DNA Oligomers - From Protein Binding To Probabilistic Modelling

DISSERTATION

zur Erlangung des akademischen Grades

Doctor of Philosophy

(Ph.D.)

vorgelegt

der Fakultät Mathematik und Naturwissenschaften

der Technischen Universität Dresden

von

Maria Helena Azevedo Teixeira de Andrade

geboren am 04. März 1986 in

Braga, Portugal

Eingereicht am

Die Dissertation wurde in der Zeit von

Mai 2013 bis Mai 2016

im B CUBE - Center for Molecular Bioengineering,

Technische Universität Dresden angefertigt.

Abstract

This dissertation focuses on rationalised DNA design as a tool for the discovery and development of new therapeutic entities, as well as understanding the biological function of DNA beyond the storage of genetic information. The study is comprised of two main areas of study: (i) the use of DNA as a coding unit to illustrate the relationship between code-diversity and dynamics of self-assembly; and (ii) the use of DNA as an active unit that interacts and regulates a target protein.

In the study of DNA as a coding unit in code-diversity and dynamics of self-assembly, we developed the DNA-Based Diversity Modelling and Analysis (DDMA) method. Using Polymerase Chain Reaction (PCR) and Real Time Polymerase Chain Reaction (RT-PCR), we studied the diversity and evolution of synthetic oligonucleotide populations. The manipulation of critical conditions, with monitoring and interpretation of their effects, lead to understanding how PCR amplification unfolding could reshape a population. This new take on an old technology has great value for the study of: (a) code-diversity, convenient in a DNA-based selection method, so semi-quantitation can evaluate a selection development and the population's behaviour can indicate the quality; (b) self-assembly dynamics, for the simulation of a real evolution, emulating a society where selective pressures direct the population's adaptation; and (c) development of high-entropy DNA structures, in order to understand how similar unspecific DNA structures are formed in certain pathologies, such as in auto-immune diseases.

To explore DNA as an active unit in Tumour Necrosis Factor α (TNF-α) inter-action and activity modulation, we investigate DNA's influence on its spatial conformation by physical environment regulation. Active TNF-α is a trimer and the protein-protein interactions between its monomers are a promising target for drug development. It has been hypothesised that TNF-α forms a very intricate network after its activation between its subunits and receptors, but the mechanism is still not completely clear. During our research, we estimate the non-specific DNA binding to TNF-α in the low micro-molar range. Cell toxicity assays confirm this interaction, where DNA consistently enhances TNF-α's cytotoxic effect. Further binding and structural studies lead to the same conclusion that DNA binds and interferes with TNF-α structure. From this protein-DNA interaction study, a new set of tools to regulate TNF-α's biological activity can be developed and its own biology can be unveiled.

To Eli and HeeChang.

Contents

List of Tables

List of Figures

Acronyms

Ab antibody.

ACN acetonitrile.

ACT-D Actinomycin D.

AFM Atomic Force Microscopy.

Ag antigen.

AIDS Acquired Immune Deficiency Syndrome.

APC Antigen Presenting Cells.

APS aminopropylsilane.

ARG amine reactive second generation.

AS-PCR Allele-Specific PCR.

AsyPCR Asymmetric PCR.

BLI Bio-Layer Interferometry.

BSA Bovine Serum Albumin.

CD Circular Dichroism.

cfDNA cell-free DNA.

CHX Cycloheximide.

CpG Cytidine phosphate Guanosine.

CV Cell viability.

dAdA D-Ala-D-Ala.

DCC Dynamic Combinatorial Chemistry.

DCM dichloromethane.

DDMA DNA-Based Diversity Modelling and Analysis.

DECL DNA-Encoded Chemical Libraries.

DIC *N,N'*-diisopropylcarbodiimide.

DIPEA *N,N*-diisopropylethylamine.

DiStRO Diversity Standard of Random Oligonucleotides.

DMEM Dulbecco's Modified Eagle's Medium.

DMF dimethylformamide.

DMSO dimethyl sulfoxide.

DNA Deoxyribonucleic acid.

dNTPs deoxynucleotides.

dsAPT double-stranded aptamer library.

dsDNA double-stranded DNA.

EDC 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide.

EDTA ethylenediaminetetraacetic acid.

eq. equivalents.

ESAC Encoded Self-Assembling Chemical.

ESI-MS Electrospray Ionisation Mass Spectroscopy.

EtBr Ethidium Bromide.

F functional factor.

Fab fragment-antigen binding.

FBS Fetal Bovine Serum.

Fc fragment constant region.

FCS Fluorescence Correlation Spectroscopy.

full-dsDNA two fully-complementary strands.

half-dsDNA two half-complementary strands.

HATU 1-[bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate.

HOAT 1-hydroxy-7-azabenzotriazole.

HPLC High Performance Liquid Chromatography.

HTS highthroughput screening.

IFN-α Interferon α.

IgG immunoglobulin G.

IL interleukins.

LPS lipopolysaccharide.

MHC major histocompatibility complex.

MS Mass Spectrometry.

MST Microscale Thermophoresis.

MTT thiazolyl blue tetrazolium bromide.

NF-κB Nuclear Factor-κB.

NHS N-hydroxysuccinimide.

NK cells Natural Killer cells.

NVOC-Cl 4,5-dimethoxy-2-nitrobenzyl chloroformate.

OD260 optical density at 260 nm.

PAGE Polyacrylamide Gel Electrophoresis.

PBS Phosphate Buffered Saline.

PCR Polymerase Chain Reaction.

PEG polyethylene glycol.

PS phosphorothioate.

PTFE polytetrafluoroethylene.

PVDF polyvinylidenefluoride.

RFU recorded fluorescence.

RNA Ribonucleic acid.

ROX carboxy-X-rhodamine.

RT-PCR Real Time Polymerase Chain Reaction.

S-DNA Salmon Sperm DNA.

SA streptavidin.

SELEX Systematic Evolution of Ligands by Exponential Enrichment.

SLE Systemic Lupus Erythematosus.

SNHS N-hydroxysulfosuccinimide.

SNP Single Nucleotide Polymorphism.

SpdCl spermidine trihydrochloride.

ssDNA single-stranded DNA.

TACE TNF-α converting enzyme.

TBE Tris/Borate/EDTA.

TEA triethylamine.

TEAA triethylammonium acetate.

TEAB triethylammonium bicarbonate.

TFA trifluoroacetic acid.

Th1 cells T helper 1 cells.

Th2 cells T helper 2 cells.

THF tetrahydrofuran.

TLR Toll-like receptor.

TNF-α Tumour Necrosis Factor α.

TNFR1 TNF-α receptor 1.

TNFR2 TNF-α receptor 2.

Tris tris(hydroxymethyl)aminomethane.

Trp tryptophan.

UHPLC Ultra-High Performance Liquid Chromatography.

VEGF Vascular Endothelial Growth Factor.

Xsmall small Holliday junction.

yDNA y-shape DNA.

Acknowledgements

I would like to thank Prof. Dr. Yixin Zhang and Prof. Dr. Carsten Werner for giving me the opportunity to work on this PhD project. In particular, I thank Prof. Dr. Yixin Zhang for giving me the opportunity to join his group and for granting me the freedom to explore the subjects I was most curious about. Also, I would like to acknowledge his mentorship, availability, and kindness. I would like to thank Prof. Dr. Carsten Werner too for his guidance and ever most useful discussions. To both, I want to thank for being part of my Thesis Advisory Committee (TAC), reviewing my dissertation, and being my examiners.

I want to acknowledge and thank the Dresden International PhD Program (DIPP), for their fantastic job that starts by finding the students the right project and ends by helping them to finish the whole PhD procedure. I was fortunate to participate in several courses and events provided by them, and for that I thank them also. Likewise, the organisation of the TAC members and follow-up evaluations was of the upmost importance for me to finish these projects, and to grow my scientific abilities. As a third part of my TAC, I want to thank Prof. Dr. Karsten Kretschmer for being always available for helpful exchanges. Also, I want to thank Dr. Michael Schlierf for participating in my last TAC evaluation, and for the valuable input and new ideas he gave me.

I would like to thank Prof. Dr. Stefan Diez and Prof. Dr. Michael Schroeder for being part of my Doctorate Commission as my examiners.

1

Introduction

1.1 Project Origin and Unexpected Observations

Rational drug design was the origin of this project. Multivalent interaction exemplified by, for example, Rao et al. paper was an inspiration for designing and improving a molecule that could, one day, be used in therapy. In this paper, the development of a trivalent system using vancomycin and D-Ala-D-Ala (dAdA) generated an affinity 25 times higher than biotin and avidin. Plus, this polyvalent system has the possibility of being further modulated [1]. This was our aim, to improve molecules already known, integrating them in a complex that would leave space for modulation.

Our initial challenge was to merge two incredible powerful DNA-based technologies: DNA-Encoded Chemical Libraries (DECL) and Systematic Evolution of Ligands by Exponential Enrichment (SELEX). The goal was clear, we would start by improving the affinity of a known ligand with an aptamer selection. An aptamer library with a large theoretical number of unique sequences, more than one trillion, would hybridise with oligonucleotides carrying a small molecule. In an effort to further optimise the aptamer selection, we used oligonucleotides with shorter hybridisation domains,

recreating the concept of dynamic combination of pharmacophores [2, 3]. This would enable the best fits to be recognised and newly generated in a large pool of possible ligands with varying affinities, leading to a more robust affinity maturation.

The choice of DNA as the moiety to modulate the affinity arose from its hydrophilicity, since active sites are usually quite hydrophobic [4]. The improved balance between hydrophobicity/hydrophilicity could bring us small molecules hybrids with improved pharmacokinetic characteristics. Also, the possibility to generate large DNA aptamers populations, by chemical synthesis, and to have a straightforward final sequencing step, to decode the pool of interest, were huge advantages.

The targets we chose were proteins involved in immunity, such as cyclophilins, interleukins (IL), and TNF-α. We started with cyclophilins, which are a family of proteins with several isoforms that can be inhibited by the immunosuppressive drug cyclosporin and its derivatives. We wanted to develop hybrid molecules able to distinguish isoforms, so a more desirable therapy with less side effects could be achieved. From this family, we focused on cyclophilin A and cyclosporin A, which have affinity in the low nano-molar.

TNF-α and IL were also interesting targets due to the lack of small molecule therapeutic options. Nowadays, only antibody (Ab)-based therapies are available. With the surge of autoimmune diseases in our world, it seems insufficient the arsenal we have to control and inhibit these targets. We selected TNF-α because it is a protein related and central to a large range of immune diseases. Given its trimeric conformation TNF-α, we hypothesised that it could be helpful to develop a better molecule if the three subunits active pockets would be explored or if the protein-protein interactions among monomers were disrupted.

These were the starting points of our project. The initial experiments were quite promising. The SELEX protocol worked and, unlike most, was a solution-based selection, without protein immobilisation. Also, the preliminary activity assays of the

hybrid molecule were interesting and the isoform specificity of the hybrid seemed possible to achieve.

However, one day while checking the DNA pool after a selection round using RT-PCR, we met an unexpected result. The amplification curves were bell-shaped, instead of the classical sigmoidal-shape. After thorough investigation, we understood it was not a mistake/contamination, but an interesting phenomenon that could be manipulated. This effect had been observed before, but was never explored. The study of this observation originated the project "DDMA".

In parallel, we continued with the main plan for the search and development of an improved molecule for immunological disease management. This time, the repeated observation of unexpected protein-DNA interaction awoke our inquiring minds. The further investigation of these unspecific physical interactions between DNA and TNF-α translated into evident increased biological activity and structural changes. We did not let this other serendipitous event pass in vain and the second half of this PhD project was born, "Physical Environment Regulation of TNF-α Activity".

Both projects started as a study of DNA and its physical-chemical properties. They evolved into the study of DNA's ability to modulate protein-protein interactions and DNA's ability to self-assemble according to pairing rules in a highly heterogeneous population.

As oligonucleotides are no longer seen as static vessels of genetic information, our work becomes more relevant under today's biotechnological developments. DNA is becoming an exemplary tool in coding non-genetic information and assembling into nanostructures.

1.2 Deoxyribonucleic Acid

1.2.1 Chemical and Physical Properties

Deoxyribonucleic acid (DNA) was first described in 1953 by Watson and Crick [5]. DNA is a deoxyoligonucleotide polymer with two antiparallel helical strands organised around one axis. The monomers are linked by covalent bonds between the 3'-end sugar and the 5'-end phosphate, the phosphodiester bonds, and form the single-stranded DNA (ssDNA). Each monomer has a nucleobase, which can be purinic (adenine (A), guanine (G)) or pyrimidinic (cytosine (C), thymine (T)), a deoxyribose, and a phosphate group. These have strict pairing rules of purine with pyrimidine, A-T and C-G, and hydrogen bonds between them in two different chains are established, generating double-stranded DNA (dsDNA).

The dsDNA forms major and minor grooves, which are determined by the distance between the nearest phosphorus atoms across the chain. A macromolecule binding to DNA usually occurs in the minor groove and shifts its conformation. The DNA molecule's torsional rigidity increases with small twisting and decreases with untwisting [6]. DNA has a diameter around 2 nm, with a structural repeat around 3.5 nm, and persistence length around 50 nm [7]. It is packed around histones, becoming highly condensed, and forming nucleosomes.

DNA interacts with other proteins at various levels. For instance, proteins regulate DNA translation, replication and transcription, which, in turn, generates proteins that participate in molecular events [8, 9].

DNA is a great biotechnological tool due to its characteristics, ease of use, and possibility to chemically synthesise in a relatively inexpensive and easy way. To start with its basic function, DNA is able to support an extremely large amount of information, e.g., cells can store one to 2 GB of data in 7 pg. Modifying the classical nucleobases expands this ability even further. For example, in vivo, nucleobases can

be modified directly by methylation, which modulates gene expression, as studied by the emerging field of epigenetics. Another very attractive feature of DNA is the possibility to be operated at the molecular level by a plethora of enzymes with high accuracy and efficiency. Examples of such tools are ligases (attachment), polymerases (elongation), endonucleases (specific site cleavage), or exonucleases (free 3'- or 5'-end cleavage) [10].

DNA chemical synthesis is based on automated solid-phase phosphoramidite chemistry synthesis, and is efficient up to around 200 nucleotides [11, 12]. The DNA sequence design is also an uncomplicated procedure and it can be done to fulfil certain pre-determined requirements. For example, the level and length of randomisation in an aptamer library or the insertion of a restriction site [13].

1.2.2 Interplay with Immune System

Oligonucleotides with sequences with unmethylated Cytidine phosphate Guanosine (CpG) motifs can induce potent immunostimulation. These motifs work as a signal to the organism, so it can act against a potential infective entity or uncontrolled growth. In eukaryotic cells, these motifs are mostly methylated (suppressed transcription). CpG motifs can [14]:

- stimulate B cells and Antigen Presenting Cells (APC),

- down-regulate macrophage class II major histocompatibility complex (MHC) antigen (Ag) presentation,

- activate the Nuclear Factor-κB (NF-κB) pathway,

- induce pro-inflammation via T helper 1 cells (Th1 cells) (Toll-like receptor (TLR) 9), and T helper 2 cells (Th2 cells),

- activate Natural Killer cells (NK cells),

- and induce polyclonal B cell activation.

CpG motifs are related to a myriad of diseases, from heart disease to Systemic Lupus Erythematosus (SLE) and Acquired Immune Deficiency Syndrome (AIDS) [15]. Insufficient methylation can lead to tumorous cells growth too [16].

CpG motifs are divided in three classes: A, B and C. A-class motifs stimulate NK cells cells and Interferon α (IFN-α) production. B-class oligonucleotides stimulate a strong B and NK cells activation and cytokine production. C-class is a mix of A and B responses, but the oligonucleotides have a phosphorothioate (PS) backbone, without poly-G, and with palindromes. All classes stimulate TLR signalling [17]. Natural CpG motifs can be hydrolysed in physiological conditions and have low cell penetration. PS-modifications make DNA more resistant to hydrolysis.

These CpG motifs activity has been studied deeply and different carrier structures have been evaluated. The size of the carrier DNA influences the activity, where longer DNA yields higher activity. Several DNA constructs carrying these motifs have been engineered, e.g., DNA origami delivery systems [18] and Y-shaped structures [19]. CpG motifs can be also used as vaccine adjuvants [20] because they stimulate TNF-α, IL-6, and IL-12 production [21].

Circulating cell-free DNA (cfDNA) is a valuable tool for diagnostics and therapy assessment in autoimmune diseases and cancers. cfDNA originates from DNA fragments formed after necrotic cell death [22]. The mean concentration in a healthy person plasma is approximately $(2.27 \pm 1.51)\,\mathrm{ng\,mL^{-1}}$. This value is elevated in autoimmune diseases, such as sarcoidosis $((3.68 \pm 2.07)\,\mathrm{ng\,mL^{-1}})$, asthma $((3.02 \pm 2.19)\,\mathrm{ng\,mL^{-1}})$, or chronic obstructive pulmonary disease (COPD) $((3.97 \pm 2.03)\,\mathrm{ng\,mL^{-1}})$. In oncological pathologies, such as non-small-cell lung cancer $((8.02 \pm 7.81)\,\mathrm{ng\,mL^{-1}})$, cfDNA is also elevated [23]. Some pathologies have been found to be associated with and classified using anti-DNA Abs, for example SLE [24].

6

1.2.3 Amplification

Polymerase Chain Reaction

PCR was initially developed by Kary Mullis to diagnose sickle cell anemia [25]. This technology made possible the specific amplification of a DNA sequence, using two shorter DNA sequences (primers) that flank the region of interest. The procedure is based on three simple steps: denaturation, annealing, and elongation of the oligonucleotide strands. Amplification follows DNA denaturation and annealing of the primers. Subsequently, the sequence is elongated by polymerase action, see Figure 1.1. The use of a thermostable polymerase, Taq polymerase, makes the PCR more reliable, sensitive, specific, and robust [26], because it is possible to achieve higher annealing temperatures.

In each amplification cycle, the DNA population doubles, 2^n, where n is the number of amplification cycles. To ensure an optimal amplification, primers should have 18 to 30 nucleotides, the same length, a 40 to 60% GC content, and no secondary structures or self-complementarity. Also, $MgCl_2$ increases specificity and DNA hybridisation [27].

Many different iterations of PCR have been developed [28]. In the next paragraphs, a few examples are described.

It is possible to generate ssDNA using PCR, such as in AsyPCR and modified nested-PCR. AsyPCR uses an excess of one primer over the other, leading to two phases of amplification. First, dsDNA production is exponential followed by linear amplification of ssDNA, which continues until enzyme depletion. Then, for ssDNA, the amplification is non-exponential and more ssDNA can be obtained with more PCR cycles [29, 30, 31]. AsyPCR has an efficiency of around 60% compared to normal PCR, which is more than 90% efficient [32].

Allele-Specific PCR (AS-PCR) is another elegant tool to manipulate DNA ampli-

fication. In this method, it is possible to amplify an allele that is different in one or more nucleotides from a template, by feeding the system a longer primer. This longer primer has, in the extra positions, nucleotides that are complementary to specific segments in a template. The primer with 3'-end mismatch is not able to extend and DNA amplification is cancelled [33, 34]. This method is used for Single Nucleotide Polymorphism (SNP) genotyping, haplotyping, and diagnostics of, for example, sickle cell anaemia [35]. AS-PCR can detect point mutations because one position mismatch will destabilise hybridisation [36].

Other technique to manipulate DNA during amplification is mutagenic PCR, where a polymerase 1000 times more error-prone than when Taq polymerase is used [37]. More recently, Dan Luo's group developed a thermostable branched DNA nanostructure that can be used as a modular primer. This can be used for labelling, and hydrogel or DNA origami formation [38].
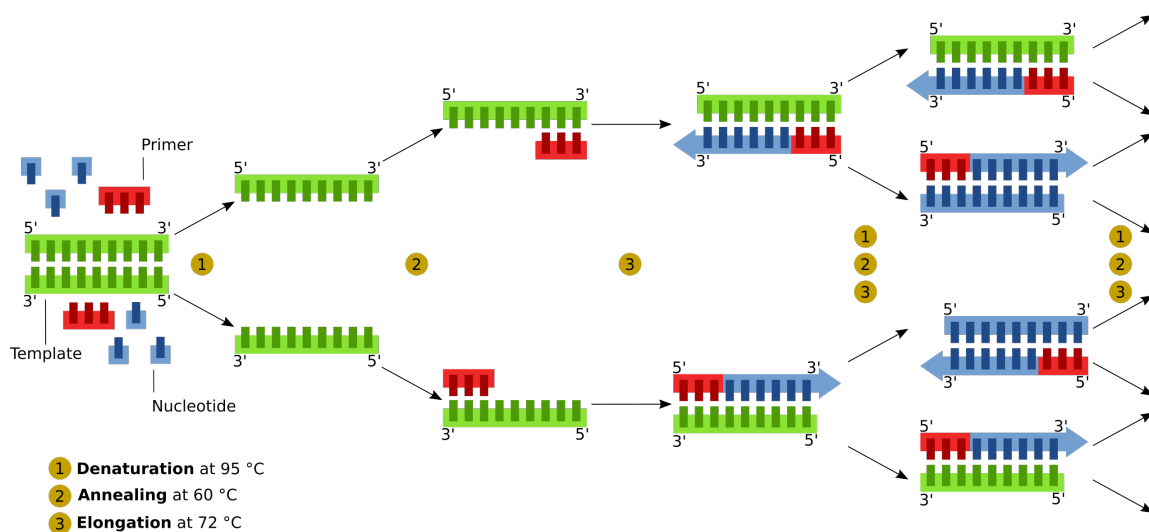
FIGURE 1.1: Schematic of a PCR cycle. The cycle starts with DNA denaturation, follows with primer annealing and elongation by polymerase action. Adapted with a permission under the Creative Commons Attribution-Share Alike 3.0 Unported license [39].

8

Real Time Polymerase Chain Reaction Another PCR type is the widely used RT-PCR, which mostly serves as a quantitative method. The setup is the same as in common PCR experiments. But, additionally to PCR, there is a correlation between fluorescence and DNA concentration. This DNA quantification based on developed fluorescence can be absolute or relative. Absolute quantification determines copy number and DNA concentration, requiring a standard curve. Relative quantification estimates concentration based on the threshold cycle, $2^{-\Delta\Delta C_t}$, where C_t is the threshold cycle and the number two represents perfect efficiency at each cycle. RT-PCR has a wide dynamic range, where C_t and DNA concentration have a linear relationship. It is a robust technique, easy to scale-up, and to multiplex [40, 41, 42]. Initial template concentration, primers, polymerase and its deactivation, deoxynucleotides (dNTPs), product reannealing and primer-dimer formation influence amplification efficiency. Thus, throughout amplification, efficiency is not constant or perfect at each cycle. For instance, efficiency is lower in later cycles as the increased DNA template concentration competes for the polymerase [43]. Gevertz et al. developed a mathematical model that simulates the PCR cycles and calculates efficiency as a function of cycle number. So,

$$[DNA]_n = [DNA]_0 \times (1 + \epsilon)^n,$$

where $[x]_n$ represents the concentration of x at cycle n and ϵ is the correction factor. Nevertheless, efficiency is assumed to be constant until plateau. Hence, it is possible to monitor dsDNA population by increase of fluorescence signal at each cycle and this, at the end of each cycle, is proportional to DNA product concentration [44]. The exponential amplification step must be well defined, so there is a more reliable concentration determination [45].

Several strategies to create fluorescence during RT-PCR have been developed. For ex-

9

ample, the concept of the dsDNA intercalation by a dye, such as SYBR Green I. Free SYBR Green I is non-fluorescent and when complexed with dsDNA increases fluorescence more than a 1,000-fold, with a sensitivity less than $10\,\mathrm{pg\,mL^{-1}}$. SYBR Green I intercalates DNA's minor groove, in a segment around 3.5 base pairs ($11.5\,\text{Å}$) [46]. It is common to use a passive reference dye, such as carboxy-X-rhodamine (ROX) to normalise the signal and cancel fluorescence variations [47]. Other strategies for fluorescence development are fluorescently-labelled oligonucleotides, hydrolysis or hybridisation probes, and molecular beacons.

Furthermore, RT-PCR can provide information about the oligonucleotides melting curve, which indicates the dsDNA pool nature. The method relies on subjecting the PCR product to a wide range of temperatures until they denature. Shorter duplexes and mismatches have a lower melting temperature than the expected size product [27].

1.2.4   Uses in Biotechnology

The DNA molecule is programmable, flexible, easily synthesised and modified, non-toxic, and biocompatible. These features make it a desirable solution for many biotechnological applications and suitable for a vast range of functions. DNA can adopt different conformations and adapt to different conditions [7]. Increasingly, it is being used as a theranostic tool, combining diagnostic and therapy [48].

In this section, three DNA-based technologies are discussed: SELEX, DECL, and organised structures. It must be emphasised that the volume of information currently available regarding other DNA uses in biotechnology is overwhelming and very frequently of upmost interest.

Systematic Evolution of Ligands by Exponential Enrichment (SELEX)

Aptamer technology was developed by two distinct groups at the same time, Tuerk and Gold at the University of Colorado [49] and Ellington and Szostak at the University of Texas in Austin [50]. These performed an in vitro selection of Ribonucleic acid (RNA) against a ligand by isolating subpopulations from a randomised oligonucleotide pool. The selected single-stranded oligonucleotides folded into a specific shape, able to interact with the target. This short oligonucleotide, up to 100 bp, was named aptamer, from the Latin *aptus*. Aptamers can be DNA or RNA molecules and form stable secondary structures, which generate hairpins, G-quartets, bulges or pseudo-knots. Aptamer interactions with a target are held through flat moieties stacking, hydrogen bonds and molecular shape complementarity [51]. In general, DNA and RNA aptamers have similar affinities, but DNA selection is simpler because it only needs direct PCR amplification. RNA must be transcribed, reverse transcribed, and then amplified [52, 30].

The concept of aptamer selection is called SELEX. SELEX is a cyclic process where large pools of randomised single-stranded oligonucleotides are selected throughout several rounds, until a final pool has the desired affinity. This randomised pool is incubated with a target and, after washing, the bound sequences are eluted and amplified by PCR. The ssDNA is regenerated from the dsDNA PCR product. The cycles are repeated until the goal specificity is achieved [53], see Figure 1.2. The assessment of pool quality can be done by different methods, mostly through affinity measures. One common way to assess complexity of a pool is by its melting curve [54]. Using Sanger or deep sequencing, it is possible to decode the sequences. For instance, Schuetze et al. showed, by sequencing every selection round, that high affinity aptamers can be found in the first selection rounds, and suggests only three cycles are necessary [55]. The same group previously used a systematic approach,

Diversity Standard of Random Oligonucleotides (DiStRO), to determine randomised oligonucleotides melting curves. They concluded that different diversities had different annealing kinetics, with melting curves able to mirror it. Thus, DiStRO can be used to compare randomised oligonucleotide populations [56].
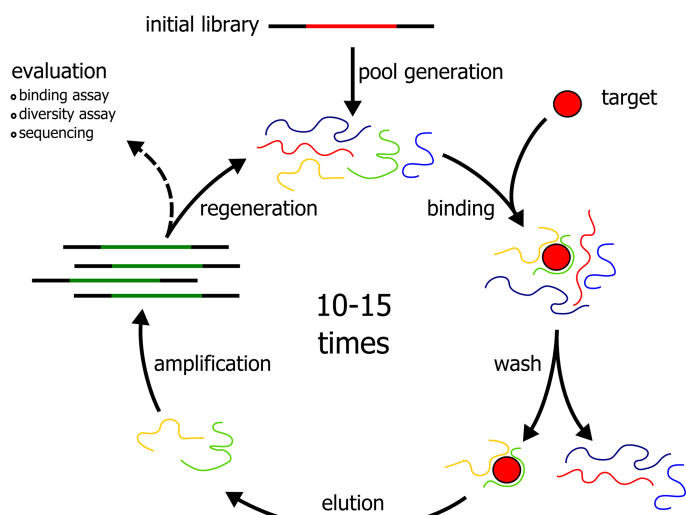
FIGURE 1.2: Representation of a SELEX cycle. A typical SELEX cycle starts with the incubation of the randomised single-stranded oligonucleotides pool with the target, followed by washing of non-binders, and elution of binders. The binders are amplified and evaluated. Then, regeneration of single-strands can be done and the cycle starts again. From Schuetze et al. [55].

Several twists in the SELEX concept have been developed [57]. For instance, blended-SELEX is based on a setup with a lead molecule linked covalently to a DNA strand and hybridised to a second oligonucleotide strand with a randomised region. This methodology was described for the selection of RNA-based irreversible inhibitors of human neutrophil elastase. With this approach, the oligonucleotide sequences which provide the lead compound with the best reaction to the target are selected [58]. The same concept is used with a randomised DNA-valine phosphonate library, an irreversible neutrophil elastase inhibitor [59]. Cell-SELEX is also possible and generates aptamers against the target molecule on the live cell. It is mostly used against surface transmembrane proteins difficult to obtain in native conformations

12

[60, 61]. In another approach, the non-homologous random recombination lets the DNA fragments randomly recombine, without sequence homology, according to their length. This concept is in contrast with SELEX, which explores the sequence space by enrichment, not by creation of new entities or rare point mutations [62].

Improved DNA sequencing technologies and bioinformatic tools helped to decrease drastically the number of selection rounds needed, as in single round enrichment of ligands [63, 64], over-represented libraries [65], an algorithm in post-SELEX to mature aptamers [66], or predefined binding parameters (smart aptamers) [67].

A vast number of aptamers have been described. Examples are human thrombin [68], ATP [69], oxytetracycline [70], thrombin [71], coagulation factors [72], ethanolamine [73], protein C [74], bisphenol A [75], chimeric aptamers [76], enterotoxin B [77], streptavidin (SA) [78, 79], tenascin-C [80], digoxin [81], and proteases [82].

Aptamers are compelling molecules to use in therapy, because of their size and ease to synthesise and modify; characteristics that overlap the advantages of DNA molecules. Most developed aptamers that are candidates for therapy inhibit protein-protein interactions, and are antagonists. They reach active pockets or disrupt the protein spacial conformation [83]. Aptamers with agonistic properties have also been developed and can co-stimulate receptors [84, 85]. The same group reported the selection of aptamers with activity higher than Abs, where two conjugated aptamers, one immunomodulatory and the other a target binder, had lower toxicity [86].

Nevertheless, the use of aptamers in therapy is complicated, because it is difficult to predict DNA pharmacokinetics. DNA's small size makes it suitable for renal clearance and endonucleases can digest it. Until now, the only aptamer used in the clinic is a modified RNA aptamer, pegaptanib (Macugen), which has low K_d for the Vascular Endothelial Growth Factor (VEGF). This aptamer is used to treat age-related macular degeneration [87].

To improve aptamers' stability in biological fluids, one can use modified nucleotides,

13

such as locked nucleic acids or spiegelmers [88], covalently bind the 3'-end to proteins, or use a PS backbone. Other strategies to improve pharmacokinetics are based on the conjugation of the aptamer to polyethylene glycol (PEG), for better tissue distribution [89], or blocking its 3- or 5-end to avoid digestion [90, 91]. In parallel, the development of new polymerases enables the use of non-natural building blocks and the continuous improvement of modified oligonucleotide synthesis is improving the technology. Also, higher affinity aptamers can be created by expanding the genetic alphabet, such as using nucleobases with higher hydrophobicity. Modifications of this nature enhance interaction probability and durability between the hydrophilic oligonucleotide and the hydrophobic active site [4].

All this potential of aptamers can be integrated in many other nanomaterials, such as gold nanoparticles and nanorods, quantum dots, carbon nanotubes [92], diagnostic and biosensor chips [93, 94], nanorulers, arrays, DNA origami [95], molecular imaging [96], hydrogels, films, microcapsules [97], and DNA micelles [98].

DNA-Encoded Chemical Libraries (DECL)

Low molecular weight molecules are ideal starting points for drug discovery and is compatible with highthroughput screening (HTS) [99]. Although HTS pans millions of molecules, it has often failed in identifying bona fide hits. In parallel, fragment-based lead discovery screens lower molecular weight libraries, less than 300 Da. These fragments are not strong inhibitors, must be used at higher concentrations, and detected with very sensitive tools. Theoretically, due to smaller size, a fragment-based approach can explore chemical spaces better than HTS.

A simple molecule is a more suitable starting point, because of less potentially unwanted effects (e.g., promiscuous interactions) [100] and the possibility of adding functions. Unlike HTS hits, fragments have high 'ligand efficiency', which makes them good candidates for optimisation.

To further expand the potency of this technology, combinatorial assembly has been used to increase exponentially the number of compounds in a screening campaign [101]. Dynamic Combinatorial Chemistry (DCC) proceeds under thermodynamic control, when all participants are in equilibrium, relying on structures self-assembly. The interconversion of library members is reversible and the final library constitution is ruled by thermodynamic stability [102, 2]. DCC virtually covers all possibilities inside a library and the target becomes the engine of this adaptive process [103].

To overcome the limitations inherent to HTS, DCC and fragment-based lead discovery, DECL strategies have been developed in recent decades [104]. DECL is a derivation of a well-known and established technology used in Ab development, phage display. Phage display was first described by George P. Smith, when he reported the expression of vectors displaying cloned Ags on a phage surface [105]. This technology relies on the selection cycles of phages that bind an Ag, thus creating soluble Ab fragments from infected bacteria. The selection process emulates the adaptive immunity [106]. There is a physical link between genotype and phenotype, so the DNA encoding the selected Ab can be retrieved. It can develop Abs against virtually every Ag, which is not possible in vivo [107].

The first concept paper about DECL was written by Brenner and Lerner and it describes the gradual *'synthesis of a polymer and an oligonucleotide (as a code) on a linker in split-and-pool cycles'*. It characterises chemical entities with genetic tags attached, which, after binding to a target and eluted, are amplified by PCR [108]. About one decade later, a series of three papers by Halpin and Harbury set the playing field for this emergent trend, and merged DNA technology with drug discovery, creating real DECL. The papers describe the improvement of DNA-programmed chemical synthesis [109], selection process [110], and how to improve unprotected DNA synthesis in solid-phase [111]. This fairly recent technology has advantages over HTS because it does not need a protein-specific assay, can be miniaturised, and

15

is comparatively low cost [112, 113].

More DECL strategies were developed around the same time [114, 115, 116]. A first basic approach of DECL is the single pharmacophore library, which can be built by split-and-pool or using complementary oligonucleotides derivatives that can carry chemical reactions (DNA-templated synthesis), with assisting complementary oligonucleotides [113]. This concept evolved and the Encoded Self-Assembling Chemical (ESAC) libraries were developed. In ESAC, there is the display of covalently linked chemical fragments at the end of self-assembled oligonucleotide strands, which carry a unique DNA sequence for each fragment. After panning, the eluted molecule carries the tag to decode [117, 118], see Figure 1.3. ESAC explores the chelate effect and the dual molecule display leads to the discovery of two binders, similarly to fragment-based approaches - dual-pharmacophore strategy [119]. This scaffold identifies synergistic binding pairs, with high affinity to the target [120].
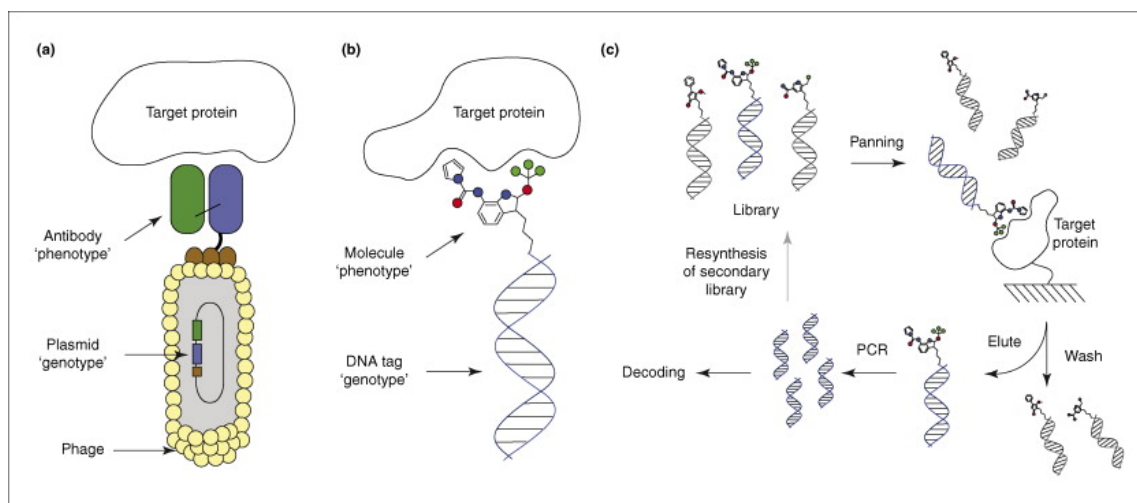
FIGURE 1.3: Comparison between phage display and DECL. A phage is the physical link between the Ab fragment (phenotype) and the gene encoding (genotype) (a). A small molecule (phenotype) is attached to a DNA strand (genotype) that carries an unique identifying code (b). Schematic of an affinity selection cycle, where the binding molecules are eluted and carry the DNA code until final identification, or resynthesis of a secondary library (c). From Melkko et al. [121].

In an advancement to this original method, the alternated stepwise assembly

of fragments carrying DNA tags was studied by high-throughput sequencing '454 technology' [122, 123]. So far, ESAC technology has been used to improve the affinity of known ligands, affinity maturation, with a second fragment [124]. Only recently, the dual-pharmacophore setting identified a de novo pair of fragments because a major hurdle in this dual-selection was how to retrieve the DNA codes. This was overcome using an inter-strand code-transfer [125]. Another progress, based on DCC, lead to the dynamic DECL strategy, where large libraries are assembled, by the dynamic reshuffling of complementary DNA strands. This reshuffle and the adaptive pressure by the target can spawn population evolution towards stronger binders [3]. Examples of published DECL campaigns are increasingly more common and a wide range of targets have been explored, such as SA [126], trypsin [121], albumin [127], carbonic anhydrase IX [113], IL-2 [128], hydrolases [129], phosphatases [130, 131], and kinase isoforms [132, 133].

Organised Structures

Nadrian Seeman's group revolutionised nano-materials with a DNA nanotechnology named DNA origami. Exploring the well-described DNA properties, the group developed a new concept to programme and manipulate structures at a nano-level, exploring DNA's predictable and controllable self-assembly [134]. These structures were stable, biocompatible, and able to carry specific sequences used for specific targeting (DNA nano-carriers). Seeman's group built the first closed polyhedral object with DNA [135], developed 2D DNA crystals [136, 137], a rotary nano-mechanical device [138], a 3D crystalline lattice [139], and DNA nanotubes with CpG motifs [140].

Currently, the widely used DNA origami technology is based on a self-assembly concept to generate low-entropy DNA structures in a simple and inexpensive way. The assembly relies on staple strands (shorter complementary ssDNA) that anneal to

specific sites, folding a scaffold strand (long ssDNA). The recently developed rolling circle amplification method is also a method for DNA origami production and generates shorter sequences and DNA nano-ribbons [95].

DNA origami nano-structures can be functionalised [141]. For example, a DNA nanorobot, able to carry and deliver drugs by using an 'aptamer-encoded logic gate', has been developed [142]. However, it has not been seen yet a consistent use of these outside of the laboratory. There are promising attempts in biosensing, enzymatic manipulation, drug delivery, and biomimetics. Meanwhile, tools like caDNAno have been developed to help DNA origami design using a graphical interface [143].

Since the advent of this technology, an incredible number of DNA-based structures has been made. A few examples are the aptamer-tethered nano-trains, diverse DNA geometries for increased cell uptake, and the development of artificial biosensors [144]. In parallel, DNA hydrogels have also been described [145], such as dendrimer-like-DNA structures [146, 147]. DNA hydrogels are 3D, biocompatible, biodegradable, low cost, easily manipulated, and formed under physiological conditions. Like DNA origami, DNA hydrogels can be used in drug delivery, tissue engineering, 3D cell culture [148], or DNA-based nano-devices [149, 150, 20].

1.2.5 Modelling, Computation and Storage

The need of better computation and modelling tools has been an evolving worry in parallel with technological development. The use of DNA for such ends is a relatively new solution and is being updated every day. In this section, there is an attempt to depict the computational world panorama and how we are on the verge of a paradigm shift, where DNA might be playing a leading role.

An innovative model for DNA computation, 'sticker model', was developed by Adleman, where DNA strands are the physical representation of information and annealing is the central mechanism. This model has a random access memory without

strand extension need, uses no enzymes and is reusable. This *'microprocessor-controlled parallel robotic workstation'* generates algorithms to be implemented using DNA-based computers. The model assumes that covalent bonds are not needed in the computation process, so the system consumes less energy, making the bridge into a molecular computer [151].

This breakthrough computation method used DNA to solve the Hamiltonian path problem, which consists of establishing a path between two cities, with an incomplete group of roads. DNA strands represent cities and roads, and the sequences are designed so a road connects to a city. When all strands are mixed, they are able to self-assemble and solve the problem by themselves [152].

Winfree also suggested that short branched DNA could self-assemble following an algorithm, making DNA molecules the simplest algorithm modules. To improve this toolbox, more complex algorithms could be assembled using sticky ends [7]. Also digital logic circuits were established [153] and a DNA model of a small neural network, using reversible strand displacement [154].

Recent exciting news reported that Microsoft bought ten million strands of synthetic DNA for data storage to Twist Biosciences, a company specialised in etching techniques similar to chip manufacturing for DNA synthesis. This data storage technology relies on the durability and density (almost a trillion gigabytes per gram) of DNA. This robustness is an advantage over current storage because it is resistant to high temperatures and magnetic fields [155]. Up to now, it is possible to code and decode a whole set of digital data from DNA [156].

This technological leap could shatter the rules of giant data storage. Today, immense data centres are being built because of the endless accumulation of data we all generate every day. Plus, the revolutionary concept of *Internet of Things* will change the way we humans interact with devices and will increase dramatically the amount of data to store in the world. It is predicted that by 2020, 50 billion devices will

be connected to the Internet and generate 44 ZB of data every year [157]. Current storage is not enough to accommodate these numbers.

Today's status of DNA storage can be seen on Figure 1.4. This technology still has several limitations holding it back: encoding and decoding (replication errors or reading speed), DNA synthesis (yield and cost), and sequencing. These limitations are being dealt with and, for example, a strategy for better DNA encoding has been already developed [158].

For better modelling and computation in a general sense, Tim Palmer suggests a very strong case for building imprecise computers, given that precise supercomputers are accurate, but energy expensive. The claim is that it is not always necessary to have deterministic Turing machines and that some randomness improves the system. *'Energy-efficient hybrid supercomputers'* with different accuracy grades could be used for different ends, where different levels of precision would be required. For instance, the human brain is able to power through fantastic levels of computation in low energy levels, less than 20 W, and give space to something as complex as creativity, jumping between low- and high-energy determinism [159, 160].

STORAGE LIMITS
Estimates based on bacterial genetics suggest that digital DNA could one day rival or exceed today's storage technology.

	Hard disk	Flash memory	Bacterial DNA
Read–write speed (µs per bit)	~3,000–5,000	~100	<100
Data retention (years)	>10	>10	>100
Power usage (watts per gigabyte)	~0.04	~0.01–0.04	<10^{-10}
Data density (bits per cm³)	~10^{13}	~10^{16}	~10^{19}

WEIGHT OF DNA NEEDED TO STORE WORLD'S DATA — ~1 kg

©nature

FIGURE 1.4: Comparison between data storage power and limits of a hard disk, a flash memory, and bacterial DNA. Adapted from Extance [161].

Population Studies Limitations   To complete this section, it is needed to discuss the roadblocks existent in modelling populations. Similarly challenging as to building a supercomputer with low energy consumption is to analyse and model a large society, in which each individual is characteristically different. It is often very difficult to collect and analyse reliable data from hundreds to thousands of individuals, not to mention a large and highly diverse society, e.g., human society.

In a more and more inter-connected world, it is less likely that we can treat any group of people as an isolated population. The concepts used nowadays plus the statistical tools are insufficient and frequently lag in regards of accuracy, precision, and/or adequacy. As a concrete example, there is an urgent need for improving the statistical models for species distribution, in order to keep up to follow drastic changes caused by climate change.

Austin reviews several methods and concludes that the tools used are limited and theory is missing. Most of all, data extension still determines the fit, which is a huge limitation. The improvement of population studies tools with more complex models could improve the selection of predictor variables [162]. Citing Caswell [37], theoretical problems in ecology can be summed as:

- exploring the consequences of a theory, determining the possibility of a theoretical model contributing to an ecological one,

- demonstrating the link between seemingly unrelated theories,

- robustness evaluation,

- developing mathematical methods for theoretical models analysis,

- finding the simplest model able to mimic the observation,

- and designing empirical and/or experimental tests of the theory.

In the concrete case of population adaptation, it is possible to summarise in two parameters: colonisation of new environments by newly adapted populations and local adaptation in a heterogeneous environment. These can include exposure to a novel host or food resource, a new biophysical environment, a new predator, or a competitor [163].

1.3 Tumour Necrosis Factor α

TNF-α is a pleiotropic pro-inflammatory cytokine mainly produced by activated macrophages and T cells. It was first described as an endotoxin that caused hemorrhagic necrosis of sarcomas. TNF-α's many biological roles have been thoroughly reviewed [164], but they can be summarised as being mostly related to inflammation and immune responses. TNF-α causes cell death through necroptosis, which is a controlled necrosis that follows signal transduction pathways [165, 166].

This cytokine has an important role in the pathological development of most autoimmune and inflammatory conditions, such as rheumatoid arthritis, inflammatory bowel disease, uveitis, SLE, or psoriasis. In the nervous system, TNF-α's inflammation induction can be linked to Alzheimer's disease, multiple sclerosis, depression, and other neurological disorders. Cardiovascular, respiratory, and metabolic diseases can also be primed by TNF-α's inflammation pathway. Plus, TNF-α is heavily involved in host defence [167].

TNF-α activates NF-κB, which up-regulates oncogenesis. Knock-outs show normal development, but are more susceptible to infections and are resistant to lipopolysaccharide (LPS) [168]. TNF-α is usually not detected in healthy individuals, but it is elevated in inflammation and infection, reaching values from 6.3 ± 3.6 to $(27.50 \pm 10.62)\,\mathrm{pg\,mL^{-1}}$ [169, 170, 171].

1.3.1 Structure and Function

Pro-TNF-α is a transmembrane protein (around $26\,\mathrm{kDa}$) that is cleaved into a soluble form (around $18\,\mathrm{kDa}$) by the TNF-α converting enzyme (TACE). TACE is also involved in processing other membrane proteins, such as TNF-α receptors, which are released from the cell membrane and generate soluble receptors that neutralise TNF-α. TACE is pro- and anti-inflammatory [172]. It is hypothesised that TNF-α is sta-

23

bilised by its soluble receptors, keeping its activity and augmenting it [173]. TNF-α activity is mediated too through high-affinity receptors TNF-α receptor 1 (TNFR1), expressed ubiquitously, and TNF-α receptor 2 (TNFR2), expressed by haematopoietic cells. Both are homodimers. TNF-α binds almost irreversibly to TNFR1, which carries a cell death domain: TNFR1 activates mostly pro-inflammatory events and apoptosis. TNFR2, without cell death domain, binds to TNF-α reversibly and is linked to tissue repair and angiogenesis.

Thus far, TNF-α activation has been attached to the macrophage, which is activated at the inflammation site [174]. Interestingly, this process is associated with release of DNA. It has been reported that high concentrations of PS oligonucleotides increase TNF-α synthesis in vitro. This is probably due to the binding of the polyanionic structure to cationic charges present on the cell surface, which also increases LPS-induced TNF-α synthesis [175, 176]. LPS and TNF-α have a synergic effect in cell death [177], which is inhibited by thalidomide and chromene derivatives [178, 179]. TNF-α and its receptors must oligomerise to initiate a biological response [180]. However, the way the interaction occurs is still not fully understood, but there are some hypothesis [181]:

- receptor aggregation: three receptors bind to one TNF-α trimer, generating a signal (Figure 1.5a),

- hexagonal array conformation: an increasing hive of TNF-α trimers bind to receptor dimers, with each ligand binding to three receptors (Figure 1.5b),

- and molecular switch formation: the ligand binds subunit B and A is released, so the receptor can bind to a second free site on the ligand, leading to conformational changes, and signal transduction (Figure 1.5c).

24
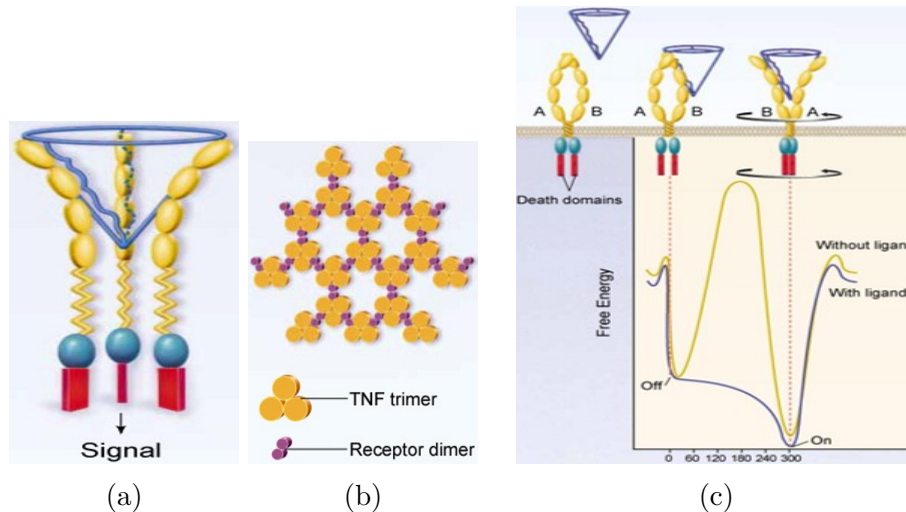
FIGURE 1.5: Trimerisation hypothesis: three receptors assemble due to the interaction with one TNF-α trimer, generating a signal (a). Expanding-network hypothesis: an expanding hive-like structure of TNF-α trimers bound to receptor dimers, capping the receptor can trigger a response (b). Molecular-switch hypothesis: each dimer can be activated when a ligand binds to B. Then, A disengages from B, so the receptor can bind through a second site to the ligand. These cause conformation changes that will alter the cytoplasmic domain of the receptor and lead to signalling (c). Adapted from Bazzoni at al. [181].

Structure-wise, TNF-α is a non-covalently linked trimer. Each monomer is an antiparallel β-sandwich and loops in a jellyroll shape [182], with an N-terminal insertion with three more β-strands. Monomers are around 60 Å long and 30 Å wide [164]. For biological activity, TNF-α must be a trimer [183] and dissociates at physiological concentrations (below nano-molar) [184]. TNF-α's trimer is highly dynamic, as the monomers assemble and disassemble [185], influenced by concentration, detergents, and pH [186]. Some drugs, such as suramin, inhibit TNF-α directly by acting on its structure and disassembling the trimer [187]. Using suramin, it was possible to model the sequence of TNF deoligomerisation (trimer to dimer to monomer) and it was estimated that each step had a similar dissociation constant of $0.2\,\mathrm{nmol\,dm^{-3}}$ [188, 189]. TNF-α also has a heparin-binding domain at the N-terminal of the trimer [190].

25

TNF-α membrane permeability can be enhanced through partial denaturation, which generates an aggregate with large amounts of α-helix at low pH, high temperatures, or in the presence of guanidine hydrochloride [191]. The α-helix is not present in the native conformation and it is hypothesised that it can interact with lipid bilayers, leading to membrane insertion [192]. In this condition, ion channels are formed, increasing Na$^+$ influx [193, 194]. These changes are ultimately linked to higher cytotoxicity.

1.3.2 Anti-TNF-α Therapy

The well-functioning of the immune system is based on a fragile equilibrium between pro- and anti-inflammatory players [195]. Unfortunately, TNF-α is very frequently engaged in the tilting of this balance and its regulation is of upmost importance. Nonetheless, the discovery and development of small molecular pharmaceutical entities able to tackle this problem has been short and TNF-α remains an elusive drug discovery target, see Figure 1.6. In this section, it will be briefly revised the advancements done in this effort.

Nowadays, TNF-α inhibitors licensed and used are mostly anti-TNF-α monoclonal Ab-based [196, 197, 198], such as:

- infliximab: a human-murine chimeric immunoglobulin G (IgG) 1,

- adalimumab: a human anti-human TNF-α, generated by phage display,

- golimumab: a fully humanised Ab,

- certolizumab pegol: a PEGylated fragment-antigen binding (Fab),

- or etanercept: a soluble receptor fusion protein, where TNFR2 is coupled to the IgG's fragment constant region (Fc).
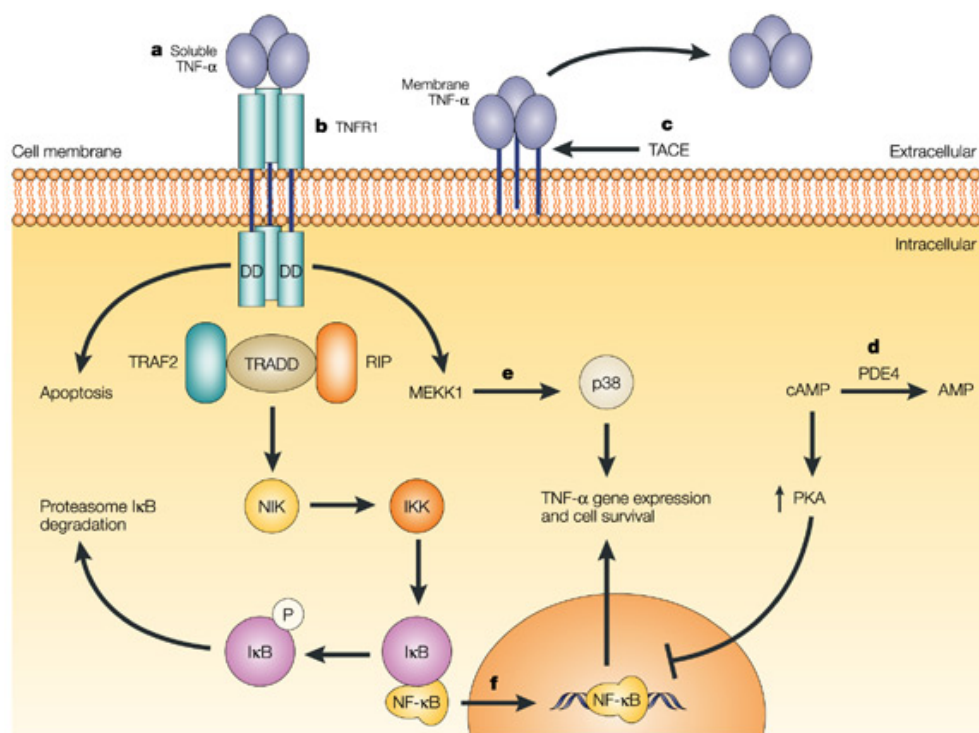
FIGURE 1.6: TNF-α signalling pathways and potential inhibition targets. TNF-α is involved in a complex net of molecular events and, thus, has a myriad of probable targets to be controlled through. Extracellularly, the inhibition of TNF-α (a) or its receptor (b) has already yield a few Ab-based therapy options. Plus, TACE (c) inhibition could prevent the generation of soluble TNF-α. Intracellularly, PDE4 (d), p38 (e) and NF-κB (f) inhibitors are also interesting. NF-κB inhibition would decrease pro-inflammatory cytokines production and promote cell death in specific populations, such as cancer cells. *Legend*: cAMP, cyclic AMP; DD, death domain; IκB, inhibitor of NF-κB; IKK, IκB kinase; MEKK1, mitogen-activated protein kinase kinase kinase 1; NIK, NF-κB-inducing kinase; PDE4, phosphodiesterase 4; PKA, protein kinase A; RIP, receptor-interacting protein; TRADD, TNFR1-associated death domain-containing protein; TRAF2, TNFR-associated factor 2. Adapted from Palladino et al. [195].

These inhibitors are used to treat inflammatory diseases, for instance, rheumatoid arthritis and Crohn's disease. However, they have typical Ab-based therapy limitations, such as high cost, plenty of side effects, and intravenous administration. In an effort to improve their pharmacokinetics, PEGylation has been used, increasing their half-lives.

The need to develop small molecules for direct TNF-α inhibition arises from the existing therapies limitations. The ideal small molecule therapy should be orally administered, have more desirable pharmacodynamic and pharmacokinetic profiles, have less side effects, be suitable for combination, and have lower cost.

Still, as referred in the previous section, it is largely unknown the exact way TNF-α interacts with its receptors, which could be the key for efficient inhibition. Protein-protein interaction or oligomerisation disruption are promising strategies for finding a more suitable target related to TNF-α's mode of action. TNF-α dynamic trimer might be the key for a small molecule target to interfere and prevent correct conformation for receptor interaction [199].

Screenings for these molecules have been made. It has been already reported a micro-molar range inhibitor by disruption of TNF-α, where trimer dissociation increases 600 times [200]. Other small molecules developed are, for example, benzobicyclooctanes [201], bicyclic pyrazolones [202], or a new class of inhibitors selected by DECL [112]. Artificial binding protein scaffolds with low molecular weight [203], peptides with low micro-molar affinity [204], bicyclic peptides [205], and other protein-based attempts [206] have also been developed. Only one anti-TNF-α DNA aptamer, VR11, (K_d of $7\,\mathrm{nmol\,dm^{-3}}$) has been reported so far [207].

1.4 Overview

This introduction is an attempt to describe DNA in all valencies. Likewise, we try to establish parallels between biological and artificial systems, see Table 1.1.

Table 1.1: DNA functions and characteristics.

	Biological System	**Artificial System**
Genetic information	Central dogma	Phage-display, ribosome display
Double helix formation	Central dogma	DNA origami, DNA hydrogel
Ligand	CpG motifs	Aptamer (SELEX)
Quaternary number	?	DNA computation/Data storage
Negatively charged polymer	?	Surface coating

- DNA's natural function is to be the basic unit of genetic materials for information storage inside cells (Subsection 1.2.1). Today, as we try to copy it, there are artificial setups, such as phage, ribosome display, and DECL (Subsection 1.2.4, DECL). These systems are able to, synthetically, create a link between genotype and phenotype.

- DNA's structure itself is also profoundly interesting. Its ability to organise extremely large amounts of data based on its four bases, which assemble into a ssDNA and then into a double helix, always following a strict pairing rule, is essential for information replication. This characteristic also makes DNA an incredible tool in the development of new biological materials, e.g., DNA hydrogels or origami. Also, its biocompatibility and overall robustness makes it an excellent tool for nanobiotechnological applications (Subsection 1.2.4, Organised Structures).

- More recently, DNA has been explored in biological systems as a ligand. Since nucleobases are entities that create interactions and establish links with their milieu, DNA is able to act directly as a ligand. For instance, the presence

of DNA carrying CpG motifs can induce immune response (Subsection 1.2.2). Exploring the possible conformations DNA can present, such as loops or G-quadruplexes, the aptamer technology arose (Subsection 1.2.4, SELEX). Using aptamer technology, one has the possibility to create a synthetic oligonucleotide ligand with very high affinity to various targets of interest.

- The physical chemistry of DNA as a highly negatively charged conductive polymer is something being explored too. However, in a biological system, we are still not very clear regarding its function as a consistent negative charged polymer chain. Nevertheless, this feature is already being used to develop coated surfaces. One can hypothesise various unspecific interactions that DNA is involved in an organism.

- As mentioned above, the capability of DNA to store large amounts of data using a quaternary system is quite amazing. In a living system, these four bases encode a limited number of amino acids. Nowadays, there is an emerging school of thought lobbying for a paradigm shift of how to store or even compute data (Subsection 1.2.5). As we know, today it is used a binary system to encode every digital information. As the world's data increases exponentially, 0 and 1 might be insufficient. Therefore, the adoption of a quaternary system might be a solution, with `A`, `G`, `C`, and `T`.

In this dissertation, I try to understand a bit further the pieces that are still missing in this puzzle. As a mix of DNA can be considered as a group of random number, we aim to develop a method to analyse the diversity. Also, we want to explore DNA's elusive functions in the biological context.

<div align="right">

2

</div>

<div align="right">

Materials and Methods

</div>

2.1 General

2.1.1 Materials

All oligonucleotides were purchased from IBA (Germany) or Eurofins (Germany) and their sequences can be consulted on Tables 2.1, 2.2, 2.3, 2.4, and 2.5. All oligo-nucleotide and protein concentrations were measured using a NanoDrop instrument (Thermo Fisher Scientific, Germany), quantifying the absorbance at 260 and 280 nm, respectively. The L929 cell line from mouse (C3H/An connective tissue) was kindly provided by the Garbe group (Center for Regenerative Therapies Dresden - CRTD). The mouse recombinant TNF-α was kindly provided by Dr. Weilin Lin. All data were plotted using OriginPro 8.6G (OriginLab, USA).

- Agilent (USA): Mx3000P RT-PCR system,

- AppliChem (Germany): Phosphate Buffered Saline (PBS), glycerol, glycine, ethylenediaminetetraacetic acid (EDTA), low molecular weight Salmon Sperm DNA (S-DNA), and tris(hydroxymethyl)aminomethane (Tris),

- Applied Photophysics (UK): Chirascan plus,

Table 2.1: DNA sequences of the templates used on the DDMA project.

Aptamer Libraries

X_{20}	GACAATTCACACACGTCCGCNNNNNNNNNNNNNNNNNNNNATGAGATCGGAAGAGCGTCG
X_{18}	GACAATTCACACACGTCCGCANNNNNNNNNNNNNNNNNNGATGAGATCGGAAGAGCGTCG
X_{16}	GACAATTCACACACGTCCGCATNNNNNNNNNNNNNNNNCGATGAGATCGGAAGAGCGTCG
X_{14}	GACAATTCACACACGTCCGCATTNNNNNNNNNNNNNNCCGATGAGATCGGAAGAGCGTCG
X_{12}	GACAATTCACACACGTCCGCATTCNNNNNNNNNNNNTCCGATGAGATCGGAAGAGCGTCG
X_{10}	GACAATTCACACACGTCCGCATTCGNNNNNNNNNNTTCCGATGAGATCGGAAGAGCGTCG
X_8	GACAATTCACACACGTCCGCATTCGTNNNNNNNNGTTCCGATGAGATCGGAAGAGCGTCG
X_6	GACAATTCACACACGTCCGCATTCGTANNNNNNAGTTCCGATGAGATCGGAAGAGCGTCG
X_4	GACAATTCACACACGTCCGCATTCGTAGNNNNCAGTTCCGATGAGATCGGAAGAGCGTCG
X_2	GACAATTCACACACGTCCGCATTCGTAGGNNTCAGTTCCGATGAGATCGGAAGAGCGTCG

High Diversity Libraries

S_1	GACAATTCACACACGTCCGCAGTCTGACTGATCACTGGACATGAGATCGGAAGAGCGTCG
S_1'	CGACGCTCTTCCGATCTCATGTCCAGTGATCAGTCAGACTGCGGACGTGTGTGAATTGTC
S_2	GACAATTCACACACGTCCGCATCTATGCCTGTATTAAAGCATGAGATCGGAAGAGCGTCG
S_3	GACAATTCACACACGTCCGCACAAGGGGTCAAGCTCCTGTATGAGATCGGAAGAGCGTCG
S_4	GACAATTCACACACGTCCGCCCCTCCCTAAACTTTGCCTAATGAGATCGGAAGAGCGTCG
S_5	GACAATTCACACACGTCCGCGGGATCAGCCTGGGGATACAATGAGATCGGAAGAGCGTCG
S_6	GACAATTCACACACGTCCGCTATAAATACTACAAGCTCATATGAGATCGGAAGAGCGTCG
S_7	GACAATTCACACACGTCCGCGGGGCACCGCCAACAGAGATATGAGATCGGAAGAGCGTCG
S_8	GACAATTCACACACGTCCGCCCGGATTATATCCCATGAGGATGAGATCGGAAGAGCGTCG
S_9	GACAATTCACACACGTCCGCCTAGTGAAGTGATTCGTTCGATGAGATCGGAAGAGCGTCG
S_{10}	GACAATTCACACACGTCCGCGAACGTTAAGGGTGACCGTCATGAGATCGGAAGAGCGTCG

S_1 Mismatched Complementary Templates

$S_1'_{M2}$	CGACGCTCTTCCGATCTCATGTCCAGTGAGTAGTCAGACTGCGGACGTGTGTGAATTGTC
$S_1'_{M4}$	CGACGCTCTTCCGATCTCATGTCCAGTGCGTGGTCAGACTGCGGACGTGTGTGAATTGTC
$S_1'_{M6}$	CGACGCTCTTCCGATCTCATGTCCAGTACGTGTTCAGACTGCGGACGTGTGTGAATTGTC
$S_1'_{M8}$	CGACGCTCTTCCGATCTCATGTCCAGCACGTGTGCAGACTGCGGACGTGTGTGAATTGTC
$S_1'_{M10}$	CGACGCTCTTCCGATCTCATGTCCAACACGTGTGTAGACTGCGGACGTGTGTGAATTGTC
$S_1'_{M12}$	CGACGCTCTTCCGATCTCATGTCCGACACGTGTGTCGACTGCGGACGTGTGTGAATTGTC
$S_1'_{M14}$	CGACGCTCTTCCGATCTCATGTCTGACACGTGTGTCAACTGCGGACGTGTGTGAATTGTC
$S_1'_{M16}$	CGACGCTCTTCCGATCTCATGTATGACACGTGTGTCACCTGCGGACGTGTGTGAATTGTC
$S_1'_{M18}$	CGACGCTCTTCCGATCTCATGCATGACACGTGTGTCACATGCGGACGTGTGTGAATTGTC
$S_1'_{M20}$	CGACGCTCTTCCGATCTCATTCATGACACGTGTGTCACAGGCGGACGTGTGTGAATTGTC

- BD Falcon (USA): 50 mL screw cap reaction tubes,

- BioTek (USA): Synergy H1 Multi-Mode Reader,

Table 2.2: DNA sequences of the templates used on the DDMA project (continuation).

Randomized Regions Libraries

$X_{2/10}$	GACAATTCACACACGTCCGCNGNCNGNCNGNTNANTNGNCATGAGATCGGAAGAGCGTCG
$X_{4/5}$	GACAATTCACACACGTCCGCAGNNTGNNTGNNCANNGGNNATGAGATCGGAAGAGCGTCG
$X_{10/2}$	GACAATTCACACACGTCCGCNNNNNGACTGNNNNNTGGACATGAGATCGGAAGAGCGTCG

Restriction Enzymes Libraries

X_{EcoRV}	GACAATTCACACACGTCCGCNNNNNNNGATATCNNNNNNNATGAGATCGGAAGAGCGTCG
X_{SmaI}	GACAATTCACACACGTCCGCNNNNNNNCCCGGGNNNNNNNATGAGATCGGAAGAGCGTCG

Primers

A	GACAATTCACACACGTCCGC
A_{18}	GACAATTCACACACGTCC
A_{15}	GACAATTCACACACG
A_{12}	GACAATTCACAC
A_{20+1}	GACAATTCACACACGTCCGCA
A_{20+5}	GACAATTCACACACGTCCGCATTCG
A_{20+10}	GACAATTCACACACGTCCGCATTCGTAGGT
B'	CGACGCTCTTCCGATCTCAT

Other Aptamer Library

Y_{20}	TAGGGAAGAGAAGGACATATGATNNNNNNNNNNNNNNNNNNNNTTGACTAGTACATGACCACTTGA
C	TAGGGAAGAGAAGGACATATGAT
D'	TCAAGTGGTCATGTACTAGTCAA

- Carl Roth (Germany): sodium acetate,

- Eppendorf (Germany): cell culture incubator Galaxy 170S (95% humidity, 5% CO_2), table top centrifuge (5424R), DNA, and protein 0.5, 1, and 2 mL low binding tubes,

- ForteBio (USA): Octet Red 384, 384-well plates with tilted bottom, anti-penta-His, aminopropylsilane (APS), and amine reactive second generation (ARG) sensors,

- GE Healthcare (UK): Streptavidin Sepharose High Performance, PD10 desalting column, and illustra Microspin G-25 columns,

- Greiner Bio-One (Austria): T25, T75, and T175 cell culture flasks with stand-

Table 2.3: DNA sequences used in the TNF-α binding and activity experiments.

ssDNA

AGGCTGATTCGGTTCATGCGGATCCA

Half-dsDNA

TGGATCCGCATGACATTCGCCGTAAG
CTTACGGCGAATGACCGAATCAGCCT

Full-dsDNA

GACAATTCACACACGTCCGCAGTCTGACTGATCACTGGACATGAGATCGGAAGAGCGTCG
CGACGCTCTTCCGATCTCATGTCCAGTGATCAGTCAGACTGCGGACGTGTGTGAATTGTC

yDNA

TGGATCCGCATGACATTCGCCGTAAG
CTTACGGCGAATGACCGAATCAGCCT
AGGCTGATTCGGTTCATGCGGATCCA

yDNA (no potential CpG motif)

TGGATCCGCATGACATTCGCCGTAAG
CTTACGGCGAATGACCGAATCAGCCT
AGGCTGATTCGGTTCATGCGGATCCA

Xsmall

GGTATACGGTCAGATCCGTA
TACGGATCTGCGTAGTAGCG
CGCTACTACGGCTAGTACGT
ACGTACTAGCACCGTATACC

VR11

GACAATTCACACACGTCCGCTGGTGGATGGCGCAGTCGGCGACAAATGAGATCGGAAGAGCGTCG

Salmon sperm DNA

Undefined, 100 to 200 bp products

ard screw cap ($25\,\text{cm}^2$), and 96-well plates with clear bottom and black walls,

- Heidolph (Germany): MR Hei-Standard stirring plate,

- Hellma Analytics (Germany): High Precision Cell $1\,\text{mm}$ light path quartz cu-

Table 2.4: DNA sequences labelled with Cy5.

Cy5-ssDNA

Cy5-TGGATCCGCATGACATTCGCCGTAAGCGACGCTCTTCCGATCTCAT

Cy5-half-dsDNA

Cy5-TGGATCCGCATGACATTCGCCGTAAGCGACGCTCTTCCGATCTCAT
CTTACGGCGAATGACCGAATCAGCCTCGACGCTCTTCCGATCTCAT

Cy5-full-dsDNA

Cy5-TGGATCCGCATGACATTCGCCGTAAGCGACGCTCTTCCGATCTCAT
ATGAGATCGGAAGAGCGTCGCTTACGGCGAATGTCATGCGGATCCA

Cy5-yDNA

Cy5-TGGATCCGCATGACATTCGCCGTAAGCGACGCTCTTCCGATCTCAT
CTTACGGCGAATGACCGAATCAGCCTCGACGCTCTTCCGATCTCAT
AGGCTGATTCGGTTCATGCGGATCCACGACGCTCTTCCGATCTCAT

Table 2.5: DNA sequences of oligonucleotides coupled to 2a.

ssDNA-2a

2a-CTTACGGCGAATGACCGAATCAGCCT

dsDNA-2a

2a-GGCAGCACTGGACACCGACGCTCTTCCGATCTC
2a-GTGTCCAGTGCTGCC

yDNA-2a

2a-TGGATCCGCATGACATTCGCCGTAAG
2a-CTTACGGCGAATGACCGAATCAGCCT
2a-AGGCTGATTCGGTTCATGCGGATCCA

vette,

- Intavis AG (Germany): peptide synthesis 6 mL columns, 5 mL syringes with filters, and ResPep SL automated solid-phase peptide synthesiser,

- Iris Biotech (Germany): 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC), *N,N*-diisopropylethylamine (DIPEA), and *N,N'*-diisopropylcarbodiimide (DIC), Fmoc-L-Trp(Boc)-OH,

- JPK Intruments AG (Germany): Nanowizard II Atomic Force Microscopy (AFM),

- Lonza (Switzerland): 1X SYBR Green I and II, and SeaKem LE Agarose,

- Lumiprobe (Germany): Cy5-NHS,

- Martin Christ (Germany): ALPHA 2-4 LD plus lyophiliser and Speedvac apparatus 2-25 CDplus,

- Merck KGaA (Germany): acetonitrile (ACN) (analysis grade), dichloromethane (DCM), piperidine, trifluoroacetic acid (TFA), triethylamine (TEA), and sodium dihydrogen phosphate dihydrate (NaH_2PO_4),

- Merck Millipore (Germany): Milli-Q instrument (Milli-Q Advantage A10) with a Milli-Q filter LCPAK0001, syringe filters ($0.22\,\mu m$ for aqueous solutions), heparin sodium salt from porcine intestinal mucosa, Fetal Bovine Serum (FBS), and O-bis-(aminoethyl)ethylene glycol trityl resin,

- NanoEntek (USA): C-Chip disposable haemocytometers,

- Nunc (USA): cryo tube vials, T25, T75, and T175 cell culture flasks with standard screw cap ($25\,cm^2$),

- Nalgene (USA): cryo freezing container,

- Nanoandmore (Germany): cantilever CP-PNPL-PS-C,

- Nanotemper (Germany): Monolith NT.Automated and standard glass capillaries,

- Nanoworld (Switzerland): PNP-TR-Au gold coated silicon nitride AFM probe ($0.08\,N\,m^{-1}$),

- New England BioLabs (USA): purified Bovine Serum Albumin (BSA),

- Plano (Germany): mica,

- Phenomex (USA): Clarity 3u Oligo-RT C18 reverse phase HPLC column,

- Promega (Germany): x-Tracta Gel Extractor,

- Quanta Bio (USA): PerfeCTa SYBR Green SuperMix,

- Raytest (Germany): Stella 3200, and Xstella 1.00 imaging and analysis system,

- Sartorius Stedtim (France): cellulose acetate, polytetrafluoroethylene (PTFE), polyvinylidenefluoride (PVDF) filters, and filter holder,

- Sigma-Aldrich (USA): acetic anhydride, succinic anhydride, N-hydroxysuccinimide (NHS), N-hydroxysulfosuccinimide (SNHS), trypan blue solution, 4,5-dimethoxy-2-nitrobenzyl chloroformate (NVOC-Cl), spermidine trihydrochloride (SpdCl), Tween 20, acetic acid, Dulbecco's Modified Eagle's Medium (DMEM) - high glucose, boric acid, Ethidium Bromide (EtBr), Cycloheximide (CHX), thiazolyl blue tetrazolium bromide (MTT), Corning Costar Spin-X centrifuge tube filters, triethylammonium bicarbonate (TEAB), Actinomycin D (ACT-D), analytical grade water, and sodium phosphate dibasic heptahydrate (Na_2HPO_4),

- Th. Geyer (Germany): ACN (purification grade), and triethylammonium acetate (TEAA),

- Thermo Fisher Scientific (Germany): PikoReal Real-Time PCR system with white 96-well Piko PCR plates and respective optical adhesive films, True-Start Hot Start *Taq* DNA Polymerase, $NaHCO_3$, SmaI, EcoRV, dNTP mix, trypsin/EDTA solution, 15% TBE-Urea Gel, Novex TBE-Urea Sample Buffer

(2X), TNF-α Recombinant Human Protein, ROX, L-glutamine, 10 and 20 bp ladders, DNA gel 6X loading dye, and glycogen,

- VACUUBRAND (Germany): vacuum pump RZ6, and PC 3000 series vacuum controller,

- VWR International (Germany): ethanol, NaCl, NaOH, power source, electrophoresis running chambers, 15 mL reaction tubes, peqSTAR2x thermocycler, 1-[bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate (HATU), 1-hydroxy-7-azabenzotriazole (HOAT), HCl, dimethylformamide (DMF), and 96-well plates for cell culture,

- Waters (USA): ACQUITY analytical Ultra-High Performance Liquid Chromatography (UHPLC) with UV light detector, the ACQUITY UPLC BEH analytical reverse phase C18 column (bead size 1.7 µm, 50 × 2.1 mm), ACQUITY TQ Electrospray Ionisation Mass Spectroscopy (ESI-MS), analytical ACQUITY UPLC OST C18 column, e2695 separation module, and 2998 photodiode array detector system,

- Zeiss (Germany): inverted microscope Zeiss Z1 Observer.

2.1.2 General Annealing Protocol of DNA Constructs

The oligonucleotides were diluted in an annealing buffer solution (10 mM Tris-HCl pH 7.5, 100 mM NaCl, and 1 mM EDTA) and heated for 5 min at 95 °C. Then, they were slowly cooled to room temperature. When indicated, the annealed products were incubated with 1X SYBR Green I.

2.2 DNA-Based Diversity Modelling and Analysis

2.2.1 *Polymerase Chain Reaction*

The PCR amplification was carried in a volume of $50\,\mu\text{L}$ on a thermocycler using the TrueStart Hot Start *Taq* DNA Polymerase (2 U), with $1.5\,\text{mM}$ $MgCl_2$, $150\,\text{nM}$ primers, $0.2\,\text{mM}$ dNTP mix, and $100\,\text{pM}$ template. The temperature protocol was: $10\,\text{min}$ at $95\,^\circ\text{C}$ (initialisation); five or 25 amplification cycles of $15\,\text{s}$ at $95\,^\circ\text{C}$ (denaturation), $30\,\text{s}$ at $60\,^\circ\text{C}$ (annealing), and $20\,\text{s}$ at $72\,^\circ\text{C}$ (elongation); $30\,\text{s}$ at $72\,^\circ\text{C}$ (final elongation); and $10\,\text{s}$ at $20\,^\circ\text{C}$ (final hold).
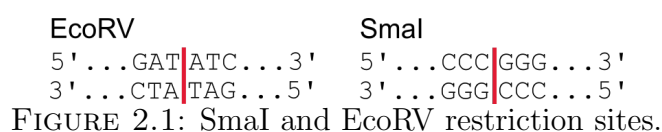
Real Time Polymerase Chain Reaction

The RT-PCR experiments were performed using the PerfeCTa SYBR Green SuperMix, which has AccuStart Taq DNA polymerase. The amplification reaction was carried on a PikoReal Real-Time PCR System. When indicated, we used a Mx3000P RT-PCR system, the associated plates, and ROX. Also when indicated, we used a homemade version of the RT-PCR mix, which consisted of: 5X TrueStart Hot Start *Taq* buffer, $8\,\text{mM}$ $MgCl_2$, $0.3\,\text{mg}\,\text{mL}^{-1}$ BSA, 10% glycerol, $1\,\text{U}\,\mu\text{L}^{-1}$ TrueStart Hot Start *Taq* DNA Polymerase, 1X SYBR Green I, and $400\,\text{nM}$ dNTP mix.

Each reaction well had $10\,\mu\text{L}$, with $150\,\text{nM}$ primers and $100\,\text{pM}$ template. In the primers concentration titration experiments, the primers A and B' final concentration ranged from $500\,\text{pM}$ to $500\,\text{nM}$. In the classic AsyPCR experiments, the primer B' final concentration ranged from $150\,\text{pM}$ to $150\,\text{nM}$. In the template titration experiments, the DNA template concentration ranged from $1\,\text{pM}$ to $10\,\text{nM}$. The temperature protocol was: $10\,\text{min}$ at $95\,^\circ\text{C}$; 20, 30 or 45 amplification cycles of $15\,\text{s}$ at $95\,^\circ\text{C}$, $30\,\text{s}$ at $60\,^\circ\text{C}$, and $20\,\text{s}$ at $72\,^\circ\text{C}$ (data acquisition point); $30\,\text{s}$ at $72\,^\circ\text{C}$; and $10\,\text{s}$ at $20\,^\circ\text{C}$. The results were collected using the PikoReal Software 2.2. All conditions were performed in triplicate and the results are an average of these values.

Restriction Enzymes Experiment

The PCR products were prepared as described before and were diluted: ten times for the five cycles of amplification sample and 100 times for the 25 cycles of amplification sample. Then, $10\,\mu L$ of this dilution were incubated with $0.5\,\mu L$ of a restriction enzyme, SmaI ($10\,U\,\mu L^{-1}$) and/or an isoschizomer of EcoRV ($10\,U\,\mu L^{-1}$), for $30\,min$ at $30\,°C$, $30\,min$ at $37\,°C$ and $20\,min$ at $80\,°C$, see Figure 2.1. After, $1\,\mu L$ of the digested sample was added to the RT-PCR mix and the amplification was monitored for 30 cycles.

```
EcoRV                    SmaI
5'...GAT|ATC...3'    5'...CCC|GGG...3'
3'...CTA|TAG...5'    3'...GGG|CCC...5'
```
FIGURE 2.1: SmaI and EcoRV restriction sites.

Library Diversity Evaluation After a Selection Experiment

The DNA libraries were assembled through annealing in a final concentration of $100\,nM$. Then, $100\,\mu L$ of $10\,\mu M$ libraries were incubated with $10\,\mu L$ (1V) SA sepharose beads, at room temperature for $1\,h$ in a shaker, in selection buffer ($150\,mM$ NaCl, $25\,mM$ NaHCO$_3$, 0.005% Tween 20, pH 9.2). The SA beads were previously washed three times with the selection buffer and the beads slurry was resuspended in selection buffer before being used. After $1\,h$, the suspension was centrifuged ($5\,min$, $1000\,rpm$, $4\,°C$) and the supernatant discarded. Following washing steps, one or two, were performed. The slurry was then subjected to alkaline denaturation ($150\,nM$ NaOH (1V), $3\,min$ at room temperature). Immediately after, the suspension was acidified with $1.5\,M$ acetic acid (0.1V) and again centrifuged. $1\,\mu L$ from the resulting supernatant was added to the RT-PCR mix and the amplification was monitored for 30 cycles.

2.2.2 Fluorescence Development Experiment

We monitored the melting curve of annealed products incubated with 1X SYBR Green I and recorded the interval between 20 and 95 °C. S_1 was used as the limiting strand, to normalise the final annealed product amount among all the constructs. The final melting curve is given as the first derivative of the recorded fluorescence (RFU), after smoothing with the Savitzky-Golay filter.

2.2.3 Electrophoresis

The gels were imaged on a Stella 3200 and Xstella 1.00 imaging and analysis system. Unless stated otherwise, the 10 µL RT-PCR reactions were loaded onto each well.

Denaturing Urea Polyacrylamide Gel Electrophoresis

For DNA resolving, 15% TBE-urea gels were pre-ran in 1X Tris/Borate/EDTA (TBE) buffer (89 mM Tris-borate, 2 mM EDTA, pH 8.4), at 160 V and 10 mA, for 30 min. The amplified samples were directly added 2X TBE Urea Sample Buffer (in a final volume of 10 µM) and heated to 70 °C for 3 min. Then, immediately placed on ice. When extreme denaturation conditions are indicated, a denaturation at 95 °C for 10 min was done instead. The gels ran for 1 h, at 160 V and 10 mA. After the run, they were read at 470 nm (SYBR Green I signal). When indicated, the gels were further stained for 20 min with 1X SYBR Green II in 1X TBE.

Agarose Gels

The samples were ran on a 4% agarose gel, which was stained with 1:20,000 EtBr when indicated. The gel was ran for 1 h, at 80 V in 1X TBE buffer. If the gel was not stained with EtBr, the SYBR Green I signal was monitored, as indicated.

2.2.4 Atomic Force Microscopy Imaging

AFM images of the RT-PCR products were deposited on freshly cleaved mica in the presence of the imaging buffer (300 mM SpdCl, 300 mM NaCl, 20 mM Tris in analytical grade water) [208]. Optimal dilutions were deposited on the mica and kept still for 90 s. The substrate was then rinsed with analytical grade water and dried with a steady flow of nitrogen. The images were obtained in air, by contact mode, using a Nanowizard II. Silicon nitride cantilevers of 200 μm length, with a resonant frequency of about 17 kHz and spring constant of about $0.08\,\mathrm{N\,m^{-1}}$ with a gold coating were used. A second order polynomial function was used to remove background slope and the images are shown as a heat map of the surface's topography. DNA origami constructs were imaged as controls to the imaging process. SpdCl is present for the binding and spreading of ssDNA and dsDNA on the mica surface [209].

2.3 Physical Environment Regulation of TNF-α Activity

2.3.1 TNF-α

Mouse recombinant TNF-α was the standard protein used in this study. When indicated, a commercial human recombinant TNF-α was used. Both domestic and commercial TNF-α bioactivity on L929 cells was tested and it was concluded they had equivalent values, see Figure 2.2.
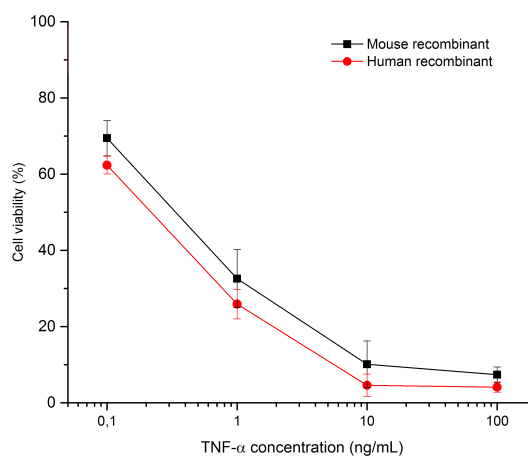

FIGURE 2.2: Cytotoxicity comparison between recombinant mouse and human TNF-α forms.

TNF-α-Cy5   Each trimer was labelled with 15 equivalents (eq.) Cy5-NHS in 0.1 mM NaHCO$_3$ pH 8, in a glass vial with a stirrer, for 2 h on ice. Then, the reaction was quenched with 0.5 mM Tris pH 9 for 20 min, and the product was purified using a PD10 desalting column.

TNF-α-NVOC   Each monomer was labelled with five or 25 eq. NVOC-Cl in 1X PBS pH 7.5, in a glass vial with a stirrer, for 2 h on ice, protected from light. Then, the reaction was quenched with 0.5 mM Tris pH 9 for 20 min and the product was purified using a PD10 desalting column. For the cell assay, the TNF-α-NVOC control was

43

kept away from light exposure and the cleaved TNF-α-NVOC was subjected to a 30 min exposure under a 366 nm lamp.

2.3.2 Anti-TNF-α Small Molecules Synthesis

The anti-TNF-α compound 2a ($K_d = 50\,\mu\text{M}$, Figure 2.3a) was synthesised as in Buller et al. [112], purified by High Performance Liquid Chromatography (HPLC) and confirmed by Mass Spectrometry (MS) ($[M - H]^- = 437.8; M_{calc} = 439.0$). Compound 2c ($K_d = 15\,\mu\text{M}$, Figure 2.3b) was prepared by coupling tryptophan (Trp) (Figure 2.3c) to 2a. 2a-PEG ($[M + H]^+ = 587.0; M_{calc} = 586.0$) and 2c-PEG ($[M + H]^+ = 855.0; M_{calc} = 854.0$) were prepared on a O-bis-(aminoethyl)ethylene glycol trityl resin, in DMF. Two eq. Trp (Fmoc-L-Trp(Boc)-OH), 2 eq. HATU and 2 eq. HOAT were activated for 30 min at room temperature, then added to the resin along with 6 eq. DIPEA. After, Fmoc was removed with a 20% piperidine solution in DMF and another reaction mixture was added to the resin, overnight at room temperature, composed of: 2 eq. 2a, 2 eq. HATU and 2 eq. HOAT. This mixture was activated for 30 min at room temperature and, then, added 6 eq. DIPEA. Considering that the first coupling with Trp was not 100% efficient, the free amino groups still present in the resin were able to react with 2a directly, generating 2a-PEG. The amino groups modified by Trp coupled to 2a and formed 2c. The day after, the resin was cleaved with 3% TFA in DCM. The final products 2a-PEG and 2c-PEG, which corresponds to 2a-Trp-PEG, were purified by reverse phase HPLC and identified by analytical reverse phase UHPLC with a C18 column and ESI-MS.

2.3.3 Oligonucleotides Assembly and Modification
DNA Structures

Several DNA structures based on a previews publication [147] were assembled for the experiments, see Figures 2.4 and Tables 2.3, 2.4, and 2.5.
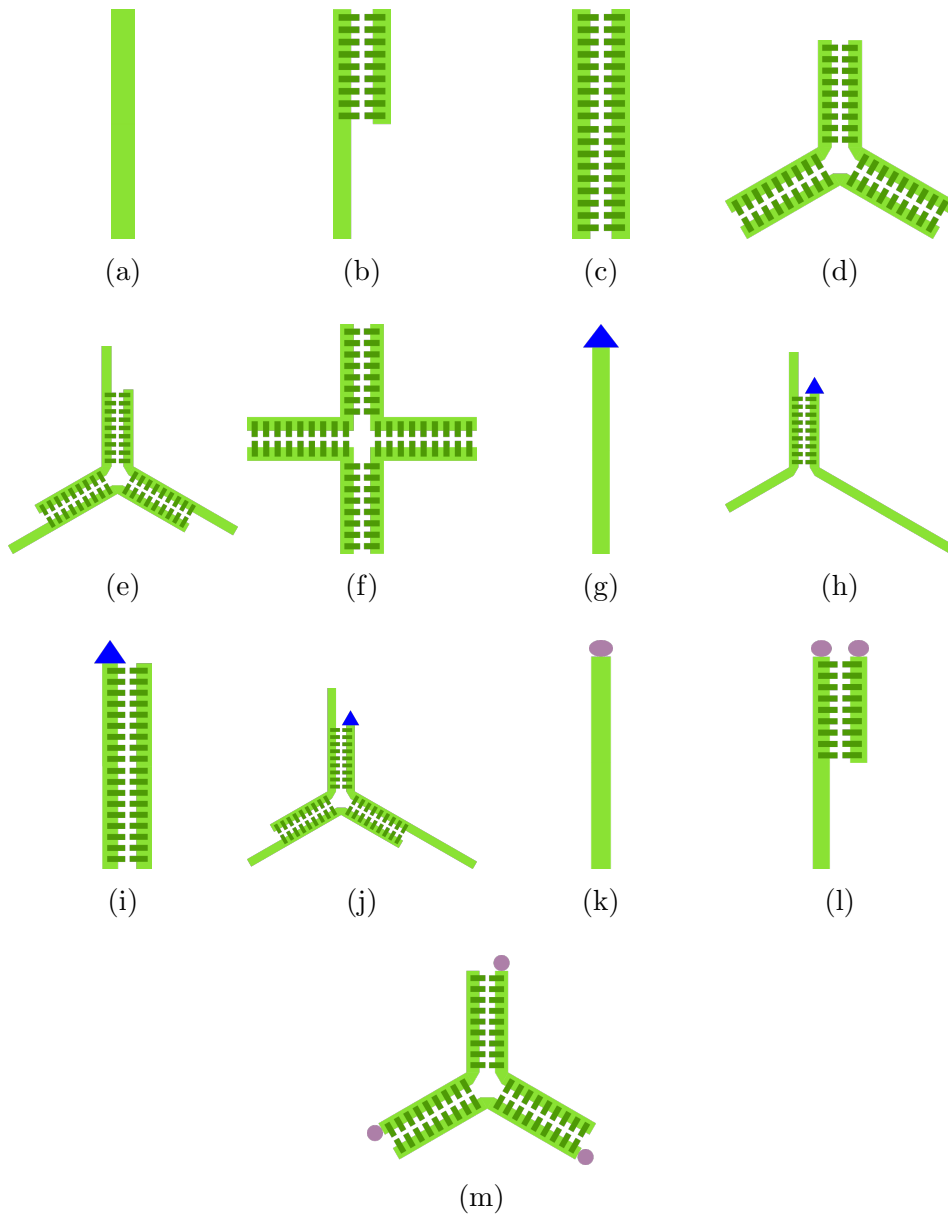
FIGURE 2.3: Small molecules 2a (a) and 2c (c) and Trp (c) structures.

In the band-shift and cell assays, the structures were annealed as described above. For the Circular Dichroism (CD) spectroscopy experiments, the DNA constructs were annealed in $10\,\text{mM}$ sodium phosphate buffer, $\text{pH}\,7.4$.

A $10\,\text{mg}\,\text{mL}^{-1}$ S-DNA solution was prepared in $50\,\text{mM}$ Tris $\text{pH}\,7.5$. S-DNA was used because it could achieve higher concentrations not possible for synthetic DNA structures. To ensure an accurate correlation between S-DNA and defined DNA structures concentrations, we titrated the cytotoxicity of S-DNA from zero to $45\times$ the optical density at 260 nm (OD260) of a $5\,\mu\text{M}$ two half-complementary strands (half-dsDNA) annealed product (Figure 2.5), with $0.1\,\text{ng}\,\text{mL}^{-1}$ of TNF-α. For consequent experiments, it was set that $10\times$ OD260 of S-DNA represented $1\times$ OD260 of half-dsDNA at $5\,\mu\text{M}$.

DNA Coupling to Anti-TNF-α Small Molecules

The oligonucleotide sequences listed on Table 2.5 were coupled through their free 5'-end amino group to the free carboxylic group present in 2a. Around $100\,\text{nM}$ DNA pellet was solved in a $300\,\text{nM}$ TEA/HCl, $\text{pH}\,9$ buffer. The carboxylic group in $20\,\text{mM}$ 2a was activated with $20\,\text{mM}$ NHS and $20\,\text{mM}$ DIC for 30 min, at room temperature on a shaker. Then, DMF and the aqueous solution with DNA were added to the activated 2a, in a final ratio of 70% DMF to 30% water. This mixture was left reacting overnight, at room temperature on a shaker.

45

FIGURE 2.4: DNA structures used in this project: ssDNA (a), half-dsDNA (b), full-dsDNA (c), yDNA (d), yDNA with a short overhang (e), quadruplex (f), Cy5-ssDNA (g), Cy5-half-dsDNA (h), Cy5-full-dsDNA (i), Cy5-yDNA (j), 2a-ssDNA (k), 2a-half-dsDNA (l), and 2a-yDNA (m).

On the next day, we performed ethanol precipitation adding 1/10V sodium acetate and 3V ethanol 100%. The tubes were put at $-80\,°C$ for 1h and then centrifuged for 30 min, 14 000 rpm at $4\,°C$. The pellet was washed with ethanol 70% and again
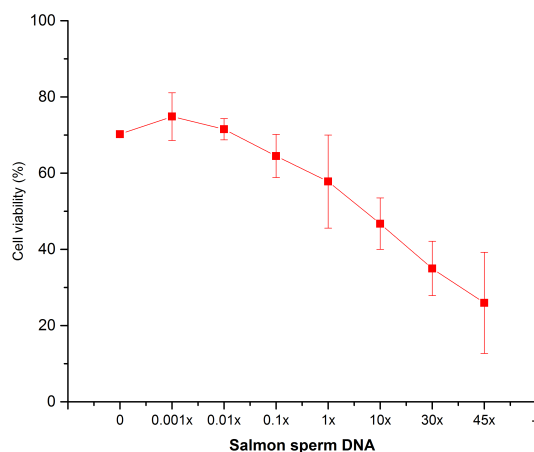
FIGURE 2.5: S-DNA titration considering $1\times$ OD260 a $5\,\mu\text{M}$ half-dsDNA annealed product. TNF-α concentration was $0.1\,\text{ng}\,\text{mL}^{-1}$.

centrifuged for $30\,\text{min}$, $14\,000\,\text{rpm}$ at $4\,^{\circ}\text{C}$. Finally, the pellet was air-dried.

TEAA was added to a final concentration of $100\,\text{mM}$, centrifuged for $10\,\text{min}$, at $14\,000\,\text{rpm}$ and at room temperature. Then, the purification on a reverse phase HPLC column. Buffer A was $100\,\text{mM}$ TEAA in water and buffer B was $100\,\text{mM}$ TEAA in ACN. An elution gradient with buffer B from 5 to 50% for $15\,\text{min}$ at $5\,\text{mL}\,\text{min}^{-1}$ was done. Before the gradient started, the column was flushed with 5% buffer B for $4\,\text{min}$. After, the column was flushed with 95% buffer B, with a fast gradient for $1\,\text{min}$. Then, the column went to the initial condition, with a gradient from 95 to 5% buffer B. The purification was monitored at $260\,\text{nm}$. The oligonucleotides were characterised by analytical reverse phase UHPLC with a C18 column and ESI-MS, with an elution gradient from 5 to 95% $50\,\text{mM}$ TEAB in ACN buffer for $5\,\text{min}$ at $2\,\text{mL}\,\text{min}^{-1}$. The positively identified samples were dried on the Speedvac and resuspended in the desired final volume.

47

2.3.4 Heparin Modification

Heparin (14 kDa) was acetylated in water with 5% acetic anhydride and 2% TEA, for 1 h at room temperature. Then, it was precipitated with tetrahydrofuran (THF) for 20 min at $-20\,^{\circ}$C and centrifuged for 30 min, 8500 rpm at $4\,^{\circ}$C. Heparin-2a and -2c conjugates were prepared by coupling acetylated heparin with 2a- and 2c-PEG. One eq. acetylated heparin, 10 eq. SNHS, and 6 eq. EDC were activated for 5 min at room temperature, in phosphate buffer, pH 7.5. Then, 5 eq. 2a- or 2c-PEG were added. The reaction was left overnight and the coupled product was retrieved by THF precipitation. The final product identity was confirmed and quantified by the increased absorbance at 280 nm.

2.3.5 TNF-α Binding to Biopolymers Assays
DNA Band-Shift

The DNA constructs were incubated with TNF-α for 1 h, at $4\,^{\circ}$C on a shaker and in selection buffer (100 mM NaCl, 20 mM Tris-HCl pH 7.6, 2 mM $MgCl_2$, 5 mM KCl, 1 mM $CaCl_2$, 0.02% Tween 20). The incubation product then was resolved with a electrophoresis run on a 4% agarose gel, which was stained with 1:20,000 EtBr when indicated. The gel was ran for 2 h and 30 min, at 80 V, $4\,^{\circ}$C and with cold 1X TBE buffer. If the gel was not stained with EtBr, the Cy5 signal was monitored. The gels were imaged on a Stella 3200, and Xstella 1.00 imaging and analysis system.

Bio-Layer Interferometry

For the Bio-Layer Interferometry (BLI) assay, anti-penta-His sensors were tested using 1X PBS with 0.02% Tween 20 as an assay buffer and 10 mM glycine, pH 1.5, as regeneration buffer, with a 96-well plate, black and non-binding. Then, increasing concentrations of S-DNA were tested against a constant concentration of TNF-α, in the following way: baseline for 50 s, loading for 100 s, and dissociation for 100 s.

48

The binding of TNF-α to APS versus ARG sensors was tested with 1X PBS with 0.02% Tween 20 as: baseline for 50 s, loading for 300 s, and dissociation for 300 s. Finally, our method was optimised for the ARG sensors in a setup that can be seen on Figure 2.6 and consisted of: baseline for 50 s, loading for 600 s, dissociation for 50 s, and regeneration for 300 s. The tested regeneration buffers were 150 mM NaOH and 10 mM glycine with 2 M NaCl, pH 1.5. The final working buffer was 1X PBS with 0.05% Tween 20. Then, two sensors were treated as controls, incubated only with TNF-α, and two other sensors were treated as test sensors, incubated with TNF-α plus DNA or heparin, being the plotted results an average of these grouped values. Each well, on a 384-well plate, had 50 µL of the assigned solution.
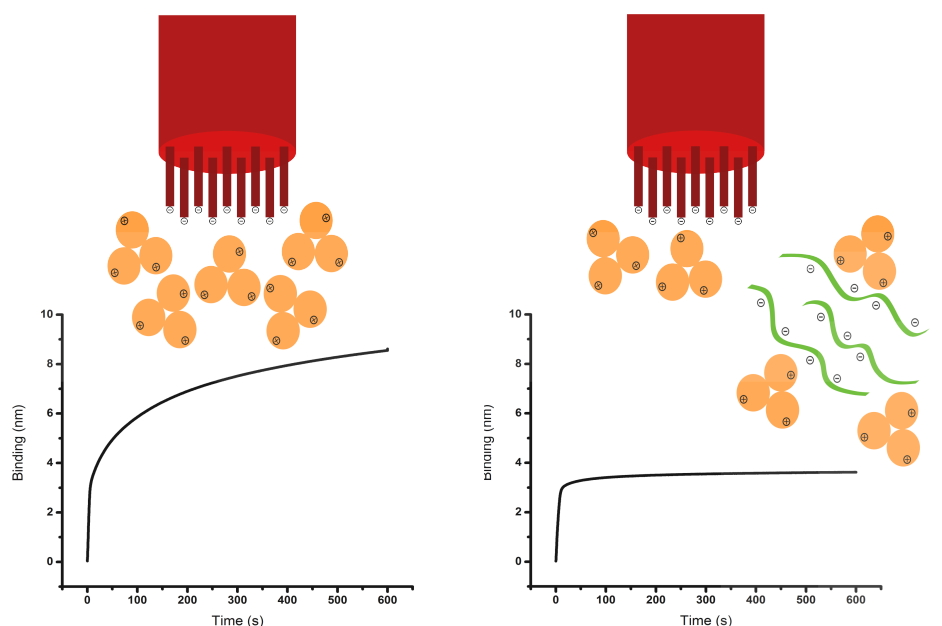
FIGURE 2.6: Scheme of the BLI assay setup. On the left, the negatively charged sensor binds the positively charged free TNF-α. On the right, TNF-α is trapped by interacting with the negatively charged S-DNA and binds less to the sensor.

49

Microscale Thermophoresis

After a short incubation of 122.5 nM to 250 µM TNF-α with 10 nM Cy5-DNA in 1X PBS with 0.05% Tween 20, the samples were loaded onto Microscale Thermophoresis (MST) standard glass capillaries. Then, the analysis was done on the red channel, with MST power 20%, and excitation power 1%, in a Monolith NT.Automated.

2.3.6 TNF-α Cytotoxicity

L929 cells were grown in DMEM supplemented with 2 mM L-glutamine and 10% (v/v) FBS. The cells were seeded in 96-well plates at 30,000 cell per well, in 100 µL complete medium for 24 h before the incubation with TNF-α. Varying concentrations of recombinant mouse TNF-α were pre-incubated for 30 min at room temperature with several testing structures (e.g., DNA, heparin) in media containing 8 µM CHX. After, all medium was removed and 90 µL of each test condition was distributed through the wells. Then, the cells were incubated for 24 h at 37 °C and 5% CO_2. On the next day, a solution of MTT in 1X PBS was added to each well to a final concentration of 0.5 mg mL^{-1} and incubated for another 90 min at 37 °C, 5% CO_2. Finally, all medium was removed and 1V dimethyl sulfoxide (DMSO) was used to solubilise the formazan crystals. Their absorbance was read at 540 nm. Cell viability (CV) was calculated as,

$$CV(\%) = \frac{X - B}{C - B} \times 100$$

Where B is the blank (wells only with medium), C is the control (wells with medium plus CHX) and X is the absorbance from the test wells.

2.3.7 TNF-α Structural Changes Monitoring

Circular Dichroism Spectroscopy

CD spectroscopy measurements were performed at 25 °C, after incubation of DNA with TNF-α in 10 mM sodium phosphate buffer, pH 7.4, for 60 min. The data was collected in the interval from 190 to 260 nm and the final spectra were an average of five scans.

The DNA spectra were collected alone. Then, a highly concentrated TNF-α solution was added directly to the 1 mm cuvette, followed by stern shaking. TNF-α and DNA interaction was monitored at 0, 30 and 60 min. The quantification of TNF-α's α-helix was done at 208 and 222 nm, and β-sheet at 217 nm [192].

Homoquenching

The incubation of 2 µM TNF-α with 10 nM TNF-α-Cy5 and 2 µM S-DNA during 120 min was done in a 384-well plate, with clear bottom and black walls. Each well had 80 µL for each tested condition and the working buffer was 1X PBS. The Cy5 signal was monitored and registered as an average of three replicates between 640 and 670 nm, with 150 gain on a Synergy H1 Multi-Mode reader.

3

Results

3.1 DNA-Based Diversity Modelling and Analysis

3.1.1 Diversity Analysis

Library Settings The oligonucleotide sequences used in this project are summarised on Tables 2.1 and 2.2, and illustrated on Figure 3.1. The standard DNA templates were 60 nucleotides-long, with the first 20 (A; forward primer region) and last 20 (B; reverse primer region) nucleotides being consistent in all sequences. In the templates middle region, there was a segment named X, which varied according to what is described below:

- aptamer libraries: 20 nucleotides that possess from two to 20 randomised positions,

- high diversity libraries: at least 15 out of 20 nucleotides were different from each other,

- S_1 mismatched complementary templates: a complementary sequence of S'_1, S_x, was increasingly degenerated, from two to 20 mismatches,

- randomised regions libraries: randomised nucleotides were distributed in every other position ($X_{2/10}$), two every two positions ($X_{4/5}$), and in two groups of five intercalated with constant nucleotides ($X_{10/2}$),

- and restriction sites libraries: the six nucleotides in the middle carried the restriction sites, for EcoRV (X_{EcoRV}) and SmaI (X_{SmaI}), and were surrounded by randomised positions.

There were also some structural variations in the primers design:

- for the non-classic asymmetric experiment, forward primers with shorter lengths: A_{18}, A_{15}, and A_{12},

- for the allele-specific experiment, primers with extra nucleotides: A_{20+1}, A_{20+5}, and A_{20+10},

- and for the experiments with longer primers, primers with ten nucleotides-long 5'-overhang: A_{10+20} and B_{10+20}.

Diverse Libraries

Diverse libraries from two to 20 randomised nucleotides were screened, see Figure 3.2. The library X_2, with 4^2 sequences, had an amplification curve similar to the standard curves where only one template is used. The sigmoid-shaped curve for lower diversity samples kept its overall trace, only lowering the plateau of final fluorescence value until the X_8 library. From X_{10} to X_{20}, the curve became bell-shaped, reaching a maximum value of fluorescence and decreasing to almost zero.
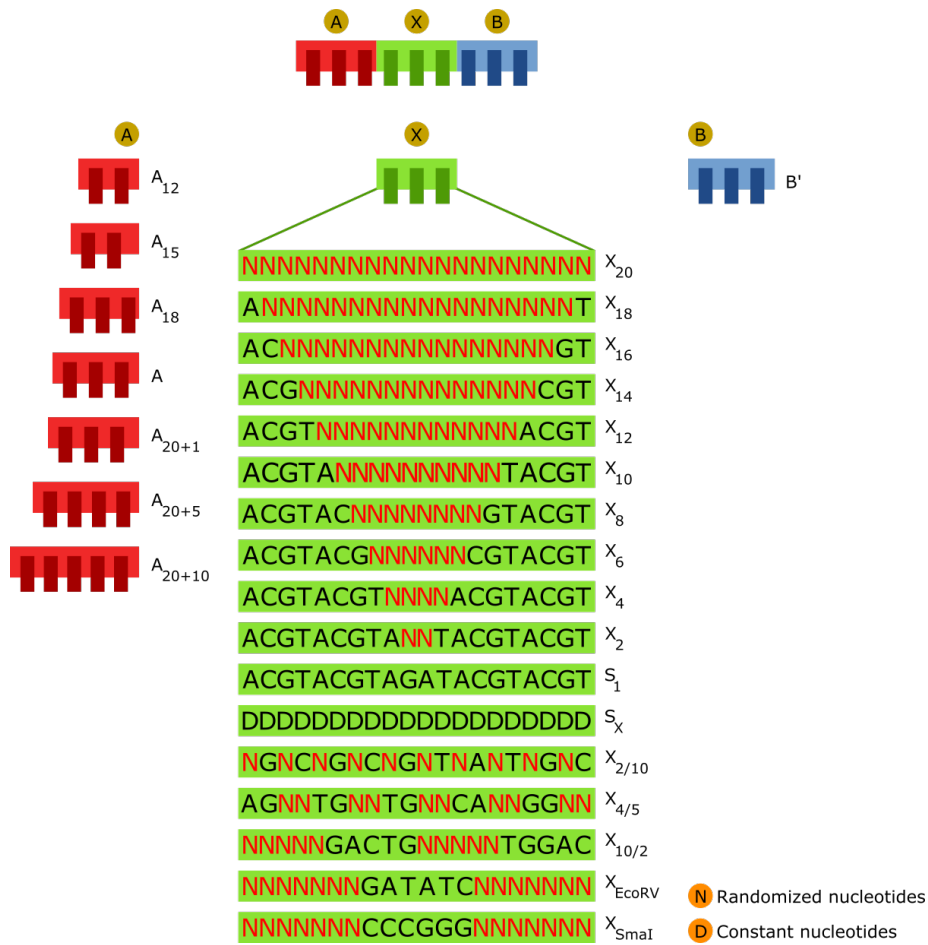
FIGURE 3.1: DNA sequences schematics.
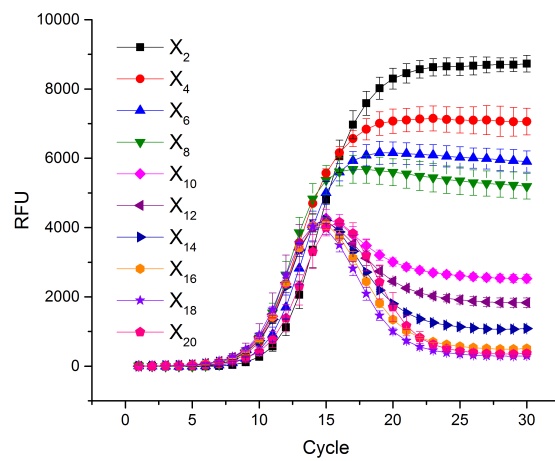
FIGURE 3.2: RT-PCR experiment using DNA libraries of different diversities (from two to 20 random positions) as templates.

54

The DNA templates were titrated from 1 pM to 10 nM and the curves' shapes were consistent for all concentrations and templates. There were only shifts proportional to the templates' initial concentration, see Figure 3.3. This demonstrates that the different curves shape were not affected by templates' concentration in any of the libraries tested, S_1, X_{10}, and X_{20}.
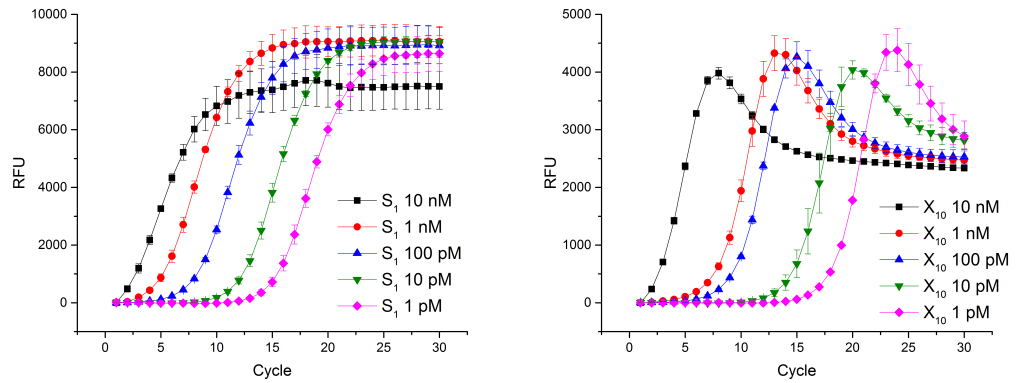
(a)   (b)
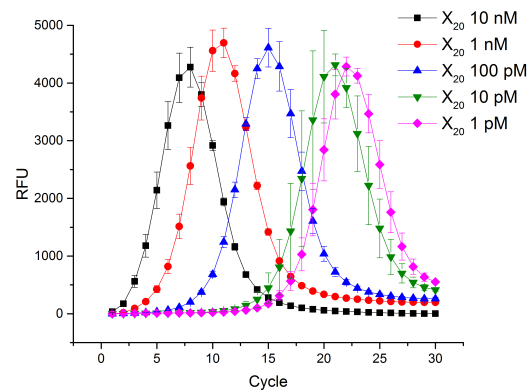
(c)

FIGURE 3.3: Library titration from 1 pM to 10 nM for S_1 (a), X_{10} (b), and X_{20} (c).

Primers Titration

Standard PCR protocols use high concentrations of primers to assure robust amplific-ation. We worked inside a range where primer concentrations had a linear correlation with the fluorescence signal. Also, we wanted to verify that this curves' effect was not influenced or was an artefact of concentration. Both primers concentration was titrated from 500 pM to 500 nM (Figure 3.4). As expected, due to the decreased con-centration, the fluorescence signal had a strong drop linked to lower primer amounts. We have chosen 100 nM primer concentration for the following experiments because it is within the range where primer concentrations are correlated linearly with the fluorescence signal.
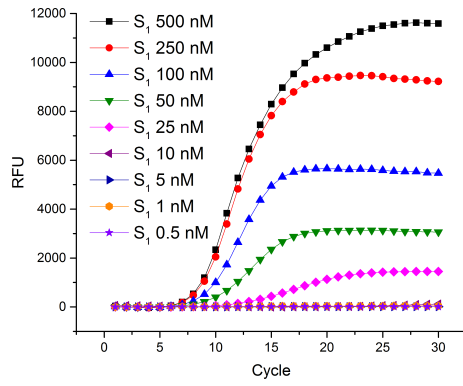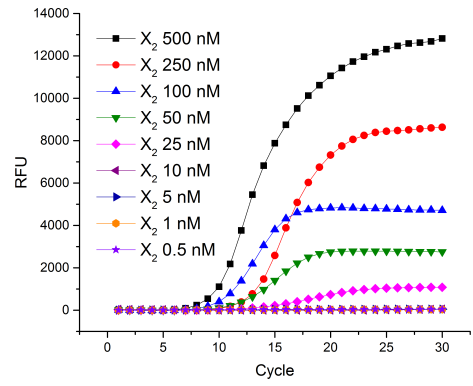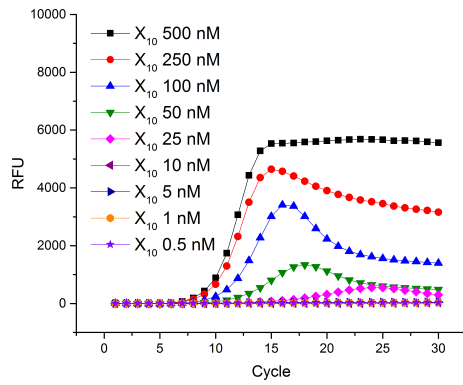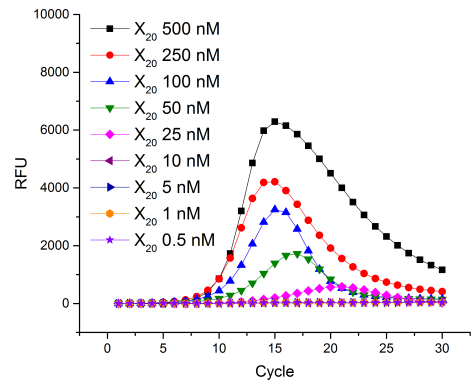
FIGURE 3.4: Forward (A) and reverse (B') primers titration for S_1 (a), X_2 (b), X_{10} (c), and X_{20} (d) libraries, with final concentrations ranging from $500\,\mathrm{pM}$ to $500\,\mathrm{nM}$ for both primers.

Annealing Step Manipulation

To better understand how our system was working and how it could be manipulated, we checked the influence of the annealing step. The annealing step seemed to be a good place to start to comprehend in more detail the amplification dynamics of diverse libraries. The standard annealing step was $30\,\mathrm{s}$, at $60\,^{\circ}\mathrm{C}$ and $7\,^{\circ}\mathrm{C\,s^{-1}}$.

Annealing Time For all templates tested, S_1, X_{10} and X_{20}, when the annealing step was 60 s long, amplification process was faster than the 30 s long experiment (Figure 3.5). Moreover, the amplification curve shape was kept.

(a)

(b)

(c)
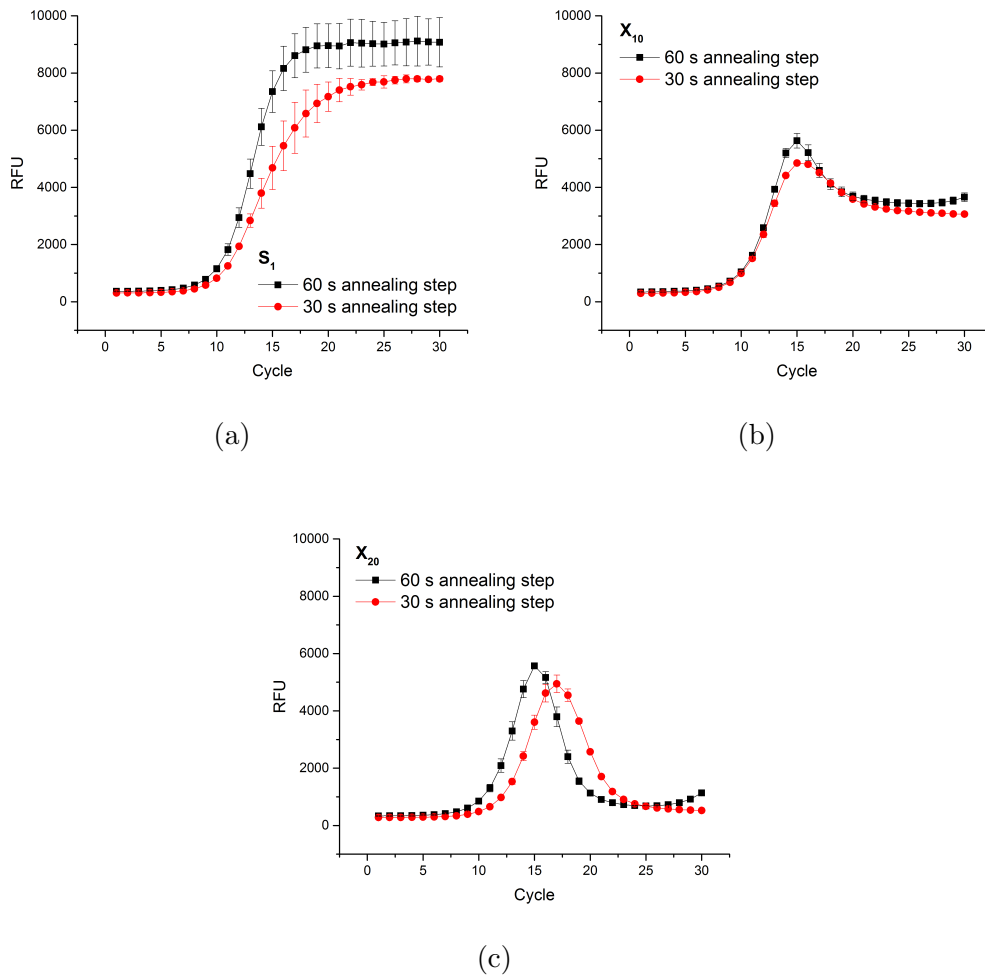
FIGURE 3.5: Annealing time of 30 s versus 60 s for S_1 (a), X_{10} (b), and X_{20} (c).

Annealing Temperature For X_2, X_{10} and X_{20}, the lower annealing temperature of 51 °C generated lower fluorescence signals and there was a shift in the amplification course (Figure 3.6). There was less amplified product at this temperature, than at 60 °C, because amplification became less efficient. The curves shapes were main-

tained.
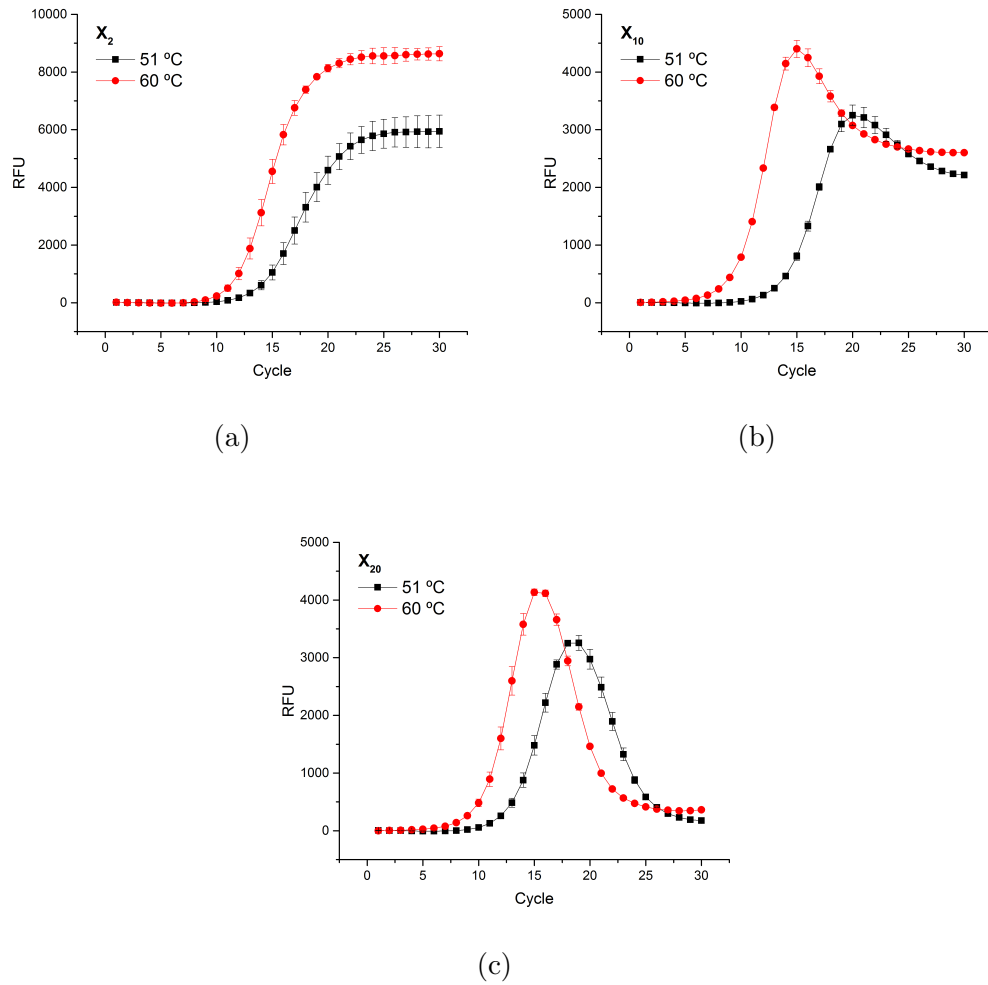
FIGURE 3.6: Testing annealing temperatures of $51\,^{\circ}\mathrm{C}$ versus $60\,^{\circ}\mathrm{C}$ for X_2 (a), X_{10} (b), and X_{20} (c).

Annealing Speed   When the annealing speed was lower, S_1 reached a higher fluorescence plateau than at a slightly faster speed (Figure 3.7). However, this effect was not seen for X_{10} and was even inverted for X_{20}. The amplification curves kept the same shape.
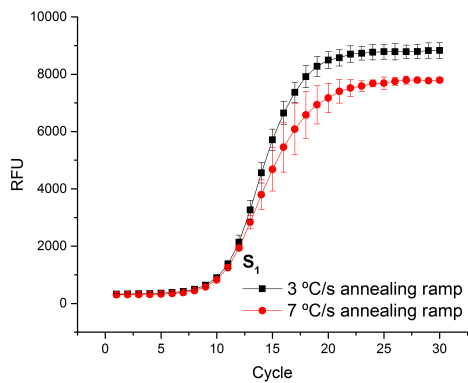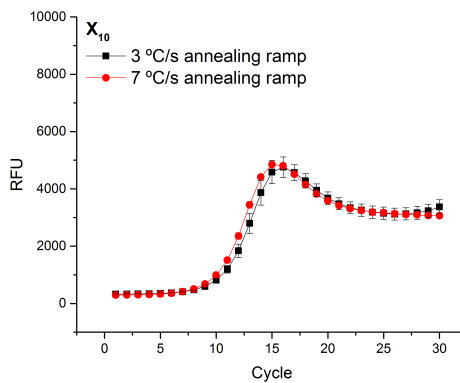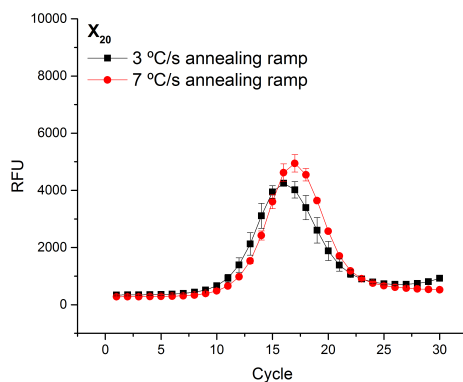
FIGURE 3.7: Annealing speed testing of $7\,^{\circ}\mathrm{C}\,\mathrm{s}^{-1}$ versus $3\,^{\circ}\mathrm{C}\,\mathrm{s}^{-1}$ for S_1 (a), X_{10} (b), and X_{20} (c).

Although the course could be slightly affected by annealing time, temperature, or speed, the correlation between curve shape and library diversity remained unchanged.

Fluorescence Development

The fluorescence development in these experiments was hypothesised to being related to the formation of increasingly not fully complementary duplexes in a partially degenerated library. As schematised on Figure 3.8, in initial PCR cycles, DNA duplexes are fully complementary because they are generated by DNA polymerase.

60

The fluorescent dye has higher affinity for duplexes and intercalates it perfectly, so double strands have a maximum fluorescence value of functional factor (F), which is assigned to 1.0. In later PCR cycles, the DNA double strands become more and more mismatched, meaning that the fluorescent dye has progressively less affinity for DNA. F becomes smaller than one (path 2). This fluorescence development process was also simulated mathematically (Figure 3.9) and it fitted the experimentally obtained results.
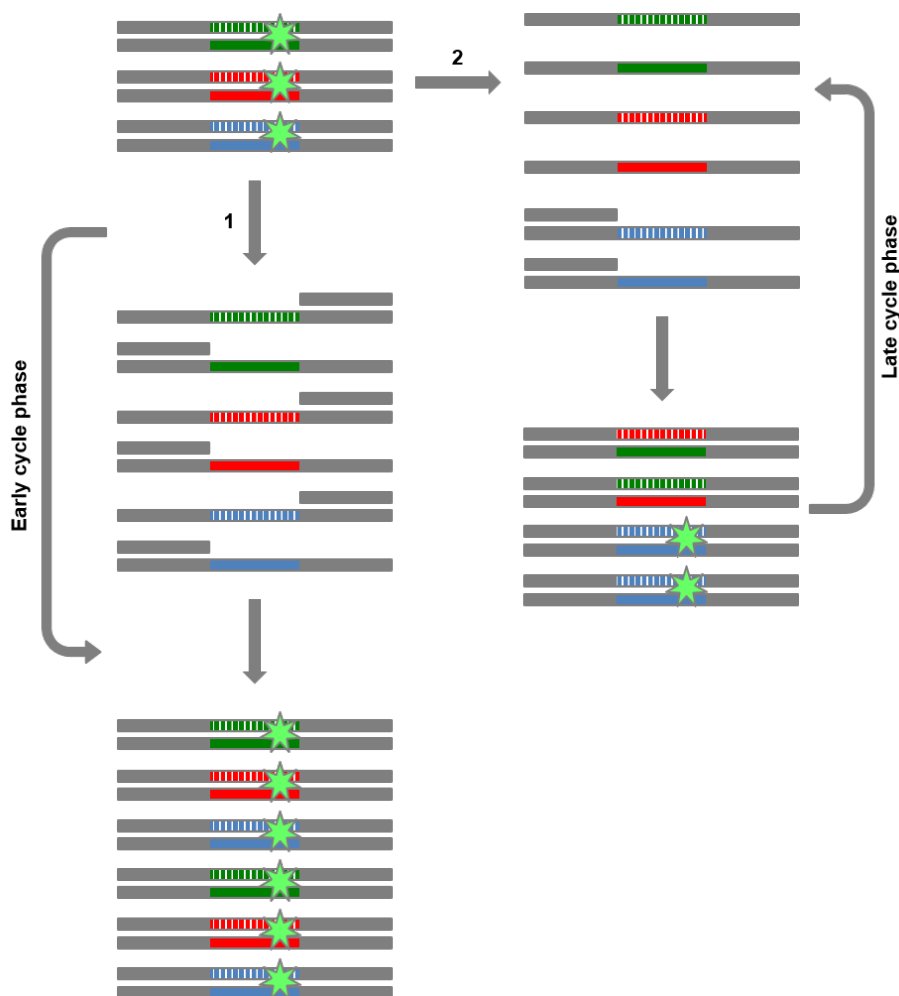
FIGURE 3.8: Formation of fully complimentary duplexes during the initial PCR cycles (1) and mismatched duplexes in later PCR cycles (2). The mismatched duplexes have lower F factor and lower affinity to fluorescence dye (green star) at high temperatures (e.g., $72\,°\mathrm{C}$ of elongation temperature).
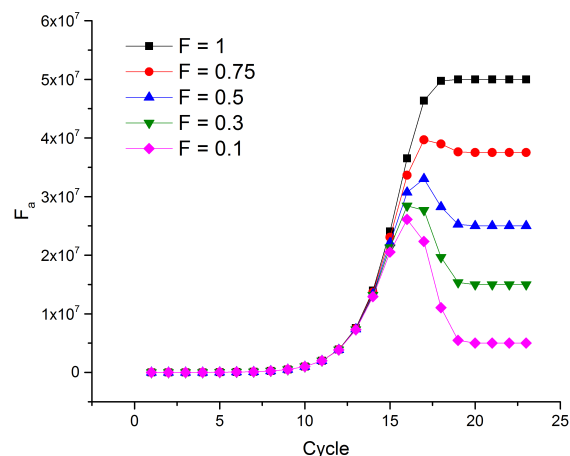
FIGURE 3.9: PCR amplification simulation with theoretical DNA libraries of different diversities.

This phenomenon was tested using S_1 *mismatched complementary templates*, see Table 2.1. Eleven annealed products were prepared: from a fully complementary sequence, S_1 with S_1', to a completely mismatched construct, S_1 with $S_1'_{M20}$. $A-X-B$ and $A'-X'-B'$ quick annealing generated mismatched duplexes that were less stable at $72\,°C$, which corresponds to the elongation temperature in PCR amplification. This lower double strand stability caused weaker dye binding, lower melting temperatures, thus, lower fluorescence, see Figure 3.10. One can observe a decreasing trend in melting temperature and fluorescence intensity from the most stable DNA duplex, which was fully complementary, to the less stable, with 20 mismatches. Single-stranded S_1 and 1X SYBR Green I showed very low to no fluorescence in the recorded interval. The melting curves mirrored the trend observed in the first diverse libraries experiment, see Figure 3.2. The sample S_1 with its complementary presented the highest melting temperature of the tested samples, provided by the highest number of matches. Then, S_1 with two or four mismatches had a similar melting profile. After, six, eight, ten, 12, 14, 16, 18, and 20 mismatches showed gradually lower melting temperatures. We observed that fluorescence intensity and melting

62

temperature are negatively correlated with the sequence mismatching, even when there was the same amount of DNA duplexes, as it had been observed by Schuetze et al. [56].
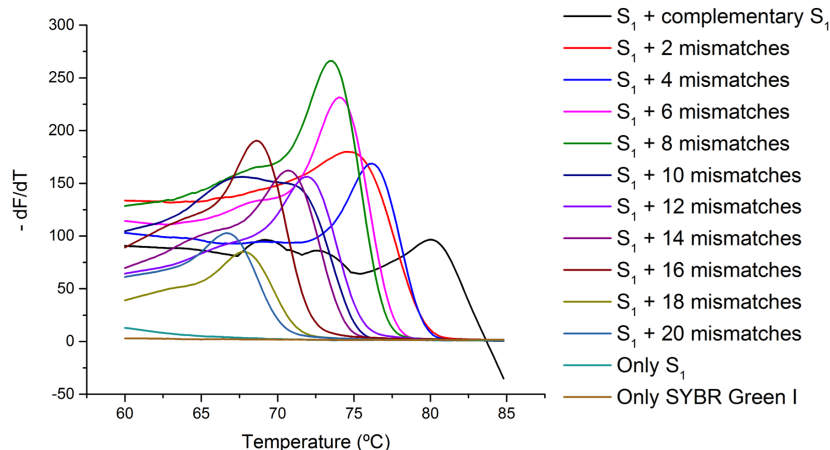

FIGURE 3.10: Savitzky-Golay filter smoothing of the first derivatives of mismatched libraries melting curves. The results correspond to a final concentration of 100 nM annealed products that were incubated with 1X SYBR Green I. The results shown are an average of three values.

Templates with Discrete Randomised Regions

Mathematically, X_{10}, $X_{10/2}$, $X_{4/5}$, and $X_{2/10}$ libraries have the same diversity of 4^{10} different sequences, but the randomised nucleotides are distributed differently (Table 2.2). As compared to X_{10}, we could demonstrate that $X_{10/2}$, $X_{4/5}$ and $X_{2/10}$ had similar amplification curves, even though the degenerated segments are disposed very differently, see Figure 3.11. For any given library X_n, when n is larger than 16, each sequence in a 10 µL 1 nM sample is statistically unique. But, each sequence in the randomised regions template is represented several thousands of times. X_{20}, X_{18}, X_{16}, $X_{2/10}$, and $X_{4/5}$ have 18 to 20 nucleotides-long randomisations, and the higher diversity libraries have curves very different from the medium diversity ones.
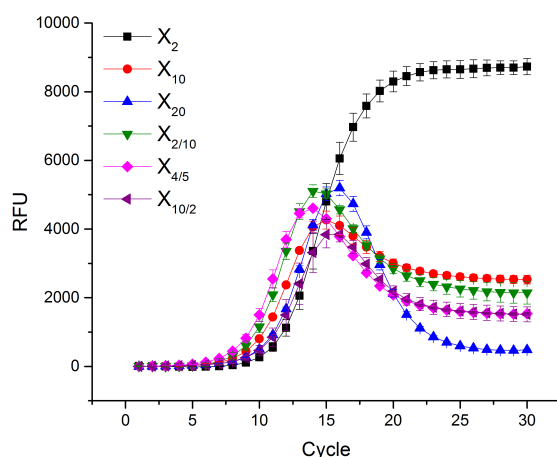
FIGURE 3.11: RT-PCR experiments using different libraries of same diversity. X_{10}, $X_{10/2}$, $X_{2/10}$, and $X_{4/5}$ have the same diversity, but different distribution of partially degenerated segments inside the sequence.

High Diversity Templates

The high diversity templates were designed so, at least, 15 out of 20 nucleotides in the X region would be different. This design is in contrast with X_{20}, for example, where the difference between two sequences can range from zero to 20. This new highly diverse library contains n sub-libraries, where members of each library are identical and the difference between sub-libraries is very high. To test the binding properties of fully matched and mismatched sequences to the dye, two annealed products were prepared: S_1 with S_1', a fully complementary duplex; and S_{10} with S_1', a highly mismatched duplex, with only four matches (Figure 3.12). Then, an agarose gel was ran. The annealed product 1 had a stronger band intensity than the product 2, which showed a higher amount of fully complementary product over the mismatched one.
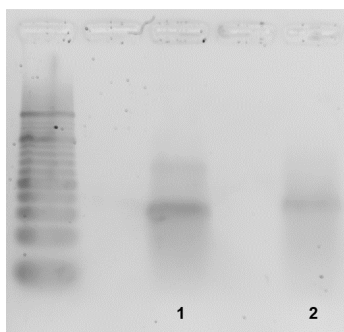
FIGURE 3.12: The correct annealed product is 60 bp: S_1 annealed with S_1' (1) and S_{10} annealed with S_1' (2). $1\,\mu M$ annealed DNA in a 4% agarose gel stained with 1X SYBR Green I with a 20 bp DNA ladder.

We then proceeded to the amplification curves monitoring, see Figure 3.13. Only one sequence template generated a sigmoidal curve with fluorescence signal higher than the two templates mixture at the plateau phase. Further mixing with high diversity templates was done, from two to ten, and the fluorescence signal emulated what happened to the partially degenerated libraries with diversities up to X_8 (4^8 different sequences). For the ten sequences mix, the fluorescence at the plateau decreased to half value of the maximum reached with one sequence.

The fluorescence signal development was influenced by the templates diversity and not the number of members present, because a mix of ten high diversity sequences, with, at least, 15 out of 20 different nucleotides, had the same amplification behaviour as an aptamer library with 65,536 sequences, with zero to eight nucleotides possibly different.
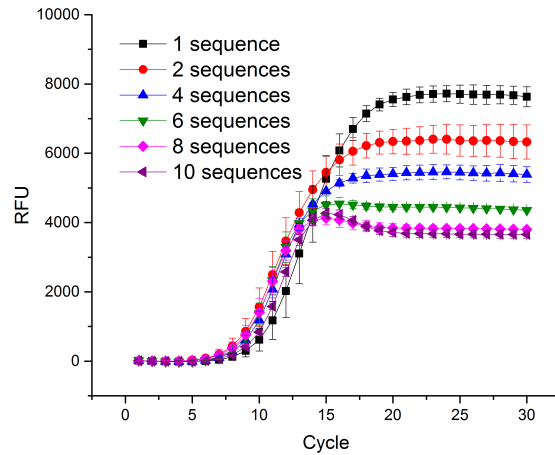
FIGURE 3.13: RT-PCR experiment using a small number of highly diverse sequences, where sequences have been designed to have a large difference (more than 15) between themselves.

3.1.2 Libraries Manipulation Through Primer Design

The standard amplification was done using 20 nucleotides-long primers, A and B', which created 60 bp products. In this section, we explore the impact of primers length, sequence and concentration imbalance on the amplification behaviour for the different templates.

Longer Primers

To understand if the amplification behaviour would be the same for a longer template using longer priming regions, with higher melting temperatures, we designed 30 nucleotide-long primers, A_{10+20} and B'_{10+20}. This way, the newly formed DNA strands were 80 bp (Figure 3.14). As expected, longer amplified products yielded higher fluorescence values, which was seen for S_1, X_{16} and X_{20}. Interestingly, for X_{16} and X_{20}, the fluorescence after the peak did not go as low as in the 60 bp control, altering the curves' shapes.

With longer primers, the resulting duplexes are more stable. Thus, their binding to

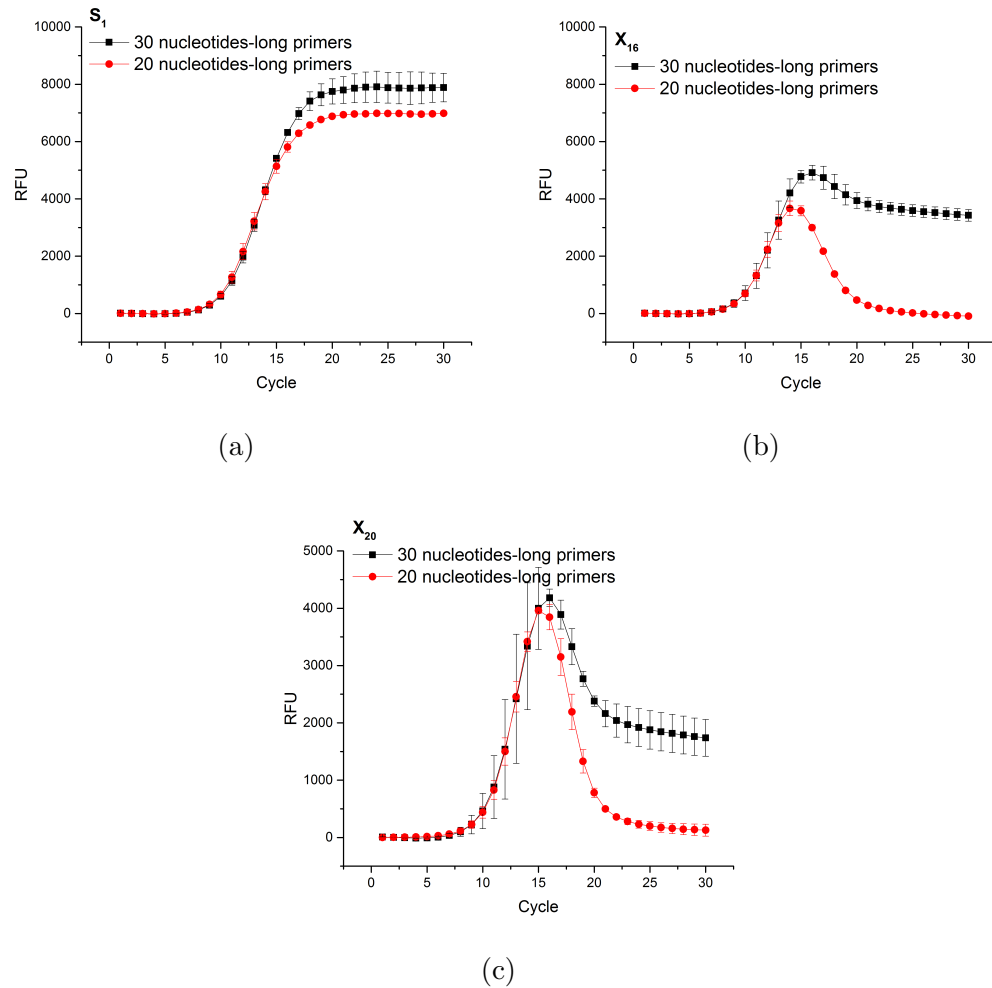the fluorescent dye at an elevated temperature is less affected by mismatching.

FIGURE 3.14: Amplification curves for S_1 (a), X_{16} (b), and X_{20} (c) while using 30 nucleotide-long primers.

Classic Asymmetric PCR

As PCR amplification is a process of population growth, the two primers needed can be seen as pivotal resources. Being these two resources inter-dependent, the limitation of only one can be detrimental to population growth. Classic AsyPCR is a common protocol for ssDNA generation and it is based on the imbalance of primer concentration. One primer is in excess in relation to the other, so one strand is

preferentially amplified [30]. In this experiment, we wanted to observe how primer imbalance could affect the amplification of diverse libraries.

For the S_1 and X_{20} templates, the decreased primer B' concentrations led to a lower molarity of the system, which generated lower fluorescence. The amplification was only efficient when B' concentration was ten times lower, $15\,\mathrm{nM}$. For $1.5\,\mathrm{nM}$ and $150\,\mathrm{pM}$, the amplification could not be detected, see Figure 3.15. The amplification was directed by the consumption of one primer, or lack thereof, hampering the comparison between libraries of different multiplicity. Further conditions were tested, from $500\,\mathrm{nM}$ to $500\,\mathrm{pM}$. In all conditions, the amplification behaviour was only similar to the standard conditions down to ten times less B' (Figure 3.16).
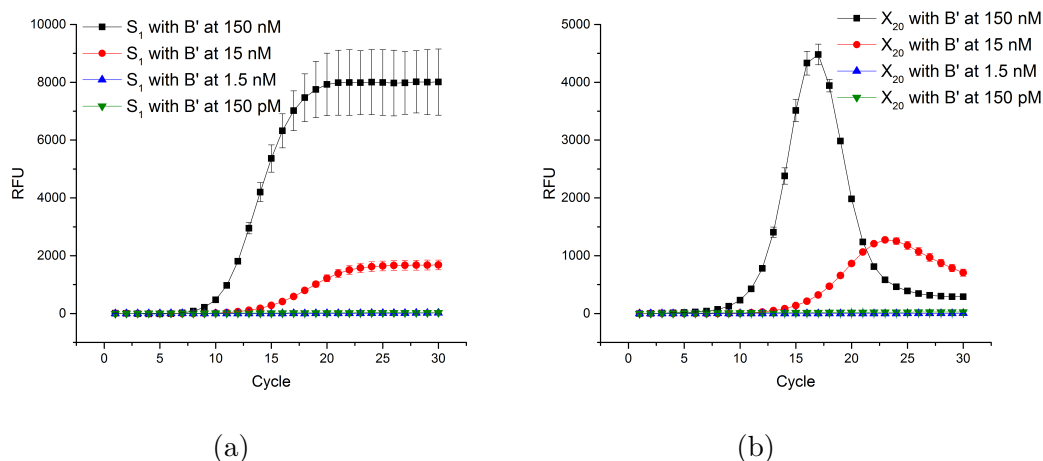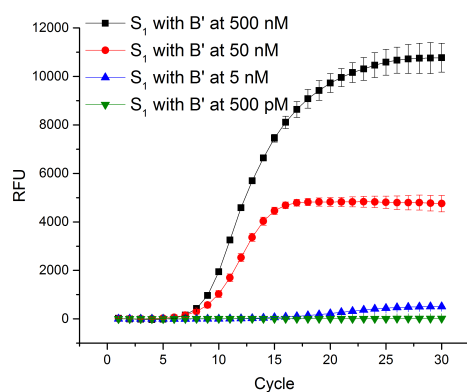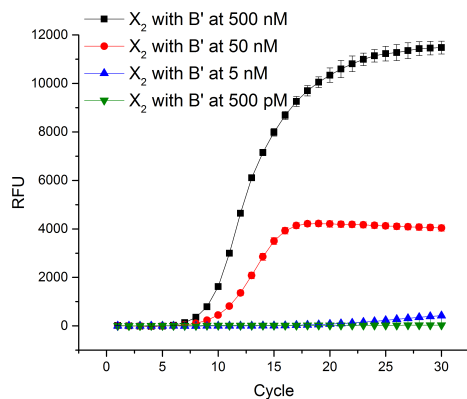

(a) (b)

FIGURE 3.15: Classic AsyPCR amplification curves for S_1 (a) and X_{20} (b). A was in $150\,\mathrm{nM}$ final concentration, while B' concentration ranged from $150\,\mathrm{pM}$ to $150\,\mathrm{nM}$.
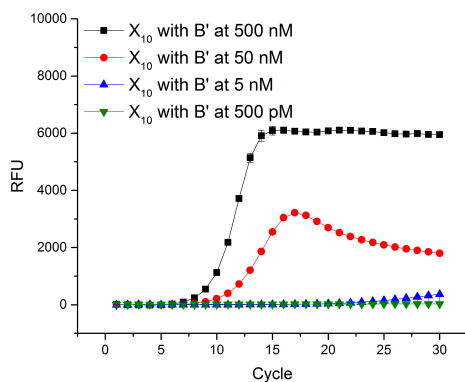
Non-Classic Asymmetric PCR

We then designed a AsyPCR experiment using primers with shorter lengths [29]. Thus, the new shorter A had lower annealing temperature. When combining it with the full-length B', the polymerisation started by shorter primers was the rate-limiting step. The primers inefficiency was then linked to their length, A_{18}, A_{15} and A_{12} (Fig-
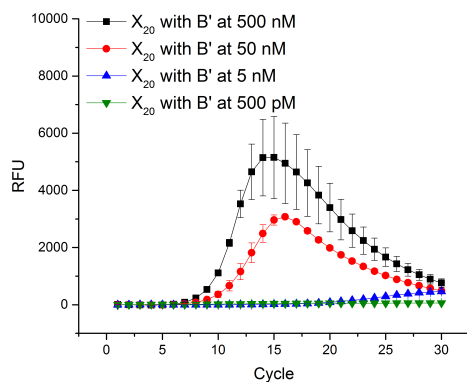
FIGURE 3.16: Classic AsyPCR amplification curves for S_1 (a), X_2 (b), X_{10} (c), and X_{20} (d). A was in $150\,\text{nM}$ final concentration, while B' concentration ranged from $500\,\text{pM}$ to $500\,\text{nM}$.

ure 3.17). In all libraries tested, S_1, X_{10} and X_{20}, the amplification with A_{18} was similar to the full-length A, only with lower fluorescence values. Probably because the amplified product was shorter by two nucleotides, leading to lower fluorescence. When A_{15} was used, amplification was possible. With A_{12}, there was no amplification in every condition.
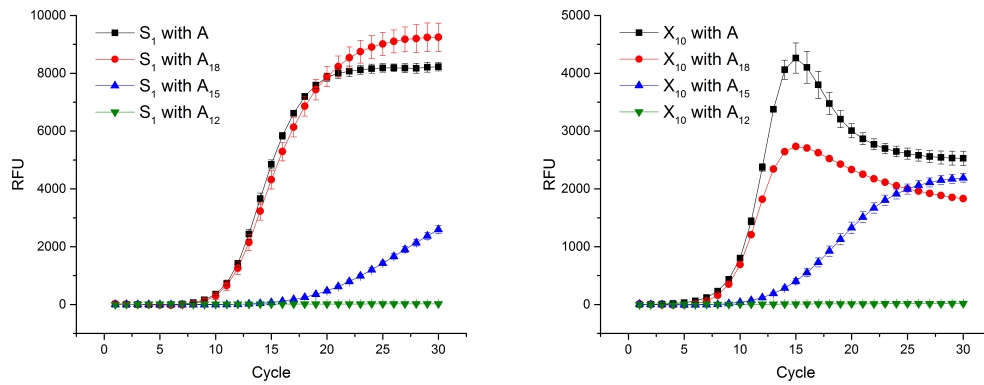
(a)

(b)
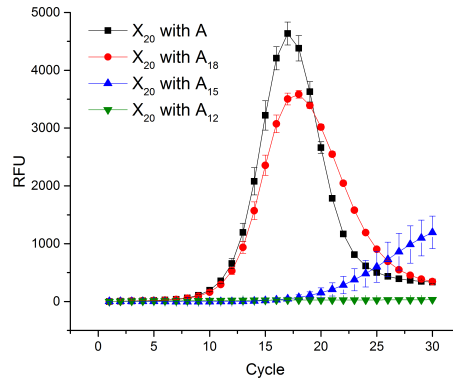
(c)

FIGURE 3.17: RT-PCR using two primers of same or different length (B' with A, A_{18}, A_{15} or A_{12}) using S_1 (a), X_{10} (b), or X_{20} (c) as templates.

In parallel, we simulated this setup with synthetic data and the results obtained were very similar to the experimental data (Figure 3.18). For one sequence, medium and large libraries, when the primer became ten times less efficient, the amplification was kept, but with reduced amplification rates. When the primers were only 1% efficient, the amplifications were remarkably suppressed. The simulation anticipated that the fluorescence peak would be delayed and with lower F. The synthesis of perfect matched duplexes became slower. So, the resulting increase of fluorescence

was overcome by the counter effect associated with DNA reshuffling at lower full length DNA concentration. This effect was particularly strong for the high diversity library.
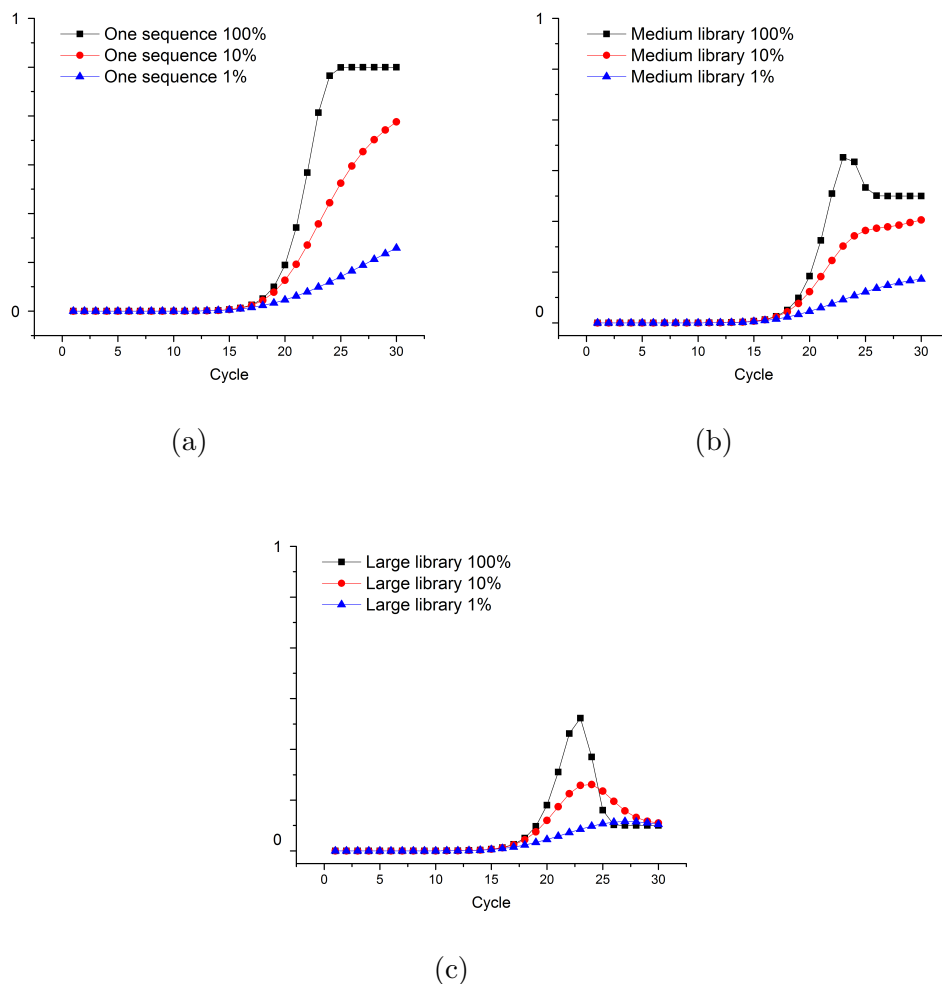

(a)

(b)

(c)

FIGURE 3.18: Simulation of amplification curves of one sequence (d), medium library (e), and large library (f), where amplification efficacy goes from 1 to 100%.

We further imaged the products of this asymmetric amplification (Figure 3.19). SYBR green I binds preferentially to dsDNA, while SYBR green II binds to both ssDNA and dsDNA. Therefore, the difference between these two staining method reflects the ratio between single and double stranded DNA. Consistent with the

amplification curves, after 20 cycles, A and A_{18} generated similarly correct size DNA, which could be stained by both SYBR Green I and II. While A_{15} and A_{12} were unable to generate detectable signal. Interestingly, when A_{15} concentration was doubled, there was some ssDNA at 60 bp being amplified. After 40 cycles, A and A_{18} generated some products, with higher to much higher molecular weights then the template, creating a smear. A_{15} generated a 60 bp product, and when in double concentration, it generated also extra products with heavier molecular weight. A_{12} continued to not be able to carry on the amplification.
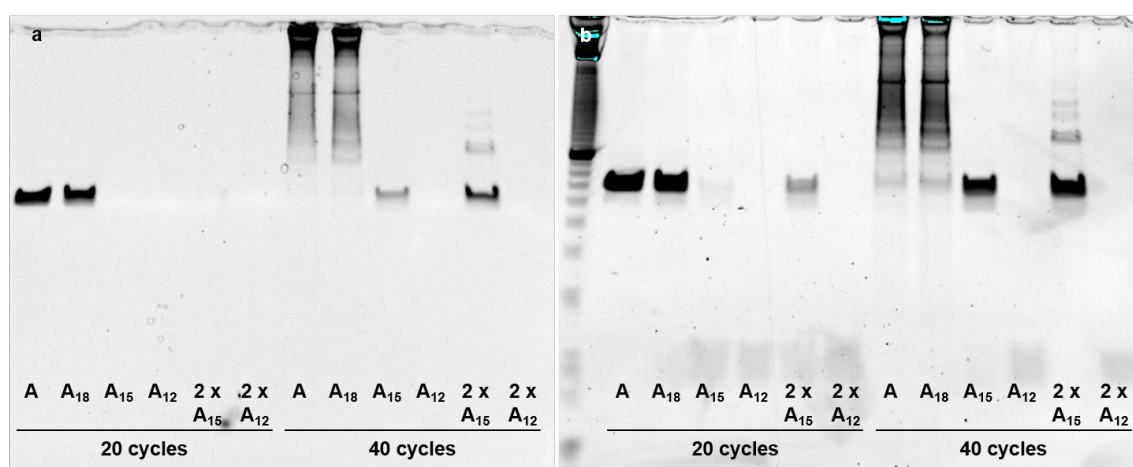

FIGURE 3.19: RT-PCR using primers of same or different length (B' with A, A_{18}, A_{15}, A_{12}, twice A_{15}, or twice A_{12}) using S_1 as template. On the left, there is the 1X SYBR Green I signal, which has more affinity to dsDNA and on the right there is the 1X SYBR Green II signal, which has more affinity to ssDNA. 10 bp DNA ladder, 15% TBE-urea gel.

Biased Population Growth

The next step in our investigation was to go in the opposite direction and use longer versions of A, in order to bias the DNA amplification in a sequence-controlled manner, just as in AS-PCR [33]. Using the principles of AS-PCR, we designed an experiment where we could manipulate and pick sequences that followed rules imposed by the primers from diverse libraries (Figures 3.20 and 3.21).
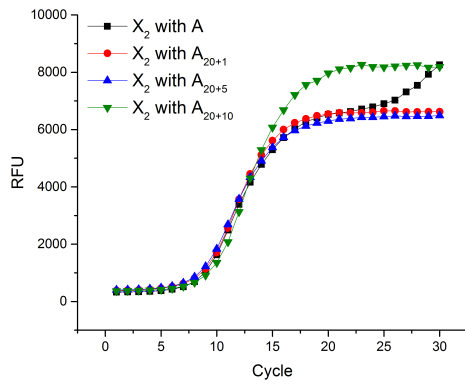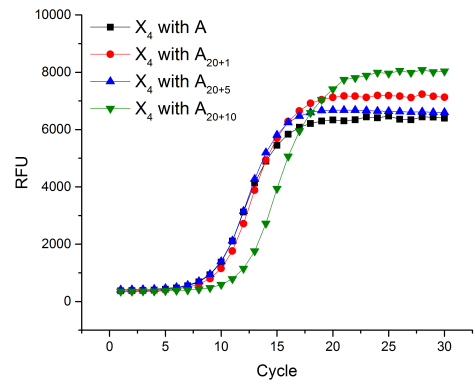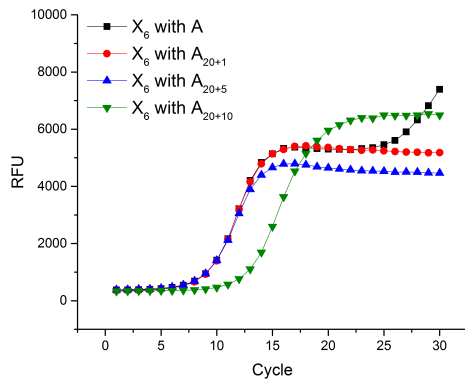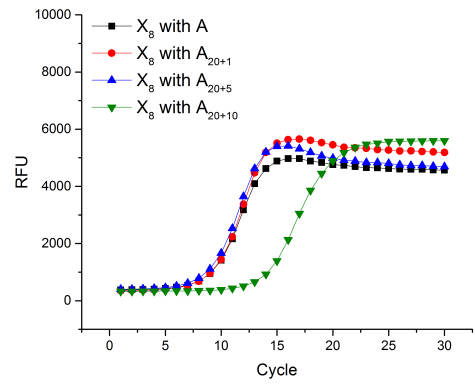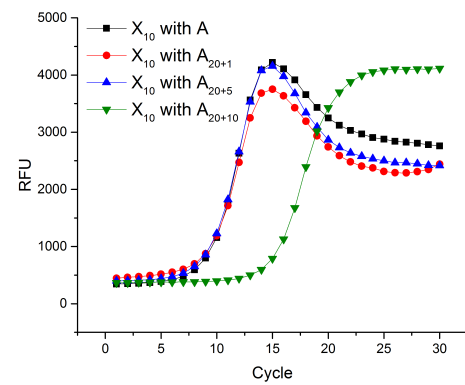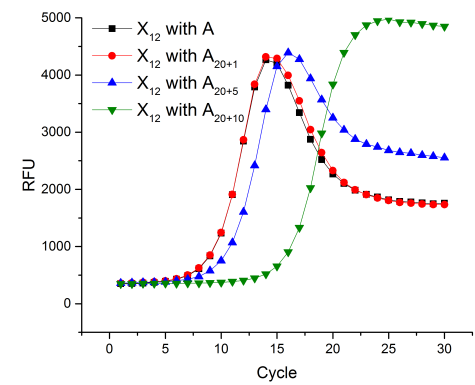
FIGURE 3.20: Screening of population growth manipulation using A, A_{20+1}, A_{20+5} or A_{20+10} as a primer with the libraries X_2 (a), X_4 (b), X_6 (c), X_8 (d), X_{10} (e), and X_{12} (f).
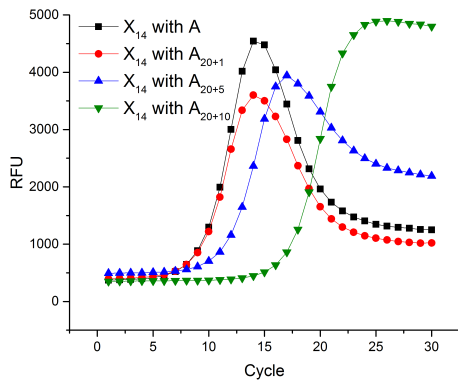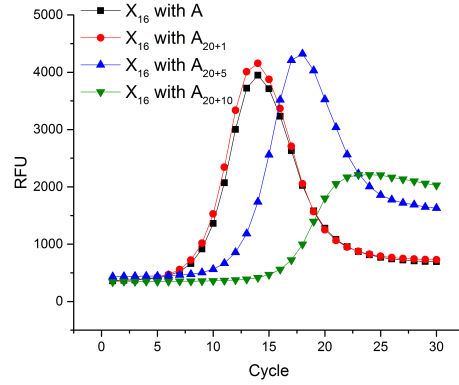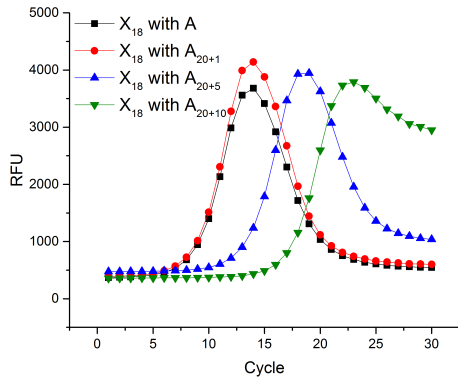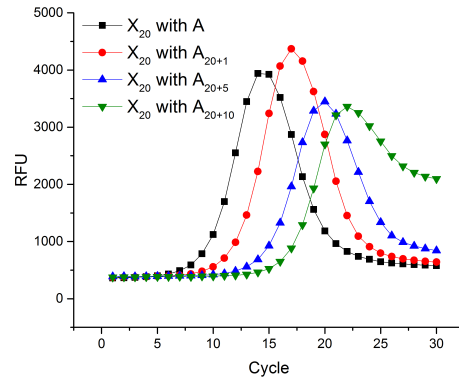
FIGURE 3.21: Screening of population growth manipulation using A, A_{20+1}, A_{20+5} or A_{20+10} as a primer with the libraries X_{14} (a), X_{16} (b), X_{18} (c), and X_{20} (d) (continuation).

When A was employed in X_{20} amplification, every template was amplified, so diversity was 4^{20}. A_{20+1}, A_{20+5}, and A_{20+10} selectively annealed to and amplified $1/4$, $1/4^5$, and $1/4^{10}$ of X_{20}, respectively. This resulted in products of decreased diversity, see Figure 3.21d. A_{20+1} shifted the amplification curve to the right (4^{19} templates) and A_{20+5} caused a greater shift because it only amplified templates which had `ATTCG` after the priming region (4^{15} templates). The diversity of both libraries was still high, so the end fluorescence signal remained similar. When A_{20+10} was used

74

with X_{20}, on top of the farther right shift, the curve shape became similar to the one obtained for X_{10}. The diversity of 4^{20} went down to 4^{10} (Figure 3.22b). In X_{10}, the allele-specific primers from A to A_{20+5} caused no difference because the same library members were amplified in every condition (Figure 3.20e). For X_{18}, A_{20+1} did not cause a shift because the primer matched all sequences (Figure 3.21c). When $X_{10/2}$ was amplified with A_{20+1} and A_{20+5}, the diversities were reduced from 4^{10} to 4^9 and 4^5, respectively, transforming its amplification curve from a bell shape to a sigmoid (Figure 3.22a). The biased amplification could be detected by a right-shift (influenced by concentration) and curve-shape (influenced by diversity).
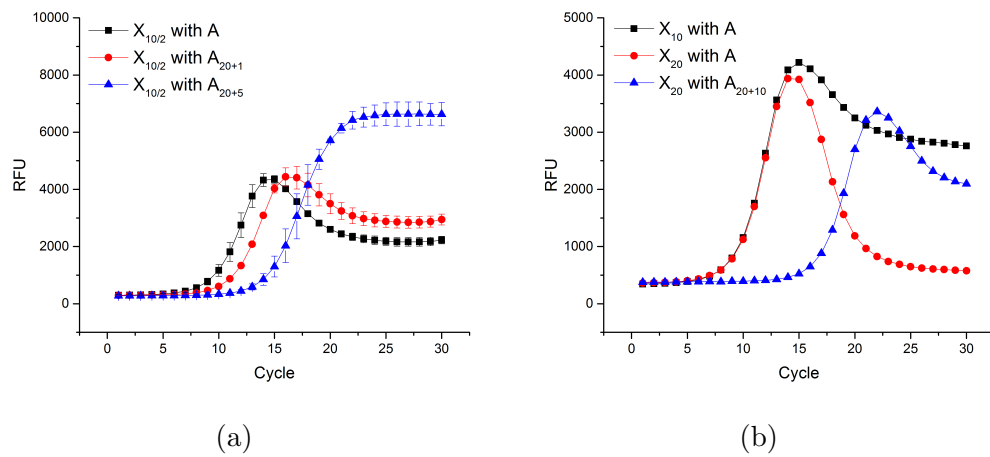
FIGURE 3.22: Comparison of biased population growth using A, A_{20+1}, A_{20+5} or A_{20+10} as a primer with the libraries $X_{10/2}$ (a) and X_{20} (b).

3.1.3 Protection by Population Diversity

Following the AS-PCR experiments, we wanted to understand how far the libraries diversity could be manipulated. For that end, we planned an experiment that would use the diversity as a selection/protection factor and that could be handled specifically. The libraries X_{EcoRV} and X_{SmaI} (Table 2.2) carried the restriction sites for EcoRV and SmaI, respectively. Both are endonucleases that create blunt ends

(Figure 2.1), and only recognise duplex restriction sites. This experiment relied on the fully-matched early duplexes as perfect restriction sites (early cycles); thus, there will be digestion. After, the progressively mismatched duplexes (later cycles) generate lower to no digestion (Figure 3.23). The restriction sites function as a selection mechanism.
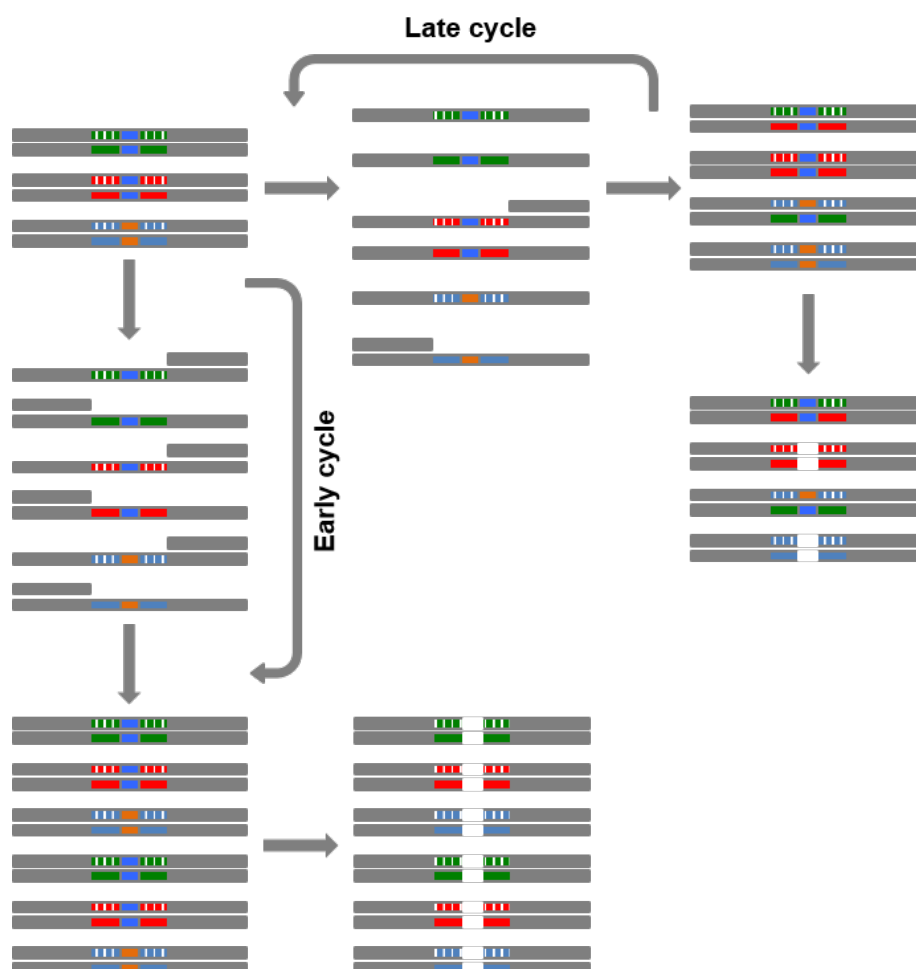
FIGURE 3.23: Diversity of libraries with restriction sites has a protective effect against endonuclease action. The fully complimentary duplex synthesised in the initial PCR cycles is more sensitive to endonucleases than the mismatched duplex generated in the late PCR cycles.

Restriction Enzymes Activity Check   To verify that the libraries carrying the restriction sites were recognised and digested by the proper enzymes, we incubated 15

cycles PCR amplified product of libraries (X_{SmaI}, X_{EcoRV}, X_{20}) with SmaI or EcoRV (Figure 3.24). X_{SmaI} and X_{EcoRV} were successfully digested by SmaI and EcoRV, respectively, originating a product of around 30 bp, and there was significantly less 60 bp original template. X_{20} was used as a control of no digestion.
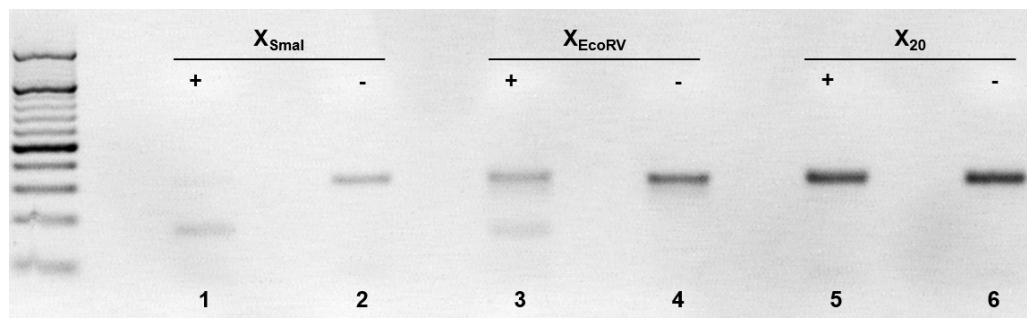
FIGURE 3.24: X_{SmaI} incubated with SmaI (1) and no enzyme control (2), X_{EcoRV} incubated with EcoRV (3) and no enzyme control (4), and X_{20} incubated with SmaI (5) and no enzyme control (6). 10 µL of PCR product treated with endonucleases were loaded into the gel. 20 bp ladder, 4% agarose gel with EtBr.

Amplification Curves The libraries were subjected to five or 25 cycles of amplification. Then, a diluted aliquot was treated with SmaI and/or EcoRV. After five cycles, most of all of newly synthesised DNA was in a perfect matched duplex and the fully complementary duplexes were recognised by the enzymes and digested, see Figure 3.25. There was a shift of around three cycles, in comparison to the no enzyme control, when one of the enzymes was used. When both enzymes were used, this shift was around seven cycles. When the population was amplified for 25 cycles, the mismatched population protected itself from digestion, because the restriction sites were not in a duplex and, thus, were not recognised. As a result, all conditions had the same amount of DNA template and the curves could not be distinguished. DNA reshuffling generated a heterogenous population that protected the templates from endonuclease digestion. The fully complimentary duplex synthesised in the initial PCR cycles was more sensitive to endonucleases than the mismatched duplex

generated in later PCR cycles.

(a)

(b)
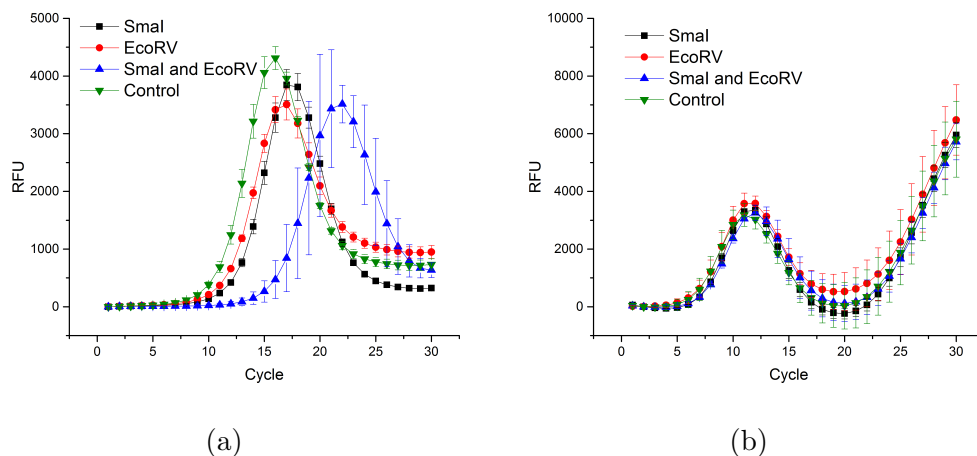
FIGURE 3.25: Amplification curves of the mixture of X_{SmaI} and X_{EcoRV} pre-treated for five (a) and 25 (b) PCR cycles and after incubation with SmaI and/or EcoRV.

Electrophoresis   To further demonstrate that diversity provides protection, we treated 20 and 35 cycles of amplification RT-PCR products of X_{SmaI}, X_{EcoRV} and S_1 with SmaI, and ran an agarose gel (Figure 3.26). SmaI was not able to digest any of the amplified products, meaning that 20 cycles was already too late for perfect matching. After 35 cycles of amplification, we saw a distribution of higher molecular weight products. This observation fits the fluorescence seen on Figure 3.25b, where there was an unexpected reoccurrence of fluorescence after 20 RT-PCR cycles on top of 25 PCR cycles of amplification. This phenomenon is discussed in the next section.
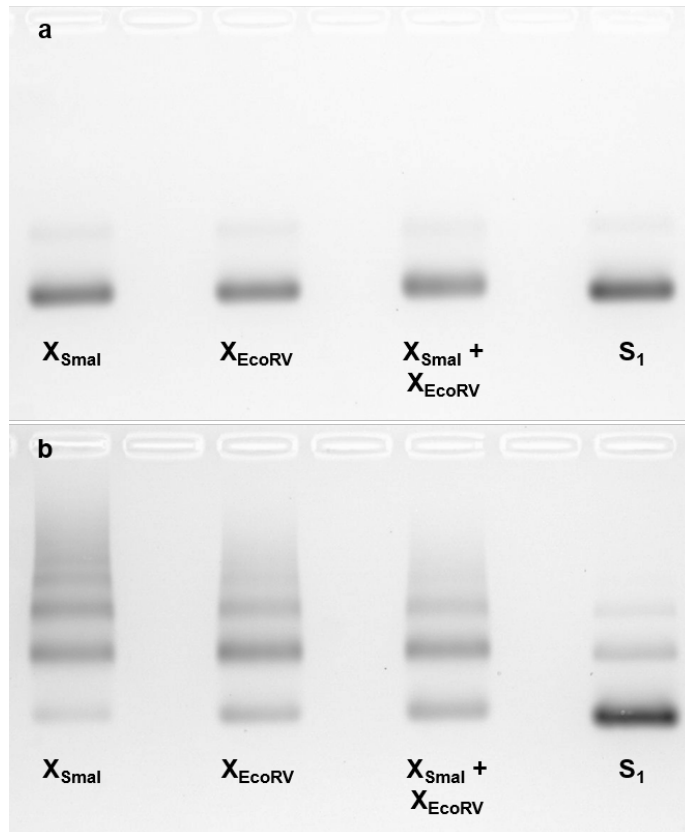
FIGURE 3.26: RT-PCR products treated with SmaI after 20 (a) and 35 cycles of amplification (b). 20 bp ladder, 4% agarose gel with 1X SYBR Green I.

3.1.4 High-Entropy Structures Formation

The later cycles of the amplification generated structures which had higher affinity to SYBR Green I, enhancing the fluorescence of the RT-PCR amplification curve to levels that were not correlated to the concentration of the DNA present in solution. We hypothesised that the large diversity of random templates gave space to unexpected interactions, which generated structures other than duplexes. For instance, X_{20} could create interactions that would go beyond the power of computation. So, a highly diverse DNA library had more probability of evolving into complex structures than a non-diverse library.

The first observation of this effect was after the 25 times PCR amplified sample, for

the endonuclease experiment, generated fluorescence values higher than the maximal fluorescent signal observed in standard RT-PCR experiments. Without the pre-PCR treatment, the fluorescence increase was only observed after 37 cycles. The initial fluorescence decrease we perceived was caused by lower dye affinity to less stable mismatched DNA duplexes, but the increase at later stages was probably caused by the enhanced dye affinity for highly convoluted and complex structures and networks. To understand how these structures are structured, we imaged them with agarose gel electrophoresis and AFM.

Temperature Effect

These ultra-structures were formed later when the handling occurred at $4\,°\mathrm{C}$ (37 cycles) and earlier if at $25\,°\mathrm{C}$ (32 cycles) (Figure 3.27). After, the samples obtained in these conditions were ran in an agarose gel, see Figure 3.28. It can be seen that the $4\,°\mathrm{C}$ condition generated several bands corresponding to higher molecular weights, heavier than 60 bp. At $25\,°\mathrm{C}$, there were no clear bands, only smears of undefined DNA products. This observation suggests that the structures formed at $25\,°\mathrm{C}$ were too complex to be separated on an agarose gel.

(a)   (b)
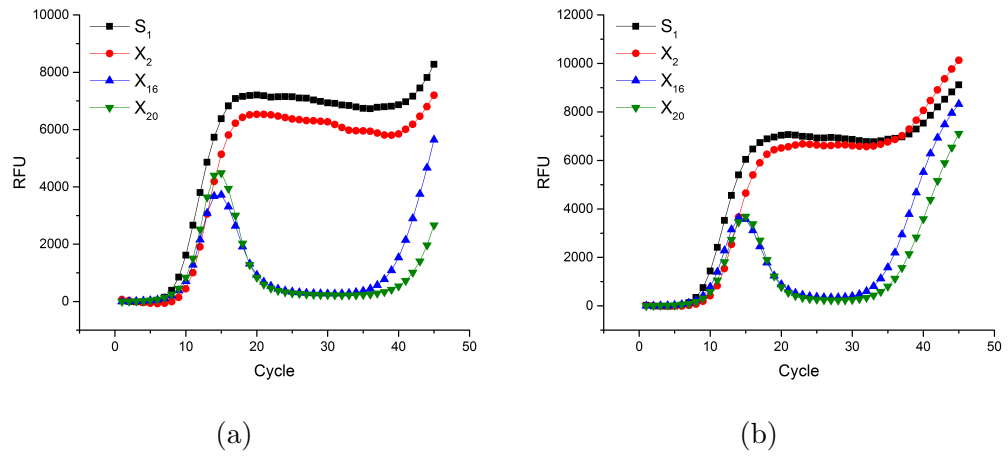
FIGURE 3.27: Amplification of S_1, X_2, X_6, X_{16}, and X_{20} during 45 cycles, with handling at $4\,^\circ$C (a) and at $30\,^\circ$C (b).
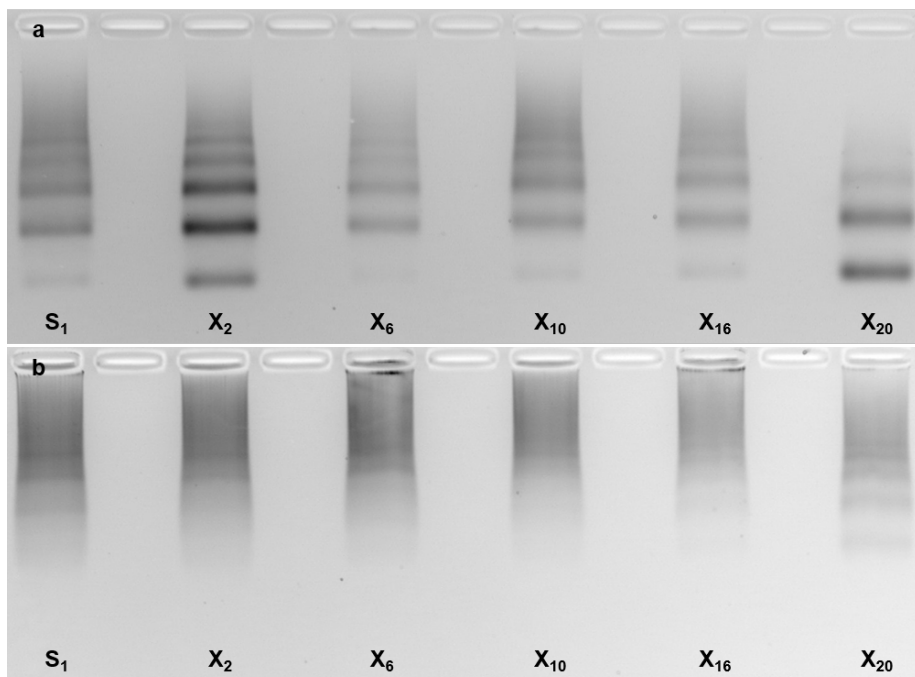
FIGURE 3.28: RT-PCR products amplified for 45 cycles at $4\,^\circ$C (a) and $25\,^\circ$C (b). The correct 60 bp product corresponds to the lowest molecular weight band. 4% agarose gel with 1X SYBR Green I.

Amplified Products Electrophoresis

To determine in more detail the structures being formed at room temperature, we further tested our samples in different conditions using electrophoresis.

Agarose Gels Samples collected after 20, 35, and 45 cycles of amplification for S_1, X_{16}, and X_{20} were ran in an agarose gel (Figure 3.29). After 20 cycles, S_1 had a strong band of the correct molecular weight DNA, 60 bp. The same was seen for the other libraries, with only a faint extra band at 80 bp. After 35 cycles, S_1 had also a faint extra band at 80 bp, while X_{16} and X_{20} showed more bands above 80 bp. After 45 cycles, all samples showed several extra bands above 60 bp.
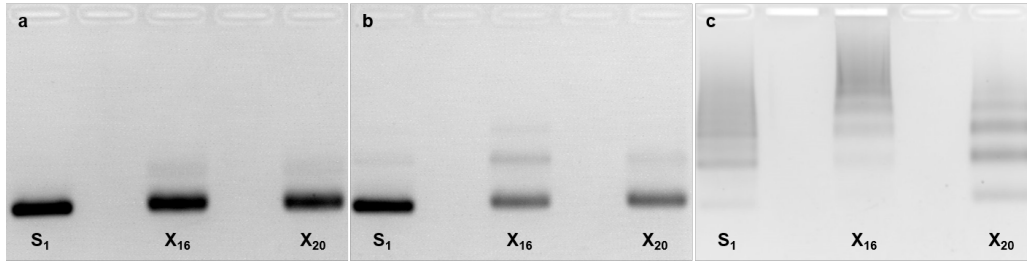
FIGURE 3.29: DNA samples of different diversity (S_1, X_{16}, and X_{20}) were subjected to 20 (a), 35 (b), or 45 (c) cycles. The correct amplified product corresponds to the band with the lowest molecular weight, 60 bp. 4% agarose gel with 1X SYBR Green I.

TBE-Urea Gels For a more detailed analysis, we prepared differently amplified samples and analysed them by Polyacrylamide Gel Electrophoresis (PAGE) under denaturing conditions. On Figure 3.30, after 20 cycles, there were still a lot of primers not consumed (20 bp) and the correct size final product (60 bp) had only a faint band. While ssDNA showed only one band, the dsDNA PCR product showed two bands, indicating the duplex was not fully denatured. Bands corresponding to DNA larger than 60 bp were observed in X_{16} and X_{20}, but not in the sample with only one sequence (S_1) with X_{16} presenting the largest shift. After 45 cycles, the 20 bp

band diminished and the 60 bp increased intensity. High molecular weight bands appeared in all samples, though S_1 still produced a strong 60 bp band. Then, higher molecular weight bands were developed up to a complete smear for the DNA sample with higher diversity, X_{20}. A similar gel was run with PCR samples that had been denatured for the double amount of time and at a higher temperature, see Figure 3.31. The results were similar, but the resolution was improved, because the dsDNA was denatured more completely.
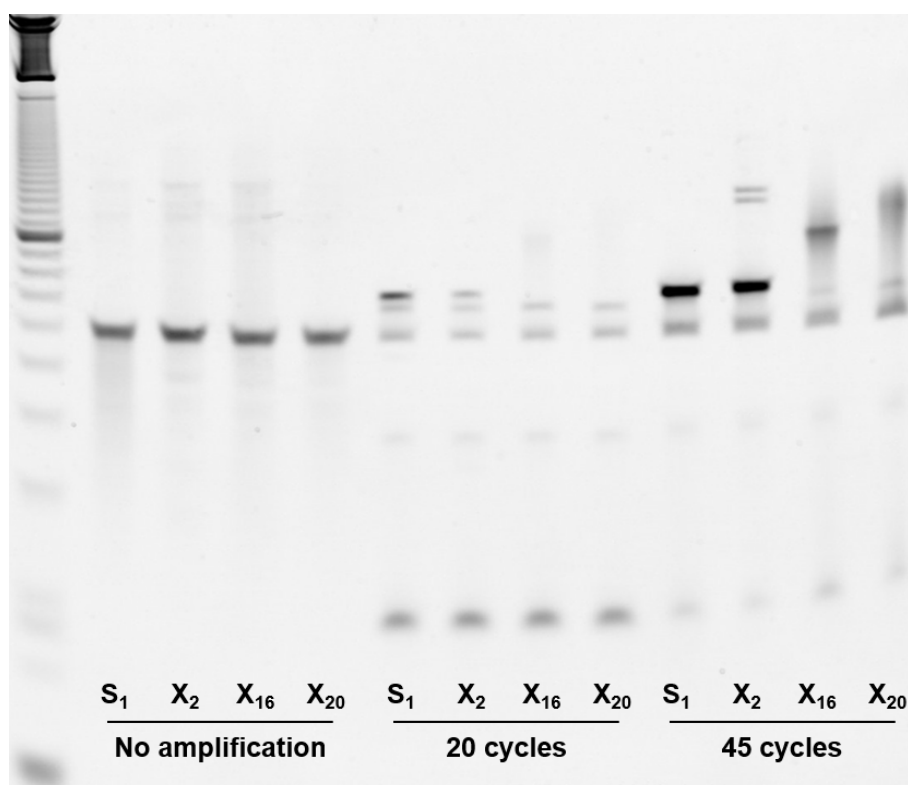
FIGURE 3.30: DNA samples of different diversity (S_1, X_2, X_{16}, and X_{20}) were subjected to 20 or 45 PCR cycles, and separated under denaturing conditions. 5 µL of PCR product were loaded into the gel. 10 bp ladder, 15% TBE-urea gel with 1X SYBR Green II.
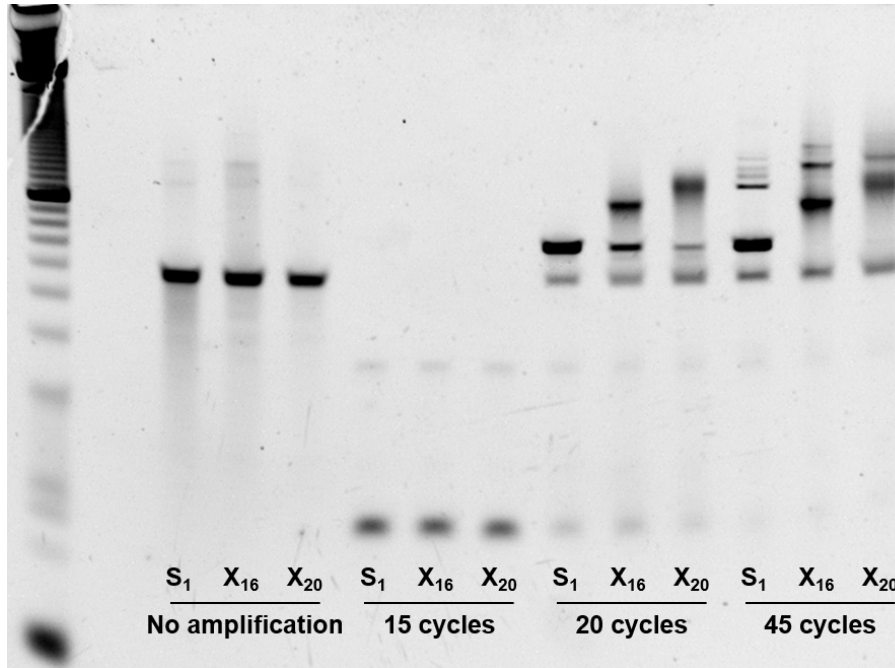
FIGURE 3.31: DNA samples of different diversity (S_1, X_{16}, and X_{20}) were subjected to none, 15, 20, or 45 PCR cycles, denatured under extreme conditions, and separated on a denaturing gel. 5 µL of PCR product were loaded into the gel. 10 bp ladder, 15% TBE-urea gel with 1X SYBR Green II.

On the next gel electrophoresis experiment, we monitored the dsDNA (SYBR Green I) and ssDNA (SYBR Green II) signals, see Figure 3.32. The dsDNA could not be seen after 15 cycles of amplification because it had been denatured by urea completely. However, it could be seen with SYBR Green II staining. After 20 cycles of amplification, the product became more stable and more difficult to break. For higher diversity libraries, there was an increased amount of bands corresponding to higher molecular weights. The ssDNA staining using SYBR Green II was also interesting, because it also became heavier according to the diversity of the original sample. After 45 cycles of amplification, the correct product became less visible and bands corresponding to higher molecular weight DNA were predominant.

It is important to note that the decrease of overall signal intensity did not reflect the loss of sample. It shows the increase of length heterogeneity, which could no longer

be resolved by gel electrophoresis.
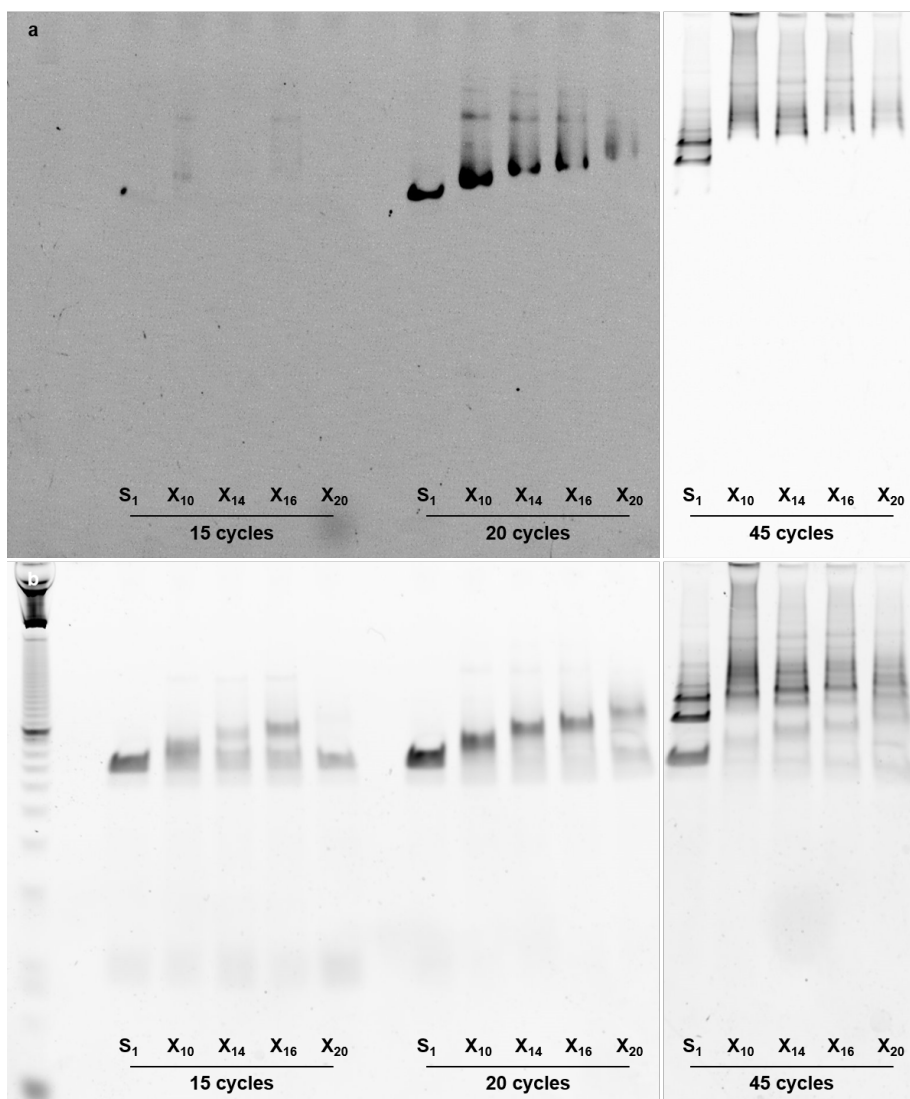
FIGURE 3.32: DNA samples of different diversity (S_1, X_{10}, X_{14}, X_{16}, and X_{20}) were subjected to 15, 20, or 45 PCR cycles, and separated on a denaturing gel. 1X SYBR Green I shows the dsDNA (a) and 1X SYBR Green II shows the ssDNA (b) signals. 5 µL of RT-PCR product were loaded into the gel. 10 bp ladder, 15% TBE-urea gel.

Atomic Force Microscopy Imaging

The RT-PCR products were also imaged by topographic analyses on the AFM, see Figure 3.33. The correct 60 bp product corresponds to around 20 nm structures, showing as a dot in the provided conditions. The height of ssDNA is around 0.4 to

0.6 nm and dsDNA's from 0.8 to 1.2 nm.

After 20 cycles, S_1's height was consistent with dsDNA and the length was around 20 nm, with some longer structures. Small structures of dsDNA were mainly observed for both S_1 and X_{20}. After 45 cycles, the dots became larger and thicker. Mostly large structures were detected for X_{12}, X_{14}, X_{16}, and X_{20}, but not S_1. Structures higher than standard DNA duplex were seen, especially for X_{16}. After 20 cycles, X_{20} generated DNA structures that were longer and higher than the ones of S_1. After 45 cycles, superstructures much higher and longer than dsDNA were observed. The prolonged PCR amplification of diverse DNA libraries led to complex 3D structures, as can be confirmed on Figure 3.34. DNA origami was used as a control for assay development. Freezing and thawing samples induced structural changes, mainly clumping (Figure 3.35).
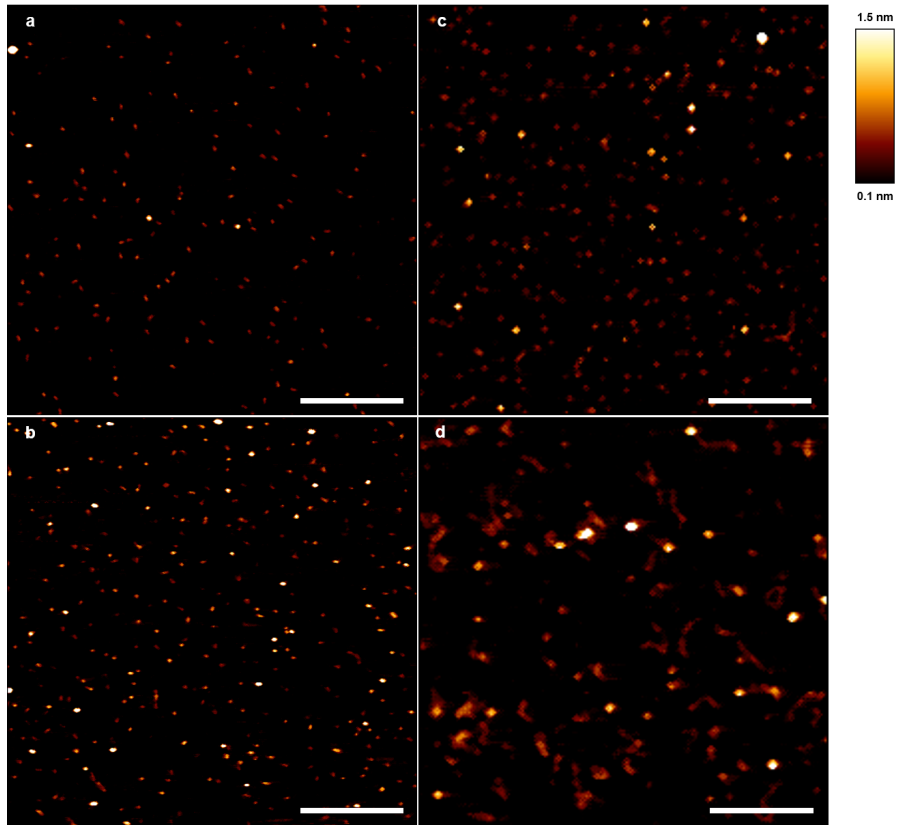
FIGURE 3.33: AFM analyses of RT-PCR products of S_1 after 20 (a) and 45 cycles (b), X_{20} after 20 cycles (c), and 45 cycles of amplification (d). Images are 1 μm × 1 μm and the scale bar is 250 nm.
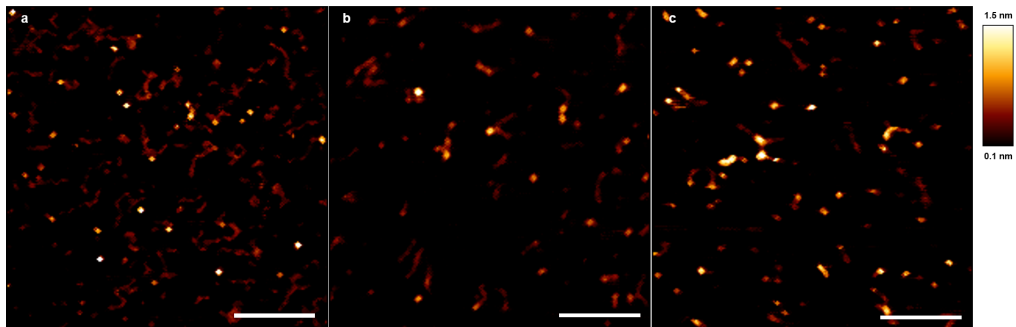
FIGURE 3.34: AFM analyses of RT-PCR products of X_{12} (a), X_{14} (b), and X_{16} (c) after 45 cycles of amplification. Images are 1 μm × 1 μm and the scale bar is 250 nm.
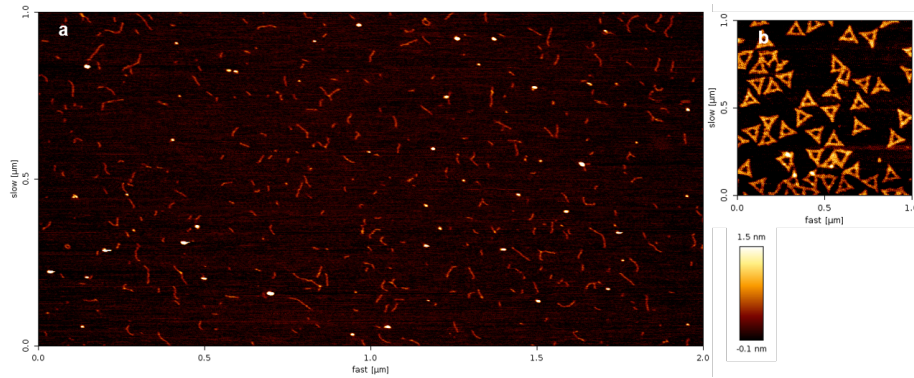
FIGURE 3.35: Frozen and thawed treated RT-PCR sample of S_1 after 45 cycles of amplification (a). Stiff in air tapping of DNA origami controls (b).

3.1.5 Library Diversity Evaluation After Selection

To emulate a real selection, we used ten highly diverse sequences (from S_1 to S_{10}) and the B' primer. This construct had a melting temperature around $54\,°\mathrm{C\,s^{-1}}$. B' was modified with biotin (a high affinity binder to SA), iminobiotin (a weak binder), or was not modified at all (Figures 3.36a and 3.36b). Using these constructs, we created four different libraries: *LibA*, which consisted of S_1 annealed with a biotin-primer and S_2 to S_{10} annealed to an unmodified primer (with some strong binders); *LibB*, which was the same as *LibA*, but with an iminobiotin-primer (with some weak binders); *LibC*, where S_1 to S_{10} were annealed to the iminobiotin-primer (with many weak binders); and *LibD*, where S_1 was annealed to biotin-primer and S_2 to S_{10} were annealed to the iminobiotin-primer (with some strong binders, but mostly weak). These libraries were incubated with SA beads. The beads were washed, once or twice to mimic different selection stringencies, and the solution pH was increased and then decreased so the DNA templates would be in solution to further continue with RT-PCR monitoring.

Figure 3.36c is the result of the selection experiment after washing the SA beads once and Figure 3.36d twice. In both graphs, *LibA* has a sigmoid amplification curve compatible with the existence of one template, or a very low diversity sample.

In *LibA*, only S_1 is annealed to the biotin-primer, so only this template was able to bind to SA. The extra wash did not alter the outcome for this library. *LibB* was the least enriched condition and the difference between one or two washes was clear. After one wash, the S_1-iminobiotin construct was still present in the final solution, but after two washes, this construct was in lower amount. Thus, the shifted curve to the right. Removing unspecific binding of unmodified DNA increased the end-point value. Since iminobiotin has lower K_d than biotin, this result fit our expectation and the increase of 5.6 amplification cycles was caused by this. With *LibC*, there was a stark decrease in the fluorescence plateau, reflecting the larger diversity of the pool monitored, because all ten sequences were present.

After a second wash, a mixture of different constructs remained. For *LibD*, even though the curve shape was similar to *LibA*, there was a lower fluorescence value at the plateau for after one wash. More DNA in *LibD* bound to the SA beads than the other libraries, because all constructs were modified with either strong or weak binders. After two washes, iminobiotin constructs were washed out, producing lower diversity. Interestingly, after one wash, its end-point was lower than *LibA*, indicating that the selection using *LibD* generated a mixture of compounds, even though the strong affinity compound was the major participant. After the second wash, only one construct was dominant. The stringency simulation demonstrated that with more stringent selection conditions, as in two washing steps, *LibB* and *LibD* curves indicated only one major binder, whereas *LibC* curve, which was composed of many compounds of similar affinity, was not altered.
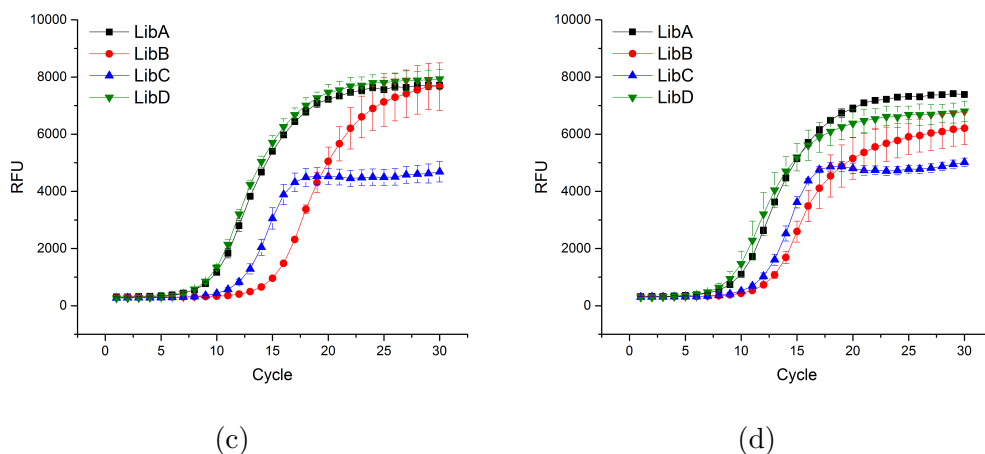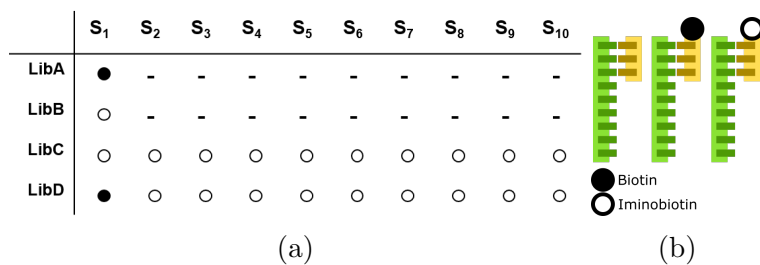
FIGURE 3.36: Libraries schematic (a): ten sequences (see Table 2.1) were annealed with a 20-mer sequence (B') (b), which was modified with biotin (full dot), imino-biotin (empty dot), or unmodified (dash). The selected mixture on SA beads after one (c) and two washing steps (d) was eluted and subjected to RT-PCR for diversity evaluation.

3.1.6 Controls

To ensure that the DDMA was robust and reproducible, we tested several control conditions.

Other Machine Another RT-PCR instrument was used to verify that our findings were not dependent on our system. The instrument used was a Mx3000P system and was ROX-dependent. ROX is used as a reference in ROX-dependent RT-PCR instruments. It does not interfere with the amplification, is not correlated with DNA concentration, is constant among wells, and normalises fluorescence variations

in optical paths. The constant ROX value integrated with the fluorescent reporter used normalises wells' fluorescence [47]. We tested our system's reproducibility after adding ROX and there was no interference, see Figure 3.37a. Then, we proceeded to amplification on the Mx3000P system and the results were similar to the ones obtained with the PikoReal system. The only difference was the range of fluorescence values, which were around five times higher with the new system (Figure 3.37b).

Other Master Mix   To verify that the effect we described was also not dependent on the commercial master mix used, we prepared a homemade version. Once again, the amplification pattern was similar, and the only difference was on the fluorescence signal, which was higher for the commercial master mix (Figure 3.37c).

Other Library   To finally address the question if this phenomenon was sequence dependent, another DNA diverse library, Y_{20}, was tested (Table 2.2). This library also had 20 randomised positions and the forward (C) and reverse (D') primers were 23 nucleotides-long, so the final DNA amplified product was of 66 bp. The amplification curve for Y_{20} was also shaped as a bell curve, but the maximum fluorescence value was lower than the one of X_{20}, see Figure 3.37d.
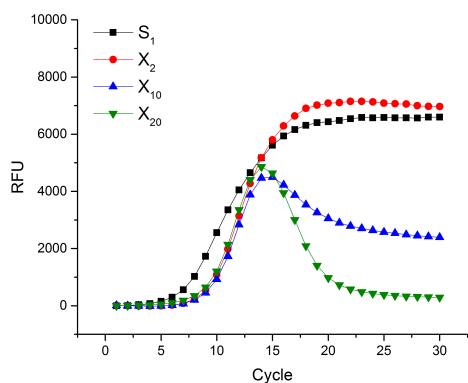
(a)
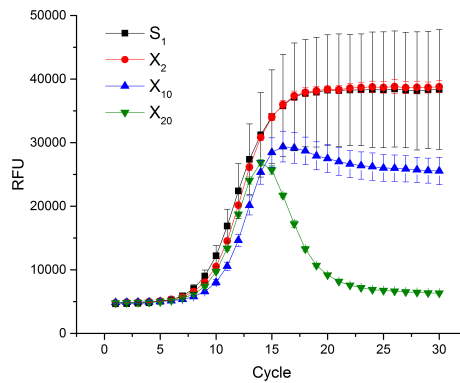
(b)
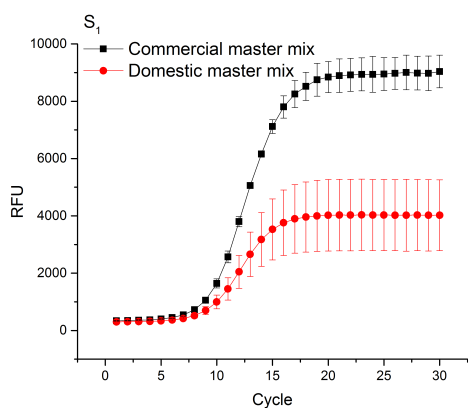
(c)
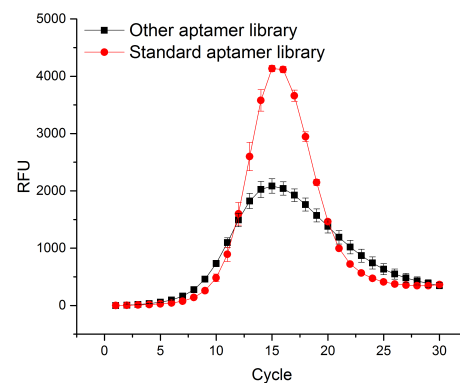
(d)

FIGURE 3.37: The addition of ROX did not alter the amplification and thereof fluorescence of the templates (a). Amplification curves obtained using another instrument (b). Comparison between commercial and homemade master mixes (c). Amplification curve obtained for another diverse library compared to the standard library used in this project (d).

3.2 Physical Environment Regulation of TNF-α Activity

3.2.1 Binding Assays

Band-Shift Assays

Unexpectedly, in the search of an improved DNA binding scaffold to TNF-α, all DNA structures tested (Tables 2.3 and 2.4, and Figure 2.4) were found to bind to TNF-α with micro-molar affinity (Figure 3.38). Various DNA constructs were incubated with a large excess of TNF-α for 1 h. Then, a band-shift assay was performed and DNA bound to it in high extent. An explanation for this effect might be related to the surface charges present on each entity present: oligonucleotides are negatively charged and TNF-α possesses a positively charged heparin-binding domain. Thus, the electrostatic attraction caused these non-covalent interactions.
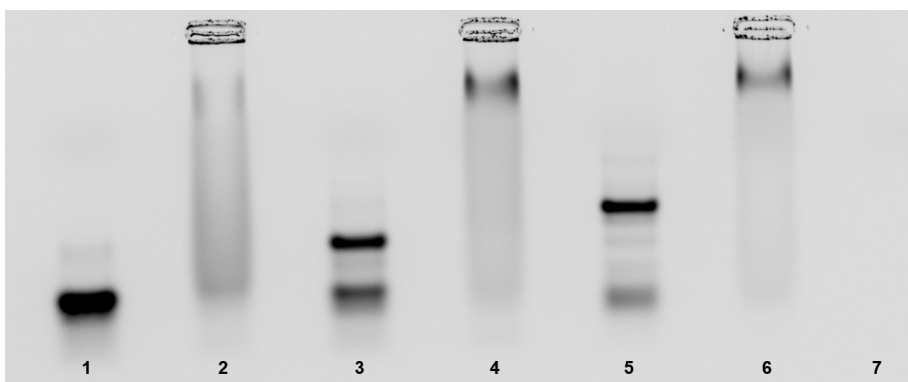
FIGURE 3.38: Band-shift assay resulting from the incubation of 50 µM TNF-α with 1 µM DNA: Cy5-ssDNA (1 and 2), Cy5-half-dsDNA (3 and 4), and Cy5-y-shape DNA (yDNA) (5 and 6). Lanes 1, 3 and 5 are the no protein control. Lane number 7 is TNF-α only control. 4% agarose gel.

Affinity Titration From several gel runs, we were able to estimate the K_d of TNF-α binding to Cy5-ssDNA, Cy5-half-dsDNA and Cy5-yDNA as below as 5 µM, 2 µM, and 1 µM, respectively (Figure 3.39). The titration of Xsmall was also performed (Figure 3.40). A Xsmall's structure K_d could be estimated to be around 1 µM.

It must be noticed that the nature of interaction between different DNA structures

and TNF-α changes. When using ssDNA or dsDNA the binding to TNF-α seemed to cause oligomerisation by cross-linking. This happened in such degree that a significant amount of the bound complex was not even able to move away from the wells. However, when using the yDNA or Xsmall, even though there was a strong binding to TNF-α, this effect did not appear because they could migrate from wells.

Bio-Layer Interferometry

In an attempt to determine more exactly the binding affinity of TNF-α to DNA and heparin, we developed a method on a BLI platform. In this section, we explore the protocol development and optimisation.

Anti-Penta-His Sensors Test   The first approach was to use anti-penta-His sensors, which captured the His-tagged TNF-α. The binding to the sensors was clear and reproducible among different experiments, see Figure 3.41a. However, when binding experiments with DNA were performed, the signal dropped after each cycle of regeneration. Due to the signal drop, it became difficult to calculate a definite K_d. Only an estimate was possible to retrieve, which was around (44.7 ± 4.7) µM (Figure 3.42).
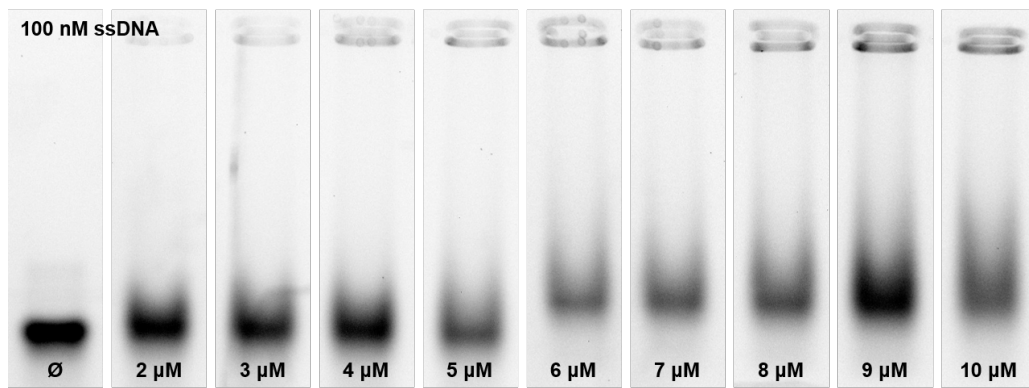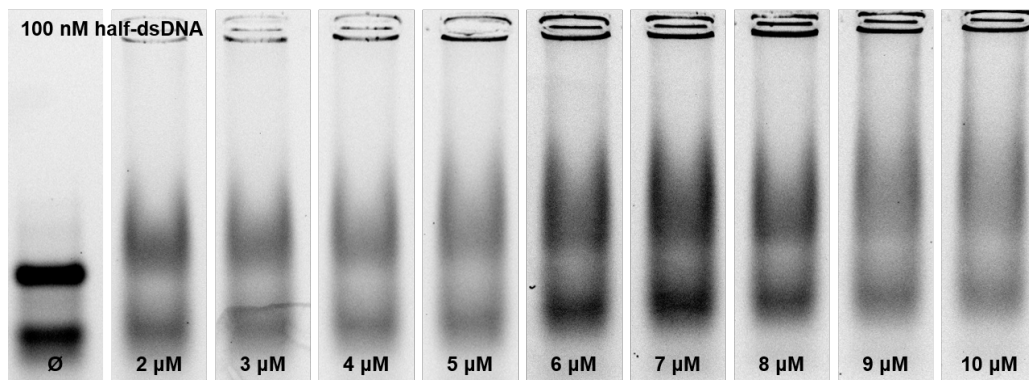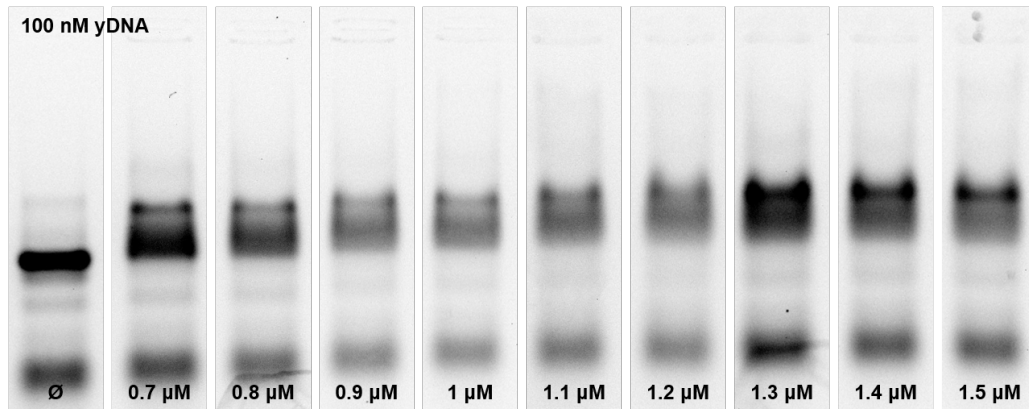
FIGURE 3.39: Band-shift assay resulting from the incubation of $100\,\text{nM}$ Cy5-ssDNA (a), Cy5-half-dsDNA (b), and Cy5-yDNA (c) with increasing TNF-α concentrations, as indicated on the image. 4% agarose gel.
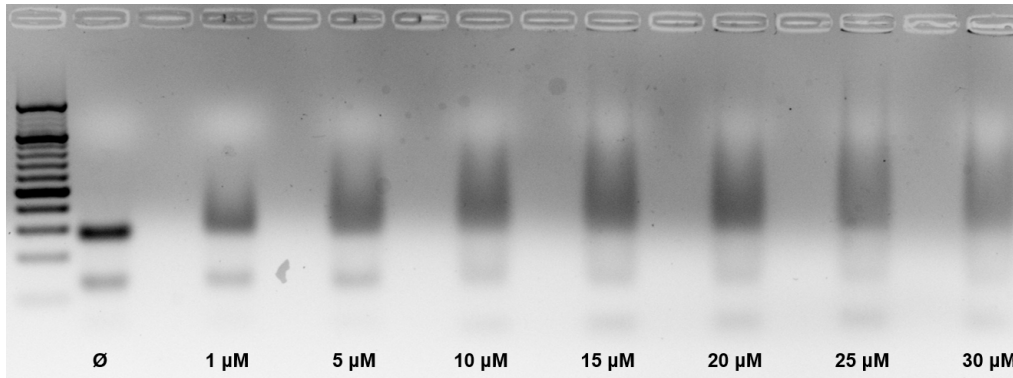
FIGURE 3.40: Incubation of 1 µM Xsmall with increasing TNF-α concentrations, as indicated on the image. 20 bp ladder, 4% agarose EtBr gel.

(a)

(b)
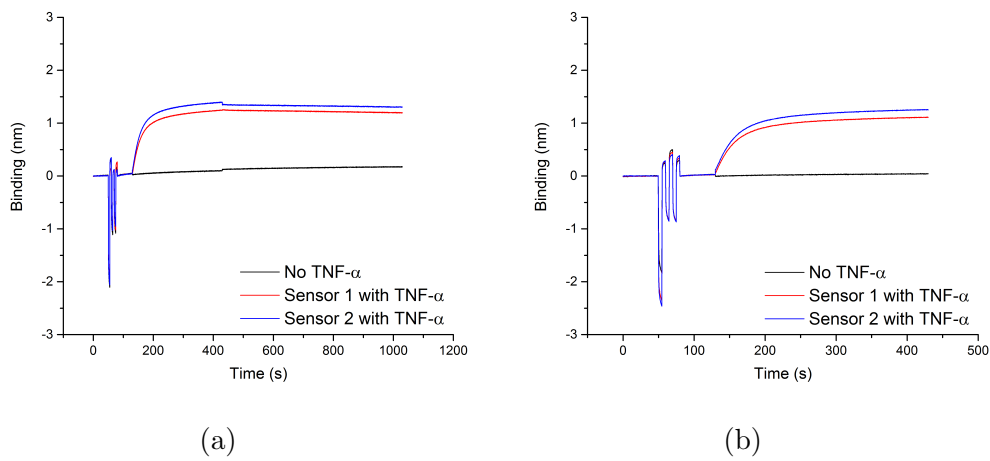
FIGURE 3.41: Regeneration and loading of three anti-penta-His sensors with 588 nM TNF-α after one (a) and two (b) consecutive loads.
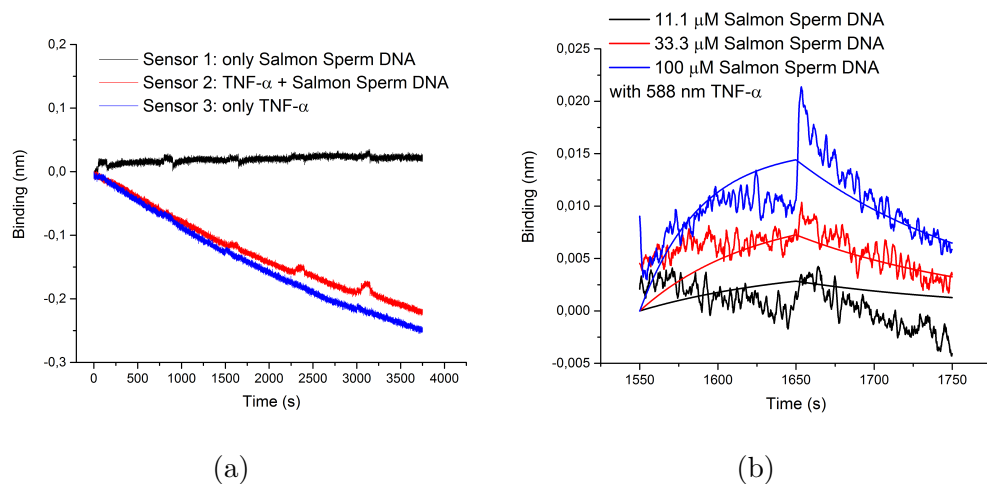
FIGURE 3.42: Binding cycles with three anti-penta-His sensors (a) and binding kinetics for the three highest concentrations of S-DNA (b). Sensor 1 was incubated only with increasing concentrations of S-DNA, from 1.23 to 100 μM. Sensor 2 was incubated with S-DNA, like sensor 1, plus 588 nM TNF-α. Sensor 3 was incubated only with 588 nM TNF-α.

Inhibition of TNF-α Binding to Amine Reactive Second Generation Sensors We then tested ARG sensors, which have a negatively charged surface (free carboxylic groups) to which TNF-α bound directly to. To explore in more detail the consequences of sensor charges, we also tested APS sensors. These sensors display amino groups on the surface, which become positively charged after protonation. APS and ARG sensors were loaded with TNF-α. The ARG sensor yield a very strong binding with a clear dissociation step and, as expected, the APS sensor did not show any binding (Figure 3.43). Thus, TNF-α interacted with ARG and was repulsed by APS.
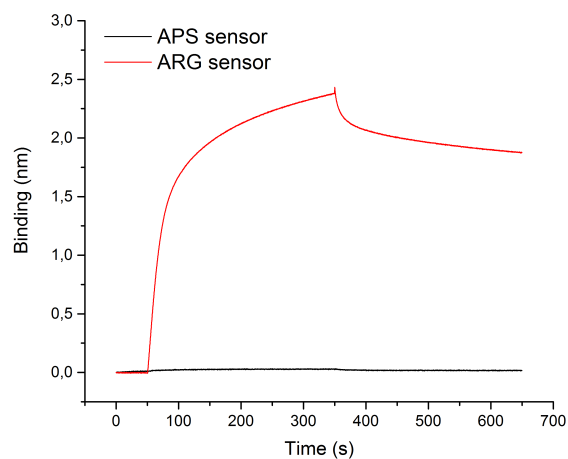
FIGURE 3.43: Loading of APS and ARG sensors with 1.44 µM TNF-α.

In a first trial of DNA and TNF-α binding measurement, we co-incubated 144 nM TNF-α with S-DNA from 0.1 to 100 µM. The obtained binding values were interesting, suggesting a decrease of TNF-α binding to the sensor correlated with the increase of DNA in the system, as seen on Figure 3.44a. However, the control sensor with no DNA had its signal decreased consistently throughout the cycles of loading/regeneration. To normalise the binding values, we calculated a ratio for each condition between the control and test sensors (Figure 3.44b). It was again observed that increasing concentrations of DNA yielded lower TNF-α binding to the unmodified sensors.
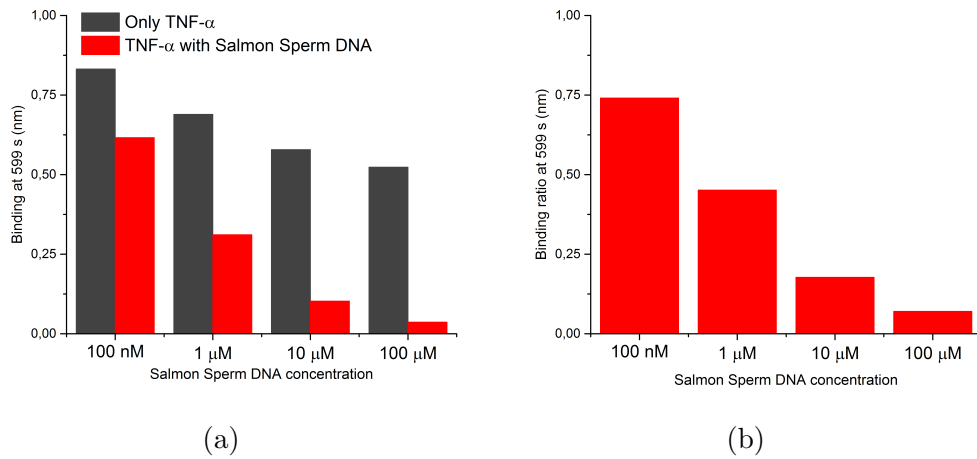
FIGURE 3.44: Co-incubation of 144 nM TNF-α with increasing concentrations of S-DNA. S-DNA at 100 nM corresponds to the first regeneration cycle, 1 μM to the second, 10 μM to the third, and 100 μM to the fourth. Total binding values (a) and binding ratios at 599 s (b).

The regeneration cycle was the next step to be improved and several conditions were tested. First, we tried a low pH buffer with a high concentration in salt plus glycine. The resulting binding and loss of signal were not optimal, see Figure 3.45, as the control sensor lost sensitivity very fast and at the last cycle had a third of its initial value. Then, we tested a regeneration buffer with high pH, 150 mM NaOH. With NaOH, the control sensor did not lose signal and the test sensor generated reliable data that was in accordance to the data obtained before (Figure 3.46a). When the binding ratio was plotted, the binding/no binding trend was obvious and, once again, the presence of DNA reduced TNF-α's binding to the ARG sensor (Figure 3.46b).
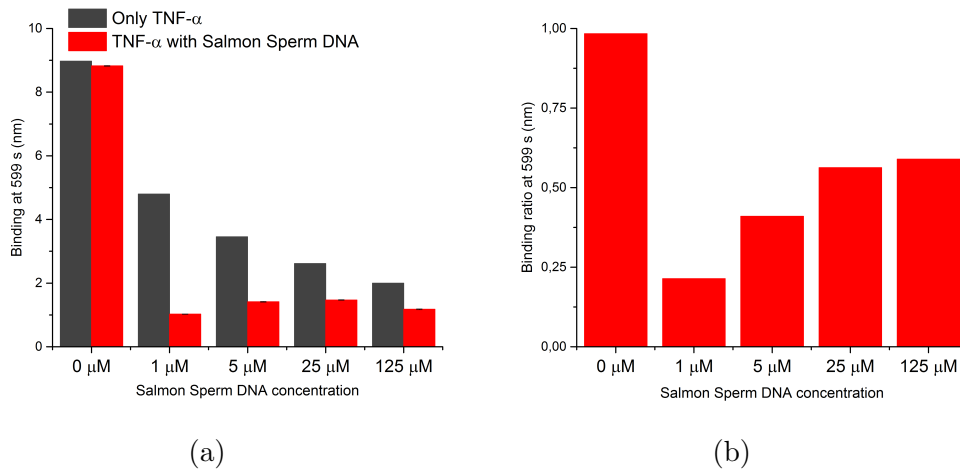
FIGURE 3.45: Co-incubation of $5\,\mu\mathrm{M}$ TNF-α with increasing concentrations of S-DNA, using as regeneration buffer $10\,\mathrm{mM}$ glycine with $2\,\mathrm{M}$ NaCl, pH 1.5. No S-DNA corresponds to the first regeneration cycle, $1\,\mu\mathrm{M}$ to the second, $5\,\mu\mathrm{M}$ to the third, $25\,\mu\mathrm{M}$ to the fourth, and $125\,\mu\mathrm{M}$ to the fifth. Total binding values (a) and binding ratios at $599\,\mathrm{s}$ (b).
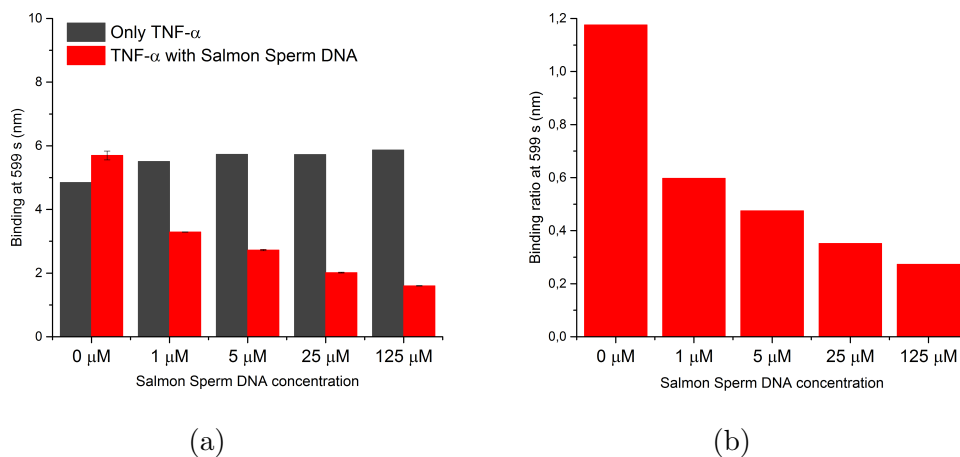
FIGURE 3.46: Co-incubation of $5\,\mu\mathrm{M}$ TNF-α with increasing concentrations of S-DNA, using as regeneration buffer $150\,\mathrm{mM}$ NaOH. No S-DNA corresponds to the first regeneration cycle, $1\,\mu\mathrm{M}$ to the second, $5\,\mu\mathrm{M}$ to the third, $25\,\mu\mathrm{M}$ to the fourth, and $125\,\mu\mathrm{M}$ to the fifth. Total binding values (a) and binding ratios at $599\,\mathrm{s}$ (b).

Binding Ratios Determination We put aside the goal of calculating a definite K_d and

aimed at comparing binding ratios between different constructs and materials. First,

we co-incubated 5 µM TNF-α with 5 µM of several DNA structures: ssDNA, half-dsDNA, yDNA and S-DNA. The binding of TNF-α to the sensor at 599 s was, at least, halved in all conditions, see Figure 3.47. The structure providing the lowest binding ratio was yDNA. When the experiment was repeated at a lower concentration of both DNA and TNF-α (1 µM) the decreased binding pattern for all structures was still visible, but it was not as strong (Figure 3.48). Then, we tested the same DNA structures coupled to the small molecule 2a (TNF-α inhibitor), using S-DNA as binding standard. TNF-α binding was reduced in a smaller extent when co-incubated with DNA-2a than it was when it was co-incubated with the same unmodified structures (Figure 3.49). Still, the binding with yDNA-2a was less than S-DNA.

(a)   (b)
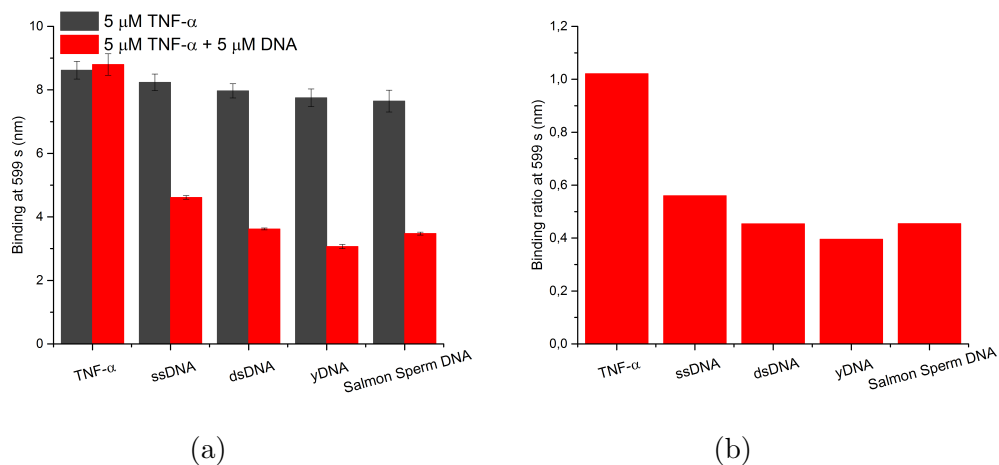
FIGURE 3.47: Co-incubation of 5 µM TNF-α with 5 µM of several DNA structures. Total binding values (a) and binding ratios at 599 s (b).

101
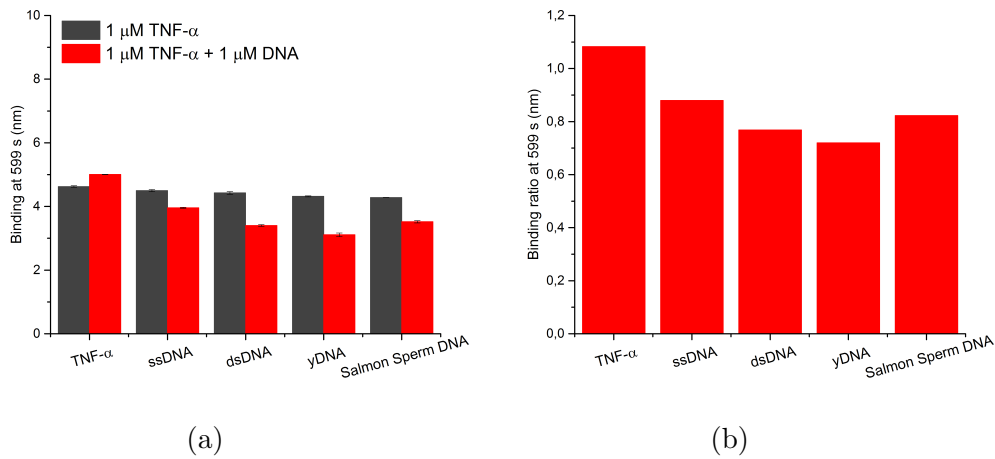
FIGURE 3.48: Co-incubation of 1 μm TNF-α with 1 μm of several DNA structures. Total binding values (a) and binding ratios at 599 s (b).
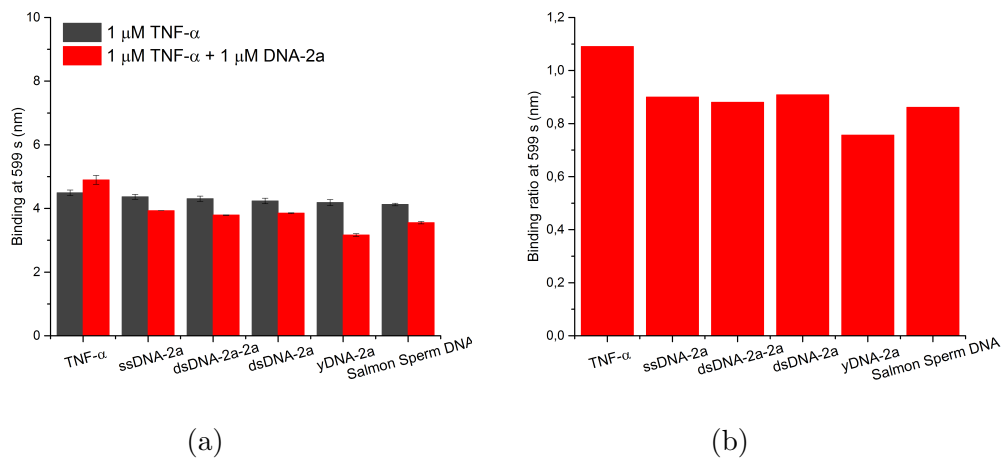
FIGURE 3.49: Co-incubation of 1 μm TNF-α with 1 μm of several DNA structures coupled to 2a and S-DNA. Total binding values (a) and binding ratios at 599 s (b).

Testing heparin and its derivatives, it was visible a very strong decrease of binding to the sensor, even stronger than with DNA: more than half less in the 5 μm TNF-α versus 5 μm heparin condition, see Figure 3.50. Even when the concentration was lowered to 1 μm, the binding reduction was more prominent than the one of DNA for

the same concentrations, see Figure 3.51. The unmodified heparin had the strongest effect on the binding of TNF-α to the ARG sensor.

(a)   (b)
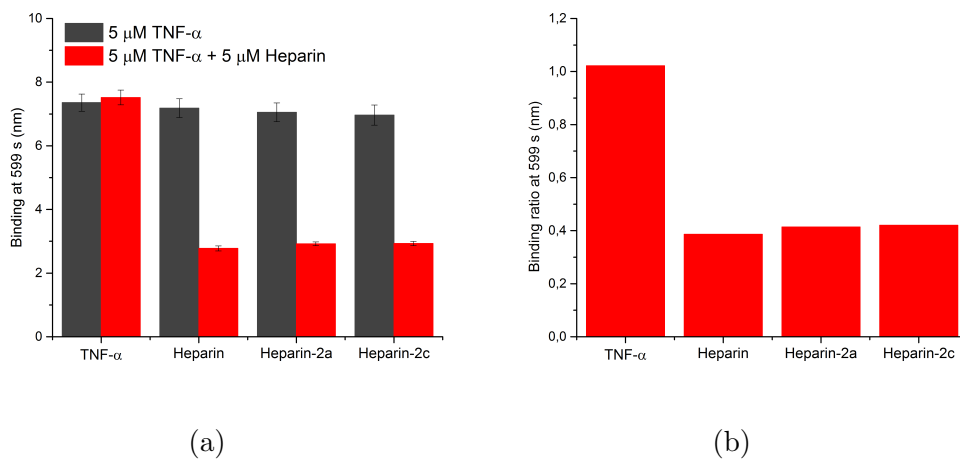
FIGURE 3.50: Co-incubation of 5 µM TNF-α with 5 µM of heparin and its derivatives. Total binding values (a) and binding ratios at 599 s (b).

(a)   (b)
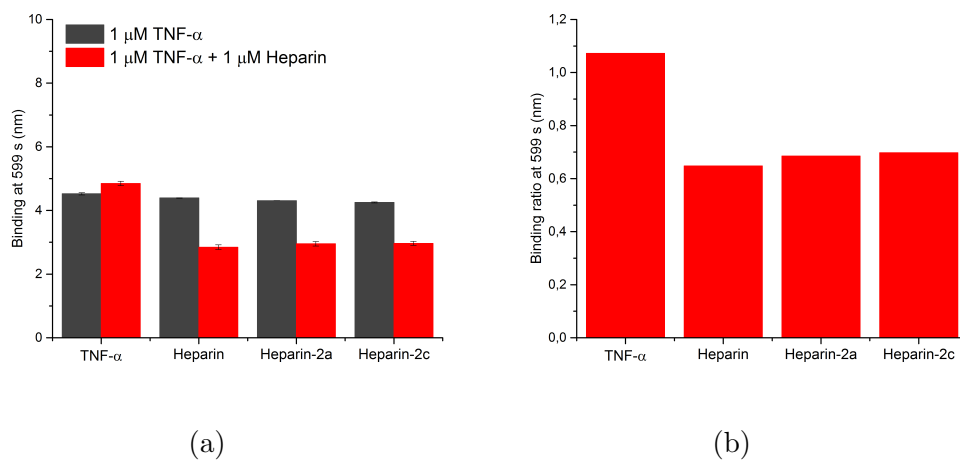
FIGURE 3.51: Co-incubation of 1 µM TNF-α with 1 µM of heparin and its derivatives. Total binding values (a) and binding ratios at 599 s (b).

Microscale Thermophoresis (MST)

The third method we applied to unveil the affinity of TNF-α to DNA was MST. The results were in accordance with the values we obtained using the band-shift assays. As it can be seen on Figure 3.52, the determined K_d of TNF-α to Cy5-full-dsDNA was $(10.7 \pm 2.0)\,\mu\text{M}$ and to Cy5-yDNA was $(4.9 \pm 1.1)\,\mu\text{M}$. At higher concentrations, TNF-α oligomerised.
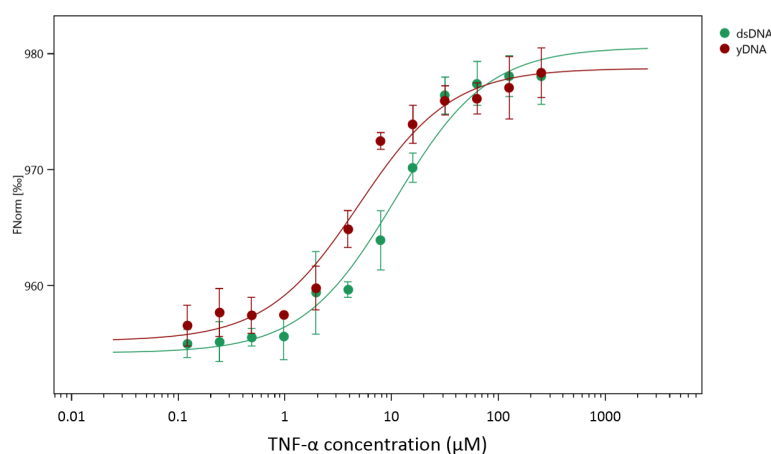
FIGURE 3.52: Dose-response-fit of DNA with TNF-α. The concentration of Cy5-full-dsDNA and Cy5-yDNA was $10\,\text{nM}$ and TNF-α concentration ranged from $122.5\,\text{nM}$ to $250\,\mu\text{M}$.

3.2.2   Cell Toxicity Assays

TNF-α's cytotoxic effect on L929 cells (*Mus musculus* fibroblasts) is the most explored in vitro model for this cytokine [210], causing either apoptosis or necrosis [211]. For TNF-α activation there must be co-incubation with ACT-D (transcription inhibitor) or CHX (translation inhibitor).

The cells were first sensitised to TNF-α with $2\,\mu\text{g}\,\text{mL}^{-1}$ ACT-D or $8\,\mu\text{M}$ CHX, see Figure 3.53a. After observing a similar toxic profile for both molecules, we proceeded to co-incubate them with $5\,\mu\text{M}$ DNA or heparin for $24\,\text{h}$ (Figure 3.53b). As reported by Stull et al. [210], ACT-D intercalated DNA, leading to optimal cell viability. As

for heparin, the effect was gone. Thus, CHX was chosen for the following experiments using DNA.

For the cell viability graphs presented in this section, the 100% cell viability corresponds to control wells, where cells were incubated only with their growth medium and 8 μM CHX. Both commercial and expressed in house TNF-α were tested in the presence or absence of DNA and generated equivalent results.

(a) (b)
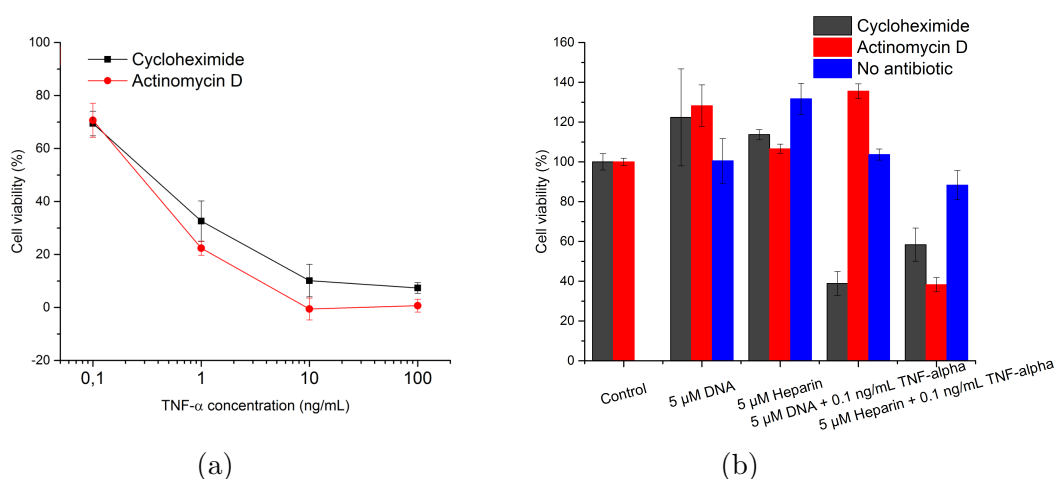
FIGURE 3.53: TNF-α co-incubated with 8 μM CHX or 2 μg mL^{-1} ACT-D (a) and DNA or heparin (b).

Oligonucleotides Co-Incubation

When short DNA sequences that were single, partially or totally annealed were co-incubated with 0.01 to 1 ng mL^{-1} TNF-α, there was an increase up to 20% of cytotoxicity (Figure 3.54). At 10 ng mL^{-1} TNF-α, cell viability was similar in every condition.
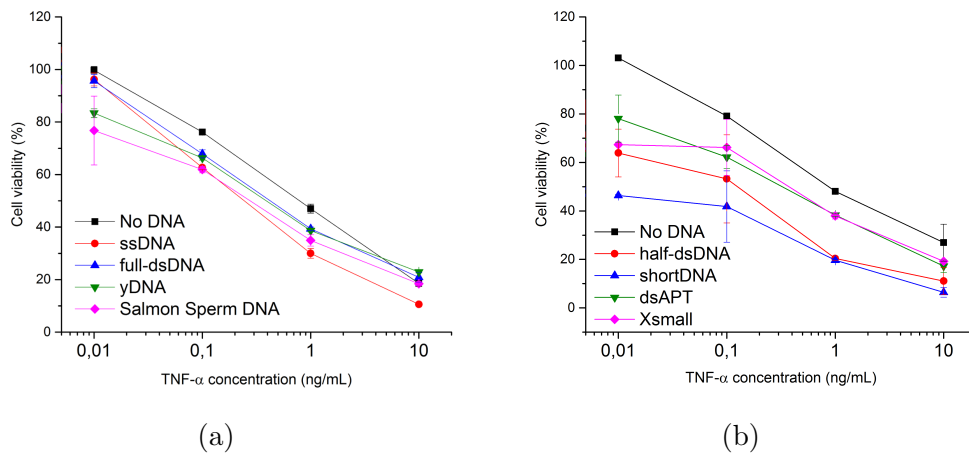
FIGURE 3.54: Each DNA construct was in a final concentration of $5\,\mu\text{M}$, in the presence of $8\,\mu\text{M}$ CHX. In (a), there is the effect of ssDNA, full-dsDNA, yDNA, and S-DNA. In (b), there is the effect of shortDNA, half-dsDNA, Xsmall and a double-stranded aptamer library (dsAPT).

To explore the time link between oligonucleotides and TNF-α enhanced cytotoxicity, some cells were pre-treated with DNA in the absence of TNF-α, see Figure 3.55. DNA pre-incubation for 2 h, followed by DNA removal was not able to induce significant increase of cell death when TNF-α was later added to the system. In parallel, DNA pre-incubation for 2 h, removal plus a new incubation for 24 h of DNA with TNF-α was able to generate enhanced cytotoxicity levels, similar to the previous observations.

106
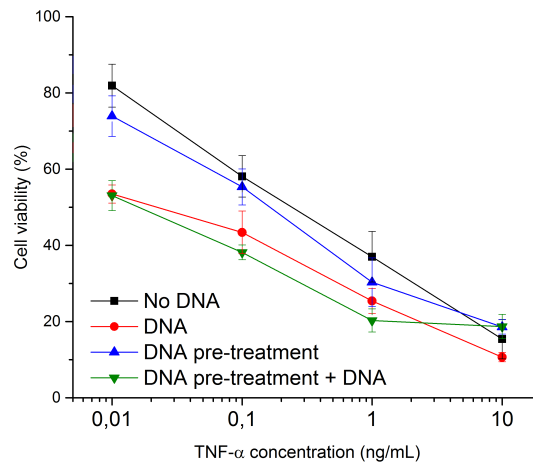
FIGURE 3.55: In the DNA condition, the incubation of TNF-α with DNA was carried for 24 h. The DNA pre-treatment condition consisted of incubating the cells, without TNF-α, with DNA for 2 h, removing it after. In the condition DNA pre-treatment plus DNA, DNA was added to the pre-treated cells with TNF-α for another 24 h. Each DNA construct was in a final concentration of 5 μM, in the presence of 8 μM CHX.

There are several reports describing enhanced cytotoxicity of DNA sequences carrying CpG motifs, which can act as immunostimulants [21]. To avoid this, we designed sequences that did not potentially present CpG motifs. As seen on Figure 3.56, the cytotoxicity was not too different from the initial sequences that could harbour CpG motifs. Meaning that either the effect was not present in any condition or, if present, was not influencing the system significantly.

107
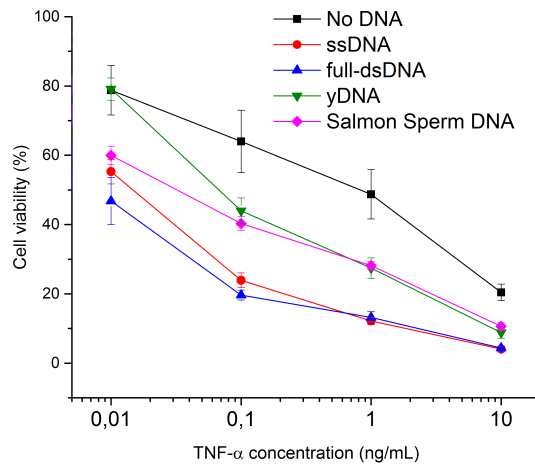
FIGURE 3.56: Cell viability assay co-incubating TNF-α with $5\,\mu$M of several DNA structures that did not harbour CpG motifs, in the presence of $8\,\mu$M CHX.

Co-Incubation with Known Cytotoxics   To explore whether DNA has a synergic effect with other cytotoxics, or whether the synergic effect is specific with TNF-α through physical interaction, we tested other cytotoxics in addition to CHX and ACT-D. The cells were incubated with $5\,\mu$M taxol and $5\,\mu$M irinotecan. It was clear that TNF-α bioactivity was independent from these two cytotoxic agents and dependent on the DNA constructs only (Figure 3.57). There was a strong cell viability decrease every time DNA was added to the cytotoxic plus TNF-α system, regardless of the initial cytotoxic effect caused by taxol or irinotecan.
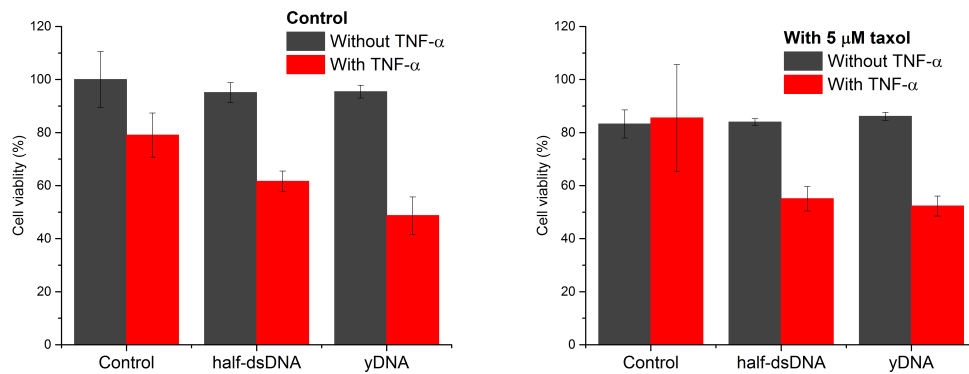
(a)

(b)
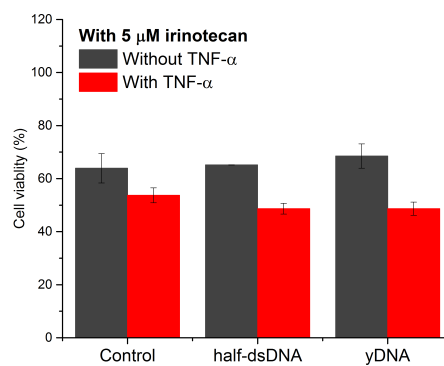
(c)

FIGURE 3.57: Co-incubation of $0.1\,\mathrm{ng\,mL^{-1}}$ TNF-α with $5\,\mu\mathrm{M}$ DNA and $5\,\mu\mathrm{M}$ taxol or irinotecan, in the presence of $8\,\mu\mathrm{M}$ CHX. The control experiment had no cytotoxic (a). In (b), taxol was added and in (c), irinotecan.

RNA Co-Incubation   We also investigated the influence of RNA in this system. Comparing with S-DNA, RNA generated a lower cytotoxic effect, see Figure 3.58.
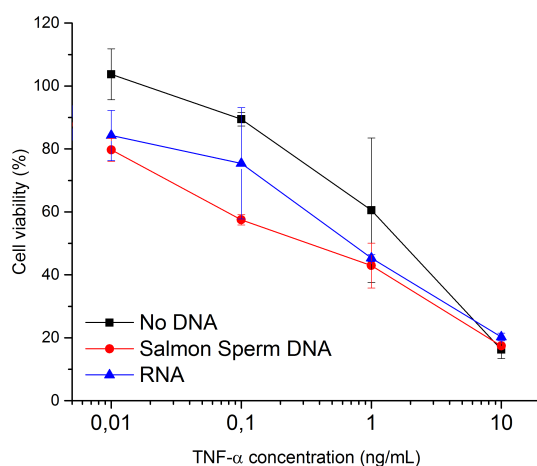
FIGURE 3.58: Cytotoxicity resulting from 5 µM RNA co-incubation with TNF-α, in the presence of 8 µM CHX.

Heparin and Derivatives Co-Incubation

Furthermore, we tested another negatively charged biopolymer effect in this system: heparin and heparin coupled to TNF-α inhibitors, 2a and 2c [112]. As tested previously on BLI, we knew that heparin and TNF-α also had a significant binding. All heparin conditions showed slightly enhanced cytotoxicity too (Figure 3.59a). Interestingly, heparin-2a and -2c had quite a strong impact on cell viability, even reversing the expected inhibition caused by the small molecules coupled to it.

As tested for the DNA constructs, we investigated the effect of the heparins on cell viability while co-incubating them with taxol and irinotecan (Figure 3.59b). 5 µM taxol or irinotecan with the heparins yielded a 20 to 30% lower cell viability, but when co-incubated with TNF-α, the value went lower, at least, 10%. This corroborates our hypothesis that DNA and TNF-α interaction was essential for cell death.

110
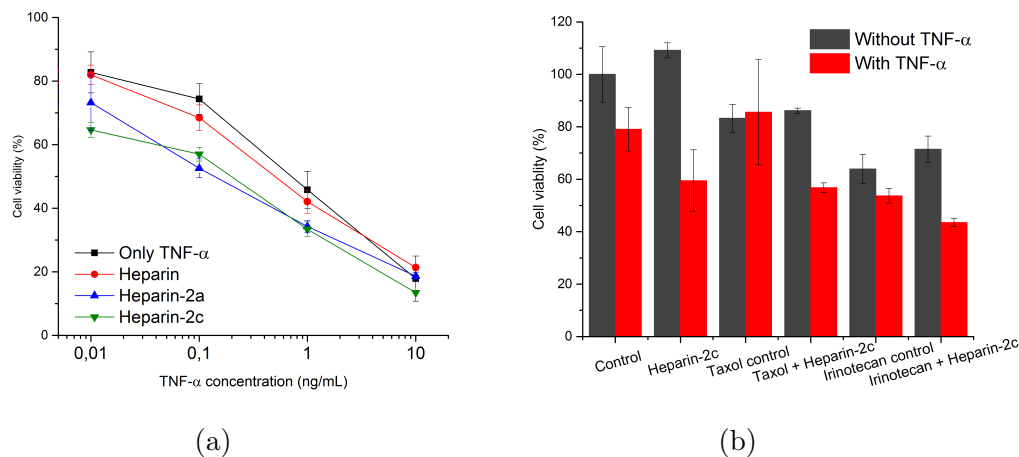
FIGURE 3.59: Heparin and its derivatives (5 μM) cytotoxic effect when incubated with increasing TNF-α concentrations (a). Co-incubation of 5 μM taxol or irinotecan with 5 μM heparin-2c, with or without 0.1 ng mL^{-1} TNF-α (b). All conditions are in the presence of 8 μM CHX.

TNF-α Modifications Effect on Cytotoxicity

TNF-α-NVOC During this study, we modified TNF-α with fluorescein and Cy5. When these modifications were performed, the affinity to DNA seemed to be diminished or even totally destroyed. As we realised this interaction could be manipulated, we modified TNF-α with a NVOC group to create a light-switch that could be used to change its interaction with DNA. Using a ratio of 5 eq. of NVOC per TNF-α monomer, we generated a TNF-α that resembled the unmodified version, with or without UV exposure (Figure 3.60). The caged TNF-α co-incubated with DNA had a lower cell death enhancing effect, when compared with the controls of unmodified TNF-α. Adding DNA, the cell viability decreased more significantly in the uncaged version of TNF-α.

When TNF-α was modified in a ratio of 25 eq. per monomer, the controls with no DNA were again similar to no UV exposure, as well as the conditions with DNA. The caging/uncaging association with DNA effect seemed to be gone, see Figure 3.61.

111
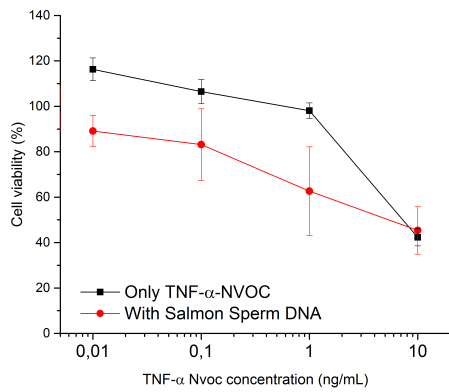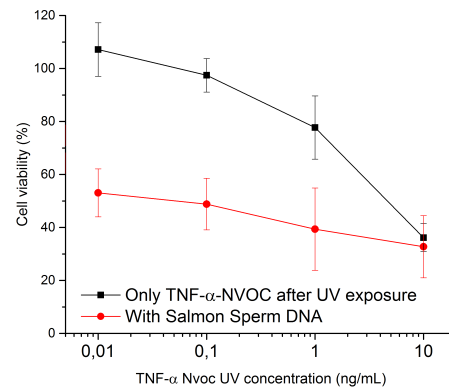
(a)   (b)

FIGURE 3.60: Influence on cell viability of TNF-α-NVOC with (a) and without UV exposure (b) with 5 µM S-DNA, in the presence of 8 µM CHX. TNF-α-NVOC is in a five times excess over each TNF-α monomer.
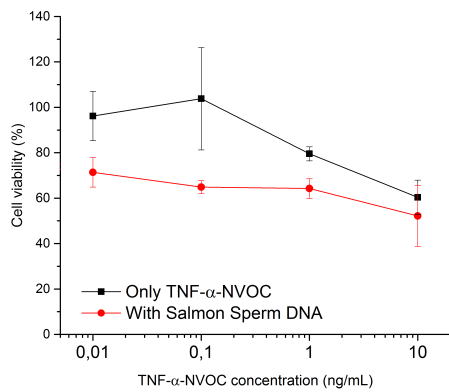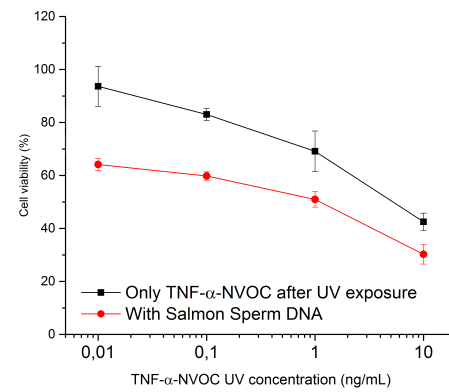
(a)   (b)

FIGURE 3.61: Influence on cell viability of TNF-α-NVOC with (a) and without UV exposure (b) with 5 µM S-DNA, in the presence of 8 µM CHX. TNF-α-NVOC is in a 25 times excess over each TNF-α monomer.

3.2.3   Circular Dichroism Spectra of TNF-α Structure

Using CD, we monitored TNF-α's structural shifts following α-helix markers, which are characterised by decreased ellipticity at 208 and 222 nm. Also, the β-sheet marker, which shows as a decrease at 217 nm, was tracked. As reference and con-

trols for this section of experiments, we provide the stability data for TNF-α (Figure 3.62a) and the standard DNA spectra of the constructs used (Figure 3.62b).
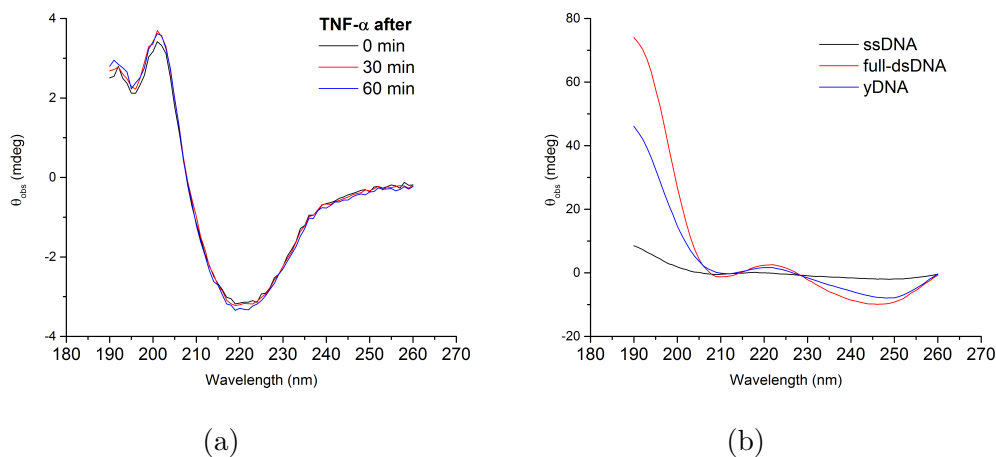
FIGURE 3.62: 10 µM TNF-α stability for 60 min (a). 10 µM DNA standard spectra (b).

On Figure 3.63a, we see that the native TNF-α conformation shifted to lower ellipticity values in the 200 to 240 nm range, when co-incubated with hybridised DNA, two fully-complementary strands (full-dsDNA), or yDNA. When the α-helix markers were plotted (Figure 3.63b), the values for 208 and 222 nm were more negative than the control when TNF-α was incubated with DNA, suggesting a larger presence of α-helix in the trimer. Also the typical β-sheet marker was decreased in comparison to the TNF-α control.

113

(a)   (b)
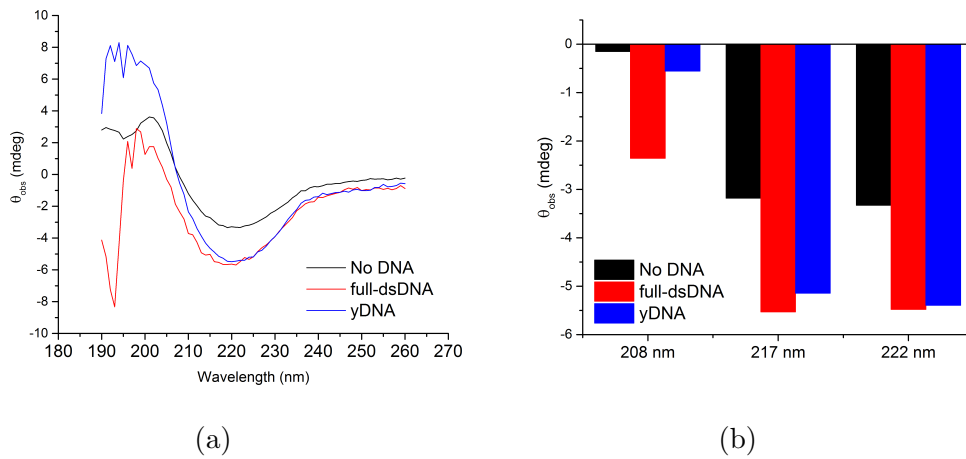
FIGURE 3.63: $10\,\mu\mathrm{M}$ DNA incubated with $10\,\mu\mathrm{M}$ TNF-α for $60\,\mathrm{min}$ (a). 208 and $222\,\mathrm{nm}$ are α-helix markers, and $217\,\mathrm{nm}$ is a β-sheet marker (b).

Altogether, these values suggest that the TNF-α native structure was indeed altered by DNA presence. Specifically, the increase in α-helix content might have been related to the increased TNF-α-induced cell death. More DNA structures (ssDNA, half-dsDNA, and S-DNA) were also tested and there were always significant shifts in TNF-α's conformation. Both α-helix and β-sheet markers were altered (Figure 3.64).

114
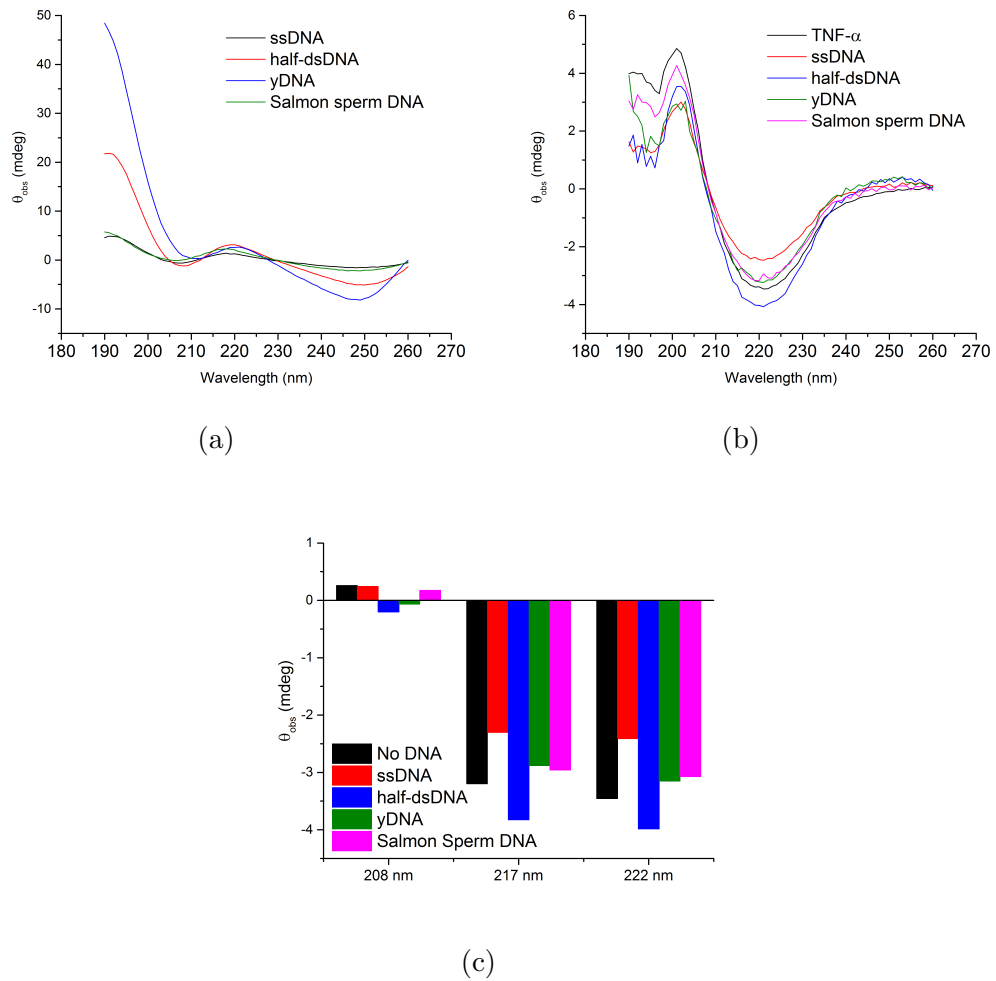
FIGURE 3.64: $10\,\mu\text{M}$ DNA spectra (a). $10\,\mu\text{M}$ DNA incubated with $10\,\mu\text{M}$ TNF-α for $5\,\text{min}$ (b). 208 and $222\,\text{nm}$ are α-helix markers, and $217\,\text{nm}$ is a β-sheet marker (c).

3.2.4 TNF-α Monomers Dynamic Study

The hypothesis was that TNF-α monomers dynamic is affected by the presence of charged biopolymers, such as DNA and heparin. For that end, we explored many possible methodologies, but ended always with inconclusive results.

The presented homoquenching experiment, albeit unsatisfactory, gave us some interesting preliminary results. The experiment was based on the concept of relief of

fluorescence homoquenching [200], meaning that the fluorescent moieties of the same kind are close enough in space that they quench themselves. For this experiment, we labelled TNF-α with Cy5, proceeded to mix unlabelled with labelled entities and co-incubated them with S-DNA (Figure 3.65a). It can be seen that 2 µM unlabelled TNF-α yielded no fluorescence signal. When 10 nM TNF-α-Cy5 was added to the unlabelled TNF-α, the fluorescence signal was higher than TNF-α-Cy5 alone, suggesting that the dynamic exchange of monomers lead to reduced self-quenching. Adding S-DNA to the system, in the presence of only labelled TNF-α, generated almost the same signal in both conditions. When TNF-α unlabelled was added, there was an increase of fluorescence, because the fluorescent signal was released. Normalising the results by the fluorescence obtained for 10 nM TNF-α-Cy5, when DNA was present, the fluorescence was less than in its absence (Figure 3.65b). Throughout 120 min, the release of fluorescence was less when DNA was present, suggesting that the system was not as dynamic as in its absence.

(a)   (b)
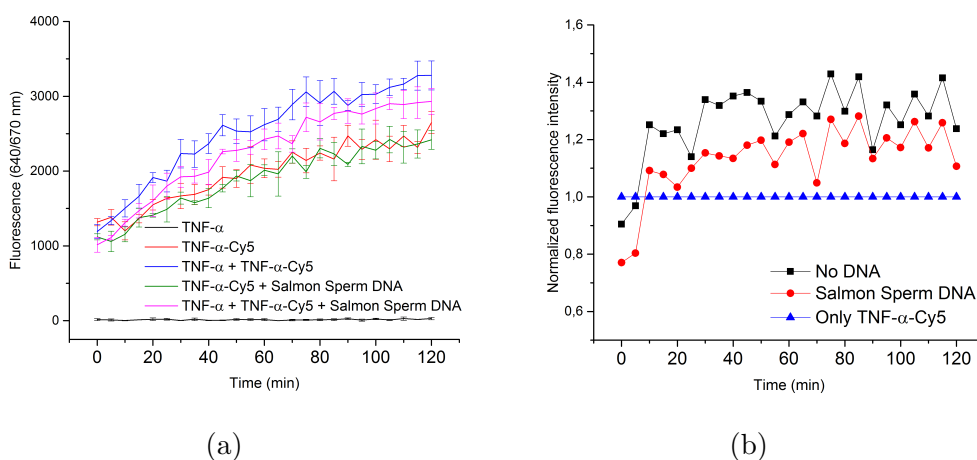
FIGURE 3.65: Incubation result of 2 µM TNF-α with 10 nM TNF-α-Cy5 and 5 µM S-DNA during 120 min (a). Fluorescence normalisation with 10 nM TNF-α-Cy5 (b).

116

4

Discussion

4.1   DNA-Based Diversity Modelling and Analysis

This project proposes a synthetic DNA library as an ideal medium for establishing experimental systems to model large populations of high complexity and dynamics. For example, a randomised 20-base sequence (N_{20}) contains more than one trillion different DNA sequences (4^{20}). Where only $10\,\mu L$ of $10\,nM$ N_{20} solution has the size of the human population and, statistically, each molecule represents a unique individual.

To explore this effect, we assigned a fully complimentary duplex with a F of 1.0. Mismatching leads to lower F values, e.g., $A - X_{10} - B/A' - X'_{10} - B'$ (medium diversity) and $A - X_{20} - B/A' - X'_{20} - B'$ (high diversity) duplexes have F values of 0.5 and 0.1, respectively (X and X' indicate mismatching in the randomised region), see Figure 3.9. Astonishingly, throughout the PCR cycles, although the production of full length DNA was not affected by using either $A - S_1 - B$ or various libraries as templates, the time courses of F values differed from each other dramatically (Figure 3.2).

We postulate that in the initial phase, primer concentrations are much higher than

those of full length DNA and probabilistically self-assembled $A - X - B/B'$ and $A' - X' - B'/A$ are more abundant than $A - X - B/A' - X' - B'$ (Figure 4.1a). $A - X - B/B'$ and $A' - X' - B'/A$ produce $A - X - B/A' - X' - B'$ in the presence of DNA polymerase (X/X' means the fully complimentary pair synthesised by DNA polymerase, according to the template). The newly synthesised fully complimentary $A - X - B/A' - X' - B'$ has an F value of 1.0. In later phases, primers are consumed and full length DNAs are abundant. Hence, $A - X - B/B'$ and $A' - X' - B'/A$ are much less prevalent than $A - X - B/A' - X' - B'$ (Figure 4.1b). The duplexes generated by annealing form mismatching pairs with lower F values, in contrast to the duplexes produced by polymerase, where F = 1.0. There is a turning point in the course of a high diversity library, as the increase of F caused by newly synthesised fully complementary duplexes equals the decrease of F caused by the re-annealing of DNA. This generates mismatched pairs through re-shuffling among the library members (Figure 3.8).

To relate F to an experimentally measurable parameter, we explored the quick annealing effect of mismatched templates that are less stable at $72\,°C$, leading to weaker dye binding and lower fluorescence (Figure 3.10). Experimentally, we demonstrated that mismatches between strands had a direct negative effect on fluorescence signal and melting temperature. It was observed that the mismatched templates were somehow paired, e.g., two and four mismatches, along the decreased melting temperature values. This similar effect for different mismatched samples reflected the fluorescence signal decrease obtained during the screening of diverse libraries. The largest difference observed in the fluorescence signal monitoring occurred between X_8 and X_{10}. As for the melting curves, there was not that sharp difference, but there was a clear cut between types of samples. These types of samples seemed to be comprised of similarly mismatched templates. As an explanation for these results grouping, we believe could be the physical interaction per se of SYBR Green
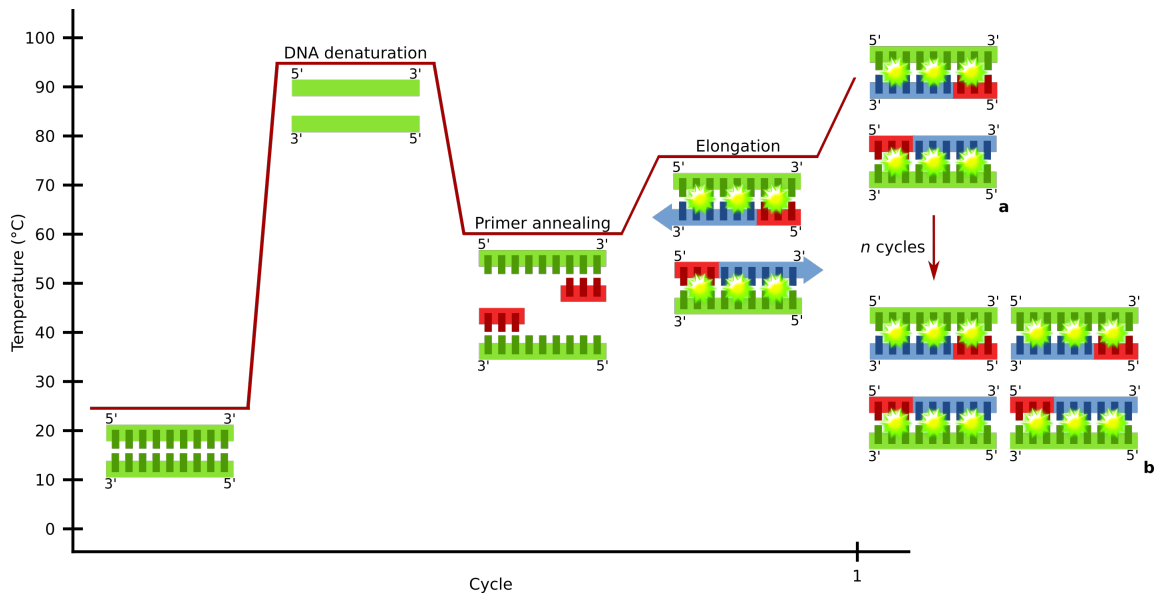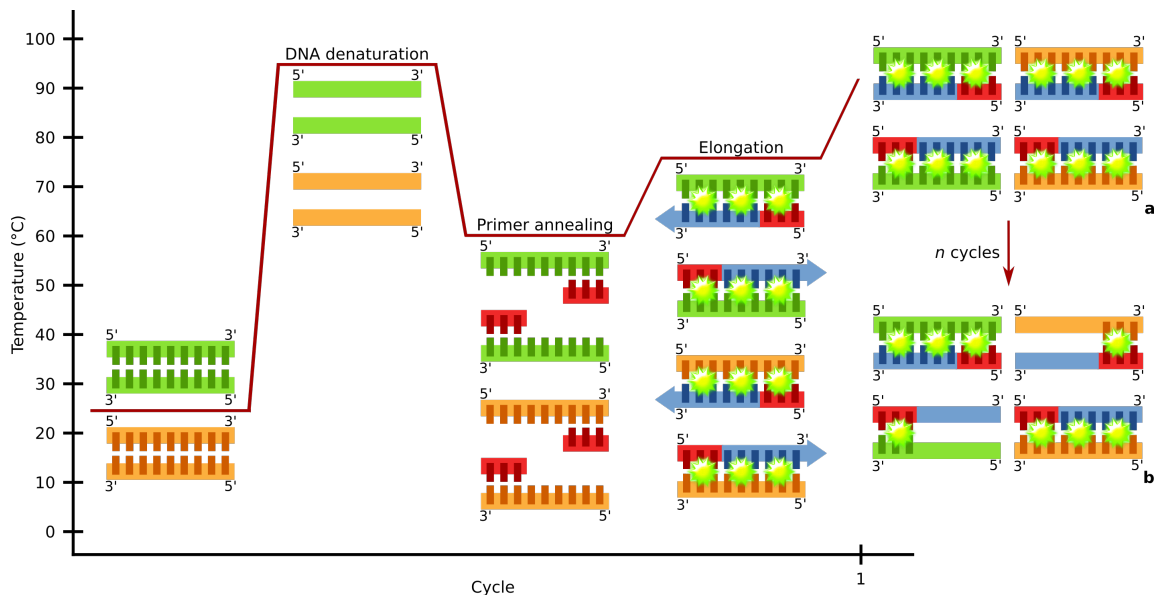
FIGURE 4.1: PCR amplification scheme when there is one (a) or many (b) templates. The green star depicts the intercalated fluorescent dye.

I to dsDNA. As referred to in the introduction, SYBR Green I intercalates dsDNA in portions of 3.5 bases and in its minor groove. The specificity of this interaction might be the explanation for such results.

119

We amplified libraries of different diversities in optimised conditions, see Figures 3.3 and 3.4 with libraries and primers concentration titration, respectively. In parallel, a simulation of the same conditions was done (Figure 3.18). The lower the diversity, the closer the amplification curve would be to the one template curve. Increasing diversity led to the transformation of a sigmoidal (lowest diversity) to a bell shape curve (highest diversity). This effect was consist among concentrations.

We continued our study with the study of the randomised sequences' nature. To start, and emulating DNA strands carrying codes in a DECL setting, we created templates that had the same total number of randomised nucleotides, 20, but distributed in different segments. For a given library X_n, when n is larger than 16, each sequence in $10\,\mu L$ at $1\,\text{nM}$ is statistically unique. However, each sequence in $X_{2/10}$ and $X_{4/5}$ is represented a few thousand times, see Figure 3.11. Even though the degeneration was very different, the amplification had similar profiles.

To understand in more detail the impact of diversity in the amplification, we designed ten different templates (S_x) with a central segment of 20 nucleotides in which, at least, 15 nucleotides were different among each other. This design was opposed to the original one (X_{20}), which had a central segment with 20 randomised positions. This randomisation could mean zero to 20 different nucleotides. This new library had a higher intrinsic diversity than X_{20}. The amplification curve for a mix of the ten highly-diverse templates should have a profile similar to X_{20}, if the dynamic duplexes reshuffling did not occur during the annealing step in each cycle. However, the curve obtained showed that thermodynamic re-equilibration played a part, even though it was too inefficient to create perfect matching between all sequences (Figure 3.13). Decreasing the sub-library number gradually generated a standard RT-PCR curve. Another way we studied the oligonucleotide population was by manipulating its growth, considering primers essential resources. Primers are inter-dependent and limiting one damages the population exponential growth, as described in AsyPCR.

120

Applying classical AsyPCR, one primer is in a much lower concentration, reducing the entire molarity of the system (Figure 3.15). The kinetics are governed by the use of one primer, preventing the comparison of libraries of different diversity. Modifying this method, we performed AsyPCR with primers shorter than the standard 20 bp (A_{12}, A_{15}, A_{18}), with lower annealing temperatures (Figure 3.17). Polymerisation with these primers, plus B', becomes the rate-limiting step. A_{18} was the only primer able to produce similar amplification to A, while A_{15} showed a strong delay and A_{12} had no amplification at all. When a highly diverse library was used as template, the simulation experiment predicted that the turning point would be later and appear at a lower F value. The synthesis of fully complementary duplex became slower, so the resulting increase of fluorescence was exceeded by the counter effect associated with DNA reshuffling at lower full length DNA concentration. In particular, $A_{15}B'$ illustrates the library diversity and the effect caused by primer imbalance and enabled the generation of kinetic-controlled DNA synthesis. This design could be used to model conditions in which the inefficient supply of one resource will become decisive for a population's progress, as in an adaptation scenario.

Based on the AS-PCR method, we elongated the primers to bias population growth. A_{20+1}, A_{20+5} and A_{20+10} selectively amplified $1/4$, $1/4^5$ and $1/4^{10}$, respectively, of X_{20}, reducing diversity (Figures 3.20 and 3.21). The amplified products have decreased diversity. This setup enables the discrimination by nucleotide. Only templates able to pair with the extra nucleotides presented by the primers are able to carry amplification. So, it is possible to pinpoint a defined sequence.

To further prove our method as a possibility to model heterogeneous populations under dynamic progress and systems that can be pressured by other selective forces, libraries with restriction sites were designed, X_{EcoRV} and X_{SmaI}, see Figure 3.25. As the full complementary duplexes are the substrate of the used endonucleases (Figure 4.2a), these segments could be used as a selection mechanism during population

growth. The crescent heterogeneity of the system caused by DNA reshuffling yield population protection against endonuclease digestion, due to the elevated number of mismatches (Figure 4.2b).

During this experiment, we observed for the first time the reocurrence of fluorescence after an extended number of amplification cycles, see Figure 4.2c. Analysing these RT-PCR products of high fluorescence signal by urea-PAGE (Figures 3.30 and 3.31) and AFM (Figures 3.33 and 3.34), we visualised large and complex 3D DNA structures. On the basis of this assembling, we hypothesise that a mixture of a large amount of random templates could have a staggering number of potential interactions. This leads to structures with partial complementarities involving more than two strands. With a library as diverse as X_{20}, the number of interactions possible surpasses the power of computation. Therefore, we suppose that a highly diverse library under polymerase action could evolve complex structures, more than a non-diverse library. Contrastingly, such structures have not been observed with genomic DNA samples [212].

To establish our method's relevance as also a tool and not only a modelling system, we designed an experiment where our method could be used as a complement of DECL. A mock-library with SA-resin plus biotin- (strong) or iminobiotin-modified (weak) oligonucleotides were used as models: *LibA* represented a library with few potent binders without weak binders; *LibB* simulated a library with few weak binders; *LibC* had many weak binders; and *LibD* represented a library with few high affinity ligand and many weak binders (Figure 3.36). Mimicking stringency by varying the number of washes, it was possible to evaluate the population behaviour towards such conditions. After one washing step, all weak and strong binders were present in the respective populations. After two washing steps, the strong binders became the main population in *LibA* and *LibD*. In *LibB*, only one type of weak binder was present and this continued to be the major binder due to lack of competition. *LibC* profile

122
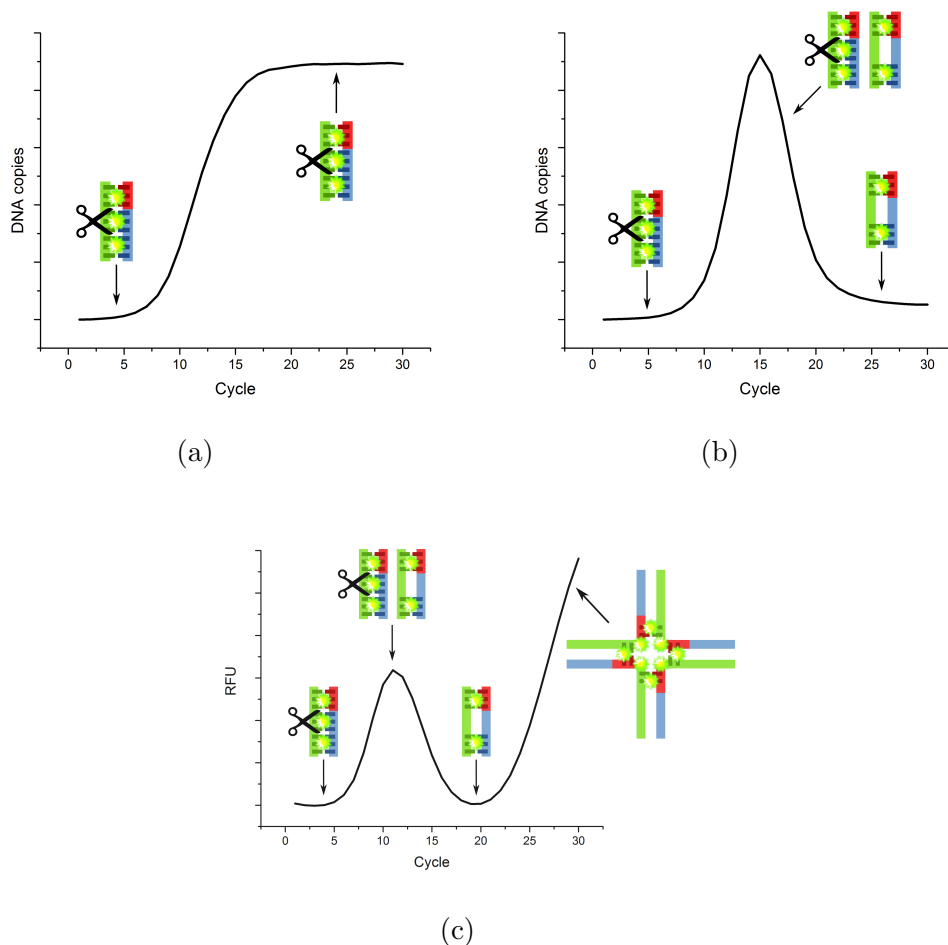
FIGURE 4.2: Amplification curve when there is one template and there is protection against restriction enzymes (a). When there are many templates, there is no protection due to mismatch (b). Reoccurrence of fluorescence for prolonged PCR amplification (c). The green star depicts the intercalated fluorescent dye and the scissors represent the restriction enzymes.

did also not change, because the whole population was composed of weak binders and, once again, no strong competition was present. Using this method, it is possible to estimate the amount of DNA conjugates captured by the protein. Furthermore, it can also distinguish different types of populations, according to their binding potency, which is even more relevant for DECL. Piggy-back ridding on the existing DECL setups, where libraries contain millions of different compounds with respective codes generated through combining sub-codes that contain tens to hundreds of

members flanked by two constant strands, we can analyse diversity of each sub-code region.

4.2 Physical Environment Regulation of TNF-α Activity

The exact mechanism by which TNF-α rules its dynamics is still blurred. It is known that this cytokine must be in a trimer conformation to exert activity. Monomers and dimers are widely assumed to be inactive, see Figure 4.3. The triggering of inflammation cascades probably plays an important role in how TNF-α finally assembles into an active trimer in vivo.
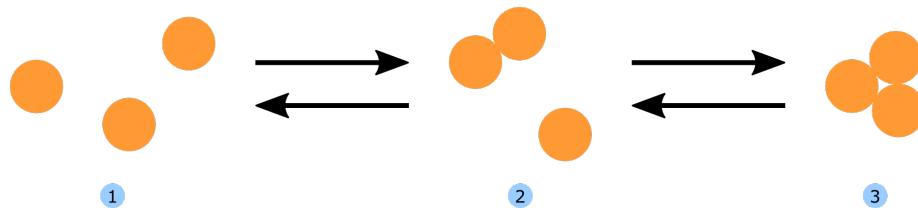
FIGURE 4.3: Proposed mechanism for TNF-α's dynamic structure. At low concentrations, it exists as a monomer (1). As the concentration increases, it assembles into a dimer (2) or even a trimer (3). All the conformations are not covalently linked and are subject to temporal variations of concentration.

The binding, bioactivity, and structural results we got of DNA's interaction with TNF-α are strong pieces of evidence supporting the idea that this direct relationship exists. We hypothesise that DNA could alter TNF-α structure and, consequentially, its activity. These interactions are thought to be linked to electrostatic interactions between the positively charged TNF-α and the negatively charged oligonucleotide. So, theoretically, this effect could be observed for any other negatively charged biopolymer, such as heparin. Some preliminary data points in this direction.

For a better understanding of TNF-α's direct interaction with DNA (Cy5-modified or wild type), a set of band-shift assays was performed. At high concentrations of TNF-α (50 µM), DNA's migration was shifted or generated smears, see Figure 3.38. Interestingly, there was often a residue amount of DNA blocked inside the wells, which seems to be caused by the non-covalent cross-linking between the trimer and DNA. The same was not observed for the yDNA and Xsmall constructs (Figures

3.39 and 3.40), suggesting its binding to TNF-α is established in a different manner. Titration experiments were performed and the estimated K_d of TNF-α to DNA was in the micro-molar range. Using another method, MST, values obtained for the binding affinity of yDNA were also in the low micro-molar range (Figure 3.52).

We proceeded to study the interaction of TNF-α with DNA in more detail. Using BLI, we tested TNF-α's interaction with a negatively charged environment. To start, we saw that TNF-α had strong binding to the carboxyl-modified surface (ARG sensor) and no binding to the amino-modified one (APS sensor), see Figure 3.43. After, we observed that co-incubation with 1 µM DNA was able to inhibit in large extent the binding of TNF-α to the negatively charged surface (Figure 3.48). The results were even more pronounced using 5 µM DNA, up to around 50% (Figure 3.47). Similar results were obtained using highly sulfated oligosaccharides, such as heparin (Figures 3.50 and 3.51).

The co-incubation with negatively charged biopolymers, such as DNA, RNA or heparin, increased TNF-α's cytotoxicity. For all DNA shapes tested, there was a consistent increase of about 20% cell death in every condition, see Figure 3.54. The same profile was obtained while using heparin and its derivatives, see Figure 3.59a. This synergy was not dependent on CpG motifs immunostimulation (Figure 3.56) or other cytotoxics lethality (Figure 3.57). We also prepared a light-switchable TNF-α, where light had little effect of TNF-α itself, but made the whole system more sensitive to DNA-mediated up-regulation of TNF-α activity (Figures 3.60 and 3.61).

We studied too how TNF-α's conformation is changed throughout these interactions by CD spectroscopy, see Figures 3.63 and 3.64. Different forms of DNA seemed to generate conformational changes, such as dsDNA and yDNA reducing the ellipticity at 217 nm strongly, an α-helix marker. TNF-α has a very weak signal at 250 nm, so the changes at this point can be assigned to the structural changes in DNA. Also, DNA shows a very weak signal at 217 nm. Thus, spectra interferences can

be excluded. There are many reports about how high temperatures and low pH can generate α-helix in TNF-α and how it is linked to higher cytotoxicity [191]. Thus, we believe the enhanced bioactivity when co-incubating with DNA could also be due to the appearance of α-helix. DNA interferes with the correct disposition and structure of TNF-α, inducing significant changes in its bioactivity.

Our final goal was to understand in detail the dynamics of TNF-α monomer assembly and how this could be influenced by the physical environment. We tried a large range of methods, e.g., homoquenching, enzyme-linked immunosorbent assay (ELISA), or Fluorescence Correlation Spectroscopy (FCS). However, none gave us a clear result explaining this effect. The main drawback was the detection limit, even for ELISA. We had to work in a range of low pico-molar concentrations, guaranteeing that all TNF-α present in solution was in a monomer conformation and not in a dimer or a trimer. In most attempts, only higher concentrations of TNF-α were detectable or the effect was so small that we could not register it. So, our intention of demonstrating how DNA interferes with these assemblies was not possible to fulfil. Nevertheless, the results from the homoquenching experiment gave us some hints about the system in accordance to our initial hypothesis (Figure 3.65). The results show a difference in the whole system dynamics with or without DNA, suggesting it is able to lock the free movement of TNF-α monomers. More experiments must be done.

As discussed in the introduction, the average cfDNA in blood is in the low micro-molar range [213] and it can be detected in several pathologies, from cancer to auto-immune diseases. For TNF-α, concentrations can range from low pico-molar to high nano-molar, according to the physiological situation, e.g., higher in inflammation [171]. The known binding affinity of anti-TNF-α Abs, current therapies, is around pico-molar and the affinity values we propose of DNA to TNF-α are in the micro-molar range. Thus, this would fit quite well the scenario in which cfDNA

would be called to interact with TNF-α, when this is elevated, see Figure 4.4. DNA would be able to bind to TNF-α, change its structure and activate it for a stronger biological response.
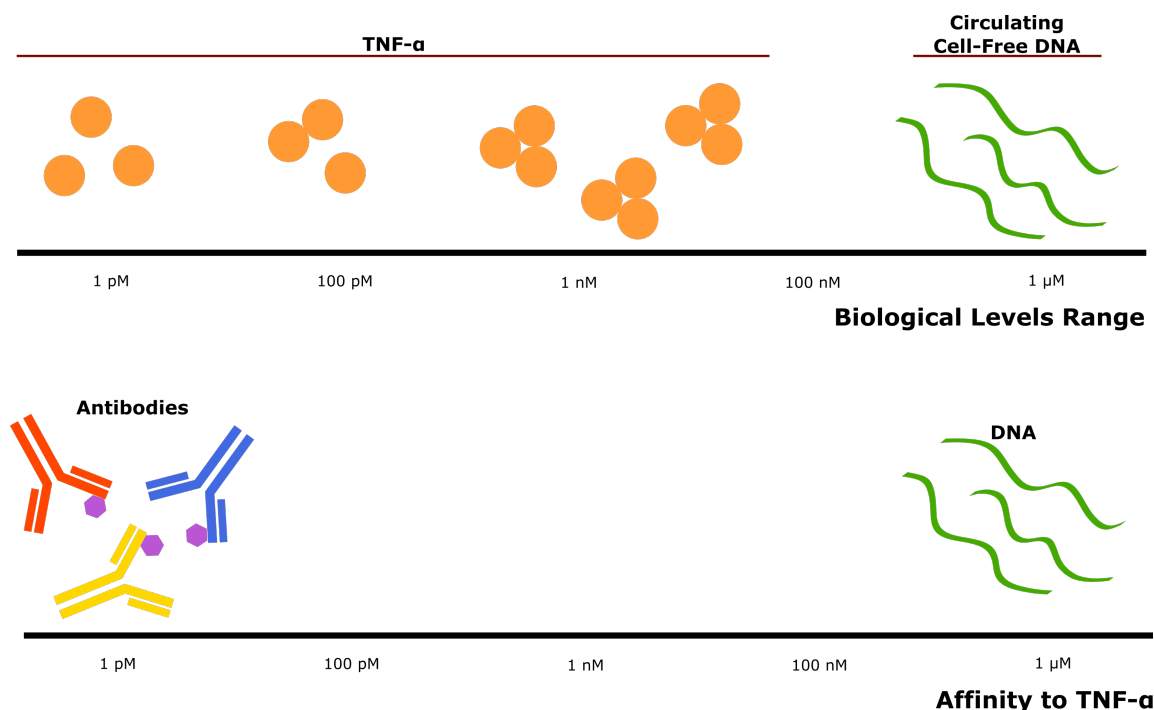
FIGURE 4.4: Comparison of biological levels in vivo and affinity to TNF-α by DNA and Abs. The specificity of Ab detection and binding is based on its ability to recognise TNF-α in very low concentrations (pico-molar). DNA's interaction with TNF-α is unspecific and only possible when both concentrations are closer in a micro-molar range. This rough regulation prevents the activation of TNF-α by DNA when concentrations in vivo do not cross the reactivity threshold.

As many different biomolecules are present in blood in a wide range of concentrations, their local concentrations can be changed and reflect whichever stimulus or signalling at a given time. These changes can result from genetic expression or interactions with other bioactive molecules. This unspecific and low-potency regulation of DNA on TNF-α can happen in vivo: if the abundant biomolecule X (e.g., DNA or heparin) interacts and regulates the signalling molecule Y (e.g., TNF-α), the concentration fluctuation of X must be in the same range of the K_d between X and

Y. In other words, Y's activity must be regulated in a subtle way in a physiological context; X is abundant and interacts with Y altering its activity. X biofunction can be influenced by its own concentration (C_Y) and X's concentration (C_X). If C_X is constantly much higher than the dissociation constant between X and Y (K_{dXY}), under physiopathological conditions, X becomes a pseudo-intrinsic feature of the biological activity of Y. If C_X is constantly much lower than K_{dXY}, it will have no effect on Y's activity. So, if C_X is in the range of K_{dXY}, a change in C_X can lead to a modification in the biological activity of Y. Other regulation mechanisms, such as post-translational modifications, can also be involved.

5

Conclusion

This dissertation explores DNA oligomers in two distinct functions: as a data modelling system and as a participant in physical interactions with other biomolecules. On both functions, DNA's abilities are explored beyond the classic storage and regulation of genetic information.

As a coding unit, we illustrate the relationship between code-diversity and dynamics of self-assembly. The DDMA method described is based on a very simple principle and its experimental measure is accurate and robust, which is corroborated by mathematical simulations. Exploring the design of different libraries and using primers with various kinetics and selectivity, we demonstrated that we could analyse, manipulate and evolve a pool of DNA strands. The DDMA method is able to:

- evaluate quantity, quality and diversity of a DNA library,

- tune kinetics of primers' binding to a template, biasing growth and modelling specific situations,

- control primers' selectivity for a template, reducing a diverse population,

- incorporate restriction sites, creating an experimental model in which heterogeneity has protective effect,

- and form large DNA structures when the system is subjected to extensive PCR cycles.

Our diverse DNA libraries behave as synthetic societies and can be used to model heterogeneous populations. They can simulate different types of diversity and function. The dynamic interactions among the individual members govern the time course and final product of a PCR reaction. These can be used to build models to verify hypotheses and lead to new discoveries. As postulated by [37], DNA could be the solution to one of the limitations existing in population modelling and be the simplest setup able to simulate an observation.

The unexpected formation of complex 3D DNA structures in extended PCR cycles showed a material fundamentally different from low-entropy structures generated by methods such as DNA origami. Clearly aware that DNA amplification under PCR conditions does not occur in cells, a related DNA amplification process in a pool of diverse sequences and the production of high-entropy structures can take place in physiological conditions, e.g., DNA autoimmune disease. DNA-associated autoimmunity has been found in many pathological conditions, but its mechanism remains unknown.

Furthermore, the RT-PCR method also has application in biochemical analysis, e.g., to analyse the mixture of oligonucleotides from selection experiments (SELEX or DECL). Plus, the use of DNA libraries as synthetic societies can open new venues to model complex and diverse populations.

In the concrete case of DECL, some of its major struggles is to unravel a few strong binders from large combinatorial chemical libraries, being able to reduce the impact of unspecific or weak interactions between DNA-conjugates and target proteins. Only

after sequencing and data analysis, it is possible to identify the hit compounds and this is a bottleneck. DDMA can evaluate almost instantly the library diversity during the selection process. Our method improves information regarding how much DNA conjugates were captured by protein and how strong their affinity might be. Since DECLs can have millions of different compounds, the codes are generated through combining sub-codes. Each sub-code region typically has tens to hundreds of members flanked by two constant strands. So, assuring there is a primer region, by its manipulation, the diversity of each sub-code region can be easily analysed through DDMA.

Our method represents a different approach to use DNA as variable and information storage [160]. As compared to more complex designs of DNA computing with up to hundreds of unique strands [159], our libraries can harbour more than one trillion different sequences and, in principle, has no size limit. Though, it is not aimed at performing classic operations, DDMA can be used to model systems beyond the power of computation. It is possible to simulate the collective behaviour of a heterogeneous and dynamic population and its underlining probabilistic structure associated with an astounding number of random interactions.

The ability of DNA to participate as an active unit in the regulation of a target protein, TNF-α, altering its physical environment also caught our attention. We explored the interaction between DNA, also heparin, and TNF-α. This binding was consistently estimated to be around the low micro-molar range and structural conformation studies confirmed changes in TNF-α structure. All these results were translated into a regulation of biological function, when we observed that TNF-α's inherent cytotoxicity was further boosted by the presence of negatively charged polymers. The proposed explanation for this binding is the existence of unspecific interactions between the positively charged TNF-α and the negatively charged DNA. We believe this observation fits the expected in vivo dynamic range of DNA. The

interaction among two biomolecules is not necessarily specific or potent because specificity relies on the biological context. These interactions are not only about specific ligand-receptor bindings, but also about the non-specific ones. As binding depends not only on the dissociation constant, but also on concentration, some interesting effects can be seen. How non-specific interactions can affect specific interactions between molecules of lower concentrations is often elusive. In this study, the variation of DNA concentration in blood would regulate TNF-α's action.

It is known that the existing therapy regarding TNF-α is insufficient and it is largely a consequence of the relative ignorance we have about how this player affects and is affected by the various molecular events in which it is involved. To develop an appropriate small molecule for this end, it is needed a detailed study of how TNF-α and its receptors are regulated, including their concentration, conformation, and oligomerisation states. Our work scratches the hypothesis of how the physical environment of this protein can be determinant for its structure and consequent activity. Summarising, this dissertation's main goal is to demonstrate DNA oligomers as active units that can be used as bionanotools in studies at the molecular level. Also, DNA as a coding unit is able to perform operations at levels that current computation methods are not able to achieve, while being useful to model systems and retrieve information from them. At last, but not the least, we generated high-entropy structures that could be similar to the ones formed in vivo in DNA-related pathologies. The link between these DNA complex and unspecific structures to the activation of proteins, such as TNF-α, by the modification of their physical environment is another great issue that this dissertation tries to unveil.

6

Outlook

6.1 DNA-Based Diversity Modelling and Analysis

To prove that the DDMA method can be used in a real DECL selection, we are planning to explore the synergic binding of two vancomycins with a dAdA modified resin. The goal is to demonstrate DDMA can distinguish populations with no (no vancomycin), weak (one vancomycin), or strong binders (two vancomycins). In this selection, each small molecule will be carrying an oligonucleotide strand with very diverse sequences, the coding regions. Therefore, the diverse oligonucleotide segments attached to the small molecules should take part in the self-assembly dynamics phenomenons we have observed.

Later, we want to further develop this method for modelling large and dynamic heterogeneous populations using DNA libraries. The goal is to develop a toolbox that can be easily adapted to create any kind of simulation. As in an *'enclosed ecosystem'*, where different variables can be applied and results can be read quickly.

We would like also to understand the molecular mechanism behind the formation of the high-entropy DNA structures and to know if these are associated with patho-

physiological conditions. And if so, how.

6.2 Physical Environment Regulation of TNF-α Activity

For this project, I envision the unravelling of the monomers' assembling dynamics as the number one goal. The follow-up of this discovery is the possibility to develop improved small molecules, or other entities, able to disrupt these protein-protein or protein-DNA interactions.

Immediately, we are focusing on determining different DNA shapes affinity to TNF-α in a more accurate way using MST. Also, we plan to image with AFM TNF-α co-incubated with DNA. We expect to see the formation or not, with or without DNA, respectively, of stabilised TNF-α oligomers.

The development of a switchable TNF-α construct is also in our plans. However, the modifications we created so far decreased its activity and binding affinity to DNA. More development must be done on this field, probably with the use of engineered proteins. For now, we will setup a system where TNF-α can interact with DNA-biotin in the presence or absence of SA beads. In cell culture, we expect to see a clear 'on/off' effect with the co-incubation or not with beads. It is not a protein switch, but more of a DNA sequester experiment.

At last, we would like to understand in more detail how TNF-α's interaction with the physical environment, e.g., negatively charged polymers, translates into the molecular biology level. We want to study the effect this interaction might have on the NF-κB pathway.

7

Summary

This dissertation focuses on rationalised DNA design as a tool for the discovery and development of new biotechnological methods, as well as understanding the biological function of DNA beyond the storage of genetic information. This study is comprised of two main areas of study: (i) the use of DNA as a coding unit to illustrate the relationship between code-diversity and dynamics of self-assembly; and (ii) the use of DNA as an active unit that interacts and regulates a target protein.

DNA-Based Diversity Modelling and Analysis   In the study of DNA as a coding unit in code-diversity and dynamics of self-assembly, we developed the DDMA method. Using PCR and RT-PCR, we studied the diversity and evolution of synthetic oligonucleotide populations. The manipulation of critical conditions, with monitoring and interpretation of their effects, helped understanding how PCR amplification can reflect the diversity of a DNA sample.

Experimentally, this study was performed by monitoring the fluorescence signal development during RT-PCR amplification. When amplifying one template, there was generation of perfectly matched duplexes, which translated into a sigmoidal curve

on the RT-PCR. However, when amplifying a diverse sample, such as an aptamer library, this sigmoid became a bell-shaped curve, see Figure 7.1. We hypothesise this phenomenon occurs because:

1. Initially, the amount of primers is higher than the template DNA. For early cycles of amplification, each primer has the possibility to create a fully complementary dsDNA by polymerase action. Then, the perfectly matched DNA duplexes intercalate the fluorescent dye. The fluorescent signal increases exponentially.

2. In later cycles, during the stationary phase there are two possibilities. If a population only has one template, the population of dsDNA will reshuffle and find a perfect match in a population of only forward and reverse strands of the same template. The fluorescent dye is able to intercalate these structures and the signal intensity remains unchanged. If a population is very diverse, when primer concentrations are lower, every time there is a temperature cycle and the dsDNA melt, there is reshuffling. As the subsequent annealing phase occurs, the strands are less likely to meet their fully complementary partner. The perfect match between complementary strands occurs in an exponentially smaller extent, thereby causing a decrease in the fluorescence signal.

DDMA has great value for the study of: (a) code-diversity, convenient in a DNA-based selection method, so semi-quantitation can evaluate a selection development and the population's behaviour can indicate the quality; (b) self-assembly dynamics, for the simulation of a real evolution, emulating a society where selective pressures direct the population's adaptation; and (c) development of high-entropy DNA structures, in order to understand how similar unspecific DNA structures are formed in certain pathologies, such as in auto-immune diseases.
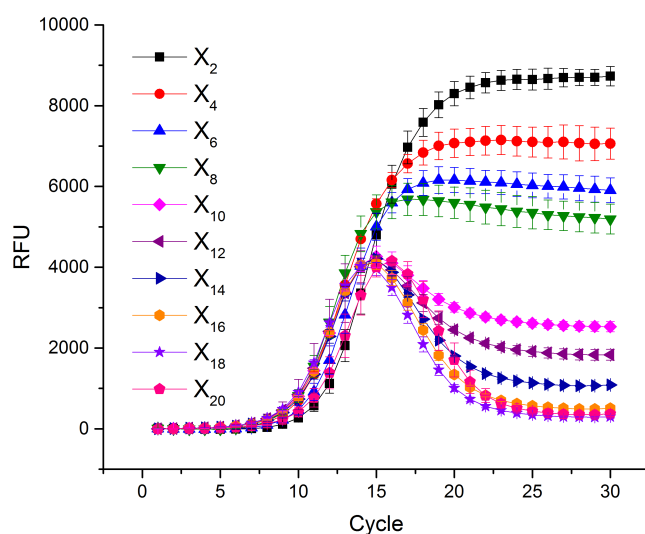
FIGURE 7.1: RT-PCR experiment using DNA libraries of different diversities. X_n corresponds to the diverse libraries, where n indicates the number of randomised nucleotides, from two to 20.

Physical Environment Regulation of TNF-α Activity   To explore DNA as a regulatory unit in TNF-α interaction and activity modulation, we investigated the DNA's influence on its spatial conformation through physical interaction. Active TNF-α is a trimer and the protein-protein interactions between its monomers are a promising target for drug development. It has been hypothesised that TNF-α forms a very intricate network after its activation between monomers and receptors, but the mechanism is still not completely clear.

During our research, we estimated the non-specific DNA binding to TNF-α in the micro-molar range, see Figure 7.2. Cell toxicity assays confirmed this interaction, where DNA consistently enhanced TNF-α's cytotoxic effect. Further binding and structural studies lead to the same conclusion that DNA binds and interferes with TNF-α structure. The probable reason for this binding is the unspecific interaction between the positively charged TNF-α and the negatively charged DNA.

The physical interaction between two biomolecules through, e.g., electrostatic inter-

138

action is often neither potent nor specific. However, a specific regulatory mechanism can occur when the fluctuation of the concentration of one component is in the range of K_d value. From the TNF-α-DNA interaction study, a new mechanism to regulate the biological activity of TNF-α has been suggested.
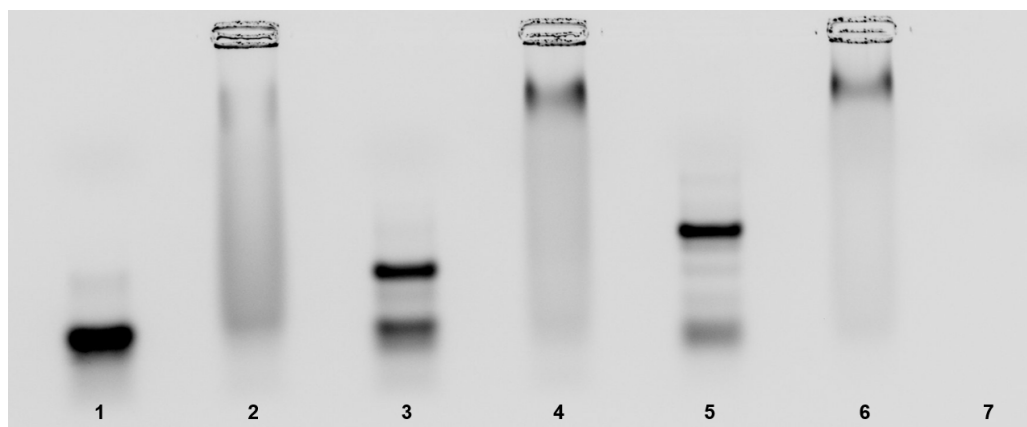
FIGURE 7.2: Band-shift assay resulting from the incubation of 50 µM TNF-α with 1 µM DNA: Cy5-ssDNA (1 and 2), Cy5-half-dsDNA (3 and 4) and Cy5-yDNA (5 and 6). Lanes 1, 3 and 5 correspond to the no protein control. Lane number 7 corresponds to TNF-α only control. 4% agarose gel.

Selbstständigkeitserklärung

Erklärung entsprechend 5.5 der Promotionsordnung

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt bernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Die vorliegende Dissertation wurde am B CUBE - Center for Molecular Bioengineering der Technischen Universität Dresden in der Arbeitsgruppe Biomolecular Interactions von Prof. Dr. Yixin Zhang in der Zeit vom Mai 2013 bis Mai 2016 unter wissenschaftlicher Betreuung von Prof. Dr. Carsten Werner angefertigt.

Meine Person betreffend erkläre ich hiermit, dass keine früheren erfolglosen Promotionsverfahren stattgefunden haben.

Ich erkenne die Promotionsordnung der Fakultät für Mathematik und Naturwissenschaften, Technische Universität Dresden an.

Datum und Unterschrift

Selbstständigkeitserklärung

Declaration according to 5.5 of the doctorate regulations

I herewith declare that I have produced this paper without the prohibited assistance of third parties and without making use of aids other than those specified; notions taken over directly or indirectly from other sources have been identified as such. This paper has not previously been presented in identical or similar form to any other German or foreign examination board.

The thesis work was conducted at B CUBE - Center for Molecular Bioengineering, Technischen Universität Dresden in the Biomolecular Interactions laboratory of Prof. Dr. Yixin Zhang from May 2013 to May 2016 under the scientific supervision of Prof. Dr. Carsten Werner.

I declare that I have not undertaken any previous unsuccessful doctorate proceedings.

I declare that I recognise the doctorate regulations of the Fakultät für Mathematik und Naturwissenschaften of the Technische Universität Dresden.

Date and signature

Bibliography

[1] J. Rao, J. Lahiri, L. Isaacs, R. M. Weis, and G. M. Whitesides, "A trivalent system from vancomycin-D-Ala-D-Ala with higher affinity than avidin-biotin," *Science*, vol. 280, pp. 708–11, 1998.

[2] P. T. Corbett, J. Leclaire, L. Vial, K. R. West, J.-L. Wietor, J. K. M. Sanders, and S. Otto, "Dynamic combinatorial chemistry with hydrazones: Libraries incorporating heterocyclic and steroidal motifs," *Chemical Reviews*, vol. 106, pp. 3652–711, 2006.

[3] F. V. Reddavide, W. Lin, S. Lehnert, and Y. Zhang, "DNA-encoded dynamic combinatorial chemical libraries," *Angewandte Chemie - International Edition*, vol. 54, pp. 7924–8, 2015.

[4] M. Kimoto, R. Yamashige, K. Matsunaga, S. Yokoyama, and I. Hirao, "Generation of high-affinity DNA aptamers using an expanded genetic alphabet," *Nature Biotechnology*, vol. 31, no. 5, pp. 453–7, 2013.

[5] J. D. Watson and F. H. C. Crick, "Molecular structure of nucleic acids," 1953.

[6] T. Dršata and F. Lankaš, "Multiscale modelling of DNA mechanics," *Journal of Physics: Condensed Matter*, vol. 27, p. 323102, 2015.

[7] N. C. Seeman, "DNA in a material world," *Nature*, vol. 421, pp. 427–31, 2003.

[8] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of the Cell*. New York: Garland Science, 4th ed., 2002.

[9] J. Bednar, R. A. Horowitz, S. A. Grigoryev, L. M. Carruthers, J. C. Hansen, A. J. Koster, and C. L. Woodcock, "Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, pp. 14173–8, 1998.

[10] S. Peng, T. L. Derrien, J. Cui, C. Xu, and D. Luo, "From cells to DNA materials," *Materials Today*, vol. 15, no. 5, pp. 190–4, 2012.

[11] A. Ellington and J. D. Pollard, "Synthesis and purification of oligonucleotides," in *Current Protocols in Molecular Biology*, vol. Chapter 2, p. Unit 2.11, 1998.

[12] A. Shivalingam and T. Brown, "Synthesis of chemically modified DNA," *Biochemical Society Transactions*, vol. 44, pp. 709–15, 2016.

[13] B. Hall, J. M. Micheletti, P. Satya, K. Ogle, J. Pollard, and A. D. Ellington, "Design, synthesis, and amplification of DNA pools for in vitro selection," in *Current Protocols in Molecular Biology*, no. Unit 24.2, pp. 1–27, 2009.

[14] H. Hemmi, O. Takeuchi, T. Kawai, T. Kaisho, S. Sato, H. Sanjo, M. Matsumoto, K. Hoshino, H. Wagner, K. Takeda, and S. Akira, "A Toll-like receptor recognizes bacterial DNA," *Nature*, vol. 408, pp. 740–5, 2000.

[15] B. Goldberg, H. B. Urnovitz, and R. B. Stricker, "Beyond danger: Unmethylated CpG dinucleotides and the immunopathogenesis of disease," *Immunology Letters*, vol. 73, pp. 13–8, 2000.

[16] R. S. Illingworth and A. P. Bird, "CpG islands - 'A rough guide'," *FEBS Letters*, vol. 583, pp. 1713–20, 2009.

[17] J. Vollmer, R. Weeratna, P. Payette, M. Jurk, C. Schetter, M. Laucht, T. Wader, S. Tluk, M. Liu, H. L. Davis, and A. M. Krieg, "Characterization of three CpG oligodeoxynucleotide classes with distinct immunostimulatory activities," *European Journal of Immunology*, vol. 34, pp. 251–62, 2004.

[18] V. J. Schüller, S. Heidegger, N. Sandholzer, P. C. Nickels, N. A. Suhartha, S. Endres, C. Bourquin, and T. Liedl, "Cellular immunostimulation by CpG-sequence-coated DNA origami structures," *ACS Nano*, vol. 5, no. 12, pp. 9696–702, 2011.

[19] M. Nishikawa, M. Matono, S. Rattanakiat, N. Matsuoka, and Y. Takakura, "Enhanced immunostimulatory activity of oligodeoxynucleotides by Y-shape formation," *Immunology*, vol. 124, pp. 247–55, 2008.

[20] D. S. Lee, H. Qian, C. Y. Tay, and D. T. Leong, "Cellular processing and destinies of artificial DNA nanostructures," *Chem Soc Rev*, vol. 45, pp. 4199–225, 2016.

[21] N. Matsuoka, M. Nishikawa, K. Mohri, S. Rattanakiat, and Y. Takakura, "Structural and immunostimulatory properties of Y-shaped DNA consisting of phosphodiester and phosphorothioate oligodeoxynucleotides," *Journal of Controlled Release*, vol. 148, pp. 311–6, 2010.

[22] C. Cheng, M. Omura-Minamisawa, Y. Kang, T. Hara, I. Koike, and T. Inoue, "Quantification of circulating cell-free DNA in the plasma of cancer patients during radiation therapy," *Cancer Science*, vol. 100, no. 2, pp. 303–9, 2009.

[23] A. Szpechcinski, J. Chorostowska-Wynimko, R. Struniawski, W. Kupis, P. Rudzinski, R. Langfort, E. Puscinska, P. Bielen, P. Sliwinski, and T. Orlowski, "Cell-free DNA levels in plasma of patients with non-small-cell lung cancer and inflammatory lung disease," *British Journal of Cancer*, vol. 113, pp. 476–83, 2015.

[24] K. Haugbro, J. C. Nossent, T. Winkler, Y. Figenschau, and O. P. Rekvig, "Anti-dsDNA antibodies and disease classification in antinuclear antibody positive patients: The role of analytical diversity," *Annals Rheumatic Diseases*, vol. 63, pp. 386–94, 2004.

[25] R. K. Saiki, S. Scharf, F. Faloona, K. B. Mullis, G. T. Horn, H. A. Erlich, and N. Arnheim, "Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia," *Science*, vol. 230, pp. 1350–4, 1985.

[26] R. K. Saiki, D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich, "Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase," *Science (New York, N.Y.)*, vol. 239, pp. 487–91, 1988.

[27] J. T. Keer, "Quantitative real-time PCR Analysis," in *Essentials of Nucleic Acid Analysis: A Robust Approach*, pp. 132–66, 2008.

[28] J. E. Hardingham, A. Chua, J. W. Wrin, A. Shivasami, I. Kanter, N. C. Tebbutt, and T. J. Price, "Polymerase chain reaction: Types, utilities and limitations," in *BRAF V600E Mutation Detection Using High Resolution Probe Melting Analysis, Polymerase Chain Reaction* (P. Hernandez-Rodriguez, ed.), pp. 157–72, intech ed., 2012.

[29] X. Tang, S. L. Morris, J. J. Langone, and L. E. Bockstahler, "Simple and effective method for generating single-stranded DNA targets and probes," *BioTechniques*, vol. 40, pp. 759–62, 2006.

144

[30] C. Marimuthu, T.-H. Tang, J. Tominaga, S.-C. Tan, and S. C. B. Gopinath, "Single-stranded DNA (ssDNA) production in DNA aptamer generation," *Analyst*, vol. 137, pp. 1307–15, 2012.

[31] M. Svobodová, A. Pinto, P. Nadal, and C. K. O' Sullivan, "Comparison of different methods for generation of single-stranded DNA for SELEX processes," *Analytical and Bioanalytical Chemistry*, vol. 404, pp. 835–42, 2012.

[32] M. Citartan, T. H. Tang, S. C. Tan, C. H. Hoe, R. Saini, J. Tominaga, and S. C. B. Gopinath, "Asymmetric PCR for good quality ssDNA generation towards DNA aptamer production," *Songklanakarin Journal of Science and Technology*, vol. 34, no. 2, pp. 125–31, 2012.

[33] L. Ugozzoli and R. Wallace, "Allele-specific polymerase chain reaction," *Methods*, vol. 2, no. 1, pp. 42–8, 1991.

[34] P. Chen, D. Pan, C. Fan, J. Chen, K. Huang, D. Wang, H. Zhang, Y. Li, G. Feng, P. Liang, L. He, and Y. Shi, "Gold nanoparticles for high-throughput genotyping of long-range haplotypes," *Nature Nanotechnology*, vol. 6, pp. 639–44, 2011.

[35] A. H. Lang, H. Drexel, S. Geller-Rhomberg, N. Stark, T. Winder, K. Geiger, and A. Muendlein, "Optimized allele-specific real-time PCR assays for the detection of common mutations in KRAS and BRAF," *Journal of Molecular Diagnostics*, vol. 13, no. 1, pp. 23–8, 2011.

[36] P.-y. Kwok, "Methods for genotyping single nucleotide polymorphisms," *Annual Review of Genomics and Human Genetics*, vol. 2, pp. 235–58, 2001.

[37] H. Caswell, "Theory and models in ecology: A different perspective," *Ecological Modelling*, vol. 43, pp. 33–44, 1988.

[38] M. R. Hartman, D. Yang, T. N. N. Tran, K. Lee, J. S. Kahn, P. Kiatwuthinon, K. G. Yancey, O. Trotsenko, S. Minko, and D. Luo, "Thermostable branched DNA nanostructures as modular primers for polymerase chain reaction," *Angewandte Chemie - International Edition*, vol. 52, pp. 8699–702, 2013.

[39] Enzoklop, "Polymerase chain reaction," 2014.

[40] C. M. Roth, "Quantifying gene expression," *Current Issues in Molecular Biology*, vol. 4, pp. 93–100, 2002.

[41] G. J. Boggy and P. J. Woolf, "A mechanistic model of PCR for accurate quantification of quantitative PCR data," *PLoS ONE*, vol. 5, no. 8, p. e12355, 2010.

[42] T. F. Scientific, "Thermo Scientific PikoReal Real-Time PCR System - User manual," 2013.

[43] S. N. Peirson, J. N. Butler, and R. G. Foster, "Experimental validation of novel and conventional approaches to quantitative real-time PCR data analysis," *Nucleic Acids Research*, vol. 31, no. 14, p. e73, 2003.

[44] J. L. Gevertz, S. M. Dunn, and C. M. Roth, "Mathematical model of real-time PCR kinetics," *Biotechnology and Bioengineering*, vol. 92, no. 3, pp. 346–55, 2005.

[45] D. Klein, "Quantification using real-time PCR technology: Applications and limitations," *Trends in Molecular Medicine*, vol. 8, no. 6, pp. 257–60, 2002.

[46] A. I. Dragan, R. Pavlovic, J. B. McGivney, J. R. Casas-Finet, E. S. Bishop, R. J. Strouse, M. A. Schenerman, and C. D. Geddes, "SYBR Green I: Fluorescence properties and interaction with DNA," *Journal of Fluorescence*, vol. 22, pp. 1189–99, 2012.

[47] L. Jordan and R. Kurtz, "Optical design of CFX96TM real-time PCR detection system eliminates the requirement of a passive reference dye," *BioRad Tech Note*, no. Bulletin 6047, pp. 1–6, 2000.

[48] T. Lammers, S. Aime, W. E. Hennink, G. Storm, and F. Kiessling, "Theranostic nanomedicine," *Accounts of Chemical Research*, vol. 44, no. 10, pp. 1029–38, 2011.

[49] C. Tuerk and L. Gold, "Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase," *Science*, vol. 249, pp. 505–10, 1990.

[50] A. D. Ellington and J. W. Szostak, "In vitro selection of RNA molecules that bind specific ligands," *Nature*, vol. 346, pp. 818–22, 1990.

[51] T. Hermann and D. J. D. Patel, "Adaptive recognition by aptamers nucleic acid," *Science*, vol. 287, pp. 820–5, 2000.

[52] A. M. Kopylov and V. A. Spiridonova, "Combinatorial chemistry of nucleic acids: SELEX," *Molecular Biology*, vol. 34, no. 6, pp. 940–54, 2000.

[53] H.-M. Meng, H. Liu, H. Kuai, R. Peng, L. Mo, and X.-B. Zhang, "Aptamer-integrated DNA nanostructures for biosensing, bioimaging and cancer therapy," *Chem. Soc. Rev.*, vol. 45, pp. 2583–602, 2016.

[54] J. Charlton and D. Smith, "Estimation of SELEX pool size by measurement of DNA renaturation rates," *RNA*, vol. 5, pp. 1326–32, 1999.

[55] T. Schütze, B. Wilhelm, N. Greiner, H. Braun, F. Peter, M. Mörl, V. A. Erdmann, H. Lehrach, Z. Konthur, M. Menger, P. F. Arndt, and J. Glökler, "Probing the SELEX process with next-generation sequencing," *PLoS ONE*, vol. 6, no. 12, p. e29604, 2011.

[56] T. Schütze, P. F. Arndt, M. Menger, A. Wochner, M. Vingron, V. A. Erdmann, H. Lehrach, C. Kaps, and J. Glökler, "A calibrated diversity assay for nucleic acid libraries using DiStRO - A diversity standard of random oligonucleotides," *Nucleic Acids Research*, vol. 38, no. 4, p. e23, 2009.

[57] R. Stoltenburg, C. Reinemann, and B. Strehlitz, "SELEX - A (r)evolutionary method to generate high-affinity nucleic acid ligands," *Biomolecular Engineering*, vol. 24, pp. 381–403, 2007.

[58] D. Smith, G. P. Kirschenheuter, J. Charlton, D. M. Guidot, and J. E. Repine, "In vitro selection of RNA-based irreversible inhibitors of human neutrophil elastase," *Chemistry & Biology*, vol. 2, pp. 741–50, 1995.

[59] J. Charlton, G. P. Kirschenheuter, and D. Smith, "Highly potent irreversible inhibitors of neutrophil elastase generated by selection from a randomized DNA-valine phosphonate library," *Biochemistry*, vol. 36, pp. 3018–26, 1997.

[60] S. Ohuchi, "Cell-SELEX technology," *BioResearch Open Access*, vol. 1, no. 6, pp. 265–72, 2012.

[61] W. Tan, M. J. Donovan, and J. Jiang, "Aptamers from cell-based selection for bioanalytical applications," *Chemical Reviews*, vol. 113, pp. 2842–62, 2013.

[62] J. A. Bittker, B. V. Le, and D. R. Liu, "Nucleic acid evolution and minimization by nonhomologous random recombination," *Nature Biotechnology*, vol. 20, pp. 1024–9, 2002.

[63] G. Gu, T. Wang, Y. Yang, X. Xu, and J. Wang, "An improved SELEX-Seq strategy for characterizing DNA-binding specificity of transcription factor: NF-κB as an example," *PLoS ONE*, vol. 8, no. 10, p. e76109, 2013.

[64] L. A. Bawazer, A. M. Newman, Q. Gu, A. Ibish, M. Arcila, J. B. Cooper, F. C. Meldrum, and D. E. Morse, "Efficient selection of biomineralizing DNA aptamers using deep sequencing and population clustering," *ACS Nano*, vol. 8, no. 1, pp. 387–95, 2014.

[65] G. V. Kupakuwana, J. E. Crill, M. P. McPike, and P. N. Borer, "Acyclic identification of aptamers for human alpha-thrombin using over-represented libraries and deep sequencing," *PLoS ONE*, vol. 6, no. 5, p. e19395, 2011.

[66] N. Savory, K. Abe, K. Sode, and K. Ikebukuro, "Selection of DNA aptamer against prostate specific antigen using a genetic algorithm and application to sensing," *Biosensors and Bioelectronics*, vol. 26, pp. 1386–91, 2010.

[67] A. P. Drabovich, M. Berezovski, V. Okhonin, and S. N. Krylov, "Selection of smart aptamers by methods of kinetic capillary electrophoresis," *Analytical Chemistry*, vol. 78, pp. 3171–8, 2006.

[68] L. C. Bock, L. C. Griffin, J. a. Latham, E. H. Vermaas, and J. J. Toole, "Selection of single-stranded DNA molecules that bind and inhibit human thrombin," *Nature*, vol. 355, pp. 564–6, 1992.

[69] M. Sassanfar and J. W. Szostak, "An RNA motif that binds ATP," *Nature*, vol. 364, pp. 550–3, 1993.

[70] J. H. Niazi, S. J. Lee, Y. S. Kim, and M. B. Gu, "ssDNA aptamers that selectively bind oxytetracycline," *Bioorganic and Medicinal Chemistry*, vol. 16, pp. 1254–61, 2008.

[71] D. Musumeci and D. Montesarchio, "Polyvalent nucleic acid aptamers and modulation of their activity: A focus on the thrombin binding aptamer," *Pharmacology and Therapeutics*, vol. 136, pp. 202–15, 2012.

[72] G. Mayer, F. Rohrbach, B. Pötzsch, and J. Müller, "Aptamer-based modulation of blood coagulation," *Haemostaseologie*, vol. 31, pp. 258–63, 2011.

[73] D. Mann, C. Reinemann, R. Stoltenburg, and B. Strehlitz, "In vitro selection of DNA aptamers binding ethanolamine," *Biochemical and Biophysical Research Communications*, vol. 338, pp. 1928–34, 2005.

[74] J. Müller, B. Isermann, C. Dücker, M. Salehi, M. Meyer, M. Friedrich, T. Madhusudhan, J. Oldenburg, G. Mayer, and B. Pötzsch, "An exosite-specific ssDNA aptamer inhibits the anticoagulant functions of activated protein C and

enhances inhibition by protein C inhibitor," *Chemistry and Biology*, vol. 16, pp. 442–51, 2009.

[75] M. Jo, J.-Y. Ahn, J. Lee, S. Lee, S. W. Hong, J.-W. Yoo, J. Kang, P. Dua, D.-K. Lee, S. Hong, and S. Kim, "Development of single-stranded DNA aptamers for specific bisphenol A detection," *Oligonucleotides*, vol. 21, no. 2, pp. 85–91, 2011.

[76] A. Boltz, B. Piater, L. Toleikis, R. Guenther, H. Kolmar, and B. Hock, "Bi-specific aptamers mediating tumor cell lysis," *Journal of Biological Chemistry*, vol. 286, pp. 21896–905, 2011.

[77] J. A. DeGrasse, "A single-stranded DNA aptamer that selectively binds to staphylococcus aureus enterotoxin B," *PLoS ONE*, vol. 7, no. 3, p. e33410, 2012.

[78] C. Wang, G. Yang, Z. Luo, and H. Ding, "In vitro selection of high-affinity DNA aptamers for streptavidin," *Acta Biochimica et Biophysica Sinica*, vol. 41, no. 4, pp. 335–40, 2009.

[79] T. Bing, X. Yang, H. Mei, Z. Cao, and D. Shangguan, "Conservative secondary structure motif of streptavidin-binding aptamers generated by different laboratories," *Bioorganic and Medicinal Chemistry*, vol. 18, pp. 1798–805, 2010.

[80] D. a. Daniels, H. Chen, B. J. Hicke, K. M. Swiderek, and L. Gold, "A tenascin-C aptamer identified by tumor cell SELEX: Systematic evolution of ligands by exponential enrichment," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 26, pp. 15416–21, 2003.

[81] Z. Kiani, M. Shafiei, P. Rahimi-Moghaddam, A. A. Karkhane, and S. A. Ebrahimi, "In vitro selection and characterization of deoxyribonucleic acid aptamers for digoxin," *Analytica Chimica Acta*, vol. 748, pp. 67–72, 2012.

[82] D. M. Dupont, L. M. Andersen, K. A. Botkjaer, and P. A. Andreasen, "Nucleic acid aptamers against proteases," *Current Medicinal Chemistry*, vol. 18, no. 27, pp. 4139–51, 2011.

[83] A. D. Keefe, S. Pai, and A. Ellington, "Aptamers as therapeutics," *Nature Reviews Drug Discovery*, vol. 9, pp. 537–50, 2010.

[84] J. O. Mcnamara, D. Kolonias, F. Pastor, R. S. Mittler, L. Chen, P. H. Giangrande, B. Sullenger, and E. Gilboa, "Multivalent 4-1BB binding aptamers

costimulate CD8 + T cells and inhibit tumor growth in mice," *Journal of Clincal Investigaton*, vol. 118, no. 1, pp. 376–86, 2008.

[85] F. Pastor, M. M. Soldevilla, H. Villanueva, D. Kolonias, S. Inoges, A. L. de Cerio, R. Kandzia, V. Klimyuk, Y. Gleba, E. Gilboa, and M. Bendandi, "CD28 aptamers as powerful immune response modulators," *Molecular therapy - Nucleic acids*, vol. 2, p. e98, 2013.

[86] E. Gilboa, J. McNamara, and F. Pastor, "Use of oligonucleotide aptamer ligands to modulate the function of immune receptors," *Clinical Cancer Research*, vol. 19, pp. 1054–62, 2013.

[87] L. Gold, N. Janjic, T. Jarvis, D. Schneider, J. J. Walker, S. K. Wilcox, and D. Zichi, "Aptamers and the RNA world, past and present," *Cold Spring Harbor Perspectives in Biology*, vol. 4, p. a003582, 2012.

[88] A. S. Jørgensen, L. H. Hansen, B. Vester, and J. Wengel, "Improvement of a streptavidin-binding aptamer by LNA- and α-L-LNA-substitutions," *Bioorganic and Medicinal Chemistry Letters*, vol. 24, pp. 2273–7, 2014.

[89] R. M. Boomer, S. D. Lewis, J. M. Healy, M. Kurz, C. Wilson, and T. G. McCauley, "Conjugation to polyethylene glycol polymer promotes aptamer biodistribution to healthy and inflamed tissues," *Oligonucleotides*, vol. 15, pp. 183–95, 2005.

[90] J. G. Bruno, "A review of therapeutic aptamer conjugates with emphasis on new approaches," *Pharmaceuticals*, vol. 6, pp. 340–57, 2013.

[91] K. N. Meek, A. E. Rangel, and J. M. Heemstra, "Enhancing aptamer function and stability via in vitro selection using modified nucleic acids," *Methods*, vol. 106, pp. 29–36, 2016.

[92] J. H. Lee, M. V. Yigit, D. Mazumdar, and Y. Lu, "Molecular diagnostic and drug delivery agents based on aptamer-nanomaterial conjugates," *Advanced Drug Delivery Reviews*, vol. 62, pp. 592–605, 2010.

[93] S. Lee, Y. S. Kim, M. Jo, M. Jin, D. ki Lee, and S. Kim, "Chip-based detection of hepatitis C virus using RNA aptamers that specifically bind to HCV core antigen," *Biochemical and Biophysical Research Communications*, vol. 358, pp. 47–52, 2007.

[94] J.-Y. Ahn, E. Kim, J. Kang, and S. Kim, "A sol-gel-integrated protein array system for affinity analysis of aptamer-target protein interaction," *Nucleic Acid Therapeutics*, vol. 21, no. 3, pp. 179–83, 2011.

[95] A. R. Chandrasekaran, N. Anderson, M. Kizer, K. Halvorsen, and X. Wang, "Beyond the fold: Emerging biological applications of DNA origami," *Chem-BioChem*, vol. 17, pp. 1081–9, 2016.

[96] H. Hong, S. Goel, Y. Zhang, and W. Cai, "Molecular imaging with nucleic acid aptamers," *Current Medicinal Chemistry*, vol. 18, no. 27, pp. 4195–205, 2011.

[97] E. Mastronardi, A. Foster, X. Zhang, and M. C. DeRosa, "Smart materials based on DNA aptamers: Taking aptasensing to the next level," *Sensors*, vol. 14, pp. 3156–71, 2014.

[98] W. Tan, H. Wang, Y. Chen, X. Zhang, H. Zhu, C. Yang, R. Yang, and C. Liu, "Molecular aptamers for drug delivery," *Trends in Biotechnology*, vol. 29, no. 12, pp. 634–40, 2011.

[99] M. M. Hann, A. R. Leach, and G. Harper, "Molecular complexity and the impact on finding leads for drug discovery," *Journal of Chemical Information and Computer Sciences*, vol. 41, pp. 856–64, 2001.

[100] B. J. Davis and D. A. Erlanson, "Learning from our mistakes: The 'unknown knowns' in fragment screening," *Bioorganic and Medicinal Chemistry Letters*, vol. 23, pp. 2844–52, 2013.

[101] R. A. E. Carr, M. Congreve, C. W. Murray, D. C. Rees, R. A. E. Carr, C. W. Murray, and D. C. Rees, "Fragment-based lead discovery: Leads by design," *Drug Discovery Today*, vol. 10, no. 14, pp. 10–5, 2005.

[102] H. M. Geysen, F. Schoenen, D. Wagner, and R. Wagner, "Combinatorial compound libraries for drug discovery: An ongoing challenge," *Nature Reviews Drug Discovery*, vol. 2, pp. 222–30, 2003.

[103] O. Ramström and J.-M. Lehn, "Drug discovery by dynamic combinatorial libraries," *Nature Reviews Drug Discovery*, vol. 1, pp. 26–36, 2002.

[104] S. Melkko, C. E. Dumelin, J. Scheuermann, and D. Neri, "Lead discovery by DNA-encoded chemical libraries," *Drug Discovery Today*, vol. 12, no. 11-12, pp. 465–71, 2007.

[105] G. P. Smith, "Filamentous fusion phage: Novel expression vectors that display cloned antigens on the virion surface," *Science*, vol. 228, pp. 1315–7, 1985.

[106] G. Winter, A. D. Griffiths, R. E. Hawkins, and H. R. Hoogenboom, "Making antibodies by phage display technology," *Annual Review of Immunology*, vol. 12, pp. 433–55, 1994.

[107] O. H. Brekke and G. Å. Løset, "New technologies in the therapeutic antibody development," *Current Opinion in Pharmacology*, vol. 3, pp. 544–50, 2003.

[108] S. Brenner and R. A. Lerner, "Encoded combinatorial chemistry," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 89, pp. 5381–3, 1992.

[109] D. R. Halpin and P. B. Harbury, "DNA display I. Sequence-encoded routing of DNA populations," *PLoS Biology*, vol. 2, no. 7, pp. 1015–21, 2004.

[110] D. R. Halpin and P. B. Harbury, "DNA display II. Genetic manipulation of combinatorial chemistry libraries for small-molecule evolution," *PLoS Biology*, vol. 2, no. 7, pp. 1022–30, 2004.

[111] D. R. Halpin, J. A. Lee, S. J. Wrenn, and P. B. Harbury, "DNA display III. Solid-phase organic synthesis on unprotected DNA," *PLoS Biology*, vol. 2, no. 7, pp. 1031–8, 2004.

[112] F. Buller, Y. Zhang, J. Scheuermann, J. Schäfer, P. Bühlmann, and D. Neri, "Discovery of TNF Inhibitors from a DNA-Encoded Chemical Library based on Diels-Alder Cycloaddition," *Chemistry and Biology*, vol. 16, no. 10, pp. 1075–1086, 2009.

[113] F. Buller, M. Steiner, K. Frey, D. Mircsof, J. Scheuermann, M. Kalisch, P. Bühlmann, C. T. Supuran, and D. Neri, "Selection of carbonic anhydrase IX inhibitors from one million DNA-encoded compounds," *ACS Chemical Biology*, vol. 6, pp. 336–44, 2011.

[114] J. Scheuermann, C. E. Dumelin, S. Melkko, and D. Neri, "DNA-encoded chemical libraries," *Journal of Biotechnology*, vol. 126, pp. 568–81, 2006.

[115] L. Mannocci, M. Leimbacher, M. Wichert, J. Scheuermann, and D. Neri, "20 years of DNA-encoded chemical libraries," *Chemical Communications*, vol. 47, pp. 12747–53, 2011.

[116] G. Zimmermann and D. Neri, "DNA-encoded chemical libraries: Foundations and applications in lead discovery," *Drug Discovery Today*, 2016.

[117] S. Melkko, J. Scheuermann, C. E. Dumelin, and D. Neri, "Encoded self-assembling chemical libraries," *Nature Biotechnology*, vol. 22, no. 5, pp. 568–74, 2004.

[118] J. Scheuermann, C. E. Dumelin, S. Melkko, Y. Zhang, L. Mannocci, M. Jaggi, J. Sobek, and D. Neri, "DNA-encoded chemical libraries for the discovery of MMP-3 inhibitors," *Bioconjugate Chemistry*, vol. 19, pp. 778–85, 2008.

[119] J. Scheuermann and D. Neri, "DNA-encoded chemical libraries," *Current Opinion in Chemical Biology*, vol. 26, pp. 99–103, 2015.

[120] S. Melkko, C. E. Dumelin, J. Scheuermann, and D. Neri, "On the magnitude of the chelate effect for the recognition of proteins by pharmacophores scaffolded by self-assembling oligonucleotides," *Chemistry and Biology*, vol. 13, pp. 225–31, 2006.

[121] S. Melkko, Y. Zhang, C. E. Dumelin, J. Scheuermann, and D. Neri, "Isolation of high-affinity trypsin inhibitors from a DNA-encoded chemical library," *Angewandte Chemie International Edition*, vol. 46, pp. 4671–4, 2007.

[122] L. Mannocci, Y. Zhang, J. Scheuermann, M. Leimbacher, G. De Bellis, E. Rizzi, C. Dumelin, S. Melkko, and D. Neri, "High-throughput sequencing allows the identification of binding molecules isolated from DNA-encoded chemical libraries," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 46, pp. 17670–5, 2008.

[123] F. Buller, M. Steiner, J. Scheuermann, L. Mannocci, I. Nissen, M. Kohler, C. Beisel, and D. Neri, "High-throughput sequencing for the identification of binding molecules from DNA-encoded chemical libraries," *Bioorganic and Medicinal Chemistry Letters*, vol. 20, pp. 4188–92, 2010.

[124] L. Mannocci, S. Melkko, F. Buller, I. Molnàr, J.-P. G. Bianké, C. E. Dumelin, J. Scheuermann, and D. Neri, "Isolation of potent and specific trypsin inhibitors from a DNA-encoded chemical library," *Bioconjugate Chemistry*, vol. 21, pp. 1836–41, 2010.

[125] M. Wichert, N. Krall, W. Decurtins, R. M. Franzini, F. Pretto, P. Schneider, D. Neri, and J. Scheuermann, "Dual-display of small molecules enables the discovery of ligand pairs and facilitates affinity maturation," *Nature Chemistry*, vol. 7, pp. 241–9, 2015.

[126] C. E. Dumelin, J. Scheuermann, S. Melkko, and D. Neri, "Selection of streptavidin binders from a DNA-encoded chemical library," *Bioconjugate Chemistry*, vol. 17, pp. 366–70, 2006.

[127] C. E. Dumelin, S. Trüssel, F. Buller, E. Trachsel, F. Bootz, Y. Zhang, L. Mannocci, S. C. Beck, M. Drumea-Mirancea, M. W. Seeliger, C. Baltes, T. Müggler, F. Kranz, M. Rudin, S. Melkko, J. Scheuermann, and D. Neri, "A portable albumin binder from a DNA-encoded chemical library," *Angewandte Chemie - International Edition*, vol. 47, pp. 3196–201, 2008.

[128] M. Leimbacher, Y. Zhang, L. Mannocci, M. Stravs, T. Geppert, J. Scheuermann, G. Schneider, and D. Neri, "Discovery of small-molecule interleukin-2 inhibitors from a DNA-encoded chemical library," *Chemistry - A European Journal*, vol. 18, pp. 7729–37, 2012.

[129] R. K. Thalji, J. J. McAtee, S. Belyanskaya, M. Brandt, G. D. Brown, M. H. Costell, Y. Ding, J. W. Dodson, S. H. Eisennagel, R. E. Fries, J. W. Gross, M. R. Harpel, D. A. Holt, D. I. Israel, L. J. Jolivette, D. Krosky, H. Li, Q. Lu, T. Mandichak, T. Roethke, C. G. Schnackenberg, B. Schwartz, L. M. Shewchuk, W. Xie, D. J. Behm, S. A. Douglas, A. L. Shaw, and J. P. Marino, "Discovery of 1-(1,3,5-triazin-2-yl)piperidine-4-carboxamides as inhibitors of soluble epoxide hydrolase," *Bioorganic and Medicinal Chemistry Letters*, vol. 23, pp. 3584–8, 2013.

[130] A. G. Gilmartin, T. H. Faitg, M. Richter, A. Groy, M. a. Seefeld, M. G. Darcy, X. Peng, K. Federowicz, J. Yang, S.-Y. Zhang, E. Minthorn, J.-P. Jaworski, M. Schaber, S. Martens, D. E. McNulty, R. H. Sinnamon, H. Zhang, R. B. Kirkpatrick, N. Nevins, G. Cui, B. Pietrak, E. Diaz, A. Jones, M. Brandt, B. Schwartz, D. a. Heerding, and R. Kumar, "Allosteric Wip1 phosphatase inhibition through flap-subdomain interaction," *Nature chemical biology*, vol. 10, pp. 181–7, 2014.

[131] S. Barluenga, C. Zambaldo, H. A. Ioannidou, M. Ciobanu, P. Morieux, J. P. Daguer, and N. Winssinger, "Novel PTP1B inhibitors identified by DNA display of fragment pairs," *Bioorganic and Medicinal Chemistry Letters*, vol. 26, pp. 1080–5, 2016.

[132] H. Yang, P. F. Medeiros, K. Raha, P. Elkins, K. E. Lind, R. Lehr, N. D. Adams, J. L. Burgess, S. J. Schmidt, S. D. Knight, K. R. Auger, M. D. Schaber, G. J. Franklin, Y. Ding, J. L. Delorey, P. A. Centrella, S. Mataruse, S. R. Skinner, M. A. Clark, J. W. Cuozzo, and G. Evindar, "Discovery of a potent class of PI3Kα inhibitors with unique binding mode via encoded library technology (ELT)," *ACS Medicinal Chemistry Letters*, no. 5, pp. 531–6, 2015.

[133] P. A. Harris, B. W. King, D. Bandyopadhyay, S. B. Berger, N. Campobasso, C. A. Capriotti, J. A. Cox, L. Dare, X. Dong, J. N. Finger, L. C. Grady, S. J. Hoffman, J. U. Jeong, J. Kang, V. Kasparcova, A. S. Lakdawala, R. Lehr,

D. E. McNulty, R. Nagilla, M. T. Ouellette, C. S. Pao, A. R. Rendina, M. C. Schaeffer, J. D. Summerfield, B. A. Swift, R. D. Totoritis, P. Ward, A. Zhang, D. Zhang, R. W. Marquis, J. Bertin, and P. J. Gough, "DNA-encoded library screening identifies benzo[b][1,4]oxazepin-4-ones as highly potent and mono-selective receptor interacting protein 1 (RIP1) kinase inhibitors," *Journal of Medicinal Chemistry*, vol. 59, pp. 2163–78, 2016.

[134] A. R. Chandrasekaran, "Programmable DNA scaffolds for spatially-ordered protein assembly," *Nanoscale*, vol. 8, pp. 4436–46, 2016.

[135] J. H. Chen and N. C. Seeman, "Synthesis from DNA of a molecule with the connectivity of a cube," *Nature*, vol. 350, pp. 631–3, 1991.

[136] E. Winfree, F. Liu, L. A. Wenzler, and N. C. Seeman, "Design and self-assembly of two-dimensional DNA crystals," *Nature*, vol. 394, pp. 539–44, 1998.

[137] C. Mao, W. Sun, and N. C. Seeman, "Designed two-dimensional DNA holliday junction arrays visualized by atomic force microscopy," *Journal of the American Chemical Society*, vol. 121, pp. 5437–43, 1999.

[138] H. Yan, X. Zhang, Z. Shen, and N. C. Seeman, "A robust DNA mechanical device controlled by hybridization topology," *Nature*, vol. 415, pp. 62–5, 2002.

[139] J. Zheng, J. J. Birktoft, Y. Chen, T. Wang, R. Sha, P. E. Constantinou, S. L. Ginell, C. Mao, and N. C. Seeman, "From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal," *Nature*, vol. 461, pp. 74–7, 2009.

[140] S. Sellner, S. Kocabey, K. Nekolla, F. Krombach, T. Liedl, and M. Rehberg, "DNA nanotubes as intracellular delivery vehicles in vivo," *Biomaterials*, vol. 53, pp. 453–63, 2015.

[141] M. R. Jones, N. C. Seeman, and C. a. Mirkin, "Programmable materials and the nature of the DNA bond," *Science*, vol. 347, no. 6224, pp. 1–11, 2015.

[142] S. M. Douglas, I. Bachelet, and G. M. Church, "A logic-gated nanorobot for targeted transport of molecular payloads," *Science*, vol. 335, pp. 831–4, 2012.

[143] S. M. Douglas, H. Dietz, T. Liedl, B. Hogberg, F. Graf, and W. M. Shih, "Self-assembly of DNA into nanoscale three-dimensional shapes," *Nature*, vol. 21, pp. 414–8, 2009.

[144] M. Cieplak and W. Kutner, "Artificial biosensors: How can molecular imprinting mimic biorecognition?," *Trends in Biotechnology*, pp. 1–20, 2016.

[145] J. B. Lee, S. Peng, D. Yang, Y. H. Roh, H. Funabashi, N. Park, E. J. Rice, L. Chen, R. Long, M. Wu, and D. Luo, "A mechanical metamaterial made from a DNA hydrogel," *Nature Nanotechnology*, vol. 7, pp. 816–20, 2012.

[146] Y. Li, Y. D. Tseng, S. Y. Kwon, L. D'Espaux, J. S. Bunch, P. L. McEuen, and D. Luo, "Controlled assembly of dendrimer-like DNA," *Nature Materials*, vol. 3, pp. 38–42, 2004.

[147] S. Chatterjee, J. B. Lee, N. V. Valappil, D. Luo, and V. M. Menon, "Probing Y-shaped DNA structure with time-resolved FRET," *Nanoscale*, vol. 4, pp. 1568–71, 2012.

[148] S. H. Um, J. B. Lee, N. Park, S. Y. Kwon, C. C. Umbach, and D. Luo, "Enzyme-catalysed assembly of DNA hydrogel," *Nature Materials*, vol. 5, pp. 797–801, 2006.

[149] C. Buranachai, S. A. McKinney, and T. Ha, "Single molecule nanometronome," *Nano Letters*, vol. 6, no. 3, pp. 496–500, 2006.

[150] A. Kuzuya, R. Watanabe, Y. Yamanaka, T. Tamaki, M. Kaino, and Y. Ohya, "Nanomechanical DNA origami pH sensors," *Sensors*, vol. 14, pp. 19329–35, 2014.

[151] S. Roweis, E. Winfree, R. Burgoyne, N. V. Chelyapov, M. F. Goodman, P. W. K. Rothemund, and L. M. Adleman, "A sticker based model for DNA computation," *Journal of Computational Biology*, vol. 5, no. 4, pp. 615–29, 1998.

[152] L. M. Adleman, "Molecular computation of solutions to combinatorial problems," *Science*, vol. 266, no. 5187, pp. 1021–4, 1994.

[153] L. Qian and E. Winfree, "Scaling up digital circuit computation with DNA strand displacement cascades," *Science*, vol. 332, pp. 1196–201, 2011.

[154] L. Qian, E. Winfree, and J. Bruck, "Neural network computation with DNA strand displacement cascades," *Nature*, vol. 475, pp. 368–72, 2011.

[155] J. Markman, "DNA is the new data," *Forbes*, 2016.

[156] D. Meyer, "Why Microsoft just bought 10 million strands of DNA," *Fortune*, 2016.

[157] D. Gil, A. Ferrández, H. Mora-Mora, and J. Peral, "Internet of things: A review of surveys based on context aware intelligent services," *Sensors*, vol. 16, p. 1069, 2016.

[158] G. M. Church, Y. Gao, and S. Kosuri, "Next-generation digital information storage in DNA," *Science*, vol. 337, p. 1628, 2012.

[159] T. Palmer, "Build imprecise supercomputers," *Nature*, vol. 526, pp. 2–3, 2015.

[160] T. N. Palmer and M. O'Shea, "Solving difficult problems creatively: A role for energy optimised deterministic/stochastic hybrid computing," *Frontiers in Computational Neuroscience*, vol. 9, pp. 1–5, 2015.

[161] A. Extance, "How DNA could store all the world's data," *Nature*, vol. 537, pp. 22–4, 2016.

[162] M. Austin, "Species distribution models and ecological theory: A critical assessment and some possible new approaches," *Ecological Modelling*, vol. 200, pp. 1–19, 2007.

[163] D. N. Reznick and C. K. Ghalambor, "The population ecology of contemporary adaptations: What empirical studies reveal about the conditions that promote adaptive evolution," *Genetica*, vol. 112-113, pp. 183–98, 2001.

[164] H. T. Idriss and J. H. Naismith, "TNF-α and the TNF receptor superfamily: Structure-function relationship(s)," *Microscopy Research and Technique*, vol. 50, pp. 184–95, 2000.

[165] N. Vanlangenakker, M. J. M. Bertrand, P. Bogaert, P. Vandenabeele, and T. V. Berghe, "TNF-induced necroptosis in L929 cells is tightly regulated by multiple TNFR1 complex I and II members," *Cell Death and Disease*, vol. 2, p. e230, 2011.

[166] S. Fulda, "Therapeutic exploitation of necroptosis for cancer therapy," *Seminars in Cell and Developmental Biology*, vol. 35, pp. 51–6, 2014.

[167] J. R. Bradley, "The role of nuclear organization in cancer," *Journal of Pathology*, vol. 214, pp. 149–60, 2008.

[168] B. B. Aggarwal, S. C. Gupta, and J. H. Kim, "Historical perspectives on tumor necrosis factor and its superfamily: 25 years later, a golden journey," *Blood*, vol. 119, no. 3, pp. 651–665, 2012.

[169] D. Intiso, M. M. Zarrelli, G. Lagioia, F. D. Rienzo, C. Checchia De Ambrosio, P. Simone, P. Tonali, and R. P. Cioffi, "Tumor necrosis factor alpha serum levels and inflammatory response in acute ischemic stroke," *Neurological Sciences*, vol. 24, pp. 390–6, 2003.

[170] V. Michalaki, K. Syrigos, P. Charles, and J. Waxman, "Serum levels of IL-6 and TNF-α correlate with clinicopathological features and patient survival in patients with prostate cancer," *British Journal of Cancer*, vol. 90, pp. 2312–6, 2004.

[171] O. Arican, M. Aral, S. Sasmaz, and P. Ciragil, "Serum levels of TNF-α, IFN-γ, IL-6, IL-8, IL-12, IL-17, and IL-18 in patients with active psoriasis and correlation with disease severity," *Mediators of Inflammation*, vol. 2005, no. 5, pp. 273–9, 2005.

[172] T. Horiuchi, H. Mitoma, S. I. Harashima, H. Tsukamoto, and T. Shimoda, "Transmembrane TNF-α: Structure, function and interaction with anti-TNF agents," *Rheumatology*, vol. 49, pp. 1215–28, 2010.

[173] D. Aderka, H. Engelmann, Y. Maor, C. Brakebusch, and D. Wallach, "Stabilization of the bioactivity of tumor necrosis factor by its soluble receptors," *The Journal of Eperimental Medicine*, vol. 175, pp. 323–9, 1992.

[174] S. Nagata, R. Hanayama, and K. Kawane, "Autoimmunity and the clearance of dead cells," *Cell*, vol. 140, pp. 619–30, 2010.

[175] G. Hartmann, A. Krug, A. Eigler, J. Moeller, J. Murphy, R. Albrecht, and S. Endres, "Specific suppression of human tumor necrosis factor-α synthesis by antisense oligodeoxynucleotides," *Antisense and Nucleic Acid Drug Development*, vol. 6, pp. 291–9, 1996.

[176] G. Hartmann, A. Krug, K. Waller-Fontaine, and S. Endres, "Oligodeoxynucleotides enhance lipopolysaccharide-stimulated synthesis of tumor necrosis factor: Dependence on phosphorothioate modification and reversal by heparin," *Molecular Medicine*, vol. 2, no. 4, pp. 429–38, 1996.

[177] H. Pfister, T. Hennet, and T. W. Jungi, "Lipopolysaccharide synergizes with tumour necrosis factor-α in cytotoxicity assays," *Immunology*, vol. 77, pp. 473–6, 1992.

[178] E. P. Sampaio, E. N. Sarno, R. Galilly, Z. A. Cohn, and G. Kaplan, "Thalidomide selectively inhibits tumor necrosis factor α production by stimulated human monocytes," *The Journal of Experimental Medicine*, vol. 173, pp. 699–703, 1991.

[179] J. F. Cheng, A. Ishikawa, Y. Ono, T. Arrhenius, and A. Nadzan, "Novel chromene derivatives as TNF-α inhibitors," *Bioorganic and Medicinal Chemistry Letters*, vol. 13, pp. 3647–50, 2003.

[180] Y. Mukai, T. Nakamura, M. Yoshikawa, Y. Yoshioka, S.-i. Tsunoda, S. Nakagawa, Y. Yamagata, and Y. Tsutsumi, "Solution of the structure of the TNF-TNFR2 complex," *Science Signaling*, vol. 3, no. 148, p. ra83, 2010.

[181] F. Bazzoni and B. Beutler, "The tumor necrosis factor ligand and receptor families," *Seminars in Medicine of the Beth Israel Hospital, Boston*, vol. 334, no. 26, pp. 1717–25, 1996.

[182] M. J. Eck and S. R. Sprang, "The structure of tumor necrosis factor-α at 2.6 Å resolution. Implications for receptor binding," *Journal of Biological Chemistry*, vol. 264, no. 29, pp. 17595–605, 1989.

[183] R. Smith and C. Baglioni, "The active form of tumor necrosis factor is a trimer," *The Journal of Biological Chemistry*, vol. 262, no. 15, pp. 6951–4, 1987.

[184] A. Corti, G. Fassina, F. Marcucci, E. Barbanti, and G. Cassani, "Oligomeric tumour necrosis factor alpha slowly converts into inactive forms at bioactive levels," *The Biochemical Journal*, vol. 284, pp. 905–10, 1992.

[185] P. Wingfield, R. H. Pain, and S. Craig, "Tumour necrosis factor is a compact trimer," *FEBS Letters*, vol. 211, no. 2, pp. 179–84, 1987.

[186] C. Poiesi, A. Albertini, S. Ghielmi, G. Cassani, and A. Corti, "Kinetic analysis of TNF-a - Oligomer monomer transition by surface plasmon resonance and immunochemical methods," *Cytokine*, vol. 5, no. 6, pp. 539–45, 1993.

[187] R. Alzani, A. Corti, L. Grazioli, E. Cozzi, P. Ghezzi, and F. Marcucci, "Suramin induces deoligomierization of human tumor necrosis factor α," *The Journal of Biological Chemistry*, vol. 268, no. 17, pp. 12526–9, 1993.

[188] R. Alzani, E. Cozzi, A. Corti, M. Temponi, D. Trizio, M. Gigli, and V. Rizzo, "Mechanism of suramin-induced deoligomerization of tumor necrosis factor α," *Biochemistry*, vol. 34, pp. 6344–50, 1995.

[189] F. Mancini, C. M. Toro, M. Mabilia, M. Giannangeli, M. Pinza, and C. Milanese, "Inhibition of tumor necrosis factor-α (TNF-α)/TNF-α receptor binding by structural analogues of suramin," *Biochemical Pharmacology*, vol. 58, pp. 851–9, 1999.

[190] M. Kenig, V. Gaberc-Porekar, I. Fonda, and V. Menart, "Identification of the heparin-binding domain of TNF-alpha and its use for efficient TNF-alpha purification by heparin-sepharose affinity chromatography," *Journal of Chromatography B*, vol. 867, pp. 119–25, 2008.

[191] W. Jeong and H.-C. Shin, "Formation of an α-helix in human tumor necrosis factor-α by guanidine hydrochloride-induced unfolding," *Molecules and Cells*, vol. 17, no. 1, pp. 62–6, 2004.

[192] L. O. Narhi, J. S. Philo, T. Li, M. Zhang, B. Samal, and T. Arakawa, "Induction of alpha-helix in the β-sheet protein tumor necrosis factor-α: Acid-induced denaturation," *Biochemistry*, vol. 35, pp. 11454–60, 1996.

[193] B. L. Kagan, R. L. Baldwin, D. Munoz, and B. J. Wisnieski, "Formation of ion-permeable channels by tumor necrosis factor-α," *Science*, vol. 255, no. 5050, pp. 1427–30, 1992.

[194] R. L. Baldwin, M. L. Stolowitz, L. Hood, and B. J. Wisnieski, "Structural changes of tumor necrosis factor α associated with membrane insertion and channel formation," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, pp. 1021–6, 1996.

[195] M. A. Palladino, F. R. Bahjat, E. A. Theodorakis, and L. L. Moldawer, "Anti-TNF-α therapies: The next generation," *Nature Reviews Drug Discovery*, vol. 2, pp. 736–46, 2003.

[196] B. Scallon, A. Cai, N. Solowski, A. Rosenberg, X. Y. Song, D. Shealy, and C. Wagner, "Binding and functional comparison of two types of tumor necrosis factor antagonists," *The Journal of Pharmacology and Experimental Therapeutics*, vol. 301, no. 2, pp. 418–26, 2002.

[197] M. S. Kim, S. H. Lee, M. Y. Song, T. H. Yoo, B. K. Lee, and Y. S. Kim, "Comparative analyses of complex formation and binding sites between human tumor necrosis factor-alpha and its three antagonists elucidate their different neutralizing mechanisms," *Journal of Molecular Biology*, vol. 374, pp. 1374–88, 2007.

[198] J. Marusic, C. Podlipnik, S. Jevsevar, D. Kuzman, G. Vesnaver, and J. Lah, "Recognition of human tumor necrosis factor α (TNF-α) by therapeutic antibody fragment: Energetics and structural features," *Journal of Biological Chemistry*, vol. 287, no. 11, pp. 8613–20, 2012.

[199] J. A. Wells and C. L. McClendon, "Reaching for high-hanging fruit in drug discovery at protein-protein interfaces," *Nature*, vol. 450, pp. 1001–9, 2007.

[200] M. M. He, A. S. Smith, J. D. Oslob, W. M. Flanagan, A. C. Braisted, A. Whitty, M. T. Cancilla, J. Wang, A. A. Lugovskoy, J. C. Yoburn, A. D. Fung, G. Farrington, J. K. Elredge, E. S. Day, L. A. Cruz, T. G. Cachero, S. K. Miller, J. E. Friedman, I. C. Choong, and B. C. Cunningham, "Small-Molecule Inhibition of TNF-," *Science*, vol. 310, pp. 1022–5, 2005.

[201] R. W. Jackson, R. Gelinas, T. a. Baughman, T. Cox, J. J. Howbert, K. a. Kucera, J. a. Latham, F. Ramsdell, D. Singh, and I. S. Darwish, "Benzobicyclooctanes as novel inhibitors of TNF-α signaling," *Bioorganic & Medicinal Chemistry Letters*, vol. 12, pp. 1093–7, 2002.

[202] M. P. Clark, S. K. Laughlin, M. J. Laufersweiler, R. G. Bookland, T. A. Brugel, A. Golebiowski, M. P. Sabat, J. A. Townes, J. C. Vanrens, J. F. Djung, M. G. Natchus, B. De, L. C. Hsieh, S. C. Xu, R. L. Walter, M. J. Mekel, S. A. Heitmeyer, K. K. Brown, K. Juergens, Y. O. Taiwo, and M. J. Janusz, "Development of orally bioavailable bicyclic pyrazolones as inhibitors of tumor necrosis factor-α production," *Society*, vol. 47, pp. 2724–7, 2004.

[203] A. Hoffmann, M. Kovermann, H. Lilie, M. Fiedler, J. Balbach, R. Rudolph, and S. Pfeifer, "New binding mode to TNF-alpha revealed by ubiquitin-based artificial binding protein," *PLoS ONE*, vol. 7, no. 2, p. e31298, 2012.

[204] J. Brunetti, B. Lelli, S. Scali, C. Falciani, L. Bracci, and A. Pini, "A novel phage-library-selected peptide inhibits human TNF-α binding to its receptors," *Molecules*, vol. 19, pp. 7255–68, 2014.

[205] W. Lian, P. Upadhyaya, C. A. Rhodes, Y. Liu, and D. Pei, "Screening bicyclic peptide libraries for protein-protein interaction inhibitors: Discovery of a tumor necrosis factor-α antagonist," *Journal of the American Chemical Society*, vol. 135, no. 32, pp. 11990–5, 2013.

[206] J. Zhang, L. Zheng, A. Zhao, B. Gao, N. L. Liu, F. Wang, J. Dong, Z. T. Xin, N. S. Shao, H. X. Wang, and Y. N. Xue, "Identification of anti-TNF-α peptides with consensus sequence," *Biochemical and Biophysical Research Communications*, vol. 310, pp. 1181–7, 2003.

[207] E. W. Orava, N. Jarvik, Y. L. Shek, S. S. Sidhu, and J. Gariépy, "A short DNA aptamer that recognizes TNF-α and blocks its activity in vitro," *ACS Chemical Biology*, vol. 8, pp. 170–8, 2013.

[208] L. Hamon, D. Pastré, P. Dupaigne, C. Le Breton, E. Le Cam, and O. Piétrement, "High-resolution AFM imaging of single-stranded DNA-binding (SSB) protein - DNA complexes," *Nucleic Acids Research*, vol. 35, no. 8, pp. 1–7, 2007.

[209] D. Pastré, L. Hamon, F. Landousy, I. Sorel, M. O. David, A. Zozime, E. Le Cam, and O. Piétrement, "Anionic polyelectrolyte adsorption on mica mediated by multivalent cations: A solution to DNA imaging by atomic force microscopy under high ionic strengths," *Langmuir*, vol. 22, pp. 6651–60, 2006.

[210] R. A. Stull, G. Zon, and F. C. Szoka, "Single-stranded phosphodiester and phosphorothioate oligonucleotides bind actinomycin D and interfere with tumor necrosis factor-induced lysis in the L929 cytotoxicity assay," *Antisense Research and Development*, vol. 3, pp. 295–300, 1993.

[211] D. T. Humphreys and M. R. Wilson, "Modes of L929 cell death induced by TNF-α and other cytotoxic agents," *Cytokine*, vol. 11, no. 10, pp. 773–82, 1999.

[212] D. Pang, A. R. Thierry, and A. Dritschilo, "DNA studies using atomic force microscopy: Capabilities for measurement of short DNA fragments," *Frontiers in Molecular Biosciences*, vol. 2, pp. 1–7, 2015.

[213] S. N. Tamkovich, O. E. Bryzgunova, E. Y. Rykova, V. I. Permyakova, V. V. Vlassov, and P. P. Laktionov, "Circulating nucleic acids in blood of healthy male and female donors," *Clinical Chemistry*, vol. 51, no. 7, pp. 1315–7, 2005.