

6. Zusammenfassung

Die Entdeckung neuer Leitstrukturen und deren Optimierung sind die grundlegenden Herausforderungen moderner Wirkstoffforschung. Die Anwendung von Methoden zum effektiven *screening* von Substanzmischungen auf biologische Aktivität ist somit der erste Schritt zur Entwicklung neuer Pharmaka. Das *high throughput screening* (HTS) ist in der Lage mehrere zehntausend Verbindungen pro Tag zu untersuchen. Als ergänzende Verfahren etablieren sich zunehmend NMR gestützte Methoden zur Untersuchung von Protein-Ligand-Wechselwirkungen. Auf theoretischer Seite gewinnt das rationale Wirkstoffdesign immer mehr an Bedeutung. Ziel ist es, durch Kombination dieser und anderer Verfahren die Suche nach neuen Wirkstoffen effektiv zu gestalten.

Das E-Selektin, das als Initiator bei der Leukocyten-vermittelten Entzündungsreaktion eine wesentliche Rolle spielt, ist das molekulare Zielprotein dieser Arbeit. Neue Leitstrukturen, die als Antagonisten gezielt diese Immunreaktion unterdrücken können, sind somit von besonderen Interesse.

Ausgangspunkt hierfür war eine zufalls-fucosylierte und zufallsulfatierte Trisaccharid-Bibliothek, in der eine oder mehrere Komponenten eine hohe Bindungsaktivität zum E-Selektin zeigten. Im Rahmen dieser Arbeit wurde dieses komplexe Gemisch von sulfatierten Trisacchariden auf seine E-Selektin Bindungsaktivität mit Hilfe von NMR-spektroskopischen Verfahren untersucht. Parallel zu den experimentellen Untersuchungen wurden verschiedene Verfahren der theoretischen Analyse zur Bindung, zum konformativen Verhalten und zum virtuellen Design der Bestandteile der Substanz-Bibliothek durchgeführt. Eine vergleichende Analyse der virtuell erzeugten Verbindungen mit dem Sialyl-Lewis^x, das den natürlichen Liganden des E-Selektins darstellt, ergab genauere Einblicke in die molekularen Erkennungsprozesse der Liganden.

Das E-Selektin wurde als Chimäre eines IgG Immunglobulins mit der IEC-Domäne

des E-Selektins und somit mit zwei Ligand-Bindungsstellen eingesetzt. Die sulfatierte Trisaccharid-Bibliothek lag als Gemisch von ca. 200 Verbindungen vor. Die Fucosylierung und die Sulfatierung erfolgte an allen Positionen des Lactoseeinheit. Die C1-Position der Lactose ist mit einem aromatischen C₈-*spacer* verknüpft, wodurch diese Position nicht modifiziert wurde. Als typische Instrumente zur Untersuchung von Protein-Ligand-Wechselwirkungen mittels der NMR-Spektroskopie wurden das *transferred* NOE-Verfahren, T₂-Relaxationszeitmessungen und ein neues Verfahren zum *screening* von Substanz-Bibliotheken auf Bindungsaktivität – das STD NMR Verfahren (*saturation transfer difference*) – eingesetzt. Besonders das zuletzt genannte Verfahren erwies sich als besonders erfolgreich. Es gelang, direkt aus der Mischung die bindenden Komponenten zu identifizieren.

Neben Standard TOCSY- und NOESY-Experimenten wurden selektive 1D-TOCSY- und 1D-COSY-Experimente zur Identifizierung der Verknüpfungsvarianten über die spezifischen ¹H-chemischen Verschiebungen der Fucose durchgeführt. Abhängig von der Position der Verknüpfung an der Lactose zeigen besonders die H1- und die H5-Protonen der Fucose charakteristische ¹H-chemische Verschiebungen. Die Analyse der 1D STD NMR-Experimente und des 2D STD TOCSY-Experiments ergab als Hauptbindungspartner Derivate der 6-Fucosyl-lactose und der 2'-Fucosyl-lactose. Spezielle Pulsprogramme für die STD TOCSY NMR-Experimente wurden entwickelt. Die Subtraktion der Spektren erfolgte bereits während der Akquisition. Mit geeigneten Phasencyclen wurde bei den eindimensionalen Experimenten die Differenzbildung nach jedem *scan* durch Drehung der Empfängerphase um 180° vorgenommen. Im STD TOCSY Experiment erfolgte die Subtraktion nach einem kompletten Phasencyclus des jeweiligen F1-Inkrementes.

Hinweise zur Position der Sulfatgruppe wurden bei der 1-2'-Fucosyl-lactose gefunden. Das H3-Proton der Fucose zeigt zumindest für eine Komponente dieses Verknüpfungstyps einen signifikanten tiefelds^{hift}

der ^1H -chemischen Verschiebung. Dieser *shift* kann ursächlich auf die räumliche Nähe einer Sulfatgruppe zurückzuführen sein.

Die T_2 -Relaxationszeiten und die STD TOCSY-Signale zeigen für die Protonen der Fucose und der Protonen der Galactose beider Verknüpfungsvarianten die stärksten Bindungseffekte.

Der genaue Bindungsmodus des natürlichen Liganden sLe^x ist nicht bekannt. Bisher ist es noch nicht gelungen das Sialyl-Lewis^x mit dem E-Selektin zusammen zu kristallisieren. Die Calcium-Abhängigkeit der Bindung und die für die Bindung wichtigen Aminosäuren sind jedoch mit Hilfe von Mutagenese-Experimente aufgeklärt. Auch die Konformation des Liganden im gebundenen Zustand konnte durch *tr*NOE-Messungen bestimmt werden.

Als Grundlage für weitere theoretische Untersuchungen wurde zunächst über ein *docking*-Verfahren (FlexiDock), das den Liganden flexibel und das Protein starr hält, ein Bindungsmodus generiert, der den Konformationsdaten der NMR-Analysen sowie der postulierten Orientierung des Sialyl-Lewis^x in der Bindungstasche entspricht. In diesem Bindungsmodell ist die Fucose über die 3-OH- und die 4-OH-Gruppe am Calcium koordiniert. Die 6-OH-Gruppe der Galactose bildet mit dem Tyr94, Asn105 und dem Glu92 mehrere Wasserstoffbrücken. Eine zusätzliche Wasserstoffbrücke ergibt sich für die 4-OH-Gruppe zum Tyr94. Die Position der negativ geladenen Carboxylgruppe liegt nahe an der Guanidiniumgruppe der Seitenkette vom Arg97. Das Glucosamin und die Seitenkette der Sialinsäure zeigen wenig direkte Kontakte zum E-Selektin.

Entscheidend für diesen Bindungsmodus ist die Position der Fucose am Calcium, die Lage der Carboxylgruppe am Arg97 und das Netzwerk an Wasserstoffbrücken der Galactose. Diese Anordnung wird nur erreicht, wenn der Abstand von ca. 10 Å vom Calcium zum N^{m2}-Atom des Arg97 im Protein, von komplementären Gruppen des Liganden überbrückt wird.

Die NMR-Analysen der sulfatierten Trisaccharid-Bibliothek ergaben als Hauptbindungspartner die 6-Fucosyl-lactose und

die 2'-Fucosyl-lactose. In einem *de novo*-Design Prozeß mit der Software LeapFrog wurden alle möglichen Verknüpfungsvarianten generiert und auf ihre Bindungsaktivität untersucht. In diesem Prozeß wurden die Liganden schrittweise aufgebaut und entsprechend ihrer Bindungsenergie klassifiziert. Zunächst konnte die Bindungsregion der Fucose am Calciumatom definiert werden. Ausgehend von dieser Position wurden alle möglichen Variationen der Fucose mit der Lactose erzeugt. Die anschließende Sulfatierung der Lactose bildete den Abschluß. Betrachtet man die Bindungsenergien der monosulfatierten Strukturen, wurde in Übereinstimmung mit den NMR-Ergebnissen, die 1-6-Fucosyl-lactose als bindungsstärkster Ligand generiert. Die Sulfatgruppe ist bei diesem Ligand mit der 3-Position der Glucose verknüpft. Die 2'-Fucosyl-lactose wurde ebenfalls als 4'-Sulfat mit einer hohen Bindungsenergie erzeugt. Die Konformation dieser Verknüpfungsvariante liefert auch ein Modell für den in den NMR-Spektren gefundenen tieffeld*shift* des H3-Protons der Fucose. Die Sulfatgruppe ist in dieser Anordnung in unmittelbarer räumlicher Nähe des Protons, dessen Resonanz um 0.15 ppm verschoben ist.

Aufgrund der kürzeren Distanzen von der am Calcium koordinierten Fucose zur Sulfatgruppe aller generierten Trisaccharide, unterscheidet sich der Bindungsmodus im Vergleich zum Sialyl-Lewis^x deutlich. Die negativ geladene Gruppe ist nicht, wie im Bindungsmodell des Sialyl-Lewis^x, am Arg97 lokalisiert, sondern ist zwischen dem Lys111 und dem Arg108 positioniert. Offensichtlich kann der Abstand vom Calcium zum Arg97 (ca. 10 Å) nicht von den sulfatierten Trisacchariden überspannt werden.

Da der gebundene Zustand der Liganden nur eine Konformation von vielen Möglichen darstellt, wurden MMC-Simulationen der bindungsaktivsten monosulfatierten Trisaccharide mit dem Programm GEGOP (*geometry of glycoproteins*) durchgeführt. Mit diesem Verfahren ist es möglich, den gesamten Konformationsraum einer chemischen Struktur abzutasten.

Zusätzlich wurde eine Clusteranalyse verwendet, die zu einer mathematisch exakten Ableitung der konformativ bevorzugten Strukturen führt. Die Hauptklassen dieser Analyse stellen dann nicht nur die energetisch günstigsten Konformationen dar, diese Hauptklassen sind durch ihren hohen Populationsanteil auch ein Maß für die tatsächliche Existenz einer Konformation.

Der Vergleich der Konformationen im gebundenen Zustand mit den gemittelten Konformationen der Hauptklassen aus der Clusteranalyse, ergab eine hohe Übereinstimmung. Somit sind die im *De novo*-Design-Prozeß generierten Strukturen energetisch günstige Konformationen und entsprechen den bevorzugten Zuständen.

Besonders deutlich wird dies bei dem 3-Sulfat der 1-6 verknüpften Fucosyl-lactose. Die LeapFrog-Struktur zeigt eine Konformation, die der gemittelten Konformation einer der Hauptklassen entspricht und das globale Energieminimum darstellt.

Mit dieser Arbeit ist es gelungen, ein NMR-gestütztes *screening* Verfahren zu etablieren mit dem die bindungsaktivsten Komponenten aus einer extrem komplexen Mischung sehr ähnlicher sulfatierter Trisaccharide identifiziert wurden. Diese Strukturen, die im Gegensatz zum natürlichen Liganden Sialyl-Lewis^x eine andere Verknüfungsposition der Fucose besitzen – statt der 3-Position, ist die Fucose dieser Liganden mit der 6- oder der 2'-Position verknüpft - stellen in Übereinstimmung mit theoretischen Methoden des Wirkstoffdesigns neue Leitstrukturen als Antagonisten des E-Selektins dar.

6.1 Summary

The discovery and optimization of new lead substances is a big challenge in modern drug research. The application of effective screening methods for testing biological activity of compound mixtures is the initial step in the development of new pharmaceuticals. High throughput screening (HTS) is able to test several 10000 compounds per day. However, NMR-spectroscopic tools are alternatives for the investigation of protein ligand interactions and are increasingly used for screening. Additionally, molecular modeling can be a useful technique for lead generation and lead optimization of biologically active compounds. Currently, computer docking studies are often used to simulate inhibitors or antagonists.

The biological target in this thesis is E-selectin which mediates the early steps of the recruitment of leukocytes from the blood stream during inflammation. New lead structures which suppress this immune response are therefore of special interest.

To find new inhibitors we tried to identify the bioactive component from a library which has high affinity in a bioassay. The library is composed of substances resulting from a random sulfatation followed by fucosylation of lactose. This complex mixture was investigated by NMR-spectroscopy and theoretical analysis, i.e. docking experiments, MMC-simulations, and methods for the virtual design and optimization of saccharide ligands. A comparative study of the virtually generated ligands and the natural ligand sialyl Lewis^x resulted in a more precise view of the binding properties.

The E-selectin used here was an IgG-chimera with the E-selectin lec-domain with two binding sites. The sulfated trisaccharide library contained about 200 compounds. The fucosylation and sulfatation was performed randomly at each position of the lactose unit. The C1-position of the lactose is joined to an aromatic C₈-spacer. The protein ligand interactions were studied by NMR-experiments such as transferred NOE (*t*NOE) and T₂-relaxation time

measurement. Additionally the binding activity of the mixture was tested with a new NMR-experiment. This technique is called STD-NMR (saturation transfer difference) and can be used for screening compound libraries for binding activity to proteins.

The chemical shifts of the various compounds with different fucose linkage patterns were assigned by TOCSY-, NOESY-, and selective COSY-, and TOCSY-experiments. Especially the H1- and H2-proton shifts of the fucose are indicative of the linkage position to the lactose. By 1D and 2D STD NMR-spectroscopy two binding ligands with 1-6 and 1-2' fucosylated lactose cores were characterized. A special TOCSY STD pulse program was developed. Subtraction was performed prior to Fourier transformation by phase cycling. In the case of the 1D STD NMR-experiment, subtraction occurs after every scan by inverting the receiver phase by 180°. In the 2D STD TOCSY NMR-experiment the subtraction took place after a complete phase cycle of the respective F1-increment.

Indications of the position of the sulfate group were found for the 1-2'-fucosyl-lactose. The H3-proton of the fucose shows at least for one compound a significant downfield shift. This shift can originate from a sulfate group which is in proximity to the H3-proton.

The T₂-relaxation time and the STD TOCSY experiments indicate that the fucose and galactose protons have the strongest interactions with the protein.

The exact binding mode of the native ligand sialyl Lewis^x has still not been determined unambiguously. Up to now, no crystallographic data of a complex of E-selectin with sialyl Lewis^x is available. But in all C-lectins the carbohydrate is complexed to the protein by forming coordination bonds with a conserved calcium, as well as by hydrogen bonding with acid and amide sidechains that also coordinate to the calcium. The essential amino acids involved in carbohydrate binding could also be determined by mutagenesis experiments. Additionally, the conformation of the sialyl Lewis^x tetrasaccharide bound to E-selectin was determined from transfer NOE

(^1H -NOE) data followed by a distance-geometry analysis.

The basis for further theoretical studies was generated with a docking procedure (FlexiDock), which keeps the protein fixed and the ligand flexible. The resulting structure is in good agreement to the NMR-analysis and shows the predicted orientation of sialyl Lewis^x in the binding pocket. In this orientation, the fucose is coordinated with the 3-hydroxy group and the 4-hydroxy group to the calcium atom. The 6-hydroxy group of the galactose formed multiple hydrogen bonds to Tyr94, Asn105 and Glu92. Additionally, a hydrogen bond was found between the 4-hydroxy group and Tyr94. The negatively charged carboxy group was in proximity to the guanidino group of Arg97. The glucosamine and the side chain of the sialic acid were not in close contact to the protein.

The crucial features of this binding mode is the coordination of the fucose to the calcium, the position of the carboxy group at Arg97, and the network of hydrogen bonds of the galactose. This orientation can only be reached, if a ligand can bridge the distance of about 10 Å between the calcium and the N^η-atom of the Arg97.

In a *de novo* design process with the program LeapFrog, all possible fucose linkage combinations to the lactose were generated. The ligands were built stepwise and classified by their binding energy. First, the binding region of the fucose to the calcium could be defined. Starting from this position, all possible structures of the fucose with the lactose were generated *in silico*. Subsequently, the lactose core was sulfated randomly. In good agreement with the NMR data, the 1-6-fucosyl lactose is the best ligand. The sulfate group of this computer generated structure was in the 3-position. The 1-2'-fucosyl-4'-O-sulfo lactose was also found as a ligand with a high binding energy. The conformation of this compound results in a good model for the downfield shift (1.15 ppm) of the proton resonance of the H3 fucosyl proton. In this configuration, the sulfate group is situated in the vicinity of the H3-proton.

Because of the smaller distances of the fucose to the sulfate group, the binding mode of all sulfated trisaccharides differs from the one of the tetrasaccharid sialyl Lewis^x, where the negatively charged group is located at the Arg97. In all sulfated trisaccharides, the negatively charged sulfate group is positioned between Lys111 and Arg108. Apparently, the distance between the calcium and the Arg97 cannot be bridged by the virtually generated sulfated trisaccharides.

The bound conformation represents only one possible conformation, therefore a MMC-simulation of the most active mono sulfated trisaccharides with the program GEGOP (geometry of glycoproteins) was carried out. The method was chosen to ensure that all sterically possible conformations were considered. Additionally a cluster analysis was performed to create a mathematically exact description of the preferred conformations. The major classes of this analysis represent energetically favorable conformations.

Comparison of the structures in the bound state with the averaged structures from the major classes of the cluster analysis resulted in a good agreement. Therefore, the virtually generated structures are energetically reasonable and correspond to the preferred states of the molecules.

The LeapFrog structure of the 1-6-fucosyl-3-O-sulfo-lactose shows a conformation which is equivalent to one of the major groups of the cluster analysis.

In this thesis, a NMR-based powerful screening tool for investigating a complex mixture composed of very similar sulfated trisaccharides was established. The bioactive compounds could be identified and characterized. These structures are different from the natural ligand sialyl Lewis^x. The linkage of the fucose is in the 6- or in the 2'-position but not in the 3-position, which is found in sialyl Lewis^x. In good agreement with theoretical methods, these structures are new leads as antagonists to E-selectin.