

Ziel der Arbeit war es, kohlenhydratbindende Proteine *in vitro* zu selektieren, deren Zielmolekül Heparin, Streptomycin, N-Acetylglucosamin bzw. Sialyl-Lewis X sein kann.

*In vitro*-Mutationsmethoden wie das „*Shuffling*“, ITCHY und SCRATCHY sollten etabliert werden, um Bibliotheken zu generieren, deren Ursprung bzw. „*Scaffold*“ in den S-layer-Proteinen verschiedener Organismen zu finden sind. Um den besten Protein-Binder zu finden, wurde die Methode des „*Ribosome Display*“ angewandt, die ebenfalls zellfrei durchgeführt werden kann und damit von zellspezifischen Eigenarten wie Toxizität, Wachstum und Vermehrung unabhängig ist. Die Kombination von solchen *in vitro*-Techniken führt in letzter Konsequenz zu einer *in vitro*-Evolution, die ein schnelleres Verfahren zur Entwicklung von verbesserten Molekülen darstellt.

## 1 Auswahl des „*Scaffolds*“ zur Generierung kohlenhydratbindender Proteine

Das „*Scaffold*“ (dt.: Gerüst) dient als Grundlage für weitere Veränderungen innerhalb der Sequenz. Typischerweise ist damit die Proteinsequenz gemeint, die auf DNA-Ebene moduliert und durch Translation in eine mutierte Proteinsequenz umgeschrieben wird. Eine Vielzahl an veränderten Sequenzen wird als Bibliothek bezeichnet.

Proteine, die die Eigenschaft besitzen Kohlenhydrate zu binden, werden als Lektine bezeichnet. Die Wechselwirkung zwischen Protein und Kohlenhydrat ist dabei von spezifischen Interaktionen geprägt, d.h. die Wechselwirkung ist strukturell begrenzt und genau lokalisierbar. Deshalb findet man oft nur Proteine mit einer Spezifität zu einem Kohlenhydrat, die eine Anwendung zur Analyse von Glykosylierungen ermöglicht (Tab. 18).

**Tabelle 18: Einige Lektine und ihre Spezifität**

Lektin (Quelle)	Spezifität
WGA ( <i>Triticum vulgare</i> )	(D-GlcNAc) <sub>n</sub>
Con A ( <i>Canavalia ensiformis</i> )	α-D-Man, α-Glc
LCA ( <i>Lens culinaris</i> )	endständige α-D-Man & α-Glc
AAA ( <i>Aleuria aurantia</i> )	-α(1-6)Fuc
MAA ( <i>Maackia amurensis</i> )	α(2-3) verknüpfte Sialinsäuren

Die S-layer Proteine, die die äußerste Begrenzung der Bakterien darstellen, müssen in der Zellwand halt finden. Schaut man sich die Struktur der Zellwand von Gram-positiven Bakterien an (I 4.2), so kann ein Zuckergrundgerüst gefunden werden, bestehend aus N-Acetylglucosamin und N-Acetylmuraminsäure. Dies wird als Peptidoglycan bezeichnet. Zellwand assoziierte Strukturen sind z.B. Teichonsäuren, Teichuronsäuren, Lipoteichonsäuren und Polysaccharidmodifikationen. Es handelt sich dabei um anionische Polymere (sekundäre Zellwandpolymere) die entsprechend aus Polyglycerolphosphat, Polyribitolphosphat oder Polyglycosylphosphat bestehen können (I 5.2). Die Verankerung der S-layer-Proteine wird vermutlich zu den sekundären Zellwandpolymeren hergestellt, insbesondere durch den N-terminalen Teil, der SLH-Domäne (Mesnage *et al.*, 2000, Mader *et al.*, 2004). Wie hoch die Spezifität der Erkennung durch die S-layer-Proteine ist, wurde bislang erst einmal gezeigt (Mader *et al.*, 2004). Danach scheint eine ähnlich begrenzte, lokalisierte Interaktion zu den Kohlenhydraten wie bei den Lektinen zu bestehen. Es ist wahrscheinlich, dass die Interaktion der S-layer-Proteine zu den Kohlenhydraten über die SLH-Domänen vermittelt wird, denn diese kommen in den meisten S-layer-Proteinen von Gram-positiven Bakterien vor, mit unterschiedlicher Sequenz aber z.T. konservierten Aminosäuren. Es war interessant zu erfahren, ob diese putative Bindung verbessert bzw. die Spezifität der Interaktion aufgehoben werden kann oder ob sie auf andere Kohlenhydrate übertragbar ist. Aufgrund dieser Überlegung wurden die folgenden Kohlenhydratmoleküle ausgewählt. **Heparin**, ein Polymer, bestehend aus sulfatierten Glucosaminomolekülen, ähnlich der sekundären Zellwandpolymere; **Streptomycin**, ebenso wie Heparin ein Aminoglycosid und unter anderem mit einem Anteil an seltenen Kohlenhydraten versehen; **Sialyl-Lewis X**, ein essentielles Molekül von Tumorzellen und Leukozyten zum Eindringen in andere Zellen, was ebenfalls Glucosamin enthält und **N-Acetylglucosamin** als Monosaccharid, was als Kontrolle diente. Die Wahl der Organismen nach verschiedenen Kriterien ist in III 4 dargestellt.

## 2 Klonierung des S-layer-Proteins aus *Bacillus stearothermophilus*

Das Surface-layer-Protein aus *Bacillus stearothermophilus* sollte in Analogie zu den anderen S-layer-Proteinen aus chromosomaler DNA amplifiziert, in pBluescript ligiert und in *E. coli* XL1 propagiert werden, um aus der Slp-Sequenz, die SLH-Sequenz (Surface-layer-homologe

Domäne) zu gewinnen. Die SLH-Sequenz ist der N-terminale Teil der Slp-Sequenz, welche in Gram-positiven Bakterien als putativer Verankerungspunkt zur Zellwand gilt.

Ein entsprechender Klon mit der Slp-Sequenz wurde nicht erhalten stattdessen lediglich die putative SLH-kodierende Sequenz. Wahrscheinlich wurden durch die Aufarbeitung der chromosomalen DNA Brüche in dieser erzeugt, was den Verlust eines zusammenhängenden ca. 3 kb langen Gen-Fragmentes bedeutet. Möglich ist auch, dass Sekundärstrukturen der chromosomalen DNA eine Amplifizierung unmöglich machten. Für die weiteren Arbeiten ist die Klonierung der Slp-Sequenz nicht von Bedeutung gewesen.

### 3 Mutagenese-Methoden

#### 3.1 „DNA-Shuffling“

Das „DNA-Shuffling“ ist eine der ersten entwickelten Mutations-Methoden, die das Einstreuen von Punktmutationen über einen Sequenzbereich ermöglichen. Dabei handelt es sich um eine sehr anspruchsvolle Methode, was sich in Optimierungen der einzelnen Reaktionen anderer Arbeitsgruppen widerspiegelt.

Die Isolierung kleiner Fragmente gleichmäßiger Qualität stellt dabei eines der Hauptprobleme dar. Schwierigkeiten bei der Gewinnung kleiner Fragmente, der Homogenität der Fragmente, Konzentration und Ausbeute dieser sind zu verzeichnen.

Anfangs wurde versucht, die von Stemmer (1994) publizierten Fragmentgrößen  $\leq 50$  bp zu separieren, was ein Problem bei der Aufreinigung und der Rekonstruktion darstellte. Die meisten Geextraktions-Kits sind so konzipiert, dass Primer und DNA-Fragmente unter 100 bp abgetrennt werden. Die Methode des Band Interception (II 2.1.11) ist für kleine Fragmente nur bedingt geeignet, da die geringe Größe der DNA-Fragmente auch zu schwächerer Bindung an die Membran führen, so dass Verluste auftreten.

Nach vielen Optimierungen wurde dazu übergangen, die DNA-Fragmente nicht aufzureinigen, sondern nur die DNase I zu inaktivieren. Es wurden DNA-Fragmente in einem Größenbereich von 50-400 bp eingesetzt, wobei die Konzentration der größeren Fragmente weit unter der Konzentration der kleineren Fragmente lag. Eingestellt wurde dieses Verhältnis durch Einsatz eines DNA-Fragment-Gemisches, das sich aus einzelnen DNase I-Proben, welche zu unterschiedlichen Zeitpunkten gestoppt wurden, zusammensetzte.

Zur Überprüfung, ob noch vollständige DNA-Sequenzen vorlagen, wurde die PCR benutzt. Wäre dies der Fall, so würde das zu einer geringeren Mutationsrate führen. Die Etablierung der Methode ist in erster Linie sehr zeitaufwändig, was verständlich wird, wenn man die Vielzahl an Parametern betrachtet. Wie in einer konventionellen PCR sind nicht nur die Konzentrationen der einzelnen Substanzen ( $Mg^{2+}$ , DNA, dNTP, Primer) ausschlaggebend, sondern in großem Maße die verwendeten PCR-Programmparameter. Kurze Annealing- und Elongationszeiten sowie eine hohe Zyklenzahl wirken sich förderlich auf die Mutationsrate aus.

### 3.2 SCRATCHY

Das SCRATCHY ist eine Mutationsmethode die das „*Shuffling*“ und das ITCHY (Incremental Truncation for the Creation of Hybrid Enzymes) miteinander vereint. Vorteil dieser Methode gegenüber anderen Mutations-Techniken ist eine hohe Mutationsrate und Sequenzunabhängigkeit. Will man sequenzübergreifend, d.h. Sequenzen einer Familie mutieren, so bot sich lange Zeit nur das „*family shuffling*“ an. Dieses konnte allerdings nur durchgeführt werden, wenn eine hohe Homologie der Sequenzen zueinander vorlag. Das SCRATCHY ist darauf nicht angewiesen und kann so beliebig angewendet werden.

Das ITCHY beinhaltet zwei einzelne Reaktionen gleichen Prinzips. Die DNA-Enden werden nach Generierung eines 5'-Überhanges sukzessive degradiert, was mit Hilfe von Restriktionsendonukleasen erfolgen kann. Die erste Degradierung findet mit Exonuklease III statt, die aus einem DNA-Doppelstrang einen partiellen Einzelstrang herstellt. Die Länge des Einzelstranges kann über die Enzymaktivität bestimmt werden, die wiederum hauptsächlich durch Parameter wie Zeit und Temperatur regulierbar ist. Die zweite Degradierung, welche mit Mung Bean Nuklease realisiert wird, führt durch Entfernung des Einzelstranges wieder zu einem DNA-Doppelstrang, der allerdings je nach Einwirkdauer der Enzyme unterschiedlich verkürzt wurde. Wird die erste Degradierungs-Reaktion zu unterschiedlichen Zeiten abgestoppt, so erhält man ein Gemisch aus unterschiedlich langen DNA-Doppelsträngen. In Hinblick auf die nachfolgende „*Shuffling*“-Reaktion sind Fragmente im Größenbereich von 50-400 bp empfehlenswert, wobei der Anteil an kleineren Fragmenten größer sein sollte, um eine große Vielfalt an Ligationsprodukten zu erhalten. Die vier SLH kodierenden Sequenzen wurden auf die gleiche Art und Weise degradiert und miteinander ligiert und dann in einen DNase I-Verdau eingesetzt, um chimäre DNA-Fragmente zu erhalten.

Ein „*Shuffling*“ mit den chimären DNA-Fragmenten allein kann aufgrund des Verlustes der DNA-Enden nicht erfolgen. Deshalb wurden die chimären bzw. hybriden DNA-Fragmente mit den Fragmenten eines DNase I-Verdau einer SLH kodierenden Sequenz gemischt. Das Verhältnis sollte dabei mehr auf Seiten der DNA-Fragmente liegen, die lediglich durch DNase I präpariert worden sind.

Ansonsten gelten die gleichen „*Shuffling*“ Bedingungen.

### 3.3 Einschätzung des Mutageneseerfolges

Für die Bestimmung der Mutationsrate wurden Sequenzen aus dem 4. „*Ribosome Display*“ Zyklus kloniert und sequenziert. 120 DNA-Sequenzen wurden zur Bestimmung der Fehlerrate herangezogen. Sowohl Deletionen, Insertionen als auch Transitionen und Transversionen traten auf. Die erhaltene durchschnittliche Mutageneserate von 1,75 % ist mehr als doppelt so groß, wie die von Stemmer (1994) für das „*Shuffling*“ berichtet von 0,7 %.

Die tatsächliche Mutationsrate vor dem „*Ribosome Display*“ kann nur abgeschätzt werden. Während der Durchführung des „*Ribosome Displays*“ wurden 240 PCR-Zyklen durchlaufen. Die *Tth*-Polymerase hat eine Fehlerrate von  $30 \times 10^{-6}$  (Frey *et al.*, 1995), damit ist sie fast doppelt so hoch wie die der *Taq*-Polymerase von  $18 \times 10^{-6}$ . Die *Tth*-Polymerase besitzt damit eine Mutationsrate von 0,003 %, d.h. 1 Fehler auf 33333 Nukleotide. Das bedeutet, wenn die Daten nur auf die 120 Sequenzen bezogen werden, die jeweils 240 PCR Zyklen durchlaufen haben und eine Größe von 776 bp haben, dass sich (240 Zyklen x 776 bp x 120 Sequenzen)  $2,23 \times 10^7$  bp =  $4,47 \times 10^7$  nt ergeben. Wenn die *Tth*-Polymerase 1 Fehler auf 33333 nt einbaut, so sind dies 1341 Fehler nur verursacht durch die *Tth*-Polymerase. Eine Mutationsrate von 1,75 % bedeutet, dass 1,75 Nukleotide/100 Nukleotide falsch eingebaut wurden (ein Strang wurde sequenziert). Auf die gesamten eingebauten Nukleotide bezogen, bedeutet dies einen Fehleinbau von 782208 Nukleotiden. Das bedeutet, dass 0,17 % durch die *Tth*-Polymerase verursacht wurden, die Mutationsrate vor dem „*Ribosome Display*“ also ca. 1,6 % betrug.

Die ermittelte Mutationsrate von durchschnittlich 1,6 % führt zu 4,1 veränderten Aminosäuren (1,6 x 776 nt/100 nt). Aufgrund der Degeneration des genetischen Codes reduziert sich die Anzahl an veränderten Aminosäuren auf eine unbestimmte Zahl. Die Auswirkungen auf das Protein können verschieden sein.

Üblicherweise bewirken Mutationen einen negativen Effekt wie beispielsweise einen Verlust an Aktivität bzw. Affinität (Loss of Function).

Da dies generell gilt, sinkt pro Mutation auch die Wahrscheinlichkeit, eine Verbesserung der Aktivität und Affinität (Gain of Function) zu erhalten. Je höher die Anzahl an Mutationen ist, desto unwahrscheinlicher wird es ein verbessertes Molekül zu finden, desto höher muss die Anzahl an Molekülen sein, um die wenigen verbesserten Proteine zu finden. Die Leistungsfähigkeit des Screening-Systems beschränkt somit die sinnvolle anwendbare Mutageneserate. Bei dem hier verwandten Screening-System handelt es sich um das „*Ribosome Display*“.

#### 4 Screening-Systeme bzw. Selektionsmethoden

Bei den Screening-Systemen bzw. Selektionsmethoden können verschiedene Funktionsweisen unterschieden werden. Basieren sie auf katalytischen Eigenschaften der Moleküle, wird meist von einem Assay gesprochen, werden Interaktionen von mindestens zwei Molekülen detektiert, wird von einem Display gesprochen. Bei den Display-Systemen handelt es sich um Selektionssysteme, die aus einer hohen Anzahl von Mitgliedern die Moleküle mit den gewünschten Bindungseigenschaften anreichern.

**Tabelle 19: Vergleich - Selektionsmethoden und ihre max. Anzahl an Bibliotheksmitgliedern**

Screen/Selektion	Screening Kapazität
Cell-based assay	10 <sup>3</sup>
Surface Display	10 <sup>6</sup>
Phage Display	10 <sup>9</sup>
Ribosome Display	10 <sup>12</sup>

Das „*Ribosome Display*“, ein zellfreies System, abgeleitet vom „*Polysome Display*“ (Mattheakis *et al.*, 1994) wurde erstmals von Hanes und Plückthun 1997 vorgestellt. Es dient der *in vitro* Selektion von Proteinen. He und Taussig zeigten, dass das „*Ribosome Display*“ auch unter Verwendung eines eukaryotischen Translationsansatzes durchführbar ist. Folgearbeiten beider Gruppen demonstrierten die Funktionalität der Methode zur Selektion einzelsträngiger Antikörper, welche zuvor aus immunisierten Mäusen gewonnen wurden (Hanes *et al.*, 1998; He *et al.*, 1999). Das mRNA-Display ist ebenfalls ein reines *in vitro* Selektionssystem, das auf der zellfreien Translation beruht.

Hier werden die Mitglieder der Proteinbibliothek von den kodierenden mRNAs präsentiert, was durch kovalente Verknüpfung von mRNA und Protein mittels Puromycin, welches vor der Translation über einen Linker an das 3'-Ende der mRNA angefügt wurde, erreicht wird. Die Vorteile dieses Ansatzes gegenüber dem „*Ribosome Display*“ liegen in der Stabilität der zu selektierenden mRNA-Protein-Fusionsmolekülen. Die Selektion von Proteinen ist eine Erweiterung zur klassischen SELEX (Systematic Evolution of Ligands by EXponential Enrichment), die von Tuerk und Gold 1990 eingeführt wurde. Mit dieser Methode ist es möglich, Nukleinsäuren (DNA und RNA) aus einem „*Pool*“ randomisierter Oligonukleotide exponentiell anzureichern. Selektierte Nukleinsäure-Liganden werden als Aptamere bezeichnet. Die hohen Affinitäten werden von Sekundärstrukturen zu der ausgewählten Zielstruktur vermittelt.

Die klassische SELEX ist gegenüber dem „*Ribosome Display*“ relativ einfach und weniger anfällig von ihren einzelnen Parametern. Das „*Ribosome Display*“ beruht auf einem biologischen Lysat, so dass RNasen und Proteasen unvermeidbar sind. Hinzu kommt die eigentliche Translationsreaktion, von deren Effizienz der Selektionserfolg direkt abhängt, und im Anschluß daran die Bindung der PRM-Komplexe über den Zeitraum der Affinitätschromatographie. Die durchaus schwierige Prozedur ist allerdings unabhängig von Toxizität der entstehenden Proteine, Wachstum und anderen zelltypischen Eigenschaften.

#### **4.1 Selektion kohlenhydratbindender Proteine mittels „*Ribosome Display*“**

Das hier verwandte „*Ribosome Display*“ wurde von Lamla und Erdmann 2001 etabliert. In Anlehnung an diese Arbeit wurden die Parameter für das verwandte „*Ribosome Display*“ abgeleitet. Wichtigste Veränderungen waren die Pufferzusammensetzungen, vor allem die des Bindungs- und Waschpuffers. Dieser sollte kein Heparin enthalten, wenn das Zielmolekül Heparin ist. Heparin wird als RNase- bzw. Protease-Inhibitor eingesetzt und verhindert unspezifische Wechselwirkungen der Ribosomen an die stationäre Phase. Des Weiteren wurde Calcium in physiologischer Konzentration zugesetzt, da viele Lektine als Kofaktor für ihre Bindung Calcium benötigen. Aufgrund dieser Tatsache und der Tatsache, dass die Dissoziation bzw. Assoziation der Ribosomenuntereinheiten über die Magnesiumkonzentration reguliert werden kann, wurde dem Elutionspuffer EDTA beigemischt.

Der Verlauf der Selektion ist in III 9.4, Abb. 31 dargestellt.

Die Selektion fand von Anfang an im „Batch“-Verfahren statt, wobei vor dem Beginn der 1. Selektionsrunde eine Inkubation der PRM-Komplexe mit dem Trägermaterial, Sepharose 6B stattfand, um mögliche Bindungen an diesem Material ausschließen zu können. Die Selektion wurde nur mit einer der vier hergestellten Bibliotheken durchgeführt, der Bsp-Bibliothek, basierend auf der SLH kodierenden Sequenz aus *Bacillus sphaericus*, da gegen vier Zielmoleküle gleichzeitig selektiert wurde. Nach vier Selektionszyklen wurde keine weitere Anreicherung mehr erreicht. Damit liegen die Anreicherung gegen die vier Kohlenhydrate Heparin, Streptomycin, N-Acetylglucosamin und Sialyl-Lewis X im Durchschnitt der durchgeführten Selektionszyklen. Mattheakis führte 1994 ein „*Ribosome Display*“ gegen Dynorphin B durch, bei dem er nach fünf Runden verschiedene bindende Peptide fand. Gerusk erreichte nach vier Runden eine Selektion gegen einen Tumormarker (PSA – prostate-specific antigen; Gerusk *et al.*, 1997), Lamla nach sieben Selektionsrunden gegen Streptavidin (Lamla & Erdmann, 2003) und nur drei Runden waren nötig, um gegen die GCN4 Leucin Zipper eine Anreicherung zu erreichen (Hanes *et al.*, 1998).

Die Selektion gegen Heparin mit der Bsp-Bibliothek wurde nach der 5. Runde beendet, da keine weitere Anreicherung mehr möglich war. Mit 18 % eluierter <sup>35</sup>S-mRNA wurde die größte Selektivität im Vergleich zu den anderen vier Molekülen erreicht, gefolgt von der Streptomycin-Selektion (10 %), N-Acetylglucosamin (7 %) und Sialyl-Lewis X (4,7 %). Die unterschiedlichen Werte könnten wie folgt interpretiert werden.

Sicherlich besitzen alle Kohlenhydrate eine eigene Struktur, gerade Heparin, was einen Größenbereich zwischen 5-40 kDa einnimmt und das eine der stärksten Säuren der Natur darstellt. Dennoch sollte eine mögliche Anreicherung gegen Glucosamin in Betracht gezogen werden. Heparin ähnelt den sekundären Zellwandpolymeren in Ladung und Größe, könnte so die stärksten Wechselwirkungen zu den veränderten SLH-Proteinen ausbilden. Streptomycin, ein Trisaccharid und Sialyl-Lewis X, ein Tetrasaccharid besitzen ebenfalls Glucosamin, das verschieden modifiziert ist. Das Glucosamin von Streptomycin ist methyliert, während das der Sialyl-Lewis X acetyliert ist und als Verzweigungspunkt dient und damit einer sterischen Abschirmung unterliegt. Dies könnte der Grund für eine geringere Anreicherung sein, wenn sich die Annahme bestätigt, dass die Selektion indirekt auf das Glucosamin gerichtet ist. Dies würde auch erklären, warum die erreichte Anreicherung gegen N-Acetylglucosamin zwischen Streptomycin und Sialyl-Lewis X liegt.



## 4.2 Sequenzvergleiche der selektierten Moleküle

Die 120 Plasmide die sequenziert wurden, wurden mit Hilfe der Programme pDraw und BioEdit untersucht.

	..... .....	..... .....	..... .....	..... .....	..... .....
	10	20	30	40	50
1.....	ATGGTGTATG	GTACCTCTAG	GGCCGGGTTT	T-CCGATGTT	GCAGGTATCG
2.....	ATGGTG-ATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTT	GCAGGTAACG
3.....	ATGGTG-ATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTT	GCAGGTAACG
4.....	ATGGTG-ATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTT	GCAGGTAACG
5.....	ATGGTG-ATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTT	GCAGGTAACG
6.....	ATGGTG-ATG	GTACCTCTAG	GGCCGGGTTT	T-CCGATGTT	GCAGGTAACG
7.....	ATGGTGTATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTT	GCAGGCAACG
8.....	ATGGTG-ATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTT	GCAGGTAACG
9.....	ATGGTG-ATG	GTACCTCTAG	AGTCGGGTTT	T-CCGGTGCT	GCAGGTAACG
10.....	ATGGTG-ATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTT	GCAGGTAACG
11.....	ATGGTG-ATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTT	GCAGGTAAC-
12.....	ATGGTGTATG	GTACCCCTAG	AGCCGGGTTT	T-CCGATGTT	GCAGGTAACG
13.....	ATGGTG-ATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTT	GCAGGTAACG
14.....	ATGGTG-ATG	GTACCTCTAG	TGCCGGGTTT	T-CCGATGTT	GCAGGTAACG
15.....	ATGGTG-ATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTC	GCAGGTAACG
16.....	ATGGTG-ATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTT	GCAGGTAACG
17.....	ATGGTGTATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTT	GCAGGTAACG
18.....	-----	-----	-----	-----	-----
19.....	ATGGTGTATG	GTCCCTC-AG	AGCCGGGTTT	T-CCGATGTT	GCAGGTAACG
20.....	ATGGTG-ATG	GTACCCCTAG	AGCCGGGTTT	T-CCGATGA-	GCAGGTAACG
21.....	ATGGTGTATG	GTACCTCTAG	AGCCGGTTTT	T-CCGATGTT	GCAGGTAACG
22.....	ATGGTG-ATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTC	GCAGGTAACG
23.....	ATGGTG-ATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTT	GCAGGTAGCG
24.....	ATGGTG-ATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTC	GCAGGTAACG
25.....	ATGGTGTATG	GTACCTCTAG	ATCCGGGTTT	T-CCGATGTT	GCAGGTAACG
26.....	ATGGTG-ATG	GTACCTCTAG	AGCCGGGCTT	T-CCGATGTT	GCAGGTAACG
27.....	ATGGTGTATG	ATACCTC-AG	AGCAGGGTTC	T-CTGATGTT	GCAGGTAACT
28.....	ATGGTGTATG	GTCCCTC-AG	AGCCGGGTTT	T-CCGATGTT	GCAGGTAACG
29.....	ATGGTG-ATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTC	GCAGGTAACG
30.....	ATGGTG-ATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTT	GCAGGTAACG
31.....	ATGGTG-ATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTT	GCTGGTAACG
32.....	ATGGTG-ATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTT	GCAGGTAACG
34.....	ATGGTGTATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTT	GCAGGTAACG
35.....	ATGGTGTATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTT	GCAGGTAACG
36.....	ATGGTG-ATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATATT	GCAGGTAACG
37.....	ATGGTG-ATG	GTACCTCTAG	AGCCGGTTTC	T-CCGATGTT	GCAGGTAACG
38.....	ATGGTGTATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTT	GCAGGTAACG
39.....	ATGGTG-GTG	GTACCTCTAG	ATC-GGGTTT	T-CAGATGTC	GCAGGTAACG
40.....	ATGGTG-ATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTT	GCAGGTAACG
41.....	ATGGTG-ATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTT	GCAGGTAACG
42.....	ATGGTG-ATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTT	GCAGGTAACA
43.....	ATGGTG-ATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTT	GCAGGTAACG
44.....	ATGGTGTATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTT	GCAGGTAACG
45.....	ATGGTGTATG	GTACCTCTAG	ACCCGGGTTT	T-CCGATGTT	GCAGGTAACG
46.....	ATGGTG-ATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTT	GCAGGTAACG
47.....	ATGGTGTATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTT	GCTGGTAACG
48.....	ATGGTGTATG	GTACCTCTAG	ACCCGGGTTT	T-CCGATGTT	GCAGGTAACG
49.....	ATGGTG-ATG	GTACCTCTAG	AGCCGGGTTT	T-CCGATGTT	GCGGGTAACG
50.....	ATGGTGTATG	GTACCTCTAG	AGCCGGGTTT	TTCCGATGTT	GCAGGTAACG
<b>Bsp orig</b>	-----	-----	-GCCGGGTTT	T-CCGATGTT	GCGGGTAACG

Abb. 37: 5'-Ende der DNA von 50 Sequenzen nach der 4. Selektionsrunde

---



---

## IV DISKUSSION

---



---

Das 5'-Ende der DNA bzw. der N-terminale Bereich von 50 Sequenzen sind in den Abb. 37/38 dargestellt.

	..... .....	..... .....	..... .....	..... .....	..... .....
	10	20	30	40	50
1	MVYGTSRAGF	SDVAGIDHEV	AINALADAGI	INGYADGTFK	PNN*
2	MVMVPLEPGF	PMLQVTTTK*			
3	MVMVPLEPGF	PMLQVTTTK*			
4	MVMVPLEPGF	RMLQVTTTK*			
5	MVMVPLEPGF	PVLQVTTTK*			
6	MVMVPLGPGF	PMLQVTTTK*			
7	MVYGTSRAGF	SDVAGNDHEV	AINALADGGI	INGYADGTFK	PNQTINPGQV
8	MVMVPLEPGF	PMLQVTTTK*			
9	MVMVPLESGF	RVLQVTTTK*			
10	MVMVPLEPGF	PMLQVTTTR*			
11	MVMVPLEPGF	PMLQVTPRSS	NQRAC*		
12	MVYGTPRAGF	SDVAGNDHEV	AINALADAGI	INGYADGTFK	PNQTINRGQV
13	MVMVPLEPGF	RMLQVTTTK*			
14	MVMVPLVPGF	RMLQVTTTK*			
15	MVMVPLEPGF	PMSQVTTTK*			
16	MVMVPLEPGF	PMLQVTTTR*			
17	MVYGTSRAGF	SDVAGNDHEV	AINALADAGI	LNGYAEGTFK	PNQTINRGQV
18	XDXKXXFXGF	PGTXXXXXXXX	XXXXXXXXXAGG	FXGXXXQFXX	FPXXXSXAKA
19	MVYGPSEPGF	RMLQVTTTRH	QSTHLLMQVS	SMDTLTAHPN	QTKQLTAVK*
20	MVMVPLEPGF	PMSR*			
21	MVYGTSRAGF	SDVAGNDHEV	AISAPADAGI	INGYADGTFK	PNQTINRGQV
22	MVMVPLEPGF	PMSQVTTTK*			
23	MVMVPLEPGF	PMLQVATTK*			
24	MVMVPLEPGF	RMSQVTTTK*			
25	MVYGTSRSGF	SDVAGNDHEV	AINALADAGI	INEYADGTFK	PNQTINRGQV
26	MVMVPLEPGF	PMLQVTTTK*			
27	MVYDTSEQGS	LMLQVTTTK*			
28	MVYGPSEPGF	RMLQVTTTRH	QSTHLLMQVS	SMDTLTAHPN	QTKQLTAVK*
29	MVMVPLEPGF	PMSQVTTTK*			
30	MVMVPLEPGF	PMLQVTTTK*			
31	MVMVPLEPGF	RMLLVTTTK*			
32	MVMVPLEPGF	PMLQVTTTK*			
34	MVYGTSRAGF	SDVAGNDHEV	AINALADAGI	INGYADATFK	PNQTINRGQV
35	MVYGTSRAGF	SDVAGNDHEV	AINALADAGI	INGYADATFK	PNQTINRGQV
36	MVMVPLEPGF	PILQVTTTK*			
37	MVMVPLEPVS	PMLQVTTTK*			
38	MCMVPLEPGF	PMLQVTTTK*			
39	MVVVPLDRVF	RCRR*			
40	MVMVPLEPGF	PMLQVTTTK*			
41	MVMVPLEPGF	PMLQVTTTK*			
42	MVMVPLEPGF	PMLQVTTTK*			
43	MVMVPLEPGF	PMLQVTTTK*			
44	MVYGTSRAGF	SDVAGNDHGV	AINALADAGI	INGYADGTFK	PNQTFNRGRV
45	MVYGTSRPGF	SDVAGNDHEV	AINALADAGI	INGYADGTFK	PNQTINRGQV
46	MVMVPLEPGF	PMLQVTTTK*			
47	MVYGTSRAGF	SDVAGNDHEV	AISALADADI	INGYADGTFK	PIQTINRGQV
48	MVYGTSRPGF	SDVAGNDHEV	AINALADAGI	INGYADGTFK	PNQTINRGQV
49	MVMVPLEPGF	PMLRVTTTK*			
50	MVYGTSRAGF	FRCCR*			
<b>Bsp orig</b>	-----GF	SDVAGNDHEV	AINALADAGI	INGYADGTFK	PNQTINRGQV

Abb. 38: Darstellung des N-terminalen Bereiches nach Translation der Sequenzen 1-50

Wie in Abb. 38 auffällt sind hauptsächlich Sequenzen selektiert wurden, die oft nur 19 Aminosäure lang sind (32 Sequenzen von 120). Daneben finden sich andere Proteine die länger, aber dennoch ein Stoppkodon aufweisen. Der Tabelle 16 in Kapitel III 9.6 ist zu entnehmen, dass es von 120 Sequenzen gerade einmal 16 Sequenzen gibt, die kein Stoppkodon besitzen. Dies ist angesichts der Tatsache, dass ein Fehlen des Stoppkodons zur Ausbildung des PRM-Komplexes unumgänglich ist, nicht nachvollziehbar. Gerade die Proteine mit 19 Aminosäuren dürften eine Ausbildung eines derartigen Komplexes nicht zulassen, da erstens der Ribosomentunnel diesen Sequenzbereich abdeckt, diese Proteine also keine Möglichkeit besitzen, mit den Zielmolekülen zu interagieren und zweitens das Ribosom am Stoppkodon in seine Untereinheiten dissoziiert und so Pheno- und Genotyp nicht miteinander gekoppelt sind.

Diese 19 Aminosäuren langen Sequenzen wurden hauptsächlich durch die Deletion des Primers Exp fow II hervorgerufen, der eine Leserahmenverschiebung erzeugte, die sich bereits in den ersten 60 Nukleotiden auswirkte. Diese konnte nur in wenigen Fällen, durch weitere Mutationen kompensiert werden, so dass 50 % der Sequenzen nach dieser PCR mit diesem Primer so entstanden sind. Nach dem Ende des „*Ribosome Displays*“ gibt es noch 26 % der 19 Aminosäuren langen Proteinen, deren Anzahl so im Verlauf des „*Ribosome Displays*“ minimiert, aber nicht komplett ausselektiert wurden konnte.

Eine Erklärungsmöglichkeit für das Auftreten von Sequenzen mit Stoppkodon ist, dass die <sup>35</sup>S-mRNA selbst mit den Zielmolekülen interagiert und somit ein Aptamer angereichert wurde. Diese Möglichkeit wurde überprüft (III 9.6). Die Ergebnisse sind jedoch sehr indifferent. Während ein Anstieg an eluierter <sup>35</sup>S-mRNA der 5. Runde von Streptomycin und N-Acetylglucosamin gegenüber dem Ausgangspool nachweisbar ist, kann dies für Heparin und Sialyl-Lewis X nicht eindeutig belegt werden. Betrachtet man die prozentualen Werte an eluierter <sup>35</sup>S-mRNA von jeweils 20 Sequenzen/Zielmolekül, so ist generell ein Anstieg zu beobachten. Diese Ergebnisse können nur so verstanden werden, dass eine Bindung von mRNA-Molekülen an die Kohlenhydrate möglich ist. Diese Bindung aber auf bestimmten Sequenzen beruhen muss, da der mRNA-Pool aus der 5. Runde mehr als 20 verschiedene Sequenzen enthält. Der prozentuale Wert ist von der Anzahl vorhandener Moleküle, die Affinität zu den Molekülen zeigen, abhängig. Dies erklärt vielleicht auch, warum die Anreicherungswerte aus dem „*Ribosome Display*“ sich reziprok zu den Werten der putativen Aptamere verhalten. Ist es möglich, dass zwei Selektionen gleichzeitig stattgefunden haben? Die RNA in Konkurrenz zu den Proteinen stand?

Die Selektion gegen Heparin zeigte eine Anreicherung von 18 %, hingegen die mRNA allein nur eine von 2,7 %. Der umgekehrte Fall liegt bei N-Acetylglucosamin vor.

Hier haben wir eine Anreicherung in der 5. Runde von 7 %, der Anteil der mRNA allein liegt bei ca. 11 %, was theoretisch nicht möglich ist, es sei denn, man zieht die Möglichkeit in Betracht, dass zwei Selektionen gleichzeitig stattgefunden haben.

Dennoch erklärt dies nicht, dass der überwiegende Teil der Sequenzen ein Stoppkodon trägt und das an unterschiedlichen Stellen, so dass Proteine unterschiedlicher Größe entstehen konnten.

### 4.3 Diskussion der Bindungsstudien im Zusammenhang mit dem Stoppkodon

Die Ergebnisse der Bindungsstudien zeigten, dass Proteine mit einer Länge von 125, 148 und ca. 200 Aminosäuren (III 9.7) an das Zielmolekül binden können, trotz des Vorhandenseins eines Stoppkodons.

Die Translation der mRNA wird durch viele Ribosomen gleichzeitig durchgeführt. Die Gruppe von Ribosomen, die an ein mRNA-Molekül angelagert ist, nennt man Polyribosom oder Polysom. Die Ribosomen dieser Einheit arbeiten unabhängig voneinander, wobei jedes eine Polypeptidkette synthetisiert. Die höchste Ribosomendichte an einem mRNA-Molekül beträgt etwa ein Ribosom pro 80 Nukleotide (Stryer). Die Faltung der entstehenden Proteine geschieht kotranslational, d.h. während der Translation.

Die Translation beim „*Ribosome Display*“ wurde nach 7,5 min durch Kälte (0-4 °C) gestoppt und die entstandenen PRM-Komplexe durch Magnesium stabilisiert.

Aufgrund dieser Tatsachen kann folgender Fall konstruiert werden. Für eine Bindung zwischen Protein und Kohlenhydrat ist ein Minimalmotiv verantwortlich. Beispiel: ein 125 Aminosäure langes terminiertes Protein wird synthetisiert, so wäre ein Minimalmotiv von ca. 100 Aminosäuren möglich, da das Protein, was durch das Stoppkodon terminiert wird, freigelassen wird, das nächste Protein aber erst ca. 26-30 Aminosäuren entfernt liegt. Dies würde bedeuten, dass eine spätere Terminierung, hervorgerufen durch längere Aminosäureketten, deren Stoppkodon entweder fehlt oder erst nach 200 Aminosäuren auftritt, stärker angereichert werden müssten, da mehrere naszierende Proteine eine kumulative Bindung erwirken. Dies ist tatsächlich der Fall. Die Sequenzierung der 120 Plasmide zeigte mehr Proteine mit einer Länge von ca. 200 Aminosäuren als mit ca. 125 Aminosäuren (VII 7).

Dieses Ergebnis würde mit der Theorie korrelieren, dass die SLH-Domänen den Verankerungspunkt zur bakteriellen Zellwand herstellen. Eine SLH-Domäne umfasst 50-60 Aminosäuren und kann bis zu 3-mal auftreten. Bei dem vorliegenden Konstrukt sind zwei dieser SLH-Motive vorhanden. Zwei SLH-Domänen sollten eine größere Bindung ermöglichen als eine SLH-Domäne. Inwiefern eine Verbesserung der Bindungsstärke durch Einführung von Mutationen erreicht werden konnte, kann nur in weiteren Untersuchungen erfolgen, die den Rahmen dieser Arbeit allerdings gesprengt hätten. Vergleichende Studien zwischen dem unveränderten und den veränderten Proteinen sollten durchgeführt werden. Die Wirkung der Mutationen auf die Affinität konnte durch die Bindungsstudien (III 9.7) eindrucksvoll gezeigt werden. Es konnten Proteine mit unterschiedlichen Affinitäten zu den Zielmolekülen angereichert werden.

Die zusätzlichen Banden in den SDS-Gelen (III 9.7) sollten eventuell auf mögliche Proteolyse der Proteine untersucht werden. Das Ergebnis könnte bei Wiederholung des Experimentes bekräftigt werden, wenn ein leistungsfähigeres S30-Lysat verwendet werden würde. Die Syntheseleistung unter Verwendung des S30-Lysates und eines Plasmides aus der Arbeitsgruppe beschränkt sich auf 75-130 µg/ml, was weit unter den kommerziell erhältlichen Lysaten z.B. der RiNA, Invitrogen von 150-200 µg/ml liegt. Für die in die Bindungsstudien eingesetzten linearen Konstrukte wurden Syntheseraten von 4-13 µg/ml erreicht. Bei leistungstärkeren Systemen könnte ein eventueller Einfluss der Lysat-Komponenten minimiert werden.

## 5 Ausblick

Um ein umfangreiches Bild über den Einfluss der SLH-Domänen bzw. Bindungsstärke der Proteine gewinnen zu können, sollten die ausstehenden drei Bibliotheken (Ea, Bst und Aki) ebenso mit dem „*Ribosome Display*“ untersucht werden.

Die Bindungskonstanten der selektierten Proteine sollten noch bestimmt werden. Dies beinhaltet aufwendige Etablierungen. Vergleichende Sequenzanalysen könnten erfolgen, um ein allgemeingültiges Bindungsmotiv zu finden. Dafür sind allerdings andere Programme nötig, als die hier verwendeten.

Studien bzw. Kreuzuntersuchungen, die die Theorie bestätigen, dass das Glucosamin als Bindungspartner dient, sollten durchgeführt werden.

Außerdem könnte die Verwendung anderer Kohlenhydrate auch noch getestet werden.