universität wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

„Chemical and technical improvement of light-directed *in situ* microarray technology“

verfasst von / submitted by

Mag. rer. nat. Matej Sack

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doktor der Naturwissenschaften (Dr. rer. nat.)

Wien, Januar 2016 / Vienna, January 2016

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 796 605 419
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Chemie
Betreut von / Supervisor:	Mark Manuel Somoza, BSc MSc PhD

Table of Contents

Acknowledgements

First of all I would like to thank Mark Somoza for the opportunity to do this PhD thesis on a relatively uncommon topic here in Austria, gaining new experience and insight as well as for all the fruitful discussions and the necessary help and time in order to achieve this goal.

Further thanks to Univ.-Prof. Dr. Christopher Gerner and Priv. Doz. Mag. Dr. Martin Bilban for taking the time to assess this work.

To my family, especially my mother, for all the needed support.

To my friends Daniel, Robert, Patrick, Mathias, Motz and Jory. For being around.

Finally to Anna, for the love and support and simply being herself.

\m/

List of figures

Abbreviations

CPG	controlled pore glass
DMD	digital micromirror device
DMT	dimethoxytrityl
DNA	deoxyribonucleic acid
dsDNA	double-stranded deoxyribonucleic acid
GWAS	genome-wide association studies
MAS	maskless array synthesis
NGS	next-generation sequencing
NPPOC	2-(2-nitrophenyl)-propoxycarbonyl
PTFE	polytetrafluoroethylene
RNA	ribonucleic acid
SNP	single-nucleotide polymorphism
SPh-NPPOC	thiophenyl-2-(2-nitrophenyl)-propoxycarbonyl
ssDNA	single-stranded deoxyribonucleic acid

Chapter 1

1 Introduction

1.1 DNA

Deoxyribonucleic acid (DNA) is a macromolecule that is accountable for storing genetic information, which in turn is crucial for the functioning development of any living organism (including viruses).

DNA consists of three chemical groups − a nitrogen base, a sugar and a phosphate − which are tethered to each other by a covalent bond, base to sugar via a glycosidic bond, sugar to phosphate via an ester bond and all is summarized as a nucleotide (Figure 1.1 a). The bases are divided into pyrimidines − a single ring structure with two nitrogen atoms − that comprise cytosine (C) and thymine (T) - uridine (U) in case of ribonucleic acid (RNA) - and purines − a two ring structure, consisting of pyrimidine fused to imidazole − that in turn comprise adenine (A) and guanine (G). The way of numbering the carbon atoms of the bases is also shown in Figure 1.1 b. Pyrimidines are anchored by the nitrogen atom 1 to the sugar, purines by the nitrogen atom 9. In contrast the sugar's carbon atoms are numbered in a clockwise manner and read "prime" to be distinguishable to the bases' carbon atoms. The sugar is called deoxyribose due to the absence of a hydroxyl group at the 2' position as in contrast to RNA where the 2'-OH group is present.

Repetitive units of nucleotides that are also connected covalently, form a polymer, one strand of DNA respectively. The right-handed double helix or double stranded structure of DNA (B-DNA) was discovered by James Watson and Francis Crick upon preliminary work of others, by analyzing X-ray diffraction data in 1953 [1]. The sequence of the nucleotides is joined together by phosphodiester bonds between the sugar group of one nucleotide and the phosphate group of an adjacent nucleotide. This sugar-phosphate backbone shapes the helical structure, is hydrophilic and negatively charged (Figure 1.1c).

Natural DNA molecules occur as a duplex in antiparallel manner in which the base pairs are joined by hydrogen bonds. In doing so A pairs with T forming two, C pairs with G forming three hydrogen bonds (Figure 1.1d). The strand and its antistrand, also known as the complementary sequence, hybridize in an exothermic process which can be reversed.

Relatively short and single-stranded DNA is referred to as oligonucleotides if it doesn't exceed a nucleotide length of about 100. Further organization leads to double-stranded units which are called the genes, which in turn become the chromosomes if elongated even further.

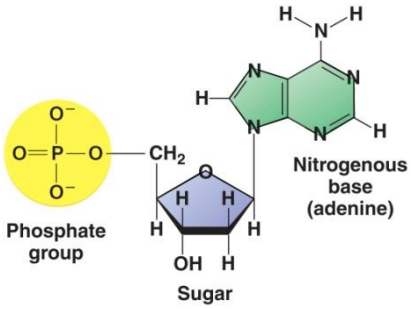


Figure 1.1a: DNA nucleotide

The base, in this case A, is connected by a glycosidic bond to the sugar which in turn is connected by an ester bond to the phosphate

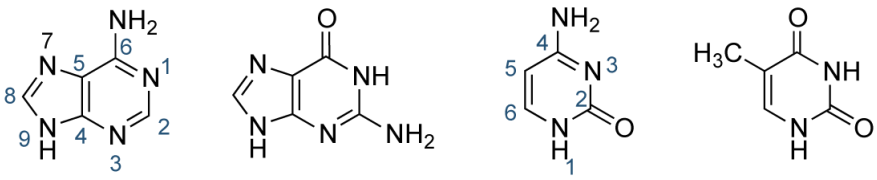


Figure 1.1b: DNA bases

Purines: A and G are tethered by the nitrogen atom 9 to the sugar

Pyrimidines: C and T (U) are anchored by the nitrogen atom 1 to the sugar

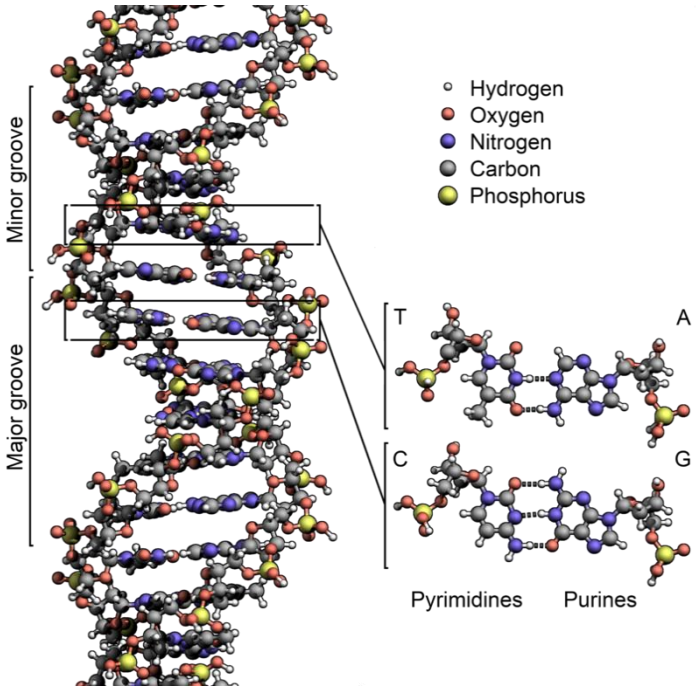


Figure 1.1c: Double helical structure of B-DNA

The asymmetrical turn of the B-DNA creates a minor and a major groove

(from wikipedia by Richard Wheeler "Zephyris", Licence: CC BY-SA 3.0, November 2015)

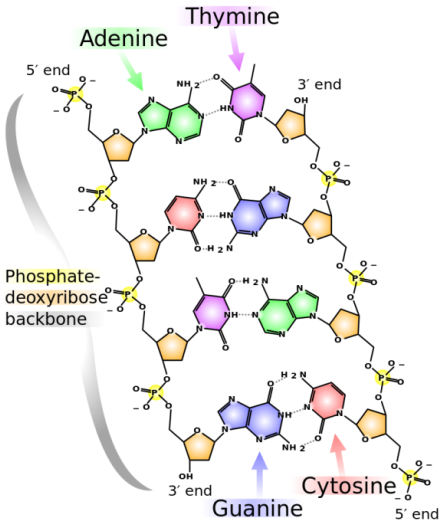


Figure 1.1d: Watson-Crick base pairing

A pairs with T via two hydrogen bonds and C pairs with G via three hydrogen bonds

(from wikipedia by Madeleine Price Ball, Licence: CC BY-SA 3.0, October 2015)

1.2 Phosphoramidites

Upon the preliminary work of Letsinger and coworkers who introduced an innovative approach of the phosphite coupling in oligonucleotide synthesis [2], Beaucage and Caruthers further developed new nucleoside phosphites. This new class of suitably protected phosphites were described to be activated by mild acid treatment, the need for a bifunctional phosphitylating agent, mild oxidation using iodine in tetrahydrofuran for generating the natural internucleotide bond and a second protected nucleoside [3]. The use of phosphoramidite chemistry makes it possible to synthesize oligonucleotides in any desired sequence out of the required phosphoramidites [4]. The initial dimethoxytrityl (DMT) -phoshoramidites apply to a pure chemical processing and consist of a nucleotide and protecting groups (Figure 1.2 a). All bases except T are protected in order ensure no unwanted reactions along the strand during synthesis and are removed after synthesis in a deprotection step. *N*,*N*-diisopropylamino on the 3'-phosphate is crucial for the activation cycle and it is removed during synthesis as in contrast to β-cyanoethyl, also on the 3'-phosphate, which is removed after synthesis. 5'-O-DMT is also removed during synthesis and along with *N*,*N*-diisopropylamino essential for the chain elongation mechanism. Figure 1.2 b depicts a DMT-driven synthesis cycle onto controlled pore glass (CPG) [5]. An entire cycle consists of four steps:

deprotection (detritylation) → coupling → capping → oxidation

Every cycle adds one DMT-phosphoramidite to the growing oligonucleotide chain. The deprotection step removes the DMT group, leaving the 5'-side with a free OH group onto which the next DMT-phosphoramidite can couple. The capping step prevents further growth by acetylation and therefore inactivating an oligonucleotide chain where the coupling step has failed, excluding it from further cycles. In doing so the error of resulting wrong sequences is decreased. The oxidation step however stabilizes the newly formed and yet unstable inter-nucleotide phosphite bond which took place during the coupling step from a trivalent to a pentavalent linkage (phosphite → phosphate).

The next step led to a replacement of the 5'-O-DMT - only chemical driven synthesis - with a light sensitive or photolabile protecting group on the 5'-side which is removed

via photodeprotection during synthesis. Upon absorption of photons at 365nm the photolabile groups like ((α -methyl-2-nitropiperonyl)-oxy]-carbonyl (MeNPOC) [6] or (2-(2-nitrophenyl)-propoxycarbonyl (NPPOC) come off [7] along with exposure solvent.

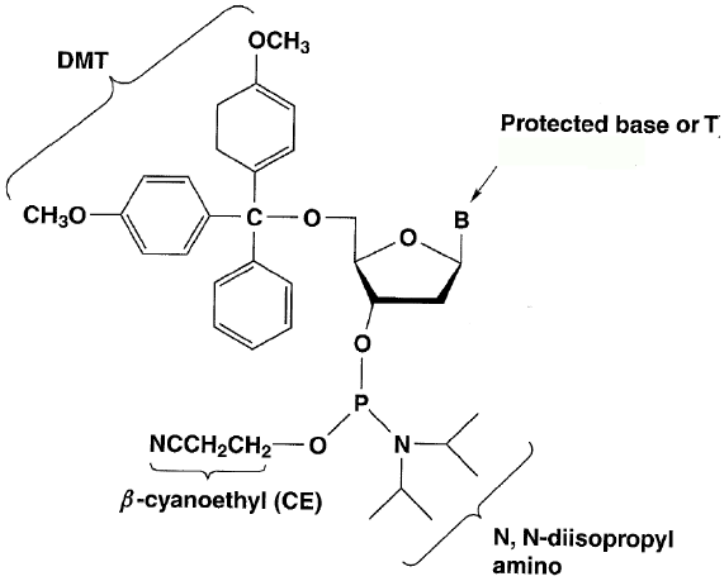


Figure 1.2 a: DMT-phosphoramidite

All bases except T are protected. *N,N*-diisopropylamino and the DMT functional group are removed during synthesis. *β*-cyanoethyl as well as the base protecting groups are removed after synthesis during the deprotection step

[modified after M. Schena, Microarray analysis, 2003]

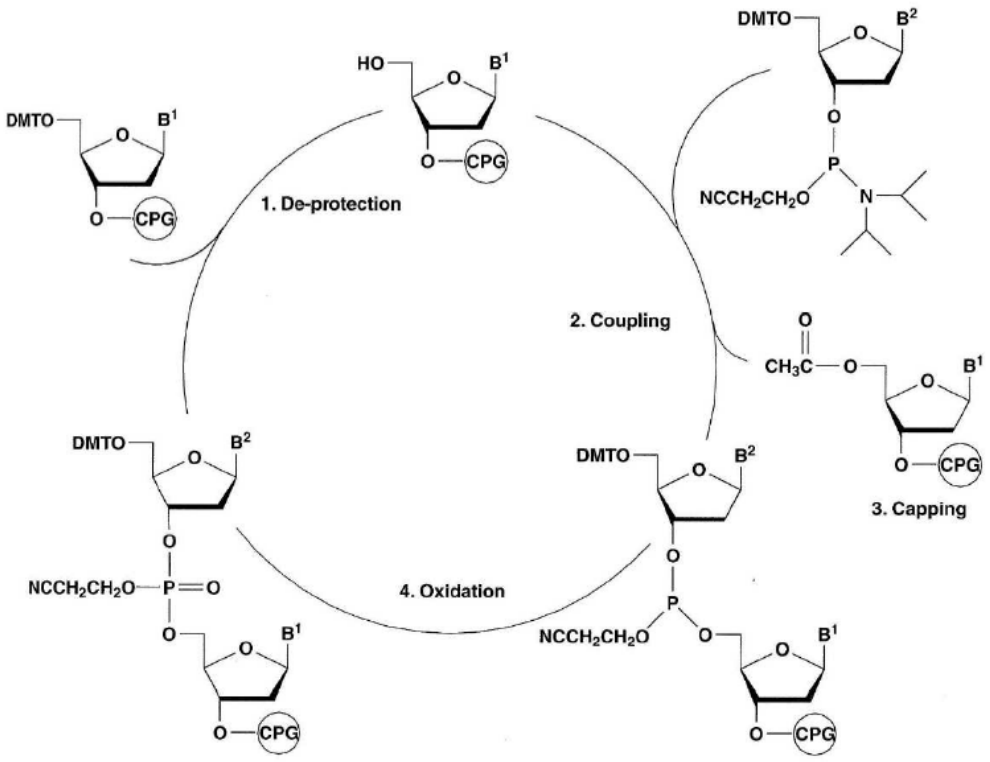


Figure 1.2 b: DMT dependent synthesis on CPG

An entire cycle (coupling to coupling) consists of repetitive deprotection (detritylation), coupling, capping and oxidation on controlled pore glass (CPG) until the desired sequence is reached

[M. Schena, Microarray analysis, 2003]

1.3 Microarray

An oligonucleotide microarray is a collection of DNA spots with well-defined positions that are anchored to a solid surface [8]. The utilized surface is a microscope glass slide from Schott AG which has to be functionalized prior to use and to create a linkage - a free hydroxyl group - between the glass slide and the future oligonucleotide.

Microarray technology evolved from Southern blotting. Pre-existing DNA microarrays, the so called controlled pore glass (CPG) technology was present back then. CPGs are porous glass beads that allow only one sequence to be synthesized and therefore not suitable for experiments where many thousands of sequences are required simultaneously for data extraction [9].

The underlying principle of the enormous efficiency of microarrays is its high selectivity due of hybridization and reliable fluorescent dye labeling.
The synthetic mounted DNA on the surface is then hybridized with its fluorescently-labeled complementary sequence and scanned with a microarray scanner or a fluorescence microscope. The resulting fluorescence is dependent on the yield of the hybridization. Less hybridization due to many different factors leads to a lower signal and vice versa.

1.4 Monohydroxysilane linker

Monohydroxysilane linkers are used as a relatively stable transition between the glass substrate and the growing oligonucleotide chains under acidic and neutral conditions [10]. Figure 1.4 shows the structure of the monohydroxysilane linker.

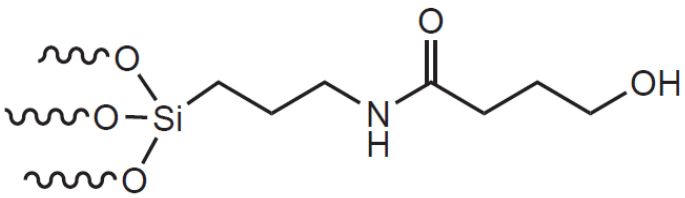


Figure 1.4: Monohydroxysilane linker

The linker's terminal OH group is attached to the 3' position of the oligonucleotide. The work of McGall in 1997 provided the method upon which the glass slides are prepared prior to the *in situ* synthesis [11].
Since NPPOC-dT has the best coupling efficiency and is the longest durable of all four (DNA)-phosphoramidites it is used to increase the distance between the glass substrate and the actual sequence. A linker-length of dT_5 to dT_{10} already helps to prevent the dye molecule to get stuck planarly to the surface, resulting in better fluorescence.

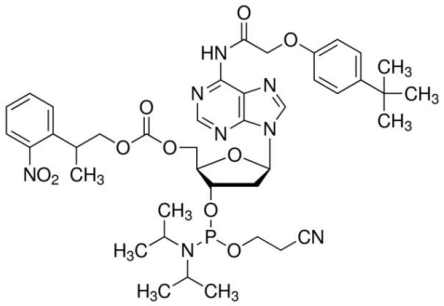
1.5 Light-directed *in situ* synthesis

Light-directed *in situ* microarray synthesis emerged from the photolithographic technology from the semiconductor industry which was combined with photolabile phosphoramidites chemistry that involved photodeprotection of the 5'-OH protection group. The main reason was to create high-density microarrays and led to the foundation of Affymetrix Inc. in 1991 [12], [13]. In contrast, to circumvent the time-intensive synthesis of many photolithographic masks that were needed for each microarray design as applied by Affymetrix, the foundation of NimbleGen in 1999 avoided those steps by introducing maskless array synthesis (MAS) whose development was triggered by Franco Cerrina and his group at UW-Madison in 1997 [14]. A digital micromirror device (DMD) is used along with the MAS technology to conduct the UV-light to the array's synthesis surface where the desired oligonucleotide sequence depends on the selective photodeprotection of the 5' protecting group of the previous phosphoramidite on the microarray. An entire light-directed phosphoramidite cycle is illustrated in Figure 1.6 b [22]. The resolution of the DMD is 1024x768 micromirrors which allow up to 786432 different sequences to be synthesized onto one single array simultaneously. More recent and higher definition DMDs have a resolution of 2560x1600, more than 4 million pixels.

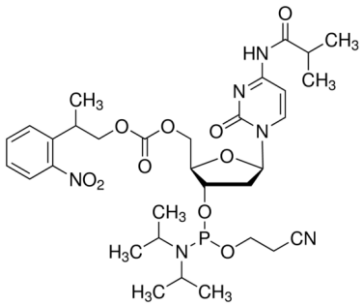
Light-directed *in situ* microarray synthesis is mostly common for DNA microarrays, but more recent developments extended its usage also for RNA microarrays [15], [16], peptide microarray synthesis [17] and has a huge impact in the field of gene expression research especially in gene expression profiling [18], [19]. Investigations that led to further optimization made by Sack et al., 2015 helped to reduce the synthesis time of an gene expression microarray from about 8 hours initially to about 1.5 hours without reduction of quality and therefore hugely decreasing the amount of time, amount of the reagents and after all a big drop in cost.

1.6 NPPOC-phosphoramidite chemistry

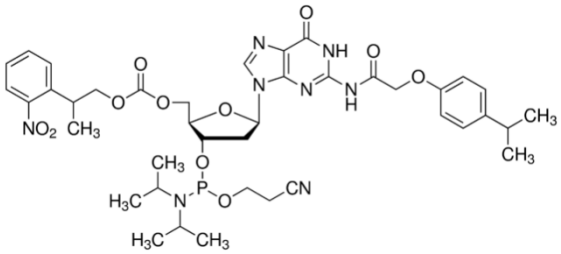
NPPOC-phosphoramidites can be divided into two main chemical components, the regular DNA/RNA molecules and chemical protecting groups. Protecting compounds are crucial to maintain the high coupling efficiency and prevent unwanted side reactions during synthesis. Figure 1.6 a illustrates the NPPOC-phosphoramidites and its protecting groups.



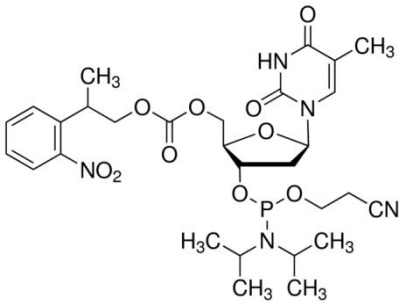
NPPOC-dA (tac)



NPPOC-dC (isobutyryl)



NPPOC-dG (ipac)



NPPOC-dT (no base protecting group)

Figure 1.6 a: Photolabile NPPOC-phosphoramidites
(dA, dC, dG and dT with base protecting group in brackets)

Purine and pyrimidine bases have to be protected in order to inhibit reactions of primary amines that may occur during synthesis. [20]. The photolabile NPPOC protecting group on the 5´ hydroxyl prevents any unintentional coupling of an incoming phosphoramidite until it's removed upon absorption of UV photons at 365nm [21].

All base protecting groups and the β-cyanoethyl group are removed after synthesis in a separate workflow by EDA/EtOH 1:1 (v:v).

During the work of this thesis attempts were made which showed that SPh-phosphoramidites work much more efficiently – less necessary energy for photodeprotection and faster times – data shown in Sack et al., 2015.

After the surface is being functionalized with the monohydroxysilane linker, the first phosphoramidite coupling takes place on the entire surface within the synthesis area. This provides an available 5´-OH after photodeptrotection for the next coupling which can be already driven by virtual masks. Every cyclic process, coupling → (capping) → oxidation → photodeprotection, ends in an addition of one oligonucleotide and therefore increasing the length. This cyclic process can be maintained until the required oligonucleotides pattern/length is achieved all over the microarray summarized and visualized in figure 1.6 b [22].

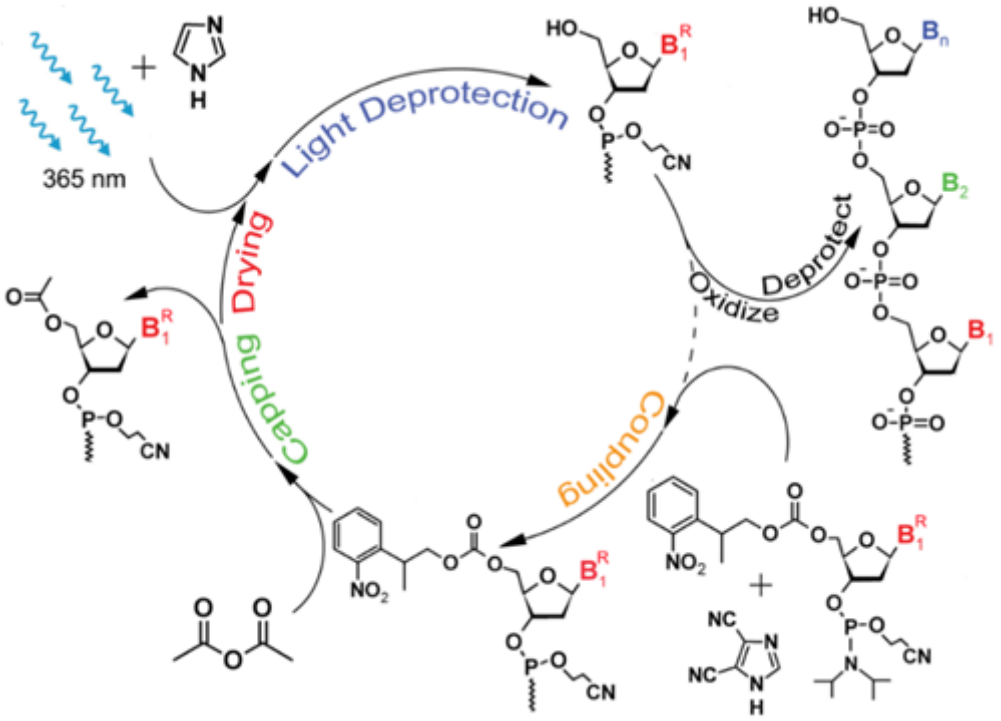


Figure 1.6 b: NPPOC-driven synthesis

This cycle depicts an entire coupling step (from coupling to coupling). Starting out with an incoming phosphoramidite whose *N*,*N*-diisopropylamino group is removed along with an activator solution in order to couple to the free 5´-hydroxyl group of the previous phosphoramidite. The capping step prevents further growth of an invalid sequence. Exposure (light deprotection) removes the 5´-NPPOC group along with UV-light at 365nm. The oxidation step turns the trivalent and instable phosphor to the pentavalent and stable phosphor (figure from Agbavwe et al., 20110) [22]

1.7 Maskless Array Synthesis (MAS)

The initial MAS was developed by F. Cerrina and coworkers in 1997. It is a maskless lithographic exposure apparatus designed for biological microarray syntheses, consisting of an optical system which conducts the UV light in a desired pattern to the microarray surface and a chemical delivering system [15]. NimbleGen Inc. made it commercially available afterwards.

A 350 W Hg arc lamp is used as the light source to deliver monochromatic UV light at 365 nm wavelength. It is adjusted to 80 mW/cm^2 in order to deliver 6 J/cm^2 for optimal NPPOC-photodeprotection. A Digital Micromirror Device (DMD) replaces the physical masks and forms an area of 1024 x 768 solely tiltable micromirrors. This leads to the essence of maskless lithography. The figure 1.7 a shows the schematics and figure 1.7 b and c an actual picture of the MAS system.

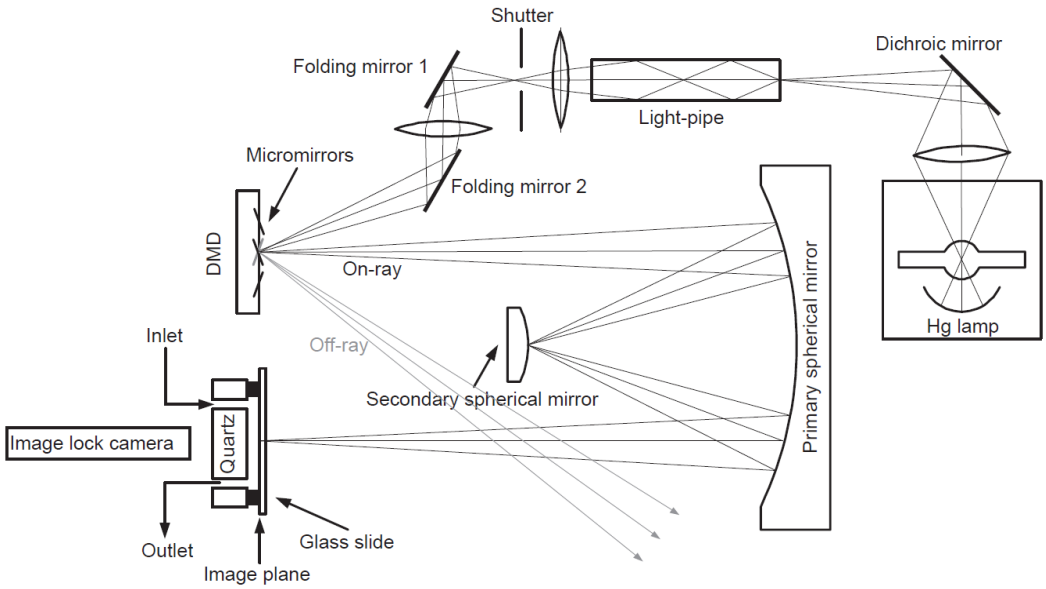


Figure 1.7 a: Schematics

The way of the light from the mercury lamp through the mirror complex up to the micromirror mounted in the cell assembly complex or conducted out of the optical system

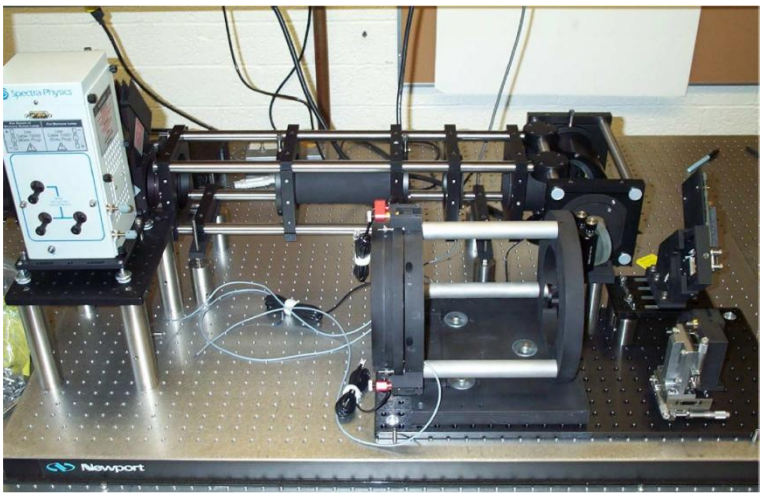


Figure 1.7b: Optical system

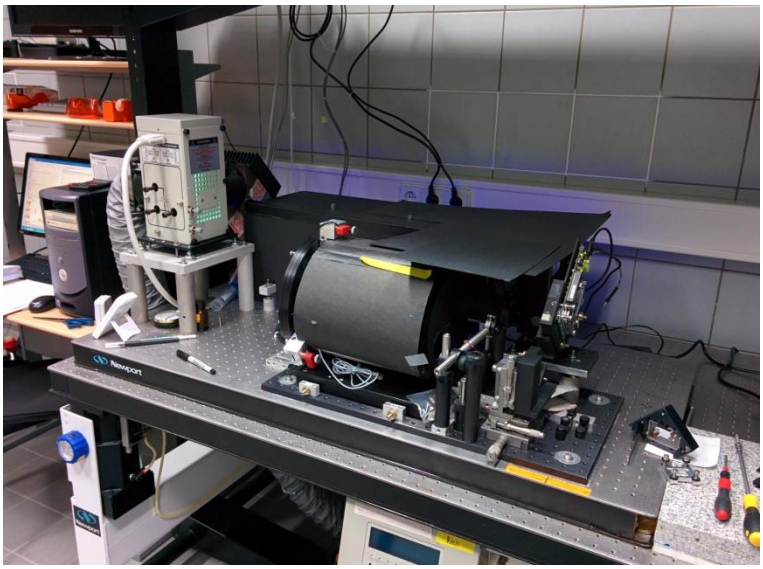


Figure 1.7c: Entire MAS system

1.8 Digital micromirror device (DMD)

The digital micromirror device (DMD) provided by Texas Instruments has an array density of 1024 x 768 pixels (Figure 1.8 a).

A micromirror is 13 µm x 13 µm and called a pixel, separated by a 1 µm gap to all adjacent micromirrors. The actual chip area - the area dimension onto which the microarray can be synthesized - is 14 mm x 10.5 mm. The micromirrors can be moved into three different positions. The ON position directs the light of the

programmed micromirror to the array. The OFF position in contrast conducts the light out of the optical system. This programmed ON/OFF position of the micromirrors results in the intended and actual chain growth and expected sequence outcome of the microarray. There also exists the FLOAT position where no electric voltage is involved and the mirrors are supposed to move freely resulting in preventing the hinge from additional/unnecessary mechanical stress (no synthesis running).

Figure 1.8 a: Digital Micromirror Device (DMD) by Texas Instruments
The picture doesn´t depict the 1024x768 version
(From wikipedia, November 2015, Licence: public domain)

1.9 Reagent delivery

A DNA synthesizer (PerSeptive Biosystems Expedite Model 8900), (Figure 1.9) delivers all required reagents along with the phosphoramidites to the reaction cell by pumping accompanied by constant He-flow. The reagent flow is controlled by a central computer along with the synchronization of the illumination shutter and the DMD display.

Figure 1.9: Expedite 8900 closed and opened

The right picture shows nine ports (transparent small vials) that can be used for different phosphoramidites, dyes etc. Below those the needed chemicals for the synthesis

1.10 Reaction cell

In order to set the reaction surface at the focal plane and to sweep the required reagents over it a reaction cell block is assembled.

However, the reaction cell was improved during this thesis. Synthesis time was halved by synthesizing two microarrays simultaneously without increasing the amount of required reagents and solvents. The microarrays are identical except being mirror images of each other (See Sack et al., 2013). For the reaction cell assembly complex see Figure 1 from this paper. The upper picture shows the exploded version of the cell assembly complex. The top microarray is separated by a 50µm PTFE (Polytetrafluoroethylene) gasket from the bottom microarray. Another gasket separates the bottom microarray from the quartz block. This second chamber can be filled with index-matching and absorbing fluids in order to minimize stray light. The bottom picture depicts the assembled cell assembly complex.

1.11 Errors during synthesis

There are many factors, starting from cell assembly, reagents preparation, chemical issues, computing communication problems and photolithography that can contribute to errors and therefore lower the sequence fidelity.
The DMT-driven phosphoramidite synthesis on a CPG-column however is mainly reduced to coupling issues resulting in capping and stopping the particular strand from further growth [23]. Light becomes an additional major error component in light-directed *in situ* synthesis causing base deletion if there is insufficient light present and unwanted base insertion due to unwanted light [24]. One aspect of light-related errors is addressed in the paper Sack et al., 2013.
It is important to understand the various sources of errors in order to guarantee the best possible outcome and maintain the high coupling efficiencies of phosphoramidites.

1.11.1 Flare, diffraction and edge scattering

Scattered light, stray light or flare are all the same things which can be summarized as light that appears unintentionally in the optical system leading to errors and further to a decreased synthesis yield. Flare emerges due to (dust)-particles in the air or on surfaces, multiple reflections in the glass and quartz, edges of the mounts and faulty mirror surfaces [25].
Diffraction affects all kind of waves if they encounter an obstacle or a slit in this case particularly if light hits edges in the optical imaging system resulting in interference and furthermore resulting in brighter or darker spots on the microarray. If the interference weakens the amount of energy for example that is needed for the photolabile NPPOC-protecting group to come off this may very well result in wrong sequence fidelity and decreases the amount of correct synthesis yield.
Although the MAS can collect scattered light at the edges by the reflective system, regions under dark exposure may receive light from its corresponding mirror edges, even if they are in OFF-position.

1.12 Applications of microarrays

The peptide-based array is based on the idea of the DNA microarrays which in turn originate from Southern blots. Geysen et al. triggered the start of combinatorial synthesis of peptides by reporting the peptide library synthesis made by multipin technology in 1984 [26]. Further development led to the usage of light-directed, spatially addressable parallel chemical synthesis method expedited by Fodor et al. in 1991 [12]. This novelty of peptide array was a success due to the excellent interaction between the desired peptide sequences and its fluorescently labeled antibody that had to be identified.

The human genome project contributed greatly to the identification of many thousands of new genes. Genes that play a crucial role in diseases [27]. Proteomics contribute to identification of gene functions which is promising having a vast impact to understand the molecular and cellular functions.

Protein microarrays have become very useful in the field of molecular biology and biochemistry. Also known as protein chips, parallel and minimized assay systems, containing small amounts of purified proteins in a high-density format [28]. Preparation of these arrays is done by either a standard contact spotter [29] or a noncontact microarrayer [30] by immobilizing proteins onto a surface, most likely microscope slides.

The most common is the antibody microarray amongst all of the analytical microarrays, hence the biggest difficulty is to synthesize antibodies which are capable of binding to the proteins of interest in a high-throughput and additionally with a high affinity and specificity. In order to circumvent the traditional method of preparing monoclonal antibodies that are difficult to make, there are new attempts like, systematic evolution of ligands by exponential enrichment (SELEX), messenger RNA (mRNA) display, ribosome display, phage antibody display and affibody display in order to accelerate the production of highly specific antibodies [31], [32], [33].

Genotyping microarrays are the most relevant field of work to identify genome related diseases in the so called microarray-based genome-wide association studies (GWAS). The use of high-throughput microarray technologies besides next-generation sequencing (NGS) unravels a fascinating overview of the entire genome

and its functions and consequences of genetic variation. While whole genomes microarrays can examine over 4 million markers per sample, NGS can interrogate the 3.2 billion bases of the human genome providing a comprehensive view of the genome. Other advantages of the genome-wide genotyping are the detection of single-nucleotide polymorphisms (SNP) beside other variations across the genome and the identification of any possible causal disease variants and therefore establishing further studies [34].

Carbohydrate microarrays can address the detection of posttranslational modification of proteins which can´t be detected by whole genome sequencing. Even though the human proteome has approximately 30000 genes available for its construction, posttranslational modifications, especially glycosylation - carbohydrate moieties are tethered by either *N*-glycosylation or *O*-glycosylation to certain sites of the protein - increase the number of functional proteins available in living organisms [35].

In order to establish reliable methods of detecting the arduous challenges of carbohydrate microarrays that have to address the structural diversity and functional complexity of carbohydrates different ways were established [36], [37], [38].

Chapter 2

2 Papers

2.1 Simultaneous light-directed synthesis of mirror-image microarrays in a photochemical reaction cell with flare suppression

Matej Sack, Nicole Kretschy, Barbara Rohm, Veronika Somoza and Mark M. Somoza

This paper was published in the Journal of Analytical Chemistry in August 2013.

Motivation

Highly complex microarrays of biopolymers, synthesized using photocleavable protecting groups, are capable of detecting about three-quarter of a million different biomolecules at the same time and therefore play an important role in the field of analysis. The goal of this paper is to further enhance the efficiency and high-throughput of light-directed *in situ* microarrays. This means in particular doubling the yield by synthesizing two identical - only mirror images of each other - microarrays onto two different substrates without increasing the amount of needed reagents. The advantages are reducing the costs per array and halving the synthesis time for otherwise two separate arrays which is useful if many microarrays of the same sort are needed. In addition a required comparison between two identical microarrays is much more reliable since each synthesis differs from another.

On average 4% of light is reflected on each glass surface due to the change in index of refraction between air and glass, resulting in unintentional photodeprotection and therefore introducing sequence errors. To address this issue a chamber behind the reaction cell can be additionally filled with a fluid matching the index of refraction of glass, and which is also light-absorbing, to minimize reflected light since unwanted light reflection, diffraction and scattered light are the main sources of decreased sequence fidelity.

analytical chemistry

Technical Note

pubs.acs.org/ac

Terms of Use

Simultaneous Light-Directed Synthesis of Mirror-Image Microarrays in a Photochemical Reaction Cell with Flare Suppression

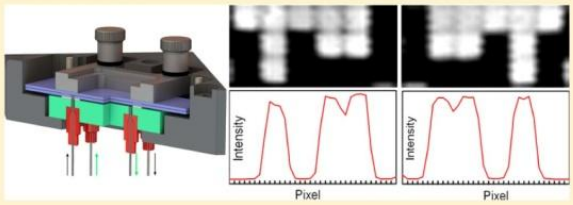
Matej Sack,[†] Nicole Kretschy,[†] Barbara Rohm,[‡,§] Veronika Somoza,[‡,§] and Mark M. Somoza*[,†]

[†]Institute of Inorganic Chemistry, University of Vienna, Währinger Strasse 42, A-1090 Vienna, Austria

[‡]Department of Nutritional and Physiological Chemistry, University of Vienna, Althanstraße 14 (UZA II), A-1090 Vienna, Austria

[§]Christian Doppler Laboratory for Bioactive Aroma Compounds, University of Vienna, Althanstraße 14, A-1090 Vienna, Austria

ABSTRACT: The use of photolabile protecting groups is a versatile and well-established means of synthesizing high complexity microarrays of biopolymers, such as nucleic acids and peptides, for high-throughput analysis. The synthesis takes place in a photochemical reaction cell which positions the microarray substrate at the focus of the optical system delivering the light and which can be connected to a fluidics system which delivers appropriate reagents to the surface in synchrony with the light exposure. Here we describe a novel photochemical reaction cell which allows for the simultaneous synthesis of microarrays on two substrates. The reaction cell positions both substrates within the limited depth-of-focus of the optical system while maintaining the necessary reagent flow conditions. The resulting microarrays are mirror images of each other but otherwise essentially identical. The new reaction cell doubles the throughput of microarray synthesis without increasing the consumption of reagents. In addition, a secondary flow chamber behind the reaction cell can be filled with an absorbent and index-matching fluid to eliminate reflections from light exiting the reaction cell assembly, greatly reducing unintended light exposure that reduces the sequence fidelity of the microarray probes.



Microarrays are versatile and widely used analytical tools with the capacity to simultaneously detect several hundred thousand to millions of different biomolecules simultaneously. Microarrays can be made by presynthesizing the probe molecule and spotting it on a surface using appropriate tethering chemistry, but modern microarrays are made with *in situ* methods in which the biomolecules are synthesized directly on the substrate from their monomer components, which allows for high probe densities, high uniformity, and high reproducibility.

Light-directed *in situ* synthesis of microarrays derives from the photolithographic technology used in the semiconductor industry in combination with combinatorial chemistry based on the selective removal of photolabile protecting groups. The technology was first commercialized by Affymetrix, which used the photolabile MeNPOC group on the 5′ end of DNA phosphoramidites to synthesize high-density DNA microarrays for genomics applications.[1] The synthesis technology was improved with the use of optical systems incorporating digital micromirror devices (DMD) to replace physical masks in the patterning of light on the microarray substrate, as well as by the use of the NPPOC photolabile group, which has significantly improved photodeprotection yield.[2−7] This maskless array synthesis (MAS) technology, originally used for DNA microarray synthesis has also been extended for the synthesis of RNA, aptamer,[8] and peptide microarrays.[9−13]

In situ microarray synthesis is robust and efficient in comparison with spotted synthesis; however, the total synthesis time and the consumption of solvents and reagents are still a significant economic constraint. In addition, the light-directed chemistry is sensitive to stray light in the system, which leads to unintended photodeprotection which degrades the sequence fidelity of the microarray probes.[7,14] Here we present an improved microfluidic photochemical reaction cell for use in light-directed synthesis that addresses both of these concerns. This reaction cell places two microarray substrates within the depth-of-focus plane of the optical system, so that two microarrays are synthesized simultaneously using the same reagents. The microarrays thus synthesized are mirror images of each other but otherwise essentially identical. The microarrays can be used independently but may have additional utility as matched pairs for experiments that would benefit from very close data comparisons; the quality of *in situ* synthesized microarrays, however, is very high and in most common applications, variations in quality between microarrays synthesized at different times are not experimentally relevant. In addition, the reaction cell assembly has a secondary chamber that can be filled with a light-absorbing and index-matching fluid to eliminate reflections that are a primary source of sequence error in light-directed synthesis.

Received: August 2, 2013
Accepted: August 22, 2013
Published: August 22, 2013

MATERIALS AND METHODS

Photochemical Reaction Cell Concept and Assembly. The reaction cell needs to position the two microarray substrates at the focal plane of the optical system. There is some tolerance to this positioning: the depth of focus of the imaging optics. The imaging optics are a 1:1 Offner relay system,[15,16] an off-axis conjugate system composed of two spherical concentric mirrors, primary and secondary. The system was designed with a numerical aperture (NA) of 0.08 to result in a resolving power of 2.7 μm. This resolving power is sufficient since it is significantly smaller than the size of individual mirrors of the digital micromirror device (DMD), 13 μm × 13 μm, separated by a 0.7 μm gap and is similar or better than those of most available microarray scanners. A low value of numerical aperture lowers the cost of the primary mirror but, more importantly, reduces the amount of scattered light originating from dust and imperfections in the optical system, which is proportional to NA^2. Unintended photodeprotection, from scattering, diffraction, and local flare, is the largest source of sequence error in light-directed microarray synthesis.[7] The depth of focus is intrinsically limited by diffraction to $<\sim\lambda/NA^2$, ~60 μm, but in practice, the positioning of the microarray substrates in the focal plane is somewhat less restricted due to limited resolution of microarray scanners. Therefore, the primary optical constraint in the simultaneous light-directed synthesis of microarray pairs is that the two substrates must be within ~60−100 μm of each other, depending on the scanner resolution.

A secondary constraint is imposed by reagent delivery. A larger reaction cell volume requires larger flow rates of solvents and reagents, the consumption of which scales with cell volume. Since our original reaction cell (for synthesizing microarrays on a single surface) had a depth of 70 μm and worked well with a standard oligonucleotide synthesizer (Expedite 8909), we took this value as a starting point. Thus, the reaction cell should consist of two standard microarray substrates (75 mm × 25 mm × 1 mm) separated by a uniform gap of ~70 μm. The microarray substrates form the entrance and exit windows for the ultraviolet light used in the synthesis. Reagents need to be introduced into this gap and to uniformly flow across the surface before exiting. We used these criteria to design and built the reaction cell shown in Figure 1. The reaction cell assembly consists of a black anodized aluminum support block, a quartz block, the two microarray substrates, two gaskets, and a clamping frame and screws to hold the parts together. Reagent delivery tubes attach to the underside of the quartz block and connect to the oligonucleotide synthesizer.

The support block forms the rigid structure for the assembly of the reaction cell and allows for the reaction cell to be precisely positioned in the focal plane. Three alignment points make contact with ball-tipped, high-precision adjustment screws (Newport AJS127-0.5H) in the optical system. After initial adjustment of the screws, the reaction cell assembly can be quickly and reproducibly positioned. The support blocks hold a quartz block. The quartz block has four 0.8 mm through-holes (two inlets, two outlets) that are countersunk on the back side to accommodate microfluidics ports. The microfluidics ports (IDEX 6-32 Coned NanoPort Assemblies) were turned on a lathe to reduce their diameter to 6.4 mm and attached within each countersunk hole with common cyanoacrylate adhesive. The front and back surfaces of the quartz block were machined to a surface parallelism error of <30 arc sec and

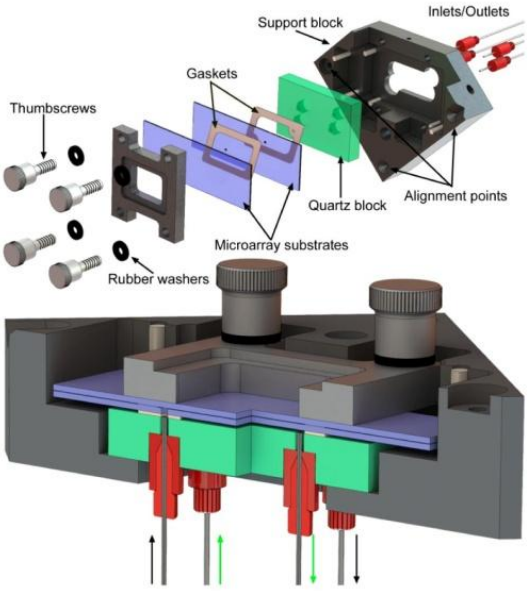


Figure 1. Exploded and section view of reaction cell assembly. The reaction cell is formed by two microarray substrates (75 mm × 25 mm × 1 mm) separated by a 50 μm PTFE gasket. Reagents enter and exit the cell via two 0.9 mm holes through the lower substrate. These holes are coupled to the inlets/outlets via an additional 250 μm thick FFKM gasket separating the lower substrate from the quartz block. The lower gasket forms a chamber that can be independently filled with a light-absorbing and index-matching fluid to reduce reflections from both quartz surfaces and from the back surface of the lower substrate. The thickness of the upper and lower gaskets in the section view have been exaggerated by a factor of 2 for visual clarity.

polished to an optical flatness of $\lambda/4$ (Mindrum Precision). During reaction cell assembly, the lower gasket is placed on the quartz surface. This gasket forms the lower chamber, which can be filled via two of the fluidics ports. Prior to microarray synthesis, this chamber can be filled with an index-matching and light absorbing fluid to prevent light reflections from light exiting the reaction chamber. In the legacy reaction cell design, an antireflective coating on the back surface of the quartz block can reduce the back reflection to a minimum of about 0.25% when new, but this value is typically larger, ~1%, due to the presence of dust, chemical films, and scratches. This 0.25−1% value is sufficient to make this unintended light exposure the largest source of error after diffraction, but unlike diffraction, the error is not confined primarily to the gaps between microarray features.[7] An alternative strategy to reduce back reflections is to fill the lower chamber with an index-matching fluid with dissolved chromophores which absorb the light exiting the reaction chamber and which either convert the light to heat or Stokes shift it beyond the absorption band of the light-labile group.

The lower gasket has two holes that align with two of holes in the quartz block. These holes couple the corresponding fluidics ports to the microarray synthesis cell. This gasket is made from 250 μm thick Chemraz 584 perfluoroelastomer (FFKM), cut to shape with a laser cutter (Spirit GX). The microarray synthesis cell is a chamber consisting of two glass substrates separated by a very thin gasket. This chamber is

accessed via two 1 mm holes, in the lower substrates, which align with the holes in the lower gasket.

The thickness of the upper gasket determines the depth of the photochemical reaction cell and therefore needs to be ~70 μm thick, chemically resistant and sufficiently elastic to form a seal for the duration of the synthesis, up to ~12 h for an array of 70mers. These requirements are quite exceptional and we were unable to find any references to such thin gaskets in the scientific or engineering literature. A perfluoroelastomer, such as Chemraz, would likely work, but the manufacturer is unable to make them thinner than 250 μm. We tried expanded polytetrafluoroethylene (ePTFE), which is commonly used in gasket applications due to its chemical resistance and ability to compress to form a seal, but found seepage through the gasket, presumably due to its porous nature. In the end we found that the common PTFE tape used for plumbing applications works well. This tape is made from unsintered PTFE and is therefore sufficiently compressible to form a seal but not porous. PTFE tape is made in many thicknesses and densities, which allowed for some experimentation. We initially used ~100 μm (120 μm uncompressed) PTFE with a density of ~1.4 g/cm^3 (Gasoila yellow tape), sintered PTFE has a density of about 2 g/cm^3, but found some loss of focus when microarrays were scanned at a resolution of 2.5 μm. Another problem with the 100 μm gap were indications that reagents were flowing in a channel through the center of the reaction cell rather than sweeping the whole surface. This was particularly apparent with the helium drying step, which was not capable of fully removing solvent from the corners of the reaction cell. Switching to thinner and lower density PTFE tape (Gasoila Industrial Strength SD, ~0.7 g/cm^3) gave a thickness of ~50 μm under compression. With this thickness, both of the paired arrays produce sharp scans with resolution limited only by the 2.5 μm pixel size of the scanner and both reagent and helium flow sweep uniformly across the entire surface of both substrates. The 50 μm PTFE gaskets are also formed with a laser cutter. Because of their thinness, they are too delicate to be reusable but can be made quickly and inexpensively.

Microarray Synthesis and Hybridization. Schott Nexterion Glass D slides functionalized with *N*-(3-triethoxysilylpropyl)-4-hydroxybutryamide (Gelest SIT8189.5). The arrays with holes were drilled with a 0.9 mm diamond bit and washed and rinsed in an ultrasonic bath prior to functionalization. The slides were loaded in a metal staining rack and completely covered with a 500 mL of a solution of 10 g of the silane in 95:5 (v/v) ethanol−water and 1 mL of acetic acid. The slides were gently agitated for 4 h and then rinsed twice for 20 min with gentle agitation in the same solution but without the silane. The slides were then drained and cured overnight in a preheated vacuum oven (120 °C). After cooling to room temperature, the slides were stored in a desiccator cabinet until use. Microarrays were synthesized directly on the slides using a maskless array synthesizer, which consists of an optical imaging system that used a digital micromirror device to deliver patterned ultraviolet light near 365 nm to the synthesis surface. Microarray layout and oligonucleotide sequences are determined by selective removal of the NPPOC photocleavable 5′-OH protecting group. Reagent delivery and light exposures are synchronized and controlled by a computer. The chemistry is similar to that used in conventional solid-phase oligonucleotide synthesis. The primary modification is the use of NPPOC phosphoramidites. Upon absorption of a UV photon, and in the presence of a weak organic base, e.g., 1% (m/v) imidazole in DMSO, the NPPOC group comes off, leaving a 5′-terminal hydroxyl which is able to react with an activated phosphoramidite in the next cycle. The DNA sequences on the microarrays in this project were synthesized with a light exposure dose of 4.5 J/cm^2, with coupling time of 40 s at monomer concentrations of 30 mM. After synthesis, the microarrays were deprotected in 1:1 (v/v) ethylenediamine in ethanol for 2 h at room temperature, washed twice with distilled water, dried with argon, and stored in a desiccator until hybridization.

Microarrays were hybridized in an adhesive chamber (SecureSeal SA200, Grace Biolabs) with a solution consisting of 0.3 pmol of 5′-Cy5-labeled probe and 200 μg of acetylated BSA in 400 μL of MES buffer (100 mM MES, 1 M NaCl, 20 mM EDTA, 0.01% Tween-20). After 2 h of rotation at 42 °C, the chamber was removed and the microarrays were vigorously washed in a 50 mL centrifuge tube with 30 mL of nonstringent wash buffer (SSPE; 0.9 M NaCl, 0.06 M phosphate, 6 mM EDTA, 0.01% Tween-20) for 2 min and then with stringent wash buffer (100 mM MES, 0.1 M NaCl, 0.01% Tween-20) for 1 min. The microarrays were then dipped for a few seconds in a final wash buffer (0.1× SSC) and then dried with a microarray centrifuge. Arrays were scanned with a Molecular Devices GenePix 4400A at a resolution of 2.5 μm.

Detection and Suppression of Reflected Light. To test the possibility of eliminating reflected light reaching the synthesis area, a small piece of radiochromic film (Far West Technology, FWT-60-20f), with a 2 mm punched hole, was placed in the reaction cell. A 9.5 mm metal disk with a 1 mm pinhole (Edmund Optics, 39730) was aligned over the hole in the film to serve as a physical mask. The entire reaction cell assembly was tilted by ~7° to move the reflection spot away from the mask hole. The lower chamber was filled with either DMSO (control) or UV absorbers dissolved in DMSO or dichloromethane. The UV absorbers (beta carotene, 9-methylanthracene, and riboflavin) were chosen for high extinction coefficients near 365 nm, high Stokes shift, low fluorescence quantum yield, and solubility in DMSO. The synthesis cell was exposed using all mirrors, with an exposure of 60 J/cm^2 (80 mW/cm^2 for 750 s).

■ RESULTS AND DISCUSSION

Synthesis of Mirror-Image Microarrays. Simultaneous synthesis of mirror-image microarrays in this microfluidic photochemical reaction chamber produces high-quality microarrays with little additional cost or effort beyond those of the single microarray synthesis of the legacy method. The primary concern with this method is that both arrays are in focus. To test the image quality of paired microarrays, we initially synthesized simple microarrays of 30mers (GTC ATC ATC ATG AAC CAC CCT GGT CTT TTT), hybridized them with labeled complementary oligonucleotides and scanned them at high resolution. The results of one such experiment is shown in Figure 2. The top row shows pixel-level close-ups from both of the arrays. Each white square corresponds to a microarray feature synthesized with a single DMD mirror. In both close-ups, the features are individually resolved, and the 0.7 μm gap between features are also clearly visible. The middle row shows plots of the scan image intensity along a horizontal line through the center of each of the pixel-level close-ups. The intensity drops by ~1000-fold between the center of hybridized pixels and unhybridized pixels, which is a typical signal/background for this type of microarray. The gap between immediately adjacent hybridized pixels is visible as a drop in intensity of

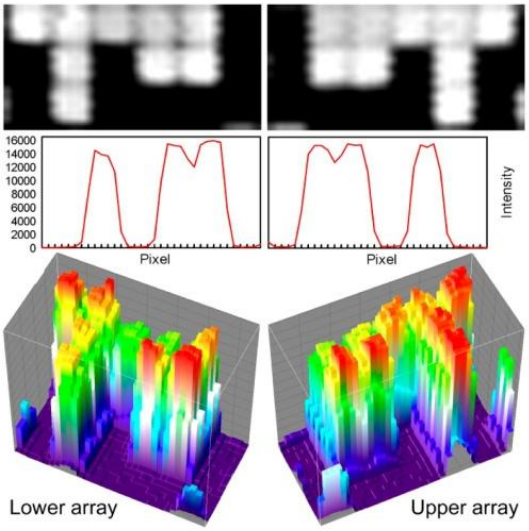


Figure 2. Scanned images and pixel intensities from two mirror-image microarrays synthesized simultaneously. Figures on the left are from the lower substrate (closest to quartz block in Figure 1), and those on the right are from the upper substrate. Top row: 3 × 6 array of features from the center of a 1024 × 768 array, scanned at 2.5 μm. Each features measures 13 μm × 13 μm and are separated by a 0.7 μm gap. Middle row: Intensity profiles of lines drawn horizontally through the close-ups above. Lower row: 3D surface intensity plots of the same close-ups.

about 20%. This interstitial intensity is due to the limited resolution of the scanner (2.5 μm), which leads to image pixels that derive most of their intensity from the adjacent bright microarray features. Diffraction also contributes significantly to intensity in gaps between microarray features, about 40% of the intensity of adjacent features when both features are exposed, and about 20% of the intensity of an adjacent feature when only one of the features is exposed.[7] The vertical sawtooth pattern probably originates from signal latency during rastering by the scanner. The microarrays are fully resolved within the constraints of scanner resolution and diffraction. The bottom row of Figure 2 shows 3-D surface intensity plots of the same close-ups. From the perspective of common microarray use, the each of the mirror image microarrays from the pair can be used as an individual microarray, but in some experimental contexts requiring close comparisons, matched pairs might be used to increase confidence in the comparison.

Blocking Reflections. The use of a light-absorbing fluid in the lower chamber resulted in the complete blockage of reflected light. Initial trials with 9-methylanthracence and riboflavin in DMSO were only partially successful due to incomplete absorption of violet light from the mercury lamp. Most of the photodeprotection of NPPOC results from the 365 nm line, but the mercury lines at 405 and 436 nm are also transmitted through the optical system and result in measurable deprotection. Beta carotene was able to completely absorb the incident light and prevent any reflection. Beta carotene is insufficiently soluble in DMSO but is highly soluble in dichloromethane,[17] which also has an index of refraction similar to that of glass. Figure 3 shows the effect of 5.5 mM beta carotene in dichloromethane. The control experiment (left film) has DMSO in the lower chamber and clearly shows the

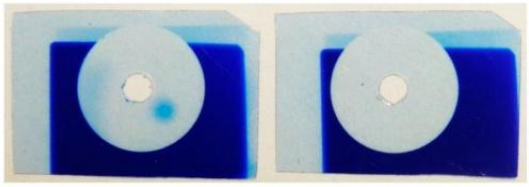


Figure 3. Visualization of light reflected into the synthesis chamber from the back surface of the quartz block and the complete suppression thereof using a light-absorbing fluid in the lower chamber. A 9.5 mm metal disk with a 1 mm diameter pinhole was used to mask radiochromic film in the synthesis chamber. The pinhole was aligned with a 2 mm hole in the film to allow the passage of light (60 J/cm^2), and the reaction cell assembly was tilted 7° to direct the reflection away from the hole. With the secondary chamber filled with a nonabsorbent fluid (left), there is a clear reflection to the lower right of the hole. When the secondary chamber is filled with a light-absorbing fluid, the reflection is completely suppressed (right).

reflection from the light transmitted through the 1 mm pinhole as a round exposed spot on the lower right-hand side. Another reflection is also apparent on the left side of the circle; this originates from transmission outside the pinhole disk that is not entirely absorbed by the radiochromic film. The film on the right shows that the beta carotene solution completely suppresses the reflections.

There are four principle sources of unintended photodeprotection: (1) global scattering, (2) edge scattering, (3) local flare (which includes reflections), and (4) diffraction.[7] Global scattering from imperfections and dust in the optical system is relatively small and results in a contrast ratio of better than 1/2500. Edge scattering originates primarily from the edges of the micromirrors and has a similar magnitude as global scattering. Diffraction is an intrinsic limitation of all imaging systems and results in partial exposure (~20%) of the area of the synthesis surface corresponding to the gaps between mirrors. Local scattering is primarily due to reflections of light exiting the reaction block but also includes scattering from bubbles in the exposure solvent. Bubbles can be eliminated by using appropriate fluidics protocols, primarily the use of helium as the blanket gas and adequate flushing of the reaction cell with exposure solvent before exposure. Reflection and diffraction remain alone as the largest sources of unintended exposure, each contributing approximately 1−2% of incident light. The use of an effective light absorber in the lower chamber, as demonstrated here, therefore reduces unintended exposure by approximately 50%. Diffraction remains as a large source of unintended exposure, but because the intensity is mostly confined to the gaps between microarray features ("spots"), it does not strongly affect the sequence fidelity within the features.

■ CONCLUSIONS

We have presented a method for doubling the efficiency of *in situ*, light directed microarray synthesis by assembling a reaction cell from two very closely spaced substrates. The method is straightforward, and we have adopted the method for routine synthesis of both DNA and RNA microarrays and for applications including gene expression and miRNA expression studies.[18,19] For microarray applications requiring high sequence fidelity, the reaction cell assembly provides a chamber that can be used to completely suppress reflections.

AUTHOR INFORMATION

Corresponding Author

*E-mail: mark.somoza@univie.ac.at.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Funding by the University of Vienna, the Faculty of Chemistry of the University of Vienna, the Austrian Science Fund (Grant FWF P23797), the Austrian Federal Ministry of Economy, Family and Youth, and the Austrian National Foundation for Research, Technology and Development is gratefully acknowledged. We thank John Wallace and Kurt Heinrich for providing assistance with material sourcing, specifications, and mechanical drawings from the original reaction cell and Walter Leuthner for machining the support block and prototypes of various components.

REFERENCES

(1) Fodor, S. P.; Read, J. L.; Pirrung, M. C.; Stryer, L.; Lu, A. T.; Solas, D. *Science* **1991**, *251*, 767−773.
(2) Singh-Gasson, S.; Green, R.; Yue, Y.; Nelson, C.; Blattner, F.; Sussman, M.; Cerrina, F. *Nat. Biotechnol.* **1999**, *17*, 974−978.
(3) Beier, M.; Hoheisel, J. D. *Nucleic Acids Res.* **2000**, *28*, e11.
(4) Hasan, A.; Stengele, K.; Giegrich, H.; Cornwell, P.; Isham, K.; Sachleben, R.; Pfleiderer, W.; Foote, R. *Tetrahedron* **1997**, *53*, 4247−4264.
(5) Pirrung, M. C.; Wang, L.; Montague-Smith, M. P. *Org. Lett.* **2001**, *3*, 1105−1108.
(6) Agbavwe, C.; Somoza, M. M. *PLoS ONE* **2011**, *6*, e22177.
(7) Agbavwe, C.; Kim, C.; Hong, D. G.; Heinrich, K.; Wang, T.; Somoza, M. M. *J. Nanobiotechnol.* **2011**, *9*, 57.
(8) Franssen-van Hal, N. L. W.; van der Putte, P.; Hellmuth, K.; Matysiak, S.; Kretschy, N.; Somoza, M. M. *Anal. Chem.* **2013**, *85*, 5950−5957.
(9) Bhushan, K. R. *Org. Biomol. Chem.* **2006**, *4*, 1857−1859.
(10) Shin, D.-S.; Lee, K.-N.; Yoo, B.-W.; Kim, J.; Kim, M.; Kim, Y.-K.; Lee, Y.-S. *J. Comb. Chem.* **2010**, *12*, 463−471.
(11) Lackey, J. G.; Mitra, D.; Somoza, M. M.; Cerrina, F.; Damha, M. J. *J. Am. Chem. Soc.* **2009**, *131*, 8496−8502.
(12) Wang, T.; Oehrlein, S.; Somoza, M. M.; Perez, J. R. S.; Kershner, R.; Cerrina, F. *Lab Chip* **2011**, *11*, 1629−1637.
(13) Lackey, J. G.; Somoza, M. M.; Mitra, D.; Cerrina, F.; Damha, M. J. *Chim. Oggi* **2009**, *27*, 30−33.
(14) Garland, P. B.; Serafinowski, P. J. *Nucleic Acids Res.* **2002**, *30*, e99.
(15) Offner, A. *Opt. Eng.* **1975**, *14*, 130−132.
(16) Offner, A. *Photogr. Sci. Eng.* **1979**, *23*, 374.
(17) Craft, N. E.; Soares, J. H. *J. Agric. Food Chem.* **1992**, *40*, 431−434.
(18) Holik, A.-K.; Rohm, B.; Somoza, M. M.; Somoza, V. *Food Funct.* **2013**, *4*, 1111−1120.
(19) Rohm, B.; Holik, A.-K.; Somoza, M. M.; Pignitter, M.; Zaunschirm, M.; Ley, J. P.; Krammer, G. E.; Somoza, V. *Mol. Nutr. Food Res.* **2013**, DOI: 10.1002/mnfr.201200846.

2.2 Express photolithographic DNA microarray synthesis with optimized chemistry and high-efficiency photolabile groups

Matej Sack, Kathrin Hölz, Ann-Katrin Holik, Nicole Kretschy, Veronika Somoza, Klaus-Peter Stengele and Mark M. Somoza

Submitted to the Journal of Nanobiotechnology December 2015, peer-reviewed, revised and resubmitted.

Motivation

The major goal of this paper was not only to improve and in this term to shorten the time of each light-directed microarray synthesis but also apply all gathered investigations for a genome-wide gene expression synthesis, a very powerful analytical tool made way more practical by decreasing the initial time of about 8 hours to just 1.5 hours. The more light-sensitive and higher photolysis quantum yield thiophenyl-NPPOC phosphoramidite was compared to the widely used NPPOC phosphoramidite. Shorter coupling to coupling times were achieved due to optimized chemistry with different activators and improved oxidation and drying protocols. Activators weren´t just checked in terms of highest signal-to-noise ratio but also the feature homogeneity. Less uniform features render the data extraction less reliable. Since different parameters of the standard protocol had room for optimization time parameters of each chemical incubation time and helium blow were improved and therefore investigated.

Express Photolithographic DNA Microarray Synthesis with Optimized Chemistry and High-Efficiency Photolabile Groups

Matej Sack[a], Kathrin Hölz[a], Ann-Katrin Holik[b], Nicole Kretschy[a], Veronika Somoza[b,c], Klaus-Peter Stengele[d] and Mark M. Somoza[a]*

a. Institute of Inorganic Chemistry, Faculty of Chemistry, University of Vienna, Vienna, Austria.

b. Department of Nutritional and Physiological Chemistry, Faculty of Chemistry, University of Vienna, Vienna, Austria.

c. Christian Doppler Laboratory for Bioactive Aroma Compounds, University of Vienna, Vienna, Austria.

d. Roche Diagnostics, Penzberg, Germany.

*Corresponding author: mark.somoza@univie.ac.at

Abstract

Background: DNA microarrays are a core element of modern genomics research and medical diagnostics, allowing the simple and simultaneous determination of the relative abundances of hundreds of thousands to millions of genomic DNA or RNA sequences in a sample. Photolithographic *in situ* synthesis, using light projection from a digitally-controlled array of micromirrors, has been successful at both commercial and laboratory scales. The advantages of this synthesis method are its ability to reliably produce high-quality custom microarrays with a very high spatial density of DNA features using a compact device with few moving parts. The phosphoramidite chemistry used in photolithographic synthesis is similar to that used in conventional solid-phase synthesis of oligonucleotides, but some unique differences require an independent optimization of the synthesis chemistry to achieve fast and low-cost synthesis without compromising microarray quality.

Results: High microarray quality could be maintained while reducing coupling time to a few seconds using DCI activator. Five coupling activators were compared, which resulted in microarray hybridization signals following the order ETT > Activator 42 > DCI >> BTT >> pyridinium chloride, but only the use of DCI lead to both high signal and highly uniform feature intensities. The photo-deprotection time was also reduced to a few seconds by replacing the NPPOC photolabile group with the new thiophenyl-NPPOC group. Other chemical parameters, such as oxidation and washing steps were also optimized.

Conclusions: Highly optimized and microarray-specific phosphoramidite chemistry, along with the use of the new high-efficiency thiophenyl-NPPOC photolabile group allow for the synthesis of high-complexity arrays of DNA 60 mers with a cycle time (coupling to coupling) of about 50 seconds, resulting in a three-fold reduction in synthesis time.

Keywords: Microarray; phosphoramidite chemistry; NPPOC; thiophenyl-NPPOC; photolabile.

Background

DNA microarrays are one of the core technologies for genomic research, allowing scientists access to the full breadth and complexity of genomes in single experiments. Typical microarray experiments focus on quantifying the abundance of nuclear DNA and RNA for insights into gene expression [1] and into the regulation of gene expression via, e.g., epigenetics [2] and micro RNA expression [3]. More recently, DNA microarrays have proved to be valuable beyond hybridization-based assays, for measuring the affinity

and specificity of DNA-binding proteins [4, 5], as platforms for aptamer-based multiplexed bioaffinity assays [6-8], and for large-scale oligonucleotide synthesis for assembly into genes [9-11], for targeted sequence capture and enrichment [12, 13], for rational design of antibody libraries via phage display [14], and for the creation of genome-wide knockout bacterial [15] and cell libraries [16, 17]. The developing technology of RNA microarrays, synthesized directly using phosphoramidite chemistry [18, 19], or synthesized enzymatically from a DNA microarray template [20], and peptide nucleic acids arrays [21] also have important applications in genomics and bioaffinity research, and share many synthesis and technological aspects.

Modern DNA microarrays are synthesized using a variety of *in situ* methods, all based on modifications of the high-efficiency phosphoramidite chemistry developed by Caruthers and coworkers [22, 23]. The original and still most common approach to *in situ* microarray synthesis is a derivative of photolithographic technology. The photolithographic method is based on the use of optical imaging systems to deliver light to the synthesis surface, where array layout and sequences are determined by selective removal of the photocleavable protecting groups on the terminus of each oligonucleotide. The primary advantages of the photolithographic approach are the very high surface density of unique DNA features that can be achieved, and the speed and flexibility of the synthesis chemistry. The flexibility of the approach results with the use of an imaging system centered on a digital micromirror device (DMD) in place of photomasks to deliver patterned light to the synthesis surface. This approach, termed Maskless Array Synthesis (MAS), allows virtual mask to control layout and the oligonucleotide sequences on the array. The speed of the photolithographic approach is due to the ability to use high efficiency photolabile groups that can be removed quickly with light exposure, and the minimal set-up time for new microarray designs. Early DNA photolithographic synthesis used 5′-(α-methyl-2-nitropiperonyl) oxycarbonyl (MeNPOC) [24] and dimethoxybenzoincarbonate (DMBOC) [25] protected phosphoramidites. The relatively low stepwise yield obtained with these groups has limited their use to microarrays of short oligomers [26]. This limitation was overcome with the discovery of the 2-(2-nitrophenyl)-propoxycarbonyl (NPPOC) group, which provides the almost quantitative coupling yield and significantly higher photolysis quantum yield necessary for the synthesis of long oligonucleotide microarrays [27, 28]. The principle limitation of the NPPOC group is its low absorptivity ($\varepsilon_{365nm/DMSO}\approx260$ $M^{-1}cm^{-1}$). These shortcomings of NPPOC have been overcome with a recently developed derivative, thiophenyl-2-(2-nitrophenyl)-propoxycarbonyl (SPh-NPPOC), which has both a higher quantum yield for photodeprotection and a much higher absorptivity [29]. The overall photodeprotection efficiency of SPh-NPPOC is 12 times greater than that of NPPOC. This allows such short exposures that the synthesis time is then dominated by the phosphoramidite coupling reaction and washing steps. Here we report on significant optimizations to the microarray synthesis chemistry, which combined with the new photolabile group, allows for very fast and efficient synthesis of high-density DNA microarrays. The optimization experiments presented here include the evaluation of alternative activators and activator concentrations, the determination of the optimal coupling time, the best oxidation strategy, and other chemical synthesis parameters.

Results

Activator optimization

The coupling reaction that extends the oligonucleotide chain by one nucleotide unit relies on the nucleophilic substitution, by the terminal 5′-hydroxyl group, of the diisopropylamino group of the phosphoramidite [30]. This reaction requires the involvement of an activator, first to first protonate, and then to displace the nitrogen of the leaving group, resulting in a reactive intermediate vulnerable to nucleophilic attack by the nucleosidic alcohol. Thus, the reaction rate should benefit from more acidic and more nucleophilic activators. Many activators have been developed with the aim to increase the speed of the reaction, particularly for sterically hindered phosphoramidites, i.e., RNA phosphoramidites. We have successfully used 4,5-dicyanoimidazole (DCI) [31] as an activator in

microarray synthesis for many years, but decided to test alternative activators that might allow for faster coupling. One of the key differences between standard solid phase synthesis of oligonucleotides and photolithographic synthesis is that solid phase synthesis relies on the use of the acid-labile dimethoxytrityl (DMT) 5'-hydroxyl protecting group, and is thus somewhat sensitive to very acidic activators, which can prematurely remove some DMT groups, leading to *n*+1 errors [32]. The photolabile groups are not limited in this manner, suggesting that very acidic activators could be used to reduce microarray synthesis time. DCI itself is known as an effective activator even though it has a relatively high pKa, presumably because it is a better nucleophile. In order to determine if alternative activators could be used to shorten the coupling time, we tested and compared five activators which have been suggested in the literature as effective activators: DCI, 5-ethylthio-1H-tetrazole (ETT) [33], 5-benzylthio-1H-tetrazole (BTT) [34], 5-[3,5-Bis(trifluoromethyl)-phenyl]-1H-tetrazole (Activator 42) [35], and pyridinium chloride [36].

The first activator experiments were to determine, for each of the five activators, the optimal coupling time. For these experiments, the same DNA sequence was synthesized four times on the microarray surface, with a different coupling time: 60 30, 15 and 6 seconds. The resulting microarrays were hybridized with the fluorescently labeled complementary sequence and scanned. High fluorescence signal was taken as a proxy for effective coupling during synthesis. Microarray synthesis using pyridinium chloride as the activator resulted in very low signal (data not shown). The results for the remaining four activators tested are graphed in Figure 1. All activators produced microarrays with high hybridization signal/noise, except for BTT, which resulted in arrays with relatively low signal and high background noise. Activator 42 resulted in the highest signal/background, but the signal was strongly dependent on coupling time, with the shorter coupling times resulting in relatively weak hybridization values. Synthesis with both DCI and ETT also resulted in strong hybridization signals, but with

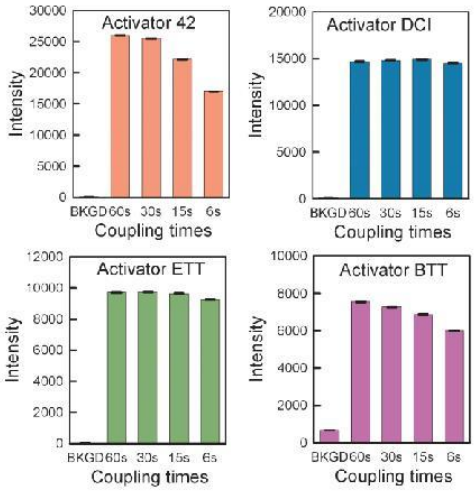


Figure 1. *The optimal coupling times for Activator 42, DCI, ETT and BTT were determined in microarray synthesis and hybridization experiments. DCI and ETT activators result in maximum hybridization signal at very short coupling times whereas the hybridization signal from arrays synthesized with BTT and 42 increases with coupling time. Error bars are the SEM.*

the advantage that the strong signal could also be obtained with very short coupling times, under 15 seconds. Because it is difficult to make sufficiently accurate absolute hybridization intensity comparisons between microarrays, a synthesis was designed to result in a microarray with a single DNA sequence, but with three or four sets of replicates, each replicate set synthesized using a different protocol. The sets were synthesized approximately in parallel, as shown in Figure 2A, in order to minimize order-of-synthesis effects. When the syntheses are performed in series, the oligonucleotides synthesized first hybridize more weakly, indicating that exposure to synthesis reagents degrades the DNA. For multiple 25mers synthesized in series, the hybridization intensity drops by ~6% relative to that of the subsequently synthesized oligonucleotide. This observed degradation indicates that minimizing synthesis time also results in higher quality microarrays.

These coupling results were then used to design an experiment allowing a direct comparison between the three best activators, Activator 42, ETT and DCI. Based on the results shown in Figure 1, the coupling times were chosen to be 30 seconds for Activator 42, and 15 sec for both DCI and ETT. The hybridization intensity results from this microarray synthesis are shown in Figure 2B.

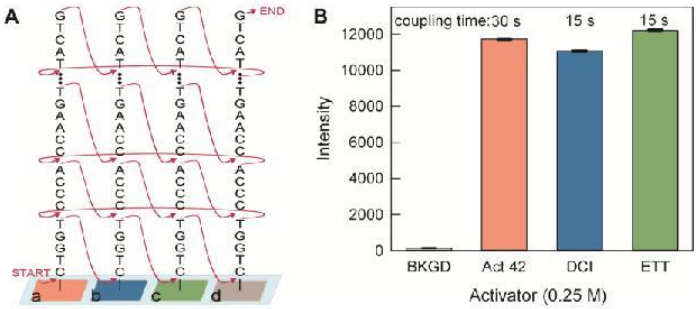


Figure 2. *Direct comparison activators by synthesizing sets of mixed base 25mers on a single microarray surface. Each oligonucleotide set was synthesized using a different activator. To avoid order-of synthesis effects, all sets were synthesized approximately in parallel using the scheme shown in (A). The results were evaluated using the intensity values obtained by hybridizing all the set with a common Cy3-labeled complementary oligonucleotide.*

The results are similar for all three activators but follow the order ETT > Activator 42 > DCI, which indicates that ETT is the better choice since it results in a slightly higher hybridization intensity than Activator 42 while requiring only half the coupling time. However, a visual examination of the scan images used to generate Figure 1 indicated activator-specific effects on the intra- and inter-feature intensity homogeneity. Figure 3 shows a detail of the center of microarrays synthesized using Activator 42, DCI, ETT and BTT. Except for the array synthesized with DCI, all of the arrays have bright spots which appear to have resulted from the final drying step after hybridization. We speculate that the more acidic activators modify the wetting properties of the surface, making it more susceptible to spot formation. The spot formation was observed consistently with multiple arrays synthesized with Activator 42, ETT and BTT. It may also be possible to avoid this issue by using an alternative surface functionalization, or an different hybridization washing and drying protocol, however, it appears from the images in Figure 3 that at least some of the hybridization intensity measured for Activator 42, ETT and BTT originates in the drying spot rather than from the actual hybridization signal. Based on these results, we decided to retain the use of DCI activator, but to reduce the coupling time from 1 minute to 15 seconds.

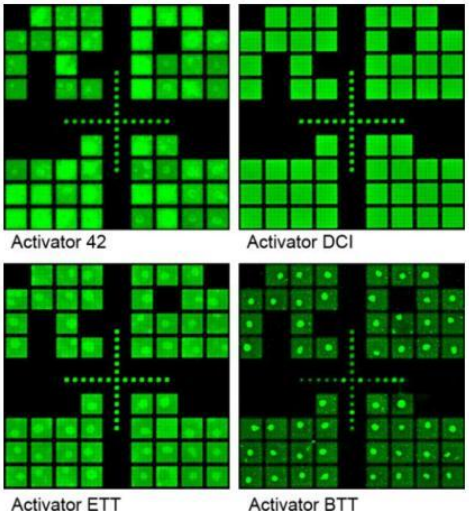


Figure 3. *Scan image details of microarrays synthesized using four different activators. Only the use of DCI activator resulted in microarrays with highly homogenous features. The small features are 14 x 14 µm and the large rectangles are made up of an array of 5 x 5 of the smaller features.*

It may be possible to reduce the coupling reaction time by increasing the concentration of the activator. Conversely, decreasing the concentration might be beneficial as well, particularly since the 0.25 M standard used in solid phase DNA synthesis may be too high given the relatively low phosphoramidite concentration (30 mM) we use in microarray synthesis. Low monomer concentrations can be used in microarray synthesis because each microarray includes only about 20 pmol [37] of oligonucleotides, and only about one quarter of these need to be extended with the corresponding

phosphoramidite during any given synthesis cycle. This synthesis scale is 3 to 4 orders of magnitude smaller than the smallest scale normally used in solid phase synthesis. A 30 mM phosphoramidite concentration is amply sufficient at this low scale while providing a margin to protect against incidental water contamination. To determine if the coupling reaction can be improved by increasing or decreasing the activator concentration, we synthesized microarrays using the scheme depicted in Figure 2A, but using different concentrations of the same activator instead of different activators. DCI and ETT were tested in separate experiments. The results are shown in Figure 4. The results indicate that a lower activator concentration (0.125 M) does not work as well. A higher concentration (0.5 M) does result in higher hybridization intensity in the case of ETT. For DCI, 0.25 M is close to optimal.

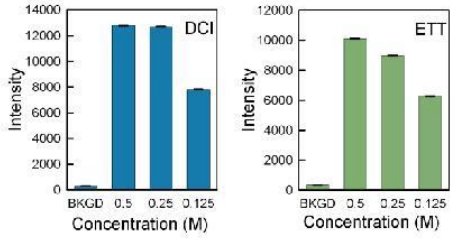


Figure 4. *The effect of DCI and ETT activator concentration on hybridization intensity. Higher ETT concentration results in higher hybridization signals, but for DCI, 0.25 M is sufficient to achieve the highest hybridization signal value.*

Oxidation

Conventional solid phase synthesis of nucleic acids requires an oxidation step preceding the removal of the DMT group because otherwise the acidic deblocking solution would cleave the phosphite triester formed by the coupling reaction. The phosphate trimester is stable and rapidly formed by iodine oxidation in the presence of water and pyridine. In photolithographic microarray synthesis, the use of the oxidizer solution can be minimized because the photodeprotection step does not affect the phosphite trimester. A final oxidation is still necessary before removing the protecting groups at the end of the synthesis. This minimal oxidation has the advantage of lowering synthesis time and reducing the risk of low coupling yield due to the high water content of the oxidizer. Previously, we had determined that intermittent oxidization or a single final oxidation was slightly preferable to an oxidation in every cycle [38], but decided to revisit this issue. While DCI is not sufficiently acidic to cleave the phosphite triester, ETT, Activator 42 and BTT are [35], therefore requiring oxidation in each cycle. To minimize the oxidation time, as well as water contamination, we compared using an oxidation in the final synthesis cycle only against the same oxidation protocol, but applied in every synthesis cycle. We also tried a very short oxidation exposure in each cycle. Figure 5 shows the results of these experiments. Since the oxidation reagent affects the entire array surface, the scheme shown in Figure 2A cannot be used to distinguish between oxidation protocols. Instead, four sets of 25mer probes with the same sequence were synthesized in series. As mentioned above, when probes are synthesized in series, rather than in parallel, the probes synthesized earlier are damaged by ongoing reagent exposure and do not hybridize as well as probes synthesized later. The dotted lines in the figure represent this ~6% trend in increased

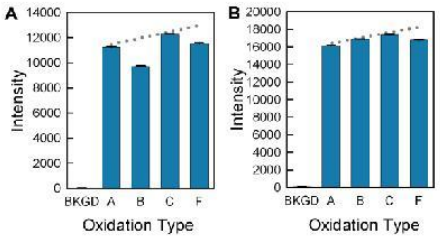


Figure 5. *Oxidation optimization. Graph A shows the hybridization intensity features on a microarray synthesized with a long oxidation step in each cycle (A), a long oxidation step in each cycle but no helium drying step (B), a sort oxidation step in each cycle (C), and a single final oxidation (F). Values below the dotted line indicate worse results. In Graph B, bars A, B and C are short oxidations steps in each cycle and F is a single final oxidation.*

hybridization intensity of probes synthesized later. The results show that very short oxidations steps, using 3 pulses (~36 µL) of oxidizer, delivered in one second, are as effective as the oxidations steps 10 times longer (~360µL in 10 sec.). Also, the short oxidations steps in each cycle result in microarrays with higher hybridization signal, as compared with the use of a single, final oxidation. The more frequent oxidation increases the total synthesis time slightly. For a typical gene expression microarray of 60mers, synthesized with 160 cycles, the total oxidation time would only amount to 160 seconds, compared 40 seconds with the previous method of a oxidizing for 10 seconds after every 40 cycles.

Drying

For reasons that remain still unknown, the microarray synthesis benefits from a drying step before the photodeprotection exposure [38]. In this step, helium or argon is flowed over the glass surface until it is dry. Previously, we have used a 30 second drying step, but decided to try to reduce this in order to be able to synthesize microarrays more quickly. Also, since the synthesis is now performed in a reaction chamber with a reduced depth (50 µm vs. 70 µm), in order to make two array simultaneously [39], the drying should be faster. To determine the optimal helium drying time, we synthesized, as with the oxidation tests, four sets of 25mer probes with the same sequence in series on the same array. For one microarray, the drying times were 5, 15, 30 and 0 seconds (Figure 6A), and for another arrays the drying time were 10, 20, 30 and 0 seconds (Figure 6B). The dotted line in each graph shows the expected ~6% trend in increased hybridization intensity of probes synthesized later. The data indicate that, while no drying (0 s) results in a drop of almost 20% in the hybridization intensity, any drying time above 5 seconds works equivalently, allowing for a significant reduction of synthesis time. Bar B of Figure 5A, representing the hybridization intensity of features also synthesized without a drying step shows a similar reduction.

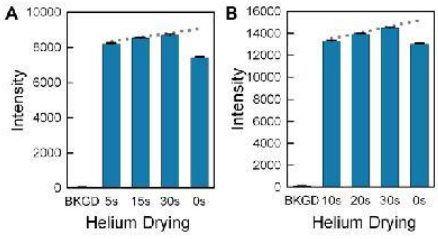


Figure 6. *Microarray drying step optimization. Two microarrays synthesized with a variety of helium drying times indicate that the drying step between coupling and light exposure significantly increases hybridization intensity, but that even short drying times are effective.*

Photodeprotection of SPh-NPPOC vs. NPPOC

The photodeprotection step of the synthesis has been the most time consuming. The NPPOC photolabile group removal requires a radiant exposure of approximately 6 J/cm^2, which can be achieved with a 75 second exposure using 80 mW/cm^2. Significantly higher irradiances are difficult to achieve because, while the total power emitted by arc sources increases with larger lamp size, the larger arc size of larger lamps actually results in less usable light due to the low numerical aperture (NA) of the synthesizer optics. Low NA is necessary to reduce synthesis errors due to scattered light, which scales approximately with NA^2 [40].

Recent results with SPh-NPPOC have demonstrated that it has a 12-fold greater

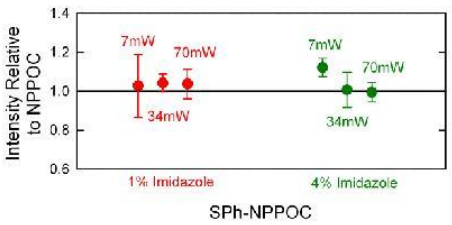


Figure 7. *Hybridization intensity values for microarrays of 25mers synthesized using SPh-NPPOC relative to the equivalent arrays synthesized using NPPOC. Synthesis was carried out using an exposure solvent consisting of either 1% or 4% imidazole in DMSO, and at radiant power values of 7, 34, or 70 mW/cm^2.*

photolytic efficiency vs. NPPOC [29]. This allows a faster deprotection or a less intense light source, or a combination of these two parameters. Figure 7 shows that the photolysis of SPh-NPPOC can be carried out successfully over a wide range of radiant power values. Specifically, we used a radiant exposure of 0.5 J/cm^2 and radiant power values of either 7, 34, or 70 mW/cm^2, corresponding to exposure times of 70, 15, and 7 seconds, respectively. We hypothesized that for the shortest exposure times with SPh-NPPOC, the exposure solvent might be more effective with a higher concentration of imidazole. This is because NPPOC—and presumably SPh-NPPOC—photolysis proceeds via a photo-induced β-elimination pathway which requires a small amount of base, preferably an amine base such as imidazole or *N*,*N*-diisopropylethylamine [41, 42]. The proton abstraction rate could be limiting under fast deprotection conditions. The data in Figure 7 shows that 1% imidazole is sufficient even for the 7 second photodeprotection.

Gene Expression Microarrays

The results from the optimization experiments described above indicate that the microarray synthesis time can be greatly reduced. Specifically, DCI still appears to be the best choice of activator, but the coupling time can be reduced to 15 seconds. The optimized oxidation protocol, oxidize for ~1 second each cycle, improves the hybridization intensity by about 10% without significantly extending the synthesis time. The helium drying time can also be reduced to 5 seconds or perhaps less. Using SPh-NPPOC allows the exposure time to be reduced to about 10 seconds or less. Since the individual optimizations were tested independently, it was important to determine if they could be successfully combined into a synthesis protocol for making useful microarrays.

Among the most complex and demanding microarray synthesis is that for high-density gene expression microarrays. Using the optimized protocol, we synthesized two sets of gene expression microarrays. One set was made using the legacy protocol with 60 seconds coupling, 30 seconds helium drying, NPPOC photodeprotection with 6 J/cm^2, and an oxidation every 40 cycles and at the end ("Legacy" synthesis). Another set gene expression microarrays of the same design was made using the new protocol with 15 seconds coupling, 10 seconds drying with helium, a short oxidation each cycle, and 10 seconds SPh-NPPOC photodeprotection ("Express" synthesis). The design of the gene expression microarray included two replicates of each of at least 3 unique 60-mer probes for more than 45000 human genes. In addition, 20 to 100 replicates of several quality control and reference sequences were also included. Using a checkerboard-like layout, one-half of the available synthesis features were used to generate a total of 382536 probe and control oligonucleotides. The microarrays were tested by hybridization with Cy3-labeled cDNA produced from mRNA extracted from a human colon adenocarcinoma cell line (Caco-2). Table 1 summarized the quality control metrics (from Cy3-labeled synthetic spike-in oligonucleotides) from these experiments. Both approaches result in high quality data, but the Express synthesis method is ~3 times faster. Figure 8 shows details of the images, along with the corresponding $\log_2$ scatter plots of the probe-level data normalized using the robust multiarray average (RMA) procedure [43].

Table 1. *Quality control data for labeled synthetic spike-in oligonucleotides used for quality control in the hybridization of the gene expression microarrays. Average intensity and coefficient of variation (c_v) for raw intensity data (top), and expression and standard error (SE) values for Robust Multi-array Average (RMA) normalized data (bottom).*

	Legacy		**Express**	
	Average	c_v	**Average**	c_v
QC-25mer	1386	0.18	2265	0.20
EcoBioA1	1437	0.27	2118	0.21
EcoBioD2	3254	0.22	2274	0.19
	Expression	**SE**	**Expression**	**SE**
QC-25mer	1066	0.97	2263	0.97
EcoBioA1	1102	0.97	2220	0.97
EcoBioD2	2741	0.97	2372	0.97

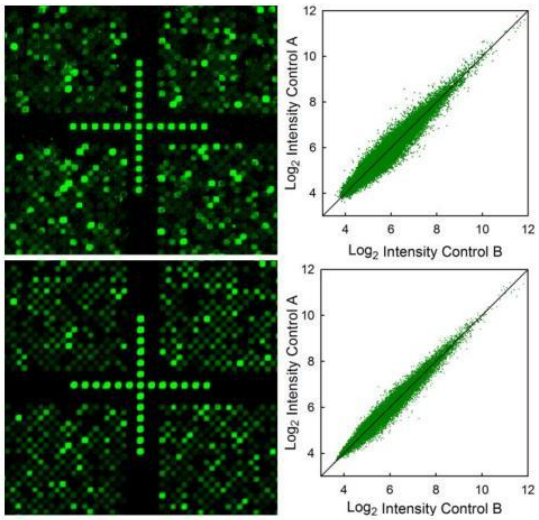


Figure 8. *Left: Details of 2.5 mm resolution scan images from gene expression microarrays synthesized with the Legacy method (top) and the Express method (bottom) and hybridized with Cy3-labeled cDNA and synthetic spike in controls. The size of each square is ~14x14 µm. Right: Scatterplots of the RMA-processed expression data from the gene expression microarrays synthesized with the legacy method (top) and the express method (bottom).*

Although the standard error values from the RMA normalized data are the same for both synthesis methods, the image quality for microarrays synthesized using the Express method appear to be consistently better than what we typically achieve using the Legacy method. This is also visible in Figure 8, where the features from the Express microarray scan have a more homogenous morphology. This difference may be a result of the shorter synthesis time for the Express synthesis, which reduces the chance that temperature drifts can cause slight changes in the alignment of the optical system. Such drifts can be mitigated by actively realigning the optics during synthesis using an image locking system [44], but such systems significantly increase the complexity of the synthesizer and increase synthesis time.

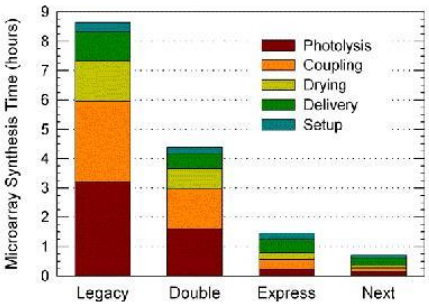


Figure 9. *Synthesis time for high-density gene expression microarrays. The optimizations presented here are labeled as "Express".*

In order to be able to use the gene expression data to evaluate microarray quality, the two microarrays synthesized with each method were hybridized with cDNA from untreated cells. The deviations from the diagonal line in the scatterplots in Figure 8 indicate the noise in the expression data rather than differential gene expression. It is clear that the express method produced microarrays that yield hybridization data with less noise. Although our experience with this new synthesis method is still limited, our experience to date suggests that the faster synthesis consistently generates microarray images with both more consistent spot morphology and reduced noise.

Discussion

Our objective with this project was to minimize the synthesis time for DNA microarrays without sacrificing quality. For both laboratory scale and industrial scale synthesis, throughput is the main determinant of both cost and productivity. At the same time, the optimization results allow for a better understanding of the underlying phosphoramidite chemistry as well as the photochemistry of the photolabile groups used in the synthesis. Using the synthesis time for human genome-wide gene expression microarrays as a metric, the legacy process we used until a few years ago required 8 to 9 hours. Such long syntheses make it impractical to produce more than one or

two gene expression microarrays per day per synthesizer. To increase the synthesis efficiency, we first introduced a method to double the efficiency by designing a photochemical reaction cell that could position two glass surfaces at the focus of the optical system [39]. One glass surface serves as the optical entrance to the reaction cell and one serves as the exit. Separated by ~50 µm, the inner surface of each glass slide receives a microarray (mirror images of each other) during the synthesis, and without any change in the synthesis chemistry or synthesis time (the setup time is ~2 minutes longer). The main limitation of this method is that the two microarrays must share a common design. With the advent of the SPh-NPPOC group, it became apparent that much shorter light exposures were possible, and served as an incentive for optimizing and shortening the remaining chemistry.

Figure 9 summarizes the typical synthesis times for gene expression microarrays, and includes the principle steps of each synthesis and their contribution to the overall use of time. Photolysis in the legacy method, along with coupling, dominate the original synthesis time ("Legacy"). The simultaneous synthesis of mirror image arrays ("Double") reduces the per array synthesis time in half even though the actual synthesis protocol is unchanged. The new optimization efforts presented here is referred to as "Express". Express synthesis uses a 15 seconds for coupling, 10 seconds for drying, and 10 second light exposures. With these short times, the previously relatively unimportant contributions to synthesis time, setup time and solvent/reagent delivery times become a significant fraction of the total. The nucleic acid synthesizer we use, an Expedite 8909 has a maximum delivery rate of 1/3 to 1/4 seconds per pulse (~12 µL/pulse), so that untimed steps, such as oxidizing, filling the reaction chamber with exposure solvent, or washing with acetonitrile, contribute measurably to the total synthesis time. For the Express synthesis, the volume of some of the washing steps was reduced, but nevertheless, "delivery" is now the largest contributor to the synthesis time. Nevertheless, the new per array synthesis time is about 1.5 hours; an almost 6-fold reduction compared the legacy method we were using until recently.

Even with these significant reductions in synthesis time, there is still considerable room for further reductions. In the case of coupling time, Figure 1 indicates that just 6 seconds is almost as good as the 15 seconds used for the express synthesis. It appears likely that relatively easy process changes, such as increasing the synthesis temperature by a few degrees, or increasing the phosphoramidite concentration above 30mM, would allow the coupling reaction to be reduced to just a few seconds. Currently, the reaction is performed at a relatively cool room temperature of 22°C. Increasing the room temperature, or partially enclosing the reaction chamber and heating it to ~30°C would likely be sufficient. The results presented in Figure 6 indicate that the drying step remains effective after another factor of 2 time reduction, from 10 to 5 seconds. Increasing the temperature seems likely to help here as well, but since the mechanism by which drying is beneficial remains unknown, this is less certain. The exposure time can also be further reduced; for the express synthesis, we used 10 second exposures at a radiant power of 50 mw/cm^2, but as Figure 7 shows, higher radiant power could be used to reduce the exposure time to 7 seconds, or less with a more intense light source. Reducing the delivery and setup times is also possible. These two categories now account for about one-half of the total synthesis time. Some or all of the washing may be unnecessary legacies of the original protocols, developed for solid phase synthesis, where the large surface areas of controlled pore glass require greater washing volumes. For example, it is known that the pyridine in the oxidation solution is a powerful quencher of the detritylation reaction [45], but in photolithographic synthesis, traces of oxidation solution probably will not interfere with photodeprotection, and therefore washing may be unnecessary between these two steps. The setup

time could also be reduced by using a second reaction chamber that could be assembled and tested while another synthesis is running. The time savings from these hypothetical reductions have been estimated and are shown in Figure 10 as "Next", with an average synthesis time for gene expression microarrays of about 45 minutes.

Beyond "Next", it may be challenging to make further large improvements without new development in the chemistry, such as even more efficient photolabile groups or better activators. Nevertheless, a further factor of ~2 can be achieved with a straightforward engineering solution. Three microarrays, of the same or different designs, can be synthesized on the same 25 x 75 mm slide surface by translating the reaction cell assembly and making 3 consecutive exposures. Since the exposure time needed by the SPh-NPPOC phosphoramidites is only a few seconds, and the remaining steps would be carried out simultaneously for all three arrays, the synthesis time would only be about 50% longer than synthesis with a single exposure. The solvent and reagent consumption would also increase about 50% due to the increased size of the reaction chamber. The resulting microarrays could then be independently hybridized.

While increasing the speed of synthesis was the primary goal of this project, it appears that microarray quality improves with the shorter synthesis time. Several of the experiments, such as those depicted in Figures 5 and 6, clearly indicated that one or more of the chemicals used in the synthesis partially degrades the oligonucleotides on the surface. This suggested that faster synthesis would result in improved microarray quality. This does appear to be the case. Although the quality assessment metrics we used to evaluate the gene expression microarrays (Table 1) indicate that the two synthesis methods yield similar microarrays, other metrics, such as spot morphology and the scatter in the gene expression data indicate that the express synthesis produces higher quality microarrays.

Conclusions

Optimizing microarray-specific phosphoramidite chemistry and using phosphoramidites with the highly efficient thiophenyl-NPPOC (SPh-NPPOC) photolabile group results in a large reduction in synthesis time without any loss of microarray quality. Combined with previous optimizations, the optimized method allows for high-density arrays of 60mers to be synthesized in about 90 minutes. The results also indicate that significant further improvements should be able to reduce synthesis time to less than 30 minutes per independently hybridizable microarray of 60mers.

Methods

Substrate preparation

All microarrays were synthesized as mirror image pairs using a method published earlier. [39] Briefly, half of the Schott Nexterion Glass D microscope slides (75x25x1mm) require two holes with a diameter of approximately 1mm that serve as the entrance and exit to the reaction chamber defined by one drilled slide and one undrilled slide separated using a 50 µm gasket (Gasoila Industrial Strength SD PTFE tape). The gasket is cut from the tape using a laser cutter. The glass slides are cleaned in an ultrasonic bath prior to functionalization.

Functionalization of substrates

The glass slides were functionalized with N-(3-triethoxysilylpropyl)-4-hydroxybutryamide [26] (Gelest SIT8189.5). The slides were placed in stainless steel rack and gently agitated in a solution of 2% (v/v) of the silane and 0.1% acetic acid in 95:5 ethanol/water. After the 4 hour silanization at room temperature, the slides were rinsed twice for 15 min in the 95:5 ethanol:deionized water and cured overnight at 120°C under vacuum. After cooling to room temperature under vacuum, the slides are stored in a desiccator until use.

Microarray synthesis

The Maskless Array Synthesizer (MAS) instrument consists of two major systems, an optical system and a chemical delivery system. The chemical side consists of an Expedite 8909 nucleic acid synthesizer, which delivers solvents and reagents to the reaction chamber where the microarray synthesis takes place. The optical system is similar to that of a photolithographic system, but it uses an array of 1024 x 768 digitally controlled mirrors (Texas Instruments 0.7 XGA DMD) in place of photomasks to pattern the ultraviolet light from a mercury lamp.

Light from a 350 W DC ultra-high pressure short arc mercury lamp (Newport 6286) is filtered by two consecutive 350-450 nm dichroic mirrors (Newport 66218). The resulting UV light, consisting of three unfiltered mercury lines (365, 405 and 436 nm), are spatially smoothed in a reflective homogenizing light pipe, and imaged onto the DMD. Light reflected by the DMD mirrors are imaged onto the two glass substrates using Offner relay optics. The intensity of UV light reaching the reaction cell image is adjusted according to the readings from a calibrated SÜSS intensity meter with a 365 nm probe (SÜSS MicroTec 1000).

The pattern of mirrors displayed on the DMD are imaged onto the two synthesis surfaces, where they determine microarray layout and oligonucleotide sequences by selectively removing the photolabile protecting groups, either NPPOC or SPh-NPPOC. Reagent delivery and the light exposures are controlled and synchronized by a computer. The phosphoramidite chemistry is similar to that used for solid-phase synthesis. The primary difference is the use of phosphoramidites with a 5'-OH photolabile protecting group. Upon absorption of a photon near UV photon, and in the presence of a weak amine base (typically 1% imidazole in DMSO), the NPPOC or SPh-NPPOC group comes off, leaving the 5' terminal hydroxyl, which reacts with activated phosphoramidite during the next coupling cycle. Representative synthesis protocols are shown in Table 2 for the legacy synthesis and in Table 3 for the express synthesis.

After synthesis, the groups are removed by immersing the microarrays in 1:1 (v/v) ethylenediamine and ethanol for 2 hours at room temperature, then washed twice in beakers filled with deionized water and dried with argon. Deprotected microarrays were stored in a desiccator cabinet until hybridized.

Table 2. *Representative chemical synthesis protocol, in Expedite 8909 format, used for the legacy syntheses. The 5'-NPPOC light-deprotection step takes place between the two "Event Out" commands, which trigger the opening and closing of a shutter blocking the UV light. The 70 sec exposure, at 80 mW/cm^2, corresponds to an exposure of 5.6 J/cm^2. The oxidation, the first two steps after the $Oxidizing header, are used intermittently (every ~40 cycles) and after the last coupling.*

Cycle NPPOC-dT (**Legacy synthesis**)				
Function	mode	pulses	sec	Description
$Coupling				
1/*Wsh	PULSE	20	0	Flush with Wsh
2/*Act	PULSE	6	0	Act
21/*T + Act	PULSE	5	0	T + Act
2/*Act	PULSE	6	0	Push with Act
1/*Wsh	PULSE	3	60	Couple monomer
1/*Wsh	PULSE	10	0	Flush with Wsh
$Capping				
40 /*Gas A	PULSE	1	30	Dry column
$Oxidizing				
15/*Ox	PULSE	30	0	Ox to column
12/*Wsh A	NA	45	0	Flush with Wsh A
17/*Aux	PULSE	15	0	Exposure solvent
130/*Event 2 Out	NA	4	1	Event 2 Out
17/*Aux	PULSE	10	50	Exposure solvent
12/*Wsh A	PULSE	5	20	Flush with Wsh A
130/*Event 2 Out	NA	4	1	Event 2 Out
12/*Wsh A	PULSE	20	0	Flush with Wsh A

Table 3. *Representative chemical synthesis protocol, in Expedite 8909 format, used for the express syntheses. The 9 sec exposure, at 50 mW/cm^2, corresponds to an exposure of 0.45 J/cm^2. The 5 pulses of Wsh A during the exposure push the exposure solvent towards the waste, but the reaction chamber remains full of exposure solvent until the next washing step. The short oxidation step in this synthesis, the first two steps after the $Oxidizing header, are used after every coupling.*

Cycle SPh-NPPOC-dT (**Express synthesis**)				
Function	mode	pulses	sec	Description
$Coupling				
1 /*Wsh	PULSE	10	0	Flush with Wsh
2 /*Act	PULSE	6	0	Act
21 /*T + Act	PULSE	5	0	T + Act
2 /*Act	PULSE	6	0	Push with Act
1 /*Wsh	PULSE	3	15	Couple monomer
1 /*Wsh	PULSE	10	0	Flush with Wsh
$Capping				
40 /*Gas A	PULSE	1	10	Dry column
$Oxidizing				
15/*Ox	PULSE	3	0	Ox to column
12/*Wsh A	NA	10	0	Flush with Wsh A
17/*Aux	PULSE	15	0	Exposure solvent
130/*Event 2 Out	NA	4	1	Event 2 Out
12/*Wsh A	PULSE	5	9	Push with Wsh A
130/*Event 2 Out	NA	4	1	Event 2 Out
12/*Wsh A	PULSE	10	0	Flush with Wsh A

NPPOC and SPh-NPPOC Phosphoramidites

NPPOC DNA phosphoramidites were purchased from Sigma-Aldrich (A112N01-01, C114N01-01, G114N01-01, T111N01-01) and diluted to 30mM with Amidite Diluent (<30ppm water) from Sigma-Aldrich (L010010-06). SPh-NPPOC phosphoramidites were manufactured by NimbleGen Systems GmbH (Waldkraiburg, Germany) and diluted as above.

Coupling time optimization

The coupling experiments were performed using the legacy protocol (Table 2), but with different coupling times, 5, 15, 30 and 60 seconds. In addition, an oxidation step, as indicated in the table, was included in every cycle due to the use of acidic activators. Each microarray was synthesized as shown in Figure 2A, with each of the replicate sets synthesized using a different coupling time (Figure 1). The same microarray design and synthesis protocol was used for each activator. Five activators were chosen, DCI in acetonitrile (Biosolve Chimie), 0.25 M ETT (Sigma L0302511), 0.25 M BTT (empBiotech NC-0102 L0302511), 0.25 M pyridine hydrochloride (Fluka 82800), 0.25 M Activator 42 (Aldrich L8300212). Each activator was used at a concentration of 0.25 M in anhydrous acetonitrile. After synthesis, the microarrays were deprotected and hybridized as described.

Activator optimization

The activator optimizations were performed using either 0.25 M DCI, 0.25 M ETT, 0.25 M BTT, 0.25 M pyridine hydrochloride, 0.25 M Activator 42. For the experiments using 0.125 M activator, the activator solution was diluted in anhydrous acetonitrile. For the experiments using 0.5 M activator, the solution was either concentrated by evaporating acetonitrile in a vacuum or (for DCI) by adding the crystalline form (Aldrich 554030).

Nucleic acid synthesizers have separate activator and phosphoramidite ports and mix the activator when needed. The Expedite 8909 accomplishes this mixing by drawing single pulses (~12 µL/pulse) of activator and phosphoramidite in an alternating fashion until reaching the desired volume, 5 pulses of each. This pulse train then mixes in the fluidics system and on the way to the synthesis area. In order to synthesize single arrays using multiple activators (Figures 3 and 4), the normal activator port was used for one of the activators and the other activators or activator concentrations were placed in unused phosphoramidite ports (the Expedite 8909 has 9 phosphoramidite ports). In order to mix the activators in these ports with the appropriate phosphoramidite, the normal mixing command in the protocol file (e.g. "21/*T + Act, PULSE, 5, 0"; monomer T and Act simultaneously) was replaced with 10 alternating single port commands, e.g.: "11/*T, PULSE, 1, 0"; Monomer T, "7/*T, PULSE, 1, 0"; Monomer 9". Command 7 delivers a single pulse out of port 9, which is filled with another activator. Several control experiments confirmed that this alternative command set resulted in equal coupling efficiency. The effectiveness of the activators was evaluated based on the homogeneity and intensity of the microarray features after hybridization.

Oxidizer and Drying optimization

Since both the oxidation and drying steps affect the entire microarray surface, it was not possible to use

the parallel synthesis scheme illustrated in in Figure 2B. Instead, replicates of the same 25mer sequence were synthesized four times in series. The location of each of several hundred synthesis replicates were randomized across the microarray surface. We observed that, when the same synthesis protocol was used for each of the syntheses, the hybridization intensity increased linearly from the first probe synthesized to the last, with an increase of ~6% per synthesis. Oxidation or drying protocols resulting in a drop in hybridization signal relative to this trend were judged as inferior. In the case of oxidation (Figure 5), we evaluated three protocols, a single final oxidation, 10 seconds oxidation per cycle (30 pulses) and 1 second oxidation per pulse (3 pulses). In all cases, the oxidation solution was tetrahyrofurane/water/pyridine/iodine 90.54/9.05/0.41/0.43 (v/v/v/w) (Sigma-Aldrich L860021). The same analysis was applied to the optimization of the helium drying time. In this case, we tried 0, 5, 10, 15, and 30 seconds of drying time after the coupling steps.

NPPOC vs. SPh-NPPOC comparison

In order to make direct comparisons between syntheses using NPPOC vs. SPh-NPPOC, we designed a microarray with probe replicates synthesized using NPPOC phosphoramidites as well as probe replicates synthesized using SPh-NPPOC phosphoramidites. Because all the monomer ports of the Expedite 8909 may not work equally well, the Sph-NPPOC synthesis was performed first, and paused while the monomer ports were cleaned out and filled with NPPOC phosphoramidites. The synthesis of the second set of microarray probes was then continued on the same surface. The location of each set of probes was randomized across the microarray surface. The experiment was performed using 7, 34 and 70 mw/cm^2 light exposures for the SPh-NPPOC amidites and using 1% or 4% (w/v) imidazole in DMSO (Figure 7). Because the probe set synthesized with SPh-NPPOC was performed first, the actual relative performance of SPh-NPPOC is likely ~6% better than indicated in Figure 7.

Genomic cDNA

The human colon adenocarcinoma cell line Caco-2 was cultured in Dulbecco's modified Eagle medium supplemented with 10% fetal bovine serum, 4mM L-glutamine and 1% penicillin/streptomycin. Cells were maintained at 37°C and 5% CO_2 in a humidified incubator and passaged at 90% confluence. For microarray experiments, 4x10^5 cells were seeded into 6-well plates and cultivated for 21 days to allow for enterocyte differentiation. Cultivation media was exchanged every second day.

Fully differentiated Caco-2 cells were incubated with serum-free media or serum-free media containing 500 µM of a test substance. After 90 min, the cells were washed with ice-cold PBS and RNA isolated using the RNeasy Mini Kit (Qiagen). The integrity of the isolated RNA was assessed by agarose gel electrophoresis in addition to photometrical analysis (Tecan, Menningen, Switzerland). A total of 10 µg RNA was reverse transcribed and simultaneously labeled using Cy3-labeled random nonamer primers (Tebu Bio, Offenbach, Germany) as described by Ouellet et al. [46].

Microarray hybridization

The microarrays used for the optimization experiments were hybridized in a solution containing 150 µl 2x MES, 110 µl nuclease free water, 13.3 µl acetylated BSA and 26.7 µl of 100 nM 5'-Cy3-labeled complementary sequence. In the case of the microarrays used for the gene expression experiment, the hybridization solution contained 3 µL herring sperm DNA (10mg/ml), 15 µL acetylated BSA (10mg/ml), 135 µL 2X MES hybridization buffer, Cy3-labeled cDNA in 85 µL water, 10 µL Cy3-labeled QC 25mer oligo (100nM), 10 µL Cy3-labeled ECO1BioA1 (100nM), and 10 µL Cy3-labeled ECO1BioD2 (100nM). All hybridizations took place in self-adhesive chambers (Grace Biolabs SA200). The microarrays were rotated (~30 Hz) in a hybridization oven at 42°C. An air bubble filling a quarter to a third of the chamber volume moves

around the hybridization chamber due to the rotation, circulating and mixing the hybridization solution, which promotes efficient hybridization (47]. After 4 hours for the optimization experiments or 22 hours for the gene expression experiments, the chamber was removed while submerged in a petri dish filled with non-stringent wash buffer pre-warmed to 42°C. The microarrays were washed with vigorous shaking in 50 ml centrifuge tubes filled with 30 ml non-stringent wash buffer (SSPE; 0.9 M NaCl, 0.06 M phosphate, 6 mM EDTA, 0.01% Tween20) for 2 min., and then similarly washed with stringent wash buffer (100 mM MES, 0.1 M Na^+, 0.01% Tween20) for 1 min. Finally, the microarrays were rinsed for about 5 seconds in final wash buffer (0.1× saline-sodium citrate buffer) to remove most of the salt before being dried using a microarray centrifuge. Microarrays were scanned at 2.5 µm resolution. Feature- or probe-level data was extracted with NimbleScan 2.1 (Roche-NimbleGen).

Gene expression analysis

The scanned images were analyzed using NimbleScan 2.1 (Roche-NimbleGen), employing robust multichip analysis (RMA) for normalization purposes. The extracted intensity data of each probe was $\log_2$ transformed and a scatter plot of control vs. treated samples created using SigmaPlot 11.0. [48, 49]

Gene expression microarray quality control

Three synthetic HPLC purified 5'-Cy3-labeled DNA oligomers, were added to the hybridization buffer at concentrations of 3.7 nM. The names and sequences of these oligonucleotides are:

GAC CAG GGT GGT TCA TGA TGA TGA C, QC_25mer

GAT TTA GGT TTA CAA GTC TAC ACC GAA TTA ACA ACA AAA AAC ACG TTT TGG AG, ECOBioA1t_53mer

GAA ATG AGG GTG TAA TTG ATT GGG CAA CTG TGC GCC ACG CTA CTT TCT TCT TCG CTT AAC, ECOBioD2_60mer

The microarray layout was designed with 100 probes for QC_25mer, 140 probes for ECOBioA1t_53mer, and 140 probes for EcoBioD2_60mer. The location of each feature on the microarray was randomized along with all the other probes. Several assessment metrics, based on these synthetic oligonucleotides, were used to evaluate the synthesis and hybridization quality the microarrays. The outcomes of the assessments are summarized in Table 1.

List of Abbreviations

BSA: Bovine serum albumin; BTT: 5-benzylthio-1H-tetrazole; DCI: 4,5-dicyanoimidazole; DMD: Digital micromirror device; DMT: Dimethoxytrityl; ETT: 5-ethylthio-1H-tetrazole; NPPOC: 2-(2-nitrophenyl)-propyloxycarbonyl; MAS: Maskless array synthesis; MES 2-(N-morpholino)ethanesulfonic acid; SPh-NPPOC: thiophenyl-2-(2-nitrophenyl)-propoxycarbonyl; SSC: Saline-sodium citrate.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All of the authors contributed substantially by performing experiments, analyzing data, contributing ideas or writing the manuscript. MS performed most of the lab work and analysis related to microarray synthesis, KH and AKH performed most of the experiments related to gene expression microarray synthesis and analysis, NK tested the performance of SPh-NPPOC phosphoramidites developed by KPS. VS and MMS supervised the experiments and data analysis and wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Funding by the University of Vienna, the Faculty of Chemistry of the University of Vienna, the Austrian

Science Fund (Grant FWF P23797), the Austrian Federal Ministry of Economy, Family and Youth, and the Austrian National Foundation for Research, Technology and Development are gratefully acknowledged.

References

1. Schena M, Shalon D, Davis RW, Brown PO: **Quantitative monitoring of gene expression patterns with a complementary DNA microarray**. *Science* 1995, **270**(5235):467-470.
2. Sandoval J, Heyn H, Moran S, Serra-Musach J, Pujana MA, Bibikova M, Esteller M: **Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome**. *Epigenetics* 2011, **6**(6):692-702.
3. Thomson JM, Parker J, Perou CM, Hammond SM: **A custom microarray platform for analysis of microRNA gene expression**. 2004, **1**(1):47-53.
4. Bulyk ML, Gentalen E, Lockhart DJ, Church GM: **Quantifying DNA-protein interactions by double-stranded DNA arrays**. *Nature Biotechnology* 1999, **17**(6):573-577.
5. Warren CL, Kratochvil NCS, Hauschild KE, Foister S, Brezinski ML, Dervan PB, Phillips GN, Ansari AZ: **Defining the sequence-recognition profile of DNA-binding molecules**. *Proceedings of the National Academy of Sciences of the United States of America* 2006, **103**(4):867-872.
6. McCauley TG, Hamaguchi N, Stanton M: **Aptamer-based biosensor arrays for detection and quantification of biological macromolecules**. *Analytical Biochemistry* 2003, **319**(2):244-250.
7. Katilius E, Flores C, Woodbury NW: **Exploring the sequence space of a DNA aptamer using microarrays** *Nucleic Acids Research* 2007 **35** (22):7626-7635
8. Franssen-van Hal NLW, van der Putte P, Hellmuth K, Matysiak S, Kretschy N, Somoza MM: **Optimized Light-Directed Synthesis of Aptamer Microarrays**. *Analytical Chemistry* 2013, **85**(12):5950-5957.
9. Richmond KE, Li M-H, Rodesch MJ, Patel M, Lowe AM, Kim C, Chu LL, Venkataramaian N, Flickinger SF, Kaysen J *et al*: **Amplification and assembly of chip-eluted DNA (AACED): a method for high-throughput gene synthesis**. *Nucleic Acids Research* 2004, **32**(17):5011-5018.
10. Tian J, Gong H, Sheng N, Zhou X, Gulari E, Gao X, Church G: **Accurate multiplex gene synthesis from programmable DNA microchips**. *Nature* 2004, **432**(7020):1050-1054.
11. Wu C-H, Lockett MR, Smith LM: **RNA-Mediated Gene Assembly from DNA Arrays**. *Angewandte Chemie International Edition* 2012, **51**(19):4628-4632.
12. Albert TJ, Molla MN, Muzny DM, Nazareth L, Wheeler D, Song X, Richmond TA, Middle CM, Rodesch MJ, Packard CJ *et al*: **Direct selection of human genomic loci by microarray hybridization**. *Nat Meth* 2007, **4**(11):903-905.
13. Hodges E, Xuan Z, Balija V, Kramer M, Molla MN, Smith SW, Middle CM, Rodesch MJ, Albert TJ, Hannon GJ *et al*: **Genome-wide in situ exon capture for selective resequencing**. *Nat Genet* 2007, **39**(12):1522-1527.
14. Larman HB, Jing Xu G, Pavlova NN, Elledge SJ: **Construction of a rationally designed antibody platform for sequencing-assisted selection**. *Proceedings of the National Academy of Sciences* 2012, **109**(45):18523-18528.
15. Warner JR, Reeder PJ, Karimpour-Fard A, Woodruff LBA, Gill RT: **Rapid profiling of a microbial genome using mixtures of barcoded oligonucleotides**. *Nat Biotech* 2010, **28**(8):856-862.
16. Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen TS, Heckl D, Ebert BL, Root DE, Doench JG *et al*: **Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells**. *Science* 2014, **343**(6166):84-87.
17. Wang T, Wei JJ, Sabatini DM, Lander ES: **Genetic Screens in Human Cells Using the CRISPR-Cas9 System**. *Science* 2014, **343**(6166):80-84.
18. Lackey JG, Mitra D, Somoza MM, Cerrina F, Damha MJ: **Acetal Levulinyl Ester (ALE) Groups for 2 '-Hydroxyl Protection of Ribonucleosides in the Synthesis of Oligoribonucleotides on Glass and Microarrays**. *Journal of the American Chemical Society* 2009, **131**(24):8496-8502.
19. Lackey JG, Somoza MM, Mitra D, Cerrina F, Damha MJ: **In-situ chemical synthesis of rU-DNA chimeras on chips and enzymatic recognition**. *Chemistry Today* 2009, **27**(6):30-33.
20. Wu C-H, Holden MT, Smith LM: **Enzymatic Fabrication of High-Density RNA Arrays**. *Angewandte Chemie International Edition* 2014, **53**(49):13514-13517.
21. Yang F, Dong B, Nie K, Shi H, Wu Y, Wang H, Liu Z: **Light-Directed Synthesis of High-Density Peptide Nucleic Acid Microarrays**. *ACS Combinatorial Science* 2015, **17**(10):608-614.
22. Beaucage S, Caruthers M: **Deoxynucleoside phosphoramidites—a new class of key**

intermediates for deoxypolynucleotide synthesis. *Tetrahedron letters* 1981, **22**(20):1859-1862.

23. McBride L, Caruthers M: **An investigation of several deoxynucleoside phosphoramidites useful for synthesizing deoxyoligonucleotides**. *Tetrahedron letters* 1983, **24**(3):245-248.
24. Pirrung MC, Bradley J-C: **Comparison of Methods for Photochemical Phosphoramidite-Based DNA Synthesis**. *The Journal of Organic Chemistry* 1995, **60**(20):6270-6276.
25. Pirrung MC, Fallon L, McGall G: **Proofing of Photolithographic DNA Synthesis with 3',5'-Dimethoxybenzoinyloxycarbonyl-Protected Deoxynucleoside Phosphoramidites**. *The Journal of Organic Chemistry* 1998, **63**(2):241-246.
26. McGall GH, Barone AD, Diggelmann M, Fodor SPA, Gentalen E, Ngo N: **The Efficiency of Light-Directed Synthesis of DNA Arrays on Glass Substrates**. *Journal of the American Chemical Society* 1997, **119**(22):5081-5090.
27. Beier M, Hoheisel JD: **Production by quantitative photolithographic synthesis of individually quality checked DNA microarrays**. *Nucleic Acids Research* 2000, **28**(4):e11-e11.
28. Pirrung MC, Wang L, Montague-Smith MP: **3'-Nitrophenylpropyloxycarbonyl (NPPOC) Protecting Groups for High-Fidelity Automated 5' → 3' Photochemical DNA Synthesis**. *Organic Letters* 2001, **3**(8):1105-1108.
29. Kretschy N, Holik A-K, Somoza V, Stengele K-P, Somoza MM: **Next-Generation o-Nitrobenzyl Photolabile Groups for Light-Directed Chemistry and Microarray Synthesis**. *Angewandte Chemie International Edition* 2015, **54**(29): 8555–8559.
30. Russell MA, Laws AP, Atherton JH, Page MI: **The mechanism of the phosphoramidite synthesis of polynucleotides**. *Organic & biomolecular chemistry* 2008, **6**(18):3270-3275.
31. Vargeese C, Carter J, Yegge J, Krivjansky S, Settle A, Kropp E, Peterson K, Pieken W: **Efficient activation of nucleoside phosphoramidites with 4,5-dicyanoimidazole during oligonucleotide synthesis**. *Nucleic Acids Research* 1998, **26**(4):1046-1050.
32. Krotz AH, Klopchin PG, Walker KL, Srivatsa GS, Cole DL, Ravikumar VT: **On the formation of longmers in phosphorothioate oligodeoxyribonucleotide synthesis**. *Tetrahedron letters* 1997, **38**(22):3875-3878.
33. Wincott F, DiRenzo A, Shaffer C, Grimm S, Tracz D, Workman C, Sweedler D, Gonzalez C, Scaringe S, Usman N: **Synthesis, deprotection, analysis and purification of RNA and ribosomes**. *Nucleic Acids Research* 1995, **23**(14):2677-2684.
34. Pitsch S, Weiss PA, Jenny L, Stutz A, Wu X: **Reliable Chemical Synthesis of Oligoribonucleotides (RNA) with 2'-O-[(Triisopropylsilyl)oxy]methyl(2'-O-tom)-Protected Phosphoramidites**. *Helvetica chimica acta* 2001, **84**(12):3773-3795.
35. Threlfall RN, Torres AG, Krivenko A, Gait MJ, Caruthers MH: **Synthesis and biological activity of phosphonoacetate- and thiophosphonoacetate-modified 2'-*O*-methyl oligoribonucleotides**. *Organic & biomolecular chemistry* 2012, **10**(4):746-754.
36. Beier M, Pfleiderer W: **Nucleotides, Part LXII , Pyridinium Salts – An Effective Class of Catalysts for Oligonucleotide Synthesis**. *Helvetica chimica acta* 1999, **82**(6):879-887.
37. Lietard J, Kretschy N, Sack M, Wahba AS, Somoza MM, Damha MJ: **Base-cleavable microarrays for the characterization of DNA and RNA oligonucleotides synthesized in situ by photolithography**. *Chemical Communications* 2014, **50**(85):12903-12906.
38. Agbavwe C, Kim C, Hong D, Heinrich K, Wang T, Somoza MM: **Efficiency, Error and Yield in Light-Directed Maskless Synthesis of DNA Microarrays**. *J Nanobiotechnol* 2011, **9**(57).
39. Sack M, Kretschy N, Rohm B, Somoza V, Somoza MM: **Simultaneous Light-Directed Synthesis of Mirror-Image Microarrays in a Photochemical Reaction Cell with Flare Suppression**. *Analytical Chemistry* 2013, **85**(18):8513-8517.
40. Kim C: **Gene synthesis from photolithographic DNA microarrays**. *Ph.D.*: University of Wisconsin-Madison; 2007.
41. Giegrich H, Eisele-Bühler S, Hermann C, Kvasyuk E, Charubala R, Pfleiderer W: **New photolabile protecting groups in nucleoside and nucleotide chemistry—synthesis, cleavage mechanisms and applications**. *Nucleosides & Nucleotides* 1998, **17**(9-11):1987-1996.
42. Hasan A, Stengele K-P, Giegrich H, Cornwell P, Isham KR, Sachleben RA, Pfleiderer W, Foote RS: **Photolabile protecting groups for nucleosides: Synthesis and photodeprotection rates**. *Tetrahedron* 1997, **53**(12):4247-4264.
43. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP: **Exploration, normalization, and summaries of high density oligonucleotide array probe level data**. *Biostatistics* 2003, **4**(2):249-264.
44. Kim C, Li M, Rodesch M, Lowe A, Richmond K, Cerrina F: **Biological lithography: Improvements in DNA**

synthesis methods. *Journal of Vacuum Science & Technology B* 2004, **22**(6):3163-3167.

45. LeProust EM, Peck BJ, Spirin K, McCuen HB, Moore B, Namsaraev E, Caruthers MH: **Synthesis of high-quality libraries of long (150mer) oligonucleotides by a novel depurination controlled process**. *Nucleic Acids Research* 2010, **38**(8):2522-2540.

46. Ouellet M, Adams PD, Keasling JD, Mukhopadhyay A: **A rapid and inexpensive labeling method for microarray gene expression analysis**. *BMC Biotechnology* 2009, **9**:97-97.

47. McQuain MK, Seale K, Peek J, Fisher TS, Levy S, Stremler MA, Haselton FR: **Chaotic mixer improves microarray hybridization**. *Analytical Biochemistry* 2004, **325**(2):215-226.

48. Rohm B, Holik A-K, Somoza MM, Pignitter M, Zaunschirm M, Ley JP, Krammer GE, Somoza V: **Nonivamide, a capsaicin analog, increases dopamine and serotonin release in SH-SY5Y cells via a TRPV1-independent pathway**. *Molecular Nutrition & Food Research* 2013, **57**(11):2008-2018.

49. Rohm B, Holik A-K, Kretschy N, Somoza MM, Ley JP, Widder S, Krammer GE, Marko D, Somoza V: **Nonivamide Enhances miRNA let-7d Expression and Decreases Adipogenesis PPARγ Expression in 3T3-L1 Cells**. *Journal of Cellular Biochemistry* 2015, **116**(6):1153-1163.

2.3 Base-cleavable microarrays for the characterization of DNA and RNA oligonucleotides synthesized *in situ* by photolithography

Jory Lietard, Nicole Kretschy, Matej Sack, Alexander S. Wahba, Mark M. Somoza and Masad J. Damha

This paper was published in the Journal of Chemical Communications in September 2014.

Motivation

The possibility to determine the source of error for *in situ* microarray syntheses, e.g. failure of deprotection steps, chemical degradation during and after synthesis and coupling inefficiency is crucial for improving the method. It was for the first time that the DNA and RNA oligonucleotide synthesis on microarrays could be analysed directly by mass spectrometry (LC-ESI-MS). For this purpose, a special custom-made cleavable dT monomer was made, and which was incorporated at the beginning of the sequence and used to separate the DNA or RNA from the glass surface. The results were used to evaluate synthesis and deprotection errors that might be occurring primarily in the synthesis of RNA microarrays. This evaluation was necessary to evaluate remaining obstacles for the synthesis of complex, high-density RNA microarrays to be used to study binding and affinity patterns of RNA-binding proteins.

ChemComm

COMMUNICATION

View Article Online
View Journal | View Issue

Cite this: *Chem. Commun.*, 2014, 50, 12903

Received 24th July 2014,
Accepted 5th September 2014

DOI: 10.1039/c4cc05771f

www.rsc.org/chemcomm

Base-cleavable microarrays for the characterization of DNA and RNA oligonucleotides synthesized *in situ* by photolithography†

Jory Lietard,[ab] Nicole Kretschy,[b] Matej Sack,[b] Alexander S. Wahba,[a] Mark M. Somoza*[b] and Masad J. Damha*[a]

Assessing synthesis efficiency, errors, failed deprotections, and chemical and enzymatic degradation of oligonucleotides on microarrays is essential for improving existing *in situ* synthesis methods, and for the development of new chemistries. We describe the use of LC-MS to analyse DNA and RNA oligonucleotides deprotected and cleaved under basic conditions from microarrays fabricated using light-directed *in situ* chemistry. The data yield essential information on array quality and sequence identity.

Arraying DNA onto chips has revolutionized the field of biomedical research,[1–4] most notably in gene expression profiling,[5] by providing an access to large nucleic acid libraries attached to one single support and by allowing the simultaneous screening of thousands of genes. These DNA libraries can originate from PCR products which are then covalently attached to the glass surface[6] or are synthesized *in situ* by ink-jet printing or photolithography,[7–9] taking advantage of the robust phosphoramidite chemistry.[10,11] The quality of the immobilized DNA is one of the crucial parameters governing the reliability of the measurement,[12] and while this parameter can be controlled to some extent for PCR products, the same level of quality assessment is less trivial for *in situ*-synthesized microarrays.

One method for quality control consists of labelling the terminus of each strand on the array with a fluorescent nucleotide and measuring the fluorescence intensity.[13,14] The decrease in intensity as the chain length increases is fitted to an exponential decay curve which then allows for the determination of a stepwise synthesis yield. In addition, this direct labelling and read-out method permits an optimization of the parameters involved in microarray synthesis, thereby enabling a relative control over array quality.[15] However, fluorescence provides at best a relative measure of sequence completion. The interpretation of the intensity can also be uncertain due to the sequence-dependence of fluorescence,[16] and it certainly cannot identify the source of error.

To be able to chemically separate the grown oligonucleotides from the glass slide and characterize the eluate using conventional analytical methods is an attractive idea, but the decisively small amount of DNA synthesized on-chip ($\sim$0.1–1 pmol mm^{-2})[17] requires the most sensitive detection techniques. In this context, radiolabelling of cleaved DNA followed by gel electrophoresis offers an overview of synthetic quality and it has been successfully applied to the monitoring of microarray synthesis defects, but like fluorescence provides primary information on the distribution of sequence lengths.[9,17,18] Mass spectrometry (MS) is another sensitive method which would provide final evidence of oligonucleotide identity but it has, to our knowledge, only been attempted on microarray surfaces suitable as matrices for MALDI-MS analyses.[19–21]

We therefore wished to develop a method that allows for MS characterization of microarrays fabricated on standard glass microscope slides. In addition to the identification of full-length products, MS would likely detect synthetic failures, degraded material and incompletely deprotected sequences; essential information for the development of new *in situ* chemistries. Indeed, we have recently embarked on the synthesis of RNA microarrays by photolithography[22,23] and the identification by MS of the synthetic RNA analytes is expected to help guide the technology to maturity. Our approach involved the incorporation of a base-labile ester functionality at the 3′-end of the oligonucleotide chain.[24] To do so, we used a custom-made NPPOC-protected dT phosphoramidite with a succinyl group attached to the 3′-OH function (cleavable dT, dT^{cleav}, Fig. 1a). Following published protocols,[25] this amidite was coupled for 1 min on silanized glass slides after the synthesis of a pentamer spacer, and the desired oligonucleotide sequence was then fabricated after NPPOC deprotection of the dT^{cleav} (Fig. 1b). To verify that dT^{cleav} coupled efficiently, we labelled the 5′-end of a dT$_{10}$ chain with a Cy3 dye. In parallel, dT decamers fabricated without dT^{cleav} were also fluorescently-labelled. Based on the difference

[a] *Department of Chemistry, McGill University, Montréal, Québec H3A 0B8, Canada. E-mail: masad.damha@mcgill.ca*
[b] *Institute of Inorganic Chemistry, University of Vienna, 1090 Vienna, Austria. E-mail: mark.somoza@univie.ac.at*

† Electronic supplementary information (ESI) available: Experimental procedures for microarray fabrication, deprotection and cleavage as well as LC-MS conditions and spectra of all array eluates. See DOI: 10.1039/c4cc05771f

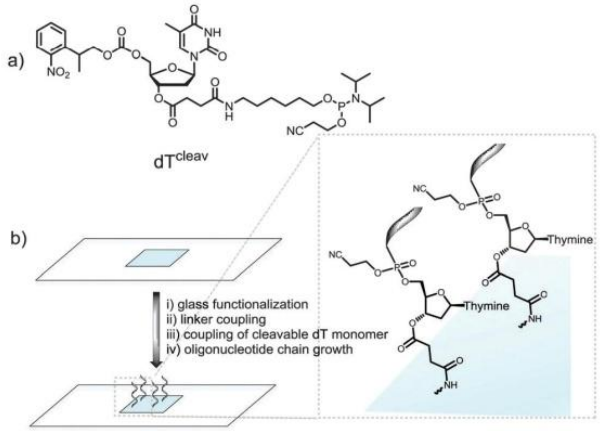


Fig. 1 (a) Chemical structure of the cleavable dT monomer; (b) schematic illustration of the synthetic steps involved in the fabrication of microarrays containing a cleavable dT unit. Glass functionalization is performed with a silanizing reagent. The linker is typically a dT or dC pentamer chain.

in fluorescence intensity between cleavable and non-cleavable sequences (Fig. S1a, ESI†), an 85% coupling yield was calculated for dT^{cleav}. Next, the same arrays were treated in concentrated ammonia at r.t. for 2 h and then scanned. The features where cleavable sequences were synthesized underwent a large drop in fluorescence intensity (Fig. S1b, ESI†), indicating that the ester function was correctly cleaved and release of the oligonucleotide in solution was almost complete.

We then attempted to collect the chemically-cleaved oligonucleotide. We chose to fabricate a simple dT_{13} model sequence according to the procedure depicted in Fig. 1a. After synthesis, the microarray was deprotected in a 1:1 mixture of ethylenediamine (EDA) and toluene (Fig. 2a), an alternative to the conventionally employed EDA/ethanol in DNA array deprotection.[8,26] After 2 h at r.t., the array was thoroughly washed with ACN, dried and the resulting DNA was collected from the surface by applying 100 μl of water (Fig. 2b).

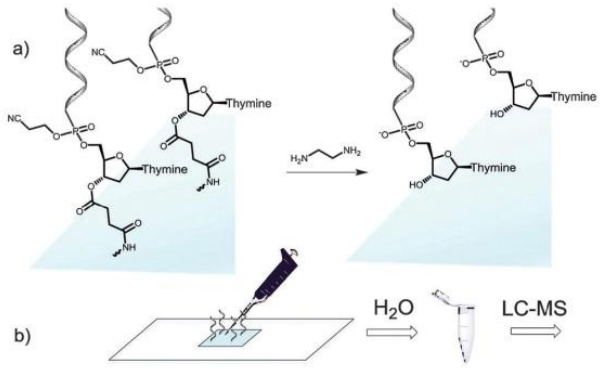


Fig. 2 Schematic illustration of the cleave-and-collect process of oligonucleotides synthesized on microarrays. (a) DNA oligonucleotides are first deprotected in EDA/toluene 1:1, 2 h, r.t. and the microarray is then washed with ACN (2 × 25 ml); (b) the DNA is then collected by pipetting 100 μl H_2O over the synthesis area. The microarray eluate is concentrated and analysed by LC-MS.

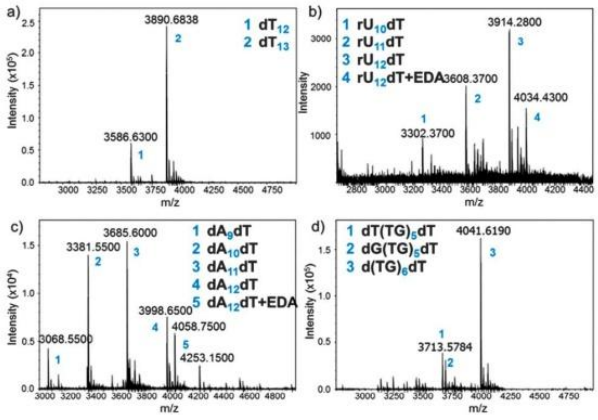


Fig. 3 MS spectra obtained after deprotection and cleave-and-collect for the following oligonucleotides: (a) dT_{13}; (b) $rU_{12}dT$; (c) $dA_{12}dT$; (d) $d(TG)_6dT$. Exact masses are shown. EDA: ethylenediamine. Numbers (blue) are referred to in the inset of each MS spectrum.

Quantification of the isolated chip eluate revealed that 20 pmol of material were obtained, consistent with the reported density of available hydroxyl groups on the silanized surface of the substrate.[17] Using a duplicating method developed earlier in our laboratory where two identical arrays are simultaneously fabricated,[27] a single automated run yielded up to 40 pmol of deprotected DNA which were subsequently analysed by liquid chromatography (LC)-electrospray ionization (ESI)-MS.

The MS trace of the cleaved dT_{13} is shown in Fig. 3a. The full-length product is detected as a 3′-OH species, demonstrating the correct cleavage at the 3′-ester functionality, together with a significant amount of a shortmer identified as dT_{12}. Since the capping step in the synthetic cycle was omitted, the $n-1$ oligonucleotides are the result of a single failed coupling. In the absence of capping, the oligonucleotide lengths follow a binomial distribution, which allows estimating the coupling yield based on the relative heights of the MS peaks. The relative peak height in Fig. 3a indicates a 98.3% coupling yield for NPPOC-dT; somewhat lower than values previously calculated by the fluorescence method.

Our cleavage method was then applied to the detection of poly dC (Fig. S11, ESI†) and poly dA (Fig. 3c) sequences. Interestingly, the amount of $n-1$, $n-2$ and $n-3$ species in crude poly dA samples exceeds those in poly dT and dC arrays. The full-length product, $dA_{12}dT$, is also present in the form of a noncovalent complex with EDA. Nucleobase deprotection is complete in both $dA_{12}dT$ and $dC_{12}dT$ cases since no trace of remaining phenoxyacetyl (Pac) or isobutyryl (iBu) groups was detected by MS. The characterization of oligonucleotide arrays was also applied to mixmers of two bases and, as shown in Fig. 3d and Fig. S13 (ESI†), MS resolution allows for the distinction between two different failure sequences.

Inspired by these results and by a previously reported procedure for the complete deprotection of RNA in EDA without facing degradation,[28] we wished to apply our method to RNA microarrays. A model $rU_{12}dT$ array was fabricated using NPPOC 2′-*O*-ALE rU amidites[22] and was then deprotected as

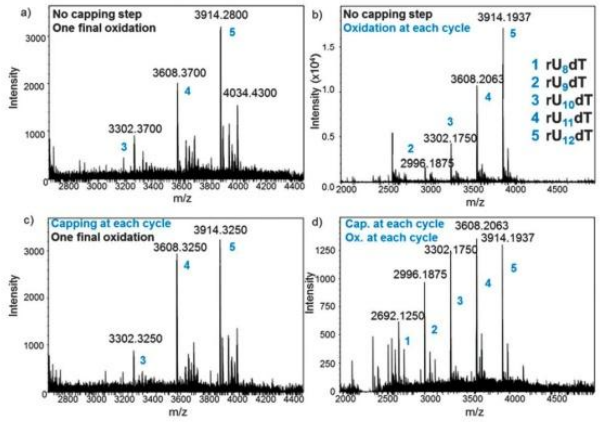


Fig. 4 MS spectra obtained after deprotection and cleave-and-collect for $rU_{12}dT$ microarrays under various fabrication conditions: (a) standard protocol without capping and oxidation; (b) an oxidation step is included; (c) a capping step is included; (d) both capping and oxidation steps are included. Exact masses are shown. Numbers 1–5 (blue) are all referred to in the MS spectrum (b).

follows: first, decyanoethylation was conducted in Et_3N/ACN 2 : 3 for 6 h at r.t. then ALE removal was performed in buffered hydrazine hydrate in pyridine/AcOH for 2 h at r.t. The intact succinyl ester was finally cleaved by treating the array with dry EDA/toluene for 2 h at r.t. The crude RNA was eluted from the surface by pipetting a small volume of sterilized water, concentrated, quantified (20 pmol per array) and injected on LC-MS. The MS spectrum is shown in Fig. 3b and the major peak corresponds to the full-length, 3′-OH $rU_{12}dT$, which is flanked by a minor peak at +60 Da resulting from a salt complex with EDA. This measurement offers, for the first time, a direct and concrete proof of correct *in situ* synthesis of RNA microarrays. Compared to dT_{13} in Fig. 3a, larger amounts of $n-1$ and $n-2$ species are also detected, which could be due to either failed couplings or to degradation products arising from cleavage at the internucleotidic phosphate. However, the presence of the *n*-mer as the main peak and the lack of 2′,3′-phosphorylated shortmers suggest that degradation is limited.

In an attempt to optimize the quality of *in situ* DNA and RNA microarray fabrication, we envisaged to modify a few key parameters in the design protocols and investigate their effect by MS. We performed this study on the dT_{13} and $rU_{12}dT$ models and considered four factors in the synthesis cycle: coupling time, the activator type, capping and oxidation steps. In DNA and RNA microarray synthesis by photolithography, the oxidation of the phosphite triester linkages can be conducted at the latest stage because deblocking the 5′-OH function does not require an acidic solution. The results as well as a representative panel are shown in Fig. 4 and Fig. S4–S10 (ESI†). Including an iodine/water-mediated oxidation or a capping step alone in the synthesis cycle seems to have little effect on array quality (compare Fig. 4b and c to the original array design in Fig. 4a), however when both steps are included, arrays of significantly lower quality were obtained (Fig. 4d). Next, the coupling time was examined and either shortened (from the standard 2 min to 1 min) or extended (5 min). In both DNA and RNA microarrays, shorter or longer coupling times resulted in arrays of poorer quality (Fig. S7, S8, S19 and S20, ESI†). Finally, the conventional 4,5-dicyanoimidazole activator was substituted with tetrazole derivatives, which afforded crude array eluates containing larger amounts of failure sequences (Fig. S9, S10 and S21, ESI†).

In summary, a reliable protocol for the deprotection and subsequent cleavage of DNA and RNA microarrays with EDA was developed using a 3′-succinylated dT phosphoramidite. The cleaved DNA microarrays or RNA microarrays are insoluble in the deprotection solution and remain on the glass surface,[28] where they can be collected with water and analysed by LC-ESI-MS. A few picomoles of crude microarray eluates are sufficient to provide a comprehensive overview of chip quality and to monitor the effect of modifying synthesis conditions. Radiolabelling or PCR amplification of the collected DNA/RNA is thus unnecessary. In addition, our approach allows for the first time the assessment of the fidelity of *in situ* RNA microarray synthesis and will have an important impact on the emergence of high-density complex RNA array technology.

The Natural Sciences and Engineering Research Council of Canada (discovery grant to M.J.D.), the Swiss National Science Foundation (Grant #PBBEP2_146174), The Austrian Science Fund (FWF P23797), ChemGenes Corporation and a McGill Fessenden Grant are gratefully acknowledged for financial support. We would also like to thank Dr Jeremy Lackey for fruitful discussions and FlexGen for the generous donation of the cleavable dT monomer.

Notes and references

1 R. B. Stoughton, *Annu. Rev. Biochem.*, 2005, **74**, 53–82.
2 H. M. Fathallah-Shaykh, *Arch. Neurol.*, 2005, **62**, 1669–1672.
3 C. Debouck and P. N. Goodfellow, *Nat. Genet.*, 1999, **21**, 48–50.
4 M. Andersen, S. Warrick and C. Adams, *Microarray Innovations*, CRC Press, 2009, pp. 215–241.
5 D. J. Duggan, M. Bittner, Y. Chen, P. Meltzer and J. M. Trent, *Nat. Genet.*, 1999, **21**, 10–14.
6 R. P. Auburn, D. P. Kreil, L. A. Meadows, B. Fischer, S. S. Matilla and S. Russell, *Trends Biotechnol.*, 2005, **23**, 374–379.
7 M. Dufva, *Biomol. Eng.*, 2005, **22**, 173–184.
8 S. Singh-Gasson, R. D. Green, Y. Yue, C. Nelson, F. Blattner, M. R. Sussman and F. Cerrina, *Nat. Biotechnol.*, 1999, **17**, 974–978.
9 E. M. LeProust, B. J. Peck, K. Spirin, H. B. McCuen, B. Moore, E. Namsaraev and M. H. Caruthers, *Nucleic Acids Res.*, 2010, **38**, 2522–2540.
10 S. L. Beaucage and M. H. Caruthers, *Tetrahedron Lett.*, 1981, **22**, 1859–1862.
11 L. J. McBride and M. H. Caruthers, *Tetrahedron Lett.*, 1983, **24**, 245–248.
12 P. Jaluria, K. Konstantopoulos, M. Betenbaugh and J. Shiloach, *Microb. Cell Fact.*, 2007, **6**, 4.
13 G. H. McGall, A. D. Barone, M. Diggelmann, S. P. A. Fodor, E. Gentalen and N. Ngo, *J. Am. Chem. Soc.*, 1997, **119**, 5081–5090.
14 M. C. Pirrung, L. Fallon and G. McGall, *J. Org. Chem.*, 1998, **63**, 241–246.
15 C. Agbavwe, C. Kim, D. Hong, K. Heinrich, T. Wang and M. M. Somoza, *J. Nanobiotechnol.*, 2011, **9**, 57.
16 C. Agbavwe and M. M. Somoza, *PLoS One*, 2011, **6**, e22177.
17 E. LeProust, H. Zhang, P. Yu, X. Zhou and X. Gao, *Nucleic Acids Res.*, 2001, **29**, 2171–2180.
18 K. E. Richmond, M.-H. Li, M. J. Rodesch, M. Patel, A. M. Lowe, C. Kim, L. L. Chu, N. Venkataramaian, S. F. Flickinger, J. Kaysen, P. J. Belshaw, M. R. Sussman and F. Cerrina, *Nucleic Acids Res.*, 2004, **32**, 5011–5018.

19 P. Kepper, R. Reinhardt, A. Dahl, H. Lehrach and S. Sauer, *Clin. Chem.*, 2006, **52**, 1303–1310.
20 C. F. W. Becker, R. Wacker, W. Bouschen, R. Seidel, B. Kolaric, P. Lang, H. Schroeder, O. Müller, C. M. Niemeyer, B. Spengler, R. S. Goody and M. Engelhard, *Angew. Chem., Int. Ed.*, 2005, **44**, 7635–7639.
21 L. Gogolin, H. Schroeder, A. Itzen, R. S. Goody, C. M. Niemeyer and C. F. W. Becker, *ChemBioChem*, 2013, **14**, 92–99.
22 J. G. Lackey, D. Mitra, M. M. Somoza, F. Cerrina and M. J. Damha, *J. Am. Chem. Soc.*, 2009, **131**, 8496–8502.
23 J. G. Lackey, M. M. Somoza, D. Mitra, F. Cerrina and M. J. Damha, *Chim. Oggi*, 2009, **27**, 30–33.
24 K. Maurer, D. Suciu and H. Gao, *US Pat.*, US20090280998 A1, 2009.
25 N. L. W. Franssen-van Hal, P. van der Putte, K. Hellmuth, S. Matysiak, N. Kretschy and M. M. Somoza, *Anal. Chem.*, 2013, **85**, 5950–5957.
26 E. F. Nuwaysir, W. Huang, T. J. Albert, J. Singh, K. Nuwaysir, A. Pitas, T. Richmond, T. Gorski, J. P. Berg, J. Ballin, M. McCormick, J. Norton, T. Pollock, T. Sumwalt, L. Butcher, D. Porter, M. Molla, C. Hall, F. Blattner, M. R. Sussman, R. L. Wallace, F. Cerrina and R. D. Green, *Genome Res.*, 2002, **12**, 1749–1755.
27 M. Sack, N. Kretschy, B. Rohm, V. Somoza and M. M. Somoza, *Anal. Chem.*, 2013, **85**, 8513–8517.
28 D. J. Dellinger, Z. Timár, J. Myerson, A. B. Sierzchala, J. Turner, F. Ferreira, Z. Kupihár, G. Dellinger, K. W. Hill, J. A. Powell, J. R. Sampson and M. H. Caruthers, *J. Am. Chem. Soc.*, 2011, **133**, 11540–11556.

2.4 Sequence-dependent fluorescence of Cy3- and Cy5-labeled double-stranded DNA

Nicole Kretschy, Matej Sack and Mark M. Somoza

This paper was submitted to the Journal Bioconjugate Chemistry in October 2015, peer-reviewed, revised and resubmitted.

Motivation

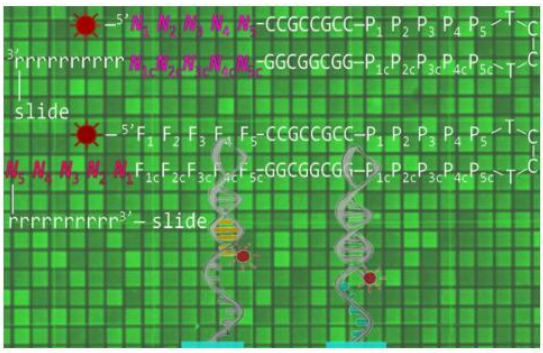
The motivation for the present paper was to determine the best and in contrast the worst fluorescent intensity of cyanine dyes for labeling nucleobase sequences synthesized by light-sensitive *in situ* microarray synthesis. Detection is the most important aspect for data extraction and in this particular case, detection of sequence-dependent fluorescence of nucleic acids also established by common analytical methods such as PCR, FISH, FRET and sequencing. We used Cy3 and Cy5 - the most common oligonucleotide cyanine dyes - to label the 5´ end of the oligonucleotide of all possible terminal DNA 5mers. *In situ* synthesis runs from 3´ to 5´. ssDNA and dsDNA were compared and it was shown that the sequence dependence of nucleobases is greater for dsDNA. Further data proves that for ssDNA as well as for dsDNA a directly adjacent guanine enhances the fluorescence than any other base. In contrast adenine is associated with the best possible fluorescence if not directly adjacent to the dye. The best possible fluorescence intensity however resulted with the GAAAA 5mer for both Cy3 and Cy5. The results obtained from these investigations can be most likely useful for enhanced signals in analytical tools such as TaqMan, PCR, FISH and next generation sequencing-by-synthesis.

Sequence-Dependent Fluorescence of Cy3- and Cy5-Labeled Double-Stranded DNA

Nicole Kretschy, Matej Sack and Mark M. Somoza

Institute of Inorganic Chemistry, Faculty of Chemistry, University of Vienna, Althanstraße 14 (UZA II), A-1090 Vienna, Austria

ABSTRACT: The fluorescent intensity of Cy3 and Cy5 dyes is strongly dependent on the nucleobase sequence of the labeled oligonucleotides. Sequence-dependent fluorescence may significantly influence the data obtained from many common experimental methods based on fluorescence detection of nucleic acids, such as sequencing, PCR, FRET and FISH. To quantify sequence dependent fluorescence, we have measured the fluorescence intensity of Cy3 and Cy5 bound to the 5' end of all 1024 possible double-stranded DNA 5mers. The fluorescence intensity was also determined for these dyes bound to the 5' end of fixed-sequence double-stranded DNA with a variable sequence 3' overhang adjacent to the dye. The labeled DNA oligonucleotides were made using light-directed, *in situ* microarray synthesis. The results indicate that the fluorescence intensity of both dyes is sensitive to all five bases or base pairs, that the sequence dependence is stronger for double- (vs. single-) stranded DNA, that the dyes are sensitive to both the adjacent dsDNA sequence and the 3'-ssDNA overhang. Purine-rich sequences result in higher fluorescence. The results can be used to estimate measurement error in experiments with fluorescent-labeled DNA, as well as to optimize the fluorescent signal by considering the nucleobase environment of the labeling cyanine dye.



INTRODUCTION

The fluorescence of molecules is always sensitive to environmental conditions, although the magnitude of changes in the fluorescence intensity of any particular fluorophore depends on its specific modes of interaction with its environment.[1] Fluorescent molecules can be used as molecular environmental probes by selecting dyes with strong responses to, for example, pH,[2] viscosity,[3] polarizability,[4] elasticity,[5] and polarity,[6] however, in applications where the fluorescent intensity is to serve as a proxy for the abundance of the labeled molecule, environmental sensitivity is a liability that can result in reduced measurement accuracy.[7] The cyanine dyes Cy3 and Cy5 are among the most widely used and versatile[8] oligonucleotide labels in, e.g., microarray experiments, fluorescent *in situ* hybridization (FISH), real-time PCR (RT-PCR), and FRET studies[9,10] and are considered to be relatively environmentally insensitive.[11] However, Cy3 and Cy5 consist of two indole rings connected by three or five carbon polymethine bridges which can undergo *cis-trans* isomerization from the first excited singlet state which competes with fluorescence.[12-15] In viscous or restrictive environments, or with conformationally-locked dye variants, the rate of isomerization is reduced or eliminated and the dyes are more fluorescent.[16] When Cy3 and Cy5 are tethered to the end of double-stranded DNA they assume a planar capping configuration similar to that of an additional base pair,[17,18] which inhibits isomerization and increases their fluorescence quantum yield and lifetime.[15] At least in the case of Cy3, the range of motions available is not fully restricted when attached to either single- and double-stranded DNA, with time-resolved fluorescence anisotropy measurements indicating decay components corresponding to rotation with DNA as well as relative to DNA. Cy3 tethered to double-stranded DNA has a lower isomerization rate resulting in a higher fluorescence quantum yield and longer lifetime relative to single-stranded DNA.[15]

Recent experiments have shown that both Cy3 and Cy5 are also quite sensitive to the particular nucleobase sequence of the ssDNA oligonucleotide to which they are attached,[19,20] with the fluorescence intensity varying by a factor of about two between the brightest and the darkest labeled oligonucleotide in the case of Cy3, and a factor of about three in the case of Cy5. The variation in fluorescence intensity for ssDNA is

strongly correlated with purine content, with purine-rich sequences associated with high intensity, and high pyrimidine content, particularly cytosine, with low intensity.[19]

The magnitude of the sequence-dependent fluorescence is large enough to affect the accuracy of experimental data derived from Cy3- and Cy5-labeled single-stranded DNA, but there is currently no data available on sequence-dependent effects in double-stranded DNA. In experimental methods based on labeled oligonucleotides, fluorescence is recorded either from the double-stranded hybrid (e.g. Sanger and next-generation sequencing, and molecular beacons[21]), or from the unhybridized strand alone (e.g. hydrolyzed labeled TaqMan probe fragments[22]). High-throughput DNA sequencing-by-synthesis is likely to be particularly vulnerable to sequence-dependent fluorescence because all short nucleobase sequences will be repeatedly encountered, and detection failures (deletion errors) from sequences highly unfavorable to fluorescence would be systematic and therefore not easily detectable with re-sequencing. Furthermore, the optical systems of sequencers need to balance dynamic range of detection with throughput, making their throughput sensitive to dyes with significant variations in fluorescence.[23] Even though our fluorescence data are obtained on microarrays, most genomics microarray data is fairly insensitive to sequence-dependent fluorescence because the labeling is typically based on reverse transcription using labeled random primers or other quasi-random methods.[24] Nevertheless, gene-specific fluorescence intensity effects, due to differences in the relative abundance of nucleobases in particular genes, have been detected.[25]

Since both Cy3- and Cy5-labeled single- and double-stranded oligonucleotides are commonly used, we present here comprehensive results for double-stranded DNA to complement and strengthen previous results for Cy3 and Cy5 5′-labeled single-stranded DNA.[19] Two types of sequence-dependent dye-dsDNA interactions, as illustrated in Figure 1, have been measured: relative intensity of the dyes at the 5′ end of each of the 1024 possible double-stranded DNA 5mers (Fig. 1B), and relative intensity of the dyes bound to the 5′ end of a fixed-sequence double helix, but with a variable 5mer sequence 3′ overhang adjacent to the dyes (Fig. 1C). The sequence-dependent contribution of the overhang is relevant since in many experimental contexts, such as PCR and FISH, a short 5′-labeled oligonucleotide is use to quantify the presence of much longer DNA or RNA molecules. Detailed data on the sequence-dependent fluorescence of cyanine dyes on single-stranded DNA (Fig. 1A) has been previously reported for Cy3, Cy5, Dy547 and Dy647;[19,26] this ssDNA data showed that over the range of all possible 5mers, the intensity of Cy3 varied by about a factor of two, and in the case of Cy5, by a factor of about three. There was also a clear pattern to the data: the fluorescence follows, to a good approximation, the cumulative distribution function of a normal distribution, with purine-rich sequences resulting in high intensities and pyrimidine-rich sequences resulting in low intensities. In addition, 5′ guanines promote higher fluorescence much more so than 5′ adenosines, and 5′ cytosines result in much lower fluorescence in comparison with 5′ thymidines. Here we will show that broadly similar trends also hold true for double-stranded DNA.

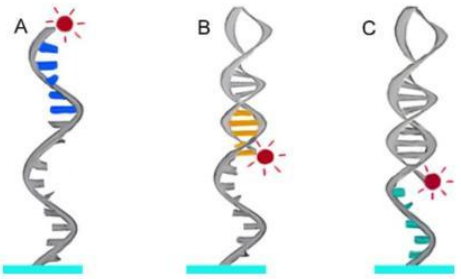


Figure 1. Interaction modes of dyes (purple) on DNA. A. 5′ dye with adjacent nucleobases (blue) in ssDNA. B. 5′ dye with base-paired nucleobases (orange) in dsDNA. C. 5′ dye with nucleobases of ssDNA (green) adjacent to a terminal dye on dsDNA.

EXPERIMENTAL PROCEDURES

Microarray synthesis

Glass slides (Schott Nexterion D, cleanroom-cleaned) were functionalized with N-(3-triethoxysilylpropyl)-4-hydroxybutyramide (Gelest SIT8189.5). The slides were loaded in a stainless-steel rack, placed in a plastic container, and covered with 500 ml of a solution consisting of 10 g of the silane in a 95:5 (v/v) ethanol:water plus 1 ml acetic acid. The slides were gently agitated for 4 hrs at room temperature and then washed twice for 20 minutes each with the above solution without the silane. The slides were drained, blown dry with argon and cured in a preheated vacuum oven (120°C) overnight and stored in a desiccator cabinet.

For the synthesis of terminally labeled oligonucleotides on microarrays we used the technique of maskless array synthesis (MAS).[27,28] MAS was developed for *in situ* synthesis of high-density DNA microarrays and consists of an optical system and a chemical delivery system. The optical system consists of a digital micromirror device (DMD), an array of individually tiltable mirrors, which direct ultraviolet light from a mercury lamp to the corresponding feature on the microarray via 1:1 imaging optics. Microarray layout and oligonucleotide sequences are determined by selective removal of the photocleavable protecting groups on the phosphoramidites at the 5′ termini of the oligonucleotides.

A computer synchronizes the light exposures pattern with solvent reagent delivery to the synthesis surface. Chemical system consists of a slightly modified Perspective Biosystems Expedite 8909 synthesizer. Oligonucleotide synthesis chemistry is similar to that used in conventional solid-phase synthesis. The standard acid-labile 5′-OH protecting group of the phosphoramidites is replaced with the photocleavable nitrophenylpropyloxycarbonyl (NPPOC) group.[29] Upon absorption of light near 365nm, the I-line of mercury, the NPPOC group comes off, leaving a free hydroxyl group that is able to react with an activated phosphoramidite in the next coupling cycle. An exposure solvent consisting of 1% (m/v) imidazole in dimethyl sulfoxide (DMSO) is needed during ultraviolet exposure to promote the cleavage of the NPPOC group.[28] The coupling reactions were performed with 30mM NPPOC phosphoramidite monomers and 0.25M dicyanoimidazole (both from SAFC) for 60 seconds. In the case of the Cy3 and Cy5 phosphoramidites (GE Healthcare 28-9172-98 and Glen Research 10-5915-95), Figure 2, the coupling reaction time was extended to 10 minutes at a monomer concentration of 15mM. Acetylation with a 1:1 mix of tertbutylphenoxyacetyl acetic anhydride in tetrahydrofuran (Cap A) and 10% N-methylimidazole in tetrahydrofuran/pyridine (8:1) (Cap B) after each coupling reaction, was used to ensure that only correctly synthesized sequences receive the fluorescent label.

After microarray synthesis the substrate was vigorously washed for 2 hrs with acetonitrile in a 50 ml Falcon tube to remove uncoupled Cy3 or Cy5 phosphoramidites, which tend to adhere non-specifically to the glass surface. The base and phosphate protecting groups were removed by immersing the glass slide into 1:1 (v/v) ethylenediamine in ethanol for two hours at room temperature. Following deprotection, the microarrays were washed twice with distilled water and dried with argon.

Microarray design

In principle, the resolution of the digital micromirror device, 768 x 1024, allows for simultaneous measurement of all possible *n*-mers up to *n*=9 (262 144), but in these experiments, only permutations of 5mers were included in order to include multiple replicates and to dedicate more microarray surface area to each sequence and therefore to achieve a good signal-to-noise ratio. The 1024 sequences were laid out in a 25 in 36 pattern, that is, each "feature" (contiguous area were a single sequence is synthesized) on the microarray corresponded to a 5 by 5 block of mirrors surrounded by a one-mirror-sized margin where no DNA was synthesized. Each of the 1024 single-sequence features was replicated 20 times on each microarray in the case of the double-stranded experiments (Fig. 1B), and 10 times in the case of in the case of the double-stranded DNA with single-stranded overhang experiments (Fig. 1C).

Double-stranded DNA annealing

To promote hairpin-loop formation and self-hybridization, after deprotection the array was incubated in 40 ml PBS buffer (0.65 M Na^+, pH 7.4) starting at 50°C and cooled to room temperature over 30 min. Then it was washed with final wash buffer for a few seconds and dried with a microarray centrifuge. Successful hairpin loop formation was then verified by hybridization of a Cy3-labeled oligonucleotide (5′-Cy3-GGC GGC GGG TTC A-3′) to two unlabeled complementary sequences on the array: (1) a sequence (TGA ACC CGC CGC CGT CCA TCCT TGG ACG GCG GCG GGT TCA) that self-hybridized via hairpin-loop formation in the previous step and is therefore blocked from hybridization with the added oligonucleotide, and (2) a sequence (TGA ACC CGC CGC C) that cannot self-

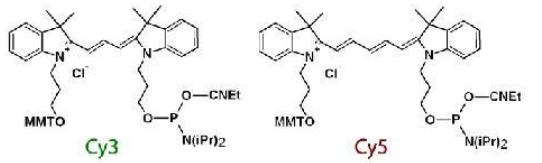


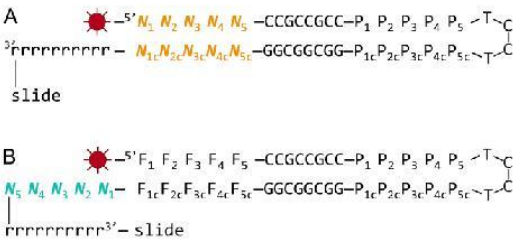
Figure 2. Molecular structures of the Cy3 and Cy5 cyanine dye phosphoramidites used in this study. After the end of the synthesis and the chemical deprotection step, the dyes are linked to the 5′ DNA nucleoside via a phosphodiester bond.

hybridize but is fully complementarity with the added labeled sequence.

Sequence design

Three principle considerations were applied to the sequence design: (1) The double-stranded sequences should all have equal melting temperatures since they must all form duplexes equally under the single hybridization condition of the microarray, (2) the melting temperature should be relatively high in order to ensure stable duplex formation, and (3) the surface density of labeled oligonucleotides should be constant for all experimental oligonucleotides on the microarray so that fluorescence intensity differences between them can be attributed to sequence-dependent effects. To meet these design principles the double-stranded oligonucleotides have the design illustrated in Scheme 1.

The sequences contain self-complementary segments to allow for duplex formation. The central TCCT sequence is known to bend easily to promote hairpin loop formation.[30] The N_i represents the 5mer experimental nucleobases that base pair with the complementary N_{ic}. On the 3′ side of the N_i is the fixed sequence CCGCCGCC which hybridizes with the GGCGGCGG sequence on the opposite side of the hairpin. This GC-rich stretch is used to increase the melting temperature. The $P_1P_2P_3P_4P_5$ sequence is derived from the experimental 5mer sequence $N_1N_2N_3N_4N_5$ using non-identity, non-complementarity logic: for all i, if N_i = dA then P_i = dC; or if N_i = dC then P_i = dT; or if N_i = dG then P_i = dA; or if N_i = dT then P_i = dG. These strands hybridize with their complementary sequences $P_{5c}P_{4c}P_{3c}P_{2c}P_{1c}$. The P_i and P_{ic} sequences have a double function: (1) they equilibrate the base composition in order to assure equal number density of all experimental sequences on the array, and (2) they increase and homogenize the melting temperatures (to T_m = 63°C, salt adjusted, 50 mM Na^+) by giving all the complementary DNA sequences on the array exactly five of each nucleobases (plus the fixed GC sequences) while retaining self-complementarity. The sequences are separated from the glass substrate with a random linker 10mer sequence synthesized from an equimolar mix of the four DNA phosphoramidites. The random linker replaces the traditional poly(dT), and linker to avoid the potential bias of any particular interaction of the dye and a dT homopolymer. An alternative perspective is that the dye is will interact with both the double-stranded and single-stranded segments, but the interaction with the single-stranded segment will be the average of all possible sequences. In the seconds set of experiments, the single-stranded sequence is permuted. The results of both data sets can be used to estimate the relative contributions, to dye intensity variation, of the single vs. double stranded segments.



Scheme 1. Sequence design for the 5′-dye self-hybridizing DNA strands. Sequence (A) is used to measure the interaction of the dyes with dsDNA and sequence (B) is used to measure the interactions of the dyes with the ssDNA overhang of dsDNA.

With these rules, all of the sequences (excluding the linker) have exactly 5 adenosines, 15 cytidines, 13 guanosines, and 7 thymidines. Since the coupling efficiency of each of the four DNA phosphoramidites can be different and can vary with time and by batch, equal numbers of each base in each of the sequences assures equal representation of the experimental oligonucleotides. This sequence design, in conjunction with acetic anhydride capping after the coupling reactions ensures equal density and melting temperature and that only accurately synthesized sequences receive the final coupling with the Cy3 or Cy5 phosphoramidite. An alternative approach, to use simpler sequences and then adjust the data for the measured coupling efficiencies is less reliable since the coupling efficiencies of the phosphoramidites used in maskless array synthesis are measured with fluorescent dye terminal labeling experiments,[31-34] which limits their accuracy due to the sequence-dependent fluorescence intensity of single-stranded DNA.[19]

The second set of experiments, with the dyes attached to fixed-sequence double-stranded DNA and a variable single stranded overhang has a similar design (Scheme 1B). Here, the permuted overhang sequence $N_1N_2N_3N_4N_5$ is added at the 3′ end to put it adjacent to the 5′ fluorescent label. The F_i and F_{ic} are complementary but are no longer permuted; $N_1N_2N_3N_4N_5$ is either GAAAA or CGTGG. GAAAA and CGTGG were chosen from the initial double-

stranded experiments as sequences resulting in high and low fluorescence intensity, respectively, for both Cy3 and Cy5.

In order to allow direct comparisons between the relative fluorescence intensity of the dyes on single- vs. double-stranded DNA, each dsDNA microarray design included sequences that cannot self-hybridize to form dsDNA, but have a very similar overall sequence design and base composition. Since most of the microarray features were needed for the dsDNA permutations, only a sampling of 57 labeled ssDNA permutations were included. These sequences were chosen to be representative of the range of expected fluorescence intensities for ssDNA found in previous experiments.[19] To prevent the self-hybridization of these sequences, the $N_{5c}N_{4c}N_{3c}N_{2c}N_{1c}$ segment was inverted to $N_{1c}N_{2c}N_{3c}N_{4c}N_{5c}$, the $P_{5c}P_{4c}P_{3c}P_{2c}P_{1c}$ segment was inverted to $P_{1c}P_{2c}P_{3c}P_{4c}P_{5c}$, palindromic N_i 5mers were avoided, and the segment GGCGGCGG was reordered to GCGGCGGG.

Data extraction and analysis

Fluorescent images of the microarrays were obtained using a GenePix 4100A scanner with resolution of 5 µM and with detector voltages set to give similar intensity ranges for both Cy3 and Cy5, and no saturated pixels, 350 and 450 volts, respectively. Dye fluorescence was excited using 532 nm and 635 nm solid-state lasers for Cy3 and Cy5, respectively. Fluorescence was collected through 550—600 nm and 655—695 nm bandpass filters for Cy3 and Cy5, respectively. Fluorescence light was collected using a 0.68 NA objective lens with a focal length of 3.1 mm. Microarray scanners are designed to provide intensity values that are highly consistent across the scanned surface. This allows highly reliable relative fluorescence comparisons between microarray features. The presence of the microarray surface, a lossless glass-air dielectric interface, close to the fluorophores does not influence the relative emission intensity or wavelength.[35]

The fluorescence intensity data was extracted from the scan image with NimbleScan v2.1 software from NimbleGen and further processed in Excel. For each microarray, fluorescence intensity values were calculated as the average of the replicates of each sequence, which were randomly located on each microarray. For the double-stranded experiment, there were 20 sequence replicates per array. For the overhang experiment there were 10 replicates per array because of the inclusion of two experimental sets, one with double-stranded sequence which strongly promotes fluorescence (dye-GAAAA) and one with a double-stranded sequence resulting in weak fluorescence (dye-CGTGG). Error was calculated as the standard error of the mean. The consensus sequence figures were generated by ranking the 1024 sequences by fluorescence intensity and then dividing the sequences into eight bins spanning equal ranges of intensity. Consensus logos for the sequences in each of these octiles of fluorescence intensity were generated using Weblogo (http://weblogo.berkeley.edu/).[36] Each of the eight consensus sequence logos per fluorescent label represents one eighth of the intensity range and are arranged together left to right in order of decreasing intensity to compactly depict the relationship between sequence and fluorescence for the entire dataset.

RESULTS AND DISCUSSION

The results for the sequence-dependent fluorescence of cyanine dyes have been highly consistent, with the adjacent purine bases promoting fluorescence relative to pyrimidine bases in single-stranded DNA,[19,26] and with the results presented here in double-stranded DNA. In addition, for both ssDNA and dsDNA, a guanine immediately adjacent to the dye consistently results in the highest fluorescence, but in the more distal positions, adenine, rather than guanine typically results in higher fluorescence. Of the pyrimidines, cytosine, rather than thymine, is most strongly associated with low fluorescence.

Cy3 and Cy5 dsDNA interactions

Figure 3 summarizes the results for both the 5' Cy3 and Cy5 terminal labeling experiments on dsDNA. These data correspond to the case where the random linker is used and the permuted nucleobases form a double strand (Scheme 1A). Here, the dye interactions with the single-stranded segment are present, but the data will reflect the average over all possible sequences. As was the case with the data from Cy3 and Cy5 labeled ssDNA,[19] the overall range of florescence intensity is about a factor of two for Cy3 and a factor of three for Cy5 (Fig. 3A). In order to be able to compare the fluorescence intensity data for dsDNA with ssDNA, the array design included reference ssDNA sequences. These sequences have a very similar design, but with bases rearranged to prevent hybridization. Figure 3A shows that both Cy3 and Cy5 on dsDNA have a somewhat extended range of fluorescence intensity in comparison to Cy3 and Cy5 on ssDNA (horizontal lines). Most of the additional range of intensity is on the lower edge of intensity, i.e., the sequences resulting in the highest fluorescence result in similar intensity for both ssDNA and dsDNA.

The fluorescence intensity of intensity of most, or perhaps all dyes, is dependent on the nucleobase environment. In many cases the mechanism is a photoinduced charge transfer between the bases and the

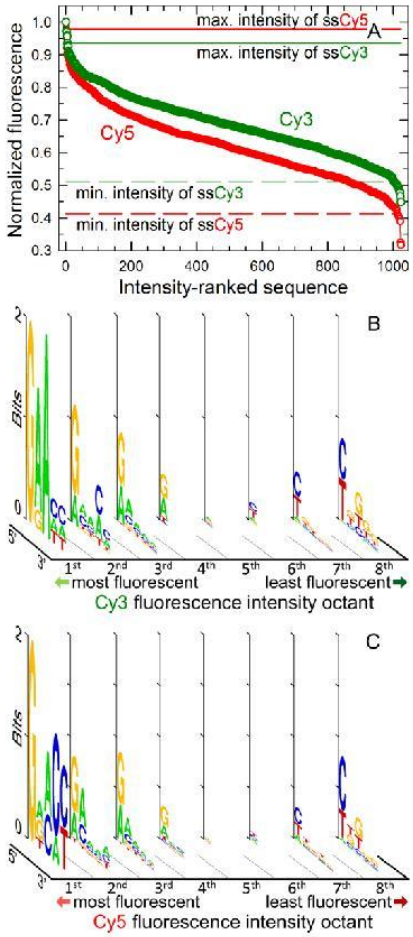


Figure 3. Double-stranded DNA labeling with Cy3 and Cy5 (Fig. 1B). (A) Relative fluorescence intensity of Cy3 and Cy5 end-labeled 5mers, ranked from most to least intense. The intensity falls by 55% for Cy3 and almost 70% for Cy5. The horizontal lines show the fluorescence intensity of single-stranded reference sequences on the same arrays. Fluorescence intensity consensus sequences of all 1024 dsDNA 5mers 5'-end-labeled using (B) Cy3 and (C) Cy5. The fluorescent range was equally divided into eight bins of equal intensity ranges, and the consensus sequence for all the 5mers is plotted for each such octile.

dye (fluorescein[37], coumarin[38], rhodamine[39] and pyrene[40]), in which case the quenching efficiency is determined by proximity and base redox potential, dG<dA<dC<dT, when the bases are reduced, or the reverse order when oxidized.[38] Ethydium bromide, another well-known dsDNA fluorescence label undergoes quenching via proton transfer to the solvent; intercalation enhances fluorescence by reducing solvent exposure.[41] In the case of the cyanine dyes, however, charge transfer is not thermodynamically favored.[42,43] Instead, the intensity of cyanine dyes conjugated with DNA is attributed to the modulation of the rotational isomerization barrier in the excited state.[12-14]

NMR data indicate that Cy3 and Cy5, 5'-linked to dsDNA, are positioned at the end of the double helix similarly in a capping configuration, in a manner similar to that of a base pair.[17,18] This arrangement should restrict the rate of *cis-trans* isomerization of the dyes, increasing fluorescence relative to the free dye. However, relative to the same dyes bound to the end of ssDNA, differences in the rate *cis-trans* isomerization are less clear since the dyes stack with the terminal base in both cases. Simulations and experiments indicate that the quantum yield of Cy3 is higher on ssDNA vs. dsDNA, and that on dsDNA the strength of the stacking interaction depends on the identity of the terminal basepair[15,44,45]. Our experiments indicate that the fluorescence of Cy3 and Cy5 is somewhat greater on dsDNA; however, the differences between our results and previously published results[15] , which show a two-fold greater fluorescence of Cy3 on ssDNA, may be due to the particular choice of cyanine dye. In particular, we conjugate with DNA using the Cy3 and Cy5 phosphoramidites (Fig. 2), rather than the sulfonated versions of these dyes, use by Sandborn et al.[15], and which are more commonly used for protein labeling. The sulfonates increase the hydrophilicity of the dyes, which could affect the strength of the stacking interactions with the nucleobases. We have previously measured the intensity of sulfonated Cy3 and Cy5 on DNA, and found a very strong pattern of sequence-specific fluorescence distinct from that of the unsulfonated dyes.[19] The sequence-specificity of sulfonated Cy3, in particular, is such that specific experimental sequence use by Sandborn et al. may have biased their results towards observing greater fluorescence for single- vs. double-stranded DNA.

[15,44,45] In order to visualize the relationship between the nucleobases sequence and the fluorescence intensity, the consensus sequences for each octant of intensity are plotted in Figure 3B and 3C for Cy3 and Cy5, respectively. These data are quite similar to those obtained with the same dyes on ssDNA[19]. The most apparent differences in the dsDNA data are that cytosine is less prominent in the weakly fluorescent sequences, and that cytosine is more prominent in the distal positions of the strongly fluorescent sequences, particularly for Cy5. If as previous studies have indicated, the fluorescence intensity of cyanine dyes is greater on ssDNA, there might be bias in

the consensus toward adenine- and thymine-rich sequences, which will tend to destabilize the double helix near the dyes, resulting in a higher locally single-stranded ("frayed" ends) population of DNA. In relationship to our previous data of Cy3 and Cy5 on ssDNA, this trend is not apparent. In the dsDNA data (Figure 3), the melting temperature of the consensus sequences for the most fluorescent intensity octants are higher than those in the equivalent octants in the ssDNA data for both Cy3 and Cy5 due to the increased population of cytosines.

Cy3 and Cy5 overhang interactions

In the results described above, the dyes must also be interacting with the immediately adjacent ssDNA overhang segment as illustrated in Figures 1C and Scheme 1B. In order to estimate how this ssDNA modulates the fluorescence, the random nucleobases linker was replaced with segments representing all possible 5mers. To avoid having too many overall permutations, only two dsDNA sequences were used, one associated with strong fluorescence (GAAAA) and one with weak fluorescence (CGTGG). About 10 replicates of each of the 2048 resulting sequences fit on a single microarray, allowing accurate relative intensity comparisons between sequences. In the dsDNA data shown in Figure 3, the sequence GAAAA resulted in the 33rd and 100th brightest fluorescence for Cy3 and Cy5, respectively. The sequence CGTGG resulted in the 1008th and 898th brightest fluorescence for Cy3 and Cy5, respectively. The results from the overhang experiment, using Cy3 as the dye, are shown in Figure 4. In Figure 4A, the intensity of each sequence has been normalized to that of the most intense sequence, which, as expected, belongs to the Cy3-dsGAAAA set. Most of the sequences with Cy3-dsCGTGG are darker than any of those with GAAAA. Figure 4A clearly shows that the intensity of the dye is similarly determined by both the dsDNA segment and the adjacent ssDNA segment since the intensity difference between the two curves is similar to the range in intensities within each curve.

The relationship between the nucleobases sequence of the permuted overhang and the fluorescence intensity is shown in using consensus logos in Figure 4A and 4B, for Cy3-dsGAAAA and Cy3-dsCGTGG, respectively. The consensus sequences show a similar pattern to those of the ssDNA data and the dsDNA data with the random overhang; the most fluorescent signal results from

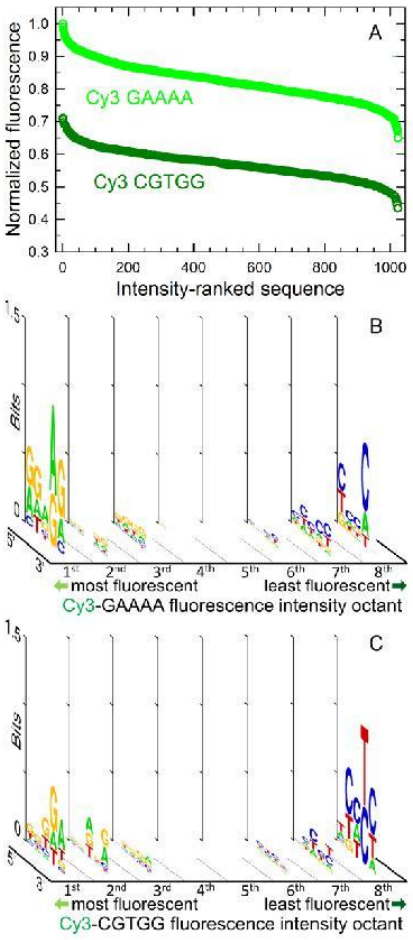


Figure 4. 5'-Cy3-dsDNA with a permuted 3' overhang (Fig. 1C). The dsDNA strand to which the Cy3 is attached has one of two sequences: GAAAA (bright) or CGTGG (dark). (A) Relative fluorescence of Cy3-GAAAA and Cy3-CGTGG, ranked from most to least intense over the range of all ssDNA 3' overhang 5mers. The intensity falls by ~35% for both Cy3-GAAAA and Cy3-CGTGG. Fluorescence intensity consensus sequences of all 1024 5mers on the 3'-overhang of (B) Cy3-dsGAAAA and (C) Cy3-dsCGTGG. The fluorescent was equally divided into eight bins of equal intensity ranges. The consensus sequence is plotted for each bin.

sequences with high purine content and the least florescence signal results from sequences with high pyrimidine content, particularly cytosine. Two additional trends are clearly visible in the consensus sequence data. First, the information content (bits) for each position is typically lower than that for the data with the random overhang. This is because in the present case, there is no single dominant base at any position, e.g., both purines are approximately equally probable in the most florescent sequences. This trend can also be anticipated by the

shape of the intensity curves in Figure 4A, which, spanning a lower range of intensity in comparison to that in Figure 3 for the same number of permuted sequences, indicate a reduced sequence dependence of fluorescence. Second, the more distal bases are more prominent in the consensus sequences, which suggests that the dye is interacting more strongly with these more distal bases. One possibility is that the presence of dye on the terminus of the double-stranded segment may tend to displace the more proximal overhang bases to conformations where they cannot affect the *cis-trans* isomerization rate. This is consistent with NMR data indicating that Cy3 occupies much of the available stacking space at the end of dsDNA.[18]

Data for Cy5 on double-stranded DNA with a permuted overhang is shown in Figure 5. These data were collected using the same methods and the same microarray design, only using Cy5 instead of Cy3. As with Cy3, the intensity difference between the two curves Figure 5A is similar to the range in intensities within each curve, clearly showing that the intensity of Cy5 is similarly determined by both the dsDNA segment and the adjacent ssDNA overhang segment. Unlike in the case of Cy3, all of the Cy5-dsCGTGG sequences are darker that the darkest of the Cy5-dsGAAAA sequences. The specific sequence Cy5-dsGAAAA in the random linker dataset resulted in an intensity of 0.8 relative to that of Cy5-dsGAACC, the most intense, suggesting that the gap between the curves in Figure 5A could be significantly increased by using GAACC as the fixed double stranded sequence. Although the two curves in Figure 5A appear to have different shapes, this is due only to the large fluorescence intensity difference between them. Independently normalizing the Cy5-dsCGTGG data would cause it to overlap very closely with the Cy5-dsGAAAA data, indicating that both double stranded sequences modulate the interaction of the dye with the overhang bases to a similar extent.

The relationship between the nucleobases sequence of the permuted overhang and the fluorescence intensity is shown in using consensus logos in Figure 5B for Cy5-dsGAAAA and in Figure 5C for Cy5-dsCGTGG. Like in the case of Cy3, the highest fluorescence is strongly associated with purines while the lowest fluorescence is strongly associated with pyrimidines. Between the purines, guanine is clearly more relevant than adenine in promoting fluorescence. Cytosine is also much more common than thymine in the sequences associated with low fluorescence. As a result of the dominance of these

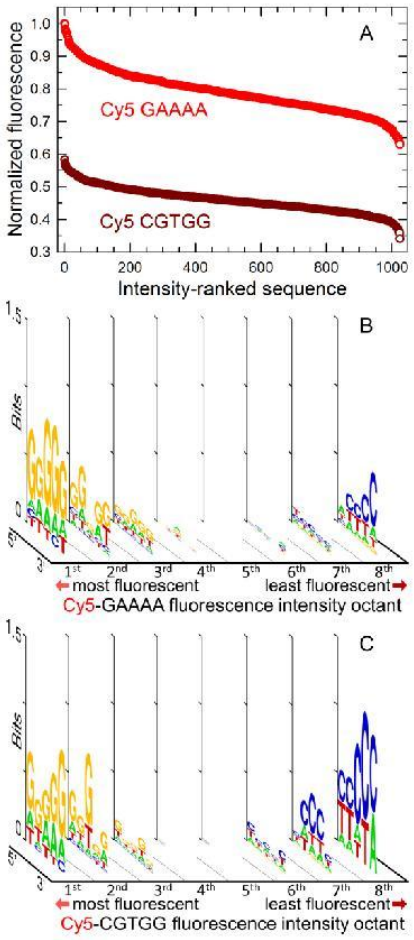


Figure 5. 5′-Cy5-dsDNA with a permuted 3′ overhang (Fig. 1C). The dsDNA strand to which the Cy5 is attached has one of two sequences: GAAAA (bright) or CGTGG (dark). (A) Relative fluorescence of Cy5-GAAAA and Cy5-CGTGG, ranked from most to least intense over the range of all ssDNA 3′ overhang 5mers. The intensity falls by ~40% for both Cy5-GAAAA and Cy5-CGTGG. Fluorescence intensity consensus sequences of all 1024 5mers on the 3′-overhang of (B) Cy5-dsGAAAA and (C) Cy5-dsCGTGG. The fluorescent was equally divided into eight bins of equal intensity ranges. The consensus sequence is plotted for each bin.

two bases, the information content of the consensus sequences is higher in the case of Cy5. The trend observed for Cy3, that the dye interacts more strongly with more distal bases is also the case with Cy5.

For both Cy3 and Cy5, sequences resulting in the lowest intensity among the dye-dsCGTGG subset have intensities similar to the darkest from the datasets with the random overhang in Figure 3. Since the use of a random nucleobases linker should be equivalent to averaging over all linker base permutations, the

expectation was that the minimum fluorescence measured in the permuted overhang experiments would be significantly lower than those measured using random overhang. One possibility is that the range over which the fluorescence intensity of Cy-dyes can be modulated via interactions with DNA is restricted. This seems reasonable since the total range over which the fluorescence quantum yield of Cy3 can be lowered by restricting the rate of *cis-trans* isomerization is about a factor of eight at room temperature, and Cy3 on DNA appears to be limited to the lower half of this range.[15] Nevertheless, some additional range of fluorescence intensity could likely be measured in permuted sequences longer than 5mers. In most of the consensus sequences in Figures 3, 4, and 5, there is information content in the 5th base, the most distal; indicating that this base also participates in modulating the intensity, so a 6th or 7th base is also likely to contribute to the modulation of fluorescence. Another perspective in this regard is that the shapes of the curves in Figures 3A, 4A, and 5A can be interpreted as cumulative distribution functions where the variable is the normalized relative fluorescence. To a good approximation, the fluorescence intensities of Cy3 and Cy5 on random DNA sequences have probability mass functions approximating those of binomial distributions, where the two results are purine or pyrimidine.[19] Most random 5mer sequences will contain a mix of purines and pyrimidines, which will result in intermediate fluorescence in the central region of the distribution. A few sequences will contain mostly or exclusively purines or pyrimidines, resulting in, respectively, fluorescence at the high and low tails of the intensity distribution. Increasing the permuted sequence length (Bernoulli trials) should result in a few sequences in the tails of the distribution that extend the range of fluorescence.

These results are consistent with previous experiments on the fluorescence of Cy3 and Cy5, which have also shown similar patterns of nucleobase dependency. Studies on the interactions of Cy3 with nucleoside monophosphates solutions have found a pattern of nucleobase-specific enhancement of fluorescence, dG>dA>dT>dC> no DNA.[46] Experiments on an intercalating cyanine dye derived from thiazole orange demonstrated a strong association of fluorescence with purine DNA homopolymers but not with pyrimidine homopolymers; the resulting fluorescence relative intensities followed the pattern dG>dA≫dC>dT> no DNA (100, 39, 2.3, 1.8 and 0.5, respectively).[47] Computer simulations in this study also indicated that the dyes associate poorly with poly(dC) and poly(dT), while binding strongly to poly(dG) and poly(dA). All these results fit well with the model that π-π interactions between cyanine dyes and nucleobases decrease the cis-trans isomerization rate. Purines, with a more extensive π system are more effective than pyrimidines. The extent of the π system follows the order dG(14)>dA(12)>dT(10)=dC(10) in terms of number of π electrons, and the order dG(153 Å^2)>dA(142 Å^2)=dT(142 Å^2)>dC(127 Å^2) in terms of surface area.[48] These results apply directly to the 5' nucleobase in our terminal labeling experiments since this is the base that is directly adjacent to the dye. We consistently observe, for both single- and double-stranded data that cyanine dye fluorescence follows this same trend, dG>dA>dT>dC, indicating that the terminal base directly affects rotational isomerization. The data also consistently shows that adjacent non-terminal bases modulate dye fluorescence, with a distance-dependent influence, indicating that sequence-dependent rigidity of the single- or double-stranded DNA also contributes to the observed fluorescence of Cy3 and Cy5. We hypothesize that the ability of the terminal base to hinder the rotational isomerization of the dye increases when part of a more rigid sequence of bases. The flexibility of DNA, particularly dsDNA, is of ongoing interest due to its role in packing and in the formation of protein-DNA complexes.[49] Many available degrees of freedom of the nucleobase contribute to DNA rigidity or flexibility, not all of which may be relevant to restricting the isomerization of the terminal dye; nevertheless, multiple experimental approaches indicate that purine stacks are more rigid that pyrimidine stacks in ssDNA.[50,51] A similar pattern is observed in dsDNA, also related to differences in base stacking area, dG (139 Å^2)>dA (128 Å^2)>dC (102 Å^2)>dT (95 Å^2), and stacking free energy, dA≫dG>dT≈dC (2.0, 1.3, 1.1 and 1.0 kcal·mol^{-1}), for B-form geometry, based on melting temperature changes.[48] Other experiments based on 5' dangling DNA hairpins and 3' RNA unpaired nucleotides give similar stability results: A≈G>T/U>C.[52,53] Sequence specificity of the flexibility of di- and tetramers[54,55], obtained from crystal structures and molecular dynamics simulations appear to be less relevant in this case because they treat paired bases symmetrically and as a single rigid unit, such that e.g. the deformability of AA(TT)=TT(AA). While this treatment is relevant to the ability of dsDNA to bend, the hydrogen bonding between Watson-Crick pairs does not contribute to duplex stabilization; instead, duplex stability is mainly determined by base-stacking interactions.[56] This suggests

that, at short length scales, the relevant modes of DNA dynamics are largely decupled from the complementary strand and interact with the cyanine dyes by restricting the available torsional volume and by changing high frequency coordinates of the potential energy surface of the excited state.[57]

CONCLUSION

With the data presented here, we have sought to clarify and quantify the impact of sequence-dependent fluorescence of Cy3 and Cy5 tethered to double-stranded DNA. The results are consistent with previous results of Cy3 and Cy5 and similar cyanide dyes tethered to single-stranded DNA.[19,26] The results are also consistent with measurements of the fluorescence yield of Cy3 in solution with each of the DNA nucleoside monophosphates, which also follows the pattern G>A>T>C.[46] The preponderance of evidence supports the hypothesis that stronger cyanine dye-nucleobase stacking interactions of the purines relative to the pyrimidines restrict the *cis-trans* isomerization rate of these dyes, enhancing fluorescence. The results can be used in the planning and analysis of experiments based on the labeling of DNA (and probably RNA) with cyanine dyes. For example, TaqMan or molecular beacon PCR probes and FISH probes using cyanine dye reporters can be designed with one or more guanines or adenines immediately adjacent to the dye for increased signal. The sequence for the latter two of these probes can also be adjusted so that the reporter dye is adjacent to a purine rich segment of the target upon hybridization. In the case of next generation sequencing-by-synthesis, where high throughput relies on maintaining the low end of the dynamic range near the noise threshold,[58,59] the data analysis pipeline can take into account the effect on measured fluorescence of adjacent nucleobases when determining the probability for a correct nucleobases assignment.

ASSOCIATED CONTENT

Supplementary Data are available: Relative Cy3 and Cy5 fluorescence intensity values in spreadsheet format.

AUTHOR INFORMATION

Corresponding Author

*E-mail: mark.somoza@univie.ac.at

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

Funding by the University of Vienna, the Faculty of Chemistry of the University of Vienna and the Austrian Science Fund (FWF P23797) is gratefully acknowledged.

REFERENCES

(1) Kupstat, A., Ritschel, T., and Kumke, M. U. (2011) Oxazine Dye-Conjugated DNA Oligonucleotides: Förster Resonance Energy Transfer in View of Molecular Dye–DNA Interactions, *Bioconjugate Chem. 22*, 2546-2557.

(2) Han, J. Y., and Burgess, K. (2010) Fluorescent Indicators for Intracellular pH, *Chem. Rev. 110*, 2709-2728.

(3) Luby-Phelps, K., Mujumdar, S., Mujumdar, R. B., Ernst, L. A., Galbraith, W., and Waggoner, A. S. (1993) A Novel Fluorescence Ratiometric Method Confirms the Low Solvent Viscosity of the Cytoplasm, *Biophysical J. 65*, 236-242.

(4) Brauns, E. B., Murphy, C. J., and Berg, M. A. (1998) Local dynamics in DNA by temperature-dependent Stokes shifts of an intercalated dye, *J. Am. Chem. Soc. 120*, 2449-2456.

(5) Somoza, M. M., Wiedersich, J., and Friedrich, J. (2007) Protein elasticity determined by pressure tuning of the tyrosine residue of ubiquitin, *J. Chem. Phys. 127*.

(6) Somoza, M. M., Andreatta, D., Murphy, C. J., Coleman, R. S., and Berg, M. A. (2004) Effect of lesions on the dynamics of DNA on the picosecond and nanosecond timescales using a polarity sensitive probe, *Nucleic Acids Res. 32*, 2494-2507.

(7) Grabolle, M., Brehm, R., Pauli, J., Dees, F. M., Hilger, I., and Resch-Genger, U. (2012) Determination of the Labeling Density of Fluorophore–Biomolecule Conjugates with Absorption Spectroscopy, *Bioconjugate Chem. 23*, 287-292.

(8) Park, J. W., Kim, Y., Lee, K.-J., and Kim, D. J. (2012) Novel Cyanine Dyes with Vinylsulfone Group for Labeling Biomolecules, *Bioconjugate Chem. 23*, 350-362.

(9) Michaelis, J., van der Heden van Noort, G. J., and Seitz, O. (2014) DNA-Triggered Dye Transfer on a Quantum Dot, *Bioconjugate Chem. 25*, 18-23.

(10) Santner, T., Hartl, M., Bister, K., and Micura, R. (2014) Efficient Access to 3'-Terminal Azide-Modified RNA for Inverse Click-Labeling Patterns, *Bioconjugate Chem. 25*, 188-195.

(11) Pawley, J. B. (2006) *Handbook of Biological Confocal Microscopy*, 3rd ed., Springer.

(12) Huang, Z. X., Ji, D. M., Wang, S. F., Xia, A. D., Koberling, F., Patting, M., and Erdmann, R. (2006) Spectral identification of specific photophysics of Cy5 by means of ensemble and single molecule measurements, *J. Phys. Chem. A 110*, 45-50.

(13) Widengren, J., and Schwille, P. (2000) Characterization of photoinduced isomerization and back-isomerization of the cyanine dye Cy5 by fluorescence correlation spectroscopy, *J. Phys. Chem. A 104*, 6416-6428.

(14) Dempster, D. N., Thompson, G. F., Morrow, T., and Rankin, R. (1972) Photochemical Characteristics of Cyanine Dyes.1. 3,3'-Diethyloxadicarbocyanine Iodide, *Photochem. Photobiol. Sci. 68*, 1479-1496.

(15) Sanborn, M. E., Connolly, B. K., Gurunathan, K., and Levitus, M. (2007) Fluorescence properties and photophysics of the sulfoindocyanine Cy3 linked covalently to DNA, *J. Phys. Chem. B 111*, 11064-11074.

(16) Hall, L. M., Gerowska, M., and Brown, T. (2012) A highly fluorescent DNA toolkit: synthesis and properties of oligonucleotides containing new Cy3, Cy5 and Cy3B monomers, *Nucleic Acids Res. 40*, e108.

(17) Iqbal, A., Wang, L., Thompson, K. C., Lilley, D. M. J., and Norman, D. G. (2008) The Structure of Cyanine 5 Terminally Attached to Double-Stranded DNA: Implications for FRET Studies *Biochemistry 47*, 7857-7862.

(18) Norman, D. G., Grainger, R. J., Uhrin, D., and Lilley, D. M. J. (2000) Location of cyanine-3 on double-stranded DNA: Importance for fluorescence resonance energy transfer studies, *Biochemistry 39*, 6317-6324.

(19) Agbavwe, C., and Somoza, M. M. (2011) Sequence-Dependent Fluorescence of Cyanine Dyes on Microarrays, *PLoS ONE 6*, e22177.

(20) Harvey, B. J., Perez, C., and Levitus, M. (2009) DNA sequence-dependent enhancement of Cy3 fluorescence, *Photochem. Photobiol. Sci. 8*, 1105-1110.

(21) Tyagi, S., and Kramer, F. R. (1996) Molecular Beacons: Probes that Fluoresce upon Hybridization, *Nat. Biotech. 14*, 303-308.

(22) Holland, P. M., Abramson, R. D., Watson, R., and Gelfand, D. H. (1991) Detection of specific polymerase chain reaction product by utilizing the 5'->3' exonuclease activity of Thermus aquaticus DNA polymerase *Proc. Natl. Acad. Sci. U. S. A. 88* 7276-7280

(23) Fuller, C. W., Middendorf, L. R., Benner, S. A., Church, G. M., Harris, T., Huang, X., Jovanovich, S. B., Nelson, J. R., Schloss, J. A., Schwartz, D. C., et al. (2009) The challenges of sequencing by synthesis, *Nat. Biotech. 27*, 1013-1023.

(24) Ouellet, M., Adams, P. D., Keasling, J. D., and Mukhopadhyay, A. (2009) A rapid and inexpensive labeling method for microarray gene expression analysis, *BMC Biotechnology 9*, 97.

(25) Jeon, H., and Choi, S. (2007) Fluorescence quenching causes systematic dye bias in microarray experiments using cyanine dye, *Genomics & Informatics 5*, 113-117.

(26) Kretschy, N., and Somoza, M. M. (2014) Comparison of the Sequence-Dependent Fluorescence of the Cyanine Dyes Cy3, Cy5, DyLight DY547 and DyLight DY647 on Single-Stranded DNA, *PLoS ONE 9*, e85605.

(27) Singh-Gasson, S., Green, R. D., Yue, Y. J., Nelson, C., Blattner, F., Sussman, M. R., and Cerrina, F. (1999) Maskless fabrication of light-directed oligonucleotide microarrays using a digital micromirror array, *Nat. Biotech. 17*, 974-978.

(28) Agbavwe, C., Kim, C., Hong, D., Heinrich, K., Wang, T., and Somoza, M. M. (2011) Efficiency, Error and Yield in Light-Directed Maskless Synthesis of DNA Microarrays, *J. Nanobiotechnol. 9*.

(29) Kretschy, N., Holik, A.-K., Somoza, V., Stengele, K.-P., and Somoza, M. M. (2015) Next-Generation *o*-Nitrobenzyl Photolabile Groups for Light-Directed Chemistry and Microarray Synthesis, *Angew. Chem., Int. Ed. 54*, 8555-8559.

(30) Warren, C. L., Kratochvil, N. C. S., Hauschild, K. E., Foister, S., Brezinski, M. L., Dervan, P. B., Phillips, G. N., and Ansari, A. Z. (2006) Defining the sequence-recognition profile of DNA-binding molecules, *Proc. Natl. Acad. Sci. U. S. A. 103*, 867-872.

(31) Lackey, J. G., Mitra, D., Somoza, M. M., Cerrina, F., and Damha, M. J. (2009) Acetal Levulinyl Ester (ALE) Groups for 2'-Hydroxyl Protection of Ribonucleosides in the Synthesis of Oligoribonucleotides on Glass and Microarrays, *J. Am. Chem. Soc. 131*, 8496-8502.

(32) McGall, G. H., Barone, A. D., Diggelmann, M., Fodor, S. P. A., Gentalen, E., and Ngo, N. (1997) The efficiency of light-directed synthesis of DNA arrays on glass substrates, *J. Am. Chem. Soc. 119*, 5081-5090.

(33) Nuwaysir, E. F., Huang, W., Albert, T. J., Singh, J., Nuwaysir, K., Pitas, A., Richmond, T., Gorski, T., Berg, J. P., Ballin, J., et al. (2002) Gene Expression Analysis Using Oligonucleotide Arrays Produced by Maskless Photolithography, *Genome Res. 12*, 1749-1755.

(34) Lackey, J. G., Somoza, M. M., Mitra, D., Cerrina, F., and Damha, M. J. (2009) In-situ chemical synthesis of rU-DNA chimeras on chips and enzymatic recognition, *Chim. Oggi-Chem. Today 27*, 30-33.

(35) Mertz, J. (2000) Radiative absorption, fluorescence, and scattering of a classical dipole near a lossless interface: a unified description, *JOSA B 17*, 1906-1913.

(36) Schneider, T. D., and Stephens, R. M. (1990) Sequence logos: a new way to display consensus sequences, *Nucleic Acids Res. 18*, 6097-6100.

(37) Crockett, A. O., and Wittwer, C. T. (2001) Fluorescein-Labeled Oligonucleotides for Real-Time PCR: Using the Inherent Quenching of Deoxyguanosine Nucleotides, *Anal. Biochem. 290*, 89-97.

(38) Seidel, C. A. M., Schulz, A., and Sauer, M. H. M. (1996) Nucleobase-specific quenching of fluorescent dyes.1. Nucleobase one-electron redox potentials and their correlation with static and dynamic quenching efficiencies, *J. Phys. Chem. 100*, 5541-5553.

(39) Widengren, J., Dapprich, J., and Rigler, R. (1997) Fast interactions between Rh6G and dGTP in water studied by fluorescence correlation spectroscopy, *Chem. Phys. 216*, 417-426.

(40) Shafirovich, V. Y., Courtney, S. H., Ya, N., and Geacintov, N. E. (1995) Proton-Coupled Photoinduced Electron Transfer, Deuterium Isotope Effects, and Fluorescence Quenching in Noncovalent Benzo[*a*]pyrenetetraol-Nucleoside Complexes in Aqueous Solutions, *J. Am. Chem. Soc. 117*, 4920-4929.

(41) Olmsted III, J., and Kearns, D. R. (1977) Mechanism of ethidium bromide fluorescence enhancement on binding to nucleic acids, *Biochemistry 16*, 3647-3654.

(42) Heinlein, T., Knemeyer, J.-P., Piestert, O., and Sauer, M. (2003) Photoinduced Electron Transfer between Fluorescent Dyes and Guanosine Residues in DNA-Hairpins, *J. Phys. Chem. B 107*, 7957-7964.

(43) Torimura, M., Kurata, S., Yamada, K., Yokomaku, T., Kamagata, Y., Kanagawa, T., and Kurane, R. (2001) Fluorescence-Quenching Phenomenon by Photoinduced Electron Transfer between a Fluorescent Dye and a Nucleotide Base, *Anal. Sci. 17*, 155-160.

(44) Pilailuk, S., Chittanon, B., and Chutintorn, P. (2015) Excited state free energy calculations of Cy3 in different environments, *Journal of Physics D: Applied Physics 48*, 205401.

(45) Spiriti, J., Binder, J. K., Levitus, M., and van der Vaart, A. (2011) Cy3-DNA Stacking Interactions Strongly Depend on the Identity of the Terminal Basepair, *Biophysical J. 100*, 1049-1057.

(46) Harvey, B. J., and Levitus, M. (2009) Nucleobase-Specific Enhancement of Cy3 Fluorescence, *J. Fluoresc. 19*, 443-448.

(47) Mikelsons, L., Carra, C., Shaw, M., Schweitzer, C., and Scaiano, J. C. (2005) Experimental and theoretical study of the interaction of single-stranded DNA homopolymers and a monomethine cyanine dye: nature of specific binding, *Photochem. Photobiol. Sci. 4*, 798-802.

(48) Guckian, K. M., Schweitzer, B. A., Ren, R. X.-F., Sheils, C. J., Tahmassebi, D. C., and Kool, E. T. (2000) Factors Contributing to Aromatic Stacking in Water: Evaluation in the Context of DNA, *J. Am. Chem. Soc. 122*, 2213-2222.

(49) Crothers, D. M. (1998) DNA curvature and deformation in protein–DNA complexes: A step in the right direction, *Proc. Natl. Acad. Sci. U. S. A. 95*, 15163-15165.

(50) Sain, A., Chen, J. Z. Y., and Ha, B.-Y. (2006) Persistency of single-stranded DNA: The interplay between base sequences and base stacking, *Phys. A 369*, 679-687.

(51) Solie, T. N., and Schellman, J. A. (1968) The interaction of nucleosides in aqueous solution, *J. Mol. Biol. 33*, 61-77.

(52) Doktycz, M. J., Paner, T. M., Amaratunga, M., and Benight, A. S. (1990) Thermodynamic stability of the 5' dangling-ended DNA hairpins formed from sequences 5'-(XY)2GGATAC(T)4GTATCC-3', where X, Y = A,T,G,C, *Biopolymers 30*, 829-845.

(53) Turner, D. H., Sugimoto, N., and Freier, S. M. (1988) RNA Structure Prediction, *Annu. Rev. Biophys. Biophys. Chem. 17*, 167-192.

(54) Fujii, S., Kono, H., Takenaka, S., Go, N., and Sarai, A. (2007) Sequence-dependent DNA deformability studied using molecular dynamics simulations, *Nucleic Acids Res. 35*, 6063-6074.

(55) Packer, M. J., Dauncey, M. P., and Hunter, C. A. (2000) Sequence-dependent DNA structure: dinucleotide conformational maps, *J. Mol. Biol. 295*, 71-83.

(56) Yakovchuk, P., Protozanova, E., and Frank-Kamenetskii, M. D. (2006) Base-stacking and base-pairing contributions into thermal stability of the DNA double helix, *Nucleic Acids Res. 34*, 564-574.

(57) Karunakaran, V., Pérez Lustres, J. L., Zhao, L., Ernsting, N. P., and Seitz, O. (2006) Large Dynamic

Stokes Shift of DNA Intercalation Dye Thiazole Orange has Contribution from a High-Frequency Mode, *J. Am. Chem. Soc. 128*, 2954-2962.

(58) Duthie, R. S., Kalve, I. M., Samols, S. B., Hamilton, S., Livshin, I., Khot, M., Nampalli, S., Kumar, S., and Fuller, C. W. (2002) Novel Cyanine Dye-Labeled Dideoxynucleoside Triphosphates for DNA Sequencing, *Bioconjugate Chem. 13*, 699-706.

(59) Bentley, D. R., and Balasubramanian, S., and Swerdlow, H. P., and Smith, G. P., and Milton, J., and Brown, C. G., and Hall, K. P., and Evers, D. J., and Barnes, C. L., and Bignell, H. R., et al. (2008) Accurate whole human genome sequencing using reversible terminator chemistry, *Nature 456*, 53-59.

Chapter 3

3 Conclusion and outlook

The first step in the direction of improving the existing procedure of photosensitive *in situ* microarray synthesis in this thesis and in terms of efficiency was by doubling the microarrays produced in each synthesis. Therefore a new assembly cell complex had to be designed and built to position both microarray substrates in the focal plane of the optical system. The enhancement results in halving the synthesis time per microarray and reduced reagent and solvent consumption. The new procedure also reduces the cost of microarray synthesis by almost a factor of two. Another major advantage of this improved method is that it results in the production of essentially identical - only mirror images - microarrays. This allows for more accurate comparisons between, for example in gene expression microarray experiments, since the control and treated samples can be hybridized to the two microarrays produced in the same synthesis. Since every new or separate synthesis is subject to inevitable small differences, a comparison with the new method becomes more reliable and significant. Further improvement concerning the cell assembly complex, e.g. tripling the microarrays per synthesis, is most unlikely due to physical and geometrical limitations with this set-up.

Stray light is an error source in light directed synthesis. This issue was also addressed in the same cell complex. Flare, diffraction, edge scattering and suspended dust particles play an important role in decreasing sequence fidelity and therefore unintended light should be suppressed and reduced to a minimum since the entire set-up of the microarray synthesis is based on a light-dependent and photolabile implementation. The cell assembly complex provides and extra chamber behind the synthesis area which can be filled with absorbing and index-matching fluid. Light is reflected by about 4% at each surface, and in order to minimize flare within the optical system, beta carotene in dichloromethane was used to fill this extra chamber as the solvent is a good match to the index of refraction of the glass microarray substrates. Beta carotene has a high extinction coefficient in the ultraviolet light near 365nm, and due to its low fluorescence quantum yield and high

Stokes shift, almost all reemitted light cannot lead to photodeprotection. Even though beta carotene seems to be working well in terms of absorption further investigations could reveal other substances that match the above mentioned criteria and work even better. In general a clean and vibration-free surrounding synthesis area should be maintained in order to guarantee a high synthesis yield. Mirrors and other optical devices and surfaces should be cleaned and therefore kept dust-free within the optical system on a regular basis in order to keep errors from stray light as low as possible.

Photosensitive *in situ* microarray synthesis is a well-established method for both laboratory and industrial scale purposes. Hence, further improvements for a higher throughput require a reduction of the synthesis time without any loss of quality and while maintaining low cost and high productivity. The new method already discussed above was introduced which halves the synthesis time by means of a newly designed photochemical reaction chamber which positions two glass substrates in the focal plane of the optical system. A drawback of this set-up is that it limits both microarrays to share the same design. The present thesis also addresses optimization with the highly-efficient thiophenyl (SPh)-NPPOC photolabile group, which requires much shorter photodeprotection times compared with the commonly used NPPOC phosphoramidites. The use of the SPh-NPPOC phosphoramidites, along with the optimization of each of the chemical and non-chemical steps during synthesis results in a much shorter synthesis time. Every synthesis benefits from shorter times but the superordinate aim of this project was to greatly reduce the synthesis time of a human genome-wide gene expression microarray which can be considered as a benchmark of modern *in situ* microarray synthesis.

The newly developed SPh-phosphoramidite profits from higher photolysis quantum yield and better absorptivity resulting in a 12 times faster photodeprotection rate compared to NPPOC, already greatly reducing time. Another approach analyzed the efficiency of different activators and activator concentration on the hybridization signals from microarray. The activator is needed for the nucleophilic substitution of the diisopropylamino group by the terminal 5’-OH group in order to extend the growing oligonucleotide chain by one incoming phosphoramidite. Hybridization signal data suggest that activator ETT followed by Activator 42 seem to work best in terms of signal intensities although both activators struggle with poor and inhomogeneous feature intensities and make data extraction less reliable. Whereas DCI activator

resulted in both high hybridization signals and feature homogeneity. Another advantage of DCI is its lower acidity compared to ETT and Activator 42 and therefore not making it mandatory to oxidize after every coupling step in order to stabilize the inter-phosphate linkage between two phosphoramidites. BTT and pyridinium chloride were excluded as usable activators since both provided very poor signal intensities and feature homogeneity. DCI as activator is most likely not optimal and it may be possible to find a better activator by mixing different activators or trying out new solvents.

Further time optimization could be achieved by adjusting other parameters, like shortening the acetonitrile washing steps and the helium blow without any sacrifice of quality during each coupling step. The effect of the helium blow and its resulting drying is yet not fully understood but nonetheless crucial for high hybridization intensities like shown in the provided data. Another time saving was achieved by decreasing the coupling time of each phosphoramidite by a factor of 4 from 60s to 15s. Determining certain thresholds where any further reduction leads to significant quality loss, the time needed for each coupling step can be still reduced by some seconds as shown by the provided data. Any further saving of time can't be accomplished with the current chemistry and would need an approach with other chemicals.

Applying all these optimizations, the synthesis time of a genome-wide gene expression microarray could be decreased from the initial ~8 hours to ~1.5 hours without any loss of quality.

A major accomplishment was achieved by Liétard and coworkers during the course of this thesis, the first syntheses of pure RNA microarrays by *in situ* photolithography and its direct proof provided by LC-(ESI-)MS. For the first time, direct detection of microarray oligonucleotides that are fabricated on glass substrates - about 20 pmol per chip - can be specifically cleaved from the surface and used for analysis by MS, was made accessible and rendered amplification by PCR obsolete. Measurement by MS depicts not only the full length products but also synthetic failures, degraded material and also the degree of incompletely deprotected sequences. The process of cleaving the oligonucleotide requires a custom-made protected NPPOC (dT) phosphoramidite (dT_{cleav}) with a base-labile ester which will cleave during the final deprotection step. Since this method was newly established there is room for improvement. The coupling efficiency of the cleavable dT

phosphoramidite is 85% determined by the difference of a sequence incorporated dT_{cleav} and without it. Different coupling times and different tetrazole based activators made the dT_{cleav} coupling yield worse. Further experimenting with coupling times or the application of a mix of different activators could improve the coupling yield making it more advantageous.

The final data presented in this work examined the resulting fluorescence of sequence-dependent cyanine dyes attached to double-stranded DNA (dsDNA). It was shown that both results, whether dyes are tethered to dsDNA or single-stranded DNA (ssDNA), behave in a similar way in terms of intensity.

The obtained evidence supports the hypothesis that nucleobase stacking interactions affect the *cis-trans* isomerization rate of the attached cyanine dyes. Purine-rich sequences, relative to pyrimidine-rich sequences result in higher dye fluorescence, indicating that purines preferentially reduce the *cis-trans* isomerization rate. These results can be considered in the sequence design of terminally-labeled oligonucleotides in order to achieve the highest possible fluorescence signal for data analysis and could be applied to sequence optimization for PCR, FISH, and FRET experiments.

4 Abstract

Highly complex and high-throughput microarrays of biopolymers, in this case nucleic acids, synthesized using maskless array synthesis (MAS) using photosensitive protecting groups, are well-established and relatively easy to use analytical tools.

MAS started out with the ability to make one microarray per synthesis. The cell assembly complex was improved so that now two identical, mirror images arrays, can be synthesized onto two different substrates simultaneously. An important aspect is to accurately and reliably position both substrates in the focal plane, which has a focal depth of about 70 µm. Several advantages are achieved, doubling the microarray synthesis rate and halving the synthesis time and costs per array. The reagent and solvents consumption is the same as for a single array synthesis. Another advantage is a more reliable comparison between experiments using the two mirror image microarrays since the microarrays are essentially identical with one another. Additionally, the new synthesis method allows for increased sequence fidelity of the microarray oligonucleotides by suppressing one of the largest sources of stray light; an extra chamber of the cell assembly complex can be filled with an absorbing and index-matching fluid in order to decrease unwanted reflection and contributes further in optimizing the sequence fidelity outcome.

The microarray synthesis chemistry was also highly optimized by using the highly light-sensitive thiophenyl-NPPOC (SPh-NPPOC) phosphoramidites, which reduce the necessary exposure time by a factor of 12. Coupling time was additionally decreased by a factor of four, to 15 seconds. Optimizing reagent delivery and incubation times of different chemicals and the crucial helium-flow also reduced synthesis time increasingly. Different phosphoramidite activators were also tested and to determine which of them are the best in terms of picture homogeneity, minimization of oxidising solution ─ needed to stabilize the growing oligonucleotide chain ─ and least expensive. Applying all improvements to the synthesis, the time for the synthesis of a gene expression microarray could be reduced from about 8 hours to about 1.5 hours.

5 Zusammenfassung (Abstract German)

Hoch komplexe *in situ* Microarrays, in diesem Fall von Nukleinsäuren, werden mittels der maskless array synthesis (MAS)-Technologie und lichtsensitiven Schutzgruppen synthetisiert und gehören zu den etablierten und relativ einfach handzuhabenden analytischen Werkzeugen. Im Zuge dieser Arbeit wurde es ermöglicht zwei idente – lediglich Spiegelbilder voneinander - Microarrays in nur einer Synthese auf zwei unterschiedlichen Substraten herzustellen. Wichtig dabei ist, dass beide Substrate genau in der optischen Brennebene positioniert werden. Die damit einhergehenden Vorteile beinhalten eine Halbierung der ursprünglichen Synthesezeit, sowie keinen Mehrverbrauch der Reagenzien, resultierend in Verringerung der Kosten. Ein weiterer Vorteil liegt im etwaigen Vergleich der beiden Microarrays miteinander. Des Weiteren wurden Versuche unternommen, einen wichtigen Faktor der Fehlerquellen, nämlich Streulicht, zu minimieren. Dabei ist es möglich eine extra Kammer des Halterungsapparates hinter dem Microarray mit lichtabsorbierenden Flüssigkeiten und dem Glas/Quartz ähnlichem Brechungsindex zu befüllen. Dadurch wird reflektiertes Licht reduziert und zeitgleich die richtige Sequenzausbeute erhöht. Im Verlauf der vorliegenden Arbeit wurde zusätzlich die gesamte involvierte Synthesechemie verbessert. Zum einen wurden Versuche mit den stark lichtsensitiven thiophenyl-NPPOC (SPh-NPPOC) Phosphoramiditen durchgeführt, welche die notwendige Belichtungszeit um einen Faktor von 12 reduzieren. Weitere Verbesserungen bei den Couplingzeiten, sowie des essentiellen Heliumbedarfs, verringerten ebenfalls die Synthesezeit. Zum anderen wurden unterschiedliche Aktivatoren auf ihre Effizienz und Bildhomogenität getestet und der beste Aktivator nachfolgend evaluiert. All jene Verbesserungen vereint, optimierten den gesamten Syntheseverlauf und reduzierten die Synthesezeit von Genexpressionsmicroarrays von den ursprünglichen ~ 8 h auf lediglich 1,5 h.

6 References

(1) J. D. Watson and F. H. C. Crick. Molecular structure of nucleic acids. *Nature*, 171(4356):737-738, 1953.

(2) R. L. Letsinger, J. L. Finnan, G. A. Heavner and W. B. Lunsford. Nucleotide chemistry. XX. Phosphite coupling procedure for generating internucleotide links. *J. Am. Chem. Soc.*, 97(11):3278-3279, 1975.

(3) S. L. Beaucage and M. H. Caruthers. Deoxynucleoside phosphoramidites—A new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Letters*, 22(20):1859-1862, 1981.

(4) S. L. Beaucage and R. P. Iyer. Advances in the synthesis of oligonucleotides by the phosphoramidite approach. *Tetrahedron*, 48(12):2223-2311, 1992.

(5) M. Schena, *Microarray analysis*, John Wiley & Sons Inc., 2003.

(6) A. C. Pease, D. Solas, E. J. Sullivan, M. T. Cronin, C. P. Holmes and S. P. Fodor. Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *PNAS*, 91(11):5022-5026, 1994.

(7) A. Hasan, K. P. Stengele, H. Giegrich, P. Cornwell, K. R. Isham, R. A. Sachleben. W. Pfleiderer and R. S. Foote. Photolabile Protecting groups for nucleosides: Synthesis and photodeprotection rates. *Tetrahedron*, 53(12):4247-4264, 1997.

(8) B. Lemieux, A. Aharoni and M. Schena. Overview of DNA chip technology. *Molecular Breeding*, 4(4):277-289, 1998.

(9) G. R. Gough, M. J. Brunden and P. T. Gilham. Recovery and recycling of synthetic units in the construction of oligodeoxyribonucleotides on solid supports. *Tetrahedron Letters*, 22(42):4177-4180, 1981.

(10) E. LeProust, H. Zhang, P. Yu, X. Zhou and X. Gao. Characterization of oligodeoxyribonucleotide synthesis on glass plates. *Nucl. Acids Res.*, 29(10):2171-2180, 2001.

(11) G. H. McGall, A. D. Barone, M. Diggelmann, S. P. A. Fodor, E. Gentalen and N. Ngo. The efficiency of light-directed synthesis of DNA arrays on glas substrates. *J. Am. Chem. Soc.*, 119(22):5081-5090, 1997.

(12) S. P. A. Fodor, J. L. Read, M. C. Pirrung, L. Styrer, A. T. Lu and D. Solas. Light-directed spatially addressable parallel chemical synthesis. *Science*, 251(4995):767-773, 1991.

(13) www.affymetrix.com, September 2015

(14) S. Singh-Gasson, R. D. Green, Y. Yue, C. Nelson, F. Blattner, M. R. Sussman and F. Cerrina. Maskless fabrication of light-directed oligonucleotide microarrays using a digital micromirror array. *Nature Biotechnology*, 17(10):974-978, 1999.

(15) J. G. Lackey, D. Mitra, M. M. Somoza, F. Cerrina and M. J. Damha. Acetal Levulinyl Ester (ALE) groups for 2´-hydroxyl protection of ribonucleosides in the synthesis of oligoribonucleotides on glass and microarrays. *J. Am. Chem. Soc.*, 131(24):8496-8502, 2009.

(16) J. Lietard, N. Kretschy, M. Sack, A. S. Wahba, M. M. Somoza and M. J. Damha. Base-cleavable microarrays for the characterization of DNA and RNA oligonucleotides synthesized *in situ* by photolithography. *Chem. Commun.*, 50:12903-12906, 2014.

(17) D.-S. Shin, K.-N. Lee, B.-W. Yoo, J. Kim, M. Kim. Y.-K. Kim and Y.-S. Lee. Automated maskless photolithography system for peptide microarray synthesis on a chip. *J. Comb. Chem.*, 12(4):463-471, 2010.

(18) E. F. Nuwaysir, W. Huang, T. J. Albert, J. Singh, K. Nuwaysir, A. Pitas, T. Richmond, T. Gorski, J. P. Berg, J. Ballin, M. McCormick, J. Norton, T. Pollock, T. Sumwalt, L. Butcher, D. Porter, M. Molla, C. Hall, F. Blattner, M. R. Sussman, R. L. Wallace, F. Cerrina and R. D. Green. Gene expression analysis using oligonucleotide arrays produced by maskless photolithography. *Genome Res.*, 12:1749-1755, 2002.

(19) A.-K. Holik, B. Rohm, M. M. Somoza and V. Somoza. *N* $^{\varepsilon}$-Carboxymethyllysine (CML), a Maillard reaction product, stimulates serotonin release and activates the receptor for advanced glycation end products (RAGE) in SH-SY5Y cells. *Food Funct.*, 4:1111-1120, 2013.

(20) H. Giegrich, S. Eisele-Bühler, C. Hermann, E. Kvasyuk, R. Charubala and W. Pfleiderer. New photolabile protecting groups in nucleoside and nucleotide chemistry—synthesis, cleavage mechanisms and applications. *Nucleosides and Nucleotides*, 17(9-11):1987-1996, 1998.

(21) S. Walbert, W. Pfleiderer and U. Steiner. Photolabile protecting groups for nucleosides : mechanistic studies of the 2-(2-nitropheny)ethyl group. *Helvetica Chimica Acta*, 84:1601-1611, 2001.

(22) C. Agbavwe, C. Kim, D. G. Hong, K. Heinrich, T. Wang and M. M. Somoza. Efficiency, error and yield in light-directed maskless synthesis of DNA microarrays. *J. Nanobiotechnol.*, 9, 57, 2011.

(23) J. Temsamani, M. Kubert and S. Agrawal. Sequence identity of the n-1 product of a synthetic oligonucleotide. *Nucleic Acids Research*, 23(11):1841-1844, 1995.

(24) C. Kim and F. Cerrina. DNA microarray: Imaging study. *J. Vac. Sci. Technol. B*, 21(6):2946-2950, 2003.

(25) K. Rosenhauer and K. Rosenbruch. Flare and optical transfer function. *Applied Optics*, 7(1):283-287, 1968.

(26) H. M. Geysen, R. H. Melven and S. J. Barteling. Use of peptide synthesis to probe viral antigen54s for epitopes to a resolution of a single amino acid. *Proc. Natl. Acad. Sci. U.S.A.*, 81:3998-4002, 1984.

(27) J. C. Venter et al., The sequence of the human genome. *Science*, 291(5507):1304-1351, 2001.

(28) H. Zhu, M. Bilgin, R. Bangham, D. Hall, A. Casamayor, P. Bertone, N. Lan, R. Jansen, S. Bidlingmaier, T. Houfek, T. Mitchell, P. Miller, R. A. Dean, M. Gerstein and M. Snyder. Global analysis of protein acitivities using proteome chips. *Science*, 293(5537):2101-2105, 2001.

(29) G. MacBeath and S. L. Schreiber. Printing proteins as microarrays for high-throughput function determination. *Science*, 289(5485):1760-1763, 2000.

(30) J. B. Delehanty and F. S. Ligler. Method for printing functional protein microarrays. *BioTechniques*, 34:380-385, 2003.

(31) D. J. Cahill. Protein and antibody arrays and their medical applications. *J. Immunol. Methods*, 250:81-91, 2001.

(32) D. Stoll, M. F. Templin, M. Schrenk, P. C. Traub, C. F. Vohringer and T. O. Joos. Protein microarray technology. *Front. Biosci.*, 7:c13-c32, 2002.

(33) M. F. Templin, D. Stoll, M. Schrenk, P.C. Traub, C. F. Vohringer and T. O. Joos. Protein microarray technology. *Trends Biotechnol.*, 20:160-166, 2002.

(34) www.illumina.com, October 2015

(35) T. K. Lindhorst. Essentials of carbohydrate chemistry and biochemistry. Weinheim, Wiley-VCH, 2003.

(36) W. G. T. Willats. Microarray for the high-throughput analysis of protein-carbohydrate interactions. *Protein Microarrays*, 57-69, 2005.

(37) D. Wang, S. Liu, B. J. Trummer, C. Deng and A. Wang. Carbohydrate microarrays for the recognition of cross-reactive molecular markers of microbes and host cells. *Nature Biotechnol.*, 20:275-281, 2002.

(38) S. Fukui, T. Feizi, C. Galustian, A. M. Lawson and W. Chai. Oligosaccharide microarrays for high-throughput detection and specificity assignments of carbohydrate-protein interactions. *Nature Biotechnol.*, 20:1011-1017, 2002.

LEBENSLAUF

BERUFSERFAHRUNG

Seit November 2015: Baxalta, Validierungsexperte Raman-Spektroskopie im Bereich Quality

AUSBILDUNG UND UNIVERSITÄRE LEISTUNGEN

2012 - 2015: Universität Wien, PhD Student

Institut für Anorganische Chemie auf dem Gebiet der *in situ* Microarray Synthese/ Nukleinsäurechemie.

2005 – 2011: Universität Wien, Diplomstudium Ernährungswissenschaften

Diplomarbeit am Institut für Medizinische Chemie (Medizinische Universität Wien)

2008: Technische Universität Wien, Bachelorstudium Elektrotechnik

1996 – 2004: Bundesgymnasium und Bundesrealgymnasium Neusiedl am See

Bilinguale Oberstufe (Deutsch - Englisch)**,** Informatik als Zusatzwahl

PUBLIKATIONEN

Express photolithographic DNA microarray synthesis with optimized chemistry and high-efficiency photolabile groups
M. Sack, A.-K. Holik, N. Kretschy, V. Somoza and M. M. Somoza
Journal of Nanobiotechnology, **2015 Eingereicht**

Sequence-dependent fluorescence of Cy3- and Cy5-labeled double-stranded DNA
N. Kretschy, M. Sack and M. M. Somoza
Journal Bioconjugate Chemistry, **2015 Eingereicht**

Base-cleavable microarrays for the characterization of DNA and RNA oligonucleotides synthesized *in situ* by photolithography
J. Lietard, N. Kretschy, M. Sack, A. S. Wahba, M. M. Somoza and M. J. Damha
Chemical Communications **2014**, 50, 12903-12906

Simultaneous light-directed synthesis of mirror-image microarrays in a photochemical reaction cell with flare suppression
M. Sack, N. Kretschy, B. Rohm, V. Somoza and M. M. Somoza
Analytical Chemistry **2013**, 85 (18), pp 8513-8517

Hard'n'Heavy: E-Guitar Sonatas
Matej Sack
Books on Demand **2014**

Hard & Heavy/ Graded Rock Guitar Solos
Matej Sack
Schell Music **2011**

KONFERENZEN: Posterpräsentationen

- Advances in microarray technology. Berlin, Deutschland, 2014
- Annual meeting of the European Iron Club. Louvain-la-Neuve, Belgien, 2011

ZIVILDIENST 2004 – 2005: Arbeiter-Samariter-Bund Österreichs

PRAKTIKA

SS 2010 und WS 2010/2011 – Tutor für biochemisches Praktikum

Juli 2010 – Arbeitsgemeinschaft der klinischen Ernährung AKE

November 2009 bis Jänner 2010 – Institut für Medizinische Chemie (MUW)

März 2009 – Institut für Medizinische Chemie (MUW)

Juli bis September 2003 – 2009 - Pioneer Hi-Bred International, Inc. (DuPont), Parndorf

SONSTIGE KENNTNISSE

- Sprachen: Slowakisch (Muttersprache), Deutsch (Muttersprache), Englisch (fließend in Wort und Schrift), Tschechisch (verhandlungsfähig), Französisch (sechsjährige Schulkenntnisse), Latein (Grundkenntnisse)
- Sprachzertifikate: Cambridge First Certificate
- EDV: MS-Office, Adobe Photoshop, MATLAB, SigmaPlot, GraphPad Prism, ChemSketch, NimbleScan
- Führerschein B Klasse