, found: 690.4;

\[ \text{IR (KBr)} \ \nu \ 1752, \ 1588 (\text{CO}) \ \text{cm}^{-1}; \]

\[ \text{UV-Vis (CH}_2\text{Cl}_2) \ \lambda \ 418, \ 515, \ 549, \ 589, \ 646 \text{ nm}; \]

\[ ^1H\text{-NMR} \ (400 \text{ MHz, CDCl}_3) \delta \ 9.06 \ (d, 3J=5.7Hz, 6H, \text{CH pyridyl}), \ 8.95 \ (d, 3J=4.7Hz, 2H, \beta\text{-pyrrole}), \ 8.86 \ (s, 4H, \beta\text{-pyrrole}), \ 8.84 \ (d, 3J=4.8Hz, 2H, \beta\text{-pyrrole}), \ 8.21 \ (d, 3J=8.3Hz, \text{CH phenoxy}), \ 8.16 \ (d, 3J=5.7Hz, \text{CH pyridyl}), \ 7.53 \ (d, 3J=8.4Hz, \text{CH phenoxy}), \ 2.81 \ (q, 3J=7.5Hz, 2H, \text{CH}_2), \ 1.43 \ (t, 3J=7.5Hz, 3H, \text{CH}_3), \ -2.89 \ (\text{br s, 2H, NH}); \]

\[ ^{13}C\text{-NMR} \ (75 \text{ MHz, CDCl}_3) \delta \ 149.4, \ 147.9, \ 134.8, \ 130.4 (\text{br}), \ 128.9, \ 119.6, \ 119.5, \ 119.5, \ 117.0, \ 27.5, \ 8.7. \]

Hydroxy tripyridylphenoxyporphyrine (18)

The propionic ester of tripyridylporphyrin 17 (105 mg, 0.152 mmol) was saponified by dissolving the compound in Tesser’s base (15ml dioxane, 4ml MeOH, 2ml 2N NaOH) and stirring the reaction mixture for 30 mins in the dark, after which the reaction was quenched by addition of 5 ml 1N HCl solution. A saturated solution of NaHCO\textsubscript{3} was added until the porphyrin precipitated out of the solution. The precipitate was isolated by centrifugation (3700 rpm, 5 mins) and washed two times with water followed by centrifugation. The porphyrin was dried under high vacuum (0.4 mbar) for a prolonged period to yield 92 mg (95%) of a purple crystalline compound which was used immediately in the next reaction:
Synthesis and characterization of maleimide functionalized porphyrin derivatives

Maldi-TOF MS (dithranol) m/z calcd: 633.8, found: 633.8;
IR (neat) ν 1591 (CO) cm⁻¹.

Azide functionalized tripyridylphenoxyporphyrine (19)
Deprotected porphyrin 18 (88 mg, 0.14 mmol) was dissolved in 10 ml DMF, followed by addition of freshly crushed sodium hydroxide (56 mg, 1.4 mmol), and the solution was stirred for 1.5 hrs in the dark. A solution of 1-mesyl-3-azidopropane (46 mg, 0.25 mmol) in 1 ml of DMF was then added to the green solution and stirred overnight in the dark under a nitrogen atmosphere. The reaction was then quenched by addition of 2 ml 1N aqueous HCl solution, poured into a saturated aqueous NaHCO₃ solution and extracted with dichloromethane and concentrated in vacuo. The residue was purified by silica column chromatography (methanol in dichloromethane: 0-4% v/v), pure fractions were isolated, concentrated in vacuo, dissolved in a minimal amount of dichloromethane and precipitated in hexane to yield 77 mg (77%) of a purple crystalline compound:
Maldi-TOF MS (dithranol) m/z calcd: 716.3, found: 716.8;
IR (neat) ν 2096 (N₃), 1592 (CO) cm⁻¹;
UV-Vis (CH₂Cl₂) λ 418, 515, 549, 589, 646 nm;
¹H-NMR (400 MHz, CDCl₃) δ 9.02 (br s, 6H, pyridyl C₆H), 8.88 (d, J=4.6, 2H, β-pyrrole), 8.82 (s, 4H, β-pyrrole), 8.19 (d, J=4.8Hz, 6H, pyridyl CH), 7.83 (d, J=8.3Hz, 2H, phenyl CH), 4.36 (t, J=5.8Hz, 2H, CH₂O), 3.79 (d, J=6.7Hz, 2H, CH₂N₃), 2.19 (qt, J=6.5Hz, 2H, CH₂CH₂);
¹³C-NMR (75 MHz, CDCl₃) δ 149.6, 149.5, 147.9, 135.2, 133.6, 130.5 (br), 128.9, 121.1, 116.9, 116.4, 112.5, 64.4 (CH₂O), 47.9 (CH₂N₃), 28.5 (CH₂CH₂).

Zinc Azide functionalized tripyridylphenoxyporphyrine acetate (Zn19)
The azide porphyrin (25 mg, 0.035 mmol) was dissolved in 10 ml chloroform with 0.05 ml (0.35 mmol) collidine and supplemented with zinc acetate dehydrate (77 mg, 0.35 mmol) in 5 ml methanol. The reaction mixture was heated to reflux in the dark for 3 hrs, after which the reaction mixture was extracted with water. The organic layer was concentrated and subjected to silica column chromatography (methanol in dichloromethane, 1:10 v/v with some collidine) and porphyrin containing fractions isolated, concentrated and precipitated in hexane from a chloroform solution to yield 25.7 mg (94%) of a purple solid:
Maldi-TOF MS (dithranol) m/z calcd: 778.2, found: 779.6;
UV-Vis (CHCl₃) λ 425, 556, 597 nm;
¹H-NMR (400 MHz, DMSO) δ 9.02 (br s, 6H, pyridyl CH), 8.88 (d, J=4.6, 2H, β-pyrrole), 8.82 (s, 4H, β-pyrrole), 8.19 (d, J=4.5Hz, 2H, pyridyl CH), 8.09 (d, J=8.3Hz, 2H, phenyl CH), 7.38 (d, J=8.3Hz, 2H, phenyl CH), 4.36 (t, J=6.0Hz, 2H, CH₂O), 3.69 (d, J=6.7Hz, 2H, CH₂N₃), 2.19 (qt, J=6.5Hz, 2H, CH₂CH₂);
¹³C-NMR did not yield sufficient signal intensity due to aggregation.

Manganese Azide functionalized tripyridylphenoxyporphyrine acetate (Mn19)
The azide porphyrin (110 mg, 0.153 mmol) was dissolved in 15 ml DMF and heated to 100 °C, after which collidine (0.2 ml, 1.53 mmol) and Mn(OAc)₂·4H₂O (376 mg, 1.53 mmol) were added. The reaction mixture was then heated to 150 °C for 3 hrs in the dark under air. UV-Vis spectroscopy was used to ascertain the complete insertion of the manganese, after which the solution was cooled, diluted with dichloromethane and extracted with water. The organic layer was evaporated under reduced pressure, followed by purification of the resulting green crystals by basic alumina column chromatography (0-2% methanol in dichloromethane v/v). The porphyrin containing fractions were isolated, concentrated in vacuo, dissolved
in a minimal amount of dichloromethane and precipitated in hexane to yield 89 mg (76%) of a green crystalline compound:

**Maldi-TOF MS** (dithranol) \( m/z \) calcd: 828.2, found: 768.7 (-OAc);
**UV-Vis** (DMF) \( \lambda \) 438, 464, 571, 606 nm.

**Clicked maleimide functionalized tripyridylporphyrin (20)**
This compound was synthesized according to the general procedure with 10 mg (0.014 mmol) porphyrin. The solvent was evaporated under reduced pressure and the crude precipitate was purified by silica column chromatography (methanol in dichloromethane: 0-10% v/v), followed by precipitation in hexane from dichloromethane to yield a mixture of free base and Cu-inserted product (with a ratio of 3:1 based on ms results) as a red solid:

**Maldi-TOF MS** (dithranol) \( m/z \) calcd: 1337.5, found: 1138.8 (M*), 1198.7 (M*+Cu), 1364.9 (M*+matrix), 1425.8 (M*+Cu+matrix);
**UV-Vis** (CH\(_2\)Cl\(_2\)) \( \lambda \) 417, 515, 541, 589, 646;
**IR** (neat) \( \nu \) 3316, 2933, 1703, 1657, 1593, 1245, 799.

**Clicked manganese maleimide functionalized tripyridylporphyrin acetate (Mn20)**
This compound was synthesized according to the general procedure with 21 mg (0.025 mmol) porphyrin and 1.0 eq CuBr. The reaction mixture was diluted with dichloromethane and extracted with saturated NaHCO\(_3\). Any precipitate in the reaction vessel was dissolved in water by addition of a few drops of 1M HCl solution, and added to the aqueous phase. The aqueous phase was extracted twice with dichloromethane. The organic layers were combined and washed with a 0.3M NaOAc buffer (pH 5.2) supplemented with EDTA. The organic layer was concentrated *in vacuo* and precipitated in hexane to yield 24.9 mg (66%) of a green solid:

**Maldi-TOF MS** (dithranol) \( m/z \) calcd: 1249.4, found: 1190.8 (M*-OAc), 1416.7 (M*+matrix), 1488.7 (M+Na*+matrix);
**UV-Vis** (10%EtOH in DMF) \( \lambda \) 465, 566, 605 nm.

**Methylated clicked manganese maleimide functionalized tripyridylporphyrin acetate (Mn21)**
A solution of Mn20 (25 mg, 0.02 mmol) in DMF was heated to 40 °C in the dark, followed by addition of methyl iodide (0.19 ml, 3.0 mmol). After one hr, the solvent was evaporated, the residue dissolved in a minimal amount of DMF and precipitated in ethyl acetate several times and extensively dried in high vacuum to yield 32 mg (120%) of a gray shiny oil.

No mass data could be obtained with Maldi-TOF or ESI-MS.
**UV-Vis** (H\(_2\)O) \( \lambda \) 466, 564, 601 nm.

**4-trimethylsilanylethynyl-benzaldehyde (22)**
This compound was synthesized according to a modified literature procedure.\(^{32}\) A solution of 4-bromobenzaldehyde (5.00 g, 27 mmol), PdCl\(_2\)(PPh\(_3\))\(_2\) (0.189 g, 0.27 mmol) and CuBr (35.9 mg, 0.27 mmol) in 120 ml distilled Et\(_3\)N was treated with trimethylsilylacetylene (5.5 ml, 38 mmol) and heated to reflux. After 16 hrs, the reaction mixture was cooled down and filtered to remove triethylamine salts. The filtrate was concentrated under reduced pressure and the crude material was purified by silica column chromatography (CH\(_2\)Cl\(_2\)/pentane 1:3 v/v) and by repeated recrystallization from petroleum ether at -20 °C to yield 3.3 g (61%) of a tan crystalline solid:

**mp** 67.2 °C;
**ESI-MS** calcd for C\(_{12}\)H\(_8\)OSi: 202.0814, found: 202.0814;
**Anal** calcd for C\(_{12}\)H\(_8\)OSi : C, 71.24; H, 6.97; found: C, 71.52; H, 7.08;
Synthesis and characterization of maleimide functionalized porphyrin derivatives

IR (neat) ν 2156 (CC), 1699 (CHO), 1600 (CC) cm⁻¹;

¹H-NMR (400 MHz, CDCl₃) δ 10.0 (s, 1H, CHO), 7.82 (d, 3J=8.4Hz, 2H, CH), 7.60 (d, 3J=8.4Hz, 2H, CH), 0.27 (s, 9H, CH₃);

¹³C-NMR (75 MHz, CDCl₃) δ 191.6 (C-CHO), 144.4, 135.7, 129.6, 129.1, 104.0 (C=C), 99.2 (C=C), -0.1 (C-CH₃).

Trimethylsilylethynyl tripyridylporphyrine (23)
This compound was synthesized according to Alder’s method. Benzaldehyde 22 (3.27 g, 16.2 mmol) and 4-pyridinecarboxaldehyde (5.19 g, 48.5 mmol) were dissolved in 140 ml propionic acid and heated until reflux, followed by addition of pyrrole (4.5 ml, 64.7 mmol). The reaction mixture was heated in the dark for 1.5 hrs, after which the solvent was evaporated under reduced pressure. The residue was dissolved in chloroform and washed with saturated NaHCO₃ solution (2x) and brine solution. The organic layer was concentrated in vacuo and subjected to silica column chromatography (methanol in dichloromethane 0-10% v/v) and neutral alumina chromatography (activity III, methanol in dichloromethane 0-1.5% v/v) multiple times to yield 0.77 g (6.5%) of a purple solid:

Maldi-TOF MS (dithranol) m/z calcd: 713.3, found: 713.0;

UV-Vis (CHCl₃) λ 418, 514, 549, 589, 645 nm;

IR (neat) ν 2156 (CC) cm⁻¹;

¹H-NMR (400 MHz, CDCl₃) δ 9.05 (d, 3J=5.1Hz, 6H, C₆H pyridyl), 8.86 (m, 8H, β-pyrrole), 8.16 (d, 3J=5.0Hz, 6H, C₆H pyridyl), 8.15 (obs d, 2H, C₆H phenyl), 7.89 (d, 3J=7.9Hz, 2H, C₆H phenyl), 0.38 (s, 9H, C₆H₃), 0.28-0.24 (br, 7H, unknown), -2.89 (br, 2H, NH);

¹³C-NMR (75 MHz, CDCl₃) δ 149.4, 147.9, 145.1, 141.2, 133.9, 130.6 (br), 130.0, 128.8, 122.6, 120.2, 117.1, 116.8, 95.6, 29.2, -0.4.

Ethynyl tripyridylporphyrine (24)
TMS protected porphyrin 23 (156 mg, 0.22 mmol) was dissolved in 15 ml of a methanol/chloroform mixture (1:2 v/v), anhydrous potassium carbonate (302 mg, 2.2 mmol) was added and the solution was stirred at 40 °C in the dark for 2.5 hrs. The organic phase was diluted with chloroform (20 ml) and extracted with 5% KHSO₄ solution, followed by neutralization with a saturated NaHCO₃ solution. The organic phase was washed with brine solution, concentrated in vacuo and subjected to silica column chromatography (methanol in dichloromethane, 0-6% v/v) to yield 125 mg (88%) of a purple crystalline compound:

Maldi-TOF MS (dithranol) m/z calcd: 641.2, found: 642.8;

IR (neat) ν 2106 (CC) cm⁻¹;

¹H-NMR (400 MHz, CDCl₃) δ 9.06 (d, 3J=5.6Hz, 6H, CH pyridyl), 8.90 (d, 3J=4.7Hz, 2H, β-pyrrole), 8.86 (s, 4H, β-pyrrole), 8.84 (d, 3J=4.8Hz, 2H, β-pyrrole), 8.17 (m, 8H, C₆H pyridyl and C₆H phenyl), 7.92 (d, 3J=8.0Hz, CH phenyl), 3.34 (s, 1H, C-CH), -2.89 (br s, 2H, NH);

¹³C-NMR (75 MHz, CDCl₃) δ 149.4, 147.9, 145.1, 141.2, 133.9, 130.6 (br), 130.0, 128.8, 122.6, 120.2, 117.1, 116.8, 95.6, 29.2, -0.4.

Manganese ethynyl tripyridylporphyrine (Mn24)
Porphyrin 24 (70 mg, 0.11 mmol) was dissolved in DMF, collidine (0.15 ml, 1.1 mmol) was added and the solution was heated at 100 °C. Then, finely crushed manganese chloride (138 mg, 1.1 mmol) was added and the solution was heated to reflux and stirred for 3 hrs in the dark. The solvent was evaporated under reduced pressure, the residue dissolved in chloroform (15 ml) and stirred with saturated aqueous sodium bromide solution for 60 hrs. The organic phase was subsequently dried with Na₂SO₄, concentrated, and purified by silica column chromatography (methanol in dichloromethane, 0-10% v/v) to yield 55 mg (65%) of a green solid:
Maldi-TOF MS (dithranol) m/z calcd: 773.1, found: 694.3 (M-Br);
UV-Vis (DMF) λ 466, 566, 613 nm.

Maleimide functionalized azide spacer (25)

**Route 1** Maleimide bromide 6 (30 mg, 0.058 mmol) was dissolved in 2 ml DMF, followed by addition of sodium azide (19 mg, 0.29 mmol). The reaction mixture was heated to 50 °C and stirred 16 hrs in the dark, concentrated *in vacuo* and subjected to silica column chromatography (methanol in ethyl acetate, 10-30% v/v). Fractions were analyzed by NMR spectroscopy, which failed to identify any pure product.

**Route 2 (a)** To a solution of Boc protected azide spacer 26 (143 mg, 0.37 mmol) in a minimal amount of ethyl acetate was added a solution of HCl in ethyl acetate (~2.5 mmol/ml, 2 ml), and stirred under nitrogen. After 10 mins, a solid precipitated from the reaction mixture and more HCl solution (1 ml) was added. The reaction mixture was stirred for 1 hr, after which the solvent was evaporated under reduced pressure and the residue was washed by addition and evaporation of EtOAc (1x), and dissolving it in *t*-butylalcohol followed by evaporation under reduced pressure (3x). In the third *t*-BuOH washing step, some dichloromethane was added. The resulting light brown sticky oil was used immediately in the next step.

**(b)** The solid prepared as described under (a) was dissolved in 10ml distilled dichloromethane and made basic by adding DIPA to a pH of 10 (as determined with indicator paper wetted with water), before addition of the solution of 5 prepared by using the DSC method on a 0.41 mmol scale. The resulting solution was stirred for 16 hrs under nitrogen in the dark. The solvent was evaporated under reduced pressure and the product was subjected to silica column chromatography (MeOH in EtOAc, 0-8% v/v). Fractions containing the product were isolated, evaporated under reduced pressure and purified with a Biobeads column (eluent: distilled THF) to remove traces of N-hydroxy succinimide. No pure product was isolated.

**Boc protected azide spacer (26)**

Bromide functionalized spacer 3 (237 mg, 0.557 mmol) was dissolved in 5 ml DMF, sodium azide (182 mg, 2.80 mmol) was added and the solution was stirred 16 hrs in the dark at 50 °C. The reaction mixture was filtered, concentrated *in vacuo* and subjected to silica column chromatography (methanol in ethyl acetate, 0-6% v/v), to yield 156 mg (72%) of a brownish oil:

**CI-MS** calcd for C17H33N5O5: 388.256, found 388.256;

**1H-NMR** (400 MHz, CDCl3) δ 6.01 and 4.99 (br s, 2H, NH), 3.61 (s, 4H, OCH2CH2O), 3.56 (t, J=3.56Hz, 2H, CH2O), 3.47 (dd, 2H, J=5.2, J=10.1Hz, CH2N), 3.33 (br m, 2H, CH2NH), 3.27 (t, J=6.81Hz, 2H, CH2CO), 2.20 (t, J=7.40Hz, 2H, CH2N3), 1.7-1.6 (m, 4H, CH2), 1.45 (s, 9H, CH3), 1.41 (obs m, 2H, CH2);

**13C-NMR** (75 MHz, CDCl3) δ 172.2 and 155.5 (C=O), 79.0 (CH3), 69.8, 69.7, and 69.5 (CH2O), 50.8 (CH2N3), 39.9 and 38.7 (NHCH2), 35.9 (CH2CO), 28.2 (CH3), 27.9, 25.9 and 24.7 (CH2).

1,11-diazo-3,6,9-trioxaundecane (27)

This compound was prepared according to a literature procedure. The obtained brown oil was purified with flash silica chromatography (ethyl acetate in hexane, 1:1 v/v) to yield a colourless oil with a minor impurity based on GC, but not observed by NMR:

**IR** (neat) v 2100 (N3) cm⁻¹;

**CI-MS** calcd for C11H16N6O3: 245.1362, found 245.1366;

**GC** (chrompack cp sil-52, 200°C, min): 5.038 (14%) and 6.837 (86%);

**1H-NMR** (400 MHz, CDCl3) δ 3.68 (m, 12H, CH2O), 3.39 (t, J=5.2Hz, 2H, CH2N3);

**13C-NMR** (75 MHz, CDCl3) δ 72.0 and 69.6 (CH2O), 50.2 (CH2N3).
1-amino, 11-diazo-3,6,9-trioxaundecane (28)
This compound was prepared according to a literature procedure. The product was purified with flash silica chromatography (ammonia in ethanol, 5% v/v) to yield 74% of a slightly coloured oil:
IR (neat) ν 3392 (NH2), 2105 (N3) cm⁻¹;
CI-MS calcd for C8H18N4O3: 219.1457, found 219.1451;
GC (chrompack cp sil-52, 200°C, min): 5.205 (100%);
1H-NMR (400 MHz, CDCl3) δ 3.68 (m, 12H, CH₂O), 3.52 (t, J=3.52Hz, 2H, CH₂NH₂), 3.40 (t, J=4.79Hz, 2H, CH₂N₃), 2.87 (br t, NH₂);
13C-NMR (75 MHz, CDCl3) δ 72.9, 70.2, 70.2, 69.8 and 69.6 (CH₂O), 50.2 (CH₂N₃), 41.3 (CH₂NH₂).

Azide maleimide spacer (29)
To a solution of 5, prepared in 10 ml CH₂Cl₂ using the DSC method on a 250mg (1.18 mmol) scale, was added amino azide 28 (284 mg, 1.30 mmol). The solution was stirred for 16 hrs in the dark, concentrated under reduced pressure and purified using silica column chromatography (methanol in ethyl acetate, 0-8% v/v). Fractions containing product were collected, concentrated and purified using a Bio-Beads size exclusion column (CH₂Cl₂ eluent), after which no product was isolated. After the silica column, combined fractions (220 mg) were analyzed with NMR:
1H-NMR (400 MHz, CDCl3) δ 6.69 (2H, s, CH), 6.48 and 6.23 (br s, 2H, NH), 5.56, 4.65 and 4.05-3.70 (m, impurities), 3.68 (s, 8H, CH₂O), 3.64-3.39 (m, 18H, CH₂O, CH₂N, CH₂CO, CH₂N₃ and impurities), 2.73 (s, integrates as 5.75, CH₂NHS), 2.19, 1.62, 1.31 (m, 6H, CH₂).

Clicked Boc-protected tripyridyl porphyrine (30)
This compound was synthesized using the general procedure using 10 mg (0.016 mmol) porphyrin and various amounts of CuBr (0.2-5 eq). The reaction mixture was concentrated in vacuo and subjected to column chromatography (methanol in dichloromethane, 0-10% v/v) and precipitation in hexane from dichloromethane to yield 13.5 mg (77%) of Cu-inserted product (using 5 eq CuBr):
Maldi-TOF MS (dithranol) calcd: 1089.4, found: 990.5 (M+-Boc), 1089.6 (M+);
UV-Vis (CHCl₃) λ 415, 539 nm.

Azide functionalized protoporphyrin IX (2H31, Mn31 and Zn31)
The synthesis of these compounds will be described elsewhere.

N-caproic maleimide propargyl amide (32)
Maleimide caproic acid (250 mg, 1.2 mmol) was dissolved in dichloromethane (15 ml), followed by the addition of DIPEA (0.25 ml, 1.5 mmol) and disuccinimidyl carbonate (304 mg, 1.4 mmol). The reaction mixture was stirred for 3.5 hrs under a nitrogen atmosphere, after which propargylamine hydrochloride (131 mg, 1.4 mmol) and DIPEA (0.25 ml, 1.5 mmol) were added. The reaction mixture was allowed to stir for 15 hrs in the dark under a nitrogen atmosphere, upon which it was diluted with dichloromethane to 50 ml, and extracted with aqueous 5% NaHSO₄ solution (two times 50 ml), saturated aqueous NaHCO₃ solution (3 times 50 ml) and brine (50 ml). The organic layer was dried with Na₂SO₄, filtered and concentrated. The residue was purified with silica column chromatography (elucent EtOAc), to yield 126 mg (43%) of a white solid:
ESI-MS calcd for C₁₃H₁₆N₃O₃Na: 271,10586, found: 271.10670;
IR (neat) ν 3290, 3258, 2100, 1694 (CO) cm⁻¹;
1H-NMR (400 MHz, CDCl₃) δ 6.68 (s, 2H, CH), 5.69 (br s, 1H, NH), 4.04 (dd, 2H, J=2.4Hz, J=5.1Hz, CH₂), 3.52 (t, J=7.2Hz, 2H, CH₂), 2.23 (t, 1H, J=2.3Hz, CCH), 2.19 (t, 2H, J=7.2Hz, CH₂), 1.71-1.57 (m, 4H, CH₂), 1.32 (m, 2H, CH₂);
**13C-NMR** (75 MHz, CDCl$_3$) δ 172.3, 170.7, 134.0, 79.9, 71.4, 37.5, 35.9, 29.0, 28.1, 26.2, 24.8.

**Maleimide functionalized protoporphyrin IX (2H33)**
This compound was synthesized according to the standard click procedure with 18.8 mg (0.025 mmol) 2H31 with 2 eq CuBr at 33 °C. Analysis of the reaction mixture with MALDI-TOF demonstrated that copper had inserted into the porphyrin during the reaction. Hence, the product was not purified.

**MALDI-TOF MS** (dithranol) calcd: 1010.5, found: 1298.7 (M+Cu+dithranol), 1360.7 (M+2Cu+dithranol);
**UV-Vis** (10% MeOH in CHCl$_3$) λ 405, 534, 572.

**Maleimide functionalized manganese protoporphyrin IX (Mn33)**
This compound was synthesized according to the standard click procedure with 5.3 mg (0.006 mmol) Mn31 at 33 °C. MALDI-TOF analysis showed the formation of multiple products. The reaction mixture was concentrated and purified by extraction or column chromatography (silica or alumina) but no pure product could be isolated.

**MALDI-TOF MS** (dithranol) calcd: 1122.4, found: 965.7, 1109.0, 1165.1.

**Maleimide functionalized zinc protoporphyrin IX (Zn33)**
This compound was synthesized according to the standard click procedure with 7.3 mg (0.009 mmol) Zn31 and 2 eq CuBr at 33 °C. After 19 hrs, the reaction mixture (3 ml) was supplemented with methanol (4 ml) and dichloromethane (80 ml) and extracted with 3 x 30 ml EDTA solution (10 g/L). The organic layer was dried with Na$_2$SO$_4$, concentrated and subjected to silica column chromatography (eluent, 5-15% v/v methanol in chloroform). The product was isolated in a yield of 5.4 mg (57%) as a brown solid:

**MALDI-TOF MS** (dithranol) calcd: 1072.4, found: 1074.3;
**IR** (neat) ν 2922, 1704, 1650;
**UV-Vis** (10% MeOH in CHCl$_3$) λ 419, 547, 584 nm;
**1H-NMR** could not be obtained due to degradation of the material.

5. **References**

Synthesis and characterization of maleimide functionalized porphyrin derivatives


11 C.W. Tornøe, C. Christensen, and M. Meldal, Peptidotriazoles on solid phase: [1,2,3]-triazoles by regiospecific copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides, J. Org. Chem. 67 (9), 3057-3064 (2002).


Chapter 6

In this chapter the expression and purification of a clamp mutant is described. The clamp mutant was labelled with two different chemical oxidation catalysts in order to obtain biohybrid clamp-based catalytic systems. One of these, i.e. the FeBABE oxidation label, was not very active as oxidation catalyst when conjugated to the clamp directly in the direct labelling approach, nor did it show considerable activity in the padlock approach when the FeBABE-labelled peptide was applied. Clamp proteins labelled with the second catalyst, i.e. a manganese(III)porphyrin, were active as oxidation catalyst, and their action on DNA substrates was investigated via the non-specific and clamp loader-mediated binding routes. In the former route macromolecular crowding reagents and the gp43 C-terminus analogue peptide were applied. A novel analysis technique was introduced, allowing the oxidized substrates to be studied at the single molecular level. This technique demonstrated that the porphyrin-labelled clamp protein is able to processively oxidize double stranded DNA substrates.
1. Introduction

This chapter describes the expression and labelling of a clamp mutant, as well as the oxidation experiments carried out with the catalysts conjugated to the clamp mutants. As outlined in Chapter 2, two different catalysts were identified as candidates for conjugation with the clamp. One is a maleimide functionalized manganese(III) trimethylpyridinium porphyrin 1 (Figure 1A), which was described in Chapter 6. The well documented interactions of cationic manganese porphyrins with DNA and the ability of methylated derivatives thereof to selectively oxidize and cleave AAA sequences of dsDNA templates after activation with an oxygen donor\(^1\) makes this catalyst an interesting candidate for conjugation. Moreover, manganese porphyrins can serve as oxidation catalysts for small organic substrates\(^2\) and synthetic polymers\(^3\) as well. In order to allow the porphyrin to bind to the minor groove of DNA, the maleimide was separated from the porphyrin by a water soluble and flexible ethylene glycol spacer (see Chapter 6). A second candidate for conjugation is the commercially available Fe-EDTA complex 2 (FeBABE). When exposed to a reducing agent and hydrogen peroxide, Fe(II)-EDTA undergoes a Fenton-type reaction in which hydroxyl radicals that can attack peptide bonds and degrade DNA are generated.\(^4\) It is

![Figure 1 - Oxidation catalysts used to label the T4 clamp protein. A. Cationic manganese(III) porphyrin maleimide (1), B. FeBABE (2), C. Cationic manganese porphyrin reference catalyst Mn-TMPyP 3.](image_url)
widely in use as reagent for DNA footprinting, a molecular biology technique that allows the
determination of protein binding sites on DNA. The small size of the generated hydroxyl radicals results in a superior resolution compared to other footprinting agents such as certain antibiotics or DNase I. Furthermore, proteins and oligonucleotides have been labelled with Fe-EDTA complexes in order to improve the sensitivity of this technique. In this chapter oxidation experiments with catalysts 1 and 2 conjugated to the clamp and Mn-TMPyP 3 are reported. The latter porphyrin is not conjugated to the T4 clamp, but serves as a positive control to relate the outcome of the oxidation experiments with 1 and 2 to other studies described in the literature.

2. Results and discussion

2.1 Protein expression and labelling

A host of cysteine clamp mutants has been developed in the past decades with the objective to unravel the clamp’s role in processive DNA synthesis and the assembly of the T4 polymerase holoenzyme (see Chapters 3, 5). Most of these mutants, however, failed to comply with our set of requirements for the construction of a catalyst-appended clamp protein. Based on the clamp crystal structure, and on the interaction sites of the clamp with the clamp loader, the catalyst label should ideally be attached to the middle of the protein body on the smooth side of gp45. In this way, a disturbance of the interactions between the clamp and DNA or the clamp and the clamp loader, respectively, is avoided. On the other hand, the label should be located in the vicinity of the DNA substrate. While the latter requirement is not so strict for the FeBABE labelled clamps (the label should be placed within 22 Å of its substrate), it is expected to be critical in the case of the porphyrin label because the porphyrin requires a certain degree of conformational freedom to be able to form a complex with the minor groove of its DNA target.

The gp45 E212C mutant, which was recently constructed in the Benkovic lab, fulfilled all the abovementioned requirements and was consequently selected because of the conveniently located cysteine residues at the middle of the protein rim on the smooth side (see Figure 2C). This mutant was overexpressed in E. coli cells and purified to homogeneity according to the well-described purification methods available for the clamp. In short, the purification procedure involved three steps. After cell lysis, nucleic acids were removed by streptomycin sulfate precipitation. The protein was subsequently purified using hydroxyapatite and repeated phenyl sepharose chromatography steps (Figure 2A and 2B, respectively). The purified protein was found to have 1.4 accessible thiol functionalities per protein trimer according to the Ellman’s assay (5,5′-dithiobis[2-nitrobenzoic acid]). Thus, about half the thiols were oxidized during the final steps of the purification or were unavailable due to steric reasons.
Figure 2 - Purification of T4 clamp protein (gp45). Coomassie stained 10% SDS-PAGE gels. A. Purification with a hydroxyapatite column. A gradient of phosphate buffer was applied to elute the protein. B. Purification by phenyl sepharose chromatography (second phenyl sepharose purification step is shown). A decreasing gradient of ammonium sulfate was applied to elute the protein. Shadows in the protein band are due to a defect of the stacking gel, not due to contaminations. C. Model of the gp45 E212C mutant, showing the backbone and shape of the protein. The introduced cysteine amino acids are shown as space-filling models and indicated by arrows.

The gp45 mutant was labelled by incubation with an excess of labelling reagent at 4 °C for prolonged periods of time (typically 5 hrs). Conjugation of the methylated manganese tripyridyl porphyrin 1 with the protein resulted in 1.3-1.4 labels per protein trimer, as determined by UV spectroscopy (the porphyrin extinction coefficient in the storage buffer was determined to be $\varepsilon = 116,000 \text{ M}^{-1}\text{cm}^{-1}$) and Bradford assays. In this particular case, protein loss due to precipitation was minimized by lowering the protein concentration from 120 to 45 $\mu$M and by adding DMSO up to a concentration of 10% v/v. Furthermore, the conjugation buffer was supplemented with 100 mM potassium acetate to exchange the iodide counter ions of the porphyrin to acetate ions, and the incubation times was shortened to 3 hrs. The absorption maximum of the porphyrin Soret band in the UV-Vis was identical for the free and conjugated porphyrin, indicating that the porphyrin did not interact significantly with the protein itself. The protein was purified from unreacted labels by size exclusion chromatography (Figure 3). Comparison of the labelling efficiency with the results of the Ellman’s assay showed that all the available cysteines had fully reacted. Incubation of other maleimide functionalized porphyrins, such as the free base, zinc and manganese tritoluyl porphyrins, as well as the manganese tripyridyl porphyrin (all of which are described in Chapter 6) with the gp45 mutant did not result in the formation of any
conjugate. Probably these labels do not react because of their lack of solubility in aqueous solutions. Furthermore, the gp45 mutant was incubated with Mn-TMPyP \( 3 \) and analyzed with size exclusion chromatography. The fractions that contained protein did not show the characteristic UV-Vis absorptions of the porphyrin, so it was concluded that the porphyrin itself does not bind strongly to the protein.

![Figure 3](image-url)

*Figure 3* – Conjugation of maleimide functionalized methylated tripyridyl manganese porphyrin with the T4 clamp. A. FPLC chromatogram of the labelling reaction mixture injected onto a Superdex 200 column to separate un conjugated label from conjugated protein by size exclusion purification. B. UV-Vis spectrum of the obtained clamp-porphyrin conjugate.

The FeBabe-labelled protein was synthesized as described above using labelling reagent \( 2 \). The conjugate had 3.6 labels per protein trimer as determined by comparing the ferrozine and Bradford assays used to determine the iron and protein concentrations, respectively. The labelling efficiency was significantly higher than in the case of the porphyrin label, which could be the result of unspecific labelling by alkylation of lysines, or the use of a different conjugation method (*i.e.* maleimide vs. iodoacetamide mediated conjugation).

To test whether the attached labels influenced the activity of the clamp, an ATPase assay (see Chapter 3 and 5) was performed to ascertain that the modified clamp could be loaded onto DNA. Indeed, both labelled clamps could replace the wild-type (wt) clamp in the assay, albeit with slightly lower activity (75-85%). This is consistent with previous studies in which gp45 was labelled at different positions with fluorophores.\(^{12}\) Thus, the catalyst appended clamp proteins are recognized by the clamp loader and are loaded onto the primer-template junction of the forked DNA substrate.

Unfortunately, during preliminary oxidation experiments it became clear that a nuclease contamination was present in the labelled clamp batches. Thus, incubation of the protein alone with a DNA plasmid already resulted in degradation of the DNA substrate. Although the nuclease contamination could not be identified unambiguously, it was presumed to be endonuclease I from *E. coli*. Relatively little is known about this endonuclease, but it is the only of about 6 endonucleases in *E. coli* that degrade double stranded DNA in a non-specific manner.\(^{13}\) Endonuclease I requires Mg\(^ {2+} \) for activity,\(^ {14} \) and when Mg\(^ {2+} \) was omitted from the reaction buffer, the nuclease activity was indeed completely abolished. Unfortunately, since the clamp loader also requires Mg\(^ {2+} \) to hydrolyze ATP, experiments in which the clamp is loaded onto a nicked plasmid cannot be performed.
in Mg\textsuperscript{2+}-deficient buffers. Munn and Alberts have described the application of anion-exchange (MonoQ) chromatography to remove a nuclease contamination from their gp45 preparations\textsuperscript{15}. In order to enable the use of clamp loader to load the clamp onto DNA substrates, the labelled clamp proteins and the unlabelled E212C mutant stocks were purified using the procedure of Munn and Alberts on a SourceQ column. Interestingly, for the porphyrin-labelled clamp proteins multiple porphyrin-containing peaks were found corresponding to clamp protein trimers that were labelled with different efficiencies, e.g. with 3, 2 or 1 porphyrin per trimer. The first fraction, which contained the highest porphyrin to protein absorption ratio, was used for further experiments. For the purified FeBABE labelled clamp proteins, the resulting fractions only contained low amounts of protein. Therefore, the protein was relabelled with the purified gp45 E212C mutant. Since the previous labelling resulted in 3.6 labels per clamp trimer, the pH of the labelling solution was now lowered to 7.0 to prevent any non-specific labelling to lysines. This did not have a positive effect, since comparison of the iron concentration of the purified conjugate to the protein concentration yielded \textasciitilde4 FeBABE labels per protein trimer. Both proteins were active in the ATPase assay.

Oxidation experiments with the catalyst labelled clamps were first carried out with small DNA substrates, such as the forked DNA (see chapters 3 and 5) and the minicircle DNA substrate\textsuperscript{16}. Oxidation events in these substrates should lead to well defined oligonucleotides which can be resolved on a denaturing PAGE gel. Furthermore, both substrates contain a unique AAA site, which is advantageous in the oxidation experiment with porphyrin labelled clamp proteins. Unfortunately, it appeared that the selected substrates did not migrate as a single band on non-denaturing gels, while the denatured (i.e. separated) bands were not easily and unambiguously assigned on denaturing gels. It should be noted that the proteins used were not yet purified from the nuclease contamination, and that this could have caused the problems described above. Nevertheless, it was decided to use a pGEM derived DNA plasmid as substrate for the oxidation reactions. Plasmids are purified in their supercoiled form (form I), which provides a sensitive way to probe for oxidation events on the DNA; introduction of a single stranded break on either of the two strands of the substrate results in the uncoiling of the plasmid to the relaxed circular form (form II). This transition is easily monitored by conventional agarose electrophoresis because of the large difference in relative mobility of the two forms. This difference is due to an increase in ethidium bromide intercalation, effectively stiffening form II compared to form I. Further oxidation events, however, do not result in a difference in mobility, unless a double stranded break occurs and the plasmid is converted to the linear form (form III). This requires two oxidation events not more than 16 nucleotides apart\textsuperscript{17}. 
2.2 Oxidation of DNA plasmids by FeBABE labelled clamps and peptides

The first set of oxidation experiments focused on using the FeBABE-labelled peptide or clamp to oxidize DNA via the Fenton reaction. Following literature protocols, ascorbic acid was used as a reducing agent for Fe(III) and hydrogen peroxide as oxygen donor. It appeared, however, that under our experimental conditions the combination of ascorbic acid and hydrogen peroxide generated oxidative damage on DNA. Therefore, in the absence of any iron source, protein, peptide, or PEG, a considerable fraction of supercoiled DNA plasmid was converted into form II, which complicated the analysis of oxidation reactions with the FeBABE-labelled peptide or clamp. The origin of this background reaction was not clear, but could have arisen from minimal amounts of transition metals present in the buffer. The buffers were remade with high-purity salts but this did not lower the level of background oxidation. Therefore, the buffers were incubated with an Amberlyst IRC-86 ion-exchange resin prior to the oxidation experiments to remove any contaminating heavy metal salts. The background oxidation persisted for a phosphate buffer but not for a Tris-HOAc buffer (without Mg²⁺-salts). This prompted us to screen different reductants (ascorbic acid, sodium ascorbate, DTT, tris(2-carboxyethyl)phosphine [TCEP], β-mercaptoethanol) for their ability to promote DNA damage in the absence of iron salts. DTT was identified as a reducing agent that was able to oxidize DNA with hydrogen peroxide in the presence but not in the absence of Fe-EDTA. The presence of PEG as macromolecular crowding reagent also caused a background oxidation reaction. Again, incubation with Amberlyst IRC-86 ion-exchange resin was not effective in inhibiting this reaction. Thus, the experiments described in this section were always carried out with DTT rather than ascorbic acid as a reductant, and in the absence of macromolecular crowding reagent.

To test the efficiency of the padlock approach (see Chapter 2) compared to the direct labelling approach, a supercoiled DNA plasmid was incubated with different concentrations of *wt* clamp and a fixed amount of the FeBABE-peptide conjugate (10 μM). The footprinting reaction was started by the addition of DTT as reducing agent and hydrogen peroxide as oxygen donor. It was expected that higher clamp concentrations would recruit more peptide to the clamp-DNA complex, hence promoting more oxidative cleavage (Figure 4). Even after a 15 minute reaction period this effect was, however, not pronounced compared to the oxidation by labelled peptide in the absence of any clamp protein. In all cases, ~71% of the total amount of DNA remained supercoiled (form I). Thus, although ~29% of the starting material indeed was oxidized, this could be attributed to oxidation by the Fe-labelled peptide in solution rather than bound to clamp proteins on DNA. The experiment was repeated with higher DTT and hydrogen peroxide concentrations, but again no pronounced effect of the clamp concentration on the oxidation efficiency could be identified (data not shown).

The direct labelling approach was first tested with different concentrations FeBABE labelled clamp protein (0 - 2.5 μM) and a supercoiled DNA plasmid (Figure 5). Although the amount of plasmid oxidized depended directly on the concentration of the protein, 85% of
Figure 4 – Oxidation of supercoiled DNA plasmid by wt gp45 and FeBABE labelled peptide. A. Schematic representation of the non-specific binding process, followed by activation of FeBABE, resulting in oxidative DNA damage. B. Ethidium bromide stained 0.8% agarose gel. The plasmid (200 ng) was incubated with the indicated amount of gp45 and 10 μM Fe-labelled peptide. The reaction was activated by addition of DTT and H2O2 and quenched after 15 mins. 0 = supercoiled plasmid, BL = background oxidation of the peptide when no protein is present. C. Conversion of supercoiled DNA to nicked DNA versus protein concentration quantified by measuring the intensity of the bands.

The DNA was still supercoiled with the highest gp45 concentration tested (total iron concentration 10 μM). An increase in hydrogen peroxide concentration resulted in more oxidative DNA damage, but also an increased background oxidation.

The experiments described above indicate that it is necessary to confirm the loading of the clamp proteins on DNA. Therefore, the clamp loader complex was used to load the clamp onto DNA. Not only could this improve the binding of clamp proteins on DNA, but also the interaction of the peptide with the clamp. Recently, it was reported that although gp45 can bind to a forked DNA substrate linked to a surface without the clamp loader, the polymerase cannot interact with such a non-specifically bound gp45.19 Since the main interactions between the clamp and the polymerase take place via the C-terminus of the polymerase and the subunit interface of the clamp, it is questionable whether the peptide can bind to non-specifically bound clamp proteins, at least in the subunit interface, which is proposed to lead to a processive gp45 species. The clamp loader, however, is able to correct
Figure 5 – Oxidation of supercoiled DNA plasmid by FeBABE labelled gp45. A. Schematic representation of the non-specific binding process, followed by activation of FeBABE metal, resulting in oxidative DNA damage. B. Ethidium bromide stained 0.8% agarose gel. The plasmid (200 ng) was incubated with the indicated amount of FeBABE labelled gp45. The reaction was activated by addition of DTT and H₂O₂ and quenched after 1 or 5 mins. Mw = molecular weight marker, 0 = supercoiled plasmid, C. Conversion of supercoiled DNA to nicked DNA versus protein concentration quantified by measuring the intensity of the bands.

the conformation of the clamp and restores the ability of the clamp to interact with the polymerase, enabling the formation of an active holoenzyme.

To ensure that the clamp protein is loaded onto DNA, wt gp45 was incubated with the clamp loader in the presence the labelled peptide and a nicked DNA plasmid. The single stranded break in the nicked plasmid (form II) is necessary to provide a recognition site for the clamp loader. Consequently, one cannot monitor the transition of form I to form II in this
A.

**Figure 6** – DNA oxidation experiment with FeBABE labelled peptide, *wt* gp45 and gp44/62. A. Ethidium bromide stained 0.8% agarose gel. Nicked DNA plasmid (200 ng) was incubated with 250 nM gp45, 125 nM gp44/62 and 10 μM Fe-labelled peptide. FeBABE was activated with 10 mM DTT and 0.1% H₂O₂, and quenched after 5 mins. Mw = molecular weight marker, 0 = supercoiled plasmid, BL = blank reaction, *i.e.* same reaction conditions without gp45 and gp44/62.

case, but progressive amounts of DNA oxidation should result in a linearized plasmid (form III). Unfortunately, no linearized DNA was generated during this experiment (Figure 6), even though a higher concentration of DTT and H₂O₂ were used (10 mM and 0.1% vs. 1 mM and 0.03% for DTT and H₂O₂, respectively). Therefore it was concluded that under the experimental conditions used, the padlock mechanism is not very efficient, even in the presence of the clamp loader.

The same experiment was subsequently performed with the FeBABE labelled clamp protein. Again, a nicked DNA plasmid was used as substrate and incubated with the protein and the clamp loader (Figure 7). In separate experiments, the reaction mixture was supplemented with the unlabelled peptide or with DNA polymerase gp43, to investigate whether more stable complexes would be formed. Additionally, reference experiments were performed with gp45 without clamp loader, and gp45 with clamp loader using a supercoiled plasmid. In all the cases, no DNA oxidation could be observed.
It can only be concluded that the FeBABE-labelled peptide and protein display a disappointing activity under the experimental conditions used. The conjugation of the iron centre to a DNA binding protein should increase the local concentration of hydroxyl radicals resulting in lower overall iron concentrations needed to obtain the same amount of DNA oxidation. If the protein is loaded by non-specific binding or by the clamp loader protein, this binding is not strong enough to oxidise the DNA efficiently. The fact that no oxidation of the plasmid takes place indicates that either the protein does not bind efficiently enough to DNA, or that the distance between the iron centre and DNA is still too large when the clamp is bound on the DNA. The calculated distance of 17 Å between the iron centre of the bound clamp and the DNA is well within the reported maximum range of hydroxyl radical damage around the iron centre (22 Å). Another possibility is that the generated hydroxyl radicals are quenched by the presence of the protein. Indeed, inhibition of radical damage on DNA is possible when the Fe-EDTA complex cleaves peptide bonds instead of generating a hydroxyl
radical; this however would require the formation of covalent contacts between the iron centre and the protein, limiting the range of this process to 12 Å.

Blanc experiments with (unconjugated) Fe(II)-EDTA complexes were effective in the oxidation of DNA substrates, although higher concentrations of catalyst were needed (typically 0.1-1 mM). Unfortunately, these concentrations are out of range for the proteins and DNA substrates involved. In DNA footprinting, the experimental conditions are aimed at generating a single break in only ~30% of the total DNA molecules, to prevent double stranded breaks to be generated. The conditions are therefore not optimized for generating as much DNA cleavage as possible, as in our case. The concentration of the oxidant can be increased, but this also results in a higher background oxidation rate.

2.3 AFM as an alternative tool for the analysis of porphyrin oxidation experiments

In contrast to the Fe(II)-EDTA mediated DNA cleavage reactions, oxidation of DNA by porphyrins occurs selectively. As discussed in Chapter 2, only AAA sites on dsDNA substrates are oxidized by cationic manganese porphyrins, which results in the formation of nicks and aldehydes. The conversion of a supercoiled DNA plasmid to a nicked DNA by an oxidation reaction is easily visualized with gel electrophoresis. Further oxidation reactions, however, are masked unless the amount of oxidative damage on the substrate results in a double stranded break, as linear fragments can be resolved from the nicked circular plasmids by gel electrophoresis. Thus, the amount of oxidation reactions per plasmid could strongly be underestimated if only nicked DNA is formed, or nicked DNA is used as starting material, e.g. in experiments with the clamp loader protein. This effect could especially be dramatic for the porphyrin mediated DNA oxidations reactions, as the oxidation sites may be spaced on the plasmid or adjacent oxidation sites may be on the same strand; in both cases their oxidation does not result in linearized plasmids.

To improve the sensitivity of the oxidation assay, we looked into alternative ways of analyzing the reactions. Based on an already existing assay, we chose to take advantage of the aldehyde groups on the DNA that result from porphyrin oxidation reactions. The aldehydes in porphyrin oxidized DNA oligonucleotides were shown to be available for chemical modification by the Meunier group, for instance for reductive amination reactions. Furthermore, a commercial assay for the quantification of the amount of abasic sites (baseless deoxyribose sugars) in DNA relies on the functionalization of the anomeric form of the sugar in DNA. The anomeric form contains an aldehyde, which is reacted with a small hydroxylamine-functionalized biotin derivative, called aldehyde reactive probe (ARP). The final amount of abasic sites can be analyzed via incubation with horseradish peroxide-functionalized streptavidin using standard enzyme-linked immunosorbent assay (ELISA) protocols. We were interested to further develop this assay for use in the oxidation reactions described in this chapter. In the fast body of literature describing DNA oxidation reactions by porphyrins, the aldehyde has not been used before for the quantitation of oxidation damage.
The ARP reagent was synthesized as described in Scheme 1, based on a literature procedure. First, the hydroxylamine group of the commercially available carboxymethylhydroxylamine hemihydrochloride was protected with a Boc-group to yield 4. The carboxylic acid of 4 was activated by the formation of the N-hydroxy succinimide (NHS) activated ester 5. The ester 5 was subsequently reacted with biotin hydrazide, resulting in Boc-protected biotin compound 6. Finally, the Boc-group of 6 was deprotected with TFA to yield the ARP reagent 7. The total yield of these four steps was 30%, slightly higher than that obtained for the literature procedure (26%).

![Scheme 1](image)

Scheme 1 – Synthesis of the ARP reagent. (i) di-t-butyl-dicarbonate, DIPEA, dioxane, 16 hrs, 78%, (ii) NHS, DCC, acetonitrile, 16 hrs, 65%, (iii) biotin hydrazide, DMF, 16 hrs, 88%, (iv) TFA, 0 °C, 1.5 hrs, 68%.

As a test, the ARP reagent 7 was first incubated with a MnTMPyP 3-oxidized pGEM plasmid for 1.5 hrs, followed by the addition of sodium cyanoborohydride to reduce the formed imine. The ARP-labelled plasmid was then incubated with streptavidin and analyzed by AFM (Figure 8). This showed that the majority of DNA strands were circular, thus not supercoiled, indicating that oxidation had taken place. Furthermore, an average of 5 globular features with the height corresponding to streptavidin was present per DNA strand (Figure 8B). This experiment demonstrated that the ARP assay indeed provides an alternative tool for analysis of the oxidation reactions. For instance, if all the streptavidin proteins are clustered near a particular region on the DNA, this would provide proof for the processivity of the oxidation reaction in the case of the porphyrin functionalized gp45. Such information would be difficult to obtain by other experimental techniques. In the case of the FeBABE-functionalized gp45 proteins, however, the processivity would be apparent by the rapid appearance of linear DNA strands rather than nicked DNA. To further investigate the streptavidin-labelling analysis procedure, it was performed as an ELISA-like assay in a microtiter plate according to the literature procedure. Unfortunately, no results were obtained due to background absorption of the ARP reagent to the microtiter plate.
Figure 8 – AFM analysis of an oxidized DNA plasmid treated with ARP and streptavidin. A. Schematic representation of the procedure. First, the plasmid is oxidized, either by loading of the labelled clamp onto DNA by the clamp loader or by non-specific binding of the clamp to DNA. Oxidation results in the generation of aldehyde functions in the sugar of the nucleotide, which can be targeted by the ARP reagent. Streptavidin is subsequently added to capture the biotin probes and facilitate the quantification of the number of oxidation events. B. Typical AFM height image showing two relaxed circular DNA plasmids with a number of globular features. C. Profile analysis (following the indicated line in the top picture) of the picture shown in B. Different features are indicated with * for a naked DNA strand (typical height 1.2 Å), and # for a DNA strand with a globular feature (3.3-5.4 Å). The apparent height of a streptavidin protein in AFM analysis is ~ 4.4 Å.
2.4 Oxidation of DNA plasmids by porphyrin labelled clamps

The ability of unconjugated catalyst 1 to oxidize the supercoiled DNA plasmid was assessed by comparison with reference catalyst Mn-TMPyP 3. Although significant less efficient than 3, 1 was indeed able to convert form I DNA into form II. The diminished reactivity could partially be due to the lower amount of positively charged substituents in 3. The crystal structure of a methylated nickel porphyrin bound to DNA is described in the literature and suggests that only two of the four methylated pyridine side groups interact with the phosphate backbone. Therefore the lower reactivity of 1 compared to 3 could arise from its lower overall charge, or the lower number of possibilities that 1 can interact with its binding site on DNA. Nevertheless, the reactivity of 1 is still sufficient to study its DNA oxidation properties when conjugated to the clamp.

The supercoiled pGEM plasmid was incubated with the clamp labelled with porphyrin 1 in complex buffer and activated with KHSO₅ as the oxygen donor. The experiment was also done in the presence of (unlabelled) peptide and in the presence of PEG as macromolecular crowding agent (Figure 9). The labelled clamp was able to convert nearly all of the supercoiled DNA into the nicked form. Interestingly, the addition of peptide or PEG resulted in a decrease in oxidation efficiency compared to the experiments without peptide or PEG. It could be that the binding of the clamp to DNA is inhibited by the binding of the peptide to the open subunit interface of a clamp trimer. PEG could lower the oxidation efficiency because its presence increases the viscosity of the solution, resulting in a lower diffusion rate of the oxidant. Prior to analysis, all the samples had to be extracted with phenol – chloroform solution to remove clamp proteins, as the porphyrin-labelled clamp gave a band shift on the gel. This can only occur when there is a sufficiently high binding affinity of the clamp for DNA. Since this band shift was not observed for FeBABE-labelled clamps, it provides further evidence that the porphyrin is still able to bind to DNA when it is conjugated to the clamp.

Quantification of the bands on the agarose gel shows that the 90% and 75% of the plasmids are converted into nicked plasmids for the reaction without PEG or peptide. For the reaction with PEG, this value drops to 60% and 40%, respectively. The streptavidin labelling analysis (Figure 10) shows that without PEG and peptide, the average number of reactions per plasmid is 1.8 ± 1.7. The histogram appears to be constituted of two Gaussian curves, with a maximum at 1 and 5 streptavidin molecules per plasmid. When analyzing the plasmids with an average of 5 streptavidins, the dots appear to be in clusters (Figure 10C, arrow). If only a distributive oxidation would take place, these dots are expected to be evenly distributed over the substrate. Therefore, the clusters could be the result of processive oxidation of the Mn-TMPyP-gp45 conjugate on that particular template. If so, the amount of processive oxidation is relatively small; around 17%. Interestingly, while gel electrophoresis shows that all the supercoiled plasmid is converted into nicked plasmid, there is a significant amount of plasmids that are not labelled at all. AFM analysis shows that most plasmids (59%) are indeed circular, and therefore have to be oxidized, so the current procedure of
Figure 9 – DNA oxidation experiments with porphyrin labelled clamp protein. A. Schematic representation of the oxidation reaction. B. Experiments in which the labelled clamp was incubated with supercoiled DNA plasmids (200 ng), activated with KHSO₅ (5 μM), and quenched at the indicated times (in minutes). Experiments were also performed in the presence of 10 wt% PEG (Mₙ = 8,000) (right column) and with peptide (10 μM) (bottom row). MW = molecular weight marker, 0 = supercoiled plasmid. C. Calculation of the amount of single stranded breaks based on the intensity of the corresponding bands on gel electrophoresis. Lines are exponential fits. D. Calculated amounts of single stranded breaks for the points starting from 2.5 minutes.
Oxidation of DNA substrates by clamp-based catalysts

A.

**Oxidation of Supercoiled Plasmid with MnTMPyP-gp45**

<table>
<thead>
<tr>
<th>ratio SAv per plasmid:</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>gp45 (n=23)</td>
<td>21</td>
<td>7</td>
<td>30</td>
<td>4</td>
<td>26</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>gp45 + peptide (n=11)</td>
<td>45</td>
<td>5</td>
<td>9</td>
<td>1</td>
<td>18</td>
<td>2</td>
<td>27</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Figure 10** – Streptavidin labelling analysis of DNA oxidation experiments with porphyrin labelled clamp protein and supercoiled plasmid. A. Quantification of the amounts of oxidation reactions based on the streptavidin labelling procedure. The ratio of streptavidin molecules found per plasmid is presented on the horizontal axis, while the relative score of the number of oxidation events is on the vertical axis. N is the number of plasmid molecules investigated. B. AFM image of the streptavidin-labelled plasmids for the reaction without peptide. The arrow indicates a region of multiple streptavidin proteins in a row, and may be an indication of processive oxidation (see text). C. Zoomed area of the image in B. containing the cluster of streptavidin proteins. C. AFM image of a typical streptavidin-labelled plasmid for the reaction with peptide.

Streptavidin labelling either does not pick up every oxidation event, or alternative oxidation pathways exist that do not result in the generation of aldehydes. Only a small amount of the investigated plasmids appear to be supercoiled (5%) when visualized by AFM, while the remaining 36% is linear. As these linear plasmids are also not seen on gel electrophoresis, they could also be the result of breaking of oxidized plasmids on the mica substrate or sample treatment (e.g. freeze-thaw cycle). In some cases, the plasmids also seem to be ‘zipped’ in certain regions (Figure 10B, top left corner). It is not known what would cause
this phenomenon, but it could arise from cross-linking between biotin-labelled aldehydes by streptavidin proteins.

For the reaction with peptide, the average number of streptavidin per plasmid is $1.3 \pm 1.3$. This is again lower than the calculated number of single stranded breaks ($S$) of $7$. The histogram does not appear to consist of Gaussian distributions, and no plasmids were seen with more than 3 streptavidins. In this case, analysis was more difficult because the DNA concentration was lower than in the case of the reaction without peptide. No supercoiled plasmids were observed, and the major part was circular (80%), while only a minor part was linear (20%). The reactions with PEG could unfortunately not be labelled and analyzed with the current procedures, due to a high background of proteins in the sample, even after gel filtration.

The oxone concentration in the previous experiment was 5 μM, while the protein concentration was 400 nM. Since each clamp trimer has three labels, this implies an oxidant to porphyrin ratio of 4:1. Under these conditions, a truly processive oxidation mechanism is hard to accomplish. Therefore, a range of oxone concentrations was tested (Figure 11). Indeed, a higher oxone concentration yields more linear plasmid, which can only occur when two oxidation events take place near each other on opposite strands. High oxone concentrations (> 20 μM), however, lead to loss of overall DNA intensity on the gel and/or extensive smearing of the bands, which makes the streptavidin labelling analysis complex and inaccurate.

![Figure 11](image)

**Figure 11** – DNA oxidation experiments with porphyrin labelled clamp protein and a supercoiled DNA plasmid (200 ng), in the presence of different amounts of KHSO$_5$ (1-20 μM). The reactions were quenched at the indicated times. Mw = molecular weight marker, 0 = supercoiled plasmid.

Finally, the clamp loader complex was used to load the clamp onto nicked DNA substrates. Because of the specificity of the porphyrin mediated oxidation reaction, the place at which the clamp is loaded strongly influences the number of oxidation events. Loading of the clamp at a region rich in AAA sequences will result in more oxidation events compared to an experiment in which the clamp is loaded at a region without any suitable oxidation sites.$^{24}$ Furthermore, the loading process also ensures that the clamp can only slide in one direction, because one side is physically blocked by the clamp loader. To test this, the pGEM T4 plasmid was nicked with the Nt.BbvCI and Nb.BbvCI endonucleases. These two enzymes
originate from the same restriction endonuclease (BbvCI) but have been specifically engineered to cut only once in either of the two strands of the BbvCI recognition site. The plasmid was digested with Nt.BbvCI and Nb.BbvCI in separate reactions, yielding two nicked plasmids in which the nicks are located near each other, but are introduced in different DNA strands. Since the clamp loader binds to the 3’ end of the nick and remains bound when the clamp is loaded, the clamp is either guided towards a region with a lot of potential oxidation sites, or a region with few oxidation sites (Figure 12), depending on which nicked plasmid is used.

Figure 12 – Schematic representation showing a segment of the pGEM T4 plasmid (nucleotide 1635 to 2113 from the origin of replication). The potential oxidation sites are indicated with black bars at the bottom. The arrows show the direction in which the clamp can diffuse when the plasmid is nicked with the indicated enzyme.

The results of the oxidation experiments with Nt.BbvCI and Nb.BbvCI nicked DNA are shown in Figure 13B and 13C, respectively. The clamp was loaded onto DNA by the clamp loader, and additional reaction solutions were supplemented with either the peptide (10 μM) or the DNA polymerase gp43. In the latter case, a relatively stable and stalled complex of gp45 and gp43 should be formed at the loading site that inhibits any oxidation unless the catalyst is able to reach a potential oxidation site from this stalled complex. This may occur in the case of the Nb.BbvCI-nicked substrate, but unfortunately the oxidation site was on the same strand as the nick and such an event cannot be detected by gel electrophoresis. A blank experiment without clamp loader protein was also performed to determine the efficiency of the clamp loader and unspecific binding approaches. Based on the intensities of the corresponding bands on the gel, in both cases only minute amounts of nicked DNA (< 2% and < 4% for the Nt.BbvCI and Nb.BbvCI-nicked substrates, respectively) were converted into linear DNA (Figure 13B, C). The highest amount of linear plasmid was formed when the nicked plasmid was incubated with the labelled clamp alone. In the absence of the clamp loader, 20% and 9% are oxidized for the Nt.BbvCI and Nb.BbvCI-nicked substrates, respectively. Thus, the clamp loading process by the clamp loader does not seem to be very efficient. Indeed, a nick is not a very efficient clamp loading site when compared to a 5’ extended single overhang. The lower efficiency, however, is probably better explained by the ~35 fold excess of clamp protein over loading sites. While gp45 in the absence of gp44/62 is apparently able to bind to the substrate and convert it into linear DNA, the presence of gp44/62 may actually slow down the catalytic turnover due to several factors. First, gp45 is more restricted in the presence of gp44/62, because the two proteins form a complex and gp45 cannot be loaded onto DNA before gp44/62 has found a loading
Figure 13 – DNA oxidation experiments with porphyrin labelled clamp protein on nicked substrates. A. Schematic representation of the oxidation reaction. B. Experiments, in which Nt.BbvCI-nicked DNA plasmid (200 ng) was incubated with the labelled clamp (250 nM), activated with KHSO₅ (20 μM) and quenched at the indicated times. Experiments were done with clamp loader (125 nM), clamp loader and peptide (10 μM), clamp loader and gp43 (250 nM) or 45 without clamp loader. C. Same experiment as in B., but with Nb.BbvCl-nicked DNA as substrate.

site. In contrast, gp45 alone is able to find an oxidation site by itself via diffusion. Second, the compared to a 5' extended single overhang. The lower efficiency, however, is probably better explained by the ~35 fold excess of clamp protein over loading sites. While gp45 in the absence of gp44/62 is apparently able to bind to the substrate and convert it into linear DNA, the presence of gp44/62 may actually slow down the catalytic turnover due to several factors. First, gp45 is more restricted in the presence of gp44/62, because the two proteins
A.

**Oxidation of Nicked Plasmid with MnTMPyP-gp45**

![Graph](image)

<table>
<thead>
<tr>
<th>ratio SAv per plasmid</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nt.BbvC (n=13)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7.7</td>
<td>7.7</td>
<td>16.4</td>
<td>30.8</td>
<td>16.4</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Nb.BbvC (n=15)</td>
<td>6.7</td>
<td>33.3</td>
<td>33.3</td>
<td>13.3</td>
<td>13.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>NtBbvC, no gp44/62 (n=6)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>33.3</td>
<td>0</td>
<td>16.7</td>
<td>16.7</td>
<td>0</td>
<td>16.7</td>
<td>16.7</td>
<td></td>
</tr>
</tbody>
</table>

B.

**Figure 14** - Streptavidin labelling analysis of the oxidation experiments on nicked plasmids. A. Quantification of the number of oxidation events based on the streptavidin labelling procedure for the Nt.BbvC and Nb.BbvC-nicked plasmids in the presence of clamp loader, and of the Nt.BbvC-nicked plasmid in the absence of clamp loader. The ratio of streptavidin molecules found per plasmid is presented on the horizontal axis, while the relative score per of the number of oxidation events is on the vertical axis. N is the number of plasmid molecules investigated. B. AFM image of the streptavidin-labelled plasmids for the reaction of Nt.BbvC-nicked DNA. The arrows indicate clusters of streptavidin molecules in a row. C. AFM image of the streptavidin-labelled plasmids of the Nb.BbvC-nicked substrate. The image is characterized by the distinct absence of any streptavidin clusters. D. and E. Zoomed images of B. and C., respectively.
form a complex and gp45 cannot be loaded onto DNA before gp44/62 has found a loading site. In contrast, gp45 alone is able to find an oxidation site by itself via diffusion. Second, the porphyrin on gp45 may bind to an oxidation site while gp45 is not on DNA. In the case of gp44/62 loaded gp45, it is hard to say whether the porphyrin can easily reach the oxidation site. Because of the excess of gp45 over nicked DNA, several gp45 proteins may be loaded in tandem, creating gp45 arrays that effectively cover the DNA and prevent oxidation reactions to take place. In this particular case, it is hard to find a more optimal concentration ratio, since lowering the amount of gp45 results in dissociation of the gp45 trimer, and increasing the plasmid concentration to equimolar amounts is not feasible due to solubility constraints.

In contrast to gel electrophoresis, the streptavidin labelling analysis was found to be successful in showing oxidation reactions on the nicked substrates (Figure 14). When the Nb.BbvCI-nicked substrate was used, the average number of streptavidin molecules per plasmid was 6.1 ± 1.6, while for the Nt.BbvCI substrate this number was significantly lower, 1.9 ± 1.2. This nicely correlates with the situation outlined in Figure 12, where the clamp is loaded on the plasmid and migrates in the direction of a high amount or low amount of oxidation sites (Nb.BbvCI and Nt.BbvCI-nicked substrate, respectively). Moreover, with the Nb.BbvCI-nicked substrate there appeared to be a consistent number of streptavidin clusters, which could very well be the result of oxidation events on the first set of oxidation sites from the loading site (Figure 12). This implies that the clamp has travelled more than 50 nt on the substrate before arriving to this area.

The electrophoresis gels show that the experiments without clamp loader yield a higher number of oxidation events per plasmid than the experiments with clamp loader (Figure 13B, C). This also results in a higher amount of streptavidin molecules found on the plasmids (6.7 ± 2.5). The distribution over the plasmids does not appear to be Gaussian, which could be due to the small number of molecules analyzed. Nevertheless, in all molecules analyzed that were oxidized in the absence of clamp loader, no clusters of more than 2 streptavidin molecules near each other could be found.

The morphology of the streptavidin-labelled DNA molecules was also investigated by AFM. In the case of the Nt.BbvCI-nicked sample, 33% of the plasmids were linear, while the remaining plasmids were circular. For the Nb.BbvCI-nicked substrate, 30% of the plasmids were linear, 40% circular and 30% appeared to be supercoiled. Although the latter plasmid did contain a very small amount of supercoiled plasmid (≤ 1%, barely visible in figure 13C, lanes 0) due to a fraction of the supercoiled starting material that was resistant to Nb.BbvCI digestion, the fraction of apparent supercoiled plasmids found with AFM analysis is much higher, and is better explained by cross-linking reactions.

3. Conclusion

A novel mutant (E212C) of the clamp protein (gp45) of the T4 bacteriophage was expressed and labelled with two different chemical oxidation catalysts to yield artificial enzymes that potentially can oxidize linear substrates such as DNA or double bond containing polymers in
a processive fashion. The proteins needed to be purified with ion-exchange chromatography in order to remove a nuclease contamination. The interaction of the labelled proteins with DNA was studied with the help of an ATPase assay and was found to be 75-85% as efficient to that of the wild-type protein, which shows that the labels do not significantly interfere with the clamp loading process.

Unfortunately, the FeBABE-labelled clamp protein and peptide displayed little or no activity on DNA plasmid substrates under the conditions used. The experiments with Fe-EDTA complexes were complicated by the occurrence of a background reaction arising from the presence of ascorbic acid and hydrogen peroxide. The experiments were therefore performed with DTT as the reductant. DNA plasmids were oxidized by the Fe-BABE labelled peptide in the presence of various concentrations of wt clamp protein. Although some of the plasmid was oxidized, the oxidation could not be shown to arise from the peptide binding to the clamp on DNA. Additionally, the FeBABE-labelled clamp was incubated at various concentrations with DNA plasmids. The amount of oxidation only showed a marginal dependence on the concentration of the clamp. The reasons for the lack of oxidation activity are unclear at the moment, and require further studies. Possibly, it can be attributed to suboptimal binding conditions for the protein on DNA, an inappropriate distance between the iron centre and the substrate, or an intrinsic low rate of radical formation in the Fenton reaction.

The MnTMPyP-labelled clamp proteins were found to oxidize DNA plasmids. When supercoiled plasmids were used as substrate, gel electrophoresis showed that the plasmids were converted into the nicked form (form II). This conversion was most efficient in the case were no PEG or peptide was added. So, although both these reagents were originally thought to enhance the amount of clamp protein bound to plasmid, these experiments now show that this is not the case, at least with the currently applied experimental conditions.

To investigate the porphyrin-mediated oxidation reactions in more detail, a novel streptavidin labelling procedure was introduced based on the commercially available ARP quantification assay. The oxidation of DNA by manganese porphyrins results in strand scissions, and aldehyde functionalities at the scission sites. Oxidized DNA plasmids were incubated with the ARP reagent, and as a result the aldehyde functions on DNA that were generated by the oxidation events were labelled with biotin. After purification by ethanol precipitation, the plasmids were incubated with streptavidin, purified by size exclusion gel chromatography and analyzed by AFM. A feasibility experiment showed that for MnTMPyP oxidized DNA, characteristic spots appeared on DNA that could be attributed to streptavidin proteins. The same results were obtained when the MnTMPyP-labelled clamp proteins were used.

The oxidized substrates for the experiments without PEG present were further analyzed with the streptavidin labelling procedure. Whereas the experiment without peptide gave two Gaussian distributions centred around 1 and 5 streptavidin molecules per plasmid, the experiment with peptide did not display a Gaussian distribution. In the former case, the oxidation is thought to proceed via a distributive and processive mechanism. Careful analysis of the resulting streptavidin-labelled DNA plasmids indeed showed that the higher
number of streptavidin proteins per plasmid are correlated with streptavidin clusters on a distinct part of the substrate.

The porphyrin-labelled clamp was also loaded onto nicked DNA substrates by the clamp loader protein. In this case, two different substrates were used that were nicked by the Nt.BbvCI and Nb.BbvCI endonucleases, which introduce a single stranded break in opposite strands but at the same recognition site. Since the clamp loader physically blocks the 3' direction, the loaded clamp is forced to move in the opposite directions. Using this approach, the clamp was directed towards a region rich in oxidation sites, and towards a region low in oxidation sites. Although gel electrophoresis was unable to show the occurrence of oxidation events for both samples as no linear plasmids were generated during the oxidation, the streptavidin labelling procedure did show distinct differences between the two samples. In the case of Nt.BbvCI-nicked DNA, the clamp is loaded in the direction of a region rich in oxidation sites. Indeed, AFM analysis showed that there are relatively high numbers of streptavidin molecules per plasmid (6.1 ± 1.6). Moreover, the streptavidin proteins were often found in clusters that only occurred once per plasmid. For the Nb.BbvCI-nicked substrate, a significantly lower number of streptavidin proteins per plasmid were found (1.9 ± 1.2), and the proteins were not present as clusters. Interestingly, in the experiment in which the clamp loader was omitted, the streptavidin labelling procedure showed that the amount of streptavidin proteins per plasmid is similar (6.1 ± 1.6 to 6.7 ± 2.5 for the reaction with and without clamp loader, respectively), while gel electrophoresis showed that ~20% of the molecules are linearized. Without clamp loader, no streptavidin clusters were observed on the substrates.

In conclusion, this chapter describes a novel concept for efficient catalytic conversions on DNA substrates. The followed procedure has a great potential as a molecular biology tool or as a new methodology in the field of DNA degrading medicines. In Chapter 2, two possible approaches were proposed for performing catalysis on DNA substrates. The first, i.e. the padlock approach, uses a labelled peptide and an unlabelled clamp, while the second one, i.e. the direct labelling approach, makes use of clamps that are labelled with the catalysts. For the FeBABE-labelled peptide and clamp, little activity was found for DNA-bound species, irrespective whether the clamp was loaded onto DNA by non-specific binding or by the clamp loader. The porphyrin-labelled clamp, in contrast, did show activity as an oxidation catalyst for DNA. With a novel streptavidin labelling procedure, which is able to identify oxidized sites on DNA on the molecular level, remarkable features of the oxidation process were identified. Clamp binding by non-specific binding yields mostly evenly distributed oxidized sites and a small amount of sites that have apparently been oxidized by a processive oxidation event (in the case without peptide or PEG). Nicked substrates however yielded a far more distinct oxidation pattern. If the clamp is loaded into the direction of a small number of potential oxidation sites, the cleavage pattern is more or less similar to that found for non-specifically loaded supercoiled templates. In contrast, if the clamp is loaded into the direction of a large number of potential oxidation sites, most plasmids display a region coated with an array of streptavidin proteins, which could be an indication that the oxidation reaction proceeds in a processive manner.
4. Experimental section

Materials

p-Bromoacetamidobenzyl-EDTA (BABE) was obtained from Toronto Research Chemicals Inc. and complexed with Fe(III) according to a literature procedure.\(^{26}\) DMF was biotech grade (Sigma-Aldrich). Reference catalyst Mn-TMPyP (3), dye content ~85%, was from Sigma. Reacti-Bind solution was from Pierce. Amberlyst IRC86 (hydrogen form) was from Fluka. Wt gp45 and gp44/62 were a gift from Prof. S. Benkovic. All other materials were obtained from standard commercial sources. AFM experiments were performed on Veeco Nanoscope III or IV AFM multimode microscopes operating in tapping mode at room temperature. Unless stated otherwise, samples were incubated for one minute on freshly cleaved and unmodified Muskovita mica, washed with 1-2 ml water (MilliQ grade) and blown dry with a nitrogen flow. Images (512 x 512 lines) were recorded at scanning speeds of ~1 Hz. Image analysis was done with the WSxM software, version 4.0 Develop 10.4.\(^{27}\) The images shown in the figures may have been zoomed. For experimental details regarding the synthetic procedures, the reader is referred to section 6.4.

Synthesis:

N-((t-Butyloxycarbonyl)-O-(carboxymethyl)hydroxylamine (4) was synthesized according to a slightly modified literature procedure.\(^{22}\) Carboxymethylhydroxylamine hemihydrate (300 mg, 2.75 mmol) was dissolved in 10 ml of dioxane followed by addition of DIPEA (0.60 ml, 3.44 mmol). After the hydroxylamine was dissolved, di-t-butyl-dicarbonate (659 mg, 3.02 mmol) was added and the solution stirred for 16 hrs under a nitrogen atmosphere. Then, 50 ml water and a few drops of 1M aqueous NaOH was added to the reaction mixture, which was subsequently extracted with 3x50 ml ethyl acetate. The aqueous phase was acidified to pH = 3 by addition of 0.5M aqueous KHSO\(_4\) and extracted with 3x50 ml ethyl acetate. The combined organic layers were washed with 0.5M KHSO\(_4\) and brine, dried over Na\(_2\)SO\(_4\) and concentrated. The product was crystallized from ethyl acetate in heptane 1:1 v/v. Yield 410 mg (78%) of a white crystalline compound:

mp 110.7°C;  
CI-MS calcd. for C\(_7\)H\(_{15}\)NO\(_5\): 191.0794, found: 192.0876 (+H\(^+\));  
\(^1\)H-NMR (400MHz, DMSO) \(\delta\) 12.76 (br s, 1H, COOH), 10.07 (br s, 1H, NH), 4.23 (s, 2H, CH\(_2\)), 1.37 (s, 9H, CH\(_3\));  
\(^{13}\)H-NMR (75MHz, DMSO) \(\delta\) 170.12, 156.40, 80.02, 71.99, 27.96.

N-Hydroxysuccinimide ester of N-((t-Butyloxycarbonyl)-O-(carboxymethyl)hydroxylamine (5). This compound was synthesized following a procedure from Ide et al.\(^{22}\) Instead of purification by crystallisation, the product was purified using column chromatography (silica, eluent acetonitrile in dichloromethane, 1:1 v/v) to yield 65% of a white crystalline compound, which was used directly in the next reaction:

ESI-MS calcd. for C\(_{11}\)H\(_{16}\)N\(_2\)O\(_7\): 288.0958, found: 311.0862 (+Na\(^+\));  
\(^1\)H-NMR (400MHz, CDCl\(_3\)) \(\delta\) 7.84 (s, 1H, NH), 4.78 (s, 2H, CH\(_2\)), 2.87 (s, 4H, CH\(_2\)), 1.49 (s, 9H, CH\(_3\));  
\(^{13}\)H-NMR (75MHz, CDCl\(_3\)) \(\delta\) 168.23, 154.55, 155.81, 82.20, 70.35, 27.65, 25.10.

Boc-protected ARP (6)

The NHS ester 5 (43 mg, 0.15 mmol) was dissolved in dry DMF, and biotin hydrazide (32 mg, 0.12 mmol) and one drop of DIPEA were added. The reaction mixture was stirred for 16 hrs under a nitrogen atmosphere, concentrated and the product was isolated by column chromatography (silica, eluent: 10% to 30% v/v MeOH in EtOAc). Yield 47.2 mg (88%) of an off-white solid:

ESI-MS calcd. for C\(_{17}\)H\(_{29}\)N\(_5\)O\(_6\)S: 431.1839, found: 454.1736 (+Na\(^+\));
**1H-NMR (400MHz, CD$_3$OD)** δ 4.48 (dd, $^3$J$_1$ = 4.9Hz, $^3$J$_2$ =7.6Hz, CH$_3$), 4.30 (dd, $^3$J$_1$ = 4.4Hz, $^3$J$_2$ = 7.8Hz, CH$_3$), 4.37 (s, 2H, CH$_2$), 3.20 (dt, 1H, $^3$J$_1$ = 6.2Hz, $^3$J$_2$ = 8.3Hz, SCH$_2$), 2.91 (dd, 1H, $^3$J = 4.9Hz, $^2$J = 12.7Hz, CHH cis), 2.69 (d, 1H, $^3$J = 12.7Hz, CHH trans), 2.28 (dt, 2H, $^3$J$_1$ = 1.6Hz, $^3$J$_2$ = 7.1Hz, CH$_3$), 1.80-1.45 (m, 6H, CH$_2$), 1.46 (s, 9H, CH$_3$);

**13C-NMR (75MHz, CD$_3$OD)** δ 172.98, 168.50, 164. 27, 157.74, 81.23 (CCH$_3$), 73.92, 61.37, 59. 77, 55.04, 39.15, 32.49, 27.62, 27.47, 26.60, 24.46.

**ARP reagent (7)**

Boc-protected ARP 6 (25.4 mg, 0.059 mmol) was dissolved in cold TFA (0 °C) and stirred for 20 min. The solution was allowed to warm to room temperature and stirred for another 1 h. The reaction mixture was concentrated, and the product was dissolved in methanol and isolated by precipitation in cold diethyl ether (-20 °C). Yield 17.7 mg (68%) of an off-white solid.

**1H-NMR (400MHz, DMSO)** δ 6.40 and 6.35 (br s, 2H, NH$_2$), 4.31 (t, 1H, J = 6.0 Hz, CH$_2$), 4.14 (obs t, 1H, CCH$_2$), 4.12 (s, 2H, CCH$_2$), 3.10 (dt, $^3$J$_1$ = 7.0 Hz, $^3$J$_2$ = 11.2 Hz, SCH$_2$), 2.81 (dd, 1H, $^3$J = 5.0Hz, $^2$J = 12.4 Hz, CHH cis), 2.58 (d, 1H, $^2$J = 12.4Hz, CHH trans), 2.13 (t, 2H, $^3$J = 7.2 Hz), 1.67-1.32 (m, 6H, CH$_2$);

**13C-NMR (75MHz, DMSO)** δ 170.98, 168.06, 162.68, 72.93, 70.87, 61.01, 59.18, 55.40, 32.92, 31.20, 28.27, 28.04, 28.01, 25.02, 21.23.

**Experimental procedures:**

**Peptide labelling**

The gp43 C-terminus peptide analogue (see Chapter 5) was labelled with FeBABE in a 20 mM sodium phosphate buffer pH 7.5 supplemented with 2 mM EDTA for 60 mins at room temperature and purified with reversed phase HPLC with 20-100% gradient of A: 0.1% TFA in acetonitrile and B: 0.1%TFA in MilliQ with 5% acetonitrile on an Alltech C8 column. The labelled peptides were analyzed with Maldi-TOF, concentrated, and dissolved in 50 mM sodium phosphate buffer pH 7.3, 1 mM EDTA, and quantified with the ferrozine assay to determine the iron content (described below in thiol, iron and protein quantification).

**Maldi-TOF MS (αCCA, m/z)** 1249.0 (unknown) , 1361.0 (M-Fe+Na+3H).

**Protein expression and labelling**

BL21(DE3) E. coli cells were transformed with a plasmid coding for the gp45 E212C mutant, grown to OD = 0.7, and induced with 0.4 mM IPTG. After 6 hrs, the cells were pelleted. The protein was purified as described previously. Conjugation with the oxidation labels was performed with 10 times excess of label in labelling buffer (50 mM Tris.HOAc pH 7.4, 1 mM EDTA, 10% glycerol), supplemented with 150 mM KOAc buffer pH 7.4 in the case of the porphyrin, at 4 °C for 5 hrs. The protein was purified from unreacted label by SEC (Superdex 200, Pharmacia) equilibrated with storage buffer (25 mM Tris.OAc pH 7.5, 50 mM KOAc, 10% glycerol), supplemented with 1 mM EDTA and 10 mM β-mercaptoethanol in the case of the porphyrin. The latter two were omitted for the FeBABE-labelled proteins, to minimize stripping of iron metals from the protein, and to avoid reduction of the iron. If stated the proteins or conjugates were further purified by a Source Q 26/10 column (Amersham) with a NaCl gradient (0-400 mM) in 40 mM Tris.HCl pH 7.4, 1 mM MgCl$_2$ followed by a HiPrep 26/10 desalting column equilibrated in the storage buffer described above.

**Thiol, iron and protein quantification**

Ellman’s test$^{11}$ was performed in 100 mM phosphate buffer pH 8.0 with an excess of 5,5′-dithiobis(2-nitrobenzoic acid). Quantification was done by UV-Vis spectroscopy ($\varepsilon_{412} = 14,150$ M$^{-1}$cm$^{-1}$). Iron concentrations were determined using the ferrozine assay$^{28,29}$ in a microwell plate, according to the protocol of Spierring et al.$^{30}$ Protein samples (20 µl) were heated to 90 °C for 30 min in 50 µl aqueous 2 M HCl. Precipitated protein was removed by centrifugation, followed by addition of 10 µl 75 mM sodium ascorbate and 50 µl of 10 mM ferrozine [3-(2-pyridyl)-5,6-bis(4-phenylsulfonic acid)-1,2,4-triazine] to the supernatant. The Fe$^{2+}$-(ferrozine) complex was formed upon the addition of 50 µl of a saturated ammonium
Oxidation of DNA substrates by clamp-based catalysts

Acetate solution. The absorption of the complex was compared to the absorption of standard samples, of which the iron concentrations were determined by UV-Vis spectroscopy after performing the test on a larger scale ($\varepsilon_{562} = 27,900 \text{ M}^{-1}\text{cm}^{-1}$). Bradford test was obtained from Bio-Rad and was used according to the instructions of the manufacturer. The concentration of clamp standards were determined by UV-Vis spectroscopy ($\varepsilon_{280} = 57,200 \text{ M}^{-1}\text{cm}^{-1}$).

ATPase assay
The ATPase activity assay was performed on a Varian Cary 100 UV-Vis spectrometer and performed as described in Chapter 5, using a 60 μl starting volume.

DNA templates
The plasmid used for oxidation experiments was a pGEM plasmid with a T4 genomic insert (3540 bp), which was purified using QiaGen maxiprep protocols. A nick was introduced by Nt.BbvCI or Nb.BbvCI (New England Biolabs), followed by purification by phenol chloroform extraction and ethanol precipitation.

Oxidation experiments
All oxidation reactions were carried out at room temperature in a 25 mM Tris.HOAc buffer pH 7.5 with 150 mM KOAc and 10 mM Mg(OAc)$_2$. For FeBABE oxidations, the samples were quenched by the addition of a quench and loading buffer, containing 25 mM Tris.HCl pH 8.0, 65 mM EDTA, 33 mM thiourea, 25% glycerol, orange G as dye. Oxidation reactions with Mn-TMPyP-conjugated gp45 proteins were quenched by the addition of HEPES-KOH pH 7.6 buffer to a final concentration of 150 mM, and extracted with 2 volumes of phenol-chloroform solution (1:1 v/v). The aqueous fraction was isolated and used for analysis. Samples from the oxidation experiments were analyzed on a 0.8% ethidium bromide-stained agarose gel made with 0.5x or 1x TBE buffer and run at ~80 mA. λ-DNA / HindIII digest (NEB) was used as molecular weight marker. When indicated, BSA that was specifically treated to remove nuclease contaminations (NEB) was added to prevent non-specific absorption.

Oxidation of supercoiled pGEM T4 with wt gp45 and FeBABE-labelled peptide
Various concentrations of wt gp45 (0 - 450 nM) were incubated with supercoiled pGEM T4 (25 ng / μl) and 10 μM FeBABE-labelled peptide. After 5 mins, the oxidation was started by the addition of DTT and H$_2$O$_2$ to a final concentration of 1 mM and 0.03% (2 μl added each), respectively. The total volume of the reaction was 20 μl. The reaction was incubated for 15 mins, after which 8 μl was quenched with 10 μl of 2.5x quench and loading buffer. The samples were analyzed on a 0.8% ethidium bromide-stained agarose gel.

Oxidation of supercoiled pGEM T4 with FeBABE-labelled gp45
Various concentrations of FeBABE-labelled gp45 (0 - 2500 nM) were incubated with supercoiled pGEM T4 (25 ng / μl). After 5 mins, the reaction was started by the addition of DTT and H$_2$O$_2$ to a final concentration of 1 mM and 0.03% (2 μl added each), respectively. The total volume of the reaction was 20 μl. After 1 and 5 mins, 8 μl of the reaction mixture was quenched with 10 μl of 2.5x quench and loading buffer. The samples were analyzed on a 0.8% ethidium bromide-stained agarose gel.

Oxidation of nicked pGEM T4 with FeBABE-labelled peptide and wt gp45
Various concentrations of wt gp45 (0 - 500 nM) were incubated with Nt.BbvCI nicked pGEM T4 (25 ng / μl) in the presence of 6 μM FeBABE-labelled peptide, 0.5 mM ATP, 50 μg/ml BSA, and 125 nM gp44/62. After 5 mins, the reaction was started by the addition of DTT and H$_2$O$_2$ to a final concentration of 10 mM and 0.1% (1 μl added each), respectively. The total volume of the reaction was 10 μl. After 5 mins the reactions were quenched with 5 μl of 2.5x quench and loading buffer.
Oxidation of supercoiled DNA with Mn-TMPyP-labelled gp45
Gp45 at a concentration of 400 nM was incubated with supercoiled pGEM T4 (25 ng / μl) in the presence of 50 μg/ml BSA. When indicated, unlabelled peptide (10 μM) and PEG (10% wt/v) were added. After 5 mins, the reaction was started by the addition of KHSO₅ to a final concentration of 5 μM. The total volume of the reaction was 60 μl. At the indicated times, the reactions were quenched, extracted, and analyzed. For the experiment with higher KHSO₅ concentrations (1 – 20 μM, 30 μl total volume), no peptide or PEG was added.

Oxidation of nicked DNA with Mn-TMPyP-labelled gp45
Gp45 at a concentration of 250 nM was incubated with Nt.BbvCI or Nb.BbvCI nicked DNA in the presence of 50 μg/ml BSA, 2 mM ATP and 125 nM gp44/62. When indicated, 250 nM gp43 or 10 μM peptide was added to the reaction solution. After 5 mins, the reaction was started by the addition of KHSO₅ to a concentration of 20 μM. The total volume of the reaction was 25 μl. At the indicated times, the reactions were quenched, extracted, and analyzed. One experiment was performed with gp45 at a concentration of 250 nM and 50 μg/ml BSA.

Calculation of the amount of single-stranded breaks
The oxidation reactions described in this chapter were analyzed by comparing the relative intensities of the ethidium bromide stained DNA bands on agarose gels. Detailed analysis, however, requires the determination of the mean number of single-stranded breaks in the DNA substrate. Although the transition from form I (supercoiled) to form II (nicked) is extremely sensitive, multiple single stranded breaks may be present on the form II plasmid that are not detected in the analysis. Several groups have developed mathematical models that calculate the mean number of single-stranded breaks (S) in DNA plasmids from the intensities of the bands on an agarose gel. First, in the absence of form III (linear), S is defined as:

\[
S = 4 \ln \frac{I_0}{I} - 3(I - I_0)
\]

Equation 1 takes into account the accessibility difference of cleavage sites between form I and form II due to differences in the DNA topology.33

\[
I + II = 1 - \left(1 - \frac{S \times (2h + 1)}{2L}\right)^{S/2}
\]

In this equation, I and II are the molar fractions of form I and II, respectively, and h is the maximal distance in nucleotides between two single-stranded breaks on opposite strands that will lead to a double-stranded break (and thus a linear DNA molecule). Estimates of this number range from \(h=16\), according to Freifeld and Trumbo35, to \(h=29\), according to Van Touw36, although the latter number was determined using a method that also introduces double stranded breaks directly. L is the number of base pairs in the DNA substrate.

DNA quantification
Ethidium bromide stained agarose gels were imaged by a Gel-Doc 1000 system (Bio-Rad). Bands were quantified by the volume analysis tool of Quantity One software version 4.6.2 (Bio-Rad). Intensities were corrected for local background and for the decreased ability to stain the supercoiled form over nicked and linear forms, applying a correction factor of 1.47.33

General procedure for the ARP labelling of oxidized DNA
Feasibility experiment (Figure 8):
Supercoiled pGEM plasmid (200 ng) was oxidized by 50 nM Mn-TMPyP 3 in the presence of 0.4 μM KHSO5 in 37.5 mM phosphate buffer pH 7.5, containing 100 mM NaCl. The oxidation was performed at 0 °C for 1 min. The reaction was quenched by the addition of 30 mM HEPES buffer pH 7.4 (final concentration). ARP was added to a final concentration of 5 mM and the mixture was incubated for 1.5 hrs. A small amount of NaCNBH3 was added and the solution was incubated for another 15 min. The reaction mixture was purified with the help of a G25 spin column. The purified solution was buffered with AFM buffer (40 mM HEPES pH 7.4, 10 mM MgCl2) (final concentration) and supplemented with 1 μl of a 50 μM streptavidin solution (a 50 fold excess with respect to the total amount of AAA sites in the DNA). This solution was incubated for 1 h, and subjected to a Sephadex G100 column equilibrated in AFM buffer. Protein containing fractions were identified via the Bradford assay. A sample (4 μl) of the first protein-containing fraction was incubated for 5 min on freshly cleaved mica, washed with 1 ml MilliQ, dried by air flow and analyzed by AFM.

Oxidized samples:
Quenched samples from the oxidation reactions were incubated with 4 mM ARP reagent. After 1.5 hours, the amination reaction was quenched by the addition of a dilute NaCNBH3 solution. The reactions were then incubated for 15 mins, and the DNA was precipitated with ethanol as described in Chapter 4. The pellets were dissolved in AFM buffer (40 mM HEPES pH 7.4, 10 mM MgCl2), and 10 fold of streptavidin was added per total amount of potential oxidation sites (~190 per plasmid). After the addition of streptavidin, the samples were incubated for another 30 mins before they were purified from excess streptavidin by a G100 sephadex column equilibrated in AFM buffer. The fraction with the highest amount of DNA was analyzed with AFM. A sample (2 μl) of the first protein-containing fraction was incubated for 5 min on freshly cleaved mica, washed with 1 ml MilliQ, dried by air flow and analyzed by AFM.

5. References

Atomic force microscopy analysis of T4 replication proteins and complexes

The T4 replication proteins are investigated with AFM in order to obtain a better understanding of the binding of these proteins to each other and to DNA. First, the T4 clamp protein (gp45), clamp loader complex (gp44/62), and DNA polymerase (gp43) were imaged individually. The complexes of gp44/62 and gp43 with gp45 were stabilized via glutaraldehyde cross-linking, purified by FPLC and imaged with AFM. This procedure was particularly successful for the gp45-gp43 complex, which was visualized as a banana-shaped structure. Subsequently, all three proteins were incubated with various DNA substrates. For gp45, the reaction conditions were optimized to facilitate binding to DNA, while gp43 formed aggregates with a long DNA substrate. Unexpectedly, multiple gp44/62 proteins were found to bind to a linear plasmid, while only one weak binding site was present. Finally, efforts were made to image the complete T4 holoenzyme and T4 replisome on DNA substrates. In the latter case, DNA synthesis reactions were imaged in the presence of the single stranded DNA binding protein gp32 and dNTPs.
1. Introduction

Since its development in 1986, the atomic force microscope has quickly gained prominence as a method to analyze nano and micro-scaled objects, including biological macromolecules such as DNA, proteins and enzymes and complexes thereof. With atomic force microscopy (AFM), these molecules can be studied at the single molecule level, revealing properties that would otherwise be masked by the fact that bulk measurements only measure the average properties of the sample. AFM has a number of advantages over other structural analysis techniques such as electron microscopy; a) Imaging can be done in air, vacuum or liquid; the latter can be a medium that allows the study of biomolecules in their native conformation and environment. b) The resolution can be as small as 1 nm in the lateral direction, but is typically 10-20 nm, while the vertical resolution is much smaller, viz. Ångströms, and excellent signal to noise ratios are achieved. c) Little or no sample treatment is required before the analysis. Lastly, changes in conformation and activity of the studied objects can be studied easily and directly.

Nucleic acids have been and still are a popular study object for AFM analysis, because sample treatment is extremely easy and different types of DNA substrates are commercially available. Short oligonucleotides, supercoiled, relaxed and linearized DNA plasmids and long DNA strands of viral or phage origin, as well as complex artificial DNA structures of biological relevance have all been studied. Designed DNA structures, lattices and devices, with potential for application in nanotechnology devices (reviewed by Seeman) are also routinely investigated with the help of AFM. In addition, DNA molecules can be stretched, teared and dissected at any position using the AFM tip. Moreover, AFM tips can also be used for force spectroscopy, in which the force of the tip or an analyte bound to the tip towards the sample can be measured at the single molecule level. In the case of DNA, untreated tips can be used to pick up individual strands from the surface, allowing the stability of DNA in the presence of DNA binding molecules to be examined. Imaging of DNA has also demonstrated that care should be taken with the interpretation of the obtained size of the objects investigated by AFM measurements. Lateral resolution is limited by the shape of the end of the imaging tip, which leads to an overestimation of the object size. The height, on the other hand, can be underestimated due to dehydration and deformation of the sample. In the case of DNA, which based on crystal structures should appear as a filament with a diameter of 2 nm, typical reported widths and heights are ~17 nm and ~0.4 nm, respectively.

Complexes of proteins and DNA have also been visualized with AFM. The E. coli RecA protein, involved in homologous DNA recombination and repair, bound to DNA showed that the protein fully covers DNA to form left-handed superhelical bundles, while in the absence of DNA it forms ring-like complexes. The binding of E. coli RNA polymerase to its promoter involves wrapping DNA around the protein, which results in the bending of the promoter site by 55° - 88°. The activity of the RNA polymerase complex could be monitored via visualizing the translocation of the protein along the DNA template. The rolling circle amplification of plasmid DNA by bacteriophage phi29 DNA polymerase was also
Atomic force microscopy analysis of T4 replication proteins and complexes

successfully studied by AFM, although only the DNA products were analyzed.\(^37\) BAL 31 exonuclease and DNase I activity was also investigated by AFM.\(^38,39\) Several adaptations, such as using short cantilevers and imaging in phase rather than height mode, allow rapid imaging of objects without applying high forces on the sample.\(^40\)

The T4 replication proteins and their complexes on DNA have not yet been subjected to analysis with AFM, with the exception of gp32.\(^41\) Several electron microscopy studies have provided useful evidence on the replisome complex and the different proteins therein (reviewed in Chapter 1).\(^42,43\) AFM experiments, however, may provide more information about the complexes in physiologic conditions (buffer, normal pressure and temperatures), and will allow activities to be monitored in time and to be correlated directly with protein conformation and stochiometry of the respective proteins in these complexes. Furthermore, the development of methods to image the T4 replication proteins would allow us to monitor the activity of the catalyst labelled clamp proteins on DNA substrates with AFM or on a combined AFM – confocal microscopy setup. The experiments herein can be considered a first approach towards this goal.

2. Results and discussion

2.1 Imaging of the T4 replication proteins

The three core components of the T4 replisome, also called the T4 holoenzyme, were observed individually by AFM after deposition on a mica surface (see experimental section). The T4 clamp protein (gp45, Figure 1A) is visualized by AFM as a globular structure, with a height of 2 nm and a diameter of 27 nm. This is consistent with a clamp protein lying flat on the surface, since the crystal structure of the clamp resembles a donut with a width of 8 nm and a height of 3 nm. The resolution is not high enough to visualize the central cavity of the protein. There seems to be some variation in the particle size (see also Figure 1B), which might be caused by the presence of gp45 proteins with lower oligomeric state (dimers and monomers), or different modes of physisorption to the surface. For the clamp loader complex (gp44/62) less variation in the size of the imaged objects was observed. The height is 1.0 - 1.5 nm, but sometimes objects are present that have twice this height. Unfortunately, the crystal structure of the complete T4 clamp protein complex is not available, and the heights observed by AFM cannot be correlated to the actual size of the complex. The T4 DNA

| Table 1 – Average size values measured for dried-in samples of T4 replication proteins on mica. |
|-----------------------------------------------|---------------|------------------|
| average diameter (nm) | average maximum height (nm) |
| gp45 | 27.4 | 2.0 |
| gp44/62 | 18.3 | 1.0 |
| gp43 | 19.4 | 0.7 |
A. T4 clamp protein, gp45

B. 120nm

C. T4 clamp loader complex, gp44/62

D. 150nm

E. T4 DNA polymerase, gp43

F. 120nm
polymerase protein (gp43) molecules are even smaller, with an average height of 0.65 nm. All T4 replication proteins have an apparent globular structure. The diameters were estimated by the flooding tool of the WSxM software\textsuperscript{44} and are reported in Table 1.

2.2 Cross-linked complexes of the T4 replication proteins

The T4 replication proteins not only assemble with DNA in the process of forming the replisome, but also form complexes in solution without any DNA template. The clamp loader protein gp44/62 binds to the clamp gp45 in solution,\textsuperscript{45,46} as does the DNA polymerase gp43.\textsuperscript{47,48} In both cases, the dissociation constants are in the nanomolar region (1-8 nM for the gp45-gp44/62 complex and 48-480 nM for the gp43-gp45 complex). In order to visualize these protein complexes with AFM, the proteins were incubated with each other at relatively high concentrations, and cross-linked with glutaraldehyde to stabilize the complexes. The complexes were subsequently purified with the help of a size-exclusion column on a FPLC system (Figure 2). In this procedure, the fluorescein-labelled gp45 E212C mutant was used to identify peaks of clamp-containing complexes. Unexpectedly, in blank experiments in which the proteins were injected separately, only an elution peak for gp45 (Figure 2A) was observed and none for gp44/62 and gp43. The mixtures of cross-linked protein complexes were nevertheless subjected to size exclusion chromatography. The main fractions that had a smaller retention volume than the clamp were analyzed, since these fractions are expected to consist of complexes that have a higher mass than the clamp trimer itself. The fluorescein absorption at 495 nm was used to identify clamp-containing fractions. The clamp loader protein was cross-linked to gp45 both in the presence and absence of ATP. In both cases, AFM analysis showed globular objects of similar height and diameter (Figure 3A). Although the diameter of these objects was even lower than for the clamp itself (20 nm vs. 27 nm), the height was higher (3.5 nm) than the average heights that were found for the clamp and clamp loader individually (2.0 nm and 1.0 nm, respectively). For the complex of gp45 with gp43, two elution fractions were analyzed. The first fraction (0.98 ml) showed globular...
A. **Cross-linking gp45**

![Cross-linking gp45](image1)

B. **Cross-linking gp45-gp44/62**

![Cross-linking gp45-gp44/62](image2)

C. **Cross-linking gp45-gp43**

![Cross-linking gp45-gp43](image3)
Atomic force microscopy analysis of T4 replication proteins and complexes

A. gp45-gp44/62 complex

B. gp45-gp43 complex

C. gp45-gp43 complex

D. gp45-gp43 complex

E. gp45-gp43 complex

F. gp45-gp43 complex
Figure 3, previous page – AFM height images of dried T4 replication protein complexes on a mica surface. A and B. Image of gp45-gp44/62 crosslinked complexes in the absence (A) and presence (B) of ATP. C. Image of the first peak fraction of the gp45-gp43 complex (0.90 ml) D. Image of the second peak fraction of the gp45-gp43 complex (1.06 ml). Banana-shaped objects are indicated with the arrows. E. Image of zoomed region in C. The insert is a model of the complex and is based on the model of the T4 holoenzyme.48 F. Height analysis of image D (trajectory shown at top).

objects (Figure 3C), whilst the second fraction (1.09 ml) contained a number of characteristic banana-shaped objects that appeared to consist of two different globular entities (Figure 3D). These objects were attributed to cross-linked gp45-gp43 complexes. Co-injection of gp45 and gp43 did not yield any peaks of higher-molecular weight than that observed for gp45. Interestingly, SDS-PAGE analysis of the cross-linked protein complexes also showed a band shift when gp43 was treated with glutaraldehyde in the absence of gp45 (data not shown), which may indicate that gp43 forms homodimers in solution, as it does in the replisome.49 Although this homodimer was not investigated with AFM, it is unlikely that the objects assigned as gp45-gp43 complexes are gp43 homodimers since these banana-shaped are formed from two objects of different size. As with the individual T4 replication proteins, the cross-linked objects were again analyzed by WSxM software to determine their sizes (Table 2).

Table 2 – Average size values measured for dried-in samples of T4 replication protein complexes on mica.

<table>
<thead>
<tr>
<th>Average diameter (nm)</th>
<th>Average maximum height (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>gp45-gp44/62 with ATP</td>
<td>20.7  3.5</td>
</tr>
<tr>
<td>gp45-gp44/62 without ATP</td>
<td>20.1  2.5</td>
</tr>
<tr>
<td>gp45-gp43 (0.90 ml)</td>
<td>16.0  1.3</td>
</tr>
<tr>
<td>gp45-gp43 (1.06 ml)</td>
<td>20.9  4.5</td>
</tr>
</tbody>
</table>

2.3 Complexes of the T4 clamp on DNA

The T4 clamp on DNA was studied using the long λ-DNA (48.5 kbp, 15-17 μm in length) as substrate, in order to render the dissociation of the clamp from the template at its ends unlikely. Entanglement of multiple strands is known to prevent single molecules to be imaged individually. The DNA strands were therefore aligned in one direction by applying a

Figure 4, opposing page – AFM height images of the T4 clamp gp45 on DNA. A. Image of complexes prepared with gp45 and λ-DNA in 10% PEG solution, showing two globular objects that were assigned as gp45 proteins. The λ-DNA was stretched by gas flow-alignment. B. Height profile of the complexes on DNA, trajectory shown at the top. C. Image of complexes made with BamHI-linearized pBR322 DNA in 10% PEG solution, showing one DNA-bound gp45 protein. D. Profile of the complex on DNA, trajectory shown at the top. E. Image of complexes prepared with supercoiled pGEM T4 plasmid as substrate. The gp45 proteins, of which two are bound to this particular DNA molecule, were cross-linked to DNA by glutaraldehyde treatment. F. Height profile of the complexes on DNA, trajectory shown at the top. Bare DNA is indicated with *, complexes with #.
Atomic force microscopy analysis of T4 replication proteins and complexes

A.  

B.  

C.  

D.  

E.  

F.  

130nm  

94nm  

140nm  

130nm  

82nm  

- 193 -
nitrogen flow.\textsuperscript{50-52} This method was also pursued in our lab to align \(\lambda\)-DNA on a glass substrate in order to create an alignment layer for liquid crystal cells that would amplify events (e.g. digestion, drug binding) to DNA at the molecular level to macroscopic changes in liquid crystal ordering visible with the naked eye.\textsuperscript{53}

The solutions containing the gp45-DNA complexes were originally made in complex buffer supplemented with 10\% PEG (\(M_N = 8,000\)), in order to enhance the binding efficiency and to stabilize the complex of the clamp on DNA via macromolecular crowding.\textsuperscript{54,55} Optimization of the reactant concentrations led to the appearance of globular features on the DNA with a height of 2 nm (Figure 4), consistent with clamp proteins bound to DNA. In this particular sample, apparent local DNA melting was observed near two clamp proteins. It is unclear at the moment if the melting was induced by the presence of the clamp proteins, or whether it is an artefact of the sample preparation. In the latter case, the proteins could have been trapped by a bubble, which prevented them to dissociate from the substrate during the final wash step.

The binding efficiency of the clamp to DNA was low. This led us to further improve the reaction conditions. First, the buffer was replaced by the commonly used buffer for DNA imaging (40 mM HEPES.KOH pH 7.4, 10 mM MgCl\(_2\)). ATPase assays (see Chapter 3) performed in this buffer confirmed that the clamp loading by the clamp loader complex on a forked DNA substrate was equally efficient as in complex buffer. Although gp44/62 was not included in the experiment described in this section, it indicates that the interaction of the clamp is most likely not significantly different in these two buffers. Second, the macromolecular crowding reagent was omitted from the solution, because it was recently found that gp45 can bind non-specifically to DNA in single molecule experiments on surface-bound fork templates.\textsuperscript{56} Finally, a pGEM derived DNA plasmid (3.5 kbp) or the pBR322 plasmid (4.4 kbp) was used as DNA substrate to avoid the need for alignment of long \(\lambda\)-DNA molecules. Ambiguous identification of protein complexes due to the supercoiled structure of the plasmids was minimized by digesting the plasmids with restriction enzymes, yielding a small, linear DNA substrate. As expected, DNA binding to the surface was very efficient in the HEPES buffer. In spite of all the above steps, the binding efficiency of the clamp to the DNA in the improved conditions was still very low. It was noted, however, that in complex buffer only linearized plasmids could be imaged, while in the HEPES buffer all DNA substrates could be imaged very efficiently.

2.4 Complexes of the T4 clamp loader and DNA polymerase on DNA

Incubation of the clamp loader protein gp44/62 with Ndel-linearized DNA resulted in the visualization of a number of globular features on the DNA strand. It is known that the clamp loader preferentially binds to 5’ extended overhangs and nicked sites.\textsuperscript{57,58} Despite the small 2-nt overhang that results from the enzymatic digestion reaction, a globular feature is seen at one end of the template (Figure 5). Surprisingly, more features are seen along the DNA strand. Possibly, gp44/62 also has some affinity with dsDNA, which could help gp44/62 to
find binding sites along the template, reminiscent of gp43\textsuperscript{59} and other DNA binding proteins.\textsuperscript{60} Gp43 complexes on \(\lambda\)-DNA were imaged as aggregated DNA structures with a central globular feature. Although the same sample preparation procedure was applied as with \(\lambda\)-DNA and gp45 (see Figure 5C and D), the DNA strands were not aligned and extended in this case. The average height of the aggregates was \(~9\) nm, with a diameter of 300-350 nm. Although the DNA substrate does not possess the preferred primer-template junction binding site, gp43 is also known to bind to dsDNA with moderate affinity (\(K_D = 150\) nM).\textsuperscript{59} Nevertheless, it was not anticipated that these condensed gp43—\(\lambda\)-DNA structures would form. Due to the absence of a well-defined structure, the imaging of gp43 on smaller substrates was not pursued further.

A. gp44/62 – DNA complex

B.

\[\text{AFM height image of gp44/62 on BamHI-linearized pBR322 DNA. Zoomed area of image in A. Height profile of the complexes on DNA, trajectory shown at the top.} \]

C. gp43 – \(\lambda\)-DNA complex

D.

\[\text{AFM height image showing four gp43 – \(\lambda\)-DNA aggregates. Zoomed image of one of the aggregates, which has a height of 9 nm and a diameter of 300 nm.} \]

\textbf{Figure 5} – A. AFM height images of gp44/62 on BamHI-linearized pBR322 DNA. B. Zoomed area of image in A. Height profile of the complexes on DNA, trajectory shown at the top. C. AFM height image showing four gp43 – \(\lambda\)-DNA aggregates. D. Zoomed image of one of the aggregates, which has a height of 9 nm and a diameter of 300 nm.
2.5 Imaging of the T4 holoenzyme and T4 replication complexes

Initially, we tried to image the holoenzyme by incubating λ-DNA with gp45 and gp43. This procedure did not result in the visualization of protein complexes on DNA. The DNA substrate itself, however, showed a remarkable feature in that it was partially denatured (Figure 6A). Although this might be attributed to an artefact of the sample preparation, for instance be the results of denaturation of the DNA due to a very low salt concentration during the wash step, it is important to note that this phenomenon was never observed in any of the control samples. Gp45 and gp43 do not have known strand displacement activities, and possible helicase impurities would require ATP or GTP for activity, which were absent in this experiment. Interestingly, partial denaturation of the DNA substrate was also observed in the case of the binding of gp45 to λ-DNA (Figure 4A).

Recently, Chamberlain and Nossal published a method to fix T4 replication proteins on a DNA template for electron microscopy analysis.43 By using a clever labelling method based on streptavidin labelling of biotinylated replication proteins, the proteins that are part of the replisome could be quantified (see Chapter 1). The cross-linking methods described in this report were also applied to analyze the T4 replication proteins with AFM. A nicked pGEM T4 plasmid (see Chapter 7) was used as a simplified replication substrate. In this substrate, replication can only start from the nicked site, and should continue until the nucleic acid triphosphate pool becomes depleted. The substrate was incubated with a mixture of the single stranded DNA binding protein gp32, gp43, gp45 and gp44/62 in the presence of ATP to assemble the replisome, and subsequently supplemented with dNTPs to start the replication reaction. Effectively, the replisome will continuously move around the closed circular template strand, displacing single stranded DNA (Figure 6E). Although no helicase (gp41) was added to the reaction mixture, the remaining proteins are still able to synthesize new DNA at a speed of ~10-30 nt/s.61,62 At specific time points, the proteins were fixed using a glutaraldehyde cross-linking method, adapted from Chamberlain and Nossal,43 or subjected to a phenol-chloroform step to remove proteins and to investigate the naked DNA intermediates. The DNA complexes were also purified from unbound proteins.
Atomic force microscopy analysis of T4 replication proteins and complexes

A.

B.

C.

D.

E.

- 197 -
proteins by a size-exclusion chromatography step. AFM analysis of the trapped reaction intermediates at 1 and 5 mins show predominantly linear DNA strands (Figures 7B and C).

In the case of the glutaraldehyde cross-linked samples, significantly more proteins are bound to the DNA than with the extracted sample. Both samples, however, are characterized by a poor background that might originate from remaining proteins that were not removed by size exclusion chromatography or phenol-chloroform extraction. More strikingly, the circular plasmid that served as reaction template was not found in the investigated areas of the sample, and the proteins are distributed over the linear DNA strands. This would not be expected if a well-defined replisome was formed. Possibly, these proteins are the single stranded binding protein gp32, but since gp32 binds cooperatively to single stranded DNA, they should form ‘pearl-necklace’ structures, which further condense into ‘bobbins’, as observed by Chastain and Nossal in their EM studies. Alternatively, these structures are clamp proteins that are loaded onto the DNA by the clamp loader. From the described experiments it cannot be determined if the imaged strands consist of single or double stranded DNA, although due to the absence of the primase gp61 only single stranded DNA can be synthesized. Samples taken from the reaction mixture at 10 mins contained complex DNA networks (Figure 6D). These networks are thought to be the result of the large amount of single stranded DNA regions that are present in the reaction resulting in annealing of two strands, or non-specific cross-linking caused by glutaraldehyde. The gp32 protein pool is probably depleted after 10 mins, which leaves the DNA regions open for annealing or cross-linking. This sample contained numerous large structures, possibly aggregated proteins, preventing the DNA networks to be investigated more closely.

3. Conclusion

The work described in this chapter can be regarded as the first step towards imaging of DNA replication process by T4 replication proteins with AFM in real-time. This is important to fully understand all the processes that occur on the molecular level during replication. As explained in Chapter 7, the oxidation of DNA by porphyrin-labelled clamp proteins could also be monitored in real-time. A straightforward procedure was followed; first, the major T4 replication proteins gp43, gp44/62 and gp45 were imaged individually. Then, we tried to visualize the complexes that are known to exist between these proteins by AFM. A glutaraldehyde-based cross-linking approach, followed by FPLC purification, did not give gp45-gp44/62 complexes. The same method, however, yielded banana-shaped objects for the complex between gp45 and gp43.

The replication proteins were also investigated in the presence of DNA substrates. In the case of gp45-DNA, globular objects on different DNA substrates were observed. Incubation of gp43 with λ-DNA resulted in aggregated structures, which were not investigated further. For gp44/62, globular-shaped objects on DNA were again observed.

Finally, replication reactions with all the replication proteins needed for efficient leading strand synthesis were performed and the intermediates were visualized with the
help of AFM. Although in the EM studies reported in the literature the replication intermediates appear as well defined objects, in our AFM study distributed proteins on linear DNA strands and complex networks of DNA were observed. Possibly, the concentration of gp32 was not high enough to cover all the newly synthesized single stranded DNA regions, which may result in the hybridization with other DNA regions.

Although our study was only preliminary, AFM appears to have potential as a method for the identification and study of DNA replication reaction intermediates.

4. Experimental section

General

AFM experiments were performed on Veeco Nanoscope III or IV AFM multimode microscopes operating in tapping mode at room temperature. Unless stated otherwise, samples were incubated for one minute on freshly cleaved and unmodified Muskovita mica, washed with 1-2 ml water (MilliQ grade) and blown dry with a nitrogen flow. Images (512 x 512 lines) were recorded at scanning speeds of ~1 Hz. Image analysis was done with the WSXM software, version 4.0 Develop 10.4.44 The images shown in the figures may have been zoomed. Lambda (λ) DNA (500 µg/ml) and pBR322 (1 mg/ml) were from New England Biolabs. The pGEM derived DNA plasmid (pGEM T4)63 was multiplied using E. coli XL1-blue cells and purified according to Qiagen Maxiprep protocols. DNA plasmids were digested with BamHI (for pBR322), Ndel, or Nt.BbvCI (pGEM T4) until completion (judged by agarose gel electrophoresis), and purified by phenol-chloroform extraction followed by ethanol precipitation (see Chapter 7). Final solutions (200-400 ng/µl) were in 10 mM Tris.HCl pH 7.6, 1 mM EDTA. Restriction and nicking enzymes were from New England Biolabs. The T4 replication proteins gp43, gp45, gp44/62 were generous gifts from the Benkovic research group. Gp32 was from New England Biolabs. Complex buffer consisted of 25 mM Tris.HOAc pH 7.5, 150 mM KOAc, 10 mM MgCl2, and AFM buffer consisted of 40 mM HEPES.HCl pH 7.4, 10 mM MgCl2.

Imaging of the T4 replication proteins

- **gp45**: T4 gp45 was diluted to 5 nM in complex buffer supplemented with 10% PEG (M₅₈ = 8,000) or in AFM buffer and imaged according to the standard protocol.
- **gp43**: T4 gp43 was diluted to 6.5 nM in AFM buffer and imaged according to the standard protocol.
- **gp44/62**: T4 gp44/62 was diluted to 4.5 nM in complex buffer and imaged according to the standard protocol.

Cross-linking of T4 replication proteins

The proteins (fluorescein-labelled gp45 to gp44/62 and gp45 to gp43), 1 µM of each protein in a total volume of 30 µl, were cross-linked by the addition of 2.25 µl 8% glutaraldehyde solution (final concentration 0,56%). After 5 mins, the reaction was quenched by the addition of 15 µl 400 mM Tris.HCl pH 7.6, 20 mM EDTA. The reaction mixtures were then injected onto a Pharmacia Superose 12 column (separating between 10 – 300 kD) on a Pharmacia SMART FPLC system with 40 mM HEPES.KOH pH 7.4, 10 mM MgCl2 as eluent at room temperature. Fractions were isolated and analyzed as described above.

Imaging of gp45 on DNA

On λ-DNA: The complex was formed in complex buffer supplemented with 10% wt/vol PEG 8,000 at 5 nM gp45 and 7.5 µg/ml λ-DNA concentrations. A small volume (1 µl) of this solution was incubated on mica, and the DNA strands were aligned by blowing the droplet in one direction using a nitrogen flow. The sample was washed and blown dry in the same direction.
Chapter 8

On linear pBR322: The complex was formed in complex buffer supplemented with 10% wt/vol PEG 8,000 and 5.5 μM of the gp43 C-terminus peptide analogue (see Chapter 5) at a 1.25 nM gp45 and 2 ng/μl BamHI-digested pBR322. Alternatively, the complex was formed on the same DNA substrate in AFM buffer without PEG and peptide.

Crosslinked: Gp45 (260 nM) was incubated for 5 mins with supercoiled pGEM T4 (25 ng/μl) in AFM buffer. Glutaraldehyde was added to 0.1 wt% final concentration, and the solution was incubated for another 3 mins. The sample was purified on a home-made Sephadex G100 spin column, and analyzed as described above.

Imaging of gp43 on DNA
The complex was formed in complex buffer supplemented with 10% wt/v PEG 8,000 at a 12 nM gp43 and 6.3 ng/μl λ-DNA concentration. Sample preparation was the same as described above for gp45.

Imaging of gp44/62 on DNA
The complex was formed in AFM buffer supplemented at a 3.5 nM gp44/62 and 2 ng/μl NdeI-linearized pGEM T4 DNA concentration. Sample preparation was the same as described above for gp45.

Imaging of the T4 holoenzyme on DNA
On λ-DNA: The complex was formed in complex buffer supplemented with PEG at a 0.8 nM gp43, 4.9 μM gp45 and 6.3 μg/ml λ-DNA concentration. Sample preparation was the same as described above for gp45.

Replication reactions with T4 replication proteins
A solution of Nt.BbvCI nicked pGEM T4 (250 nM) in complex buffer (50 mM HEPES.KOH pH 7.5, 100 mM KCl, 5 mM MgCl2, 1 mM EDTA and 10 mM β-mercaptoethanol) and 22.5% glycerol was incubated at room temperature with gp32 (2 μM), gp43, gp44/62 and gp45 (250 nM each), in the presence of ATP (2 mM), and BSA (100 μg/ml). After 6 mins, dNTPs were added to a concentration of 250 μM of each nucleotide. The total volume of the reaction mixture was 125 μl. After 1, 5 and 10 mins, 20 μl of the reaction mixture was added to 1.5 μl of 8% glutaraldehyde solution (0.1%), incubated for 5 mins, and supplemented with 11 μl 0.4 M Tris.HCl pH 7.5, 20 mM EDTA. Alternatively, 20 μl of the reaction mixture was added to 40 μl of phenol-chloroform solution (1:4 v/v) buffered with 100 mM Tris-HCl pH 7.6. The solutions were vortexed, separated by centrifugation, and the aqueous layer was isolated. The cross-linked or extracted solutions were then purified on home-made Sephadex G100 spin columns (600 μl of 80% slurry centrifuged 2 mins at 2000 rpm). AFM imaging was carried out as described above.

5. References

Atomic force microscopy analysis of T4 replication proteins and complexes


Chapter 8

47. G.J. Latham, D.J. Bacheller, P. Pietroni, and P.H. von Hippel, Structural analyses of gp45 sliding clamp interactions during assembly of the bacteriophage T4 DNA polymerase holoenzyme. III. The Gp45 DNA polymerase binds to the same face of the sliding clamp as the clamp loader, *J. Biol. Chem.* **272** (50), 31685-31692 (1997).
55. G.M. Sanders, G.A. Kassavetis, and E.P. Geiduschek, Use of a macromolecular crowding agent to dissect interactions and define functions in transcriptional activation by a DNA-tracking protein:


Appendix

1. Fluorescence resonance energy transfer (FRET)

Resonance energy transfer (RET) is the transfer of energy from a fluorophore donor to an acceptor that occurs when the donor and acceptor are within an appropriate distance and when there is significant overlap between the emission spectrum of the donor and the absorption spectrum of the acceptor. RET is not the result of emission from the donor being absorbed by the acceptor. It is non-radiative, thus the donor does not emit a photon, but rather transfers its energy via dipole-dipole interactions, curiously, this process is commonly referred to as fluorescence energy transfer (FRET), yet there is no fluorescence transfer involved. Since the donor emission quenching is strongly dependent on the distance between the donor and acceptor (see equation 1 below), this technique can be used as a spectroscopic ruler (Figure 1).

\[
E_T = \frac{R_0^6}{R_0^6 + R^6}
\]

In this equation, \( R \) is the distance between the two fluorophores and \( R_0 \) is the Förster distance. The Förster distance is different for every donor-acceptor pair and strongly depends on the nature and spectral overlap between the donor and the acceptor.

Figure 1 – The effect of the distance between the donor and acceptor (\( R \)) on the energy transfer (\( E_T \))
2. Calculating donor-acceptor distances from stopped-flow fluorescence spectroscopy experiments

This section describes how the distance between a donor and an acceptor fluorophore is calculated with stopped-flow FRET experiments. The technique and has mainly been adapted from Alley et al. and Trakselis et al.\textsuperscript{4-6} Distances between the donor (tryptophan) and the coumarin phenyl maleimide (CPM) acceptor can be derived using the following equations. First, the general dependence of FRET on the distance has a $1/R^6$ relationship\textsuperscript{7} (equation 1) which can be rearranged to equation 2:

$$R = R_0 \times \frac{1}{\sqrt{E_T} - 1}$$

in which $E_T$ is the transfer efficiency of the FRET process. $E_T$ can be expressed in the fluorescence intensities of the donor in the presence and absence of acceptor (CPM): $I_{AD}$ and $I_D$, respectively, and in the fluorescence intensity of the acceptor in the presence and absence of donor ($I_{AD}$ and $I_A$):

$$E_T = \left(1 - \frac{I_{AD}}{I_D}\right)$$

Equation 3 is the relation used for quenching of the tryptophan residue, and equation 4 describes the acceptor sensitization. The constants $\varepsilon_A$ and $\varepsilon_D$ are the extinction coefficients of the acceptor and donor at 290nm, which in the case of tryptophan and CPM is 3340 and 4100 M$^{-1}$·cm$^{-1}$, respectively, as determined on unlabelled CPM and L-tryptophan. The $I_{AD}/I_A$ ratio was determined by dividing the fluorescence intensities obtained from the corresponding double and triple mutants under identical conditions after excitation at 290 nm, and was used to calculate the starting point distance between the donor and acceptor.

The Förster distance $R_0$ is the distance at which 50% of the energy is transferred, and this distance strongly depends on the nature and the spectral overlap between the acceptor and donor, and on the medium:

$$R_0 = 0.211 \left(\frac{\phi_D \kappa \eta^{-2} J}{J}\right)^{1/6}$$

where $\phi_D$ is the quantum yield of the donor, \textit{i.e.} the ratio between the fluorescence emission intensity and the UV absorption of the probe. The parameter $\phi_D$ was calculated by comparing the absorption (A) and fluorescence emission (F) intensities of wt gp45 (subscript D) to those of the reference compound L-tryptophan (subscript R).\textsuperscript{4}
Appendix

\[ \phi_D = \varphi_A \left( \frac{F_D}{F_R} \right) \left( \frac{A_R}{A_D} \right) \]

6

The orientation factor \( \kappa^2 \) is assumed to be 2/3 for a freely rotating probe.\(^7\) The latter assumption was validated by Alley et al., using fluorescence anisotropy measurements\(^4\), and resulted in less than 10% error in the calculated distance between acceptor and donor.\(^8\) The parameter \( \eta \) is the refractive index of the medium (assumed to be 1.4; normally it ranges from 1.3 - 1.4), and \( J \) the overlap integral of the emission spectrum of the donor and excitation spectrum of the acceptor, which can be calculated with equation 7:

\[
J = \int F_D(\lambda) \times \varepsilon_A(\lambda) \times \lambda^4 \times \delta \lambda
\]

7

in which \( F_D(\lambda) \) is the fluorescence intensity of the donor as a fraction of the total integrated intensity, \( \varepsilon_A(\lambda) \) the extinction coefficient of the acceptor, and \( \lambda \) the wavelength of the overlap. In the case of the double mutants, \( R_0 \) was calculated to be 29 Å.

Using stopped flow fluorescence experiments, no absolute fluorescence data can be obtained. In order to correlate the normalized fluorescence changes from these experiments to real fluorescence intensities, which are needed to accurately calculate distances, the values from steady state fluorescence spectroscopy experiments were used as a reference allowing the fluorescence intensities (\( I_{AD} \) and \( I_A \)) to be quantified. The values of the reference state were determined at several points in the assembly process. The normalized fluorescence values are then corrected for the amount of interprotein FRET, i.e. excitation of the acceptor by donors of other proteins, hence the \( F_{AD^{290}} \) and \( F_{AD^{390}} \) are subtracted from each other, yielding the amount of interprotein FRET (\( X \)). Excitation at 390 nm results in fluorescence without any contribution from interprotein FRET, thus \( F_{AD^{390}} \) is directly related to \( I_A \). Since \( F_{AD^{290}} \) does have a contribution from the interprotein FRET, comparison of these values leads to a direct measurement of this component. It is important to note that the intensity of the normalized emission spectra does not change when different excitation wavelengths are used.\(^9\) The determination of fluorescence intensities from normalized emission intensities, and the determination of the amount of interprotein FRET is described in detail by Alley et al.\(^5\)

Gp45 consists of three identical subunits and thus has three subunit interfaces. Two of these interfaces are closed in solution, while the other one is opened.\(^4,\(^10\) In the FRET experiments, the average of the transfer across the open and closed subunit interfaces is measured, and consequently the efficient transfer between the tryptophan donor and CPM across the closed subunit interface has to be subtracted from that of the open subunit interface. The latter distance is assumed to be fixed, at a distance of 19 Å for the V162C mutant. The transfer efficiency across the closed subunit interface (\( E_C \)) is assumed to be 0.95, and the efficiency across the open interface (\( E_O \)) is calculated as follows:\(^4\)
Appendix

\[ E_o = 3E_r - 2E_c \] 8

The actual measured transfer efficiency is the average of the transfer across two closed subunits and one open:

\[ E_r = \frac{(2E_c + E_o)}{3} \] 9

By inserting equation 9 into equation 2, the precise distance between the open subunit interfaces can be measured.

3. Analysis of FRET titration experiments

The fluorescence titration experiments with the dansyl labelled peptide and the T4 clamp, described in chapter 5 was fitted to a 1:1 binding model. This model was derived as follows:

Consider the reversible binding reaction of the peptide I to the clamp protein E:

\[ E + I \rightleftharpoons E \cdot I \] 1

for which the equilibrium constant is described as:

\[ K_a = \frac{[E \cdot I]}{[E][I]} \quad \text{or} \quad K_d = \frac{[E][I]}{[E \cdot I]} \] 2

where the brackets indicate the concentration.

Biological binding events range from weak interactions (\( K_D > 10^{-3} \)) to very strong interactions, such as binding of hormones to receptors, antigens to antibodies or biotin to streptavidin (\( K_D = 10^{-11}-10^{-13} \)). Except for the latter cases, in which the binding is virtually not an equilibrium process, the ensemble of ligands (in this case peptide I) is either in a bound or unbound state. The fraction of protein containing a bound I, \( i.e. \) the fraction of filled binding sites is given by:

\[ \nu = \frac{[E \cdot I]}{[E]} \] 3

Next, the concentration of complex \([E \cdot I]\) will be expressed in the total concentrations of peptide \([I]\) and protein \([E]\). Since \([I]_T = [I]_U + [E \cdot I]\), \([I]_U\) can be expressed as:

\[ [I]_U = [I]_T - [E \cdot I] \] 4

and, likewise,

\[ [E]_T = [E]_U + [E \cdot I] \] 5

Thus, suffix T represents total concentration of a given species, and suffix U represents the free, or unbound concentration of a given species.
The association constant $K_a$ (equation 2) can be rearranged to:

$$[E]_u = \frac{[E \cdot I]}{K_a \times [I]_u} \tag{6}$$

Substitution of equation 6 into 5 yields:

$$[E]_f = [E \cdot I] + \frac{[E \cdot I]}{K_a \times [I]_u} \tag{7}$$

in which equation 4 can also be substituted:

$$[E]_f = [E \cdot I] + \frac{[E \cdot I]}{K_a \times ([I]_f - [E \cdot I])} \tag{8}$$

This equation, which removes the necessity of knowing the free (unbound) concentrations of protein and peptide, can be rearranged to:

$$K_a [E \cdot I]^2 - (K_a [E]_f + K_a [I]_f + 1)[E \cdot I] + K_a [E]_f [I]_f = 0 \tag{9}$$

and solved for $[E \cdot I]$ using the standard solution for a quadratic equation:

$$[E \cdot I] = \frac{1 + K_a [E]_f + K_a [I]_f \pm \sqrt{4K_a [E]_f [I]_f + (K_a [E]_f + K_a [I]_f + 1)^2}}{2K_a} \tag{10}$$

Using the definition for the binding ratio $\nu$ (equation 3) we can write:

$$\nu = \frac{1 + K_a [E]_f + K_a [I]_f - \sqrt{4K_a [E]_f [I]_f + (K_a [E]_f + K_a [I]_f + 1)^2}}{2K_a [E]_f} \tag{11}$$

The change in fluorescence depends on the amount of bound peptide, and because we are dealing with a quenching process and the initial fluorescence is set at 1, we can calculate the fluorescence value as follows:

$$F_{calc} = 1 - \nu \times \Delta F_{max} \tag{12}$$

$\Delta F_{max}$ and $K_a$ were fitted in Microsoft Excel to the observed fluorescence change using nonlinear regression (least squares method).
4. Scatchard plot analysis

Substituting equation 2 into equation 3 gives:

\[
\frac{v}{[E]} = \frac{[E]_U}{[E]_U + [E]_U} = \frac{K_a [E]_U [I]_U}{1 + K_a [I]_U},
\]

which can be rearranged to:

\[
v + vK_a [I]_U = K_a [I]_U
\]

and be divided by \([I]_\text{free}\) and rearranged into:

\[
\frac{v}{[I]_U} = K_a - K_a v
\]

This is the Scatchard equation for a 1:1 binding of ligand to protein. If there are multiple binding sites, \(n\), this equation rearranges to:

\[
\frac{v}{[I]_U} = nK_a - K_a v
\]

A plot of \(v/[I]_U\) versus \(v\) will give a straight line, with slope \(-K_a\) and \(x\)-intercept \(n\). The \(y\)-intercept is \(nK_a\).

5. References

Summary

DNA replication is a highly important and intriguing process that lies at the very heart of life. Decades of investigation have revealed that complexes comprising various proteins and enzymatic functions are responsible for faithfully copying DNA at astonishing speeds. The efficiency of replication is dramatically enhanced by clamp proteins, which are circular proteins that thread onto DNA and tether DNA polymerases to their templates. The structures of clamp proteins of various organisms from all domains of life show striking similarities, suggesting a similar mode of action. Nevertheless, variations are observed in the amount of subunits in the ring, as well as the stability of clamp complexes on DNA. The T4 bacteriophage replication system is well-described and serves as a simplified model for the mammalian DNA replication process. Like in other replication systems, the T4 clamp protein is essential for efficient DNA synthesis. The interactions between the T4 DNA polymerase and the clamp primarily take place via the open subunit interface of the clamp, which seems to be a unique feature of the T4 bacteriophage and T4-related phages. The clamp is loaded onto DNA at primer-template junctions by the clamp-loader protein complex, and is relatively unstable on DNA when not in the holoenzyme complex with the polymerase.

This thesis aims to develop processive oxidation catalysts based on the T4 clamp protein. Different strategies to obtain the proposed biohybrid oxidation systems were explored. First, two distinct procedures to link the chemical catalyst to the protein are introduced, i.e. the direct labelling and the “padlock” approach, respectively. Second, mechanisms by which the catalyst can be loaded onto DNA are identified and discussed. Furthermore, two catalysts are introduced that will be used to construct the T4 clamp-based artificial enzyme. Whereas cationic manganese porphyrins are able to specifically cut AAA sequences of DNA, footprinting reagents like FeBABE cause random oxidative damage. The interaction of the T4 clamp with synthetic polymers and a DNA plasmid lacking any clamp loader sites was investigated via ATPase assays and stopped-flow fluorescence spectroscopy measurements. Two poly(isocyanopeptides), PIAA and PIAAA, with negatively charged carboxylates on the periphery of the polymer, were proposed to mimic the physical properties of DNA and might therefore be suitable templates for the T4 clamp. Under the conditions used, i.e. in the presence of the clamp loader complex or macromolecular crowding agents, it was concluded that the clamp neither binds to the polymer templates nor to the reference DNA plasmid with significant affinity. Nevertheless, promising FRET studies suggested a weak interaction with PIAA in macromolecular crowded solutions, while a ATPase assay experiment showed interaction with PIAAA. Due to lack of time, these interactions were not studied further, but these preliminary results justify further research into the binding of clamp proteins to synthetic substrates.

Subsequently, DNA block-copolymers that consisted of a DNA and a synthetic polymer block were aimed as templates for the T4 clamp. In these polymers, the clamp may bind to the DNA block, and subsequently translocate to the synthetic block facilitating
oxidation reactions to take place on the latter. Such a strategy would overcome the problem of too weak interactions as observed for the synthetic PIAA and PIAAA polymers. Initial model studies were carried out focussing on the synthesis of DNA-PEG conjugates via the Michael addition of thiol-functionalized oligonucleotides and a commercially available PEG-maleimide polymer. Quantitative yields were obtained with 5 equivalents of polymer and any residual impurities could be removed by size exclusion chromatography. A maleimide appended polybutadiene polymer was also synthesized, but did not give rise to conjugates under the conditions tested. Other linking procedures were also tested but failed to give polybutadiene-DNA conjugates as well.

Based on these initial studies, it was concluded that the approach to oxidize synthetic polymers or DNA block co-polymers with the proposed bio-hybrid clamp catalyst was not feasible, and it was decided to focus subsequent studies on the use of DNA itself as substrate for the bio-hybrid catalyst.

The C-terminal domain of the T4 DNA polymerase is essential for interaction of the polymerase with the T4 clamp, as is the case for many other clamp-interacting proteins. In previous work described in the literature, two different binding sites on the clamp have been identified in studies with peptide analogues of the polymerase C-terminus. These peptides were found to bind to the open subunit interface and the interdomain connecting loop of the clamp. The binding affinity of the peptides for the subunit interface have been studied and determined. We have extended these studies by focussing on the influence of the presence of DNA and macromolecular crowding reagents (e.g. PEG) on the binding process, which was monitored via FRET based titration experiments. It was found that the binding of the peptide to the clamp is not affected by the addition of DNA, whereas the presence of PEG leads to a weaker interaction. Attempts were made to determine the number of binding sites for the peptide on the clamp via Scatchard plots. This led to the estimation of 2-9 binding sites per clamp trimer, which turned out to be clamp-concentration dependent. It was concluded that the interactions of the clamp with the peptide follow a very complex mechanism, involving multiple distinct binding sites of which the stoichiometry is dependent on the concentration of the clamp. The experiments described in this thesis were not extensive enough to fully elucidate this mechanism. ATPase assay measurements revealed that the peptide inhibits ATP hydrolysis both when the clamp loader-clamp complex is formed in solution and when it is bound to the DNA template. There is no experimental evidence, therefore, that the peptide can stabilize the clamp-DNA complex.

Maleimide-functionalized porphyrin derivatives were synthesized and characterized as potential catalysts that can be conjugated to the T4 clamp protein. The porphyrin was coupled to a water-soluble linker containing a maleimide group for bioconjugation. Different synthetic routes were explored, of which the Cu(I) catalyzed [3+2] cycloaddition of azides and alkynes proved to be an efficient method to link porphyrins to the spacers. The scope of this method, which led to the synthesis of a small library of novel bioconjugatable porphyrins, was explored.
The E212C clamp mutant was expressed, purified, and provided with two different chemical oxidation complexes in order to obtain clamp-based biohybrid catalysts. The FeBABE-labelled clamp protein prepared via the direct labelling approach turned out to be a rather unreactive oxidation catalyst. The same was found to be the case for the FeBABE-labelled peptide that was attached via the padlock approach. In contrast, the porphyrin-labelled clamp proteins were active as oxidation catalysts, and their action on DNA substrates was investigated by the non-specific and clamp loader-mediated binding routes. Additionally, the effects of the presence of macromolecular crowding reagents and the gp43 C-terminus analogue peptide were investigated. In the case of a supercoiled plasmid, on which the porphyrin-labelled clamp protein was allowed to bind via a non-specific binding approach, it was found that the DNA was readily oxidized to the nicked form, but that the oxidative damage was not extensive enough to yield the linear form of DNA. Based on agarose gel analysis, it was concluded that the addition of the peptide did not significantly affect the oxidation efficiency, while the addition of PEG led to a decrease in oxidative cleavage damage, probably due to the increase in viscosity of the solution. It was estimated, based on the intensities of the corresponding bands on the gel, that the average number of oxidized sites on the plasmid was 11. A novel analysis technique was developed allowing the determination of the number of oxidation events per plasmid at the single molecular level. This technique demonstrated that the porphyrin-labelled clamp protein was able to processively oxidize double stranded DNA substrates in ~15-20% of the cases. The histogram of the number of oxidation events per plasmid revealed a distribution with two maxima corresponding to plasmids with low and high amount of oxidation events per DNA chains. In the latter case, the oxidation events were found to be present in clusters, indicative of a processive oxidation mechanism.

Nicked plasmids were also oxidized in experiments in which the catalyst was loaded onto the substrate with the help of the clamp loader complex. The loading site was adapted to specifically send the clamp into the direction of a region with either a high number or a low number of oxidizable DNA sequences (AAA). Analysis of these reactions with agarose gels did not identify any reaction products, as the reaction did not lead to significant amounts of linear DNA fragments. The single-molecule analysis technique, however, convincingly demonstrated that the direction in which the clamp was allowed to diffuse correlated well with the number of oxidized AAA sites. In addition, in the case in which the plasmid was sent to a region with a high number of AAA sequences, the oxidized sites were found to appear in clusters. This was in contrast to the complementary experiment, in which the clamp was forced to move to regions with a low number of AAA sites. In this case, the clusters were absent.

Finally, the T4 replication proteins were studied with the help of AFM, in order to increase our understanding of the interactions between these proteins and between these proteins and DNA. First, the T4 clamp protein, the clamp loader complex, and the DNA polymerase were imaged individually. The complexes of the clamp loader and the clamp, as well as the DNA polymerase and the clamp, were stabilized via glutaraldehyde cross-linking, followed
by purification by FPLC and imaged with AFM. This procedure was particularly successful for the complex of the clamp with the DNA polymerase, which was visualized as a banana-shaped structure. Subsequently, all three proteins were incubated with various DNA substrates. For the clamp, the reaction conditions were optimized to facilitate binding to DNA, yielding distinct globular shapes of the clamp on the DNA structure. The DNA polymerase formed aggregates with a long DNA substrate. Unexpectedly, multiple clamp loader proteins were found to be bound to a linear plasmid, while only one weak binding site was present. Finally, efforts were made to image the complex of the clamp and the polymerase (the T4 holoenzyme) and the T4 replisome, which also includes the clamp loader complex and the single stranded DNA binding protein, on DNA substrates. In the latter case, DNA synthesis reactions were imaged in the presence of deoxyribonucleotide triphosphates.
Samenvatting

DNA-replicatie is een belangrijk proces, dat de basis vormt voor de instandhouding en vermenigvuldiging van leven. Jaren van wetenschappelijk onderzoek naar de biochemische beginselen van dit proces hebben aangetoond dat het genetisch materiaal (DNA) wordt gekopieerd door complexen van eiwitten, waarvan elk onderdeel een specifieke taak heeft. Deze complexen zijn in staat om zeer snel DNA te repliceren, zonder daarbij veel fouten te maken. De snelheid wordt vaak gestimuleerd door de aanwezigheid van zogenaamde ‘clamp’ of klem-eiwitten, die zich om het DNA heen vouwen en zorgen dat het onderdeel dat het DNA kopieert, het DNA-polymerase, niet loslaat zolang het replicatieproces nog niet voltooid is. Deze klem-eiwitten komen bij veel verschillende organismen voor en hebben een vergelijkbare ringvormige structuur. Toch zijn er subtiele verschillen zoals het aantal subeenheden waaruit de ring bestaat en de stabiliteit van het complex van de ring en DNA.

Het DNA-replicatiesysteem van de T4-bacteriofaag is zeer goed bestudeerd en wordt gebruikt als een model om het veel ingewikkeldere DNA-replicatiesysteem bij zoogdieren te begrijpen. Het T4-klem-eiwit is essentieel voor efficiënte DNA-replicatie. De interacties tussen het T4-klem-eiwit en het DNA-polymerase vinden hoofdzakelijk plaats via een proces dat uniek is voor de T4-bacteriofaag en gerelateerde fagen, namelijk via de open ruimte van het hoefijzervormige klem-eiwit. Het klem-eiwit zelf wordt geladen op zogenaamde primer-template uiteinden van het DNA door het klem-laad-eiwitcomplex, tenzij het deel uitmaakt van een complex met de polymerase. Het complex van DNA-polymerase en klem-eiwit wordt het holoenzym complex genoemd.

Het onderzoek dat beschreven is in dit proefschrift heeft tot doel om zogenaamde processieve oxidatiekatalysatoren te ontwikkelen die gebaseerd zijn op het T4-klem-eiwit. Verschillende methodes om dit doel te bereiken worden eerst geïntroduceerd, namelijk het koppelen van de chemische katalysatoren via covalente bindingen aan het eiwit zelf en het bewerkstelligen van de koppeling door gebruik te maken van de specifieke interacties tussen de polymerase en het klem-eiwit. Verschillende manieren worden besproken waarop het beoogde klem-eiwit-katalysator-conjugaat op het substraat gebonden kan worden. Voorts worden twee verschillende katalysatoren beschreven, die gebruikt zijn om het kunstmatige enzym op basis van het T4-klem-eiwit te construeren. Deze katalysatoren zijn geselecteerd om het DNA te knippen, hetzij op specifieke locaties, hetzij op willekeurige plaatsen.

Om te onderzoeken of de beoogde klem-eiwit-katalysatoren ook gebruikt kunnen worden om synthetische polymeersubstraten te oxideren, is de interactie tussen het T4-klem-eiwit en verschillende synthetische polymeren - alsmede een DNA-plasmide zonder de natuurlijke locaties waarop het eiwit geladen kan worden - bestudeerd door middel van ‘ATPase-assays’ en ‘stopped-flow’ fluorescentiemetingen. Voor deze studies zijn twee poly(isocyanopeptiden) (PIAA en PIAAA) geselecteerd omdat zij een rigide structuur en negatieve lading bezitten. Door deze eigenschappen werden zij verondersteld de fysieke eigenschappen van het DNA te benaderen, waardoor ze geschikte substraten zouden.
kunnen zijn voor het klem-eiwit. Zowel in de aanwezigheid van het klem-laad-eiwit als in
oplossingen die de binding van het klem-eiwit stimuleren (door middel van zogenaamde
‘macromolecular crowding’ agentia) vond er geen significante interactie plaats tussen het
klem-eiwit en het polymeer of het klem-eiwit en het DNA-substraat. Desalniettemin werd er
een zwakke interactie gevonden voor PIAA in de oplossing die binding stimuleert, terwijl de
ATPase-assay tevens een zwakke, nog niet helemaal bevestigde interactie liet zien met
PIAAA. Deze verkennende studies tonen aan dat verder onderzoek noodzakelijk is om de
interacties tussen het klem-eiwit en negatief geladen polymeersubstraten te begrijpen.

Vervolgens werd getracht DNA-blok-copolymeren te synthetiseren, bestaande uit een
dNA- en een polymeergedeelte. Hiermee zou het klem-eiwit op het DNA-gedeelte geladen
c kunnen worden om zich vervolgens naar het polymeergedeelte te verplaatsen, waardoor er
op het laatste een oxidatiereactie plaats zou kunnen vinden. Ondanks de zwakke interacties
van het klem-eiwit de met negatief geladen synthetische polymeer zoals PIA en PIAAA,
zouden op deze wijze deze synthetische substraten toch geoxideerd kunnen worden. Als
modelsysteem werd gekozen voor DNA-PEG-polymeren, die via een Michael-additie van
thiol-gefunctionaliseerde oligonucleotides en een commercieel verkrijgbare PEG-maleimide-
polymer gemaakt kunnen worden. Kwantitatieve opbrengsten konden worden gehaald
door vijfvoudige overmaat van het polymeer te gebruiken. De zuivering van het product kon
worden bewerkstelligd met behulp van ‘size-exclusion’-chromatografie. Een maleimide-
gefunctionaliseerd polybutadienopolymer werd ook gesynthetiseerd, maar dit polymeer
kon onder de gebruikte condities niet gekoppeld worden met het bovengenoemde DNA-
fragment. Daarnaast werden andere koppelingsroutes onderzocht, die uiteindelijk niet
leidden tot de succesvolle ontwikkeling van polybutadien-DNA-conjugaten.

Op basis van de bovengenoemde initiële studies werd de conclusie getrokken dat het
oxideren van synthetische polymeer of DNA-blokopolymeren door middel van de
voorgestelde klem-eiwit-katalysator-hybrides niet haalbaar was. Daarom werd besloten de
vervolgstudies te richtten op het gebruik van DNA als substraat voor de biohybride
katalysator.

Zoals het geval is bij andere eiwitten die interacties aangaan met klem-eiwitten, is het
C-terminusdomein van de T4-bacteriofaag essentieel voor de interactie van de polymerase
met het T4-klem-eiwit. Studies met peptideanalogen van de C-terminus van de polymerase
hebben aangetoond dat deze peptiden op twee verschillende plaatsen aan het T4-klem-eiwit
c kunnen binden: enerzijds in de ruimte tussen de open subeenheden en anderzijds aan de
zogenaamde interdomein-verbindingbrug op het midden van de subeenheid. De
bindingsaffiniteit met betrekking tot de eerste interactie is reeds bepaald in eerder onderzoek
dat is beschreven in de literatuur. In dit proefschrift hebben wij deze studies verder
uitgebreid door de invloed van de aanwezigheid van DNA en van ‘macromolecular
crowding’ condities op het bindingsproces te bestuderen met behulp van zogenaamde FRET-
titratie-experimenten. Uit deze experimenten bleek dat de binding van de peptiden aan het
klem-eiwit niet wordt beïnvloed door de aanwezigheid van DNA, terwijl het toevoegen van
‘macromolecular crowding’-verbinding PEG leidde tot een zwakkere interactie. Voorts
werden pogingen ondernomen om het aantal bindingsplaatsen voor het peptide op het klem-eiwit te bepalen met behulp van zogenaamde Scatchard-curves. Afhankelijk van de concentratie van het klem-eiwit kon een schatting van 2 tot 9 bindingsplaatsen per klem-eiwit-trimeer gemaakt worden. Geconcludeerd werd dat de interacties van het peptide met het klem-eiwit complex zijn, waarbij - afhankelijk van de concentratie van het klem-eiwit - meerdere bindingssites een rol spelen. Helaas konden de uitgevoerde experimenten dit proces niet geheel ophelderen. ATPase-assayexperimenten lieten zien dat het peptide het complex van het klem-ladereiwit en het klem-eiwit niet alleen in oplossing maar ook op een DNA-substraat kan verhinderen, waardoor er geen experimenteel bewijs gevonden kon worden dat het peptide het klem-eiwitcomplex op DNA kan stabiliseren.

Als onderdeel van het onderzoek werden verder maleimide-gefunctionaliseerde porphyrinederivaten gesynthetiseerd en gekarakteriseerd met als doel deze vast te maken aan het T4-klem-eiwit. Het porphyrine werd gekoppeld via een wateroplosbaar verbindingssluk dat tevens een maleimidegroep bevat die gebruikt kan worden om de koppelingsreactie uit te voeren. Verschillende synthetische routes werden onderzocht, waaruit bleek dat de koper(I) gekatalyseerde [3+2] cycloaddities van azides aan alkynen een efficiënte methode is om porphyrines aan het maleimideverbindingsslu te koppelen. De algemene toepasbaarheid van deze methode is onderzocht hetgeen heeft geleid heeft tot de synthese van een kleine bibliothek van niet eerder verkregen bio-conjugeerbare porphyrines.

Een E212C-mutant van het T4-klem-eiwit werd tot expressie gebracht, gezuiverd, en voorzien van twee verschillende chemische oxidatiekatalysatoren teneinde biohybride klem-eiwitkatalysatoren te verkrijgen. Het klem-eiwit waar het FeBABE complex direct aan het eiwit gekoppeld was, bleek geen efficiënte katalysator te zijn. Dit was ook het geval voor het klem-eiwit waaraan het FeBABE complex via een peptide gekoppeld was. Dit peptide had eenzelfde sequentie als de C-terminus van het T4-DNA-polymerase. Daarentegen was het klem-eiwit voorzien van een porphyrine wel actief als oxidatiekatalysator. Deze activiteit is vervolgens uitvoerig bestudeerd door middel van oxidatie-experimenten met verschillende substraten gebaseerd op DNA-plasmiden. Het klem-eiwit is op deze DNA-substraten vastgezet via niet-specifieke interacties, alsook met behulp van het klem-laad-eiwitcomplex. Tevens werd de invloed van de aanwezigheid van PEG en het peptide op de activiteit van de katalysator bepaald. Wanneer een zogenaamde ‘supercoiled’ plasmide werd gebruikt, en het klem-eiwit geladen werd door middel van niet-specifieke interacties, werd het plasmide snel geoxideerd tot zijn zogenaamde ‘genickte’ vorm. De oxidatieve schade was echter niet zodanig dat er lineaire fragmenten gevormd werden. De reactieproducten werden geanalyseerd met behulp van agarosegel-electroforese, waaruit bleek dat de toevoeging van het peptide aan het reactiemengsel geen signifieante invloed had op de efficiëntie van de reactie, terwijl de toevoeging van PEG leidde tot een vermindering van de oxidatieve schade. Dit laatste werd waarschijnlijk veroorzaakt door een toename van de viscositeit van de oplossing. De hoeveelheid geoxideerde plaatsen op het plasmide werd geschat op 11 en wel
aan de hand van de intensiteiten van de verschillende reactieproducten op de agarose gel. Om de hoeveelheid oxidatiereacties per plasmide meer in detail te bestuderen, werd een nieuwe analysetechniek ontwikkeld, waarmee deze hoeveelheid met moleculaire resolutie bepaald kan worden. Deze techniek liet zien dat in ~15-20% van de gevallen het substraat processief geoxideerd werd, hetgeen betekent dat meerdere plaatsen op het DNA achtereenvolgens omgezet worden. Onder de condities waarbij het peptide niet aan het reactiemengsel was toegevoegd, konden reactiemechanismen worden onderscheiden die respectievelijk weinig of veel oxidatieschade per plasmide opleverden. In het laatste geval bleken de oxidatiereacties voor te komen in clusters, die alleen maar gevormd konden zijn via een processief reactiemechanisme.

Naast de supercoiled plasmiden werden ook genickte plasmides onderzocht als substraat in de oxidatiereactie. Hierbij werd de aan het klem-eiwit gekoppelde katalysator op het DNA geplaatst met behulp van het klem-laad-eiwit. De locatie waar het katalysatoreiwit geladen werd, was zo gekozen dat de katalysator ofwel in de richting van veel oxideerbare plaatsen, danwel in de richting van weinig oxideerbare plaatsen gestuurd werd. Analyse van deze experimenten met behulp van agarosegels lieten geen vorming van reactieproducten zien, omdat de oxidatiereactie niet leidde tot een significante toename in lineaire DNA-fragmenten. Met de nieuwe analysetechniek werden de reactieproducten echter ook op single molecule niveau geanalyseerd, waarmee overtuigend kon worden aangetoond dat de richting waarin het klem-eiwit wordt gestuurd bepalend is voor de hoeveelheid oxidatiereacties die plaatsvindt op het plasmide. Wanneer de bio-hybride katalysator gedwongen werd zich te verplaatsen naar een gebied met veel oxideerbare DNA-sequenties (AAA), werd een relatief grote hoeveelheid geoxideerde plaatsen op het DNA waargenomen, die vaak ook in clusters voorkwamen. Dit was in tegenstelling tot de reacties waarbij de katalysator zich alleen naar de tegenovergestelde regio met weinig oxideerbare DNA-sequenties kon verplaatsen. Hier werden geen clusters van geoxideerde plaatsen gevonden en het aantal oxidatiereacties was lager.

maar één enkele zwakke bindingsplaats voor het eiwit aanwezig was. Tenslotte werd getracht om het complex van het klem-eiwit en het DNA-polymerase (het zogenaamde T4-holoenzyme) en het T4-replisome, waarin ook het klem-laad-eiwit en het enkelstrengs-DNA-bindend-eiwit aanwezig zijn, op DNA zichtbaar te maken. In het laatste geval werd de synthese van DNA bestudeerd in de aanwezigheid van deoxyribonucleotidetrifosfaten. Als functie van de tijd waren hierbij langer wordende DNA-strengen zichtbaar, hetgeen de mogelijkheid opent om DNA-synthese met behulp van AFM te bestuderen.
Jawel, beste lezer, dit proefschrift is bijna uit. Alhoewel, voor de meesten begint het lezen van een proefschrift met het dankwoord, dus tegen hen zou ik willen zeggen: veel succes, want je hebt nog heel wat voor de boeg! Dat is tenminste mijn gevoel, als ik terugkijk op de maanden die ik aan het schrijven van dit boekje heb besteed. Voor het dankwoord moeten we echter nog verder teruggaan, namelijk de dikke vier jaar die als basis dienen voor dit schrijven!

In oktober 2002 begon ik met mijn promotie, in de groep van Prof. Roeland Nolte en Prof. Alan Rowan. Roeland, ik wil je van harte bedankt voor het vertrouwen dat je in mij had om mij aan te nemen, en dit op het einde van mijn promotietijd een tweede keer te doen. Ondanks dat je het vaak (veel te) druk had, nam je toch altijd de tijd voor een goed gesprek over het onderzoek. Ik kan me nog goed jouw peptalk herinneren een paar dagen voor kerst in een koud en verlaten Huijgensgebouw. Helaas heb je toen niet meteen gelijk gekregen met je optimisme, maar uiteindelijk is het toch nog goed gekomen. Ik hoop dat je in de komende jaren tijd gaat krijgen om alleen de leuke taken te doen en de rest te laten liggen.

Alan, begeleid worden door jou was niet altijd even makkelijk. Terugkijkend doet je kantoor mij denken aan het kasteel in Das Schloß van Kafka: het is soms moeilijk om een afspraak met je te maken, en normale communicatiemiddelen zoals telefoons, e-mails etc. lijken niet te werken in jouw domein. Ook geldt er een apart soort definitie van tijd (twee minuutjes lijken vaak twee uur). Toch is jouw kantoor een walhalla van inspiratie. Onderzoeksproblemen worden opzij geschoven en er is volop ruimte voor nieuwe ideeën en persoonlijke motivatie. Dit laatste heb ik altijd als zeer prettig ervaren, en ik wil je dan ook succes wensen in je verdere carrière als hoogleraar.

Jeroen, jij was aangesteld als mijn dagelijkse begeleider tijdens het laatste gedeelte van mijn promotie. Ik vond het leuk om af en toe bij je binnen te stappen en je op de hoogte te houden van de vorderingen. Ik wens ook jou veel succes toe in de toekomst.

De manuscriptcommissie (prof. Ger Pruijn, prof. Andreas Herrmann en Dr. Maarten Merkx) wil ik bedanken voor het bestuderen van mijn proefschrift en voor hun nuttige suggesties. Voorts wil ik de topschool NRSC-C bedanken voor het financieren van mijn onderzoek en in het bijzonder Dr. Ad Kolen voor zijn hulp bij het realiseren van mijn onderzoeksreis naar Amerika.

I could not have done my PhD without the cooperation and support of Professor Stephen Benkovic. Steve, thanks for your support during the last years, and for having me in your lab. I still remember your words “we should make this work for your thesis” one of the first times we met back in 2002. Indeed, it took a considerable amount of work and time, but I’m happy that we succeeded in the end. I’m also greatly indebted for the help of some of the Benkovic lab members. First of all, Dr. Michael Trakselis, who has been involved in this research from the start and introduced me into the world of clamp proteins. Michael, I wish you all the best with your research group in Pittsburgh and great scientific successes in the future. Furthermore, there’s Dr. Michelle Spiering who has been my main target for all the
updates and troubleshooting requests, and Dr. Zhihao Zhuang who supported me in the Benkovic lab. My stay at Penn State was a great success, for which I would also like to thank the other Benkovites with special mentions to Mike, Jun, Ali, and Scott.

Wat het lab tot zo’n speciale plaats maakt, is de mix van mensen die er in hun reactiekolfjes roeren. Om niemand tekort te doen, wil ik iedereen van de organisch chemische familie (ik bedoel moleculair chemische clustergenoten) die heeft bijgedragen aan de gezelligheid op het lab bedanken voor de leuke tijd. Toch ontkom ik er natuurlijk niet aan om sommigen bij naam te noemen. Ten eerste zijn er natuurlijk de directe labgenoten, begonnen in het UL (Pieter, Dennis, Matthijs, Irene) en later in iets andere samenstelling in het Huijgensgebouw (Marta, Mark, Linda, Marco, Nikos, Friso). Ook heb ik vele Spaanse studenten voorbij zien komen Van hen wil ik Raquel, Ribera (uit La Pobla de Segur) en Maria Marza bedanken voor hun gezelligheidsverhogende aanwezigheid. Ik ben blij dat ik tijdens de promotie bijgestaan wordt door Marta en Mark. Marta, het was fijn om je als buurvrouw te hebben in ons rustige cubicle rijtje en op het lab. De aanwezigheid van Mark op het lab heeft geleid tot het ontstaan van verscheidene nutteloze maar vermakelijke ‘labsporten’ die een aangename afleiding waren van het kolomwerk. Ik moet alleen niet vergeten je af en toe op je plek te zetten tijdens het squashen!

Gelukkig stond ik er in het lab niet helemaal alleen voor, en ben ik bijgestaan door twee studenten. Als eerste was daar Maria Blasco Castells. Maria, I owe you so much for laying the foundations for the synthetic work described in this thesis. After a period in which pursued reaction pathways all turned out to be dead-end streets, you enthusiastically started the work on click-reactions with porphyrins. What was at that time just an idea became the rescue of my thesis. I think that initially you had some problems with being away from home, but in the end I’m sure that you are will look back with fond memories and great pleasure. I especially liked to see how you gained confidence during your stay in Nijmegen, and I hope that you still enjoy this in your present job and life. ¡Muchos gracias!

Dan was daar ook nog Dennis van de Mosselaar, die zijn afstudeerstage aan het HLO kwam doen en voor mij en Irene heel nuttig werk heeft gedaan. Dennis, je benaderde de ideale stage-student: heel enthousiast en zo zelfstandig dat ik je rustig alleen kon laten werken. Ondanks het feit dat je te maken had met een van de lastigste verbindingen op aarde (protoporphyrines), was je chemische strijdlust altijd aanwezig en heb je zwaarbevochten resultaten geboekt! Jammer genoeg had ik niet veel tijd meer om wat met je verbindingen te doen. Ik ben ook blij dat je mij geholpen hebt met het maken van de belangrijke uitgangsverbindingen voor de stofjes die ik in Amerika nodig had. Ik weet zeker dat je bij Mercachem op je plaats bent en ik wens je daar dan nog veel (chemisch) werkplezier toe!

Peter Schön and Kasper Nørgaard, thanks for your experimental help with the AFM work. We had to put a lot of effort into the experiments, but in the end it laid the foundations for the greatest results of my thesis. Special thanks to Kasper, for playing around with the image rendering software and producing fine pictures along the way!
Tijdens mijn promotie ben ik regelmatig naar de meer bio-gerichte omgeving van het NCMLS gevlecht voor proefjes en raad. Ik wil Prof. Henk Stunnenberg, Siebe van Genese, Carla Schwartz, Prof. Ger Pruijn, Dr. Frank Wagener, en Dr. Mark Lambermon (ook vaak dichterbij te vinden) bedanken voor de discussies en experimentele hulp. In het Huijgensgebouw ben ik veel dank verschuldigd aan Rien van der Gaag voor de FPLC-ondersteuning, Frank Nelissen voor hulp bij de agarosegels, Dr. Richard Feron bij het fotograferen van de agarosegels, en Hans Adams voor hulp bij het maken van de peptiden. Dr. Huub op den Camp wil ik bedanken voor de Maldi-TOF-ondersteuning en Helene, Peter en Ad voor het analyses en hulp met de NMR experimenten. En natuurlijk mag ik Désirée en Peter van Dijk niet vergeten, bedankt voor alle ondersteuning!

Gelukkig was er naast het werk op het lab ook ruimte voor ontspanning. Paul en Jaap, Irene en Matthijs, Ton (en Susanne), bedankt voor de gezelligheid en de etentjes. Paul in het bijzonder, vanwege de wekelijkse fitness-sessies die vaak uitmonden in grondige analyses van de afdeling en het doornemen van de laatste roddels. Daarnaast was er natuurlijk de voetbalclub, de Californië-reisgenoten (Hans – ook bedankt voor alle porphyrine-ondersteuning – Johan, Richard, Joost O), en de zondagochtend squash-fanatici. Ook mijn vrienden van de middelbare school zijn onmisbaar om het chemische wereldje soms even te ontvluchten. Luc, Marijn, Jasper, ik waardeer het dat we ondanks het feit dat we elk onze eigen weg zijn gegaan, nog steeds onze vriendschap koesteren. Bedankt voor alle uurtjes Gonzo, ik bedoel NDRGRND.

Het afgelopen jaar (bijna) heb ik een fijne tijd gehad bij bioMérieux. Ik wil hiervoor al mijn collega’s daar bedanken! Speciale dank is er voor kamergenoot Peter, die als microbioloog en zelfaangewezen vrijwilliger nog wat hoofdstukken van mijn proefschrift heeft nagelezen op fouten. Helaas ga ik jullie alweer verlaten, en ik wens jullie het beste voor de toekomst. Zoals ze zeggen, in een land niet ver van hier: “C’est la vie!”.

Mijn familie en schoonfamilie wil ik bedanken voor hun interesse en ondersteuning. Jullie zullen niet altijd begrepen hebben waar ik die vier jaar aan besteed heb, toch is het fijn dat jullie interesse hebben getoond in mijn gefröbel op de universiteit.

Het laatste – en belangrijkste - woord is voor Annemiek. Door de jaren heen heb je de frustraties van het academisch onderzoek (en het schrijven van een proefschrift) met mij moeten delen. Veel te vaak heb jij je opgeofferd zodat ik goed kon werken en schrijven. Gelukkig dwong je me ook om af en toe het werk te relativeren en wat afstand te nemen van mijn promotie. Bedankt dat je mij gesteund hebt en voor alle jaren die nog komen gaan.

Joost
Joost Clerx was born on January 15\textsuperscript{th} 1979 in Nijmegen, The Netherlands. He attended Dominicus College secondary school, which he finished in 1997. He subsequently started his chemistry study at the Radboud University Nijmegen, where he obtained his M.Sc. degree in 2002, graduating \textit{cum laude}. He did his major in Organic Chemistry, under the supervision of Prof. Dr. R.J.M. Nolte, studying the synthesis and applications of divalent polymer-functionalized biotin ligands for the construction of macromolecular bio-hybrid assemblies. His minor involved a close collaboration with the Molecular Biology Department of Prof. Dr. H.G. Stunnenberg and the Bio-organic Chemistry Department of Prof. Dr. J.C.M. van Hest on the cloning and expression of \textit{de novo} designed $\beta$-sheet and $\alpha$-helical proteins inspired from structural proteins in Nature.

In October 2002, he started his Ph.D. project as junior researcher in the Physical Organic and Supramolecular Chemistry group under supervision of Prof. Dr. A.E. Rowan and Prof. Dr. R.J.M. Nolte. In this project, he studied the construction and oxidation properties of catalyst-labelled T4 clamp proteins. The results thereof are presented in this thesis.

From June 2007 to April 2008, Joost was employed at bioMérieux BV as a Senior Scientist. Starting from May 2008, he will move to Organon, part of Schering-Plough corporation, to become a project manager in the Biotechnology Department.