

---

Die Verbindung von Kleinwinkelstreuung und  
*in silico*-Methodik zur Aufklärung von  
Konformationswechseln großer Proteinkomplexe

---

Dissertation  
Zur Erlangung des Grades  
Doktor der Naturwissenschaften

Am Fachbereich Biologie  
Der Johannes Gutenberg-Universität Mainz

Christian Meesters  
geboren am 31. März 1976 in Goch

Dekan:                    xxxxxxxxxxxx  
                             xxxxxxxxxxxx  
                             xxxxxxxxxxxx

1. Berichterstatter:    xxxxxxxxxxxx  
                             xxxxxxxxxxxx  
                             xxxxxxxxxxxx

2. Berichterstatter:    xxxxxxxxxxxx  
                             xxxxxxxxxxxx  
                             xxxxxxxxxxxx

Tag der mündlichen Prüfung: 28. November 2008



It is easy to get carried away with the infinite number of possible adjustments offered by POVray. One should attempt nevertheless to maintain a careful balance between artistic ambition and scientific message. The beauty of an image can always be improved, its scientific content not.

Ansgar Philippsen

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung – Konformationswechsel in Proteinkomplexen</b>	<b>1</b>
1.1	Proteine und Konformationen . . . . .	1
1.2	Hämocyanine . . . . .	2
1.2.1	Arthropodenhämocyanine . . . . .	3
1.2.1.1	Struktureller Aufbau . . . . .	3
1.2.1.2	Kooperativität & Allosterie . . . . .	4
1.2.1.3	Das Hämocyanin von <i>Eurypelma californicum</i> . . . . .	6
1.2.1.4	Das Hämocyanin von <i>Pandinus imperator</i> . . . . .	7
1.2.1.5	Die Hämocyanine von <i>Palinurus elephas</i> und <i>Palinurus argus</i> . . . . .	8
1.2.1.6	Das Hämocyanin von <i>Porcellio scaber</i> . . . . .	9
1.2.1.7	<i>Upogebia pusilla</i> -Hämocyanin . . . . .	10
1.2.2	<i>Nautilus pompilius</i> -Hämocyanin . . . . .	11
1.3	Bla-g Hexamerin von <i>Blattella germanica</i> . . . . .	12
1.4	Struktur & Funktion des $\alpha$ -Hämolysins . . . . .	13
<b>2</b>	<b>Einführung in die Theorie der Kleinwinkelstreuung</b>	<b>16</b>
2.1	Erzeugung von Röntgen- und Neutronenstrahlung . . . . .	16
2.1.1	Röntgenquellen . . . . .	16
2.1.1.1	Synchrotronquellen . . . . .	16
2.1.1.2	Röntgeneratoren . . . . .	16
2.1.2	Neutronenquellen . . . . .	17
2.2	Eigenschaften von Röntgen- und Neutronenstrahlung . . . . .	17
2.2.1	Energie und Impulsübertrag . . . . .	17
2.2.2	Streuung & die Debye-Gleichung . . . . .	19
2.2.3	Streulänge und Wirkungsquerschnitt . . . . .	21
2.2.4	Kohärenter und inkohärenter Wirkungsquerschnitt bei Neutronen . . . . .	21
2.2.5	Kontrast . . . . .	22
2.3	Geometrie der Kratky-Kamera und Konvolution . . . . .	22
2.4	Indirekte Fouriertransformation . . . . .	24
2.5	Der Gyrationradius und die Guinierapproximation . . . . .	25
<b>3</b>	<b>Material und Methoden</b>	<b>28</b>
3.1	Verwendete Chemikalien und Geräte . . . . .	28
3.1.1	Chemikalien . . . . .	28
3.2	Gelelektrophoresen . . . . .	28
3.2.1	SDS-Gelelektrophorese und Coomassie-Färbung . . . . .	28
3.2.2	Native Polyacrylamidgele . . . . .	29
3.2.3	Anfärben von Phenoxidasen nach Gelelektrophorese . . . . .	29
3.3	Aufreinigung der Hämocyanine . . . . .	29
3.3.1	Verwendete Puffer . . . . .	29
3.3.2	Tierhaltung . . . . .	30
3.3.3	Gewinnung und Aufreinigung . . . . .	30
3.3.4	Kontrolle der Integrität nach Bestrahlung . . . . .	30
3.3.5	Deoxygenierung . . . . .	31

3.4	Lichtstreuung . . . . .	31
3.4.1	Statische Lichtstreuung . . . . .	31
3.4.2	Dynamische Lichtstreuung . . . . .	31
3.5	Analytische Ultrazentrifugation . . . . .	32
3.6	Messungen an der Kratky-Kamera . . . . .	32
3.6.1	Aufbau der Kratky-Kamera . . . . .	32
3.6.1.1	Der Röntngenerator . . . . .	32
3.6.1.2	Der Detektor . . . . .	32
3.6.1.3	Aufbau von Zusatzgeräten . . . . .	32
3.6.2	Vorgehensweise . . . . .	33
3.6.3	Wartung und Kalibrierung . . . . .	34
3.6.4	Auswertung der Streudaten . . . . .	34
3.6.4.1	Auswertung der Primärdaten . . . . .	34
3.6.4.2	Entschmierung der Daten . . . . .	34
3.6.5	Hitzedenaturierung an der Kratky-Kamera . . . . .	36
3.6.6	Kratky-Plots zum Verfolgen der Denaturierung . . . . .	36
3.6.7	Massenbestimmung mittels Kratky-Plots . . . . .	37
3.6.8	Berechnung theoretischer Streukurven aus atomaren Proteinmodellen . . . . .	37
3.6.9	Berechnung theoretischer Streukurven von 3D-Modellen aus der Elektronenmikroskopie . . . . .	37
3.6.10	Entwicklung von dreidimensionalen <i>ab initio</i> Modellen . . . . .	38
3.6.11	Die Verwendung von <b>parmscan</b> zur Konformationsberechnung von 24meren . . . . .	39
3.7	Elektronenmikroskopie . . . . .	40
3.7.1	Probenpräparation und Aufnahmen . . . . .	40
3.7.2	Bildbearbeitung . . . . .	40
3.7.2.1	Einzelpartikelanalyse . . . . .	40
3.7.2.2	Elektronenkristallographie . . . . .	41
3.8	Molekulare Modellierung und Molekulardynamiksimulation . . . . .	41
3.8.1	Molekulare Modellierung mit <b>MODELLER</b> . . . . .	41
3.8.2	Auswertung und Beurteilung der Strukturmodelle . . . . .	42
3.8.3	Molekulardynamiksimulationen . . . . .	42
3.8.4	„ <i>Morphing</i> “ . . . . .	43
3.8.5	<i>rigid body</i> -Modellierung des <i>Palinurus elephas</i> -Hämocyanins auf der Basis von SAXS-Daten . . . . .	43
3.9	Charakterisierung von Hämocyaninen unter SDS-Einfluss . . . . .	44
3.9.1	Besondere Erwägungen zur Röntgenkleinwinkelstreuung . . . . .	44
3.9.2	Nachweismöglichkeiten für Phenoloxidasen . . . . .	44
3.9.2.1	Kontrolle der Phenoloxidaseaktivität im Polyacrylamidgel . . . . .	44
3.9.2.2	„Dot-Blot“-Assay zur Aktivitätskontrolle . . . . .	44
3.9.3	Verifikation der Detergens-CMC . . . . .	45
3.10	$\alpha$ -Hämolysin . . . . .	46
3.10.1	Expression und Aufreinigung . . . . .	46
3.10.2	Konzentrationsbestimmung des $\alpha$ -Hämolysins . . . . .	47
3.10.3	Gefriertrocknung des aufgereinigten Proteins . . . . .	47
3.10.4	Rekonstitution des $\alpha$ -Hämolysins . . . . .	48
3.10.4.1	Bestimmung des Toxintiters . . . . .	48
3.10.4.2	Reinigung der Rundkolben zur Lipidpräparation . . . . .	48
3.10.4.3	Präparation von Lipiden . . . . .	49
3.10.4.4	Herstellung von Vesikeln . . . . .	49
3.10.4.5	Bestimmung des Cholesterolgehaltes von Vesikeln . . . . .	49
3.10.5	Röntgenkleinwinkelstreuung am Monomer des $\alpha$ -Hämolysins . . . . .	50
3.10.6	Neutronenkleinwinkelstreuung am $\alpha$ -Hämolysin . . . . .	50
3.10.6.1	Bestimmung des Matchpunktes . . . . .	50

3.10.6.2	Durchführung der Kleinwinkelstreuung . . . . .	51
3.10.6.3	Berechnung der Neutronenstreuendichte . . . . .	51
3.10.7	Berechnung der Massenverteilung innerhalb geometrischer Körper . . . . .	51
3.10.8	Errechnen und Darstellung der Sequenzkonservierung . . . . .	52
3.11	Hitzedenaturierung mittels <i>Differential Scanning Calorimetry</i> . . . . .	53
3.12	Fluoreszenzspektroskopie . . . . .	53
3.13	Circulardichroismusspektroskopie . . . . .	53
3.14	Darstellung von Molekülgraphiken . . . . .	54
3.15	Darstellung von Dichten . . . . .	54
3.16	Darstellung von Phylogrammen . . . . .	54
3.17	Darstellung von Kleinwinkelstreuungsdaten mit SPlot . . . . .	54
<b>4</b>	<b>Resultate</b>	<b>55</b>
4.1	Konformationsübergänge von Hämocyaninen . . . . .	55
4.1.1	Die Struktur 24merer Hämocyanine von <i>Eurypelma californicum</i> und <i>Pandinus imperator</i> unter Einfluss verschiedener Effektoren . . . . .	55
4.1.1.1	Reinigung und Kontrolle des Hämocyanins von <i>Eurypelma californicum</i> und <i>Pandinus imperator</i> . . . . .	55
4.1.1.2	Oxygenierung und Deoxygenierung . . . . .	56
4.1.1.3	Der Einfluss von SDS auf die Konformation . . . . .	57
4.1.2	Die Struktur des Hämocyanins von <i>Palinurus elephas</i> unter Einfluss verschiedener Effektoren . . . . .	63
4.1.2.1	Modellierung der Struktur . . . . .	63
4.1.2.2	SAXS am Hämocyanin von <i>Palinurus elephas</i> und <i>Palinurus argus</i> unter Einfluss verschiedener Effektoren . . . . .	65
4.1.2.3	<i>rigid body</i> -Modellierung auf Basis der SAXS-Daten . . . . .	68
4.1.3	Die Struktur des Hämocyanins von <i>Upogebia pusilla</i> . . . . .	74
4.1.3.1	SAXS-Daten . . . . .	74
4.1.3.2	Elektronenmikroskopische Daten . . . . .	75
4.1.4	Die Struktur des Hämocyanins von <i>Porcellio scaber</i> . . . . .	78
4.2	Vergleich elektronenmikroskopische Daten und Kleinwinkelstreuungsdaten von <i>N. pompilius</i> -Hämocyanin . . . . .	79
4.3	Elektronenmikroskopie am Bla-g Hexamerin von <i>Blattella germanica</i> . . . . .	81
4.4	Konformationen des $\alpha$ -Hämolsins von <i>Staphylococcus aureus</i> . . . . .	83
4.4.1	Reinigung des Toxins . . . . .	83
4.4.2	Die monomere Struktur des Hämolsins . . . . .	84
4.4.2.1	Kleinwinkelstreuungsdaten . . . . .	84
4.4.2.2	Modellierung und Verifikation der Monomerstruktur . . . . .	85
4.4.2.3	Rückgratflexibilität des $\alpha$ -Hämolsins . . . . .	89
4.4.2.4	Evolutionäre Aspekte . . . . .	90
4.4.3	Die Präporenstruktur des $\alpha$ -Hämolsins . . . . .	91
4.4.3.1	Vergleich der Struktur zwischen $\alpha$ -Hämolsinwildtyp und H35R-Mutante . . . . .	91
4.4.3.2	Matchpunktbestimmung der Lipide . . . . .	92
4.4.3.3	Aufnahme und Interpretation von SANS-Daten . . . . .	93
4.4.3.4	Elektronenmikroskopische Daten . . . . .	95
4.4.3.5	Konformationswechsel in den Monomeren während der Oligomerisierung . . . . .	97
4.5	Molekulargewichtsbestimmung mit Hilfe des Kratky-Plots . . . . .	102
4.6	Verfolgung einer Hitzedenaturierung mit der Kratky-Kamera am Beispiel von Lysozym . . . . .	104
4.7	Eine <i>Software</i> zur <i>rigid body</i> -Modellierung von Proteinen auf SAS-Datenbasis . . . . .	105
4.7.1	Implementierung . . . . .	107
4.7.1.1	Programdesign . . . . .	107

4.7.1.2	Implementation von Standardalgorithmen . . . . .	107
4.7.1.3	Berechnung der theoretischen Abstandsverteilungsfunktion und Streukurven . . . . .	108
4.7.1.4	Softwaremetrik . . . . .	108
4.7.1.5	Die Verwirklichung von Symmetrieoperationen . . . . .	109
4.7.2	Ein <i>Monte Carlo</i> -Algorithmus mit <i>Simulated Annealing</i> . . . . .	110
4.7.3	Ein <i>grid search</i> -Algorithmus zur verfeinerten Suche im Parameterraum . . . . .	111
4.7.4	Vergleich mit <code>parmscan</code> . . . . .	113
4.7.5	Ein <i>Monte Carlo</i> -Algorithmus zur Darstellung eines Wahrscheinlichkeitsraums . . . . .	115
4.7.6	Ein Algorithmus zur Berechnung der Neutronenstreu­längendichte von Proteinen . . . . .	116
4.7.7	Webhosting der Software . . . . .	116
4.8	Eine graphische Benutzeroberfläche zum schnellen Plotten von Kleinwinkelstreu­daten	117
<b>5</b>	<b>Diskussion</b>	<b>119</b>
5.1	Die Struktur 24merer Hämocyanine von <i>Eurypelma californicum</i> und <i>Pandinus imperator</i> unter Einfluss verschiedener Effektoren . . . . .	119
5.1.1	Einfluss der Sauerstoffbindung auf die Konformation . . . . .	119
5.1.2	Einfluss von SDS auf die Konformation . . . . .	120
5.1.2.1	Verifikation der SDS-CMC . . . . .	120
5.1.2.2	Einfluss von SDS auf das Hämocyanin von <i>Eurypelma californicum</i>	121
5.1.2.3	Einfluss von SDS auf das Hämocyanin von <i>Pandinus imperator</i> . . . . .	122
5.2	Konformationsübergänge in den hexameren Hämocyaninen von <i>Palinurus elephas</i> und <i>Palinurus argus</i> . . . . .	123
5.2.1	Homologiemodellierung für das Hämocyanin von <i>Palinurus elephas</i> . . . . .	123
5.2.2	Einfluss von Urat und Koffein auf die Struktur der Hämocyanine . . . . .	124
5.3	Strukturelle Charakterisierung des Hämocyanins von <i>Porcellio scaber</i> . . . . .	128
5.4	Hitzedenaturierung mit der Kratky-Kamera . . . . .	129
5.5	Massenbestimmung mit dem Kratky-Plot . . . . .	129
5.6	Die Struktur des Bla-g Hexamerins von <i>Blattella germanica</i> . . . . .	130
5.7	Die strukturelle Charakterisierung des Hämocyanins von <i>Upogebia pusilla</i> . . . . .	131
5.8	Der Vergleich von elektronenmikroskopischen Daten und Kleinwinkelstreu­daten bei dem Hämocyanin von <i>Nautilus pompilius</i> . . . . .	132
5.9	Strukturcharakterisierungen des $\alpha$ -Hämolysins . . . . .	135
5.9.1	Charakterisierung des Monomers in Lösung . . . . .	135
5.9.2	Vergleich von Präpore (H35R) und Wildtyp . . . . .	138
5.9.3	Spekulationen zum Konformationsübergang von $\alpha_7^*$ nach $\alpha_7$ . . . . .	139
5.10	<i>rigid body</i> -Modellierung von Proteinen auf SAXS-Datenbasis – eine Softwarebibliothek . . . . .	142
5.10.1	Beurteilung der Qualität . . . . .	142
5.10.2	Vergleich mit anderen Programmen zur <i>rigid body</i> -Modellierung . . . . .	142
5.10.3	Ausblick auf zukünftige Versionen der Software . . . . .	143
5.11	Das Programm SPlot . . . . .	144
<b>6</b>	<b>Zusammenfassung und Ausblick</b>	<b>146</b>
	<b>Literaturverzeichnis</b>	<b>156</b>
<b>A</b>	<b>Softwareskripte zur Evaluierung von Röntgenkleinwinkelstreuungsdaten</b>	<b>157</b>
A.1	Skripte zur Steuerung der <i>EMBL-SAXS</i> Software Suite . . . . .	157
A.1.1	DAMMIN . . . . .	157
A.1.2	DAMAVR . . . . .	159
A.1.3	dammin2dino . . . . .	160
A.2	Evaluation von <code>parmscan</code> -Läufen . . . . .	163

---

A.3	MODELLER Skripte . . . . .	163
A.3.1	<i>Loop Modelling</i> . . . . .	165
A.4	Vergleich von Elektronenmikroskopie und SAXS-Daten . . . . .	166
A.5	Umgang mit Pipealign . . . . .	166
A.6	PreToExcel . . . . .	169
<b>B</b>	<b>Hinweise für Betrieb und Wartung der Kratky-Kamera und des Zubehörs</b>	<b>170</b>
B.1	Besonderheiten bei der Wartung . . . . .	170
B.2	Proben in die Kapillare einspritzen und entnehmen . . . . .	170
<b>C</b>	<b>Bisherige Publikationen</b>	<b>171</b>
C.1	Originalarbeiten . . . . .	171
C.2	Erwähnungen in Danksagungen . . . . .	172
C.3	Veröffentlichte Software . . . . .	173
C.4	Reporte für Experimente an Großforschungsanlagen . . . . .	173
C.5	Betreute Arbeiten . . . . .	173
C.6	Vorträge . . . . .	173
C.7	Poster . . . . .	174
<b>D</b>	<b>Bildrechte</b>	<b>175</b>

# Abkürzungsverzeichnis

SI-Einheiten (Rümcker, 1978) und deren dezimale Vielfache werden nicht aufgeführt.

Å	$10^{-10}$ m
AFM	<i>Atomic Force Microscopy</i> bzw. Rasterkraftmikroskopie
AG	Arbeitsgruppe
a. u.	<i>arbitrary unit</i> – „willkürliche Einheit“ angegeben, wenn die Angabe absoluter Einheiten physikalisch nicht sinnvoll ist.
bar	veraltete Einheit für den Luftdruck ( $\hat{=} 10^5$ Pa)
$\beta$ -CFTs	<i><math>\beta</math>-channel forming toxins</i> bzw. $\beta$ -kanalformende Toxine
CPC	1-Hexadecylpyridiniumchlorid oder Cetylpyridiniumchlorid
CPK	Moleküldarstellung nach Corey, Pauling und Koltun (siehe <a href="http://www.netsci.org/Science/Compchem/feature14b.html">http://www.netsci.org/Science/Compchem/feature14b.html</a> )
CTF	<i>contrast transfer function</i> (Kontrasttransferfunktion eines Mikroskops)
2D / 3D	zweidimensional / dreidimensional
$D_{max}$	maximaler Abstand im Molekül
Da	Dalton; Äquivalent zur atomaren Masseneinheit u $1 \text{ Da} = 1,66 \cdot 10^{-27} \text{ kg}$
DAM	<i>Dummy Atom Modell</i>
DOPE	diskrete optimierte Protein Energie (Shen & Sali, 2006)
EM	Elektronenmikroskop
FE	funktionale Einheit (eines Molluskenhämocyanins)
FPLC	<i>Fast Performance Liquid Chromatography</i>
FRET	<i>Fluorescence Resonance Energy Transfer</i> Fluoreszenzresonanzenergietransfer
FWHM	<i>Full Width Half Maximum</i> (Halbwertsbreite)
GBIF	<i>Global Biodiversity Information Facility</i> - <a href="http://www.gbif.org">www.gbif.org</a>
HAC	<i>hierarchical ascendant classification</i> - hierarchische, „auf Vorfahren basierende“ Klassifizierung
IFT	Indirekte Fourier Transformation
ITC	Isotherme Titrationskalorimetrie
KLH	<i>Keyhole Limpet Hemocyanin</i>
LB	<i>lysogeny broth</i> - ein komplexes Nährmedium für Bakterien
$M_{wt}$	Molekulargewicht
MBTH	3-Methyl-2-benzo-thiazolin Hydrazon Hypochlorid Hydrat
MOPS	3-[N-Morpholin]-Propansulfonsäure
NMR	<i>Nuclear Magnetic Resonance</i> Kernspinresonanzspektroskopie
OD	Optische Dichte
$P(r)$	Abstandsverteilungsfunktion (siehe Abschnitt 2.2.2 auf Seite 19)
PAGE	Polyacrylamid-Gelelektrophorese
PDB	Dateiformat der Protein Data Bank (Berman <i>et al.</i> , 2000)
PSD	Position Sensitive Detector
$R_g$	Gyrationsradius (siehe Abschnitt 2.5 auf Seite 25)

---

RMSD	<i>root mean square distance</i> (mittlere quadratische Abweichung)
RMSF	<i>root mean square fluctuation</i> (mittlere quadratische Fluktuation)
SANS	<i>Small Angle Neutron Scattering</i> (Neutronenkleinwinkelstreuung)
SAXS	<i>Small Angle X-ray Scattering</i> (Röntgenkleinwinkelstreuung)
UAc	Uranylacetat
UV	Ultraviolett
$v/d$ Waals	van der Waals
Vis	Visible; sichtbarer Bereich des Lichtes
2 x YT	zweifache Menge von Hefeextrakt ( <i>yeast extract</i> ) und Trypton ( <i>tryptone</i> ) im Vergleich zu LB-Medium

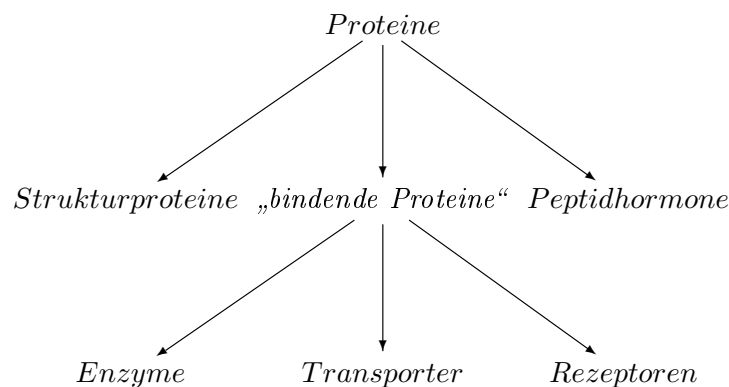


# Kapitel 1

## Einleitung – Konformationswechsel in Proteinkomplexen

### 1.1 Proteine und Konformationen

Proteine spielen eine Schlüsselrolle für fast alle Funktionen der Zelle. Um einen Einstieg in die Thematik der Proteine zu finden, lassen sie sich ganz grob in einige Klassen einteilen:



Zu den Strukturproteinen zählen beispielsweise die Proteine des Zytoskeletts oder das bekannte Kollagen. Peptidhormone sind vor allem bekannt durch Vertreter wie das Insulin und Somatotropin. Sie werden oft nicht als Proteine im engeren Sinn betrachtet. Die übrigen Proteine lassen sich gliedern in Enzyme, Transporter und Rezeptoren. Die Trennung ist nicht scharf<sup>1</sup>, so können einige Enzyme auch Transporter sein oder einige Rezeptoren auch enzymatische Funktion aufweisen. In dieser Arbeit werden u. a. Sauerstofftransporter behandelt, die sekundär zu Enzymen konvertieren können. Dies wird eingehend in den nächsten Abschnitten behandelt.

Der Aufbau von Proteinen ist wie folgt zu beschreiben:

*Primärstruktur* → *Sekundärstruktur* → *Tertiärstruktur* → *Quartärstruktur*

Primärstruktur bezeichnet lediglich die Abfolge der Aminosäuren, verknüpft über Peptidbindungen. Mit Sekundärstruktur ist die räumliche Anordnung der Primärstruktur, insbesondere durch Wasserstoffbrücken zwischen  $C_{\alpha}$ -Atomen und Stickstoffatomen des Peptidrückgrats, gemeint. Bekannte Sekundärstrukturelemente sind  $\alpha$ -Helices und  $\beta$ -Faltblätter. Dazu zählen aber auch eine Reihe von spezifischen und klassifizierten „Schleifen“. Von Tertiärstruktur spricht man bei der räumlichen Organisation von Sekundärstruktureinheiten zueinander. Und eine Quartärstruktur schließlich liegt vor, wenn sich verschiedene Proteine, man spricht hier von Untereinheiten, zu einem übergeordneten Komplex zusammenfinden.

---

<sup>1</sup>Und auch nicht vollständig: Viele Proteine, wie Antikörper und die Proteine der Lichtsammelkomplexe von Pflanzen, sind einfach unterschlagen worden.

Der Sinn einer Quartärstruktur besteht oftmals in einer besonderen Fähigkeit zur Regulation durch Effektoren. Der Einfluss von Effektoren zwingt solchen Proteinkomplexen oftmals eine andere Konformation auf. Eine Konformation ist *eine* Struktur *eines* Proteins, womit deutlich wird, dass ein Protein in mehreren Strukturvariationen existieren kann.

Verschiedene Konformationen sind oftmals die strukturelle Manifestation eines regulativen Eingriffs durch einen Effektor, beispielsweise für die Regulation einer Stoffwechsellkaskade (z. B. die Pyruvatkinase in der Glykolyse) oder infolge der Sauerstoffbindung bei Hämoglobinen und Hämocyaninen (siehe folgende Abschnitte). Auch die Assoziation zu einem Komplex, die Bildung einer Quartärstruktur, kann einen Konformationswechsel in den Untereinheiten induzieren. Dies wird u. a. beim  $\alpha$ -Hämolsin von *Staphylococcus aureus*, einem porenbildenden Toxin, in dieser Arbeit besprochen.

Im Zentrum dieser Arbeit stehen Fragestellungen rund um die Konformationswechsel von Proteinkomplexen. Hierzu wurden vor allem die Methoden der Kleinwinkelstreuung und der rechnergestützten *in silico*-Modellbildung angewendet. Mit Hilfe der Kleinwinkelstreuung können Proteine in Lösung untersucht werden, was es erlaubt unter Bedingungen zu arbeiten, die denen eines zellulären Milieus zumindest nahekommen. *In silico*-Methodik, wie Homologiemodellierung oder das Modellieren von Konformationen auf Basis der experimentellen Daten aus der Kleinwinkelstreuung, erlaubt es viel schneller Fragen zu klären als dies experimentell möglich ist. Andererseits wird im Laufe dieser Arbeit immer wieder deutlich werden, wie wichtig der experimentelle Abgleich der „Antworten“ aus dem Computer ist.

In den folgenden Abschnitten der Einleitung werden die wesentlichen Fragen dieser Arbeit eingeführt: Was passiert mit einem Hämocyanin bei der Sauerstoffbindung? Was passiert strukturell wenn ein Hämocyanin zur Tyrosinase wird? Was charakterisiert die Konformationen des  $\alpha$ -Hämolsins von *Staphylococcus aureus*?

## 1.2 Hämocyanine

Die Fähigkeit von Kupferproteinen Sauerstoff ( $O_2$ ) im „gemäßigten“ chemischen Umfeld zu binden, zieht die Aufmerksamkeit vieler Gruppen der chemischen Forschung auf sich. Neben den katalytischen Fähigkeiten (s. u.) interessiert aber auch die Regulation dieser Proteine und wie diese Regulation auf einer strukturellen Ebene realisiert ist.

Eine erste Einteilung von Kupferproteinen unterschied drei grundlegende Typen. Inzwischen werden, mit Hilfe moderner spektroskopischer und kristallographischer Techniken, sieben Typen klassifiziert (Koval *et al.*, 2006). Zu den Proteinen mit einem „Typ-3 Kupferzentrum“ zählen Hämocyanine, Tyrosinasen und Catecholoxidasen. Das aktive Zentrum enthält zwei Kupferionen, die durch je drei Stickstoffdonoren von angrenzenden Histidinresten koordiniert werden (Gaykema *et al.*, 1986; Hazes *et al.*, 1993; Magnus *et al.*, 1994; Solomon *et al.*, 1993; Solomon *et al.*, 1996; Klabunde *et al.*, 1998; Matoba *et al.*, 2006). Die elektronische Konfiguration der Sauerstoffbindung an die Kupferatome ist bei den Typ-3 Kupferzentren  $\mu\text{-}\eta^2\text{:}\eta^2$ . Dies erklärt die geometrische Konfiguration der Sauerstoffbindung, bei der jedes Atom des  $O_2$ -Moleküls von jedem der beiden Kupferionen koordiniert wird. Man bezeichnet diese Konfiguration deshalb auch als *side-on* Konfiguration.

Die wesentliche Eigenschaft und Aufgabe von Hämocyaninen ist es an diesem aktiven Zentrum  $O_2$  reversibel zu binden. Hämocyanin erfüllt so in vielen Arthropoden und Mollusken die Aufgabe des Sauerstofftransportes. Im Gegensatz dazu verwenden Tyrosinasen und Catecholoxidasen den gebundenen Sauerstoff zur Oxidation von phenolischen Substraten zu Catecholen (bei Tyrosinasen) und anschließend zu *o*-Chinonen (wozu Tyrosinasen und Catecholoxidasen fähig sind; vgl. Abbildung 1.1). Diese *o*-Chinone können zur Produktion des Pigmentes Melanin spontan polymerisieren (eine detaillierte Übersicht bietet Riley, 1997).

Tyrosinasen und Catecholoxidasen erfüllen eine wesentliche Aufgabe in allen Bereichen des Lebens: Das von ihnen indirekt produzierte Melanin wirkt beispielsweise bakteriostatisch und spielt

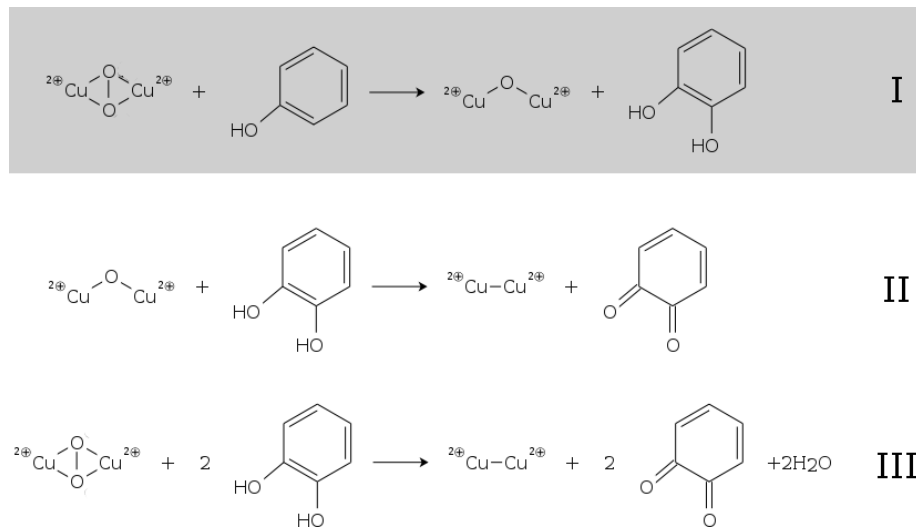


Abbildung 1.1: *Reaktionen der Tyrosinasen und Catecholoxidasen*. Tyrosinase können die Reaktionen I-III katalysieren, Catecholoxidasen lediglich Reaktionen II+III. Die Abbildung wurde mit GChemPaint generiert (Bréfort, 2005) und ist an die entsprechende Abbildung in der Dissertation von T. Schweikardt (Schweikardt, 2008) angelehnt. Nachdruck mit Genehmigung von T. Schweikardt.

eine bedeutende Rolle in der Immunabwehr der Invertebraten (Johansson & Söderhäll, 1996). Darüber hinaus ist es wesentlich für die Färbung von Haut und Federn und damit zur Tarnung und zum Schutz vor ultravioletter Strahlung verantwortlich (Riley, 1997). Schließlich spielt es über die Quervernetzung der Chinone eine bedeutende Rolle bei der Sklerotisierung des Exoskeletts der Arthropoden. Phenoloxidasen im Allgemeinen spielen auch eine direkte Rolle im Immunsystem vieler Organismen, die ganz unterschiedlich wahrgenommen werden kann (Lai-Fook, 1970; Decker *et al.*, 2001; Nagai *et al.*, 2001).

### 1.2.1 Arthropodenhämocyanine

Die Superfamilie der Arthropodenhämocyanine umfasst nicht nur die Hämocyanine im engeren Sinne (als Sauerstofftransporter), sondern daneben auch die Pseudo-Hämocyanine der Crustaceen (Markl & Decker, 1992; Terwilliger *et al.*, 2005), die Hexamerine der Insekten (siehe Seite 12 in dieser Einleitung) und die Phenoloxidasen (Burmester, 2001).

#### 1.2.1.1 Struktureller Aufbau

Strukturell basieren alle diese Proteine auf ähnlichen, nierenförmigen Untereinheiten von etwa 72–75 kDa (Volbeda & Hol, 1988; Decker & Terwilliger, 2000; siehe rechtsstehende Abbildung). Diese Untereinheit besitzt drei Domänen: Die N-terminale Domäne I umfasst 6–7  $\alpha$ -Helices (van Schaick *et al.*, 1982; Gaykema *et al.*, 1984; Volbeda & Hol, 1989). Domäne II enthält das Typ-3-Kupferzentrum eingebettet in  $\alpha$ -Helices (Markl & Decker, 1992; Gaykema *et al.*, 1984). Domäne III schließlich ist wesentlich von einem  $\beta$ -Barrel<sup>2</sup> aufgebaut. Bei den Arthropoden werden diese Untereinheiten zu Hexameren assembliert. Höhere Oligomerisierungsstufen, in Form von Kombinationen von Hexameren, kommen häufig vor und sind artabhängig (vgl. Abbildung 1.3 auf der nächsten Seite). Eine Erläuterung zur Struktur bestimmter Arthropodenhämocyanine folgt in den entsprechenden Abschnitten (Seite 6 ff.).

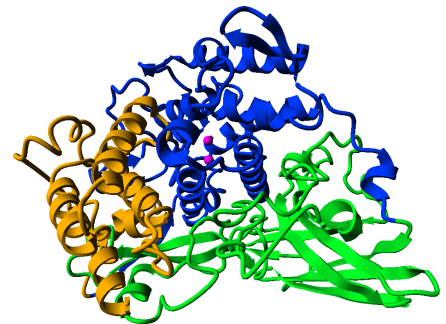


Abbildung 1.2: *Domänenstruktur der Arthropodenhämocyaninuntereinheiten nach Hazes et al., 1993; PDB-code: 1LLA*. Die Kupferionen der Domäne II sind violett eingefärbt.

<sup>2</sup>Englische Fachausdrücke der Strukturbiologie werden direkt übernommen.

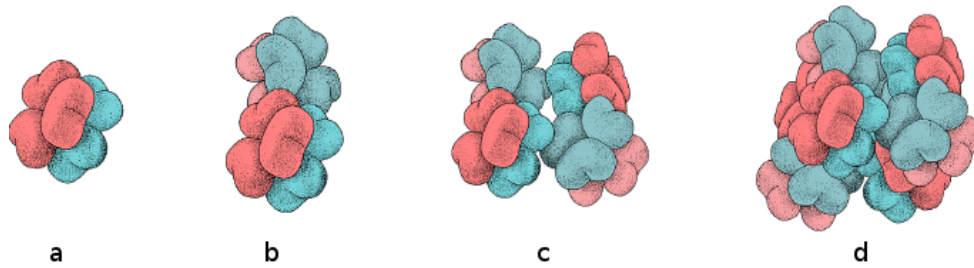


Abbildung 1.3: *Oligomerisierungsstufen von Arthropodenhämocyaninen*. a) Hexamere, wie sie bei vielen Arthropoden vorkommen können. b) 2 x 6mere, die typisch für Crustaceen sind, aber auch bei einigen Cheliceraten vorkommen (Markl, 1980; Stöcker *et al.*, 1988). c) 4 x 6mere kommen bei den Thalassinidea (Roxby *et al.*, 1974; Miller & Holde, 1974; Miller *et al.*, 1977; Taylor *et al.*, 2000) und bei Cheliceraten vor (Lamy *et al.*, 1981; Markl *et al.*, 1981). d) 48mere schließlich kommen bei Pfeilschwanzkrebse (Limulidae) vor (Lamy *et al.*, 1982). Myriapoda stellen eine weniger beachtete Ausnahme dar: Bei diesen Tieren gibt es auch 36mere (Magnum, 1985; Jaenicke *et al.*, 1999). Bildquelle: K. Reh binder für Savel-Niemann *et al.*, 1988.

### 1.2.1.2 Kooperativität & Allosterie

Die spezifische Assoziation der Untereinheiten wird dadurch möglich, dass diese – abhängig von der Art – unterschiedlich aufgebaut sind (siehe z. B. van Holde & Miller, 1995; Markl & Decker, 1992). Dadurch können Untereinheiten spezifische Kontakte zu anderen formen. Die Zusammenlagerung verschiedener sauerstoffbindender Untereinheiten befähigt die Hämocyanine nun dazu, die Bindungseigenschaften für Sauerstoff abhängig vom Beladungsgrad mit Sauerstoff zu modulieren (siehe z. B. Decker *et al.*, 1979; Decker *et al.*, 1989; Savel-Niemann *et al.*, 1988). Die Beeinflussung der Bindeaffinität zu einem Liganden durch ebendiesen Liganden bezeichnet man als Kooperativität. Die theoretische Beschreibung der Kooperativität erfordert eine von zwei Möglichkeiten: Ein konzertiertes, beziehungsweise MWC-artiges, Modell (nach Monod, Wyman & Changeux; Monod *et al.*, 1965) oder ein sequenzielles bzw. KNF-Modell (nach Koshland, Nemethy & Filmer; Koshland *et al.*, 1966). Beim KNF-Modell wird davon ausgegangen, dass die Bindung eines Liganden an eine Unterheit einen Konformationswechsel dieser Untereinheit induziert, wodurch benachbarte Untereinheiten die Affinität zu diesem Liganden ändern. Für große Proteinkomplexe, wie Hämocyanine, ist solch ein Modell aber nicht einzusetzen: Der Ligand kann zunächst an jede der Untereinheiten binden. Bei einer sequentiellen Bindung weiterer Liganden müssten so viele Konformationen berücksichtigt werden, dass diese physikalisch nicht mehr nachgewiesen werden können. Solch ein Modell besitzt also zu viele Freiheitsgrade für eine aussagekräftige Analyse. Im Gegensatz dazu geht das MWC-Modell davon aus, dass ein solcher Proteinkomplex grundsätzlich in zwei Konformationen vorkommen kann: In einer hoch- und einer niederaffinen Konformation. Ist kein Ligand vorhanden, so liegt mit großer Wahrscheinlichkeit nur die niederaffine Konformation vor. Durch die Bindung des Liganden verschiebt sich das Gleichgewicht von der niederaffinen Form stärker zur hochaffinen Form. Das Bindungsverhalten des hexameren Hämocyanins von *Panulirus interruptus* konnte beispielsweise mit diesem Modell gut beschrieben werden (Connelly *et al.*, 1989; vgl. Abschnitt 1.2.1.5 für eine weitere Beschreibung hexamerer Hämocyanine). Insbesondere bei noch komplexeren Proteinverbänden hat sich dieses Modell jedoch als unzureichend herausgestellt. Das „nesting-Modell“, angeregt durch Wyman (Wyman, 1970; Wyman, 1984), berücksichtigt die Hierarchieebenen der Quartärstruktur, denen unterschiedliche allosterische (s. u.) Gleichgewichte zugeordnet werden. So kann das Sauerstoffbindeverhalten der dodekameren Crustaceenhämocyanine mit der Verknüpfung zweier Hexamere und von 24meren Cheliceratenhämocyanine unter Wechselwirkung zweier Dodekamere beschrieben werden (Robert *et al.*, 1987; Decker *et al.*, 1986; Decker & Sterner, 1990).

Die Sauerstoffaffinität der Arthropodenhämocyanine kann aber auch noch durch andere physiologische Faktoren beeinflusst werden: Protonen (über Modulation des pH-Wertes, ein Effekt, der als Bohr-Effekt bezeichnet wird; (Decker *et al.*, 1989; Decker & Sterner, 1990; Hellmann, 2004)), Lak-

tat, Urat und auch Koffein (als nicht-physiologisches Uratanalogon; Truchot, 1980; Bridges *et al.*, 1984; Graham, 1985; Nies *et al.*, 1992; Jaenicke, 1998; Hellmann *et al.*, 2001; Hartmann *et al.*, 2001; Hellmann *et al.*, 2004). Sogar Wasser wird als Modulator der Sauerstoffbindeeigenschaften von Hämocyaninen diskutiert (Hellmann *et al.*, 2003). Diese Effekte – ein Ligand bindet und beeinflusst die Affinität zu einem anderen Liganden an einer anderen Bindestelle – subsummiert man unter dem Begriff der Allosterie. Konformationsübergänge sind die Grundvoraussetzung für Kooperativität. Als allosterische Einheit bezeichnet man die Untereinheit(-engruppe), die den Konformationsübergang (gemeinsam) durchführt. Im Falle der Arthropodenhämocyanine (s. u.) können dies Hexamere oder Vielfache davon sein.

### 1.2.1.3 Das Hämocyanin von *Eurypelma californicum*



Abbildung 1.4: *Eurypelma californicum*, die verbreitetste Tarantel Nordamerikas. Quelle: Photographie aus Institutsbeständen, Hintergrund entfernt.

*Eurypelma californicum* ist eine Vogelspinne (Theraphosidae; engl.: Tarantula), die vor allem in den US-Bundesstaaten Californien, Arizona, New Mexico und im nördlichen Mexico verbreitet ist. Die Spinnen leben in Erdhöhlen dieser ariden Gebiete, sind sehr träge, ortsgebunden und erreichen eine Körpergröße von etwa 8 cm (Conze, 2005). Es sind Jagdspinnen, die ihrer Beute nachts vor ihren Höhlen auflauern. Ihr Gift ist für Menschen ungefährlich. Die Tiere atmen über sogenannte Buchlungen (Paul *et al.*, 1989; Paul & Fincke, 1989). Von hier gelangt Sauerstoff direkt in die Hämolymphe der Tiere und kann vom Hämocyanin gebunden werden. Neuere Ergebnisse unserer Gruppe zeigen, dass die Tiere in der Lage sind, je nach den physiologischen Anforderungen, den Hämolympfluss durch Schließung des Pedunculums zwischen Ophistoma und Prosoma aktiv zu steuern (Pohlmann *et al.*, 2007). Der Sauerstoffbedarf im Gewebe der Spinnen wird in erster Linie durch das in der Hämolymphe gelöste Hämocyanin gedeckt: Während die Untereinheiten des Hämocyanins von *E. californicum* sich nicht kooperativ verhalten (Decker *et al.*, 1979; Markl *et al.*, 1981), weist das vollständige 24mer eine sehr

starke Kooperativität auf (mit einem Hill-Koeffizienten von bis zu 7; Loewe, 1978). Zusammen mit der Empfindlichkeit auf einige allosterische Effektoren, ermöglicht dies eine fein abgestimmte Sauerstoffversorgung der Gewebe.

Das Hämocyanin der Arachniden ist ein 24meres Hämocyanin, was auch auf das Hämocyanin von *E. californicum* zutrifft. Wenngleich eine 3D-Rekonstruktion des Hämocyanins aus kryoelektronenmikroskopischen Daten vorliegt (Stohr, 2007), so ist diese nicht publiziert und alle bisher zugängliche strukturelle Information über das Hämocyanin der Vogelspinne beruht auf weniger hochauflösenden Daten. Die Kristallstruktur eines *in vitro* generierten artifiziellen hexameren Hämocyanins von *Limulus polyphemus* (Hazes *et al.*, 1993), zusammen mit den Sequenzdaten des Hämocyanins der Vogelspinne, erlaubte es T. Schweikardt ein Modell des Hämocyanins von *E. californicum* zu errechnen (Voit *et al.*, 2000; Schweikardt, 2008). Daten aus der Röntgenkleinwinkelstreuung geben hierbei die Konformation – ebenfalls auf Basis der Röntgenstruktur von *L. polyphemus*-Hämocyanin – vor (Hartmann & Decker, 2002). Im Rahmen dieser Kleinwinkelstreuarbeiten konnte auch die Vermutung zweier grundlegender Konformationen für den oxygenierten und deoxygenierten Zustand bestätigt werden (Decker *et al.*, 1996; Hartmann & Decker, 2002; Hartmann & Decker, 2004; Hartmann *et al.*, 2004).

Zwei der Hexamere setzen sich aus den Untereinheiten a,b, d–g zusammen, während sich die anderen beiden Hexamere aus den Untereinheiten a, c–g zusammensetzen (Markl *et al.*, 1980; Voit *et al.*, 2000). Es gibt folglich insgesamt sieben Untereinheiten. Die Verbindung zwischen den Dodekameren wird durch das Paar b-c vermittelt (vgl. Abbildung 1.5). Dies wird auch für die analogen Untereinheiten des Pfeilschwanzkrebse *Limulus polyphemus* in einer aktuellen elektronenmikroskopischen Rekonstruktion bestätigt (Martin *et al.*, 2007).

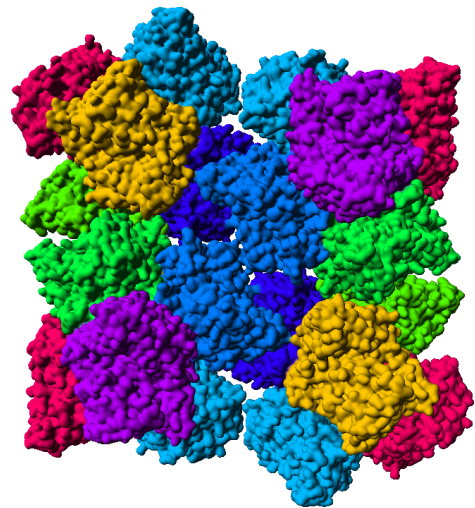


Abbildung 1.5: Das Hämocyanin von *Eurypelma californicum* in einer Oberflächendarstellung. Identische Untereinheiten sind gleichartig eingefärbt; lediglich die beiden Grüntöne stehen für denselben Untereinheitentyp (farblich abgesetzt für bessere Übersichtlichkeit). Die Abbildung zeigt die oxygenierte Konformation nach Hartmann & Decker, 2002.

### *Ziele dieser Arbeit*

Mit der Kratky-Kamera des Institutes – einer Apparatur zur Röntgenkleinwinkelstreuung, vgl. Abschnitt 2.3 & 3.6 – waren vor Beginn dieser Doktorarbeit nur wenige Messungen vorgenommen worden (Hönig, 2004). Ziel war es zunächst ältere Daten zum Konformationswechsel bei der Oxygenierung zu reproduzieren (Decker *et al.*, 1996) und diese in Verbindung zu Daten des Hämocyanins des Kaiserskorpions (siehe folgender Abschnitt) zu bringen. Dieses wurde zusammen mit B. Pairet durchgeführt (siehe die Diplomarbeit von B. Pairet (Pairet, 2007)).

Bekannt war zudem, dass das Hämocyanin in Gegenwart von SDS Phenoloxidaseaktivität entwickeln kann (Decker & Rimke, 1998; Decker *et al.*, 2001; Jaenicke & Decker, 2008). Aufgrund von *Molecular Modelling* auf Basis von kristallographischen Daten anderer Hämocyanine wird vermutet, dass ein konserviertes Phenylalanin der Domäne I den Zugang phenolischer Substrate zum aktiven Zentrum unterbindet (Decker *et al.*, 2007). Nicht bekannt war, ob mit der Bindung von SDS ein Konformationswechsel auf Ebene des 24meres stattfindet und welcher Art dieser sein kann. Aufgabe in dieser Arbeit war es dieses mittels Röntgenkleinwinkelstreuung zu untersuchen. Dieser Teil der Arbeit ist bereits veröffentlicht: Baird *et al.*, 2007; vgl. auch das Kapitel zu den Originalarbeiten C.1.

#### 1.2.1.4 Das Hämocyanin von *Pandinus imperator*



Abbildung 1.6: Ein Exemplar von *Pandinus imperator*. (Quelle: Photo aus Institutsbeständen, Hintergrund entfernt.)

Der Kaiserskorpion *Pandinus imperator* (linke Abbildung) kommt in Zentralafrika vor und ernährt sich vorwiegend von Schaben, Grillen und anderen Insekten. Gelegentlich stehen auch kleinere Säuger auf dem Speiseplan. Muttertiere brüten Eier im Körper aus und sind lebendgebärend. *P. imperator* ist leicht zu halten und stellt daher ein bevorzugtes Untersuchungsobjekt in der Zoologie dar (Casper, 1985).

Das Hämocyanin des Kaiserskorpions ist ebenfalls ein 24mer und die Zahl der Untereinheiten des Hämocyanins beträgt wie bei *Eurypelma californicum* sieben, möglicherweise acht (Heimerl & Decker, 1989; Decker, 1990). Elektronenmikroskopische Aufnahmen belegen die große strukturelle Ähnlichkeit zum Hämocyanin von *E. californicum* (Heimerl & Decker, 1989). Wenngleich die Struktur zum Beginn meiner Arbeit weniger gut untersucht war, so lagen jedoch eine *single particle*-Rekonstruktion für das Hämocyanin des Skorpions *Androctonus australis* (Boisset *et al.*, 1990) und strukturelle Studien für das Hämocyanin der verwandten Art *Buthus indicus* (Ali *et al.*, 1995; Ali *et al.*, 2000) vor. Überdies erschien während der Zeit meiner Dissertation eine Studie auf der Basis von Kleinwinkelstreudaten zum Hämocyanin von *Buthus indicus* (Ali *et al.*, 2007).

*Ziele dieser Arbeit*

### *Ziele dieser Arbeit*

Die Ziele hinsichtlich des Hämocyanins von *Pandinus imperator* sind nahezu deckungsgleich mit denen zum Hämocyanin von *E. californicum* formulierten (vgl. den vorigen Abschnitt): Durch die Arbeiten unseres Institutes war bekannt, dass auch die Tyrosinaseaktivität dieses Protein durch SDS aktiviert werden kann (Nillius, 2007; Decker *et al.*, 2007; Nillius *et al.*, 2008). In Zusammenarbeit mit Wah Chiu & Yao Cong, Houston, wurden zwei noch nicht veröffentlichte *single particle*-Rekonstruktionen vom oxygenierten Protein, mit und ohne SDS, angefertigt (unveröffentlicht: Cong *et al.*, 2008). Der Frage, inwieweit die Strukturen der Hämocyanine von *E. californicum* und *P. imperator* bei der Aktivierung von SDS in Lösung korrelieren, wurde in dieser Arbeit mittels Röntgenkleinwinkelstreuung nachgegangen. Überdies wurde zusammen in der Diplomarbeit von B. Pairet (Pairet, 2007) der Konformationswechsel vom oxygenierten zum deoxygenierten Zustand mittels Röntgenkleinwinkelstreuung charakterisiert.

### 1.2.1.5 Die Hämocyanine von *Palinurus elephas* und *Palinurus argus*

Die Europäische Languste *Palinurus elephas* kommt vor allem im Mittelmeer und in einem Bogen von Nordwestafrika bis in die vom Golfstrom erwärmten Gewässer Südnorwegens vor (siehe Abbildung 1.7). Die Tiere werden etwa 50 cm groß und tragen charakteristische weiße Flecken auf dem Hinterleib. Üblicherweise leben sie in Tiefen zwischen 5 und 150 Metern, am häufigsten sind sie zwischen 10 und 70 Meter Meerestiefe anzutreffen (Hunter, 1999). Die Karibische Languste *Palinurus argus* (auch *Panulirus argus* in einigen Quellen) ist etwas kleiner als ihr europäischer Vetter und kommt im Westatlantik zwischen Nordostbrasilien und dem US-Bundestaat North Carolina vor, wobei die gesamte Karibik ebenfalls zu ihrem Lebensraum gehört (siehe Abbildung 1.7). Es ist allerdings unklar, inwieweit sich die Art in den sehr warmen Gebieten der Karibik fortpflanzen kann (Gregory *et al.*, 1982).

Die Hämocyanine der beiden Langusten gehören zu den hexameren Hämocyaninen der Crustaceen. Hexamere Hämocyanine (1 x 6mere) bilden den Grundbaustein der Arthropodenhämocyanine (vgl. Abschnitt 1.2.1.1) und das Verständnis der Umsetzung der Konformationsänderungen in Folge kooperativer Interaktionen (vgl. Abschnitt 1.2.1.2) muss hier, beim grundlegenden Baustein für höhere Aggregationsstufen, ansetzen. Bisher existieren von hexameren Strukturen Röntgenstrukturen für die Hämocyanine von *Panulirus interruptus* (Gaykema *et al.*, 1984; Gaykema *et al.*, 1986; Volbeda & Hol, 1989), des artifiziellen *in vitro*-Hexamers von *Limulus polyphemus* (Hazes *et al.*, 1993; Magnus *et al.*, 1994) und eine *single particle*-Rekonstruktion des Hämocyanins von *P. elephas* (Meissner *et al.*, 2003).

Es ist bekannt, dass Urat und Koffein als Ureatanlagon die Sauerstoffbindung eng verwandter Hämocyanine modulieren können. Ihre Bindeeigenschaften für 12mere Hämocyanine wurden charakterisiert (Nies *et al.*, 1992; Jaenicke, 1998; Menze *et al.*, 2000; Menze *et al.*, 2001; Hellmann *et al.*, 2004). Die Bindung von Urat senkt die Sauerstoffaffinität der Hämocyanine, was zu verstehen ist, da die Tiere Urat im Zuge von metabolischen Stress (O<sub>2</sub>-Mangel) in die Hämolymphe ausscheiden (Dykens, 1991). Für hexamere Hämocyanine wurden Bindungsstudien jedoch erstmals in der Diplomarbeit von A. Rabenhorst in unserer Gruppe durchgeführt (Rabenhorst, 2007).

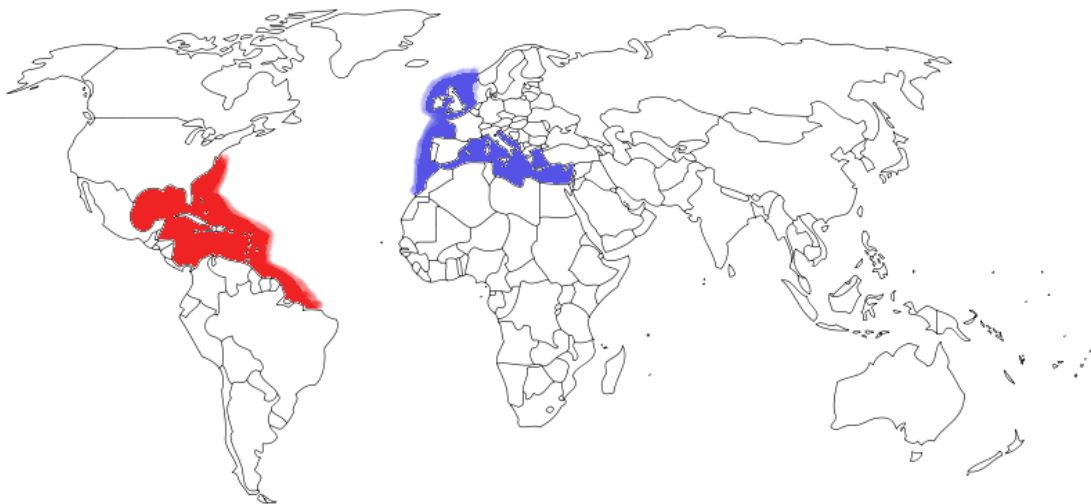


Abbildung 1.7: Verbreitungsgebiete der europäischen Languste (*Palinurus elephas*, blau) und der karibischen Languste (*Palinurus argus*, rot) nach Informationen des GBIF-Projektes (Edwards *et al.*, 2000).

#### Ziele dieser Arbeit

Zusammen mit A. Rabenhorst wurden zusätzlich zu den Bindungsexperimenten mittels Isothermer Titrationskalorimetrie (ITC) Röntgenkleinwinkelstreuemessungen durchgeführt (Rabenhorst, 2007). Ziel dieser Messungen war es die Konformationsänderung nachzuvollziehen, welche durch

Bindung von Urat bzw. Koffein induziert werden muss. Eine vergleichbare Charakterisierung mittels Röntgenkleinwinkelstreuung wurde in unserer Gruppe bereits für das 2 x 6mere Hämocyanin von *Homarus americanus* unter Laktateinfluss durchgeführt (Hartmann *et al.*, 2001). Im Zuge dieser Arbeit konnte die Auswertung der Daten detaillierter erfolgen, als dies damals möglich war, insbesondere auch durch eigene, im Zuge dieser Dissertation erstellte, Software (siehe Kapitel 4.7).

### 1.2.1.6 Das Hämocyanin von *Porcellio scaber*

Die Gruppe von M. Zimmer (Arbeitsgruppe Evolutionsökologie von Asseln an der Universität Kiel) befasst sich mit der Frage wie bestimmte Isopoda ihr Immunsystem organisieren. Eines der Hauptuntersuchungsobjekte hierzu ist die wohlbekannte Kellerassel (*Porcellio scaber*; siehe rechtsstehende Abbildung). Diese Assel kommt vor allem in Zentral- und Mitteleuropa vor. Es gibt insgesamt sieben Unterarten, die natürlicherweise in den gemäßigten Breiten vorkommen, aber – in der Nähe menschlicher Behausungen – auch auf anderen Teilen der Welt anzutreffen sind (Edwards *et al.*, 2000). Die Tiere ernähren sich vor allem von Detritus und sind wichtige Elemente des Ökosystems bei der Bereitstellung von Spurenelementen aus pflanzlichem Detritus (Kautz & Topp, 2000).



Abbildung 1.8: *Porcellio scaber*, die deutsche Kellerassel in Nahaufnahme. Bildquelle: [www.gardensafari.net](http://www.gardensafari.net); Nachdruck für diese Arbeit mit Genehmigung von Hania Berdys.

#### *Ziele dieser Arbeit*

Im Zuge der Diplomarbeit von S. May (AG Zimmer, Kiel) entstand eine Kooperation mit der Arbeitsgruppe der Molekularen Biophysik mit dem Ziel eine nähere Charakterisierung des Hämocyanins von *P. scaber* durchzuführen. Bei der Trennung mit der analytischen Gelfiltration unter Einbeziehung der statischen Lichtstreuung (vgl. 3.4.1) fiel auf, dass die Hämocyaninprobe aus zwei verschiedenen großen Komponenten, Peak 1 & 2 genannt, besteht. Mittels statischer Lichtstreuung und analytischer Ultrazentrifugation (vgl. Abschnitt 3.5 auf Seite 32 im Methodenkapitel) konnte gezeigt werden, dass diese beiden „Peaks“ in ihrer Größe und ihrem Sedimentationsverhalten hexameren bzw. dodekameren Arthropodenhämocyaninen entsprechen. Es ergaben sich folgende Fragestellungen, welche in dieser Arbeit mit Hilfe der Elektronenmikroskopie beantwortet wurden: Gibt es beide Formen in beiden Proben? Stehen diese in einem Gleichgewicht? Wie ist das Hämocyanin aufgebaut? Wie stehen die beiden Hexamere in der dodekameren Form räumlich zueinander?

### 1.2.1.7 *Upogebia pusilla*-Hämocyanin

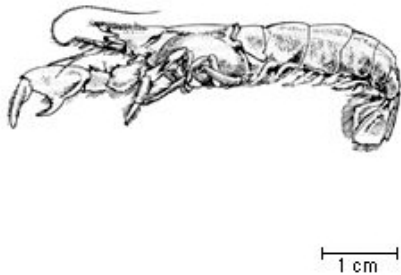


Abbildung 1.9: *Upogebia pusilla*. Zeichnung von *U. pusilla* aus Cottiglia, 1983.

Während dieser Doktorarbeit arbeitete M. Paoli aus Padua in der Gruppe der Molekularen Biophysik als Gaststudent. Sein Untersuchungsgegenstand war das Hämocyanin der mediterranen Schlammmakrele *Upogebia pusilla*. Ziel war es den Unterschied zwischen den Hämocyaninen der Crustaceen und Cheliceraten näher zu charakterisieren: Für das Hämocyanin der Thalassinidea, einer Familie der Malacostraca oder „höheren Krebse“, zu denen *U. pusilla* gehört, wurde eine ähnliche Größe wie für das Hämocyanin der Crustaceen angenommen – jedoch mit einer anderen strukturellen Organisation. Diese Arbeitshypothese bestand mit Beginn der Arbeiten und wird in Paoli *et al.*, 2007, als bestätigt angesehen. Hämocyanine in der Gruppe der Crustaceen sind in der Regel 2 x 6mere (gelegentlich 1 x 6, Markl, 1986). Makrelen der Thalassinidea jedoch besitzen ein 4 x 6meres Hämocyanin (Roxby *et al.*, 1974; Miller & Holde, 1974; Miller *et al.*, 1977; Taylor *et al.*, 2000). Allerdings sind für *U. pugettensis* (Miller *et al.*, 1977) und *U. deltaura* (Taylor *et al.*, 2000) auch 2 x 6mere Zwischenformen beschrieben.

4 x 6mere Hämocyanine gibt es ebenfalls bei den Cheliceraten (siehe oben, Kapitel 1.2.1.3 und 1.6). Zum Zeitpunkt der Arbeit gab es zwei prinzipielle Modelle für die strukturelle Organisation der ikosatetrameren Hämocyanine: Bei Cheliceraten das Hämocyanins des Skorpions *Androctonus australis* (Lamy *et al.*, 1981; Sizaret *et al.*, 1982) beziehungsweise das für *E. californicum* (Voit *et al.*, 2000; Erker *et al.*, 2006). Das zweite Modell wurde für die Hämocyanine der Thalassinidea am Beispiel von *Callinassa californiensis* vorgeschlagen (Cavellec *et al.*, 1990). Während im ersten Model die Hämocyanine als quasi-planar aufgefasst werden, mit 2 x 6meren antiparallel zueinander (vgl. Abbildung 1.5), kommen Paoli *et al.*, 2007, zu dem Schluss, dass die Hexamere der Thalassinidea quasi-tetraedrisch assemblieren.

#### *Ziele dieser Arbeit*

Die Messungen von M. Paoli in Mainz dienten der Erhärtung u. a. der eben vorgestellten Hypothesen. Zusätzlich wurde von ihm auch das Sauerstoffbindeverhalten dieses Hämocyanins charakterisiert. Mein Beitrag waren elektronenmikroskopische Qualitätskontrollen und Röntgenkleinwinkelstreuung. Letzteres mit dem Fernziel den Konformationswechsel des *U. pusilla*-Hämocyanins bei der Oxy-/Deoxygenierung nachvollziehen zu können.

### 1.2.2 *Nautilus pompilius*-Hämocyanin

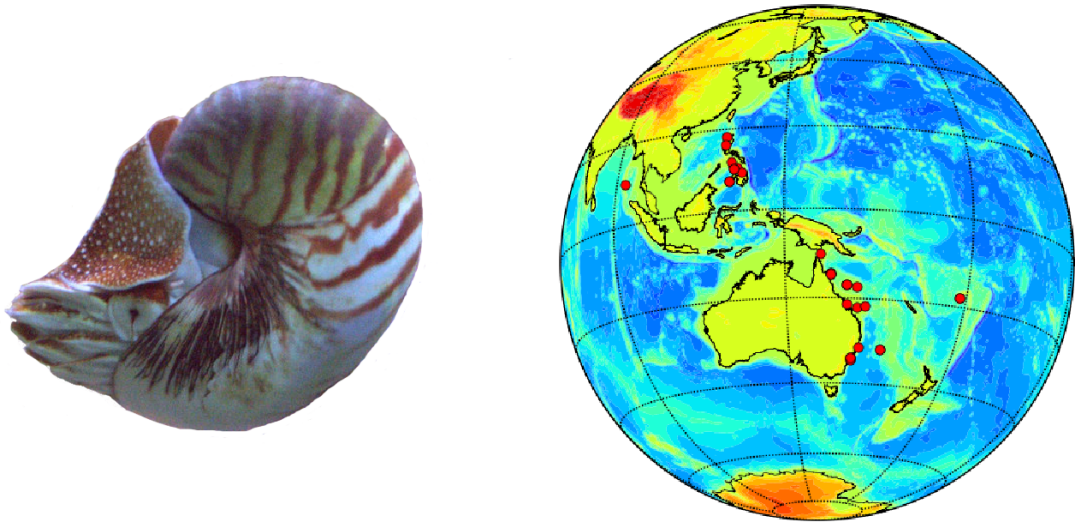


Abbildung 1.10: Links: *Nautilus pompilius* in einem Aquarium. (Quelle: Eigene Aquariumaufnahme, Hintergrund entfernt) Rechts: Das Verbreitungsgebiet der Art. Aufgezeigt an von GBIF registrierten Fundstellen (rote Punkte; Edwards *et al.*, 2000). Abbildung erstellt mit einem eigenen Skript auf Basis von matplotlib (Hunter, 2007).

*Nautilus pompilius* (siehe Abbildung 1.10) ist ein „lebendes Fossil“. Die Art gehört zur Familie der „Perlboote“, Nautiloidea, und ist eine von insgesamt acht rezenten Arten. Die Gruppe gehört zu den Kopffüßern, Cephalopoda, und ist ähnlich den ausgestorbenen Ammoniten. *N. pompilius* ist die am weitesten verbreitete Art der Cephalopoda und kommt von den Andamanen bis zu den Fidschi-Inseln und Südjapan bis zum Großen Barriere-Riff an den Steilhängen der Korallenriffe vor. Tagsüber tauchen die Tiere 300 – 400 Meter tief ab, des Nachts dringen sie bis in Tiefen von 70 Meter vor (Wells *et al.*, 1992). Die Schale von *Nautilus* ist gekammert und mit Wachstum der Tiere scheiden sich immer weitere Kammern ab, wobei die Tiere jedoch nur in der jüngsten Kammer leben.

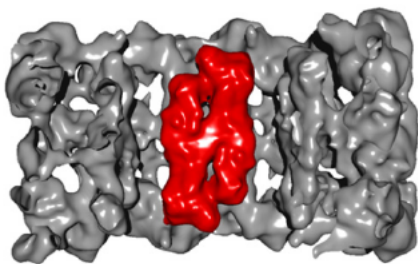


Abbildung 1.11: Darstellung des *Nautilus* Hämocyanins aus der Elektronenmikroskopischen Struktur. Das Paar der g-Untereinheiten ist rot eingezeichnet. Zu sehen ist ein Querschnitt durch die Struktur mit der inneren Kragendomäne (Erläuterung im Text).

Die Hämocyanine von Mollusken und den zuvor angesprochenen Arthropoden unterscheiden sich auf allen strukturellen Ebenen und werden als distinkte Proteinfamilien aufgefasst (Markl & Decker, 1992; van Holde & Miller, 1995; Harris & Markl, 1999; Decker *et al.*, 2007). Die meisten Molluskenhämocyanine (insbesondere von Gastropoden, (Poly-)Placophora, Protobranchia) basieren auf einem 400 kDa großen Proteinmotiv, welches sich in acht paraloge funktionale Einheiten (FEs) unterteilen lässt und folgende Grundstruktur aufweist: a → b → c → d → e → f → g → h. Jede dieser FEs besitzt ein Molekulargewicht von 45 – 50 kDa (bis auf FE h mit etwa 60 kDa) und trägt je ein binukleares Kupferzentrum, wie wir es bereits von den Arthropoden der vorherigen Abschnitte kennen. In Cephalopoden fehlt die C-terminale FE h, übrig bleibt eine Untereinheit von etwa 350 kDa. Dies trifft beispielsweise auf *Octopus* und *Nautilus* zu. Einige andere Cephalopoden, wie *Sepia* und *Loligo*, weisen eine Duplikation einer zentralen FE auf, was zu einer

abweichend organisierten 400 kDa-Untereinheit führt. In der grundlegenden Quartärstruktur der Molluskenhämocyanine formen zehn Untereinheiten einen Hohlzylinder, der sich in eine externe Hülle und eine interne „Kragendomäne“ gliedert. In Gastropoden und Protobranchia sind zwei solcher Dekamere aneinandergelagert, so dass sich Didekamere formen, die manchmal sogar tubu-

läre Multi-Dekamere formen können (Markl *et al.*, 2001). In (Poly-)Placophora und Cephalopoden hingegen kommen ausschließlich einfache Dekamere vor (van Holde & Miller, 1995).

#### *Ziele dieser Arbeit*

In der Gruppe von U. Meissner (Institut für Zoologie der Universität Mainz) wurden zu Beginn dieser Arbeit an *N. pompilius*-Hämocyanin elektronenmikroskopische (EM) Messungen und 3D-Rekonstruktionen der Molekülstruktur basierend auf diesen Daten durchgeführt (resultierend in zwei Veröffentlichungen: Meissner *et al.*, 2006; Gatsogiannis *et al.*, 2007; siehe auch Abbildung 1.11). Daran anknüpfend sollte mittels Röntgenkleinwinkelstreuung (zum Prinzip der Methode siehe Kapitel 2) an oxygeniertem und deoxygeniertem *N. pompilius*-Hämocyanin der Konformationswechsel im Zuge der Sauerstoffbindung nachvollzogen werden. Eine ähnliche Arbeit existiert bereits von unserer Gruppe zum Hämocyanin der Schlüssellochschnecke *Megathura crenulata* (Hartmann *et al.*, 2001; Hartmann *et al.*, 2004). Letztlich konnte nur ein Datensatz oxygenierten Proteins gewonnen werden und das Ziel dieses Teilprojektes erfuhr dahingehend eine Veränderung, dass der Datensatz der Kleinwinkelstreuung zur Validierung der gewonnenen EM-Rekonstruktion verwendet wurde. Die Kleinwinkelstreuung wurde zusammen mit B. Pairet durchgeführt und die Daten sind in seiner Diplomarbeit (Pairet, 2007) eingehend diskutiert. Der Vergleich EM-Rekonstruktion versus Kleinwinkelstreuendaten ist Bestandteil dieser Arbeit.

### **1.3 Bla-g Hexamerin von *Blattella germanica***

In der Arbeitsgruppe der Molekularen Biophysik stehen vor allem die gerade vorgestellten Hämocyanine von Arthropoden im Zentrum der Forschungsarbeiten. Aber auch verwandte Proteine, wie die Speicherhexamerine der Insekten, rückten in den letzten Jahren in den Fokus. Diese Proteine besitzen – nach bisherigem Wissen – einen eher „passiven“ Charakter: Sie besitzen keine enzymatische Funktion und binden keine Effektoren, lediglich eine Funktion als Transporter für Hormone wird diskutiert (Telfer & Kunkel, 1991). Die Speicherhexamerine gehören zur Hämocyaninsuperfamilie der Arthropoden (Burmester, 2001). Sie werden in vier Gruppen gegliedert: Calliphorine, die im Zuge der larvalen Entwicklung in die Hämolymphe sezerniert werden und in der weiteren Entwicklung der Tiere in den Fettkörpern gespeichert werden, ein spezielles Hexamerin der Diptera, welches sich ähnlich verhält wie die Calliphorine (Telfer & Kunkel, 1991), die Arylphorine, welche ebenfalls in ihrem Expressionsprofil den Calliphorinen gleichen und schließlich ein spezielle Hexamerine der Lepidoptera, deren Expressionsprofil von den übrigen Hexamerinen abweicht, aber ebenfalls mit der larvalen Entwicklung korreliert (die Eigenschaften der Hexamerine sind besonders gut in folgenden Arbeiten dargestellt: Beusch, 2006; Lafargue, 2008; Telfer & Kunkel, 1991. Abgesehen vom evolutionären Interesse an der Hämocyaninsuperfamilie und ihren physiologischen Eigenschaften (Burmester, 2001), stehen einige der Speicherhexamerine in Verdacht allergenes Potential zu entfalten. Bisher gab es in der Arbeitsgruppe des Institutes für Molekulare Biophysik drei Diplomarbeiten zu diesem Thema (Cirak, 2000; Lafargue, 2003; Beusch, 2006), in denen die Aufreinigung und physikochemische Charakterisierung stattfand.

#### *Ziele dieser Arbeit*

Ziel der Diplomarbeit von A. Beusch (Beusch, 2006) war es u. a. das Hexamerin der Deutschen Schabe (*Blattella germanica*) aufzureinigen und zu charakterisieren. Im Rahmen meiner Arbeit wurde mit Hilfe elektronenmikroskopischer Daten die Annahme, dass es sich um ein Hexamerin handelt, verifiziert und zugleich weitere Informationen zur strukturellen Organisation geliefert.

## 1.4 Struktur & Funktion des $\alpha$ -Hämolytins

*Staphylococcus aureus* ist ein grampositives Bakterium, welches im Zuge einer mikroskopischen Untersuchung im Allgemeinen in Paaren, kurzen Ketten oder haufenartig vorkommt (vgl. rechtsstehende Abbildung). Das Bakterium wird häufig auf menschlichen Oberflächen, wie zum Beispiel der Nasenschleimhaut (Ostfeld *et al.*, 1983; Nishijima *et al.*, 1995), der Haut (Noble, 1998) oder – zumindest transient – dem Oropharynx (Smith *et al.*, 2001) oder Faeces (Arvola *et al.*, 2006) angetroffen. Daher wird davon ausgegangen, dass der Organismus selbst keinerlei Krankheiten verursachen kann, sondern quasi opportunistisch von Verletzungen der angesprochen (Schleim-) Häute profitiert (Veerachato, 2006). Doch sogar im Deutschland der Gegenwart kommt es immer wieder zu Ausbrüchen lokal begrenzter Epidemien, verursacht von *Staphylococcus aureus* (Wiese-Posselt *et al.*, 2007). Ist eine Kolonie erst in ein Wirtsgewebe eingedrungen, so kann sie eine Vielzahl von Krankheiten auslösen. Diese reichen von kleineren Hautinfektionen und Abzessen, bis zu lebensbedrohlichen wie Pneumonie, Meningitis, Endocarditis, toxischem Schock Syndrom (TSS) und Sepsis. Während der letzten Jahre kam es zudem zu einer Häufung von antibiotikaresistenten Stämmen des Bakteriums, die eine ernstzunehmende Gefahr im Gesundheitswesen darstellen, insbesondere in Krankenhäusern (Arvola *et al.*, 2006; Rice, 2006).

Neben der opportunistischen Fähigkeit viele Antibiotikaresistenzen zu erwerben, resultiert die Gefährlichkeit des Bakteriums aus dem breiten Arsenal von Toxinen, die es sekretieren kann. Hierzu zählen Agentien, welche die Integrität von Zellmembranen des Wirtes angreifen, Superantigene und eine Vielzahl hydrolytischer Enzyme. Diese Toxine stellen für das Bakterium einerseits eine Verteidigung gegen das Immunsystem der Wirtes und andererseits eine effektive Möglichkeit durch Nekrose des Wirtsgewebes „Nährboden“ zu gewinnen dar.

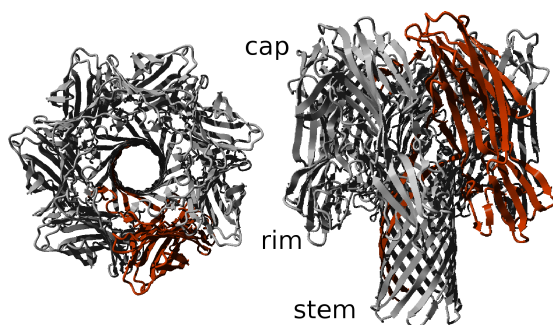


Abbildung 1.13: Aufsicht auf das  $\alpha$ -Hämolytin von der extrazellulären Seite und eine Seitenansicht; schematische Abbildung nach der Struktur von Song *et al.*, 1996. Jeweils eines der sieben Monomere ist eingefärbt. An der Seitenansicht sind die Domännennamen angegeben.

1968, und von Bhakdi & Trantum-Jensen, 1991. Das Gen des  $\alpha$ -Hämolytins, *hla*, wurde bereits vor mehr als zwanzig Jahren kloniert und sequenziert (Gray & Kehoe, 1984). Die ersten 26 Aminosäuren stellen eine Signalsequenz dar, die abgeschnitten wird, wenn das Toxin sekretiert wird. Das reife Protein besitzt 293 Aminosäuren und hat ein Molekulargewicht von 33248 Da. Das Gen ist unter Kontrolle eines akzessorischen Genregulators (*agr*), der bewirkt, dass das Toxin erst in der späten Phase exponentiellen Wachstums exprimiert wird. Die Monomere des  $\alpha$ -Hämolytins

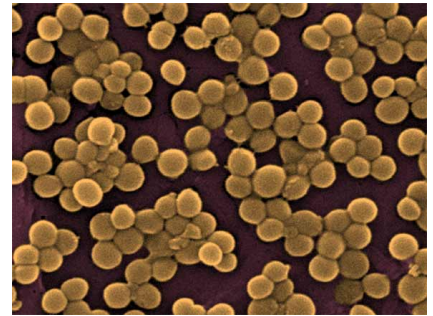


Abbildung 1.12: Falschfarbendarstellung von Zellen von *Staphylococcus aureus* in der Rasterelektronenmikroskopischen Aufnahme bei 9500-facher Vergrößerung. (Quelle: www.biology4kids.com)

Eines der bedeutendsten Toxine von *Staphylococcus aureus* ist das  $\alpha$ -Hämolytin (oder  $\alpha$ -Toxin; vgl. die linksstehende Abbildung). Dieses Toxin wird von einer großen Zahl von Stämmen des Bakteriums (wenngleich nicht allen; Bhakdi & Trantum-Jensen, 1991) sekretiert und wirkt auf viele Säugetierzellen toxisch. Neuere Arbeiten konnten einerseits zeigen, dass die Virulenz bedeutender *S. aureus*-Stämme vor allem auf dem  $\alpha$ -Hämolytin beruht, andererseits aber auch eine Impfung in den Bereich des Möglichen rückt (Wardenburg & Schneewind, 2008). Der Name des Toxins stammt von seiner hämolytischen Wirkung. Kaninchenerythrocyten sind besonders anfällig. Das Toxin wirkt überdies dermonekrotisch und neurotoxisch. Bereits die Menge von einem Mikrogramm wirkt auf Kaninchen lethal, sofern das Toxin intravenös appliziert wird. Eine breite Übersicht bieten die Artikel von Freer *et al.*,

werden ausgeschieden und assemblieren auf den Membranen von Wirtszellen zu heptameren Poren (Gouaux *et al.*, 1994; Vandana *et al.*, 1997). Die Genese der Pore wurde ausgiebig studiert und das  $\alpha$ -Hämolyisin avancierte zum Modellprotein für vergleichbare porenbildende Proteine. Es wurde zur Lösung der Frage wie Proteine eine Membran penetrieren und in sie inserieren können herangezogen. Folgendes Schema mag den bisherigen Konsens zu den strukturellen Aspekten der Porengnese illustrieren (modifiziert nach Gouaux, 1998):



Die bisherige Vorstellung zur Porengnese ist folgendermaßen zusammenzufassen: Das Toxinmonomer (Status  $\alpha_1$ ) bindet an die Membran ( $\alpha_1^*$ ). Diese Bindung kann offenbar über zwei verschiedene Wege erfolgen (Hildebrand *et al.*, 1991). Liegt das Toxin in geringer Konzentration vor, so bindet es an einen spezifischen, wenngleich unbekanntem, Rezeptor. In höherer Konzentration aber lagert es sich unspezifisch an die Testmembran. Valeva *et al.*, 2006, schlagen vor, dass der Rezeptor als Quasirezeptor zu betrachten ist, welcher durch die räumliche Konzentration spezifischer, im Grunde niedrigaffiner, Lipide in sogenannten *rafts* generiert wird. Diffusionsgetrieben finden dann die Monomere auf der Membran (Status  $\alpha_1^*$ ) zueinander und assemblieren zu einer „Präpore“ (Valeva *et al.*, 1997; Status  $\alpha_7^*$ ). Biochemische Experimente machen deutlich, dass das Präporenstadium in mindestens drei Schritte (a, b und c) unterteilt werden muss (Valeva *et al.*, 1997). Diese Schritte unterscheiden sich in Kontaktflächenformung zwischen den Protomeren, insbesondere am Rest H35, und dem Grad zu dem der glycinreiche Transmembranteil einer hydrophoben Umgebung ausgesetzt wird. Zustand  $\alpha_{7,a}^*$  ist bereits oligomerisiert, aber empfindlich gegenüber dem Detergens SDS (Natriumdodecylsulfat) – im Gegensatz zur Porenkonformation  $\alpha_7$ . Für den folgenden Zustand,  $\alpha_{7,b}^*$  konnte gezeigt werden, dass S130, derjenige Rest, der in der Porenstruktur von Song *et al.*, 1996, den „tiefsten“ Punkt der *stem*-Domäne markiert, bereits in den hydrophoben Bereich der Membran eintaucht. Und im Zustand  $\alpha_{7,c}^*$  schließlich ändert sich die Konfiguration des Restes H35 so, dass dieser ebenfalls in den hydrophoben Bereich der Membran taucht.

Während des Umfaltens vom kompakten Monomer zur Pore (Status  $\alpha_7$ ) muss zwar ein erheblicher Teil des Proteins bewegt werden, jedoch ändert sich der Anteil der Sekundärstrukturelemente kaum (Ikigai & Nakae, 1985; Tobkes *et al.*, 1985). Die Struktur der Pore ( $\alpha_7$ ) konnte geklärt werden (Song *et al.*, 1996; siehe Abbildung 1.13). Die Gestalt dieser Pore erinnert stark an den Fruchtkörper eines Pilzes. Die Pore hat eine Höhe von ca. 100 Å und in etwa denselben Durchmesser. Der innere Durchmesser der Pore variiert von 15 Å bis 46 Å. Leitfähigkeitsanalysen deuten darauf hin, dass der effektive Durchmesser mit  $11.4 \text{ Å} \pm 0.4 \text{ Å}$  etwas geringer ausfällt als der kleinste sterisch gegebene Durchmesser (Menestrina, 1986). Die Pore wird durch drei Domänen charakterisiert: *Cap*, *Rim* und *Stem* (vgl. Abbildung 1.13). Die *Stem*-Domäne ist ein 14-strängiges antiparalleles  $\beta$ -Barrel. Die *Cap*-Domäne stellt einen großen, hydrophilen Bereich dar, der in das extrazelluläre Medium ragt. Und die *Rim*-Domäne schließlich stellt die Unterseite der *Cap*-Domäne dar und besitzt eine definierte Phosphocholinbindestelle (Galdiero & Gouaux, 2004).

Die Zahl der Untereinheiten konnte kristallographisch und biochemisch mit 7 festgelegt werden (Gouaux *et al.*, 1994; Song *et al.*, 1996), doch deuteten ältere biochemische und elektronenkristallographische Arbeiten auf eine hexamere Struktur hin (Bhakdi *et al.*, 1981; Ikigai & Nakae, 1985; Tobkes *et al.*, 1985; Ward & Leonard, 1992). Diese Diskrepanz wird von Gouaux als Möglichkeit der Bildung von Hexameren unter besonderen Umständen diskutiert (Gouaux, 1998). Elektronenkristallographische Arbeiten (Olofsson *et al.*, 1990; Olofsson *et al.*, 1991) deuteten auf Oktamere hin. Eine Arbeit (Czajkowsky *et al.*, 1998), die auf Rasterkraftmikroskopie (Englisch: *atomic force microscopy* bzw. AFM) basiert und mit verschiedenen Lipidzusammensetzungen arbeitet, bestätigt die Vermutung (Gouaux, 1998), dass insbesondere die Zusammensetzung der Membran den vorherrschenden Oligomerisierungsgrad vorgibt.

*Ziele dieser Arbeit*

Am Beginn dieser Arbeit stand die Frage im Raum, welche strukturellen Rahmenbedingungen während der Porenbildung gelten: Wenngleich vieles zum Porenbildungsmechanismus bereits bekannt ist, so können doch die strukturellen Veränderungen in diesem Prozess nur unzureichend beschrieben werden. So ist beispielsweise die Struktur des Monomers (Status  $\alpha_1$ ) unbekannt; keine Gruppe hat es bisher geschafft Kristalle des monomeren Proteins zu züchten. Auf Basis der Struktur bekannter und verwandter Toxine (Olson *et al.*, 1999; Pédelacq *et al.*, 1999; Guillet *et al.*, 2004), die ebenfalls zur Gruppe der  $\beta$ -kanalformenden Toxine ( *$\beta$ -channel forming toxins* oder  $\beta$ -CFTs) gehören, wurde ein Versuch unternommen die Struktur des monomeren  $\alpha$ -Hämolytins vorzuschlagen. Eine solche Struktur würde den Startpunkt der Transition  $\alpha_1 \rightarrow \alpha_7$  festlegen. Zum Zeitpunkt dieser Arbeit fehlte jegliche strukturelle Information zum Präporenstatus ( $\alpha_7^*$ ) – gleich zu welchem dieser Zwischenstadien (a, b oder c). Mit H35R stand eine Mutante des  $\alpha$ -Hämolytins zur Verfügung, für die nachgewiesen ist auf der Membran zu assemblieren, aber doch in einem Präporenstadium „stecken bleibt“ und nicht in die Membran inseriert (Jursch *et al.*, 1994; Panchal & Bayley, 1995). Dieses Präporenstadium ist wahrscheinlich gleichzusetzen mit  $\alpha_{7,a}^*$  (Valeva *et al.*, 1997). Zusätzliche Strukturen dieser beiden Stadien würden nicht nur das strukturell-sterische Verständnis der Membraninsertion verbessern helfen, sondern auch zu einem besseren physikochemischen Verständnis führen und klären, wie genau die Energiebarriere der Insertion überwunden werden kann.

# Kapitel 2

## Einführung in die Theorie der Kleinwinkelstreuung

Kleinwinkelstremethoden sind wesentlicher Bestandteil dieser Arbeit, daher gibt dieses Kapitel eine kurze Einführung in ihre Theorie. Überdies werden Aspekte angesprochen, die für die Interpretation der Ergebnisse und der algorithmischen Umsetzung in der Software wesentlich sind, die im Rahmen dieser Arbeit entstand. Viele Aspekte der Röntgen- und Neutronenkleinwinkelstreuung gleichen sich. Daher wird im Folgenden nur dann auf Unterschiede hingewiesen, wenn diese bedeutsam sind.

### 2.1 Erzeugung von Röntgen- und Neutronenstrahlung

#### 2.1.1 Röntgenquellen

##### 2.1.1.1 Synchrotronquellen

Die beliebtesten Quellen intensiver Röntgenstrahlen sind gegenwärtig sogenannte Synchrotrons. Zur Erzeugung der Röntgenstrahlung werden Elektronen in einer (nahezu) kreisförmigen Anlage (ein sogenannter „Speicherring“) stark beschleunigt. Dieser hochenergetische Elektronenstrahl emittiert bei der Ablenkung in der Kurve des Speicherrings, meist zusätzlich provoziert durch sogenannte Ablenkmagnete, brillantes Röntgenlicht. Aufgrund der möglichen Größe der angegliederten Experimentiereinrichtungen (bis zu 200 m Länge für ein Strahlrohr) kann der Strahl beliebig gut kollimiert<sup>1</sup> werden. Daher stellen Synchrotrons gegenwärtig in der strukturellen biologischen Forschung unverzichtbare Instrumente dar.

##### 2.1.1.2 Röntgeneratoren

Im Falle eines Röntgenerators werden Elektronen beschleunigt und auf eine Metallplatte gelenkt. Um den Röntgenfluss konstant zu halten, werden die Anoden gedreht und sind mit einer Wasserkühlung versehen. Die Konstruktion dieser Anodenröhren hat den Namen „Röntgenröhren“ geprägt. Röntgenstrahlung wird in dem Moment generiert, in dem Elektronen auf die Anode treffen. Dabei entsteht Bremsstrahlung. Diese entsteht stets, wenn ein geladenes Teilchen seine kinetische Energie, ausgelöst durch eine starke Geschwindigkeitsänderung, in elektromagnetische Strahlung umwandelt. Eine genaue Betrachtung gaben Bethe & Heitler (Bethe & Heitler, 1934). Ein Aspekt der Bremsstrahlung ist, dass ein breites Spektrum entsteht (siehe Abbildung Tsuruta & Johnson, 2001), das erst durch Filter monochromatisiert werden muss (siehe Abschnitt 3.6 auf Seite 32).

Moderne Röntgeneratoren mit entsprechenden Kollimationsverfahren erreichen in vielen Fällen bereits die Qualität der in Synchrotrons aufgenommenen Daten (z. B. Round *et al.*, 2005; siehe Produktinformation von Bruker-ASX, 2008; Anton-Paar-GmbH, 2008). Lediglich bei zeitaufgelösten Messungen ist ein Synchrotron zu bevorzugen.

---

<sup>1</sup>Kollimation bezeichnet in der Optik die parallele Ausrichtung von Strahlung.

### 2.1.2 Neutronenquellen

Neutronen sind neutrale, subatomare, elementare Partikel, die durch Rutherford postuliert und von Chadwick experimentell nachgewiesen wurden (Chadwick, 1932a; Chadwick, 1932b). Außer in Wasserstoff  $^1\text{H}$  sind sie Teil aller Atomkerne, haben eine Masse ähnlich wie Protonen und einen Kernspin von  $1/2$  sowie einen magnetischen Impuls (für eine allgemeine Einleitung zu Eigenschaften von Neutronen siehe Harrown *et al.*, 2007). Ein Neutronenstrahl kann tief in Materie eindringen (siehe auch Abschnitt 2.2.3). Diese Fülle von Eigenschaften wird auf vielfältige Weise für die Klärung biologischer Fragestellungen genutzt. Eine sehr gute Übersicht bietet Fitter *et al.*, 2007. Hier werden im Folgenden ausschließlich Aspekte mit Relevanz für die Kleinwinkelstreuung erörtert. Neutronenstrahlung mit einer für Streuexperimente ausreichenden Intensität werden gegenwärtig entweder in Kernreaktoren oder in sogenannten Zertrümmerungs-Quellen (Englisch: *spallation sources*) generiert. Im Falle der Reaktoren erzeugt die Spaltung von Uran Neutronen im Energiebereich zwischen 0,5 und 3 MeV. Bei den Zertrümmerungsquellen treffen beschleunigte subatomare Partikel (z. B. Protonen) auf Schwermetalle wie Wolfram, Blei oder flüssiges Quecksilber und erzeugen dabei Neutronen. Die Entwicklung dieser Quellen erlaubt es erst in den letzten Jahren an einen Einsatz in biologischer Forschung zu denken.

#### *Geesthacht Neutron Facility - GeNF*

Die Neutronenstreudaten dieser Arbeit wurden an der GKSS („Gesellschaft für Kernenergieverwertung in Schiffbau und Schifffahrt mbH“), Geesthacht, generiert. Die GKSS beherbergt mit der sogenannten *Geesthacht Neutron Facility* (GeNF) einen Kernreaktor (FRG1) aus den 1950er-Jahren. Der Reaktor nutzt  $\text{U}_3\text{Si}_2$  mit einer Anreicherung von 20 % als Kernbrennstoff (Ulrich & Schreiner, 2007). Im Kern ist eine Quelle für „kalte Neutronen“ integriert. Diese ermöglichen eine Vielzahl von Experimenten, vornehmlich für materialwissenschaftliche Anwendungen. Daneben gibt es auch eine für organische Proben spezialisierte Neutronenkleinwinkelstreueinrichtung, SANS1, deren Schema in der folgenden Abbildung gezeigt ist.

Die Parameter der Experimentiereinrichtung (vgl. Abbildung 2.1) sind im Material & Methoden-Kapitel unter 3.10.6.2 zu finden. Prinzipiell funktioniert das Instrument jedoch so, dass sogenannte thermische Neutronen aus dem Reaktorkern austreten und durch einen Behälter mit kaltem Wasserstoff ( $T = 25 \text{ K}$ ) geleitet werden. Dieser Behälter dient als Quelle kalter (= abgebremster, siehe den folgenden Abschnitt) Neutronen und wird auch als Moderator bezeichnet. Da diese Neutronen aber noch eine zu breite Wellenlängenverteilung aufweisen, werden sie durch einen Geschwindigkeitsselektor geleitet: Dieser besteht aus einer rotierenden Trommel mit Schlitzen. Wenn die Neutronen mit der „richtigen“ Geschwindigkeit (bzw. Wellenlänge; siehe Gleichung 2.2) ankommen, so passieren sie die Schlitze ungehindert; ansonsten verhindert die Krümmung der Schlitze das Passieren und die Neutronen werden absorbiert. Anschließend wird der Strahl kollimiert, das heißt er wird „zurechtgestutzt“, so dass möglichst auf einem definierten Fleck mit parallelen Strahlen belichtet werden kann. Wenngleich sich die Kollimationsgeometrie bei der SANS1-Anlage deutlich von der Geometrie in der Kratky-Kamera unterscheidet, so gelten die in Abschnitt 2.3 erläuterten grundsätzlichen mathematischen Überlegungen auch hier. Nachdem der Strahl durch die Probe tritt, fällt nicht gestreute Strahlung auf einen Primärstrahlfänger und der gestreute Anteil auf den nachgeschalteten Detektor. Der Detektor selber ist in einer evakuierten Röhre fahrbar gelagert und kann so auf verschiedene Distanzen zur Probe eingestellt werden. Damit wird erreicht, dass die gestreute Strahlung über einen möglichst weiten Winkelbereich aufgezeichnet werden kann.

## 2.2 Eigenschaften von Röntgen- und Neutronenstrahlung

### 2.2.1 Energie und Impulsübertrag

In einem Streuexperiment erfährt ein einfallendes Teilchen eine Veränderung des Impulses nach Wechselwirkung mit der Probe. Das bedeutet, dass dieses Teilchen eine Veränderung in Richtung und/oder Geschwindigkeit erfährt. Im Falle von Photonen (Röntgenstrahlung) ist die Wellenlänge

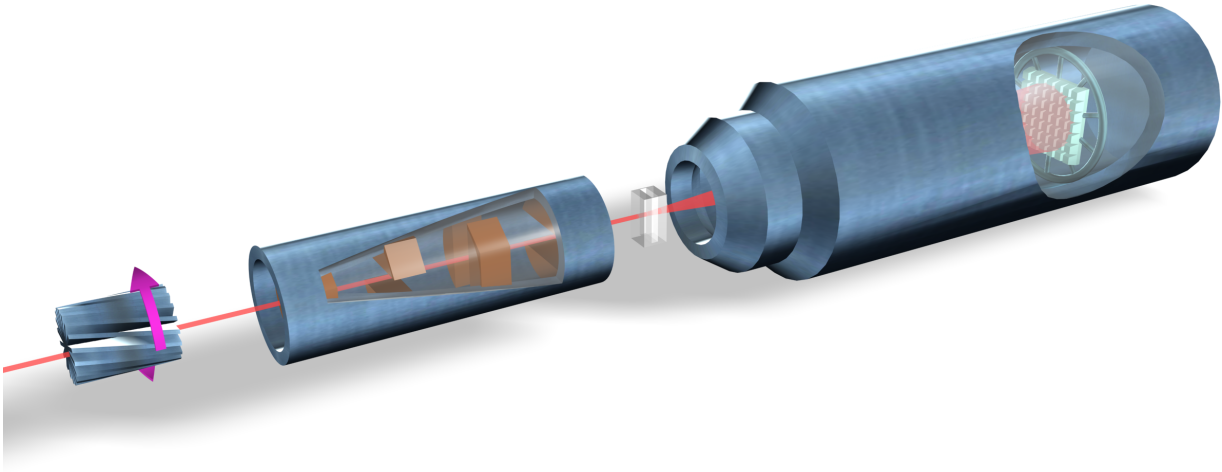


Abbildung 2.1: *Schema der Neutronenkleinwinkelstreuanlage am FRG1 der GKSS.* Der Neutronenstrahl fliegt von links nach rechts: Ausgehend vom Moderator trifft er auf das erste Element des Schemas, den rotierenden Geschwindigkeitsselektor. Im Anschluss folgt das Kollimationssystem mit einer Reihe von Schlitzblenden variabler Größe, worauf der Strahl auf die Probe in einer Küvette trifft. Der nicht gestreute Primärstrahlanteil wird durch einen Primärstrahlblocker ausgeblendet (nicht dargestellt), bevor der gestreute Anteil auf den Detektor trifft. Der Detektor ist fahrbar in einer Röhre gelagert, so dass verschiedene Streuwinkelbereiche erfasst werden können. Die Abbildung wurde von A. Neu erstellt und freundlicherweise für diese Arbeit zur Verfügung gestellt.

gegeben durch die Lichtgeschwindigkeit ( $c$ ) und die Frequenz ( $f$ ) der Strahlung:

$$\lambda = \frac{c}{f} . \quad (2.1)$$

Für Neutronen ist der Impuls durch  $p = \hbar k$  gegeben, wobei  $\hbar$  die Plancksche Konstante  $h/2\pi$  ist und  $k$  der Wellenvektor der Neutronen  $|k| = 2\pi/\lambda$ . Die Wellenlänge der Neutronen ergibt sich aus

$$\frac{h^2}{2m\lambda^2} = 2k_B T , \quad (2.2)$$

wobei  $k_B$  die Boltzmannkonstante und  $T$  die Moderatortemperatur (siehe 2.1.2) darstellen. Die Impulsänderung kann durch den Streu- oder Impulsübertragsvektor  $q$  ausgedrückt werden. Dieser ist als Vektordifferenz zwischen den ein- und ausgehenden Streuwellenvektoren definiert:

$$\vec{q} = k_0 - k_1 , \quad (2.3)$$

wobei  $k_0$  und  $k_1$  dem ein- bzw. ausgehenden Streuwellenvektoren entsprechen (siehe 2.2.2). Die Veränderung des Impulses ist damit  $\hbar\vec{q}$ . Im Gegensatz zu der elastischen Streuung, bei welcher der einfallende Strahl keine Energie verliert und die Wellenlänge konstant bleibt, gibt es auch noch die inelastische Streuung. Im Fall der inelastischen Streuung verlieren (oder gewinnen) die gestreuten Teilchen Energie, durch Wechselwirkung mit Atomkernen im Falle der Neutronenstreuung oder durch Wechselwirkung mit den Elektronenhüllen der Atome im Falle der Photonen (bei der Röntgenstreuung). Im Folgenden wird allerdings ausschließlich der elastische Fall betrachtet, da nur bei einer konstanten Wellenlänge die vorgestellten Analysemethoden aussagekräftig sind.

### 2.2.2 Streuung & die Debye-Gleichung

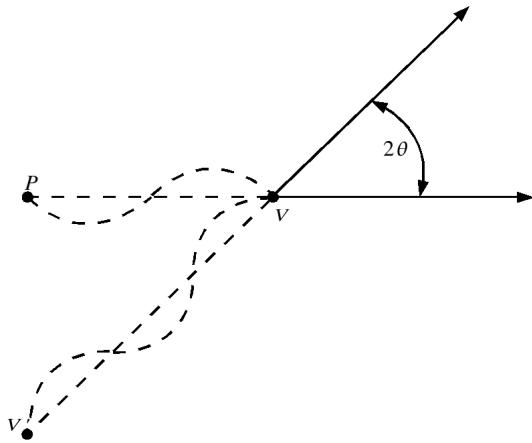


Abbildung 2.2: *Schema der Geometrie von Streueignissen.* Für jeden Punkt  $V$ , der in einem Winkel  $2\theta > 0$  streut, gibt es einen Punkt  $V'$ , der genau ( $180^\circ$ ) außer Phase streut. Folglich löscht sich alle Streuung aus, es sei denn, das Streuvermögen in  $V$  und  $V'$  unterscheidet sich. Abbildung nach Fitter *et al.*, 2007.

Im Allgemeinen streut einfallende Strahlung in ein Medium, welches unabhängig von seiner Position streut, nur in Vorwärtsrichtung ( $\phi = 2\theta = 0$ ). Jedoch gibt es für jedes Volumenelement ( $V$ ) ein Volumenelement ( $V'$ ), welches genau ( $180^\circ$ ) außer Phase streut, wobei gilt:  $PV - VV' = \lambda/2$  (siehe Abbildung 2.2). Folglich löscht sich alle weitere Streuung aus, es sei denn, das Streuvermögen in  $V$  und  $V'$  unterscheidet sich. Röntgenstrahlung wechselwirkt mit den Elektronen einer Probe und ist folglich durch Fluktuationen der Elektronendichte gestreut. Neutronen wechselwirken nicht mit Elektronen, sondern mit den Kernen der Atome (siehe 2.2.1).

Für die elastische Streuung ergibt sich aus Gleichung 2.3, dass  $|k_1| = |k_0|$  und folglich ist  $|\vec{q}| = 2k_0 \sin \theta$ . Für kristalline Materialien resultieren sogenannte Bragg peaks bei Werten von  $\vec{q}$  gleich dem reziproken Gitterabstand:

$$|\vec{q}| = \frac{2\pi}{d}, \quad (2.4)$$

wobei  $d$  der Abstand im Kristall zwischen zwei Netzebenen ist. Da  $k_0 = 2\pi/\lambda$ , ergibt sich durch Ersetzen das Braggsche Gesetz:

$$\lambda = 2d \sin \theta. \quad (2.5)$$

Vereinfacht ausgedrückt ist dies die Bedingung für konstruktive Interferenz (vgl. Abbildung 2.3) von Wellen mit dem Einfallswinkel  $\theta$  auf eine Menge equidistanter Ebenen mit dem Abstand  $d$ .

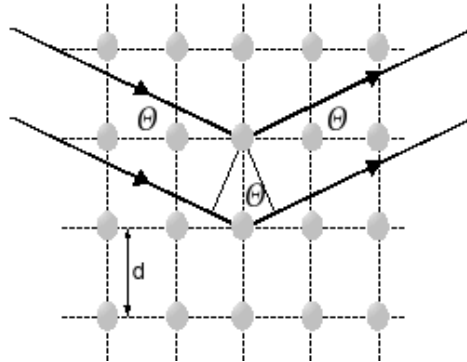


Abbildung 2.3: *Darstellung zum Braggschen Gesetz.* Ein einfallender Neutronen- oder Röntgenstrahl trifft von links auf eine regelmäßige Anordnung von Atomen und wird nach rechts gestreut: Der horizontale Abstand der Atome ist  $d$ . Sowohl der einfallende als auch der gestreute Strahl stehen in einem Winkel  $\theta$  zu den Ebenen der Atome. Abbildung nach Fitter *et al.*, 2007.

Obwohl das Braggsche Gesetz sich auf kristalline Materialien beschränkt, gilt der Fourier- oder inverse Zusammenhang zwischen einer Struktur im realen Raum ( $r$ ) und der Streuung im reziproken Raum: Gleichung 2.3 kann präziser folgendermaßen ausformuliert werden:

$$q = |\vec{q}'| = |k_0 - k_1| = |k| \cdot |e_0 - e_1| = 4\pi \hbar \frac{\sin \theta}{\lambda}. \quad (2.6)$$

Für die Bedeutung von  $e$  siehe Abschnitt 2.2.3.

In der Literatur erfolgt jedoch meist eine Darstellung ohne  $\hbar$ :

$$q = |\vec{q}| = \frac{|\vec{q}'|}{\hbar} = 4\pi \frac{\sin \theta}{\lambda}. \quad (2.7)$$

Die Amplitude einer Streuwelle von zwei Atomen,  $x$  und  $y$ , ergibt sich aus den Atomformfaktoren der Atome:

$$A_{x,y} = A_e \cdot f_{x,y} . \quad (2.8)$$

Hierbei ist  $A_e$  die Streuamplitude des freien Elektrons im Vakuum und  $f_x$  bzw.  $f_y$  die Zahl der Elektronen im jeweiligen Atom. Die Phasendifferenz zwischen den beiden Streuwellen lässt sich aus der Projektion des Ortsvektors  $\vec{r}$  auf den Streuvektor  $\vec{q}$  bestimmen, dabei ergibt sich:

$$\phi = \vec{r} \cdot \vec{k}_0 - \vec{r} \cdot \vec{k}_1 = \vec{r} \cdot \vec{q} . \quad (2.9)$$

Moleküle jedoch bestehen üblicherweise aus sehr viel mehr Atomen. Gleichung 2.8 muss also erweitert werden zu:

$$A(\vec{q}) = \sum_i A_i \cdot e^{-i\vec{r}_i \cdot \vec{q}} = A_e \sum_i f_i \cdot e^{-i\vec{r}_i \cdot \vec{q}} , \quad (2.10)$$

wobei  $i$  der Laufindex über die Atome ist.

Die Amplitude einer Welle ist jedoch nicht messbar, sondern nur die Intensität. Es gilt:

$$I(\vec{q}) = |A(\vec{q})|^2 . \quad (2.11)$$

Eine eingehendere mathematische Erläuterung liefert Pairat, 2007. Werden Moleküle in Lösung betrachtet, so ist – bei einer großen Zahl von Molekülen – die Streuung isotrop, d. h. es gibt keine bevorzugte Richtung für die Streuung, da die Moleküle keine bevorzugte Orientierung gegenüber dem Strahl einnehmen. Gleichung 2.11 muss also über alle Raumrichtungen gemittelt werden:

$$I(q) = \langle I(\vec{q}) \rangle_\Omega = A_e^2 \langle \sum_{x,y} f_x \cdot f_y \cdot \cos((\vec{r}_x - \vec{r}_y) \cdot \vec{q}) \rangle_\Omega , \quad (2.12)$$

wegen  $e^{ix} = \cos x + i \sin x$ . Um das Mittel über die Raumrichtungen auszuführen, dient folgender Zusammenhang, bei dem der Abstandsvektor  $\vec{r} = \vec{r}_x - \vec{r}_y$  gesetzt ist:

$$\begin{aligned} \langle \cos(\vec{r} \cdot \vec{q}) \rangle_\Omega &= \frac{1}{2} \int_0^\pi \cos(\vec{r} \cdot \vec{q}) \cdot \sin(\varphi) d\varphi \\ &= \frac{1}{2} \int_0^\pi \cos(r \cdot q \cdot \cos(\varphi)) \cdot \sin(\varphi) d\varphi \\ &= \int_{-qr/2}^{qr/2} \cos(x) \cdot \frac{1}{qr} \cdot dx \\ &= \frac{\sin(qr)}{qr} , \end{aligned}$$

wobei  $\phi$  der Winkel zwischen  $\vec{r}$  und  $\vec{q}$  ist und  $r$  und  $q$  die Beträge der Vektoren darstellen. Hiermit ergibt sich aus Gleichung 2.12 die Debye-Gleichung:

$$I(q) = A_e^2 \cdot \sum_{x,y} f_x \cdot f_y \cdot \frac{\sin(qr)}{qr} . \quad (2.13)$$

Die Debye-Gleichung beschreibt den Zusammenhang zwischen den Formfaktoren der beteiligten Atome und ihren Abständen. Sie gilt für alle zufällig orientierten Moleküle – sofern diese sich in einem perfekten Vakuum befinden. Tatsächlich aber müssen Proteine und andere Biomoleküle in einem geeigneten Puffer gemessen werden. Analog zu Gleichung 2.22 verringert sich der Kontrast des Systems um den Kontrast des Puffers. Zusätzlich muss berücksichtigt werden, dass die Streuung in einem Molekül nicht an definierten Streuzentren stattfindet, sondern an der gesamten Elektronendichte,  $\rho(r_p)$ , des Moleküls. Mit folgender Substitution

$$A_e \cdot f_p = \rho(r_p) \cdot dV_p , \quad (2.14)$$

wobei  $p$  für das Molekül – ein Protein – steht, geht Gleichung 2.12 in folgendes Integral über:

$$I(\vec{q}) = \int_{V_1} \int_{V_2} \rho(\vec{r}_1) \cdot \rho(\vec{r}_2) \cdot \cos((\vec{r}_1 - \vec{r}_2) \cdot \vec{q}) d^3r_1 \cdot d^3r_2 , \quad (2.15)$$

wobei  $V_1$  und  $V_2$  für die eingangs erwähnten Volumenelemente stehen. Setzt man  $\vec{r} = \vec{r}_1 - \vec{r}_2$ , so ergibt sich:

$$I(\vec{q}) = \int_{V_2=V} \underbrace{\int_{V_1} \rho(\vec{r}_1) \cdot \rho(\vec{r}_1 - \vec{r}) \cdot d^3\vec{r}_1 \cdot \cos(\vec{r} \cdot \vec{r}\vec{q}) \cdot d^3\vec{r}}_{\gamma(r)}. \quad (2.16)$$

Hierbei ist  $\gamma(r)$  die Patterson-Funktion, eine Autokorrelation der Elektronendichte ( $\rho(r) \times \rho(r)$ ). Analog zu den gerade durchgeführten Umformungen ergibt sich mit  $\gamma(r) = \rho(r) \times \rho(r)$ :

$$I(q) = \int_V \gamma(r) \cdot \frac{\sin(qr)}{qr} \cdot d^3\vec{r}. \quad (2.17)$$

Wiederum sind  $r$  und  $q$  die Beträge der Vektoren  $\vec{r}$  und  $\vec{q}$ . Analog zu 2.12 kann auch hier die Integration über alle Raumrichtungen ausgeführt werden:

$$I(q) = 4\pi \int r^2 \cdot \gamma(r) \cdot \frac{\sin(qr)}{qr} \cdot dr. \quad (2.18)$$

Der Ausdruck  $r^2 \cdot \gamma(r)$  wird als Abstandsverteilungsfunktion  $P(r)$  zusammengefasst. Man erhält

$$I(q) = 4\pi \int P(r) \cdot \frac{\sin(qr)}{qr} \cdot dr. \quad (2.19)$$

Abstandsverteilungsfunktion  $P(r)$  und Streuintensität  $I(q)$  lassen sich also ineinander überführen<sup>2</sup>. Dieser Sachverhalt ist zentral für die Praxis der Kleinwinkelstreuung und findet auch bei der indirekten Fouriertransformation (siehe Abschnitt 2.4 auf Seite 24 dieses Kapitels) Verwendung.

### 2.2.3 Streulänge und Wirkungsquerschnitt

Typischerweise wird in einem Streuexperiment die Zahl von Teilchen, die in einer bestimmten Richtung gestreut werden, festgestellt – ohne Änderungen der Energie, also ohne den inelastischen Fall (siehe 2.2.1) zu berücksichtigen. Der *differentielle Wirkungsquerschnitt* ist bei einem festen Winkel  $d\Omega$  folgendermaßen definiert:

$$\frac{d\sigma}{d\Omega} = \frac{(\text{Teilchen } s^{-1} \text{ die in } d\Omega \text{ gestreut werden})}{\Phi d\Omega}. \quad (2.20)$$

Hierbei ist  $\Phi$  der einfallende Teilchenflux (Zahl der Teilchen pro Fläche und Sekunde). Der *gesamte Wirkungsquerschnitt* ist nun definiert als die in einer Sekunde gestreute Gesamtteilchenzahl, normalisiert durch den Flux:

$$\sigma_s = \int \left( \frac{d\sigma}{d\Omega} \right) d\Omega. \quad (2.21)$$

Bei bekanntem Flux ergeben sich also alle diese Parameter zwangsläufig. Während bei Neutronenquellen der Flux üblicherweise bekannt ist, kann er für Röntgenquellen bestenfalls geschätzt werden. Bei Röntgenquellen wird daher meist nur ein relativer Flux angegeben. Auch Flächendetektoren für Neutronen können absolute Angaben über die detektierten Teilchen machen, während mit Röntgendetektoren keine genauen Angaben gewonnen werden. Folglich sind Angaben über Neutronenstreuintensitäten in dieser Arbeit auch absolut skaliert, während Angaben für Röntgenstreuintensitäten dies nicht sein können.

### 2.2.4 Kohärenter und inkohärenter Wirkungsquerschnitt bei Neutronen

Atomkerne sind u. a. durch eine inkohärente und eine kohärente Neutronenstreulänge  $b$  charakterisiert. Wenn ein Neutron mit einem Spin von  $1/2$  auf ein Isotop mit einem Kernspin  $I$  trifft, so kann der Spin des Neutronen-Kern-Systems zwei Werte annehmen:  $I \pm 1/2$ . Die Streulänge

<sup>2</sup>Bei der hierzu notwendigen Transformation handelt es sich nicht wie in Pairet, 2007 und Hartmann & Decker, 2004 erörtert um eine Sinustransformation. Eine detaillierte Betrachtung einer Sinustransformation liefern z. B. Press *et al.*, 2002.

der zwei möglichen Systeme ist mit  $b^+$  und  $b^-$  angegeben. Die möglichen Spinzustände, die mit den beiden Werten assoziiert sind, sind  $2(I + 1/2) + 1 = 2I + 2$  und  $2(I - 1/2) + 1 = 2I$ . Die Gesamtzahl der möglichen Zustände ist also  $4I + 2$ . Sind die Neutronen nicht polarisiert und die Kernspins zufällig orientiert, so haben alle Spinzustände die gleiche Wahrscheinlichkeit. Folglich ist die Häufigkeit von  $b^+$  mit  $\frac{I+1}{2I+1}$  und von  $b^-$  mit  $\frac{I}{2I+1}$  zu wichten. Die kohärente Streulänge für jedes Isotop ist mit  $\sigma_c = 4\pi\bar{b}^2$  gegeben, wobei  $\bar{b}$  dem Mittel über die Streulängen bei allen Temperaturen der + und - Populationen entspricht. Analog dazu ist der totale Wirkungsquerschnitt mit  $\sigma_s = 4\pi\bar{b}^2$  gegeben. Wenn ein Isotop keinen Spin hat (z. B.  $^{12}\text{C}$ ), dann ist  $b^2 = \bar{b}^2 = \bar{b}^2$  und inkohärente Streuung tritt nicht auf.

Nur der kohärente Anteil der Streuung enthält Informationen über die räumliche Anordnung im untersuchten Molekül. Der inkohärente Anteil enthält keine strukturelle Information und wird üblicherweise, sofern isotrop, als „Hintergrund“ abgezogen. Für  $^1\text{H}$  ergibt sich eine Streulänge von  $0,374^{-12}$  cm, während sie für  $^2\text{H}$  (Deuterium)  $-0,667^{-12}$  cm beträgt. Andere kleine Elemente haben Streulängen in derselben Größenordnung. Hieraus folgt, dass bei einem System mit organischer Materie (Wasserstoff, Deuterium, Kohlenstoff, Sauerstoff und Phosphor als die häufigsten Elemente) eine Variation des Anteils von Deuterium gegenüber Wasserstoff im Lösungsmittel genutzt werden, um den Kontrast bestimmter anderer Moleküle zu minimieren, wobei ein hoher Wasserstoffanteil üblicherweise einen geringen Kontrast bedeutet (vgl. den folgenden Abschnitt zum Thema „Kontrast“).

### 2.2.5 Kontrast

Für ein System mit zwei Komponenten gilt

$$I(q, \Delta\rho_A, \Delta\rho_B) = \Delta\rho_A^2 I_A(q) + \Delta\rho_A \Delta\rho_B I_{AB}(q) + \Delta\rho_B^2 I_B(q) . \quad (2.22)$$

Hierbei stehen die Indizes  $A$  und  $B$  für die beiden Komponenten und  $\Delta\rho_X = \rho_X - \rho_S$ , wobei  $\rho_X$  der mittleren Streulängendichte der individuellen Komponenten  $A$  und  $B$  entspricht und  $\rho_S$  die mittlere Streulängendichte des Lösungsmittels ist. In Gleichung 2.22 wird vorausgesetzt, dass die Unterschiede in der Streulängendichte zwischen den zwei Komponenten größer ist als die Dichtefluktuationen innerhalb der Komponenten. Mit der Gleichung ist zum Beispiel die Situation hinsichtlich des Kontrastes in einem System aus Lösungsmittel, Lipiden und Proteinen zu beschreiben.

## 2.3 Geometrie der Kratky-Kamera und Konvolution

Ein grundsätzliches Problem von Laborgeräten ist die relativ geringe Intensität von Röntgeneratoren (siehe Abschnitt 2.1.1.2) gegenüber Synchrotrons (vgl. Abschnitt 2.1.1.1). Mit einer geschickten Wahl der Kollimationsgeometrie kann erreicht werden, dass relativ viel Probe bestrahlt wird und so ein größerer Anteil gestreuten Röntgenlichtes erhalten wird im Vergleich zu einer punktförmig bestrahlten Probe. Im Falle der Kratky-Kamera besitzt der Röntgenstrahl ein rechteckiges Profil beim Auftritt auf die Probe. Die Probe selber ist bei Flüssigkeiten in eine Kapillare gefüllt oder bei Feststoffen in einen Zylinder mit rechteckigem Ausschnitt gespannt. Der Nachteil der angesprochenen Kollimation ist jedoch, dass diese im Zuge der Datenauswertung berücksichtigt werden muss, denn sie „verschmiert“ die Streuung ausgehend von der Probe.

Die erste der nun folgenden Abbildungen (Abbildung 2.4) definiert die  $t$ - und die  $x$ -Richtung des Primärstrahls. Hierbei ist  $t$  die Breite des Primärstrahls in der Ausdehnungsrichtung und  $x$  die Höhe des Strahls.

Die angesprochene „Verschmierung“ basiert auf der nicht vernachlässigbaren Breite und Höhe des Strahls: Würde der Strahl punktförmig auf die Probe treffen und der Detektor einen „unendlichen“ Abstand von der Probe haben, so würde alle detektierte Streuung nur von diesem Punkt der Probe ausgehen. Da der Primärstrahl eine Ausdehnung in beide Richtungen besitzt, kommt auch die detektierte Streuung von mehreren Punkten der Probe zugleich. Man kann sich zudem

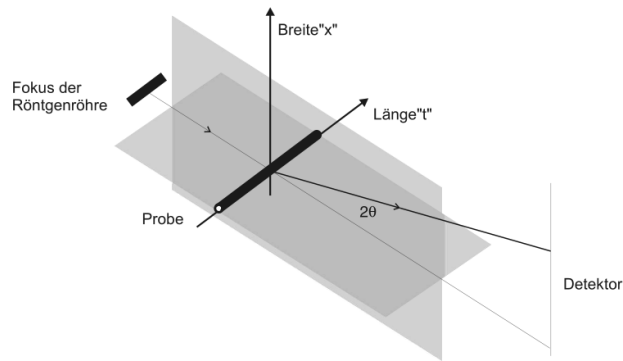


Abbildung 2.4: Skizze Primärstrahlgeometrie der Kratky-Kamera und der Definition der Achsen. Abbildung aus Schmid, 2005.

sehr leicht klarmachen, dass dies auch eine Dämpfung des Signals bedeutet: Gemäß der Betrachtung in Abschnitt 2.2.2 gilt, dass mit der Ausdehnung des Primärstrahls auch mehr und mehr Volumenelemente in der Probe hinzukommen, die  $180^\circ$  außer Phase sind. Außerdem muss ggf. auch noch eine „Verschmierung“ aufgrund der Wellenlänge berücksichtigt werden, für den Fall, dass der Strahl nicht monochromatisch ist. Die folgenden Betrachtungen sind Feigin & Svergun, 1987, entnommen. Es können drei Aspekte unterschieden werden, wobei die Längengrößen in  $q$  (siehe Abschnitt 2.2.2) gegeben sind:

- Abbildung 2.5 skizziert die Verschmierung, die aus der Primärstrahlhöhe resultiert. Eine gemessene oder experimentelle Intensität ( $I(q)_{exp}$ ) ergibt sich hiernach aus der tatsächlichen, nachdem diese mit dem Primärstrahlprofil konvolviert (= gefaltet) wurde:

$$I(q)_{exp} = \int_{-\infty}^{\infty} P(x) \cdot I(q-x) \cdot dx := T_3 . \quad (2.23)$$

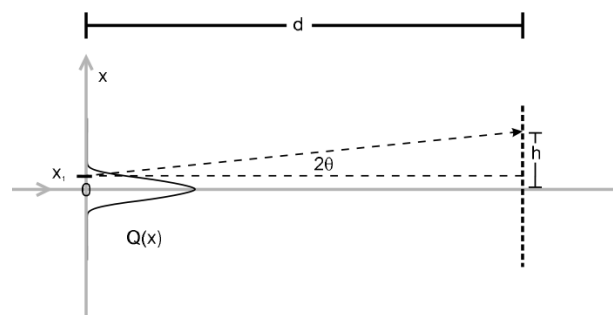


Abbildung 2.5: Skizze der Primärstrahlhöhenverschmierung der Daten. Abbildung aus Schmid, 2005.

- Analog ergibt sich die Breitenverschmierung und Abbildung 2.6 zeigt, warum die Koordinaten hierbei quadratisch addiert werden müssen.

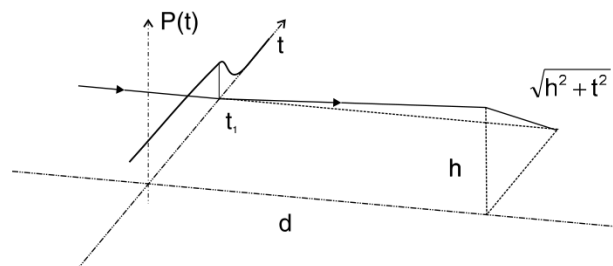


Abbildung 2.6: Skizze der Breitenverschmierung der Daten. Abbildung aus Schmid, 2005.

$$I(q)_{exp} = \int_{-\infty}^{\infty} P(x) \cdot I(\sqrt{q^2 + t^2}) \cdot dt := T_4 . \quad (2.24)$$

- Schließlich gilt für die Wellenlängenverteilung  $P(\lambda)$  um  $\lambda_0$  mit  $\lambda' = \frac{\lambda}{\lambda_0}$  nach Harrown *et al.*, 2007, folgende Verschmierung des Signals

$$I(q)_{exp} = \int_{-\infty}^{\infty} P(\lambda') \cdot I\left(\frac{q}{\lambda'}\right) \cdot dt := T_2 . \quad (2.25)$$

Die hier definierten Konvolutionsintegrale  $T_2$  bis  $T_4$  werden im folgenden Abschnitt verwendet und die Abkürzungen werden in Abbildung 2.7 erläutert.

## 2.4 Indirekte Fouriertransformation

Schauen wir uns Gleichung 2.19 noch einmal an:

$$I(q) = 4\pi \int P(r) \cdot \frac{\sin(qr)}{qr} \cdot dr .$$

Es wird deutlich, dass bei bekannter Abstandsverteilungsfunktion,  $P(r)$ , die Streuintensität,  $I(q)$ , einfach berechnet werden kann. Die umgekehrte Berechnung stellt jedoch in der Praxis eine Hürde dar, da die Intensität nur für einen eingeschränkten Bereich von  $q$  bekannt ist:  $q \rightarrow \infty$  kann nicht gemessen werden und  $q \rightarrow 0$  ist nicht zugänglich, da der Primärstrahl zum Schutz des Detektors blockiert wird. Der Bereich  $q \rightarrow \infty$  kann nach Porod extrapoliert (Porod, 1951) und der Bereich  $q \rightarrow 0$  nach Guinier erschlossen werden (Guinier, 1939). Eine direkte Transformation auf diesem Weg hat den Nachteil relativ großer Unsicherheit für die extrapolierten Bereiche: Man verliert Reproduzierbarkeit, da in die Extrapolation der experimentelle Fehler einfließt. Gleichzeitig bedeutet dies, dass so eine korrekte Fehlerbetrachtung nicht mehr möglich ist.

Die Indirekte Fouriertransformation (IFT nach dem Entwickler O. Glatter; Glatter, 1977) ermöglicht dieses Problem zu umgehen. Sie erlaubt die simultane Berechnung der entschmierten Streukurve und der Abstandsverteilungsfunktion aus einem experimentellen Datensatz. Zur Durchführung wird in einem ersten Schritt eine hypothetische  $P(r)$  geschätzt, zweitens über verschiedene Operationen verschmiert und drittens an die experimentellen Daten angepasst. Abbildung 2.7 zeigt den Ablauf der IFT schematisch.

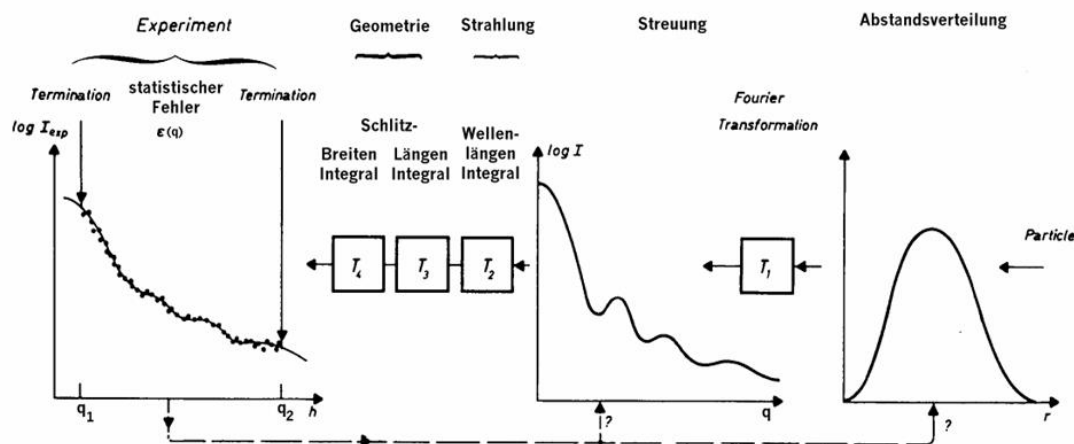


Abbildung 2.7: Schematischer Verlauf einer IFT; modifiziert nach Glatter & Kratky, 1982. Der Term  $T_1$  ist in Gleichung 2.26 erläutert.

Wie weiter oben (siehe 2.3) dargestellt, werden die verschiedenen Effekte der geometrischen Verschmierung mit  $T_3$  (vgl. Gleichung 2.23) und  $T_4$  (vgl. Gleichung 2.24) symbolisiert.  $T_2$  steht für die Wellenlängenverschmierung (vgl. Gleichung 2.25). Die entschmierte, theoretische Streukurve erhält man aus Gleichung 2.19:

$$I(q) = 4\pi \int P(r) \cdot \frac{\sin(qr)}{qr} \cdot dr = T_1 \cdot P(r) , \quad (2.26)$$

wobei  $T_1$  die Transformation abkürzt. Zur Durchführung der IFT muss zunächst ein maximaler Abstand,  $D_{max}$ , der  $P(r)$ , bzw. des untersuchten Moleküls, abgeschätzt werden, für den gilt:

$$P(r) = 0 \quad \forall \quad r \geq D_{max} . \quad (2.27)$$

Die Abstandsverteilungsfunktion wird nun auf dem Intervall  $0 \leq r \leq D_{max}$  durch  $N$  orthogonale Funktionen  $\varphi(r)$  approximiert:

$$P(r) = \sum_{i=1}^N c_i \cdot \varphi_i(r) . \quad (2.28)$$

Es gibt zwei Ansätze dieses algorithmisch nachzuvollziehen: Das Programm ITP verwendet kubische  $B$ -Splines (Glatter, 1980), während das Program GNOM (Svergun, 1992) Stufenfunktionen verwendet. Anstelle einer gesamten  $P(r)$  lassen sich nun die einzelnen Funktionen  $\varphi_i(r)$  mit den Operationen  $T_1$  bis  $T_4$  transformieren. Es gilt

$$\begin{aligned} I(q)_{exp} &= T_1 T_2 T_3 T_4 P(r) \\ &= \sum_{i=1}^N c_i T_1 T_2 T_3 T_4 \varphi_i(r) \\ &= \sum_{i=1}^N c_i \psi_i(q) , \end{aligned}$$

wobei  $\psi_i(q) = T_1 T_2 T_3 T_4 \varphi_i(r)$  definiert wurde. Mit dieser Gleichung lassen sich nun die Koeffizienten  $c_i$  bestimmen, wenn

$$\chi^2 = \sum_{k=1}^M \frac{[I_{exp}(q_k) - \sum_{i=1}^N c_i \psi_i(q_k)]^2}{\sigma_k^2} \quad (2.29)$$

minimiert wird. Hierbei ist  $M$  die Zahl der Messpunkte und  $\sigma_k$  der zugehörige Messfehler. Mit dem Ziel eine stabile Lösung zu erhalten wird nicht  $\chi^2$  minimiert, sondern  $\Gamma = \chi^2 + \alpha_c \cdot N_c$ .  $\alpha_c$  ist ein Wichtungsfaktor und  $N_c$  eine Nebenbedingung. Beispielsweise sind physikalisch sinnvolle Lösungen ausschließlich glatte Kurven. Andere Nebenbedingungen hängen von der erwarteten Form oder fraktalen Dimension der untersuchten Moleküle ab, die ebenfalls Einfluss auf den Verlauf der Kurve haben. Die Realisierung der Nebenbedingungen ist jedoch programm- und algorithmusspezifisch, weshalb auf die beiden in dieser Arbeit verwendeten Programme verwiesen sei (Glatter, 1980; Svergun, 1992).

## 2.5 Der Gyrationradius und die Guinierapproximation

Ein wichtiger Parameter, der sich direkt aus einem Streudatensatz gewinnen lässt, ist der Gyrationradius  $R_g$ . Er ist ein Maß für die Kompaktheit eines Moleküls. Der Gyrationradius ist zunächst einmal definiert durch den Trägheitsmoment  $I$  und die Masse  $M$  eines Moleküls:

$$R_g := \sqrt{\frac{I}{M}} . \quad (2.30)$$

Im Falle eines Polymers wird der  $R_g$  auch spezifiziert über die Zahl  $N$  und Position  $r$  der Atome:

$$R_g^2 := \frac{1}{2N^2} \sum_{i,j} (r_i - r_j)^2 . \quad (2.31)$$

Weil die Zahl der Konformationen eines Polymers quasi unendlich groß ist und zudem noch über die Zeit variieren kann, wird der Gyrationradius in der Polymerphysik als Mittel aller Polymerzustände einer Probe und über die Zeit verstanden:

$$R_g^2 := \frac{1}{N} \left\langle \sum_{k=1}^N (r_k - \bar{r})^2 \right\rangle . \quad (2.32)$$

Diese letzte Definition – welche zu den beiden vorherigen equivalent ist – findet Anwendung im Softwarebereich der Proteinmodellierung. In der von mir geschriebenen und während dieser Dissertation eingesetzten Software (siehe Abschnitt 4.7 auf den Seite 105 ff. des Resultatekapitels) wird jedoch, da eine direkte Umsetzung von Gleichung 2.32 etwas rechenintensiver ist, eine Variation der Gleichung eingesetzt:

$$R_g = \sqrt{\frac{\sum_{k=1}^N f_k \cdot r_k^2}{\sum_{k=1}^N f_k}}, \quad (2.33)$$

hierbei ist  $f_k$  die Streuamplitude der beteiligten Atome bzw. ein willkürlicher Faktor proportional zur Masse. Gleichung 2.33 wird in der eigenen Programmbibliothek zum *rigid body*-Modelling (siehe Resultatekapitel, Abschnitt 4.7) und bei der Berechnung des  $R_g$  aus Elektronendichten einer *single particle*-Rekonstruktion (siehe Material & Methode, Abschnitt 3.6.9) verwendet. Gleichung 2.33 kann überdies in verschiedene Komponenten zerlegt werden. Ist das betrachtete Molekül beispielsweise ein Zylinder und die zentrale Achse ist  $z$ , so ist die radiale Komponente  $R_{xy}$  gegeben durch

$$R_{xy} = \sqrt{\frac{\sum_{k=1}^N f_k \cdot (x_k^2 + y_k^2)}{\sum_{k=1}^N f_k}}, \quad (2.34)$$

wobei  $x_k$  und  $y_k$  zwei der Komponenten der Distanzvektoren im Raum, senkrecht zu  $z$ , sind. Diese Zerlegung ist von besonderem Interesse, wenn die „Verdrillung“ eines zylindrischen Moleküls beschrieben werden soll und ist Teil der beiden oben genannten Softwarepakete. Die Betrachtung stammt von H. Hartmann (Hartmann *et al.*, 2004) und wurde für das Hämocyanin von *Megatoura crenulata* eingeführt..

Für die Kleinwinkelstreuung hat der Gyrationradius eine herausragende Bedeutung, weil er direkt aus den Daten ermittelt wird. Hierzu kann  $I(q)$  zunächst nur in der Nähe von  $q = 0$  betrachtet werden. Entwickelt man  $\frac{\sin(qr)}{qr}$  nach McLaurin (vgl. Feigin & Svergun, 1987)

$$\frac{\sin(qr)}{qr} = 1 - r^2 q^2 / 6 + r^4 q^4 / 120 - \dots, \quad (2.35)$$

so wird klar, dass bei  $q = 0$  für die ersten beiden Terme gilt:

$$I(q) = I(0) \cdot (1 - q^2 R_g^2 / 3), \quad (2.36)$$

wobei

$$I(0) = 4\pi \int_0^{d_{\max}} \gamma(r) r^2 dr \quad (2.37)$$

und

$$R_g^2 = \frac{\int_0^{d_{\max}} \gamma(r) r^4 dr}{2 \cdot \int_0^{d_{\max}} \gamma(r) r^2 dr}, \quad (2.38)$$

mit  $d_{\max}$  dem maximalem Abstand im untersuchten Molekül einer homodispersen Lösung. Der Ausdruck auf der rechten Seite von Gleichung 2.36 entspricht folglich den ersten beiden Termen der McLaurinreihe. Daher gilt, mit einer Genauigkeit proportional zu  $q^4$ , für den Beginn einer Streukurve

$$I(q) = I(0) \cdot e^{-\frac{q R_g^2}{3}}. \quad (2.39)$$

Dies ist die Guiniernäherung (Guinier, 1939). Diese Näherung ist jedoch nur gültig im  $q$ -Wertebereich, für den gilt  $q \cdot R_g \lesssim 1$ . Wenn man Gleichung 2.35 in die Debyegleichung (Gleichung 2.13) einsetzt so ergibt sich

$$I(q) = \int \int \rho(r_1) \rho(r_2) dr_1 dr_2 - \frac{q^2}{6} \int \int \rho(r_1) \rho(r_2) |r_1 - r_2|^2 dr_1 dr_2. \quad (2.40)$$

Ein Vergleich der letzten Gleichung mit Gleichung 2.36 zeigt, dass

$$I(0) = \left| \int \rho(r) dr \right|^2 . \quad (2.41)$$

Folglich ist  $I(0)$  das Quadrat der Gesamtstretlänge des betrachteten Moleküls. Werden nun die Koeffizienten der  $q^2$ -Terme gleichgesetzt, so ergibt sich

$$R_g^2 = \frac{\int_0^\infty P(r) \cdot r^2 \cdot dr}{2 \cdot \int_0^\infty P(r) \cdot dr} . \quad (2.42)$$

Diese letzte Gleichung wird eingesetzt zur präzisen Berechnung des Gyrationradius aus einer gegebenen Abstandsverteilungsfunktion.

# Kapitel 3

## Material und Methoden

### 3.1 Verwendete Chemikalien und Geräte

#### 3.1.1 Chemikalien

Chemikalien wurden im Allgemeinen von der Firma Carl-Roth (Carl-Roth GmbH, Karlsruhe) bezogen. Die Herkunft aller übrigen Chemikalien wird in den entsprechenden Abschnitten explizit angegeben.

„Wasser“ im Kontext dieses Kapitels ist als deionisiertes Wasser, gereinigt mit einer Milliporeanlage (MilliQ *bioce*l-Anlage von der Millipore GmbH, Bad Schwalbach), zu verstehen. Das Wasser hatte einem Leitwiderstand von 18,2  $M\Omega$ .

### 3.2 Gelelektrophoresen

#### 3.2.1 SDS-Gelelektrophorese und Coomassie-Färbung

Die hier verwendete SDS-Gelelektrophorese basiert auf dem System von Laemmli (Laemmli, 1970) in der Modifikation nach See und Jakowski (See & Jackowski, 1989). Die Puffer- und Gelzusammensetzungen zeigt Tabelle 3.1.

Tabelle 3.1: *Lösungen zur Gelelektrophorese*. Für native Gele wurde auf die Zugabe von SDS verzichtet.

<i>Trenngel</i>	125 mM Tris; pH 6.8 bei 20 °C 7,5 % (v/v) Rotiphorese 0,2 % (w/v) SDS; 8.8 bei 20 °C
<i>Sammelgel</i>	375 mM Tris 3,0 % (v/v) Rotiphorese 0,2 % (w/v) SDS; 8.8 bei 20 °C
<i>Laufpuffer</i>	20 mM Tris 150 mM Glycin 0,08 % (w/v) SDS; pH 6.8 bei 20 °C
<i>Probenpuffer</i>	25 % (v/v) <i>Sammelgelpuffer</i> 4 % (w/v) SDS 20 % (v/v) Glycerin 10 % $\beta$ -Mercaptoethanol 1 Spatelspitze Bromphenolblau
<i>Färbelösung</i>	0,5 % (w/v) Coomassie G 250 50 % (v/v) Methanol 10 % (v/v) Essigsäure

Vorgehensweise: Es fanden ausschließlich Minigele (8 x 10 cm Fläche, circa 1 mm Dicke) Verwendung. Die Gele wurden in zwei Stufen gegossen: Erst das *Trenn-* und dann das *Sammelgel*, wobei den Gelen je ca. 45 Minuten Zeit zur Polymerisierung gegeben wurde. Die Polymerisierung erfolgte durch Zugabe von TEMED (4 ‰ (v/v); Tetramethylethyldiamin) und APS (6 ‰ (w/v); Ammoniumpersulfat). Für die Taschen wurden Käämme mit 8–10 Taschen verwendet.

Vor dem Auftragen auf das Gel wurden Proben 1:1 mit dem *Probenpuffer* versetzt und anschließend für 5 Minuten bei 95 °C denaturiert. Die anschließende Elektrophorese erfolgte bei konstanten 110 V. Die Färbung erfolgte für etwa eine Stunde in der *Färbelösung*. Die Entfärbung des Hintergrundes fand im selben Puffer ohne Zugabe des Farbstoffes Coomassie statt.

### 3.2.2 Native Polyacrylamidgele

Die native Polyacrylamidgelelektrophorese (PAGE) dient der Detektion von Proteinen in Gelen, ohne dass diese denaturiert werden und damit einhergehend zu Monomeren zerfallen. Das Protokoll orientiert sich an dem Protokoll zur Gelelektrophorese von See und Jackowski (See & Jackowski, 1989). Vor Einfüllen in die Geltaschen wurden Proben 1:1 mit einem Puffer aus 50 % Glycerin und 0,1 % Bromphenolblau versetzt. Die Elektrophorese erfolgte mit Minigelen (siehe oben) bei 4 °C und einer konstanten Spannung von 100 V. Sofern nicht eine Färbung hinsichtlich der Phenoloxidaseaktivität (im folgenden Abschnitt beschrieben) durchgeführt wurde, wurden die Proben mittels Coomassie angefärbt (siehe Abschnitt 3.2.1).

### 3.2.3 Anfärben von Phenoloxidasen nach Gelelektrophorese

Im Falle der Hämocyanine von *Eurypelma californicum* und *Pandinus imperator* wurde zum Nachweis der Phenoloxidaseaktivität vor Aufnahme von Kleinwinkelstreudaten (siehe Abschnitt 3.9.1) Substrat mit Protein in nativen Gelen (siehe Abschnitt 3.2.2) inkubiert. Auf diese Weise kann enzymatische Aktivität direkt mit Reinheit und Oligomerisierungszustand in Verbindung gebracht werden.

Die Methode basiert auf der Publikation von Nellaiappan (Nellaiappan & Vinayakam, 1993) und ist dieser nachempfunden: Nach der Gelelektrophorese wurden die Gele für 10–15 Minuten in Petrischalen gelegt und mit 100 mM Phosphatpuffer (pH 7,0) inkubiert, um den alkalischen Elektrophoresepuffer zu ersetzen. Zudem wird so eine Autooxidation von Diphenolen und somit eine Hintergrundfärbung vermieden. Anschließend erfolgte eine Inkubation in Färbelösung für etwa 10 Minuten. Die Färbelösung bestand aus 10 ‰ (w/v) SDS (als Aktivator), 0,3 ‰ (w/v) 3-Methyl-2-Benzo-Thiazoenone-Hydraxon (als Verstärker der Färbung) und 10 mM Dopamin (als Substrat) im oben genannten Phosphatpuffer. Die Gele wurden ausschließlich über einen Gelscanner (BioRad, München, Typ *Densitometer GS-800*) dokumentiert, da die Färbung nicht dauerhaft ist.

## 3.3 Aufreinigung der Hämocyanine

In dieser Arbeit wurden ausschließlich Hämocyanine der Arthropoden *Eurypelma californicum* und *Pandinus imperator* selbst aufgereinigt. Die anderen Hämocyanine wurden von anderen Quellen bezogen und wie in den entsprechenden Abschnitten beschrieben aufgereinigt.

### 3.3.1 Verwendete Puffer

Sofern nicht anders angegeben wurden für die Aufreinigung der Hämocyanine standardmäßig die folgenden Puffer verwendet und daher auch mit einer eigenen Kurzbezeichnung versehen:

Tabelle 3.2: Zur Aufreinigung und Lagerung der Hämocyane von *E. californicum* und *P. imperator* verwendete Puffer.

<i>Eurypelma</i> -Stabilisierungspuffer	100	mM Tris; pH 7,8 bei 20 °C
	5	mM MgCl <sub>2</sub>
	5	mM CaCl <sub>2</sub>
<i>Pandinus</i> -Stabilisierungspuffer	100	mM Tris; pH 7,8 bei 20 °C
	10	mM MgCl <sub>2</sub>
	10	mM CaCl <sub>2</sub>

Nach ihrer Herstellung wurden die Puffer gefiltert (Carl Roth, Karlsruhe, Typ *A009.1*) und mit dem Vakuum einer Membranpumpe (vacuubrand, Wertheim, Typ *MZ 2C*) entgast.

### 3.3.2 Tierhaltung

Die Tiere (*E. californicum* und *P. imperator*) wurden einzeln in verschlossenen Plastikwannen gehalten, deren Boden mit Sand und Mulch bedeckt war. In den Wannen befinden sich zusätzlich kleine Pflanztöpfe aus Ton, um den Tieren eine Rückzugsmöglichkeit zu bieten. Den Tieren wurden zweimal pro Woche Wasser gegeben und alle zwei Wochen wurde ein Heimchen gefüttert. Der Licht-Dunkel-Wechsel erfolgte automatisch im 12-Stunden Rhythmus.

### 3.3.3 Gewinnung und Aufreinigung

Die Hämolymphe aus *E. californicum* und *P. imperator* wurde durch Punktion des Perikards erhalten. Die austretende Hämolymphe wurde mit einer Pasteurpipette aufgenommen und in ein gekühltes Reaktionsgefäß gegeben, in welches etwa 100  $\mu\text{l}$  des entsprechenden Puffers (siehe 3.3.1) vorgelegt waren. Unmittelbar vor Auftragung auf die Gelfiltrationssäule wurde die Hämolymphe 15 Minuten bei 15000 x g und 4 °C zentrifugiert (Tischzentrifuge von Hettich Zentrifugen, Tuttlingen, Typ *Universal 16 R*). Die Aufreinigung des Hämocyans erfolgte mittels Gelfiltration (GE Healthcare, Säulenmaterial Typ *Sephacryl S-300* und Säule Typ *XK26/60* mit einer FPLC von BioRad, München, Typ *BioLogic DuoFlow*) bei 4 °C. Die Flussrate betrug 1,3 mL/Minute. Es wurden Fraktionen von 3 mL Größe gesammelt. Die Fraktionen wurden so zusammengeführt, dass ausschließlich Hämocyanin in der Lösung einhalten war.

Abschließend wurden die zusammengeführten Fraktionen mit Membranfiltern („Vivaspin“-Membranfilter der Sartorius AG, Göttingen, Typ *VS2001* mit einer Molekulargewichtssperre von 10 kDa) bei knapp 40000 x g konzentriert. Das Konzentrat wurde in ein 1,5 mL Reaktionsgefäß überführt und nochmals 15 Minuten bei 22000 x g und 4 °C zentrifugiert, um ggf. co-konzentrierte Verunreinigungen wie Bakterien und Staub zu pelletieren. Der Überstand wurde aufbewahrt und für weitere Experimente aliquotiert. Die Konzentration des Hämocyans wurde über die optische Dichte bei 280 nm bestimmt. Hierbei wurden spezifische Extinktionskoeffizienten von  $\varepsilon = 1.04 \text{ mL} \cdot \text{mg}^{-1} \cdot \text{cm}^{-1}$  für das Hämocyanin von *E. californicum* (Loewe & Linzen, 1973; Hübler, 2001) und  $\varepsilon = 1.1 \text{ mL} \cdot \text{mg}^{-1} \cdot \text{cm}^{-1}$  für das Hämocyanin von *P. imperator* (Waschbüsch, 2004) angenommen.

### 3.3.4 Kontrolle der Integrität nach Bestrahlung

Um sicherzustellen, dass die Arthropodenhämocyane durch die Behandlung in der Kratky-Kamera (siehe 3.6) keinen Schaden nahmen, wurde Proben von *Eurypelma californicum*-Hämocyanin und *Pandinus imperator*-Hämocyanin mit einer Konzentration von ca. 5 mg/mL folgendermaßen behandelt:

- Ein Teil blieb unbehandelt und diente als Negativkontrolle.
- Ein weiteres Aliquot (ca. 50  $\mu\text{l}$ ) wurde mehrfach durch die Kapillare, die als Probenhalter in der Kratky-Kamera dient, hin und her gepresst und anschließend aufgefangen.

- Ein drittes Aliquot wurde in die Kapillare gefüllt und für 4 Stunden bestrahlt. Dies entspricht auch der Strahldauer während eines gewöhnlichen Experimentes mit der Kratky-Kamera.

Abschließend wurden alle drei Aliquots auf ein natives Gel mit Calcium- und Magnesiumionen (siehe 3.2.2) gegeben.

### 3.3.5 Deoxygenierung

Die Deoxygenierung der Hämocyanine fand statt wie geschildert (Pairet, 2007). Proteinlösungen wurden innerhalb eines „*Atmosbags*“ (Sigma-Aldrich, München; Typ *Z530220*) auf ein Uhrglas gegeben. Der „*Atmosbag*“ wurde wiederholt mit Stickstoff geflutet bis ein Sauerstoffpartialdruck  $\leq 0,3\%$ , beziehungsweise  $\leq 2,3$  Torr, (kontrolliert über ein Messgerät von Greisinger Electronic, Typ *GMH 3690 Gl*) erreicht war. Dieser Stickstoffatmosphäre wurde die Hämocyaninlösung für 90 Minuten ausgesetzt.

Um die Deoxygenierung der Hämocyanine zu gewährleisten und in einer Konformation zu arretieren (Decker & Sterner, 1990), war der pH der Lösungen von 7,8 auf 7,1 erniedrigt. Der Erfolg der Deoxygenierung wurde mittels UV/Vis-Spektroskopie (Varian Deutschland, Darmstadt, Typ *Cary 300 Bio*) verfolgt: Nach der Deoxygenierung im *Atmosbag* wurde ein Teil des Hämocyanins in eine verschließbare Küvette (Hellma, Müllheim, Typ *104-10-40*) gefüllt und die Absorption von 200 – 700 nm vermessen. Ein fehlender *Peak* bei 340 nm zeigte die erfolgreiche Deoxygenierung an. Hierauf wurde der Küvettendeckel entfernt und die Küvette mit Luft beblasen, wieder mit Parafilm abgedeckt und invertiert. Der Absorptionspeak bei 340 nm des derart re-oxygenierten Proteins hatte sich mit der Kontrolle des unbehandelten Hämocyanins zu decken, andernfalls wurde die begonnene SAXS-Messung abgebrochen.

## 3.4 Lichtstreuung

### 3.4.1 Statische Lichtstreuung

Molekulargewicht und Homogenität einiger Hämocyaninproben wurden mittels der Vielwinkellichtstreuung (oder MALLS = **M**ulti **A**nge **L**aser **L**ight **S**cattering) an einem Lichtstreuophotometer kontrolliert. Die Messungen wurden ausschließlich von Prof. Dr. E. Jaenicke durchgeführt und ausgewertet.

Das Lichtstreuophotometer (Wyatt Technology, Santa Barbara, CA; Typ *Dawn DSP Laser Lichtstreuophotometer*) verfügt über einen He/Ne-Laser (Wellenlänge  $\lambda = 632,8$  nm). Proben werden in die Anlage injiziert, ggf. über eine Gelfiltration (GE Healthcare, NJ, USA; Superose 6 HR 10/30) getrennt und über drei Detektoren geleitet (HPLC-Pumpe von Hewlett Packard; Typ *HP 1100*): Ein UV/Vis-Detektor (Bio-Rad, USA, Typ *UV-1806*), der Detektor zur Vielwinkellichtstreuung und ein Refraktometer. Zur Vermeidung von Luftblasen in den Detektoren verfügt die Anlage zusätzlich über einen *Online*-Entgaser (ERC Incorporation, Tokio, Japan, Typ *ERC-3215*), der vor die Gelfiltrationssäule geschaltet ist. Die Auswertung der Daten erfolgte mit der Software ASTRA (Version 4.90). Für das Brechungsindexinkrement  $dn/dc$  wurde für die Hämocyanine von *E. californicum* und *P. imperator* ein Wert von 0,1827 mL/g angenommen (Waschbüsch, 2004).

### 3.4.2 Dynamische Lichtstreuung

Dynamische Lichtstreuung (auch Photonenkorrelationsspektroskopie) dient zur Größenbestimmung von (meist kolloiden) Partikeln und wird vor allem in der Rheometrie eingesetzt. Bei der Dynamischen Lichtstreuung wird die zeitliche Fluktuation des Streulichtes verfolgt. Diese gibt Aufschluss über die Partikelgrößen, da die Diffusionskonstante eines Teilchens bei einer bestimmten Temperatur von Masse und Form des Teilchens abhängt. Messungen erfolgten bei 20 °C im Zetasizergerät (Malvern, Worcestershire, UK; Typ *Nano - ZS*) mit Niedrigvolumenküvetten (Wyatt Technology, Santa Barbara, CA; Typ *DTS 2145*). Die Auswertung erfolgte mit der Software DTS-Nano 3.00 unter der vereinfachenden Annahme, dass sich die Partikel wie Kugeln verhalten.

## 3.5 Analytische Ultrazentrifugation

Analytische Ultrazentrifugation wurde ausschließlich unter Verwendung der Absorptionsoptik der Zentrifuge (Beckmann Coulter, Fullerton USA; Gerät Optima XL-I) durchgeführt. Die Proteinkonzentration wurde so eingestellt, dass die  $OD_{280\text{nm}}$  ein Wert von ca. 0,5 nicht überstieg. Die Auswertung erfolgte mit dem Programm Ultrascan (Version 9, Demeler, 2005), mit dem vorab auch die optimale Zentrifugationsgeschwindigkeit für einen Sedimentationslauf über eine Simulation ermittelt werden konnte. Das Programm besitzt die Möglichkeit die Methode von van Holde & Weischet (van Holde & Weischet, 1978) zur Auswertung von Sedimentationsläufen zu nutzen (Demeler *et al.*, 1997; Demeler & Saber, 1998). Für diese Arbeit wurden ausschließlich Analysen der Sedimentationsgeschwindigkeit durchgeführt.

## 3.6 Messungen an der Kratky-Kamera

### 3.6.1 Aufbau der Kratky-Kamera

Die Kratky-Kamera (Anton Paar GmbH, Graz, Österreich; Typ KKK; für die Theorie vgl. Abschnitt 2.3 auf Seite 22) wurde unverändert nach Hönig, 2004, übernommen. Im Verlauf der Arbeit wurde eine temperierbare Zelle (Anton Paar GmbH, Graz, Österreich; Typ *TCU 50*) installiert (siehe Abschnitt 3.6.1.3 in diesem Kapitel). Die Position des „ES“-Schlitzes wurde nachkalibriert (siehe Abschnitt 3.6.3 in diesem Kapitel), so dass eine optimale Kollimation gewährleistet war.

#### 3.6.1.1 Der Röntngenerator

Der Röntngenerator (Rich Seiffert GmbH, Ahrensberg; Typ *ISO-Debyelex 3003*) wurde mit einer konstanten Leistung von 1,05 kW (30 mA x 35 kV) betrieben und während der gesamten Betriebsdauer mit Wasser auf einer Temperatur von  $21 \pm 1$  °C gehalten, um eine möglichst konstante Röntgenintensität zu erreichen. Der Generator wurde stets stufenweise auf die Ausgangsleistung gestellt, um das Gerät zu schonen.

#### 3.6.1.2 Der Detektor

Die Anlage wurde mit einem **Positionsensitivem Detektor** (M. Braun GmbH, Ahrensburg, inzwischen Hecus GmbH, Graz, Österreich; Typ *PSD 50*) betrieben. Die Streudaten wurden in 2048 Kanälen aufgenommen, das Pulshöhenspektrum wurde in 1024 Kanäle unterteilt. Um Ausreißer, die u. a. auf inelastischer Streuung beruhen können, nicht mit aufzuzeichnen, wurden die unteren 200 und die oberen 50 Kanäle nicht berücksichtigt.

Der Detektor wurde vor dem Betrieb ausgiebig mit Zählgas (90 % Argon / 10 % Methan) gespült und während des Betriebs mit konstanten  $\sim 7,5$  bar geflutet. Die am Zähldraht angelegte Hochspannung betrug  $\sim 3,265$  kV. Da die freie Weglänge der Ionen mit dem Druck und damit indirekt mit der Temperatur variierte, wurden die Parameter jeden Tag vor Beginn der Messung so optimiert, dass das Pulshöhenspektrum der Wanderspaltmessungen (siehe Abschnitt 3.6.2) ein Maximum bei Kanal 450–500 aufwies.

Gesteuert wurde der Detektor mit der Software ASA (Version 2.3), ebenfalls von Hecus aus Graz.

#### 3.6.1.3 Aufbau von Zusatzgeräten

Im Laufe dieser Doktorarbeit wurde in die Kratky-Kamera ein temperierbarer Probenhalter (Anton Paar GmbH, Graz, Österreich; Typ *TCU 50*) eingebaut. Dieser ermöglicht es mittels eines Peltier-Elements die Kapillare mit der Probe auf eine Temperatur zwischen  $-30$  °C und  $120$  °C einzustellen. Damit ist es möglich empfindliche Proben während der Messung zu kühlen oder auch Experimente zur Hitzedenaturierung durchzuführen und so strukturelle Veränderungen der Proteine zu verfolgen. Das Peltierelement wurde mit einer Wasserkühlung bei einem Durchfluss von ca. 500 mL pro Minute und einer Wassertemperatur von  $10$  °C gegengekühlt.

### 3.6.2 Vorgehensweise

Alle Röntgenkleinwinkelstreuungsmessungen wurden an der unter Abschnitt 3.6.1 beschriebenen Kratky-Kamera durchgeführt, deren Betriebsweise in diesem Abschnitt besprochen wird. Jede Messung an der Kratky-Kamera gliederte sich in folgende Arbeitsstufen:

**Probenpräparation:** Die Proben wurden vorbereitet, wie in den Abschnitten zu den einzelnen Proteinen beschrieben und anschließend mit einer Spritze über einem speziell gefertigten Adapter in die Probenkapillare gefüllt (siehe Abbildung 3.1). Der Adapter wurde von der Werkstatt der Biologie durch W. Kern nach meiner Vorlage gefertigt und erlaubte, bei einer Oberflächenspannung, die in etwa der des Wassers entspricht, das Einfüllen bis zu 50  $\mu\text{l}$  kleiner Probenmengen. Eine verringerte Oberflächenspannung erforderte größere Probenmengen (mindestens etwa 100  $\mu\text{l}$ ). Der Probenhalter mit Kapillare wurde in der Kamera befestigt.

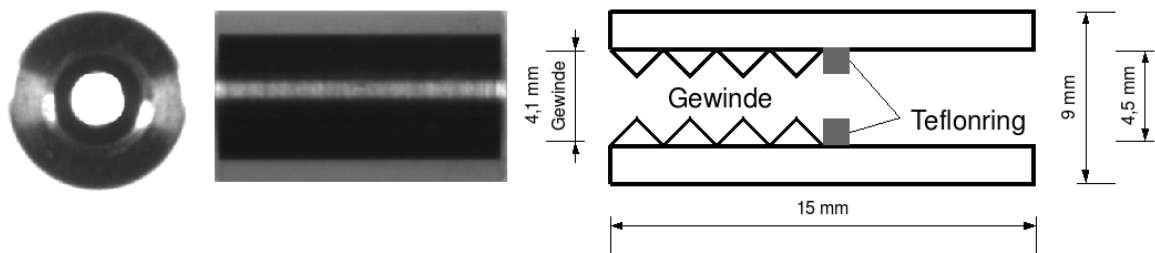


Abbildung 3.1: *Adapter zum Befüllen der Probenkapillare.* Von links nach rechts: Aufsicht, Seitenansicht und Schemazeichnung. Der Adapter wird auf eine Seite der Kapillare geschraubt und die eine Spritze mit Probe kann auf der anderen Seite aufgesetzt werden. Der Teflonring verhindert ein seitliches Austreten der Probe bei leichtem Druck.

**Messungen:** Messungen fanden üblicherweise in der Abfolge *Puffermessung*  $\rightarrow$  *Probenmessungen*  $\rightarrow$  *Puffermessung* statt. Hierdurch sollten insbesondere Hystereseeffekte in der Detektorsensibilität, bedingt durch Druckschwankungen im Zählgas über den Messverlauf, ausgeglichen werden. Sollten Messungen an deoxygenierten Hämocyaninen durchgeführt werden, so musste die Abfolge zwangsläufig auf *Probenmessungen*  $\rightarrow$  *Puffermessung* angepasst werden, da die Deoxygenierung zeitaufwendig war und die Probenkapillare währenddessen im sauerstofffreien Medium aufbewahrt werden musste. Die Gesamtdauer der *Puffermessungen* entsprach immer der Dauer der *Probenmessungen*.

Die einzelnen Messphasen wurden in 0,5 – 1 h lange Einzelmessungen unterteilt. Zwischen diesen Einzelmessungen, die zugleich auch Teildatensätze (siehe Abschnitt 3.6.4.1) darstellen, befanden sich jeweils 12 Wanderspaltmessungen von je 20 Sekunden Länge. Bei diesen Messungen wurde eine Blende mit einem Längsspalt in die Kamera, kurz hinter der Röntgenröhre, eingebracht. Der Längsspalt wird von einem Schrittmotor angetrieben, so dass innerhalb der 20 Sekunden der Spalt einmal vor dem Strahl hin- und herwandert. Auf diese Weise kann die Position des Primärstrahls und die relative Intensität festgestellt werden. Beide können sich im Laufe eines Messtages durch thermische Veränderungen (Druck des Zählgases (s. o.) und Verziehung des Kameragehäuses) verschieben. Durch die Wanderspaltmessungen ist sichergestellt, dass derartige Änderungen berücksichtigt werden können.

**Probenentnahme:** Proben konnten, beispielsweise zur anschließenden Konzentrationsbestimmung, so aus der Kapillare entnommen werden, dass das gesamte Volumen in einem Mikroreaktionsgefäß aufgefangen wurde: Hierzu wurde zunächst der Deckel an der Einfüllöffnung ab- und der Einfülladapter dort angeschraubt. Eine leere Spritze mit einem kleinen Luftvolumen konnte nun angebracht werden, so dass mit der Spritze ein Unterdruck aufgebaut wurde. Erst dann wurde der Deckel am anderen Ende abgeschraubt, so dass die Probe in die Spritze gesaugt wurde.

### 3.6.3 Wartung und Kalibrierung

Die einzelnen Bestandteile des Kamerasystems wurden in regelmäßigen Abständen gemäß Anleitung gewartet. Besondere Hinweise sind in Anhang B angegeben.

Die Kratky-Kamera wurde, wie in der Diplomarbeit von D. Hönig (Hönig, 2004) installiert, übernommen. Anfänglich wurden auch alle Einstellungsparameter übernommen. Im Laufe dieser Arbeit mussten verschiedene Parameter nachkalibriert werden. Hierzu wurde von mir das Programm „**Calibrate**“ mit einer Kommandozeilenbenutzeroberfläche geschrieben<sup>1</sup>. „**Calibrate**“ benötigt einen Pythoninterpreter (van Rossum, 1998) in der Version 2.4-2.6 sowie die Module **scipy** (Jones *et al.*, 2007) und **matplotlib** (Hunter, 2007). Folgende Kalibrierungen wurden mit diesem Programm durchgeführt:

- Ermittlung des Konversionsfaktor „Detektorkanäle zu Millimetern“.
- Feststellen der Primärstrahlposition (auch implementiert im Programm zum Export nach EXCEL, siehe Abschnitt 3.6.4.1).
- Die Einstellung des „ES“-Schlitzes (Teil des Kollimationssystems).

Zusätzlich wurde Silberbehenat dazu verwendet, über die Braggreflexe der Kristalle mit bekanntem Kristallgitter den Abstand vom Detektor zur Probe präzise zu berechnen. Die Dekonvolution dieser Daten wurde mit dem Programm **swax2xy** (Version 0.33; Unruh *et al.*, 1999) durchgeführt.

### 3.6.4 Auswertung der Streudaten

Die Auswertung der Streudaten gliedert sich in Aufbereiten und Zusammenfassen der Streudaten, die Korrektur für die optischen Effekte und alle weitergehenden Interpretationen der Daten (in den folgenden Unterkapiteln).

#### 3.6.4.1 Auswertung der Primärdaten

Die Teilmessungen für einen Datensatz (zu den Teilmessungen siehe Abschnitt 3.6.2) wurden für die Schwankungen der Primärstrahlintensität und auf eine Stunde Messzeit und Berücksichtigung der Detektortotzeit normiert. Anschließend erfolgte die Mittlung der Datensätze der Proben und des Puffers. Die Differenz der beiden Datensätze wurde gebildet, so dass man das Streusignal des Proteins erhält:

$$I_{\text{Protein}} = I_{\text{Protein+Puffer}} - I_{\text{Puffer}} \quad (3.1)$$

Die Daten aus der **ASA**-Software (Version 2.3) wurden mit der Software **EXCEL** (Microsoft Corp., Redmont, WA; Version: EXCEL2003, SP3) weiter verarbeitet. Der Export zum EXCEL95-Format wurde mit dem selbstgeschriebenen Programm **PreToExcel**<sup>2</sup> durchgeführt (siehe Anhang A.6, Seite 169). Neben dem Pufferabzug erfolgte mit **EXCEL** eine Mittelung der Teildatensätze eines Tages und die Zusammenfassung der Datensätze eines Versuches zu einem Gesamtdatensatz, jeweils unter Berücksichtigung der Fehlerfortpflanzung<sup>3</sup>. Außerdem wurden die apparenten Gyrationradien mittels Guinierplots und eine erste Betrachtung der Streudaten in **EXCEL** durchgeführt.

#### 3.6.4.2 Entschmierung der Daten

Die Dekonvolution oder Entschmierung der Daten wurde mit Hilfe der Programme **GNOM** (Svergun, 1992) bzw. **ITP** (Glatter, 1977) durchgeführt. Zumeist wurde **GNOM** verwendet, da dieses komfortabler zu verwenden ist. Da **GNOM** für das Primärstrahlprofil lediglich die Approximation eines Trapezes verwendet, während **ITP** das gemessene Trapez einliest, wurden die Entschmierungen

<sup>1</sup>Das Programm ist auf der CD, die der Dissertation beiligt, dokumentiert.

<sup>2</sup>Das Programm stellt eine wesentliche Arbeitserleichterung dar und ist dokumentiert auf der CD mit den Daten der Dissertation zu finden.

<sup>3</sup>Beispieltabellen mit den Datensätzen befinden sich ebenfalls auf der CD mit den Daten der Dissertation.

stehts auch mit Hilfe von ITP kontrolliert. Die Standardparameter für den Gebrauch der Entschmierungsprogramme für Daten von der Kratky-Kamera des Institutes sind in den folgenden Tabellen aufgelistet. Für das Programm GNOM wurden diese Parameter zusammen mit B. Pairet gewonnen (Pairet, 2007).

Tabelle 3.3: Die Standard-Parameter für GNOM.

Parameter	q [ $\text{\AA}^{-1}$ ]	Parameter	q [ $\text{\AA}^{-1}$ ]
AH1	0,5352	LH1	0,248
AW1	0,002	LW1	0,001

Tabelle 3.4: Die Standard-Parameter für ITP.

Kontrollparameter	Parameter	Beschreibung	Wert
	ERREL		-1.0
	IFB		0
	ASQ		0.0
	REX		0.0
Konstanten	Parameter	Beschreibung	Wert
	ITYP	Entschmierungseinstellung	0
	ICO		0
	IDEL		0
	NFORM		0
	NTH		0
Zahl der Datenpunkte	Parameter	Beschreibung	Wert
	NU	erste Punkte von *.ITP	0
	NO	letzte Datenpunkte von *.ITP	290
	NHI	Intervalle zwischen q=0 und dem ersten Datenpunkt	10
	NTH	Zahl der Interpolationspunkte	0
Realraumkonstanten	Parameter	Beschreibung	Wert
	TRANS	Transformationsfaktor für die Winkelgrößen	1.0
	CSQRT	Normalisierungsfaktor für die Fourierintegrale	1.0
	DCON	Inkrement für die Dekonvolution im reziproken Raum	4.0
Lagrangefaktor	Parameter	Beschreibung	Wert
	ILE	Schritte zur Berechnung des Stabilisierungsfaktors	19
	SLAM	Obergrenze für den Stabilisierungsfaktors	6.5
	DLAM	Dekrement für die Berechnung des Stabilisierungsfaktors	0.25

Das Programm SPLOT (eigenes Programm) bietet sowohl für GNOM als auch für ITP Datenmasken an (siehe Resultate, Abschnitt 4.8).

### 3.6.5 Hitzedenaturierung an der Kratky-Kamera

Die Installation der temperierbaren Probenzelle ermöglicht es nicht nur Proben während der Messung effektiv zu kühlen (siehe z. B. Abschnitt 3.10.5), sondern auch die Hitzedenaturierung eines Proteins in Lösung strukturell zu verfolgen.

Lyophilisiertes Lysozym wurde in Phosphatpuffer (50 mM, pH 7.0 bei 20 °C) in einer finalen Konzentration von 10 mg/mL aufgenommen. Um die ungefähre Schmelztemperatur des Proteins zu ermitteln, wurde eine DSC-Messung durchgeführt (siehe Abschnitt 3.11)<sup>4</sup>. Zunächst wurde 1 h bei 20 °C bestrahlt, anschließend je eine halbe Stunde bei 30, 40, 50, 60 und 70 °C. Abschließend wurde nochmals 1 h bei 80 °C bestrahlt. Puffermessungen wurde je eine halbe Stunde vor und nach den Probenmessungen durchgeführt.

Diese Vorgehensweise geht zu Lasten der Genauigkeit, ermöglicht es jedoch an einem Tag einen vollständigen Teildatensatz einer Denaturierung zu gewinnen (siehe Abschnitt 5.4 im Diskussionskapitel). Für einen vollständigen Datensatz einer Denaturierung wurden, je nach Fragestellung, zwei alternative Abläufe gewählt:

1. Entweder wurde ausschließlich unterhalb und überhalb der Schmelztemperatur bestrahlt: Dies ermöglicht es, für je eine Proteinkonzentration, zwei Teildatensatz (nativ und denaturiert) innerhalb eines Tages zu erhalten.
2. Oder es wurde in kleinen Schritten, wie oben beschrieben, eine Temperaturreihe gemessen: Dies ermöglichte es, die Dynamik einer Denaturierung zu verfolgen.

### 3.6.6 Kratky-Plots zum Verfolgen der Denaturierung

Der sogenannte *Kratky-Plot* ist die Auftragung von  $I_{(q)} \cdot q^2$  über  $q$ . Sie ist historisch begründet, denn die Integration unter der Kurve ergibt die sogenannte Porodinvariante  $Q$ , welche proportional dem Volumen eines Moleküls ist (Glatter & Kratky, 1982).  $Q$  ergibt sich aus dem Integral

$$Q = \int_0^{\infty} I_{(q)} \cdot q^2 \cdot dr = \int_0^{q^*} I_{(q)} \cdot q^2 \cdot dr + \frac{k_1}{q^*}. \quad (3.2)$$

Da es offensichtlich nicht möglich ist die gestreute Intensität bei unendlichem Winkel zu messen, wird die Integration zu einem relativ großen Streuvektor  $q^*$  aufgetragen und numerisch nach der Simpsonschen Regel (Horwitz, 2001) ausgeführt. Der verbleibende Rest der Kurve wird analytisch integriert (Glatter & Kratky, 1982).

Gemäß Porod (Porod, 1951) oszilliert das Ende der Kurve um  $k_1 q^{-4}$ . Durch einen Plot von  $I_{(q)}$  über  $q^4$  (dem sogenannten Porod-Plot; Glatter & Kratky, 1982) ist  $k_1$  zu bestimmen: Der Plot sollte in einer Geraden auslaufen, die regelmäßig um  $k_1$  oszilliert. Der Anteil der analytischen Integration beträgt im Allgemeinen weniger als 10 %. Für Partikel mit homogener Elektronendichte – wie Proteinen – erlaubt diese Methode eine Bestimmung des Volumens und bei bekannter Dichte damit auch der Masse.

Dieses ist in der vorliegenden Arbeit aufgrund der Intensitätsschwankungen der Röntgenanlage des Institutes nicht möglich gewesen, da die Methode entweder eine absolute Intensität oder zumindest eine sehr hohe Reproduzierbarkeit der Intensitätswerte<sup>5</sup> erfordert. Eine Massenbestimmung wurde dennoch mit dem Kratky-Plot versucht (siehe 3.6.7).

Außerdem wird aus Gleichung 3.2 deutlich, dass ein Kratky-Plot den Unterschied zwischen nativem und denaturiertem Protein gut visualisieren kann: Eine Denaturierung kann beispielsweise infolge einer innermolekularen Aggregation mit einer Verdichtung des Proteins einhergehen. Dies

<sup>4</sup>Die Schmelztemperatur ist nur als „ungefähre“ Schmelztemperatur zu verstehen, da die DSC-Messung mit ca. 1 mg/mL durchgeführt wurde, während die SAXS-Messung bei einer Konzentration von 10 mg/mL stattfand. Beide Konzentrationen waren durch die jeweilige Methode bedingt.

<sup>5</sup>In diesem Fall kann die Kalibrierung auch durch externe Standards durchgeführt werden.

bedeutet eine relative Erhöhung der Streuintensität bei großen Werten von  $q$ ; ein Effekt der im Kratky-Plot überhöht wird (Segel *et al.*, 1998; Segel *et al.*, 1999; Lipfert & Doniach, 2007). Zur Kontrolle während der SDS-Inkubation in der Kratky-Kamera wurde genau dieses Phänomen genutzt (siehe Abschnitt 3.9.1).

### 3.6.7 Massenbestimmung mittels Kratky-Plots

Angeregt durch einen Jahresbericht des japanischen Synchrotrons *Photon Factory* wurde ein Versuch unternommen, die Möglichkeit zu klären, ob trotz der Intensitätsschwankungen der instituteigenen Kleinwinkelstreuanlage eine Bestimmung des Molekulargewichts von Proteinen durchführbar ist. Eine russisch-japanische Arbeitsgruppe (Semisotnov *et al.*, 2002) hat im erwähnten Jahresbericht zeigen können, dass die Position des Maximums im Kratky-Plot (siehe Abschnitt 3.6.6) umgekehrt proportional zur Masse eines Proteins ist. Für eine Kugel gilt:

$$I_{(q)} = (\rho \cdot V \cdot \frac{3}{q^3} \cdot (\sin(q \cdot r) - (q \cdot r \cdot \cos(q \cdot r))))^2, \quad (3.3)$$

wobei  $r$  dem Radius,  $V$  dem Volumen und  $\rho$  der Elektronendichte der Kugel entsprechen. Folglich gilt bei Auftragung von  $I_{(q)} \cdot q^2$  gegen  $q$ , dass die Position des Maximums ausgedrückt in  $q$  umgekehrt proportional zum Radius der Kugel ist:

$$\max_q \propto \frac{1}{r}. \quad (3.4)$$

Die Linearisierung durch doppelte Logarithmierung (eigentlich: Anwendung des *logarithmus naturalis*) ergibt also eine zu größerem Molekulargewicht abfallende Gerade. Da dies jedoch in strikter Weise nur für Kugeln gilt, ist eine präzise Bestimmung des Molekulargewichtes von Proteinen mittels des Kratky-Plots ausschließlich bei globulären Proteinen möglich. Bei Proteinen mit Hohlräumen anderer Dichte, mit Komponenten anderer Dichte (z.B. Zucker, Lipid) oder großer Abweichung von der Kugelform gibt es einen systematisch falschen Beitrag, der nicht einzuschätzen ist. Sofern eine Kalibrierung mit bekannten Proteinen durchgeführt und eine Regressionsgerade errechnet ist, ergibt sich das Molekulargewicht eines unbekanntes Proteins somit aus

$$M_{wt} = e^{(\text{Achsenabschnitt} + \text{Steigung} \cdot \ln(\max_q))} \quad (3.5)$$

(vgl. Abschnitte 4.5 und 5.5).

### 3.6.8 Berechnung theoretischer Streukurven aus atomaren Proteinmodellen

Die Berechnung theoretischer Streukurven aus atomaren Proteinmodellen erfolgte mit dem Programm **CRY SOL** (Svergun *et al.*, 1995). **CRY SOL** berücksichtigt die Atomformfaktoren der einzelnen Atome eines Proteins, während ein Vorläuferprogramm (Pavlov & Fedorov, 1983) Approximationen für kleine würfelförmige Volumina durchführte. **CRY SOL** nutzt eine Multipolentwicklung (Stuhrmann, 1970) und ist daher ebenfalls nur approximativ. Insbesondere bei größeren Strukturen (beispielsweise Hämocyanine ab 430 kDa, was einem hexameren Hämocyanin von Arthropoden entspricht) ist es sehr schwer akkurate Kurven zu errechnen, da die Zahl der Oberschwingungen mit der Ordnung des Fibonaccigitters ab dieser Größe nicht oder nur schwer in Einklang zu bringen ist: Die entsprechenden Kurven sehen „eckig“ aus und können dann nicht mehr verwendet werden. In diesem Fall wurde auf die Berechnung mit der Debye-Gleichung (siehe Theorie Abschnitt 2.2.2) ausgewichen. Der Vorteil von **CRY SOL** jedoch ist die Berücksichtigung der Lösungsmittelumgebung eines Proteins, die in die Berechnung mit der Debye-Gleichung nicht eingeht.

### 3.6.9 Berechnung theoretischer Streukurven von 3D-Modellen aus der Elektronenmikroskopie

Die Abstandsverteilungsfunktion eines Moleküls  $P(r)$  ist aufzufassen als ein Histogramm der Zahl der Abstände innerhalb eines Moleküls über die Abstände und folgendermaßen definiert:

$$P(r) = r^2 \cdot (\rho(\vec{r}) * \rho(\vec{r})). \quad (3.6)$$

Die Abstandsverteilungsfunktion ergibt sich folglich aus dem Quadrat der Abstände multipliziert mit der Patterson-Funktion der Elektronendichte ( $\rho(\vec{r})$  steht für die Elektronendichte zwischen zwei Punkten). Ist die Abstandsverteilungsfunktion bekannt, so ist die Streukurve des Moleküls durch folgende Transformation zu errechnen:

$$I(q) = 4\pi \cdot \int_0^{D_{\max}} P(r) \cdot \frac{\sin(qr)}{qr} \cdot dr . \quad (3.7)$$

Hierbei ist  $q$  definiert als der Streuvektor

$$q \equiv \frac{4 \cdot \pi}{\lambda} \cdot \sin(\theta) ; \quad (3.8)$$

$\theta$  ist der halbe Streuwinkel. Gleichung 3.7 ist analog der *Debye-Gleichung* (siehe Theorie, Abschnitt 2.2.2), welche den Zusammenhang zwischen den abstandsabhängigen Formfaktoren ( $f$ ) und der resultierenden Streuintensität herstellt (siehe auch Glatter & Kratky, 1982; Feigin & Svergun, 1987).

Im Fall einer *single-particle* basierten elektronenmikroskopischen Rekonstruktion (EM-Rekonstruktion) eines Moleküls gibt es keinerlei Information über die Elektronendichten oder die Formfaktoren der Atome in Abhängigkeit des Ortes. Allerdings ist die Dichte einer EM-Rekonstruktion für jedes Teilvolumen der Rekonstruktion proportional zur Elektronendichte des entsprechenden Teilvolumens<sup>6</sup>. Solange keine atomare Auflösung der EM-Rekonstruktion vorliegt gilt:

$$\rho_{EM} \propto \rho_{\text{Elektronendichte}} . \quad (3.9)$$

Zur Zeit liegen zwei Programme für die Berechnung der Streukurven auf Basis von EM-Rekonstruktionen vor:

- **em2p2i** (geschrieben in FORTRAN77) von Dr. H. Hartmann (Inst. f. Mol. Biophysik) und
- **EM2SAXS** (geschrieben in Python) von mir. Eine detaillierte Beschreibung von EM2SAXS befindet sich im Anhang A.4. Die Implementierung des Algorithmus zur Berechnung von Streukurven gleicht der Implementierung in meiner Bibliothek zum *rigid body*-Modellierung (siehe Resultatekapitel 4.7.1.2).

Die Programme unterscheiden sich nicht in der Implementierung des eigentlichen Algorithmus, jedoch im Umfang der Funktionalität und durch die Programmiersprache – während die kompilierte Sprache FORTRAN77 Vorteile hinsichtlich der Geschwindigkeit bietet, ist die in Python programmierte Bibliothek wesentlich benutzerfreundlicher.

Das Vorgehen bei der Berechnung theoretischer Röntgenstreukurven, ausgehend von elektronenmikroskopischen Dichtekarten, ist nicht vergleichbar mit der Berechnung von theoretischen Streukurven auf Basis atomarer Modelle (siehe Abschnitt 3.6.8 Seite 37 dieses Kapitels), da über Voxel variabler Dichte iteriert werden muss. Die Berechnung ist fehleranfällig für den Schwellenwert, ab welchem ein Voxel als zum Volumen zugehörig erachtet wird (siehe Abschnitt 4.2 auf Seite 79 ff. des Resultatekapitels).

### 3.6.10 Entwicklung von dreidimensionalen *ab initio* Modellen

*Ab initio* Modelle stellen für die Modellierung von Proteinstrukturen auf Basis von Kleinwinkelstreuungsdaten die Methode der Wahl dar, wenn es gilt Modelle zu validieren, beziehungsweise ohne (oder mit wenigen) Annahmen überhaupt erste Modelle zu erstellen.

In der vorliegenden Arbeit wurden verschiedene Programme verwendet, auf deren spezifischen Einsatz hier eingegangen wird. Für die Erläuterung allgemeiner Aspekte der Programme sei auf

<sup>6</sup>Genauer: Die Dichte der EM-Rekonstruktion ist proportional dem Coulombpotential des entsprechenden Volumens. Da Ladungverschiebungen nicht zwangsläufig mit der Position von Atomen einhergehen, sind Coulombpotential und Elektronendichte bei atomarer Auflösung nicht proportional.

die folgenden Referenzen verwiesen.

#### DAMMIN:

Das Programm DAMMIN (Svergun, 1999) ist der Quasistandard für die Erstellung von *ab initio* Modellen. Das Programm erstellt ein Multiphasenmodell eines Partikels mit dicht gepackten „Dummy Atomen“. Diese sind durch einen Vektor charakterisiert, der den Atomen eine spezifische Phase (Protein oder Lösungsmittel, 1 oder 0) zuweist. DAMMIN führt ein sogenanntes „Simulated Annealing“ aus, um einen optimalen Fit an die Daten zu erreichen. Zugleich wird versucht eine kompakte Struktur zu erzielen. Selbst geschriebene Skripte gewährleisten eine möglichst effiziente Recherauslastung (siehe Anhang A.1.1).

Sofern nicht anders angegeben wurden 12 initiale Modelle erstellt und gemittelt (Volkov & Svergun, 2003). Das Mittel diente im folgenden Schritt als Muster für die Generierung weiterer 12 Modelle, die anschließend wieder gemittelt wurden und das finale Modell ergaben. Sofern weitere Parameter abweichend von den Standardeinstellungen gegeben wurden, werden diese bei den gezeigten Modellen genannt.

#### GASBOR:

GASBOR (Svergun *et al.*, 2001) funktioniert ähnlich wie DAMMIN, jedoch gibt es statt „Dummy Atomen“ „Dummy Reste“, ebenfalls in Form von Kugeln. Diese vereinen den mittleren Formfaktor von Aminosäuren. Die „Dummy Reste“ werden in Art einer Kette, mit oder ohne Verzweigungen, angelegt und das Programm versucht, ähnlich wie bei DAMMIN, die Kette in eine Konfiguration zu bringen, welche die experimentellen Daten optimal widerspiegelt.

Das Programm ist insbesondere für kleinere Proteine geeignet, da hier mit der Kette der „Dummy Reste“ die Sekundärstruktur des Proteins nachempfunden werden kann – bei größeren Proteinen (>2000 Aminosäuren) kann das Programm nicht mehr sinnvoll eingesetzt werden. Es wurde ähnlich wie DAMMIN mit einem Skript (siehe Anhang A.1.1) gefüttert, jedoch entfiel die Mittelung der Strukturen, da diese Mittelung die Annäherung an die Sekundärstruktur mit den Ketten wieder verschmieren würde.

#### SUPCOMB:

SUPCOMB (Kozin & Svergun, 2001) wurde verwendet, um *ab initio* Modelle, welche mit DAMMIN generiert wurden, so übereinanderzulegen, dass die räumliche Diskrepanz zwischen den Modellen minimal wird. Das Programm verwendet keine globale Minimierung räumlicher Diskrepanzwerte, wie üblicherweise bei strukturellen Alignments (Konagurthu *et al.*, 2006). Stattdessen wird eine lokale Minimierung der Hauptträgheitsachsen der Modelle angestrebt.

Das Programm wurde zudem dazu verwendet, um eine Superposition atomarer Modelle mit *ab initio* Modellen zu erreichen. Da es unmöglich ist strukturelle Elemente, wie  $C_{\alpha}$ -Atome, in *ab initio* Modellen zu identifizieren, stellte SUPCOMB hier die einzige Alternative dar.

#### DAMAVER:

DAMAVER (Volkov & Svergun, 2003) ist kein eigenständiges Programm, sondern ein Paket von Programmen bzw. Skripten. Es dient zur quasi-parallelen Mittelung der mit DAMMIN erstellten Modelle. Herzstück dieser Mittelung ist das Programm SUPCOMB, welches die Superposition der einzelnen Modelle garantiert. Damit die Programme DAMAVERs automatisch hintereinander geschaltet werden können und es nicht menschlicher Eingriffe bedarf, wurden von mir auch hierzu Skripte geschrieben (siehe Anhang A.1.1).

### 3.6.11 Die Verwendung von parmscan zur Konformationsberechnung von 24meren

Das Programm parmscan von Dr. H. Hartmann wurde in der Vergangenheit dazu verwendet die oxygenierte und deoxygenierte Konformation des Hämocyanins von *E. californicum* auf der Basis von SAXS-Daten zu errechnen (Hartmann & Decker, 2002; Hartmann & Decker, 2004). Es ist auf die besondere Symmetrie 24merer Hämocyanine angepasst und berücksichtigt neben

der zweifachen Rotationssymmetrie auch die Relationen auf den Hierarchieebenen der Trimere, Hexamere und Dodekamere.

Als Eingabe für das Programm diente die Struktur des artifiziellen Hexamers von *L. polyphemus* (Hazes *et al.*, 1993, PDB-Code: 1LLA). Das Hexamer wurde durch Symmetrieoperationen in ein 24mer überführt. Da die Berechnung der theoretischen  $P(r)$  oder  $I(q)$  zum Vergleich mit den experimentellen Daten mit den Atomen eines 24mers viel zu lange dauern würde – die Komplexität<sup>7</sup> des Algorithmus ist wegen der Berücksichtigung aller Atomabstandspaare (siehe Gleichung 2.12 auf Seite 20 des Theoriekapitels)  $O(n^2)$  – wurde das Modell folgendermaßen angepasst: Sämtliche Seitenkettenatome wurden entfernt. An ihre Stelle tritt ein „formelles“ Atom, das in den geometrischen Schwerpunkt der Seitenkettenatome platziert wird. Die Atome des Peptidrückgrates werden ebenfalls, bis auf das  $C_\alpha$ -Atom entfernt (persönliche Mitteilung Dr. H. Hartmann).

Im Übrigen wurde die Vorgehensweise der Publikation zum Hämocyanin von *E. californicum* (Hartmann & Decker, 2002) imitiert: Das Programm verwendet eine Reihe von Rahmenbedingungen für Bewegungsparameter. Innerhalb dieser Grenzen bewegt **parmscan** die Struktur und evaluiert die theoretischen  $P(r)$  und  $I(q)$  im Vergleich zu den experimentellen Daten. Bei der Festlegung der Parameter halfen die Erfahrungen während der Publikation zu Hartmann & Decker, 2002: Da die gemessenen Daten für die oxygenierten und deoxygenierten Hämocyanine ähnlich den für diese Publikation zugrundeliegenden Daten (Decker *et al.*, 1996) waren, konnte von ähnlichen Bewegungsparametern ausgegangen werden. Dies schränkte die Zahl der abzusuchenden Bewegungsparameter erheblich ein.

## 3.7 Elektronenmikroskopie

### 3.7.1 Probenpräparation und Aufnahmen

Elektronenmikroskopische Aufnahmen wurden ausschließlich im Negativkontrast angefertigt. Hierzu fanden Netzchen mit 300 Mesh (=Maschen pro Inch), die mit Kohlefilm überschichtet wurden (Meesters, 2001), Verwendung. Als Kontrastmittel kamen entweder Uranylacetat (UAc) 2 % (w/v) oder Phosphowolframsäure 2 % (w/v), pH 7,0, zum Einsatz. Proteine wurden vor dem Auftrag auf circa 0,1 mg/mL verdünnt und dann in Volumina von 4  $\mu$ l auf das Netzchen pipettiert. Die Probe hatte hierauf etwa eine Minute Zeit auf dem Netzchen zu adsorbieren. Anschließend wurden die Netzchen zweimal mit je 4  $\mu$ l Wasser gewaschen und mit 4  $\mu$ l des Kontrastmittels benetzt. Nach einer weiteren Minute wurde der größte Teil des Kontrastmittels mit einem Filter (Whatman, Dassel, Typ 1004090) abgezogen und der Rest luftgetrocknet. Die Färbung erfolgte nach der Methode von Jacques Dubochet (Dubochet *et al.*, 1982).

Die Aufnahmen wurden mit einem Tecnai 12F-Mikroskop bei 120 kV Beschleunigungsspannung getätigt. Die Bilder wurden entweder mit einer integrierten CCD-Kamera oder auf Film (Kodak, Stuttgart, Typ *SO163*) aufgenommen.

### 3.7.2 Bildbearbeitung

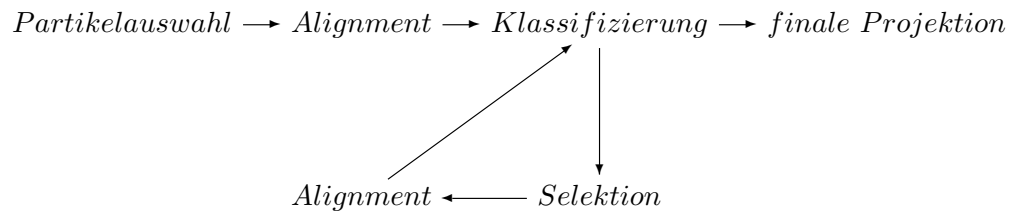
#### 3.7.2.1 Einzelpartikelanalyse

Alle Schritte der Einzelpartikelanalyse erfolgten mit dem Programm SPIDER (Frank *et al.*, 1981) in der Version 14.18<sup>8</sup>. Alle Schritte des Alignments (Ausrichten der Partikel aufeinander) wurden referenzfrei durchgeführt. Lediglich in Teilprojekten (siehe dort) kamen einfache geometrische Figuren zum leichteren Sortieren von Partikeln zum Einsatz.

Die Bildbearbeitungsschritte gliederten sich wie folgt:

<sup>7</sup>Für die sogenannte O-Notation der Komplexität von Algorithmen siehe Loudon, 1999.

<sup>8</sup>SPIDER liegt mittlerweile in einer neueren Version vor. Diese erwies sich in Tests allerdings als fehleranfällig und wurde daher nicht verwendet.



- a) **Partikelwahl:** Die Partikelwahl geschah stets per Hand. Gegebenenfalls wurde vorab der Kontrast der Bilder erhöht, aber diese veränderten Bilder wurden nur für die Bestimmung der Partikelkoordinaten, nicht zum Ausschneiden der Partikel an sich verwendet. Die Selektion geschah manuell, weil bisherige automatische Auswahlalgorithmen zuviele scheinbare Partikel auswählen, die tatsächlich keine sind. Dies ist bei großen Projekten mit vielen Partikeln unproblematisch, würde aber bei Projekten mit kleinen Partikelzahlen, wie in dieser Arbeit, die Statistik der Alignments signifikant verschlechtern.
- b) **Alignment:** Für die ersten Schritte des Alignments wurden Partikel vorgezogen, die einem Fourier-*low-pass*-Filter (bei etwa dem zwei- bis vierfachen der Nyquistfrequenz) unterzogen wurden. Weitere Alignments wurden mit den Originalpartikeln durchgeführt.
- c) **Klassifizierung:** Hierarchische Klassifizierung (*hierarchical ascendant classification* oder HAC nach Ward, 1982) wurde zur Klassifizierung der Partikel nach dem ersten Alignment eingesetzt. Die Zahl der Klassen wurde eingeschränkt, so dass die Größenunterschiede der Eigenvektoren der Klassenmittel von Klasse zu Klasse signifikant groß waren<sup>9</sup>.
- d) **Selektion „guter“ Partikel:** Partikel, die beispielsweise aufgrund irgendwelcher Artefakte zu stark von ihrem jeweiligen Klassenmittel abwichen (beispielsweise unterhalb von Korrelationskoeffizientengrößen von 0,5 bis 0,7), wurden ausgesondert und das Alignment ohne diese Partikel wiederholt.
- e) **Erstellung der finalen Projektionen:** Sowie das Alignment keine Verbesserung mehr gegenüber einem vorhergehendem Alignment brachte wurde abgebrochen und die finalen Klassenmittel (Projektionen) wurden erstellt.

### 3.7.2.2 Elektronenkristallographie

Vollständige elektronenkristallographische Analysen kamen in dieser Arbeit nicht zum Tragen, weil die Qualität der Kristalle nicht ausreichend war (für Details: siehe die Diskussion zur Präpore des  $\alpha$ -Hämoglobins 5.9.2). Eine Beschreibung der Kristallgitter und erste Korrelationsmittel wurden jedoch mit der Programmsuite IPLT (Philippsen *et al.*, 2003; Philippsen *et al.*, 2007) durchgeführt.

## 3.8 Molekulare Modellierung und Molekulardynamiksimulation

### 3.8.1 Molekulare Modellierung mit MODELLER

MODELLER ist ein Programm zur sogenannten „vergleichenden“ oder Homologie-Modellierung von Proteinstrukturen (Sali & Blundell, 1993; Martí-Renom *et al.*, 2000; Fiser & Sali, 2003). Die Bezeichnung „vergleichende“ Modellierung stammt von der Vorhersage der Struktur von Proteinen mit Sequenzhomologie zu Musterproteinen mit experimentell ermittelter Struktur. Die Vorgehensweise bei einer solchen Modellierung folgt stets dem gleichen Muster:

1. Alignment („Ausrichten“) der Sequenzen von Mustern und Zielprotein.
2. Festlegen der räumlichen Randbedingungen (Abstände von Atomen, Sekundärstruktur, etc.).

<sup>9</sup>Wählt man zu viele Klassen, so laufen die Größen der Eigenvektoren auf einem niedrigen Niveau aus. Dies würde indizieren, dass kein signifikanter Unterschied von Klasse zu Klasse besteht.

3. Modellierung der Zielstruktur durch Applikation dieser Randbedingungen auf die Sequenz des Zielproteins.
4. Gegebenenfalls eine Anpassung oder Relaxation der Struktur des Zielproteins durch eine abschließende Optimierung hinsichtlich einer Wahrscheinlichkeitsfunktion: MODELLER generiert eine Wahrscheinlichkeitsfunktion für bestimmte Atomensembles (z. B. Aminosäureseitenketten oder prostetische Gruppen) auf Basis bekannter Röntgenstrukturen. Hierbei wird für jede dieser Atomgruppen ermittelt welche Konfigurationen mit welcher Häufigkeit vorkommen. MODELLER ermittelt mit einer *steepest descent* Simulation, bei der nahestehende Atomgruppen hinsichtlich ihrer Konfiguration so lange optimiert werden, bis eine größtmögliche Übereinstimmung für die Wahrscheinlichkeitsfunktionen hinsichtlich aller Atomgruppen erreicht ist.

Das Alignment der Sequenzen wurde mittels CLUSTAL X (Jeanmougin *et al.*, 1998) ermittelt und Fehler per Hand korrigiert. Waren Bereiche des Alignments in keiner der Musterstrukturen angegeben, etwa weil es in der kristallographischen Analyse an diesen Stellen keine gut definierte Elektronendichte gab, so waren diese Bereiche in den resultierenden Modellen schlecht bestimmt. Es wurde versucht mittels eines sogenannten „Loopmodellings“ (Modellierung von Schleifen; Fiser *et al.*, 2000) eine möglichst wahrscheinliche Struktur für jene Bereiche zu finden. Das Loopmodelling in MODELLER erlaubt, ungeachtet des Namens der Methode, strukturelle Randbedingung sogar in Form von Sekundärstrukturelementen vorzugeben.

### 3.8.2 Auswertung und Beurteilung der Strukturmodelle

Eine erste Beurteilung von Modellstrukturen fand über die in MODELLER integrierte Möglichkeit der Berechnung „diskreter optimierter Protein Energien“ (kurz: DOPE; Shen & Sali, 2006) statt. Je niedriger der Wert, desto näher liegt eine Struktur im Mittel ihrer Elemente an den am häufigsten vorkommenden Strukturelementen in der Proteindatenbank (Berman *et al.*, 2000). DOPE-Werte sind relative Werte: Hohe Werte bedeuten mit großer Wahrscheinlichkeit die Verletzung von Strukturprinzipien. Eine Wichtung erfolgt über die Häufigkeit der Strukturelemente in der Proteindatenbank. Ungeachtet der Qualität der Beurteilung durch DOPE-Werte kann es dazu kommen, dass niedrige DOPE-Werte bei „unsinnigen“ Strukturen vorkommen. Unsinnig heißt hier beispielweise, dass topologisch unmögliche Konfigurationen festgestellt werden. Dies kann vorkommen, wenn nur Teile der Struktur (z. B. Untereinheiten oder Schleifen eines Proteins) auf in einem bestimmten Lauf des Programms modelliert werden, da das Programm nicht in der Lage ist Strukturen auf topologische Plausibilität zu prüfen.

Diejenige Struktur mit den besten DOPE-Werten ohne offensichtliche Fehler wurde für die weitere Auswertung verwendet. Eine abschließende Kontrolle fand mit dem in YASARA (Krieger *et al.*, 2002) integrierten Programm WhatIf (Vriend, 1990; Hooft *et al.*, 1996) statt.

Dem Vergleich von Strukturmodellen ging eine Superposition verschiedener Modelle mit dem MUSTANG-Algorithmus (Konagurthu *et al.*, 2006), welcher in YASARA integriert ist, voran. Ein wichtiges Kriterium für die Unterschiede zwischen Modellen ist die mittlere, quadratische Abweichung (*root mean square distance* bzw. RMSD) zweier Modellen  $x$  und  $y$ :

$$\text{RMSD} = \sqrt{\sum_{j=1}^N (P_x^2 - P_y^2)}, \quad (3.10)$$

wobei  $N$  alle betrachteten Atome sind und  $P$  für die Postionsvektoren steht.

### 3.8.3 Molekulardynamiksimulationen

Molekulardynamiksimulationen wurden mit YASARA (Krieger *et al.*, 2002) durchgeführt. Hierbei wurde das Kraftfeld YAMBER2 (Krieger *et al.*, 2004) verwendet, das eine Anpassung des Amber-Kraftfeldes (Cornell *et al.*, 1995) an sich wiederholende Raumzellen in einer Simulation darstellt.

Die Simulationen am  $\alpha$ -Hämolyisin sollten nach Möglichkeit den experimentellen Bedingungen entsprechen. Daher wurden pH-Wert und Ionenstärke entsprechend angepasst. Das zeitliche Inkrement für eine Simulation betrug 2 ps. Alle 25 ps erfolgte ein Ausschreiben der gesamten Simulationszelle in eine „Schnappschuss“-Datei. Diese Dateien dienten zur späteren Analyse. Alle Simulationen und Analysen wurden mit den Skripten durchgeführt, die im YASARA-Download enthalten sind.

Die Berechnung der mittleren quadratischen Fluktuation (*root mean square fluctuation* bzw. RMSF-Werte) diente dazu die Beweglichkeit innerhalb von Modellen abzubilden. Die Werte werden über die „Schnappschüsse“ einer Simulation für jedes  $C_\alpha$ -Atom gemittelt, wobei der Index  $j$  über die karthesischen Koordinaten  $x, y, z$  des Atomvektors  $P$  und der Index  $k$  über die Zahl der „Schnappschüsse“ ( $N$ ) läuft:

$$\text{RMSF} = \sqrt{\sum_{j=1}^3 \left( \frac{1}{N} \sum_{k=1}^N P_{jk}^2 - \bar{P}_j^2 \right)}. \quad (3.11)$$

### 3.8.4 „Morphing“

Zur Simulation von Konformationsübergängen wurde auch sogenanntes „Morphing“ eingesetzt. *Morphing* wurde mit YASARA (Krieger *et al.*, 2002) durchgeführt. Hierbei wurde eine Start- und eine Zielstruktur vorgegeben. Das Programm versucht die Start- in die Zielstruktur zu überführen, ohne das hierbei eine Molekulardynamiksimulation im engeren Sinn ausgeführt wird. Kräfte in Richtung der Atomkoordinaten der Zielstruktur ziehen an den Atomen der Startstruktur, wobei die Kräfte im Bereich weniger Pikonewton angewendet werden, also eine reale Größenordnung nie verlassen. Zudem ist während eines *Morphings* gewährleistet, dass es zu keinen Verletzungen sterischer (z. B. Überlappungen von Van-der-Waals-Radien) und topologischer (z. B. „Verknäulungen“ der Primärstruktur) Bedingungen kommen kann. Hierdurch ist implizit sichergestellt, dass „unmögliche“ Bereiche des Ramachandranplots nur mit geringer Wahrscheinlichkeit eingenommen werden.

### 3.8.5 *rigid body*-Modellierung des *Palinurus elephas*-Hämocyans auf der Basis von SAXS-Daten

Für die *rigid body*-Modellierung auf der Basis von SAXS-Daten wurde eine eigene Software geschrieben. Die Interna der Software sind wesentliches Resultat dieser Dissertation und im Resultatekapitel beschrieben (siehe Abschnitt 4.7, Seite 105). Alle Schritte der Modellierung folgten der Anleitung (Meesters, 2008b). Es wurden zwei Modellierungsmodi angewendet:

1. Ein Monte Carlo-Lauf mit anschließender systematischer Suche nach einer Konformation, welche die SAXS-Daten am besten reflektiert (siehe Abschnitt 4.7.2, Seite 110 ff.).
2. Ein Monte Carlo-Lauf mit dem Ziel durch einen nicht-negativen Algorithmus der kleinsten Quadrate diejenigen Konformationen zu finden, die zusammen die SAXS-Daten ausmachen können (siehe Abschnitt 4.7.5, Seite 115 ff.).

Für das Hämocyans von *Palinurus elephas* wurde allen Modellen die D3-Symmetrie des Hexamers aufgezwungen, so dass alle Untereinheiten letztlich in gleicher Weise bewegt wurden. Zudem wurde die Struktur vor Aufnahme der Modellierungen so zentriert, dass der geometrische Schwerpunkt aller Atome im Ursprung des verwendeten karthesischen Koordinatensystems war und die Rotationsachse des Moleküls mit der Achse „Z“ in der verwendeten Software zusammenfiel.

## 3.9 Charakterisierung von Hämocyaninen unter SDS-Einfluss

### 3.9.1 Besondere Erwägungen zur Röntgenkleinwinkelstreuung

Im Wesentlichen wurde die Röntgenkleinwinkelstreuung durchgeführt wie unter Abschnitt 3.6 beschrieben. Auch wenn unsere schottischen Kooperationspartner zeigen konnten, dass insbesondere Arthropodenhämocyanine unter SDS-Einwirkung große strukturelle Stabilität aufweisen (Baird *et al.*, 2007), war doch stets die Möglichkeit der Denaturierung zu berücksichtigen. Da Denaturierung hier in erster Linie die Dissoziation in Untereinheiten (oder Oligomere von Untereinheiten, beispielsweise 6mere im Fall eines 24mers) bedeutet, würde dies einerseits erhebliche Auswirkung auf eine SAXS-Messung haben, andererseits aber dadurch auch gut zu erkennen und zu verfolgen sein.

Folglich mussten die Teilmessungen auf eine Strahlzeit von 30 Minuten beschränkt werden. Aufzeichnungen über eine halbe Stunde erlauben es zu kontrollieren, ob eine signifikante Veränderung, sprich ein Zerfall oder eine starke Formveränderung des Proteins, auftritt. Zur Kontrolle wurden die Daten direkt untereinander verglichen und Kratky-Plots (siehe 3.6.6) für die einzelnen Teilmessungen angefertigt.

Eine Mittelung der Daten wurde nur dann vorgenommen, wenn im direkten Vergleich keine signifikanten Veränderungen auftraten und die Kratky-Plots im Verlauf aller Teilmessungen keinen eindeutigen Trend zeigten<sup>10,11</sup>.

Die verwendeten Proteinkonzentrationen für die Gewinnung der Teildatensätze waren 2 x 4,5 mg/mL, 6,8 mg/mL, 6,9 mg/mL, 9,2 mg/mL und 30 mg/mL.

### 3.9.2 Nachweismöglichkeiten für Phenoloxidasen

#### 3.9.2.1 Kontrolle der Phenoloxidaseaktivität im Polyacrylamidgel

Es ist möglich die Aktivität von Phenoloxidasen innerhalb einer nicht-denaturierenden Gelmatrix nachzuweisen. Das ursprüngliche Protokoll hierzu wurde von Nellaiappan & Vinayakam, 1993, entwickelt.

Das native Gel wurde hierzu 10 – 15 Minuten in 0,1 M Phosphatpuffer (62 mM Na<sub>2</sub>HPO<sub>4</sub>, 38 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 7,0) inkubiert, um den alkalischen Elektrophoresepuffer (pH 8,8) zu ersetzen. So wird Hintergrundfärbung durch starke Autooxidation vermieden. Anschließend wurde das Gel im Phosphatpuffer mit 3,3 mM Substrat und 0,1 mM SDS als Aktivator inkubiert. Erste Ergebnisse zeigen sich in der Regel nach einer dreißigminütigen Inkubationszeit. Eine Zugabe von MBTH (3-Methyl-2-Benzo-Thiazolinone-Hydrazon; 0,3 % [w/v] in Ethanol; Sigma-Aldrich, München, Produktnummer *M8006-5G*) kann die Färbung verstärken.

#### 3.9.2.2 „Dot-Blot“-Assay zur Aktivitätskontrolle

Der „Dot-Blot“-Assay wurde zur schnellen Verifikation der Aktivität von Hämocyaninen eingesetzt. Diese Vorgehensweise ermöglicht vor allem eine schnellen und unkomplizierte qualitative Überprüfung einer Vielzahl von Proben. Das Prozedere ist ähnlich wie unter 3.9.2.1 beschrieben und ist auf Basis von Winder & Harris, 1991, und Nellaiappan & Vinayakam, 1993, entwickelt worden (Winder, 1994). Auf einer PVDF-Membran (Fa. BioRad, München, Typ *Sequi-Blot PVDF-Membran*) werden 2 µl der Färbelösung bestehend aus 500 µl des Aktivators (z. B. eine

<sup>10</sup>Die zusätzliche Kontrolle mittels Kratky-Plots wurde zu Anfang insbesondere für die Interpretation der Daten herangezogen. Es zeigte sich aber, dass Kratky-Plots empfindlich auch auf kleine Veränderungen im untersuchten Molekül reagieren. Daher wurden diese Plots im folgenden zur routinemäßigen Qualitätskontrolle herangezogen.

<sup>11</sup>Bei Korrekturen kam die Frage auf was mit „Trend“ gemeint ist. Für einen Trend ist hierbei unerheblich, welcher Natur er ist: Sofern Daten regelmäßig voneinander abweichen ist die Berechnung eines arithmetischen Mittels nicht mehr plausibel. Er kann hierbei durch Dissoziation, Aggregation oder kinetisch langsame Konformationswechsel ausgelöst werden.

bestimmte SDS-Konzentration), 500  $\mu\text{l}$  MBTH (0,3 % [w/v] in Ethanol, Zugabe ist optional) und 3,3 mM des Substrates gegeben. Sofern nicht anders angegeben handelt es sich bei dem Substrat in dieser Arbeit um Dopamin<sup>12</sup>. Anschließend wird der Tropfen mit 2  $\mu\text{l}$  der Probe durchmischt. Die Negativkontrolle wurde mit Wasser anstelle der Probe durchgeführt.

Bei bereits aktiviertem Protein wurde die MBTH-Lösung so mit der Dopaminlösung vermischt, so dass die oben genannten Konzentrationen erhalten blieben und abschließend die Probe zupipettiert.

### 3.9.3 Verifikation der Detergens-CMC

Detergenzien lagern sich oberhalb einer detergensspezifischen Konzentration, der sogenannten kritischen mizellären Konzentration oder CMC, zu Mizellen zusammen<sup>13</sup>. Die thermodynamische Beschreibung des Zusammenlagerungsprozesses basiert im Allgemeinen auf zwei Modellen, welche beide einen hochkooperativen Prozess annehmen, der durch hydrophobe Interaktionen angetrieben wird (Roux *et al.*, 1984).

Gemäß des geschlossenen Assoziationsmodells besteht ein chemisches Gleichgewicht zwischen Detergensmonomeren und Micellen. Dieses Gleichgewicht ist analog dem Massenwirkungsgesetz (Blandamer *et al.*, 1995): Ist  $N$  die Aggregationszahl eines Detergens und  $S$  seine Konzentration, so gilt

$$N \cdot S \rightleftharpoons S_N \quad (3.12)$$

und die Gleichgewichtskonstante ergibt sich klassisch zu

$$K = \frac{S_N}{S^N} . \quad (3.13)$$

Die Verifikation der kritischen micellären Konzentration (CMC) von Detergenzien stellt eine große experimentelle Herausforderung dar. Tabellenwerte werden u. a. von den Herstellern zusammengefasst angeboten. Für einige Detergenzien besteht aufgrund umfangreicher experimenteller Arbeiten auch die Möglichkeit die CMC zu errechnen. So bieten Dutkiewicz & Jakubowska, 2002 diese Möglichkeit für SDS (Sodium Dodecylsulfat) an. Das andere in dieser Arbeit als Aktivator verwendete Detergens ist CPC (1-Hexadecylpyridiniumchlorid oder Cetylpyridiniumchlorid; Sigma-Aldrich, München, Produktnummer C9002).

In dieser Arbeit wurde für die verwendeten Puffer eine Verifikation der CMC-Werte angestrebt, da die CMC für die verwendeten Puffer zu diesem Zeitpunkt nicht feststand (erst die Diplomarbeit von C. Dietze (Dietze, 2008) lieferte detaillierte Werte). Eine einfache Kontrolle ist die „Tröpfchenmethode“ (Ringler *et al.*, 2000). Hierbei wird eine Verdünnungsreihe des Detergens im Puffer hergestellt. Von jeder Verdünnung wird ein definiertes Volumen in eine Spritze mit stumpfer Nadel (verwendet wurde eine 250  $\mu\text{l}$ -Spritze von Hamilton, Bonaduz, CH) aufgenommen. Die Spritze wird so eingespannt, dass ihre Spitze senkrecht zum Boden zeigt, dann wird das aufgenommene Volumen abgetropft. Die Zahl der dabei entstehenden Tropfen wird für jede Verdünnung des Detergens gezählt und gegen die Konzentration des Detergens aufgetragen.

Da die Oberflächenspannung bei einer Zunahme der Detergenzkonzentration kleiner wird, reißen die Tropfen mit der Erhöhung der Detergenzkonzentration früher ab: Die Zahl der Tropfen nimmt linear mit der Detergenzkonzentration zu. Wird jedoch der Wert der CMC erreicht, so erhöht sich die Zahl freier Detergensmonomere nicht mehr. Die Zahl der Tropfen erhöht sich nur noch wenig, aber immer noch linear mit der Detergenzkonzentration. Die CMC kann nun über den Schnittpunkt der beiden Geraden zwischen Anstieg und Plateau ermittelt werden.

<sup>12</sup>Detektiert wurde also ausschließlich die Diphenolaseaktivität des jeweiligen Proteins. Diese Einschränkung wurde gemacht, da in dieser Arbeit nicht die enzymatische Aktivität der Protein im Vordergrund stand.

<sup>13</sup>Es gibt die Anekdote, nach welcher der ursprüngliche Vorschlag von McBain, dass sich Detergensmoleküle überhalb einer kritischen micellären Konzentration (CMC) zu Aggregaten zusammenlagern (McBain, 1913) mit einem schlichten „Nonsense, McBain“ abgetan wurde (Menger, 1979).

### 3.10 $\alpha$ -Hämolyisin

Für die Behandlung des  $\alpha$ -Hämolyisins wurden die folgenden Puffer und Medien verwendet:

Tabelle 3.5: Puffer für die Reinigung und Lagerung des  $\alpha$ -Hämolyisins

<i>Reinigungspuffer 1</i>	20 mM Ammoniumacetat pH 5,7 bei 20 °C
<i>Reinigungspuffer 2</i>	500 mM Ammoniumacetat pH 5,7 bei 20 °C
<i>Phosphatpuffer</i>	20 mM Phosphatpuffer pH 7,2 bei 20 °C
<i>PBS</i> (Phosphat Buffer Saline)	136 mM NaCl 2,6 mM KCl 9,3 mM Na <sub>2</sub> HPO <sub>4</sub> 1,4 mM KH <sub>2</sub> PO <sub>4</sub> pH 7,2 bei 20 °C

Tabelle 3.6: Expressionsmedien für das  $\alpha$ -Hämolyisin

LB	10 g NaCl 5 g Hefeextrakt 10 g Trypton 12,5 $\mu$ g Chloramphenicol auf 1L mit vollentsalztem Wasser
2 x YT	5 g NaCl 10 g Hefeextrakt 16 g Trypton 12,5 $\mu$ g Chloramphenicol auf 1L Aqua dest.

#### 3.10.1 Expression und Aufreinigung

Mit einer in Glycerin gefrorenen *Staphylococcus aureus*-Kultur (Stamm Y 1090) wurden 4 mL LB-Medium angeimpft und über Nacht bei 37 °C geschüttelt. Am Folgetag wurde diese Kultur in weitere 80 mL LB-Medium gegeben und ca. 7 bis 8 Stunden weiter bei 37 °C inkubiert. Anschließend wurde diese Vorkultur auf 4 Kolben mit je 1 L 2 x YT-Medium überführt und auf 8 2L-Kolben mit je 500 mL Inhalt aufgeteilt. Diese Kulturen wurden über Nacht bei 37 °C geschüttelt.

Am darauf folgenden Morgen wurden die Kulturen in Zentrifugenbecher gefüllt und in einer Sorvall-Zentrifuge (Sorvall RC-5B Zentrifuge; ehemals Sorvall, heute von Fisher Scientific, Schwerfen) für 20 min bei 13700 x g zentrifugiert. Der Überstand wurde mit einem Ultrafiltrationsmodul (Satorius AG; Typ *SM 17566*) konzentriert. Auf das Konzentrat kamen in drei Schritten insgesamt 2,5 L *Reinigungspuffer 1*, um die Lösung umzupuffern und zu entsalzen, damit sie auf die Säule (s. u.) gegeben werden konnte. Das hieraus resultierende Konzentrat wurde abermals zentrifugiert und der Überstand anschließend auf die Säule gesaugt.

Die Aufreinigung erfolgte mit Hilfe einer FPLC-Anlage (FPLC=*fast protein liquid chromatography*; von BioRad, München, Typ *BioLogic DuoFlow*) gekühlt bei 4 °C. Beim Säulenmaterial handelte sich um 30 mL CM-Sepharose FF (ein Kationenaustauscher: GE Healthcare, Solingen,

Typ 17-0729-10). Die Säule wurde nach Aufbringen der Probe mit 3 Säulenvolumen des Waschpuffers (*Reinigungspuffer 1*) gewaschen und anschließend wurde die Konzentration von *Reinigungspuffer 2* in einem linearen Gradienten innerhalb von 82 min auf 100 % erhöht. Die Flussrate betrug konstante 2 mL/min. Mit Start des Gradienten wurden Fraktionen von 1 mL Größe gesammelt. Bestimmte Fraktionen wurden in sogenannten „Pools“ vereinigt (siehe Abbildung 4.32, Seite 4.32).

Die Säule wurde anschließend mit 10 Säulenvolumen Wasser, 10 Säulenvolumen 0,1 % (w/v) SDS mit 200 mM NaOH und 10 Säulenvolumen Wasser gespült und abschließend unter Wasser mit 3 mM  $\text{NaN}_3$  gelagert.

### 3.10.2 Konzentrationsbestimmung des $\alpha$ -Hämolytins

Zur Konzentrationsbestimmung des  $\alpha$ -Hämolytins fanden zwei Methoden Verwendung:

- a) die Methode nach Bradford (Bradford, 1976) und
- b) die Bestimmung über den theoretischen, spezifischen Extinktionskoeffizienten bei 280 nm.

Zur Bestimmung nach Bradford wurde das Bio-Rad Protein-Assay-Reagenz 1:5 mit Wasser gemischt. 10  $\mu\text{L}$  Probe wurden mit 1 mL des verdünnten Reagenz gemischt. Nach 5 min Inkubationszeit wurde die  $\text{OD}_{595 \text{ nm}}$  gemessen. Eine Negativkontrolle durch Vermessen einer Mischung von Wasser und Reagenz, im selben Verhältnis wie oben beschrieben, wurde durchgeführt. Eine Kalibrationsgerade wurde unter Verwendung verschiedener Verdünnungen (von 0 bis 1 mg/mL) von Rinderserumalbumin (BSA) erstellt.

Der theoretische, spezifische Extinktionskoeffizient wurde mit dem Program `pepstats` der EMBOSS-Software-Suite (Version 4.1.0, Rice *et al.*, 2000) errechnet:  $\epsilon_{\text{mol}, 280 \text{ nm}} = 63440 \cdot \text{mol}^{-1}$  bzw.  $\epsilon_{\text{Masse}, 280 \text{ nm}} = 1,77 \cdot \text{mg} \cdot \text{mL}^{-1}$ . Das Protein besitzt 8 Tryptophanreste und 14 Tyrosinreste, so dass diese Bestimmung leicht möglich ist.

Das  $\alpha$ -Hämolysin ist ein relativ kleines Protein (293 Aminosäuren), so dass der theoretische Wert wahrscheinlich nicht weit vom tatsächlichen Wert abweicht (Rice *et al.*, 2000). Zudem wurde die Messung vor allem für die schnelle Kontrolle der Konzentration zur SAXS-Messung verwendet. Hierbei spielt vor allem die Konzentration einer Teilmessung relativ zu anderen eine Rolle und der Verzicht auf eine chemische Proteinbestimmung verringert den Pipettierfehler.

### 3.10.3 Gefriertrocknung des aufgereinigten Proteins

Das gereinigte  $\alpha$ -Hämolysin wurde gefriergetrocknet (lyophilisiert), um a) den Ammoniumacetatpuffer aus der Probe zu holen<sup>14</sup> und b) beliebige Konzentrationen durch Variation der Puffermenge beim Lösen des Lyophilisates einstellen zu können. Ersteres ist wichtig, da für SANS-Messungen ausschließlich ein Puffer mit auswechselbaren Protonen verwendet werden sollte (siehe Abschnitt 3.10.6 auf Seite 50). Hohe Konzentrationen für die SAXS-Messungen (siehe Abschnitt 3.10.5 auf Seite 50) ließen sich so ebenfalls leicht einstellen.

Die Fraktionen der Aufreinigung (siehe Abschnitt 3.10.1) wurden im flüssigen Stickstoff (ca.  $-190^\circ\text{C}$ ) gefroren. Der Deckel des 1,5 mL-Reaktionsgefäßes wurde hierzu punktiert, um den Druckausgleich zu ermöglichen. Das Reaktionsgefäß wurde anschließend mit dem Probevolumen voran schnell in den Flüssigstickstoff gegeben. Es wurde darauf geachtet, dass die Aliquots nicht größer als 1000  $\mu\text{L}$  waren.

<sup>14</sup>Ammoniumacetatpuffer gehört zu den „volatilen“ Puffer: Er ist bei geringen Konzentrationen nicht stabil, da Teile des Puffers verdampfen. Wird der Puffer eingefroren und einem Unterdruck ausgesetzt, so verdampft der Puffer zu großen Teilen. Dies ist ein Effekt, der auch bei der Molekulargewichtsbestimmung mittels Scanning-Transmissionselektronenmikroskopie genutzt wird: Hier kann Protein in diesem Puffer gefroren und im Vakuum eines Elektronenmikroskops getrocknet werden. Die Deflektion des Elektronenstrahl ist zu nahezu 100 % die Masse des Proteins und nicht des Salzes zurückzuführen (Müller & Engel, 2001). Im Fall der  $\alpha$ -Hämolytinsproben war die Konzentration des Puffers so groß, dass noch einer kleiner Teil des Salzes als „weißer Schnee“ zurückblieb.

Vor der Gefriertrocknung wurde ein Aluminiumblock mit Löchern ebenfalls in Flüssigstickstoff gekühlt und die Reaktionsgefäße mit der Probe wurden hineingestellt. Dieser Block stellte sicher, dass die Probe während des Transfers in die Gefriergetrocknungsanlage (Martin Christ, Osterode, Typ *ALPHA 1-4 LSC*) nicht tauen konnte. Anschließend wurde der Block mit den Proben in die Gefriergetrocknungsanlage gestellt und Vakuum ( $\leq 1$  mbar) über Nacht angelegt. Am folgenden Tag wurde die Probe wieder bei  $-20^\circ\text{C}$  gelagert.

### 3.10.4 Rekonstitution des $\alpha$ -Hämolytins

#### 3.10.4.1 Bestimmung des Toxintiters

Zur Qualitätskontrolle des  $\alpha$ -Hämolytins wurde eine Bestimmung des Titers vorgenommen und mit Referenzwerten der AG Bhakdi / Valeva (Institut für Medizinische Mikrobiologie und Hygiene, Uniklinikum Mainz) verglichen. Der übliche Titer, bei dem es zu einer Lyse von Kaninchenerythrozyten kommt, ist etwa  $10^{-8}$  bis  $10^{-9}$  M<sup>15</sup>. Die Kaninchenerythrozyten wurden von der AG Valeva bezogen.

Für die eigentliche Bestimmung des Titers wurde PBS-Puffer (siehe Tabelle 3.10 auf Seite 46) in Mikrotiterplatten mit Rundboden vorgelegt. 10 – 20  $\mu\text{g}$  des Toxins wurden anschließend in eine der Vertiefungen gegeben. Hiervon ausgehend wurde eine Verdünnungsreihe in anderen Vertiefungen der Mikrotiterplatte angefertigt. Eine Negativkontrolle mit PBS-Puffer ohne Zusatz des  $\alpha$ -Hämolytins existierte ebenfalls. Zur Verdünnungsreihe wurde eine 5 %ige Erythrocytenlösungen gegeben (die Stammlösung der Erythrocyten gilt hierbei als 100 %ige Lösung).

Nach einer Inkubation von 30 – 60 Minuten im Wärmeschrank bei  $37^\circ\text{C}$  konnte der Titer durch den Lysegrad der Erythrocyten visuell geschätzt werden.

#### 3.10.4.2 Reinigung der Rundkolben zur Lipidpräparation

Um sicherzustellen, dass jeder Rundkolben zur Präparation der Lipide den größtmöglichen Anforderungen an die Sauberkeit entspricht, wurde jeder Kolben nach jeder Präparation folgendem Spülprotokoll unterzogen:

1. Spülen mit Wasser
2. Spülen mit Chloroform
3. Spülen mit Aceton
4. In 2% (v/v) Hellmanex-Lösung inkubiert, davon 30 Minuten im angewärmten Ultraschallbad (ca.  $50^\circ\text{C}$ )
5. Ausspülen mit Wasser und abschließendes lufttrocknen

Um sicherzustellen, dass nach dem letzten Schritt keine Detergensreste neue Präparationen im Kolben stören, wurden die Kolben unmittelbar vor dem Gebrauch nochmals mit etwas Chloroform ausgespült.

---

<sup>15</sup>Persönliche Mitteilung A. Valeva

### 3.10.4.3 Präparation von Lipiden

Alle Lipidvesikel für die hier dargestellten Experimente hatten folgende Zusammensetzung:

Tabelle 3.7: *Lipidzusammensetzung der verwendeten Vesikel*. Die Prozentangaben stehen für das molare Verhältnis.

Komponente	Gehalt [%]	Sigma-Produktnummer
Cholesterol	30	C8667
L- $\alpha$ -Phosphatidylethanolamin	15	P6386
1,2-Diacyl- <i>sn</i> -glycero-3-phospho-L-serin	15	P6641
N-acyl-4-sphingenyl-1-O-phosphorylcholin*	40	S0756

\* hauptsächlich 16:0

Alle Lipide wurden in Chloroform gelöst, sofern diese nicht bereits in Chloroform geliefert wurden. Die Lipide wurden in kleine Rundkolben gefüllt, gemischt und anschließend unter einem Stickstoffstrom getrocknet. Eine Trocknung dieser Lipidpräparate erfolgte über Nacht im Vakuum ( $\leq 5 \cdot 10^{-1}$  mbar) bei 50 °C im Sandbad.

Entweder wurden die getrockneten Lipide nun unter Argon gelagert oder direkt in Puffer aufgenommen und zu Vesikeln weiterverarbeitet. Die Puffer werden in den entsprechenden Abschnitten der Resultate angegeben.

### 3.10.4.4 Herstellung von Vesikeln

Zur Herstellung von Vesikeln wurden die Lipide (siehe voriger Abschnitt) mit Puffer überschichtet. Die Kolben wurden anschließend so lange gevortext bis sich sämtliche Lipide vom Glas gelöst hatten. Hiernach wurde das Lipid-Puffer-Gemisch mittels eines Extruders (Avanti Polar Lipids, Alabaster, AL, Typ 610000) durch eine Membran (Avanti Polar Lipids, 50 nm Durchmesser: Typ 800308; 100 nm Durchmesser: Typ 800309; Filterunterlage: Typ 610014) gepresst, um unilamellare Vesikel einer Größe herzustellen. Um sicherzustellen, dass auch eine Phasentrennung der Lipide zur verbesserten Interaktion der Toxine mit der Membran stattfindet, fand die Extrusion bei 50 °C statt. Hierzu wurde der Extruder auf eine Heizplatte gestellt und es wurde darauf geachtet, dass die Lipide möglichst kurze Zeit der Umgebungstemperatur ausgesetzt waren. Nach der Extrusion wurden die Vesikel dann langsam auf 20 °C im Wasserbad (Thermo Haake, Karlsruhe, Typ C25P) abgekühlt. Die Abkühlung erfolgte, startend von 50 °C, in einem linearen Gradienten für 12 Stunden.

Das Protokoll lehnt sich an ein älteres Protokoll der AG Valeva / Bhakdi (Institut für Medizinische Mikrobiologie und Hygiene, Uniklinikum Mainz) an (Valeva *et al.*, 1995) und wurde für unsere Arbeitsgruppe angepasst: Die Ultraschallbehandlung des ursprünglichen Protokolls wurde nicht durchgeführt.

### 3.10.4.5 Bestimmung des Cholesterolgehaltes von Vesikeln

Zur Bestimmung der Lipidkonzentration einer Emulsion von Vesikeln kann man u. a. den Cholesterolgehalt in einer gegebenen Lösung bestimmen. Bei bekanntem relativem Cholesterolgehalt ist so auf die Gesamtlipidkonzentration zurückzuschließen.

Für diese Bestimmung wurde das CHOD-PAP-Assay (Roche Diagnostics, Mannheim, Produktnummer 12016630 122) verwendet. Es basiert auf der Bestimmung des Gehaltes von  $\Delta^4$ -Cholesteron nach enzymatischem Abbau des Cholesterols durch eine Cholesterolesterase (Allain *et al.*, 1974; Roeschlau *et al.*, 1974). 1 mL der CHOD-PAP-Reagenz des Assays wurden in eine Küvette pipettiert und der Leerwert der Extinktion wurde bei 500 nm bestimmt. 10  $\mu$ l der zu testenden Liposomenpräparation wurden hinzugegeben und bei 15 °C für 15 Minuten inkubiert. Anschließend

wurde erneut die Extinktion bei 500 nm bestimmt. Der Cholesteringehalt ergibt sich aus

$$c[mM] = 14,9 \cdot (\text{OD}_{\text{Leerwert}} - \text{OD}_{\text{Endwert}}) . \quad (3.14)$$

Die Messungen fanden bei Raumtemperatur statt. Der Koeffizient (14,9) entstammt den Angaben von Roche Diagnostics.

### 3.10.5 Röntgenkleinwinkelstreuung am Monomer des $\alpha$ -Hämolytins

Im Wesentlichen wurde hier dasselbe Vorgehen gewählt wie bereits beim Hämocyanin von *E. californicum* unter SDS-Einfluss (siehe Abschnitt 3.9.1 auf Seite 44): Auch hier ist das Protein in Lösung nicht sehr stabil und diesem Umstand musste Rechnung getragen werden. Die Messintervalle wurden auf 30 Minuten beschränkt. Der Aggregationsneigung des Proteins wurde dadurch entgegengewirkt, indem sämtliche Schritte entweder auf Eis oder bei sehr niedriger Temperatur stattfanden:

1. Dem gefriergetrockneten Protein wurde eiskalter Puffer hinzugegeben und die Lösung wurde vorsichtig mehrfach auf- und abpipettiert, um sämtliches Protein in Lösung zu überführen.
2. Anschließend wurde die Lösung zentrifugiert, um bereits entstandene Aggregate zu entfernen. Dieser Schritt erfolgte 15 Minuten lang bei 1 °C und 22000 x g<sup>16</sup> (HERMLE Labor-technik, Wehingen, Typ Z233 MK-2).
3. Die Lösung wurde daraufhin auf Eis zur Kratky-Kamera überführt und schnellstmöglich in die vorgekühlte Kapillare überführt.
4. Die eigentlichen Messungen erfolgten bei 0,5 °C.

### 3.10.6 Neutronenkleinwinkelstreuung am $\alpha$ -Hämolytins

#### 3.10.6.1 Bestimmung des Matchpunktes

Vor der Durchführung des Neutronenkleinwinkelstreuexperimentes mit der H35R-Mutante (siehe Abschnitt 1.4 in der Einleitung auf Seite 13) des  $\alpha$ -Hämolytins stand die Überlegung, dass der Kontrast der Lipide minimiert werden muss. Auf diese Weise sollte nur noch das Protein zu den Streudaten beitragen (siehe z. B. Bu *et al.*, 2003).

Sämtliche Messungen wurden am SANS-1 Strahlrohr des FRG-1 Reaktors der GeNF an der GKSS durchgeführt (Stuhrmann *et al.*, 1995; siehe Abschnitt 2.1.2 auf Seite 17 des Theoriekapitels). Die Proben wurden in Quartzküvetten von 1 mm Dicke (Firma Hellma, Mühlheim) gefüllt und in einen temperierten Probenhalter gestellt. Für die Bestimmung des Matchpunktes war es wichtig, dass der Durchmesser der Liposomen  $\leq 50$  nm blieb. Daher wurde die Temperatur auf relativ kühle 20 °C  $\pm$  0,5 °C eingestellt. Die Liposomen wurden wie unter Abschnitt 3.10.4.4 beschrieben hergestellt, wobei hier eine Membran mit 50 nm Porengröße verwendet wurde.

Es wurden Puffer mit einem Deuteriumgehalt von 99,2 %, 60,3 %, 40,7 %, 17,2 % und 0,15 % verwendet. Von ganzen, „glatten“ Zahlen abweichende Beträge beruhen auf der Berücksichtigung des in allen Puffersubstanzen (Wasser, Phosphat / Hydrogenphosphat) enthaltenen Deuteriums.

Die Streudaten wurden für den Hintergrund des Puffers, der Probenzelle und anderer Quellen korrigiert (Wignall & Bates, 1986). Vier Proben-Detektor-Abstände zwischen 9700 mm und 700 mm wurden eingestellt, so dass ein q-Bereich von 0.004 bis 0.24 Å<sup>-1</sup> abgedeckt war. Die Wellenlänge der Neutronen betrug 8 Å mit einer Wellenlängenauflösung von 10 % (Halbwertsbreite der Wellenlängenverteilung). Für jedes Lipid-Puffer-Gemisch wurden der Cholesterolgehalt bestimmt (siehe Abschnitt 3.10.4.5). Die Streudaten wurden mit dem so gewonnenen Normalisierungsfaktor für die Lipidkonzentration korrigiert und die Streuung bei I(0) wurde mittels Guinier-Analyse

<sup>16</sup>Eine höhere Zentrifugalgeschwindigkeit wäre wünschenswert gewesen, aber aufgrund des kleinen Volumens und der Notwendigkeit eines schnellen Transfers war dies nicht möglich.

(Glatter & Kratky, 1982) bestimmt. Wird die Wurzel der Streuung der Nullintensität ( $I(0)$ ) gegen den Deuteriumgehalt des Puffers aufgetragen (siehe Abschnitt 4.4.3.2 auf Seite 92), so ergibt sich die Konzentration an Deuterium für die der Kontrast des Lipidgemisches gleich Null ist. Puffer für die Neutronenkleinwinkelstreuung (im nächsten Abschnitt) wurden auf diese Deuteriumkonzentration eingestellt.

### 3.10.6.2 Durchführung der Kleinwinkelstreuung

Die in Abschnitt 2.1.2 (Seite 17) vorgestellte Experimentiereinrichtung der GKSS wurde in Zusammenarbeit mit Vasil Gharamus<sup>17</sup> durchgeführt. Streudaten wurden wie im vorigen Absatz beschrieben aufgenommen und für die verschiedenen Detektorabstände zusammengefasst. Die Auswertung erfolgte mit einer Programmsuite von O. Glatter (Glatter, 1977) unter Annahme verschiedener geometrischer Körper, da eine referenzfreie Auswertung nicht mehr möglich schien (vgl. die Diskussion auf Seite 138).

### 3.10.6.3 Berechnung der Neutronenstrelängendichte

Die Neutronenstrelängendichte (nSLD) eines Proteins ist definiert als die Summe der kohärenten Streulängen über alle Atome innerhalb eines gegebenen Volumens  $\delta V$ , geteilt durch  $\delta V$  (Harrown *et al.*, 2007) bzw.

$$\rho(r)\delta V = \sum_i b_i . \quad (3.15)$$

Hierbei ist  $b$  die Streulänge kovalent gebundener Atome. Der Laufindex  $i$  geht über alle Atome eines Proteins, inklusive der Wasserstoffatome.

Die Möglichkeit der Berechnung wurde in meine Softwarebibliothek (`sas-rigid`) zur *rigid body*-Modellierung (siehe Abschnitt 4.7 auf Seite 105) integriert.

Die Werte für  $b$  wurden Tabellenwerken entnommen (Engelmann & Moore, 2001; Harrown *et al.*, 2007). Da der Fehler der Volumenschätzung in jedem Fall wesentlich größer ist als der Fehler für  $b$ , wurden Fehler für  $b$  nicht berücksichtigt. Überdies berücksichtigt kein bekanntes Programm und auch der eigene Algorithmus nicht die Lösungsmittelzugänglichkeit aller Reste der Aminosäuren eines Proteins, so dass grundsätzlich davon ausgegangen wird, dass sämtliche austauschbaren Wasserstoffatome im Gleichgewicht mit dem Lösungsmittel stehen. Dies gilt insbesondere auch in Anbetracht extrem langer Wartezeiten (mehrere Tage) von Proben nach ihrer Präparation und dem Beginn der Messung an einem Reaktor.

### 3.10.7 Berechnung der Massenverteilung innerhalb geometrischer Körper

Die für das  $\alpha$ -Hämolyisin gewonnenen Neutronenstreudaten wurden unter der Annahme geometrischer Körper (Scheiben oder Zylinder) ausgewertet (siehe Abschnitt 3.10.6.2). Die Auswertung erfolgte durch Vasil Gharamus (GKSS). Es gilt für die Abstandsverteilungsfunktion bzgl. der Schnittfläche eines stäbchenförmigen Aggregates (cross-section pair distribution function oder  $\rho_{cs}(r)$ ):

$$\frac{d\Sigma(q)}{d\Omega} = \frac{2\pi^2}{q} \int_0^\infty \rho_{cs}(r) J_0(qr) dr , \quad (3.16)$$

wobei  $J_0$  eine Besselfunktion nullter Ordnung ist (Glatter, 1977; Pedersen, 1997). Die Masse pro Längeneinheit des stäbchenförmigen Aggregates ergibt sich dann aus

$$\rho_{cs}(r) = \frac{c}{2\pi M_L} \int r \Delta\rho(r') \Delta\rho(r+r') dr' , \quad (3.17)$$

wobei  $\Delta\rho(r)$  für den Kontrast (siehe Abschnitt 2.2.5 zum Thema Kontrast im Theorieteil auf Seite 22) steht und die Vektoren  $r$  und  $r'$  in der Schnittflächenebene liegen (Glatter, 1977). Analog zu

<sup>17</sup>Dies ist die phonetische Schreibweise, in englischsprachigen Referenzen wird von ihm größtenteils die Schreibweise „Haramus“ verwendet.

Gleichung 2.42 im Theorieteil (Seite 25) ergibt sich ein Gyrationradius für den Querschnitt einer Mizelle beziehungsweise eines Liposoms (Pedersen, 1997):

$$R_{cs,g}^2 = \frac{\int_0^\infty r^2 \rho_{cs}(r) dr}{2 \int_0^\infty \rho_{cs}(r) dr} . \quad (3.18)$$

Die Streuintensitäten von scheibenförmigen Aggregaten werden hingegen in Form der Dicke der Abstandsverteilungsfunktion  $\rho_T(r)$  ausgedrückt:

$$\frac{d\Sigma(q)}{d\Omega} = \frac{2\pi^2}{q^2} \int_0^\infty \rho_T(r) \cos(qr) dr . \quad (3.19)$$

$\rho_T(r)$  ist durch

$$\rho_T(r) = \frac{c}{2\pi M_S} \int r \Delta\rho(r') \Delta\rho(r+r') dr' \quad (3.20)$$

gegeben, wobei  $r$  der Vektor in „x-Richtung“ und  $M_S$  eine Masse pro Flächeneinheit ist. Die Berechnung des entsprechenden Gyrationradius  $R_{T,g}$  und von  $M_S$  erfolgt analog der Berechnung bei stäbchenförmigen Aggregaten.

Wie aus Gleichung 3.17 erkennbar ist gibt es einen maximalen Wert für  $r'$ : Ein Stab – oder eine Scheibe – hat eine gegebene Dicke, die nicht überschritten werden kann. Diese Dicke kann als die korrespondierende Dicke der zugrundeliegenden Membran angesehen werden.

### 3.10.8 Errechnen und Darstellung der Sequenzkonservierung

Nach Schneider & Stephens, 1990 kann der Grad der Konservierung an bestimmten Positionen innerhalb eines Alignments,  $R_{seq}$ , als Differenz der maximalen Entropie ( $S_{max}$ ) und der Entropie an der beobachteten Position  $S_{beobachtet}$  aufgefasst werden:

$$R_{seq} = S_{max} - S_{beobachtet} = \log_2 N - \left( - \sum_{n=1}^N p_n \log_2 p_n \right) . \quad (3.21)$$

Hierbei ist  $p_n$  die beobachtete Häufigkeit eines Symbols  $n$  an einer gegebenen Position in der Sequenz und  $N$  ist die Zahl der verschiedenen Symbole für eine bestimmte Art von Sequenz, entweder 4 bei DNA/RNA-Sequenzen oder 20 bei Proteinsequenzen. Folglich ist die maximale Sequenzkonserviertheit bei Proteinen  $\log_2 20 \approx 4,32$  bit (weitere Ausführungen finden sich auch in Crooks *et al.*, 2004.).

Mittels `Pipealign` (Plewniak *et al.*, 2003) wurden insgesamt 107 Sequenzen der Proteinfamilie des  $\alpha$ -Hämolytins ermittelt. Ein eigenes Skript (Erläuterung im Anhang A.5) entfernte alle Sequenzen aus dem Alignment, die entweder doppelt vorhanden oder in einer anderen Sequenz enthalten waren (was der Fall sein kann, wenn Vorläufersequenzen in der Datenbank enthalten sind). Das Skript errechnete auf Basis der restlichen 75 Sequenzen den Informationsgehalt in Bezug auf die Sequenzpositionen des  $\alpha$ -Hämolytins. Diese Werte wurden in ein Histogramm mit 10 Einheiten unterteilt und zur Darstellung in 3D-Molekülabbildungen (siehe Seite 90 ff.) verwendet.

### 3.11 Hitzedenaturierung mittels *Differential Scanning Calorimetry*

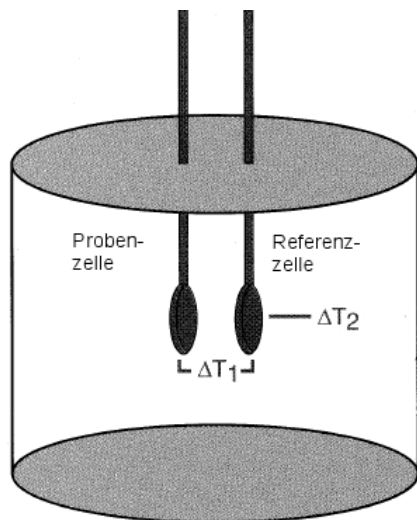


Abbildung 3.2: Schematische Darstellung des DSC-Gerätes von MicroCal. Abbildung Hübler, 2001

Eine Hitzedenaturierung zur Bestimmung des Schmelzpunktes wurde mit einem VP-DSC Kalorimeter (MicroCal, Northampton, MA; Typ *VP-DSC*) durchgeführt. Das Kalorimeter ist hochsensitiv und erlaubt die Verwendung kleiner Probenmengen. Es besitzt zwei münzförmige Zellen aus Tantalum 61<sup>TM</sup>, einer besonders hitzeresistenten Tantalverbindung, mit einem Probevolumen von 0,5 mL. Die Zellen sind durch einen adiabatischen Mantel umgeben. Die linksstehende Abbildung zeigt ein Schema des Aufbaus.

Die Temperaturdifferenz zwischen Proben- und Referenzzelle wird durch ein Peltierelement bestimmt. Ein weiteres Peltierelement ermittelt die Differenz zwischen den Zellen und dem umgebenden Mantel. Werden die Zellen aufgeheizt und kommt es zu einer temperaturinduzierten Reaktion in der Probenzelle (beispielsweise einem Phasenübergang), so wird die Temperaturdifferenz zwischen den Zellen ( $\Delta T_1$ ) durch Anpassung des Heizstromes der Probenzelle so ausgeglichen, dass die Temperaturdifferenz möglichst gering ist. Die Wär-

memenge, welche während der Messzeit benötigt wird, um  $\Delta T_1 = 0$  zu halten, also das Integral über die Leistungsdifferenz während des Heizens, ist das Maß für die Wärmemenge der Reaktion in der Probenzelle.

Vor dem Experiment wurde Protein gegen einen Probenpuffer dialysiert. Die Proteinkonzentration betrug stets etwa  $1\text{mg} \cdot \text{mL}^{-1}$ . Vor jeder Messung wurde das gesamte Messprogramm ausschließlich mit Puffer in beiden Zellen, d. h. ohne Probe, durchgeführt, um eine Basislinie zu erhalten und Hystereseeffekte klein zu halten. Die bei den Probenscans erhaltenen Werte (in [cal/min]) wurden für den jeweiligen Pufferwert korrigiert und das Signal auf die Heizrate (in [cal/°C]) normalisiert. Dies erfolgte mit dem Program „Origin für DSC“.

Eine detaillierte Auswertung der gewonnenen Daten stand nicht im Vordergrund der Arbeit, da diese Experimente nur mit Blick auf die Schmelzpunktbestimmung für die entsprechenden Experimente mit der Kratky-Kamera (siehe Abschnitt 3.6.5 auf Seite 36) durchgeführt wurden. Es wurde nur die Entfaltungstemperatur des Proteines ( $T_m$ ), welcher der Temperatur am Peakmaximum entspricht, ermittelt.

### 3.12 Fluoreszenzspektroskopie

Fluoreszenzmissionsspektren (aufgezeichnet mit den Spektrometern von Hitachi, Tokyo, Japan; Typ *F4500* oder von HORIBA JobinYvon, München, Typ *SPEX<sup>®</sup> FluoroLog<sup>®</sup>-3*) wurden zur Vermeidung des Inner-Filter-Effektes (Schmid, 1994) bei Proteinkonzentration zwischen 0,05 und 0,1 mg/mL aufgezeichnet. Die Proben wurden in Halbmikroküvetten mit 10 mm optischer Weglänge (Helma, Mühlheim) gefüllt. Sämtliche Spektren wurden gegen die Streuung des Puffers korrigiert. Die Anregungswellenlänge beträgt 280 nm oder 295 nm (zur Aufnahme der Tryptophanfluoreszenz). Das Wellenlängeninkrement betrug 1 nm bei einer Filterbandbreite von 5 nm.

### 3.13 Circular dichroismusspektroskopie

Circular dichroismusspektroskopie bzw. CD-Spektroskopie diente zur Kontrolle der Sekundärstruktur und damit der Integrität des  $\alpha$ -Hämolytins (mit dem CD-Spektrometer von Jobin Yvon;

Typ *CD6*). Spektren wurden in Schritten von  $0,5^\circ$  von 191 – 260 nm bei  $20^\circ\text{C}$  aufgezeichnet. Die Integrationszeit für einen Datenpunkt betrug 1 s und die Bandbreite des Monochromators 2 nm. Die Proben wurden in spezielle Küvetten mit 1 mm Dicke (Helma, Mühlheim) gegeben. Insgesamt wurden je 20 Proben- und Pufferspektren aufgezeichnet. Das Wellenlängeninkrement betrug 0,5 nm.

Da das Spektrometer zum Zeitpunkt der Messungen nicht bei kleineren Wellenlängen messen konnte und nicht genügend Protein für Wiederholungen zur Verfügung stand, wurden die Daten nur für einen qualitativen Vergleich genutzt. Stattdessen wurden die Pufferspektren gemittelt und von den ebenfalls gemittelten Probenspektren abgezogen sowie die molare Elliptizität durch Normierung auf die Konzentration bestimmt.

### 3.14 Darstellung von Molekülgraphiken

Moleküle wurden entweder mit den Programmen DINO (Philippsen, 2001) oder YASARA (Krieger *et al.*, 2002) dargestellt. Während DINO für alle Moleküle zum Einsatz kam, die zusammen mit *ab initio*-Rekonstruktionen (siehe Abschnitt 3.6.10 dieses Kapitels) gezeigt wurden, konnte YASARA für alle übrigen Molekülgraphiken verwendet werden. Um eine einheitliche Darstellung zu gewährleisten, wurde für jede Serie von Abbildungen in denen DINO einmal zum Einsatz kam das Programm auch für alle anderen Abbildungen der Serie verwendet.

### 3.15 Darstellung von Dichten

Die Darstellung von Dichten (siehe Abschnitt 4.7.5 auf Seite 115 des Resultatekapitels zur Erstellung dieser Dichten) erfolgte mit Chimera (Pettersen *et al.*, 2004).

### 3.16 Darstellung von Phylogrammen

Phylogramme wurden auf Basis von Sequenzalignments (siehe Abschnitt 3.10.8 dieses Kapitels) mit dem Programm Dendroscope (Huson *et al.*, 2007) generiert. Eine graphische Anpassung erfolgte nur zur optimierten Darstellung von Farbe und Text.

### 3.17 Darstellung von Kleinwinkelstreuungsdaten mit SPlot

Die Erstellung aller Diagramme von Kleinwinkelstreuungsdaten erfolgte mit dem selbstgeschriebenen Programm SPlot (siehe Resultate, Abschnitt 4.8, Seite 117).

# Kapitel 4

## Resultate

### 4.1 Konformationsübergänge von Hämocyaninen

#### 4.1.1 Die Struktur 24merer Hämocyanine von *Eurypelma californicum* und *Pandinus imperator* unter Einfluss verschiedener Effektoren

##### 4.1.1.1 Reinigung und Kontrolle des Hämocyanins von *Eurypelma californicum* und *Pandinus imperator*

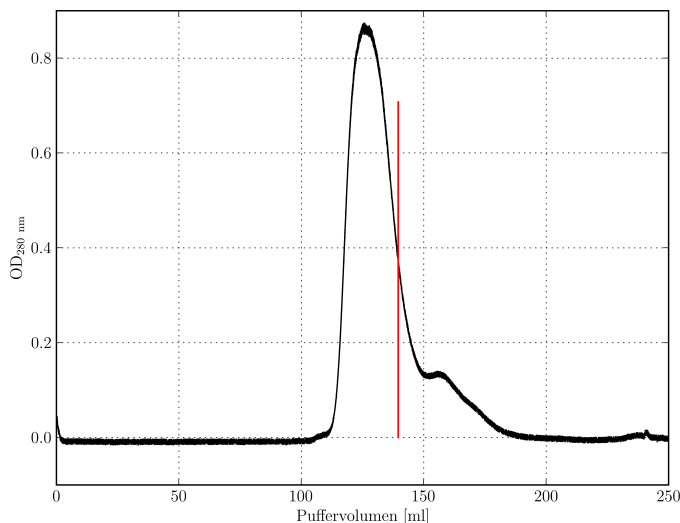


Abbildung 4.1: *Elutionsprofil einer Aufreinigung von P. imperator-Hämocyanin.* Der rote Strich markiert den Punkt, ab welchem kein Hämocyanin (erster Peak) mehr gesammelt wurde, weil es mit „Beiprotein“ (zweiter Peak) verunreinigt gewesen wäre.

ler können so als Referenz für die Dispersität und strukturelle Integrität der Proben herangezogen werden. Problematische Proben wurden verworfen.

Zusätzlich zur Überprüfung der Integrität durch die statische Lichtstreuung wurde getestet, inwieweit eine Hämocyaninprobe durch die Behandlung in der Kratky-Kamera in Mitleidenschaft gezogen wird. Hierzu wurden eine Probe für jedes der Hämocyanine in drei Teile aliquotiert: Ein Teil diente als unbehandelte Referenz, ein Teil wurde acht Stunden in der Kratky-Kamera bestrahlt und ein dritter Teil wurde mehrfach in der Probenkapillare der Kamera auf- und abpipettiert. Mit der Bestrahlung wird etwa die doppelte Dosis der Strahlung erreicht, die für die übrigen Experimente erreicht wurde. Das Pressen durch die Kapillare initiiert den sterischen Stress beim Befüllen der Probenkapillare der Kamera ebenfalls in übergroßer Weise: Für die eigentlichen Experimente wurde weitaus vorsichtiger gearbeitet. Mit allen Aliquots wurde ein

Die Aufreinigung der 24meren Hämocyanine erfolgte mittels Gelfiltration (siehe linksstehende Abbildung). Die auf die Gelfiltrationssäule aufgetragene Hämolymphe ist nicht rein, sondern enthält zusätzlich „Beiprotein“ (Conze, 2005). Für SAXS-Experimente musste sichergestellt werden, dass keinerlei „Beiprotein“ im gereinigten Hämocyanin enthalten ist. Zusammen mit B. Pairet konnte mittels statischer Lichtstreuung gezeigt werden, dass die verwendeten Hämocyaninlösungen von *E. californicum* und *P. imperator* auch nach längerer Lagerung ( $\geq 1/2$  a) homodispers sind (Pairet, 2007).

Da die Messung der statischen Lichtstreuung zeitaufwändig ist, wurde zusätzlich eine Aufnahme der dynamischen Lichtstreuung, mit den Proben aus der statischen Lichtstreuung als Referenz, durchgeführt (Pairet, 2007). Der hydrodynamische Durchmesser und der assoziierte Fehler

natives Gel gefahren. Zwischen den Proben konnte kein Unterschied festgestellt werden, so dass davon auszugehen ist, dass die Hämocyane von *E. californicum* und *P. imperator* durch die Behandlung zur Bestrahlung nicht in Mitleidenschaft gezogen werden.

#### 4.1.1.2 Oxygenierung und Deoxygenierung

Ebenfalls in der Diplomarbeit von B. Pairet (Pairet, 2007) wurde der Konformationswechsel der 24meren Hämocyane von *E. californicum* und *P. imperator* bei der Oxygenierung / Deoxygenierung mittels SAXS nachvollzogen. Die hier gezeigten Daten stellen ein Resümé der zitierten Diplomarbeit dar und wurden zu großen Teilen von B. Pairet unter meiner Anleitung generiert. Die Abbildungen wurden dort bereits gezeigt.

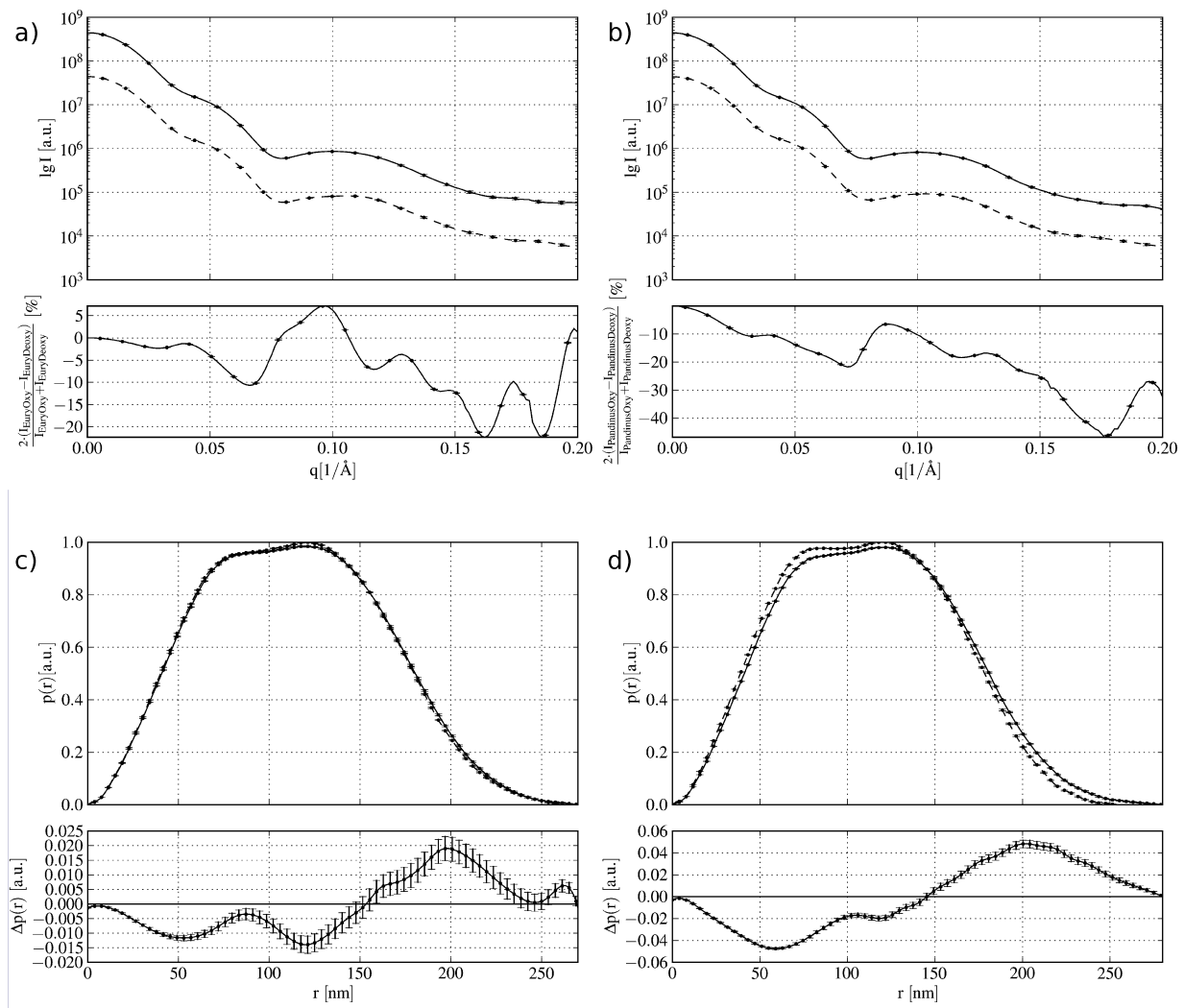


Abbildung 4.2: Stredaten der Hämocyane von *E. californicum* und *P. imperator* oxygeniert (—) und deoxygeniert (---). a) Entschmierte Stredaten von *E. californicum*-Hämocyanin. Die Kurven sind zur besseren Übersicht um eine logarithmische Einheit gegeneinander versetzt. b) Entschmierte Stredaten von *P. imperator*. Die Kurven sind ebenfalls um eine logarithmische Einheit gegeneinander versetzt. c) Abstandsverteilungsfunktionen von *E. californicum* Hämocyanin im oxygenierten und deoxygeniertem Zustand. d) Abstandsverteilungsfunktionen von *P. imperator* Hämocyanin. Unter den Daten befindet sich jeweils ein Differenzplot; für die Streukurven ist dieser prozentual nach Skalierung auf  $I(0)$  berechnet. Fehlerbalken sind exemplarisch angegeben. Die Teilabbildungen sind bereits in der Diplomarbeit von B. Pairet gezeigt (Pairet, 2007).

Tabelle 4.1 auf der Folgeseite zeigt die physikalischen Parameter zusammenfassend für beide Proteine und beide Konformationen. Es wird deutlich, dass sich die Konformationen der oxygenierten und deoxygenierten Hämocyane von *E. californicum* und *P. imperator* gleichen. Und

auch, dass das Hämocyanin von *P. imperator* den Konformationswechsel etwas „extremer“ durchzuführen scheint. Jedenfalls sind die Unterschiede zwischen oxygeniertem und deoxygeniertem Zustand signifikant größer als beim Hämocyanin von *E. californicum*.

Tabelle 4.1: *Physikalische Parameter für die oxygenierten und deoxygenierten Hämocyanine von E. californicum und P. imperator.* Referenzdaten früherer Messungen dienen der Übersicht.

Protein	Zustand	$R_g$ [Å]	$D_{max}$ [Å]	Quelle
<i>E. californicum</i> Hämocyanin	oxygeniert	$87,3 \pm 0,2$	$270 \pm 2$	eigene Daten*
	deoxygeniert	$86,6 \pm 0,1$	$260 \pm 2$	eigene Daten*
	oxygeniert	$88,0 \pm 0,5$	$270 \pm 5$	Decker <i>et al.</i> , 1996
	deoxygeniert	$86,5 \pm 0,5$	$250 \pm 5$	Decker <i>et al.</i> , 1996
<i>P. imperator</i> Hämocyanin	oxygeniert	$87,5 \pm 0,3$	$280 \pm 2$	eigene Daten*
	deoxygeniert	$85,6 \pm 0,2$	$254 \pm 2$	eigene Daten*

\* größtenteils von B. Pairet im Rahmen seiner Diplomarbeit unter meiner Anleitung aufgenommen (Pairet, 2007).

Auf Basis der Streudaten wurde versucht durch *rigid body*-Modellierung Strukturvorschläge für die oxygenierten und deoxygenierten Konformationen zu machen. Bereits in der Vergangenheit wurde eine solche Arbeit für das Hämocyanin von *E. californicum* durchgeführt. Das Programm `parmscan` von Dr. H. Hartmann wurde speziell für diesen Zweck entwickelt (Hartmann & Decker, 2002; Hartmann & Decker, 2004; siehe auch Abschnitt 3.6.11 auf Seite 39 des Methodenkapitels). Das Programm berücksichtigt, neben der zweifachen Rotationsymmetrie des Proteins, die verschiedenen weiteren Beziehungen der Untereinheiten. So berücksichtigt es gesondert die Position der Trimere innerhalb der Hexamere und die Position der Hexamere innerhalb der Dodekamere. Auch wird die relative Position der Dodecamere zueinander als eigener Parameter berücksichtigt. Das Programm ist also hochspezialisiert für *rigid body*-Modellierung an 24meren Hämocyaninen. Ursprünglich wurde ausschließlich die  $P(r)$  ausgewertet.

Die Version des Programms, die B. Pairet und mir zur Verfügung stand, wertete nicht nur  $P(r)$ , sondern auch  $I(q)$  aus. Da in einem Lauf von `parmscan` hunderttausende Konformationen durchlaufen werden und das Programm die Zwischenergebnisse in einer Logdatei speichert, habe ich ein Skript zur Auswertung geschrieben (siehe Abschnitt A.2 auf Seite 163).

Die zusätzliche Berechnung der theoretischen Streudaten ( $I(q)$ ) führte zu Unstimmigkeiten: Für die Übereinstimmung hinsichtlich der  $P(r)$  wurden andere Werte ausgegeben als für die Streudaten. Dies führte dazu, dass, orientierte man sich an der  $P(r)$  eine gänzlich andere Konformation als passend zu den experimentellen Daten erachtet wurde, als bei den theoretischen Streudaten im Vergleich zu den experimentellen. In Abschnitt 4.7.4 auf Seite 4.7.4 dieses Kapitels wird der Versuch einer Erklärung unternommen.

#### 4.1.1.3 Der Einfluss von SDS auf die Konformation

Zum Einfluss von SDS auf die 24meren Hämocyanine von *E. californicum* und *P. imperator* gibt es im Institut für Molekulare Biophysik eine Vielzahl von Experimenten, da SDS als *in vitro*-Aktivator der Phenoloxidaseaktivität dieser Proteine identifiziert wurde (Decker & Rimke, 1998). Dies erlaubte eine ganze Reihe von kinetischen und biochemischen Charakterisierungen der Enzymeigenschaften (u. a. Decker *et al.*, 2001; Jaenicke & Decker, 2004a; Decker & Jaenicke, 2004; Nillius, 2007). Eine strukturelle Charakterisierung der Aktivierung durch SDS setzte aber den Nachweis der Stabilität der Proteine in SDS voraus, der unsere Gruppe in Kollaboration gelang (Baird *et al.*, 2007).

Die Bestimmung der cmc-Werte für CPC und SDS in den für die SAXS-Messungen verwendeten Lösungen (siehe Tabelle 4.2) mit Hilfe der Tröpfchenmethode (siehe Abschnitt 3.9.3 auf Seite 45)

sollte sicherstellen, dass eine Detergenzkonzentration gewählt wurde, bei der alle Hämocyaninmoleküle in einer Konformation vorliegen. SAXS-Messungen mit CPC waren nicht möglich, da CPC zusammen mit den Hämocyaninen von *E. californicum* und *P. imperator* große unspezifische Aggregate formte. Die folgende Tabelle zeigt die ermittelten Werte und belegt, dass mit der gewählten Konzentration von 5 mM SDS im Puffer die cmc um das 4 – 5-fache überschritten ist. Damit ist eine Konzentration gewählt, bei der die enzymatische Aktivität maximal und die Proteine immer noch stabil sind (Baird *et al.*, 2007).

Tabelle 4.2: Die kritische Mizellare Konzentration für CPC und SDS in den Puffern zur Aufnahme der Röntgenkleinwinkelstreuung mit Hämocyanin.

Detergens	Puffer	cmc [mM]
CPC	100 mM TRIS, pH 7,8 10 mM MgCl <sub>2</sub>	0,83
CPC	100 mM TRIS, pH 7,8 5 mM MgCl <sub>2</sub>	0,83
SDS	100 mM TRIS, pH 7,8 10 mM MgCl <sub>2</sub>	0,68
SDS	100 mM TRIS, pH 7,8 5 mM MgCl <sub>2</sub>	0,76

Zur Klärung der Frage, ob SAXS-Daten mit SDS verwendet werden können, musste zunächst sichergestellt sein, dass ein Datensatz von SDS und Hämocyanin überhaupt ausgewertet werden kann. Die folgende Abbildung (4.3) zeigt die Pufferzusammensetzung (5 mM SDS, 5 mM MgCl<sub>2</sub>, 100 mM TRIS, pH 7,8) die sicherstellt, dass das Protein maximale Aktivität bei größtmöglicher Stabilität aufweist.

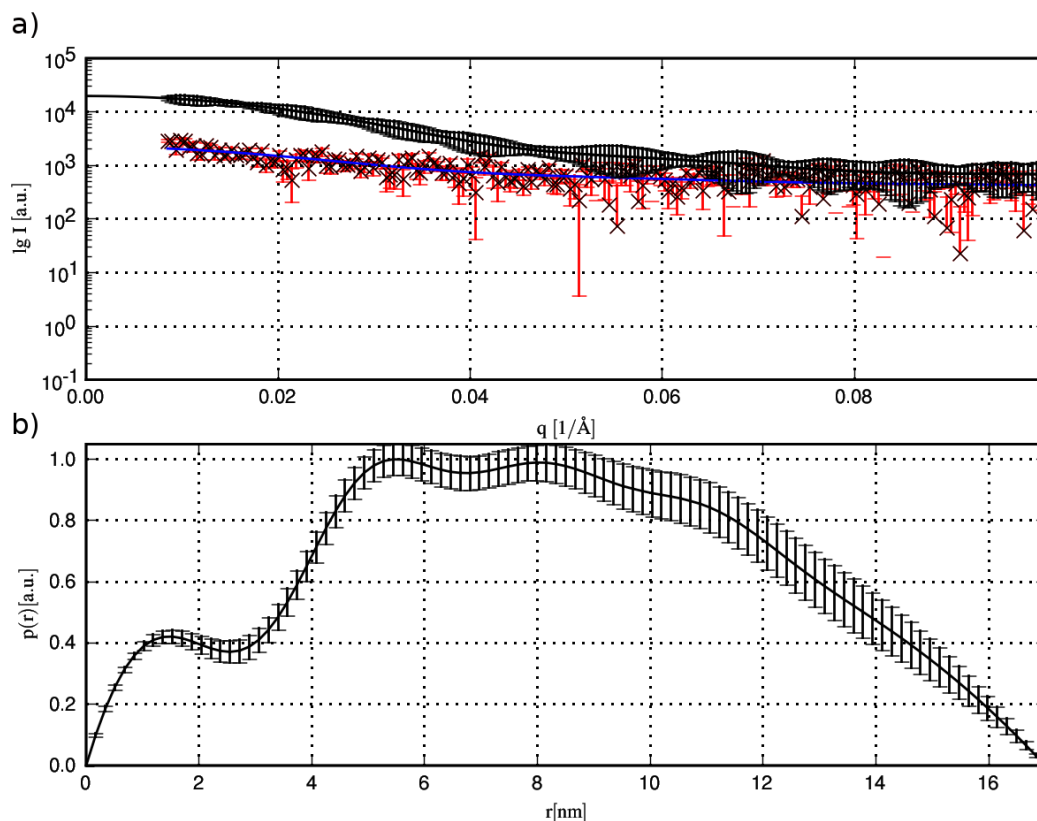


Abbildung 4.3: Streudaten für 5 mM SDS. a) Die primären Daten ( $\times$  mit roten Fehlerbalken), der apparente Fit durch diese Daten (blaue Linie) und die entschmierten Daten (schwarze Linie); jeweils mit exemplarischen Fehlerbalken. b) Die Abstandsverteilungsfunktion mit exemplarischen Fehlerbalken.

Die Entschmierung (siehe Abschnitt 3.6.4.2, Seite 3.6.4.2) konnte am besten unter der Annahme einer homodispersen Lösung durchgeführt werden. Deutlich zu erkennen ist der große experimentelle Fehler der Streudaten in Abbildung 4.3, aus dem auch der große Fehler der Abstandsverteilungsfunktion resultiert. Der Wert für den  $D_{max}$  von etwa 17 nm ist äußerst unsicher, sehr wahrscheinlich sind die SDS-Aggregate wesentlich kleiner: Eine Entschmierung unter Annahme eines kleineren Wertes ließ sich dennoch besser<sup>1</sup> durchführen, als unter der Annahme größerer Werte. Die relativ geringe Intensität der Streuung durch die SDS-Aggregate führt zwar zu großen Unsicherheiten in der Bestimmung der Aggregatgröße, erlaubt aber einen standardgemäßen Pufferabzug (siehe Abschnitt 3.6.4.1 auf Seite 34 des Methodenkapitels). Eine Wiederholung des Experimentes mit längerer Strahlzeit wird statistisch sicherere Daten bringen, aber da dies nicht wesentlich war, um das Streuexperiment an Hämocyaninen unter SDS-Einfluss durchzuführen, wurde darauf verzichtet.

Abbildung 4.4 zeigt die SAXS-Daten des Hämocyanins unter SDS-Einfluss im Vergleich mit Referenzdaten. Es wurden insgesamt drei Proteinkonzentrationen vermessen: 4,7 mg/mL, 9,4 mg/mL und 63,5 mg/mL, jeweils in Gegenwart von 5 mM SDS. Die zugehörigen Abstandsverteilungsfunktionen und die Kratky-Plots werden in Abbildung 4.5 auf der nächsten Seite gezeigt. Tabelle 4.3 fasst die Ergebnisse für beide Hämocyanine zusammen.

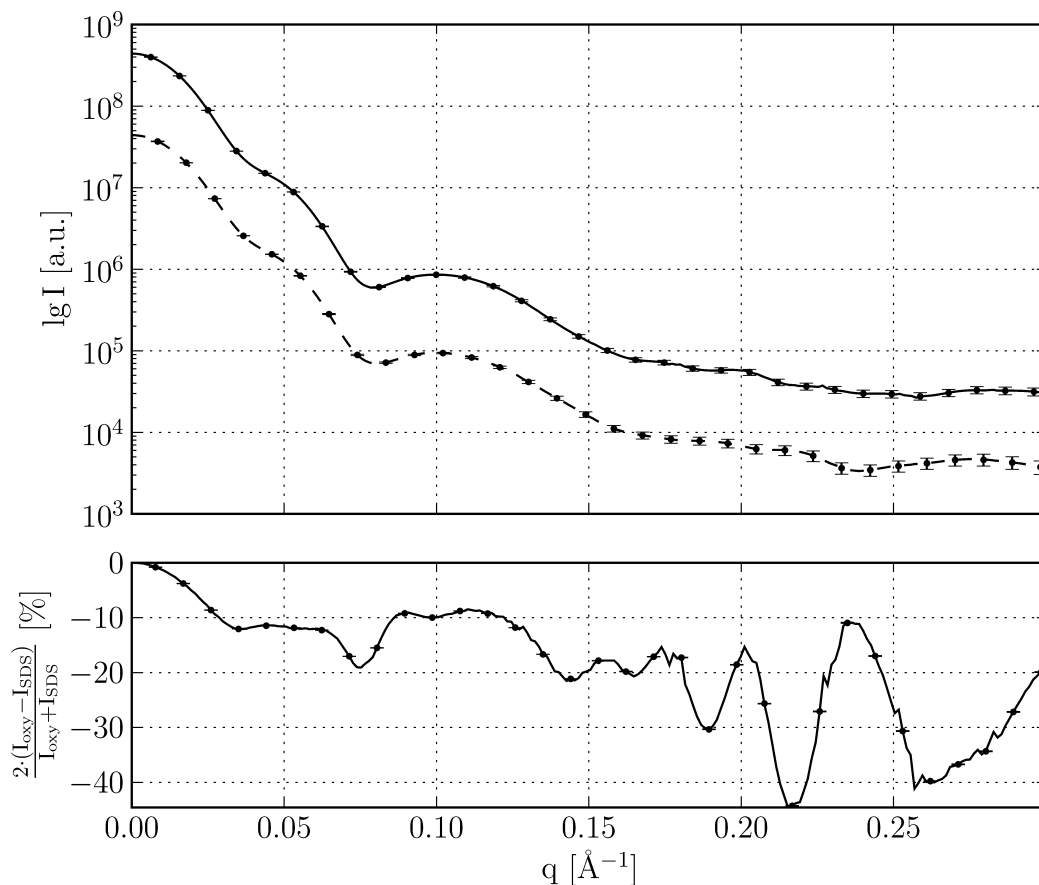


Abbildung 4.4: Streudaten des *E. californicum* Hämocyanins ohne SDS (—; mit „oxy“ bezeichnet) und mit 5 mM SDS (---). Die Daten sind zur besseren Übersichtlichkeit um eine logarithmische Einheit versetzt. Der untere Teil der Abbildung zeigt die prozentuale Differenz der Daten nach Skalierung auf  $I(0)$ .

<sup>1</sup>Keine Über- / Unterschwingungen der dekonvolvierten Streukurven, bessere Qualitätsparameter nach GNOM; siehe Abschnitt 3.6.4.2, Seite 34 im Methodenkapitel.

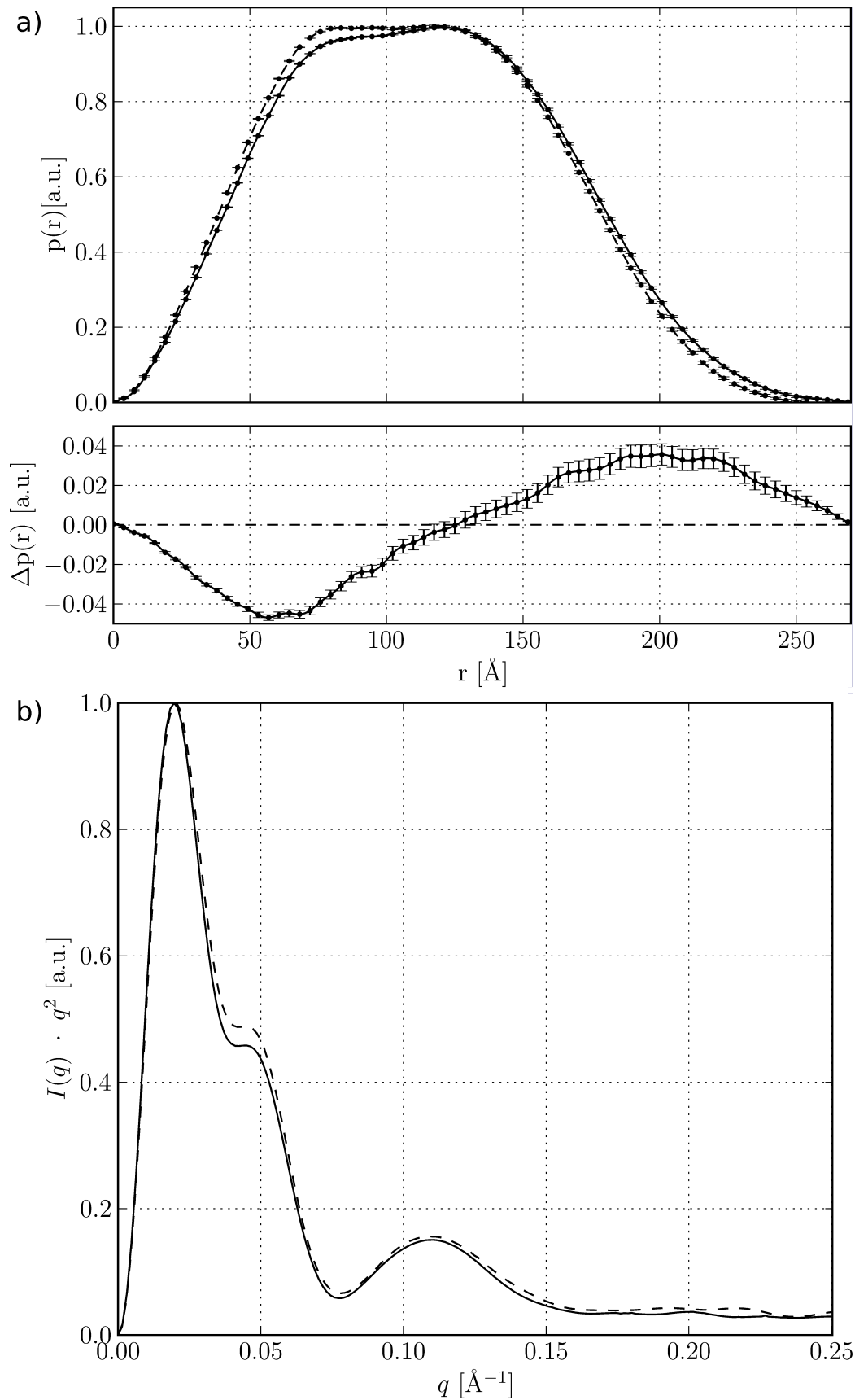


Abbildung 4.5: Daten des *E. californicum* Hämocyanins ohne SDS (—) und mit 5 mM SDS (---). a) Abstandsverteilungsfunktion mit Differenzplot. b) Kratky-Plot. Die Positionen der Maxima sind 0,01963 (ohne SDS) und 0,02053 (mit SDS), also nahezu identisch. Dies ist ein Indikator für nahezu gleiche Molekulargewichte (siehe Abschnitt 3.6.7, Seite 37). Abbildung a) ist bereits in Baird *et al.*, 2007, publiziert.

In gleicher Weise wurde mit dem Hämocyanin des Kaiserskorpions (*P. imperator*) verfahren. Obwohl die Langzeitstabilität in SDS geringer ist als die des *E. californicum* Hämocyanins (Baird *et al.*, 2007), konnten die Streudaten unter der Annahme einer homodispersen Lösung entschmiert werden. Es wurden insgesamt drei Proteinkonzentrationen vermessen: 7,1 mg/mL, 16,2 mg/mL und 36,9 mg/mL; jeweils in Gegenwart von 5 mM SDS.

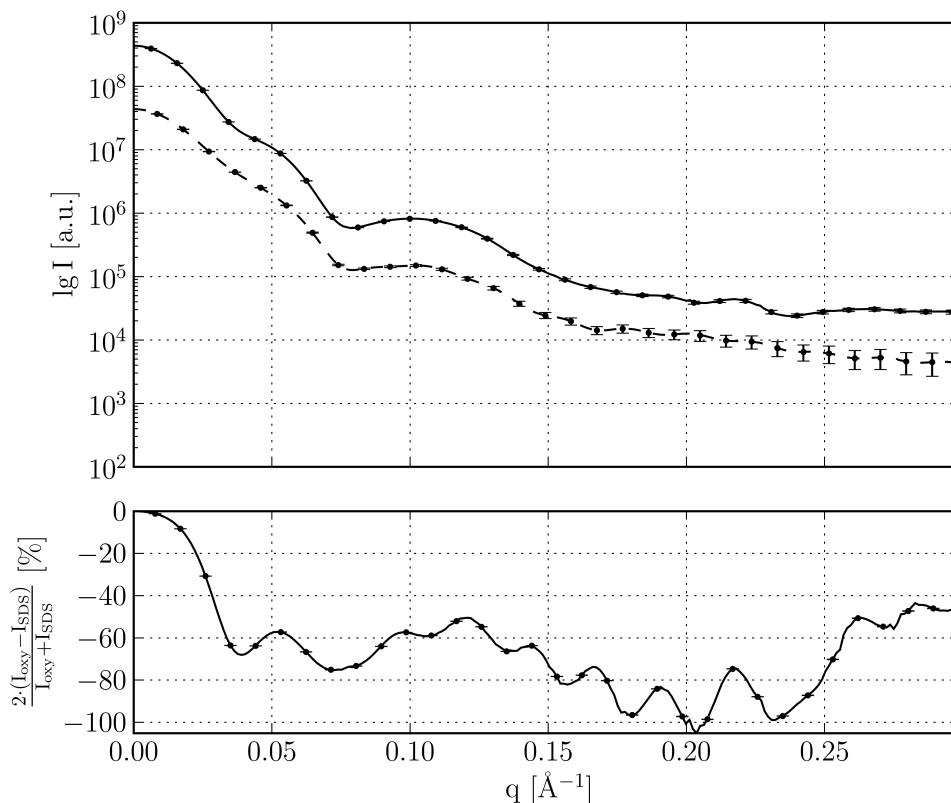


Abbildung 4.6: Streudaten des *P. imperator* Hämocyanins ohne SDS (—; mit „oxy“ bezeichnet) und mit 5 mM SDS (---). Die Daten sind zur besseren Übersichtlichkeit um eine logarithmische Einheit versetzt. Der untere Teil der Abbildung zeigt die prozentuale Differenz der Daten nach Skalierung auf  $I(0)$ .

Abbildung 4.7 auf der nächsten Seite zeigt die Abstandsverteilungsfunktion von *P. imperator*-Hämocyanin unter SDS-Einfluss sowie die zugehörigen Kratky-Plots. Die Stellung der beiden Dodekamere eines 24mers beeinflusst das Profil der Abstandsverteilungsfunktion maßgeblich (Hartmann & Decker, 2002; Hartmann & Decker, 2004). Eine systematische Rotation des einen Dodekamers im Verhältnis zum zweiten Dodekamer in Schritten von  $5^\circ$   $\varphi_{dode}$  führte jedoch zu keiner Annäherung an die experimentelle Abstandsverteilungsfunktion unter SDS-Einfluss in Abbildung 4.7.  $\varphi_{dode}$  ist hierbei der Kippwinkel zwischen den beiden Dodekamern nach Hartmann & Decker, 2002.

Tabelle 4.3: Physikalische Parameter für die Hämocyanine von *E. californicum* und *P. imperator* unter SDS-Einfluss. Referenzdaten anderer Messungen dienen dem Vergleich.

Protein	Zustand	$R_g$ [Å]	$D_{max}$ [Å]
<i>E. californicum</i> Hämocyanin	oxygeniert	$87,3 \pm 0,2$	$270 \pm 2$
	deoxygeniert	$86,6 \pm 0,1$	$260 \pm 2$
	mit SDS	$85,2 \pm 0,3$	$256 \pm 2$
<i>P. imperator</i> Hämocyanin	oxygeniert	$87,5 \pm 0,3$	$280 \pm 2$
	deoxygeniert	$85,6 \pm 0,2$	$254 \pm 2$
	mit SDS	$83,5 \pm 0,4$	$269 \pm 2$

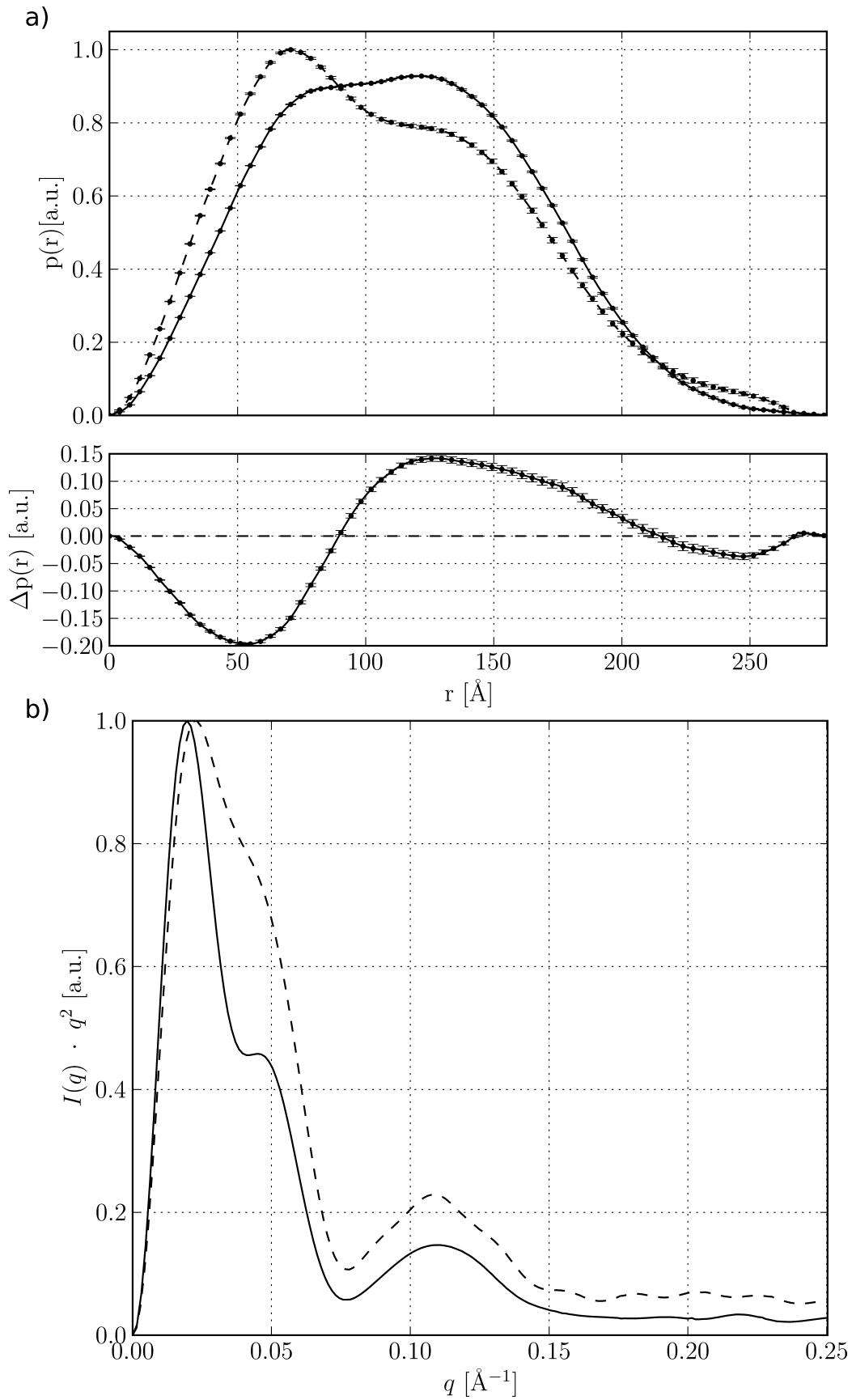


Abbildung 4.7: Daten des *P. imperator* Hämocyanins ohne SDS (—) und mit 5 mM SDS (---). a) Abstandsverteilungsfunktion mit Differenzplot. b) Kratky-Plot. Die Positionen der Maxima sind 0,01963 (ohne SDS) und 0,02317 (mit SDS).

## 4.1.2 Die Struktur des Hämocyanins von *Palinurus elephas* unter Einfluss verschiedener Effektoren

### 4.1.2.1 Modellierung der Struktur

Für die Modellierung des *P. elephas*-Hämocyanins wurde folgende Struktur nach sogenanntem *Profiling* der Proteindatenbank PDB (Berman *et al.*, 2000) durch MODELLER (Sali & Blundell, 1993) aus möglichen Musterstrukturen ausgewählt: Die Struktur des hexameren Hämocyanins von *Panulirus interruptus* (PDB-Code: 1HCY, Volbeda & Hol, 1989), das artifizielle Hexamer der Type II-Untereinheit des Hämocyanins von *Limulus polyphemus* in der oxygenierten (PDB-Code: 1OXY; Magnus *et al.*, 1994), der deoxygenierten Form (PDB-Code: 1LLA; Hazes *et al.*, 1993) und die Met-Form des Komplexes (PDB-Code: 1LL1; Liu *et al.*, 1997). Insgesamt also zwei Proteine mit vier Strukturen.

Die Modellierung erfolgte für eine Untereinheit des Hämocyanins mit der Sequenz nach T. Burmester (Meissner *et al.*, 2003)<sup>2</sup>. Das strukturelle Alignment ist in Abbildung 4.8 auf der folgenden Seite gezeigt.

Da das YAMBER-Kraftfeld (Krieger *et al.*, 2004) keine Parameter für Metallionen in variabler Koordination aufweist, wurde darauf verzichtet einen Strukturvorschlag für die oxygenierte Form des Proteins zu erstellen. Zudem wurde auch nur eine Untereinheit modelliert: Die Vorlagestrukturen stehen alle unter dem Einfluss unterschiedlicher Effektoren. Da davon auszugehen ist, dass diese, aufgrund leicht unterschiedlicher Konformationen die Kontaktflächen der Untereinheiten beeinflussen, wäre die Modellierung der Kontaktflächen höchst willkürlich. Stattdessen wurde eine D3-Symmetrie auf die modellierte Untereinheit appliziert und das Oligomer für die Dauer von 2 ps mit dem NOVA-Kraftfeld (Krieger *et al.*, 2002), ohne die Annahme expliziten Wassers, in YASARA relaxiert. Die Relaxierung sollte lediglich sicherstellen, dass es zu keinen sterischen Konflikten kommt.

Tabelle 4.4: Übersicht über die DOPE-Werte der Modellierung (jeweils bezogen auf ein Monomer).

Protein	DOPE-Wert
modellierte Monomere	-76923,6 – -73823,1
1hcy	-75643,0
1ll1	-76261,5
1oxy	-73733,8
1lla	-79263,2

Das beste Modell hat einen DOPE-Wert von -76923,6 und liegt damit im Bereich der Werte für die Musterstrukturen (siehe Tabelle 4.1.2.1, Seite 63). Dieser Wert reflektiert damit den Normbereich für die bisher bekannten Hämocyaninstrukturen.

*Beurteilung der Qualität:* Die Struktur ist in Abbildung 4.9 auf Seite 65 zu sehen. Es ist sichergestellt, dass die Struktur keine überlappenden van-der-Waals-Radien aufweist. Eine Kontrolle mit WhatIf (Vriend, 1990; Hooft *et al.*, 1996) zeigt eine Reihe von Problemen mit den Torsionswinkeln innerhalb der Peptidbindungen an, diese sind auch schon in den Musterstrukturen überwiegend schlecht gesetzt. Insgesamt gibt es 235 Punkte in „verbotenen“ Bereichen des Ramachandranplots (Ramachandran *et al.*, 1963) für das gesamte Protein von 3996 möglichen Verbindungen. Zusätzlich sind die B-Faktoren mit 10–20 relativ hoch (normal sind B-Faktoren zwischen 3 und 5 für kristallographische Strukturen, deren Daten unter Kryo-Bedingungen ermittelt wurden).

<sup>2</sup>Professor E. Jaenicke stellte mir diese Sequenz zur Verfügung.

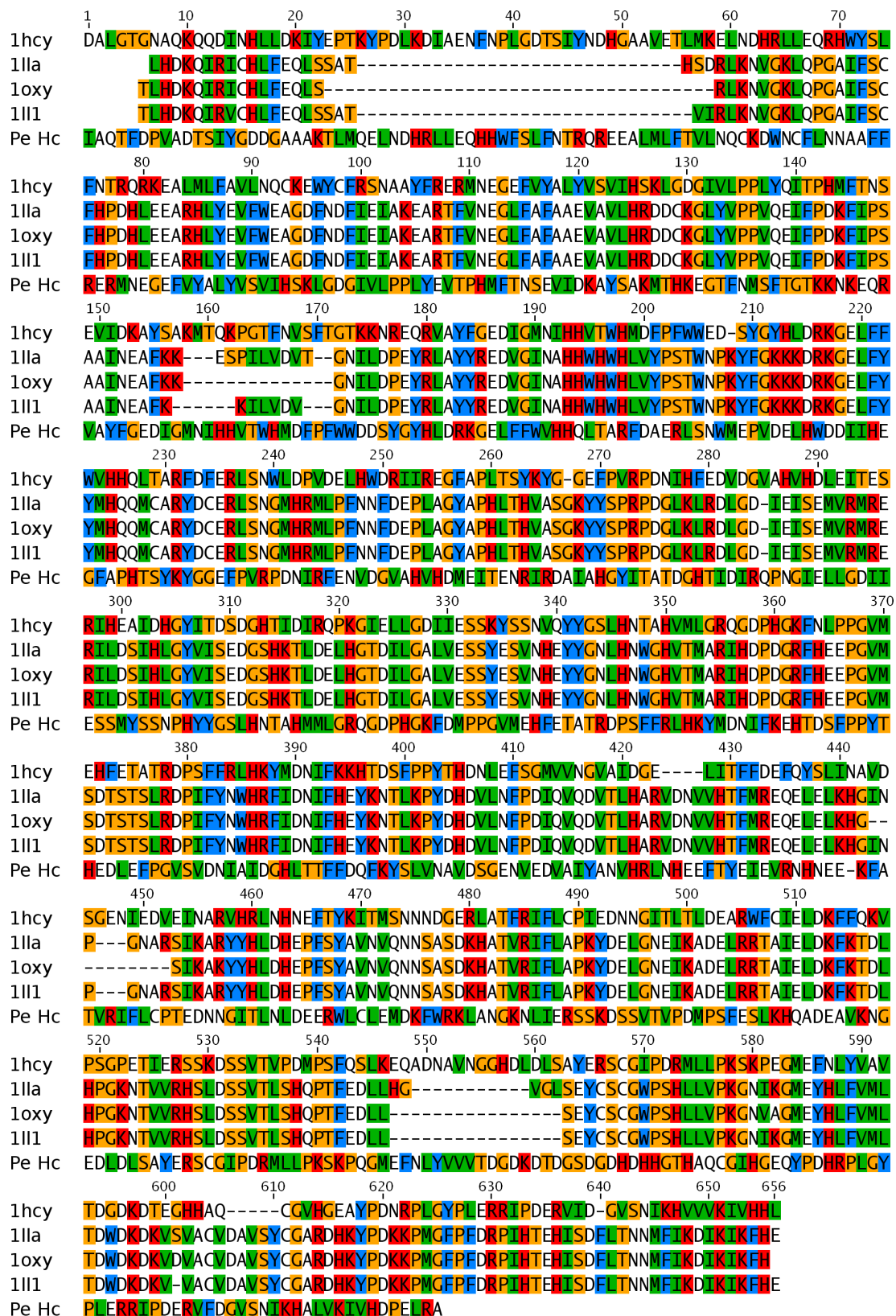


Abbildung 4.8: Ergebnis des strukturellen Alignments der zur Modellierung verwendeten Hämocyanine. Die linke Spalte zeigt die PDB-Codes der verwendeten Strukturen. Die Zeile „Pe Hc“ steht für die Sequenz des *P. elephas*-Hämocyanins. „-“ steht nicht zwangsläufig für eine Lücke im Sequenzalignment, sondern ebenso für eine Lücke in den Musterstrukturen. Die PDB-Codes 1LLA, 1OXY und 1LL1 sind Strukturen des Hämocyanins von *L. polyphemus* in verschiedenen Oxydationszuständen (siehe Text). Die Farben der Aminosäuren entsprechen dem CLUSTAL-Farbschema (Jeanmougin *et al.*, 1998) und dienen der besseren Unterscheidung.

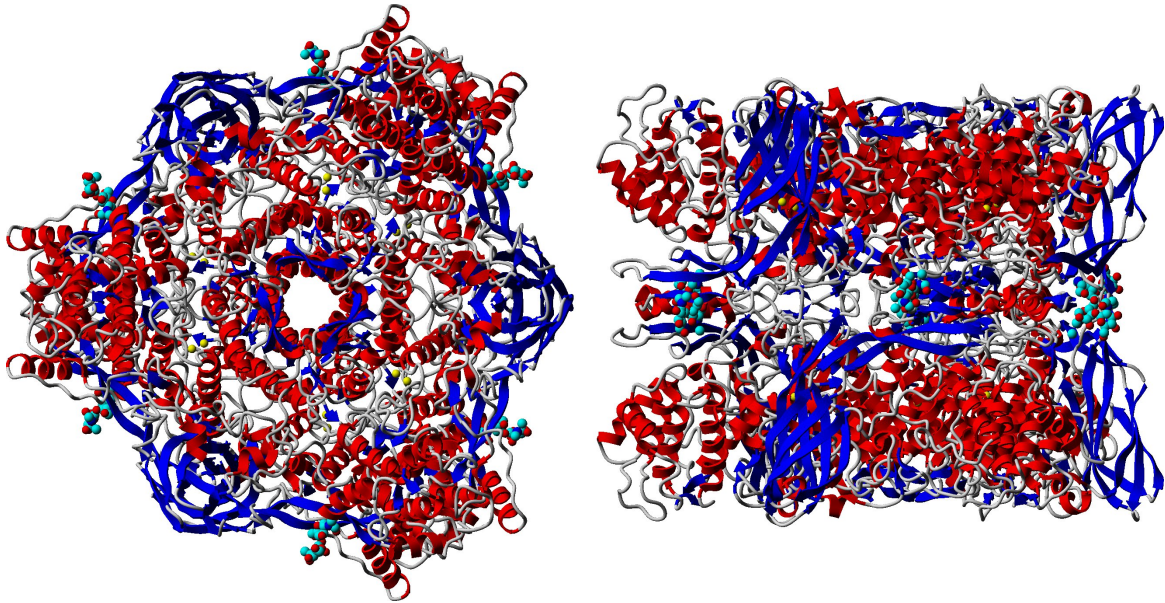


Abbildung 4.9: Modell für das Hämocyanin von *P. elephas*. Links: Sichtachse entlang der dreizähligen Symmetrieachse. Rechts: Nach Drehung um  $90^\circ$  um die Achse in Papierebene. Die Kupferatome sind Gelb eingefärbt. Die Zuckerreste sind raumfüllend in CPK-Darstellung gerendert.  $\alpha$ -Helices sind rot,  $\beta$ -Faltblätter blau eingefärbt.

#### 4.1.2.2 SAXS am Hämocyanin von *Palinurus elephas* und *Palinurus argus* unter Einfluss verschiedener Effektoren

In der Diplomarbeit von A. Rabenhorst wurde erstmalig der Einfluss von Urat und – als Ura-  
tanalogon – Koffein auf die hexameren Hämocyanine von *Palinurus elephas* und *Palinurus argus*  
untersucht (Rabenhorst, 2007). Es konnte mittels isothermer Titrationskalorimetrie (ITC) ge-  
zeigt werden, dass die Proteine wahrscheinlich je zwei Bindungsstellen für die Effektoren besitzen.  
Überdies lagen Werte für die Affinität zu den Effektoren vor und eine strukturelle Charakteri-  
sierung konnte angegangen werden: Die Messungen erfolgten in Gegenwart von 4,5 mM Koffein  
bzw. 2 mM Urat. Bei diesen Konzentrationen ist unter Berücksichtigung der Dissoziationskon-  
stanten sichergestellt, dass die Bindungsstellen zu 99,9 % besetzt sind (Rabenhorst, 2007). Aus  
den SAXS-Messungen, zur Überprüfung des Einflusses der Effektoren auf die Struktur in Lösung,  
ergaben sich die in der folgenden Tabelle dargestellten Parameter:

Tabelle 4.5: Physikalische Parameter für die Hämocyanine von *P. elephas* und *P. argus* unter dem Einfluss von Urat und Koffein<sup>†</sup>.

Protein	Effektor	$R_g$ [Å]	$D_{max}$ [Å]
<i>P. elephas</i> Hämocyanin	ohne <sup>‡</sup>	$48,4 \pm 0,3$	$136 \pm 2$
	Urat	$49,0 \pm 0,4$	$144 \pm 2$
	Koffein	$49,2 \pm 0,5$	$144 \pm 2$
<i>P. argus</i> Hämocyanin	ohne <sup>‡</sup>	$48,1 \pm 0,4$	$134 \pm 2$
	Urat	$49,0 \pm 0,3$	$138 \pm 2$
	Koffein	$49,2 \pm 0,3$	$136 \pm 2$

<sup>†</sup> im Gegensatz zur Darstellung in Rabenhorst, 2007, wird von einem größeren Fehler für den  $D_{max}$ -Wert ausgegangen.

<sup>‡</sup> Die Hämocyanine waren zum Zeitpunkt der Messung oxygeniert.  $O_2$  war als Effektor gegenwärtig.

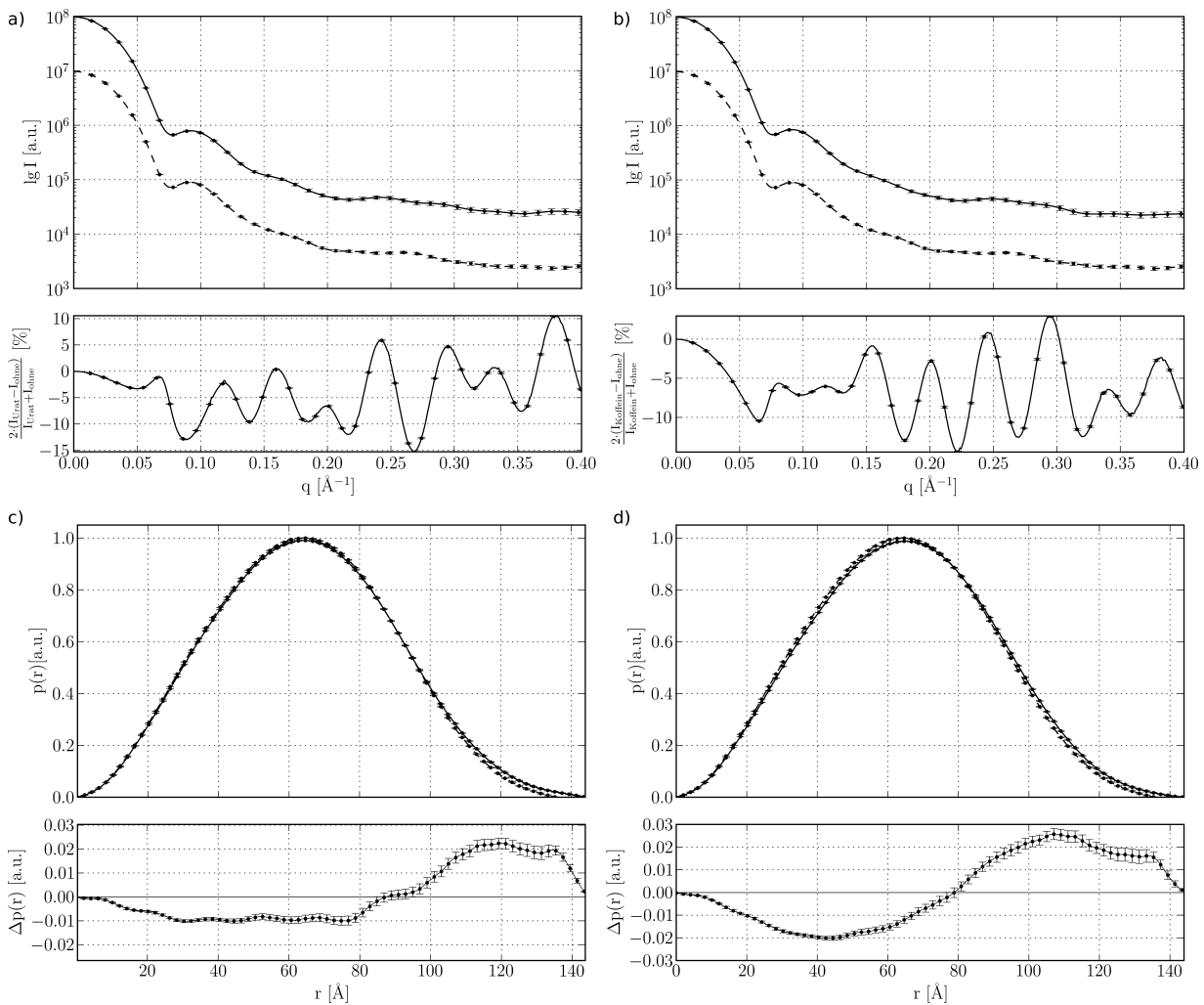


Abbildung 4.10: Streudaten des Hämocyanins von *P. elephas* unter dem Einfluss von Urat und Koffein (mit Effektor: —; ohne Effektor ---). a) Entschmierte Streudaten mit und ohne Urat. b) Entschmierte Streudaten mit und ohne Koffein. Die Kurven in a) und b) sind zur besseren Übersicht um eine logarithmische Einheit gegeneinander versetzt. c) Abstandsverteilungsfunktionen der Daten aus a). d) Abstandsverteilungsfunktionen der Daten aus b).

Unter den Daten befindet sich jeweils ein Differenzplot; für die Streukurven ist dies die prozentuale Differenz nach Skalierung auf  $I(0)$ . Fehlerbalken sind exemplarisch angegeben.

Die Abbildungen c) und d) sind bereits in der Diplomarbeit von A. Rabenhorst gezeigt (Rabenhorst, 2007).

Die Abbildungen 4.10 (oben) und 4.11 (folgende Seite) zeigen die SAXS-Daten für die Hämocyanine von *P. elephas*, respektive von *P. argus*. Bei der Darstellung der Abstandsverteilungsfunktionen wurde davon ausgegangen, dass sich die Masse der Proteine durch die Effektorbindung nicht, bzw. in zu vernachlässigender Weise, verändert<sup>3</sup>.

Offenbar gleichen sich die durch die beiden Effektoren induzierten Veränderungen sowohl in ihrer Amplitude, als auch hinsichtlich der Art: Der Kurvenverlauf ist jeweils sehr ähnlich. Wie die Differenzdarstellungen der Streudaten zeigen, betrifft die Effektorbindung auch sämtliche Abstandsebenen in den Molekülen. Auch ist aus dem Differenzplot zur  $P(r)$  unmittelbar zu schlussfolgern, dass vor allem Bewegungen größerer Domänen oder ganzer Untereinheiten den Konformationswechsel unter Urat- bzw. Koffeinfluss dominieren: Bei kleinen Abständen werden nur geringe Veränderungen beobachtet, dennoch summieren sich diese bei den größeren Abständen auf. Folglich gibt es auf Sekundärstrukturebene nur geringe Veränderungen. Viel wahrscheinlicher scheint die Bewegung ganzer Untereinheiten. Die Herkunft des Hämocyanins – *P. elephas* oder *P. argus* – spielt für diese Aussagen keine Rolle, sie gelten bezogen auf beide Hämocyanine bei beiden

<sup>3</sup>Zwei Urat- bzw. Koffeinmoleküle auf über 450 kDa Masse entsprechen etwas weniger als einem Gewichtsprozent bezogen auf das jeweilige Hämocyanin.

Effektoren.

Diese Feststellung ist keineswegs eine Selbstverständlichkeit, denn die ITC-Daten zeigen leicht unterschiedliche Stöchiometrien für die Bindung der Effektoren bei den beiden Arten (Rabenhorst, 2007) und auch in der Vergangenheit wurde bei 12meren Hämocyanin bereits konstatiert, dass die Stöchiometrie der Bindung von Urat und Koffein artabhängig sein muss (Menze *et al.*, 2000; Hellmann *et al.*, 2004).

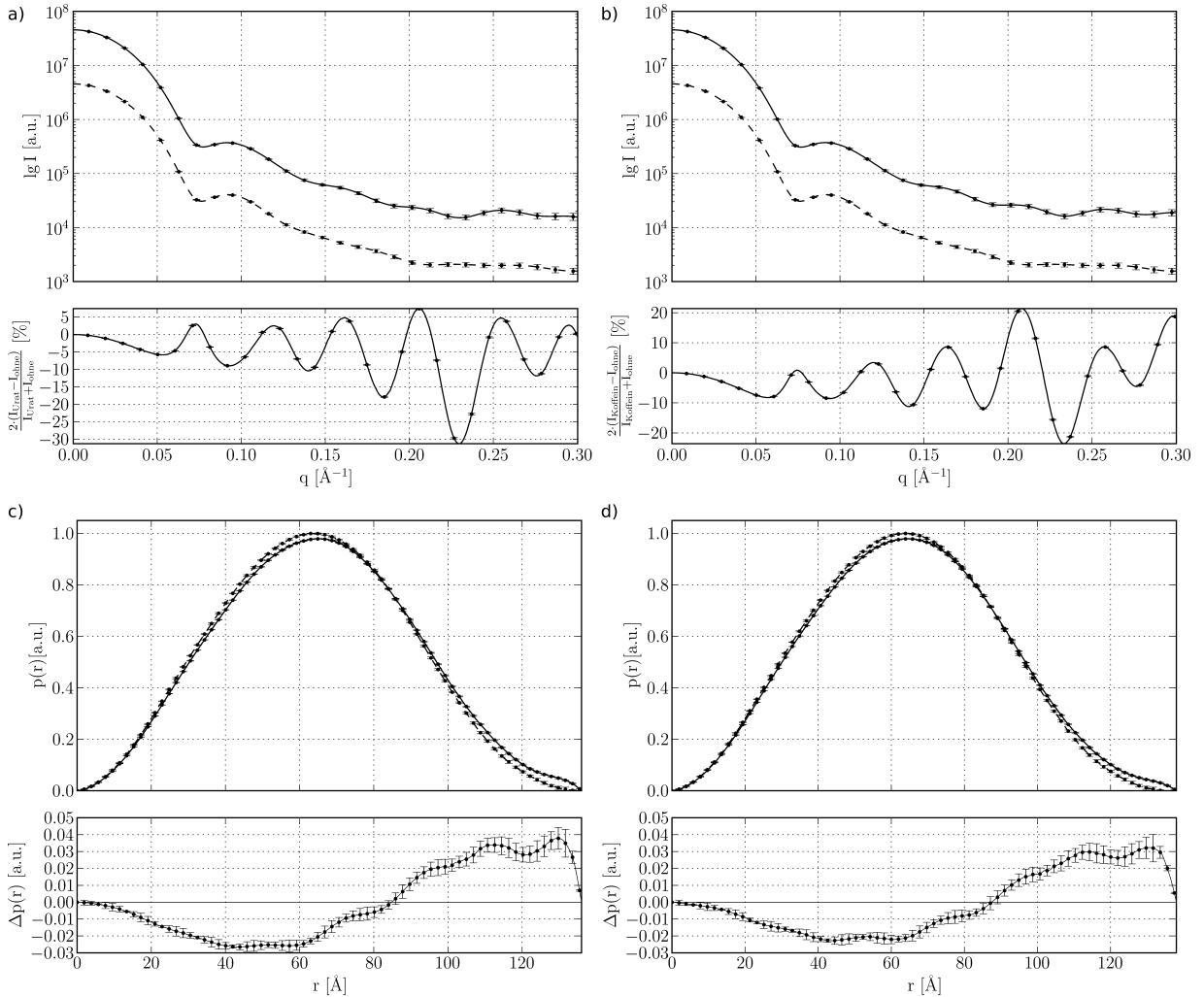


Abbildung 4.11: Streudaten des Hämocyanine von *P. argus* unter dem Einfluss von Urat und Koffein (mit Effektor: —; ohne Effektor ---). a) Entschmierte Streudaten mit und ohne Urat. Die Kurven sind zur besseren Übersicht um eine logarithmische Einheit gegeneinander versetzt. b) Entschmierte Streudaten mit und ohne Koffein. Die Kurven sind ebenfalls um eine logarithmische Einheit gegeneinander versetzt. c) Abstandsverteilungsfunktionen der Daten aus a). d) Abstandsverteilungsfunktionen der Daten aus b). Unter den Daten befindet sich jeweils ein Differenzplot; für die Streukurven ist dies die prozentuale Differenz nach Skalierung auf  $I(0)$ . Fehlerbalken sind exemplarisch angegeben.

Die Abbildungen c) und d) sind bereits in der Diplomarbeit von A. Rabenhorst gezeigt (Rabenhorst, 2007).

### 4.1.2.3 *rigid body*-Modellierung auf Basis der SAXS-Daten

Für das Hämocyanin von *Palinurus elephas* liegt ein Homologiemodell vor. Damit war es naheliegend für die im vorigen Abschnitt vorgestellten SAXS-Daten diejenigen Konformationen zu finden, die am besten den Daten entsprechen. Hierzu wurde eine eigene Software unter der Bezeichnung **sas-rigid** programmiert (siehe 4.7 auf Seite 105 dieses Kapitels). Zunächst wird das Ergebnis einer *Monte-Carlo*-Simulation mit *simulated annealing* vorgestellt (siehe Abschnitt 4.7.2, Seite 110), auf die ein systematisches Suchen der Bewegungsparameter folgte. Dies bezweckt einem optimalen Fit an die experimentellen Daten (siehe Abschnitt 4.7.3, Seite 111). Der *Monte-Carlo*-Algorithmus diente dazu eine ungefähre Konformation zu finden, die den Daten entspricht. Hieraus ließen sich Bewegungsparameter eingrenzen, die in einer systematischen Suche dann schneller abgearbeitet werden konnten. Das Prinzip sieht in jedem Fall vor, dass eine Bewegung auf die Untereinheiten des Hämocyanins appliziert wurde, worauf ein Vergleich der Abstandsverteilungsfunktion der so generierten Konformation mit den experimentellen Daten erfolgte. Um die Berechnung zu beschleunigen, wurden die Atome der Aminosäureseitenketten entfernt. An ihre Stelle trat ein „Atom“ im geometrischen Schwerpunkt der gelöschten Atome. Auf diese Weise schrumpft das Protein nicht wesentlich, so dass gewährleistet blieb, dass Bewegungen nicht überhöht erscheinen (siehe auch Hartmann & Decker, 2002, für einen ähnlichen Ansatz).

In allen folgenden Simulationen werden die Bewegungsparameter bezogen auf die Untereinheit A des Proteins angegeben. Alle weiteren Untereinheiten sind über Symmetrieoperationen mit der Untereinheit A verbunden (siehe Abschnitt 4.7.1.5 auf Seite 109 in diesem Kapitel zur Anwendung von Symmetrieoperation in **sas-rigid**).

Die folgende Abbildung zeigt das Ergebnis der Modellierung für die Konformationen ohne Effektoreinfluss und unter Urateinfluss im Vergleich. Zu beachten ist, dass auch für die Struktur ohne Effektoreinfluss eine Modellierung durchgeführt wurde, da die Musterstrukturen für die Homologiemodellierung überwiegend ohne gebundenen Sauerstoff bestimmt und die eigenen Daten unter Sauerstoffeinfluss gewonnen wurden.

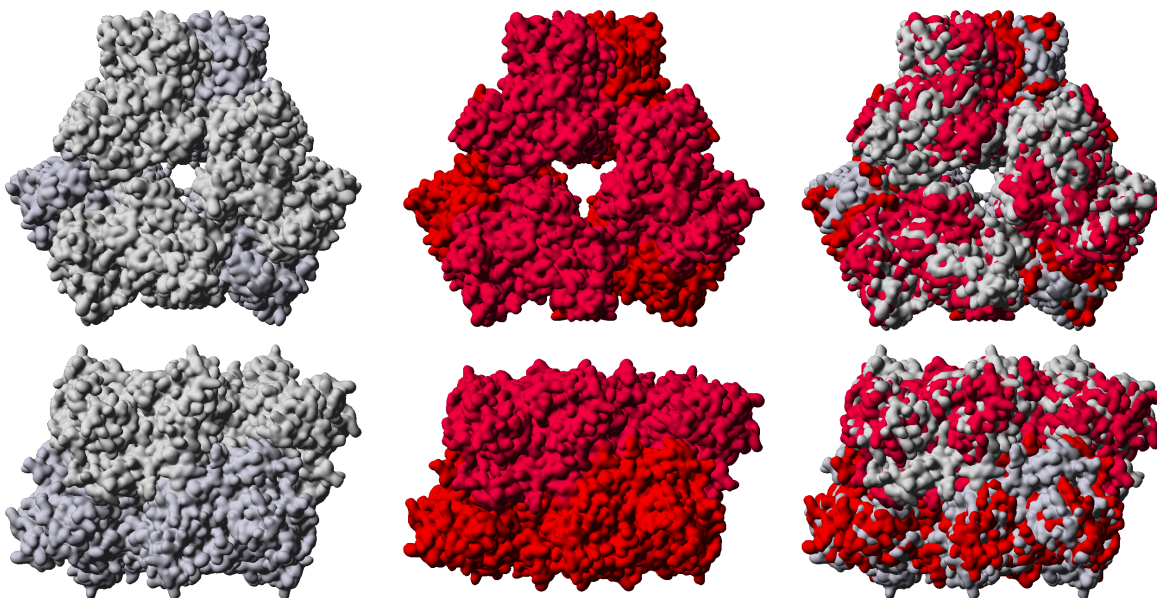


Abbildung 4.12: Molekulare Oberflächen der atomaren Strukturen nach einer systematischen Parametersuche (siehe Abschnitt 4.7.3) für die Konformationen ohne und mit Urateinfluss. Grau: Ohne Effektor. Rot: Unter Urateinfluss. Die rechte Spalte zeigt beide Konformationen übereinandergelegt. Oben: Sicht entlang der Rotationsachse. Unten: „Seitenansicht“, 90° gekippt zur oberen Ansicht. - Die beiden Trimere (jeweils „oben“ und „unten“) sind zur besseren Unterscheidung farblich voneinander abgesetzt.

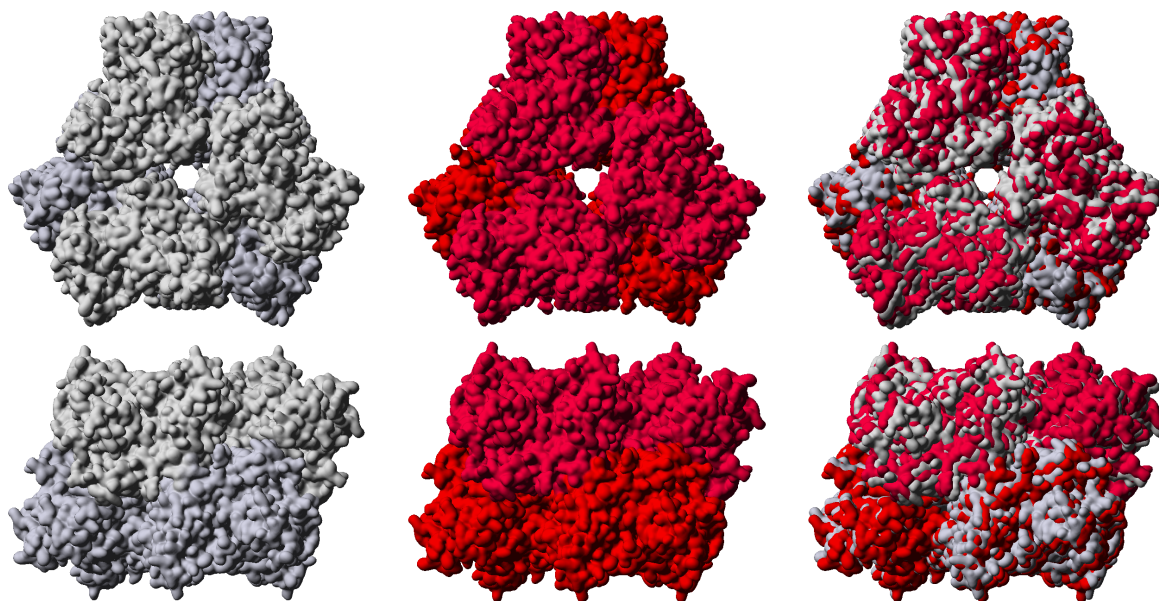


Abbildung 4.13: Molekulare Oberflächen der atomaren Strukturen nach einer systematischen Parametersuche (siehe Abschnitt 4.7.3) für die Konformationen ohne und mit Koffeineinfluss. Grau: Ohne Effektor (dieselbe Struktur wie in Abbildung 4.12). Rot: Unter Koffeineinfluss. Die rechte Spalte zeigt beide Konformationen übereinandergelegt. Oben: Sicht entlang der Rotationsachse. Unten: „Seitenansicht“, 90° gekippt zur oberen Ansicht. - Die beiden Trimere (jeweils „oben“ und „unten“) sind zur besseren Unterscheidung farblich voneinander abgesetzt.

Tabelle 4.6 zeigt die primären Bewegungsparameter für die Untereinheit A des Hämocyanins. Das Homologiemodell wurde vor dem Start der *rigid body*-Modellierung im Koordinatensystem zentriert, so dass der geometrische Schwerpunkt des Moleküls im Ursprung des Koordinatensystems lag.

Tabelle 4.6: Primäre Bewegungsparameter nach systematischer Suche. Aufgeführt sind die Translationsparameter  $dx$ ,  $dy$  und  $dz$  und die Rotationsparameter  $d\theta$ ,  $d\varphi$  und  $d\psi$  in Bezug auf die Startstruktur (das Homologiemodell des Hämocyanins). Translations- und Rotationsparameter sind aus Abschnitt 4.7.1.5 auf Seite 109 zu entnehmen.

verglichene Strukturen	$dx$ [Å]	$dy$ [Å]	$dz$ [Å]	$d\theta$ [°]	$d\varphi$ [°]	$d\psi$ [°]
ohne Effektor	$0,0 \pm 2,0$	$-2,0 \pm 2,0$	$0,0 \pm 2,2$	$-9,0 \pm 2,3$	$-9,0 \pm 2,6$	$-2,0 \pm 4,1$
unter Urat einfluss	$0,0 \pm 2,0$	$-0,4 \pm 2,3$	$0,0 \pm 2,0$	$-8,7 \pm 2,3$	$-7,2 \pm 2,8$	$-3,6 \pm 3,6$
unter Koffeineinfluss	$0,0 \pm 2,0$	$0,0 \pm 2,0$	$0,0 \pm 2,0$	$-9,0 \pm 2,0$	$-6,3 \pm 3,0$	$0,6 \pm 3,7$

Mit Tabelle 4.6 ist bereits die Ähnlichkeit der Strukturen unter Urat- und Koffeineinfluss ersichtlich. Aber auch die Abbildungen 4.12 und 4.13 lassen dies erwarten. Zugleich wird deutlich, dass das Modell ohne Einfluss von Effektoren nach erfolgter *rigid body* Modellierung signifikant vom ursprünglichen Homologiemodell abweicht. Dies stellt keineswegs die Qualität des Algorithmus infrage: Das Homologiemodell ist, aufgrund der Musterstrukturen, das Modell eines deoxygenierten Proteins. Überdies werden die Vorlagestrukturen als qualitativ nicht hochwertig erachtet<sup>4</sup> und für andere Hämocyanine wurde bereits die Existenz mehrerer Konformationen unter Sauerstoffeinfluss aufgrund von ITC-Daten angenommen (Menze *et al.*, 2000; Hellmann *et al.*, 2004).

Tabelle 4.7 ist der Versuch die relativ abstrakten Bewegungsparameter aus Tabelle 4.6 in Parameter umzusetzen, die bezogen auf das spezielle Molekül einen verständlicheren Zugang bieten. Tabelle 4.8 bietet die Gyrationradien der Modelle im Vergleich mit den experimentellen Daten. Sie zeigt, dass die Gyrationradien der Modelle systematisch unter den experimentell ermittelten

<sup>4</sup>Dies wurde wiederholt von verschiedenen Personen berichtet: Professor H. Decker, Dr. H. Hartmann, Dr. K. Magnus, Professor J. Markl. Aber auch eine Analyse mit WhatIf (Vriend, 1990; Hooft *et al.*, 1996) bestätigt, dass die Vorlagestrukturen keine gute Qualität aufweisen.

Gyrationsradien liegen, allerdings nur wenig. Zudem ist der Trend derselbe: Unter Urat- oder Koffeineinfluss vergrößert sich der Gyrationsradius auch bei den Modellen.

Tabelle 4.7: *Geometrische Vergleichsdaten zum systematischen Parameterscan für das Hämocyanin von Palinurus elephas unter Urat- und Koffeineinfluss.* Alle Werte sind auf eine Nachkommastelle gerundet. Die Fehlerschätzung aus Tabelle 4.6 konnte nicht auf diese Darstellung der Parameter übertragen werden. Dies in **sas-rigid** zu implementieren ist aber vorgesehen (siehe Abschnitt 5.10.3 auf Seite 143 der Diskussion).

vergichene Strukturen	Bewegung der Monomere von der Rotationsachse <sup>†</sup> [Å]
ohne Effektor <i>versus</i> Urateinfluss	1, 1
ohne Effektor <i>versus</i> Koffeineinfluss	0, 5
Urateinfluss <i>versus</i> Koffeineinfluss	0, 6
ohne Effektor <i>versus</i> Homologiemodell <sup>††</sup>	2, 5
vergichene Strukturen	mittlere Distanz aller Atome [Å]
ohne Effektor <i>versus</i> Urateinfluss	3, 0
ohne Effektor <i>versus</i> Koffeineinfluss	2, 0
Urateinfluss <i>versus</i> Koffeineinfluss	1, 6
ohne Effektor <i>versus</i> Homologiemodell	14, 9
vergichene Strukturen	von Flächen begrenzter Winkel <sup>‡</sup> [°]
ohne Effektor <i>versus</i> Urateinfluss	14, 9
ohne Effektor <i>versus</i> Koffeineinfluss	4, 9
Urateinfluss <i>versus</i> Koffeineinfluss	4, 0
ohne Effektor <i>versus</i> Homologiemodell	18, 8
vergichene Strukturen	Winkel zwischen Regressionsgeraden <sup>*</sup> [°]
ohne Effektor <i>versus</i> Urateinfluss	7, 6
ohne Effektor <i>versus</i> Koffeineinfluss	1, 5
Urateinfluss <i>versus</i> Koffeineinfluss	6, 1
ohne Effektor <i>versus</i> Homologiemodell	12, 8

<sup>†</sup> Dies entspricht einer Weitung eines Trimers im Hexamer, also einer Dislozierung von der Rotationsachse.

<sup>††</sup> Verglichen wird hier das Homologiemodell aus Abschnitt 4.1.2.1 mit dem Modell nach Durchlauf der *rigid body*-Modellierung.

<sup>‡</sup> Hierzu wurden durch jede Untereinheit „A“ eine Fläche durch den geometrischen Schwerpunkt gelegt, so dass die Untereinheit möglichst so halbiert wird, dass auf beiden Seiten der Flächen die gleiche Anzahl Atome sind. Die Flächen sind im Einzelnen beschrieben durch einen Punkt im Koordinatensystem und einen Normalenvektor (mit Länge 1) senkrecht zur Fläche: Für das Homologiemodell:  $x = -31,0 \text{ Å}$ ,  $y = 13,9 \text{ Å}$ ,  $z = 23,1 \text{ Å}$  mit einem Normalenvektor bei  $x = -0,4$ ,  $y = 0,5$ ,  $z = -0,8$ ; Für das Modell ohne Effektoreinfluss der *rigid body*-Modellierung:  $x = -36,6 \text{ Å}$ ,  $y = 0,8 \text{ Å}$ ,  $z = 18,4 \text{ Å}$  mit einem Normalenvektor bei  $x = -0,4$ ,  $y = 0,5$ ,  $z = -0,8$ ; für das Modell unter Urateinfluss:  $x = -37,6 \text{ Å}$ ,  $y = -1,9 \text{ Å}$ ,  $z = 17,9 \text{ Å}$  mit einem Normalenvektor bei  $x = -0,3$ ,  $y = 0,3$ ,  $z = -0,9$ ; für das Modell unter Koffeineinfluss:  $x = -37,1 \text{ Å}$ ,  $y = -1,0 \text{ Å}$ ,  $z = 19,2 \text{ Å}$  mit einem Normalenvektor bei  $x = -0,5$ ,  $y = 0,5$ ,  $z = -0,7$ . Die Berechnung erfolgte in YASARA (Krieger *et al.*, 2002) (vgl. Abbildung 4.14).

<sup>\*</sup> Für die Regressionsgerade werden alle  $C_{\alpha}$ -Atome jeder Untereinheit „A“ als Punkte aufgefasst (vgl. Abbildung 4.14).

Tabelle 4.8: *Gyrationsradien der Modelle im Vergleich mit den experimentellen Daten.*

vergichene Strukturen	$R_{g,\text{experimentell}}$ [Å]	$R_{g,\text{des Modells}}$ [Å]
Homologiemodell vor Modellierung	–	48, 3
ohne Effektor	$48, 4 \pm 0, 3$	48, 1
unter Urateinfluss	$49, 0 \pm 0, 4$	48, 8
unter Koffeineinfluss	$49, 2 \pm 0, 5$	48, 7

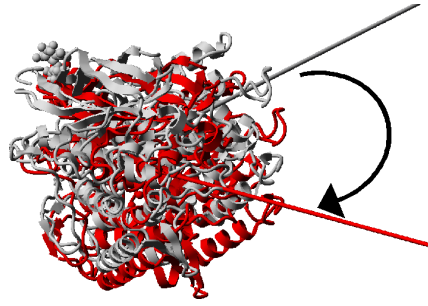


Abbildung 4.14: Winkel zwischen zwei identischen Untereinheiten des Hämocyanins von *P. elephas* in verschiedenen Konformationen. Die Abbildung dient zur Veranschaulichung der Sachverhalte in Abbildung 4.7: Der Winkel zwischen zwei Untereinheiten ist formal als der Winkel zwischen zwei Geraden zu beschreiben. Entweder werden die Geraden als Regressionsgeraden durch eine Punktwolke im dreidimensionalen Raum (z. B. alle  $C\alpha$ -Atome einer Untereinheit) beschrieben oder die Punktwolke wird durch eine Fläche geteilt und der Normalenvektor der Fläche wird so verlängert, bis sich die so entstandenen Geraden schneiden. In beiden Fällen kann es dazu kommen, dass sich die Geraden nicht schneiden. Dies wird durch Translationsbewegungen ausgeglichen, denn diese beeinflussen den eingeschlossenen Winkel nicht.

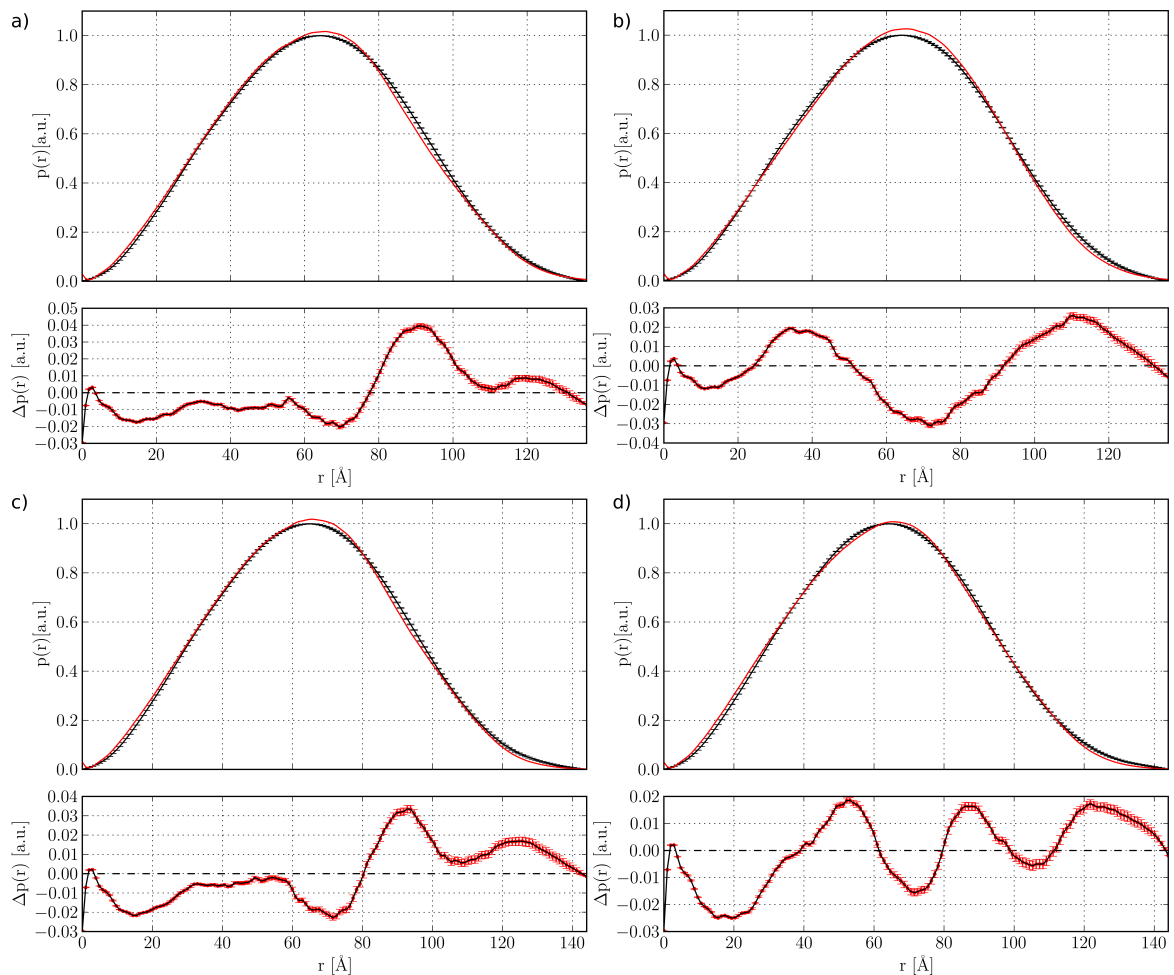


Abbildung 4.15:  $P(r)$  für alle Modelle, welche mit dem Parameterscan (siehe Abschnitt 4.7.3) generiert wurden. Experimentelle Daten sind schwarz, die Daten der Modelle rot. a) Die theoretische  $P(r)$  des Homologiemodells vor der Modellierung bezogen auf die experimentellen Daten ohne Effektorzugabe. b) Die theoretische  $P(r)$  des Modells ohne Effektor mit den experimentellen Daten. c) und d) Die theoretischen  $P(r)$  mit den jeweiligen experimentellen Daten für Urat und Koffein, respektive.

Im Folgenden werden nun Abbildungen gezeigt, die auf einer modifizierten Annahme für die *rigid body*-Modellierung beruhen: Während für die Abbildungen 4.12 und 4.13 nach *einer* Konformation gesucht wurde, die den Daten am besten entspricht, so gilt für die folgenden Abbildungen die

Annahme, dass ein Molekül in Lösungen verschiedene Konformationen einnehmen kann. Die gemessenen Daten entsprechen somit der Streuung eines Konformationenensembles. Die Darstellung der übereinandergelagerten Konformationen erfolgt in einer Dichtedarstellung. Den Berechnungen liegt der in Abschnitt 4.7.5 auf Seite 115 dieses Kapitels vorgestellte Algorithmus zugrunde. Es sei bereits vorweggenommen, dass die Dichten durch eine Übereinanderlagerung verschiedener Konformationen generiert werden, daher weisen die Dichten unterschiedliche Volumina auf.

Die Dichten setzen sich aus verschiedenen Konformationen zusammen, deren theoretische Abstandsverteilungsfunktionen zusammen, gewichtet durch einen Koeffizienten nach Gleichung 4.8, der experimentellen Abstandsverteilungsfunktion bestmöglich entsprechen. Damit nicht abwegige Konformationen in die finalen Dichten eingehen, wurden Konformationen, deren Gyrationradius oder  $D_{max}$  um mehr als 2 Å von den Referenzdaten aus Abschnitt 4.1.2.2 abwichen, nicht für die weiteren Berechnungen herangezogen.

Insgesamt wurden jeweils 15000 Konformationen berechnet und damit etwa 1% der möglichen Zustände<sup>5</sup>. Dass die Zahl der tatsächlich ausgewählten Koeffizienten im Vergleich dazu sehr klein ist, belegt dass der Algorithmus robust ist: Es wird nicht versucht den Fit an die experimentellen Daten durch Addition vieler theoretischer Daten – und damit eine Vielzahl von Koeffizienten – zu optimieren (vgl. Abschnitt 4.7.5, Seite 115).

Tabelle 4.9: Die von den Dichten der Abbildungen 4.16 und 4.17 eingenommenen Volumina. Zusätzlich ist die Zahl der Koeffizienten gezeigt, die die Zahl der Konformationen wiedergibt, die in die Dichtendarstellung nach Gleichung 4.8 eingingen.

Struktur	Volumen <sup>†</sup> [Å <sup>3</sup> ]	Zahl der Koeffizienten
Homologiemodell <sup>‡</sup>	536540	—
nach Modellierung ohne Effektoreinfluss	541536	2
nach Modellierung unter Urateinfluss	616848	5
nach Modellierung unter Koffeineinfluss	581264	4

<sup>†</sup> Berechnet mit YASARA (Krieger *et al.*, 2002).

<sup>‡</sup> Das Modell aus Abschnitt 4.1.2.1.

Im Grunde bestätigen die Abbildungen 4.16 und 4.17 die aus den Daten aus den Abbildungen 4.12 und 4.13: Auch für die Dichten ergibt sich dieselbe apparente Verdrehung der Untereinheiten gegeneinander. Eine detailliertere Analyse befindet sich im Abschnitt 5.2.2 auf Seite 126.

<sup>5</sup>Die Zahl der möglichen Permutationen für die Bewegungsparameter beträgt etwas mehr als eine Millionen.

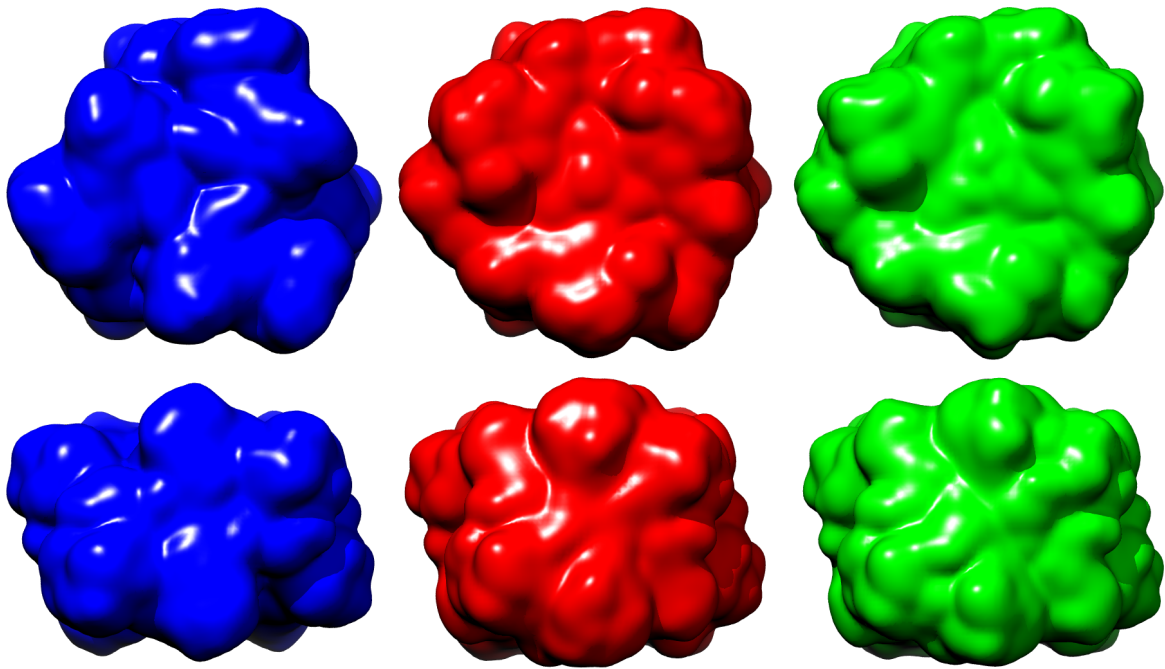


Abbildung 4.16: *Dichtendarstellung einer Monte-Carlo-Simulation (siehe Abschnitt 4.7.5)*. Oben: Sicht entlang der Rotationsachse. Unten: „Seitenansicht“, 90° gekippt zur oberen Ansicht. Blau: Ohne Effektor. Rot: Unter Urateinfluss. Grün: Unter Koffeineinfluss. Die Dichten hatten eine Auflösung von 2 Å, daraufhin wurde zur besseren Übersichtlichkeit jede der Dichten mit einem Gaußfilter mit einem Kernel von 4 Voxeln konvolviert (mittels Chimera; Pettersen *et al.*, 2004). Die Dichte ohne Effektoreinfluss wurde so skaliert, dass sich die nierenförmigen Untereinheiten gut abzeichneten. Den anderen Dichten wurde dieselbe Skalierung zugewiesen. Sie weisen aber ein anderes Volumen auf, da hier eine unterschiedliche Anzahl von Strukturen einging (vgl. Tabelle 4.9 auf der vorherigen Seite).

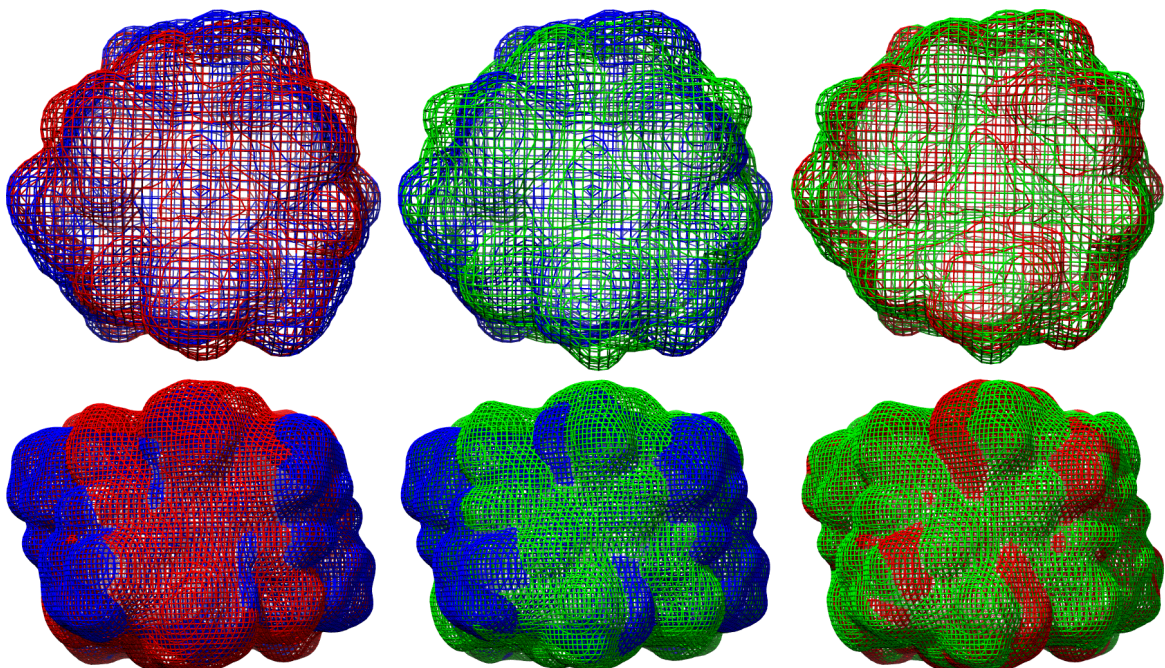


Abbildung 4.17: *Dichtendarstellung aus Abbildung 4.16 übereinandergelagert*. Oben: Sicht entlang der Rotationsachse. Unten: „Seitenansicht“, 90° gekippt zur oberen Ansicht. Blau: Ohne Effektor. Rot: Unter Urateinfluss. Grün: Unter Koffeineinfluss. Dargestellt sind die Dichten als Netze, um eine Tiefenwirkung zu erzielen.

### 4.1.3 Die Struktur des Hämocyanins von *Upogebia pusilla*

Mit dem Fernziel den Konformationswechsel des *U. pusilla*-Hämocyanins bei der Oxy-/Deoxygenierung nachvollziehen zu können, wurde das Projekt an diesem Hämocyanin begonnen. Hierzu wurden mittels SAXS Daten am oxygenierten Protein aufgenommen. Daten des deoxygenierten Proteins konnten später, mangels Material, nicht mehr hinzugefügt werden. Zusätzlich wurden elektronenmikroskopische Referenzdaten aufgenommen. Diese dienten in erster Linie zur Kontrolle der Integrität der Proben.

#### 4.1.3.1 SAXS-Daten

Zur Aufnahme der Kleinwinkelstreudaten wurden folgende Proteinkonzentrationen verwendet: 5,2 mg/mL, 10,9 mg/mL, 20,0 mg/mL, 41,7 mg/mL und 71,3 mg/mL. Der  $D_{max}$  beträgt  $252,0 \text{ \AA} \pm 2 \text{ \AA}$  und der  $R_g$   $84,4 \text{ \AA} \pm 0,1 \text{ \AA}$ .

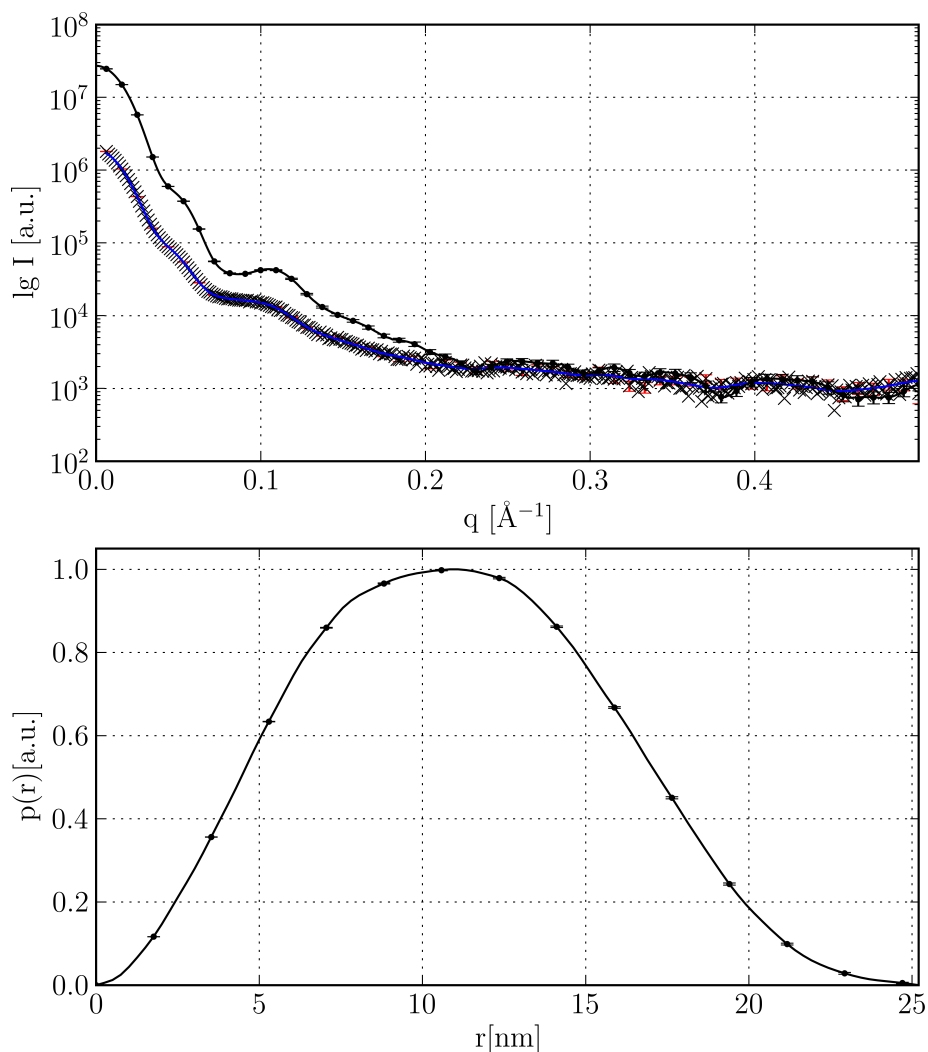


Abbildung 4.18: Streudaten des Hämocyanins von *U. pusilla*. a) Die primären Daten ( $\times$  mit roten Fehlerbalken), der apparente Fit durch diese Daten (blaue Linie) und die entschmierten Daten (schwarze Linie); jeweils mit exemplarischen Fehlerbalken. b) Die Abstandsverteilungsfunktion mit exemplarischen Fehlerbalken.

Abbildung 4.19 auf der nächsten Seite zeigt eine *ab initio*-Rekonstruktion, die mit dem Programm DAMMIN ohne die Annahme einer Symmetrie (P1) erstellt wurde. Andere Symmetrien (P2 oder P4) lieferten ebensowenig überzeugende Rekonstruktionen wie die Annahme einer oblaten oder prolaten Struktur. Alle Kombinationen wurden in einem kugelförmigen Suchvolumen getestet. Als „überzeugend“ wurden hier Rekonstruktionen angenommen, wenn Substrukturen erkennbar waren, die Hexameren entsprechen könnten.

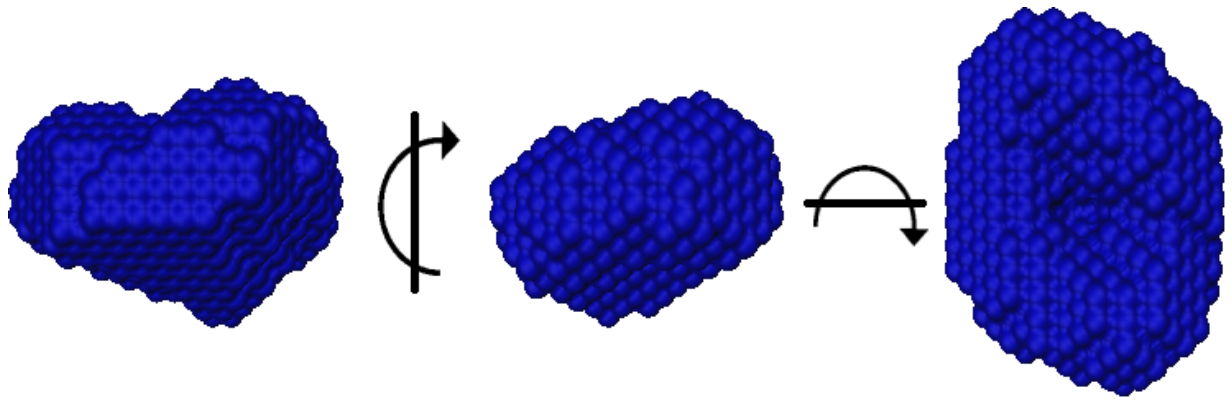


Abbildung 4.19: *DAMMIN*-basierte *ab initio*-Rekonstruktion des *U. pusilla*-Hämocyanins. Von links nach rechts ist jede Darstellung um  $90^\circ$  um die eingezeichnete Achse gedreht. Der Rekonstruktion wurde keine Symmetrie aufgezwungen.

#### 4.1.3.2 Elektronenmikroskopische Daten

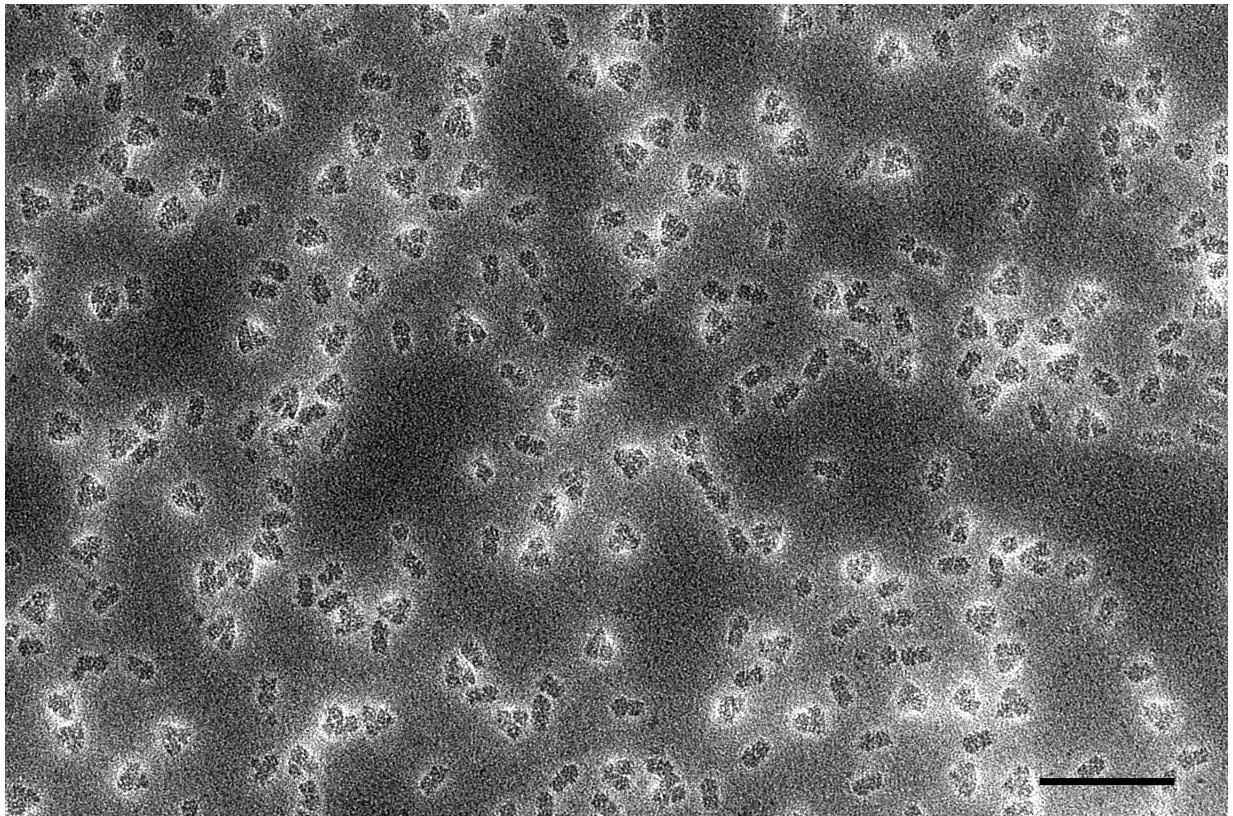


Abbildung 4.20: Übersichtsaufnahme des *U. pusilla*-Hämocyanins im Negativkontrast mit Uranylacetat. Der Maßstabsbalken gibt eine Länge von 200 nm an.

Abbildung 4.20 zeigt eine Übersichtsaufnahme der Probe des *U. pusilla*-Hämocyanins von M. Paoli. Die Aufnahmen wurden so gescannt, dass ein Pixel eine Kantenlänge von  $4 \text{ \AA}$  hat. Insgesamt konnten 1599 Partikel aus den Aufnahmen „gepickt“ werden. Offenbar gab es drei Formen von Partikeln: Solche mit einem, zwei oder drei „Klecksen“. Die erste Kategorie kam kaum vor, wurde als Zerfallsprodukt erachtet und für die weitere Prozessierung nicht verwendet. Repräsentative Partikel der letzten beiden Kategorien wurden ausgewählt. Diese wurden mit einem Tiefpassfilter konvolviert und zwei-, beziehungsweise dreifache Rotationssymmetrie wurde aufgezwungen. Die so entstandenen Referenzartikel (siehe Abbildung 4.21) wurden verwendet, um die übrigen Partikel in zwei Klassen einzuteilen.

Alle Partikel wurden auf die rechtsstehenden Referenzen (Abbildung 4.21) ausgerichtet und nach dem Kreuzkorrelationskoeffizienten zwischen der Referenz und dem jeweiligen Partikel gruppiert: Nur Partikel mit einem Kreuzkorrelationskoeffizienten  $\geq 80\%$  zu einer Referenz wurden in der jeweiligen Klasse belassen. Auf diese Weise wurde zugleich eine große Zahl von Partikeln von der weiteren Betrachtung ausgeschlossen. Dies ist vor allem auf schlechte Defokuswerte und Astigmatismus zurückzuführen und nicht auf die Orientierung der Partikel. So ist sichergestellt, dass alle wesentlichen Orientierungen in den Projektionsmitteln erhalten blieben.

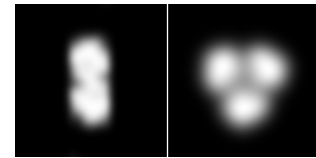


Abbildung 4.21: Zwei Referenzen für das referenz-basierte Sortieren. Die Kantenlänge der Kästchen beträgt 51,2 nm.

318 Partikel konnten keiner der beiden vorgegebenen Gruppen zugeordnet werden. 551 Partikel konnten der „Zwei-Klecks-Gruppe“ und 721 Partikel der „Drei-Klecks-Gruppe“ zugeordnet werden. Abbildung 4.22 zeigt das Ergebnis der Gruppierung dieser erneut ausgerichteten Partikel. Partikel, die in eine der beiden Gruppen sortiert wurden, konnten anschließend einem weiteren – *referenzfrei* – Alignment unterzogen werden.

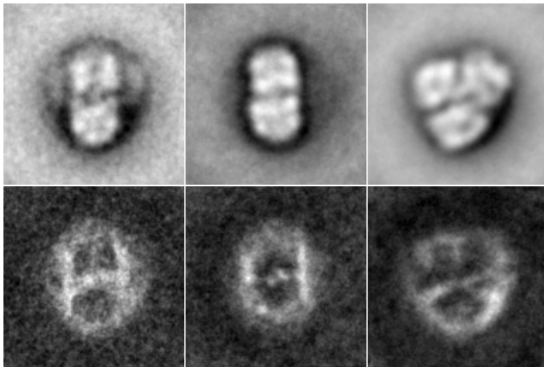


Abbildung 4.22: *Resultat der Klassifizierung nach Alignment.* Obere Reihe von links nach rechts: Mittel aller Partikel, die keine der beiden anderen Klassen zugeordnet werden konnte. Mittel der Partikel, die Referenz 1 zugeordnet wurden. Und das Mittel der Partikel, die Referenz 2 zugeordnet wurden. Die untere Reihe zeigt die korrespondierenden Varianzwerte zu den Klassenmitteln, wobei hohe Grauwerte hohe Varianzen repräsentieren. Die Kantenlänge der Kästchen beträgt 51,2 nm.

Um sicherzustellen, dass keine Konvolutionsartefakte die Abbildungen dominieren, wurde eine CTF-Korrektur durchgeführt. Im Wesentlichen bedeutete dies die Applizierung eines Tiefpassfilters, da keine Thonringe erkennbar waren<sup>6</sup>. Daher werden diese Ergebnisse nicht gezeigt, denn diese sind bis auf eine scheinbare Verschmierung identisch mit den Bildern in den Abbildungen 4.23, 4.24 und 4.25.

Die folgende Abbildung (Nr. 4.23) zeigt das Ergebnis der abschließenden Alignments nach Klassifizierung der ersten Gruppe von Partikeln. Nur die Klassen 2, 5 und 6 sind so gut definiert, dass ein abschließendes Alignment möglich war. Klasse 5 war die best-definierte Klasse, auch erkennbar an der Varianz der Klasse. Das Mittel wird in Abbildung 4.24 auf der nächsten Seite gesondert dargestellt.

<sup>6</sup>Dies ist wahrscheinlich auf die Qualität des Mikroskopes zurückzuführen und war bei allen elektronenmikroskopischen Arbeiten in dieser Dissertation problematisch.

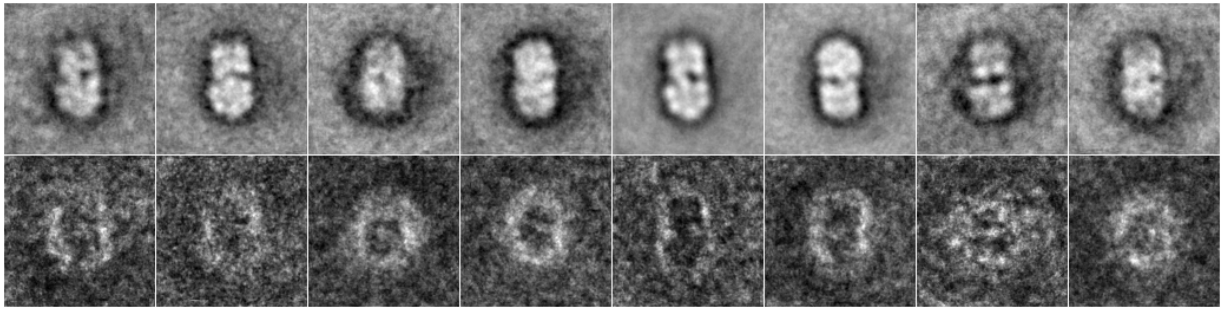


Abbildung 4.23: *Klassenmittel der ersten Gruppe nach erneutem Alignment.* Wie in Abbildung 4.22 zeigt die obere Reihe die Klassenmittel und die untere die korrespondierenden Varianzen. Nur die Klassen 2, 5 und 6 wurden einem abschließenden Alignment unterzogen. Die Kantenlänge der Kästchen beträgt 51,2 nm.

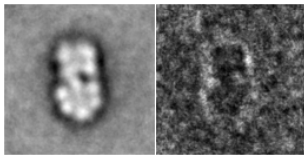


Abbildung 4.24: *Klassenmittel der Klasse 5 aus Abbildung 4.23 mit korrespondierender Varianz.* Die Kantenlänge der Kästchen beträgt 51,2 nm.

Die erste Partikelgruppe scheint aus zwei Hexameren zusammengesetzt zu sein, wie in Abbildung 4.23 deutlich zu sehen ist. Dabei entspricht ein „Klecks“ wahrscheinlich einem Hexamer. Die Rotationsachse des unteren steht annähernd senkrecht zur Abbildungsebene, während die Rotationsachse des oberen um etwa  $90^\circ$  in Paperebene gedreht scheint, so dass die zweizählige Symmetrieachse des Hexamers zu erkennen ist. Die Länge des Partikels beträgt etwa 27 nm und seine Breite etwa 14 nm. Damit ist ein Hexamer apparent größer als vergleichbare Strukturen. Allerdings scheint das verwendete Mikroskop, ebenso wie das andere Tecnaikroskop für die 3D-Rekonstruktionen der AG Markl, nicht gut hinsichtlich der Vergrößerung skaliert zu sein (siehe Seite 79 ff. im folgenden Abschnitt).

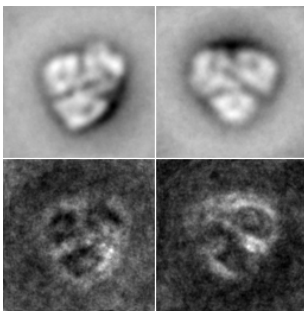


Abbildung 4.25: *Klassenmittel der Gruppe 2-Partikel nach dem abschließendem Alignment mit korrespondierender Varianz.* Die Kantenlänge der Kästchen beträgt 51,2 nm.

Abbildung 4.25 zeigt abschließend die alignierten Klassenmittel der Gruppe 2- oder „Drei-Klecks“-Partikel. Die Partikel haben in der Fläche einen ungefähren Durchmesser von 29 nm. Beide Klassen können als ein Trimer von Hexameren, beziehungsweise ein 18mer, interpretiert werden. Jedoch zeigt keiner der „Hexamerklecks“ die charakteristischen Merkmale eines Hexamers im Negativkontrast. Eine C3-Symmetrie scheint naheliegend, allerdings kann auf Basis dieser Projektion auch keine R3-Symmetrie, wie für das Hämocyanin vorgeschlagen (siehe Abschnitt 1.2.1.7 der Einleitung auf Seite 10), ausgeschlossen werden.

Wenngleich also eine quasi-tetraedrische R3-Symmetrie auf Basis der elektronenmikroskopischen Daten aus Abbildung 4.25 für einen Teil der Partikel nicht ausgeschlossen werden kann, so ist doch durch das zahlenmäßige Verhältnis der Partikelgruppen mit zwei und drei Hexameren deutlich, dass nicht alle Partikel die postulierte R3-Symmetrie (Paoli *et al.*, 2007) aufweisen können. Eine detaillierte Diskussion befindet sich in Abschnitt 5.7 auf Seite 131.

#### 4.1.4 Die Struktur des Hämocyanins von *Porcellio scaber*

Im Rahmen einer Kollaboration mit der Arbeitsgruppe Evolutionsökologie von Asseln an der Universität Kiel wurde eine Probe des Hämocyanins der Kellersassel *P. scaber* gereinigt und charakterisiert. Vor dieser Arbeit gab es keine strukturelle Charakterisierung dieses Hämocyanins. Ziel war also im engeren Sinne nicht den Konformationswechsel des Hämocyanins unter irgendwelchen Bedingungen nachzuvollziehen, sondern eine erste strukturelle Charakterisierung zu erreichen.

Den Asseln wurde Hämolymphe durch dorsale Punktion des Herzens mit einer Glaskapillare entnommen. Die Aufreinigung des Hämocyanins erfolgte über die Gelfiltrationssäule an der MALLS-Anlage (vgl. Abschnitt 3.4.1, Seite 31; Durchführung: Professor E. Jaenicke; zur gesamten Prozedur der Aufreinigung siehe May, 2005). Hierbei konnten zwei Elutionspeaks identifiziert werden. Beide Peaks wiesen Absorption bei 340 nm – typisch für ein Typ-III-Kupferzentrum – auf. Von beiden Peaks wurden elektronenmikroskopische Aufnahmen mit einer CCD-Kamera gemacht (siehe Abschnitt 3.7, Seite 40). In Peak 2 waren ausschließlich 6mere Hämocyanine zu finden, während in Peak 1 ein Gemisch von Hexameren und Dodekameren vorlag. Dies war in guter Übereinstimmung mit den Daten der analytischen Ultrazentrifugation (unveröffentlichte Daten von Professor E. Jaenicke).

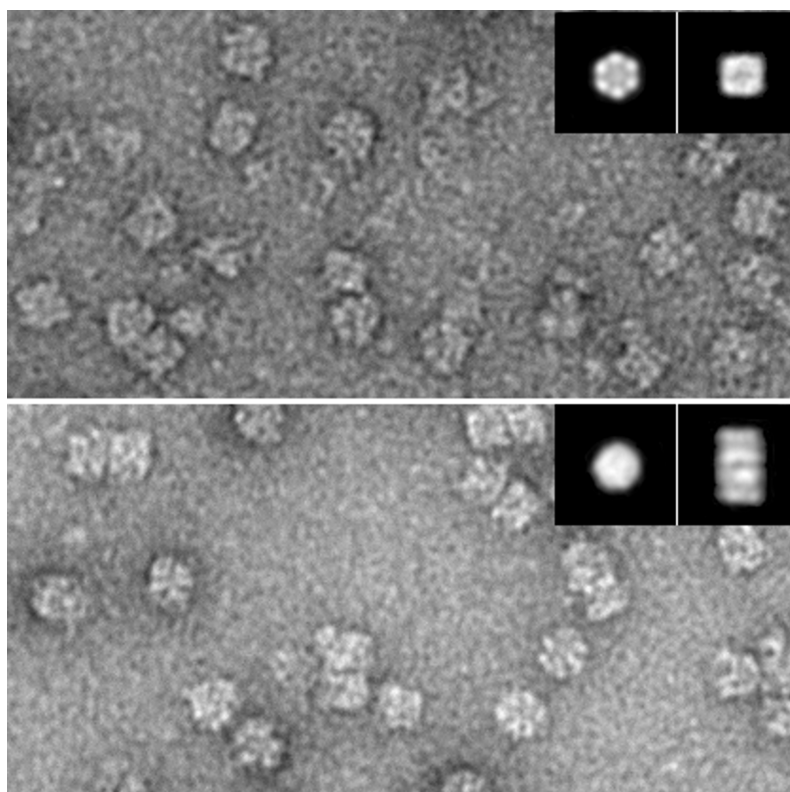


Abbildung 4.26: Elektronenmikroskopische Aufnahmen des *P. scaber*-Hämocyanins im Negativkontrast mit Uranylacetat. Oben: Aufnahme der Peak 2-Probe. Unten: Peak 1-Probe. Die Einschübe zeigen das Mittel für eine Hexamerklasse (oben) und eine Dodekamerklasse (unten) mit dem Blick entlang der Symmetrieachse und um 90° gedreht. Den Mitteln wurde keine Symmetrie auferlegt. Die Kantenlänge der Kästchen beträgt 33 nm. Diese Abbildung ist zur Veröffentlichung eingereicht (Jaenicke *et al.*, 2008).

Abbildung 4.26 zeigt die finalen Mittel in „Auf-“ und „Seitenansicht“ für das Hexamer und das Dodekamer nach Abschluss der Bildbearbeitung (siehe Abschnitt 3.7.2, Seite 40 für eine Erklärung zur Bildbearbeitung). Eine Analyse der Aufnahmen erfolgt in Abschnitt 5.3, Seite 128. Bereits die Mittel lassen erkennen, dass das Hexamer eine D<sub>3</sub>-Symmetrie besitzt, während das Dodekamer durch Translation der Hexamere – ohne Rotation – abgebildet werden kann.

## 4.2 Vergleich elektronenmikroskopische Daten und Kleinwinkelstreudaten von *N. pompilius*-Hämocyanin

Anfang 2006 lag eine *single particle*-Rekonstruktion des Hämocyanins von *N. pompilius* vor (Meissner *et al.*, 2006). In Zusammenarbeit mit B. Pairet wurden SAXS-Daten des Proteins mit dem Ziel aufgenommen, den Konformationsübergang von Oxygenierung und Deoxygenierung nachzuvollziehen (Pairet, 2007). Letztlich aber konnte nur Material für den Teildatensatz des oxygenierten Hämocyanins zur Verfügung gestellt werden. Die existierenden Daten konnten jedoch herangezogen werden, um einen Vergleich zwischen elektronenmikroskopischen und SAXS-Daten durchzuführen.

Professor Meissner stellte die Dichte der Rekonstruktion aus Meissner *et al.*, 2006, zur Verfügung. Die Kantenlänge der Voxel betrug 1,8553 Å. Der Schwellenwert für die Dichte, bei welcher ein Volumen umschlossen wird, welches in etwa dem Molekulargewicht des Proteins entspricht, 0,05 (persönliche Mitteilung Professor Meissner). Die nominelle Masse betrug etwa 3,6 MDa. Wird mit diesen Parametern eine Berechnung der theoretischen Streukurve des Proteins durchgeführt (siehe Abschnitt 3.6.9 auf Seite 37) so ergibt sich eine erhebliche Diskrepanz der Daten (siehe Abbildung 4.27).

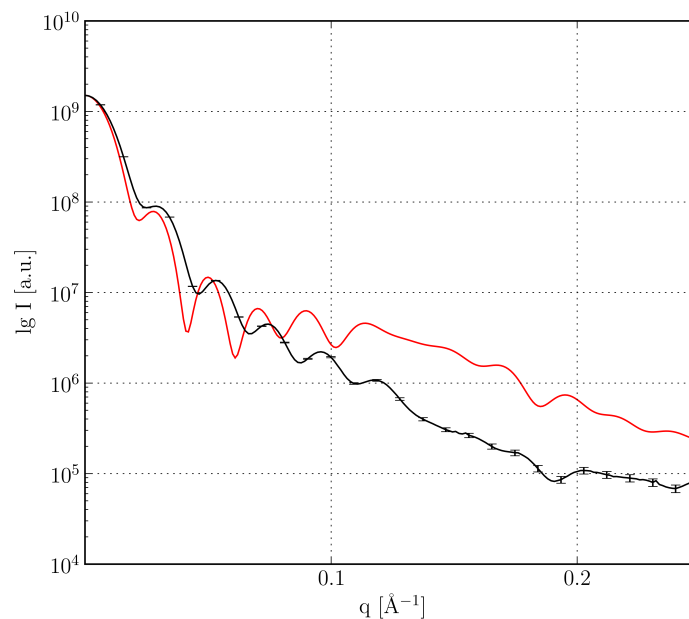


Abbildung 4.27: Entschmierte Streukurve für das *N. pompilius*-Hämocyanin (schwarz) zusammen mit der theoretischen Streukurve der *single particle*-Rekonstruktion (rot).

Es ist offensichtlich, dass die Skalierung hinsichtlich der Werte für  $q$  differiert. Die Abweichung beträgt etwa 12 %. Bei gleichem Schwellenwert für die Dichte wäre die Kantenlänge für die Voxel 1,7069 Å für den bestmöglichen Fit – unter der Annahme, dass die SAXS-Daten korrekt skaliert sind.

Wird sowohl der Schwellenwert, als auch die Größenskalierung systematisch variiert (vgl. Abbildung 4.28), so ergibt sich der beste Fit (d. h. die kleinstmögliche Diskrepanz zwischen experimentellen Werten und theoretischen Werten) bei einem Schwellenwert von 0,07 und einer Voxelkantenlänge von 1,6237 Å. Aufgrund der langen Rechenzeit und der schlechten Fitwerte wurde keine feinere Suche durchgeführt.

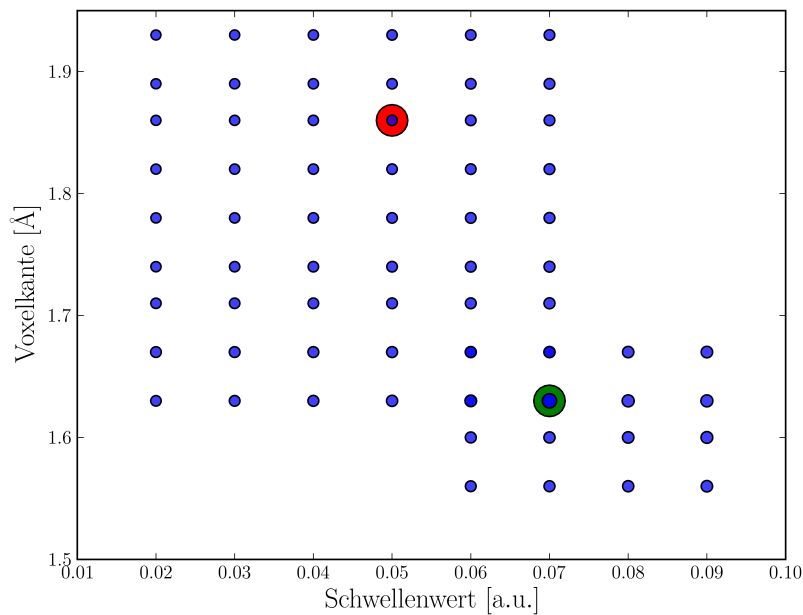
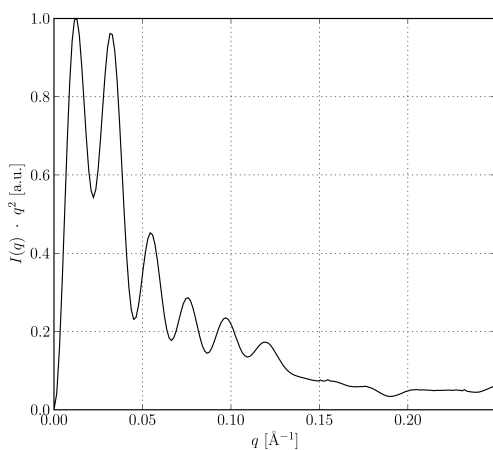


Abbildung 4.28: *Untersuchter Parameterraum der Fitwerte.* Aufgetragen sind die angenommenen Voxelkantenlängen über den Schwellenwert für die Volumenbestimmung (Erläuterung siehe Text). Die Punktgröße ist  $\propto \frac{1}{\chi^2}$ . Die Parameter für den roten Punkt entsprechen den Angaben von Professor U. Meissner. Die Parameter für den grünen Punkt spiegeln den besten Fit wieder. Die Punktgrößen für den roten und grünen Punkt stehen in keinem Zusammenhang mit dem Fitwerten.

Wie bereits angedeutet sind ist die Übereinstimmung zwischen experimenteller und theoretischer Kurve insgesamt schlecht – und zwar unabhängig vom Schwellenwert und Skalierungsfaktor. Abbildung 4.27 bestätigt dies. Die Größenordnung für den Gütefaktor

$$\chi^2 = \frac{1}{n-1} \sum_{i=1}^n \left[ \frac{c \cdot I_{\text{theoretisch}}(q_i) - I_{\text{experimentell}}(q_i)}{\sigma(q_i)} \right]^2, \quad (4.1)$$

wobei  $c$  ein arbiträrer Faktor für die Skalierung auf einen gemeinsamen Wert von  $I(0)$  ist und  $\sigma(q_i)$  für den experimentellen Fehler des  $i$ -ten  $q$ -Wertes steht, liegt im Bereich von  $10^4 - 10^8$  †. Die Erklärung hierfür liegt im unterschiedlichen Verlauf der Kurven: Während beide Kurven Minima und Maxima an ähnlichen Punkten im reziproken Raum aufweisen, ist doch insbesondere bei hohen Werten von  $q$  eine erhebliche Diskrepanz für den Intensitätsabfall festzustellen.



Wie im Kratky-Plot (vgl. Abschnitt 3.6.6, Seite 36; linksstehende Abbildung) zu sehen ist, läuft die Kurve in ein Plateau aus. Dies zeigt, dass der Pufferabzug (vgl. Abschnitt 3.6.4.1, Seite 34) korrekt war und damit die einzige Fehlerquelle für einen fehlerhaften (eigentlich überhöhten) Intensitätsverlauf ausgeschlossen werden kann. Überdies kann geschlußfolgert werden, dass sich das untersuchte Protein in einem nativem Zustand befand (Lipfert & Doniach, 2007).

†Solche Werte gelten als Indikator für eine äußerst schlechte Übereinstimmung. Gute Übereinstimmung würde durch  $\chi^2$ -Werte wesentlich kleiner als  $10^2$  indiziert werden.

Abbildung 4.29: *Kratky-Plot der Nautilusstreudaten.*

### 4.3 Elektronenmikroskopie am Bla-g Hexamerin von *Blattella germanica*

In der Diplomarbeit von A. Beusch (Beusch, 2006) wurde das Hexamerin der Deutschen Schabe (*Blattella germanica*) aufgereinigt und charakterisiert. Meine Aufgabe war die Anfertigung elektronenmikroskopischer Bilder, um eine strukturelle Charakterisierung des Proteins zu erzielen.

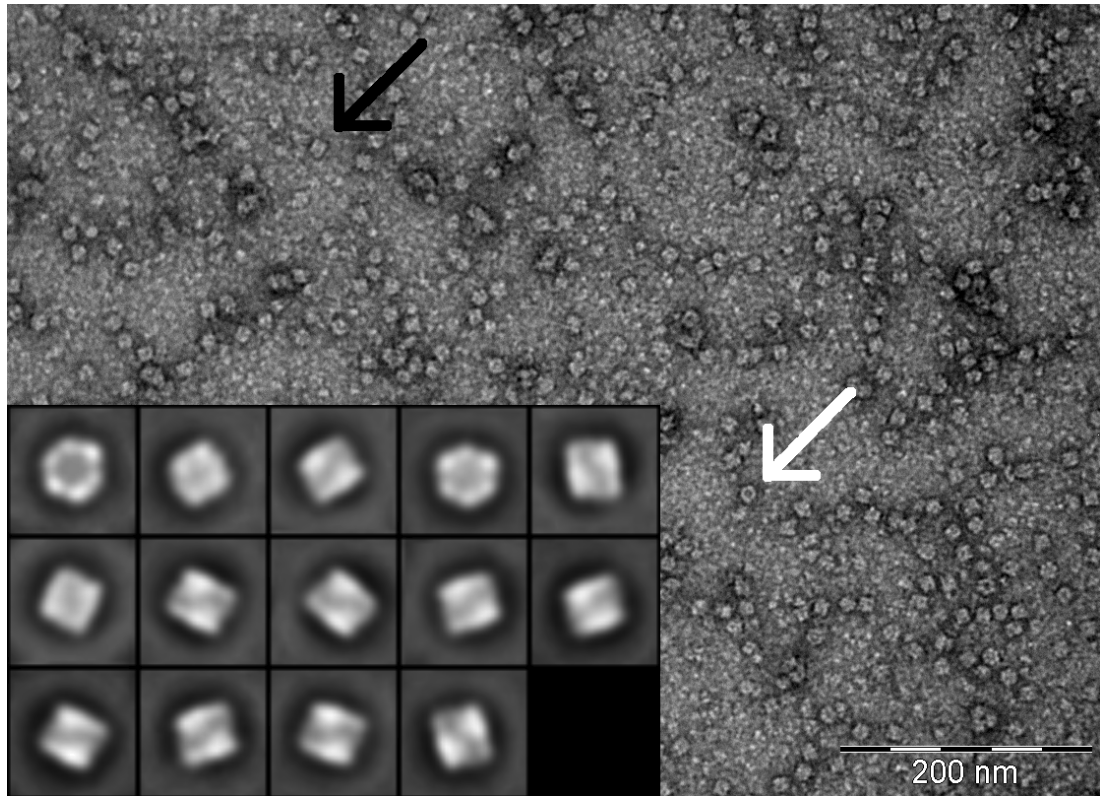


Abbildung 4.30: *Elektronenmikroskopische Aufnahme und Mittelung von Bla-g Hexamerin*. Übersichtsaufnahme einer Aufnahme des Bla-g Hexamerins mit Einschub. Der Einschub zeigt Klassenmittel aus insgesamt 5272 Partikeln. Die Kantenlänge eines Kastens beträgt 24,6 nm. Der weiße Pfeil deutet auf ein Partikel mit Sicht entlang der Rotationsachse (siehe Text) und der schwarze Pfeil auf ein Partikel, das dazu um  $90^\circ$  gedreht scheint. Die Bilder wurden bereits in der Diplomarbeit von A. Beusch gezeigt (Beusch, 2006).

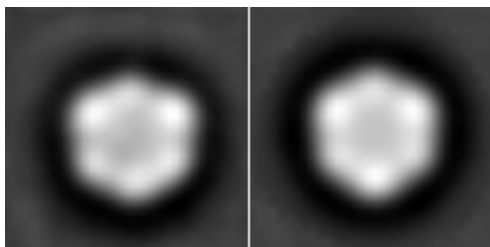


Abbildung 4.31: *Sicht entlang der Symmetrieachse*. Links: Das nicht symmetrisierte Klassenmittel einer Aufsicht (Klassenmittel Nr. 4 von links im Einschub der Abbildung 4.30). Rechts: Dasselbe Klassenmittel mit aufgezwungener dreifacher Rotationssymmetrie. Die Bilder wurden bereits in der Diplomarbeit von A. Beusch gezeigt (Beusch, 2006).

Abbildung 4.30 zeigt eine Übersichtsaufnahme einer Probe des Bla-g Hexamerins von *B. germanica*. Die Details der Aufreinigung sind in der Diplomarbeit von A. Beusch beschrieben (Beusch, 2006). Es ist deutlich zu sehen, dass die Probe homogen verteilt ist. Insgesamt 5272 Partikel konnten „gepickt“ werden (vgl. Abschnitt 3.7.2.1, Seite 40). Es ergaben sich nach einer Klassifizierung insgesamt 14 Klassen; die Einteilung in so viele Klassen ist wahrscheinlich eher ungleichmäßiger Kontrastmittelverteilung als tatsächlicher Unterschiedlichkeit zuzuschreiben. In diesen Klassen gibt es zwei Hauptgruppen: Eine Sicht entlang der vermuteten Rotationsachse des Hexamers („Aufsicht“) und eine Sicht, bei der die Partikel um  $90^\circ$  gedreht sind („Seitenansicht“).

Die Aufsicht ist scheinbar sechsfach symmetrisch, aber eine dreifache Symmetrie (siehe Abbildung 4.31) zeigt, in Hinblick auf ein Rotationspowerspektrum, eine bes-

sere Übereinstimmung mit der besten Aufsichtsklassenmittel (Beusch, 2006). Die bestätigt die Erwartungen hinsichtlich der Symmetrie von Hexamerinen: Das Bla-g Hexamerin weist wahrscheinlich ebenfalls eine D<sub>3</sub>-Symmetrie auf. Der mittlere Durchmesser der Partikel dieser Klasse beträgt  $120 \text{ \AA} \pm 2,5 \text{ \AA}$  und ist damit ebenfalls in guter Übereinstimmung mit der bisher bestimmten Kristallstrukturen hexamerer Hämocyanine (Volbeda & Hol, 1989; Hazes *et al.*, 1993; Magnus *et al.*, 1994). Eine detaillierte Diskussion befindet sich in Abschnitt 5.6.

## 4.4 Konformationen des $\alpha$ -Hämolytins von *Staphylococcus aureus*

Die vorhergehenden Abschnitte befassten sich mit strikt löslichen Proteinen. Im Folgenden wird das  $\alpha$ -Hämolytin von *Staphylococcus aureus* behandelt, ein Toxin, welches von dem Bakterium als Monomer sekretiert wird, auf den Zellmembranen von Wirtsorganismen assemblieren Poren bilden kann. Eine detaillierte Einführung gibt diese Arbeit auf Seite 13 ff. Ziel der Experimente war es das Toxin strukturell zu charakterisieren. Während die Struktur der Pore bekannt ist, waren zu Beginn dieser Arbeit die Strukturen des Monomers und die Strukturen der „Präpore“ unbekannt.

### 4.4.1 Reinigung des Toxins

Bei vielen strukturellen Experimenten ist die Reinheit eines Proteins ausschlaggebend für die Aussagekraft des Experimentes. Im Fall der Kleinwinkelstreuung ist zudem Homodispersität unabdingbare Voraussetzung.

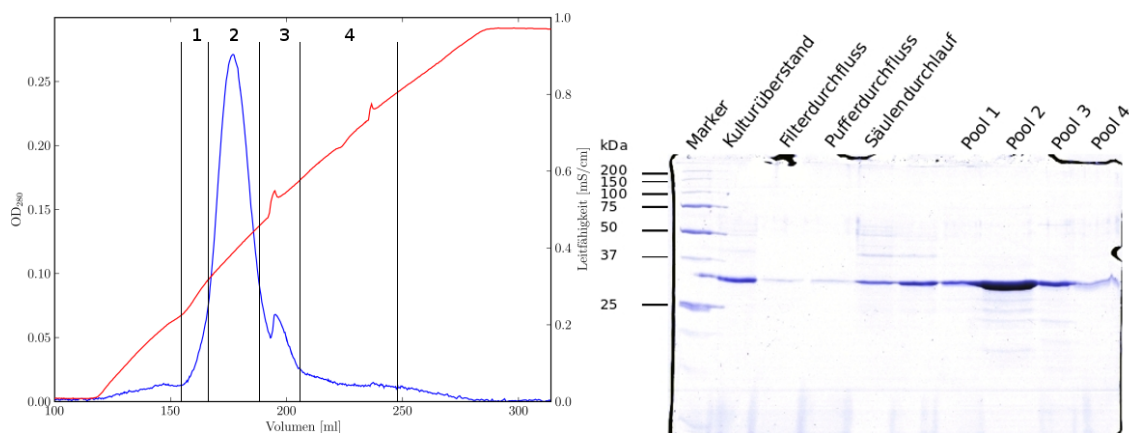


Abbildung 4.32: *Reinigung des  $\alpha$ -Hämolytins wildtyp Proteins*. Links: Elutionsprofil nach Beginn der Elution. Blau: Absorption ( $OD_{280}$ ), Rot: Leitfähigkeit. Die „Pools“ in der rechten Teilabbildung durch die Ziffern und senkrechte Ziffern indiziert. Rechts: Mit Coomassie gefärbtes Gel dieser Reinigung. Von links nach rechts: Eine Bahn des Molekulargewichtstandards, Auszug aus dem Kulturüberstand, Verlust durch das Ultrafiltrationsmodul (direkt nach Applikation des Kulturüberstands = „Filterdurchfluss“ und indirekt nach erfolgter Umpufferung = „Pufferdurchfluss“), Verlust nach Applikation auf die CM-Sepharosesäule (= „Säulendurchlauf“; zweimal aufgetragen) und verschiedene zusammengeführte Fraktionen des gereinigten Proteins (= „Pool 1–4“). Die Beschreibung des Protokolls findet sich im Abschnitt 3.10.1, Seite 46.

Abbildung 4.32 zeigt ein Gel einer Reinigung des  $\alpha$ -Hämolytins wildtyp Proteins (wt) (siehe Abschnitt 3.10.1, Seite 46 zur Beschreibung des Reinigungsprotokolls). Die Reinigung ging mit enormen Verlusten von Protein einher: Alle Proben, die in der Abbildung als „Durchfluss“ bezeichnet werden sind einem Volumen von etwa zwei Litern entnommen. Dennoch ergab eine Reinigung aus etwa acht Litern Kultur 20 – 40 mg (Trockengewicht - extrapoliert von lyophilisierten Proben) des wt-Proteins und etwas weniger als 10 mg für die Mutante H35R. Die Expression der Mutante H35R ist nur etwa ein Viertel der Menge für den Wildtyp bezogen auf das Kulturvolumen<sup>7</sup>.

Das Protokoll für die Reinigung des Wildtyps und der Mutante war identisch. Die Reinheit der Proteine ( $\alpha$ -Hämolytins wildtyp und die Mutante H35R) unterschied sich nicht und betrug mehr als 98%. Überladene Bahnen eines Gels (in Abbildung 4.32 beispielsweise die Bahn „Pool 2“) zeigten stets ein leichtes Bandenmuster unterhalb der Toxinbande.

<sup>7</sup>Persönliche Mitteilung A. Valeva.

## 4.4.2 Die monomere Struktur des Hämolymins

### 4.4.2.1 Kleinwinkelstreudaten

Zur Aufnahme der Kleinwinkelstreudaten wurden folgende Proteinkonzentrationen verwendet:  $2 \times 4,5$  mg/mL, 6,8 mg/mL, 6,9 mg/mL, 9,2 mg/mL und 30 mg/mL. Die Daten liessen sich gut zusammenfassen, wobei jedoch Daten, die bei den kleineren Proteinkonzentrationen aufgenommen wurden, nicht bis zu kleinen Werten von  $q$  reichen. Daher rührt der relativ große Fehler bei kleinen  $q$ -Werten (siehe Abbildung 4.33 (a)). Insgesamt konnten Werte von  $0,0138 \text{ \AA}^{-1}$  bis  $0,4003 \text{ \AA}^{-1}$  berücksichtigt werden.

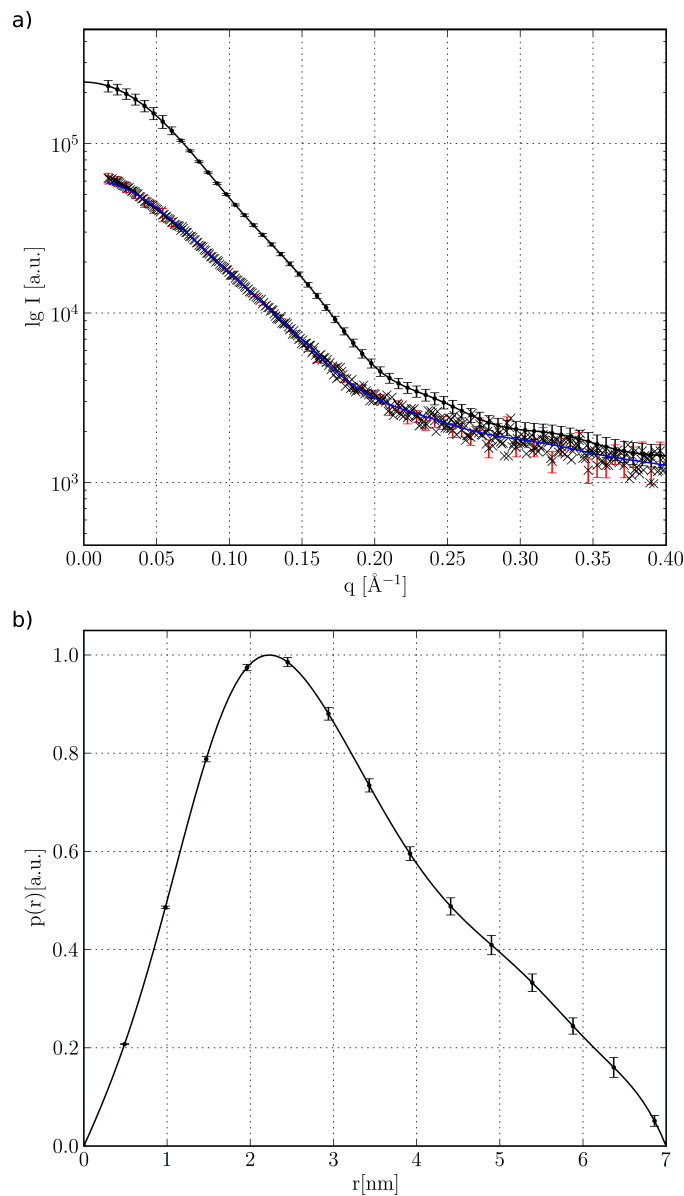


Abbildung 4.33: Kleinwinkelstreudaten des wt- $\alpha$ -Hämolymins. a) Die primären Daten ( $\times$  mit roten Fehlerbalken), der apparente Fit durch diese Daten (blaue Linie) und die entschmierten Daten (schwarze Linie); jeweils mit exemplarischen Fehlerbalken. b) Die Abstandsverteilungsfunktion mit exemplarischen Fehlerbalken.

Die Abstandsverteilungsfunktion des wt- $\alpha$ -Hämolymins entspricht einem typischen prolaten Partikel (Glatter & Kratky, 1982) und bestätigt die Erwartungen (Olson *et al.*, 1999). Der maximale Abstand im Toxin beträgt  $70 \text{ \AA} \pm 2 \text{ \AA}$  und der Gyrationradius ist  $23,8 \text{ \AA} \pm 0,6 \text{ \AA}$ .

#### 4.4.2.2 Modellierung und Verifikation der Monomerstruktur

Beim sogenannten *Profiling* der Proteindatenbank PDB (Berman *et al.*, 2000) durch die Modellierungssoftware MODELLER (Sali & Blundell, 1993) (siehe Abschnitt 3.8.1, Seite 41) wurden folgende Sequenzen, für die eine röntgenkristallographische Struktur vorlag, als ausreichend ähnlich ermittelt, um als Muster für die Modellierung dienen zu können: Das Leukocidin F (bzw. LukF; PDB-Codes: 1LKF, 2LKF und 3LKF; Olson *et al.*, 1999), das Panton-Valentin Leukocidin F (bzw. LukF-PV; PDB-Code: 1PVL; Pédelacq *et al.*, 1999) und schließlich das Panton-Valentin Leukocidin S (bzw. LukS-PV; PDB-Code: 1T5R; Guillet *et al.*, 2004). Insgesamt liegen also drei Proteine mit fünf Strukturen als Muster vor. Strukturen anderer Toxine, wie z. B. des Cholera-toxins (PDB-Code: 1XTC, Zhang *et al.*, 1995), wurden aufgrund von zu großen Unterschieden in der Primärstruktur nicht als Muster für eine Modellierung in Erwägung gezogen. Abbildung 4.34 auf der nächsten Seite zeigt das strukturelle Alignment.

Insgesamt wurden 500 Modelle für das monomere  $\alpha$ -Hämolytin mit MODELLER (siehe Abschnitt 3.8.1, Seite 41) generiert. Das beste Modell für das  $\alpha$ -Hämolytin hat eine DOPE-Wert von -28960,7 und scheint damit schlechter als die Referenzwerte der Musterstrukturen (siehe Tabelle 4.10). Es wurde nicht dasjenige Modell mit dem niedrigsten (= besten) DOPE-Wert ausgewählt, sondern das Modell mit dem niedrigsten Wert, welches zugleich keine unsinnigen Strukturelemente aufwies (z. B. topologisch unmögliche Verläufe des Peptidrückgrats).

Tabelle 4.10: Übersicht über die DOPE-Werte der Modellierung

Protein	DOPE-Wert
modellierte Toxine	-29133,2 – -27273,8
„beste“ Struktur	-28960,7
bestes Modell nach MD-Simulation	-30197,7
LukF (PDB-Code: 1LKF)	-32132,3
LukF (PDB-Code: 2LKF)	-32100,7
LukF (PDB-Code: 3LKF)	-32161,4
LukF-PV (PDB-Code: 1PVL)	-32153,6
LukS-PV (PDB-Code: 1T5R)	-34707,9

Die Ursache für den schlechten DOPE-Wert kann u. a. in der Schleife zwischen den Resten 135-140 liegen (L3 in Abbildung 4.35 a)), da diese in keiner der Musterstrukturen enthalten war. Überdies besteht, aufgrund fehlender Information in den Musterstrukturen, große Unsicherheit über die Position und Form der Schleifen L1 (Reste 12 bis 17) und L2 (Reste 119 bis 127; ebenfalls Abbildung 4.35 a)). Es wurde daher versucht, mit der in MODELLER integrierten „Loopmodellierung“-Funktionalität (Fiser *et al.*, 2000) Klarheit über die Struktur und Lage dieser Schleifen zu gewinnen.

Der Ansatz hierzu war im Wesentlichen „Versuch und Irrtum“. Insbesondere bei Schleife L3 war es naheliegend, dass diese in  $\beta$ -Form an die  $\beta$ -Faltblätter in der unmittelbaren Nachbarschaft anschließt und so wurde versucht dieser Schleife diese Struktur in der Modellierung aufzuzwingen. Zudem wurden verschiedene Einschränkungen der Abstände (siehe Fiser *et al.*, 2000) der  $C_{\alpha}$ -Atome der Schleifen L1-L3 zu ihren Nachbarn ausprobiert. Keiner dieser Versuche führte zu einer signifikanten Verbesserung des DOPE-Wertes. Es ist also naheliegend diese Schleife als tatsächlich strukturell nicht klassifizierbar zu betrachten.

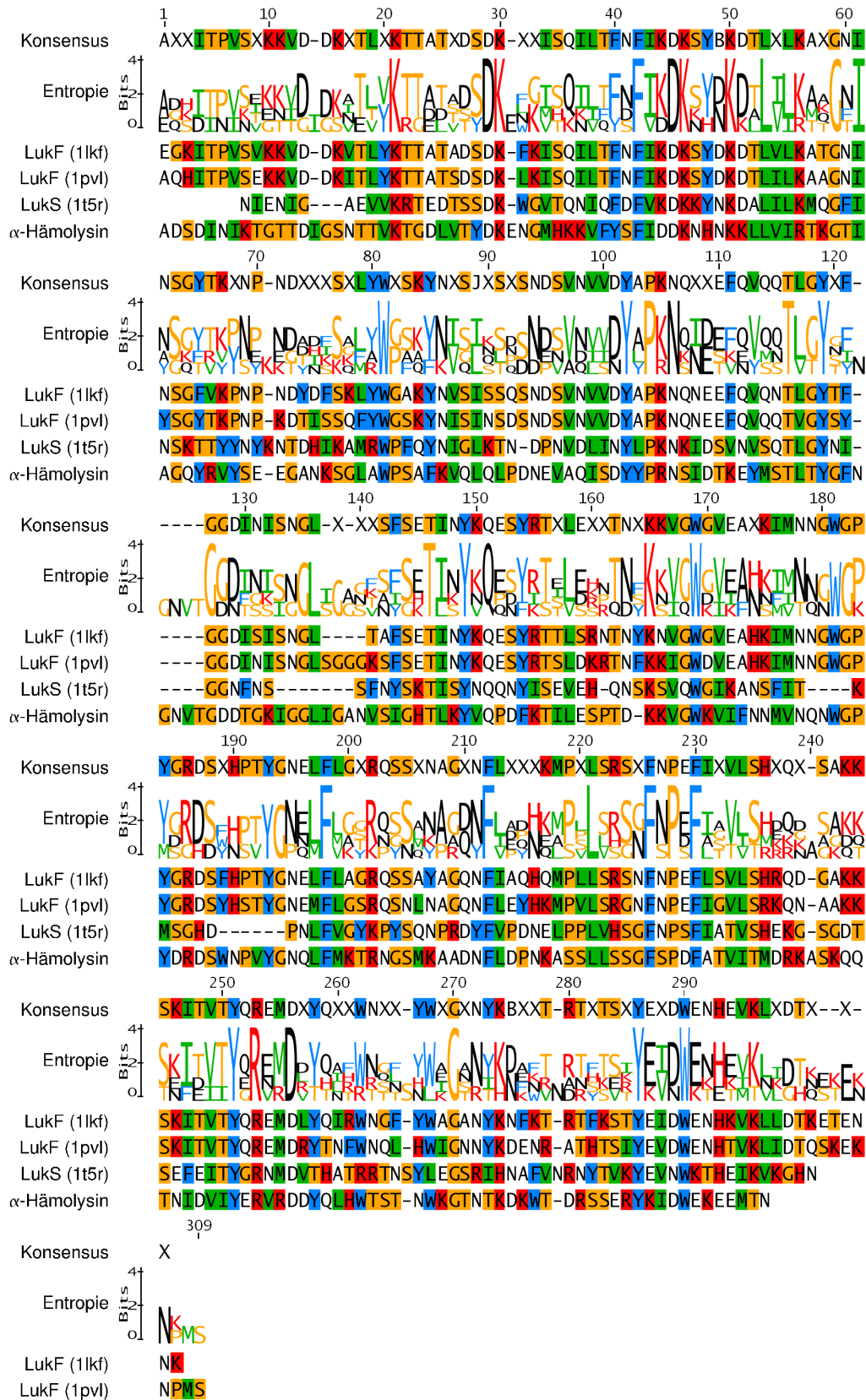


Abbildung 4.34: Das für die molekulare Modellierung verwendete Alignment. Für jede Zeile ist die Konsensussequenz, die Sequenzentropie (nach Schneider & Stephens, 1990) und, in einzelnen Reihen, die Mustersequenzen zusammen mit der Sequenz des  $\alpha$ -Hämolyisins angegeben. Die Farben der Aminosäuren entsprechen dem CLUSTAL-Farbschema (Jeanmougin *et al.*, 1998) und dienen der besseren Unterscheidung.

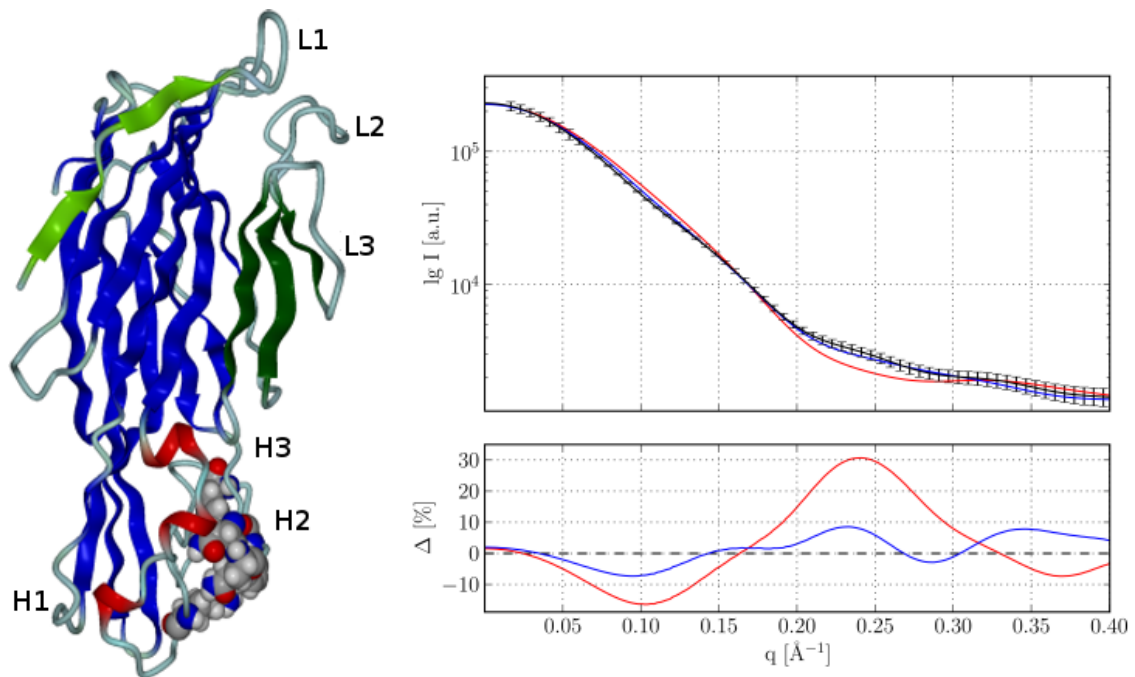


Abbildung 4.35: *Modell des  $\alpha$ -Hämolytins und SAXS-Daten.* a) Das Modell des Toxins mit schematischen  $\beta$ -Faltblättern und  $3_{10}$ -Helices (H1-H3) nach Durchführung der Molekulardynamiksimulation mit YASARA. L1-L3 bezeichnen drei spezielle „loops“ bzw. Schleifen. Mit H1-H3 werden drei  $3_{10}$ -Helices bezeichnet (siehe Text). Die Cholinbindedomäne nach Galdiero & Gouaux, 2004, ist in CPK-Darstellung besonders hervorgehoben. Die hellgrünen  $\beta$ -Faltblattsegmente liegen in der Porenstruktur wie ein Arm zu den Nachbarprotomeren (siehe Abbildung 4.51). Die dunkelgrün eingefärbten  $\beta$ -Faltblätter formen in der Porenkonformation des Toxins die *stem*-Domäne. b) Experimentelle Streudaten (schwarz) mit exemplarischen Fehlerbalken. Dazu die theoretische Streukurve des Toxinmodells (blau) und einer Musterstruktur (von LukF, rot). Im unteren Diagramm ist die Differenz der beiden theoretischen Streukurven in den jeweiligen Farben zur experimentellen Kurve direkt aufgetragen. Die Abbildung ist bereits in Meesters *et al.*, 2008, veröffentlicht.

Zusätzlich zu den bioinformatischen Methoden wurden SAXS-Experimente durchgeführt um das Modell des monomeren  $\alpha$ -Hämolytins in Lösung zu verifizieren. Ein Vergleich der Modellstruktur und den experimentellen Daten geschieht sowohl über die theoretischen Streudaten der Modelle, als auch über *ab initio*-Modelle auf Basis der experimentellen Streudaten.

Der Vergleich der theoretischen Streukurve für das Homologiemodell des  $\alpha$ -Hämolytins, bestimmt mittels CRY SOL (siehe Abschnitt 3.6.8, Seite 37), zeigt eine bemerkenswerte Übereinstimmung zu den experimentellen Daten mit einem  $\chi$ -Wert von 1,1 (siehe Tabelle 4.11). Im Gegensatz dazu ist die Übereinstimmung der theoretischen Streudaten mit den experimentellen Streudaten für die Musterstrukturen wesentlich schlechter. Abbildung 4.35 b) zeigt die experimentellen Daten und die theoretischen Daten für das  $\alpha$ -Hämolytin und beispielhaft für die LukF-Struktur. Die theoretischen Daten für die anderen Leukocidine sind äußerst ähnlich.

Tabelle 4.11: *Übersicht über die  $\chi$ -Werte der theoretischen Streuung verschiedener Modelle zu den experimentellen Streudaten des monomeren  $\alpha$ -Hämolytins.*

Protein	$\chi$ -Wert
$\alpha$ -Hämolytin	1,1
LukF (alle Strukturen)	4,3
LukF-PV	5,1
LukS-PV	4,0
LukF (nach „Loopmodellung“)	3,9
LukF $\Delta$ C-Terminus (nach „Loopmodellung“)	4,1

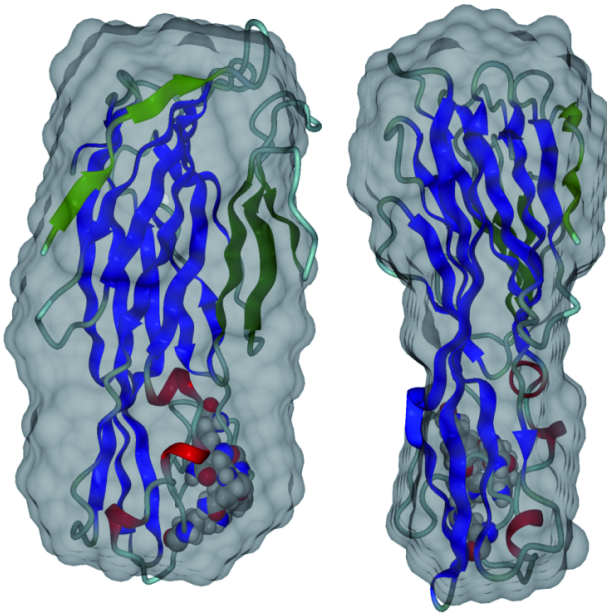


Abbildung 4.36: *Ab initio*-Modell generiert mit DAMMIN. Gezeigt sind zwei Ansichten: Die rechte ist gegen den Uhrzeigersinn um  $90^\circ$  zur longitudinalen Achse der linken Ansicht gedreht. In die DAMMIN-Dichten ist das Modell aus Abbildung 4.35 gelegt. Die Abbildung ist bereits in Meesters *et al.*, 2008, veröffentlicht.

Molekulardynamik drei  $3_{10}$ -Helices. Die ersten beiden, in der Reihenfolge der Aminosäuren benannt (H1 und H2 in Abbildung 4.35 a)), traten nur in ungefähr  $\frac{2}{3}$  der Simulationszeit von 2,5 ns auf. Helix H3 trat nach ihrer Ausbildung nach ungefähr 200 ps konstant auf. Eine nähere Betrachtung der Strukturen von LukF und LukF-PV zeigt, dass diese an den entsprechenden Positionen zwar keine Helices, aber „Ansätze“ von Helices aufweisen: Das Peptidrückgrat folgt in etwa einer idealen  $3_{10}$ -Helix. Nach der Molekulardynamiksimulation verbesserte sich der DOPE-Wert erheblich auf -30197,7 (siehe Tabelle 4.10).

Zusätzliche strukturelle Unterschiede zwischen den Musterstrukturen und dem  $\alpha$ -Hämolyisin entstehen durch „Deletionen“ und „Insertionen“, sowohl auf genetischem als auf strukturellem Niveau: Die Musterstrukturen, mit Ausnahme des LukS-PV, zeigen einen zusätzlichen N-terminal Abschnitt von 9 – 11 Aminosäuren, der größtenteils  $\beta$ -Faltblattcharakter aufweist. Andererseits zeigt das  $\alpha$ -Hämolyisin eine Schleife zwischen den Resten 119 bis 128 (L2 in Abbildung 4.35 a)), welche in den Musterstrukturen nicht gelöst ist.

Die Reste 135–140 (L3 in in Abbildung 4.35 a)) sind sowohl in der Sequenz des  $\alpha$ -Hämolyisins als auch in den Sequenzen der Musterstrukturen enthalten. Aber in den Kristallstrukturen von 1lkf, 3lkf und 1t5r ist dieser Bereich nicht gelöst. Die anderen Musterstrukturen zeigen hier ebenfalls einen sehr ungeordneten Bereich. Ein Versuch diesen Bereich mit „Loopmodellierung“ zu lösen (s. o.) ergab weder hinsichtlich der DOPE-Werte noch hinsichtlich der Übereinstimmung theoretischer und experimenteller Streudaten eine Verbesserung.

In der übrigen Struktur kann eine große Übereinstimmung zwischen den Musterstrukturen und dem  $\alpha$ -Hämolyisin festgestellt werden. Olson *et al.*, 1999, errechneten einen RMSD-Wert für die C $\alpha$ -Atome von 1lkf von 2,1 Å für den  $\beta$ -Sandwichbereich und von 2,4 Å für die *rim*-Domäne.

Eine Schwierigkeit in der Berechnung der theoretischen Daten für die Struktur des LukF entsteht dadurch, dass strukturelle Information für die Positionen 129–135 fehlt, da diese Aminosäuren in der Kristallstruktur nicht gelöst sind. Ein Versuch diese über „Loopmodellierung“ zu generieren zeigte ebensowenig einen besseren  $\chi$ -Wert wie die anschließende *in silico*-Deletion des C-Terminus nach Position 289 (vgl. Tabelle 4.11, die unteren beiden Zeilen). Aminosäuren nach Position 289 des LukF fehlen in der Primärsequenz des  $\alpha$ -Hämolyisins, das insgesamt 293 Aminosäuren lang ist. Ein zusätzlicher Hinweis für die Qualität des  $\alpha$ -Hämolyisin-Modells ist der Vergleich mit dem *ab initio*-Modell, welches auf den SAXS-Daten basiert (linke Abbildung). Dichte, die im *ab initio*-Modell scheinbar „leer“ ist, ist tatsächlich durch die Reste der Aminosäure nahezu ideal gefüllt. Diese Seitenketten werden zur besseren Übersicht nicht gezeigt.

Während keine der Musterstrukturen Helices aufwies, entstanden beim  $\alpha$ -Hämolyisin durch die Modellierung und die anschließende Mole-

### 4.4.2.3 Rückgratflexibilität des $\alpha$ -Hämolytins

Für das  $\alpha$ -Hämolytin wurde eine Molekulardynamiksimulation (siehe Abschnitt 3.8.3, Seite 42) von 2,5 ns Dauer durchgeführt. Nach einer Equilibrierungszeit von 300 ps wurden alle  $C\alpha$ -Atompositionen der „Schnappschüsse“ (vgl. Abschnitt 3.8.3, Seite 42) dazu verwendet die RMSF-Werte zu errechnen.

Im Allgemeinen bleibt das Modell des  $\alpha$ -Hämolytins während der Simulation stabil. Die meisten Aminosäurereste zeigen niedrige RMSF-Werte (vgl. Abbildung 4.37 a)). Einige Bereiche jedoch zeigten erhöhte Flexibilität während der Simulation, insbesondere bezüglich der  $C\alpha$ -Atome der Peripherie (vgl. Abbildung 4.37 b)). Ausnahmen der generellen Stabilität des Modells sind der N-Terminus, die Schleifen L1-L3 und die gesamte Phosphocholinbindestelle. Es ist außerdem auffällig, dass Aminosäuren, welche die Transmembrandomäne der Pore ausmachen, ebenfalls eine insgesamt erhöhte Flexibilität aufweisen.

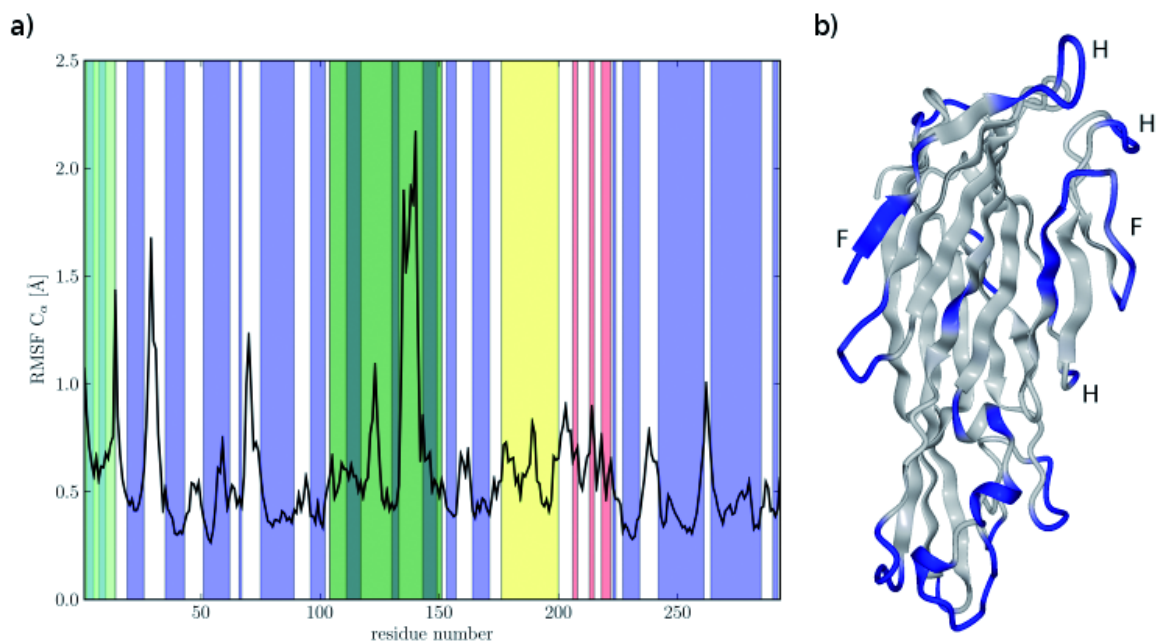


Abbildung 4.37: Flexibilität des  $\alpha$ -Hämolytins. a) RMSF-Werte der Simulation bezogen auf die  $C\alpha$ -Atome der Aminosäurereste. Hintergrundfarben: Blau –  $\beta$ -Faltblattstruktur, Rot –  $3_{10}$ -Helices, Gelb – die Cholinbindedomäne nach Galdiero & Gouaux, 2004, Hellgrün – N-terminale Domäne, Dunkelgrün – Transmembrandomäne in der Porenstruktur. b) Darstellung der RMSF-Werte auf der Modellstruktur. Je kräftiger das Blau, desto höher der RMSF-Wert. Auf die Abkürzungen „F“ & „H“ wird in der Diskussion, Seite 135, eingegangen.

Die Abbildung ist bereits in Meesters *et al.*, 2008, veröffentlicht.

#### 4.4.2.4 Evolutionäre Aspekte

Wenngleich keine Korrelation der Sequenzkonservierung mit den RMSF-Werten besteht, so fällt diese doch an einigen Schlüsselpositionen doch zusammen (vgl. rechtsstehende Abbildung).

Die Darstellung der Sequenzkonservierung (rechts) und des Phylogramms (unten) beruht auf 75 Sequenzen, ermittelt mittels Pipealign (siehe Abschnitt 3.10.8, Seite 52).

Glycin 134 in der  $\alpha$ -Hämolyysinsequenz ist häufig im Gesamtalignment an dieser Position. Glycin 15 ist charakteristisch für LukS-Sequenzen, während die entsprechende Position in LukS-Sequenzen durch Aspartat dargestellt wird. Asparagin 123 ist der dritte Rest eines NGN-Motifs im  $\alpha$ -Hämolyysin, welches YGG für die meisten LukF- und GGN für die meisten LukS-Sequenzen ist. Schlüsselpositionen innerhalb der Phosphocholinbindesstelle sind innerhalb der Sequenzen von LukF-Toxinen und dem  $\alpha$ -Hämolyysin konserviert. Es konnte gezeigt werden, dass einige Seitenketten dieser Domäne in der Abwesenheit von Lipiden ungeordnet erscheinen (Song *et al.*, 1996; Olson *et al.*, 1999), während sie in Anwesenheit von Phosphocholin geordnet sind (Galdiero & Gouaux, 2004).

Das zirkuläre Phylogramm ist auf der Folgeseite dargestellt und generiert wie auf Seite 54 des Methodenkapitels beschrieben. In der Abbildung werden, mit Ausnahme des  $\alpha$ -Hämolyisins, keine Proteinbezeichnungen, sondern PDB-Codes (Berman *et al.*, 2000) oder Zugangscodes der Swiss-Prot-Datenbank ([www.expasy.org/sprot](http://www.expasy.org/sprot)) angegeben. Es ist deutlich zu sehen, dass das  $\alpha$ -Hämolyysin in den ermittelten Sequenzen relativ isoliert steht: Es gibt keine direkt nahestehende Sequenz im Alignment. Wenngleich Sequenzen in den Datenbanken gefunden wurden, die dem  $\alpha$ -Hämolyysin größtenteils entsprachen, so waren dies ausschließlich Vorläufersequenzen (vom  $\alpha$ -Hämolyysin wird eine Signalsequenz abgeschnitten) oder doppelte Einträge. Keiner dieser Einträge kann Beiträge zum Informationsgehalt hinsichtlich der Struktur des  $\alpha$ -Hämolyisins liefern.

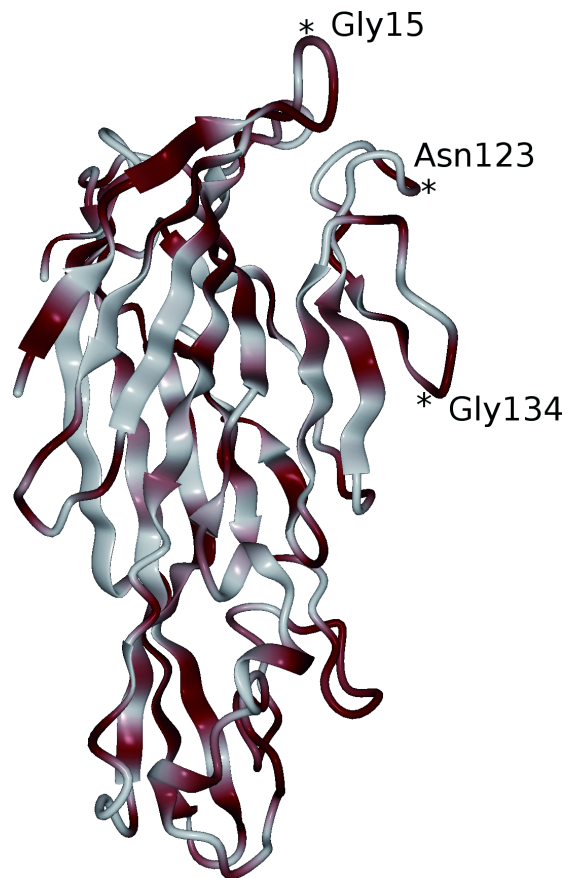


Abbildung 4.38: Sequenzkonservierung im  $\alpha$ -Hämolyysin nach Abschnitt 3.10.8, Seite 52. Je stärker der Rot-Ton, desto stärker ist der Beitrag der  $\alpha$ -Hämolyysinsequenz zur Sequenzinformation des Gesamtalignments der Toxinfamilie.

Die Abbildung ist bereits in Meesters *et al.*, 2008, veröffentlicht.



konnten gar nicht erst bis zu 190 nm oder besser aufgenommen werden<sup>8</sup>.

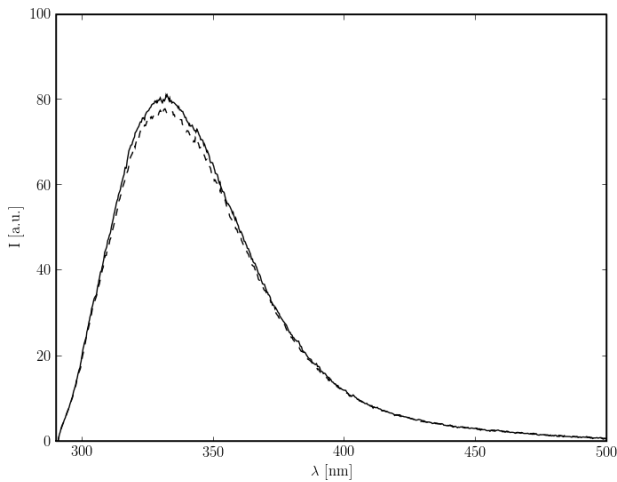


Abbildung 4.41: Fluoreszenzspektren für die beiden  $\alpha$ -Hämolysinvarianten. Pufferkorrigierte Spektren für Wildtyp (---) und H35R (—).

Neben der Charakterisierung der Sekundärstruktur durch CD-Messungen ist die Exposition von Tryptophan- und Tyrosinresten zum wässrigem Medium ein sensibler Indikator für die strukturelle Identität zweier Proteine: Eine Polaritätsverschiebung der Umgebung eines Fluorophors führt zu einer Veränderung der Emissionswellenlängen (siehe auch Galla, 1988).  $\alpha$ -Hämolysin besitzt mit 8 Tryptophan- und 14 Tyrosinresten eine ganze Reihe von Indikatoren für eine eventuelle Fehlfaltung des Proteins.

Abbildung 4.41 zeigt die entsprechenden Fluoreszenzspektren. Die Fluoreszenzmaxima ergeben sich für den Wildtyp zu 331,6 nm und für H35R zu 330,8 nm. Für die Bestimmung der Maxima wurden die Kurven nicht geglättet.

#### 4.4.3.2 Matchpunktbestimmung der Lipide

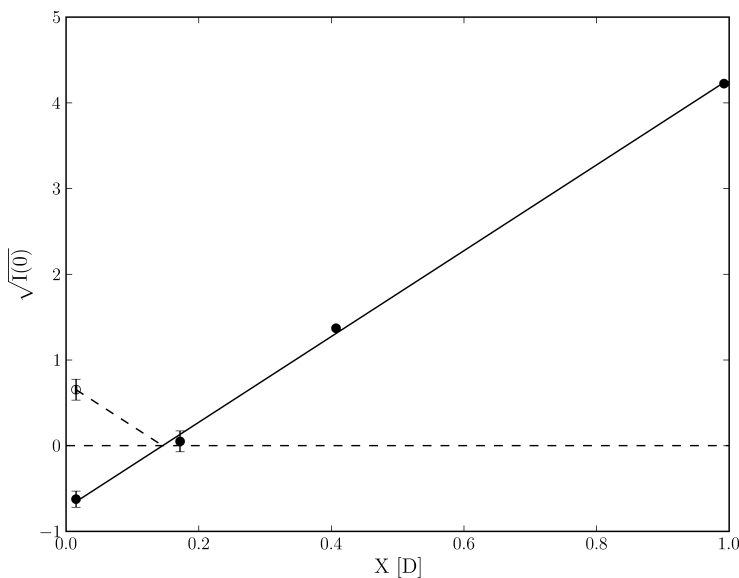


Abbildung 4.42: Matchpunktbestimmung für die Lipidmischung des SANS-Experimentes. Aufgetragen ist der Kontrast ( $\sqrt{I(0)}$ ) über dem relativen Deuteriumanteil. Der gestrichelte Teil spiegelt die tatsächlichen Daten bei kleinem Deuteriumanteil wieder, der als „negativer“ Kontrast gilt.

Zur Bestimmung des Matchpunktes wurde dieselbe Lipidzusammensetzung verwendet, wie beim SANS-Experiment (im nächsten Abschnitt). Ziel war es eine Deuteriumkonzentration im Puffer zu finden, bei der die Lipide keinen Kontrast mehr zum Medium bieten, so dass bei Zugabe des  $\alpha$ -Hämolyins nur noch die Streuung des Proteins aufgenommen würde (vgl. Theoriekapitel „Wirkungsquerschnitt bei Neutronen“ auf Seite 21 ff.).

Die linke Abbildung 4.42 zeigt das Experiment (siehe Methodenkapitel Seite 50): Eine Lipidmischung wurde in verschiedenen Puffern gemessen, deren Deuteriumanteil variiert wurde. Die Intensität bei  $I(0)$  wurde mittels eines Guinierplots ermittelt (siehe Gleichung 2.39 auf Seite 26).

Wie zu erkennen ist, liegt der Matchpunkt für die Lipide bei einem Deuteriumanteil von unter 20 %, genauer bei  $14,272\% \pm 2,367\%$ . Der Wert für 60,3 % Deuterium konnte nicht berücksichtigt werden, da aufgrund eines Lecks vor der Messung zuviel Material verloren ging.

<sup>8</sup>Dies ist auf den Zustand des CD-Spektrometers zurückzuführen: Durch Abnutzung optischer Elemente konnte kein klares Signal bei kürzeren Wellenlängen gewonnen werden.

### 4.4.3.3 Aufnahme und Interpretation von SANS-Daten

Bereits die ersten SANS-Daten zeigten, dass das „Matching“ des Lipids gelungen war. Wie deutlich in Abbildung 4.43 zu erkennen ist, zeigen die Kurven nicht das typische Profil für Vesikel: Eine Streukurve mit vielen Minima<sup>9</sup>.

Für die Auswertung hinsichtlich der Masseverteilung innerhalb der Vesikel (siehe Abschnitt 3.10.7 des Methodenkapitels auf Seite 51) wurde ein molekulares Volumen von 40349,83 Å für das Protomer (bestimmt mit YASARA (Krieger *et al.*, 2002)) angesetzt. Damit ergibt sich für die Massenverteilung innerhalb der Vesikel die folgende Tabelle:

Tabelle 4.12: Gewichtsverteilung in Vesikeln nach den SANS-Daten.

Mutante	Angenommener Körper	$M_S$ [g/cm <sup>2</sup> ]	$M_L$ [g/cm]	Heptamer <sup>†</sup> [g/cm <sup>2</sup> ]
wt	Scheibe	$7,1 \cdot 10^{-9} \pm 0,7 \cdot 10^{-9}$		$3,6 \cdot 10^{-7}$
H35R	Zylinder <sup>†</sup>		$3,6 \cdot 10^{-15} \pm 0,4 \cdot 10^{-15}$	

<sup>†</sup> Obwohl Abbildung 4.45 im nächsten Abschnitt Vesikel in der Probe für die Mutante eindeutig nachweist, konnte kein scheibenförmiger Körper an das entsprechende SANS-Profil gefittet werden. Die Steigung liegt zu nahe bei 1 (siehe Abbildung 4.43).

<sup>‡</sup> theoretisch bei dichtester Packung

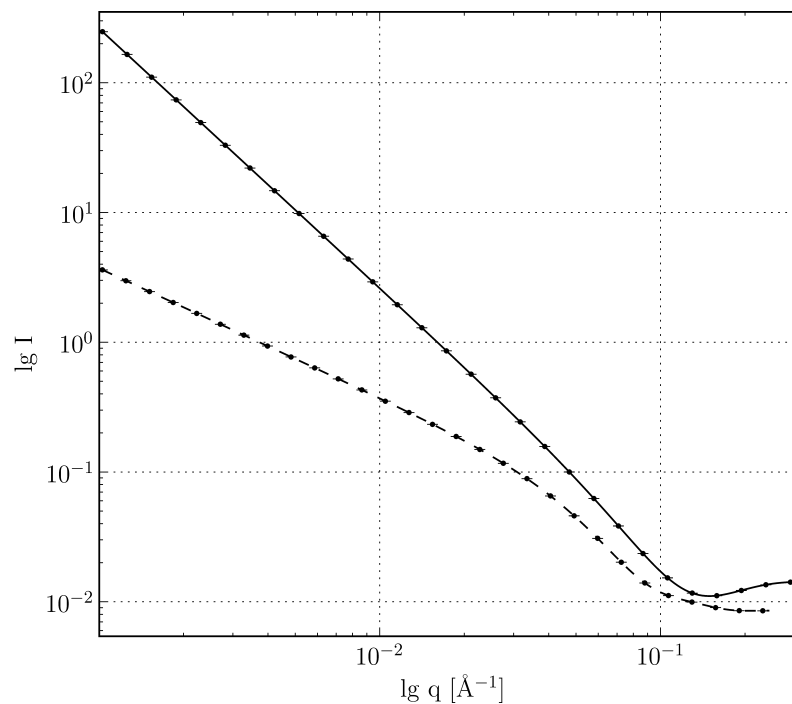


Abbildung 4.43: SANS-Daten des wt-Hämolytins in Vesikeln (—) und der Mutante H35R in Vesikeln (---). Zu beachten ist die doppelt logarithmische Skala im Gegensatz zu den bisherigen Darstellungen. Gegeben sind exemplarische Fehlerbalken.

Abbildung 4.43 zeigt die beiden Streukurven für den rekonstituierten Wildtyp und die präporformende Mutante H35R. Wie deutlich zu sehen ist, hat das „Matching“ sehr gut funktioniert, denn Vesikel streuen mit ausgeprägten Minima, ähnlich wie bei Hohlkugeln (Svergun & Koch, 2003). Es ist also ausschließlich die Streuung des Proteins zu erkennen.

<sup>9</sup>Vesikelstreuung ist ähnlich einer Streuung für Hohlkugeln (Svergun & Koch, 2003).

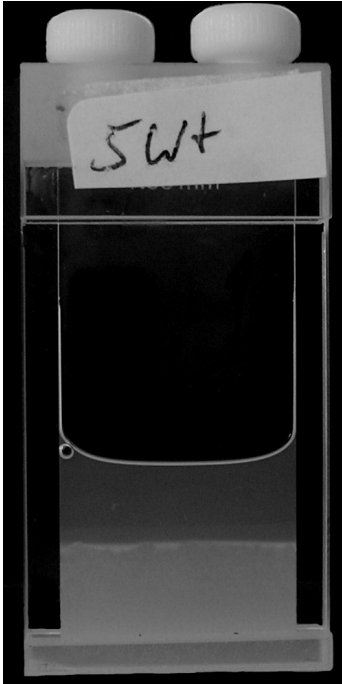


Abbildung 4.44: Trennung von Probe (unten) und Puffer (oben) nach etwa einem Tag unter Gravitationseinfluss. Hier am Beispiel der Toxinwildtypprobe.

Dennoch zeigen die Streukurven nicht „freies“ Protein, sondern definierte Aggregate an. Oder besser gesagt: Die Proteine sind in großen „Aggregaten“ organisiert. Das linksstehende Photo ist ein Indiz dafür, dass diese Annahme richtig ist: Offenbar sind die Proben – die Probe für die Mutante H35R zeigte ein ähnliches Bild – unter Gravitationseinfluss innerhalb eines Tages sedimentiert. Die Probe war vorher homogen verteilt, die Trübung beruht auf der Größe der Vesikel. Damit ist naheliegend, dass vor allem das Protein für die Sedimentation verantwortlich ist.

Die Wildtypprobe sedimentierte etwas schneller als die Probe mit der H35R-Mutante. Dies ist womöglich auf die unterschiedlichen Proteinkonzentrationen zurückzuführen. Für die Probe zu der das wt- $\alpha$ -Hämolyisin zugesetzt wurde, konnte eine effektive Proteinkonzentration von 10,3 mg/mL bestimmt werden. Für die zweite Probe, zu der H35R- $\alpha$ -Hämolyisin zugesetzt wurde ergab sich die Proteinkonzentration zu 5,9 mg/mL. In beiden Fällen wurde die Konzentration mit dem Bradford-Assay bestimmt (siehe Abschnitt 3.10.2 auf Seite 47 des Methodenkapitels).

Die Interpretation der Gyrationradien für die verschiedenen Vesikeltypen (siehe Abschnitt 3.10.7, Seite 51) zeigt schließlich aber doch noch ein interessantes Resultat in der folgenden Tabelle (Nr. 4.13). Die Membrandicke wird – aufgrund des Lipidmatchings – nicht durch die Lipiddicke vorgegeben, sondern durch die auf der Membran sitzenden Proteine. Die zu den Gyrationradien korrespondierende Dicke ist ungefähr die Hälfte der Länge der Protomere in der Pore, respektive Präpore.

Tabelle 4.13: Gyrationradien und Membrandicken nach den SANS-Daten.

Mutante	Angenommener Körper	$R_{T,g}$ [Å]	$R_{CS,g}$ [Å]	korrespondierende Membrandicke [Å]
wt	Scheibe	$11,5 \pm 0,4$		$39 \pm 2$
H35R	Zylinder		$24,1 \pm 0,9$	$34 \pm 2$

## 4.4.3.4 Elektronenmikroskopische Daten

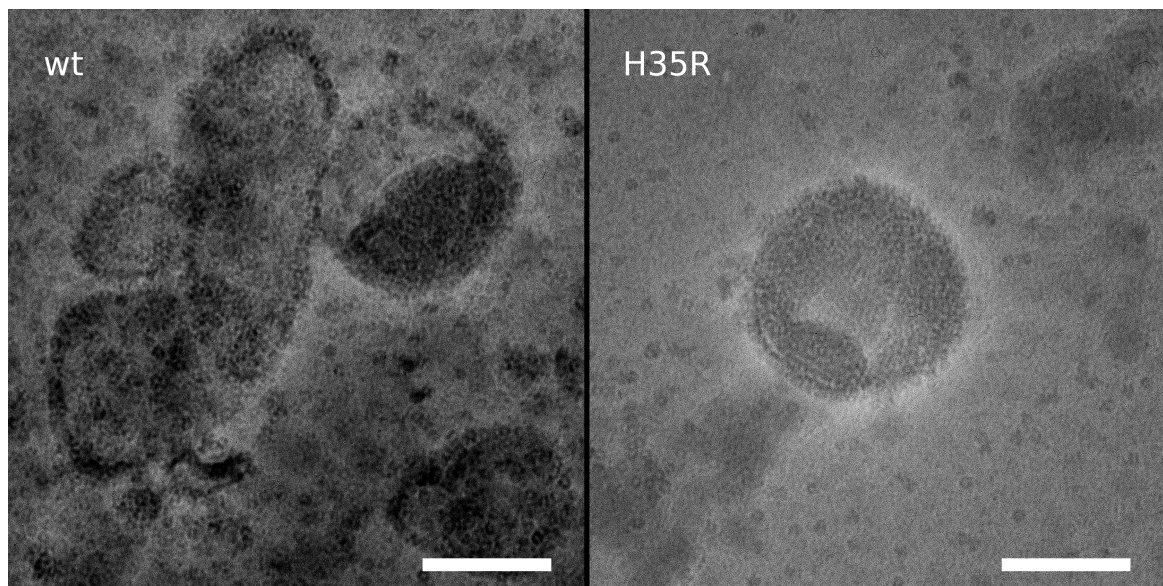


Abbildung 4.45: Elektronenmikroskopische Übersichtsaufnahme über die Proben nach Neutronenbestrahlung. Der Maßstabsbalken ist jeweils 100 nm lang.

Nach Abschluss der SANS-Messungen musste geklärt werden aus welchem Grund, dass das Protein, nicht wie in einer Referenzarbeit mit vergleichbarer Methodik (Bu *et al.*, 2003), als homodisperse Lösung vorlag, sondern in Form bestimmter „Aggregate“. Zur Klärung dieser Frage wurden elektronenmikroskopische Aufnahmen angefertigt. Abbildung 4.45 zeigt die bereits bestrahlten Proben in Übersichtsaufnahmen. Zudem wurde das Verhalten der Proben in der dynamischen Lichtstreuung untersucht. Da die Vesikel, zur Kontrolle ihrer Größe, bereits vor Zugabe des Proteins mit dieser Methode untersucht worden waren, existierten Referenzdaten.

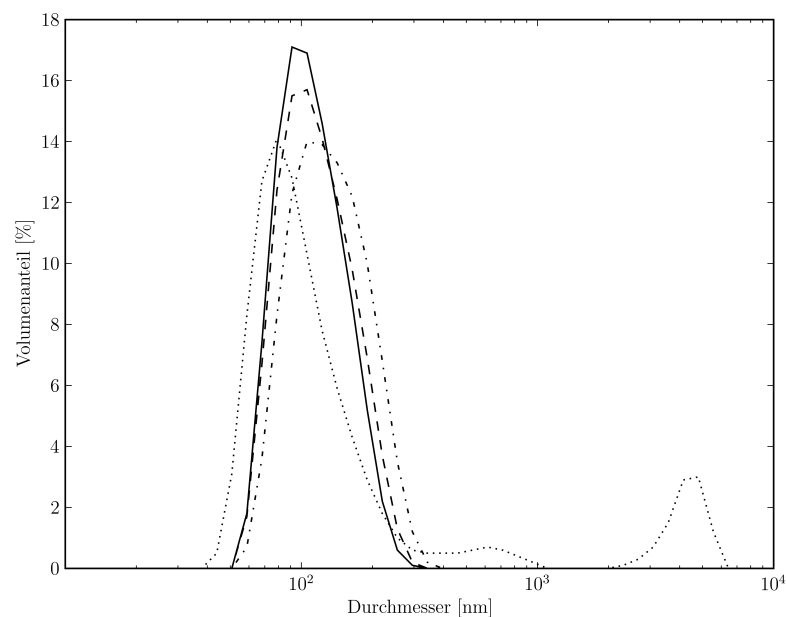


Abbildung 4.46: Daten der Dynamischen Lichtstreuung für Vesikelproben. Gezeigt ist der Volumenanteil für Vesikel vor (—) und nach (---) Zugabe der H35R-Mutante, sowie die Daten für Vesikel vor (····) und nach (· · ·) Zugabe des Wildtyp-Hämolytins.

Abbildung 4.46 zeigt die DLS-Daten vor und nach Zugabe von Protein für beide Proben. Sie

bestätigen, dass die Vesikel durch die Zugabe des Proteins im Wesentlichen unverändert blieben. Lediglich kleine Teile der mit wt- $\alpha$ -Hämolyysin behandelten Vesikel wurden zu größeren Partikeln umgesetzt. Der Peak bei  $3,5 \cdot 10^3$  nm beruht wahrscheinlich auf Zusammenlagerungen kleinerer Vesikel. Jedenfalls zeigen elektronenmikroskopische Bilder keine Vesikel oder Aggregate dieser Größe, sondern allenfalls aneinandergelagerte Vesikel.

Wie Abbildung 4.45 schon erkennen ließ, sind in beiden Proben die Proteine auf den Vesikeln dicht gepackt. Es kam die Frage auf, ob die Proteine in den Vesikeln sogar derart dicht gepackt sind, dass sie 2D-Kristalle formen. Daher wurden diese einer ersten elektronenkristallographischen Analyse unterzogen. Die folgende Abbildung 4.14 zeigt den besten Kristall. Deutlich ist zu sehen, dass das Objekt nur Reflexionen erster Ordnung zeigt. Und auch mit dem bloßen Auge wird deutlich, dass die Ordnung der Proteine nicht sehr gut ist. Zudem ist die Ordnung in diesen parakristallinen Vesikeln uneinheitlich. Dies ist auch in der folgenden Tabelle 4.14 zu sehen, in der einige Kristalle hinsichtlich der Größe der Einheitszellen ausgewertet wurden.

Es ist somit deutlich, dass eine kristallographische Analyse im engeren Sinne nicht möglich war.

Tabelle 4.14: *Einheitszellen auswertbarer 2D-Toxinkristalle*. Gegeben sind die Vektoren der Einheitszellen ( $\vec{a}$  &  $\vec{b}$ ) und der Winkel  $\gamma$ , der von diesen Vektoren eingeschlossen wird.

Mutante	$\vec{a}$ [Å]	$\vec{b}$ [Å]	$\gamma$
wt	62,6	60,1	117,0°
	57,8	59,2	114,1°
	66,0	72,3	132,6°
H35R	54,1	60,5	115,3°

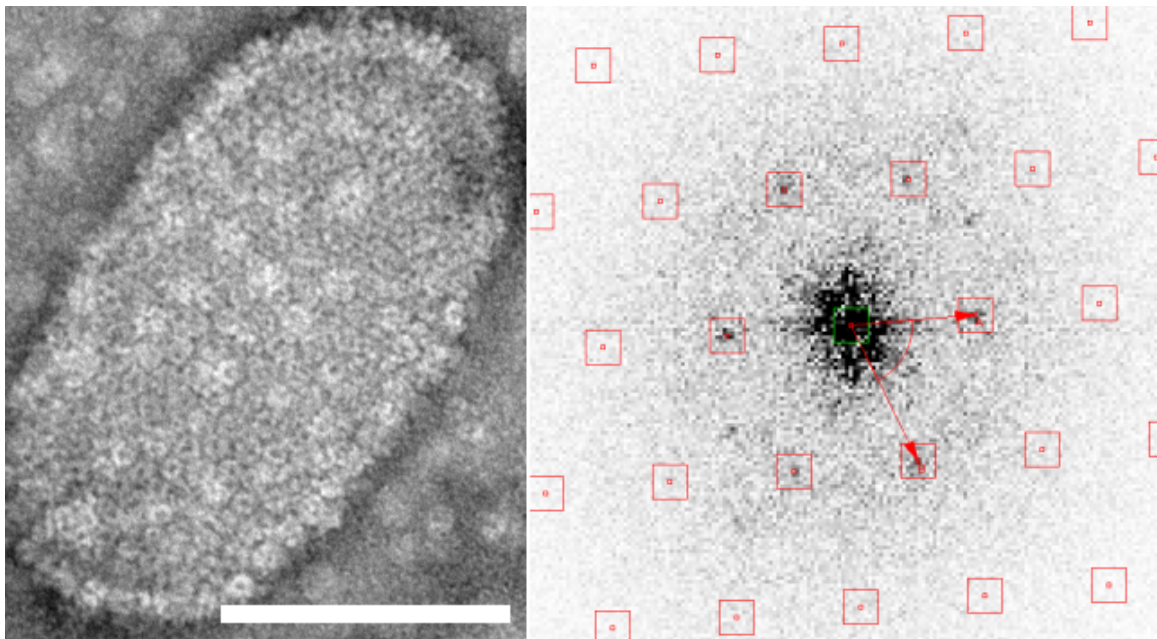


Abbildung 4.47: *Parakristallines  $\alpha$ -Hämolyysin im Negativkontrast (kontrastinvertiert) und das korrespondierende Powerspektrum (kontrastinvertiert)*. Der Maßstabsbalken entspricht etwa 100 nm. Das Gitter im Powerspektrum ist eingezeichnet. Reflexionen erster Ordnung sind erkennbar; ein zweites Gitter von der Unterseite des Vesikels ist in den nicht indizierten Reflexionen zu sehen. Die Vektoren des Gitters entsprechen dem ersten Eintrag aus Tabelle 4.14.

Die Abbildung zur Rechten ist das Ergebnis einer *single particle*-Mittelung. Die Partikel mussten manuell gewählt werden, da der Kontrast nicht gut und die Proben mit dem Lipid gemischt waren. Es konnten 1495 Partikel für die Wildtypprobe eindeutig identifiziert werden. Für die H35R-Probe waren dies 1087 Partikel. Die Aufnahmen wurden so eingescannt, dass ein Pixel zwei Ångström auf der Probenebene entsprach.

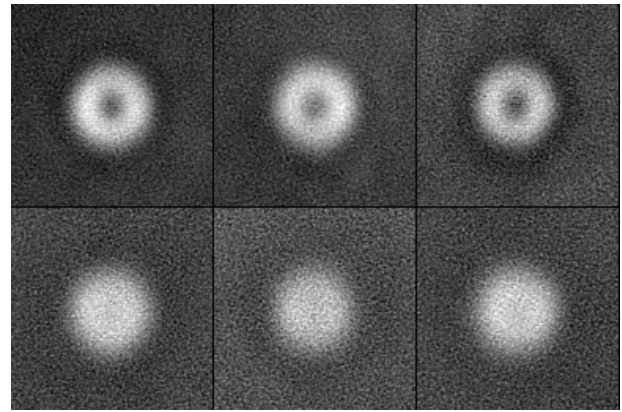


Abbildung 4.48: Projektionen des  $\alpha$ -Hämolytins auf Vesikeln; wt (obere Reihe) und H35R (untere Reihe). Für beide Proben ergaben sich drei Klassen. Die zugehörigen Varianzen (vgl. Abbildung 4.23) sind unimodal. Den Projektionen wurde keine Symmetrie aufgezwungen. Die Kantenlänge der Kästchen beträgt 25,6 nm.

Die Projektionen zeigen wenig Detail. Dennoch ist offensichtlich, dass für den Wildtyp kaum Dichte im Zentrum des Partikels festgestellt werden kann. Für die Mutante findet sich zwar ebenfalls eine leichte Dichtedepression im Zentrum, jedoch keineswegs so ausgeprägt. Der Durchmesser (gemessen von der Kante des Dichteanstiegs) der Partikel beträgt in beiden Fällen etwa  $100 \text{ \AA} \pm 8 \text{ \AA}$  und entspricht damit in etwa dem Durchmesser von  $100 \text{ \AA}$ , der auch kristallographisch für den Porenkomplex ermittelt wurde (Song *et al.*, 1996).

Alle Projektionen stellen „Aufsichten“ entlang der Rotationschse des Poren- beziehungsweise Präporenkomplexes dar. Die Unterschiede, die eine Einordnung in verschiedene Klassen ergaben, ergeben sich nicht auf den ersten Blick aus den Bildern. Sie sind wahrscheinlich Rauschen in den Daten sowie einer leichten Variabilität in der Kontrastmittelverteilung zuzuschreiben.

#### 4.4.3.5 Konformationswechsel in den Monomeren während der Oligomerisierung

Ausgehend von der Monomerstruktur wurde eine fiktive Präporenstruktur gebaut, in dem sieben Monomere in die Porenstruktur (Song *et al.*, 1996; PDB-Code: 7AHL) gelegt wurden. Nach einer kurzen Energieminimierung (Krieger *et al.*, 2002) ohne explizites Wasser zeigten die Aminosäurereste, die den Kontakt formen, keine sterische Überlappung. Die Kontaktfläche zwischen den Monomeren entspricht weitgehend der Kontaktfläche der Protomere der Porenstruktur. Einige der kontaktformenden Aminosäurereste sind identisch für diese fiktive Präpore und die Porenstruktur. Andere, welche nicht in der Lage sind Kontakt mit den korrespondierenden Stellen des benachbarten Protomers zu formen, müssen einen Konformationswechsel vollziehen, um in die finale Porenstruktur gelangen zu können (vgl. Abbildung 4.49).

Die Abbildung 4.49 auf der nächsten Seite zeigt den Vergleich der kontaktformenden Aminosäuren bezogen auf ein Monomer in der fiktiven Präporenkonformation und in der Porenkonformation. Als „kontaktformend“ wurde eine Aminosäure eingestuft, sofern diese nicht mehr als 5 Ångström von kontaktformenden Aminosäuren im benachbarten Protomer entfernt ist. Es ist evident, dass ein großer Teil der kontaktformenden Aminosäuren für beide Konformationen identisch ist. Eine quantitative Beurteilung der Kontakte wäre allerdings willkürlich, da die Aminosäurekonfigurationen der fiktiven Präpore zu keinem Zeitpunkt hinsichtlich der Kontakte mittels einer Molekulardynamiksimulation optimiert wurden. Dies ist auch nicht mit den Computerressourcen des Institutes für Molekulare Biophysik in vertretbarer Zeit zu berechnen, da die Bindung der Protomere der Präpore an eine zuvor energieminimierte Membran zu berechnen wäre

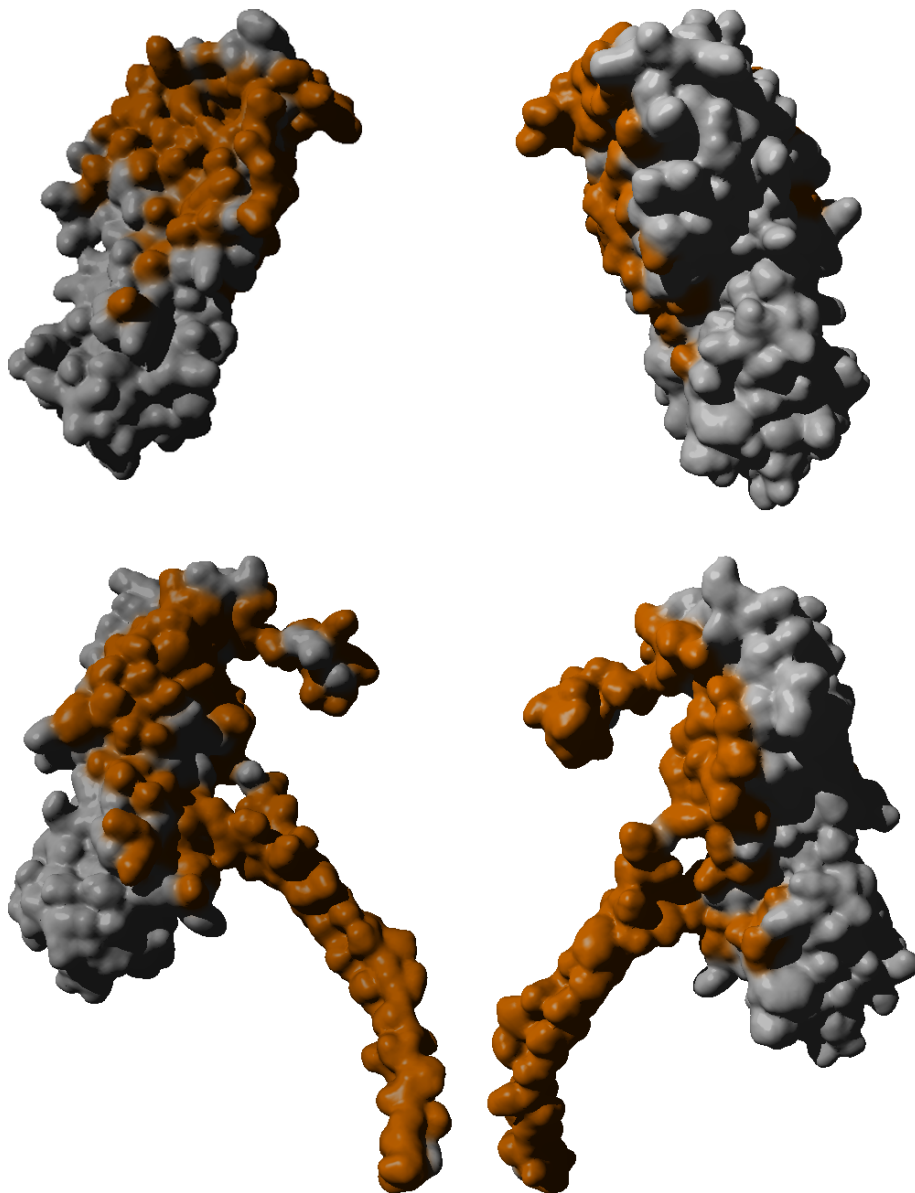


Abbildung 4.49: *Protomer-Protomer Kontaktflächen in der fiktiven Präpore und der ausgeprägten Porenstruktur*. Gezeigt ist eine Oberflächendarstellung der Protomere. Aminosäuren, die zur Kontaktfläche in den Strukturen beitragen sind farblich hervorgehoben. Oben: Flip- und Flop-Darstellung eines Präporenprotomers. Unten: Flip- und Flop-Darstellung eines Protomers in der Porenkonformation.

Abbildung 4.51 auf Seite 100 zeigt das Ergebnis eines *Morphings* (siehe Abschnitt 3.8.4 auf Seite 43 des Methodenkapitels). Die fiktive Präpore wurde hierbei in die Porenstruktur überführt. Dies entspricht dem Übergang  $\alpha_7^* \rightarrow \alpha_7$  unter der Voraussetzung, dass die fiktive Präpore einem Zustand in  $\alpha_7^*$  entspricht (siehe Einleitung Seite 13 ff.). Klar zu erkennen ist, dass sich der Komplex übergangsweise weiten muss, damit sich die *stem*-Domäne ausbilden kann (vgl. Abbildung 1.13 auf Seite 13). Diese Domäne schraubt sich quasi in die Membranebene hinein.

Abbildung 4.50 zeigt den energetischen Verlauf des Konformationswechsels. Der geplottete Energiebetrag ist mit dem YAMBER2-Kraftfeld (Krieger *et al.*, 2004) ohne Lipide in der *stem*-Domäne berechnet. Berücksichtigt wurden hierbei Bindungsenergien innerhalb des Proteins (Bindungsenergien, Winkelspannungen, Spannung in dihedralen Winkeln, Beiträge von Planaritätsabweichungen), also sterische Komponenten, und zusätzlich Van-der-Waals- und Coulomb-Wechselwirkungen.

Abbildung 4.50 ist nur im direkten Vergleich mit den Abbildungen 4.51 und 4.52 auf den nächsten Seiten zu verstehen. Außerdem muss Berücksichtigung finden, dass die Energiebeiträge in simuliertem Wasser ermittelt wurden. Daher können die Energien der letzten Schritte des *Morphing*

nicht dargestellt werden, da die Energiebeträge extrem groß werden, weil hydrophobe Seitenketten unrealistischer Weise in Wasser statt in Lipiden exponiert sind. Hierauf ist auch der Anstieg gegen Ende der Kurve zurückzuführen. Realistisch ist die Kurve somit nur bis zum Beginn der Penetration in die Membranebene (vgl. Abbildung 4.51) zu interpretieren.

Es ist dennoch hiermit bewiesen, dass zur Umfaltung in die Porenkonformation Energie aufgewendet werden muss. Folglich ist die fiktive Präpore energetisch zumindest semistabil. Da bekannt ist, dass das Präporenstadium in drei verschiedene intermediäre Stadien aufgeteilt ist (Valeva *et al.*, 1997), ist es naheliegend energetische Hürden zumindest teilweise für diese Aufspaltung in Stadien verantwortlich zu machen (siehe Abschnitt 5.9.3 auf Seite 139 des Diskussionskapitels).

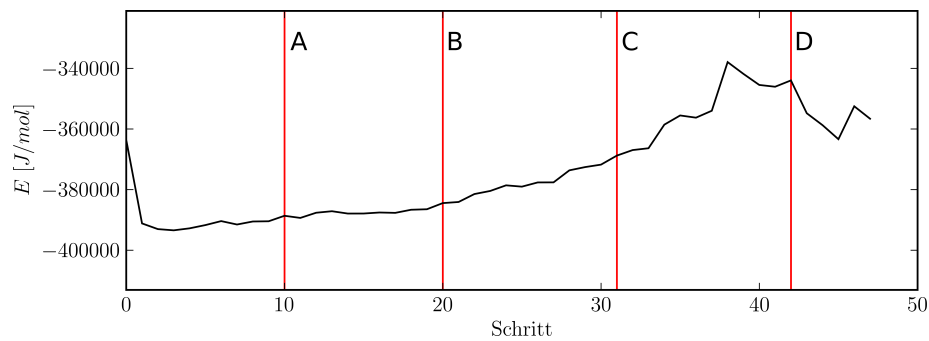


Abbildung 4.50: *Energieprofil der einzelnen Morphing-Schritte in Wasser gemessen mit dem YAMBER2-Kraftfeld.* Die roten Linien markieren von links nach rechts: A. Start der Weitung der Präpore, B. Beginn der Umfaltung des Transmembransegmentes im Innern der Präpore, C. Das Transmembransegment trifft auf die Membran, zugleich müssen sich die Segmente verdrillen, D. Das Transmembransegment relaxiert auf den finalen Durchmesser.

Die einzelnen Schritte des durch *Morphing* simulierten Konformationswechsels sind folgendermaßen zu beschreiben:

- A. Die Sequenzabschnitte des späteren Transmembransegmentes entfalten sich in die Mitte des Hohlzylinders, der durch die Präpore dargestellt wird. Durch gegenseitige sterische Hinderung wird eine leichte Weitung der Protomere an der Membran induziert.
- B. Die weitere Entfaltung des späteren Transmembransegmentes induziert eine Verdrehung der Protomere. Von der Membranseite aus gesehen erfolgt diese Verdrehung im Uhrzeigersinn (siehe Abbildung 4.52).
- C. Das Transmembransegment trifft auf die Membran und die Verdrehung der Protomere und die gegenseitige Verdrehung der entstehenden antiparallelen  $\beta$ -Faltblätter des Transmembransegmentes bedingen einander gegenseitig.
- D. Mit dem Fortschreiten der Ausstülpung des Transmembransegmentes spielt sich dieses auf den finalen Durchmesser ein. Damit verbunden beginnen die N-terminalen Aminosäurereste den erforderlichen Raum ebenfalls umzufalten und den Kontakt zum benachbarten Protomer zu formen.
- E. In den letzten Schritten des *Morphings* erreicht die Konformation des Komplexes die Struktur des Porenzustandes (Song *et al.*, 1996). Hierbei findet nur ein Re-Arrangieren der Aminosäurekonfigurationen statt.

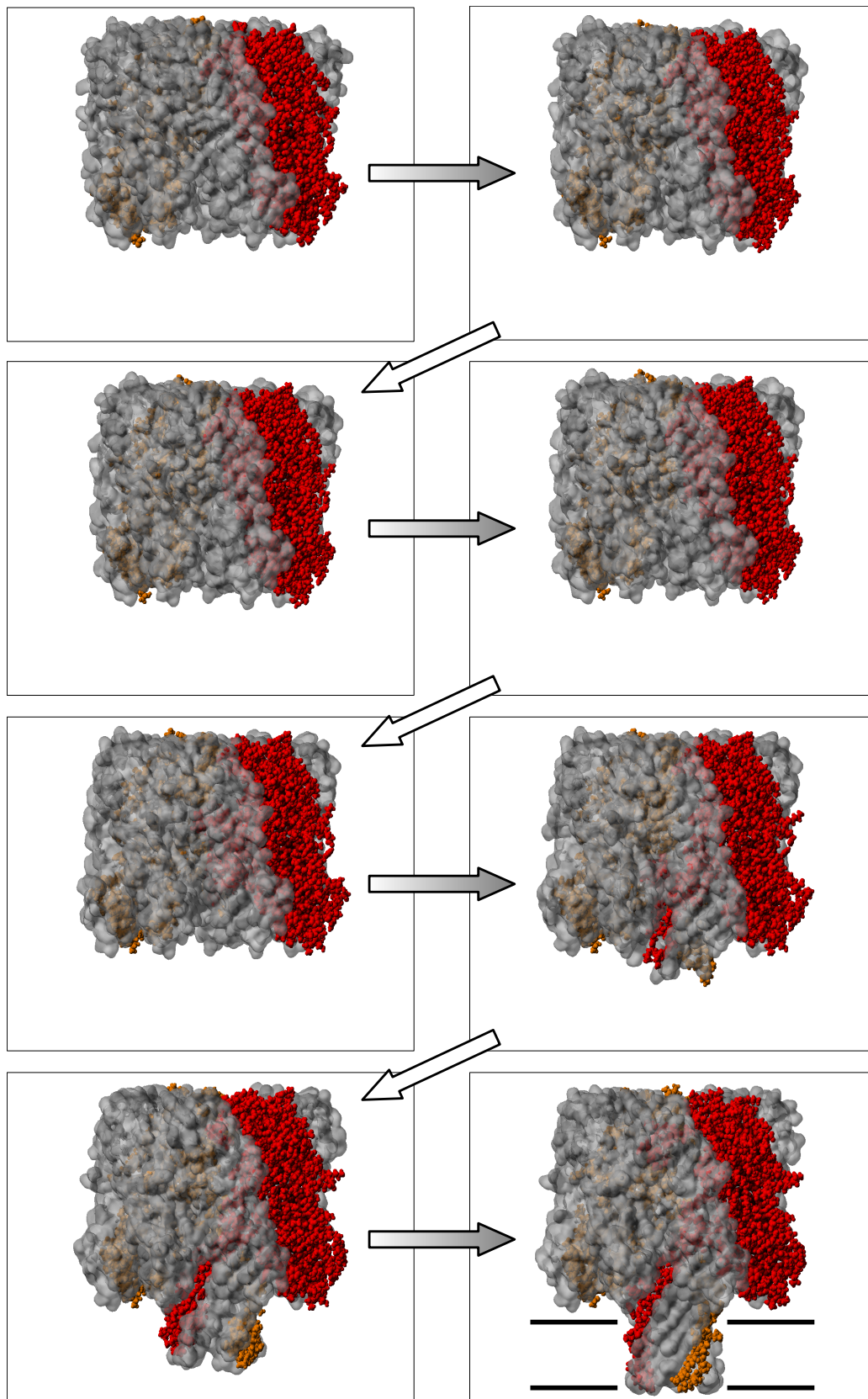


Abbildung 4.51: *Resultat des Morphings für den Konformationsübergang  $\alpha_7^* \rightarrow \alpha_7$ . Zwei Untereinheiten sind farblich hervorgehoben, die restliche Struktur wird durch eine transparente Oberfläche repräsentiert. Die einzelnen Abbildungen stellen gleichmäßige Inkremente des Morphings gemessen an den RMSD-Werten dar. Die Membranebene ist in der letzten Abbildung der Sequenz angedeutet.*

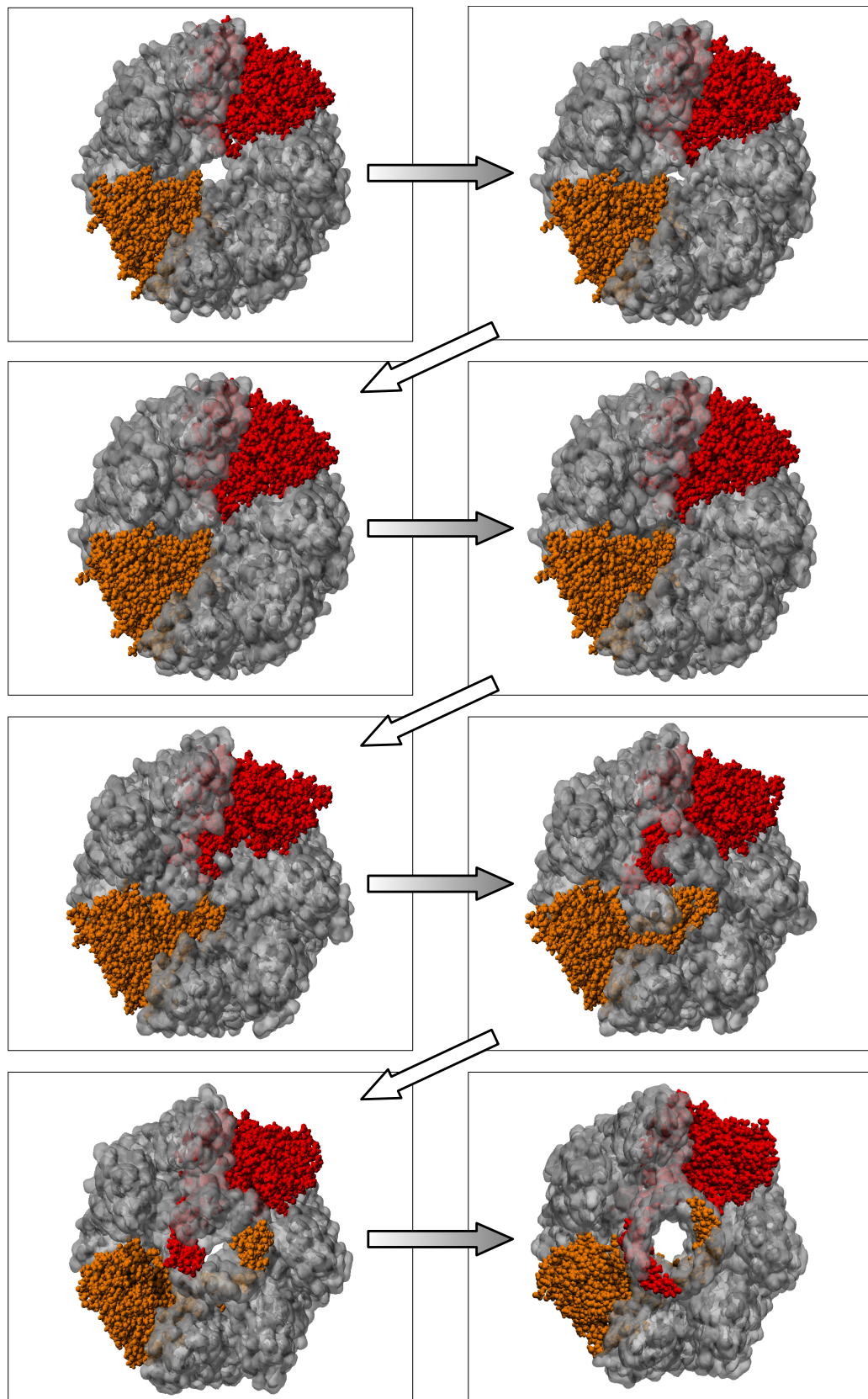


Abbildung 4.52: *Resultat des Morphings für den Konformationsübergang  $\alpha_7^* \rightarrow \alpha_7$ . Zwei Untereinheiten sind farblich hervorgehoben, die restliche Struktur wird durch eine transparente Oberfläche repräsentiert. Die einzelnen Abbildungen stellen gleichmäßige Inkremente des Morphings gemessen an den RMSD-Werten dar. Alle Einzelabbildungen sind im Vergleich zu Abbildung 4.51 um  $90^\circ$  gedreht, so dass der Betrachter von der zellulären Seite der Membran blickt.*

## 4.5 Molekulargewichtsbestimmung mit Hilfe des Kratky-Plots

Wie in Abschnitt 3.6.7 auf Seite 37 beschrieben ist es prinzipiell möglich, durch Streudaten ohne absolute Intensitätskalibrierung (vgl. Abschnitt 3.6.6, Seite 36) mit Hilfe des Kratky-Plots das Molekulargewicht eines Proteins zu ermitteln – sofern es Referenzdaten für Proteine mit bekanntem Molekulargewicht gibt. Ziel der hier beschriebenen Ergebnisse ist es

- eine Datengrundlage für die Experimente im Institut für Molekulare Biophysik zu haben, die es ermöglichen die Molekulargewichte von Proteinen mittels SAXS an der Kratky-Kamera zu bestimmen und
- das Prinzip dieser speziellen Molekulargewichtsbestimmungsmethode zu verifizieren: Ist es möglich und wie gut funktioniert diese Methode?

Gemäß dem in Abschnitt 3.6.7 beschriebenen Vorgehen wurden die notwendigen Daten gesammelt und durch Logarithmieren linearisiert; siehe Abbildung 4.53. Es handelt sich ausschließlich um entschmierte Daten (vgl. 2.4), um eine Konzentrationsabhängigkeit der Daten ausschließen zu können. Die Datenpunkte liegen wie erwartet auf einer Geraden. Der lineare Korrelationskoeffizient ist 0.997. Für die Regressionsgerade ergibt sich eine Steigung von  $-2,6 \pm 0,1$  und ein Achsenabschnitt von  $4,2 \pm 0,2$ . Die Abweichung des theoretischen oder anderweitig experimentell bestimmten Molekulargewichtes mit demjenigen Molekulargewicht, dass mit dieser Methode bestimmt wurde, liegt bei etwa 10 % bis 20 %, siehe Tabelle 4.5.

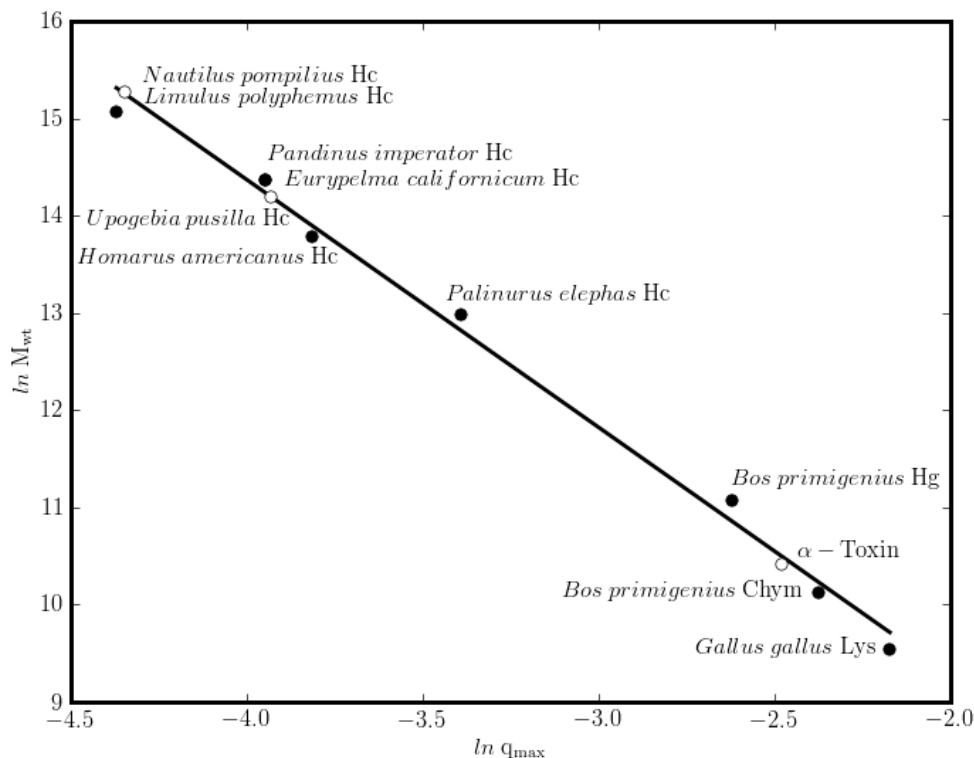


Abbildung 4.53: Eichgerade zur Bestimmung des Molekulargewichtes mit Hilfe des Kratky-Plots. Aufgetragen sind der natürliche Logarithmus des Molekulargewichtes gegenüber dem natürlichen Logarithmus des Wertes, für den Streuvektor für den der Kratky-Plot maximal ist. Mit ausgefülltem Symbol aufgetragen sind Proteine, die zur Kalibrierung verwendet wurden. Die durchgezogene Linie ist das Ergebnis einer linearen Regression; Einzelheiten im Text. Die Abkürzungen im einzelnen: „Chym“  $\triangleq$  Chymotrypsin, „Hc“  $\triangleq$  Hämocyanin, „Hg“  $\triangleq$  Hämoglobin und „Lys“  $\triangleq$  Lysozym.

Tabelle 4.15: Übersicht über die bei der Ergebnis der Kalibrierung in Abbildung 4.53. Angegeben ist ob ein Protein in der Kalibrierung verwendet wurde (Spalte „kal.“ ja oder nein), das Protein und seine Herkunft, der q-Wert bei welchem der Kratky-Plot sein Maximum hat, das Molekulargewicht nach der hier beschriebenen Methode, eine Referenzangabe für das Molekulargewicht, die Quelle für diese Referenzangabe (die Angabe „Packungsaufschrift“ bezieht sich auf die Molekulargewichtsangaben gekaufter Proteine) und die prozentuale Abweichung der beiden Molekulargewichte.

kal.?	Protein	max q	M <sub>w</sub> [Da] nach SAXS	M <sub>w</sub> [Da] durch alter- native Me- thode	Methode & Quelle	ΔM <sub>w</sub> [%]
nein	Hämocyanin von <i>Nautilus pom- pilus</i>	0.01293	4230701	?	entfällt	entfällt
nein	Hämocyanin von <i>Upogebia pusil- la</i>	0.01963	1460783	1500000	Statische Lichtstreuung, im Rahmen von Paoli <i>et al.</i> , 2007 <sup>†</sup>	3
nein	α-Hämolysin von <i>Staphylococcus aureus</i>	0.0836	36458	332484	Sequenz	9,7
ja	Lysozym von <i>Gallus gallus do- mesticus</i>	0.1139	16584	14000	Packungsaufschrift	18,5
ja	Chymotrypsin von <i>Bos primige- nius</i>	0.09303	27770	25000	Packungsaufschrift	11,1
ja	Hämoglobin von <i>Bos primigeni- us</i>	0.072685	52066	64500	Packungsaufschrift	19,3
ja	Hämocyanin von <i>Palinurus ele- phas</i>	0.034556	345967	432368	Sequenz (siehe 4.1.2.1)	20,0
ja	Hämocyanin von <i>Homarus ame- ricanus</i>	0.02207	1083900	965000	Statische Lichtstreuung, Hübner, 2001	12,3
ja	Hämocyanin von <i>Pandinus impe- rator</i>	0.01929	1527259	1735000	Statische Lichtstreuung, Waschbüsch, 2004	12
ja	Hämocyanin von <i>Eurypelma ca- lifornicum</i>	0.01929	1527259	1730000	Sequenz	11,7
nein	Hämocyanin von <i>Eurypelma ca- lifornicum</i>	0.01929	1527259	1709000	Statische Lichtstreuung, Hübner, 2001	10,6
ja	Hämocyanin von <i>Limulus poly- phemus</i>	0.012645	4477819	3501000	Statische Lichtstreuung, Hübner, 2001	11,7

<sup>†</sup> Der Wert wurde mittels statischer Lichtstreuung bestimmt, schwankte aber stark. Angegeben ist ein unsicherer Mittelwert.

## 4.6 Verfolgung einer Hitzedenaturierung mit der Kratky-Kamera am Beispiel von Lysozym

Will man die Hitzedenaturierung eines Proteines mittels SAXS strukturell verfolgen, so ist es hilfreich komplementäre Daten, wie die Schmelzkurve des Proteins (zu Schmelzkurven mittels DSC siehe Abschnitt 3.11, Seite 53), zu gewinnen. Abbildung 4.54 zeigt eine solche Schmelzkurve beispielhaft für das Lysozym des Hühnereiweiß (zu Lysozym siehe Blake *et al.*, 1965) in Phosphatpuffer (siehe Abschnitt 3.11, Seite 53). In der vorliegenden Arbeit geht es lediglich darum zu zeigen, dass solche Experimente an der Kratky-Kamera des Institutes nach dem Einbau der temperierbaren Zelle (siehe 3.6.1.3, 32) möglich sind. Also wurde der Schmelzpunkt des Proteins bestimmt und Streudaten gemäß Abschnitt 3.6.5 bei 20 °C und bei 80 °C bestimmt.

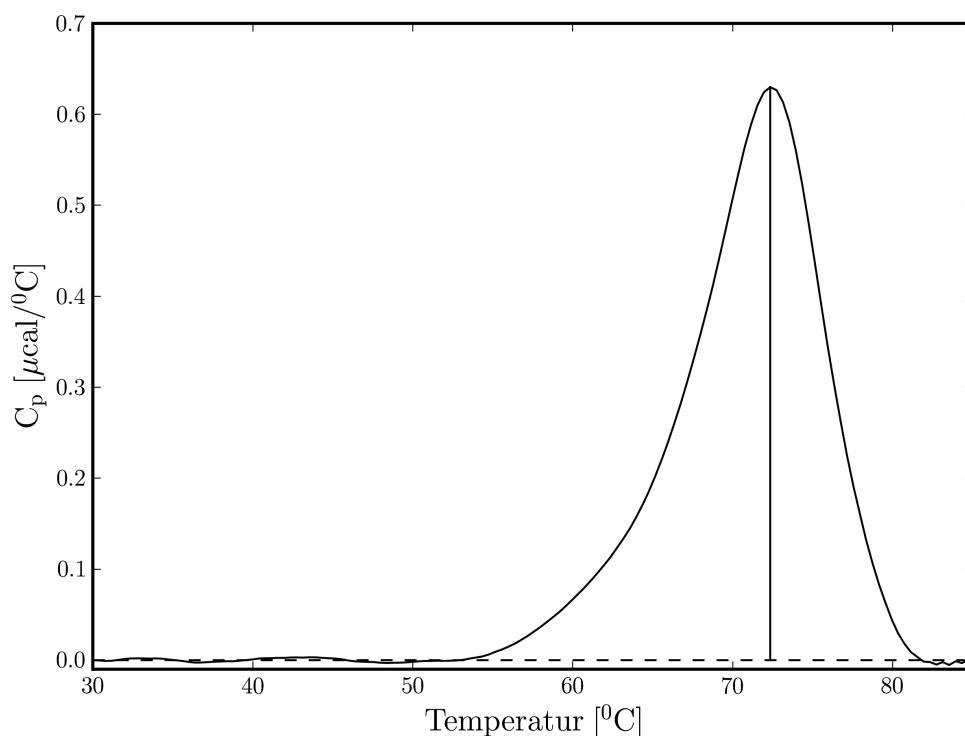


Abbildung 4.54: Schmelzkurve von Hühnerlysozym (1 mg/mL) in Phosphatpuffer gemessen mit DSC. Deutlich zu erkennen ist der Schmelzpunkt bei  $\sim 72.3^\circ\text{C}$  (kenntlich gemacht durch die Senkrechte beim Maximum der Kurve). Die gestrichelte Linie ist die Nulllinie nach Abzug des Hintergrundes.

Das Ergebnis ist in Form eines Vergleiches der Abstandsverteilungsfunktionen und zweier Modelle in Abbildung 4.55 auf der nächsten Seite dargestellt. Die Abstandsverteilungsfunktionen zeigen denselben maximalen Abstand innerhalb des Moleküls ( $D_{\text{max}} = 70 \text{ \AA} \pm 2 \text{ \AA}$ ). Es ist jedoch anzumerken, dass die Fehler(-balken) aufgrund der kurzen Strahlzeit (siehe Abschnitt 3.6.5, Seite 36) relativ groß sind. Die Veränderungen, welche auch die Modelle zeigen, sind klein, aber deutlich: Das Protein erfährt bei der Hitzedenaturierung offenbar eine „Komprimierung“ ohne jedoch dabei völlig zusammenzufallen, also apparent zu „schrumpfen“. Die allgemeine Form eines prolaten Ellipsoides wird beibehalten. Dennoch zeigen ab initio-Modelle (siehe Abschnitt 3.6.10, Seite 38), welche mit dem atomaren Modell (PDB-Code: 2LYZ, Diamond, 1974) übereinandergelegt wurden, die angesprochene Veränderung sehr deutlich (siehe Abbildung 4.55 auf der nächsten Seite).

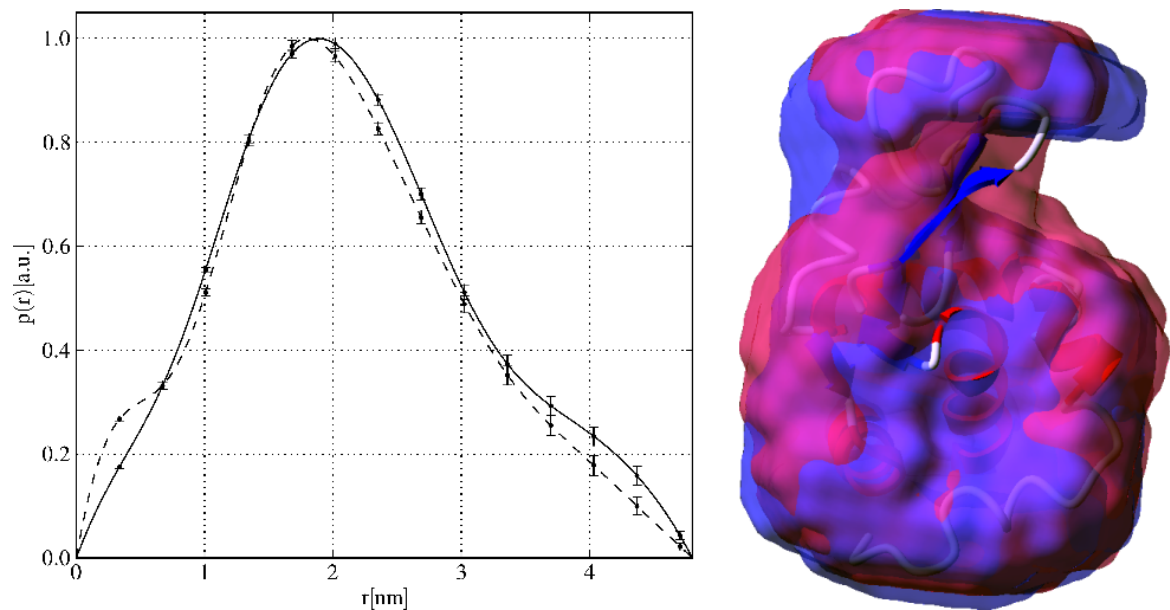


Abbildung 4.55: *Abstandverteilungsfunktion  $P(r)$  und Modell nativen und hitzedenaturierten Hühnerlysozyms.* In der  $P(r)$  ist das native, bei 20 °C bestrahlte Lysozym mit einer durchgezogenen Linie dargestellt, während das bei 80 °C bestrahlte mit einer gestrichelten Linie dargestellt ist; jeweils mit exemplarischen Fehlerbalken. Die Modelle im rechten Teil der Abbildung sind mit DAMMIN gerechnete *ab initio*-Modelle. Blau gezeigt ist das native, bei 20 °C temperierte Lysozym und rot das bei 80 °C bestrahlte. Zu beachten ist, dass das erhitzte Protein wesentlich kompakter wirkt, obwohl sich seine Größe nicht verändert.

Die SAXS-Messung einer Temperaturreihe in Inkrementen von 10 °C wurde ebenfalls durchgeführt. Die Kratky-Kamera erfordert jedoch für gewöhnlich lange Strahlzeiten. Dies erklärt die schlechte statistische Güte der durchgeführten Temperaturreihe. Es war nicht möglich, die Daten in irgendeiner Weise zu interpretieren oder auch nur in eine übersichtliche Darstellungsform zu überführen (Hilmenyuk & Schwiering, 2006). Eine detaillierte Diskussion findet sich unter Abschnitt 5.4, Seite 129.

## 4.7 Eine Software zur *rigid body*-Modellierung von Proteinen auf SAS-Datenbasis

Bereits vor Beginn dieser Doktorarbeit gab es eine ganze Reihe von Programmen zur *rigid body*-Modellierung auf Basis von Kleinwinkelstreudaten. Ein vergleichender Überblick über diese Programme findet sich in der Diskussion unter Abschnitt 5.10, Seite 142 ff. Die wesentlichen Nachteile dieser Programme sind jedoch an dieser Stelle bereits erwähnenswert, denn die Funktionen der hier vorgestellten Software füllen diese Lücken:

- Die meisten Programme können nur mit einer eingeschränkten Zahl von Symmetrien umgehen.
- Eine numerische Analyse der geometrischen Verhältnisse von Proteinuntereinheiten zueinander gestaltet sich schwierig oder ist gar nur über Programme anderer Anbieter zu erreichen.
- Sämtliche Programme sind „Black Boxes“, d. h. dem Nutzer ist nicht klar, was das Programm im Hintergrund ausführt. Dies kann im Zusammenhang mit unzureichender Dokumentation leicht zu einer fehlerhaften Interpretation der Daten führen.
- kryptische Bedienungsanleitungen und schlecht dokumentierte Datenformate führen dazu, dass die Programme z. T. schlicht unbrauchbar sind<sup>10</sup>.

<sup>10</sup>Diese Programme werden hauptsächlich durch die Autoren der Programme zitiert; beispielsweise DALAI (Chacón *et al.*, 1998; Chacón *et al.*, 2000).

- Einige dieser Programme sind bestenfalls semiautomatisch und erwarten z. T. fehleranfällige Eingriffe der Nutzer oder komplementäre Daten anderer Methoden (z. B. Daten aus der analytischen Ultrazentrifugation).

Die selbstgeschriebene Software ist bewusst nicht als eigenständiges Programm, sondern als ein Modul der Programmiersprache Python geschrieben worden. Da Python leicht erlernbar ist, ermöglicht dieser Ansatz es einem Nutzer flexible Skripte, angepasst an die Problematik der eigenen Forschungsfragen, zu schreiben. Diese Software weist folgende Eigenschaften auf:

1. Keine Beschränkung auf bestimmte Symmetrien. Durch einen sehr einfachen Ansatz ist es möglich jedes Verhältnis – auch ein eigentlich nicht symmetrisches – in der Datenstruktur eines Skriptes abzubilden. Die Einschränkung auf Schönfliessymmetrien (Tsuruta & Johnson, 2001) entfällt völlig. Die einzige Einschränkung ist, dass nur eine endliche Zahl von Untereinheiten berücksichtigt werden kann<sup>11</sup>.
2. Eine numerische Analyse der vollzogenen Bewegung, inklusive einer Fehlereinschätzung für die Bewegungsparameter ist möglich.
3. Die Software fasst jede Untereinheit eines Proteins als einen zu bewegendem Körper auf. Einzelne Körper können auch während eines Programmlaufs festgehalten werden. Domänen einer Untereinheit können unabhängig voneinander bewegt werden, wenn diese Untereinheit vor Beginn eines Programmlaufs getrennt wurden. Die Software bietet Unterstützung bei der Trennung und Zusammenführung von Untereinheiten.
4. Anbindung an YASARA (Krieger *et al.*, 2002). Durch YASARAs einfache Handhabung ist es leicht, für Nutzer Strukturen so zu manipulieren, dass Domänen eines Proteins als Untereinheiten aufgefasst werden und gegeneinander bewegt werden können (s. o.). Überdies hilft YASARA bei der Analyse der Ergebnisse.
5. Obwohl die Software intern im kartesischen Koordinatensystem operiert ist die Analyse der Ergebnisse auch in einem polaren Koordinatensystem möglich. Dieses ist insbesondere bei Strukturen mit Punktsymmetrie anzuwenden.
6. Die Software ist umfangreich dokumentiert. Die Anleitung erhält jeder Nutzer mit dem Download der Software. Der Quellcode der Software ist offen (siehe Abschnitt 4.7.7).

Die Software wurde auf den Namen `sas-rigid` „getauft“. Damit wird zum Ausdruck gebracht das sie mit Kleinwinkelstreudaten (*small angle scattering data*) arbeitet und auf starre Körper (*rigid bodies*) appliziert werden kann.

Die weiteren Aspekte der Software werden in den folgenden Abschnitten erläutert.

---

<sup>11</sup>Die maximale Zahl der Untereinheiten hängt von der Rechnerarchitektur ab. Es ist nicht getestet worden, wo hier Grenzen liegen können.

## 4.7.1 Implementation

### 4.7.1.1 Programdesign

Wie in der nebenstehenden Abbildung schematisch dargestellt, ist die nutzerexponierte Seite der Software eine sogenannte Skriptingschicht: Nutzer der Software haben die Möglichkeit, durch eigene Skripte auf die Funktionalität zuzugreifen. Um dieses zu vereinfachen, stehen diverse „Nützlichkeitsfunktionen“ (*utility functions*) und Beispiele zur Verfügung. Die „Nützlichkeitsfunktionen“ subsumieren verschiedene Schritte häufig vorkommender, komplexer Operationen zur Präparation der Daten vor dem Start eines der zentralen Algorithmen (siehe Abschnitt 4.7.2, Seite 110 ff.). Diese Skriptingschicht stellt also quasi das Interface für eine Klasse zur Behandlung atomarer Modelle (**PDB-Model class** in der Abbildung) dar. Diese ihrerseits ist abgeleitet von einer allgemeinen Modellklasse. Damit wird erreicht, dass weitere abzuleitenden Klassen garantiert dieselbe Benutzeroberfläche haben werden. Hierzu zählt die nur teilweise implementierte **EM-Model class** für Elektronendichten niedriger Auflösung. Der Name der **EM-Model class** legt nahe, dass diese Dichten aus der Elektronenmikroskopie kommen sollen, jedoch sind auch niedrigaufgelöste Elektronendichten der makromolekularen Kristallographie denkbar. Beiden Klassen stehen diverse Module zur Gewährleistung der Funktionalität zur Verfügung. Die **PDB-Model class** beispielsweise nutzt das PDB-Modul des biopython-Projektes (Hamelryck & Manderick, 2003), während für die **EM-Model class** eigene Voxel- und Vektorklassen definiert wurden. Zur Berechnung der Abstandsverteilungsfunktion oder der Streukurven wird auf in C geschriebene Module oder externe Programme zurückgegriffen (siehe Abschnitt 4.7.1.3, Seite 108).

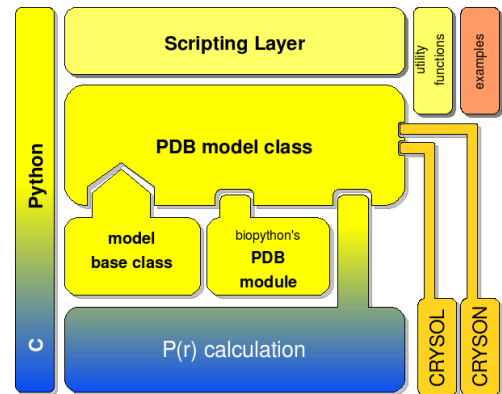


Abbildung 4.56: Aufbau des Softwarepaketes. Diese Abbildung ist Teil der Anleitung und bereits online publiziert (siehe Abschnitt 4.7.7).

### 4.7.1.2 Implementation von Standardalgorithmen

Sämtlichen Algorithmen basieren auf der Annahme, dass experimentelle Daten und theoretische Daten (Streukurven und Abstandsverteilungsfunktionen) für eine Auswertung von Proteinkomplexen gleicher Größe stammen. Somit muss gelten, dass

$$I(0)_{\text{experimentell}} = I(0)_{\text{Modell}} \quad (4.2)$$

beziehungsweise die gleichwertige Bedingung

$$\int P(r)_{\text{experimentell}} dr = \int P(r)_{\text{Modell}} dr \quad (4.3)$$

Die Integration wird nach der Simpsonschen Regel (Press *et al.*, 2002) ausgeführt. Da die Inkremente für  $q$  und  $r$  nicht notwendiger Weise für die experimentellen und die errechneten Daten gleich sind, wird vor einer Skalierung mittels C-Splines (Press *et al.*, 2002) interpoliert. Beide Teilalgorithmen sind in der `scipy`-Bibliothek integriert (Jones *et al.*, 2007). Gütefaktoren für die Abweichung experimenteller und theoretischer Werte ergeben sich aus

$$\chi^2 = \frac{1}{n-1} \sum_{i=1}^n \left[ \frac{c \cdot I_{\text{Modell}}(q_i) - I_{\text{experimentell}}(q_i)}{\sigma(q_i)} \right]^2 \quad (4.4)$$

beziehungsweise

$$\chi^2 = \frac{1}{n-1} \sum_{i=1}^n \left[ \frac{c \cdot P_{\text{Modell}}(r_i) - P_{\text{experimentell}}(r_i)}{\sigma(r_i)} \right]^2 \quad (4.5)$$

für  $n$  Datenpunkte.  $\sigma$  steht hierbei für den assoziierten experimentellen Fehler und  $c$  für den arbiträren Skalierungsfaktor zur Erfüllung der Bedingung 4.2 bzw. 4.3.

### 4.7.1.3 Berechnung der theoretischen Abstandsverteilungsfunktion und Streukurven

Während der überwiegende Teil der Software in Python (van Rossum, 1998) implementiert ist, wurde, aufgrund der größeren Geschwindigkeit kompilierter Sprachen, die Berechnung der Abstandsverteilungsfunktionen,  $P(r)$ , und Streukurven der Modelle,  $I(q)$ , in C implementiert. Dem Beispiel von Hartmann & Decker, 2002, folgend wird die  $P(r)$  berechnet durch

$$P(r) = \sum_i^N \sum_{j \leq i}^N |\vec{r}_i - \vec{r}_j| \quad (4.6)$$

für Bereiche  $\Delta r$ , für die gilt:

$$r - \frac{\Delta r}{2} \leq |\vec{r}_i - \vec{r}_j| < r + \frac{\Delta r}{2} . \quad (4.7)$$

Die Streufunktion,  $I(q)$ , ergibt sich danach aus Gleichung 2.19 (Seite 21 im Theoriekapitel). Sowohl die Berechnung von  $P(r)$  als auch von  $I(q)$  sind in der Software integriert. Zusätzlich haben die Nutzer die Wahl, die  $I(q)$  mit den Programmen CRY SOL (für Röntgenstreudaten) oder CRYSON (für Neutronenstreudaten; Svergun *et al.*, 1995) zu errechnen.

### 4.7.1.4 Softwremetrik

In der Informatik gibt es hunderte von Metriken für die Beurteilung einer Softwarequalität. Oftmals führt eine Orientierung an einer Metrik aber zur Dysfunktion: Um die Software an die Metrik anzupassen, wird entlang der Bedingungen des metrischen Systems programmiert. Deshalb folgen an dieser Stelle „roh-metrische“ Angaben zum Programmcode mit dem Ziel sorgfältige Programmierung zu belegen.

Sämtliche Pythonmodule wurden mittels `pylint` (<http://www.logilab.org/857>), einer Software zur Kontrolle von Codestandards in Python, kontrolliert. Dies half auch offensichtliche Fehler zu eliminieren. Im Übrigen folgt der Code, soweit möglich, PEP 8 (<http://www.python.org/dev/peps/pep-0008/>) einem Standard für die Pythonprogrammierung. Der C-Code folgt den Standards der GNU-Programmierung (<http://www.gnu.org/prep/standards/>).

Python offeriert die Anlage von sogenannten *Docstrings*. Diese erlauben die automatische Generierung von Hilfen, auch interaktiv im Pythoninterpreter, und haben zugleich bedeutende Dokumentarfunktion, die über übliche Codekommentare hinausgeht: Sie zwingen zu einer allgemeinen Beschreibung der einzelnen Funktionen, wodurch alle Parameter eingehend beschrieben werden. Zusätzlich können auch andere Kommentare im Code gesetzt werden. Tabelle 4.16 zeigt, dass sämtlicher Code ausführlich dokumentiert ist.

Zusätzlich zu den in Tabelle 4.16 aufgeführten Beispielen bietet die Software eine über zwanzigseitige Anleitung. Alle Dateien, inklusive der Anleitung, sind in einem „*subversion repository*“ <http://subversion.tigris.org/> archiviert. *subversion* ist eine Software zur Versionsverwaltung. Der veröffentlichte Teil von *sas-rigid* (siehe Abschnitt 4.7.7), auf dem auch die gezeigten Resultate beruhen, ist unter <http://svn.origo.ethz.ch/viewvc/sas-rigid/> unmittelbar einzusehen.

Tabelle 4.16: *Metrik der gesamten Software gegliedert nach Sprachen und Aufgaben.* Der von 100% abweichende Anteil beruht auf Leerzeilen und Rundungsfehlern<sup>†</sup>.

Sprache	Aufgabe	Dateien	Zeilen	Anteil		
				Code [%]	Dokumentation <sup>‡</sup> [%]	Kommentar [%]
Python	Kernbereich	17	3800	63	25	8
	Tests	8	478	67	17	9
	Beispiele	5	216	50	21	16
	Installationsroutine	1	214	75	13	8
C	Kernbereich	1	381	55	–	16
bash	Helferroutinen	2	58	38	–	43

<sup>†</sup> Gezeigt ist die Statistik für Version 0.3.3 der Software. Spätere Versionen werden davon abweichen. Version 0.4 (zum Zeitpunkt der Einreichung dieser Arbeit noch nicht fertiggestellt) umfasst den Code zum im Abschnitt 4.7.5 vorgestellten Algorithmus mit 3 zusätzlichen Dateien, sowie ca. 700 Zeilen mehr Code.

<sup>‡</sup> Dokumentation innerhalb von Programmcode (in sogenannten *Docstrings*) ist eine besondere Eigenschaft der Sprache Python.

#### 4.7.1.5 Die Verwirklichung von Symmetrieoperationen

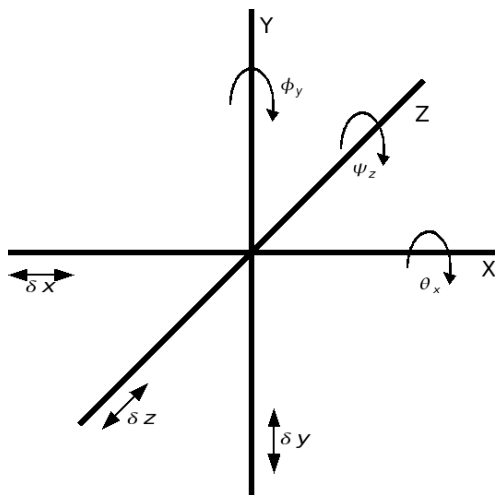


Abbildung 4.57: *Das kartesische Koordinatensystem mit den in sas-rigid verwendeten Bezeichnungen für die Rotationswinkel und Translationen.* Diese Abbildung ist Teil der Anleitung und bereits online publiziert (siehe Abschnitt 4.7.7).

Wie bereits ausgeführt sollte die Software u. a. zwei Dinge leisten: a) Es sollen während einer Simulation alle Untereinheiten gegeneinander bewegt werden können und b) soll die Software arbiträre Symmetrieangaben berücksichtigen können. Diese Bedingungen sind am leichtesten unter Verwendung eines kartesischen Koordinatensystems zu erfüllen: Sämtliche Daten in der PDB-Datenbank (Berman *et al.*, 2000) sind aus naheliegenden Gründen in diesem Koordinatensystem abgelegt und Symmetrieoperationen, die nicht nur eine Rotation erfordern, sind in diesem Koordinatensystem ebenfalls einfach umzusetzen.

Die nebenstehende Abbildung zeigt das verwendete Koordinatensystem. Die Software berücksichtigt Translationen  $dx$ ,  $dy$  und  $dz$  entlang der Achsen X, Y und Z, sowie Rotationen  $d\theta$ ,  $d\varphi$  und  $d\psi$  um diese Achsen<sup>12</sup>.

Symmetrieoperationen nach Bewegung einer Untereinheit werden innerhalb von *sas-rigid* umgesetzt, in dem eine Untereinheit aus dem Speicher des Rechners entfernt wird. Diejenige Untereinheit, für die angegeben ist, dass sie durch eine Symmetrieoperation in die gelöschte Untereinheit zu überführen ist, wird im Speicher dupliziert.

Und abschließend wird die zuvor spezifizierte Symmetrieoperation auf die duplizierte Untereinheit angewendet. Auf diese Weise sind bislang nur echte Symmetrieoperationen im Code realisiert. Es ist aber geplant, in einer zukünftigen Version (siehe Abschnitt 5.10.3 auf Seite 143 der Diskussion) Symmetrieoperationen im weiteren Sinn auch für nicht-identische Untereinheiten zu ermöglichen.

<sup>12</sup>Nutzer der Software können Grenzen für die Bewegungen in 'dx', 'dy', 'dz', 'theta', 'phi' und 'psi' angeben beziehungsweise innerhalb der Pythonskripte ansprechen.

### 4.7.2 Ein Monte Carlo-Algorithmus mit *Simulated Annealing*

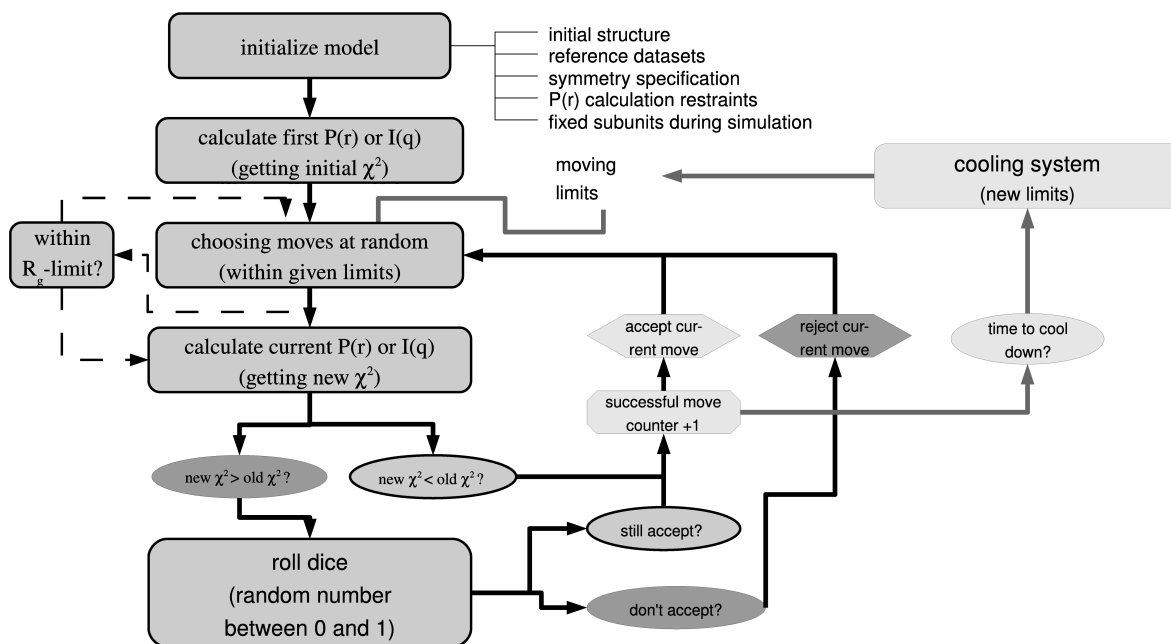


Abbildung 4.58: *Schema des Monte Carlo-Algorithmus*. Der Programmfluss wird durch ein vorab eingestelltes Abbruchkriterium eingeleitet. Weitere Erläuterungen siehe Text. Diese Abbildung ist bereits in Meesters, 2008a, publiziert.

Abbildung 4.58 zeigt eine schematische Darstellung des *Monte Carlo-Algorithmus* mit *Simulated Annealing*<sup>13</sup>. Dieser Algorithmus dient vor allem zur Einschränkung des Parameterraumes, so dass eine systematische Suche, wie sie in Abschnitt 4.7.3 erläutert wird, effizienter durchgeführt werden kann.

Der Algorithmus verlangt, dass ein Nutzer ein Modell initialisiert. Bei dieser Initialisierung können Rahmenbedingungen vorgegeben werden. Neben einem ersten atomaren Modell, das vorliegen muss, braucht es ebenfalls einen Kleinwinkelstreureferenzdatensatz. Alle weiteren Parameter sind optional:

- Die Symmetrie kann vorgegeben werden. Bislang können Symmetrie im engeren Sinne nur Homooligomere aufweisen (bezogen auf Version 0.3.3 der Software).
- Der Nutzer muss wählen, ob  $I(q)$  oder  $P(r)$  mit den experimentellen Daten verglichen werden soll. Im Falle von  $P(r)$  kann eine Toleranz für den Wert von  $D_{max}$  und das Inkrement für  $r$  vorgegeben werden.  $I(q)$  kann auf verschiedene Weise berechnet werden: CRY SOL (für SAXS-Daten), CRYSON (für SANS-Daten; für beide Programme: Svergun *et al.*, 1995) und die interne Methode nach Gleichung 2.19.
- Einzelne Untereinheiten können auch „fixiert“ werden. Das heißt, dass diese Untereinheiten während der Simulation nie ihre Koordinaten ändern. Dies ist auch bei einer systematischen Parametersuche möglich (siehe Abschnitt 4.7.3, Seite 111) und dient zur Beschleunigung. Als Beispiel mag ein dimeres Homooligomer gelten: Wird eine Untereinheit auf diese Weise fixiert, so fällt die Hälfte der zu berechnenden Bewegungen weg. Andererseits gilt auch, dass alle Bewegungen der anderen Untereinheit relativ zur fixierten in Bewegungen überführt werden können, die beide zueinander machen.

Außerdem kann der Nutzer der Software globale Grenzen für die Bewegungen der einzelnen Untereinheiten setzen. Sodann wird der eigentliche Algorithmus gestartet. Vor dem ersten Durchlauf

<sup>13</sup> *Monte Carlo-Algorithmen mit Simulated Annealing* (Kirkpatrick *et al.*, 1983) finden in der statistischen Physik weite Verwendung (Landau & Binder, 2000).

jedoch wird eine erste  $I(q)$  oder  $P(r)$  berechnet und ein erster  $\chi^2$ -Wert zu den experimentellen Referenzdaten errechnet. Dann startet die Hauptschleife:

1. Als nächstes wird gewürfelt, welche Untereinheiten in welcher Weise wie weit bewegt werden sollen. Ein Nutzer der Software hat die Möglichkeit, alle Untereinheiten gleichzeitig bewegen zu lassen oder nur eine zufällige Untereinheit pro Durchlauf.
2. Die Bewegung wird ausgeführt. Dies kann auch parallelisiert geschehen: Die Software erkennt automatisch die zur Verfügung stehenden Prozessoren.
3. Optional kann nun der  $R_g$  berechnet werden. Ist der Wert ausserhalb einer vorgegebenen Toleranzgrenze, so wird die gerade ausgeführte Bewegung verworfen und bei Schritt 1 wieder weitergerechnet. Der Grund für diesen Schritt ist, dass der nächste Schritt sehr rechenintensiv ist.
4. Nach Wahl des Nutzers der Software wird nun  $P(r)$  oder  $I(q)$  berechnet<sup>14</sup>, woraus der  $\chi^2$ -Wert im Vergleich mit den experimentellen Referenzdaten errechnet werden kann. Ist der  $\chi^2$ -Wert kleiner dem vorherigen Wert, so wird die gerade ausgeführte Bewegung akzeptiert und 1 zu einem internen Zähler addiert. Der neue  $\chi^2$ -Wert wird gespeichert und bei nächsten Runde als Referenz herangezogen. Ist der neue  $\chi^2$ -Wert aber schlechter als der vorherige, so wird die aktuell durchgeführte Bewegung doch mit einer Wahrscheinlichkeit von

$$\text{Zufallswert} \leq e^{\left( -\frac{\chi_{\text{vorherig}}^2 - \chi_{\text{neu}}^2}{\sqrt{\chi_{\text{vor}}^2 \text{ Kühlung}}} \right)}$$

akzeptiert, wobei der Zufallswert zwischen 0 und 1 liegt. Dies soll verhindern, dass das System in ein lokales Minimum fällt (vgl. auch Petoukhov & Svergun, 2005).

5. Hiernach fängt der Algorithmus wieder bei Schritt 1 an, bis die voreingestellte Zahl von Durchläufen erreicht wird.

Zusätzlich wird immer wieder überprüft, ob bereits eine gewisse Zahl erfolgreicher Durchläufe, also Durchläufe, bei denen ein besserer  $\chi^2$ -Wert erreicht wurde, überschritten wurde. Ist dies der Fall, so kühlt das System ab: Es werden neue, engere Grenzen für die auszuführenden Bewegungen einer Untereinheit gesetzt und zugleich wird das Inkrement für jeden Bewegungsparameter kleiner. Üblicherweise werden alle Parameter um 10% kleiner, aber Nutzer der Software haben auch die Möglichkeit einen anderen Wert zu wählen.

### 4.7.3 Ein *grid search*-Algorithmus zur verfeinerten Suche im Parameterraum

Mit dem gerade vorgestellten Algorithmus ist nicht sichergestellt die „beste“ Lösung zu finden, es ist sogar relativ unwahrscheinlich: Die Vorkehrungen, die das *simulated annealing* ausmachen (das Akzeptieren „schlechter“ Bewegungen und das systematische Kühlen) erreichen nicht zugleich, dass ein finales Optimum erreicht wird. Der *Monte Carlo*-Algorithmus mit *simulated annealing* ergibt aber für einen Nutzer der Software einen engen Rahmen für eine verfeinerte Suche im Parameterraum. Mit Abschluss der *Monte Carlo*-Funktion wird das finale Optimum für die Bewegung jeder Untereinheit des untersuchten Moleküls zurückgegeben. In der Nähe dieser Parameter kann nun schrittweise jede Kombination von Bewegungsparameter daraufhin überprüft werden, wie gut die theoretische  $I(q)$  oder  $P(r)$  mit den experimentellen Daten übereinstimmt. Am Ende dieser systematischen Suche kann eine Fehlerabschätzung für die einzelnen Bewegungsparameter erreicht werden, in dem die Abweichung der besten Lösungen mit der optimalen, finalen Lösung verglichen werden. Hartmann & Decker schlagen vor die Lösungen mit den 30 % besten  $\chi^2$ -Werten zu nehmen, sowie Mittelwerte und Standardabweichung zu errechnen (Hartmann & Decker, 2002).

<sup>14</sup>Dies ist wegen Gleichung 2.19 gleichwertig, aber der Teufel steckt im Detail: Bei kleinen Protein fällt der Lösungsmittel einfluss auf die Streukurve relativ stärker ins Gewicht als bei großen Proteinen, weil das Verhältnis von Oberfläche zu Masse größer wird.

Zur Kontrolle der Qualität der beiden hintereinandergeschalteten Algorithmen wurde die Struktur des oxygenierten 24meren Hämocyanins von *E. californicum* auf die eben vorgestellte Weise modelliert. Zur Vereinfachung der Simulation, wurde nur Trimere als unabhängig bewegte Einheiten berücksichtigt, da `sas-rigid` im Gegensatz zu `parmscan` nicht für die speziellen strukturellen Beziehungen eines 24meren Hämocyanins adaptiert ist (s. u.). Als Referenzdaten dienen die SAXS-Daten dieser Arbeit beziehungsweise von B. Pairet (Pairet, 2007). Für die oxygenierte Konformation des Proteins existiert bereits ein Modell auf der Basis von SAXS-Daten (Hartmann & Decker, 2002). Die SAXS-Daten für dieses Referenzmodell sind bereits etwas älter und standen nicht zur Verfügung (Decker *et al.*, 1996). Das Ergebnis für die mit der eigenen Software durchgeführten Rekonstruktion zeigt die folgende Abbildung, womit belegt ist, dass `sas-rigid` unter ähnlichen Voraussetzungen zu ähnlichen Ergebnissen gelangt, wie das Programm `parmscan` (Hartmann & Decker, 2002).

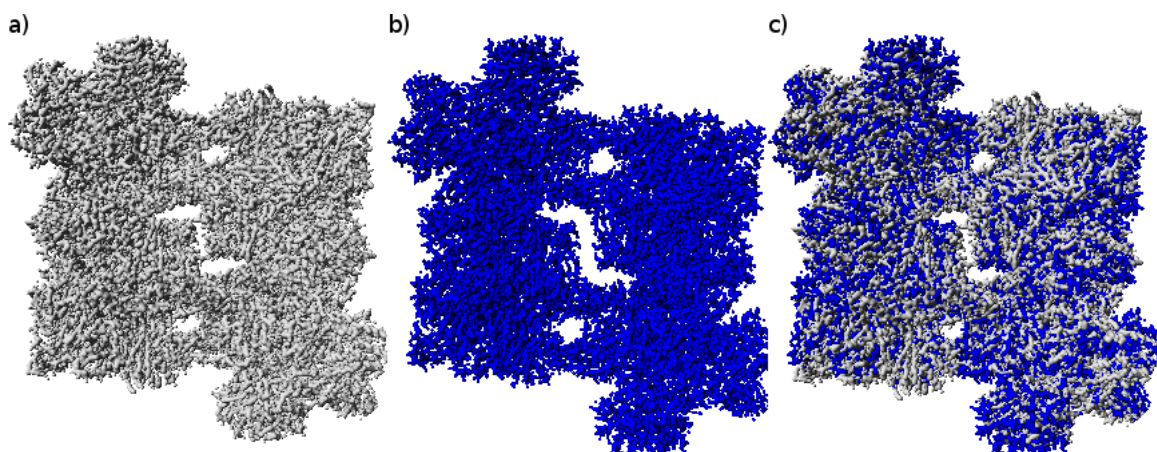


Abbildung 4.59: Vergleich von Modellierungen der oxygenierten Hämocyaninkonformation von *E. californicum*. a) Struktur nach Hartmann & Decker, 2002. Das Modell hat einen  $R_g$  von 88 Å und einen  $D_{max}$  von 270 Å. b) Struktur berechnet durch sukzessive Ausführung der Algorithmen der Abschnitte 4.7.2 und 4.7.3. Das Modell hat einen  $R_g$  von 87,3 Å und ebenfalls einen  $D_{max}$  von 270 Å. c) Superposition beider Modelle.

Bislang, wieder bezogen auf Version 0.3.3 der Software, ist es nicht möglich Bewegungsparameter bezogen auf die Symmetrie zu berechnen (wie in Hartmann & Decker, 2002), sondern nur für die einzelnen Untereinheiten bezogen auf ein absolutes Koordinatensystem. Allerdings ist bereits implementiert diese Parameter in Kugelkoordinaten umzurechnen, was sinnvoll sein kann, sofern es sich um ein Protein mit einer Rotationssymmetrie handelt.

Abbildung 4.60 zeigt die beiden theoretischen Abstandsverteilungsfunktionen der Strukturen aus Abbildung 4.59 zusammen mit dem Datensatz für das oxygenierte *E. californicum*-Hämocyanin. Da der Datensatz, welcher der Struktur aus Hartmann & Decker, 2002, zu Grunde liegt, nicht zur Verfügung stand, wurde ausschließlich der Datensatz aus Abschnitt 4.1.1.2 auf Seite 56 zum Vergleich verwendet. Die beiden Datensätze müssen allerdings sehr ähnlich sein (Pairet, 2007). Es wird deutlich, dass die in Hartmann & Decker, 2002, vorgeschlagene Struktur, die experimentellen Daten etwas besser beschreibt. Dies ist wahrscheinlich auf die auf die speziellen Anpassungen des `parmscan`-Programms auf die Symmetrie eines 24meren Hämocyanins zurückzuführen: Wie in Hartmann & Decker, 2002, beschrieben, gibt es mehr Beziehungen als die für die eigene Simulation verwendete C2-Symmetrie. Die folgenden Beziehungen fanden in der Arbeit von Hartmann & Decker Berücksichtigung, wurden in der eigenen Software aber nicht implementiert:

- Der Abstand der Trimere wird berücksichtigt und für symmetrieverwandte Trimere gleich gehalten.
- Für den Abstand der Hexamere innerhalb eines Dodekamers gilt Gleiches.

- Der Abstand der Dodekamere fließt als gesonderter Parameter ein.
- Gleiches gilt für die verschiedenen Kippwinkel für die Hexamere und Dodekamere zueinander.

Alle diese Aspekte bleiben in der eigenen Software unberücksichtigt. Das Programm **parmscan** ist also sehr viel spezieller an die Gegebenheiten von 24meren Hämocyaninen angepasst, wie auch in Hartmann & Decker, 2004, erläutert. Die eigene Software kann, wie oben beschrieben, zwar beliebige Symmetrieeigenschaften berücksichtigen, jedoch keine komplexen Abstandsbeziehungen wie bei 24meren Hämocyaninen. Dennoch stimmt das mit der eigenen Software erzielte Modell sowohl visuell (Abbildung 4.59) als auch hinsichtlich der assoziierten Abstandsverteilungsfunktion (Abbildung 4.60) gut mit der Referenzstruktur und den Daten überein. Die Abstandsverteilungsfunktionen in Abbildung 4.60 wurden mit **sas-rigid** errechnet.

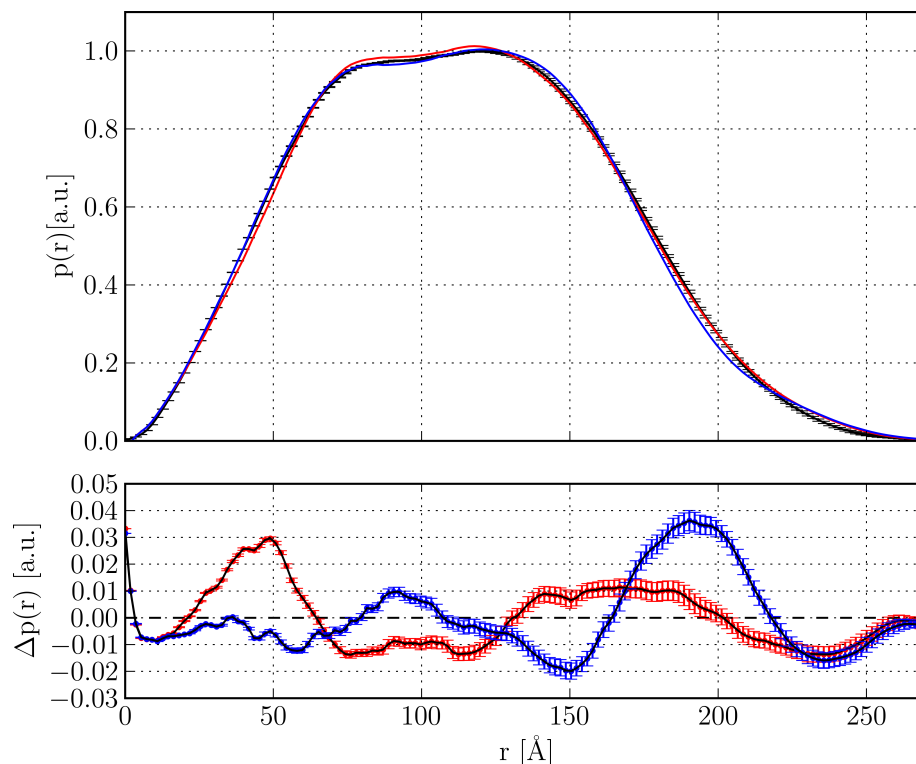


Abbildung 4.60: Vergleich der theoretischen  $P(r)$  der Strukturen aus Abbildung 4.59. Gezeigt ist die experimentelle  $P(r)$  für das oxygenierte *E. californicum*-Hämocyanin (—) mit exemplarischen Fehlerbalken. Zusätzlich dazu die theoretische Streuung der Struktur aus Hartmann & Decker, 2002 (rot) und die  $P(r)$  der selbst errechneten Struktur (blau). Die untere Tafel zeigt den Differenzplot für die experimentellen und die theoretischen Daten.

#### 4.7.4 Vergleich mit parmscan

Wie im vorigen Abschnitt eingeleitet ist das Programm **parmscan** von Dr. H. Hartmann (siehe Abschnitt 3.6.11 auf Seite 39 des Methodenkapitels) darauf optimiert eine *rigid body*-Modellierung von 24meren Hämocyaninen über einen *grid search*-Algorithmus durchzuführen. Allerdings konnte keine voll zufriedenstellende Modellbildung für die Hämocyanine von *E. californicum* und *P. imperator* im oxygenierten und deoxygenierten Zustand durchgeführt werden (siehe Daten im Abschnitt 4.1.1.2 auf Seite 56 dieses Kapitels).

Die Schwierigkeit besteht darin, die beste Übereinstimmung zwischen theoretischen und experimentellen Daten für einen Datensatz zu erlangen (Abbildung 4.61 zeigt die Diskrepanz für die  $I(q)$  einer Konformation).

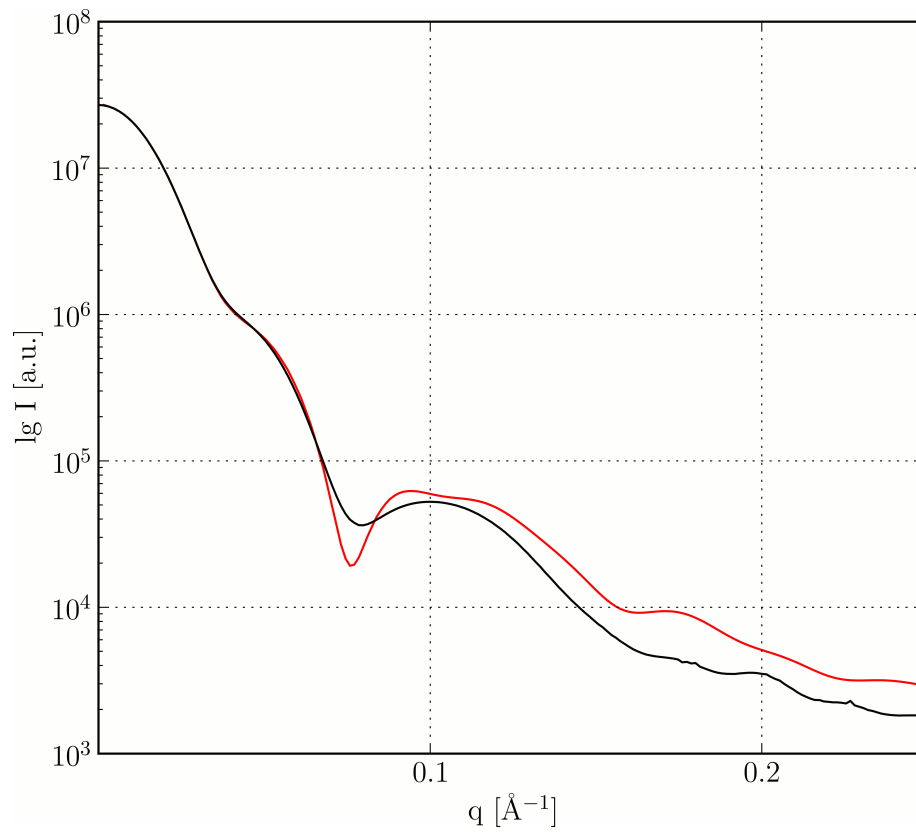


Abbildung 4.61: Vergleich der theoretischen Streukurven nach *parmscan* (rot) und experimentell (schwarz)

### 4.7.5 Ein *Monte Carlo*-Algorithmus zur Darstellung eines Wahrscheinlichkeitsraums

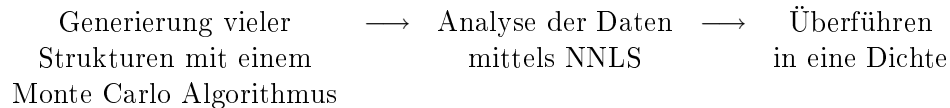
Die bisherige Implementation nimmt eine beste Lösung für eine Proteinstruktur auf Basis der gemessenen Daten an. Im Grunde ist dies aber als unwahrscheinlich zu erachten, da

- a) ein Ligand, welcher die Konformation eines Proteines ändert, immer im Gleichgewicht mit der Lösung steht und das Protein daher ebenfalls im Gleichgewicht verschiedener Konformationen stehen muss und
- b) weil die Domänen großer Proteinkomplexe, wie sie üblicherweise mittels SAXS charakterisiert werden, ohnehin auch dem Lösungsmitteldruck ausgesetzt sind und dieser fluktuiert mit der Zeit.

Folglich können „beste Lösungen“ für eine Struktur nur als eine wahrscheinliche Lösung gelten. Besser ist es, das Messsignal als gegeben durch eine lineare Kombination verschiedener Strukturen aufzufassen:

$$I(q)_{\text{experimentell}} = \sum_{i=0}^N c_i I(q)_i . \quad (4.8)$$

Hierbei sind  $c_i$  die Koeffizienten für die theoretischen Streukurven  $I(q)$  der  $i$ -ten Strukturen. Um dieses algorithmisch im Computer abzubilden, wurden in jüngerer Zeit folgende Teilalgorithmen implementiert:



„NNLS“ steht hierbei für einen nicht-negativen Algorithmus der kleinsten Quadrate (englisch: *non-negativity-constrained least squares algorithm*). Gegeben sind  $N$  Datensätze mit je  $M$  Datenpunkten (theoretische Streudaten oder Abstandsverteilungsfunktionen für jede Struktur aus dem Lauf eines Monte Carlo Algorithmus). Diese ergeben eine Matrix  $\mathbf{T}$  mit den Dimensionen  $(N \times M)$ . Zusätzlich gibt es die experimentellen Daten, die einer Matrix  $\mathbf{E}$  mit den Dimensionen  $(1 \times M)$  entsprechen<sup>15</sup>. Der NNLS-Algorithmus versucht nun den Ausdruck

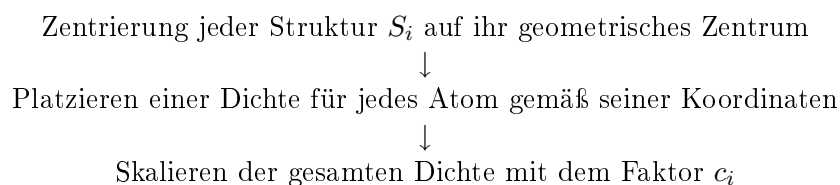
$$\|\mathbf{E} - \mathbf{T} \cdot \mathbf{c}\|^2 \quad (4.9)$$

zu minimieren, wobei gilt, dass

$$c_i \geq 0 \quad \forall i . \quad (4.10)$$

Hierbei ist  $c_i$  das  $i$ -te Element von  $\mathbf{c}$ .  $\mathbf{c}$  ist somit die Koeffizientenmatrix mit der Dimension  $(N \times 1)$ , die für die optimale Lösung der Gleichung 4.8 gesucht wird. Die Implementierung des NNLS-Algorithmus erfolgte nach der Beschreibung von Bro & de Jong (Bro & de Jong, 1997)<sup>16</sup>. Die negativen Koeffizienten werden nicht nur, wie in anderen NNLS-Algorithmen, auf null gesetzt, sondern positive Koeffizienten werden iterativ hinsichtlich Ausdruck 4.9 optimiert. Der Algorithmus hat sich als sehr stabil erwiesen: In Testläufen wurden von  $1,5 \cdot 10^4$  Datensätzen nur 2 bis 6 Koeffizienten ausgewählt. Damit wird zugleich verhindert, dass die Dichte für eine Konformation (s. u.) unrealistisch „verschmiert“ erscheint.

Die Überführung in eine Dichte erfolgt so, dass Nutzer die Voxelgröße auswählen können. Die übrigen Schritte sind wie folgt:



<sup>15</sup>Da experimentelle Daten grundsätzlich unterschiedlich viele Datenpunkte haben können, wird die Dimensionalität von  $\mathbf{E}$  über einer *Spline*-Interpolation (Press *et al.*, 2002) gewährleistet.

<sup>16</sup>Die Implementierung des Algorithmus von Bro & de Jong für nicht-kommerzielle Zwecke ist erlaubt.

Die Zentrierung jeder Struktur ist optional und sollte möglichst vor Aufruf des Monte Carlo Algorithmus erfolgen. Der Monte Carlo Algorithmus funktioniert hierbei ohne das in den vorhergehenden Abschnitten vorgestellte *simulated annealing*. Es werden lediglich mögliche Strukturen „gewürfelt“, wobei Strukturen außerhalb eines Toleranzwertes für die Abweichung von dem experimentellen Gyrationradius ausgeschlossen werden können, um Speicherplatz zu sparen. Für alle anderen Strukturen werden die Bewegungsparameter und die zugehörigen theoretischen Kurven zu  $I(q)$  oder  $P(r)$  in eine Logdatei geschrieben. Die Auswertung hinsichtlich Gleichung 4.8 erfolgt im Anschluss daran.

#### 4.7.6 Ein Algorithmus zur Berechnung der Neutronenstreulängendichte von Proteinen

Der Algorithmus kann die Neutronenstreulängendichte (nSLD) bei verschiedenen Anteilen von Deuterium im Puffer berücksichtigen: Es wird bei Proteinen angenommen, dass austauschbare Wasserstoffatome (Harrown *et al.*, 2007) im Gleichgewicht mit  $^1H$  (Protonen) und  $^2H$  (Deuterium) des Puffers stehen. Die Schätzung des Volumens erfolgte mit YASARA (Krieger *et al.*, 2002). Zusätzlich kann die `sas-rigid` die Annahme treffen, dass ein Atom in etwa den Raum von  $10 \text{ \AA}^3$  einnimmt. Dieser Durchschnittswert ist durch die Van-der-Waals-Radien der verschiedenen chemischen Gruppen gegeben (Hendricks, 1930; Huggins, 1932; Sidgwick, 1933). Bei Metallen fließt stattdessen das Volumen für eine einfache Metallbindung des entsprechenden Metalls ein (Pauling, 1960).

Die Schätzung des Volumens auf diese Weise erfolgt jedoch nur, wenn eine externe Angabe, welche YASARA und andere Programme bei vorliegender Struktur sehr gut berechnen können, fehlt. Die gerade beschriebene Schätzung ist ausreichend für den Fall, dass keine Struktur eines Proteins vorliegt und das eingenommene Volumen nicht präzise berechnet werden kann.

Im Übrigen folgt der Algorithmus den Angaben von Abschnitt 3.10.6.3, Seite 51 im Methodenkapitel.

#### 4.7.7 Webhosting der Software

Die gesamte Software – inklusive einer umfangreichen Dokumentation, einer Bedienungsanleitung und diversen Nutzungsbeispielen – ist unter <http://sas-rigid.origo.ethz.ch/> öffentlich zugänglich und für jedermann herunterzuladen. Sie steht unter der *Gnu Public License* (<http://www.gnu.org/licenses/gpl-3.0.html>), einer Lizenz, die sicherstellen soll, dass der Code der Software jederzeit frei verfügbar ist.

## 4.8 Eine graphische Benutzeroberfläche zum schnellen Plotten von Kleinwinkelstreudaten

SPlot ist ein selbstgeschriebenes Plottingprogramm für SAXS-Daten. Es basiert auf wxPython ([www.wxpython.org](http://www.wxpython.org)), Python (van Rossum, 1998), numpy & scipy (Jones *et al.*, 2007; numerische und wissenschaftliche Libraries), sowie matplotlib (Hunter, 2007).

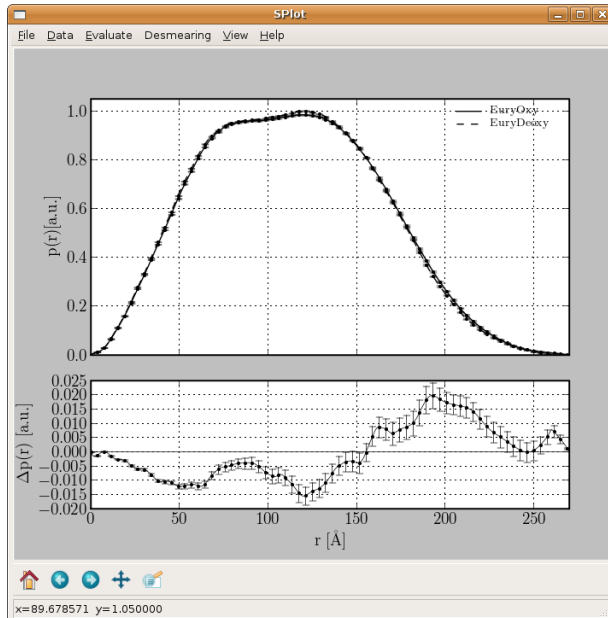


Abbildung 4.62: Das Hauptfenster von SPlot. Gezeigt wird, wie die Abstandverteilungsfunktion zweier Datensätze differiert, inklusive einem Panel für die Darstellung der Abweichung mit exemplarischen Fehlerbalken.

wesentlich kürzt.

Neben den gerade beschriebenen Operationen mit Rohdatensätzen bietet das Programm weiterhin Optionen für den Umgang mit entschmierten Datensätzen:

- Sowohl die eigentlichen Streudaten als auch die Abstandsverteilungsfunktionen zweier Datensätze können verglichen werden. Hierbei wird die Fehlerfortpflanzung automatisch berücksichtigt. Die Unterschiede zweier Datensätze können so direkt visualisiert werden.
- Es ermöglicht das Auslesen von Metadaten von entschmierten Datensätzen. Insbesondere der Abgleich verschiedener Datensätze fällt so leichter.
- Das Programm bietet die Möglichkeit Kratky- und Porodplots direkt zu plotten. Dies erlaubt Veränderungen in der Konformation oder des Volumens direkt zu visualisieren. Zugleich wird eine numerische Analyse offeriert (wo liegt das Maximum?; wie entwickeln sich die Fehler?; etc.).
- Fits theoretischer Daten können ebenfalls eingeladen werden. Der Vergleich ist graphisch vom Vergleich der experimentellen Datensätze abgehoben.
- Schließlich sind noch diverse Funktionen zur Manipulation der Darstellung eingebunden, die wenige Wünsche der Darstellung offen lassen.

SPlot, die Abbildung zeigt das nebenstehende Hauptfenster der Anwendung, bietet verschiedene Leistungen an:

- Die schnelle Darstellung von ASA-Dateien (dem Format der Software zur Kratky-Kamera), inklusive des Pulshöhenspektrums. Dies erlaubt den schnellen Abgleich von Datensätzen, noch während die Messung läuft und kann so helfen fehlerhafte Einstellungen schneller zu finden.
- Interfaces zu den Entschmierprogrammen GNOM und ITP/ITO (s.u.) und erlaubt so ein schnelles und weniger fehlerträchtiges Entschmieren verglichen mit der strikt manuellen Einstellung der Entschmierprogramme.
- Die Position des Primärstrahls bei Wanderspaltmessungen kann für eine Vielzahl von Datensätzen gleichzeitig ermittelt und ausgedruckt werden. Dies erleichtert die Dokumentation wesentlich.
- Indirekt bietet das SPlot über das Skript PreToExcel auch den direkten Export in EXCEL-Tabellen an, was die Aufarbeitung

GNOM Feeding Interface

Please Specify Parameters for GNOM. [? Help](#)

Printer  Formfactor

EXPERT  INPUT1

NSKIP1  NSKIP2

OUTPUT  ISCALE

PLOINP  PLORES

EVAERR  PLOERR

LKERN  JOBTYP

RMIN  RMAX

LZRMIN  LZRMAX

KERNEL  DEVIAT

IDET  FWHM1

AH1  LH1

AW1  LW1

ALPHA  NREAL

RAD56  NEXTJOB

Abbildung 4.63: *SPlot* Interface zur Dateneingabe in *GNOM*.

Die linksstehende Abbildung zeigt beispielsweise das Interface für *GNOM* (Svergun, 1992). Das Interface bietet eine Hilfefunktion basierend auf der Dokumentation von *GNOM*. Es ist möglich, die Parameterdatei *gnom.cfg*, welche von *GNOM* ausgelesen wird, zu speichern. Dann allerdings muss das Programm *GNOM* in der Shell von Hand gestartet werden. Dies ist nötig, da *GNOM* eine nicht lösbare Verbindung zu *Gnuplot* ([www.gnuplot.info](http://www.gnuplot.info)) enthält, welche den Programmablauf von aufrufenden Programmen wie *SPlot* stören würde. Der Vorteil, den das Interface bietet, besteht darin, dass *GNOM*-Ausgabedateien schnell in *SPlot* eingelesen werden können. Verschiedene Resultate können übereinander geplottet werden und die Eingabeparameter können online geändert werden. Die parallele Nutzung eines dritten Programms, wie ein Texteditor, entfällt.

Ein sehr ähnliches Interface besteht auch für die Programme *ITP* und *ITO* (Glatter, 1977). Hier wiegen die Vorteile noch stärker: Da die Programme *ITP* und *ITO* ursprünglich für *VAX/VMS* geschrieben wurde, sind sie durch fehlende Benutzeroberflächen (auch keine Kommandozeilenoberfläche) nicht mehr zeitgemäß. Das *SPlot*-Interface bietet jedoch direkte Hilfefunktionen an und stellt sicher, dass die sehr restriktiven Formen der Parameterdateien für *ITP* und *ITO* eingehalten werden. Letztere sind eine sehr große Fehlerquelle bei der Nutzung von *ITP* und *ITO*.

# Kapitel 5

## Diskussion

### 5.1 Die Struktur 24merer Hämocyanine von *Eurypelma californicum* und *Pandinus imperator* unter Einfluss verschiedener Effektoren

#### 5.1.1 Einfluss der Sauerstoffbindung auf die Konformation

Konformationsänderungen des Hämocyanins von *Eurypelma californicum* unter dem Einfluss von Sauerstoff wurden hinsichtlich des Konformationsüberganges bei der Oxygenierung, beziehungsweise Deoxygenierung, bereits in der Vergangenheit untersucht (Decker *et al.*, 1996; Hartmann & Decker, 2002; Hartmann & Decker, 2004; Hartmann *et al.*, 2004). Die SAXS-Daten, denen diese Arbeiten zugrunde liegen, wurden mit einer Kratky-Kamera am Institut für Physikalische Chemie in Graz generiert. Die Kratky-Kamera des Institutes für Molekulare Biophysik im Zuge der Diplomarbeit von D. Hönig installiert (Hönig, 2004). Ziel meiner Arbeiten war es zunächst, mit Hilfe des Hämocyanins von *E. californicum* als Referenz, die Qualität der Daten der Kratky-Kamera zu evaluieren. Die Kamera wurde seit der Diplomarbeit von D. Hönig nur in einem Fortgeschrittenenpraktikum genutzt und es weder klar, wie reproduzierbar die Daten sind, noch wie gut die Dekonvolution, beziehungsweise Entschmierung, der Daten gelingt.

Im Zuge der Diplomarbeit von B. Pairet (Pairet, 2007) wurden, von mir betreut, Daten des Proteins unter atmosphärischen Bedingungen (Sauerstoffeinfluss) aufgenommen. Es zeigt sich deutlich, dass die primären physikalischen Parameter zwischen den eigenen Daten und den Daten des Grazer Institutes im Wesentlichen übereinstimmen (siehe Tabelle 4.1, Seite 57). Die Differenz im Gyrationradius von ca. 0,7 Å kann dem experimentellen Fehler „angelastet“ werden. Ebensogut ist es möglich, ja sogar wahrscheinlich, dass die Unterschieden in den Datensätzen struktureller Heterogenität zuzuschreiben sind: Für sämtliche Experimente wurden Hämocyaninproben verschiedener Tiere zusammengeführt, um eine ausreichende Menge der Probe für die Messung zu haben. Damit entsteht eine Heterogenität innerhalb der Hämocyanine, die sich auch in leicht unterschiedlichen Werten für eine SAXS-Messung auswirken kann und die kleinen angesprochenen Differenzen zu erklären vermag. Überdies ist seit langem bekannt, dass das Hämocyanin von *E. californicum* auf Ebene der 24mere eine inhärente Heterogenität besitzt (persönliche Mitteilung Professor H. Decker).

Es ist festzuhalten, dass die institutseigene Kratky-Kamera interpretierbare und reproduzierbare Daten liefern kann. Es war möglich an verschiedenen Tagen mit verschiedenen Experimentatoren Ergebnisse zu erzielen, die im Rahmen des experimentellen Fehlers identisch waren. Da die physikalischen Parameter der Daten nach, beziehungsweise während, der Dekonvolution ermittelt wurden und die gemessenen Kurvenverläufe den publizierten entsprechen (vgl. Decker *et al.*, 1996), ist überdies davon auszugehen, dass die Dekonvolution korrekt durchgeführt werden kann. Die in den Tabellen 3.3 und 3.4 ab Seite 35 des Methodenkapitels aufgeführten Parameter für die Entschmierung der SAXS-Daten der institutseigenen Kratky-Kamera können also als korrekt für alle weiteren Arbeiten an der Kratky-Kamera des Institutes erachtet werden<sup>1</sup>.

<sup>1</sup>Mit der im Juli 2008 fällig gewordenen Sanierung des Röntgenlabors und der damit verbundenen Stilllegung

Zusätzlich zu den Streudaten des oxygenierten Proteins konnten zusammen mit B. Pairet auch SAXS-Daten zum deoxygenierten Protein gewonnen werden (Pairet, 2007). Auch hier stimmen Gyrationradius und Abstandsverteilungsfunktion im Großen und Ganzen mit den Referenzdaten aus Decker *et al.*, 1996, überein. Allerdings gibt es eine wesentliche Diskrepanz: Während Decker *et al.* für das deoxygenierte Protein von *E. californicum* von einem maximalen Abstand  $250 \text{ \AA} \pm 5 \text{ \AA}$  berichteten, wurde für die eigenen Daten ein Wert von  $D_{max} = 260 \text{ \AA} \pm 2 \text{ \AA}$  ermittelt. Diese Diskrepanz ist nicht mit der Heterogenität der Proben zu erklären. Während bei den Grazer Daten ein pH-Wert von 7,2 gewählt wurde (Decker *et al.*, 1996) herrschte bei den eigenen Messungen ein pH-Wert von 7,1. D. Hönig konnte zeigen, dass auch kleine Änderungen des pH-Wertes einen großen Effekt hervorrufen können (Hönig, 2004). Ob dieser allerdings alleine ausreichend ist, den Unterschied im  $D_{max}$  zu erklären, ist nicht klar. Insgesamt wird bei der Referenzarbeit von einer wesentlich größeren experimentellen Unsicherheit ausgegangen (Decker *et al.*, 1996). Es ist also nicht auszuschließen, dass der Wert für den  $D_{max}$  dort unterschätzt beziehungsweise bei den eigenen Daten leicht überschätzt wird.

Ähnlich wie die SAXS-Daten von *E. californicum*-Hämocyanin dazu verwendet wurden, mittels **parmscan** (siehe Abschnitt 3.6.11, Seite 39) Strukturvorschläge für die oxygenierte und die deoxygenierte Konformation zu machen (Hartmann & Decker, 2002), sollte dies auch für die hier gezeigten Daten durchgeführt werden. Überdies war geplant **parmscan** auch auf die SAXS-Daten für das Hämocyanin von *P. imperator* anzuwenden. Warum dies nicht funktioniert hat wird in den Abschnitten 4.1.1.2 auf Seite 56 und 4.7.4 auf Seite 113 erläutert (beide Abschnitte im Resultatekapitel): Offenbar enthielt das Programm **parmscan** Fehler.

In der Diplomarbeit von B. Pairet konnte jedoch mittels *ab initio*-Modellierung mit dem Programm **MCSAS** (Hartmann & Decker, 2004) gezeigt werden, dass sich die Konformationen von *E. californicum* und *P. imperator* in ihren Konformationen bei Oxygenierung und Deoxygenierung weitgehend gleichen (Pairet, 2007). Es ist naheliegend anzunehmen, dass sich die strukturellen Reaktionen auf Veränderungen im Sauerstoffpartialdruck bei allen Cheliceratenhämocyaninen gleichen: Vergleichbare Arbeiten zum 24meren Hämocyanin des Skorpions *Buthus indicus* stützen diese Annahme (Ali *et al.*, 2007). Die Arbeit von Ali *et al.* zeigt für das Hämocyanin eines Skorpions einen Konformationswechsel von oxygeniert zu deoxygeniert, wie er im Wesentlichen auch für das Hämocyanin der Spinne *E. californicum* beschrieben wurde (Decker *et al.*, 1996; Hartmann & Decker, 2002) – hierzu siehe auch Ali *et al.*, 1995 und Ali *et al.*, 2000.

## 5.1.2 Einfluss von SDS auf die Konformation

### 5.1.2.1 Verifikation der SDS-CMC

Die in dieser Arbeit vorgestellte Tröpfchenmethode zur Bestimmung der kritischen mizellären Konzentration (CMC) eines Detergens (siehe Abschnitt 3.9.3 auf Seite 45 des Methodenkapitels) ist nicht sonderlich exakt, stellt jedoch eine kostengünstige und schnelle Alternative zu anderen Methoden dar (eine Übersicht über diese anderen Methoden bietet Kaufmann *et al.*, 2006). Tabelle 4.2 auf Seite 58 zeigt die ermittelten Werte für die CMCs von SDS für die Puffer, die zur Aufnahme der im Folgenden besprochenen Kleinwinkelstreudaten verwendet wurden.

Die ermittelten Werte decken sich weitgehend mit anderweitig bestimmten Werten (Nillius, 2007; Jaenicke & Decker, 2008; Dietze, 2008). Es ist also festzuhalten, dass die Tröpfchenmethode zur ungefähren Bestimmung der CMC eines Detergens eingesetzt werden kann, wenn es gilt die CMC für einen veränderten Puffer schnell zu ermitteln. Die Methode scheint jedoch etwas weniger präzise, was unter Umständen aber durch größere Spritzenvolumina ausgeglichen werden kann.

Für die Aufnahme der im folgenden Abschnitt besprochenen SAXS-Daten wurde eine SDS-Konzentration von 5 mM eingestellt. Unabhängig von der Präzision der CMC-Bestimmung ist damit gesichert, dass in etwa die 4 – 5fache Menge der CMC eingesetzt wurde. Ziel war es, die

---

der Kamera ist allerdings davon auszugehen, dass diese bei Inbetriebnahme nach der Sanierung erneut kalibriert werden muss.

Konformation der Hämocyanine unter SDS-Einfluss zu charakterisieren, da SDS eingesetzt werden kann, um die Phenoloxidaseaktivität der Proteine zu aktivieren. Bei einer Charakterisierung mittels SAXS sollte die Proteinprobe möglichst in einer Konformation sein, denn dies erleichtert die Interpretation der Daten. Baird *et al.*, 2007, konnten zeigen, dass die Hämocyanine von *E. californicum* und *P. imperator* nicht ihre strukturelle Integrität verlieren und die Phenoloxidaseaktivität hoch ist. Da ein Mehrfaches der CMC des Aktivators SDS eingesetzt wurde, ist es wahrscheinlich, dass alle Hämocyaninmoleküle in der Lösung mit SDS gesättigt waren. Somit sind die aufgenommenen SAXS-Daten unter SDS-Einfluss mit einer Konformation des jeweiligen Hämocyanins zu interpretieren.

### 5.1.2.2 Einfluss von SDS auf das Hämocyanin von *Eurypelma californicum*

Für das Hämocyanin von *E. californicum* konnte gezeigt werden, dass es mit SDS zu einer Phenoloxidase zu aktivieren ist (Decker & Rimke, 1998; Decker *et al.*, 2001). Eine detaillierte kinetische Charakterisierung folgte (Jaenicke & Decker, 2008) und auch für eine ganze Reihe von Crustaceen-Hämocyaninen konnte gezeigt werden, dass diese sich mit SDS zu Catecholoxidasen konvertieren lassen (Jaenicke & Decker, 2004a). Molekulare Modellierung auf Basis von kristallographischen Daten anderer Hämocyanine lässt vermuten, dass ein konserviertes Phenylalanin der Domäne I den Zugang phenolischer Substrate zum aktiven Zentrum unterbindet (Decker & Tuczek, 2000; Decker *et al.*, 2007). Es wird spekuliert, inwieweit SDS die Struktur des Hämocyanins verändert, so dass phenolische Substrate Zugang zum aktiven Zentrum finden können. Bis zum Beginn dieser Dissertation war nicht bekannt, ob mit der Bindung von SDS ein Konformationswechsel auf Ebene des 24meres stattfindet und welcher Art dieser sein kann.

Wie in Abschnitt 4.1.1.3, Seite 57 des Resultatekapitels in Abbildung 4.3 gezeigt, trägt SDS signifikant zur Hintergrundstreuung des Puffers bei. Es ist anzunehmen, dass der größere Teil der SDS-Moleküle in Mizellen aggregiert ist. Die Frage, die sich mit diesem Resultat verbindet ist, inwieweit sich die Streuung von Puffer mit SDS von einem Streudatensatz von Hämocyanin, Puffer und SDS abziehen lässt. Inwieweit ist also der konventionelle Pufferabzug (siehe Abschnitt 3.6.4.1, Seite 34 im Methodenkapitel) durchzuführen?

Es ist plausibel anzunehmen, dass SDS in großem Überschuss in der Lösung vorhanden war: Maximal wurden 63,5 mg/mL des Hämocyanins vermessen. Dies entspricht einer etwa  $4,2 \cdot 10^{-5}$  M Lösung, während stets 5 mM SDS in der Lösung waren. Dies entspricht im ungünstigsten Fall einem Verhältnis vom Hämocyanin zu SDS von etwa 1:120. Da sich die Teildatensätze der verschiedenen Proteinkonzentrationen übereinanderlagern ließen, muss auch für die höchste Proteinkonzentration das Hämocyanin mit SDS gesättigt sein. Ein unbekannter Teil des SDS wird sich an das Protein anlagern und ein weit größerer Teil frei in Lösung sein: Teilweise als Monomer, größtenteils in Mizellen.

Damit ist der Pufferabzug prinzipiell möglich, da sich die Pufferlösung und die Probenlösung hinsichtlich des SDS gleichen werden. Es ist jedoch zu berücksichtigen, dass Daten, die mit dieser Annahme behandelt werden, einem systematischen Fehler unterliegen. Dieser Fehler ist dadurch bedingt, dass ein unbekannter Teil des SDS an das Protein angelagert und damit aus der Lösung entzogen ist. Aufgrund des Verhältnisses von Hämocyanin zu SDS sollte dieser Fehler jedoch klein sein (weit unter einer Promille, weil das ungünstigste molare Verhältnis des Hämocyanins zu SDS besser als 1:100 ist und das Massenverhältnis etwa 1:1650 beträgt).

Die SAXS-Daten des Hämocyanins von *E. californicum* zusammen mit SDS bestätigen die Annahme, dass sich SDS an das Protein lagerte: Mit zunehmendem Streuvektor  $q$  steigt die Streuintensität des mit SDS assoziierten Proteins relativ zum Protein ohne SDS (vgl. Abbildung 4.4). Da bis zu einem Wert von  $q = 0,05$  der Intensitätsunterschied gering ist, kann angenommen werden, dass die Massenzunahme ebenfalls gering ist. Eine quantitative Schätzung ist auf diese Weise nicht zu erreichen, da die Intensitätsschwankungen einer Kratky-Kamera diese mit einem zu großen Fehler versehen würde. Die Maxima des Kratky-Plots aus Abbildung 4.5 (b) in Verbindung mit den Ergebnissen aus Abschnitt 4.5, Seite 102 im Resultatekapitel, lassen aber vermuten, dass die Massenzunahmen unter 10 % liegt.

Für die Skalierung der Abstandsverteilungsfunktion (vgl. Abbildung 4.5 (a)) wurde dementsprechend der Massebeitrag durch das SDS vernachlässigt. Dies ändert nichts an den Aussagen: Es erscheint naheliegend die Konformation des Hämocyanins unter SDS-Einfluss mit der deoxygenierten Konformation zu vergleichen (vgl. Abbildung 4.2 b)), die Kurven sind sehr ähnlich. Ein Blick auf den Gyrationradius und den  $D_{max}$  läßt ebenfalls schließen, dass diese Annahme zutreffend ist (vgl. Tabelle 4.3). Die Daten von Baird *et al.* belegen aber, dass das Protein nicht deoxygeniert sein kann: Das Hämocyanin verändert unter Anlagerung von SDS nicht den Grad der Oxygenierung (Baird *et al.*, 2007), was auch durch eigene Absorptionsmessungen bestätigt ist (nicht gezeigt). Auch die Sekundärstrukturanteile, bestimmt über CD-Spektroskopie, ändern sich nur wenig. Überdies bestätigt der Kratky-Plot (Abbildung 4.5 b)), dass das Protein nicht entfaltet wird – jedenfalls nicht bis zu einer Auflösung von etwa 20 Å (die auszuwertenden Daten reichen bis etwa  $q = 0,3$  Å).

Zusammenfassend ist festzuhalten, dass das Hämocyanin von *E. californicum* unter SDS-Einfluss eine Konformation annimmt, die der Konformation unter Deoxybedingungen gleicht. Hierbei aber wird das Protein nicht deoxygeniert und es lagert sich nur wenig SDS an.

### 5.1.2.3 Einfluss von SDS auf das Hämocyanin von *Pandinus imperator*

Ziel bei den SAXS-Messungen von *P. imperator*-Hämocyanin unter SDS-Einfluss war dasselbe wie beim *E. californicum*-Hämocyanin: Die Frage nach der Konformation des 24mers. Aus den Abbildungen 4.6 und 4.7 ist zu schlussfolgern, dass der Konformationswechsel weit drastischer ist als beim *E. californicum*-Hämocyanin. Aus den Streudaten und dem Kratky-Plot wird deutlich, dass sich weit mehr SDS-Moleküle an das Protein angelagert haben. Es ist apparent, dass insbesondere die Minima der Streukurve unter SDS-Einfluss weniger stark ausgeprägt sind. Diese werden durch die relative Stellung der 6mere und 12mere des 24mers zueinander geprägt (Hartmann & Decker, 2004). Die Abstandsverteilungsfunktion in Abbildung 4.7 a) bestätigt diese Interpretation: Die  $P(r)$  verändert sich derart, dass dies nicht durch eine Verkippung der Dodekamere zueinander, wie im Fall der Konformationen des *E. californicum*-Hämocyanin unter Oxy- und Deoxybedingungen, zu erklären ist (siehe Abschnitt 4.1.1.3, Seite 57 ff. im Resultatekapitel). Dr. H. Hartmann bestätigte, dass eine solche Abstandsverteilungsfunktion durch keine Simulation, die bisher als strukturell plausibel betrachtet wurde, zu erklären ist (persönliche Mitteilung).

Festzuhalten ist an dieser Stelle aber, dass die Entschmierung der Daten unter der Annahme einer monodispersen Lösung durchgeführt werden konnte. Eine Entschmierung unter der Annahme einer polydispersen Lösung wurde ebenfalls versucht, brachte aber kein Ergebnis (die Beurteilung der Entschmierungsversuche durch das Programm GNOM waren durchweg signifikant schlechter als bei der Entschmierung unter der Annahme einer monodispersen Lösung).

Eine bisher unveröffentlichte *single particle*-Rekonstruktion von Cong *et al.* (Veröffentlichung in Vorbereitung) für das Hämocyanin von *P. imperator* unter dem Einfluss von SDS deutet darauf hin, dass es nur geringe Veränderungen in der Quartärstruktur des Proteins gibt (persönliche Mitteilung von Prof. H. Decker). Die Daten bestätigen zwar die Vermutung, dass kleine Veränderungen der Struktur den Zugang phenolischer Substrate zum aktiven Zentrum eröffnen, stehen aber in krassem Widerspruch zu den SAXS-Daten.

Daten der dynamischen Lichtstreuung bestätigen einerseits die Stabilität des Proteinkomplexes unter SDS-Einfluss im Großen und Ganzen und zeigen andererseits, dass gerade das Hämocyanin von *P. imperator* nicht so stabil ist, wie das Hämocyanin von *E. californicum* (Baird *et al.*, 2007): Der hydrodynamische Durchmesser des Moleküls nahm bei Inkubation über Nacht signifikant ab. Auch bei eigenen Versuchen elektronenmikroskopischer Aufnahmen des *P. imperator*-Hämocyanins (siehe Nillius *et al.*, 2008) wurde festgestellt, dass das Protein relativ empfindlich ist und bei Absorption auf einem Netz zur Elektronenmikroskopie (siehe Abschnitt 3.7, Seite 40 ff. im Methodenkapitel) leicht in Dodekamere oder gar Hexamere zerfällt.

Vor der Aufnahme der Daten für die *single particle*-Rekonstruktion in Cong *et al.* wurde das Protein fünf Minuten mit SDS inkubiert. In dieser Zeit entfaltet das Hämocyanin seine enzymatische Aktivität (Nillius, 2007), zerfällt aber nicht (Baird *et al.*, 2007). Die Messdauer während der

SAXS-Messung war wesentlich länger, vor allem aber muss eine Probe für die Messung in eine Probenkapillare gefüllt werden. Man kann nun spekulieren, dass die bei der Befüllung der Kapillare auftretenden Scherkräfte die, in der dynamischen Lichtstreuung beobachteten Veränderungen, beschleunigten.

Insofern ist nicht eindeutig zu schlussfolgern, welche strukturellen Veränderungen bei der Einwirkung von SDS auf das Hämocyanin von *P. imperator* vor und während der SAXS-Messung stattgefunden haben. Ausreichende Rechenkapazität vorausgesetzt ist allerdings denkbar die Situation mit dem neuen *Monte-Carlo*-Algorithmus in **sas-rigid** (siehe Abschnitt 4.7.5, Seite 115 im Resultatekapitel) zu simulieren, um zumindest mögliche Positionen der Hexamere zueinander herauszufinden.

## 5.2 Konformationsübergänge in den hexameren Hämocyaninen von *Palinurus elephas* und *Palinurus argus*

Die hexameren Hämocyanine der Langusten *P. elephas* und *P. argus* erfahren besonderes Interesse, da ein Verständnis der Konformationsübergänge innerhalb der „Basiseinheit“ der Arthropodenhämocyanine, dem Hexamer, von zentraler Bedeutung für das Verständnis auch höherer Hierarchiestufen ist. Überdies ist der experimentelle Zugang bei den relativ kleinen Komplexen auch einfacher und es gibt bereits einige kristallographisch und elektronenmikroskopisch aufgeklärte Strukturen (siehe Abschnitt 1.2.1.5, Seite 8 in der Einleitung).

### 5.2.1 Homologiemodellierung für das Hämocyanin von *Palinurus elephas*

Für die *rigid body*-Modellierung auf Basis der SAXS-Daten (siehe nächster Abschnitt) musste ein atomares Modell des hexameren Hämocyanins zur Verfügung stehen. Es gibt bereits eine ganze Reihe von kristallographisch aufgeklärten hexameren Hämocyaninstrukturen (Gaykema *et al.*, 1984; Gaykema *et al.*, 1986; Volbeda & Hol, 1989; Hazes *et al.*, 1993; Magnus *et al.*, 1994). Aufgrund der inhärent geringen Auflösung der Röntgenkleinwinkelstreuung wäre eine dieser Strukturen möglicherweise hinreichend gewesen, um als Startstruktur für eine *rigid body*-Modellierung zu dienen. Für das Hämocyanin von *P. elephas* wurde jedoch die Sequenz von Prof. T. Burmester ermittelt (Meissner *et al.*, 2003), so dass es naheliegend ist mit der eigentlichen Sequenz ein Homologiemodell (vgl. Abbildung 4.9 auf Seite 65) zu entwickeln, um die Qualität der *rigid body*-Modellierung zu erhöhen.

Es wurde, aufgrund der phylogenetischen Nähe, die Annahme getroffen, dass das Hämocyanin von *P. elephas* in gleicher Weise glykosiliert ist wie das Hämocyanin in der Struktur von *P. interruptus*. Die Struktur von *P. interruptus* zeigt für jede Untereinheit zwei Zuckerreste am Asparagin 167: 2-Acetylamino-2-deoxy- $\alpha$ -D-Glucopyranose und darauf folgend ein N-Acetyl-D-Glucosamin (Gaykema *et al.*, 1986).

Neben den Zuckerresten wurde auch versucht fehlende Schleifen in den Musterstrukturen (vgl. das Sequenzalignment in Abbildung 4.8) zu modellieren. Ziel ist es nicht etwa *die* einzig korrekte Struktur zu ermitteln, sondern eine wahrscheinliche Konfiguration für diese fehlenden Schleifen zu bestimmen, denn fehlende Strukturelemente wirken sich viel stärker auf eine Streukurve aus, als leicht falsch positionierte (Petoukhov *et al.*, 2002). Es wurde also nicht angestrebt das *Loop-modelling* zu optimieren, sondern die beste Struktur eines direkten Anlaufs mit dem Programm **MODELLER** wurde übernommen – selbstverständlich nicht ohne zumindest die beschriebene Energieminimierung zur Vermeidung sterischen Überlapps von Atomen auszuführen (siehe Abschnitt 4.1.2.1, Seite 63).

Ein qualitativer Vergleich mit dem Homologiemodell aus Meissner *et al.* ist nicht möglich, da das Homologiemodell nicht veröffentlicht wurde (Meissner *et al.*, 2003). Das eigene Homologiemodell zeigt jedoch, dass es sich qualitativ an den Musterstrukturen orientiert, für die eine schlechte Qualität bezeugt wird (siehe Abschnitt 4.1.2.3, Seite 68 im Resultatekapitel). Und so gibt die **WhatIf**-Analyse (Vriend, 1990; Hooft *et al.*, 1996) auch einige Schwächen für das Homologiemodell

an (siehe Abschnitt 4.1.2.1, Seite 63). Dennoch stellt es das bestmögliche Modell dar, das für eine *rigid body*-Modellierung zur Verfügung steht: Alle Aminosäuren – und damit alle Atome – sind im Gegensatz zu anderen Strukturen plaziert. Kleine Fehlstellungen werden bei einem Vergleich mit SAXS-Daten nicht ins Gewicht fallen, weil die Auflösung von SAXS-Daten gering ist.

### 5.2.2 Einfluss von Urat und Koffein auf die Struktur der Hämocyanine

Auf Basis der SAXS-Daten (siehe Abbildungen 4.10 und 4.11 auf Seite 66 ff.) der Hämocyanine von *P. elephas* und *P. argus* kann geschlussfolgert werden, dass die Effektoren Urat und Koffein die Konformationen der Hämocyanine in ähnlicher Weise beeinflussen. Dies ist eine keineswegs triviale Aussage, da für die beiden Effektoren bei anderen Hämocyaninen unterschiedliche Bindestöchiometrien gefunden wurden (Hellmann *et al.*, 2001; Hellmann *et al.*, 2004). Andererseits wurden in der Diplomarbeit von A. Rabenhorst für beide Effektoren bei beiden hier besprochenen Hämocyaninen eine ähnliche Anzahl von Bindestellen von 1 – 2 Bindestellen pro Hexamer nachgewiesen (Rabenhorst, 2007). Überdies wird davon ausgegangen, dass Urat und Koffein die Sauerstoffaffinität von Hämocyaninen mit ähnlicher Wirkung modulieren (Hellmann *et al.*, 2001; Menze *et al.*, 2001).

Eine genauere Analyse des Konformationswechsels wurde mittels *rigid body*-Modellierung angestrebt. Zunächst wurde versucht durch systematisches Bewegen der Untereinheiten herauszufinden, welche Konformation am besten zu den SAXS-Daten passt. Um die Rechenzeit zu verkürzen, wurde vorab mit Hilfe einer Monte Carlo-Simulation eine Näherungslösung gesucht (siehe Abschnitt 4.7.2, Seite 110 ff. im Resultatekapitel), dann folgte die systematische Lösung ausgehend von der Näherungslösung.

Die Abbildungen 4.12 und 4.13 auf den Seite 68 ff. im Resultatekapitel zeigen die Resultate dieser Modellierung in Form molekularer Oberflächen. Es wird deutlich, dass die Untereinheiten der Konformationen unter Effektoreinfluss gegenüber der Konformation ohne Effektoren leicht gedreht erscheinen – und zwar beide in ähnlicher Weise. Tabelle 4.6 mit den angewendeten finalen Bewegungsparametern bestätigt diesen Eindruck. Hierbei ist zu bemerken, dass auch die Konformation des oxygenierten Hämocyanins ohne Einfluss von Urat oder Koffein gegenüber dem Homologiemodell verdreht erscheint. Die Ursache hierfür ist unter Umständen darin zu sehen, dass das Homologiemodell, aufgrund der Musterstrukturen, eher der deoxygenierten Struktur entspricht und in seiner Konformation ebenfalls auf die SAXS-Daten angepasst werden muss<sup>2</sup>. Eine abschließende Bewertung der Plausibilität dieser Vermutung ist erst durch einen Referenzdatensatz mit dem deoxygenierten Protein zu erreichen.

Die Interpretation der relativen Unterschiede zwischen den von Urat und Koffein beeinflussten Strukturen und der effektorfreien Referenz ist durch diese Überlegung aber nicht beeinträchtigt. Wenngleich durch die Bewegungsparameter, welche die Software `sas-rigid` fand (vgl. Tabelle 4.6), suggeriert wird, dass keine Weitung oder Verengung der Trimere (als Halbmoleküle des Hexamers rund um die Rotationsachse) stattgefunden hat, belegt Tabelle 4.7 auf Seite 70 das Gegenteil. Eine Weitung der durch Urat oder Koffein beeinflussten Konformationen gegenüber der Konformation der effektorfreien Konformation ist wesentlich durch die Rotationsparameter erreicht. Dies ist in Übereinstimmung mit der Vergrößerung der Gyrationradii (siehe Tabelle 4.5 auf Seite 65 für eine Übersicht der ermittelten Gyrationradii). Zugleich wird in Tabelle 4.7 deutlich, dass die apparente Verdrehung der Untereinheiten unter Urateinfluss etwas stärker ausfällt, als unter Koffeieinfluss. Tabelle 4.6 macht deutlich, dass der Bewegungsparameter  $d\psi$  (siehe Abschnitt 4.7.1.5, Seite 109 im Resultatekapitel) hierfür wesentlich sein muss. Dieser Rotationsparameter bestimmt die Drehung um die Rotationsachse des Moleküls (siehe Abschnitt 3.8.5, Seite 43 im Methodendenkapitel).

<sup>2</sup>Die Konformation ohne Einfluss von Urat und Koffein mag gegenüber dem Homologiemodell stark verändert erscheinen, aber dieser Eindruck trügt: Durch die Verdrehung der Untereinheiten scheinen die trimeren Halbmoleküle in der Sicht entlang der Rotationsachse fast genau übereinander zu liegen. Aber dies ist ein Artefakt der Darstellung, wie ein genauer Blick auf die Seitenansicht erkennen lässt.

Und in der Tat: Ein direkter Vergleich der Abbildungen 4.12 und 4.13 zeigt, dass der zentrale „Kanal“ entlang der Rotationsachse bei der Konformation unter Urateinfluss größer ist als bei der Konformation unter Koffeineinfluss und bei beiden größer als bei der effektorfreien Konformation.

Dieser Umstand lässt trefflich über die möglichen Bindungsstellen für Urat und Koffein spekulieren. Ein sogenanntes *Docking*, also das Einpassen eines Liganden in seine Bindetasche in einem Protein im Computer, um die Bindeenthalpie und die Bindekonfiguration der Ligandenatome sowie der Atome der Bindetasche vorherzusagen, ist aufgrund der Qualität und Größe der Hämocyaninstrukturen noch niemandem möglich gewesen. Aber einige theoretische Überlegungen können Hinweise auf die Position der möglichen Bindungsstellen geben: Die Bindungsstudien von A. Rabenhorst (Rabenhorst, 2007) belegen, dass es maximal zwei mögliche Bindestellen für Urat und Koffein bei dem Hämocyanin von *P. elephas* geben muss. Das Wesentliche hierbei ist: Diese Bindungsstellen sind equivalent und damit wahrscheinlich auch strukturell gleichwertig. Die naheliegende Möglichkeit strukturell equivalente Bindungsstellen in einem Molekül mit D3-Symmetrie zu verwirklichen ist das Molekül in zwei Halbmoleküle mit C3-Symmetrie zu teilen (vgl. Abbildung 5.1).

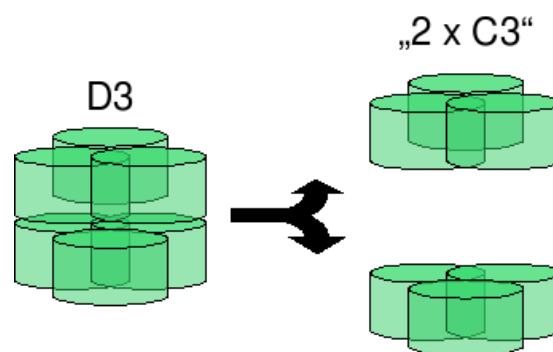


Abbildung 5.1: Schema zu den Symmetrieüberlegungen der Bindestelle für Urat und Koffein an Hämocyanin von *P. elephas* und *P. argus*. Nähere Erläuterungen im Text.

Die Effektoren müssen demnach an die einzige Stelle binden, die in beiden Halbmolekülen, den Trimeren, jeweils identisch ist: Eine Stelle im Bereich der Rotationsachse (siehe linksstehende Abbildung). Es ist damit noch keineswegs klar, wo genau die Bindestelle ist, aber allein jene theoretischen Erwägungen bringen die Lösung unter Umständen näher. Und sie erklären möglicherweise, warum die Zahl der Bindungsstellen bei steigender Temperatur des ITC-Experimentes so stark absinkt: Mit steigender Temperatur steigt prinzipiell die Flexibilität des Proteins (Tehei & Zaccari, 2007). Besetzt nun ein Effektormolekül eine Bindetasche, so kann das Protein als Reaktion darauf eine neue Konformation einnehmen. Im Falle des 6meren Hämocyanins von *P. elephas* ergäbe sich hierdurch eine veränderte Bindeenthalpie für das zweite Effektormolekül in der zweiten Bindetasche. Vor allem aber sollte sich die Dissoziationskonstante vergrößern, weil sich das Molekül „geweitet“ hat. Wenn die Symmetrie erhalten bleibt – wie in den *rigid body*-Modellierungen – und sich die trimeren Halbmoleküle bei Bindung des ersten Effektormoleküls nicht unterschiedlich verhalten, ändert sich die Bindetasche für das erste Effektormolekül in gleicher Weise: In einer ITC-Studie würde nach wie vor nur ein Typus von Bindestellen erkannt werden. Dies alles deckt sich mit den experimentellen Beobachtungen aus den ITC-Versuchen zu den Hämocyaninen von *P. elephas* und *P. argus* (Rabenhorst, 2007), die Erklärung ist aber hochspekulativ.

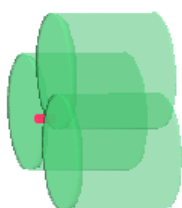


Abbildung 5.2: Schema zur Bindestelle für Urat und Koffein an Hämocyanin von *P. elephas* und *P. argus*. Rot ist ein Effektormolekül angedeutet. Nähere Erläuterungen im Text.

Es sei noch angemerkt, dass diese Überlegungen nicht im Gegensatz zu früheren Überlegungen bezüglich der Uratbindestellen am Hämocyanin von *Astacus leptodactylus* steht (Hellmann *et al.*, 2001): Hierbei handelt es sich um ein Dodekamer mit vier immunologisch unterschiedlichen Untereinheiten ( $\alpha'$ ,  $\alpha$ ,  $\beta$  und  $\gamma$  mit einer Stöchiometrie von 2:2:4:4) (Stöcker *et al.*, 1988). Hellmann *et al.* schreiben eine der Bindestellen der Untereinheit  $\beta$  zu und die andere  $\gamma$ , was zusammen den acht gefundenen Bindestellen entsprechen würde. Diese Interpretation wird aber in der selben Publikation in Zweifel gezogen. Deutlich scheint nur, dass bei diesem Hämocyanin eine Lokalisierung der Bindestellen über Überlegungen zur Symmetrie nicht möglich ist.

Wie gut ist nun die Qualität einer solchen *rigid body*-Modellierung? Die Unsicherheiten beim Homologiemodell (s. o.) einmal bei Seite gelassen, ist offensichtlich, dass eine solche Modellierung niemals eine kristallographisch bestimmte Struktur ersetzen kann. Hierzu ist die Auflösung der Kleinwinkelstreuung viel zu niedrig und der Ansatz schlicht zu grob. Es ist weiterhin unmittelbar plausibel, dass eine *rigid body*-Modellierung „falsche“ Strukturen ergeben muss: Die Kontakte der Untereinheiten und diverse energetische Beiträge werden ignoriert. Letztlich muss sogar zwingend eine gewisse Toleranz gegenüber Verletzungen sterischer Bedingungen herrschen, weil es zu Überlappungen von Untereinheiten kommen kann, um einen SAXS-Datensatz möglichst gut zu „fitten“ (siehe auch Hartmann & Decker, 2002).

Aber andererseits konnte hier gezeigt werden, dass auch mit solchen Modellen Interpretationen auf physiologischer Ebene gelingen können. Überdies wurde bereits plausibel gemacht, dass es zu einer Bewegung größerer Strukturelemente – wie z. B. Untereinheiten – kommen muss (siehe Abschnitt 4.1.2.2, Seite 65 im Resultatekapitel). Also ist es naheliegend, mit Hilfe einer *rigid body*-Modellierung zu versuchen einen solchen Konformationswechsel nachzuvollziehen. Überdies findet die Bewegung jeder Unterheit nur im Bereich weniger Ångström in jede Richtung statt. Dies ist in Übereinstimmung mit den primären Daten, zeigt aber auch, dass die Modellierung nicht aus dem Ruder gelaufen ist und keine Übereinstimmung mit den SAXS-Daten durch unsinnige Translationen von Untereinheiten zu erreichen versucht wurde<sup>3</sup>.

Man könnte nun argumentieren, dass nicht notwendigerweise ganze Untereinheiten bewegt werden müssen, sondern dass es näherliegender ist, nur Domänen des Hämocyanins gegeneinander zu bewegen, um die SAXS-Daten optimal wiederzugeben. Generell gilt hier der Einwand, dass eine Erhöhung von Freiheitsgraden bei einer Simulation zwar die Qualität von Datenfits erhöht – aber zugleich die Aussagen, die sich schlussfolgern lassen, in die Beliebigkeit abgleiten. Andererseits gilt auch, dass die Verletzungen der sterischen und energetischen Rahmenbedingungen mit der „Feinkörnigkeit“ der bewegten Einheiten abnimmt. Im Fall des Hämocyanins von *P. elephas* bot sich an, die Domäne I unabhängig von den beiden anderen Domänen zu bewegen, weil bei den Kristallstrukturen eine solche apparente Bewegung festgestellt wurde (Volbeda & Hol, 1989; Magnus *et al.*, 1994) und diese Domäne ohnehin im „Verdacht steht“ eine besondere Flexibilität in Bezug auf die Domäne II mit dem aktiven Zentrum aufzuweisen (Decker & Jaenicke, 2004; Jaenicke & Decker, 2004b; Decker *et al.*, 2007; Decker *et al.*, 2007; Jaenicke & Decker, 2008).

Eine solche *rigid body*-Modellierung mit einer unabhängigen ersten Domäne wurde versucht, allerdings konnte aufgrund mangelnder Rechenkapazität kein abschließender Konformationsvorschlag erhalten werden.

Zusätzlich zu den bisher vorgestellten Modellen wurde erwogen, dass ein Protein von der Größe eines Hämocyanins mehrere Konformationen im oxygenierten Zustand aufweisen kann (Bernhard *et al.*, 1983; Robert *et al.*, 1987; Decker *et al.*, 1989; Decker, 1990; Leidescher & Decker, 1990; Menze *et al.*, 2000; Hartmann *et al.*, 2001; Hellmann *et al.*, 2001; Menze *et al.*, 2001; Hellmann *et al.*, 2003; Spinozzi *et al.*, 2003; Colangelo *et al.*, 2004; Hellmann *et al.*, 2004 – stellen einen kleinen Auszug von Veröffentlichungen dar, die argumentativ konformationelle Heterogenität explizit oder implizit für Hämocyanine im oxygenierten Zustand fordern oder für deren Aussagen dies zu fordern ist). Es ist naheliegend eine Modellierung durchzuführen, die diesem Umstand Rechnung trägt. In der Gruppe um D. Svergun (EMBL, DESY-Outstation in Hamburg) gab es im vergangenen Jahr erstmals Ansätze konformationelle Heterogenität von Proteinstrukturen mit SAXS zu messen und softwaremäßig nachzuvollziehen (Bernadó *et al.*, 2007; Bernadó *et al.*, 2008). Diese Arbeiten nutzen einen genetischen Algorithmus um ein Ensemble von Konformationen zu generieren, dessen theoretische Streuung zusammen einen SAXS-Datensatz optimal widerspiegeln soll. Abgesehen von der methodischen Unterschiedlichkeit (zum Algorithmus, der die Dichten generiert siehe Abschnitt 4.7.5, Seite 115), sucht dieser Algorithmus aber nach unterschiedlichen Konformationen in intrinsisch besonders flexiblen Proteinen; Proteinen also, deren

<sup>3</sup>Denkbar wäre eine Translation der Untereinheiten, so dass alle Untereinheiten in der Nähe der Ebene orientiert sind, die durch die Grenze der beiden Halbmoleküle beschrieben ist. Testläufe während der Programmierphase zeigten, dass hier ein lokales Minimum für die Modellierung liegen kann, das nicht die Symmetriebedingungen verletzt. Dies ist kein ausgeprägtes Minimum und eine äußerst schlechte Lösung für das Modellierungsproblem, weil strukturell undenkbar. Aber es ist ein Beispiel für ein lokales Minimum, das es zu vermeiden gilt.

räumliche Diskrepanz von Konformation zu Konformation besonders groß ist. Im Folgenden jedoch wird davon ausgegangen, dass aus dem SAXS-Datensatz eines relativ stabilen Moleküls, eines 6meren Hämocyanins, strukturelle Heterogenität abzuleiten ist. Dies ist auch plausibel, weil

- a) ein Ligand, welcher die Konformation eines Proteins ändert, immer im Gleichgewicht mit der Lösung steht und das Protein daher ebenfalls im Gleichgewicht verschiedener Konformationen stehen muss und
- b) weil die Domänen großer Proteinkomplexe, wie sie üblicherweise mittels SAXS charakterisiert werden, ohnehin auch dem Lösungsmitteldruck ausgesetzt sind – und dieser fluktuiert mit der Zeit

(siehe auch Abschnitt 4.7.5, Seite 115).

Der in Abschnitt 4.7.5, Seite 115 des Resultatekapitels vorgestellte Algorithmus versucht geringe Flexibilität in Form einer Dichte von Aufenthaltsmöglichkeiten wiederzugeben. Das Ergebnis ist in den Abbildungen 4.16 und 4.17 auf Seite 73 des Resultatekapitels gezeigt. Es wird deutlich wie sehr sich die Konformationen mit den zuvor vorgestellten Konformationen in der Wiedergabe durch die molekularen Oberflächen ähneln, wenn sich die Sicht auf die äußeren Bereiche konzentriert (vgl. Abbildungen 4.12 und 4.13). Andererseits scheint der innere „Kanal“ entlang der Rotationsachse verschlossen.

Im Wesentlichen bestätigt Abbildung 4.17 aber das Bild der Abbildungen 4.12 und 4.13: Die Konformationen entsprechen einander weitgehend unabhängig von den Details des Modellierungsalgorithmus. Der Vorteil einer Dichterepresentation wird in Abbildung 5.3 deutlich: Mit ihr ist es möglich eine Wahrscheinlichkeit für die verschiedenen „Unterkonformationen“ abzuschätzen, denn je größer ein Dichtewert ist, desto sicherer ist anzunehmen, dass an dieser Position tatsächlich ein Teil des Moleküls im zeitlichen Mittel anzutreffen ist.

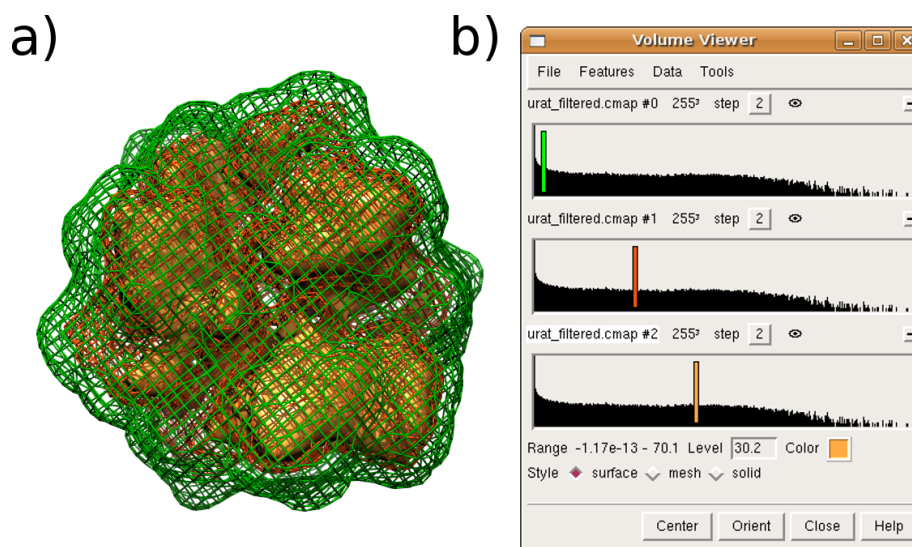


Abbildung 5.3: Dichte für das Hämocyanin von *P. elephas* unter Urateinfluss aus Abbildung 4.16 bei verschiedenen Skalierungen. a) Die Dichte des Proteins dargestellt in verschiedenen Farben und bei verschiedenen Schwellenwerten. b) Ein Bildschirmfoto das den Verlauf der Dichtewerte und die verschiedenen Einstellungen zeigt. Die Farben der Dichte in a) entsprechen den Farben der Schwellenwert-Regler in b).

### Ausblick

Genügend Rechenkraft vorausgesetzt ist es sehr lohnenswert den Algorithmus zur Generierung der Dichten (siehe Abschnitt 4.7.5, Seite 115) einzusetzen, um detailliertere Dichten zu erhalten. Insbesondere mit diesem Algorithmus ist die gesonderte Bewegung der Domäne I gegen die anderen beiden Domänen der Untereinheit sinnvoll: Da zwar ebenfalls eine Erhöhung der Freiheitsgrade durch die Verdoppelung der bewegten Körper geschähe, aber durch die Abbildung in Form einer Dichte, kann eine sinnvolle Abschätzung der tatsächlichen relativen Aufenthaltswahrscheinlichkeit der Domänen zueinander gelingen.

Darüber hinaus kann eine detaillierte Dichte dazu verwendet werden die Konformationsänderung mit Modellierungstechniken wieder auf die eigentliche atomare Struktur zu beziehen (Wriggers & Chacón, 2001). Hierbei wird die Dichte mit einem Laplace-Operator konvolviert, so dass die *de facto*-Auflösung zum Docking in eine Dichte bis 25 Ångström betragen kann. Dies ist eine Auflösung, die auch in den SAXS-Daten „stecken kann“.

### 5.3 Strukturelle Charakterisierung des Hämocyanins von *Porcellio scaber*

Die elektronenmikroskopischen Daten des Hämocyanins der Kellerassel *P. scaber* sollten zur Klärung der Frage nach der strukturellen Organisation des Hämocyanins beitragen. Während die ursprüngliche Forschungsfrage in Richtung des Immunsystems der Asseln und der Phenoloxida-seaktivität des Hämocyanins zielte, fiel während der Diplomarbeit von S. May an der Universität Kiel (AG Zimmer, „Evolutionsökologie von Asseln“) auf, dass das Hämocyanin der Kellerasseln

- a) offenbar Phenoloxidaseaktivität aufweist und
- b) bei der Reinigung über die Gelfiltration der MALLS-Anlage (siehe Abschnitt 3.4.1, Seite 31) offenbar zwei Peaks aufweist.

Die elektronenmikroskopischen Aufnahmen konnten die Daten der analytischen Ultrazentrifugation (von Professor E. Jaenicke; diese Daten sind nicht Teil dieser Arbeit) bestätigen: Ein Gelfiltrationspeak enthält nur hexameres Hämocyanin, der zweite ein Gemisch aus hexameren und dodekameren Hämocyanin. Hexameres und dodekameres Hämocyanin stehen wahrscheinlich nicht miteinander in Gleichgewicht, da auch bei langer Lagerung des separierten Hexamers kein Dodekamer in der Probe nachgewiesen werden konnte.

Das hexamere Hämocyanin weist D<sub>3</sub>-Symmetrie auf. Wie Abbildung 4.26 auf Seite 78 deutlich macht ist es ähnlich anderen hexameren Hämocyaninen. Das dodekamere Hämocyanin hingegen zeigt eindeutig (siehe Abbildung 4.26, unten) eine Besonderheit: Die Teil-Hexamere können durch einfache Translation ineinander überführt werden. Andere dodekamere Hämocyanine zeigen Stellungen bei denen die Rotationsachsen der Hexamere senkrecht zueinander stehen oder die Hexamere, durch Disulfidbrücken vernetzt, lose aufeinander sitzen (für eine Übersicht siehe van Holde & Miller, 1982). Dieser Befund stellt ein Novum dar und ist auch physiologisch von besonderem Interesse, da die Assel wahrscheinlich zwei verschiedene Untereinheiten exprimiert (persönliche Mitteilung Professor E. Jaenicke). Deren Bedeutung ist allerdings noch nicht geklärt und bedarf weiterer Untersuchungen.

Die Ergebnisse sind bereits zur Veröffentlichung eingereicht (Jaenicke *et al.*, 2008).

## 5.4 Hitzedenaturierung mit der Kratky-Kamera

Proteinfaltung wird häufig mittels Kleinwinkelstreuung untersucht (siehe z. B. die rezenten Arbeiten von Putnam *et al.*, 2007; Bernadó *et al.*, 2007; Lipfert & Doniach, 2007). Insbesondere kleine Proteine und Peptide stehen im Mittelpunkt dieser Aktivitäten, weil diese rechnerisch am einfachsten beschrieben werden können (Berger *et al.*, 1977; Dainese *et al.*, 2000; Zhu *et al.*, 2003; Zhu *et al.*, 2004; Hsu *et al.*, 2007). Dissoziationsstudien wurden aber bereits auch an großen Proteinkomplexen, wie beispielsweise Hämocyaninen und Ceruloplasmin durchgeführt (Vachette *et al.*, 2002; Spinozzi *et al.*, 2003). Hitzedenaturierungsstudien sind ebenfalls ein Ansatz, um über die Entfaltung auf die Faltung von Proteinen bzw. die Assoziation von Proteinkomplexen schließen zu können. Auch diese Aspekte werden mit Hilfe der Kleinwinkelstreuung untersucht (Kozak, 2006). Und schließlich kann mit Hitzebehandlung durchaus auch Einblick in physiologische Prozesse gewonnen werden (Skouri-Panet *et al.*, 2006).

In Hinblick auf derartige Studien wurde an der Kratky-Kamera des Institutes ein temperierbarer Probenhalter installiert. In einem Praktikum wurde die Installation vorgenommen und erste Denaturierungsstudien am Beispiel von Lysozym durchgeführt (Hilmenyuk & Schwiering, 2006). Wie im entsprechenden Abschnitt (4.6, Seite 104) gezeigt, kann die Hitzedenaturierung eines Proteins mit Hilfe des temperierbaren Probenhalters nachvollzogen werden, vorausgesetzt es werden zwei Datensätze mit dem nativen und dem hitzedenaturierten Protein aufgenommen. Die temperierbare Zelle wurde also erfolgreich und gewinnbringend installiert – auch in Hinblick auf die Kühlung von Proben, wie es für die Messungen am  $\alpha$ -Hämolyisin (siehe Abschnitt 3.10.5, Seite 50 im Methodenkapitel) notwendig war.

Interessant ist, dass das Protein sich während der Hitzedenaturierung kaum veränderte, was im Gegensatz zu älteren Studien steht (Nicoli & Benedek, 1976; diese Arbeit verfolgt die Denaturierung mittels statischer Lichtstreuung). Eine Arbeit, die sich detailliert mit der Hitzedenaturierung von Hühnereilysozym, verfolgt mit SAXS, beschäftigt, kommt jedoch zu ganz ähnlichen ganz ähnlichen Ergebnissen, wie sie auch mit der Kratky-Kamera des Institutes für Molekulare Biophysik gefunden wurde (Hirai *et al.*, 2000).

## 5.5 Massenbestimmung mit dem Kratky-Plot

Nach der Debye-Gleichung (Gleichung 2.13 im Theoriekapitel; Abschnitt 2.2.2) ist

$$I(0) \propto \sum_{x,y} f_x \cdot f_y \quad (5.1)$$

( $\sum_{x,y} f_x \cdot f_y$  ist die Summe aller Formfaktorenprodukte aller Atome eines Moleküls) und damit auch

$$I(0) \propto M_{wt} . \quad (5.2)$$

Sofern es also möglich ist die Streuintensität bei  $I(0)$  zu bestimmen, ist auch das Molekulargewicht von Partikeln zu berechnen. Wie in der theoretischen Einleitung (Abschnitt 2.2.3, Seite 21) ausgeführt ist dies jedoch bei Röntgenstrahlung äußerst schwierig, da sich der Flux nur ungenau ermitteln lässt. Überdies müsste in einem naiven Verfahren gegen  $I(0)$  extrapoliert werden, da es wegen der Primärstrahlblocker nicht möglich ist  $I(0)$  zu messen. Dies ist möglich, aber nicht präzise, da zusätzlich auch Kenntnis der Proteinkonzentration notwendig ist und ggf. eine zusätzliche Korrektur der Instrumenteneigenschaften durchgeführt werden muss (Guinier, 1939).

In einem Synchrotron ist es, wegen des relativ konstanten und zugleich hohen Fluxes, üblich eine Skalierung mit externen Standards durchzuführen. Eine Normalisierung, z. B. mit GNOM (Svergun, 1992), ermöglicht dann eine sehr genaue Bestimmung von  $I(0)$ .

Für Kratky-Systeme ist dieses nur schwer möglich: Die geringe Intensität verlangt lange Strahlzeiten auf externe Standards und in dieser Zeit kann die Intensität leicht schwanken, was zusammengekommen eine äußerst dürftige Kalibrierung darstellen würde. Dennoch ist gezeigt worden,

dass eine sogenannte „absolute Skalierung“ mit Wasser als externem Standard möglich ist (Ort-haber *et al.*, 2000). Der Aufwand hierzu ist jedoch groß, so dass dieses Verfahren nicht tauglich für den Laboralltag ist.

Die Arbeiten am japanischen „*Photon Factory*“-Synchrotron von Semisotnov und Mitarbeitern zeigen jedoch eine Alternative auch für Kratky-Systeme auf: Die Bestimmung des Molekulargewichtes mit Hilfe des Kratky-Plots (Semisotnov *et al.*, 2002). Die Arbeit am Synchrotron erlaubt es mit relativ kleinen Proteinmengen und kurzen Strahlzeiten umzugehen. Die Fragestellung für die Kratky-Kamera war also: Kann man mit den Bedingungen denen die Arbeit an der Kamera unterliegt (Geometrie der Kamera, Strahlzeiten, Proteinbedarf) ein ähnliches Ergebnis erzielen? Und: Gelten die Rahmenbedingungen auch für große Proteinkomplexe?

Festzuhalten ist, dass der postulierte lineare Zusammenhang zwischen dem Logarithmus des Molekulargewichtes und dem Logarithmus des maximalen  $q$ -Wertes im Kratky-Plot auch bei den Messungen mit der Kratky-Kamera bestätigt wurde. Der lineare Korrelationskoeffizient ist sogar besser als bei Semisotnov *et al.* (0,997 gegenüber 0,988). Doch während bei dem Synchrotronexperiment eine Steigung von 3,0 gefunden wurde, liegt diese mit den eigenen Werten bei 2,5. Der Wert von 3,0 entspricht der Steigung, wie sie für die Streuung idealer Kugeln uniformer Dichte zu erwarten ist.

Relativ kleine, globuläre Proteine stehen einer Kugelform relativ nahe. Der untere Teil der Eichgerade (Abbildung 4.53) bestätigt dies; eine Gerade durch diese Punkte würde eine entsprechend höhere Steigung aufweisen. Doch Proteine entsprechen nicht idealen geometrischen Körpern. Deformationen können entweder einen prolaten oder oblaten Körper erzeugen. Diese Deformation wirkt sich umso stärker aus je größer ein Protein ist. Aber auch kleine Proteine, wie das sehr prolate  $\alpha$ -Hämolyisin (zur Struktur siehe Abschnitt 4.4.2, Seite 84 ff. im Resultatekapitel) können schon relativ weit von der Eichgerade abweichen (siehe Abbildung 4.53 auf Seite 102).

Starke Deformationen in Vergleich zu idealen geometrischen Körpern der zur Eichung verwendeten Proteine können also die Eichgerade signifikant beeinflussen. Dies erklärt die relativ großen Fehler bei der Bestimmung des Molekulargewichtes (siehe Tabelle 4.5 auf Seite 103).

Zusammenfassend ist also festzuhalten, dass die Molekulargewichtsbestimmung mit dem Kratky-Plot zwar mit einer Kratky-Kamera möglich, aber insbesondere bei großen Proteinen fehleranfällig ist. Der Vorteil der Methode liegt in der einfachen Anwendbarkeit: Ist einmal eine Eichgerade aufgezeichnet, so kann eine Probe während der Messung – quasi instantan – hinsichtlich ihres Aggregationszustandes beurteilt werden. Dies ermöglicht eine einfache, aber wirkungsvolle Kontrolle während eines Experimentes. Es kann unmittelbar beurteilt werden, ob ein Proteinkomplex beispielsweise aggregiert oder dissoziiert vorliegt.

### *Ausblick*

Eine kleinere Veröffentlichung im Sinne einer Laborhandreichung ist zwar möglich und wünschenswert, aber mit dem Aufkommen sehr viel intensiverer Laborröntgenquellen für die Kleinwinkelstreuung (SAXSess der Anton Paar GmbH, [www.anton-paar.com/x-ray-structure-analysis/saxs\\_CXSN-5RHEZG.en.0.jsp](http://www.anton-paar.com/x-ray-structure-analysis/saxs_CXSN-5RHEZG.en.0.jsp) und Nanostar der Firma Bruker AXS [www.bruker-axs.de/nanostar.html](http://www.bruker-axs.de/nanostar.html)) ist diese Arbeit eigentlich obsolet. Die neuen Röntgenquellen erlauben sehr viel kürzere Messzeiten und bieten – laut Herstellerangabe – auch einen viel konstanteren Flux. Daher ist die Kalibrierung der Intensität mit einem externen Standard einfach und im Laboralltag durchzuführen. Die Molekulargewichtsbestimmung auf diese Weise ist auch sehr viel präziser, da sie nicht von Faktoren wie der (im Zweifel unbekannt) Struktur eines Proteins abhängig ist.

## **5.6 Die Struktur des Bla-g Hexamerins von *Blattella germanica***

Das Hexamerin Bla-g der Schabe *B. germanica* wurde im Rahmen der Diplomarbeit von A. Beusch gereinigt und charakterisiert (Beusch, 2006). Vor Beginn dieser Diplomarbeit gab es bereits zwei

Diplomarbeiten zum Hexamerin Per a 3 aus *Periplaneta americana* (Cirak, 2000; Lafargue, 2003). Aufgrund der phylogenetischen Nähe der zwei Schabenarten bestand die Vermutung, dass auch die deutsche Schabe *B. germanica* ein vergleichbares Hexamerin exprimieren sollte. Zudem stehen beide Hexamerine, von *B. germanica* wie von *P. americana* im Verdacht allergenes Potential zu besitzen (Lafargue, 2008).

Gleichgewichtsläufe in der Ultrazentrifuge indizierten ein Molekulargewicht von  $369,7 \text{ kDa} \pm 19 \text{ kDa}$ . Dies ist ein für Hexamerine geringer Wert (vgl. Hübler, 2001; Lafargue, 2003). Der Sedimentationskoeffizient von  $17 \text{ S} \pm 0,2 \text{ S}$  ist ebenfalls relativ gering. Dennoch zeigen die elektronenmikroskopischen Aufnahmen deutlich eine hexamere Struktur mit der für Hexamerine typischen Symmetrie D<sub>3</sub> (siehe Abschnitt 4.3, Seite 81). Es ist daher zu vermuten, dass der Bla-g-Komplex relativ kompakt ist und keine posttranslationalen Modifikationen wie Glykosilierungen aufweist<sup>4</sup>.

### Ausblick

Mit Hilfe der elektronenmikroskopischen Daten konnten bestehende Vermutungen hinsichtlich der Struktur von Bla-g bestätigt werden. Eine vollständige *single particle*-Rekonstruktion zur Ermittlung der 3D-Struktur ist möglich – der zusätzliche Erkenntnisgewinn in Bezug auf das allergene Potential des Hexamerins ist aber wahrscheinlich gering. Eine Publikation der Proteincharakterisierung ist in Vorbereitung.

## 5.7 Die strukturelle Charakterisierung des Hämocyanins von *Upogebia pusilla*

Für das Hämocyanin von *Upogebia pusilla* wurde die Quartärstruktur eines 4 x 6meren Hämocyanins postuliert, dass ähnlich wie das Hämocyanin von *Callinassa californiensis* in Form von zwei 2 x 6meren organisiert ist, die quasi-tetraedisch assemblieren (Cavellec *et al.*, 1990). In einer Publikation unserer Gruppe wurde diese Annahme als bestätigt aufgefasst (Paoli *et al.*, 2007).

Die eigenen SAXS-Daten widersprechen der formulierten Annahme nicht (siehe Abschnitt 4.1.3.1, Seite 74). Und auch das aus den SAXS-Daten abgeleitete Molekulargewicht von ca. 1,4 MDa weist in diese Richtung (siehe Abschnitt 4.5, Seite 102). Dennoch ist das so bestimmte Molekulargewicht etwas niedrig und vor allem die elektronenmikroskopischen Daten (vgl. Abschnitt 4.1.3.2, Seite 75) lassen ein Problem vermuten: Die elektronenmikroskopischen Daten belegen klar, dass zumindest ein Teil der Probe in Form eines 2 x 6mers vorlag<sup>5</sup>. Auch eine *ab initio*-Rekonstruktion auf Basis der SAXS-Daten (siehe Abbildung 4.19 auf Seite 75) zeigt keine Dichte, die vier Hexameren entsprechen könnten. Die *ab initio*-Rekonstruktion war frei von Annahmen zur Symmetrie.

Ich habe in Folge der Auswertung der Daten auf eventuelle Probleme mit der Probe hingewiesen<sup>6</sup>. Es war denkbar, dass der apparente Widerspruch der Daten zu der Annahme über die Symmetrie des Komplexes auf Probleme bei der Probenbereitung zurückzuführen waren. Andernfalls muss davon ausgegangen werden, dass die Annahme zur Symmetrie des Komplexes falsch ist. Ich hielt es für wahrscheinlicher, dass es Probleme mit der Probenbereitung oder Lagerung gab, weil auf die Zugabe divalenter Kationen zur Stabilisation verzichtet wurde (persönliche Mitteilung M. Paoli). Dies kann auch einen Einfluss auf weitere Daten (beispielsweise aus der Ultrazentrifugation und Sauerstoffbindeexperimenten) gehabt haben. So sind beispielsweise auch mehrere Experimente mit der statischen Lichtstreuung ausgeführt worden, deren Ergebnisse hinsichtlich der Masse

<sup>4</sup>Letzteres ist bislang experimentell nicht nachgewiesen und einzig aus den zitierten Daten geschlussfolgert.

<sup>5</sup>Ich wurde im Laufe meiner Arbeit häufiger mit der Annahme konfrontiert, dass Proteinkomplexe während der Adsorption an den Kohlefilm des Netzchens für die Elektronenmikroskopie oder während des Prozesses der Kontrastierung zerfallen und dass aufgrund dessen Teilstrukturen eines solchen Komplexes abgebildet werden. Aus der Literatur ist mir kein solcher Fall bekannt. Wenn ein solcher Zerfall während des Einwirkens des Kontrastmittels auftritt, so ist jedoch zwingend davon auszugehen, dass die Teile des Komplexes überwiegend in räumlicher Nähe auf dem Netzchen absorbiert bleiben (persönliche Mitteilung Dr. S. Müller, Basel). Es wurde keine große Zahl von 2 x 6meren gefunden.

<sup>6</sup>Persönliche Mitteilung an M. Paoli, Dr. N. Hellman und Professor H. Decker.

des Proteins große Schwankungen aufweisen (persönliche Mitteilung Professor E. Jaenicke). Die Publikation von Paoli *et al.* berücksichtigt diese Einwände nicht (Paoli *et al.*, 2007).

## 5.8 Der Vergleich von elektronenmikroskopischen Daten und Kleinwinkelstreudaten bei dem Hämocyanin von *Nautilus pompilius*

Zu Beginn dieser Dissertation war die Struktur des Hämocyanins von *N. pompilius* noch unbekannt. Durch elektronenmikroskopische Bilder im Negativkontrast konnte aber bereits vor Jahrzehnten erkannt werden, dass sich dieses Hämocyanin von den Hämocyaninen anderer Mollusken unterscheidet (Bonaventura *et al.*, 1981). Es stellt ein Dekamer dar, zusammengesetzt aus zwei Pentameren, deren Kragendomäne auf halber Höhe des flachen Zylinders in das Innere des Zylinders ragt. Inzwischen ist überdies bekannt, dass das Tier nur ein Gen für sein Hämocyanin hat (Bergmann *et al.*, 2006), wodurch naheliegend ist, dass die Symmetrie des Hämocyanins D<sub>5</sub> ist<sup>7</sup>.

Ziel der Projektes am *N. pompilius*-Hämocyanin war es ursprünglich mittels SAXS-Messungen an oxygeniertem und deoxygeniertem Protein die Konformationsänderungen des Hämocyanins im Zuge der Sauerstoffbindung nachzuvollziehen. Eine ähnliche Arbeit existiert bereits von unserer Gruppe zum Hämocyanin der Schlüssellochschnecke *Megathura crenulata* (Hartmann *et al.*, 2001; Hartmann *et al.*, 2004). Hier konnte mittels *ab initio*-Modellierung eindrucksvoll belegt werden, welche strukturellen Auswirkungen der Konformationswechsel bei Sauerstoffbeladung beziehungsweise -entladung auf das Hämocyanin hat. Letztlich konnte aber nur, wegen Materialmangels, ein SAXS-Datensatz für das oxygenierte Protein gewonnen werden. Die erhaltenen Daten konnten aber zur Validierung der *single particle*-Struktur von Professor U. Meissner (Institut für Zoologie der Universität Mainz) herangezogen werden.

Wie Abbildung 4.27 deutlich zeigt, gibt es eine Diskrepanz in der Skalierung zwischen den elektronenmikroskopischen Daten für das Hämocyanin von *N. pompilius* und den SAXS-Daten: Die Position der Minima ist gegeneinander verschoben. Welcher Datensatz weist nun die bessere beziehungsweise korrekte Skalierung auf? Die Kratky-Kamera wurde gewissenhaft kalibriert (siehe Abschnitt 3.6.3 im Kapitel Material & Methoden), während die elektronenmikroskopischen Daten in ihrer Skalierung erst nachträglich durch Vergleich mit atomaren Modellen angeglichen wurden (persönliche Mitteilungen von Professor Meissner, Dr. Depoix und Professor Markl). Hierbei wurden ebenfalls Unterschiede von ca. 10 % festgestellt. Dies stimmt gut mit den geschätzten 12 % Differenz (siehe Abschnitt 4.2) überein. Der nominelle Vergrößerungsfaktor für die elektronenmikroskopischen Bilder ist also falsch (gewesen)<sup>8</sup>.

Wesentliche Merkmale der experimentellen und der von den *single particle*-Daten abgeleiteten Streukurven sind ähnlich: Der Kurvenverläufe entsprechen einander in ihrer Form. Die tiefere Lage der errechneten Minima gegenüber den experimentell bestimmten Daten kann dadurch erklärt werden, dass eine *single particle*-Rekonstruktion durch die ihr auferlegte Symmetrie einem „perfekten“ geometrischen Körper zumindest nahekommt. Dieses führt zur stärkeren Ausprägung von Minima (Glatter & Kratky, 1982; Feigin & Svergun, 1987). Das Protein in Lösung ist aber immer ein leicht deformiertes Molekül, dessen Form mit der Zeit leicht fluktuiert – die Minima in einem SAXS-Datensatz sind also nicht so stark ausgeprägt. Diese Überlegung hat aber keinen Einfluss auf die Position der Minima.

Neben der Diskrepanz hinsichtlich der Skalierung der Daten, gibt es auch eine mit Blick auf den Intensitätsabfall. Wie in Abschnitt 4.2 angedeutet, so ist ein inadäquater Abzug der Pufferstreuintensitäten weitgehend auszuschließen. Folglich ist es naheliegender anzunehmen, dass die *single particle*-Rekonstruktion Fehler in der Dichteskalierung hatte oder die Dynamik der zur Verfügung

<sup>7</sup>Anderfalls müssten in einem Genprodukt mindestens zwei Kontaktflächen verschiedener Affinität der Untereinheiten zueinander realisiert sein.

<sup>8</sup>Es ist mir nicht bekannt, ob die Mikroskope inzwischen ausreichend für systematische Fehler in der Angabe der Vergrößerung kalibriert sind. Das Elektronenmikroskope einen solchen Fehler aufweisen ist nicht ungewöhnlich (persönliche Mitteilung Prof. Stahlberg, UC Davis, Kalifornien). Der Grund für diesen systematischen Fehler liegt in kleinen Abweichungen der Projektionsgeometrie beim Zusammenbau der Mikroskope vor Ort.

stehenden Grautiefe nicht ausgeschöpft wurde. Das rekonstruierte Volumen, welche von Professor Meissner zur Verfügung gestellt wurde, besaß ein Dichteminimum von -0,019, ein Maximum von 0,337 und einen Mittelwert von 0,008. Abgesehen von der Frage weshalb ein solches Volumen eine negative Dichte aufweisen kann<sup>9</sup>, wird durch diese Statistik deutlich, dass die Dichten nicht gleichmäßig verteilt sind. Es ergeben sich folgende Probleme eine theoretische Kurve zu errechnen:

- Ein Volumen von 224 x 224 x 224 Voxeln Kantenlänge, wie das für das Hämocyanin von *N. pompilius*, besitzt insgesamt 11239424 Voxel. Davon wurden bei einem Schwellenwert von 0,05 698508 Voxel ausgewählt. Folglich können sich Fehler durch die Darstellung der Fließkommazahlen im Rechner stark fortpflanzen (Denise & Zimmermann, 1999). Derartige Fehler treten aber ebenso bei der Berechnung der *single particle*-Rekonstruktion auf und können sich hier durch fehlerhafte „Rückprojektionen“ auch in einer insgesamt fehlerhaften Rekonstruktion niederschlagen (Frank, 2006).
- Weil die Dichtewerte der *single particle*-Rekonstruktion schief verteilt sind (vgl. Abschnitt 4.2, Seite 79) ist es schwierig einen Schwellenwert festzulegen, bei dem das umschlossene Volumen möglichst exakt mit der Masse des Proteins korreliert. Ein Fehler bei der Festlegung des Schwellenwertes kann dazu führen, dass die absolute Intensität, aber auch der generelle Verlauf fehlerhaft sind. Doch eine Variation der Schwellenwerte (vgl. Abbildung 4.28) führte zu keiner signifikanten Verbesserung des Fits, daher ist der Einfluss eines falschen Schwellenwertes zwar nicht ganz auszuschließen, aber als eher gering anzusehen.

Neben diesen numerischen Aspekten gibt es noch eine weitere Erklärung dafür, warum der Intensitätsverlauf zwischen experimenteller und theoretischer Kurve so stark voneinander abweicht: Ein grundlegender Fehler in der *single particle*-Rekonstruktion. Dieses wurde in einem Treffen mit Professor Meissner angesprochen. Es wurde von mir angeregt beim *Alignment* (Abschnitt 3.7.2.1, Seite 40) durch Einführen spezieller Dichten, die sich wieder herausmitteln müssen (eine übliche Methode der Kontrolle, Frank, 2006), sicherzustellen, dass die Rekonstruktion stimmig ist. Professor Meissner teilte mir mit, dass dies bereits geschehen sei.

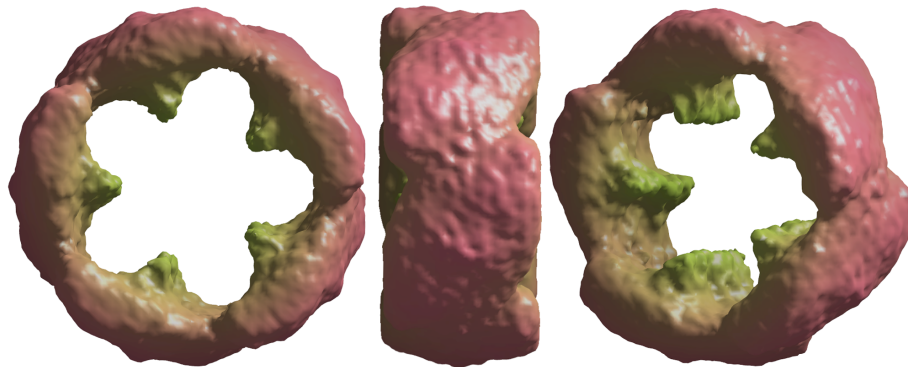


Abbildung 5.4: *Ab initio*-Modellierung des Hämocyanins mit MCSAS (Hartmann & Decker, 2004) auf der Basis von SAXS-Daten des *Nautilus pompilius*-Hämocyanins. Gezeigt ist die Rekonstruktion für das Dekamer mit D5-Symmetrie. Von links nach rechts: Sicht entlang der Symmetrieachse, die Symmetrieachse um 90° gedreht und in derselben Richtung um 45° weiter gedreht. Diese Abbildung ist bereits in Païret, 2007 vorgestellt und wurde durch Dr. H. Hartmann generiert.

Die Struktur des Proteins stellt einen relativ flachen Zylinder dar, an dessen Innenseite sich Ausstülpungen, die sogenannte Kragendomänen, befinden (vgl. Abbildung 5.4). Die erste Struktur wurde mit einer D5-Schönfliessymmetrie publiziert (Meissner *et al.*, 2006). In der darauffolgenden Publikation (Gatsogiannis *et al.*, 2007) wird dezidiert unterschieden zwischen einer D5-Symmetrie für die Wand und einer gebrochenen Symmetrie für die Kragendomäne, die damit nur mehr C5-Symmetrie aufweist. Die Autoren diskutieren ihren Fehler aus der ersten Publikation eingehend,

<sup>9</sup>Bei einer *single particle*-Rekonstruktion werden niemals Dichten von dem finalen Volumen abgezogen, also kann es theoretisch keine negativen Werte aufweisen.

erwähnen aber die Diskrepanz der ersten publizierten Struktur zu den SAXS-Daten, die den Fehler offensichtlich machte, nicht.

### *Ausblick*

Eine Aufnahme von SAXS-Daten unter Deoxybedingungen, wie ursprünglich geplant, scheint aufgrund von Materialmangel nicht mehr möglich. *N. pompilius* ist eine relativ seltene Art (vgl. Edwards *et al.*, 2000). Eine *single particle*-Rekonstruktion des Hämocyanins unter Deoxybedingungen wird demnächst von der Arbeitsgruppe um Professor Markl (Institut für Zoologie der Universität Mainz) eingereicht (persönliche Mitteilung von Professor Decker).

Würde die zweite Struktur des *N. pompilius*-Hämocyanins (aus Gatsogiannis *et al.*, 2007) zur Verfügung stehen, könnte auch diese mit den SAXS-Daten verglichen werden. Streudaten sind primäre Daten. Die Rekonstruktionen einer Dichte aus *single particle*-Daten stellen sekundäre Daten, beziehungsweise die Interpretation primärer Daten, dar. Der Vergleich interpretierter Daten mit unabhängig gewonnenen primären Daten kann einerseits zur Verifikation oder Falsifizierung der Interpretation (der Dichte aus den *single particle*-Daten; z. B. wie in Svergun *et al.*, 1994) und andererseits zu einer Verbesserung der Interpretation beitragen (z. B. wie in Gabashvili *et al.*, 2000). Für eine Verbesserung der elektronenmikroskopischen Rekonstruktion ist eine Korrektur der Amplituden im Fourierraum denkbar (siehe Gabashvili *et al.*, 2000). Ein systematischer Vergleich der Skalierungsfaktoren für das Volumen der Rekonstruktion und ihrer Voxelgröße wie in Abbildung 4.28 auf Seite 80 des Resultatekapitels ermöglicht überdies eine Einschätzung der Größe der Hydrathülle des *N. pompilius*-Hämocyanins (Svergun *et al.*, 1998).

Umgekehrt kann aber auch der direkte Vergleich der Dichten aus der *ab initio*-Rekonstruktion aus SAXS-Daten (vgl. Abbildung 5.4) mit der Dichte aus der *single particle*-Rekonstruktion zur Verifizierung der *ab initio*-Rekonstruktion, beziehungsweise des verwendeten MCSAS-Programms (Pairet, 2007), beitragen.

## 5.9 Strukturcharakterisierungen des $\alpha$ -Hämolytins

### 5.9.1 Charakterisierung des Monomers in Lösung

Um den Prozess der Porenenstehung des  $\alpha$ -Hämolytins auf struktureller Ebene nachvollziehen zu können, müssen die Strukturen des Toxins in den verschiedenen Stadien entlang des Pfades zur Pore (vgl. Abschnitt 1.4, Seite 13) im Detail bekannt sein. Während die Struktur der heptameren Pore gelöst ist (Song *et al.*, 1996), gilt dies bisher nicht für die übrigen Schritte. Alle Versuche kristallographisch arbeitender Gruppen die Struktur des Monomers des Toxins zu lösen sind bislang gescheitert (persönliche Mitteilung Prof. Bhakdi, Institut für Medizinische Mikrobiologie und Hygiene, Uniklinikum Mainz). An dieser Stelle setzt der Versuch ein, die Struktur des  $\alpha$ -Hämolytins mittels indirekter Methoden zu charakterisieren.

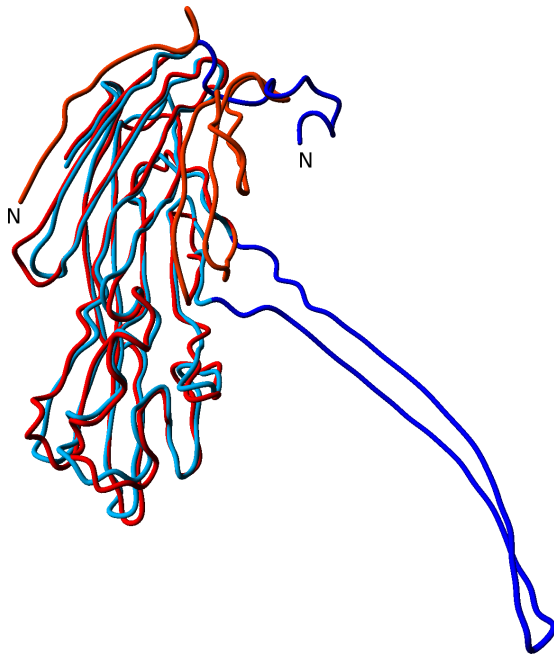


Abbildung 5.5: Superposition der Strukturen des  $\alpha$ -Hämolytins Monomers (rötlich; aus Abbildung 4.35 a)) und des Protomers (bläulich) aus der heptameren Porenstruktur (Song *et al.*, 1996). Diejenigen Abschnitte, die für den Konformationswechsel die größte Bewegung ausführen müssen, sind farblich abgesetzt. Die Abbildung ist bereits in Meesters *et al.*, 2008, veröffentlicht.

Im Gegensatz zum  $\alpha$ -Hämolytins sind für drei (eng verwandte) Leukocidine die kristallographischen Strukturen ihrer Monomerformen bekannt (Olson *et al.*, 1999; Pédelacq *et al.*, 1999; Guillet *et al.*, 2004). Diese Proteine sind eng verwandt mit dem  $\alpha$ -Hämolytins und erzeugen ähnliche Poren (Gouaux, 1998). Der Vergleich dieser Strukturen mit der Struktur des  $\alpha$ -Hämolytins in der Porenkonformation zeigt eine Reihe distinkter struktureller Unterschiede. Diese Unterschiede sind vor allem in der Umfaltung der Transmembrandomäne und des N-Terminus zu finden (siehe linke Abbildung). Andererseits besitzen die Toxine aber auch eine bemerkenswerte Ähnlichkeit, trotz einer Sequenzidentität von nur 23 – 32 %.

Alle bisherigen Strukturvergleiche sind auf statische Modelle beschränkt, die auf röntgenkristallographischen Daten beruhen. In dieser Arbeit steht der Versuch im Vordergrund, die Struktur und das Verhalten des Toxins in Lösung zu beschreiben. Hierzu dient ein theoretisches, komparatives Modell zusammen mit experimentellen Belegen. Dieser Ansatz folgt anderen Beispielen, in welchen SAXS-Daten dazu verwendet wurden, modellierte Strukturen zu verwerfen oder zu bestätigen (Perkins *et al.*, 1998; Trapani *et al.*, 2001; Egea *et al.*, 2001; Dimasi *et al.*, 2006; King-Scott *et al.*, 2007).

Die theoretische Streukurve des  $\alpha$ -Hämolytins Modells zeigt eine sehr gute Übereinstimmung mit den experimentellen Daten (siehe Abbildung 4.35 auf Seite 87 des Resultatekapitels). Die theoretischen Streukurven der Musterstrukturen für die Modellierung zeigten signifikante Differenzen zur experimentellen Kurve des  $\alpha$ -Hämolytins (siehe auch Tabelle 4.11 auf Seite 87). Folglich kann die in Olson *et al.*, 1999, geäußerte Annahme, dass die Strukturen der Leukocidine und des  $\alpha$ -Hämolytins grundsätzlich gleichartig sind, nicht unwiderrufen stehen bleiben und muss relativiert werden.

Auf dem ersten Blick sind „Deletionen“ und „Insertionen“ struktureller (nicht gelöste Teilabschnitte in den Musterstrukturen) und genetischer (Differenzen in den Primärstrukturen) Natur die wahrscheinliche Erklärung für die Unterschiede der Streukurven. Doch weder die Kürzung des N-Terminus, einzigartig für die Leukocidinstrukturen, noch die Modellierung der nicht gelösten Reste zwischen Position 129 und 135 in den Leukocidinstrukturen konnten die Musterstrukturu-

ren in eine Form überführen, die besser mit den SAXS-Daten in Einklang zu bringen war. Im Gegensatz dazu konnte gezeigt werden, dass das Insert von 10 Aminosäuren im  $\alpha$ -Hämolyisin an den Positionen 119 – 128 (L2 in Abbildung 4.35 a)), besonders wichtig für die Übereinstimmung theoretischer und experimenteller Daten, insbesondere in der Region  $q \approx 0, 2 - 0, 3 \text{ \AA}^{-1}$ , ist (siehe Abbildung 4.35 b) auf Seite 87). Dieser Abschnitt zeigt wahrscheinlich eine Schleifenkonfiguration. Andere Unterschiede können für die Teile der Modellstruktur gefunden werden, die definierte Sekundärstruktur aufweisen: Die postulierte  $\alpha$ -Hämolyisin Struktur besitzt drei  $3_{10}$ -Helices, deren Verlauf in den Musterstrukturen nur angedeutet ist.

Abgesehen von diesen Unterschieden ist das modellierte  $\alpha$ -Hämolyisinmonomer den Röntgenstrukturen der Leukocidinmonomere sehr ähnlich. Alle diese monomeren Strukturen zeigen einen deutlichen konformationellen Unterschied verglichen mit der Protomerstruktur der  $\alpha$ -Hämolyisinpore. Folglich sind es bestimmte Teile des Proteins, die flexibel sein müssen, um den Konformationswechsel  $\alpha_1 \rightarrow \alpha_7$  (siehe Abschnitt 1.4, Seite 14 in der Einleitung) auszuführen. Basierend auf dem Modell der Monomerstruktur wurde das Flexibilitätsprofil mittels einer Molekulardynamiksimulation bestimmt (siehe Abbildung 4.37 auf Seite 89). Die Regionen höchster Flexibilität können wie folgt klassifiziert werden:

- a) Bereiche, die man als „Gelenke“ (H für Englisch *hinge* in Abbildung 4.37 b)) des Konformationswechsels auffassen bzw. vermuten kann
- b) Bereiche, die ihrerseits die Konfiguration im Zuge des Konformationswechsels ändern müssen, wie z. B. der N-Terminus und die antiparallelen  $\beta$ -Faltblätter, welche das membranüberspannende  $\beta$ -Barrel in der Porenstruktur formen (beide als F für *flexible* in Abbildung 4.37 b) bezeichnet)

Die Phosphocholinebindedomäne (Gelb in Abbildung 4.37 a)) zeigt ebenfalls einige Flexibilität. Interessanterweise trifft dies auch auf die postulierten  $3_{10}$ -Helices zu. In der Porenstruktur ist nur die erste der Helices ebenfalls vorhanden, während die beiden anderen (von Rest 213 bis 215 und 218 bis 221) in der Porenstruktur Schleifen sind, die nur ungefähr dem Verlauf einer  $3_{10}$ -Helix folgen.

An dieser Stelle muss klargestellt werden, dass keinerlei Korrelation zwischen struktureller Flexibilität und Sequenzkonservierung gefunden werden konnte. Dennoch gibt es ein Zusammenfallen an einigen Positionen der  $\alpha$ -Hämolyisinsequenz: Es ist offensichtlich, dass der Winkel zwischen benachbarten Protomeren in der Membranebene der Porenstruktur zwischen  $\alpha$ -Hämolyisinporen und Leukocidinporen differieren muss, da erstere Homoheptamere sind, während Leukocidinporen wahrscheinlich Tetramere von Heterodimeren sind (Miles *et al.*, 2002; Jayasinghe & Bayley, 2005). Glycin 15 versetzt den N-terminalen Riegel in die Lage, einen weiten Winkelbereich für intermediäre Konformationen abzudecken, im Vergleich zu den Resten der anderen Sequenzen (siehe 4.38 auf Seite 90, Jayasinghe *et al.*, 2005; siehe auch Abbildung 5.6 a) auf der nächsten Seite). Das NGN-Motiv von Position 121 bis 123 ist zumindest ähnlich anderen Motiven im Alignment von  $\beta$ -CFTs, wo es oftmals zu GGN wird (siehe Abbildung 5.6 b)). Es ist ungefähr in der Mitte der Membranebene der *stem*-Domäne in der Porenkonformation und befindet sich in L2 des  $\alpha$ -Hämolyisinmodells (siehe Abbildung 4.35 a)). In vielen anderen Toxinen wird das Motiv zu GGN oder GGD. Der erste und dritte Aminosäurerest ragt in das Innere der Pore und ist hier wesentlich für das elektrostatische Potential verantwortlich (Aksimentiev & Schulten, 2005). Dieser gesamte Sequenzbereich, inklusive der flankierenden Reste sequenzab- und aufwärts, ist wahrscheinlich in einer Art *turn*- oder *coil*-Konfiguration (Garnier *et al.*, 1978), was wiederum die Idee der intrinsischen Flexibilität stützt. Und schließlich ist auch Glycin 134 hoch konserviert – es ist sogar in allen untersuchten Sequenzen vorhanden (siehe Abbildung 5.6 c)). Zusammen mit den anderen Glycinen und Prolinen, die in diesem Abschnitt gehäuft vorkommen, sorgt es für die besonders hohe Flexibilität in demjenigen Bereich, der die größte Konformationsänderung beim Konformationswechsel  $\alpha_1 \rightarrow \alpha_7$  durchzuführen hat.

Ausgehend von der Struktur (vgl. Abbildung 5.5 auf der vorherigen Seite) ist offensichtlich, dass auch dieser stark konservierte Bereich inhärent flexibel sein muss, um die Umfaltung zur Porenkonformation zu gewährleisten.

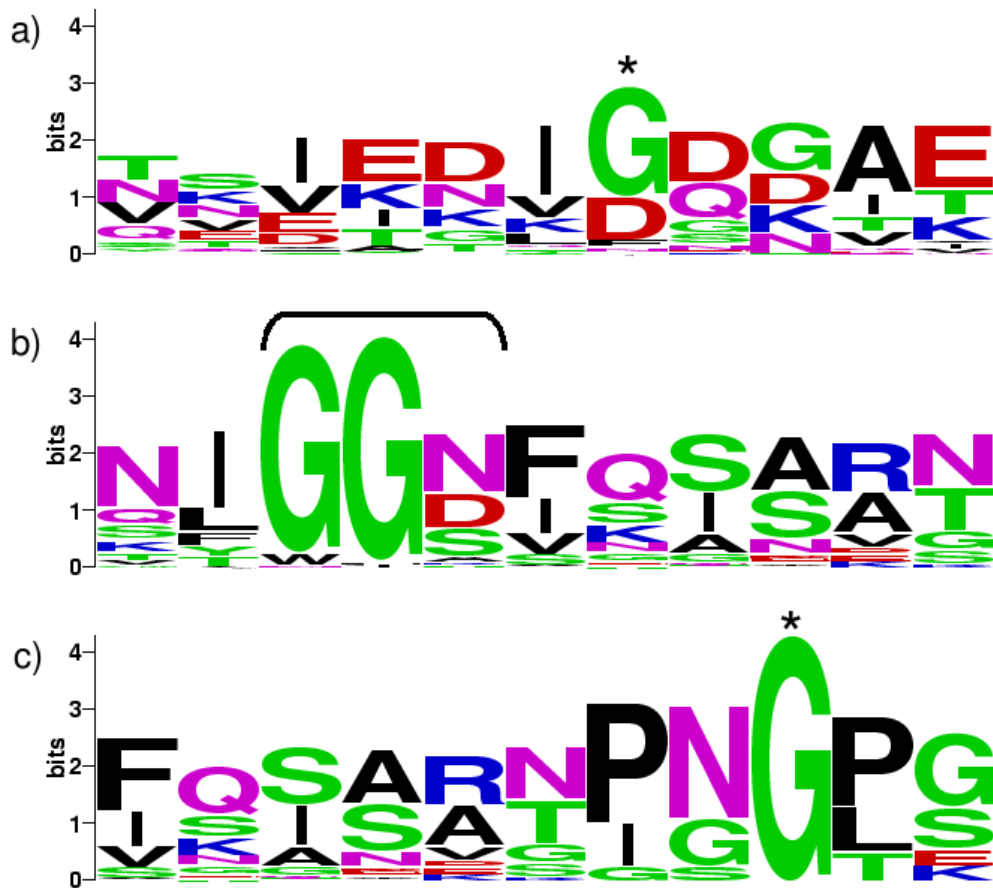


Abbildung 5.6: Sequenzlogos generiert mit *weblogo* (Crooks *et al.*, 2004) für ein Alignment von 107 Sequenzen der  $\beta$ -CFT-Familie (siehe Seite 52). a) Das Sternchen bezeichnet die Position von Glycin 15 des  $\alpha$ -Hämolytins. b) Die Klammer zeigt das NGN-Motiv des  $\alpha$ -Hämolytins. c) Das Sternchen bezeichnet die Position von Glycin 134 des  $\alpha$ -Hämolytins.

Abbildung 4.39 auf Seite 91 zeigt ein Phylogramm für die zum Alignment verwendeten Sequenzen. Es wird deutlich, dass das  $\alpha$ -Hämolytins in diesem Phylogramm relativ isoliert ist. Folglich ist der Informationsgehalt nach Gleichung 3.21 bezogen auf Abbildung 4.38 auf Seite 90 als relativ sicher anzusehen.

Wie kann die Flexibilität der Membran- oder Phosphocholinbindedomäne erklärt werden? Das  $\alpha$ -Hämolytins ist offenbar in der Lage, verschiedene Lipide in verschiedenen Zuständen zu binden (Watanabe *et al.*, 1987; Valeva *et al.*, 2006; allerdings mit unterschiedlichen Affinitäten). Dieses verlangt unmittelbar einige konformationelle Flexibilität. Der Zustand der membranbindenden Domäne kann also gewissermaßen mit einem „Saugnapfen“ verglichen werden, welcher vom Protein verwendet wird, um die Membran zu finden und sich zu einem „Anker“ umzufalten. Weitere Aspekte werden zu dieser Frage werden in Abschnitt 5.9.3 erörtert.

### 5.9.2 Vergleich von Präpore (H35R) und Wildtyp

Mit der Mutante H35R stand eine Variation des  $\alpha$ -Hämolytins zur Verfügung, die nicht in der Lage ist in die Porenkonformation zu wechseln (Jursch *et al.*, 1994; Panchal & Bayley, 1995). Die Konformation, welche diese Mutante final einnimmt, ist wahrscheinlich gleichzusetzen mit  $\alpha_{7,a}^*$  (Valeva *et al.*, 1997, siehe Abschnitt 1.4, Seite 13).

Für alle Experimente, die den experimentellen Vergleich zweier Punktmutanten<sup>10</sup> eines Proteins betreffen, muss sichergestellt sein, dass diese Mutanten – mit Ausnahme des untersuchten Aspektes – identisch sind, andernfalls können Schlussfolgerungen der Experimente in Zweifel gezogen werden. Ziel der Experimente mit H35R- $\alpha$ -Hämolytins war die Struktur der eingenommenen Präpore bei einer Auflösung zu ermitteln, die es erlaubt sterische Rahmenbedingungen für eine Zwischenstufe des Konformationswechsels  $\alpha_1 \rightarrow \alpha_7$  abzuleiten. Also musste geklärt sein, dass zumindest die Monomere – das H35R- $\alpha$ -Hämolytins und das Wildtyp- $\alpha$ -Hämolytins – strukturell identisch sind. Die Abbildungen 4.40 und 4.41 auf Seite 91 des Resultatekapitels belegen, dass die CD- und Fluoreszenzspektren der Proteine sehr ähnlich sind<sup>11</sup>. Damit ist festzuhalten, dass die Toxine als Monomere gleiche Sekundärstrukturanteile aufweisen und ihre aromatischen Aminosäuren die gleiche Lösungsumgebung erfahren.

Primäres Ziel der Messungen an H35R- $\alpha$ -Hämolytins war es eine SANS-Kurve zu erzielen, die es erlaubt mittels *rigid body*-Modellierung oder / und *ab initio*-Modellierung einen Strukturvorschlag für die Präpore zu entwickeln. Hierzu wurde eine für SANS-Messungen ausreichende Menge Protein auf zuvor hergestellte Vesikel gegeben (zur Herstellung der Vesikel siehe Abschnitt 3.10.4.3, Seite 49 ff.). Vorab wurde sichergestellt, dass die Messungen bei einer Deuteriumkonzentration stattfinden, bei der das Lipid der Vesikel nicht zum Kontrast beiträgt (siehe Abbildung 4.42 auf Seite 92). Die Idee für diese Vorgehensweise entstammt einer Arbeit für Sec A (Bu *et al.*, 2003). Bu *et al.* verwendeten einen Matchpunkt bei einem Deuteriumgehalt von 14,3% im Puffer, während der eigenständig ermittelte Matchpunkt bei 14,272%  $\pm$  2,367% liegt. Damit sind die beiden Punkte trotz unterschiedlicher Lipidkompositionen der Vesikel annähernd gleich.

Die eigentlichen SANS-Streukurven (Abbildung 4.43 auf Seite 93) zeigen deutlich, dass das „Matching“ sehr gut funktionierte: Die Lipide tragen sehr wenig bis gar nicht zum Kontrast bei. Das Signal stammt somit fast ausschließlich vom rekonstituierten Protein. Ebenso deutlich ist aber auch, dass die Streukurve nicht einer Streukurve freien Proteins entspricht, sondern größeren Zusammenlagerungen. Die Betrachtung der Probenküvetten für die SANS-Messung ließ bereits vermuten, dass sich verhältnismäßig schwere Komplexe gebildet haben müssen (vgl. Abbildung 4.44). Dies ist bestätigt durch eine anschließende Kontrolle der Lösungen im Elektronenmikroskop (siehe Abbildung 4.45 auf Seite 95): Das Toxin hat sich in beiden Fällen sehr dicht in den Vesikeln zusammengelagert, lediglich kleinere Mengen scheinen als frei in der Lösung befindliche Poren, beziehungsweise Präporen, befindlich. Damit war klar, warum eine „klassische“ Auswertung der SANS-Daten nicht möglich war.

Möglich war hingegen eine Auffassung der Vesikel als geometrische Körper (Zylinder oder Scheiben). Eine Bestimmung des mittleren Gewichtes pro Flächeneinheit war für das H35R- $\alpha$ -Hämolytins dennoch nicht möglich, im Gegensatz zum Wildtyp (vgl. Tabelle 4.12 auf Seite 93): Der entscheidende Unterschied zwischen den beiden Proben war, dass aus einem bisher unbekanntem Grund das Wildtyp- $\alpha$ -Hämolytins „flächigere“ Vesikel hinterließ als H35R- $\alpha$ -Hämolytins. Eine mögliche Ursache hierfür ist darin zu sehen, dass H35R die *stem*-Domäne nicht ausbildet und damit kein Lipid bei der Penetration der Membran verdrängt (Valeva *et al.*, 1997), so dass die freie, nicht von Toxin besetzte, Oberfläche von Lipiden in den Vesikeln geringer war als beim Wildtyp.

Ein interessanter Aspekt der Auswertung der SANS-Daten teilt sich aber in Tabelle 4.13 auf Seite 94 Resultatekapitels mit: Die Membrandicke, die zur Dekonvolution für die geometrischen

<sup>10</sup>Der Wildtyp ist hier als „Punktmutante“ der Punktmutante aufgefasst.

<sup>11</sup>Das CD-Spektrum scheint leichte Unterschiede aufzuweisen, aber sehr wahrscheinlich kam es zu einem fehlerhaften Abzug des Puffers oder zu einer falschen Skalierung auf die molare Elliptizität durch einen Fehler in der Konzentrationsbestimmung.

Körper angenommen werden muss, ist für das Wildtyp- $\alpha$ -Hämolytin größer als für das H35R- $\alpha$ -Hämolytin. Dies stützt bereits die Annahme, dass sich die *stem*-Domäne für das H35R nicht ausprägt und diese Mutante somit eine geringere Ausdehnung als der Wildtyp besitzt.

Um diese Vermutungen zu erhärten, wurden elektronenmikroskopische Aufnahmen angefertigt. Zunächst bestand Hoffnung, dass eine elektronenkristallographische Analyse durch die dichte Zusammenlagerung der  $\alpha$ -Hämolytinheptamere möglich ist. Abbildung 4.47 auf Seite 96 des Resultatekapitels zeigt aber, dass die so zu erzielende Auflösung gering bleiben würde: Selbst relativ gut streuende Kristalle zeigen nur Reflektionen erster Ordnung. Die Anordnung ist also nicht als kristallin, sondern als parakristallin einzustufen. Dies wird erhärtet durch die Tatsache, dass jedes betrachtete (para-)kristalline Vesikel eine andere Einheitszelle aufweist und die zugrundeliegende Symmetrie nur  $p1$  ist, also im Grunde keine Symmetrie vorliegt (siehe Tabelle 4.14 auf Seite 96). Im nächsten Schritt wurde auf einen zweidimensionalen *single particle*-Ansatz ausgewichen: Gut erkennbare (Prä-)Porenkomplexe wurden von den elektronenmikroskopischen Aufnahmen ausgewählt und aufeinander ausgerichtet (siehe Abschnitt 3.7, Seite 40 ff. im Methodenkapitel). Die Auflösung ist sogar für Arbeit im Negativkontrast durch den Kontrast relativ beschränkt<sup>12</sup>: Die Partikel sind klein und der Kontrast gering, da die Proteinpartikel mit Lipidvesikeln assoziiert waren. Daher war nur die Extramembrandomäne dem Kontrastmittel exponiert und ist daher „suboptimal“ abgebildet. Dennoch ist evident, dass sich Wildtyp- $\alpha$ -Hämolytin und H35R- $\alpha$ -Hämolytin unterscheiden. Während der Wildtyp eine zentrale Depression der Proteindichte aufweist, was mit der Pore korreliert werden kann, zeigt die Mutante nur eine geringe Dichtevariation (vgl. Abbildung 4.48 auf Seite 97).

Damit können zwei Dinge konstatiert werden:

1. Die Mutante H35R des  $\alpha$ -Hämolytins sitzt senkrecht auf der Membran. Der Beweis hierfür ist der für die Präpore in Abbildung 4.48 konstatierte Durchmesser, der mit etwa 100 Å genau dem Durchmesser der Pore in der Kristallstruktur entspricht (Song *et al.*, 1996). Dies stellt eine Bestätigung von älteren, nicht ausgewerteten elektronenmikroskopischen Daten dar (Jursch *et al.*, 1994). Im Gegensatz zu Jursch *et al.* konnte für die eigenen Daten kein signifikant unterschiedlicher Durchmesser der Komplexe gefunden werden<sup>13</sup>.
2. Bei der Mutante H35R ist die *stem*-Domäne, nicht wie bei der Monomerstruktur, am Rumpf der Protomere gehalten. Dies bedeutet, dass durch die Mutation ein Ausstülpfen des *stem*-Domäne zunächst nicht verhindert wird, die Penetration der Membran aber durchaus.

Eine mögliche Erklärung für den zweiten Befund gibt der folgende Abschnitt.

### 5.9.3 Spekulationen zum Konformationsübergang von $\alpha_7^*$ nach $\alpha_7$

Um den Konformationsübergang von einer Präpore ( $\alpha_7^*$ ) zur finalen Pore ( $\alpha_7$ ) besser nachvollziehen zu können, wurde eine geometrische Analyse mittels *Morphing* (siehe Abschnitt 3.8.4, Seite 43 im Methodenkapitel) versucht. Hierzu wurden die modellierten  $\alpha$ -Hämolytinmonomere auf die einzelnen Protomere der Porenstruktur ausgerichtet. Beim eigentlichen *Morphing* wird die so erhaltene fiktive Präpore sukzessive in die Porenkonformation überführt. Es ist sichergestellt, dass es zu keinen Verletzungen sterischer Bedingungen kommt und der Struktur keine geometrisch unmöglichen Bindungswinkel aufgezwungen werden.

Es ist plausibel, sich eine mögliche Präpore des  $\alpha$ -Hämolytins als einen heptameren Komplex vorzustellen, dessen Rotationsachse senkrecht zur Membran steht und dessen Protomer-Protomer-Kontakte ähnlich derjenigen in der Porenstruktur sind. Zudem ist dies durch eine Vielzahl biochemischer Arbeiten gestützt (Bayley, 1994; Bhakdi *et al.*, 1994; Valeva *et al.*, 1995; Walker *et al.*,

<sup>12</sup>Üblicherweise ist Kontrastarmut nur unter Kryo-Bedingungen problematisch.

<sup>13</sup>Jursch *et al.* haben keine *single particle*-Auswertung, sondern lediglich eine Bestimmung über einen Längenmaßstab durchgeführt. Sie geben für H35R-Präporen einen Durchmesser von 105 Å an gegenüber einem Durchmesser von 85 Å für die Pore des Wildtyps. Damit ist der Durchmesser der Wildtyp-Pore gegenüber dem kristallographisch bestimmten Durchmesser (100 Å, siehe Song *et al.*, 1996) hier so stark abweichend, dass auch für die Größenangaben aus Jursch *et al.* von einem nicht kalibrierten Elektronenmikroskop ausgegangen werden muss.

1994a; Walker *et al.*, 1994b; Walker & Bayley, 1995a; Walker & Bayley, 1995b; Valeva *et al.*, 1997). Insbesondere Arbeiten mit elektronenmikroskopischen Bildern (Jursch *et al.*, 1994) und AFM-generierten Konturen (Fang *et al.*, 1997; Malghani *et al.*, 1999) unterstützen diese Annahme direkt durch visuelle Belege für verschiedene präporenformende Mutanten. Damit ist noch keine Aussage darüber getroffen, in welchem Zustand diese Präpore genau ist, beziehungsweise welche exakte Struktur für einen Präporenkomplex vorzuschlagen ist.

In der Literatur gibt es also einige Arbeiten und Überlegungen zum Präporenstadium, doch keine Arbeit zum Konformationsübergang  $\alpha_7^* \rightarrow \alpha_7$  an sich. Die rechtsstehende Abbildung fasst verschiedene Wege zum Präporenstadium für  $\beta$ -channel forming toxins ( $\beta$ -CFTs;  $\beta$ -kanalformende Toxine) zusammen. Sie beschreibt die Wege zu einem Präporenstadium und macht auch klar, worin eine Schwierigkeit für den Konformationsübergang mancher Toxine liegen kann: Ist der Durchmesser der späteren Pore relativ zu der Länge der Transmembrandomäne klein, so kann der Konformationsübergang von der Präpore zur Pore nicht stattfinden, wenn zu wenige Monomere den Präporenkomplex bilden. Auch wenn die Präpore aus zu vielen Untereinheiten besteht, kann der Konformationsübergang verhindert werden, weil kein Ring geschlossen werden kann, der zur Bildung einer stabilen Pore unerlässlich ist.

Die Präporenstadien anderer Toxine können zum  $\alpha$ -Hämolyisin oder den LukF-/LukS-Präporen der  $\gamma$ -Hämolyisine sehr verschieden sein. Perfringolysin und das Cytolysin von *Clostridium perfringens* sind beispielsweise sehr viel größere Komplexe und unterliegen deshalb wahrscheinlich nicht den sterischen Einschränkungen des Konformationsüberganges des  $\alpha$ -Hämolyisins (Dang *et al.*, 2005; Rossjohn *et al.*, 2007)<sup>14</sup>.

Das *Morphing* (siehe Abschnitt 3.8.4, Seite 43 im Methodenkapitel) als geometrische Überlegung belegt zwei Dinge (vgl. Abbildungen 4.51 und 4.52 auf Seite 100 ff.):

1. Die Transmembrandomäne verhält sich beim Konformationsübergang zur Pore nicht wie ein rigider Körper, sondern äußerst flexibel. Die Sekundärstrukturanteile der Domäne verändern sich während des Übergangs fortwährend, lediglich der Bereich um die Reste 118 bis 120 liegt während der meisten *Morphing*-Schritte als  $\beta$ -Faltblatt vor.
2. Der Konformationsübergang ist möglich ohne, dass es zu einer wesentlichen Veränderung der Position der Protomer-Rümpfe kommen muss. Zwar drehen sich die Protomere während des *Morphings* ein wenig um ihre longitudinale Achse, aber die Protomer-Protomer-Kontakte bleiben im Wesentlichen erhalten.

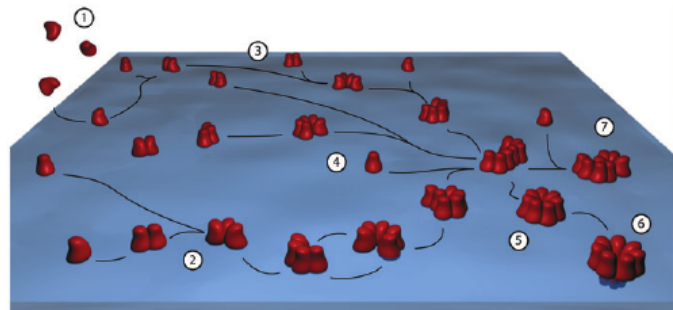


Abbildung 5.7: Schema zu Membraninsertion von  $\beta$ -CFTs.

① Bindung löslicher Monomere an die Membran mit oder ohne der Vermittlung durch spezifische Bindestellen. Laterale Diffusion und Protomer-Protomer-Interaktionen führen zur Oligomerisierung. Während der Oligomerisierung können sich zusätzliche Protomere an das Ende einer wachsenden Kette anlagern (Pfad ②) bis die richtige Länge zur Formung einer Präpore ⑤ erreicht ist. Manchmal mag es zu einem „Überschiessen“ der Oligomerisierung kommen, so dass sich zuviele Untereinheiten für ein Zusammenschließen zusammenlagern ⑦. Eine anderer möglicher Pfad (siehe ③ und ④) besteht in der Zusammenlagerung von Dimeren und kleinen Oligomeren bis zur Präpore. Diese alternativen Pfade konnten noch für kein  $\beta$ -CFT auseinandergehalten werden. Die Bildung einer Präpore ist wahrscheinlich irreversibel (Walker *et al.*, 1995). Der anschließende Konformationswechsel zur Pore ist langsam und, mit der Ausnahme sehr niedriger Monomerkonzentrationen, wahrscheinlich auch geschwindigkeitslimitierend. Die Abbildung und die Informationen hierzu sind Bayley *et al.*, 2005, entnommen. Die Erlaubnis für die Verwendung der Abbildung in dieser Arbeit durch die *Nature Publishing Group* liegt vor.

<sup>14</sup>Das hier angesprochene Cytolysin ist nicht mit der Struktur des Cytolysins von *Vibrio cholerae* zu verwechseln, dessen Porenstruktur wahrscheinlich in der Größenordnung der  $\alpha$ -Hämolyisinpore liegt (Olson & Gouaux, 2005; Valeva *et al.*, 2005; Pantano & Montecucco, 2006)

Die notwendige Verdrehung der Protomere gegeneinander, um Raum für die passierende Transmembrandomäne zu gewinnen, ist auch eine mögliche Erklärung dafür, warum H35R- $\alpha$ -Hämolytins nicht in der Lage ist eine Pore auszubilden. Das Arginin ist in einer hochpolaren, partiell negativ geladenen Region eingebettet (siehe Abbildung 5.8). Damit wird möglicherweise die energetische Hürde für die erforderliche Drehung der Untereinheiten zu groß.

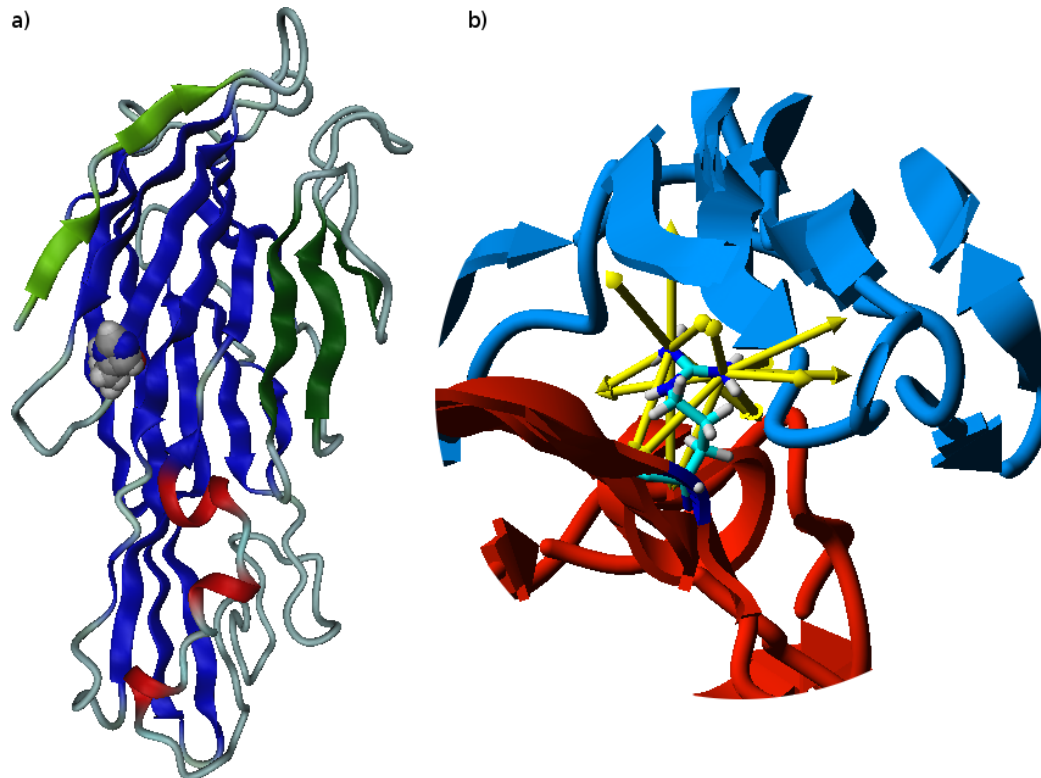


Abbildung 5.8: *Position des Arginins in der H35R-Mutante des  $\alpha$ -Hämolytins.* a) Die Struktur des Monomers mit der raumfüllenden Darstellung des Histidin 35. b) H35R in der Präporenstartstruktur des Morphings (siehe Abschnitt 3.8.4, Seite 43). Die Pfeile ausgehend von den Stickstoffatomen des Arginins zeigen auf alle Sauerstoffatome in einem Abstand  $\leq 5,5 \text{ \AA}$ . Rot einzeichnet ist die Matrix des Protomers des Arginins; Blau die Matrix des benachbarten Protomers.

Mit dem Morphing ist außerdem belegt, dass der Konformationswechsel  $\alpha_7^* \rightarrow \alpha_7$  an sich eine energetische Hürde darstellt (vgl. Abbildung 4.50 auf Seite 99). Dies erklärt, warum der Konformationsübergang der geschwindigkeitsbestimmende Schritt für die Ausbildung der Pore ist (Bayley *et al.*, 2005).

Eine abschließende energetische Analyse der *Morphing*-Schritte ist nicht möglich, da mit der zur Verfügung stehenden Rechenkraft kein Eintauchen der *stem*-Domäne in die Membran simuliert werden konnte. Doch der kontinuierliche Verlauf der Kurve in Abbildung 4.50 suggeriert, dass bis zum Eintauchen der *stem*-Domäne in die Membran beständig zusätzliche Energie benötigt wird. Jeder der *Morphing*-Schritte war energetisch unvorteilhafter als der vorhergehende.

Es ist denkbar, dass ein Teil der hierzu notwendigen Energie vom Druck der umgebenden Lösung stammt: Sobald die sich ausprägende Transmembrandomäne auf die Membran auftrifft, befindet sich auf der entgegengesetzten Seite eine leichte Einstülpung. Diese Situation tritt zumindest im *Morphing* auf und ist geometrisch bedingt, wenn alle Transmembranabschnitte der Monomere in dichter Folge umfallen.

Das *Morphing* kann schließlich auch die Frage beantworten, wie es dem  $\alpha$ -Hämolytins gelingt, die Lipide im Bereich der *stem*-Domäne beiseite zu schieben: Beim Auftreffen der *stem*-Domäne auf die Membran ist durch die Verdrehung der Untereinheiten gewährleistet, dass nur ein Strang einer Untereinheit auf die Membran trifft. Erst mit den weiteren Schritten wird nach und nach der volle Durchmesser erreicht.

## 5.10 *rigid body*-Modellierung von Proteinen auf SAXS-Datenbasis – eine Softwarebibliothek

### 5.10.1 Beurteilung der Qualität

Die `sas-rigid`-Bibliothek zur *rigid body*-Modellierung (siehe Abschnitt 4.7, Seite 105 ff.) stellt einen Versuch dar, *rigid body*-Modellierung auf der Basis von SAXS-Daten einfach und ohne Einschränkung auf besondere Fälle durchführen zu können. Im Großen und Ganzen ist es gelungen diesem Anspruch gerecht zu werden. Die erzielten Ergebnisse sind durchaus überzeugend und liefern Modelle, deren Interpretation im Kontext bekannter Rahmenbedingungen möglich ist (siehe Abschnitt 5.2.2, Seite 124).

Dennoch ist auch nicht zu verschweigen, dass ein jeder Algorithmus zur *rigid body*-Modellierung auf der Basis von SAXS-Daten von seiner Natur her heuristisch ist und sein muss: Jeder bekannte Algorithmus macht Annahmen über die Eigenschaften des zu modellierenden Systems, um die Rechenzeit einzuschränken.

Auch die `sas-rigid`-Bibliothek beinhaltet solche Ansätze. Im Prinzip sogar eine ganze Reihe davon: Nutzer der *Software* haben beispielsweise verschiedene Möglichkeiten, die Zahl der berücksichtigten Atome zu verringern, um Speicherplatz und Rechenzeit zu sparen. Es ist im Ermessen des Nutzers zu entscheiden, inwieweit das für den Einzelfall zu rechtfertigen ist. Dieser Umstand schränkt die leichte Nutzbarkeit sicherlich ein wenig ein.

Aber der einfache Zugang zu der *Software* ist durch eine Vielzahl von Beispielen und eine detaillierte Anleitung sicherlich gegeben.

### 5.10.2 Vergleich mit anderen Programmen zur *rigid body*-Modellierung

Die *rigid body*-Modellierung auf der Basis von SAXS-Daten ist im Allgemeinen datenorientiert: Die Orientierung von Körpern zueinander wird durch direktes Fitten an die Daten gefunden. Manche Programme nutzen aber zusätzliche Randbedingungen was ihre Nutzung einschränkt, wenn die erforderlichen Experimente nicht ausgeführt werden können. Die folgende Liste gibt einen kurzen Vergleich zu den mir bekannten Programmen zur *rigid body*-Modellierung:

- Es gibt den Vorschlag die Rechenzeit für eine Modellierung durch Repräsentation, der zu bewegend Unterheiten, in Form dreidimensionaler Ellipsoide zu verkürzen (Wall *et al.*, 2000). Die Streuung dieser einfachen Körper und ihre Orientierung zueinander ist schnell zu berechnen. Und auch die Positionierung ganzer Körper ist schneller berechnet, als die Positionierung hundertauser Atome. Diese beiden Punkte beschleunigen die gesamte Modellierung wesentlich. Das Programm von Wall *et al.* ist aber nicht mehr verfügbar, so dass ein direkter Vergleich ausgeschlossen ist. Es ist aber zu vermuten, dass diese Software durch zu die zu großen Vereinfachungen räumlich wenig präzise Modelle ergeben wird.
- In der *constraint fit procedure* (Boehm *et al.*, 1999; Sun *et al.*, 2004) werden Proteinuntereinheiten durch Kugeln repräsentiert und die theoretische Streuung tausender Variationen der Kugelpositionen wird auf den Fit zu den experimentellen Daten getestet. Zusätzliche Information durch Analytische Ultrazentrifugation wird benötigt, damit die finalen Modelle ausreichend akkurat für eine Interpretation sind. Diese Methode ist nicht in Form eines selbständigen Programms implementiert, so dass ein Test von meiner Seite nicht möglich war. Sie wird aber z. Zt. im Programm `Ultrascan` implementiert (s. u.; persönliche Mitteilung Professor B. Demeler (*University of Texas in San Antonio*)).
- Nöllmann *et al.* wenden einen *Monte Carlo*-Algorithmus mit *simulated annealing* an, der dem eigenen Algorithmus (vgl. Abschnitt 4.7.2, Seite 110) sehr stark gleicht (Nöllmann *et al.*, 2005). Im Grunde ist er besser als der eigene Algorithmus, da auch der sterische Überlapp von Untereinheiten bessere Berücksichtigung findet. Die Autoren stellen das Programm zur Verfügung, ich konnte es aber nicht testen, weil es sehr viele Abhängigkeiten an Drittsoftware und zusätzliche experimentelle Daten (aus FRET und / oder NMR) aufweist.

- Wieder eine andere Software berechnet die theoretische Streuung der Modelle durch sogenannte *spherical harmonics* und erlaubt es dem Nutzer in die Fitprozedur interaktiv einzugreifen (Kozin & Svergun, 2001). Das Programm ist wenig präzise und läuft nicht auf den heute gängigen Betriebssystemen.
- Das Programm *sas-ref* beinhaltet ebenfalls einen *Monte Carlo*-Algorithmus mit *simulated annealing*, sehr ähnlich dem eigenen Algorithmus (Petoukhov & Svergun, 2005). Es besitzt eine gute Dokumentation, funktionierte aber in meinen Händen nicht: Nach einer Rechenzeit von mehr drei bis fünf Tagen produzierte das Programm immer Strukturen, in der Untereinheiten „explodiert“ waren. Die Untereinheiten wiesen nach einer solchen Simulation keine Kontakte mehr auf und waren über mehrere Untereinheitendurchmesser deplaziert.
- Das Programm *parmscan* (Hartmann & Decker, 2002) wird an dieser Stelle nicht besprochen. Der Abschnitt 4.7.4 auf Seite 113 des Resultatekapitels bietet eine detaillierte Übersicht. Wichtig ist noch zu konstatieren, dass dieses Programm ausschließlich mit 24meren Hämocyaninen arbeiten kann.

Aus der Liste wird ersichtlich, dass es einige vielversprechende Ansätze gibt *rigid body*-Modellierung auf der Basis von SAXS-Daten in einer *Software* umzusetzen. Alle diese Ansätze haben aber für den Nutzer erhebliche Schwächen, so dass ich mich gezwungen sah eine eigenständige Alternative zu entwickeln.

### 5.10.3 Ausblick auf zukünftige Versionen der Software

Die *sas-rigid*-Software ist bei weitem noch nicht ausgereift. Einige Fähigkeiten müssen noch implementiert werden und insbesondere die Fehlerbeseitigung<sup>15</sup> (*Debugging*) wird noch Zeit in Anspruch nehmen. Folgende Fähigkeiten sind bereits implementiert, aber noch nicht Teil einer neuen Paket-Veröffentlichung (*software release*):

1. Modelle können jetzt auch beim Sicherheitsspeichern in einem Zug komprimiert werden. (Bislang war der Speicherplatzbedarf der Dateien hoch.)
2. Die Streukurve kann mittels der Debye-Gleichung (Gleichung 2.19) ermittelt werden.
3. Die neue *Monte Carlo*-Methode (siehe Abschnitt 4.7.5, Seite 115 ff.) wird in der nächsten Version enthalten sein.

Hinzu kommen diverse Fehlerbereinigungen.

Fähigkeiten, die eine gute Software zur *rigid body*-Modellierung haben sollte sind sicherlich

- a) die Möglichkeit auch eine Quasisymmetrie zwischen nicht identischen Untereinheiten eines Proteinkomplexes abzubilden und
- b) Lösungsmiteleinflüsse auf die Streukurve adäquat zu berücksichtigen. Dieser Punkt kann sehr gut bei einer Implementierung in *Ultrascan* eingefügt werden (s. u.).

Auf der Konferenz *From Computational Biophysics to Systems Biology* im Mai 2008 in Jülich (CBSB08: <http://www.fz-juelich.de/conference/cbsb08>) wurde ich von Professor B. Demeler (*University of Texas in San Antonio*, <http://www.demeler.uthscsa.edu/>), dem Hauptautor des Programms *Ultrascan* ((Demeler, 2005); <http://www.ultrascan.uthscsa.edu/>) zur Analyse von Daten der Analytischen Ultrazentrifugation angesprochen. Von ihm stammt die Anregung zur Implementierung des *Monte Carlo*-Algorithmus zur Darstellung von *rigid body*-Modellierungen auf SAXS-Datenbasis in Form von Dichten (vgl. Abschnitt 4.7.5, Seite 115 ff.). Er machte mir das Angebot die Komponenten von *sas-rigid* in *Ultrascan* zu implementieren. Auf diese Weise können Daten komplementärer Techniken (SAXS und Analytische Ultrazentrifugation) in einem Programm verarbeitet werden. Unserer Einschätzung nach ist es möglich diese

<sup>15</sup>Es wird davon ausgegangen, dass sich in *jedem* Quellcode einer Software vor einer systematischen Fehlersuche etwa 1 Fehler auf 100 Zeilen befindet und dass die Zahl der zu findenden Fehler bis zum Start systematischer Fehlerbeseitigung proportional zur Zahl gefundener Fehler ist (Tajima & Matsubara, 1981).

Implementierung innerhalb von einem halben Jahr umzusetzen, jedoch erfordert dies den Code in C++ (Stroustrup, 2007) umzuschreiben.

In *Ultrascan* ist die Repräsentation eines Proteins in Form von sogenannten *bead models* (Chacón *et al.*, 2000) bereits implementiert. Hierbei wird die gesamte Dichte eines Proteins in Form von Kugeln, den *beads*, ausgedrückt. Dies macht es einfach das Volumen des Proteins so zu vergrößern, dass auch die Hydrathülle umschlossen wird, wodurch Forderung b) der oben aufgeführten Liste sehr einfach zu erfüllen ist.

## 5.11 Das Programm SPlot

Wie im Resultatekapitel (Abschnitt 4.8) beschrieben, hat das Programm *SPlot* eine Reihe von Funktionen, welche die Arbeit mit der institutseigenen Kratky-Kamera vereinfachen können. Ziel war es in erster Linie eine graphische Benutzeroberfläche zur Behandlung der Kleinwinkelstreudaten zur Verfügung zu haben. Der gesamte Arbeitsfluss im Umgang mit Daten der Kratky-Kamera stellt sich in der Kombination aus *SPlot* & *PreToExcel* (siehe Anhang A.6) wie folgt dar (Abbildung 5.9):

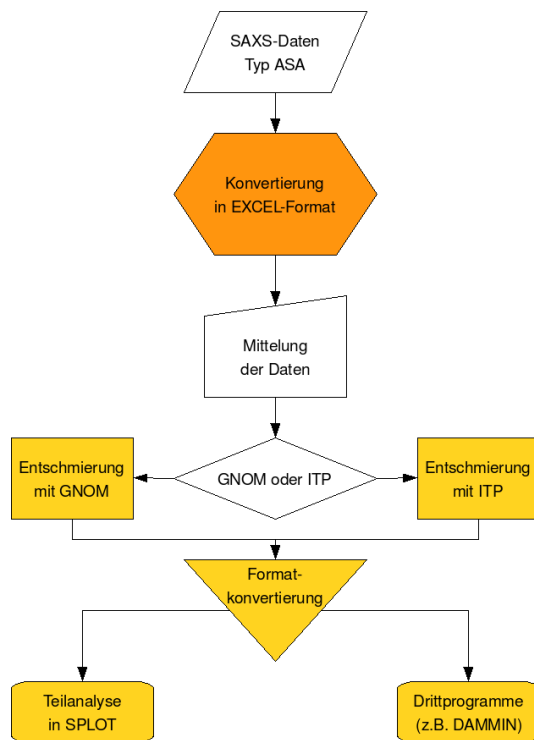


Abbildung 5.9: Arbeitsfluss für SAXS-Daten der Kratky-Kamera (ASA-Format) mit den Programmen *SPlot* & *PreToExcel*. Orange ist der Arbeitsschritt mit *PreToExcel* unterlegt, Gelb diejenigen Arbeitsschritte, bei denen *SPlot* assistiert.

*SPlot* bietet folgende Einsatzmöglichkeiten:

- Schnelle („*online*“) Überprüfung der von der Kamera aufgenommenen Daten. So ist es möglich einen Datensatz sehr schnell mit vorherigen Datensätzen zu vergleichen und festzustellen, ob sich eine Veränderung (z. B. Aggregation, Dissoziation, verändertes Plateauniveau (durch veränderten Zählgasfluss) usw.) ergeben hat.
- Arbeit mit Studenten, die oftmals nicht mit Kommandozeilenprogrammen umgehen können.
- Einfaches Generieren professioneller Abbildungen verschiedener Formate.
- Einfache Vergleiche verschiedener SAXS-Datensätze hinsichtlich SAXS-spezifischer Aspekte.
- Konvertieren von SAXS-Datensätzen in verschiedene Dateiformate.

Das Programm ist also vielfältig einsetzbar und ist zudem auch schnell und einfach zu erweitern. Es ist vergleichbar mit **Primus** (Konarev *et al.*, 2003), einem Programm zum Prozessieren von SAXS-Daten. Während der Funktionalitätsumfang ähnlich ist, bietet **Primus** im Gegensatz zu **SPlot** Schnittstellen zu anderen Programmen an und unterstützt andere Datenformate.

Der Hauptnachteil von **SPlot** ist die bisherige Beschränktheit auf das ASA-Datenformat als das primäre Eingabeformat. Dies kann aber bei Bedarf geändert werden, in dem ein zusätzlicher Parser für andere Formate geschrieben wird. Darüber hinaus unterliegt **SPlot** auch der Einschränkung auf UNIXartige Betriebssysteme (bisher getestet unter Linux und OS X). Die Ursache liegt in der Verwendung von wxPython (Rappin & Dunn, 2007), einer Pythonschnittstelle für die wx-Bibliothek (Smart *et al.*, 2005), einer Bibliothek für die plattformübergreifende Programmierung graphischer Benutzeroberflächen. Obwohl wxPython eigentlich Plattformunabhängigkeit garantiert ist es mir bislang nicht gelungen **SPlot** unter Windows ohne vielfache Abstürze auszuführen.

# Kapitel 6

## Zusammenfassung und Ausblick

In der vorliegenden Arbeit wurden die Methoden Homologiemodellierung und Molekulardynamik genutzt, um die Struktur und das Verhalten von Proteinen in Lösung zu beschreiben. Mit Hilfe der Röntgenkleinwinkelstreuung wurden die mit den Computermethoden erzeugten Vorhersagen verifiziert.

Für das  $\alpha$ -Hämolysin, ein Toxin von *Staphylococcus aureus*, das eine heptamere Pore formen kann, wurde erstmalig die monomere Struktur des Protein in Lösung beschrieben. Homologiemodellierung auf Basis verwandter Proteine, deren monomere Struktur bekannt war, wurde verwendet, um die monomere Struktur des Toxins vorherzusagen. Flexibilität von Strukturelementen in einer Molekulardynamiksimulation konnte mit der Funktionalität des Proteines korreliert werden: Intrinsische Flexibilität versetzt das Protein in die Lage den Konformationswechsel zur Pore nach Assemblierung zu vollziehen. Röntgenkleinwinkelstreuung bewies die Unterschiede der monomeren Struktur zu den Strukturen der verwandten Proteine und belegt den eigenen Vorschlag zur Struktur. Überdies konnten Arbeiten an einer Mutante, die in einer sogenannten Präporenkonformation arretiert und nicht in der Lage ist eine Pore zu formen, zeigen, dass dieser Übergangszustand mit der Rotationsachse senkrecht zur Membran gelagert ist. Eine geometrische Analyse beweist, dass es sterisch möglich ist ausgehend von dieser Konformation die Konformation der Pore zu erreichen. Eine energetische und kinetische Analyse dieses Konformationswechsels steht noch aus.

Ein weiterer Teil der Arbeit befasst sich mit den Konformationswechseln von Hämocyaninen. Diese wurden experimentell mittels Röntgenkleinwinkelstreuung verfolgt. Konformationswechsel im Zusammenhang mit der Oxygenierung konnten für die 24meren Hämocyanine von *Eurypelma californicum* und *Pandinus imperator* beschrieben werden. Für eine Reihe von Hämocyaninen ist nachgewiesen, dass sie unter Einfluss des Agens SDS Tyrosinaseaktivität entfalten können. Der Konformationswechsel der Hämocyanine von *E. californicum* und *P. imperator* bei der Aktivierung zur Tyrosinase mittels SDS wurde experimentell bestätigt und die Stellung der Dodekamere der Hämocyanine als wesentlich bei der Aktivierung festgestellt. Im Zusammenhang mit anderen Arbeiten gilt damit die Relaxierung der Struktur unter SDS-Einfluss und der sterische Einfluss auf die verbindenden Untereinheiten b & c als wahrscheinliche Ursache für die Aktivierung zur Tyrosinase.

Eigene Software zum sogenannten *rigid body*-Modellierung auf der Basis von Röntgenkleinwinkelstreuendaten wurde erstellt, um die Streudaten des hexameren Hämocyanins von *Palinurus elephas* und *Palinurus argus* unter Einfluss der Effektoren Urat und Koffein strukturell zu interpretieren. Die Software ist die erste Implementierung eines Monte Carlo-Algorithmus zum *rigid body*-Modelling. Sie beherrscht zwei Varianten des Algorithmus: In Verbindung mit *simulated annealing* können wahrscheinliche Konformationen ausgefiltert werden und in einer anschließenden systematischen Analyse kann eine Konformation geometrisch beschrieben werden. Andererseits ist ein weiterer, reiner Monte Carlo-Algorithmus in der Lage die Konformation als Dichteverteilung zu beschreiben.

# Literatur

- Aksimentiev, A. & Schulten, K. (2005). *Biophysical Journal* **88** (6), 3745–3761.
- Ali, S. A., Abbasi, A., Stoeva, S., Kayed, R., Dolashka-Angelova, P., Schwarz, H., & Voelter, W. (2000). *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **126** (3), 361–376.
- Ali, S. A., Grossmann, J. G., Abbasi, A., & Voelter, W. (2007). *Protein Pept Lett* **14** (5), 481–488.
- Ali, S. A., Zaidi, Z. H., & Abbasi, A. (1995). *Comparative Biochemistry and Physiology - Part A: Molecular and Integrative Physiology* **112** (1), 225–232.
- Allain, C. C., Poon, L. S., Chan, C. S., Richmond, W., & Fu, P. C. (1974). *Clinical Chemistry* **20** (4), 470–475.
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). *Nucleic Acids Research* **25** (17), 3389–3402.
- Anton-Paar-GmbH (2008). [www.anton-paar.com](http://www.anton-paar.com).
- Arnold, K., Bordoli, L., Kopp, J., & Schwede, T. (2006). *Bioinformatics* **22** (2), 195–201.
- Arvola, T., Ruuska, T., Keränen, J., Hyöty, H., Salminen, S., & Isolauri, E. (2006). *Pediatrics* **117** (4), e760–e768.
- Baird, S., Kelly, S. M., Price, N. C., Jaenicke, E., Meesters, C., Nillius, D., Decker, H., & Nairn, J. (2007). *Biochimica et Biophysica Acta* **1774** (11), 1380–1394.
- Bayley, H. (1994). *Journal of Cellular Biochemistry* **56** (2), 177–82.
- Bayley, H., Jayasinghe, L., & Wallace, M. (2005). *Nature Structural and Molecular Biology* **12** (5), 385–386.
- Berger, J., Pilz, I., Witters, R., & Lontie, R. (1977). *Monatshefte für Chemie* **108**, 1387–1396.
- Bergmann, S., Lieb, B., Ruth, P., & Markl, J. (2006). *Journal of Molecular Evolution* **62** (3), 362–374.
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., & Bourne, P. E. (2000). *Nucleic Acids Research* **28** (1), 235–242.
- Bernadó, P., Mylonas, E., Petoukhov, M. V., Blackledge, M., & Svergun, D. I. (2007). *Journal of the American Chemical Society* **129** (17), 5656–5664.
- Bernadó, P., Pérez, Y., Svergun, D. I., & Pons, M. (2008). *Journal of Molecular Biology* **376** (2), 492–505.
- Bernhard, H., Pilz, I., Meisenberger, O., Witters, R., & Lontie, R. (1983). *Biochimica et biophysica acta* **748**, 28–33.
- Bethe, H. A. & Heitler, W. (1934). *Proceedings of the Royal Society London* **A146**, 83–112.
- Beusch, A. (2006). *Isolierung und Charakterisierung von Hexamerin der Deutschen Schabe Blattella germanica*. Diplomarbeit, Universität Mainz.
- Bhakdi, S., Fussle, R., & Trantum-Jensen, J. (1981). *Proceedings of the National Academy of Sciences of the United States of America* **78** (9), 5475–9.
- Bhakdi, S., Jursch, R., Bröker, M., Ronneberger, H., & Hungerer, K. D. (1994). *Behring Institute Mitteilungen* (95), 80–84.
- Bhakdi, S. & Trantum-Jensen, J. (1991). *Microbiol Rev* **55** (4), 733–51.
- Blake, C. C., Koenig, D. F., Mair, G. A., North, A. C., Phillips, D. C., & Sarma, V. R. (1965). *Nature* **206** (986), 757–761.
- Blandamer, M. J., Cullis, P. M., Soldi, L. G., Engberts, J. B., Kacperska, A., Os, N. M. V., & Subha, M. C. (1995). *Advances in Colloid and Interface Science* **58** (2-3), 171–209.
- Boehm, M. K., Woof, J. M., Kerr, M. A., & Perkins, S. J. (1999). *Journal of Molecular Biology* **286** (5), 1421–1447.

- Boisset, N., Taveau, J. C., Lamy, J., Wagenknecht, T., Radermacher, M., & Frank, J. (1990). *Journal of Molecular Biology* **216** (3), 743–760.
- Bonaventura, C., Bonaventura, J., Miller, K. I., & Holde, K. E. V. (1981). *Archives of Biochemistry and Biophysics* **211** (2), 589–598.
- Bradford, M. M. (1976). *Analytical Biochemistry* **72**, 248–254.
- Bridges, C. R., Morris, S., & Grieshaber, M. K. (1984). *Respiratory Physiology* **57** (2), 189–200.
- Bro, R. & de Jong, S. (1997). *Journal of Chemometrics* **11**, 393–401.
- Bruker-ASX (2008). [www.bruker-axs.com/nanostar.html](http://www.bruker-axs.com/nanostar.html).
- Bréfort, J. (2005). <http://www.nongnu.org/gchempaint>.
- Bu, Z., Wang, L., & Kendall, D. A. (2003). *Journal of Molecular Biology* **332** (1), 23–30.
- Burmester, T. (2001). *Molecular Biology and Evolution* **18** (2), 184–195.
- Casper, G. S. (1985). *Journal of Arachnology* **13**(3), 277–283.
- Cavellec, A., Boisset, N., Taveau, J.-C., & Lamy, J. (1990). *Invertebrate Dioxygen Carriers*. Leuven University Press. CM45.
- Chacón, P., Díaz, J. F., Morán, F., & Andreu, J. M. (2000). *Journal of Molecular Biology* **299** (5), 1289–1302.
- Chacón, P., Morán, F., Díaz, J. F., Pantos, E., & Andreu, J. M. (1998). *Biophysical Journal* **74** (6), 2760–2775.
- Chadwick, J. (1932a). *Nature* **192**, 312.
- Chadwick, J. (1932b). *Proceedings of the Royal Society of London, A* **136**, 692–708.
- Cirak, S. (2000). *Isolierung und Charakterisierung des Schabenallergens Per a 3*. Diplomarbeit, Universität Mainz.
- Colangelo, N., Hellmann, N., Giomi, F., Bubacco, L., Muro, P. D., Salvato, B., Decker, H., & Beltramini, M. (2004). *Micron* **35** (1-2), 53–54.
- Cong, Y., Schweikardt, T., Woolford, D., Zhang, Q., Respress, J., Khant, H., Decker, H., Ludtke, S., & Chiu, W. (2008). noch nicht veröffentlicht: Sds induced conformational change of scorpion hemocyanin as revealed by cryo-electron microscopy at 8 Å resolution.
- Connelly, P. R., Johnson, C. R., Robert, C. H., Bak, H. J., & Gill, S. J. (1989). *Journal of Molecular Biology* **207** (4), 829–832.
- Conze, W. (2005). *Untersuchung zur Temperaturabhängigkeit der O<sub>2</sub>-Bindungseigenschaft von Hämocyanin aus Arthropoden verschiedener Biotope*. Doktorarbeit, Universität Mainz, Institut für Molekulare Biophysik.
- Cornell, W., Cieplak, P., Bayly, C., Gould, I., Merz, K., Ferguson, D., Spellmeyer, D., Fox, T., Caldwell, J., & Kollman, P. (1995). *Journal of the American Chemical Society* **117**, 5179–5197.
- Cottiglia, M. (1983). *Crostacei Decapodi lagunari. Guide per il riconoscimento delle specie animali delle acque lagunari e costiere Italiane*.
- Crooks, G. E., Hon, G., Chandonia, J.-M., & Brenner, S. E. (2004). *Genome Res* **14** (6), 1188–1190.
- Czajkowsky, D. M., Sheng, S., & Shao, Z. (1998). *Journal of Molecular Biology* **276** (2), 325–330.
- Dainese, E., Svergun, D., Beltramini, M., Muro, P. D., & Salvato, B. (2000). *Archives of Biochemistry and Biophysics* **373** (1), 154–162.
- Dang, T. X., Hotze, E. M., Rouiller, I., Tweten, R. K., & Wilson-Kubalek, E. M. (2005). *Journal of Structural Biology* **150** (1), 100–108.
- Decker, H. (1990). *Biophysical Chemistry* **37** (1-3), 257–263.
- Decker, H., Hartmann, H., Sterner, R., Schwarz, E., & Pilz, I. (1996). *FEBS Letters* **393** (2-3), 226–230.
- Decker, H., Hellmann, N., Jaenicke, E., Lieb, B., Meissner, U., & Markl, J. (2007). *Integrative and Comparative Biology* **47**, 631–644.
- Decker, H. & Jaenicke, E. (2004). *Developmental & Comparative Immunology* **28** (7-8), 673–687.
- Decker, H., Markl, J., Loewe, R., & Linzen, B. (1979). *Hoppe Seylers Z Physiological Chemistry* **360** (10), 1505–1507.
- Decker, H. & Rimke, T. (1998). *Journal of Biological Chemistry* **273** (40), 25889–25892.
- Decker, H., Robert, C., & Gill, S. (1986). *Invertebrate Oxygen Carriers* chapter Nesting – An extension of the allosteric model and its application to tarantula hemocyanin, pp. 383–388. Springer Verlag, Berlin.

- Decker, H., Ryan, M., Jaenicke, E., & Terwilliger, N. (2001). *Journal of Biological Chemistry*, **276** (21), 17796–17799.
- Decker, H., Savel-Niemann, A., Körschenhausen, D., Eckerskorn, E., & Markl, J. (1989). *Biol Chem Hoppe Seyler*, **370** (6), 511–523.
- Decker, H., Schweikardt, T., Nillius, D., Salzbrunn, U., Jaenicke, E., & Tucek, F. (2007). *Gene*, **398** (1-2), 183–191.
- Decker, H. & Sterner, R. (1990). *Naturwissenschaften*, **77**, 561–568.
- Decker, H. & Terwilliger, N. (2000). *Journal of Experimental Biology*, **203** (Pt 12), 1777–1782.
- Decker, H. & Tucek, F. (2000). *Trends in Biochemical Sciences*, **25** (8), 392–397.
- Demeler, B. (2005). *Modern Analytical Ultracentrifugation: Techniques and Methods*, chapter UltraScan: A Comprehensive Data Analysis Software Package for Analytical Ultracentrifugation Experiments, pp. 210–229. Royal Society of Chemistry (UK).
- Demeler, B. & Saber, H. (1998). *Biophysical Journal*, **74** (1), 444–454.
- Demeler, B., Saber, H., & Hansen, J. C. (1997). *Biophysical Journal*, **72** (1), 397–407.
- Denise, A. & Zimmermann, P. (1999). *Theoretical Computer Science*, **218**, 233–248.
- Diamond, R. (1974). *Journal of Molecular Biology*, **82** (3), 371–391.
- Dietze, C. (2008). *Untersuchungen zur Aktivierung von Hämocyanin zur Phenoloxidase*. Diplomarbeit, Universität Mainz, Institut für Molekulare Biophysik.
- Dimasi, N., Roessle, M., Moran, O., Candiano, G., Svergun, D. I., & Biassoni, R. (2006). *International Journal of Biological Macromolecules*, .
- Dubochet, J., Groom, M., & Müller-Neuteboom, S. (1982). *Advances in optical and electron microscopy*, volume 8, chapter The mounting of macromolecules for electron microscopy with particular reference to surface phenomena and the treatment of support films by glow discharge. In *Advances in optical and electron microscopy*, pp. 107–135. Academic Press, London, New York.
- Dutkiewicz, E. & Jakubowska, A. (2002). *Colloid Polymer Sciences*, **280**, 1009–1014.
- Dykens, J. A. (1991). *Comparative Biochemistry and Physiology*, **98B**, 579–582.
- Edwards, J. L., Lane, M. A., & Nielsen, E. S. (2000). *Science*, **289** (5488), 2312–2314.
- Egea, P. F., Rochel, N., Birck, C., Vachette, P., Timmins, P. A., & Moras, D. (2001). *Journal of Molecular Biology*, **307** (2), 557–576.
- Engelmann, D. M. & Moore, P. B. (2001). *International Table for Crystallography*, volume F, chapter Small-angle neutron scattering, pp. 438–443. Kluwer Academic Publishers, 1 edition.
- Erker, W., Scheumann, V., Möller, M., Knoll, W., Rühle, J., & Decker, H. (2006). *Micron*, **37** (8), 735–741.
- Fang, Y., Cheley, S., Bayley, H., & Yang, J. (1997). *Biochemistry*, **36** (31), 9518–9522.
- Feigin, L. & Svergun, D. (1987). *Structure Analysis by Small-Angle X-Ray and Neutron Scattering*. Plenum Press.
- Fiser, A., Do, R. K., & Sali, A. (2000). *Protein Science* **9** (9), 1753–1773.
- Fiser, A. & Sali, A. (2003). *Methods in Enzymology* **374**, 461–491.
- Fitter, J., Gutberlet, T., & Katsaras, J. (2007). *Neutron Scattering in Biology*. Springer Verlag.
- Frank, J. (2006). *Three-Dimensional Electron Microscopy of Macromolecular Assemblies*. Oxford University Press.
- Frank, J., Shimkin, B., & Dowse, H. (1981). *Ultramicroscopy* **6**, 343–358.
- Freer, J. H., Arbuthnott, J. P., & Bernheimer, A. W. (1968). *Journal of Bacteriology* **95** (3), 1153–68.
- Gabashvili, I. S., Agrawal, R. K., Spahn, C. M., Grassucci, R. A., Svergun, D. I., Frank, J., & Penczek, P. (2000). *Cell* **100** (5), 537–549.
- Galdiero, S. & Gouaux, E. (2004). *Protein Science* **13** (6), 1503–11.
- Galla, H.-J. (1988). *Spektroskopische Methoden in der Biochemie*, volume 1. Georg Thieme Verlag.
- Garnier, J., Osguthorpe, D. J., & Robson, B. (1978). *Journal of Molecular Biology* **120** (1), 97–120.
- Gatsogiannis, C., Moeller, A., Depoix, F., Meissner, U., & Markl, J. (2007). *Journal of Molecular Biology* **374** (2), 465–486.
- Gaykema, W. P., Volbeda, A., & Hol, W. G. (1986). *Journal of Molecular Biology* **187** (2), 255–275.
- Gaykema, W. P. J., Hol, W. G. J., Vereijken, J. M., Soeter, N. M., Bak, H. J., & Beintema, J. J. (1984). *Nature* **309**, 23–29.

- Glatter & Kratky (1982). *Small Angle X-ray Scattering*. ACADEMIC PRESS INC.
- Glatter, O. (1977). *Journal of Applied Crystallography* **10**, 415–421.
- Glatter, O. (1980). *Acta Physica Austria* **52**, 243–256.
- Gouaux, E. (1998). *Journal of Structural Biology* **121** (2), 110–122.
- Gouaux, J. E., Braha, O., Hobaugh, M. R., Song, L., Cheley, S., Shustak, C., & Bayley, H. (1994). *Proceedings of the National Academy of Sciences of the United States of America* **91** (26), 12828–31.
- Graham, R. A. (1985). *Comp Biochem Physiol B* **81** (4), 885–887.
- Gray, G. S. & Kehoe, M. (1984). *Infection and Immunity* **46** (2), 615–618.
- Gregory, D. R., Ronald, J., Labisky, F., & Combs, C. L. (1982). *Transactions of the American Fisheries Society* **111**(5), 575–584.
- Guillet, V., Roblin, P., Werner, S., Coraiola, M., Menestrina, G., Monteil, H., Prévost, G., & Mourey, L. (2004). *Journal of Biological Chemistry* **279** (39), 41028–41037.
- Guinier, A. (1939). *Annals of Physics* **12**, 161–237.
- Hamelryck, T. & Manderick, B. (2003). *Bioinformatics* **19** (17), 2308–2310.
- Harris, J. R. & Markl, J. (1999). *Micron* **30** (6), 597–623.
- Harrow, T. A., Wignall, G. D., & Katsaras, J. (2007). *Neutron Scattering in Biology* chapter Neutron Scattering for Biology, pp. 1–17. Springer Verlag.
- Hartmann, H., Bongers, A., & Decker, H. (2001). *European Biophysical Journal*, **30** (6), 471–475.
- Hartmann, H., Bongers, A., & Decker, H. (2004). *Journal of Biological Chemistry*, **279** (4), 2841–2845.
- Hartmann, H. & Decker, H. (2002). *Biochimica Biophysica Acta*, **1601** (2), 132–137.
- Hartmann, H. & Decker, H. (2004). *Methods in Enzymology*, **379**, 81–106.
- Hartmann, H., Lohkamp, B., Hellmann, N., & Decker, H. (2001). *Journal of Biological Chemistry*, **276** (23), 19954–19958.
- Hartmann, H., Müller, T., & Decker, H. (2004). *Micron*, **35** (1-2), 11–13.
- Hazes, B., Magnus, K. A., Bonaventura, C., Bonaventura, J., Dauter, Z., Kalk, K. H., & Hol, W. G. (1993). *Protein Science*, **2** (4), 597–619.
- Heimerl, R. & Decker, H. (1989). Technical Report 82, Verhandlungen der Deutschen Zoologischen Gesellschaft.
- Hellmann, N. (2004). *Biophysical Chemistry*, **109** (1), 157–167.
- Hellmann, N., Hörnemann, J., Jaenicke, E., & Decker, H. (2004). *Micron*, **35** (1-2), 109–110.
- Hellmann, N., Jaenicke, E., & Decker, H. (2001). *Biophysical Chemistry*, **90** (3), 279–299.
- Hellmann, N., Raitzel, K., & Decker, H. (2003). *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, **136** (3), 725–734.
- Hendricks, S. B. (1930). *Chemical Reviews*, **7**, 431.
- Hildebrand, A., Pohl, M., & Bhakdi, S. (1991). *Journal of Biological Chemistry*, **266** (26), 17195–200.
- Hilmenyuk, T. & Schwiering, M. (2006). F2-praktikumsprotokoll, Institut für Molekulare Biophysik, Universität Mainz.
- Hirai, M., Arai, S., & Iwase, H. (2000). *Thermochimica Acta*, **344**, 95–102.
- Hooft, R. W., Vriend, G., Sander, C., & Abola, E. E. (1996). *Nature*, **381** (6580), 272.
- Horwitz, A. (2001). *Journal of Computational and Applied Mathematics*, **134**, 1–11.
- Hsu, I.-J., Shiu, Y.-J., Jeng, U.-S., Chen, T.-H., Huang, Y.-S., Lai, Y.-H., Tsai, L.-N., Jang, L.-Y., Lee, J.-F., Lin, L.-J., Lin, S.-H., & Wang, Y. (2007). *Journal of Physical Chemistry A*, **111** (38), 9286–9290.
- Huggins, M. L. (1932). *Chemical Reviews*, **10**, 427.
- Hunter, E. (1999). *Crustaceana*, **72**(6), 545–565.
- Hunter, J. D. (2007). *Computing in Science and Engineering*, **9** (3), 90–95.
- Huson, D. H., Richter, D. C., Rausch, C., DeZulian, T., Franz, M., & Rupp, R. (2007). *BMC Bioinformatics*, **8**, 460.
- Hönig, D. (2004). *Aufbau einer Kratky-Kamera und Testversuche am Vogelspinnen-Hämocyanin*. Diplomarbeit, University of Mainz, Germany.
- Hübler, R. (2001). *Charakterisierung und Stabilität von Arthropoden-Hämocyanin*. Doktorarbeit, Institut für Molekulare Biophysik, Universität Mainz.

- Ikgai, H. & Nakae, T. (1985). *Biochemical and Biophysical Research Communications*, **130** (1), 175–81.
- Jaenicke, E. (1998). *Untersuchungen zur Bindung der Liganden Urat, Koffein, L-Laktat und Ecdyson an Hämocyanin*. Diplomarbeit,, Universität Mainz, Institut für Molekulare Biophysik.
- Jaenicke, E. & Decker, H. (2004a). *Micron*, **35** (1-2), 89–90.
- Jaenicke, E. & Decker, H. (2004b). *Chembiochem*, **5** (2), 163–169.
- Jaenicke, E. & Decker, H. (2008). *FEBS Journal*, **275** (7), 1518–1528.
- Jaenicke, E., Decker, H., Gebauer, W., Markl, J., & Burmester, T. (1999). *Journal of Biological Chemistry*, **274** (41), 29071–29074.
- Jaenicke, E., Fraune, S., May, S., Irmak, P., Augustin, R., Meesters, C., Decker, H., & Zimmer, M. (2008). Phenoloxidase is not expressed during the immune response in woodlice while their hemocyanin has phenoloxidase activity. zur Veröffentlichung einreicht.
- Jayasinghe, L. & Bayley, H. (2005). *Protein Science*, **14** (10), 2550–2561.
- Jayasinghe, L., Miles, G., & Bayley, H. (2005). *Journal of Biological Chemistry*, .
- Jeanmougin, F., Thompson, J. D., Gouy, M., Higgins, D. G., & Gibson, T. J. (1998). *Trends in Biochemical Sciences*, **23** (10), 403–405.
- Johansson, M. W. & Söderhäll, K. (1996). *Progress in Molecular and Subcellular Biology*, **15**, 46–66.
- Jones, E., Oliphant, T., Peterson, P., et al. (2001–2007).
- Jursch, R., Hildebrand, A., Hobom, G., Tranum-Jensen, J., Ward, R., Kehoe, M., & Bhakdi, S. (1994). *Infection and Immunity*, **62** (6), 2249–2256.
- Kabsch, W. & Sander, C. (1983). *Biopolymers*, **22** (12), 2577–2637.
- Kaufmann, T. C., Engel, A., & Rémy, H.-W. (2006). *Biophysical Journal*, **90** (1), 310–317.
- Kautz, G. & Topp, W. (2000). *Biology and Fertility of Soils*, **31**(2), 102–107.
- King-Scott, J., Nowak, E., Mylonas, E., Panjikar, S., Roessle, M., Svergun, D. I., & Tucker, P. A. (2007). *Journal of Biological Chemistry*, **282** (52), 37717–37729.
- Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). *Science*, **220** (4598), 671–680.
- Klabunde, T., Eicken, C., Sacchettini, J. C., & Krebs, B. (1998). *Nature Structural Biology*, **5** (12), 1084–1090.
- Konagurthu, A. S., Whisstock, J. C., Stuckey, P. J., & Lesk, A. M. (2006). *Proteins*, **64** (3), 559–574.
- Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., & Svergun, D. I. (2003). *Journal of Applied Crystallography*, **36**, 1277–1282.
- Koshland, D. E., Némethy, G., & Filmer, D. (1966). *Biochemistry*, **5** (1), 365–385.
- Koval, I. A., Gamez, P., Belle, C., Selmeczi, K., & Reedijk, J. (2006). *Chemical Society Reviews*, **35** (9), 814–840.
- Kozak, M. (2006). *Biopolymers*, **83** (6), 668–674.
- Kozin, M. & Svergun, D. I. (2001). *Journal of Applied Crystallography*, **34**, 33–41.
- Krieger, E., Darden, T., Nabuurs, S. B., Finkelstein, A., & Vriend, G. (2004). *Proteins*, **57** (4), 678–683.
- Krieger, E., Koraimann, G., & Vriend, G. (2002). *Proteins*, **47** (3), 393–402.
- Laemmli, U. K. (1970). *Nature*, **227** (5259), 680–685.
- Lafargue, B. (2003). *Isolierung und Charakterisierung der allergenen Per a 3- Isoformen von Periplaneta americana*. Diplomarbeit,, Universtität Mainz.
- Lafargue, B. (2008). *Isolierung und Charakterisierung allergener Proteine aus der amerikanischen Schabe (Periplaneta americana)*. Doktorarbeit,, Universität Mainz.
- Lai-Fook, J. (1970). *Journal of Morphology*, **130** (3), 297–313.
- Lamy, J., Bijlholt, M. C., Sizaret, P. Y., Lamy, J., & van Bruggen, E. F. (1981). *Biochemistry*, **20** (7), 1849–1856.
- Lamy, J., Sizaret, P. Y., Frank, J., Verschoor, A., Feldmann, R., & Bonaventura, J. (1982). *Biochemistry*, **21** (26), 6825–6833.
- Landau, D. P. & Binder, K. (2000). *A Guide to Monte Carlo Simulations in Statistical Physics*. Cambridge University Press.
- Leidescher, T. & Decker, H. (1990). *European Journal of Biochemistry* **187** (3), 617–625.
- Lipfert, J. & Doniach, S. (2007). *Annual Reviews of Biophysics and Biomolecular Structure* **36**, 307–327.
- Liu, S., Ton-that, H., & Magnus, K. (1997). Crystallographic studies of hydroxo, nitrosyl, azido and fluoro met form hemocyanin subunit ii from limulus.

- Loewe, R. (1978). *Journal of Comparative Physiology* **128**, 161–168.
- Loewe, R. & Linzen, B. (1973). *Hoppe Seylers Z Physiological Chemistry* **354** (2), 182–188.
- Loudon, K. (1999). *Mastering Algorithms in C*. O'Reilly.
- Magnum, C. P. (1985). *American Journal of Physiology*, R505–R514.
- Magnus, K. A., Hazes, B., Ton-That, H., Bonaventura, C., Bonaventura, J., & Hol, W. G. (1994). *Proteins* **19** (4), 302–309.
- Malghani, M. S., Fang, Y., Cheley, S., Bayley, H., & Yang, J. (1999). *Microsc Res Tech* **44** (5), 353–356.
- Markl, J. (1980). *Journal of Comparative Physiology B* **140**, 199–207.
- Markl, J. (1986). *The Biological Bulletin* **171**, 90–115.
- Markl, J. & Decker, H. (1992). *Adv. Comp. Environ. Physiol.* **13**, 325–376.
- Markl, J., Lieb, B., Gebauer, W., Altenhein, B., Meissner, U., & Harris, J. R. (2001). *J Cancer Res Clin Oncol* **127 Suppl 2**, R3–R9.
- Markl, J., Savel, A., Decker, H., & Linzen, B. (1980). *Hoppe Seylers Z Physiological Chemistry* **361** (5), 649–660.
- Markl, J., Savel, A., & Linzen, B. (1981). *Hoppe Seylers Z Physiological Chemistry* **362** (9), 1255–1262.
- Martin, A. G., Depoix, F., Stohr, M., Meissner, U., Hagner-Holler, S., Hammouti, K., Burmester, T., Heyd, J., Wriggers, W., & Markl, J. (2007). *Journal of Molecular Biology* **366** (4), 1332–1350.
- Martí-Renom, M. A., Stuart, A. C., Fiser, A., Sánchez, R., Melo, F., & Sali, A. (2000). *Annu Rev Biophys Biomol Struct* **29**, 291–325.
- Matoba, Y., Kumagai, T., Yamamoto, A., Yoshitsu, H., & Sugiyama, M. (2006). *Journal of Biological Chemistry* **281** (13), 8981–8990.
- May, S. (2005). *Enthält die Hämolymphe von Landasseln (Isopoda: Oniscidea) Hämocyanin?* Diplomarbeit, Universität Kiel.
- McBain, J. W. (1913). *Transactions of the Faraday Society* **9**, 93–107.
- Meesters, C. (2001). *Identifizierung der Extramembrandomänen im Na<sup>+</sup>/H<sup>+</sup>-Antiporter NhaA von Eschericia coli durch Markierung mit Goldpartikeln und Antikörperfragmenten an 2D-Kristallen*. Diplomarbeit, Universität Mainz.
- Meesters, C. (2008a). *From Computational Biophysics to Systems Biology (CBSB08)* volume 40 chapter A Software Library for Monte Carlo-Based Rigid Body Modelling Against Small Angle Scattering Data, pp. 321–323. John von Neumann Institute for Computing (NIC).
- Meesters, C. Anleitung: *Monte Carlo Rigid body modelling with SAS data - A Manual for the sas\_rigid Python package*. Molecular Biophysics Department, University of Mainz.
- Meesters, C., Brack, A., Hellmann, N., & Decker, H. (2008). Structural characterization of the alpha-hemolysin from staphylococcus aureus. akzeptiert zur Veröffentlichung in *Proteins: Structure, Function, and Bioinformatics*.
- Meissner, U., Gatsogiannis, C., Moeller, A., Depoix, F., Harris, J. R., & Markl, J. (2006). *Micron*, .
- Meissner, U., Stohr, M., Kusche, K., Burmester, T., Stark, H., Harris, J. R., Orlova, E. V., & Markl, J. (2003). *Journal of Molecular Biology*, **325** (1), 99–109.
- Menestrina, G. (1986). *Journal of Membrane Biology*, **90** (2), 177–90.
- Menger, F. M. (1979). *Accounts of Chemical Research*, **12** (4), 111–117.
- Menze, M., Hellmann, N., Decker, H., & Grieshaber, M. (2001). *Journal of Experimental Biology*, **204** (Pt 5), 1033–1038.
- Menze, M. A., Hellmann, N., Decker, H., & Grieshaber, M. K. (2000). *Biochemistry*, **39** (35), 10806–10811.
- Miles, G., Movileanu, L., & Bayley, H. (2002). *Protein Science*, **11** (4), 894–902.
- Miller, K., Eldred, N., Arisaka, F., & van Holde, K. (1977). *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology*, **115**, 171–184.
- Miller, K. & Holde, K. E. V. (1974). *Biochemistry*, **13** (8), 1668–1674.
- Monod, J., Wyman, J., & Changeux, J. P. (1965). *Journal of Molecular Biology*, **12**, 88–118.
- Müller, S. A. & Engel, A. (2001). *Micron*, **32** (1), 21–31.
- Müller, S. A., Pozidis, C., Stone, R., Meesters, C., Chami, M., Engel, A., Economou, A., & Stahlberg, H. (2006). *Molecular Microbiology*, **61** (1), 119–125.
- Nagai, T., Osaki, T., & Kawabata, S. (2001). *Journal of Biological Chemistry*, **276** (29), 27166–27170.

- Nellaiappan, K. & Vinayakam, A. (1993). *Biotech Histochem*, **68** (4), 193–195.
- Nicoli, D. F. & Benedek, G. B. (1976). *Biopolymers*, **15** (12NA-NA-770103-770104), 2421–2437.
- Nies, A., Zeis, B., Bridges, C. R., & Grieshaber, M. (1992). *Journal of Experimental Biology*, **168**, 111–124.
- Nillius, D. (2007). *Aktivierung von Hämocyanin zur Tyrosinase*. Doktorarbeit, Universität Mainz.
- Nillius, D., Jaenicke, E., & Decker, H. (2008). *FEBS Letters*, **582** (5), 749–754.
- Nishijima, S., Namura, S., Kawai, S., Hosokawa, H., & Asada, Y. (1995). *J Am Acad Dermatol*, **32** (4), 677–679.
- Noble, W. C. (1998). *Br J Dermatol*, **139 Suppl 53**, 9–12.
- Nöllmann, M., Stark, W. M., & Byron, O. (2005). *Journal of Applied Crystallography*, **38**.
- Olofsson, A., Hebert, H., & Thelestam, M. (1991). *Journal of Structural Biology*, **106** (3), 199–204.
- Olofsson, A., Kavéus, U., Hacksell, I., Thelestam, M., & Hebert, H. (1990). *Journal of Molecular Biology*, **214** (1), 299–306.
- Olson, R. & Gouaux, E. (2005). *Journal of Molecular Biology*, **350** (5), 997–1016.
- Olson, R., Nariya, H., Yokota, K., Kamio, Y., & Gouaux, E. (1999). *Nature Structural Biology*, **6** (2), 134–40.
- Orthaber, D., Bergmann, A., & Glatter, O. (2000). *Journal of Applied Crystallography*, **33**, 218–225.
- Ostfeld, E., Segal, J., Segal, A., & Bogokovski, B. (1983). *Isr J Med Sci*, **19** (12), 1046–1049.
- Pairet, B. (2007). *Röntgenkleinwinkelstreuung an Hämocyaninen – Untersuchung von Konformationsübergängen bei der Oxygenierung der Proteine*. Doktorarbeit, Fachbereich Physik, Johannes Gutenberg-Universität Mainz. cm83.
- Panchal, R. G. & Bayley, H. (1995). *Journal of Biological Chemistry*, **270** (39), 23072–23076.
- Pantano, S. & Montecucco, C. (2006). *Toxicon*, **47** (1), 35–40.
- Paoli, M., Giomi, F., Hellmann, N., Jaenicke, E., Decker, H., Muro, P. D., & Beltramini, M. (2007). *Gene*, **398** (1-2), 177–182.
- Paul, R. & Fincke, T. (1989). *Journal of Comparative Physiology B*, **159**, 419–432.
- Paul, R., Fincke, T., & Linzen, B. (1989). *Journal of Comparative Physiology B*, **159**, 409–418.
- Pauling, L. (1960). *The Nature Of The Chemical Bond*. Cornell University Press.
- Pavlov, M. Y. & Fedorov, B. A. (1983). *Biopolymers* **22**, 1507–1522.
- Pedersen, J. S. (1997). *Advances in Colloid and Interface Science* **70**, 171–210.
- Perkins, S. J., Ashton, A. W., Boehm, M. K., & Chamberlain, D. (1998). *International Journal of Biological Macromolecules* **22** (1), 1–16.
- Petoukhov, M. V., Eady, N. A. J., Brown, K. A., & Svergun, D. I. (2002). *Biophysical Journal* **83** (6), 3113–3125.
- Petoukhov, M. V. & Svergun, D. I. (2005). *Biophysical Journal* **89** (2), 1237–1250.
- Petterson, E. F., Goddard, T. D., Huang, C. C., Couch, G., Greenblatt, D., Meng, E., & Ferrin, T. (2004). *Journal of Computational Chemistry* **25** (13), 1605–1612.
- Philippsen, A. (2001). [www.dino3d.org](http://www.dino3d.org).
- Philippsen, A., Schenk, A. D., Signorell, G. A., Mariani, V., Berneche, S., & Engel, A. (2007). *Journal of Structural Biology* **157** (1), 28–37.
- Philippsen, A., Schenk, A. D., Stahlberg, H., & Engel, A. (2003). *Journal of Structural Biology* **144** (1-2), 4–12.
- Plewniak, F., Bianchetti, L., Brelivet, Y., Carles, A., Chalmel, F., Lecompte, O., Mochel, T., Moulinier, L., Muller, A., Muller, J., Prigent, V., Ripp, R., Thierry, J.-C., Thompson, J. D., Wicker, N., & Poch, O. (2003). *Nucleic Acids Research* **31** (13), 3829–3832.
- Plewniak, F., Thompson, J. D., & Poch, O. (2000). *Bioinformatics* **16** (9), 750–759.
- Pohlmann, A., Möller, M., Decker, H., & Schreiber, W. G. (2007). *Magn Reson Imaging* **25** (1), 129–135.
- Porod, G. (1951). *Kolloidzeitschrift* **124**, 83–114.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (2002). *Numerical Recipes in C++ The Art of Scientific Computing*. Cambridge University Press.
- Putnam, C. D., Hammel, M., Hura, G. L., & Tainer, J. A. (2007). *Q Rev Biophys* **40** (3), 191–285.
- Pédrelacq, J. D., Maveyraud, L., Prévost, G., Baba-Moussa, L., González, A., Courcelle, E., Shepard, W., Monteil, H., Samama, J. P., & Mourey, L. (1999). *Structure* **7** (3), 277–287.

- Rabenhorst, A. (2007). *Zur Kristallisation von multimeren Arthropoden-Hämocyaninen*. Diplomarbeit, Universität Mainz, Institut für Molekulare Biophysik.
- Ramachandran, G. N., Ramakrishnan, C., & Sasisekharan, V. (1963). *Journal of Molecular Biology* **7**, 95–99.
- Rappin, N. & Dunn, R. (2007). *wxPython in Action*. Manning Publications.
- Rice, L. B. (2006). *Am J Infect Control* **34** (5 Suppl 1), S11–9; discussion S64–73.
- Rice, P., Longden, I., & Bleasby, A. (2000). *Trends Genet* **16** (6), 276–277.
- Riley, P. A. (1997). *Int J Biochem Cell Biol* **29** (11), 1235–1239.
- Ringler, P., Heymann, B., & Engel, A. (2000). *Membrane Transport* chapter Two-Dimensional Crystallization of Membrane Proteins, pp. 229–268. Oxford University Press.
- Robert, C. H., Decker, H., Richey, B., Gill, S. J., & Wyman, J. (1987). *Proceedings of the National Academy of Sciences in the USA*, **84** (7), 1891–1895.
- Roeschlau, P., Bernt, E., & Gruber, W. (1974). *Z Klin Chem Klin Biochem*, **12** (5), 226.
- Rossjohn, J., Polekhina, G., Feil, S. C., Morton, C. J., Tweten, R. K., & Parker, M. W. (2007). *Journal of Molecular Biology*, .
- Round, A. R., Wilkinson, S. J., Hall, C. J., Rogers, K. D., Glatter, O., Wess, T., & Ellis, I. O. (2005). *Phys Med Biol*, **50** (17), 4159–4168.
- Roux, A. H., Héту, D., Perron, G., & Desnoyers, J. E. (1984). *Journal of Solution Chemistry*, **13** (1), 1–25.
- Roxby, R., Miller, K., Blair, D. P., & Holde, K. E. V. (1974). *Biochemistry*, **13** (8), 1662–1668.
- Rümcker, B. (1978). *SI-Einheiten - Gesetzliche Einheiten und ihre Anwendungspflicht in der Praxis ab 1978*. Kissing: WEKA Verlag.
- Sali, A. & Blundell, T. L. (1993). *Journal of Molecular Biology* **234** (3), 779–815.
- Savel-Niemann, A., Markl, J., & Linzen, B. (1988). *Journal of Molecular Biology* **204** (2), 385–395.
- Schmid, B. A. (2005). *Aufbau einer Röntgenkleinwinkelstreuanlage und Untersuchung von Zirkoniumoxid-Solen und weiterer Systeme*. Doktorarbeit, Universität Stuttgart.
- Schmid, W. (1994). *Optische Spektroskopie*. VCH Weinheim.
- Schneider, T. D. & Stephens, R. M. (1990). *Nucleic Acids Research* **18** (20), 6097–6100.
- Schweikardt, T. (2008). *Zur Struktur und Funktion der Typ-3-Kupferproteine sowie der Lichtsammler-komplexe LHCI und LHCII*. Doktorarbeit, Universität Mainz.
- Schweikardt, T., Jaenicke, E., & Decker, H. (2004). *Micron* **35** (1-2), 97–98.
- See, Y. & Jackowski, G. (1989). *Protein Structure* chapter Estimating molecular weights of polypeptides by SDS gel electrophoresis, pp. 1—21. Oxford, IRL Press.
- Segel, D. J., Eliezer, D., Uversky, V., Fink, A. L., Hodgson, K. O., & Doniach, S. (1999). *Biochemistry*, **38** (46), 15352–15359.
- Segel, D. J., Fink, A. L., Hodgson, K. O., & Doniach, S. (1998). *Biochemistry*, **37** (36), 12443–12451.
- Semisotnov, G. V., Timchenko, A. A., Melnik, B. S., Kimura, K., & Kihara, H. (2002). Technical Report 20B, Photon Factory.
- Shen, M.-Y. & Sali, A. (2006). *Protein Science*, **15** (11), 2507–2524.
- Sidgwick, N. V. (1933). *Annual Repetitions of the Chemical Society*, **29**, 64.
- Sizaret, P. Y., Frank, J., Lamy, J., Weill, J., & Lamy, J. N. (1982). *European Journal of Biochemistry*, **127** (3), 501–506.
- Skouri-Panet, F., Quevillon-Cheruel, S., Michiel, M., Tardieu, A., & Finet, S. (2006). *Biochim Biophys Acta*, **1764** (3), 372–383.
- Smart, J., Hock, K., & Csomor, S. (2005). *Cross-Platform GUI Programming with wxWidgets*. Prentice Hall PTR.
- Smith, A. J., Jackson, M. S., & Bagg, J. (2001). *J Med Microbiol* **50** (11), 940–946.
- Solomon, E., Sundaram, U., & Machonkin, T. (1996). *Chem Rev* **96** (7), 2563–2606.
- Solomon, E. I., Lowery, M. D., LaCroix, L. B., & Root, D. E. (1993). *Methods in Enzymology* **226**, 1–33.
- Song, L., Hobaugh, M. R., Shustak, C., Cheley, S., Bayley, H., & Gouaux, J. E. (1996). *Science* **274** (5294), 1859–66.
- Spinozzi, F., Maccioni, E., Teixeira, C. V., Amenitsch, H., Favilla, R., Goldoni, M., Muro, P. D., Salvato, B., Mariani, P., & Beltramini, M. (2003). *Biophysical Journal* **85** (4), 2661–2672.

- Stallman, R., Fox, B., & Ramey, C. (1988). [www.gnu.org/software/bash/bash.html](http://www.gnu.org/software/bash/bash.html).
- Stohr, M. (2007). *3D Cryo-Elektronenmikroskopie des 4x6-Hämocyanins der Vogelspinne Eurypelma californicum unter Oxy- und Deoxybedingungen*. Doktorarbeit, Universität Mainz, Institut für Zoologie.
- Stroustrup, B. (2007). *HOPPL III: Proceedings of the third ACM SIGPLAN conference on History of programming languages* chapter Evolving a language in and for the real world: C++ 1991-2006, pp. 4–14–59. New York, NY, USA: ACM.
- Stuhrmann, H., Burkhardt, N., Dietrich, G., Jünemann, R., Meerwinck, W., Schmitt, M., Wadzack, J., Willumeit, R., Zhao, J., & Nierhaus, K. (1995). *Nucl. Instrum. Methods A*, **A 356**.
- Stuhrmann, H. B. (1970). *Acta Crystallographica*, **A26**, 297–306.
- Stöcker, W., Raeder, U., Bijholt, U., Wichertjes, T., van Bruggen, E. F. J., & Markl, J. (1988). *Journal of Comparative Physiology*, **158B**, 271–289.
- Sun, Z., Reid, K. B. M., & Perkins, S. J. (2004). *Journal of Molecular Biology*, **343** (5), 1327–1343.
- Svergun, D. (1992). *Journal of Applied Crystallography*, **25**, 495–503.
- Svergun, D., Barberato, C., & Koch, M. (1995). *Journal of Applied Crystallography*, **28**, 768–773.
- Svergun, D. & Koch, M. H. J. (2003). *Reports on Progress in Physics*, **66**, 1735–1782.
- Svergun, D. I. (1999). *Biophysical Journal*, **76** (6), 2879–2886.
- Svergun, D. I., Pedersen, J. S., Serdyuk, I. N., & Koch, M. H. (1994). *Proceedings of the National Academy of Sciences of the United States of America*, **91** (25), 11826–11830.
- Svergun, D. I., Petoukhov, M. V., & Koch, M. H. (2001). *Biophysical Journal*, **80** (6), 2946–2953.
- Svergun, D. I., Richard, S., Koch, M. H., Sayers, Z., Kuprin, S., & Zaccai, G. (1998). *Proceedings of the National Academy of Sciences of the United States of America*, **95** (5), 2267–2272.
- Tajima, D. & Matsubara, T. (1981). *Software State-of-the-Art: Selected Papers*, chapter The Computer Software Industry in Japan, pp. 76–86. Dorset House Publishing, New York.
- Taylor, A. C., Astalla, C. M., & Atkinson, R. J. A. (2000). *Journal of Experimental Marine Biology and Ecology*, **244**, 265–283.
- Tehei, M. & Zaccai, G. (2007). *FEBS J*, **274** (16), 4034–4043.
- Telfer, W. H. & Kunkel, J. G. (1991). *Annu Rev Entomol*, **36**, 205–228.
- Terwilliger, N. B., Ryan, M. C., & Towle, D. (2005). *Journal of Experimental Biology*, **208** (Pt 13), 2467–2474.
- Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). *Nucleic Acids Research*, **22** (22), 4673–4680.
- Thompson, J. D., Prigent, V., & Poch, O. (2004). *Nucleic Acids Research*, **32** (4), 1298–1307.
- Thompson, J. D., Thierry, J. C., & Poch, O. (2003). *Bioinformatics*, **19** (9), 1155–1161.
- Tobkes, N., Wallace, B. A., & Bayley, H. (1985). *Biochemistry*, **24** (8), 1915–1920.
- Trapani, S., Linss, J., Goldenberg, S., Fischer, H., Craievich, A. F., & Oliva, G. (2001). *Journal of Molecular Biology*, **313** (5), 1059–1072.
- Truchot, J. P. (1980). *Journal of Experimental Zoology*, **214**(2), 205–208.
- Tsuruta, H. & Johnson, J. E. (2001). *International Tables for Crystallography*, volume F, chapter Small-angle X-ray scattering, pp. 428–437. Kluwer Academic Publishers, 1 edition.
- Ulrich, I. & Schreiner, P. (2007). Technical report, GKSS.
- Unruh, T., Bunjes, H., & Westesen, K. (1999). *Journal of Physical Chemistry B*, **103** (47), 10373–10377.
- Vachette, P., Dainese, E., Vasyliov, V. B., Muro, P. D., Beltramini, M., Svergun, D. I., Filippis, V. D., & Salvato, B. (2002). *Journal of Biological Chemistry*, **277** (43), 40823–40831.
- Valeva, A., Hellmann, N., Walev, I., Strand, D., Plate, M., Boukhallouk, F., Brack, A., Hanada, K., Decker, H., & Bhakdi, S. (2006). *Journal of Biological Chemistry*, **281** (36), 26014–26021.
- Valeva, A., Palmer, M., & Bhakdi, S. (1997). *Biochemistry*, **36** (43), 13298–13304.
- Valeva, A., Palmer, M., Hilgert, K., Kehoe, M., & Bhakdi, S. (1995). *Biochimica et Biophysica Acta*, **1236** (2), 213–218.
- Valeva, A., Walev, I., Boukhallouk, F., Wassenaar, T. M., Heinz, N., Hedderich, J., Lautwein, S., Möcking, M., Weis, S., Zitzer, A., & Bhakdi, S. (2005). *Molecular Microbiology*, **57** (1), 124–131.
- van Holde, K. & Weischet, W. (1978). *Biopolymers*, **17**, 1387–1403.
- van Holde, K. E. & Miller, K. I. (1982). *Quarterly Reviews of Biophysics*, **15** (1), 1–129.
- van Holde, K. E. & Miller, K. I. (1995). *Advances in Protein Chemistry*, **47**, 1–81.

- van Rossum, G. (1998). [www.python.org](http://www.python.org).
- van Schaick, E. J., Schutter, W. G., Gaykema, W. P., Schepman, A. M., & Hol, W. G. (1982). *Journal of Molecular Biology*, **158** (3), 457–485.
- Vandana, S., Raje, M., & Krishnasastri, M. V. (1997). *Journal of Biological Chemistry*, **272** (40), 24858–24863.
- Veerachato, G. (2006). *On The Cellular Stress Response To Staphylococcus aureus alpha-Toxin*. Doktorarbeit,, Universität Mainz.
- Voit, R., Feldmaier-Fuchs, G., Schweikardt, T., Decker, H., & Burmester, T. (2000). *Journal of Biological Chemistry*, **275** (50), 39339–39344.
- Volbeda, A. & Hol, W. G. (1988). *Prog Clin Biol Res*, **274**, 291–307.
- Volbeda, A. & Hol, W. G. (1989). *Journal of Molecular Biology*, **209** (2), 249–279.
- Volkov, V. V. & Svergun, D. I. (2003). *Journal of Applied Crystallography*, **36**, 860–864.
- Vriend, G. (1990). *Journal of Molecular Graphics*, **8** (1), 52–6, 29.
- Walker, B. & Bayley, H. (1995a). *Journal of Biological Chemistry*, **270** (39), 23065–23071.
- Walker, B. & Bayley, H. (1995b). *Protein Engineering*, **8** (5), 491–495.
- Walker, B., Braha, O., Cheley, S., & Bayley, H. (1995). *Chemistry and Biology*, **2** (2), 99–105.
- Walker, B., Kasianowicz, J., Krishnasastri, M., & Bayley, H. (1994a). *Protein Engineering*, **7** (5), 655–662.
- Walker, B., Kasianowicz, J., Krishnasastri, M., & Bayley, H. (1994b). *Protein Engineering*, **7** (5), 655–662.
- Wall, M. E., Gallagher, S. C., & Trehella, J. (2000). *Annual Reviews in Physical Chemistry*, **51**, 355–380.
- Ward, J. H. (1982). *American Statistical Association*, **58**, 236–244.
- Ward, R. J. & Leonard, K. (1992). *Journal of Structural Biology*, **109** (2), 129–141.
- Wardenburg, J. B. & Schneewind, O. (2008). *The Journal of Experimental medicine*, **205**, 287–294.
- Waschbüsch, D. (2004). *Charakterisierung von Hämocyaninen und Untersuchung der Proteininstabilität des Hämocyanins des Kaiserskorpions Pandinus imperator mit hydrodynamischen Methoden*. Diplomarbeit,, Institut für Molekulare Biophysik, Universität Mainz.
- Watanabe, M., Tomita, T., & Yasuda, T. (1987). *Biochimica et Biophysica Acta*, **898** (3), 257–65.
- Wells, M. J., Wells, J., & O’Dor, R. K. (1992). *Journal of Marine Biology Ass. U.K.*, **72**, 313–328.
- Wicker, N., Dembele, D., Raffelsberger, W., & Poch, O. (2002). *Nucleic Acids Research*, **30** (18), 3992–4000.
- Wicker, N., Perrin, G. R., Thierry, J. C., & Poch, O. (2001). *Molecular Biology and Evolution*, **18** (8), 1435–1441.
- Wiese-Posselt, M., Heuck, D., Draeger, A., Mielke, M., Witte, W., Ammon, A., & Hamouda, O. (2007). *Clin Infect Dis*, **44** (11), e88–e95.
- Wignall, G. D. & Bates, F. S. (1986). *Journal of Applied Crystallography*, **20**, 28–40.
- Winder, A. J. (1994). *J Biochem Biophys Methods*, **28** (3), 173–183.
- Winder, A. J. & Harris, H. (1991). *European Journal of Biochemistry*, **198** (2), 317–326.
- Wriggers, W. & Chacón, P. (2001). *Structure*, **9** (9), 779–788.
- Wyman, J. (1970). *Ciba Found Symp*, , 167–196.
- Wyman, J. (1984). *Quarterly Reviews of Biophysics*, **17** (4), 453–488.
- Zhang, R. G., Scott, D. L., Westbrook, M. L., Nance, S., Spangler, B. D., Shipley, G. G., & Westbrook, E. M. (1995). *Journal of Molecular Biology*, **251** (4), 563–573.
- Zhu, L., Kihara, H., Kojima, M., Zhou, J.-M., & Perrett, S. (2003). *Biochemical and Biophysical Research Communications*, **311** (2), 525–532.
- Zhu, L., Qin, Z.-J., Zhou, J.-M., & Kihara, H. (2004). *Biochimie*, **86** (2), 127–132.

# Anhang A

## Softwareskripte zur Evaluierung von Röntgenkleinwinkelstreuungsdaten

Die in der Zeit meiner Doktorarbeit erstellten Skripte wurden entweder in Python (van Rossum, 1998, ab Version 2.4) oder BASH (Stallman *et al.*, 1988, ab Version 3.0) geschrieben und dienten vornehmlich der Vereinfachung und Beschleunigung der Auswertung von SAXS- bzw. SANS-Daten. Sämtliche Skripte sind auch auf CD gespeichert und im Institut für Molekulare Biophysik der Universität Mainz erhältlich. Dennoch sind sie hier abgedruckt, um ggf. eine Reproduktion und Modifikation unabhängig ihrer elektronischen Version zu ermöglichen. Insbesondere meinen Nachfolgern in der Arbeitsgruppe möchte ich diese Möglichkeit nicht nur einräumen, sondern sie auch dazu ermuntern, diese Vorlagen zu nutzen. Der hier dargestellte Quellcode ist jedoch verkürzt wiedergegeben. Der eigentliche Code ist wesentlich reichhaltiger kommentiert und auf der CD, die zu dieser Arbeit archiviert wird, gespeichert.

### A.1 Skripte zur Steuerung der *EMBL-SAXS* Software Suite

#### A.1.1 DAMMIN

Das folgende Skript ist in BASH geschrieben und erlaubt es, das Programm DAMMIN (Svergun, 1999) so zu füttern, dass dieses gemäß der vorgegebenen Parameter eine Serie unabhängiger Rekonstruktionen erstellt. Die Rekonstruktionen werden stets unter einem neuen Dateinamen abgespeichert und können in einem späteren Schritt (siehe A.1.2) gemittelt werden.

Im Prinzip lassen sich die Programme CRED0 und GASBOR (Petoukhov *et al.*, 2002; Svergun *et al.*, 2001) in vergleichbarer Weise füttern. In der Praxis hat sich dieses allerdings als zu umständlich erwiesen und die Programme wurden per Hand gefüttert. Dennoch kann das folgende BASH-Skript als Muster dienen (Skript 1 auf der Folgeseite).

Bei allen diesen Skripten, die sich auch auf der CD mit den Daten dieser Arbeit befinden, ist es erforderlich bestimmte Parameter für den Einzelfall anzupassen<sup>1</sup>.

---

<sup>1</sup>In BASH und in Python sind '#' Kommentarzeichen, ab denen die Interpreter die Zeile nicht mehr ausführen.

---

**Skript 1** *Beispiel für ein Skript mit dem sich semi-automatisch verschiedene Läufe des Programms DAMMIN sequentiell starten lassen.*


---

```

1  #!/bin/bash
  #
  #####
  #
  # Input data the user is encouraged to change:
6  #
  # define mode: [F]ast, [S]low, [J]ag, [K]eep, [E]xpert
  MODE="S"
  # default input data file (as gnom.out):
  INPUT="gnom.out"
11 # project description string:
  PROJECT_DESCRIPTION="Eury_Hc_oxy"
  # Angular units in the input file (q=1/Aa (1) or q=1/nm (2))
  ANGULAR_UNIT="1"
  # Portion of the curve to be fitted
16 # PORTION="0.5"
  # Initial DAM: type S for sphere, E for ellipsoid, C for cylinder, P for parallelepiped or start file name
  INITIAL_DAM="S"
  # Symmetry to be imposed: P1...P19 or P23 or P432 or PICO
  SYMMETRY="P2"
21 # Particle shape: <P>rolate, <O>blate, or <U>nknown
  SHAPE="U"
  # if shape is oblate specify direction of expected anisotropy a<L>ong Z, a<C>ross Z, or <U>nknown
  ANIS="U"
  # Repeat over a certain number of Iterations
26 ITERATIONS="1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24"
  #
  #####
  #
  #define a help function
31 function usage()
  {
    echo "Available options are:"
    echo "<mode>: [F]ast, [S]low, [J]ag, [K]eep, [E]xpert"
    echo "<logfile>: somefilename, per default it will be generated as"
36    echo "as abbreviation from other input"
    echo "<input data file>: (as gnom.out) including path, if necessary"
    echo ""
    echo "invoke like:"
    echo "damminfeed [-m <mode>][-l <logfile>][-i <input data file>]"
41 }

  while getopts ":m:l:i:" opt; do
    case $opt in
46      m ) MODE=$OPTARG;;
        l ) LOGFILE=$OPTARG;;
        i ) INPUT=$OPTARG;;
        \? ) usage ; exit 1
    esac
  done
51 shift $((SOPTIND -1))

  for ITER in $ITERATIONS; do
    #create logfile name:
    TEMP=${INPUT##*/} #take only the input file name itself - not its path
56    #LOGFILE=${MODE} ${TEMP%.*} ${INITIAL_DAM} ${SYMMETRY} ${SHAPE} ${PORTION:0:1} ${PORTION:2:2} ${ITER}
    LOGFILE=${INITIAL_DAM} ${SYMMETRY} ${SHAPE} ${PORTION:0:1} ${PORTION:2:2} ${ITER}
    #check in which mode dammin should be started and run the appropriate routine]
    if [ $MODE = 'S' ]; then
61      echo "Starting dammin with the following parameters"
      echo ""
      printf "Mode: %3s          log file name: %15s\n" $MODE $LOGFILE
      printf "input file: %15s\n" $INPUT
      printf "project description: %s\n" $PROJECT_DESCRIPTION
      printf "portion of the curve to be fitted: %s\n" $PORTION
66      printf "initial volume shape: %s\n" $INITIAL_DAM
      printf "symmetry: %s\n" $SYMMETRY
      printf "particle shape %s\n" $SHAPE
      if [ $SHAPE = 'O' ]; then
71        printf "expected direction of anisotropy: %s\n" $ANIS
      fi

      sleep 3

      (dammin) <<- EOF
76          ${MODE}
          ${LOGFILE}
          ${INPUT}
          ${PROJECT_DESCRIPTION}
          ${ANGULAR_UNIT}
81          ${PORTION}
          ${INITIAL_DAM}
          ${SYMMETRY}
          ${SHAPE}
          EOF
86      else
        echo ""
        echo "CHOSEN MODE IS CURRENTLY NOT SUPPORTED"
        exit 1
      fi
91 done

  echo "damminfeed finished"

  exit 0

```

---

## A.1.2 DAMAVER

Das folgende Skript ist in BASH geschrieben und erlaubt das Füttern der DAMAVER-Suite (Volkov & Svergun, 2003). Hiermit können mit DAMMIN (Svergun, 1999) erstellte Rekonstruktionen (siehe A.1.1) gemittelt werden.

---

**Skript 2** *Beispiel für ein Skript mit dem sich semi-automatisch Ausgaben des Programms DAMMIN mitteln lassen.*

---

```

#! /bin/bash

# header information for script2html.py:

5 # bash comments for script2html.py starting with ';'
# keywords for script2html.py starting with '#~ [keyword]'
#~ [topic] SAXS
#~ [description] preparing the merger of dammin files
#~ [requirements] linblock
10 #~ [version] 0.1
#~ [author] C. Meesters
#~ [date] 7. July 2006

#version history
15 # 0.1 original implementation - merges feeds damaver

#define a help function
function usage()
{
20     echo "merge_prepare.sh"
     echo "This scripts creates a directory 'merger' and copies all *-1.pdb"
     echo "created by dammin from the local directory into this freshly created directory."
     echo "File names will be truncated to round1-<number>.pdb, where <number> is the number"
     echo "of the original pdb file before the "-1" - assuming several pdb files like:"
25     echo "name<number>-1.pdb"
     echo ""
     echo "Available options:"
     echo "-h - will display this help"
     echo "-d - can be used to specify a different directory name to be created."
30     echo "-n - can be used to specify a different file name pattern to be used for"
     echo "-a - will feed ALL found pdb files into damaver (usefull for MCSAS)"
     echo "      the output files"
     echo "-x <number> - will take the first x pdb files , in system order"
     echo "-r <referece file> - will use the given file as a reference for damsupsup"
35 }

#defining default directory to be created
DIRECTORY="merger"
#defining default file name pattern to be used
40 NAME="round2-"

while getopts ":d:n:x:r:" opt; do
45     case $opt in
         d ) DIRECTORY=$OPTARG;;
         n ) NAME=$OPTARG;;
         x ) NUMBER=$OPTARG;;
         r ) REFERENCE=$OPTARG;;
         \?) usage ; exit 1
50     esac
done
shift $((OPTIND -1))

55 #if NUMBER is not given by the user, set it to an insane one
declare -i NUMBER=$NUMBER
if [ "$NUMBER" -le 0 ]; then
    declare -i NUMBER=10000
fi
60

#message informing the user of what is specified
linblock "Calling merge_damming: Calling damaver"
echo "The following parameters are specified:"
echo "Directory to write into: ${DIRECTORY}"
65 echo "File pattern looks like: ${NAME}"

#check whether the output directory exists, if not create it
if [ ! -d ${DIRECTORY} ]; then
70     \mkdir ${DIRECTORY}
     cd ${DIRECTORY}
else
     echo "ERROR: The output directory, ${DIRECTORY}, already exists!"
     exit 1
75 fi

#copy all files ending on -1.pdb to the current directory and rename them according to
#the given pattern:
#for X in ../*.pdb; do
80 for X in ../*-1.pdb; do #DAMMIN
     declare DUMMY='ls ${X} | cut -d\ . -f3 | cut -d\ / -f2 '
     #echo $DUMMY
     declare FILENUM=${DUMMY:12:2} # for DAMMIN
     #echo $FILENUM
85     #declare FILENUM=${DUMMY:3:2} # for MCSAS
     echo "Copying ${X} to ${NAME}${FILENUM}.pdb"

```

---

```

cp ${X} ${NAME}${FILENUM}.pdb
#decrease number counter
NUMBER=$NUMBER-1
90 #exit this loop if the counter reaches zero; counter is a file counter
    if [ "$NUMBER" -le 0 ]; then
        break
    fi
done
95 #starting batch runs for damaver programs

#if no reference is given:
if [ ! -n "$REFERENCE" ]; then
100 damsel /b
else
    #create a simple damsups logfile with $REFERENCE as the reference
    touch damsups.log
    echo ${REFERENCE} >> damsups.log
105 #create a simple damsups logfile with $REFERENCE as the reference
    for X in *.pdb; do
        echo ${X} >> damsups.log
    done
fi
110 damsups /b
damaver /b
damfilt /b
damstart /b
115 exit 0

```

### A.1.3 dammin2dino

Nachdem mit DAMMIN (Svergun, 1999; Volkov & Svergun, 2003) und die verwandten Programme CREDO (Petoukhov *et al.*, 2002) und GASBOR (Svergun *et al.*, 2001) Modelle generiert wurden, galt es diese adäquat darzustellen. Hierzu musste das darstellende Programm in der Lage sein, Kugeln mit arbiträren Radius transparent darzustellen. `dammin2dino` ist ein einfaches Werkzeug, für Arbeiten in der `shell`, das überdies auch in der Lage ist, die Informationen über die Sekundärstruktur eines Proteines aus einer PDB-Datei einzulesen oder diese aus einer `dssp`-Logdatei (Kabsch & Sander, 1983) zu generieren:

#### Skript 3 *Das Skript dammin2dino zur Darstellung der ab initio-Modellierungen der Programme DAMMIN, GASBOR oder CREDO.*

```

1 from optparse import OptionParser
import os
import re
46
    likelihood = float(l[8])
    if likelihood > maximum:
        maximum = likelihood
        number = l[1]
        likelihoods[number] = likelihood

__author__ = "Christian Meesters"
6 __version__ = "0.2"
__date__ = "16. March 2007"
51
__history__ = """
9. October 2006 - 0.1 - original implementation
11 16. March 2007 - 0.2 - support for reading
    actual pdb files
    (reads out position
    for sec. structure)
56
20. March 2007 - 0.3 - enabled support for
    reading dssp files
16 """

__doc__ = """
21 This file accepts an arbitrary dammin or gasbor
    or damaver pdb output files and writes out dino
    source files for them.
    """
66

def _damminfile_reader(inputfile):
26 """
    will read in the likelihood values from
    the inputfile
    """
    print "reading likelihoods from input file"
31
    try:
        fname = open(inputfile, 'r')
    except IOError, msg:
        raise IOError(msg)
36
    likelihoods = {}
    #find out the overall maximum value
    maximum = 0
41
    for line in fname:
        l = line.split()
        if l[0] == 'ATOM':
86
            likelihoods[number] = likelihood

    fname.close()
51
    return likelihoods, maximum

def _pdb_reader(inputfile):
56 """
    will read in a pdb file and returns
    two lists:
    - tuples of residues (# of first and
    last AA) in helices
61
    - numbers of residues (# of first and
    last AA) in beta-sheets
    """
    try:
        fname = open(inputfile, 'r')
    except IOError, msg:
        raise IOError(msg)
66
    np = re.compile("\S{3}\sA\s+\d+",
                    re.IGNORECASE)
    n = re.compile("\d+")
71
    helix, sheet = list(), list()
    for line in fname:
        if 'HELIX' in line:
            tmp1 = re.search(n,
                np.findall(line)[0]).group()
            tmp2 = re.search(n,
                np.findall(line)[1]).group()
            helix.append((int(tmp1), int(tmp2)))
61
        if 'SHEET' in line:
            tmp1 = re.search(n,
                np.findall(line)[0]).group()
            tmp2 = re.search(n,
                np.findall(line)[1]).group()
            sheet.append((int(tmp1),
                int(tmp2)))
86

```

```

fname.close()
#anything found?
if len(helix) == 0 and len(sheet) == 0:
    raise AssertionError("No secondary
190     structure information read
        from pdb file")
return helix, sheet

95 def __dssp_reader(inputfile):
    """
        will read in a dssp log file and
        return two lists:
        - numbers of residues (# of first and
100     last AA) in helices
        - numbers of residues (# of first and
        last AA) in beta-sheets

        Note: This script does not distinguish
105     between E, S, and
        B labels - all are considered 'sheets'.
    """
    try:
        fname = open(inputfile, 'r')
110    except IOError, msg:
        raise IOError(msg)

    flag, first, last = False, 0, 0
    sfirstflag = False
    hfirstflag = False
115    helix, sheet = list(), list()
    for line in fname:
        l = line.split()
        if flag:
120            sheetcondition = ('E' in line or 'S'
                                in line or 'B' in line)
            helixcondition = ('H' in line)
            if sheetcondition:
                if not sfirstflag:
                    first = int(l[1]) #AA number
                    sfirstflag = True
                elif not sheetcondition and
225                    sfirstflag:
                    sfirstflag = False
                    last = int(l[1]) - 1
                    sheet.append((first, last))
            elif helixcondition:
                if not hfirstflag:
                    first = int(l[1]) #AA number
                    hfirstflag = True
                elif not helixcondition and
235                    hfirstflag:
                    hfirstflag = False
                    last = int(l[1]) - 1
                    helix.append((first, last))
140            # header finished ?
            if l[0] == '#': flag = True

    fname.close()
145    return helix, sheet

def script_writer(**option_dict):
    """
        will write out dino script files
150     according to the options
    """
    #input file name minus extension
    input = option_dict['inputfile'][-4]
155    print "Starting to write out dino
        script file."

    try:
        fname = open(option_dict['outputfile'],
160            'w')
    except IOError, msg:
        raise IOError(msg)

    if option_dict['pdbfile'] and not
165    option_dict['dssp']:
        helix, sheet = __pdb_reader(
            option_dict['pdbfile'])
    elif option_dict['pdbfile'] and
    option_dict['dssp']:
170        helix, sheet = __dssp_reader(
            option_dict['dssp'])
    if option_dict['pdbfile']:
        fname.write('set solvent rname=HOH,
175     H2O,WAT,WTR,CA,ZN,FE,SE,
        MN,SO4\n')
        fname.write('set base aname=N1,C2,N2,
        O2,N3,C4,N4,O4,C5,C5M,C6,
180     N6,O6,N7,C8,N9\n')
        fname.write('set protein rname=ALA,CYS,
        ASP,GLU,PHE,GLY,HIS,ILE,LYS,
        LEU,MET,ASN,PRO,GLN,ARG,SER,
        THR,VAL,TRP,TYR,5HP,ABA,PCA,
185     FGL,BHD,HTR,MSE,CEA,ALS,TRO,
        TPQ,MHO,IAS,HYP,CGU,CSE,RON,
        3GA,TYS\n')
        fname.write('set na $dna or $rna\n')
        fname.write('set backbone aname=CA,C,N,O,
        OXT,HA,HN\n')
        fname.write('set hydrophobic rname=TRP,
        ILE,MET,VAL,TYR,LEU,PHE\n')
        fname.write('set basic rname=ARG,LYS\n')
        fname.write('set acidic rname=ASP,GLU\n')
        fname.write('set polar rname=SER,THR,TYR,
        HIS,CYS,ASN,GLN\n')
        fname.write('set aromatic rname=PHE,TYR,
        HIS,TRP\n')
        fname.write('set aliphatic rname=ALA,GLY,
        ILE,LEU,MET,PRO,VAL\n')
        fname.write('set purin rname=A,ADE,G,
        GUA\n')
        fname.write('set pyrimidin rname=C,CYT,T,
        THY,U,URA\n')
        fname.write('scene hide\n')
        fname.write('load %s -name dflt\n' %
205     option_dict['pdbfile'])
        if len(helix) > 0:
            fname.write('set helix rnum=')
            nl = 0
            for t in helix[-1]:
                fname.write('%d:%d, ' %
210     (t[0], t[1]))
            fname.write('%d:%d\n' %
                (helix[-1][0],
                helix[-1][1]))
        if len(sheet) > 0:
            fname.write('set strand rnum=')
            nl = 0
            for t in sheet[-1]:
                fname.write('%d:%d, ' %
220     (t[0], t[1]))
            fname.write('%d:%d\n' %
                (sheet[-1][0],
                sheet[-1][1]))
        fname.write('\n')
        fname.write('. dflt set
225     rtype=strand -sel $strand\n')
        fname.write('set h hydrophobic\n')
        fname.write('. dflt new -name ca
            -type trace\n')
        fname.write('. dflt.ca set
230     color=grey\n')
        fname.write('. dflt.ca set
            color=blue -sel
            rtype=strand\n')
        fname.write('. dflt.ca set
235     color=red -sel
            rtype=helix\n')
        fname.write('. dflt.ca show\n')

        fname.write('. dflt.ca render
240     hsc, t=1.\n')

        fname.write('. dflt new -name
        sell -sel rnum=35\n')
        fname.write('. dflt.sell
        render cpk\n')

        fname.write('//load in all data\n')
        fname.write('load %s.pdb\n' % input)
        if option_dict['surface']:
            fname.write('load %s -type msms
250     -name surf\n' %
                option_dict['surface'])
        fname.write('\n//center the scenes\n')
        fname.write('scene center [%s]\n' % input)
        if option_dict['pdbfile']:
            fname.write('scene center [%s]\n' %
255     'dflt')
        fname.write('scene autoslab\n')
        fname.write('scene set bg=%s\n' %
            option_dict['background'])
        fname.write('scene set near=-40,
260     far+=40\n')
        fname.write('\n//rendering\n')
        fname.write('%s new -name all\n' % input)
        fname.write('%s.all set color=%s\n' %
            (input, option_dict['ball_color']))
        fname.write('%s.all render custom,
265     sr=%s,bw=0.3,detail=3,t=0.4\n' %
            (input, option_dict['radius']))
        fname.write('\n\n')
        if option_dict['surface']:
            fname.write('. surf attach %s\n' % input)
            fname.write('. surf new -name env\n')
            fname.write('. surf.env set color=blue\n')
            #fname.write('. surf.env render custom\n')
        fname.write('\n//manipulate the scene\n')
        fname.write('\n//put backwards\n')
        fname.write('scene transz -150\n')
        fname.write('scene center [%s]\n' % input)
        fname.write('scene autoslab\n')

```

```

if option_dict['likelihood_flag']:
    likelihoods, maximum =
    _damminfile_reader(
285         option_dict['inputfile'])
        fname.write('\n\n//setting the
            confidence rendering\n')
        for key, value in likelihoods.iteritems():
            value = int(value * 255 / maximum)
290             fname.write('%s.all set
                color={0,0,%d}
                -sel anum=%s\n' %
                    (input,value,key))
            fname.write('scene set bg=white\n')
295             fname.write('scene show\n')
            fname.close()

def main():
    parser = OptionParser(usage=__doc__)
300     parser.add_option("-i", "--input",
        action="store",
        type="string",
        dest="input",
        help="use to specify
305         the input file name,
            default is 'damaver.pdb'")
        parser.add_option("-o", "--output",
            action="store",
            type="string",
            dest="output",
310             help="use to specify
                the output file name,
                    default is 'msmsin.dat'")
            parser.add_option("-l", "--likelihood",
                action="store_true",
                dest="likelihood_flag",
                default=False,
                help="coloring according to
315             likelihood")
            parser.add_option("-s", "--surface",
                action="store",
                type="string",
                dest="surface",
320             help="use to specify
                msms surface file (without
                    extension!)")
            parser.add_option("-c", "--color",
                action="store",
                type="string",
                dest="ball_color",
330             help="use to specify
                the ball color (defaults to
                    red)")
            parser.add_option("-b", "--background",
                action="store",
                type="string",
                dest="background",
335             help="use to specify
                the background color
                    (defaults to white)")
            parser.add_option("-r", "--radius",
                action="store",
                type="string",
                dest="radius",
340             help="use to specify
                the ball radius
                    (defaults to 8)")
            parser.add_option("-p", "--pdb",
                action="store",
                type="string",
                dest="pdb",
345             help="use to specify
                an additional actional
                    pdb file, will be
                        superimposed with
                            hsc rendering")
            parser.add_option("-d", "--dssp",
                action="store",
                type="string",
                dest="dssp",
350             help="use to specify
                a dssp logfile to read
                    sec. structure information")
            (options, args) = parser.parse_args()
355             #print out a little header:
            print ""

    print "Running dammin2dino..."
    print ""
    if options.input:
        inputfile = options.input
        print "specified input file: "
            + inputfile
    else:
        inputfile = "damfilt.pdb"
        print "as input file '" + inputfile +
            "' will be tried (default)."
        inputfile = inputfile
    if options.output:
        outputfile = options.output
        print "specified output file prefix: "
            + outputfile
    else:
        outputfile = 'dino.scr'
        print "no output file prefix specified,
            using default file name
            ('dino.scr')"
    if options.surface:
        surface = options.surface
        print "specified surface file name
            (without extension!): " + surface
    else:
        surface = None
        print "no surface file specified"
    if options.ball_color:
        ball_color = options.ball_color
        print "specified ball color = %s" %
            ball_color
    else:
        ball_color = 'red'
        print "ball color = %s (default)" %
            ball_color
    if options.background:
        background = options.background
        print "specified background color: " +
            background
    else:
        background = 'white'
        print "background color is " +
            background + " (default)"
    if options.radius:
        radius = options.radius
        print "specified radius = %s" % radius
    else:
        radius = '8'
        print "ball radius = %s" % radius
    if options.pdb:
        pdbfile = options.pdb
        print "trying to read in %s, too" %
            pdbfile
    else:
        pdbfile = None
    if options.dssp:
        dssp = options.dssp
        print "retrieving secondary structure
            information from %s" % dssp
    else:
        dssp = None
    likelihood_flag = options.likelihood_flag
    option_dict =
        {'likelihood_flag': likelihood_flag,
         'inputfile': inputfile, 'outputfile':
         outputfile, 'surface': surface,
         'background': background, 'ball_color':
         ball_color, 'radius': radius, 'pdbfile':
         pdbfile, 'dssp': dssp}
    #call the actual functional function
    script_writer(**option_dict)
    if __name__ == "__main__":
        os.nice(1)
        main()

```

Skript A.1.3 erzeugt eine Eingabedatei für das Programm Dino (Philippsen, 2001).

## A.2 Evaluation von parmscan-Läufen

Zusätzlich zu den Bedienskripten zu Programmen der *ab initio*-Modellierung wurde auch ein Skript zur vereinfachten Auswertung von **parmscan**-Logdateien (zu **parmscan** siehe Abschnitt 3.6.11, Seite 39) geschrieben. Das Skript liest die Logdateien ein und gibt die Bewegungsparameter, die zu Konformationen führten, welche am besten mit den experimentellen SAXS-Daten übereinstimmten, aus. **parmscan** gibt für jede Konformation die zugrundeliegenden Bewegungsparameter und die Gütefaktoren für die Übereinstimmung der theoretischen, auf Basis der Konformation errechneten,  $P(r)$  und  $I(q)$  zu den jeweiligen experimentellen Daten in die Logdatei aus. Das Evaluationskript zeigt für jede Konformation beide Gütefaktoren an.

Das Skript ist auf der Begleit-CD zu dieser Dissertation enthalten.

## A.3 MODELLER Skripte

Die folgende Serie kleiner Skripte gibt die wesentliche Prozedur des Modellierens mit **MODELLER** ab Version 9.1 (Sali & Blundell, 1993; Martí-Renom *et al.*, 2000) wieder, nachdem die Software auf ein Python-Interface (van Rossum, 1998) umgestellt wurde. Diese Skripte entbinden nicht vom aufmerksamen Lesen der **MODELLER**-Anleitung, liefern jedoch subtile Änderungen im Pythoncode, der in den Beispielen von **MODELLER** relativ amateurhaft zusammengesetzt ist. Neben den Kommentaren befinden sich auch auskommentierte Codezeilen in den Beispielen, die optionale Erweiterungen des Codes darstellen.

Die Skripte erscheinen in der Reihenfolge, wie sie tatsächlich auch für das Erstellen der Modelle verwendet wurden.

Das erste Skript zeigt wie man das Profil für zwei Moleküle, hier die Mutante H35R und den Wildtyp des  $\alpha$ -Hämolytins, bildet. Dieses Alignmentprofil wird intern von **MODELLER** benötigt, sofern keines per Hand erstellt wird<sup>2</sup>.

---

### Skript 4 Generierung des Alignmentprofils für **MODELLER**

---

```

# import standard modules
import os
3 # import Modeller's modules
from modeller import log, environ
from modeller import sequence_db, alignment

if __name__ == "__main__":
8     #-- since this script takes cpu power
    os.nice(10)

    log.verbose()
    env = environ()
13
    #-- Prepare the input files

    #-- Read in the sequence database
    sdb = sequence_db(env)
18 #sdb.read(seq_database_file='pdball.pir',
    #       seq_database_format='PIR',
    #       chains_list='ALL',
    #       minmax_db_seq_len=(30, 4000),
    #       clean_sequences=True)
23
    #-- Write the sequence database
    # in binary form
    #sdb.write(seq_database_file='pdball.bin',
28 #         seq_database_format='BINARY',
    #         chains_list='ALL')

    #-- Now, read in the binary database
    sdb.read(seq_database_file='pdball.bin',
33
    seq_database_format='BINARY',
    chains_list='ALL')

    #-- Read in the target sequence/alignments
    aln = alignment(env)
    aln.append(file='wt.pir',
38             alignment_format='PIR',
             align_codes='ALL')
    aln.append(file='H35R.pir',
             alignment_format='PIR',
             align_codes='ALL')

    #-- Convert the input sequence/alignment
    # into profile format
    prf = aln.to_profile()
    #-- Scan sequence database to pick
    # up homologous sequences
    prf.build(sdb, matrix_offset=-450,
43             rr_file='${LIB}/blosum62.sim.mat',
             gap_penalties_id=(-500, -50),
             n_prof_iterations=1,
             check_profile=False,
             max_aln_evalue=0.01)

    #-- write to output file
    prf.write(file='build_profile.prf')
    #--convert back to alignment format
    aln = prf.to_alignment()

    aln.write(file='build_profile.ali',
58             alignment_format='PIR')

```

---

Das folgende Skript erstellt einen Vergleich zwischen allen möglichen Kandidaten für ein Muster („*Template*“ in der **MODELLER**-Sprache) für das strukturelle Modellieren. Die `align_codes` in Zeile 14 des Beispiels stehen für die PDB-Codes der besten Modelle, die im ersten Skript in der Datei `build_profile.prf` abgelegt wurden.

<sup>2</sup>Das Alingment muss auf jedem Fall überprüft werden. Bei komplexen Sequenzen empfiehlt es sich die Aufgabe des Alignments nicht **MODELLER**, sondern beispielsweise mit **ClustalX** (Thompson *et al.*, 1994) durchführen zu lassen und ggf. auch per Hand zu verbessern (Schweikardt *et al.*, 2004).

**Skript 5** *Analyse möglicher Musterstrukturen.*

```

# import standard modules
import os
3
# import Modeller's modules
from modeller import log, environ, alignment
8
if __name__ == "__main__":
    os.nice(19)
    log.verbose()
    #create enviroment
env = environ()
aln = alignment(env, file='pdball.pir',
                align_codes=('11kfA', '21kfA',
                            '31kfA', '1pv1', '1t5rA'))
aln.malign()
aln.malign3d()
aln.compare_structures()
aln.id_table(matrix_file='family.mat')
env.dendrogram(matrix_file='family.mat',
               cluster_cut=-1.0)

```

Das nächste Skript zeigt wieder ein Alignment: Ein zweidimensionales Alignment, das von MODELLER benötigt wird. Die beiden Sequenzen (die Mutante H35R und den Wildtyp des  $\alpha$ -Hämolytins) wurden gegen LukF (PDB-Code 1lkf) ausgerichtet.

**Skript 6** *Ausführen eines zweidimensionalen Alignments in MODELLER.*

```

import os
# import Modeller's modules
from modeller import environ
5 from modeller import alignment, model, log
if __name__ == "__main__":
    os.nice(1)
    log.verbose()
10 #creating enviroment
    env = environ()
    aln = alignment(env)
15 mdl = model(env, file='1lkf')
    #optional append to model parameters:
    # , model_segment=('FIRST:A', 'LAST:A')
aln.append_model(mdl, align_codes='11kfA',
                 atom_files='1lkf.pdb')
20 aln.append(file='wt.pir',
            align_codes='7ah1A')
5 aln.append(file='H35R.pir',
            align_codes='7ah1M')
10 aln.align2d()
    aln.write(file='wt_1lkf.ali',
             alignment_format='PIR')
    aln.write(file='wt_1lkf.pap',
             alignment_format='PAP')
15 aln.write(file='h35r_1lkf.ali',
            alignment_format='PIR')
    aln.write(file='h35r_1lkf.pap',
            alignment_format='PAP')

```

Nun folgt das eigentliche Erstellen der Modelle. Ebenfalls für beide Sequenzen. Das `del(a)` ist in normalem Pythoncode nicht zwingend notwendig, stellt hier aber sicher, dass die Referenzen zum `automodel`-Objekt zerstört sind. Dieses wird in manchen Codes mit MODELLER-Importen nicht automatisch durchgeführt und kann zu einer erheblichen Störung der Evaluierung der zweiten Sequenz führen (hier: `h35r_1lkf.ali`). Es ist hierbei irrelevant an welchen Namen das `automodel`-Objekt gebunden wird.

**Skript 7** *Erstellen der Modelle.* Bei tatsächlichen Skripten sind die Zahlen für die Anzahl der zu generierenden Modelle anzupassen.

```

import os
2 from modeller import environ, refine
from modeller.automodel import automodel
17
if __name__ == "__main__":
    os.nice(1)
    env = environ()
    a = automodel(env,
                 alnfile='wt_1lkf.ali',
                 knowns='11kfA',
                 sequence='7ah1A')
7
12 a.starting_model = 1
    a.ending_model = 5
a.md_level = refine.very_slow
a.make()
17 del(a)
a = automodel(env,
              alnfile='h35r_1lkf.ali',
              knowns='11kfA',
              sequence='7ah1M')
22 a.starting_model = 1
    a.ending_model = 5
    a.md_level = refine.very_slow
    a.make()

```

An die Erstellung der Modelle schließt sich die Evaluierung der Modelle an: Es gilt eine Einschätzung der Qualität zu finden. Diese gelingt über eine Einschätzung des DOPE-Wertes (siehe 3.8.1). An dieser Stelle wurden die Beispiele aus der MODELLER-Anleitung wesentlich erweitert.

---

### Skript 8 *Evaluation der Modelle über ihren DOPE-Wert.*

---

```

import os
import glob

# import Modeller's modules
5 from modeller import environ, selection, log
from modeller.scripts import complete_pdb

if __name__ == "__main__":
    os.nice(19)

    # request verbose output
    log.verbose()
    env = environ()
    # read topology
15 env.libs.topology.read(
        file='${LIB}/top_heav.lib')
    # read parameters
    env.libs.parameters.read(
        file='${LIB}/par.lib')

    # read model files
    models = list()
    names = glob.glob('*B9999*.pdb')
    for name in names:
25 models.append(complete_pdb(env, name))
    # Assess with DOPE:
    report = ''
    for model, name in zip(models, names):
30 # don't consider in case of
        # a rotational fit

        if 'fit' in name: continue
        # all atom selection
        s = selection(model)
        dope = s.assess_dope(
            output='ENERGY_PROFILE NO_REPORT',
            file='7ahl.profile',
            normalize_profile=True,
            smoothing_window=15)
        report += 'DOPE score for %s: %f\n' %
            (name, dope)

        # read model file
        mdl2 = complete_pdb(env, '1lkf.pdb')

        # Assess with DOPE:
        s = selection(mdl2) # all atom selection
        dope = s.assess_dope(
            output='ENERGY_PROFILE NO_REPORT',
            file='1lkf.profile',
            normalize_profile=True,
            smoothing_window=15)
        report += 'DOPE score for 1lkf: %f\n' % dope

        # open file for a report
        try:
            f = open('DOPE_report.log', 'w')
        except IOError, msg:
            raise IOError(msg)
        f.write(report)
        f.close()

```

---

#### A.3.1 Loop Modelling

Modellierung von Schleifen („*Loop Modelling*“) wird üblicherweise verwendet, um Lücken in einem Modell sinnvoll zu schließen. Solche Lücken sind meist auf eine „verschmierte“ Elektronendichte aus kristallographischen Daten zurückzuführen. In der vorliegenden Arbeit wurde die Methode auf Bereiche des  $\alpha$ -Hämolybins angewendet, um eine bessere Übereinstimmung mit den SAXS-Daten zu erreichen. Die Methode von Fiser *et al.*, 2000, ist in MODELLER implementiert und fand auf folgende Weise Anwendung:

---

### Skript 9 *Loop Modelling in MODELLER.*

---

```

import os

from modeller import log, environ,
5 selection, forms, physical,
features, secondary_structure
from modeller.automodel import loopmodel,
refine

# Create a new class based on 'loopmodel'
# so that we can redefine
# select_loop_atoms (mandatory!)
class hemolysin(loopmodel):
    def special_restraints(self, aln):
15 rsr = self.restraints
        at = self.atoms

        # loop residues to be beta strand
        rsr.add(secondary_structure.strand(
            self.residue_range('147:', '148:')))
20 rsr.add(secondary_structure.strand(
            self.residue_range('164:', '165:')))
        # This routine picks the residues to be
        # refined by loop modeling
    def select_loop_atoms(self):
25 # One loop from residue 147 to

        # 28 inclusive
        return selection(
            self.residue_range('147:', '164:'))

30 if __name__ == '__main__':
    os.nice(19)

    log.verbose()
    env = environ()

    # directories for input atom files
    env.io.atom_files_directory =
        '.././atom_files'

    m = hemolysin(env,
        inmodel='7ahlA.B99990015.pdb',
        sequence='../7ahlA')
    # index of the first loop model
    m.loop.starting_model = 1
    # index of the last loop model
    m.loop.ending_model = 20
    # loop refinement method
    m.loop.md_level = refine.slow_large

50 m.make()

```

---

Wichtig ist vor allem die **refinement method** auf **slow\_large** zu setzen, da sonst die Ausbeute physikalisch sinnvoller Modelle sehr gering und vom Zufall bestimmt ist. Zudem sind auch die flankierenden Sekundärstrukturelemente als Randbedingungen in **special\_restraints()** aufgenommen. Die Methoden der **hemolysin**-Klasse werden bei der Initialisierung der Klasse automatisch aufgerufen. Dies erledigt MODELLER im Hintergrund ohne das Vorhandensein einer pythontypischen **\_\_init\_\_**-Methode.

## A.4 Vergleich von Elektronenmikroskopie und SAXS-Daten

Die Berechnung theoretischer Streukurven von Single-Particle-Modellen erfolgte entweder mit `em2p2i` von H. Hartmann oder mittels des selbstgeschriebenen `EM2SAXS` (siehe Abschnitt 3.6.9, Seite 37). `EM2SAXS` ist im engeren Sinne kein Programm, sondern eine in Python ([www.python.org](http://www.python.org)) geschriebene Bibliothek. Da Python eine interpretierte Sprache darstellt, sind Programme, die `EM2SAXS` nutzen, zwar langsamer als `em2p2i`, dafür bietet `EM2SAXS` jedoch einige andere Vorteile:

- Die Bibliothek nutzt die Python Imaging Library ([www.pil.org](http://www.pil.org)) und ist daher in der Lage verschiedene Formate (z. B. ein Paket von `.tiff`-Dateien oder eine serielle Darstellung der EM-Rekonstruktion als Rohdatenformat) einzulesen.
- Es ist möglich die Streuung ab einem bestimmten Schwellenwert zu berechnen, genauso wie es möglich ist einen bestimmten Schwellenwert abzuziehen.

Der Code zu diesem Skript ist bereits im „*subversion repository*“ <http://subversion.tigris.org/> zu `sas-rigid` (siehe Abschnitt 4.7, Seite 105) archiviert.

## A.5 Umgang mit Pipealign

`Pipealign` (Plewniak *et al.*, 2003) kann verwendet werden, um die Arbeit mit Sequenzdatenbanken zu erleichtern: Das Webinterface verwendet `Ballast` (Plewniak *et al.*, 2000) (ein Programm zum Nachbearbeiten von `BLAST`-Ergebnissen (Altschul *et al.*, 1997)), `DBClustal` (ein Datenbankbasiertes `Clustal`-Pendant (Thompson *et al.*, 1994)), `Rascal` (Thompson *et al.*, 2003) (ein Programm zur automatischen Korrektur von `Clustal`-Alignments), `Leon` (Thompson *et al.*, 2004) (ein Programm, welches Nachbarn innerhalb von Alignments suchen kann) und `Secator` (Wicker *et al.*, 2001; Wicker *et al.*, 2002) (zum Bau phylogenetischer Bäume). Dies stellt einen klaren Fortschritt gegenüber der manuellen Suche, Alignments und Gruppierung vergangener Jahre dar. Die Schwäche von `Pipealign` ist jedoch, dass im finalen Alignment doppelte Einträge oder teilweise überlappende Einträge (weil ein Protein beispielsweise der Vorläufer eines anderen Proteines ist) enthalten sein können – dies ist sicherlich für viele Datenbanksuchen sinnvoll, nicht jedoch, wenn aufgrund der Resultate der Konservierungsgrad der Sequenzen oder ähnliche Dinge berechnet werden sollen. Hier nämlich täuschen doppelte Einträge eine übermäßig starke Konservierung einer Proteinfamilie vor, die unter Umständen gar nicht vorherrscht<sup>3</sup>.

Daher habe ich mir ein eigenes Programm geschrieben. Dieses durchsucht das Ergebnis von `Pipealign`, entfernt doppelte Einträge, berechnet den Konservierungsgrad des Musterproteins gegenüber der Proteinfamilie für jede Sequenzposition nach Schneider & Stephens, 1990 und wandelt diese numerische Analyse in Ausgaben um, die mit `Dino` (Philippsen, 2001) oder einem Plot dargestellt werden können:

```

#!/usr/bin/python                25
from optparse import OptionParser
import os
5 import pylab
from math import log            30
import numpy

__author__ = "Christian Meesters"
__version__ = "0.1"
__date__ = "4. July 2007"      35

class InputError(Exception):
    """if user gives wrong input"""
15 def __init__(self, value):
    Exception.__init__(self, value)
    self.value = value
    def __str__(self):
        return repr(self.value)
20 def __SavitzkyGolaySmoothing(data,
    kernel=11, order=4):
    """
        applies a Savitzky-Golay filter
                                45
                                input parameters:
                                - data => data as a 1D numpy array
                                - kernel => a positiv integer >
                                  2*order giving the kernel size
                                - order => order of the polynomial
                                returns smoothed data as a numpy array

                                invoke like:
                                smoothed = _savitzky_golay(<rough>,
                                  [kernel]=value], [order= value]
                                """
                                try:
                                    kernel = abs(int(kernel))
                                    order = abs(int(order))
                                except ValueError, msg:
                                    raise ValueError("kernel and order have
                                        to be of type int (floats will be
                                        converted).")
                                if kernel % 2 != 1 or kernel < 1:
                                    raise TypeError("kernel size must be a
                                        positive odd number, was: %d" %
                                        kernel)
                                if kernel < order + 2:
                                    raise TypeError("kernel is to small for

```

<sup>3</sup>Dennoch kann natürlich das Arbeiten mit doppelten Einträgen im Einzelfall sinnvoll sein, z. B. wenn diese doppelten Einträge von verschiedenen Organismen stammen. Der Ansatz, der mit `Pipealign` verfolgt wird, ist also nicht falsch.

```

50         the polynomials\should be >
           order + 2")
155
# a second order polynomial has 3
# coefficients
order_range = range(order+1)
55 half_window = (kernel -1) // 2
b = numpy.mat([[k**i for i in order_range]
               for k in range(-half_window,
                               half_window+1)])
# since we don't want the derivative,
60     else choose [1] or [2], respectively
m = numpy.linalg.pinv(b).A[0]
window_size = len(m)
half_window = (window_size-1) // 2
65
# precompute the offset values for
# better performance
offsets = range(-half_window, half_window+1)
offset_data = zip(offsets, m)
70
smooth_data = list()
# temporary data, with padded values for
# the 'end values' (since we want the
# same length after smoothing)
75 start = numpy.linspace(numpy.mean(data[:3]),
                          numpy.mean(data[:3]), half_window)
end = numpy.linspace(numpy.mean(data[-3:]),
                     numpy.mean(data[-3:]), half_window)
data = numpy.concatenate((start, data, end))
80 for i in range(half_window, len(data) -
                  half_window):
    value = 0.0
    for offset, weight in offset_data:
        value += weight * data[i + offset]
85 return numpy.array(smooth_data)

def read_sequences(**option_dict):
90 """
    will read in sequences from a file in
    FASTA format and put them in a directory
    as {identifier: sequence}

    invoke like:
95 read_sequences(<**option_dict>)
"""
try:
    input = open(option_dict['inputfile'], 'rU')
100 except IOError, msg:
    raise IOError(msg)

sequences = dict()
105 identifier = None
sequence = ''

for line in input:
    # identifier line hit?
    if line.startswith('>'):
        # already sequence in memory?
110         if identifier:
            sequences[identifier] = sequence
            identifier = line.rstrip('\n').rstrip('>')
            sequence = '' # erase sequence
        else:
115         sequence += line.rstrip('\n')

input.close()

# since it might happen that empty sequences
# are read in named 'group', these entries
# should be deleted
120 for key in sequences.keys():
    if not(len(sequences[key].strip('-'))):
        del sequences[key]

125 if option_dict['verbose']:
    print "Found a total
          of %d sequences" % len(sequences)
return sequences

130 def check_doubles(sequences, **option_dict):
"""
    find those sequences which occur more
    than once
135 """
returns dict{identifier:[identifiers]}

doubles = dict()

140 interns = option_dict['containing']

for key in sorted(sequences):
    for key2 in sorted(sequences):
        if interns and sequences[key].rstrip('-')
145         in sequences[key2] and key != key2:
            if doubles.has_key(key):
                doubles[key].append(key2)
            else:
                doubles[key] = [key2]
        elif not interns and sequences[key] ==
150         sequences[key2] and key != key2:
            if doubles.has_key(key):
                doubles[key].append(key2)
            else:
                doubles[key] = [key2]

155 if option_dict['verbose']:
    print "Found the following
          sequences as doubles in sequence:"
    for count, key in
160         enumerate(sorted(doubles.keys())):
            print "#%4d: %s is equal to %s" %
                  (count+1, key, doubles[key])

return doubles

165 def compare_entries(sequences, query1, query2,
**option_dict):
"""
    compare two queries with each other:
    prints those sequence positions which
    are identical

    invoke like:
170 compare_entries(sequences, query1,
                  query2, **option_dict)
    , where query1 and 2 are strings
"""
number = 0
175 for seq1, seq2 in zip(sequences[query1],
                        sequences[query2]):
    if seq1 != '-':
        number += 1
        if seq1 == seq2:
            print "%d: %s == %s" %
180                 (number, seq1, seq2)

def delete_sequences(sequences, doubles,
**option_dict):
185 """
    delete those sequences which occur
    doubled, returns cleared list of
    sequences
"""
190 if option_dict['verbose']:
    print "deleting entries"
# for security reasons ...
copy = sequences.copy()

done = list()
195 for key in sorted(sequences.keys()):
    if doubles.has_key(key):
        try:
            for entry in doubles[key]:
                done.append(entry)
                if not entry in done:
                    del copy[entry]
                    if option_dict['verbose']:
                        print "deleting " + entry + " ",
                    except: pass
200 if option_dict['verbose']:
    print "\nafter deleting double entries
          %d sequences are left" % len(copy)
return copy

205 def save_file(sequences, **option_dict):
"""
    save file in FASTA format
"""
210 if option_dict['verbose']:
    print "saving sequences to %s" %
          option_dict['outputfile']

try:
    out = open(option_dict['outputfile'], 'w')
215 except TypeError:
    raise TypeError("No file name specified")
except IOError, msg:
    raise IOError(msg)

for key in sequences.keys():
    out.write('>%s\n' % key)
    count = 50
    for x in sequences[key]:
        out.write(x)
        count -= 1
        if count == 0:
            count = 50
            out.write('\n')
220 out.write('\n')
out.close()

225 def calculate_entropy(sequences, **option_dict):
"""
    returns the 'entropy' as list of
    [actual position in the alignment,
    Amino acid,
    entropy value]
230 """
if option_dict['verbose']:
    print "Calculating entropies for
          positions in the query"
# get the query sequence again
raw_sequences = read_sequences(**option_dict)
query = raw_sequences['query']
# frequencies at a given position in query
frequencies = list()

235 for pos, letter in enumerate(query):
    # total number of occupied AAs
    total = 0

```

```

260     # total of AAs equal to query at pos
        occ = 0
        if letter != '-': # no gap?
            for key in sequences.keys():
                value = sequences[key][pos]
                if value.isalpha():
265                     total += 1
                     if value == letter:
                         occ += 1
                     frequencies.append((pos + 1, letter,
270                                     float(occ) / float(total)))

# calculate information content according to
# Schneider & Stephens (1990):
smax = log(20, 2) # 20 possible AAs
entropy = list()
275 for AA in frequencies:
    try:
        entropy.append((AA[0], AA[1],
                        (smax - (-1 * AA[2] * log(AA[2], 2)))
                        * AA[2]))
280    except OverflowError: # log(0) ?
        entropy.append((AA[0], AA[1], 0))
    return entropy

def write_dino_color_scr(entropy, **option_dict):
285     """
    """
    # first: find maximum & minimum entropy
    maximum, minimum = -1e99, 1e99 # dummy values
290     for pos in entropy:
        if pos[2] > maximum: maximum = pos[2]
        if pos[2] < minimum: minimum = pos[2]
        if abs(maximum - minimum) < 1.0:
            raise AssertionError("Values for the
295                 entropy are too close for sensible
                    division into histograms")
    # next: find increment for equal binning
    increment = (maximum - minimum) /
300                (option_dict['bins'] - 1)
    # next: create histogram
    histo = [list() for x in
305                range(option_dict['bins'])]
    for pos, aa in enumerate(entropy):
        step = int(aa[2] / increment)
        histo[step-1].append(pos + 1)
    # next: write DINO color_scr-file
    try:
        color = open('color_scr', 'w')
310    except IOError, msg:
        raise IOError, msg
    for pos in range(len(histo)):
        tmp = str(pos+1).zfill(2)
        color.write('set c%$ rnum=' % tmp)
315        for number in histo[pos]:
            if number == histo[pos][-1]:
                color.write('%d\n' % number)
            else:
                color.write('%d,' % number)
        color.write('\n//-----\n')
320    # write ca-selection code
    # color intensity increment
    increment = 1.0 / (option_dict['bins'] - 1)
    color.write('. dflt. ca set color=white\n\n')
    for sel in range(len(histo)):
325        intensity = sel * increment
        color.write('. dflt. c
                        set color={%.2f,%.2f,1.0}
                        -sel $c' % (intensity, intensity))
        tmp = '%d' % (sel+1)
330        color.write(tmp.zfill(2))
        color.write('\n')
        color.write('\n//-----\n')
    # repeat with surface
    color.write('. surface. all
335                set color=white\n\n')
    for sel in range(len(histo)):
        intensity = sel * increment
        color.write('. surface. all
                        set color={%.2f,%.2f,1.0}
                        -sel $c' % (intensity, intensity))
340        tmp = '%d' % (sel+1)
        color.write(tmp.zfill(2))
        color.write('\n')
        color.close()
345 def plot_entropies(entropy, **option_dict):
    """
    """
    # first: find maximum & minimum entropy
    maximum, minimum = -1e99, 1e99 # dummy values
350     for pos in entropy:
        if pos[2] > maximum: maximum = pos[2]
        if pos[2] < minimum: minimum = pos[2]

355     # next: find increment for equal binning
    increment = (maximum - minimum) /
                (option_dict['bins'] - 1)
    positions, ent = list(), list()
    for pos, aa in enumerate(entropy):
360        step = int(aa[2] / increment)
        positions.append(pos+1)
        ent.append(step)

        smooth_ent = __SavitzkyGolaySmoothing(ent,
365             kernel=21)
        #since the smoothed data can drop below zero
        smooth_ent = numpy.where(smooth_ent < 0,
            0, smooth_ent)
        # scale between 0 and 4.32 (max. info content)
370        m = max(max(ent), max(smooth_ent))
        ent = [(x/float(m))*log(20, 2) for x in ent]
        smooth_ent = [(x/float(m))*log(20, 2)
            for x in smooth_ent]
        pylab.plot(positions, smooth_ent, 'k-',
375             linewidth=1.5)
        pylab.plot(positions, ent, 'k-')
        pylab.xlim(xmin=1, xmax=len(positions))
        pylab.ylim(ymin=0, ymax=log(20, 2))

#amino terminal latch
#pylab.axvspan(1, 49, ymax=log(20, 2),
380             facecolor='gray', alpha=0.4)
#transmembrane part
#pylab.axvspan(104, 151, ymax=log(20, 2),
385             facecolor='gray', alpha=0.4)

#individual sheets
#pylab.axvspan(3, 4, ymax=log(20, 2),
390             facecolor='yellow', alpha=0.4)
# same for other such regions ...

#choline binding site
pylab.axvspan(175, 200, ymax=log(20, 2),
395             facecolor='red', alpha=0.4)

# contact sides
pylab.axvspan(1, 23, ymax=log(20, 2),
400             facecolor='yellow', alpha=0.4)
# same for other such regions ...

pylab.xlabel('residue number')
pylab.ylabel('bits')
pylab.show()
405 def main():
    """
    """
    parser = OptionParser(usage=__doc__)
    parser.add_option("-i", "--input",
410                    action="store",
                    type="string",
                    dest="input",
                    help="specify input file
                        (default = 'input.aln')")
    parser.add_option("-o", "--output",
415                    action="store",
                    type="string",
                    dest="output",
                    help="specify output file
                        (default = no output file)")
    parser.add_option("-v", "--verbose",
420                    action="store_true",
                    dest="verbose",
                    help="verbosity flag")
    parser.add_option("-c", "--containing",
425                    action="store_true",
                    dest="containing",
                    help="not only find exact double
                        entries, but also those
                        which contain others")
    parser.add_option("-b", "--bins",
430                    action="store",
                    type="int",
                    dest="bins",
                    help="define number of bins
                        for color table histogram,
                        defaults to 10")
    (options, args) = parser.parse_args()
435
    if options.input:
        inputfile = options.input
        print "specified input file: %s" %
440            inputfile
    else:
        inputfile = "input.aln"
        print "looking for 'input.aln' as
445            the alignment file"

    if options.output:
        outputfile = options.output
        print "Output will be directed to : %s"
450            % outputfile
    else:
        outputfile = 'default.aln'

    if options.bins:
        bins = options.bins
        if bins > 99:
455            raise InputError("More than 99
                                cannot be supported")
    else:
        bins = 10
    print "Color table will be divided
460        into %d bins." % bins

    print "Verbose mode: ", options.verbose

    option_dict = {'inputfile': inputfile,

```

```

470     'outfile': outfile,
        'verbose': options.verbose,
        'containing': options.containing,
        'bins': bins}

475     # read in all sequences
    sequences = read_sequences(**option_dict)
    # check those with a double occurrence
    doubles = check_doubles(sequences,
                            **option_dict)
480     # delete those which occur more than once
    sequences = delete_sequences(sequences,
                                doubles, **option_dict)
    # save to alignment file
    save_file(sequences, **option_dict)

485     # calculate the information content
    # relative to the query
    entropy = calculate_entropy(sequences,
                                **option_dict)
    # write DINO color.scr file
    write_dino_color_scr(entropy, **option_dict)
490     # generate plot of entropies
    plot_entropies(entropy, **option_dict)
    #
    compare_entries(sequences, 'PDB|7ahl_A',
                   'PDB|1l1kf_A')
495     if __name__ == "__main__":
        os.nice(1)
        main()

```

Nicht alles in diesem Programm ist automatisiert: So müssen z.B. Einträge in der Funktion `plot_entropies` per Hand korrigiert, beziehungsweise angepasst werden.

## A.6 PreToExcel

Das Programm `PreToExcel` ist verfasst worden, um Streudaten der Kratky-Kamera des Institut für Molekulare Biophysik schneller weiterverarbeiten zu können. Es transferiert Daten aus der `ASA`-Software (Version 2.3) in das Dateiformat `EXCEL95`, das mit `EXCEL` (Microsoft Corp., Redmont, WA; bis Version: `EXCEL2003`, SP3) lesbar ist (siehe Abschnitt 3.6.4.1, Seite 34). Mit `OpenOffice` ([www.openoffice.org](http://www.openoffice.org); Version  $\geq 2.2$ ) ist eine Verarbeitung der so transferierten Daten ebenfalls möglich.

`PreToExcel` ist in Python (van Rossum, 1998) geschrieben und auf verschiedenen Rechnern im Institut für Molekulare Biophysik installiert. Zusammen mit `SPlot` (siehe Abschnitt 4.8, Seite 117) stellt `PreToExcel` eine wesentliche Arbeitserleichterung im Umgang mit den SAXS-Daten der Kratky-Kamera dar. Da der Quellcode für diesen Anhang zu umfangreich ist, ist er lediglich auf der CD mit den Daten zu dieser Arbeit enthalten.

## Anhang B

# Hinweise für Betrieb und Wartung der Kratky-Kamera und des Zubehörs

Zweck dieses Anhangs ist es Besonderheiten und Fallstricke im Umgang mit der Kratky-Kamera zu dokumentieren, so dass auch eventuelle Nachfolger an diesem Gerät leichten Zugang finden.

### B.1 Besonderheiten bei der Wartung

- Das Peltierelement der temperierbaren Kapillare muss in jedem Fall gegengekühlt werden, auch dann, wenn es selber kühlen soll. Das Wasserbad ist hierzu auf eine Temperatur von 10 °C gestellt worden, was aus Sicherheitsgründen am unteren Rand der vorgegebenen 10 – 12,5 °C liegt. Es ist darauf zu achten, dass die Wasserleitungen stets frei bleiben und sich nicht zusetzen. Falls sich doch eine Leitung zusetzte, hat es sich als günstiger herausgestellt, die Fehlersuche vom Wasserbad ausgehend zu starten (auch das Ventil, welches die Flusstärke steuert, kann sich verstellen)<sup>1</sup>.
- Auch das Wasser des Wärmeaustauschers muss regelmäßig ausgetauscht werden. Es ist darauf zu achten, dass auch der Filter gut gesäubert wird. Hierzu wurde der Wärmeaustauscher aufgebockt, so dass das Wasser einfach abzulassen ist.
- Das Öl der Vakuumpumpe ist in regelmäßigen Abständen zu wechseln (siehe Handbuch der Pumpe). Zwischen Pumpe und Kamera ist der Zeolith-Ölfilter installiert. Es ist spätestens bei jedem Ölwechsel darauf zu achten, dass kein Öl in den Filter gelangt ist. Dies ist ein Zeichen für einen Defekt des Rückschlagventils an der Pumpe. In diesem Fall kann Öl in die Kamera dringen, was in jedem Fall zu vermeiden ist. Es gilt dann das Ventil und das Zeolith im Filter auszuwechseln, bevor an einen Weiterbetrieb der Anlage zu denken ist. Der Ölfilter bedarf in diesem Fall zusätzlich einer Grundreinigung mit Benzin.

### B.2 Proben in die Kapillare einspritzen und entnehmen

Die in Abschnitt 3.6.2 auf Seite 33 beschriebene Vorgehensweise beim Einfüllen und Entnehmen der Proben in die Kapillare der Kratky-Kamera verlangt manchmal erhebliches Fingerspitzengefühl. Es sollte daher mit Wasser geübt werden – insbesondere bei „wertvollen“ Proben, die nur in geringem Volumen vorliegen. Proben mit veringertener Oberflächenspannung (aufgrund von Detergenzien oder besonders amphiphilen Proteinen) verlangen ebenfalls besondere Vorsicht.

---

<sup>1</sup>Keinesfalls versuchen die kleinen Schläuche abzulösen, in dem man die Adapter abschraubt: Die Verbindungen sind geklebt und brechen ab! Besser ist es den Halter aus der Kamera zu schrauben und dann mit einer Zange die Schläuche abzuziehen. **Andernfalls kann die Kamera Schaden nehmen!**

# Anhang C

## Bisherige Publikationen

### C.1 Originalarbeiten

Die während der Zeit meiner Doktorarbeit veröffentlichten oder in Veröffentlichung befindlichen Originalarbeiten:

- Shareghe Mehraeen, Henning Stahlberg, James E. Evans, *Christian Meesters*, Ilke Arslan, Paul A. Midgley, and Nigel D. Browning (2005) Testing 3-D Reconstruction Methods for Electron Tomography (*Microscopy and Microanalysis* **11**(2), 336–337)

Diese Arbeit stellt den Startpunkt der Doktorarbeit von S. Mehraeen dar. Meine Aufgabe war es die statistische Analyse von  $\text{Sn}_x\text{Si}_{1-x}$  Quantum Dots eingebettet in einer mesoporen Silicon Matrix zu programmieren. Hierzu wurden von mir diverse Skripte in SPIDER (Frank *et al.*, 1981) sowie Python (van Rossum, 1998) erstellt und S. Mehraeen in den Umgang mit diesen Skripten eingewiesen. – Diese Publikation ist **nicht** Teil meiner Dissertation.

- Shirley A. Müller, Charalambos Pozidis, Remington Stone, *Christian Meesters*, Mohamed Chami, Andreas Engel, Anastassios Economou, and Henning Stahlberg (2006) Double hexameric ring assembly of the type III protein translocase ATPase HrcN (*Molecular Microbiology* **61**(1), 119–125)

Mein Beitrag zu dieser Publikation liegt in der bioinformatischen Auswertung der strukturellen und phylogenetischen Aspekte von HrcN. Hierzu wurden die Ergebnisse eines `pipealign`-Laufs (Plewniak *et al.*, 2003) statistisch ausgewertet und ähnlich wie unter 3.10.8 beschrieben graphisch auf ein mit SWISS-Model (Arnold *et al.*, 2006) erstelltes Strukturmodell projiziert. – Diese Publikation ist **nicht** Teil meiner Dissertation.

- Sharon Baird, Sharon M. Kelly, Nicholas C. Price, Elmar Jaenicke, *Christian Meesters*, Dorothea Nillius, Heinz Decker, and Jacqueline Nairns (2007) Hemocyanin conformational changes associated with SDS-induced phenol oxidase activation (*Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1774**(11), 1380–1394)

In dieser Arbeit wird die Stabilität von in SDS inkubierten Hämocyaninen nachgewiesen und mittels diverser spektroskopischer Methoden Indikationen für einen SDS-assoziierten Konformationswechsel gefunden. Mein Beitrag zu dieser Publikation ist es, Röntgenkleinwinkelstreudaten des Hämocyanins der Vogelspinne *Eurypelma californicum* mit und ohne SDS hinsichtlich eines Konformationswechsels zu interpretieren. Eine detaillierte Besprechung der eigenen Resultate findet sich in Kapitel 4.1.1 und der Diskussion (Kapitel 5.1).

- *Christian Meesters* (2008) A software Library for Monte Carlo-Based Rigid Body Modelling Against Small Angle Scattering Data (From Computational Biophysics to Systems Biology 2008 (40), 321–323)

Diese Arbeit wird intensiv in den Resultaten in Abschnitt 4.7 sowie der Diskussion in Abschnitt 5.10 behandelt.

- *Christian Meesters*, Antje Brack, Nadja Hellmann, and Heinz Decker (2008) Structural Characterization of the  $\alpha$ -Hemolysin Monomer from *Staphylococcus aureus* (*Proteins: Structure, Function, and Bioinformatics* – akzeptiert zur Veröffentlichung)

Diese Arbeit wird intensiv in den Resultaten unter Abschnitt 4.4 sowie der Diskussion in Abschnitt 5.9.1 behandelt.

- Elmar Jaenicke, Sebastian Fraune, Sandra May, Pinar Irmak, Rene Augustin, *Christian Meesters*, Heinz Decker, and Martin Zimmer (2008), Phenoloxidase is not expressed during the immune response in woodlice while their hemocyanin has phenoloxidase activity  
Zur Veröffentlichung eingereicht.

## C.2 Erwähnungen in Danksagungen

Neben Originalveröffentlichungen wurde meine Arbeit auch in folgenden Publikationen in Danksagungen für die verschiedensten Beiträge erwähnt:

- Charalambos Pozidis, Aggeliki Chalkiadaki, Amalia Gomez-Serrano, Henning Stahlberg, Ian Brown, Anastasia P. Tampakaki, Ariel Lustig, Giorgos Sianidis, Anastasia S. Politou, Andreas Engel, Nickolas J. Panopoulos, John Mansfield, Anthony P. Pugsley, Spyridoula Karamanou, and Anastassios Economou (2003) Type III Protein Translocase HrcN is a peripheral membrane ATPase that is activated by oligomerization (*The Journal of Biological Chemistry* **278**(28), 25816–25824)

Diese Arbeit ist die Vorläuferpublikation zu der oben genannten ersten Publikation zu HrcN (Müller *et al.*, 2006, siehe C.1). In dieser Arbeit ging es sowohl um die biochemische als auch die erste strukturbiochemische Charakterisierung von HrcN. Mein Anteil lag in der Hilfestellung für C. Pozidis. Im Zuge dessen habe ich geholfen die ersten elektronenmikroskopischen Aufnahmen zu machen und habe maßgeblich dazu beigetragen die zweite Publikation von 2006 zu ermöglichen: Eigentlich sollten die elektronenmikroskopischen Arbeiten, wegen anscheinender Aggregation des Proteins eingestellt werden, ich konnte beweisen, dass der Kontrast einzelner Partikel in dem „Scheinaggregat“ ausreichend für weiterführende Analysen ist.

- Elmar Krieger (1993–2008) YASARA - Software für Proteindarstellung, Modelling und Molekulardynamik

Elmar Kriegers Programm YASARA (Krieger *et al.*, 2002) habe ich während meiner Arbeit intensiv verwendet. Mit Hilfe von detaillierten Fehlermeldungen half ich das Programm zu verbessern. Überdies habe ich Herrn Krieger einige Änderungen am Pythoninterface des Programs vorschlagen können, was die Nutzerfreundlichkeit und Stabilität des Interfaces wesentlich erhöhte.

- Greg Wilson (2006–2008) Software Carpentry Course using Python

Wie Greg Wilson von der University of Toronto in einem Interview (<http://www.hpcwire.com/hpc/771596.html>) beschreibt, gibt es in den Augen vieler Wissenschaftler eine Diskrepanz zwischen den akkuraten Arbeiten im Labor und einer eher fahrlässigen Herangehensweise bei von durch Wissenschaftlern erstellten Softwareprogrammen. Die Reaktion hierauf ist spezieller Kurs, den er auch online stellte: <http://swc.scipy.org/>. Mein Beitrag waren diverse didaktische Hinweise und kurze „Codesnippets“.

- Dorothea Nillius, Elmar Jaenicke, and Heinz Decker (2008) Switch between tyrosinase and catecholoxidase activity of scorpion hemocyanin by allosteric effectors (*FEBS Letters*) **582**, 749–754

Diese Arbeit stellt wesentliche Ergebnisse der Dissertation von D. Nillius vor. Sie hat sämtliche biochemischen Arbeiten hierzu durchgeführt. Mein Anteil war die Anfertigung der elektronenmikroskopischen Aufnahme innerhalb dieser Publikation, deren Zweck darin liegt, die strukturelle Integrität des untersuchten Proteins unter SDS-Einfluss zu indizieren. Eine weitere Analyse der Daten wurde von mir – unter Berücksichtigung einer Kooperation mit der Gruppe von Wha Chiu in Houston – nicht vorgenommen.

## C.3 Veröffentlichte Software

Neben Originalveröffentlichungen wurde durch mich geschriebene Software auch an folgenden Orten veröffentlicht:

Webseite bzw. Organisation	URL	Titel / Inhalt	Art des Lektorates
Scipy (Scientific Python)	<a href="http://www.scipy.org/Cookbook/SavitzkyGolay">www.scipy.org/Cookbook/SavitzkyGolay</a>	Savitzky Golay Filtering of Data	Wiki (Gemeinschaftslektorat)
Deutsches Pythonforum	<a href="http://www.python-forum.de/topic-6891.html">www.python-forum.de/topic-6891.html</a>	Refactoring von Menüeinträgen in wxPython	Forum (Gemeinschaftslektorat)

## C.4 Reporte für Experimente an Großforschungsanlagen

Folgende Berichte für Experimente an Großforschungsanlagen wurden verfasst und sind über die Archive der betreffenden Anlagen verfügbar:

Publikationsreihe	Datum	Titel	Autoren
ESRF experiment report MD163-1	2005	Imaging of immobilised magnetic target-nanoparticles: liposomes and ferrofluids entrapping target for magneto-photodynamic X-ray therapy of cancer M-PXT	T. Nawroth C. Meesters H. Decker M. Rusp G. Le Duc S. Corde A. Bravin H.Requardt T. Brochard
GKSS experiment report 2007, ISSN: 0344-9629	2007	SANS measurement of a $\alpha$ -hemolysin pre-pore mutant	C. Meesters

## C.5 Betreute Arbeiten

Folgende Diplomarbeiten wurden von mir betreut oder mitbetreut:

Autor	Fertigstellung	Titel	Betreuer
Bruno Pairet	16. Januar 2007	Röntgenkleinwinkelstreuung an Proteinen – Untersuchung von Konformationsübergängen bei Hämocyaninen	Dr. H. Hartman & Professor H. Decker
Anja Rabenhorst	September 2007	Zur Kristallisation von multimeren Arthropoden- Hämocyaninen	Prof. Dr. E. Jaenicke
Claudia Dietze	2. Juni 2008	Untersuchungen zur Aktivierung von Hämocyanin zur Phenoloxidase	Prof. Dr. E. Jaenicke

Die Diplomarbeit von Bruno Pairet wurde wesentlich von mir betreut. Die Diplomarbeit von Anja Rabenhorst wurde von Professor Jaenicke betreut. Mein Anteil bei dieser Diplomarbeit umfasste die Betreuung des experimentellen Teils zur Kleinwinkelstreuung. Die Diplomarbeit von Claudia Dietze wurde ebenfalls von Professor Jaenicke betreut. Die Daten zur Kleinwinkelstreuung wurden von mir aufgenommen, während sie begleitende biochemische Tests durchführte.

## C.6 Vorträge

An dieser Stelle sind nur Vorträge aufgeführt, die auf Einladung zu Stande kamen.

Anlass	Datum	Titel
Seminarreihe der <i>Structural Biology</i> Abteilung der Universität Davis, Kalifornien	September 2003	Preliminary studies on the human $\text{Na}^+/\text{H}^+$ -exchanger, NHE1
„Lightning Talk“ im Rahmen der EuroSciPy-Konferenz 2008	Juli 2008	<code>sas-rigid</code> – rigid body modelling on small angle scattering data

## C.7 Poster

Mit folgenden Postern konnte ich mein Arbeit vorstellen:

Tagung	Datum	Titel des Posters	Autoren
Microscopy & Microanalysis, Hawaii	31. Juli - 4. August 2005	Testing 3-D Reconstruction Methods for Electron Tomography	S. Mehraeen H. Stahlberg J. E. Evans <i>C. Meesters</i> I. Arslan P. A. Midgley N. D. Browning
MWFZ-Tagung Mainz	2006	Small Angle Scattering on the Hemocyanin of the Scorpion <i>Pandinus imperator</i>	B. Pairet <i>C. Meesters</i> H. Hartmann E. Jaenicke H. Decker
ESRF user meeting, Grenoble	7. / 8. Februar 2006	Therapeutic imaging and indirect radiation therapy (IRT) with target nanoparticles for cancer therapy	T. Nawroth G. Le Duc <i>C. Meesters</i> H. Decker S. Corde H. Requardt A. Bravin
Tagung der Deutschen Biophysikalischen Gesellschaft in Mainz	24. - 27. September 2006	Robust Molecular Weight Determination using Small Angle X-Ray Scattering without absolute intensity scaling	<i>C. Meesters</i> H. Decker
Tagung der Deutschen Biophysikalischen Gesellschaft in Mainz	24. - 27. September 2006	Small Angle Scattering on the Hemocyanin of the Scorpion <i>Pandinus imperator</i>	B. Pairet <i>C. Meesters</i> H. Hartmann E. Jaenicke H. Decker
Tagung der Deutschen Biophysikalischen Gesellschaft in Mainz	24. - 27. September 2006	Physicochemical comparison of the cockroach allergen Per a 3 with the cross-reactive protein Bla Hx	A. Beusch B. Lafargue <i>C. Meesters</i> S. Tenzer H. Decker
SNI-Tagung, Hamburg	4. - 6. Oktober 2006	Small Angle Scattering on the Hemocyanin of the Scorpion <i>Pandinus imperator</i>	<i>C. Meesters</i> B. Pairet H. Hartmann E. Jaenicke H. Decker
MWFZ-Tagung Mainz	29. Juni 2007	Conformational Changes of Scorpion Hemocyanin upon De-/Oxygenation	B. Pairet <i>C. Meesters</i> H. Hartmann E. Jaenicke H. Decker
MWFZ-Tagung Mainz	29. Juni 2007	X-ray solution structure of the $\alpha$ -Hemolysin monomer	<i>C. Meesters</i> A. Brack H. Decker N. Hellmann
TETHMEM 2007 (Novel Model Systems for Bimolecular Lipid Membranes), Schloss Ringberg	19.-22. September 2007	Interaction of $\alpha$ -toxin from <i>Staphylococcus aureus</i> with lipid vesicles	A. Brack <i>C. Meesters</i> A. Valeva R. Stork S. Bhakdi H. Decker N. Hellmann
Workshop des SFB 625, Mainz	27. September 2007	Structural Characterization of the $\alpha$ -Hemolysin monomer	<i>C. Meesters</i> A. Brack H. Decker N. Hellmann
CBSB 08, Jülich	19. - 21. Mai 2008	A Software Library for Monte Carlo-Based Rigid Body Modelling Against Small Angle Scattering Data	<i>C. Meesters</i>

# Anhang D

## Bildrechte

Das Layout für Abbildung 1.1 ist dem Layout der Abbildung 8 in der Dissertation von Thorsten Schweikardt (Schweikardt, 2008) nachempfunden. Die Erlaubnis für die Kopie liegt vor.

Das Copyright für Abbildung 2.1 liegt bei Alf Neu. Die Erlaubnis für die ausschließliche Verwendung in dieser Dissertationsschrift liegt vor.

Das Copyright für den Nachdruck des Fotos der Kellerrassel *Porcellio scaber* (Abbildung 1.8) liegt bei Hania Berdys. Die Erlaubnis für die ausschließliche Verwendung in dieser Dissertationsschrift liegt vor.

Die Verwendung von Abbildung 5.7 für diese Dissertationsschrift ist durch die Nature Publishing Group durch die Copyrightlizenz 1984261508632 des Copyright Clearance Center, MA, USA gedeckt.

Bilder aus wissenschaftlichen Veröffentlichungen wurden unter Angabe der Quelle gemäß §51 UrhG verwendet. Die Erlaubnis für die Abbildungen 4.10 c) und d) und 4.11 c) und d) wurde explizit von Anja Rabenhorst eingeholt. Gleiches gilt für Abbildung 4.7 a): Die Erlaubnis zur Verwendung in dieser Arbeit wurde von Claudia Dietze erteilt. Abbildung 4.30 wurde bereits in der Diplomarbeit von Alexandra Beusch gezeigt. Die Erlaubnis für die Wiederverwendung in dieser Arbeit liegt vor. Und schließlich liegt die Erlaubnis zur Verwendung von Abbildung 5.4 und der gemeinsam gewonnenen SAXS-Daten von Bruno Pairet vor.

Die Erlaubnis für den Nachdruck von Bildern und Photographien aus Institutsbeständen wurde von Herrn Professor Decker eingeholt. Auch diese Abbildungen dürfen nicht außerhalb dieser Dissertationsschrift verwendet werden.

*Sofern nicht anders angegeben sind alle weiteren Abbildungen durch mich für diese Dissertationsschrift erstellt worden und dürfen ohne Erlaubnis des Autors und des Institutes für Molekulare Biophysik der Universität Mainz nicht anderweitig verwendet oder nachgedruckt werden.*