

Einführung

Ziel dieses Übungsblatts ist die Analyse von Unordnung und Ungenauigkeiten in Makromolekülstrukturen. Im Folgenden liegt eine Proteinstruktur vor, welche sowohl durch Kristallstrukturanalyse als auch Kernspinresonanzspektroskopie (NMR) ermittelt wurde. Anhand dieser wird betrachtet, wie sich Unsicherheiten in den Koordinaten der jeweiligen Modelle widerspiegeln und wie sie quantifiziert und visualisiert werden können.

Struktur-Ensembles

Laden Sie das mit NMR-Spektroskopie ermittelte Modell einer mutierten monomerischen Superoxiddismutase 1ba9 in *Chimera*. Hierbei handelt es sich um ein Struktur-Ensemble von 36 unterschiedlichen aus den NMR Daten berechneten Strukturen, welche das Spektrum möglicher Konformationen beschreiben.

Clustering von Strukturen

Zur besseren Übersicht über die Strukturvariationen soll die Visualisierung auf eine Untermenge von Strukturen beschränkt werden, in welcher die Elemente sich stärker von einander unterscheiden. Ein Ansatz hierzu ist das Clustering der Strukturen basierend auf durchschnittlichen Distanzen der Atomkoordinaten. Aus jedem der resultierenden Cluster wird dann die Struktur ausgewählt, welche allen anderen innerhalb des Clusters am ähnlichsten ist (Cluster Repräsentant). Starten Sie dazu Ensemble Cluster:

```
Tools -> MD/Ensemble Analysis -> Ensemble Cluster
```

Beschränken Sie unter `Parts to Match` die berücksichtigten Atome auf das Rückgrat mit einem entsprechenden Spezifikationsausdruck und klicken Sie auf OK.

Im Ergebnisfenster sehen Sie nun eine Liste von Modell-Clustern mit jeweiliger Anzahl der Mitglieder und dem Repräsentant-Modell. Aktivieren Sie die Checkbox `Choose ... in Model Panel` im Feld `Treatment of Chosen Clusters` mit der Option `representatives`. Wählen Sie nun alle Cluster aus. Im `Model Panel` sind nun die Cluster-Repräsentanten ausgewählt. Lassen Sie diese sichtbar (`show only`), **alle Strukturen sollen jedoch aktiviert bleiben**. Das `Ensemble Cluster` Fenster wird nicht mehr benötigt.

Visualisierung lokaler Unordnung

Untersuchen Sie die fünf Cluster Repräsentanten. Was beobachten Sie hinsichtlich der Überlagerung? Existieren starke lokale Unterschiede (Unordnung)? Können Sie für die Unterschiede Tendenzen zu bestimmten strukturellen Elementen beobachten? Was lässt sich über das Molekül an diesen Stellen aussagen?

Soweit haben Sie lokale Unordnung in der Struktur visuell evaluiert. Mit Hilfe der überlagerten Strukturen lässt sich diese auch quantifizieren. Hierfür kann für jedes Atom des Proteinrück-

grates der „*root-mean-square deviation*“ (RMSD) Wert zwischen allen Strukturen als Maß für die durchschnittliche Abweichung der Positionen folgendermaßen berechnet werden:

- Öffnen Sie das Sequenzfenster (`Tools -> Sequence -> Sequence`). Wählen sie das erste Modell aus (die Sequenzen aller Modelle sind zwangsläufig identisch).
- Starten Sie im Sequenzfenster über das Menü `Structure -> Associations` und klicken Sie im erschienenen Fenster auf `Propagate` und anschließend auf `OK`. Dieser Schritt ist notwendig, damit alle 36 Strukturen mit der aktuell angezeigten Sequenz assoziiert werden.
- Es ist wichtig, dass die Strukturen am Rückgrat überlagert sind, da die aktuellen Koordinaten aller Modelle verwendet werden. Sofern Sie bisher alle Strukturen gleichzeitig bewegt haben und keine deaktiviert war, sollte das zutreffen.
- Berechnen Sie die Rückgrat RMSD Werte im Sequenzfenster über `Headers -> RMSD : backbone`. Diese sind nun intern gespeichert und werden im Sequenzfenster als Balkendiagramm über den Residues angezeigt.

Für weitere numerische Analysen können die berechneten Daten über andere *Chimera* Fenster als Datei exportiert werden. Es ist jedoch auch möglich die Daten direkt in Chimera zur Visualisierung zu nutzen. Beschränken Sie dazu zunächst die Sichtbarkeit auf eines der angezeigten Modelle. Starten Sie dann über das Menü des Hauptfensters:

```
Tools -> Depiction -> Render by Attribute
```

In diesem Menü ist es möglich die Visualisierung (unter dem Reiter `Render`) oder Auswahl (`Select`) von Strukturen von gewissen Attributen abhängig zu machen. Wählen Sie unter `Models` das Modell aus, auf welches die Visualisierung angewendet werden soll. Unter `Attributes of` wird bestimmt, aus welcher Ebene der Strukturhierarchie die Attribute verwendet werden und die Visualisierung angewendet wird. Wählen Sie hier `residues`.

Unter `Attribute` wird das verendete Attribut ausgewählt. Die Rückgrat RMSD Werte heißen hierbei `maxRMSDbackbone`. Nach entsprechender Auswahl erscheint darunter nun ein Histogramm, welches aus den Attributwerten erstellt wurde, und in welchem sich mehrere Farbbalken befinden. Diese definieren einen stetigen Farbverlauf, bei dem die Farbe einem entsprechenden Wert im Histogramm zugeordnet ist. Sie können die Position und Farbe der Balken (und somit den definierten Farbverlauf) steuern. Darunter unten befinden sich weitere diverse Einstellungen zur Visualisierung. Behalten Sie dort die vorgegebenen Einstellungen, schauen Sie sich den kleinsten und größten RMSD Wert im Histogramm an und klicken sie auf `Apply`.

Im Hauptfenster ist die Struktur nun nach berechnetem Rückgrat RMSD Wert eingefärbt. Experimentieren Sie im noch offenen `Render/Select by Attribute` Fenster mit unterschiedlichen Darstellungsformen (z.B. bewegen Sie die Farbbalken, ändern Sie die Farben...) und klicken Sie jeweils auf `Apply` um die Visualisierung zu aktualisieren.

Notieren Sie sich den längsten Abschnitt mit der größten lokalen Unordnung der Proteinkette mit Hilfe der vorgestellten Visualisierungen. Deaktivieren Sie abschließend die Sichtbarkeit aller NMR Modelle (schließen Sie die Modelle aber nicht).

Elektronendichte

Dichtekarten

Laden Sie nun die mit Röntgenstrukturanalyse ermittelte Struktur des gleichen Moleküls 1mfm und **bewegen Sie es nicht**. Laden Sie nun die dazugehörige Elektronendichtekarte mit dem Befehl: `open edsID:1mfm`. Die Karte und die Struktur sind nun korrekt überlagert und können zusammen bewegt werden. Wählen für das Molekül eine Darstellungsart, unter der sich die Elektronendichte um Atome herum gut untersuchen lässt — die initiale Schleifen-Darstellung zeigt hierfür zu wenige einzelne Atome.

Beim Laden der Karte erscheint Ihnen das `Volume Viewer` Fenster. In diesem steuern Sie die Darstellung der Elektronendichte. Das Histogramm zeigt die Verteilung verschiedener Dichtewerte an und mit dem Marker steuern Sie den Wert, für den die Isolinien angezeigt werden. Mit dem `step` Parameter lässt sich die Auflösung der Isolinien steuern - setzen Sie diesen Wert auf 1. Es wird bedingt durch übriges Lösungsmittel und Rauschen viel störende Elektronendichte innerhalb der Einheitszelle zu sehen sein. Um nur die Dichte anzuzeigen, welche die Atome des Moleküls direkt umgibt, kann im Menü des `Volume Viewer` das `Zone` Feld aufgerufen werden:

```
Features -> Zone
```

Wenn Sie nun auf den erschienenen `Zone` Button klicken, wird die Darstellung der Dichte auf einen angegebenen Radius (in Ångström) um die ausgewählten Atome im Hauptfenster beschränkt. Sie sollten hierbei nur die Atome des Proteins und ggf. der Ionen auswählen (d.h. ignorieren Sie die Wassermoleküle).

Überlegen Sie, was die Änderung der Isolinien beim Verschieben des Markers im Histogramm über die Elektronendichte aussagt. Untersuchen Sie die Verteilung der Elektronendichte um die Struktur, und experimentieren Sie mit den Darstellungsoptionen im `Volume Viewer` Fenster. Schauen Sie sich insbesondere die Elektronendichte in dem Bereich an, den Sie zuvor notiert haben (lokale Unordnung) und vergleichen Sie diese mit der Dichte im Rest des Moleküls. Sie können die Dichtekarte im `Volume Viewer` durch Klicken auf das Augensymbol verstecken / hervorbringen.

B-Faktoren & Vergleich

Aus der Elektronendichte können B-Faktoren — ein Maß für lokale Ungenauigkeit der Atompositionen — berechnet werden. Im vorliegenden Modell wie in vielen anderen sind die berechneten B-Faktoren bereits in der Strukturdatei hinterlegt und können direkt zur Visualisierung im `Render/Select by Attribute` Menü als Attribut verwendet werden. Achten Sie darauf, dass das entsprechende Modell (1mfm) ausgewählt ist. Das B-Faktor Attribut lässt sich für Atome (`bfactor`) oder gemittelt für die Residues (`average -> bfactor`) darstellen.

Experimentieren Sie mit der Visualisierung der B-Faktor Werte. Vergleichen Sie Strukturabschnitte mit hohen B-Faktor Werten mit den entsprechenden Abschnitten in den NMR Modell und achten Sie dabei auf die berechneten RMSD Werte. Um zu sehen, wie B-Faktoren, Elektronendichte und NMR-Ensemble Abweichungen mit einander korrelieren, kann man das NMR

Ensemble über die Röntgen-ermittelte Struktur überlagern. Aktivieren Sie wieder die Sichtbarkeit eines NMR Modells, färben Sie es nach RMSD Wert ein und rufen Sie das MatchMaker Tool auf:

```
Tools -> Structure Comparison -> MatchMaker
```

Wählen Sie hier 1mfm als Referenzstruktur und alle 1ba9 Modelle zum Matchen aus. Alle vorgegebenen Einstellungen können behalten werden (da die Strukturen sehr ähnlich sind und nichts besonders optimiert werden muss). Klicken Sie auf OK. Das MatchMaker Tool führt im Hintergrund ein Sequenzalignment und Distanzminimierungen durch. Alle Strukturen sollten dann überlagert sein.

Alternative Konformationen

Im Folgenden liegt der Fokus auf einem bestimmten Teil des Moleküls, so dass es hilfreich ist, eine lokale Strukturüberlagerung durchzuführen. Wählen Sie dazu zunächst die **Rückgratatome** der Residues 145 bis 147 aller Modelle aus. Starten Sie dann das MatchMaker Tool (unter Tools -> Structure Comparison) und wählen Sie dort 1mfm als Referenzmodell und alle 1ba9 Modelle zum Matchen aus. Aktivieren Sie unbedingt beide Further restrict matching to current selection Checkboxen, bevor sie auf OK klicken.

Die Strukturen sind im ausgewählten Bereich nun gut überlagert (jedoch nicht im Rest des Moleküls). Um diesen besser im Detail zu betrachten, wechseln Sie zunächst in die Atomare Darstellung, wählen Sie **alle** Atome der Residues 145 bis 147 aus und lassen Sie nur diese anzeigen (show selection). Verstecken Sie Modell 1mfm und die Dichtekarte und betrachten Sie zunächst nur das NMR Ensemble. Zur Übersicht ist es auch hilfreich die Wasserstoffatome zu verstecken (~show H). Betrachten Sie die Seitenkette des Cysteins 146. Was fällt Ihnen auf? Gibt es gewisse Tendenzen zu bestimmten Konformationen? Betrachten Sie nun die selbe Seitenkette in Modell 1mfm. Sehen Sie etwas ungewöhnliches?

Im Röntgen Modell sehen Sie *alternative Konformationen* der selben Seitenkette. Liegt die Seitenkette in unterschiedlichen Konformationen in den kristallisierten Proteinenmolekülen vor, so wird die Elektronendichte entsprechend verteilt im Experiment gemessen. Sind die unterschiedlichen Varianten ausgeprägt genug, so wird die Software, welche die Atome an die Dichte anpasst, unterschiedliche Konformationen erstellen und sie dem Modell hinzufügen. I.d.R. werden diese direkt in die Modelldateien geschrieben und bei der Visualisierung in *Chimera* entsprechend angezeigt.

Verifizieren Sie anhand der Dichtekarte, dass die Elektronendichte für die Konformationen plausibel erscheint (vorher ist es zur Übersicht sinnvoll die Darstellung der Dichte mit dem zone Feature auf das Cystein 146 zu beschränken). Prüfen Sie ebenfalls, ob die unterschiedlichen Konformationen sich im NMR Ensemble widerspiegeln. Schätzen Sie mit Hilfe der Dichtekarte und dem NMR Ensemble ein, welche der Konformationen wahrscheinlicher ist.

Disulfidbrücken

Lassen Sie nun nur Modell 1mfm ohne Elektronendichte sichtbar und wählen Sie die Residues 56 bis 58 aus. Mit `display selection` werden die ausgewählten Atome zusätzlich zu den bisher angezeigten sichtbar gemacht. Sie sehen nun, dass die entfernten Teile der Sequenz lokal bei einander liegen und laut Visualisierung über eine Disulfidbrücke zwischen den Cysteinen 57 und 146 verbunden sind, wobei für jede der Konformationen des Cysteins 146 eine Bindung angezeigt wird.

Betrachten Sie die Bindungen. Was fällt Ihnen bzgl. ihrer Längen auf? Sie können die Distanz zwischen je zwei ausgewählten Atomen mit dem Befehl `distance selection` als Label anzeigen lassen. Finden Sie damit die Längen der Bindungsvarianten heraus. Die Disulfidbrücke ist eine kovalente Bindung von etwa 2.05 Å Länge. Welche der vorgeschlagenen Bindungen ist realistischer, bzw. überhaupt möglich?

Offensichtlich liegt hier ein Problem vor, da für eine der Bindungsvarianten der Abstand zwischen den Schwefelatomen viel zu groß ist. Sie haben zuvor jedoch anhand der Elektronendichte und dem NMR Ensemble verifiziert, dass diese Konformation der Cystein 146 Seitenkette relativ wahrscheinlich ist. Spekulieren Sie mit Hilfe der Analysemethoden und Erkenntnisse aus dem ersten Teil dieser Übung über mögliche Erklärungen für Ihre Beobachtungen. Hierfür ist es notwendig den Rest des Moleküls wieder sichtbar zu machen und zu veranschaulichen, wo genau die betrachtete Disulfidbrücke liegt.

Fragen

1. Das Modell 1mfm wurde mit einer Auflösung von 1.02Å erzeugt. Was sagt dies über die Qualität der ermittelten Struktur aus?
2. Welche ist (generell) die genauere Methode zur Strukturermittlung und warum? Warum bedarf es einer Alternative?
3. Im NMR Modell haben Sie mindestens eine Loop-Region mit hoher lokaler Variation in den Atomkoordinaten beobachtet. Schätzen Sie mit Hilfe der anderen gewonnenen Erkenntnisse ein, ob die Unsicherheit durch einen Mangel von NMR Daten an den Regionen oder durch lokale Flexibilität und Bewegung zustande kommt.
4. Im NMR Modell wurden die Rückgrat RMSD Werte zwischen den einzelnen Strukturen als Maß für Unordnung verwendet. Spekulieren Sie über andere mögliche Anwendungen für Atomare RMSD Werte im Kontext von Molekülstrukturen.
5. Ist es möglich aus einem NMR Ensemble allein alternative Konformationen innerhalb von Proteinstrukturen zu bestimmen? Begründen Sie.
6. Um aus einem NMR Ensemble eine repräsentative Struktur zu ermitteln, wird vorgeschlagen alle Atomkoordinaten zu mitteln. Ist dies ein guter Ansatz? Warum bzw. warum nicht?
7. Überlegen Sie, wie die Qualität einer mit Röntgenstrukturanalyse ermittelten Struktur unabhängig von der Auflösung eingeschätzt werden kann.