Surface Plasmon Resonance Applications in Drug Discovery

with an Emphasis on Small Molecule and Low Affinity Systems

Inauguraldissertation
zur Erlangung der Würde eines Doktors der Philosophie
vorgelegt der Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Daniel Ricklin
aus Zürich, Schweiz

Referent: Prof. Dr. Beat Ernst
Korreferent: Prof. Dr. Ueli Aebi

Basel, Juni 2005

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von

Prof. Dr. Beat Ernst, Institut für Molekulare Pharmazie, Universität Basel
Prof. Dr. Ueli Aebi, M. E. Müller Institut, Biozentrum, Universität Basel

Basel, den 5. Juli 2005

Prof. Dr. Hans-Jakob Wirz
Dekan

A discovery is said to be an accident
meeting a prepared mind…
(Albert Szent-Györgyi, 1893-1986)

Meinen Eltern
Johanna und Albert

Acknowledgement

This thesis was performed at the Institute of Molecular Pharmacy of the University of Basel under the supervision of Prof. Dr. Beat Ernst, and was generously supported by the Swiss National Science Foundation.

First and foremost, I thank Prof. Dr. Beat Ernst for his great scientific support, the generous and modern infrastructure, and his constructive and fruitful discussions. With its multidisciplinary and international atmosphere, the institute created a motivating, challenging and encouraging environment. The integration of scientific seminars, project meetings, teaching opportunities and supervision of diploma theses was very stimulating for the development of skills beyond pure science.

I sincerely thank Prof. Dr. Ueli Aebi for being the co-referee of my thesis.

My deep and special thanks are also going to all the former and present colleagues at the institute, who created a very comfortable working atmosphere and provided me with proteins, analytes and good ideas. Daniel Strasser and Steven Knecht helped me forming a 'Biacore team' and gave me many new inputs. Caroline Bellac with her diploma thesis and Svenja Landweer in a 'summer project' were a tremendous support for the experimental part of this thesis. Rita Born, Karin Johansson, Daniela Stokmaier, Andrea Frey, Claudia Riva, Oleg Khorev, and Daniel Kreyenbühl were not only responsible for many of the biological and chemical work in the asialoglycoprotein-receptor project, but also supported me with critical and fruitful discussions during project meetings. I also like to thank Dr. Said Rabbani, Dr. Oliver Schwardt, Dr. Brian Cutting, Gabriela Pernter, and Bea Wagner, for their administrative and technical help as well as Matthias Studer and Andreas Stöckli for their computer support.

Dr. Angelo Vedani and Dr. Markus Lill from Biographics Laboratories in Basel helped me in many aspects of molecular modeling, and Prof. Dr. Paul Jenö and Thierry Mini from the Biocenter of the University of Basel performed the mass spectrometric analysis of asialoglycoprotein. Prof. Thomas Peters, Dr. Hanne Peters, Dr. Thomas Weimar, Thies Köhli, and Dr. Lars Herfurth from the Medical University of Lübeck, Germany greatly facilitated my entrance in the field of Biacore analysis. I would like to thank them for the collaboration in the GSLA-2 project, as well as Dr. John Magnani from GlycoTech Inc. in Rockville, USA, for his donations of the diagnostic antibody. I also want to thank Prof. Dr. Alex Eberle from the Department of Research of the University Hospital Basel for his collaboration in the hexahistidine project.

Hence performing a PhD thesis is not solely about science; it needs help and support from many other sides. Therefore, I primarily want to thank my parents, who not only supported me financially and morally throughout the whole course of my educational career, but also let me feel their love and care. My special thanks are going to Salome Lichtsteiner, who closely accompanied me during this thesis, shared my ups and down, and always understood in motivating me to carry on. Finally, I would like to thank the many friends and relatives, who created the social ground and network for this work.

Summary

Surface plasmon resonance (SPR) technology evolved into a key technology for the characterization of biomolecular interactions, and is integrated in many stages of the drug discovery process. Despite recent developments in the area of instrument sensitivity and data processing, working with small molecules and low affinity interactions still remains a major challenge. The aim of this thesis was therefore to evaluate and develop different methods for the accurate and reliable determination of thermodynamic and kinetic information of such interaction systems.

Through participation in the international ABRF-MIRG'02 study, the instrument used for this thesis was validated for small molecular analyses. The results obtained for a small sulfonamide analyte binding to bovine carbonic anhydrase II were very close to the study average and corresponded well with solution-based methods. Screening experiments with human serum albumin and a set of known plasma protein binders further confirmed the effectiveness of SPR for small molecule assays. However, the albumin assay also revealed some limitations; while neutral and cationic drugs generated very reproducible K_D values, the deviations were usually larger for free acids. Some compounds like diazepam or L-tryptophan showed a more complex binding behavior. Most of these atypical signal effects could be attributed to ligand- or *pH*-induced structural changes of albumin, which are well described in literature. Finally, a new immobilization method for human serum albumin was developed by targeting its single free cysteine residue for a reversible coupling to the sensor chip.

The interaction of monovalent carbohydrates with their protein targets is one of the most prominent examples of small molecule/low affinity systems. They play an important role in many biological processes from cellular recognition to infection diseases. In order to characterize such carbohydrate-protein interactions, a diagnostic monoclonal antibody (GSLA-2) directed against a carbohydrate epitope was investigated using a combination of SPR and nuclear magnetic resonance. By screening the tetrasaccharide antigen sialyl Lewis^a and a set of structurally related compounds, the thermodynamic and kinetic binding properties as well as the recognition pattern could be successfully described. With a K_D in the low micromolar range and fast kinetic on- ($\sim 10^4$ $M^{-1}s^{-1}$) and off-rates (>0.1 s^{-1}), the interaction correlated very well with earlier reports about carbohydrate-protein interactions. Truncation of the antibody to its antigen-binding parts led to a significant increase in binding activity and reduced non-specific binding.

The human hepatic asialoglycoprotein receptor served as a more complex example of carbohydrate-binding proteins. This C-type lectin is involved in the clearance of desialylated glycoproteins from blood and is regarded as an important target for selective delivery of genes and drugs to the liver. After expression of the carbohydrate recognition domain in *E.coli*, the lectin could be successfully purified using a combination of different chromatographic steps and was immobilized on a SPR sensor chip. Binding of the physiological glycoprotein ligands asialofetuin and

asialoorosomucoid was characterized by a polyvalent interaction pattern with slow dissociation rates and sub-nanomolar K_D values. In contrast, monovalent sugars like galactose or *N*-acetyl galactosamine showed fast kinetics and affinities in the micro- to millimolar range. In addition, the processed SPR signals of all small sugar analytes had a negative sign and had to be mirrored before analysis. The negative binding signals were clearly concentration-dependent and could be fitted to a single binding site model. The resulting affinity values were validated by a competitive ELISA method and with literature values. Furthermore, the interaction was found to be strongly calcium- and *pH*-dependent, as it is reported for the receptor. Ligand-induced conformational changes or interactions of the immobilized lectin with the dextran matrix of the sensor chip were evaluated as the most likely explanation of the negative SPR signals. Whether this is an isolated behavior of the asialoglycoprotein receptor or whether these observations could be applied to other lectins with shallow, surface-accessible binding sites has to be investigated in more detail.

A combined analysis of all protein studies performed in this thesis clearly reveals the benefits and limitations of SPR technology for the analysis of small molecules and low affinity interactions. The label-free detection and the simultaneous evaluation of both thermodynamic and kinetic parameters allow a rapid and deep insight into molecular binding mechanisms, even at the limit of detectability. Careful assay design and proper data processing are a prerequisite for the analysis of small molecules, since even small signal deviations might significantly influence the binding constants. The studies of human serum albumin and the asialoglycoprotein also revealed, that SPR detection cannot be solely regarded as a mass detector. Structural changes of the immobilized proteins or matrix-effect could also influence the detected SPR signal and should always be considered in the planning and evaluation of experiments.

In a small pilot project, the molecular mechanism of the interaction between the hexahistidine tag, which is widely used for the purification of recombinant proteins, and immobilized nickel surfaces was investigated using SPR. By injecting a series of oligohistidine peptides (His2-His10), the influence of the number of histidine residues to the binding behavior could be evaluated. As expected, the His6 peptide revealed the best compromise between rebinding and entropic effects, resulting in the lowest K_D of the series (34 nM). The distance between the two simultaneously binding imidazole rings was also found to play an important role for the binding strength, as is could be shown by screening a series of His_2Ala_4 peptides.

Abbreviations

ADME(T)	Absorption, distribution, metabolism, excretion (toxicology)
ABRF	Association of biomolecular research faculties
ACN	Acetonitrile
AFM	Atomic force microscopy
AGP	α_1-Acid glycoprotein
Ala	L-Alanine
ASF	Asialofetuin
ASGP-R	Asialoglycoprotein receptor
Asn	L-Asparagine
ASOR	Asialoorosomucoid
Asp	L-aspartic acid
AU	Absorbance units
AUC	Analytical ultracentrifugation
BSA	Bovine serum albumin
BXM	(+)-Biotinyl-3-maleimidopropionamidyl-3,6-dioxaoctanediamine (Biotin-PEO$_3$-maleimide)
CA II	Carbonic anhydrase II
CBS	4-Carboxybenzenesulfonamide
CD	Circular dichroism
CM5	Carboxy-methylated chip surface (100 nm)
CMD	Carboxymethyl dextran
CRD	Carbohydrate recognition domain
Cys	L-Cysteine
Da	Dalton
DEAE	Diethylaminoethyl
DIPEA	*N*,*N*’-Diisopropylethylamine
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
DNA	Desoxyribonucleic acid
DPI	Dual polarization interferometry
DTNB	5,5’-Dithio-bis(2-nitrobenzoic acid); Ellmans Reagent
DTT	Dithiothreitol
E. coli	Escherichia coli
EDC	1-Ethyl-3-(3-dimethylaminopropyl)-carbodiimide hydrochloride
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
ESI	Electrospray ionization
F(ab’)$_2$	Antigen-binding fragment of an antibody, dimeric
Fab’	Antigen-binding fragment of an antibody, monomeric
Fmoc	9-Fluorenylmethoxycarbonyl group
FPLC	Fast protein liquid chromatography
Gal	D-Galactose
GalNAc	*N*-Acetyl-D-galactosamine
Glc	D-Glucose
GlcNAc	*N*-Acetyl-D-glucosamine
Gln	L-Glutamine

Glu	L-Glutamic acid
GMP	Good manufacturing practice
GPCR	G-Protein coupled receptor
GSH	Glutathione, reduced form
GSLA-2	Monoclonal antibody (mouse IgG1) recognizing sialyl Lewis a
GSSG	Glutathione, oxidized form
H1 / HL-1	Hepatic lectin 1 (ASGP-R)
H2 / HL-2	Hepatic lectin 2 (ASGP-R)
hAGT	Human O^6-alkylguanine-DNA-alkyltransferase
HBS	HEPES buffered saline
HCl	Hydrochloric acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	L-Histidine
HMA	Human mercaptalbumin
HNA	Human non-mercaptalbumin
HOBt	1-Hydroxybenzotriazole
HPLC	High performance liquid chromatography
HSA	Human serum albumin
HTS	High-throughput screening
HUGO	Human genome organization
IC_{50}	50% inhibition constant
IDA	Iminodiacetic acid
IFC	Integrated fluidic cartridge
IgG	Immunoglobulin G
IMAC	Immobilized-metal affinity chromatography
ITC	Isothermal titration calorimetry
K_A	Equilibrium association constant
K_D	Equilibrium dissociation constant
K_i	Inhibition constant
k_{off}	Dissociation rate constant
k_{on}	Association rate constant
Lem	Lemieux spacer
mAb	Monoclonal antibody
MALDI	Matrix-assisted laser desorption ionization
MBP	Maltose-binding protein
MEA	2-Mercaptoethylamine
MES	2-Morpholinoethanesulfonic acid
Met	L-Methionine
MIRG	Molecular interactions research group
MS	Mass spectrometry
MW	Molecular weight
NA	Neutravidin
NaOH	Sodium hydroxide
NEM	*N*-Ethylmaleimide
NHS	*N*-Hydroxysuccinimide
NME	New molecular entities
NMR	Nuclear magnetic resonance
NTA	Nitrilotriacetic acid
OSM	Orosomucoid
P20	Polysorbate-20 (Tween-20)

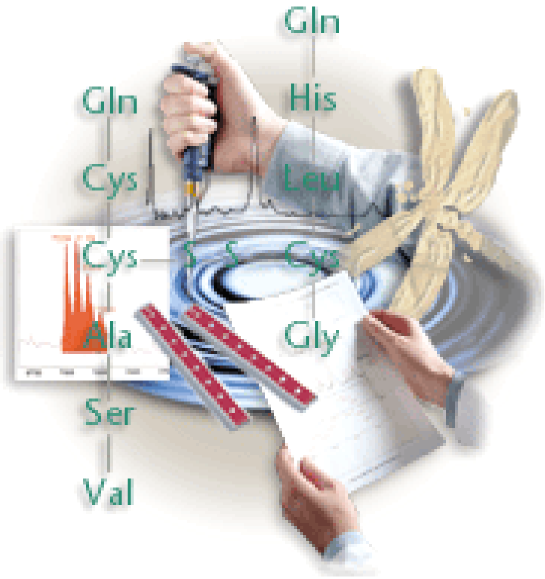
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PDEA	2-(2-Pyridinyl-dithio)- ethaneamine
PPB	Plasma protein binding
QCM	Quartz crystal microbalancing
QUASAR	Quasi-atomistic receptor (modeling)
R	Response
R_{eq}	Equilibrium response
R_{max}	Maximum response
RP	Reversed phase
RU	Resonance unit
SA	Streptavidin
SAR	Structure-activity relationship
SDS	Sodium dodecylsulfate
SEC	Size exclusion chromatography
SFS	Stopped-flow spectrometry
SKR	Structure-kinetics relationship
sLe^a	Sialyl Lewis^a
sLe^x	Sialyl Lewis^x
SpA	Staphylococcal protein A
SpG	Streptococcal protein G
SPR	Surface plasmon resonance
STD	Saturation transfer difference
TBTU	2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborate
TDC	Target definition compound
TFA	Trifluoroacetic acid
TIR	Total internal reflection
Tris	Tris(hydroxymethyl)aminomethane
trNOE	Transfer nuclear Overhauser effect
Trp	L-Tryptophan

Contents of the Thesis

Chapter 1: General Introduction	Drug Discovery on the Move	2
	The Need for Biosensors	6
	Structure and Aim of the Thesis	8
Chapter 2: Biacore Technology	Introduction	12
	Materials and Methods	39
	Results and Discussion	44
	Conclusions	48
Chapter 3: Human Serum Albumin	Introduction	54
	Materials and Methods	63
	Results and Discussion	74
	Conclusions	100
Chapter 4: Diagnostic Antibody GSLA-2	Introduction	106
	Materials and Methods	112
	Results and Discussion	118
	Conclusions	134
Chapter 5: Asialoglycoprotein Receptor	Introduction	140
	Materials and Methods	149
	Results and Discussion	160
	Conclusions	199
Chapter 6: Hexahistidine Tag	Introduction	206
	Materials and Methods	210
	Results and Discussion	214
	Conclusions	224
Chapter 7: General Considerations	Working with Small Molecules	228
	Carbohydrate-Protein Interactions	236
	Negative Binding Responses	239
Appendices	A (CA II)	248
	B (HSA)	249
	C (GSLA-2)	255
	D (ASGP-R)	256
	E (HisTag)	261

Chapter 1

General Introduction


Gln
His
Leu
Cys
Gly
Gln
Cys
Cys
S S
Ala
Ser
Val

1.1 Drug Discovery on the Move

Drug discovery has gone through many changes in the last few decades. While it was first dominated by traditional organic (medicinal) chemistry, screening of natural products, and standard pharmacological assays, it changed dramatically with the development of new technology in both chemistry and biology, as well as in computational sciences and engineering. Molecular biology and biotechnology offered a deeper insight into drug targets (enzymes, receptors, ion channels) and greatly facilitated their production and mutation [1]. Molecular modeling technologies allowed calculation and simulation of drug/protein interactions *in silico*, in some cases without even knowing the structure of the target (e.g. QSAR studies [2]). Rational drug design using protein models or surrogates was believed to revolutionize the process of developing new drugs. Combinatorial chemistry opened the possibility to get access to much larger compound libraries in shorter times than it was possible with rational medicinal chemistry. The accessibility of large numbers of compounds triggered the need for faster testing and screening, which led to the development of high-throughput screening (HTS) or even ultra-HTS methodologies, where far more than ten-thousand molecules could be screened in a single day. This field especially profited from improvements in automation and miniaturization. The human genome projects competitively performed by the international *human genome organization* (HUGO) [3] and the company *Celera Genomics* [4] as well as gene chips by the company *Affymetrix* [5] induced a shift of interest towards finding new targets on the gene level. With genomics still in progress, proteomics emerged as a new field looking no longer at genes but on differences in the expression pattern of proteins in cells or tissues. Proteomics combined traditional electrophoretic techniques (2D-PAGE) with new developments in protein mass spectrometry (ESI, MALDI) to characterize and identify protein targets. While each of these technological developments was first expected to change the way of designing new drugs completely, enthusiasm was set back after a while. Even worse, the number of new molecular entities (NME) on the market remained constant or even decreased while development cost increased dramatically in the last years [6]. Nowadays, the trend is turning to the combination of methods from the fields mentioned above, from medicinal to combinatorial chemistry, from biophysical methods to HTS, or from natural product screening to rational drug design.

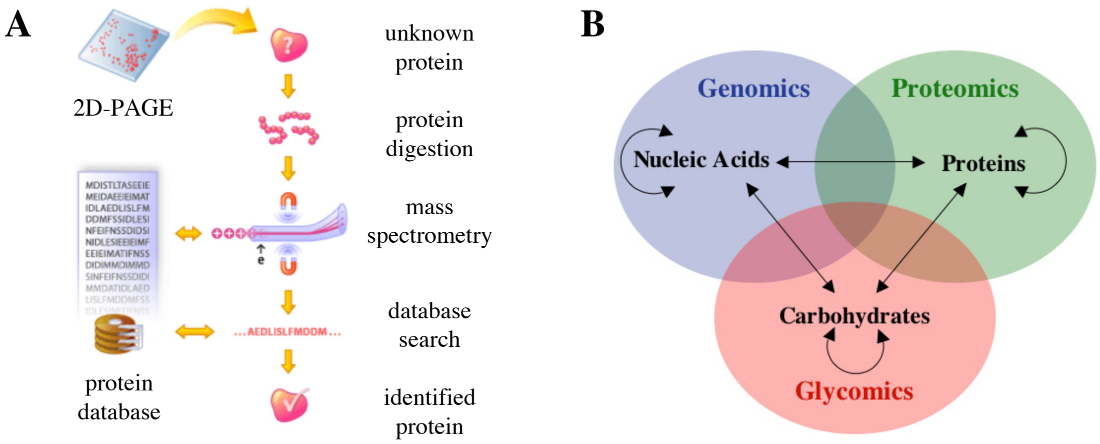

A
2D-PAGE
unknown protein
protein digestion
mass spectrometry
database search
identified protein
protein database
...AEDLISLFMDDM...
B
Genomics
Proteomics
Glycomics
Nucleic Acids
Proteins
Carbohydrates

Figure 1-1: OMICS disciplines. **A:** Identification of unknown proteins by proteomics methods (adapted from the BioTeach website [7]). **B:** Interplay of genomics, proteomics and glycomics. Arrows indicate interactions, which are of interest for both life science and drug discovery (adapted from Ratner *et al.* [8]).

Even though genes and proteins are by far the largest groups of possible target structures, carbohydrates also begin to establish as promising structures. Some sources even speculate about glycomics as the emerging field in the future (*Fig*. 1-1). Carbohydrates offer a few interesting properties, which can be used for drug discovery. They are usually smaller than other biopharmaceuticals, rather stable, less immunogenic, highly specific, and can be formulated more easily. Unfortunately, large-scale production of carbohydrates is still a demanding and expensive task. In addition, carbohydrate-protein interactions are less suitable for screening since 1:1 complexes usually show only weak binding and the observed strength is provided through multivalent interactions [9]. Mediated by glycoproteins and glycolipids, carbohydrate-protein interactions are involved in many physiological and pathological conditions, from microbial infection to inflammation and transplantation medicine. Most of these contacts are mediated by cell surface receptors. Despite their importance, these receptors are fairly neglected as drug targets. With nearly 50%, enzymes clearly dominate the list of targets of drugs on the market, followed by G-protein coupled receptors (GPCR; *Fig*. 1-2) [10].

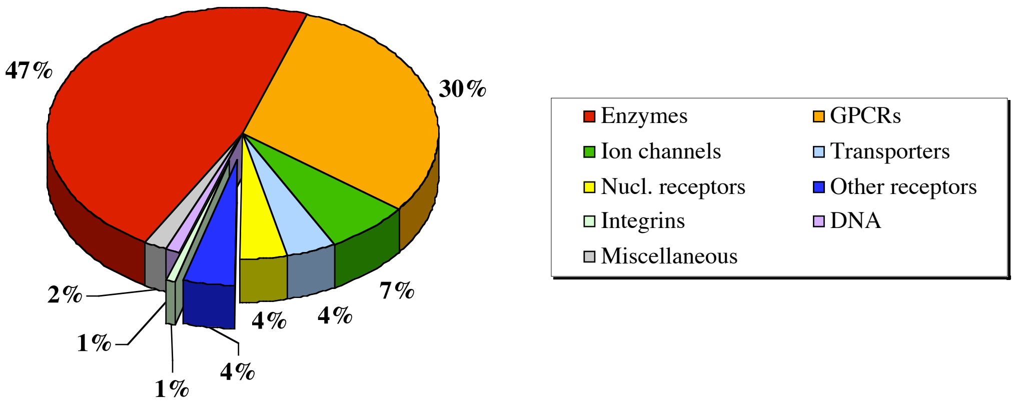


Figure 1-2: Molecular targets of marketed small molecule drugs by biochemical class (chart adapted from Hopkins *et al.* [10]).

Modern drug discovery was not only influenced by the development of new methods and instrumentations in an increasingly rapid pace, but also by some important paradigm shifts. One of the most important conceptual changes was the integration of pharmacokinetic properties in library selection and early screening. Pharmacokinetics, often abbreviated by the term ADME, which stands for *'Absorption, Distribution, Metabolism, Excretion'*, includes all processes in which the body reacts to the administered drug compound. By measuring concentrations in different body fluids it follows the time course of a drug in the body. Retrospective studies showed that nearly 40 % of all drug development failures could be associated to insufficient pharmacokinetic properties (*Fig.* 1-3A) [11].

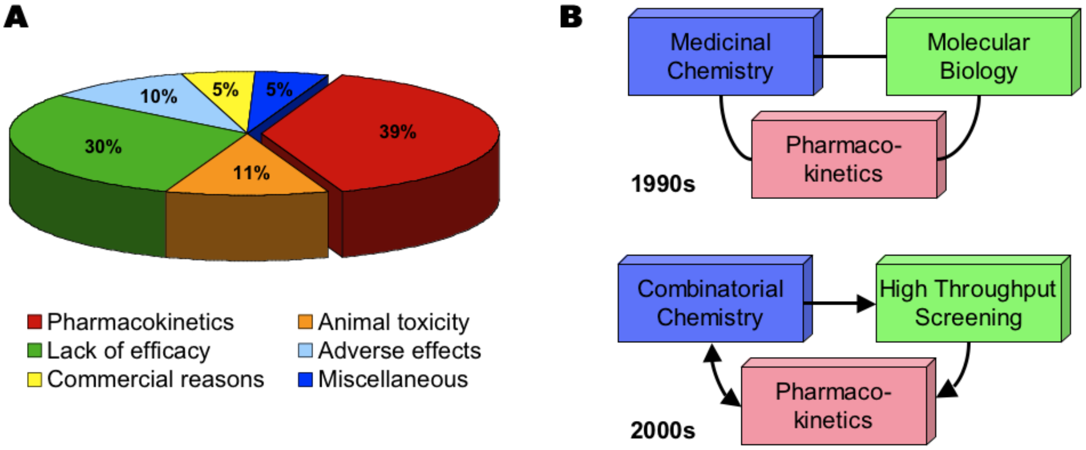
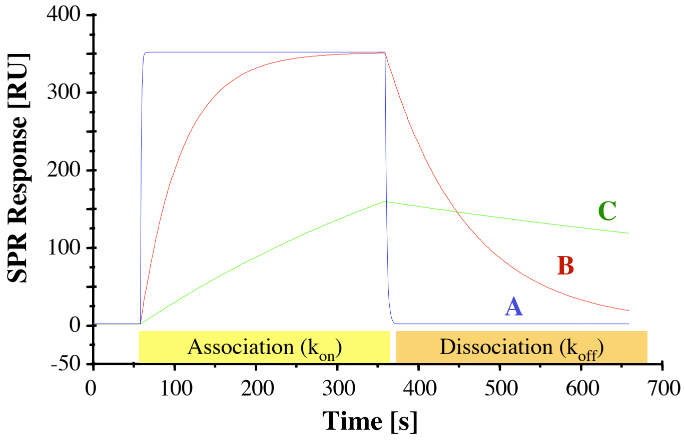


Figure 1-3: Role of pharmacokinetics in drug discovery. **A:** Reasons for drug attrition in 1997. **B:** Interplay between disciplines from the classical project-collaboration approach in the 1990s to a more streamlined and automated approach after 2000 (illustrations adapted from van de Waterbeemd *et al.* [12]).

In the past, the main effort in drug discovery was to optimize the efficacy and specificity of a drug candidate. Problems concerning the pharmacokinetic properties of the candidate drug were often detected at a later stage during pharmaceutical development, in preclinical animal experiments or even during clinical studies. Failures at this state, however, are very costly and have to be avoided. Therefore, pharmacokinetic studies were integrated already in the discovery process, e.g. by *in silico* prediction of ADME properties during library design or early *in vitro* screening for undesired physico-chemical properties (*Fig*. 1-3B). These developments inspired Van de Waterbeemd to redefine ADME as *'Automated Decision-Making Engine'* [12].

Traditionally, drug leads are optimized by just looking at their overall affinities or activities for a target, i.e. their thermodynamic properties. These values just summarize different entropic and enthalpic effects (e.g. hydrophobic or electrostatic interactions, desolvatation, or induced fit). Similar affinities, however, can be the result of completely different kinetic profiles (*Fig*. 1-4). Therefore, a much deeper insight into drug-target interaction is available when the kinetic behavior of a compound is taken into account, resulting in meaningful structure-kinetic relationships [13]. This opens new possibilities for the development of drug candidates with tailor-made properties. For example, an effective enzyme inhibitor should have a fast association rate constant (rapid binding), but a very slow dissociation rate constant (sticking in the binding site).



Equilibrium Constant $(K_D) = \frac{k_{off}}{k_{on}}$

Curve	k_{on} [$M^{-1}s^{-1}$]	k_{off} [s^{-1}]	K_D [μM]
A	$5 \cdot 10^4$	$5 \cdot 10^{-1}$	10
B	$1 \cdot 10^3$	$1 \cdot 10^{-2}$	10
C	$1 \cdot 10^2$	$1 \cdot 10^{-3}$	10

Figure 1-4: Simulated binding curves for a surface plasmon resonance (SPR) biosensor. Even though all three curves (A, B, C) result in the same binding affinity (K_D = 10 μM), the corresponding kinetic rate constants (k_{on}, k_{off}) might deviate significantly.

1.2 The Need for Biosensors

The application of novel and efficient technologies is of high importance to the drug discovery process, since they will lower development costs and decrease the time to market [14]. Even though developments in the field of high-throughput screening and computational chemistry greatly accelerated and facilitated the drug finding process, there are significant limitations to overcome. An example are fluorescence-based HTS assays, which may generate false positive (e.g. binding to the reporter enzyme [14] or direct hydrophobic interaction of the label with the target [15]) or false negative results (e.g. occluding of the binding site [15]). New technologies to confirm or refute screening hits are therefore highly needed. Biosensors have attracted a great deal of attention in this field in recent years.

The definition of the term *biosensor* is not very sharp and can be explained in different ways. In principle, it is a device consisting of a biological part (e.g. DNA, protein, cell) and a physical transducer (semiconductor, electrode, optical component). Biosensor platforms are often miniaturized and work on small chips. First biosensor systems were developed for clinical diagnostics and tailor-made for one specific target or assay. For drug discovery, however, biosensors had to become much more flexible, allowing the screening of a broad variety of compounds from different sources with a reasonable throughput. Research and development in biosensors lead to many experimental or commercial systems on different biological levels (cell, membranes, proteins) and detection principles (electrochemical, optical) [16].

Surface plasmon resonance-based instruments are nowadays the most popular class of biosensors. Their label-free detection, the real-time data acquisition possibilities, their high degree of automation and throughput, as well as the ease of use made them to a valuable tool in drug discovery. An extremely wide range of molecules can be analyzed, from small drugs, DNA, peptides, or proteins up to virus particles or even whole cells [17]. Compared to classical endpoint assays, which are mainly based on competition or inhibition experiments, SPR sensors provide much more information and properties simultaneously (*Table* 1-1).

Table 1-1: Parameters available from SPR-biosensor analyses

Property	**Parameter**
Thermodynamics	K_A, K_D, ΔH, ΔS
Kinetics	k_{on}, k_{off}, $t_{1/2}$
Competition / Inhibition	IC_{50}, K_i
Concentration	[C]
Function / Activity	(yes/no decisions)

Working in the field of drug discovery also means dealing with small molecules. While first SPR biosensor were primarily designed for protein-protein interactions such as antibody-antigen binding, recent improvement in the instrument hardware, experimental design and data analysis extended their application to the routine investigation of low-molecular-weight compounds [14]. Even though the number of reports about successful assays with small molecules is increasing [18], there are still problems and limitations that have to be solved. Moreover, little is known about the impact of biosensors on special topics like surface receptors.

1.3 Structure and Aim of the Thesis

The application of SPR biosensors for the analysis of small molecules is still a relatively new field and under constant development. Many studies were conducted to improve the data quality but there are still limitations and problems to tackle with. Especially when working with carbohydrate-surface receptor interactions, which usually show only weak binding and very fast kinetics, maximum sensitivity and data quality is required. Therefore, the aim of this PhD thesis was to develop different strategies to improve small molecule assays using a Biacore 3000 instrument and to apply these findings to different carbohydrate-protein systems.

For this purpose, different model systems were examined, which are presented as own projects in individual chapters. *Chapter* 2 gives an introduction into the detection principle, function and application of Biacore instruments. In addition, methods generally used for all projects as well as the results from the participation in an international study for a system validation are also presented in this chapter. In *chapter* 3, human serum albumin is investigated as a first model system. It focuses on data analysis and artifacts caused by conformational changes of this target protein. In *chapter* 4, carbohydrate-protein interactions are first investigated using a therapeutic monoclonal antibody (GSLA-2) recognizing a tetrasaccharide epitope (sLe^a). With this model, the influence of several factors like the presence of labels, protein size or non-specific binding could be studied. Binding of mono- and oligosaccharides to the asialoglycoprotein receptor (ASGP-R) was analyzed in *chapter* 5. In *chapter* 6, the binding properties and specificities of the commonly used hexahistidine tag are investigated. Finally, general findings and recommendations for the work with small molecules, carbohydrate-protein interactions, and negative binding signals are summarized in *chapter* 7.

1.4 References

[1] J. Drews, *Science* 2000, *287*, 1960.

[2] A. Vedani, M. Dobler, *J Med Chem* 2002, *45*, 2139.

[3] E. S. Lander, L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J. P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J. C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R. H. Waterston, et al., *Nature* 2001, *409*, 860.

[4] J. C. Venter, M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural, G. G. Sutton, H. O. Smith, M. Yandell, C. A. Evans, R. A. Holt, J. D. Gocayne, P. Amanatides, R. M. Ballew, D. H. Huson, J. R. Wortman, Q. Zhang, C. D. Kodira, X. H. Zheng, L. Chen, M. Skupski, G. Subramanian, P. D. Thomas, J. Zhang, G. L. Gabor Miklos, C. Nelson, S. Broder, A. G. Clark, J. Nadeau, V. A. McKusick, N. Zinder, A. J. Levine, R. J. Roberts, M. Simon, C. Slayman, M. Hunkapiller, R. Bolanos, A. Delcher, I. Dew, D. Fasulo, M. Flanigan, L. Florea, A. Halpern, S. Hannenhalli, S. Kravitz, S. Levy, C. Mobarry, K. Reinert, K. Remington, J. Abu-Threideh, E. Beasley, K. Biddick, V. Bonazzi, R. Brandon, M. Cargill, I. Chandramouliswaran, R. Charlab, et al., *Science* 2001, *291*, 1304.

[5] *Affymetrix Inc., Santa Clara, U.S.A. (www.affymetrix.com).*

[6] E. F. Schmid, D. A. Smith, *Drug Discov Today* 2004, *9*, 18.

[7] C. Antler, *Investigating the cellular machinery: Protein identification*, BioTech website (http://www.bioteach.ubc.ca/Bioinformatics/InvestigatingTheCellularMachinery), 2005.

[8] D. M. Ratner, E. W. Adams, M. D. Disney, P. H. Seeberger, *Chembiochem* 2004, *5*, 1375.

[9] Z. Shriver, S. Raguram, R. Sasisekharan, *Nat Rev Drug Discov* 2004, *3*, 863.

[10] A. L. Hopkins, C. R. Groom, *Nat Rev Drug Discov* 2002, *1*, 727.

[11] T. Kennedy, *Drug Discovery Today* 1997, *2*, 436.

[12] H. van de Waterbeemd, E. Gifford, *Nat Rev Drug Discov* 2003, *2*, 192.

[13] P. O. Markgren, W. Schaal, M. Hamalainen, A. Karlen, A. Hallberg, B. Samuelsson, U. H. Danielson, *J Med Chem* 2002, *45*, 5430.

[14] D. G. Myszka, R. L. Rich, *Pharm. Sci. Technol. Today* 2000, *3*, 310.

[15] M. A. Cooper, *Nat Rev Drug Discov* 2002, *1*, 515.

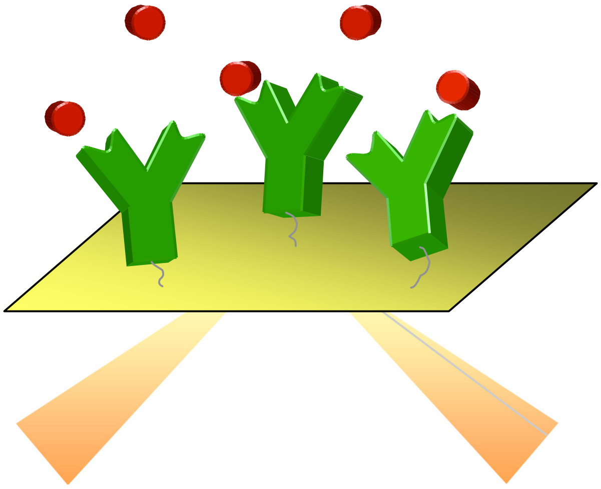
[16] M. Keusgen, *Naturwissenschaften* 2002, *89*, 433.

[17] R. L. Rich, D. G. Myszka, *Drug Discovery Today: Technologies* 2004, *1*, 301.

[18] R. L. Rich, D. G. Myszka, *J Mol Recognit* 2005, *18*, 1.

Chapter 2

Biacore Technology



2.1 Introduction

This chapter first gives an overview of the detection principle (surface plasmon resonance, SPR) and the function of Biacore instruments. After an introduction in assay design, immobilization techniques and data analysis, it covers the applications of Biacore in the drug discovery and development process. Finally, alternative and complementary biophysical methods are discussed and compared with the Biacore technology.

2.1.1 Surface plasmon resonance and Biacore technology

Surface plasmon resonance is an electron charge density wave phenomenon first observed as early as in the late 1950s [1]. But it took another ten years until its mechanism and versatility was recognized. The first commercially available SPR detection systems only appeared in the 1980s.

The underlying principles of this phenomenon are total internal reflection (TIR), evanescence electric field (E), and surface plasmon waves. *Total internal reflection* occurs when a light beam propagates through two non-absorbing media of different refractive index (e.g. glass-air or glass-buffer). Above a critical incidence angle (Θ), the light beam is no longer refracted when it hits the interface of the two media, but is fully reflected and propagates back into the source medium (*Fig.* 2-1A). Even though the light beam keeps its net energy upon reflection, an electric field intensity called *evanescence wave* (E) leaks into the other medium. This wave is exponentially decreasing with distance from the interface (*Fig.* 2-1B). When the interface is coated with a thin metal film the p-polarized component of the evanescence field penetrates this layer and induces electromagnetic *surface plasmon waves* in the conducting metal. Plasmons represent electron density fluctuations in a conducting metal and can be regarded as the equivalent of photons in the case of light. A non-magnetic metal like gold is normally used for these metal layers and the thickness has to be lower than the wavelength of the incident light beam [2-4].

Since both photons and surface plasmons are a form of electromagnetic energy, they can be fully described only by quantum physics. However, their properties can be explained in a simplified manner as vector quantities. The light photon momentum at the interface can be resolved into two vector components (parallel and perpendicular to the interface). The magnitude of these incident light vectors (ilv) directly depends on the light angle. The surface plasmon wave can be similarly described as a vector, which depends on a number of factors (metal properties, layer thickness, surrounding media). Only when the energy and momentum of the incident light vector exactly correspond to the one of the surface plasmon vector (spv), a resonance phenomenon occurs and photons are converted into plasmons (*Fig.* 2-1D). Otherwise, there is no such conversion and the light is fully reflected (*Fig.* 2-1C).

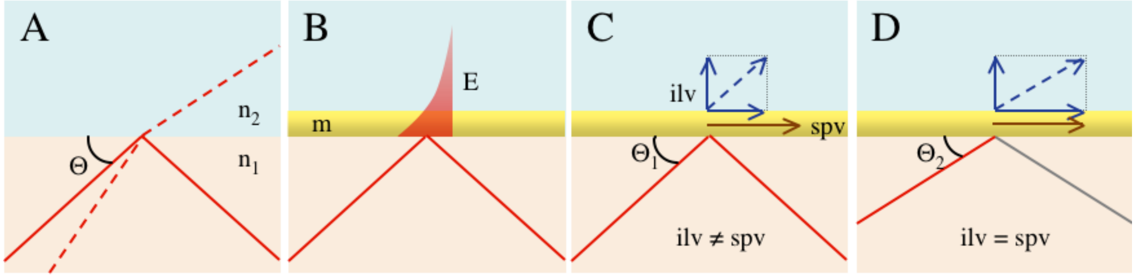


Figure 2-1: Principles of SPR. **A:** Total internal reflection (solid line) and refraction (dashed line) of a light beam in dependence of the incidence angle Θ at the interface of two different media (n_1, n_2). **B**: Evanescence field wave (E) leaking through a thin metal film (m). **C, D:** SPR in the gold surface. If the incident light vectors (ilv) have another value than the surface plasmon vector (spv) light is fully reflected (**C**). Only a specific angle leads to a matching of the two vectors and a resulting resonance (**D**).

When the metal nature and thickness as well as the properties of one medium are kept constant during an experiment, resonance can be obtained only by variation of incident light angle and the refractive index of the second medium. Therefore, changes in this medium can be followed be adjusting the incident light angle until a dip in light intensity (resonance) is detectable [2-4].

Biacore biosensors (with BIA standing for *biomolecular / biospecific interaction analysis*) use this phenomenon for the detection of mass differences in a sample cell. A *sensor chip* carries a thin gold layer (50 nm) on a glass support. The chip is in direct contact with a flow cell (sample side) and a glass prism (detector side). A monochromatic, plane-polarized light beam at a wavelength of 760 nm is sent through the prism and is totally internally reflected at the interface. The generated evanescence field wave penetrates into the sample cell and allows the detection of refractive index properties to a distance of about 1 μm from the surface. The angle of minimum light intensity is detected using a two-dimensional detector array. Biomolecular interactions around the surface cause a change of the solute concentration and therefore of the refractive index of the medium, which can be detected as a change in the incidence light angle and converted into a response signal (*Fig.* 2-2). The unit of the response signal is called *resonance unit* (RU) and represents a shift in the resonance angle of approximately 10^{-4} ° [3]. Since the mass of the molecules directly influences the refractive index, SPR biosensors are often referred as *mass detectors*. In the case of proteins, the correlation between sensor signal and mass increase was experimentally determined (*Eq.* 1) [5]:

$$1\ \mathrm{RU} = 1\ \mathrm{pg/mm^2} \qquad [\text{E-1}]$$

This correlation is practically constant for molecules with high protein and low lipid and carbohydrate content [3]. Even though there might be some minor deviations, the relationship is applicable to other biomolecules such as nucleic acids, carbohydrates, lipids or conjugate molecules. As a consequence, mass concentration can be detected with high sensitivity for nearly all molecules, regardless of their nature [4]. On the other side, sensitivity is dependent on the distance from the surface. Therefore, other changes around the interface, e.g. electrostatic attraction or conformational changes, might also induce a shift of the incident light angle [6].

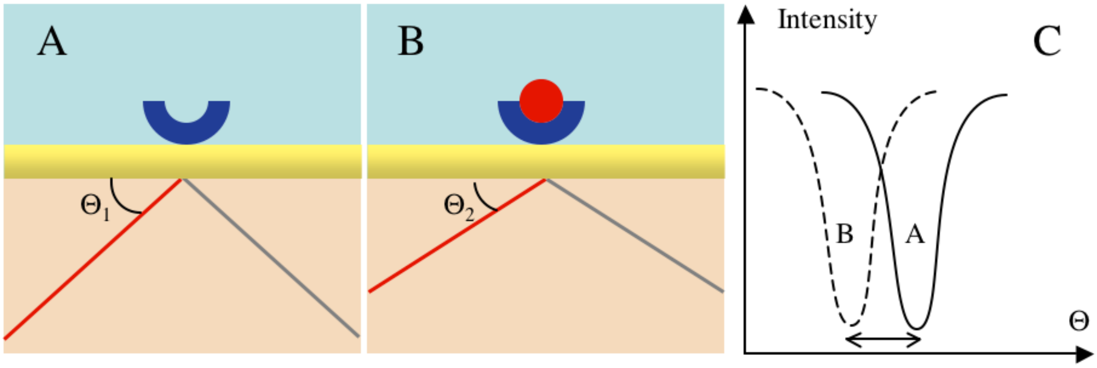


Figure 2-2: Detection of biomolecular interaction by SPR. **A, B:** Sensor surface before and after interaction of two molecules. **C**: Shift of light intensity dip upon interaction.

In a typical Biacore experiment, one binding partner (e.g. a receptor or enzyme) is immobilized on the sensor chip and the other is injected in solution. In the case of proteins, the direct surface-attachment to a solid (gold) support often leads to uncontrollable binding and loss in biological activity. To overcome this problem, a special surface chemistry was developed involving a 'protecting polymer', which carries functional groups for easy immobilization. This polymer consists of thiolated *carboxymethyl dextran* chains, which are directly complexed by the gold surface via the sulfur atom and form a self-assembled monolayer. Free carboxyl groups in this matrix can be used to immobilize various synthetic and biological molecules using well-defined chemistry. As a result, the molecules are embedded in a highly hydrophilic hydrogel and are kept in a quasi-solvent environment [7]. Electrostatic effects caused by remaining carboxyl groups can be suppressed in most cases by adding salts (e.g. 150 mM NaCl) to the running buffer [3]. A schematic overview of the experimental setup is visualized in *figure* 2-3.

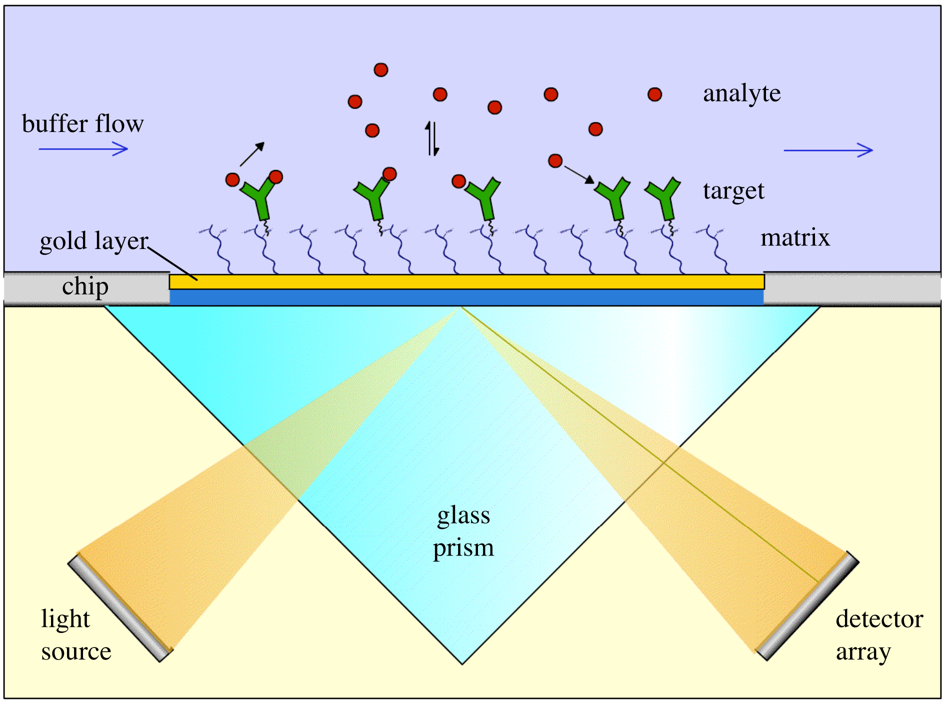


Figure 2-3: Experimental setup of Biacore instruments. A target molecule is immobilized on a gold-coated sensor chip via a hydrogel matrix. Binding of analyte molecules in solution is detected by SPR phenomenon.

Another problem that had to be addressed was the sample delivery system. In stationary systems, mass transport of molecules to the surface is governed by diffusion and convection processes. To ensure reliable results, incubation times of several hours were necessary in this case, which is not suitable for real-time systems. A flow system in a micro-flow cell offers a continuous transport of sample to and from the surface, therefore minimizing diffusion and convection effects. Developments in miniaturization led to an *integrated fluidic cartridge* (IFC; *Fig*. 2-5C), which further reduced sample consumption and sample plug dispersion after injection [3].

The shift in resonance angle can be monitored in real-time and plotted in dependence of time. From such a signal vs. time plot, called *sensorgram*, the different stages of a binding event can be visualized and evaluated (*Fig*. 2-4). With only buffer running through the flow system, the signal forms a stable baseline. Upon injection of the analyte solution, the sensorgram is dominated by the *association phase*, where analyte molecules bind to the target on the chip. However, bound molecules already start dissociating again during injection. After a certain injection time, a *steady state* is reached, where binding and dissociating molecules are in equilibrium. As soon as the injection is stopped, running buffer replaces the analyte cloud and only the pure *dissociation phase* is visible. Some assays require an additional regeneration step to reach the baseline again (*Fig*. 2-4).

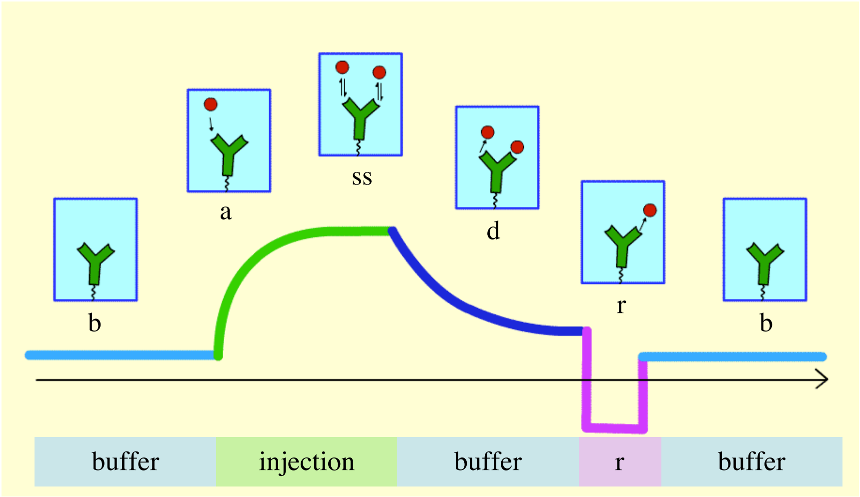


Figure 2-4: Schematic representation of a sensorgram as the time course of a binding event. Baseline signal (b), association (a), steady state (ss), dissociation (d), and regeneration (r) of tightly bound molecules.

From the shape of the binding curve kinetic parameters like the association and dissociation *rate constants* (k_{on}, k_{off}) can be fitted and calculated. The *equilibrium dissociation constant* (K_D) can be directly calculated from the kinetic rate constants using *equation* 2 or independently determined from the steady state signals at different concentrations. This steady state affinity can be calculated using *equation* 3, where R_{eq} is the equilibrium response signal, K_A is the *equilibrium association constant* (*Eq*. 2), C the concentration, R_{max} the maximum possible response, and n a steric interference factor.

$$K_D = \frac{k_{off}}{k_{on}} \ [\mathrm{M}] \quad \text{and} \quad K_A = \frac{k_{on}}{k_{off}} \ [\mathrm{M}^{-1}] \qquad \text{[Eq. 2]}$$

$$R_{eq} = \frac{K_A \cdot C \cdot R_{\max}}{1 + K_A \cdot C \cdot n} \ [\mathrm{RU}] \qquad \text{[Eq. 3]}$$

2.1.2 Biacore 3000

Biacore AB [8] introduced the first SPR biosensor in 1990 [9, 10]. It was primarily designed for the analysis of protein-protein interactions (e.g. antibody-antigen) and had limitations both in regard of sensitivity and automation. With the introduction of Biacore 2000 in 1994 these problems were addressed and it was even possible to investigate small molecules (< 500 Da). Another improvement of sensitivity was realized with Biacore 3000 in 1998. Thanks to its flexibility and the ease of automation it soon gained interest both in pharmaceutical industry and academic laboratories [11, 12].

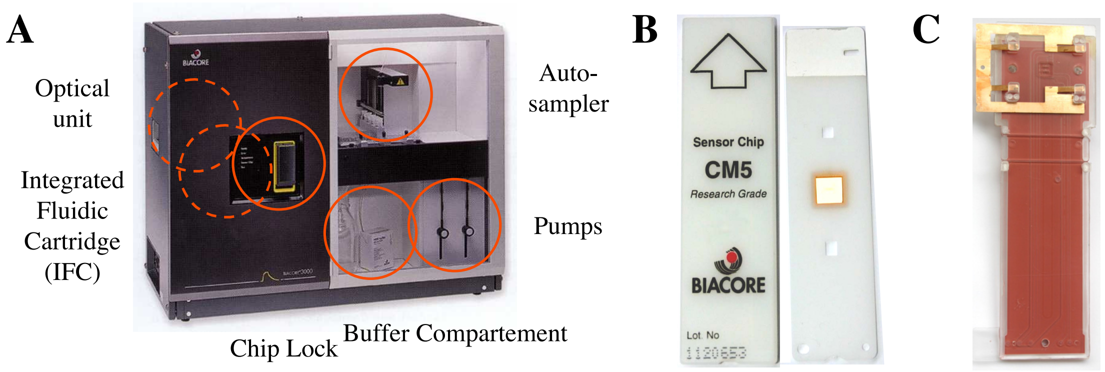


Figure 2-5: Biacore 3000 instrument. **A:** Front view of the instrument with important parts indicated by circles. Dashed lines represent parts that are inside the instrument and not visible. **B:** CM5 sensor chip (right) with its cartridge (left). **C:** Integrated fluidic cartridge (IFC).

Biacore 3000 is basically built of three parts (*Fig*. 2-5A); an *autosampler* for sample delivery and injection, the *optical unit*, and the sensor chip compartment with the *integrated fluidic cartridge* (IFC; *Fig*. 2-5C). The IFC divides the sensor chip into four individually addressable flow cells (1.2 mm^2, 0.02 μl per cell) and controls the buffer flow with different valves. One of these flow cells is usually used as a control surface to subtract bulk signals of the buffer or non-specific binding. The other flow cells can be used for the immobilization of target molecules. The flow rate is variable in a range from 1 to 100 μl/min and the whole IFC and optical unit is thermostatically controlled (4-40°C). Samples are injected by a movable autosampler needle, which can deliver samples from vial racks or 96 well plates. The range of injectable volumes is between 1 and 400 μl, depending on the injection mode [13].

2.1.3 Assay Design

In every Biacore experiment one of the binding partners has to be attached to the sensor chip surface (see also *section* 2.1.1). Biacore calls this molecule the ‘ligand’. However, this term is more often used to describe molecules binding to receptors and is part of many expressions like ‘ligand-induced conformational changes’. Therefore, the immobilized molecule is always referred to as the *target* in this thesis. In agreement with the Biacore nomenclature, the interacting molecule in solution is called *analyte* (*Fig*. 2-6A). The expression '*surface*' and derived terms like 'surface density' are normally referred to the sensor chip with the immobilized target.

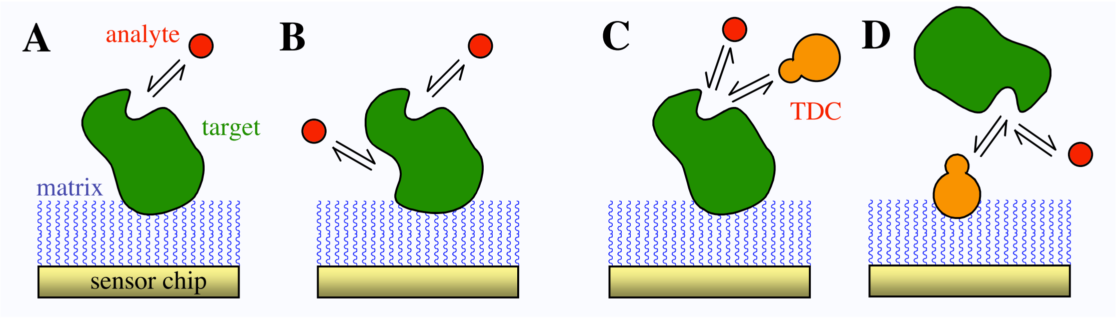


Figure 2-6: Comparison of different assay types. Direct binding assays fitting to a single-site model (**A**) or a two independent-sites model (**B**). Surface competition assay (**C**) and inhibition in solution assay (**D**), in which the analyte competes with a 'target definition compound' (TDC) for the same binding site.

The first strategic decision to make is which of the binding partners is immobilized. For most of the systems, however, there is no real choice since multiple analytes will be screened against a single target molecule. In order to get maximum signal responses, systems with an immobilized small molecule and a bigger analyte (e.g. protein) in solution are preferred. The maximum response for a SPR signal can be estimated using *equation* 4.

$$R_{\max} = \frac{MW_{analyte}}{MW_{target}} \cdot density_{target} \cdot valency \qquad \text{[Eq. 4]}$$

Unfortunately, for the majority of drug discovery applications the large molecule (receptor, enzyme, etc.) will be the target and small molecules (MW < 500 Da) are used for screening. This often leads to very small signals around the detection limit, especially when the coupling results in a low density or a reduced activity of the target. In addition, immobilization of the small molecule might change the binding event dramatically, since multivalency or rebinding effects are often observed [14].

Second, several assay formats can be performed, of which the *direct binding assay* is by far the most popular (*Fig*. 2-6A&B). However, competition assay formats might be preferable for different reasons (e.g. small analyte size). In the *surface competition assay* (*Fig*. 2-6C) the analyte is mixed with a constant concentration of a *target definition compound* (TDC), which is normally a tight inhibitor. The TDC should form a complex with a half-life of more than 20 s and should be at least 5 to 10 times larger than the compounds to be screened. Changes of the overall binding response are then evaluated and compared with the signal of the TDC alone. Both the direct and the competition assay are sensitive to non-specific binding. A third format is the *inhibition in solution assay* (*Fig.* 2-6D), where the TDC is immobilized and the analyte solutions are mixed with a constant concentration of the target. The signal reflects the concentration of free target and is therefore site-related - only analytes that interact directly with the binding site inhibit the interaction. On the other hand, much higher amounts of the target molecule are needed for this assay, which is often a problem in the case of proteins. While all three formats are suitable for ranking experiments only the direct assay can provide high-quality equilibrium and kinetic data [15].

2.1.4 Immobilization

Target immobilization is one of the most important and crucial steps in a Biacore binding assay. Loss of target activity and many artifacts are directly related to unfavorable coupling procedures. The unique properties of the hydrogel matrix used for Biacore experiments (CM5 chip; *Fig.* 2-5B) offers many alternative strategies for covalent immobilization of proteins, oligosaccharides, nucleotides, or small molecules (*Fig.* 2-7).

While *covalent coupling* approaches usually generate stable surfaces with high density, *capturing* techniques have the advantage of being fully regenerable and allow immobilization from (crude) protein mixtures. These two general approaches also show different results in respect of target orientation. They also lead to oriented and therefore highly active surfaces but often show a lower density and stability (surface bleeding).

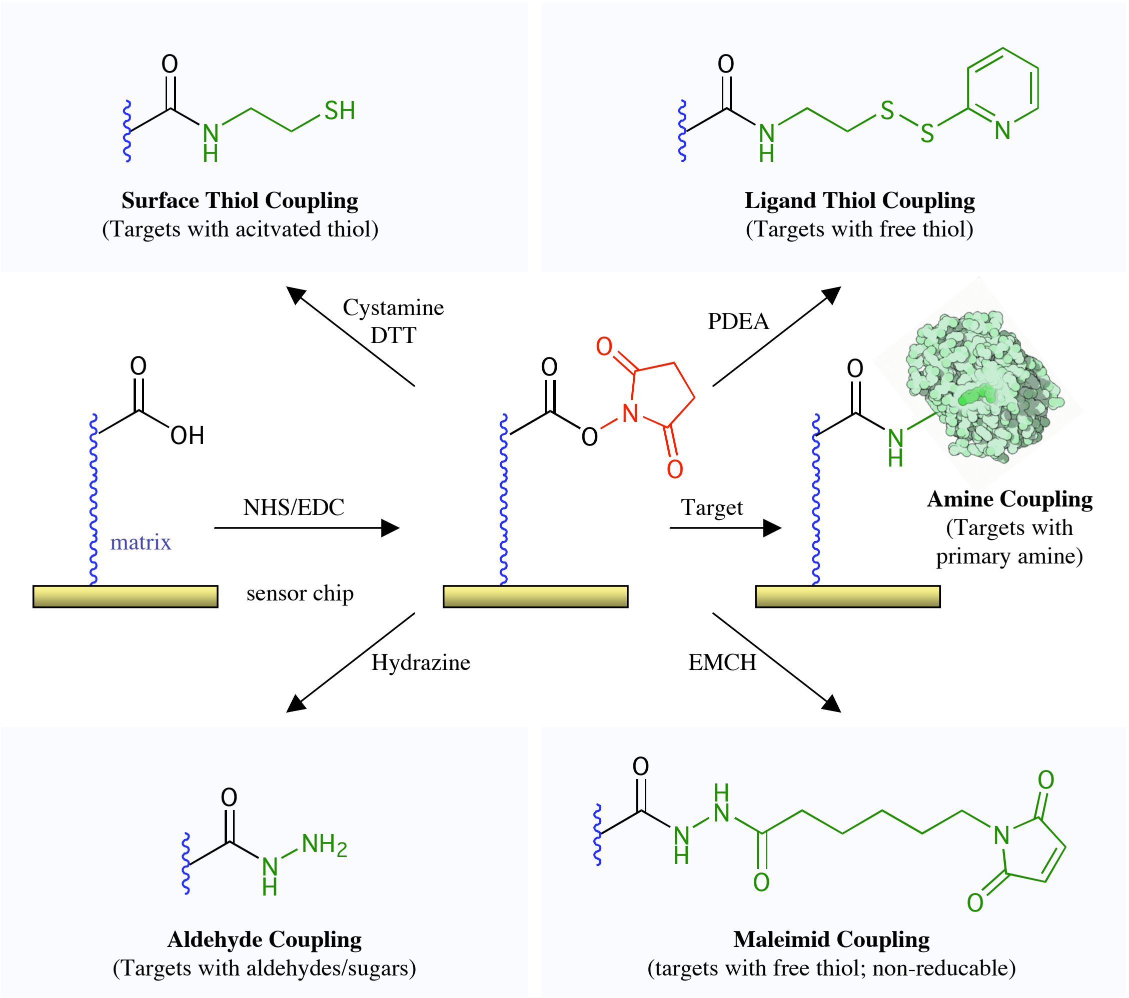

Surface Thiol Coupling
(Targets with acitvated thiol)
Ligand Thiol Coupling
(Targets with free thiol)
Cystamine
DTT
PDEA
matrix
NHS/EDC
sensor chip
Target
Amine Coupling
(Targets with
primary amine)
Hydrazine
EMCH
Aldehyde Coupling
(Targets with aldehydes/sugars)
Maleimid Coupling
(targets with free thiol; non-reducable)

Figure 2-7: Coupling methods for Biacore sensor chip CM5. After activation of the matrix-based carboxyl groups by NHS and EDC, targets can be directly immobilized via primary amine groups or the surface can be functionalized for alternative strategies (NHS = *N*-hydroxysuccinimide,, EDC = 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide, PDEA = 2-(2-pyridinyldithio)ethaneamine), EMCH = *N*-(ε-maleimidocaproic acid)-hydrazide,

Covalent immobilization often attacks multiple attachment sites and therefore leads to *randomized* coupling, which often results in a loss of activity (due to direct modification of residues in the binding site, steric hindrance or conformational changes) and surface heterogeneity (*Fig*. 2-8A). However, if a specific functional group is available at a defined location of the target, *site-specific* coupling might allow the generation of an oriented, homogeneous surface (*Fig*. 2-8B). Finally, *capturing* approaches (*Fig*. 2-8C) make use of specific biomolecular interactions between the target and an immobilized capturing protein.

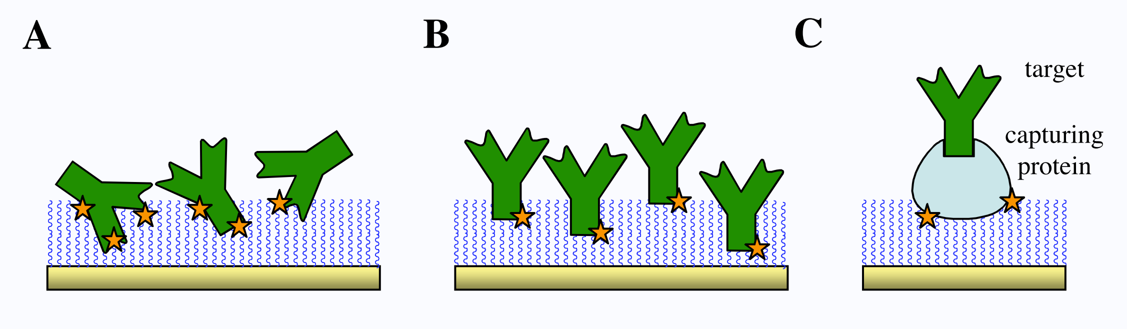


Figure 2-8: Target orientation after different coupling procedures. Randomized (**A**) and site-directed (**B**) covalent immobilization compared to a capturing approach (**C**). Sites of covalent attachment are marked with a star.

While the standard CM5 sensor chip is suitable for most applications, some experiments require modifications of the matrix chemistry or a pre-coated capturing structure. Biacore therefore offers a selection of sensor chip with alternative surface coatings to extend the possible applications or to reduce experimental artifacts (*Table* 2-1).

Table 2-1: Available sensor chip surfaces from Biacore with their surface modification and principal applications (from Biacore Sensor Surface Handbook [16]).

Chip	Type	Application / Improvement
CM5	normal carboxymethyl dextran	general purpose
CM4	lower carboxymethylation	reduced non-specific binding
CM3	shorter dextran matrix	large molecules
C1	flat carboxyl (no dextran matrix)	dextran interference (e.g. lectins)
SA	immobilized streptavidin	biotinlyated molecules (e.g. RNA)
NTA	immobilized nitrilotriacetic acid	histidine-tagged proteins
L1	lipophilic groups on dextran	liposomes, bilayers
HPA	flat hydrophobic surface	lipid monolayers
Au	plain gold surface	custom design

The most widely used immobilization techniques are further discussed and compared in the following sections.

Amine Coupling

The amine coupling procedure makes use of the primary amine groups on the protein surface (lysine residues and the *N*-terminus), which directly react with active esters generated by NHS/EDC activation (*Fig.* 2-9). In order to reach the highest efficiency of the reaction, proteins have to be pre-concentrated on the sensor-chip surface. This surface attraction is reached by lowering the *pH* of the immobilization buffer just below the *pI* value of the protein, where amine groups are positively charged and get attracted by the negatively charged carboxyl group of the matrix. However, since the reaction only takes place with uncharged amines, the immobilization *pH* should not be too low [17].

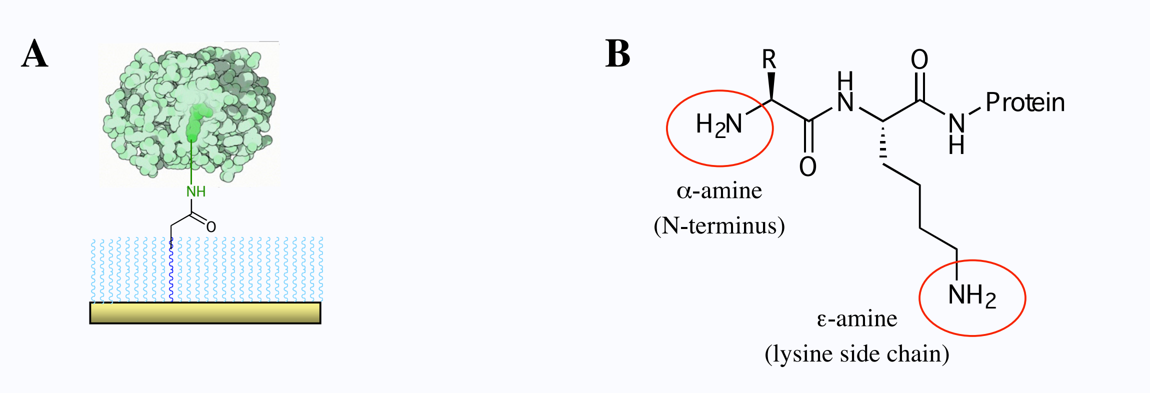


Figure 2-9: Amine coupling: Surface chemistry (**A**) and targeted amine groups at the *N*-terminus and the lysine side chain (**B**).

The major advantage of amine coupling lies in its universality, stability and speed. Nearly all proteins and peptides possess multiple primary amine groups (*N*-terminus and lysine residues), which are often surface-exposed due to their hydrophilicity. On the other hand, since targetable lysine amines often are randomly distributed over the protein surface, amine coupling leads to a random and non-predictable immobilization of the molecule (*Fig.* 2-8A). This is especially problematic in the case of surface receptors since their binding sites are directly accessible to the solvent and charged residues like lysine are often involved in ligand binding. Therefore, amine coupling sometimes leads to a massive decrease of surface activity, e.g. more than 80% loss is reported for some antibodies, and might also influence binding affinity and kinetics [18]. Different strategies have therefore been developed to overcome this problem. For example, differences in the reactivity of the α-amino group of the *N*-terminus and ε-amino groups of the lysine side chain (*Fig.* 2-9B) were used for a site-specific PEGylation of a somatostatin-analogue peptide at different *pH* values [19]. Other groups tried to reversibly protect reactive lysines with 2,3-dimethylmaleic anhydride [20, 21] or masking binding site amines by immobilizing the protein in the presence of a known binder [22]. However, none of these approaches seem to be generally applicable and neither method has been used for Biacore assays. Another drawback of amine coupling is the requirement of acidic conditions for surface attraction. Some proteins are not stable in the immobilization buffers required for amine coupling and only inactive protein is therefore immobilized. Acidic proteins with *pI* values below 3.5 can hardly be immobilized via amine coupling. Finally, popular buffers and reagents bearing primary amines like Tris cannot be used due to competition with the amino groups of the protein. Sodium azide, which is frequently used as a preservative in protein preparations, also might interfere with amine coupling and should therefore be removed [23].

Thiol coupling

Immobilization of thiol-bearing targets can be performed either by formation of disulfide bridges or by covalent reactions with maleimides. Since free thiol groups are very rare compared to primary amines, these approaches often lead to a site-directed and therefore oriented immobilization of the target. Disulfide bridge formation offers the additional advantage that such bonds can be reduced leading to a fully regenerable chip surface. However, since the spontaneous formation of disulfides is thermodynamically not favored and takes very long, activation of the thiol group either on the chip surface or in the target is needed (*Fig*. 2-10A).

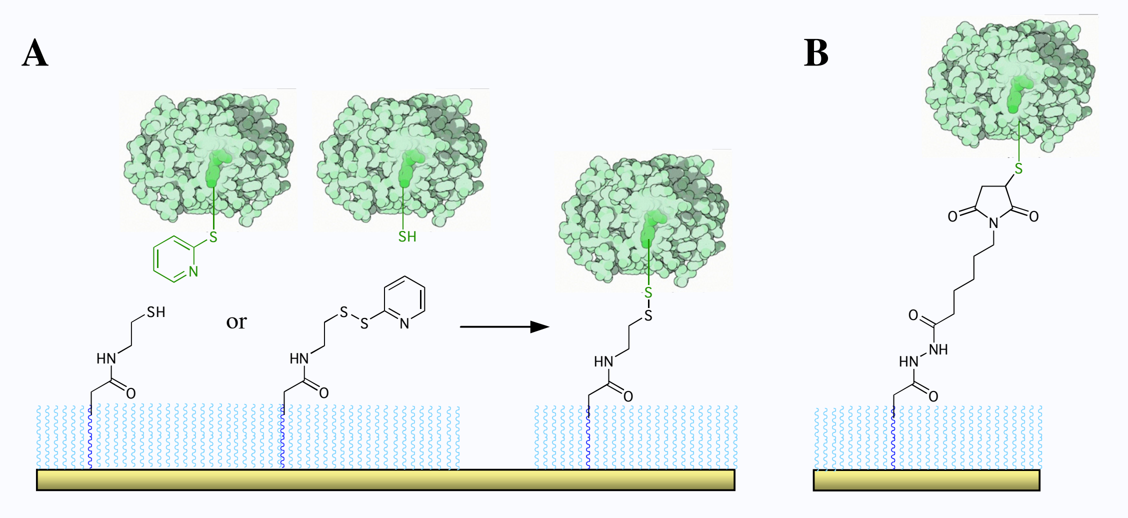


Figure 2-10: Immobilization methods for thiol-bearing targets: surface and ligand thiol coupling (**A**) and maleimide coupling (**B**).

Even though disulfide bridges are very stable under physiological conditions they are cleaved under reducing conditions as well as at higher *pH*. These limitations can be circumvented by using maleimide coupling, which forms a covalent non-reducible thioether bond with free thiol groups of a target (*Fig*. 2-10B). Both methods are usually less susceptible to buffer and reagent additions than amine coupling [16].

The low frequency of free thiols in proteins is one of the major drawbacks of this method. Though active thiols can be introduced by functionalizing amine or carboxyl groups (surface thiol coupling), the advantage of a site-directed attachment is usually lost. Only a few natural proteins contain a free and surface-accessible cysteine residue (e.g. albumin). The introduction of additional cysteines into recombinant proteins by site-directed mutagenesis may disturb protein structure and function (e.g. oligomerization) [24]. A very elegant approach of introducing *N*-terminal cysteine residues was recently reported by Gentle *et al.* [25].

Other covalent coupling methods

Aldehyde coupling is mainly used for immobilizing carbohydrate molecules or glycoproteins (e.g. antibodies). First, a reactive aldehyde group has to be generated by oxidation of *cis*-diols, which can then be immobilized on a hydrazine-activated surface. A final reduction step with cyanoborohydride is usually needed to stabilize the surface. Since glycosylation of proteins is often limited to a few well-known sites, this approach usually leads to a site-directed immobilization. However, the necessary (mild) oxidation and reduction steps might influence the activity of the target.

Of course, carboxyl groups can also be used for coupling procedures, but this approach is limited by the applied surface chemistry. While amine groups can simply be introduced to the chip surface by immobilizing ethylenediamine, the activation of carboxyl groups in the protein is much more problematic since they readily react with protein amines and form oligomers. By an activation with NHS and EDC in an excess of PDEA carboxyl groups can be functionalized with activated thiols and immobilized by surface thiol coupling (see above).

Capturing

Capturing approaches are widely used in biomolecular interaction measurement. They rely on non-covalent protein-protein or protein-small molecule interactions and are especially suitable for experiments where both target and analyte have to be screened simultaneously. In addition, capturing often serves as an easy way for oriented coupling, since binding occurs at a well defined site of the target. Three major coupling classes can be defined: antibody-antigen systems, interactions between proteins and naturally occurring sites (e.g. Protein A/IgG) and capturing of artificially introduced affinity tags (e.g. biotin or hexahistidine).

Antibody-antigen systems offer many advantages over other capturing approaches. Interactions show normally high affinities (nanomolar range) and specificity. However, production of antibodies against a new target can be very time and cost consuming and care has to be taken to avoid overlaps between antibody and analyte binding sites. Therefore, antibody systems used for Biacore analysis are often directed against well-known antigens like tags or conserved domains of protein families.

Affinity tags are short peptide sequences or whole protein domains, which show high affinity to a specific target structure. This could be another protein, a small molecule or a metal ion. Tags are an established method in protein expression and purification, and plasmids for the production of fusion proteins are readily available. Expressed tags can be used for purification (affinity chromatography) as well as for immobilization on a sensor chip. However, not every expression system tolerates a newly introduced domain and special elution conditions might have to be applied during purification. This might lead to reduced yield or decreased activity of the proteins. An overview of several important tag systems can be found in *table* 2-2 and in Terpe [26].

Table 2-2: Popular tag systems used for target capturing in Biacore experiments and other assays.

Tag Name	Residues	Captured by	K_D [nM]	Ref [a]
His-tag	6(-10)	Ni^{2+}-NTA, anti-$His_{5/6}$	nM-μM[b]	[27]
Strep-tag II	8	Streptactin	n.d.	[28]
FLAG	8	Anti-FLAG mAb	412	[29]
SBP (streptavidin binding protein)	38	Streptavidin	2.5	[30]
Z/ZZ-domain (Protein A)	53/123	Human IgG1 (Fc)	17/1.5	[31]
Glutathione S-Transferase	211	Glutathione, anti-GST	n.d.	[32]

[a] All references and K_D values refer to SPR experiments, except SBP (spin-filter binding inhibition assay).

[b] The isolated hexahistidine peptide was shown to bind differently to Ni-NTA than tagged proteins. At least two His-tags were found to be necessary for a stable binding on a Biacore NTA-chip.

A common disadvantage of all capturing approaches is their non-covalent character. While mid-range affinities and non-physiological buffer conditions might be tolerable for purification purposes, this might be a problem for the creation of stable sensor surfaces. Captured surfaces often show a certain degree of bleeding (surface decay) and finding selective regeneration conditions can be very difficult.

Recently, a new and elegant way of using an affinity-tag approach for the generation of covalent surfaces was introduced by the company *Covalys Biosciences AG* [33]. The so-called *SNAP-tag* is expressed as a fusion protein of mutated human O^6-alkylguanine-DNA-alkyltransferase (hAGT) with the protein to be immobilized. hAGT is a DNA repair protein, which transfers the alkyl group from O^6-alkylguanine-DNA to one of its cysteine residues. Derivatized O^6-benzylguanine fixed on the sensor chip can be used for site directed and covalent immobilization of the fusion protein. The main advantages of this approach are the defined orientation of the fusion protein, the covalent and therefore stable character of the coupling and the possibility to immobilize the protein directly from the crude cell extract [34]. On the other side, hAGT is a protein of some 200 amino acids, whose cloning and expression as a fully functional domain might be difficult in some expression systems. A successful application of this approach can be found in a recent study by Huber *et al.* [35], where a hAGT fusion protein of cyclophilin D was expressed in *E. coli* and immobilized with high activity and stability.

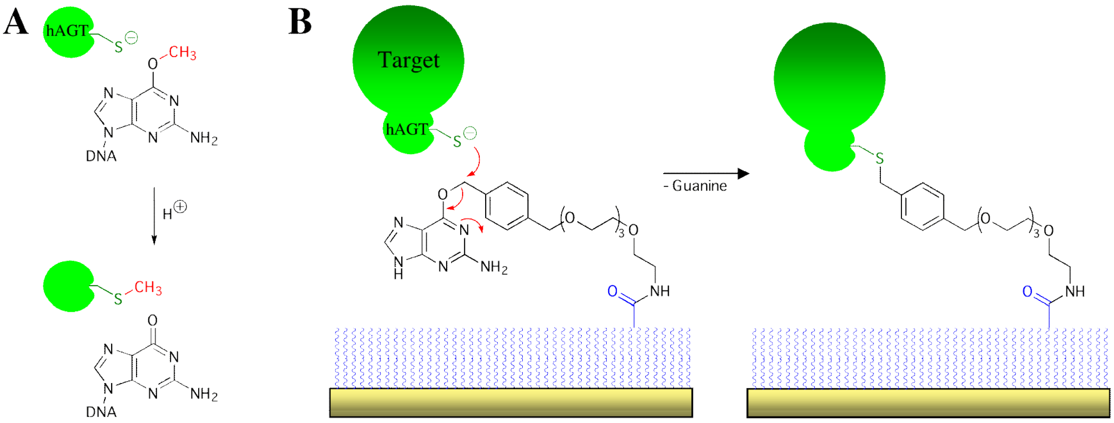


Figure 2-11: Protein immobilization by SNAP-tag technology. **A:** Naturally occurring DNA repair by hAGT. **B:** Coupling of a hAGT fusion protein to an immobilized guanidine derivative (adapted from Kindermann *et al.* [34]).

2.1.5 Assay conditions

Biacore 3000 accepts a wide range of conditions and variation of parameters. On the other hand, most of the experiment are conducted under near-physiological conditions using water-based buffer systems and temperatures between 20 and 37°C. Buffers used for Biacore experiments are normally amine-free (to avoid conflicts in amine-coupling) and contain a certain amount of salt for suppressing electrostatic effects on the carboxylated matrix [3] (e.g. 10 mM HEPES or phosphate buffer at *pH* 7.4 with 150 mM NaCl). Reagents such as EDTA or polysorbate are often added to reduce non-specific binding, but only after possible interferences with the binding experiment have been excluded.

One of the unique features of the Biacore technology is its flow system. This ensures a fast delivery of the sample to and from the surface. Variations of flow rate are suitable for the detection of any mass transport effects. This phenomenon might occur when the interaction between analyte and target is comparable or faster than the diffusion of analyte from bulk solution to the surface. Mass transport is dependent on the flow rate, cell dimensions and diffusion coefficient of the analyte [36, 37]. High flow rates (50-100 μl/min) and a low surface densities are therefore recommended for the reduction of these effect and highest data quality. However, the flow rate is often limited by the sample consumption or the required injection time. Experimental series with variation of surface density and flow rate could therefore be helpful for the detection of such effects and for finding a suitable compromise between sample consumption, signal intensity and mass transport [38].

In order to clean all parts of the injection system and to equilibrate the surface, a series of buffer blanks should be injected before each experiment [39]. Injection modes especially designed for highest volume accuracy and high-resolution dissociation phases (*kinject* command) should always be used for sample injections during analyte screening. Injections of buffer blanks before and within binding experiment, inclusion of positive and negative controls, washing steps, as well as a proper maintenance further increase the accuracy and quality of the binding data [13]. Sample injection should be done randomized and in replicates to eliminate the total experimental noise.

Regeneration is one of the most critical parts of a binding assay, especially when dealing with proteins. Too soft conditions lead to remaining analyte and a possible carry-over effect, while too harsh conditions might denature the protein. Specific methods like the removal of calcium ion in the case of C-type lectins are always preferred to unspecific approaches (acidic, basic or chaotropic conditions, detergents and high salt concentrations). Sometimes a cocktail of different regeneration compounds is needed and approaches to find a suitable combination are described in literature [40]. To avoid any carry-over of the regeneration solution a buffer blank injection should be performed at the end of each cycle [39].

2.1.6 Data analysis

Although generating Biacore data is fairly easy, the accurate interpretation of the equilibrium, kinetic and thermodynamic data has proven to be more difficult. Deviations from an expected binding model do not always represent a more complex interaction but are often caused by experimental design.

Data processing should be done in an accurate and reproducible way in order to remove matrix and bulk effects of the binding signals. This is especially necessary when working with small molecules, since even small changes of the signal might lead to variations in the binding constants. Therefore, advanced processing steps like blank subtraction (double referencing) should be performed to remove even minor experimental errors [39]. If no literature data are available about an interaction, data should be fitted to a simple 1:1 binding model first (*Eq.* 5). Since some targets posses more than one binding site, the equation has to be extended to a two independent binding site model. If mass transport effects (see *section* 2.1.5) are suspected or reported, a mass transport coefficient (k_m) might be introduced (*Eq.* 6).

$$A + B \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} AB \qquad \text{[Eq. 5]}$$

$$A_0 \underset{k_m}{\overset{k_m}{\rightleftharpoons}} A + B \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} AB \qquad \text{[Eq. 6]}$$

Unfortunately, using the sum of two or more equations or increasing the number of parameters will almost always lead to a better fit, regardless of the underlying binding mechanism [38]. Careful validation with additional experimental or literature data is therefore recommended before relying on a new binding model. Additional models for surface heterogeneity or a drifting baseline are available in the evaluation software. Even though a better fitting might represent a real effect on the surface, more time should be invested to avoid such drifts or heterogeneities by changing the experimental setup.

A proper data processing is especially important for fitting kinetic data. Initially, different algorithms using curve transformation [41] or nonlinear last squares analysis [42] were used for the evaluation of the binding kinetics. However, these methods only fitted single binding curves (or even portions thereof) and were found to be often insufficient to discriminate between different binding mechanisms [38]. In the *global analysis* approach, the association and dissociation phases of the entire data set are fitted to a model simultaneously, resulting in very accurate and robust data [43]. Therefore, this method is implemented in the current evaluation software tools and should always be used for kinetic fits.

2.1.7 Applications of Biacore in drug discovery

The analysis of molecular interactions is a key part of the drug discovery process. Though the scientific community and pharmaceutical industry first hesitated to accept SPR-based interaction studies [12], Biacore instruments and similar biosensors were validated as an important biophysical method and are now well integrated. Currently, these instruments are used in nearly every aspect of the drug discovery process, from target identification, compound screening and lead optimization to supporting clinical trials, regulatory approval and biopharmaceutical manufacturing (*Fig*. 2-12).

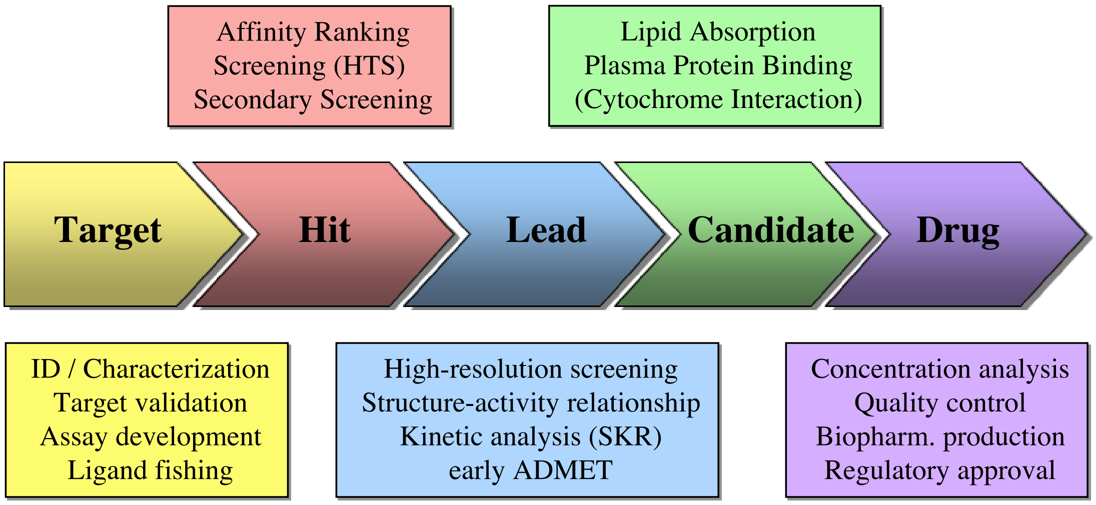

Affinity Ranking
Screening (HTS)
Secondary Screening
Lipid Absorption
Plasma Protein Binding
(Cytochrome Interaction)
Target
Hit
Lead
Candidate
Drug
ID / Characterization
Target validation
Assay development
Ligand fishing
High-resolution screening
Structure-activity relationship
Kinetic analysis (SKR)
early ADMET
Concentration analysis
Quality control
Biopharm. production
Regulatory approval

Figure 2-12: Application examples for Biacore 3000 and other SPR biosensor platforms in drug discovery. (ID = identification, ADMET = pharmacokinetics and toxicology, HTS = high throughput screening, SKR = structure-kinetic relationship).

Genomic and proteomic investigations generated new dimensions of possible disease-related targets. However, their characterization and validation is often difficult and time consuming. Biacore assays might be performed to get first qualitative information e.g. for the identification of key binding subunits in a multiprotein complex [44]. By combining Biacore with mass spectrometry (BIA-MS) it is possible to identify unknown ligands binding to a target of interest. In an approach often referred as ligand fishing, crude samples (plasma, synovial fluid, etc.) are injected over the immobilized target. Captured ligands can either be eluted, digested on the chip and analyzed by electrospray MS or the chip can be removed from the instrument and directly used for MALDI-MS [45] (*Fig*. 2-13). A similar method can be applied for identifying unknown target proteins by capturing them on a chip coated with known ligands. For example, *Graffinity Pharmaceuticals AG* [46] is specialized on screening for new targets or generating interaction fingerprints of known proteins against small molecule-arrays using SPR technology [47].

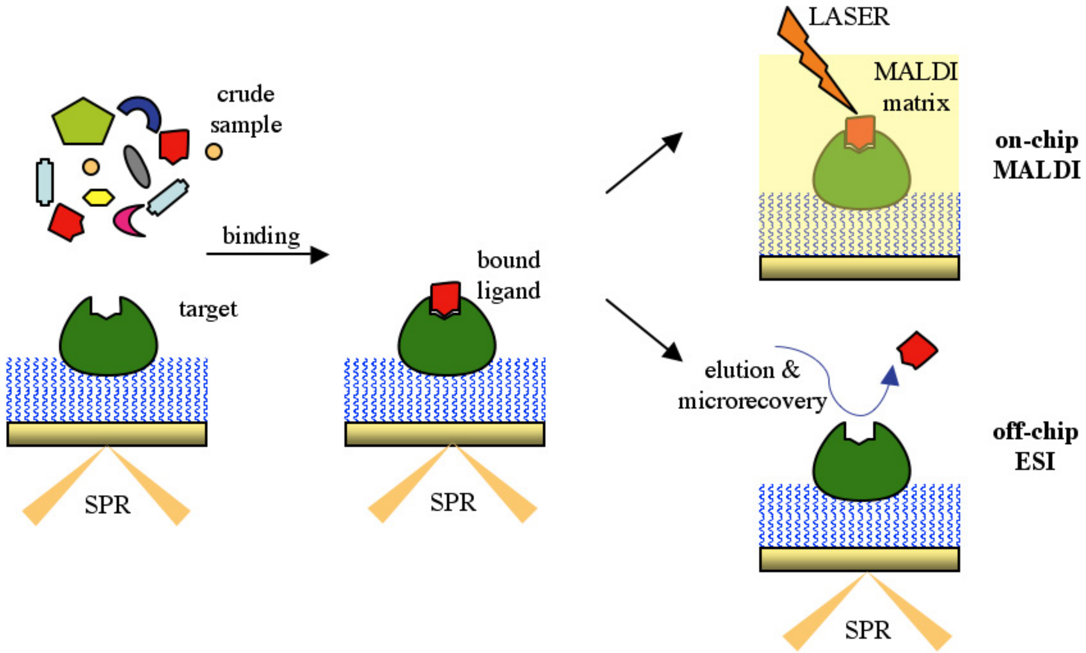


Figure 2-13: Combination of Biacore with mass spectrometry (BIA-MS) allows to capture ligands out of a crude sample and identify them by MALDI-MS on the removed chip or by ESI-MS after elution and recovery of the ligand (illustration adapted from Nedelkov *et al*. [45]).

After a new target is identified and characterized, assays for high-throughput screening have to be developed. Since Biacore offers real-time monitoring of biomolecular interaction, changes in assays conditions can be investigated very quickly and this information can be transferred to ELISA or fluorescence assays. While those types of assays are better suited for (ultra)high-throughput screening, they often generate a certain amount of false positive and negative results. One of the major applications of SPR assays in drug discovery is therefore the *secondary screening* of interesting or uncertain hits without the interference with any labels, reporter enzymes or intrinsic fluorescence.

Once a lead compound is selected, high-resolution assays can be performed to get more detailed information about thermodynamics and kinetics of an interaction. Variations of association and dissociation rate constants on changes of the lead scaffold can be examined for the generation of structure-kinetic relationships. Two types of graphs were found to be useful: k_{on}/k_{off} plots (*Fig*. 2-14A) and response/stability plots, where the response at the equilibrium is plotted against the signal after a certain dissociation time (*Fig*. 2-14B).

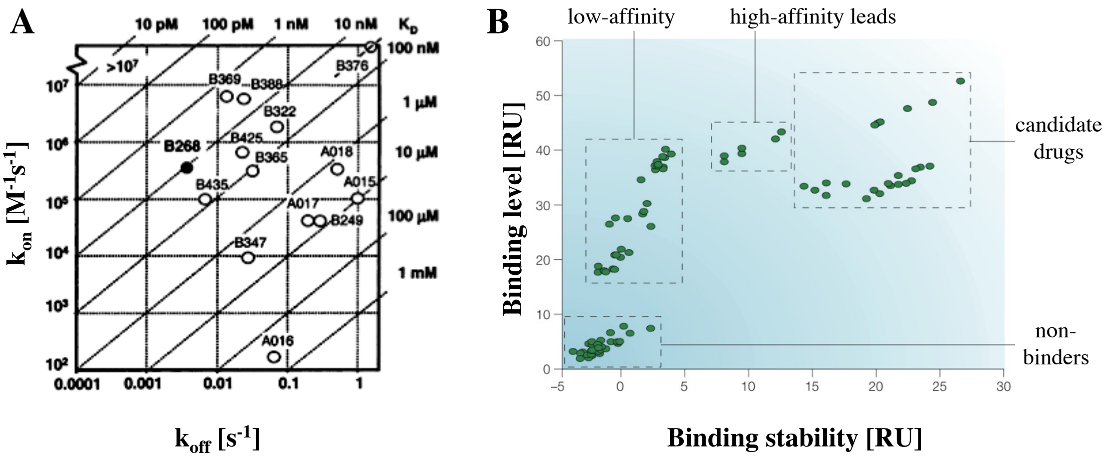


Figure 2-14: Methods for representing kinetic data. **A**: k_{on}/k_{off} plot, where the on and off rates are plotted on the two axes and the K_D can be determined from the diagonals (image from Markgren *et al.* [48]). **B**: Response/stability plot, in which the signal intensities at the binding equilibrium are plotted against the remaining signal at a given time point of the dissociation phase (image from Cooper [44]).

More than half of the drugs currently on the market are binding to G-protein-coupled receptors (*GPCR*), ion channels or other transmembrane receptors (*Fig*. 1-2). Even though sensor chips for the attachment of lipid mono- and bilayers are available, the development of membrane protein assays is extremely difficult. While some surface receptors can be expressed as soluble forms, this is not possible for GPCR. However, some examples of the immobilization and on-chip reconstitution of a GPCR (rhodopsin) are reported [49-52] (*Fig*. 2-15).

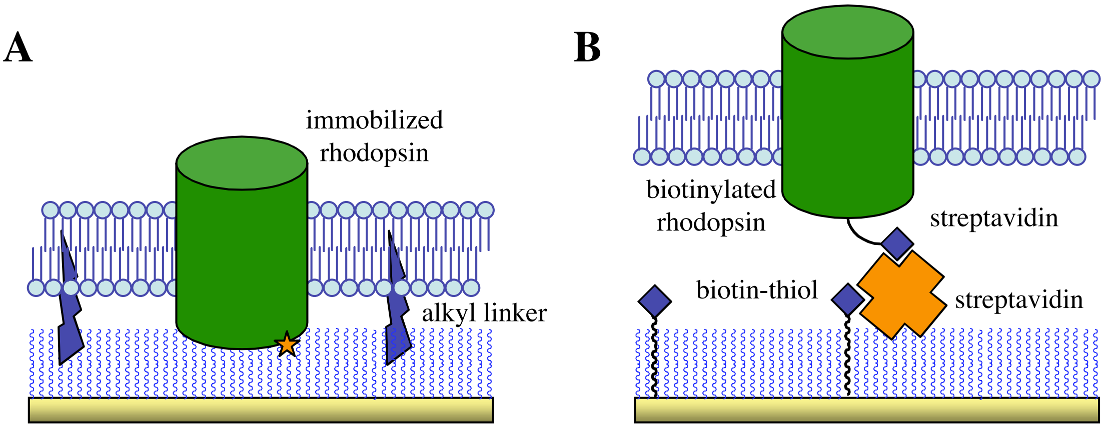


Figure 2-15: Immobilization and reconstitution of a GPCR (rhodopsin) on a sensor chip. **A:** Method of Karlsson *et al.* [49] with covalent coupling on a L1 chip. **B:** Method of Bieri *et al.* [51] using immobilized biotin molecules on a self-assembled monolayer, where streptavidin was captured. For site-directed labeling, biotinylated rhodopsin was used.

Pharmacokinetic characterization of lead and candidate compounds gets increasingly important (see chapter 1) and is performed at early stages of the discovery process (early ADMET). The level of *plasma protein binding* is an important factor in the delicate balance between intended physiological activity, long-term efficacy and potential side effects of a drug [44]. Several assays are reported of drugs binding to human serum albumin (HSA; see chapter 3) or α_1-acid glycoprotein (AGP). Recent publications tend to combine such investigations and plot HSA against AGP binding to get a quick overview of pharmacokinetic properties [11]. *Absorption* and *membrane permeability* can also be estimated by screening drugs against membranes and liposomes, which were captured on a sensor chip [53, 54]. Cytochrome P450 (CyP450) is a key player in drug *metabolism* and *toxicity*. Even though there is no direct binding assay available so far, Biacore technology has successfully been used to determine drug-induced induction of CyP450 mRNA by direct hybridization on immobilized DNA oligomers [55].

Finally, SPR biosensors can be used for diagnostic assays during clinical studies (e.g. determination of antibody serum titers), for monitoring expression levels or for batch control in the manufacturing of biopharmaceutical compounds. For this purpose, hardware (Biacore C) and software had to be adapted to fulfill the requirements for good manufacturing practice (GMP) [12, 44]. The advantages of biosensors in this field are their high accuracy and precision, the minimal time consumption for sample preparation, short assay times and their ability to detect even low affinity antibodies.

Moreover, not only the total concentrations but also the active analyte fraction can be detected [12].

Not only the possible application expanded over the years but also the available instruments. In 2001, Biacore S51 was introduced. This instrument was especially designed for the detection of small molecules at higher throughput. For this purpose, the number of flow cells was increased from four to six while reducing the individual flow cell area from 1 mm^2 to 0.1 mm^2. A hydrodynamic flow system lacking any valves replaced the conventional fluidic cartridge leading to much cleaner sensorgrams, especially during dissociation phase. With the possibility to use 384-well plates and the completely automated (wizard-driven) software, throughput was increased [11, 56]. New developments in the field of SPR biosensors go into direction of microarray formats, where the detection of multiple spots at the same time is possible. Applied Biosystems Inc. introduced the first of these SPR arrays (8500 affinity chip analyzer) in 2003, which is capable of analyzing up to 400 spots on a single large flow cell but has a detection limit of only 5000 Da [12]. Biacore recently announced to release a similar system in early 2005 [12, 57].

2.1.8 Biophysical methods used in drug discovery

Apart from traditional high-throughput screening mostly based on fluorescence read-outs, biophysical methods still are an invaluable area in drug discovery. They allow much closer insight into drug-target interactions and are an alternative where HTS is not possible. While several methods are also generating information about binding affinity or kinetics and are therefore competing with SPR technology, others can deliver a closer insight in more specific areas of molecular interactions. In principle, two groups of methods can be differentiated: methods with a main focus on *interaction analysis* and methods from *structural biology* that also provide information about binding events (*Table* 2-3).

Among the group of interaction analysis methods, the merit of *SPR biosensors* are their lack of any labeling, the real-time detection, independence on any spectroscopic properties, as well as the high degree automation. However, analyzing small molecules can be very challenging since the signal intensity is directly dependent on the size. In addition, differences between specific and non-specific binding can hardly be distinguished.

Table 2-3: Comparison of biophysical methods in drug discovery with their analytical focus and some important properties.

Method	Main application focus	Label-free	Buffer-based	Non destructive	Membrane systems
SPR	Affinity, kinetics	+	+	+	+/-
ITC	Affinity, thermodynamics	+	+	+	-
AUC	Size, stoichiometry, affinity	+	+	+	-
SFS	Kinetics	+/-	+	+	-
DPI	Density, layer thickness, size	+	+	+	+
QCM	Mass changes, viscosity	+	+	+	+/-
MS	Structure, identity	+	+/-	-	+/-
NMR	Structure, dynamics	+/-	+	+	-
X-ray	Structure	+	(+)	-	-
AFM	Structure, imaging	+	+	+	+
CD	Structure, conformation	+	+	+	+

Isothermal titration calorimetry (ITC) clearly overcomes the size limitations, since it detects small changes in temperature caused by the absorption or release of heat upon molecular interaction. It is also label-free, does not require an immobilization (solution-based) and is the only method to differentiate between entropic (ΔS) and enthalpic (ΔH) binding components [58]. Applications in drug discovery and life science range from protein folding and protein-protein interactions to the analysis of protein binding of small molecules [59]. The ability to get equilibrium binding data makes it a valuable tool for the validation of Biacore results, especially when an influence of immobilization effects is suspected. On the other hand, no kinetic information is provided and the sample consumption is much higher (often in the milligram range for both target and analyte). *Analytical ultracentrifugation* (AUC) is not primarily a method for validating SPR results but generates useful supportive information about the molecular mass, size, and shape of molecules in solution as well as about binding stoichiometry and association energy. Improved data analysis and the integration of fluorescence detection greatly increased the sensitivity and AUC sees therefore a revival in drug discovery applications [60]. With *dual-polarization* *interferometry* (DPI) a new detection technique for optical biosensors is under development. With DPI changes in surface density and thickness can be detected, which makes it especially interesting for the validation of immobilization processes [61]. Since many protein-ligand interactions induce a conformational change, such binding events might be quantified using DPI. However, only proof-of-concepts studies are available in this area at the moment [62]. Another way for a label-free detection of interactions is the use of *quartz crystal microbalances* (QCM), where frequency shifts of a piezoelectric crystal upon changes in mass, viscosity, stiffness, conductivity, and dielectric constant around a surface are measured. Even though the detection is not very selective and analysis in liquid environments might be challenging, some application in drug/small molecule detection are described [63, 64].

There are far less alternative methods available for getting kinetic information compared to equilibrium binding. *Stopped-flow spectroscopy* (SFS), which can be coupled to UV/VIS and fluorescence spectroscopy but also to NMR [65] and MS [66], is the most important technique in this area. Two (or more) solutions are prepared in separate syringes and simultaneously injected into a mixing chamber under high pressure. The flow is then suddenly stopped and the reaction is measured at different time points in the observation chamber (e.g. by fluorescence quenching) [67]. SFS is often used for the investigation of protein folding and other macromolecular changes but can also be applied to small molecules.

Some of the methods from structural biology can also be used to get binding parameters. For example, *mass spectrometry* (MS) was used predominantly for the analysis of small molecules for many years and only the development of soft ionization methods (ESI, MALDI) allowed its application for the characterization of biomacromolecules. The combination of these methods with two-dimensional SDS-PAGE initiated one of the most important fields in drug discovery, called proteomics, where the identity and abundance of disease-related proteins can be determined from tissue extracts, body fluids, etc. Besides protein identification, information about protein structure, posttranslational modifications or binding epitopes can also be obtained [68]. Newer developments made it possible to characterize enzyme/substrate interactions [69] or to get relative binding affinities of drug mixtures binding to a protein (in combination with micro-size exclusion chromatography) [70]. Mass spectrometry can also be combined with Biacore instruments (BIA-SPR) for ligand-fishing experiments (see *section* 2.1.7). One of the most powerful and emerging biophysical methods in drug discovery, however, is *nuclear magnetic resonance* spectroscopy (NMR). The unique feature of NMR is its ability to get information about both protein structure and dynamics. Therefore, NMR is not only used as an analytical tool for molecule characterization but also for the description of binding events and the validation of HTS hits. Technological developments like T1-relaxation, trNOESY or STD-NMR opened the methods for the closer evaluation of ligand binding to proteins [71]. For example, STD-NMR was used to determine the binding epitope of the tetrasaccharide sLe^x on E-selectin [72]. NMR is especially sensitive for weak affinities with K_D up to 1-10 mM but has limitation for high-affinity interactions. For some applications there are also limits for proteins with a size of more than 40-50 kDa or a special isotopic labeling is required [71]. In many areas NMR is used complementary with *X-ray crystallography*. Structural data from one (or both) of these methods are a prerequisite for rational drug design, modeling and ligand docking studies. The bottleneck of this technology has been the preparation of suitable crystals for X-ray analysis. However, a great deal of effort has recently been invested to develop methods that allow high-throughput crystallography [73]. Structural data from X-ray and NMR are very helpful for the development of Biacore assays, e.g. for the selection of suitable immobilization methods or pre-selection of interesting analytes.

Atomic force microscopy (AFM) is primarily an imaging tool, but has an interesting potential for single molecule manipulation as well as in the analysis of intra- and intermolecular binding forces. The method uses a thin sharp tip attached to a cantilever for scanning a molecular probe (proteins, complexes, cells) positioned on a piezoelectrical crystal, which moves the sample. Bending of the cantilever during scanning is detected by a reflected LASER beam resulting in a topological map of the probe [74]. Since the analysis can be performed under near-physiological conditions, there are many applications in drug discovery [75] and the method was e.g. used for monitoring the growth of amyloid fibrils in Alzheimer's disease [76]. Force curve experiments can be performed using AFM in order to determine the binding strength of molecules, which was applied to the characterization of the biotin-streptavidin interaction [77]. In addition, cantilever technology can be used in a different way for the determination of binding events. Cantilever arrays are now under development, where one binding partner is fixed on an oscillating cantilever tip and changes in the amplitude are detected upon binding of analytes in solution [78].

Circular dichroism (CD) is another biophysical method from structural biology with possible applications in the analysis of molecular interactions. It detects the unequal absorption of left- and right-handed circularly polarized light by optically active molecules. Even small changes in the secondary structure (e.g. helicity) and conformation of proteins can be monitored [79]. Even though this technique is usually used to monitor protein folding, denaturation or conformational changes induced by temperature, *pH* and other environmental factors, it can also be used for the evaluation of binding constants, or to determine the number (and location) of amino acids involved in the binding event [80]. However, since the absorption is monitored in the near and far UV range, buffer components, ligands or other additives might interfere with the signal and even optically inactive analyte show a signal when fixed in the binding site.

2.1.9 Comparison of SPR technology (Biacore) with other methods

While the unique features of SPR-based biosensors were acknowledged rapidly in life science research as well as in pharmaceutical industry, there was much more skepticism whether biosensor data would match the results from solution-phase methods [12]. Biosensor-based reaction constants are obtained from surface-based experiments and their reliability was initially questioned due to a variety of potential artifacts [38]. Target immobilization could lead to restrictions in its rotational freedom and accessibility, which could affect binding parameters (affinity, kinetics). In addition, the analyte has to be transported to and from the target surface in a rapid and uniform manner to avoid concentration gradients at the surface [81].

Meanwhile, different studies have been performed, which compare Biacore data with other biophysical methods, such as ITC, AUC and SFS. Carbonic anhydrase II was chosen as a target in two of these studies [82, 83], because of its good characterization, commercial availability and the formation of simple 1:1 complexes with arylsulfonamide compounds. In both studies, equilibrium and thermodynamic parameters (SPR and ITC) as well as kinetic constants (SPR and SFS) were in very good agreement, showing that immobilization and sample delivery did not have a significant influence on thermodynamic properties and data quality. The MIRG'02 study (see also *section* 2.3) [83] also compared ITC and SPR, but included AUC for the analysis of molecular mass, homogeneity and assembly state. They not only tested comparability but also reproducibility, since the same pair of target and analyte was provided to several independent groups. Again, an excellent agreement for the binding parameters was seen between the methods. In addition, having a universal and detailed protocol as well as applying careful experimental handling were found to be essential for a comparable and reproducible assay with all technologies [83].

In general, interaction constants obtained from Biacore experiment are not different from those obtained by solution-based methods. However, especially the immobilization step might lead to inactivation or structural changes of the binding site. Therefore, Biacore results should always be validated, if a suitable method is available. This could be done by comparing the results with data from other direct binding assays or by measuring reference compounds known from literature. Result from competition or inhibition assays (ELISA, fluorescence assays) might also give an idea about the reliability of Biacore data, but should be used with care.

Biophysical methods are not only important and useful for data validation and the development of Biacore assays. All of them have special foci, strengths and drawbacks. Combining Biacore results with data from other methods may contribute to a very detailed and accurate picture of a binding event. Therefore, those methods should be regarded as complementary rather than alternative, and their integration in the drug discovery process is increasingly important for many stages from target and hit validation to lead optimization.

2.2 Materials and Methods

2.2.1 Materials

Reagents and proteins

NHS, EDC and ethanolamine solutions (prepared from the amine coupling kit), PDEA, immobilization buffers (10 mM sodium acetate *pH* 4.0, *pH* 4.5, *pH* 5.0, and *pH* 5.5), 10 mM glycine regeneration solutions (*pH* 1.5 - 3.0) as well as *BIAdesorb*, *BIAdisinfectant*, *Normalizing*, and *Test* solutions (maintenance kit) were directly purchased from Biacore AB (Freiburg i. Br., Germany). All other reagents were from Sigma (Fluka Holding AG, Buchs, Switzerland).

Equipment

All SPR analysis were performed on a Biacore 3000 system using research grade CM5 sensor chips (Biacore AB, Freiburg i.Br., Germany). The system was additionally equipped with a Thermo Haake C10/K10 water bath system (Digitana AG, Horgen, Switzerland) for temperature control of the Biacore 3000 autosampler. All vials, caps, and other consumables were directly ordered at Biacore AB. A Branson 2510 ultrasonic water bath (Merck Schweiz AG, Dietikon, Switzerland) and a vacuubrand MZ-2C vacuum pump cooled by a Huber polystat cc1 system (E. Renggli AG, Rotkreuz, Switzerland) were used for buffer degassing. Buffer *pH* values were controlled using a Metrohm 691 *pH*-meter equipped with a combined *pH* glass electrode with built-in temperature probe (No. 8.109.1236; Metrohm AG, Herisau, Switzerland).

Software

All Biacore results were acquired using the instrument-bundled software *Biacore control 3000* (version 3.1). Data processing and steady state analysis were performed in *BIAevaluation* software (version 4.0; Biacore AB, Freiburg i.Br., Germany) or in *Scrubber* (BioLogic Software Pty Ltd., Campbell, Australia). *BIAevaluation* or *CLAMP XP* (Center for Biomolecular Interaction Analysis, University of Utah, USA) [84] was used for kinetic analyses. *Prism* (GraphPad Software Inc., San Diego, USA) was used for the generation and fitting of some data plots. Certain curve processing steps, method generation and calculations were done in Microsoft *Excel 2000* (Microsoft Schweiz GmbH, Wallisellen, Switzerland).

Visualizations of crystal structure data were prepared in an open-source version of *PyMol* for MacOS X (version 0.97; DeLano Scientific, San Carlos, USA) [85].

2.2.2 Preparation of running buffers

All buffers used for Biacore experiments were filtered to reduce particle load and avoid clogging of the IFC (using nitrocellulose membranes with a pore size of 0.44 μm). In addition, buffers were degassed every day by keeping them in an ultrasonic bath for at least 10 min under reduced pressure (< 50 mbar). Biacore experiments were conducted at 25°C unless otherwise noted. Samples were filled in 7 mm polypropylene vials, capped, and centrifuged before each run. Autosampler racks were kept at a constant temperature of 20°C for reducing evaporation effects.

2.2.3 Instrument maintenance procedures

To ensure maximum instrument lifetime but also high data quality, maintenance procedures were applied at a higher frequency than recommended by Biacore. A correlation between maintenance and data quality was for example demonstrated in Cannon *et al.* [86]. *Desorb* routines (cleaning of the instrument with 0.5% SDS followed by 50 mM glycine *pH* 9.5) were performed at least once per week but also between experiment series and at first indications of decreased data quality. Microbial growth was inhibited by applying the *Sanitize* procedure at least monthly using diluted BIAdisinfectant solution (sodium hypochlorite). In addition, the system was rarely turned off but kept under constant flow conditions (*run* or *standby*). Manual cleaning of the needle, syringes, and the injection port were done on a regular basis. Instrument performance was tested using the internal system check routine.

2.2.4 Initial preparation of new CM5 chips

In order to remove minor impurities from the chip surface, each CM5 chip was treated with a selection of regeneration solutions prior to use (preconditioning). For this purpose, a new sensor chip CM5 was inserted and primed three times with water. At a flow rate of 50 μl/min solutions containing 50 mM NaOH, 10 mM HCl, 0.1 % SDS, and 100 mM phosphoric acid, respectively, were injected twice for 20 s each. The system was then *primed* at least three times with running buffer. To ensure highest sensitivity, the SPR detector response was normalized before running assays, by calibrating the detector at various light intensities under conditions of total internal reflection (automated *normalize* procedure included in Biacore control software; injection of 70% (v/v) glycerol in water).

2.2.5 Amine coupling

The standard amine coupling procedure was based on the descriptions in Johnsson *et al.* [17]. Solutions of 0.1 M NHS, 0.4 M EDC and 1 M ethanolamine were prepared in water as recommended by the manufacturer and stored at -20°C. Aliquots of these solutions were just thawed shortly before use. The *pH* of the immobilization buffer was chosen to reach a balance between the need for the biomolecule to be sufficiently positively charged to be concentrated at the surface yet retain uncharged amino groups for reaction with the activated matrix. This was done by injecting constant concentrations of the target protein at *pH* values at or below its *pI* to a non-activated flow cell and selecting the highest *pH*, which showed sufficiently high attraction signals (*pH scouting*).

Sensor chip surface was activated with a freshly prepared 1:1 mixture of NHS and EDC solution for a specified *activation time* (3-15 min). Proteins to be immobilized were diluted in an appropriate immobilization buffer to a concentration typically in the range of 10-100 μg/ml and injected over the activated surface (immobilization time). Deactivation of remaining active esters on the surface was done by injecting an ethanolamine solution for at least the same time as used for activation (*deactivation time*). Typically, the surface was washed with short pulses of a regeneration solution in order to remove protein, which was bound only by electrostatic interactions. Finally, the surface was allowed to stabilize and hydrate by exposition to pure running buffer for several minutes or even hours, until a stable baseline could be detected.

2.2.6 Data processing and double referencing

Biosensor result files were directly loaded into *Scrubber* for further processing. After the baseline of all flow cells were overlaid just before injection start, sensorgrams were cropped and injection start points were aligned. The sensorgram of an empty flow cell was subtracted from all the other sensorgrams (*referencing*). To reduce systematic errors in curve shape, the average of an ensemble of blank injections was subtracted from the binding curves (*double referencing*). Finally, steady state signals of all sensorgrams were plotted against their concentration and fitted to an appropriate binding model.

2.2.7 Preparation of thiol surfaces

CM5 chips were preconditioned as described above. Standard amine coupling (see *section* 2.2.5) was used for the immobilization of a thiol-bearing molecule. After activation (5 min) 40 mM cystamine dihydrochloride in 0.1 M borate buffer *pH* 8.5 was injected. The surface was then reduced by injecting a 3 min pulse of 0.1 M DTT in 0.1 M borate buffer *pH* 8.5 and deactivated for 5 min. This procedure typically yielded in a baseline increase of 100-250 RU. Substitution of cystamine by GSSG was also tried but showed much lower yields (≤ 80 RU) despite the higher molecular weight (data not shown).

2.2.8 Instrument validation (ABRF-MIRG'02 study)

Through the participation in an international study [83] initiated by the *Molecular Interactions Research Group* (MIRG) of the *Association of Biomolecular Research Facilities* (ABRF), the instrument could be validated and the results could be compared with 28 other Biacore laboratories. Each participant was provided with vials containing 75 μg freeze-dried bovine carbonic anhydrase II (CA II; *Fig*. 2-16A) and 0.6 ml of a 2 mM 4-carboxybenzenesulfonamide (CBS; *Fig*. 2-16B) solution in PBS. A detailed assay protocol describing all steps from instrument cleaning and buffer preparation to data analysis was also provided.

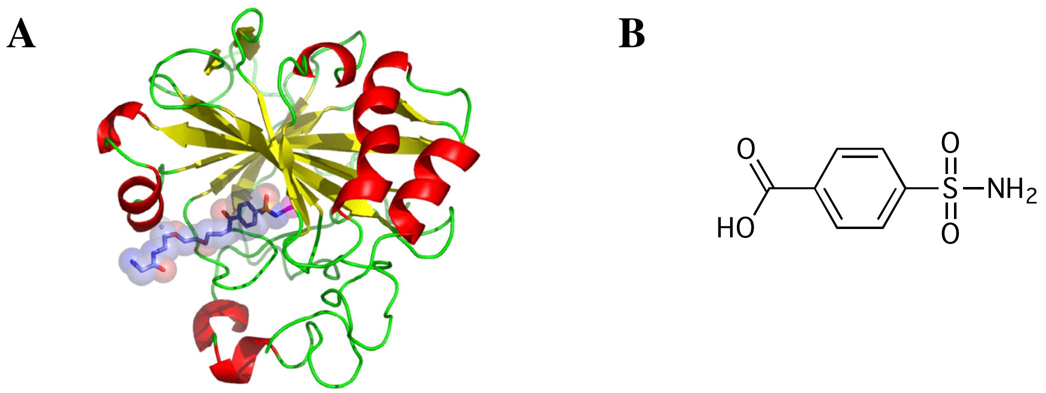


Figure 2-16: Interaction system used in the ABRF-MIRG'02 study. **A:** Crystal structure of carbonic anhydrase II in complex with the sulfonamide inhibitor EG1 (PDB entry: 1CNW). **B:** Structure of the analyte 4-carboxybenzenesulfonamide (CBS).

A tenfold phosphate buffer stock was prepared by mixing 0.2 M solutions of each NaH_2PO_4 and Na_2HPO_4 in a ratio of 1:4 (v/v) and adjusting the *pH* to 7.4. Final 20 mM PBS running buffer was obtained by a tenfold dilution of the stock solution, addition of NaCl to a concentration of 150 mM and degassing (see *section* 2.2.2) just before use. A new CM5 sensor chip was primed several times with water. After an initial *desorb* step (see *section* 2.2.3) and chip preconditioning (see *section* 2.2.4), the system was primed three times with freshly degassed PBS and the *normalize* procedure was performed. To test the instrument cleanness and performance as well as to further equilibrate the system, 30 buffer blanks were injected at a flow rate of 100 μl/min using the instruments *kinject* command (1 min injection, 3 min dissociation).

CA II was reconstituted in 10 mM sodium acetate *pH* 5.0 to a concentration of 0.125 mg/ml. Standard amine coupling (see *section* 2.2.5) was used to generate CA-II surfaces (10 μl/min, 7 min activation/deactivation, 10 min immobilization). One high-density (11 kRU) and one medium-density (7.5 kRU) surface was created in this way. Unmodified flow cells served as reference and control surfaces.

Six CBS solutions between 0.08 and 20.0 μM were prepared by threefold serial dilution of the CBS stock solution in running buffer. Each sample was dispensed into triplicate single-use, snap-capped plastic vials and randomized (order according to the protocol) in the autosampler block. Samples were injected at 100 μl/min for 1 min (3 min dissociation) in three series separated by a blank. Nine additional blanks were injected before the run but excluded from data processing. Since bound CBS dissociated completely from the surface within the dissociation time, no regeneration step was required. Data collection rate was set to 'high' for all experiments and data processing was performed in *BIAevaluation* software. The complete dataset was baseline-averaged, cropped and aligned manually. Sensorgrams of the reference surface was subtracted from the sample cell for each injection separately. An average sensorgram of all intermediate blanks was used for double referencing. Equilibrium and kinetic rate constants were calculated by fitting processed sensorgrams to a simple bimolecular reaction model in both *CLAMP* and *BIAevaluation*. Results obtained in this way and raw data were sent to the study coordinator and compared with the whole panel of Biacore users [83].

2.3. Results and Discussion

2.3.1 Instrument validation (ABRF-MIRG'02 study)

Working with small molecules requires special care and procedures through all the steps from experimental setup to data evaluation. Even small signal errors can have a dramatic influence on the results of equilibrium or kinetic binding experiments. Furthermore, there was much concern, whether surface-based methods like Biacore can be really compared to experiments in solution. The ABRF-MIRG study was therefore initiated in order to investigate the reproducibility of a small-molecule assay among different laboratories and instruments from all over the world and to compare the results with solution-based methods. A participation in this study not only allowed to contribute to this important question but also to have a system for internal system validation.

One of the key features of this study was that every participant was provided with exactly the same batch of protein and analyte (CA II and CBS), as well as with a detailed assay protocol. This eliminates many of the systematic errors in data acquisition and evaluation, but also reduces artifacts due to sample impurities.

Since instrument purity is one of the prerequisites to obtain reliable data, the reproducibility and scattering of repeated buffer injections was first investigated (*Fig*. 2-17). Even though the chip surfaces were unmodified, raw data of the 30 blank injection was not flat but showed a typical curve shape with a rapid signal increase to 10 RU and a constant decrease to around -15 RU suggesting pressure or matrix effects to be involved (*Fig*. 2-17A). Subtracting the signals of a reference flow cell (*referencing*), which is the standard procedure recommended by Biacore, obviously increased the signal quality by eliminating many on the bulk and matrix effects but blank curves still showed deviations from the ideal flat shape (*Fig*. 2-17B). Only the introduction of the double referencing procedure, where the average of all blanks is subtracted from the individual curves, resulted in flat sensorgrams for all flow cells (*Fig*. 2-17C).

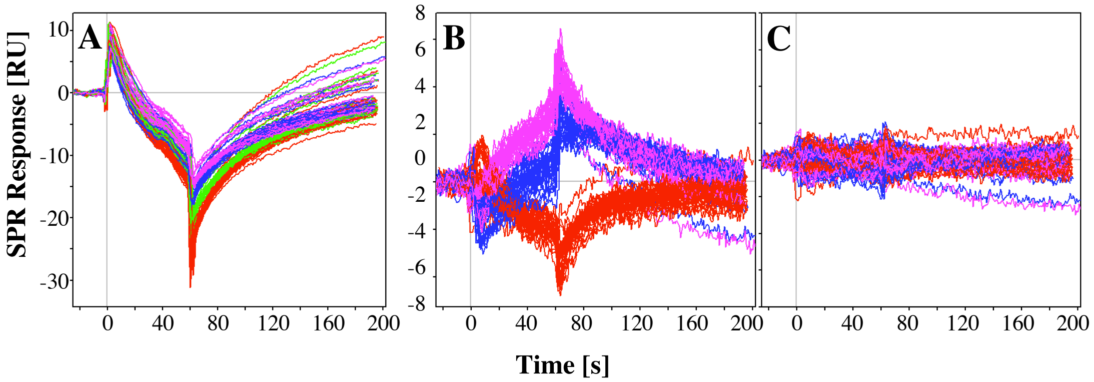


Figure 2-17: Test of instrument cleanness by repeated buffer injections (30 blanks). **A:** Raw data after baseline normalization, alignment and cropping. **B:** Blanks after referencing (subtraction of reference flow cell). **C:** Blanks after double referencing (average of all blanks subtracted from each individual blank). Flow cells 1 (red), 3 (blue) and 4 (magenta) were used as control cells, flow cell 2 (green) as reference cell.

Double referenced signals showed only small deviations from the zero baseline of approximately ±1 RU, which is mainly caused by the instruments noise (±0.3 RU [13]). Only the first injection showed a slightly higher deviation, which could be explained by impurities in the autosampler system, lack of equilibration or other artifacts. Comparison of all study data also confirmed that responses collected early in the run often showed a certain drift that was eliminated over time [83]. Therefore, these results clearly showed that reproducible flat blank signals could be obtained by applying the additional double referencing step and that the instrument was clean and in good shape.

The same data processing procedure was then applied to the sensorgrams of the CBS injection series. Here again, only double referencing generated high quality data and flat blank signals (*Fig*. 2-18A-C). Simple referencing resulted in proper triplicates but the curve shape implied that the binding equilibrium was not yet reached even at high concentrations (*Fig*. 2-18B).

High-density and medium-density surfaces showed the same binding behavior (*Fig*. 2-18D&E) and no binding could be detected on the unmodified control surface (*Fig*. 2-18F). Sensorgrams processed in this way could easily be fitted to a simple bimolecular binding model either by a global kinetic fit of the whole curve or by equilibrium binding plots (*Fig*. 2-19).

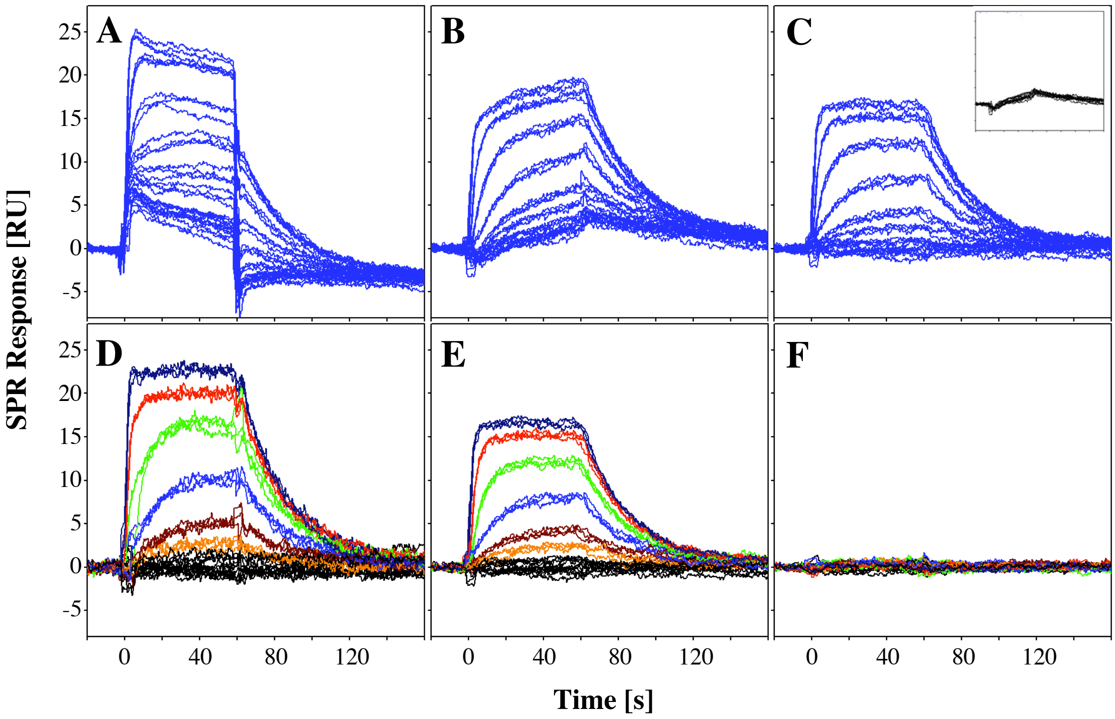


Figure 2-18: Data processing of CBS-CA II sensorgrams. **A:** Raw data after baseline normalization, alignment and cropping. **B:** Data set after referencing (signal subtraction of reference flow cell). **C:** Data set after double referencing, i.e. subtraction of blank average (insert). Sensorgrams of the high-density (**D**), medium-density (**E**) and control surface (**F**).

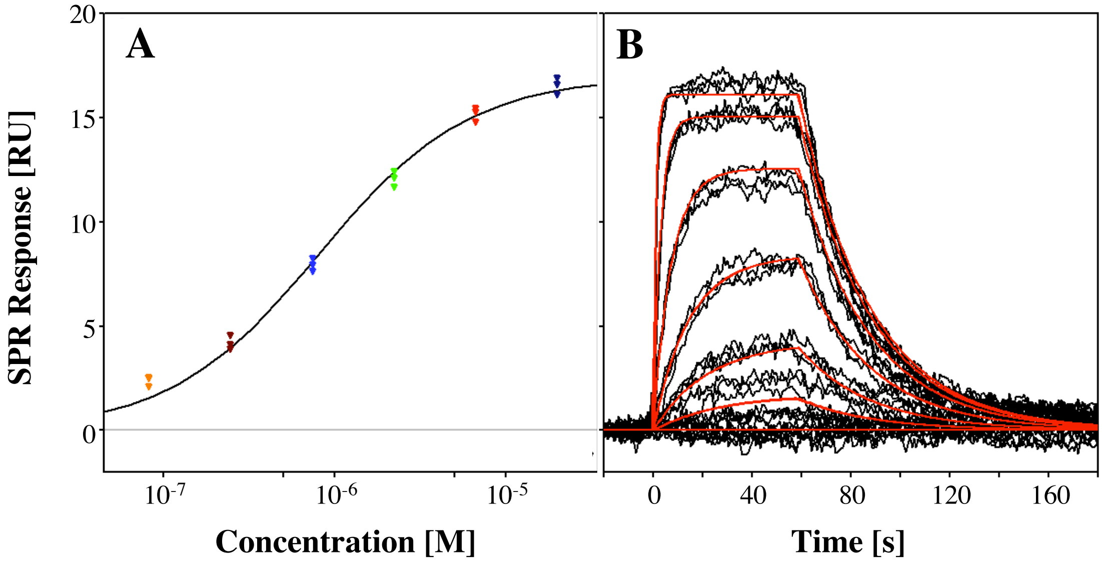


Figure 2-19: CBS-CA II Data Analysis (medium-density surface). **A:** Determination of the equilibrium dissociation constant by fitting steady state response to a simple bimolecular binding model. **B:** Kinetic analysis of sensorgrams by a global simultaneous fit of k_{on} and k_{off} in CLAMP. Simulated data from the fitting algorithm (red curves) are superimposed to the processed binding data (black sensorgrams).

While the whole data processing was performed in *BIAevaluation*, *CLAMP* was used as alternative software for the kinetic analysis. Even though the signal intensity was rather low (with maximum responses around 20 RU) triplicate injections were fully superimposable and fitted very well to the simple bimolecular binding model (*Fig.* 2-19). Only the response of the lowest concentration was slightly higher than calculated. The results were then compared with the averaged equilibrium and kinetic Biacore data from the ABRF study as well as with the K_D value from ITC experiments (*Table* 2-4).

Table 2-4: Equilibrium and kinetic parameters of the CBS/CA II binding compared to the average of the ABRF-MIRG'02 study [83].

Parameter	**Own results SPR**	**Study Average SPR**	**Study Average ITC**
Replicates [a]	6	59	14
K_D [μM]	0.86±0.14 [b] 0.93±0.11 [d]	0.90±0.22 [b]	1.00±0.22 [c]
k_{on} [$M^{-1}s^{-1}$]	$4.1\pm0.5 \times 10^4$	$4.0\pm0.7 \times 10^4$	-
k_{off} [s^{-1}]	0.034±0.003	0.036±0.007	-

[a] Number of complete analyses (triplicate injections). [b] Value calculated by $K_D=k_{off}/k_{on}$, [c] Value calculated by $K_D=1/K_A$, [d] Value determined from the equilibrium plot (simple bimolecular binding model). Separate data from all six replicates (own results) can be found in *appendix* A.

All values from our own experiments were very close to the study average and well within the standard deviations, both for kinetic as well as for equilibrium data. The K_D value from the fit of equilibrium plots was only slightly higher than the K_D calculated from the kinetic rate constants. The variation of the binding parameters within the entire panel of data in the study (59 data sets) was reported to be less than 20%, which is rather low considering the large number of different users and instruments [83]. This clearly illustrates that Biacore results from different laboratories are very reproducible if samples and protocols are standardized. Furthermore, the study also demonstrated the validity of surface-based assays by comparing SPR data to those derived from ITC. Both methods resulted in the same standard deviation and their K_D values matched within the experimental error. Immobilization to the sensor surface might lead to loss of activity but does not seem to alter the thermodynamic binding properties [83]. In a similar study carried out by the same group [86], 36 Biacore users investigated the interaction between azetazolamide and CA II in the same way (identical batches of protein/analyte and protocols). Here again, the reproducibility and accuracy was very high and the largest deviations were found to be the immobilization densities (due to activation reagent quality) and signal artifacts (due to bad instrument maintenance) [86].

2.4 Conclusions

Biacore instruments are versatile tools with a huge potential both in life science and drug discovery. Their ability to generate not only equilibrium binding but also kinetic and thermodynamic data offer a much more detailed insight into a molecular interaction compared to other screening and biophysical methods. Recent developments in the field of data acquisition and processing expanded their applications for the analysis of low molecular weight compounds.

In order to get most accurate data for small molecules on a Biacore 3000, every step from sample preparation and assay design to data processing and maintenance must be performed with care. Unfortunately, no internal test routines or reference systems are available from Biacore to verify the performance and sensitivity of an instrument in this area. An interaction pair consisting of the enzyme carbonic anhydrase II and the sulfonamide inhibitor CBS as used in the ABRF MIRG'02 study showed to be a suitable model systems for this purpose. Both molecules are commercially available in high purity and the interaction follows a simple 1:1 binding model. Even though the molecular weight of CBS (201 Da) is very close to the detection limit, reliable equilibrium and kinetic binding data can be obtained when applying advanced data processing steps.

The participation in the ABRF MIRG'02 study helped establishing a small molecule test system and contributed to the investigation of the reproducibility of Biacore data and their comparability with solution-based methods. The data sets obtained on different SPR instruments by independent groups showed a very high reproducibility and the resulting binding equilibrium data were comparable to ITC. The performance of the instrument used in this thesis was found to be very good when the results were compared with the study average.

2.5 References

[1] T. Turbadar, *Proc Phys Soc (London)* 1959, *73*, 40.

[2] *Surface Plasmon Resonance (Technology Note 1)*, Biacore AB, 2001.

[3] U. Jonsson, L. Fagerstam, B. Ivarsson, B. Johnsson, R. Karlsson, K. Lundh, S. Lofas, B. Persson, H. Roos, I. Ronnberg, *et al.*, *Biotechniques* 1991, *11*, 620.

[4] K. Nagata, H. Handa, *Real-Time Analysis of Biomolecular Interactions: Applications of BIACORE*, Springer, Tokio, 2000.

[5] E. Stenberg, B. Persson, H. Roos, C. Urbaniczky, *J Colloid Interface Sci* 1991, *143*, 513.

[6] T. Mannen, S. Yamaguchi, J. Honda, S. Sugimoto, A. Kitayama, T. Nagamune, *Anal Biochem* 2001, *293*, 185.

[7] S. Lofas, B. Johnson, *J Chem Soc Chem Commun* 1990, 1526.

[8] *Biacore AB, Uppsala, Sweden (www.biacore.com).*

[9] L. G. Fagerstam, A. Frostell, R. Karlsson, M. Kullman, A. Larsson, M. Malmqvist, H. Butt, *J Mol Recognit* 1990, *3*, 208.

[10] M. Malmqvist, *Nature* 1993, *361*, 186.

[11] R. Karlsson, *J Mol Recognit* 2004, *17*, 151.

[12] R. L. Rich, D. G. Myszka, *Drug Discovery Today: Technologies* 2004, *1*, 301.

[13] *Biacore® 3000 Instrument Handbook*, March 1999 ed., Biacore AB, Uppsala, 1999.

[14] Y. Shinohara, Y. Hasegawa, H. Kaku, N. Shibuya, *Glycobiology* 1997, *7*, 1201.

[15] R. Karlsson, M. Kullman-Magnusson, M. D. Hamalainen, A. Remaeus, K. Andersson, P. Borg, E. Gyzander, J. Deinum, *Anal Biochem* 2000, *278*, 1.

[16] *Biacore Sensor Surface Handbook*, Biacore AB, Uppsala, 2003.

[17] B. Johnsson, S. Lofas, G. Lindquist, *Anal Biochem* 1991, *198*, 268.

[18] A. A. Kortt, G. W. Oddie, P. Iliades, L. C. Gruen, P. J. Hudson, *Anal Biochem* 1997, *253*, 103.

[19] M. Morpurgo, C. Monfardini, L. J. Hofland, M. Sergi, P. Orsolini, J. M. Dumont, F. M. Veronese, *Bioconjug Chem* 2002, *13*, 1238.

[20] T. Morcol, A. Subramanian, W. H. Velander, *J Immunol Methods* 1997, *203*, 45.

[21] N. Endo, N. Umemoto, Y. Kato, Y. Takeda, T. Hara, *J Immunol Methods* 1987, *104*, 253.

[22] A. Subramanian, W. H. Velander, *J Mol Recognit* 1996, *9*, 528.

[23] G. A. Canziani, S. Klakamp, D. G. Myszka, *Anal Biochem* 2004, *325*, 301.

[24] R. David, M. P. Richter, A. G. Beck-Sickinger, *Eur J Biochem* 2004, *271*, 663.

[25] I. E. Gentle, D. P. De Souza, M. Baca, *Bioconjug Chem* 2004, *15*, 658.

[26] K. Terpe, *Appl Microbiol Biotechnol* 2003, *60*, 523.

[27] K. M. Muller, K. M. Arndt, K. Bauer, A. Pluckthun, *Anal Biochem* 1998, *259*, 54.

[28] K. Busch, J. Piehler, H. Fromm, *Biochemistry* 2000, *39*, 10110.

[29] A. Einhauer, A. Jungbauer, *J Chromatogr A* 2001, *921*, 25.

[30] A. D. Keefe, D. S. Wilson, B. Seelig, J. W. Szostak, *Protein Expr Purif* 2001, *23*, 440.

[31] L. Jendeberg, B. Persson, R. Andersson, R. Karlsson, M. Uhlen, B. Nilsson, *J Mol Recognit* 1995, *8*, 270.

[32] D. B. Jones, M. H. Hutchinson, A. P. Middelberg, *Proteomics* 2004, *4*, 1007.

[33] *Covalys Biosciences AG, Witterswil, Switzerland (www.covalys.com).*

[34] M. Kindermann, N. George, N. Johnsson, K. Johnsson, *J Am Chem Soc* 2003, *125*, 7810.

[35] W. Huber, S. Perspicace, J. Kohler, F. Muller, D. Schlatter, *Anal Biochem* 2004, *333*, 280.

[36] P. Schuck, *Biophys J* 1996, *70*, 1230.

[37] R. W. Glaser, *Anal Biochem* 1993, *213*, 152.

[38] D. G. Myszka, *Curr Opin Biotechnol* 1997, *8*, 50.

[39] D. G. Myszka, *J Mol Recognit* 1999, *12*, 279.

[40] K. Andersson, M. Hamalainen, M. Malmqvist, *Anal Chem* 1999, *71*, 2475.

[41] R. Karlsson, A. Michaelsson, L. Mattsson, *J Immunol Methods* 1991, *145*, 229.

[42] D. J. O'Shannessy, M. Brigham-Burke, K. K. Soneson, P. Hensley, I. Brooks, *Anal Biochem* 1993, *212*, 457.

[43] L. D. Roden, D. G. Myszka, *Biochem Biophys Res Commun* 1996, *225*, 1073.

[44] M. A. Cooper, *Nat Rev Drug Discov* 2002, *1*, 515.

[45] D. Nedelkov, R. W. Nelson, *Trends Biotechnol* 2003, *21*, 301.

[46] *Graffinity Pharmaceuticals AG, Heidelberg, Germany (www.graffinity.com).*

[47] C. L. Baird, D. G. Myszka, *J Mol Recognit* 2001, *14*, 261.

[48] P. O. Markgren, W. Schaal, M. Hamalainen, A. Karlen, A. Hallberg, B. Samuelsson, U. H. Danielson, *J Med Chem* 2002, *45*, 5430.

[49] O. P. Karlsson, S. Lofas, *Anal Biochem* 2002, *300*, 132.

[50] J. Northup, *Methods Mol Biol* 2004, *261*, 93.

[51] C. Bieri, O. P. Ernst, S. Heyse, K. P. Hofmann, H. Vogel, *Nat Biotechnol* 1999, *17*, 1105.

[52] P. Stenlund, G. J. Babcock, J. Sodroski, D. G. Myszka, *Anal Biochem* 2003, *316*, 243.

[53] A. Frostell-Karlsson, H. Widegren, C. E. Green, M. D. Hamalainen, L. Westerlund, R. Karlsson, K. Fenner, H. van de Waterbeemd, *J Pharm Sci* 2005, *94*, 25.

[54] Y. N. Abdiche, D. G. Myszka, *Anal Biochem* 2004, *328*, 233.

[55] M. Oyama, T. Ikeda, T. Lim, K. Ikebukuro, Y. Masuda, I. Karube, *Biotechnol Bioeng* 2000, *71*, 217.

[56] D. G. Myszka, *Anal Biochem* 2004, *329*, 316.

[57] E. Walladen, in *Biacore Press Release, Vol. 2004*, Biacore AB, 2004, Biacore SPR array.

[58] E. Freire, *Drug Discovery Today: Technologies* 2004, *1*, 295.

[59] P. C. Weber, F. R. Salemme, *Curr Opin Struct Biol* 2003, *13*, 115.

[60] T. Laue, *Drug Discovery Today: Technologies* 2004, *1*, 309.

[61] G. H. Cross, A. A. Reeves, S. Brand, J. F. Popplewell, L. L. Peel, M. J. Swann, N. J. Freeman, *Biosens Bioelectron* 2003, *19*, 383.

[62] M. J. Swann, L. L. Peel, S. Carrington, N. J. Freeman, *Anal Biochem* 2004, *329*, 190.

[63] C. K. O'Sullivan, G. G. Guilbault, *Biosens Bioelectron* 1999, *14*, 663.

[64] K. S. Carmon, R. E. Baltus, L. A. Luck, *Biochemistry* 2004, *43*, 14249.

[65] W. A. McGee, L. J. Parkhurst, *Anal Biochem* 1990, *189*, 267.

[66] B. M. Kolakowski, L. Konermann, *Anal Biochem* 2001, *292*, 107.

[67] P. W. Atkins, *Physical chemistry*, 6th ed., Oxford University Press, Oxford, 1998.

[68] G. L. Glish, R. W. Vachet, *Nat Rev Drug Discov* 2003, *2*, 140.

[69] J. M. Wiseman, Z. Takats, B. Gologan, V. J. Davisson, R. G. Cooks, *Angew Chem Int Ed* 2005, *44*, 913.

[70] P. A. Wabnitz, J. A. Loo, *Rapid Commun Mass Spectrom* 2002, *16*, 85.

[71] H. Widmer, W. Jahnke, *Cell Mol Life Sci* 2004, *61*, 580.

[72] M. Rinnbauer, B. Ernst, B. Wagner, J. Magnani, A. J. Benie, T. Peters, *Glycobiology* 2003, *13*, 435.

[73] T. L. Blundell, H. Jhoti, C. Abell, *Nat Rev Drug Discov* 2002, *1*, 45.

[74] M. Stolz, D. Stoffler, U. Aebi, C. Goldsbury, *J Struct Biol* 2000, *131*, 171.

[75] J. Legleiter, T. Kowalewski, *Drug Discovery Today: Technologies* 2004, *1*, 163.

[76] C. Goldsbury, U. Aebi, P. Frey, *Trends Mol Med* 2001, *7*, 582.

[77] C. Yuan, A. Chen, P. Kolb, V. T. Moy, *Biochemistry* 2000, *39*, 10219.

[78] J. Fritz, M. K. Baller, H. P. Lang, H. Rothuizen, P. Vettiger, E. Meyer, H. Guntherodt, C. Gerber, J. K. Gimzewski, *Science* 2000, *288*, 316.

[79] F. X. Schmid, in *Protein structure : a practical approach* (Ed.: T. E. Creighton), IRL Press, Oxford, 1989, pp. XVIII.

[80] B. A. Wallace, R. W. Janes, *Biochem Soc Trans* 2003, *31*, 631.

[81] D. G. Myszka, X. He, M. Dembo, T. A. Morton, B. Goldstein, *Biophys J* 1998, *75*, 583.

[82] Y. S. Day, C. L. Baird, R. L. Rich, D. G. Myszka, *Protein Sci* 2002, *11*, 1017.

[83] D. G. Myszka, Y. N. Abdiche, F. Arisaka, O. Byron, E. Eisenstein, P. Hensley, J. A. Thomson, C. R. Lombardo, F. Schwarz, W. Stafford, M. L. Doyle, *J Biomol Tech* 2003, *14*, 247.

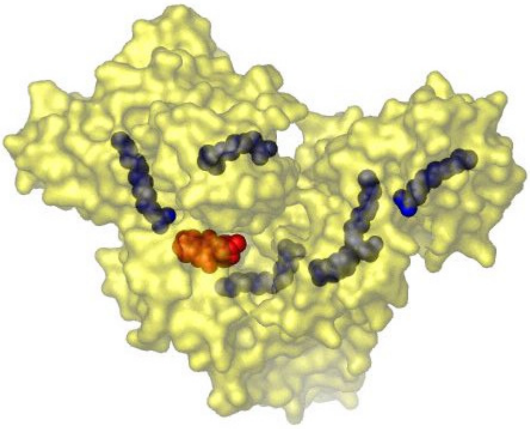
[84] D. G. Myszka, T. A. Morton, *Trends Biochem Sci* 1998, *23*, 149.

[85] W. L. Delano, *Drug Discov Today* 2005, *10*, 213.

[86] M. J. Cannon, G. A. Papalia, I. Navratilova, R. J. Fisher, L. R. Roberts, K. M. Worthy, A. G. Stephen, G. R. Marchesini, E. J. Collins, D. Casper, H. Qiu, D. Satpaev, S. F. Liparoto, D. A. Rice, Gorshkova, II, R. J. Darling, D. B. Bennett, M. Sekar, E. Hommema, A. M. Liang, E. S. Day, J. Inman, S. M. Karlicek, S. J. Ullrich, D. Hodges, T. Chu, E. Sullivan, J. Simpson, A. Rafique, B. Luginbuhl, S. N. Westin, M. Bynum, P. Cachia, Y. J. Li, D. Kao, A. Neurauter, M. Wong, M. Swanson, D. G. Myszka, *Anal Biochem* 2004, *330*, 98.

Chapter 3

Human Serum Albumin



3.1 Introduction

Human serum albumin (HSA) was chosen as a model system for small molecule-binding proteins. It is very well characterized, commercially available, and has a high relevance for drug discovery and development. Many physiological and synthetic binders are known and readily availably, usually covering an affinity range from high nanomolar to low millimolar, the same range many small molecules are binding to surface receptors.

3.1.1 Albumin: A key player in pharmacokinetics and drug development

Drug compounds not only act on the human body (pharmacodynamics) but are also processed by the body when they are administered (pharmacokinetics). First, they have to reach circulation by absorption, permeation and transport processes before they are distributed to tissues and organs. Furthermore, metabolic processes and excretion constantly eliminate the active species (*Fig*. 3-1).

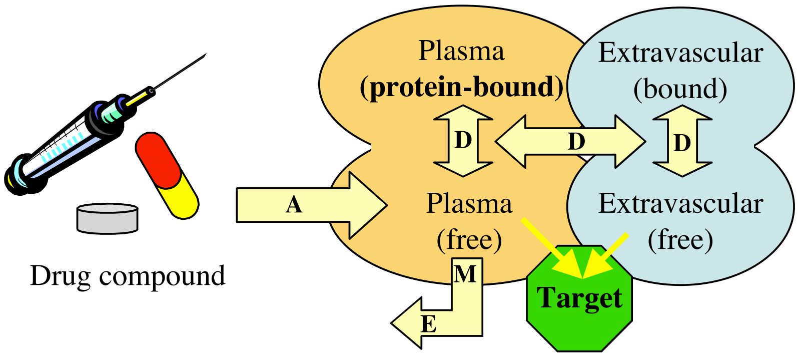


Figure 3-1: Pharmacokinetic processing of a drug with absorption (A), distribution (D), metabolism (M) and excretion (E). Only free fractions of the drug can bind to the target.

Since a drug compound has to reach a specific target in a certain organ or tissue, distribution is crucial for its efficacy. Besides active and passive transportation and diffusion, binding to protein receptors is playing an important role in this process. One of the major contributors stems from *plasma protein binding*, i.e. the interaction with soluble protein in the blood. With a concentration of approximately 680 μM, i.e. 40 g/L plasma, HSA is by far the most abundant protein in this compartment. Alongside with HSA, α_1-acid glycoprotein (AGP), α-globulins, and β-lipoproteins are also showing significant drug-binding behavior. While HSA predominantly binds acidic aromatic compounds, basic drugs often interact with AGP.

Plasma protein binding affects nearly all stages of pharmacodynamics and pharmacokinetics. Since only unbound molecules are able to interact with their targets [1], plasma binding directly affects metabolism, rapid clearance and toxicity. Furthermore, many drugs competitively interact with the same binding site on HSA, leading to displacement and sudden changes of a drugs' plasma concentration. Therefore, plasma protein binding was regarded as a somewhat dangerous property for many years and was tried to keep on a moderate level. Newer developments in drug discovery changed the view of the role of HSA and plasma binding completely. Especially the contribution of HSA to an increase in elimination half time helps developing drugs, which act longer (retard effect) and have to be administered less often (leading to a better compliance). An attractive example of increased HSA binding is the insulin derivative *detemir* (Novo Nordisk). By attaching a fatty acid moiety a long-acting form of insulin could be developed [2].

Albumin was recognized as a principal blood component as early as 1839. The name 'albumin' was derived from the white color (lat: albus) of protein precipitates. Although the most outstanding property is its ability to bind an broad variety of endogenous and exogenous ligands, it performs many additional functions. For example, HSA contributes 80% to colloid osmotic blood pressure, is mainly responsible for the maintenance of blood *pH*, plays a major role in detoxification, and sequesters oxygen free radicals [3].

3.1.2 Structure and properties of HSA

HSA is synthesized by the liver as a single peptide chain and is exported to circulation, where it remains with a plasma half-life of 19 days. The non-glycosylated protein consists of 585 amino acids resulting in a molecular weight of 66,500 Da. HSA possesses an unusually high number of disulfide bridges; 34 of its 35 cysteine residues are involved in 17 disulfide bridges while a single cysteine (Cys-34) remains free. Due to its high amount of acidic amino acids, the net charge of albumin is clearly negative [3, 4].

Even though HSA structure and binding features were explored for several decades, it was not possible to obtain any crystal structure at a reasonable resolution for years. The long period of frustration was ended in 1992, when a group from the NASA science center presented a structure at 2.8 Å resolution [5], obtained under microgravity environment in a U.S. space shuttle [4]. It shows HSA as a heart-shaped protein, which is organized in three homologous domains (I-III; *Fig*. 3-2A). Each of these domains can be further divided into two subdomains (A, B) consisting of six and four α-helices, respectively. This leads to an unusually high content of helical structures of around 67% (*Fig*. 3-2B), which are stabilized by the large number of disulfide bridges forming nine double loops. Such disulfide pairings between helical motifs are very unusual and are believed to be responsible for some of the unique features of albumin. No β-sheets are present in the crystal structure [3, 5].

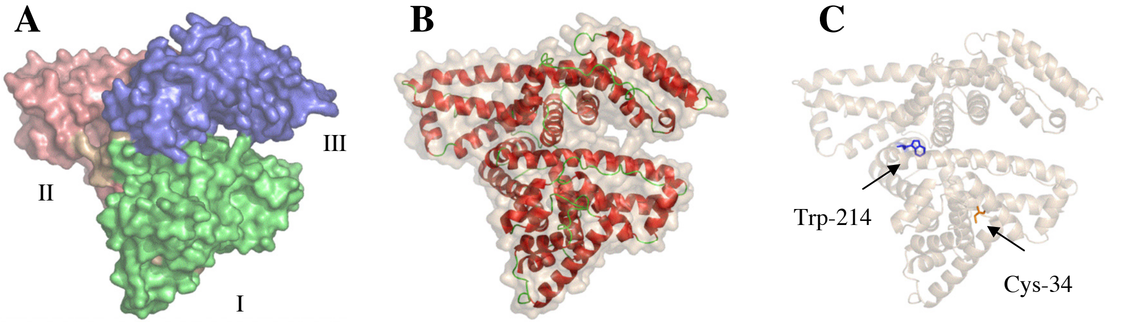


Figure 3-2: Crystal structure of HSA. **A:** Surface view with three domains (I in green, II in red, III in blue), **B:** Secondary structure (helices in red, turns in green), **C:** Unique amino acids (single tryptophan in blue, single free cysteine in orange)

Albumin shows an enormously high stability, and even exposure to very low or high *pH* values (*pH* 1.2-9), heat (up to 72°C; albumin preparations usually are pasteurized for 10 h at 60°C), or 8 M urea have no deleterious effects. The structure of HSA is nevertheless rather flexible and very sensitive to environmental factors such as *pH*, ionic strength or temperature. For example, five *pH*-dependant conformations are described in literature for human and bovine serum albumin (BSA) [4]. Besides the *N-form* (normal) at physiological conditions, albumin undergoes two changes each at lower and higher *pH* values. Around *pH* 4 the *F-form* (fast) is predominant, in which the two halves of the molecule separate leading to a lengthening. Further decrease of *pH* below 3.5 initiates the so-called *E-form* (expanded), where the domains are unfolded as much as the disulfide pattern permits. The structural changes at higher *pH* values are more subtle and gradual starting with the *B-form* (basic) between *pH* 7 and 9, which is discussed to have physiological importance. Finally, the *A-form* (aged) is usually detected above *pH* 9-10, involves both ionic forces and hydrogen bonding, and is not fully reversible (*Table* 3-1) [4].

Table 3-1: *pH*-induced isomeric forms and conformational transitions of albumin (adapted from Peters [4])

Isomeric Form	**E**	**F**	**N**	**B**	**A**
Name	expanded	fast	normal	basic	aged
Transition Name	E → F	F → N	N → B	B → A	
Transition *pH*	2.7	4.3	8	10	

Not only conformational isoforms but also other microheterogeneities are observed for albumin. The unpaired cysteine residue in position 34 has been detected in all avian and mammalian albumins with known structure and makes up for most of the mercaptan in plasma. This residue is buried in a crevice formed by a seven-residue turn in domain I under physiological conditions and is therefore partially protected from oxidation. As a result, the majority of plasma albumin exists in the free *mercaptalbumin* form (HMA) while about one-third carries a mixed disulfide (*non-mercaptalbumin;* HNA) with predominantly cysteine but also glutathione (in a ratio of four to one). A minor fraction of HSA is oxidized to a higher degree (sulfenic/sulfonic/sulfinic acid; HNA_{oxi}) [4, 6]. Dimeric and higher oligomeric forms of HSA are almost exclusively detected in commercial preparations and are believed to be generated by the pasteurization process during production. While some of these dimers can be easily reduced to the monomeric form, some others withstand this procedure indicating another aggregation mechanism than the formation of S-S bonds [4]. Finally, up to 10% of the circulating albumin is glucosylated at certain lysine residues. In patients with diabetes, this amount is significantly increased [3].

3.1.3 Ligand binding to HSA

HSA is one of the major transport proteins in human plasma. It binds a number of poorly soluble endogenous compounds such as non-esterified fatty acids, bilirubin, and bile acids and thus facilitates their transport in the plasma. In addition, a wide range of drugs and other exogenous substances also show a strong binding to albumin.

Table 3-2: Extended Sudlow binding site classification of Carter [3] with the site names and locations as well as some examples of known binders.

Site	**I**	**II**	**III + IV**	**V**	**VI**
Name	Warfarin-azapropazone	Indole-benzodiazepine	Long/medium fatty acid	Metal site 1	Metal site 2
Location	subdomain IIA	subdomain IIIA	uncertain[a]	Cys-34	*N*-terminus
Binders	warfarin salicylate bilirubin[b]	diazepam naproxen digitoxin[b]	oleate myristate palmitate	auranophan cisplatin Au(I), Hg(II)	Cu(II) Ni(II) Co(II)

[a] The exact number and location of fatty acid binding sites is still not known. [b] Separate binding sites for bilirubin and digitoxin are proposed but not confirmed.

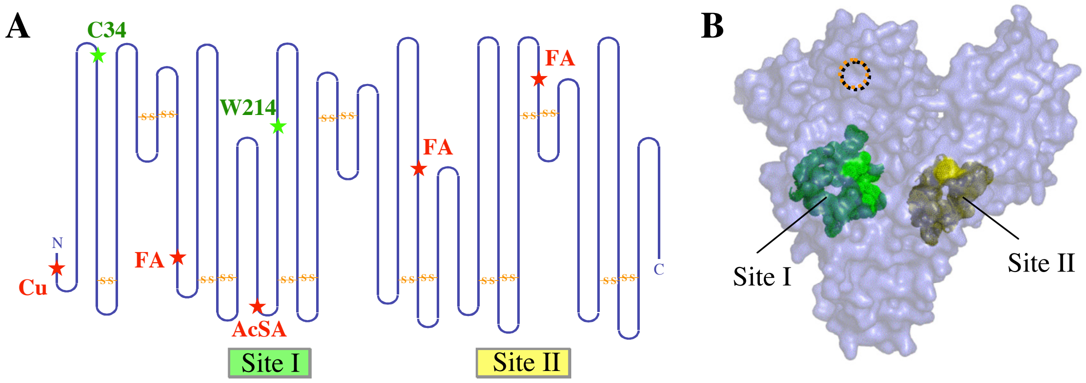


Figure 3-3: Location of Sudlow binding sites in HSA. **A:** Schematic overview adapted from Peters [4]. Several specific binding sites are symbolized with a red star (Cu: copper, FA: fatty acids, AcSA: acetylsalicylic acid). Positions of disulfide bridges, Cys-34 and Trp-214 are also indicated. **B:** Crystal structure (PDB code: 1BM0) of HSA with colored Sudlow binding sites and (hidden) location of Cys-34 (dotted orange circle).

Many investigations have been performed about the location and properties of the drug binding sites. Pioneering work was done by Sudlow and coworkers [7, 8], who identified two major binding sites located on subdomains IIA and IIIA, respectively (*Table* 3-2, *Fig*. 3-3). The anticoagulant drug warfarin was described as a marker compound for Sudlow site I (subdomain IIA), which is therefore often referred as *warfarin-azapropazone site*. Ligands for this binding site typically are bulky heterocyclic anions with the charge situated in fairly central position of the molecule (*Fig*. 3-4A). They are mostly non-aromatic except for some phenyl groups. The large specificity of site I is responsible for the cosmopolitan reputation of albumin among transport proteins [4]. The surface of the binding site is described as 'an elongated sock-shaped pocket wherein the foot region is primarily hydrophobic and the leg is primarily hydrophilic' [3]. Site I shows a high degree of adaptability and ligands can induce conformational changes [4]. The other major binding site (Sudlow site II, subdomain IIIA) is called *indole-benzodiazepine site* or *diazepam site*, according to its marker drug. Site II ligands often are aromatic and neutral or anionic with a more peripheral charge (*Fig*. 3-4B). Typical binders are indoles (e.g. L-tryptophan), benzodiazepines (e.g. diazepam) and non-steroidal anti-inflammatory drugs (e.g. naproxen). Binding to site II is more affected by HSA dimerization (non-thiol dimers) and glycation than binding to site I [4]. Additional specific binding sites for digitoxin, tamoxifen, and bilirubin are also proposed but their very existence is still subject to some debate [4, 5, 9, 10]. Even though albumin binding of some cationic drugs (e.g. imipramine, quinidine) is also described in literature [11], these molecules primarily bind to α_1-acid glycoprotein.

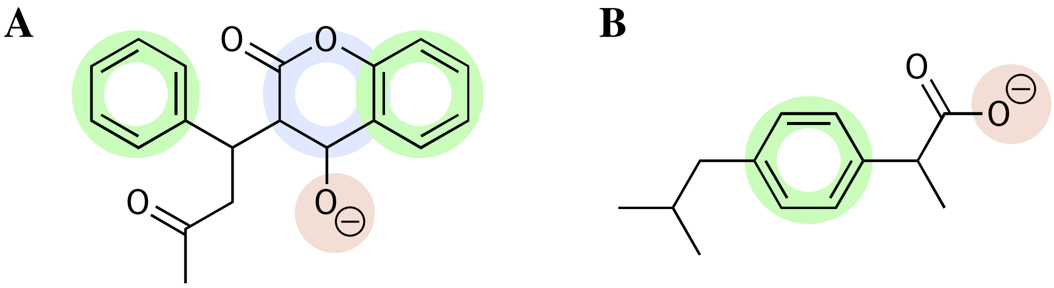


Figure 3-4: Pharmacophore preference of the two major drug binding sites. **A:** Bulky heterocyclic ligand for Sudlow site I with central negative charge (example: warfarin). **B:** Anionic aromatic ligand for Sudlow site II with peripheral charge (example ibuprofen). Aromatic rings are highlighted in green, heterocycles in blue and anionic groups in red.

Unfortunately, ligand-binding characterization is far more complicated in the case of HSA compared to most other targets. Traditionally, localization of binding sites for a specific drug has been done by competition experiments using known binders. However, one drug can bind to more than one site simultaneously and two drugs can bind at different location within the same binding site. Even more complicating is the fact that binding of a first drug at a specific site can influence the activity of a second drug at the same or even another site by allosteric effects [12]. Furthermore, it has been shown that endogenous ligands such as fatty acids [13] as well as environmental factors (*pH*, ions) can induce huge conformational changes thus influencing a drugs binding behavior.

Cys-34 also does also play a role in drug binding properties of HSA. Thiol-active endogenous compounds such as L-cysteine or glutathione, but also drugs like D-penicillamine, aurothiomalate, or cisplatin can form covalent bonds with the free cysteine residue. More recent studies showed that Cys-34 also interacts with nitric oxide forming *S*-nitrosoalbumin (~ 1% of the total albumin fraction) [4]. This supports the theory of albumin being a key carrier or reservoir of nitric oxide [3]. Whether the oxidation state of Cys-34 directly influences the affinities of reversibly binding molecules is still discussed controversially.

Some metal ions like copper(II) and nickel(II) bind strongly at a defined site at the *N*-terminus (Asp-Ala-His), other metals (Zn, Cd, Ca, Mn) also show a certain specific but far weaker affinity to HSA [3, 4]. Chloride is the most important anionic element binding specifically to HSA. Even though its K_D value is only 1.4 mM [14], the high abundance of chloride in the plasma (~ 100 mM) leads to significant binding and competition effects (especially with site II ligands) [4]. In addition to the endogenous ligands, HSA can also bind an impressive array of drug molecules from different classes with a rather broad specificity.

3.1.4 HSA analysis on Biacore

Acquiring data about pharmacokinetic properties has evolved into one of the most important and interesting applications for SPR instruments in drug discovery. Traditionally, most of the available binding data were generated from equilibrium dialysis experiments. However, these data give only information about the free or bound fraction of a drug. If the drug is dialyzed against whole plasma (including AGP, globulins and other proteins) these $\%bound_{PPB}$ values might be significantly larger than values derived from dialyses against purified albumin solutions ($\%bound_{HSA}$). This is especially true for cationic compounds, which predominantly bind to AGP. Other methods used for acquiring %bound or affinity data are fluorescence quenching [15], ITC [16], affinity chromatography (HPAC) [17], capillary electrophoresis [18], or differential CD [19]. While some of the published values are very consistent between those methods, others differ from each other by orders of magnitude. SPR-based analysis offers an interesting alternative, since only low amounts of HSA and unlabeled compounds are needed. In addition, assays can be automated easily and may provide information about kinetics as well as equilibrium affinity. Furthermore, differences of high affinity binders (low or sub-μM range) can not been described accurately by %bound values (> 97 $\%\mathrm{bound}_{\mathrm{HSA}}$) [20].

Three publications using Biacore for the characterization of HSA/drug interactions have been released so far, with focus on proof-of-concept [20] and on improvement of accuracy and robustness [21, 22], respectively. Ranking experiments to discriminate between low, medium, and strong binders, high-resolution screening for exact affinity values [20, 21], as well as competition experiments for characterizing primary binding sites [22] were covered by these publications. Since HSA is described to react sensitive on environmental changes (see *section* 3.1.2) assay parameters should be chosen with care. Different parameters have been investigated so far: HSA stability on the sensor chip [20, 21], effects of DMSO concentration and temperature [21], as well as immobilization parameters, species differences of albumins, and the influence of different batches and suppliers of HSA [22]. Some marker substances were investigated in all three studies (e.g. warfarin), others only in one or two of them (e.g. salicylic acid) and of some important markers there are no SPR data at all (e.g. diazepam, L-tryptophan). There are not only some significant differences in the assay conditions between the three reports, but the affinity values of some of the investigated drugs also show a high variability (*Table* 3-3).

Table 3-3: Comparison of assay conditions and affinity data of the three available HSA-drug binding studies using Biacore [20-22].

Year / Reference	**2000 [20]**	**2001 [21]**	**2003 [22]**
Temperature	25°C	25°C	25°C
Phosphate	67 mM	65.5 mM	10 mM
NaCl	70 mM	70 mM	150 mM
pH	7.4	7.4	7.4
DMSO conc.	5%	3%	3%
K_D Warfarin	9 μM	2.5 μM	3.7 μM
K_D Naproxen	26 μM	1.5 μM	10.6 μM
K_D Digitoxin	28 μM	38 μM	24.8 μM
K_D Salicylate	n.d.	48 μM	141 μM

3.1.5 Aims in this project

Fast determination of HSA-binding data on drug-like substances and lead compounds rely on a robust assay format. While the proof-of-concept for HSA assays on Biacore has already been performed, there are still some open questions and deviations between values from different publications. The first aim of this project was therefore to establish, validate, and optimize a HSA assay in terms of throughput and robustness. Possible environmental factors such as *pH* shifts, buffer ionic strength, and oxidation state of Cys-34 should be investigated more intensively. Finally, the possibility of using this single free cysteine residue to perform a site-specific immobilization should be explored.

3.2 Materials and Methods

This section only describes materials, equipment and procedures specifically used for the albumin project. Materials and general methods used in all Biacore assays are described in *section* 2.2.

3.2.1 Materials

Reagents and proteins

Human serum albumin used in all assays was from Sigma (Fluka AG, Buchs, Switzerland) and was essentially globulin and fatty acid-free (No. A-3782). Different batches of DMSO were used for the experiments. Dried DMSO from Riedel-de Haën (No. 34943) was found to be especially suitable. All other reagents were purchased from Sigma. Freshly bidistilled water delivered through quartz tubes was used for the preparation of all Biacore buffers and solutions.

Drug Compounds

Diazepam, nitrazepam, and phenprocoumon were a gift from Roche AG (Basel, Switzerland). Warfarin, digitoxin, indoprofen, cholic acid, L-tryptophan, L-kynurenine and quinidine were from Sigma (Fluka AG, Buchs, Switzerland). Salicylic acid, acetylsalicylic acid, quinine, and prednisone were from Siegfried AG (Zofingen, Switzerland). Sodium salicylate was purchased from Hänseler AG (Herisau, Switzerland). Naproxen (sodium salt) and salbutamol were a gift from Prof. Hans Leuenberger (Institute of Pharmaceutical Technology, University of Basel, Switzerland).

Equipment

HPLC separation were done on an Agilent 1100 purification system, equipped with a quaternary pump, a cooled well-plate autosampler, a column thermostat, a DAD detector, and a cooled analytical fraction collector (Agilent AG, Basel, Switzerland). Size exclusion chromatography was performed on a TSK Gel G2000SW column (7.5 x 600 mm; Tosoh Biosciences GmbH, Stuttgart, Germany). A Molecular Devices SpectraMax Plus UV absorbance plate reader was used for the Ellmans assays and a SpectraMax Gemini XP fluorescence plate reader for measuring intrinsic tryptophanyl emission (Bucher Biotec, Basel, Switzerland). A Bio-Rad SmartSpec 3000 spectrophotometer equipped with a quartz cuvette was used for absorption measurement in the UV range.

3.2.2 Preparation of running buffer

10 mM PBS with 3% DMSO was used as running buffer during the assays if not differently described. A 1.03-fold PBS (10.3 mM sodium phosphate, 155 mM NaCl) was prepared and adjusted to *pH* 7.30. After filtration and degassing (see *section* 2.2.2), 3% DMSO were added (using a glass pipette) and the final *pH* was verified to be 7.40 after DMSO addition. The prepared running buffer was used no longer than 24 hours to avoid shifts in DMSO concentration.

3.2.3 Amine coupling of HSA

The protocol for the immobilization of albumin was the result of a combination and an adaptation of different sources [20-22]. HSA was dissolved in water to a stock concentration of 5 mg/ml and further diluted to 50 μg/ml in 10 mM sodium acetate buffer *pH* 5.0 just before immobilization. Standard amine coupling (see *section* 2.2.5) was used to reach surface densities of 8-18 kRU. Typical activation/deactivation as well as immobilization times ranged from 5-7 min each. All surfaces were washed with three consecutive pulse injections (20 s) of 50 mM NaOH to remove noncovalently bound protein. Unmodified carboxymethyl dextran flow cells were used as reference surfaces.

3.2.4 Stabilization of HSA surface

Even though HSA is immobilized covalently, surface decay of up to 80% within the first few hours has been observed in previous studies [21, 22]. The reason for this behavior is still not known, and various attempts for stabilization have failed so far. The most effective procedure was found to be repetitive buffer or sample injections over several hours [22]. Therefore, 20 injections of each running buffer, 20 μM warfarin and 6 μM naproxen were performed over 10 hours. DMSO calibration (see *section* 3.2.5) was done before and after the stabilization process. Baseline signals as well as difference signals for warfarin and naproxen were monitored and plotted against time.

3.2.5 DMSO calibration

Since even small shifts in DMSO concentration cause huge changes in bulk signal intensity, any difference between the volumes of the sample and reference surface are clearly visible. Therefore, these DMSO signals have to be subtracted from the sample binding response [20]. Calibration solutions ranging from approximately 2.9% to 3.1% DMSO were prepared by adding 1 μl DMSO (A) or 50 μl PBS 1.03x (B) to 1 ml of running buffer. Solutions A and B were mixed to get five calibration solutions (100%, 75%, 50%, 25%, 0% A). The calibration procedure was performed before each binding experiment. Difference signals between sample and reference cell were plotted against the signal on the reference cell and fitted using linear regression. Sample responses were then corrected according to [20].

3.2.6 General optimization of the HSA assay

The aim of the general optimization experiments was to reach improvements in throughput and handling while maintaining accuracy and reproducibility. Special care was applied to the handling of DMSO-containing solutions. Therefore, different batches of DMSO have been tested and a panel of sample containers (glass or plastic vials, 96 well plates) was evaluated. Additional parameters such as vial capping, autosampler temperature, the number of sample dilutions, and carry-over effects were investigated. Available software for data evaluation (*BIAevaluation*, *Scrubber*, *Excel*) was compared concerning ease of use, automation, reproducibility and time consumption. Finally, an application for the automated sample randomization and method generation was developed in *Excel* based on visual basic scripts (see *appendix* B1).

For all further assays, airtight rubber caps and polypropylene vials were used. The autosampler area was maintained at a constant temperature of 20±2°C, while keeping the IFC and detector at 25°C. *Scrubber* was used for data processing and evaluation. All sensorgrams were double referenced and the whole curves were corrected for DMSO effects. Equilibrium binding data were fitted either to a single-site or to a two independent-sites model.

3.2.7 HSA affinity ranking

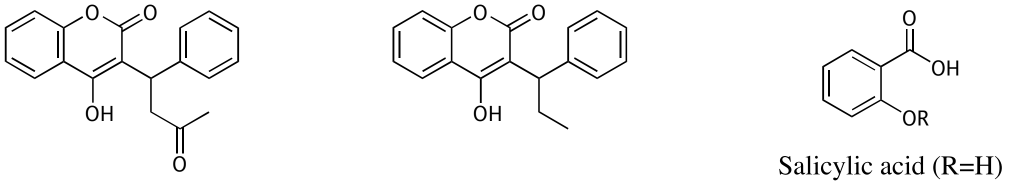
The binding strength of different known binders have been tested with a set chosen by criteria such as commercial availability, structural diversity, covering of binding sites, functional groups, and available results. A list of the drugs used for different experiments alongside with some important properties can be found in *table* 3-4, their structures are shown in *Fig*. 3-5.

Table 3-4: Drug compounds used for HSA binding experiments with some important physico-chemical properties and known plasma protein binding values.

Compound	Abbr	MW	log P [a]	pK_a [a]	$\%bound_{PPB}$ [b]	Site	Charge
Acetylsalicylic acid	AcSA	180.2	1.19	3.49	37	I	anionic
Cholic acid	ChoA	408.6	2.02	4.98	48	I[c]	anionic
Corticosterone	Cort	346.5	1.94	-	68	n.d.[c]	neutral
Diazepam	Dzpm	284.7	2.82	3.40	93.2	II	neutral
Digitoxin	Dgtx	765.0	1.85	-	97	II[c]	neutral
Indoprofen	Indo	281.3	2.77	4.38	98.5	II	anionic
Naproxen	Napr	230.3	3.18	4.15	99.5	II	anionic
Nitrazepam	Nitra	281.3	2.25	-	87	II	neutral
Phenprocoumon	Ppro	280.3	3.62	-	99.5	I	neutral
Prednisone	Pred	358.4	1.46	-	38	n.d.[c]	neutral
Quinidine	Qdne	324.4	3.44	8.56	88	II	cationic
Quinine	Quin	324.4	3.44	8.56	93	II	cationic
Salbutamol	Salb	239.3	0.64	-	8	n.d.	neutral
Salicylic acid	ScyA	138.1	2.26	2.97	73	I	anionic
L-Tryptophan	LTrp	204.2	-1.05	7.38	80-90	II	zwitter
Warfarin	Warf	308.3	2.60	5.08	99.5	I	neutral

[a] log P and pK_a values were retrieved from the online PhysPro database [23] and are determined experimentally (octanol/water distribution), except for the log P of salbutamol, which is only estimated and pK_a of indoprofen (calculated). [b] %bound values were collected from different sources. An extended list with all references can be found in *appendix* B2. [c] Special binding sites for bile salts (bilirubin, cholic acid) and steroids (especially digitoxin) are discussed, but not confirmed.

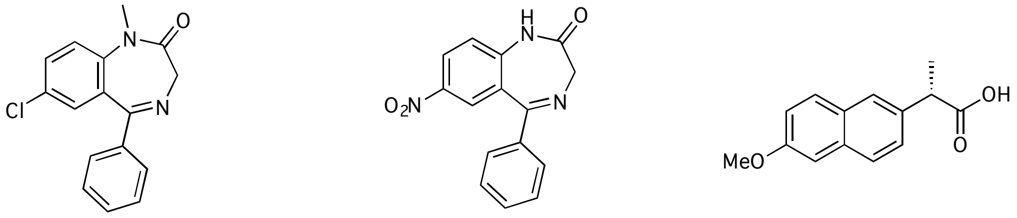
For affinity ranking, 30 mM stock solutions of each compound in DMSO were diluted in PBS 1.03x to a concentration of 1 mM in 3% DMSO. These solutions were further diluted in running buffer to a final concentration of 333 μM and injected as randomized triplicates as described in *section* 3.2.6. Steady state binding signals was averaged between 10 and 20 s after injection start, divided by the compounds molecular weight, and evaluated in *Excel*. Average values from triplicate injections of all drugs were ranked according to their normalized SPR signal and plotted against the corresponding log P value or the bound fraction ($\%bound_{PPB}$). Finally, the shape of each binding curve was examined for deviations from expected behavior (rapid kinetics, block signal) and classified into different groups.



Warfarin

Phenprocoumon

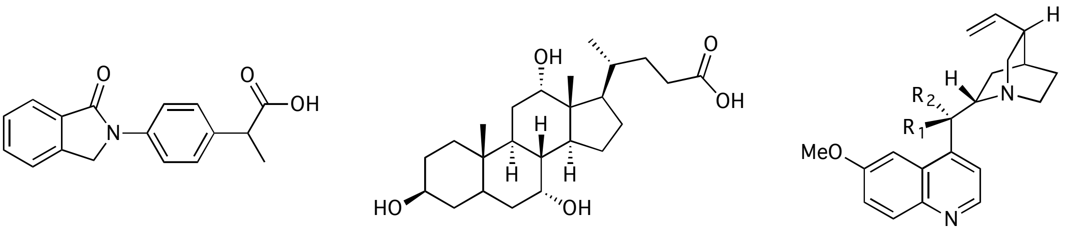
Salicylic acid (R=H)
Acetylsalicylic acid ($COCH_3$)



Diazepam

Nitrazepam

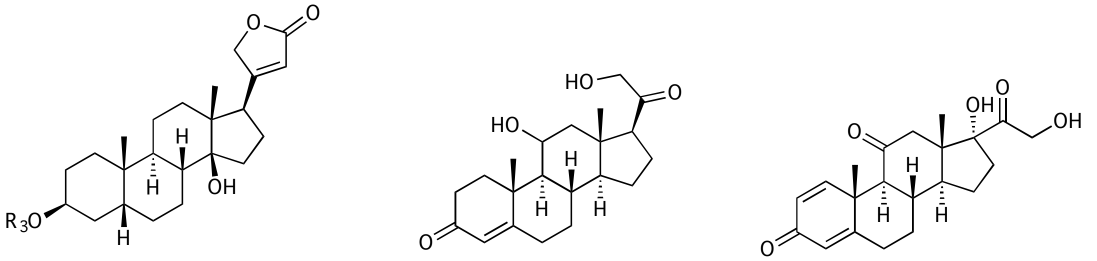
Naproxen



Indoprofen

Cholic acid

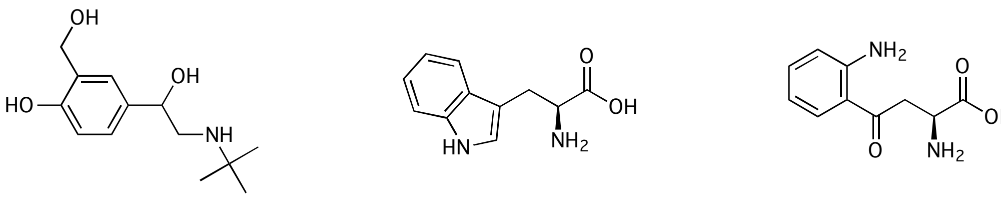
Quinine (R_1=OH, R_2=H)
Quinidine (R_1=H, R_2=OH)



Digitoxin
(R = Digitoxose)

Corticosterone

Prednisone



Salbutamol

L-Tryptophan

L-Kynurenine

Figure 3-5. Structures of known HSA-binding drugs selected for HSA experiments. Top row: marker drugs for site I. Second row: Marker compounds for site II. Third row: Examples for anionic and cationic substances. Fourth row: Steroids as marker for a possible digitoxin binding site. Bottom row: Weakly binding and zwitter-ionic compounds.

3.2.8 High-resolution binding experiments

High-resolution binding experiments were used to determine a drugs K_D value. These binding studies were performed with a selection of compounds in different experiments, i.e. assay dependence from buffer strength (see *section* 3.2.9), oxidation state (see *section* 3.2.10), ligand-induced conformational changes (see *section* 3.2.11), overlay effects (see *section* 3.2.12) or thiol coupling (see *section* 3.2.13).

30 mM stock solutions of each compound were prepared in dry DMSO. 30 μl of these stock solutions were diluted in 970 μl PBS 1.03x to reach concentrations of 1 mM drug and 3% DMSO. DMSO stock solutions were stable at 4 °C for several days, except for L-tryptophan, which was dissolved directly in running buffer and had to be prepared freshly every day [10]. Concentration series were prepared by threefold serial dilution in running buffer covering a range from 1000-0.15 μM. The solutions were injected as triplicates in a randomized manner. Injections were done using the *kinject* command with an association time of 30 s and 20 s of undisturbed dissociation. Five buffer blanks were injected before and one between each series. No regeneration was needed but additional wash steps for the needle and the IFC were introduced to avoid carry-over. In the case of L-tryptophan wash steps alone were not sufficient for avoiding carry-over. Therefore, a short buffer pulse of 12 s after the IFC wash was added. In addition, the dissociation phase was extended to 60 s and tryptophan binding was not averaged during steady state but immediately after injection end.

3.2.9 Influence of ionic strength of sample/running buffer

Since some of the compounds investigated in screening contain free carboxyl or amine groups and HSA is known to react on shifts in *pH* (see *section* 3.1.2), the influence of ionic strength / buffer capacity in a HSA assay was investigated. For this purpose, buffer systems with variable anion and cation concentrations were prepared (10/150, 67/93, 67/150 mM phosphate/NaCl) at *pH* 7.40 and 3% DMSO. Three non-ionic (diazepam, warfarin, digitoxin) and three anionic (naproxen, indoprofen, cholic acid) drugs as well as one cationic drug (quinine) were screened in all three buffer systems. Absolute and relative (to 10/150 mM PBS) K_D values were determined.

3.2.10 Influence of HSA redox state on drug binding

Oxidation state of the single free cysteine residue of HSA has shown to be highly variable in commercial preparations [24]. Therefore, the influence of the HSA redox state on binding affinities has been investigated. For preparing oxidized and reduced HSA samples, 4 mg/ml HSA (60 μM) in sample buffer (10.1 mM Na_2HPO_4, 1.76 mM KH_2PO_4, 137 mM NaCl, 1 mM EDTA, *pH* 7.4) were mixed with 10fold molar ratios of reduced (GSH) or oxidized glutathione (GSSG) or with an equimolar amount or a 5fold molar excess of DTT. Incubation at 25°C was 1 hour for GSH and 3 hours for GSSG and DTT [6, 25]. After the incubation time, excess of thiol-active reagents was removed by size exclusion chromatography (TSK Gel G2000SW, isocratic in sample buffer, 1 ml/min, ambient temperature, 30 min). Dimer and monomer peaks were collected separately and latter was used for Ellmans assay.

The degree of HSA oxidation was determined by measuring the concentration of free thiol groups using an enhanced Ellmans assay protocol [26]. HSA preparations were diluted twofold in sample buffer to get an appropriate signal intensity. Ellman/cystamine reagent was prepared by dissolving 77.1 mg DTNB and 56.3 mg cystamine dihydrochloride (10 mM both) in 25 ml buffer 7.0 (100 mM NaH_2PO_4, 0.2 mM EDTA, *pH* 7.0) and readjusting the *pH* to 7.0. To 1 ml of HSA sample 200 μl of strong buffer 8.2 (100 mM boric acid, 0.2 mM EDTA, *pH* 8.2) and 20 μl of Ellman/cystamine reagent were added and mixed immediately. After 5 min the adsorption at 412 nm (A_{412s}) was read against a water blank. Protein (A_{412p}) and reagent blanks (A_{412r}) were measured in parallel. Concentration of free thiols was calculated according to *equation* 7 with ε_{412} = 14,150 $M^{-1}cm^{-1}$ and d = 1 cm.

$$[SH] = (A_{412s} - A_{412p} - A_{412r}) / (\Delta\varepsilon_{412} \text{ x } d) \qquad \text{[Eq. 7]}$$

Concentration of HSA in the same sample was determined by measuring absorption of the protein blank at 279 nm (A_{279p}) in a quartz cuvette and calculated according to *equation* 8 with ε_{412} = 0.531 $gL^{-1}cm^{-1}$ [27].

$$[HSA] = A_{279p} / (\Delta\varepsilon_{279} \text{ x } 1 \text{ cm}) / 66{,}500 \text{ Da} \qquad \text{[Eq. 8]}$$

Finally, the sulfhydryl concentration was divided by the HSA concentration to get the molar ratio of free thiol groups.

HSA treated with DTT (1:1) and L-cysteine (1:10) as well as an untreated but separated HSA sample were immobilized on different flow cells of a sensor chip using amine coupling (see *sections* 2.2.5 and 3.2.3). Warfarin, diazepam, digitoxin, naproxen, indoprofen, and cholic acid were tested on the surface in a high-resolution assay (see *section* 3.2.8).

3.2.11 Monitoring of *pH*-induced conformational changes

Conformational transitions induced by *pH* changes have been monitored using a variety of methods [4]. We therefore investigated whether *pH* shifts could also influence Biacore signals of immobilized HSA. A sensor surface containing 18 kRU of human serum albumin was screened for conformational changes. For this purpose, 10 mM buffer systems covering a *pH* range from 1.5 to 9.5 were prepared in steps of 0.5 units (Glycine *pH* 1.5-3.0; Citrate *pH* 2.0-7.0; Phosphate *pH* 6.0-8.5; Borate *pH* 8.0-9.5). A 10 mM phosphate buffer at *pH* 7.5 was used as running buffer without any additives (NaCl, EDTA, etc.). Each buffer was injected for 5 min and monitored for another 5 min after injection end. Sensorgrams were referenced against an untreated surface and steady state signals were plotted against *pH*. In order to reduce buffer salt-induced differences in signal intensity, values for citrate and borate injections were normalized to the signals of phosphate buffer in the overlapping areas. Glycine injections were normalized to the processed citrate curve in the same way.

To validate the observations made by Biacore analysis, the *pH*-induced transition of HSA was monitored by fluorescence spectroscopy using the method of Ruker *et al.* [28]. Samples containing 4.5 μM HSA were prepared by mixing 5 μl of a 450 μM stock solution of HSA in water and 495 μl of 10 mM citrate or phosphate buffer at different *pH* values. 350 μl of each sample were transferred to a black-walled 96 well plates. All samples were measured at ambient temperature using an excitation wavelength of 295 nm and scanning tryptophanyl emission from 320 - 350 nm in 1 nm steps. The following instrument parameters were chosen for the analysis: no cut-off, medium sensitivity, and highest precision (30 reads/well). All signals were corrected for buffer baseline fluorescence and all measurements were performed six times. Averaged λ_{max} values were plotted against *pH*.

3.2.11 Monitoring of ligand-induced conformational changes

Binding modes of diazepam and L-tryptophan were further analyzed using high-resolution screening (see *section* 3.2.8). As a metabolite of L-tryptophan, L-kynurenine was selected to confirm the tryptophan results and was analyzed under the same conditions (direct dissolvation in running buffer, fresh preparation). All sensorgrams for tryptophan and kynurenine were processed normally but mirrored by multiplication of each data point with -1. In addition, average binding was determined not only during steady state (time range 20-25 s) but also just after injection end (time range 31-35 s). Binding plots obtained by this adapted method could be fitted normally to a single-site model.

3.2.12 Overlay of signal effects

1 mM of salicylic acid, sodium salicylate, and L-tryptophan were dissolved directly in two running buffers of varying capacity (10/150, 67/93 mM phosphate/NaCl *pH* 7.4, 3% DMSO). The *pH* value of each solution as well as of the running buffers alone was determined in order to identify any shifts in *pH*. Salicylic acid was screened under the same conditions in both buffer systems (high-resolution screening).

An overlay of ligand- and *pH*-induced signal changes was prepared by dissolving 1 mM L-tryptophan in 10/150 mM phosphate/NaCl buffer 3% DMSO at *pH* 7.40 and 7.14 (equals a *pH* shift induced by 1 mM salicylic acid). Both solutions as well as a buffer blank at *pH* 7.14 were diluted (threefold, serial) in running buffer *pH* 7.40 and analyzed by high-resolution screening.

3.2.13 Thiol coupling

While thiol immobilization of proteins featuring natural or introduced thiol groups are already described in literature [29], no application for albumin has been reported so far. A novel *in situ* coupling method targeting the free Cys-34 in HSA was therefore developed. A combination of a mild reducing agent (reduced glutathione) and a thiol-reactive activating agent (DTNB) was used to generate free and reactive thiols in the protein (*Fig* 3-6). HSA was dissolved in 10 mM phosphate buffer *pH* 7.40 at a concentration of 2 mg/ml (30 μM). 60 μl of the HSA solution were mixed with 5 μl of a 10 mM GSH and 5 μl of a 10 mM DTNB solution and incubated for 15 min at ambient temperature. After incubation, 200 μl of a 10 mM formate buffer *pH* 3.5 were added and the mixture was incubated for another 90 min. A thiol-functionalized surface on a CM5 sensor chip was prepared as described in *section* 2.2.7. The thiol surface was reduced again just before immobilization with a 2 min pulse of 100 mM DTT in 100 mM borate buffer *pH* 8.5. The protein was immobilized by injection of the HSA/GSH/DTNB mixture for 5 min at a flow rate of 10 μl/min and resulted in an immobilization density of 8-16 kRU. Non-specifically bound protein was removed at a flow rate of 50 μl/min with three consecutive pulses (12 s) of 50 mM NaOH. Thiol-coupled HSA could be removed completely by injecting a 3 min pulse of each 100 mM DTT and 50 mM NaOH, and immobilized again on the same flow cell (see evaluation below).

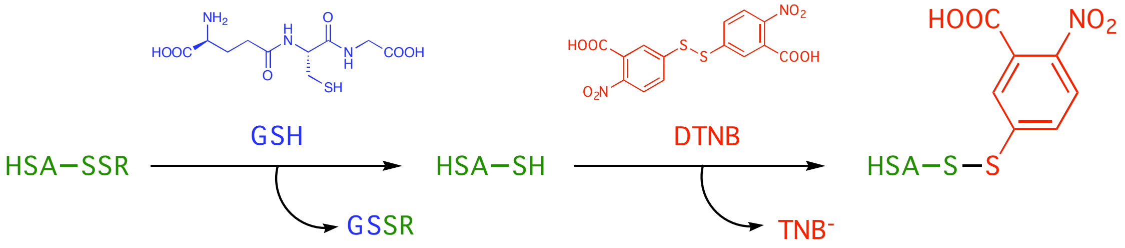


Figure 3-6: Mild reduction and activation of HSA. Mixed disulfides at Cys-34 are first cleaved by GSH and then activated with DTNB leaving a (yellow colored) TNB anion. R represents a mixed disulfide moiety such as L-cysteine and glutathione, or another HSA molecule (in the case of HSA dimers).

For showing the relevance of the two additives (GSH, DTNB) in the interaction mixture, one or both of these reagents were replaced by phosphate buffer (blank) and each mixture was immobilized under the same conditions. Since fatty acids are known to induce conformational changes of albumin resulting in a better accessibility of Cys-34 [30], a possible effect of fatty acids on surface density or coupling efficacy was investigated. Oleic acid was used for this purpose and was dissolved in pure ethanol to a concentration of 240 μM. 20 μl of the solution were transferred to an empty vial and the ethanol was evaporated for several hours at room temperature. HSA mixtures with and without additives were pipetted into the oleic acid-containing vial and immobilized as described above.

Thiol-coupled surfaces were evaluated and compared to amine immobilization. Stability was tested using the same procedure as described for the amine coupling (see *section* 3.2.4) and compared to the stability of the amine surface. For this purpose, absolute baseline signals and warfarin signals at 20 μM were plotted against time. Reuseability of thiol-functionalized sensor chips and reproducibility of the coupling procedure was tested by repeated injections of DTNB-activated HSA (3 min) followed by 2 min pulses of each 100 mM DTT in 100 mM borate buffer *pH* 8.5 (regeneration) and 50 mM NaOH (wash) after 3 min. Finally, high-resolution screening was performed for a selection of HSA binders (warfarin, diazepam, digitoxin, naproxen, indoprofen, L-tryptophan, quinine) and K_D values were compared with those obtained from amine-coupled surfaces.

Lysine distribution on the surface of the HSA crystal structure (*Fig*. 3-21) was visualized with PyMol using the PDB file *1BM0*. A semitransparent (80%) Connolly surface [31] was generated. Residues defining binding site I (K199, L219, R222, F223, L234, L238, R257, L260, A261, I290, A291, E292) and binding site II (P384, L387, I388, F395, L407, R410, L430, V433, A449, L453, E450, R485) [5] were selected, highlighted (dots), and colored green and yellow respectively. Finally, amine nitrogens were colored red. Effects of oleic acid on the accessibility of Cys-34 was also studied in PyMol. Crystal structures with (PDB: *1GNI*) and without oleic acid (PDB: *1BM0*) were overlaid and domain II of both structures were fitted. A Connolly surface was generated for both molecules and the sulfur atom of Cys-34 was highlighted by an orange color.

3.3 Results & Discussion

Human serum albumin assays on Biacore systems are already described in literature and are an important tool when developing drug-like substances. However, there are some differences and limitations in the protocols so far available. Various aspects and improvements were investigated in the present thesis:

- Assay conditions were further optimized in respect of throughput and data quality (see *section* 3.3.1).
- The assay quality was evaluated by performing different high-resolution and ranking studies (see *sections* 3.3.1 and 3.3.2).
- While some parameters (DMSO concentration, temperature, immobilization conditions) have already been investigated, no attention has been paid on the redox state of HSA and on the ion strength of the running buffer. These effects were investigated in *section* 3.3.3 (buffer strength) and *section* 3.3.4 (redox state).
- Since conformational changes can have a direct influence on signal intensity [32], experiments on *pH*- and ligand-induced conformational changes were performed (see *sections* 3.3.5 to 3.3.7).
- Finally, a new coupling method was developed for the reversible immobilization of HSA by its single free cysteine residue (see *section* 3.3.8).

3.3.1 General optimization of HSA assay conditions

To fulfill the requirements of a drug discovery environment, an assay has to be accurate, reliable and fast. For this purpose, a HSA assay protocol was established on the base of available protocols [20-22, 33-35] and optimized for fast and reliable screening of drug-like substances. Many different parameters, from the number of concentrations or the type of vials to buffer preparation data evaluation were tested. By using polypropylene vials, which were tightly capped (rubber caps) and by cooling autosampler racks to 15-20 °C reproducibility of triplicate injections was very high even when injecting from the same vial three times. Vial capping was found to be critical in earlier studies due to buffer dehydration after the hard-plastic caps were penetrated the first time [36]. Nine concentrations prepared by threefold dilution were found to be sufficient to characterize a drug in high resolution. This approach reduced preparation time and the number of vials needed. In addition, up to eight high-resolution assays could be prepared on the two racks in parallel for screening over night. Typical preparation times were around 30 minutes and high-resolutions screening of one compound was slightly above three hours. Carry-over was only found to be a problem with certain compounds. Injection of a short buffer pulse as described in [22] did not always completely eliminate this behavior. Therefore, additional wash steps for the needle and the IFC had to be included. For L-tryptophan, which showed higher carry-over effect, an additional blank buffer injection after each injection was mandatory.

DMSO quality – and particularly its water content – was found to be crucial for high quality experiments. Some compounds with high log P values (i.e. digitoxin, medazepam, progesterone) readily precipitated during dilution into running buffer when using DMSO batches, which were not kept under dry conditions. Different types of sample containers were tested for chemical resistance against DMSO containing buffers. Standard polypropylene vials (7 mm) provided by Biacore AB had a good resistance profile and were available at low price. Glass vials showed no benefit over polypropylene vials. None of the tested 96 well plates were of practical use for DMSO assays, since they all showed large signal intensities when buffer blank were injected repeatedly. Not even well plates declared as 'DMSO resistant' passed the requirements. Glass well plates are commercially available but very expensive. Therefore, drug stock solutions at 100% DMSO were prepared in glass vials and all samples at final dilution in sample buffer were measured in polypropylene vials. For preparation of running buffer it had to be considered that DMSO will raise the *pH*. As a consequence, DMSO-free buffers were prepared at slightly lower *pH* values (*pH* 7.30) to reach *pH* 7.4 after addition of DMSO. In any case, final *pH* should be checked after adding DMSO.

For data evaluation, the instrument-bundled software *BIAevaluation* was compared to the newly available software tool *Scrubber*. *BIAevaluation* is a flexible tool that allows the combination of different result files into a single project file. It also offers several curve manipulation tools and many fitting algorithms for kinetic and steady state analysis. Unfortunately, there are no tools for automated DMSO correction or double referencing of whole sensorgrams, curve alignment has to be done manually. In addition, no model for two binding sites is included and fitting is restricted to 24 curves. While a two independent- sites model could easily be added to the software, curves had to be exported to *Excel* for DMSO correction and double referencing and re-imported for further analysis. The restriction to 24 binding curves is not convenient for many high-quality analyses, since they normally consist of at least triplicate injections spanning a concentration range of three or more orders of magnitude (8-16 concentrations). As a result, data analysis of a complete data set using *BIAevaluation* usually takes several hours. *Scrubber*, on the other hand, is limited on single result files but offers a much higher degree of automation. Tools for curve alignment and cropping, as well as for DMSO correction and double referencing are applied in series, and the processed sensorgrams can be fitted to a single-site or two independent-sites binding model (without the limitation to 24 curves). This reduced analysis time to less than an hour. Since all processing steps can be saved in a method file and applied to another result file, this procedure can be further reduced to only a few minutes. A full kinetic analyis is not possible in *Scrubber*, but k_{off} values (and dissociation half-times) are calculated and curves can be exported to *BIAevaluation* or *CLAMP* [37].

All these measures taken together resulted in highly reproducible injection triplets, which could be fitted using a single-site or two independent-sites model. Sensorgrams and data of a typical assay are shown in *figure* 3-7 for warfarin. Association and dissociation phases were very fast and equilibrium data fitted very well to a two independent site model, which is in full agreement with previous Biacore studies [21, 22]. Assay preparation usually took less than 30 minutes and a complete data set for one compound was acquired and evaluated within four to five hours.

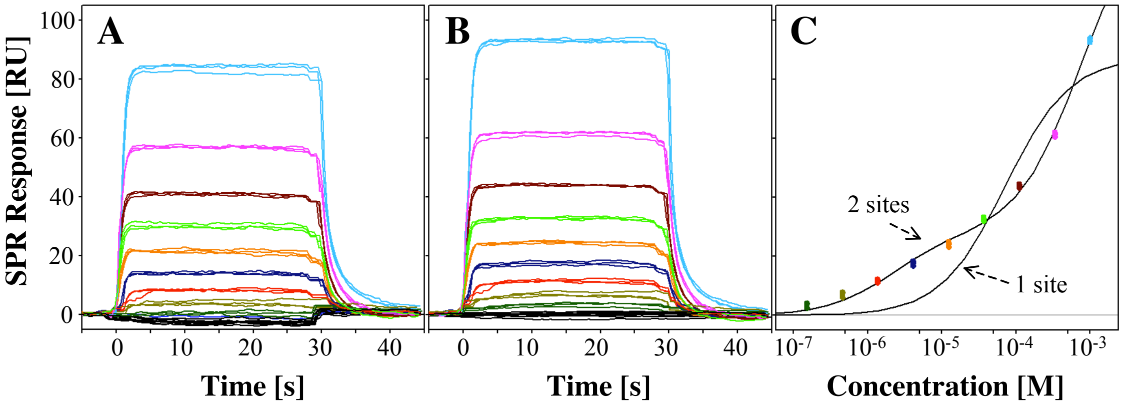


Figure 3-7: Optimized binding assay for drug-HSA interaction using warfarin as an example (threefold dilution series, 0.15-1000 μM). Sensorgrams after normal referencing (**A**) and after DMSO correction and double referencing (**B**) are shown. **C:** Steady state affinity plot of warfarin with fits generated by a single-site or two independent-sites binding model.

3.3.2 HSA-drug binding studies

In principle, there are two major types of binding experiments, i.e. ranking and high-resolution studies, that can be performed to get information about a drugs affinity for HSA. In *ranking studies*, SPR signals at a single concentration are divided by the compounds molecular weight. This normalized signal intensities give only qualitative information about binding strength. To evaluate the optimized HSA assay, 16 different compounds (*Table* 3-4, *Fig*. 3-5) were injected at a fixed concentration of 333 μM and ranked by their normalized responses (*Fig*. 3-8).

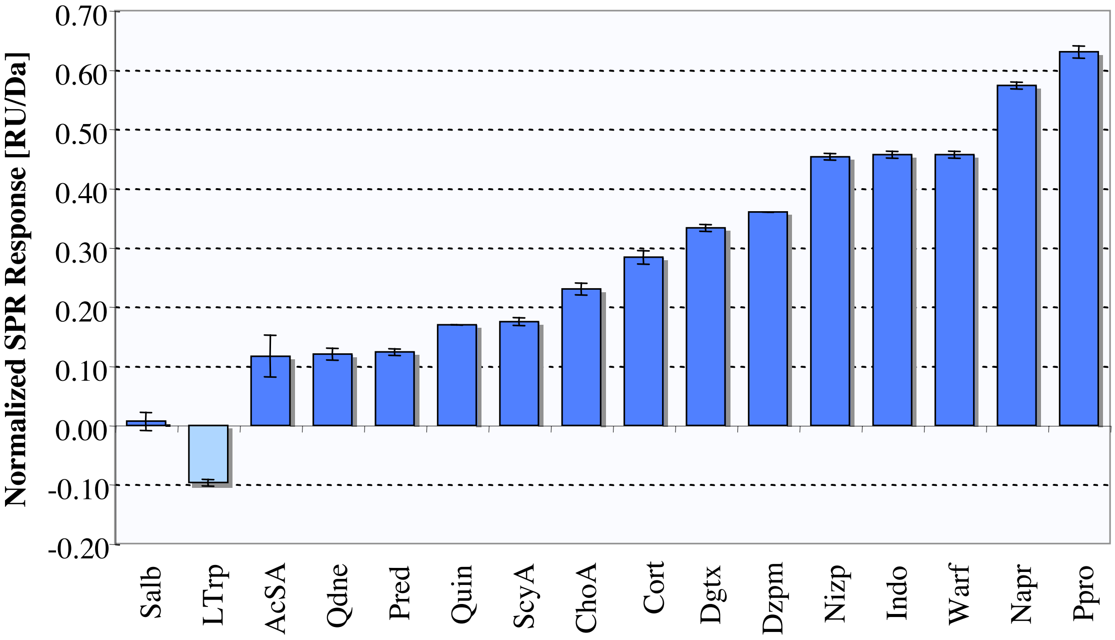


Figure 3-8: Ranking of 16 known HSA-binding drugs according to their normalized response at a concentration of 333 μM (SPR signal divided by molecular weight). Abbreviations are listed in *table* 3-4.

HSA is reported to predominantly bind hydrophobic (and anionic) drugs with high affinity [4]. To correlate the ranked compounds with their physico-chemical properties and biological relevance, the normalized responses were plotted against log P values (*Fig*. 3-9A) and the bound fraction on plasma proteins (*Fig*. 3-9B).

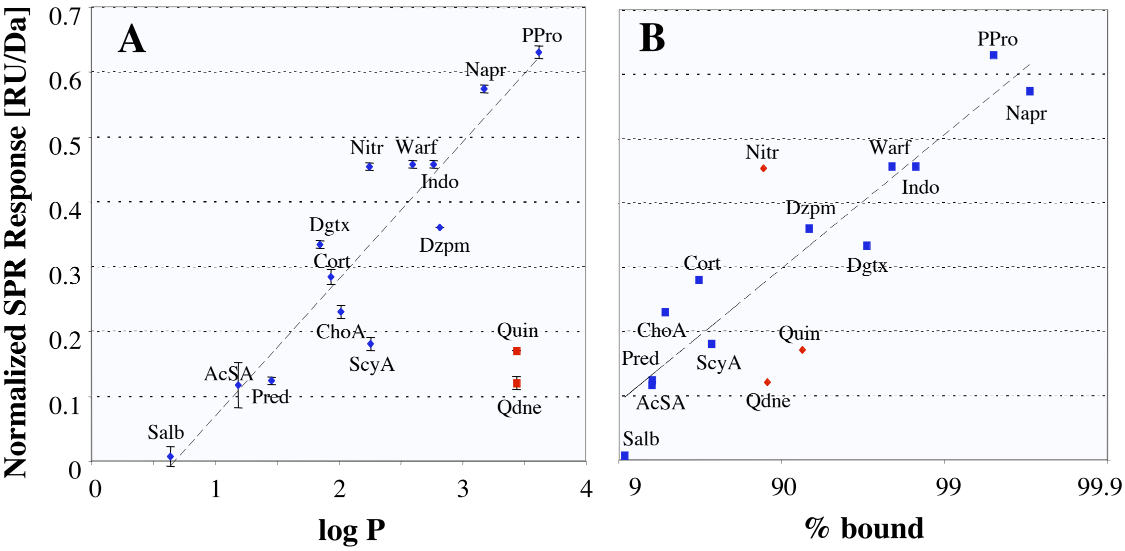


Figure 3-9: Correlation of normalized SPR responses from ranking experiments and literature log P (**A**) and %bound values (**B**). To increase the resolution of high-affinity compounds %bound values were put on a logarithmic scale. Abbreviations, references, log P and %bound values are listed in *table* 3-4.

Data from ranking experiments are in good agreement with literature [20]. Even small differences between similar structures are clearly visible. For example, quinine is ranked higher than quinidine, which is in agreement with plasma protein binding studies of the two drugs [38]. Interestingly, these substances showed the most prominent deviation from the correlation in both correlation plots. The differences in the %bound-plot can be explained by the fact that the literature values refer to binding to total plasma protein. As cationic drugs, quinine and quinidine bind strongly to α_1-acid glycoprotein (AGP). HSA only contributes to around 35% of the total 70-90% protein binding in the case of quinine [39]. Removing these drugs from the plot raises the correlation coefficient significantly from 0.76 to 0.85, when additionally removing nitrazepam it even increases to 0.91. The relatively high deviation in the log P-plot might also be influenced by the cationic nature of these drugs, since binding of positively charges compounds is unusual for HSA and therefore weaker than expected for the whole plasma protein fraction. The higher standard deviation in the ranking signal of acetylsalicylic acid might be explained by ionic effects or by transacetylation upon binding, i.e. transfer of the acetyl group to Lys-199 of HSA [4].

All molecules (except L-tryptophan) showed very fast binding kinetics with a rapid association phase and typical dissociation half-times at or below one second. Analyzing the ranking sensorgrams in more detail led to a classification into four clearly distinguishable classes:

A) Drugs showing a proper block signal (acetylsalicylic acid, cholic acid. corticosterone, digitoxin, prednisone, quinidine, quinine)

B) Compounds with a multiphase association (diazepam, nitrazepam, progesterone)

C) Compounds leading to a decreased baseline (naproxen) after the injection

D) Substances showing negative binding signals (L-tryptophan).

Examples of ranking curves are shown in *figure* 3-10. A complete overview can be found in *appendix* B3.

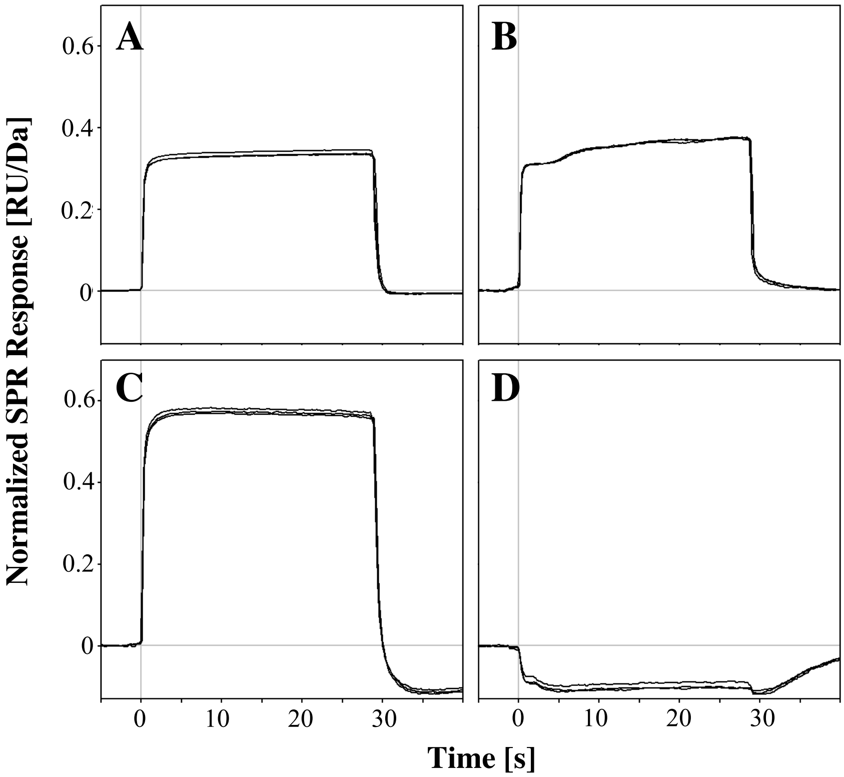


Figure 3-10: Examples of drug compound injections during ranking experiments (333 μM, triplicate injections). Four classes of curve shapes can be distinguished: proper block signals (**A**, digitoxin), multiphase association (**B**, diazepam), negative post-injection baseline (**C**, naproxen), negative binding signals (**D**, L-tryptophan).

Salbutamol could not be classified because of its low binding signal intensity. The two coumarin derivatives warfarin and phenprocoumon only showed a small negative effect on the post-injection baseline, while naproxen induced the most pronounced changes. Multiphasic binding of nitrazepam was less obvious than those of diazepam and progesterone, but resulted in an increased post-injection baseline. Interestingly, curve shape classes did not correlate clearly with binding site preferences, substance class, acidity, or log P value. However, site II ligands tended to show more deviations from an expected block signal. The binding behavior of such ligands with atypical binding behavior was further investigated (see *section* 3.3.5).

3.3.3 Influence of ionic strength of sample/running buffer

The choice of an appropriate buffer system is crucial for any type of biosensor assay. While two of the published HSA assay studies used 67 mM PBS buffer [20] the most recent one switched to 10 mM PBS [22]. Biacore usually recommends buffers of quite low strength (e.g. 10 mM HEPES buffer) with an addition of 150 mM NaCl (for suppressing electrostatic effects with the carboxyl surface) for their assays. Interestingly, the published Biacore binding data for some analytes are very consistent, while showing larger differences for others like naproxen (*Table* 3-3; a detailed list with available K_D values can be found in *appendix* B2). When taking a closer look at these data, anionic compounds seem to be especially affected. In order to evaluate any dependence of binding affinity on changes in buffer strength, a set of six substances was measured in three buffer systems (10/150, 67/150, 67/93 mM phosphate/NaCl). All compounds generated regular binding curves in all three systems and deviations between the experiments were within an order of magnitude (*Table* 3-6).

Table 3-6: Absolute and relative affinities of several marker compounds in PBS buffers of different ionic strength (10/150, 67/150, 67/93 mM phosphate/NaCl). All buffers contained 3% DMSO and were at *pH* 7.4.

Compound	**K_D [μM] 10/150 mM**	**K_D [μM] 67/93 mM**	**K_D [μM] 67/150 mM**
Warfarin	2.9 ± 0.4	2.8 ± 0.6	2.1 ± 0.03
Diazepam[a]	-	-	48.2 ± 20.0
Digitoxin	74.2 ± 1.6	72.1 ± 1.7	71.5 ± 0.2
Naproxen	12.2 ± 1.9	22.9 ± 4.4	27.7 ± 1.3
Indoprofen	0.7 ± 0.2	4.7 ± 1.2	2.1 ± 0.3
Cholic acid	21.2 ± 6.4	105.7 ± 6.7	46.9 ± 4.5
Quinine	1013.0 ± 25.6	1150.0 ± 78.1	1613.3 ± 32.2

[a] Reliable data for diazepam could only be collected in a single experiment.

To visualize the susceptibility on buffer changes more clearly, all K_D values were divided (normalized) by the value in 10 mM PBS (*Fig*. 3-11).

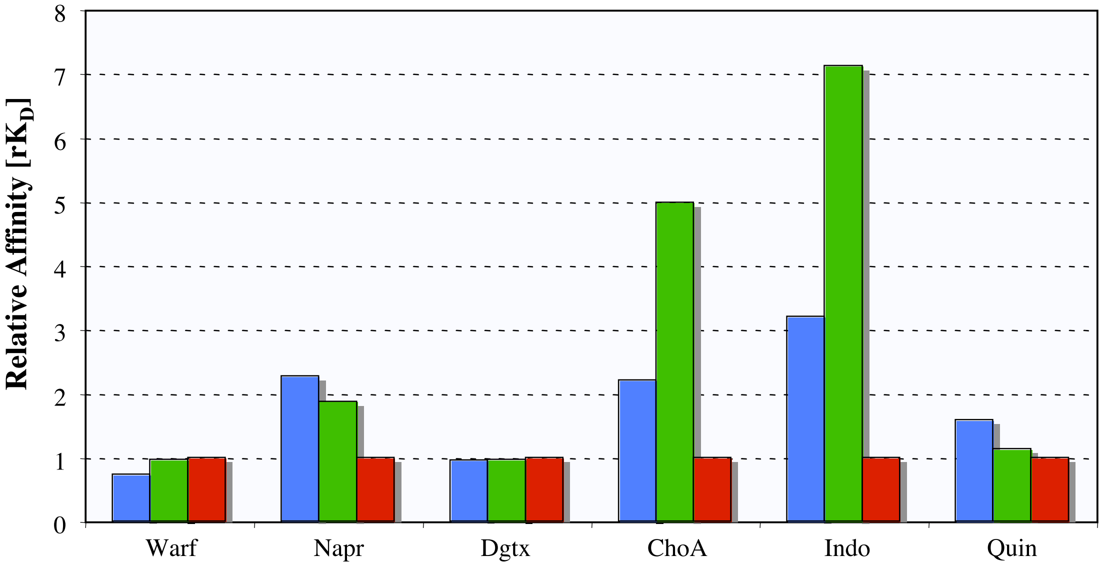


Figure 3-11: Relative affinities (rK_D) in PBS buffers *pH* 7.4 3% DMSO with different ionic strength (red = 10 mM phosphate/150 mM NaCl, green = 67/93 mM, blue = 67/150 mM). All affinity values were divided by the K_D in 10/150 mM PBS.

Comparison of the relative binding affinities (*Fig*. 3-11) shows that uncharged ligands like warfarin or digitoxin are nearly unaffected by changes in ion strength. Charged ligands, however, are clearly dependant on both buffer capacity and ionic strength. Two effects might be responsible for this behavior. First, the existence of a counter-ion like sodium directly determines the charge and acidity of the anionic molecules, therefore influencing ionic interactions or even conformational isoforms. Such a charge dependant behavior was described for several indole derivatives [40]. Second, chloride ions are known to interact with HSA (see *section* 3.1.3) and compete with several ligands. Especially site II ligands often show a significant competition behavior with chloride ions [4]. For example, the medium-chain fatty acid octanoate showed only one-third of the affinity in 130 mM chloride buffer compared to pure phosphate buffer [41]. Similar effects have been demonstrated for L-tryptophan [40, 42] and diazepam [4]. These findings demonstrate the importance of carefully matching both buffer capacities and ionic strength between experiments.

3.3.4 Influence of HSA redox state on drug binding

Around 30% of the unpaired cysteine residue (Cys-34) of HSA forms mixed disulfides with GSH and cysteine, dimers or higher oxidized states (see *section* 3.1.2).

Commercial preparations of HSA were shown to contain an increased fraction of oxidized HSA (up to 75%, depending on the manufacturer) [24]. Shifts from mercaptalbumin to non-mercaptalbumin have also been observed during storage [24], in elderly patients [43], and after physical exercise [44]. While drugs directly binding to Cys-34 (see *section* 3.1.3) are clearly affected by the redox state, its contribution on other binding sites is still not clear. The influence of HSA oxidation was therefore investigated by artificial reduction or oxidation of Cys-34. Treatment with different thiol-reactive reagents as described in *section* 3.2.9 resulted in both oxidized and reduced forms (*Fig*. 3-12, *Table* 3-7).

Table 3-7: Modification of HSA Cys-34 using different thiol-active reagents. All samples were incubated under the stated conditions and purified using size exclusion chromatography. Total HSA concentration was determined by measuring UV absorbance at 279 nm and total free thiol concentration was determined by Ellmans assay.

Reagent	none [a]	GSH	GSSG	Cys	DTT	DTT
Excess	-	1:10	1:10	1:10	1:1	1:5
Incubation	3 h	3 h	3 h	3 h	1 h	2 h
Temp.	37°C	37°C	37°C	37°C	37°C	37°C
Free Cys-34	29%	59%	20%	5%	88%	165%

[a] HSA from Sigma without any thiol-modifying treatment but also separated by size exclusion.

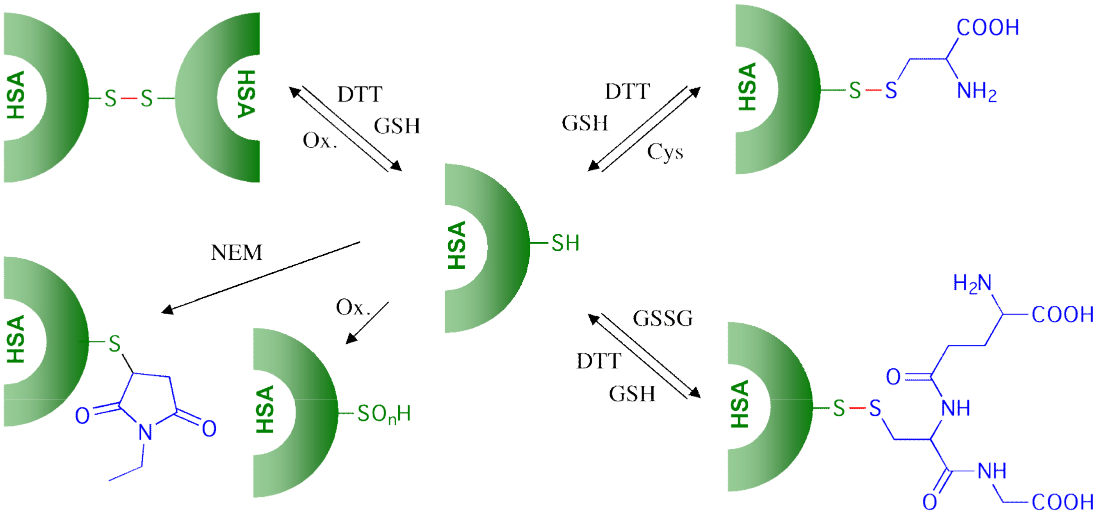


Figure 3-12: HSA oxidation states and their conversion. DTT and the reduced form of glutathione (GSH) are able to selectively reduce Cys-34 while oxidized glutathione (GSSG) and L-cysteine (Cys) increase the fraction of mixed disulfides. N-Ethylmaleimide (NEM) covalently blocks all residual free cysteine residues. Oxidative stress conditions like metal ions, hydrogen peroxide, etc. (Ox.) are reported to increase dimeric and higher oxidized forms of HSA [45].

With a fraction of around 30% free Cys-34, the untreated HSA from Sigma is exactly within the usual ranges of commercial batches (25-59% [24]). As expected, the reduced form of glutathione (GSH) significantly increased the free fraction while the oxidized form (GSSG) slightly decreased it. The most effective oxidizing agent was found to be L-cysteine, which formed mixed disulfides with Cys-34 very effectively after 3 hours at 37°C. *N*-Ethylmaleimide could also successfully be used to oxidize Cys-34 (less than 5% free Cys), but this would lead to a non-natural form of the HSA thiol and might influence binding properties in a different way (size, charge state) than glutathione and cysteine. Incubation of HSA with a fivefold excess of DTT for several hours [6] clearly led to the reduction of more than one cysteine (i.e. 165% instead of 100%). However, less than one disulfide bridge was cleaved in average under these conditions, since this led to three free cysteine residues (i.e. 300%) per molecule HSA. To fully avoid cleavage of internal disulfide bridges, both the DTT excess and the incubation time were reduced to yield a free fraction of nearly 90%. The relatively high tolerability of reducing agents is in good agreement with the observation that the helical environment of HSA protects the disulfide bridges from reduction [3].

After immobilization of untreated, reduced (DTT, 1:1) and oxidized (Cys 1:10) HSA to different flow cells of a sensor chip, binding properties were evaluated by injection of marker compounds targeting the two major drug binding sites (Sudlow sites I and II, *Table* 3-2). To avoid any ionic masking effects, the experiments were repeated in three buffer systems with varying buffer strength (10/150, 67/150, 67/93 mM phosphate/NaCl). Binding affinities were determined in the same way as in high-resolution experiments and compared between the HSA surfaces (*Table* 3-8).

Table 3-8: Absolute and relative affinities of several marker compounds on native, reduced, and oxidized HSA surfaces. All values are averages over three experiments.

Compound	**K_D [μM] native**	**K_D [μM] reduced**	**K_D [μM] oxidized**	**$rK_{D\,red}$[1]**	**$rK_{D\,ox}$[1]**
Warfarin	2.7 ± 0.5	2.9 ± 0.6	2.2 ± 0.2	1.1	0.8
Diazepam[2]	27.4	49.8	67.3	1.8	2.5
Digitoxin	72.6 ± 1.2	73.2 ± 2.3	72.0 ± 1.9	1.0	1.0
Naproxen	23.3 ± 7.8	20.5 ± 8.4	19.0 ± 8.2	0.9	0.9
Indoprofen	2.3 ± 1.3	2.5 ± 2.2	2.7 ± 2.7	1.2	0.9
Cholic Acid	58.8 ± 36.0	57.2 ± 48.6	57.8 ± 45.5	0.9	1.0

[1] $rK_{D\,red/ox}$ = relative affinities of reduced/oxidized surfaces compared to the native HSA surface.

[2] reliable data for diazepam could only be collected in a single experiment; therefore, no standard deviation is specified.

Only Diazepam showed a significant difference in affinity between the surfaces. However, this is caused mainly by the unusual binding mode of this substance (see *section* 3.3.2) and is more than ten times higher than in literature [3]. No one of the other compounds showed a significant change in its binding affinity in dependence of the redox state. On the other side, the redox stability of the surfaces can hardly be tested. Even though the reduced and oxidized HSA in solution was found to be stable at 4°C for several days (only a minor increase of less than 3% within one week was detected for untreated HSA), redox processes - predominantly oxidation - could occur on the chip during the course of experiments. The higher standard deviations for anionic compounds are mainly caused by the differences in buffer strength and are in agreement with the experiments in *section* 3.3.3.

3.3.5 Ligand-induced conformational changes

When comparing the sensorgrams of the investigated drug compounds from ranking assays (see *section* 3.3.2), diazepam and L-tryptophan showed atypical curve shapes and were therefore analyzed more extensively. Interestingly, no sensorgrams for these two compounds have been published so far, despite their importance as known binders or even marker drugs. One focus of our studies was the possibility of conformational changes induced by these ligands. Several HSA-binding drugs are reported to show allosteric effects [12] and ligand-induced conformational changes were recently expected to cause changes in SPR signal intensity [32] (see *chapter* 7).

L-tryptophan was the only compound causing negative binding signals during ranking. This unusual behavior could be confirmed in a high-resolution assay (*Fig*. 3-13A). However, mirroring the sensorgrams (*Fig*. 3-13B) led to concentration-dependant binding curves, which could be plotted and fitted using steady state affinity analysis (*Fig*. 3-13C). Triplicates showed a much higher accuracy and fitted better to a 1:1 binding model when the signal was evaluated just after injection end (i.e. 32-35 s from injection start) rather than on the steady state phase (25-27 s). A possible explanation of this behavior might be an overlay of a negative signal effect caused by a conformational change and a smaller positive signal contribution of increasing surface concentrations of L-tryptophan. After injection end, the dissociation signal is generated only by one component (most likely the conformational change). The K_D value before and after injection end were 165±5.6 μM and 173±7.4 μM, respectively. These values are in good agreement with results from capillary electrophoresis studies (150 μM) [18].

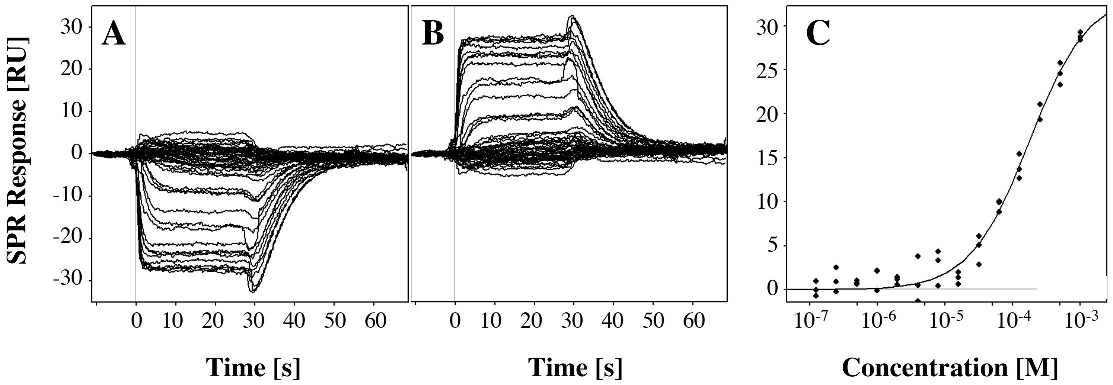


Figure 3-13: L-Tryptophan binding on HSA. High-resolution curves (0.1-1000 μM) after double referencing (**A**) and multiplication by -1 (mirroring, **B**). **C:** Steady state affinity plot evaluated just after injection end.

Since shifts in *pH* were shown to influence the SPR signal of HSA (see *section* 3.3.6) the *pH* of the sample buffer with and without 1 mM L-tryptophan was determined (*Table* 3-9). The observed shift from *pH* 7.40 to 7.38 was negligible and would rather lead to a very small positive than a negative signal effect (see *section* 3.3.6). If the tryptophan molecule would cause the negative binding signal, other known binders with similar structure should induce the same effect. To show this correlation, HSA binding of a metabolic product of tryptophan, L-kynurenine, was investigated in parallel (*Fig*. 3-14).

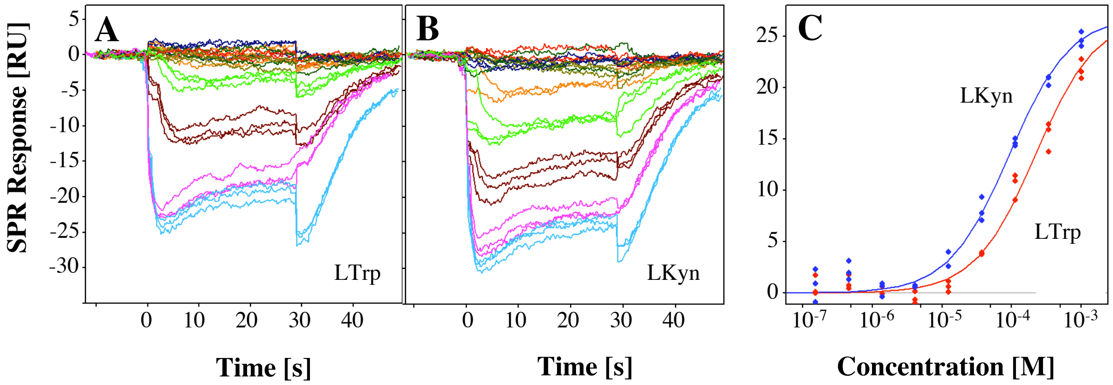


Figure 3-14: Binding curves of L-tryptophan (**A**, LTrp) and L-kynurenine (**B**, LKyn) on HSA (triplicate injections, 0.15-1000 μM). **C:** Steady state affinity plot of both compounds after mirroring.

L-Kynurenine shows the same negative binding behavior as L-tryptophan. With a K_D value of 89±2 μM the HSA affinity is higher compared to tryptophan (192±5 μM). These findings are in agreement with literature [46], where a %bound$_{HSA}$ value of 92% for L-tryptophan and more than 95% for L-kynurenine are reported. Affinity values of

L-kynurenine are only available for BSA binding [47], which is much stronger than the value obtained in the Biacore assay (BSA: $K_D \sim 5\ \mu M$, HSA/Biacore: $K_D \sim 90\ \mu M$). The lower affinities for both compounds compared to literature might also be caused by competition reactions with chloride ions as described in Peters [4].

Diazepam binding curves have a clearly different shape compared to warfarin or naproxen as well as to L-tryptophan. Even though the sensorgrams were all positive in sign, at least two time-dependant stages (i and ii) were observed. A first rapid association phase is followed by a further curve increase after approximately five seconds (*Fig*. 3-15A). However, depending on the surface, an additional phase (iii) after 20 s could be detected (*Fig*. 3-15B) in some cases, which was also manifested in a slightly prolonged dissociation phase. Neither buffer strength nor HSA oxidation state clearly influenced the curve shape.

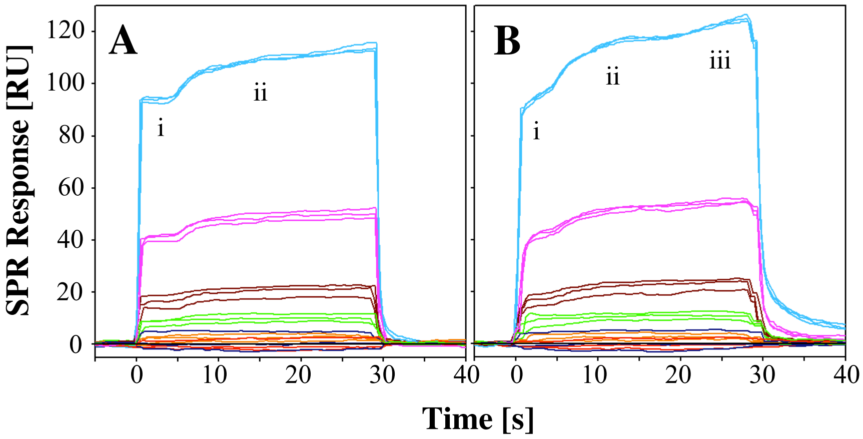


Figure 3-15: Diazepam binding curves (amine coupling, triplicate injections, 0.15-1000 μM). Multiphase association with two (A) or three (B) time-dependant stages (i, ii, iii) depending on the surface.

Two hypotheses for this phenomenon might be reasonable - first, the existence of two separate binding sites and second, the induction of a conformational change upon binding. In the case of the binding site model, binding to the first and readily accessible site was rapid while binding to the second site was not favored or even only possible through an allosteric effect and therefore slower and retarded. Indeed, earlier studies proposed the existence of one or more secondary diazepam binding sites [48]. However, simultaneous binding to two distinct sites is expected to occur simultaneously, even if allosteric effects are involved. If a direct competition was responsible for this effect, saturation and, therefore, a clear change in ratio of the two phases would be expected. Furthermore, nearly all known binders show rapid association and dissociation phases, whereas the second phase of diazepam binding is clearly slower. Conformational changes, on the other side, may vary from fast (e.g. 100 ms for the N-F transition) to rather slow (e.g. in the case of N-B transitions). The time lapse between the two phases observed in the binding curve might be explained more easily with the conformational model. Since not every change in conformation is expected to be readily visible as an SPR shift (see *chapter* 7), only later stages might lead to a detectable change in shape or charge distribution.

Interestingly, both L-tryptophan and diazepam show one slow binding phase, indicating a possibly common mechanism in the signal effect. Whereas the slow step occurs during association in the case of diazepam, L-tryptophan possesses a decelerated dissociation rate. The interference by diazepam with tryptophan binding has been demonstrated [4, 49]. This leads to the conclusion that this specific location within Sudlow site II might be especially susceptible to conformational changes upon ligand binding.

However, such a hypothesis should be verified by another method. One very reliable way of proofing conformational changes upon ligand binding is the use of X-ray crystallography. Unfortunately, HSA crystal structure determination is rather difficult [3] and there are only a few co-crystallization structures available. The only single drug molecules bound to HSA are the anesthetics halothane and propofol, which are not readily available for SPR analysis due to size and volatility issues and seem to rather stabilize than change HSA conformation [50]. In addition, there is a HSA structure in complex with warfarin, but it also contains fatty acid molecules (myristic acid), which themselves induce a large conformational change and enhance warfarin binding. Therefore, only small conformational changes could be detected when compared to the HSA-myristate complex without warfarin [13]. No high-resolution crystals are available for site II binders like diazepam or L-tryptophan.

Circular dichroism and fluorescence measurement are also two well-established methods for the detection of conformational changes. However, since fluorescence assays monitor changes of the single intrinsic tryptophan residue (Trp-214), binding of extrinsic tryptophan molecules will interfere with such an analysis. Circular dichroism, on the other side, detects the optical activity of asymmetric molecules in the far and near UV range. Even though it is possible to see conformational changes using this method, differentiation between changes of the protein and the drug is very difficult. Even molecules that are achiral show optical activity when immobilized at a certain conformation upon protein binding [19].

A biointeraction chromatography study published recently [51] supports the theory of ligand-induced conformational changes by L-tryptophan. While an excess of phenytoin in the mobile phase only showed a competitive interaction behavior on L-Trp binding, the reversed experiment (excess L-Trp) altered phenytoin binding in a negative allosteric way, as it is typical for conformational effects. It was also proposed that the binding site of L-tryptophan might be located deeper in the pocket of Sudlow site II than the one of phenytoin [51].

While no reference literature could be found about negative binding signals induced by ligand binding, there is one study [32] that also reported negative post-injection baselines similar to those observed in the case of naproxen and other anti-inflammatory and coumarin drugs. This effect was demonstrated for the interaction of carbohydrates with mannose binding protein and was also attributed to conformational changes. More evidence for this hypothesis came from the example of tissue transglutaminase binding, where binding of calcium ions (35 Da) induced much higher binding signals (> 1,000 RU) than expected from a simple mass increase [32]. Since transglutaminase is known to undergo significant conformational changes upon binding of calcium ions [52], this was explained to be the main reason for this behavior.

The observed anomalies in the binding curves were highly reproducible over different experiments with changes of flow cells, chips, or analyte preparation. However, depending on the HSA surface, effects like multiphasic association and especially negative post-injection signals were more or less pronounced. No obvious correlation between changed parameters (flow cell, immobilization density, HSA oxidation state) has been found. A possible explanation is that small fluctuations in the immobilization procedure (activation time and efficacy, immobilization buffer *pH*, HSA purity) might influence the way HSA is fixed on the carboxymethyl dextran surface and, as a consequence, its susceptibility to conformational changes. Interestingly, both Biacore-HSA studies performed in 67 mM PBS [20, 21] only showed ranking and K_D values but no sensorgrams (except for warfarin in [21]). The binding curves published in the most recent Biacore-HSA study [22] show nearly no negative post-injection effect for naproxen and warfarin but clearly for salicylic acid. All those studies used a slightly higher immobilization *pH* (5.2 instead of 5.0), which might have an influence on this effect.

3.3.6 *pH*-induced conformational changes

A possible contribution of conformational changes of albumin on SPR signals as suspected in the case of ligand binding was further evaluated using a different approach. HSA is known to undergo several conformational changes upon alteration of *pH* [3]. Different studies are published, where this change was followed by circular dichroism and intrinsic fluorescence [28], changes in Cys-34 reactivity [53], or modified ligand binding behavior [4]. Since such changes of the protein isoform might also change the electron density around the gold surface and therefore the SPR signal, this effect was studied by injecting buffer blanks at various *pH* values on a HSA surface (*Fig*. 3-16).

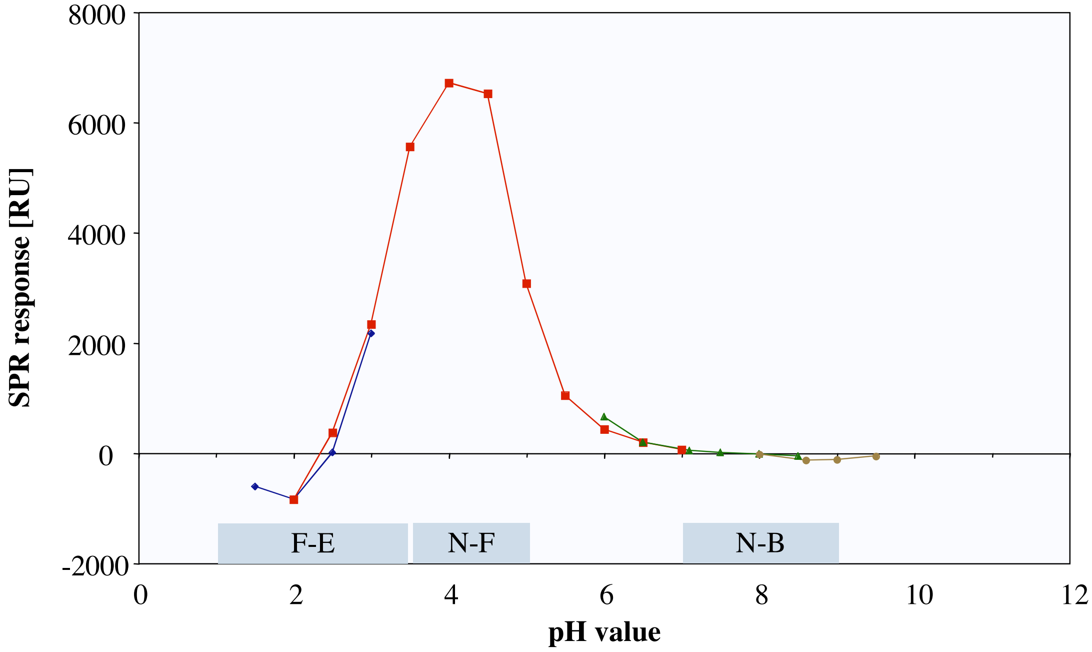


Figure 3-16: *pH*-induced conformational changes monitored using Biacore. Blank injections were performed with buffer systems covering a *pH* range from 1.5 to 9.5 (*pH* 1.5 - 3.0: 10 mM glycine; blue, *pH* 2.0 - 7.0: 10 mM citrate; red, *pH* 6.0 - 8.5: 10 mM phosphate; green, *pH* 8.0 - 9.5; 10 mM borate; brown). The three known transitions are indicated at the bottom (F-E, N-F, N-B).

The observed *pH*-induced changes in SPR signal-intensity correlate very well with the described mechanism of conformational changes. The negative signals below *pH* 3.5 can be attributed to an acid induced expansion (or F-E transition), where the expanded structure leads to a decrease in electron density at the surface. The N-B transition between *pH* 7 and 9 was also visible as a (smaller) decrease in signal below baseline. With 7,000 RU the most remarkable change was a positive peak between *pH* 3.0 and 6.0. Here again, a conformational change (N-F transition) is described in literature [28].

The huge structural changes in this *pH* range were confirmed by measuring alterations of the intrinsic fluorescence of Trp-214 (*Fig*. 3-17).

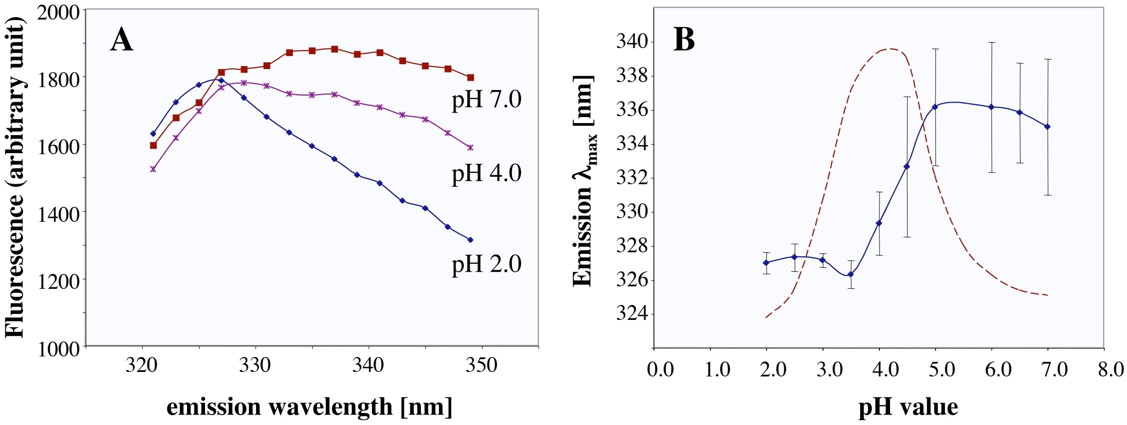


Figure 3-17: *pH*-induced conformational changes of HSA in 10 mM citrate buffer, monitored by intrinsic fluorescence measurement of Trp-214. **A:** Scan of Trp emission after excitation at 295 nm. Emission spectra in citrate buffer *pH* 2, 4, and 7 are overlaid. **B:** Shift in emission maxima between *pH* 2 and 7 (solid line). SPR curve in the same range is shown as dashed line (simplified from *Fig*. 3-16).

The emission maximum shifts based on the change in the microenvironment of Trp-214 correlate very well with the signals observed with Biacore (*Fig*. 3-17B). The relatively high standard deviations in the emission scan are caused by the shallow maxima at higher *pH* values compared to those at *pH* 2 (*Fig*. 3-17A). However, since the peak maximum exactly lies below the *pI* value of HSA, this effect could also be caused by higher protonation of amine groups and, as a result, increased surface attraction between HSA and the carboxymethyl dextran matrix. This second hypothesis is supported by the study of Paynter *et al.* [54], where electrostatic interactions around the *pI* values of three different proteins could be correlated with changes in SPR signal intensity. A possible electrostatic surface attraction effect was investigated by injecting HSA at low concentration in the same series of citrate buffer over an unmodified sensor-chip surface (*Fig*. 3-18).

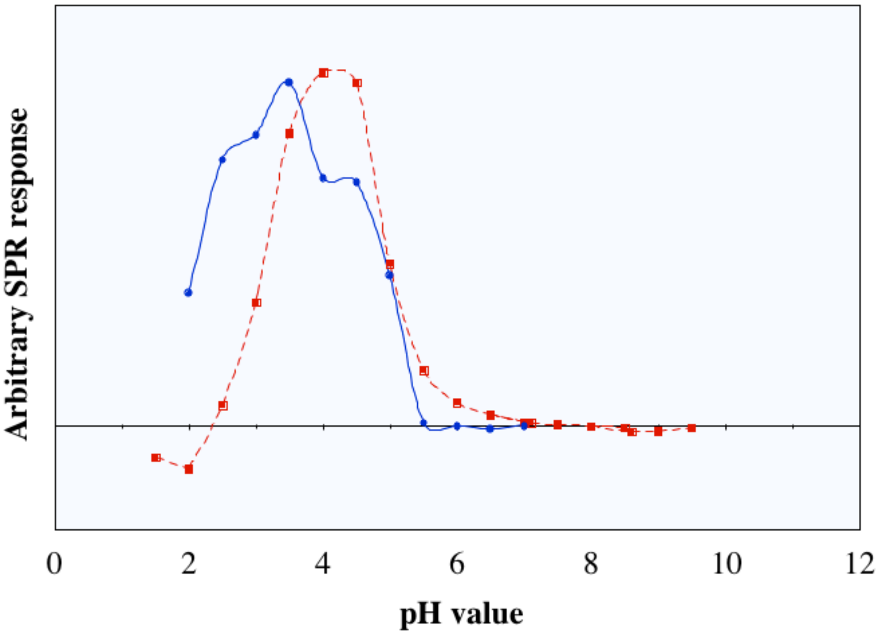


Figure 3-18: Normalized SPR responses of HSA surface attraction (blue solid line) and *pH*-induced signal effects (red dashed line; simplified from *Fig*. 3-16) at different *pH* values (10 mM citrate buffer).

As expected, surface attraction effects start just around the *pI* of HSA (5.2), i.e. between injections at *pH* 5.5 and 5.0. The course of the plot shows some similarities to the one of immobilized HSA but also some major differences. The signal intensity increases only below the *pI*, while there is no attraction at all above *pH* 5.2. Even though there is also a signal decrease at lower *pH* values, this effect is less pronounced and starts at lower *pH*. In addition, there seem to be at least two stages around the maximum. This might be explained by conformational changes in HSA, which led to a change in charge distribution on the surface. However, it is difficult to state, if these effects of HSA in solution are comparable in any way to the effects of immobilized HSA in its microenvironment. In conclusion, the *pH*-dependant signal changes on the HSA surface are likely to be an overlay of electrostatic and conformational effects.

3.3.7 Overlay of different signal effects

Since changes in *pH* or in ionic strength were shown to directly influence the SPR signal intensity, further experiments were performed to investigate how much these effects affect affinity data. Many of the current Biacore studies are performed in buffers of low ionic strength (e.g. 10 mM phosphate, HEPES, etc.). Especially when dealing with substances featuring free carboxyl or amine groups, this might lead to a shift in *pH* of the sample solution. Salicylic acid and L-tryptophan were chosen as model compounds for an acidic and a zwitterionic structure. While a 1 mM solution of salicylic acid in 10 mM PBS 3% DMSO significantly decreased the *pH* by more than 0.25 units, the *pH* of the tryptophan solution at the same concentration remained stable. As expected, increasing the buffer capacity from 10 to 67 mM reduced *pH* shifts remarkably (*Table* 3-9).

Table 3-9: Ligand-induced shifts in buffer *pH* at two different phosphate buffer capacities.

Compound	**10 mM PBS 3% DMSO**	**67 mM PBS 3% DMSO**
no addition (blank)	*pH* 7.40	*pH* 7.40
1 mM salicylic acid	*pH* 7.14	*pH* 7.35
1 mM sodium salicylate	*pH* 7.38	*pH* 7.39
1 mM L-tryptophan	*pH* 7.39	*pH* 7.39

Salicylic acid was therefore screened on the same HSA surface in the low and high capacity PBS buffer in order to discriminate between real binding signals and SPR shifts possibly caused by *pH*-induced conformational changes (*Fig.* 3-19).

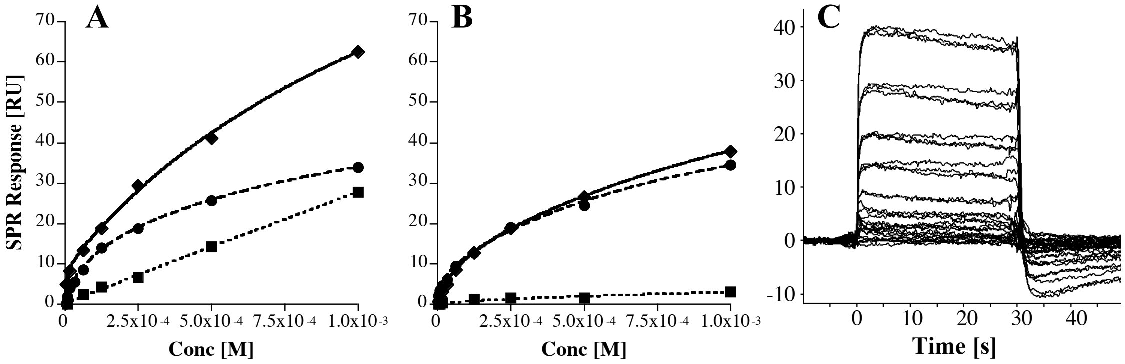


Figure 3-19: HSA binding of salicylic acid. Steady state affinity plots for salicylic acid (υ; solid line) and sodium salicylate (λ; dashed line) in 10 mM PBS *pH* 7.4 (**A**) and 67 mM PBS *pH* 7.4 (**B**) over a concentration range of 1 mM - 4 μM. Dotted lines (ν) indicate a dilution series of buffer blank at *pH* 7.14 (in A) and 7.35 (in B). An overlay of typical sensorgrams for salicylic acid in 67 mM PBS is shown in **C**.

Despite its low molecular weight of only 138 Da, salicylic acid binding generated detectable sensorgrams in both buffers. Injection triplicates were slightly more accurate in 67 mM PBS than in 10 mM PBS (*Fig*. 3-19C). At higher concentrations deviations between signals at high and low buffer capacity were getting more pronounced. Both data sets were fit to a two independent-sites binding model to determine K_D values.

While salicylic acid in 67 mM PBS had a K_D of 124±87 μM, the affinity for the same interaction in 10 mM PBS was significantly lower (1460±771 μM). At the highest concentration (1 mM salicylic acid) the signal difference between the two systems is approximately 30-40 RU. A linear signal increase was observed on the same HSA surface, when a dilution series of a 10 mM PBS buffer at *pH* 7.14 in running buffer (*Fig*. 3-19A&B). This clearly demonstrated that the binding data for free salicylic acid in 10 mM PBS are an overlay of a mass increase and a *pH*-induced signal change as observed in *section* 3.3.6. As expected, there was no *pH* effect of sodium salicylate in either buffer system (*Table* 3-9).

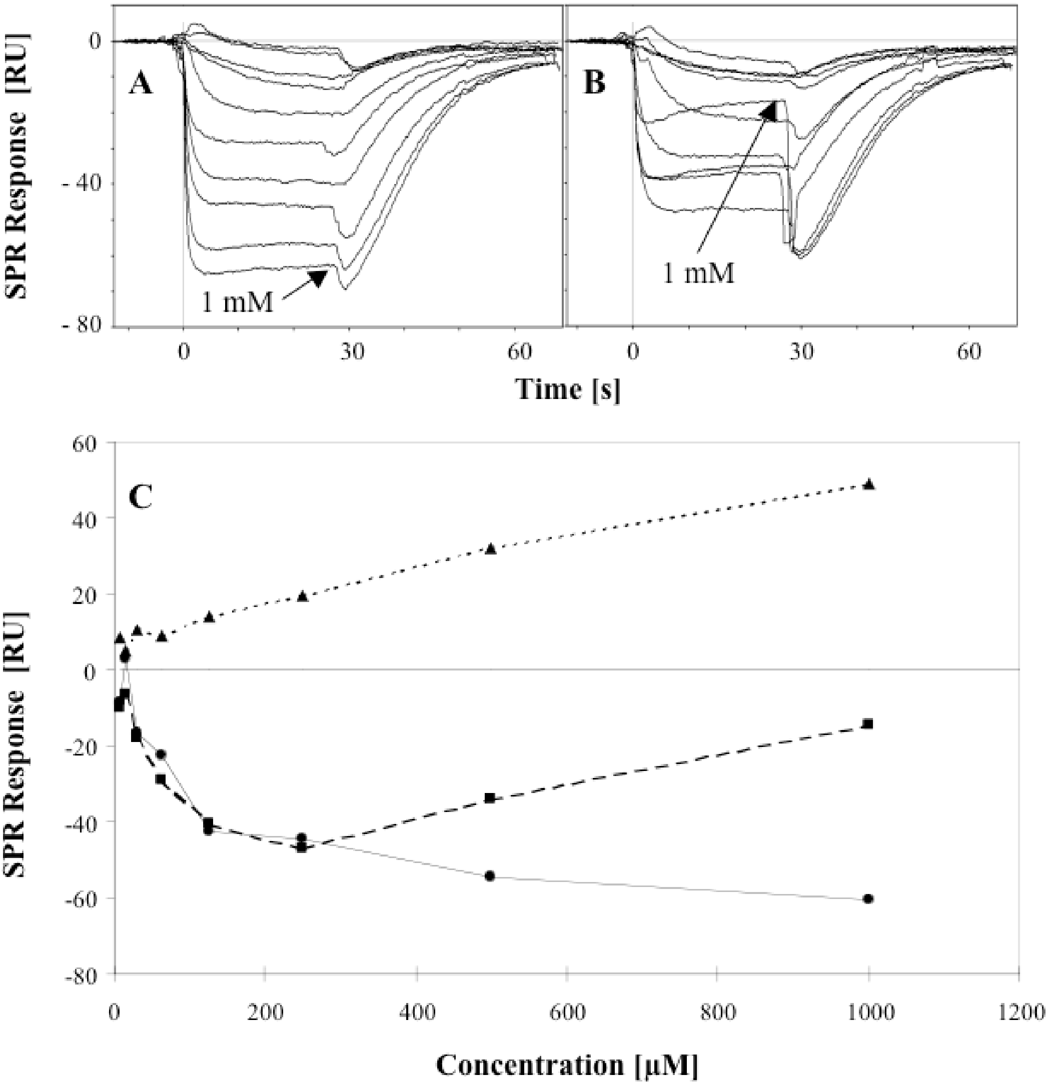


Figure 3-20: L-Tryptophan/HSA assay in 10 mM PBS with 3% DMSO (PBS3D) *pH* 7.40 as running buffer. Sensorgrams at various concentrations (8-1000 μM) in sample buffer PBS3D at *pH* 7.40 (A) and *pH* 7.14 (B). Highest concentrations (1 mM) are indicated by an arrow. C: Average responses of samples at *pH* 7.40 (solid line) and *pH* 7.14 (dashed line) are plotted against concentration. A twofold dilution series of PBS3D *pH* 7.14 in running buffer is shown as dotted line.

In order to demonstrate this effect more clearly, the experiment was extended to L-tryptophan, which typically generated negative binding signals (see *section* 3.3.5). For this purpose 1 mM tryptophan was directly dissolved in 10 mM PBS *pH* 7.14 (equivalent to the sample buffer *pH* of 1 mM salicylic acid). This solution was further diluted and screened in 10 mM PBS *pH* 7.4 as running buffer. Dilution series of tryptophan at *pH* 7.4 and a buffer blank at *pH* 7.14 in running buffer were also screened as controls (*Fig*. 3-20).

While L-tryptophan at *pH* 7.4 showed a normal but negative binding curve, a serial dilution of PBS *pH* 7.14 generated an almost linear signal increase up to 40 RU. An overlay of normal binding curve (up to 111 μM) and a *pH*-induced signal increase (111 - 1000 μM) could be detected in the case of tryptophan at *pH* 7.14. While the overlay of analyte-induced signals and *pH*-induced conformational changes are obvious in the case of L-tryptophan because of different directionalities, it is impossible to separate the two events in the case of salicylic acid. This implies that free acids should be avoided if possible and that the *pH* value of starting concentration should be carefully controlled.

3.3.8 Thiol Immobilization of HSA

Since oxidation of the unpaired cysteine residue 34 did not change the binding behavior of HSA (see *section* 3.3.4) this site was chosen for an oriented immobilization approach on a thiol-functionalized sensor chip. Thiol-coupled surfaces are expected to be fully regenerable, therefore saving preparation time and cost. In addition, a change of immobilization site from primary amine to a single thiol might have additional benefits. Cys-34 is located in domain I of HSA while the two primary drug binding sites are positioned in domains II and III respectively. Binding site II (domain III) is reported to consist of a hydrophobic cavity next to a cationic patch. If this patch consists of lysine residues, they could be blocked by the immobilization, which would directly influence the binding signals (either by steric effects, conformational changes or blocked functional groups). The crystal structure around the binding sites was therefore analyzed in more detail (*Fig*. 3-21).

Potential target groups for amine coupling reactions can be identified around both major drug binding sites. However, since there are much more surface-accessible lysine residues distributed over the whole protein surface, a direct effect on the binding signals may be possible, but difficult to quantify. Binding of atypical ligands like diazepam or L-tryptophan is therefore of special interest.

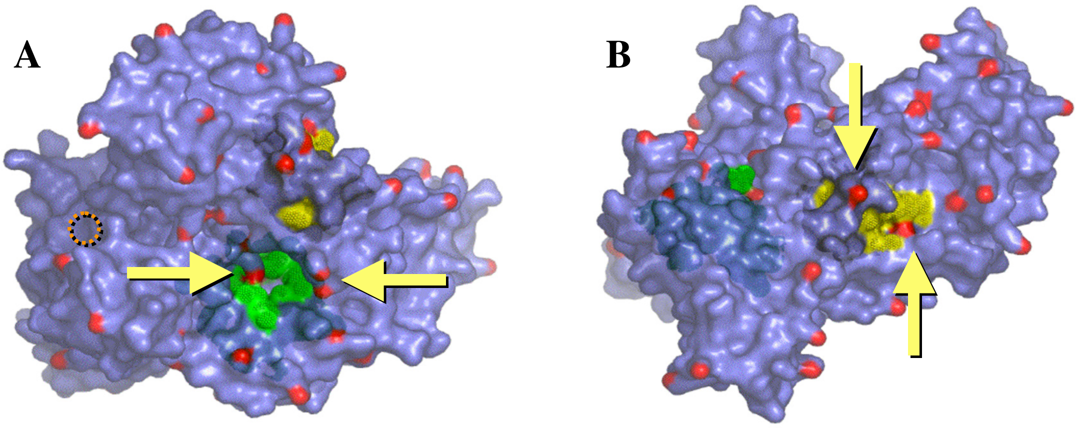


Figure 3-21: Distribution of surface-accessible lysine residues over the HSA crystal structures with focus on Sudlow binding site I (**A**, green) and site II (**B**, yellow). Primary amine groups of lysine residues are highlighted in red. Arrows indicate potential amine coupling targets around the binding sites and a dotted orange circle (in A) represents the approximate position of Cys-34 (hidden on the opposite side of the protein).

Optimization of coupling conditions

Standard ligand thiol immobilization using immobilized PDEA [29] did only lead to low-density surfaces (*appendix* B4). Therefore, a novel in-situ immobilization method was developed. In this approach, the reduced form of glutathione (GSH) was used as a mild reducing agent (see *section* 3.3.4) and DTNB as an activator group. Incubation of HSA with a 28-fold excess of both GSH and DTNB at low *pH* generated a reactive HSA fraction, which could be directly immobilized on the thiol chip (*Fig*. 3-22). Removal of one or both agents dramatically decreased the HSA surface-density (*Fig*. 3-22B).

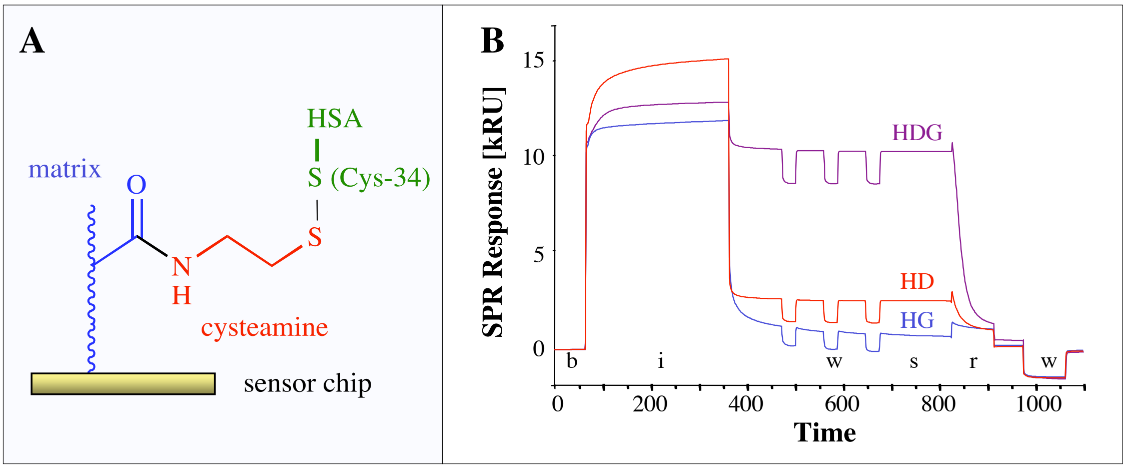


Figure 3-22: Thiol immobilization of HSA on a thiol-functionalized sensor chip. **A:** Schematic overview of the assay surface. **B:** Immobilization sensorgrams of HSA mixtures with GSH (blue, HG), DTNB (red, HD) and a combination of both agents (purple, HDG) showing different phases of the immobilization process: b = baseline, i = injection of the mixture, w = NaOH wash, s = stable HSA surface, r = DTT reduction.

Even though glutathione is involved in the formation of disulfide bonds both *in vivo* and *in vitro* its main function seems to be the one of a reducing agent in this case. This was supported by the finding that another mild reducing agent, 2-mercaptoethylamine (MEA), led to similar result. DTT could also be used, but surfaces generated this way were less stable (*appendix* B5). Lower *pH* values during immobilization increased surface density and stability much more than expected from a simple surface attraction. The reason for this tendency might be an expansion of the HSA molecule during the *pH*-induced N-F or F-E transition (see *section* 3.1.2), which makes Cys-34 more accessible for a reaction with the surface [4]. In addition, Cys-34 has an unusual low pK_{SH} value of 7 (compared to pK_{SH} 8.5 for free cysteine) making it to the most reactive thiol in human serum [53]. This atypical property of Cys-34 was demonstrated by its reactivity to 2,2-dithiopyridine, which is structurally related to DTNB. The authors reported a minimum in rate constant at *pH* 5 and two maxima around *pH* 2.5 (caused by the N-F-E transitions) and above *pH* 8, where the thiols are fully ionized [53].

Binding of long-chain fatty acids to HSA is known to induce massive conformational changes [55]. These transitions also affect the crevice where Cys-34 is situated and protected from oxidation. When adding three or more molar equivalents of oleic acid to HSA, the solvent accessibility of Cys-34 is remarkably increased, as demonstrated by fluorescent methods [56]. This effect is also visible when superimposing the crystal structures of HSA with and without fatty acids bound (*Fig*. 3-23A). Cys-34 is only accessible on the Connolly surface in the co-crystal with fatty acid molecules (*Fig*. 3-23B). Addition of oleic acid to the HSA immobilization mixture was therefore believed to increase immobilization density. Surprisingly, oleic acid did not improve but decreased immobilization efficacy, resulting in lower surface densities (*Fig*. 3-23C).

Finally, different ratios of HSA and the two additives (GSH and DTNB) as well as variations in immobilization *pH* and incubation time were evaluated. A ratio of 1:28 was found to be best suited, and two incubation steps (10 min without and 90 min in immobilization buffer) were necessary to yield optimum densities.

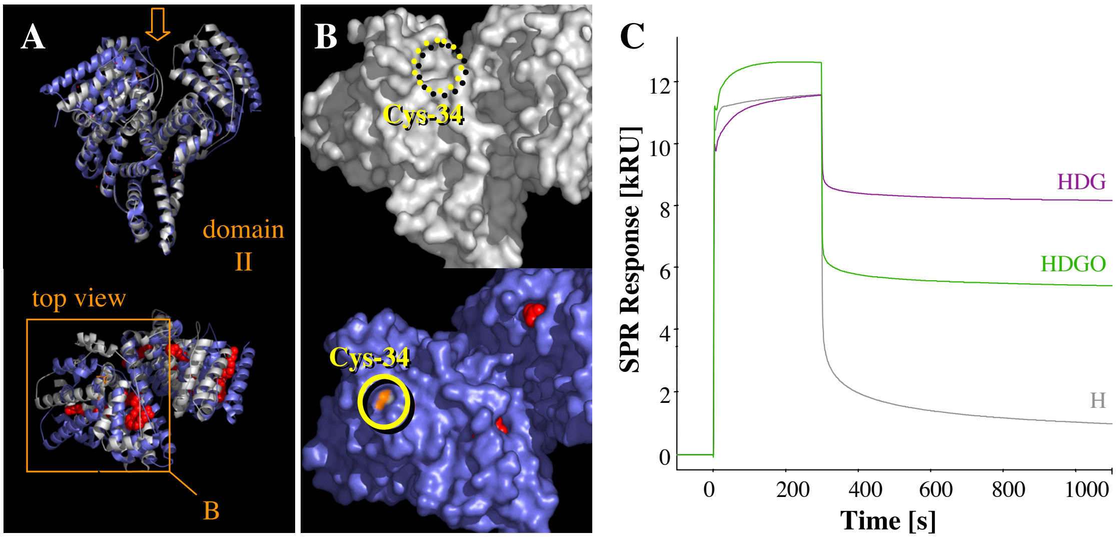


Figure 3-23: Effect of oleic acid on HSA structure and thiol coupling efficacy. **A:** Overlay of crystal structures of HSA alone (white) and in complex with oleic acid (blue). Oleic acid molecules are shown in red in a top view albumin (bottom). **B:** View on domain I of the protein surface of HSA with the position (yellow circles) and accessibility of Cys-34 (orange) in albumin with (bottom) and without oleic acid (top). **C:** Overlay plot of thiol coupling sensorgrams of untreated HSA (grey, H), and in the mixture with GSH/DTNB (purple, HDG) or GSH/DTNB/oleic acid (green, HDGO).

Evaluation of thiol-coupled surfaces

The stability of thiol-coupled surfaces was very high in running buffer. Compared to amine-coupled surfaces, where the surface density decreases during several hours and a special stabilization procedure is needed (see *section* 3.2.4), nearly no surface bleeding was detected in the case of thiol immobilization (*Fig*. 3-24A). The same trend was visible for warfarin activity: while the affinity after 10 hours was 87% of the initial value in the case of amine-coupled surfaces, thiol-coupled surfaces retained 98% activity after the same time period. The surprisingly high stability might be caused by the protective microenvironment of Cys-34 (see *section* 3.1.2). Nevertheless, surfaces generated in this way could easily be reduced with DTT and re-immobilized without further treatment (*Fig*. 3-24B).

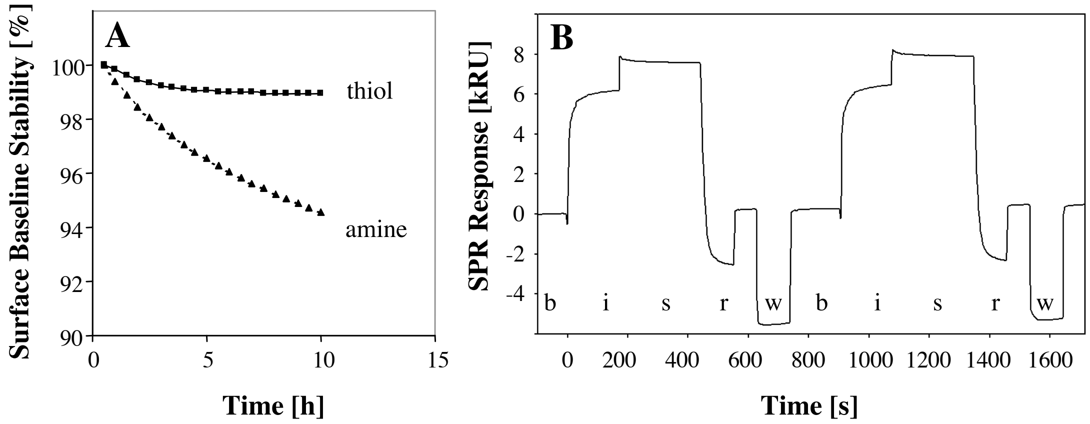


Figure 3-24: Evaluation of thiol-coupled surfaces. **A:** Surface stability during repeated buffer injection of amine-coupled and thiol-coupled HSA surfaces over 10 hours. **B:** Repeatability of thiol immobilization procedure (b = baseline, i = injection, s = HSA surface, r = DTT regeneration, w = NaOH wash).

While thiol-coupled surfaces fulfilled all prerequisites for a stable immobilization, the most important property is its ligand binding behavior for the various binding sites. Different marker compounds were screened over the thiol-immobilized HSA to monitor any differences in binding affinity induced by the immobilization method. Binding data obtained from the thiol-coupled surface were compared to those from amine-coupled HSA and to values from the literature (*Table* 3-10).

Table 3-10: Comparison of affinity data obtained on amine and thiol coupled HSA surfaces and reference values.

Compound	**K_D [μM] Amine**	**K_D [μM] Thiol**	**K_D [μM] Ref [b]**
Warfarin	2.9	2.9	2.5
Diazepam[a]	n.d.	n.d.	n.d.
Digitoxin	74.2	38.5	38
Naproxen	12.2	20.5	1.5
Indoprofen	0.7	1.7	n.d.
L-Tryptophan	173	149	n.d.
Quinine	1013	3090	2500

[a] No reliable data could be obtained by fitting diazepam plots.

[b] Comparable Biacore binding data from [22].

All of the tested marker compounds showed reproducible binding curves on the thiol-coupled surface. Binding affinities were very similar on both surfaces and deviations were usually lower than a factor of two. In addition, nearly all the values agreed very well with the ones found in literature. Compared to amine coupling, quinine affinity was threefold lower on the thiol surface but less than twofold from the reference value. The most prominent deviation to the literature value was detected in the case of naproxen, which was around or more then tenfold lower, respectively. Since naproxen generally showed higher K_D values over all experiments, this phenomenon might be caused by the substance quality. On the other hand, naproxen data also deviated throughout the available Biacore studies and a K_D of 26 μM can be found in Frostell-Karlsson *et al.* [20], which is even higher than the value obtained on the thiol surface.

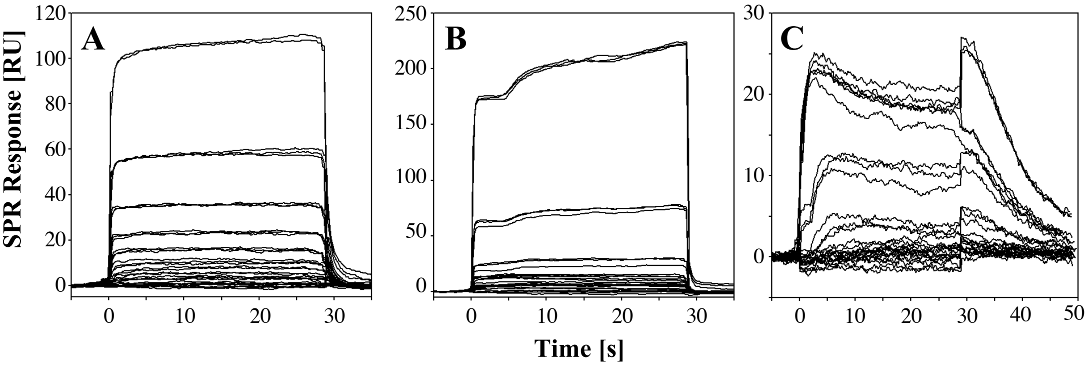


Figure 3-25: Triplicate injections of warfarin (**A**), diazepam (**B**) and L-tryptophan (**C**) on a thiol-coupled HSA surface (concentration range =0.15-1000 μM; L-Tryptophan were multiplied by -1).

Neither diazepam nor L-tryptophan showed a significantly different behavior on the thiol-coupled surface compared to the amine-coupled one (*Fig.* 3-25). L-Tryptophan signals were also negative after double referencing and had to be mirrored first (*Fig.* 3-25C). The resulting K_D value corresponds very well with the affinity obtained on the standard amine surface. Diazepam curves showed a typical multiphase shape (*Fig.* 3-25B) and could not be interpreted by standard fitting algorithms. These findings clearly demonstrate that the signal effects are not primarily caused by the immobilization method.

3.4 Conclusion

Human serum albumin is one of the most important proteins in the human body. Its huge concentration in the plasma and its unique possibility to bind a tremendous panel of structurally diverse compound with rather high affinity not only make it to an interesting but also to an indispensable target in drug discovery. Available assays showed that Biacore is a versatile tool for the acquisition of HSA-drug binding interactions, which offers many applications from ranking over high-resolution assays to competition studies. However, not all available data are consistent between those studies and with values obtained by alternative methods. Drugs binding to Sudlow site II and compounds with a carboxyl group showed higher deviations. For some important site II marker drugs like diazepam or L-tryptophan there are no Biacore data available at all.

Assay optimization and evaluation

Different assay parameters that might be involved in affinity or signal artifacts were therefore investigated in an optimized HSA assay system. Since DMSO is essential for the solubilization of hydrophobic drugs, the quality of DMSO and the resistance of vials had turned out to be highly important. When using tightly closing rubber caps and a cooled autosampler, triplicate injections were accurate when injected three times from a single vial instead of using separate vials. Together with a threefold dilution series ranging from 0.15 to 1000 μM and by using *Scrubber* instead of *BIAevaluation* the time for sample preparation, data acquisition and data evaluation could be remarkably decreased. Ranking and high-resolution data obtained with a panel of known binders on such an optimized assay correlated very well with literature and compound properties (log P and %bound$_{\mathrm{PPB}}$). Even small differences between structurally similar compounds could be evaluated this way.

Effect of changes in HSA redox state and running buffer ionic strength

The oxidation state of the single free cysteine residue (Cys-34) did not influence the affinity of the investigated drug compounds. However, drugs interacting directly with thiol groups (captopril, cisplatin, auranofan) are surely affected and a possible contribution for another set of compounds has been suspected in single experiments described in literature. Furthermore, HSA dimers (or higher oligomers) may have an influence on binding site accessibilities. Such an effect is reported e.g. for

L-tryptophan [57]. An experiment where the monomer and the oligomer fraction of a commercial HSA batch after size exclusion chromatography is immobilized on two separate flow cells might give new insights in this phenomenon.

The influence of buffer ionic strength was much more pronounced and resulted in rather large deviations in affinity values when changing buffer composition. This might be explained by shifts in charge state of both the ligands and the protein as well as with a direct competition of chloride ions with site II binders. These compounds, and especially structures carrying a carboxyl group, were especially affected by buffer changes. Variation in ionic strength should therefore be investigated more systematically and experiment in a chloride-free environment should be performed. In any case, running buffer composition should be carefully matched between different experiments. Furthermore, substances with free acid groups should be avoided, since they may change the *pH* of the sample buffer and lead to *pH*-induced conformational changes. This effect could be shown on the example of salicylic acid.

Signal effects induced by pH and ligands

Variations of *pH* around immobilized HSA also induced large positive and negative SPR signals. Response curves obtained during a *pH* screening correlated very well with known *pH*-induced structural transitions as described in literature or measured using intrinsic fluorescence and circular dichroism. Though, especially the large signal increase between *pH* 4 and 5 might also be influenced by charge effects (surface attraction). Another observation concerning a possible contribution of conformational changes to binding signals came from experiments with L-tryptophan and diazepam, which both showed unusual binding curves. While diazepam binding was characterized by a multiphasic binding, L-tryptophan generated negative binding signals, which nevertheless could be mirrored and fitted to a single-site binding model. These effects can hardly be explained by simple ligand binding and there are strong hints for a contribution of conformational effects. At least the phenomenon of negative post-injection baseline as observed in the case of naproxen and other HSA binders have been reported in one study [32]. Both *pH* and ligand-induced conformational effects may even overlay and mask a regular binding signal. Therefore, care has to be taken at sample preparation (sample solution *pH*) as well as during data evaluation and interpretation.

Negative binding signals as observed in the case of L-tryptophan are often regarded as a general problem with the binding assay and therefore rejected. However, as clearly shown in the current study, such data can be mirrored and fitted to concentration-dependant binding models. The albumin assay used in this study was successfully evaluated using different model compounds and validated with literature values, both from other Biacore experiments and from other methods.

Thiol-coupling of HSA

The single free cysteine residue of HSA (Cys-34) could successfully be used for a site-directed immobilization approach. When incubating HSA with a combination of reduced glutathione and DTNB, surfaces of high density and stability were generated on a thiol-functionalized sensor surface. The stability was even higher than for comparable amine-coupled surfaces. Since neither activation and deactivation steps nor a special stabilization procedure is needed, surface preparation times can be considerably decreased. Ligand binding properties for the major binding sites did not change significantly between the immobilization methods. One of the major advantages of surfaces generated by thiol-coupling is their ability to completely remove the protein under reducing conditions and re-immobilize a new batch without additional procedures. This reusability of sensor-chip surfaces reduces experiment costs and increases the flexibility in assay design. Possible applications for HSA are studies of albumin mutants, effects of glycation or lysine acetylation, or simply the replacement of ‘aged’ HSA between experimental series.

Albumin - a model protein for biochemistry?

Due to its high availability, low costs, its enormous stability and unusual ligand binding properties, HSA is often utilized as a substitute for a typical protein in many assays. However, as a protein, albumin is far away from typical. Its high sensitivity to changes in *pH*, charge or temperature as well as it complex binding patterns involving allosteric effects, direct competition and binding to multiple binding sites can lead to an overlay of effects that can hardly be evaluated.

3.5 References

[1] H. van de Waterbeemd, E. Gifford, *Nat Rev Drug Discov* **2003**, *2*, 192.

[2] D. R. Owens, *Nat Rev Drug Discov* **2002**, *1*, 529.

[3] D. C. Carter, J. X. Ho, *Adv Protein Chem* **1994**, *45*, 153.

[4] T. Peters, *All about albumin: biochemistry, genetics, and medical applications*, Academic Press, San Diego (Calif.), **1996**.

[5] X. M. He, D. C. Carter, *Nature* **1992**, *358*, 209.

[6] M. Sogami, S. Nagoka, S. Era, M. Honda, K. Noguchi, *Int J Pept Protein Res* **1984**, *24*, 96.

[7] G. Sudlow, D. J. Birkett, D. N. Wade, *Mol Pharmacol* **1975**, *11*, 824.

[8] G. Sudlow, D. J. Birkett, D. N. Wade, *Mol Pharmacol* **1976**, *12*, 1052.

[9] I. Sjoholm, B. Ekman, A. Kober, I. Ljungstedt-Pahlman, B. Seiving, T. Sjodin, *Mol Pharmacol* **1979**, *16*, 767.

[10] J. Chen, C. Ohnmacht, D. S. Hage, *J Chromatogr B Analyt Technol Biomed Life Sci* **2004**, *809*, 137.

[11] U. Kragh-Hansen, *Pharmacol Rev* **1981**, *33*, 17.

[12] B. Honore, *Pharmacol Toxicol* **1990**, *66 Suppl 2*, 7.

[13] I. Petitpas, A. A. Bhattacharya, S. Twine, M. East, S. Curry, *J Biol Chem* **2001**, *276*, 22804.

[14] G. Scatchard, W. T. Yap, *J Am Chem Soc* **1964**, *86*, 3434.

[15] H. H. Parikh, K. McElwain, V. Balasubramanian, W. Leung, D. Wong, M. E. Morris, M. Ramanathan, *Pharm Res* **2000**, *17*, 632.

[16] H. Aki, M. Yamamoto, *J Pharm Sci* **1994**, *83*, 1712.

[17] D. S. Hage, *J Chromatogr B Analyt Technol Biomed Life Sci* **2002**, *768*, 3.

[18] J. Ostergaard, C. Schou, C. Larsen, N. H. Heegaard, *Electrophoresis* **2002**, *23*, 2842.

[19] C. Bertucci, E. Domenici, *Curr Med Chem* **2002**, *9*, 1463.

[20] A. Frostell-Karlsson, A. Remaeus, H. Roos, K. Andersson, P. Borg, M. Hamalainen, R. Karlsson, *J Med Chem* **2000**, *43*, 1986.

[21] R. L. Rich, Y. S. Day, T. A. Morton, D. G. Myszka, *Anal Biochem* **2001**, *296*, 197.

[22] Y. S. Day, D. G. Myszka, *J Pharm Sci* **2003**, *92*, 333.

[23] in *PhysPro Database, Vol. 2004*, Syracuse Research Corporation, **1999-2004**.

[24] T. Hayashi, H. Imai, K. Kuwata, M. Sogami, S. Era, *Clin Chim Acta* **2002**, *316*, 175.

[25] T. Hayashi, K. Suda, H. Imai, S. Era, *J Chromatogr B Analyt Technol Biomed Life Sci* **2002**, *772*, 139.

[26] C. K. Riener, G. Kada, H. J. Gruber, *Anal Bioanal Chem* **2002**, *373*, 266.

[27] F. B. Edwards, R. B. Rombauer, B. J. Campbell, *Biochim Biophys Acta* **1969**, *194*, 234.

[28] M. Dockal, D. C. Carter, F. Ruker, *J Biol Chem* **2000**, *275*, 3042.

[29] B. Johnsson, S. Lofas, G. Lindquist, A. Edstrom, R. M. Muller Hillgren, A. Hansson, *J Mol Recognit* **1995**, *8*, 125.

[30] Y. A. Gryzunov, A. Arroyo, J. L. Vigne, Q. Zhao, V. A. Tyurin, C. A. Hubel, R. E. Gandley, Y. A. Vladimirov, R. N. Taylor, V. E. Kagan, *Arch Biochem Biophys* **2003**, *413*, 53.

[31] M. L. Connolly, *Science* **1983**, *221*, 709.

[32] J. E. Gestwicki, H. V. Hsieh, J. B. Pitner, *Anal Chem* **2001**, *73*, 5732.

[33] S. Nyberg, H. Nordin, R. Karlsson, M. Hamalainen, *Characterization of drug-plasma protein interactions using surface plasmon resonance*, Biacore AB, Uppsala, **2002**

[34] *Evaluation of small molecule assays in Biacore® 2000 / 3000 (Biacore Knowledge and Training Products Note)*, Biacore AB, Uppsala, **2003**.

[35] *Buffer and sample preparation for direct binding assay in DMSO using Biacore® (Biacore Knowledge and Training Products Note)*, Biacore AB, Uppsala, **2003**.

[36] R. J. Ober, E. S. Ward, *Anal Biochem* **1999**, *271*, 70.

[37] D. G. Myszka, T. A. Morton, *Trends Biochem Sci* **1998**, *23*, 149.

[38] G. W. Mihaly, M. S. Ching, M. B. Klejn, J. Paull, R. A. Smallwood, *Br J Clin Pharmacol* **1987**, *24*, 769.

[39] S. Wanwimolruk, J. R. Denton, *J Pharm Pharmacol* **1992**, *44*, 806.

[40] R. H. McMenamy, *J Biol Chem* **1964**, *239*, 2835.

[41] B. Honore, R. Brodersen, *Anal Biochem* **1988**, *171*, 55.

[42] B. G. Jenkins, R. B. Lauffer, *Mol Pharmacol* **1990**, *37*, 111.

[43] S. Era, T. Hamaguchi, M. Sogami, K. Kuwata, E. Suzuki, K. Miura, K. Kawai, Y. Kitazawa, H. Okabe, A. Noma, et al., *Int J Pept Protein Res* **1988**, *31*, 435.

[44] H. Imai, T. Hayashi, T. Negawa, K. Nakamura, M. Tomida, K. Koda, T. Tajima, Y. Koda, K. Suda, S. Era, *Jpn J Physiol* **2002**, *52*, 135.

[45] M. Anraku, K. Yamasaki, T. Maruyama, U. Kragh-Hansen, M. Otagiri, *Pharm Res* **2001**, *18*, 632.

[46] E. W. Holmes, P. M. Russell, G. J. Kinzler, E. W. Bermes, Jr., *Clin Chim Acta* **1994**, *227*, 1.

[47] J. E. Churchich, *Biochim Biophys Acta* **1972**, *285*, 91.

[48] O. J. Bos, M. J. Fischer, J. Wilting, L. H. Janssen, *Biochim Biophys Acta* **1988**, *953*, 37.

[49] U. Kragh-Hansen, *Biochem J* **1983**, *209*, 135.

[50] A. A. Bhattacharya, S. Curry, N. P. Franks, *J Biol Chem* **2000**, *275*, 38731.

[51] J. Chen, D. S. Hage, *Nat Biotechnol* **2004**, *22*, 1445.

[52] R. Casadio, E. Polverini, P. Mariani, F. Spinozzi, F. Carsughi, A. Fontana, P. Polverino de Laureto, G. Matteucci, C. M. Bergamini, *Eur J Biochem* **1999**, *262*, 672.

[53] A. O. Pedersen, J. Jacobsen, *Eur J Biochem* **1980**, *106*, 291.

[54] S. Paynter, D. A. Russell, *Anal Biochem* **2002**, *309*, 85.

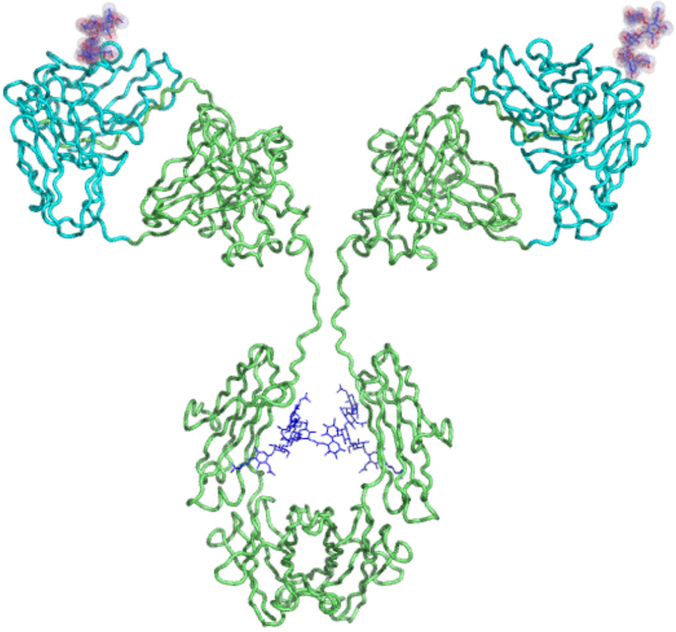
[55] S. Curry, H. Mandelkow, P. Brick, N. Franks, *Nat Struct Biol* **1998**, *5*, 827.

[56] R. Narazaki, T. Maruyama, M. Otagiri, *Biochim Biophys Acta* **1997**, *1338*, 275.

[57] N. P. Sollenne, H. L. Wu, G. E. Means, *Arch Biochem Biophys* **1981**, *207*, 264.

Chapter 4

Diagnostic Antibody GSLA-2



4.1 Introduction

Monoclonal antibodies belong to the first proteins that have been investigated using Biacore technology [1]. Their high specificity and strong affinity not only made them interesting targets for analytical methods but also founded their importance for the diagnosis and treatment of various diseases. Together with the boost in their medical application, the need for analytical tools for the characterization of antibody-antigen interactions increased. In this project, a combination of Biacore and NMR was used to investigate the binding properties and specificity of the diagnostic anti-tumor antibody GSLA-2. Furthermore, the antibody served as a model protein for investigating the influence of factor like immobilization, protein size, or non-specific binding on the assay quality.

4.1.1 The importance of antibodies in diagnosis and therapy

The last few years have seen a growing acceptance and steady increase in the application of therapeutic and diagnostic monoclonal antibodies (mAbs). Following the success of recombinant proteins, mAbs represent the second wave of innovation created by the biotechnology industry and a promise of new versatile therapeutic agents to fight cancer, autoimmune diseases or infection [2-4]. Due to their important role in the immune system and their promising applications, many investigations have been performed to broaden the scientific knowledge about their structure and functions. Nowadays, the 'anatomy' of the antibody molecule is well known (*Fig*. 4-1) [5] and various crystal structures of whole antibodies or their parts are available in structural database [6]. Immunoglobulin G (IgG), which represents the largest class of mammalian antibodies, consists of two heavy and two light chains. Both chains are further divided into highly conserved constant domains (C_H1-3, C_V) and terminal variable domains (V_H, V_L), which are responsible for the antigen-binding (*Fig*. 4-1A). Several β-sheets and various inter- and intrachain disulfide bridges, especially in the flexible hinge region, stabilize the structure of the antibody. IgG are glycoproteins, which contain a conserved *N*-linked glycosyl chain at C_H2 domain (residue Asn-297). The minimal structure of the attached oligosaccharide of the complex biantennary type is a heptasaccharide. Glycosylation is reported to be important for both the structure and function of the immunoglobulins (*Fig*. 4-1B) [7].

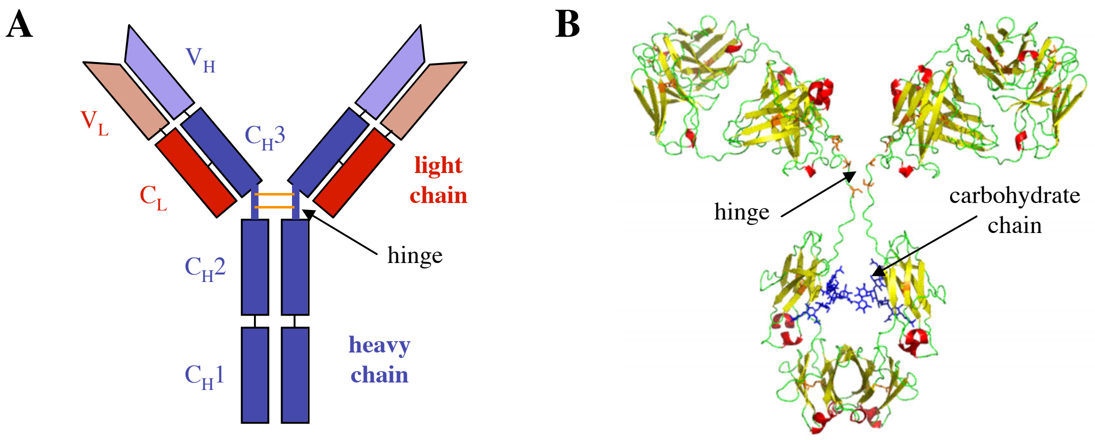


Figure 4-1: Structure of an antibody (IgG1). **A:** Schematic representation showing the heavy chains (blue) and light chains (red) with constant (C_H1-3/C_L) and variable domains (V_H/V_L). **B:** Crystal structure of a human IgG1 (helices are colored in red, sheets in yellow). Cysteine residues of the hinge region are highlighted in orange, the carbohydrate chain in blue. The image was created from the file IgG1.pdf (available at http://www.path.cam.ac.uk/~mrc7/pdb/index.html) by Clark [8].

Since antibodies are the natural ‘line of defense’ and specialized in the detection and elimination of pathogens, their use in the diagnosis and treatment of human diseases has been propagated as a promising area. However, the selection and production of antibodies limited their application for a long time. A first boost in the field of clinical antibodies began with the development of the hybridoma technology by Köhler and Milstein [9], which allowed the expression of monoclonal antibodies in mice. While the production was greatly facilitated, the mouse mAbs were found to induce severe immune responses when applied to men. Therefore, further developments led to chimeric mAbs, where the constant region of the antibody was replaced by the human sequence, and to humanized mAbs with only the binding loops originating from mouse (*Fig*. 4-2). As a final step in this development so far, only the epitope-recognizing parts of the antibody (single chain variable fragments; scFv) are recombinantly produced in *E. coli* [4].

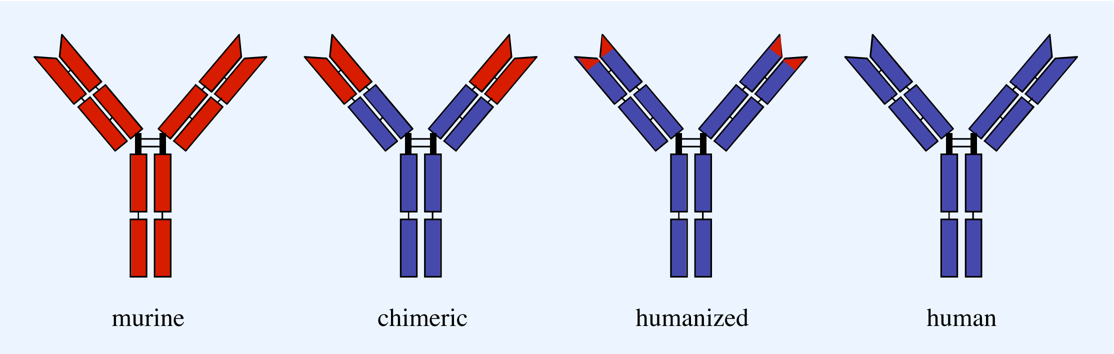


Figure 4-2: Engineering of monoclonal antibodies. The immunogenicity of murine mAbs can be reduced by replacing the constant regions of the mouse sequence (red) by their human counterparts (blue). While chimeric mAbs contain the whole variable domain, only the antigen-binding regions remain murine in the humanized mAbs.

Particularly in the case of life-threatening diseases like cancer or autoimmune disorders, a high specificity and accuracy of the antibody-based diagnosis is an absolute prerequisite. False negative results may prevent or retard an essential treatment of the patient, while false positive diagnoses may burden the organism with unnecessary cures and induce an enormous psychological distress [10]. In addition, therapeutic antibodies with low specificity may bind to healthy tissues and cause severe side effects. Therefore, a profound knowledge about the antibody-antigen interaction on the molecular level helps to develop and select highly specific antibodies. One of the most detailed methods for this purpose is X-ray crystallography, and several hundred antibody-antigen complexes are already available in the PBD database [6, 11]. However, not all molecules can be resolved by crystallography and some important structures such as carbohydrate-antibody complexes were found to be especially challenging [12]. Nuclear magnetic resonance (NMR) evolved into a valuable alternative for the analysis of such structures and was successfully applied to antibody-antigen complexes [13]. While *transfer nuclear Overhauser enhancement* (trNOE) experiments were used to describe the bioactive conformation of carbohydrate epitopes bound to monoclonal antibodies [14], *saturation transfer difference* (STD NMR) [15] was used for mapping the binding epitopes of sugar-based antigens [16].

4.1.2 GSLA-2: a diagnostic monoclonal antibody

In this study, the tumor-diagnostic monoclonal antibody GSLA-2 was investigated. Since GSLA-2 is exclusively used for in-vitro analytical and diagnostic purposes, the immunogenic potential of the murine IgG1 mAb (see *section* 4.1.1) is not relevant.

GSLA-2 recognizes the tumor antigen CA19-9, which was identified as a monosialoganglioside specifically present in patients with colon and pancreas cancer but not in healthy individuals or in patients with other diseases [17, 18]. The molecular binding epitope of this specific recognition was shown to be the tetrasaccharide sialyl Lewis^a (sLe^a), which is a major determinant of pancreatic and gastrointestinal cancer [19, 20]. The carbohydrate epitope sLe^a, as well as the closely related tetrasaccharide sialyl Lewis^x (sLe^x) are both composed of a 5-acetylneuraminic acid (Neu5Ac), a galactose, a *N*-acetylglucosamine (Glc*N*Ac), and a fucose moiety (*Fig*. 4-3A). The only difference between the two molecules is the branching of Glc*N*Ac, where the connections of the 3 and 4 position are reversed (*Fig*. 4-3B) [21].

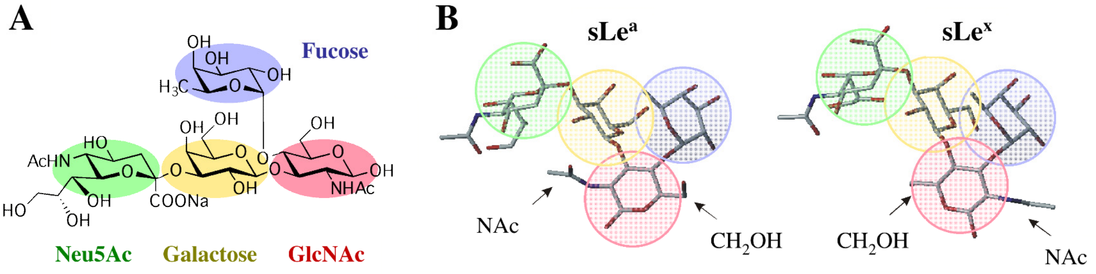


Figure 4-3: Structure of the sialyl Lewis^a (sLe^a) epitope. **A:** Representation of the tetrasaccharide sLe^a with the individual sugar moieties highlighted in different colors. **B:** Three-dimensional structure of sLe^a in comparison with the related sialyl Lewis^x (sLe^x).

While most of the monoclonal antibodies, including GSLA-2, distinguish between the two tetrasaccharide epitopes, there are rare exceptions such as the antibody HECA 452 [22]. E-selectin, an endothelial receptor that is involved in leukocyte trafficking and tumor metastasis, has also been reported to interact with both Lewis antigens [23]. Adhesion of cancer cells to E-selectin can be mediated by sLe^a-carrying glycoproteins (e.g. mucins) or glycolipids (i.e. gangliosides). While sLe^a was found to be mainly responsible for adhesion of human colon, pancreas, and gastric cancer cells, the interaction with lung, liver, and ovarian tumor cells was mediated by sLe^x [24].

The high specificity for tumors of the colon and pancreas made GSLA-2 an useful tool for the diagnosis for these cancers. For example, the antibody was successfully used in diagnostic ELISA against mucin and conjugated sLe^a [25, 26] and for western blot detection of mucins [27, 28]. However, little is known about the exact binding mode and the specificity towards sLe^a derivatives. We, therefore, investigated the interaction between GSLA-2 and a set of carbohydrates on the molecular level.

4.1.3 Preparing and using antibodies for Biacore analysis

Due to affinities in the nanomolar range, the enormous stability and the large mass increase upon binding, interactions between antibodies and protein antigens served as ideal models for the development and proof-of-concept studies of Biacore biosensors [1]. Together with other protein-protein systems, this type of interaction remained the major application of SPR analysis for many years [29]. Only the increase in the instrument’s sensitivity as well as the development of enhanced methods for data acquisition and evaluation allowed the application of the technology for the analysis of small molecule interactions (see *chapters* 2 and 7). This development also offered novel applications for the detection of small haptens by immobilized antibodies [30, 31]. In contrast to the detection of protein antigens, the small SPR signal intensities of low molecular weight compound do not tolerate artifacts from non-specific binding and surface heterogeneity. One way of solving this problem is a reduction of the antibody size by removing the constant domains. Different enzymes are known, which are able to cleave antibodies around their flexible hinge region. Papain cleaves above the disulfide-rich area resulting in an intact constant fragment (Fc) and two monomeric antibody-binding fragments (Fab’). In contrast, pepsin and ficin cleave below the disulfides and produce a dimeric F(ab’)$_2$ fragment and two cut, non-functional Fc parts (*Fig*. 4-4).

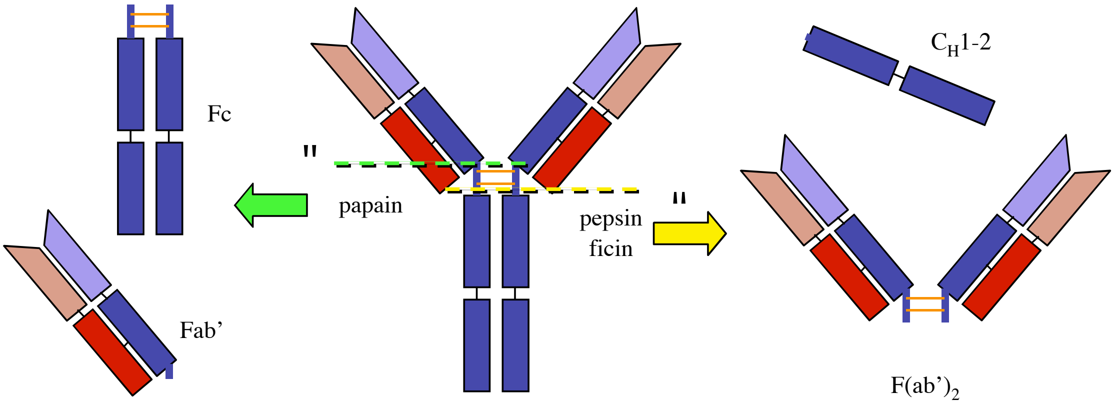
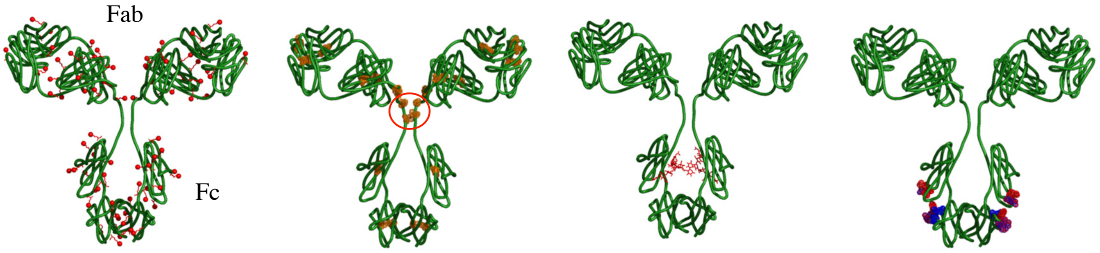


Figure 4-4: Enzymatic cleavage of an antibody and produced, functional fragments. Papain digestion creates constant (Fc) and antigen-binding fragments (Fab’), while cleavage by pepsin or ficin generate F(ab’)$_2$ dimers (as well as non-functional halves of Fc).

In addition to non-specific binding, the immobilization-related inactivation of a significant portion of the antibodies is a major problem for the detection of small molecules. Therefore, different methods have been developed for the oriented immobilization of monoclonal antibodies using either capturing or site-directed coupling techniques [32-34]. Antibody molecules possess different distinct sites for a possible attachment to the surface, such as amino groups of lysines, cysteines of the hinge region (after mild reduction), the carbohydrate chain, or protein A binding sites (*Fig*. 4-5).



amine groups **thiol groups** **carbohydrate chain** **protein A binding site**

Figure 4-5: Attachment sites for the immobilization of monoclonal antibodies (green). While the amine groups of lysine residues (red spheres) are widely distributed, cysteine residues are usually less frequent (the hinge region is highlighted by a red circle). Most mAbs only possess a single glycosylation site per heavy chain (red lines), and a binding site for staphylococcal protein A and streptococcal protein G (red/blue areas on the Fc part).

Alternatively to the enzymatic or biochemical approaches, the generation of smaller antibody fragments like scFv (see *section* 4.1.1) or the introduction of specific attachment sites can also be performed by recombinant expression of engineered antibodies [35, 36].

4.1.4 Aims in this project

The primary aim in this project is to get a deeper insight into the binding specificity of the diagnostically relevant monoclonal antibody GSLA-2 (mouse IgG1), which recognizes the sialyl Lewis^a (sLe^a) epitope of some colon and pancreas cancer cells. For this purpose, a Biacore assay based on covalently immobilized GSLA-2 is developed and validated with sLe^a. A set of structurally related compounds is then screened and the binding data are compared with results from STD NMR experiments. Since mouse IgG1 molecules are the most important subclass of monoclonal antibodies, GSLA-2 is further used as a model protein for the investigation of additional experimental parameters, such as surface activity, protein size, immobilization density, or non-specific binding.

4.2 Materials and Methods

4.2.1 Materials

This section describes materials, equipment and procedures specifically used for the GSLA-2 project. Materials and general methods used in all Biacore assays are described in *section* 2.2.

Reagents and proteins

Murine monoclonal antibody GSLA-2 (IgG1, ~ 4 mg/ml in PBS 7.4, 0.005% sodium azide) was a kind gift from Dr. John L. Magnani (GlycoTech, Gaithersburg, USA). Staphylococcal protein A (SpA; soluble, essentially salt-free, Sigma P6031) and streptococcal protein G (SpG; recombinant, expressed in E. coli, Sigma P4689) were from Sigma (Fluka AG, Buchs, Switzerland). A murine monoclonal IgG1 directed against sheep myoglobin (from the *'Biacore 3000 getting started'* kit) as well as the polyclonal rabbit anti-mouse IgG1 antibody was from Biacore (Biacore AB, Uppsala, Sweden). Sialyl Lewis^a attached to a Lemieux spacer (sLe^a-Lem) was a gift from Prof. Dr. Ole Hindsgaul (Department of Chemistry, Carlsberg Laboratory, Copenhagen, Danmark) and Dr. Oliver Schwardt (Institute of Molecular Pharmacy, University of Basel, Switzerland). Label-free sialyl Lewis^a (sLe^a) was purchased from Sigma (Fluka AG, Buchs, Switzerland). The sources of the compounds tested in the ranking experiments as well as reagent required for NMR experiments are specified in Herfurth *et al.* [37]. Ficin from fig tree latex (lyophilized, Sigma F6008), 2-mercaptoethylamine hydrochloride, (+)-biotinamidohexanoic acid hydrazide (BHZ; Sigma B3770) and *N*-ethylmaleimide were from Sigma (Fluka AG, Buchs, Switzerland), and (+)-biotin-PEO$_3$-maleimide was from Molecular BioSciences, Boulder, USA.

Equipment

HPLC separations were done on an Agilent 1100 purification system, equipped with a quaternary pump, a cooled well-plate autosampler, a column thermostat, a DAD detector, and a cooled analytical fraction collector (Agilent AG, Basel, Switzerland). Size exclusion chromatography was performed on a TSK Gel G2000SW column (7.5 × 600 mm; Tosoh Biosciences GmbH, Stuttgart, Germany). Hi-Trap 5 ml desalting columns (Amersham Biosciences, Otelfingen, Switzerland) were used for buffer exchange.

4.2.2 Immobilization of GSLA-2

GSLA-2 was immobilized using standard amine coupling (see *section* 2.2.5). For this purpose, the surface of a CM5 sensor chip was activated for 7 min. GSLA-2 was diluted in 10 mM sodium acetate buffer *pH* 5.0 to a final concentration of 50 μg/ml and injected for 5-10 min, dependent on the required surface density. After deactivating the surface for another 7 min, non-covalently bound antibody was removed by a short pulse (30 s) of 0.5% SDS. Typically, surface densities between 4 and 14 kRU were obtained by this method.

4.2.3 High-resolution screening of sialyl Lewis^a

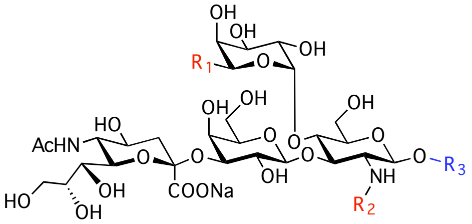
A stock concentration of sLe^a was prepared by dissolving the compound in water to a concentration of 20 mM. This solution was further diluted in running buffer (10 mM HEPES *pH* 7.4, 150 mM NaCl, 3 mM EDTA, 0.005% polysorbate 20; HBS-EP) either as a linear twofold series between 0.2 and 200 μM (11 concentrations) or by combining a twofold series from 200 to 25 μM and a threefold series from 20 to 0.25 μM (9 concentrations). All concentrations were injected as randomized triplicates for 2 min with a dissociation phase of 30 s at a flow rate of 50 μl/min. No regeneration or wash steps were required between the injections. Five buffer blanks were injected before each experiment and one blank between the single series. In order to facilitate the x-axis alignment of the SPR signals and increase the signal quality, buffer spiking was introduced by adding 1% water to the running buffer but not to the sample buffer [38]. An untreated flow cell was used as a reference surface for the referencing process and the signals of the buffer blanks were used for double referencing (see *section* 2.2.6). *Scrubber* was used for data processing and analysis.

4.2.4 Screening and ranking of sLe^a and related compounds

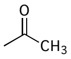
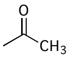
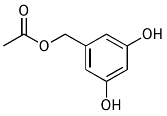
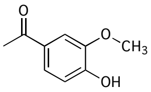
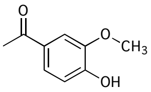
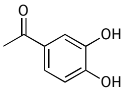
In order to determine the binding specificity of GSLA-2 towards carbohydrate epitopes, sLe^a and a set of structurally related compounds (Fig. 4-6) were investigated by a combination of Biacore and STD NMR.

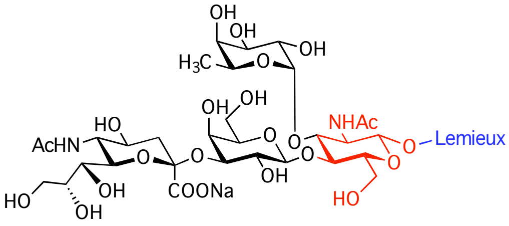
All NMR experiments and the major part of the Biacore ranking (analytes **1b, 2-8**) were performed by Dr. Lars Herfurth (Institute of Chemistry, Medical University of Lübeck, Germany). For the Biacore ranking, concentration series of the compounds (0.2-800 μM) were injected for 4 min as single injections at a flow rate of 10 μl/min in

HBS-P buffer (10 mM HEPES *pH* 7.4, 150 mM NaCl, 3 mM EDTA, 0.005% polysorbate 20). The signals were referenced against an immobilized anti-myoglobin mAb and no double referencing was performed before fitting the steady state responses to a single binding site model. Data processing and analysis was performed in *BIAevaluation [37, 39].*

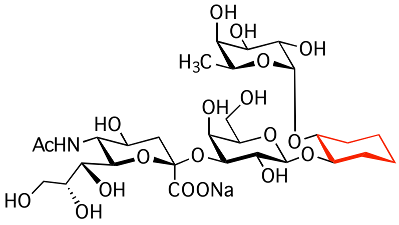


1 (sLe^a) - **5**

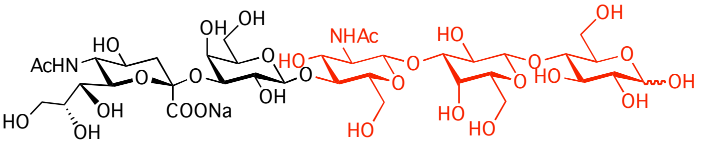
No.	R_1	R_2	R_3
1a	CH_3		H
1b	CH_3		Lemieux
2	H		Lemieux
3	CH_3		Lemieux
4	H		Lemieux
5	CH_3		Lemieux



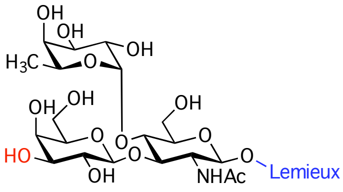
6 (sLe^x)



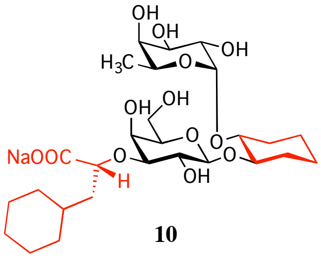
7



8 (LSTa)



9 (Le^a)



10

Lemieux-Spacer =
$(CH_2)_8COOCH_3$

Figure 4-6: Sialyl Lewis^a (sLe^a) and structurally related compounds analyzed by Biacore and STD NMR. Conserved parts of the sLe^a core structure are represented in black, modifications in red, and the attachment of a Lemieux spacer at the reducing end in blue.

In order to complete the data set, analytes **1a**, **9**, and **10** (*Fig*. 4-6) were analyzed using the high-resolution assay described in section 4.2.3. The influence of the Lemieux spacer was further investigated by a comparative analysis of the thermodynamic and kinetic binding properties of analytes **1a** and **1b** using *Scrubber* and *CLAMP*.

4.2.5 Production and evaluation of antibody fragments

60 μl of GSLA-2 was mixed with 30 μl Tris/EDTA buffer (50 mM Tris·HCl *pH* 7.0 2 mM EDTA). 8 μl of a ficin stock solution (1 mg/ml in Tris/EDTA buffer; 1.8 units/ml) and 2 μl cysteine activation solution (6.1 mg/ml in Tris/EDTA buffer) were added and the mixture was incubated for 4 hours at 37°C. $F(ab')_2$ were separated from Fc fragments and whole mAb using size exclusion chromatography (isocratic conditions for 40 min, flow rate 1 ml/min, ambient temperature, running buffer: 50 mM sodium acetate *pH* 5.0, 150 mM sodium chloride). 1 ml fractions were collected and maximum concentrated $F(ab')_2$ fractions were around 0.1 mg/ml (*Fig*. 4-7).

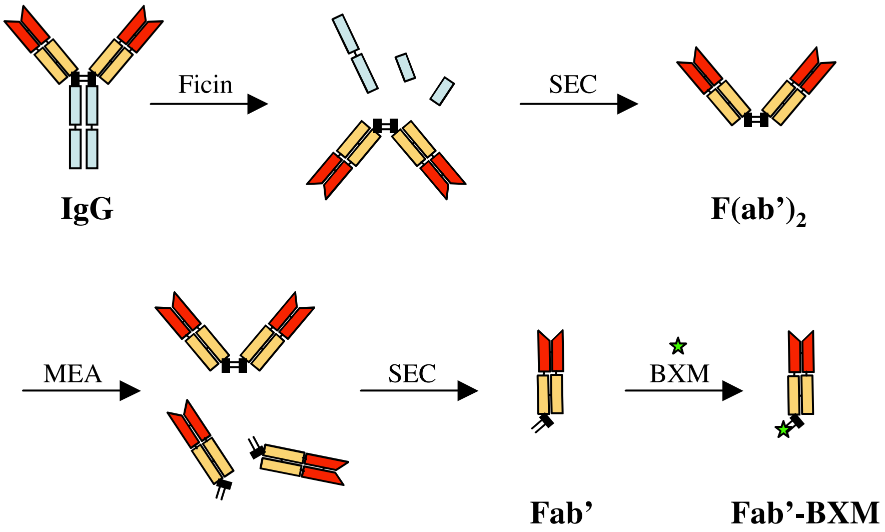


Figure 4-7: Schematic overview of the key steps in the antibody fragment preparation. After the ficin digest, the $F(ab')_2$ fragments were isolated by size-exclusion chromatography (SEC). A part of these fragments were further cleaved by a mild reduction with 2-mercaptoethylamine (MEA). The free cysteine residues were blocked either with biotin-maleimide (BXM) or with *N*-ethylmaleimide (NEM).

Purified $F(ab')_2$ fractions (~ 0.1 mg/ml) were further cleaved by buffer exchange on a HiTrap desalting column to phosphate/EDTA buffer (0.1 M sodium phosphate *pH* 6.0 5 mM EDTA) and addition of a 200 mM MEA solution in phosphate/EDTA buffer to final MEA concentration 5 mM. After incubation for 90 min at 37°C, any free thiol groups at the hinge region were either blocked by adding NEM solution or biotinylated by adding biotin-maleimide solution to a concentration of 20 μM for 2 hours at room temperature. Finally, buffer was exchanged twice to 10 mM sodium phosphate *pH* 5.0 to completely remove any free MEA and thiol active reagents (*Fig*. 4-7).

Purified preparations of whole GSLA-2 as well as of its F(ab')$_2$ and Fab' fragments in 10 mM sodium acetate *pH* 5.0 were immobilized on separate flow cells of a CM5 sensor chip as described in *section* 4.2.2. One flow cell was left untreated as a reference surface. All surfaces were simultaneously evaluated by high-resolution screening with sLe^a-Lem (see section 2.1.3). Surface activity was calculated by dividing the measured R$_{max}$ value by the theoretical R$_{max}$ (*Eq*. 4, see *section* 2.1.3). For comparative plots, all fragment densities were normalized to the weight and valency of the whole antibody using *equation* 9:

$$\text{Normalized Density} = \frac{Valency_{Fragment} \times MW_{GSLA-2}}{MW_{Fragment} \times Valency_{GSLA-2}} \times Density_{Surface} \qquad \text{[Eq. 9]}$$

In-process controls of the enzymatic digest, the mild reduction and the separation steps were performed by SDS-PAGE and dot blot analysis. A modified protocol from Lämmli [40] was used for the preparation of SDS-PAGE (8-12% gels) and silver staining or an improved protocol for Coomassie staining [41] was applied for detection.

4.2.6 Evaluation of non-specific binding

In order to investigate the contribution of different production batches or sources of sLe^a-Lem (**1b**) for non-specific binding, two different batches of the compound were screened according to the high-resolution protocol described in *section* 4.2.3. In addition, the effect of buffer additions (3 mM EDTA, 0.005% polysorbate) was evaluated by screening sLe^a-Lem (**1b**) in various combinations of running and sample buffers (HBS-N, HBS-P, HBS-EP). One sLe^a-Lem (**1b**) batch showing high non-specific binding was diluted to a concentration of 200 μM in each buffer and injected for 2 min at a flow rate of 50 μl/min over GSLA-2 and its fragments (see *section* 4.2.5). The injection series was repeated for all three running buffers.

4.2.7 Comparison of capturing approaches

Immobilization protocols for staphylococcal protein A (SpA) and streptococcal protein G (SpG) were developed based on methods from literature [33, 42, 43]. Stock solutions of SpA and SpG were prepared by dissolving 1 mg lyophilized protein in 1 ml water. While SpA could directly be diluted in 10 mM sodium acetate *pH* 5.0, SpG solution had to be dialyzed extensively against immobilization buffer (10 mM sodium acetate *pH* 4.0) to remove the remaining Tris salt. The final concentration for both SpA and SpG immobilization solutions was 50 μg/ml. Immobilization to a CM5 sensor chip was done using standard amine coupling (see *section* 2.2.5) at a flow rate of 10 μl/min. Activation and deactivation contact times were 10 min, the proteins were injected for 15 min. Both surfaces were evaluated by injecting a fivefold GSLA-2 concentration series between 0.2 and 666 μM for 10 min at flow rate of 10 μl/min with a dissociation phase of 10 min. The SpA surface was then regenerated by a 60 s pulse of 10 mM glycine *pH* 1.5. The binding data were analyzed using *Scrubber* (affinity) and *CLAMP* (kinetics). Finally, GSLA-2 was captured on immobilized SpA for 20 min at a concentration of 0.1 mg/ml and a flow rate of 5 μl/min (mAb density ~ 3,000 RU), and a 200 μM sLe^a-Lem (**1b**) solution was injected for 2 min. HBS-EP buffer was used throughout immobilization, evaluation and capturing experiments.

In an alternative capturing approach, a rabbit anti-mouse IgG1 antibody was immobilized to a CM5 sensor chip using amine coupling. A 30 μg/ml solution of the antibody in 10 mM sodium phosphate buffer *pH* 5.0 was injected for 7 min over the activated surface at a flow rate of 5 μl/min. After deactivation for 7 min and a 2 min wash step with 10 mM glycine *pH* 2.0, the surface showed a density of 12,000 RU. The surface was evaluated by injecting GSLA-2 at concentrations of 3, 0.3, and 0.03 μM for 5 min. Following a dissociation phase of 5 min, GSLA-2 was removed by injecting 10 mM glycine *pH* 2.0 three times for 120 min. For the sLe^a screening, GSLA-2 was captured by injecting the antibody at a concentration of 400 μg/ml for 5 min (resulting in a density of 2,200 RU). sLe^a-Lem (**1b**) was injected between 0.2 and 200 μM as single injections for 1 min and the equilibrium responses were fitted to a single binding site model.

4.3 Results & Discussion

4.3.1 Development of a GSLA-2 Biacore assay

Covalent immobilization of monoclonal antibodies by standard amine coupling is a well-established procedure and often results in high-density surfaces. Even though the coupling procedure may induce a significant loss in activity, the remaining binding sites are usually sufficient for an interpretable signal. Therefore, GSLA-2 was directly immobilized on a CM5 sensor chip using amine coupling, which yielded in high surface densities (≥ 10,000 RU). Injection of a concentration series of sialyl Lewis^a with an attached Lemieux spacer (sLe^a-Lem, **1b**) over the antibody surface generated small (< 100 RU) but clearly detectable, and highly reproducible SPR signals. As expected for carbohydrate-protein interactions, yet unusual for antibody-antigen binding, very fast binding kinetics could be observed (*Fig*. 4-8A). Steady state responses were concentration-dependent and could be fitted to a single binding site model (*Fig*. 4-8B). Despite the fast association and dissociation rates, the complete data set could be fitted kinetically to a simple Langmuir 1:1 binding model (*Fig*. 4-8C).

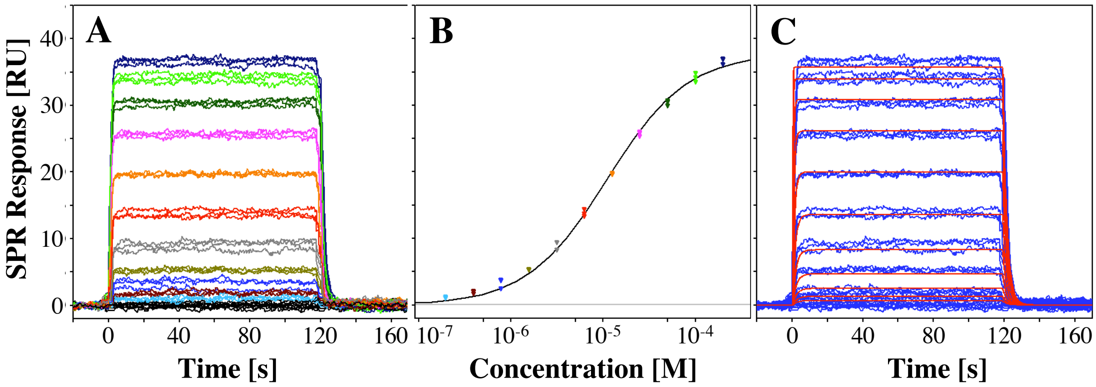


Figure 4-8: High-resolution GSLA-2 binding assay. **A:** Overlaid sensorgrams of a typical sLe^a-Lem (**1b**) concentration series (0.2-200 μM), injected as randomized triplicates. **B:** Steady state responses of the same data set, fitted to a single binding site model. **C:** Kinetic evaluation of the data set using *CLAMP*. The sLe^a signals (blue) are overlaid with the simulated kinetic curves (red).

Due to the fast kinetics of the interaction and the rapid recurrence to the baseline, no regeneration or wash steps were required. The steady state phase was reached within a few seconds and was found to be very stable over several minutes. Fitting the signals from the high-resolution assay to a single binding site model led to K_D values in the range of 8 to 15 μM. Again, such K_D's are rather high for typical antibody-binding events, but expected for the interaction of monovalent sugars with their protein targets.

In order to detect a dependency of the binding affinity from the surface density, GSLA-2 was immobilized at three different levels on the same sensor chip. After screening sLe^a-Lem (**1b**) simultaneously over all three surfaces, the binding affinities and the surface activity was evaluated and correlated with the antibody density (*Fig*. 4-9, *Table* 4-1).

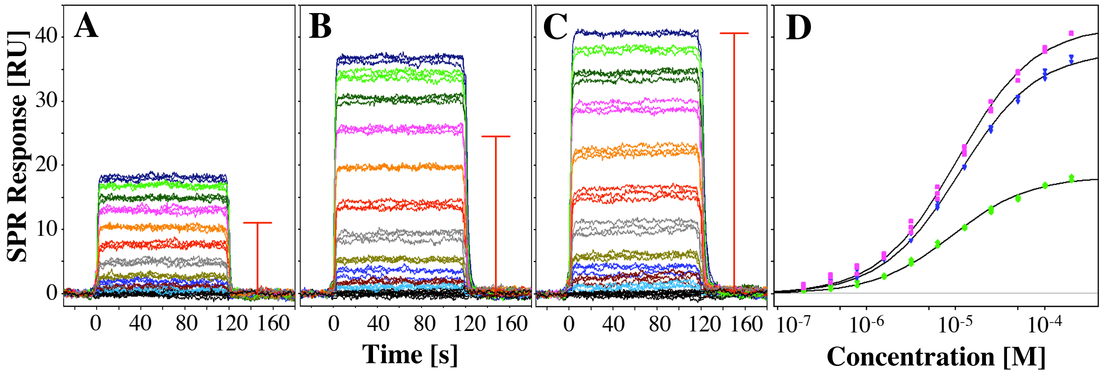


Figure 4-9: Influence of the GSLA-2 surface density on the binding activity. **A-C:** Sensorgrams of a sLe^a-Lem series (0.2-200 μM) on a low-density (**A**; 3,600 RU), a medium-density (**B**; 8,400 RU) and a high-density surface (**C**; 13,600 RU). Red bars indicate the expected maximum response when normalized to the high-density surface by assuming a constant surface activity. **D**: Equilibrium binding plots of the low (green), medium (blue), and high-density surface (magenta), fitted to a single binding site model.

Table 4-1: Evaluation of the surface activity and binding affinity of sLe^a-Lem (**1b**; 0.2-200 μM, triplicate injections) at three different GSLA-2 surface densities.

Surface	**Density [RU]**	**Exp. K_D [i] [μM]**	**Exp. R_{max} [i] [RU]**	**Calc. R_{max} [ii] [RU]**	**Activity [iii]**
high-density	13,600	10.3	41.8	179.7	23%
medium-density	8,400	11.0	37.8	111.0	34%
low-density	3,600	9.2	18.3	47.6	39%

[i] Experimental values were obtained by fitting the steady state responses to a single binding site model. [ii] Calculated using *equation* 4 (see section 2.1.3) with valency = 2, MW_{Target} = 150 kDa, $MW_{Analyte}$ = 991 Da. [iii] Calculated by division of the experimental with the calculated R_{max} value.

As expected, higher immobilization densities of GSLA-2 using amine coupling led to increased signal intensity. However, when calculating the surface activity, a significant decrease was visible with increasing surface densities (*Fig*. 4-9, *Table* 4-1). A similar effect has been reported recently by Huber *et al*. [44], who observed a negative trend in surface activity with increasing density for the cyclophilin D system. Steric hindrance and a reduced accessibility of the binding sites had been discussed as the most plausible explanation for this effect [44]. However, the reduced surface activity had no significant influence on the binding affinity of sLe^a-Lem (**1b**) to the remaining GSLA-2 sites, since the K_D value remained within the experimental error (Table 4-1). Therefore, surface densities between 10-15 kRU are regarded as a good compromise between surface activity and signal intensity.

4.3.2 Screening of sialyl Lewis^a and derivatives

The ability to distinguish between sLe^a and sLe^x or to recognize both equally well plays an important role in many biological processes. A profound knowledge about the binding specificity and the binding mode of this kind of interactions is therefore a prerequisite for the development of specific ligands, e.g. in the case of E-selectin. GSLA-2, which is able to specifically recognize sLe^a, was used as a model antibody by screening a set of sLe^a derivatives and related compounds (*Fig*. 4-6). In a first step, all analytes were injected over immobilized GSLA-2 and the K_D values were determined by fitting the equilibrium responses to a single binding site model (*Fig*. 4-10A) [37].

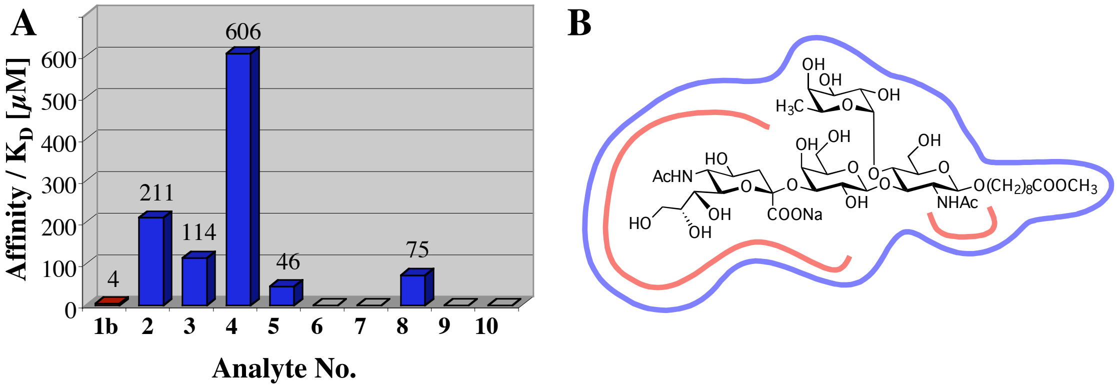


Figure 4-10: Ranking of sLe^a and some derivatives. **A:** Comparison of binding affinities of the Biacore screening assay. The corresponding analyte structures are visualized in *figure* 4-6 (see *section* 4.2.4). **B:** Binding hypothesis developed from the screening results and STD NMR data.

The Biacore screening clearly identified the sialyl group as an absolute requirement for a detectable binding event. Replacement of this group by cyclohexyl-lactic acid moiety (analyte **10**) or a complete removal (Le^a, **9**) led to a complete loss in activity. In addition, the *N*-acetyl group of the Glc*N*Ac moiety seems to be of similar importance, since any modification of this group led to a significant decrease of the binding affinity (**2-5**). This effect is also visible in the case of sLe^x (**6**), where the position of fucose and galactose are changed, leading to a radically reduced affinity. In addition, the shifted position of the Lemieux-spacer might also contribute to this effect. Likewise, no binding could be detected when replacing the whole Glc*N*Ac by a simple cyclohexane moiety (**7** and **10**). Interestingly, the linear pentasaccharide LSTa (**8**) still showed a small affinity to GSLA-2, despite the lack of fucose. The higher flexibility of LSTa compared to sLe^x due to the absence of a branched fucose seems to allow a positioning of the *N*-acetyl group that is similar to sLe^a. Finally, the removal of the methyl group of the fucose moiety also had a negative influence on the binding affinity, as it is illustrated by the comparison of analytes **3** and **4**. These findings led to the binding hypothesis, where the *N*-acetyl group of Glc*N*Ac and the Neu5Ac moiety (sialyl) are essential for a high-affinity interaction with GSLA-2 [37] (*Fig*. 4-10B, red areas).

STD NMR experiments [37] offered an even deeper insight into the binding mode of sLe^a and its derivatives. It could be shown that GSLA-2 recognizes the whole tetrasaccharide epitope, since all protons of the four pyranose rings received saturation (*Fig*. 4-10B, blue area). The importance of the two essential areas identified during Biacore screening was confirmed by the NMR experiments. The *N*-acetyl groups from both the Neu5Ac and the Glc*N*Ac moiety received a large fraction of the saturation transfer. In contrast, only weak contacts were found for the Lemieux spacer. When investigating the binding of sLe^x (**6**), it was observed that the antibody mainly recognized the sialyl group, while no signal could be detected anymore for Glc*N*Ac. A major difference to E- or P-selectin might be the fact that their binding site is rather shallow [45] and interacts predominantly with the galactose and fucose moieties [46], whereas the binding site of GSLA-2 seems to be deeper and therefore less tolerant toward structural modifications of the tetrasaccharides [37].

Even though the Lemieux spacer [47] was found to interact only slightly with GSLA-2 by NMR experiments (see above), it might nevertheless influence the binding affinity during Biacore experiments. Therefore, sialyl Lewis^a in presence and absence of a Lemieux-spacer (**1a** and **1b**, *Fig*. 4-6) were injected at the same concentration range (*Fig*. 4-11A&B). Both their equilibrium binding responses (*Fig*. 4-11B) and their kinetic properties were compared and analyzed (*Table* 4-2).

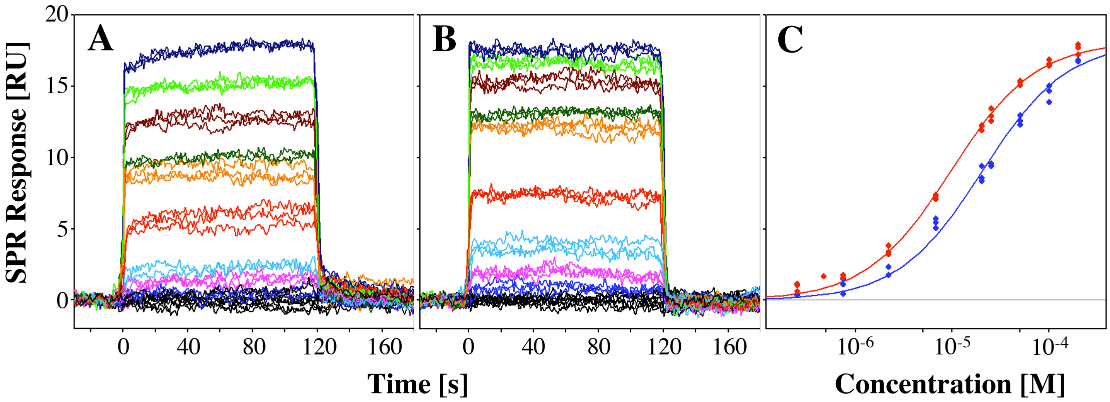


Figure 4-11: Influence of the Lemieux spacer on sLe^a on its binding to GSLA-2. The pure tetrasaccharide (**A**, analyte 1a) and sLe^a with spacer (**B**, analyte **1b**) are injected over the same concentration range (0.25-200 μM) as randomized triplicates. **C:** Overlay plot of the steady state binding responses of sLe^a with (red) and without (blue) Lemieux spacer.

Table 4-2: Kinetic evaluation of sialyl Lewis^a as an unlabeled tetrasaccharide (sLe^a, **1a**) and with a Lemieux spacer at the reducing end (sLe^a-Lem, **1b**). All data are average values from three independent experiments.

Analyte	**k_{on} [10^4 $M^{-1}s^{-1}$]**	**k_{off} [s^{-1}]**	**$K_{D\ kin}$ [μM]** [i]	**$K_{D\ eq}$ [μM]** [ii]	**$t_{1/2}$ [s]**[iii]
sLe^a	2.26 ± 0.50	0.51 ± 0.08	22.6	23.1 ± 2.8	1.4
sLe^a-Lem	5.94 ± 2.26	0.61 ± 0.17	10.3	10.6 ± 1.2	1.1

[i] Calculated using $K_D = k_{off}/k_{on}$. [ii] Derived from fitting the steady state responses to a single binding site model. [iii] Calculated using $t_{1/2} = \ln 2/k_{off}$.

The Lemieux derivative showed a significantly higher affinity for GSLA-2, most probably due to an increased hydrophobicity. Interestingly, the on-rate almost exclusively contributes to the shift in affinity, while the off-rate remains constant. This might be explained by an entropic effect, where the hydrophobic spacer group is pushed out of the hydrophilic solvent into the binding pocket, leading to a faster association rate. The slightly higher K_D of sLe^a-Lem (10 μM) from the high-resolution experiments used for this kinetic evaluation compared to the value from the ranking experiment (4 μM) might be caused by a different data processing. These findings clearly illustrate that spacer groups usually used for purification of oligosaccharides using reversed-phase chromatography may influence the binding behavior of the labeled compound.

4.3.3 Production and evaluation of antibody fragments

Antibodies offer an attractive opportunity to specifically cleave the molecule in order to generate smaller fragments with preserved binding activity. Since the SPR signal is dependent on the ratio between the molecular weights of the target and the analyte, binding to smaller fragments often generate higher signal intensities. In principle, two enzymes are predominantly used for this purpose: papain, which generates monovalent Fab' fragments and pepsin for the preparation of dimeric $F(ab')_2$ fragments (*Fig*. 4-4, see *section* 4.1.3). The pepsin-type digestion leaves the hinge region intact and the $F(ab')_2$ can be further cleaved by mild reduction of the hinge disulfides. The resulting free thiol groups can then be functionalized by thiol-active reagents and used for a site-directed coupling approach [32, 48]. However, pepsinolysis of intact antibodies has found to be difficult, and could only be improved by adding a deglycosylation step prior to the digestion. The deglycosylation is rather time-consuming (24-48 h) and involves different glycosidases. Therefore, a method was developed to cleave GSLA-2 without the need for deglycosylation and to biotinylate the reduced fragments [49]. In this study, the cysteine protease ficin (EC 3.4.22.3) was identified as a valuable alternative to pepsin, since it readily cleaves the intact mAb with high specificity and could also be triggered to directly generate Fab' fragments by increasing the activator concentration (cysteine). Furthermore, an optimal concentration of mercaptoethylamine as a mild reducing agent could be determined for the mild reduction of GSLA-2 $F(ab')_2$ [49]. These results were then combined for an efficient generation of GSLA-2 fragments. Size exclusion chromatography was found to be a suitable method for monitoring the ficinolysis and to isolate the $F(ab')_2$ (*Fig*. 4-12).

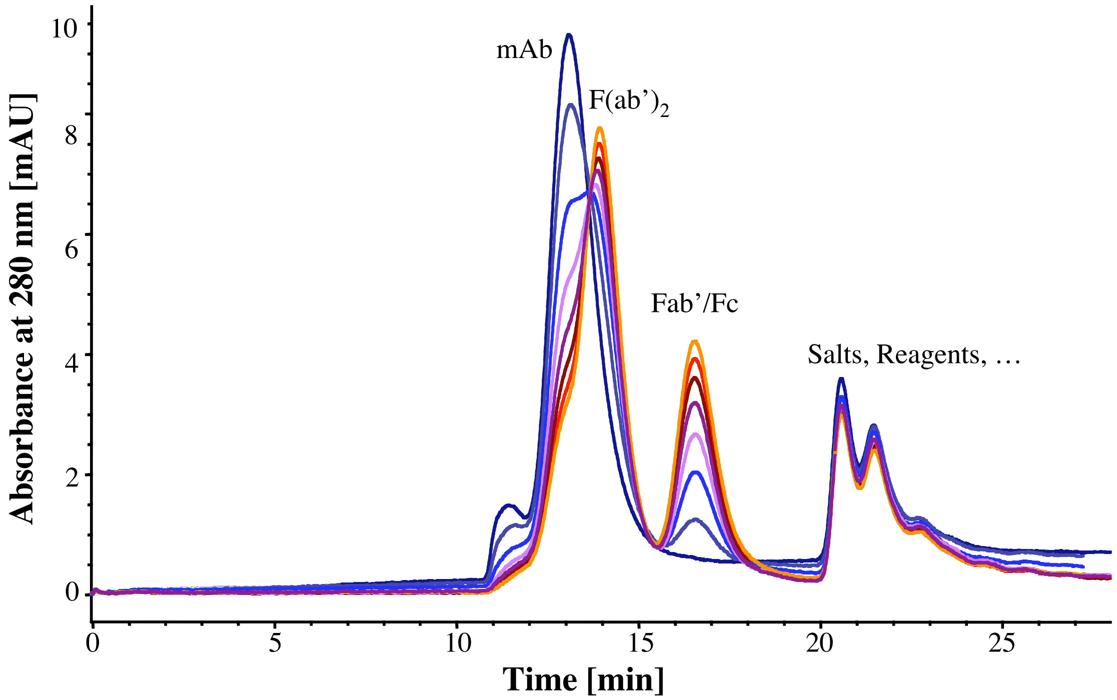


Figure 4-12: Monitoring of the ficin digestion by size exclusion chromatography from start (dark blue) to 4 hours of incubation at 37°C (orange) in 30 min steps. While the first two peaks could be identified as the intact antibody and the F(ab')$_2$ fragment, the third peak developed over the incubation time and are most likely Fc or Fab' fragments. The last peak group did not change significantly over time and consists of buffer salts and reagents (ficin, cysteine, etc.).

By incubating the isolated F(ab')$_2$ fragments with mercaptoethylamine (MEA), the hinge region was mildly reduced. The free thiols were either blocked or functionalized with a thiol-active biotinylation reagent (*Fig.* 4-13A). Dot-blot analysis of the biotinylated Fab' with labeled extravidin (a derivative of avidin with reduced non-specific binding) confirmed the successful addition of a biotin group to the protein (*Fig.* 4-13B). Unfortunately, the biotinylated fragment could not be captured on a custom-made neutravidin sensor chip. Neutravidin (a deglycosylated form of streptavidin) was used instead of streptavidin because of its lower non-specific binding [50]. It is not yet clear, whether the lack of capturing efficiency was caused by the rather low concentration of the Fab'-biotin construct, by remaining impurities from the biotinylation reagent, or by inactive neutravidin. Therefore, the experiment has to be repeated on a commercially available streptavidin sensor chip.

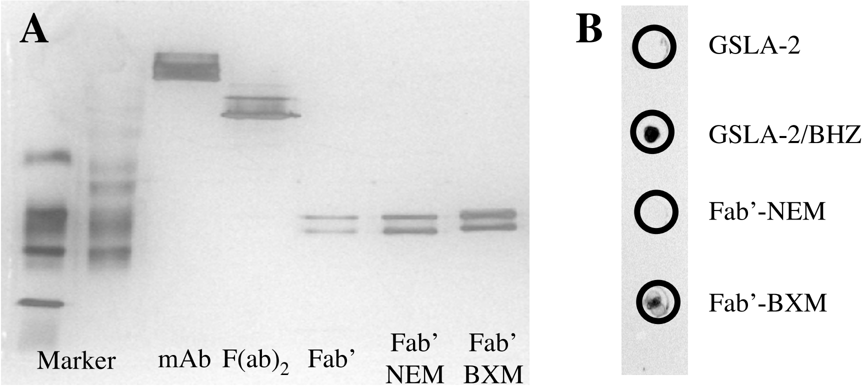


Figure 4-13: Separation and biotinylation of GSLA-2 fragments. **A:** SDS-PAGE analysis (10%, non-reducing, silver staining) of the untreated mAb, F(ab')$_2$ and Fab' fragments as well as of Fab' functionalized with either *N*-ethylmaleimide (NEM) or biotin-spacer-maleimide (BXM). **B:** Dot blot analysis using peroxidase-labeled extravidin. The functionalized Fab' were compared with intact GSLA-2 (untreated or biotinylated with biotin-hydrazine (BHZ).

In order to see whether the reduced protein size of the GSLA-2 fragments has a beneficial effect on the signal quality of amine-coupled surfaces, the whole mAb as well as its F(ab')$_2$ and Fab' fragments were immobilized on separate flow cells. Compared to the intact mAb, both fragments could only be immobilized at rather low densities even when considering the reduced molecular weight. Again, the most probable explanation of this effect is the lower protein concentration after the various separation and desalting steps. Another reason might be the reduction of surface-accessible lysine residues by removing the Fc part, which results in lower surface attraction and reduced coupling efficiency. Nevertheless, all surfaces showed sufficient signal intensity for the screening of sLe^a-Lem (**1b**, *Fig*. 4-14A-C) and the resulting equilibrium responses fitted to a single site model in all cases (*Fig*. 4-14D).

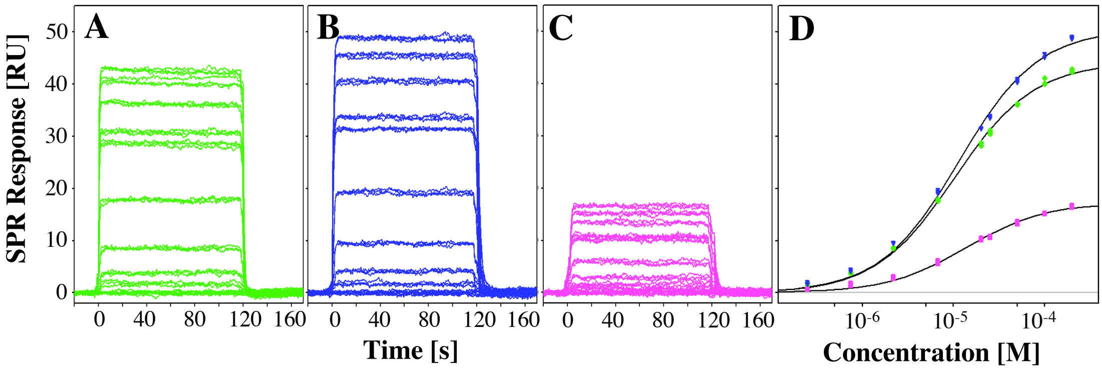


Figure 4-14: Screening of sLe^a (0.25-200 μM) on immobilized GSLA-2 (**A**, green) and its F(ab')$_2$ (**B**, blue) and Fab' fragments (**C**, magenta). An overlay of the corresponding steady state plot is shown in **D**. All data sets were fitted to a single binding site model.

The K_D values obtained by fitting the steady state responses from the different GSLA-2 surfaces were all in a close range (*Table* 4-3), indicating that the fragment production process (ficinolysis, reduction, separation, desalting) did not significantly alter the activity of the antibody binding site. Whether the slight trend to higher K_D values with reduced fragment size is really caused by a change in binding affinity or other effects like surface packing or the experimental error are involved, could only be decided after additional experiments. However, the slightly decreased affinity of the monomeric Fab' fragment compared to the bivalent $F(ab')_2$ and mAb molecules could be based in the reduced local concentration of carbohydrate binding sites.

Table 4-3: Evaluation of the surface activity and binding affinity of sLe^a-Lem (**1b**, 0.2-200 μM, triplicate injections) on three different GSLA-2 surfaces featuring the whole mAb, the $F(ab')_2$, or the Fab' fragment.

Fragment	MW [kDa]	Valency	Density [RU]	Exp. K_D [a] [μM]	Exp. R_{max} [a] [RU]	Calc. R_{max} [b] [RU]	Activity [c]
mAb	150	2	12,750	9.8	43.5	168.5	26%
$F(ab')_2$	100	2	5,530	11.5	50.5	109.6	46%
Fab'	50	1	1,770	13.9	17.4	35.1	49%

[a] Experimental values were obtained by fitting the steady state responses to a single binding site model. [b] Calculated using *equation* 4 (see section 2.1.3) with $MW_{Analyte}$ = 991 Da. [c] Calculated by division of the experimental with the calculated R_{max} value.

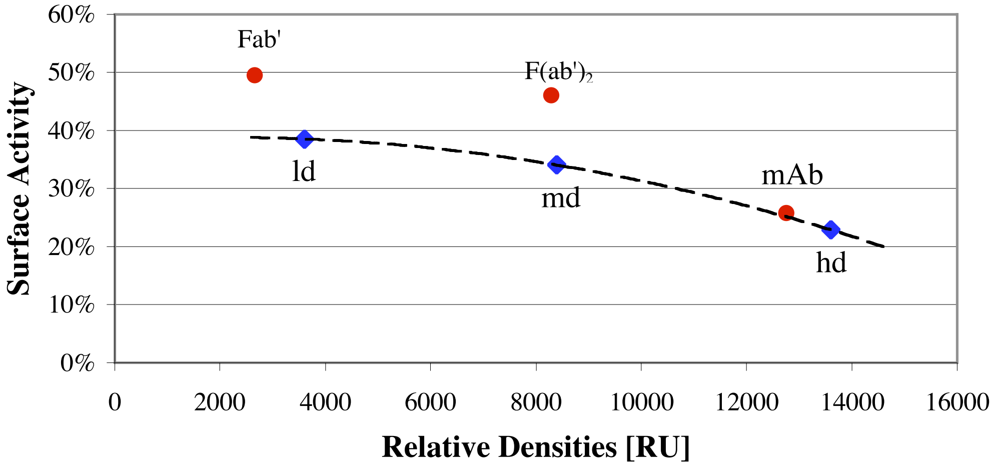


Figure 4-15: Comparison of surface density (blue) and fragment size (red) effects on the surface activity. Calculated activity values for high-density (hd), medium-density (md) and low-density (ld) GSLA-2 surface are overlaid by the values from surfaces of immobilized GSLA-2 fragments (mAb, $F(ab')_2$, Fab'). The relative surface densities were calculated by *equation* 9 (see *section* 4.2.5). The expected trend in the density-activity correlation is represented by a dashed line.

By comparing the experimental and calculated R_{max} values, a remarkable increase in the surface activity could be observed for the two smaller fragments of GSLA-2. While this effect could be partly induced by the lower surface density (see *section* 4.3.1, *Table* 4-1), it does not explain the entire activity improvement. This becomes evident, when the activities of the fragment experiment are directly compared with those from the surface density experiment (*Fig.* 4-15). It clearly illustrates that the improved surface activity caused by the cleavage of GSLA-2 to its $F(ab')_2$ and Fab' fragments was not only caused by a reduced density effects. While the surface activity for the whole antibody lies exactly on the trend line, both fragment are shifted towards higher surface activity. This effect might be caused by a better accessibility of the binding sites or by an immobilization closer to the surface, which induces a higher SPR signal. Hence, truncated mAb preparations seem to have an advantage over the intact molecule and should be always considered for Biacore experiments. On the other side, the fragment preparation procedure is still very time-consuming and usually results in decreased target concentrations and lower surface densities. If the fragments are not functionalized and are immobilized by standard amine coupling, the direct generation of Fab' fragments by papain might be advantageous since it requires less preparation steps.

4.3.4 Evaluation of non-specific binding

During the high-resolution experiments, some sLe^a-Lem (**1b**) samples showed a clearly two-phased binding signal (*Fig.* 4-16A). After a first rapid increase, a slower association phase could be detected. As soon as the injection was stopped, the SPR signal dropped rapidly with a subsequent slower dissociation period. The intensity of the rapid phases corresponds with the block signals usually observed for the binding of sLe^a. This suggested two independent binding events, one of which is believed to occur non-specifically. One way of reducing non-specific binding is the immobilization of a similar but inactive protein to the reference cell [30, 51]. Therefore, an anti-myoglobin antibody of the same immunoglobulin class (mouse IgG1) was immobilized at a similar density and used as a reference cell. Large non-specific signal effects as shown in *figure* 4-16A could not be eliminated (see *appendix* C1), but the use of an antibody reference often enhanced the signal quality in case of smaller signal impurities (*Fig.* 4-16B&C). While the relative binding intensities and therefore the K_D value remained nearly constant, the reproducibility and shape of the signal were improved.

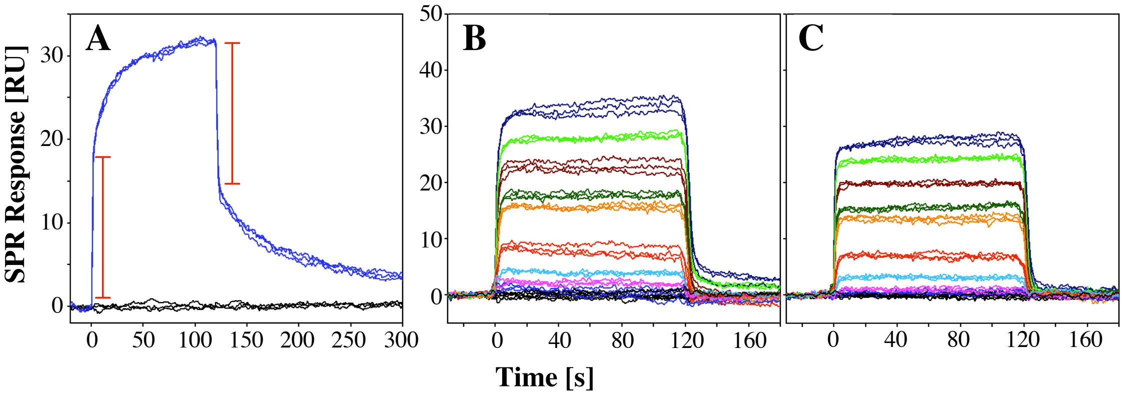


Figure 4-16: Non-specific binding of some sLe^a injections. **A:** Two-phased sensorgram of a sLe^a injection at 20 μM (triplicate). The estimated intensity of the specific binding event is indicated by red bars. **B, C:** Overlay plot of the same concentration series (0.25-200 μM) referenced against a blank sensor surface (**B**) or an anti-myoglobin antibody (**C**).

Not all samples induced the same non-specific effect and not all surfaces seemed to be equally susceptible to non-specific binding. Therefore, two sLe^a-Lem (**1b**) batches from distinctive sources were screened over surfaces at three different densities (*Fig.* 4-17). Depending on its source, the samples induced a lower or higher degree of non-specific binding on the same surface (*Fig.* 4-17AB). Interestingly, NMR analysis could detect no impurity in the specific sample, which showed the most prominent signal effect. The signal-inducing component could therefore not be identified so far, and a deeper investigation of analytical or purification methods is required. In addition, the non-specific binding effect was larger with increasing surface density (*Fig.* 4-17C&D).

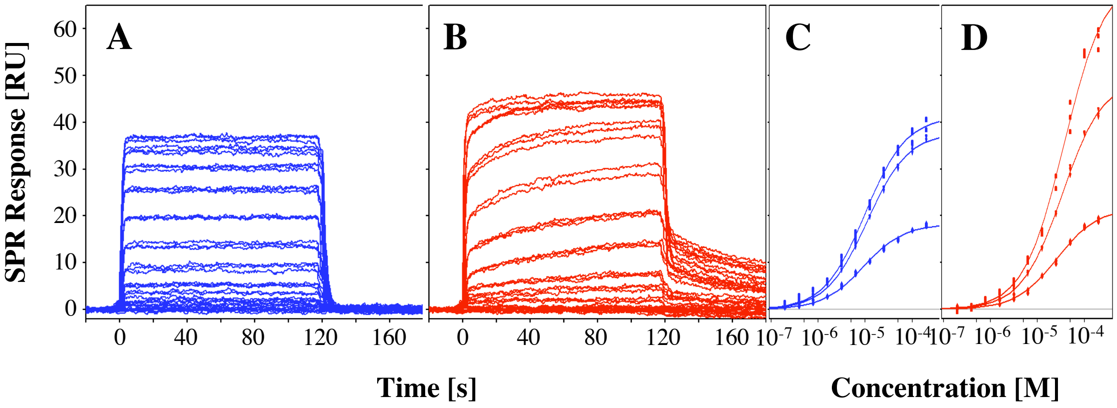


Figure 4-17: Non-specific binding of different sLe^a-Lem (**1b**) batches. **A, B:** Sensorgrams of two batches from different sources, screened (200 nM-200 μM) over a medium-density GSLA-2 surface (8,400 RU; amine coupling). **C, D:** Steady state plots of the same batches on a low (3,600 RU), medium (8,400 RU) and high-density (13,600 RU) surface. All data sets are fitted to a single binding site model.

Since F(ab')$_2$ and Fab' fragments lack the Fc part, which could contribute to the non-specific binding effect, a 'contaminated' batch of sLe^a-Lem (**1b**) was also tested on these fragments. In the same experiment, the influence of the buffer compositions was tested for both the sample and running buffer. For this purpose, a plain HBS buffer (10 mM HEPES, 150 mM NaCl, *pH* 7.4; HBS-N) was compared to HBS containing 3 mM EDTA and 0.05% polysorbate (HBS-EP). The same concentration (20 μM) of sLe^a-Lem (**1b**) was screened in all combinations of running/sample buffers and immobilized fragments (*Fig.* 4-18).

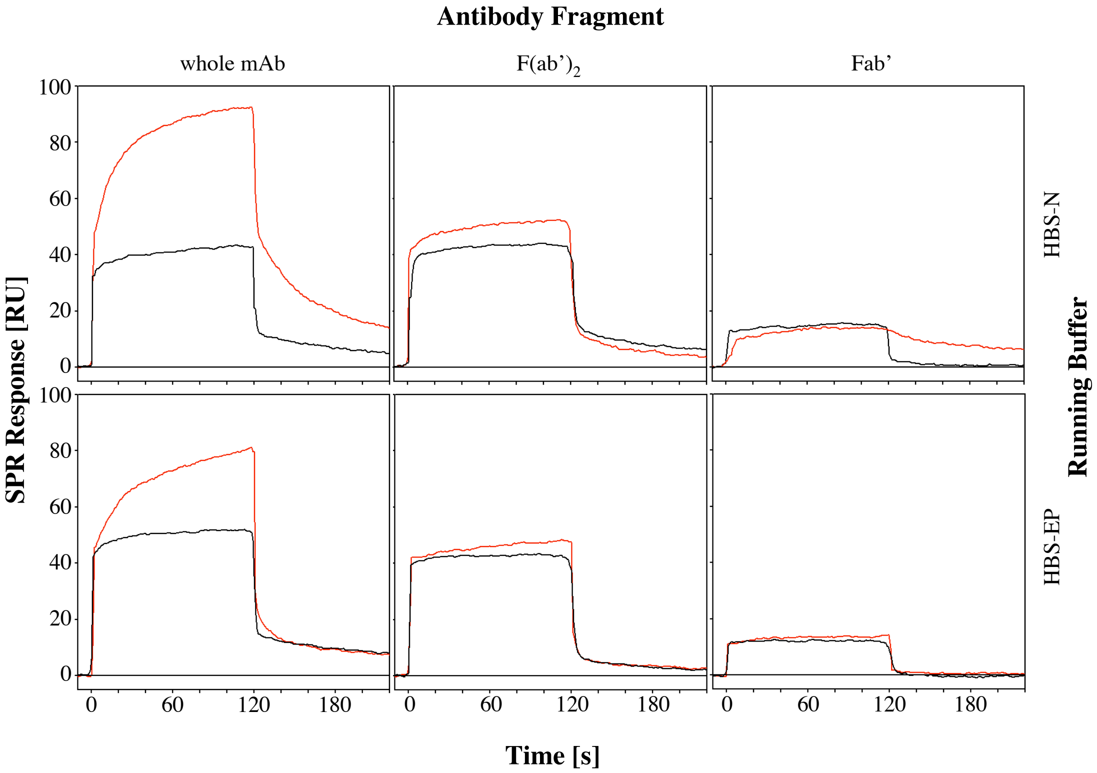


Figure 4-18: Influence of buffer additions and antibody fragment size on the non-specific binding of sLe^a-Lem (**1b**). The tetrasaccharide was diluted in either HBS-N (red) or HBS-EP (black) as *sample buffer* and injected over whole antibody, F(ab')$_2$ and Fab' surfaces. Either HBS-N (top row) or HBS-EP (bottom row) was used as *running buffer* throughout the experiments.

A reduction of the antibody molecule to its antigen-binding parts was able to reduce but not completely eliminate non-specific binding. This indicates that the Fc part is at least partly involved in this effect. In addition, the choice of running and sample buffers also influenced the degree of non-specific binding. When HBS-EP was used as sample buffer, the undesired signal effect could be reduced to a minimum, especially in combination with HBS-EP as running buffer.

4.3.5 Comparison of capturing approaches

The detection of small molecules on large targets usually requires a densely packed surfaces and high protein activity. A high-density immobilization of GSLA-2 would facilitate the detection of weakly binding derivatives or of smaller analytes. Capturing approaches lead to an oriented attachment of the target molecules and are therefore a flexible way of generating highly active surfaces (see *section* 2.1.4). Staphylococcal protein A (SpA) and streptococcal protein G (SpG) are the most widely used natural capturing proteins for the purification and immobilization of antibodies. Both proteins are produced by bacteria as a part of their defense strategy to circumvent the immune system. They bind to similar sites on the constant region (Fc) of the immunoglobulins of various species and classes. However, the binding affinity differs significantly between these classes and the widely used class of murine IgG1 is reported to bind only weakly to SpA and very weakly to SpG [52]. While these unfavorable binding properties might be sufficient for some separation and purification steps, immobilization is more demanding in terms of stability. In order to evaluate the applicability to capturing approaches on Biacore, both proteins were covalently immobilized and a concentration series of GSLA-2 was injected over both surfaces (*Fig.* 4-19A&B). The immobilization densities obtained by standard amine coupling were rather low (≤ 4000 RU), but corresponded with those reported in literature [33, 42, 43]. Coupling of SpG was even more challenging. One reason for this limitation are the low *pI* values of the two proteins (5.1 for SpA [53] and even 4.2 for recombinant SpG [54]), which have a negative influence on the preconcentration by surface attraction (see *section* 2.1.4).

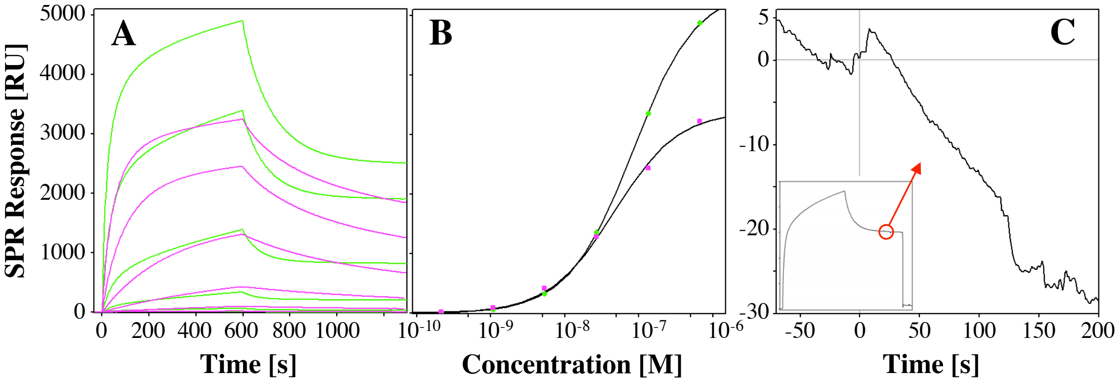


Figure 4-19: Capturing of GSLA-2 using bacterial proteins. **A:** Binding of GSLA-2 (0.2-666 nM) to either SpA (green) or SpG (magenta). **B:** Concentration plot of SpA and SpG using the binding signal after a 600 s injection period. C: Binding signal of a 200 μM sLe^{a} solution on 3000 RU GSLA-2 captured with SpA. The inlet shows the capturing and sLe^{a} binding (red circle).

GSLA-2 clearly bound to SpA as well as to SpG, but showed significant deviations in the binding kinetics. The dissociation phase of the SpG/GSLA-2 interaction showed a uniform but very moderate stability, and could therefore not be used for a stable capturing of the antibody. Binding to SpA, on the other hand, seemed to be composed of two separate dissociation processes. Especially in case of higher concentrations, a first fast decay could be observed, which was then followed by a very stable binding phase (*Fig*. 4-19A). However, when the binding affinity of the two interactions was estimated by fitting the signal intensity after 10 min of injection to a single binding site model (*Fig*. 4-19B), the antibody seemed to bind stronger to SpG (K_D = 46 nM) than to SpA (K_D = 83 nM). Therefore, the binding was also evaluated kinetically and both interactions corresponded best though not perfectly with a heterogeneous binding model (*Table* 4-4, for plots see *appendix* C2).

Table 4-4: Kinetic evaluation of the interaction of GSLA-2 (0.2-666 nM) with SpA and SpG, when fitted to a heterogeneous binding model.

	k_{on} [10^5 $M^{-1}s^{-1}$]		**k_{off} [10^{-3} s^{-1}]**		**K_D [nM] [a]**		**$t_{1/2}$ [min] [b]**	
Site [c]	**1**	**2**	**1**	**2**	**1**	**2**	**1**	**2**
SpA	1.68	0.18	8.06	0.01	48	0.6	1.4	1070
SpG	1.07	0.11	1.30	0.54	12	51	9	21

[a] Calculated using $K_D = k_{off}/k_{on}$. [b] Calculated using $t_{1/2} = \ln 2/k_{off}$. [c] The expression 'site' refers to a heterogeneity in the binding behavior but not necessarily to separate binding sites on the capturing protein.

Since the deviations in the association rate constants are similar between the two proteins for both binding sites (factor 10), the major difference was found in the off-rate. Indeed, while SpG shows only a twofold deviation between the two binding sites, the rates vary by more than a factor of 800 in case of SpA. A possible explanation of the more complex binding mode of SpA might be the source of the two proteins. While SpA was used in its natural form, SpG was purchased as a recombinant form lacking binding domains for albumin and the constant part of IgG Fab fragments [54]. Since SpA is also reported to bind Fab fragments [55], this interaction might overlay with the specific Fc binding. Comparing the kinetic analysis with the estimated 'equilibrium' plot (*Fig*. 4-19B) illustrates, that such estimations might be dangerous if the steady state phase is not reached. Only few literature data are available about the determination of dissociation constants for SpA and SpG. Two Biacore studies with human IgG1 determined K_D values of 7 nM [56] and 47 nM [57] for SpA and SpG, respectively.

While the dissociation rate of SpG is not sufficient for a stable capturing, the second phase of the SpA dissociation seemed to fulfill these requirements. Therefore, sLe^a-Lem (**1b**) was injected at a concentration that usually resulted in a saturation of the binding sites (200 μM). However, the detected SPR signal (≤ 5 RU) was much lower than expected and corresponded to an activity of only ~10%. In addition, the surface showed a significant baseline drift, which makes an appropriate evaluation of small signal intensities even more difficult (*Fig.* 4-19C). Non-specific binding of the Fab' fragments of GSLA-2 (see above) could be one reason of the low activity, since steric hindrance might inhibit an interaction of sLe^a-Lem (**1b**) with the variable regions. Another explanation might be found in a recent study by Sagawa *et al.* [55], where they reported conformational changes of the antibody upon binding, which led to a significant weakening of antigen interactions. A possible workaround for the rather low surface stability of SpA capturing was presented by Catimel *et al.* [33], who cross-linked the antibody with SpA after the capturing by injecting dimethylpimelidate. However, a decreased activity was also observed for this approach, because the cross-linking reagent could also target lysine residues in the binding site [58].

In order to increase the specificity and stability of the IgG1 capturing, the experiment was repeated with an immobilized polyclonal rabbit anti-mouse IgG1 antibody. Even though the capturing antibody could be immobilized at a high surface density (> 15,000 RU), the capturing density of GSLA-2 was rather low (< 5,000 RU). It is not yet clear whether this limitation is caused by an immobilization-induced deactivation of the primary antibody, steric hindrance, or by a SPR phenomenon (greater distance of GSLA-2 from the surface). However, the captured GSLA-2 fraction was much more stable than those on SpA or SpG, and was therefore suitable for the screening of small molecules (*Fig.* 4-20A). The activity of the captured GSLA-2 surface was evaluated by injecting a concentration series of sLe^a-Lem (**1b**, *Fig.* 4-20B&C).

Compared to SpA capturing, the baseline was more stable and the binding activity was significantly increased (*Fig.* 4-20B). However, with slightly more than 33% the activity was still far beyond the expectations for an oriented surface. The reason for this phenomenon is not yet clear and has to be investigated further. Despite the low signal intensities, the steady state responses could be fitted to a single site binding model (*Fig.* 4-20C). The resulting K_D of 15 μM was only slightly higher than the values obtained from covalent coupling approaches (4-12 μM). Due to their low capturing densities and binding activities, the investigated capturing approaches were not suitable for the screening of small molecules on GSLA-2 and were therefore not developed any further.

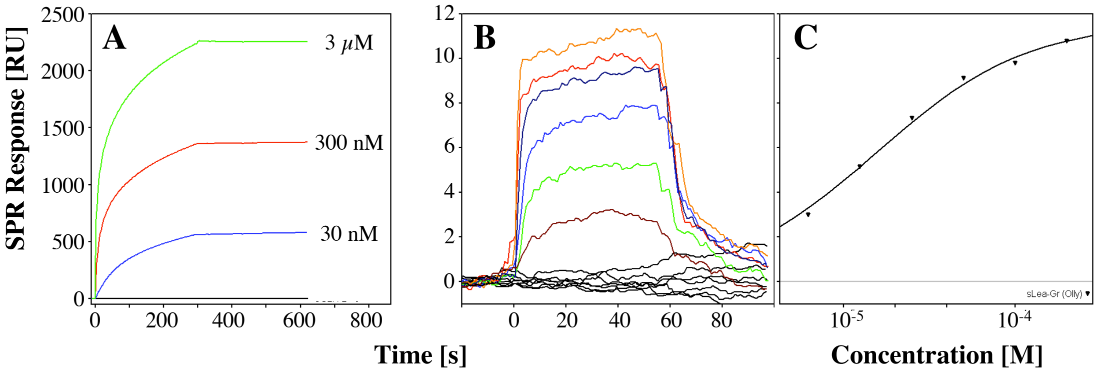


Figure 4-20: Capturing of GSLA-2 with a rabbit anti-mouse IgG1 antibody. **A:** Injections (5 min) of GSLA-2 at three different concentrations to the immobilized capturing antibody. **B, C:** Test of the surface activity by injecting sLe^a-Lem (**1b**) on a captured GSLA-2 surface. Overlaid sensorgrams are shown in **B**, the corresponding steady state plot in **C**.

4.4 Conclusions

The monoclonal antibody GSLA-2 shows a high selectivity of the sialyl Lewis^a epitope (sLe^a), and is therefore clinically used as a diagnostic tool for the detection of colon and pancreas cancer, However, little is known about the molecular binding specificity and its binding to related carbohydrates. Therefore, a stable and reproducible Biacore assay was developed based on the covalent immobilization of GSLA-2 on the sensor chip surface. Screening of sLe^a in solution resulted in kinetic properties, which are typical for carbohydrate-protein interaction, such as rapid association and dissociation rate constants and a rather low affinity (low micromolar range). This Biacore assay, in combination with data from STD NMR experiments was used to screen a set of structurally related compounds for binding to the antibody. Evaluation of Biacore data identified the presence and orientation of the Neu5Ac and Glc*N*Ac moieties as essential for a detectable binding event. Any replacement, removal, or reversal of Glc*N*Ac (e.g. in the case of sLe^x) led to a loss of affinity. Furthermore, the methyl group of the fucose moiety seems not to be essential but beneficial for binding. STD NMR experiments showed that GSLA-2 recognized the whole tetrasaccharide epitope and confirmed the high importance of the two *N*-acetyl groups. This allows a much deeper insight into the binding mode of sLe^a to the antibody and helps explaining the difference to the selectins, which recognize both sLe^a and sLe^x.

As a representative of the most popular class of monoclonal antibodies, GSLA-2 also allowed to investigate different aspects of Biacore assay design and optimization. For example, the influence of a labeling group to the binding affinity could be analyzed. While a Lemieux spacer, attached to the reducing end of sLe^a for a facilitated purification of the molecules, did only receive minor parts of the saturation transfer in the STD NMR experiment, its contribution to the binding affinity was nevertheless detectable in the Biacore assay. The slightly improved affinity is most likely caused by hydrophobic interactions or by an improved k_{on} (e.g. by ‘pushing’ the molecule out of the solvent), and illustrates the great importance of the label-free SPR technology. As expected, increasing the surface density of the immobilized antibody led to enlarged signal intensities, while the K_D value remained constant. The active protein fraction also got smaller with increasing surface densities, indicating a limited accessibility of binding sites.

In an additional experiment, the benefit of using antibody fragments was illustrated by comparing the whole GSLA-2 antibody with its F(ab’)$_2$ and Fab’ fragments. For this purpose, a method was developed for the fast and reliable digestion, cleavage, purification and functionalization of such fragments. Both truncated antibody molecules could be immobilized by amine coupling. While the relative immobilization density was lower compared to untreated GSLA-2, the surface activity was remarkably increased. In addition, $F(ab')_2$ and Fab' fragments also showed a lower susceptibility for non-specific binding, as it occurred with some sLe^a batches. No-specific binding could also be reduced by adding EDTA (and polysorbate) to the running and sample buffers. While the immobilization of fragmented antibodies was of an advantage in the case of GSLA-2, the time-consuming procedure and the loss of protein through the preparation process has to be considered.

Finally, three different methods for the capturing of GSLA-2 were compared in terms of stability and activity. The affinity of the mouse IgG1 towards staphylococcal protein A and streptococcal protein G was too weak for a stable capturing baseline and generated rather low surface densities. In contrast, a specific rabbit anti-mouse IgG1 antibody was able to capture GSLA-2 with slightly higher density and a remarkably increased stability. However, none of the tested capturing surfaces showed an improved activity as it had been expected for an oriented immobilization. Due to the limitations of the capturing approach, covalent coupling is clearly recommended, since it generated highly dense and suitably active antibody surfaces. If higher signal intensities are required for future experiments, an optimization of the Fab' immobilization density seems to be most promising, since these fragments showed an excellent combination of specificity and activity in the screening assays.

4.5 Literature

[1] M. Malmqvist, *Nature* **1993**, *361*, 186.

[2] J. M. Reichert, A. K. Pavlou, *Nat Rev Drug Discov* **2004**, *3*, 383.

[3] S. Y. Tetin, S. D. Stroupe, *Curr Pharm Biotechnol* **2004**, *5*, 9.

[4] O. H. Brekke, I. Sandlie, *Nat Rev Drug Discov* **2003**, *2*, 52.

[5] E. A. Padlan, *Mol Immunol* **1994**, *31*, 169.

[6] *Protein Data Bank: The worldwide repository for macromolecular structures*, Research Collaboratory for Structural Bioinformatics (RCSB), http://www.pdb.org/, **2005**.

[7] R. Jefferis, J. Lund, *Chem Immunol* **1997**, *65*, 111.

[8] M. R. Clark, *Chem Immunol* **1997**, *65*, 88.

[9] G. Kohler, C. Milstein, *Nature* **1975**, *256*, 495.

[10] J. Wardle, R. Pope, *J Psychosom Res* **1992**, *36*, 609.

[11] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Res* **2000**, *28*, 235.

[12] D. R. Bundle, in *Glycosciences: Status and Perspectives* (Eds.: H. J. Gabius, S. Gabius), Chapman and Hall, Weinheim, **1997**, pp. 311.

[13] H. Maaheimo, P. Kosma, L. Brade, H. Brade, T. Peters, *Biochemistry* **2000**, *39*, 12778.

[14] T. Haselhorst, J. F. Espinosa, J. Jimenez-Barbero, T. Sokolowski, P. Kosma, H. Brade, L. Brade, T. Peters, *Biochemistry* **1999**, *38*, 6449.

[15] M. Mayer, B. Meyer, *Angew Chem Int Ed* **1999**, *38*, 1784.

[16] M. A. Johnson, B. M. Pinto, *J Am Chem Soc* **2002**, *124*, 15368.

[17] J. L. Magnani, M. Brockhaus, D. F. Smith, V. Ginsburg, M. Blaszczyk, K. F. Mitchell, Z. Steplewski, H. Koprowski, *Science* **1981**, *212*, 55.

[18] J. L. Magnani, B. Nilsson, M. Brockhaus, D. Zopf, Z. Steplewski, H. Koprowski, V. Ginsburg, *J Biol Chem* **1982**, *257*, 14365.

[19] T. Nakayama, M. Watanabe, T. Katsumata, T. Teramoto, M. Kitajima, *Cancer* **1995**, *75*, 2051.

[20] T. Nakagoe, T. Sawai, T. Tsuji, M. Jibiki, M. Ohbatake, A. Nanashima, H. Yamaguchi, T. Yasutake, H. Ayabe, K. Arisawa, *Tumour Biol* **2001**, *22*, 115.

[21] U. Spohr, O. Hindsgaul, R. U. Lemieux, *Can J Chem* **1985**, *63*, 2644.

[22] A. M. Duijvestijn, E. Horst, S. T. Pals, B. N. Rouse, A. C. Steere, L. J. Picker, C. J. Meijer, E. C. Butcher, *Am J Pathol* **1988**, *130*, 147.

[23] E. L. Berg, M. K. Robinson, O. Mansson, E. C. Butcher, J. L. Magnani, *J Biol Chem* **1991**, *266*, 14869.

[24] M. Ugorski, A. Laskowska, *Acta Biochim Pol* **2002**, *49*, 303.

[25] G. Gorrin, S. Christensen, M. Moore, K. Trujillo, M. Farwell, Y. Teramoto, J. L. Magnani, *Tumour Biol* **1997**, *18*, Supp. 2: 79.

[26] P. D. Rye, N. V. Bovin, E. V. Vlasova, A. A. Molodyk, A. Baryshnikov, F. T. Kreutz, W. I. Garinther, B. C. Schultes, A. A. Noujaim, R. Madiyalakan, J. Magnani, O. Nilsson, K. Nilsson, K. Nustad, L. Norum, H. Bell, Y. Cao, M. R. Suresh, D. L. Very, J. V. Freeman, K. K. Yeung, J. Hilgers, *Tumour Biol* **1998**, *19*, 390.

[27] H. S. Silverman, S. Parry, M. Sutton-Smith, M. D. Burdick, K. McDermott, C. J. Reid, S. K. Batra, H. R. Morris, M. A. Hollingsworth, A. Dell, A. Harris, *Glycobiology* **2001**, *11*, 459.

[28] H. S. Silverman, M. Sutton-Smith, K. McDermott, P. Heal, S. H. Leir, H. R. Morris, M. A. Hollingsworth, A. Dell, A. Harris, *Glycobiology* **2003**, *13*, 265.

[29] R. L. Rich, D. G. Myszka, *J Mol Recognit* **2000**, *13*, 388.

[30] M. Strandh, B. Persson, H. Roos, S. Ohlson, *J Mol Recognit* **1998**, *11*, 188.

[31] M. Adamczyk, J. A. Moore, Z. Yu, *Methods* **2000**, *20*, 319.

[32] P. Peluso, D. S. Wilson, D. Do, H. Tran, M. Venkatasubbaiah, D. Quincy, B. Heidecker, K. Poindexter, N. Tolani, M. Phelan, K. Witte, L. S. Jung, P. Wagner, S. Nock, *Anal Biochem* **2003**, *312*, 113.

[33] B. Catimel, M. Nerrie, F. T. Lee, A. M. Scott, G. Ritter, S. Welt, L. J. Old, A. W. Burgess, E. C. Nice, *J Chromatogr A* **1997**, *776*, 15.

[34] B. Johnsson, S. Lofas, G. Lindquist, A. Edstrom, R. M. Muller Hillgren, A. Hansson, *J Mol Recognit* **1995**, *8*, 125.

[35] A. Skerra, *J Mol Recognit* **2000**, *13*, 167.

[36] W. M. Mullett, E. P. Lai, J. M. Yeung, *Methods* **2000**, *22*, 77.

[37] L. Herfurth, B. Ernst, B. Wagner, D. Ricklin, D. S. Strasser, J. L. Magnani, A. J. Benie, T. Peters, *J Med Chem* **2005**, *submitted.*

[38] M. J. Cannon, G. A. Papalia, I. Navratilova, R. J. Fisher, L. R. Roberts, K. M. Worthy, A. G. Stephen, G. R. Marchesini, E. J. Collins, D. Casper, H. Qiu, D. Satpaev, S. F. Liparoto, D. A. Rice, Gorshkova, II, R. J. Darling, D. B. Bennett, M. Sekar, E. Hommema, A. M. Liang, E. S. Day, J. Inman, S. M. Karlicek, S. J. Ullrich, D. Hodges, T. Chu, E. Sullivan, J. Simpson, A. Rafique, B. Luginbuhl, S. N. Westin, M. Bynum, P. Cachia, Y. J. Li, D. Kao, A. Neurauter, M. Wong, M. Swanson, D. G. Myszka, *Anal Biochem* **2004**, *330*, 98.

[39] L. Herfurth, PhD thesis, *Molekulare Erkennung von Kohlenhydraten durch monoklonale Antikörper und ein Lektin - NMR-Experimente*, Medical University of Lübeck (Lübeck, Germany), **2002**.

[40] U. K. Laemmli, *Nature* **1970**, *227*, 680.

[41] D. Kang, Y. S. Gho, M. Suh, C. Kang, *Bull Korean Chem Soc* **2002**, *23*, 1511.

[42] R. L. Rich, D. G. Myszka, *J Mol Recognit* **2001**, *14*, 223.

[43] R. J. Darling, U. Kuchibhotla, W. Glaesner, R. Micanovic, D. R. Witcher, J. M. Beals, *Biochemistry* **2002**, *41*, 14524.

[44] W. Huber, S. Perspicace, J. Kohler, F. Muller, D. Schlatter, *Anal Biochem* **2004**, *333*, 280.

[45] W. S. Somers, J. Tang, G. D. Shaw, R. T. Camphausen, *Cell* **2000**, *103*, 467.

[46] M. Rinnbauer, B. Ernst, B. Wagner, J. Magnani, A. J. Benie, T. Peters, *Glycobiology* **2003**, *13*, 435.

[47] R. U. Lemieux, D. R. Bundle, D. A. Baker, *J Am Chem Soc* **1975**, *97*, 4076.

[48] D. S. Wilson, J. Wu, P. Peluso, S. Nock, *J Immunol Methods* **2002**, *260*, 29.

[49] C. Bellac, Diploma thesis, *Herstellung von Methoden zur selektiven Kopplung von Targetproteinen an Biosensoren,* University of Basel (Basel, Switzerland), **2002**.

[50] Y. Hiller, J. M. Gershoni, E. A. Bayer, M. Wilchek, *Biochem J* **1987**, *248*, 167.

[51] R. J. Ober, E. S. Ward, *Anal Biochem* **1999**, *271*, 70.

[52] E. Harlow, D. P. Lane, *Using antibodies: a laboratory manual*, 2nd rev. ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, **1999**.

[53] I. Bjork, B. A. Petersson, J. Sjoquist, *Eur J Biochem* **1972**, *29*, 579.

[54] C. R. Goward, J. P. Murphy, T. Atkinson, D. A. Barstow, *Biochem J* **1990**, *267*, 171.

[55] T. Sagawa, M. Oda, H. Morii, H. Takizawa, H. Kozono, T. Azuma, *Mol Immunol* **2005**, *42*, 9.

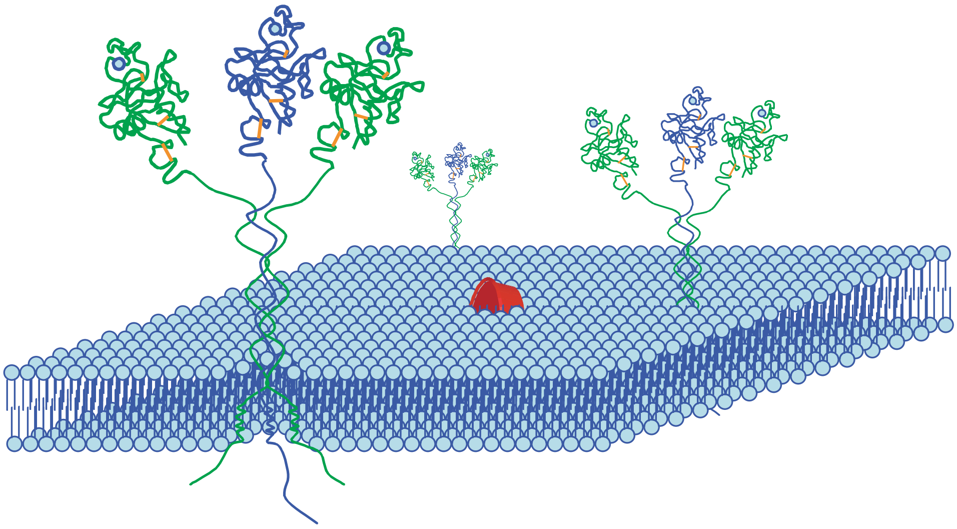
[56] H. G. Svensson, H. R. Hoogenboom, U. Sjobring, *Eur J Biochem* **1998**, *258*, 890.

[57] S. Radaev, P. D. Sun, *J Biol Chem* **2001**, *276*, 16478.

[58] E. C. Nice, B. Catimel, *Bioessays* **1999**, *21*, 339.

Chapter 5

Asialoglycoprotein Receptor



5.1 Introduction

Drug and gene targeting to specific organs is a promising approach for the development of highly effective therapies while reducing side-effects. Carbohydrate-lectin interactions between liver-cell receptors and their physiological (or synthetic) ligands have been described as an efficient method showing high specificity. However, for the development of new carbohydrate-based agents, a detailed knowledge about binding modes and affinities is crucial. Therefore, the interaction of the asialoglycoprotein receptor (ASGP-R) with some of its natural ligands has been investigated using Biacore.

5.1.1 The asialoglycoprotein receptor as a promising drug target

The liver is the major metabolic organ in the human body and is responsible of clearing the blood from undesired endo- and exogenous compounds. Therefore, a plethora of enzymes and receptors is specialized in metabolizing and excreting different classes of molecules. With some 500,000 receptors/hepatocyte, the asialoglycoprotein receptor is one of the most abundant receptors and an important contributor to this hepatic machinery [1, 2]. It was first described by Ashwell and Morell [3], when they discovered that glycoproteins featuring a terminal galactose (or *N*-acetyl galactosamine) were rapidly removed from the circulation by a receptor-mediated mechanism. Normally, sialic acid residues mask the penultimate galactose of many oligosaccharide chains. However, these masking groups might be removed by sialidases, resulting in asialoglycoproteins. Many studies have been performed in recent years in order to investigate the differential biological function and physiological role between sialylated and desialylated glycoproteins. Suppression of antigenicity, masking and metastasis of tumor cells, clearance of apoptotic liver cells or regulation of glycoprotein hormones are some of the proposed functions. However, receptor-deficient mice did not show significantly elevated plasma level of asialoglycoproteins nor was their life span influenced in any way [1]. Therefore, the exact physiological relevance of the receptor is still not clear and has to be investigated further.

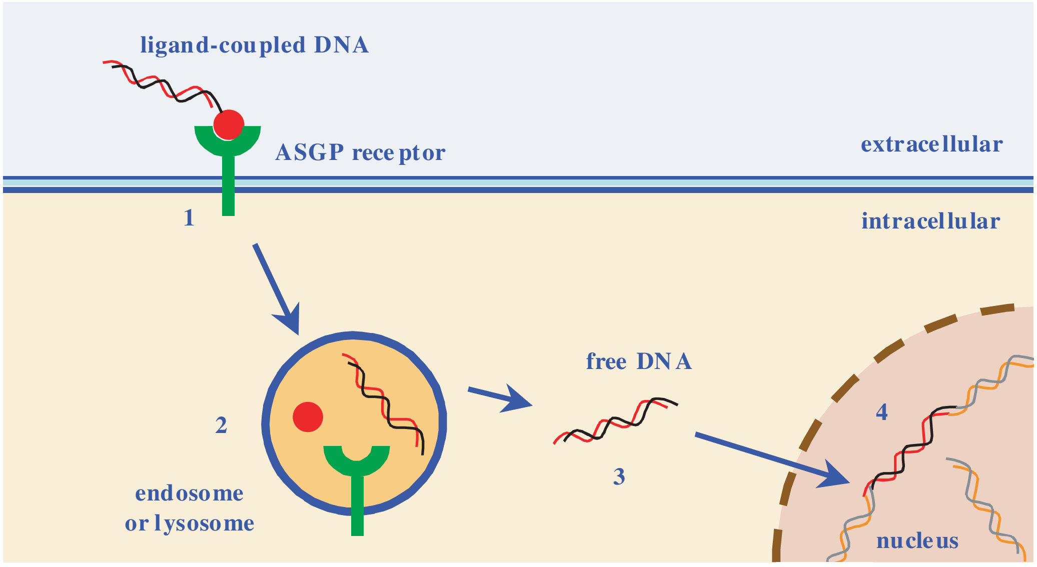


Figure 5-1: Liver-directed gene targeting using ASGP receptors. DNA is coupled to natural or synthetic ligands of ASGP-R via a cleavable linker (1). After binding and endocytosis, the DNA is released in the endosome or lysosome (2). The free DNA is released (3) and can pass to the nucleus where it is inserted into the host DNA (4).

Hence the liver is not only the center of metabolism but is often affected by genetic disorders, intoxication or tumor growth and metastasis. Due to the great importance of a fully functioning liver, such diseases often dramatically reduce a patient's quality or even expectance of life. Bringing drugs, radionuclides, or genes directly into hepatocytes is therefore a major aim for an effective therapy of liver disorders. Receptor-mediated endocytosis could be the key function for selectively transporting therapeutic agents from the blood to the hepatocytes, and ASGP-R is one of the most promising targets for this purpose. Several methods have been published to deliver genes [4, 5], drugs [6], anti-tumor or anti-viral agents [7], radiolabels or lipoproteins [8] to the hepatocytes via ASGP-R uptake (for reviews see Wu *et al*. [9] and Nishikawa [10]). By coupling the diagnostic or therapeutic agents to physiological or synthetic ligands through cleavable linkers, a release of the agents after the uptake can be achieved (*Fig*. 5-1). Compared to alternative approaches of liver-targeting such as viral vectors (e.g. by attenuated hepatitis B viruses [11]) or antibodies, sugar-lectin interactions have the advantage of being less immunogenic while retaining their specificity.

5.1.2 Structure and function of ASGP-R

Isolation of ASGP-R from liver preparations showed that the receptor consists of two related proteins in a concentration ratio of 3:1. Each polypeptide chain is in the range of 40-60 kDa and is glycosylated with two or three N-linked oligosaccharide chains. The two receptor subtypes are named HL-1 (i.e. hepatic lectin 1) and HL-2, the human receptors are usually simply referred to as H1 and H2 [12]. The sequence identity of the two subunits is 55 % for the human receptor [13]. Different studies suggested that both subunits are required to build a functional receptor, forming a hetero-oligomeric receptor complex. Hence the exact stoichiometry is still not fully solved, with suggestion from tri- to hexameric complexes of various H1/H2 ratios [12]. Since triantennary ligands were found to bind with a very high affinity (see *section* 5.1.4), a trimeric receptor consisting of two H1 and one H2 subunits is expected to be the minimum requirement (*Fig*. 5-2A) [14].

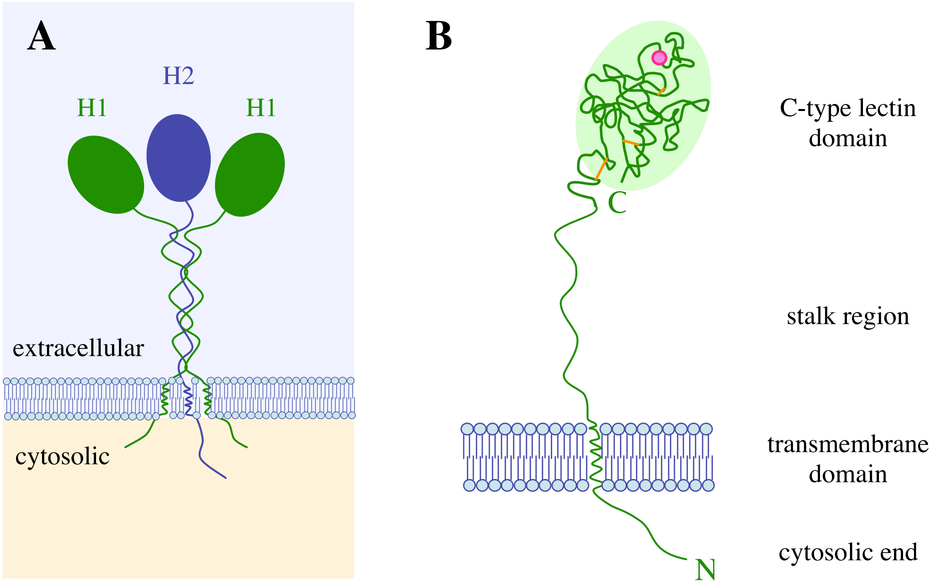


Figure 5-2: Structure of the asialoglycoprotein receptor. **A:** Hetero-oligomeric complex of two H1 and one H2 subunits, which is believed to be the minimum requirement of a functional receptor. **B:** Anatomy of a receptor subunit with different intra- and extracellular domains. Disulfide bridges are indicated as orange lines, the calcium ion involved in ligand binding is indicated by a pink circle.

Both ASGP-R subtypes are membrane-bound proteins, which consist of approximately 300 amino acids. Their general structure can be differentiated in a small *N*-terminal cytoplasmic end, a single transmembrane domain (~20 hydrophobic residues), and a large *C*-terminal exoplasmic part, consisting of a stalk region and the *carbohydrate* *recognition domain* (CRD; *Fig*. 5-2B) [15]. ASGP-R, and especially its CRD, is classified as a C-type lectin (in contrast to S-type lectins). This class of proteins is able to bind carbohydrate structures in calcium-dependent manner [16].

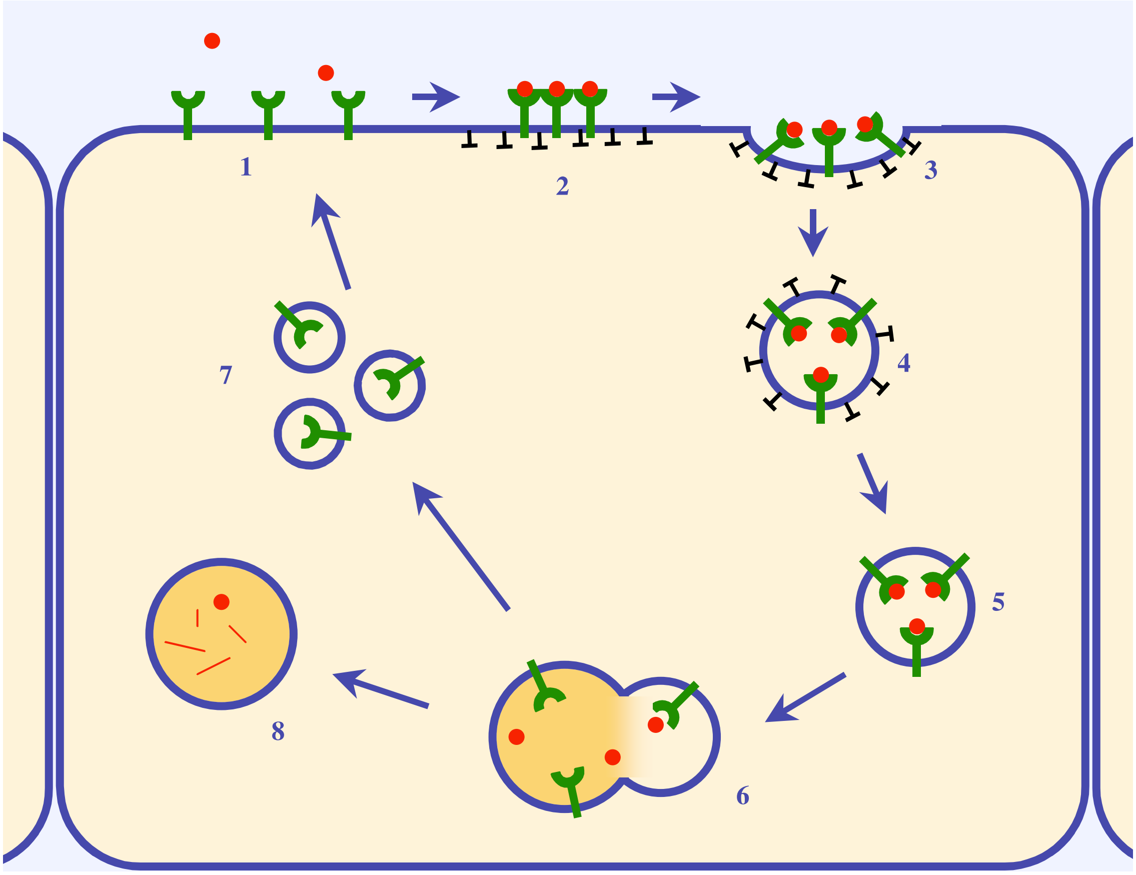


Figure 5-3: Receptor-mediated endocytosis. Membrane-bound receptors (1) cluster upon ligand binding (2), forming clathrin-coated pits (3). These are internalized as coated vesicles (4). After uncoating, the vesicles (5) fuse with endosomes and release the ligand (6). While the receptors are transported to the surface in recycling vesicles (7), the ligand is degraded (8).

For a successful clearing of asialoglycoproteins from the circulation, these ligands have to be internalized and degraded. This process is called *receptor-mediated endocytosis* and involves several membrane-based steps. First, the receptor clusters into specialized clathrin-coated domains, first forming coated pits and later coated vesicles, which are internalized. After removal of the clathrin coat the vesicles fuse with intracellular endosomes. The significantly lower *pH* (6.0-6.2 in early endosomes) leads to a dissociation of both calcium ions and bound ligands. While the ASGP receptor is recycled and transported back to the plasma membrane by recycling vesicles, the ligand-containing endosomes further fuse with lysosomes, where the degradation process takes place (*Fig*. 5-3) [12].

In contrast to signal-transducing receptors, transport receptors like ASGP-R perform receptor-mediated endocytosis as a continuous process, i.e. independent of any ligands. However, ligand binding was shown to increase the internalization rate by a factor of ~2 [17]. In principle, this triggering might be caused by a ligand-induced conformational change or by a clustering after cross-linking by the ligand. Since monovalent glycopeptides are internalized with essentially the same kinetics as multivalent ligands, the hypothesis of a conformational change across the membrane is more likely [17]. While natural ASGP receptors always contain both H1 and H2 subtypes, H1 seem to be the major requirement for receptor-mediated endocytosis. A tyrosine residue (Tyr-5) in H1 is essential for triggering the uptake signal, while the corresponding phenylalanine 5 in H2 is not able to induce the same effect [18, 19]. Since the occurrence of two subunits is strictly conserved in mammalian species, hetero-oligomers seem to have some distinct advantage over homo-oligomers (e.g. by enhancing affinity or stabilizing the complex) [2].

5.1.3 H1-CRD as a molecular target

For many years after its discovery, the crystal structure of the ASGP-R subunits was not available and little was known about the carbohydrate binding site. Meier *et al.* published the first crystal structure of a recombinant form of the H1 subunit of human ASGP-R in 2000 [13]. For this purpose, they expressed only the CRD (residues 147-290) in *E. coli*, solubilized and refolded the protein, and determined its structure by X-ray crystallography at a resolution of 2.3 Å (in the presence of lactose and calcium). While most of the protein could be resolved, some residues at the termini as well as the lactose ligand could not be interpreted. The structure shows a globular protein, containing two α-helices and eight β-strands. Of the seven cysteine residues of the H1-CRD, six were forming disulfide bridges while the seventh is not described by the structural data. Furthermore, three calcium-binding sites were detected, which pinned together several loops. One of these calcium ions (at site 2) is also part of the carbohydrate binding site. The calcium ions are coordinated by eight oxygen atoms, five of which are forming the basis of a pentagonal bipyramidal geometry. While the other two oxygens are those of protein carboxylate groups for site 1 and 3, two water molecules saturate the coordination at site 2. Upon ligand binding, they are replaced by the oxygen atoms of the 3’ and 4’ hydroxyl groups of the sugar molecule. Even though there was no ligand visible in the crystal structure, homology studies with the rat mannose binding protein made it possible to identify the residues involved in sugar binding (Gln-239, Asp-241, Trp-243, Glu-252, and Asn-264) [13]. The crystal structure (PDB code 1DV8) with some of the important sites as well as the amino acid sequence of ASGP-R H1-CRD are shown in *figure* 5-4.

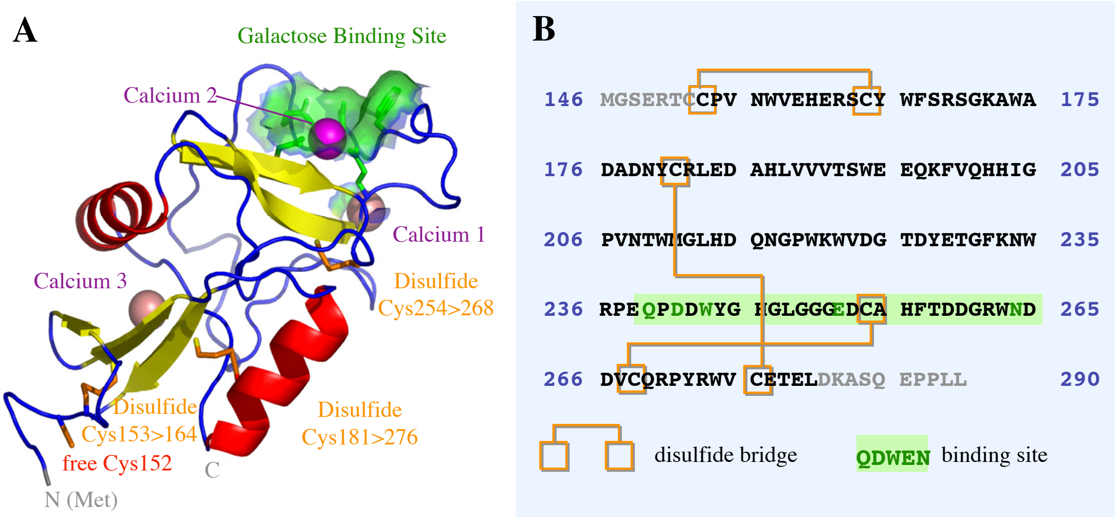

A
Galactose Binding Site
Calcium 2
Calcium 1
Disulfide
Cys254>268
Calcium 3
Disulfide
Cys153>164
free Cys152
Disulfide
Cys181>276
C
N (Met)
B
146 MGSERTCCPV NWVEHERSCY WFSRSGKAWA 175
176 DADNYCRLED AHLVVVTSWE EQKFVQHHIG 205
206 PVNTWMGLHD QNGPWKWVDG TDYETGFKNW 235
236 RPEQPDDWYG HGLGGGEDCA HFTDDGRWND 265
266 DVCQRPYRWV CETELDKASQ EPPLL 290
disulfide bridge
QDWEN binding site

Figure 5-4: Structure of H1-CRD with important sites. **A:** Crystal structure (PDB entry 1DV8). Helices are colored in red, strands in yellow, disulfide bridges are represented as orange sticks, calcium ions as magenta spheres, and the sugar-binding site is highlighted in green. The *N*-terminus up to Cys-152 is not part of the PDB file and was added for illustration purpose. **B:** Amino acid sequence of the ASGP-R H1-CRD with indicated disulfide bridges and binding site. Residues highlighted in grey are not visible in the published crystal structure file.

ASGP-R H1-CRD is one of the rare C-type lectins with three calcium sites as an integral parts of the structure. Two of them (sites 1 and 2) are located in close proximity (*Fig*. 5-4) and even share a single glutamate residue (Glu-252). While site 2 is essential for sugar binding and is present in all C-type lectins, site 1 is described for some collectin-related proteins (e.g. mannose-binding protein) and site 3 was only found in the subunits of the coagulation factor IX/X binding protein. The three calcium binding sites of H1-CRD share some common features while showing clear differences in other aspects. First, they all form a pentagonal bipyramidal geometry, where the calcium is coordinated by eight oxygen atoms. All sites involve one main-chain carbonyl group (1 = Glu-252, 2 = Asp-265, 3 = Val-190) as well as several side-chain carboxyl groups (Asp, Asn, Glu, Gln) and water (*Fig*. 5-5). While the sugar-binding site 2 only coordinates two water molecules, the other sites involve three waters. In the case of site 3, both carboxyl oxygens of Glu-196 and Glu-277 are involved in calcium coordination [13].

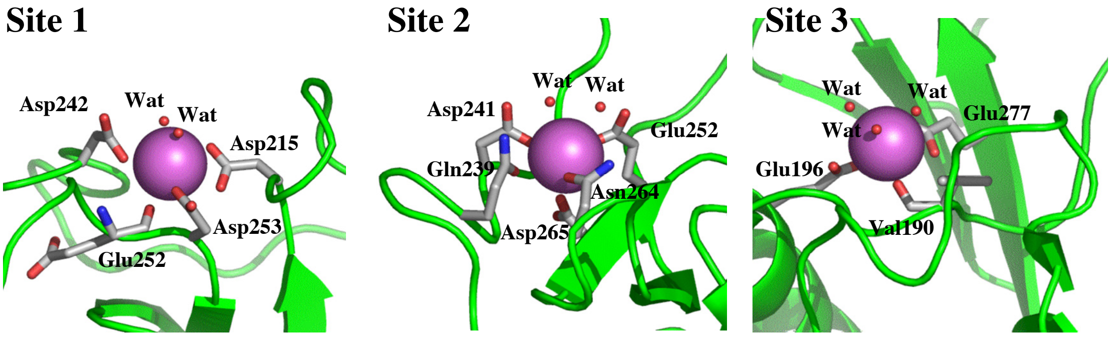

Site 1
Asp242
Wat
Wat
Asp215
Asp253
Glu252
Site 2
Wat
Wat
Asp241
Glu252
Gln239
Asn264
Asp265
Site 3
Wat
Wat
Wat
Glu277
Glu196
Val190

Figure 5-5: Comparison of the three calcium binding sites of the ASGP-R H1-CRD. The protein backbone is represented in green and the calcium ion as a purple sphere. Amino acids involved in calcium binding are shown as sticks and water molecules as small red spheres.

5.1.4 Ligand binding to H1-CRD and ASGP-R

The ASGP receptor recognizes terminal galactose and *N*-acetyl galactosamine moieties (*Fig*. 5-6A), as they occur naturally on desialylated glycoproteins. Binding of terminal Gal*N*Ac was found to more than tenfold stronger than galactose [20]. It soon became evident that the valency of these moieties on a ligand is an important determinant for its binding efficacy. While the affinity for a single galactose ligand is rather low (K_D in the low millimolar range), bi- and triantennary oligosaccharides logarithmically increase the affinity to low micromolar and nanomolar dissociation constants, respectively (*Fig*. 5-6B). Adding a forth galactose only slightly increases the affinity [2, 21]. Many different linkers have been synthesized in order to optimize the binding geometry and affinity for liver-targeting approaches (example in *Fig*. 5-6C).

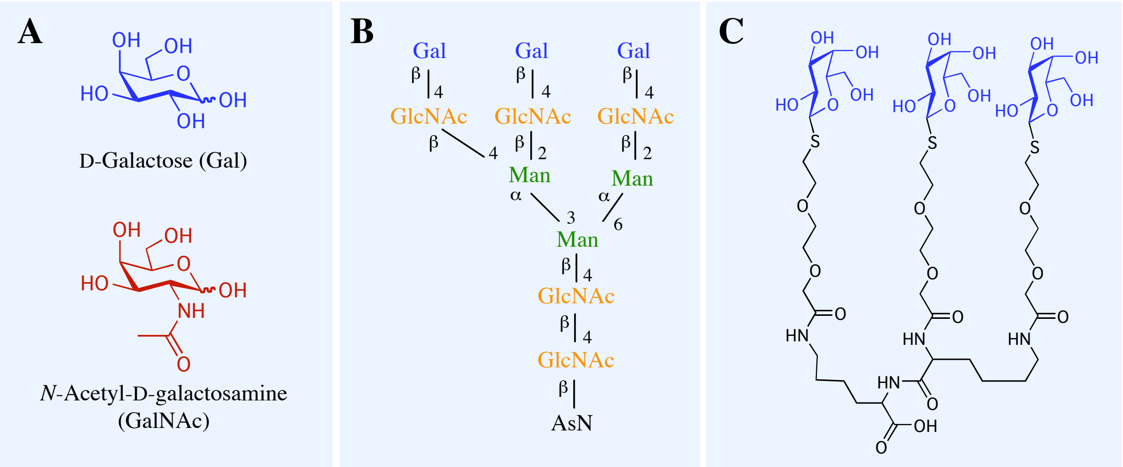

A
D-Galactose (Gal)
N-Acetyl-D-galactosamine
(GalNAc)
B
Gal
GlcNAc
Man
AsN
C
S
O
HN
OH

Figure 5-6: Carbohydrate ligands for the ASGP receptor. **A:** Monosaccharide ligands galactose and *N*-acetylgalactosamine. **B:** Natural triantennary ligand TRI-GP (from Lee *et al.* [21]). **C:** Example of a synthetic triantennary ligand (Gal_3Lys_2-II from Kichler *et al.* [22]).

The observation that the change to triantennary oligosaccharides generated the most remarkable increase in binding affinity, led to more hypothesis about a possible receptor and binding geometry (*Fig*. 5-7). Antibody cross-linking studies and lateral diffusion measurements suggested a heterohexa- (*Fig*. 5-7B) or pentameric geometry (*Fig*. 5-7C) [12], while oxidative cross-linking of H1 and H2 peptides with terminal cysteine residues resulted in homotrimers (H1), homotetramers (H2) and 2:2 heterotetramers (H1/H2 mixtures) [23]. Investigations with a triantennary glycopeptide on rat ASGP-R suggested a triangular arrangement of the three galactose residues with distances of 15, 22 and 25 Å [24, 25] (*Fig*. 5-7A). Further studies showed that two galactose moieties specifically interact with HL-1 subunits and the third galactose with HL-2 [26].

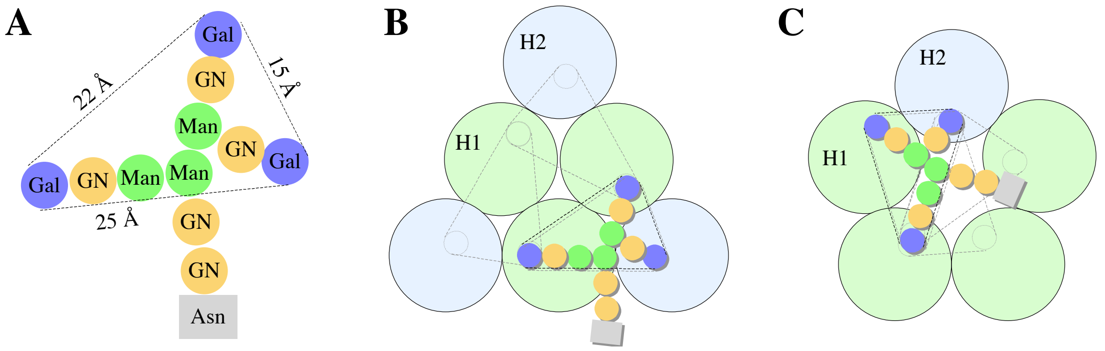


Figure 5-7: Possible binding geometries of triantennary ASGP-R ligands. **A:** Dimensions of an optimal triantennary oligosaccharide (GN = *N*-Acetyl glucosamine). **B, C:** Hypothetical models for the arrangement of the hetero-oligomeric receptor complex (all images adapted from Geffen and Spiess [12]).

As it is typical for C-type lectins, binding to the H1-CRD of the receptor shows an absolute requirement for Ca^{2+}[16]. Two to three calcium ions were found to be bound per CRD in the case of rabbit ASGP-R, with equilibrium dissociation constants of 0.3-1 mM for the individual calcium ions [27]. Since the bound ligand has to be released after fusion with hepatic endosomes, the binding affinity is also strongly *pH*-dependent. Conformational changes within the CRD lead to the release of both the ligand and the calcium ions when lowering the *pH* below 6.5 [28].

Some of the best-characterized ligands are the desialylated glycoproteins asialoorosomucoid (ASOR) and asialofetuin (ASF) with five and three *N*-linked glycans, respectively [17]. ASOR shows equilibrium dissociation constants of ~ 40 nM for cells expressing only H1 and ~ 10 nM for the functional receptor [29]. Even though binding modes and affinities for many natural ligands and multivalent Gal or Gal*N*Ac conjugates with synthetic spacers have been investigated, little is known about the potential of carbohydrate mimics and no SPR-based binding data are available.

5.1.5 Aims in this project

The primary aim in this project is the development of a Biacore assay for the characterization of the molecular interaction between ASGP-R H1-CRD and their natural (and synthetic) ligands. For this purpose, the expression and purification of the protein is optimized and the product is characterized by different methods. H1-CRD is then immobilized on the sensor chip and the assay conditions have to be optimized. Different known physiological binding properties such as the C-type lectin specific calcium-dependency or the *pH*-dependent release of ligands are validated using the molecular protein assay. Finally, the assay is validated with a panel of different ligands (glycoproteins, mono-, di- and oligosaccharides) and used to characterize a set of monoclonal anti-H1-CRD antibodies.

5.2 Materials and Methods

This section describes materials, equipment and procedures specifically used for the ASGP-R project. Materials and general methods used in all Biacore assays are described in *section* 2.2.

5.2.1 Materials

Reagents and proteins

Reversed-phase HPLC was performed with gradient-grade acetonitrile and trifluoroacetic acid (Riedel-de Haën). Gradient-grade HPLC water (Fluka) was used for weak anion exchange chromatography and freshly bidistilled water delivered through quartz tubes was used for the preparation of all other buffers and solutions. Neuraminidase agarose from *Clostridium perfringens* (EC 3.2.1.18, Type VI-A, Sigma N5254) was used for enzymatic desialylation of orosomucoid (α_1-acid glycoprotein from human plasma, Fluka 50646). Asialofetuin from fetal calf serum was commercially available as desialylated product (Type I, acidic desialylation, Sigma A4781). D-Galactose was from Senn Chemicals AG (Dielsdorf, Switzerland), *N*-acetyl-D-galactosamine from Acros (Chemie Brunschwig AG, Basel, Switzerland), D-galactosamine from New Zealand Pharmaceuticals (Palmerston North, New Zealand), and methyl α-D-galactopyranoside from Sigma. D-Lactose, D-glucose, and methyl β-D-galactopyranoside were gifts from various sources. Polyacrylamide-type glycoconjugates containing 20 mol% of either β-D-glucose or β-D-Gal*N*Ac and 5% biotin (No. 0022-BP and 0031-BP; Lectinity Holdings Inc., Moscow, Russia) was used for both Biacore experiments and the solid-phase competition assay. Carboxymethyl dextran (Fluka 86524, sodium salt, 12 kDa, 1.5 mmol carboxyl groups/g) as well as all other reagents from Sigma, Riedel-de Haën and Fluka were from Fluka AG, Buchs, Switzerland.

Equipment

HPLC separations were done on an Agilent 1100 purification system, equipped with a quaternary pump, a cooled well-plate autosampler, a column thermostat, a DAD detector, and a cooled analytical fraction collector (Agilent AG, Basel, Switzerland). Size exclusion chromatography was performed on a TSK Gel G2000SW column (7.5 × 600 mm; Tosoh Biosciences GmbH, Stuttgart, Germany). Two identical columns (polymer-based DEAE columns, 7.5 × 80 mm) from different brands, i.e. Phenomenex

Biosep-DEAE-P and Shodex IEC DEAE-825 (both from Brechbühler AG, Schlieren, Switzerland), were used for weak anion exchange separation. For affinity chromatography, empty Bioscale MT-2 columns (2 ml volume; Bio-Rad Laboratories AG, Reinach, Switzerland) were filled with differently derivatized sepharose. Reversed-phase HPLC was performed on a SpectraSystem HPLC system consisting of a vacuum degasser, an autosampler, a binary pump and a UV detector (Thermo Separation) using Vydac 214TP54 columns (RP-C4. 4.6 × 150 mm).

5.2.2 Expression and purification of recombinant H1-CRD (PhD thesis of R. Born [30])

Human hepatic asialoglycoprotein receptor H1-CRD was produced recombinantly based on a published method [13, 31]. The method was further optimized by Rita Born [30], in order to increase both yield and purity. Briefly, a truncated form of the H1 subunit of the ASGP receptor including the whole CRD domain (amino acid residues 147-291) was expressed in *E. coli*. After lysis and solubilization of the inclusion bodies, the protein was refolded by dialysis and affinity-purified on a galactose sepharose column. A schematic overview of the major expression and purification steps is visualized in *figure* 5-8A&B.

5.2.3 Separation of H1-CRD monomers and dimers

Since the H1-CRD fractions after expression and purification (see *section* 5.2.2) always contained a variable amount of a dimeric species as well as some minor impurities, different HPLC methods for a further separation and purification of H1-CRD were developed. Four different chromatographic principles (size exclusion, reversed-phase, weak anion exchange and affinity chromatography) were applied and compared,

Separation according on the molecular weight of the monomer and dimer fractions was performed with *size exclusion* chromatography. Unprocessed H1-CRD samples were injected on the column equilibrated with 10 mM HEPES *pH* 7.4 and separated under isocratic conditions at a flow rate of 1 ml/min and ambient temperature. Specific lectin interactions with the silica-based column material was tested by adding 2 mM $CaCl_2$ to the running buffer.

In a *reversed-phase* separation approach, a water-to-acetonitrile gradient was used to isolate H1-CRD monomers from dimers. 0.1% trifluoroacetic acid (TFA) was added to all solvents to improve separation and peak shapes. After equilibration of the C4 RP-column in water/0.1% TFA, the sample was injected and a 10 min linear gradient from 0 to 40% acetonitrile (ACN) in 0.1% TFA was performed, followed by a shallow gradient from 40 to 52.5% ACN/TFA within 10 min. Finally, the gradient was raised to 95% ACN/TFA and the column was washed for 5 min at these conditions before returning to 100% water/TFA. Collected fractions were analyzed using SDS-PAGE adapted from the method by Laemmli [32] (15% SDS-PAGE, non-reducing conditions, silver staining).

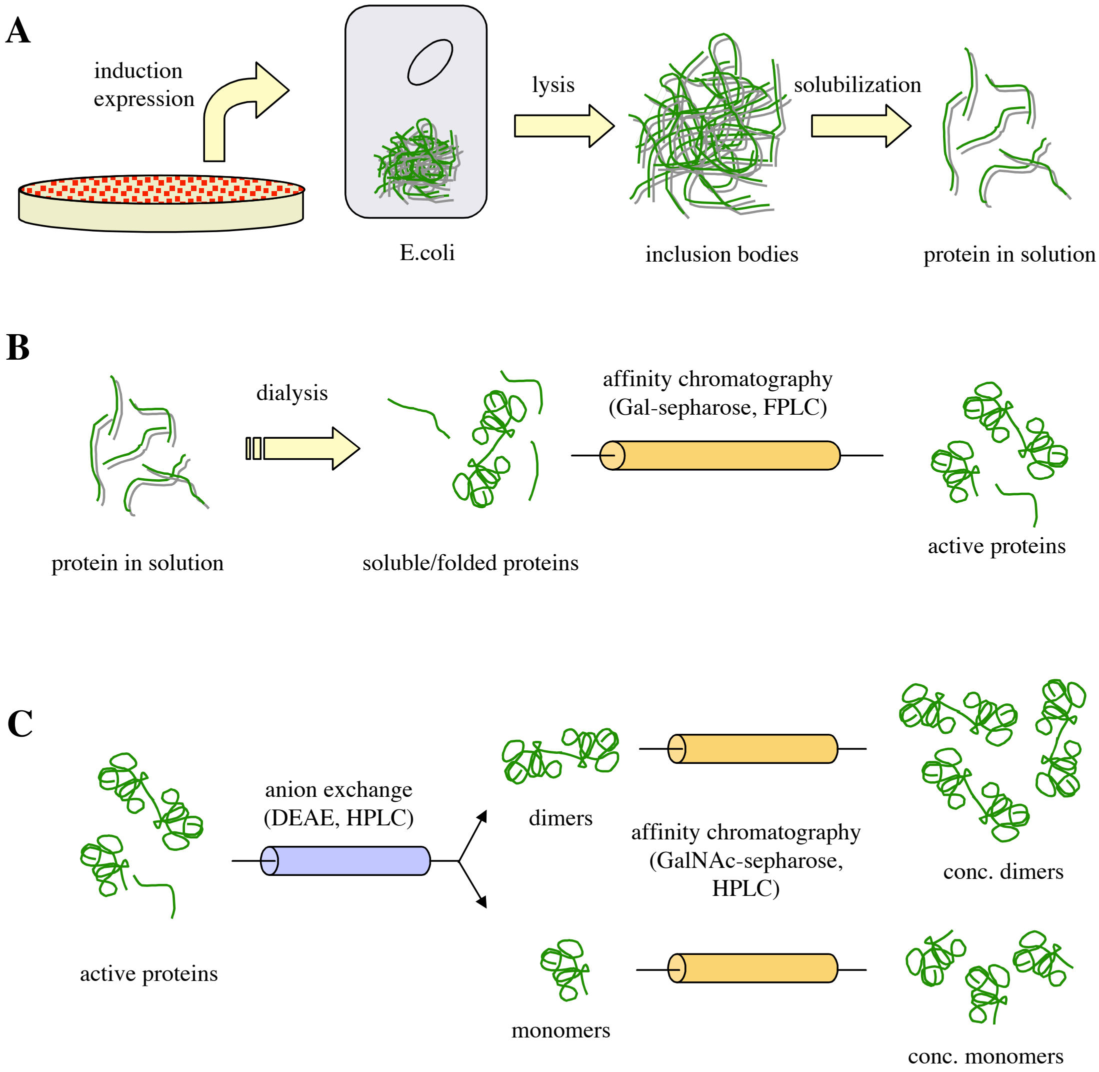

A
induction
expression
lysis
solubilization
E.coli
inclusion bodies
protein in solution
B
dialysis
affinity chromatography
(Gal-sepharose, FPLC)
protein in solution
soluble/folded proteins
active proteins
C
anion exchange
(DEAE, HPLC)
dimers
affinity chromatography
(GalNAc-sepharose,
HPLC)
conc. dimers
active proteins
monomers
conc. monomers

Figure 5-8: Simplified overview of expression (A), primary purification (B) and monomer/dimer separation (C) in the production cycle of H1-CRD.

A polymer-based DEAE column was chosen for *weak anion exchange* (*Fig*. 5-8C) in order to have a higher flexibility for elevated *pH* conditions. Different ionic strength, running buffer *pH* and temperatures were evaluated. Finally, 25 mM Tris-HCl *pH* 8.0 (running buffer A) was identified as an ideal starting condition. Running buffer B was prepared by adding 250 mM calcium chloride to running buffer A. In order to ensure a maximum resolution, the column was washed with a cycle of water, 0.1N NaOH, water, and 100% B for at least one hour each at a reduced flow rate of 0.1-0.2 ml/min. After these washing and regeneration steps, the column was extensively equilibrated with running buffer at 15% B for at least 8 hours (overnight). Since purified H1-CRD samples already contained 120 mM NaCl (see *section* 5.2.2), the high salt load had to be reduced prior to injection on the DEAE column. Therefore, all H1-CRD samples underwent a buffer exchange on a HiTrap desalting column into running buffer A. The separation of monomers was performed at 20°C and a flow rate of 0.5 ml/min. After injection, the sample was washed for 2 min at 15% B and the linear gradient was raised from 15% to 35% B within 25 min. All peaks were detected at 280 nm and collected for further analysis. In case of automated separation cycles, both the autosampler and the fraction collector were kept at 10-15 °C in order to prevent the samples from evaporation and (proteolytic) degradation.

For HPLC-based *affinity chromatography*, galactose or Gal*N*Ac were coupled to sepharose using the divinylsulfone method [33]. 20 ml of packed sepharose 6B were washed three times with 20 ml water and centrifuged for 5 min at 3000 rpm. After suspension in 20 ml 0.5 M carbonate buffer *pH* 11.0, the sepharose was activated by addition of 2 ml divinylsulfone and incubation for 1 hour under constant shaking at 1000 rpm. The activated material was extensively washed in a glass frit with 2 l water and split up into three parts. While two fractions were resuspended in 1 volume of 0.5 M carbonate buffer *pH* 10.0 containing either 20% (m/v) galactose or 20% (m/v) Gal*N*Ac, the third fraction served as blank control and was suspended in pure carbonate buffer. The coupling process was performed overnight under constant shaking at 300 rpm. Following one centrifugation (5000 rpm for 5 min) and three wash steps with water, the fractions were suspended in 1 volume part 0.5 M carbonate buffer *pH* 8.5 and blocked with 0.02 volume parts of 2-mercaptoethanol (shaking for 2 hours at 300 rpm). The fraction were washed three times with a tenfold volume of water and packed into empty columns at a flow rate of 2 ml/min. After connection to the instrument, the affinity columns were properly equilibrated with running buffer (10 mM HEPES *pH* 7.4, 10 mM $CaCl_2$) for at least one hour. For purification and affinity tests, samples were injected at a flow rate of 1 ml/min and kept at isocratic condition for 5 min in order to wash away any inactive proteins and other impurities. A linear gradient to an elution buffer (10 mM HEPES *pH* 7.4, 2 mM EDTA) was performed for 10 min with a following plateau phase of 2 min at 100% elution conditions. Blank runs were subtracted from the sample injections in order to increase the sensitivity of the method. Samples before and after anion exchange (see above) were tested for any loss in activity. For this purpose, constant volumes of the unprocessed sample and the monomer fraction of the same batch after DEAE separation were compared. Denatured H1-CRD was used as a negative control. The specificity of the interaction was further evaluated by injecting the same sample on all three columns (Gal-, Gal*N*Ac, and blank sepharose). When affinity chromatography was used to concentrate samples after DEAE separation (*Fig*. 5-8C), multiple injections were performed under isocratic running buffer conditions, before a single elution gradient was started. Collected fractions were analyzed using SDS-PAGE as described earlier.

5.2.4 Characterization of H1-CRD by mass spectrometry

H1-CRD samples were either desalted by reversed-phase HPLC (see *section* 5.3.3) or by elution from ZipTips. Protein mass analysis was performed using ESI or MALDI technology before and after reduction (DTT) and alkylation (iodoacetamide) of the proteins. For the sequence confirmation experiments, H1-CRD monomer was digested using trypsin or LysC protease and analyzed by ESI-MS. All mass spectrometry experiments were performed by Thierry Mini in the laboratory of Prof. Paul Jenö (Biocenter, University of Basel, Switzerland).

3.2.5 Development and optimization of a H1-CRD Biacore assay

Immobilization of H1-CRD

The distribution and accessibility of potential sites for immobilization on the H1-CRD molecule was investigated using crystal structure data (PDB code 1DV8) [13]. Six residues of the *N*-terminus, which were not represented in the crystal structure (147-152), were directly added in *PyMol*. The accessible areas of the protein were visualized by a Connolly surface model [34], in which the lysine residues, the *N*-terminus, Cys-152, as well as the residues involved in carbohydrate binding were highlighted in different colors (*Fig*. 5-22, see *section* 5.3.4).

Separated H1-CRD monomer and dimer fractions were concentrated and buffer-exchanged using affinity chromatography on HPLC (see *section* 5.2.3). Since the used sample buffer (10 mM HEPES *pH* 7.4, 2 mM EDTA) did not contain any primary amines, those samples could be directly used for immobilization. Monomers and dimers were immobilized on separate flow cells of the same CM5 sensor chip using standard amine coupling procedure (see *section* 2.2.5). All samples were diluted in 10 mM acetate buffer *pH* 4.5 to a final concentration of approximately 20 μg/ml. After activating the surface for 5 to 10 min at at flow rate of 5 μl/min, the sample was injected for 5 to 15 min and the surface was deactivated for a time corresponding to the activation phase. Typical surface densities ranged from 1800 to 2500 RU for the monomer and 1,900 to 3,000 RU for the dimer.

Alternatively, a thiol immobilization method was developed based on the procedure used for human serum albumin (see *section* 3.2.13). First, a thiol-functionalized surface was prepared (see *section* 2.2.7). Usually, 20 μl H1-CRD monomer (~ 200 μg/ml) were mixed with 5 μl each of 10 mM GSH and 10 mM DTNB. After 2 hours of incubation at ambient temperature, 200 μl 10 mM sodium acetate buffer *pH* 4.0 were added and the sample was incubated for another 2 hours. Finally, the sample was injected for 60 to 120 minutes at a flow rate of 2 μl/min and washed with three consecutive 30 s pulses of 0.1% SDS at a flow rate of 50 μl/min. Surface densities up to 4,000 RU could be obtained with this method. In contrast to the thiol immobilization procedure for albumin, the modified method for ASGP-R was far less reproducible and concentrations as well as incubation times had to be slightly adapted for each immobilization.

Buffer composition

A buffer system based on the Biacore HBS-N buffer (10 mM HEPES *pH* 7.4, 150 mM NaCl) was used as running buffer. The optimal calcium concentration was determined by injecting a constant concentration of asialofetuin (2.5 μM ASF) in running buffer containing 2, 20, 50, or 100 mM $CaCl_2$ and the signal intensity was evaluated. The affinity of H1-CRD for calcium ions was further investigated by injecting a twofold serial dilution of $CaCl_2$ between 30 μM and 50 mM in calcium-free running buffer. Steady state data were fitted to a single site binding model as well as to a two independent sites model. In order to identify the essential components of the running buffer, Gal*N*Ac was screened as described in *section* 5.2.6 in four different buffer compositions (*Table* 5-1).

Table 5-1: Buffer compositions for the identification of the components essential for ligand binding.

System	Buffer	*pH*	NaCl	$CaCl_2$
1	10 mM HEPES	7.4	150 mM	50 mM
2	10 mM HEPES	7.4	150 mM	5 mM
3	10 mM imidazole	7.4	150 mM	50 mM
4	10 mM HEPES	7.4	no	50 mM

A hypothetical binding of HEPES to the carbohydrate binding site was further investigated with docking studies by Dr. Markus Lill in the group of Dr. Angelo Vedani (Biographics Laboratories 3R, Basel, Switzerland).

Influence of pH

The influence of buffer acidity was studied by screening galactose and Gal*N*Ac (see *section* 5.2.6) in buffer system 4 (*Table* 5-1) at *pH* 7.4, 7.0, 6.5, 6.0, and 5.5. Both the signal intensity and the binding affinity were calculated and compared. Since some *pH* values were below the buffer capacity range of HEPES, the buffer was carefully controlled before and after the experiment. In addition, the *pH* of the buffer at *pH* 5.5 was measured after addition of 10 mM galactose in order to detect any analyte-induced shifts in *pH*.

In order to investigate a potential release of calcium ions at lower *pH* values, a series of HEPES buffer blanks was injected over immobilized H1-CRD monomer. The buffers ranged from *pH* 4.0 to 7.5 in steps of 0.25 units and contained either 50 mM $CaCl_2$ or 100 mM NaCl. All solution were injected in triplicates for 1 min at a flow rate of 50 μl/min. An empty flow cell served as a reference surface. Equilibrium data of the induced SPR signals were averaged and plotted against the buffer *pH*.

DMSO tolerability

Galactose and Gal*N*Ac were screened (see *section* 5.2.6) in buffer system 4 (*Table* 5-1), which was either DMSO-free or contained 5% DMSO. In case of DMSO addition, a calibration procedure as described in *section* 3.2.5 was performed. Differences in the binding affinity for the two ligands between the buffer systems were evaluated.

5.2.6 H1-CRD ligand binding

The newly developed Biacore assay was validated by screening of some known natural ligands. For this purpose, two different asialoglycoproteins as well as small mono- and disaccharides were analyzed by high-resolution screening and the values were compared with a solid-phase competition assay, which was developed by Daniela Stokmaier in our laboratory [35]. Finally, the Biacore assay was also used to characterize a panel of monoclonal anti-H1-CRD antibodies. Data processing of all Biacore data as well as evaluation of equilibrium data was done in *Scrubber* while kinetic analyzes were performed in *CLAMP* (see *section* 2.2.1).

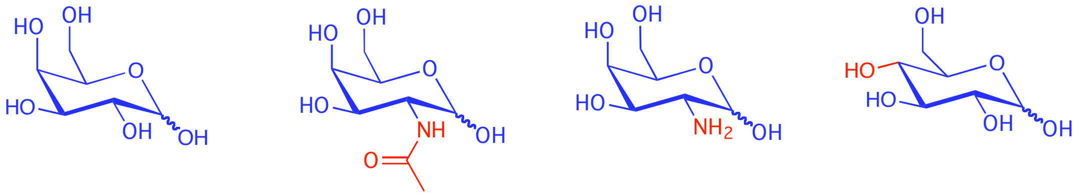
Asialoglycoproteins

While asialofetuin (ASF) was commercially available, no desialylated derivative of orosomucoid (OSM; i.e. α_1-acid glycoprotein) is commercially available. Therefore, orosomucoid had to be desialylated before screening. In principle, two methods are available for this purpose: the acidic desialylation using sulfuric acid [36] and the enzymatic cleavage of sialic acid moieties by a neuraminidase [37]. Both methods were performed and compared with each other. For the acidic desialylation, the glycoprotein was dissolved in water to a concentration of 1 mg/ml and acidified by adding an equal volume of 0.1 N H_2SO_4 to a final concentration of 0.5 mg/ml OSM in 0.05 N sulfuric acid. After incubation for 1 hour at 80 °C and constant shaking at 500 rpm, the solution was neutralized by adding 1N NaOH. An OSM blank was prepared by substituting the acid by pure water. In the enzymatic desialylation approach, the protein was dissolved in reaction buffer (0.1 M sodium acetate *pH* 5.0, 2 mM $CaCl_2$) to a concentration of 12.5 mg/ml. 1 ml of the OMS solution was mixed with 300 μl of the neuraminidase-agarose suspension (0.125 U) and incubated for 4 hours at 37 °C under shaking at 800 rpm. After centrifugation for 5 min at 13,000 rpm, the ASOR-containing supernatant was collected and the enzyme was washed twice with reaction buffer. Since the *pI* of orosomucoid is reported to increase from 2.7 to 5.0 after removal of its sialic acids [37], a weak anion exchange method was developed to separate ASOR from OSM. For an in-process monitoring, small samples of the desialylation mixture were taken immediately after enzyme addition and after 1, 2, 3, and 4 hours of incubation. All samples were injected to a DEAE column equilibrated at 25 mM HEPES *pH* 7.4, 2 mM $CaCl_2$, 12.5 mM NaCl at a flow rate of 0.5 ml/min at 20 °C. After 2 min at starting conditions a linear NaCl gradient was applied to 175 mM NaCl within 15 min. OSM and ASOR samples of the acidic desialylation were analyzed using the same method.

Both asialoglycoproteins were dissolved in running buffer (10 mM HEPES *pH* 7.4, 50 mM $CaCl_2$) to a concentration of 25 μM (with assumed molecular weights of 4 kDa for ASF [36] and 40 kDa for ASOR [37]). A threefold dilution series between 1000 and 4 nM was screened against all three H1-CRD surfaces described in *section* 5.2.5 (monomer, dimer, thiol). In order to avoid non-specific binding and carry-over effects, 0.005% polysorbate-20 were added to the running buffer. The samples were injected for 15 min with a subsequent dissociation phase of 15 min at a flow rate of 20 μl/min. The surface was regenerated by a 2 min pulse of HBS-EP buffer (10 mM HEPES *pH* 7.4, 150 mM NaCl, 3 mM EDTA, 0.005% polysorbate-20) and two short pulses (30 s) of 0.1% SDS. After the regeneration procedure, the surface was equilibrated in running buffer for 15 min before injecting the next sample. Signals of an untreated flow cell were subtracted to avoid bulk effects (referencing). In addition, blank injections were performed before the experiment as well as between different samples and were subtracted from the glycoprotein sensorgrams (double referencing [38]).

Mono- and disaccharides

Stock solutions of the carbohydrate analytes (*Fig*. 5-9) were prepared by dissolving the compounds in running buffer (10 mM HEPES *pH* 7.4, 50 mM $CaCl_2$) to a concentration of 100 mM. High-resolution screening was performed by injecting twofold serial dilutions between 5 mM and 5 μM as randomized triplicates. Each sample was injected for 30 s with an undisturbed dissociation phase of 20 s using the instruments *kinject* command at a flow rate of 50 μl/min. No regeneration or washing steps had to be applied. Five buffer blanks were injected at the beginning and one between the triplicate series. Signals of an untreated flow cell and averaged blank injections were subtracted from the sample sensorgrams. Since referenced sensorgrams showed negative SPR signals (see *section* 5.3.5, *Fig*. 5-29), the whole data set had to be *mirrored* by multiplication of each data point with -1. Mirrored steady state data were evaluated between 10 and 20 s of the injection period and were fitted to a single site binding model. In addition, the sensorgrams were also evaluated kinetically using a simple 1:1 Langmuir binding isotherm [39]. Binding to the different H1-CRD surfaces (monomer, dimer, thiol; see *section* 5.2.4) was evaluated and compared.

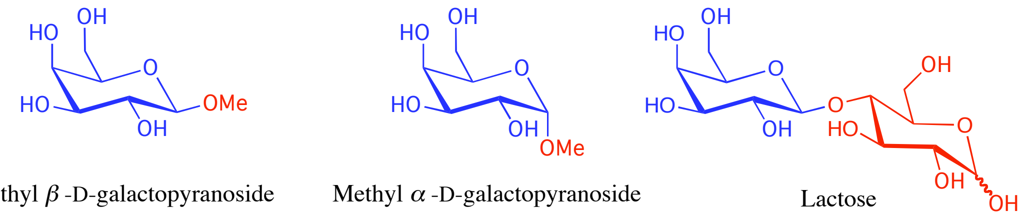


D-Galactose
(180.2 Da)

N-Acetyl-D-galactosamine
(Gal*N*Ac; 221.2 Da)

D-Galactosamine
(179.2 Da)

D-Glucose
(180.2 Da)



Methyl β -D-galactopyranoside
(194.2 Da)

Methyl α -D-galactopyranoside
(194. 2 Da)

Lactose
(342.3 Da)

Figure 5-9: Mono- and disaccharides used for the validation of the H1-CRD Biacore assay. The galactose core structure is visualized in blue, modifications are highlighted in red.

Polyvalent, polymer-bound forms of Gal*N*Ac and glucose were investigated by injecting the biotinylated glycoconjugates used for the solid-phase competition assay over amine-coupled H1-CRD monomer and dimer surfaces under the same buffer conditions. Both polymers (10 μg/ml) were injected for 3 min followed by another 3 min of undisturbed dissociation at a flow rate of 20 μl/min. In order to test for competition effects, a 10 mM galactose solution was injected for 1 min and the (negative) signal intensity as well as the post-injection drop of the signal was evaluated. Finally, the surface was regenerated by a 1 min pulse of HBS-EP buffer (see above) and equilibrated in running buffer for 5 min.

Solid-phase competition assay

Biacore-derived K_D values were compared with a solid-phase competition assay, which was developed in a diploma thesis by Daniela Stokmaier [30, 35], based on a similar polymer-assay for E-selectin [40].

Monoclonal anti-H1-CRD antibodies

A set of six monoclonal mouse anti-human H1-CRD antibodies, which were produced with hybridoma technology and preselected using an ELISA [30], were ranked and screened with the Biacore assay. All antibodies were purified by affinity chromatography and were classified as mouse IgG1 [30]. In a first ranking assay, each antibody was injected at a concentration of 2 μg/ml for 5 min with a dissociation phase of 5 min. The assay was performed in 10 mM HEPES *pH* 7.4, 50 mM $CaCl_2$ at a flow rate of 10 μl/min. Potential blocking properties of the antibodies were tested by injecting a 1 min pulse of 5 mM galactose before mAb injection and after the dissociation phase. Calcium dependency was evaluated with an injection of HBS-EP buffer (see above) for 1 min. Finally, the surface was regenerated by a 12 s pulse of 10 mM HCl at a flow rate of 50 μl/min and was equilibrated in running buffer for 10 min. The same assay was repeated using HBS-EP as the running buffer. The maximum intensities at injection end were plotted against the remaining signal intensity after 5 min of dissociation (stability plot). In addition, blocking properties were evaluated by comparing the galactose signal before and after antibody injection.

The three antibodies with the most promising binding properties selected for kinetic screening and were prepared as twofold dilution series between 140 nM and 140 pM in HBS-EP as a running buffer. Samples were injected in a randomized order for 10 min at a flow rate of 25 μl/min. After a 10 min dissociation phase, the surface was regenerated by two 12 s injections of 10 mM HCl at a flow rate of 50 μl/min and was equilibrated in running buffer for 5 min. The full data sets (140 nM - 140 pM) as well as a reduced data set (9 nM - 140 pM) were kinetically evaluated in CLAMP by applying different interaction models.

5.2.6 Evaluation of negative binding signals

In order to evaluate a weak specific interaction of the immobilized lectin with the carboxymethyl dextran (CMD) hydrogel of the sensor chip, soluble CMD was added to the running buffer. For this purpose, galactose and Gal*N*Ac were screened (see *section* 5.2.5) in 10 mM HEPES *pH* 7.4, 50 mM $CaCl_2$ containing 1 mg/ml CMD, 10 mg/ml CMD, or no CMD addition. Both the signal intensities and binding affinities were compared.

5.3 Results & Discussion

5.3.1 Expression and purification of ASGP-R H1-CRD

The ability to test physiological and synthetic ligands on an isolated form of a receptor is a prerequisite for a deep understanding of molecular binding events. Since only the extracellular carbohydrate recognition domain of the human hepatic asialoglycoprotein receptor is involved in ligand binding and posttranslational glycosylation does not influence its activity, the lectin domain can be expressed as a truncated protein in *E. coli*. Even though a protocol for the expression and purification already existed [13], the yields were rather low and the protein fraction contained several populations. Therefore, the expression system as well as the purification protocols were further optimized. By selecting a different strain of *E. coli* (AD494 instead of JM109), the expression yield could be remarkably increased (*Fig.* 5-10A). The combination of cell lysis by sonication, solubilization of the inclusion bodies and refolding by dialysis followed by affinity purification on a galactose-sepharose column (*Fig.* 5-8A&B) generated a highly active protein fraction. However, non-reducing SDS-PAGE analysis always showed two protein bands at approximately 17 and 34 kDa, indicating a dimerization. In addition, the ratio between the two populations was found to highly variable between different production batches (*Fig.* 5-10B). Some additional bands around the monomer and dimer fractions were visible ('satellite bands', see *appendix* D3, lanes 4 and 5).

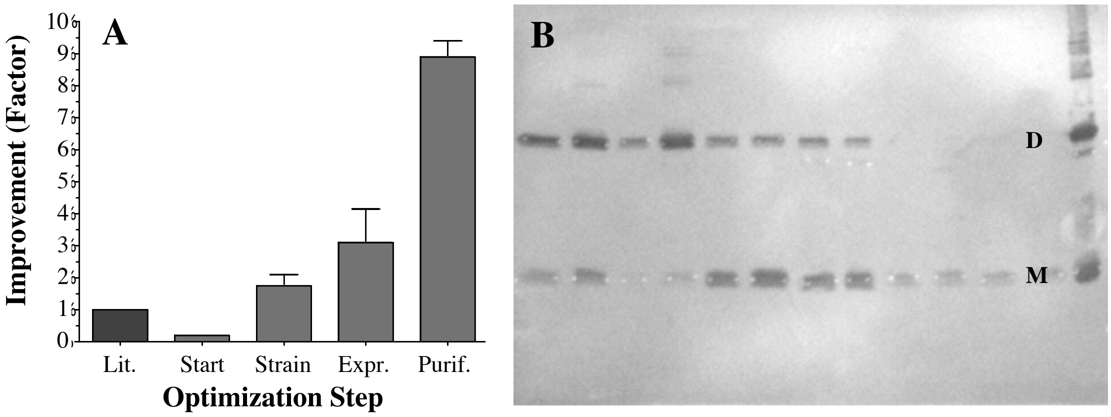


Figure 5-10: Optimized expression of ASGP-R H1-CRD. **A:** Comparison of the expression levels obtained by the existing protocol published in literature (Lit.; [31]) and after its application (Start), as well as after the optimization of the E. coli strain (Strain), the expression method (Expr.), and the purification protocol (Purif.). **B:** Western blot of different H1-CRD production batches, demonstrating the high variability in the monomer (M) to dimer (D) ratio (blotted from a non-reducing 15% SDS-PAGE, primary antibody = monoclonal anti-H1-CRD B01.4 (see *section* 5.3.5), secondary antibody = anti-mouse Fc, HRP-conjugate). All images are adapted from Born [30].

5.3.2 Separation of H1-CRD monomers and dimers

In contrast to the purification of the active protein by separating it from other cellular impurities, the separation of monomer and dimer populations was found to be more complicated. While the major difference of the two proteins is their size, only little variations are expected in case of their net charge or affinity. In order to get maximum resolution and short analysis/separation times, high-performance liquid chromatography (HPLC) was used for the evaluation of different separation methods.

Size exclusion HPLC

Since the H1-CRD dimer should have twice the molecular mass of the monomer, a separation by size seems to be the method of choice. Unfortunately, the absolute masses and therefore also the mass difference are still relatively low for a successful separation on size exclusion columns. The separation is mainly dependent on the material and the length of the column, while parameters such as the column diameter, flow rate, or solvent properties are less important [41]. Even though a rather large column (60 cm) was used for the separation, both fractions appeared as a single peak. In some cases, a small shoulder of the monomer peak was visible, which might contain H1-CRD dimer. For a complete baseline separation, extremely large columns or another column material (e.g. pore size) was required.

In addition, the lectin was found to interact with the column material in a specific manner. When calcium-containing buffers were used for the separation, the protein remained on the column and only eluted after a switch to calcium-free buffers (Fig. 5-11). This phenomenon could be explained by a calcium-dependent interaction of the CRD with hydroxyl-groups of the silica-based column material. Addition of EDTA to the running buffer as well as an increase of total ion concentration (e.g. by adding NaCl) could reduce such interferences. However, highly ionic solutions are not feasible for Biacore immobilization, since they suppress preconcentration of the protein on the sensor chip surface (see section 2.1.4). Due to its low resolution and the undesired interaction properties, the size exclusion method was regarded as not suitable for the separation of H1-CRD monomers and dimers.

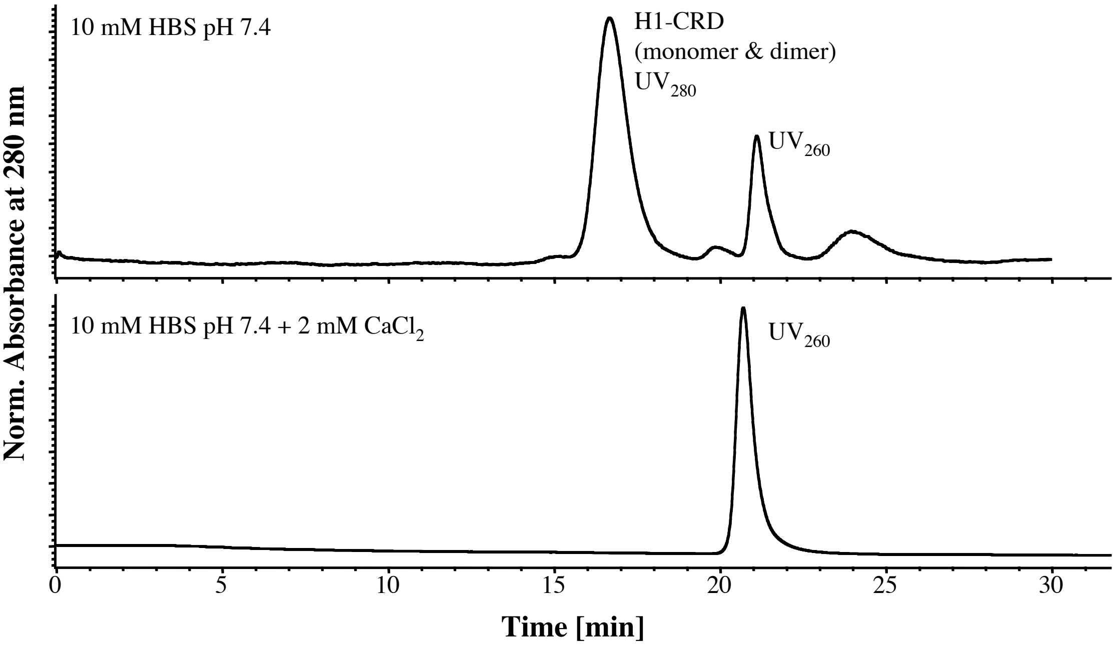


Figure 5-11: Size exclusion chromatography of H1-CRD in HBS buffer with and without calcium addition. Peaks with an UV maximum correspond to a protein fraction (H1-CRD), peaks with maxima at 260 nm are likely to salts or other impurities.

Reversed-phase HPLC

In contrast to the reversed-phase separation of small molecules, where the molecules constantly interacts with the material, separation of proteins follows a different principle. Under hydrophilic starting conditions the protein tightly binds to column with its most lipophilic part ('hydrophobic foot') and is completely eluted above a certain concentration of organic solvent [42]. Therefore, a successful separation of H1-CRD monomer and dimer proteins would require that the dimerization influences the hydrophobic foot region. Indeed, it was possible to develop a shallow water/acetonitrile gradient, which resulted in a baseline separation of the two fractions (*Fig.* 5-12A). The resulting fractions were analyzed by SDS-PAGE and confirmed the isolation of both populations (*Fig.* 5-12B).

Unfortunately, the separation range was found to be rather narrow and highly dependent on environmental parameters such as the temperature or the sample matrix (salts etc.). Therefore, the reproducibility was rather low and led to unpredictable results, especially when changing batches or injection volumes. Furthermore, the addition of organic solvent (acetonitrile) and the required acidity (0.1% TFA) might decrease the proteins activity and was therefore not recommended for the large-scale production of functional protein. This method was therefore only used for analytical purposes, e.g. the preparation of fully desalted samples for mass spectrometry.

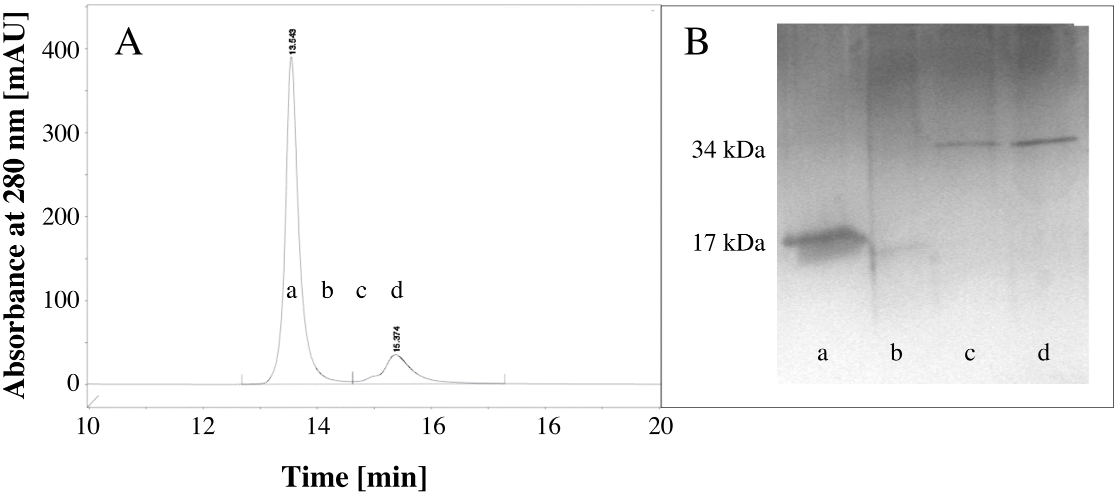


Figure 5-12: Reversed-phase separation of H1-CRD monomer and dimer fractions. **A:** Chromatogram showing two clearly separated peaks. **B:** SDS-PAGE analysis of the collected fractions (a-d) of chromatogram A (15% gel, non-reducing, silver staining).

Weak anion exchange HPLC (DEAE column)

Anion exchange separates different species according to their surface charge. While completely different proteins with similar charge distributions might elute in a very similar manner, even single residue mutations in the same protein can be separated in other cases [43]. A polymer-based DEAE column equilibrated with Tris buffer was used for the separation of H1-CRD. The polymer material has the advantage of being stable in a much wider *pH* range than silica-based columns are. The *pH* optimum for a proper separation was found to be at *pH* 8.0. This is higher than expected, since values just one unit above the protein's *pI* (i.e. approximately *pH* 6 for H1-CRD) are usually sufficient for weak anion exchange. However, since the aim was not to separate H1-CRD from structurally different impurities but to separate closely related oligomeric forms, a higher *pH* seems to be necessary to gain additional resolution. Higher *pH* values had been avoided to exclude a possible disruption of intramolecular disulfide bridges. Calcium chloride was chosen for the gradient elution due to its high chloride activity and since calcium containing buffers were expected to retain the proteins activity after elution. Reproducible separations of H1-CRD monomers and dimers in less than 30 min could be established with a linear gradient from 37.5 to 100 mM $CaCl_2$. Different batches of H1-CRD were separated and the collected peak fractions were analyzed by SDS-PAGE (*Fig.* 5-13).

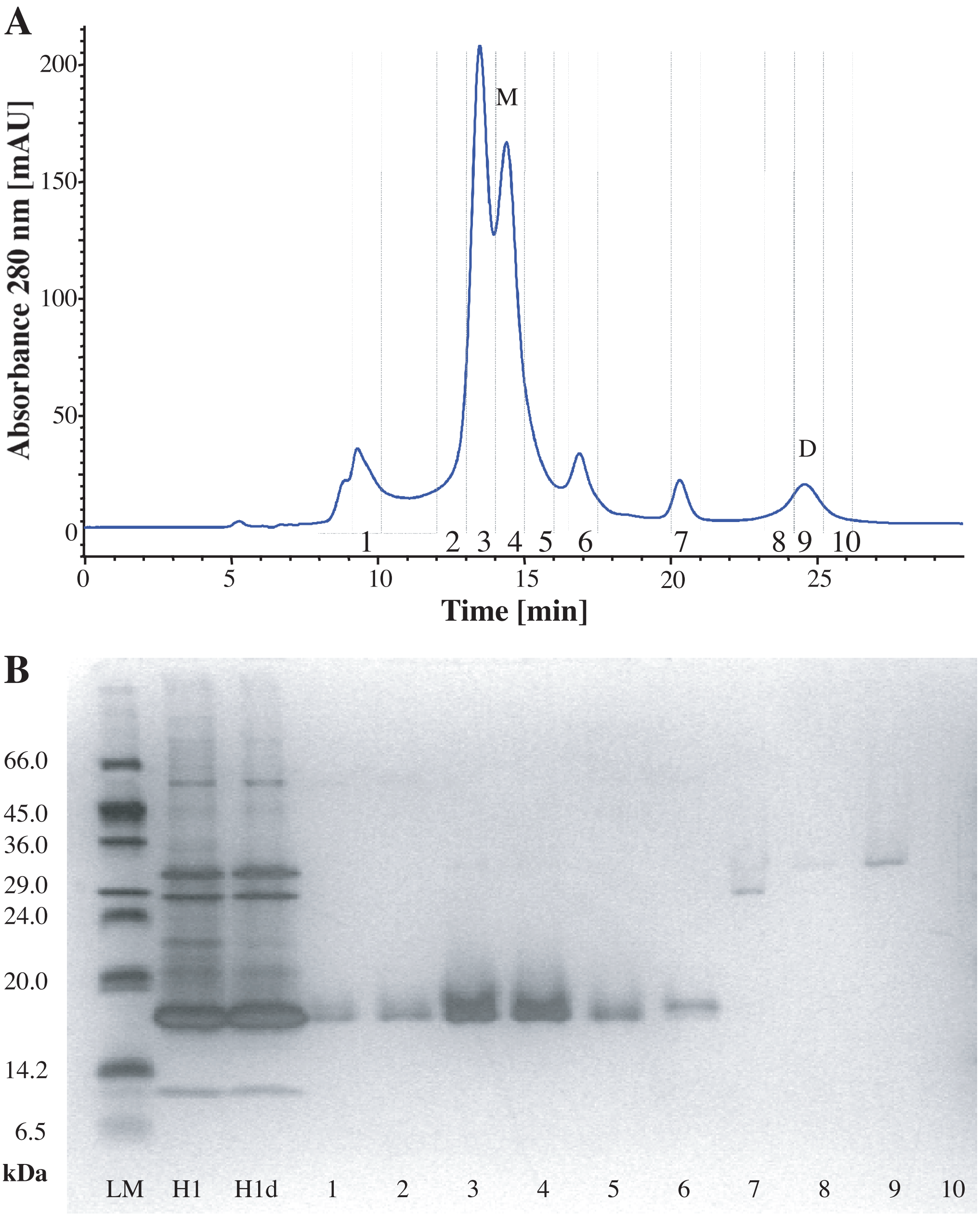


Figure 5-13: DEAE separation of H1-CRD monomers and dimers. **A**: Example of a chromatogram for the separation on a polymer-based DEAE column. Monomer fractions (M) can be clearly separated from the dimer (D). Collected fractions (1-10) were analyzed using SDS-PAGE (15% gel, non-reducing conditions, silver staining) and compared to the unprocessed sample before (H1) and after desalting (H1d) as well as to a low molecular weight marker (LM; **B**).

Typical chromatograms obtained by H1-CRD injections showed one peak group between 12 and 16 min. While some batches showed a single peak, other groups consisted of two or even three peaks or shoulders, which were all identified as H1-CRD monomers by SDS page. An additional peak at 16.5 min might also be monomeric CRD, but had a slightly increased molecular weight on SDS-PAGE. A single peak at

20.5 min corresponded to a satellite band below the dimer on SDS-PAGE but could not be identified by mass spectrometry. Finally, H1-CRD dimer eluted at 24 min as a single peak. The DEAE separation method therefore proofed to be suitable for isolating pure H1-CRD monomer and dimer fraction as well as for removing minor impurities.

The reason for the heterogeneity in the monomeric region is not yet clear and possible causes include folding or disulfide isomerization, proteolytic degradation during production and storage, or mutations. Since the uneven number of cysteine residue seems to promote the formation of mixed disulfides (see *section* 5.3.3), these small heterogeneities in retention on HPLC and SDS could be explained by different thiol adducts, which alter surface charge and molecular weight of the protein. The slightly increased *pH* value of 7.8-8.0 during storage and separation could further promote disulfide-shuffling reactions. Since DEAE is able to separate monomeric and dimeric forms of the same protein, the region around the dimerization site (e.g. free cysteine residues) seems to directly influence the interaction of the protein with the column resin. Similar effects are also described in literature. For example, DEAE anion exchange chromatography was used for the separation of different oxidation states of human serum albumin, which also only differ in the state of one cysteine residue (i.e. free Cys, mixed disulfides, and oxidized Cys) [44, 45]. The matrix of the polymer-based DEAE column was also identified to be crucial for a successful separation in the case of HSA [44].

While the relative peak intensities varied significantly between batches, they were very consistent within one batch (*Fig*. 5-14A). In contrast to RP-HPLC, the retention times changed only slightly when increasing the injection volume. Small shifts in retention times (± 0.5 min) were sometimes observed upon buffer changes, most likely due to small differences in buffer *pH* or calcium concentrations. However, properly equilibrated systems showed a very high reproducibility (*Fig*. 5-14B). Desalting was found to be a crucial step in the process. Since high NaCl concentrations (500 mM) were used during refolding and FPLC separation, direct injection of those samples on the DEAE column essentially inhibited binding of the protein to the column and a large fraction eluted already after the void volume. The desalting step on a HiTrap desalting column did not change the composition of H1-CRD samples (*Fig*. 5-13, lanes H1 and H1d) nor did it influence retention times during DEAE separation. Water quality was found to be crucial for this type of chromatography and only bidistilled or gradient-grade HPLC water could be used. Deionized and filtered water always generated a significant background signal, which often covered dimer peaks.

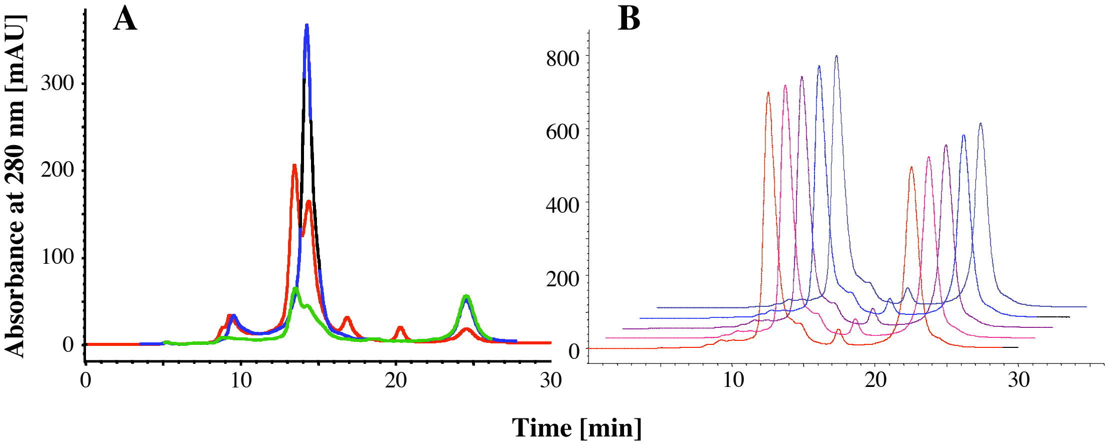


Figure 5-14: Reproducibility of H1-CRD separation. **A:** Different batches during production development showing varying concentrations and impurities but very reproducible retention times. **B:** Repetitive injections of the same batch leads to absolute reproducibility of intensities and retention times.

Weak anion exchange separation was found to be best suited for the analysis and production of pure H1-CRD monomers and dimers. By extending the gradient down to 12.5 mM $CaCl_2$ it was even possible to fully separate H1-CRD monomers and dimers from H2-CRD monomers and dimers [46] using a single column/buffer system (*Fig*. 5-15). This expands the analytical applications of this method while avoiding possible cross-contaminations of the two subunits.

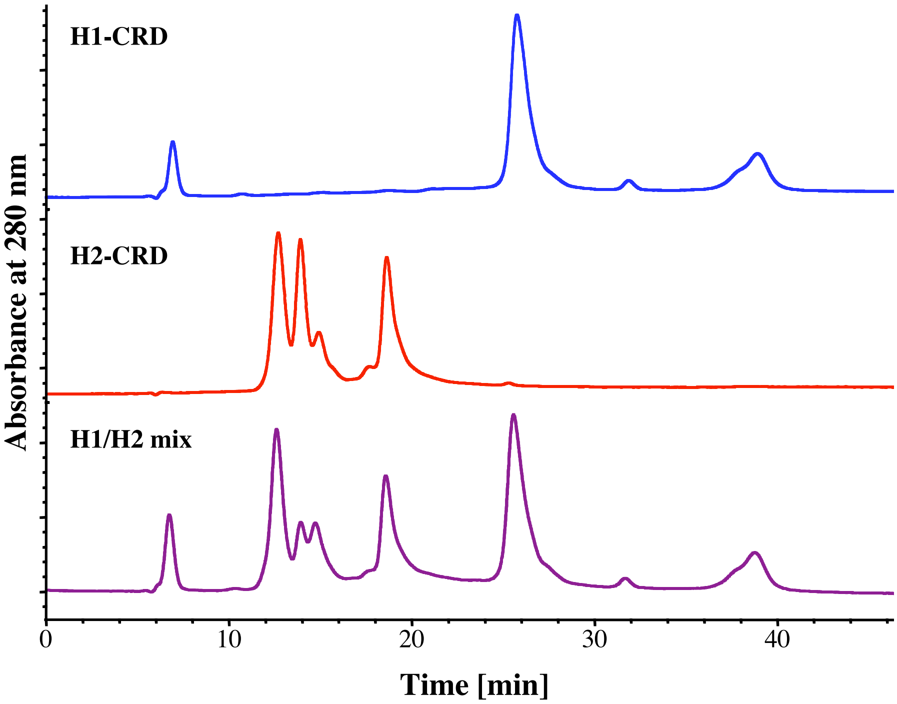


Figure 5-15: Monomer/dimer separation of H1-CRD (blue, top) and H2-CRD (red, middle) as well as of a H1/H2-CRD mixture (purple, bottom) using a single gradient run on a DEAE column.

Compared to H1-CRD, the separation of H2-CRD was found to be more complex and always resulted in different monomer peaks. The ratio of peak intensities could even change within different injections of the same batch (see chromatograms 'H2-CRD' and 'H1/H2 mix' on *figure* 5-15). SDS-PAGE analysis showed the existence of small amounts of dimer within the monomer peaks with significant variations in concentration. This indicates that additional, less stable dimers may be formed before and after separation.

Affinity HPLC

FPLC-based affinity chromatography on a galactose-sepharose column has already been used in the protein production process (see *sections* 5.2.2 and 5.3.1). It makes use of the lectin identity of the protein, which binds to galactose with a sufficiently high affinity. However, a separation of H1-CRD monomers and dimers is only possible if the process of dimerization alters the binding properties of the CRD binding site, e.g. due to steric hindrance or direct blocking of essential residues. Similar to the experiments on FPLC, it was not possible to discriminate between the two protein populations with this method, even when a linear EDTA gradient was used instead of step gradient. Yet it could be used to detect and remove non-active protein fractions.

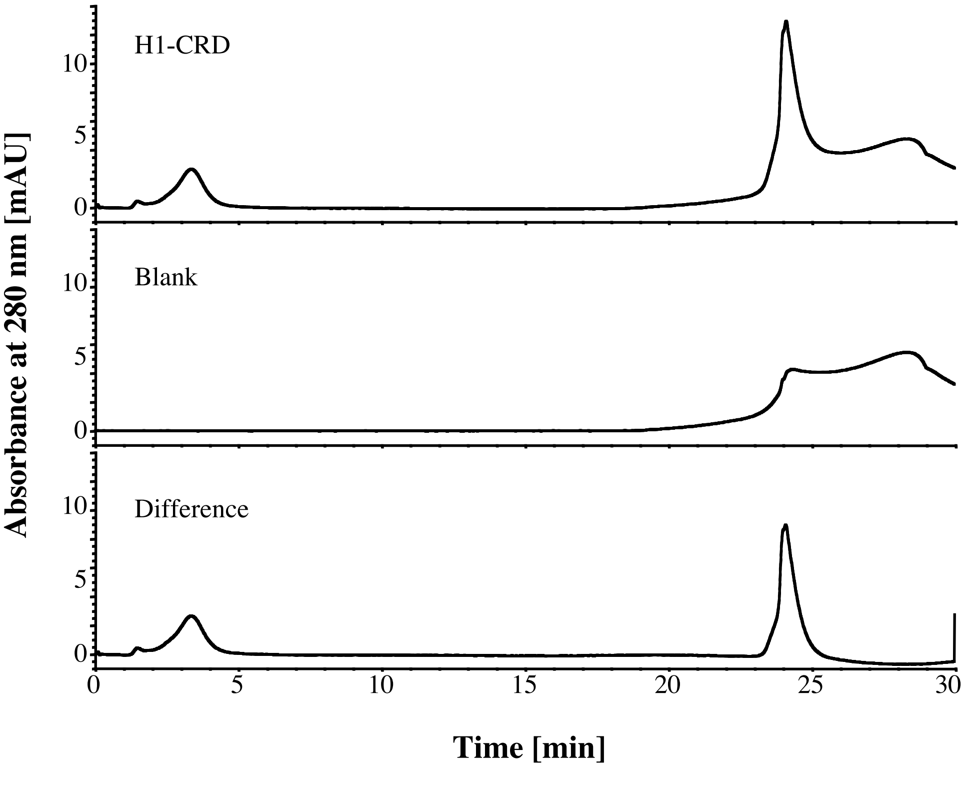


Figure 5-16: Processing of galactose-sepharose affinity chromatograms. The chromatogram from a blank injection was subtracted from the H1-CRD chromatogram in order to get the pure sample/protein signals.

Every change to the elution buffer caused a significant peak-like increase in the baseline. Therefore, a blank run had to be subtracted from the sample run for highest sensitivity, especially at low protein concentration (*Fig*. 5-16). A small signal shortly after injection start was usually detected and is believed to contain salts and inactive protein.

A first application of this method was the verification of H1-CRD activity before and after DEAE separation (*Fig*. 5-17A). No loss of activity could be detected when the samples were separated on the DEAE column. The injection peak was even slightly decreased indicating a higher purity of the sample. In contrast, a deactivated H1-CRD sample (before refolding) did not bind to the affinity column (*Fig*. 5-17A). Since the sepharose material itself was suspected to interact non-specifically with the lectin, a blank sepharose column was compared to galactose and Gal*N*Ac-coupled columns (*Fig*. 5-17B). While the protein clearly bound to both derivatized columns and did not elute until EDTA was added, there was no active protein peak detectable on the blank sepharose. However, the protein did not appear in the injection peak but eluted widely distributed within the first ten minutes. An extended washing phase is therefore recommended. No major difference in the separation properties and efficiency of galactose and Gal*N*Ac columns could be detected.

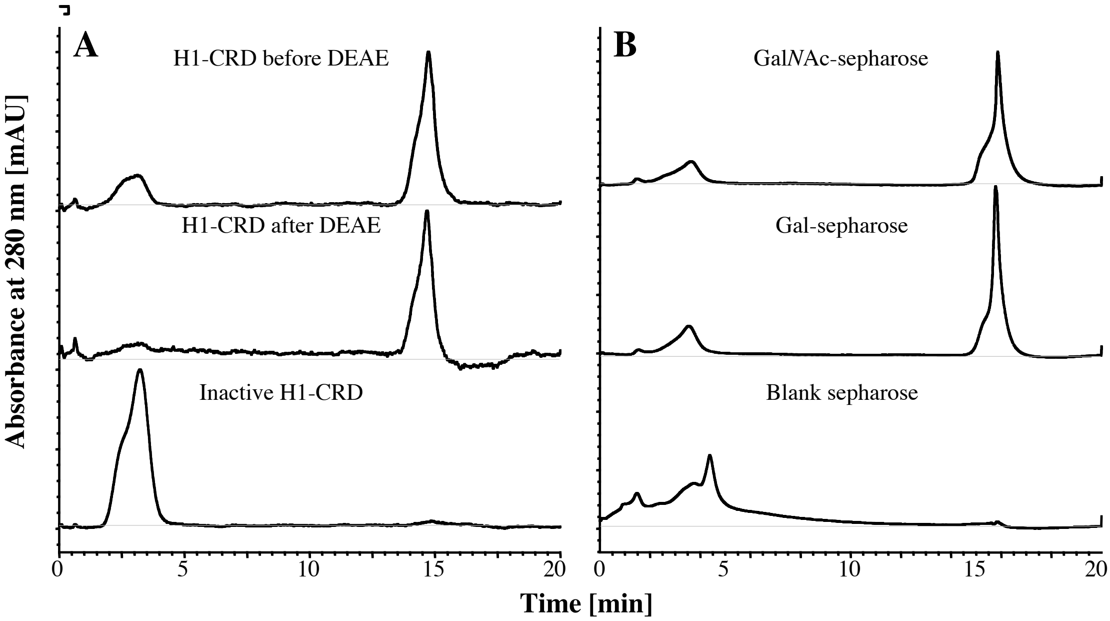


Figure 5-17: Affinity chromatography experiments with H1-CRD. **A:** Verification of the protein activity before and after DEAE separation. An inactive H1-CRD fraction was used as negative control. **B:** Binding specificity test on Gal- and Gal*N*Ac-sepharose columns compared to a blank sepharose column.

While it was not possible to separate monomers and dimers using affinity chromatography on a Gal- or Gal*N*Ac sepharose column, the method had its benefit in the post-production phase. In one run it was possible to check for the activity of DEAE-separated fractions, increase their concentration by multiple injections prior to a single elution, and simultaneously exchange the buffer to an alternative system compatible with Biacore immobilization or other applications (i.e. HEPES buffer).

5.3.3 Characterization of H1-CRD using mass spectrometry

H1-CRD samples, which were separated and desalted with RP-HPLC (see *section* 5.3.2), were further characterized using mass spectrometry. Since *E. coli* was used as an expression system, no glycosylation or phosphorylation is performed posttranslationally, and the measured masses were expected to correlate with the theoretical ones (*Table* 5-2). However, isolated monomer showed a single main peak on ESI-MS with a mass of 16,934 Da, which was significantly lower than expected (*Fig*. 5-18A). The mass difference of -57 Da did not correspond to the deletion or proteolysis of an amino acid. If the protein monomer has been expressed with a lower mass, the doubled difference should also be visible in the dimer fraction. ESI and MALDI analysis of the dimer fraction showed relatively wide peaks with maxima between 33,690 and 33,700 Da (*Fig*. 5-18B). This peak widening might be caused by disulfide shuffling. Here again, the difference to the expected mass (33,982 Da) was much larger than expected (approximately -280 Da) and clearly different from the 114 Da difference expected when doubling the monomer mass. This led to the hypothesis that the resulting mass differences might be the sum of an amino acid deletion and some adducts. The large discrepancy between monomer and dimer deviations suggested that a free cysteine residue (only 6 of the 7 cysteines of H1-CRD are involved in disulfide bridges; see *section* 5.1.3) could be involved in both dimerization and mass adducts. Therefore, an additional ESI analysis of the monomer fraction was performed under reducing conditions (*Fig*. 5-18C). The resulting mass was significantly lower than the one measured under non-reducing conditions. The difference of 131 Da exactly corresponded to the deletion of methionine residue (*Fig*. 5-20). Indeed, it is reported in literature that E. coli is able to remove a *N*-terminal methionine under certain conditions (the second amino acid has to be small) [47, 48]. Since these conditions are given in the case of H1-CRD (the second amino acid is a glycine), a posttranslational cleavage of the terminal methionine is highly likely.

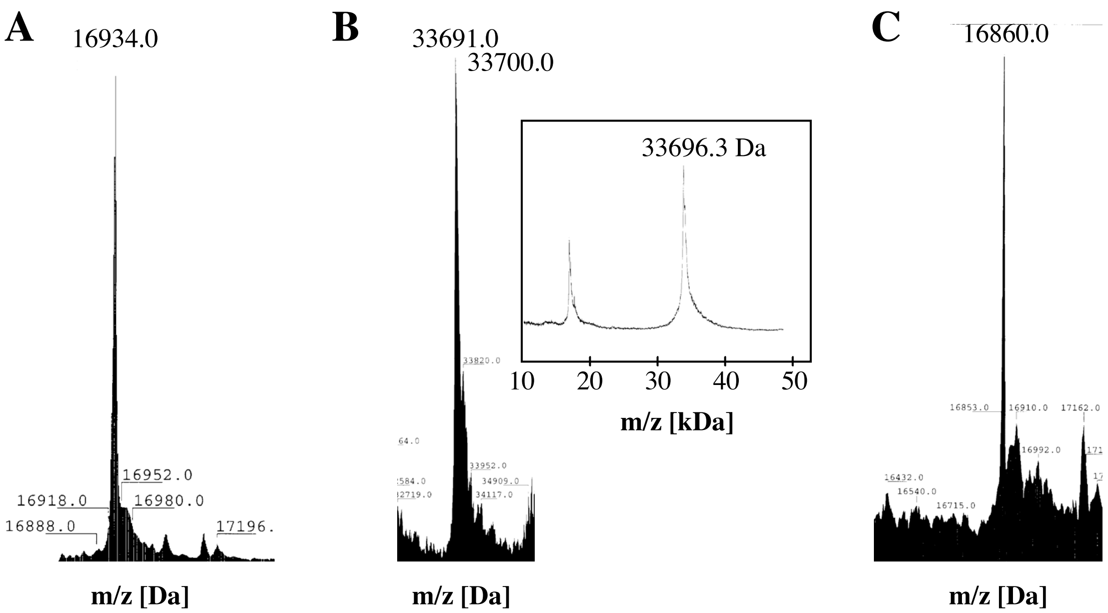


Figure 5-18: Mass analysis of H1-CRD fractions. **A:** ESI spectrum of the monomer fraction under non-reducing conditions. **B:** ESI and MALDI (insert) analysis of the dimer fraction under non-reducing conditions. **C:** ESI spectrum of the monomer under reducing conditions (DTT).

Table 5-2: Results of the mass spectrometric analysis of H1-CRD compared to the calculated theoretical mass (ProtParam analyis; see *appendix* D1).

H1-CRD Fraction	Experimental Conditions	Calculated Mass [Da]	Experimental Mass [Da]	Difference [Da]
Monomer	non-reducing	16,991	16,934	-57
Monomer	reducing	16,991	16,860	-131
Dimer	non-reducing	33,980	~ 33,700	-280

In order to confirm the absence of the terminal Met and to proof the sequence in more detail, a tryptic digest of the monomeric protein was performed after reduction and alkylation. Of the 12 expected tryptic fragments, 9 could be clearly identified (*Fig*. 5-19A). The remaining three peptides were too small to be visible in MS. Therefore, the digest was repeated with a LysC protease, which also cleaves after lysine but not after arginine residues (like trypsin), leading to larger fragments. This method led to a full sequence alignment and also confirmed the absence of the *N*-terminal methionine (*Fig*. 5-19B).

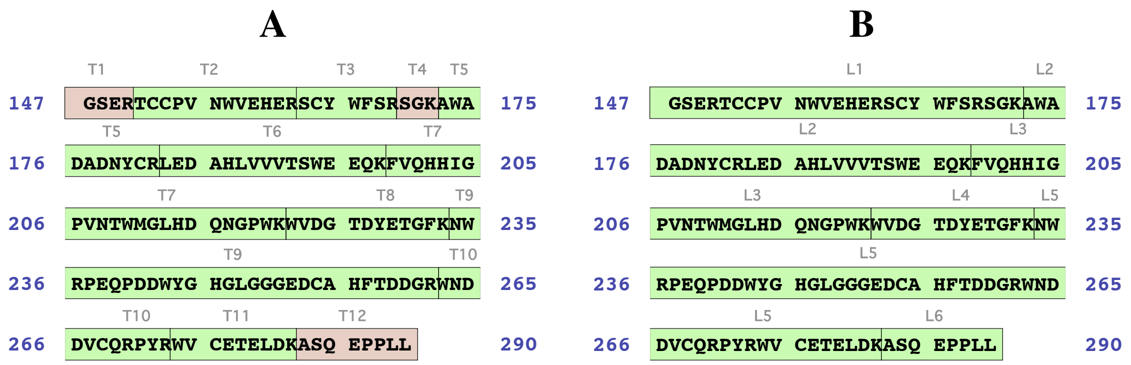

A
B
T1 T2 T3 T4 T5
147 GSERTCCPV NWVEHERSCY WFSRSGKAWA 175
T5 T6 T7
176 DADNYCRLED AHLVVVTSWE EQKFVQHHIG 205
T7 T8 T9
206 PVNTWMGLHD QNGPWKWVDG TDYETGFKNW 235
T9 T10
236 RPEQPDDWYG HGLGGGEDCA HFTDDGRWND 265
T10 T11 T12
266 DVCQRPYRWV CETELDKASQ EPPLL 290
L1 L2
147 GSERTCCPV NWVEHERSCY WFSRSGKAWA 175
L2 L3
176 DADNYCRLED AHLVVVTSWE EQKFVQHHIG 205
L3 L4 L5
206 PVNTWMGLHD QNGPWKWVDG TDYETGFKNW 235
L5
236 RPEQPDDWYG HGLGGGEDCA HFTDDGRWND 265
L5 L6
266 DVCQRPYRWV CETELDKASQ EPPLL 290

Figure 5-19: Proteolytic digests of H1-CRD analyzed by ESI-MS. Reduced and alkylated monomer fractions were digested with trypsin (**A**; T1-T12) or LysC protease (**B**; L1-L6). Identified peptidic fragments are highlighted in green, unidentified in red. No fragments with an *N*-terminal methionine residue could be detected.

With the knowledge about the absence of the terminal Met, the differences between calculated and measured masses were recalculated (Table 5-3). While the monomer under reducing condition very well correlates with the calculated mass, there is now a positive difference under non-reducing conditions. This strongly supports the hypothesis of an adduct on the free cysteine residue of H1-CRD, which is readily cleavable when adding reducing agents. Comparison of the mass difference (74 Da) with the ABRF DeltaMass database [49] indicated an addition of 2-mercaptoethanol [50] (*Fig*. 5-20). Since this reducing agent was used during the refolding process, an addition seems to be reasonable. Following this hypothesis, the mass increase changed to +305 Da when a combination of reducing and oxidizing glutathione was used in the refolding step during production development (*Fig*. 5-20; see *appendix* D2). Again, this increase corresponds to an addition of glutathione to the free sulfhydryl group. This behavior of a glutathiolation during refolding had been also observed in another study [51]. The small difference of ~ 18 Da in case of the monomer is rather caused by the wide peak maximum or by disulfide shuffling than by a real change of mass. An additional MS analysis of the dimer fraction in the laboratory of Prof. Jasna Peter-Katalinic (University of Münster, Germany) revealed a mass of 33,706 Da, which further reduces the difference to the calculated mass.

Table 5-3: Recalculated comparison of the results of the mass spectrometric analysis of H1-CRD compared and the calculated theoretical mass without the *N*-terminal methionine residue.

H1-CRD Fraction	Experimental Conditions	Calculated Mass [Da]	Experimental Mass [Da]	Difference [Da]
Monomer	non-reducing	16,860	16,934	+74
Monomer	reducing	16,860	16,860	±0
Dimer	non-reducing	33,718	~ 33,700	-18

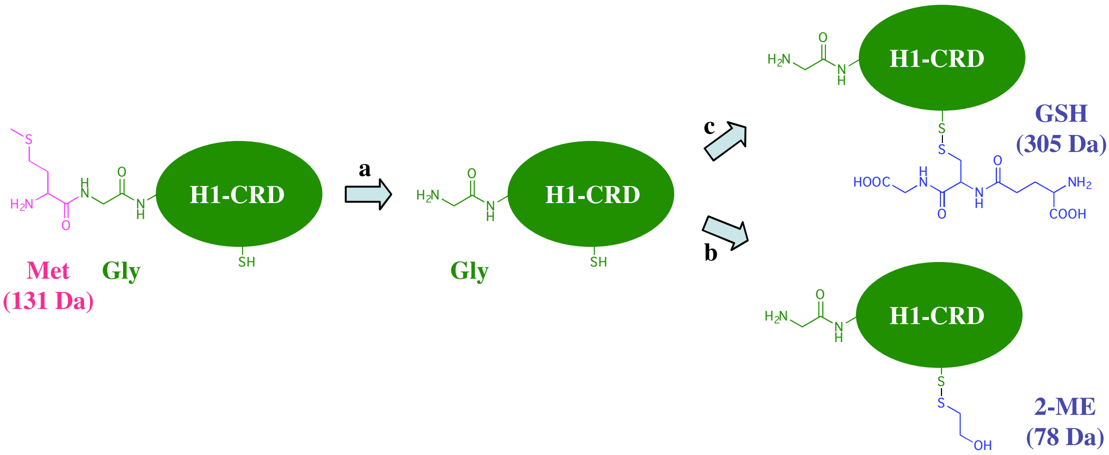


Figure 5-20: Mass changes by derivatization of H1-CRD at the *N*-terminus or cysteine residues. After a posttranslational removal of the *N*-terminal methionine (a), the free cysteine could be conjugated by 2-mercaptoethanol (2-ME; b) or glutathione (GSH; c) during refolding.

SDS-PAGE analysis showed further evidence for a disulfide-based dimerization, since the protein showed two bands under non-reducing conditions (17 and 33 kDa), while a single band (17 kDa) was visible in reducing buffer (see *appendix* D3). This leads to the hypothesis that a single free cysteine residue is responsible for dimerization. Formation of mixed disulfides during refolding could act as a protection group for the cysteine and therefore inhibits the dimerization process (*Fig*. 5-21). This also helps explaining the rather large variations in the monomer-dimer ratio during production, since a competitive reaction between monomers and thiol-active agents is expected to be highly dependent on concentrations, *pH*, salts, or temperature.

In the crystal structure of H1-CRD, only one of the two vicinal cysteine residues (Cys-152 and Cys-153) is resolved [13]. Since Cys-153 was reported to be involved in disulfide bridge formation, the position of Cys-152 should be rather fixed and therefore visible in the crystal. A derivatization of this residue at its thiol group might be a possible reason why Cys-152 could not be positioned into electron density. However, the *N*-terminal disulfide bridge (Cys-153/Cys-163) could not be confirmed by mass spectrometric analysis [31]. Therefore, it is rather difficult to determine, which cysteine residues are effectively involved in mixed disulfides or dimerization. Mutation experiments seem to be the most promising way to learn more about the disulfide structure of H1-CRD. Deletion or mutation of the single free cysteine could successfully prevent dimer formation. On the other hand, this cysteine residue could also be used for a site-specific labeling or immobilization of the protein by adding thiol-active compounds, which allow labeling of the protein (e.g. by biotin or a fluorescent dye) during refolding steps.

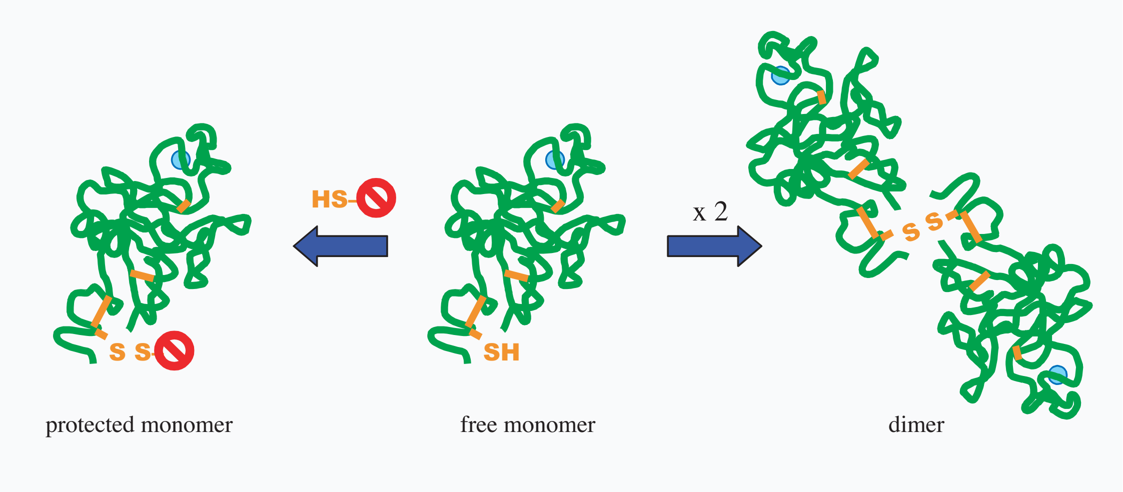


Figure 5-21: Dimerization hypothesis for H1-CRD. During refolding, the single free cysteine residue of H1-CRD (middle) could form dimers with another CRD (right) or might be protected from dimerization by thiol-active reagents like 2-mercaptoethanol or glutathione (left).

5.3.4 Development and optimization of a Biacore assay

Immobilization of H1-CRD

Standard amine coupling usually leads to a randomized immobilization on the sensor chip, since it targets all primary amines (lysines, *N*-terminus) that are available on the surface (see *section* 2.1.4). Therefore, an analysis of the Connolly surface was done to estimate the risk of an immobilization-induced deactivation of the binding site (*Fig*. 5-22). Neither the lysine residues nor the *N*-terminus are directly involved in ligand binding, which makes a direct blocking of the binding site very unlikely. Nevertheless, immobilization might lead to conformational changes or steric hindrance and, as a consequence, influence the affinity indirectly.

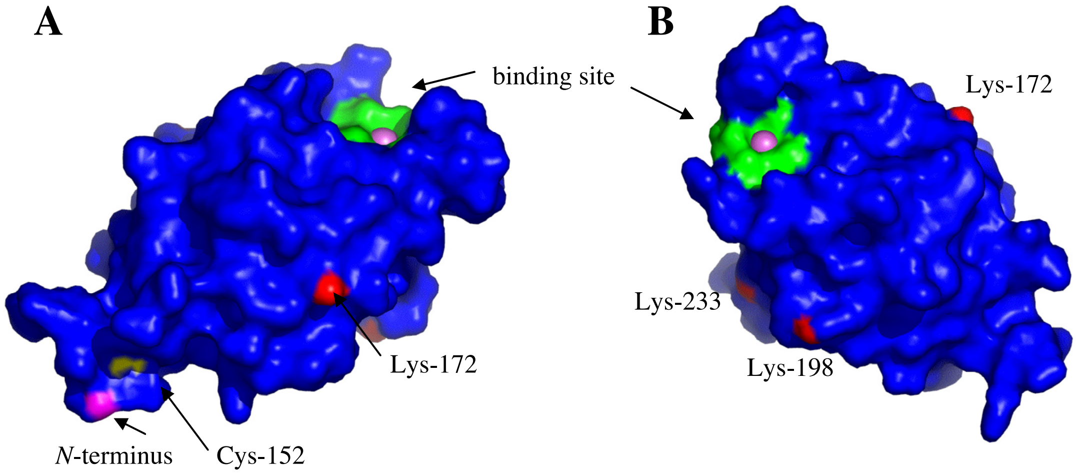


Figure 5-22: Analysis of possible immobilization sites on the H1-CRD structure. **A:** The *N*-terminus (Gly-147; magenta) and the hypothetical free cysteine residue (Cys-152; yellow) are located opposite to the binding site. **B:** All three accessible lysine residues (Lys 172, 198, 223; red) are not in close proximity of the sugar binding site. Amino acids involved in ligand binding are highlighted green, the calcium ion as a purple sphere. *N*-terminal six residues (GSERTC) were added manually to the PDB crystal structure 1DV8.

Of the four lysine residues in the H1-CRD crystal structure, three are accessible from the surface (*Fig*. 5-22B). They are all located clearly distant from the carbohydrate-binding site and an immobilization at these sites is not expected to decrease binding activity. The other possibility for amine coupling is the *N*-terminus, which is in even larger distance from the binding site (*Fig*. 5-22A). Targeting the supposed single free cysteine residue (Cys-152) in a thiol-coupling approach should not influence the activity either, since it also is situated in the same area opposite the binding site (*Fig*. 5-22A). This investigation indicates that H1-CRD immobilization through both amine and thiol coupling is not expected to cause any loss in activity. However, changes in salt conditions or decreased *pH* values, as it is necessary for amine coupling, might lead to a conformational change and a different accessibility pattern.

H1-CRD monomer and dimer fractions were immobilized separately on different flow cells of the sensor chip. The buffer exchange during the final affinity chromatography step (see *section* 5.3.2) allowed a direct immobilization without further processing. Surface attraction analysis by *pH* scouting showed an optimal immobilization *pH* of 4.5 to 5.0 (see *appendix* D4) at a concentration of 20 μg/ml protein. No $CaCl_2$ was added during the immobilization process, since the H1-CRD does not bind calcium under acidic conditions (see *'influence of pH'* later in this section). After immobilization end, the buffer was changed to calcium-containing running buffers, leading to active

H1-CRD. Using this method, H1-CRD fractions could be reproducibly immobilized on activated sensor chips (*Fig*. 5-23A&B).

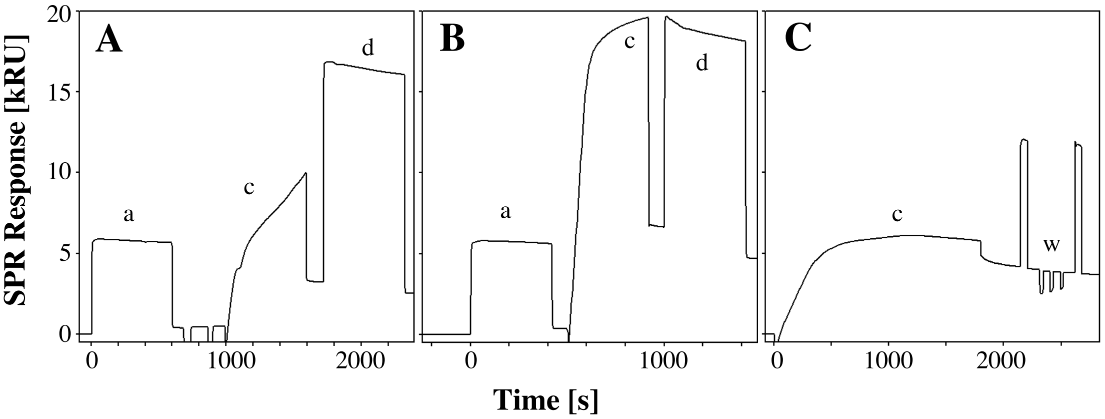


Figure 5-23: Immobilization of H1-CRD on CM5 sensor chips. **A:** Amine coupling of the monomer fraction with activation (a), coupling (c) and deactivation step (d). **B:** Amine coupling of the dimer fraction. **C:** Thiol coupling of H1-CRD after in-situ reduction/activation on a thiol-functionalized sensor chip surface (w = washing step).

As expected by the larger molecular weight of the H1-CRD dimer, the immobilization of dimer fractions usually led to higher surface densities. In addition, surface attraction and immobilization efficiency was also significantly increased in case of the dimer. This might be caused by the higher local concentration of primary amine groups on the molecule. The binding activities of immobilized monomer and dimer surfaces were evaluated by injecting a Gal*N*Ac concentration series (see *section* 5.3.5) and comparing the signal intensities and K_D values (*Table* 5-4).

In contrast to the binding affinity (K_D), which only showed slight variations between the two fractions, the calculated activity deviated more clearly. While both surfaces generated negative binding signals (see *section* 5.3.5), the signal intensity was unexpectedly high in case of the H1-CRD dimer and even exceeded the calculated R_{max} value. This indicates that the signal generation of the Gal*N*Ac/H1-CRD interaction is not solely caused by a mass increase upon binding but also includes another component (see *section* 5.3.6). The calculated affinity of 150% also shows that *equation* 4 (see *section* 2.1.3) does not sufficiently describe R_{max} for atypical binding signals as in the case of immobilized H1-CRD dimer.

Table 5-4: Surface activities and binding affinities of the immobilized H1-CRD monomer and dimer fractions towards an injected Gal*N*Ac concentration series (5 μM - 5 mM; see *section* 5.3.5).

Surface	**Surface Density [RU]**	**Calculated R_{max} [RU]** [a]	**Experimental R_{max} [RU]**	**Activity** [b]	**K_D [μM]** [c]
Monomer	2500	33	28	85%	150 μM
Dimer	4100	53	79	(150%)	100 μM
Factor	1.6	1.6	**2.8**	(1.8)	1.5

[a] Calculated from the surface density and the molecular weights of the target (H1-CRD monomer/dimer) and the analyte (Gal*N*Ac) by *equation* 4 (see *section* 2.1.3). [b] Calculated by dividing the experimental by the calculated R_{max}. [c] Calculated from the mirrored steady state plot using a single binding site model (see *section* 5.3.5).

The unpaired cysteine residue of the H1-CRD opened the possibility for a thiol coupling approach. Though, since the thiol group is expected to be involved in the formation of either a dimer or a mixed disulfide, a mild reduction step had to be performed. Therefore, the *in situ* method developed for the thiol immobilization of HSA (see *sections* 3.2.13 and 3.3.8) was adapted and led to a successful immobilization on a thiol-functionalized sensor flow cell (*Fig.* 5-23C). However, the generated surfaces were found to be less stable and usually required more than one hour before a stable baseline was reached. In addition, the immobilization conditions, especially the incubation and immobilization times, were also less reproducible and had to be empirically adapted for each immobilization. Since typical immobilization times were more than one hour, there is clearly no advantage concerning preparation speed. Even though thiol coupling seems to be an interesting alternative for a flexible and fully reversible immobilization of ASGP-R, the parameters for this method have to be developed further in order to improve the surface stability and the coupling time.

Buffer composition

C-type lectins bind their ligands in a calcium-dependent manner. In order to find an optimal calcium concentration, the signal intensity of asialofetuin was studied in buffers of increasing calcium concentration. The optimum was determined to be 50 mM $CaCl_2$, whereas higher concentrations (100 mM) decreased the intensity again (*Fig.* 5-24A). Any mismatch in the calcium concentration of the running and sample buffer led to an overlay of different signal effect and injection of calcium in a calcium-free buffer system generated positive binding signals showing rapid kinetics. Therefore, a dilution series of $CaCl_2$ was screened and could be fitted to two independent sites model (*Fig.* 5-24B).

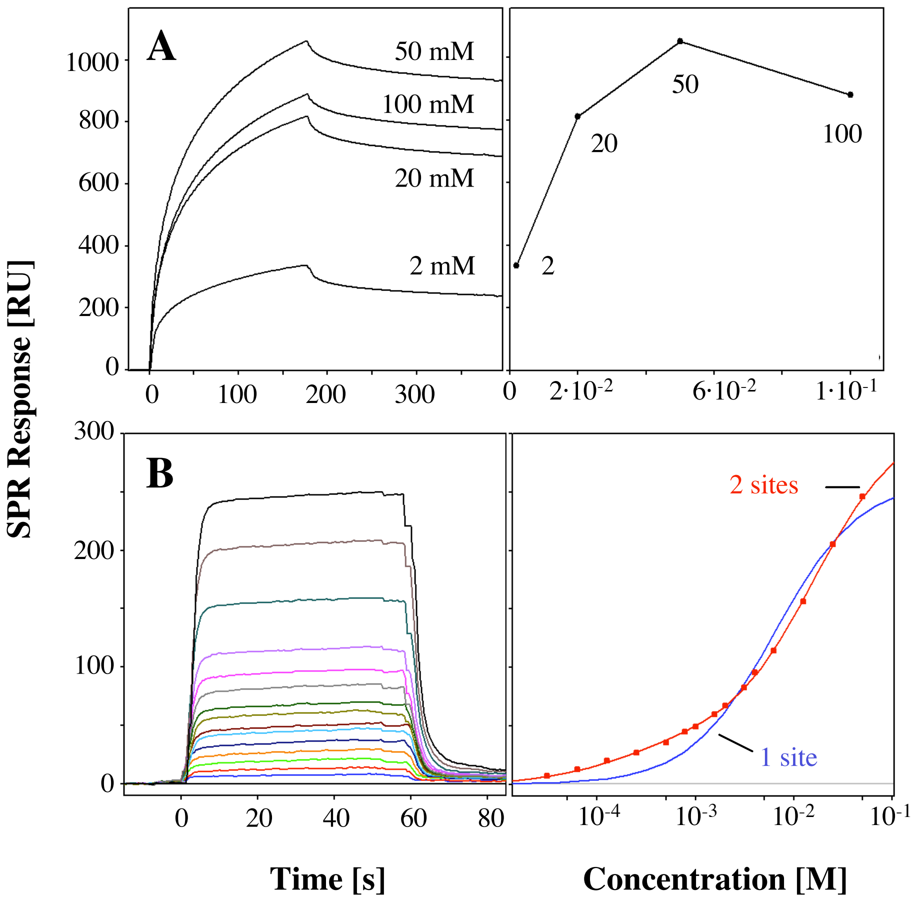


Figure 5-24: Calcium dependency of H1-CRD. **A:** Optimization of the calcium concentration for the binding assays. Asialofetuin (0.1 mg/ml) was injected at different calcium concentrations. **B:** Direct binding of calcium chloride to H1-CRD. $CaCl_2$ was injected from 30 μM to 50 mM over a H1-CRD surface in calcium-free buffer and fitted to single site (blue) and two independent binding sites models (red).

The steady state analysis indicated a high affinity (K_D = 170 μM) and a low affinity binding site (K_D = 16 mM). This is in good agreement with a study by Andersen *et al.* [27], in which they determined K_D's of 350 μM and 7 mM for the interaction of calcium with isolated rabbit ASGP-R subunits. In agreement with the crystal structure, they found three calcium ions per polypeptide chain and, after competition experiments with $MgCl_2$, stated two high affinity and one low affinity site.

The requirement of a rather high calcium concentration was confirmed by screening Gal*N*Ac as a monovalent carbohydrate (see *section* 5.3.5) in buffer containing 5 mM or 50 mM $CaCl_2$. While a clear and reproducible binding signal could be detected in the high calcium buffer, reducing the calcium concentration led to nearly complete signal loss (*Fig*. 5-25).

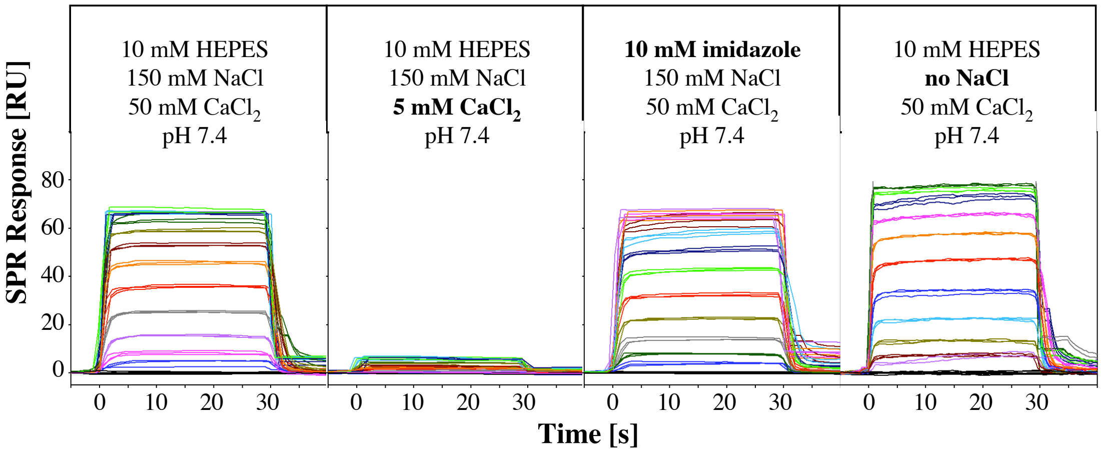


Figure 5-25: Influence of different buffer components on the signal intensity of Gal*N*Ac (5 μM-5 mM) to immobilized H1-CRD (mirrored signals; see *section* 5.3.5).

HEPES buffer was used for the experiments due to its high capacity at physiological *pH* and its ability to dissolve calcium chloride (in contrast to phosphate buffer). However, modeling experiments showed that HEPES (*Fig*. 5-26A) might interact with ASGPR and was also suspected to be involved in the generation of negative binding signals (see *section* 5.3.6). Docking studies for HEPES suggested that two of the sulfonic acid oxygen atoms might interact with the calcium. In this position, the hydrophobic piperazine core might interact with the aromatic side chain of Trp-243 (*Fig*. 5-26B). Therefore, HEPES was replaced by imidazole, which also tolerates calcium, and Gal*N*Ac was screened (see *section* 5.3.5) in both buffer systems. No significant change in binding affinity or signal intensity could be detected indicating that there is no competition for the binding site (*Fig*. 5-25).

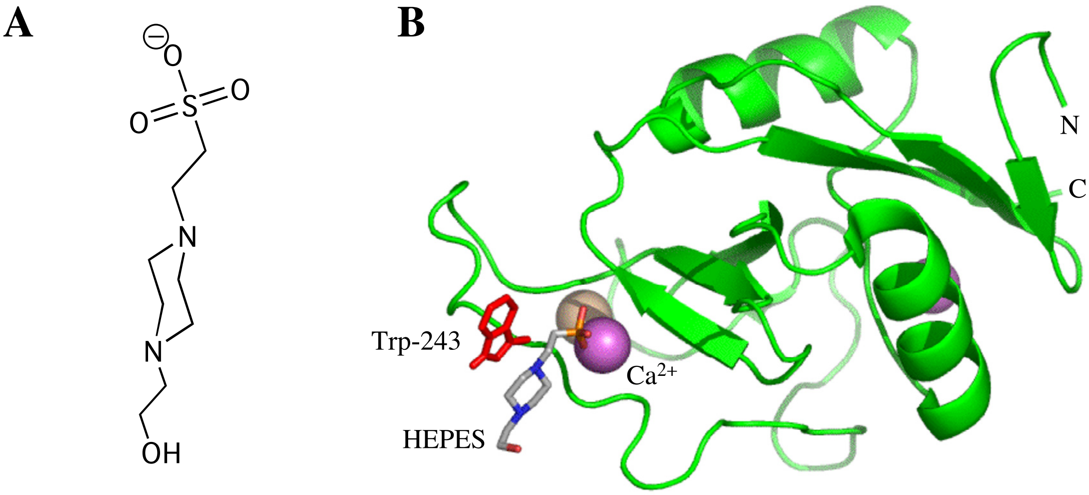


Figure 5-26: HEPES binding hypothesis for H1-CRD. **A**: Structure of HEPES (2-[4-(2-hydroxyethyl)-piperazin-1-yl]ethanesulfonic acid). **B**: Hypothetical binding mode of HEPES on H1-CRD after docking study (performed by Dr. Markus Lill, Biographics Laboratories, Switzerland).

In order to prevent ionic interactions with the carboxymethyl dextran matrix, NaCl is usually added to Biacore buffers. Since the buffer system is already highly ionic due to its $CaCl_2$ content, NaCl was removed from the buffer system. No decrease in affinity or intensity could be observed (*Fig.* 5-25, and NaCl was therefore avoided for future analyte screening experiments.

Influence of pH

The ASGP-R has to release its ligand upon fusing with the endosome. This mechanism is caused by a decrease of the surrounding *pH*, which leads to a conformational change of the receptor followed by a release of the calcium ions and the ligand. In rat ASGP-R, charged amino acids in and next to the binding site were identified as a molecular 'switch', which initiates the ligand release. The key step in this mechanism is the protonation of His-202 (corresponds to His-256 in human H1-CRD) [52, 53]. Therefore, H1-CRD is expected to show a significant *pH*-dependency in a critical *pH* range below 6.5. In order to test this relationship, binding of Gal*N*Ac and galactose to H1-CRD was analyzed in running buffers at various *pH* values (7.4-5.5). While a decrease of *pH* from 7.4 to 6.5 showed only a slight shift in affinity, a significant activity drop was visible at *pH* 6.0. Finally, essentially no binding could be detected when lowering the buffer *pH* to 5.5 (*Fig.* 5-27A&B).

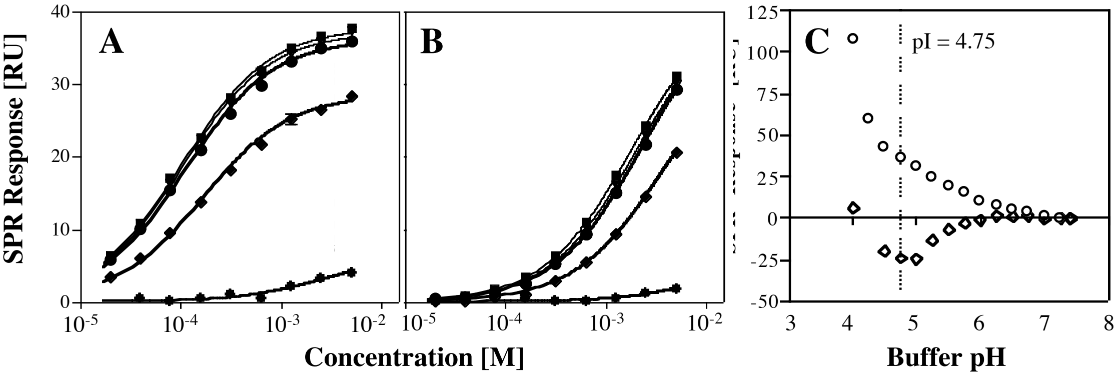


Figure 5-27: Influence of buffer *pH* on the binding of Gal*N*Ac (**A**) and galactose (**B**) to H1-CRD. Both analytes were injected in 10 mM HEPES buffer including 50 mM $CaCl_2$ at *pH* 7.4 (■), 7.0 (▼), 6.5 (●), 6.0 (◆) and 5.5 (✻). **C:** Injections of HEPES buffer blanks between *pH* 4.0-7.4 with 50 mM $CaCl_2$ (″) or 100 mM NaCl (ϒ) over a H1-CRD surface. The *pI* value of H1-CRD is indicated by a dotted line. Vertical error bars of the triplicate injections are shown in all curves but not visible in most cases due to low standard deviation.

This observation is in good agreement with literature. In case of the rat ASGP-R, a rapid decrease of ligand binding was reported when lowering the *pH* from 8.8 to 4.8. The midpoint of ligand release was found to be *pH* 7.1 and almost all ligand was released at the endosomal *pH* of 5.4. The release mechanism was clearly calcium-dependent. After an increase from 1 mM (physiological) to 5 mM $CaCl_2$ the ligand binding remained stable until a *pH* of ~7.4 while the endpoint stayed at *pH* ~ 5 (midpoint < *pH* 6.5) [52]. The significantly higher calcium concentration in the present Biacore assay is therefore expected to cause H1-CRD releasing its ligand even later. From an experimental point of view, these data also imply only a minor sensitivity towards small shifts in *pH* as they might occur during buffer preparation or as they could be induced by acidic analytes (see *section* 3.3.7). When lowering the *pH* during buffer blank injections at 50 mM $CaCl_2$, the signal began to drop significantly below *pH* 6.5 indicating that calcium is released from the protein. After further decreasing the *pH* below the CRD's *pI* (4.75) the SPR signal increased again, because surface attraction effects between the negatively charged matrix and the protonated protein became dominant. This effect could be confirmed by replacing the $CaCl_2$ with 100 mM NaCl, leading to a slow but steady signal increase until the *pI*, below which the signal raise was accelerated significantly (*Fig.* 5-27C). This indicates that the signal drop is selectively caused by the release of calcium ion. Whether the steady increase between *pH* 7 and 4.75 is solely generated by attraction effects or whether a reported conformational change of ASGP-R at lower *pH* values [28] is involved, can hardly be determined by Biacore experiments alone.

DMSO tolerability

The development and synthesis of novel carbohydrate mimics or conjugates is often accompanied by a significant increase in hydrophobicity. Therefore, organic modifiers have to be added to the solvent and DMSO has proofed to be ideally suited for most applications due to its miscibility with water and its biocompatibility. However, addition of DMSO nevertheless might influence protein activity or binding properties. In order to test H1-CRD for any DMSO sensitivity, carbohydrate screening (galactose and Gal*N*Ac) was performed in DMSO-free buffer and buffer containing 5 % DMSO (*Table* 5-5). Both the binding affinities and the signal intensities (R_{max}) were not significantly influenced by the addition of 5% DMSO. This makes the Biacore assay suitable for the screening of hydrophobic, drug-like substances.

Table 5-5: Binding of Gal*N*Ac and galactose to H1-CRD in buffer without and with 5 % DMSO.

Analyte	no DMSO K_D [μM]	5% DMSO K_D [μM]
Gal*N*Ac	72.2 ± 0.33	96.5 ± 0.55
Galactose	1300 ± 10	1610 ± 10

5.3.5 H1-CRD ligand screening

Asialoglycoproteins

Asialofetuin (ASF) and asialoorosomucoid (ASOR) are the best-characterized natural ligands for the ASGP-R. They both contain several branched glycan chain with terminal galactose residues and are reported to bind in the nanomolar range. Therefore, the binding assay was validated with these glycoproteins. While ASF is commercially available as desialylated product, the sialic acid groups of orosomucoid (i.e. acid α-glycoprotein, see *section* 3.1.1) had to be removed first. Two desialylation methods were performed and compared for this purpose: chemical cleavage in diluted H_2SO_4 and enzymatic desialylation by neuraminidase. In order to monitor the desialylation progress, an anion exchange separation method was developed. Since the glycoprotein looses the negative charges of the sialic acid groups, this causes a shift to earlier retention times on a DEAE column (*Fig*. 5-28).

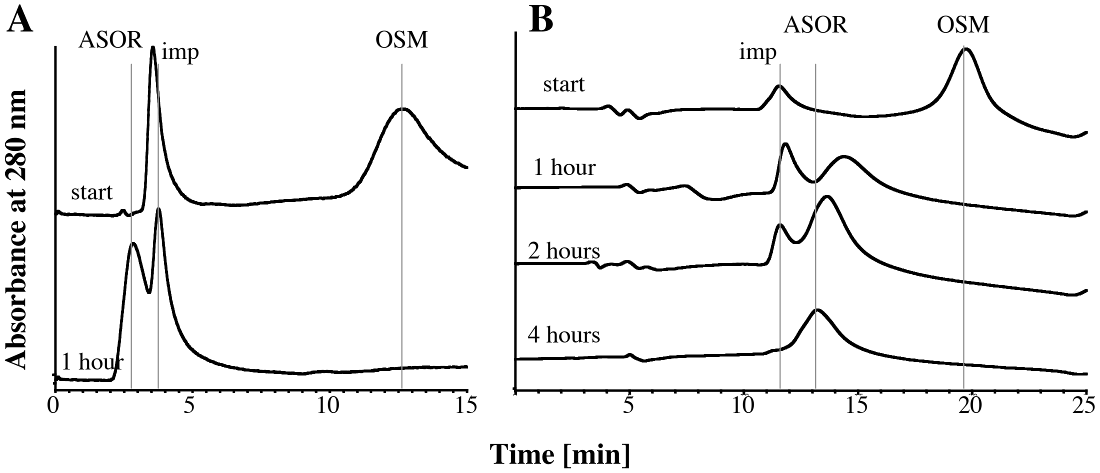


Figure 5-28: Monitoring of the desialylation process of orosomucoid (OSM) to asialoorosomucoid (ASOR). **A:** Sample before and after incubation in 0.05 M H_2SO_4 for 60 min at 80 °C. **B:** Enzymatic desialylation of OSM shortly after neuraminidase addition and after 1, 2, and 4 hours at 37°C. There was also an unknown impurity visible in both separations with a λ_{max} at 260 nm (imp).

While the process was complete after one hour in the case of the acidic desialylation (*Fig.* 5-28A), the enzymatic method was significantly more time-consuming (4 hours; *Fig.* 5-28B). In both cases, an unknown impurity with an absorption maximum at 260 nm was detected even at start conditions. Interestingly, the impurity seems to vanish during enzymatic but not during acidic desialylation.

Both asialoglycoproteins were screened over different H1-CRD surfaces in order to evaluate their kinetic properties and to detect differences between the immobilization methods described in *section* 5.3.4. The binding specificity for desialylated glycoproteins bearing terminal Gal/Gal*N*Ac moieties was tested by injecting similar concentrations of ASF, ASOR as well as orosomucoid (OSM) before desialylation over a H1-CRD surface. While latter showed no detectable binding, both asialoglycoproteins bound clearly and specifically to the lectin with rather slow binding kinetics (*Fig.* 5-29A). The two active glycoproteins were then screened over three H1-CRD surfaces (amine-coupling of monomers and dimers, thiol-coupling of monomers) and the resulting sensorgrams were kinetically evaluated (*Fig.* 5-29B&C).

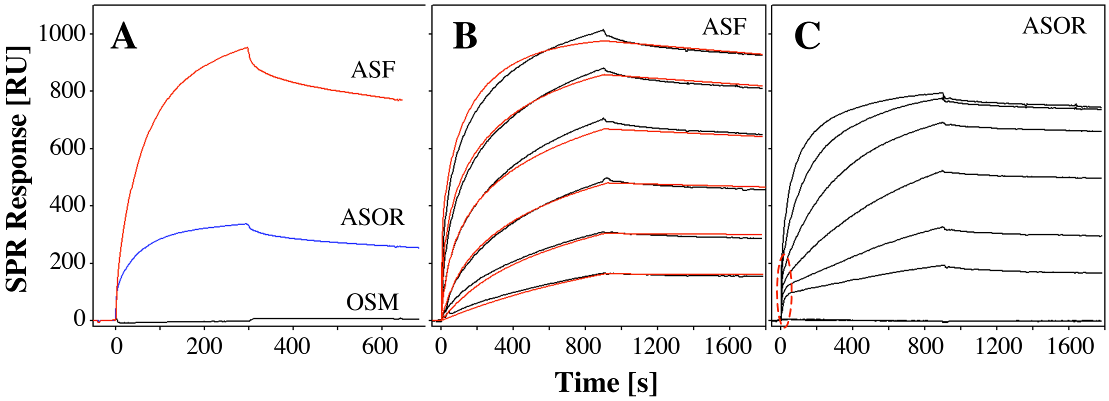


Figure 5-29: Binding of asialoglycoproteins to immobilized H1-CRD. **A:** Comparison of the binding specificity of asialofetuin (ASF), asialoorosomucoid (ASOR) and orosomucoid (OSM) at 2.5 μM on a low-density surface. **B:** Injection series of ASF (1000-4 nM) in 10 mM HBS. 50 mM $CaCl_2$, 0.005% polysorbate 20, *pH* 7.4 over a thiol-coupled H1-CRD surface. Simulated kinetic curves of a bivalent binding model are overlaid in red. **C:** Injection series of ASOR under the same conditions as in B. An unusually fast increase of the signal at injection start is indicated by a red circle.

Asialofetuin showed clearly concentration-dependent and reproducible binding signals with slow kinetic off-rates. As expected, the sensorgrams could not be fitted to a simple Langmuir 1:1 binding model, because polyvalent binding, rebinding effects, microheterogeneities of the glycosylation pattern as well as possible surface heterogeneity from protein immobilization are influencing the binding behavior.

Therefore, acceptable fits resulted when using a *surface heterogeneity* model while the best matches were realized with a *bivalent analyte* model (*Fig*. 5-29B, *Table* 5-6).

Table 5-6: Kinetic evaluation of asialofetuin binding to three different H1-CRD surfaces using a *bivalent analyte* binding model (4-1000 nM in 10 mM HBS, 50 mM $CaCl_2$, 0.005% polysorbate 20, *pH* 7.4). Overlay sensorgrams of all binding curves and fits can be found in *appendix* D5.

Surface	Density [RU]	k_{on} [$M^{-1}s^{-1}$]	k_{off} [s^{-1}]	R_{max} [RU]	K_D [nM] [c]	$t_{1/2}$ [min] [d]
Monomer [a]	1900	$76 \cdot 10^3$	$6 \cdot 10^{-5}$	705	0.79	190
Dimer [a]	1850	$154 \cdot 10^3$	$3 \cdot 10^{-5}$	1235	0.20	375
Thiol [b]	2400	$123 \cdot 10^3$	$4 \cdot 10^{-5}$	1010	0.31	300

[a] Prepared by amine coupling of monomer or dimer fractions. [b] Prepared by thiol coupling of the monomer fraction. [c] Calculated as $K_D = k_{off}/k_{on}$. [d] Calculated as $t_{1/2} = \ln 2/k_{off}$.

Asialoorosomucoid showed an even more complex binding pattern. After a rapid signal increase at injection start, the association phase significantly slowed down. The dissociation part was monophasic and similar to the one of ASF. Therefore, an overlay of two independent binding events as observed in the case of the GSLA-2 antibody (see *section* 4.3) can be excluded. Depending on the surface, higher concentrations of ASOR (0.3-1 μM) generated large negative drifts during injection (see *appendix* D6). As a consequence, the sensorgrams could not appropriately be fitted to a common kinetic model. The reason for this atypical binding behavior is not known. An accumulation or competition of impurities as seen during desialylation cannot be excluded. Therefore, ASOR experiments have to be repeated after a more intense purification and characterization of the glycoprotein.

As mentioned above, a *bivalent analyte* model generated the most appropriate kinetic description of the ASF binding behavior, even though no complete fit could be achieved (*Fig*. 5-29B). Fetuin is reported to bear three sites for *N*-linked glycosylation, each carrying either a bi- or triantennary glycan [54, 55]. In contrast to ASOR, which only carries terminal galactose, ASF also shows heterogeneities in the carbohydrate moieties with a Gal:Gal*N*Ac ratio of 4:1 [20]. Therefore, multivalent binding is expected for both glycoproteins and does also explain, why a bivalent analyte model describes the interaction rather well but not completely. Even though bivalent binding implies two sets of rate constants for the primary and secondary binding event, respectively, only one set can be obtained by the Biacore experiment. When an already bound analyte molecule interacts with its second (or third) binding site, no increase of mass concentration and therefore no additional SPR signal can be detected [56].

Significant variations between the monomer and dimer surfaces were found for the kinetic rate constants. While the on-rate of the dimer was increased by a factor of two, its dissociation rate was slower by the same factor. This resulted in a four times higher affinity and clearly prolonged dissociation half-life of more than six hours.

The thiol-coupled surface was expected to show properties more similar to the monomer rather than the dimer fraction, due to its mode of immobilization. However, all values were between the two amine-coupled surfaces and clearly closer to the dimer fraction. This might be explained by the slightly higher surface density obtained by this coupling method resulting in an increased rate of rebinding and multivalency. In addition, an immobilization solely around the N-terminus and therefore opposite to the binding region (*Fig.* 5-22) could improve the accessibility of the binding sites. The obtained K_D's between 200 and 800 pM are lower than the IC_{50} values reported for the interaction with isolated human ASGP-R lectins (17 nM) [20]. Dissociation constants for the interaction with overexpressed H1 in the absence of H2 were only obtained for ASOR (40 nM) but not for ASF [29].

Mono- and disaccharides

The asialoglycoprotein receptor shows a high selectivity for galactose moieties and its derivatives (see *section* 5.1). Gal*N*Ac is the natural monosaccharide with the highest affinity for the ASGP-R, at least ten times better than galactose [20, 57]. The expected affinity in the micromolar range is well suited for SPR detection, whereas the small molecular weight around 200 Da was identified as a potential problem. First test injections with 1 mM solutions of different carbohydrate analytes provided a confusing result, since all injections caused negative binding signals. This behavior usually indicates problems in the assay design and are therefore mostly rejected. However, comparison of the samples showed a tendency in the negative signal intensity, which directly correlated to the expected affinity of the tested sugars. Therefore, a concentration series of Gal*N*Ac was injected and analyzed more deeply. While the SPR signals on the reference flow cell always had a positive sign, the H1-CRD flow cell generated responses both in the positive and negative range (*Fig.* 5-30A). Referencing of the signals (i.e. subtraction of the reference flow cell) resulted in negative sensorgrams for each sample injection (*Fig.* 5-30B). Since *Scrubber* allows the multiplication of the sensorgrams with a constant value (included for the normalization of values from different assay types), the binding curves could be mirrored by multiplication with -1 (*Fig.* 5-30C). Processed sensorgrams from randomized triplicate injections not only were fully reproducible, but also showed the fast binding kinetics expected for protein-carbohydrate interactions [58, 59] resulting in stable steady state phases within a few seconds (*Fig.* 5-30D). When fitting these equilibrium values to a single binding site model, the data points matched the model with high accuracy (*Fig.* 5-30E) and led to dissociation constants (K_D in the rage of 90-150 μM for Gal*N*Ac), which were well within the expected range. In addition, the sensorgrams could also be kinetically fitted to a simple Langmuir 1:1 interaction model (*Fig.* 5-30F). These results indicated that the Biacore binding assay generated reliable interaction data despite the unusual occurrence of negative SPR signals. A detailed discussion of possible causes for the negative responses can be found in *section* 5.3.6.

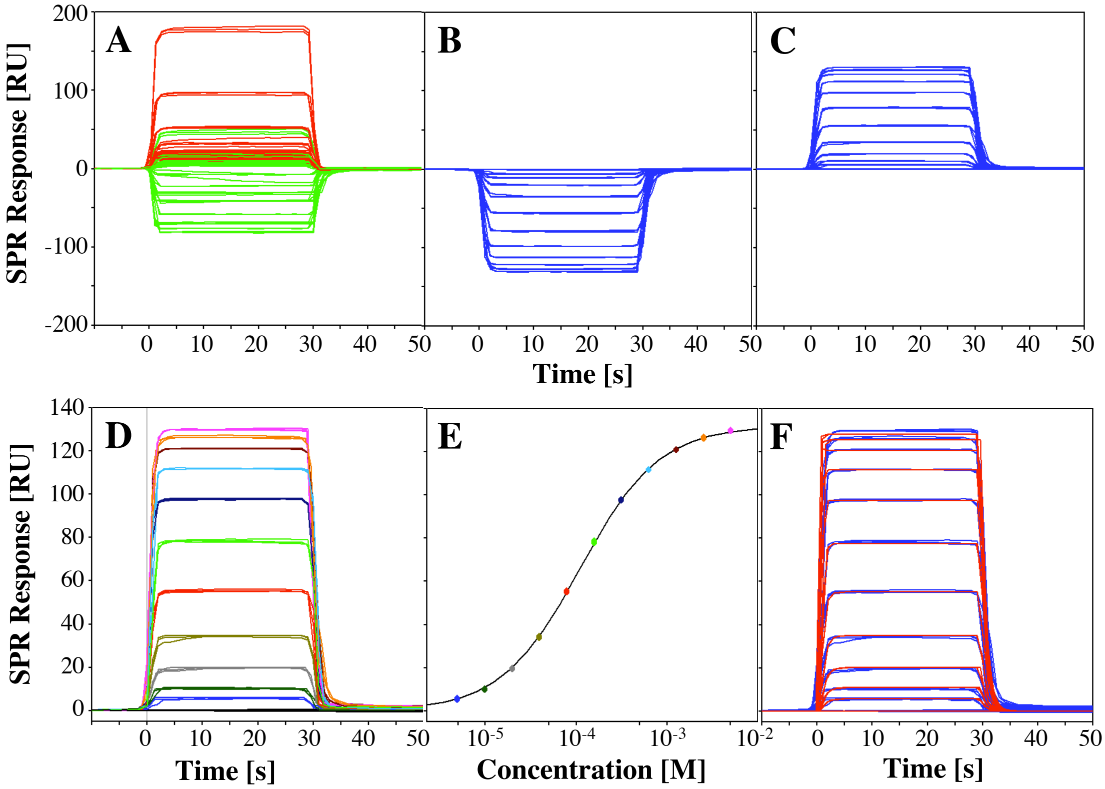


Figure 5-30: Data processing (**A**-**C**) and data evaluation (**D**-**F**) of Gal*N*Ac binding to H1-CRD. Referencing of the biosensor raw data (**A**) by subtracting the signals of an empty reference cell (red) from the measurement cell (green) leads to negative binding signals (**B**) that can be mirrored to get interpretable data (**C**). Resulting overlay plots (**D**) can be either evaluated by fitting the equilibria to a single binding site model (**E**) or by kinetic fitting of the whole sensorgrams to a Langmuir 1:1 interaction model (**F**).

In order to validate the binding assay and get more detailed information about the binding motif of the ligands, a set of galactose derivatives as well as glucose (negative control) was screened. All carbohydrate analytes showed negative binding signals with fast kinetic properties. After mirroring, all binding equilibria could be fitted to a single binding site model and resulted in affinities between 0.1 and 5 mM (*Table* 5-7).

As expected, Gal*N*Ac was identified as the best binding monosaccharide, approximately 10 to 15 times better than galactose and lactose. Galactose derivatives bearing a methyl residue at the reducing end bound with a lower affinity than galactose, with the beta form showing a slightly better K_D. While galactosamine bound with detectable but clearly weaker affinities, glucose binding was beyond detection limit. The K_D values obtained in the Biacore assay were compared to IC_{50} values from a solid-phase competition assay, were a polymeric Gal*N*Ac competes with the analytes for immobilized H1-CRD (*Fig*. 5-32A) [35].

Table 5-7: Equilibrium binding constants of a panel of carbohydrate analytes binding to immobilized H1-CRD monomer and dimer fractions (triplicate injections, 5 μM - 5 mM) compared with IC_{50} values of a solid-phase competition assay [35] and of literature data [57, 60, 61] (SPR fits see *appendix* D7).

Carbohydrate Analyte	**Monomer K_D [μM]**	**Dimer K_D [μM]**	**SPCA [a] [35] IC_{50} [μM]**	**Literature Values**
Gal*N*Ac	150 ± 0.9	104 ± 0.3	97	90 [b]
Lactose	2090 ± 50	1390 ± 10	1100	300 [b]
Galactose	1460 ± 10	1590 ± 10	n.d.	1700 [b]
Methyl β-D-galactopyranoside	2200 ± 40	1800 ± 10	1800	1000 [b]
Methyl α-D-galactopyranoside	2760 ± 60	2350 ± 20	2600	1600 [b]
Galactosamine	4170 ± 110	4630 ± 50	2600	> 25000 [c]
Glucose	> 10 mM	> 10 mM	n.d.	60000 [b]

[a] Solid-phase competition assay [35]. [b] IC_{50} values [μM] from isolated rabbit ASGP-R lectin [57, 60] [c] IC_{50} value [μM] from rat hepatic membranes [61]. The assays were performed with radioactively labeled ligands.

Not only the general ranking of the analytes was congruent, but also the absolute affinity values were found to be very close. The rather high calcium concentration of the Biacore assay does not seem to clearly influence binding, since a much lower concentration of 5 mM was used in the solid-phase competition assay. When comparing the binding affinities to published literature values [57, 60, 61], the general ranking and the affinity ranges are similar again. While lactose is reported to bind better than in the competition assay [35] or the Biacore assay, the biggest deviation was found in case of galactosamine with a tenfold weaker reported affinity. These variations may be attributed to a different experimental setup or, more likely, to species differences between human and rabbit/rat lectin. Furthermore, galactosamine was measured on rat liver plasma membranes, which contain the whole ASGP-R [61] and nothing is reported about the subtype of the isolated rabbit lectin used for the other compounds [57, 60]. The otherwise excellent agreement with the two methods (competition assay and literature IC_{50}) further validates the Biacore assay. Most of the analytes bound slightly better to the dimer surface with 10-50% increased affinity values. A possible reason for this difference might be the higher local concentration of binding sites, favoring repetitive binding events and subsequently the kinetic on-rate. To analyze this effect more deeply, all data sets were also kinetically evaluated by globally fitting the curves from the triplicate injections to a simple 1:1 Langmuir binding isotherm (*Table* 5-8).

Table 5-8: Kinetic analysis of the carbohydrate binding data on H1-CRD dimer (steady state affinity data in *Table* 5-7). All curves were globally fitted to a simple 1:1 binding model.

Carbohydrate Analyte	k_{on} [$M^{-1}s^{-1}$]	k_{off} [s^{-1}]	$t_{1/2}$ [s] [a]	K_D [μM] [b]
Gal*N*Ac	6676 ± 83	0.68 ± 0.01	1.0	100
Lactose	779 ± 17	1.09 ± 0.02	0.6	1400
Galactose	694 ± 15	1.11 ± 0.02	0.6	1590
Methyl β-D-galactopyranoside	534 ± 10	0.98 ± 0.02	0.7	1840
Methyl α-D-galactopyranoside	396 ± 9	0.94 ± 0.02	0.7	2380
Galactosamine	378 ± 8	1.76 ± 0.03	0.4	4650
Glucose [c]	n.d.	n.d.	n.d.	n.d.

[a] Calculated as $t_{1/2} = \ln 2/k_{off}$; [b] Calculated as $K_D = k_{off}/k_{on}$: [c] SPR signals were too low for a kinetic analysis.

When comparing the kinetic on- and off-rates of the interactions, the k_{on} values seem to contribute more to the overall affinity, indeed. This is especially obvious in the case of Gal*N*Ac: while the off-rate differs only by less than a factor of 2, the on-rate was almost ten times higher than those of galactose. In addition, the association rates followed the ranking of the overall affinity, whereas the tendency of the dissociation rates is less pronounced. These results therefore support the theory stated above that on-rates are the predominant factor for the differences in the K_D values. On the other hand, galactosamine showed a slightly lower on- and a significantly faster off-rate than the other analytes resulting in a weaker affinity. However, fitting of fast kinetics is rather difficult and all values are around or below the detection limits stated by Biacore. The absolute values have therefore to be interpreted with care.

Unexpectedly, H1-CRD surfaces generated by thiol coupling showed a dramatically reduced signal intensity compared to the amine-coupled surfaces (*Fig.* 5-31A&B). In addition, the binding mode seems to get more complex since the obtained equilibrium data fitted better to a two independent sites than to a single binding site model (*Fig.* 5-31C; Gal*N*Ac data of the same experiment is visualized in *Fig.* 5-38C). While the K_D for the first site was likely to correspond with an active protein interaction (140 μM), the second K_D was rather high (6 mM). This may indicate, that a substantial amount of the immobilized protein might be deactivated or fixed in an unfavorable position. On the other side, it is also possible that the exact position of attachment influences the generation of the binding signal.

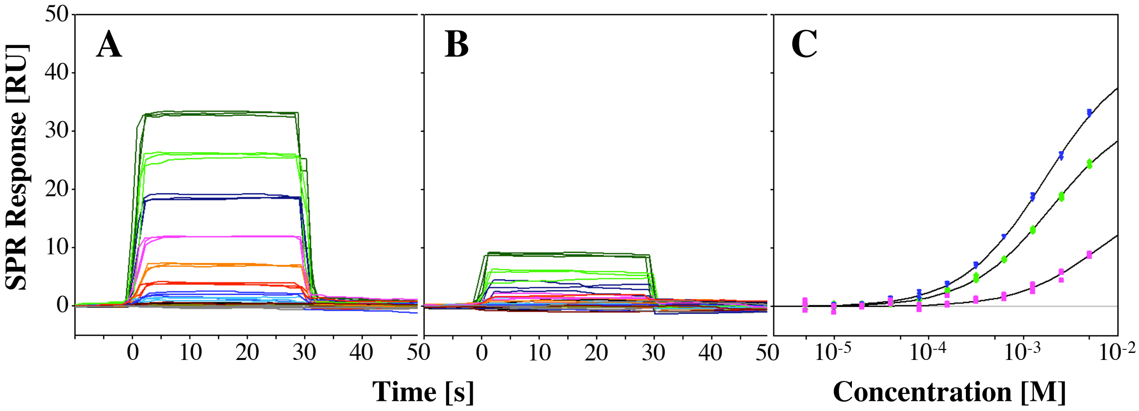


Figure 5-31: Effect of the immobilization method on the binding of monosaccharides using galactose as an example. Overlay plots of sensorgrams typically obtained on amine-coupled (**A**) and thiol-coupled (**B**) H1-CRD monomer surfaces. **C:** Comparison of the equilibrium dissociation curves of galactose on amine-coupled monomer (green) and dimer (blue) as well as thiol-coupled (magenta) H1-CRD surfaces. While latter was fitted to a two independent sites model, the two data sets from amine coupling fitted to a single site model.

Polymeric forms of the monosaccharides as they are used e.g. in the H1-CRD solid-phase competition assay, are expected to show higher affinities due to polyvalent binding and slower kinetic off-rates caused by rebinding effects. To evaluate this behavior, a constant concentration of polymer-bound Gal*N*Ac and glucose was injected over the H1-CRD surfaces (*Fig.* 5-32B). While poly-glucose did not or only very weakly bind to H1-CRD, poly-Gal*N*Ac showed a strong activity for both monomer and dimer surfaces. In contrast to the very fast kinetic profile of the free monosaccharide, the polymer-bound sugar featured a very slow dissociation rate. Injection of monovalent galactose clearly competed for the same binding site and caused a detectable drop in the polymer binding. Finally, the surface could be completely regenerated with EDTA, indicating a calcium-dependent and therefore specific binding. The difference in the signal intensity between the monomer and dimer surface is mainly caused by variations in the surface density of both proteins.

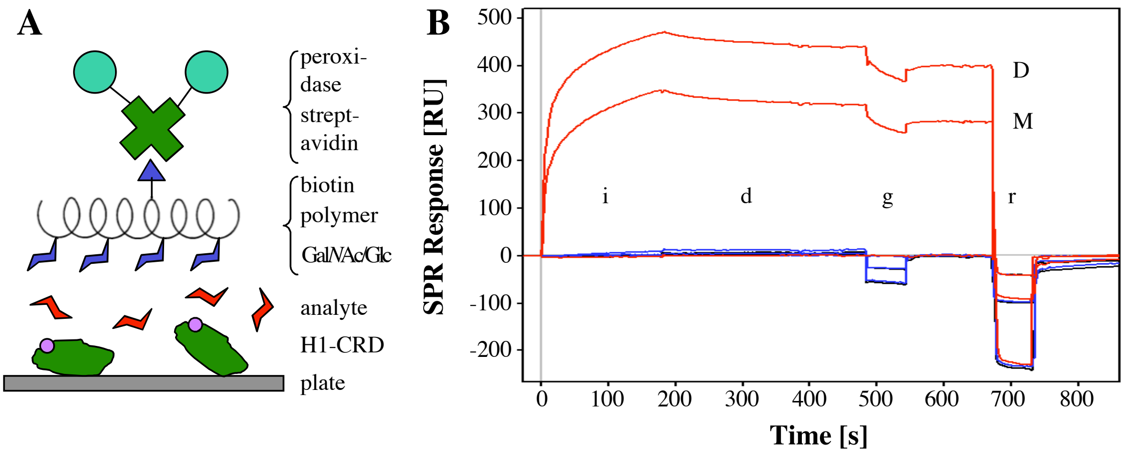


Figure 5-32: H1-CRD polymer assay. **A:** Overview of the solid-phase polymer competition assay [35] used for the screening of H1-CRD ligands (adapted from Bovin [62]). **B:** Injection of poly-Gal*N*Ac (red) and poly-glucose (blue) over immobilized H1-CRD monomer (M) and dimer (D) fractions. After an injection (i) and dissociation phase (d), 10 mM galactose (g) were injected to test for competition effects. Finally, the surface was regenerated by injecting EDTA buffer (r).

anti-H1-CRD antibodies

Monoclonal antibodies have developed into some of the most powerful tools in both molecular biology and medicine. Their high selectivity and usually strong affinities make them also interesting for the purification and characterization of recombinant proteins. For the development of Biacore assays, monoclonal antibodies are important in two major areas, i.e. capturing of the proteins for oriented immobilization (see *section* 2.1.4) as well as for blocking the active protein in order to evaluate binding specificity.

A set of monoclonal antibodies directed against the human ASGP-R H1-CRD [30] was screened for their activity towards monomer and dimer fractions. In a first step, six antibodies were injected at a constant concentration and their signal intensities, dissociation stability, blocking capabilities and calcium/EDTA sensitivity were evaluated (*Fig*. 5-33A).

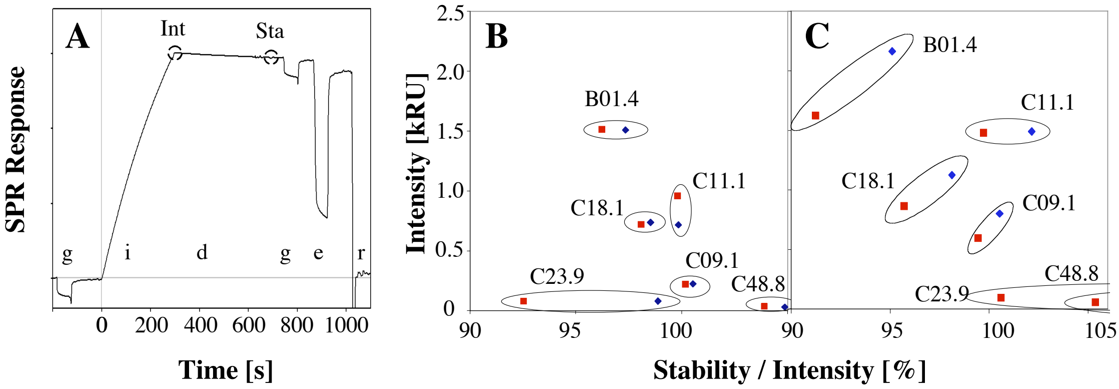


Figure 5-33: Ranking of a set of H1-CRD-binding antibodies. A: Setup of the ranking assay with a preliminary injection of 5 mM galactose (g), antibody injection (i), dissociation phase (d), a second galactose injection for testing blocking properties, an EDTA pulse (e) and the regeneration with HCl (r). Data points used for the ranking (Int, Sta) are indicated by dashed circles. **B, C:** Ranking of the antibody samples by plotting the signal intensity after the 5 min injection (Int) against the quotient of the stability after 5 min dissociation (Sta) and the intensity in buffer containing 50 mM $CaCl_2$ (**B**) or 3 mM EDTA (**C**). Values for the monomer surface are indicated in red and for the dimer surface in blue.

Even though some of the antibodies seem to block sugar binding in the solid-phase competition assay [30], none of the screened antibodies showed a clear blockage of the binding site in the Biacore assay. The differences between the galactose signal before and after the antibody injection were below 10% for all samples (see *appendix* D8). This discrepancy to the results of the competition assay might e.g. be caused by a steric hindrance of the Gal*N*Ac-polymer. While the general ranking of the antibodies did not change when switching from calcium- to EDTA-containing buffer, there were nevertheless some significant variations visible in the binding behavior of the samples. Particularly, the stability of the dimer binding seemed to increase slightly, which also led to higher signal intensities (shift to the upper right corner in *figure* 5-33B&C). These differences might be explained by small conformational changes of the H1-CRD when adding or removing calcium, resulting in a changed accessibility of the binding epitopes. However, the relatively high calcium concentration might also influence the binding in a non-specific manner. Therefore, the following screening experiments were performed in calcium-free buffer. Since antibodies with a fast and tight binding profile are preferred for diagnostic or analytical applications, the three antibodies that matched best to this profile (B01.4, C11.1, C18.1) were selected for a further high-resolution screening on both H1-CRD surfaces (*Fig*. 5-34).

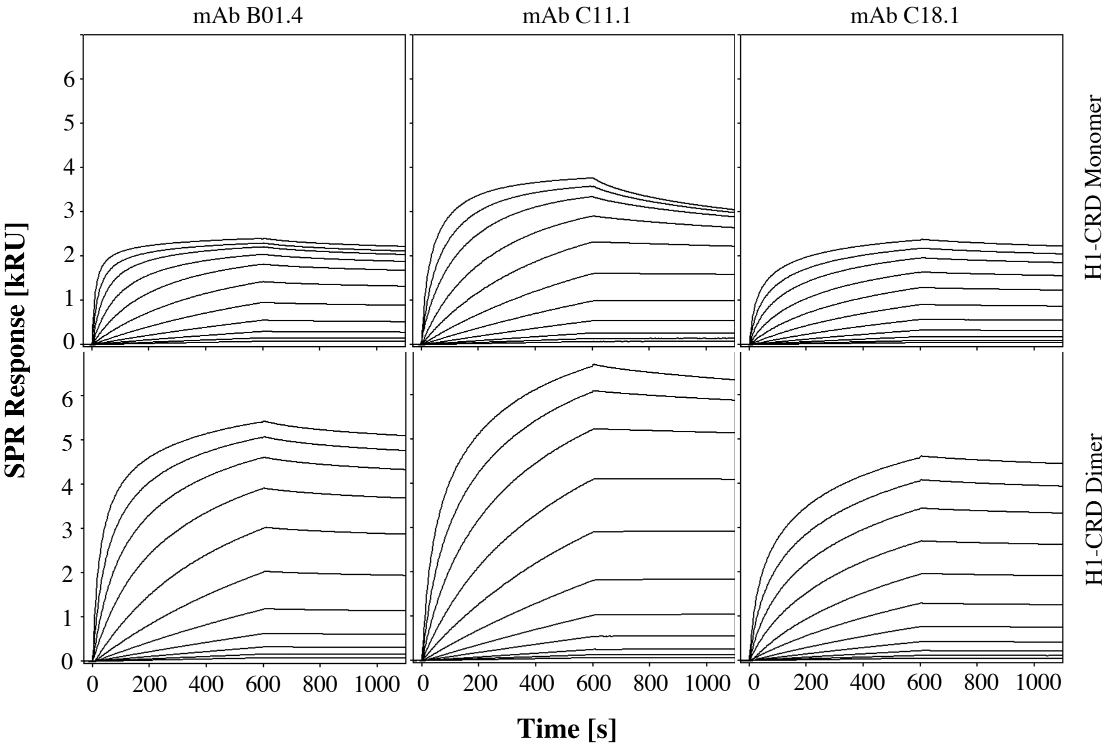


Figure 5-34: Screening of the three monoclonal antibodies B01.4, C11.1, and C18.1 on H1-CRD monomer and dimer surfaces over a concentration range between 140 nM and 140 pM in HBS buffer containing 3 mM EDTA. Association and dissociation phases were acquired for ten minutes each.

In correlation with the data from the ranking experiments, the selected antibodies selectively bound to both H1-CRD surfaces in a clearly concentration-dependent manner. While all antibodies showed a very slow dissociation phase, the signal intensity was significantly different between the samples and the surfaces, despite the molecular weight and concentration was the same for all samples. Interestingly, none of the tested antibodies showed a significantly faster off-rate. This could be influenced by the fact that the antibodies were selected by an ELISA method [30], where only antibodies with a tight binding to the protein withstood the washing conditions. In order to estimate and compare the kinetic properties of the antibodies, a reduced data set (140 pM - 9 nM; six data points) was fitted to a Langmuir 1:1 binding model (*Table* 5-9, *Fig*. 5-35).

Table 5-9: Binding properties of the three screened antibodies calculated from kinetic fits using a simple 1:1 Langmuir binding model with data sets between 140 pM and 9 nM. Values for interactions with the monomer and dimer surface are compared.

Antibody	**K_D [pM] [a]**		**$t_{1/2}$ [min] [b]**	
	Monomer	**Dimer**	**Monomer**	**Dimer**
B01.4	308	334	71	129
C11.1	78	n.d. [c]	447	n.d. [c]
C18.1	233	124	133	346

[a] Calculated as $K_D = k_{off}/k_{on}$. [b] Calculated as $t_{1/2} = \ln 2/k_{off}$. [c] No binding properties could be calculated for antibody C11.1 on the dimer surface because of its very slow dissociation phase.

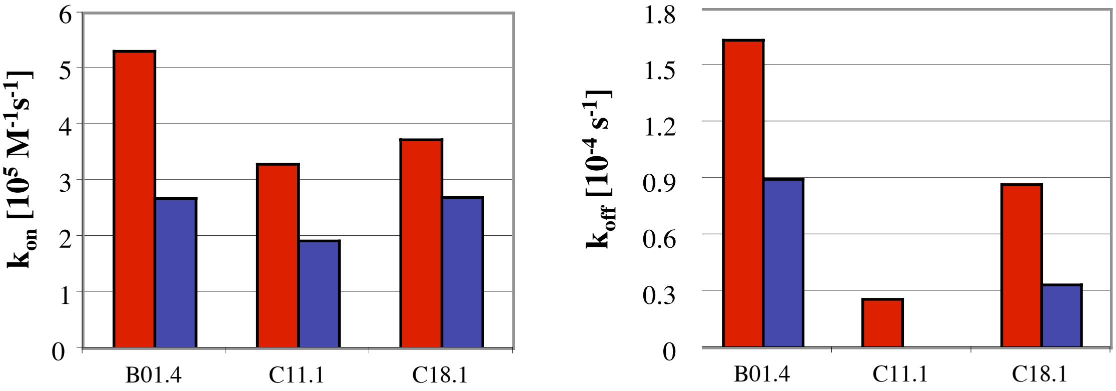


Figure 5-35: Comparison of kinetic on- and off-rates of the three selected antibodies on the H1-CRD monomer (red bars) and dimer (blue bars) surfaces after fitting to a Langmuir 1:1 binding model. No off-rate could be calculated for the antibody C11.1 on the dimer surface, due to the slow dissociation phase.

While the difference in K_D between the monomer and dimer fractions was rather low in the case of B01.4, the dimer affinity of C18.1 was almost doubled. Since the antibody C11.1 showed essentially no dissociation from the dimer surface in the measured time period, no off-rate and therefore no binding properties could be calculated. However, this exceptionally slow dissociation also indicates an increased affinity for the dimer. A closer look at the on- and off-rates showed a general tendency of lower values in case of the dimer interaction. Translated into a mechanic model, this means that the antibodies bind slower but tighter to the dimers. Reduced accessibility (e.g. due to steric hindrance) might be the cause for the reduced on-rate, while the prolonged dissociation may be influenced by the enriched local concentration around the initial binding site, which enforces rebinding or concerted binding. However, while both monomer rate constants for B01.4 are approximately twice as high compared to the dimer values (*Fig.* 5-35), therefore compensating in the calculated K_D (*Eq.* 2, see *section* 2.1.1), the differences in the off-rate seem to be much larger than for the on-rates in case of the other two samples. The comparison of the intensities in the ranking assay (*Fig*. 5-33) and the kinetically obtained affinities also indicates that samples with a fast k_{on} (i.e. B01.4) are overestimated during ranking, since they reach higher intensities in a short injection period.

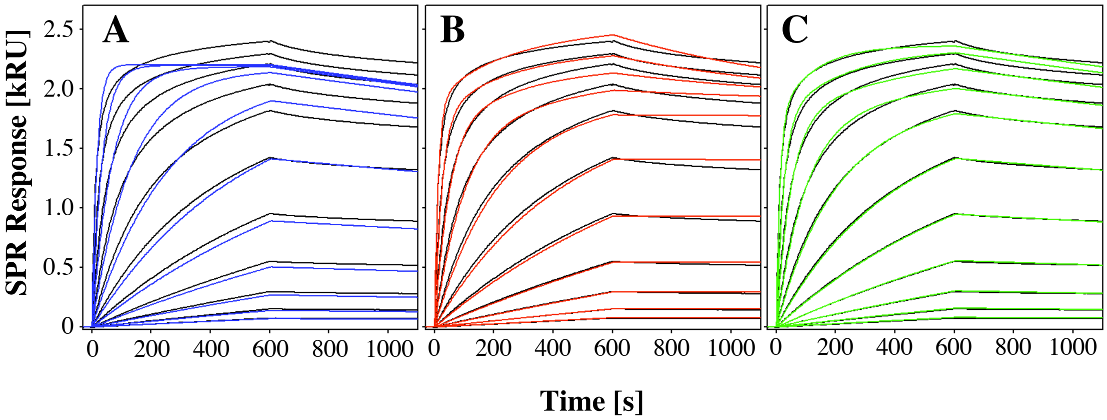


Figure 5-36: Kinetic evaluation of antibody B04.1 on the H1-CRD monomer surface over the full concentration range (140 nM - 140 pM). The same set of sensorgrams was fitted to a 1:1 Langmuir (**A**; blue), bivalent analyte (**B**; red), or surface heterogeneity model (**C**; green).

When the whole data sets of B01.4 (140 pM - 140 nM; 11 data points) were kinetically fitted, a simple Langmuir 1:1 binding model did not adequately describe the interaction anymore (*Fig*. 5-36A). From a theoretical point of view, the interaction should follow a bivalent analyte model, since both Fab' parts of the antibody are able to bind to H1-CRD in a concerted reaction. Indeed, the fitting of the sensorgrams was clearly improved when using this binding model (*Fig*. 5-36B). However, some curves still showed some significant deviations. Therefore, a binding model for a heterogeneous surface was applied in parallel, resulting in even better fits (*Fig*. 5-36C). This might either be explained by variations in accessibility of the binding epitopes due to differences in the immobilizations or by the existence of multiple binding epitopes. Latter hypothesis is supported by mass spectrometric epitope mapping experiments, which showed several linear or structural epitopes [30]. As a consequence, binding of anti-H1-CRD antibodies could be regarded as a complex interaction event, which contains an overlay of both bivalent and heterogeneous elements. Therefore, immobilization of the antibody to the sensor chip should be preferred for further characterization, since the two Fab' domains bind independently to the protein, reducing the kinetic model to a simple 1:1 Langmuir interaction in most cases. A first attempt in this direction was done by capturing B01.4 with a polyclonal rabbit anti-mouse IgG1 antibody and injecting H1-CRD in solution (*Fig*. 5-37).

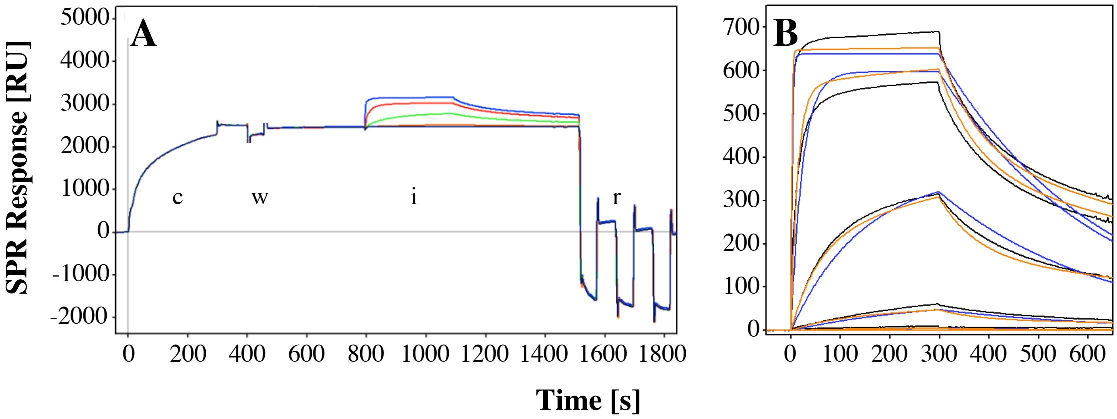


Figure 5-37: H1-CRD screening on the captured antibody B01.4. **A:** Setup of the screening assay beginning with the capturing of B01.4 (c) with a rabbit anti-mouse IgG1 antibody, followed by a short wash step (w), the injection of H1-CRD (i) and three regeneration steps (r). **B:** Kinetic evaluation of the H1-CRD sensorgrams (250 pM - 2.5 μM). Overlay of the binding data (black) with simulated curves of a *Langmuir 1:1* (blue) and a *conformational change* model (red) are shown.

While the capturing assay was very reproducible and generated interpretable sensorgrams (*Fig*. 5-37A), the data did not fit accurately to a standard kinetic model (Langmuir, surface heterogeneity, or bivalent analyte models). A model optimized for conformational changes resulted in the closest fit (*Fig*. 5-37B), although it is highly questionable if this corresponds to the real binding mode. Injection of H1-CRD on the polyclonal capturing antibody alone also generated significant binding responses, indicating a considerable amount of non-specific binding. Therefore, a direct immobilization of the antibodies to be investigated should be preferred.

5.3.6 Evaluation of negative binding responses

By far the most unusual property of the H1-CRD Biacore assay was the occurrence of negative binding responses during the screening of small carbohydrate analytes (see *section* 5.3.5). Nearly no information is available about negative SPR signals upon ligand binding. Observation of such effect often leads to rejection of the data since bad assay design is suspected to be the reason. However, in the case of monosaccharide binding to H1-CRD, data could be mirrored, reproducibly fitted to a single binding site model and were comparable to a solid-phase competition assay as well as to literature data (*Table* 5-7). Since Biacore not only detects mass increases from ligand binding but every change in mass concentration around the gold surface, the protein itself could also contribute to the observed overall signal. Several mechanistic hypotheses seem to be reasonable, including (i) conformational changes, (ii) lectin-interaction with the dextran matrix or (iii) electrostatic interactions with matrix-bound carboxyl groups.

Hypothesis I: conformational change

Essentially two observations support the hypothesis of a conformational change as a possible cause for the negative binding signals in the case of ASGP-R. First, lowering the *pH* to slightly acidic conditions (from *pH* 7.4 to 5.6) induces a conformational change, during which calcium ions and ligands are released. This has been shown e.g. by iodination [28] and mutation experiments [52, 53]. More importantly, ligand binding is also reported to induce a conformational change across the membrane in order to induce the receptor-mediated endocytosis (see *section* 5.1.2) [17, 19]. However, nothing is known about the exact mechanism nor are there any structural data available, which describe these changes.

At the moment, there is no proof that conformational changes can be monitored on SPR instruments, and only sporadic reports of this phenomenon are available [63-68]. The most interesting study in this context has been performed with another lectin, i.e. the maltose-binding protein (MBP) [66]. Injection of different sugars on the immobilized MBP induced a drop of the post-injection baseline, which could be correlated with the known analyte activity. However, no negative responses for the binding equilibria during analyte injection were reported. Even though the signals were rather low ($\leq$ 20 RU), they were fitted to a single site model and resulted in a maltose K_D similar to values obtained by other methods. Interestingly, MBP is known to undergo a substantial conformational change ('venus fly trap'-like hinge twist) upon mannose binding, as it has been shown by crystal structure analysis [69]. Based on the SPR data of MBP and another protein (transglutaminase; see also *section* 3.3.5), they concluded that a decrease in the hydrodynamic radius or volume of a protein might lead to a negative SPR signal and vice versa. However, the question remains open, if the directionality of such changes plays a role for the intensity and sign of the resulting SPR response. For example, a longitudinal stretch of an immobilized protein is expected to induce a more significant change than internal changes or expansions in parallel to the chip surface (*Fig.* 5-38A&B). An additional hint for structural changes being involved in the biding of carbohydrates to H1-CRD comes from the differences between amine- and thiol-coupled lectin surfaces. While Gal*N*Ac bound clearly worse to the thiol-coupled surface than to the amine-coupled surface, the interaction with ASF seemed to be nearly not affected (*Fig*. 5-38C&D).

The overall binding signal is believed to consist of a negative component generated by a conformational change and a positive mass contribution. In the case of ASF, the small negative part is negligible due to the large molecular weight of the glycoprotein (> 40 kDa), while it seems to be dominant for the small monosaccharide (221 Da). Since the overall activity of H1-CRD is not lost after thiol immobilization, as it is demonstrated by ASF-binding, the immobilization method seems to selectively influence the generation of the negative component. This might correlate well with the dramatic loss of signal intensity for Gal*N*Ac (*Fig*. 5-38C) but also for the slightly increased response of ASF on the thiol surface compared to the amine-coupled monomer surface, which should be still capable to combine negative and positive signals (*Fig*. 5-38D).

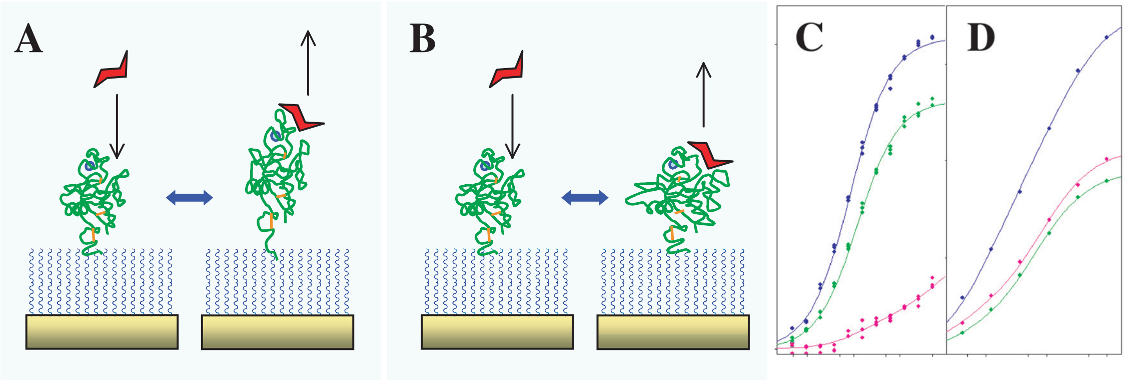


Figure 5-38: Conformational change hypothesis for the generation of negative binding signals. While some conformational changes could induce a negative SPR signal by decreasing the mass concentration (**A**), others are not detectable (**B**). Binding of a monosaccharide (Gal*N*Ac; **C**) and an asialoglycoprotein (ASF; **D**) on amine-coupled monomer (green) and dimer surfaces (blue) as well as to a thiol-coupled H1-CRD surface (magenta).

As in the case of human serum albumin (see *section* 3.3), monitoring of conformational changes with alternative methods like circular dichroism (CD) could confirm the SPR results. However, since H1-CRD contains 11 tryptophan and 6 tyrosine residues distributed over the whole protein chain, small changes at a specific site are much more difficult to follow than in the case of HSA with only one tryptophan residue. Indeed, tryptophan fluorescence experiments did not show any calcium or analyte-induced changes of the CD spectrum in the case of H1-CRD. Again, co-crystallization of the lectin with Gal*N*Ac or another analyte should provide a deeper insight in the binding mechanism.

Hypothesis II: Competitive interactions with the dextran matrix

Another plausible explanation of the negative signal effect is the replacement of a heavier or multiple components by the small analyte. Since the signal returns to the initial value after injection end, this component has to be available in the running buffer. Different buffer ingredients were therefore considered causing the signal effect. HEPES, which was suspected to interact with the calcium binding site, was excluded as a possibility by substituting it with imidazole (see *section* 5.3.4). Calcium ions were the second component to investigate, because they were found to be essential for high signal intensities (*Fig*. 5-25). Since the values between the Biacore assay at 50 mM $CaCl_2$ and the solid-phase competition assay at 1 mM $CaCl_2$ are very similar, the calcium concentration seems to be mainly responsible for the signal generation rather than the affinity. It is therefore possible that several calcium ions (MW = 40 Da) are competitively released upon ligand binding, which results in a negative SPR signal. On the other hand, binding signals of ASF were also dependent on the calcium concentration, indicating that calcium contributes to the binding activity of the proteins. Binding to dextran was therefore considered as another plausible hypothesis, since the carboxymethyl dextran hydrogel of the sensor chip mainly consists of a linear glucose polysaccharide [70]. Even though direct injection of H1-CRD in 10 mM HEPES 50 mM $CaCl_2$ did not show significant binding to a plain CM5 sensor chip surface and injections of free glucose and polymeric glucose resulted in barely detectable signals, the enhanced local concentration around the lectin might lead to a weak but constant binding. Upon ligand binding, the protein is released from the dextran and changes its orientation. After injection end, dextran binding becomes predominant again and the protein switches back (*Fig*. 5-39A). If this hypothesis was true, binding results had to be regarded as competitive (IC_{50}) rather than direct binding (K_D). This hypothesis was further investigated by screening Gal*N*Ac and galactose in running buffer at three different concentrations (0, 1, 10 mg/ml) of soluble carboxymethyl dextran (CMD; *Fig*. 5-39B&C). CMD is also recommended to avoid non-specific binding in Biacore-based ligand fishing assays, e.g. for recovering samples for mass spectrometry [71].

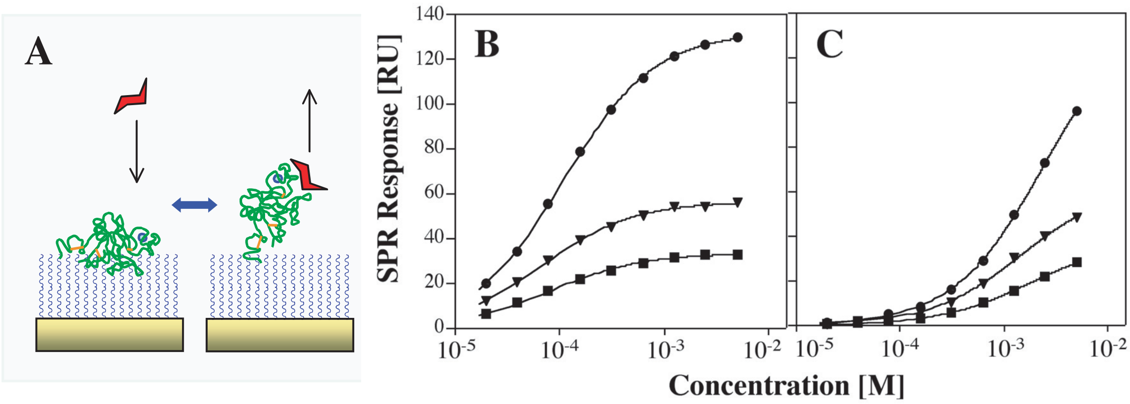


Figure 5-39: Influence of carboxymethyl dextran (CMD) on the negative SPR signals. **A:** Hypothetical model of H1-CRD (green) weakly interacting with the CMD-matrix (blue) of the sensor chip. Upon binding of a carbohydrate analyte (red) the protein is released from the matrix. **B, C**: Mirrored negative responses of Gal*N*Ac (**B**) and D-galactose (**C**) without CMD (■), with 1 mg/ml CMD (▼), and with 10 mg/ml CMD (●) added to the buffer.

The addition of CMD clearly influenced the generation of the negative binding responses for both galactose and Gal*N*Ac (*Fig.* 5-39B&C). While 1 mg/ml CMD increased R_{max} of Gal*N*Ac approximately twofold, the amplification was much greater for 10 mg/ml CMD (~ factor 5-6). It is therefore believed that soluble CMD interacts with the sugar binding site in absence of any carbohydrate analytes. During sugar injection, the weakly binding but heavy CMD is rapidly replaced by the small analyte, resulting in a signal drop, but interacts again after injection end. Unfortunately, neither the dissociation constant nor the exact molecular weight of the CMD used for the experiments is known (approximately 12 kDa, according to the manufacturer). However, an experimental series of constant concentrations of pure dextran at different well-defined size ranges might give additional evidence for this hypothesis. In addition, the hydroxyl groups of the sensor chip's CMD-matrix could be derivatized prior to protein immobilization in order to inhibit lectin-interactions. On the other side, addition of soluble CMD might be useful to amplify signal intensities of analytes, especially if their molecular weight is in a critical range, where negative and positive signal components are compensating each other.

5.4 Conclusions

Targeting the human asialoglycoprotein receptor (ASGP-R) is a powerful approach for a specific treatment of the liver by conventional or genetic means. Even though several methods for the generation of potential ligand molecules have been described, little is known about the molecular level of the receptor's binding properties. Therefore, a Biacore assay was developed based on the carbohydrate recognition domain (CRD) of the H1 subunit of ASGP-R. By optimizing an existing method for the expression of the H1-CRD of the human hepatic asialoglycoprotein receptor, the lectin was obtained in remarkably increased yields, full activity and high purity. However, dimerization of the H1-CRD with significant variations in the monomer/dimer ratio was still observed. Therefore, different chromatographic methods were developed for separating the two species using high-performance liquid chromatography (HPLC). While size exclusion and reversed-phase chromatography were found to be less suitable for this purpose, weak anion exchange on a DEAE column showed a successful and highly reproducible separation of the two fractions. In addition, remaining minor impurities after expression and purification could also be removed by this method. Even though affinity chromatography could not be used for separation purposes, it was identified as an important method for simultaneously concentrating H1-CRD fractions, testing their activities and performing a buffer exchange.

Mass spectrometric characterization of H1-CRD monomer and dimer fractions clearly proofed their identity and sequence. It also showed that the *N*-terminal methionine residue was posttranslationally removed during expression in *E. coli*. Differences between the monomer spectra under reducing and non-reducing conditions further indicated that a single free cysteine residue, which is not involved in disulfide bridge formation, seems to be responsible for dimerization. Addition of thiol-active compounds (2-mercaptoethanol, glutathione) during the refolding process leads to the formation of mixed disulfides, which is believed to prevent dimerization.

H1-CRD monomer and dimer fraction could be successfully and actively immobilized to a Biacore sensor chip. While amine coupling generated stable surfaces in a highly reproducible manner, thiol coupling could be also applied to immobilize the protein at its free cysteine residue. However, thiol-coupled surfaces required elongated immobilization times and were less stable, when compared to amine-coupled ones. Based on the immobilized H1-CRD, a Biacore binding assay could be developed, which not only allowed the screening of carbohydrate ligands but also generated information about the binding mechanism and properties of the protein.

Two asialoglycoproteins and a set of mono- and disaccharides were used to evaluate the binding specificity of the Biacore assay. As expected, all binding events were found to be strongly calcium-dependent. While the two glycoproteins showed positive SPR signals in the expected range of intensity, the small carbohydrates reproducibly induced a negative binding signal. Even though this effect is completely atypical for small molecule interactions, the negative sensorgrams were generated in a clearly concentration-dependent manner. Mathematical mirroring of the sensorgrams led to steady state binding plots, which could perfectly described by a single binding site model. The obtained K_D values were in a very good agreement with data from both a solid-phase competition assay and literature. In addition, the rather low affinities in the micromolar range as well as the fast kinetic properties corresponded well with the common knowledge about carbohydrate-protein interactions. With a K_D of 100-150 μM, Gal*N*Ac was found to be the monosaccharide showing the highest affinity for H1-CRD. Galactose and some of its derivatives bound clearly weaker (1-4 mM), while glucose did not bind to the lectin with reasonable affinity. Only small differences between the binding strength to the H1-CRD monomer and dimer fractions could be detected (less than a factor 2). The addition of DMSO did not influence the binding properties considerably, therefore qualifying the assay for the screening of more hydrophobic, drug-like molecules. The genesis of the negative SPR signals could not be determined with absolute certainty. However, there is strong evidence that both calcium ions and weak interactions of the CRD with the dextran matrix of the sensor chip might contribute to this phenomenon. In addition, conformational changes as they are reported to occur upon ligand binding or during a change of environmental *pH*, are also a possible cause.

Due to their multivalent binding properties, the interaction profiles of the two asialoglycoproteins was clearly different from the monovalent sugars. Slightly increased on-rates combined with clearly slower dissociation rates led to affinities in the picomolar range, as shown for asialofetuin. With a factor of four, the affinity difference between the monomer and dimer surfaces was clearly more pronounced, indicating a strong contribution of local concentration, avidity and rebinding. Asialoorosomucoid could be successfully produced by desialylation of α_1-acid glycoprotein and showed a similar but more complex binding mode than asialofetuin.

Therefore, no exact affinity and kinetic values could be obtained for this interaction. A more extended purification or characterization of the glycoprotein seems to be required.

Calcium concentration and *pH* were shown to be important for a successful interaction of lectin. Any reduction or removal of calcium dramatically decreased the binding signals of all tested ligands. This effect was more prominent in the case of the small, monovalent carbohydrates. With 50 mM $CaCl_2$, the optimal calcium concentration for the Biacore assay was found to be clearly higher than the physiological concentration (1 mM). By injecting calcium chloride to immobilized H1-CRD, two independent binding sites with K_D values of 170 and 16000 μM could be detected, which is in agreement with literature. When the *pH* was lowered from 7.4 (plasma) to 5.5 (endosome), the protein basically lost its binding activity. Further experiments showed that a release of calcium ions below *pH* 6 is responsible for this effect. These observations clearly fit with the described mechanism of receptor-mediated endocytosis, where calcium and ligands are released in the early endosome (*pH* ~ 6).

Furthermore, the Biacore assay was successfully used to rank and characterize a set of monoclonal antibodies binding to the human ASGP-R H1-CRD. Since these antibodies were pre-selected by an ELISA method, all tested samples showed a very tight binding to the lectin with dissociation half-times of 1-7.5 hours and K_D values in the picomolar range. Similar to the asialoglycoproteins, the kinetic profile did not follow a simple 1:1 binding mode and the interaction was significantly stronger for the dimer surface. On the other side, the calcium dependency was far smaller compared to the carbohydrates and glycoproteins. None of the antibodies showed a significant blocking behavior in the ranking assay. The H1-CRD Biacore assay was therefore shown to be a valuable alternative to the ELISA for the evaluation and characterization of monoclonal antibodies during production.

The evaluation of the assay with a broad spectrum of natural ligands, from small sugars to asialoglycoproteins and monoclonal antibodies, clearly showed that immobilized H1-CRD can be used for the screening and characterization of carbohydrate-lectin interactions as well as antibody-antigen interactions. The kinetic profiles of the substance classes showed significant differences, which all were within expectations. Comparison of the affinity values with available literature data and alternative binding assays resulted in a very good agreement. Together with the observed interaction properties like calcium and *pH* dependency, these findings confirm the validity of the H1-CRD Biacore assay, despite the unusual occurrence of negative SPR signals or the high calcium requirements. Further development steps should therefore concentrate on the origin of the negative binding responses and on the reduction of the calcium concentration to near physiological values.

5.5 References

[1] S. Ishibashi, R. E. Hammer, J. Herz, *J Biol Chem* 1994, *269*, 27803.

[2] Y. C. Lee, R. T. Lee, in *Carbohydrates in Chemistry and Biology, Vol. 4*, 1st edition ed. (Eds.: B. Ernst, G. W. Hart, P. Sinay), Wiley-VCH, Weinheim, 2000, pp. 549.

[3] G. Ashwell, A. G. Morell, *Adv Enzymol Relat Areas Mol Biol* 1974, *41*, 99.

[4] J. M. Wilson, M. Grossman, C. H. Wu, N. R. Chowdhury, G. Y. Wu, J. R. Chowdhury, *J Biol Chem* 1992, *267*, 963.

[5] U. Westerlind, J. Westman, E. Tornquist, C. I. Smith, S. Oscarson, M. Lahmann, T. Norberg, *Glycoconj J* 2004, *21*, 227.

[6] T. Yao, S. Degli Esposti, L. Huang, R. Arnon, A. Spangenberger, M. A. Zern, *Am J Physiol* 1994, *267*, G476.

[7] E. A. Biessen, A. R. Valentijn, R. L. De Vrueh, E. Van De Bilt, L. A. Sliedregt, P. Prince, M. K. Bijsterbosch, J. H. Van Boom, G. A. Van Der Marel, P. J. Abrahams, T. J. Van Berkel, *Faseb J* 2000, *14*, 1784.

[8] P. C. Rensen, S. H. van Leeuwen, L. A. Sliedregt, T. J. van Berkel, E. A. Biessen, *J Med Chem* 2004, *47*, 5798.

[9] J. Wu, M. H. Nantz, M. A. Zern, *Front Biosci* 2002, *7*, d717.

[10] M. Nishikawa, *Biol Pharm Bull* 2005, *28*, 195.

[11] L. Wang, S. Kaneko, M. Honda, K. Kobayashi, *Virus Res* 2002, *85*, 187.

[12] I. Geffen, M. Spiess, *Int Rev Cytol* 1992, *137B*, 181.

[13] M. Meier, M. D. Bider, V. N. Malashkevich, M. Spiess, P. Burkhard, *J Mol Biol* 2000, *300*, 857.

[14] H. F. Lodish, *Trends Biochem Sci* 1991, *16*, 374.

[15] K. B. Chiacchia, K. Drickamer, *J Biol Chem* 1984, *259*, 15440.

[16] K. Drickamer, *J Biol Chem* 1988, *263*, 9557.

[17] M. D. Bider, M. Spiess, *FEBS Lett* 1998, *434*, 37.

[18] C. Fuhrer, I. Geffen, K. Huggel, M. Spiess, *J Biol Chem* 1994, *269*, 3277.

[19] C. Fuhrer, I. Geffen, M. Spiess, *J Cell Biol* 1991, *114*, 423.

[20] J. U. Baenziger, Y. Maynard, *J Biol Chem* 1980, *255*, 4607.

[21] Y. C. Lee, R. R. Townsend, M. R. Hardy, J. Lonngren, J. Arnarp, M. Haraldsson, H. Lonn, *J Biol Chem* 1983, *258*, 199.

[22] A. Kichler, F. Schuber, *Glycoconj J* 1995, *12*, 275.

[23] M. D. Bider, J. M. Wahlberg, R. A. Kammerer, M. Spiess, *J Biol Chem* 1996, *271*, 31996.

[24] Y. C. Lee, R. R. Townsend, M. R. Hardy, J. Lonngren, K. Bock, in *Biochemical and Biophysical Studies of Proteins and Nucleic Acids* (Eds.: T. B. Lo, T. Y. Liu, C. H. Li), Elsevier, New York, 1984, pp. 349.

[25] R. T. Lee, P. Lin, Y. C. Lee, *Biochemistry* 1984, *23*, 4255.

[26] K. G. Rice, O. A. Weisz, T. Barthel, R. T. Lee, Y. C. Lee, *J Biol Chem* 1990, *265*, 18429.

[27] T. T. Andersen, J. W. Freytag, R. L. Hill, *J Biol Chem* 1982, *257*, 8036.

[28] M. DiPaola, F. R. Maxfield, *J Biol Chem* 1984, *259*, 9163.

[29] M. D. Bider, R. Cescato, P. Jeno, M. Spiess, *Eur J Biochem* 1995, *230*, 207.

[30] R. Born, *Production and Characterization of Antibodies against the Human Hepatic Asialoglycoprotein Receptor,* PhD thesis, University of Basel (Basel, Switzerland), 2005.

[31] M. D. Bider, M. Spiess, *Expression, characterization, and crystallization of the carbohydrate recognition domain of subunit H1 of the asialoglycoprotein receptor*, unpublished manuscript, 1997.

[32] U. K. Laemmli, *Nature* 1970, *227*, 680.

[33] N. Fornstedt, J. Porath, *FEBS Lett* 1975, *57*, 187.

[34] M. L. Connolly, *Science* 1983, *221*, 709.

[35] D. Stokmaier, *Receptor-mediated endocytosis in human hepatocytes,* diploma thesis, University of Basel (Basel, Switzerland), 2004.

[36] R. G. Spiro, *J Biol Chem* 1960, *235*, 2860.

[37] E. A. Popenoe, R. M. Drew, *J Biol Chem* 1957, *228*, 673.

[38] D. G. Myszka, *J Mol Recognit* 1999, *12*, 279.

[39] *BIAevaluation Software Handbook: version 3.0*, July 1997 ed., Biacore AB, Uppsala, 1997.

[40] G. Weitz-Schmidt, D. Stokmaier, G. Scheel, N. E. Nifant'ev, A. B. Tuzikov, N. V. Bovin, *Anal Biochem* 1996, *238*, 184.

[41] P. Stanton, *Methods Mol Biol* 2004, *251*, 55.

[42] D. Carr, *The Handbook of Analysis and Purification of Peptides and Proteins by Reversed-Phase HPLC*, Third Edition ed., Grace Vydac, Hesperia CA, 2002.

[43] P. Stanton, *Methods Mol Biol* 2004, *251*, 23.

[44] S. Era, T. Hamaguchi, M. Sogami, K. Kuwata, E. Suzuki, K. Miura, K. Kawai, Y. Kitazawa, H. Okabe, A. Noma, *et al.*, *Int J Pept Protein Res* 1988, *31*, 435.

[45] T. Hayashi, S. Era, K. Kawai, H. Imai, K. Nakamura, E. Onda, M. Yoh, *Pathophysiology* 2000, *6*, 237.

[46] S. Rabbani, I. Gasser, B. Ernst, *unpublished results.*

[47] T. Meinnel, Y. Mechulam, S. Blanquet, *Biochimie* 1993, *75*, 1061.

[48] I. E. Gentle, D. P. De Souza, M. Baca, *Bioconjug Chem* 2004, *15*, 658.

[49] *Delta Mass: A Database of Protein Post-Translational Modifications*, The association of biomolecular research facilities, http://www.abrf.org/index.cfm/dm.home, 2005.

[50] V. V. Papov, S. A. Gravina, J. J. Mieyal, K. Biemann, *Protein Sci* 1994, *3*, 428.

[51] S. J. Stahl, P. T. Wingfield, J. D. Kaufman, L. K. Pannell, V. Cioce, H. Sakata, W. G. Taylor, J. S. Rubin, D. P. Bottaro, *Biochem J* 1997, *326 (Pt 3)*, 763.

[52] S. Wragg, K. Drickamer, *J Biol Chem* 1999, *274*, 35400.

[53] H. Feinberg, D. Torgersen, K. Drickamer, W. I. Weis, *J Biol Chem* 2000, *275*, 35176.

[54] M. G. Yet, C. C. Chin, F. Wold, *J Biol Chem* 1988, *263*, 111.

[55] E. D. Green, G. Adelt, J. U. Baenziger, S. Wilson, H. Van Halbeek, *J Biol Chem* 1988, *263*, 18253.

[56] K. Nagata, H. Handa, *Real-Time Analysis of Biomolecular Interactions: Applications of BIACORE*, Springer, Tokio, 2000.

[57] D. T. Connolly, R. R. Townsend, K. Kawaguchi, W. R. Bell, Y. C. Lee, *J Biol Chem* 1982, *257*, 939.

[58] T. Weimar, B. Haase, T. Kohli, *J Carbohydrate Chemistry* 2000, *19*, 1083.

[59] Y. Shinohara, Y. Hasegawa, H. Kaku, N. Shibuya, *Glycobiology* 1997, *7*, 1201.

[60] R. T. Lee, R. W. Myers, Y. C. Lee, *Biochemistry* 1982, *21*, 6292.

[61] D. Yi, R. T. Lee, P. Longo, E. T. Boger, Y. C. Lee, W. A. Petri, Jr., R. L. Schnaar, *Glycobiology* 1998, *8*, 1037.

[62] N. V. Bovin, *Glycoconj J* 1998, *15*, 431.

[63] H. Sota, Y. Hasegawa, M. Iwakura, *Anal Chem* 1998, *70*, 2019.

[64] Z. Salamon, S. Cowell, E. Varga, H. I. Yamamura, V. J. Hruby, G. Tollin, *Biophys J* 2000, *79*, 2463.

[65] S. Boussaad, J. Pean, N. J. Tao, *Anal Chem* 2000, *72*, 222.

[66] J. E. Gestwicki, H. V. Hsieh, J. B. Pitner, *Anal Chem* 2001, *73*, 5732.

[67] T. Mannen, S. Yamaguchi, J. Honda, S. Sugimoto, A. Kitayama, T. Nagamune, *Anal Biochem* 2001, *293*, 185.

[68] S. Paynter, D. A. Russell, *Anal Biochem* 2002, *309*, 85.

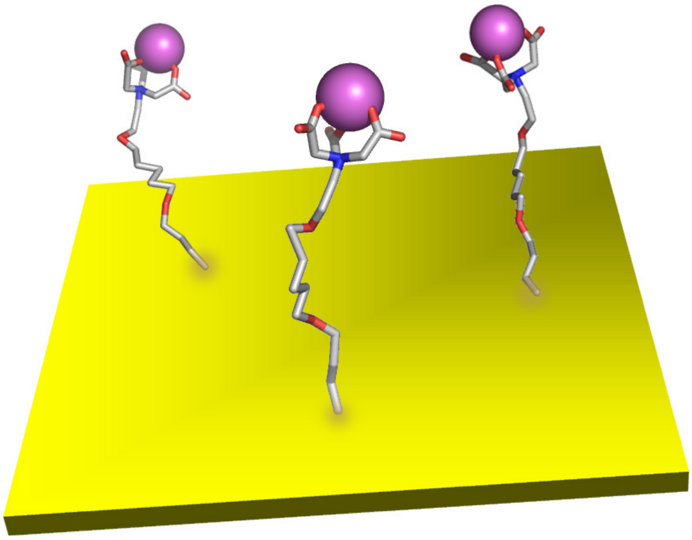
[69] S. L. Mowbray, M. O. Sandgren, *J Struct Biol* 1998, *124*, 257.

[70] S. Lofas, B. Johnson, *J Chem Soc Chem Commun* 1990, 1526.

[71] *Recovering samples from Biacore® 3000 for mass spectrometry (Technology Note 2)*, Biacore AB, 2002.

Chapter 6

Hexahistidine Tag



6.1 Introduction

The purification of recombinant proteins and peptides is an important challenge in life science. A widely used approach in this context is the introduction of short peptide tags for a selective capturing of the protein on a chromatographic support. For this purpose, the hexahistidine tag shares the widest acceptance. Despite the successful usage of this tag in many applications, little is known about the exact binding mechanism.

6.1.1 The hexahistidine tag – a successful story

Some thirty years ago, the idea arose of using surface-bound chelators like iminodiacetic acid (IDA; *Fig.* 6-1A) to capture bivalent transition metal ions (Ni^{2+}, Zn^{2+}, Co^{2+}, and Cu^{2+}) for the single-step isolation and purification of proteins [1]. This method relies on the coordinative binding of certain amino acid residues of the protein to be purified (most likely histidine and to a lower extent also cysteine residues) to the free valencies of the transition metal ion [2-4]. Today known as *immobilized-metal affinity chromatography* (IMAC), this technology is not limited to chromatographic purposes but can also be applied to other areas such as surface plasmon resonance (SPR) [5] or electrochemistry [6], where it allows to predict and control the orientation of proteins immobilized at different interfaces. The use of IMAC exhibits a number of advantages over other biospecific interactions such as enzymes and their cofactors, receptors and their ligands, or antibodies and their antigens. The benefits are high protein loading, mild elution conditions, simple regeneration and low costs. These factors are decisive when large-scale purification procedures for industrial applications have to be developed.

The pioneering work of Hochuli and his coworkers [7, 8] broadened the technique for the efficient purification of recombinant proteins with metal-affinity. First, they described a new chelator, nitrilotriacetic acid (NTA; *Fig.* 6-1B), which is able to coordinate Ni^{2+} with high stability leaving two coordination sites for ligand binding [7]. Second, Hochuli investigated a number of fusion proteins bearing 2-6 consecutive histidine residues at either the *C*- or *N*-terminus of dihydrofolate reductase [8]. A tag consisting of six histidines was found to result in the highest purity and yield of the protein, and was therefore recommended for the purification of recombinant proteins [8]. Since it is rather uncommon that numerous histidines are expressed in close proximity at the surface of naturally occurring proteins, such oligohistidine affinity tags guarantee a high selectivity. Nowadays, the hexahistidine tag is the most commonly used system for the purification of recombinant proteins and is applied to proteins from several expression systems, from bacteria and yeast to insect and mammalian cells [9]. As a consequence, several hundred His-tagged structures are deposited in the PDB [10] and other protein databases.

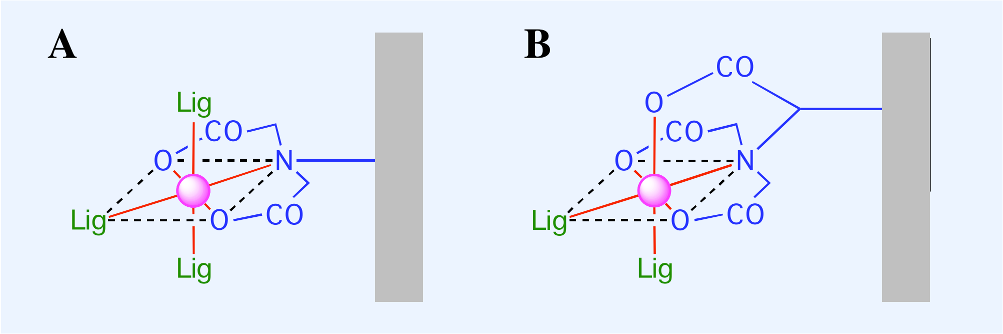


Figure 6-1: Coordination of a hexavalent metal ion like Ni^{2+} to iminodiacetic acid (IDA; **A**) and nitrilotriacetic acid (NTA; **B**). The two chelators (blue) are bound to the metal (magenta) and are fixed on a solid support (grey), leaving two and three coordination sites for ligands (green), respectively. Metal coordination bonds are represented in red. The illustrations are adapted from Hochuli *et al.* [7].

Although the His tag technology has become a standard procedure for the purification of recombinant proteins, the molecular basis of the metal ion chelating properties is still not fully understood. Only rare cases are known, where dissociation constants or even kinetic rate constants (k_{on} / k_{off}) of oligohistidines have been determined [5]. More data are available from experiments with His tag-fusion proteins captured on Ni^{2+}-NTA biosensor chips [5, 11]. Single-molecule experiments using scanning force microscopy revealed that His tags form different types of complexes with significantly different stabilities and energy landscapes along their force-driven dissociation pathways [12, 13]. Only recently, elucidation of a possible binding mechanism of metal ions to various His tag motifs (ranging from His2 to His6) was performed by molecular simulations [14].

6.1.2. The need for further developments

Although IMAC technology is widely used in protein purification, the application still has its limits. The purification and especially the immobilization of recombinant proteins often suffers from the relatively low affinity of the His tag due to unfavorable steric conditions. Furthermore, the introduction of extended tags might elicit undesired changes in protein properties such as decrease of solubility [15], misfolding [16] or dimerization [17]. For pharmaceutical-grade proteins, where the native structure is usually required, the His tag has to be removed chemically or enzymatically. The fact that the methodology for the final removal of the tag by chemical means (i.e. a chemically cleavable linker) is not yet developed, is a evident drawback. In addition, the His tag strategy is hardly ever used for the production of synthetic peptides, because additional synthetic efforts were necessary for the introduction of the tag. A possible application of a peptide purification system, which makes use of improved metal-affinity and linker technologies, is shown in *figure* 6-2.

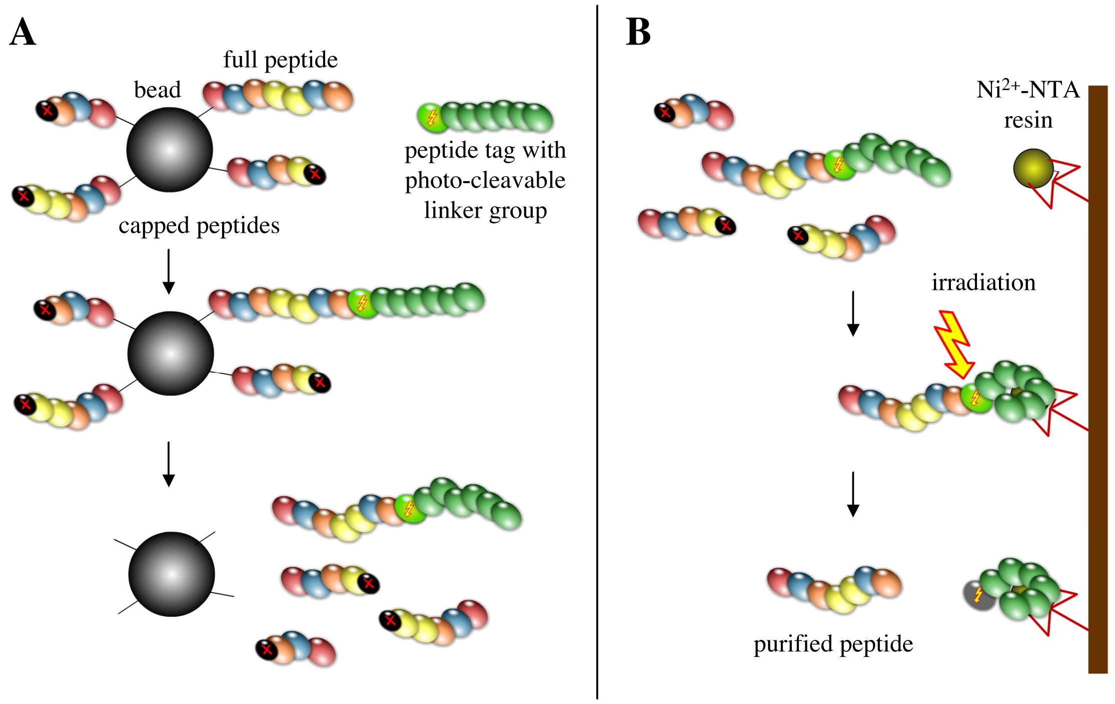


Figure 6-2: Application of improved peptide tags for the purification of synthetic peptides from solid-phase synthesis. A metal-affinity peptide tag (e.g. hexahistidine) with an attached photo-cleavable linker group is coupled as a final step of the synthesis. Incomplete, end-capped peptides are not coupled to the tag by this method (**A**). After cleavage of the peptides from the resin, the mixture is applied to a metal-affinity column (e.g. Ni^{2+}-NTA), where only the tagged peptides are binding. By irradiating the column, the linker releases the peptide in a pure and active manner (**B**).

While the His tag technology was successfully applied to the immobilization of recombinant proteins on sensor chips of SPR biosensors [5], the stability of the captured surfaces is still a problem. With an estimated K_D of 10^{-6} M, a single 6-His tag was found to be insufficient for a stable binding and the use of two consecutive hexahistidine tags was recommended [5]. Even with this modification, surface bleeding might remain an obstacle, especially when measuring small molecules. As a consequence, capturing on NTA sensor chips only contribute to 2% of all SPR experiments published in 2003 [18]. A possible workaround is the use of anti-hexahis or anti-pentahis antibodies [19], which represented 10% of the capturing approaches in 2003 [18]. Interactions of histidine tags with monoclonal antibodies usually show K_D values in the range of 10^{-7} to 10^{-8} M [19]. Hence, an improvement of the chelating properties would certainly increase the use of IMAC for SPR experiments.

An advanced tag technology would have a major impact on drug design and development. Tags with improved chelating properties would not only lead to improved purification procedures of recombinant proteins but also open up the possibility for the site-directed, oriented, and stable immobilization of recombinant proteins on analytical surfaces. This would allow rapid determinations of thermodynamic and kinetic binding properties of small molecules, e.g. by target-based binding assays, surface plasmon resonance experiments or atomic force microscopy.

6.1.3 Aims in this project

Increasing the knowledge about the binding process of oligohistidines to Ni^{2+}-NTA surfaces is the major aim in this project. Biacore technology can be used for the determination of dissociation constants and the qualitative analysis of the binding event. For this purpose, oligohistidine tags are not investigated as fusion proteins but as isolated peptides. The effect of peptide length on binding properties is examined in order to proof the existence of an ideal length concerning the binding affinity. In addition, the optimal distance of two histidine residues is determined by the preparation and analysis of several mixed histidine-alanine hexapeptides. This project was performed in collaboration with Steven Knecht (Institute of Molecular Pharmacy, University of Basel, Switzerland).

6.2 Materials and Methods

This section describes materials, equipment and procedures specifically used for the HisTag project. Materials and general methods used in all Biacore assays are described in *section* 2.2.

6.2.1 Materials

Reagents

9-Fluorenylmethoxycarbonyl- (Fmoc) protected His(Trt)-NovaSyn TGT resin and 1-hydroxybenzotriazole (HOBt) were purchased from NovaBiochem (VWR International AG, Lucerne, Switzerland), dihistidine (His2), Fmoc-protected His(Trt) and Ala as well as 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborate (TBTU) from Bachem (Bachem AG, Bubendorf, Switzerland). All solvents used for the peptide synthesis were purchased from PerSeptive Biosystems or Applied Biosystems. HPLC-grade water, acetonitrile, and trifluoroacetic acid (TFA) were used during peptide purification (Fluka AG, Buchs, Switzerland).

Equipment

A Pioneer automated peptide synthesizer was used for the synthesis of both oligohistidine and His_2Ala_4-peptides. Oligohistidine peptides were purified on a combination of a Jasco PV-980 intelligent HPLC pump system and a Jasco UV-1570 intelligent UV/VIS detector (Jasco GmbH, Gross-Umstadt, Germany) equipped with a preparative C18 column (SymmetryPrep, 19 × 150 mm, 7 µm; Waters AG, Rupperswil, Switzerland). HPLC purification of His_2Ala_4 peptides was performed on an Agilent 1100 purification system, equipped with a quaternary pump, a cooled well-plate autosampler, a column thermostat, a DAD detector, and a cooled analytical fraction collector (Agilent AG, Basel, Switzerland) using a Phenomenex Jupiter C18 column (4.6 × 150 mm, 5 μm; Brechbühler AG, Schlieren, Switzerland). Mass spectrometric analysis of the peptides was performed either on a Waters micromass ZQ (negative ionization mode) or a Finnigan LCQ Deca system (positive ionization mode). NTA sensor chips were used for all Biacore assays (Biacore AB, Uppsala, Sweden).

6.2.2 Synthesis and purification of oligohistidine peptides

Except for the dihistidine, the whole series of oligohistidines (His3-His10), each of them with a free *N*-terminal amine and a *C*-terminal acid, were synthesized using continuous-flow technology and Fmoc-strategy. The Fmoc-group was removed with 20% piperidine (v/v) in dimethylformamide (DMF) and the resin was subsequently washed with pure DMF. Coupling steps were performed using 0.5 M DIPEA and TBTU/HOBt in 0.5 M DMF as activator solutions. Fully protected peptide products were cleaved from the resin using a cleavage mixture containing 5% thioanisole, 4.5% water and 0.5% ethane-1,2-dithiol in TFA (all v/v) and then concentrated in the rotary evaporator before precipitation with ice-cold *tert*-butyl methyl ether. Crude peptides were purified by reversed-phase HPLC. All peptides except trihistidine were purified with a gradient of acetonitrile in water (0-35%, containing 0.1 % TFA). For H-$(His)_3$-OH the aqueous phase had to be changed to 10 mM ammonium acetate *pH* 8.8 to get longer retention on the column with the same gradient. Major peaks were collected and analyzed by mass spectrometry in the negative or positive ionization mode.

6.2.3 Oligohistidine binding assay

In order to avoid non-specific binding and increase the detection sensitivity, increasing amounts of EDTA and polysorbate were added to the different running buffers. 10 mM HEPES, 150 mM NaCl, 50 μM EDTA, *pH* 7.4 was used as *eluent buffer* and 10 mM HEPES, 150 mM NaCl, 3 mM EDTA, 0.005% polysorbate 20 (HBS-EP) as *dispensor buffer*. While the eluent buffer was connected to the left pump of Biacore 3000, responsible for the constant flow and sample injection, the right pump (sample preparation, wash steps) was attached to the dispensor buffer. Tenfold dilution series of oligohistidines were done in eluent buffer according to *table* 6-1 and were freshly prepared before each experiment.

Table 6-1: Oligohistidine peptides used in the Biacore assay with their molecular weight (MW) and the injected concentration range (tenfold linear dilution series).

Analyte	Sequence	MW [Da]	Conc. Range
Dihis	HH	292.3	500 nM - 5 mM
Trihis	HHH	429.4	50 nM - 500 μM
Tetrahis	HHHH	566.6	5 nM - 50 μM
Pentahis	HHHHH	703.7	500 pM - 5 μM
Hexahis	HHHHHH	840.9	500 pM - 5 μM
Heptahis	HHHHHHH	978.0	500 pM - 5 μM
Octahis	HHHHHHHH	1115.1	500 pM - 5 μM
Nonahis	HHHHHHHHH	1252.3	500 pM - 5 μM
Decahis	HHHHHHHHHH	1389.6	500 pM - 5 μM

At the beginning of each cycle, a single flow cell of the NTA chip was loaded with nickel by injecting 500 μM $NiCl_2$ in eluent buffer for 1 min at a flow rate of 20 μl/min. A second flow cell was left unloaded and served as a reference cell. After injection and dissociation of the sample for 5 min each at a flow rate of 20 μl/min, the surface was regenerated by injecting two consecutive 1 min pulses of 500 mM imidazole in water and regeneration buffer (10 mM HEPES *pH* 7.4, 150 mM NaCl, 350 mM EDTA, 0.005% polysorbate 20) at a flow rate of 100 μl/min. In case of the larger peptides (His7-10) an additional 1 min injection of 0.5% SDS in water and the whole washing procedure had to be performed twice for each cycle to avoid carry-over effects. Each concentration was measured in triplicates in a randomized manner. Three buffer blanks were injected before each experiment and one between each series, and were used for double referencing in *Scrubber* (see *section* 2.2.6).

6.2.4 Synthesis and purification of His_2Ala_4 peptides

The synthesis of the His_2Ala_4 peptides was performed as described for the oligohistidine peptides. A series of peptides composed of two histidine and four alanine residues was prepared in this way. While one of the histidine residues had a fixed position at the *C*-terminus (His-resin), the second histidine was consecutively shifted from a vicinal position to the *N*-terminus (*Table* 6-2).

Table 6-2: Mixed His_2Ala_4 peptide series with their sequence and molecular weight.

Analyte	Sequence	MW [Da]
HisAla1	AAAAHH	576.6
HisAla2	AAAHAH	576.6
HisAla3	AAHAAH	576.6
HisAla4	AHAAAH	576.6
HisAla5	HAAAAH	576.6

Due to their small size and relatively high hydrophilicity, a separation of these peptides with a traditional reversed-phase peptide purification approach (water / acetonitrile / TFA) was not possible. Therefore, the peptides were separated under basic conditions, above their theoretical *pI* of 6.92 (calculated using the *PeptideMass* tool [20]). For this purpose, a silica-based C18 column with extended *pH* tolerability was selected and equilibrated with 10 mM ammonium acetate buffer *pH* 8.8. After sample injection and an isocratic phase of 2 min, a linear gradient to 5% acetonitrile was applied within 20 min. The relatively wide peaks were completely collected and the solvents and ammonium acetate were removed by lyophilization over night.

6.2.5 His_2Ala_4 peptide binding assay

For the screening of the His_2Ala_4 peptides, the same experimental setup as described in *section* 6.2.3 was used. The purified peptides were diluted in eluent buffer to a stock concentration of 25 mM, and fivefold linear dilution series between 0.32-5000 μM were prepared. After loading a single flow cell with nickel, the peptide samples were injected for 1 min with a dissociation phase of 20 s at a flow rate of 20 μl/min. The surface was regenerated with a 1 min pulse of regeneration buffer. The signals of an unloaded NTA flow cell were used for referencing and buffer blank injection before (3 blanks) and between (1 blank) the cycles for double referencing (see *section* 2.3.6).

6.3 Results & Discussion

6.3.1 Synthesis and purification of oligohistidine peptides

For any synthesized peptide, a yield of at least 40% was achieved after HPLC purification. Much higher concentrations of the injected solution could be reached with precipitated peptides than with directly lyophilized products, which led to a shorter purification process. HPLC chromatograms of crude peptides normally consisted of one major peak and a few byproducts. The addition of 0.1% TFA to both separation buffers lowered the *pH* specifically below the *pI* of histidine (7.6) and therefore led to a partial protonation of the oligohistidines. This gave the peptides a very hydrophilic character and aggravated their retention on C18 columns. On different scales of hydrophobicity, histidine is placed in the middle among the 20 naturally occurring amino acids [21]. As a consequence, it has also a small lipophilic character allowing a proper retention on a C18 column. A direct correlation between peptide length and retention time was observed for the HPLC purification (*Table* 6-3).

Table 6-3: Data from the HPLC purification (retention time) and the mass spectrometric analysis of the oligohistidine peptide series (calculated and measured monoisotopic masses).

Peptide	Retention Time [min] [a]	Calculated Mass [Da]	Experimental Mass [Da]
His3	7.9 [b]	428.2	428.2
His4	4.5	566.3	565.2
His5	4.7	703.3	703.2
His6	6.0	840.4	840.3
His7	9.0	977.4	977.4
His8	12.5	1114.4	1114.4
His9	13.4	1250.5	1250.7
His10	14.2	1388.6	1388.5

[a] His4-10 were separated using a water/acetonitrile/TFA gradient. [b] For His3, water/TFA was substituted with ammonium acetate buffer *pH* 8.8.

While His4-10 could be easily separated with the water/acetonitrile/TFA system, the trihistidine peptide eluted already with the injection peak under these conditions. By using ammonium acetate at *pH* 8.8 instead of 0.1% TFA, the average charge of the trihistidine could be decreased resulting in a stronger lipophilic character. The increase of buffer *pH* prolonged the retention time on the C18 column from 2.5 min (injection peak) up to 7.9 min using the same gradient. ESI/MS analysis was successful using cone voltages from 50 to 80 V, which led to single and double charged hydrogen adducts and small amounts of single charged sodium adducts in positive mode (*Table* 6-3). Only single charged species and small fragments, arisen from the applied ionization energy, were visible in the negative mode.

6.3.2 Oligohistidine binding assay

The method of the binding assay was based on the study of Nieba *et al.* [5] and the recommendations of Biacore [22]. However, the eluent buffer was prepared without any addition of polysorbate 20, since a slow but steady decrease of binding activity of the peptides was observed, when a polysorbate-containing sample buffer was used. The reason for this behavior is not yet understood and could include complexation effects. Furthermore, two additional washing steps (500 mM imidazole and 0.1% SDS) were introduced besides the regeneration with 350 mM EDTA in order to fully avoid any carry-over effects. After the complete regeneration, the surface could be reproducibly loaded with nickel before each sample injection. While the amount of captured nickel ions was highly constant within one experimental series, the nickel capacity could vary between individuals flow cells and was found to decrease slightly over several weeks of chip usage. Using the modified assay procedure, all peptide samples could be measured in triplicates with high signal intensity and reproducibility (*Fig*. 6-3).

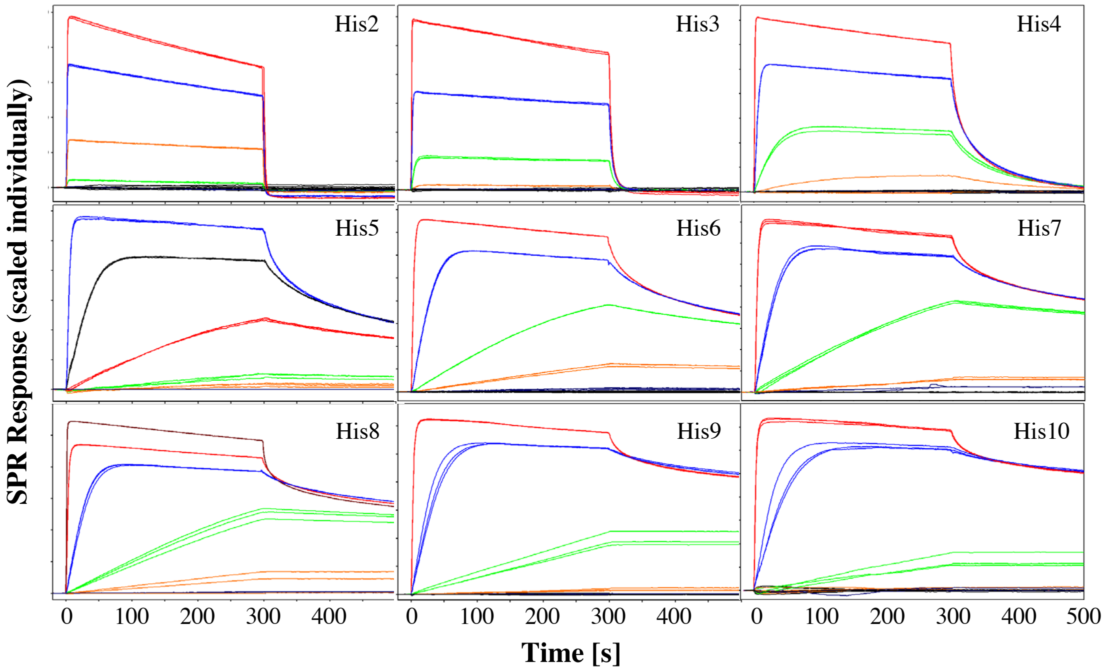


Figure 6-3: Sensorgrams of the oligohistidine peptides His2-His10. All samples were injected as randomized triplicates. The concentration range of individual samples is listed in *table* 6-1.

The reproducibility of the triplicate injections was very good, except for the peptides containing seven or more histidine residues. In these sensorgrams, concentrations around the saturation level of the chip surface showed a significantly higher signal deviation within the triplicate injections. No apparent trend could be detected within an injection series, and neither additional wash steps nor a change in the injection order showed any improvement. Therefore, sample carry-over and loss of binding activity over assay time are rather unlikely. Furthermore, impurities from the synthesis or degradation of peptides could be excluded by HPLC analyses of the peptide solutions before and after the assay. Finally, no mass transfer effect could be detected when running the experiment at different flow rates between 10 and 100 µl/min (see *appendix* E1). Since the effect only occurred with the larger peptides, time-dependent changes in the conformation could be regarded as a possible explanation. Therefore, additional experiments had to be performed to investigate this behavior in more detail.

When analyzing the whole peptide series, an additional effect became apparent: all oligohistidines showed a significant shift during ‘steady state’. In addition, the post-injection baseline signal dropped under the initial level, which is clearly visible in the sensorgrams of His2 and His3 (*Fig*. 6-3). This phenomenon was also observed for His_2Ala_4 hexapeptides (see *section* 6.3.3). The drift during injection as well as the strong rebinding effects made the mathematical determination of binding kinetics in terms of k_{on} and k_{off} impossible. On the other hand, the binding phases could be evaluated qualitatively and showed significant differences within the peptide series. When the sensorgrams were normalized by dividing the SPR response with the peptide mass, a clear trend to slower dissociation with increased peptide length could be observed (*Fig*. 6-4A). While the baseline of His2 and His3 rapidly returned to the baseline, the dissociation became steadily slower from His4 to His10 (*Fig*. 6-3), most likely caused by avidity and rebinding effects. Clear indicators of rebinding are the facts that dissociation doesn’t follow a normal exponential decay, dissociation rates seem to vary with the concentration of analyte, and the baseline is not reached during dissociation phase. These effects have already been demonstrated by Nieba *et al.* [5] for the synthesized hexahistidine peptide. For peptides with more than eight histidine residues, a very stable capturing could be achieved at concentrations below saturation (*Fig*. 6-4A, His10 peptide).

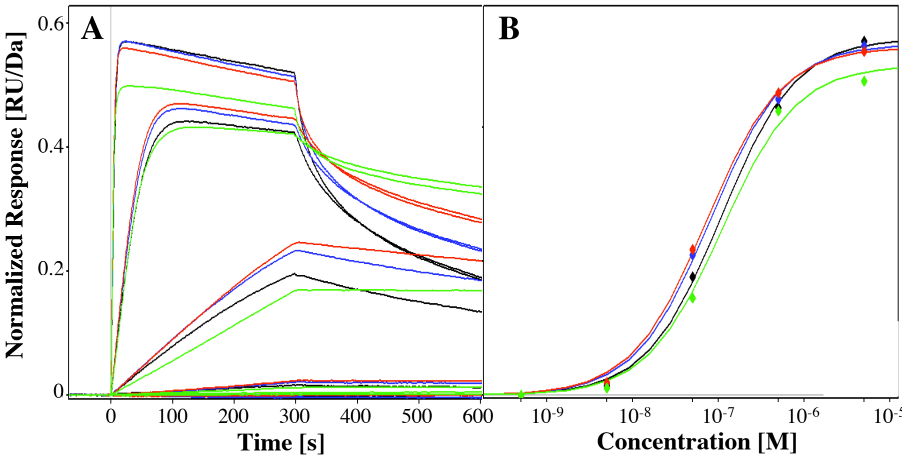


Figure 6-4: Overlay plots of oligohistidine responses normalized by the molecular weight of the corresponding peptide. Both the sensorgrams (**A**) and the steady state affinity plot (**B**) were recorded over a concentration range of 0.5-5000 nM (tenfold linear dilution series) for His5 (black), His6 (blue), His7 (red), and His10 (green).

Even though the sensorgrams cannot be described by standard binding models, the affinity could be estimated by fitting the SPR response at a narrow time range before injection end to a single binding site model (Fig. 6-4B). The resulting K_D values are listed in *table* 6-4 and are further visualized in *figure* 6-5.

Table 6-4: Estimated dissociation constants of the screened oligohistidine peptides.

Peptide	K_D [μM]
His2	123.0±7.0
His3	2.89±0.20
His4	0.836±0.081
His5	0.067±0.004
His6	0.034±0.002
His7	0.039±0.002
His8	0.044±0.004
His9	0.120±0.005
His10	0.165±0.018

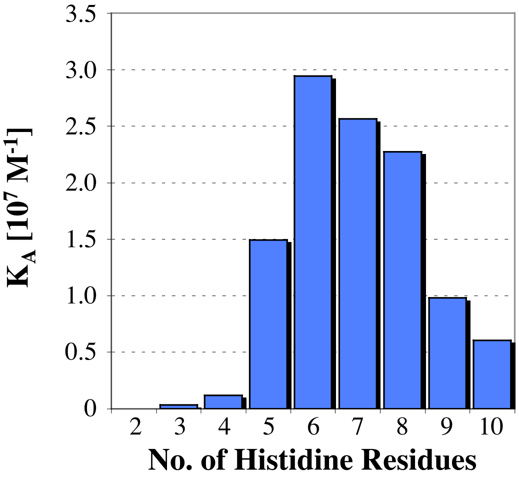


Figure 6-5: Association constants of the oligohistidine peptides calculated from the values in *table* 6-4 by *equation* 2 (see *section* 2.1.1).

The shortest peptide (His2) of the series also showed the most unfavorable binding properties with a K_D in the high micromolar range. By addition of a histidine residue (i.e. His3), the interaction could be improved by a factor of 40. This trend was continued until a length of six consecutive histidine residues. With a K_D of 34 nM, the hexahistidine showed the highest affinity among the peptide series. Additional histidine residues did not improve the affinity any further, but led to an increase in K_D until it reached a value of 165 nM for the decahistidine. The initial improvement in K_D (His2-His6) can be explained by thermodynamic means (*Eq.* 10). An increasing number of interacting groups directly improve the enthalpy term (ΔH) due to a higher probability of simultaneous electrostatic interactions (e.g. by rebinding). The higher flexibility of longer peptides might further increase this probability up to a certain point. This finally leads to a slower dissociation rate with increasing length. However, the free binding enthalpy (ΔG) is also dependent on entropy (ΔS), and when a peptide length of six residues is transgressed, the entropy seems to become the dominating term of the equation. With increasing length, peptides have more possibilities to adopt different conformations. The loss of entropy by forcing the molecule into a binding conformation increases with each additional residue and the free enthalpy is increasing, which leads to weaker binding. Therefore, a peptide length of six histidine residues seem to be an optimal compromise between the enthalpic and entropic components.

$$\Delta G = \Delta H - T\Delta S \qquad \text{[Eq. 10]}$$

The estimated K_D values for the free His6 peptide in this study are remarkably stronger than those for hexahistidine-tagged proteins reported by Nieba *et al.* [5] (~1 μM). Even though these authors also investigated the interaction of the free peptide with the nickel surface, no K_D value had been reported. Limited accessibility of the tag, steric hindrance, or electrostatic interactions with the tag are possible explanations for the discrepancy between the free peptide and the tagged proteins. The strength of this technology lies in the stability, meaning a strong rebinding during dissociation. Whereas the short peptides, especially di- and trihistidine, show a very rapid dissociation without rebinding, the rebinding effect appears with tetrahistidine and becomes more and more obvious with increasing peptide length.

6.3.3 Synthesis and purification of His_2Ala_4 peptides

The Ni^{2+}-NTA surface allows the simultaneous binding of two histidine residues (*Fig.* 6-1, see *section* 6.1.1). In order to determine the optimal distance between the two interacting imidazole groups, a series of His_2Ala_4 hexapeptides was prepared, where the position of one histidine was consecutively shifted while the second was fixed at the *C*-terminus. While the synthesis of the peptide series was comparable to those of the oligohistidine peptides (see *section* 6.3.1), the purification was much more challenging.

Similar to His3, the peptide peaks overlaid with the injection peak of the HPLC separation when standard conditions were used (water/acetonitrile/TFA). When these peaks were collected and used for Biacore screening, the SPR signal always showed a strong non-specific component (see *appendix* E2). Therefore, ammonium acetate was used instead of water/TFA due to its higher *pH* range, low UV cut-off, and volatility. With this method, it was possible to fully separate the peptides from the reagents used for the synthesis.

6.3.4 His$_2$Ala$_4$peptide assay

Based on the findings of the oligohistidine system (see section 6.3.2), a similar assay was developed for the screening of the His$_2$Ala$_4$ peptides. As expected and similar to the His2 peptide (see *Fig*. 6-3), the kinetic rate constants of the whole peptide series was very fast (*Fig*. 6-6).

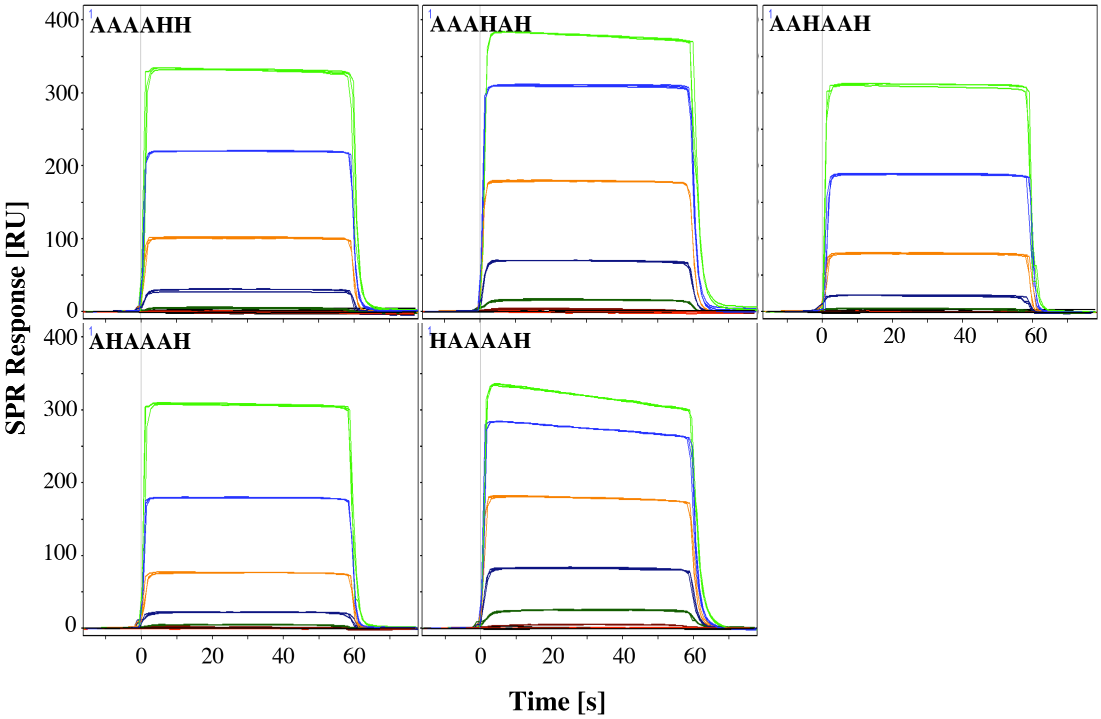


Figure 6-6: Sensorgrams of the different His$_2$Ala$_4$ peptides over a concentration range of 0.32-5000 μM (fivefold linear dilution; randomized triplicate injections).

While the overall shape of the sensorgrams approximately remained constant for all peptides, there were differences visible in terms of signal intensity and equilibrium drift (*Fig*. 6-6). Due to the fast kinetics and the complete return to the baseline, less regeneration/wash steps were required compared to the oligohistidine assay (only a single EDTA injection). Since the amino acid composition and the peptide length was the same for the whole series, the molecular weight of the peptides remained constant. This facilitated the direct comparison of the peptides in terms of affinity and kinetics. The reproducibility of the randomized triplicate injections was excellent for all peptides (*Fig.* 6-7A). The steady state response could acceptably be fitted to a single binding site model (*Fig.* 6-7B). Finally, all data sets also fitted kinetically to a 1:1 Langmuir binding model with acceptable accuracy (*Fig.* 6-7C). The minor deviations are most likely caused by small rebinding effects, the involvement of both or only a single histidine residue in the binding, and the negative drift during injection.

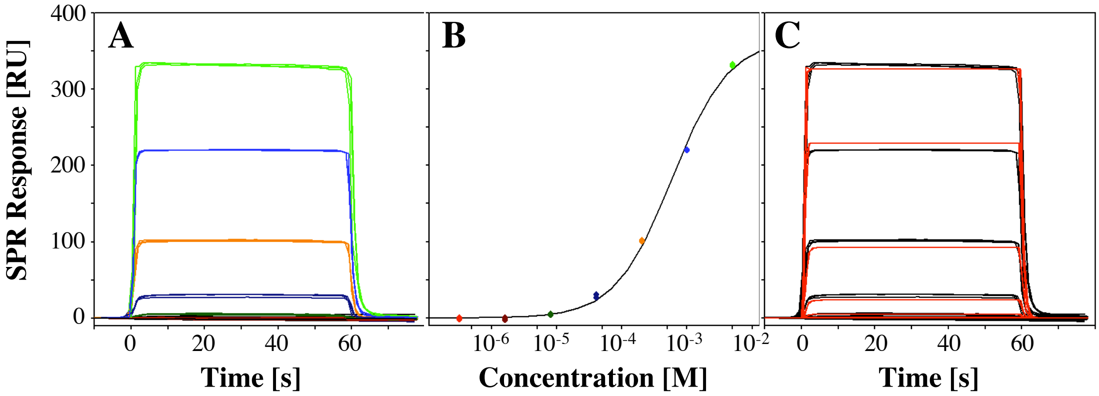


Figure 6-7: Binding assay for His_2Ala_4 peptides on the example of AAAAHH. **A:** Overlay plot of triplicate sensorgrams between 0.32 and 5000 μM. **B:** Equilibrium binding plot fitted to a single binding site model. **C:** Kinetic evaluation of the data set by fitting it to a 1:1 Langmuir binding model.

The whole peptide series was therefore evaluated both in terms of affinity and kinetics (*Table* 6-5), and the values were compared with each other (*Fig.* 6-8A). As demonstrated for AAAAHH (*Fig.* 6-5), all curves could be described with a 1:1 Langmuir model, while the equilibrium responses were fitted to a single binding site model (see also *appendix* E3).

Table 6-5: Evaluation of kinetics (1:1 Langmuir model) and affinity (single binding site model) of the His_2Ala_4 peptide series.

Sequence	k_{on} [$M^{-1}s^{-1}$]	k_{off} [s^{-1}]	$t_{1/2}$ [a] [s]	$K_{D\ kin}$ [b] [μM]	$K_{D\ ss}$ [c] [μM]
AAAAHH	1650	0.96	0.7	582	600
AAAHAH	3067	0.69	1.0	225	230
AAHAAH	1400	1.20	0.6	857	820
AHAAAH	1170	1.05	0.7	897	900
HAAAAH	4265	0.61	1.1	143	140

[a] Calculated by $t_{1/2} = \ln 2/k_{off}$. [b] Calculated by $K_D = k_{off}/k_{on}$. [c] Determined from the steady state binding plot by fitting the data to a single binding site model.

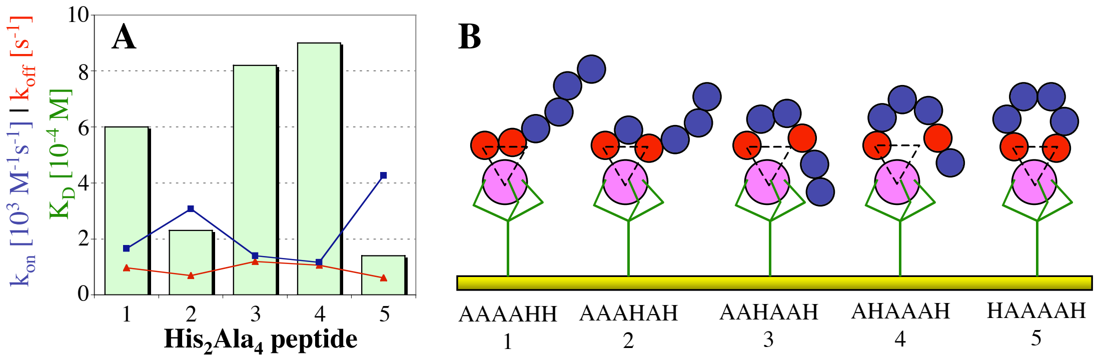


Figure 6-8: Binding properties and geometries of the His_2Ala_4 peptide series. **A:** Kinetic properties as an overlay of k_{on} (blue lines), k_{off} (red lines) and K_D (green columns) for the individual peptides (1-5). **B:** Hypothetical binding geometries of the five investigated peptides. Histidine residues are represented in red and alanine in blue, nickel as a purple sphere and NTA as green lines. The black, dashed triangle symbolizes the optimal binding geometry.

Comparison of the kinetic properties revealed some interesting similarities and differences. While there were only minor deviations in the kinetic off-rate between the five peptides, the calculated association phase showed significantly higher deviations. The significant deviation between the K_D values of the HH peptide (123 μM; *Table* 6-4) and the AAAAHH peptide (582 μM; *Table* 6-5) could be explained by steric or entropic effects. While peptides AAAAHH, AAHAAH, and AHAAAH were rather similar with a slight trend to weaker affinities, AAAHAH and HAAAAH bound much stronger to the nickel surface (*Fig.* 6-8A). This effect was influenced by both on- and off-rates, but with a much stronger contribution from the association rate. Interestingly, these two peptides also showed the most prominent drift during steady state (*Fig.* 6-6; see next paragraph). These findings indicate that a rather small gap between the two binding residues is preferred. Both a vicinal position as well as a larger spacer of 2-3 alanines led to a significant loss of affinity. HAAAAH, with four linking alanine residues is believed to possess enough flexibility for establishing an optimal binding geometry again (*Fig.* 6-8B). This hypothesis also leads to suggestions for further experiments. First, the gap between the two histidine residues could be fine-tuned by substituting the intermediate alanine by other moieties, e.g. with different sizes or limited conformational flexibility. Second, the effect of larger spacers (i.e. more than 4 alanines) should be investigated. Finally, the histidine positions of peptide 2 and 5 could be combined resulting in a peptide with the sequence HAAHAH.

Similar to the experiments with oligohistidine peptides (see section 6.3.2), a significant drift during the steady state phase was observed (*Fig.* 6-6). This effect was most obvious for AAAHAH and HAAAAH, which also showed the strongest binding affinity (*Table* 6-5). In contrast to most of the oligohistidines, the signal of the His_2Ala_4 peptides always returns to the baseline after injection, even at high sample concentrations. A possible post-injection drop in the baseline signal due to removed nickel could therefore be detected more easily. In order to quantify this effect, the drop in the nickel baseline was determined from the sensorgram (Δ baseline; *Fig*. 6-9C) and used for the calculation of the expected signal decrease (ΔR_{calc}; *Table* 6-6). This value could then be compared with the steady state drift in the sensorgram (ΔR_{exp}; *Fig*. 6-9B).

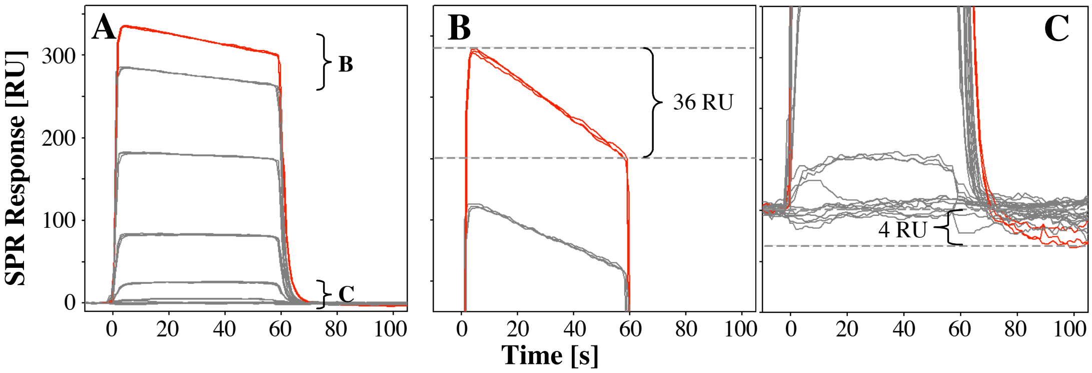


Figure 6-9: Removal of Ni^{2+} from the surface by complexation with the HAAAAH peptide. **A:** Overlay plot of HAAAAH with the investigated concentration (5 mM) highlighted in red. **B:** Close-up of the steady state area showing the typical drift resulting in a lower signal at injection end (ΔR_{exp}). **C:** Close-up of the baseline area with the characteristic drop of the baseline signal at injection end (Δ baseline).

Table 6-6: Evaluation of the complexation-induced signal drops in the baseline and the binding signal at a constant peptide concentration (5 mM).

Peptide	Δ baseline [a] [RU]	ΔR_{calc} [b] [RU]	ΔR_{exp} [a] [RU]	Slope [c] [RUs^{-1}]
AAAAHH	< ±1	n.d.	-6	-0.10
AAAHAH	< ±1	n.d.	-12	-0.20
AAHAAH	< ±1	n.d.	-4	-0.07
AHAAAH	< ±1	n.d.	-6	-0.10
HAAAAH	-4 RU	-38	-36	-0.60

[a] Determined from the sensorgram (*Fig*. 6-9). [b] Calculated using *equation* 4 (see *section* 2.1.3) with MW_{target}= 60 Da (nickel), $MW_{analyte}$= 576.6 Da (HAAAAH), density = Δ baseline, valency = 1. [c] Calculated as ΔR_{exp}/injection time (60s).

Due to the rather low signals for the baseline drop (≤ 4 RU), the expected steady state shift could only be calculated for HAAAAH. Longer injection times or higher peptide concentrations might improve the signal intensity. For HAAAAH, the loss of Ni^{2+} from the surface corresponded very well with the signal drop during the injection phase (*Table* 6-6). A removal of nickel ions leads to a decreased binding capacity of the surface. Therefore, a lower number of interactions with peptide molecules occurs per time unit resulting in an equilibrium drift. While the baseline drop was too small to detect for most of the peptides, the steady state shift could be determined for all samples. Similar to the affinity measurements (*Table* 6-5), AAAHAH and HAAAAH showed the strongest effect. The complexation effect might also be explained in terms of solution-based stability constants [23]. While the affinity for the Ni^{2+}-NTA in solution ($-\log K_D = 11.26$) is approximately 350 times stronger than the Ni-His interaction ($-\log K_D = 8.69$), additional histidine residues could decrease this difference and lead to a strong competition between NTA and the peptide for the nickel ion. Therefore, a complexation of nickel seems to be a reasonable explication of the steady state drift.

6.4 Conclusions

The hexahistidine tag has found its way to the most important tagging technology for the purification of recombinant proteins through recent years. Despite their numerous applications, little is known about the exact mechanism of the metal binding. In addition, the technology is only of limited use for some important applications such as the immobilization of tagged proteins to Ni^{2+}-NTA sensor chips, due to insufficient stability of the capturing. The hexahistidine tag had been selected as an ideal structure by empirical screening and has been used without any changes ever since. Therefore, the rational alteration of the tag structure based on binding geometries and kinetic data might improve the properties and broaden the application of metal-affinity tags.

In a first step, the ideal peptide length of a consecutive oligohistidine peptide series between two and ten His residues was tested on a Ni^{2+}-NTA surface using Biacore. With an increasing number of histidine moieties, the kinetic properties changed significantly to slower dissociation rates. Due to the complex binding mode, which is believed to include avidity, rebinding, and removal of nickel from the surface, a quantitative evaluation of this series was not possible. However, the trend to increased surface stability for larger peptides was clearly visible and could be demonstrated when overlaying the individual dissociation phases. Despite the steady increase in dissociation stability, the estimated overall affinity only improved until a peptide length of six histidine residues and became weaker when the series was continued to ten histidines. This indicates a strong contribution from the association phase, which might be influenced by the increased conformational flexibility and a loss of entropy for longer peptides. Therefore, the hexahistidine unit, identified by Hochuli et al. [8] from the screening of tagged model proteins, seems to represent the best compromise between flexibility and stability, indeed.

Even though the best binding properties were obtained with six consecutive histidine residues, only two moieties can bind simultaneously to a captured nickel ion. Knowing the ideal distance and binding geometry might therefore help identifying alternative spacers for metal affinity tags. For this purpose, a series of His_2Ala_4 hexapeptides with varying distance of the two His was tested using the same assay. While all peptides featured fast kinetic properties, the steady state affinity showed significant variations. The K_D values did not follow a linear trend to stronger or weaker binding with increasing distance of the imidazole rings, but seem to clearly prefer a spacer length of either one or four alanine residues between the histidines. Similar to the oligohistidine peptides, the K_D seemed to be strongly influenced by the association rate constant, underlining the importance of a predefined binding geometry. Based on these results, novel combinations of histidine with other natural or synthetic amino acids could be developed using Biacore and molecular modeling in order to identify improved binding motifs.

A common phenomenon in both assays was a strong negative drift in the equilibrium responses during peptide injection. In case of the His_2Ala_4 peptides, this drift became stronger with increasing affinity of the peptide. Since the nickel baseline was found to be reproducibly decreased after peptide injection, a removal of nickel ions from the surface through complexation by His-containing peptides in solution seems to be a reasonable explanation. Indeed, the signal drop during injection correlated very well with the decreased binding capacity of the nickel surface.

As a consequence, the strong interaction of oligohistidine tags with surface-bound nickel seems to be a combination of rebinding effects, an optimal binding geometry, and a limited flexibility (i.e. peptide length). Most of these points have a great potential for improvement and could lead to the development of short tags suitable for both peptide/protein purification and surface immobilization.

6.5 Literature

[1] J. Porath, J. Carlsson, I. Olsson, G. Belfrage, *Nature* **1975**, *258*, 598.

[2] E. K. Ueda, P. W. Gout, L. Morganti, *J Chromatogr A* **2003**, *988*, 1.

[3] G. S. Chaga, *J Biochem Biophys Methods* **2001**, *49*, 313.

[4] V. Gaberc-Porekar, V. Menart, *J Biochem Biophys Methods* **2001**, *49*, 335.

[5] L. Nieba, S. E. Nieba-Axmann, A. Persson, M. Hamalainen, F. Edebratt, A. Hansson, J. Lidholm, K. Magnusson, A. F. Karlsson, A. Pluckthun, *Anal Biochem* **1997**, *252*, 217.

[6] J. Madoz-Gurpide, J. M. Abad, J. Fernandez-Recio, M. Velez, L. Vazquez, C. Gomez-Moreno, V. M. Fernandez, *J Am Chem Soc* **2000**, *122*, 9808.

[7] E. Hochuli, H. Dobeli, A. Schacher, *J Chromatogr* **1987**, *411*, 177.

[8] E. Hochuli, H. Bannwarth, R. Döbeli, R. Gentz, D. Stüber, *Bio/Technology* **1988**, *9*, 1321.

[9] K. Terpe, *Appl Microbiol Biotechnol* **2003**, *60*, 523.

[10] *Protein Data Bank: The worldwide repository for macromolecular structures*, Research Collaboratory for Structural Bioinformatics (RCSB), http://www.pdb.org/, **2005**.

[11] P. D. Gershon, S. Khilko, *J Immunol Methods* **1995**, *183*, 65.

[12] M. Conti, G. Falini, B. Samori, *Angew Chem Int Ed Engl* **2000**, *39*, 215.

[13] L. Schmitt, M. Ludwig, H. E. Gaub, R. Tampe, *Biophys J* **2000**, *78*, 3275.

[14] H. L. Liu, Y. Ho, C. M. Hsu, *J Biomol Struct Dyn* **2003**, *21*, 31.

[15] E. A. Woestenenk, M. Hammarstrom, S. van den Berg, T. Hard, H. Berglund, *J Struct Funct Genomics* **2004**, *5*, 217.

[16] A. Chant, C. M. Kraemer-Pecore, R. Watkin, G. G. Kneale, *Protein Expr Purif* **2005**, *39*, 152.

[17] J. Wu, M. Filutowicz, *Acta Biochim Pol* **1999**, *46*, 591.

[18] R. L. Rich, D. G. Myszka, *J Mol Recognit* **2005**, *18*, 1.

[19] K. M. Muller, K. M. Arndt, K. Bauer, A. Pluckthun, *Anal Biochem* **1998**, *259*, 54.

[20] M. R. Wilkins, I. Lindskog, E. Gasteiger, A. Bairoch, J. C. Sanchez, D. F. Hochstrasser, R. D. Appel, *Electrophoresis* **1997**, *18*, 403.

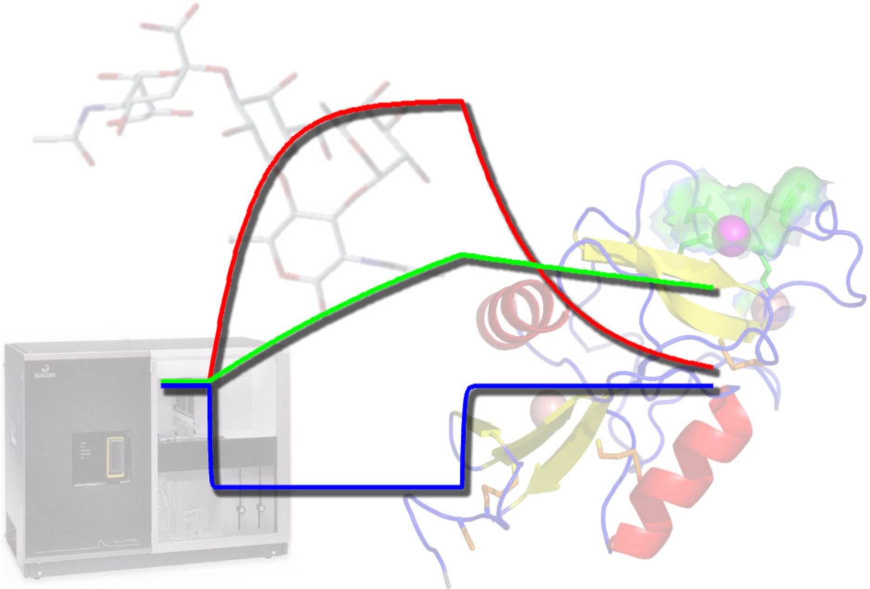
[21] J. L. Cornette, K. B. Cease, H. Margalit, J. L. Spouge, J. A. Berzofsky, C. DeLisi, *J Mol Biol* **1987**, *195*, 659.

[22] *Capture of histidine-tagged proteins using NTA or anti-histidine antibodies (Application Note 12)*, Biacore AB, **2002**.

[23] T. E. Furia, in *CRC handbook for food additives*, 2nd ed. (Ed.: T. E. Furia), The Chemical Rubber Co., Cleveland, **1972**.

Chapter 7

General Considerations



The previous chapters showed different applications of the surface plasmon resonance (SPR) technology for characterizing molecular interactions with emphasis on small molecules and carbohydrate-protein interactions. Both areas are rather challenging compared to the classical Biacore applications such as antibody-antigen interactions, and provide some pitfalls and limitations. The aim of this chapter is therefore to summarize general considerations and recommendations both from the literature and our own experience. After focusing on general assay design and applications to small molecules, the importance and special features of carbohydrate-protein interactions will be discussed. Negative SPR binding signals were one of the most unusual and interesting finding of this thesis, and are therefore reconsidered from a general point of view.

7.1 Working with small molecules

The major component of the SPR signal generation is an increase of the local electron density around the gold surface upon ligand binding. Since larger molecules induce a more intense shift of the resonance angle, Biacore instruments are often referred to as 'mass detectors'. As a consequence, small molecules show only SPR signals of low intensity and are more difficult to detect. While larger responses often tolerate small heterogeneities and artifacts, these inaccuracies can lead to severe deviations of the binding parameters in case of low molecular weight analytes. Alongside with the improvement of instrument sensitivity, a number of protocols and recommendations for such assays have been developed. A good overview about small molecule assays on Biacore and other SPR instruments is given in reviews by Rich and Myszka [1-5]. Careful planning of a Biacore assay is a prerequisite for a successful analysis, and includes the selection of targets and analytes, assay design, data processing, evaluation of the results, and instrument maintenance (*Fig*. 7-1).

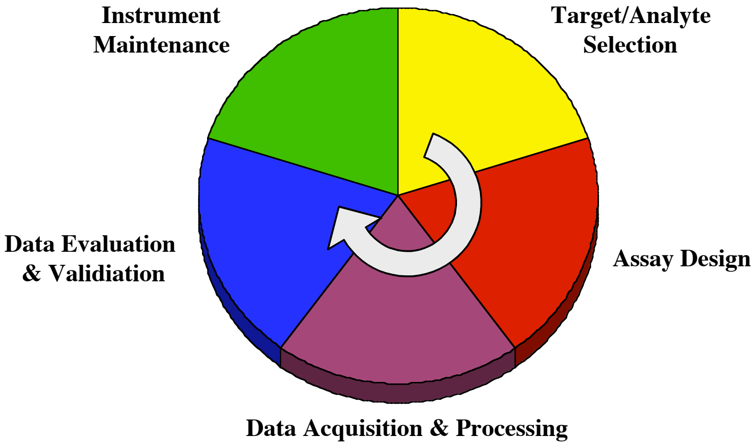


Figure 7-1: Key steps in the planning of a Biacore assay. Instrument maintenance is an important part of the whole cycle and should be considered throughout the entire process.

7.1.1 Target and analyte selection

When an interaction between two different molecules is measured, the molecular properties and the quality of the binding partners are vital for the outcome of the assay. As much information as possible about target and analytes should be collected to avoid inactivation effects, binding interferences and artifacts. Some of the most important parameters to consider are listed in *table* 7-1.

Table 7-1: Molecular properties of analytes and targets that are of importance for assay development.

Protein (Target)	Small Molecule (Analyte)
• Size / Molecular weight	• Molecular weight
• *pI* value	• pK_a values
• Purity (SDS-PAGE, SEC)	• Purity (HPLC)
• Stability (pH, SDS, ...)	• Stability
• Free amines, thiols, ...	• Spacers, ...
• Tags, capturing domains	• log P, solubility

Size / Molecular weight

The size of the binding partners is directly responsible for the intensity of the SPR signal (see *equation* 4, *section* 2.1.3). Therefore, the target size should be kept as low as possible, and protein domains that are not involved in the interaction should be removed. However, any removal of protein domains has to be validated carefully, since even domains distant from the binding site may contribute to the protein activity (stabilization etc.). Smaller protein fragments might also show a better binding activity and reduced non-specific binding, as it could be demonstrated for the monoclonal antibody GSLA-2 (see *sections* 4.3.3 and 4.3.4). In case of the analyte in solution, the molecule should exceed at least 200 Da, but higher molecular weights are clearly preferred. However, for most drug discovery applications the lead or candidate compounds are kept below 500 Da to fulfill the requirements of the Lipinski 'rule of 5' [6].

pI / pK_a

While the *pI* value of a protein target is important for its immobilization (see *section* 2.1.3), the pK_a of a small analyte molecule can also influence the outcome of a Biacore experiment. As demonstrated for the interaction of salicylic acid with human serum albumin (see *section* 3.3.7), acidic compounds might exceed the capacity of the running buffer and lead to a shift in *pH*. In case of *pH*-sensitive proteins like HSA, this might generate an artificial SPR signal that overlays with the binding response. Therefore, free carboxylic group should be avoided if possible, and the *pH* of the injected sample with the highest concentration should be carefully controlled. An increase of the running buffer capacity might be required, if ligand-induced pH shifts are detected.

Target and analyte purity

The purity of both the target and the analytes is one of the most important parameters in a SPR experiment. Kapoor *et al.* suggest a purity of at least 90% [7], but these values are highly dependent on the detection method and the type of assay. Since any increase in mass or electron density around the gold surface is equally detected, a differentiation between specific and non-specific signals is nearly impossible. In some cases, non-specific binding can be detected by a clearly biphasic sensorgram as it had been observed for some sLe^a batches in case of GSLA-2 (see *section* 4.3.4). Chromatographic purification and sensitive analytics should therefore be performed for each analyte molecule (e.g. LC-MS). Similarly, impurities in the protein fraction might lead to heterogeneous binding signals and non-specific binding. Again, highly selective purification steps (e.g. affinity chromatography) and analysis by non-reducing SDS-PAGE are an important way for avoiding such complications. Oligomerization of the target should also be considered as an impurity, since the increased local concentration is likely to influence the kinetic and thermodynamic properties of the interaction [8]. While these deviations may be small for monovalent analytes, polyvalent ligands usually bind significantly stronger to oligomeric forms of the target. This effect was also visible in the asialoglycoprotein receptor project, where monomeric and dimeric fractions of the receptor were analyzed separately (see *section* 5.3).

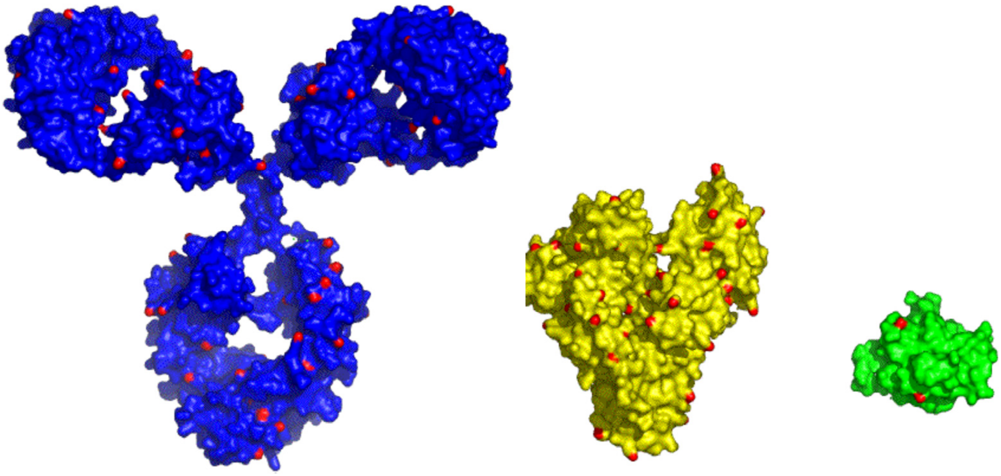


Figure 7-2: Comparison of the size and lysine distribution of different targets. An IgG1 antibody (blue; PDB IgG1, see *chapter* 4), human serum albumin (yellow; PDB 1BM0, see *chapter* 3), and ASGP-R H1-CRD (green; PDB 1DV8, see *chapter* 5) are visualized at the same scale. Amine groups of lysine residues are highlighted in red.

Stability

In many cases, important parameters about the stability and other features of a protein target are already available from its expression and purification. For example, elution conditions during affinity chromatography are a good starting point for the development of appropriate regeneration conditions.

Functional groups and spacers

If a crystal structure of the target is available, the immobilization success and the expected surface heterogeneity can be estimated by visualizing the surface-accessible functional groups such as lysine or cysteine residues (*Fig*. 7-2). If a small molecule has to be immobilized, the introduction of a spacer group might be necessary. In this case, crystal structure data might also be beneficial for deciding the length and position of such a group. Spacer groups and tags might also facilitate the purification of the analyte molecules. Since these additional groups might interfere with the interaction, as it has been shown for the Lemieux spacer on sLe^a (see *section* 4.3.2), they should be avoided if possible.

7.1.2 Assay design

In order to guarantee a maximum of reliability and sensitivity, a small molecule assay has to be planned carefully. The points to consider include the appropriate choice of sensor chips, immobilization chemistry and density, reference surfaces, buffers and analyte concentrations, injection and regeneration conditions, wash steps and control experiments. The various decisions to take are highly dependent on the individual experiment and are beyond the scope of this thesis. Therefore, only some important points are listed below:

- Despite the expected surface heterogeneity, amine coupling is usually the *immobilization method* of choice for small molecule assays. Thiol coupling might be a valuable alternative, but requires a free cysteine residue (see *sections* 3.3.8 and 5.3.4). Capturing approaches usually lead to oriented surfaces but are often too unstable for high-performance screening purposes (see *section* 4.3.5). Recent developments like the SNAP-tag [9, 10] might help combining surface stability with target orientation.

- The choice of an appropriate reference surface might be crucial in some cases [11, 12], e.g. when non-specific binding is involved. Different approaches have been suggested depending on the experiment and the available proteins (*Table* 7-2). While Biacore originally recommended using activated/deactivated surfaces in order to mimic the changed surface charges, this approach had later been found to generate higher deviations than untreated reference surfaces [13]. Mutated or blocked target surfaced are usually preferred over similar proteins or denatured targets, since latter may show remaining or even altered binding activity.

Table 7-2: Comparison of different approaches for the generation of reference surfaces.

Approach	Specificity	Stability	Availability
Untreated chip surface	poor	high	always
Activated/deactivated surface	poor	high	always
Similar, inactive protein	moderate	high	rarely
Deactivated target (denaturation)	variable	variable	high
Deactivated target (mutation)	high	high	variable
Blocked target (inhibitor)	high	variable	variable

- A well-known or already validated analyte with a rather high molecular weight should be chosen for *assay development*. This assures sufficient signal intensities and facilitates the detection of irregularities.
- If DMSO is required to dissolve the analytes (as in case of human serum albumin, see *section* 3.3), its concentration should be kept as low as possible (1-5% preferred, max. 8%) and a DMSO calibration is mandatory [14]. In addition, the effect of different DMSO concentration on the binding affinity should be controlled carefully [15].
- Injection of sample *triplicates* and *randomization* of the samples increases the reliability and is an important tool for the recognition of carry-over effects.

Additional information to assay design can be found in the Biacore sensor surface handbook [16] as well as in some reviews [7, 17-20].

7.1.3 Data acquisition and processing

The following procedures were found to be helpful or even essential for the processing of low-intensity SPR signals, as they are often observed when dealing with small molecules:

- Extensive *priming*, *normalizing*, *degassing*, frequent *exchange of buffers*, and short *centrifugation of all samples* are important steps for avoiding signal disturbances by air bubbles and particles.
- *Tight capping* of the sample vials (rubber caps) and *temperature control* of the autosampler compartment (~20°C). This usually allows the reproducible measurement of sample triplicates even from the same vial (see *chapters* 3-5).
- *Buffer spiking* [21] facilitates the alignment of small signals before referencing and leads to more accurate difference sensorgrams (see *section* 4.2.3).
- *Data acquisition rates* should be set to the maximum to assure most accurate fits and avoid loss of information at injection start and end due to the de-spiking procedure. This is especially important when kinetic investigations are performed.
- *Referencing* to an appropriate surface (see also *section* 7.1.2) can be vital, especially when non-specific binding is observed. During method development, two or three reference surfaces on the same chip (e.g. untreated vs. activated/deactivated surface, or inactive target surface) should be created and evaluated.

- *Double referencing* [18] eliminates most of the systemic noise during data processing and was used in all projects. Five to ten 'warm-up' blanks at the beginning of each experiment and one or two blank between each series are normally sufficient.
- *De-spiking* (included in *Scrubber*) automatically removes signal spikes caused by small differences in the detection time as a result of the serially connected flow cells or by air bubbles.

Reproducible data processing is essential for a high-quality analysis of low-molecular weight data sets. Unfortunately, the current software tool provided by Biacore (*BIAevaluation*) requires most of the processing steps to be done manually, which could lead to small deviations. In addition, important steps like double referencing and DMSO correction are not implemented. The kinetic models included in *BIAevaluation* cover most of the application and can be extended easily. However, the limitation to 24 simultaneous binding curves is not suitable for high quality experiments (triplicate injections etc.). *Scrubber*, a software tool released in 2003, overcomes most of the limitations of *BIAevaluation* by offering standardized and highly automated data processing including double referencing and DMSO correction. Since Scrubber does not include a kinetic module, *CLAMP* [22] is a valuable alternative for the evaluation of kinetic data. Future versions of *Scrubber* are planned to include the functional capabilities of *CLAMP* [23]. Therefore, *Scrubber* (and *CLAMP*) should be preferred for the analysis of small molecule assays on Biacore 3000.

7.1.4 Data evaluation and assay validation

Results obtained by Biacore experiments were shown to be very reproducible and to correspond with solution-based experiments in many cases [21, 24-26]. However, SPR signals represent changes in the electron density around the gold surface, which are usually caused by analyte binding but could also include conformational changes, non-specific binding, or bulk effects. Therefore, a critical evaluation of the data is evident and includes the selection of appropriate binding models, correlation with literature data and validation with other analytical methods. In many cases, more complex binding models lead to better fit results, but this effect can be simply a cause of higher number of mathematical parameters to define a curve. As a consequence, complex binding models should always be questioned critically and correlated with known mechanistic properties of a binding event. For example, a bivalent binding model is very unlikely for the interaction between two monovalent molecules. In any case, kinetic analysis should be performed by globally fitting the entire data set [27, 28]. In addition, mass transport to and from the surface [29-31] should always be considered and tested by variation of the flow rates. Finally, the Biacore assay should always be validated by different methods or experiments. This could include a comparison with known binding data from literature, as it was done in the case of albumin (see *chapter* 3) or with data from ELISA (e.g. ASGP-R, see *chapter* 5) or other biophysical methods (e.g. GSLA-2, *chapter* 4). Alternatively, the confirmation of a targets' known binding properties such as the calcium- or pH-dependency of ASGP-R (see *chapter* 5) can also be included as part of the validation process.

7.1.5 Instrument maintenance and validation

Unspecific adsorption of proteins and analytes as well as microbial growth can dramatically influence the performance and reproducibility of SPR experiments. As a consequence, a comparative study with several Biacore instruments investigating the same enzyme/inhibitor pair was able to attribute a lack of reproducibility to bad instrument maintenance [21]. Biacore recommends to clean the system at least once a week and to perform a disinfection routine every month. However, working with small molecules might require higher cleaning frequencies and the use of additional wash solutions [18, 19]. Additional *desorb* routines between experiments are therefore suggested if the target tolerates a removal of the sensor chip. In addition, the running buffer should be switched to pure water after the experiment, in order to avoid salt build-up in the instrument [18].

Even when an instrument is purified and maintained regularly, small deviations in the detection unit cannot be excluded. Especially when dealing with small molecules, a maximum in terms of performance and sensitivity is required. Unfortunately, neither the Biacore *system check* procedure nor the myoglobin/anti-myoglobin mAb system from the *getting-started-kit* are appropriate for testing the instruments' sensitivity. Therefore, an established low molecular weight system should be selected for this purpose. Both the targets and a set of analytes should be commercially available in good quality, and the interaction should be well characterized. Two of the investigated systems in this thesis essentially fulfill these requirements, i.e. bovine carbonic anhydrase II (see *section* 2.3) and human serum albumin (see *chapter* 3). Even though the interaction of small drugs with albumin is one of the most intensely described interaction systems, the complexity of its binding events, the rather high sensitivity towards environmental factors and the requirement of DMSO for most analytes reduces its suitability for validation purposes. The carbonic anhydrase system, on the other hand, was found to be very robust and its reproducibility has been validated using Biacore instruments from various laboratories [21, 26]. In addition, this drug-enzyme interaction not only provides affinity constants but also kinetic parameters, which can be compared with literature.

7.2 Carbohydrate-protein interactions

Among the biomolecular interactions, those between carbohydrates and their protein targets belong to the most unusual but also most interesting ones. They are involved in many critical processes like inflammatory responses, cellular trafficking and signaling, fertilization, tumor metastasis, or infection diseases. Proteins that specifically interact with carbohydrates are referred to as *lectins*, which include many subclasses like selectins, annexins, or siglecs. Some of them are soluble in the cytosol or in body fluids, while many others are embedded in membranes. Carbohydrates are usually attached to other structures like proteins or lipids, and the glycosylation of proteins is one of the most important posttranslational modifications. Compared to proteins and nucleic acids, the information content of carbohydrates is much larger due to their complex stereochemistry and their branched structures [32]. Their involvement in many pathogenic conditions makes lectins an interesting target in drug discovery. However, the generation of inhibitors for lectin-interactions was found to be very challenging. Carbohydrate binding sites are often rather shallow and solvent-exposed, and large areas of the lectin might be involved in the binding. The forces involved in lectin interactions with monovalent carbohydrates are generally weak and often feature equilibrium dissociation constants in the micro- to millimolar range. The low affinity is a consequence of very rapid binding kinetics with dissociation half-times of less than a second. Even interactions with specific antibodies, usually characterized by very strong affinities, might remain in the micromolar range when a carbohydrate epitope is involved [12]. This effect was also observed in the case of sLe^a binding to GSLA-2 (see *section* 4.3.2), where the interaction resulted in a K_D of ~ 10 μM. In contrast to the affinity, the binding specificity is very high. For example, the carbohydrate recognition domain (H1-CRD) of human ASGP-R only recognizes galactose and its derivatives while it does not bind to glucose (see *section* 5.3.5). This selectivity is achieved by multiple polar (hydroxyl groups) and non-polar interactions (carbon face of the sugar ring) [33-35]. In contrast to the strong monovalent interactions of most peptides and drugs, carbohydrates increase their affinity mainly by multivalent binding to surface-bound targets. Glycoproteins usually contain multiple glycan chains and are able to address more than one receptor. This receptor-clustering is an important mechanism in cell-cell recognition and signaling pathways [8, 36, 37]. Avidity and rebinding effects predominantly influence the dissociation rate and often lead to binding affinities in the nanomolar range [8]. This effect has also been observed for the ASGP receptor, where monovalent galactose bound with a K_D of 10^{-3} M while galactose-bearing glycoproteins were able to binding around 10^{-9} M (see *section* 5.3.5; *Fig*. 7-3B&C). Carbohydrate-protein interactions are far more complex than many other biological binding event, ranging from mono- to multivalent, from fast to slow, and from weak to strong. As a consequence, the knowledge about kinetic and thermodynamic properties of carbohydrate-lectin interactions becomes of a paramount importance [7, 8, 38].

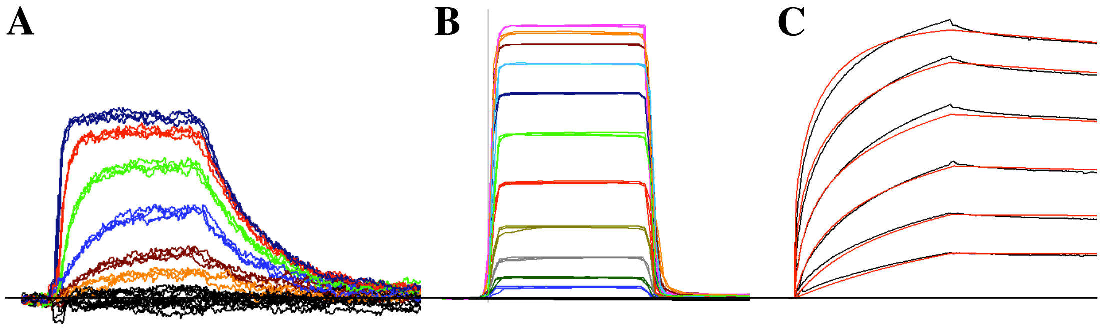


Figure 7-3: Comparison of typical sensorgrams for enzyme-drug (**A**; carbonic anhydrase II vs. CBS, see *section* 2.3), lectin-monosaccharide (**B**; ASGP-R H1-CRD vs. GalNAc, see *chapter* 5.3.5), and lectin-glycoprotein interactions (**C**; ASGP-R H1-CRD vs. asialofetuin, see *section* 5.3.5).

The use of SPR technology for the characterization of carbohydrate-protein interactions provides many advantages. Sugars usually lack any chromophore or fluorophore, making the detection with traditional analytical methods rather challenging. Since SPR is independent on specific functional groups, Biacore technology is a valuable alternative for the detection of carbohydrates [7]. In addition, data about affinity and kinetics of an interaction can be addressed in a single experiment, therefore facilitating the assessment of structure-activity relationships. Compared to peptides and synthetic molecules, the synthesis of carbohydrates is often more complicated and time-consuming, and automation processes are still under development [32]. As many of the biologically relevant lectins and carbohydrates are embedded in membranes or attached to larger structures, surface-based SPR assays are believed to be much closer to reality than solution-based experiments [36]. This advantage might be used to investigate clustering effects and multivalent binding, e.g. by reverse the molecules on the surface and in solution, or by varying the immobilization density [37, 39-41]. Variation of the immobilization density cannot only influence the binding stoichiometry [37], but also lead to a change in the selectivity of the lectin [41]. For example, the same interaction pair can lead to completely different sensorgrams, dependent on which partner is immobilized on the sensor chip [39]. Finally, the flow system might induce an additional critical component, since some of the interactions between sugars and lectins occur in the blood circulation and were found to be flow-dependent (e.g. selectin-induced tethering and rolling during inflammation processes) [42, 43]. However, evaluation of flow-dependency might interfere with mass transport effects (see *section* 7.1.4). The fast kinetics for monovalent sugars has clear benefits for Biacore experiments, since binding equilibria are reached within seconds (*Fig*. 7-3B), therefore reducing the injection and assay time. In addition, the rapid return to the baseline eliminates the need for any regeneration conditions, preventing the target from any damage and increasing the life-time of the protein surface. Even if multivalent, tight binding event are analyzed, the carbohydrates can usually be removed by specific conditions (e.g. removal of calcium in the case of C-type lectins; see *chapter* 5) or by competition with monovalent sugars [7]. Kinetic analysis of the binding data is often more challenging in case of carbohydrate-protein interaction. While the rapid kinetic rate constants of monovalent sugars are usually close to the detection limit of the Biacore instrument, the properties of multivalent ligands are influenced by rebinding effects and heterogeneities. Immobilization density is an important factor in this context and an optimal compromise between signal intensity and rebinding/avidity has to be elaborated [8]. Finally, some lectins might interact with the hydrogel that covers Biacore sensor chips, since is consists of a glucose polymer (dextran). For example, the sensor surface had to be changed to a glycolipid layer in a competition assay with soluble concanavalin A against immobilized sugar, since the tetrameric lectin was found to interact with the dextran matrix [40]. A similar effect might also be involved in the generation of negative binding signals, as they have been observed in the case of the ASGP-R H1-CRD (see *sections* 5.3.6 and 7.3). Furthermore, negative binding has also been detected in our laboratory with other lectins such as the myelin-associated glycoprotein (MAG) or E-selectin though not fully investigated [44, 45]. Such matrix interactions might therefore be a general problem for the characterization of lectin interactions, both when immobilized and in solution.

7.3 Negative binding responses and signal overlay

The most unusual phenomenon observed during the Biacore experiments in this thesis was the generation of negative SPR responses in the albumin (see *section* 3.3.5) and the ASGP-R project (see *section* 5.3). In both cases, these negative responses were clearly concentration dependent and highly reproducible. While the effect was restricted to a special group of analytes in the case of human serum albumin (zwitterionic amino acids, i.e. L-tryptophan and L-kynurenine), the H1-CRD of ASGP-R showed negative responses for the whole set of binding monosaccharides. In contrast, heavier molecules like asialoglycoproteins or anti-H1-CRD antibodies always generated highly positive signals. Since the asialoglycoproteins and the monovalent sugars are likely to bind to the same site of the lectin (see *section* 5.1.4), the effect cannot be simply attributed to independent binding mechanisms but seems rather to be influenced by the molecular weight of the analyte. As a consequence, the resulting sensorgram is believed to be a product of two overlaying SPR signals of opposite sign.

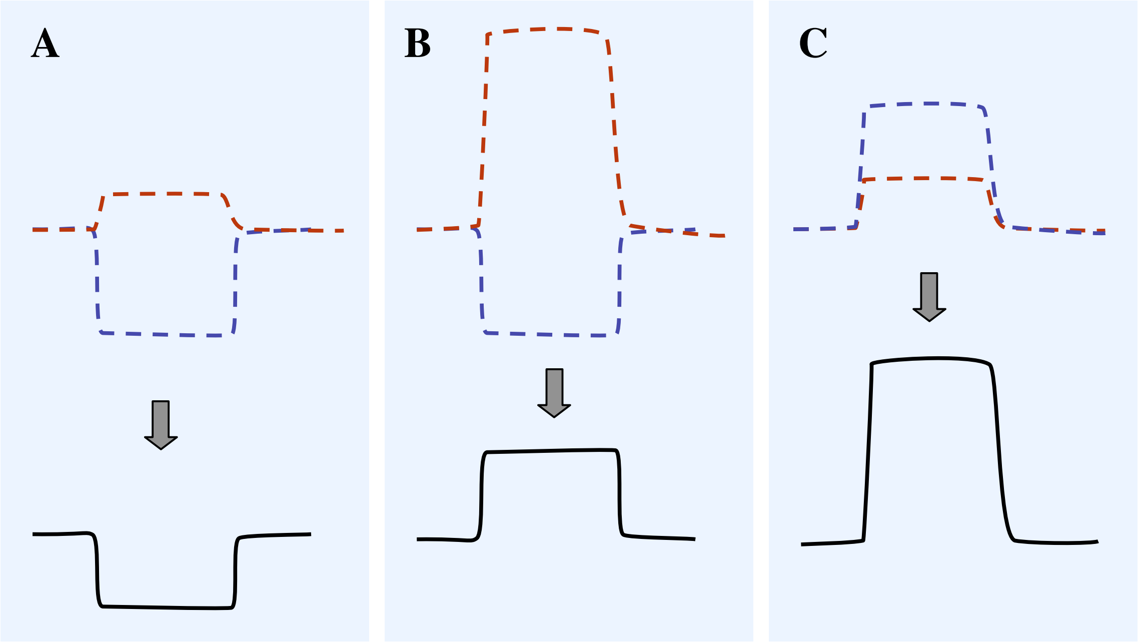


Figure 7-4: Generation of an apparent sensorgram (black solid line) by an overlay of two signal components (red and blue dashed lines). If the two signals have opposite signs, the ratio of their intensities decides whether the resulting sensorgram is negative (**A**) or positive (**B**). The addition of two positive signals might amplify the intensity of the apparent sensorgram (**C**).

The binding of an analyte molecule to the immobilized target increases the electron density around the surface in any case and therefore mandatory induces a positive signal effect, which is directly dependent on the molecular weight of the analyte. If the second negative signal component is independent on the analyte size, the resulting sensorgram could be either positive (large analyte) or negative (small analyte), as it is shown in *figure* 7-4A&B. An experimental series with H1-CRD analytes of constant affinity but varying molecular weight might help confirming this hypothesis. As a consequence of this signal generation model, an overlay of two positive signal component should amplify the intensity of the resulting sensorgram (*Fig*. 7-4C). Indeed, a corresponding effect was observed in the case of sialic acid binding to HSA (see *section* 3.3.7), where a pH-induced conformational change was suspected to overlay with the binding signal.

The question remains, what might cause the additional signal component? In general, Biacore is often referred to as a 'mass detector', since any binding of an analyte leads to an increased mass and electron concentration, which then induces the SPR signal. However, any other change in the electron density like fluctuations in the matrix length or the protein layer could also influence the response. Different models might be considered (*Fig*. 7-5), of which the most reasonable are discussed in more detail.

- Since interactions between two molecules do not occur statically but include many dynamic steps (induced fit of the protein, conformational adaptation of the analyte), the protein layer might change its shape, thickness, and density upon ligand binding. Such *conformational changes* (*Fig*. 7-5B) might well change the electron density and the SPR response.

- Even if only minor changes in the protein structure are induced, they still might influence the exposition of charged residues on the protein surface. Since the carboxymethyl dextran surface is negatively charged under physiological conditions, this could lead to an *electrostatic attraction or repulsion* of the protein (*Fig*. 7-5C) resulting in a changed distance to the gold surface.

- When a buffer component is able to interact with the target, the addition of an analyte molecule might lead to a competitive removal of this component (solution competition, *Fig*. 7-4D). If the buffer substance has a higher molecular weight than the analyte, this removal induces a negative signal.

- Interaction of a lectin with the glucose-based dextran matrix could also influence the packing density of the surface. Upon analyte binding, the protein (or matrix chains) are competitively released resulting in a decreased density and a negative signal (*surface competition*, *Fig*. 7-5E).

The latter two models could be the reason for the effects observed in the atypical binding properties of ASGP-R H1-CRD (see *section* 5.3). When calcium-containing running buffer is used for the screening, the lectin domain binds to the dextran matrix and pulls the protein closer to the surface. As soon as an analyte molecule reaches the lectin, a competition with the binding to the surface takes place and the protein flips back, leading to the negative SPR signal (see *section* 5.3.5). When soluble dextran is added to the running buffer, the model might change from a surface to a solution competition situation. This is supported by the fact that increasing dextran concentrations amplify the negative response while retaining the overall affinity (see *section* 5.3.6). The good correlation with reported affinities and the dependency from calcium and buffer pH indicate a specific reaction. Even though the interaction of H1-CRD with monovalent glucose was found to be very weak (see *section* 5.3.5), the high local concentration might lead to a considerable binding affinity for the dextran matrix. As a consequence of this possible interaction, it is hard to determine if the experiment is closer to a direct (K_D) or competitive (IC_{50}) determination of the affinity. Exchange or derivatization of the surface matrix could give a deeper insight in the underlying principles and might confirm the stated hypothesis. In addition, other lectins should be analyzed for the same effect.

In case of human serum albumin, conformational changes or an influence of surface charges are more likely to induce the effects observed for L-tryptophan and L-kynurenine (see *section* 3.3.5). Since no other class of analytes showed negative responses, the effect has to be correlated to the analytes themselves or their binding site. However, since both the signal intensities and the molecular weight of the two substances are rather low, the negative binding effect could be masked by an overlay with a more intense positive binding signal in case of other analytes. For salicylic acid, the overlay of two positive signals induced by the binding event and a pH-induced conformational change is highly likely. An increase of the running buffer capacity or the use of sodium salicylate successfully eliminated the pH-dependent signal component (see *section* 3.3.7).

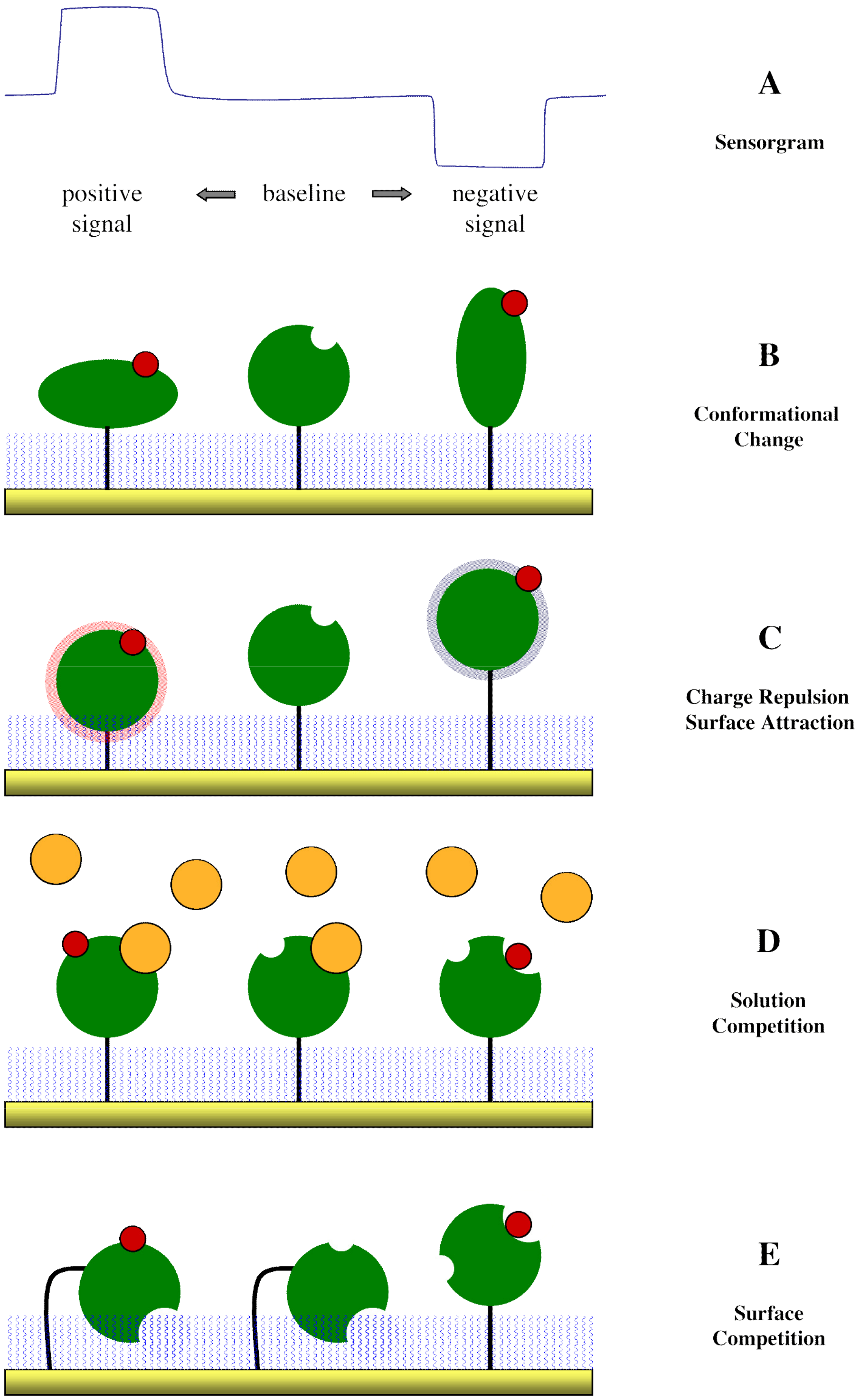

positive signal
baseline
negative signal
A
Sensorgram
B
Conformational Change
C
Charge Repulsion Surface Attraction
D
Solution Competition
E
Surface Competition

Figure 7-5: Comparison of different hypothetical model for the generation of positive and negative sensorgrams (**A**) besides the simple mass addition. **B:** Binding of an analyte (red) to the immobilized protein (green) induces a conformational change. **C:** Analyte binding leads to the exposition of positive (pale red circle) or negative surface charges (pale blue circle). **D:** A large buffer component (orange) binds to the protein and is competitively removed upon ligand binding. **E:** The protein interacts with the dextran matrix (blue) of the sensor chip (yellow) and is released when the analyte competes for the binding site.

Unfortunately, little is reported in literature about signal abnormalities or changes of the protein/matrix layer during Biacore experiments [46-51]. The first description of negative responses was published by Gestwicki *et al.* [49], who observed negative post-injection responses when injecting different saccharides to maltose-binding protein (MBP). The detected post-injection signals were concentration-dependent but rather small, and nothing was stated about negative equilibrium responses. The author attributed this effect to ligand-induced conformational changes, and emphasized this theory with a second example, where small calcium ions induced a large SPR signal when interacting with transglutaminase [49]. Since MBP is also a lectin, a competitive mechanism as postulated for H1-CRD might also be possible. Conformational changes were also stated by Mannen *et al.* [50] for the signal generation of a series of immobilized proteins at varying pH. However, Paynter *et al.* [51] questioned the conclusions of this study and suggested an electrostatic interaction with the surface as a more plausible mechanism, after they did some additional experiments with different proteins and peptides. The unspecific detection of any changes in the electron density around the gold surface by SPR makes a clear distinction between different mechanisms very difficult and only carefully planned control experiments might give rise to a specific model. Furthermore, the simultaneous occurrence of more than one effect (e.g. conformational and electrostatic changes) is very likely. As a consequence, more experiments have to be performed in order to bring more light in the complex effect that might be involved in the SPR signal generation.

7.4 References

[1] R. L. Rich, D. G. Myszka, *J Mol Recognit* **2000**, *13*, 388.

[2] R. L. Rich, D. G. Myszka, *J Mol Recognit* **2001**, *14*, 273.

[3] R. L. Rich, D. G. Myszka, *J Mol Recognit* **2002**, *15*, 352.

[4] R. L. Rich, D. G. Myszka, *J Mol Recognit* **2003**, *16*, 351.

[5] R. L. Rich, D. G. Myszka, *J Mol Recognit* **2005**, *18*, 1.

[6] C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney, *Adv Drug Deliv Rev* **2001**, *46*, 3.

[7] M. Kapoor, C. J. Thomas, K. Bachhawat-Sikder, S. Sharma, A. Surolia, *Methods Enzymol* **2003**, *362*, 312.

[8] E. Duverger, N. Frison, A. C. Roche, M. Monsigny, *Biochimie* **2003**, *85*, 167.

[9] M. Kindermann, N. George, N. Johnsson, K. Johnsson, *J Am Chem Soc* **2003**, *125*, 7810.

[10] W. Huber, S. Perspicace, J. Kohler, F. Muller, D. Schlatter, *Anal Biochem* **2004**, *333*, 280.

[11] R. J. Ober, E. S. Ward, *Anal Biochem* **1999**, *271*, 70.

[12] M. Strandh, B. Persson, H. Roos, S. Ohlson, *J Mol Recognit* **1998**, *11*, 188.

[13] R. Karlsson, *Working with small molecules in Biacore*. in *Biacore Symposium 2002*, Chicago, USA, **2002**.

[14] A. Frostell-Karlsson, A. Remaeus, H. Roos, K. Andersson, P. Borg, M. Hamalainen, R. Karlsson, *J Med Chem* **2000**, *43*, 1986.

[15] R. L. Rich, Y. S. Day, T. A. Morton, D. G. Myszka, *Anal Biochem* **2001**, *296*, 197.

[16] *Biacore Sensor Surface Handbook*, Biacore AB, Uppsala, **2003**.

[17] D. G. Myszka, *Curr Opin Biotechnol* **1997**, *8*, 50.

[18] D. G. Myszka, *J Mol Recognit* **1999**, *12*, 279.

[19] D. G. Myszka, *Strategies for kinetic analysis of small molecules using Biacore*, in *Biacore Workshop 2001*, Fulda, Germany, **2001**.

[20] R. Karlsson, M. Kullman-Magnusson, M. D. Hamalainen, A. Remaeus, K. Andersson, P. Borg, E. Gyzander, J. Deinum, *Anal Biochem* **2000**, *278*, 1.

[21] M. J. Cannon, G. A. Papalia, I. Navratilova, R. J. Fisher, L. R. Roberts, K. M. Worthy, A. G. Stephen, G. R. Marchesini, E. J. Collins, D. Casper, H. Qiu, D. Satpaev, S. F. Liparoto, D. A. Rice, Gorshkova, II, R. J. Darling, D. B. Bennett, M. Sekar, E. Hommema, A. M. Liang, E. S. Day, J. Inman, S. M. Karlicek, S. J. Ullrich, D. Hodges, T. Chu, E. Sullivan, J. Simpson, A. Rafique, B. Luginbuhl, S. N. Westin, M. Bynum, P. Cachia, Y. J. Li, D. Kao, A. Neurauter, M. Wong, M. Swanson, D. G. Myszka, *Anal Biochem* **2004**, *330*, 98.

[22] D. G. Myszka, T. A. Morton, *Trends Biochem Sci* **1998**, *23*, 149.

[23] T. A. Morton, BioLogic Pty Ltd, personal communication, **2005**.

[24] J. Deinum, L. Gustavsson, E. Gyzander, M. Kullman-Magnusson, A. Edstrom, R. Karlsson, *Anal Biochem* **2002**, *300*, 152.

[25] Y. S. Day, C. L. Baird, R. L. Rich, D. G. Myszka, *Protein Sci* **2002**, *11*, 1017.

[26] D. G. Myszka, Y. N. Abdiche, F. Arisaka, O. Byron, E. Eisenstein, P. Hensley, J. A. Thomson, C. R. Lombardo, F. Schwarz, W. Stafford, M. L. Doyle, *J Biomol Tech* **2003**, *14*, 247.

[27] T. A. Morton, D. G. Myszka, I. M. Chaiken, *Anal Biochem* **1995**, *227*, 176.

[28] L. D. Roden, D. G. Myszka, *Biochem Biophys Res Commun* **1996**, *225*, 1073.

[29] R. W. Glaser, *Anal Biochem* **1993**, *213*, 152.

[30] P. Schuck, *Biophys J* **1996**, *70*, 1230.

[31] D. G. Myszka, X. He, M. Dembo, T. A. Morton, B. Goldstein, *Biophys J* **1998**, *75*, 583.

[32] D. B. Werz, P. H. Seeberger, *Chemistry* **2005**.

[33] U. Spohr, O. Hindsgaul, R. U. Lemieux, *Can J Chem* **1985**, *63*, 2644.

[34] R. U. Lemieux, *Chem Soc Rev* **1989**, *18*, 347.

[35] W. I. Weis, K. Drickamer, *Annu Rev Biochem* **1996**, *65*, 441.

[36] H. Sota, R. T. Lee, Y. C. Lee, Y. Shinohara, *Methods Enzymol* **2003**, *362*, 330.

[37] C. W. Cairo, J. E. Gestwicki, M. Kanai, L. L. Kiessling, *J Am Chem Soc* **2002**, *124*, 1615.

[38] C. R. Bertozzi, L. L. Kiessling, *Science* **2001**, *291*, 2357.

[39] Y. Shinohara, Y. Hasegawa, H. Kaku, N. Shibuya, *Glycobiology* **1997**, *7*, 1201.

[40] D. A. Mann, M. Kanai, D. J. Maly, L. L. Kiessling, *J Am Chem Soc* **1998**, *120*, 10575.

[41] N. Horan, L. Yan, H. Isobe, G. M. Whitesides, D. Kahne, *Proc Natl Acad Sci U S A* **1999**, *96*, 11782.

[42] R. Alon, S. Chen, K. D. Puri, E. B. Finger, T. A. Springer, *J Cell Biol* **1997**, *138*, 1169.

[43] M. W. Nicholson, A. N. Barclay, M. S. Singer, S. D. Rosen, P. A. van der Merwe, *J Biol Chem* **1998**, *273*, 763.

[44] D. S. Strasser, S. Shelke, B. Ernst, *unpublished results.*

[45] D. Ricklin, Z. Dragic, B. Ernst, *unpublished results.*

[46] H. Sota, Y. Hasegawa, M. Iwakura, *Anal Chem* **1998**, *70*, 2019.

[47] Z. Salamon, S. Cowell, E. Varga, H. I. Yamamura, V. J. Hruby, G. Tollin, *Biophys J* **2000**, *79*, 2463.

[48] S. Boussaad, J. Pean, N. J. Tao, *Anal Chem* **2000**, *72*, 222.

[49] J. E. Gestwicki, H. V. Hsieh, J. B. Pitner, *Anal Chem* **2001**, *73*, 5732.

[50] T. Mannen, S. Yamaguchi, J. Honda, S. Sugimoto, A. Kitayama, T. Nagamune, *Anal Biochem* **2001**, *293*, 185.

[51] S. Paynter, D. A. Russell, *Anal Biochem* **2002**, *309*, 85.

Appendices

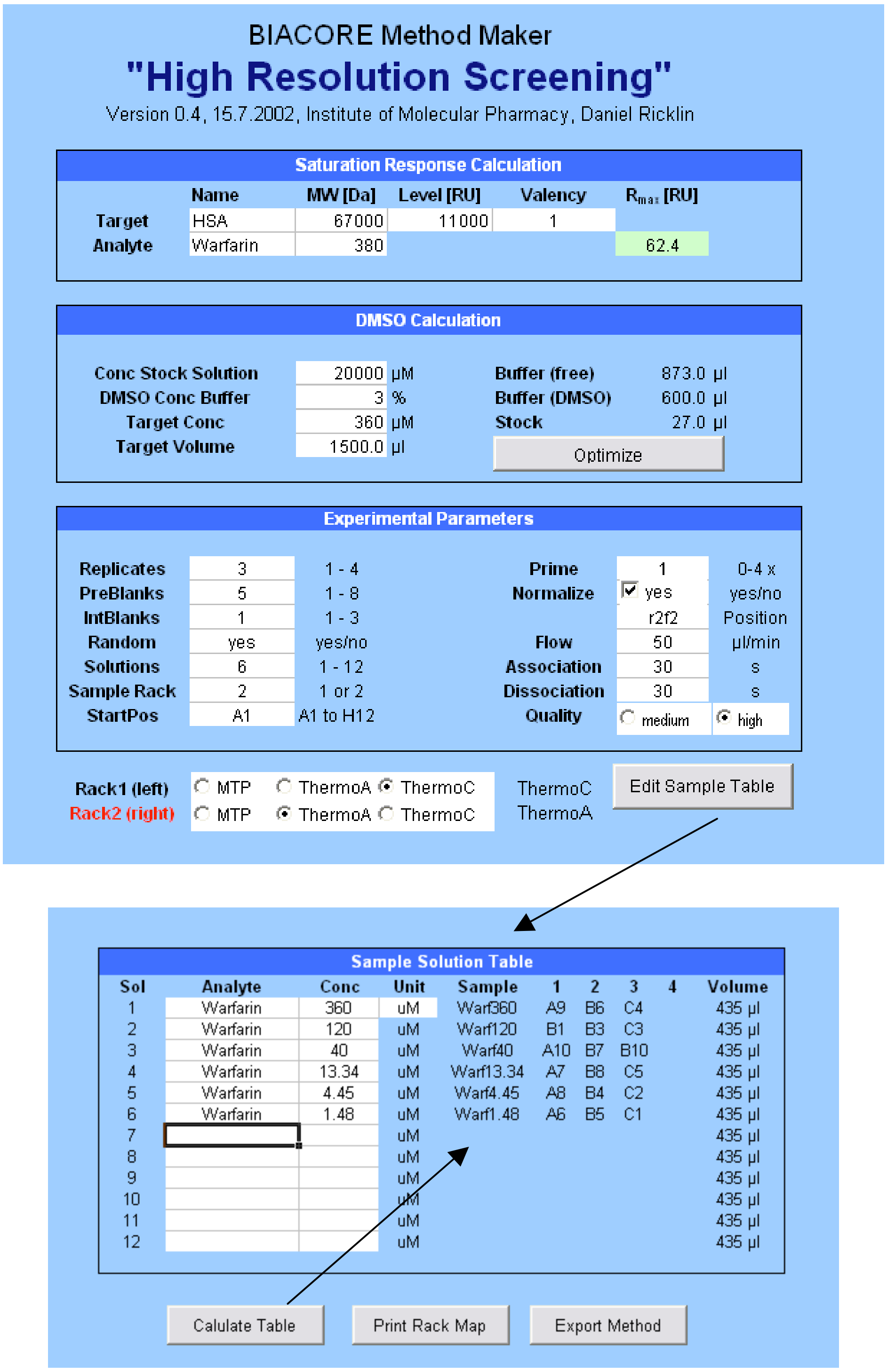
A: Carbonic Anhydrase II (CA II)
B: Human Serum Albumin (HSA)
C: Diagnostic Antibody GSLA-2
D: Asialoglycoprotein Receptor (ASGP-R)
E: Hexahistidine Tag (HisTag)

Appendix A1: Individual replicates of the CA II-CBS assay

	k_{on} [10^4 $M^{-1}s^{-1}$]	k_{off} [s^{-1}]	K_D (kin) [nM]	K_D (equil) [nM]
Run2 Fc1	3.62	0.039	1072	1110
Run2 Fc3	3.7	0.029	792	969
Run3 Fc1	3.58	0.035	969	897
Run3 Fc3	4.35	0.032	738	852
Run4 Fc1	4.31	0.036	838	949
Run4 Fc3	4.74	0.035	730	808
Average	4.05	0.034	856.5	930.8

Appendix B1: Biacore Method Maker "High Resolution Screening" (Excel)

The following screenshots illustrate the automated routine for the generation of randomized sample tables, DMSO concentration calculations, and Biacore 3000 methods (developed in Excel using Visual Basic).



BIACORE Method Maker

"High Resolution Screening"

Version 0.4, 15.7.2002, Institute of Molecular Pharmacy, Daniel Ricklin

Saturation Response Calculation

	Name	MW [Da]	Level [RU]	Valency	R_{max} [RU]
Target	HSA	67000	11000	1	
Analyte	Warfarin	380			62.4

DMSO Calculation

Conc Stock Solution	20000 µM	**Buffer (free)**	873.0 µl
DMSO Conc Buffer	3 %	**Buffer (DMSO)**	600.0 µl
Target Conc	360 µM	**Stock**	27.0 µl
Target Volume	1500.0 µl		

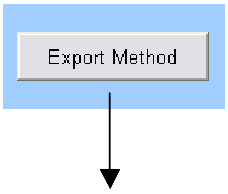
Experimental Parameters

Replicates	3	1 - 4	**Prime**	1	0-4 x
PreBlanks	5	1 - 8	**Normalize**	☑ yes	yes/no
IntBlanks	1	1 - 3		r2f2	Position
Random	yes	yes/no	**Flow**	50	µl/min
Solutions	6	1 - 12	**Association**	30	s
Sample Rack	2	1 or 2	**Dissociation**	30	s
StartPos	A1	A1 to H12	**Quality**	○ medium	● high

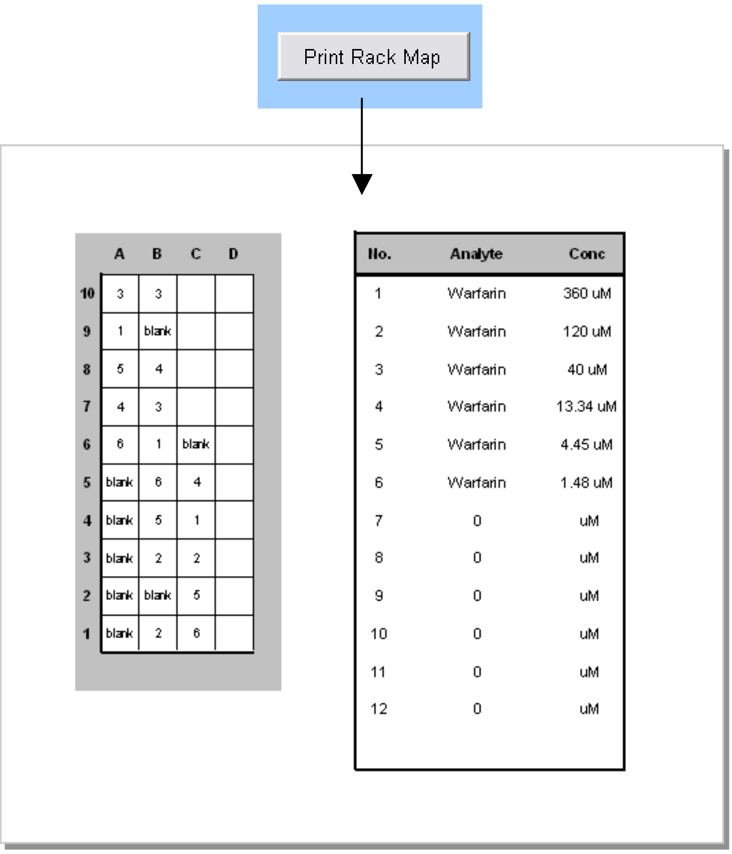
Rack1 (left)	○ MTP	○ ThermoA	● ThermoC	ThermoC
Rack2 (right)	○ MTP	● ThermoA	○ ThermoC	ThermoA

Sample Solution Table

Sol	Analyte	Conc	Unit	Sample	1	2	3	4	Volume
1	Warfarin	360	uM	Warf360	A9	B6	C4		435 µl
2	Warfarin	120	uM	Warf120	B1	B3	C3		435 µl
3	Warfarin	40	uM	Warf40	A10	B7	B10		435 µl
4	Warfarin	13.34	uM	Warf13.34	A7	B8	C5		435 µl
5	Warfarin	4.45	uM	Warf4.45	A8	B4	C2		435 µl
6	Warfarin	1.48	uM	Warf1.48	A6	B5	C1		435 µl
7			uM						435 µl
8			uM						435 µl
9			uM						435 µl
10			uM						435 µl
11			uM						435 µl
12			uM						435 µl



```
! This method was automatically created by HiResMaker Excel sheet
! © 2002 Daniel Ricklin, University of Basel
! Date = 09/23/2003

MAIN
PRIME
NORMALIZE r2f2
LOOP samples STEP
APROG sampleInject %sample %pos %conc
ENDLOOP
APPEND STANDBY
END

DEFINE LOOP samples
LPARAM %sample %pos %conc
TIMES 1
blank R2A1 0u
blank R2A2 0u
blank R2A3 0u
blank R2A4 0u
blank R2A5 0u
Warf1.48 R2A6 1.48u
Warf13.34 R2A7 13.34u
Warf4.45 R2A8 4.45u
Warf360 R2A9 360u
Warf40 R2A10 40u
Warf120 R2B1 120u
blank R2B2 0u
Warf120 R2B3 120u
Warf4.45 R2B4 4.45u
Warf1.48 R2B5 1.48u
Warf360 R2B6 360u
Warf40 R2B7 40u
Warf13.34 R2B8 13.34u
blank R2B9 0u
Warf40 R2B10 40u
Warf1.48 R2C1 1.48u
Warf4.45 R2C2 4.45u
Warf120 R2C3 120u
Warf360 R2C4 360u
Warf13.34 R2C5 13.34u
blank R2C6 0u
END

DEFINE APROG sampleInject
PARAM %sample %pos %conc
MODE -d0.1  !Data collection rate = high
KEYWORD conc %conc
FLOW 50
CAPTION High Resolution Screening (HSA/Warfarin), sample = %sample
* KINJECT  %pos 50 30
-20 RPOINT baseline -b
15 RPOINT plateau
QUICKINJECT r2f7 10
END
```



	A	B	C	D
10	3	3		
9	1	blank		
8	5	4		
7	4	3		
6	6	1	blank	
5	blank	6	4	
4	blank	5	1	
3	blank	2	2	
2	blank	blank	5	
1	blank	2	6	

No.	Analyte	Conc
1	Warfarin	360 uM
2	Warfarin	120 uM
3	Warfarin	40 uM
4	Warfarin	13.34 uM
5	Warfarin	4.45 uM
6	Warfarin	1.48 uM
7	0	uM
8	0	uM
9	0	uM
10	0	uM
11	0	uM
12	0	uM

Appendix B2: Literature data for HSA-drug interactions

Compound	Abbr	MW [Da]	log P	pKa	Plasma Protein Binding (%bound; with references)					
					1	2	3	4	5	6
Acetylsalicylic Acid	AcSA	180.2	1.19	3.49				49	37.3	10
Cholic Acid	ChoA	408.6	2.02	4.98						48
Corticosterone	Cort	346.5	1.94						67.8	
Diazepam	Dzpm	284.7	2.82	3.4	86.1			99	93.2	96.5
Digitoxin	Dgtx	765	1.85		96.6	97.9	94.7	97	76.4	92.5
Indoprofen	Indo	281.3	2.77							98.5
Naproxen	Napr	230.3	3.18	4.15	99.7	99.9	99.8	99.7	99	99.5
Nitrazepam	Nitra	281.3	2.25					87	82.3	87.5
Phenprocoumon	Ppro	280.3	3.62							99.5
Prednisone	Pred	358.4	1.46		53.7			75	37.6	
Quinidine	Qdne	324.4	3.44	8.56				87	61.7	
Quinine	Quin	324.4	3.44		34.6		22			
Salbutamol	Salb	239.3	0.64		8	8				10
Salicylic Acid	ScyA	138.1	2.26	2.97	81.7	79.9	93		73	85
Warfarin	Warf	308.3	2.6	5.08	98	99.4	99.6	99	97.9	99

Compound	Abbr	MW [Da]	Equilibrium Dissociation Constants (μM; with references)					
			1	2	3	7	8	9
Acetylsalicylic Acid	AcSA	180.2						7.7
Cholic Acid	ChoA	408.6						
Corticosterone	Cort	346.5			1400			
Diazepam	Dzpm	284.7				5	2	
Digitoxin	Dgtx	765	28	24.8	38		14–23	
Indoprofen	Indo	281.3						
Naproxen	Napr	230.3	26	10.6	1.5			
Nitrazepam	Nitra	281.3						
Phenprocoumon	Ppro	280.3						
Prednisone	Pred	358.4		288				
Quinidine	Qdne	324.4				122	714	
Quinine	Quin	324.4	566		2500		133	
Salbutamol	Salb	239.3		4300				
Salicylic Acid	ScyA	138.1		141	48	42	5–14	7.8
Warfarin	Warf	308.3	9	3.7	2.5	5–8	4–11	7.1
Method			Biacore	Biacore	Biacore	CE	Dialysis	ITC
Phosphate [mM]			67	65.5	10	67	-	33
NaCl [mM]			70	70	150	0	-	0
Temp [°C]			25	25	25	37	-	37
DMSO [%]			5	3	3	no [a]	-	n.a.
pH			7.4	7.4	7.4	7.4	-	7.4

[a] Acetonitrile and methanol were used for CE experiments instead of DMSO

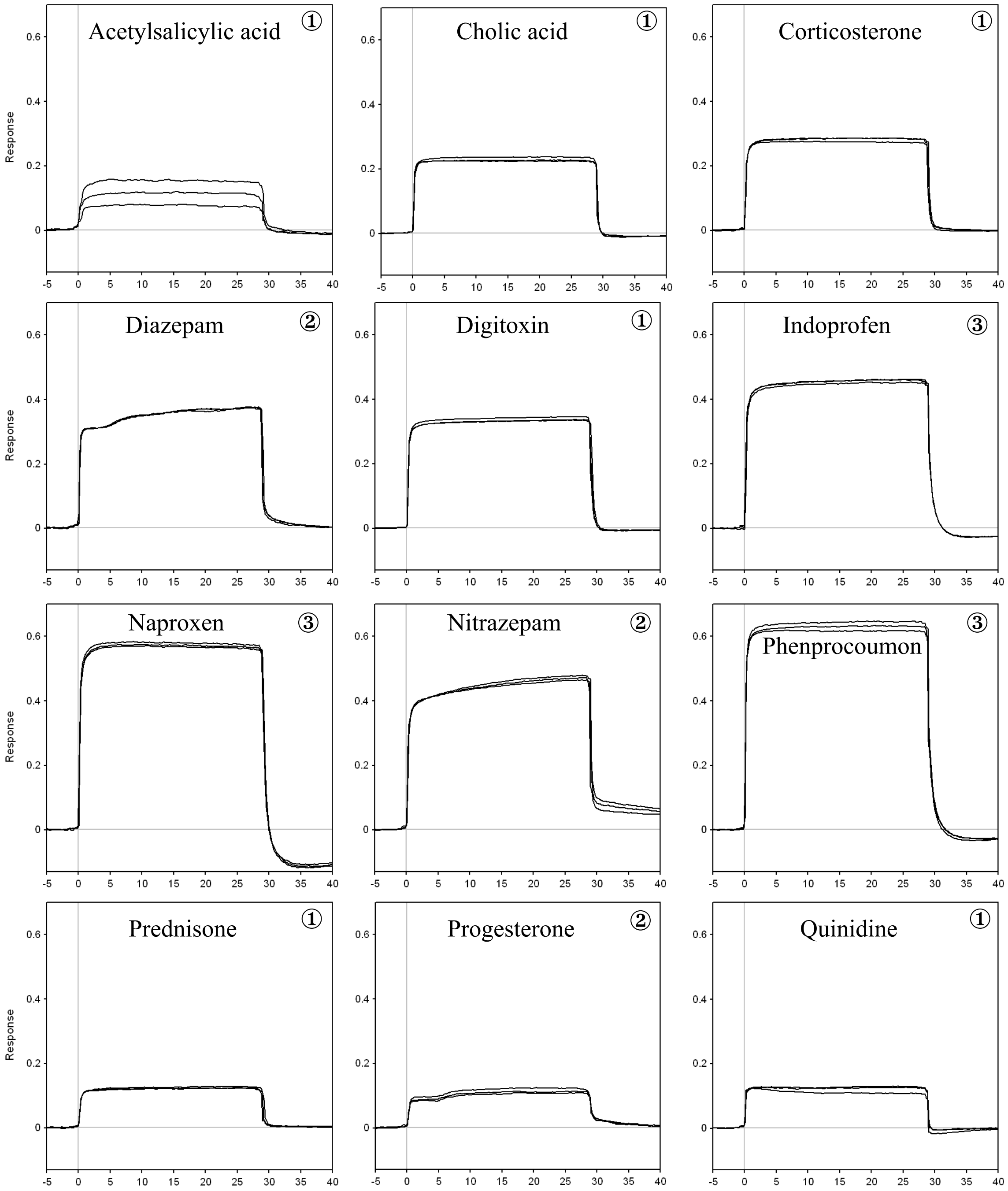
All log P and pKa values were from the PhysPro database (http://www.syrres.com/esc/physdemo.htm)

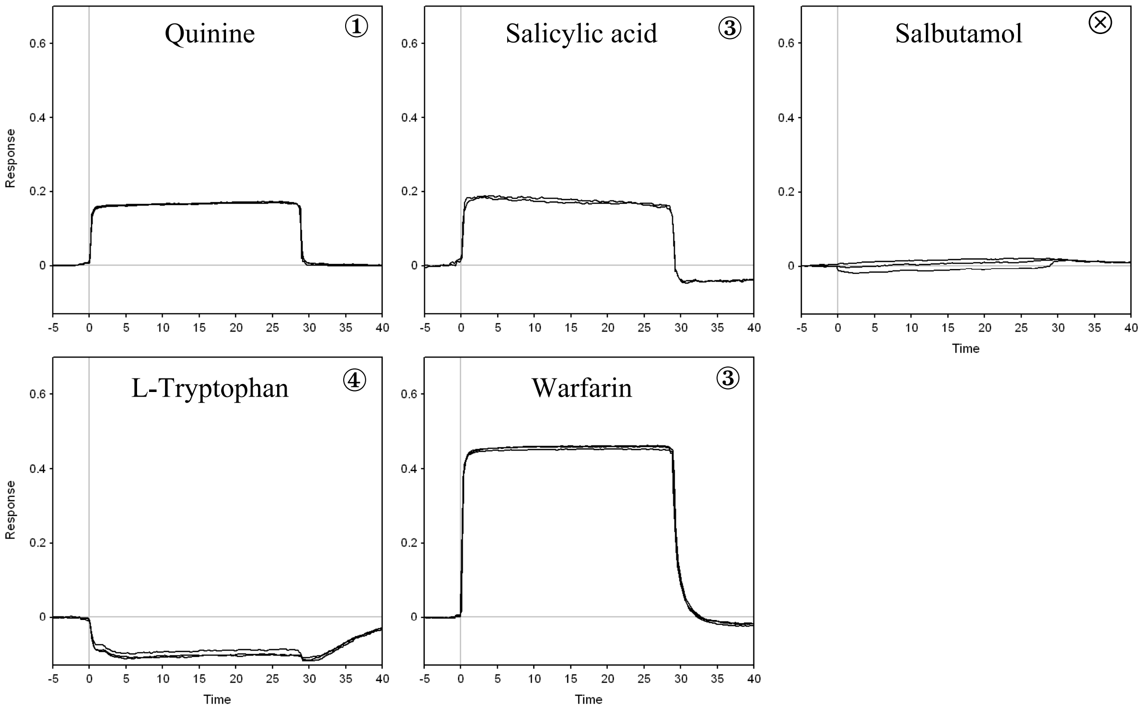
Values from Carter and Ho [8] are from a summary table of different equilibrium dialysis experiments.

1. A. Frostell-Karlsson, A. Remaeus, H. Roos, K. Andersson, P. Borg, M. Hamalainen, R. Karlsson, *J Med Chem* **2000**, *43*, 1986.
2. R. L. Rich, Y. S. Day, T. A. Morton, D. G. Myszka, *Anal Biochem* **2001**, *296*, 197.
3. Y. S. Day, D. G. Myszka, *J Pharm Sci* **2003**, *92*, 333.
4. A. G. Gilman, L. S. Goodman, J. G. Hardman, L. Goodman, A. Gilman, *Goodman and Gilman's the pharmacological basis of therapeutics*, 9th ed., McGraw-Hill, New York etc., **1996**.
5. K. Valko, S. Nunhuck, C. Bevan, M. H. Abraham, D. P. Reynolds, *J Pharm Sci* **2003**, *92*, 2236.
6. *Arzneimittel-Kompendium der Schweiz*, Documed, Basel, **2004**.
7. J. Ostergaard, C. Schou, C. Larsen, N. H. Heegaard, *Electrophoresis* **2002**, *23*, 2842.
8. D. C. Carter, J. X. Ho, *Adv Protein Chem* **1994**, *45*, 153.
9. H. Aki, M. Yamamoto, *J Pharm Sci* **1994**, *83*, 1712.

Appendix B3: Sensorgrams from the ranking assay

A panel of known HSA binders was injected at a constant concentration (333 μM in 10 mM PBS, 3% DMSO) as triplicates and classified according to their curve shapes (see legend at the end of this section).

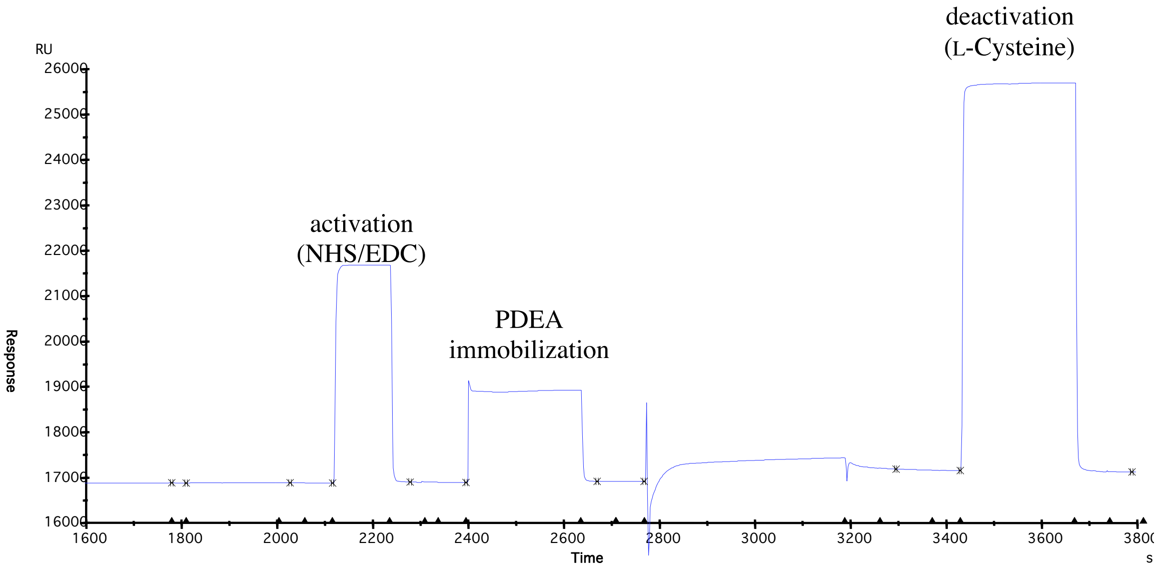




Classifications of the sensorgrams according to their shape:

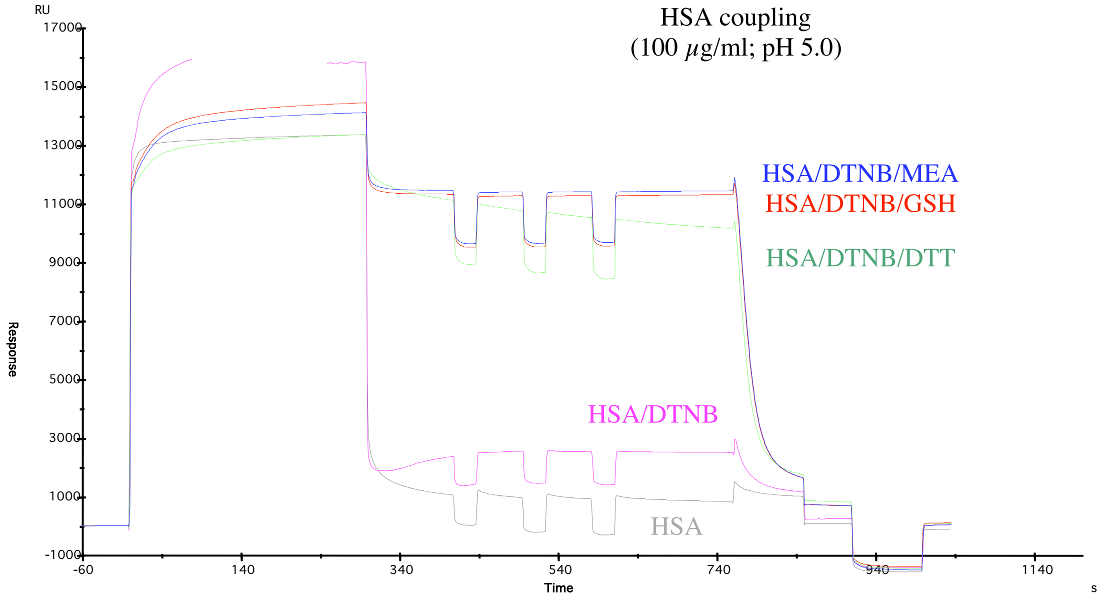
① Proper block signals
② Multiphase association
③ Negative post-injection baselines
④ Negative binding signals
⊗ Not determined (low signal intensity)

Appendix B4: Standard thiol immobilization (PDEA; no DTNB/GSH)



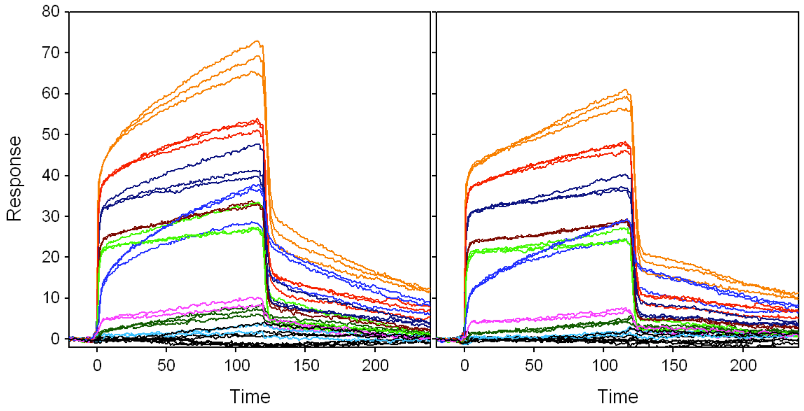
The ligand thiol coupling procedure as recommended by Biacore (see *Biacore Sensor Surface Handbook*) was applied to HSA, but did not result in significant surface densities.

Appendix B5: Influence of reducing agents on thiol coupling (GSH/MEA/DTT)



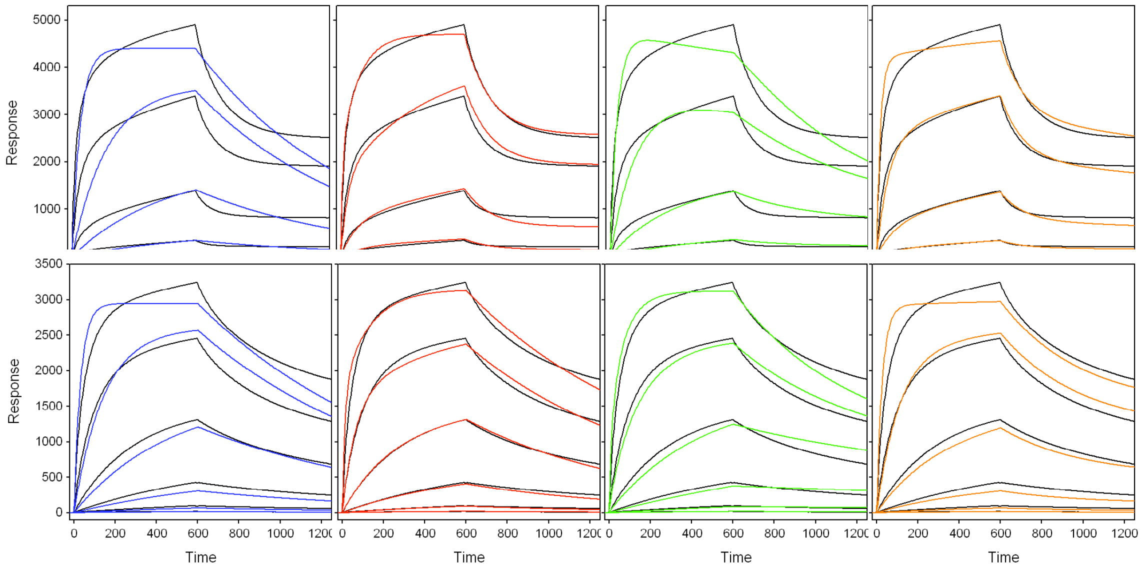
The in situ reduction and activation procedure developed for the thiol coupling of HSA was further evaluated by replacing glutathione (GSH) as the standard reducing agent with mercaptoethylamine (MEA) or dithiothreitol (DTT). Ellmans reagent (DTNB) was added to activate the free thiol group (Cys-34).

Appendix C1: Influence of the reference surface on non-specific binding



Triplicate injection series (0.2-200 μM) of sialyl Lewis^a with Lemieux spacer (sLe^a-Lem) on immobilized antibody GSLA-2. The same data set was referenced against a untreated sensor chip surface (left) or against an immobilized anti-myoglobin antibody (from the 'Biacore 3000 Getting Started Kit'). Even though the antibody reference was able to reduce the non-specific binding signal, it could not be removed completely.

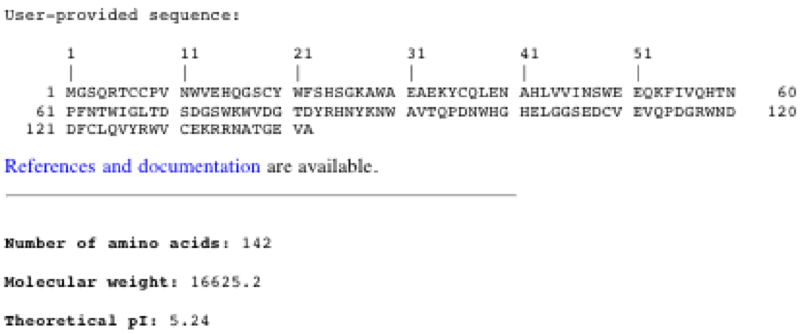
Appendix C2: Kinetic evaluation of GSLA-2 binding to SpA and SpG



Binding of GSLA-2 on immobilized staphylococcal protein A (SpA; top row) and streptococcal protein G (SpG; bottom row). SPR signals (black) are overlaid by kinetic simulations of a Langmuir 1:1 (blue), surface heterogeneity (red), bivalent (green), or conformational change (orange) binding model).

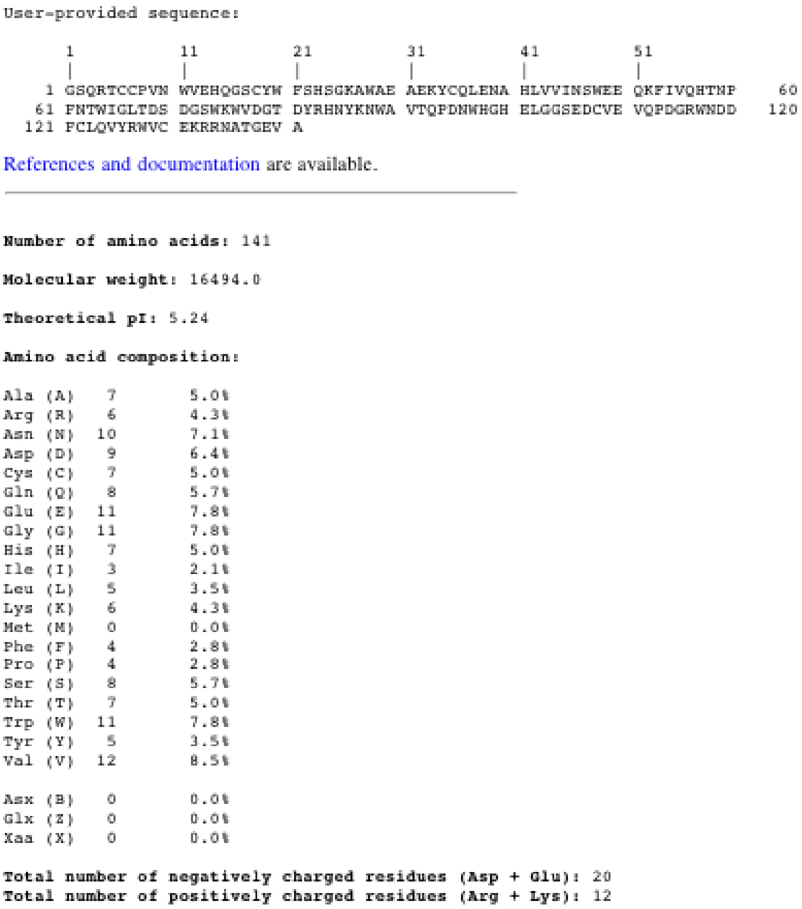
Appendix D1: H1-CRD data from ProtParam (www.expasy.org)

Parameters with the *N*-terminal methionine residue:



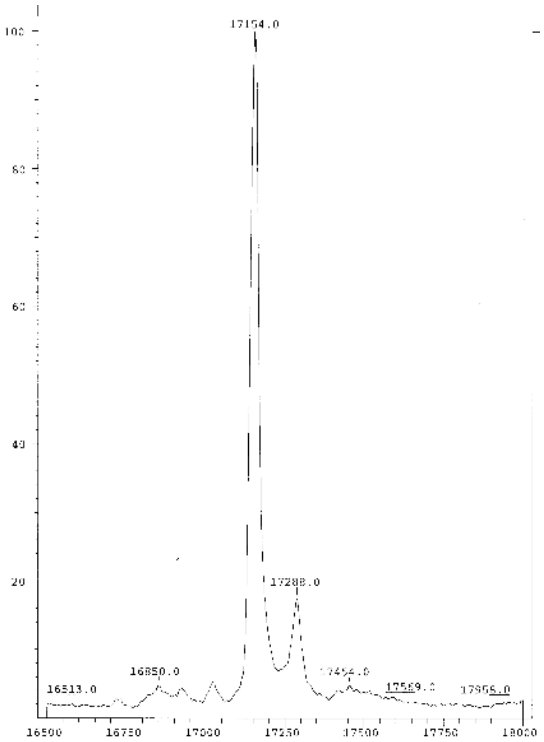
```
User-provided sequence:

        1          11         21         31         41         51
        |          |          |          |          |          |
      1 MGSQRTCCPV NWVEHQGSCY WFSHSGKAWA EAEKYCQLEN AHLVVINSWE EQKFIVQHTN    60
     61 PFNTWIGLTD SDGSWKWVDG TDYRHNYKNW AVTQPDNWHG HELGGSEDCV EVQPDGRWND   120
    121 DFCLQVYRWV CEKRRNATGE VA

References and documentation are available.
________________________________________

Number of amino acids: 142

Molecular weight: 16625.2

Theoretical pI: 5.24
```

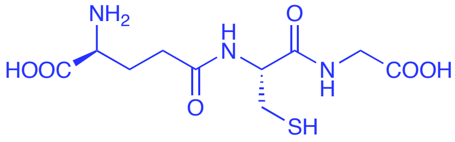
Parameters after removal of the terminal methionine residue:



```
User-provided sequence:

        1          11         21         31         41         51
        |          |          |          |          |          |
      1 GSQRTCCPVN WVEHQGSCYW FSHSGKAWAE AEKYCQLENA HLVVINSWEE QKFIVQHTNP    60
     61 FNTWIGLTDS DGSWKWVDGT DYRHNYKNWA VTQPDNWHGH ELGGSEDCVE VQPDGRWNDD   120
    121 FCLQVYRWVC EKRRNATGEV A

References and documentation are available.
________________________________________

Number of amino acids: 141

Molecular weight: 16494.0

Theoretical pI: 5.24

Amino acid composition:

Ala (A)   7        5.0%
Arg (R)   6        4.3%
Asn (N)  10        7.1%
Asp (D)   9        6.4%
Cys (C)   7        5.0%
Gln (Q)   8        5.7%
Glu (E)  11        7.8%
Gly (G)  11        7.8%
His (H)   7        5.0%
Ile (I)   3        2.1%
Leu (L)   5        3.5%
Lys (K)   6        4.3%
Met (M)   0        0.0%
Phe (F)   4        2.8%
Pro (P)   4        2.8%
Ser (S)   8        5.7%
Thr (T)   7        5.0%
Trp (W)  11        7.8%
Tyr (Y)   5        3.5%
Val (V)  12        8.5%

Asx (B)   0        0.0%
Glx (Z)   0        0.0%
Xaa (X)   0        0.0%

Total number of negatively charged residues (Asp + Glu): 20
Total number of positively charged residues (Arg + Lys): 12
```

Appendix D2: Mass analysis of H1-CRD with glutathione adduct

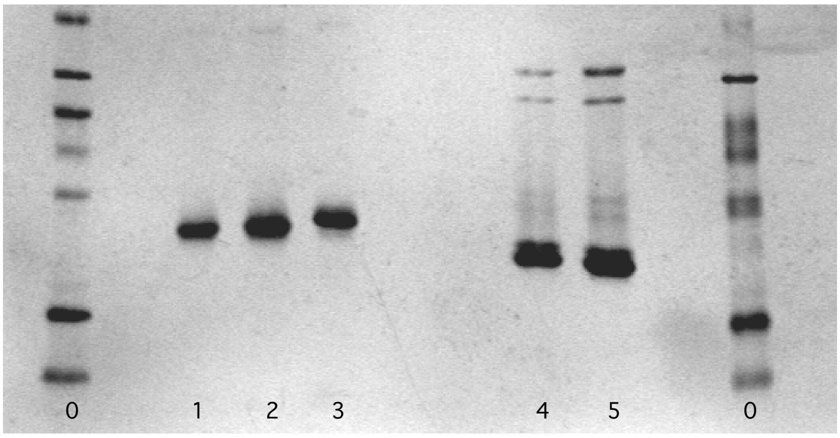




Glutathione (GSH; 307 Da)

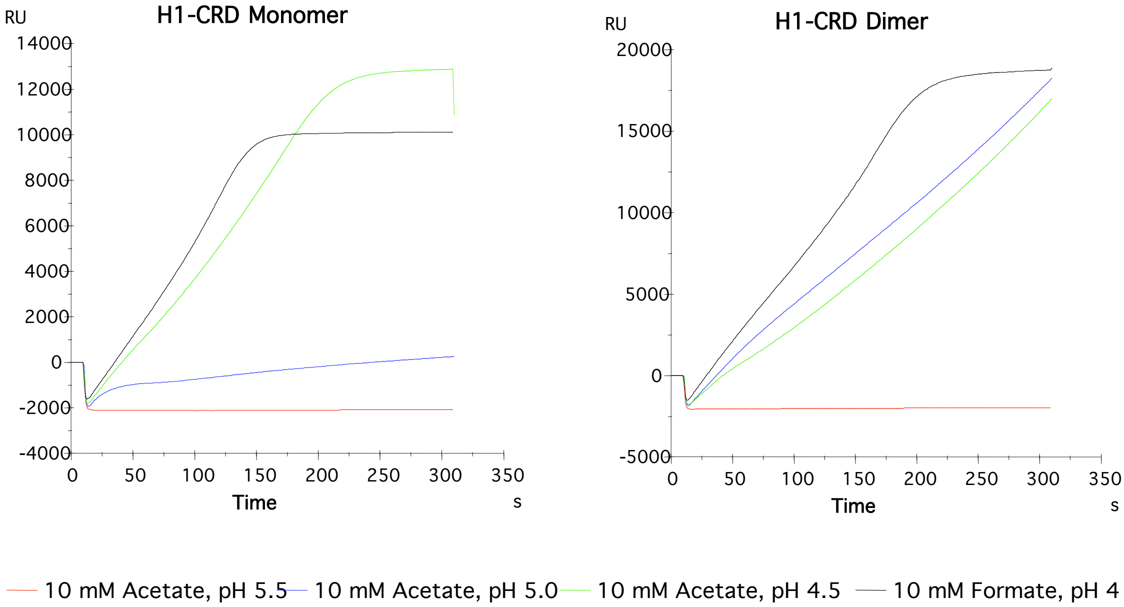
When glutathione was used during the dialysis steps of the H1-CRD refolding procedure, the mass of the native monomer protein shifted from 16,934 Da (2-mercaptoethanol adduct) to 17154 Da. The mass difference of 304 Da to the reduced protein (16,860 Da) corresponds to a glutathione adduct.

Appendix D3: SDS-PAGE of H1-CRD under reducing and non-reducing conditions



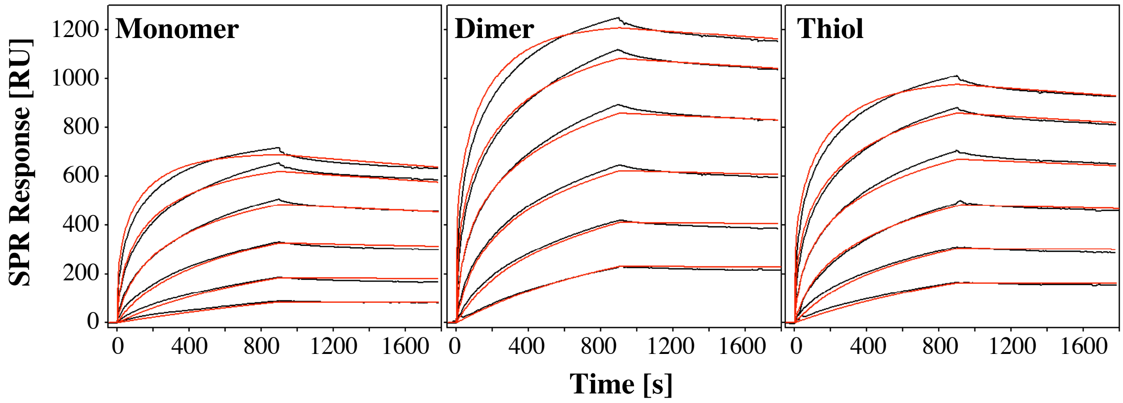
0) LMW marker reduced, 1) H1-CRD reduced, 2) reduced after reduction and alkylation, 3) reduced after alkylation, 4) non-reduced, 5) non-reduced after alkylation, 0) LMW marker non-reduced

Appendix D4: Surface attraction of H1-CRD



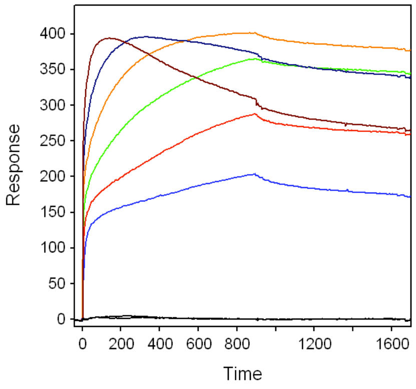
A constant concentration of H1-CRD monomers and dimers (~10-20 μg/ml) was prepared in sodium acetate buffer at different pH (4.0-5.5) and injected over a plain sensor chip surface in order to determine the optimal immobilization buffer (pH scouting).

Appendix D5: Sensorgrams and kinetic fits of ASF binding to different H1 surfaces



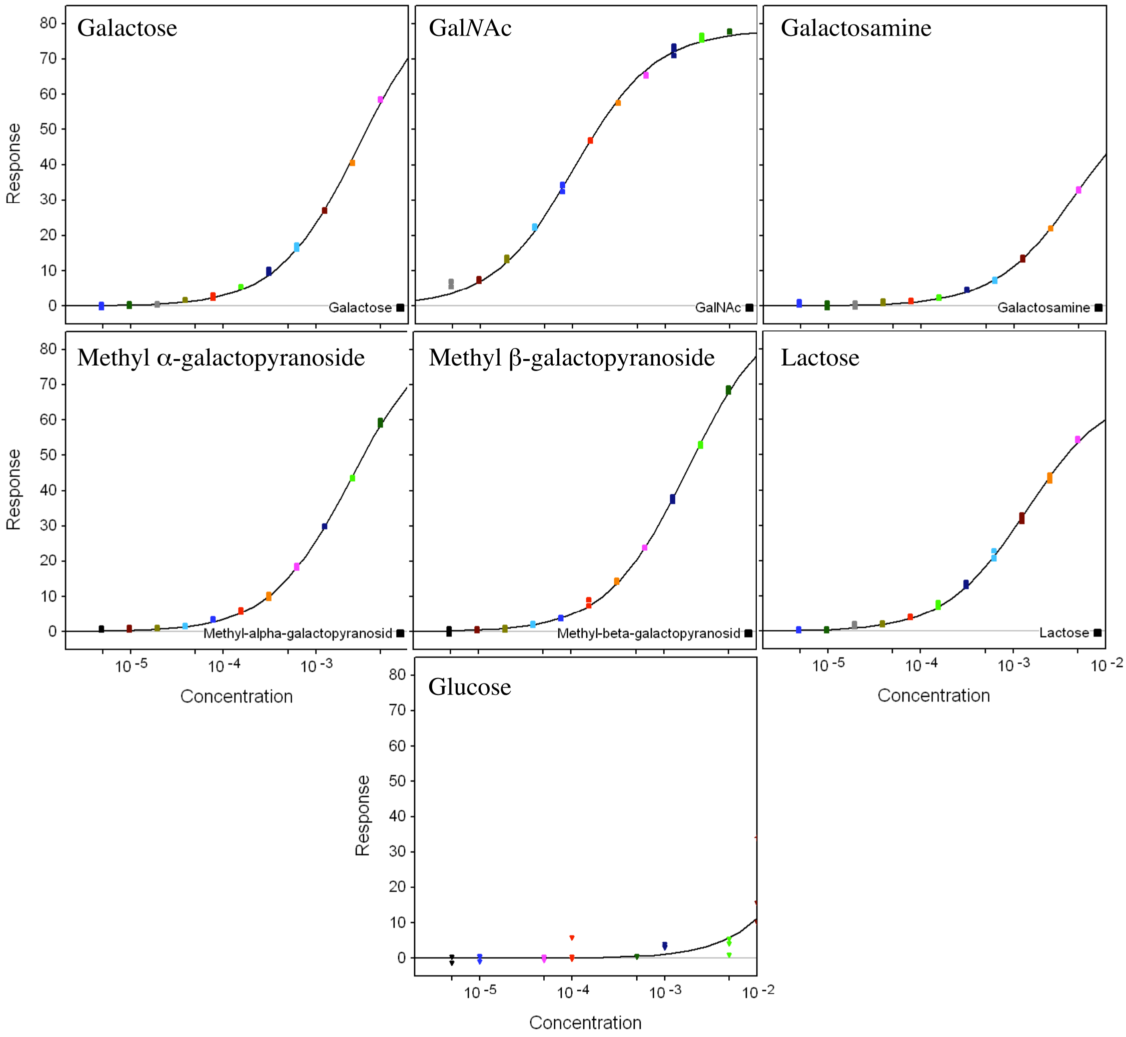
Injection series of asialofetuin (4-1000 nM in 10 mM HBS, 50 mM $CaCl_2$, 0.005% polysorbate 20, *pH* 7.4) on three different H1-CRD surfaces (amine coupling of H1 monomers and dimers, as well as thiol immobilization of H1 monomers). All data sets were kinetically fitted to a bivalent analyte model using *CLAMP* (red curves) and overlaid with the sensorgrams (black curves). The resulting values are shown in *table* 5-6.

Appendix D6: Sensorgrams of ASOR binding to H1-CRD



Binding of asialoorosomucoid (ASOR; 4-1000 nM in 10 mM HEPES pH 7.4, 50 mM $CaCl_2$) to immobilized H1-CRD dimer. In contrast to asialofetuin, ASOR shows a complex binding mode with a biphasic association phase, and a negative equilibrium drift at high concentrations.

Appendix D7: Equilibrium response plots of monovalent galactose derivatives



Affinity analysis of galactose and derivatives (monovalent carbohydrates) after screening over immobilized H1-CRD dimer. Triplicate injections between 5 μM and 5 mM were fitted to a single binding site model.

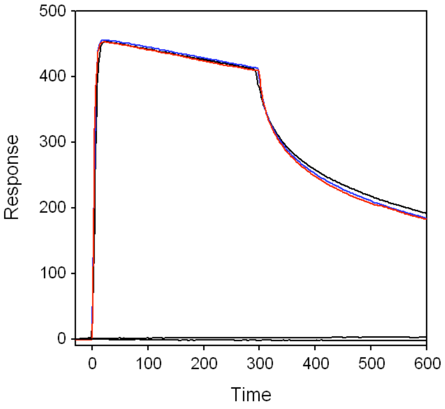
Appendix D8: Blocking properties of anti-H1-CRD mAbs

Binding of anti-H1-CRD antibodies to the carbohydrate binding site was expected to block galactose interaction. A constant concentration of galactose (5 mM) was therefore injected before and after antibody injection (start/end) and the signal difference was calculated (difference). However, the relatively low signal difference made a sensitive detection of blocking properties difficult. None of the antibodies was able to significantly alter the galactose binding signal (difference ≤ 10%).

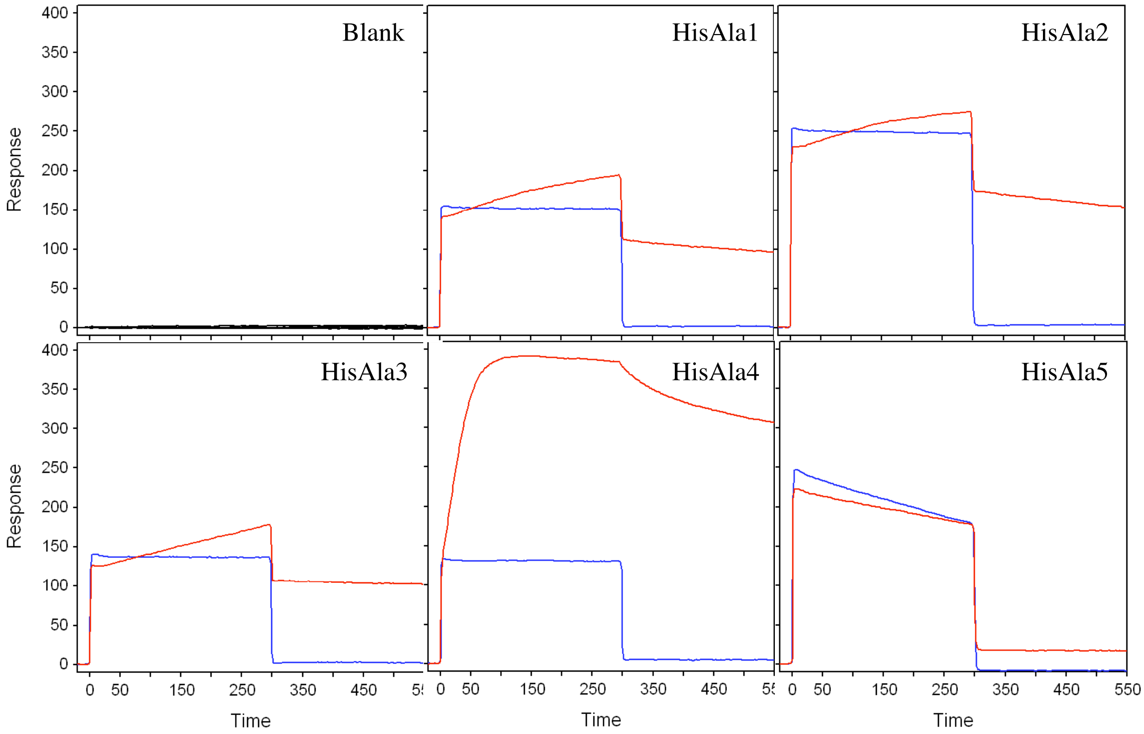
	Monomer Surface			Dimer Surface		
mAb	Start [RU]	End [RU]	Difference [%]	Start [RU]	End [RU]	Difference [%]
Blank	-64.6	-62.8	-2.8	-99.5	-93.2	-6.3
B01.4	-67.8	-64.2	-5.3	-107.0	-114.5	7.0
C09.1	-67.4	-63.9	-5.2	-108.0	-99.4	-8.0
C11.1	-67.7	-64.6	-4.6	-107.7	-103.1	-4.3
C14.6	-68.0	-63.7	-6.3	-107.8	-96.8	-10.2
C18.1	-66.7	-62.0	-7.0	-107.7	-104.9	-2.6
C23.8	-66.4	-64.7	-2.6	-107.9	-98.8	-8.4
C48.9	-65.5	-64.0	-2.3	-114.5	-106.7	-6.8

Appendix E1: Mass transport analysis of the hexahistidine peptide

Mass transport analysis of the hexahistidine peptide (5 μM) at flow rates of 5 (black), 20 (blue), and 50 μl/min (red).

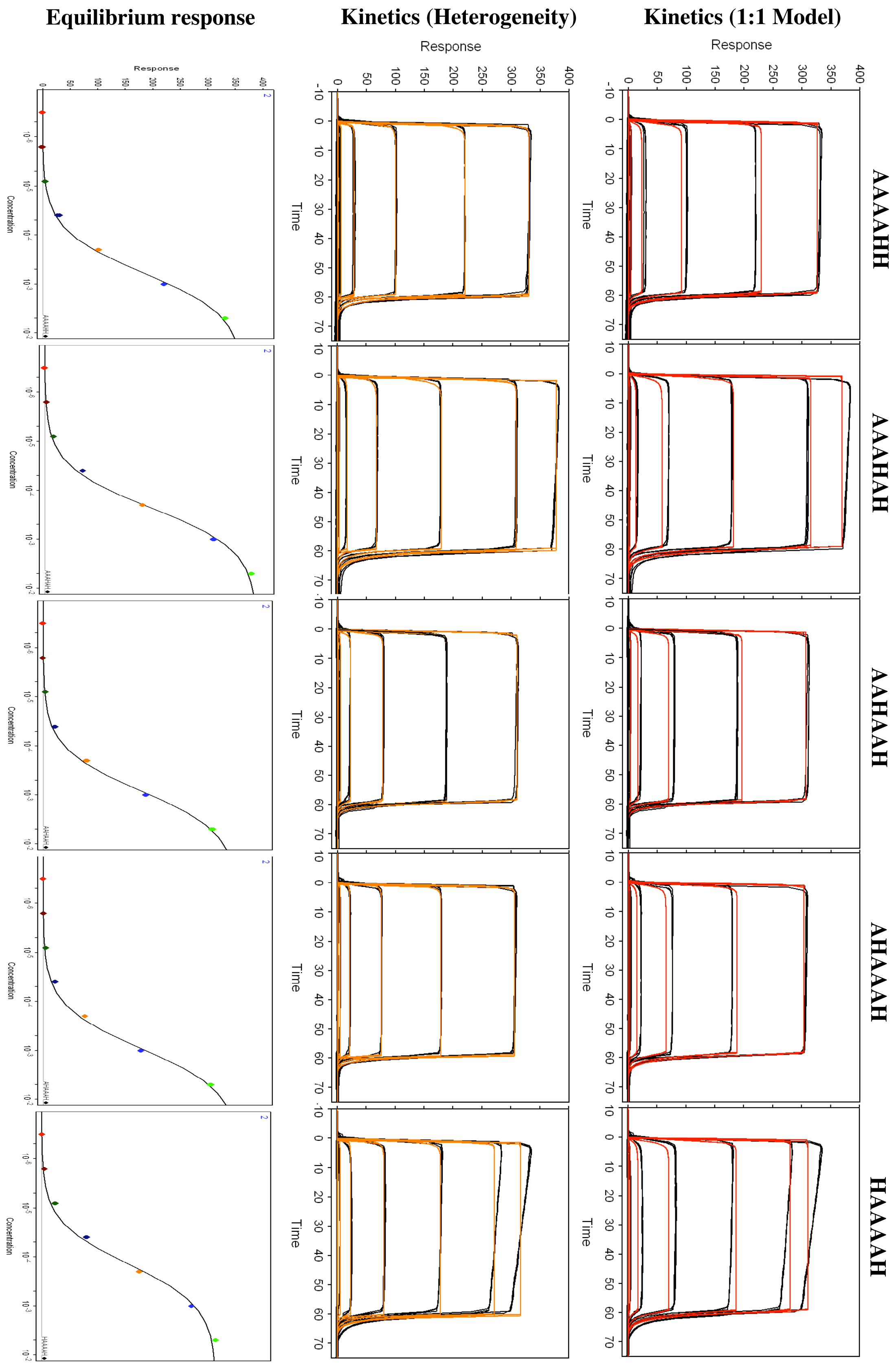


Appendix E2: Non-specific binding of HisAla peptides after standard purification



When the HisAla peptides were purified on HPLC using under standard conditions (water, acetonitrile, 0.1% TFA), the peptides eluted close to the injection peak. Fractions from such separation steps caused a significant non-specific signal on the NTA sensor chip (red sensorgrams). The shift to basic elution conditions (ammonium acetate pH 8.8) allowed the collection of pure peptide fractions, which showed a clear SPR binding signal (blue sensorgrams).

Appendix E3: Kinetic evaluation of HisAla peptides



DANIEL RICKLIN
Metzerstrasse 29 • CH-4056 Basel • Phone +41 613821556 • email: daniel.ricklin@unibas.ch

CURRICULUM VITAE

Date and place of birth	26 September 1974, Zurich
Nationality	Swiss

Education

2001 - 2005	University of Basel, Institute of Molecular Pharmacy PhD thesis in the group of Prof. B. Ernst '*Surface Plasmon Resonance Applications in Drug Discovery (with an emphasis on small molecule and low affinity systems)*'
1993 - 1999	Swiss Federal Institute of Technology (ETH Zurich) Diploma in *Pharmaceutical Sciences*
1988 - 1993	Kantonsschule Wiedikon (grammar school), Zurich

Practical Experience

2000	Proteome Center Rostock, Germany Internship in the group of Prof. M. O. Glocker Study about '*Analysis of antibody profiles in autoimmune diseases using mass spectrometry*' Development of a database for proteomic data
1998 (April – May)	Tumor Biology Center, Freiburg, Germany Practical course in pharmaceutical chemistry in the group of Dr. F. Kratz Development of macromolecular anticancer drugs
1996 - 1997	Albis Apotheke, Zurich Practical year in a public pharmacy

Additional Skills

Languages	German: native speaker English: CAE; certificate in scientific writing French: working knowledge
PC skills	Advanced knowledge in Windows, MacOS, Unix, MS Office applications, database, internet, HTML

Daniel Ricklin
Metzerstrasse 29 • CH-4056 Basel • Phone +41 613821556 • email: daniel.ricklin@unibas.ch

Scientific Qualifications

Papers

The following results from my thesis have been submitted or are in the final stage of preparation:

- Lars Herfurth, Beat Ernst, Beatrice Wagner, Daniel Ricklin, Daniel S. Strasser, John L. Magnani, Andrew J. Benie, Thomas Peters, *Comparative epitope mapping with STD NMR of sialyl Lewis a and derivatives bound to a monoclonal antibody*, J. Med. Chem. **2005,** 48, 6879.
- Daniel Ricklin, Beat Ernst, *pH- and ligand-induced conformational changes of human serum albumin monitored by surface plasmon resonance*, J. Biol. Chem.; submitted.
- Steven Knecht, Daniel Ricklin, Alex N. Eberle, Beat Ernst, *The Oligohis-Tag: Mechanism of Binding to* Ni^{2+}*-NTA surfaces*; Biochemistry, submitted.
- Rita Born, Daniel Ricklin, Karin Johansson, Said Rabbani, Beat Ernst, *Expression and purification of the human hepatic asialogylcoprotein receptor H1-CRD*; in preparation.
- Daniel Ricklin, Daniel S. Strasser, Beat Ernst, *Negative binding signals in Biacore small molecule assays: are they just artefacts or important structural information*; in preparation

Posters

Daniel Ricklin, Beat Ernst, *Optimization of human serum albumin-drug binding assays using surface plasmon resonance biosensors*, PharmaDay 2004, ETH Zurich.

Daniel Strasser, Daniel Ricklin, Beat Ernst, *Analysis of the carbohydrate binding properties of the therapeutic antibody GSLA-2*, Fall Meeting of the Swiss Chemical society 2004, University of Zurich.

Oral presentations

Daniel Ricklin, Marcus Bantscheff, Hans-Jürgen Thiesen, Michael O. Glocker, *Protein analysis by mass spectrometry in clinical research*, Technology Conference 'Biosystems and Medical Technologies', September 2000, Warnemünde, Germany.

Daniel Ricklin, Beat Ernst, *Optimizing Biosensor Data for Low Molecular Weight Assays on Biacore 3000*, Fall Meeting of the Swiss Chemical society 2004, University of Zurich

Educational experience

Lectures offered on '*Pharmaceutically relevant information retrieval*' for students of pharmaceutical sciences

Supervisor of several practical courses in molecular modeling

Supervision of a diploma thesis about '*Development of methods for the selective coupling of target proteins on biosensors*'

Initial supervision and training of several doctoral students in the use of Biacore

Initiation and organization 'PharmaDay 2003' (information event for students interested in pharmaceutical sciences)

Courses

Graduate course 'Key issues in drug discovery and development'

Numerous lectures and seminars (drug discovery, analytics, molecular medicine) as well as attendance at congresses