

# Titel der Bachelorarbeit

Bachelorarbeit  
zur Erlangung des akademischen Grades  
Bachelor of Science

vorgelegt von  
Markus Stabrin  
geboren in Hamburg

Angefertigt am Max-Planck-Institut für molekulare Physiologie in Dortmund  
Geschrieben am Lehrstuhl für Experimentelle Physik I  
Fakultät Physik  
Technische Universität Dortmund  
2014

1. Gutachter : Prof. Dr. Metin Tolan
2. Gutachter : Prof. Dr. Stefan Raunser

Datum des Einreichens der Arbeit: 09.07.2014

## **Kurzfassung**

Hier folgt eine kurze Zusammenfassung des Inhalts und der Ergebnisse der Arbeit in deutscher Sprache.

## **Abstract**

A brief summary of the content and the results of the thesis is given here in english.

# Inhaltsverzeichnis

<b>Inhaltsverzeichnis</b>	<b>IV</b>
<b>Abbildungsverzeichnis</b>	<b>V</b>
<b>Tabellenverzeichnis</b>	<b>VI</b>
<b>1 Einleitung</b>	<b>1</b>
<b>2 Theoretische Grundlagen und Messvorbereitung</b>	<b>2</b>
2.1 Hämocyanin . . . . .	2
2.2 Transmissionselektronenmikroskopie . . . . .	3
2.3 Probenpräparation für die negative Kontrastierung . . . . .	6
2.4 Kryopräparation . . . . .	7
<b>3 Computerbasierte Einzelpartikelanalyse</b>	<b>8</b>
3.1 Digitale Bildbearbeitung . . . . .	9
3.2 Extraktion der Partikel . . . . .	10
3.3 Ausrichtung und Klassifizierung . . . . .	10
3.4 Rekonstruktion . . . . .	11
3.5 Verbesserung der Auflösung . . . . .	11
<b>4 Trickreiche Symmetrieproblem</b>	<b>13</b>
<b>5 Ergebnisse</b>	<b>14</b>
<b>6 Zusammenfassung und Ausblick</b>	<b>15</b>
<b>Quellenverzeichnis</b>	<b>15</b>

# Abbildungsverzeichnis

2.1	Schematische Aufbau eines Transmissionselektronenmikroskop . . . . .	3
2.2	Beispiel einer Kontrast-Transfer-Funktion . . . . .	5

# Tabellenverzeichnis

2.1	Mikroskopdaten . . . . .	4
3.1	Verwendete EDV . . . . .	9

# Kapitel 1

## Einleitung

Hier folgt eine kurze Einleitung in die Thematik.

Hier folgt eine kurze Einleitung in die Thematik.

Die Einleitung muss kurz sein, damit die vorgegebene Gesamtlänge der Arbeit von 25 Seiten nicht überschritten wird. Die Beschränkung der Seitenzahl sollte man ernst nehmen, da überschreitung zu Abzügen in der Note führen kann. Selbstverständlich kann jede(r) für sich eine Version der Arbeit mit beliebig vielen und langen Anhängen und Methodenkapiteln erstellen. Das ist dann aber Privatvergnügen; die einzureichende und zu beurteilende Arbeit muss der Längenbeschränkung genügen. Um der Längenbeschränkung zu genügen, darf auch nicht an der Schriftgröße, dem Zeilenabstand oder dem Satzspiegel (bedruckte Fläche der Seite) manipuliert werden.

## Kapitel 2

# Theoretische Grundlagen und Messvorbereitung

### 2.1 Hämocyanin

(Abbildung)

Mit einer Größe von 3-8 MDa gehören Hämocyanine zu den größten bekannten Proteinen der Natur (Habe im Evolution of Hämocyanin structure gesucht, aber nichts über die Größe direkt gefunden). Diese transportieren sowohl den Sauerstoff von marinen lebenden wirbellosen Tieren als auch einiger Spinnenarten. Sie liegen ohne an Blutzellen gebunden zu sein direkt in der Hämolymphe dieser Tierarten vor. In den aktiven Zentren bindet der Sauerstoff an Kupfer und oxidiert diesen von Cu(I) zu Cu(II). Dieses emittiert, durch Kupfer-Peroxo-Komplexbildung mit dem Sauerstoff, Lichtquanten im sichtbaren Wellenlängenbereich von 480-420 nm. (<http://www.chemie-unterricht.de/dc2/komplexe/farbe.html>)(<http://archimed.uni-mainz.de/pub/2001/0109/diss.pdf>) Damit sind Lebewesen mit Hämocyaninen Blaublüter.

Hämocyanine besitzen mit einem Radius von ca. 35 nm und einer Höhe von ca. 15 nm eine hohlzylindrische Struktur (Quelle Paper von Julian). Sie bestehen aus 10 strukturell ähnlichen Untereinheiten, und sind somit Decamere (Oligomere bestehend aus 10 Untereinheiten) (Quelle Paper Julian). Jede Untereinheit besteht weiterhin aus sieben bis acht funktionellen Einheiten (FE), die artspezifisch sind. Das Hämocyanin des *Sepia Officinalis* (*gemeiner Tintenfisch*), das in dieser Arbeit untersucht wird, besteht beispielsweise aus acht dieser FEs (benannt N' term-a-b-c-d-d'-e-f-g-C' term) (Quelle Paper von Julian). Jede FE ist ca. 50 kDa schwer, wodurch das gesamte Protein ein Gewicht von ca. 4 MDa besitzt.

Die FEs FE-a, FE-b, FE-c, FE-d, FE-e und FE-f sind bei allen bekannten Hämocyaninen vorhanden. Dabei ähnelt sich deren Architektur und es liegt nahe, dass sich diese FEs vor der Aufspaltung der einzelnen Arten entwickelt haben. Daher ist anzunehmen, dass diese in der Struktur dieselben Positionen besetzen und somit dieselbe Rolle beim Bilden der Quartärstruktur ihres Proteins spielen, obwohl dies noch nicht direkt gezeigt werden konnte (Habe die Paper gesucht aber nicht gefunden, waren das die, die du mir



noch schicken wolltest?). FE-d' hingegen entwickelte sich vor etwa 740 Millionen Jahren durch eine Duplizierung der FE-d.

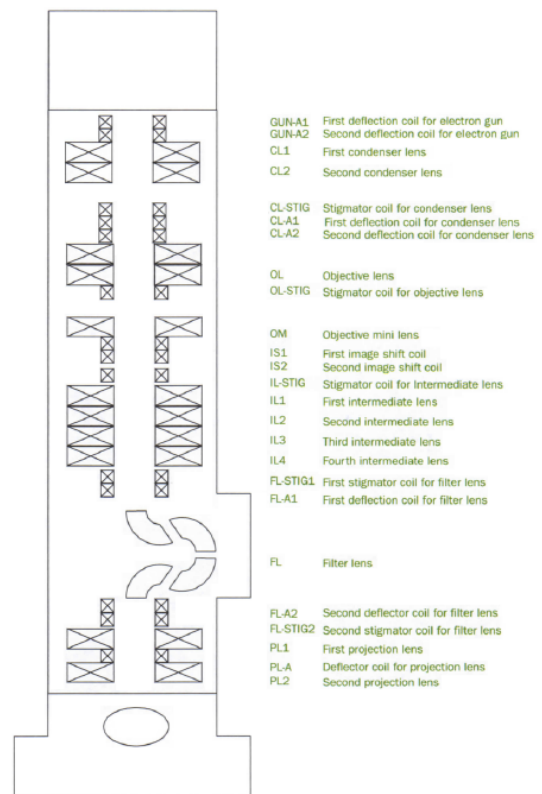
Geometrisch betrachtet besteht der Hohlzylinder aus zwei kovalent miteinander verbundenen Bereichen, der äußeren Wand und dem inneren Kragen. Die Wand ist charakteristisch für alle Hämocyane und setzt sich aus den oben beschriebenen FEs a,b,c,d,e und f zusammen. Der Kragen hingegen ist bei jeder Hämocyaninart anders aufgebaut und besteht beim *Sepia Officinalis* aus FE-d' und FE-g.

Da die Funktionsweise von Proteinen durch ihre Quartärstruktur festgelegt ist, ist es notwendig eine Strukturanalyse durchzuführen. Eine Möglichkeit wäre die hochauflösende Kristallographie, doch lassen sich besonders große Proteine nur sehr schwer kristallisieren. Zudem liegen diese nicht mehr in ihrer nativen Konformation vor, wie es in Lösung der Fall ist. Daher bietet sich die Kryotransmissionselektronenmikroskopie an, welche eine Hochauflösung der Struktur ermöglicht.

## 2.2 Transmissionselektronenmikroskopie

Das Prinzip der Bildentstehung bei einem Transmissionselektronenmikroskop ist dem eines Lichtmikroskops sehr ähnlich: Ein Elektronenstrahl erzeugt eine Projektion der Probe im Realraum. Im Vergleich zum Lichtmikroskop werden jedoch keine optischen, sondern elektrische Linsen verwendet. Ein schematischer Aufbau der verwendeten Mikroskope ist in Abbildung 2.1 dargestellt. Die dazugehörigen technischen Daten können der Tabelle (2.1) entnommen werden.

Als Elektronenquelle wird je nach Mikroskop entweder ein LaB<sub>6</sub>-Kristall (Lanthanhexaborid) oder eine Feldemissionsquelle genutzt. Bei der Feldemission ( $\Delta E = 0.6-1.2 \text{ eV}$ ) ist ein schmaleres Energiespektrum der Elektronen zu beobachten als beim LaB<sub>6</sub>-Kristall ( $\Delta E = 1.5-3 \text{ eV}$ ), wodurch ein höheres Auflösungsvermögen ermöglicht wird. Nach dem Beschleunigen der Elektronen werden diese mithilfe von elektronischen Linsen auf die Probe und anschließend auf die Kamera fokussiert. Diese Linsen bestehen aus Spulen, die Magnetfelder erzeugen. Die bewegten Elektronen werden von diesen durch die Lorentz-Kraft



**Abbildung 2.1:** Schematische Aufbau eines Transmissionselektronenmikroskops. (Quelle Benutzerhandbuch, JEOL Ltd.)

	JEM1400	JEM3200-FSC
Elektronenquelle	LaB <sub>6</sub> -Kristall	Feldemission (Schottky)
Typ	TEM	Feldemission-Kryo-TEM
Beschleunigungsspannung	40-120 kV	100-300 kV
Vergrößerung	200 - 1.200.000	100-1.200.000
Probenkühlung	Stickstoff	Stickstoff, Helium
Energiefilter	keiner	in-column-Energiefilter

**Tabelle 2.1:** Kenngrößen der verwendeten Mikroskope. (Quelle Benutzerhandbuch, JEOL Ltd.)

abgelenkt ohne deren Energie zu beeinflussen (Nur die senkrechte Komponente der Kraft wirkt auf die Elektronen). Um Wechselwirkungen der Elektronen mit anderer Materie als der Probe zu verhindern, ist das gesamte Spulensystem evakuiert. Abschließend wird das Transmissionsbild im Objektsystem vergrößert und auf einem Fluoreszenzschirm abgebildet oder direkt mit einer Kamera oder einem Photofilm detektiert. Abhängig von der Intensität erzeugt dieser ein Bild mit Graustufen. Je dunkler der Bildpunkt ist, umso größer ist die Intensität.

Damit auf dem Bild Strukturen erkennbar werden, muss ein ausgeprägter Amplitudenkontrast zwischen Probe und Lösung vorhanden sein. Eine Schwierigkeit tritt dabei bereits dadurch auf, dass die Elektronendosis bei biologischen Proben nicht zu groß gewählt werden darf um Schäden zu vermeiden. Der Amplitudenkontrast kann sowohl durch Teilchen- als auch durch Welleneigenschaften der Elektronen entstehen. Bei der Betrachtung als Teilchen ergibt sich der Kontrast durch Streueffekte. Dafür dürfen die Ordnungszahlen der Elemente in der Probe und der Lösung nicht in derselben Größenordnung sein. Dies ist auf die Wechselwirkung der Elektronen an Atomkernen zurückzuführen. An großen Kernen werden diese vermehrt elastisch gestreut, während an kleinen Kernen bevorzugt inelastische Streuung auftritt. Erstere führt zu einer Richtungsänderung der Elektronen. Mithilfe einer Blende werden diese aus dem Elektronenstrahl entfernt.

Die inelastisch gestreuten Elektronen tragen jedoch nur zum Rauschen im Bild bei und tragen keine Information. Der Amplitudenkontrast kommt dementsprechend durch die geringere Intensität bei Elementen größerer Ordnungszahl zustande.

Befinden sich nun die Elemente der Probe und der Lösung in derselben Größenordnung, werden an beiden etwa gleich viele Elektronen elastisch gestreut und es ist kein Kontrast zu erkennen. Wird das Elektron jedoch als Welle betrachtet kann eine andere Art von Kontrast hergeleitet werden. Wenn elektromagnetische Wellen mit Materie wechselwirken, kommt es zu einer Phasenverschiebung. Dies führt wiederum zu Interferenzerscheinungen zwischen den gebeugten und den ungebeugten Elektronen und wird als Phasenkontrast bezeichnet. Die Verschiebung wird durch die Gleichung (2.1) (Quelle) beschrieben.

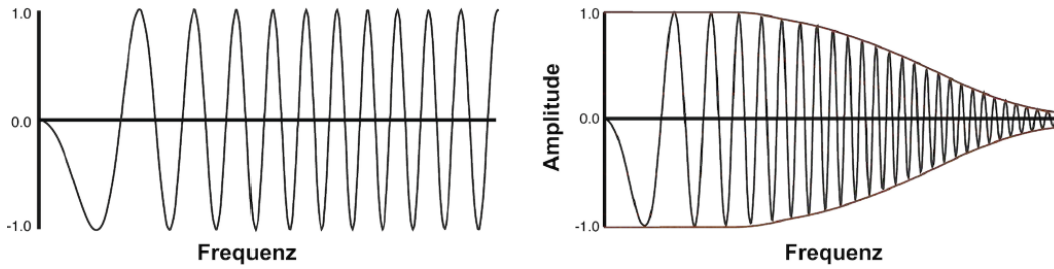
$$\gamma(k, \Delta z) = \frac{\pi}{2} (\lambda^3 C_s k^4 - 2\lambda \Delta z k^2). \quad (2.1)$$

Dabei ist  $\Delta z$  der Unterfokus,  $C_s$  die mikroskopspezifische sphärischen Aberati-

onskonstante,  $\lambda$  die Wellenlänge der Elektronen und  $k$  der reziproke Abstand vom Bildmittelpunkt im Frequenzraum. Biologische Proben führen nur zu einem schwachen Phasenkontrast, weshalb der Unterfokus relativ hoch gewählt werden muss. Zur Rückrechnung auf den Amplitudenkontrast wird die Kontrast-Transfer-Funktion (KTF, Gleichung (2.2), Quelle) verwendet, die sowohl Anteile von Phasen-, als auch Amplitudenkontrast beinhaltet. Sie ist die Fouriertransformierte der Punktspreizfunktion, die angibt, wie ein punktförmiges Objekt durch das System dargestellt wird.

$$\text{KTF}(\gamma(k, \Delta z)) = \sqrt{1 - A^2} \sin[\gamma(k, \Delta z)] - A \cos[\gamma(k, \Delta z)]. \quad (2.2)$$

Ein Beispiel für eine KTF ist in Abbildung ?? dargestellt. Die Funktion alterniert zwischen  $+1$  und  $-1$ , wobei die Nullstellen vom Unterfokus abhängen. An den Nullstellen gibt es keinen Kontrast und das Bild enthält keine Informationen bei den entsprechenden Frequenzen. Es ist daher wichtig während der Bildaufnahme verschiedene Defokuswerte zu wählen, um möglichst alle Frequenzen im gesamten Fourierraum abzudecken. Die Funktion wird durch eine einhüllende Funktion gedämpft und ist auf Störungen (Drift, Energieaufspaltung im Strahl usw.) zurückzuführen. Beschrieben wird diese dabei durch den Exponentialkoeffizienten (B-Faktor) der approximierten Gaußfunktion.



**Abbildung 2.2:** Beispiel einer Kontrast-Transfer-Funktion. Links ist der theoretische, rechts ist der reale Verlauf dargestellt. Der reale Verlauf wird von einer dämpfenden Funktion eingehüllt. (Quelle Behrmann, 2012)

Für das Auflösungsvermögen eines Mikroskops gilt: Zwei Gegenstände können nicht voneinander getrennt dargestellt werden, wenn das Objekt kleiner ist, als die Wellenlänge der eingesetzten Strahlung. Bei Elektronen, die mit einer Beschleunigungsspannung von 300 kV beschleunigt werden, liegt diese bei  $\lambda = 0.02 \text{ \AA}$ . Die wichtigsten begrenzenden Faktoren für die Auflösungsgrenze sind das Abbe-Kriterium, Abbildungsfehler und die Pixelgröße des Photochips. Das Abbe-Kriterium ergibt sich aus der Gleichung (2.3) (Quelle).

$$d = 0.61 \frac{\lambda}{n \sin(\alpha)} \quad (2.3)$$

Es ist abhängig von der Brechzahl  $n$  des Mediums zwischen Objektiv und Probe und dem halben Öffnungswinkel des Objektivs (Quelle). Abbildungsfehler, wie die sphärische Aberration, führen zu Anteilen in der KTF (Gleichung (2.2)). Bei höheren

Frequenzen führt der geringer werdende Abstand zwischen zwei Nulldurchgängen dazu, dass das Signal nicht mehr fehlerfrei ausgewertet werden kann. Auch ein Driften der Probe während der Aufnahme führt zu einer Abschwächung des Signals. Zudem spielt die Pixelgröße  $A_{px}$  des Photochips eine Rolle, das Abtasttheorem (Quelle) besagt, dass die maximale Auflösung dem doppelten Wert der Pixelgröße entspricht. Eine Fouriertransformation überträgt die Bildinformation in den Frequenzraum, in dem die Intensität bei verschiedenen Frequenzen  $f_s$  (spatial frequency). Die Frequenz der maximalen Auflösung wird als Nyquist-Frequenz bezeichnet. Es wird meist die Einheit der absoluten Frequenz  $f_a$  verwendet, die eine Verbindung zwischen der Pixelgröße und der Frequenz durch Gleichung (2.4) erstellt.

$$f_a = A_{px}/f_s \quad (2.4)$$

Es ist zu erkennen, dass dieser Wert nicht kleiner als 0.5 werden sollte, da dies der Nyquist Frequenz entspricht und eine höhere Auflösung nicht möglich ist.

## 2.3 Probenpräparation für die negative Kontrastierung

Um Proben mit einem Transmissionselektronenmikroskop untersuchen zu können, müssen zunächst ein paar Vorbereitungen getroffen werden. Die Probe muss strukturerhaltend auf dem Probenträger fixiert werden, damit diese sich nicht bewegen kann. Dies hat zudem den Vorteil, dass die Probe nicht in das Vakuum des Mikroskops evaporieren kann. Zudem muss ein Kontrast vorhanden sein, um Strukturen der Probe erkennen zu können.

Die Methode der negativen Kontrastierung ist eine Möglichkeit zum Erfüllen dieser Anforderungen. Dabei wird die Probe zusammen mit einer Schwermetalllösung (0.75 % Uranylformiat) auf einem Probenträger aufgetragen. Dieser ist ein (größe hier einfügen) großes, rundes Kupfergitter (G2400C, Plano GmbH). Die Probe wird zur Verarbeitung in einem Puffer auf die benötigte Konzentration verdünnt (TRIS 50 mM pH 7.4, NaCl 150 mM, MgCl<sub>2</sub> 5 mM, CaCl<sub>2</sub> 5 mM). Nach der Benetzung mit einem Polymerfilm wird das Gitter mit Kohlenstoff bedampft, auf dem die Probe später haften wird. Das Gitter wird in ein Plasma eingebracht, das aus beschleunigten geladenen und ungeladenen Atomen und Molekülen besteht. Durch die Wechselwirkung entstehen Ionen auf der Oberfläche und diese wird polarisiert. Dies ist notwendig, da die Schwermetall- und Probenlösung hydrophil (polar) ist und die Probe ansonsten nicht haftet. Nach dem Auftragen der Probe wird diese gewaschen, mit der Schwermetalllösung benetzt und anschließend luftgetrocknet.

Die negative Kontrastierung ist eine gute Möglichkeit zur Kontrastverstärkung in einem Transmissionselektronenmikroskop bei Proben, die primär aus leichten Elementen aufgebaut sind. Wie in Kapitel 2.2 beschrieben streuen größere Atomkerne im Vergleich vermehrt elastisch, während kleinere vermehrt inelastisch streuen. Biologische Proben bestehen aus leichten Elementen (Kohlenstoff, Stickstoff, Sauerstoff etc.), während Schwermetalle schwere Elemente beinhalten. Der auftretende Intensitätsunterschied führt in dem erzeugten Bild zu einem Amplitudenkontrast. Da an der Lösung

und nicht an der Probe gestreut worden ist, wird diese Methode als negative Kontrastierung bezeichnet.

Die Nachteile der negativen Kontrastierung wirken sich negativ auf die Auflösung aus. Durch das Trocknen der Probe kann es zu Deformationen der nativen Konformation und Artefakten der Kontrastlösung kommen. Dies führt besonders bei großen Molekülen zu deutlichen Fehlern in der Rekonstruktion. Auch gibt es Probleme bei inneren Strukturen. Die Schwermetalle kommen nicht in die Proteine und an diesen Stellen ist der Kontrast kaum bis gar nicht vorhanden.

## 2.4 Kryopräparation

Hämocyanine gehören zu den größten Proteinen der Natur und sind daher besonders anfällig für Rekonstruktionsfehler bei der negativen Kontrastierung. Bei der Kryopräparation wird die Probe in amorphem, vitrifiziertem Eis schockgefroren. Damit erfüllt diese Methode die Anforderungen aus Kapitel 2.3. Da die Probe in Lösung gefriert und nicht getrocknet wird, verbleibt diese in ihrer nativen Konformation. Für den Vitrifizierungsvorgang ist das Gerät *Cryo Plunge 3* verwendet worden. Ein Gefrieren ohne Bildung von Eiskristallen ist mit flüssigem Stickstoff nicht möglich, da der Leidenfrost-Effekt ein langsames Gefrieren bedingt. Daher findet das eigentliche Vitrifizieren in einem mit Ethan gefüllten Behälter statt, der mit Stickstoff gekühlt wird.

Biologische Proben bestehen wie Wasser aus leichten Elementen und streuen bevorzugt inelastisch. Wie in Kapitel 2.2 beschrieben, sind keine Strukturen bei einem Amplitudenkontrast etwa gleich großer Elemente zu erkennen. Daher ist es besonders wichtig auf den Phasenkontrast nach Gleichung (2.2) zu nutzen.

Der Probenträger ist ein etwa (größe hier einfügen) großes Metallgitter (z.B. Quantifoil R2/1 Cu-300, Plano GmbH). Dieses ist bereits mit einem Kohlenstofffilm überzogen, in dem in regelmäßigen Abständen industriell Löcher geätzt worden sind (<http://www.cp-download.de/plano11/Kapitel-1.pdf>). Die Oberfläche des Gitters wird, wie in Kapitel 2.3 beschrieben, durch ein Plasma polarisiert. Zum Auftragen der Probe wird das Gitter in eine Pinzette eingespannt und in das Gerät gehängt. Durch eine sich seitlich befindende Öffnung wird nun die Probe auf das Gitter aufgetragen. Das Gerät übernimmt den Ablösch- und Gefriervorgang vollautomatisch. Vom Probenträger wird die Probenlösung mit Filterpapier fast vollkommen entfernt, sodass nur noch ein dünner Film übrig bleibt. Anschließend wird das Gitter in den mit flüssigem Ethan gefüllten Behälter fallen gelassen. Dabei ist es wichtig, dass die Luftfeuchtigkeit erhöht ist. Dadurch wird ein Verdunsten während des Fallens verringert. Abschließend kommt das Gitter direkt in die sich in flüssigem Stickstoff befindliche Probenaufbewahrungsbox und dies wird wiederum bis zur Verwendung in flüssigem Stickstoff gelagert.

## Kapitel 3

# Computerbasierte Einzelpartikelanalyse

(Bild von Vorgehensweise hier!)

Ein Problem bei der rekonstruktion eines 3D Modells mithilfe eines Transmissions-elektronenmikroskops ist, dass sich nur 2D Abbildungen der 3D Elektronendichte erzeugen lassen. Es ist daher eine Rückprojektion nötig, um das Objekt darstellen zu können. Eine etablierte Methode für Rücktransformationen ist die Tomographie, doch ist diese bei biologischen Proben nicht vorteilhaft, da die relativ hohe Elektronendosis das Material zerstört und nicht alle Raumwinkel abgedeckt werden können. Alternativ wird daher auf die Einzelpartikelanalyse zurückgegriffen. Für diese ist es notwendig, dass die Probe ungeordnet vorliegt, sodass viele 2D Projektionen bei verschiedenen Projektionsparametern der Probe auftreten. Diese sind neben den drei Euler-Winkeln  $\Psi$ ,  $\Phi$  und  $\Theta$  auch die Verschiebung in x- und y-Richtung, da die ausgewählten Partikel nicht immer im Schwerpunkt des Bildes liegen. Während bei der Tomographie diese bekannt sind, müssen bei dieser Methode diese erst iterativ berechnet werden. Die Methode arbeitet nach dem in Abbildung ?? dargestellten Schema.

Da die Anzahl der verwendeten Partikel mehrere Tausend umfasst, muss auf die Rechenleistung von Computern zurückgegriffen werden. Eine Liste der verwendeten Systeme und Programme ist in Tabelle 3.1 gegeben.

Um eine effizientere Methode der Verarbeitung zu gewährleisten ist es nötig das Bildformat der Bilddateien von TIFF (*tagged image file format*) auf HDF (*hierarchical data format*) zu ändern. Diese Art von Dateien besitzen eine Kopfzeile (*header*), in der Bildinformationen (Projektionsparameter, Verschiebungen, Bildzugehörigkeit etc.) gespeichert werden können. Zudem gibt es die Möglichkeit die Speicherung in einer Datenbank (BDB, *Berkeley-Datenbank*) vorzunehmen. In dieser sind Bild und Kopfzeile getrennt voneinander gespeichert, sodass diese getrennt geladen werden können um die Zugriffszeit zu minimieren.

In diesem Kapitel wird besonders auf die Datenauswertung der bei der Kryo-TEM entstandenen Daten eingegangen. Bei der negativen Kontrastierung sind die Schritte analog. Der schematische Rekonstruktionsprozess kann wie folgt unterteilt werden:

Name	Verwendung	Quelle
Eman/Eman2	Prozessierung	<a href="http://blake.bcm.edu">http://blake.bcm.edu</a>
Sparx	Prozessierung	<a href="http://sparx-em.org">http://sparx-em.org</a>
CTF-GUI	KTF & Defokus bestimmung	MPI-Dortmund (R.Efremov, C.Gatsogiannis)
GUI	Klassifizierung (K-Means)	MPI-Dortmund (C.Gatogiannis)
USFC Chimera	3D Modellierung	<a href="http://www.cgl.ucsf.edu/chimera">http://www.cgl.ucsf.edu/chimera</a>
Python	Programmierung	<a href="https://www.python.org">https://www.python.org</a>
CLB cluster	Datenverarbeitung	MPI Dortmund

**Tabelle 3.1:** Verwendete Software, Programmiersprachen und Systeme.

1. **Digitale Bildbearbeitung:** Bestimmung des Defokus aus der KTF-Information; Aussortieren von Bildern (*micrograph*); Filterung der Bilder
2. **Extraktion der Partikel:** Auswahl der Partikel aus den Bildern (*particle picking*); Speichern der Einzelnen Partikel in einem Datensatz (*stack*)
3. **Ausrichtung und Klassifizierung:** Zentrierung und Drehung der Partikel (*alignment*); Zuordnung der Partikel in verschiedene Klassen; Aussortierung schlechter Einzelbilder
4. **3D Rekonstruktion:** Iterative Bestimmung der Projektionsparameter; Rekonstruktion
5. **Verbesserung der Auflösung:** Filterung; Maskierung; Symmetrisierung; KTF Korrektur

### 3.1 Digitale Bildbearbeitung

(Wrapfigure: vorher/nachher untereinander)

Bei der Kryo-TEM sind von derselben Position 15 Bilder in 200 ms Abschnitten gemacht worden. Um das Signal-zu-Rausch Verhältnis zu verbessern, werden diese zunächst miteinander verglichen und aufsummiert.

Die Bilder des Mikroskops sollten in das HDF Format umformatiert werden, um effizienter bearbeitet werden zu können. In der Kopfzeile werden die Bilder zunächst numeriert (*micrograph ID*), sodass später eine Zuordnung möglich ist. Anschließend wird die KTF der einzelnen Bilder analysiert und in der Kopfzeile als KTF-Objekt gespeichert. Dafür wird ein KTF-Fit durch das 1D Powerspektrum des Bildes gefittet, wobei z.B. der Defokus durch die Fitparameter bestimmt wird. Dabei wurde das Programm *CTF-GUI* verwendet. Dieses gibt neben den Fitparametern auch die dazugehörigen Bilder aus, welche bei der Qualitätskontrolle eine Rolle spielen. Stimmt die Qualität nicht, kann das Bild sofort aussortiert werden.

Für die Extrahierung der Partikel im nächsten Schritt wird sowohl ein gaußförmiger Highpass-, als auch ein gaußförmiger Lowpassfilter auf eine Kopie der Bilder angewendet. Dieser bewirkt, dass die besonders kleinen und großen Frequenzen (große und kleine Objekte) unterdrückt werden und die Strukturen der Proteine besser zu erkennen sind.

## 3.2 Extraktion der Partikel

(Wrapfigure: Oben micrograph, da unter 3 einzelpartikel)

Aus den bearbeiteten Bildern müssen nun die einzelnen Partikel extrahiert werden. Dafür ist das EMAN2 Programm *e2boxer.py* verwendet worden. Es wird dafür jeder Partikel auf jedem Bild einzeln ausgewählt und die Koordinaten der Partikelmittelpunkte in einer Textdatei gespeichert. Anschließend werden die Koordinaten auf den ungefilterten Datensatz angewendet und von dort extrahiert. Dabei ist eine quadratische Boxgröße mit einer Seitenlänge von 512 px bei einer Pixelgröße von 0.83 Å gewählt worden. Diese ist nur knapp größer als das Protein und zudem ist die Boxgröße von 256 px eine von EMAN2 empfohlene Boxgröße um eine erste Rekonstruktion mit einem verkleinerten Datensatz durchzuführen. Zudem werden die extrahierten Partikel in einem HDF Datensatz (*stack*) gespeichert. Abschließend wird der Datensatz normalisiert, invertiert und KTF-Korrigiert.

## 3.3 Ausrichtung und Klassifizierung

Klassifizieren bedeutet, dass Gruppen gleicher Projektionsrichtungen gefunden werden. Durch Aufsummieren der Bilder zu einer Klassensumme wird das Signal-zu-Rausch Verhältnis verbessert, wodurch die geringe Elektronendosis ausgeglichen werden kann. Eine Klassensumme entspricht dabei einer Projektionsrichtung. Können manche Partikel keinen Klassen zugeordnet werden, werden diese aus dem Datensatz entfernt, da sie zum Rauschen beitragen.

**1. Ausrichtung (Referenzfrei):** Zunächst werden alle Bilder der Partikel zu einem Bild aufsummiert und als Referenz benutzt. Das Sparxmodul *sxali2d.py* vergleicht zunächst alle Partikel mit dieser und richtet durch Rotation an diesem aus. Anschließend werden die Partikel noch zueinander verschoben. Mathematisch wird dazu eine Kreuzkorrelation angewandt. Nach jedem Iterationsschritt wird eine Überlagerung der neu gefundenen Partikel als Referenz genutzt. (Abbildung der Alignierung ?)

**2. K-Means Klassifizierung:** Eine mögliche Klassifizierungsmethode ist der K-Means Algorithmus (SPARX: *sxk\_means.py*). Dieser erstellt eine festgelegte Anzahl  $K$  an Klassen, indem er die Bilder in einem Hyperraum in Clustern zusammenfasst. Dazu werden zunächst  $K$  zufällige Schwerpunkte in die Hyperebene gesetzt und durch Verschieben ebendieser die Cluster gefunden. Dieser Algorithmus ist zwar einfach und schnell, jedoch nicht reproduzierbar, da das Ergebnis stark von den Anfangsbedingungen und der Qualität der Partikel abhängig ist (Quelle Paper).



**3. Ausrichtung (*multi reference alignment*):** Im Gegensatz zur referenzfreien Ausrichtung stehen nun die Erzeugten Klassen als Referenz zur Verfügung. Die Partikel werden der am besten passenden Klassensumme zugeordnet und nach dieser Ausgerichtet. Anschließend werden die zugeordneten Partikel aufsummiert.

**4. ISAC:** Der ISAC (*Iterative Stable Alignment and Clustering*) Algorithmus ist eine gute Möglichkeit zur Überprüfung der Klassenergebnisse (SPARX:*sxisac.py*). In diesem Fall wird nicht die Anzahl der Klassen, sondern die der Partikel pro Klasse vorgegeben. Nach jedem Iterationsschritt werden die Klassen auf Stabilität und Reproduzierbarkeit im Hinblick auf Alignierung und Klassifizierung geprüft (Quelle Paper).

### 3.4 Rekonstruktion

Um eine erste 3D Rekonstruktion als Startreferenz zu erhalten, werden die erzeugten Klassensummen genutzt. Deren Signal-zu-Rausch Verhältnis ist im Vergleich zu den Einzelpartikeln durch die Summation verbessert worden. Mithilfe des Sparx-Programms *sxviper.py* (Bisher unveröffentlicht) ist so ein erstes Modell erzeugt worden. Als Inertialmodell wurde hierbei eine zylindrische Struktur verwendet.

Eine iterative Findung der Projektionsparameter führt das Sparx-Programms *sxali3d.py* durch. Dafür wird die 3D Referenz bei festen Winkelabschnitten in 2D Projektionen überführt und mit den Partikeln verglichen (*multi reference alignment*). Die Partikel werden nach der am besten passenden Projektion ausgerichtet und die Projektionsrichtungen in die Kopfzeile geschrieben. Aus den bestimmten Parametern wird erneut eine 3D Rekonstruktion aus den Partikeln erstellt und als Referenz an den nächsten Iterationsschritt übergeben. Dabei werden die Größe der Winkelabschnitte stetig verringert, sodass mehr Projektionsrichtungen möglich sind.

Um die Auflösung einer Rekonstruktion zu bestimmen ist das FSK<sub>0.5</sub>-Kriterium angewendet worden. Dazu wird der Datensatz der verwendeten Partikel in zwei gleich große Teile aufgespalten und von jedem eine Rekonstruktion erstellt. Diese werden Mithilfe einer Kreuzkorrelation miteinander verglichen und die Übereinstimmung in Abhängigkeit der Frequenz dargestellt. Als noch aufgelöst wird die Frequenz bezeichnet, bei der die Übereinstimmung noch 50 % beträgt.

### 3.5 Verbesserung der Auflösung

Um nach einem erfolgreichen Iterationsschritt eine verbesserte Referenz zu erhalten ist es von Vorteil, die zuletzt erzeugte Rekonstruktion zu bearbeiten. Dies kann durch Maskierung, Filterung, Symmetrisierung oder KTF-Korrektur möglich gemacht werden.

**Maskierung:** Bei einer 3D Rekonstruktion entstehen Fragmente (Rauschen) außerhalb des Dichtevolumens des Partikels. Diese können beim nächsten Iterationsschritt

zu fehlerhaften Zuordnungen der Projektionsparameter führen. Ein Entfernen des Rauschens ist durch das Maskieren des Partikels mit einer binären Maske möglich. Wird das Volumen mit der Maske multipliziert bleibt nur der Teil übrig, der von der Maske umschlossen ist.

**Filterung:** Stimmen die Partikelprojektionsrichtungen nicht genau mit der Realität überein, kommt es zu Überlappungen der Dichten. Besonders bei den großen Frequenzen (kleinen Strukturen) entsteht so ein Rauschen im Volumen. Daher ist es von Vorteil die diese aus der Dichte mithilfe eines Lowpassfilters herauszufiltern, um das Rauschen zu unterdrücken.

**Symmetrisierung:** Viele Strukturen von Proteinen weisen Symmetrien auf. Dabei treten folglich genau dieselben strukturellen Anordnungen auf, welche auch in der Dichte übereinstimmen sollten. Es ist in solchen Fällen nicht nur möglich während des Rekonstruktionsprozesses die Partikel mit der Referenz zu vergleichen, sondern auch innerhalb der Rekonstruktion die symmetrischen Bereiche. Besitzt ein Partikel beispielsweise eine  $C5$  Symmetrie und werden die symmetrischen Bereiche miteinander verglichen ist das etwa so, als wären 5 mal so viele Partikel zur Rekonstruktion verwendet worden. Sind symmetriebrechende Abschnitte der Struktur vorhanden, ist es nötig diese zuvor mit einer extra angepassten Maske auszuschneiden. Eine Symmetrisierung von nicht symmetrische Bereiche führt nämlich zu einer Zerstörung der Struktur.

**KTF-Korrektur:** Die KTF beinhaltet alle Strukturinformationen des Partikels, doch sind besonders die hohen Frequenzen nicht einfach auszulesen. Wird nicht viel Wert auf eine Hochauflösung gelegt ist es zum Beispiel möglich die KTF-Informationen ab der ersten Nullstelle abzuschneiden, sodass nur die große Strukturen sichtbar werden. Da jedoch für die Hochauflösung die großen Frequenzen relevant sind, ist es unter anderem möglich das Vorzeichen der negativen Amplituden zu ändern (*phase flip*). Dadurch werden die KTFs, welche durch verschiedene Defokus Werte erzeugt worden sind einheitlich und eine Ausrichtung der Partikel wird erleichtert (<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3166661/>). Zudem können gesondert Struktureigene Frequenzen verstärkt werden (*pw adjustment*). Dafür werden Röntgenstrukturen in die Dichte gefittet und ist man sich sicher, dass diese an der richtigen Stelle sind werden deren Frequenzen extrahiert und in der KTF verstärkt. Zudem können die großen Frequenzen mithilfe der Modifikation des B-Faktors besser betrachtet werden (Vergleich Kapitel 2.2). Jedoch verschlechtert sich das Signal-zu-Rausch Verhältnis, da bei der Verstärkung nicht zwischen Information und Rauschen unterschieden werden kann.

## Kapitel 4

# Trickreiche Symmetrieproblem

## Kapitel 5

# Ergebnisse

## Kapitel 6

# Zusammenfassung und Ausblick

Hier sollen die Ergebnisse zusammengefasst und weiterführende Untersuchungen diskutiert werden.

## Eidesstattliche Versicherung

Ich versichere hiermit an Eides statt, dass ich die vorliegende Bachelorarbeit mit dem Titel "Titel der Bachelorarbeit" selbständig und ohne unzulässige fremde Hilfe erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

---

Ort, Datum

---

Unterschrift

### Belehrung

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden (§ 63 Abs. 5 Hochschulgesetz - HG - ).

Die Abgabe einer falschen Versicherung an Eides statt wird mit Freiheitsstrafe bis zu 3 Jahren oder mit Geldstrafe bestraft.

Die Technische Universität Dortmund wird ggf. elektronische Vergleichswerkzeuge (wie z.B. die Software "turnitin") zur Überprüfung von Ordnungswidrigkeiten in Prüfungsverfahren nutzen.

Die oben stehende Belehrung habe ich zur Kenntnis genommen.

---

Ort, Datum

---

Unterschrift