

# Proteinstabilität webbasiert analysieren

Dennis M. Krüger, Holger Gohlke

Der CNA-Webserver ist ein Online-Service, der biologisch relevante Proteineigenschaften basierend auf Rigiditätsanalysen vorhersagt und visualisiert. Er erlaubt, Struktur und Flexibilität eines Proteins mit dessen (Thermo-)Stabilität und Funktion zu verknüpfen.

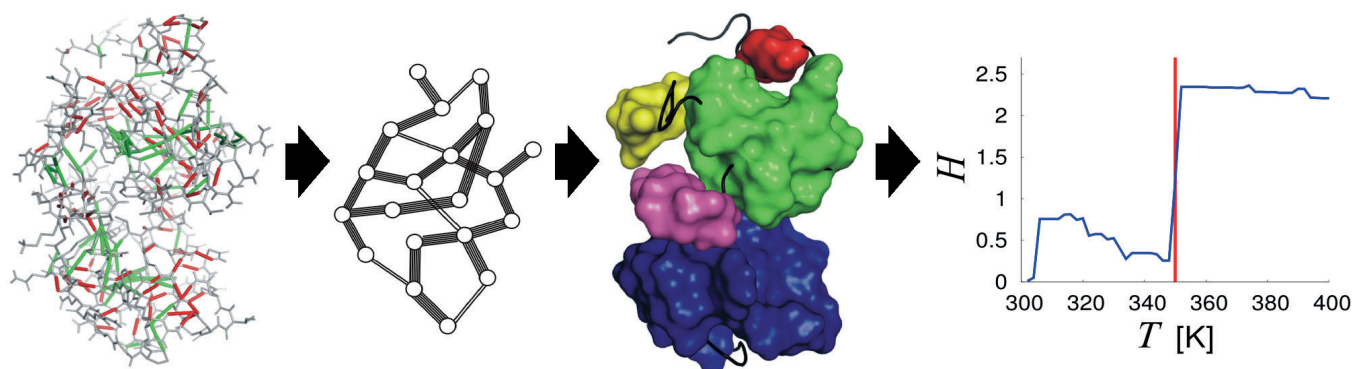


Abb. 1. CNA-Berechnung, von links nach rechts: Startstruktur mit nichtkovalenten Wechselwirkungen (rote und grüne Linien), Repräsentation als Netzwerk, Zerlegung in rigide Cluster, Identifikation von Phasenübergängen nach der Simulation einer thermischen Entfaltung.

● Proteine üben ihre Funktion durch Interaktion mit anderen Biomakromolekülen oder Kleinmolekülliganden aus. Hierzu ist konformationelle Anpassung nötig, um den Bindungspartner zu akkommodieren. Molekulare Erkennung und Katalyse sind auf diese Weise unmittelbar mit der Rigidität beziehungsweise der Flexibilität der Bindungspartner verknüpft. Die Begriffe Rigidität und Flexibilität beschreiben hierbei die prinzipielle Möglichkeit (oder Unmöglichkeit) einer internen Bewegung in einer Gruppe von Atomen.

Weiterhin sind Proteinrigidität und (Thermo-)Stabilität miteinander verknüpft: Proteine aus thermophilen Organismen sind bei gleicher Temperatur häufig rigider als Homologe aus mesophilen Organismen. So ist das Wissen über die Verteilung rigider und flexibler Bereiche in einem

Protein wesentlich, um dessen Struktur mit seiner Funktion und/oder (Thermo-)Stabilität zu verknüpfen. Darüber hinaus sind Informationen über rigide und flexible Bereiche eines Proteins grundlegend für pharmazeutische Wirkstoffforschung.

Moleküldynamik(MD)-Simulation ist der am weitesten verbreitete computergestützte Ansatz, um Proteindynamik und daraus abgeleitet Rigiditäts- und Flexibilitätseigenschaften von Proteinen zu untersuchen. MD-Simulationen liefern Atombewegungen unter Anwendung der Gesetze der klassischen Mechanik und benötigen erheblichen Zeit- und Rechenaufwand. Alternativ lassen sich Proteine mit einer Graphentheorie-basierten Rigiditätsanalyse untersuchen. Die Methode modelliert ein Protein als ein Netzwerk aus Knoten (Atomen) und Distanzbeschränkun-

gen (kovalente und nichtkovalente Wechselwirkungen). Sodann wird für jedes Subnetzwerk bestimmt, ob die mit den Knoten assoziierte Zahl (interner) Freiheitsgrade die der Beschränkungen überschreitet (flexible

## ● QUERGELESEN

- » Anhand eines Proteindatenbank(PDB)-Codes, einer PDB-Datei oder eines Strukturensambles sagt der CNA-Webserver Proteineigenschaften basierend auf Rigiditätsanalysen vorher. Er untersucht dabei, wie starr oder flexibel die Moleküle sind.
- » Globale und lokale Flexibilitätseigenschaften visualisiert der CNA-Webserver mit Graphen und interaktiven Darstellungen.
- » (Thermo-)Stabilität und Zusammenhänge zwischen der Rigidität oder Flexibilität und der Funktion von Proteinen lassen sich untersuchen sowie strukturelle Schwachstellen identifizieren.

Region) oder nicht (rigide Region). Dieser von Kuhn, Thorpe und Mitarbeitern entwickelte Ansatz benötigt für ein Protein mittlerer Größe nur wenige Sekunden.<sup>1)</sup>

Die Constraint-Network-Analyse (CNA) simuliert die thermische Entfaltung eines Proteins durch schrittweises Entfernen nichtkovalenter Wechselwirkungen (Abbildung 1).<sup>2,3)</sup> So können nicht nur rigide und flexible Regionen eines Proteins identifiziert werden; CNA erlaubt darüber hinaus die Vorhersage von Thermostabilitäten, die

Identifikation von strukturellen Schwachstellen (Entfaltungskerne) sowie die Verknüpfung von Proteinflexibilität mit (allosterischer) Proteinfunktion.<sup>4,5,6)</sup> Da eine Rigiditätsanalyse sensitiv bezüglich der verwendeten Eingabestruktur ist, bietet CNA die Möglichkeit, ensemblebasierte Analysen durchzuführen, was zu robusteren Ergebnissen führt. Zum einen kann hierzu ein Strukturensemble analysiert werden, zum anderen existiert ein Ansatz, der die thermische Fluktuation von nichtkovalenten Wechselwirkungen berücksichtigt.<sup>5,7)</sup>

Als Online-Service stehen die Möglichkeiten von CNA im Rahmen des CNA-Webservers zur Verfügung.<sup>8,9)</sup>

### In drei Schritten zur Stabilitätsanalyse

● CNA simuliert die thermische Entfaltung eines Proteins in drei Schritten.<sup>2)</sup> Im ersten Schritt entsteht aus der 3D-Struktur des zu untersuchenden Proteins eine Netzwerksrepräsentation; es lassen sich zudem Kleinmolekülliganden für den Netzerkaufbau berücksichtigen. Als nichtkovalente Wechselwirkungen werden Wasserstoff- und Salzbrücken sowie hydrophobe Wechselwirkungen einbezogen. Einen Temperatureinfluss modelliert die Methode, indem sie schwächere nichtkovalente Wechselwirkungen ab einer gegebenen Temperatur nicht mehr berücksichtigt. Mit steigender Temperatur entsteht so eine Serie von Netzwerken mit einer immer geringeren Dichte an nichtkovalenten Wechselwirkungen. Dieser Prozess spiegelt ein Aufschmelzen der Proteinstruktur mit zunehmender Temperatur wider.

Im zweiten Schritt zerlegt ein kombinatorischer Algorithmus das Netzwerk in rigide Cluster und flexible Regionen.<sup>1)</sup> Ein rigider Cluster ist eine Gruppe von Atomen, die sich nur als Starrkörper bewegen kann. Wird der Algorithmus auf die Serie von Netzwerken angewendet, so ergibt sich eine Trajektorie des

Stabilitätsverlustes des Proteins mit zunehmender Temperatur.

Im dritten Schritt werden globale und lokale Indizes berechnet.<sup>3)</sup> Globale Indizes beschreiben die makroskopische Rigidität oder Flexibilität für jedes untersuchte Netzwerk. Damit lassen sich Phasenübergänge (Schmelzpunkte) entlang der thermischen Entfaltung identifizieren. Diese Phasenübergänge können mit der (Thermo-)Stabilität eines Proteins in Beziehung gesetzt sowie zur Identifikation von strukturellen Schwachstellen im Protein verwendet werden.<sup>4,5)</sup> Lokale Indizes charakterisieren dagegen die Flexibilitäts- und Rigiditätseigenschaften auf der Bindungsebene eines Proteins. Sie dienen im Wesentlichen zur Untersuchung von Stabilitätsänderungen durch Mutationen oder Ligandenbindung.

### Webbasierte Constraint-Network-Analyse

● Das Eingabeformular des Webservers ist benutzerfreundlich. Als Eingabe benötigt er entweder den Code eines Proteins aus der Proteindatenbank (PDB), eine selbst erzeugte PDB-Datei oder ein komprimiertes Archiv (erlaubte Formate: TGZ, TAR.GZ, ZIP), das ein Strukturensemble in Form von PDB-Dateien enthält. Der Nutzer wählt dann eine Berechnungsart:

- Analyse eines einzelnen Netzwerks basierend auf einer einzelnen Eingabestruktur
- Analyse eines Ensembles aus Netzwerken basierend auf einem Strukturensemble
- Analyse eines Ensembles aus Netzwerken basierend auf einer einzelnen Eingabestruktur.

Im letzten Fall werden die Zahl und Verteilung der nichtkovalenten Wechselwirkungen in der Netzwerksrepräsentation innerhalb gewisser Grenzen zufällig variiert. Dies simuliert thermische Fluktuationen des Proteins, ohne Atome zu bewegen.

Nach dem Absenden eines CNA-Auftrags erscheint ein Weblink zur Ergebniseite. Die Darstellung im

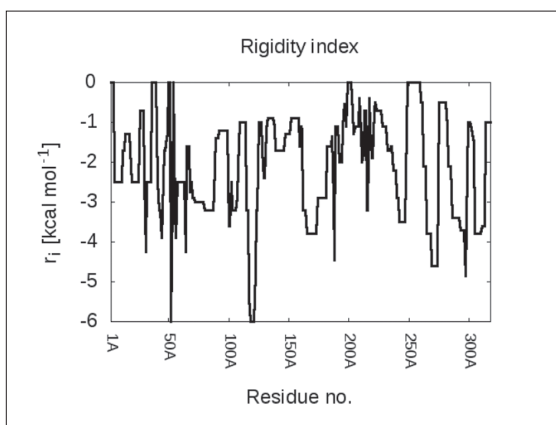


Abb. 2. Lokaler Rigiditätsindex des Proteins.

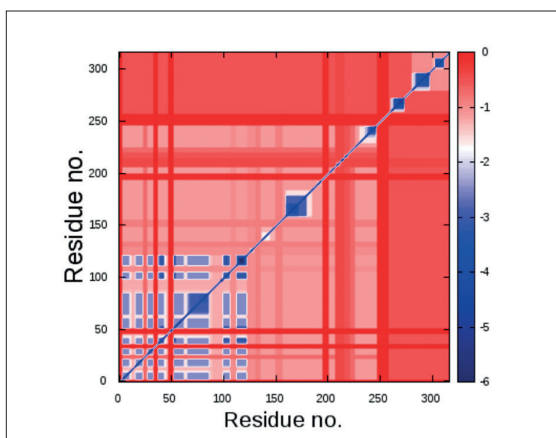


Abb. 3. Stabilitätskarte des Proteins.

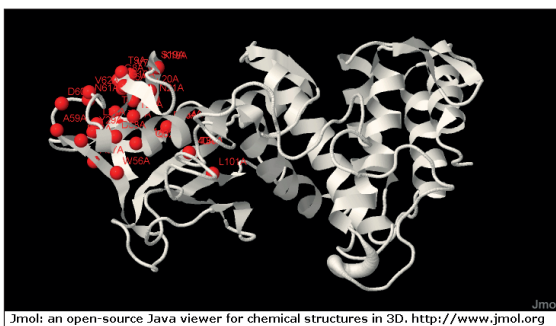


Abb. 4. Visualisierung von strukturellen Schwachstellen im Protein mit Jmol.

Webbrowser erfolgt in einem separaten Fenster. Dieses gibt Aufschluss über den aktuellen Stand und die noch zu erwartende Dauer der CNA-Berechnung. Nach Ende der Berechnung erscheint hier die tabellarische und grafische Auswertung der Ergebnisse. Die Ergebnisausgabe enthält eine Tabelle mit Informationen über alle identifizierten Phasenübergänge. Sofern die Berechnung nur für ein einzelnes Netzwerk durchgeführt wurde, werden zusätzlich Diagramme für alle globalen Indizes angezeigt. Entsprechende Diagramme gibt es ebenfalls für die lokalen Indizes (Abbildung 2).

Weitere Ergebnisse sind Stabilitätskarten (Abbildung 3) und Informationen über strukturelle Schwachstellen (Abbildung 4). Diese werden mit dem Jmol-Plugin auf die Eingabestruktur projiziert und im Webbrowser visuell dargestellt.<sup>9)</sup> Für die Ergebnisausgabe der lokalen Indices und der strukturellen Schwachstellen gibt es zusätzlich eine Startstruktur, die in der ursprünglichen B-Faktoren-Spalte entsprechend modifiziert wurde. Mit externer Visualisierungssoftware lässt sich die Struktur dann gemäß dieser Information kolorieren.

Die Daten einer Analyse sind zehn Tage lang vom Server abrufbar. Häufig gestellte Fragen, Hinweise zu Parametern oder Anforderungen an die Proteinstrukturen sowie eine Beispielrechnung stehen auf Hilfeseiten.

### ● Auf einen Blick

Der CNA-Webserver ist unter Microsoft Internet Explorer, Firefox und Google Chrome nutzbar.<sup>8)</sup> Eine Bildschirmauflösung von mindestens 1024x768 ist empfehlenswert. Da das Jmol-Plugin die Ergebnisse visualisiert, muss die Java-Unterstützung im Browser aktiviert sein.<sup>10)</sup> Alternativ lassen sich die Ergebnisse mit externer Software darstellen, zum Beispiel

### Nicht nur für Bioinformatiker

● Der CNA-Webserver bietet eine einfache und vergleichsweise schnelle Möglichkeit, um Einblicke in die Stabilitätseigenschaften von Proteinstrukturen zu erhalten. Die Simulation der thermischen Entfaltung eines Proteins benötigt in der Regel nur wenige Minuten. Die benutzerfreundliche Umgebung des Online-Dienstes spricht auch solche Forscher an, deren Hauptinteresse nicht die Bioinformatik ist.

Dennis M. Krüger ist Bioinformatiker und arbeitet im Arbeitskreis von Holger Gohlke am Institut für Pharmazeutische und Medizinische Chemie der Heinrich-Heine-Universität Düsseldorf. <http://cpclab.uni-duesseldorf.de>

#### Literatur und Internet

- 1) D.J. Jacobs, A.J. Rader, L.A. Kuhn, M.F. Thorpe, *Proteins* 2001, 44, 150–165.
- 2) C. Pfleger, P.C. Rathi, D. Klein, S. Radestock, H. Gohlke, *J. Chem. Inf. Model.* 2013, 53, 1007–1015.
- 3) C. Pfleger, S. Radestock, E. Schmidt, H. Gohlke, *J. Comput. Chem.* 2013, 34, 220–233.
- 4) S. Radestock, H. Gohlke, *Proteins* 2011, 79, 1089–1108.
- 5) P.C. Rathi, S. Radestock, H. Gohlke, *J. Biotechnol.* 2012, 159, 135–144.
- 6) H. Gohlke, L.A. Kuhn, D.A. Case, *Proteins* 2004, 56, 322–327.
- 7) C. Pfleger, H. Gohlke, *Structure* 2013, DOI:10.1016/j.str.2013.07.012.
- 8) D.M. Krüger, P.C. Rathi, C. Pfleger, H. Gohlke, *Nucl. Acids Res.* 2013, 41, W340–348.
- 9) [www.cnanalysis.de](http://www.cnanalysis.de), <http://cpclab.uni-duesseldorf.de/cna>
- 10) <http://jmol.sourceforge.net/>
- 11) [www.pymol.org/](http://www.pymol.org/)
- 12) [www.ks.uiuc.edu/Research/vmd/](http://www.ks.uiuc.edu/Research/vmd/)

durch Pymol und VMD.<sup>11,12)</sup> Der Webserver ist frei verfügbar und erfordert keine Registrierung oder Benutzerangaben. Missbrauch durch automatische Programme soll ein Captcha-Sicherheitscode verhindern: Der Nutzer muss eine automatisch erzeugte, schwer lesbare Buchstaben-Zahlenkombination eingeben.

[www.cnanalysis.de](http://www.cnanalysis.de)

## Kurz notiert

### Open Science

● Bill Gates und das Finanzierungsunternehmen Tenaya Capital investieren in den Netzbetreiber Researchgate, um die Idee von Open Science voranzutreiben. Die Investoren sind Teil der dritten Finanzierungsrunde. Das Unternehmen entwickelt mit den insgesamt 35 Mio. US-Dollar neue Wege zum Publizieren wissenschaftlicher Ergebnisse, inklusive Forschungsdaten und Erkenntnissen aus misslungenen Versuchen. Zudem plant Researchgate das etablierte System wissenschaftlicher Reputationsmessung um digitale und dynamische Komponenten zu ergänzen.

[www.researchgate.net](http://www.researchgate.net)

### Primärdaten veröffentlichen

● Autoren können ihre Primärdaten online publizieren und für andere Wissenschaftler recherchierbar machen. Wie eine wissenschaftliche Publikation erhalten diese Daten einen Digital Object Identifier (DOI). Wie Primärdaten über Verlagspublikationen der Wissenschaftsgemeinschaft zugänglich sein sollten – dafür erarbeitete die Vertretung der internationalen Wissenschaftsverlage STM mit 15 wissenschaftlichen Bibliotheken und Informationszentren Empfehlungen.

[www.thieme-connect.de](http://www.thieme-connect.de)

### Material sucht Charakterisierung

● Wissenschaftler, die neue Materialien hergestellt haben, finden auf einer Internetplattform diejenigen Kollegen, die Materialien charakterisieren. Die Ergebnisse einer so entstandenen Kooperation können die Partner im assoziierten Journal *Sample of Science Bulletin* nach dem Open-Access-Prinzip als zitationsfähigen Artikel veröffentlichen.

[www.sampleofscience.net](http://www.sampleofscience.net)

### Offen wie ein Scheunentor

● Eine Weltkarte mit industriellen Kontrollsystemen, die vor Spionageangriffen nicht geschützt sind, zeigen Informatiker der Freien Universität Berlin.

[www.scadacs.org/iram.html](http://www.scadacs.org/iram.html)