

**Untersuchungen zur Erkennung,  
zum Nachweis und zur simultanen  
Genotypisierung von  
Hepatitis C-Viren (HCV)  
mittels Mikroarrays**

Dissertation

Zur Erlangung des Grades eines Doktors der Naturwissenschaften

– Dr. rer. nat. –

Fachbereich Biologie/Chemie der Universität Bremen

vorgelegt von

**Denja Katharina Drutschmann**

Bremen, im Mai 2004

Tag des Promotionskolloquiums: 23. Juli 2004

1. Gutachter: Prof. Dr. Dietmar Blohm
2. Gutachter: PD Dr. Andreas Dotzauer

# Inhaltsverzeichnis

<b>1</b>	<b>EINLEITUNG.....</b>	<b>1</b>
<b>1.1</b>	<b>Mikroarrays .....</b>	<b>1</b>
1.1.1	DNA-Mikroarrays .....	3
1.1.1.1	Oberflächen .....	3
1.1.1.2	DNA-Oligonukleotide .....	4
1.1.2	Tissue-Mikroarrays und Proteinarrays .....	6
1.1.3	Aufbringen der Fänger-moleküle .....	7
1.1.3.1	Spotten.....	7
1.1.3.2	Photolithographie .....	8
1.1.4	Messdatenanalyse .....	9
1.1.5	Interpretation der Messdaten.....	10
1.1.6	Bioinformatik .....	10
1.1.6.1	Bayessche Netze .....	13
<b>1.2</b>	<b>Hepatitis C-Virus (HCV) .....</b>	<b>15</b>
1.2.1	Das HCV Genom.....	16
1.2.2	Klassifizierung.....	18
1.2.2.1	Geographische Verteilung des Hepatitis C-Virus .....	19
1.2.3	Verwendete Verfahren zur Hepatitis C-Virus-Diagnose.....	20
<b>2</b>	<b>AUFGABENSTELLUNG.....</b>	<b>25</b>
<b>3</b>	<b>MATERIAL UND METHODEN.....</b>	<b>26</b>
<b>3.1</b>	<b>Design der Oligonukleotid-Bibliotheken.....</b>	<b>26</b>
3.1.1	Manuell designte HCV-Oligonukleotid-Bibliothek .....	26
3.1.1.1	Datenquelle .....	26
3.1.1.2	Alignments .....	26
3.1.1.3	Design der Oligonukleotide.....	26
3.1.2	Software-generierte HCV-Oligonukleotid-Bibliothek .....	27
<b>3.2</b>	<b>Primer, Oligonukleotide und HCV-Proben.....</b>	<b>28</b>
<b>3.3</b>	<b>Verwendete Lösungen und Puffer .....</b>	<b>31</b>
<b>3.4</b>	<b>Herstellung der HCV-Targets .....</b>	<b>34</b>
3.4.1	Gene-Elute Mammalian total RNA Kit (Sigma).....	34
3.4.2	QIAamp viral Mini Kit (QIAGEN).....	34
3.4.3	Two-Step RT-PCR .....	34
3.4.4	Reinigung .....	35
3.4.5	Targetherstellung mittels PCR .....	35
3.4.6	Quantifizierung der Proben .....	35
3.4.7	Einzelstrang-Isolation.....	36

<b>3.5</b>	<b>Sequenzierung .....</b>	<b>36</b>
3.5.1	Fällung der DNA für die Sequenzierung.....	36
<b>3.6</b>	<b>Herstellung der Mikroarrays .....</b>	<b>37</b>
3.6.1	Slides .....	37
3.6.2	Roboter.....	37
3.6.2.1	Ultramikro-Pipettierroboter, CAG, Bremen.....	37
3.6.2.2	SpotArray 24 (Packard BioScience).....	37
<b>3.7</b>	<b>Hybridisierungen.....</b>	<b>38</b>
3.7.1	Anbinden .....	38
3.7.2	Abblocken der freien reaktiven Gruppen.....	38
3.7.3	Prähybridisierung und Hybridisierung .....	38
3.7.4	Auswertung.....	39
<b>3.8</b>	<b>Optimierung der Hybridisierungsbedingungen .....</b>	<b>39</b>
3.8.1	Slidehomogenität.....	39
3.8.2	Variation der Hybridisierungstemperatur .....	39
3.8.3	Variation der Formamidkonzentration.....	39
3.8.4	Variation der Salzkonzentration in dem Waschpuffer .....	39
<b>4</b>	<b>ERGEBNISSE.....</b>	<b>40</b>
<b>4.1</b>	<b>Hybridisierung .....</b>	<b>40</b>
4.1.1	Optimierung.....	40
4.1.1.1	Schmelztemperatur .....	41
4.1.1.2	Formamid .....	42
4.1.1.3	Waschbedingungen.....	43
4.1.1.4	Spottingkonzentration.....	44
4.1.1.5	Zeit.....	45
4.1.1.6	Untersuchungen von Einzelstrang- gegen Doppelstrang Hybridisierungen	46
4.1.1.7	Mismatch.....	47
4.1.2	Weitere Parameter bei der Analyse der Hybridisierungsdaten.....	48
4.1.2.1	Bleaching .....	48
4.1.3	Untersuchung der Slidehomogenität .....	50
4.1.4	Qualität der Hybridisierungen.....	53
4.1.4.1	Intra-Array.....	53
4.1.4.2	Intra-Slide .....	56
4.1.4.3	Inter-Slide .....	57
<b>4.2</b>	<b>HCV-Target Charakterisierung.....</b>	<b>58</b>
4.2.1	RNA-Isolation .....	59
4.2.2	Sequenzierung .....	60
<b>4.3</b>	<b>Oligonukleotidbibliotheken.....</b>	<b>60</b>
4.3.1	Manuell erstellte Oligonukleotidbibliothek.....	60
4.3.1.1	Universelle Oligonukleotide.....	61

4.3.1.2	Hybridisierung mit Genotyp 1a aus der Targetgruppe 1 .....	63
4.3.1.3	Hybridisierung mit Genotyp 1b aus der Targetgruppe 1 .....	63
4.3.1.4	Hybridisierung mit Genotyp 3 aus der Targetgruppe 1 .....	64
4.3.1.5	Hybridisierung mit Genotyp 4 aus der Targetgruppe 1 .....	65
4.3.1.6	Zusammenfassung .....	65
4.3.2	Software-generierte Oligonukleotid-Bibliothek .....	67
4.3.2.1	Hybridisierung mit Genotyp 1a aus der Targetgruppe 1 .....	68
4.3.2.2	Hybridisierung mit Genotyp 1b aus der Targetgruppe 1 .....	69
4.3.2.3	Hybridisierung mit Genotyp 3 aus der Targetgruppe 1 .....	70
4.3.2.4	Hybridisierung mit Genotyp 4 aus der Targetgruppe 1 .....	70
4.3.2.5	Zusammenfassung .....	71
<b>4.4</b>	<b>Verifizierung mit weiteren HCV-Targets .....</b>	<b>75</b>
4.4.1	Manuelle Oligonukleotid-Bibliothek.....	76
4.4.2	Software-generierte Oligonukleotidbibliothek .....	78
4.4.2.1	Analyse mit Intra-Array Normierung.....	78
4.4.2.2	Analyse mit iQomega-Software .....	79
<b>5</b>	<b>DISKUSSION .....</b>	<b>82</b>
<b>5.1</b>	<b>Etablierung und Optimierung der Mikroarrays .....</b>	<b>82</b>
<b>5.2</b>	<b>Vergleich und Charakterisierung der Oligonukleotid-Bibliotheken.....</b>	<b>86</b>
<b>5.3</b>	<b>Design.....</b>	<b>91</b>
<b>6</b>	<b>ZUSAMMENFASSUNG UND AUSBLICK.....</b>	<b>95</b>
<b>7</b>	<b>LITERATURVERZEICHNIS.....</b>	<b>98</b>
<b>8</b>	<b>ANHANG.....</b>	<b>108</b>
<b>9</b>	<b>DANKSAGUNG.....</b>	<b>114</b>

# 1 EINLEITUNG

## 1.1 Mikroarrays

Seit dem Jahr 2001 ist das menschliche Genom sequenziert <sup>[1]</sup> und nun gilt es, in diesem „Buch“ lesen zu können, d.h. Gene zu identifizieren und ihr komplexes Zusammenspiel zu entschlüsseln. Mit Hilfe von Mikroarrays können entsprechende Experimente, z.B. die differentielle Genomanalyse im Hochdurchsatz, die Veränderung der Genexpression, die Wechselwirkungen von Genen, sowie Untersuchungen einzelner Gene oder Allele, durchgeführt werden. Das Prinzip der Mikroarrays beruht auf einer spezifischen Erkennung zweier DNA oder RNA Stränge: Ist die Sequenz einer Probe zu der des zu analysierenden RNA- oder DNA-Fragments komplementär, so kann sie auf Grund der Watson-Crick Basenpaarung selektiv hybridisieren und dann anhand von Markierungen, z.B. radioaktiv oder mit (Fluoreszenz-) Farbstoffen detektiert werden. Die Grundlage ist die spezifische Detektion von DNA, immobilisiert auf Nitrocellulose-Membranen, wie sie Edwin Southern als erster beschrieben hat <sup>[2]</sup>. Der technologische Fortschritt wie beispielsweise positioniergenaue Robotersysteme und die Bioinformatik ermöglichen eine enorme Miniaturisierung <sup>[3]</sup>. Während mit dem Southern-Blot nur ein Gen zur Zeit untersucht werden konnte, ist es nun möglich, über eine Million Spots parallel auf kleinstem Raum zu untersuchen <sup>[4]</sup>. Die Herstellung von Mikroarrays erfolgt dabei über Kontaktprinting, Pipettiersysteme <sup>[5]</sup> oder der *in-situ* Synthese von Oligonukleotiden auf der Oberfläche z.B. mit photolithographischen Verfahren <sup>[6]</sup>. Die Hauptanwendungsgebiete der DNA-Mikroarrays sind die Bestimmung des Expressionslevels von Genen (siehe Abbildung 1-1) und die Identifizierung von Sequenzen z.B. für Mutationsuntersuchungen.

Für die Genexpressionsanalyse gibt es zwei Varianten von Mikroarrays. Zum Einen können lange DNA-Fragmente auf dem Slide immobilisiert werden, z.B. PCR-Fragmente oder cDNAs. Sie erzielen hohe Signalintensitäten nach der Hybridisierung, was für die Detektion schwach exprimierter Gene wichtig ist. Zum Anderen gibt es Oligonukleotid-Mikroarrays, deren Nukleotide eine Länge zwischen 20 und 100 Basen aufweisen. Die langen Oligonukleotide (ab 50mere) kommen ebenfalls bei der Expressionsanalyse zum Einsatz. Hierbei repräsentiert ein Oligonukleotid ein Gen oder ein Transkript <sup>[7]</sup>. Die Firma Affymetrix stellt Mikroarrays für die Genexpressions-Analyse mit kurzen Oligonukleotiden (25mere) her, die sie *in-situ* (photolithographisch) synthetisieren. Jedes Gen wird hierbei

von 11-20 einzelnen Oligonukleotid-Paaren abgedeckt. Ein Oligonukleotid-Paar besteht aus einer Probe ohne Basenfehlpaarung und einer mit einer mittelständigen Basenfehlpaarung.

Auch in der Entdeckung und Erforschung neuer Medikamente spielt die DNA-Mikroarray-Technik eine Rolle [8, 9]. Mit Hilfe der Expressionsanalyse können anhand veränderter Expressionsprofile der behandelten Zelle(n) Anhaltspunkte auf Wirksamkeit oder Nebenwirkungen der Medikamente ermittelt werden. Bei Untersuchungen von Krankheitserregern werden Mikroarrays eingesetzt [10, 11]. So hat z.B. Troesch *et al.* [12] basierend auf der 16S rRNA 26 verschiedene Spezies des Mycobacteriums, Erreger der Tuberkulose untersucht. Auch bei der Diagnose von Krebs werden DNA Mikroarrays verwendet [13, 14]. Hierbei wird der Zellstatus auf einer molekularen, genomweiten Ebene untersucht. So kann jeder einzelne Zelltyp anhand seines Expressionsprofils identifiziert werden. Bisher steht diese Art der Diagnostik nur den Forschern zur Verfügung. Um sie auch in der Klinik anwenden zu können, sind weitere klinische Studien notwendig.

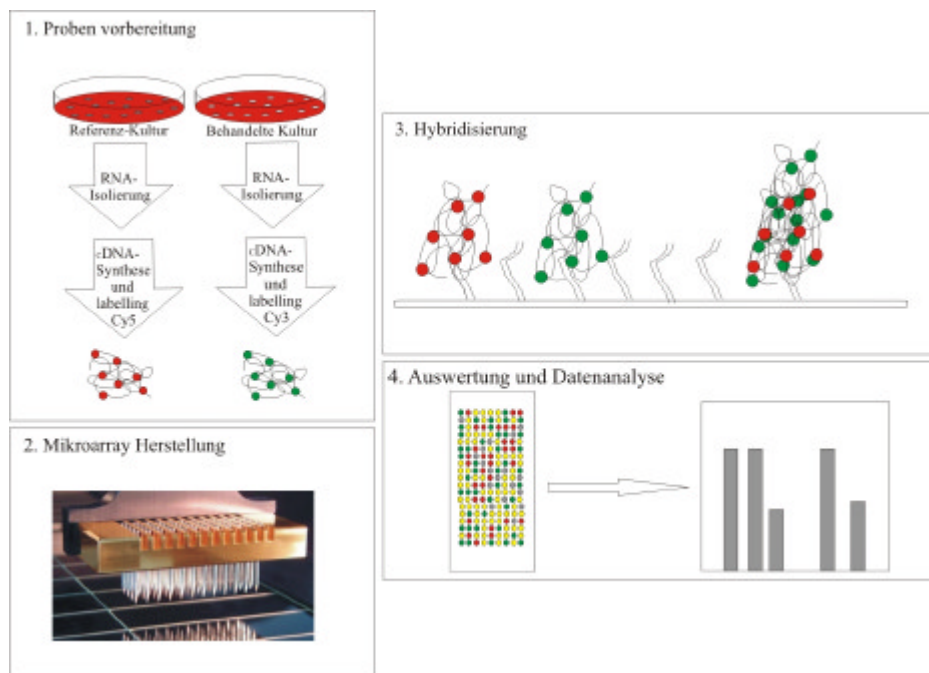


Abbildung 1-1: Flussdiagramm einer Genexpressionsanalyse

Die Mikroarray-Technologie wird ebenfalls für die Genotypisierung eingesetzt, z.B. zur Detektion von Einzelbasenaustauschen (Single Nucleotide Polymorphism, SNPs) über Single Base Extension (SBE) [15]. Hierbei werden kurze DNA-Stränge definierter Sequenzen, sog. TAGs, auf die aktivierte Oberfläche, den Slide, gespottet. Diese dienen der

Positionierung auf dem Slide. Die Primer für die SBE bestehen aus den revers-komplementären Sequenzen der TAGs und der Sequenz des nachzuweisenden Allels. Letztere wird so ausgewählt, dass das 3' Ende eine Base vor der SNP-Position endet. In der SBE-Reaktion lagern sich die SBE-Primer spezifisch an die entsprechende Allelsequenz und eine Thermosequase verlängert den SBE-Primer gemäß der Watson-Crick Basenpaarung um das entsprechende ddNTP, welches durch Fluoreszenzfarbstoffe eindeutig zugeordnet werden kann. Das ddNTP verhindert eine weitere Verlängerung der Primer und durch die spezifische Markierung des ddNTP mit einem Fluoreszenzfarbstoff, kann bei der anschließenden Hybridisierung die Art des Polymorphismus ermittelt werden.

Neben Experimenten mit markierten Targetmolekülen, gibt es auch Mikroarrays, auf denen markierungsfrei detektiert werden kann. Als Beispiel wird das Pyrosequencing auf DNA-Mikroarrays zur Analyse von SNPs beschrieben <sup>[16]</sup>. Gänzlich ohne Markierung kommt die reflectometric-interference-spectroscopy (RifS) Methode aus <sup>[17]</sup>. Mit Hilfe der Interferenz-Differenz einer Oberfläche mit aufgetragenen Fängeroligonukleotiden vor und nach der Hybridisierung mit dem Targetmolekül, lässt sich ein Hybridisierungsereignis ermitteln. Hashimoto *et al.* zeigten, dass Minor Groove Binder, der spezifisch DNA bindet, elektrochemisch aktiv ist, und somit zur elektrochemischen Detektion von hybridisierter DNA eingesetzt werden kann <sup>[18]</sup>. Jianbing *et al.* stellten die Eignung von Glasfaserarrays bei der funktionellen Genomanalyse vor <sup>[19]</sup>.

### **1.1.1 DNA-Mikroarrays**

#### *1.1.1.1 Oberflächen*

Es gibt verschiedene Konzepte zur Bindung von Oligonukleotiden auf aktivierten Oberflächen. Auf nichtkovalente Oberflächen, beispielsweise poly-L-Lysin oder Aminosilan, binden die unmarkierten Oligonukleotide über Wasserstoffbrücken oder über ionische Wechselwirkung der funktionellen Gruppen mit dem negativ geladenen Phosphatrückgrat des gespotteten Oligonukleotides. Bei kovalenten Oberflächen müssen die Oligonukleotide eine funktionelle Gruppe, z.B. eine Aminogruppe, tragen, mit der sie eine chemische Bindung auf Oberflächen mit z.B. Aldehyd-, Isothiocyanat- (siehe Abbildung 1-2) oder Epoxy-Gruppen ausbilden. Zudem gibt es Dreidimensionale Oberflächen, die eine baumartige Polymerschicht aufbauen und durch ihre dreidimensionale Struktur eine größere Anzahl an reaktiven Gruppen für die Bindung der



gespotteten Oligonukleotiden bereitstellen [20, 21]. Es hat sich bisher keine Oberflächenmodifikation vollständig durchgesetzt.

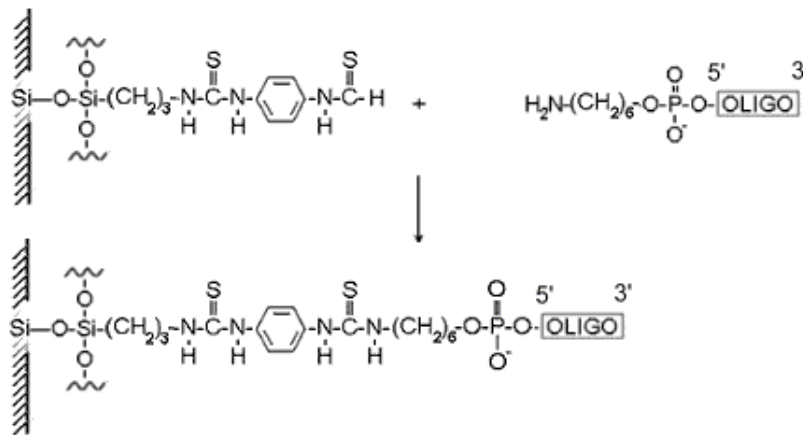


Abbildung 1-2: Schematische Darstellung einer Slideoberfläche nach Aktivierung mit Phenylendiisothiocyanat und die Ausbildung einer Amidbindung mit einem aminomodifizierten Oligonukleotid zu einer kovalenten Bindung.

#### 1.1.1.2 DNA-Oligonukleotide

Die verschiedenen Anwendungen der DNA-Mikroarrays stellen unterschiedliche Ansprüche an die Sensitivität bzw. Spezifität der Oligonukleotide. Beide Parameter beeinflussen sich gegenseitig. Kürzere Oligonukleotide haben eine höhere Sensitivität, die längeren erreichen eine höhere Spezifität gegenüber der Ziel-DNA. Oligonukleotide zwischen 20 und 35 Nukleotiden Länge werden sowohl für die Detektion von SNPs als auch in der Expressionsanalyse verwendet. Bei einem Affymetrix-Chip werden zwischen 11 und 20 Shortmers *in-situ* synthetisiert, die ein einziges Gen repräsentieren. Das ist notwendig, um die erforderliche Spezifität zu erreichen. Darüber hinaus ist ein Oligonukleotid mit einer mittelständigen Basenfehlpaarung zu jedem dieser Oligonukleotide vorhanden. Diese mismatch Oligonukleotide dienen ebenfalls dazu die Spezifität der Hybridisierung zu gewährleisten. Eine Analysesoftware errechnet aus den Signalintensitäten der mismatch Oligonukleotide und der Perfect Match Oligonukleotide einen endgültigen Wert. Oligonukleotide zwischen 50 und 70 Nukleotiden Länge werden in der Genexpressionsanalyse angewendet. Ihre Länge macht sie spezifisch für ihre Ziel-Gene. Für SNP-Nachweise sind sie auf Grund ihrer Länge ungeeignet, da einzelne Basenfehlpaarungen nicht mehr abgetrennt werden können.

Die Begriffe Sensitivität und Spezifität werden hier im biologischen Sinn verwendet <sup>[22]</sup>. In Kapitel 1.1.6 werden sie im bioinformatischen Sinne verwendet. Im biologischen Sinn wird die Sensitivität hier als Nachweisgrenze bei der Auswertung der Signalintensitäten nach der Hybridisierung beschrieben. Durch die Länge eines gebildeten cDNA-DNA-Hybrids wird hier eine stärkere Bindung erreicht als bei den kürzeren cDNA-Oligonukleotid-Hybriden. Mit Spezifität wird die Übereinstimmung der Reihenfolge der Nukleotidbasen bei der Bindung beschrieben. Dabei bildet nach der Watson-Crick Basenpaarung ein Cytosin mit einem Guanin drei Wasserstoffbrücken aus, Adenin mit Thymin dagegen zwei. Ist die Basenpaarung des Hybrids korrekt, so ist die Bindungsstärke optimal. Sie verringert sich, wenn eine Basenpaarung z.B. aufgrund einer Punktmutation nicht mehr komplementär ist. Je länger ein Oligonukleotid ist, desto höher ist die Wahrscheinlichkeit, dass diese Sequenz kein zweites Mal in einem Pool von Sequenzen vorkommt. Das vermindert die Möglichkeit der Kreuzhybridisierung, d.h. Bindung an Nicht-Zielsequenzen. Kreuzhybridisierungen können bei kurzen Oligonukleotiden häufiger auftreten. Aufgrund ihrer geringeren Anzahl an Basen und somit relativ geringen Stabilität der Hybridbindung können cDNA-Oligonukleotid-Hybride mit Basenfehlpaarungen durch stringente Hybridisierungs- oder Waschbedingungen im Experiment von Hybriden ohne Basenfehlpaarung abgetrennt werden. Dabei ist die Lage der Basenfehlpaarung von Bedeutung. Mittelständige Basenfehlpaarungen können leichter abgetrennt werden als endständige.

Eine verbesserte Diskriminierung von einzelnen Basenfehlpaarungen wird durch den Einsatz von künstlichen Basenanaloga erreicht, wie z.B. PNA und LNA <sup>[23, 24]</sup>.

Peptide Nucleotide Acids (PNA) unterscheiden sich im Aufbau zur DNA durch ein N-(2-Aminoethyl)glycin Proteinrückgrat anstelle des Desoxyribosephosphat-Rückgrates. Sie bilden mit der DNA über die Watson-Crick Basenpaarung B-DNA ähnliche Strukturen aus <sup>[25]</sup>. Diese PNA-DNA-Interaktionen sind spezifischer als entsprechende DNA-DNA oder RNA-RNA Bindungen. Trotz des unterschiedlichen Aufbaus des Rückgrades gleichen sie in der Struktur der DNA. Neben der enormen biochemischen Stabilität zeigt PNA eine hohe Bindungsaffinität und hohe Spezifität. PNA-DNA Duplexe ermöglichen die Detektion eines einzelnen Basenaustauschs bei niedrigeren Raumtemperaturen <sup>[26]</sup> und daher sind hier sichere Unterscheidungen z.B. bei Mutationen möglich <sup>[27]</sup>.

Eine weitere Form synthetischer Moleküle sind Locked Nucleic Acids (LNA). Dabei handelt es sich um bicyclische RNA-Analoga. Wie PNAs, haben LNAs eine höhere

Affinität zu ihren komplementären Targetmolekülen <sup>[28]</sup> und können somit zur Detektion von SNP-Analysen, <sup>[29, 30]</sup> aber auch in der Genexpression eingesetzt werden <sup>[31]</sup>.

### 1.1.2 Tissue-Mikroarrays und Proteinarrays

Neben DNA-Mikroarrays gibt es noch eine Reihe weiterer Arten von Mikroarrays. Bei den Tissue-Mikroarrays (TMA) werden bis zu 1000 Gewebeproben auf einen Slide gebracht und parallel analysiert <sup>[32]</sup>. Bereits identifizierte Gene, die eine bestimmte Rolle in der Entwicklung oder dem Fortschreiten einer Krankheit spielen sollen, können mit Tissue-Mikroarrays *in-situ* nachgewiesen werden <sup>[33]</sup>. Auf diese Weise können die Erkenntnisse aus den cDNA Mikroarrays auf die TMAs übertragen und darüber hinaus noch erweitert und validiert werden. Mehr noch, die TMAs können als Qualitätskontrolle der Immunhistochemie und der *in-situ* Hybridisierung dienen <sup>[34]</sup>.

DNA-Mikroarrays werden für Expressionsuntersuchungen auf genomischen und dem Transkriptom Level, also der mRNA, herangezogen. Diese Daten lassen nicht unbedingt auf die Proteinexpression schließen. Die Funktionalität der Proteine hängt häufig von post-translationalen Prozessen des Vorläuferproteins oder von Regulationsmechanismen des Stoffwechselweges ab. Für diese Art von Untersuchungen gibt es verschiedene Anwendungen von Protein-Mikroarrays oder Ligand-Bindungs-Assays. Die Interaktion von Proteinen wurde bisher mit dem sog. Yeast-two-Hybrid System untersucht <sup>[35]</sup>. In einem genomumfassenden Versuch haben Zhu *et al.* <sup>[36]</sup> 5800 verschiedene rekombinante Proteine von *S. cerevisiae* auf einem Proteom-Chip untersucht. Dazu wurden Genprodukte auf den Chip gebracht und über 90% der Proteine dieses Organismus konnten untersucht werden. Untersuchungen zur Charakterisierung und Identifizierung von DNA-bindenden Proteinen, z.B. Transkriptionsfaktoren, können mit DNA-Protein-Assays durchgeführt werden <sup>[37]</sup>. Darüberhinaus können mit Enzym-Substrat-Assays verschiedene Enzyme beschrieben werden, z.B. Restriktionsenzyme, Peroxidasen, Phosphatasen und Proteinkinasen <sup>[35, 38, 39]</sup>. Mit Hilfe solcher Assays ist es möglich noch unbekannte Aktivitäten einzelner Enzyme zu identifizieren. Für Rezeptor-Ligand-Assays werden kleine organische Moleküle (small molecules) durch kombinatorische Festphasenchemie produziert <sup>[40]</sup> und in einem Arrayformat immobilisiert. Diese Mikroarrays werden mit fluorophormarkiertem Targetproteinen inkubiert, um so neue Liganden zu identifizieren <sup>[35]</sup>. Immunoassays sind von generellem Interesse vor allem für diagnostische Applikationen <sup>[41]</sup>. Antigen-Antikörper Interaktionen können so gescreent werden. Templin *et al.* haben achtzehn

verschiedene Autoantigene, die als diagnostische Marker für Autoimmunkrankheiten z.B. rheumatische Erkrankungen dienen, immobilisiert und so konnte eine parallele Detektion der verschiedenen Arten der Autoantikörpern durchgeführt werden <sup>[42]</sup>.

### 1.1.3 Aufbringen der Fängermoleküle

#### 1.1.3.1 Spotten

Für das Aufbringen von Oligonukleotiden oder Peptiden im Nanolitermaßstab ist die Verwendung von Pipettierrobotern notwendig. Diese ermöglichen die erforderliche Positioniergenauigkeit und exakte Dosierung kleinster Volumina. Es gibt zwei grundsätzliche Prinzipien von Pipettierrobotern oder Arrayern. Beim Kontaktverfahren werden zum Aufbringen der Oligonukleotide Stahlnadeln, sog. pins, verwendet. Sie tauchen in die zu spottende Lösung ein und setzen die Lösung auf der Slideoberfläche ab, vermittelt durch Adhäsionskräfte, die bei direktem Kontakt der Flüssigkeit mit der Slideoberfläche wirken. Durch Variation in der Dauer des Slidekontaktes kann die Spotgröße beeinflusst werden. Je länger die Verweildauer, desto größer ist der Spot. Jedoch ist anzumerken, dass für die Auswertung ein kleiner (bis zu 200  $\mu\text{m}$ ), klar abzugrenzender Spot vorteilhafter ist als ein großer Spot. Der Vorteil der Kontaktverfahren liegt in der Automatisierung. Zum Beispiel gibt es den SpotArray der Firma Perkin-Elmer® in zwei Ausführungen, mit 24 und 72 Nadeln (split pins). Im non-Kontaktverfahren werden die Oligonukleotide entweder durch Druckluft auf den Slide gebracht (z.B. mit dem Topspotter von IMTEK, Freiburg, Deutschland) oder durch Piezotechnik (z.B. GeSiM, Dresden, Deutschland; siehe Abbildung 1-3).

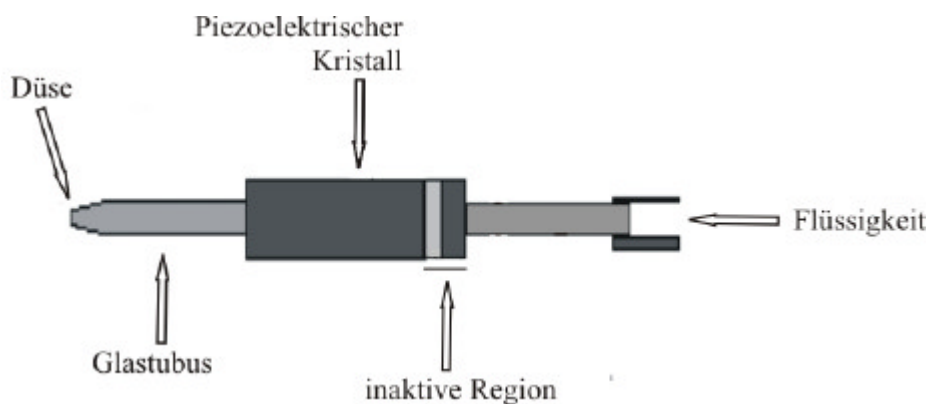


Abbildung 1-3: Schematische Zeichnung eines Piezo-Dispersers <sup>[43]</sup> stellvertretend für ein non-Kontaktverfahren.

Die Piezokeramik wird auf Glaskapillaren gebracht, durch Anlegen von Spannung zieht sich diese Keramik zusammen und ermöglicht so die Abgabe kleinster, definierter Flüssigkeitsmengen. Mit diesem Verfahren wird die Oberfläche der Slides nicht berührt und kann daher auch nicht beschädigt werden. Die Spotqualität, d.h. die Spotmorphologie ist bei dieser Technik, verglichen mit dem Kontakt-Verfahren, sehr gut. Kennzeichnend für eine gute Spotqualität ist eine homogene Morphologie. Diese wird durch die Spottinglösung und die Feuchtigkeit während des Spotvorgangs erreicht und muss optimiert werden. Allerdings wird mit diesen Pipettierrobotern kein Hochdurchsatz erreicht.

1.1.3.2 Photolithographie

Affymetrix hat 1993 die ersten sogenannten high density Mikroarrays mit über 100.000 Spots pro Mikroarray auf den Markt gebracht, die *in-situ* auf dem Mikroarray photolithographisch synthetisiert werden [6].

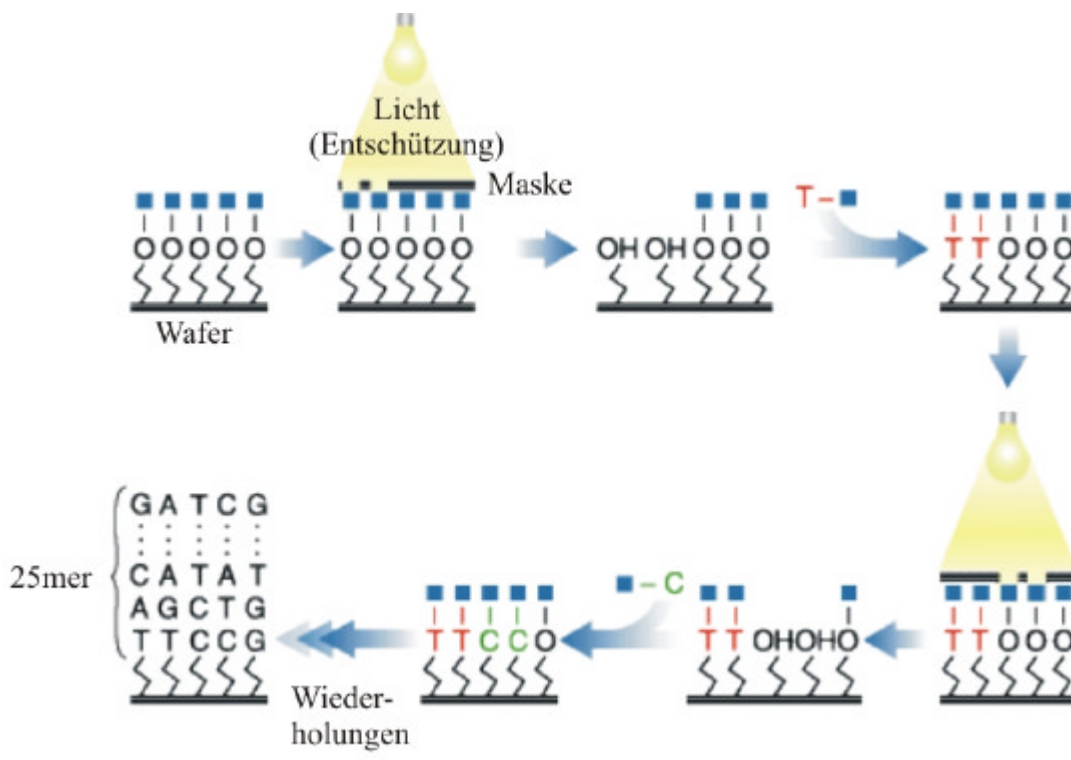


Abbildung 1-4: Schematische Darstellung der photolithographischen *in-situ* Synthese von Oligonukleotiden von Affymetrix, verändert nach [44].

Dazu entschützt ultraviolettes Licht die Stellen auf der Oberfläche, die für den anstehenden Syntheseschritt vorgesehen sind. Der kritische Schritt hierbei ist das präzise Übereinanderlagern der Maske über der Oberfläche vor dem nächsten Syntheseschritt. Ist

die entsprechende Stelle auf der Oberfläche aktiviert, wird eine Lösung eines Desoxynukleotides über die Oberfläche gespült. Diese Desoxynukleotide tragen ihrerseits wiederum Schutzgruppen, die im nächsten Schritt durch UV-Einstrahlung entschützt werden. In einem zyklischen Verfahren entstehen so bis zu 25mere. Längere Sequenzen sind aufgrund einer schlechten Kupplungseffizienz während der Synthese nicht möglich. Ein weiterer Effekt der *in-situ* Synthese sind die höheren Fehlerraten in der Basenabfolge. Die Folge sind nicht korrekt synthetisierte Sequenzen, was zu höherer Kreuzreaktivität führen kann.

#### **1.1.4 Messdatenanalyse**

Typischerweise werden für das Auslesen hybridisierter Mikroarrays Fluoreszenz-Scanner eingesetzt. Das Detektionsprinzip beruht auf der Verwendung von Fluorophoren, die durch Anregung mit einer bestimmten Wellenlänge in einen angeregten Zustand überführt werden. Bedingt durch die aromatische Struktur der Fluorophore werden delokalisierte  $\pi$ -Elektronen in ein höheres Energieniveau angehoben. Beim Zurückfallen auf den Grundzustand wird die Energiedifferenz spontan als Strahlung (Fluoreszenz) emittiert. Die Hybridisierungen können durch Anregung der an die Targetmoleküle gekoppelten Fluorophore mit Hilfe einer Lichtquelle selektiv und quantitativ ermittelt werden. Es gibt verschiedene Arten von Scannern. Bei Geräten mit polychromatischer Lichtquelle erfolgt die Anregung universell z.B. durch eine Weißlichtquelle, bei der alle Fluorophore gleichzeitig angeregt werden. Die Detektion erfolgt über ausgewählte, einstellbare Filter für die jeweils verwendeten Farbstoffe. Mit Hilfe einer CCD Kamera wird das emittierte Licht eingefangen. Die Belichtungszeit bestimmt die Sensitivität der Aufnahme. Daneben werden auch Laser als Lichtquelle eingesetzt. Hier erfolgt die Anregung über definierte Wellenlängen, die dem Anregungsoptimum des jeweiligen Fluorophores entsprechen. Detektiert wird das emittierte Licht der Fluorophore ebenfalls durch Filter, um den sog. Crosstalk zu minimieren. Mit Hilfe von Photomultipliertubes (PMTs) wird hier die Sensitivität der Aufnahme bestimmt. Hierbei handelt es sich um eine elektronische Vervielfältigung des eingefangenen Lichtsignals. Die Anzahl der Laser ist geräteabhängig. Die Art dieser Scanner kann ebenfalls in zwei Kategorien unterteilt werden: konfokal und nichtkonfokal. Bei den konfokalen Scannern wird nur eine bestimmte Schicht der Spots ausgelesen, während bei den nichtkonfokalen Geräten die gesamte Schichtdicke erfaßt wird.

### 1.1.5 Interpretation der Messdaten

Um Messdaten interpretieren zu können, müssen sie zunächst aufbereitet werden. Eine Normalisierung von Messdaten ist für den Vergleich zwischen Ergebnissen mehrerer Mikroarrayexperimente notwendig. Sie beseitigt die Messfehler und Ungleichmäßigkeiten. Der Hauptteil der Publikationen bezüglich DNA-Mikroarrays beschäftigt sich mit der Genexpressionsanalyse <sup>[45]</sup>, und von den dort behandelten Themen lassen sich nicht alle auf die hier angewendete Art von Mikroarrays übertragen. Eine Übersicht, welche Daten bei cDNA Mikroarrays normalisiert werden sollen, ist bei Brazma *et al.* in <sup>[46]</sup> beschrieben. Die dort beschriebenen Daten dienen dazu, die cDNA Mikroarray Experimente zu standardisieren und somit unter den Forschergruppen vergleichbar zu machen.

Auf zwei Aspekte der Normalisierung soll hier näher eingegangen werden: ab wann ist ein Signal als Signal zu werten und wie sollen zwei Arrays miteinander verglichen werden? Die Frage, was ein Signal ist, d.h. wie hoch muss der Unterschied zwischen Hintergrund und schwachem Signal sein, damit ein Signal als solches gewertet werden kann, diskutiert Quackenbush in <sup>[47]</sup>. Danach sollte ein Signal mindestens so hoch sein wie der Slide-Hintergrund nach der Hybridisierung plus zwei Standardabweichungen. Bei der Frage nach dem Vergleich zweier Arrays gibt es in der Genexpression zwei Möglichkeiten. Man kann global, d.h. über die Signalintensitäten des gesamten Arrays gemittelt <sup>[48]</sup> normieren, oder über housekeeping Gene, deren Signalintensität unverändert bleiben sollte. Diese Art wird als interne Normierung bezeichnet. Externe Normierung wird mit sog. spiking Kontrollen durchgeführt. Artifizielle Sequenzen oder Proben werden in bekannter Konzentration zu der Hybridisierungslösung gegeben und sollten dann immer gleiche Signale liefern (sofern die Hybridisierungsbedingungen unverändert bleiben) <sup>[49, 50]</sup>.

### 1.1.6 Bioinformatik

Der Anteil der Bioinformatik in der Molekularbiologie nimmt stetig zu. Es gibt Programme zur Auswahl von Primern für die PCR, von Oligonukleotiden, zur Untersuchung von Homologien zwischen Sequenzen (z.B. Blast, FASTA) und Alignmentprogramme, um nur einen Bruchteil derer zu nennen, die zur Ermittlung geeigneter Oligonukleotide verwendet werden. Auch nach der Hybridisierung, zur Analyse der Daten und zur Integration der neuen Daten in die bestehenden Erkenntnisse, hat die Bioinformatik Einzug gehalten. Dies liegt vor allem an den enormen Datenmengen, die be- und verarbeitet werden. Zunächst gibt es bioinformatische Ansätze, wie die Hybridisierungsdaten am besten ausgewertet und

vor allem normiert werden sollten. Auch hierzu behandelt der größte Anteil der veröffentlichten Publikationen die Genexpressionsanalyse, dem derzeitigen Hauptanwendungsgebiet von Mikroarrays.

Für die Berechnung der für die Hybridisierung wichtigen Schmelztemperatur gibt es verschiedene Programme, die z.T. auch unterschiedliche Formeln angeben. In den meisten Fällen fließt die Länge des Oligonukleotides ein, so dass für unterschiedliche Längen auch unterschiedliche Formeln verwendet werden sollen.

Eine vereinfachte Gleichung nur unter Berücksichtigung der Bindungsstärke zwischen den Basenpaarungen ist die Wallace Regel. Breslauer *et al.* <sup>[51]</sup> entwickelte und SantaLucia *et al.* <sup>[52, 53]</sup> erweiterte unter Berücksichtigung thermodynamischer Parameter eine Formel, die heute als Standard verwendet wird. Dabei wird jeweils einem Paar aufeinanderfolgender Basen entsprechende thermodynamische Parameter (Enthalpie  $\Delta H$  und Entropie  $\Delta S$ ) zugeordnet. So werden in dieser Gleichung sowohl die Sequenz, die Anzahl der einzelnen Nukleotide (A, T, C und G) wie auch die Salzkonzentration berücksichtigt. Durch Einbeziehung der Thermodynamik in die Berechnung der Schmelztemperatur können die sog. nearest neighbor Interaktionen berücksichtigt werden. Sie berücksichtigt nicht nur die Wasserstoffbrückenbindung zwischen den gegenüberliegenden Nukleinbasen, sondern auch die Kräfte entlang der Helixachse. Im Internet gibt es verschiedene Serviceseiten, die die Schmelztemperatur für das eingegebene Oligonukleotid berechnen <sup>[54]</sup>.

Enthalpie und Entropie sind Parameter, die auch bei der Ermittlung von Sekundärstrukturen involviert sind. Eine Minimierung interner Faltungen soll die freie Zugänglichkeit der Oligonukleotide während der Hybridisierung gewährleisten. Dabei werden  $\Delta G$  Werte ermittelt, die die Stabilität der Oligonukleotide widerspiegeln. Je weniger stabil ein Oligonukleotid ist, desto geeigneter ist es für die Hybridisierung. Je größer der negative Wert, desto wahrscheinlicher ist die Struktur. Zuker <sup>[55]</sup> hat seine Forschungsergebnisse zur Ermittlung der Sekundärstruktur als Service ins Internet gestellt unter <sup>[56]</sup>. Die Ermittlung des  $\Delta G$  erfolgte in Kooperation mit SantaLucia <sup>[52]</sup>.

BLAST (Basic Local Alignment Search Tool) ist ein Programm, um z.B. Sequenzdatenbanken zu durchsuchen. Der BLAST Algorithmus erkennt sowohl lokale als auch globale Alignments. Lokale Alignment Methoden wurden entwickelt, um nach sehr



ähnlichen Regionen in zwei Sequenzen zu suchen, wobei die ähnlichen Regionen nicht notwendigerweise dieselbe Orientierung haben müssen. BLAST-ähnliche Methoden suchen zunächst sehr kurze gemeinsame Segmente und erweitern dann diese Region so weit wie möglich. Algorithmen, wie z.B. von Smith-Waterman <sup>[57]</sup> identifizieren die wahrscheinlichste signifikante Übereinstimmung anhand von Evolutionsmodellen. Globale Alignment Methoden, z.B. von Needleman und Wunsch <sup>[58]</sup>, sollten angewendet werden, wenn zusätzliche Annahmen gemacht werden können, z.B., dass die ähnlichsten Regionen in den Sequenzen sollen in derselben Orientierung vorliegen. Diese Methode ermittelt ein „global map“ zwischen Sequenzen, wobei Alignments, die sich überlappen oder kreuzen, entfernt werden.

Bei der Auswahl der Software-generierten Bibliothek spielt die Sensitivität und Spezifität eine große Rolle. Die Bedeutung des Begriffes Sensitivität bei der Auswahl der Oligonukleotide unterscheidet sich von der molekularbiologischen Definition. Wie in Kapitel 3.1 beschrieben, wurden die Oligonukleotide aus einem Alignment von 944 Kompletengenom-Sequenzen des HCV aus den Datenbanken ausgewählt. Die Anzahl der einzelnen Sequenzen pro Genotyp ist dabei nicht gleichmäßig verteilt. So gibt es elf Sequenzen, die dem Subtyp 1b zugeordnet werden, aber nur vier für den Genotyp 4. Die bioinformatische Sensitivität gibt an, wie viele Sequenzen das betrachtete Oligonukleotid erkennt. Eine hohe Sensitivität hat ein Oligonukleotid, das viele Sequenzen erkennt.

Unter der Spezifität eines Oligonukleotides im bioinformatischen Sinn ist, wie in der molekularbiologischen Definition, die Erkennung eines bestimmten HCV-Genotypes oder Subtypes zu verstehen. Ist ein Oligonukleotid spezifisch, so erkennt es nur diesen HCV-Typ, für den es ausgewählt wurde.

Regeln für die Auswahl von Oligonukleotiden hat unter anderem Li *et al.* 2001 aufgestellt <sup>[59]</sup>. So ist neben der Schmelztemperatur und der freien Energie ( $\Delta G$ ), die Position des Oligonukleotides auf der Target-DNA wichtig, ebenso wie das Vorkommen einzelner Basen in der ausgewählten Oligonukleotid-Sequenz. So sollte die Anzahl der einzelnen Basen nicht mehr als 50 % der Probenlänge überschreiten, die Länge aufeinanderfolgender, gleicher Basen (sog. stretches) sollte nicht mehr als 25 % betragen und keine selbst-komplementären Regionen innerhalb des Oligonukleotides vorkommen.

Diese oder ähnliche Kriterien werden heute in Programmen für das Design von Oligonukleotiden oder Primern angewendet, siehe hier zu z.B. [60, 61, 62].

1.1.6.1 Bayessche Netze

Bayessche Netze werden verwendet, um die Wahrscheinlichkeit eines Ereignisses zu finden, wenn andere Ereignisse bekannt sind. Komplexität und Unsicherheiten in der Datenlage können mit diesem Werkzeug in Auswerteanalyse Software implementiert werden. Bayessche Netze werden häufig als Graphen dargestellt, bei denen verschiedene Knoten miteinander in Verbindung stehen. Jeder dieser Knoten repräsentiert eine zufällige Variable. Oft wird auch von einem Eltern-Kind-Verhältnis gesprochen, wobei der Elternknoten den Kindknoten direkt beeinflusst. Diese Einflüsse werden durch sogenannte bedingte Wahrscheinlichkeiten quantifiziert.

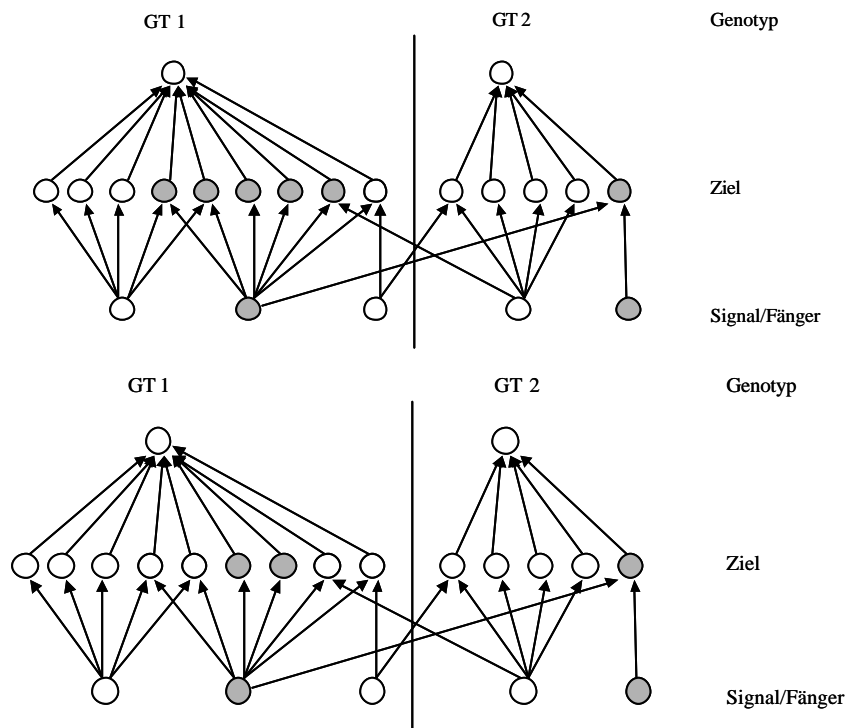


Abbildung 1-5: Oben: Darstellung des prinzipiellen Vorgehens bei der Auswertung mit Bayesschen Netzen. Es sind Vorkenntnisse aus der Konfigurierung der Bibliothek vorhanden, die durch die Pfeile verdeutlicht werden. Die ausgefüllten Kreise entsprechen den Fluoreszenzsignalen, die nach der Hybridisierung erhalten wurden. Somit könnten die ausgefüllten Kreise auf der Zielebene (Konfiguration) auf die entsprechenden Genotypen hindeuten. Unten: Ausschluss von drei Zielnukleinsäuren durch Bayessche Netze, da das zweite erwartete Signal nach der Hybridisierung ausgeblieben ist.

Durch entsprechende Programmierung und Vorkenntnisse durch die Konfigurierung der Software-generierten Bibliothek (Redundanzniveau, Lage und Anzahl der Mismatches etc.) konnten die durch die Hybridisierung erhaltenen Signalintensitäten durch die Verwendung der Bayesschen Netze in der Auswertung mit der Software *iOmega* die wahrscheinlichsten Ergebnisse ermittelt werden. Abbildung 1-5 aus Volkmann *et al.* soll dies deutlich machen<sup>[63]</sup>. Für jeden Genotyp oder Subtyp wurde mehr als ein Oligonukleotid ausgewählt. Aufgrund dessen wurde erwartet, dass pro Hybridisierung entsprechend dem verwendeten HCV-Geno- oder Subtyp mehr als ein Oligonukleotid detektiert sollte. Auf der Ebene der Zielnukleinsäuren wird dies durch die Pfeile bzw. der dunkel eingefärbten Felder deutlich (Abbildung 1-5). Das Bayessche Netz wurde mit dem Wissen aus der Konfigurierung ausgestattet werden, d.h. welche Oligonukleotide für welche Genotypen designed wurden, welche sie *in-silico* treffen und welche Genotypen sie mit bzw. ohne mismatch erkennen<sup>[64]</sup>.

Die Software *iOmega*<sup>[65]</sup> von der Firma *iSenseIt*, Bremen, hat ihr Auswerteprogramm auf den Bayesschen Netzen aufgebaut. Mit Hilfe dieser Software war es möglich, die Hybridisierungsdaten nicht nur mit manuellen Werkzeugen, sondern auch bioinformatisch auszuwerten.

## 1.2 Hepatitis C-Virus (HCV)

Hepatitis C-Viren wurden erstmals 1989 von Choo *et al.* entdeckt <sup>[66]</sup>. Es gilt als Meilenstein in der Molekularbiologie, da die Identifizierung dieser Non-A und Non-B Hepatitiden (NANBH), mit molekularen Klonierungstechniken gelang. Die HCV Infektion hat epidemische Ausmaße angenommen. HCV ist weltweit verbreitet - ca. 1 – 2 % der Weltbevölkerung sind chronisch infiziert. Das sind ca. 170 Millionen Menschen – viermal mehr als Patienten mit HIV <sup>[67]</sup>. Eine Studie der Weltgesundheitsorganisation (WHO) hat gezeigt, dass vor allem in Afrika (5,3 % der Bevölkerung), im östlichen Mittelmeerraum (4,6 %) und in der Westpazifikregion (3,0 %) die Zahl der Infizierten signifikant höher ist als in Nordamerika und Europa (Abbildung 1-6). Schätzungen für Europa gehen von 1,2 bis zu 5 Millionen HCV-positiven Personen aus. Bei Blutspendern in europäischen Ländern wurden Antikörper gegen das Hepatitis C-Virus (Anti-HCV-Antikörper), die als Marker einer Durchseuchung dienen, mit einer Häufigkeit von 0,23 % in Skandinavien bis 1,15 % in Italien beschrieben. In Deutschland liegt die Durchseuchung in der Bevölkerung bei mindestens 0,4 %, 84 % davon sind auch Virusträger.

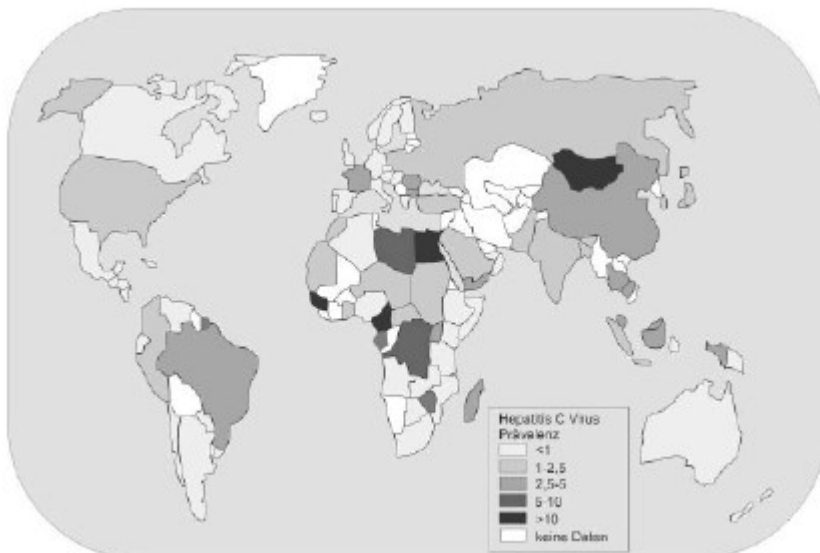


Abbildung 1-6: Globale HCV Prävalenz aus Zein <sup>[68]</sup>.

60-80 % der Infektionen mit HCV werden chronisch. Im Langzeitverlauf von 10-20 Jahren entwickeln 20-30 % der Patienten eine Leberzirrhose. 20-25 Jahre nach Infektion treten gehäuft Leberzellkarzinome auf. Bis heute gibt es keine zuverlässige Therapie für Hepatitis C. Es gibt Ansätze die in einigen Fällen erfolgversprechend erscheinen: Ribavirin

in Kombination mit pegyliertem Interferon alpha 2b oder alpha 2a. Das Ribavirin ist ein Basenanalogen und verlangsamt durch den Einbau in das Virusgenom die virale Reproduktion. Diese Wirkungsweise macht deutlich, dass die Therapie nicht ohne schwere Nebenwirkungen angewendet werden kann. Auch schlägt die Therapie nicht immer an. Es gibt Voraussetzungen, die erfüllt werden müssen, um die Therapie zu beginnen: der Patient sollte weiblich sein, nicht älter als 40 Jahre, keine Leberzirrhose haben, weniger als 2 Mio. Kopien/ml Blut haben und mit Genotyp 2 oder 3 infiziert sein. Der letzte Punkt macht die Notwendigkeit einer Klassifizierung des HCV deutlich, auf die in Abschnitt 1.2.2 näher eingegangen wird.

### 1.2.1 Das HCV Genom

HCV gehört zur Familie der Flaviviridae und zeigt markante Ähnlichkeit zu Mitgliedern aus den Gattungen der humanen Pestiviren und tierischen Flaviviren. Das Genom des HCV besteht aus einem positiven, einzelsträngigen RNA-Strang und ist etwa 9400 Basen lang. Es kodiert für ein Offenes Leseraster (open reading frame, ORF), welches am 3' und am 5' Ende von nichttranslatierenden Regionen (UTR) flankiert wird. Das Polyprotein wird nach der Translation in die einzelnen, funktionellen Proteine gespalten.

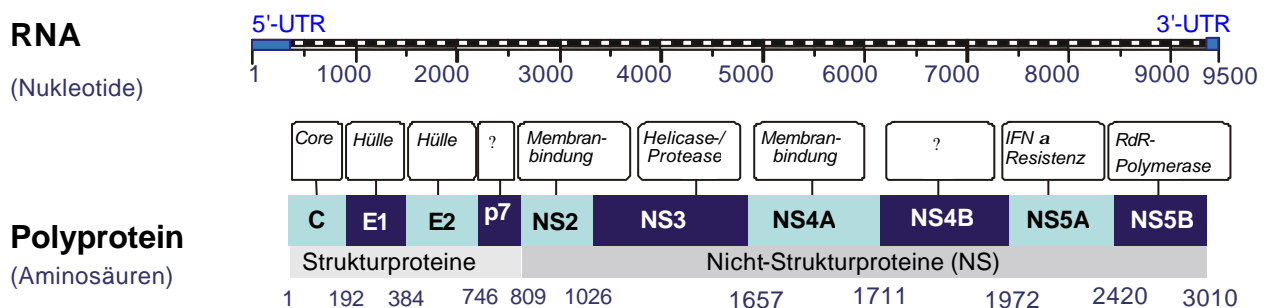


Abbildung 1-7: Schematische Darstellung des HCV Genoms.

Das HCV-Genom ist in Abbildung 1-7 schematisch dargestellt und soll im Folgenden erläutert werden. Die hochkonservierte Region der **5' untranslatierten Region (5' UTR)** ist ca. 340 Nukleotide lang. Sie beinhaltet mehrere stem-loop-Strukturen, die die internal ribosome binding site (IRES) bilden, die wiederum die CAP-unabhängige Translation der viralen RNA vermittelt <sup>[69]</sup>. Eine Promoterstruktur wurde gefunden, was darauf hindeutet, dass während des viralen Replikationszyklus DNA mit Hepatitis C-Virus Sequenzen synthetisiert wird <sup>[70]</sup>. Das **C-Protein** ist in der Lage, RNA zu binden und zu polymerisieren, daher könnte es sich mit der genomische RNA verbinden, um als

Komponente für das virale Nukleokapsid zu dienen <sup>[71]</sup>. Darüber hinaus ist es in verschiedene zelluläre Prozesse involviert. So konnte gezeigt werden, dass es die Transkription von Hepatitis B-Viren (HBV) und die Replikation des human immunodeficiency virus (HIV) inhibiert <sup>[72]</sup>. Die beiden Proteine **E1** und **E2** haben an ihrem C-Terminus hydrophobe Domänen, die wahrscheinlich als Membrananker dienen. E1 und E2 formen nicht-kovalente Heterodimere, welche eine funktionale Untereinheit der Virushülle repräsentieren. Sequenzvergleiche von E2 haben zwei hypervariable Regionen, HVR1 von aa 390-aa 410 und HVR2 von aa 474-aa 480, aufgezeigt <sup>[73]</sup>. Die hohe Mutationsrate, die mit dieser Region assoziiert ist, könnte das Produkt des selektiven Drucks durch das Immunsystem des Wirts sein <sup>[74]</sup>. Bei **NS2** handelt es sich um ein hydrophobes Transmembranprotein, das aus dem Polyprotein durch zwei proteolytische Reaktionen freigesetzt wird. Der N-Terminus wird durch eine Signalpeptidase des Wirtes vom E2/p7 Polyprotein getrennt <sup>[75]</sup>. Die **NS2/NS3** Bindung wird durch eine viruscodierte Proteinase, der sogenannten NS2-3 Proteinase, gespalten <sup>[76]</sup>. NS3 ist ein multifunktionales Molekül. Der N-Terminus enthält eine Serinproteinase, während der C-Terminus für die RNA-abhängige NTPase/Helikase von Bedeutung ist. Die Serinproteaseaktivität von **NS3** ist ein attraktives Ziel für die Entwicklung von Medikamenten für Antikörper gegen HCV (anti-HCV) <sup>[74, 77]</sup>. **NS4A** wird als Cofaktor der Proteinase für eine effiziente Proteolyse benötigt. **NS4B** gehört zu den sehr hydrophoben Proteinen, wahrscheinlich ist es membranassoziiert. Seine Funktion ist noch unbekannt. Aus **NS5A** entstehen zwei cytoplasmatische Phosphoproteine. Zwar ist die Funktion von NS5A in der viralen Replikation unbekannt, aber es gibt Beweise, dass NS5A in die Resistenz von HCV für Interferon alpha (IFN $\alpha$ ) involviert ist <sup>[78]</sup>. Nach Sequenzvergleichen wurde eine Region zwischen aa 2209 und aa 2248 in Verbindung mit der Antwort auf IFN $\alpha$ -Behandlung und Mutationen in dem Bereich gefunden. Diese Region ist als ISDR (interferon sensitivity determining region) bekannt <sup>[79]</sup>. NS5A ist ein potenter Inhibitor der Protein Kinase (PKR) <sup>[80]</sup>. Die PKR gehört zu den IFN-induzierten Proteinkinasen. PKR phosphoryliert die  $\alpha$ -Untereinheit des eukaryotischen Translationsinitiationsfaktors 2 nach Aktivierung durch doppelsträngige RNA. Das Resultat ist die Inhibition der Proteinsynthese und der Tod der infizierten Zelle. NS5A blockiert die Kinaseaktivität durch die Behinderung der PKR Dimerisierung. Beim **NS5B** handelt es sich um eine RNA abhängige RNA-Polymerase (RdRp). Die **3'UT-Region** kann wiederum in drei Teile untergliedert werden. Zunächst gibt es einen Bereich von 28-42 Nucleotiden, die schwach konserviert sind. Anschließend

folgt ein Polypyrimidin Bereich, der vor allem aus poly-Uracil und einzelnen Cytosinen besteht. Der dritte Teil ist der sogenannte X-trail. Es handelt sich um eine gut konservierte Region, die 98 Nukleotide enthält und wahrscheinlich eine sehr stabile stem-loop-stem Struktur hervorbringt <sup>[81]</sup>. Die eigentliche Funktion konnte noch nicht geklärt werden. Der X-trail scheint für die virale Replikation absolut notwendig zu sein <sup>[82]</sup>. Neuere Untersuchungen deuten auf die Existenz eines weiteren Proteins namens alternate reading frame protein (ARFP) hin. Walewski *et al.* <sup>[83]</sup> entdeckten in der C-Region des HCV Genoms ein cluster mit unüblich konservierten Codons, die auf ein potentielles überlappendes ORF hindeuten. Aufgrund seines Ursprungs in der Core-Region könnte dieses Protein die spezifische Immunantwort stimulieren und wäre so wertvoll für diagnostische Methoden und als Komponente für Impfstoffe. Das Protein soll hoch basisch sein und könnte eine Rolle in der HCV-Replikation, Pathogenese und in der Entwicklung von Leberkrebs spielen.

### **1.2.2 Klassifizierung**

HCV-Isolate aus der ganzen Welt wurden untersucht und zeigten substantielle Unterschiede in der Nukleotidsequenz <sup>[84]</sup>. Anhand verschiedener Sequenzabschnitte innerhalb des HCV-Genoms kann eine Klassifizierung vorgenommen werden. Die phylogenetische Einteilung beruht auf der Analyse der NS5-Region des HCV nach Simmonds *et al.* <sup>[85]</sup>. Eine Verfeinerung der Methode zur Klassifizierung schlagen Robertson *et al.* vor <sup>[86]</sup>. Hierbei soll nicht mehr nur die NS5-Region eine Rolle bei der Klassifizierung spielen, sondern auch das C-Protein, die E1-Region und die NS5B. Simmonds hat die verschiedenen Isolate des HCV in Genotypen und Subtypen eingeteilt. Die HCV Genotypen sind mit arabischen Nummern nach ihrer Entdeckung durchnummeriert (z.B. Genotyp 1). Subtypen sind nahe verwandte HCV Stämme innerhalb der Genotypen, die durch Kleinbuchstaben nach der genotypspezifischen Nummerierung gekennzeichnet sind (z. B. Subtyp 1a oder 1b). Eine weitere Unterteilung sind die sogenannten Quasispecies. Sie bilden einen Komplex aus genetischen Variationen, die durch Anhäufung von Mutationen während der viralen Replikation innerhalb eines Wirtsorganismus hervorgerufen werden können <sup>[87]</sup>.

Die Zahl der Sequenzeinträge steigt ständig. So sind zur Zeit 25584 Nukleotidsequenzen in NCBI eingetragen. Davon sind die meisten dem Genotyp 1 (7,2%) und dem Subtyp 1b (6,5 %) zugeordnet. In Genotypen werden solche Isolate eingeteilt, deren Sequenzähnlichkeit zwischen 65,7 % und 68,9 % liegt. Diese Ähnlichkeit bezieht sich auf

das Kompletengenom des HCV und nicht auf einzelne Abschnitte. Heute geht man von sechs verschiedenen Genotypen und ~90 verschiedenen Subtypen aus, die den Genotypen untergeordnet sind. Subtypen sind eng verwandte Isolate innerhalb eines Genotypes. Ihre Ähnlichkeit auf der Nukleotidebene liegt zwischen 76,9 % und 80,1 %.

Wie in anderen RNA-Viren auch (z.B. HIV), können in einem infizierten Individuum eine Population eng verwandter, jedoch heterogener Sequenzen (sog. Quasispezies) vorkommen [88]. Sie sind genetische Varianten innerhalb verschiedener Isolate aus einem Wirtsorganismus. Ihre Ähnlichkeit zueinander liegt zwischen 90,8 % und 99 %. Die biologische Konsequenz wird zum Einen die Etablierung einer chronischen Infektion als Folge einer Mutantenselektion, die den neutralisierenden Antikörpern oder den zytotoxischen T-Lymphozyten entkommen konnten und zum Anderen dem Versagen von Impfstoffen zugeschrieben [89]. Die Verteilung der Quasispezies ist ein Ergebnis der Mutationen, die sich während der viralen Replikation ereignen. Die Mutationsrate wird bei HCV mit  $1,44 \cdot 10^{-3}$  bis  $1,92 \cdot 10^{-3}$  Basensubstitutionen pro Genomposition pro Jahr angegeben [90, 91].

#### 1.2.2.1 Geographische Verteilung des Hepatitis C-Virus

Die Verbreitung der verschiedenen Genotypen unterscheidet sich regional. Darüberhinaus ist sie auch von der Art der Virusübertragung abhängig. Wie aus Abbildung 1-8 zu

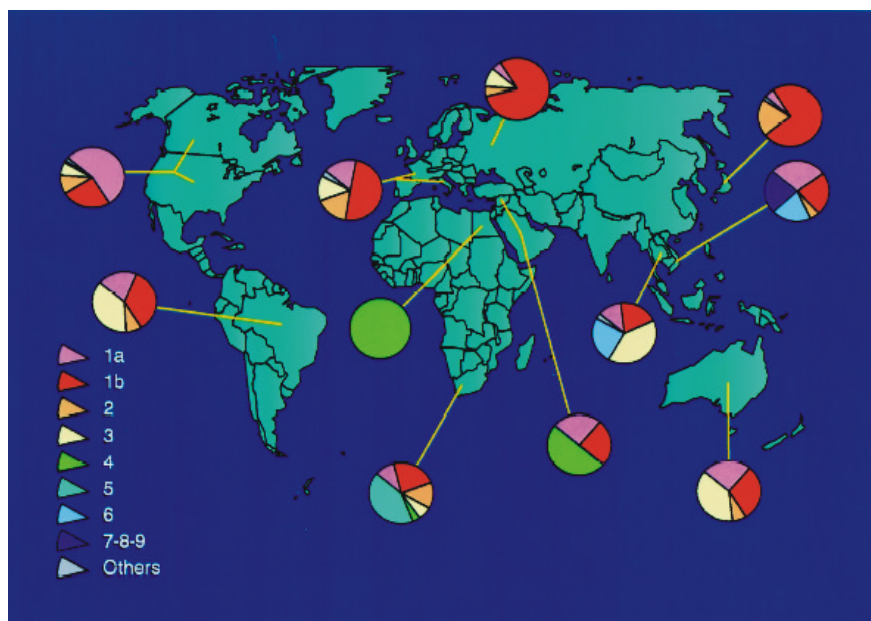


Abbildung 1-8: Weltweite geographische Verteilung der HCV Geno- und Subtypen, „Others“ sind unklassifizierte Sequenzen [87].



erkennen ist, treten die Subtypen 1a und 1b vor allem in den Vereinigten Staaten und in Europa auf. In Japan ist der Subtyp 1b sogar für 73 % der HCV Infektionen verantwortlich. Subtypen 2a und 2b kommen hauptsächlich in Nordamerika, Europa und Japan vor. Subtyp 2c hingegen in Norditalien. Der Genotyp 3 ist am häufigsten unter einer der Risikogruppen für die Übertragung von HCV, Drogenabhängigen, die sich ihre Drogen intravenös verabreichen, verbreitet. In Nordafrika und dem Mittleren Osten ist Genotyp 4 vorherrschend, während die Genotypen 5 und 6 auf Südafrika und Hongkong beschränkt sind. Die nun folgenden Varianten werden zwar als Genotypen bezeichnet, werden jedoch als Varianten desselben Genotypes angesehen und dem Genotyp 6 zugeordnet. Zu diesen Varianten zählen die Genotypen 7, 8 und 9, die vor allem in Vietnam vorkommen.

Die geographische Verteilung und die Diversität der HCV Genotypen könnten Schlüsse auf den Ursprung zulassen. Simmonds <sup>[92]</sup> datiert den Ursprung des Virus auf vor 500-2000 Jahren.

### 1.2.3 Verwendete Verfahren zur Hepatitis C-Virus-Diagnose

Die Tests zur Diagnose der Hepatitis C-Viren sind nicht standardisiert. Es gibt zwei Kategorien von Tests für die virologische Diagnose und Überwachung. Die serologischen Assays detektieren spezifisch anti-HCV Antikörper (indirekte Tests) <sup>[93]</sup>. Direkte Tests detektieren, quantifizieren oder charakterisieren Komponenten der Hepatitis C-Virus Partikel, z.B. Hepatitis C-Virus RNA oder das Core-Antigen.

#### Serotypisierung

- Anti-HCV Detektion über EIAs (Enzyme Immuno Assays), z.B. Ortho Hepatitis C-Virus 3.0 von Ortho-Clinical Diagnostics, oder über ELISAs (Enzyme Linked Immuno Sorbant Assays), z.B. Abbott Hepatitis C-Virus EIA 2.0 von Abbott Diagnostic: diese Tests detektieren eine Mischung aus Antikörpern gegen verschiedene HCV Epitope der Core-, NS3-, NS4- und NS5-Proteine. Virale Antigene werden auf verschiedene Oberflächen, z.B. Mikrotiterplatten, Microbeads aufgebracht. Erfahrungen aus klinischen Studien deuten darauf hin, dass EIAs in 99 % der Fälle immunkompetenter Patienten positiv sind. Die Tests könnten jedoch bei Hämolyse-Patienten oder Patienten mit geschwächtem Immunsystem falsch negativ ausfallen.
- Bei der serologischen Bestimmung der Hepatitis C-Virus Genotypen (serotyping) (Murex Diagnostics) mittels kompetitiven EIAs mit Hilfe typspezifischer Antikörper

für die NS4-Region können die HCV Genotyp zugeordnet, nicht aber nach Subtypen klassifiziert werden. Bei 90 % der immunkompetenten Patienten mit chronischer Hepatitis C konnten interpretierbare Ergebnisse erzielt werden.

Nachteilig bei den indirekten Tests ist das Unvermögen, zwischen einer vergangenen und akuten Infektion zu unterscheiden.

### Genotypisierung

Molekulare Bestimmung der Hepatitis C-Virus Genotypen (genotyping):

- Sequenzierung der NS5B oder E1 Region gilt als „goldener Standard“. Die Sequenz wird anschließend mit Referenzsequenzen und phylogenetisch analysiert <sup>[87, 94]</sup>.
- Direkte Sequenzanalyse mit z.B. Trugene HCV 5'NC von Visible Genetix. Nach der Isolation der viralen RNA dient diese als Templat in einer RT-PCR Reaktion. Das amplifizierte Material wird anschließend in einer CLIP Reaktion markiert: das in einer PCR entstehende Amplifikat wird parallel sequenziert. Mit dieser Methode entstehen durch den Einsatz zweier fluorophormarkierter Primer durch zyklische Amplifizierung bidirektionale Sequenzen. Diese beiden Sequenzen sind 183 Nukleotide lang <sup>[84]</sup> werden kombiniert und gegen eine Referenzsequenz mit ClustalW aligniert. Da die Methode auf der Untersuchung der 5' UTR basiert, ist die Zuordnung zu den Subtypen nicht immer möglich <sup>[95, 96]</sup>. Ein Trugene Kit für die NS5B ist in Entwicklung.
- RFLP (Restriktions Fragment Längen Polymorphismus) ist eine enzymatische Methode, die durch den Einsatz von Restriktionsendonukleasen die 5'UTR des HCV anhand von Bandenmustern in einem Agarosegel genotypisiert. Davidson <sup>[97]</sup> benutzte hierfür die Endonukleasen HaeIII-RsaI und HinfI-MvaI, gefolgt von einem Verdau mit BstU1 oder ScrFI, um die Genotypen 1a, 1b, 2a, 2b, 3a, 3b, 4, 5 und 6 zu identifizieren und klassifizieren. Diese Methode muß kontinuierlich angepasst werden, um den stetig wachsenden Sequenzdaten gerecht zu werden.
- Pyrosequencing ist ein schnelle Methode zur preisgünstigen, für den Hochdurchsatz geeigneten und akkuraten Genotypisierung <sup>[98]</sup>. Ein 237 Nukleotide langes Fragment aus der 5' UTR wird mit einem biotinmarkierten und einem unmarkierten Primer isoliert. Es folgt eine Abtrennung des biotinmarkierten Stranges über Streptavidin-Magnetic Beads, eine Einzelstrangisolation, die Hybridisierung der Sequenzierprimer und anschließend das Pyrosequencing. Ein Multiplex-Ansatz wird von Nader Pourmand beschrieben <sup>[99]</sup>.

- DEIA (DNA enzyme immunoassay, von Gen-Eti-K DEIA, Sorin Biomedica). Hierbei werden Fragmente der Core Region mit einer Nested PCR amplifiziert und die resultierenden Amplikons (250 Bp) werden gegen Typ- oder Subtypen spezifische Oligonukleotide hybridisiert, die in den wells von Mikrotiterplatten adsorbiert sind <sup>[100]</sup>. Ein Vergleich mit InnoLipa <sup>[101]</sup> zeigte eine Diskrepanz bei der Subtypisierung von Genotyp 1 und 2 auf. Mit DEIA konnten die Subtypen der Genotypen 2 zweifelsfrei zugeordnet werden, während dies mit INNO-LiPA (s. u.) nicht gelang.
- Reverse Hybridisierung an Genotyp-spezifischen Oligonukleotiden (INNO-LiPA-Hepatitis C-Virus II). Hierbei handelt es sich um ein System, bei dem auf Nitrocellulosestreifen genotypspezifische Oligonukleotide, ausgewählt anhand der 5'UTR des Hepatitis C-Virus aufgebracht sind. Amplifizierte und biotinylierte DNA wird mit diesen Oligonukleotiden hybridisiert, die als parallele Banden auf den Membranstreifen immobilisiert sind. Während der Amplifikation wird die zu untersuchende Sequenz mit Hilfe eines 5'-biotinylierten Primers biotinyliert. Das markierte Produkt aus der 5' UTR hybridisiert nur mit Sonden, deren Sequenz genau passt, so dass auch Subtypen eindeutig voneinander unterschieden werden können. Für diese hohe Spezifität ist das Arbeiten unter stringenten Hybridisierungsbedingungen von  $50\text{ °C} \pm 2\text{ °C}$  erforderlich. Nach der Hybridisierung wird mit alkalischer Phosphatase markiertes Streptavidin zugegeben, das an zuvor biotinylierte Hybride bindet. Die darauf folgende Inkubation mit dem Chromogen BCIP/NBT führt zu einem bräunlich-violetten Niederschlag. Mit Hilfe einer Interpretationsvorlage aus Kunststoff kann aus der Summe aller positiven Banden der Genotyp des Hepatitis C-Virus bestimmt werden. Mit diesem Test können die sechs wichtigsten Genotypen und deren häufigste Subtypen gleichzeitig bestimmt werden. Die theoretisch zuverlässigste Möglichkeit, die Genotypen in Hepatitis C-Virus-Isolaten zu bestimmen, würde in der Sequenzierung des gesamten RNA-Genoms bestehen. Der INNO-LiPA Hepatitis C-Virus II Test basiert auf der Nukleotidvariation in der 5'UTR, und dieser Test erzielt eine hundertprozentige Übereinstimmung mit Sequenzierungen, die an dieser Region vorgenommen wurden <sup>[102]</sup>. Vergleichende Studien haben gezeigt, dass dieser Test zwischen den Subtypen 2a und 2c sowie in 5-10 % der Fälle zwischen den Subtypen 1a und 1b nicht unterscheiden kann <sup>[103]</sup>. Darüber hinaus sind verschiedene neue <sup>[104]</sup> oder atypische Muster <sup>[105]</sup> beschrieben worden. Dieses macht die Variabilität des Virus deutlich und zeigt, dass ständige Kontrolle der Tests notwendig ist.

- Mittels eines Light-Cycler Systems etablierten Schröter *et al.* <sup>[106]</sup> eine Methode zur Klassifizierung der Genotypen 1, 2, 3 und 4 mit drei verschiedenen Paaren von Hybridisierungsprimern aus der 5'UTR. Der Typisierung geht eine Amplifizierung voraus.

All diese Methoden sind PCR-basiert und daher verlässlich, sodass die Daten für die molekulare Pathogenese relevant sind. Nachteilig ist allerdings, dass sie teuer und zeitaufwendig sind und spezielle Anschaffungen nötig machen, um die Verlässlichkeit der Ergebnisse zu gewährleisten, und um Kontaminationen zu vermeiden.

### Quantifizierung

- Branched DNA (bDNA) (Quantiplex HCV RNA assay, Chiron Diagnostics). Bei dieser Technik wird keine PCR zur Amplifikation der RNA durchgeführt. Die notwendige Sensitivität wird durch Signalverstärkung erreicht. Hierbei werden verschiedene DNA Proben verwendet, für die Detektion und für die Erhöhung der Anzahl potentieller Bindungsstellen, von denen ggf. das Reportersignal generiert wird <sup>[107]</sup>. Die Methode basiert auf einer vergleichenden Amplifikation des viralen Templates mit einer bekannten Menge eines synthetischen Standards, der zu jedem Reaktionsgefäß gegeben wird. Die relativen Mengen an Virustemplet und Standardamplikon werden am Ende der Reaktion gemessen und die Resultate werden mit Kurven des parallel gelaufenen Standards verglichen. Falsch positive Resultate von 3 % zwischen 0.2 und 0.5 MEq/ml (= 1000 Equivalente pro ml), definiert als die Menge an HCV RNA, die die gleiche Menge Licht erzeugen wie  $10^6$  Kopien des HCV-RNA-Referenzstandards <sup>[108]</sup> werden beschrieben <sup>[109, 110]</sup>. Die dynamische Bandbreite bei der Quantifikation liegt bei Versant HCV-RNA 3.0 bei 615-7.700.000 IU/ml <sup>[93]</sup>. Jedoch sind spezielle Vorschriften bei der Probengewinnung einzuhalten, um falsch negative Ergebnisse zu vermeiden.
- Reverse Transkriptase PCR (Cobas Amplicor HCV Monitor, Roche Diagnostic Systems). Dieser quantitative Test amplifiziert die HCV-RNA aus Serum oder Plasma. Der Test basiert auf einem vergleichenden PCR-Assays. Die untere Nachweisgrenze liegt bei 600 IU/ml. Hierbei werden die Viruslast im Serum oder Plasma nicht am Ende der Reaktion, sondern während der Reaktion abgelesen. Diese Methode ist sensitiver als die o.g. Methode der bDNA.

Beide Methoden sind weniger sensitiv als qualitative Tests. Sie sollten nicht für die Diagnose einer HCV Infektion oder für die Endpunktbestimmung der Therapie eingesetzt werden.

- Eine Quantifizierung mit TaqMan Proben wurde 1999 von Kawai *et al.* beschrieben <sup>[111]</sup>. Auch hier lagen die Primer und die Probe in der 5' UTR.

#### Qualitative Tests mittels RT-PCR

- Inhouse Methoden mittels RT-PCR wurden von verschiedenen Gruppen beschrieben. Die Methode von Lin *et al.* <sup>[112]</sup> detektiert HCV-RNA der 5'UTR in Blut mittels einer reversen Transkription und anschließender Amplifikation mit nested Primern in einem einzelnen Reaktionsgefäß.
- Cobas Amplicor HCV, Roche Diagnostic Systems. Die Detektion des Virus kann bereits ein bis zwei Wochen nach der Infektion erfolgen. Falsch positive und falsch negative Ergebnisse können auftreten. Die untere Nachweisgrenze liegt hier bei 50 UI/ml <sup>[113]</sup>.
- Versant HCV-RNA Qualitative Assay, Bayer Corp., ist ein Assay, der auf der Amplifizierung des Targets (TMA; engl: target amplification assay) beruht <sup>[114]</sup>. Die Nachweisgrenze liegt bei 10 IU/ml <sup>[115]</sup>. Diese Reaktion ist isothermal und verwendet zwei Enzyme: eine Reverse Transkriptase und eine T7 RNA-Polymerase. Die Reaktion läuft über immobilisierte Oligonukleotide, deren Hybridbildung eine enzymatische Reaktion hervorruft, die durch ein farbiges oder lumineszentes Signal detektiert wird.

Cobas und Versant erreichen eine Spezifität von 98 % - 99 % <sup>[116]</sup>.

Eine Co-Detektion von u.a. HCV Antikörpern und Nukleinsäuren wurde 2003 von Perrin beschrieben <sup>[117]</sup>. Dieser kombinierte Immuno- und Oligosorbent Mikroarray (ComboOLISA) besteht aus einer 96-well Mikrotiterplatte, in die spezifische Oligonukleotide aus der 5'UTR zur Detektion des Virusgenoms, virale Proteine, um die Serumantikörper zu fangen und nichtspezifische Proteine zur Verifizierung der Spezifität durch Spotten aufgebracht sind. Die Resultate waren mit Standard ELISAs vergleichbar.

## 2 AUFGABENSTELLUNG

HCV mit seiner enormen Variabilität auf verschiedenen taxonomischen Ebenen (Genotypen und Subtypen) dient hier als Modell für Untersuchungen zum Einsatz von Mikroarrays bei der Identifikation von Organismen auf Einzelbasenaustauschen. Die Möglichkeit, auf der Größe eines Objektträgers mit relativ einfachen Mitteln schnell und sicher das Vorhandensein und gleichzeitig die Klassifizierung von Hepatitis C-Viren aus einer Serum- oder Plasmaprobe festzustellen, könnte mit Hilfe der Mikroarrays gegeben sein. Für die Diagnose von HCV wäre dies ein entscheidender Vorteil. Diese Arbeit soll zeigen, ob Mikroarrays dafür geeignet sind. Ein Vergleich manuell ausgewählter Oligonukleotide und bioinformatisch determinierter Sonden soll darüber hinaus zeigen, ob und inwiefern die Bioinformatik die Generierung von Oligonukleotid-Bibliotheken übertreffen kann. Eine Übersicht, wie die Arbeit aufgebaut ist, ist in Abbildung 2-1 skizziert.

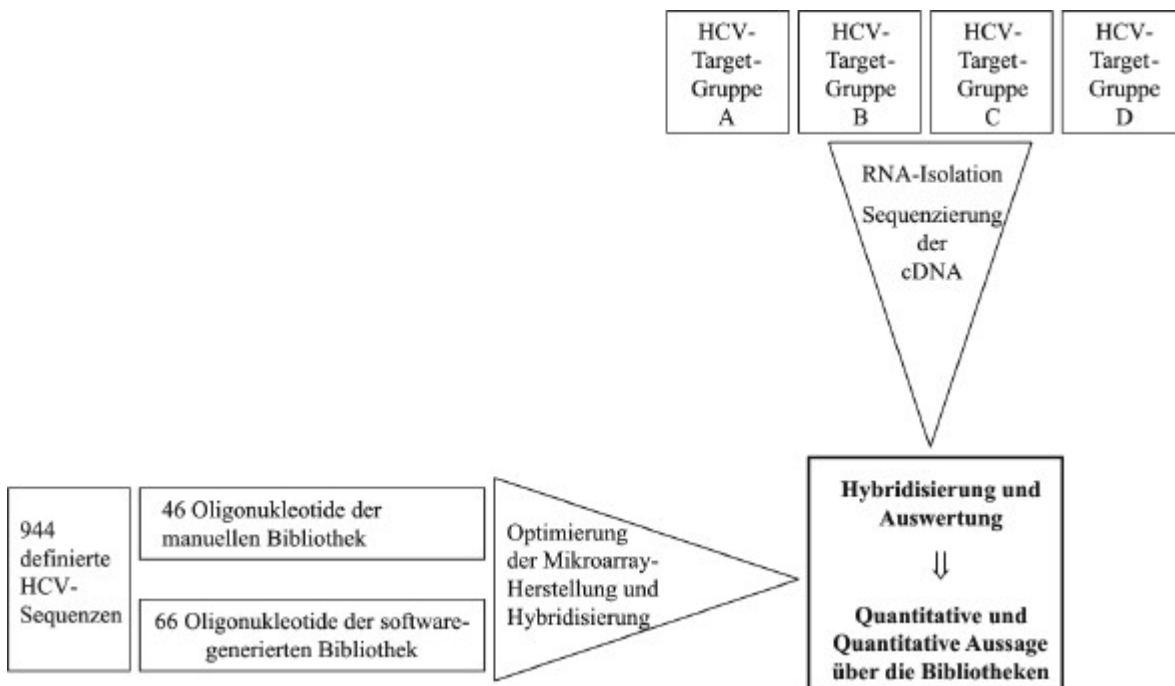


Abbildung 2-1: Übersichtsskizze über den Aufbau der Arbeit. Das Ziel ist ein Mikroarray zur Identifizierung und Klassifizierung von Hepatitis C-Viren. Dazu wurden auf verschiedenen Wegen, aber derselben Datenbasis Oligonukleotid-Bibliotheken generiert. Nach Optimierung der Spotting- und Hybridisierungsparameter wurden verschiedene HCV-Isolate unterschiedlicher Geno- und Subtypen mit den Oligonukleotiden hybridisiert und die erhaltenen Signale anschließend analysiert und bewertet.

## 3 MATERIAL UND METHODEN

### 3.1 Design der Oligonukleotid-Bibliotheken

Die Auswahl der Oligonukleotide ist zur Differenzierung der verschiedenen Genotypen und Subtypen des HCV von entscheidender Bedeutung. Für diese Arbeit wurden Oligonukleotide mit zwei verschiedenen Methoden designt. Zum Einen manuell über Alignments und Programme zur Oligonukleotid-Auswahl, die Vorarbeiten wurden hierzu von Dr. H. Gersdorf durchgeführt. Zum Anderen über eine für diese Aufgabe entwickelte Software.

#### 3.1.1 Manuell designte HCV-Oligonukleotid-Bibliothek

##### 3.1.1.1 Datenquelle

Die Basis für die Auswahl der manuell designten Oligonukleotid-Bibliothek 944 Sequenzen der Kompletengenome des Hepatitis C-Virus, die 1998 in den Datenbanken im National Center for Biotechnology Information (NCBI) <sup>[118]</sup> und der DNA Data Bank of Japan (DDBJ) <sup>[119]</sup> enthalten waren.

##### 3.1.1.2 Alignments

Die Sequenzen aller 5'UTRs wurden über das SRS des European Molecular Biology Laboratory (EMBL) <sup>[120]</sup> eingeholt und die Sequenzlänge auf 360 bp angepasst. Es wurden verschiedene Alignments mit ClustalW <sup>[121]</sup> erarbeitet. Ein umfassendes Alignment, das alle 944 5'UTR-Sequenzen enthielt und mehrere Teilalignments, die nur die genotypspezifischen Sequenzen enthielten. Die Oligonukleotide wurden anhand der sieben literaturbekannten Regionen innerhalb der 5'UTR positioniert <sup>[122]</sup>.

##### 3.1.1.3 Design der Oligonukleotide

Die Bindungsstellen HCV-universeller sowie Genotypspezifischer „Master-Oligonukleotide“, die diesen Regionen zugeordnet wurden, wurden innerhalb der Alignments auf vorkommende Sequenzvarianten überprüft. Varianten dieser Master-Oligonukleotide wurden aus dem Alignment extrahiert und in einer Liste zusammengestellt. Die erhaltenen Oligonukleotid-Kandidaten wurden auf folgende Parameter optimiert:

- Angleichen der Schmelztemperatur auf  $60\text{ °C} \pm 2\text{ °C}$  nach Breslauer <sup>[51]</sup> durch Veränderung der Oligonukleotidlänge mittels der Software *Primer calculator* <sup>[123]</sup>.
- Ermittlung der Sekundärstruktur mittels *mfold* <sup>[124]</sup>. Vermeidung niedriger  $\Delta G$ -Werte, die stabile intramolekulare Bindungen der Oligonukleotide ermöglichen und so die Zugänglichkeit während der Hybridisierung verringern.
- Spezifitätskontrolle der Oligonukleotide mittels FASTA <sup>[125]</sup> und BLAST <sup>[126]</sup> des NCBI.

### 3.1.2 Software-generierte HCV-Oligonukleotid-Bibliothek

Die Software *OptiNA* wurde von Dr. M. Nölte entwickelt <sup>[64]</sup> und steht durch eine Benutzerverwaltung mit Login und Passwort unter <sup>[127]</sup> zur Verfügung.

Als zugrundeliegender Algorithmus wurde der Greedy-Algorithmus <sup>[128]</sup> verwendet. Insgesamt 16 Parameter gehen in die Auswahl und Bewertung der Oligonukleotid-Kandidaten ein. Diese sind die Schmelztemperatur, die Länge, der GC-Gehalt, die Temperatur, der Salzgehalt, die DNA/RNA Parametersätze der „nearest neighbor interactions“, „ $\Delta G$  correction terms“, „Mikroarray correction term“ für die Bewertung der Hybridisierungseffizienzen. Toleranzniveau für ggf. falsch-positive Treffer, Redundanzniveau für die Anzahl der Zielsequenzen, die das Oligonukleotid erkennt, Anzahl der Treffer der Zielsequenzen, Anzahl der Treffer auf Nichtzielsequenzen und die Hierarchie der gesuchten Organismen. Diese Daten sind die Grundlage für die Auswertung der Hybridisierungsdaten mit Hilfe der Software *iOmega*, der Firma *iSenseIt*, Bremen.

Die Datengrundlage sind hier die 944 HCV-Sequenzen, wie sie für die manuelle Bibliothek zusammengetragen wurden. Bei der Erzeugung von Oligonukleotid-Kandidaten wurden die Bindungsregionen nicht berücksichtigt. Es wurden alle Oligonukleotide einer vorgegebenen Länge (15 bis 30 nt) betrachtet und für  $T_M$ -Intervalle von 58-62 °C die Oligonukleotid-Kandidaten berechnet. Die Optimierung der Oligonukleotide bezüglich der Sekundärstruktur wurde der sog.  $\Delta\Delta G$ -Ansatz verwendet. Dazu wird mit dem Oligonukleotid, der Zielsequenz und der Position auf der Zielsequenz ein  $\Delta G_1$  berechnet, unter der Bedingung, dass keine Basenpaarungen an der Position der Zielsequenz zugelassen sind. Ein zweites  $\Delta G_2$  wird berechnet, bei dem jedoch diese Bedingung aufgehoben ist.  $\Delta\Delta G(x,t) = \Delta G_1 - \Delta G_2$ . Die Oligonukleotide sollten möglichst in einem Bereich mit wenig Sekundärstruktur liegen, diese sind primär loop-Bereiche. Die Optimierung des Redundanz-Niveaus sollte soweit möglich erfolgen. Bevorzugt wurden



Oligonukleotide mit mindestens zwei Basenfehlpaarungen zu den Nichtziel-Sequenzen ausgewählt.

### 3.2 Primer, Oligonukleotide und HCV-Proben

Die Synthese aller Oligonukleotide und der Primer HCV 90 und HCV 88 wurde bei Thermo Electron GmbH, Ulm, in Auftrag gegeben. Die Primer für die RT-PCR wurden von MWG Biotech synthetisiert.

Tabelle 1: Verwendete Oligonukleotide.

Oligonukleotid	Sequenz 5' -> 3'	Markierung
<b>Manuelle Bibliothek</b>	Siehe Tabelle 8	5' Amino (C6/MMT)
<b>Software Bibliothek</b>	Siehe Tabelle 9	5' Amino (C6/MMT)
<b>A-M13cA-Cy5</b>	CGG ATA ACA ATT TCA ACA AGG A	5'Cy5; 3' Amino (C7)
<b>M13A-Cy5</b>	TCC TGT GTG AAA TTG TTA TTC G	5'Cy5

Tabelle 2: Verwendete Primer.

Primer	Sequenz 5' -> 3'	Lage auf M62321	Markierung
<b>HCV 90 (nested)</b>	GCA CTC GCA AGC ACC CTA	300-317	5' Cy5
<b>HCV 88 (nested)</b>	TGC AGC CTC CAG GAC C	106-121	5' Biotin
<b>RT-antisense</b>	GAG GTT TAG GAT TCG TGC TC	347-366	Keine
<b>RT-sense</b>	CTC CAC MAT GAA TGA GTC CC	25-44	Keine
<b>RT-n-antisense</b>	GAT GCA CGG TCT ACG AGA CC	325-344	Keine
<b>RT-n-sense</b>	CGT TAG TAT GAG TGT CGT GC	89-108	Keine

#### HCV-Targetgruppe A

In der Abteilung Biotechnologie und Molekulare Genetik (BMG) waren HCV-Proben in Form von PCR-Produkten vorhanden, die ursprünglich aus dem Labor Dr. Gärtner stammen. Mit vier Proben der Genotypen 1a, 1b, 3 und 4 wurden die Bibliotheken getestet.

#### HCV-Targetgruppe B

Dr. R. Frodl, Labor Dr. Gärtner, Weingarten, Deutschland, hat freundlicherweise mehrfach Serumproben HCV-infizierter Personen zur Verfügung gestellt.

### HCV-Targetgruppe C

HCV-Genotypisierungs-Panel des Referenzzentrums für Hepatitis C, Essen, Deutschland <sup>[129]</sup>. Die Isolate wurden mit verschiedenen Methoden quantifiziert und klassifiziert.

### HCV-Targetgruppe D

Frau Dr. B. Becker hat PCR-Produkte von HCV-positiven Plasmaproben zur Verfügung gestellt, die vom Blutspendedienst des DRK bzw. von einem Spender zur Verfügung gestellt wurden. Darüberhinaus konnte HCV-positives Serum, bereitgestellt von der Universität Göttingen, untersucht werden. In Tabelle 3 sind alle HCV-Proben aufgelistet.

Tabelle 3: Auflistung der verwendeten HCV Proben, ihre Spezifizierung, Herkunft und ggf. Quantifizierung und Genotypisierung.

Targetgruppe	Probe	Genotyp
A	HG 1a	1a
A	HG 1b	1b
A	HG 3a	3a
A	HG 4	4
B	RF 1b	1b
B	RF 4	4
B	RF 1a	1a
B	RF 3a	3a
B	RF 3a	3a
B	RF 1a	1a
B	RF 2a	2a
B	RF 1b	1b
B	RF 1a	1a
B	RF 3a	3a
B	RF 1b	1b
B	RF 3a	3a
B	RF 1b	1b
B	RF 2b	2b
B	RF 3s	3b
B	RF 3s	3b
B	RF 4	4

Targetgruppe	Probe	Genotyp	Quantifizierung		Genotypisierung	
			Cobas [Kopien/ml]	Quantiplex [Kopien/ml]	InnoLipa	Sequenzierung
C	P 1a	1a	$1,7 \times 10^5$	$2,1 \times 10^5$	1a/1b	1a
C	P 1b	1b	$5,3 \times 10^5$	$4,2 \times 10^5$	1b	1b
C	P 2a	2a	$2,6 \times 10^5$	$3,4 \times 10^5$	2a/c	2a
C	P 2b	2b	$6,4 \times 10^5$	$9,0 \times 10^5$	2b	2b
C	P 2i	2i	$3,7 \times 10^5$	$3,7 \times 10^5$	2i	2i
C	P 3a	3a	$2,1 \times 10^6$	$3,2 \times 10^6$	3a	3a
C	P 5a	5a	$3,9 \times 10^5$	$2,9 \times 10^6$	5a	5a
D	BB#1	unbekannt	HCV-positives Frischplasma vom DRK			
D	BB#2	unbekannt	HCV-positives Frischplasma vom DRK			
D	BB#3	unbekannt	HCV-positives Frischplasma vom DRK			
D	BB#4	unbekannt	HCV-positives Frischplasma eines Spenders			
D	BB#5	1b	HCV-positives Serum Universität Göttingen			
D	BB#6a	1a	Infektiöses HCV-Plasmid (H77)			
D	BB#6b	1a	Infektiöses HCV-Plasmid (H77)			

### 3.3 Verwendete Lösungen und Puffer

#### 10x TE

100 mM Tris-HCl pH 8,0 (Sigma)

1 mM Na<sub>2</sub>-EDTA pH 8,0 (Merck)

#### TBE

0,89 M Tris (Sigma)

0,89 M Borsäure (Fluka)

0,02 M Na<sub>2</sub>-EDTA pH 7-7,4 (Merck)

#### Spottingpuffer

H<sub>2</sub>O<sub>dd</sub> mit 1 % Glycerin

#### Blocking-Lösung

0,15 M Glutarsäure-Anhydrid (Sigma) in 1-Methyl-2-Pyrrolidon (Merck)

10 % 0,2 M Borsäure pH 8,0 (Fluka)

#### Prähybridisierungslösung

10 mM Tris-HCl pH 7,5 (Sigma)

1 mM EDTA pH 7,0 (Merck)

600 mM NaCl (Fluka)

1x Denhardt´s Lösung

500 µg/ml Heringsperma-DNA (Roche)

50 % (v/v) deionisiertes Formamid

#### 100x Denhardt

für 5 ml

0,1 g BSA

0,1 g Ficoll 400

0,1 g PVP 40000

4,7 g H<sub>2</sub>O

Deionisieren von Formamid

50 g AG501-X8D Resin (BioRad) mit einem Liter Formamid (Fluka) über Nacht unter Rühren deionisieren, filtrieren, aliquotieren und bis zum Gebrauch bei  $-20\text{ }^{\circ}\text{C}$  einfrieren.

Hybridisierungslösung 1

10 mM Tris-HCl pH 7,5 (Sigma)  
1 mM EDTA pH 7,0 (Merck)  
600 mM NaCl (Fluka)  
1x Denhardt´s Lösung  
500  $\mu\text{g/ml}$  Heringsperma-DNA (Roche)  
50 % (v/v) deionisiertes Formamid  
5 nM der HCV-Lösung

Hybridisierungslösung 2

10 mM Tris-HCl pH 7,5 (Sigma)  
1 mM EDTA pH 7,0 (Merck)  
600 mM NaCl (Fluka)  
1x Denhardt´s Lösung  
500  $\mu\text{g/ml}$  Heringsperma-DNA (Roche)  
50 % (v/v) deionisiertes Formamid  
10 nM der HCV-Lösung

Hybridisierungslösung 3

10 mM Tris-HCl pH 7,5 (Sigma)  
1 mM EDTA pH 7,0 (Merck)  
600 mM NaCl (Fluka)  
1x Denhardt´s Lösung  
500  $\mu\text{g/ml}$  Heringsperma-DNA (Roche)  
30 % (v/v) deionisiertes Formamid  
10 nM der HCV-Lösung

Hybridisierungslösung 4

10 mM Tris-HCl pH 7,5 (Sigma)  
1 mM EDTA pH 7,0 (Merck)  
600 mM NaCl (Fluka)  
1x Denhardt´s Lösung  
500 µg/ml Heringsperma-DNA (Roche)  
60 % (v/v) deionisiertes Formamid  
10 nM der HCV-Lösung

Waschpuffer 1

3x SSC + 10 mM EDTA

Waschpuffer 2

2x SSC +10 mM EDTA

Waschpuffer 3

1x SSC + 10 mM EDTA

Waschpuffer 4

0,2x SSC + 10 mM EDTA

20x SSC

3 M NaCl (Fluka)  
300 mM Trinatriumcitrat\*2H<sub>2</sub>O

### 3.4 Herstellung der HCV-Targets

Zur Isolation der Virus RNA wurden zwei verschiedene, kommerziell erhältliche Kits verwendet und miteinander verglichen. Mittels einer reversen Transkription wurden cDNA herstellt, die dann vervielfältigt wurden. Es schloss sich eine Markierung, Reinigung und Quantifizierung der HCV-Targets an. Zur weiterführenden Charakterisierung wurde eine Sequenzierung durchgeführt.

#### 3.4.1 Gene-Elute Mammalian total RNA Kit (Sigma)

Es wurde nach Herstellerangaben vorgegangen.

#### 3.4.2 QIAamp viral Mini Kit (QIAGEN)

Es wurde nach Herstellerangaben vorgegangen.

#### 3.4.3 Two-Step RT-PCR

Für die Transkription der RNA wurde zunächst eine RT-PCR durchgeführt im Doppelansatz mit einer Amplifikation mit nested Primern nach Schmidt *et al.* <sup>[130]</sup>. Für die reverse Transkription wurden ein 50 µl Ansatz mit je 30 pmol des RT-antisense und RT-sense Primers, 200 µM dNTP, 10x Puffer und 10 µl isolierte RNA in ein Reaktionsgefäß pipettiert und 5 min bei 95 °C inkubiert. Die Proben wurden anschließend sofort auf Eis gestellt und mit einem Enzymgemisch aus 1,25 U Expand-Revers Transkriptase (Roche) und 1 U Taq DNA-Polymerase versetzt. Die reverse Transkription wurde bei 42 °C für 30 min durchgeführt. Nach einer initiativen Denaturierung bei 95 °C für 60 s, folgten 30 Zyklen mit der Denaturierung bei 95 °C für 60 s, der Anlagerung bei 60 °C für 60 s und einer Polymerisierung bei 72 °C für 120 s. Ein abschließender Polymerisierungsschritt bei 72 °C für 600 s vervollständigte die Amplifikationsprodukte. Eine Amplifikation der HCV cDNA mit nested Primern folgte. Dazu wurden jeweils in einem 50 µl Ansatz 30 pmol n-antisense und n-sense Primer, 200 µM dNTP, 10x Puffer, 1,5 mM MgCl<sub>2</sub>, 1 U Taq-Polymerase und 1 µl cDNA aus der RT-PCR in ein Reaktionsgefäß pipettiert. Folgendes Temperaturprofil wurde durchgeführt: ein einmaliger initiativer Denaturierungsschritt bei 95 °C für 120 s, gefolgt von einer Denaturierung bei 95 °C für 60 s, einer Anlagerungstemperatur bei 60 °C für 60 s, dem Polymerisierungsschritt bei 72 °C für 120 s. Dieser Zyklus wurde 30x wiederholt. Eine Schlusspolymerisierungsphase bei 72 °C für 600 s komplettierte die PCR.

10 µl des PCR-Gemisches wurden auf ein 1,8 %iges Agarosegel geladen und durch Vergleich mit einer 1 kb Leiter charakterisiert.

### **3.4.4 Reinigung**

#### QIAquick PCR Purification Kit (QIAGEN)

Es wurde nach Herstellerangaben vorgegangen. Zusätzlich wurde bei der Reinigung von Cy5-markierten Primern nach dem Anbinden des Amplifikats an die Säule mit 750 µl Guanidiumhydrochlorid zusätzlich gewaschen.

#### FPLC-Aufreinigung

Die PCR-Reaktion-Gemisch wurde auf der FPLC mittel Ionenaustausch-chromatographie (Säule: MonoQ® HR 5/5) über einen Salzgradienten aufgereinigt. Nach der FPLC-Aufreinigung wurden die Produktfraktionen gesammelt und in einen Centricon-30-Zentrifugenkonzentrator (Millipore) überführt und durch Zentrifugation bei 4000 rpm eingengt. Die Fraktion mit der DNA wurde bei 4000 rpm 10 min zentrifugiert, 3x mit 500 µl 1x TE gewaschen, dazwischen jeweils 3 min bei 4000 rpm zentrifugiert und DNA in ein neues Reaktionsgefäß überführt.

### **3.4.5 Targetherstellung mittels PCR**

Die Amplifikate aus Abschnitt 3.4.3 wurden in einer nested PCR für die Hybridisierung mit einem Cy5 am 5' Ende über den entsprechenden Primer (Primer 90) markiert. Für einen Ansatz wurden jeweils 400 nM HCV 90, 400 nM HCV 88, 1,5 mM MgCl<sub>2</sub>, 1 x PCR-Puffer, 200 µM dNTP, 2 U Taq-Polymerase und als Template eine 1:100 Verdünnung aus der two-step Reversen Transkriptase-PCR verwendet. Der Reaktionsansatz wurde mit Mineralöl als Verdunstungsschutz überschichtet. Das Temperaturprofil wurde nach Boldt <sup>[131]</sup> in einem Thermocycler (Landgraf) durchgeführt.

Es wurde ein 2 %iges Agarosegel (Seakam) gegossen. Nach dem Aushärten wurden 10 µl der vorbereiteten Proben aus der PCR auf das Gel aufgetragen. Die Elektrophorese wurde bei 120 V und 45 min durchgeführt.

### **3.4.6 Quantifizierung der Proben**

Die Proben wurden durch photometrische Doppelbestimmung mit je einer 1:50 bzw. 1:100 Verdünnung der Probe im Spectrophotometer (Hitachi U-3000) quantifiziert. Durch die Optische Dichte (OD) kann die Konzentration der amplifizierten DNA ermittelt werden. Dabei wurde folgende Umrechnung verwendet: 1 OD entspricht 33 µg einzelsträngiger



DNA. Anhand dessen wurde die Konzentration über die molare Masse (hier 330 g/mol) ermittelt.

### **3.4.7 Einzelstrang-Isolation**

Die Einzelstrangisolierung doppelsträngiger DNA erfolgte über Streptavidin-Magnetic-Beads (Boehringer Mannheim). Nach Angabe des Herstellers wurde die entsprechende Menge Beads dreimal in  $H_2O_{dd}$  gewaschen, dazwischen 30 s bei 13.200 rpm zentrifugiert. Die Waschlösung wurde abgenommen, DNA-Lösung aufpipettiert und die DNA 30 min unter Schütteln (GFL 3005) bei Raumtemperatur (RT) an die Beads gebunden. Anschließend wurde dreimal kurz mit  $H_2O_{dd}$  gewaschen. Die Denaturierung erfolgte mit 50  $\mu$ l einer 100 mM Natriumhydroxid-Lösung, die auf die Beads gegeben wurde. Es wurde 3 min bei RT geschüttelt. Der Überstand wurde abgenommen und in ein neues Reaktionsgefäß überführt. Nach einer Wiederholung wurden die Überstände mit 50  $\mu$ l 200 mM HCl neutralisiert.

## **3.5 Sequenzierung**

Für die Sequenzierung wurde das Cycle-PCR-Reaktionsgemisch folgendermaßen angesetzt: 1  $\mu$ l  $H_2O_{dd}$ , 2  $\mu$ l Big Dye Ready Mix (Applied Biosystems), 1  $\mu$ l HCV 88 und HCV 90 und 6  $\mu$ l PCR Produkt aus Absatz 3.4.3. Der Ansatz wurde mit Mineralöl überschichtet und bei 95 °C für 15 s denaturiert. Die Anlagerungsphase bei 60 °C für 15 s und die Polymerisierungsphase bei 60 °C für 240 s im Gene Amp PCR System 9600 (Perkin Elmer) durchgeführt. Dieser Zyklus wurde 26x wiederholt.

### **3.5.1 Fällung der DNA für die Sequenzierung**

Das Produkt aus der Cycle-PCR wurde ölfrei in neue Reaktionsgefäße überführt, die DNA mit 34  $\mu$ l 3 M Natriumacetat und 100  $\mu$ l Ethanol (absolut) (Riedel-de-Häen) gefällt. Bei 13200 rpm wurde für 20 min zentrifugiert. Das Pellet wurde mit 70 %igem Ethanol gewaschen und der Zentrifugationsschritt wiederholt. Der Überstand wurde abgenommen und das Pellet vollständig getrocknet. Anschließend wurde in 30  $\mu$ l  $H_2O_{dd}$  resuspendiert und die Lösung in farblose 0,5 ml Reaktionsgefäße überführt. Der Deckel der Reaktionsgefäße wurde abgeschnitten und durch Gummisepten ersetzt. Die Sequenzierung erfolgte im Kapillarsequenzierer ABI Prims 310 (Applied Biosystems) durch die Sanger-Coulson-Methode. Die Auswertung der Spektren erfolgte durch die Software-Tools Sequence Analyse 3.4.1 und Chromas 1.45.

## 3.6 Herstellung der Mikroarrays

### 3.6.1 Slides

Die Oberflächenmodifizierung wurden von Frau B. Meyer-Schlosser durchgeführt. Es handelt sich um mit 1,4-Phenylendiisothiocyanat aktivierte Objektträger, die wie in [20] beschrieben, hergestellt wurden.

### 3.6.2 Roboter

#### 3.6.2.1 Ultramikro-Pipettierroboter, CAG, Bremen

Das Besspotten der Slides wurde im nonkontakt-Verfahren durchgeführt. Der Ultramikro-Pipettierroboter mit Piezo-Pipette wurde im Centrum für Angewandte Sensorik entwickelt.

Die Oligonukleotide wurden im Spottingpuffer gelöst und 70 µl einer 10 µM Lösung staubfrei in die Kavitäten einer 96-well Mikrotiterplatte (Nunc) überführt. Das Aufbringen der Oligonukleotide auf die Slides erfolgte stets nach dem gleichen Muster. Die Oligonukleotide wurden so auf dem Slide abgelegt, dass ein Mikroarray aus drei identischen Arrays entstand (vgl. Abbildung 3-1). Als Positivkontrolle diente das Oligonukleotid S\_U23G.

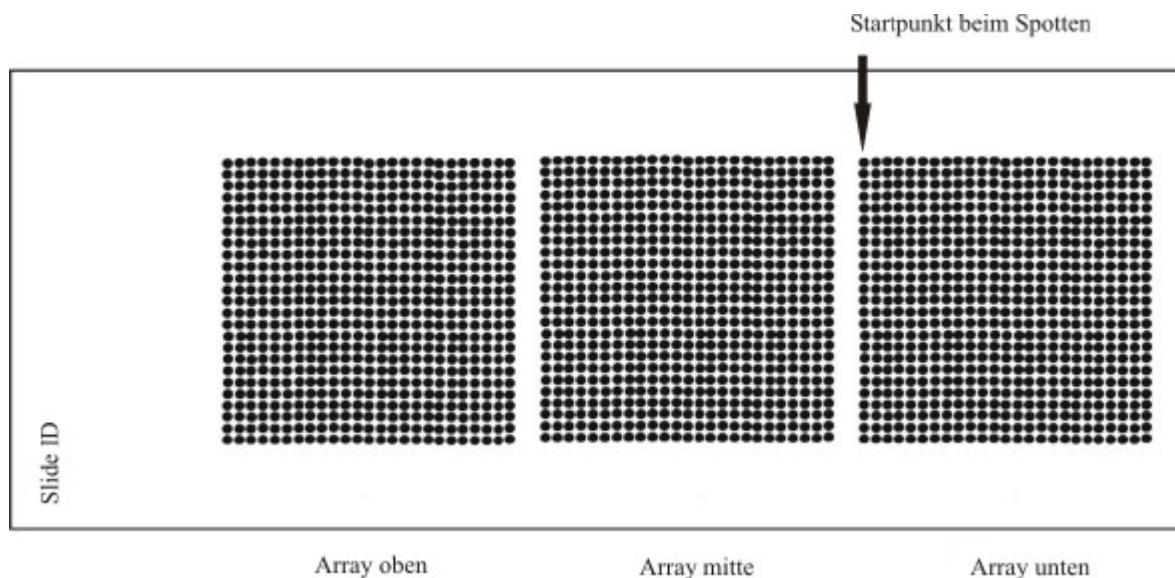


Abbildung 3-1: Schematische Darstellung des Spottingmusters auf einem Slide.

#### 3.6.2.2 SpotArray 24 (Packard BioScience)

Die Oligonukleotide wurden in der Spottinglösung resuspendiert und 20 µl einer 10 µM Lösung wurden in je eine Kavität der Mikrotiterplatten (Nunc) überführt. Die

Luftfeuchtigkeit während des Spottingvorgangs wurde auf 65 % eingestellt, die Anzahl der Testspots auf 15 festgelegt und die Verweildauer der Pins auf dem Mikroarray auf 700 ms eingestellt. Die Oligonukleotide, ebenso wie die Kontrollen, wurden 10-fach in aufeinanderfolgenden Spots abgelegt. Die 10er Blöcke der Kontrollen wurden redundant aufgebracht. Als Positivkontrolle diente das Oligonukleotid S\_U23G.

## **3.7 Hybridisierungen**

### **3.7.1 Anbinden**

Die Oligonukleotide wurden nach dem Aufbringen in einer feuchten Kammer über Nacht bei Raumtemperatur inkubiert, um die Anbindungsreaktion zwischen der Aminomodifizierung der Oligonukleotide und der Phenylendiisothiocyanatgruppe der aktivierten Slides vollständig ablaufen zu lassen. Anschließend wurden mit Waschlösung 1 die nicht angebondenen Oligonukleotide von der Oberfläche abgespült und die Slides anschließend im Druckluftstrom getrocknet.

### **3.7.2 Abblocken der freien reaktiven Gruppen**

Die freien reaktiven Phenylendiisothiocyanatgruppen wurden für 15 min auf einem Orbitalschüttler bei 100 rpm (GFL 3005) in der Blockinglösung abblockiert. Nach 15 min wurden die Slides entnommen, in H<sub>2</sub>O<sub>dd</sub> kurz gespült und mit Druckluft getrocknet. Die Slides können so entweder bei -20 °C in Objektträgerboxen für ca. 4 Wochen eingefroren oder direkt weiter verarbeitet werden.

### **3.7.3 Prähybridisierung und Hybridisierung**

Die Prähybridisierungslösung wurde frisch hergestellt und die ggf. eingefrorenen bespotteten Slides bei Raumtemperatur aufgetaut. In dieser Lösung wurden die Slides für eine Stunde schüttelnd inkubiert, anschließend kurz mit H<sub>2</sub>O<sub>dd</sub> gespült und mit Druckluft getrocknet. Während der Inkubationszeit wurde die Hybridisierungslösung 10 min bei 95 °C denaturiert und sofort in Eiswasser 10 min abgekühlt. 200 µl der Hybridisierungslösung wurden auf den Slide pipettiert und mit einem Deckglas (24 x 60 cm) luftblasenfrei abgedeckt. In einer feuchten Kammer mit der Prähybridisierungslösung wurde die Lösung für eine Stunde bei Raumtemperatur inkubiert. Die Slides wurden anschließend in Waschpuffer 2 kurz abgespült, dann 10 min in Waschpuffer 2 schüttelnd (100 rpm) inkubiert. In Waschpuffer 2 wurde nochmals für

10 min unter Schütteln inkubiert. Nach dem Trocknen unter Druckluft wurden die Slides im Scanner Axon 4000B (Axon) bei einer Wellenlänge von 635 nm und einer PMT von 600 ausgelesen. Die Bildanalyse wurde mit dem Programm Genepix (Axon) durchgeführt.

### **3.7.4 Auswertung**

Wenn nicht anders erwähnt, wurde folgende Auswertung gewählt: Jeder der drei Arrays auf dem Mikroarray wurde einzeln ausgewertet. Die Median-Werte für die einzelnen Spots des jeweiligen Oligonukleotids wurden gemittelt. Der Wert der Positivkontrolle wurde als 100 % gesetzt und die anderen Mittelwerte der Signalintensität auf die Positivkontrolle bezogen. Darüber hinaus wurde eine Intra-Array-Normierung eingeführt, vgl. Kapitel 4.1.4.1, oder die Software *iQomega* verwendet, siehe Kapitel 4.4.2.2.

## **3.8 Optimierung der Hybridisierungsbedingungen**

### **3.8.1 Slidehomogenität**

Zur Untersuchung der Spothomogenität wurden 10  $\mu$ M der Oligonukleotide Amino-M13A und Amino-cM13A-Cy5 mit dem Ultramikro-Pipettierroboter auf die aktivierten Glasslides gespottet. Die Hybridisierung wurde mit einer Konzentration von 5 nM M13A-Cy5 in 30 ml Prähybridisierungslösung bei Raumtemperatur für 60 min schüttelnd durchgeführt. Dreimal wurde der Mikroarray gescannt: nach der Inkubation der Oligonukleotide, nach dem Abspülen mit Waschlösung 1 und nach der Hybridisierung.

### **3.8.2 Variation der Hybridisierungstemperatur**

Die Hybridisierungstemperatur wurde zu Optimierungszwecken variiert, vgl. Kapitel 3.7.3, jedoch bei Hybridisierungstemperaturen von 22 °C, 30 °C und 35 °C.

### **3.8.3 Variation der Formamidkonzentration**

Die Formamidkonzentration im Hybridisierungspuffer wurde zu Optimierungszwecken variiert, vgl. Kapitel 3.7.3, jedoch mit 30 %, 50 % oder 60 % Formamid.

### **3.8.4 Variation der Salzkonzentration in dem Waschlösung**

Die Salzkonzentration im Waschlösung wurde zur Optimierungszwecken variiert, vgl. Kapitel 3.7.3, jedoch mit verschiedenen Verdünnungen der Standard Sodium Citrat - Lösung (SSC), vgl. Waschlösung 1, Waschlösung 2 und Waschlösung 3.

## **4 ERGEBNISSE**

Das Ziel der Arbeit ist die beiden mit unterschiedlichen Methoden hergestellten Oligonukleotid-Bibliotheken zur Erkennung von Hepatitis C-Viren auf Mikroarrays zu evaluieren und zu verifizieren. Dazu wurde die Mikroarrayherstellung etabliert und die optimalen Hybridisierungsbedingungen ermittelt. Zudem wurden zwei Oligonukleotid-Bibliotheken auf verschiedene Weise erstellt. Die 54 Oligonukleotide der manuellen Bibliothek wurden mit Hilfe von Alignments und Programmen zum Primerdesign erstellt. Für die Software-generierte Bibliothek wurde ein Programm entwickelt, mit dem 66 Oligonukleotide für die Identifizierung und Klassifizierung ausgewählt wurden. Anhand von vier HCV-Targets und 46 Oligonukleotide der manuellen Bibliothek wurden die Mikroarrays etabliert und eine Auswertung der Hybridisierungsdaten durch Intra-Array-Normierung eingeführt. Die Versuche wurden auf die Oligonukleotiden der Software-generierten Bibliothek ausgeweitet und zur Auswertung der Hybridisierungssignale auch die Analyse-Software IQMega eingesetzt. Für die Evaluierung der Daten wurden weitere HCV-Targetmoleküle isoliert, sequenziert und anschließend auf entsprechenden Mikroarrays hybridisiert.

### **4.1 Hybridisierung**

Im Folgenden wird auf unterschiedliche Parameter eingegangen, die bei der Hybridisierung bzw. der Auswertung der Hybridisierungsdaten eine große Rolle spielen. In dieser Arbeit wurden Vergleiche zwischen verschiedenen Mikroarrays durchgeführt. Diese sind mit den gleichen Oligonukleotiden bespottet, jedoch mit verschiedenen Targetmolekülen hybridisiert worden. Aus diesem Grund ist die Vergleichbarkeit sowohl zwischen den einzelnen Arrays auf einem Slide als auch zwischen den Slides sehr wichtig. Entscheidend dabei ist die Homogenität der Slideoberfläche. Anschließend wurden verschiedene Parameter, die Einflüsse auf die Stringenz und Kinetik der Hybridisierung ausüben, optimiert.

#### **4.1.1 Optimierung**

Hybridisierungen sind dynamische Prozesse, die durch verschiedene Parameter beeinflusst werden (vgl. Abbildung 4-1). Einige Parameter wurden variiert, um die optimalen Hybridisierungsparameter zu ermitteln.

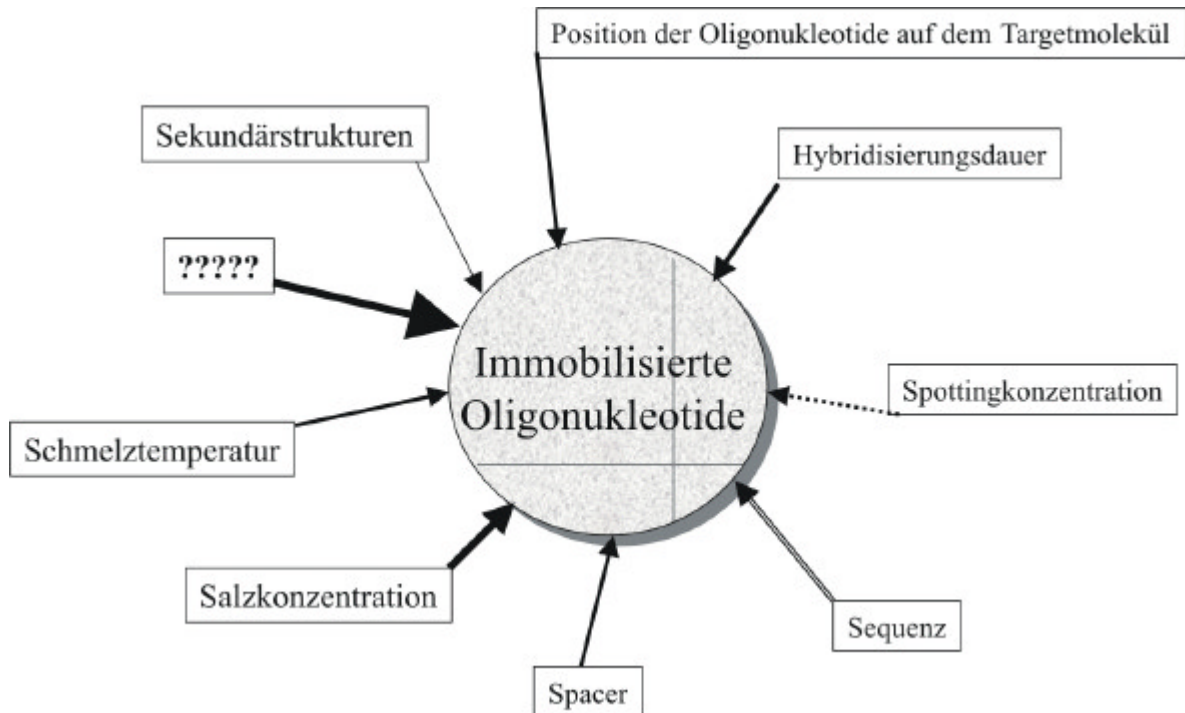


Abbildung 4-1: Skizze von Parametern, die die Hybridisierung auf einem DNA-Mikroarray beeinflussen können. Die unterschiedliche Form der Pfeile symbolisiert die verschiedenen Einflussgrößen der einzelnen Parameter auf die Hybridisierung.

#### 4.1.1.1 Schmelztemperatur

Die Hybridisierungstemperatur wird durch die Schmelztemperatur des DNA-Doppelstranges bestimmt. Dabei gilt die Faustregel: Schmelztemperatur minus 5 °C ist in der Regel die Hybridisierungstemperatur <sup>[132]</sup>. Die Schmelztemperatur für die Oligonukleotide der manuellen sowie der Software-generierten Bibliothek liegt bei 60 °C ± 5 °C.

Für die Optimierung der Hybridisierungstemperatur wurden verschiedene Oligonukleotide aus der manuellen Bibliothek herangezogen und bei 22 °C, 30 °C bzw. 35 °C hybridisiert. In Abbildung 4-2 ist ein Vergleich der relativen Signalintensitäten nach Hybridisierung bei den genannten Temperaturen dargestellt. Die einzelnen Signalintensitäten wurden auf die Positivkontrolle bezogen, um die Signalintensitäten miteinander vergleichen zu können. Deutlich zu erkennen sind hier höhere Signalintensitäten, vor allem bei den universellen Oligonukleotiden bei einer Hybridisierungstemperatur von 22 °C. Bei den spezifischen Oligonukleotiden ist dieser Unterschied nicht vorhanden. Hier erreichen die Oligonukleotide bei einer Hybridisierungstemperatur von 30 °C die höchsten Signale. Bei

diesen Oligonukleotiden liegen die Signalintensitäten näher zusammen, wie aus den Überlappungen der Fehlerbalken z.B. bei M\_3u-5B bei allen drei Hybridisierungstemperaturen hervorgeht. Werden nur die (Genotyp)-spezifischen Oligonukleotide betrachtet, würde die optimale Hybridisierungstemperatur bei 30 °C liegen. Allerdings sind hier die Standardabweichungen deutlich höher als bei den universellen Oligonukleotiden. Aus diesem Grund wurde 22 °C (Raumtemperatur) als optimal gewertet.

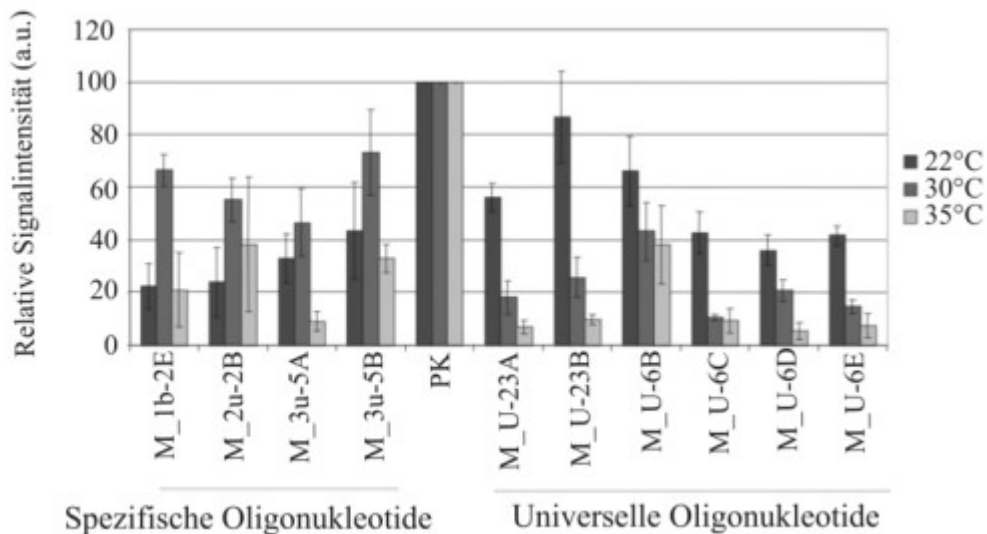


Abbildung 4-2: Diese Darstellung zeigt die relative Signalintensität bei Hybridisierungstemperaturen von 22 °C (RT), 30 °C und 35 °C. Getestet wurden universelle und (Genotyp-) spezifische Oligonukleotide der manuellen Bibliothek. Bei den universellen Oligonukleotiden ist eine deutlich stärkere Signalintensität bei Hybridisierung bei Raumtemperatur zu sehen. Während bei den genotyp-spezifischen Oligonukleotiden höhere Werte bei 30 °C zu erkennen sind, aber auch bei 22 °C sind die Signale deutlich. Die Werte wurden auf die Positivkontrolle (PK) bezogen, die als 100 % gesetzt wurde.

#### 4.1.1.2 Formamid

Formamid wird bei Hybridisierungen dem Hybridisierungspuffer zugesetzt, um Sekundärstrukturen der DNA aufzuschmelzen. Dieses chaotrope Reagenz destabilisiert die Wasserstoffbrückenbindungen zwischen den Strängen. Beim Einsatz von Formamid kann die Hybridisierungstemperatur um 2,4 °C - 2,9 °C pro mol Formamid <sup>[133]</sup> herabgesenkt werden.

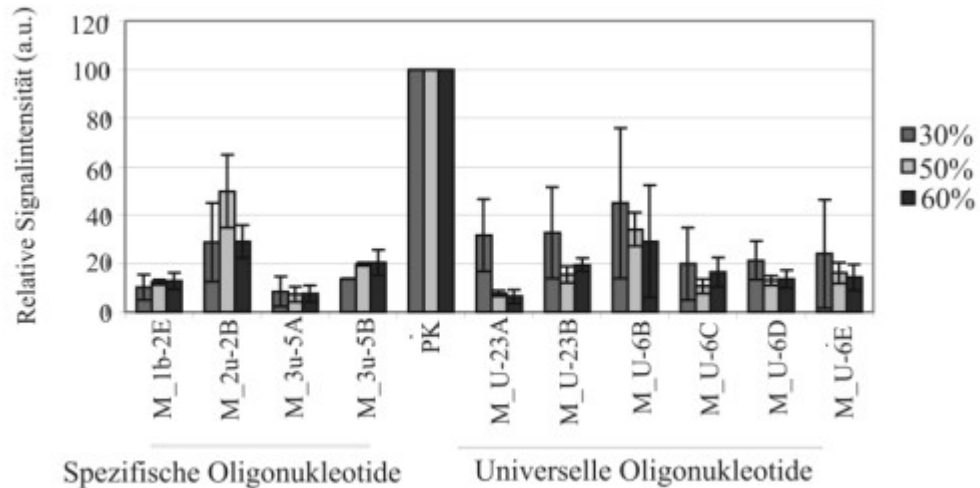


Abbildung 4-3: Darstellung der relativen Signalintensitäten ausgewählter Oligonucleotide der manuellen Bibliothek normiert auf die Positivkontrolle (PK) nach Hybridisierung bei Formamidkonzentrationen von 30 %, 50 % und 60 % im Hybridisierungspuffer. Bei den (Genotyp-) spezifischen Oligonucleotiden ist die Signalintensität bei allen drei Versuchsbedingungen vergleichbar, während bei den universellen Oligonucleotiden deutliche höhere Signale bei 30 % Formamid detektiert werden konnten, jedoch bei einem sehr großen Fehler bis zu 54 % bei M\_U-6B. Bei den Formamidkonzentrationen von 50 % und 60 % sind dagegen vergleichbare Signalintensitäten und deutlich geringere Fehler von 29 % bei M\_2u-2B zu erkennen.

Die Experimente, die den Daten aus Abbildung 4-3 zugrunde liegen, zeigen, dass die gewählten Formamidkonzentration keinen erheblichen Einfluß auf die Hybridisierung haben. Zwar sind auf den ersten Blick die Signalintensitäten bei den universellen Oligonucleotiden bei einer Formamidkonzentration von 30 % höher als bei 50 % bzw. 60 %, jedoch sind auch die Fehler bis zu 54 % bei M\_U-6B maximal. In den nachfolgenden Experimenten wurde eine optimale Formamidkonzentration von 50 % aufgrund der geringeren Fehlerbalken gewählt.

#### 4.1.1.3 Waschbedingungen

Die Waschbedingungen, vor allem die NaCl-Konzentration in der Waschlösung, ist wichtig für die Stringenz und für die Minimierung des Hintergrundes. Die Stringenz determiniert die Stabilität einer Bindung, d.h. Hybride ohne Basenfehlpaarungen sind stabiler als mit Basenfehlpaarungen. Die Position der Basenfehlpaarung spielt dabei eine weitere Rolle, wie in Kapitel 4.1.1.7 näher betrachtet wird. Die NaCl-Konzentration wurde zwischen 200 mM, 100 mM und 20 mM Endkonzentration in der Waschlösung variiert. In Abbildung 4-4 sind die relativen Signalintensitäten für verschiedene Oligonucleotide der manuellen Bibliothek gezeigt. Deutlich zu erkennen sind die signifikant höheren



Signalintensitäten bei 100 mM NaCl in der Waschlösung. Fast alle Oligonukleotide haben hier die höchsten Werte, außer M\_U-23A und M\_U23-B.

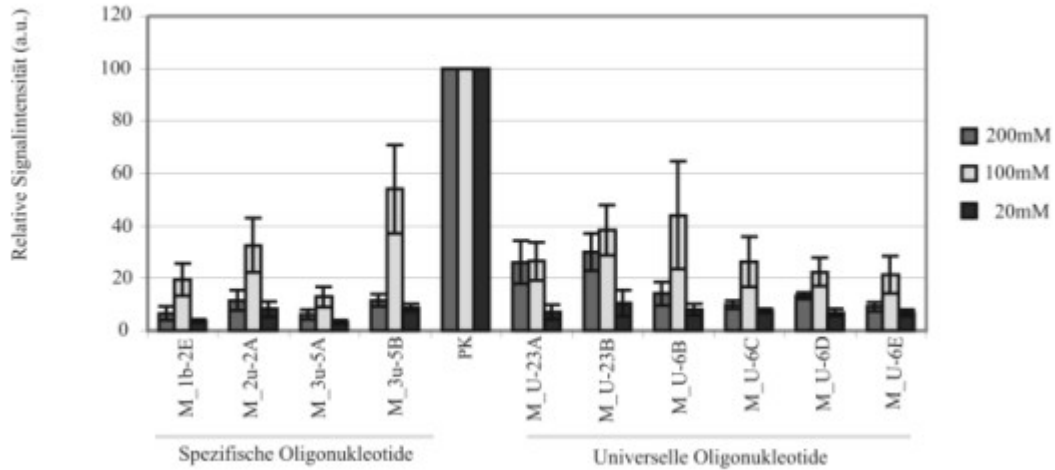


Abbildung 4-4: Das Diagramm zeigt relative Signalintensitäten verschiedener Oligonukleotide der manuellen Bibliothek bei der Variation der NaCl-Konzentrationen in der Waschlösung. Die Signalintensitäten der einzelnen Oligonukleotide nach der Hybridisierung wurden auf die Positivkontrolle (PK) normiert. Starke Signalintensitäten sind bei einer NaCl-Konzentration von 100 mM zu erkennen, während bei 200 mM und 20 mM deutlich schwächere Signale zu erkennen sind.

#### 4.1.1.4 Spottingkonzentration

Die Spottingkonzentration ist abhängig von der Bindungskapazität der aktivierten Slideoberfläche, aber auch von der Art und Struktur der Oligonukleotide. Die hier verwendeten PDITC-Slides (Abbildung 1-2) haben eine Bindungskapazität mit  $45 \text{ fmol/mm}^2$  [134]. Die aufgetragenen Tropfen hatten einen durchschnittlichen Durchmesser von  $250 \mu\text{m}$ . D.h. diese  $250 \mu\text{m}$  könnten  $2 \text{ fmol}$  Oligonukleotide aufnehmen. Bei dem Ultramikropipettierroboter werden pro Tropfen ca.  $250 \text{ pl}$  Flüssigkeit abgegeben, bei 5 Tropfen dementsprechend  $1,25 \text{ nl}$ . Bei einer Spottingkonzentration von  $10 \mu\text{M}$  werden  $12,5 \text{ fmol}$  aufgetropft, d.h. um sechsfach mehr als theoretisch abgebunden werden könnten. In Abbildung 4-5 sind die absoluten Signalintensitäten der gespotteten Oligonukleotidkonzentrationen von  $10 \mu\text{M}$ ,  $1 \mu\text{M}$ ,  $0,1 \mu\text{M}$ ,  $0,01 \mu\text{M}$  und  $0,001 \mu\text{M}$  gezeigt.

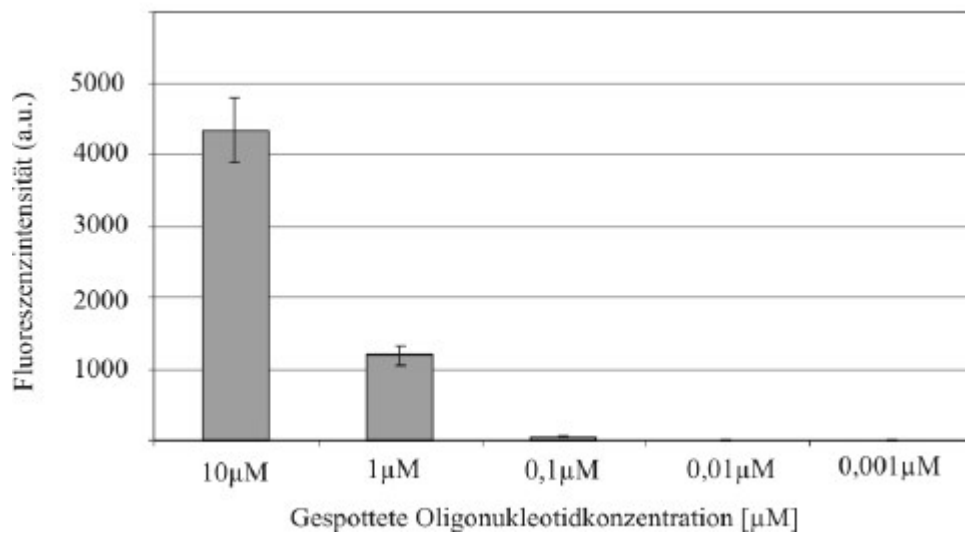


Abbildung 4-5: Variation der Spottingkonzentration. Aufgetragen sind verschiedene Spottingkonzentrationen von 0,001 µM bis 10 µM und die absolute Signalintensität. Diese Werte wurden nach Hybridisierung mit 50 % Formamid bei Raumtemperatur erzielt. Deutlich zu erkennen ist die höchste Signalintensität bei 10 µM.

#### 4.1.1.5 Zeit

Die Hybridisierung ist ein dynamischer Prozess, in dem sich ein Gleichgewicht zwischen Oligonukleotiden und Target einstellt. Die Länge des Target spielt in diesem Gleichgewicht eine große Rolle, da dieses frei beweglich in der Lösung und sein Zielmolekül auf der Oberfläche immobilisiert ist. Bei längeren Targetmolekülen von über 1000 bp, z.B. cDNAs, wird die Hybridisierung oft über Nacht durchgeführt, um den Targets ausreichend Zeit für eine spezifische Erkennung der entsprechenden Oligonukleotide zu geben.

Die hier verwendeten Targets sind nur 209 bp lang. Oligonukleotids aus verschiedenen Bereichen der Targets wurden ausgesucht und bei 30 min, 60 min, 120 min bzw. 180 min hybridisiert. In Abbildung 4-6 sind die absoluten Signalintensitäten nach der Hybridisierung von drei Oligonukleotiden aus der manuellen Bibliothek nach verschiedene Hybridierungszeiten aufgetragen. Zu erkennen sind zunächst unterschiedliche Signalintensitäten, abhängig von der Identität des Oligonukleotides. Für M\_1b-2E ist ein deutliches Optimum der Hybridisierungsdauer bei 60 min zu sehen.

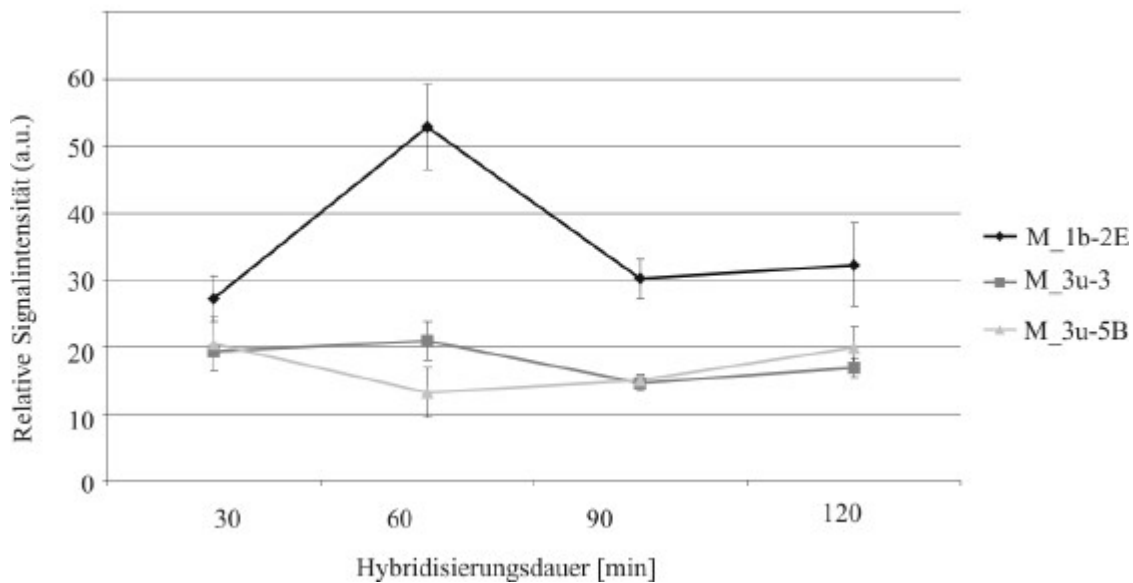


Abbildung 4-6: Das Diagramm zeigt eine Auftragung der relativen Signalintensität gegen die verschiedenen Hybridisierungszeiten für drei verschiedene Oligonukleotide. Für M\_1b-2E ist ein deutliches Optimum der Hybridisierungsdauer bei 60 min zu sehen.

#### 4.1.1.6 Untersuchungen von Einzelstrang- gegen Doppelstrang Hybridisierungen

Aufgrund der hier angewendeten Reversen Transkription der Virus RNA in cDNA wird doppelsträngiges Hybridisierungsmaterial hergestellt. Bei dem Design der Oligonukleotide wurde jedoch nur der sense-Strang des Virus, der in den Datenbanken hinterlegt ist, verwendet. Untersuchungen sollten zeigen, ob die Oligonukleotide spezifisch sind und keine falsch positiven Signale aufgrund von Kreuzreaktivität auftreten. In Abbildung 4-7 sind die verschiedenen Signalintensitäten der Hybridisierungsereignisse für einige genotyp-spezifische Oligonukleotide der Software-generierten Bibliothek dargestellt. Deutlich zu erkennen ist die höhere Signalintensität bei fünf Oligonukleotiden bei Verwendung von doppelsträngigem Target. Außerdem sind die unterschiedliche Signalintensitäten der verschiedenen Oligonukleotide auffällig, obwohl die Schmelztemperatur angepasst, die Sekundärstruktur minimiert und die Oligonukleotide für die Erkennung eines Genotypes entwickelt wurden.

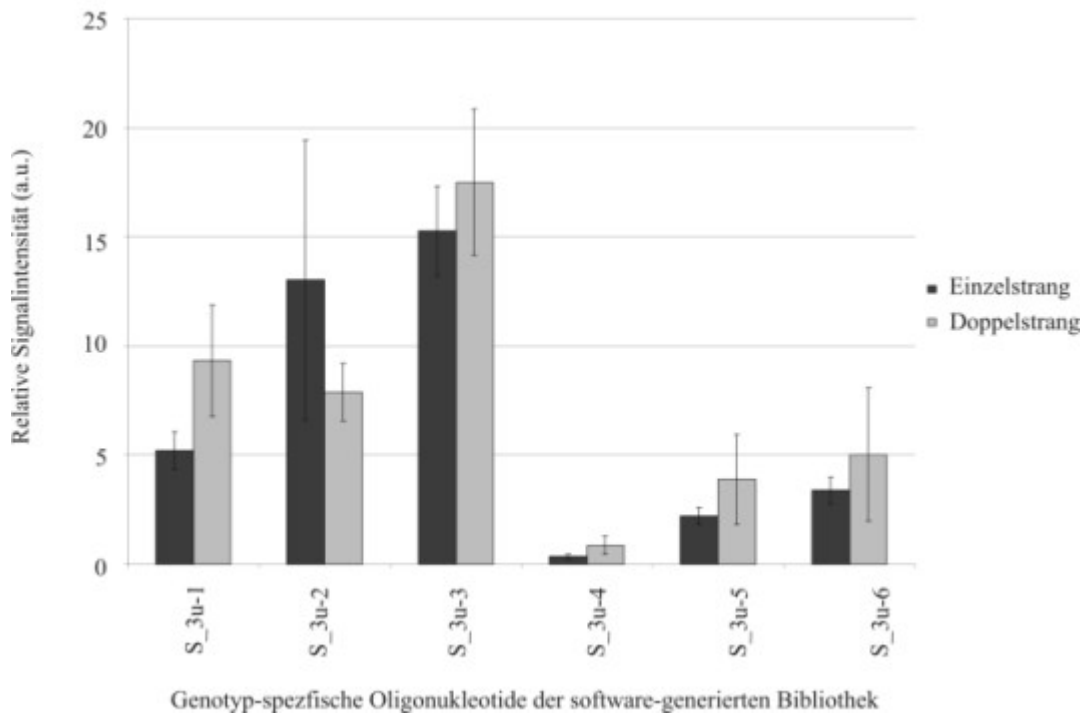


Abbildung 4-7: Vergleich der Signalintensitäten normiert auf die Positivkontrolle (nicht dargestellt) für genotyp-spezifische Oligonukleotide nach Hybridisierung mit einzelsträngigem und doppelsträngigem HCV-Target am Beispiel des Genotypes 3a.

#### 4.1.1.7 Mismatch

Bei der Analyse von Einzelnukleotid Polymorphismen (SNPs) ist es notwendig *in-vitro* zwischen einem sogenannten Perfect Match (PM), die entspricht 100 % Homologie von Fänger und Target - und einem Fänger-Target-Paar mit einem oder mehreren Basenfehlpaarungen unterscheiden zu können. In Abbildung 4-8 sind exemplarisch Hybridisierungsereignisse zwischen dem HCV Genotyp 1a und sechs verschiedenen Fängern dargestellt. Die verwendeten Oligonukleotide haben zu dem Target 100 % Homologie (U23G) bis hin zu drei Basenfehlpaarungen (mismatches, MM). Deutlich zu sehen ist die Abnahme der Signalintensität von der Perfect Match (U23G) zu der Mismatch Probe U23D mit einem Mismatch von 61 %. Hat die Targetsequenz drei Mismatches zu der Oligonukleotidsequenz, wie bei U23J, so beträgt die Abnahme der Signalintensität 93 %. Es ist also ein deutlicher Abfall der Signalintensität zu erkennen. Dabei verhalten sich die Oligos mit Basenfehlpaarungen unterschiedlich, abhängig von der Position des Mismatches. Der Mismatch bei U23C ist an sechster Position vom 5'Ende, U23D an der sechster Position vom 3' Ende. Bei U23C handelt es sich um einen Austausch von T→C, bei U23D von T→G. Energetisch gesehen ist die AT-Bindung die stabilste Bindung. Eine AG-Bindung ist schwächer und eine AC-Bindung ist noch schwächer.

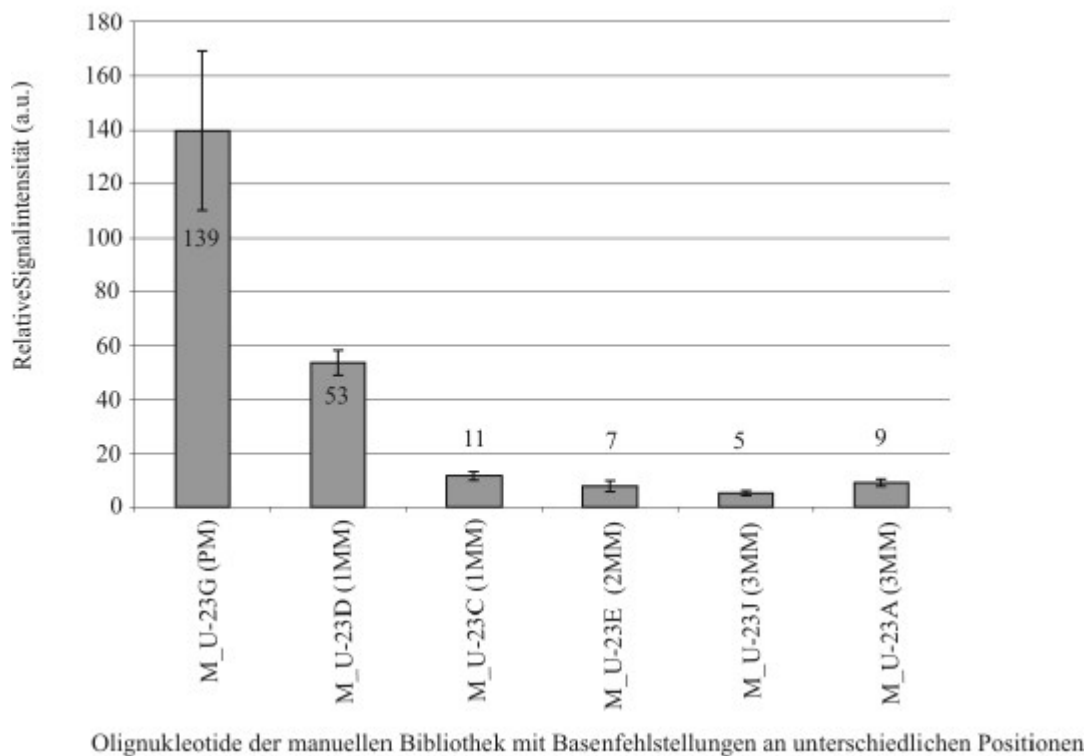


Abbildung 4-8: Graphische Darstellung der Signalintensitäten verschiedener Oligonucleotide mit einem Perfect Match und ein oder mehreren Basenfehlpaarungen an unterschiedlichen Positionen.

## 4.1.2 Weitere Parameter bei der Analyse der Hybridisierungsdaten

### 4.1.2.1 *Bleaching*

Cy-Farbstoffe sind lichtempfindlich. Zu häufiges Scannen kann zu Bleaching-Effekten führen. Dargestellt ist in der Abbildung 4-9 das Verhalten des Farbstoffs Cy5 nach 15 Scanvorgängen, die hintereinander durchgeführt wurden (dunkle Balken). Dazu wurden doppeltmarkierte Oligonucleotide gespottet, direkt gescannt und anschließend über Nacht eingefroren. Nach Lagerung über Nacht bei  $-20\text{ }^{\circ}\text{C}$  und Auftauen auf Raumtemperatur wurde erneut 10 mal hintereinander gescannt (dargestellt in hellen Balken). Deutlich zu sehen ist eine Abnahme der Signalintensität um durchschnittlich 38 %. Darüber hinaus ist ein starker Verlust der Signalintensität nach der Lagerung über Nacht bei  $-20\text{ }^{\circ}\text{C}$  von 74 % zu erkennen. Der Abfall über die folgenden zehn Messungen liegt bei 44 %. Insgesamt beläuft sich die Abnahme der Signalintensität um  $\geq 90\text{ }%$ .

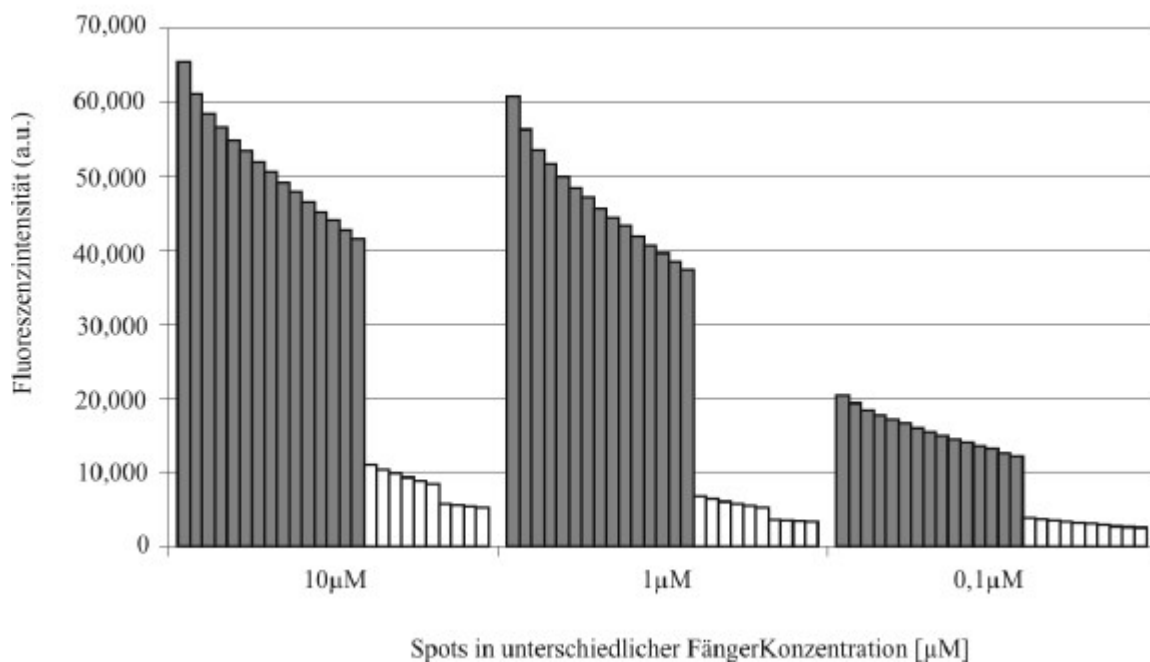


Abbildung 4-9: Bleaching Effekt durch die Laserabtastung im Laser-Scanner. Die dunklen Balken sind aufeinanderfolgende Messungen direkt nach der Hybridisierung, während die hellen Balken die selben Slides nach einem Einfrier-Auftau-Zyklus markieren.

Nach diesen Untersuchungen lag die optimale Oligonukleotidkonzentration bei 10  $\mu\text{M}$ . Die optimale Konzentration der zu hybridisierenden Targetlösung betrug 10 nM. Die grundlegende Optimierung der Hybridisierungsbedingungen ergab eine Formamidkonzentration von 50 % im Hybridisierungspuffer (Hybridisierungslösung 1) und eine Hybridisierungstemperatur von 22 °C bei doppelsträngigen Targetmolekülen. Eine Waschlösung mit 100 mM NaCl (Waschlösung 3) zeigt sich für die Abtrennung von unspezifischen Hybridisierungsereignissen optimal. Bei den untersuchten Oligonukleotiden konnte unter diesen Bedingungen ein Oligonukleotid mit mindestens einer Basenfehlpaarung zum Target sicher von dem Oligonukleotid ohne Basenfehlpaarung unterschieden werden. Sie unterscheiden sich durch eine 60 %ige Abnahme der Signalintensität bei dem Oligonukleotid mit Basenfehlpaarung.

Mit diesen Parametern konnte gezeigt werden, dass sich mittelständige Basenfehlpaarungen von Perfect Match Oligonukleotiden abtrennen lassen. Somit sind die Bedingungen stringent, um mit den ausgewählten Oligonukleotiden SNPs, die für die Determinierung der Genotypen von entscheidender Bedeutung sind, sicher unterscheiden zu können.

### 4.1.3 Untersuchung der Slidehomogenität

Eine homogene Signalintensität auf der Oberfläche ist zu gewährleisten, um Signalintensitäten auf dem Slide verteilt sind sowie die Slides verschiedener Experimente miteinander vergleichen zu können. Befinden sich diese Spots an unterschiedlichen Positionen auf dem Array, muss eine Homogenität der Signalintensitäten über den ganzen Array gegeben sein. Zur experimentellen Durchführung wurde ein doppelmarkiertes Oligonukleotid, mit einer Aminomodifizierung am 3'-Ende und einer Cy5-Markierung am 5'-Ende auf dem Slide immobilisiert und zur Überprüfung der Gleichmäßigkeit der Oberfläche verwendet. Die verschiedenen Schritte der Mikroarrayherstellung wurden dabei einzeln untersucht.

#### Spotting

Auf dem Mikroarray wurden die zwei verschiedene Oligonukleotide in drei identische Arrays positioniert. Bei den Oligonukleotiden handelt es sich um A-M13cA-Cy5 und um A-M13A. Die Arrays werden nach ihrer Position auf dem Slide als „Array oben“, „Array mitte“ und „Array unten“ bezeichnet (vgl. Abbildung 3-1).

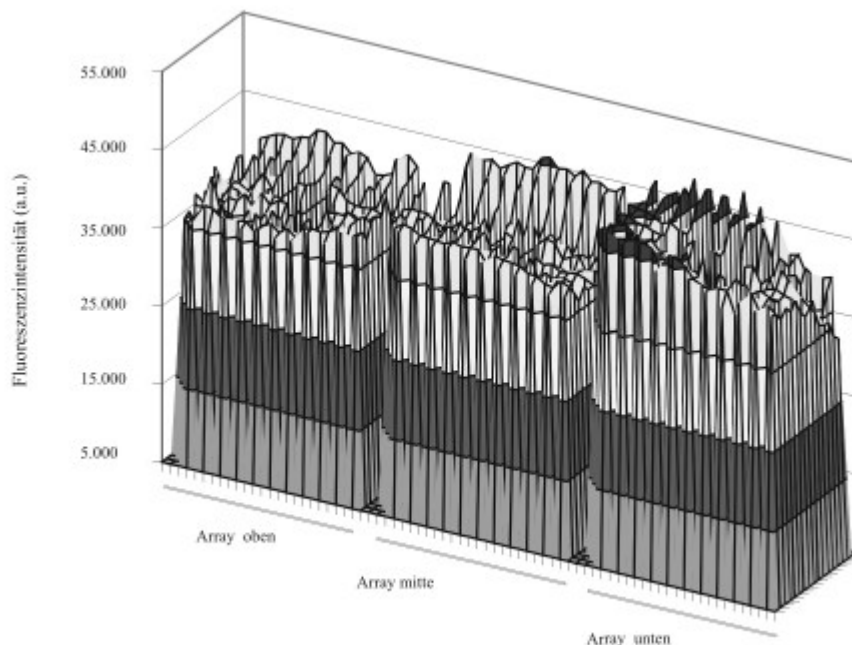


Abbildung 4-10: Dreidimensionale Darstellung der einzelnen Fluoreszenzintensitäten eines Mikroarrays, auf den 10  $\mu$ M Oligonukleotidlösung gespottet wurden. Das Oligonukleotid ist doppelmarkiert. Am 3' Ende trägt es eine Aminomodifikation und am 5' Ende eine Cy5-Markierung. Die höchsten Fluoreszenzintensitäten von 55000 relativen Einheiten erreichen die Spots am linken Rand des unteren Arrays. Der obere und mittlere Array erreichen Signalintensitäten zwischen 35000 und 45000 relativen Einheiten.

In Abbildung 4-10 ist die Verteilung der Signalintensität nach der Anbindung des doppelt markierten Oligonukleotids M13MP18 dargestellt. Zu erkennen ist deutlich eine höchsten Signalintensität mit einem Maximalwert von 48358 a.u. am linken Rand des unteren Arrays. Deutlich zu sehen ist ein Abfall der Werte zum rechten hinteren Rand hin bis auf 33025 a.u.. Die Standardabweichung liegt hier bei 46 %. Die Signalintensitäten des oberen Arrays liegen zwischen 31274 a.u. und 42263 a.u.. Sie variieren um 35 %. Im mittleren Array erreichen die Werte zwischen 34740 a.u. und 45941 a.u.. Die Standardabweichungen liegen hier bei 32 %. Bei dem oberen Array ist kein deutlicher Gradient zu erkennen, während beim mittleren Array zum Rand des Slides die Fluoreszenzwerte am höchsten sind.

### Oligonukleotid- Immobilisierung

Nach der Immobilisierung werden bei der Mikroarray-Herstellung die nicht angebondenen Oligonukleotide abgewaschen und die Anbindungsrate an die aktivierte Oberfläche quantifiziert. Die prozentualen Abweichungen im oberen und mittleren Array liegen unter 10 % . Eine Abnahme der Signalintensität im Vergleich zum ersten Scanvorgang liegt für den oberen und mittleren Array bei unter 1 % und ist daher vernachlässigbar. Beim unteren Array liegt eine Abnahme der Signalintensität um 18 % vor.

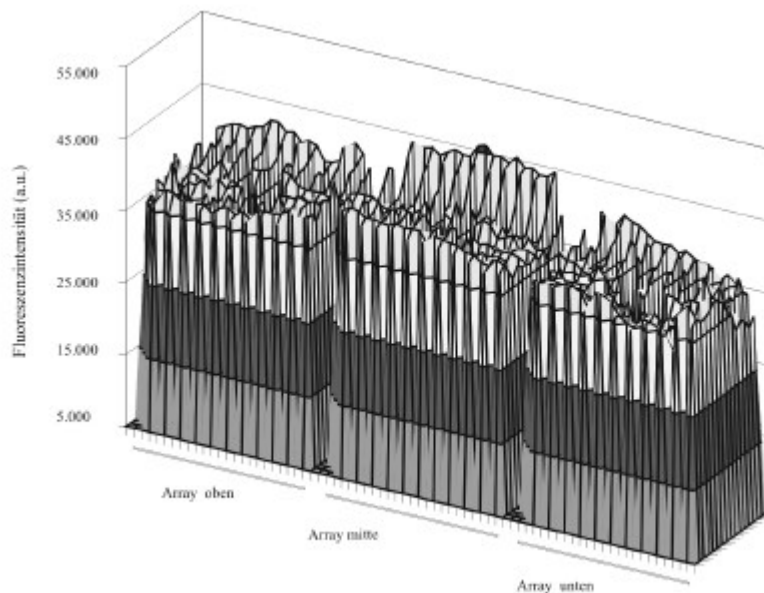


Abbildung 4-11: Dreidimensionale Darstellung der einzelnen Signalintensitäten des immobilisierten, doppeltmodifizierten Oligonukleotides nach dem Abwaschen der nicht angebondenen Oligonukleotide. Die höchste Signalintensität von über 45000 relativen Einheiten wird im mittleren Array erreicht. Der untere Array zeigt eine Abnahme der Signalintensität von etwa 18 %, während der obere sowie der untere Array eine Abnahme der Signalintensität um weniger als 1 % zeigen.



### Hybridisierung mit markiertem Oligonukleotid

Nach dem Hybridisieren einer 5  $\mu\text{M}$  Oligonukleotidlösung des komplementären Stranges ist vor allem im oberen und mittleren Array eine deutliche Erhöhung der Fluoreszenz auf über 45000 a.u. zu erkennen (vgl. Abbildung 4-12). In Abbildung 4-11 sind an korrespondierender Stelle die Werte zwischen 35000 a.u. und 43000 a.u.. Daraus kann geschlossen werden, dass nicht alle freien Bindungsplätze der gebundenen Oligonukleotide durch den komplementären Strang genutzt werden.

Deutlich zu sehen ist eine immer stärkere Abnahme der Signalintensitäten der markierten, nicht hybridisierten Oligonukleotide v.a. im unteren Array, was vermutlich auf das Abwaschen der immobilisierten Oligonukleotide während der Hybridisierung zurückzuführen ist. Grund hierfür könnte der noch zu optimierende Herstellungsprozess bei der Funktionalisierung der Oberflächen sein.

Beschränkt sich die Betrachtung auf die hybridisierten Oligonukleotide in dem oberen bzw. mittleren Array sind die gleichen Phänomene wie in den beiden vorangegangenen Abbildungen zu beobachten. Auch hier ist die Schwankung der Signalintensität bei 14 % beim oberen und 16 % mittleren Array, während sie im unteren Array 45 % erreicht.

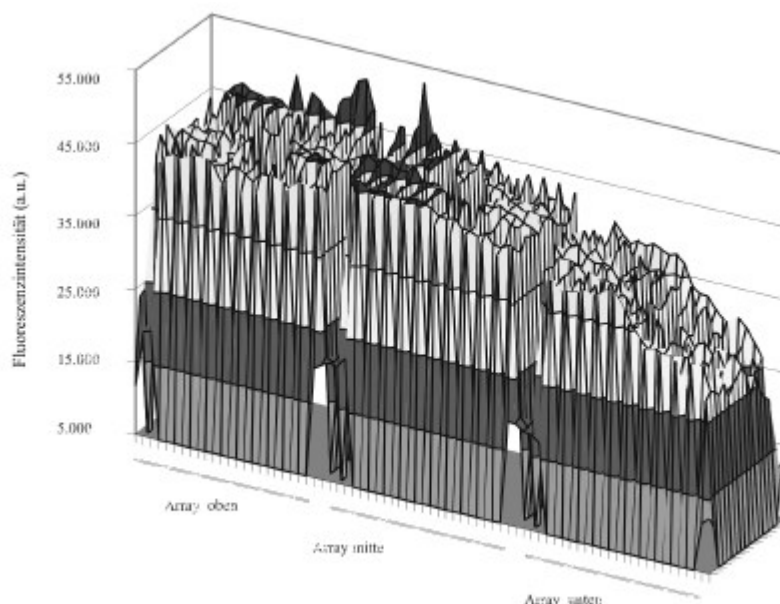


Abbildung 4-12: Verteilung der Signalintensität nach Hybridisierung mit Cy5 markiertem komplementären Oligonukleotid. Unterschiede in der Signalintensität zwischen den einzelnen Arrays auf dem Slide werden immer deutlicher: im unteren Array betragen die Schwankungen bei fast 45 %, während das obere und das mittlere Array Standardabweichungen etwa 15 % aufweisen.

Zusammenfassend kann festgestellt werden, dass die Schwankungsbreite der Signalintensität vom Auftropfen bis zur Hybridisierung für ein bestimmtes Oligonukleotid minimal sind, d.h. die Schwankungsbreite bei der Erstellung der Mikroarrays von Spotter und Bindekapazität ist akzeptabel. Allerdings konnte bei diesen Untersuchungen gezeigt werden, dass die Oberfläche des Slides, vor allem im unteren Array, inhomogen ist.

Dieses Ergebnis führte zu der Entwicklung eines Pipettierschemas und einer Auswertungstrategie, die nun näher betrachtet werden.

#### **4.1.4 Qualität der Hybridisierungen**

Um die verschiedenen Mikroarrays miteinander vergleichen zu können und die versuchsbedingten Abweichungen von den genotypbedingten Abweichungen in der Signalintensität bestimmen zu können, wurden zum einen die Standardabweichung innerhalb eines Arrays bestimmt („Intra-Array“), weiter die Abweichungen innerhalb eines Slides („Intra -Slide“) und die Abweichungen von verschiedenen Mikroarrays bei gleichem Target und gleichen Hybridisierungsbedingungen („Inter-Slide“).

##### *4.1.4.1 Intra-Array*

Bedingt durch die geringe Slidehomologie, wie sie in Kapitel 4.1.3 beschrieben ist, wurde ein Bepottungsmuster entwickelt, um die Inhomogenität der Slideoberfläche auszugleichen. Jedes Oligonukleotid wurde insgesamt sechsmal aufgebracht und zwar in je zwei Blöcken, die vervollständigt durch eine Positivkontrolle (PK) ein Quadrupol bilden (vgl. Abbildung 4-13). Mit Hilfe dieser Normierung war es möglich, innerhalb eines Arrays die Signalintensitäten der Oligonukleotide miteinander zu standardisieren.

In Abbildung 4-14 ist ein Vergleich der Auswertemethode mit und ohne Intra-Array Normierung dargestellt. Hybridisiert wurden Oligonukleotide der Software-generierten Bibliothek gegen HCV-Genotyp 1a. Zu erkennen sind die verschiedenen Signalintensitäten für die verschiedenen Oligonukleotide. Dargestellt sind die Signalintensitäten ohne die Intra-Array Normierung (helle Linie) sowie Intra-Array Normierung (dunkle Linie). Deutlich zu erkennen ist ein Oligonukleotid (S\_U-10), dessen Signalintensitäten extreme Differenzen aufzeigen. Jedoch ist anhand der enormen Fehler bei der Auswertung ohne die Intra-Array Normierung eindeutig, dass der Wert mit Intra-Array Normierung vermutlich der Korrekte ist. Weniger drastisch, aber ebenso deutlich, ist dies auch bei sechs weiteren Oligonukleotiden.

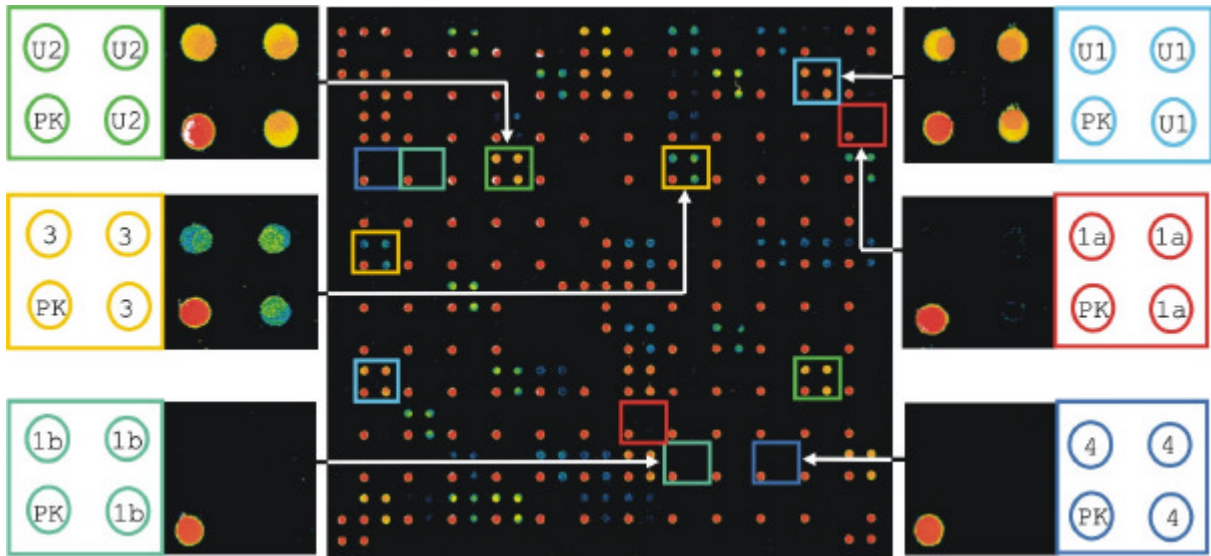


Abbildung 4-13: Hybridisierter Array zur Verdeutlichung des Spottingsmusters. Für jedes dreifach gespottete Oligonukleotid wird eine Positivkontrolle (PK) mitgeführt und jedes Quadrupol ist pro Chip zweimal vorhanden. Beispielpfand vergrößert sind die Quadrupole für die universellen Oligonukleotide U1 und U2 sowie für die genotyp- bzw. subtyp-spezifischen Oligonukleotide für 1a, 1b, 3 und 4 <sup>[135]</sup>.

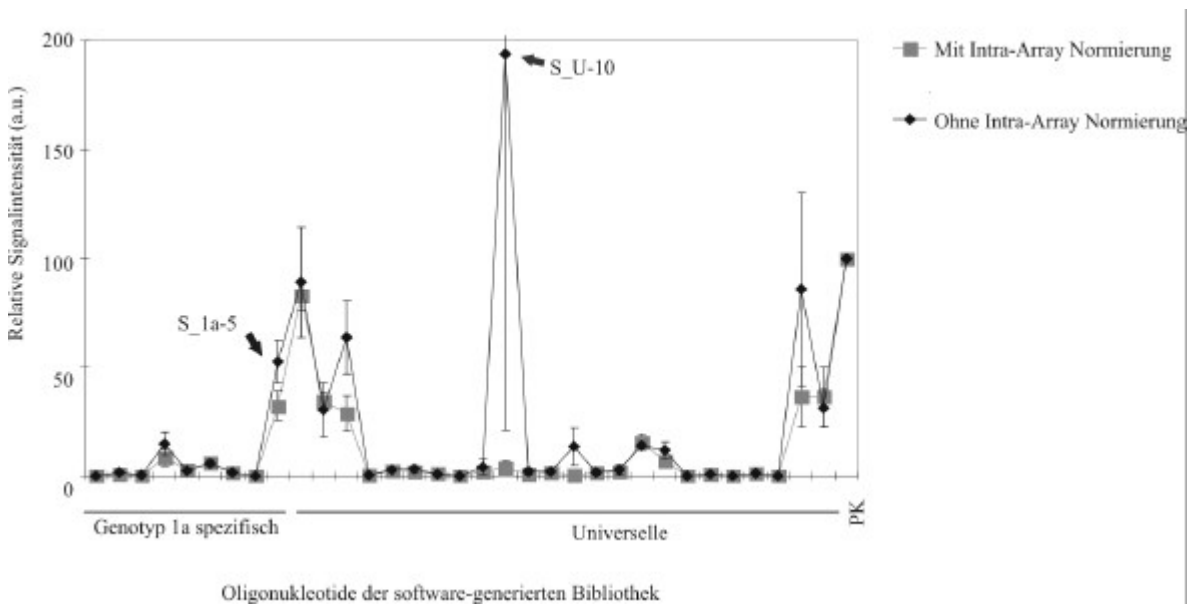


Abbildung 4-14: Darstellung zweier Auswertestrategien mit (◆) und ohne (■) Intra-Array Normierung. Gegeneinander aufgetragen sind relative Signalintensitäten verschiedener Oligonukleotide der Software-generierten Bibliothek, die mit HCV Genotyp 1a hybridisiert wurden. Deutlich ist eine Diskrepanz der resultierenden Signalintensitäten zu erkennen. Der Vorteil der Intra-Array Normierung zeigt sich in den wesentlich kleineren Fehlern deutlich.

Abbildung 4-15 illustriert diese Ergebnisse auf einem entsprechenden Mikroarray. Die vier dargestellten Profile zeigen ausgewählte Reihen bzw. Spalten des Arrays. Anhand derer zeigt sich die Inhomogenität des Slides, die die Ergebnisse aus Abbildung 4-14 widerspiegeln. Die beiden rot umrandeten Quadrupole verdeutlichen die Position von S\_U-10, die grün umrandeten die Lage von S\_1a-5. Beispielhaft ist für S\_U-10 gezeigt, dass die hohe Signalintensität und der große Fehlerbalken auf schwache Signalintensität des Quadrupols unten links in Abbildung 4-15, 1 zurückzuführen ist.

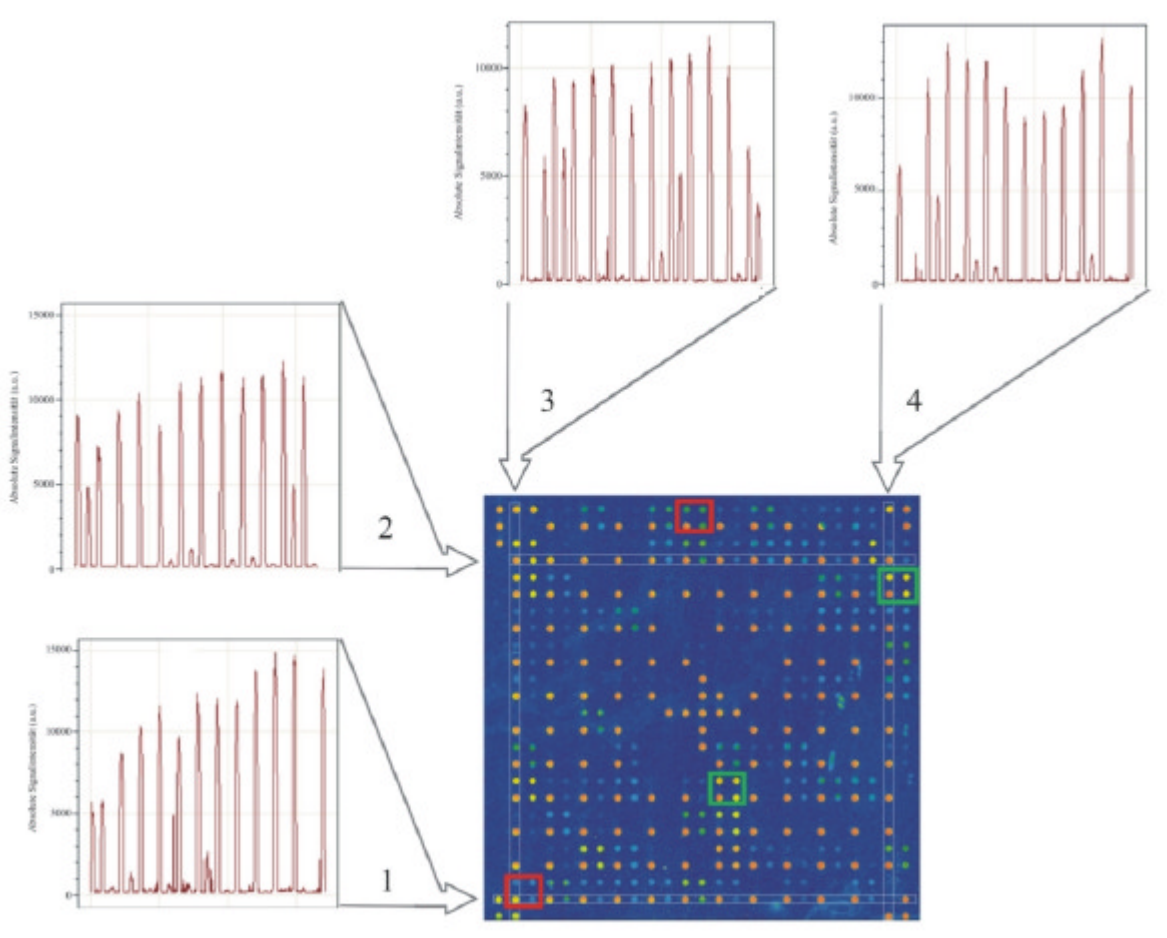


Abbildung 4-15: Darstellung einer Slideoberfläche anhand ausgewählter Intensitätsprofile und der Positivkontrollen. Vor allem im unteren Bereich, gekennzeichnet mit dem Profil 1, ist eine deutliche Signalintensitätszunahme von links nach rechts von ca. 62 % zu erkennen. Auch bei Profil 3 ist eine entsprechende Abnahme der Signalintensität von ca. 70 % zu sehen. Eine Schwankung von ca. 33 % ist bei Profil 2 zu erkennen, was in diesem Beispiel die geringste Abweichung darstellt. In Profil 4 liegt die Schwankung bei ca. 38 %. Die rote Umrandung der Quadrupole zeigt die Position von S\_U-10, die grüne Umrandung weist die Lage von S\_1a-5.

4.1.4.2 *Intra-Slide*

Pro Slide sind drei Arrays aufgebracht, die als Parallelexperimente gezählt werden. Ein Vergleich dieser Arrays wird in den folgenden beiden Abbildungen gezeigt und ermöglicht Aussagen über die Effizienz der Hybridisierung. Mit Hilfe der Intra-Array Normalisierung können die drei Arrays, die auf einem Slide aufgebracht sind, zueinander in Beziehung gebracht werden und man erhält vergleichbare Daten. Für die Erhebung der endgültigen Signalintensitäten wurden zunächst die Spots in jedem Quadrupol auf die dazugehörige Positivkontrolle bezogen, wobei die Signalintensität der Positivkontrolle auf 100 % gesetzt wurde. Dann wurden die beiden Quadrupole miteinander in Beziehung gesetzt, um für einen Array einen Wert pro Oligonukleotid zu erhalten. Wichtig ist hierbei nicht nur die Präsenz eines Signals, sondern auch die Signalintensität. Das wird in Abbildung 4-16 für die genotypspezifischen Oligonukleotide mit geringer Signalintensität gezeigt. Auch gibt es für die universellen Oligonukleotide, deren Signalintensität höher ist als der genotypspezifischen Oligonukleotide. Die Abbildung zeigt die hohe Korrelation innerhalb des Slides nach der Intra-Array Normierung.

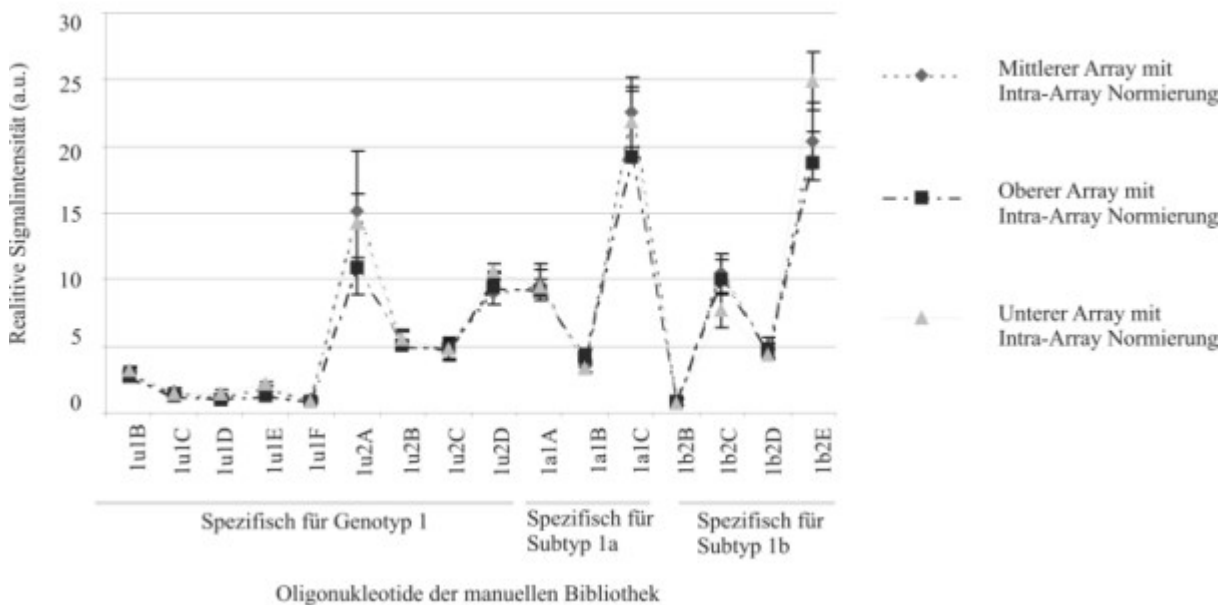


Abbildung 4-16: Darstellung der relativen Signalintensitäten genotypspezifischer Oligonukleotide zur Untersuchung der Slidehomogenität innerhalb eines Arrays nach Hybridisierung der Oligonukleotide aus der manuellen Bibliothek mit HCV Genotyp 1a. Die Ergebnisse korrelieren miteinander und somit ist eine Vergleichbarkeit der Ergebnisse innerhalb eines Slides möglich. Dabei ist die Anwendung der Intra-Array Normierung von entscheidender Bedeutung, da sie zu vergleichbaren Ergebnissen führt.

4.1.4.3 *Inter-Slide*

Die Inter-Slide-Korrelation ist wichtig, um die Ergebnisse miteinander vergleichen zu können. In Abbildung 4-17 ist beispielhaft eine solche Wechselseitigkeit zwischen zwei verschiedenen Slides graphisch dargestellt. Sie wurden mit demselben Target (HG 1a) gegen die manuell ausgewählten Oligonukleotide hybridisiert. Deutlich zu erkennen ist hier die Korrelation der Signalintensitäten der genotyp-spezifischen Oligonukleotide, die somit einen Vergleich auch zwischen verschiedenen Slides zulässt. Es ist allerdings eine höhere Diskrepanz bezüglich der Signalintensitäten einzelner Oligonukleotide zwischen den Slides zu erkennen als bei den Intra-Slide Vergleichen, z.B. bei M\_1u2B. Die hier vorkommenden Unterschiede sind signifikant, während die Abweichungen bei den restlichen Oligonukleotiden nicht signifikant sind, wie durch die überlappenden Fehlerbalken erwiesen ist.

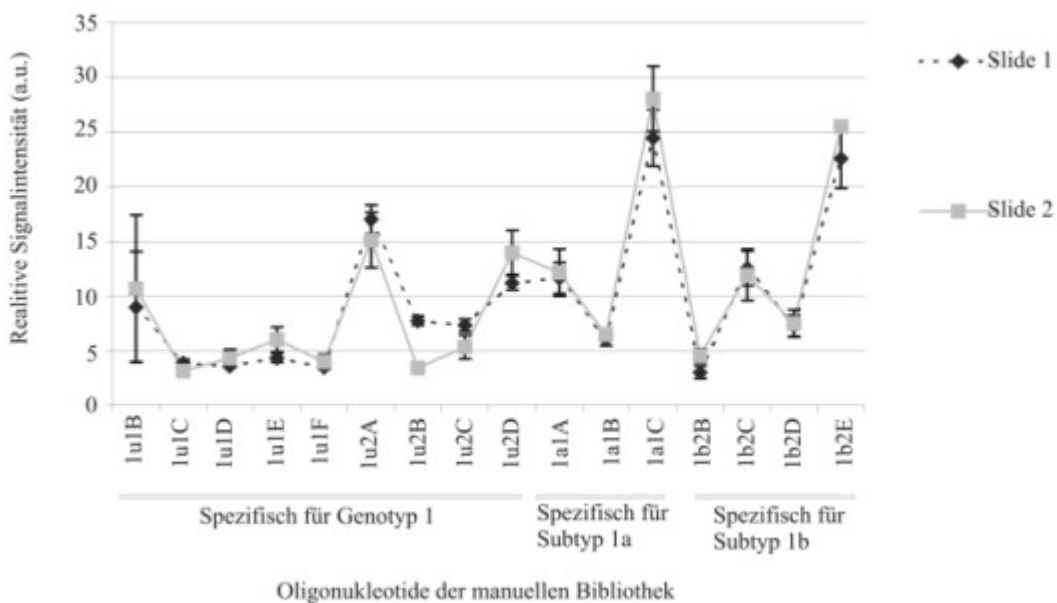


Abbildung 4-17: Graphische Darstellung der relativen Signalintensitäten spezifischer Oligonukleotide der manuellen Bibliothek im Vergleich auf zwei verschiedenen Slides. Es sind nicht-signifikante Unterschiede in der Signalintensität in den meisten Fällen zu erkennen. Eine Ausnahme bildet Oligonukleotide M\_1u-2B. Bei diesem Oligonukleotid sind die Unterschiede in der Signalintensität signifikant.

## 4.2 HCV-Target Charakterisierung

Verschiedene HCV-Targets wurden für die Untersuchung bereitgestellt (vgl. Tabelle 3). Um sie für die Mikroarray-Hybridisierung einsetzen zu können, wurden die RNA zunächst aus dem Serum oder Plasma isoliert, in cDNA umgeschrieben und die gewünschte Region mittels einer nested PCR amplifiziert. Zur Markierung der cDNA für die Hybridisierung werden ein unmarkierter und ein Cy5-markierter Primer eingesetzt. Da eine Zuordnung der Targets zu einem Genotyp nicht immer vorhanden war, wurde eine Sequenzierung angeschlossen. Mit der erhaltenen Sequenz und einem Alignment wurde der Genotyp des Targets durch Vergleiche mit Referenzsequenzen ermittelt (vgl. Abbildung 4-18).

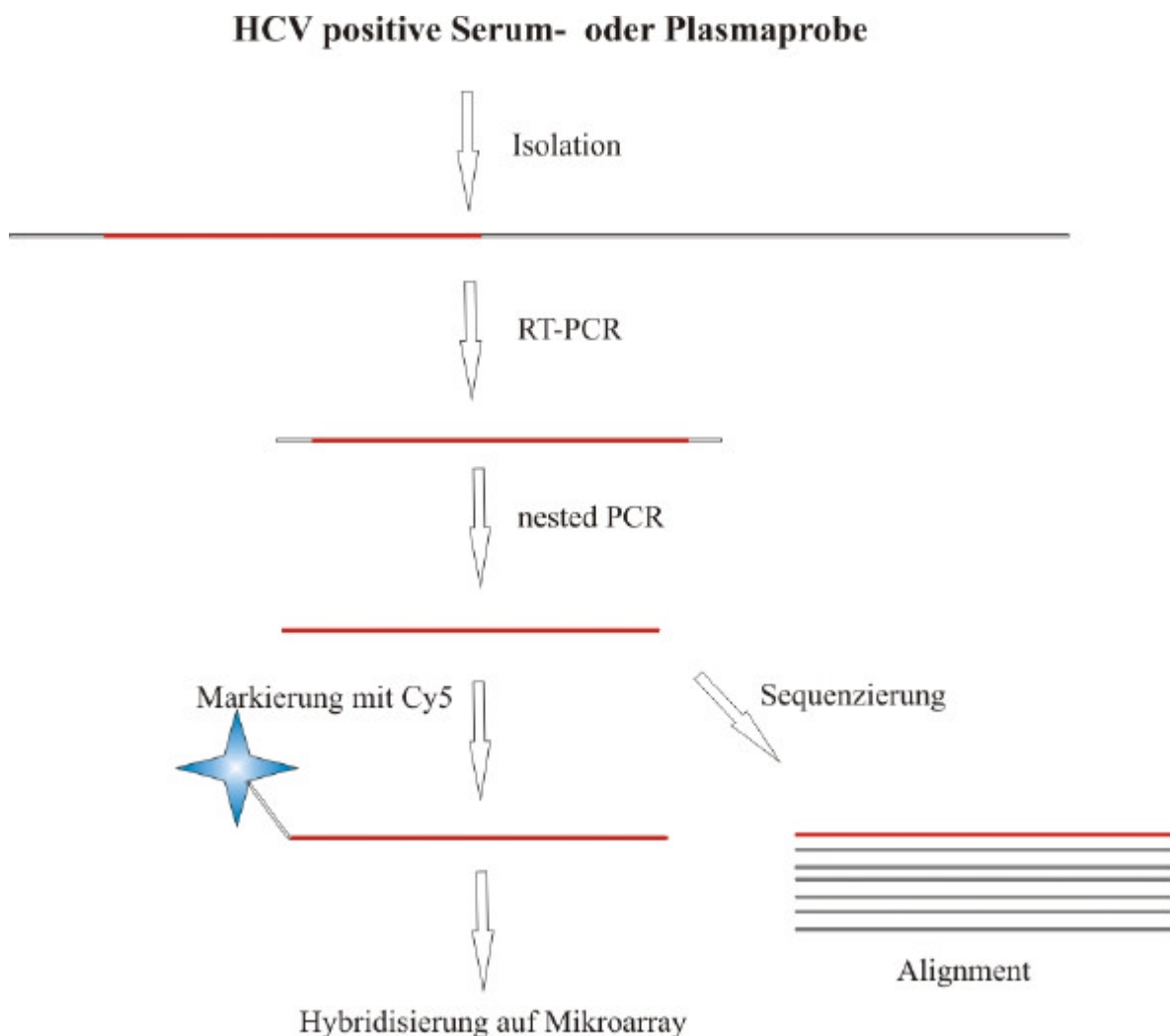


Abbildung 4-18: Schematische Darstellung des Ablaufs bei der HCV-Target Charakterisierung.

### 4.2.1 RNA-Isolation

Ein Vergleich der beiden RNA-Isolationskits von QIAGEN (QIAamp viral RNA) bzw. Sigma (GenElut Mammalian total RNA Kit) wurde durchgeführt. Abbildung 4-19 zeigt bei der Reversen Transkription (RT) PCR nach der Isolation über das QIAamp viral RNA Kit deutliche Banden in der erwarteten Größe von ca. 300 Basenpaaren. Die errechnete Größe der cDNA ist 336 bp, korreliert also mit den Erwartungen. Bei der Aufreinigung über GenElut Mammalian total RNA Kit von Sigma sind keine Banden zu sehen. Bei der „nested PCR“ ist nach der Aufreinigung über den GenElut ebenfalls keine Bande zu erkennen, bei der Isolation über QIAamp viral RNA sind hingegen Doppelbanden zu sehen. Diese sind im erwarteten Bereich, der bei 252 Basenpaaren liegt. Parallel dazu wurden sowohl Positiv- als auch Negativkontrollen mitgeführt. Die Positivkontrolle war eine bereits isolierte und in cDNA umgeschriebene RNA, die in die nested PCR eingesetzt wurde. Als Negativkontrollen dienten PCR-Ansätze ohne Template.

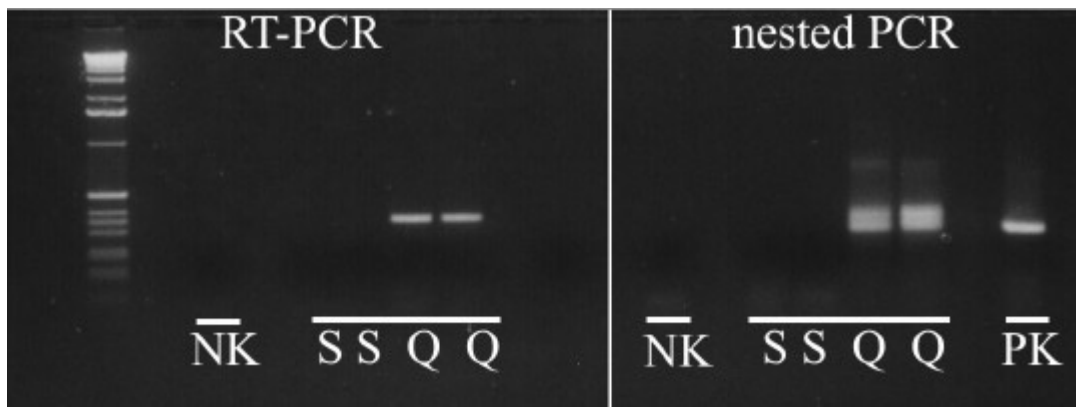


Abbildung 4-19: Vergleich zweier Kits zur DNA-Isolierung, wobei „S“ die Abkürzung für das Isolationskit von Sigma ist und „Q“ für das QIAGEN Kit steht. Die Proben wurden in Doppelbestimmung durchgeführt. Die Reverse Transkription (RT-PCR) zeigt nur bei der Isolation mit QIAGEN Banden in der erwarteten Größe von 336 bp bei der RT-PCR und von 252 bp bei der nested PCR. Die Positivkontrolle (PK) ist ein PCR-Produkt aus einer früheren Aufreinigung. Die Negativkontrollen (NK) sind entsprechende PCR-Reaktionen ohne RNA-Template.



### 4.2.2 Sequenzierung

Nach der Aufreinigung der RNA und dem Umschreiben in cDNA wurden die PCR-Produkte sequenziert. Dies wurde durchgeführt, da zwar z.T. der Geno- bzw. Subtyp bekannt war, jedoch keine Sequenzinformationen vorlagen. Nach der Sequenzierung wurden die Sequenzen editiert und gegen publizierte HCV-Referenzen in einem Alignment verglichen. Dieses ist als Abbildung 8-1 im Anhang dargestellt. Deutlich zu sehen sind hier die Übereinstimmungen der sequenzierten HCV-Targets im Vergleich zu den HCV-Referenzsequenzen aus den öffentlich zugänglichen Datenbanken JJDB und NCBI. Besonders wichtig ist dieses übereinstimmende Ergebnis in den sieben Regionen, die Stuyver zur Klassifikation herangezogen hat, da in diesen Bereich die Sequenzunterschiede zur Identifikation der Geno- und Subtypen liegen <sup>[122]</sup>.

## 4.3 Oligonukleotidbibliotheken

Ziel der Untersuchungen ist es, anhand von zwei unterschiedlicher generierter HCV-Oligonukleotid-Bibliotheken eine Einteilung in die Geno- bzw. Subtypen des Hepatitis C-Virus zu ermöglichen. Dazu wurden Oligonukleotide, wie in Kapitel 4.3.1 beschrieben, manuell ausgewählt und eine zweite Bibliothek wie in Kapitel 4.3.2. beschrieben mit der Software OptiNA erstellt. Zunächst wurden die beiden Bibliotheken mit vier HCV Proben der Targetgruppe 1 (GT 1a, 1b, 3a und 4) getestet. Anschließend wurden die Ergebnisse mit zusätzlichen HCV-Proben verifiziert.

### 4.3.1 Manuell erstellte Oligonukleotidbibliothek

Die auf manuell ausgewähltem Weg erstellte Bibliothek umfasst 54 Oligonukleotide. Eine ausführliche Liste mit entsprechenden Daten zur Länge, zur Schmelztemperatur und der Lage auf der Referenzsequenz mit der Accessionnummer M62321 nach Choo *et al.* <sup>[84]</sup> ist im Anhang aufgeführt (siehe Tabelle 8).

Diese Tabelle umfasst 22 universelle Oligonukleotide und neun spezifisch für Genotyp 1, drei für Subtyp 1a, vier für Subtyp 1b, zwei spezifisch für Genotyp 2, zwei für Subtyp 2a, zwei für Subtyp 2b; sieben spezifisch für Genotyp 3; drei spezifisch für Genotyp 4. Die Lage der Oligonukleotide ist in Abbildung 4-20 schematisch dargestellt.

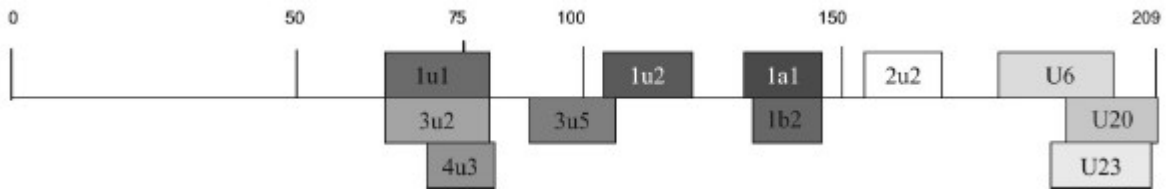


Abbildung 4-20: Vergleichende Darstellung der Target-Positionen der Oligonukleotide aus der manuellen Bibliothek. Diese spiegelt die Lage der Master-Oligonukleotide auf dem cDNA-Strang wieder, mit dem hybridisiert wurde. Die Varianten dieser Master-Oligonukleotide sind in dieser Abbildung nicht berücksichtigt.

Die Nomenklatur der Oligonukleotide spiegelt die Vorgehensweise bei der Auswahl wieder. Es wurden elf Master-Oligonukleotide ausgewählt, wie sie in Abbildung 4-20 aufgezeichnet sind, d.h. diese haben unterschiedliche Bindungsregionen auf dem HCV-Target. Für die Nomenklatur wurde nun zunächst der Genotyp mit einer Nummer benannt, anschließend der Subtyp, danach eine Nummerierung, um die Bindungsregionen unterscheiden zu können. Schließlich folgt ein weiterer Buchstabe – fortlaufend im Alphabet, der die Variante bezeichnet. Am Beispiel **1a1** soll die Nomenklatur verdeutlicht werden: **1a1** steht für Genotyp 1. Am **1a1** ist der Subtyp festgemacht. Bei Oligonukleotiden, die keinen spezifischen Subtyp erkennen, wird hier „u“ für universell an diese Stelle gesetzt. **1u1** bezeichnet die „Bindungsregion“ 1 in der das Oligonukleotid lokalisiert ist. Für dieses Oligonukleotid gibt es fünf Varianten (siehe Tabelle 8), die mit den Buchstaben B-F gekennzeichnet sind. Die universellen Oligonukleotide haben statt der Geno- und Subtypbezeichnung ein „u“ für universell. Die restliche Buchstabenfolge ist analog.

Die Hybridisierungen wurden alle mit den entsprechenden HCV-Targets aus der Targetgruppe 1 (vgl. Kapitel 3.2) durchgeführt.

#### 4.3.1.1 Universelle Oligonukleotide

Aus Abbildung 4-21 ist ersichtlich, dass die universellen Oligonukleotide Hepatitis C-Viren in der Probe detektieren. Die Ergebnisse sind nach Hybridisierung mit Genotyp 1a ermittelt worden. Gut zu sehen sind dabei die verschiedenen Abtrennungsmöglichkeiten der Mismatches abhängig von der Position, an der sie liegen. Die dunkel eingefärbten Balken haben 100 % Targethomologie und zeigen mit 60 % bis 100 % die stärksten Signalintensitäten. Der rot markierte Balken (M\_U6C) hat einen Mismatch an der zweiten

Basenposition. Es ist aber klar zu erkennen, dass sich hier die Signalintensität kaum von der des Perfect Match Oligonukleotides unterscheidet. Liegt die Mismatch Position weiter in der Mitte des Oligonukleotides, kann der Einfluss an der Abnahme der Signalintensität abgelesen werden. Die Oligonukleotide mit zwei oder mehr Mismatch Positionen haben die geringsten Signalintensitäten.

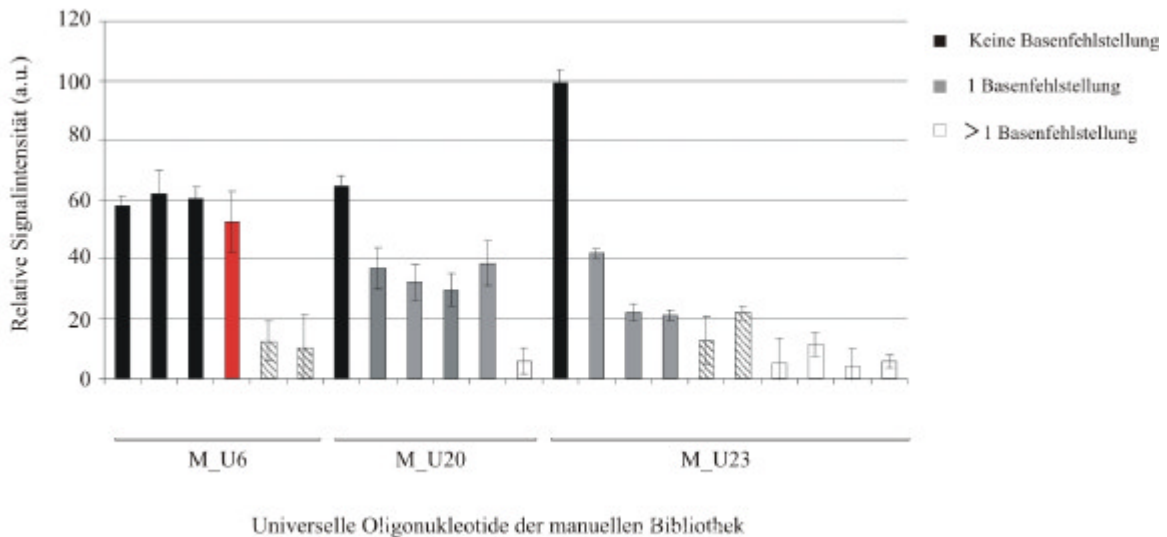


Abbildung 4-21: Graphische Darstellung der Signalintensitäten der universellen Oligonukleotide, abhängig von der Position bzw. der Anzahl und der Lage der Basenfehlpaarungen. Hybridisiert wurden die Oligonukleotide gegen Genotyp 1a aus der Targetgruppe. Schwarz gefärbt sind die Balken mit 100 % Target-Homologie; dunkelgraue Balken repräsentieren Oligonukleotide, die eine endständige Basenfehlpaarungen haben, die schraffierten weisen mittelständige Mismatches auf.

Betrachtet man nur die schwarzen Balken, bei denen die Oligonukleotide keine Basenfehlpaarung zu dem Targetmolekül haben und nach denselben Designparametern ausgewählt wurden, sieht man deutliche Unterschiede in der Signalintensität von 41 %. Die Oligonukleotide mit Basenfehlpaarungen zum Perfect Match haben geringere Signalintensität im Vergleich zu dem Perfect Match. Das ist für alle hier betrachteten Oligonukleotide der Fall. Bei den drei Oligonukleotiden in M\_U6, die auch eine Signalintensität von ca. 60 a.u. erreichen, ist die Basenfehlpaarung an der letzten Base bzw. an der drittletzten Position. Das deutet daraufhin, dass es auf die Position des Mismatches ankommt. Sie muß mindestens sechs Basen vom Ende der Sequenz entfernt sein. Je mehr nicht-passende Positionen in der Sequenz eingebaut sind, desto besser ist die Abtrennung.

4.3.1.2 Hybridisierung mit Genotyp 1a aus der Targetgruppe 1

In Abbildung 4-22 sind die relativen Signalintensitäten aller Oligonukleotide zu sehen, die für Genotyp 1 und die Subtypen 1a und 1b ausgewählt wurden. Die dunklen Balken zeigen Oligonukleotide, die nach einem Sequenzvergleich 100 % Homologie zum Target besitzen, also einen Perfect Match darstellen. Auffällig ist die unterschiedliche Signalintensität der verschiedenen Perfect Match Oligonukleotide 1u1B, 1u2A, 1u2C, 1a1A und 1b2E. Ein eventueller Zusammenhang mit der Zugänglichkeit, bedingt durch die Sekundärstruktur wird in Kapitel 5 diskutiert.

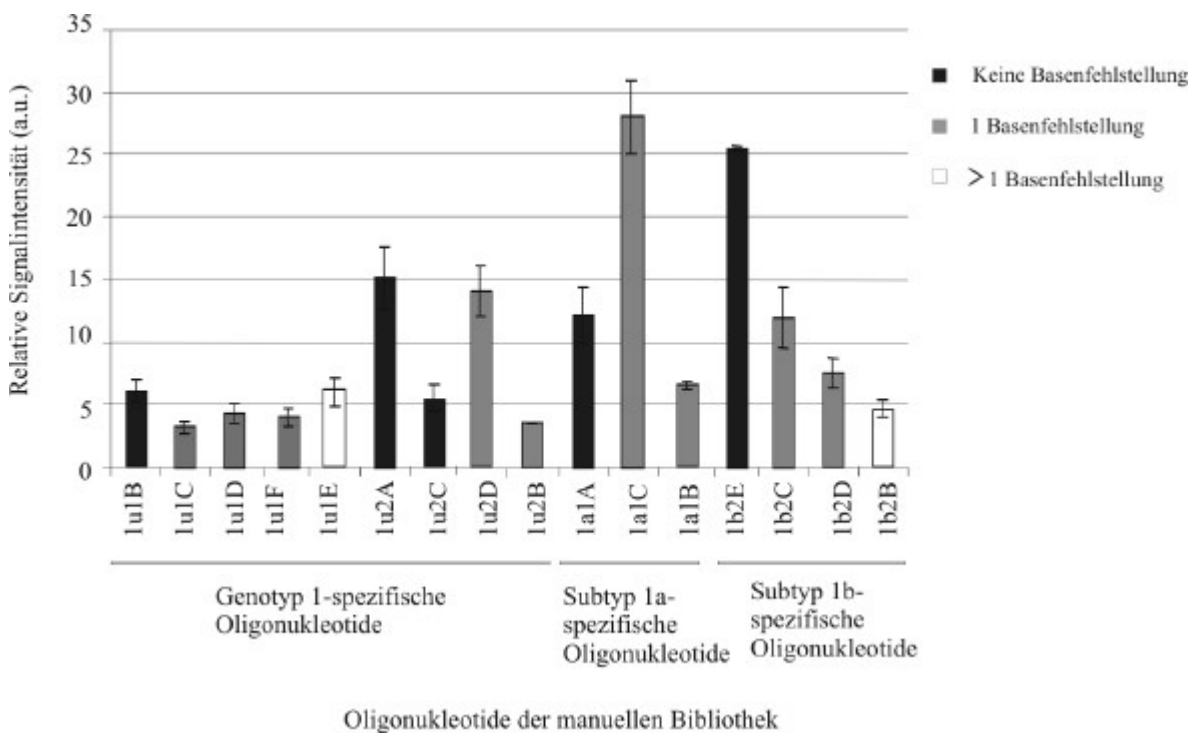


Abbildung 4-22: Vergleichende graphische Darstellung der Signalintensitäten spezifischer Oligonukleotide für Genotyp 1 und Subtypen 1a und 1b der manuellen Bibliothek nach Hybridisierung mit Genotyp 1a aus der Targetgruppe 1. Alle Oligonukleotide für den Genotyp 1 und für die Subtyp 1a und 1b zeigen Hybridisierungsereignisse. Das heißt, dass hier die Spezifität der Oligonukleotide nicht ausreichend ist, um zwischen den Subtypen 1a und 1b zu differenzieren.

4.3.1.3 Hybridisierung mit Genotyp 1b aus der Targetgruppe 1

In diesem Versuch wurden dieselben Oligonukleotide wie in Abbildung 4-22 verwendet, um das Hybridisierungsverhalten eines HCV-Targets mit Genotyp 1b zu untersuchen. In Abbildung 4-23 sind die verschiedenen Signalintensitäten, bezogen auf die nichtdargestellte Positivkontrolle, in einem Balkendiagramm illustriert. Die dunklen Balken zeigen die

Oligonukleotide, die keine Basenfehlpaarung zum Target haben, die Grauen sind Sequenzen mit einen oder mehreren Mismatches. Auch hier ist gut die Abtrennung der Oligonukleotide mit Basenfehlpaarung in Bezug auf die entsprechende Lage zu erkennen. So lässt sich z.B. 1u1D mit seiner Mismatchposition an Position 5 auf der Sequenz anhand der Signalintensität nicht von dem Oligonukleotid ohne Basenfehlpaarung (1u1B) unterscheiden. Auffällig ist, dass Oligonukleotid 1a1B, das zum Targetmolekül perfekt passt, ein deutlich geringeres Hybridisierungssignal zeigt als die entsprechenden Varianten dieses Oligonukleotides mit einer oder mehreren Basenfehlpaarungen.

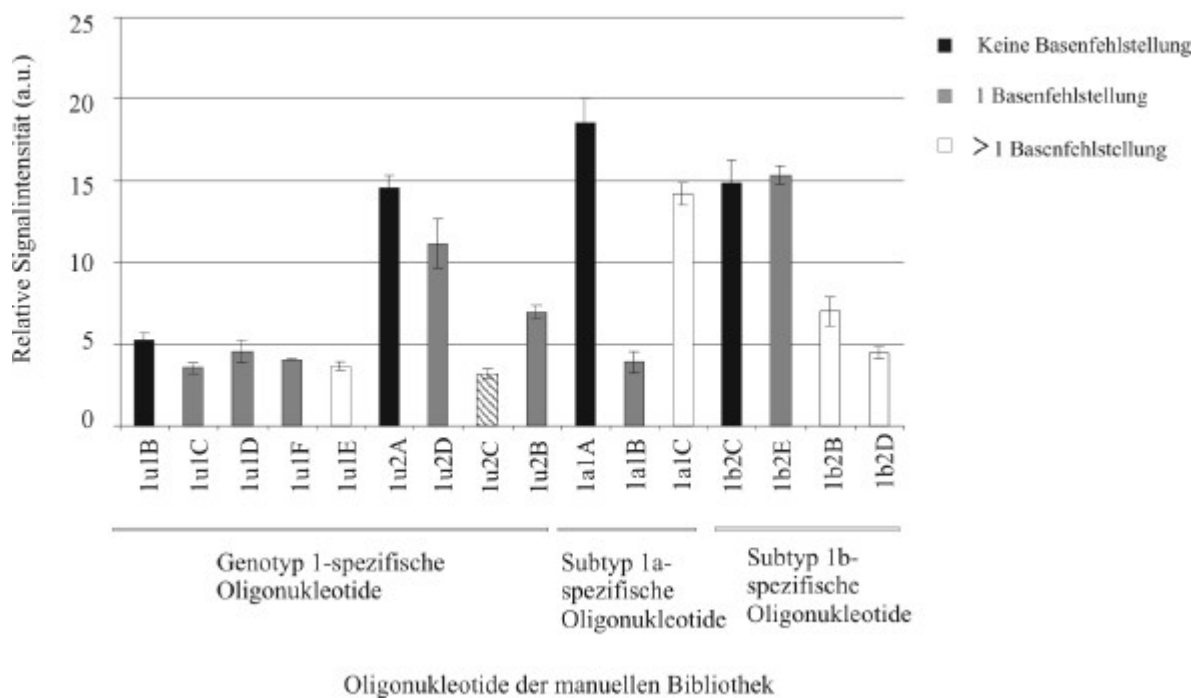


Abbildung 4-23: Graphische Darstellung zum Vergleich der Signalintensitäten spezifischer Oligonukleotide für Genotyp 1 und die Subtypen 1a und 1b der manuellen Bibliothek nach Hybridisierung mit Genotyp 1b. Grau markierte Balken zeigen Oligonukleotide, die eine endständige Basenfehlpaarungen haben, die schraffierten weisen mittelständige Mismatches auf. Alle Oligonukleotide für den Genotyp 1 und für die Subtypen 1a und 1b zeigen Hybridisierungssignale. Mit diesen Oligonukleotiden kann zwischen den Subtypen 1a und 1b nicht differenziert werden.

#### 4.3.1.4 Hybridisierung mit Genotyp 3 aus der Targetgruppe 1

In der manuellen Bibliothek sind vier Oligonukleotide für die Detektion von Genotyp 3 ausgewählt worden. Dabei handelt es sich um Oligonukleotide, die nur den Genotyp erkennen, nicht aber die Subtypen diskriminieren können. Die Abbildung 4-24 zeigt die Signalintensitäten dieser Oligonukleotide nach Hybridisierung mit Subtyp 3a aus der

Targetgruppe 1. Hierbei ist zu sehen, dass die Sequenzen deutlich geringere Signalintensitäten, bezogen auf die Positivkontrolle, erreichen als die Beispiele der Hybridisierungen mit Subtyp 1a oder 1b. Auch eine Abtrennung der Oligonukleotide mit einer Basenfehlpaarung ist hier nicht zu erkennen, obwohl deren Lage nicht innerhalb der ersten fünf Nukleotide liegt.

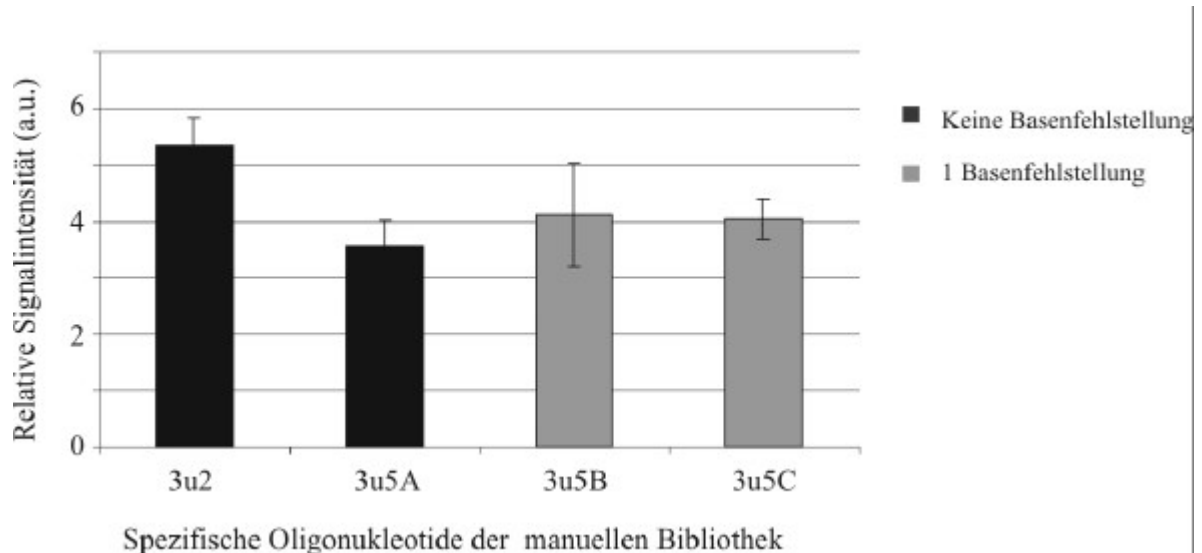


Abbildung 4-24: Vergleichende graphische Darstellung der Signalintensitäten spezifischer Oligonukleotide für Genotyp 3 aus der Targetgruppe 1 der manuellen Bibliothek nach Hybridisierung mit Genotyp 3a. Alle vier Oligonukleotide zeigen niedrige, aber gleichmäßige Hybridisierungssignale, trotz Einzelbasenaustauschen bei zweien der vier Oligonukleotide im Vergleich zu der dazugehörigen Perfect Match Sequenz.

#### 4.3.1.5 Hybridisierung mit Genotyp 4 aus der Targetgruppe 1

Für diesen Genotyp waren die spezifizierten Einträge in den Datenbanken limitiert, sodass nur ein Oligonukleotid M\_4u3 ausgewählt werden konnte. Die Hybridisierungseigenschaften sind in der Zusammenfassung im folgenden Kapitel berücksichtigt.

#### 4.3.1.6 Zusammenfassung

In den vorausgegangenen Abbildungen wurde gezeigt, ob bei den verschiedenen Oligonukleotide nach Hybridisierung mit ihrem Target ein Hybridisierungsereignis gibt oder nicht und ob dieses Signal spezifisch ist. Gezeigt wurden die Ergebnisse der geno- oder subtypspezifischen Oligonukleotide. Ein Vergleich der vier Geno- bzw. Subtypen Abbildung 4-25 verdeutlicht die Möglichkeit der Differenzierung. Für den

Vergleich wurden jeweils die Oligonukleotide ohne Basenfehlpaarung zu ihrem Target herangezogen. Die Auftragung erfolgt in einer logarithmischen Skala, um die hohen Signalintensitäten bei den universellen Oligonukleotiden und die geringen Signalintensitäten z.B. bei Genotyp 3 übersichtlich darzustellen. Zu erkennen ist, dass bei fast allen Oligonukleotiden, auch bei denen mit Basenfehlpaarungen, Hybridisierungssignale bei verschiedenen HCV Targets zu erkennen sind. Die Hybridisierungssignale fallen unterschiedlich stark aus und sind daher deutlich zu unterscheiden, sodass eine Genotypisierung möglich ist. Zwar ergeben die Oligonukleotide für Genotyp 1 immer Hybridisierungssignale, jedoch erscheinen bei Genotyp 3 und 4 zusätzlich die spezifischen Oligonukleotide. Aufgrund dieser Muster können somit die HCV-Genotypen zugeordnet werden. Anders verhält es sich bei der Unterscheidung der Subtypen 1a und 1b. Anhand dieser Oligonukleotide lassen sich diese mit der manuell ausgewählten Oligonukleotid-Bibliothek nicht diskriminieren.

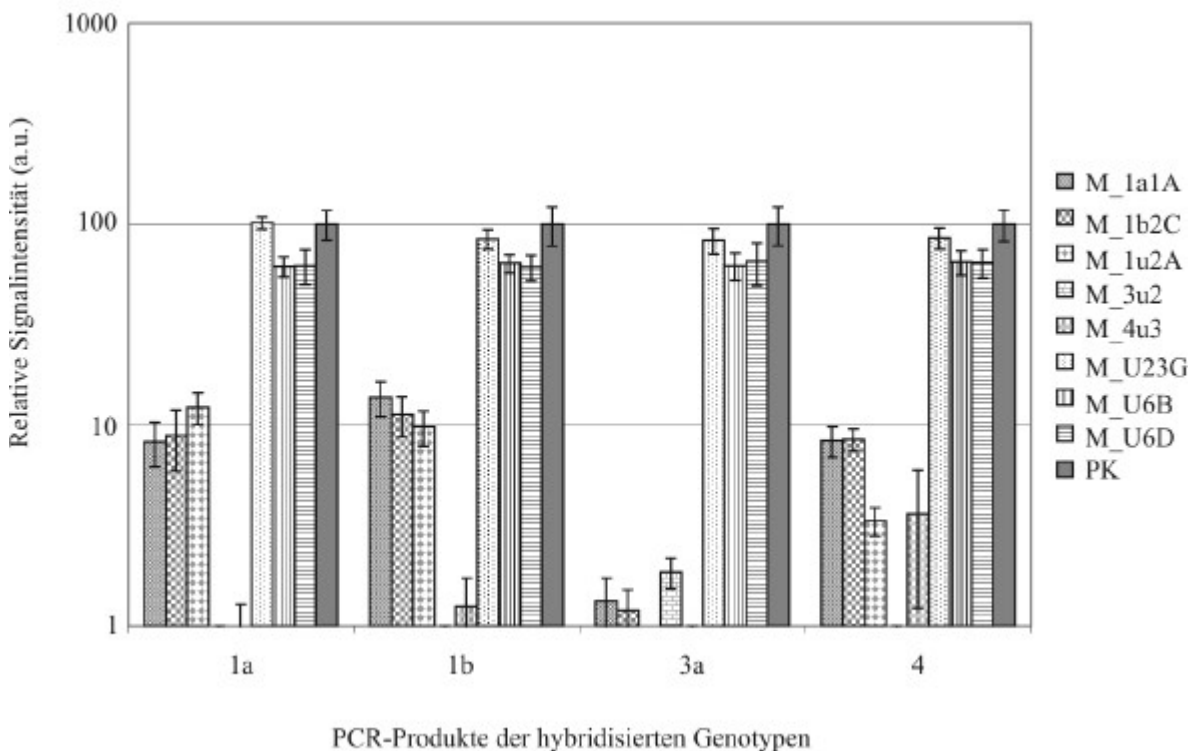


Abbildung 4-25: Vergleich von Signalintensitäten ausgewählter Oligonukleotide der manuellen Bibliothek nach Hybridisierung mit entsprechenden PCR-Produkten aus der Targetgruppe 1 auf verschiedenen, identisch bespotteten Slides. Das Balkendiagramm zeigt ausgewählte Oligonukleotide, die zu ihrem Target 100 % Homologie, zu den anderen jedoch mindestens eine Basenfehlpaarung haben. Die universellen Oligonukleotide (M\_U23G, M\_U6B, M\_U6D) zeigen die höchsten Signalintensitäten von etwa 100 a.u., wie sie für die Positivkontrolle gesetzt wurden. Anhand der spezifischen Sequenzen kann eine Unterscheidung in Genotypen, nicht aber in Subtypen getroffen werden.

### 4.3.2 Software-generierte Oligonukleotid-Bibliothek

Diese Bibliothek beinhaltet 66 Oligonukleotide. Davon werden 20 als universelle Oligonukleotide eingeordnet. Um die Gesamtheit der Hepatitis C-Viren mit Oligonukleotiden zu erfassen, müssen acht dieser zwanzig den genotypspezifischen Oligonukleotiden zugeordnet werden. Im Detail wird darauf in der Diskussion eingegangen. Eine Auflistung der Oligonukleotide mit zusätzlichen Informationen bzgl. der Lage der Oligonukleotide zu der Referenz mit der Accessionnummer M62321 <sup>[84]</sup>, die Schmelztemperatur und die Länge des jeweiligen Oligonukleotides befindet sich im Anhang in Tabelle 9. Darüber hinaus sind neun Oligonukleotide spezifisch für Subtyp 1a generiert (plus zwei für Genotyp 1 aus den universellen Oligonukleotiden), sechzehn für Subtyp 1b; sechs sind spezifisch für Genotyp, acht für Genotyp 3, davon einer für Subtyp 3a und einer für Subtyp 3b, sechs für Genotyp 4. Eine Übersicht der Lage der Oligonukleotide der Software-generierten Bibliothek ist Abbildung 4-26 enthalten.

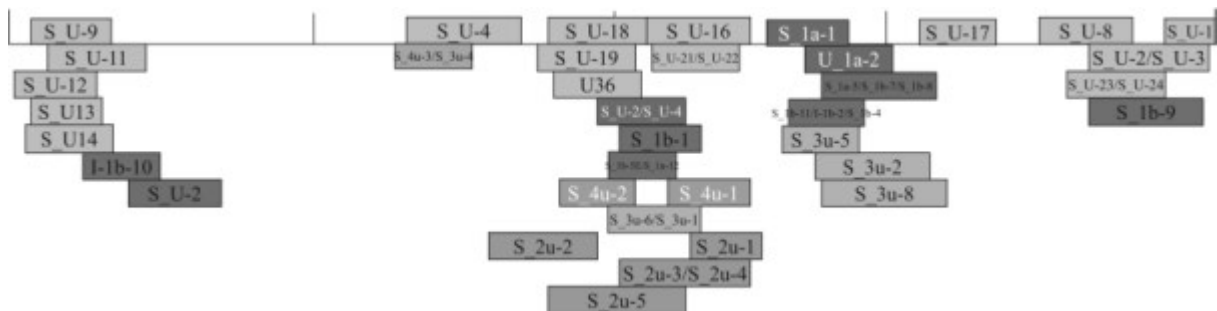


Abbildung 4-26: Vergleichende Darstellung der Target-Positionen der Oligonukleotide aus der Software-generierten Bibliothek. Die Position und die Größe der einzelnen Kästchen korrespondiert mit der Länge des Oligonukleotids und dessen Lage auf dem PCR-Produkt.

Die Nomenklatur bezieht sich hier nicht auf bestimmte Regionen. Der Name der Oligonukleotide besteht aus drei Teilen. Am Beispiel von S<sub>2u-3</sub> soll dies erläutert werden. S<sub>2u-3</sub> steht für die Software-generierte Bibliothek. S<sub>2u-3</sub> zeigt die Geno- bzw. Subtyp-Spezifität an, in diesem Fall für Genotyp 2. S<sub>2u-3</sub> ist eine Durchnummerierung der Oligonukleotide, die für diesen Genotyp ausgewählt wurden. S<sub>U-3</sub> bezeichnet die universellen Oligonukleotide.

Die Hybridisierungen wurden alle mit den entsprechenden HCV-Targets aus der Targetgruppe 1 durchgeführt.



4.3.2.1 Hybridisierung mit Genotyp 1a aus der Targetgruppe 1

Für Genotyp 1 beinhaltet die Software-generierte Bibliothek 25 Oligonukleotide, deren Signalintensität nach Hybridisierung mit Genotyp 1a aus Targetgruppe 1 in Abbildung 4-27 dargestellt ist. Die dunklen Balken markieren die Oligonukleotide, die 100 % Homologie zu dem eingesetzten Target aufweisen. Klar zu erkennen ist, dass die spezifischen Oligonukleotide für Genotyp 1a mehr Perfect Match Oligonukleotide hat. So könnte zumindest in der Theorie die Software-generierten Oligonukleotide spezifischer sein als die manuellen. Der experimentelle Teil zeigt sowohl bei den Sequenzen spezifisch für Subtyp 1a als auch für 1b schwache, aber deutliche Signale. Dabei ist die Signalintensität sehr variabel. Deutlich zu erkennen ist, dass ein Perfect Match Oligonukleotid (z.B. bei S\_1a-4) eine sehr viel geringere Signalintensität aufweist, als ein Oligonukleotid (S\_1b-9), welches zwei Mismatches hat. Die Signale an sich haben relative Intensitäten zwischen 0,2 a.u. und 11 a.u. und sind somit verhältnismäßig schwach.

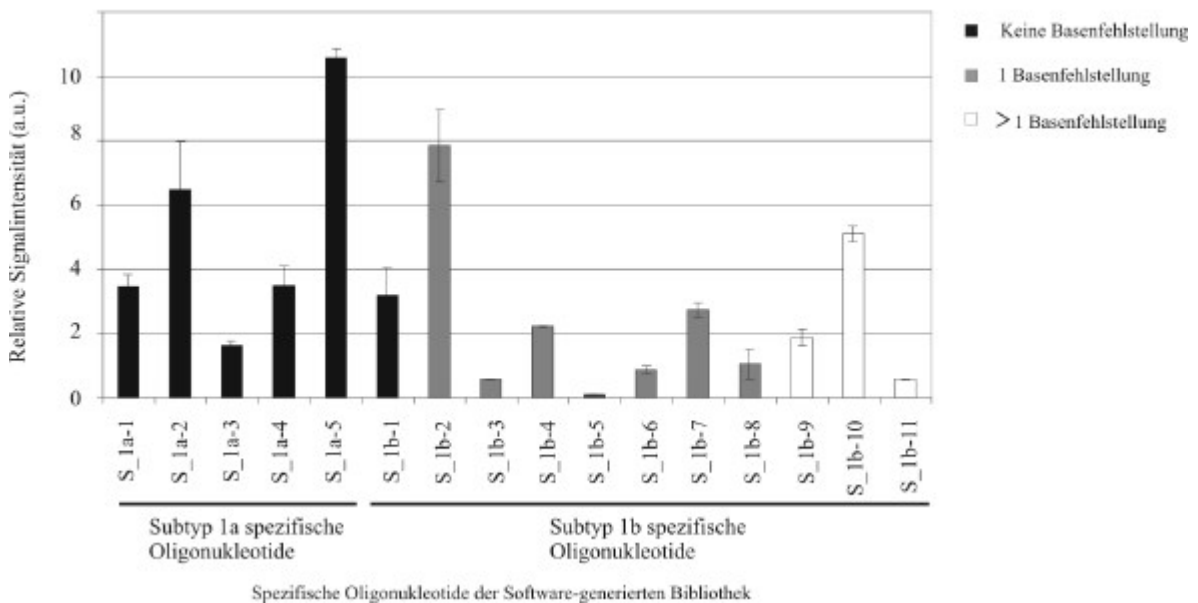


Abbildung 4-27: Graphische Darstellung der Signalintensitäten einer Hybridisierung mit Genotyp 1a aus der Targetgruppe 1. Die schwarz dargestellten Balken repräsentieren Sequenzen ohne Basenfehlpaarung, die Grauen, die mit einer Mismatch Position, während die Weißen Oligonukleotide mit mehr als einer Basenfehlpaarung zwischen Oligonukleotide und Targetsequenz darstellen. Zu erkennen sind auch hier unterschiedliche Signalintensitäten für die verschiedene Oligonukleotide.

4.3.2.2 Hybridisierung mit Genotyp 1b aus der Targetgruppe 1

Für Genotyp 1 beinhaltet die Software-generierte Bibliothek 25 Oligonukleotide, deren Signalintensität nach Hybridisierung mit Genotyp 1b aus Targetgruppe 1 in Abbildung 4-28 dargestellt ist. Schwarz markiert sind Oligonukleotide, die 100 % Homologie zu dem verwendeten Target aufweisen. Von fünf Oligonukleotiden, die für Subtyp 1a ausgewählt wurden, hat eines (S\_1a-5) keine Basenfehlpaarung und weist eine Signalintensität von 16 a.u. auf, während die verbleibenden vier Oligonukleotide eine oder mehr Basenfehlpaarungen haben und deren Signalintensität zwischen 2,1 a.u. und 6,5 a.u. liegt. Von den zehn Subtyp 1b-spezifischen Oligonukleotiden sind fünf Perfect Match zum HCV-Target 1b, zeigen jedoch schwankende Signalintensitäten zwischen 5,7 a.u. und 15,8 a.u. Die Oligonukleotide mit einer Basenfehlpaarung weisen ebenfalls unterschiedliche Signalintensitäten auf. Das Oligonukleotid mit zwei Basenfehlpaarungen zeigt eine Signalintensität von 0,6 a.u..

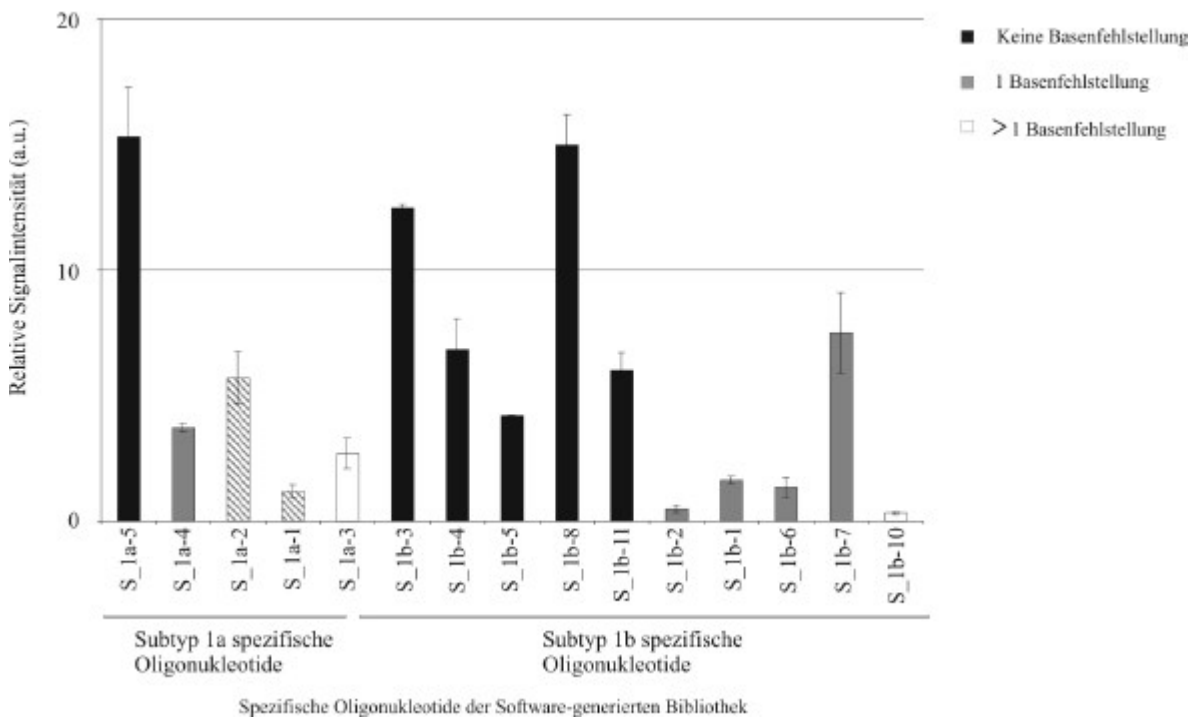


Abbildung 4-28: Graphische Darstellung der Signalintensitäten einer Hybridisierung mit Genotyp 1b aus der Targetgruppe 1. Die schwarz dargestellten Balken repräsentieren Sequenzen ohne Basenfehlpaarung, die Grauen, die mit einer endständigen Mismatch Position, die schraffierten Balken stellen die Oligonukleotide mit mittelständigen Basenfehlpaarungen dar. Die weißen Balken repräsentieren Oligonukleotide mit mehr als einer Basenfehlpaarung zwischen Oligonukleotid und Targetsequenz darstellen. Zu erkennen sind unterschiedliche Signalintensitäten für verschiedene Oligonukleotide.

#### 4.3.2.3 Hybridisierung mit Genotyp 3 aus der Targetgruppe 1

Für Genotyp 3 beinhaltet die Software-generierte Bibliothek zehn Oligonukleotide, deren Signalintensität nach Hybridisierung mit Genotyp 3a aus Targetgruppe 1 in Abbildung 4-29 dargestellt ist. Zu erkennen sind verschiedene Signalintensitäten für die einzelnen Oligonukleotide im vergleichsweise niedrigen Bereich von 0 bis 5,9 a.u.. Dunkel markiert sind Oligonukleotide, die 100 % Homologie zu dem verwendeten Target aufweisen. Die Schwankungsbreite bei den PM Oligonukleotiden reicht von 0,8 a.u. bei S\_3u-4 bis 5,8 a.u. bei S\_3u-2.

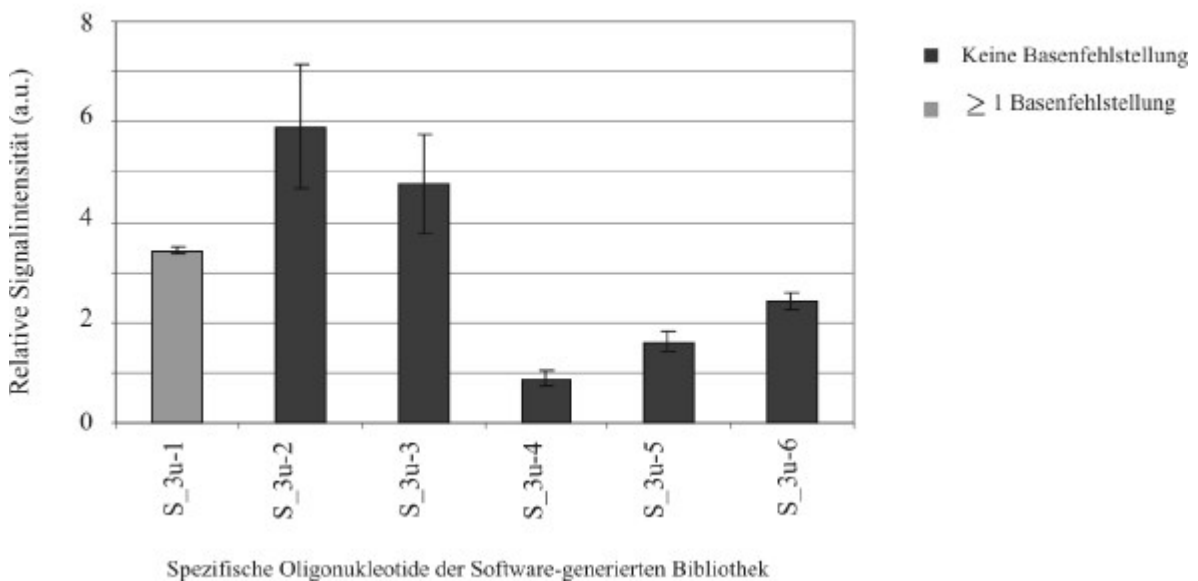


Abbildung 4-29: Graphische Darstellung der Signalintensitäten einer Hybridisierung mit Genotyp 3a aus der Targetgruppe 1. Die dunklen Balken repräsentieren Sequenzen ohne Basenfehlpaarung, die Hellen die mit mindestens einer Mismatch position.

#### 4.3.2.4 Hybridisierung mit Genotyp 4 aus der Targetgruppe 1

Für Genotyp 4 beinhaltet die Software-generierte Bibliothek nur fünf Oligonukleotide, deren Signalintensität nach Hybridisierung mit Genotyp 4 aus Targetgruppe 1 in Abbildung 4-30 dargestellt ist. Das Oligonukleotid S\_U-21 ist spezifisch für Genotyp 4, wurde aber als universelles Oligonukleotid klassifiziert, um Hepatitis C-Viren zu erkennen. Diese Abbildung zeigt Signalintensitäten für die einzelnen Oligonukleotide von null bis drei relativen Signalintensitäten. Die schwarz Balken markieren Oligonukleotide, die 100 % Homologie zu dem eingesetzten Target aufweisen. Nur eine Sequenz, S\_U-21, zeigt Signalintensitäten, die sich vom Hintergrund abgrenzen lassen und so in die Auswertung einfließen können.

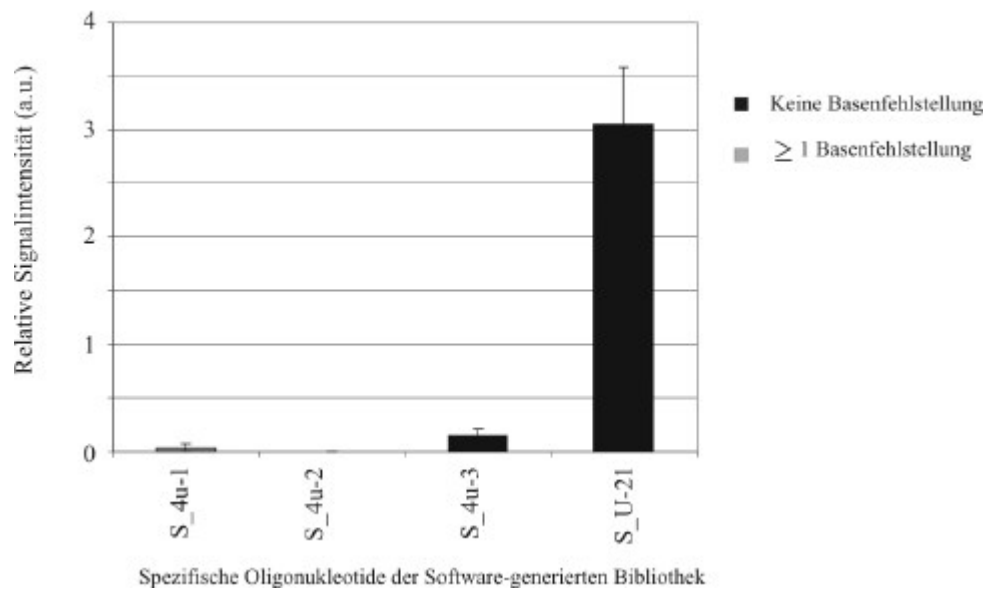


Abbildung 4-30: Graphische Darstellung der Signalintensitäten einer Hybridisierung mit Genotyp 4 aus der Targetgruppe 1. Die dunklen Balken repräsentieren Sequenzen ohne Basenfehlpaarung, die Hellen die mit mindestens einer Mismatch Position. Zusätzlich ist ein universelles Oligonukleotid in die Liste aufgenommen worden. Wie bereits erwähnt, wurden einzelne Genotyp-spezifische Sequenzen als universell gekennzeichnet, um die Detektion von HCV zu gewährleisten. Für die Genotypisierung ist dieses Oligonukleotid jedoch wichtig, da die anderen Oligonukleotide nur sehr schwache Signalintensitäten zeigen.

#### 4.3.2.5 Zusammenfassung

Anhand der ermittelten Hybridisierungsintensitäten und –spezifitäten wurden für die Software-generierte Bibliothek Oligonukleotide ausgewählt, die zu ihrem Target keine Basenfehlpaarungen haben, jedoch zu den anderen HCV-Targets.

Aus Abbildung 4-31 wird deutlich, dass bei Einsatz der Oligonukleotide der Software-generierten Bibliothek eine Einteilung der Genotypen und der Subtypen möglich ist. Die Oligonukleotide, zur Erkennung von Subtyp 1a ausgewählt, zeigen Hybridisierungssignale bei Hybridisierung mit HCV Target 1a, 1b und 4. Hingegen sind die 1b-spezifischen Oligonukleotide spezifisch für ihr Target. Das Oligonukleotid, das für Genotyp 3 ausgewählt wurde, wird auch vom HCV-1b-Target erkannt, während das Oligonukleotid für Genotyp 4 spezifisch sein Target erkennt. Aufgrund der falsch-positiven Signale, die jedoch typisch und für dieses HCV-Target reproduzierbar sind, kann mit einem Hybridisierungsmuster gearbeitet werden. Die Software-generierte Bibliothek ist somit in der Lage, sowohl die Geno- als auch die Subtypen richtig voneinander zu unterscheiden.

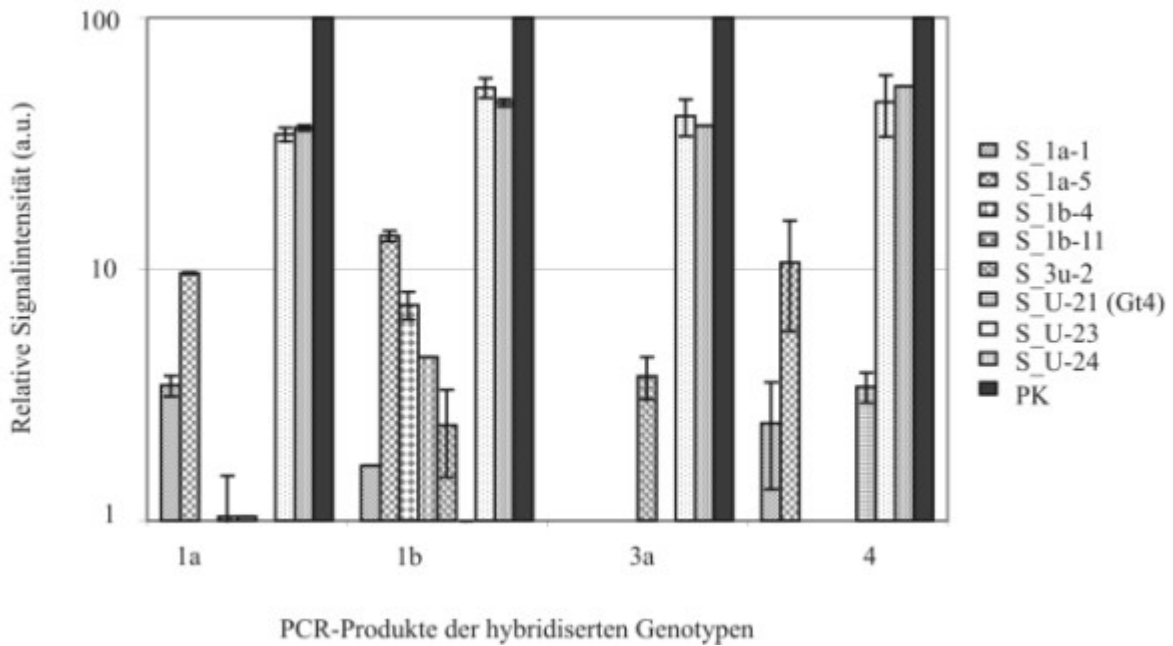


Abbildung 4-31: Vergleich von Signalintensitäten ausgewählter Oligonukleotide der Software-generierten Bibliothek nach Hybridisierung mit entsprechenden PCR-Produkten aus der Targetgruppe 1 auf verschiedenen, identisch bespoteteten Slides. Das Balkendiagramm zeigt Oligonukleotide, die zu ihrem Target 100 % Homologie, zu den anderen jedoch mindestens eine Basenfehlpaarung haben. Die universellen Oligonukleotide (S\_U-23 und S\_U-24) haben die höchsten Werte, sind aber mit Signalintensitäten zwischen 34 a.u. bei S\_U-23 nach Hybridisierung mit HCV-Target 1a und 53 a.u. bei S\_U-24 nach Hybridisierung mit HCV-Target 4 geringer als die auf 100 % gesetzte Positivkontrolle. Anhand der spezifischen Sequenzen kann eine Unterscheidung in Genotypen und in Subtypen getroffen werden.

Multitarget-Hybridisierung (MTH)

MTH sind Hybridisierungen mit zwei verschiedenen HCV Geno- bzw. Subtypen. Die jeweiligen Geno- bzw. Subtypen wurden in equimolaren Mengen hybridisiert. Beispielhaft sind hier die Ergebnisse der Hybridisierungen mit den Subtypen 1a und 1b in Abbildung 4-32 und von Genotyp 1a und Genotyp 3 in Abbildung 4-33 dargestellt. Gezeigt sind die relativen Signalintensitäten ausgewählter Oligonukleotide, die zu ihrem Target keine Basenfehlpaarung zu den anderen Targets jedoch Basenfehlpaarungen haben. Die Hybridisierungsmuster für die Subtypen 1a und 1b aus Abbildung 4-31 sind hier in einer Abbildung zu erkennen. Ebenso wie dort beschrieben, sind unspezifische Hybridisierungssignale bei S\_1b-4 und S\_1b-11 zu erkennen. Ohne Vorkenntnisse welche Genotypen oder Subtypen eingesetzt wurden, ist es nicht möglich anhand der Hybridisierungsmuster die Geno- sowie Subtypen zu unterscheiden.

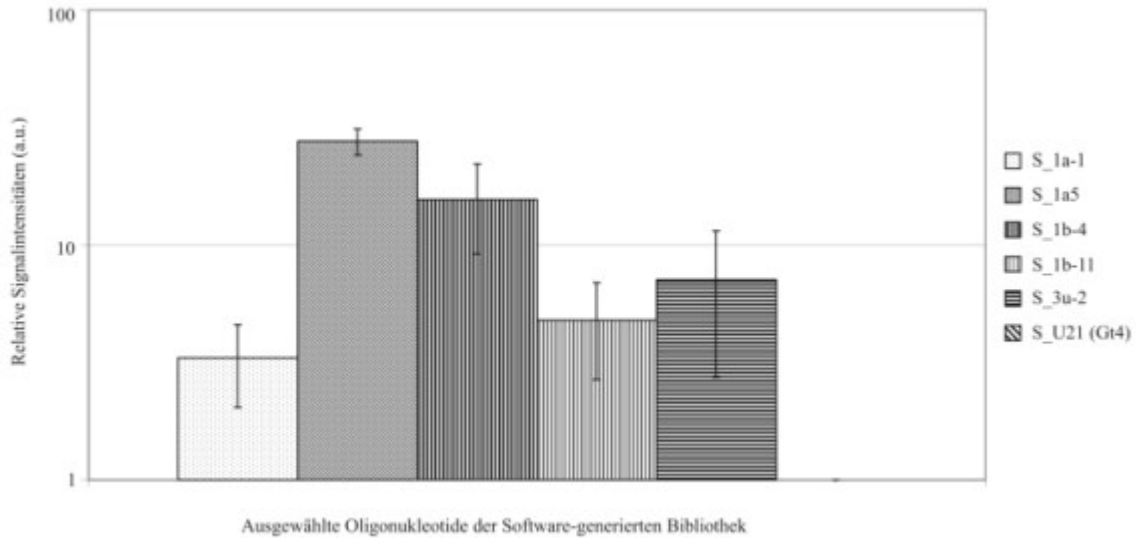


Abbildung 4-32: Graphische Darstellung der relativen Signalintensitäten für ausgewählte Oligonukleotide der Software-generierten Bibliothek nach Hybridisierung mit equimolaren Mengen von Subtyp 1a und Subtyp 1b.

Anders sieht es bei einer Multitarget Hybridisierung mit Subtyp 1a und Subtyp 3 aus. Die Hybridisierungsmuster aus Abbildung 4-31 finden sich hier nicht übereinandergelagert wieder. Für Oligonukleotid S\_1a-1 ist hier kein Hybridisierungssignal detektierbar, dahingegen ist für S\_1b-4 ein Signal nach der Hybridisierung zu erkennen.

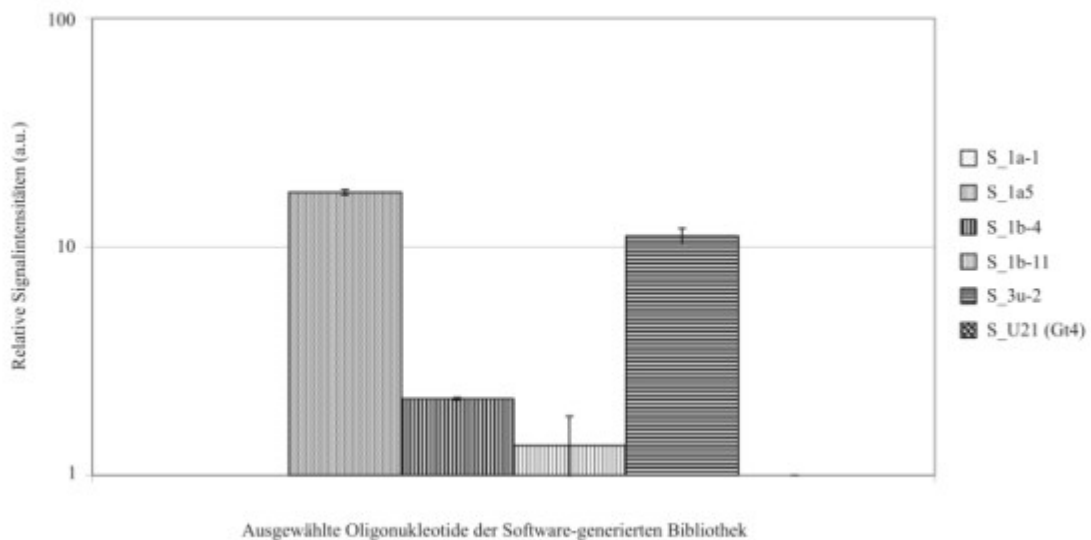


Abbildung 4-33: Graphische Darstellung der relativen Signalintensitäten für ausgewählte Oligonukleotide der Software-generierten Bibliothek nach Hybridisierung mit equimolaren Mengen von Subtyp 1a und Genotyp 3.

Konfigurationsbasierte Auswertung mit der Software iOmega

Neben der Auswertung mit Hilfe der Intra-Array Normierung war es bei der Software-generierten Bibliothek zusätzlich möglich, eine weitere Auswertung anzuwenden. Basierend auf dem Design der Oligonukleotide und dem dahinter steckenden Informationsgehalt wurde eine Software entwickelt. Das Programm „iOmega“ von iSenseIt AG, Bremen <sup>[63]</sup>, baut auf den Daten für das Oligonukleotid-Design auf. Bei der Selektion der Oligonukleotide wurde gegen die Datenbanken geblastet und die hits und scores der einzelnen Oligonukleotide in einer Tabelle gelistet. Diese Informationen sind in die Software als Grundlage für die Bayessche Netze eingeflossen. Daher ist es möglich, eventuelle Kreuzreaktionen und Mismatches zu berücksichtigen und entsprechend herauszurechnen. Drei verschiedene Modelle stehen in der Software zur Verfügung. Beim Simple-model wird die Hierarchie des Organismus und die Signalintensitäten nach der Hybridisierung berücksichtigt. Das One-target-model zieht die Daten zur Kreuzreaktivität und Informationen zu den Mismatches aus dem Design zu Berechnung des wahrscheinlichsten HCV-Genotypes heran. Das Complex-model sollte verwendet werden, wenn mehr als eine Zielsequenz zur Hybridisierung im selben Ansatz verwendet wurde. Die Auswertung mit dem one-target-model für die vier Genotypen (GT) ist in Tabelle 4 aufgezeigt.

Tabelle 4: Auswertung der Hybridisierungsereignisse der Targetgruppe A (HG) mit der IQMega-Software unter Verwendung des one-target-models.

Identifizierung	Genotyp	Ermittelte Wahrscheinlichkeit [%]				
		GT 1a	GT 1b	GT 2	GT 3	GT 4
HG 1a	1a	89	1	1	1	1
HG 1b	1b	3	87	1	1	1
HG 3	3	1	1	1	89	1
HG 4	4	89	1	1	1	1

Mit einer Wahrscheinlichkeit von 87 % bzw. 89 % lassen sich alle Genotypen richtig zuordnen.

#### 4.4 Verifizierung mit weiteren HCV-Targetts

Die Ergebnisse der bisherigen Hybridisierungsexperimente zeigten, dass Hybridisierungsmuster zur Identifizierung der Geno- bzw. Subtypen herangezogen werden könnten. Wie in Kapitel 1.2.2 einleitend beschrieben, sind die Sequenzunterschiede im Genom des HCV hoch, sodass die bisherigen Ergebnisse mit weiteren HCV-Isolaten verifiziert werden müssen. Für die Überprüfung der Hybridisierungsmuster mit klinischen HCV Proben 4.4 wurden die Slides mit dem Spotarray24 von Perkin Elmer bespottet. Dieses Spottingssystem ist schneller als der Ultramikro-Pipettierroboter und erlaubt es, alle zu untersuchenden Spots auf einem Slide abzulegen, was wegen der Gerätecharakteristik mit dem Ultramikro-Pipettierroboter nicht möglich war. Der Spotarray hingegen ist in der freien Positionierung der Oligonukleotide limitiert, sodass die Intra-Array Normierung nicht mehr angewendet werden konnte. Untersuchungen zeigten jedoch, dass sich die Slidehomogenität durch Optimierungen im Herstellungsprozeß wesentlich verbessert hat, sodass auf die Intra-Array-Normierung verzichtet werden konnte. Es ist ausreichend, die Werte der einzelnen Oligonukleotide in Beziehung zum Mittelwert aller Positivkontrollen zu setzen, wie deutlich in Abbildung 4-34 abzulesen ist.

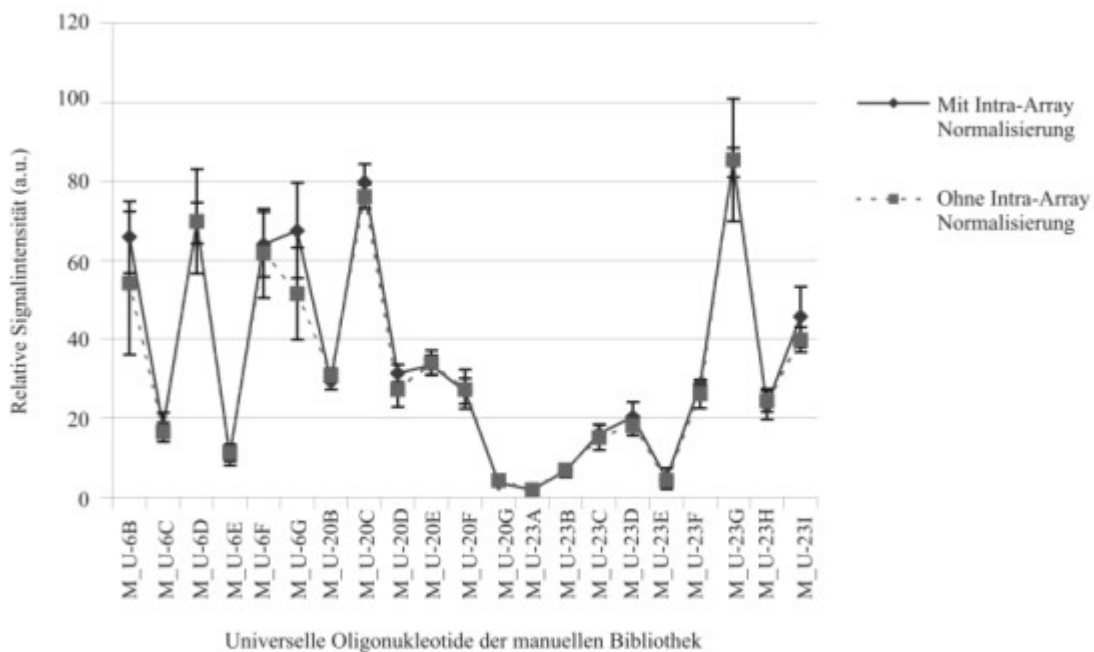


Abbildung 4-34: Vergleich der Signalintensitäten universeller Oligonukleotide der manuellen Bibliothek mit und ohne Intra-Array Normalisierung. Für die Normalisierung ohne Intra-Array Normalisierung wurde aus den Signalintensitäten der gleichen Oligonukleotiden der Mittelwert gebildet und dieser in Relation zum Mittelwert aller Positivkontrollen gesetzt. Hier ist eine sehr gute Korrelation zwischen den beiden Auswertemethoden zu sehen, mit Ausnahme der beiden Oligonukleotiden M\_U6B und M\_U6G.



Die ermittelten Hybridisierungsmuster aus Targetgruppe 1 wurden mit weiteren Plasma und Serumproben (vgl. Targetgruppen 2, 3, 4) verifiziert. Diese zusätzlichen Proben enthalten nicht nur die Geno- und Subtypen 1a, 1b, 3 und 4, sondern auch noch den Genotyp 2, vertreten durch die Subtypen 2a und 2b. Darüber hinaus enthalten sie die Subtypen 3a und 3b vom Genotyp 3. Oligonukleotide für die Subtypen 2a und 2b sind ausgewählt worden, konnten aber bisher nicht getestet werden. Eine Klassifizierung in die Subtypen 3a und 3b ist aufgrund nicht verfügbaren Oligonukleotide nicht möglich.

#### 4.4.1 Manuelle Oligonukleotid-Bibliothek

Zehn verschiedene spezifische Oligonukleotide wurden für diese Auswertung näher untersucht. Dessen Hybridisierungsmuster wurde für insgesamt 15 verschiedene HCV-Proben verglichen, wobei mit einer Ausnahme (GT-5) zwei HCV-Isolate desgleichen Subtyps analysiert wurden. In Abbildung 4-35 sind diese Ergebnisse zusammengefasst. Werden zunächst die beiden Proben desselben Subtyps miteinander verglichen, fällt auf, dass die beiden Hybridisierungsmuster nicht übereinstimmen. Dabei variiert nicht nur die Stärke des Signals, sondern auch das Vorhandensein des Signals. Beispielsweise ist bei GT-1a-I ein spezifisches Signal bei M\_1u2A und M\_1a1A zu erkennen und damit der Subtyp exakt beschrieben. Hingegen zeigen sich zusätzliche Signale bei den Oligonukleotiden M\_1b2C und M\_2u2B nach Hybridisierung mit der Probe GT-1a-II. Das wiederum lässt keine eindeutige Einteilung zu. Ebenso verhält es sich mit den Isolaten GT-1b-I und GT-1b-II. Beim ersteren zeigen sowohl die Genotyp 1-spezifischen Proben (M\_1u2A, M\_1a1A, M\_1b2C) und das Genotyp 2-spezifische Oligonukleotid M2u2B und die Sequenzen für Genotyp 3 (M\_3u2; M\_3u5A; M\_3u5B) ein Hybridisierungsereignis an, während bei der Probe GT-1b-II nur die genannten Fänger für Genotyp 1 und 2 hybridisieren. Ein Vergleich der Hybridisierungsmuster von Probe GT-1a-II und GT-1b-II zeigt, dass beide identisch sind. Die in Kapitel 4.3.1 erhaltenen Resultate, dass mit den Oligonukleotiden der manuellen Bibliothek die Subtypen 1a und 1b nicht diskriminiert werden können, bestätigten sich hier. Bei den Proben für Genotyp 2 (GT-2a-I, GT-2a-II, GT-2b-I, GT-2b-II) sind spezifische Signale bei den entsprechenden Oligonukleotiden zu erkennen. Auch die zusammengehörigen Hybridisierungsmuster sind sich ähnlicher als bei den Oligonukleotiden für Genotyp 1. Eine Unterscheidung anhand der jeweils spezifischen Fängermoleküle ist ebenfalls möglich, da bei beiden Hybridisierungen mit Isolaten desselben Subtyps das dafür ausgewählte Oligonukleotid ein Signal zu detektieren war.

Darüber hinaus unterscheiden sich die Hybridisierungsmuster zwischen Genotyp 1 und Genotyp 2, was eine Unterscheidung zwischen diesen beiden Genotypen ermöglicht. Ein Vergleich der Daten der Hybridisierungen bei den Isolaten für Genotyp 3 zeigt eine gute Korrelation der Hybridisierungsmuster der jeweiligen Subtypen untereinander. Dabei fällt auf, dass M\_3u5B bei beiden Experimenten mit Isolaten der Subtypen 3a ein Signal ergibt, jedoch nicht nach Hybridisierung mit den Subtypen 3b. Nach Analyse der Versuche mit Genotyp 4 (GT-4a-I und GT-4a-II) zeigt sich ein inhomogenes Muster. Vor allem das spezifische Oligonukleotid M\_4u3 erscheint nur bei GT-4a-II. Die Muster von den Genotypen 1 und 4 sind sich sehr ähnlich. Für Probe 5 (GT-5-I) ist kein eindeutiges, sich von den anderen unterscheidbares Hybridisierungsmuster zu sehen. Es zeigen sich Signalintensitäten bei den Oligonukleotiden für Genotyp 1.

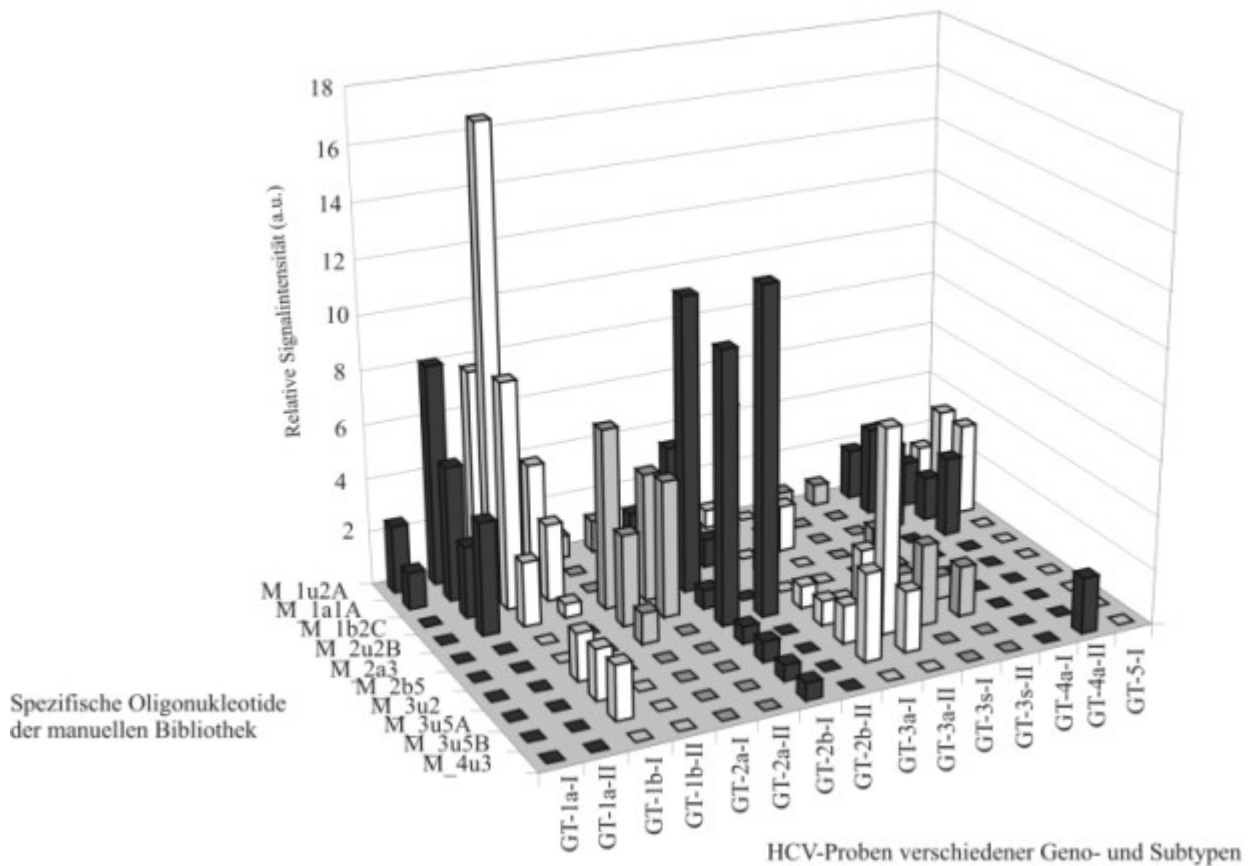


Abbildung 4-35: Überblick über die verschiedenen Hybridisierungsmuster der einzelnen Genotypen mit unterschiedlichen Proben derselben Geno-bzw. Subtyps der manuellen Bibliothek.

### 4.4.2 Software-generierte Oligonukleotidbibliothek

Die Software-generierte Oligonukleotid-Bibliothek wurde auf zweierlei Weisen analysiert: Zum einen mit der Intra-Array Normierung wie Kapitel 4.4.2.1 beschrieben und mit der IQ-Mega-Software, vgl. Kapitel 4.4.2.2.

#### 4.4.2.1 Analyse mit Intra-Array Normierung

Aus der Software-generierten Bibliothek wurden elf Oligonukleotide selektiert, die zu ihrem Target keine Basenfehlpaarung haben, jedoch zu den anderen, gegen die sie diskriminiert werden sollen. Aufgetragen in Abbildung 4-36 sind die Hybridisierungsmuster dieser elf Oligonukleotide bei verschiedenen Genotypen. Dabei sind pro Geno- bzw. Subtyp jeweils zwei Isolate miteinander zu vergleichen.

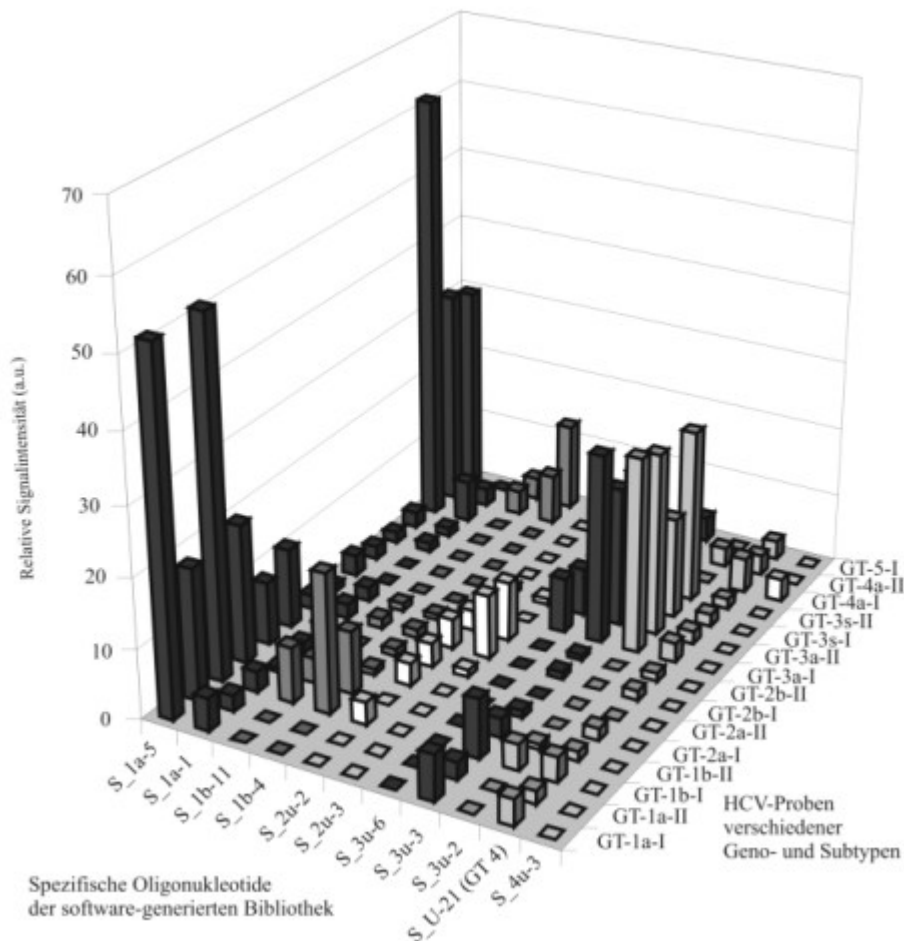


Abbildung 4-36: Überblick über die verschiedenen Hybridisierungsmuster der einzelnen Genotypen mit unterschiedlichen Proben desselben Geno- bzw. Subtyps bei Oligonukleotiden der Software-generierten Bibliothek.

Mit diesen Oligonukleotiden sind die Hybridisierungsmuster bei beiden Isolaten pro Geno- bzw. Subtyp ähnlicher als bei der manuellen Bibliothek. Die Isolate des Subtyps 1a zeigen selektiv spezifische Signale bei S\_1a-1 und S\_1a-5, während die des Subtyps 1b zusätzliche Hybridisierungsereignisse bei S\_1b-11 und S\_1b-4 aufweisen. Somit sind die Subtyp 1a spezifischen Oligonukleotide nicht selektiv für 1b, jedoch kann eine Zuordnung über das Hybridisierungsmuster erfolgen. Auch die Subtypen 2a und 2b zeigen eine charakteristische Verteilung der Signale. Eine Unterscheidung in die beiden Subtypen, könnte möglich sein, da bei Oligonukleotid S\_2u-2 ein Hybridisierungsereignis nach allen vier Isolaten detektiert werden kann. Jedoch wird ein Hybridisierungsereignis bei Sequenz S\_2u-3 nur bei dem Subtyp 2b detektiert. Die Proben mit den Subtypen 3a und 3b werden selektiv erkannt, können aber nicht diskriminiert werden. Für Subtyp 4 ist eine Abgrenzung zu anderen HCV-Typen schwierig, denn das spezifische Oligonukleotid S\_U-21 (GT 4) erscheint bei allen anderen Proben ebenfalls und an die zweite Sequenz (S\_4u-3) hybridisiert nur eines der zwei Isolate.

#### 4.4.2.2 Analyse mit iOmega-Software

Für die Analyse mit Hilfe der Software wurde die GenePix Result (gpr)-Datei mit den Rohdaten aus der Auswertesoftware des GenePix Programmes importiert. Durch die Wahl des Models und des Diskrimination values wurde eine Wahrscheinlichkeit des vorhandenen Genotypes ermittelt. Aufgeteilt nach den untersuchten Targetgruppen sind die ermittelten Wahrscheinlichkeiten in den nachfolgenden Tabellen (siehe Tabelle 5, Tabelle 6 und Tabelle 7) aufgeführt. In der Software wurden die Wahrscheinlichkeiten unter Berücksichtigung der Subtypen 1a und 1b, sowie der Genotypen 2, 3, 4 sowie „others“ zugeordnet. Eine Aufteilung in die Subtypen 2a, 2b oder 3a, 3b kann nicht vorgenommen werden.

Die Serumproben aus Tabelle 5 wurden vom Labor Dr. Gärtner, Weingarten, zur Verfügung gestellt. Die Zuordnung der Genotypen erfolgte bereits in diesem Labor mittels einer vertraulichen *in-house* PCR Methode <sup>[136]</sup>. Zu erkennen ist in fast allen Fällen eine Übereinstimmung der ermittelten Wahrscheinlichkeiten mit der tatsächlichen Klassifizierung. Bei den Subtypen 1a und 1b, sowie den Genotypen 2 und 3 waren die ermittelten Genotypen mit der vorher bestimmten Klassifizierung kongruent. Die Isolate des Genotypes 4 wurden mit der Auswerte-Software fälschlicherweise dem Subtyp 1b zugeordnet.

Tabelle 5: Auswertung der Hybridisierungsereignisse der Targetgruppe B (RF) mit der IQMega-Software unter Verwendung des one-target-models.

Identifizierung	Genotyp	Ermittelte Wahrscheinlichkeit [%]				
		GT 1a	GT 1b	GT 2	GT 3	GT 4
510638	1b	1	89	1	1	1
580431	4	1	89	1	1	1
520100	1a	87	1	1	1	3
519916	3a	1	1	1	89	1
590546	3a	1	1	1	89	1
538102	1a	87	1	1	1	3
549500	2a	2	1	87	1	1
509365	1b	3	87	1	1	1
529996	1a	88	1	1	1	2
508932	3a	1	1	1	89	1
509450	1b	3	87	1	1	1
538716	3a	1	1	1	89	1
509365	1b	3	87	1	1	1
582363	2b	1	1	89	1	1
507760	3b	1	1	1	89	1
514046	3b	1	1	1	89	1
02298	4	1	89	1	1	1

Tabelle 6: Auswertung der Hybridisierungsereignisse der Targetgruppe C (Panel) mit der IQMega-Software unter Verwendung des one-target-models.

Identifizierung	Genotyp	Ermittelte Wahrscheinlichkeit [%]				
		GT 1a	GT 1b	GT 2	GT 3	GT 4
Panel	1a	89	1	1	1	1
Panel	1b	24	66	1	1	1
Panel	2a	2	1	88	1	1
Panel	2b	1	1	89	1	1
Panel	2i	2	1	84	1	4
Panel	3a	6	1	1	84	1

Die Targetgruppe C umfasst Proben, die vom Referenzzentrum für Hepatitis C an der Universität Essen erworben werden können<sup>[129]</sup>. Sie beinhaltet elf HCV-positive Serum-Proben mit neun unterschiedlichen Geno- bzw. Subtypen. Die Klassifizierung wurde mit verschiedenen Methoden durchgeführt (Vgl. Tabelle 3). Mit der IQMega-Software konnte diesen Isolaten der Subtyp 1a und die Genotypen 2 und 3 korrekt zugeordnet werden. Bei dem Isolat 1b ist die Wahrscheinlichkeit mit 66 % deutlich niedriger als bei den anderen Isolaten mit > 84 %.

Die Isolate aus Tabelle 7 wurden von Dr. Britta Becker, Abteilung für Virologie, Universität Bremen, zur Verfügung gestellt. Bei einem Teil der Proben war der Genotyp unbekannt. Die Zuordnung zu einem bestimmten Geno- bzw. Subtyp konnte mit Hilfe dieser Software vorgenommen werden. Die Wahrscheinlichkeiten für diese Fälle sind hierbei wieder > 82 % sehr hoch. Für die Proben mit bekanntem Geno- bzw. Subtyp war die Zuordnung korrekt.

Tabelle 7: Auswertung der Hybridisierungsereignisse der Targetgruppe 4 (BB) mit der *IQMega*-Software unter Verwendung des one-target-models.

Identifizierung	Genotyp	Ermittelte Wahrscheinlichkeit [%]				
		GT 1a	GT 1b	GT 2	GT 3	GT 4
BB#1	unknown	5	85	1	1	1
BB#2	unknown	1	6	82	1	1
BB#3	unknown	87	1	1	1	2
BB#4	unknown	1	1	1	88	1
BB#5	1b	5	85	1	1	1
BB#6a	1a	84	1	1	1	3
BB#6b	1a	75	1	1	1	10

## 5 DISKUSSION

Die Anwendungsgebiete von Mikroarrays sind sehr breit gefächert. Aktuell werden sie hauptsächlich für die Untersuchung differentiell exprimierter Gene in sogenannten Genexpressions-Studien eingesetzt. Im Gegensatz dazu befasst sich diese Arbeit mit der Organismen-Identifikation mit Hilfe von DNA-Mikroarrays. In dieser Fragestellung bieten Mikroarrays den großen Vorteil, verschiedene Proben im hochparallelen Ansatz durchführen zu können.

Das Hepatitis C-Virus wurde als Modellorganismus ausgesucht, wobei die Klassifizierung in verschiedene Geno- und Subtypen eine besondere Herausforderung darstellte. So wurden in dieser Arbeit insgesamt 122 verschiedene Oligonukleotide untersucht. 56 davon wurden per Hand ausgewählt und in der manuellen Bibliothek zusammengefasst. 66 Oligonukleotide wurden mittels eines eigens entwickelten Programms ausgewählt und als Software-generierte Bibliothek in dieser Arbeit analysiert. Die Generierung der verschiedenen Oligonukleotide war ein Hauptaspekt, wie in Kapitel 3.1 beschrieben wird. Der zweite wichtige Punkt, der in dieser Arbeit behandelt wurde, war die Etablierung und Optimierung der Mikroarrayherstellung und -hybridisierung, wie in Kapitel 4.1 beschrieben. Dabei wurde eine Auswertestrategie entwickelt, um Fehler im Herstellungsprozess zu minimieren, wie in Unterkapitel 4.1.4.1 beschrieben.

### 5.1 Etablierung und Optimierung der Mikroarrays

Eine Herausforderung war die parallele Etablierung des Roboters, der aktivierten Oberflächen und die Optimierung der Hybridisierung. Der Vorteil des verwendeten Roboters war die freie Positionierbarkeit jedes Oligonukleotids an jedem beliebigen Ort auf der aktivierten Oberfläche. Jedoch war die Piezopipette anfällig für Staubpartikeln, die sich an der Düse festsetzten und somit die Tropfenabgabe stark beeinträchtigten. Dadurch war das Spotten einer großen Anzahl von Oligonukleotiden nicht möglich. Infolgedessen wurden die Oligonukleotide zum Teil nicht zuverlässig abgelegt, so dass zunächst keine vergleichbare Datenlagen erzeugt werden konnten. Nach innovativer Optimierung ist es jedoch gelungen und es wurden optimale Hybridisierungsparameter für die Oligonukleotide ermittelt. Wie die einzelnen Parameter bei der Hybridisierung jedoch zusammenwirken, ist nicht geklärt. Es gibt Literatur zu den verschiedenen Einflüssen. Southern *et al.* <sup>[137]</sup>

beschreiben den Einfluß in der Zusammensetzung der Basen und der Sequenz des Oligonukleotides. Shchepinov *et al.* <sup>[138]</sup> beschreiben den Einsatz von Spacern zur Erhöhung der Signalintensität. Sie fanden heraus, dass die optimale Spacerlänge bei 40 Atomen liegt, was zu einer 150fachen Verstärkung des Signals führt. Peterson *et al.* haben Untersuchungen zum Effekt der Probendichte auf der Oberfläche auf die Hybridisierung durchgeführt und konnten zeigen, dass mit einer geringen Probendichte von  $3 \times 10^{12}$  Molekülen/cm<sup>2</sup> alle Proben binden konnten und die Hybridisierungskinetik vergleichsweise schneller verlief als bei hohen Konzentrationen ( $1,2 \times 10^{13}$  Molekülen/cm<sup>2</sup>). Mir *et al.* haben den Einfluß der Struktur auf die Hybridisierung bei Oligonukleotid-Mikroarrays untersucht <sup>[139, 140]</sup> und zeigten die bevorzugte Bindung der Oligonukleotide an die loop-Bereiche in einer DNA-Sekundärstruktur. Diese Sekundärstruktur kann mit Hilfe von Helferoligonukleotiden aufgebrochen werden. Sie werden so positioniert, daß sie die Sekundärstruktur in dem Bereich, indem das Oligonukleotid binden soll, aufbrechen. Untersuchungen hierfür wurden von <sup>[141, 142]</sup>. Peplies *et al.* <sup>[143]</sup> haben darüber hinaus die Einflüsse der Sekundärstruktur und die Orientierung der Oligonukleotide auf der Slideoberfläche untersucht. Aber es könnte noch weitere Größen geben, die noch nicht erkannt und somit noch nicht erforscht werden konnten. Optimiert wurden in dieser Arbeit die Hybridisierungstemperatur, die Formamidkonzentration im Hybridisierungspuffer, die Salzkonzentration im Waschpuffer, die Spottingkonzentration und die Dauer der Hybridisierung. Darüberhinaus wurden die Targetkomplexität und der Einfluß von Basenfehlpaarungen untersucht. Selbstverständlich ist hier ebenfalls die Kombination der einzelnen Parameter verantwortlich für das Ergebnis. Die Hybridisierungstemperatur ist abhängig von der Länge und der Schmelztemperatur der Oligonukleotide. Der Einsatz von Formamid spielt dabei ebenfalls eine Rolle. Formamid, in wässrigen Lösungen, verringert die Stabilität von DNA durch Hemmung der Ausbildung von Wasserstoffbrückenbindungen, schwächt damit die Helixstabilität und führt somit zu einer Abnahme der Schmelztemperatur <sup>[133]</sup>, sowie der Stabilität von Sekundärstrukturen. Es kann aber auch die Diskriminierung von Perfect Match und Mismatch Proben erleichtern <sup>[144]</sup>. So konnte zum Einen die Hybridisierungstemperatur auf Raumtemperatur (22 °C) erniedrigt werden (Kapitel 4.1.1.1). Die errechnete Schmelztemperatur beträgt 60 °C. Zum anderen konnten die starken Sekundärstrukturen der verwendete 5' UTR des HCV aufgebrochen werden. Begründet ist die starke interne Faltung durch die Lage der IRES (internal ribosomal entry site) in dieser Region <sup>[145]</sup>. Diese Sekundärstrukturen



beeinflussen die Hybridisierungseffizienz und sollen daher durch den Einsatz von Formamid aufgebrochen werden, um die Zugänglichkeit des Targetmoleküls für die immobilisierten Oligonukleotide zu erhöhen. Die optimale Formamidkonzentration wurde mit 50 % ermittelt (Kapitel 4.1.1.2).

Die gespottete Oligonukleotidkonzentration (Abbildung 4-4) zeigt bei 10  $\mu\text{M}$  ein Optimum. Nicht abgebildet ist die Signalintensität bei einer Konzentration von 100  $\mu\text{M}$ , da es hier zu sogenannten „Schweif“ kommt. Diese entstehen vermutlich beim Abwaschen der nicht gebundenen Oligonukleotide, die dann in Kontakt mit den noch reaktiven Phenylendiisothiocyanat-Gruppen kommen und dort gebunden werden. Somit ist eine genaue Lokalisierung der Spots und somit deren Analyse nicht mehr möglich. Darüber hinaus kann es zu Interaktionen von benachbarten Spots kommen, die die Ergebnisse verfälschen könnten.

In Abbildung 4-6 ist die Signalintensität von drei Oligonukleotiden der manuellen Bibliothek bei einer Hybridisierungsdauer von 30 min, 60 min, 90 min und 120 min dargestellt. Entgegen den Erwartungen ist ein klares Optimum bei 60 min bei dem Oligonukleotid M\_1b2E zu sehen und eine anschließende Abnahme der Signalintensität mit zunehmender Hybridisierungsdauer. Nach den Annahmen für ein Modell zur Kinetik ist die Hybridisierung eine Gleichgewichtsreaktion, in der sich nach einer bestimmten Zeit ein Equilibrium einstellt. Allerdings liegen hierbei die Angaben in der Literatur bei mehreren Stunden, bis sich das Gleichgewicht einstellt <sup>[146]</sup>. Die dargestellten Ergebnisse zeigen jedoch bei den Oligonukleotiden M\_3u-3 und M\_3u5B keine Zunahme und nur geringe Schwankungen. Ku *et al.* <sup>[147]</sup> beschreibt, dass u.a. Formamid die Kinetik einer Mikroarrayhybridisierung beeinflussen könnte. Dies könnte daraufhin deuten, dass die Hybridisierung bereits nach 30 min, und somit vor Beginn der Messung, ihr Gleichgewicht erreicht hat. Dagegen hat das Oligonukleotid M\_1b-2E ein Maximum nach 60 min und fällt anschließend wieder ab. Da die Werte normalisiert und reproduzierbar sind, wurde dieses Phänomen berücksichtigt und alle folgenden Hybridisierungen für 60 min durchgeführt.

In den Datenbanken ist meist nur die cDNA der Sequenzen hinterlegt. Beim Design der Oligonukleotide wurde daher nur die sense Orientierung der HCV-Sequenz beachtet. Zwar wurde eine Homologieüberprüfung durchgeführt, jedoch galt es zudem zu testen, ob es bei Hybridisierungen mit doppelsträngigen PCR Produkten nicht zu Kreuzhybridisierungen kommt. Wie in Abbildung 4-7 exemplarisch zu sehen ist, deutet nichts auf eine derartige Kreuzhybridisierung hin. Bei einer Kreuzhybridisierung würde ein anderes

Hybridisierungsmuster erwartet werden. Die Signalintensitäten für die Einzelstrang-Hybridisierung ist in den meisten Fällen sogar etwas schwächer als die Hybridisierung mit dem Doppelstrang.

In Abbildung 4-8 ist eine Hybridisierung gezeigt, die mit den ermittelten Parametern eine erwartete Verteilung bei der Detektion der Basenfehlpaarungen wiedergibt. Das Oligonukleotid M\_U23D mit einem mittelständigen Mismatch erreichte nur 61 % der Signalintensität des Perfect Match Oligonukleotides. Das spricht für die Qualität der Hybridisierung und der ermittelten Parameter.

Mittels der Bleaching-Experimente sollte gezeigt werden, wie empfindlich die Farbstoffe auf die Laserbestrahlung im Scanner reagieren. So wird deutlich, dass die Ermittlung der optimalen Laserpower und Verstärkung (z.B. Photomultiplier) zwar wichtig sind, dabei die Signalintensität jedoch merklich abnimmt. Nach der Lagerung über Nacht war dieses Phänomen sogar noch stärker. Es ist anzunehmen, dass dies mit der Stabilität des Fluoreszenzfarbstoffes Cy5 bei Luftfeuchtigkeit und der Wärmeentwicklung während des Scanvorgangs zusammenhängt <sup>[148]</sup>. Die hybridisierten Mikroarrays sollten also möglichst zügig und ohne weitere Lagerung gescannt werden.

Die größten Herausforderungen dieser Arbeit waren die Etablierung und Optimierung des Ultramikro-Pipettierroboter, bzw. der Piezo-Pipetten für den Roboter, sowie der Slides. Nach Durchführung geeigneter Experimente zeigte sich, dass der Ultramikro-Pipettierroboter und die Pipetten zufriedenstellend arbeiteten. Dies zeigte sich durch gute Spothomologie, wie in Abbildung 4-10 gezeigt ist. Optimierungsbedarf bestand anfangs bei den aktivierten Slideoberflächen, da durch die Art der Herstellung die reaktiven Gruppen auf der Oberfläche inhomogen verteilt waren. Dieses Problem wurde durch Optimierung im Herstellungsprozeß behoben. Bis zu diesem Zeitpunkt wurde eine alternative Lösung angestrebt. Dazu wurde eine Array-interne Normierung etabliert. Redundantes Aufbringen der Oligonukleotide ermöglichte die Minimierung von Fehlerquellen und lieferte somit standardisierte Ergebnisse. Mögliche Fehlerquellen, die ausgeglichen werden konnten waren fehlende Spots oder suboptimale Spotmorphologie, heterogene Hybridisierung und Verunreinigungen durch Staub oder andere fluoreszierende Partikel. Durch Aufbringen der gleichen Oligonukleotide an verschiedenen Positionen werden einige dieser Fehlerquellen minimiert. Um parallel eine Lösung für die inhomogene Anbindungseffizienz der

Slideoberfläche zu erarbeiten, wurde ein spezielles Spottingmuster entwickelt. Dazu wurden jeweils drei Oligonukleotide an benachbarten Positionen gespottet und eine Positivkontrolle komplettierte die quadratische Anordnung zu einem sog. Quadrupol (vgl. Abbildung 4-13). Dieser wurde bei der Software-generierten Bibliothek zweimal bzw. dreimal bei der manuellen Bibliothek, auf den Slide gespottet. Die Annahme dabei war, dass die Signalintensität der integrierten Positivkontrolle bei den verschiedenen Hybridisierungen annähernd konstant bleibt, sodass die Positivkontrolle als Bezugsgröße herangezogen werden kann, ähnlich der Normalisierung durch Housekeeping Genes bei der Genexpressions-Analyse <sup>[49]</sup> oder der RT-PCR <sup>[149]</sup>. Für die Positivkontrolle wurde das universelle Oligonukleotid S\_U23G ausgewählt, welches den Primer 88 (vgl. Tabelle 2) erkennt, und somit das HCV-Target, dass in der PCR (vgl. Kapitel 3.4.5) amplifiziert wurde, detektiert. Mit Hilfe dieser Intra-Array-Normierung war es möglich, sowohl innerhalb der einzelnen Arrays (Intra-Array), der Arrays auf einem Slide (Intra-Slide) und der Slides untereinander (Inter-Slide) zu vergleichen. Das ist von essentieller Bedeutung für den Vergleich von Hybridisierungsmustern. In Kapitel 4.1.4 ist deutlich gezeigt, dass Vergleiche mit dieser Art der Normalisierung möglich sind.

## 5.2 Vergleich und Charakterisierung der Oligonukleotid-Bibliotheken

Als Zielorganismus für die Detektion durch Oligonukleotide wurde das Hepatitis C-Virus ausgewählt. Die enorme Variabilität des Virus wird durch die stetig wachsende Zahl an Einträgen in die NCBI Datenbank deutlich. 2001 gab es ca. 11000 HCV Sequenzeinträge, im Mai 2004 hatte sich die Zahl auf über 25000 Sequenzeinträge verdoppelt. Aufgrund dieser Variabilität wurde das Virus in zu unterscheidende Geno- und Subtypen unterteilt <sup>[85, 86]</sup>. Die Klassifizierung stellte eine große Herausforderung bei dem Design der Oligonukleotiden und der Auswertung der Hybridisierungsereignisse dar. Basierend auf vergleichenden Sequenzuntersuchungen von Stuyver *et al.* <sup>[122, 150]</sup> und anhand eigener Sequenzvergleiche, wurden Bereiche innerhalb der 5'UTR ausgewählt, die den einzelnen Geno- bzw. Subtypen zugeordnet werden konnten. Die Grundlage für diese Einteilung sind zum Teil einzelne Punktmutationen. Die so ausgewählten Oligonukleotide wurden in der manuellen Bibliothek zusammengefasst (vgl. Abbildung 5-1). Sie beinhaltet 54 Sequenzen, von denen 22 zur generellen Detektion von HCV (universelle Oligonukleotide) und 32 spezifische Oligonukleotide für ausgewählte Geno- und Subtypen dienen.

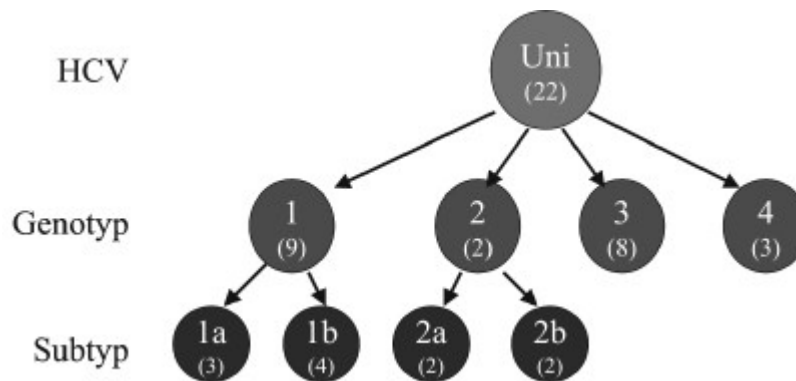


Abbildung 5-1: Hierarchische Übersichtsskizze der manuellen Oligonukleotid-Bibliothek. In den Klammern ist die Anzahl der jeweiligen Oligonukleotide angegeben.

Zunächst wurden nur die 46 Oligonukleotide untersucht, die sich aus den universellen und spezifischen für Genotyp 1 und dessen Subtypen, sowie Genotyp 3 und 4 zusammensetzten, da zunächst nur vier Proben zur Verfügung standen (1a, 1b, 3 und 4). Bei der Verifizierung der erhaltenen Ergebnisse wurden dann die Oligonukleotide für den Genotyp 2 sowie die entsprechenden Subtypen in die Experimente eingebunden.

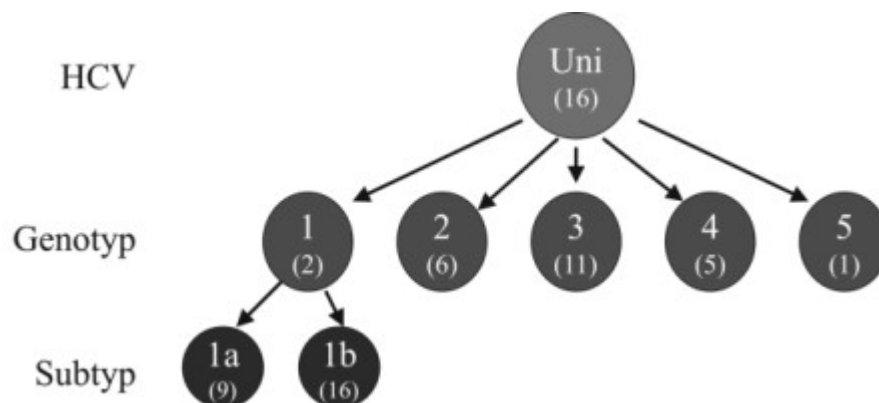


Abbildung 5-2: Hierarchische Übersichtsskizze der Software-generierten Oligonukleotid-Bibliothek. In den Klammern ist die Anzahl der jeweiligen Oligonukleotide angegeben.

Die Software-generierte Bibliothek basiert auf den gleichen Sequenzgrundlagen. Die Oligonukleotide wurden anhand von Sensitivitäts- und Spezifitätsparametern ausgewählt. Diese Bibliothek umfasst insgesamt 66 Oligonukleotide. In Abbildung 5-2 ist eine hierarchische Übersicht der verschiedenen Oligonukleotide und ihre Zuordnung dargestellt. In dieser Abbildung sind die universellen Oligonukleotide ihrem Geno- und Subtyp zugeordnet. Wie in Kapitel 4.3.2 erläutert, sind insgesamt sechs Oligonukleotide, die

spezifisch für einen Genotyp sind, den universellen zugeordnet worden. Damit wurden gleichzeitig die Anforderungen alle Genotypen zu detektieren und eine Klassifizierung in die Geno- bzw. Subtypen zu gewährleisten, erfüllt.

Obwohl für das Design der Oligonukleotide die gleichen Design-Parameter wie Länge und Schmelztemperatur verwendet wurden, hat sich gezeigt, dass es kein optimales Hybridisierungsprotokoll für alle Oligonukleotide gibt. Die Hybridisierungseffizienzen der einzelnen Oligonukleotide konnten mit den hier untersuchten Parametern nicht angeglichen werden und lieferten somit unterschiedliche Signalintensitäten, abhängig von den Charakteristika der einzelnen Oligonukleotide. Es wurde bei der Auswahl der Oligonukleotide darauf geachtet, bekannte Einflußgrößen wie  $\Delta G$ -Wert,  $T_M$ -Wert und Länge der Oligonukleotide möglichst uniform zu halten. Einige Parameter, vor allem die Länge des Oligonuklotides wurden jedoch oft aufgrund der Eigenschaften der Ziel-Sequenz relaxiert, da es anderenfalls zu Kreuzreaktivität mit Nicht-Ziel-Sequenzen geführt hätte. Natürlich sollte dies im Hinblick auf die ohnehin schon hohe Varianz in der 5' UTR des viralen Genoms vermieden werden.

Die experimentellen Untersuchungen zeigen die Varianz in der Hybridisierungseffizienz der einzelnen Oligonukleotide. Am Beispiel der Abbildung 4-22 wird die unterschiedliche Hybridisierungseffizienz der einzelnen Oligonukleotide deutlich. Dargestellt sind dort die Signalintensitäten für 16 unterschiedliche Oligonukleotide der manuellen Bibliothek. Von diesen weisen fünf Oligonukleotide keine Basenfehlpaarung zur Zielsequenz auf und zehn Oligonukleotide hatten eine und zwei Oligonukleotide zwei Mismatches. Unterteilt sind die 16 Oligonukleotide in vier Gruppen, den Masteroligonukleotiden entsprechend, und auf drei Regionen innerhalb der Zielsequenz. Aufgrund der Einteilung dieser Gruppen liegen nur drei Regionen für diese vier Gruppen vor. Zwei verschiedene Gruppen umfassen zwei verschiedene Regionen für die Erkennung des Genotyps 1, während die anderen beiden Gruppen in derselben Region liegen und die Subtypen unterscheiden sollen. Betrachtet man ausschließlich die Signalintensitäten der Perfect Match Oligonukleotide für die verschiedenen Gruppen, so fällt auf, dass eine Gruppe (M\_1u1B) deutlich geringere Signalintensitäten aufweist als die anderen. Jedoch ist auch hier die Lage des Oligonukleotides zum 5' Ende der Zielsequenz orientiert. Die Länge der Oligonukleotide beträgt jeweils 15 bis 18 nt. In der Literatur wird von Niemeyer *et al.* <sup>[151]</sup> ein Zusammenhang mit der Sekundärstruktur aufgezeigt. Eine Region mit vorrangig Stamm-

Strukturen ist für Sekundärstrukturen nicht so zugänglich wie loop-Strukturen. Bringt man die Lage der Oligonukleotide in Zusammenhang mit der Sekundärstruktur der Zielsequenz (vgl. Abbildung 8-2), wird deutlich, dass dies hier nicht der alleinige Grund sein kann. Für alle Oligonukleotide gibt es auch für die verschiedenen Variationen in den Sekundärstrukturen Stamm-Bereiche mit gleicher Anzahl an Nukleotiden. Untersuchungen von Peplies *et al.* <sup>[143]</sup> zeigen, dass die Orientierung der immobilisierten Probe eine Rolle bei der Hybridisierungseffizienz spielt. Ist der Fänger so orientiert, dass der Hauptteil der Targetsequenz in die Lösung ragt und somit minimale sterische Hinderung bei der Hybridisierungsreaktion auftreten kann, ist die Hybridisierungseffizienz erhöht. Der umgekehrte Fall trifft in diesem Beispiel bei den Oligonukleotiden mit geringer Signalintensität zu. Diese liegen so, dass der Hauptteil der Sequenz bei der Hybridisierung Richtung Slideoberfläche zeigt. Dies könnte also die geringere Hybridisierungseffizienz erklären. Auch bei den Fängeroligonukleotiden für Genotyp 3 (vgl. Abbildung 4-24) könnte dies der Grund sein. Denn die Oligonukleotide liegen genau in der Mitte der Fängersequenz, sodass es hier zu sterischen Effekten kommen könnte.

Auch bei den Software-generierten Oligonukleotiden waren die Hybridisierungseffizienzen sehr heterogen. Beispielhaft kann dies an den dargestellten Signalintensitäten der subtypspezifischen Oligonukleotide in Abbildung 4-27 links erläutert werden. Alle diese Fänger haben keine Basenfehlpaarung zu der Targetsequenz. Sie liegen jedoch an unterschiedlichen Positionen. Dies lässt vermuten, je näher sie am 3' Ende des Zielmoleküls liegen, desto höher ist ihre Signalintensität. Hier ist also eine Korrelation zu den Oligonukleotiden der manuellen Bibliothek zu erkennen. Vergleicht man dazu die universellen Oligonukleotide der manuellen Bibliothek, die am 3' Ende liegen, so haben diese, sofern sie keine Mismatches zum Zielmolekül aufweisen, die höchsten Signalintensitäten. Deutlich zu sehen ist das in den Abbildung 4-25 und Abbildung 4-30. Diese Diagramme zeigen Vergleiche ausgewählter Oligonukleotide, die die bestmögliche Differenzierung zwischen den Geno- und Subtypen erlauben. Auf Grund ihrer unterschiedlichen Positionen und somit unterschiedlichen Hybridisierungseffizienzen wird die Auswertung anhand von Hybridisierungsmustern durchgeführt und kann nicht ausschließlich über ein positives Hybridisierungssignal ausgewertet werden. In Abbildung 4-25 sind die Hybridisierungen der manuellen Bibliothek zu sehen. Hier kann zwischen den Genotypen, nicht aber zwischen den Subtypen 1a und 1b unterschieden werden. Im Gegensatz dazu ist dies mit der Software-generierten Bibliothek möglich.

Wie in Abbildung 4-35 mit Oligonukleotiden der manuellen Bibliothek zu erkennen ist, ist das detektierbare Hybridisierungsereignis der Spots mit den Oligonukleotiden an den Positionen auf dem Mikroarray zwar erwartet, konnte jedoch nicht validiert werden. Dem gegenüber entstehen Signalmuster, die für bestimmte Geno- bzw. Subtypen charakteristisch sind. Je weiter die Geno- bzw. Subtypen phylogenetisch voneinander entfernt sind bzw. die erkannten Sequenzregionen einzigartig sind und sich nicht durch einzelne Mismatches voneinander unterscheiden, desto klarer ist die Abgrenzung gegenüber den anderen Varianten und desto klarer ist daher die Zuordnung zu dem jeweiligen Genotyp. Sehr deutlich wird dies am Beispiel Genotyp 2 im Vergleich mit Genotyp 1 und Genotyp 4. In Abbildung 5-3 wird deutlich, dass Genotyp 1 und 4 relativ eng miteinander verwandt sind. Dagegen hebt sich Genotyp 2 deutlich von Genotyp 1 und 4 ab. Die Abbildung 5-3 zeigt die Phylogenie anhand der kompletten Sequenz. Cristina *et al.* <sup>[152]</sup> zeigen ein übereinstimmendes Ergebnis nach Analyse von Sequenzen aus der 5'UT-Region.

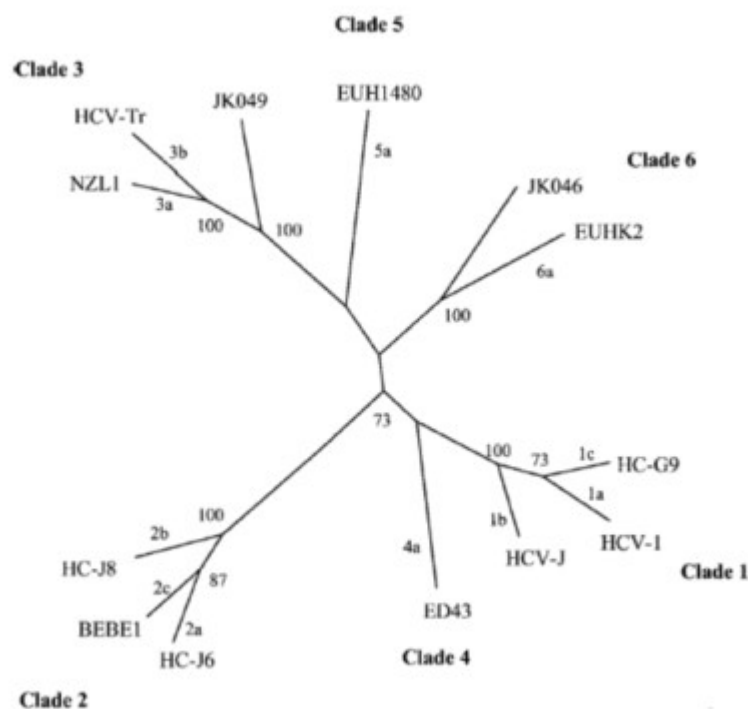


Abbildung 5-3: Parsimony Analyse repräsentativer HCV Komplet-Sequenzen aus Robertson *et al.* <sup>[86]</sup>.

Die klinisch wichtige Diskriminierung von Subtyp 1a und 1b von allen anderen Geno- und Subtypen kann mit der Software-generierten Bibliothek eindeutig durchgeführt werden. Dies liegt vor allem an der Tatsache, dass für Subtyp 1b eine viel größere Anzahl von

Sequenzen in den öffentlich zugänglichen Datenbanken enthalten ist und daher eine größere Sicherheit bei der Auswahl der Oligonukleotide besteht. Ein weiterer Aspekt ist, dass für Subtyp 1b die größte Anzahl an Oligonukleotiden vorhanden ist, sowohl bei der manuellen als auch bei der software-generierten Bibliothek. Diese konnten aufgrund der größeren Sequenzinformation in den Datenbanken ausgewählt werden. Zwar treten bei der Software-generierten Bibliothek auch die positionsbedingten Schwierigkeiten auf, z. B. ein stärkeres Signal eines Oligonukleotides, das eine Basenfehlpaarung aufweist im Vergleich zu einem Perfect Match Oligonukleotid an anderer Position, dies kann jedoch durch die wissensbasierte Auswertung kompensiert werden. Darüber hinaus helfen Bayessche Netze trotz der Unsicherheiten, die durch leicht schwankende Hybridisierungssignale auftreten können, eine klare Aussage über die Wahrscheinlichkeit des Genotyps einer unbekannt Probe zu treffen <sup>[63, 153]</sup>. Bei der Auswertung ist bemerkenswert, dass für keine der hybridisierten HCV-Proben die entsprechenden Sequenzen in das Design mit einbezogen wurden, da sie nicht in den verwendeten Datenbanken hinterlegt waren.

Die Ergebnisse der Multitarget-Hybridisierungen zeigen nicht die simple Addition der Hybridisierungsmuster der eingesetzten Genotypen. Untersucht wurde dieses Verhalten, da in der Literatur ein natürlich vorkommendes rekombinantes Virus beschrieben wurde <sup>[154]</sup>. Allerdings wird dieses Phänomen kontrovers diskutiert, da diese gemischten Infektionen, die zur Entstehung eines intragenotypischen Virus führen, wahrscheinlich ein seltenes Ereignis darstellen <sup>[155]</sup>.

### 5.3 Design

Als Grundlage für das Design der Oligonukleotide dienten Sequenzen aus der NCBI und DDBJ. 1998 beinhalteten beide Datenbanken 944 Einträge für Kompletengenome von HCV. Zu diesem Zeitpunkt wurden diese Sequenzen extrahiert und dienten als Basis für die Auswahl der Oligonukleotide in den in dieser Arbeit charakterisierten Bibliotheken. Im März 2004 waren es 25125 Nukleotid-Einträge. Die meisten davon (1673) sind dem Subtyp 1b zugeordnet. Daran ist das Interesse an diesem Subtyp zu erkennen.

Für die Berechnung der Schmelztemperatur wurde in beiden Fällen das nearest-neighbor Modell zugrunde gelegt, bei dem angenommen wird, dass die Stabilität eines Basenpaares von der Identität und der Orientierung der benachbarten Basenpaare abhängt <sup>[51, 52]</sup>. Die



Grundlage für dieses Modell ist die Annahme der Änderung der freien Energie bei 37°C ( $\Delta G_{37}$ ) bei der Bildung der Doppelhelix-Struktur. Diese Werte sind (1) die Änderung der freien Energie bei der Ausbildung der Doppelhelix, (2) die Änderung der freien Energie bei Initiierung einer Helix-Struktur und (3) die Änderung der freien Energie als Entropie-Effekt, wenn der Duplex aus selbstkomplementären Strängen aufgebaut wird. Im Internet sind verschiedene Programme frei verfügbar, mit denen die Schmelztemperatur ermittelt werden kann. Bei der Kalkulation der Schmelztemperaturen, sollte darauf geachtet werden, dass immer die gleiche Formel verwendet wird, da es sonst zu enormen Schwankungen von  $\pm 17$  °C für dieselbe Sequenz bei gleichen Parametern kommen kann. Es ist daher wichtig, dass die Schmelztemperaturen aller Oligonukleotide mit derselben Methode ermittelt werden. Die Hybridisierungsbedingungen sollten aber immer optimiert werden.

Obwohl die Oligonukleotide anhand gleicher Parameter wie Schmelztemperatur, GC-Gehalt und minimale Sekundärstruktur ( $\Delta G$ -Werte) ausgewählt wurden, konnten unterschiedliche Hybridisierungseffizienzen nach der Hybridisierungsreaktion gemessen werden. Beispielhaft soll dies an den Oligonukleotiden M\_1u1B, M\_1u2A und M\_1u2D aus Abbildung 4-23 erläutert werden. M\_1u1B liegt am 5' Ende des Targets und hat keine Basenfehlpaarung, M\_1u2A liegt hingegen am 3' Ende des Targets, hat ebenfalls keine Basenfehlpaarung (PM) und M\_1u2D liegt an ähnlicher Position wie M\_1u2A, hat aber eine Basenfehlpaarung. Nach der Hybridisierung zeigte sich folgende Verteilung der Signalintensität: Die höchste hatte M\_1u2A (PM), gefolgt von M\_1u2D (ein MM) und M\_1u1B (PM). Somit konnte zwar M\_1u2D deutlich von M\_1u2A unterschieden werden, die Signalintensität dieses Mismatch Oligonukleotides lag aber immer noch deutlich höher als das Signal von M\_1u1B ohne Basenfehlpaarung. Bei unbekanntem Proben könnte dies ein Problem bei der Zuordnung zum Genotyp ergeben. Aus diesem Grund war die Zuordnung zu den Genotypen erschwert, da auch falsch-positive Signale detektiert werden konnten. Die zur Auswahl der Oligonukleotide betrachteten Parameter scheinen also nicht ausreichend zu sein, um einheitliche(re) Signalintensitäten zu erzielen. Peplies <sup>[143]</sup> hat die bereits etablierte in-situ Hybridisierungssonden für die 16S rRNA auf Mikroarrays übertragen und konnte mit Hilfe verschiedener Modifikationen, z.B. Einsatz von Spacern oder Inversion der Position auf dem Mikroarray, die Signalintensitäten angleichen. In diesem Fall müsste experimentell überprüft werden, ob das zu einer Erhöhung der Signalintensität bei M\_1u1B führen würde. Ein anderes Phänomen ist eine höhere Signalintensität des Oligonukleotides mit einem Mismatch im Vergleich zum

Oligonukleotid ohne Basenfehlpaarung. Beispielhaft könnte hier aus Abbildung 4-22 die Oligonukleotide M\_1a1A und M\_1a1C genannt werden. Das Oligonukleotid mit einer Basenfehlpaarung M\_1a1C hat eine deutlich höhere Signalintensität als das Oligonukleotid M\_1a1A ohne Basenfehlpaarung. Eine Erklärung hierfür könnten Gutell *et al.* <sup>[156]</sup> liefern. Nicht alle Basenpaarungen haben dieselbe Energie, d.h. manche Basenpaarungen sind schwächer als andere. Das gilt zum einen für G-C Bindung mit drei Wasserstoffbrücken im Vergleich zu zwei bei A-T Bindungen, aber auch für nichtkanonische Wechselwirkungen.

Ein ganz ähnliches Vorgehen zum Design wie bei der Software-generierte Oligonukleotid-Bibliothek stellen Wang *et al.* vor <sup>[157]</sup>. Ihr Ziel war eine Mikroarray-basierte Erkennung und Genotypisierung viraler Pathogene. Sie verwendeten als Datengrundlage GenBank. Die dort hinterlegten Sequenzdaten werden für die verschiedenen Virusfamilien geclustert und anschließend in 70 nt lange Segmente aufgeteilt, die jeweils mit 25 nt überlappen. Zwischen jedem 70mer und jedem Virusgenom der entsprechenden Familie wurde ein paarweises Alignment durchgeführt. Die 70mere wurden dann in der Reihenfolge ihrer Treffer aufgelistet und anschließend die fünf Oligonukleotide mit dem höchsten score-Wert und ihr revers-komplementärer Strang ausgewählt. Mit diesem Vorgehen wurden auch fünf Oligonukleotide für HCV ausgewählt von denen eines zwischen der 5'UTR und der core-Region liegt (auf D90208 an nt 331), eines in der core-Region (nt 813) und drei in der NS5B Region (nt 7978, nt 8865 und nt 8971). In der Publikation von Wang *et al.* wurden diese Oligonukleotide jedoch nicht experimentell überprüft.

Die Eignung der Mikroarrays zur Identifizierung und Klassifizierung von Viren wurde bereits für andere Viren gezeigt. Zur molekularen Bestimmung und Detektion von Influenzaviren verwendeten Sengupta *et al.* <sup>[158]</sup> 463 Oligonukleotide zwischen 17 und 29 nt Länge, die aus publizierten Oligonukleotiden ausgesucht wurden. Die Signalintensitäten der Oligonukleotide variierte stark und so mußte jedes individuell betrachtet werden, was Erfahrung mit den Hybridisierungseigenschaften des Oligonukleotides voraussetzt. Auch für die Detektion und Diskriminierung von Orthopoxviren wurden Mikroarrays eingesetzt <sup>[159]</sup>. Die hier eingesetzten Oligonukleotide waren zwischen 13 und 21 nt lang, und zur Identifikation für jede Virusspezies wurden mehrere Oligonukleotide verwendet. Zur Differenzierung der einzelnen Spezies wurde ein Gen verwendet und die entsprechenden Sequenzen alignet, um geeignete Regionen für das Design der Oligonukleotide zu identifizieren. Zusätzlich wurde ein eigens entwickeltes Programm "Oligoscan" verwendet, um entsprechende Oligonukleotide auszuwählen. Die

Parameter für das Design waren weniger detailliert als in der hier vorliegenden Arbeit. So wurden Oligonukleotide ausgewählt, die mindestens ein Mismatch zu den anderen Orthopoxvirusspezies hatte, die Länge der Oligonukleotide wurde auf 13 bis 21 nt beschränkt und die Schmelztemperatur sollte zwischen 36 °C und 58 °C liegen. Mit den ausgewählten Oligonukleotide konnte die Spezies jeweils korrekt zugeordnet werden.

Für die Identifizierung von bestimmten Mutationen, die bei der Therapie von HCV mit Interferon einen Einfluß haben, haben Takahashi *et al.* erst kürzlich einen elektrochemischen DNA-Array entwickelt <sup>[160]</sup>. Die Effektivität der Interferon Therapie ist abhängig vom Genotyp und der Viruslast, sowie individueller Charakteristika der Patienten. Zwei Mutationen an bestimmten Nukleotidpositionen des Gens MxA scheinen für die Antwort auf die Therapie verantwortlich zu sein. Für diesen und zwei weitere SNPs wurden jeweils Oligonukleotide mit mittelständigen Mismatches auf den elektrochemischen Array immobilisiert. Die Hybridisierung mit entsprechenden DNA-Fragmenten wurde durchgeführt und der Array anschließend mit einem elektrochemisch-aktiven Minor-Groove Binder, der spezifisch doppelsträngige DNA bindet reaktiviert. Die Auswertung erfolgte über einen elektrochemischen Analysator. Es konnte gezeigt werden, daß mit dieser Methode die korrekte Diskriminierung der SNP möglich war.

## 6 ZUSAMMENFASSUNG UND AUSBLICK

In der vorgestellten Arbeit wird der Einsatz von DNA-Mikroarrays zur simultanen Erkennung und Klassifizierung des Hepatitis C-Virus (HCV) untersucht. Es konnte gezeigt werden, dass es mit dieser Technologie möglich ist, die Varianten des HCV anhand der Sequenzunterschiede in der nichttranslatierten Region am 5'-Ende (5'UTR) zu identifizieren und gleichzeitig zu genotypisieren. Auf Basis dieser methodischen Untersuchungen wurde zur Optimierung der Konfiguration ein Vergleich zweier Oligonukleotid-Bibliotheken durchgeführt, die entweder mit der Software „*OptiNa*“ generiert oder manuell ausgewählt wurden.

Die Etablierung der Mikroarray-Technologie umfasste Experimente zur Homogenität der Slides, zur Homogenität der Tropfenabgabe des Roboters sowie zur optimalen Spottingkonzentration der Oligonukleotid- und der hybridisierten Targetlösungen. Dabei zeigte sich, dass die Inhomogenität der Oberflächen zu einer hohen Standardabweichung bei der Auswertung führt. Als Lösungsansatz wurde ein definiertes Spottingmuster und damit verbunden eine Intra-Array Normierung etabliert. Die optimale Oligonukleotidkonzentration lag bei 10  $\mu\text{M}$ . Die optimale Konzentration der zu hybridisierenden Targetlösung betrug 10 nM. Die grundlegende Optimierung der Hybridisierungsbedingungen ergab eine Formamidkonzentration von 50 % im Hybridisierungspuffer und eine Hybridisierungstemperatur von 22 °C bei doppelsträngigen Targetmolekülen. Bei den untersuchten Oligonukleotiden konnte unter diesen Bedingungen ein Oligonukleotid mit mindestens einer Basenfehlpaarung zum Target sicher von dem Oligonukleotid ohne Basenfehlpaarung unterschieden werden. Sie unterscheiden sich durch eine mindestens 60 %ige Abnahme der Signalintensität bei dem Oligonukleotid mit Basenfehlpaarung.

Für die manuelle Bibliothek wurden Oligonukleotide innerhalb bestimmter Regionen in der 5'UTR ausgewählt <sup>[150]</sup>. Die ausgewählten Oligonukleotide, die die meisten Sequenzen für einen Genotypen abdeckten, wurden als Masteroligonukleotide bezeichnet. Die HCV-Sequenzvarianten wurden durch zusätzliche Oligonukleotide, sogenannte „Varianten“, abgedeckt. Durch einen Homologievergleich mit anderen Sequenzen in den Datenbanken konstatierte die notwendige Spezifität der Oligonukleotide. Diese Bibliothek umfasst insgesamt 54 Oligonukleotide. Bei der Software-generierten Bibliothek wurden die Regionen innerhalb der 5' UTR nicht berücksichtigt. Hier wurden anhand des Alignments die Oligonukleotide aufgrund von physikalischen Parametern, wie z.B. Länge der

Oligonukleotide sowie Schmelztemperatur und  $\Delta G$ -Wert, ausgewählt. Ein Homologievergleich zu Sequenzen in den NCBI-Datenbanken wurde durchgeführt. Auf diese Weise wurden 66 Oligonukleotide ermittelt. Die Oligonukleotide beider Bibliotheken wurden *in-silico* miteinander verglichen. Die Lage aller Oligonukleotide wurde auf allen 36 untersuchten und sequenzierten Hepatitis C-Virus Zielmolekülen und einer Referenz (HCV-1, Accessionnummer M62321) ermittelt, sowie ggf. die Anzahl und Lage der Basenfehlpaarungen (vgl. Tabelle 8 und Tabelle 9). Für den experimentellen Vergleich wurden HCV-PCR-Produkte in einer nested PCR mit Cy5 am 5'-Ende markiert, aufgereinigt und sowohl in Hybridisierungsansätzen mit einem Target als auch in einer Multitarget Hybridisierung mit jeweils zwei unterschiedlichen Geno- bzw. Subtypen gegen die Software-generierte Bibliothek hybridisiert. Die Auswertung erfolgte durch charakteristische Hybridisierungsmuster der jeweiligen Geno- bzw. Subtypen. Bei der Auswertung dieser Muster konnten mittels der Oligonukleotide aus der manuellen Bibliothek die Genotypen 1, 2 und 3 korrekt identifiziert werden, ebenso die Subtypen 2a, 2b, 3a und 3b. Eine Unterteilung der klinisch relevanten Subtypen 1a und 1b war nicht möglich. Eine Validierung der Hybridisierungsmuster mit weiteren HCV-Proben zeigte jedoch, dass bei unterschiedlichen HCV-Proben desselben Typs die Muster nicht übereinstimmen und somit eine Typisierung für einen Außenstehenden schwierig sein würde. Dies war bei der Software-generierten Bibliothek anders. Die Oligonukleotide wurden hier so ausgewählt, dass eine Unterscheidung der Subtypen 1a und 1b möglich ist. Allerdings konnte eine Unterscheidung der Subtypen 2a, 2b, 3a und 3b nicht gezeigt werden, da keine spezifischen Oligonukleotide generiert wurden. Jedoch ermöglicht die eigens entwickelte Auswertesoftware *iOmega* der Firma *iSenseIt* GmbH, Bremen, die zuverlässige Zuordnung zu den korrekten HCV-Typen in Prozent. Mit den Software-generierten Oligonukleotiden konnte eine korrekte Zuordnung zu den HCV-Typen nach Hybridisierungen zweier equimolarer Geno- bzw. Subtypen mit der *iOmega*-Software nicht erreicht werden, jedoch anhand der Hybridisierungsmuster.

Es konnte gezeigt werden, dass die Software zur Auswahl von Oligonukleotiden für die Erkennung und Klassifizierung der Genotypen 1a, 1b, 2, 3 und 4 von Hepatitis-C Viren geeignet ist. Für umfassendere, klinische Studien wäre es wünschenswert, ergänzende Oligonukleotide für weitere Geno- und Subtypen auszuwählen. Die für die Therapie relevante Klassifizierung, die Erkennung von HCV-Genotyp 1 und die Klassifizierung in die Subtypen 1a und 1b wurde mit dieser Version des Hepatitis C-Virus Mikroarrays jedoch erreicht werden.

Mit Hilfe der vorgestellten Oligonukleotid-Bibliotheken, insbesondere der Software-generierten, ist es möglich Hepatitis C-Viren sicher zu identifizieren und zu genotypisieren. Eine Implementierung zusätzlicher Oligonukleotide für eine weitere Genotypisierung der anderen Subtypen wäre von wissenschaftlichem Interesse. Zusammen mit der Auswertesoftware hat man ein einfaches und intuitives Werkzeug zur Hand. Für den medizinischen Aspekt reicht die vorhandene Klassifizierung in die Genotypen 1-5 und Subtypen 1a und 1b aus. Der noch zu optimierende Aspekt dieser Methode ist die vorgeschaltete PCR, denn mit den verwendeten Primern wird bereits eine Selektion durchgeführt. Eventuell wäre hier die Etablierung einer Multiplex-PCR von Vorteil. Darüberhinaus könnte eventuell mit dem Einsatz von Spacern oder durch die Inversion des Oligonukleotides auf dem Slide eine gleichförmigere Hybridisierungseffizienz erreicht werden.

Die bisherigen Heilungschancen für Menschen mit chronischer Hepatitis sind je nach Genotyp unterschiedlich. Ist man mit Genotyp 1b infiziert, der wie in Abbildung 1-8 dargestellt, vor allem in Europa und Asien verbreitet ist, hat der Erkrankte mit den bisherigen Therapeutika kaum Aussicht auf Heilung. Neue Hoffnung für die Betroffenen lässt eine ganz neue Technik zu: das Gene-Silencing. Wenn auf natürlichem Weg doppelsträngige RNA in die Zelle gelangt, wird sie von der RNase III (DICER) in 21 nt kurze Stücke gespalten, die als small interfering RNAs oder siRNAs bezeichnet werden. Die siRNA kann auch synthetisch hergestellt und in die Zelle eingebracht werden. Sie wird dort von einem Proteinkomplex, dem sogenannten RISC (RNA Induced Silencing Complex) gebunden. Ein Strang des siRNA Duplexes wird abgestreift, der komplementäre Strang bleibt im RISC zurück. Der Komplex erkennt und bindet nun an die target RNA, den er mit dem komplementären siRNA-Strang erkennt, und schneidet ihn enzymatisch in kleine Stücke. Die virale Replikation wird so unterbrochen. Im Juni 2003 publizierten Sen *et al.* <sup>[161]</sup> und Yokota *et al.* <sup>[162]</sup> erfolgreiche Experimente mit siRNA für die NS5A bzw. die 5'UTR in der menschlichen Hepatoma Zelllinie HepG2. Eine nächste Herausforderung wird es sein, die siRNA zu den entsprechenden Zellen bzw. Geweben im Körper zu bringen. Es besteht somit Hoffnung auf Heilung für die weltweit 170 Millionen HCV-Infizierten.

## 7 LITERATURVERZEICHNIS

1. Lander, E.S., *et al.*, *Initial sequencing and analysis of the human genome*. Nature, 2001. **409**(6822): p. 860-921.
2. Southern, E., *Detection of specific sequences among DNA fragments separated by gel electrophoresis*. J Mol Biol, 1975. **98**: p. 503-517.
3. Blohm, D.H. und Guiseppi-Elie, A., *New developments in microarray technology*. Curr Opin Biotechnol, 2001. **12**(1): p. 41-47.
4. [http://www.affymetrix.com/support/technical/datasheets/human\\_datasheet.pdf](http://www.affymetrix.com/support/technical/datasheets/human_datasheet.pdf), *Datasheet GeneChip Genome U133 V2.0 Array*. 2004.
5. Heller, M.J., *DNA microarray technology: devices, systems, and applications*. Annu Rev Biomed Eng, 2002. **4**: p. 129-153.
6. Fodor, S.P., Rava, R.P., Huang, X.C., Pease, A.C., Holmes, C.P., und Adams, C.L., *Multiplexed biochemical assays with biological chips*. Nature, 1993. **364**(6437): p. 555-556.
7. [http://oligos.qiagen.com/arrays/oligosets\\_overview.php](http://oligos.qiagen.com/arrays/oligosets_overview.php), AROS. 2004, QIAGEN.
8. Debouck, C. und Goodfellow, P.N., *DNA microarrays in drug discovery and development*. Nat Genet, 1999. **21**(1 Suppl): p. 48-50.
9. Butte, A., *The use and analysis of microarray data*. Nat Rev Drug Discov, 2002. **1**(12): p. 951-960.
10. Xiang, Z., Yang, Y., Ma, X., und Ding, W., *Microarray expression profiling: analysis and applications*. Curr Opin Drug Discov Devel, 2003. **6**(3): p. 384-395.
11. Dudda-Subramanya, R., Lucchese, G., Kanduc, D., und Sinha, A.A., *Clinical applications of DNA microarray analysis*. J Exp Ther Oncol, 2003. **3**(6): p. 297-304.
12. Troesch, A., Nguyen, H., Miyada, C.G., Desvarenne, S., Gingeras, T.R., Kaplan, P.M., Cros, P., und Mabilat, C., *Mycobacterium species identification and rifampin resistance testing with high-density DNA probe arrays*. J Clin Microbiol, 1999. **37**(1): p. 49-55.
13. Nakamura, Y., *Isolation of p53-target genes and their functional analysis*. Cancer Sci, 2004. **95**(1): p. 7-11.
14. Schmidt, U. und Begley, C.G., *Cancer diagnosis and microarrays*. Int J Biochem Cell Biol, 2003. **35**(2): p. 119-124.
15. Hirschhorn, J.N., Sklar, P., Lindblad-Toh, K., Lim, Y.M., Ruiz-Gutierrez, M., Bolk, S., Langhorst, B., Schaffner, S., Winchester, E., und Lander, E.S., *SBE-TAGS: an array-based method for efficient single-nucleotide polymorphism genotyping*. Proc Natl Acad Sci U S A, 2000. **97**(22): p. 12164-12169.
16. Ronaghi, M., Uhlen, M., und Nyren, P., *A sequencing method based on real-time pyrophosphate*. Science, 1998. **281**(5375): p. 363, 365.
17. Jung, A., Stemmler, I., Brecht, A., und Gauglitz, G., *Covalent strategy for immobilization of DNA-microspots suitable for microarrays with label-free and time-resolved optical detection of hybridization*. Fresenius J Anal Chem, 2001. **371**(2): p. 128-136.
18. Hashimoto, K. und Ishimori, Y., *Preliminary evaluation of electrochemical PNA array for detection of single base mismatch mutations*. Lab Chip, 2001. **1**(1): p. 61-63.

19. Jiangbing, F.D., C.; Chanfeng, Z.; Lixin, Z., Wenyi, F, *High-density fiber optic array technology and its application in functional genomic studies*. Chinese Science Bulletin, 2003. **48**(18): p. 1903-1905.
20. Benters, R., Niemeyer, C.M., Drutschmann, D., Blohm, D., und Wohrle, D., *DNA microarrays with PAMAM dendritic linker systems*. Nucleic Acids Res, 2002. **30**(2): p. E10.
21. Le Berre, V., Trevisiol, E., Dagkessamanskaia, A., Sokol, S., Caminade, A.M., Majoral, J.P., Meunier, B., und Francois, J., *Dendrimeric coating of glass slides for sensitive DNA microarrays analysis*. Nucleic Acids Res, 2003. **31**(16): p. e88.
22. Ekins, R. und Edwards, P., *On the meaning of "sensitivity"*. Clin Chem, 1997. **43**(10): p. 1824-1831.
23. Brandt, O., Feldner, J., Stephan, A., Schroder, M., Schnolzer, M., Arlinghaus, H.F., Hoheisel, J.D., und Jacob, A., *PNA microarrays for hybridisation of unlabelled DNA samples*. Nucleic Acids Res, 2003. **31**(19): p. e119.
24. Nielsen, P.E., Egholm, M., und Buchardt, O., *Peptide nucleic acid (PNA). A DNA mimic with a peptide backbone*. Bioconjug Chem, 1994. **5**(1): p. 3-7.
25. Egholm, M., Buchardt, O., Christensen, L., Behrens, C., Freier, S.M., Driver, D.A., Berg, R.H., Kim, S.K., Norden, B., und Nielsen, P.E., *PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen-bonding rules*. Nature, 1993. **365**(6446): p. 566-568.
26. Schwarz, F.P., Robinson, S., und Butler, J.M., *Thermodynamic comparison of PNA/DNA and DNA/DNA hybridization reactions at ambient temperature*. Nucleic Acids Res, 1999. **27**(24): p. 4792-4800.
27. Lohse, J., Dahl, O., und Nielsen, P.E., *Double duplex invasion by peptide nucleic acid: a general principle for sequence-specific targeting of double-stranded DNA*. Proc Natl Acad Sci U S A, 1999. **96**(21): p. 11804-11808.
28. Kauppinen, S., Nielsen, P.T., Mouritzen, P., Nielsen, A.T., Vissing, H., Moller, S., und Ramsing, N.B., *LNA Microarrays in Genomics*. PharmaGenomics, 2003: p. p. 24-34.
29. Mouritzen, P., Nielsen, A.T., Pfundheller, H.M., Choleva, Y., Kongsbak, L., und Moller, S., *Single nucleotide polymorphism genotyping using locked nucleic acid (LNA)*. Expert Rev Mol Diagn, 2003. **3**(1): p. 27-38.
30. Igloi, G.L., *Single-nucleotide polymorphism detection using peptide nucleic acids*. Expert Rev Mol Diagn, 2003. **3**(1): p. 17-26.
31. Tolstrup, N., Nielsen, P.S., Kolberg, J.G., Frankel, A.M., Vissing, H., und Kauppinen, S., *OligoDesign: Optimal design of LNA (locked nucleic acid) oligonucleotide capture probes for gene expression profiling*. Nucleic Acids Res, 2003. **31**(13): p. 3758-3762.
32. Kononen, J., Bubendorf, L., Kallioniemi, A., Barlund, M., Schraml, P., Leighton, S., Torhorst, J., Mihatsch, M.J., Sauter, G., und Kallioniemi, O.P., *Tissue microarrays for high-throughput molecular profiling of tumor specimens*. Nat Med, 1998. **4**(7): p. 844-847.
33. Simon, R., Mirlacher, M., und Sauter, G., *Tissue microarrays*. Biotechniques, 2004. **36**(1): p. 98-105.
34. Van De Rijn, M. und Gilks, C.B., *Applications of microarrays to histopathology*. Histopathology, 2004. **44**(2): p. 97-108.
35. MacBeath, G. und Schreiber, S.L., *Printing proteins as microarrays for high-throughput function determination*. Science, 2000. **289**(5485): p. 1760-1763.



36. Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P., Lan, N., Jansen, R., Bidlingmaier, S., Houfek, T., Mitchell, T., Miller, P., Dean, R.A., Gerstein, M., und Snyder, M., *Global analysis of protein activities using proteome chips*. Science, 2001. **293**(5537): p. 2101-2105.
37. Bulyk, M.L., Gentalen, E., Lockhart, D.J., und Church, G.M., *Quantifying DNA-protein interactions by double-stranded DNA arrays*. Nat Biotechnol, 1999. **17**(6): p. 573-577.
38. Arenkov, P., Kukhtin, A., Gemmell, A., Voloshchuk, S., Chupeeva, V., und Mirzabekov, A., *Protein microchips: use for immunoassay and enzymatic reactions*. Anal Biochem, 2000. **278**(2): p. 123-131.
39. Zhu, H., Klemic, J.F., Chang, S., Bertone, P., Casamayor, A., Klemic, K.G., Smith, D., Gerstein, M., Reed, M.A., und Snyder, M., *Analysis of yeast protein kinases using protein chips*. Nat Genet, 2000. **26**(3): p. 283-289.
40. Kohn, M., Wacker, R., Peters, C., Schroder, H., Soulere, L., Breinbauer, R., Niemeyer, C.M., und Waldmann, H., *Staudinger ligation: a new immobilization strategy for the preparation of small-molecule arrays*. Angew Chem Int Ed Engl, 2003. **42**(47): p. 5830-5834.
41. Schweitzer, B., Wiltshire, S., Lambert, J., O'Malley, S., Kukanskis, K., Zhu, Z., Kingsmore, S.F., Lizardi, P.M., und Ward, D.C., *Inaugural article: immunoassays with rolling circle DNA amplification: a versatile platform for ultrasensitive antigen detection*. Proc Natl Acad Sci U S A, 2000. **97**(18): p. 10113-10119.
42. Templin, M.F., Stoll, D., Schrenk, M., Traub, P.C., Vohringer, C.F., und Joos, T.O., *Protein microarray technology*. Trends Biotechnol, 2002. **20**(4): p. 160-166.
43. Cooley, P.W., D.; Antohe, B., *Applications of Ink-Jet Printing Technology to BioMEMS and Microfluidic Systems*. 2001, MicroFab Technologies, Inc.: Plano, Texas, USA.
44. <http://www.affymetrix.com/technology/manufacturing/index.affx>.
45. Schuchhardt, J., Beule, D., Malik, A., Wolski, E., Eickhoff, H., Lehrach, H., und Herzel, H., *Normalization strategies for cDNA microarrays*. Nucleic Acids Res, 2000. **28**(10): p. E47.
46. Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P., Stoeckert, C., Aach, J., Ansorge, W., Ball, C.A., Causton, H.C., Gaasterland, T., Glenisson, P., Holstege, F.C., Kim, I.F., Markowitz, V., Matese, J.C., Parkinson, H., Robinson, A., Sarkans, U., Schulze-Kremer, S., Stewart, J., Taylor, R., Vilo, J., und Vingron, M., *Minimum information about a microarray experiment (MIAME)-toward standards for microarray data*. Nat Genet, 2001. **29**(4): p. 365-371.
47. Quackenbush, J., *Microarray data normalization and transformation*. Nat Genet, 2002. **32 Suppl**: p. 496-501.
48. Yang, Y.H., Dudoit, S., Luu, P., Lin, D.M., Peng, V., Ngai, J., und Speed, T.P., *Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation*. Nucleic Acids Res, 2002. **30**(4): p. e15.
49. Eickhoff, B., *Normalization of array hybridization experiments in differential gene expression analysis*. Nucleic Acids Res, 1999. **27**(22): p. e33.
50. Hill, A.A., Brown, E.L., Whitley, M.Z., Tucker-Kellogg, G., P., H.C., und Slonim, D., *Evaluation of normalized procedures for oligonucleotide array data based on spiked cRNA controls*. Genome Biol, 2001. **2**(12): p. 1-13.

51. Breslauer, K.J., Frank, R., Blocker, H., und Marky, L.A., *Predicting DNA duplex stability from the base sequence*. Proc Natl Acad Sci U S A, 1986. **83**(11): p. 3746-3750.
52. SantaLucia, J., Jr., *A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics*. Proc Natl Acad Sci U S A, 1998. **95**(4): p. 1460-1465.
53. SantaLucia, J., Jr., Allawi, H.T., und Seneviratne, P.A., *Improved nearest-neighbor parameters for predicting DNA duplex stability*. Biochemistry, 1996. **35**(11): p. 3555-3562.
54. <http://alces.med.umn.edu/rawtm.html>, *Virtual Genome Center*. 1995.
55. Zuker, M., *Mfold web server for nucleotic acid folding and hybridization prediction*. Nucleic Acids Res, 2003. **31**(13): p. 2406-3415.
56. <http://www.bioinfo.rpi.edu/applications/mfold/>. *Mfold*. 2004.
57. Smith, T.F. und Waterman, M.S., *Identification of common molecular subsequences*. J Mol Biol, 1981. **147**(1): p. 195-197.
58. Needleman, S.B. und Wunsch, C.D., *A general method applicable to the search for similarities in the amino acid sequence of two proteins*. J Mol Biol, 1970. **48**(3): p. 443-453.
59. Li, F. und Stormo, G.D., *Selection of optimal DNA oligos for gene expression arrays*. Bioinformatics, 2001. **17**(11): p. 1067-1076.
60. ArrayDesigner, <http://www.premierbiosoft.com/dnamicroarray/dnamicroarray.html>. 2004, Premier Biosoft.
61. OligoArray, <http://www.premierbiosoft.com/dnamicroarray/dnamicroarray.html>. 2003.
62. Rouillard, J.M., Herbert, C.J., und Zuker, M., *OligoArray: genome-scale oligonucleotide design for microarrays*. Bioinformatics, 2002. **18**(3): p. 486-487.
63. Volkmann, G., *Interpretation von DNA Microarray Daten mit Bayesschen NETzen*, in *Fachbereich 3 (Mathematik und Informatik)*. 2003, Universität BRremen: Bremen. p. 181.
64. Nölte, M.V., G.; Drutschmann, D.; Waschulzik\*, T.; Blohm, D., *Bioinformatik-System zur Optimierung von Oligonukleotid-bibliotheken für DNA-Mikroarrays*. Statusseminar Dechema, 2001.
65. Volkmann, G.W., D.; Trutnau, H.-H., *Software-Plattform für Design und Analyse von DNA-Microarrays*. Transkript, 2003. **1-2**: p. 58-59.
66. Choo, Q.L., Kuo, G., Weiner, A., Overby, L.R., Bradley, D.W., und Houghton, M., *Isolation of a cDNA clone derived from a blood-born non-A, non-B viral hepatitis genome*. Science, 1989. **244**: p. 359-362.
67. Cohen, J., *The scientific challenge of hepatitis C*. Science, 1999. **285**(5424): p. 26-30.
68. Organization, W.H., *Weekly Epidemiological Record*, in *Weekly Epidemiological Record*. 2000.
69. Wang, T.H., Rijnbrand, R.C., und Lemon, S.M., *Core protein-coding sequence, but not core protein, modulates the efficiency of cap-independent translation directed by the internal ribosome entry site of hepatitis C virus*. J Virol, 2000. **74**(23): p. 11347-11358.
70. Dumas, E., Staedel, C., Colombat, M., Reigadas, S., Chabas, S., Astier-Gin, T., Cahour, A., Litvak, S., und Ventura, M., *A promoter activity is present in the DNA sequence corresponding to the hepatitis C virus 5' UTR*. Nucleic Acids Res, 2003. **31**(4): p. 1275-1281.
71. Santolini, E., Migliaccio, G., und La Monica, N., *Biosynthesis and biochemical properties of the hepatitis C virus core protein*. J Virol, 1994. **68**(6): p. 3631-3641.

72. Srinivas, R.V., Ray, R.B., Meyer, K., und Ray, R., *Hepatitis C virus core protein inhibits human immunodeficiency virus type 1 replication*. *Virus Res*, 1996. **45**(2): p. 87-92.
73. Weiner, A.J., Brauer, M.J., Rosenblatt, J., Richman, K.H., Tung, J., Crawford, K., Bonino, F., Saracco, G., Choo, Q.L., Houghton, M., und et al., *Variable and hypervariable domains are found in the regions of HCV corresponding to the flavivirus envelope and NS1 proteins and the pestivirus envelope glycoproteins*. *Virology*, 1991. **180**(2): p. 842-848.
74. De Francesco, R., *Molecular virology of the hepatitis C virus*. *J Hepatol*, 1999. **31**(Suppl 1): p. 47-53.
75. Mizushima, H., Hijikata, M., Tanji, Y., Kimura, K., und Shimotohno, K., *Analysis of N-terminal processing of hepatitis C virus nonstructural protein 2*. *J Virol*, 1994. **68**(4): p. 2731-2734.
76. Santolini, E., Pacini, L., Fipaldini, C., Migliaccio, G., und Monica, N., *The NS2 protein of hepatitis C virus is a transmembrane polypeptide*. *J Virol*, 1995. **69**(12): p. 7461-7471.
77. Kuo, G., Choo, Q.L., Alter, H.J., Gitnick, G.L., Redeker, A.G., Purcell, R.H., Miyamura, T., Dienstag, J.L., Alter, M.J., Stevens, C.E., und et al., *An assay for circulating antibodies to a major etiologic virus of human non-A, non-B hepatitis*. *Science*, 1989. **244**(4902): p. 362-364.
78. Keskinen, P., Melen, K., und Julkunen, I., *Expression of HCV structural proteins impairs IFN-mediated antiviral response*. *Virology*, 2002. **299**(2): p. 164-171.
79. Enomoto, N., Sakuma, I., Asahina, Y., Kurosaki, M., Murakami, T., Yamamoto, C., Izumi, N., Marumo, F., und Sato, C., *Comparison of full-length sequences of interferon-sensitive and resistant hepatitis C virus 1b. Sensitivity to interferon is conferred by amino acid substitutions in the NS5A region*. *J Clin Invest*, 1995. **96**(1): p. 224-230.
80. Gale, M.J., Jr., Korth, M.J., Tang, N.M., Tan, S.L., Hopkins, D.A., Dever, T.E., Polyak, S.J., Gretch, D.R., und Katze, M.G., *Evidence that hepatitis C virus resistance to interferon is mediated through repression of the PKR protein kinase by the nonstructural 5A protein*. *Virology*, 1997. **230**(2): p. 217-227.
81. Tanaka, T., Kato, N., Cho, M.J., Sugiyama, K., und Shimotohno, K., *Structure of the 3' terminus of the hepatitis C virus genome*. *J Virol*, 1996. **70**(5): p. 3307-3312.
82. Yi, M. und Lemon, S.M., *3' nontranslated RNA signals required for replication of hepatitis C virus RNA*. *J Virol*, 2003. **77**(6): p. 3557-3568.
83. Walewski, J.L., Keller, T.R., Stump, D.D., und Branch, A.D., *Evidence for a new hepatitis C virus antigen encoded in an overlapping reading frame*. *Rna*, 2001. **7**(5): p. 710-721.
84. Choo, Q.L., Richman, K.H., Han, J.H., Berger, K., Lee, C., Dong, C., Gallegos, C., Coit, D., Medina-Selby, R., Barr, P.J., und et al., *Genetic organization and diversity of the hepatitis C virus*. *Proc Natl Acad Sci U S A*, 1991. **88**(6): p. 2451-2455.
85. Simmonds, P., Holmes, E.C., Cha, T.A., Chan, S.W., McOmish, F., Irvine, B., Beall, E., Yap, P.L., Kolberg, J., und Urdea, M.S., *Classification of hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS-5 region*. *J Gen Virol*, 1993. **74**(Pt 11): p. 2391-2399.
86. Robertson, B., Myers, G., Howard, C., Brettin, T., Bukh, J., Gaschen, B., Gojobori, T., Maertens, G., Mizokami, M., Nainan, O., Netesov, S., Nishioka, K., Shin i, T., Simmonds, P., Smith, D., Stuyver, L., und Weiner, A., *Classification, nomenclature, and database development for hepatitis C virus (HCV)*

- and related viruses: proposals for standardization. International Committee on Virus Taxonomy. Arch Virol*, 1998. **143**(12): p. 2493-2503.
87. Zein, N.N., *Clinical significance of hepatitis C virus genotypes. Clin Microbiol Rev*, 2000. **13**(2): p. 223-235.
88. Duarte, E.A., Novella, I.S., Weaver, S.C., Domingo, E., Wain-Hobson, S., Clarke, D.K., Moya, A., Elena, S.F., de la Torre, J.C., and Holland, J.J., *RNA virus quasispecies: significance for viral disease and epidemiology. Infect Agents Dis*, 1994. **3**(4): p. 201-214.
89. Martell, M., Esteban, J.I., Quer, J., Genesca, J., Weiner, A., Esteban, R., Guardia, J., and Gomez, J., *Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasispecies nature of HCV genome distribution. J Virol*, 1992. **66**(5): p. 3225-3229.
90. Ogata, N., Alter, H.J., Miller, R.H., and Purcell, R.H., *Nucleotide sequence and mutation rate of the H strain of hepatitis C virus. Proc Natl Acad Sci U S A*, 1991. **88**(8): p. 3392-3396.
91. Okamoto, H., Kojima, M., Okada, S., Yoshizawa, H., Iizuka, H., Tanaka, T., Muchmore, E.E., Peterson, D.A., Ito, Y., and Mishiro, S., *Genetic drift of hepatitis C virus during an 8.2-year infection in a chimpanzee: variability and stability. Virology*, 1992. **190**(2): p. 894-899.
92. Simmonds, P., *Reconstructing the origins of human hepatitis viruses. Philos Trans R Soc Lond B Biol Sci*, 2001. **356**(1411): p. 1013-1026.
93. Pawlotsky, J.M., *Use and interpretation of virological tests for hepatitis C. Hepatology*, 2002. **36**(5 Suppl 1): p. S65-73.
94. Schroeter, M., *Golden Standard for Genotyping of HCV is sequencing of NS5 Region and comparing to reference sequences*, D. Drutschmann, Editor. 2002: Hamburg.
95. Halfon, P., Trimoulet, P., Bourliere, M., Khiri, H., de Ledinghen, V., Couzigou, P., Feryn, J.M., Alcaraz, P., Renou, C., Fleury, H.J., and Ouzan, D., *Hepatitis C virus genotyping based on 5' noncoding sequence analysis (Trugene). J Clin Microbiol*, 2001. **39**(5): p. 1771-1773.
96. Ross, R.S., Viazov, S.O., Holtzer, C.D., Beyou, A., Monnet, A., Mazure, C., and Roggendorf, M., *Genotyping of hepatitis C virus isolates using CLIP sequencing. J Clin Microbiol*, 2000. **38**(10): p. 3581-3584.
97. Davidson, F., Simmonds, P., Ferguson, J.C., Jarvis, L.M., Dow, B.C., Follett, E.A., Seed, C.R., Krusius, T., Lin, C., Medgyesi, G.A., and et al., *Survey of major genotypes and subtypes of hepatitis C virus using RFLP of sequences amplified from the 5' non-coding region. J Gen Virol*, 1995. **76**(Pt 5): p. 1197-1204.
98. Pourmand, N., Elahi, E., Davis, R.W., and Ronaghi, M., *Multiplex Pyrosequencing. Nucleic Acids Res*, 2002. **30**(7): p. e31.
99. Elahi, E., Pourmand, N., Chaung, R., Rofoogaran, A., Boisver, J., Samimi-Rad, K., Davis, R.W., and Ronaghi, M., *Determination of hepatitis C virus genotype by Pyrosequencing. J Virol Methods*, 2003. **109**(2): p. 171-176.
100. Viazov, S., Zibert, A., Ramakrishnan, K., Widell, A., Cavicchini, A., Schreier, E., and Roggendorf, M., *Typing of hepatitis C virus isolates by DNA enzyme immunoassay. J Virol Methods*, 1994. **48**(1): p. 81-91.

101. Le Pogam, S., Dubois, F., Christen, R., Raby, C., Cavicchini, A., und Goudeau, A., *Comparison of DNA enzyme immunoassay and line probe assays (Inno-LiPA HCV I and II) for hepatitis C virus genotyping*. J Clin Microbiol, 1998. **36**(5): p. 1461-1463.
102. van Doorn, L.J., Kleter, B., Voermans, J., Maertens, G., Brouwer, H., Heijntink, R., und Quint, W., *Rapid detection of hepatitis C virus RNA by direct capture from blood*. J Med Virol, 1994. **42**(1): p. 22-28.
103. Smith, D.B., Mellor, J., Jarvis, L.M., Davidson, F., Kolberg, J., Urdea, M., Yap, P.L., und Simmonds, P., *Variation of the hepatitis C virus 5' non-coding region: implications for secondary structure, virus detection and typing*. The International HCV Collaborative Study Group. J Gen Virol, 1995. **76**(Pt 7): p. 1749-1761.
104. Mitchell, P.S., Sloan, L.M., Majewski, D.W., Rys, P.N., Heimgartner, P.J., Rosenblatt, J.E., Cockerill, F.R., 3rd, Smith, T.F., und Patel, R., *Comparison of line probe assay and DNA sequencing of 5' untranslated region for genotyping hepatitis C virus: description of novel line probe patterns*. Diagn Microbiol Infect Dis, 2002. **42**(3): p. 175-179.
105. Matera, G., Barreca, G.S., Lamberti, A.G., Quirino, A., Foca, D., und Liberto, M.C., *'An atypical pattern of HCV INNO-LiPA subtype'*. J Clin Virol, 2002. **25**(2): p. 237-238.
106. Schroter, M., Zollner, B., Schafer, P., Landt, O., Lass, U., Laufs, R., und Feucht, H.H., *Genotyping of Hepatitis C Virus Types 1, 2, 3, and 4 by a One-Step LightCycler Method Using Three Different Pairs of Hybridization Probes*. J Clin Microbiol, 2002. **40**(6): p. 2046-2050.
107. Read, S.J., Burnett, D., und Fink, C.G., *Molecular techniques for clinical diagnostic virology*. J Clin Pathol, 2000. **53**(7): p. 502-506.
108. Urdea, M.S., Wuestehube, L.J., Laurenson, P.M., und Wilber, J.C., *Hepatitis C--diagnosis and monitoring*. Clin Chem, 1997. **43**(8 Pt 2): p. 1507-1511.
109. Pawlotsky, J.M., *Measuring hepatitis C viremia in clinical samples: can we trust the assays?* Hepatology, 1997. **26**(1): p. 1-4.
110. Germer, J.J., Heimgartner, P.J., Ilstrup, D.M., Harmsen, W.S., Jenkins, G.D., und Patel, R., *Comparative evaluation of the VERSANT HCV RNA 3.0, QUANTIPLEX HCV RNA 2.0, and COBAS AMPLICOR HCV MONITOR version 2.0 Assays for quantification of hepatitis C virus RNA in serum*. J Clin Microbiol, 2002. **40**(2): p. 495-500.
111. Kawai, S., Yokosuka, O., Kanda, T., Imazeki, F., Maru, Y., und Saisho, H., *Quantification of hepatitis C virus by TaqMan PCR: comparison with HCV AmpliCor Monitor assay*. J Med Virol, 1999. **58**(2): p. 121-126.
112. Lin, H.J., Shi, N., Mizokami, M., und Hollinger, F.B., *Polymerase chain reaction assay for hepatitis C virus RNA using a single tube for reverse transcription and serial rounds of amplification with nested primer pairs*. J Med Virol, 1992. **38**(3): p. 220-225.
113. Wiest, M., *Diagnostik der Hepatitis C Infektion*. 2004, [http://www.hepatitis-care.de/content/literatur/artikel/diagnostik\\_wiest/diagTab1.htm](http://www.hepatitis-care.de/content/literatur/artikel/diagnostik_wiest/diagTab1.htm)
114. Pawlotsky, J.M., *Molecular diagnosis of viral hepatitis*. Gastroenterology, 2002. **122**(6): p. 1554-1568.
115. Sarrazin, C., Teuber, G., Kokka, R., Rabenau, H., und Zeuzem, S., *Detection of residual hepatitis C virus RNA by transcription-mediated amplification in patients with complete virologic response according to polymerase chain reaction-based assays*. Hepatology, 2000. **32**(4 Pt 1): p. 818-823.

116. Pawlowsky, J.M., *Use and interpretation of hepatitis C virus diagnostic assays*. Clin Liver Dis, 2003. **7**(1): p. 127-137.
117. Perrin, A., Duracher, D., Perret, M., Cleuziat, P., und Mandrand, B., *A combined oligonucleotide and protein microarray for the codetection of nucleic acids and antibodies associated with human immunodeficiency virus, hepatitis B virus, and hepatitis C virus infections*. Anal Biochem, 2003. **322**(2): p. 148-155.
118. <http://www.ncbi.nlm.nih.gov>.
119. <http://www.ddbj.nig.ac.jp>.
120. <http://srs.ebi.ac.uk/>.
121. <http://www.ebi.ac.uk/clustalw/>.
122. Stuyver, L., Wyseur, A., van Arnhem, W., Hernandez, F., und Maertens, G., *Second-generation line probe assay for hepatitis C virus genotyping*. J Clin Microbiol, 1996. **34**(9): p. 2259-2266.
123. Rozen, S. und Skaletsky, H.J., *Primer3 on the WWW for general users and for biologist programmers*, in *Bioinformatics Methods and Protocols: Methods in Molecular Biology*, S. Krawetz and S. Misener, Editors. 2000, Human Press: Totowa, NJ. p. 365-386.
124. <http://www.bioinfo.rpi.edu>.
125. <http://www.ebi.ac.uk/fasta33/>.
126. <http://www.ncbi.nlm.nih.gov/BLAST/>.
127. <http://home/zait/uni-bremen.de/~ellola/gensensorik/configTool/>.
128. Kosaraju, S.R., Schaffer, A.A., und Biesecker, L.G., *Approximation algorithms for a genetic diagnostics problem*. J Comput Biol, 1998. **5**(1): p. 9-26.
129. Roggendorf, M., *HCV Panel*. 2002, <http://www.hepatitis-c-referenzzentrum.de/html/hcvpanel.htm>
130. Schmidt, W.N., Klinzman, D., LaBrecque, D.R., Macfarlane, D.E., und Stapleton, J.T., *Direct detection of hepatitis C virus (HCV) RNA from whole blood, and comparison with HCV RNA in plasma and peripheral blood mononuclear cells*. J Med Virol, 1995. **47**(2): p. 153-160.
131. Boldt, L., *Methodische Untersuchungen im Vorfeld der Entwicklung miniaturisierter DNA-analytischer Verfahren*, in *Fachbereich Biologie/Chemie*. 1999, Bremen: Bremen. p. 117.
132. Innis MA and Gelfand DH (1990). Optimization of PCRs. pp. 3-12 in: PCR Protocols (Innis, G. und Sninsky and White, e.A.P., New York., *Optimization of PCRs*. PCR Protocols, ed. G. Innis and S.a. White. 990, New York: Academic Press. 3-12.
133. Blake, R.D. und Delcourt, S.G., *Thermodynamic effects of formamide on DNA stability*. Nucleic Acids Res, 1996. **24**(11): p. 2095-2103.
134. Guo, Z., Guilfoyle, R.A., Thiel, A.J., Wang, R., und Smith, L.M., *Direct fluorescence analysis of genetic polymorphisms by hybridization with oligonucleotide arrays on glass supports*. Nucleic Acids Res, 1994. **22**(24): p. 5456-5465.
135. Drutschmann, D., *Virendiagnostik per Chip*. Nachrichten aus der Chemie, 2002. **50**: p. 54.
136. Frodl, R., *in-house PCR*, D. Drutschmann, Editor. 2002: Bremen.
137. Southern, E., Mir, K., und Shchepinov, M., *Molecular interactions on microarrays*. Nat Genet, 1999. **21**(1 Suppl): p. 5-9.

138. Shchepinov, M.S., Case-Green, S.C., und Southern, E.M., *Steric factors influencing hybridisation of nucleic acids to oligonucleotide arrays*. Nucleic Acids Res, 1997. **25**(6): p. 1155-1161.
139. Mir, K.U. und Southern, E.M., *Determining the influence of structure on hybridization using oligonucleotide arrays*. Nat Biotechnol, 1999. **17**(8): p. 788-792.
140. Riccelli, P.V., Merante, F., Leung, K.T., Bortolin, S., Zastawny, R.L., Janeczko, R., und Benight, A.S., *Hybridization of single-stranded DNA targets to immobilized complementary DNA probes: comparison of hairpin versus linear capture probes*. Nucleic Acids Res, 2001. **29**(4): p. 996-1004.
141. Fuchs, B.M., Glockner, F.O., Wulf, J., und Amann, R., *Unlabeled helper oligonucleotides increase the in situ accessibility to 16S rRNA of fluorescently labeled oligonucleotide probes*. Appl Environ Microbiol, 2000. **66**(8): p. 3603-3607.
142. Niemeyer, C.M., Boldt, L., Ceyhan, B., und Blohm, D., *DNA-Directed immobilization: efficient, reversible, and site-selective surface binding of proteins by means of covalent DNA-streptavidin conjugates*. Anal Biochem, 1999. **268**(1): p. 54-63.
143. Peplies, J., Glockner, F.O., und Amann, R., *Optimization strategies for DNA microarray-based detection of bacteria with 16S rRNA-targeting oligonucleotide probes*. Appl Environ Microbiol, 2003. **69**(3): p. 1397-1407.
144. Urakawa, H., Noble, P.A., El Fantroussi, S., Kelly, J.J., und Stahl, D.A., *Single-base-pair discrimination of terminal mismatches by using oligonucleotide microarrays and neural network analyses*. Appl Environ Microbiol, 2002. **68**(1): p. 235-244.
145. Brown, E.A., Zhang, H., Ping, L.H., und Lemon, S.M., *Secondary structure of the 5' nontranslated regions of hepatitis C virus and pestivirus genomic RNAs*. Nucleic Acids Res, 1992. **20**(19): p. 5041-5045.
146. Bhanot, G., *The Importance of Thermodynamic Equilibrium for High Throughput Gene Expression Arrays*. 124 Biophysical Journal, 2003. **84**: p. 124-135.
147. Ku, W.C., Lau, W.K., Tseng, Y.T., Tzeng, C.M., und Chiu, S.K., *Dextran sulfate provides a quantitative and quick microarray hybridization reaction*. Biochem Biophys Res Commun, 2004. **315**(1): p. 30-37.
148. Bull, M. *Stability of Cy5.0 dye-labelled oligonucleotides and decreasing Cydye Signal loss in microarray experiments*. in Northwest gene expression conference. 2003. washington.
149. Vandesompele, J., *Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes*. Genome Biol, 2002. **3**(7).
150. Stuyver, L., Wyseur, A., van Arnhem, W., Lunel, F., Laurent-Puig, P., Pawlotsky, J.M., Kleter, B., Bassit, L., Nkengasong, J., van Doorn, L.J., und et al., *Hepatitis C virus genotyping by means of 5'-UR/core line probe assays and molecular analysis of untypeable samples*. Virus Res, 1995. **38**(2-3): p. 137-157.
151. Niemeyer, C.M., Boldt, L., Ceyhan, B., und Blohm, D., *Evaluation of single-stranded nucleic acids as carriers in the DNA-directed assembly of macromolecules*. JBSD, 1999. **17**(3): p. 527-538.
152. Cristina, J., Mukomolov, S., Colina, R., Kalinina, O., Garcia, L., Khan, B., Mogdasy, C., und Karayiannis, P., *Hepatitis C virus phylogeny: a useful clinical tool*. Acta Virol, 2002. **46**(3): p. 179-182.
153. Trutnau, H.-H.N., M.; Volkmann, G.; Drutschmann, D.; Blohm, D., *Genotypisierung von Hepatitis-C-Viren mittels DNA-Mikroarray-Technik durch Bioinformatik verbessert*. Biospektrum, 2003. **1**: p. 93-95.

154. Kalinina, O., Norder, H., Mukomolov, S., und Magnius, L.O., *A natural intergenotypic recombinant of hepatitis C virus identified in St. Petersburg*. J Virol, 2002. **76**(8): p. 4034-4043.
155. Viazov, S., Widell, A., und Nordenfelt, E., *Mixed infection with two types of hepatitis C virus is probably a rare event*. Infection, 2000. **28**(1): p. 21-25.
156. Gutell, R.R., Larsen, N., und Woese, C.R., *Lessons from an evolving rRNA: 16S and 23S rRNA structures from a comparative perspective*. Microbiol Rev, 1994. **58**(1): p. 10-26.
157. Wang, D., Coscoy, L., Zylberberg, M., Avila, P.C., Boushey, H.A., Ganem, D., und DeRisi, J.L., *Microarray-based detection and genotyping of viral pathogens*. Proc Natl Acad Sci U S A, 2002. **99**(24): p. 15687-15692.
158. Sengupta, S., Onodera, K., Lai, A., und Melcher, U., *Molecular detection and identification of influenza viruses by oligonucleotide microarray hybridization*. J Clin Microbiol, 2003. **41**(10): p. 4542-4550.
159. Laassri, M., Chizhikov, V., Mikheev, M., Shchelkunov, S., und Chumakov, K., *Detection and discrimination of orthopoxviruses using microarrays of immobilized oligonucleotides*. J Virol Methods, 2003. **112**(1-2): p. 67-78.
160. Takahashi, M., Okada, J., Ito, K., Hashimoto, M., Hashimoto, K., Yoshida, Y., Furuichi, Y., Ohta, Y., Mishiro, S., und Gemma, N., *Electrochemical DNA array for simultaneous genotyping of single-nucleotide polymorphisms associated with the therapeutic effect of interferon*. Clin Chem, 2004. **50**(3): p. 658-661.
161. Sen, A., Steele, R., Ghosh, A.K., Basu, A., Ray, R., und Ray, R.B., *Inhibition of hepatitis C virus protein expression by RNA interference*. Virus Res, 2003. **96**(1-2): p. 27-35.
162. Yokota, T., Sakamoto, N., Enomoto, N., Tanabe, Y., Miyagishi, M., Maekawa, S., Yi, L., Kurosaki, M., Taira, K., Watanabe, M., und Mizusawa, H., *Inhibition of intracellular hepatitis C virus replication by synthetic and vector-derived small interfering RNAs*. EMBO Rep, 2003. **4**(6): p. 602-608.



## 8 ANHANG

### Abkürzungsverzeichnis

5' UTR	Nichttranslatierte Region am 5' Ende ( <i>engl.</i> untranslated region)
a.u.	Relative Einheiten
bp	Basenpaare
BSA	Rinderserumalbumin ( <i>engl.</i> Bovine Serum Albumin)
DDBJ	DNA Data Bank of Japan
ddNTP	Didesoxynukleotidtriphosphat
DRK	Deutsches Rotes Kreuz
DNA	Desoxyribonukleinsäure
FPLC	Fast Protein Liquid Chromatography
GT	Genotyp
HCV	Hepatitis C-Virus
IRES	Internal Ribosomale Entry Site
Kb	Kilobasen
MM	Mismatch
NaCl	Natriumchlorid
NCBI	The National Center for Biotechnology Information
nt	Nukleotiden
NTP	Nukleotidtriphosphat
PCR	Polymerase Ketten Reaktion ( <i>engl.</i> Polymerase Chain Reaction)
PDITC	Phenylendiisothiocyanat
PM	Perfect Match
PVP	Polyvinylpyrrolidon
RiFS	Reflectometric Interference Spectroscopy
RT	Raumtemperatur
SBE	Einzelstrangverlängerung ( <i>engl.</i> Single Base Extension)
SNP	Einzelbasenaustausch ( <i>engl.</i> Single Nucleotide Polymorphism)
SRS	Sequence Retrieval System
ÜN	Über Nacht

Tabelle 8: Auflistung der manuell ausgewählten Oligonukleotide

Bezeichnung	Länge	Schmelztemperatur [°C]	Oligonukleotid Sequenz 5'→ 3'	Lage auf M62321
<b>Genotyp 1</b>				
M-1u1B	17	60,4	AATTGCCAGGACGACCG	172-188
M-1u1C	17	57,5	AATTGCCAGGATGACCG	172-188
M-1u1D	17	60,4	AATTGGCAGGACGACCG	172-188
M-1u1E	16	57,4	ATTGGCCAGGACGACC	172-187
M-1u1F	15	58,7	ATTGCCAGGGCGACC	173-187
M-1u2A	17	59,3	CGCTCAATGCCTGGAGA	209-225
M-1u2B	17	59,5	CGCTCAGTGCCTGGAGA	209-225
M-1u2C	16	59,9	CGCTCGATGCCTGGAG	209-224
M-1u2D	18	57,9	CCACTCAATGCCTGGAGA	208-225
<b>Subtyp 1a</b>				
M-1a1A	16	60,7	CCCCCGCAAGACTGCT	236-251
M-1a1B	15	59,2	CCCCGCGAGACTGCT	237-251
M-1a1C	14	57,3	CCCGCAAGACCGCT	238-251
<b>Subtyp 1b</b>				
M-1b2B	15	60,2	CCCGCCAGACCGCTA	238-252
M-1b2C	16	59,3	CCCCGCGAGACTGCTA	237-252
M-1b2D	15	57,4	CCCGCAAGACCGCTA	238-252
M-1b2E	17	57,9	CCCGCAAGACTGCTAGC	238-254
<b>Genotyp 2</b>				
M-2u2B	16	58,0	TAGCGTTGGGTTGCGA	259-274
M-2u2C	16	57,9	TAGGGTTGGGTCGGGA	259-274
<b>Subtyp 2a</b>				
M-2a1	17	61,5	GCCGGGAAGACTGGGTC	276-192
M-2a3	17	59,2	ACCCACTCTATGCCCGG	206-222
<b>Subtyp 2b</b>				
M-2b4	21	61,3	ATTACCGGAAAGACTGGGTCC	173-193
M-2b5	21	60,5	ACCCACTCTATGTCCGGTCAT	206-226
<b>Genotyp 3</b>				
M-3u2	16	57,4	AATCGCTGGGGTGACC	172-187
M-3u3	21	59,4	CGCTCAATACCCAGAAATTG	209-229
M-3u4	17	57,6	CCGCGAGATCACTAGCC	239-255
M-3u5A	16	59,1	TCTTGGAGCAACCCGC	196-211
M-3u5B	17	57,8	TTCTTGGAAACAACCCGC	195-211
M-3u5C	19	59,1	TCTTGGAGTAACCCGCTCA	196-214
M-3u7	18	59,5	TCTTGGAAACAACCCGCTC	196-213
<b>Genotyp 4</b>				
M-4u1	16	60,1	AATCGCCGGGATGACC	172-187
M-4u2	16	58,9	GCTCAATGCCCGGAAA	210-225
M-4u3	14	61,7	CGCCGGGATGACCG	176-188
<b>Universelle Oligos</b>				
M-U20B	18	58,2	TAGGGTGCTTGCAAGTGC	297-314
M-U20C	16	59,8	GGGTGCTTGCGAGTGC	299-314
M-U20D	18	59,5	TAGGGTGCTTGGGAGTGC	297-314
M-U20E	16	59,6	TGGGGTGCTTGCGAGT	297-312
M-U20F	17	59,2	TAGGGTGCTGGCGAGTG	297-313
M-U20G	17	57,7	TAGGGTGCTGGGGAGTG	297-313
M-U23A	18	60,1	GATAGGGCGCTTCGGAGT	295-312
M-U23B	20	58,7	TGATAGGGTGCTCGAGAGTG	294-313
M-U23C	17	59,3	ATAGGGCGCTTGCGAGT	296-312
M-U23D	18	58,5	GATAGGGTGCTGGCGAGT	295-312
M-U23E	19	60,3	TGATAGGGTGCTGGGGAGT	294-312
M-U23F	19	56,5	TGATAGGGTGCTTGCAAGT	294-312
M-U23G	19	59,2	TGATAGGGTGCTTGCGAGT	294-312
M-U23H	19	57,8	TGATAGGGTGCTTGGGAGT	294-312
M-U23I	17	60,0	ATGGGGTGCTTGCGAGT	296-312
M-U23J	19	56,6	TGATAGGGTGCTCGAGAGA	294-311
M-U6B	23	58,1	GTAAGGCTGATAGGGTGCTT	286-306
M-U6C	16	57,6	GCCTGATGGGGTGCTT	291-306
M-U6D	21	58,6	GTAAGGCTGATAGGGTGCTC	286-305
M-U6E	16	57,0	GCCTGATAGGGCGCTT	291-306
M-U6F	22	59,2	GGTACTGCTGATAGGGTGTTT	285-306
M-U6G	21	59,0	GTAAGGCTGATAGGGTGCTG	286-305

Tabelle 9: Auflistung der Software-generierten Oligonukleotide

Bezeichnung	Länge	Schmelztemperatur [°C]	Oligonukleotid Sequenz 5' -> 3'	Lage auf M62321
<b>Genotyp 1</b>				
<b>Subtyp 1a</b>				
I 1a-1	16	60,6	GTGCCCCGCAAGACT	233-248
I 1a-2	17	59,1	CGCAAGACTGCTAGCCG	240-256
I 1a-3	17	59	TCAAACCCGCTCAATGC	205-218
I 1a-4	18	59	TAAACCCGCTCAATGCCT	203-220
I 1a-5	22	57,2	AAGACTGCTAGCCGAGTAGTGT	243-264
<b>Subtyp 1b</b>				
I 1b-1	16	59,1	ACCCGCTCAGTGCCTG	206-221
I 1b-2	17	57,8	CTCCCGGAGAGCATAG	126-142
I 1b-3	17	60,5	CCCCGCGAGACTGCTAG	237-253
I 1b-4	17	60,6	CCGCGAGACTGCTAGCC	279-255
I 1b-5	17	60,6	CAACCCGCTCAATGCCT	204-220
I 1b-6	19	59,5	ATTAACCCGCTCAGTGCCT	202-220
I 1b-7	19	58,5	GAGACCGCTAGCCGAGTAGT	244-262
I 1b-8	20	57,8	CGAGACTGCTAGCCGAGTAG	242-261
I 1b-9	21	60,6	CCTGATAGGGTGCTCGAGAGT	292-312
I 1b-10	15	58,3	GCCAGCCCCTAACGG	118-132
I 1b-11	15	59,2	CCCCGCGAGACTGCT	237-251
<b>Genotyp 2</b>				
I 2u-1	14	56,7	TCCGGTCATTGGGC	219-232
I 2u-2	21	59,9	AAGACTGGGTCTTTCTTGGGA	182-202
I 2u-3	20	58,7	ACTCTATGTCCGGTCATTGG	211-230
I 2u-4	21	60,5	ACCCACTCTATGTCCGGTCAT	206-226
I 2u-5	25	59,3	CTTTCTTGGATAAACCCACTCTATG	193-217
<b>Genotyp 3</b>				
I 3u-1	19	59,6	AACAACCCGCTCAATACCC	202-220
I 3u-2	22	58,9	CGAGATCACTAGCCGAGTAGTG	242-263
I 3u-3	24	58,9	AAGATCACTAGCCGAGTAGTGTG	243-266
I 3u-4	15	58,7	CGGAATCGCTGGGGT	169-183
I 3u-5	13	59,3	CCCCCGGAGATCAC	236-250
I 3u-6	17	60,3	CAACCCGCTCAATACCCA	204-221
<b>Genotyp 4</b>				
I 4u-1	16	57,3	ATGCCCGGAGATTGG	215-230
I 4u-2	18	56,6	TTTCTTGGATCAAACCCG	194-211
I 4u-3	15	59,1	GGAATCGCCGGGATG	170-184
<b>universelle Oligos</b>				
I U-1	19	59,9	ATAGGGTGCTTGGGAGTGC	296-314
I U-2	21	60	CCTGATAGGGTGTTGGGAGT	292-312
I U-3	21	59,8	CTGATAGGGTGCTCGAGAGTG	293-313
I U-4	17	57,5	AATTGCCAGGATGACCG	172-188
I U-5	17	59,8	TTGTGCAGCCTCCAGGA	103-119
I U-6	18	60,4	GAACAGCCTCCAGGACCC	105-122
I U-7	17	60	CGTACAGCCTCCGGGAC	104-120
I U-8	18	58,9	GTCGTACAGCCTCCAGGG	102-119
I U-9	16	59,9	CAGCCTCCAGGACCC	108-123
I U-10	18	57,7	GTACAGCCTCCAGGACCC	105-122
I U-11	15	59,4	CCTCCAGGACCC	111-125
I U-12	15	60,4	TGCAGCCTCCAGGCC	106-120
I U-13	14	57,6	ACAGCCTCCAGGCC	108-121
I U-14	17	58,2	ACAGCCTCCAGAACCC	107-123
I U-15	16	60,4	CGTGCAGCCTCCAGGA	104-119
I U-16	20	60,4	CTCAATGCCTGGAGATTGG	211-230
I U-17	15	58,4	GCGTTGGGTGCGAA	261-275
I U-18	19	60,4	CTTTCTTGGAAACAACCCGC	193-211
I U-19	19	59,8	TCCTTTCTCGGAACAACCC	191-209
I U-20	17	57,9	TTTCTTGGAGCAACCCG	194-210
I U-21	17	57	TCAATGCCCGGAAATTT	212-228
I U-22	17	58,7	TCAATGCCCGGAGATTG	212-228
I U-23	20	58,6	TACTGCCTGATAGGGTGCTG	287-306
I U-24	20	57,2	TACTGCCTGATAGGGTGCTT	287-306







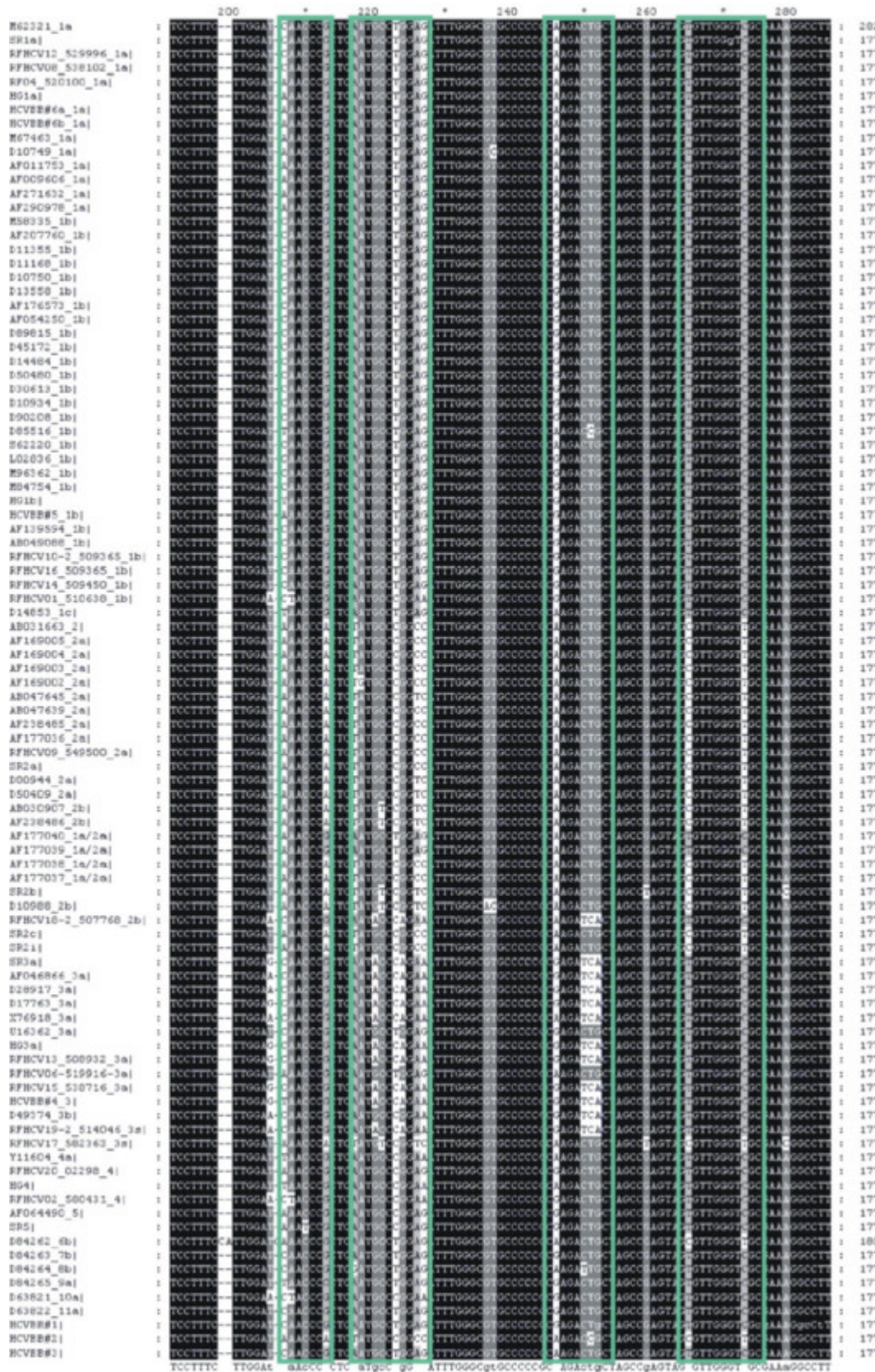


Abbildung 8-1: Alignment der sequenzierten HCV-Targets mit publizierten Sequenzen aus der NCBI. Die grün-umrandeten sieben Regionen stellen die Bereiche dar, die Stuyver für die Genotypisierung innerhalb der 5' UTR herangezogen hat <sup>[122]</sup>.

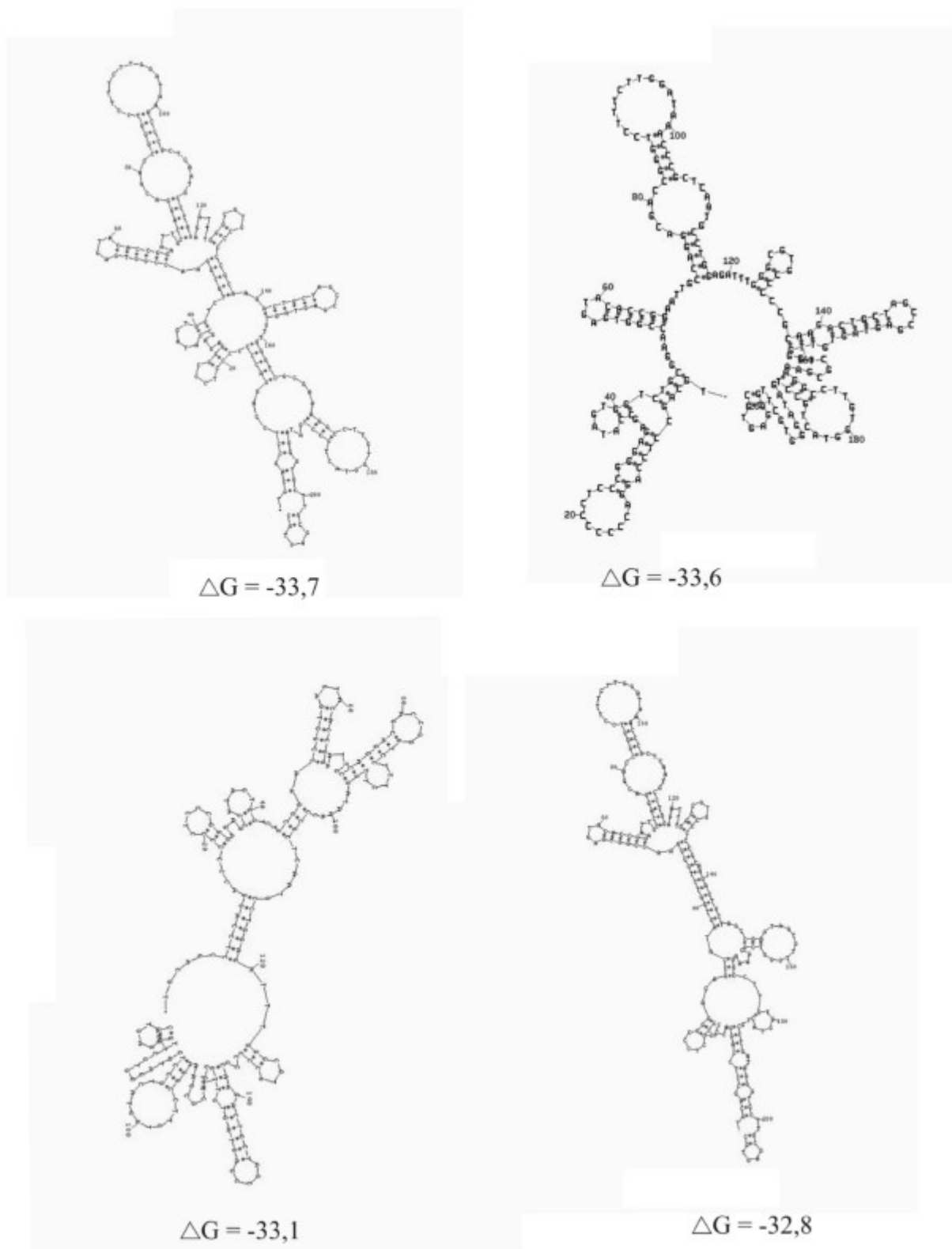


Abbildung 8-2: Graphische Darstellung der verschiedenen Sekundärstrukturen nach Zuker *et al.* <sup>[55]</sup>, die bei 22 °C ausgebildet werden können..

## 9 DANKSAGUNG

Ich möchte meinem Doktorvater, Prof. Dr. Dietmar Blohm, danken, der im CAG federführend tätig war und ohne den diese interessante, interdisziplinäre Arbeit nicht möglich gewesen wäre!

Herrn Dr. Andreas Dotzauer möchte ich sehr gerne für die Übernahme des Zweitgutachtens danken.

Ich danke allen Kollegen aus den Teilprojekten, deren tatkräftige Unterstützung sehr zum Gelingen dieser Arbeit beigetragen hat. Besonders nennen möchte ich hier: Thorsten Lilienkamp (Roboter), Rüdiger Benters und Britta Meyer-Schlosser (aktivierte Slides), Manfred Nölte (Oligos), Gerald Volkmann (Auswertung), Frank Jäger (IQmega), Tom Wetjen (Genotypisierung), Jörg Peplies (Hybridisierung und Spotter), Lars Ziegenhagen (Pipetten).

Ganz besonderer Dank gilt der Arbeitsgruppe BMG: Zu Eurer Kollegialität, Kameradschaftlichkeit und Hilfsbereitschaft kann man Euch nur beglückwünschen. Es war schön mit Euch zu arbeiten! Michael Adler möchte ich für die Bereitstellung der Abbildung 1-7 danken.

Bei Herrn Dr. Reinhard Frodl, Labor Dr. Gärtner in Weinheim und Herrn Dr. Stefan Roß, Virologie der Universität Essen, möchte ich mich für ihre Unterstützung mit HCV-Proben bedanken.

Britta Becker möchte ich auch sehr herzlich für ihre Hilfe mit den Serumproben danken.

Gerne möchte mich auch für die Mühe mit Korrekturlesen bei Björn, Christiane, Heike und Sonja bedanken

Heike, Dorothea, Regina, Christiane und Ilka möchte ich sehr danken – ich denke gerne an unsere Abende in der Kartoffeldiele!

Meinen Eltern möchte ich für die jahrelange finanzielle und vor allem mentale Unterstützung danken!

Und vor allem möchte ich Ron danken – er musste mich beim Zusammenschreiben ertragen ;-)

