



**Entwicklung einer neuen Methode zur Selektion von synthetischen  
Wirkstoffbanken**

**Dissertation**

zur Erlangung des akademischen Grades des  
Doktors der Naturwissenschaften (Dr. rer. nat.)

eingereicht im Fachbereich Biologie, Chemie, Pharmazie  
der Freien Universität Berlin

vorgelegt von  
Diplom Biologin Julia Knospe  
aus Oranienburg  
2015

Diese Arbeit wurde in der Zeit von Juli 2010 bis November 2014 am Max-Planck-Institut für molekulare Genetik Berlin, in der Abteilung Vertebratengenomik (Prof. Dr. Hans Lehrach) unter der Leitung von Dr. Jörn Glökler angefertigt.

Gutachter: Prof. Dr. Beate Koksch (Freie Universität, Berlin)  
Prof. Dr. Hans Lehrach (Max-Planck-Institut für molekulare Genetik, Berlin)

Disputation am 07.10.2015

*Meinen Liebsten*

*Johanna und Karsten*

# INHALTSVERZEICHNIS

<b>INHALTSVERZEICHNIS .....</b>	<b>I</b>
<b>ABKÜRZUNGSVERZEICHNIS UND EINHEITEN .....</b>	<b>III</b>
<b>1 EINLEITUNG .....</b>	<b>1</b>
1.1 <i>IN VITRO</i> EVOLUTION .....	2
1.1.1 <i>IN VITRO</i> SELEKTIONSMETHODEN .....	2
1.2 DNA-KODIERTE BANKEN (DELS) .....	5
1.2.1 DELS OHNE EVOLUTIONÄRE PROZESSE .....	6
1.2.1.1 DIE „SPLIT & POOL“-METHODE .....	7
1.2.1.2 DIE ESAC-TECHNOLOGIE (ENCODED SELF-ASSEMBLING CHEMICAL LIBRARIES) ..	8
1.2.2 DELS MIT EVOLUTIONÄREN PROZESSEN .....	9
1.2.2.1 DNA-TEMPLATBASIERTE SYNTHESE (DNA-TEMPLATED SYNTHESIS) .....	10
1.2.2.2 SEQUENZABHÄNGIGE SORTIERUNG UND SYNTHESE (DNA-ROUTED SYNTHESIS) ..	11
1.3 SYNTHESE UND SELEKTION VON NUKLEINSÄURE-KODIERTEN CHEMISCHEN WIRKSTOFFBANKEN .....	12
1.4 NUKLEINSÄURE-BANKEN .....	15
1.5 HOCHDURCHSATZSEQUENZIERUNG DER NÄCHSTEN GENERATION .....	16
1.6 ZIELSTELLUNG .....	18
<b>2 MATERIAL UND METHODEN .....</b>	<b>19</b>
2.1 MATERIAL .....	19
2.2 METHODEN .....	26
2.2.1 NECLYS (NUCLEIC ACID-ENCODED CHEMICAL LIBRARY AND SELECTION, SYNTHESE UND SELEKTION VON NUKLEINSÄURE-KODIERTEN CHEMISCHEN WIRKSTOFFBANKEN) .....	26
2.2.2 ROBOTER KINGFISHER FLEX .....	26
2.2.3 KOPPLUNG VON NUKLEINSÄURE AN MIKROPARTIKEL .....	28
2.2.3.1 KOPPLUNG VON NUKLEINSÄUREN AN MAGNETISCHE MIKROPARTIKEL .....	28
2.2.3.2 IONISCHE BINDUNG VON NUKLEINSÄURE AN MAGNETISCHE MIKROPARTIKEL .....	30
2.2.4 LIGATION .....	31
2.2.5 SEPARATION .....	33
2.2.6 SPALTUNG .....	35
2.2.7 SYNTHESE .....	37
2.2.8 FMOC-SCHUTZGRUPPENABSPALTUNG .....	39
2.2.9 ELUTION VON DEAE-MIKROPARTIKELN .....	39
2.2.10 AUFREINIGUNG VON NUKLEINSÄUREN .....	40
2.2.11 ZWEITE SYNTHESERUNDE .....	41
2.2.12 SELEKTION .....	41
2.2.13 SEQUENZIERUNG .....	42
2.2.14 POLYMERASEKETTENREAKTION (PCR) .....	44
2.2.15 EMULSION-POLYMERASEKETTENREAKTION (EPCR) .....	46
2.2.16 GELELEKTROPHORESE .....	47
2.2.16.1 AGAROSE – GELELEKTROPHORESE .....	47
2.2.16.2 POLYACRYLAMID – GELELEKTROPHORESE .....	48
2.2.17 SPEKTROPHOTOMETRISCHE QUANTIFIZIERUNG VON DNA .....	49
2.2.18 NACHWEISMETHODEN FÜR NUKLEINSÄUREN .....	49

<b>3</b>	<b>ERGEBNISSE</b> .....	<b>51</b>
3.1	STRATEGIEN ZUR KOPPLUNG VON NUKLEINSÄURE AN MIKROPARTIKELN.....	52
3.1.1	KOVALENTES KOPPLUNGSVERFAHREN MITTELS CARBODIIMID-METHODE.....	52
3.1.2	KOPPLUNG VON BIOTINYLIERTER NUKLEINSÄURE AN DYNABEADS M-280 STREPTAVIDIN .....	54
3.1.3	NUKLEINSÄUREKOPPLUNG MITTELS IONISCHER BINDUNG.....	56
3.2	LIGATIONSREAKTIONEN DER DNA-BANK MIT ADAPTERFRAGMENTEN.....	58
3.2.1	EINFACHE LIGATION IM VERGLEICH ZUR ZYKLISCHEN LIGATION .....	60
3.2.2	ZYKLISCHEN LIGATION IM VERGLEICH ZUR ISOTHERMALEN LIGATION .....	62
3.3	SORTIERUNG DER LIGIERTEN DNA – BANK .....	63
3.3.1	MANUELLE SORTIERUNG.....	64
3.3.2	AUTOMATISIERTE SORTIERUNG PER KINGFISHER FLEX ROBOTER.....	66
3.3.3	ANWENDUNG DER SANGER-SEQUENZIERUNG FÜR DIE ANALYSE DER SORTIERUNG .....	69
3.4	SPALTUNGSMECHANISMUS AN DER SORTIERTEN DNA – BANK .....	71
3.5	SYNTHESE DER DNA-KODIERTEN SUBSTANZ – BANK UND SELEKTION GEGEN DAS ZIELMOLEKÜL STREPTAVIDIN .....	74
3.5.1	SYNTHESEREAKTION MIT D-BIOTIN UND SELEKTION GEGEN DAS ZIELMOLEKÜL STREPTAVIDIN .....	77
3.5.2	ZWEI SYNTHESERUNDEN MIT GLYCIN & D-BIOTIN UND SELEKTION GEGEN DAS ZIELMOLEKÜL STREPTAVIDIN .....	78
3.5.3	SYNTHESE MIT GLYCIN & D-BIOTIN UNTER 16 VERSCHIEDENEN SYNTHESEBEDINGUNGEN UND SELEKTION GEGEN STREPTAVIDIN.....	81
3.6	ANWENDUNG VON HOCHDURCHSATZSEQUENZIERUNG FÜR DIE ANALYSE VON DNA-KODIERTEN SUBSTANZ – BANKEN .....	84
3.6.1	DAS 454-PYROSEQUENZIERUNGSSYSTEM VON ROCHE APPLIED SCIENCES.....	85
<b>4</b>	<b>DISKUSSION</b> .....	<b>94</b>
4.1	MAGNETISCHE MIKROPARTIKEL UND DEREN EIGENSCHAFTEN.....	94
4.2	DIE EFFIZIENZRATE DER EINZELNEN ITERATIVEN ARBEITSSCHRITTE.....	96
4.2.1	DIE LIGATION DER DNA-BANK MIT DEN ADAPTEROLIGONUKLEOTIDEN.....	97
4.2.2	SORTIERUNGSMECHANISMUS .....	100
4.2.3	SPALTUNG AN EINER DEFINIERTEN POSITION IM DNA-FRAGMENT .....	101
4.2.4	DIE CHEMISCHE SYNTHESE UND DIE SELEKTION .....	103
4.3	HOCHDURCHSATZSEQUENZIERUNG VS. SANGER-SEQUENZIERUNG .....	104
4.4	DELS ALS NEUE PLATTFORM IN DER PHARMAZEUTISCHEN UND MEDIZINISCHEN INDUSTRIE .....	106
4.5	SCHLUSSFOLGERUNG UND AUSBLICK.....	109
<b>5</b>	<b>ZUSAMMENFASSUNG</b> .....	<b>112</b>
<b>6</b>	<b>SUMMARY</b> .....	<b>114</b>
<b>7</b>	<b>LITERATURVERZEICHNIS</b> .....	<b>116</b>
<b>6</b>	<b>ANHANG</b> .....	<b>125</b>
6.1	KINGFISHER FLEX PROTOKOLLE .....	125
6.2	SEQUENZIERUNGSDATEN .....	129
6.3	PUBLIKATIONEN.....	135
6.4	LEBENS LAUF .....	136
6.5	DANKSAGUNG .....	137
6.6	SELBSTSTÄNDIGKEITSERKLÄRUNG .....	138

## ABKÜRZUNGSVERZEICHNIS UND EINHEITEN

A	Purinbase Adenin
Abb.	Abbildung
Abk.	Abkürzung
APS	Ammoniumperoxidsulfat
ATP	Adenosintriophosphat
bidest H <sub>2</sub> O	aqua bidestillata (lat.) = doppelt destilliertes Wasser
bp	Basenpaare
BSA	bovines Serumalbumin
bzw.	beziehungsweise
C	Pyrimidinbase Cytosin
°C	Grad Celsius
CDI	1, 1'-Carbonyldiimidazole
DELS	DNA-kodierte Banken (DNA-encoded libraries)
dH <sub>2</sub> O	aqua destillata (lat.) = destilliertes Wasser
DMF	Dimethylformamid
DMT-MM	4-(4,6-Dimethoxy-1,3,5 triazin-2-yl)-4-methylmorpholiniumchlorid
DNA	Desoxyribonukleinsäure
dNTP	Desoxyribonukleosidtriphosphat
ds	doppelsträngig
EDC	1-Ethyl-3-(3-Dimethylaminopropyl)carbodiimid
EDTA	Ethylendiamintetraessigsäure
et al.	et alteres (lat.: und andere)
F	Fluoreszenzgruppe
Fmoc	Flourenylmethoxycarbonyl
G	Purinbase Guanin
g	Gramm
h	Stunde
HCl	Chlorwasserstoff (Salzsäure)
K	Kontrolle
l	Liter
lt.	laut
M	molar (mol/l)
MES	2-(N-Morpholino)ethansulfonsäure
min	Minute
Mio.	Millionen
MPIMG	Max-Planck-Institut für molekulare Genetik
NaCl	Natriumchlorid
NAD <sup>+</sup>	Nicotinamidadenindinukleotid
NaOH	Natriumhydroxid
NGS	next-generation sequencing
NHS	N-Hydroxysuccinimid
nt	Nukleotide
P	Phosphatgruppe
PAGE	Polyacrylamid-Gelelektrophorese
PBS	Phosphatgepufferte Salzlösung
PCR	Polymerase-Kettenreaktion

pH	pH-Wert
RNA	Ribonukleinsäure
rpm	Umdrehung pro Minute
RT	Raumtemperatur
s	Sekunden
SDS	Natriumlaurylsulfat
ss	einzelsträngig
SSC	Natriumchlorid-Natriumcitrat-Puffer
Std.	Stunde
T	Pyrimidinbase Thymidin
Tab.	Tabelle
TBE	Tris-Borsäure-EDTA
TE	Tris-EDTA
TEMED	Tetramethylethylendiamin
Tris	Tris-(hydroxymethyl-)aminomethan
U	Unit
V	Volt
w/v	Gewicht/Volumen

### Präfixe vor Einheiten

p	Piko ( $10^{-12}$ )
η	Nano ( $10^{-9}$ )
μ	Mikro ( $10^{-6}$ )
m	Milli ( $10^{-3}$ )
k	kilo ( $10^3$ )
M	Mega ( $10^6$ )
G	Giga ( $10^9$ )

# 1 EINLEITUNG

Für das Verständnis des lebenden Systems ist das Auffinden von funktionellen, molekularen Liganden, welche an biologischen Zielen binden, eine fundamentale Voraussetzung (Clark, 2010) und somit essentiell für die Entdeckung von neuen biologischen-, therapeutischen- bzw. medizinischen Komponenten.

Eine Vielzahl von Methoden sind für die Suche und das Auffinden von molekularen Liganden verfügbar und können grob in zwei Kategorien gegliedert werden. Zum einen gibt es die so genannte Screening-Methode und zum anderen die biochemische Selektion (Clark, 2010).

Die Screening-Methode basiert auf der Durchsichtung von umfangreichen Molekülbanken, auf der Suche nach biologisch bzw. pharmakologisch aktiven Substanzen. Heutzutage findet die Screening-Methode vor allem als automatisiertes Hochdurchsatz-Verfahren (High-Throughput-Screening, HTS) statt, bei dem eine gleichzeitige Durchsichtung von tausenden bis Millionen Komponenten möglich ist (Buller *et al.*, 2009). Dabei werden riesige Substanzdatenbanken in automatisierten *in vitro* Test-Assays auf Bindung an einen Zielrezeptor überprüft.

Die biologische Selektion unterscheidet sich in einem wichtigen Punkt von der Screening-Methode. Biologische Methoden ermöglichen die so genannte „Darwinsche Evolution“. Dabei wird sich der durch Charles Darwin (1809-1882) erstmals in der Veröffentlichung seines Buches im Jahre 1859 „*On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle of Life*“ beschriebene Selektionsdruck, zur Entdeckung von bindenden Substanzen an einem Zielrezeptor, zu eigen gemacht. Das bedeutet, bei der Durchsichtung bzw. Entdeckung von neuen an Zielsubstanzen bindenden Liganden werden durch den natürlichen Selektionsdruck die besten Kandidaten bevorzugt. Dabei müssen die Selektionsprozesse nicht an Zellen oder Organismen gebunden sein, sondern können auch *in vitro* in künstlichen Systemen ablaufen.

Dieser Prozess wurde in den letzten Jahrzehnten als *in vitro* Selektion- bzw. Evolution geprägt. Die traditionelle Selektion ist gegenüber der Screening-Methode durch Flexibilität, einfache Handhabung und geringen Kosten zwar überlegen, (Clark, 2010) weist allerdings auch einige Einschränkungen auf.



## 1.1 *IN VITRO* EVOLUTION

Durch die Entwicklung des Forschungsbereiches der Molekularbiologie wurde es möglich, Selektionen auch auf der molekularen Ebene durchzuführen. Die ersten Evolutionsexperimente auf der Ebene von Nucleinsäuren wurden durch Sol Spiegelman und seinen Kollegen in den 1960er Jahren durchgeführt. Die Arbeiten beruhten auf der Frage wie es Viren möglich ist, ihr RNA-Genom in Zellen mit einem DNA-Genom zu replizieren (Joyce, 2007). Im Laufe der Arbeiten isolierte und identifizierte Spiegelman und seine Kollegen *in vitro* die so genannten RNA-abhängigen RNA-Replikasen für die Replikation bestimmter RNA-Spezien (Mills *et al.*, 1967). Einen großen Einfluss auf die Molekularbiologie und vor allem auf die *in vitro* Evolution hatte die 1986 von Mullis und seinen Kollegen entdeckte und etablierte Polymerasekettenreaktion (PCR). Diese Methode ermöglicht es Nucleinsäuren spezifisch unter der Verwendung von Primern, Desoxynucleotiden und DNA-Polymerasen *in vitro* zu vervielfältigen, nahezu unabhängig von ihrer Länge oder Sequenz (Mullis *et al.*, 1986).

Die Etablierung der Polymerasekettenreaktion hatte einen starken Einfluss auf die Molekularbiologie im Allgemeinen und auf Evolutions- bzw. Selektionsexperimente im Besonderen. Nun war es möglich, *in vitro* jegliches einzel- bzw. doppelsträngige DNA- oder RNA-Fragment zu vervielfältigen (Joyce, 2007). Die Bedeutung der *in vitro* Evolution geht dabei vor allem auf die Möglichkeit zurück, dass während der iterativen Vorgänge Mutationen erfolgen können und es somit zu evolutionäre Prozesse kommen kann.

### 1.1.1 *IN VITRO* SELEKTIONSMETHODEN

Alle Selektionsmethoden basieren auf einem grundlegenden Prinzip, dem Prinzip der Genotyp-Phänotyp-Kopplung, unabhängig von der Art der Anwendung. Darunter versteht man die Verknüpfung einer replikationsfähigen Nucleinsäure (Genotyp) mit einer durch diese Nucleinsäure-kodierten Eigenschaft (Phänotyp) (Paschke, 2006).

Auch ohne die Etablierung der PCR-Methode konnte im Jahre 1985 Phage-Display als erste Selektionsmethode mit dem Prinzip der Genotyp-Phänotyp-Kopplung etabliert werden (Smith, 1985). Smith entwickelte eine Methode, um Polypeptide auf der Oberfläche filamentöser Bakteriophagen zu präsentieren.

Dabei handelt es sich um ein Virus, der *E. coli* Zellen infiziert allerdings nicht abtötet, sondern kontinuierlich in das Medium abgegeben wird. Infiziert man Bakterien mit einem Phagen, so kann dieser sich vermehren und präsentiert in Folge das neu eingeführte Peptid an seiner Oberfläche, während er gleichzeitig die Information für die entsprechende Proteinvariante in Form der DNA in seinem Inneren verpackt enthält (Paschke, 2006). Mit dieser Methode ist es somit möglich, aus großen rekombinanten Genbanken (Phagen-genomen) Polypeptide auf der Oberfläche der Bakteriophagen darzustellen und geeignete Bindungspartner für einen bestimmten Liganden zu identifizieren und zu isolieren.

Gerade vier Jahre nach der Etablierung der Polymerasekettenreaktion, wurde von drei verschiedenen Arbeitsgruppen unabhängig voneinander im Jahre 1990 die Isolierung von funktionellen kurzen Nukleinsäureliganden beschrieben.

Der Prozess der Selektion wurde von Tuerk und Gold mittels an T4 DNA-Polymerase bindenden RNA-Molekülen beschrieben. Dieser Prozess wurde als SELEX (Systematic Evolution of Ligands by EXponential Enrichment) bezeichnet (Tuerk & Gold, 1990). Ellington und Szostak berichteten über spezifisch an kleine organische Farbstoffe bindenden RNA-Molekülen, welche als Aptamere bezeichnet wurden (Ellington & Szostak, 1990). Als drittes beschrieben Robertson und Joyce die Anwendung der *in vitro* Selektion zur Anpassung der Gruppe-I-Ribozym-Spaltungsreaktion, bei der bevorzugt ssDNA anstelle von ssRNA gespalten wird (Robertson & Joyce, 1990). Diese Methoden dienen der Isolierung und Identifizierung von funktionellen, kurzen Nukleinsäuren, die durch einen definierten Versuchsaufbau festgelegte Eigenschaften aufweisen.

Im Jahre 1997 wurde eine weitere Selektionsmethode publik. Die von Hanes und Plückthun beschriebene Ribosomen-Display-Methode. Die Autoren beschreiben eine *in vitro* Proteinbiosynthese, bei der der mRNA-Ribosomen-Proteinkomplex nicht wie normalerweise nach der Translation zerfällt, sondern intakt gehalten wird. Dabei kodiert jede mRNA für ein spezielles Protein und ist mit diesem über ein einzelnes Ribosom verbunden. Anschließend erfolgt die Selektion der eventuellen Bindungspartner gegen Liganden, wobei durch Waschschrte unspezifische Ribosomen-Komplexe beseitigt werden. Die RNA wird von dem Komplex isoliert und mithilfe einer Reversen Transkription in cDNA und anschließend durch die PCR in DNA umgeschrieben. Nun kann die Identifizierung bzw. ein neuer Selektionszyklus erfolgen (Hanes & Plückthun, 1997).

Mit dieser Methode erhält man nicht nur Bindungspartner gegen Liganden, sondern die exprimierten Proteine können zudem den Nukleinsäure-Abschnitten, die sie kodieren, zugeordnet werden.

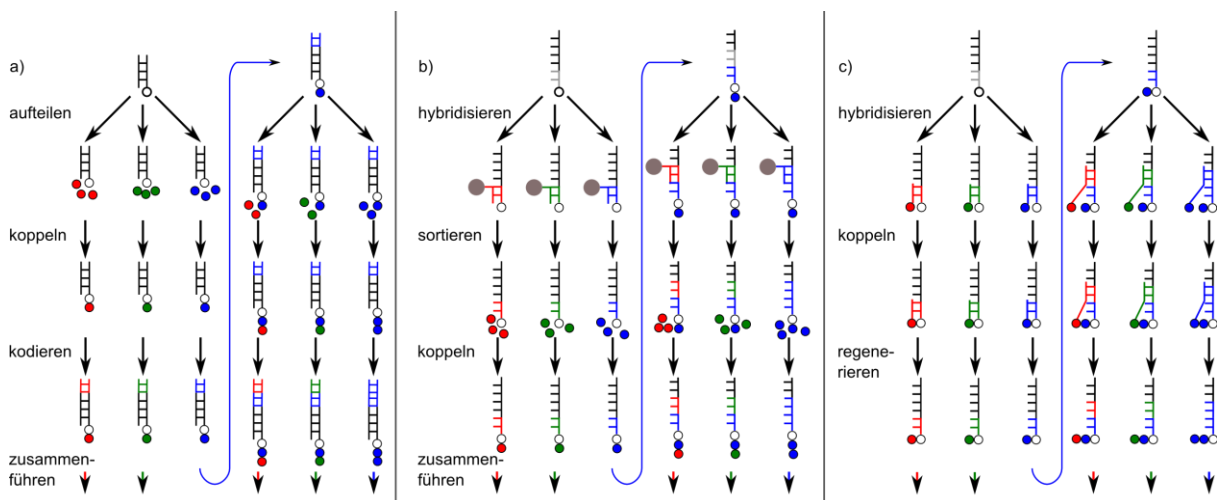
Obwohl Ribosomen komplexe biologische Systeme darstellen, sollte im gleichen Jahr noch eine Optimierung des Ribosomen-Display-Verfahrens erfolgen. Dabei stand im Vordergrund, ein kleines organisches Molekül zu finden, das diese hochmolekularen Strukturen reduziert und somit jedes Protein bzw. Peptid direkt mit seiner kodierenden mRNA verknüpft ist. Im Jahr 1997 wurde nahezu zeitgleich sowohl von Szostak (Roberts & Szostak, 1997) als auch von Yanagawa (Nemoto *et al.*, 1997) das Prinzip des mRNA-Display dargestellt. Dieses Prinzip beruht darauf, Peptide bzw. Proteine mit der für sie kodierenden mRNA kovalent über ein Brückenmolekül zu verknüpfen. Als Brückenmolekül wird dabei das Antibiotikum Puromycin verwendet, da dieses aufgrund einer strukturellen Ähnlichkeit das 3'-Ende einer aminoacrylierten tRNA nachahmt und so in die Bindungsstelle (A-Stelle) des Ribosoms eintreten kann. So kann mittels enzymatischer Synthese eine mRNA-Bank mit einem DNA-Puromycin-Oligomer (Linkermolekül) verbunden werden und es liegt anschließend ein mRNA-DNA-Puromycin-Molekül vor. Nun kann eine *in vitro* Translation erfolgen und das resultierende Fusionsmolekül besteht aus dem Protein, dem Puromycin, dem Linker und der mRNA. Am Übergang zwischen der mRNA und der DNA hat das Puromycin Gelegenheit die A-Stelle des Ribosoms zu besetzen, die Translation bricht ab und der Komplex zerfällt so, dass nun das Puromycin an dem Protein und über den Linker an der mRNA gebunden ist. Anschließend kann nach einer Reversen Transkription des RNA-Anteiles die Selektion durchgeführt werden (Richard, 1999).

Die *in vitro* Selektionsmethoden sind heutzutage gut etabliert und stellen eine exzellente Plattform für das Auffinden von Bindungspartnern gegen Liganden dar (Clark *et al.*, 2009). Allerdings zeigen sie auch einige Einschränkungen auf, vor allem bei der Entdeckung von niedermolekularen Wirkstoffen. Daher hat sich in den letzten Jahren ein neues Methodenfeld entwickelt, die so genannten DNA-kodierten Banken (DELs, DNA-encoded libraries).

## 1.2 DNA-KODIERTE BANKEN (DELS)

Das Hochdurchsatz-Screening und die biochemischen Selektionsmethoden von funktionalen Komponenten, welche als Liganden an biologischen Zielen binden, ist in der Pharmaindustrie ein wichtiger Zweig zur Identifizierung von neuen Wirkstoffen (Buller *et al.*, 2008). Neben den gängigen Methoden (siehe 1.1.1) hat sich in den letzten Jahren eine weitere Methode entwickelt, die DELs (DNA-encoded libraries bzw. DNA-kodierte Banken).

DELs präsentieren eine neue Technik, die in zwei Methoden unterteilt werden können. Bei dem ursprünglichen Modell handelte es sich zunächst um ein theoretisches Konzept ohne evolutionäre Prozesse, entwickelt von Sydney Brenner und Richard A. Lerner im Jahre 1992. In diesem Konzept wird eine chemische Synthese durch das Anhängen eines DNA-Codes protokolliert (Brenner & Lerner, 1992). Bei der zweiten Methode können auch evolutionäre Prozesse vollzogen werden. Hier werden die Nukleinsäuren genutzt, um die Synthese von kleinen Molekülen zu erfassen (Abb. 1). Beide Methoden haben gemein, dass sie die Kopplung zwischen einem Phänotyp (Molekül) und dem Genotyp (Nukleinsäure) voraussetzen.



**Abb. 1: Konzepte der DNA-kodierten Banken.** Originalkonzept nach Brenner & Lerner ohne evolutionäre Prozesse (a). Zweite Methode der DELs mit möglichen evolutionären Prozessen (b-c). Sequenzabhängige Sortierung und anschließende Synthese (b). DNA-templatbasierte Synthese (c).

Die neu entwickelten Methoden erlauben es, die gängigen Selektionsmethoden um wichtige Bereiche zu erweitern und somit steigt stetig das Interesse für den Einsatz neuer Strategien für die Selektion von DNA-kodierten Banken. DELs dienen der Identifizierung und Isolierung von kleinen Molekülen, welche in der Lage sind, an spezifischen biologischen Zielen zu binden.

Dabei handelt es sich um Sammlungen von Komponenten, die in der Akademie als auch in der Pharmazie verwendet werden, um Bindungspartner gegen jedes beliebige Ziel zu entdecken.

DNA-kodierte Substanz-Banken bestehen aus kleinen Molekülen, welche durch kombinatorische Chemie synthetisiert werden und kovalent an DNA-Fragmente gekoppelt sind. Die DNA-Fragmente besitzen einen identifizierbaren Barcode, wodurch die Kopplung von Genotyp (DNA) und Phänotyp (kleine Moleküle) ermöglicht wird.

Die Selektion von unnatürlichen Molekülen erfordert zudem Methoden, um einen synthetischen Chemotyp an einen lesbaren Genotyp zu koppeln. Für den geeigneten Genotyp erweist sich dabei DNA, da diese sehr stabil und zugänglich für Manipulationen, Amplifikationen und Sequenzierungen ist (Clark, 2010). DNA-kodierte chemische Wirkstoffbanken stellen ein vielversprechendes Werkzeug in der Wirkstoffforschung dar (Franzini *et al.*, 2014). Gegenwärtig arbeitet eine Vielzahl von Arbeitsgruppen an der Entwicklung kombinatorischer Selektionsmethoden, wobei die DELs eine ganz neue Plattform darstellen.

### 1.2.1 DELS OHNE EVOLUTIONÄRE PROZESSE

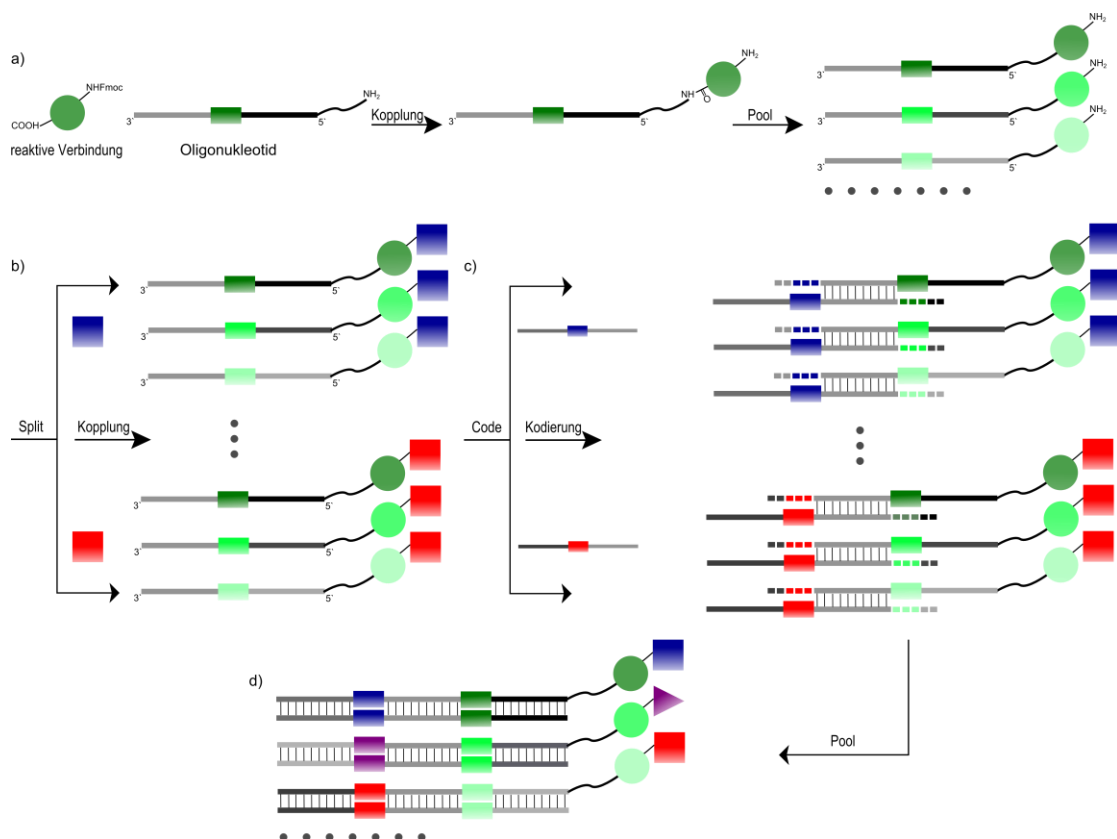
DNA-kodierte Banken bei denen es zu keinerlei evolutionären Prozessen kommt, stellen eine Technologie für die Synthese und das Screening von Sammlungen chemischer Komponenten dar, bei welcher der DNA-Code während der Synthese aufgebaut wird (Clark, 2010). Somit kann bei Technologien die auf diesem Konzept basieren nicht davon ausgegangen werden, dass evolutionäre Prozesse bzw. Mutationen vollzogen werden (Kleiner *et al.*, 2011). Die Herstellung von DNA-kodierten Banken, die für *in vitro* Selektionen verwendet werden können, beschäftigt viele akademische als auch industrielle Arbeitsgruppen.

Die DNA-kodierten Banken ohne evolutionäre Prozesse können in zwei Prinzipien aufgeteilt werden. Die „Split & Pool“-Methode wurde unter anderen von Wissenschaftlern der Firma GlaxoSmithKline und Nuevolution entwickelt und die so genannte ESAC (Encoded self-assembling chemical libraries)-Technologie wurde in den Laboratorien von Professor Dr. Neri (Institute of Pharmaceutical Science, Zürich, Schweiz) konstruiert (Scheuermann & Neri, 2010).

### 1.2.1.1 DIE „SPLIT & POOL“-METHODE

Die „Split & Pool“-Methode basiert immer auf einem ähnlichen Prinzip. Zunächst wird ein erstes Set von chemisch reaktiven Verbindungen an aminomodifizierten Oligonukleotiden gekoppelt (Mannocci *et al.*, 2011). Anschließend erfolgen eine Fmoc-Schutzgruppenabspaltung und eine Aufreinigung der Fragmente, die letztendlich vereinigt werden (Abb. 2, a). Die vereinte Reaktion wird nun in einzelne Reaktionsgefäße aufgeteilt („split“), die der Anzahl an unterschiedlichen reaktiven chemischen Verbindungen entspricht.

Nach erfolgter Kopplung (Abb. 2, b) wird die Kodierung der zweiten chemischen Verbindung vorgenommen (Scheuermann & Neri, 2010). Diese erfolgt mittels Hybridisierung eines teilweise komplementären Oligonukleotids und anschließender Elongation (Abb. 2, c) durch das Klenow-Fragment (Buller *et al.*, 2011). Die Schritte können iterativ vorgenommen werden, so dass DNA-kodierte Banken mit mehr als zwei chemisch gekoppelten Verbindungen entstehen (Mannocci *et al.*, 2011). Das Resultat ist eine doppelsträngige DNA-kodierte Bank (Abb. 2, d).



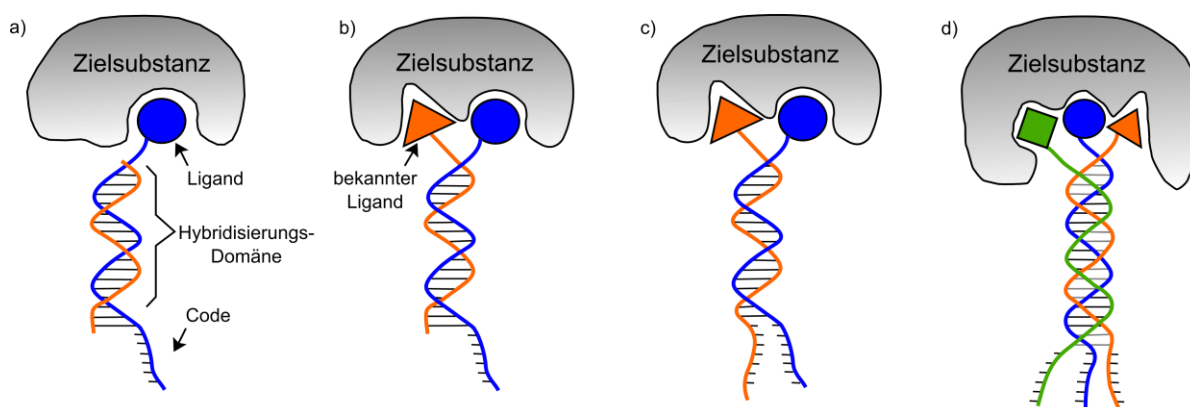
**Abb. 2:** Die „Split & Pool“-Methode der DNA-kodierten Bank ohne evolutionäre Prozesse.

Kopplung der ersten reaktiven Verbindung, anschließende Schutzgruppenabspaltung sowie Aufreinigung und Vereinigung der Reaktion (a). Split-Schritt der Reaktion in unterschiedliche Reaktionsgefäße, anschließende Kopplung (b) und Kodierung (c) der chemischen Verbindung. Das Resultat ist eine doppelsträngige DNA-kodierte Bank (d) (nach Buller *et al.*, 2010).

Die Kombination von „Split & Pool“-Chemie und die schrittweise DNA-Kodierung stellt eine effiziente und robuste Technik dar, um DNA-kodierte Banken herzustellen (Mannocci *et al.*, 2011).

### 1.2.1.2 DIE ESAC-TECHNOLOGIE (ENCODED SELF-ASSEMBLING CHEMICAL LIBRARIES)

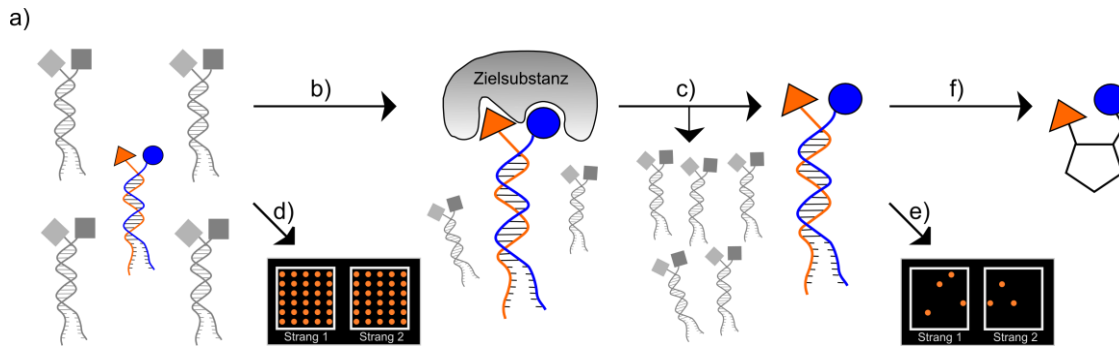
Die Konstruktion der ESAC-Technologie erfolgte sowohl in den Laboratorien von Prof. Dr. Neri (Melkko *et al.*, 2004) (Institute of Pharmaceutical Science, Zürich, Schweiz), als auch von Prof. Dr. Andrew Hamilton (Sprinz *et al.*, 2005) von der Yale Universität. Die ESAC-Technologie basiert auf dem Prinzip, dass zwei (oder mehr) DNA-Banken aus einer variablen Kodierungsregion, die durch eine Hybridisierungsregion flankiert wird und einer chemischen Modifikation, die für die Konjugation von reaktiven Verbindungen nötig ist, besteht (Melkko *et al.*, 2007). In der Abbildung 3 sind die verschiedenen Formate der ESAC-Banken dargestellt.



**Abb. 3: Schematische Darstellung der verschiedenen ESAC-Formate.** Einzel-Pharmakophor Selektion (a). Affinitätsreifung mittels eines bekannten Binders (b). DNA-Bank als Duplex (c). DNA-Bank als Triplex (d) (nach Scheuermann *et al.*, 2008).

Dabei wird mittels Hybridisierung durch die komplementären Hybridisierungsregionen ein DNA-Duplex bzw. auch -Triplex etc. gebildet, welches sowohl die reaktive Verbindung als auch die Kodierungsregion für die Ermittlung der gekoppelten Komponente beinhaltet (Melkko *et al.*, 2007). Im Anschluss kann eine Selektion gegen biologische Ziele durchgeführt werden, wobei die angereicherten Komponenten mittels PCR amplifiziert und durch ihren DNA-Code mithilfe der Sequenzierung bzw. DNA-Microarrays identifiziert werden können (Melkko *et al.*, 2004).

In der Abbildung 4 ist eine typische Selektion und Auswertung der ESAC-Banken dargestellt.



**Abb. 4: Darstellung einer typischen Selektion der ESAC-Bank und anschließende Auswertung.** Die DNA-Duplex (a) wird mit der Zielsubstanz der Wahl inkubiert (b) und anschließend gewaschen (c), um die nicht gebundenen Moleküle zu verwerfen. Anschließend erfolgen eine PCR auf den DNA-Code des verbleibenden Binders und ein Vergleich zwischen der unselektierten und der selektierten Bank mittels DNA-Microarray (d, e). Am Schluss erfolgt die Identifizierung der Liganden (f) durch den zugehörigen DNA-Code (nach Scheuermann *et al.*, 2008).

### 1.2.2 DELS MIT EVOLUTIONÄREN PROZESSEN

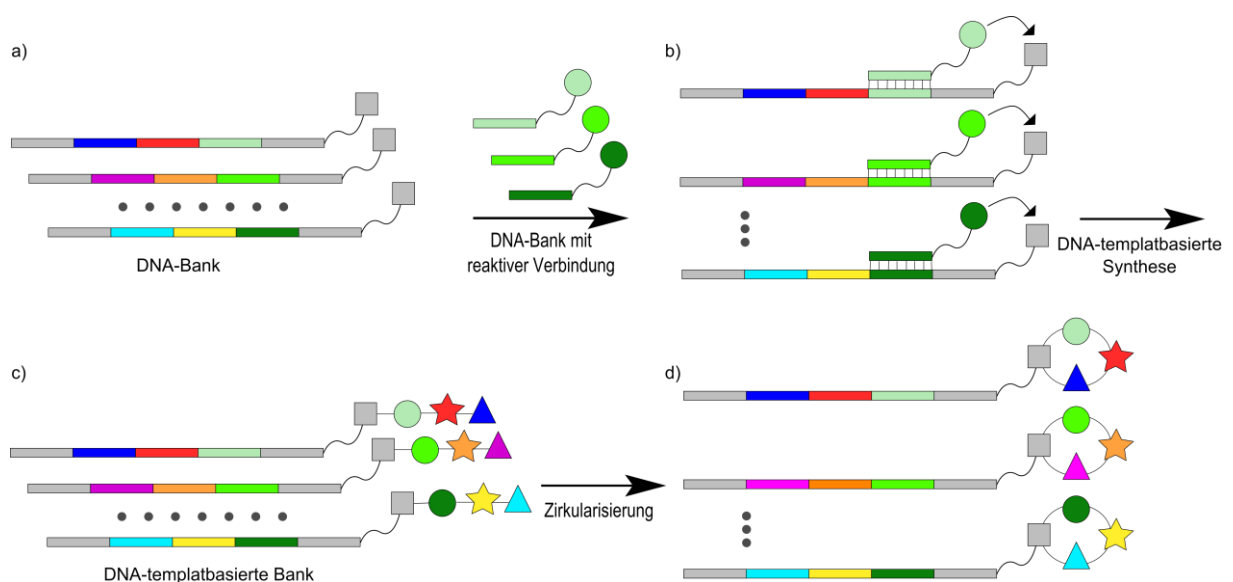
DNA-kodierte Banken, bei denen evolutionäre Prozesse vollzogen werden, können nach zwei Prinzipien eingeteilt werden. Zum einen gibt es die DNA-templatbasierte Synthese (DNA-templated synthesis), die von der Arbeitsgruppe um Liu (Gartner & Liu, 2001) im Jahre 2001 als Pionierarbeit entwickelt wurde. Unter diesen Zweig fällt auch die durch Wissenschaftlern von Vipergeren ApS (Kopenhagen, Dänemark) entwickelte Methode YoctoReactor (Hansen *et al.*, 2009). Zum anderen gibt es die sogenannte sequenzabhängige Sortierung mit anschließender Synthese (DNA-routed synthesis), die im Jahre 2004 von Halpin and Harbury (Halpin & Harbury, 2004a) dargelegt wurde.

Ein besonderes Merkmal der hier beschriebenen Methoden ist die Möglichkeit der Entstehung von evolutionären Prozessen. Da die Abläufe zumeist aus iterativen Prozessen bestehen, kann es während diesen zu Mutationen kommen, welche die Diversität der Banken erhöhen können. Dies ist möglich, da bei diesen Prinzipien der DNA-Code nicht während der Synthese aufgebaut wird, sondern schon generiert ist.



### 1.2.2.1 DNA-TEMPLATBASIERTE SYNTHESE (DNA-TEMPLATED SYNTHESIS)

Die DNA-templattbasierte Synthese nach Gartner und Liu (Gartner & Liu, 2001) basiert auf der Methode, dass komplementäre Sequenzen mit reaktiven Verbindungen in die räumliche Nähe der DNA-Banken gebracht werden (Clark, 2010). Die DNA-Bank besitzt verschiedene Kodierungsregionen, zu denen DNA-Sequenzen, an denen reaktive Verbindungen gekoppelt sind, komplementär sind. Durch eine Hybridisierung wird die reaktive Verbindung in die räumliche Nähe der DNA-Banken gebracht, wodurch anschließend die Übertragung auf die DNA erfolgen kann. Dieser Vorgang wird in Abhängigkeit der vorhandenen Kodierungsregionen wiederholt, so dass eine DNA-gerichtete Bank entsteht. Der Arbeitsablauf der DNA-gerichteten Synthese nach Gartner und Liu ist in Abbildung 5 dargestellt. Im Anschluss können Selektionen durchgeführt werden, wobei die bindenden Sequenzen durch PCR amplifiziert und Sequenzierung detektiert werden können.



**Abb. 5: Darstellung der DNA-templattbasierten Synthese (DNA-templated synthesis).** Die DNA-Bank enthält drei Kodierungsregionen, die mit einer komplementären DNA-Bank hybridisiert wird (a). Anschließend wird die reaktive Verbindung auf die komplementäre DNA-Bank übertragen (b). Dieser Vorgang wird korrespondierend zu der Anzahl der Kodierungsregionen wiederholt, wobei eine DNA-gerichtete Bank entsteht (c). Abschließend werden die reaktiven Verbindungen zirkularisiert (d) und eine Selektionsreaktion kann durchgeführt werden (nach Kleiner *et al.*, 2010).

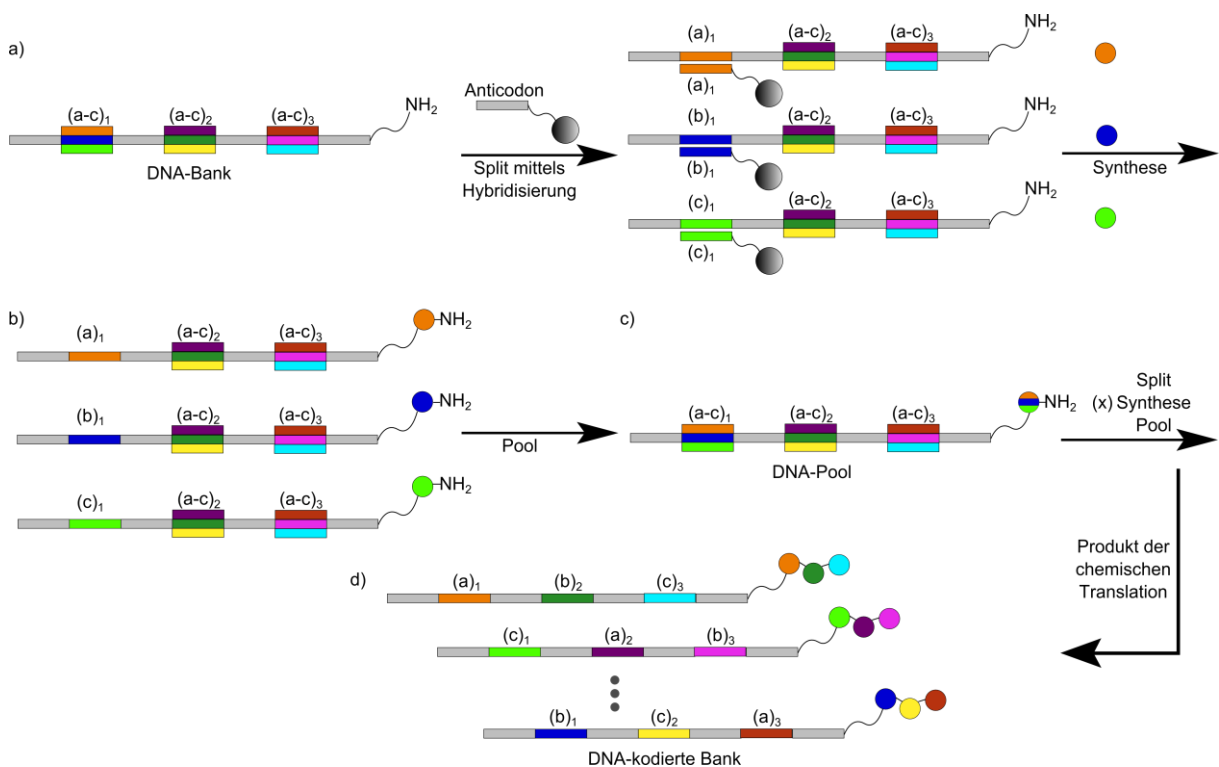
Die YoctoReactor-Methode von Vipergen ApS schließt die Bildung einer dreidimensionalen DNA-Verzweigung mit ein. Die DNA-Abschnitte enthalten eine reaktive Verbindung und eine zugehörige Kodierungsregion. Diese Reaktionspartner werden mithilfe eines Oligonukleotids in räumliche Nähe gebracht. Anschließend können die DNA-Abschnitte mittels Ligation vereinigt werden, wodurch auch die reaktiven Verbindungen gekoppelt sind.

Das Helferoligonukleotid kann nun entfernt und der Prozess mit dem nächsten Baustein fortgeführt werden (Heitner & Hansen, 2009).

### 1.2.2.2 SEQUENZABHÄNGIGE SORTIERUNG UND SYNTHESE (DNA-ROUTED SYNTHESIS)

Im Jahre 2004 präsentierten D.R. Halpin und P.B. Harbury (Halpin & Harbury, 2004a) eine neue interessante Methode zur Konstruktion DNA-kodierter Banken, die so genannte DNA-geleitete Synthese (DNA-routed synthesis). Erstmals dienen die DNA-Banken sowohl für die DNA-Kodierung als auch für die Programmierung der „Split & Pool“-Synthese.

Die Methode besteht aus einer Reihe von hintereinander geschalteten Säulen, welche zu den Kodierungsregionen der DNA-Bank komplementäre Anticodons beinhalten. Dadurch können eine Sortierung der DNA-Bank mittels Hybridisierung und eine anschließende Synthese erfolgen. Der Ablauf ist schematisch in Abbildung 6 dargestellt.



**Abb. 6: DNA-geleitete Synthese (DNA-routed synthesis).** Die DNA-Bank besteht aus Kodierungsregionen und einem 5'-Aminolinker für die Synthese. Die Sortierung erfolgt durch eine Hybridisierung mittels Anticodon, die an Mikropartikel gekoppelt sind (a). Anschließend kann die Festphasen-Synthese in verschiedenen Reaktionsgefäßen, spezifisch für jede Kodierungsregion erfolgen (b). Die DNA wird vereinigt und der Prozess wird wiederholt (c). Das Resultat der chemischen Translation ist eine DNA-kodierte Bank (nach Halpin & Harbury, 2004b)

Die Gruppe um Halpin und Harbury entwickelte eine Methode, bei der die DNA-Bank mittels komplementären Oligonukleotiden, welche an Partikeln gekoppelt sind, in verschiedene Reaktionsgefäße sortiert werden können. In diesen verschiedenen Reaktionskammern können dann die Festphasen-Synthesen stattfinden (Halpin *et al.*, 2004).

Durch die „Split & Pool“-Methode wird immer eine Synthese in einem Reaktionsgefäß durchgeführt, ganz im Gegenteil zu der DNA-gerichteten Synthese (DNA-templated synthesis), bei der alle Synthesen in einer Reaktionskammer stattfinden (Clark, 2010). Mittlerweile wurde die Methode zur Herstellung DNA-kodierter Banken von der Gruppe um Harbury erweitert, so dass viele Reaktionen parallel durchgeführt und ein hoher Durchsatz erreicht werden kann (Weisinger *et al.*, 2012).

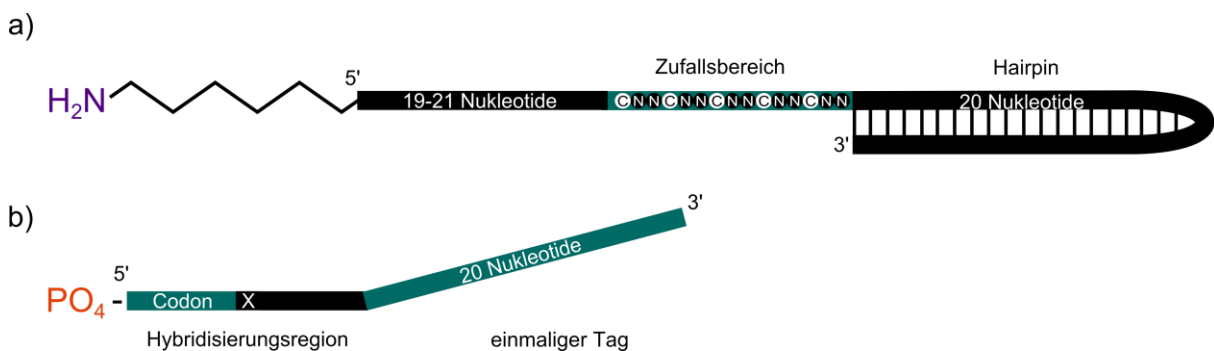
### **1.3 SYNTHESE UND SELEKTION VON NUKLEINSÄURE-KODIERTEN CHEMISCHEN WIRKSTOFFBANKEN**

Die hier entwickelte Methode zur Selektion von synthetischen Wirkstoffbanken, beziehungsweise die Entwicklung einer Methode zur Herstellung von DNA-kodierten Banken, wurde in dieser Arbeit als Neclys-Methode (Nucleic acid-encoded chemical library synthesis and selection, Synthese und Selektion von Nukleinsäure-kodierten chemischen Wirkstoffbanken) bezeichnet. Dabei sollen niedermolekulare Wirkstoffbanken durch kombinatorische chemische Synthese als Konjugate an dafür kodierende Nukleinsäuren hergestellt werden. Diese Technik basiert auf dem Prinzip der sequenzabhängigen Sortierung und anschließenden Synthese (siehe Abbildung 1), bei der die Nukleinsäure (Genotyp) benutzt wird, um die Synthese von kleinen Molekülen (Phänotyp) zu erfassen. Somit erfolgt auch hier die Kopplung von Phänotyp und Genotyp und es können zudem mögliche evolutionäre Prozesse vollzogen werden.

Im Detail: es soll eine Methode zur Synthetisierung von Nukleinsäure-kodierten kombinatorischen Banken entwickelt werden. Im Anschluss ist jede spezifische Komponente kovalent an eine Nukleinsäure, zur Dekodierung, gebunden. Bei den Komponenten kann es sich theoretisch um jedes Beliebige handeln, welches chemisch an die Nukleinsäure gekoppelt werden kann. Einige Beispiele wären Aminosäuren, Peptide, Polypeptide, Oligosaccharide, Nukleinsäuren oder organische Komponenten.

Die Herstellung der chemischen kombinatorischen DNA-kodierten Substanz-Bank soll durch iterative Abläufe, bestehend aus: Ligation, Separation, Spaltung, Synthese und Aufreinigung erfolgen. Der Schwerpunkt liegt dabei auf der Separation der DNA-Bank, die in dieser Arbeit über einen sehr kleinen, einzigartigen DNA-Abschnitt mittels Hybridisierung nach einer Ligationsreaktion erfolgen soll.

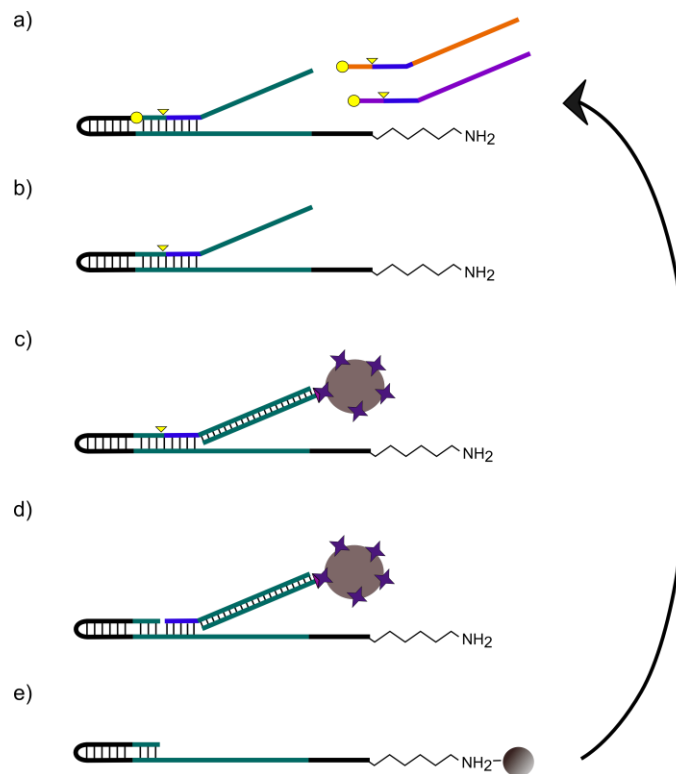
Als DNA-Bank soll ein Hairpin-Oligonukleotid verwendet werden, wodurch das Fragment teilweise doppelsträngig vorliegt und somit für die Ligation direkt verwendet werden kann. An dem 5'-Ende befindet sich eine funktionelle Gruppe für die Synthese. Zudem ist der Barcode des Oligonukleotids von konstanten Regionen flankiert, welche für Analysezwecke verwendet werden können (Abb. 7, a). Die Adapteroligonukleotide wurden so gestaltet, dass diese aus einer Hybridisierungsregion, welcher die Kodierungsregion enthält, einer Tagsequenz, welche zu der Kodierungsregion einmalig ist und einer Modifikation für die Spaltung bestehen (Abb. 7, b).



**Abb. 7: Schematische Darstellung der Nucleinsäure-Bank und der Adapteroligonukleotide.** Die DNA-Bank besitzt einen Zufallsbereich der von konstanten Regionen flankiert ist, wodurch eine Hairpinstruktur ausgebildet werden kann. Die funktionelle Aminogruppe ist in lila dargestellt, der Zufallsbereich in türkis (a). Die Adapteroligonukleotide besitzen an ihrem 5'-Ende eine Phosphatgruppe (orange), welche für die Ligationsreaktion benötigt wird. Anschließend folgen die Kodierungsregion und die Modifikation (X= abasische Position), die in dem Hybridisierungsbereich eingebettet sind. Hinter dieser Region folgt die einmalige Tagsequenz, welche zu der Kodierungsregion spezifisch ist (b).

Im ersten Schritt wird ein Adapter teilweise an die DNA-Matrize ligiert. Die Adapteroligonukleotide hybridisieren mit der einzelsträngigen DNA-Bank in der Nähe des Hairpins (Abb. 8, a) und die Ligase führt eine Verknüpfung unter stringenten Hybridisierungskonditionen durch, so dass die ersten dekodierten Basen perfekt zusammen passen (Abb. 8, b). So wird gewährleistet, dass nur die spezifischen Adapter direkt an das Ende der DNA-Matrize ligiert werden. Jeder Adapter besitzt zudem eine einzigartige 20-mer Tagsequenz, die für die spezifische Sortierung verantwortlich ist.

Im folgenden Schritt wird die ligierte DNA-Bank durch Hybridisierung mit immobilisierten Oligonukleotiden (Abb. 8, c) unter Verwendung des Magnetpartikelprozessor KingFisher Flex von Thermo Scientific sortiert. Dabei sind an den magnetischen Streptavidin-Mikropartikeln Hybridisierungsoligonukleotide konjugiert, welche zu den Tagsequenzen der Adapter komplementär sind. Durch die magnetischen Partikel kann die DNA-Bank sortiert und räumlich nach ihren Sequenzen separiert werden.



**Abb. 8: Schematische Darstellung des Arbeitsablaufes für die Synthese und Selektion von synthetischen Wirkstoffbanken.** Hybridisierung des spezifischen Adapters an der DNA-Bank (a). Ligation mittels der T4 DNA Ligase, Verknüpfung der beiden Fragmente unter stringenten Bedingungen (b). Separation der ligierten DNA-Bank mittels komplementären Oligonukleotiden, die an magnetischen Partikeln immobilisiert sind (c). Spaltung an der abasischen Position des Adapters durch die Endonuklease IV. So wird die DNA-Bank um drei Nucleotide verlängert und gleichzeitig der Adapter entfernt (d). Anschließend erfolgt die Synthese (e) und der Zyklus kann von vorne beginnen.

Nach der Separation werden die Partikel mit der sortierten DNA-Bank in separate Reaktionsgefäße überführt und der Adapter an seiner spaltbaren Modifikation mittels einer Endonuklease entfernt (Abb. 8, d). Bei der Modifikation handelt es sich um eine abasische Position, die durch die Endonuklease IV so gespalten werden kann, dass ein Einzelstrangbruch in der DNA (Boiteux & Guillet, 2004) und somit ein 5'-Phosphat- und ein 3'-Hydroxylende generiert wird.

Die dekodierte Sequenz bleibt an der DNA-Bank gebunden und die Synthese (Abb. 8, e) kann durchgeführt werden. Zu diesem Zweck wurden die Fragmente an magnetischen DEAE-Partikeln mittels ionischer Bindung gebunden, so dass eine Festphasenpeptidsynthese nach dem Carbodiimid-Verfahren (Altmann & Mutter, 1993) durchgeführt werden konnte. Die DEAE-Partikel mit der sortierten Bank werden in eine neue Mikrotiterplatte überführt, wobei in jedem Behälter eine andere Reaktionsbedingung vorliegt. Die Nukleinsäuren kodieren für die Synthese der gleichen Verbindung (Glycin für den 1. Synthesezyklus, d-Biotin für den 2. Zyklus), aber unter verschiedenen Konditionen. Es ist auch möglich unter gleichen Reaktionsbedingungen verschiedene Komponenten an die DNA zu synthetisieren. Nach der Synthese erfolgt die Schutzgruppenabspaltung und die DNA kann von den DEAE-Mikropartikeln eluiert werden. Schlussendlich wird die erhaltene DNA-kodierte Substanz-Bank vereint und für einen neuen Synthesedurchgang aufgereinigt. Nach der vollständigen chemischen Synthese ist es möglich eine Selektion gegen jedes beliebige Zielmolekül, nach dem bekannten Prinzip für Phage-Display oder SELEX (Glökler, *et al.*, 2010), durchzuführen. Auf diesem Wege können neue funktionelle Liganden aufgefunden werden.

## 1.4 NUKLEINSÄURE-BANKEN

Für die chemische Synthese von Oligonukleotiden wird vor allem die Phosphoramidit-Methode mittels Festphasensynthese verwendet (Beaucage & Iyer, 1992). Diese synthetisierten DNA-Banken besitzen im Allgemeinen einen Zufallsbereich, welcher von konstanten Regionen flankiert wird. Diese konstanten Regionen können sowohl für die Amplifikation als auch für die Ausbildung einer Hairpinstruktur des Oligonukleotids genutzt werden.



**Abb. 9: Schematische Darstellung der Nukleinsäure-Bank H<sub>2</sub>NhpIIIc6V16 bzw. H<sub>2</sub>NhpSOLiC6V16.** Die DNA-Bank besteht aus einer Zufallsregion, bei der jedes dritte Nukleotid durch ein konstantes Cytosin definiert ist. Flankiert wird diese Region von den konstanten Bereichen, die als Primersequenzen als auch zur Ausbildung einer Hairpinstruktur am 3'-Ende dienen. N= Nukleotide, C= Cytosin.

Die theoretische Diversität der Nukleinsäure-Bank ergibt sich aus der Anzahl der variablen Nukleinbasen innerhalb des Zufallsbereiches. In diesem Fall besitzt die DNA-Bank in dieser

Region gezielt eingebaute Cytosinbausteine, die sich nach zwei unbekanntem Nucleotiden wiederholen. Aus diesem Grund ergeben sich für jedes Codon  $4^2$ , somit 16 verschiedene Synthesemöglichkeiten. Dadurch wird bei der Verwendung von fünf Codons eine Diversität von  $16^5$  und somit eine DNA-kodierte Substanz-Bank bestehend aus circa eine Millionen verschiedenen Molekülen erreicht.

Diese Diversität kann zum einen mit einem Sequenzierungslauf abgedeckt werden und zum anderen bewegt sich die Diversität in einer Größenordnung, dass diese auch experimentell erreicht werden kann (International Human Genome Sequencing Consortium, 2004).

## 1.5 HOCHDURCHSATZSEQUENZIERUNG DER NÄCHSTEN GENERATION

Seit Einführung der Sanger-Sequenzierung im Jahre 1977 (Sanger *et al.*, 1977) entwickelten sich im Laufe der Zeit eine Vielzahl von Sequenzierungs-Technologien mit dem Anspruch der stetigen Effizienzsteigerung, Parallelisierung und natürlich Verringerung der Kosten. Heutzutage spricht man von der zweiten bzw. dritten Generation der Sequenzierungstechnologien (NGS – next generation sequencing bzw. NNGS – next next generation sequencing). Zu der ersten Generation der Sequenzierungssysteme gehört die automatisierte Sanger-Methode, die über mehrere Jahrzehnte dominieren konnte.

Seither wurden die Technologien stetig verbessert und weiter entwickelt, da diese zu einer der wichtigsten Methoden zählen, um biologische Systeme zu untersuchen. Zudem erreichte die Optimierung und Verbesserung der Sanger-Methode in den letzten Jahren ein Plateau und daraus resultierte die Entwicklung und Entstehung von neuen Systemen, die sich anderen Methoden der DNA-Sequenzierung bedienen (Bonetta, 2006).

Zu dem jetzigen Zeitpunkt gibt es drei kommerzielle Systeme von DNA-Sequenzierungsmethoden, die alle zu der zweiten Generation zählen. Darunter fallen zum einen Roche's 454 GS FLX Genome Analyzer vermarktet durch Roche Applied Sciences, das Solexa-Sequenzierungssystem von Illumina, welches auf dem Konzept der Sequenzierung durch Synthese (Mardis, 2008b) basiert und das durch Applied Biosystems vermarktete SOLiD-System (siehe Tab. 1). Alle diese Methoden haben gemein, dass die zu sequenzierende DNA-Bank per PCR hergestellt wird und den massiven parallelen Durchsatz (Lee & Tang, 2012). Unterschiede zeigen sich deutlich in der verwendeten Sequenzierungs-Chemie und dass jede Technologie andere Vor- bzw. Nachteile aufzeigt.

**Tab. 1: Vergleich der Sequenzierungsplattformen der zweiten Generation (Liu *et al.*, 2012).**

Plattform			
	Roche (454)	Illumina/Solexa	SOLiD
Jahr der Markteinführung	2005	2006	2007
Sequenzierungsmethode	Pyrosequenzierung	Sequenzierung durch Synthese	Sequenzierung durch Ligation
Amplifikation	Emulsions-PCR	Brückenamplifikation	Emulsions-PCR
Genauigkeit	99,9%	98%	99,94%
Datenmenge (Gb/Lauf)	0,7	600	120
Leselänge (bp)	700	35 - 150	35 - 50

Die 454-Methode von Roche zeigt klare Stärken in der Leselänge der DNA-Fragmente die bisher keine andere NGS-Methode erreicht. Allerdings zeigen sich prinzipbedingt Schwächen beim Lesen langer Homopolymere. Nachteile der Solexa- und SOLiD-Methode bestehen ganz klar in der Weite der Leselänge, im Gegenzug dazu werden aber sehr hohe Datenmengen pro Lauf hergestellt.

Zusammenfassend kann festgestellt werden, dass diese neuen Sequenzierungstechnologien zwar kürzere Leseweiten der DNA-Fragmente produzieren als die traditionelle Kapillarsequenzierung (~ 800 bp), allerdings ist es möglich Millionen von Sequenzen parallel zu sequenzieren und zu analysieren. Dadurch ergeben sich riesige Mengen an Sequenzierungsdaten, wodurch der Anwendungsbereich erweitert wird und auch Genomprojekte realisierbar sind, die vorher als nicht ökonomisch angesehen wurden.

Den Übergang zur dritten Generation stellt die Halbleitersequenzierung des Unternehmens Ion Torrent aus Guilford in Connecticut, das zu Life Technologies gehört, dar (Schadt *et al.*, 2010). Mittlerweile gibt es eine Vielzahl neuer Technologien, die vor allem von der klonalen DNA-Amplifikation per PCR Abschied nehmen, die für die Sequenzierungsmethoden der zweiten Generation unverzichtbar für die Herstellung ausreichender Mengen an Sequenziermatrizen sind. Somit wird eine weitere Kosten- und Arbeitsreduktion erzielt und zudem die Einzelmolekül-Sequenzierung erreicht. Firmen die sich mit der dritten Generation von Sequenzierungsmethoden beschäftigen sind unter anderen Pacific Biosciences (SMRT – single molecule real time sequencing), Helicos Biosciences (tSMS – true single molecule sequencing), Oxford Nanopore (nanopore sequencing) (Metzker, 2010) und viele mehr.



## 1.6 ZIELSTELLUNG

Ziel dieser Arbeit ist die Entwicklung einer neuen Methode, mit welcher niedermolekulare Wirkstoffbanken durch kombinatorische chemische Synthese als Konjugate an dafür kodierenden Nucleinsäuren hergestellt werden können. Nach der Synthese der kombinatorischen Wirkstoffbank ist eine Selektion von neuartigen Wirkstoffen möglich.

Das Hochdurchsatz-Screening und die biochemischen Selektionsmethoden von funktionalen Komponenten, welche als Liganden an biologischen Zielen binden, ist in der Pharmaindustrie ein wichtiger Zweig zur Identifizierung von neuen Wirkstoffen, deren Bedeutung stetig steigt.

Zunächst sollen in dieser Arbeit die einzelnen Schritte zur Herstellung von niedermolekularen Wirkstoffbanken erarbeitet und protokolliert werden. Im Anschluss erfolgen die Vereinigung der Schritte und das Übertragen auf einen programmierbaren Roboter, so dass ein halbautomatisiertes Verfahren entsteht. Ziel ist es, zwei spezifische Syntheserunden zu vollziehen und im Anschluss eine Selektion gegen ein Zielmolekül durchzuführen. Mittels Hochdurchsatzsequenzierung soll so gezeigt werden, dass die Kopplung des Genotyps (Nucleinsäure) und des Phänotyps (Molekül) spezifisch erfolgt und jedes beliebige gekoppelte Molekül durch den einzigartigen DNA-Code der Nucleinsäure aus einer riesigen DNA-Bank ermittelt werden kann.

Im Vordergrund dieser Arbeit steht dabei die Separation, also die räumliche Trennung einer DNA-Bank anhand seiner Nucleinsäureabfolge, die im Zusammenhang mit DNA-kodierten Banken den sogenannten Genotyp darstellt. Das heißt, es soll im Laufe der Arbeit erreicht werden, dass eine Nucleinsäure-Bank mittels Hybridisierung eines nur sehr kleinen einzigartigen Abschnittes in unterschiedliche Gefäße sortiert bzw. separiert werden kann. Bisher ist dies nur über sehr lange Hybridisierungsregionen spezifisch möglich, wodurch die Diversität eingeschränkt ist und hohe Kosten entstehen.

## 2 MATERIAL UND METHODEN

### 2.1 MATERIAL

**Tab. 2: Chemikalien und deren Bezugsfirmen**

Chemikalien	Hersteller
ABIL WE 09	Evonik Goldschmidt, Essen
Acrylamid/Bisacrylamid (19:1)	Bio Rad Laboratorius, München
Ammoniumperoxidsulfat (10%, w/v)	Merck, Darmstadt
Betain	Sigma-Aldrich®, Steinheim
Bromphenolblau	Sigma-Aldrich®, Steinheim
Calciumchlorid	Merck, Darmstadt
1,1'-Carbonyldiimidazole	Sigma-Aldrich®, Steinheim
d-Biotin	Sigma-Aldrich®, Steinheim
Dimethylformamid	Sigma-Aldrich®, Steinheim
4-(4,6-Dimethoxy-1,3,5 triazin-2-yl)-4-methylmorpholiniumchloride	Sigma-Aldrich®, Steinheim
1-Ethyl-3-(3-Dimethylaminopropyl)carbodiimid	Sigma-Aldrich®, Steinheim
Ethylendiamintetraessigsäure	Merck, Darmstadt
Ethanol	Merck, Darmstadt
Essigsäure	Merck, Darmstadt
Ethanolamine	Sigma-Aldrich®, Steinheim
Ethidiumbromid, 10 mg/ml	Invitrogen™, Karlsruhe
Fmoc-Glycin-OH	Sigma-Aldrich®, Steinheim
Glycerol	Merck, Darmstadt
Harnstoff	Sigma-Aldrich®, Steinheim
Hydroxylapatit	Sigma-Aldrich®, Steinheim
Isobutanol	Merck, Darmstadt
Isopropanol	Merck, Darmstadt
Kaliumchlorid	Merck, Darmstadt
Kaliumdihydrogenphosphat	Merck, Darmstadt
Kaliumhydroxid	Merck, Darmstadt
Magnesiumchlorid	Merck, Darmstadt
Methanol	Merck, Darmstadt
1-Methyl-2-Pyrrolidon, 99%	Sigma-Aldrich®, Steinheim
Mineralöl	Sigma-Aldrich®, Steinheim
2-(N-Morpholino)ethansulfonsäure	Sigma-Aldrich®, Steinheim
Natriumborat	Merck, Darmstadt
Natriumchlorid	Merck, Darmstadt
Natriumcitrat	Merck, Darmstadt
Natriumdihydrogenphosphat	Sigma-Aldrich®, Steinheim
Natriumhydroxid	Merck, Darmstadt
Natriumlaurylsulfat	Roth, Karlsruhe
Natriumphosphat	Merck, Darmstadt
N-Hydroxysuccinimid	Sigma-Aldrich®, Steinheim
Piperidin, 99%	Sigma-Aldrich®, Steinheim
2-Pyrrolidon, 99%	Sigma-Aldrich®, Steinheim
Tegosoft DEC	Evonik Goldschmidt, Essen
Tetramethylethyldiamin	Invitrogen™, Karlsruhe
Tris	Merck, Darmstadt

Triton X-100 Tween® 20 UltraPure Agarose Wasserstoffperoxid, 30% Xylencyanol	Sigma-Aldrich®, Steinheim Riedel-de Haen®, Seelze Invitrogen™, Karlsruhe Sigma-Aldrich®, Steinheim Merck, Darmstadt
------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------

**Tab. 3: Lösungen und Puffer**

Lösungen bzw. Puffer	Zusammensetzung
2× B&W-Puffer	10 mM Tris-HCl (pH 7,5) 1 mM EDTA 2 M NaCl
Emulsions-Öl	7% ABIL WE 09 20% Mineral Öl 73% Tegosoft DEC
SiMAG-DEAE-Beads Bindungspuffer	20 mM Tris-HCl (pH 8)
SiMAG-DEAE-Beads Elutionspuffer	20 mM Tris-HCl (pH 8) 0,5 / 1 / 1,5 M Natriumchlorid
Hybridisierungspuffer/Separation	0,5 M Natriumchlorid 50 mM Natriumcitrat 0,1% Tween® 20
	0,75 M Natriumchlorid 75 mM Natriumcitrat 0,1% Tween® 20
	1,5 M Natriumchlorid 150 mM Natriumcitrat 0,1% Tween® 20
10× Ladepuffer für Agarose-Gelelektrophorese	200 mM Tris-HCl (pH 7,4) 50 mM MgCl <sub>2</sub> 20 mM CaCl <sub>2</sub> 20 mM KCl 1,4 mM NaCl 0,5% Tween® 20
	0,1% Xylencyanol 0,1% Bromphenolblau 0,5% SDS 0,1 M EDTA (pH 8) 50% Glycerin

6× Ladepuffer für PAGE	10 mM Tris-HCl (pH 7,6) 60% Glycerin 0,03% Xylencyanol 0,03% Bromphenolblau 60 mM EDTA
Lagerungspuffer (SiMAG-Carboxyl-Beads)	1x PBS (pH 7,4) 0,1% Tween® 20 10 mM Ethanolamin
0,1 M MES (pH 5) / (pH 6)	0,976 g MES ad 50 ml bidest H <sub>2</sub> O
0,1 M MES (pH 5), 1,25 M EDC	120 mg EDC 500 µl 0,1 M MES (pH 5)
20% Piperidin in DMF	2 ml Piperidin 8 ml DMF
1× PBS (pH 7,4)	1,7 mM Kaliumdihydrogenphosphat 5,2 mM Natriumdihydrogenphosphat 150 mM Natriumchlorid
10% SDS-Puffer	1 g SDS-Pellets ad 10 ml bidest H <sub>2</sub> O
20× SSC-Puffer	3 M Natriumchlorid 0,3 M Trinatriumcitrat
1× TBE-Puffer	89 mM Tris 89 mM Borsäure 2 mM EDTA (pH 8)
1× TE-Puffer	10 mM Tris-HCl (pH 8) 1 mM EDTA
1 M Tris-HCl-Puffer (pH 7,4 / 7,6 / 8)	30,3 g Tris ad 250 ml bidest H <sub>2</sub> O
TT-Puffer	0,25 M Tris-HCl (pH 8) 0,01% Tween® 20
TT-Puffer, 10 mM Ethanolamin	0,25 M Tris-HCl (pH 8) 0,01% Tween® 20 10 mM Ethanolamin

**Tab. 4: Laborgeräte und Verbrauchsmaterialien**

<b>Laborgeräte bzw. Verbrauchsmaterial</b>	<b>Hersteller</b>
DynaMag Magnetseparator	Invitrogen™, Karlsruhe
Gel-Elektrophoresis EPS 601	Pharmacia biotech, Amersham
KingFisher Flex Roboter	Thermo Scientific®, Vantaa
Laborwaage LC 2200	Sartorius, Göttingen
Laborwaage 2004 MP	Affimetrix®, High Wycombe
Laser Scanner Affimetrix 428	Sartorius, Göttingen
pH-Meter 762 calimatic	Knick®, Berlin
Rotator stuart SB3	BioCote®, Staffordshire
Spectrophotometer, NanoDrop ND-1000	Thermo Scientific®, Vantaa
Thermal Cycler PTC-200	Biozyme®, Oldendorf
Thermomixer compact 5355	Eppendorf, Hamburg
UVT-28 M Transilluminator	Herolab GmbH, Wiesloch
Vortex-Genie 2	Scientific Industries™, Karlsruhe
Zentrifuge 5810R	Eppendorf, Hamburg
Zentrifuge 5415D	Eppendorf, Hamburg

**Tab. 5: Enzyme, Lösungen und Kits**

<b>Enzyme, Lösungen und Kits</b>	<b>Hersteller</b>
ATP, 10 mM	Fermentas, St. Leon-Rot
bovines Serumalbumin (10mg/µl)	New England BioLabs®, USA
dNTPs, 2,5 mM	Roboklon, Berlin
Dynabeads® M-280 Streptavidin, Ø 2,8 µm, 10 mg/ml	Invitrogen™, Karlsruhe
Dynabeads® MyONE Carboxylic Acid, Ø 1 µm, 10 mg/ml	Invitrogen™, Karlsruhe
Endonuklease IV, 10 U/µl	New England BioLabs®, USA
<i>E. coli</i> Ligase (10 U/µl)	New England BioLabs®, USA
GeneMATRIX PCR/DNA Clean-Up DNA Purification-Kit	Promega, Mannheim
O'RangeRuler™ 20 bp DNA Ladder	Fermentas, St. Leon-Rot
SiMAG-Carboxyl, Ø 0,5 µm, 600 µmol COOH/g, 50 mg/ml	Chemicell, Berlin
SiMAG-DEAE, Ø 0,5 µm, 50 mg/ml	Chemicell, Berlin
SV40 Large T Antigen (Helikase), 1 µg/µl	Roboklon, Berlin
Taq DNA-Polymerase, 5 U/µl	Roboklon, Berlin
T4 DNA Ligase, 268 U/µl	Roboklon, Berlin
Wizard SV Gel and PCR Clean-up System	Roboklon, Berlin

## OLIGONUKLEOTIDE

Die verwendeten Oligonukleotide sind in den folgenden Tabellen mit ihrer jeweiligen Sequenz, Länge in Nukleotiden und dem jeweiligen Hersteller aufgelistet. Die Konzentration der verwendeten Oligonukleotide wurde immer auf 100 µM eingestellt.

**Tab. 6: DNA-Bank.** Zufällige Positionen (N=A,T,G,C) sind blau gekennzeichnet

Name	Sequenz ( 5' → 3' )	Länge (nt)	Hersteller
htmpi	(C <sub>6</sub> NH <sub>2</sub> )-TGACACCGTACCTGCTCTATGCCTAT CATGGGCAAGCACGCCAGGGACGTCCCTGGCG TGCTT	63	Purimex
htmpc	(C <sub>6</sub> NH <sub>2</sub> )-TGACACCGTACCTGCTCTTGCAATAC CAGGCTTAAGCACGCCAGGGACGTCCCTGGCG TGCTT	63	Purimex
H <sub>2</sub> NhpSOLiC6V16	(C <sub>6</sub> NH <sub>2</sub> )-CTGCTGTACGGCCAAGGGCGNNCNN CNNCNNCNNCNNATCACC GACTGCCCATAGA GCTCTATGGGCAGTCGGTGAT	77	Purimex
H <sub>2</sub> NhpIIIc6V16	(C <sub>6</sub> NH <sub>2</sub> )-GTGACTGGAGTTCAGACGTTGCNNCNN CNNCNNCNNCNNAGATCGGAAGAGCGTCGTG TACACGACGCTCTTCCGATCT	79	Purimex
SOLiB30	CCTCTCTATGGGCAGTCGGTGATNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNCGCCTTGGCCGTA CAGCAG	72	Purimex

**Tab. 7: Ligationsadapter.** Abasische Positionen (X) sind in grün und zufällige Positionen (N=A,T,G,C) in blau markiert

Name	Sequenz ( 5' → 3' )	Länge (nt)	Hersteller
aAAG	PO <sub>4</sub> -AAG-X- NGNNGCGGTGGAAGTATATGTACCT	29	Purimex
aACG	PO <sub>4</sub> -ACG-X- NGNNGGATGTGAAGTGTGTATCCAC	29	Purimex
aAGG	PO <sub>4</sub> -AGG-X- NGNNGGAGCGTGTAATCCGATCTAA	29	Purimex
aATG	PO <sub>4</sub> -ATG-X- NGNNGGCCGTGTTACGACCCGTTAA	29	Purimex
aCAG	PO <sub>4</sub> -CAG-X- NGNNGGCACTATATGCACAGCTCGA	29	Purimex
aCCG	PO <sub>4</sub> -CCG-X- NGNNGCGGATAAAGACACAGTTCGC	29	Purimex
aCGG	PO <sub>4</sub> -CGG-X- NGNNGGCTTCATTACTACTATGCGC	29	Purimex
aCTG	PO <sub>4</sub> -CTG-X- NGNNGCGCGTAACTTTATCTAAGGC	29	Purimex
aGAG	PO <sub>4</sub> -GAG-X- NGNNGCCTTACACATACCTAGTGAG	29	Purimex

aGCG	PO <sub>4</sub> -GCG-X- NGNNGCGACAATACTTGACAGCACG	29	Purimex
aGGG	PO <sub>4</sub> -GGG-X- NGNNGCGAAGCGTATGGAAGCTACC	29	Purimex
aGTG	PO <sub>4</sub> -GTG-X- NGNNGCTGAGAATATGAGAGACGCC	29	Purimex
aTAG	PO <sub>4</sub> -TAG-X- NGNNGCTCACTAAGTTCATAGACCG	29	Purimex
aTCG	PO <sub>4</sub> -TCG-X- NGNNGCCAGAAGGATATACACGCGC	29	Purimex
aTGG	PO <sub>4</sub> -TGG-X- NGNNGCCTAAGACCGTAGAATAGCC	29	Purimex
aTTG	PO <sub>4</sub> -TTG-X- NGNNGACCTGCGTAAAGGATAGCAC	29	Purimex

Tab. 8: Primer

Name	Sequenz ( 5' → 3' )	Länge (nt)	Hersteller
Pro3	ATAGTCCCTGGCGTGCTT	18	Sigma-Aldrich®
Pro5	TGACACCGTACCTGCTCT	18	Sigma-Aldrich®
SOLi3GCG	GCTGTACGGCCAAGGCGGCG	20	Sigma-Aldrich®
SOLi3CGG	GCTGTACGGCCAAGGCGCGG	20	Sigma-Aldrich®
SOLi5GCG	TATGGGCAGTCGGTGATGCG	20	Sigma-Aldrich®
SOLi5CGG	TATGGGCAGTCGGTGATCGG	20	Sigma-Aldrich®
SOLiB3	CTGCTGTACGGCCAAGGCG	19	Sigma-Aldrich®
SOLiB5	CCTCTCTATGGGCAGTCGGTGAT	23	Sigma-Aldrich®
P3Illumina	ACACGACGCTCTTCCGATCT	20	BioTeZ
P5Illumina	GTGACTGGAGTTCAGACGTTG	21	BioTeZ
P3IIIGCG	CGACGCTCTTCCGATCTGCG	20	BioTeZ
P3IIIGGG	CGACGCTCTTCCGATCTGGG	20	BioTeZ
P3IIIAAG	CGACGCTCTTCCGATCTAAG	20	BioTeZ
P3IIITTG	CGACGCTCTTCCGATCTTTG	20	BioTeZ
P3IIICGG	CGACGCTCTTCCGATCTCGG	20	BioTeZ
P3IIICAG	CGACGCTCTTCCGATCTCAG	20	BioTeZ

Tab. 9: Sequenzierungsprimer. Barcodesequenzen sind in orange gekennzeichnet

Name	Sequenz ( 5' → 3' )	Länge (nt)	Plattform	Hersteller
MID1fS	CGTATCGCCTCCCTCGCGCCATCAGACGAGT GCGTACACGACGCTCTTCCGATCT	55	Roche (454)	Metabion
MID2fS	CGTATCGCCTCCCTCGCGCCATCAGACGCTC GACAACACGACGCTCTTCCGATCT	55	Roche (454)	Metabion
MID3fS	CGTATCGCCTCCCTCGCGCCATCAGAGACGC ACTCACACGACGCTCTTCCGATCT	55	Roche (454)	Metabion
MID4rS	CTATGCGCCTTGCCAGCCCGCTCAGAGCACT GTAGGTGACTGGAGTTCAGACGTTG	56	Roche (454)	Metabion
MID5rS	CTATGCGCCTTGCCAGCCCGCTCAGATCAGA CACGGTACTGGAGTTCAGACGTTG	56	Roche (454)	Metabion

MID6rS	CTATGCGCCTTGCCAGCCCGCTCAGATATCG CGAGGTGACTGGAGTTCAGACGTTG	56	Roche (454)	Metabion
--------	--------------------------------------------------------------	----	-------------	----------

**Tab. 10: Hybridisierungsoligonukleotide**

Name	Sequenz ( 5' → 3' )	Länge (nt)	Hersteller
hybGC	H <sub>2</sub> N-CGTGCTGTCAAGTATTGTCG	20	Purimex
hybCG	H <sub>2</sub> N-GCGCATAGTAGTAATGAAGC	20	Purimex
hybAAG	biotin-AGGTACATATACTTCCACCG	20	Purimex
hybACG	biotin-GTGGATACACACTTCACATC	20	Purimex
hybAGG	biotin-TTAGATCGGATTACACGCTC	20	Purimex
hybATG	biotin-TTAACGGGTCGTAACACGGC	20	Purimex
hybCAG	biotin-TCGAGCTGTGCATATAGTGC	20	Purimex
hybCCG	biotin-GCGAACTGTGTCTTTATCCG	20	Purimex
hybCGG	biotin-GCGCATAGTAGTAATGAAGC	20	Purimex
hybCTG	biotin-GCCTTAGATAAAGTTACGCG	20	Purimex
hybGAG	biotin-CTCACTAGGTATGTGTAAGG	20	Purimex
hybGCG	biotin-CGTGCTGTCAAGTATTGTCG	20	Purimex
hybGGG	biotin-GGTAGCTTCCATACGCTTCG	20	Purimex
hybGTG	biotin-GGCGTCTCTCATATTCTCAG	20	Purimex
hybTAG	biotin-CGGTCTATGAACTTAGTGAG	20	Purimex
hybTCG	biotin-GCGCGTGTATATCCTTCTGG	20	Purimex
hybTGG	biotin-GGCTATTCTACGGTCTTAGG	20	Purimex
hybTTG	biotin-GTGCTATCCTTTACGCAGGT	20	Purimex

**Tab. 11: Fluoreszenzoligonukleotid**

Name	Sequenz ( 5' → 3' )	Länge (nt)	Hersteller
Cy5aGCG	(Cy5)CGACAATACTTGACAGCACG	20	Purimex



## 2.2 METHODEN

### 2.2.1 NECLYS (NUCLEIC ACID-ENCODED CHEMICAL LIBRARY AND SELECTION, SYNTHESE UND SELEKTION VON NUKLEINSÄURE-KODIERTEN CHEMISCHEN WIRKSTOFFBANKEN)

Die Entwicklung einer neuen Methode, mit welcher niedermolekulare Wirkstoffbanken durch kombinatorische chemische Synthese als Konjugate an dafür kodierenden Nukleinsäuren hergestellt werden (Neclys), erfolgte halbautomatisiert mit dem Magnetpartikelprozessor KingFisher Flex von Thermo Scientific®. Im Laufe der Arbeit wurden die einzelnen Arbeitsschritte für eine effizientere Vorgehensweise auf den Roboter übertragen.

### 2.2.2 ROBOTER KINGFISHER FLEX

Um die Synthese und Selektion von Nukleinsäure-kodierten chemischen Wirkstoffbanken zu parallelisieren, effizienter und reproduzierbarer zu gestalten, wurden einzelne Arbeitsschritte wie die Separation, die Synthesereaktion, die Fmoc-Entschützung sowie die anschließende Elution auf den Magnetpartikelprozessor KingFisher Flex übertragen.

Der Roboter stellt eine vielseitige offene Plattform für das automatisierte Arbeiten mit magnetischen Partikeln dar. Dabei können die Protokolle selbst gefertigt und dem jeweiligen Arbeitsanliegen frei angepasst werden. Der Roboter ist aus  $8 \times 12$  Stabmagneten aufgebaut, die am oberen Ende miteinander verbunden und durch einen Schwenkkopf als Gesamtes frei beweglich sind. Insgesamt gibt es sieben Steckplätze für Mikrotiterplatten, wobei immer eine so genannte Tip-Platte, welche die Plastikhülle für die Magneten beinhaltet, mitgeführt wird.

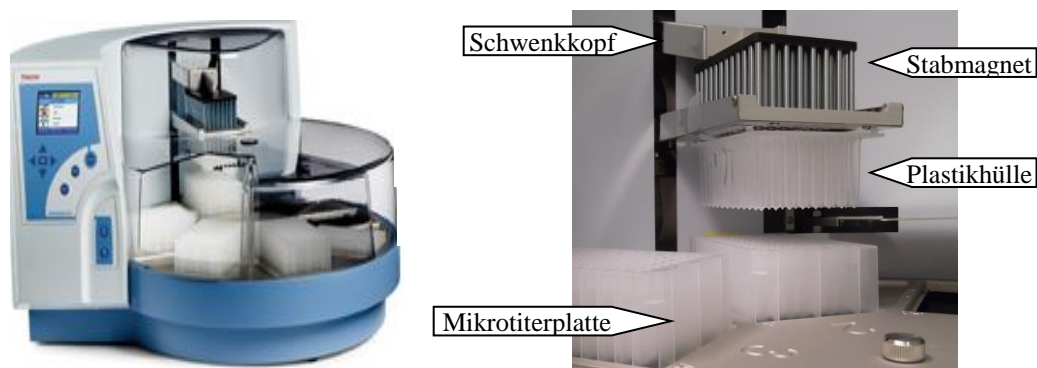


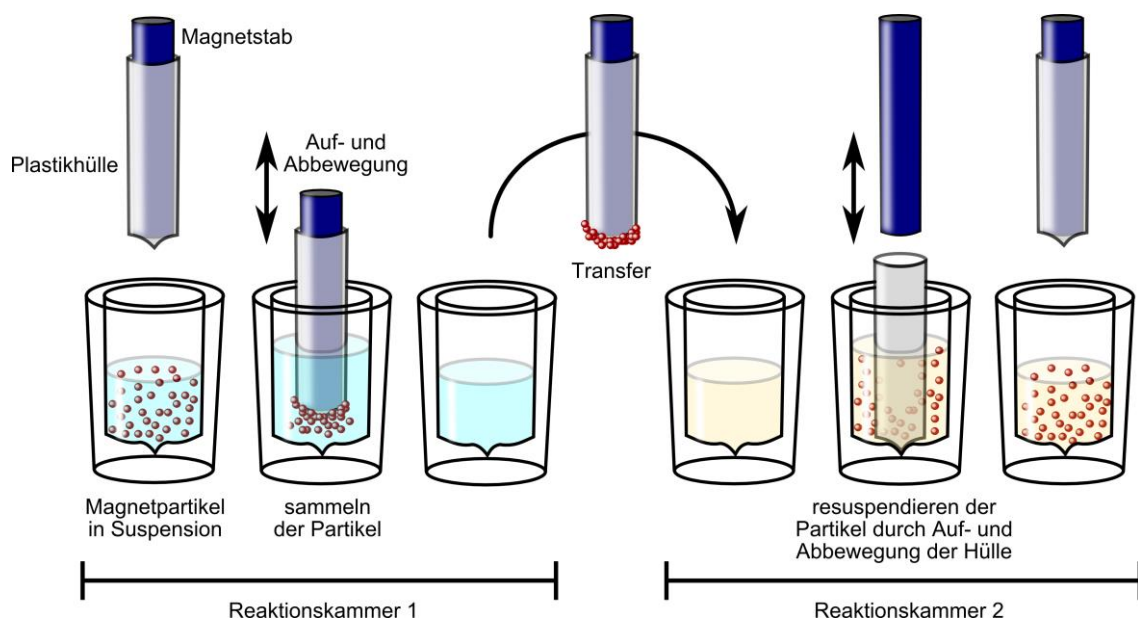
Abb. 10: Außen- und Innenansicht des KingFisher Flex Roboters.

Durch den Magnetkopf kann im 96er Mikrotiterplattenformat gearbeitet werden, wodurch eine hohe Parallelisierung erreicht wird.

Durch die Magnetstäbe können magnetische Partikel an die Stäbe separiert und von einer Reaktionskammer in die nächste Kammer einer Mikrotiterplatte transferiert werden. So können die Partikel voll automatisch in verschiedenen Lösungen, unter verschiedenen Reaktionsbedingungen sowie Zeiten inkubiert werden.

Der Arbeitsablauf des Transfers von magnetischen Partikeln von einer Kammer in die nächste ist in Abb. 11 schematisch dargestellt. Zunächst wird der Magnetstab von einer Plastikhülle umgeben. Anschließend kann der Magnetstab in die Reaktionskammer, in der sich die magnetischen Partikel in Lösung befinden, eingetaucht werden.

Dort werden die Partikel durch Auf- und Abbewegungen des Magneten an die Stäbe gebunden und können in eine neue Reaktionskammer transferiert werden. In der neuen Lösung können die Partikel durch schnelle Bewegungen der Plastikhülle resuspendiert werden. Dazu sind allerdings vorher die Magnetstäbe aus der Hülle zu ziehen. Des Weiteren sind die verschiedensten Reaktionsbedingungen möglich, z.B. die Partikel an den Magnetstäben gebunden zu lassen und durch Bewegungen in der Suspension zu inkubieren.



**Abb. 11: Funktionsweise des KingFisher Flex.** Transfer von magnetischen Partikeln von einer Reaktionskammer in die nächste (nach QIAGEN).

### 2.2.3 KOPPLUNG VON NUKLEINSÄURE AN MIKROPARTIKEL

In dieser Arbeit wurden verschiedene Arten von Mikropartikeln verwendet. Für die Kopplung von modifizierten Nukleinsäuren an magnetische Partikel wurden die Dynabeads<sup>®</sup> M-280 Streptavidin (Ø 2,8 µm, 10 mg/ml), die Dynabeads<sup>®</sup> MyONE Carboxylic Acid (Ø 1 µm, 10 mg/ml), beides von Invitrogen<sup>™</sup> oder die SiMAG-Carboxyl (Ø 0,5 µm, 50 mg/ml) von Chemicell verwendet. Für die ionische Bindung von Nukleinsäure an magnetische Partikel wurden die SiMAG-DEAE Mikropartikel (Ø 0,5 µm, 50 mg/ml) von Chemicell benutzt.

#### 2.2.3.1 KOPPLUNG VON NUKLEINSÄUREN AN MAGNETISCHE MIKROPARTIKEL

Für die kovalente Kopplung von 5'-aminomodifizierten Oligonukleotiden wurden magnetische Partikel mit carboxylierter Oberfläche nach leicht modifizierten Herstellerangaben verwendet.

Bei der Immobilisierung mittels SiMAG-Carboxyl (Chemicell) wurden 200 µl Magnetpartikelsuspension zweimal mit je 1 ml 2-(N-Morpholino)ethansulfonsäure (MES, 0,1 M, pH 5) für 5 min gewaschen. Die Separation erfolgte mithilfe des Magneten DynaMag von Invitrogen<sup>™</sup>. Nach dem zweiten Waschschrift wurden die Magnetpartikel in 250 µl 0,1 M MES (pH 5) mit 10 mg frisch hinzugefügtes 1-Ethyl-3-(3-Dimethylaminopropyl) carbodiimid (EDC) resuspendiert und für 10 min bei RT (22°C) auf dem Rotator Stuart SB3 von BioCote<sup>®</sup> inkubiert. Anschließend wurden 50 µg der 5'-aminomodifizierten DNA zu der Reaktion hinzugefügt und die Immobilisierung für 2 h bei RT auf dem Rotator inkubiert. Nach dreimaligem Waschen mit 1 ml Lagerungspuffer (1× PBS, 0,1% Tween<sup>®</sup> 20, 10 mM Ethanolamin) wurden die Partikel in 250 µl des Puffers resuspendiert und bei 4°C gelagert. Für den Nachweis der Immobilisierung wurden jeweils 6,25 µl (0,3 mg) Partikel vom Lagerungspuffer separiert und in 50 µl Hybridisierungspuffer (750 mM NaCl, 75 mM Natriumcitrat, 0,1% Tween<sup>®</sup> 20) für 30 min bei 40°C und 800 rpm im Thermomixer compact 5355 von Eppendorf inkubiert. Dem Ansatz wurden zur Fluoreszenzdetektion 1 µl (5 pmol) komplementäres Cy5-gekoppeltes Oligonukleotid hinzugefügt. Anschließend wurde der Überstand abgenommen und die Partikel zweimal mit 1× TE-Puffer (150 µl) gewaschen und in 25 µl dieses Puffers resuspendiert. Die Auswertung erfolgte mit dem Laser Scanner Affimetrix 428 von Sartorius. Zu diesem Zweck wurden 2 µl der Partikelsuspension und des

zuvor abgenommenen Überstandes auf einen Glasobjektträger aufgetragen und nach etwas Trocknungszeit gescannt.

Bei der kovalenten Kopplung mit Dynabeads<sup>®</sup> MyONE Carboxylic Acid von Invitrogen<sup>™</sup> wurden 20 µl (0,2 mg) von der Partikelsuspension genommen und zweimal mit 200 µl MES (0,1 M, pH 5) gewaschen und anschließend mit dem Magneten separiert und in 100 µl des MES-Puffers resuspendiert. In einem separaten Reaktionsgefäß wurden zu 40 µl Inkubationspuffer (0,1 M MES (pH 5), 1,25 M EDC) 1 nmol des zu immobilisierenden Oligonukleotids hinzugefügt und alles mit Millipore-Wasser auf 100 µl aufgefüllt. Anschließend wurden die 100 µl der Partikelsuspension und der Nukleinsäure vereinigt, für 10 s gevortext und alles auf dem Rotator Stuart SB3 entweder für 3 h oder Übernacht bei RT (22°C) inkubiert. Die Kopplung wurde dann zweimal mit 200 µl TT-Puffer (0,25 M Tris-HCl (pH 8), 0,01% Tween<sup>®</sup> 20) und einmal mit TT-Puffer, 10 mM Ethanolamin für jeweils 30 min gewaschen. Nach den Waschschrritten wurden die Partikel separiert anschließend in 20 µl 1× TE-Puffer resuspendiert und bei 4°C gelagert. Nach Herstellerangaben werden von 1 mg Partikeln 1,8 nmol Oligonukleotid kovalent gekoppelt. Für den Nachweis der Immobilisierung wurden jeweils 5 µl Partikelsuspension (0,05 mg Partikel) gekoppelt mit circa 0,1 nmol Nukleinsäure vom Lagerungspuffer separiert und in 50 µl Hybridisierungspuffer (750 mM NaCl, 75 mM Natriumcitrat, 0,1% Tween<sup>®</sup> 20), welcher ein komplementäres Fluoreszenz-oligonukleotid in äquimolaren Mengen enthielt, für 30 min bei 40°C und 800 rpm im Thermomixer inkubiert. Anschließend wurde der Überstand abgenommen und die Partikel zweimal mit 1× TE-Puffer (150 µl) gewaschen und in 5 µl dieses Puffers resuspendiert. Die Auswertung erfolgte mit dem Laser Scanner Affimetrix 428. Zu diesem Zweck wurden 3 µl der Partikelsuspension und des zuvor abgenommenen Überstandes auf einen Glasobjektträger aufgetragen und nach etwas Trocknungszeit gescannt.

Für die Kopplung von 5'-biotinylierter Nukleinsäure wurden die Dynabeads<sup>®</sup> M-280 Streptavidin (Invitrogen<sup>™</sup>) verwendet. Auch hier erfolgte die Immobilisierung weitestgehend nach den Herstellerangaben.

Es wurden 100 µl (1 mg) der Partikelsuspension mit 100 µl 2× B&W-Puffer (10 mM Tris-HCl (pH 7,5), 1 mM EDTA, 2 M NaCl) gewaschen, mit dem Magneten separiert und anschließend in 200 µl 2× B&W-Puffer resuspendiert. Anschließend wurden 200 pmol Oligonukleotid, gelöst in 200 µl Millipore-Wasser, zu den Partikeln gegeben und bei RT (22°C) für 30 min auf dem Rotator Stuart SB3 inkubiert. Nach der Kopplung wurden die

Partikel zweimal mit 400  $\mu\text{l}$  1 $\times$  B&W-Puffer gewaschen, anschließend in 100  $\mu\text{l}$  1 $\times$  TE-Puffer resuspendiert und bei 4°C aufbewahrt. Nach Herstellerangaben sind so an 1 mg Streptavidinpartikel 200 pmol Hybridisierungsoligonukleotid gekoppelt.

Für den Nachweis der Immobilisierung wurden jeweils 10  $\mu\text{l}$  Partikelsuspension (0,1 mg Partikel) gekoppelt mit 20 pmol Nukleinsäure vom Lagerungspuffer separiert und in 50  $\mu\text{l}$  Hybridisierungspuffer (1,5 M NaCl, 150 mM Natriumcitrat, 0,1% Tween<sup>®</sup> 20), welcher ein komplementäres Fluoreszenzoligonukleotid in äquimolaren Mengen enthält, für 30 min bei 46°C und 800 rpm im Thermomixer inkubiert. Anschließend wurde der Überstand abgenommen und die Partikel zweimal mit 1 $\times$  TE-Puffer (100  $\mu\text{l}$ ) gewaschen und in 10  $\mu\text{l}$  dieses Puffers resuspendiert. Die Auswertung erfolgte mit dem Laser Scanner Affimetrix 428. Zu diesem Zweck wurden 3  $\mu\text{l}$  der Partikelsuspension und des zuvor abgenommenen Überstandes auf einen Glasobjektträger aufgetragen und nach etwas Trocknungszeit gescannt.

#### **2.2.3.2 IONISCHE BINDUNG VON NUKLEINSÄURE AN MAGNETISCHE MIKROPARTIKEL**

Bei den SiMAG-DEAE Mikropartikeln von Chemicell handelt es sich um schwache Anionenaustauscher, die in dieser Arbeit verwendet wurden, um die ligierten und sortierten Oligonukleotide nach der Spaltungsreaktion an eine feste Phase zu binden, so dass anschließend die Synthese durchgeführt werden konnte.

Nach der Spaltung lag das zu bindende Oligonukleotid in 1 $\times$  NEB3-Puffer (New England BioLabs<sup>®</sup>) mit einem Volumen von 30  $\mu\text{l}$  vor. Zu diesem Ansatz wurden 120  $\mu\text{l}$  Millipore-Wasser hinzugefügt, so dass sich die Salzkonzentration stark verringerte und ein Endvolumen von 150  $\mu\text{l}$  entstand. Für jede Spaltungsreaktion wurden 10  $\mu\text{l}$  Partikelsuspension einmal mit 20 mM Tris-Puffer (pH 8) gewaschen und mit dem Magneten separiert. Anschließend wurden die Partikel und das Oligonukleotid vereinigt und bei RT (22°C) für 30 min auf dem Rotator Stuart SB3 inkubiert.

Nach der Bindung wurden die Beads einmal mit 150  $\mu\text{l}$  Tris-Puffer (20 mM, pH 8) gewaschen und in 50  $\mu\text{l}$  des selbigen Puffers resuspendiert. Die an den Partikeln gebundene Nukleinsäure konnte nun für die Synthesereaktion verwendet werden.

## 2.2.4 LIGATION

Eine Ligation ist die enzymkatalysierte Verknüpfung zweier DNA- oder RNA-Fragmente an ihren jeweiligen 3'-Hydroxy bzw. 5'-Phosphatenden. Das Enzym, welches diese Reaktion katalysiert ist die sogenannte Ligase. Die DNA-Ligasen spielen bei der Replikation, Rekombination und der Reparatur der DNA eine entscheidende Rolle (Lehman, 1974). Generell werden die DNA Ligasen in zwei Kategorien eingeteilt. Zum einen gibt es die ATP-abhängigen und zum anderen die NAD<sup>+</sup>-abhängigen Ligasen. Dabei sind alle bekannten eukaryotischen zellulären DNA Ligasen ATP-abhängig (Shuman, 2009). Die am häufigsten verwendete DNA Ligase ist die T4 DNA Ligase, welche die Bildung einer Phosphodiesterbindung zwischen einer freien 5'-Phosphat und einer 3'-Hydroxygruppe der Desoxyribose katalysiert (Kuhn & Frank-Kamenetskii, 2005).

Um die beste Effizienz und eine hohe Spezifität der Ligation zwischen der DNA-Bank und den Ligationsadaptern zu erhalten, wurden mehrere Methoden untersucht. Zum einen wurde die Ligrationsreaktion nach dem Standardprotokoll (einfache Ligation) zum anderen zyklisch (Wiedmann *et al.*, 1994) im Thermozykler mittels der T4 DNA Ligase (Roboklon) durchgeführt. Die DNA-Bank wurde in einem Verhältnis von 1:20 zu den Ligationsadaptern eingesetzt, um eine höchstmögliche Effizienz zu erhalten. Somit lag die DNA-Bank mit einer Endkonzentration von 4 µM und jeder Ligationsadapter mit 5 µM im Reaktionsansatz vor. Bei 16 verschiedenen Ligationsadaptern ergab sich somit ein 20-facher Überschuss zur DNA-Bank.

**Tab. 12: Zusammensetzung der einfachen und zyklischen Ligrationsreaktion**

	Volumen	Endkonzentration
DNA-Bank (100 µM)	4 µl	4 µM
16 Ligationsadapter (100 µM je Oligo)	5 µl je Adapter (80 µl)	5 µM je Adapter
10× T4 DNA Ligase Puffer	10 µl	1×
T4 DNA Ligase (268 U/µl)	5 µl	13,4 U/µl
Millipore-H <sub>2</sub> O	ad 100 µl	

Sowohl bei der einfachen als auch der zyklischen Ligation wurde das Pipettierschema in Tabelle 12 beibehalten. Bei der einfachen Ligation wurden die Proben zunächst bei 95°C für 3 min denaturiert und anschließend erfolgte bei 37°C die Anlagerung des Hairpins der DNA-

Bank. Nachdem die Proben auf die Inkubationstemperatur heruntergekühlt waren, wurde die T4 DNA Ligase hinzugefügt und die Inkubation erfolgte bei 22°C für 30 min. Anschließend folgte die Inaktivierung der Ligase bei 80°C für 10 min. Bei der zyklischen Ligation wurde der Denaturierungsschritt als auch die Inaktivierung beibehalten. Zwischen diesen Schritten, wurde allerdings zyklisch gearbeitet. Die Anlagerung des Hairpins als auch die Inkubation wurden in 10 Zyklen wiederholt. Dabei variierte die Anlagerungstemperatur zwischen 40°C und 45°C für 30 Sekunden bis 1 Minute. Die Inkubationstemperatur betrug konstant 22°C, die Zeit variierte aber zwischen 3 – 5 min.

Zudem wurde in einem weiteren Versuch überprüft, ob auch eine isothermale Ligation möglich wäre. Zu diesem Zweck wurde die zyklische Ligation zum Vergleich mit einbezogen. Die isothermale Ligation erfolgte bei 37°C für 45 min. Dabei wurde neben der T4 DNA Ligase (Roboklon) auch mit der *E. coli* Ligase (New England BioLabs®) gearbeitet. Ein weiteres Enzym, die SV40 Large T Antigen (Helikase) von Roboklon, sollte die Effizienz der Ligation noch steigern. Dazu mußte dieses Enzym zusätzlich zu den jeweiligen Ligasen hinzugefügt werden. Die DNA-Bank wurde in dieser Versuchsreihe in einem Verhältnis von 1:10 zu den Ligationsadaptern eingesetzt.

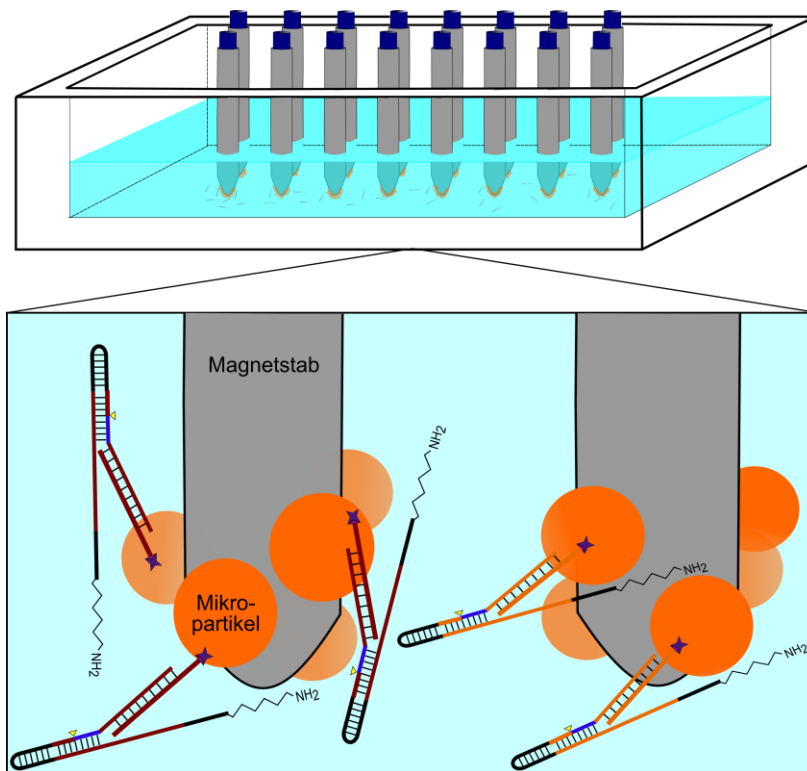
**Tab. 13: Zusammensetzung der zyklischen und isothermalen Ligationsreaktion**

	Volumen [µl]						Endkonzentration
	T4 DNA Ligase Ansätze			<i>E. coli</i> Ligase Ansätze			
DNA-Bank (100 µM)	0,2	0,2	0,2	0,2	0,2	0,2	0,8 µM
16 Ligationsadapter (100 µM je Adapter)	0,125 (16×)	0,125 (16×)	0,125 (16×)	0,125 (16×)	0,125 (16×)	0,125 (16×)	0,5 µM je Adapter
10× T4 DNA Ligase Puffer	2,5	2,5	2,5				1×
T4 DNA Ligase (268 U/µl)	0,5	0,5					5,36 U/µl
10× <i>E. coli</i> Ligase Puffer				2,5	2,5	2,5	1×
<i>E. coli</i> Ligase (10 U/µl)				0,5	0,5		0,2 U/µl
SV40 Large T Antigen (1 µg/µl)	0,1		0,1	0,1		0,1	4 ng/µl
Millipore-H <sub>2</sub> O	ad 25						

Bei dieser Versuchsreihe wurde die zyklische Ligation (Ablauf siehe oben) mit einer isothermalen Ligation unter dem Einsatz verschiedener Enzyme verglichen. Der Nachweis der Ligationen erfolgte entweder in Agarose- oder Polyacrylamidgelen.

## 2.2.5 SEPARATION

Die ligierte DNA-Bank wurde mittels magnetischen Mikropartikeln separiert. An diesen Mikropartikeln befinden sich gekoppelte Hybridisierungsoligonukleotide (siehe Abschnitt 2.2.3.1). Diese Oligonukleotide hybridisieren mit der Adaptersequenz der ligierten DNA-Bank, wodurch diese hinsichtlich ihrer Sequenz sortiert werden können. Insgesamt ergaben sich somit 16 verschiedene Sortierungsvarianten. Dieser spezifische Hybridisierungsschritt erfolgte nach mehreren Vorversuchen auf dem Magnetpartikelprozessor KingFisher Flex von Thermo Scientific®. Dadurch ist jede Variante der immobilisierten Hybridisierungsoligonukleotide durch die magnetischen Mikropartikel an einer festen Position verankert. Zum genauen Verständnis dient die Abbildung 12.



**Abb. 12: Separation der ligierten DNA-Bank an magnetische Mikropartikel.** Die Hybridisierung erfolgte in einem Reaktionsgefäß (siehe Abbildung oben) an 16 verschiedenen Magnetstäben, welche mit magnetischen Mikropartikeln bestückt waren. Durch die Hybridisierungsoligonukleotide erfolgte eine Sortierung der ligierten DNA-Bank aufgrund deren komplementären Adaptersequenzen (siehe Abbildung unten).

Bevor allerdings dieser Schritt auf den Roboter übertragen werden konnte, erfolgte in einer Reihe von Vorversuchen die genaue Analyse dieses Prozesses. Es wurde neben den Pufferbedingungen sowie der Inkubationszeit und -temperatur auch die Hybridisierungseffizienz überprüft. Dabei variierte die Natriumchloridkonzentration des



Puffers zwischen 500 mM – 1,5 M und die Natriumcitratkonzentration zwischen 50 mM – 150 mM. Lediglich die Zugabe von 0,1% Tween<sup>®</sup> 20 wurde konstant gehalten. Diesem Puffersystem wurde zusätzlich noch 1-Methyl-2-Pyrrolidon (99%) mit einer Endkonzentration von 5 – 10 % hinzugefügt. Ein zusätzlicher Puffer (200 mM Tris (pH 7,4), 50 mM MgCl<sub>2</sub>, 20 mM CaCl<sub>2</sub>, 20 mM KCl, 1,4 mM NaCl, 0,5% Tween<sup>®</sup> 20) wurde für die Hybridisierung auch in Betracht gezogen. Die Temperatur variierte zwischen RT (22°C) und 46°C. Letztendlich wurden Inkubationszeiten von 20 min – 1 Stunde untersucht. Die Auswertung über die Qualität und Quantität der Separation erfolgte mit dem Laser Scanner Affimetrix 428 von Sartorius. In diesem Fall diente ein Fluoreszenzoligonukleotid, welches mit dem Farbstoff Cy5 gekoppelt war, als Hybridisierungsoligo und die Inkubation erfolgte im Thermomixer compact 5355 von Eppendorf bei 800 rpm. Im Anschluss wurden jeweils Proben abgenommen und auf einen Glasobjektträger aufgetragen und zum Scannen leicht getrocknet. Durch den Einsatz einer Titrationsreihe des Fluoreszenzoligos konnte zudem die Kopplungseffizienz der immobilisierten Oligonukleotide bestimmt werden. Die Nachweismethode mittels Fluoreszenzoligo wurde für die unterschiedlichen Mikropartikel, wie unter Abschnitt 2.2.3.1 beschrieben durchgeführt.

Nach Abschluss dieser Vorversuche konnte das manifestierte Protokoll auf den KingFisher Flex übertragen werden. Bevor das endgültige Protokoll festgelegt war, erfolgten auch hier zunächst einige Versuche, um die besten Parameter am Roboter zu bestimmen. Die Separation erfolgte nun mit einer definierten ligierten DNA-Bank.

Wie unter Ligation (2.2.4) beschrieben, wurden 400 pmol DNA-Bank eingesetzt und die Ligationsadapter im 20-fachen Überschuss zur Bank verwendet. Somit ergaben sich 16 verschiedene Ligationsvarianten mit einer Konzentration von jeweils 25 pmol. In die 96er Deepwell-Platte (Thermo Scientific<sup>®</sup>, Vantaa, U design, 2 ml) wurden 16 Kammern mit den Mikropartikeln, welche die kovalent gekoppelten Hybridisierungsoligonukleotide trugen, beladen (siehe Abb. 13).

H	G	F	E	D	C	B	A	
hybAAG	hybACG	hybAGG	hybATG	hybCAG	hybCCG	hybCCG	hybCTG	1
hybGAG	hybGCG	hybGGG	hybGTG	hybTAG	hybTCG	hybTGG	hybTTG	2

**Abb. 13. Schematische Darstellung von der Beladung der Bindungsplatte.** In alle 16 Kammern wurden 50 pmol vom Hybridisierungsoligonukleotid, gekoppelt an magnetische Mikropartikel (0,25 mg), geladen.

Die Hybridisierungsoligonukleotide wurden wieder im doppelten Überschuss zu der Ligation eingesetzt (50 pmol), so dass auch wirklich die gesamte ligierte DNA-Bank separiert werden

konnte. Anschließend wurde die ligierte DNA-Bank in 10 ml des Hybridisierungspuffers (1,5 M Natriumchlorid, 150 mM Natriumcitrat, 0,1% Tween<sup>®</sup> 20) in einer 6er Deepwell-Platte (10 ml, MPI-Werkstatt) vorgelegt. Die Platten wurden anschließend in den Roboter geladen und die Reaktion gestartet. Die Mikropartikel wurden dabei an den Magnetstab geankert und in die 6er Deepwell-Platte überführt. Die Reaktion wurde für 1 h bei 46°C durch Auf- und Abbewegung inkubiert (siehe Abb.10). Im Anschluss an die Separation wurden die Mikropartikel mit der sortierten DNA-Bank in 100 µl 0,1 M 2-(N-Morpholino)ethansulfonsäure (pH 5) für 1 min gewaschen und in 100 µl 0,1 M MES (pH 5) einer neuen Mikrotiterplatte eluiert. Im Anschluss kann mit der sortierten DNA-Bank die Spaltungsreaktion durchgeführt werden. Das Kingfisher Flex Protokoll zur Separation ist im Anhang nachlesbar. Der Nachweis der Separation erfolgte mittels spezifischer Amplifikationsverfahren und Detektion per Agarosegelelektrophorese oder auch zur Qualitätskontrolle per Sanger-Sequenzierung (siehe Abschnitt 2.2.13). Die 6er Deepwell-Platte (10 ml, MPI-Werkstatt) wurde nach einem Durchgang der Separation recycelt. Dazu mußte die Platte mit 5%-igen Wasserstoffperoxid befüllt und über Nacht einwirken gelassen werden. Am nächsten Tag wurde die Platte mehrmals mit Millipore-Wasser gespült und für ihre Weiterverwendung getrocknet.

### 2.2.6 SPALTUNG

Bis zu diesem Schritt ist die ligierte DNA-Bank mittels einem Hybridisierungsoligonukleotids hinsichtlich ihrer Sequenz sortiert worden und an magnetischen Mikropartikeln fest geankert. Für eine zweite Ligationsrunde, muss das Oligonukleotid an seiner abasischen Position gespalten werden, so dass anschließend eine erneute Ligation und Separation erfolgen kann. Als Ergebnis sollte die DNA-Bank nach der Spaltung um drei Nukleotide verlängert sein und in der Lösung frei vorliegen. Die Synthese erfolgte erst nach der Spaltung, da das Oligonukleotid für die Synthese mittels einer ionischen Bindung an Mikropartikeln geankert werden sollte (siehe Abschnitt 2.2.3.2). In dieser Arbeit wurden zum einen die enzymatische und zum anderen die chemische Spaltung untersucht. Dabei sollte ermittelt werden, welche Methode effizienter ist und somit für die Spaltungsreaktion eher in Frage kommt.

Die chemische Spaltung wurde zum einen mit 0,05 M Natriumhydroxid und zum anderen mit 2 mM EDTA vorgenommen. Hierbei ist es wichtig zu erwähnen, dass im Vordergrund lag,

das Oligo von den Partikeln zu lösen, so dass diese frei in der Lösung vorliegen und die Synthese durchgeführt werden kann. Dabei sollte vor allem die Effizienz der Rückgewinnung des Oligos von den Partikeln sowohl chemisch als auch enzymatisch verglichen werden. Die sortierten Partikel wurden in ein neues 0,2 ml Reaktionsgefäß überführt, mittels eines magnetischen Separators DynaMag (Invitrogen™) separiert und der Überstand verworfen. Anschließend wurden die Mikropartikel mit der gebundenen DNA-Bank zum einen mit 0,05 M NaOH und zum anderen mit 2 mM EDTA in einem Endvolumen von 30 µl bei 70°C für 10 min im Thermozykler PTC-200 (Biozyme®) inkubiert.

Bei der enzymatischen Spaltung wurden gleich zwei Ansätze verfolgt. Zum einen wurde das Oligo von den magnetischen Partikeln gelöst und zum anderen gleich das Fragment an seiner abasischen Position gespalten, so dass direkt ein neuer Durchgang gestartet werden könnte. Als das geeignete Enzym erweist sich dabei die Endonuklease IV von *E. coli* (New England BioLabs®). Die Endonuklease IV ist an der DNA-Reparatur beteiligt. Um genau zu sein, initiiert dieses Enzym die Reparatur von abasischen Positionen in der DNA durch Spaltung des DNA-Gerüsts an der 5'-Position der abasischen Stelle (Puri *et al.*, 2013). Dabei handelt es sich um eine apurinic/apyrimidinic (AP) Endonuklease, welche in der Lage ist, die DNA an der AP-Stelle zu hydrolysieren und dabei ein 5'-Phosphat- und ein 3'-Hydroxylende zu generieren (Hosfield *et al.*, 1999). Da dieses Enzym spezifische Spaltungsfragmente generiert, wurde es für die Spaltungsreaktion in dieser Arbeit verwendet. Die Mikropartikel mit der sortierten DNA-Bank wurden auch hier in ein neues 0,2 ml Reaktionsgefäß überführt und der Überstand verworfen. Anschließend wurden zu der DNA-Bank zwischen 30 U – 50 U Endonuklease IV in 1×NEB Puffer 3 mit einem Endvolumen von 30 µl gegeben. Die Reaktionsansätze wurden vor der Zugabe des Enzyms bei 95°C für 3 min denaturiert und anschließend in 5 min die Temperatur auf 37°C (0,2°C/s) gesenkt. Nachdem eine Inkubationstemperatur von 22°C erreicht war, wurde das Enzym dem Ansatz beigefügt und die Reaktion für 1 h inkubiert. Danach wurde der Ansatz für 20 min auf 80°C erhitzt und so die Reaktion gestoppt. Anschließend wurden die Partikel von der Lösung separiert und die Rückgewinnung der ligierten DNA-Bank über eine PCR nachgewiesen. Diese wurde auch auf den Mikropartikeln durchgeführt, auf denen sich im besten Fall nach der Spaltung bzw. Lösung des Fragments keine Rückstände mehr befinden sollten. Die Detektion erfolgte mittels einer Agarosegelelektrophorese.

### 2.2.7 SYNTHESE

Die Synthesereaktion beinhaltet die Konjugation von niedermolekularen Stoffen durch kombinatorische chemische Synthese an kodierende Nukleinsäuren. Die DNA-Fragmente besitzen einen Barcode, wodurch die Identifizierung des gekoppelten Wirkstoffes ermittelt werden kann.

Das Verfahren zum Synthetisieren einer kombinatorischen DNA-Bank kann nicht nur für die Synthese von unterschiedlichen chemischen Banken verwendet werden. Es kann auch für das Screening für optimale Reaktionsbedingungen verwendet werden, da jedes Codon für spezielle Reaktionsbedingungen stehen kann. Das heißt zum einen können an die sortierte DNA-Bank unter denselben Reaktionsbedingungen verschiedene Reagenzien gekoppelt werden und zum anderen kann das gleiche Reagenz unter verschiedenen Bedingungen an die DNA synthetisiert werden, um so eine Syntheseoptimierung zu ermitteln. In dieser Arbeit wird die Synthesereaktion an dem Beispiel der Syntheseoptimierung dargestellt.

Nach der Spaltungsreaktion wurde die sortierte DNA-Bank an SiMAG-DEAE Mikropartikel (siehe 2.2.3.2) gebunden und in 50 µl 20 mM Tris-Puffer (pH 8) bis zur Synthesereaktion aufbewahrt. Es wurden zwei Syntheseyklen durchgeführt. In der ersten Runde wurde Glycin und in der zweiten d-Biotin an die DNA-Bank unter verschiedenen Bedingungen synthetisiert. Im Anschluss konnte eine Selektion gegen das Zielmolekül Streptavidin durchgeführt werden und mittels NGS-Sequenzierung eine Anreicherung eines Barcodes und somit einer Reaktionsbedingung ermittelt werden. Die Synthese selber erfolgte auf den Grundlagen einer Peptidsynthese, dabei kann die Aminogruppe der DNA-Bank mit der Carboxy-Gruppe des Glycins bzw. die entschützte Aminogruppe des Glycins mit der Carboxy-Gruppe des Biotins reagieren (Bayer, 1991).

Nach der Spaltung lagen 16 sortierte DNA-Varianten, sortiert nach ihrem Barcode vor. Jedem Barcode wurde im Folgenden eine Reaktionsbedingung zugeordnet, die für beide Syntheserunden aufrechterhalten wurde. Die zufällige Zuordnung ist in Tabelle 14 nach zu lesen.

**Tab. 14: Synthesebedingungen.** 16 verschiedene Synthesebedingungen. Jedem Codon ist eindeutig ein Kopplungsreagenz und ein Lösungsmittel zugeordnet.

Codon	Kopplungsreagenz	Lösungsmittel
aAAG	DMT-MM (10 mM)	DMF
aACG	DMT-MM (10 mM)	Methanol
aAGG	DMT-MM (10 mM)	Ethanol
aATG	DMT-MM (10 mM)	bidest H <sub>2</sub> O
aCAG	EDC (50 mM)	DMF
aCCG	EDC (50 mM)	Methanol
aCGG	EDC (50 mM)	Ethanol
aCTG	EDC (50 mM)	bidest H <sub>2</sub> O
aGAG	CDI (500 mM)	DMF
aGCG	CDI (500 mM)	Methanol
aGGG	CDI (500 mM)	Ethanol
aGTG	CDI (500 mM)	bidest H <sub>2</sub> O
aTAG	/	DMF
aTCG	/	Methanol
aTGG	/	Ethanol
aTTG	/	bidest H <sub>2</sub> O

Zur Vorbereitung der Synthese wurde Glycin mit einer Konzentration von 10 mM und d-Biotin mit 5 mM im entsprechenden Lösungsmittel DMF/Methanol/Ethanol oder bidest H<sub>2</sub>O hergestellt. Davon wurden 4 bzw. 8 µl in einem Gesamtvolumen von 100 µl Kopplungsreagenz eingesetzt, wodurch eine Endkonzentration des Glycins bzw. d-Biotins von 400 µM entstand. Die Kopplungsreagenzien setzten sich aus 10 mM DMT-MM, 50 mM EDC oder 500 mM CDI in DMF/Ethanol/Methanol oder bidest H<sub>2</sub>O zusammen. Somit ergaben sich 16 verschiedene Kopplungsreagenzien, die jeweils einem Codon zugeordnet waren. Zunächst wurde die sortierte DNA-Bank, gebunden an DEAE-Beads in eine 96er Deepwell-Platte (Thermo Scientific<sup>®</sup>, Vantaa, U design, 2 ml), beladen. In eine andere Deepwell-Platte wurden die 16 verschiedenen Synthesebedingungen in einem Endvolumen von 100 µl mit dem Glycin bzw. d-Biotin äquivalent zu dem Barcode befüllt. Die Synthese erfolgte bei RT (22°C) für 30 min. Anschließend wurde den Mikropartikeln mit der DNA-Bank ein dreimaliger Waschschrift mit 200 µl 20 mM Tris-Puffer (pH 8) in Deepwell-Platten unterzogen (siehe Anhang: Synthese). Nach der Synthese mit Glycin wurden die Mikropartikel nach den Waschschriften direkt in eine neue Mikrotiterplatte überführt und die Schutzgruppenabspaltung (siehe 2.2.8) durchgeführt. Nach der Synthese mit d-Biotin hingegen konnte direkt die Elution von den DEAE-Partikeln (siehe 2.2.9) erfolgen.

### 2.2.8 FMOC-SCHUTZGRUPPENABSPALTUNG

Während der Synthese wurde die Aminosäure Glycin an die DNA-Bank synthetisiert. Für eine neue Syntheserunde musste die Fmoc-Schutzgruppe der gekoppelten Aminosäure zunächst entfernt werden, um so ein neues Molekül an die DNA-Bank synthetisieren zu können. Diese Schutzgruppe macht die funktionelle Aminogruppe des Glycins temporär gegenüber Reagenzien unreaktiv, durch die sie normalerweise umgewandelt werden würde. Gleichzeitig kann die Carbonsäurefunktion der Aminosäure zur Kopplung aktiviert werden. Ohne diese Schutzgruppe würde ein Homopolymer des Glycins unbekannter Länge entstehen. Die Reaktion wurde auf dem KingFisher Flex Roboter in 96er KF Mikrotiterplatten (Thermo Scientific<sup>®</sup>, U design, 200µl) durchgeführt. Zu diesem Zweck wurden die Mikropartikel, an denen die reaktive DNA-Bank gebunden war, aus einer 96er Deepwell-Platte in 200 µl 20% Piperidin in DMF für 3 min inkubiert. Im Anschluss wurden die Partikel in frische 200 µl 20% Piperidin in DMF überführt und dieses Mal für 17 min inkubiert. Die Partikel wurden dann zweimal für 30 min mit 100 µl 20 mM Tris-Puffer (pH 8) gewaschen und dann in eine separate 96er KF Mikrotiterplatte überführt (für Details siehe Anhang: Schutzgruppenabspaltung und Elution)

### 2.2.9 ELUTION VON DEAE-MIKROPARTIKELN

Die Elution der DNA-kodierten Bank von den SiMAG-DEAE Mikropartikeln von Chemicell erfolgte direkt nach der Schutzgruppenabspaltung bzw. nach der Biotinsynthese, um in Anschluss eine neue Syntheserunde bzw. eine Selektion starten zu können. Zu diesem Zweck wurden die Mikropartikel mittels dem KingFisher Flex Roboter in eine neue 96er KF Mikrotiterplatte überführt. Für die Elution wurden die Mikropartikel in 50 µl 20 mM Tris-Puffer (pH 8), 1 M NaCl für 30 min durch Auf- und Abbewegungen inkubiert (siehe Anhang: Elution von DEAE-Mikropartikeln). Nach Beendigung der Inkubation wurden die Mikropartikel an den magnetischen Pins des Roboters gehalten und das so erhaltene Eluat aufgenommen und vereinigt, um es für die neue Syntheserunde/Selektion aufzureinigen. Die besten Elutionsbedingungen wurden mittels verschiedener NaCl-Konzentrationen und anschließender Auswertung mittels eines Agarosegels überprüft.

### 2.2.10 AUFREINIGUNG VON NUKLEINSÄUREN

Um nur die gewünschten Produkte nach einer Reaktion vorliegen zu haben, wurden PCR-Produkte und DNA-Proben nach enzymatischen aber auch chemischen Reaktionsschritten aufgereinigt. Dabei wurden restliche Primer, dNTPs, Enzyme und sonstige Reaktionsbestandteile beseitigt. Für die Aufreinigung der Produkte existieren eine Menge Verfahren, die in Kit-Form angeboten werden. Zumeist handelt es sich dabei um ein säulenbasiertes Verfahren, bei dem die Säulen mit einer Silikatmembran versehen sind. Auch die in dieser Arbeit verwendeten Säulen-basierten Verfahren funktionieren nach diesem Prinzip. Mit dem GeneMATRIX PCR/DNA Clean-Up DNA Purification-Kit von Roboklon wurde vor allem gearbeitet, um die PCR-Produkte nach einer Emulsions-PCR aufzureinigen. Des Weiteren wurde mit dem Wizard SV Gel and PCR Clean-up System von Promega gearbeitet, welches sich für die Aufreinigung von Fragmenten über 100 bp eignet. Beide Kits wurden für diese Zwecke nach den Herstellerangaben verwendet.

Da in dieser Arbeit vor allem mit Fragmenten gearbeitet wurde, die kleiner als 100 bp waren, wurde eine modifizierte Form des Wizard SV Gel and PCR Clean-up System von Promega angewendet. Zu 100 µl DNA-Probe wurden 100 µl Bindungspuffer und 300 µl Isopropanol gegeben. Alles wurde gut gemixt, in ein SV Minicolumn überführt und für 1 min bei RT (22°) inkubiert. Anschließend wurde die Probe bei  $16,000 \times g$  für 1 min zentrifugiert und der Überstand verworfen. Da in dieser Arbeit nach einem Syntheszyklus die Proben vereint wurden, ergab sich eine Probengröße von 800 µl. Der Ansatz wurde achtmal in 100 µl über die gleiche Säule geschickt, bevor nach dem Protokoll weiter gearbeitet wurde. Danach folgte ein Waschschrift mit dem Waschpuffer des Kits. Allerdings wurde die Zusammensetzung des Puffers geändert. Auf 2 ml Waschpuffer wurden 8 ml Ethanol hinzugefügt. Die Probe wurde mit 500 µl von diesem Waschpuffer versetzt und wieder wie oben angegeben zentrifugiert, der Puffer wurde anschließend verworfen. Die Probe wurde dann mit 50 µl Millipore-Wasser für 1 min bei RT inkubiert und anschließend zentrifugiert. In dem Überstand befand sich dann die aufgereinigte Probe.

### 2.2.11 ZWEITE SYNTHESERUNDE

Nach einem Synthesedurchlauf bestehend aus den Schritten 2.2.4 – 2.2.10, konnte die vereinte und aufgereinigte DNA-kodierte Bank für einen neuen Durchlauf und somit der Synthese von d-Biotin verwendet werden. Der Ablauf wurde wie beschrieben beibehalten, lediglich einzelne Schritte wie die Schutzgruppenabspaltung wurden in diesem Fall nicht angewendet.

### 2.2.12 SELEKTION

Nach der Synthese der kombinatorischen Wirkstoffbank ist eine Selektion von neuartigen Wirkstoffen möglich, beziehungsweise kann durch die Selektion eine Syntheseoptimierung ermittelt werden. In dieser Arbeit wurde als Zielmolekül der Selektion Streptavidin gewählt, da zuvor zwei Synthesesyklen mit Glycin und d-Biotin durchgeführt wurden. Da an allen Molekülen der DNA-kodierten Bank d-Biotin unter verschiedenen Bedingungen synthetisiert wurde, können über eine Anreicherung einer oder einiger DNA-Fragmente an dem Zielmolekül die besten Synthesebedingungen ermittelt werden. Jeder Barcode steht dabei für eine Synthesebedingung und kann nach erfolgter Selektion mittels Sequenzierung ermittelt und somit identifiziert werden. Für die Selektion wurden die Dynabeads<sup>®</sup> M-280 Streptavidin (Invitrogen<sup>™</sup>) verwendet. So befand sich das Zielmolekül an einer festen Phase und durch die magnetischen Mikropartikel ist eine einfache Handhabung gewährleistet. Für die Selektion wurden 1,5 mg (150 µl) Streptavidin-Partikel mit 150 µl 2× B&W-Puffer (10 mM Tris-HCl (pH 7,5), 1 mM EDTA, 2 M NaCl) gewaschen. Anschließend wurden die Partikel in 300 µl 2× B&W-Puffer resuspendiert. Zu 40 µl der aufgereinigten DNA-kodierten Substanz-Bank wurden 260 µl bidest H<sub>2</sub>O gegeben, so dass ein Endvolumen von 300 µl entstand. Die Mikropartikel und die DNA-kodierte Bank wurden anschließend vereinigt und bei RT (22°C) für 30 min auf dem Rotator Stuart SB3 von BioCote<sup>®</sup> inkubiert. Danach erfolgte zweimaliges Waschen mit 1 ml 1× B&W-Puffer. Die Separation der Partikel erfolgte jeweils mit dem Magnetpartikelseparator DynaMag von Invitrogen<sup>™</sup>. Die selektierte DNA-kodierte Bank wurde in 150 µl 1× TE-Puffer resuspendiert und im Anschluss auf die Mikropartikel, beladen mit der DNA-kodierten Bank, eine PCR-Reaktion durchgeführt. Ziel war die Ermittlung einer Anreicherung durch Sequenzierung bzw. Gelelektrophorese.



### 2.2.13 SEQUENZIERUNG

Heutzutage sind neue und vor allem mehrere Technologien für die DNA-Sequenzierung verfügbar. Dabei stellt die Sequenzierung eine der wichtigsten Plattformen zur Untersuchung von biologischen Systemen dar (Ronaghi, 2001). In dieser Arbeit wurden verschiedene Sequenzierungsverfahren wie das Roche (454) GS FLX Sequenzierungssystem von Roche Applied Science und die DNA-Kapillar-Sequenzierung nach dem Kettenabbruchmodell von Sanger (Sanger *et al.*, 1977) zu Analyse Zwecken verwendet.

Die Bestimmung der Nukleotidabfolge in einem DNA-Molekül wird als DNA-Sequenzierung bezeichnet. Bis Mitte der siebziger Jahre schien dies ein unlösbares Problem, bis 1977 Fred Sanger und Alan R. Coulson eine neue Sequenziertechnologie veröffentlichten. Dieses Kettenabbruch-Verfahren (Sanger *et al.*, 1977) verbesserte frühere Techniken von Sanger und Coulson und die im selben Jahr entwickelte Sequenziermethode von Maxam und Gilbert (Maxam & Gilbert, 1977). Durch die Reduzierung von toxischen Chemikalien sowie der Verwendung von Radioisotopen bei gleichzeitiger Verbesserung der Leseweite und Qualität der Sequenzen, wurde die Sanger-Sequenzierung die Methode, die sich für die nächsten 30 Jahre durchsetzen konnte (Schuster, 2008). Die Methode basiert auf dem gezielten Einbau von Didesoxynukleotiden (ddNTPs) während der Synthese eines DNA-Stranges, wodurch die Unterbrechung der Reaktion erreicht wird (Bonetta, 2006). Im Laufe der Zeit wurde die Kettenabbruchsequenzierung weiter entwickelt und durch den Einsatz von Fluoreszenzfarbstoffen konnte eine Automatisierung des Systems erfolgen. Die automatisierte Sanger-Methode wird auch als die erste Generation von Sequenzierungssystemen bezeichnet (Metzker, 2010). Um die Proben beim Sequenzierungsteam des MPI für molekulare Genetik abgeben zu können, wurde zunächst eine PCR durchgeführt, so dass die Proben im Anschluss eine definierte Primersequenz an ihrem 3'- und 5'-Ende besaßen. Durch die verwendeten Primer wurden die Sequenzen zum einen verlängert und zum anderen dienen diese als Ansatzstelle für die Sequenzierungsprimer. Zur Sequenzierung abgegeben wurde ligierte DNA-Bank nach der Separation, um die Spezifität dieser kontrollieren zu können. Zu diesem Zweck wurde die DNA-Bank nicht von den Mikropartikeln gelöst, sondern die PCR direkt auf die separierte DNA-Bank, immobilisiert an magnetische Mikropartikel, durchgeführt. So befand sich im Anschluss in der wässrigen Phase das amplifizierte Fragment. Das Pipettierschema und PCR-Programm sind in Tabelle 15 und 16 nachzuvollziehen.

**Tab. 15: Zusammensetzung der PCR-Reaktion für alle 16 Separationen**

	Volumen [ $\mu$ l]	Endkonzentration
10× Puffer B (15 mM MgCl <sub>2</sub> , ROBOKLON)	5	1×
dNTP-Mix (2,5 mM)	4	200 $\mu$ M
Multiplexing Primer 1.0 (100 $\mu$ M)	1	2 $\mu$ M
Primer Sindex 1 (100 $\mu$ M)	1	2 $\mu$ M
Partikel mit Matrizen-DNA	0,25 mg	
Taq-Polymerase (5 U/ $\mu$ l, Roboklon)	2	0,2 U/ $\mu$ l
Millipore-H <sub>2</sub> O	ad 50	

Die PCR wurde in zwei Schritten durchgeführt, da zu Beginn der Reaktion nur ein kleines Stück des Primers komplementär war und sich so zunächst eine geringere Temperatur für die Primeranlagerung ergab. Nach einigen Zyklen konnte diese erhöht werden, da nun genügend amplifiziertes Fragment für die spezifische Primeranlagerung vorliegen sollte.

**Tab. 16: PCR-Programm**

	Temperatur	Zeit	Zyklen
Initiale Denaturierung	94°C	30 s	1×
Denaturierung	94°C	1 min	3×
Primeranlagerung	42°C	1 min	
Elongation	72°C	1 min	
Denaturierung	94°C	1 min	14×
Primeranlagerung	69°C	1 min	
Elongation	72°C	1 min	
Kompletterierung	72°C	5 min	1×

Die 16 verschiedenen Proben inklusive DNA-Bank wurden dann dem Sequenzierungsteam übergeben, wo die eigentlichen Arbeiten für die Sequenzierung durchgeführt wurden. Im Anschluss wurden die Daten übermittelt und mithilfe des Programmes GENTLE veranschaulicht bzw. ausgewertet.

Neben der traditionellen Sequenzierungsmethode nach Sanger sind eine Vielzahl von neuen Technologien entstanden, deren Leistungsfähigkeit um ein deutliches höher ist. Diese werden als Hochdurchsatzsequenzierungen der nächsten Generation bezeichnet (Stangier *et al.*, 2007) und sind seit 2004 verfügbar (Mori *et al.*, 2013).

Darunter fällt auch die in dieser Arbeit verwendete Methode der Pyrosequenzierung (454). Das Roche (454) GS FLX System basiert auf der Technologie der Pyrosequenzierung. Dies ist eine Technik, welche auf der Detektion von freigesetzten Pyrophosphat (PPi) während der DNA-Synthese basiert (Ronaghi, 2001). Dabei verursacht der Einbau eines Nukleotids durch die DNA-Polymerase die Freisetzung von Pyrophosphat, welches eine Serie von enzymatischen Reaktionen verursacht (Mardis, 2008a). Das freigesetzte Pyrophosphat wird anschließend in Adenosintriphosphat umgewandelt. Dies ist ein Cofaktor des Enzyms Luciferase, wodurch diese in der Lage ist, Luciferin in Oxyluciferin zu oxidieren. Als Resultat dieser Enzymkette entsteht ein Lichtsignal, welches detektiert werden kann. Die Stärke des Lichtsignals ist dabei proportional zu der Anzahl der eingebauten Nukleotide. Nachdem die Apyrase alle nicht eingebauten Nukleotide degradiert hat, kann die Reaktion mit den nächsten dNTP durchgeführt werden (Harrington *et al.*, 2013).

Nach zwei vollständigen Synthesedurchgängen und anschließender Selektion auf das Zielmolekül Streptavidin wurden die selektierten Fragmente zur Sequenzierung abgegeben, um eine Anreicherung und somit bevorzugte Synthesemethode zu ermitteln. Da die DNA-Bank so konstruiert wurde, dass die konstanten Regionen komplementär zu den Sequenzierungsprimern (siehe Tab. 9) sind, konnte durch eine PCR-Reaktion (siehe Abschnitt 2.2.14) bzw. Emulsions-PCR (siehe 2.2.15) der sogenannte Barcode und benötigte Primer direkt an die selektierte Bank angefügt werden. So wurde der DNA-kodierten Bank ein einmaliger Barcode zugeordnet. Anschließend wurden die Proben wieder beim Sequenzierungsteam des MPIMG abgegeben und per Pyrosequenzierung (454) sequenziert. Die erhaltenen Daten wurden anschließend mithilfe des Programmes X-WIN 32 ausgewertet.

### **2.2.14 POLYMERASEKETTENREAKTION (PCR)**

Die Mitte der achtziger Jahre durch Mullis und seinen Kollegen entwickelte Polymerasekettenreaktion (*polymerase chain reaction*, PCR) ist eine schnelle und sehr sensitive Methode in der DNA-Analytik (Mullis *et al.*, 1986). Es handelt sich dabei um eine durch spezifische Primer-definierte enzymatische *in vitro* Replikation, bei der durch sich wiederholende Zyklen eine annähernd exponentielle Amplifikation der Zielsequenz erreicht wird. Die Technik beruht auf zwei Oligonukleotide (Primer), die die zu amplifizierende DNA

flankieren und einer thermostabilen Taq-Polymerase, welche in der Lage ist, die DNA-Matrize durch wiederholende Prozesse zu vervielfältigen. Der zyklische Ablauf besteht aus Denaturierung der DNA, Anlagerung der Primer an die komplementäre Sequenzen und der Elongation der angelagerten Primer mittels der Polymerase (Saiki *et al.*, 1988). Diese Schritte werden zyklisch wiederholt, wobei die Produkte vorheriger Zyklen als Ausgangsmaterial für den nächsten Zyklus dienen. So kann eine annähernd exponentielle Vervielfältigung erreicht werden. Die Reaktion erfolgte zumeist in einem Endvolumen von 50 µl (s. Tab. 17) und die Zyklenanzahl variierte zw. 10 und 35. Das entsprechende PCR-Programm ist in Tabelle 18 nachzulesen.

**Tab. 17: Zusammensetzung der PCR-Reaktion**

	Volumen [µl]	Endkonzentration
10× Puffer B (15 mM MgCl <sub>2</sub> , Roboklon)	5	1×
dNTP-Mix (2,5 mM)	8	400 µM
3'-Primer (100 µM)	2	4 µM
5'-Primer (100 µM)	2	4 µM
Matrizen-DNA	0,5-2	
Taq-Polymerase (5 U/µl, Roboklon)	1	0,1 U/µl
Millipore-H <sub>2</sub> O	ad 50	

Die PCR-Ansätze wurden auf Eis pipettiert und anschließend ein entsprechendes Programm gewählt und gestartet.

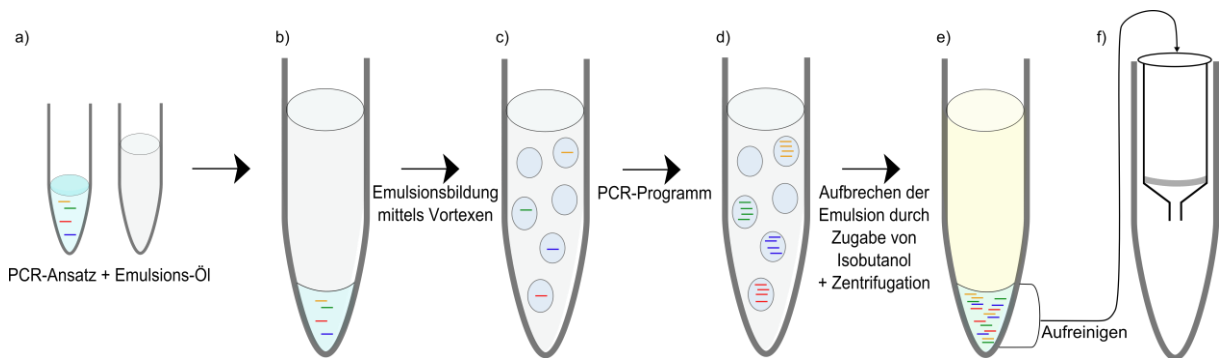
**Tab. 18: Programm zur Durchführung der PCR**

	Temperatur	Zeit	Zyklen
Initiale Denaturierung	95°C	30 s	1×
Denaturierung	95°C	1-2 min	10-35×
Primeranlagerung	58-69°C	1 min	
Elongation	72°C	1 min	
Kompletterung	72°C	5 min	1×

Die Analyse des PCR-Produktes erfolgte anschließend über die Gelelektrophorese und die Konzentrationsbestimmung des erhaltenen Produktes wurde mithilfe des NanoDrop ND-1000 ermittelt.

### 2.2.15 EMULSION-POLYMERASEKETTENREAKTION (EPCR)

Die Kompartimentierung der Polymerasekettenreaktion reduziert die Bildung von Artefakten, vor allem wenn komplexe DNA-Banken vervielfältigt werden sollen (Schütze *et al.*, 2011). Bei dieser Technik werden die zu amplifizierenden Fragmente mit den zugehörigen Reagenzien kompartimentiert und befinden sich in winzigen Reaktoren (Mizellen), wobei im Idealfall in jeder Mizelle nur ein Fragment vorliegt. Somit steht jedem Fragment die gleiche Menge an Ressourcen zur Verfügung.



**Abb. 14: Arbeitsablauf der Emulsions-PCR.** Herstellung der PCR-Reaktion und des Emulsions-Öl (a) auf Eis. Vereinigung der wässrigen Phase und des Ölgemisches (b) und Emulsionsbildung mittels Vortexen (c). Amplifikation (d) und anschließendes Aufbrechen der Mizellen mittels Isobutanol, sowie Aufreinigung der PCR-Produkte (e-f) (nach Schütze *et al.*, 2011).

Die ePCR besteht aus einer wässrigen Phase, die sich aus den Komponenten lt. Tabelle 17 zusammensetzt und aus einem Wasser-in-Öl-Gemisch, welches aus 7% ABIL WE 09, 20% Mineral-Öl und 73% Tegosoft DEC besteht. Zu einem 50 µl PCR-Ansatz wurden zusätzlich 2,5 µl bovines Serumalbumin (10 mg/µl) zur Stabilisierung der Taq-Polymerase und 300 µl Emulsions-Öl gegeben. Das Emulgieren erfolgte mithilfe eines Vortexers bei 4°C für 5 min bei maximaler Geschwindigkeit. Anschließend wurden die Reaktionsansätze auf PCR-Stripes aufgeteilt und das PCR-Programm (siehe Tab. 18) gestartet. Im Anschluss wurden die Proben wieder vereinigt und die Mizellen durch Zugabe von 1 ml Isobutanol aufgebrochen. Die Aufreinigung der PCR-Produkte erfolgte im nächsten Schritt mit dem GeneMATRIX PCR/DNA clean-up Purification Kit von Roboklon.

## 2.2.16 GELELEKTROPHORESE

Die Gelelektrophorese ist ein gängiges Trennverfahren, bei dem Gemische von Proteinen, DNA und RNA aufgetrennt werden können. Das Prinzip beruht auf der Wanderung der geladenen Moleküle im elektrischen Feld. Dabei wandern die positiv geladenen Moleküle zur Kathode und negativ geladene Moleküle zur Anode. Das Trennverfahren wird dabei in einer so genannten Trägermatix durchgeführt (Brody & Kern, 2004). Dabei macht man sich die Mobilität der Moleküle, die durch ihre Form und Größe bestimmt wird, zunutze. Die elektrophoretische Trennung wird durch die im Gel vorhandenen Poren erreicht, die wie ein Molekularsieb wirken. Dabei bewegen sich die Moleküle aufgrund ihrer unterschiedlichen Größe bei gleicher Ladung unterschiedlich schnell im elektrischen Feld und werden so elektrophoretisch aufgetrennt (Seyffert, 2003). Als Trägermedien werden bei der Gelelektrophorese hauptsächlich Agarose oder Polyacrylamid eingesetzt. Diese Stoffe bilden ein engmaschiges Netz, das die zu trennenden Moleküle bei ihrer Wanderung im elektrischen Feld hindert.

### 2.2.16.1 AGAROSE – GELELEKTROPHORESE

Agarose ist ein lineares Polysaccharid, mit dem relativ großporige Gele hergestellt werden. Dabei lässt sich die Größe der Poren durch die Agarosekonzentration steuern (Müllhardt, 2013). Die elektrophoretische Auftrennung von DNA-Fragmenten erfolgte in dieser Arbeit hauptsächlich über 3%-ige Agarosegele in 1× TBE-Puffer, denen 2 µl (10 mg/ml) Ethidiumbromid zur Detektion der Fragmente zugesetzt wurde. Diese Konzentration der Agarosegele erlaubt einen Auftrennungsbereich, der kleiner als 200 bp ist. Die Proben wurden vor dem Gelauftrag mit 10× Ladepuffer versetzt und als Größenstandard wurde der O'RangeRuler™ 20 bp (Fermentas) verwendet. Als Laufpuffer diente 1× TBE-Puffer und die anschließende elektrophoretische Auftrennung erfolgte bei 80 V für 30 min.

### 2.2.16.2 POLYACRYLAMID – GELELEKTROPHORESE

Als gängiges Gelmaterial wird neben Agarose vor allem Polyacrylamid verwendet, wodurch vor allem kleinere Nukleinsäuren (Müllhardt, 2013) aufgetrennt werden können. Polyacrylamidgele bestehen aus dem Monomer Acrylamid und dem Quervernetzer N,N'-Methylenbisacrylamid. In einer radikalischen Reaktion wird die Gelmatrix durch die Ausbildung von Polymeren gebildet. Dabei dient Ammoniumperoxidsulfat (APS) als Polymerisationsstarter und N,N,N',N'-Tetramethylethyldiamin (TEMED) als Radikalstabilisator. Der Vernetzungsgrad und somit der Trennbereich des Gels kann durch die Acrylamid/Bisacrylamid-Konzentration (Chrambach & Rodbard, 1971) variiert werden (siehe Tab. 19). In dieser Arbeit wurden für analytische Zwecke vor allem 12%-ige denaturierende Polyacrylamidgele verwendet. Denaturierende Polyacrylamidgele werden vor allem genutzt, um die Ausbildung von Sekundär- und Tertiärstrukturen zu unterbinden, die das Laufverhalten von Nukleinsäuren beeinflussen können. Zu diesem Zweck wurde der Polymerisationsansatz mit 7 M Harnstoff versetzt. Zur Herstellung der Gele wurden Acrylamid/Bisacrylamid (19:1) mit Harnstoff und 1× TBE-Puffer versetzt. Die Polymerisation wurde durch die Zugabe von APS und TEMED gestartet. Die Proben wurden mit 6× Ladepuffer für Polyacrylamidgele versetzt und als Laufpuffer diente auch hier 1× TBE-Puffer, dazu wurde der 20 bp Größenstandard von Fermentas verwendet. Die Gele liefen in vertikalen Kammern bei 80 V für 1 h 40 min. Zur Detektion der Fragmente wurden die Gele im Anschluss in einem Ethidiumbromidbad (3 µl/ 30 µg in 100 ml 1× TBE-Puffer) für 15 min inkubiert und anschließend unter UV-Licht analysiert.

**Tab. 19: Zusammensetzung und Trennbereich von denaturierenden Polyacrylamidgelen.** Dargestellt wurde die Zusammensetzung für 50 ml Gellösung.

Acrylamidkonzentration	6%	10%	12%
Trennbereich	50-200 bp	20-80 bp	12-60 bp
40% Acrylamid/Bisacrylamid (19:1)	7,5 ml	12,5 ml	15 ml
Harnstoff	21 g	21 g	21 g
1× TBE-Puffer	ad 50 ml	ad 50 ml	ad 50 ml
APS (10%)	500 µl	500µl	500µl
TEMED	40 µl	40 µl	40 µl

### 2.2.17 SPEKTROPHOTOMETRISCHE QUANTIFIZIERUNG VON DNA

Die spektrophotometrische Analyse von Nukleinsäure ist ein wesentlicher und wichtiger Teil der Laborroutine. Diese Methode gibt Hinweise auf die Qualität und Menge gelöster DNA bzw. RNA. Für die Konzentrations- und Reinheitsbestimmung der vorliegenden DNA wurde das NanoDrop ND-1000 von Thermo Scientific® verwendet. Mithilfe dieses Instruments können besonders kleine Volumina für die Analyse der DNA genutzt werden. Zudem ist es möglich Messungen in einem Spektrum von 220-750 nm vorzunehmen. Das Prinzip dieser Methode basiert darauf, dass die Eigenschaft der Basen, ultraviolettes Licht zu absorbieren, ausgenutzt wird. Für die Messung wurde 1 µl der Probe auf das Analysefenster gegeben und das Absorptionsspektrum in Relation zu einer Referenzlösung gemessen. DNA hat sein Absorptionsmaximum bei einer Wellenlänge von 260 nm, Proteine hingegen bei 280 nm. Mit der Berechnung des Quotienten der Messung bei 260 nm und 280 nm ( $OD_{260/280}$ , optische Dichte) lässt sich neben der Konzentrationsbestimmung auch die Reinheit der Nukleinsäure bestimmen (Hauk, 2013). Liegt der Quotient bei etwa 1,8 handelt es sich um eine reine DNA-Lösung. Ein Quotient größer als 2,0 deutet auf RNA-Rückstände hin. Verunreinigungen durch Proteine oder Phenol, die ein Absorptionsmaximum bei 280 nm haben, machen sich durch einen Quotienten, der bei etwa 1,5 liegt, bemerkbar. Für die Berechnung der Nukleinsäuremengen gelten folgende Gleichungen (Wollenberger *et al.*, 2003):

dsDNA:        1  $OD_{260} = 50 \mu\text{g/ml}$

ssDNA:        1  $OD_{260} = 33 \mu\text{g/ml}$

RNA:            1  $OD_{260} = 40 \mu\text{g/ml}$

### 2.2.18 NACHWEISMETHODEN FÜR NUKLEINSÄUREN

Die schnellste und gängigste Methode um Nukleinsäure in Gelen nachzuweisen ist die Ethidiumbromidfärbung. Ethidiumbromid ist ein solvatochromer Farbstoff, der seine Farbe ändert, wenn es an DNA gebunden wird (Zimmermann, 1980). Dabei handelt es sich um einen allgemein gebräuchlichen, unter UV-Licht fluoreszierenden Farbstoff, der in der Lage ist, sich reversibel in die Basenabfolge der DNA einzulagern. Bei Bestrahlung mit UV-Licht



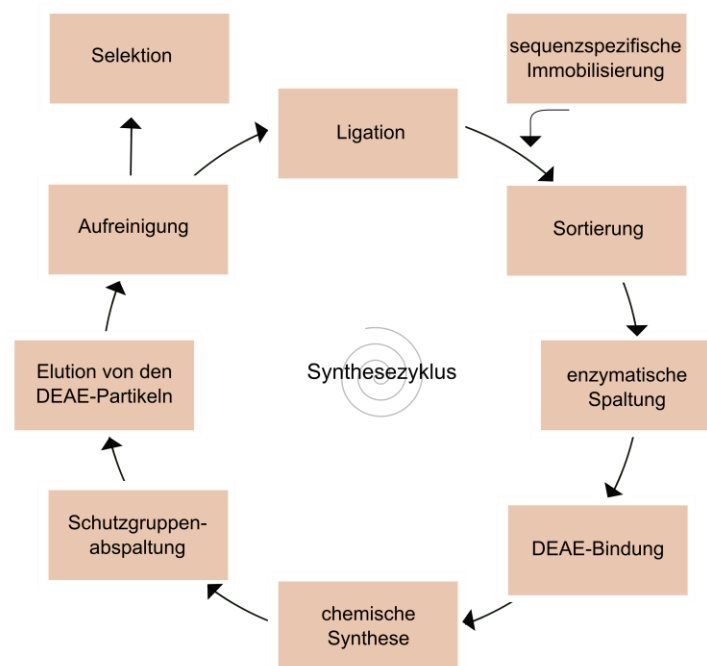
der Wellenlänge zwischen 260-360 nm kommt es zu einer Fluoreszenz, wodurch die DNA unter dem Transilluminator sichtbar und somit detektierbar wird.

In dieser Arbeit wurde den Agarosegelen der Farbstoff bereits bei der Herstellung hinzugefügt. Dabei wurden auf 50 ml Gellösung 2  $\mu$ l (10 mg/ml) Ethidiumbromid verwendet. Polyacrylamidgele wurden nach dem Lauf für 15 min in einem Ethidiumbromidbad (3  $\mu$ l/ 30  $\mu$ g in 100 ml 1 $\times$  TBE-Puffer) unter leichtem Schütteln inkubiert.

In dieser Arbeit wurde ein Fluoreszenznachweis mittels Cy5 nach einer Hybridisierungsreaktion genutzt, um die Immobilisierung eines Oligonukleotids an magnetische Partikel bzw. die Effizienz der Separation nachzuweisen. Zu diesem Zweck besaß das Hybridisierungsoligo Cy5aGCG an seinem 5'-Ende den Fluoreszenzfarbstoff Cy5. Nach der Immobilisierung des Oligos hybGCG an die magnetischen Partikel wurde der Ansatz mit einer äquimolaren Menge des Hybridisierungsoligos bzw. einer Titrationsreihe zu dem immobilisierten Oligo in Hybridisierungspuffer inkubiert. Anschließend wurde die Probe zweimal mit dem jeweiligen Hybridisierungspuffer gewaschen und in 1 $\times$  TE-Puffer resuspendiert (siehe Abschnitt 2.2.3.1). Die Visualisierung der Hybridisierung und somit der Detektion der Immobilisierung an magnetische Partikel erfolgte über das Fluoreszenzsignal von Cy5aGCG mithilfe des Laser Scanner Affimetrix 428 von Sartorius. Für das Scannen der Proben wurden diese auf einen Glasobjektträger aufgetragen, kurz eingetrocknet und anschließend gescannt. Dabei wurden die Cy5-Moleküle mithilfe des Lasers bei einer Wellenlänge von 633 nm angeregt und die Emissionssignale bei 665 nm (Waggoner *et al.*, 1993) detektiert und digitalisiert.

### 3 ERGEBNISSE

In den letzten Jahren hat sich neben dem Hochdurchsatz-Screening und den gängigen Verfahren der biochemischen Selektionsmethoden eine Vielzahl von neuen Techniken für das Auffinden von funktionellen, molekularen Liganden von biologischen Zielen entwickelt. Dieses neue Methodenfeld wird als DNA-kodierte Banken (DELs) bezeichnet. In diesem Teil der Arbeit sollen die einzelnen Schritte einer neuartigen Synthesestrategie für DNA-kodierte Substanz-Banken beschrieben werden. Nach erfolgten Selektionsexperimenten soll zudem die Anwendung der Hochdurchsatzsequenzierung der nächsten Generation für die Analyse von DNA-kodierten Substanz-Banken herangezogen werden. In dem folgenden Diagramm ist eine Übersicht der einzelnen iterativen Schritte zur Herstellung DNA-kodierter Substanz-Banken dargestellt.



**Abb. 15: Schematische Darstellung des Arbeitsablaufes der Neclys-Technologie.** Die DNA-kodierte Substanz-Bank wird durch iterative Reaktionsschritte aufgebaut. Im Anschluss erfolgt die Selektion gegen ein Zielmolekül, das von Interesse ist.

Mithilfe dieses Diagramms kann genau nachvollzogen werden, aus welchen Bausteinen sich die Neclys-Technologie zusammenbaut. Hauptbestandteile der Technologie sind unter anderem die Ligation, Separation, Spaltung, Synthese und Elution, die sich iterativ wiederholen bis die gewünschte Größe der DNA-kodierten Bank erreicht wurde. Nach den Syntheserunden folgt schlussendlich die Selektion gegen ein Zielmolekül.

### **3.1 STRATEGIEN ZUR KOPPLUNG VON NUKLEINSÄURE AN MIKROPARTIKELN**

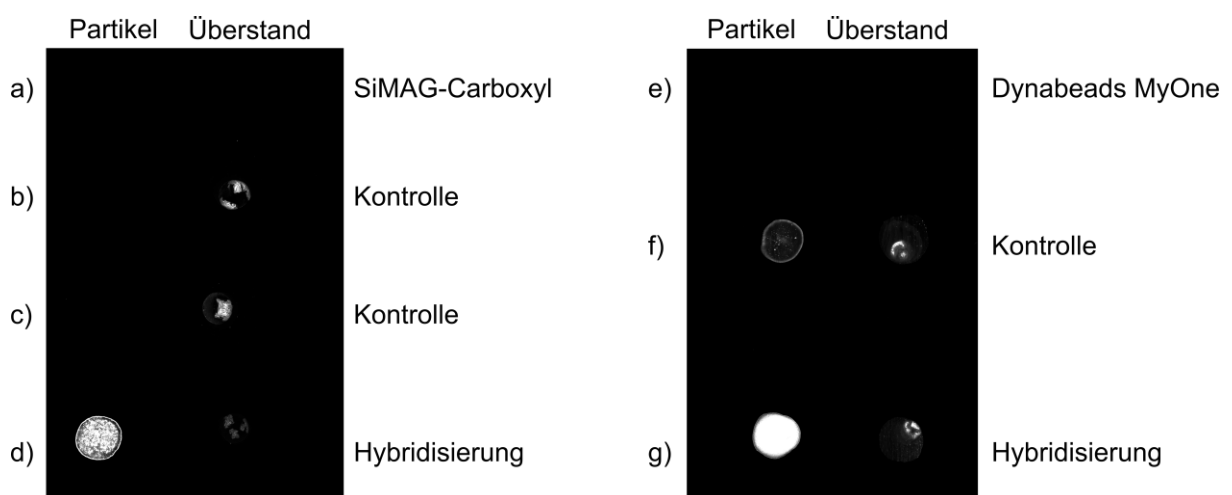
In dieser Arbeit wurden magnetische Mikropartikel für verschiedene Verfahren verwendet. Sowohl für die Sortierung von Nukleinsäure mittels immobilisierter Oligonukleotide, für die Synthese- als auch die Selektionsreaktion. Bei magnetischen Partikeln handelt es sich um kleine paramagnetische Kügelchen, die mit den verschiedensten Materialien beschichtet werden können und so ihre gewünschten Eigenschaften erhalten. Diese eignen sich für viele Techniken so gut, da sie in der Reaktion frei beweglich sind und sich durch ihre magnetischen Eigenschaften aber auch schnell von der Flüssigphase separieren lassen (Müllhardt, 2013). In dieser Arbeit wurden verschiedene Varianten von Mikropartikeln und somit auch Kopplungsmethoden getestet, um so die beste Effizienz, Handhabung und Technik für den Sortierungsschritt zu ermitteln.

In der Separation wird die ligierte DNA-Bank mittels magnetischen Mikropartikeln getrennt. Dafür befinden sich an den Partikeln gekoppelte Hybridisierungsoligonukleotide. Diese Oligos hybridisieren mit der Adaptersequenz der ligierten DNA-Bank und so kann diese hinsichtlich ihrer Nukleotidabfolge sortiert werden.

#### **3.1.1 KOVALENTES KOPPLUNGSVERFAHREN MITTELS CARBODIIMID-METHODE**

Dieses Verfahren beschreibt die kovalente Kopplung von aminomodifizierter Nukleinsäure an magnetische carboxylierte Mikropartikel mittels der Carbodiimid-Methode. Bei der kovalenten Bindung oder auch Atombindung handelt es sich um eine chemische Bindung, die für den Zusammenhalt von chemischen Verbindungen verantwortlich ist. Dabei wird angenommen, dass jeweils zwei Atome ein oder mehrere Elektronenpaare gemeinsam nutzen (Jakubke & Karcher, 1998). Das einzelne Elektron verlässt dabei nicht seinen Atomverband (Beyer & Walter, 1988). In dieser Arbeit wurden Mikropartikel, deren Oberfläche aktivierbare Carboxylgruppen tragen, von zwei Herstellern getestet. Zum einen wurden die Dynabeads MyONE Carboxylic Acid (Ø 1 µm, 10 mg/ml) von Invitrogen und zum anderen die SiMAG-Carboxyl (Ø 0,5 µm, 50 mg/ml) von Chemicell verwendet. Mittels der Carbodiimid-Methode wurden an diesen Partikeln Hybridisierungsoligonukleotide, welche eine Aminogruppe an ihrem 5'-Ende tragen, gekoppelt (hybGC / hybCG). Carbodiimide haben vor allem eine große

Bedeutung und Aufmerksamkeit durch ihre vielseitige Verwendung bei der organischen Synthese (Kurzer & Douraghi-Zadeh, 1967). Carbodiimide sind Vernetzungsreagenzien, die verwendet werden um die Bildung einer Amidbindung zwischen einer Carboxyl- und Aminogruppe zu vermitteln. Oft wird ein zweites Reagenz wie N-Hydroxysuccinimid (NHS) zu der Reaktion hinzugegeben, um diese zu modifizieren (Sinz, 2003). Das zweite Reagenz NHS fängt dabei den intermediär gebildeten und zur Racemisierung neigenden O-Acylisoharnstoff ab. Die Carbodiimid-Methode beruht auf der Aktivierung der Carboxykomponente. Durch den elektronenanziehenden Charakter des Carbodiimid wird die positive Partialladung auf dem Carbonylkohlenstoffatom erhöht, wodurch dessen Reaktivität gegenüber einer Aminogruppe stark ansteigt (Altmann & Mutter, 1993). Die Kopplung der 5'-aminomodifizierten Oligonukleotide hybGC und hybCG an die carboxylierten magnetischen Mikropartikel erfolgte nach Herstellerangaben und ist unter 2.2.3.1 nachzulesen. Nach erfolgter Kopplung sollte die Effizienz mittels Fluoreszenzdetektion ermittelt werden. Zu diesem Zweck wurden die immobilisierten Oligonukleotide mit einem komplementären Oligo, welches die Fluoreszenzgruppe Cy5 an seinem 5'-Ende trägt, hybridisiert und anschließend mit einem Laser Scanner das Signal ausgewertet. Um die Effizienz der Kopplung ermitteln zu können, wurde der Überstand der Hybridisierung mit auf den Glasobjektträger aufgetragen. So konnte über die Stärke des Signals ermittelt werden, wie viel Oligonukleotid an den Partikeln gebunden hatte.



**Abb. 16: Kovalente Kopplung mittels Carbodiimid-Methode.** Die Kopplung der Oligonukleotide hybGC und hybCG erfolgte an SiMAG-Carboxyl Partikel (a-d) und Dynabeads MyOne Mikropartikel (e-g). Probe der Mikropartikel (a, e). Bei den Reaktionen (b, c und f) handelt es sich um Kontrollen, um die Hintergrundaktivität der Immobilisierung bzw. Hintergrundfluoreszenz zu ermitteln. Ansatz (d, g) zeigt die Hybridisierung des immobilisierten hybGC mit dem Fluoreszenzoligo Cy5aGCG an unterschiedlichen Mikropartikeln.

In der Abbildung 16 ist die Detektion der Immobilisierung eines aminomodifizierten Oligonukleotids an carboxylierte magnetische Mikropartikel zu erkennen. Die Visualisierung erfolgte mit dem Laser Scanner Affimetrix 428 von Sartorius mithilfe des Fluoreszenzsignals des hybridisierten Oligonukleotids.

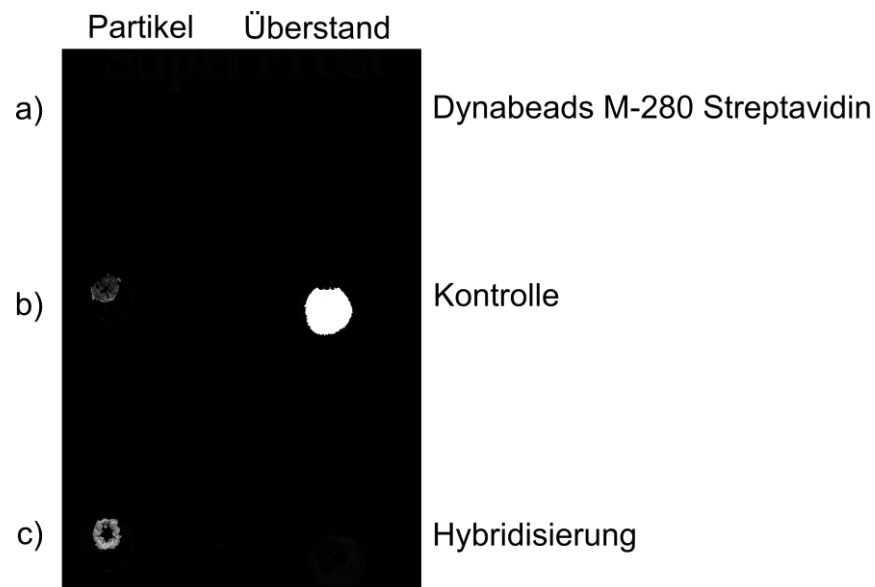
Um Hintergrundaktivitäten der Mikropartikel ausschließen zu können, wurden Proben der Partikel mit aufgeführt (Abb. 16, a & e). Die Hybridisierung (d & g) zeigt bei der Verwendung beider Partikelvarianten ein starkes Fluoreszenzsignal, im Überstand hingegen ist ein sehr schwaches Signal wahrzunehmen. So kann darauf geschlossen werden, dass die Immobilisierung effizient funktioniert hat. Um auszuschließen, dass das Fluoreszenzoligo auch ohne immobilisiertes Hybridisierungsoligo an den Partikeln bindet, wurden Kontrollen mitgeführt. Bei der Kontrolle Abb. 16, b ist kein Signal an den Partikeln wahrzunehmen, lediglich im Überstand. So scheint das Oligo nicht an den Partikeln ohne gekoppelte, komplementäre Oligonukleotide binden zu können. In den Kontrollen Abb. 16, c & f sind die verwendeten Oligonukleotide nicht komplementär und auch in diesen Ansätzen ist nur ein Signal in dem Überstand zu erkennen. So ist auch die Hybridisierung spezifisch. Im Vergleich der beiden verwendeten magnetischen Mikropartikel ist zu verzeichnen, dass die Dynabeads MyONE Carboxylic Acid von Invitrogen im Vergleich zu den SiMAG-Carboxyl von Chemicell ein stärkeres Signal aufzeigen und somit deren Effizienz höher ist.

### **3.1.2 KOPPLUNG VON BIOTINYLIERTER NUKLEINSÄURE AN DYNABEADS**

#### **M-280 STREPTAVIDIN**

Eine weitere Kopplungsmethode von Nukleinsäuren an magnetische Partikel sollte über das Schlüssel-Schloss-Prinzip zwischen dem Vitamin d-Biotin und dem Protein Streptavidin erfolgen. Zu diesem Zweck ist die Oberfläche der verwendeten magnetischen Partikel mit Streptavidin beschichtet und die Hybridisierungsoligonukleotide sind an ihrem 5'-Ende biotinyliert. Streptavidin ist ein Protein, welches von der Bakterienspezies *Streptomyces* hergestellt wird (Chaiet & Wolf, 1964). Dabei handelt es sich um ein tetrameres Protein, das aus vier identischen Untereinheiten aufgebaut ist. Jedes der vier Untereinheiten ist in der Lage, ein Molekül des Vitamins d-Biotin mit sehr hoher Affinität zu binden (Weber *et al.*, 1989). Diese Bindung ist eine der stärksten nichtkovalenten Bindung in der Biologie.

Die Kopplung der 5'-biotinylierten Oligonukleotide an die Dynabeads M-280 Streptavidin erfolgte nach den Herstellerangaben von Invitrogen. Auch in diesem Fall erfolgte die Detektion der Hybridisierung mithilfe eines Laser Scanners und einem zu detektierenden Fluoreszenzsignal.



**Abb. 17: Kopplung biotinylierter Nucleinsäure an Dynabeads M-280 Streptavidin.** Probe der Partikel (a), um Hintergrundaktivitäten auszuschließen. Die Reaktion (b) diente als Kontrolle, um die Spezifität der Hybridisierung zu ermitteln. Ansatz (c) zeigt die Hybridisierung des immobilisierten Oligos mit dem komplementären Fluoreszenzoligo Cy5aGCG.

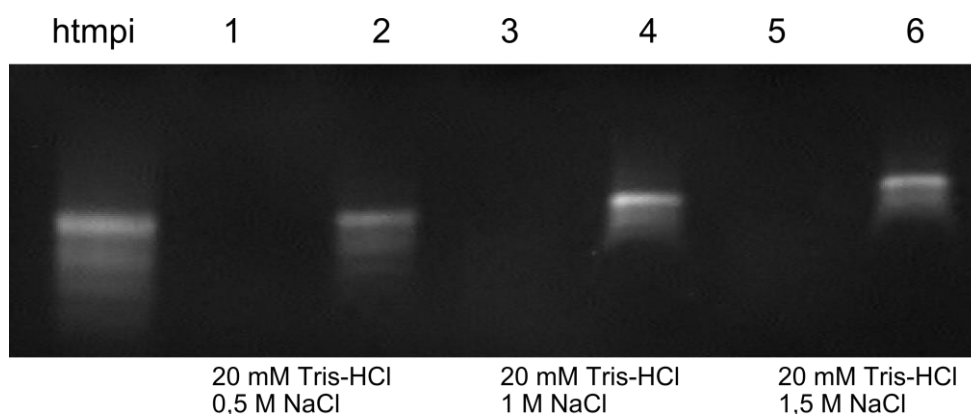
In der Abbildung 17 sind die Ergebnisse der Immobilisierung einer biotinylierten Nucleinsäure an Streptavidin-Partikel dargestellt. Auch hier wurde die Hintergrundaktivität der Partikel durch eine Probe (Abb. 17, a) ausgeschlossen. Die Hybridisierung der immobilisierten, biotinylierten Oligonukleotide mit dem komplementären Fluoreszenzoligo (c) zeigte ein starkes Signal bei den Partikeln, keines aber im Überstand. Somit kann darauf geschlossen werden, dass sowohl die Immobilisierung als auch die Hybridisierung funktioniert haben. Auch die Spezifität wurde durch die Kontrolle (b) getestet. In dieser Probe sind die Oligonukleotide nicht komplementär zueinander und es sollte so zu keiner Hybridisierung kommen. Da in dem Überstand ein starkes Signal zu verzeichnen war und keines an den Partikeln, ist auch dies nachgewiesen worden. Des Weiteren wurde die Effizienz der Immobilisierung getestet. Nach Literaturangabe von Invitrogen können an 1 mg Magnetpartikel 200 pmol Nucleinsäure gekoppelt werden. Da in dem Ansatz Abb. 17, c kein Fluoreszenzsignal im Überstand zu verzeichnen war, muss davon ausgegangen werden, dass eine Hybridisierung zwischen den immobilisierten Oligos und dem Fluoreszenzoligo stattgefunden hat. Da beide Oligos in äquimolaren Mengen eingesetzt wurden, kann

festgestellt werden, dass die Literaturangaben reproduziert werden konnten. Somit können an 1 mg Magnetpartikeln 200 pmol Nukleinsäure gekoppelt und die gleiche Menge an komplementären Oligonukleotiden mittels Hybridisierung gebunden werden.

Da die Streptavidin-Partikel eine so hohe Effizienz und Spezifität aufgezeigt haben, wurden diese für die Versuchsreihen der Synthese und Selektion von Nukleinsäure-kodierten chemischen Wirkstoffbanken verwendet.

### 3.1.3 NUKLEINSÄUREKOPPLUNG MITTELS IONISCHER BINDUNG

Für die ionische Bindung von Nukleinsäure an magnetische Partikel wurden die SiMAG-DEAE Mikropartikel ( $\varnothing$  0,5  $\mu$ m, 50 mg/ml) von Chemicell verwendet. Die Ionenbindung beruht auf dem Prinzip der elektrostatischen Anziehung entgegengesetzt geladener Ionen. Ionen werden dabei aus Atomen durch Abgabe bzw. Aufnahme von Elektronen gebildet (Jakubke & Karcher, 1998). Diese Methode der Kopplung von Nukleinsäure an magnetische Partikel wurde für die Festphasensynthese verwendet. Die sortierte DNA-Bank konnte für die Synthese an eine feste Phase gebunden werden und erleichtert deren Handhabung. Zunächst wurden einige Konditionen für die Kopplung und Elution überprüft, um die optimalen Bedingungen zu bestimmen. Die Kopplung erfolgte wie unter 2.2.3.2 beschrieben. Die Effizienz der Kopplung konnte über die Ausbeute der Rückgewinnung bzw. Elution des Fragments von den Partikeln mit bestimmt werden.



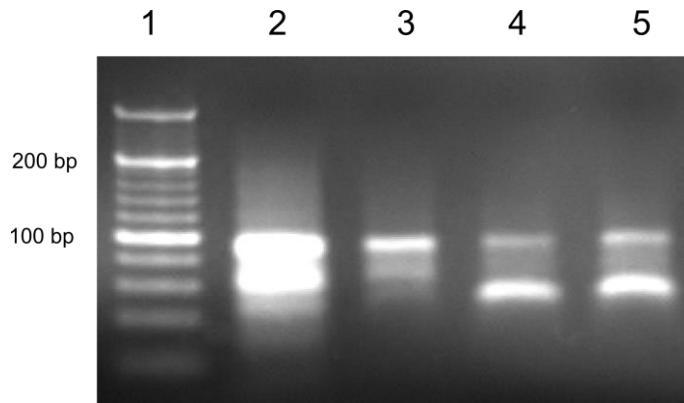
**Abb. 18: Detektion des Elutionsertrag des gebundenen Oligonukleotids htmpi unter verschiedenen NaCl-Konzentrationen auf einem Ethidium-gefärbten Agarosegel (3%).** Vergleichskonzentration des eingesetzten Oligos (htmpi). Probe der Kopplungslösung nach erfolgter Bindung des Fragments an die Partikel (Spur 1, 3, 5). Eluate der DEAE-Mikropartikel unter verschiedenen Elutionsbedingungen (Spur 2, 4, 6).

Das Oligonukleotid htmpi wurde zur Bestimmung der besten Kopplungs- und Elutionsbedingungen verwendet. Zunächst erfolgte die Kopplung wie beschrieben, anschließend wurde die Elution in 20 mM Tris-HCl und verschiedenen NaCl-Konzentrationen getestet. Über ein 3%-iges Agarosegel wurde anschließend die Ausbeute der Elution im Vergleich zu der eingesetzten Ausgangskonzentration des Oligonukleotids qualitativ bestimmt.

Nachdem das Oligo htmpi an die DEAE-Mikropartikel gebunden hatte, wurden die Partikel in ein neues Reaktionsgefäß überführt und die Elution vorgenommen. Von der Kopplungslösung wurde jeweils eine Probe auf das Agarosegel aufgetragen, um so die Bindungseffizienz bestimmen zu können. In dem Gelbild (Abb.18) ist in allen Ansätzen in der Kopplungslösung kein Fragment zu verzeichnen und es kann festgestellt werden, dass die Bindung unter diesen Bedingungen (siehe 2.2.3.2) sehr effizient ist. Anschließend erfolgte die Elution mit verschiedenen Salzkonzentrationen. Im Vergleich zwischen der Elution mit 0,5 M NaCl und 1 M NaCl kann man sehen, dass die Bande zwar aufgrund des Salzgehaltes höher läuft, aber auch die Elutionsausbeute größer ist. Zwischen 1 M und 1,5 M NaCl ist allerdings kein Unterschied zu erkennen. Somit erfolgt die Elution im Folgenden mit 20 mM Tris (pH 8), 1 M NaCl. Die einzelnen Arbeitsschritte wurden im Laufe der Arbeit auf den Magnetpartikelprozessor KingFisher Flex von Thermo Scientific übertragen, um so ein halbautomatisiertes Verfahren herzustellen (Anhang: Elution von DEAE-Mikropartikeln).

Da die Bindung und Elution von den DEAE-Mikropartikeln nicht direkt hintereinander erfolgen, sondern ein Schritt in Lösungsmittel für die kombinatorische chemische Synthese dazwischen liegt, wurde ein Stabilitätstest mit der DNA-Bank durchgeführt. Dabei sollte überprüft werden, ob die DNA-Bank im Lösungsmittel stabil bzw. an den Partikeln gebunden bleibt und erst während der Elution gelöst werden kann. Zu diesem Zweck wurde die DNA-Bank wieder an die DEAE-Partikel gebunden und vor dem Elutionsschritt in 1 Methyl-2-Pyrrolidon (99%) und Dimethylformamid (DMF) inkubiert. Anschließend wurde die Elution mit 20 mM Tris-HCl (pH 8), 1 M NaCl durchgeführt.





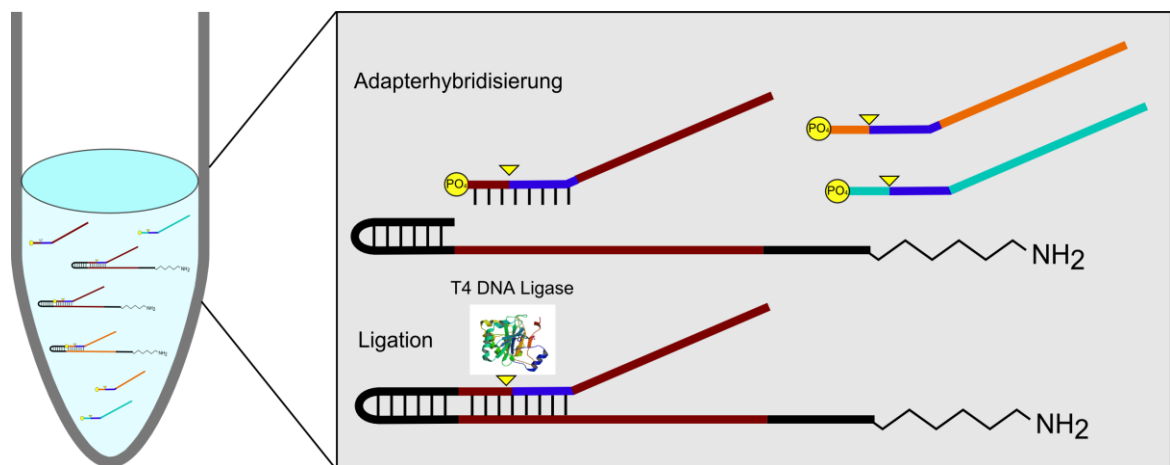
**Abb. 19: Elutionsertrag der immobilisierten DNA-Bank auf einem Ethidium-gefärbten Agarosegel (3%).** O'RangeRuler 20 bp DNA Ladder (Spur 1). Konzentration des eingesetzten Ausgangsoligos, welches an den SiMAG-DEAE Partikeln gebunden werden sollte (Spur 2). Elution erfolgte bei allen mittels 20 mM Tris-HCl, 1 M NaCl. Ohne Inkubationsschritt (Spur 3). Zwischen der Bindung und der Elution erfolgte eine Inkubation in 1 Methyl-2-Pyrrolidon (Spur 4) und DMF (Spur 5).

Die immobilisierte DNA-Bank bleibt an den DEAE-Partikeln sowohl bei einer Inkubation mit 1 Methyl-2-Pyrrolidon (Abb. 19, Spur 4) als auch mit DMF (Spur 5) an den Partikeln stabil gebunden und löst sich nicht ab. Ansonsten wäre in dem Gelbild kein Eluat zu verzeichnen. Im Vergleich zum Ansatz ohne Inkubation (Spur 3) ist kein Unterschied zu erkennen. Die Elution von den magnetischen Partikeln erfolgte erst durch die Inkubation mit dem Elutionspuffer.

### 3.2 LIGATIONSREAKTIONEN DER DNA-BANK MIT ADAPTERFRAGMENTEN

Die DNA-Bank wird mit einer Adaptersequenz für den anschließenden sequenzspezifischen Sortierungsschritt ligiert. Der Aufbau der DNA-Bank und der Adaptersequenzen ist unter 1.3 nachzulesen. Es ergeben sich in jeder Ligationsrunde 16 verschiedene Ligationsvarianten, da der Barcode der DNA-Bank aus zwei Basen besteht und vier mögliche Basen zur Verfügung stehen ( $4^2$ ). Die dritte Position ist ein konstanter Cytosinbaustein und somit ergeben sich 16 Ligationsvarianten für eine Runde. Nach der zweiten Runde ergeben sich schon 256 Molekülvarianten und so weiter. Nach fünf Runden wäre so eine DNA-kodierte Substanz-Bank mit einer Größe von etwa 1 Mio. verschiedener Molekülen aufgebaut. Die Adaptersequenzen besitzen eine abasische Position hinter dem ersten Codon (5'-Codon-abasische Position), um nach der Sortierung eine Spaltungsreaktion vornehmen zu können und den Verbleib des Codons an der DNA-Bank für eine neue Ligationsrunde zu garantieren. Für die Ligation mittels T4 DNA Ligase (Roboklon) müssen einige Voraussetzungen erfüllt

sein. Unter anderem muss hinter der Ligationsstelle eine Mindestanzahl hybridisierter Nukleotide vorhanden sein. Diese Voraussetzung (hier nicht dargestellt) wurde in einigen Ligationsversuchen ermittelt. Dabei ergab sich, dass die beste Ligationseffizienz entsteht, wenn der Hybridisierungsbereich aus neun Positionen besteht. Dem Codon, der abasischen Position und anschließend folgen nochmals fünf Basen bevor die einmalige 20-mer Tagsequenz beginnt. Die schematische Übersicht der Ligationsreaktion zwischen der DNA-Bank und den Adaptersequenzen mittels T4 DNA Ligase ist zum besseren Verständnis in der folgenden Abbildung dargestellt.



**Abb. 20: Schematische Darstellung der Ligationsreaktion.** Der Adapter hybridisiert teilweise mit der DNA-Bank. Dabei wird nur der komplementäre und somit spezifische Adapter von der T4 DNA Ligase direkt an das 3'-Ende der DNA-Bank ligiert.

Um die optimalen Ligationsbedingungen zu erreichen, wurden zunächst einige Vorversuche durchgeführt. Unter anderem wurde die Inkubationszeit, Temperatur und der Ablauf spezifiziert. Auch der Versuch mit zusätzlichen oder anderen Enzymen als der T4 DNA Ligase eine Steigerung in der Ligationseffizienz zu erreichen, wurden unternommen.

Bei einem Einsatz der DNA-Bank zu den Adapterfragmenten im Verhältnis 1:1 und 1:2 konnte eine geringere Effizienz der Ligation festgestellt werden, als wenn ein Verhältnis von 1:20 gewählt wurde. Dieses Verhältnis wurde für alle weiteren Ligationsversuche beibehalten.

### 3.2.1 EINFACHE LIGATION IM VERGLEICH ZUR ZYKLISCHEN LIGATION

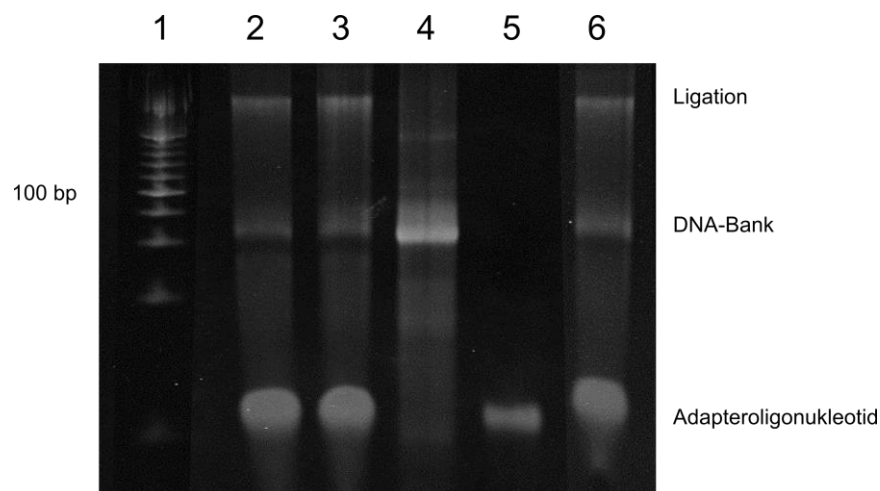
Ligationsreaktionen bestehen zumeist aus einem linearen Temperaturverlauf. Dies bedeutet, nach einem ersten Denaturierungs- und eventuellen Anlagerungsschritt folgt über eine bestimmte Zeit ein konstanter Temperaturverlauf, in der die eigentliche Ligation abläuft. Auch in dieser Arbeit wurde zunächst nach diesem Modell gearbeitet. Da die Effizienz der Ligationsreaktion stetig erhöht werden sollte, wurde der Versuch unternommen diese durch einen zyklischen Temperaturverlauf zu steigern. In der Tabelle 20 ist der Temperaturverlauf der einfachen und zyklischen Ligation dargestellt.

**Tab. 20: Programm der einfachen und zyklischen Ligation.**

Einfache Ligationsreaktion		Zyklische Ligationsreaktion	
94°C	2 min	95°C	2 min
37°C	5 min	40°C – 45°C	1 min
22°C	30 min	22°C	5 min
80°C	10 min	80°C	10 min

} 10 ×

Über die anschließende denaturierende PAGE konnte qualitativ festgestellt werden, unter welchen Reaktionsbedingungen die beste Ligationseffizienz erzielt werden kann.

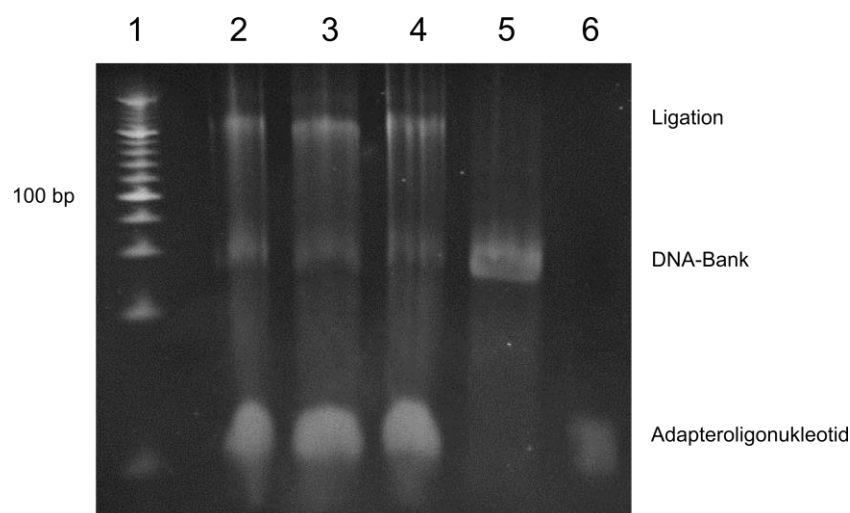


**Abb. 21: Ligationsreaktion zwischen der DNA-Bank und den Adapteroligonukleotiden auf einer den. PAGE (Ethidiumbromid-gefärbt, 12%).** O'RangeRuler 20 bp DNA Ladder (Spur 1). Zyklische Ligation mit 40°C (Spur 2) und 45°C (Spur 3) Annealingtemperatur. DNA-Bank (Spur 4). Adapteroligonukleotid (Spur 5). Einfache Ligation (Spur 6) bei 37°C Annealingtemperatur ohne zyklischen Verlauf.

In dem Gelbild kann man deutlich die unterschiedliche elektrophoretische Auftrennung zwischen den einzelnen Proben erkennen. Bei den Ansätzen in der Spur 2 und 3 handelt es

sich um die zyklischen Ligationen mit unterschiedlichen Annealingtemperaturen. Hier ist zunächst eine höhere Ligationseffizienz bei einer verwendeten Annealingtemperatur von 45°C zu verzeichnen. Im Vergleich der beiden Ligationmethoden konnte gezeigt werden, dass mit der zyklischen Ligation (Spur 3) eine höhere Ausbeute erzielt werden konnte als mit der einfachen Ligation (Spur 6). In der Spur 4 und 5 wurden die zu ligierenden Komponenten einzeln auf die PAGE aufgetragen, um so genau analysieren zu können, auf welcher Höhe die Komponenten laufen und ob diese aufgrund der Ligation nicht mehr einzeln in der Reaktion nachzuweisen sind. Dies wäre ein Indiz für eine effiziente Ligation. Auch hier ist zu sehen, dass in der zyklischen Ligation (Spur 3) die einzelnen Komponenten am geringsten und die Ligrationsbande am stärksten vorhanden ist.

Durch diesen Versuch wurde festgelegt, dass die Ligrationsreaktionen zyklisch bei einer Annealingtemperatur von 45°C und einer Inkubationstemperatur von 22°C erfolgen. Nachdem die Temperaturen und der Verlauf bestimmt wurden, sollten auch die Annealing- und Inkubationszeiten optimiert werden.



**Abb. 22:** Ligrationsreaktion zwischen der DNA-Bank und den Adapteroligonucleotiden auf einer den. PAGE (Ethidiumbromid-gefärbt, 12%). O'RangeRuler 20 bp DNA Ladder (Spur 1). Zyklische Ligationen mit gleichem Temperaturverlauf aber unterschiedlichen Annealing- und Inkubationszeiten (Spur 2-4). DNA-Bank (Spur 5). Adapteroligonucleotid (Spur 6).

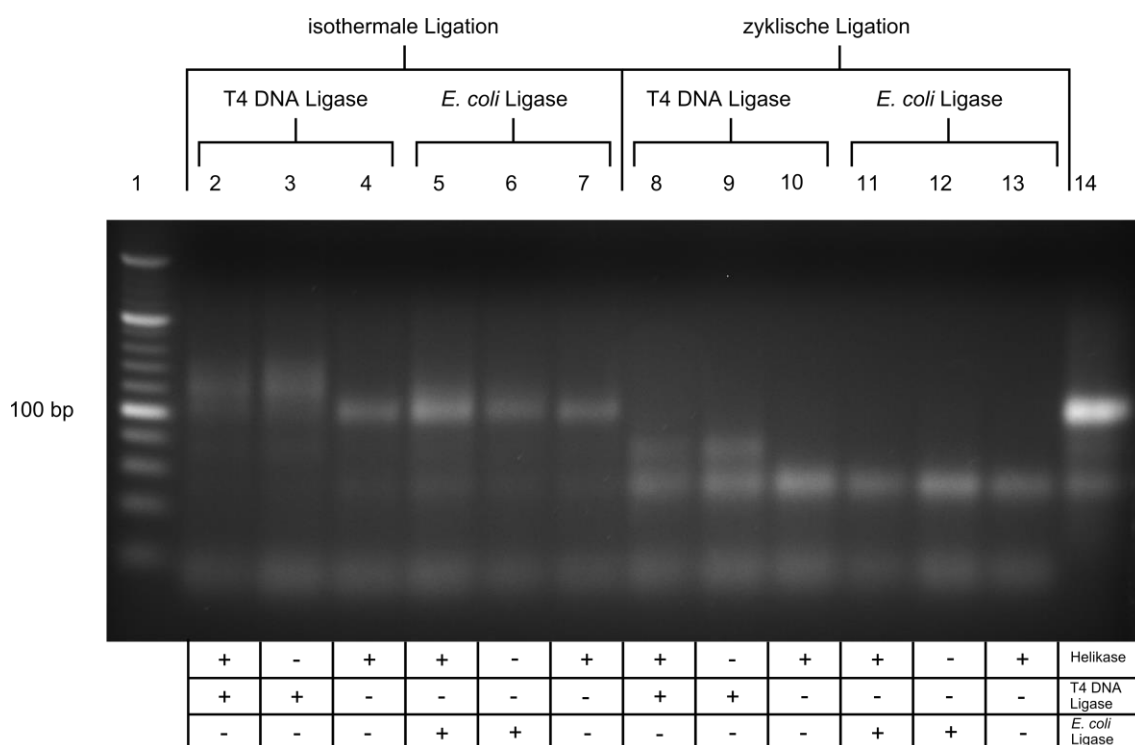
**Tab. 21:** Inkubationstemperatur und -zeiten der zyklischen Ligation (Abb. 22). Jeder Ansatz ist mit der gleichen Tempertur aber mit unterschiedlichen Inkubationszeiten behandelt worden. Rot umrandet: Ligrationsbedingung mit der besten Effizienz.

Temperatur	Spur 2	Spur 3	Spur 4
94°C	2 min	2 min	2 min
10 × { 45°C	1 min	30 s	30 s
22°C	5 min	5 min	3 min
80°C	10 min	10 min	10 min

In der denaturierenden PAGE zeigen alle Ansätze eine eindeutige Ligationsreaktion zwischen der DNA-Bank und den Adapteroligonukleotiden. Die beste Effizienz ist bei einer Annealingzeit von 30 s und Inkubationszeit von 3 min zu verzeichnen (Abb. 22, Spur 4). In dieser Reaktion ist gegenüber den anderen Reaktionen (Spur 1 & 2) die stärkste Ligationsbande zu verzeichnen. Es konnte so die Ligationseffizienz nochmals gesteigert werden, obwohl die Reaktionszeiten gegenüber den vorangegangenen Ergebnissen (Abb.21, Spur 3) verringert wurden. Die Ligationseffizienz liegt somit bei circa 50% und wird unter diesen Konditionen (siehe Tab. 21, rot umrandet) fortgesetzt.

### 3.2.2 ZYKLISCHEN LIGATION IM VERGLEICH ZUR ISOTHERMALEN LIGATION

Durch die Verwendung zweier weiterer Enzyme, als auch den Versuch die Ligationsreaktion isothermal durchzuführen, sollte die Reaktion im Allgemeinen vereinfacht bzw. nochmals versucht werden deren Effizienz zu steigern. Als Vergleich diente die zyklische Ligation, die nach dem Programm wie in Tabelle 21 (rot umrandet) durchgeführt wurde.



**Abb. 23: Isothermale und zyklische Ligationsreaktion zwischen der DNA-Bank und den Adapteroligonukleotiden auf einem Ethidium-gefärbten Agarosegel (3%).** O'RangeRuler 20 bp DNA Ladder (Spur 1). Isothermale Ligation (Spur 2-7). Zyklische Ligation (Spur 8-13). Jede Ligationsmethode wurde mit der T4 DNA Ligase und der *E. coli* Ligase, sowie der SV 40 Large Antigen Helikase durchgeführt. DNA-Bank (Spur 14).

Die isothermale Ligation wurde konstant bei 37°C für 45 min ohne vorherigen Denaturierungs- oder Annealingschritt durchgeführt. Durch die Verwendung der SV40 Large T Antigen Helikase von Roboklon sollte die Effizienz der Ligation gesteigert werden. Das Enzym sollte die falsch gepaarten abasischen Adapteroligonukleotide wieder von der DNA-Bank lösen und so freien Raum für die passenden Adapter herstellen. Gleichzeitig sollte der Einsatz der *E. coli* Ligase von New England BioLabs als Vergleichsenzym zur T4 DNA Ligase untersucht werden.

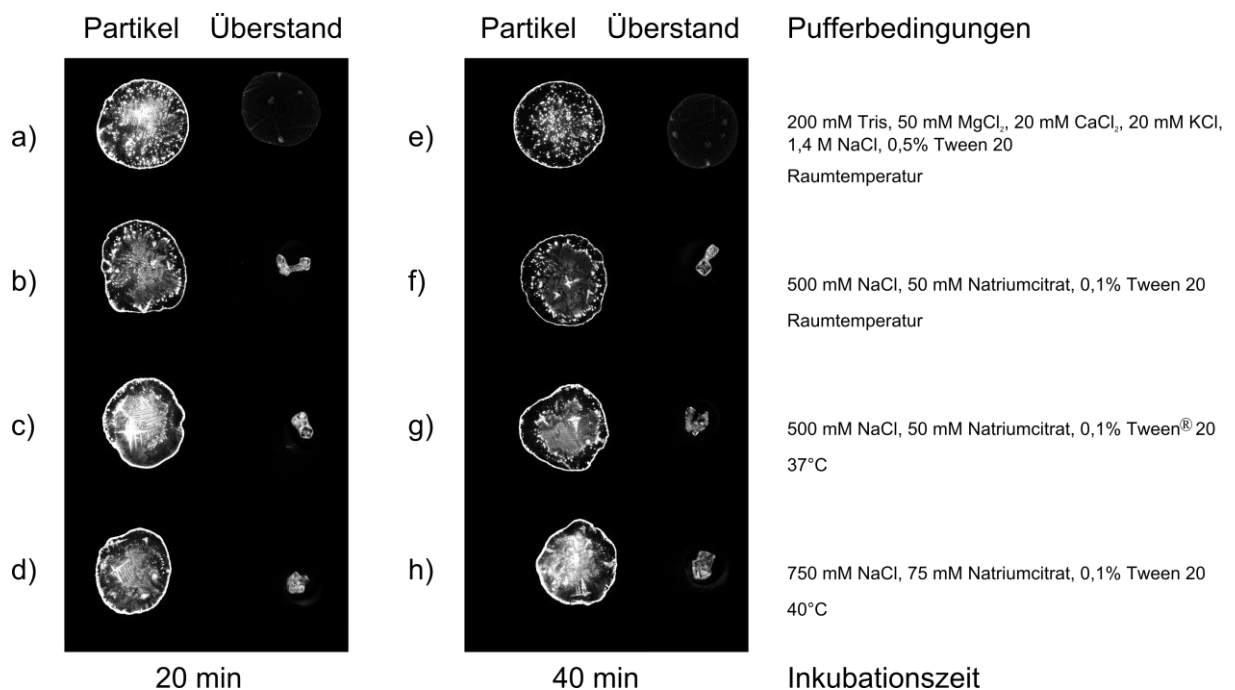
Unter isothermalen Umständen (Abb. 23, Spur 2-7) zeigen die Fragmente eine andere elektrophoretische Auftrennung als bei der zyklischen Ligation (Spur 8-13). Die DNA-Banden laufen auf einer anderen Höhe. Bei den Proben der T4 DNA Ligase (Spur 2, 3, 8 & 9) ist jeweils eine Ligationsbande zu verzeichnen. Dabei scheint der Einsatz der SV40 Large T Antigen Helikase keinen Unterschied in der Effizienz der Ligation zu zeigen. Die Reaktionen der *E. coli* Ligase (Spur 5, 6, 11 & 12) zeigen keine Ligationsbande, dabei spielen die Reaktionsbedingungen keine Rolle. Die beste Ligationseffizienz ist in Spur 9 zu erkennen. Dabei handelt es sich um die Reaktion ohne Zugabe der Helikase unter zyklischen Bedingungen, bei der die Ligation von der T4 DNA Ligase durchgeführt wird.

### **3.3 SORTIERUNG DER LIGIERTEN DNA – BANK**

Der Hauptkern dieser Arbeit, die Separation bzw. Sortierung der ligierten DNA-Bank, erfolgte auf der Grundlage von Hybridisierungsreaktionen, so dass sich nach diesem Schritt jede DNA-Bank-Variante in einem separaten Gefäß befand. Realisiert wurde dies durch die einmalige 20-mer Tagsequenz, die jedes DNA-Molekül nach der Ligation mit den Adapteroligonukleotid trug. Der spezifische Sortierungsschritt erfolgte zwischen den an magnetischen Partikeln immobilisierten Hybridisierungsoligonukleotiden und der Adaptersequenz der DNA-Bank. Die Separation wurde nach einigen manuellen Versuchen auf den Magnetpartikelprozessor KingFisher Flex für eine einfache und vor allem automatisierte Handhabung übertragen. Durch den Magnetpartikelprozessor erhält jedes Hybridisierungsoligonukleotid eine feste Position und somit kann die räumliche Separation aus einem DNA-Pool gewährleistet werden.

### 3.3.1 MANUELLE SORTIERUNG

Die Sortierung wurde durch einige Vorversuche optimiert und manifestiert. Zunächst handelte es sich um manuelle Hybridisierungsexperimente, um die passenden Parameter für die Separation zu ermitteln. In einem ersten Schritt wurden verschiedene Konditionen wie die Inkubationszeit, -puffer und -temperatur für eine optimale Reaktion bestimmt. Zu diesem Zweck wurden Hybridisierungsoligonukleotide an die magnetischen Partikel gekoppelt (siehe Abschnitt 2.2.3.1). Anschließend erfolgten die Hybridisierung mit dem fluoreszierenden Oligonukleotid Cy5aGCG und die Auswertung mit dem Laser Scanner Affimetrix 428 von Sartorius (siehe Abschnitt 2.2.18).

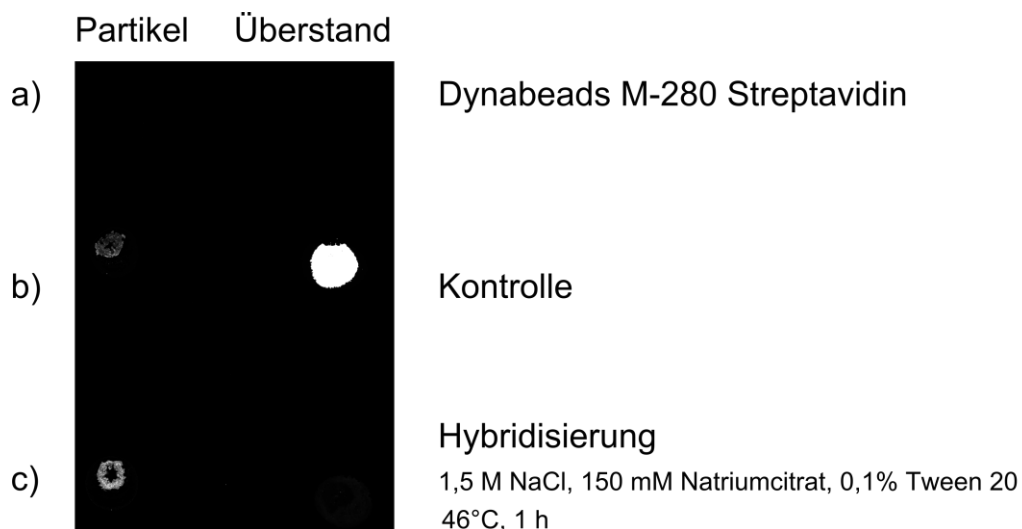


**Abb. 24: Hybridisierungssignal der magnetischen Mikropartikel.** Kopplung der Hybridisierungsoligonukleotide an SiMAG-Carboxyl Partikel (Chemicell) und anschließende Hybridisierung des komplementären Fluoreszenzoligos Cy5aGCG unter verschiedenen Pufferbedingungen, Inkubationszeiten und -temperaturen. Detektion der Hybridisierungseffizienz der Partikel und einer Probe der Hybridisierungslösung (Überstände) mittels Laser Scanner 428.

Die Inkubation der Hybridisierung erfolgte unter verschiedenen Bedingungen, deren Optimum anhand der Signalstärke der Partikel in Bezug zu den Überständen ermittelt wurde. Bei den Überständen handelt es sich um Hybridisierungslösung, die vor dem Waschschrift separiert und mit auf den Glasobjektträger aufgetragen wurde. So kann ermittelt werden, ob sich noch fluoreszierendes Oligonukleotid in der Lösung befindet oder ob alles durch die Hybridisierung abgefangen wurde. Das heißt, auf diesem Wege kann die Effizienz bestimmt

werden. Ein starkes Signal der Partikel ist bei einer Inkubationszeit von 40 min mit dem stärksten Salzgehalt des Puffers zu verzeichnen (Abb. 24, h). Allerdings ist in allen Überständen der Proben ein Fluoreszenzsignal detektiert worden. Somit ist die Hybridisierung noch nicht zu 100% effektiv.

Die Reaktionsbedingungen schienen noch nicht optimal gewählt worden zu sein. Aus diesem Grund wurde eine erneute Hybridisierung vorgenommen. Dieses Mal wurde der Salzgehalt des Puffers nochmal erhöht und die Inkubationszeit auf 1 h erhöht. Auch die Temperatur wurde von 40°C auf 46°C angehoben. Für eine bessere Effizienz wurden zudem Streptavidin-Partikel verwendet. Die Hybridisierungsoligonukleotide waren aus diesem Grund 5'-biotinyliert.



**Abb. 25: Hybridisierungssignal der Streptavidin-Mikropartikel.** Probe der Partikel (a), um Hintergrundaktivitäten auszuschließen. Die Reaktion (b) diente als Kontrolle, um die Spezifität der Hybridisierung zu ermitteln. Ansatz (c) zeigt die Hybridisierung des immobilisierten Oligos mit dem komplementären Fluoreszenzoligo Cy5aGCG.

Ein Fluoreszenzsignal ist in der Hybridisierungsreaktion (Abb. 25, c) klar zu verzeichnen, keines allerdings im Überstand. Somit hat die Reaktion sehr stringent funktioniert und die Pufferbedingungen scheinen optimal zu sein. Auch die Kontrolle (b) bestätigt dieses Ergebnis. In dieser sind die verwendeten Oligonukleotide nicht komplementär und so sollte kein Signal bei den Partikel aber ein starkes im Überstand vorhanden sein. Dieses Ergebnis wurde im Folgenden auf das automatisierte System übertragen.

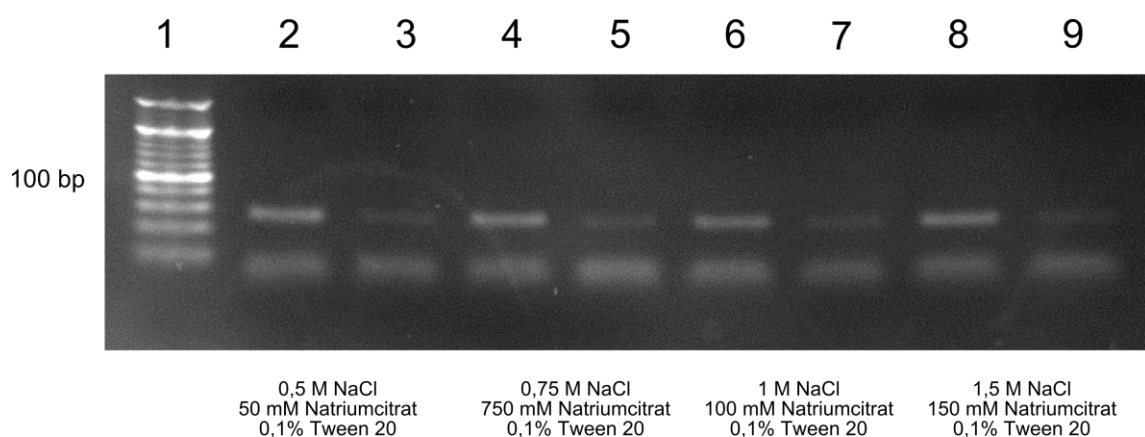


### 3.3.2 AUTOMATISIERTE SORTIERUNG PER KINGFISHER FLEX ROBOTER

Für ein automatisiertes Verfahren müssen die einzelnen Schritte im Laufe der Arbeit auf den Magnetpartikelprozessor übertragen werden. Bevor die Separation einer Ligationsreaktion aus einem Pool von Ligationen durchgeführt wurde, erfolgten definierte Separationen mit einzelnen Ligationen, um die Übertragung der Pufferbedingungen auf das automatisierte System vorzunehmen.

Zu diesem Zweck wurden zwei Hybridisierungsoligonukleotide verwendet, um diese an die magnetischen Partikel zu koppeln. Anschließend wurde eine Ligationsreaktion durchgeführt. Diese ist nur zu einem der Hybridisierungsoligos komplementär, der andere dient als Kontrolle, um die Spezifität zu ermitteln. Insgesamt wurden vier Separationen in einzelnen Kammern vorgenommen, um so vier Pufferbedingungen testen zu können. Die Inkubation erfolgte bei 40°C für 40 min durch Auf- und Abbewegungen der Magnetspins in 10 ml Hybridisierungslösung. An den Magnetspins haften die immobilisierten Hybridisierungsoligonukleotide, die nun die komplementäre Ligation mittels Hybridisierung binden sollen. In jeder Kammer wurden zwei Pins der insgesamt 16 zur Verfügung stehenden mit Hybridisierungsoligos belegt. Ein komplementäres und ein nicht-komplementäres Oligo zur definierten Ligation.

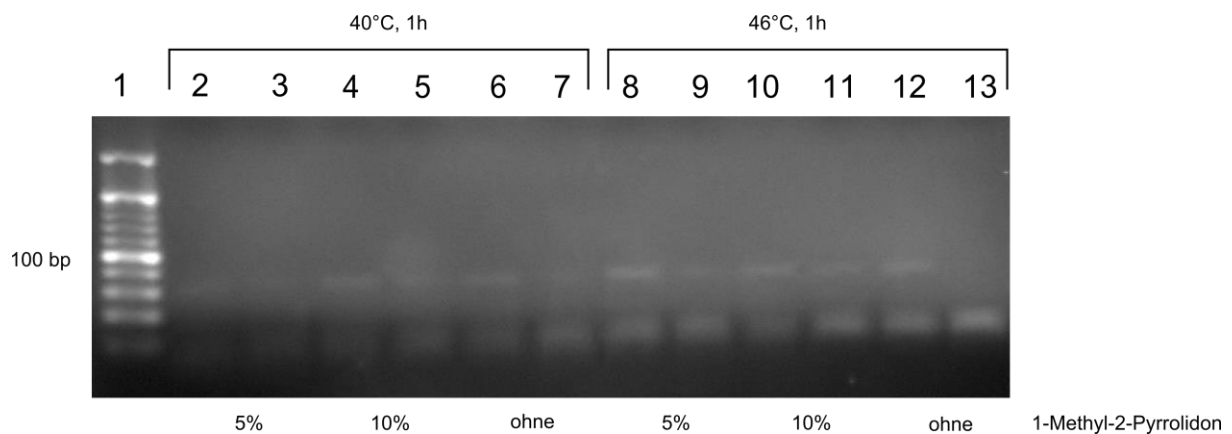
Im Anschluss wurden die Partikel aufgenommen und durch eine PCR konnte nachgewiesen werden, ob die Ligation an den komplementären Oligonukleotiden hybridisiert hatte oder ob es zu Hintergrundbindungen gekommen ist.



**Abb. 26: PCR-Ausbeute der Separation auf einem Ethidium-gefärbten Agarosegel (3%).** O'RangeRuler 20 bp DNA Ladder (Spur 1). PCR-Fragmente der Separation mit komplementären Hybridisierungsoligo (Spur 2, 4, 6, 8) und nicht-komplementären Oligo (Spur 3, 5, 7, 9) mit vier Pufferbedingungen. Jede Pufferbedingung wurde mit passenden und nicht-passenden Hybridisierungsoligo durchgeführt.

In der Abbildung 26 ist die PCR-Ausbeute der Separation unter verschiedenen Pufferbedingungen dargestellt. Zu jeder Pufferkondition gibt es zwei Proben. Bei der einen Probe (Spur 2, 4, 6, 8) handelt es sich um die PCR-Ausbeute nach erfolgter Separation. Diese Ansätze beinhalteten das komplementäre immobilisierte Hybridisierungsoligonukleotid zur Ligation. Bei der anderen Probe (Spur 3, 5, 7, 9) war das immobilisierte Oligo nicht-komplementär zur Ligation. Die beste PCR-Ausbeute (Spur 8) mit dem geringsten Hintergrund (Spur 9) ist mit den Pufferbedingungen (1,5 M NaCl, 150 mM Natriumcitrat, 0,1% Tween 20) zu verzeichnen.

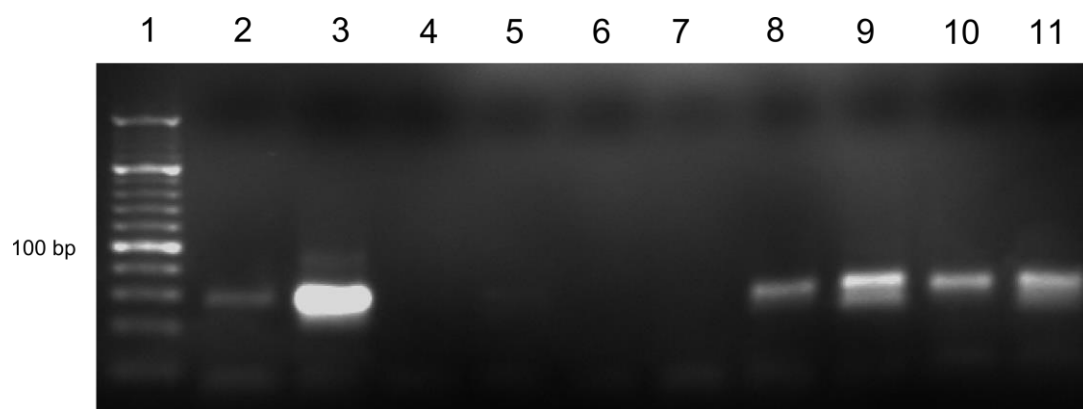
Die Stringenz sollte durch den Einsatz von 1 Methyl-2-Pyrrolidon (5% & 10%) im vorher analysierten Puffer bei einer Inkubationstemperatur von 40°C und 46°C für 1 h erhöht werden. Der Versuchsaufbau wurde, wie oben beschrieben, beibehalten.



**Abb. 27: PCR-Ausbeute der Separation auf einem Ethidium-gefärbten Agarosegel (3%).** O'RangeRuler 20 bp DNA Ladder (Spur 1). PCR-Fragmente der Separation mit komplementären Hybridisierungsoligo (Spur 2, 4, 6, 8, 10, 12) und nicht-komplementären Oligo (Spur 3, 5, 7, 9, 11, 13) mit gleichem Ausgangspuffer, aber der Zugabe von 5% und 10% 1-Methyl-2-Pyrrolidon. Jede Reaktion wurde bei 40°C und 46°C für 1 h durchgeführt.

Die Zugabe von 5% oder auch 10% 1-Methyl-2-Pyrrolidon konnte den Hintergrund sowohl bei 40°C als auch bei 46°C nicht verringern. Die beste Ausbeute der Separation (Spur 12) mit dem geringsten Hintergrund (13) ist bei der Verwendung des Puffers ohne Zusatz von 1-Methyl-2-Pyrrolidon zu verzeichnen. Durch die Erhöhung der Temperatur wurde die Stringenz und Spezifität der Separation soweit erhöht, dass kein Hintergrund in diesem Ansatz zu erkennen ist. Somit konnte die Ligation spezifisch an die komplementären Hybridisierungsoligonukleotide separiert werden. Die folgenden Reaktionsbedingungen (1,5 M NaCl, 150 mM Natriumcitrat, 0,1% Tween 20) bei 46°C für 1 h werden für alle zukünftigen Separationsreaktionen beibehalten (siehe Anhang: Separation).

Da die Pufferbedingungen, Inkubationszeiten und -temperaturen optimiert wurden, erfolgte eine Separation mit allen 16 Ligationsvarianten an alle 16 immobilisierten Hybridisierungsoligonukleotide in einer Inkubationskammer auf dem Magnetpartikelprozessor. Die Zusammensetzung und Konzentrationen der Ligationsreaktionen sind unter Abschnitt 2.2.4, Tab. 12 nachzulesen. Anschließend wurde in die 6-Well-Platte 10 ml Hybridisierungspuffer und der Pool der Ligationsreaktionen gegeben. In die 96er Deepwell-Platte (Thermo Scientific) wurden 16 Kammern mit den Mikropartikeln, welche die kovalent gekoppelten Hybridisierungsoligonukleotide tragen, beladen (siehe Abb. 13). Alle Platten wurden auf den Roboter geladen und die Separation konnte gestartet werden (siehe Anhang: Separation). Zum genauen Verständnis dient Abbildung 12. Im Anschluss wurden die magnetischen Partikel mit der sortierten Ligation in neue Reaktionsgefäße überführt. Mittels einer PCR sollte nun die Spezifität der Separation analysiert werden. Zu diesem Zweck wurde eine PCR auf einige ausgewählte Separationen mit einem spezifischen Primer P3IIIGCG durchgeführt. Dieser Primer sollte nur ein Amplifikat erzeugen, wenn auch die passende Ligation (mit dem Adapter aGCG) separiert wurde. Bei allen andern separierten Ligationen sollte kein Amplifikat entstehen.



**Abb. 28: PCR-Ausbeute der Separation auf einem Ethidium-gefärbten Agarosegel (3%).** O'RangeRuler 20 bp DNA Ladder (Spur 1). Spezifische Amplifikation der Separation (aGCG) mit dem Primer P3IIIGCG nach 10-Zykeln (Spur 2) und 15-Zyklen (Spur 3). Amplifikation mit dem spezifischen Primer auf nicht-komplementäre Separationen (Spur 4-7). Unspezifische Amplifikation auf die Separationen (Spur 8-11). Mit allen Proben wurden 10- (Spur 2, 4, 6, 8, 10) bzw. 15-Zyklen durchgeführt (Spur 3, 5, 7, 9, 11).

In dem Gelbild (Abb. 28) sind die PCR-Fragmente der Separation dargestellt. In der Spur 2 (10-Zyklen) und 3 (15-Zyklen) ist die PCR-Reaktion des spezifischen Primers auf die passende Separation zu erkennen. Bereits nach 10-Zyklen ist eine starke distinkte Bande auf der richtigen Höhe (circa 60 bp) zu verzeichnen. Durch die Kontrollen (Spur 4-7) kann letztendlich erst die Spezifität dieser Reaktion festgestellt werden. Wenn sich auch dort PCR-

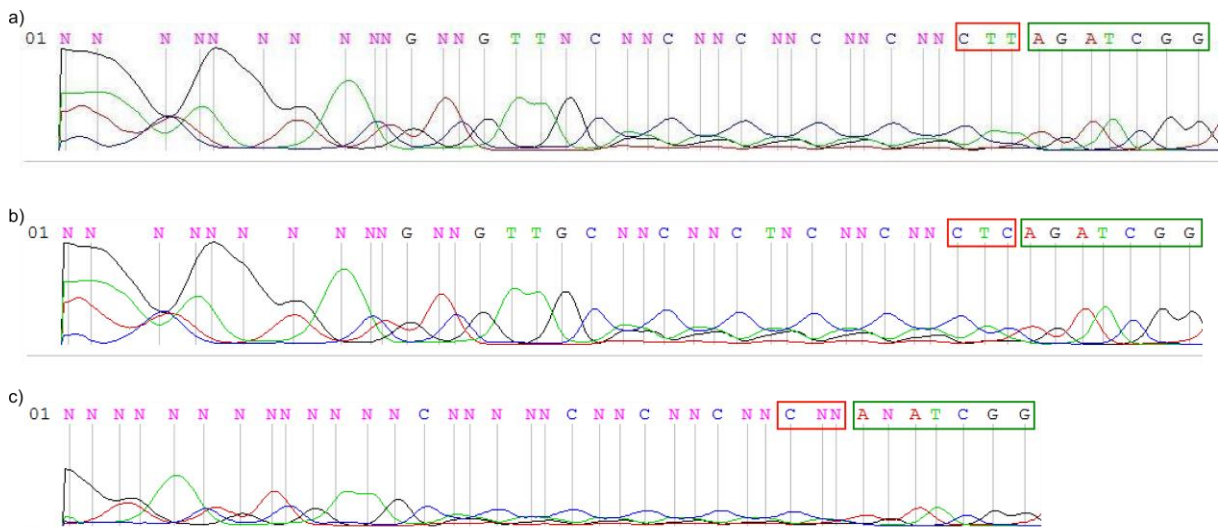
Fragmente wiederfinden, dann erfolgt die Sortierung nicht spezifisch. Bei diesen PCR-Ansätzen handelt es sich um Separationen, bei denen der spezifische Primer kein Amplifikat erzeugen sollte. Da keinerlei Fragment in diesen Ansätzen wahrzunehmen ist, konnte gezeigt werden, dass die Sortierung sehr stringent verläuft und keine Hintergrundhybridisierungen wahrzunehmen sind. Vor allem die Kontrolle (Spur 5, 6) zeigt diese starke Spezifität auf. Bei dieser Separation handelt es sich um eine Ligation mit dem Adapter aGGG, somit unterscheiden sich die beiden Proben lediglich um eine Base und trotzdem hat der spezifische Primer kein Amplifikat erzeugt. Die PCR-Reaktionen in Spur 8-11 wurden mit unspezifischen Primern durchgeführt und passen auf alle Separationen. Diese dienen lediglich dem allgemeinen Nachweis der Sortierung falls die spezifische Amplifikation kein Ergebnis erzeugt hätte. Die Sortierung der ligierten DNA-Bank verläuft sehr stringent unter diesen Inkubationsbedingungen auf dem Magnetpartikelprozessor.

### **3.3.3 ANWENDUNG DER SANGER-SEQUENZIERUNG FÜR DIE ANALYSE DER SORTIERUNG**

Für die Analyse der Sortierung wurde die Sequenzierungsmethode nach Sanger angewendet (Sanger *et al.*, 1977). Es sollten genaue Aussagen über die Spezifität und Hintergrundhybridisierungen gemacht werden. Zu diesem Zweck bot sich die Sanger-Sequenzierung nach einer Separationsrunde an. Es erfolgte eine Sortierung (siehe Abschnitt 2.2.5) mit allen 16 immobilisierten Hybridisierungsoligonukleotiden und den Ligationsreaktionen (siehe Abschnitt 2.2.4) in einer 6-Well-Platte. Nach Überführung der 16 sortierten Ligationen in neue Reaktionsgefäße erfolgte auf den Partikeln eine PCR-Reaktion, um das sortierte DNA-Fragment zu amplifizieren und gleichzeitig von den magnetischen Partikeln zu eluieren. Abschnitte der verwendeten Primer (Sindex1 und Multiplexing Primer 1.0) dienen dem Sequenzierungsteam als Sequenzierungsprimer. Die Sequenzierung wurde vom Service-Team des MPI für molekulare Genetik vorgenommen.

Durch die Amplifizierung des Genotyps (DNA-Bank) kann ermittelt werden, welches Adapteroligonukleotid zuvor ligiert wurde. Diese Erkenntnis läßt den Schluß zu, ob die Separation spezifisch an dem komplementären Hybridisierungsoligonukleotid erfolgte. Die einmalige 20-mer Tagsequenz der Adapter ist komplementär zu den Hybridisierungsoligonukleotiden und jede Tagsequenz ist einmalig zu deren

Kodierungsregion (siehe Abb. 7). Nach der Separation befinden sich alle DNA-Fragmente an den bekannten Hybridisierungsoligonukleotiden vor. Durch eine Sequenzierung kann bestimmt werden, ob die DNA-Bank auch richtig sortiert wurde. Bei der Sequenzierung wurde der *sense* Strang der DNA-Bank H<sub>2</sub>NhpIIIc6V16 generiert. Somit muss das gesuchte Barcodon zum Adaptertriplett direkt komplementär sein. Die DNA-Bank besitzt eine konstante Region, die den Barcode flankiert und somit beginnt das gesuchte Barcodon direkt hinter dieser konstanten Region.



**Abb. 29: Sanger-Sequenzierungsdaten.** Der sense-Strang der DNA-Bank H<sub>2</sub>NhpIIIc6V16 wurde generiert. Durch das Basentriplett der DNA-Bank (Genotyp) kann auf die Basenabfolge des ligierten Adapteroligonukleotids rückgeschlossen und so Aussagen über die Spezifität der Separation an die komplementären Hybridisierungsoligonukleotide getroffen werden. Ausschnitt der Sequenzierungsdaten für die Adapteroligonukleotide aAAG (a) und aGAG (b) an deren komplementären Hybridisierungsoligonukleotide hybAAG und hybGAG. Sequenzdaten der initialen DNA-Bank (c). Die dargestellten Sequenzen verlaufen in 5'→ 3' Richtung. Rote Umrandung: identifizierbares Barcodon. Grüne Umrandung: konstante Region der DNA-Bank.

Die dargestellten Sequenzen verlaufen in 5'→3'-Richtung und somit befindet sich am 3'-Ende die konstante Region [AGATCGG]. In 5'-Richtung schließt direkt das Barcodon an. Die konstante Region wurde in der Abbildung 29 grün umrandet, die gesuchte Triplettssequenz wurde rot markiert. Es wurde von den 16 Sequenzierungsdaten der Sortierung nur zwei gezeigt, sowie die initiale DNA-Bank als Vergleichsprobe. Die DNA-Bank besitzt an jeder dritten Stelle des Barcodes eine Konstante. Bei dieser konstanten Base handelt es sich um den Baustein Cytosin. In der Abbildung 29 kann man bei allen drei Proben deutlich erkennen, dass der Cytosinbaustein konstant an jeder dritten Position im Barcodon vorkommt. Bei den Sequenzierungsdaten in Abb. 29, a handelt es sich um die Separation an das Hybridisierungsoligonukleotid hybAAG. Bei einer spezifischen Separation sollte das

Adapteroligonukleotid aAAG hybridisiert und die gesuchte Sequenz im Genotyp somit folgende Abfolge haben 5' TTC (Abb.29, a rot umrandet). Auch bei der Separation an das Hybridisierungsoligonukleotid hybGAG konnte eine eindeutige und sehr spezifische Separation (b, rot umrandet) gezeigt werden. Die gesuchte Sequenz ist hier 5' CTC. Bei der initialen DNA-Bank (Abb. 29, c) hingegen sollte nur der konstante Baustein Cytosin im Barcode an jeder dritten Stelle detektierbar sein, an den anderen Positionen müssen Wobble-Basen vorhanden sein. Ansonsten handelt es sich nicht um eine DNA-Bank. Durch die Sanger-Sequenzierung konnte auch dies nachgeprüft und aufgezeigt werden. Hier wurden die eindeutigsten Sequenzierungsdaten dargestellt, die anderen Datensätze sind im Anhang nachzulesen.

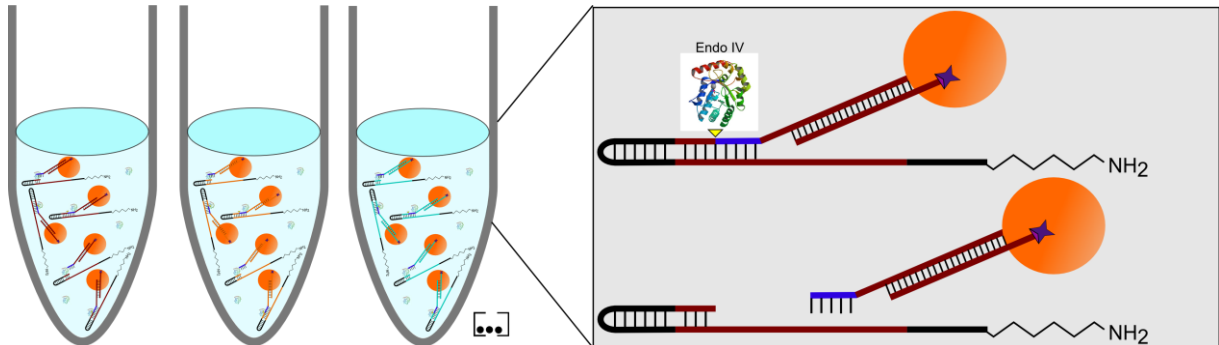
### **3.4 SPALTUNGSMECHANISMUS AN DER SORTIERTEN DNA – BANK**

Für eine zweite Syntheserunde muss an der DNA-Substanz-Bank eine Spaltungsreaktion erfolgen. Auf diesem Wege wird die DNA-kodierte Bank von dem Adapter gelöst und für eine zweite Ligations- als auch Separationsrunde mit anschließender Synthese freigegeben. Es wurden in dieser Arbeit zwei Varianten getestet. Zum einen die chemische Elution und zum anderen die enzymatische Spaltung.

Bei der chemischen Elution sollte das sortierte DNA-Fragment durch den Einsatz von 0,5 M EDTA oder 0,05 M NaOH bei 70°C für 10 min von den Streptavidin-Mikropartikeln gelöst werden. In diesem Fall würde im Anschluss die Synthese erfolgen und erst danach die Spaltungsreaktion an der abasischen Position der DNA-kodierten Bank, um eine neue Syntheserunde starten zu können.

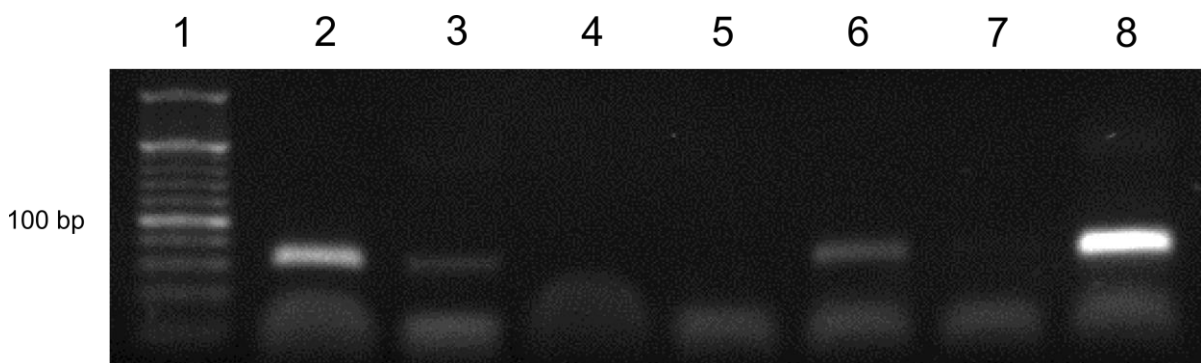
Durch die enzymatische Spaltungsreaktion an der abasischen Position der sortierten DNA-Bank wird das Fragment zum einen von den Streptavidin-Partikeln gelöst und gleichzeitig um drei Nukleotidbausteine verlängert. Bei der Nukleotidverlängerung handelt es sich um den Kodierungsbereich des Adapteroligonukleotids, welcher komplementär zu dem identifizierbaren Barcode der DNA-Bank ist (siehe Abbildung 30). Nach der Synthese kann dem angefügten Molekülbaustein dieses Barcodon eindeutig zugeordnet werden. Durch diese Spaltungsreaktion an der ligierten DNA-Bank wird die Ligationsreaktion in dem zweiten Zyklus um drei Basen versetzt und das nächste dekodierende Barcodon der DNA-Bank zur Ligation des Adapteroligos verwendet. Für diese Variante der Spaltungsreaktion wurde mit

der Endonuklease IV (New England BioLabs) gearbeitet. Dabei handelt es sich um eine AP-Endonuklease, welche durch ihre Aktivität eine Spaltung an der 5'-Position der abasischen Stelle im DNA-Fragment vornimmt (Bogenhagen & Pinz, 1998).



**Abb. 30: Spaltungsreaktion der sortierten DNA-Bank mittels Endonuklease IV.** Die sortierte DNA-Bank befindet sich in separate Gefäße, wo die Spaltungsreaktion durchgeführt wird (Abb.links). Die Endonuklease IV nimmt eine Spaltung an der abasischen Position im DNA-Fragment vor (Abb. rechts, oben). Auf diese Weise wird die dekodierte DNA-Bank von den Mikropartikeln gelöst und gleichzeitig um drei Nucleotidbausteine verlängert (Abb. rechts, unten) und liegt im Anschluss frei im Puffer vor. An den Mikropartikeln befindet sich das von der DNA-Bank gespaltene Adapteroligo, verkürzt um drei Basen.

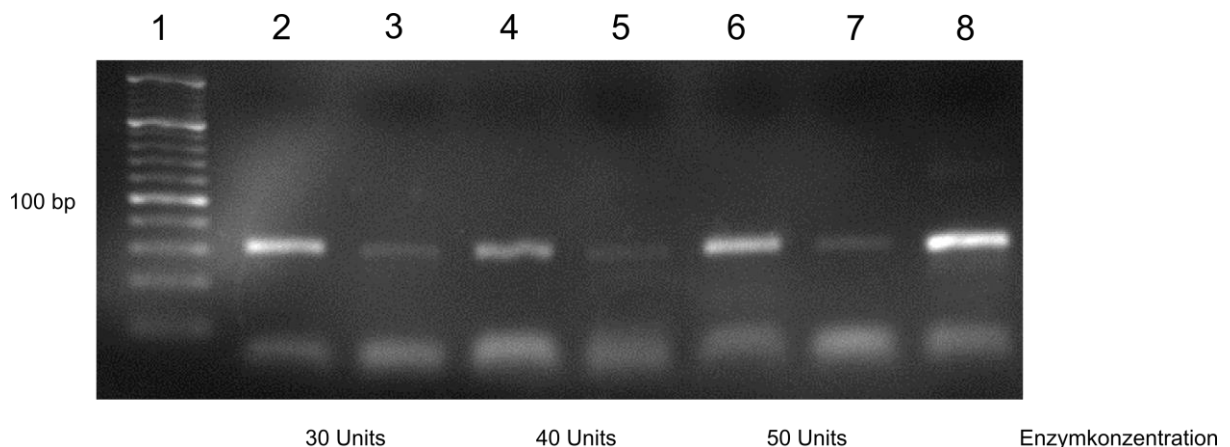
Um genaue Aussagen über die Effizienz der Spaltungsvarianten machen zu können, wurde die chemische Elution mit der enzymatischen Spaltung verglichen. Nach erfolgter Inkubation wurden die Partikel von der Lösung separiert und die Rückgewinnung der separierten DNA-Fragmente von den Mikropartikeln mittels PCR nachgewiesen. Zu diesem Zweck erfolgte sowohl eine PCR-Reaktion auf die gelöste DNA-Bank als auch auf die Mikropartikel. Auf diesem Wege können Aussagen über den Verbleib der DNA-Bank auf den Mikropartikeln gemacht werden, also wie effizient die Elution bzw. die Spaltungsreaktion ist.



**Abb. 31: PCR-Produkte der Spaltungsreaktion und Elution der sortierten DNA-Bank von den magnetischen Stretavidin-Partikeln auf einem Ethidium-gefärbten Agarosegel (3%).** O'RangeRuler 20 bp DNA Ladder (Spur 1). PCR-Produkt der DNA-Bank (Spur 2) und der dazugehörigen Mikropartikel (Spur 3) nach der enzymatischen Spaltung. PCR-Produkt der DNA-Bank nach der Elution mit EDTA (Spur 4) und NaOH (Spur 6). PCR-Produkte von den Mikropartikeln nach der Elution mit EDTA (Spur 5) und NaOH (Spur 7). Eingesetztes Ausgangsmaterial der DNA-Bank als Vergleichskontrolle der Effizienz (Spur 8).

In der zweiten Spur befindet sich die enzymatische Spaltung der separierten Ligation an Streptavidin-Beads. Diese Spaltung an der abasischen Position des separierten Fragments wurde mittels der Endonuklease IV durchgeführt und im Anschluss eine PCR gefahren, um die Fragmente sichtbar zu machen. Die sichtbare Bande (Spur 2) läuft auf derselben Höhe wie die Vergleichskontrolle (Spur 8). Bei dieser Vergleichskontrolle handelt es sich um die PCR-Produkte der separierten DNA-Bank, ohne dass eine Spaltungs- oder Elutionsreaktion durchgeführt wurde. Auf diese Weise kann verglichen werden, wieviel DNA-Fragment von den Mikropartikeln rückgewonnen werden kann. Im Vergleich zu der Kontrolle ist nach der Spaltung mit der Endonuklease IV (Spur 2) eine gute Rückgewinnung der separierten DNA-Fragmente zu verzeichnen. Vor allem im Vergleich zu den Elutionsreaktionen (Spur 4 und 6) war eine wesentlich höhere DNA-Ausbeute möglich. Auch die Rückstände der DNA-Bank auf den Mikropartikeln (Spur 3, 5 und 7) sind sehr gering und zeigen ein sehr schwaches Signal. Nach diesem Ergebnis wurde für die Spaltungsreaktion ausschließlich mit der Endonuklease IV gearbeitet.

Um die Spaltungsreaktion zu optimieren, wurde eine Titrationsreihe der Enzymkonzentration vorgenommen. Nach der Inkubation erfolgte auch hier eine PCR-Reaktion auf die rückgewonnene DNA-Bank und die freigewordenen Mikropartikel.



**Abb. 31: PCR-Produkte nach der Spaltungsreaktion der sortierten DNA-Bank von den magnetischen Streptavidin-Partikeln mit unterschiedlicher Enzymkonzentration auf einem Ethidium-gefärbten Agarosegel (3%).** O'RangeRuler 20 bp DNA Ladder (Spur 1). PCR-Produkt der DNA-Bank und der dazugehörigen Mikropartikel (Spur 3, 5, 7) nach der enzymatischen Spaltung mit einer Enzymkonzentration von 30 Units (Spur 2), 40 Units (Spur 4) und 50 Units (Spur 6). Eingesetztes Ausgangsmaterial der DNA-Bank als Vergleichskontrolle der Effizienz (Spur 8).

Zu den 25 pmol separierter Ligrationsreaktion (hybridisiert mit den immobilisierten Hybridisierungsoligonukleotiden, siehe Abschnitt Separation 2.2.5) wurden zwischen 30 U – 50 U Endonuklease IV in 1× NEB Puffer 3 mit einem Endvolumen von 30 µl gegeben.



Die Detektion der Spaltungseffizienz erfolgte auch hier über ein 3%-iges Agarosegel. Eine Erhöhung der Enzymkonzentration von 30 Units (Spur 2) auf 50 Units (Spur 6) hatte keine höhere Spaltungseffizienz, also Rückgewinnung der sortierten DNA-Bank zur Folge. Die eingesetzte Enzymkonzentration von 30 Units auf 25 pmol separierter Ligation zeigte die beste Rückgewinnung der DNA-Fragmente von den Streptavidin-Mikropartikeln.

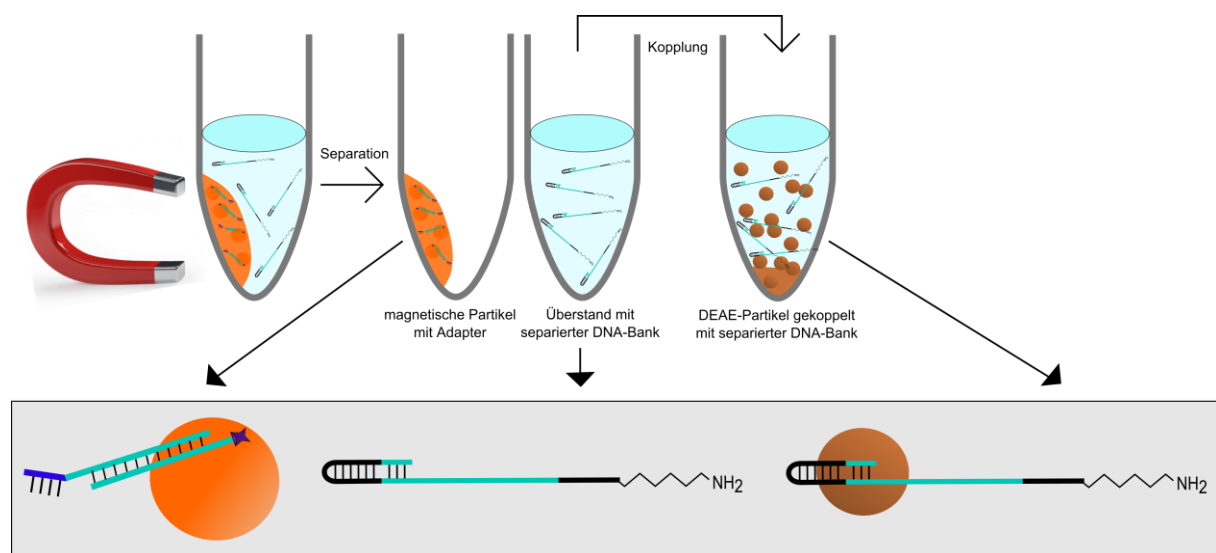
### **3.5 SYNTHESE DER DNA-KODIERTEN SUBSTANZ – BANK UND SELEKTION GEGEN DAS ZIELMOLEKÜL STREPTAVIDIN**

Die Synthesereaktion beinhaltet die Konjugation der niedermolekularen Stoffe Glycin und d-Biotin durch kombinatorische chemische Synthese an die sortierte DNA-Bank. Der Barcode dient der Identifizierung des gekoppelten Wirkstoffes bzw. in unserem Fall der Zuordnung der verwendeten Synthesebedingungen.

Die verwendete Synthesereaktion entspricht der Festphasenpeptidsynthese, die mittels der Carbodiimid-Methode vorgenommen wurde. Die 1963 beschriebene Peptidsynthese am festen Träger (Merrifield, 1963) hat die Peptidsynthese bzw. -chemie revolutioniert. Durch die Festphasenankerung können sämtliche Reagenzien und Nebenprodukte schnell und einfach isoliert werden (Altmann & Mutter, 1993) und bei jedem Schritt des Zyklus können chemische Schutzgruppen unerwünschte Reaktionen verhindern (Cox & Nelson, 1993). Das Grundprinzip der Peptidsynthese beruht darauf, dass die Aminogruppe der zu koppelnden Aminosäure durch Einführung einer reversiblen schützenden Gruppe stabilisiert wird. Dann wird das Carboxyl der Aminosäure so verändert, dass es zur Kopplung mit einer zweiten Aminosäure befähigt ist (Bergmann & Zervas, 1932). Nach der Synthese kann die schützende Gruppe entfernt und ein neuer Synthesedurchgang gestartet werden. Heutzutage gibt es ein umfangreiches Arsenal verschiedener Schutzgruppen, dennoch haben sich nur einige in der Praxis durchgesetzt (Altmann & Mutter, 1993). Zu den klassischen Schutzgruppen gehört auch die Fluorenylmethoxycarbonyl (Fmoc-)-Gruppe, die in dieser Arbeit verwendet wurde, um die N-terminale Aminosäure des zu koppelnden Glycins zu schützen. Nach dem Synthesedurchgang konnte durch eine basische Abspaltung mittels Piperidin die Fmoc-Schutzgruppe wieder entfernt werden. Dabei bildet sich durch Protonenabstaktion des Fluorensystems zunächst ein Carbanion, welches durch eine Eliminationsreaktion zerfällt. Die Decarboxylierung führt letztendlich zur Freisetzung der Aminogruppe (Altmann & Mutter,

1993). Zur Aktivierung der Carboxykomponente wurde in dieser Arbeit die Carbodiimid-Methode gewählt. Carbodiimide sind ungesättigte Verbindungen mit einer Allenstruktur (Nakajima & Ikada, 1995). Durch den elektronenanziehenden Charakter dieser Verbindungen wird die Partialladung auf dem Carbonylkohlenstoff erhöht und dessen Reaktivität gegenüber der Aminogruppe gesteigert (Altmann & Mutter, 1993). Die verwendeten Kopplungsreagenzien sind EDC, DMT-MM und CDI. Diese wurden in verschiedenen Konzentrationen in unterschiedlichen Lösungsmitteln wie Methanol, Ethanol, DMF und bidest H<sub>2</sub>O verwendet. Auf diesem Wege ergaben sich 16 verschiedene Synthesebedingungen (siehe Tabelle 14), deren Effizienz ermittelt werden sollte.

Die sortierte DNA-Bank wurde nach der Spaltungsreaktion für die Synthesereaktion an SiMAG-DEAE Mikropartikeln (siehe Abbildung 32) immobilisiert, wodurch eine einfache Handhabung der Waschschrte und anschließenden Schutzgruppenabspaltung ermöglicht wurde.



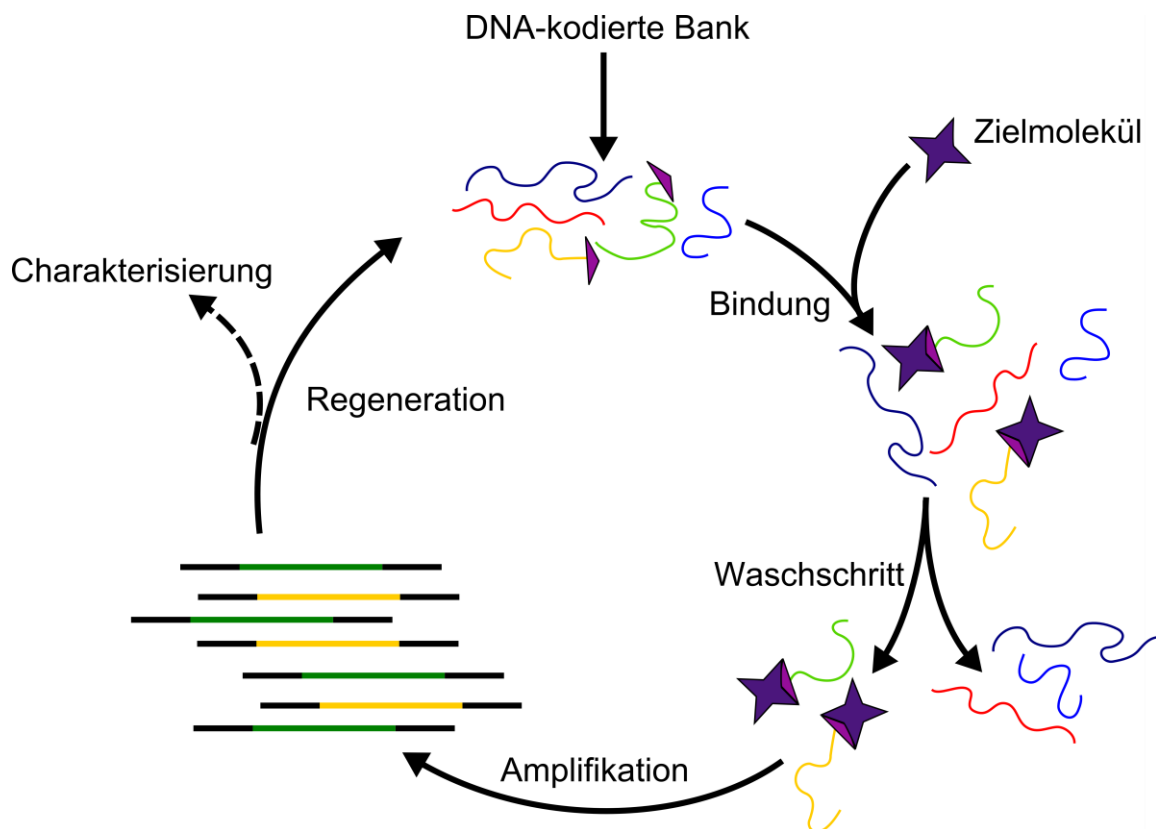
**Abb. 32: Immobilisierung der gespaltenen DNA-Bank an SiMAG-DEAE Mikropartikeln.** Mithilfe des Magnetpartikelseparators DynaMag wurden die magnetischen Partikel, welche die abgespaltenen Adapter tragen von der DNA-Bank separiert. Anschließend wurden zu dem Überstand mit der DNA-Bank die ionischen Partikel gegeben und so konnte eine Kopplung an SiMAG-DEAE Mikropartikeln erfolgen.

Bevor zwei vollständige Syntheserunden durchgeführt wurden, bei der zuerst die Aminosäure Glycin und nach einem weiteren Ligations-, Spaltungs- und Sortierungsschritt d-Biotin unter 16 verschiedenen Synthesebedingungen an die DNA-Bank geankert wurden, erfolgten einige Vorversuche. Auf diesem Wege wurde die Effizienz der Synthese und Spezifität bestimmt.

Bei den zu koppelnden niedermolekularen Stoffen handelte es sich zum einen um Glycin und zum anderen um d-Biotin. In der ersten Syntheserunde wurde Glycin und in der zweiten

d-Biotin an die DNA-kodierte Bank gekoppelt. Wenn nur eine Syntheserunde durchgeführt wurde, erfolgte direkt die Kopplung mit d-Biotin.

Auf diese Weise konnte im Anschluss zur Charakterisierung der besten Synthesebedingungen eine Selektion gegen das Zielmolekül Streptavidin durchgeführt werden. Der Ablauf der Selektion ist in Abbildung 33 dargestellt.



**Abb. 33:** *In vitro* Selektion der DNA-kodierten Substanz-Bank gegen ein Zielmolekül.

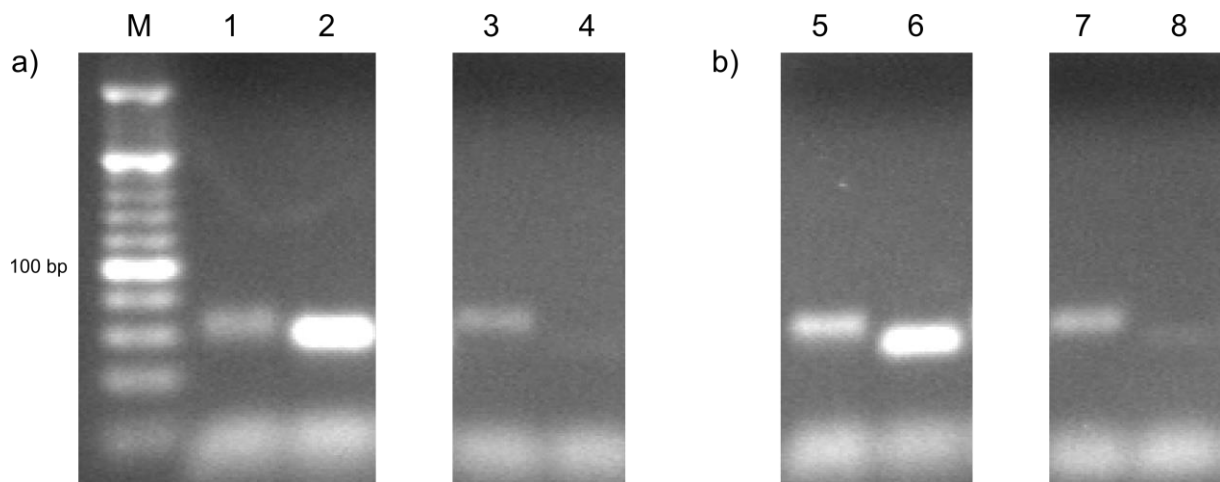
Für die Selektion wurde die synthetisierte DNA-kodierte Substanz-Bank mit einem Zielmolekül (Dynabeads M-280 Streptavidin von Invitrogen) für 30 min bei RT (22°C) inkubiert. Nach einem Waschschritt, bei dem alle nicht gebundenen DNA-Fragmente separiert wurden, konnte eine Amplifikation erfolgen, ohne die DNA-Fragmente vorher von den Partikeln isolieren zu müssen. Mittels Hochdurchsatzsequenzierung erfolgte im Anschluss eine Charakterisierung der besten Synthesebedingungen.

### 3.5.1 SYNTHESEREAKTION MIT D-BIOTIN UND SELEKTION GEGEN DAS ZIELMOLEKÜL STREPTAVIDIN

Zunächst wurde getestet, ob die Synthesereaktion am 5'-Aminolinker der DNA-Bank erfolgt oder ob es auch zu einer chemischen Konjugation des d-Biotins an den Nukleotidbausteinen der DNA-Fragmente kommt.

Zu diesem Zweck wurde zum einen mit der DNA-Bank H<sub>2</sub>NhpIIIc6V16 und zum anderen mit der DNA-Bank SOLiB30, welche keinen 5'-Aminolinker besaß, gearbeitet. Für diese Versuchsreihe wurde ohne vorige Ligations-, Spaltungs- und Sortierungsreaktion gearbeitet. Die Synthese erfolgte direkt an der DNA-Bank. Als Kopplungsreagenz wurde DMT-MM (10 mM) und EDC (50 mM) in DMF verwendet. Die DNA-Banken wurden für die Synthesereaktion an SiMAG-DEAE Mikropartikel gebunden und anschließend in 100 µl Kopplungsreagenz (DMT-MM in DMF und EDC in DMF), welche eine Endkonzentration von 400 µM d-Biotin besaß, inkubiert. Die Reaktion wurde bei RT (22°) für 30 min vorgenommen. Nach der Elution der DNA-Banken von den Mikropartikeln mit 20 mM Tris-HCl, 1 M NaCl wurden diese aufgereinigt und eine Selektion gegen das Zielmolekül Streptavidin konnte vorgenommen werden. Da das Streptavidin an Mikropartikeln geankert war, konnte auch hier direkt nach der Selektion ein Waschschriff zur Beseitigung von nicht-gebundenen Fragmenten eingefügt werden. Mit einer PCR-Reaktion auf die Streptavidin-Partikeln mit den DNA-Banken konnte ermittelt werden, ob bei beiden DNA-Banken eine Synthese stattgefunden hat oder nur bei der mit dem 5'-Aminolinker. Zudem wurde analysiert, ob DMT-MM oder EDC das bessere Aktivierungsreagenz ist.

Die Banken wurden sowohl mit dem Kopplungsreagenz DMT-MM in DMF (Abb. 34, a) als auch EDC in DMF (Abb. 34, b) inkubiert. Für beide Reaktionsbedingungen gab es zum einen eine d-Biotinsynthese für beide Banken (Spur 1 & 2 / 5 & 6) aber als Kontrollansätze keine Biotinsynthese, um den Hintergrund der Banken bestimmen zu können (Spur 3 & 4 / 7 & 8). Im Idealfall sollte nur bei der Bank eine Biotinsynthese stattgefunden haben, die den 5'-Aminolinker trägt und in der anschließenden Selektion gegen Streptavidin nur die Fragmente angereichert sein, bei denen auch d-Biotin in der Reaktion zugegeben wurde. Man kann jeweils in der Abbildung 34 erkennen, dass eine Anreicherung bei der DNA-Bank H<sub>2</sub>NhpIIIc6V16 mit Biotinsynthese (Spur 2 & 6) im Vergleich zu der Bank SOLiB30 ohne Aminolinker (Spur 1 & 5) zu erkennen ist.



**Abb. 34: PCR-Produkte der Synthese mit d-Biotin und Selektion gegen Streptavidin auf einem Ethidiumgefärbten Agarosegel (3%).** O'RangeRuler 20 bp DNA Ladder (Spur M). Synthesereaktion erfolgte mit dem Kopplungsreagenz DMT-MM (10 mM) in DMF (a) und EDC (50 mM) in DMF (b). Synthese mit d-Biotin und anschließende Selektion gegen Streptavidin der Kontrollbank SOLiB30 (Spur 1 & 5) und H<sub>2</sub>NhpIIIc6V16 (Spur 2 & 6). Kontrollreaktionen ohne Biotinsynthese der Bank SOLiB30 (Spur 3 & 7) und H<sub>2</sub>NhpIIIc6V16 (Spur 4 & 8).

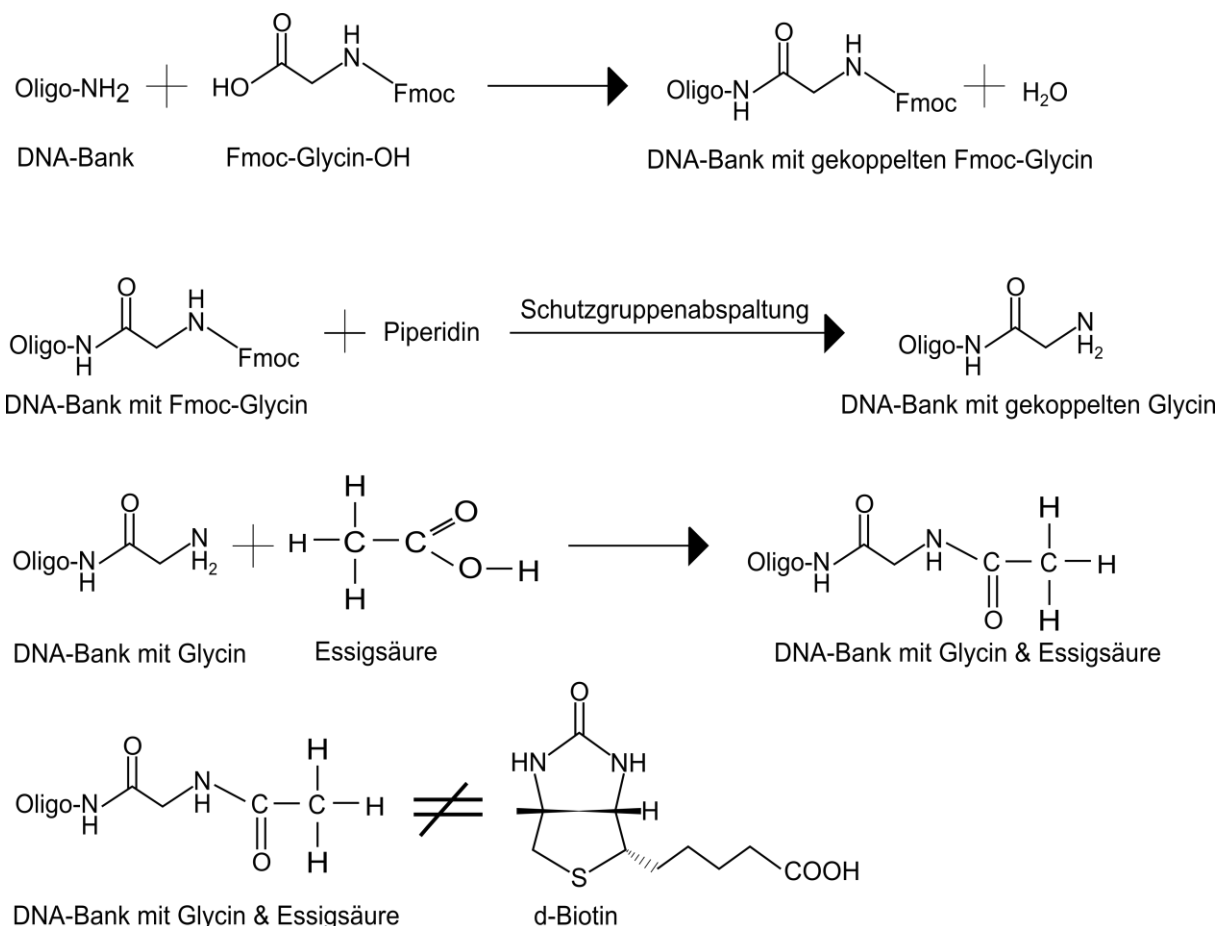
Die Kontrollen bei denen keine Biotinsynthese stattgefunden hatte, zeigten auch kaum PCR-Produkte. Vor allem die 5'-Aminolinker DNA-Bank zeigt minimalen Hintergrund (Spur 4 & 8). Die beste Syntheseausbeute ist in Spur 2 zu verzeichnen. Dabei handelt es sich um die d-Biotinsynthese an der DNA-Bank H<sub>2</sub>NhpIIIc6V16 mit DMT-MM in DMF.

### 3.5.2 ZWEI SYNTHESERUNDEN MIT GLYCIN & D-BIOTIN UND SELEKTION GEGEN DAS ZIELMOLEKÜL STREPTAVIDIN

In dieser Versuchsreihe sollte erstmals gezeigt werden, dass zwei Syntheserunden spezifisch erfolgen. Dabei sollte in dem ersten Zyklus die Aminosäure Glycin und im Zweiten das Vitamin d-Biotin an die DNA-Bank synthetisiert werden. Auf diesem Wege sollte gezeigt werden, dass eine DNA-kodierte Substanz-Bank erzeugt werden kann.

Zu diesem Zweck wurde eine Ligation mit allen 16 verschiedenen Adaptoren hergestellt. Anschließend erfolgte die Separation an die immobilisierten Hybridisierungsoligonukleotide, die Spaltungsreaktion (siehe 2.2.4 – 2.2.6) und Kopplung auf SiMAG-DEAE Mikropartikeln. Die Synthese wurde mit einigen ausgewählten Proben in DMF mit einer Endkonzentration von 10 mM DMT-MM für 30 min bei RT (22°C) vorgenommen. Dem Ansatz wurde die Aminosäure Fmoc-Glycin-OH mit einer Endkonzentration von 400 µM beigefügt. Im Anschluss erfolgte nach der ersten Syntheserunde die Schutzgruppenabspaltung (siehe 2.2.8)

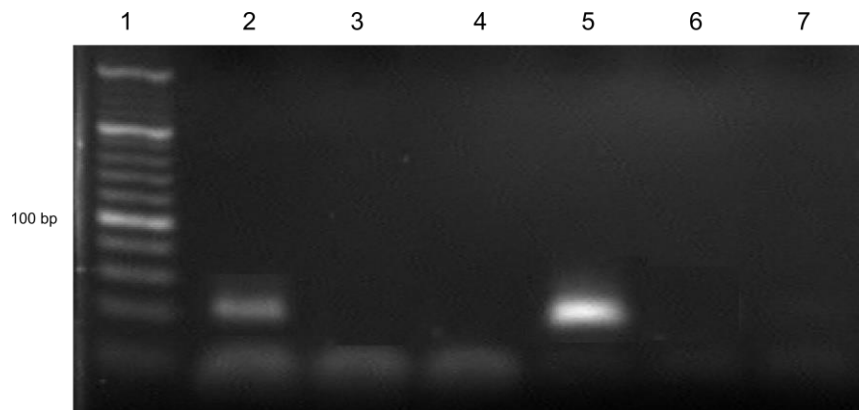
und der Zyklus konnte von vorne beginnen. In der zweiten Runde wurde d-Biotin an die DNA-Banken gekoppelt und mittels einer Selektion gegen das Zielmolekül Streptavidin sollte gezeigt werden, dass eine Synthese mit Glycin und d-Biotin stattgefunden hatte. Dies sollte durch eine Anreicherung der erfolgreichen Synthesen gegenüber Kontrollen nachgewiesen werden. Dabei können nur die DNA-Banken an das Streptavidin-Molekül binden, bei denen auch vorher d-Biotin gekoppelt wurde. Es muss aber auch nachgewiesen werden, dass in dem vorangegangenen Zyklus Glycin an die Bank synthetisiert wurde und nicht nur im zweiten Zyklus d-Biotin. Durch eine Kontrollreaktion sollte dies ermittelt werden. In diesen Kontrollreaktionen wurde nach der Glycinsynthese eine Essigsäurebehandlung vorgenommen. Die Proben wurden zu diesem Zweck mit einer Endkonzentration der Essigsäure von 400 mM in 100  $\mu$ l DMF für 30 min bei RT (22°C) inkubiert. Anschließend wurden die Proben gewaschen und es erfolgte wie bei den anderen Ansätzen die Elution von den DEAE-Mikropartikeln.



**Abb. 35: Vereinfachte Darstellung der zwei Syntheserunden mit Essigsäurebehandlung.** Nachdem die Synthese mit Fmoc-Glycin-OH stattgefunden hatte, erfolgte eine Behandlung mit Essigsäure. Dabei konnte die Essigsäure an das entschützte Glycin koppeln und blockierte somit die freie Aminogruppe für die Biotinsynthese. Eine Synthese des d-Biotins an die DNA-Bank und eine Selektion gegen das Zielmolekül Streptavidin war somit nicht möglich.

Wie in der Abbildung dargestellt, reagiert die entschützte Aminogruppe des Glycins mit der Carboxyl-Gruppe der Essigsäure. Im Anschluss steht keine freie Aminogruppe für die Synthese mit d-Biotin zur Verfügung, somit sollte diese DNA-Bank auch nicht an Streptavidin koppeln können. Auf diesem Wege kann die die Glycinsynthese nachgewiesen werden.

Nach der Selektion wurde für eine Detektion der Ergebnisse wieder eine PCR-Reaktion mit spezifischen Primern durchgeführt und durch die Agarosegelelektrophorese sichtbar gemacht.



**Abb. 36: PCR-Produkte der Synthesereaktionen mit Fmoc-Glycin-OH und d-Biotin mit anschließender Selektion gegen Streptavidin auf einem Ethidium-gefärbten Agarosegel (3%).** O'RangeRuler 20 bp DNA Ladder (Spur 1). Zwei Synthesedurchläufe mit Glycin und d-Biotin mit anschließender Selektion und Amplifikation mit einem spezifischen Primer (Spur 2) und unspezifischen Primern (Spur 5). Im Kontrollansatz (Spur 3) passt der Primer nicht. PCR-Produkte der Synthese und anschließenden Selektion nach der Essigsäurebehandlung mit passenden Primer (Spur 4) und nicht-passenden Primer (Spur 6). In der letzten Kontrolle (Spur 7) hat keine Synthese stattgefunden.

Im Gelbild sollte bei einer positiven Anreicherung eine Bande zwischen 50-60 bp zu erkennen sein. Dies ist in den Ansätzen (Spur 2 & 5) deutlich zu erkennen. Der Unterschied zwischen diesen beiden Proben liegt in der Verwendung der Primer. In der ersten Probe (Spur 2) wurde ein spezifischer Primer verwendet, der auf Grund der Sortierung ermittelt werden konnte. In der zweiten Probe (Spur 5) wurden die Primer für die konstante Region der DNA-Bank verwendet. In beiden Proben ist ein deutliches Signal zu erkennen, in den Kontrollen hingegen nicht und somit ist die Synthese in zwei Runden gelungen.

Die Kontrollen zeigten kein Signal. Nach der Essigsäurebehandlung (Spur 4 & 6) blieb das Signal aus und somit konnte auf diesem Wege die Kopplung von Glycin nachgewiesen werden. Auch die anderen Kontrollen zeigten kein Signal. Da in Spur 3 keine Bande zu verzeichnen war, konnte die Spezifität der Sortierung nochmals aufgezeigt und durch die Probe in Spur 7 die Bildung von Artefakten ausgeschlossen werden, da dort keine Synthesereaktion stattgefunden hatte.

### 3.5.3 SYNTHESE MIT GLYCIN & D-BIOTIN UNTER 16 VERSCHIEDENEN SYNTHESEBEDINGUNGEN UND SELEKTION GEGEN STREPTAVIDIN

In den vorangegangenen Versuchen konnten zwei Synthesedurchgänge mit Glycin und d-Biotin gezeigt werden. Um genaue Aussagen über die Spezifität und Effizienz der Synthese machen zu können, erfolgten die Reaktionen mit demselben Kopplungsreagenz und mit spezifischen Adaptern und Primern.

Im Folgenden sollen zwei Synthesedurchgänge mit allen 16 Adaptern unter 16 verschiedenen Reaktionsbedingungen erfolgen. Nach nur zwei Synthesedurchgängen würde eine DNA-kodierte Substanz-Bank mit einer Diversität von 256 verschiedenen Molekülen entstehen. Durch die Verwendung von 16 verschiedenen Kopplungsreagenzien ergaben sich nach zwei Runden auch 256 verschiedene Synthesebedingungen, die durch den Barcode der DNA-kodierten Bank ermittelt werden können.

Die Syntheserunden erfolgten halbautomatisiert mit dem Magnetpartikelprozessor KingFisher Flex von Thermo Scientific. In der folgenden Tabelle sind alle Arbeitsschritte aufgelistet, die zum einen manuell und zum anderen mit dem Roboter vorgenommen wurden. Die manuellen Arbeitsschritte fanden jeweils in einem Reaktionsgefäß statt und somit konnten diese auch ohne den Roboter durchgeführt werden. Der Roboter wurde vor allem für parallele Schritte die explizit eingehalten werden mussten angewendet. So kann eine genaue Reproduzierbarkeit der Arbeitsschritte garantiert werden (Schütze *et al.*, 2011). Die Protokolle sind im Material und Methoden-Abschnitt als auch im Anhang nachzulesen.

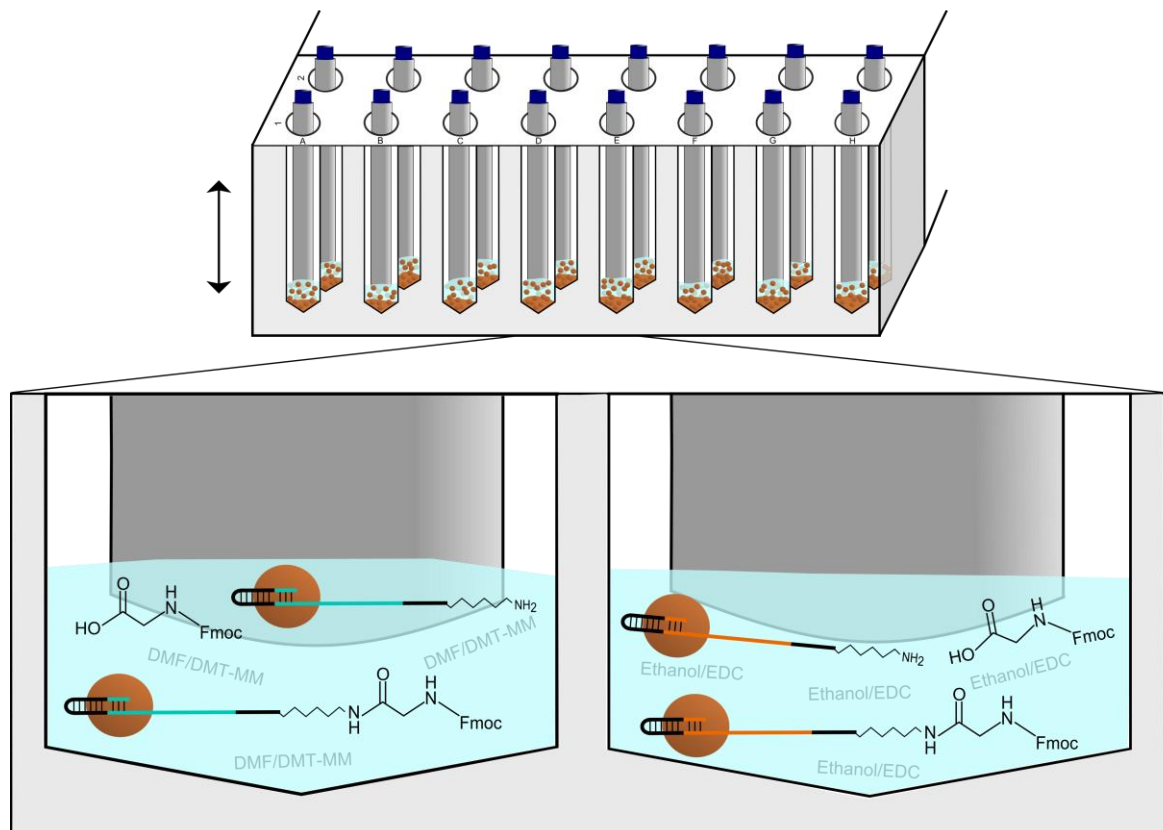
**Tab. 22: Arbeitsablauf der Neclys-Technologie.** Sowohl manuelle als auch automatisierte Protokolle wurden für die Synthese der DNA-kodierten Substanz-Bank angewendet.

Reihenfolge der Arbeitsschritte	Methode
Ligationsreaktion	Manuell
Separation an immobilisierte Oligonukleotide	Roboter
Spaltungsreaktion mittels Endo IV	Manuell
Kopplung der Bank an DEAE-Partikel	Manuell
Synthesereaktion	Roboter
Schutzgruppenabspaltung	Roboter
Elution von den DEAE-Partikeln	Roboter
Poolen und aufreinigen	Manuell
Zweite Syntheserunde	



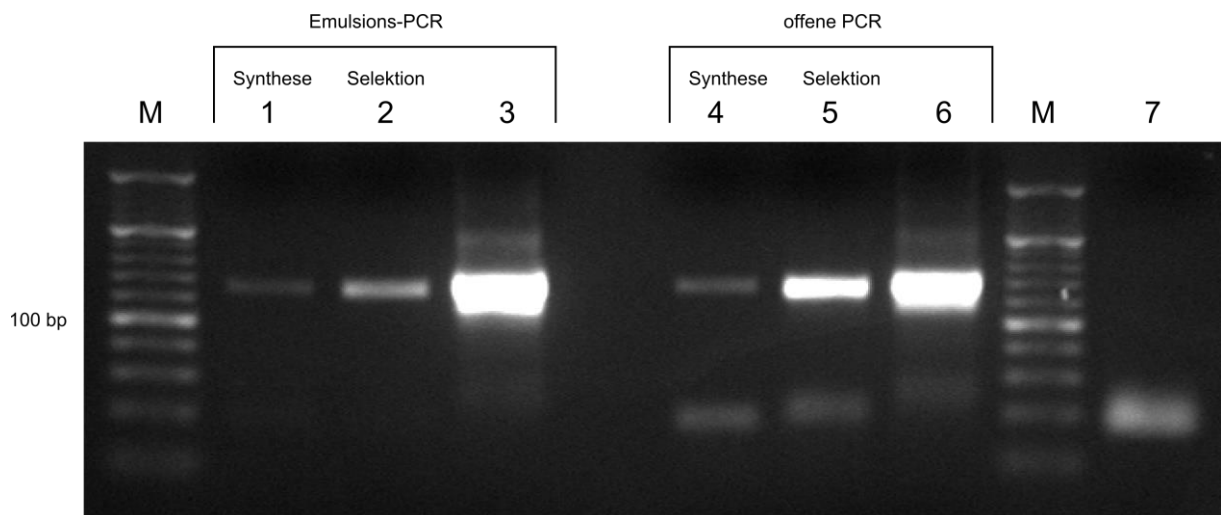
Nachdem die Ligations-, Separations- und Spaltungsreaktion als auch die Bindung der separierten DNA-Bank an die DEAE-Partikeln erfolgt sind, konnte die Synthese von Fmoc-Glycin-OH unter 16 verschiedenen Synthesebedingungen (siehe Tab. 14) erfolgen.

In der folgenden Abbildung soll am Beispiel der Fmoc-Glycin-OH Synthese dargestellt werden, wie die Synthese auf dem Kingfisher Flex Magnetpartikelprozessor abläuft. Auch die Biotinsynthese in der zweiten Runde erfolgte unter den gleichen Bedingungen.



**Abb. 37: Synthese von Fmoc-Glycin-OH an die separierte DNA-Bank mit dem Kingfisher Flex.** Nachdem die separierte, gesplante DNA-Bank an die DEAE-Mikropartikel gebunden war, konnte die Synthese erfolgen. Zu diesem Zweck wurde die DNA-Bank in eine Mikrotiterplatte geladen und mit dem jeweiligen Kopplungsreagenz inkubiert. Die Synthese erfolgte in 16 verschiedenen Reaktionsgefäßen unter 16 verschiedenen Synthesebedingungen. Vergrößert dargestellt wurde die Synthese von Glycin mit dem Kopplungsreagenz DMT-MM in DMF (links) und EDC im Lösungsmittel Ethanol.

Nach den zwei Syntheserunden wurden die Proben vereint und für weitere Analysezwecke aufgeteilt. Alle PCR-Reaktionen erfolgten zum einen unter dem Standard-Programm (offen) und zum anderen als Emulsions-PCR. Zunächst wurde eine PCR-Reaktion nach der Synthese durchgeführt. Eine weitere PCR-Reaktion nach der Selektion auf das Zielmolekül Streptavidin. Durch den Vergleich der PCR-Produkte der Synthese und der Selektion sollte eine Anreicherung aufgezeigt werden, wenn die Synthesedurchgänge effizient gewesen sind. In der Abbildung 38 sind die PCR-Produkte in einem 3%-igen Agarosegel dargestellt.



**Abb. 38: PCR-Produkte nach zwei Synthesedurchgängen auf einem Ethidium-gefärbten Agarosegel (3%).** O'RangeRuler 20 bp DNA Ladder (Spur M). Die PCR wurde einmal nach dem Standard-Programm und zum anderen nach dem Emulsions-Protokoll durchgeführt. In beiden Methoden ist nach der Selektion (Spur 2 & 5) eine Anreicherung gegenüber den PCR-Produkten nach der Synthese (Spur 1 & 4) zu erkennen. Ausgangskonzentration der initialen DNA-Bank (Spur 3 & 6). Negativkontrolle der PCR-Reaktion (Spur 7).

In dem Gelbild sind die PCR-Produkte nach der Synthese und der Selektion einmal unter dem Standard PCR-Programm und zum anderen nach einer Emulsions-PCR aufgezeigt. Durch die verwendeten Primer wurden Banden bei einer Größe von circa 130 bp erwartet. Als Vergleichsgröße wurden auch die PCR-Produkte der DNA-Bank ohne Synthesedurchgänge aufgetragen (Spur 3 & 6). Zunächst konnte gezeigt werden, dass die Emulsions-PCR weniger Artefakte gegenüber der Standard PCR-Methode erzeugt. Unter beiden Methoden ist eine klare Bande nach der Synthese (Spur 1 & 4) zu verzeichnen, die nach der Selektion angereichert werden konnte (Spur 2 & 5).

Somit konnte eindeutig eine Anreicherung der DNA-kodierten Substanz-Bank nach der Selektion gegen Streptavidin gezeigt werden. Welche Synthesebedingungen dabei die effizientesten bzw. welche Synthesebedingungen angereichert vorliegen, kann erst nach einer Hochdurchsatz-Sequenzierung ermittelt werden.

### 3.6 ANWENDUNG VON HOCHDURCHSATZSEQUENZIERUNG FÜR DIE ANALYSE VON DNA-KODIERTEN SUBSTANZ – BANKEN

Die neue Generation der Sequenzierung ist eine leistungsfähige Hochdurchsatz-Technologie, die seit dem Jahre 2004 verfügbar ist. In der Praxis zeigt sich dies durch die Kopplung von Millionen DNA-Fragmenten an einer Festphase und gleichzeitiger Sequenzierung dieser in parallelen Schritten. Auf diesem Wege wird der hohe Sequenzierungsdurchsatz erreicht (Mori *et al.*, 2013). Der Arbeitsablauf einer Standard DNA-Sequenzierung gliedert sich traditionell in drei Schritte auf: Proben-Präparation, Sequenzierung und Datenanalyse (Morozova *et al.*, 2009). Der größte Vorteil der neuen Hochdurchsatztechnologien ist die Fähigkeit, eine enorme Menge an Datensätzen bei relativ geringen Kosten zu erzeugen (Metzker, 2010). In dieser Arbeit wurde das Verfahren von 454 / Roche für die Analyse der DNA-kodierten Substanz-Bank verwendet. Mit dieser Methode kann ein hoher Datensatz erzielt werden, wodurch eine Anreicherung bei der Selektion an das Zielmolekül und somit eine bevorzugte Synthesebedingung ermittelt werden konnte.

Dazu wurden wieder zwei Synthesedurchgänge wie unter 3.5.3 beschrieben, durchgeführt. Nachdem die Selektion gegen das Zielmolekül erfolgt war, wurden auf die initiale DNA-Bank, die Synthesefragmente und die selektierten DNA-Fragmente, eine offene PCR als auch eine Emulsions-PCR mit den Barcodes versehenen Primern durchgeführt (siehe Tab. 9). So können die erhaltenen Daten der jeweiligen Reaktionsbedingung zugeordnet werden. Die Proben wurden nach der Polymerasekettenreaktion an das Sequenzierungsteam des MPI übergeben, wo die eigentliche Sequenzierung durchgeführt wurde. Im Anschluss wurden die Rohdaten durch einen Bioinformatiker des MPI nach ihren Barcodes sortiert. Mit diesen vorsortierten Daten konnte anschließend eine Feinanalyse durchgeführt werden. Zu diesem Zweck wurde mit dem Programm X-Win 32 gearbeitet. Dadurch konnten die Sequenzen nach ihren Barcodes und deren Häufigkeit analysiert werden. Die Barcodes erlauben wieder einen Rückschluss auf die Synthesebedingung und durch deren vorkommende Häufigkeit konnten Anreicherungen und somit optimale Synthesebedingungen ermittelt werden. Es können aber auch Rückschlüsse auf Schädigungen an der DNA-Bank ermittelt werden. Die Häufigkeitsverteilungen der Barcodes für die initiale DNA-Bank, für die Synthese als auch Selektion, für die offene und geschlossene PCR sind im Anhang nachzulesen.

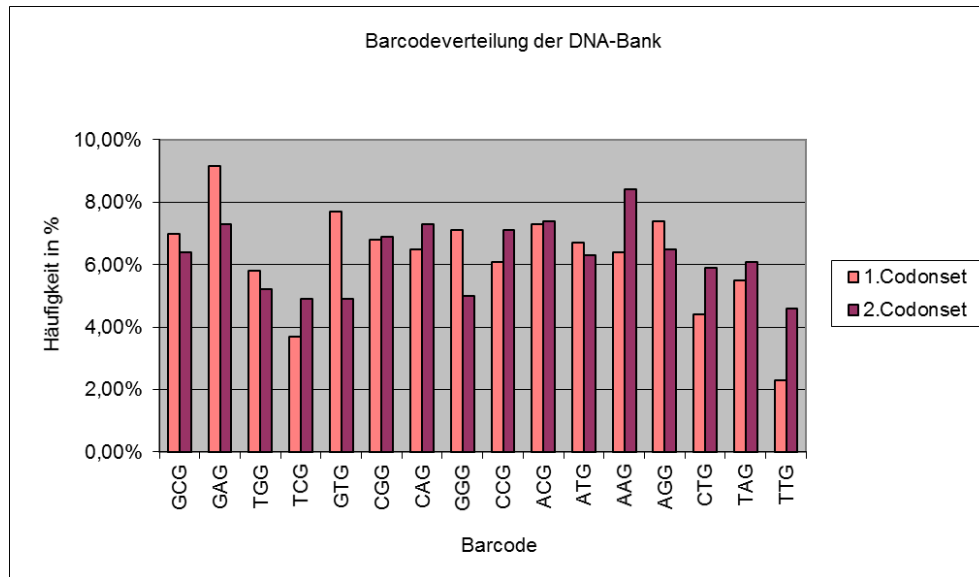
### 3.6.1 DAS 454-PYROSEQUENZIERUNGSSYSTEM VON ROCHE APPLIED SCIENCES

Die DNA-kodierte Substanz-Bank wurde durch das Serviceteam des MPI für molekulare Genetik mittels 454-Pyrosequenzierung sequenziert. Die erhaltenen Rohdaten wurden so bearbeitet, dass im Anschluss die Amplicons mit der richtigen Länge von 18 bp zur Auswertung vorlagen. Diese Amplicons wurde mithilfe des X-Win 32 Programmes nach ihren Barcodes sortiert. Dass heißt, das erste Codonset, das Zweite und zwei Codonsets zusammen wurden im Vorkommen ihrer Häufigkeit sortiert und anschließend die Häufigkeit gegeneinander mittels graphischer Vorstellung präsentiert und analysiert.

Dabei wurden immer die Daten der DNA-Bank gegen die Synthese-Daten und die Synthese-Daten gegen die Selektions-Daten dargestellt und ausgewertet. Insgesamt ergaben sich 70708 Sequenzen durch die Pyrosequenzierung von denen 52808 zur Auswertung herangezogen werden konnten, da diese die 18 bp langen Amplicons enthielten. Die anderen Sequenzen waren entweder zu lang oder aber zu kurz (siehe Anhang), da sie beispielsweise noch die Primersequenzen enthielten. Zudem wurden zur Auswertung nur die Codons verwendet, die auch aus unseren bekannten Basentriplett NNG bestanden. Abweichungen, die in der dritten Base vorhanden waren, wurden nicht in die Auswertung mit einbezogen, waren aber auch äußerst gering.

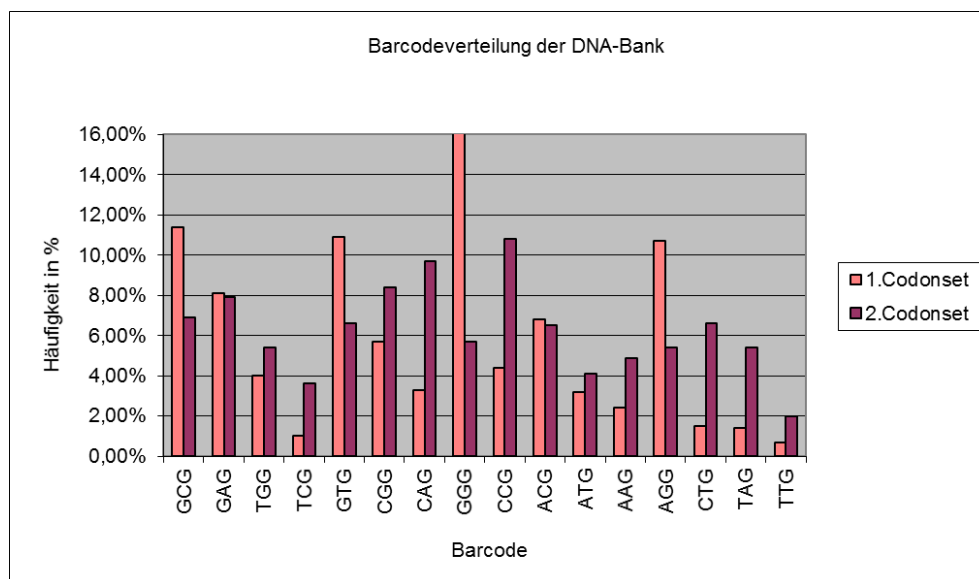
Zunächst wurde aus den vorsortierten Rohdaten die Häufigkeit (%) der 16 verschiedenen Barcodons der DNA-Bank, der Synthese und der Selektion für das erste und zweite Codonset, sowie zwei Codonsets tabellarisch dargestellt, bevor die graphischen Auswertungen für genaue Analysen hinzugezogen wurden. Die Rohdaten sind im Anhang nachzulesen.

Die Häufigkeitsverteilung der verschiedenen Barcodons der initialen DNA-Bank sollte gleichmäßig sein, da die DNA-Fragmente keine chemische Behandlung erfahren haben und somit die Diversität gewahrt sein sollte. Dargestellt wurde exemplarisch die Codonverteilung des ersten und zweiten Sets für die DNA-Bank, einmal offen und einmal in Emulsion amplifiziert.



**Abb. 39: Graphische Darstellung der Häufigkeitsverteilung (%) des ersten und zweiten Barcodons der DNA-Bank.** Die Amplifizierung erfolgte in einer offenen PCR-Reaktion, anschließend wurden die Fragmente sequenziert und die Häufigkeit der Codonverteilung ermittelt.

In der Abbildung kann man deutlich sehen, dass es kaum Schwankungen in der Häufigkeitsverteilung der Barcodons gibt. Die 16 verschiedenen Codons kommen sowohl im ersten Triplet als auch im Zweiten gleichmäßig vor und zeigen zudem eine konstante Präsenz. Somit scheint die Diversität der DNA-Bank gewährleistet zu sein. Abweichungen von diesem Bild nach der Synthese und Selektion zeigen Anreicherungen durch optimale Synthesebedingungen bzw. Verluste und somit schlechte Synthesebedingungen wieder.

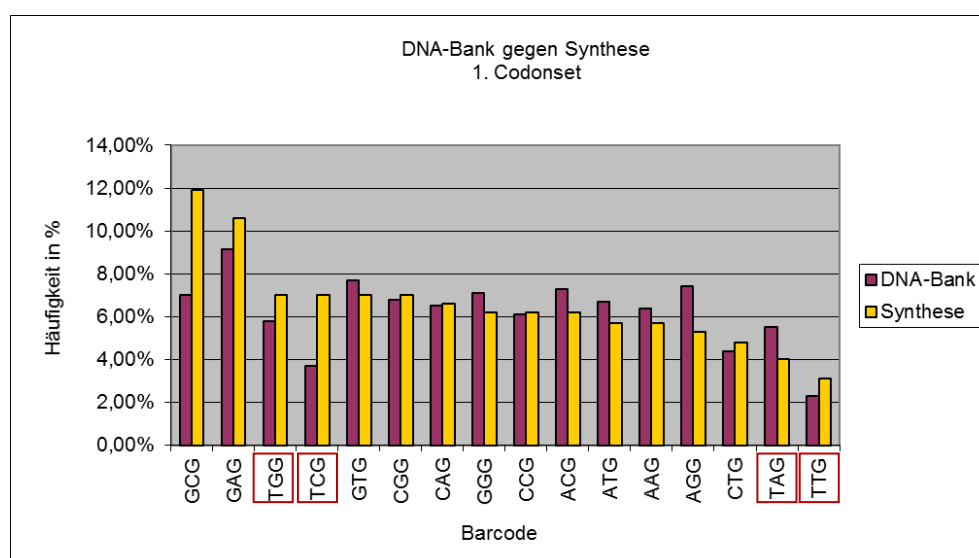


**Abb. 39: Graphische Darstellung der Häufigkeitsverteilung (%) des ersten und zweiten Barcodons der DNA-Bank.** Die Amplifizierung erfolgte in einer Emulsions-PCR.

Die Häufigkeitsverteilung der DNA-Bank aus der Emulsions-PCR hingegen zeigt stärkere Schwankungen innerhalb des ersten Codonsets als auch im Zweiten, sowie untereinander auf. Die Häufigkeitsverteilung schwankt zwischen der offenen und der Emulsions-PCR, die Diversität innerhalb der unbehandelten DNA-Bank konnte aber dargestellt werden.

Nach der Synthesereaktion sollte gegenüber der ursprünglichen DNA-Bank eine leichte Anreicherung für Barcodons vorhanden sein, bei denen die Reaktion nur in Lösungsmittel aber ohne Kopplungsreagenz erfolgte (siehe Tab. 14). Diese Anreicherungen sollten entstehen, da durch das Fehlen der Kopplungsreagenzien die DNA-Bank weniger geschädigt wird.

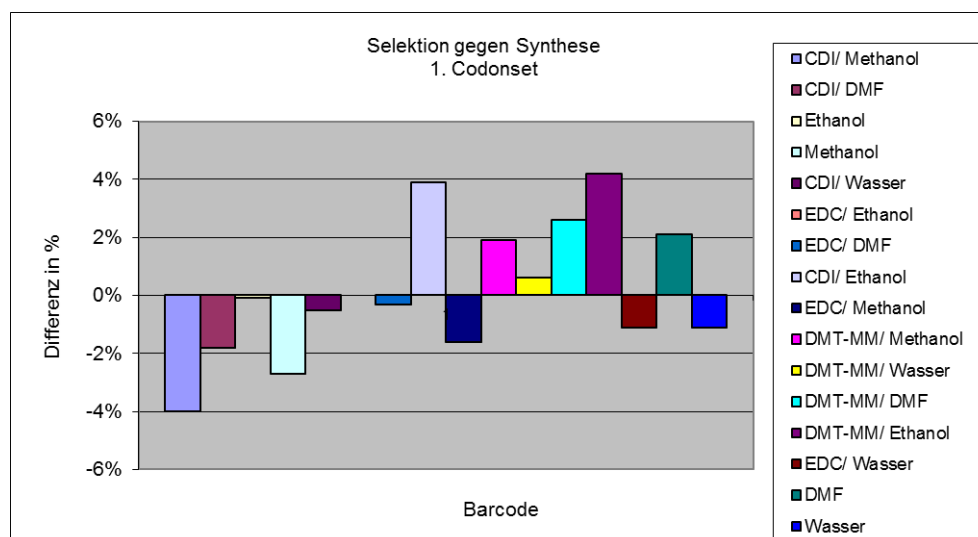
Zu diesem Zweck, wurden die Barcodons aus der Synthese denen der DNA-Bank gegenübergestellt. Folgende Barcodons entsprechen Synthesereaktionen ohne Kopplungsreagenzien: TAG, TCG, TGG und TTG. Alle vier Codons sind nach der Synthese sowohl für das erste Codonset als auch für das zweite Codonset entweder angereichert oder in ihrer Häufigkeitsverteilung ähnlich zur DNA-Bank (siehe Abb. 40). Diese Daten sind sowohl für die offene PCR als auch für die Emulsions-PCR zu finden (siehe Anhang Pyrosequenzierungsdaten). Somit konnte gezeigt werden, dass die Barcodons ohne Kopplungsreagenz nach der Synthese leicht angereichert waren und keine Schädigung der DNA stattgefunden hat. Exemplarisch ist eine Häufigkeitsverteilung zwischen der Synthese und der DNA-Bank aus dem ersten Codonset der offenen PCR aufgezeigt.



**Abb. 40: Graphische Darstellung der Häufigkeitsverteilung (%) des ersten Barcodesets der DNA-Bank und der Synthese.** Barcodes bei denen die Synthese ohne Kopplungsreagenz und nur in Lösungsmittel erfolgte, sind rot umrandet.

Nach der Selektion sollte eine Anreicherung der Barcodes vorliegen, bei denen die Synthese mit Kopplungsreagenzien erfolgte. Bei diesen sollte die Kopplung von d-Biotin und Glycin im Vergleich zu den Barcodes, die der Synthese ohne Kopplungsreagenzien entsprechen, besser funktioniert haben. Barcodes die der Synthese ohne Kopplungsreagenz entsprechen sollten im Vergleich in geringeren Mengen auftreten. Dieser Vorteil sollte im Anschluss durch die Selektion gegen Streptavidin erkennbar sein, da hier die besseren Kopplungsbedingungen vorliegen.

Im Idealfall sind die gleichen Barcodons nach der ersten wie nach der zweiten Synthese angereichert. Zudem sollten die vorher aufgezeigten Barcodons ohne Kopplungsreagenz nach der Selektion in einer geringeren Häufigkeit auftreten als es nach der Synthese der Fall war. Bei der Auswertung wurden die Barcodons nach der Synthese mit denen nach der Selektion in ihrer Häufigkeit verglichen. Konnte eine Anreicherung festgestellt werden, wurde zudem darauf geachtet, dass dieses Barcodons vorher im Vergleich zwischen der Synthesereaktion und der DNA-Bank keine Schädigung aufwies. Es ist exemplarisch für das erste Codonset der Vergleich aus der offenen PCR dargestellt. Um eine bessere Übersicht zu erhalten, wurde die Differenz aus der Häufigkeitsverteilung zwischen der Synthese und nach der Selektion für das jeweilige Codon gezogen und graphisch dargestellt.



**Abb. 41: Graphische Darstellung der Häufigkeitsdifferenz (%) der Synthese gegen die Selektion für das erste Codonset.** Um eine Anreicherung gegen Streptavidin festzustellen, wurde die Häufigkeit der Barcodons für die Synthese- und die Selektionsreaktion gegeneinander gestellt. Die den Barcodons entsprechenden Synthesebedingungen sind in der Legende dargestellt. Eine Häufigkeitsdifferenz im positiven Skalabereich bedeutet eine Anreicherung gegen das Zielmolekül Streptavidin und somit optimale Synthesebedingungen.

Man kann deutlich erkennen, dass eine Anreicherung für die Barcodons, bei denen in der Synthese Kopplungsreagenzien enthalten waren, stattgefunden hat. Eine gute Anreicherung ist bei den Barcodes bzw. bei den Synthesebedingungen, die in Tabelle 23 aufgezeigt sind zu verzeichnen. Diese Dateninformation ist aus der gesamten Auswertung entstanden. Im Gegensatz dazu zeigen die Barcodes ohne Kopplungsreagenz eine verminderte Häufigkeit. Es konnte aufgezeigt werden, dass die Synthese von d-Biotin und Glycin durch die Zugabe von Kopplungsreagenzien effizienter als die Synthese in reinem Lösungsmittel ist.

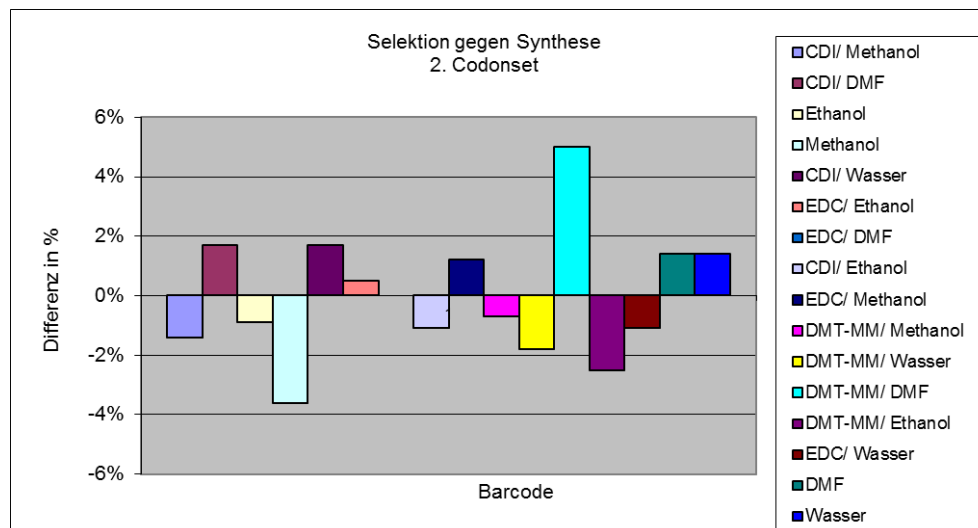
**Tab. 23: Angereicherte Barcodes und deren entsprechenden Kopplungs- und Lösungsmittel.** Durch die dekodierten Barcodes kann auf die entsprechenden Synthesebedingungen rückgeschlossen werden. Dargestellt sind die angereicherten Barcodes von allen erhaltenen Daten (siehe Anhang).

Barcode	Kopplungsreagenz	Lösungsmittel
GGG	CDI	Ethanol
ACG	DMT-MM	Methanol
AAG	DMT-MM	DMF
AGG	DMT-MM	Ethanol
GAG	CDI	DMF
CAG	EDC	DMF

Nach genauen Analysen konnte festgestellt werden, dass vor allem Barcodes angereichert waren, die der Synthesebedingung mit dem Kopplungsreagenz DMT-MM und dem Lösungsmittel DMF entsprachen.

Ideal wäre es, wenn die angereicherten Barcodes aus dem ersten Set auch im zweiten Codonset vorkämen, dann könnte man eine Präferenz für ein Kopplungsreagenz und Lösungsmittel identifizieren. Exemplarisch ist wieder die Auswertung des zweiten Barcodesets zwischen der Synthesereaktion und nach der Selektion aus der offenen PCR dargestellt.





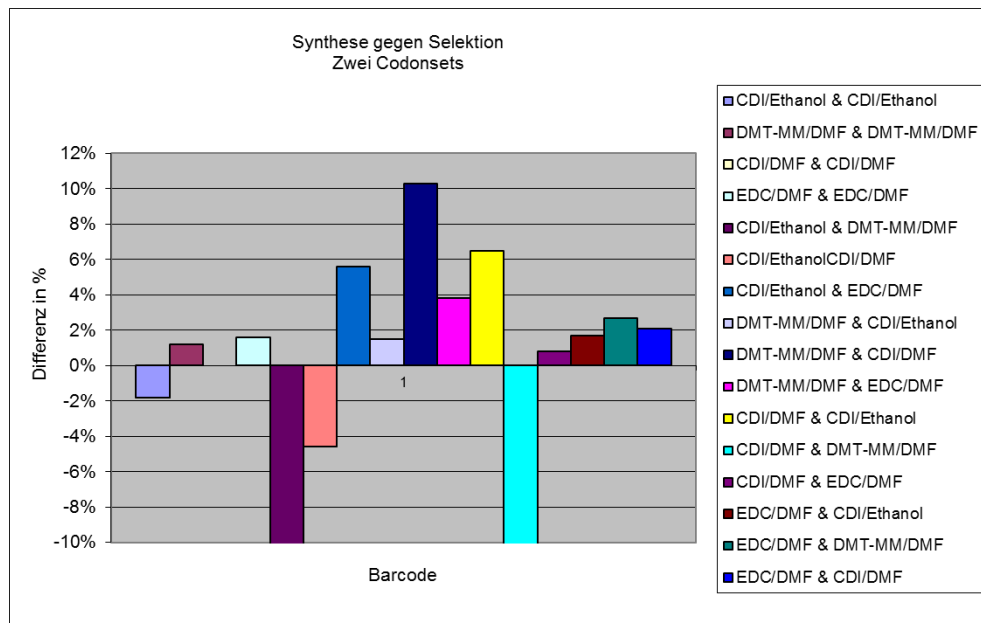
**Abb. 42: Graphische Darstellung der Häufigkeitsdifferenz (%) der Synthese gegen die Selektion für das zweite Codonset.** Die Häufigkeiten der Barcodons für die Synthese- und die Selektionsreaktion wurden gegen einander gestellt, um eine Anreicherung fest zu stellen. Die den Barcodons entsprechenden Synthesebedingungen sind in der Legende dargestellt.

Nach der Auswertung des zweiten Codonsets ist deutlich eine Anreicherung des Barcodes AAG und somit der Synthesebedingung mit dem Kopplungsreagenz DMT-MM und dem Lösungsmittel DMF zu verzeichnen.

Somit konnte zunächst festgestellt werden, dass die Barcodes bei denen in der Synthesereaktion ein Kopplungsreagenz zugegeben wurde, deutlich gegenüber denen ohne Zugabe eines Kopplungsreagenz angereichert waren. Das weist auf eine effektivere Synthesereaktion hin. Besonders die Kombination DMT-MM in DMF zeigte sich als optimales Kopplungsreagenz.

Im Anschluss wurden die Daten hinsichtlich zweier Barcodesets sortiert. Da zwei Synthesen durchgeführt wurden kann anhand der zwei hintereinander folgenden Barcodons analysiert werden, welche Reagenzkombination am besten arbeitet und zu einem guten Syntheseergebnis führt.

Zunächst wurden die zwei Barcodekombinationen selbst zusammengestellt und nach ihrer Häufigkeit sortiert. Dabei wurden die Barcodons gegeneinander dargestellt, die im ersten und zweiten Set am häufigsten vorkamen. Graphisch dargestellt wurden die Daten aus der Emulsions-PCR.



**Abb. 43: Graphische Darstellung der Häufigkeitsdifferenz (%) der Synthese gegen die Selektion für zwei Codonsets, die von Interesse waren.** Die Häufigkeit der Barcodons für die Synthese- und die Selektionsreaktion wurden gegen einander gestellt, um eine Anreicherung und somit bevorzugte Synthesebedingung fest zu stellen. Die den Barcodons entsprechenden Synthesebedingungen sind in der Legende dargestellt.

Man kann sowohl Anreicherungen als auch Verluste deutlich erkennen. Die Kombinationen wurden selber zusammengestellt, da die Barcodons einzeln angereichert vorkamen. In Tabelle 24 sind die angereicherten Codonsets und die dazugehörigen Synthesebedingungen aufgelistet.

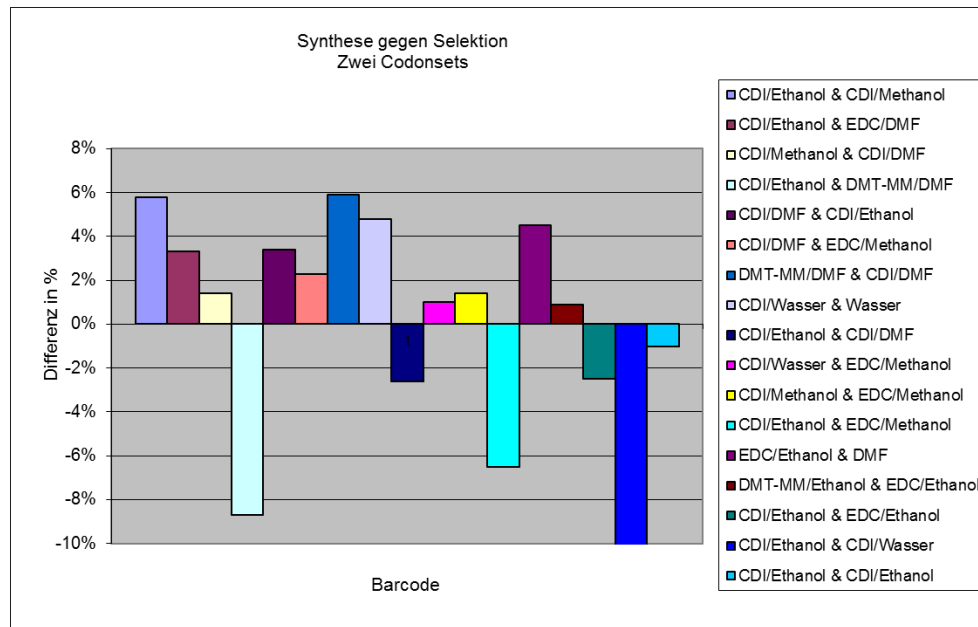
**Tab. 24: Angereicherte Codonsets und deren entsprechenden Kopplungs- und Lösungsmittel nach zwei Syntheserunden.** Durch die dekodierten Barcodes kann auf die entsprechenden Synthesebedingungen rückgeschlossen werden.

Barcode	Kopplungsreagenz und Lösungsmittel
AAGGAG	DMT-MM/DMF & CDI/DMF
AAGCAG	DMT-MM/DMF & EDC/DMF
GAGGGG	CDI/DMF & CDI/Ethanol
GGGCAG	CDI/Ethanol & EDC/DMF

Festgestellt werden muss, dass es keine Anreicherungskombination gibt die jeweils die gleichen Reagenzien enthält. Es sind lediglich immer Kombinationen von verschiedenen Kopplungsreagenzien angereichert. Insgesamt können vier verschiedene Barcodes identifiziert werden, die in unterschiedlichen Kombinationen auftreten (siehe Tabelle 24). Zu diesen Barcodes zählen AAG, GAG, CAG und GGG. Dabei fällt auf, dass vor allem das

Kopplungsreagenz DMT-MM und das Lösungsmittel DMF angereichert sind. Diese scheinen gute Synthesebedingungen darzustellen.

Anschließend wurden die Codonsets ausgewertet, die tatsächlich bei der Sequenzierung am häufigsten vorkamen, unabhängig von den vorher identifizierten einzeln angereicherten Barcodes. Dargestellt wurden wieder die Ergebnisse aus der Emulsions-PCR.



**Abb. 44: Graphische Darstellung der Häufigkeitsdifferenz (%) der Synthese gegen die Selektion für zwei Codonsets.** Die Häufigkeit der Barcodons für die Synthese- und die Selektionsreaktion wurden gegen einander gestellt, um eine Anreicherung und somit bevorzugte Synthesebedingungen fest zu stellen. Die den Barcodons entsprechenden Synthesebedingungen sind in der Legende dargestellt.

Die am häufigsten vorkommenden Codonsets stimmen fast vollständig mit denen, die von Interesse waren überein. In der Tabelle sind wieder die angereicherten Barcode-sets mit ihren zugehörigen Synthesebedingungen dargestellt.

**Tab. 25: Angereicherte Codonsets und deren entsprechenden Kopplungs- und Lösungsmittel nach zwei Syntheserunden.** Durch die dekodierten Barcodes kann auf die entsprechenden Synthesebedingungen rückgeschlossen werden.

Barcode	Kopplungsreagenz und Lösungsmittel
GGGGGG	CDI/Ethanol & CDI/Methanol
GGGTAG	CDI/Ethanol & DMF
GAGGGG	CDI/DMF & CDI/Ethanol
AAGGAG	DMT-MM/DMF & CDI/DMF
CGGTAG	EDC/Ethanol & DMF

Es konnte nach der Auswertung festgestellt werden, dass auch Kombinationen vorkamen bei denen entweder kein Kopplungsreagenz dabei war oder Wasser als Lösungsmittel diente. Allerdings gibt es keine Kombination, bei der Wasser bzw. kein Kopplungsreagenz in dem ersten Codonset vorkam.

Somit kann abschließend festgestellt werden, dass bei der Verwendung von Kopplungsreagenzien in Lösungsmittel gegenüber der Verwendung von ausschließlich dem Lösungsmittel ein eindeutiger Vorteil in der Syntheseeffizienz während der chemischen Synthese entstanden ist. Dies konnte durch die Anreicherung an das Zielmolekül Streptavidin festgestellt werden. Zudem konnte analysiert werden, dass besonders oft das Lösungsmittel DMF und das Kopplungsreagenz DMT-MM eine Anreicherung in der Codonsequenz gezeigt haben.

## 4 DISKUSSION

Die dargelegte Technologie stellt ein neues Verfahren für die Synthese von kombinatorischen Banken, wobei jede Verbindung der Bank kovalent an eine einzigartige Nukleinsäure verknüpft ist, dar. Die Nukleinsäure dient dabei als Anleitung für die Synthese der Verbindung und besitzt gleichzeitig einen molekularen Barcode zur Identifizierung des gebundenen Moleküls. Durch das Auslesen der Nukleinsäure wird die Selektion und gerichtete Evolution von Verbindungen möglich gemacht. Im Gegensatz zu anderen Methoden zur Generierung von DNA-kodierten Substanz-Banken (Halpin & Harbury, 2004) kann mit dieser Methode mit sehr kurzen Nukleinsäurefragmenten und einer hohen und flexiblen Codierungsdichte gearbeitet werden.

### 4.1 MAGNETISCHE MIKROPARTIKEL UND DEREN EIGENSCHAFTEN

Durch die Entwicklung der Nano- und Mikropartikel als Suchmaschine für Arzneimittel ist einer der wegweisenden biomedizinischen Anwendungen entstanden. (Ruiz-Hernández *et al.*, 2011). In dieser Arbeit sollten magnetische Mikropartikel für die verschiedensten Arbeitsschritte verwendet werden. Zum einen für die kovalente- und biotinylierte Kopplung und zum anderen für die ionische Bindung von Nukleinsäure an magnetische Partikel. Durch diese Kopplungsverfahren konnte im Anschluss mit der Nukleinsäure praktisch gearbeitet werden, da aufgrund der magnetischen Partikel deren Handhabung vereinfacht wurde. Darunter fallen sowohl Waschschrte als auch die eigentliche Synthesereaktion. Die gekoppelte bzw. gebundene DNA liegt auf diese Weise nicht frei in Lösung, sondern immobilisiert an Mikropartikeln vor und lässt sich durch den magnetischen Effekt der Partikel innerhalb eines Reaktionsgefäßes einfach bearbeiten, bewegen und wieder aufnehmen.

Für die kovalente Kopplung mittels dem Carbodiimid-Verfahren wurden die Dynabeads MyONE Carboxylic Acid von Invitrogen oder die SiMAG-Carboxyl von Chemicell verwendet. Die Mikropartikel von SiMAG haben dabei einen halb so großen Durchmesser im Vergleich zu den Dynabeads MyONE ( $\varnothing$  1  $\mu\text{m}$ ). Die Dichte der Carboxylgruppen liegt bei den SiMAG-Partikeln bei 2,25 g/cm<sup>3</sup> und bei den Dynabeads MyONE bei 1,8 g/cm<sup>3</sup> (Herstellerangaben). Die kleineren Mikropartikel zeigen somit auch eine höhere Dichte der Carboxylgruppen auf deren Oberfläche. Aufgrund der höheren Dichte sollte die

Bindungskapazität etwas höher sein und die Hintergrundaktivitäten der Partikel geringer. In der Praxis (siehe Abb. 16) konnte gezeigt werden, dass die Hintergrundaktivitäten bei beiden Mikropartikelvarianten gering ist und die größeren Partikeln (Dynabeads MyONE) trotz ihrer geringeren Dichte ein höheres Bindungssignal aufgezeigt haben.

Eine weitere Variante um Nukleinsäure, in diesem Fall biotinylierte, an magnetische Kügelchen zu koppeln ist die Verwendung von streptavidin-derivatisierten Partikeln (Wang *et al.*, 2001), welche sich ausgezeichnet für Automatisierungen eignen (Wochner *et al.*, 2007). Die hier verwendeten Dynabeads M-280 Streptavidin von Invitrogen besaßen einen Durchmesser von 2,8  $\mu\text{m}$  und eine Streptavidindichte auf deren Oberfläche von 1,4  $\text{g}/\text{cm}^3$  (Herstellerangaben). Somit handelte es sich bei diesen Mikropartikeln um die Größten mit der geringsten Dichte. Trotz alledem haben diese Partikel die besten Bindungseffizienzen gezeigt (siehe Abb. 17). Die Dichte des Streptavidins auf der Oberfläche der Partikel reicht völlig aus, um für die Nukleinsäure unzugänglich zu sein. Aus diesem Grund wurde entschieden, die Hybridisierungsoligonukleotide an streptavidin-derivatisierten Mikropartikeln zu immobilisieren, um anschließend die Separation der DNA-Bank vornehmen zu können.

Um auch die Synthesereaktion handhabbar machen zu können und eine Festphasensynthese durchführen zu können, wurde die sortierte DNA-Bank nach der Spaltungsreaktion wieder an magnetische Partikel gebunden. Zu diesem Zweck wurden die SiMAG-DEAE Mikropartikel ( $\emptyset$  0,5  $\mu\text{m}$ ) von Chemicell verwendet. Bei diesen Mikropartikeln handelt es sich um Kügelchen, die auf ihrer Oberfläche mit Diethylaminoethylgruppen ausgestattet sind und so als schwacher Anionenaustauscher fungieren. Die Dichte dieser funktionellen Gruppe ist recht hoch und liegt bei 2,25  $\text{g}/\text{cm}^3$ . Die Kopplung der Nukleinsäure erfolgte hier über eine ionische Bindung. In der Arbeit konnte gezeigt werden, dass die Bindung als auch die Elution von den SiMAG-DEAE Mikropartikel sehr effizient erfolgte (siehe 3.1.3). Dabei war es nicht nötig eine Aufreinigung zwischen der Spaltungsreaktion und der folgenden Bindung an die Mikropartikel vorzunehmen, sondern durch Verdünnung des Spaltungsansatzes konnte die passende Salzkonzentration für die Bindung an die DEAE-Partikel eingestellt werden. Auf diesem Wege wurde im iterativen Ablauf dieser Technologie ein weiterer Zwischenschritt eingespart.

Durch neuartige therapeutische und diagnostische Verfahren in der Entwicklung oder in der klinischen Praxis können die frühen Erkenntnisse der Nanotechnologie auf die Medizin realisiert werden (Farokhzad & Langer, 2006). Darunter zählt natürlich auch die Verwendung von magnetischen Mikropartikeln, die auf ihrer Oberfläche (je nach Anwendung) mit den unterschiedlichsten funktionellen Gruppen bestückt sein können.

## 4.2 DIE EFFIZIENZRATE DER EINZELNEN ITERATIVEN ARBEITSSCHRITTE

Die vorgestellte Arbeit „Die Entwicklung einer neuen Methode zur Selektion von synthetischen Wirkstoffbanken“ besteht aus einer Vielzahl von einzelnen iterativen Arbeitsschritten, die im Laufe optimiert und vor allem auf den Magnetpartikelprozessor KingFisher Flex von Thermo Scientific übertragen werden sollten. Durch dieses Gerät können die Mikropartikel automatisch nach jedem Arbeitsschritt in ein neues Gefäß transferiert werden, so dass Verklumpungen der Partikel verhindert werden können (Eulberg *et al.*, 2005). Zudem ist es möglich mit Mikrotiterplatten zu arbeiten, wodurch ein hoher paralleler Durchsatz, eine einfache Handhabung und reproduzierbare Bedingungen gewährleistet werden (Glökler *et al.*, 2010). Das Gerät erlaubt völlig flexible Einstellungen und eigen programmierbare Protokolle, so dass diese je nach Bedarf angefertigt und angepasst werden können.

Somit bestand ein primäres Ziel dieser Arbeit darin, die Arbeitsschritte, soweit diese es zuließen, auf den Roboter für eine einfache Bedienung zu transferieren. Es konnte gezeigt werden (siehe Tabelle 22), dass die Hälfte der iterativen und parallelen Arbeitsschritte auf den Roboter für ein halbautomatisiertes Verfahren übertragen werden konnten. Auf diesem Wege konnten auch die einzelnen Arbeitsabläufe optimiert und deren Effizienz gesteigert werden, da so vor allem manuelle Fehler vermindert wurden. Die hier gezeigte Methode der Herstellung von DNA-kodierten Substanz-Banken entspricht der sequenzabhängigen Sortierung und Synthese (DNA-routed synthesis) und gewährleistet die Kopplung zwischen Genotyp (DNA) und Phänotyp (Molekül) und ermöglicht somit die *in vitro* Evolution.

#### 4.2.1 DIE LIGATION DER DNA-BANK MIT DEN ADAPTEROLIGONUKLEOTIDEN

Ein Hauptmerkmal dieser Arbeit ist die Ligationsreaktion zwischen der DNA-Bank und den Adapteroligonukleotiden. Durch die Ligation zwischen diesen beiden Komponenten wird zum einen die Dekodierung und zum anderen die Sortierung gewährleistet. Die Adapter bestehen aus einer 20-mer Tagsequenz, die zur Sortierung verwendet wurde und gleichzeitig einmalig zu deren Barcodon (siehe Abb. 7) war. Während der Ligation hybridisiert dieser Barcode mit der komplementären Sequenz der DNA-Bank und wird mittels der T4 DNA Ligase ligiert. Die Spaltung erfolgte an der abasischen Position der Adapteroligonukleotide, wodurch der Barcode die DNA-Bank um drei Nucleotide verlängerte und gleichzeitig die Dekodierung erfolgte. Die Ligationsreaktion wurde mit 16 verschiedenen Adaptermolekülen durchgeführt, da der Barcode aus fünf Codons mit je zwei unbekanntem und einer konstanten Base bestand. Somit ergaben sich pro Ligationsreaktion 16 verschiedene Oligos ( $4^2$ ).

Die T4 DNA Ligase verbindet effektiv zwei benachbarte, kurze Oligonukleotide (Deugau & van de Sande, 1978). Dabei katalysierte die Ligase die ATP-abhängige Phosphodiesterbindung zwischen der 5'-Phosphat- und der 3'-Hydroxylgruppe der benachbarten DNA-Stränge (Wu & Wallace, 1989). Der Mechanismus beinhaltet die Bildung eines Enzym-Adenylat-Komplexes, gefolgt vom Transfer des Adenylates vom Enzym auf die Phosphatgruppe des Substrates. Schlussendlich wird die Phosphodiesterbindung zwischen den benachbarten Strängen gebildet (Weiss & Richardson, 1967; Olivera *et al.*, 1968; Harvey *et al.*, 1971). Die T4 DNA Ligase benötigt als Ligationssubstrat doppelsträngige komplementäre DNA-Abschnitte, ein Einzelstrang-Fragment kann der Ligase nicht als Substrat dienen (Higgins & Cozzarelli, 1978). Die Effizienz der Ligation wird dabei noch durch weitere Faktoren beeinflusst. Der komplementäre Bereich sollte nicht zu kurz gewählt sein, da ansonsten die Effizienz leidet. In dieser Arbeit konnte festgestellt werden, dass mindestens acht komplementäre Basen für eine effiziente Ligation vorhanden sein sollten. Verstärkend kommt hinzu, dass das Vorhandensein von Basenpaar-Mismatches die Effizienz der Ligationsreaktion vermindert (Wu & Wallace, 1989), gleichzeitig zeigt die T4 DNA Ligase im Vergleich zu anderen Ligasen überhaupt eine Toleranz gegenüber Basenpaar-Mismatches (Lohman *et al.*, 2001). In diesem Fall handelt es sich um abasische Positionen, die sich im Adapteroligonukleotid für die Spaltungsreaktion befanden. Diese abasischen Positionen wurden von der T4 DNA Ligase toleriert. Trotz alledem wurde aus diesem Grund der Hybridisierungsbereich so groß wie möglich gewählt, denn die Ligationsreaktion sollte



mit der bestmöglichen Effizienz durchgeführt werden. Ein interessanter Aspekt ist dabei, dass die T4 DNA Ligase sogar eine AP (apurinic/aprimidinic) Lyaseaktivität besitzt. Diese wird durch die Präsenz von ATP inhibiert (Bogenhagen & Pinz, 1998). Das Konzentrationsverhältnis der DNA-Bank zu den Adapteroligonukleotiden wurde in einem Verhältnis von 1:20 für die Ligationsreaktion gewählt. Da es sich um 16 verschiedene Adapter handelte, konnte auch nur jedes 16te Molekül hybridisieren. Um ein äquimolares Verhältnis zu erzielen, wurden die Adapter in einem hohen Überschuss zu der Reaktion gegeben. Des Weiteren konnte im Verlauf der Arbeit festgestellt werden, dass die Ligationseffizienz in der ersten Runde geringer war als in der Zweiten (nicht dargestellt). Dass heißt die Ligationseffizienz im zweiten Synthesedurchgang zeigte eine bessere Ausbeute als in der ersten Runde. In der ersten Runde wurde die gesamte DNA-Bank für die Ligation eingesetzt bzw. dargestellt. Da die DNA-Bank aus konstanten Positionen im Barcode (Cytosin) bestand, könnte es sein, dass nach der ersten Ligationsrunde alle korrekt synthetisierten DNA-Fragmente selektiert wurden und die Fehlerhaften durch die nicht vorhandene Ligation anschließend bei Waschschrritten verworfen wurden. So müßten sich im zweiten Durchgang bessere Effizienzen zeigen, da die fehlerhaften Oligonukleotide beseitigt wurden. Zudem liegt die DNA-Bank in der zweiten Ligationsrunde in einer geringeren Konzentration im Vergleich zu der ersten Ligationsrunde durch den vollzogenen Syntheszyklus vor. Die Konzentration der T4 DNA Ligase bleibt aber gleich. Eventuell ist die bessere Effizienz in der zweiten Runde auch auf diesen Zusammenhang rückzuschließen. Das Ziel dieser Arbeit bestand auch darin, die einzelnen Schritte so effektiv, kostengünstig und handhabbar wie möglich zu gestalten. Aus diesem Grund sollte durch die Verwendung von weiteren Enzymen als auch der Änderung des Reaktionsablaufes getestet werden, ob die Ligationseffizienz noch gesteigert bzw. die Reaktion vereinfacht werden kann.

Im Laufe der Arbeit wurde der Ablauf der Ligationsreaktion von einer einfachen zu einer zyklischen Variante geändert. Das Konzept basierte dabei auf einer Ligationskettenreaktion (ligase chain reaction) (Wiedmann, 1994). Die hier angewendete Methode unterscheidet sich dabei in ein paar wichtigen Punkten von der eigentlichen LCR-Methode. Es wird nach einer einmaligen initialen Denaturierung zwischen der Primertemperatur von 45°C und der Ligationstemperatur von 22°C zyklisch zwischen den Temperaturen gewechselt. Durch eine erneute Denaturierung würde die T4 DNA Ligase inaktiviert werden und müßte jedes Mal neu hinzugefügt werden. Dies wäre sehr kostenintensiv und auch aufwendig – zudem auch nicht notwendig – da die Temperatur von 45°C ausreicht, so dass sich die falsch angelagerten Adapteroligonukleotide wieder von der DNA-Bank lösen und Platz für die komplementären

Oligos machen. Dieser Aspekt war der Hauptgrund für die Einführung der zyklischen Ligationsreaktion. Nicht-ligierte Adapter sollten sich auch diese Weise wieder von der DNA-Bank lösen und die blockierten Bindungsstellen wieder freigeben. Durch die zyklische Reaktion konnten sich die Adapter immer wieder neu sortieren und an den richtigen Bindungsstellen der DNA-Bank hybridisieren. Die Effizienz konnte mit diesem Verfahren nochmal gegenüber der einfachen Ligationsreaktion gesteigert werden (siehe Abb. 21).

Die hier dargestellte zyklische Ligationsmethode ist eine Alternative zu dem eigentlichen LCR-Verfahren (Wiedmann, 1994), wenn man eine ATP-abhängige Ligase verwenden möchte, die zudem nicht thermostabil ist. Bei der thermostabilen Ligase, wie sie eigentlich verwendet wird, handelt es sich um eine NAD<sup>+</sup>-abhängige Ligase, die keine abasischen Positionen im zu ligierenden Oligonukleotid akzeptieren würde (Lohman *et al.*, 2001).

Darüber hinaus wurde getestet, ob die Verwendung von weiteren Enzymen auch eine Ligationsreaktion ermöglicht. Dabei sollte zudem versucht werden, die Ligationsreaktion unter isothermalen Bedingungen zu gestalten. Die *E. coli* Ligase sollte als Alternative zur T4 DNA Ligase eingesetzt werden. Dabei handelt es sich um die DNA Ligase von *E. coli*. Dieses Enzym katalysiert die Reparatur von Einzelstrangbrüchen in Doppelstrangfragmenten. Anders als die T4 DNA Ligase ist diese nicht in der Lage auch „glatte Enden“ zu ligieren, sondern kann dies nur bei „klebrigen Enden“ (Lohman *et al.*, 2001).

Die *E. coli* Ligase konnte in keiner Reaktion ein Ligationsfragment aufzeigen (siehe Abb. 23) und eignet sich nicht als Ligationsenzym, da diese die abasischen Positionen bei einer Ligationsreaktion nicht akzeptiert. Als zusätzliches Enzym wurde sowohl bei der isothermalen als auch der zyklischen Ligationsreaktion das SV40 Large T Antigen (Helikase) eingesetzt. Durch die Aktivität dieser Helikase sollte erreicht werden, die Reaktion unter isothermalen Bedingungen (Jeong *et al.*, 2009) gestalten zu können bzw. die Ligase bei der zyklischen Reaktion zu unterstützen. Die Aktivität der Helikase entwindet DNA-Fragmente (Baran *et al.*, 1997), wodurch eine isothermale Reaktion möglich wäre oder aber das Enzym löst falsch angelagerte Adapteroligonukleotide wieder von der DNA-Bank und würde somit die Effizienz der Ligation steigern. Schlussendlich konnte festgestellt werden, dass der Einsatz der Helikase keine Steigerung in der Effizienz der Ligationsreaktion zeigte (siehe Abb. 23) und der Einsatz nur die Kosten erhöhen würde. Das vorher etablierte Protokoll der zyklischen Ligationsreaktion mittels T4 DNA Ligase zeigte die besten Ligationseffizienzen unter Verwendung der geringsten Kosten.

Die Familie der DNA Ligasen sind zentrale Enzyme in der Molekularbiologie, Nukleinsäureforschung und der Hochdurchsatzsequenzierung. Die Enzyme zeigen alle eigene

Eigenschaften (Lohman *et al.*, 2001), wobei sich die der T4 DNA Ligase für unser Verfahren als die Effizienteste herausgestellt hatte.

#### 4.2.2 SORTIERUNGSMECHANISMUS

Ein weiterer wichtiger Arbeitsschritt bestand in der sequenzspezifischen Separation („aufspreizen“) der ligierten DNA-Bank anhand der Sequenzabfolge. Das hier dargestellte Verfahren beruht auf dem Prinzip der sequenzabhängigen Sortierung und anschließenden Synthese. Bei dieser Methode wird ganz im Gegenteil zu anderen Techniken (Li & Liu, 2004; Hansen *et al.*, 2009; Melkko *et al.*, 2004) die DNA-Bank vor der Synthese anhand ihrer Sequenz sortiert und in einzelne Reaktionsgefäße überführt. Anschließend erfolgt die eigentliche Festphasensynthese in separaten Behältern in einem hohen parallelen Durchsatz. In dieser Arbeit wurde zur Sortierung eine 20-mer Tagsequenz verwendet, die das 3'-Ende der Adapteroligonukleotide bildete. Komplementäre Hybridisierungsoligonukleotide wurden zu diesem Zweck an magnetische Mikropartikel gekoppelt und an den Magnetpartikelprozessor KingFisher Flex an den einzelnen Magnetspins geankert. Anschließend erfolgte die Inkubation, wobei die Adapteroligonukleotide anhand ihrer Sequenz an den magnetischen Partikeln mittels Hybridisierung abgefangen werden konnten. Eine anschließende Sanger-Sequenzierung konnte zeigen, dass diese recht spezifisch verläuft und dass eine Separation mit einer sehr kurzen einmaligen Tagsequenz möglich ist. Durch die Verwendung von solchen kurzen Oligonukleotiden können die Kosten reduziert und in kleinen Volumina gearbeitet werden.

Im Vergleich zu anderen Technologien, die auch mit der „Split & Pool“-Methode mit evolutionären Prozessen (Halpin & Harbury, 2004a;b;c) arbeiten, wird in der hier vorgestellten Arbeit mit kleineren und stabileren Oligonukleotiden gearbeitet. Die hier präsentierte DNA-Bank ist lediglich ca. 80 bp lang und liegt teilweise doppelsträngig vor. Durch die doppelsträngige Struktur ist die DNA stabiler (Henning, 2002) als sie im Einzelstrangzustand wäre. Durch die geringe Größe können Kosten vermindert und die Handhabung verbessert werden. Ein besonderer Unterschied liegt in dem Vorgang der Sortierung. Bei Halpin & Harbury ist die DNA-Bank 340 bp groß und besteht aus 20-mer DNA-Abschnitten, die als Kodierungsregionen bezeichnet werden. Diese hybridisieren

während der Sortierung mit komplementären DNA-Abschnitten, den sogenannten Anticodons, und werden dabei durch die Kopplung an Resinpartikeln separiert (Halpin & Harbury, 2004a). Es gibt zwei grundlegende Unterschiede in der hier vorgestellten Technik. Zum einen besteht die Kodierungsregion bei der hier beschriebenen Technik aus lediglich neun Basenpaaren, wobei nur das erste Triplet für die Spezifität bzw. die Dekodierung verantwortlich ist. Zum anderen wird hier eine strikte Reihenfolge während der Hybridisierung beibehalten. In jedem Zyklus rutscht das komplementäre Oigo (Adapteroligonukleotid) zur DNA-Bank um ein Triplet weiter. Bei Halpin & Harbury kann die Hybridisierung irgendwo im 340-mer Fragment erfolgen, ohne dass rückgeschlossen werden kann an welcher Stelle diese erfolgte. Durch die hier vorgestellte Arbeit kann genau analysiert werden, welcher Genotypbaustein zu welchem Phänotypbaustein gehört. Dies wird durch den anschließenden Spaltungsmechanismus an der abasischen Position (hinter dem Triplet) im Adapteroligonukleotid gewährleistet. Die 20-mer Tagsequenz ist dabei einzigartig zu der Kodierungsregion und dient der Separation. Durch die anschließende Spaltung wird diese aber wieder entfernt und lediglich das Triplet bleibt erhalten. Somit kann mit wesentlich kleineren Oligonukleotiden gearbeitet werden und der kurze Hybridisierungsbereich erlaubt eine Vielzahl von Hybridisierungs- und somit Syntheseyklen, ohne dass die DNA-kodierte Substanz-Bank ihre geringe Größe und somit flexible Handhabung verliert.

#### **4.2.3 SPALTUNG AN EINER DEFINIERTEN POSITION IM DNA-FRAGMENT**

Bei der hier vorgestellten Methode wurde nach der Ligationsreaktion der DNA-Bank mit den Adaptermolekülen eine Sortierung mittels der Tagsequenz des Adapters durchgeführt. Nach erfolgreicher Separation erfolgte an der ligierten DNA-Bank eine Spaltung. Diese diente nicht nur einem Zweck, sondern gleich mehreren. Zum einen wurde so das DNA-Fragment von den magnetischen Partikeln gelöst und konnte für die Synthesereaktion an die lösungsmittelstabilen DEAE-Partikel gebunden werden. Zum anderen diente die Spaltung der Freisetzung der Tagsequenz. Die Reaktion erfolgte zu diesem Zweck an einer definierten Position – an der abasischen Stelle – im Fragment. Das hybridisierte Triplet blieb an der DNA-Bank verknüpft, so dass im nächsten Zyklus die Ligation um drei Basen versetzt erfolgen konnte.

Da die Triplets zur DNA-Bank komplementär sind, konnte durch Auslesen der DNA-Bank bestimmt werden, welches Molekül/Synthesebedingungen gekoppelt/angewendet wurde. Um diese distinkte Arbeitsabfolge gewährleisten zu können, musste die Spaltung an der richtigen Stelle erfolgen, als auch die passenden Fragmentenden für einen weiteren Zyklus hergestellt werden.

Das gewählte Enzym musste in der Lage sein das Doppelstrangfragment an einer definierten Position so zu spalten, dass im Anschluss ein 5'-Phosphat- und 3'-Hydroxylende vorliegt. Bei dem gewählten Enzym sollte es sich also um eine Endonuklease handeln. Es gibt die verschiedensten Möglichkeiten solch eine Spaltung vorzunehmen. Es könnte die RNase HII von *E. coli* verwendet werden. Dabei handelt es sich um eine Endoribonuklease, welche in der Lage ist 5'-Phosphat- und 3'-Hydroxylenden durch die Hydrolyse der Phosphordiesterbindung zu generieren. Dies wäre allerdings nur möglich, wenn in dem zu spaltenden Fragment sowohl RNA als auch DNA enthalten wäre (Nichols & Yue, 2008). Da in dieser Arbeit nicht mit RNA-DNA-Bausteinen gearbeitet wurde, wurde die RNase HII auch nicht für die Spaltungsreaktion weiter in Betracht gezogen. Bei der eingebauten Modifikation handelte es sich um eine abasische Stelle direkt hinter dem ligierten Triplet. Somit muss das gewählte Enzym in der Lage sein, diese abasische Position zu spalten. Ein Enzym, das diese Fähigkeit zeigt, ist die Endonuklease APE 1. Dieses Enzym ist in der Lage abasische Positionen in DNA-Fragmenten zu reparieren. Dabei generiert das Enzym einen DNA-Einzelstrangbruch mit einem 5'-Phosphat- und 3'-Hydroxylende (Li *et al.*, 2014).

In vorangegangenen Studien (hier nicht gezeigt) konnte allerdings gezeigt werden, dass eine andere Endonuklease die besten Spaltungseffizienzen aufzeigen konnte. Dabei handelt es sich um die Endonuklease IV von *E. coli*. Dieses Enzym ist an der Reparatur der DNA beteiligt, indem es das DNA-Gerüst an der apurinic/apyrimidinic (AP)-Stelle durch Hydrolyse spaltet und dabei ein 5'-Phosphat- und 3'-Hydroxylende (Hosfield *et al.*, 1999) generiert. Neben der Reparaturfunktion spaltet dieses Enzym doppelsträngige DNA-Fragmente, wenn diese synthetische, abasische Positionen besitzen (Takeuchi *et al.*, 1994). Bei abasischen Stellen handelt es sich um die häufigsten DNA-Lesionen, die spontan in der DNA auftauchen (Puri *et al.*, 2013). In dieser Arbeit wurde nachgewiesen, dass die Spaltungsreaktion mit einer Effizienz von etwa 70% (siehe Abb. 45) durchgeführt werden konnte. Durch die iterativen Abläufe wurde zudem gezeigt, dass diese Spaltungsreaktion an der definierten abasischen Position erfolgt und dabei ein 5'-Phosphat- und 3'-Hydroxylende generiert wurde, da ansonsten keine zweite Ligationsreaktion möglich gewesen wäre.

Die Endonuklease IV hat sich als ein hervorragendes Spaltwerkzeug mit großer Genauigkeit und hoher Effizienz in dieser Arbeit hervorgetan.

#### 4.2.4 DIE CHEMISCHE SYNTHESE UND DIE SELEKTION

Ein wichtiges Arbeitsfeld bei der Herstellung von DNA-kodierten Substanz-Banken stellt die Synthese von kleinen Molekülen an die DNA-Fragmente dar. Seit der Festphasensynthesemethode, die im Jahre 1963 von Merryfield eingeführt wurde, konnte sich das Konzept der kombinatorischen Chemie entwickeln (Shin *et al.*, 2005). Auch in dieser Arbeit wurde sich für eine Festphasenpeptidsynthese die mittels der Carbodiimid-Methode vorgenommen wurde, entschieden. Bei der Festphasensynthese, wird das herzustellende Peptid an einem festen Träger hergestellt. In dieser Studie wurde zu diesem Zweck mit DEAE-Magnetpartikeln von Chemicell gearbeitet. Diese können aufgrund ihrer magnetischen Eigenschaft gut gehandhabt werden. Somit können Reagenzien und Nebenprodukte schnell und einfach isoliert werden und gleichzeitig können Reagenzien in großem Überschuss für eine bessere Kopplungsausbeute verwendet werden. Vorteile der Immobilisierung liegen vor allem auch in der Automatisierbarkeit des Prozesses. Das Prinzip der Peptidsynthese beruht auf der Kopplung zweier Aminosäuren, wobei die Aminogruppe der Aminosäure geschützt und gleichzeitig das Carboxyl der Aminosäure zur Kopplung aktiviert wird (Bergmann & Zervas, 1932). Somit stellt die Peptidsynthese ein Wechselspiel aus Schutzgruppeneinführung und -Abspaltung sowie Aktivierung der Carboxylkomponenten dar. Zur Einführung von Schutzgruppen haben sich verschiedene Methoden etabliert. Wichtig dabei ist zu beachten, dass einige Aminosäuren eine Seitenkettenschutzgruppe benötigen. Darunter fallen die Aminosäuren Lysin, Arginin, His, Glutaminsäure, Asparaginsäure, Serin, Threonin, Tyrosin und Cystein. Diese wird benötigt, um unerwünschte Nebenprodukte zu vermeiden (Altmann & Mutter, 1993). Um diesen zusätzlichen Schritt zu vermeiden, wurde in dieser Arbeit für die Synthesereaktion mit der Aminosäure Glycin gearbeitet. Diese benötigt keine Seitenkettenschutzgruppe. Somit ist es bei der Wahl der Schutzgruppe wichtig, dass die terminale Schutzgruppe im Laufe des Zyklus entfernt werden kann, die Seitenkettenschutzgruppen allerdings nicht geschädigt werden. (Altmann & Mutter, 1993). Zu den wichtigsten Schutzgruppeneinführungen zählen die Benzyloxycarbonyl-Methode, die *tert.*-Butyloxycarbonyl-Methode, die Fluoren-9-ylmethoxycarbonyl-Methode und die

Allylester-Methode. Diese sind unter anderem mittels Essigsäure, Trifluoressigsäure oder sekundären Aminen wie Piperidin spaltbar (Altmann & Mutter, 1993; Beyer & Walter, 1988). Auch bei der Knüpfung der Peptidbindung gibt es zahlreiche Methoden. Die wichtigsten darunter basieren auf der Verwendung energiereicher N-acylierter Aminosäurederivate. Dazu gehört die Säurechlorid-, Azid-, Carbodiimid-, Azolid-, und N-Carbonsäureanhydrid-Methode (Beyer & Walter, 1988).

Die Entwicklung in der Identifizierung von neuen Wirkstoffzielen sowie die Verbesserung bei der Herstellung von Peptid-Technologien haben die Nachfrage nach hochwirksamen synthetischen Peptid-Banken stets erhöht. Zur Erzeugung von synthetischen Peptidbanken hat sich die kombinatorische Chemie als ein effektives Werkzeug etabliert und stellt ein wichtiges Werkzeug bei der Entdeckung neuer Arzneimittel (Shin *et al.*, 2005) dar. Unerwähnt soll nicht bleiben, dass es neben der Peptidsynthese auch alternative Formen wie die Click-Chemie, Harnstoffsynthese, reduktive Aminierung, nukleophile aromatische Substitution, die Michael- und Staudinger-Addition, die Diels-Alder-Reaktion, die Suzuki-Kopplung und die Wittig-Reaktion gibt.

Zur Selektionsreaktion ist zu sagen, dass diese dem eigentlichen Auffinden von biologisch relevanten Liganden, die an einem Zielmolekül binden, dienen. Durch die Etablierung der DNA-kodierten Banken wurde auch dieses Verfahren vereinfacht. Im Vergleich zu anderen Technologien können DELs einfachen Selektionsverfahren mit minimalen Einsatz der generierten DNA-kodierten Bank und dem immobilisierten Zielmolekül unterzogen werden. Dem gegenüber stehen die kosten- und laborintensiveren Technologien wie das Phage-Display oder Hochdurchsatzscreening.

### 4.3 HOCHDURCHSATZSEQUENZIERUNG VS. SANGER-SEQUENZIERUNG

Neben der traditionellen Sanger-Sequenzierung sind längst Verfahren entwickelt worden, deren Leistungsfähigkeit die der Sanger-Sequenzierung um ein Vielfaches überragen und den Begriff der Hochdurchsatzsequenzierung neu definieren (Stangier *et al.*, 2007). Die neue Generation der Hochdurchsatzsequenzierung (NGS) steigert dramatisch die Anzahl der klinisch verfügbaren genetischen Tests und ist einer der wichtigsten Arbeitsfelder in der biomedizinischen Forschung (Fox *et al.*, 2009; Swanson *et al.*, 2014; Zhu *et al.*, 2013). Durch die Entwicklung der Sequenzierungstechnologien in den letzten Jahren wurde die

Anzahl der verschiedensten Anwendungen in der Biologie als auch der Medizin erhöht (Berglund *et al.*, 2011). Die Sanger-Sequenzierung, auch als „Sequenzierung der ersten Generation“ bekannt, dominierte die letzten Jahrzehnte den Markt (Zhu *et al.*, 2013). Dabei handelt es sich um eine sehr akkurate Methode, um lange DNA-Fragmente auszulesen. Mit ihr wurde von 1998-2003 mit einem Milliarden-Dollar-Einsatz die erste Version des drei Milliarden Nucleotide umfassenden Humangenoms erstellt (Swanson *et al.*, 2014). Die Nachteile der Sanger-Sequenzierung gegenüber den neuen Hochdurchsatzsequenzierungen sind offensichtlich. Im Vergleich zur Sanger-Sequenzierung ist die NGS-Plattform fähig, einen hohen parallelen Durchsatz von Datenmengen zu erschaffen (Woollard *et al.*, 2011). Zudem sind die Probenvorbereitung und Auftrennung der Nucleotidsäurefragmente im elektrischen Feld sehr aufwendig. Hinzu kommt die geringe Parallelisierbarkeit der Sanger-Sequenzierung. NGS kann auf mehreren kommerziell erhältlichen Plattformen durchgeführt werden. Dazu zählen Roche`s 454 GS FLX Genome Analyzer, vermarktet durch Roche Applied Sciences, das Solexa-Sequenzierungssystem von Illumina und das durch Applied Biosystems vermarktete SOLiD-System. Alle drei Systeme basieren auf dem gleichen Prinzip. Zunächst erfolgt die Preparation der Fragmentbank, danach eine massive parallele klonale Amplifikation und im Anschluss die eigentliche Sequenzierung (Woollard *et al.*, 2011). Somit wird auf langsame und kostspielige Abläufe beim NGS-Verfahren verzichtet. Es wird zum Beispiel auf die klassische Klonierung von DNA verzichtet (Morozova *et al.*, 2009) und auf die langsame Auftrennung der Sequenzierprodukte durch die Gelelektrophorese. Zwar wird die Sanger-Sequenzierung nicht vollständig abgeschafft und weiterhin zur Entschlüsselung kleiner DNA-Moleküle verwendet werden, doch gelangen immer fortschrittlichere und neuere Sequenzierungstechnologien auf den Markt. Auch in dieser Arbeit wurde mithilfe der Sanger-Sequenzierung die Spezifität der Separation nachgewiesen (siehe 3.3.3) und konnte für eine genaue Analyse dieser herangezogen werden. Für die Analyse und Auswertung der DNA-kodierten Substanz-Bank konnte in dieser Studie mit der Hochdurchsatzsequenzierung gearbeitet werden. Eingefügte Barcode-Sequenzen, welche über Primer eingefügt wurden, erlauben die parallele Sequenzierung der gesamten DNA-kodierten Bank inklusive der initial eingesetzten DNA-Bank. Durch diese Methode ist eine spätere Zurodnung der erhaltenen Datensätze möglich. So wurde das Potenzial der 454-Pyrosequenzierungsmethode ausgenutzt. Die Stärken dieser Methode liegen aber vor allem in der langen Leselänge und der hohen Qualität der Daten (Timmermann *et al.*, 2009). Die Sequenzierung lieferte 70.000 Sequenzen und nach genauen Analysen konnten in dieser Studie mittels Hochdurchsatzsequenzierung optimale Synthesebedingungen für die Synthese



von d-Biotin und Glycin für die Herstellung der DNA-kodierten Bank ermittelt werden. Für die Analyse von DNA-kodierten Substanz-Banken können Hochdurchsatzsequenzierungssysteme der nächsten Generation effektiv herangezogen werden.

#### **4.4 DELS ALS NEUE PLATTFORM IN DER PHARMAZEUTISCHEN UND MEDIZINISCHEN INDUSTRIE**

Die Entwicklung von DNA-kodierten Substanz-Banken wird von einer Vielzahl von Arbeitsgruppen verfolgt. Dabei handelt es sich um ein wichtiges Technologiefeld, um neuartige Liganden, welche an Zielmoleküle binden, zu finden. Diese Liganden sind von großem pharmazeutischem und medizinischem Interesse. Die Erkennung eines Zielmoleküls auf molekularer Ebene stellt dabei nicht nur die Basis biologischer Regulation und Kommunikation sondern auch die Grundlage für die Entwicklung von Arzneimitteln dar. Bei vielen Krankheiten werden Zielsysteme erkannt, deren Blockade das Fortschreiten der Krankheit mindern würde (Kessler, 1993). Da diese Zielsysteme allerdings unzureichend identifiziert sind, ist es wichtig eine „Screening-Methode“ zu entwickeln, bei der eine Wirkstoffbank mit einer sehr hohen Diversität angewendet werden kann und so eine Bindung an das Zielsystem ermöglicht wird. Somit haben sich über die letzten Jahre die Methoden zur Durchsichtung von Liganden, die an Zielmoleküle binden, immer weiter entwickelt. Aus den gängigen Methoden heraus, wie das Phage-Display oder der SELEX-Technologie, haben sich die DNA-kodierten Banken entwickelt. Auch in dieser Arbeit wurde das Ziel verfolgt, eine neue Technologie für die Entwicklung von DNA-kodierten Substanz-Banken zu entwickeln. Es konnte nicht nur die Methode aufgezeigt werden – sondern auch – dass es mit der hier vorgestellten Technologie möglich ist, eine Selektion für optimale Synthesebedingungen durchzuführen.

Die in dieser Studie verwendete DNA-Bank besaß fünf Barcodons, wobei ein Barcode aus zwei unbekanntem und einer konstanten Base an der dritten Position bestand (3'-NNCNNCNNCNNC-5'). Somit wurde die Ligationsreaktion pro Syntheszyklus mit 16 verschiedenen Adapteroligonukleotiden (4<sup>2</sup>) durchgeführt. Nach zwei Runden ergab sich auf diese Weise bereits eine DNA-kodierte Substanz-Bank mit einer Diversität von 256 verschiedenen Molekülen. Nach fünf Runden könnten mit dieser Methode und der hier verwendeten Bank circa eine Millionen verschiedene synthetische Moleküle hergestellt

werden. Im Bezug zu anderen Technologien wird hier nur mit einer initialen DNA-Bank, wie es die Arbeitsgruppe um Li (Li *et al.*, 2013) macht, gearbeitet. Mit dieser initialen Bank ist es möglich eine DNA-kodierte Substanz-Bank mit einer Diversität von einer Millionen Molekülen in kleinen Volumina herzustellen, wobei durch Erhöhung der verwendeten Barcodons die Diversität noch weiter gesteigert werden könnte. Bei der hier vorgestellten Methode handelt es sich um ein flexibles System, welches durch weitere Anpassung in der Chemie der Synthese oder der verwendeten Komponenten für den Anwender frei einstellbar und auf den KingFisher Flex Roboter übertragbar ist.

Viele Strategien wurden mittlerweile für die Generierung von DNA-kodierten Banken vorgestellt (Litovchick *et al.*, 2014). Diese unterscheiden sich vor allem in der Größe der verwendeten DNA-Bank, der Synthesestrategie und der Anzahl der dargestellten chemischen Einheiten am DNA-Fragment (Buller *et al.*, 2011). Auch kommerziell verfügbar sind mittlerweile die Technologien um die DNA-kodierten Banken. DNA-templatbasierte Verfahren werden von den Firmen Ensemble Discovery und Vipergen entwickelt und solche, die auf der „Split-und Pool“-Technik basieren von GlaxoSmithKline, Nuevolution und Philochem (Scheuermann & Neri, 2010). Besonders die Gruppe um Scheuermann und Neri konnten in den vergangenen Jahren mit ihrer Technik einige Liganden gegen interessante Zielmoleküle aufzeigen. Dazu gehört ein Albuminbinder (Dumelin *et al.*, 2008), ein Ligand für das anti apoptotische Protein Bcl-xL (Melkko *et al.*, 2010), ein Inhibitor für das Molekül Interleukin 2 (Leimbacher *et al.*, 2012) und noch einige mehr (Dumelin *et al.*, 2006; Buller *et al.*, 2009).

Während Antikörper heutzutage praktisch gegen jedes beliebige Ziel unter Verwendung von Phage-Display-Methoden ermittelt werden können, fehlte lange ein ähnliches Verfahren für die einfache Entdeckung von niedermolekularen Verbindungen, die spezifisch an Zielproteine binden. Durch die Entwicklung der DNA-kodierten chemischen Banken kombiniert mit geeigneten Selektions- und Hochdurchsatzsequenzierungsstrategien scheint diese Lücke gefüllt worden zu sein (Scheuermann *et al.*, 2006). Phage-Display, Ribosomen-Display und ähnliche Technologien erlauben die Amplifikation der genetischen Information des selektierten Binders nach einer Affinitäts-basierten Selektion. All diese Technologien sind nur für das Anzeigen und der Selektion von Polypeptiden kompatibel, sie können aber nicht für die Selektion und das Aufzeigen von kleinen organischen Molekülen verwendet werden. (Buller *et al.*, 2010). Das Konzept der Genotyp-Phänotyp-Kopplung wurde auf das Konzept der DNA-kodierten Banken übertragen. Dabei ist ein kleines Molekül an einem einzigartigen

DNA-Fragment gekoppelt. Dieses dient als indentifizierbarer Barcode und erlaubt auf diesem Wege auch die Selektion und das Aufzeigen von kleinen organischen Molekülen. Die DNA-Kodierung präsentiert dabei ein attraktives Feld für die Synthese von großen kombinatorischen DNA-Banken, da die DNA-Tagsequenz die Konstruktion der Bank erleichtert und vorgibt. Zudem wird bei den Verfahren der SELEX-Technologie und dem Phage-Display lediglich die DNA-Bank bzw. nur ein Protein auf der Oberfläche der Phage präsentiert (Franzini *et al.*, 2014). Bei den DNA-kodierten Banken ist eine große Molekülvielfalt bzw. sind Mehrfachmolekülkombinationen durch die verschiedensten Synthesstrategien möglich. Auch die Kopplung von chemischen Elementen ist vorstellbar. Die DELs sind zudem einfacher hanbhabbar. Beim Phage-Display ist auf der Oberfläche der Phagen immer eine Antikörperbank (Franzini *et al.*, 2014) dargestellt. Somit stellt die DNA der Phage den Genotyp und der Antikörper den Phänotyp dar. Der Phage selbst ist auch noch vorhanden. Dieses zusätzliche Element wird bei den DNA-kodierten Banken vermieden und macht deren Verwendung leichter.

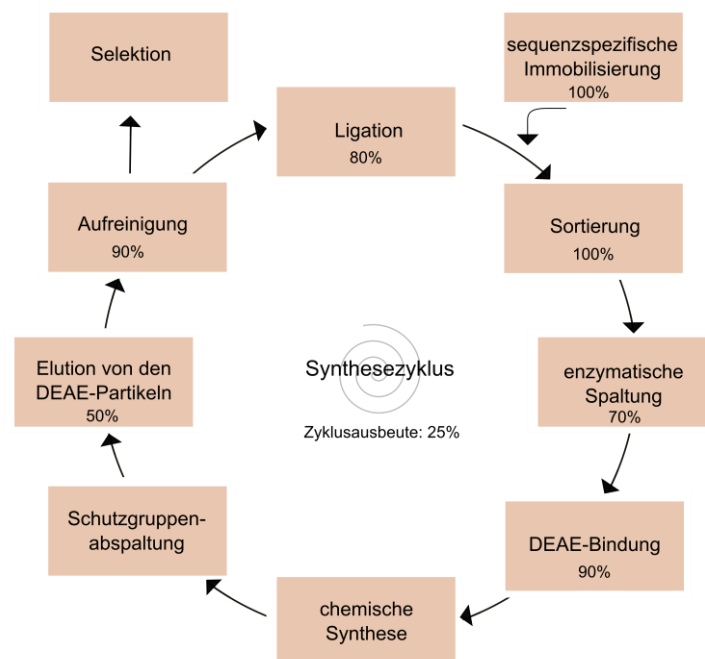
Konventionelle Technologien – neben dem Phage-Display und der SELEX-Technologie – die zur Entdeckung von niedermolekularen Liganden gegen Zielmoleküle eingesetzt werden, verwenden oft Sammlungen von chemischen Komponenten, die einzeln gegen das Zielmolekül selektiert werden. Solch ein Hochdurchsatzscreening ist nicht nur sehr teuer und komplex, sondern resultiert zudem oft in unzureichende Qualitäten (Buller *et al.*, 2010) und schlechten Ergebnissen. Zudem bestehen bei konventionellen Hochdurchsatzscreenings die DNA-Banken aus einer Komplexität von höchstens  $10^8$  Molekülen, oftmals sogar nur aus  $10^3$ - $10^6$ . Im Vergleich, DNA-kodierte Banken können aus einer Komplexität die bis zu  $10^{15}$  Moleküle besitzt, bestehen (Kleiner *et al.*, 2011). DNA-kodierte Substanz-Banken können einfachen Affinitäts-basierten Selektionen durch die Verwendung von Einwegröhrchen und wenigen Mikrogramm des immobilisierten Zielmoleküls im Vergleich zu dem laborintensiven Aufwand des Hochdurchsatzscreening unterzogen werden (Kleiner *et al.*, 2010). In dieser Studie konnte bestätigt und gezeigt werden, dass es sich bei der hier vorgestellten Arbeit um ein schnelles und kostengünstiges Verfahren handelt, um DNA-kodierte Banken mit einer hohen Diversität für Affinitäts-basierte Selektionen herzustellen.

Die Entdeckung von biologisch aktiven kleinen Molekülen steht im Focus der akademischen und industriellen Einrichtungen, um neue Wirkstoffe und Therapien gegen jegliche Krankheiten aufzufinden. Besonders bewußt wird einem die Bedeutsamkeit dieser Arbeit, wenn man betrachtet, dass die Pharmaindustrie jährlich eine Milliarde Dollar (Kleiner *et al.*, 2011) für dieses Forschungs- und Entwicklungsfeld spendet.

## 4.5 SCHLUSSFOLGERUNG UND AUSBLICK

Die Nukleinsäure wird in der Natur als kodierendes Material verwendet. Warum sollte diese Eigenschaft nicht auch technologisch genutzt werden, um DNA-kodierte Substanz-Banken herzustellen. Man unterscheidet bei der Herstellung DNA-kodierter Banken in zwei grundlegende Techniken. Zum einen gibt es protokollierende und zum anderen instruierende Verfahren. Der grundlegende Unterschied liegt im Aufbau der Bank. Wenn man nur protokolliert, also die DNA-Bank aufbaut, dann ist diese nicht reproduzierbar und somit kann keine Selektion gegen ein Zielmolekül durchgeführt werden. Mit der instruierenden Methode sind Selektionen gegen ein Molekül das von Interesse ist möglich. Zudem wird bei diesem Verfahren die Möglichkeit der *in vitro* Evolution berücksichtigt.

Die hier vorgestellte Methode (Neclys-Technologie) stellt ein neues Verfahren für die Konstruktion von DNA-kodierten chemischen Banken, die dem Prinzip der Phänotyp-Genotyp Kopplung folgt, dar. Der Arbeitsablauf dieser Technologie besteht aus vielen einzelnen iterativen Arbeitsschritten, die im Laufe dieser Arbeit optimiert, handhabbar und automatisiert wurden.



**Abb. 45: Schematische Darstellung des Arbeitsablaufes der Neclys-Technologie.** Die DNA-kodierte Substanz-Bank wird durch iterative Reaktionsschritte aufgebaut. Im Anschluss erfolgt die Selektion gegen ein Zielmolekül, das von Interesse ist. Die Effizienzen der einzelnen Schritte wurden im Laufe der Arbeit ermittelt und stetig gesteigert. So kann bei der Immobilisierung und Separation bereits eine 100%-ige Effizienz aufgezeigt werden. Die anderen Schritte liegen bereits bei etwa 80%, so dass von einer Gesamtausbeute von mindestens 25% ausgegangen werden kann.

Mit der Neclys-Methode wird es möglich sein, *in vitro* Evolution mit kombinatorischen Banken mit bis zu  $10^7$  verschiedenen synthetischen Molekülen in kleinen Volumina durchzuführen. Diese Technologie ermöglicht es  $10^7$  verschiedene Varianten herzustellen, da bei dieser Technologie 16 Adapteroligos verwendet werden und fünf Barcodons vorhanden sind ( $16^5$ ). Diese Molekülvarianten liegen anschließend teilweise mehrfach vor, da mehr als  $10^7$  Moleküle rückgewonnen werden. Das heißt, die Diversität liegt nach fünf Syntheserunden bei  $10^7$ , die Anzahl der Moleküle aber um ein vielfaches höher. Somit liegt jedes synthetische Molekül mehrfach vor. Da es mit den Hochdurchsatzsequenzierungssystemen heutzutage möglich ist  $10^7$  Molekülvarianten in einem Lauf abzudecken, kann unter Umständen die Selektion auf eine Runde reduziert werden.

Mit dem vorgestellten Verfahren ist es aber auch möglich, höhere Diversitäten auf einem einfachen Wege zu erreichen. Dass erreicht man, indem die Adapterzahl erhöht und dadurch nicht mit einem Tripletts bzw. zwei Basen und einer Konstanten, sondern mit drei oder vier Basen pro Dekodierung gearbeitet wird. So kann die Diversität schnell und effektiv bei Bedarf erhöht werden. Bei einem Einsatz von drei Basen würde sich nach einer Runde die Diversität von 16 Molekülvarianten auf 64 und nach zwei Runden bereits von 256 auf 4096 Molekülvarianten erhöhen. Dabei ist zu beachten, dass nicht nur die Herstellung von DNA-kodierten Banken – bei denen immer andere Moleküle an die DNA-Bank gekoppelt werden – sondern eine Selektion auf die besten bzw. optimalen Synthesebedingungen möglich ist. Im Gegensatz zu den anderen Arbeitsgruppen konnte eine Selektion auf die besten bzw. optimalen Synthesebedingungen aufgezeigt werden.

Ein weiterer Aspekt, der die Vielfältigkeit dieser Technologie aufzeigt, ist eine zusätzliche Anwendungsmöglichkeit. Mit dieser Methode kann nicht nur die Selektion von niedermolekularen Wirkstoffen erfolgen. Insbesondere durch die hier vorgestellte Sortierungsmethode könnte es auch möglich sein, Stoffe auf diesem Wege zu transportieren. Somit stellt das Verfahren auch eine Transportmöglichkeit von niedermolekularen Stoffen dar. Der zielgerichtete Wirkstofftransport von niedermolekularen Stoffen im Gewebe ist ein wichtiger Faktor in der pharmazeutischen Industrie. Zu den bekannten Systemen, wie dem Antikörpertransport, ist es wichtig neue Trägersysteme zu entwickeln, da die bekannten Systeme Limitationen aufzeigen und zudem sehr teuer in der Herstellung sind (Krall *et al.*, 2013)

Die Neclys-Technologie stellt eine neue Plattform dar, mit der die verschiedensten Methoden durchgeführt und weiterentwickelt werden können. Zum einen können DNA-kodierte Substanz-Banken auf chemischem (Peptidbindung, Harnstoffsynthese, Click-Chemie, Diels-Alder Reaktion etc.) aber auch auf enzymatischem Wege hergestellt werden, wodurch die Selektion von niedermolekularen Stoffen erfolgen kann. Zum anderen können optimale Reaktionsbedingungen für die Synthese von Peptiden ermittelt bzw. selektiert werden. Dabei könnte jedes Barcodon während jeder Zyklusrunde für bestimmte Reaktionspartner wie Temperatur, Druck, pH, Lösungsmittel und dergleichen kodieren.

Schlussendlich wäre sogar die Selektion gegen neue Farbstoffe, Medikamenten und ähnlichen Stoffen vorstellbar. Der Arbeitsprozess ist durch die frei programmierbare Plattform des KingFisher Flex Roboter automatisierbar und mit der hier vorgestellten Methode wäre es sicherlich möglich viele neue Medikamente oder Liganden die von Interesse sind, für viele interessante Anwendungen in der Wissenschaft und Pharmaindustrie zu entdecken.

## 5 ZUSAMMENFASSUNG

Die Identifizierung kleiner Moleküle, wie chemische Verbindungen oder Antikörper, welche spezifische Bindungsaffinitäten für Zielproteine aufzeigen, bleibt eine der wichtigsten Herausforderungen in der biologischen, chemischen und medizinischen Forschung. Besonders für die Entdeckung neuer Arzneimittel stellt es ein zentrales Thema dar. Die sogenannten DNA-kodierten Substanz-Banken stellen eine neue Technik für die Synthese und die Selektion einer großen Anzahl von kleinen Liganden, die spezifisch an relevante Proteine binden, dar. Nach einer Affinitäts-basierten Selektion können DNA-kodierte Moleküle mithilfe des eindeutigen DNA-Barcodes identifiziert und analysiert werden.

In dieser Arbeit wird die Technik einer neuartigen Synthesestrategie für DNA-kodierte Substanz-Banken beschrieben. Diese wurde auf dem Prinzip der Festphasenpeptidsynthese generiert und basiert auf dem Konzept der sequenzabhängigen Sortierung und anschließenden Synthese (DNA-routed synthesis). Dabei wird die Nukleinsäure (Genotyp) benutzt, um die Synthese von kleinen Molekülen (Phänotyp) zu erfassen.

Die hier etablierte Methode basiert auf einem iterativen Ablauf, der sich aus den Hauptkomponenten der Ligations-, Separations, Spaltungs und Synthesereaktion zusammensetzt. Dabei konnte das Verfahren auf den Magnetpartikelprozessor KingFisher Flex von Thermo Scientific übertragen werden, so dass eine halbautomatisierte Technologie entstand. Bereits nach zwei Syntheserunden können 256 verschiedene Molekülvarianten aufgrund der Verwendung von 16 unterschiedlichen Adaptermolekülen, generiert werden. Es konnte mittels Sanger-Sequenzierung gezeigt werden, dass die Sortierung bzw. Separation der ligierten DNA-Bank durch eine sehr kurze einmalige Tagsequenz mittels Hybridisierung möglich und vor allem spezifisch ist.

Nach zwei vollständigen Syntheserunden, bei denen die Aminosäuren Glycin und das Vitamin d-Biotin unter 16 verschiedenen Synthesebedingungen an die DNA-Bank synthetisiert wurden, konnten Selektionsexperimente gegen das Zielmolekül Streptavidin durchgeführt werden.

Durch die Anwendung der Hochdurchsatzsequenzierung zur Analyse der DNA-kodierten Substanz-Bank konnten die optimalen Synthesebedingungen ermittelt werden. Die Dekodierung des Barcodes gab dabei Aufschlüsse über die verwendeten Synthesebedingungen. Als besonders effizientes Kopplungsreagenz konnte DMT-MM und als Lösungsmittel DMF mittels 454-Pyrosequenzierung ermittelt werden. Dieses Ergebnis stärkt die Tatsache, dass Hochdurchsatzsequenzierung der neuen Generation für die Analyse von DNA-kodierten Substanz-Bank geeignet sind. Der hohe Durchsatz dieser Sequenzierungsart gewährleistet die Analyse von großen Datenmengen DNA-kodierter Substanz-Banken und verringert dabei die Dekodierungskosten.

Mit der hier dargestellten Methode konnte gezeigt werden, dass es möglich ist, die neuartigen DNA-kodierten Wirkstoffbanken zur Selektion von optimalen Synthesebedingungen zu verwenden als auch schnell und kostengünstig diese aufzubauen, um dann als Liganden gegen pharmazeutisch und medizinisch bedeutsame Zielmoleküle eingesetzt zu werden.



## 6 SUMMARY

The identification of small molecules, such as chemical compounds that with high specificity and affinity for given target proteins remains a key challenge in the biological, chemical and medical research. There is a need to discover new drugs, especially with novel properties without limiting constraints such as rational design. The so-called DNA-encoded libraries provide a new approach for the synthesis and the selection of a large number of small ligands that bind specifically to target proteins of interest. DNA-encoded molecules can be identified and analyzed using unique DNA barcodes following an affinity-based selection.

In this work, the technique of a novel synthetic strategy for DNA-encoded libraries is described. This was based on the principle of solid-phase peptide synthesis and on the concept of DNA-routed synthesis. The nucleic acid (genotype) is used to detect the synthesis of small molecules (phenotype) following the natural example of protein translation.

The established method is based on an iterative process, composed of the main steps of ligation, separation, cleavage and synthesis reaction. The method was adapted to the magnetic particle processor KingFisher Flex from Thermo Scientific, resulting in a semi-automated procedure. Already after two rounds of synthesis 256 different molecule variants can be generated by using 16 different adapter molecules. Shown by Sanger sequencing, specific sorting and separation of the ligated DNA library by hybridization is possible by using a very short and unique tag sequence.

After two full synthesis rounds in which the amino acids glycine and the vitamin d-biotin were synthesized to the encoding DNA library using 16 different synthesis conditions, selection experiments against the target molecule streptavidin were performed.

Optimal synthesis conditions were determined by using high-throughput sequencing for the analysis of the enriched DNA-encoded library. The decoding of the barcode elucidated the respective synthesis conditions. Through the use of the 454 pyrosequencing system a particularly efficient coupling reagent DMT-MM in combination with the solvent DMF could be determined. This result reinforces the fact that the new generation of massive parallel sequencing is highly suitable for analysis of DNA-encoded libraries.

The high throughput of these sequencing systems ensures the analysis of large data sets of DNA-encoded chemical libraries and thereby reduces the costs of decoding.

With the described method it was demonstrated that it is possible to use the novel DNA-encoded chemical libraries for the selection of optimal synthesis conditions. These can be constructed quickly and inexpensive and applied in high throughput screening processes for ligand discovery of pharmaceutically and medically important target molecules.

## 7 LITERATURVERZEICHNIS

- Altmann, K.-H., & Mutter, M. (1993). Die chemische Synthese von Peptiden und Proteinen. *Chemie in Unserer Zeit* 27(6), 274–286.
- Baran, N., Pucshansky, L., Marco, Y., Benjamin, S., & Manor, H. (1997). The SV40 large T-antigen helicase can unwind four stranded DNA structures linked by G-quartets. *Nucleic Acids Research* 25(2), 297–303.
- Bayer, E. (1991). Auf dem Weg zur chemischen Synthese von Proteinen. *Angewandte Chemie* 103(2), 117–133.
- Beaucage, S. L., & Iyer, R. P. (1992). Advances in the Synthesis of Oligonucleotides by the Phosphoramidite Approach. *Tetrahedron* 48(12), 2223–2311.
- Berglund, E. C., Kiialainen, A., & Syvänen, A.-C. (2011). Next-generation sequencing technologies and applications for human genetic history and forensics. *Investigative Genetics* 2, 23.
- Bergmann, M., & Zervas, L. (1932). Über ein allgemeines Verfahren der Peptid-Synthese. *Berichte Der Deutschen Chemischen Gesellschaft (A and B Series)* 65(7), 1192–1201.
- Beyer, H. & Walter, W. (1988). Lehrbuch der organischen Chemie. S. Hirzel Verlag. Stuttgart 21, 231–232.
- Bogenhagen, D. F., & Pinz, K. G. (1998). The action of DNA ligase at abasic sites in DNA. *The Journal of Biological Chemistry* 273(14), 7888–7893.
- Boiteux, S., & Guillet, M. (2004). Abasic sites in DNA: repair and biological consequences in *Saccharomyces cerevisiae*. *DNA Repair* 3(1), 1–12.
- Bonetta, L. (2006). Genome sequencing in the fast lane. *Nature Methods* 3(2), 141–147.
- Brenner, S. & Lerner, R. A. (1992). Encoded combinatorial chemistry. *Proceedings of the National Academy of Sciences of the United States of America* 89(12), 5381–5383.
- Brody, J. R., & Kern, S. E. (2004). History and principles of conductive media for standard DNA electrophoresis. *Analytical Biochemistry* 333(1), 1–13.
- Buller, F., Mannocci, L., Scheuermann, J., & Neri, D. (2010). Drug discovery with DNA-encoded chemical libraries. *Bioconjugate Chemistry* 21(9), 1571–1580.
- Buller, F., Mannocci, L., Zhang, Y., Dumelin, C. E., Scheuermann, J., & Neri, D. (2008). Design and synthesis of a novel DNA-encoded chemical library using Diels-Alder cycloadditions. *Bioorganic & Medicinal Chemistry Letters* 18(22), 5926–5931.

- Buller, F., Steiner, M., Frey, K., Mircsof, D., Scheuermann, J., Kalisch, M., ... Neri, D. (2011). Selection of Carbonic Anhydrase IX Inhibitors from One Million DNA-Encoded Compounds. *ACS Chemical Biology* 6(4), 336–344.
- Buller, F., Zhang, Y., Scheuermann, J., Schäfer, J., Bühlmann, P., & Neri, D. (2009). Discovery of TNF inhibitors from a DNA-encoded chemical library based on diels-alder cycloaddition. *Chemistry & Biology* 16(10), 1075–1086.
- Chalet, L., & Wolf, F. J. (1964). The properties of streptavidin, a biotin-binding protein produced by Streptomyces. *Archives of Biochemistry and Biophysics* 106, 1–5.
- Chrambach, A., & Rodbard, D. (1971). Polyacrylamide Gel Electrophoresis. *Science* 172(3982), 440–451.
- Clark, M. A. (2010). Selecting chemicals: the emerging utility of DNA-encoded libraries. *Current Opinion in Chemical Biology* 14(3), 396–403.
- Clark, M. A., Acharya, R. A., Arico-Muendel, C. C., Belyanskaya, S. L., Benjamin, D. R., Carlson, N. R., ... Morgan, B. A. (2009). Design, synthesis and selection of DNA-encoded small-molecule libraries. *Nature Chemical Biology* 5(9), 647–654.
- Cox, M., & Nelson, D. (2009). Lehninger Biochemie. *Springer-Verlag, Berlin, Heidelberg* 4, 130–133.
- Deugau, K. V., & van de Sande, J. H. (1978). T4 polynucleotide ligase catalyzed joining of short synthetic DNA duplexes at base-paired ends. *Biochemistry* 17(4), 723–729.
- Dumelin, C. E., Scheuermann, J., Melkko, S., & Neri, D. (2006). Selection of streptavidin binders from a DNA-encoded chemical library. *Bioconjugate Chemistry* 17(2), 366–370.
- Dumelin, C. E., Trüssel, S., Buller, F., Trachsel, E., Bootz, F., Zhang, Y., ... Neri, D. (2008). A portable albumin binder from a DNA-encoded chemical library. *Angewandte Chemie (International Ed. in English)* 47(17), 3196–3201.
- Ellington, A. D. & Szostak, J. W. (1990). In vitro selection of RNA molecules that bind specific ligands. *Nature* 346(6287), 818–822.
- Eulberg, D., Buchner, K., Maasch, C., & Klussmann, S. (2005). Development of an automated in vitro selection protocol to obtain RNA-based aptamers: identification of a biostable substance P antagonist. *Nucleic Acids Research* 33(4), e45.
- Farokhzad, O. C., & Langer, R. (2006). Nanomedicine: developing smarter therapeutic and diagnostic modalities. *Advanced Drug Delivery Reviews* 58(14), 1456–1459.
- Fox, S., Filichkin, S., & Mockler, T. C. (2009). Applications of ultra-high-throughput sequencing. *Methods in Molecular Biology (Clifton, N.J.)* 553, 79–108.

- Franzini, R. M., Neri, D., & Scheuermann, J. (2014). DNA-Encoded Chemical Libraries: Advancing beyond Conventional Small-Molecule Libraries. *Accounts of Chemical Research* 47(4), 1247-1255.
- Gartner, Z. J. & Liu, D. R. (2001). The generality of DNA-templated synthesis as a basis for evolving non-natural small molecules. *Journal of the American Chemical Society* 123(28), 6961–6963.
- Glökler, J., Schütze, T., & Konthur, Z. (2010). Automation in the High-throughput Selection of Random Combinatorial Libraries—Different Approaches for Select Applications. *Molecules* 15(4), 2478–2490.
- Halpin, D. R. & Harbury, P. B. (2004a). DNA display I. Sequence-encoded routing of DNA populations. *PLoS Biology* 2(7), E173.
- Halpin, D. R. & Harbury, P. B. (2004b). DNA display II. Genetic manipulation of combinatorial chemistry libraries for small-molecule evolution. *PLoS Biology* 2(7), E174.
- Halpin, D. R., Lee, J. A., Wrenn, S. J., & Harbury, P. B. (2004). DNA display III. Solid-phase organic synthesis on unprotected DNA. *PLoS Biology* 2(7), E175.
- Hanes, J. & Plückthun, A. (1997). In vitro selection and evolution of functional proteins by using ribosome display. *Proceedings of the National Academy of Sciences of the United States of America* 94(10), 4937–4942.
- Hansen, M. H., Blakskjaer, P., Petersen, L. K., Hansen, T. H., Højfeldt, J. W., Gothelf, K. V. & Hansen, N. J. V. (2009). A yoctoliter-scale DNA reactor for small-molecule evolution. *Journal of the American Chemical Society* 131(3), 1322–1327.
- Harrington, C. T., Lin, E. I., Olson, M. T., & Eshleman, J. R. (2013). Fundamentals of pyrosequencing. *Archives of Pathology & Laboratory Medicine* 137(9), 1296–1303.
- Harvey, C. L., Gabriel, T. F., Wilt, E. M., & Richardson, C. C. (1971). Enzymatic Breakage and Joining of Deoxyribonucleic Acid IX. SYNTHESIS AND PROPERTIES OF THE DEOXYRIBONUCLEIC ACID ADENYLATE IN THE PHAGE T4 LIGASE REACTION. *Journal of Biological Chemistry* 246(14), 4523–4530.
- Hauk, A. (2013). Quantifizierung von DNA durch Absorptionsmessung. *Biologie in unserer Zeit* 43(5), 278–278.
- Heitner, T. R., & Hansen, N. J. V. (2009). Streamlining hit discovery and optimization with a yoctoliter scale DNA reactor. *Expert Opinion on Drug Discovery* 4(11), 1201–1213.
- Henning, W. (2002). Genetik. *Springer-Verlag. Berlin, Heidelberg* 3, 285–289.
- Higgins, N.P., & Cozzarelli, N. R. (1978). DNA-joining enzymes: a review. *Methods in Enzymology* 68, 50–71.

- Hosfield, D. J., Guan, Y., Haas, B. J., Cunningham, R. P., & Tainer, J. A. (1999). Structure of the DNA repair enzyme endonuclease IV and its DNA complex: double-nucleotide flipping at abasic sites and three-metal-ion catalysis. *Cell* 98(3), 397–408.
- International Human Genome Sequencing Consortium. (2004). Finishing the euchromatic sequence of the human genome. *Nature* 431(7011), 931–945.
- Jakubke, HD. & Karcher, R. (1998). Lexikon der Chemie. *Spektrum Akademischer Verlag, Berlin, Heidelberg*, 231–232.
- Jeong, Y.-J., Park, K., & Kim, D.-E. (2009). Isothermal DNA amplification in vitro: the helicase-dependent amplification system. *Cellular and Molecular Life Sciences: CMLS* 66(20), 3325–3336.
- Joyce, G. F. (2007). Forty years of in vitro evolution. *Angewandte Chemie (International Ed. in English)* 46(34), 6420–6436.
- Kessler, H. (1993). Peptide – Schlüsselverbindungen für einen neuen Weg zur Entwicklung von pharmazeutischen Wirkstoffen. *Angewandte Chemie*, 105(4), 572–573.
- Kleiner, R. E., Dumelin, C. E. & Liu, D. R. (2011). Small-molecule discovery from DNA-encoded chemical libraries. *Chemical Society Reviews* 40(12), 5707-5017.
- Kleiner, R. E., Dumelin, C. E., Tiu, G. C., Sakurai, K. & Liu, D. R. (2010). In vitro selection of a DNA-templated small-molecule library reveals a class of macrocyclic kinase inhibitors. *Journal of the American Chemical Society* 132(33), 11779–11791.
- Krall, N., Scheuermann, J., & Neri, D. (2013). Entwicklung zielgerichteter niedermolekularer zytotoxischer Wirkstoffverbindungen mit DNA-codierten chemischen Bibliotheken. *Angewandte Chemie* 125(5), 1424–1443.
- Kuhn, H., & Frank-Kamenetskii, M. D. (2005). Template-independent ligation of single-stranded DNA by T4 DNA ligase. *The FEBS Journal* 272(23), 5991–6000.
- Kurzer, F., & Douraghi-Zadeh, K. (1967). Advances in the chemistry of carbodiimides. *Chemical Reviews* 67(2), 107–152.
- Lee, H. & Tang, H. (2012). Next-generation sequencing technologies and fragment assembly algorithms. *Methods in Molecular Biology (Clifton, N.J.)* 855, 155–174.
- Lehman, I. R. (1974). DNA ligase: structure, mechanism, and function. *Science (New York, N.Y.)* 186(4166), 790–797.
- Leimbacher, M., Zhang, Y., Mannocci, L., Stravs, M., Geppert, T., Scheuermann, J., ... Neri, D. (2012). Discovery of small-molecule interleukin-2 inhibitors from a DNA-encoded chemical library. *Chemistry (Weinheim an Der Bergstrasse, Germany)* 18(25), 7729–7737.

- Li, X., & Liu, D. R. (2004). DNA-templated organic synthesis: nature's strategy for controlling chemical reactivity applied to synthetic molecules. *Angewandte Chemie (International Ed. in English)* 43(37), 4848–4870.
- Li, M., Völker, J., Breslauer, K. J., & Wilson, D. M. (2014). APE1 incision activity at abasic sites in tandem repeat sequences. *Journal of Molecular Biology* 426(11), 2183–2198
- Li, Y., Zhao, P., Zhang, M., Zhao, X., & Li, X. (2013). Multistep DNA-Templated Synthesis Using a Universal Template. *Journal of the American Chemical Society* 135(47), 17727–17730.
- Litovchick, A., Clark, M. A., & Keefe, A. D. (2014). Universal strategies for the DNA-encoding of libraries of small molecules using the chemical ligation of oligonucleotide tags. *Artificial DNA, PNA & XNA* 5(1).
- Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R. & Law, M. (2012). Comparison of next-generation sequencing systems. *Journal of Biomedicine & Biotechnology*, 2012, 251364.
- Lohman, G. J. S., Tabor, S., & Nichols, N. M. (2001). DNA Ligases. In *Current Protocols in Molecular Biology* 94 (3), 3.14.1–3.14.7.
- Mannocci, L., Leimbacher, M., Wichert, M., Scheuermann, J. & Neri, D. (2011). 20 years of DNA-encoded chemical libraries. *Chemical Communications (Cambridge, England)* 47(48), 12747–12753.
- Mardis, E. R. (2008a). Next-generation DNA sequencing methods. *Annual Review of Genomics and Human Genetics* 9, 387–402.
- Mardis, E. R. (2008b). The impact of next-generation sequencing technology on genetics. *Trends in Genetics: TIG* 24(3), 133–141.
- Maxam, A. M., & Gilbert, W. (1977). A new method for sequencing DNA. *Proceedings of the National Academy of Sciences of the United States of America* 74(2), 560–564.
- Melkko, S., Dumelin, C. E., Scheuermann, J. & Neri, D. (2007). Lead discovery by DNA-encoded chemical libraries. *Drug Discovery Today* 12(11-12), 465–471.
- Melkko, S., Mannocci, L., Dumelin, C. E., Villa, A., Som mavilla, R., Zhang, Y., ... Neri, D. (2010). Isolation of a small-molecule inhibitor of the antiapoptotic protein Bcl-xL from a DNA-encoded chemical library. *ChemMedChem* 5(4), 584–590.
- Melkko, S., Scheuermann, J., Dumelin, C. E., & Neri, D. (2004). Encoded self-assembling chemical libraries. *Nature Biotechnology* 22(5), 568–574.
- Merrifield, R. B. (1963). Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *Journal of the American Chemical Society* 85(14), 2149–2154.
- Metzker, M. L. (2010). Sequencing technologies - the next generation. *Nature Reviews. Genetics* 11(1), 31–46.

- Mills, D. R., Peterson, R. L. & Spiegelman, S. (1967). An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. *Proceedings of the National Academy of Sciences of the United States of America* 58(1), 217–224.
- Mori, A., Deola, S., Xumerle, L., Mijatovic, V., Malerba, G., & Monsurrò, V. (2013). Next generation sequencing: new tools in immunology and hematology. *Blood Research* 48(4), 242–249.
- Morozova, O., Hirst, M., & Marra, M. A. (2009). Applications of new sequencing technologies for transcriptome analysis. *Annual Review of Genomics and Human Genetics* 10, 135–151.
- Müllhardt, C. (2013). Der Experimentator: Molekularbiologie/Genomics. *Springer-Verlag, Berlin, Heidelberg* 7, 58–70.
- Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. & Erlich, H. (1986). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harbor Symposia on Quantitative Biology* 51 Pt 1, 263–273.
- Nakajima, N., & Ikada, Y. (1995). Mechanism of Amide Formation by Carbodiimide for Bioconjugation in Aqueous Media. *Bioconjugate Chemistry* 6(1), 123–130.
- Nemoto, N., Miyamoto-Sato, E., Husimi, Y. & Yanagawa, H. (1997). In vitro virus: bonding of mRNA bearing puromycin at the 3'-terminal end to the C-terminal end of its encoded protein on the ribosome in vitro. *FEBS Letters* 414(2), 405–408.
- Nichols, N. M., & Yue, D. (2008). Ribonucleases. *Current Protocols in Molecular Biology / Edited by Frederick M. Ausubel ... [et Al.]*, Chapter 3, Unit3.13.
- Olivera, B. M., Hall, Z. W., & Lehman, I. R. (1968). Enzymatic joining of polynucleotides, V. A DNA-adenylate intermediate in the polynucleotide-joining reaction. *Proceedings of the National Academy of Sciences of the United States of America* 61(1), 237–244.
- Paschke, M. (2006). Phage display systems and their applications. *Applied Microbiology and Biotechnology* 70(1), 2–11.
- Puri, R. V., Singh, N., Gupta, R. K., & Tyagi, A. K. (2013). Endonuclease IV Is the major apurinic/apyrimidinic endonuclease in Mycobacterium tuberculosis and is important for protection against oxidative damage. *PloS One* 8(8), e71535.
- Richard W, R. (1999). Totally in vitro protein selection using mRNA-protein fusions and ribosome display. *Current Opinion in Chemical Biology* 3(3), 268–273.
- Roberts, R. W. & Szostak, J. W. (1997). RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proceedings of the National Academy of Sciences of the United States of America* 94(23), 12297–12302.
- Robertson, D. L. & Joyce, G. F. (1990). Selection in vitro of an RNA enzyme that specifically cleaves single-stranded DNA. *Nature* 344(6265), 467–468.



- Ronaghi, M. (2001). Pyrosequencing sheds light on DNA sequencing. *Genome Research* 11(1), 3–11.
- Ruiz-Hernández, E., Baeza, A., & Vallet-Regí, M. (2011). Smart Drug Delivery through DNA/Magnetic Nanoparticle Gates. *ACS Nano* 5(2), 1259–1266.
- Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., ... Erlich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science (New York, N.Y.)* 239(4839), 487–491.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* 74(12), 5463–5467.
- Schadt, E. E., Turner, S. & Kasarskis, A. (2010). A window into third-generation sequencing. *Human Molecular Genetics* 19(R2), R227–240.
- Scheuermann, J., Dumelin, C. E., Melkko, S., & Neri, D. (2006). DNA-encoded chemical libraries. *Journal of Biotechnology* 126(4), 568–581.
- Scheuermann, J., Dumelin, C. E., Melkko, S., Zhang, Y., Mannocci, L., Jaggi, M., Sobek, J. & Neri, D. (2008). DNA-encoded chemical libraries for the discovery of MMP-3 inhibitors. *Bioconjugate Chemistry* 19(3), 778–785.
- Scheuermann, J. & Neri, D. (2010). DNA-encoded chemical libraries: a tool for drug discovery and for chemical biology. *Chembiochem: A European Journal of Chemical Biology* 11(7), 931–937.
- Schuster, S. C. (2008). Next-generation sequencing transforms today's biology. *Nature Methods* 5(1), 16–18.
- Schütze, T., Rubelt, F., Repkow, J., Greiner, N., Erdmann, V. A., Lehrach, H., Konthur, Z. & Glökler, J. (2011). A streamlined protocol for emulsion polymerase chain reaction and subsequent purification. *Analytical Biochemistry* 410(1), 155–157.
- Schütze, T., Wilhelm, B., Greiner, N., Braun, HJ., Peter, F., Mörl, M., ... Glökler, J. (2011). Probing the Selex Process with Next-Generation Sequencing. *PLoS ONE* 6(12), e29604
- Seyffert, W. (2003). Lehrbuch der Genetik. *Spektrum Akademischer Verlag. Heidelberg, Berlin* 2, 1070–1075.
- Shin, D.-S., Kim, D.-H., Chung, W.-J., & Lee, Y.-S. (2005). Combinatorial solid phase peptide synthesis and bioassays. *Journal of Biochemistry and Molecular Biology*, 38(5), 517–525.
- Shuman, S. (2009). DNA ligases: progress and prospects. *The Journal of Biological Chemistry* 284(26), 17365–17369.







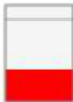




- Sinz, A. (2003). Chemical cross-linking and mass spectrometry for mapping three-dimensional structures of proteins and protein complexes. *Journal of Mass Spectrometry* 38(12), 1225–1237.
- Smith, G. P. (1985). Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science (New York, N.Y.)* 228(4705), 1315–1317.
- Sprinz, K. I., Tagore, D. M. & Hamilton, A. D. (2005). Self-assembly of bivalent protein-binding agents based on oligonucleotide-linked organic fragments. *Bioorganic & Medicinal Chemistry Letters* 15(17), 3908–3911.
- Stangier, K., Bauser, C., Regenbogen, J. (2007). Praxisbericht: Next Generation DNA-Sequenzierung. *Laborwelt* 8(3), 13–17.
- Swanson, A., Ramos, E., & Snyder, H. (2014). Next Generation Sequencing is the Impetus for the Next Generation of Laboratory-Based Genetic Counselors. *Journal of Genetic Counseling*.
- Takeuchi, M., Lillis, R., Demple, B., & Takeshita, M. (1994). Interactions of Escherichia coli endonuclease IV and exonuclease III with abasic sites in DNA. *The Journal of Biological Chemistry* 269(34), 21907–21914.
- Timmermann, B., Albrecht, M.W., Amstislavskiy, V., Borodina, T., Dahl, A., Herwig, R., ... Lehrach, H. (2009). Systematische Analyse der genetischen Variabilität des Menschen. *Laborwelt* 10(3), 8–10.
- Tuerk, C. & Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science (New York, N.Y.)* 249(4968), 505–510.
- Waggoner, A. S., Ernst, L. A., Chen, C.-H., & Rechtenwald, D. J. (1993). Pe-Cy5. *Annals of the New York Academy of Sciences* 677(1), 185–193.
- Wang, J., Kawde, A. N., Erdem, A., & Salazar, M. (2001). Magnetic bead-based label-free electrochemical detection of DNA hybridization. *The Analyst* 126(11), 2020–2024.
- Weber, P. C., Ohlendorf, D. H., Wendoloski, J. J., & Salemme, F. R. (1989). Structural origins of high-affinity biotin binding to streptavidin. *Science* 243(4887), 85–88.
- Weisinger, R. M., Marinelli, R. J., Wrenn, S. J. & Harbury, P. B. (2012). Mesofluidic Devices for DNA-Programmed Combinatorial Chemistry. *PLoS ONE* 7(3), e32299.
- Weiss, B., & Richardson, C. C. (1967). Enzymatic breakage and joining of deoxyribonucleic acid, I. Repair of single-strand breaks in DNA by an enzyme system from Escherichia coli infected with T4 bacteriophage. *Proceedings of the National Academy of Sciences of the United States of America* 57(4), 1021–1028.
- Wiedmann, M., Wilson, W. J., Czajka, J., Luo, J., Barany, F., & Batt, C. A. (1994). Ligase chain reaction (LCR)--overview and applications. *PCR Methods and Applications* 3(4), 51–64.

- Wochner, A., Cech, B., Menger, M., Erdmann, V. A., & Glökler, J. (2007). Semi-automated selection of DNA aptamers using magnetic particle handling. *BioTechniques* 43(3), 344, 346, 348 passim.
- Woollard, P. M., Mehta, N. A. L., Vamathevan, J. J., Van Horn, S., Bonde, B. K., & Dow, D. J. (2011). The application of next-generation sequencing technologies to drug discovery and development. *Drug Discovery Today*, 16(11-12), 512–519.
- Wollenberger, U., Renneberg, R., Bier, F. F. & Scheller, F. W. (2003). Analytische Biochemie: Eine praktische Einführung in das Messen mit Biomolekülen. *Wiley-VCH Verlag, Weinheim*, 222p.
- Wu, D. Y., & Wallace, R. B. (1989). Specificity of the nick-closing activity of bacteriophage T4 DNA ligase. *Gene* 76(2), 245–254.
- Zhu, Z., Zhang, Y., Ji, Z., He, S., & Yang, X. (2013). High-throughput DNA sequence data compression. *Briefings in Bioinformatics*.
- Zimmermann, H. (1980). Bindung von Ethidiumbromid an DNA. *Nachrichten Aus Chemie, Technik Und Laboratorium* 28(11), 797–800.

## 6 ANHANG

### 6.1 KINGFISHER FLEX PROTOKOLLE

#### Separation

 <p><b>Separation</b> Instrument type Description</p>	KingFisher Flex	
 <p><b>Tip1</b> Tip Pick-Up plate Leave plate</p>	96er Deepwell-Platte Tip-Platte Tip-Platte	 <empty>
 <p><b>CollectBeads</b> Plate Collect count Collect time [s]</p>	Bindungsplatte 5 30	 <p>Reagents: 1x TE-Puffer Volume [µl]: 50</p>
 <p><b>Hybridization</b> Plate Beginning of step: Precollect Release beads Mixing/heating parameters: Heating temperature [°C] Preheat Mixing time [hh:mm:ss] Mixing speed End of step: Postmix Collect beads</p>	Hybridisierungsplatte No No 46 Yes 01:00:00 Slow No No	 <p>Reagents: 1,5 M NaCl, 150 mM Na-Citrat, 0,1% Tween Volume [ml]: 10</p>
 <p><b>Wash</b> Plate Beginning of step: Precollect Release beads Mixing/heating parameters: Heating during mixing Mixing time [hh:mm:ss] Mixing speed End of step: Postmix Collect beads</p>	Waschplatte No No No 00:01:00 Slow No No No	 <p>Reagents: 0,1 M MES, pH 5 Volume [µl]: 100</p>
 <p><b>Elution</b> Plate Beginning of step: Precollect Release beads Mixing/heating parameters: Heating during mixing Mixing time [hh:mm:ss] Mixing speed End of step: Postmix Collect beads</p>	Elutionsplatte No Yes No 00:02:00 Slow No No No	 <p>Reagents: 0,1 M MES, pH 5 Volume [µl]: 100</p>

## Synthese



### Synthese

Instrument type  
Description

KingFisher Flex



#### Tip1

Tip  
Pick-Up plate  
Leave plate

96erDeepwell-Platte  
Tip-Platte  
Tip-Platte



<empty>



#### CollectBeads

Plate  
Collect count  
Collect time [s]

Bindungsplatte  
5  
30



Reagents: 20mMTris (pH 8)  
Volume [μl]: 50



#### Synt

Plate  
Beginning of step:  
Precollect  
Release beads  
Mixing/heating parameters:  
Heating during mixing  
Mixing time [hh:mm:ss]  
Mixing speed  
End of step:  
Postmix  
Collect count  
Collect time [s]

Syntheseplatte  
No  
Yes  
No  
00:30:00  
Slow  
No  
3  
1



Reagents: Lösungsmittel  
Volume [μl]: 100



#### Wash1

Plate  
Beginning of step:  
Precollect  
Release beads  
Mixing/heating parameters:  
Heating during mixing  
Mixing time [hh:mm:ss]  
Mixing speed  
End of step:  
Postmix  
Collect beads

Waschplatte\_1  
No  
No  
No  
00:30:00  
Slow  
No  
No  
No



Reagents: 20mMTris (pH 8)  
Volume [μl]: 200


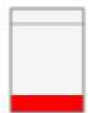
 **Wash2**

Plate	Waschplatte_2		Reagents: 20mMTris (pH 8) Volume [µl]: 200
Beginning of step:			
Precollect	No		
Release beads	No		
Mixing/heating parameters:			
Heating during mixing	No		
Mixing time [hh:mm:ss]	00:30:00		
Mixing speed	Slow		
End of step:			
Postmix	No		
Collect beads	No		


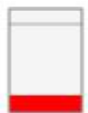


 **Wash3**


Plate	Waschplatte_3		Reagents: 20mMTris (pH 8) Volume [µl]: 200
Beginning of step:			
Precollect	No		
Release beads	No		
Mixing/heating parameters:			
Heating during mixing	No		
Mixing time [hh:mm:ss]	00:30:00		
Mixing speed	Slow		
End of step:			
Postmix	No		
Collect beads	No		

## Schutzgruppenabspaltung und Elution

 **Deprotection and Elution**

Instrument type	KingFisher Flex
Description	

 **Tip1**

Tip	96er Deepwell-Platte		<empty>
Pick-Up plate	Tip-Platte		
Leave plate	Tip-Platte		













 **Depr1**

Plate	Deprotection1		Reagents: 20%Piperidin in DMF Volume [µl]: 100
Beginning of step:			
Precollect	No		
Release beads	Yes		
Mixing/heating parameters:			
Heating during mixing	No		
Mixing time [hh:mm:ss]	00:03:00		
Mixing speed	Slow		
End of step:			
Postmix	No		
Collect count	3		
Collect time [s]	1		

	<b>Depr2</b>	Deprotection2		Reagents: 20%Piperidin in DMF
	Plate			Volume [μl]: 100
	Beginning of step:	No		
	Precollect	No		
	Release beads	Yes		
	Mixing/heating parameters:			
	Heating during mixing	No		
	Mixing time [hh:mm:ss]	00:17:00		
	Mixing speed	Slow		
	End of step:			
	Postmix	No		
Collect count	3			
Collect time [s]	1			
	<b>Wash1</b>	Waschplatte_1		Reagents: 20mM Tris (pH8)
	Plate			Volume [μl]: 100
	Beginning of step:	No		
	Precollect	No		
	Release beads	No		
	Mixing/heating parameters:			
	Heating during mixing	No		
	Mixing time [hh:mm:ss]	00:30:00		
	Mixing speed	Slow		
	End of step:			
	Postmix	No		
Collect beads	No			
	<b>Wash2</b>	Waschplatte_2		Reagents: 20mM Tris (pH8)
	Plate			Volume [μl]: 100
	Beginning of step:	No		
	Precollect	No		
	Release beads	No		
	Mixing/heating parameters:			
	Heating during mixing	No		
	Mixing time [hh:mm:ss]	00:30:00		
	Mixing speed	Slow		
	End of step:			
	Postmix	No		
Collect beads	No			
	<b>Elution</b>	Elutionsplatte		Reagents: 20mM Tris (pH8), 1M NaCl
	Plate			Volume [μl]: 50
	Beginning of step:	No		
	Precollect	No		
	Release beads	Yes		
	Mixing/heating parameters:			
	Heating during mixing	No		
	Mixing time [hh:mm:ss]	00:30:00		
	Mixing speed	Slow		
	End of step:			
	Postmix	No		
Collect count	3			
Collect time [s]	1			

## Elution von DEAE-Mikropartikeln

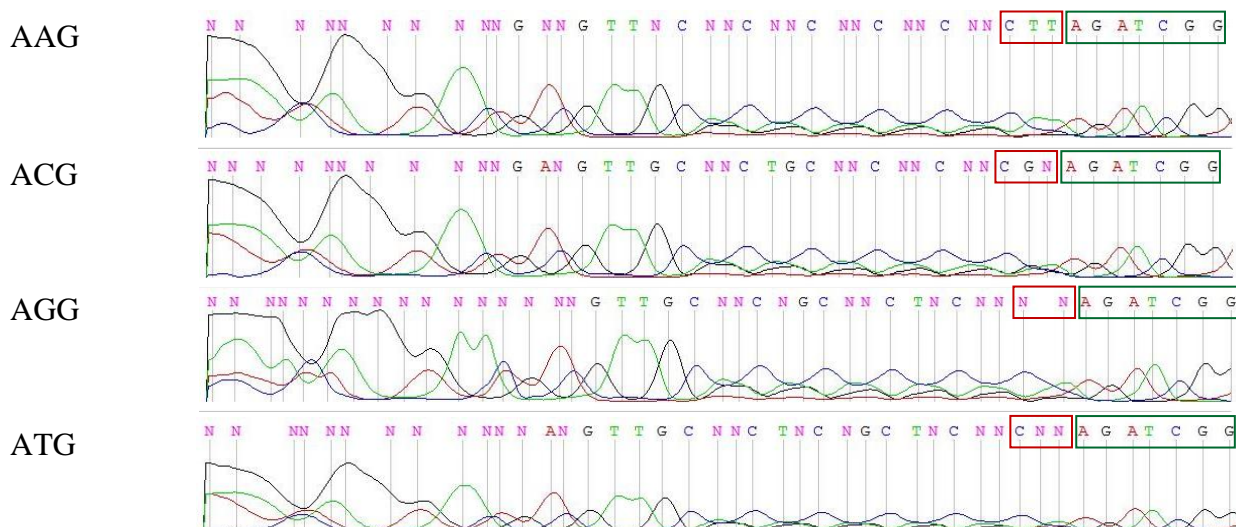
<b>KF Flex</b>	<b>Elution</b>		
Instrument type		KingFisher Flex	
Description			
	<b>Tip1</b>		
Tip		96er Deepwell-Platte	<empty>
Pick-Up plate		Tip-Platte	
Leave plate		Tip-Platte	
	<b>Elution</b>		
Plate		Elutionsplatte	Reagents: 20mM Tris (pH8), 1M NaCl
Beginning of step:			Volume [µl]: 50
Precollect		No	
Release beads		Yes	
Mixing/heating parameters:			
Heating during mixing		No	
Mixing time [hh:mm:ss]		00:30:00	
Mixing speed		Slow	
End of step:			
Postmix		No	
Collect count		3	

## 6.2 SEQUENZIERUNGSDATEN

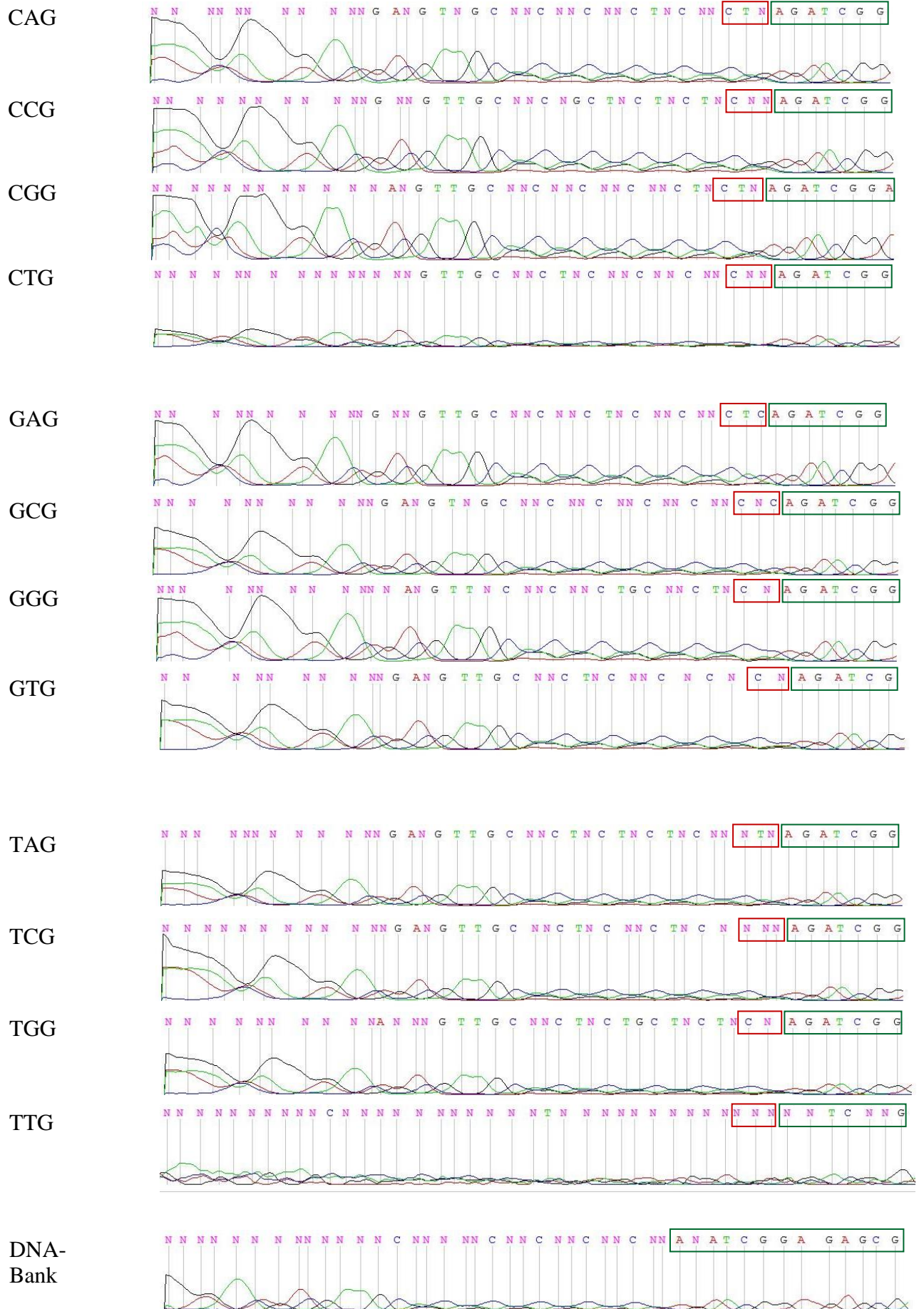
### SANGER-SEQUENZIERUNGSDATEN

Es wurde lediglich ein Ausschnitt aus den Sequenzierungsmaterial, mit den gesuchten Sequenzen dargestellt und nicht die gesamte sequenzierte DNA-Sequenz.

Adapter:







## 454-PYROSEQUENZIERUNGSDATEN

**Tab. A1: 454-Pyrosequenzierungsdaten.** Anzahl der vorkommenden MID-Kombinationen.

MID-Kombination	Bezeichnung	Reads
MID1fS_MID4rS	Initiale DNA-Bank (ePCR)	21549
MID2fS_MID5rS	DNA-Bank nach Synthese (ePCR)	3884
MID3fS_MID4rS	DNA-Bank nach Selektion (offene PCR)	8513
MID1fS_MID5rS	DNA-Bank nach Synthese (offene PCR)	1052
MID2fS_MID6rS	DNA-Bank nach Selektion (ePCR)	17195
MID3fS_MID6rS	Initiale DNA-Bank (offene PCR)	18515

∑ 70708 Reads

**Tab. A2: 454-Pyrosequenzierungsdaten.** Ampliconsequenzlänge und deren Häufigkeit. Amplicons mit der richtigen Länge wurden rot markiert.

Ampliconlänge in bp	Reads
10	61
11	16
12	22
13	35
14	104
15	212
16	881
17	5844
18	52808
19	7012
20	2248
21	521
22	213
23	48
24	16
25	5
26	3
27	1
32	1
42	6
43	3
44	6
45	9
46	3
47	10
48	3
49	8
50	12
51	30
52	48

53	54
54	26
55	20
56	4
57	4
58	2
59	1
[...]	[...]

**Tab. A3: 454-Pyrosequenzierungsdaten.** 1. Codonset und deren Häufigkeit in Prozent für die offenen und geschlossenen PCR-Daten der initialen DNA-Bank, der DNA-Bank nach Synthese und der DNA-Bank nach Selektion.

Codon	offene PCR-Daten			Emulsions-PCR-Daten		
	<i>DNA-Bank</i>	<i>Synthese</i>	<i>Selektion</i>	<i>DNA-Bank</i>	<i>Synthese</i>	<i>Selektion</i>
GCG	7,00%	11,90%	7,90%	11,40%	10,00%	9,70%
GAG	9,15%	10,60%	8,80%	8,10%	7,60%	11,40%
TGG	5,80%	7,00%	6,90%	4,00%	10,40%	4,10%
TCG	3,70%	7,00%	4,30%	1,00%	0,60%	3,10%
GTG	7,70%	7,00%	6,50%	10,90%	6,20%	7,00%
CGG	6,80%	7,00%	7,00%	5,70%	6,60%	5,40%
CAG	6,50%	6,60%	6,30%	3,30%	2,80%	7,40%
GGG	7,10%	6,20%	10,10%	24,60%	17,00%	17,10%
CCG	6,10%	6,20%	4,60%	4,40%	11,20%	5,10%
ACG	7,30%	6,20%	8,10%	6,80%	5,00%	4,50%
ATG	6,70%	5,70%	6,30%	3,20%	2,60%	2,70%
AAG	6,40%	5,70%	8,30%	2,40%	3,60%	7,90%
AGG	7,40%	5,30%	9,50%	10,70%	7,90%	6,00%
CTG	4,40%	4,80%	3,70%	1,50%	4,10%	3,10%
TAG	5,50%	4,00%	6,10%	1,40%	3,60%	3,40%
TTG	2,30%	3,10%	2,00%	0,70%	1,10%	2,00%

**Tab. A4: 454-Pyrosequenzierungsdaten.** 2. Codonset und deren Häufigkeit in Prozent für die offenen und geschlossenen PCR-Daten der initialen DNA-Bank, der DNA-Bank nach Synthese und der DNA-Bank nach Selektion.

Codon	offene PCR-Daten			Emulsions-PCR-Daten		
	<i>DNA-Bank</i>	<i>Synthese</i>	<i>Selektion</i>	<i>DNA-Bank</i>	<i>Synthese</i>	<i>Selektion</i>
GCG	6,40%	6,60%	5,20%	6,90%	5,70%	6,30%
GAG	7,30%	6,60%	8,30%	7,90%	4,80%	9,00%
TGG	5,20%	5,20%	4,30%	5,40%	6,20%	5,80%
TCG	4,90%	7,90%	4,30%	3,60%	5,60%	4,20%
GTG	4,90%	3,50%	5,20%	6,60%	5,90%	2,40%
CGG	6,90%	6,10%	6,60%	8,40%	6,90%	7,30%
CAG	7,30%	8,30%	8,30%	9,70%	5,30%	7,80%
GGG	5,00%	8,30%	7,20%	5,70%	9,60%	7,00%

CCG	7,10%	5,70%	6,90%	10,80%	9,90%	9,70%
ACG	7,40%	7,40%	6,70%	6,50%	7,50%	6,70%
ATG	6,30%	7,40%	5,60%	4,10%	5,00%	5,40%
AAG	8,40%	3,90%	8,90%	4,90%	5,40%	5,60%
AGG	6,50%	10,00%	7,50%	5,40%	8,40%	7,30%
CTG	5,90%	5,70%	4,60%	6,60%	4,40%	4,40%
TAG	6,10%	5,20%	6,60%	5,40%	8,20%	5,50%
TTG	4,60%	2,20%	3,60%	2,00%	2,30%	5,70%

**Tab. A5: 454-Pyrosequenzierungsdaten.** Zwei Codonsets die von Interesse waren und deren Häufigkeit in Prozent für die Emulsions-PCR-Daten der initialen DNA-Bank, der DNA-Bank nach Synthese und der DNA-Bank nach Selektion.

Codon	Emulsions-PCR-Daten		
	<i>DNA-Bank</i>	<i>Synthese</i>	<i>Selektion</i>
GGGGGG	6,50%	9,70%	7,90%
AAGAAG	0,40%	1%	2,20%
GAGGAG	7,10%	3,90%	3,90%
CAGCAG	4,60%	1,90%	3,50%
GGGAAG	12,90%	27,20%	11,80%
GGGGAG	18,70%	14,60%	9,90%
GGGCAG	20,30%	6,80%	12,40%
AAGGGG	1,80%	1%	2,50%
AAGGAG	1,60%	1%	11,30%
AAGCAG	2,60%	1%	4,80%
GAGGGG	6,70%	5,20%	11,70%
GAGAAG	3,90%	18,40%	2,70%
GAGCAG	6,30%	4,90%	5,70%
CAGGGG	1,90%	1%	2,70%
CAGAAG	0,90%	1%	3,70%
CAGGAG	2,10%	1%	3,10%

**Tab. A6: 454-Pyrosequenzierungsdaten.** Zwei Codonsets und deren Häufigkeit in Prozent für die Emulsions-PCR-Daten der initialen DNA-Bank, der DNA-Bank nach Synthese und der DNA-Bank nach Selektion.

Codon	Emulsions-PCR-Daten		
	<i>DNA-Bank</i>	<i>Synthese</i>	<i>Selektion</i>
GGGGCG	9,50%	2,20%	8,00%
GGGCAG	12,60%	3,90%	7,20%
GCGGAG	5,80%	5,50%	6,90%
GGGAAG	8,00%	15,50%	6,80%
GAGGGG	4,20%	3,30%	6,70%
GAGCCG	5,50%	4,40%	6,70%
AAGGAG	1%	0,60%	6,50%
GTGTTG	1,20%	1,10%	5,90%

GGGGAG	11,70%	8,30%	5,70%
GTGCCG	7,40%	4,40%	5,40%
GCGACG	4,20%	3,90%	5,30%
GGGCCG	0,10%	11,60%	5,10%
CGGTAG	1,90%	0,60%	5,10%
AGGCGG	5,80%	3,90%	4,80%
GGGCGG	11%	7,20%	4,70%
GGGGTG	8%	18,20%	4,50%

## 6.3 PUBLIKATIONEN

### Artikel

Schütze T., Rubelt F., **Repkow J.**, Greiner N., Erdmann V. A., Lehrach H., Konthur Z., Glökler J. (2011). A streamlined protocol for emulsion PCR and subsequent purification. *Analytical Biochemistry*. 410 (1):155-157.

### Poster

**Repkow, J.**, Lehrach, H., Glökler, J. (2011). Neclys (Nucleic acid-encoded chemical library synthesis and selection). *Day of Science*. Berlin, Germany

## **6.4 LEBENSLAUF**

Der Lebenslauf ist in der Online-Version aus Gründen des Datenschutzes nicht enthalten.

## 6.5 DANKSAGUNG

Für die Betreuung, Begutachtung und Möglichkeit meine Dissertation am Max-Planck-Institut für molekulare Genetik durchführen zu können, möchte ich mich zunächst bei Herrn Prof. Dr. Hans Lehrach bedanken.

Frau Prof. Dr. Beate Koksch von der Freien Universität Berlin danke ich für die unkomplizierte externe Betreuung und Übernahme der Begutachtung der vorliegenden Arbeit.

Besonderer Dank gilt Dr. Jörn Glökler, der durch die Bereitstellung des Themas maßgeblich zu dieser Arbeit beigetragen hat. Auch für die fortwährende Betreuung und Hilfe, sowie den zahlreichen Gesprächen und Erklärungen bei der Erarbeitung der Dissertation bedanke ich mich.

Bei meinen Eltern möchte ich mich für ihre immer anhaltende Unterstützung in allen Lebenslagen bedanken und dass sie mir stets mit Rat und Tat zur Seite stehen. Meinem Vater möchte ich für die grammatikalische als auch kausale Korrektur meiner Arbeit nochmals danken und meiner Mutter für ihren immer anhaltenen Optimismus.

Mein größter Dank gilt natürlich meiner kleinen Familie!

Danke Johanna, dass du so ein lieber Engel bist und es mir durch deine fröhliche und unkomplizierte Art einfach gemacht hast, die nötige Zeit für die Beendigung dieser Arbeit zu finden. Du bist das beste Teilergebnis, das im Laufe dieser Arbeit entstanden ist.

Danke Karsten, dass wir zusammen immer ein Team sind!



## **6.6 SELBSTSTÄNDIGKEITSERKLÄRUNG**

Hiermit versichere ich, Julia Knospe, dass die vorliegende Arbeit selbständig verfasst und ohne Benutzung anderer als der angegebenen Hilfsmittel und Quellen angefertigt wurde.

Oranienburg, den 10.01.2015

\_\_\_\_\_  
Unterschrift