



Numerik-Aufgabe 2: Proteinfaltung

Digitale Vorstellung und Besprechung der Ergebnisse:
Do 09. Januar 2024 in der Vorlesung

Alle bekannten Proteine werden aus einem Satz von 20 verschiedenen Aminosäuren zusammengesetzt. Die Abfolge der N Aminosäuren ist die Primärstruktur des Proteins. In ihrer biologischen Umgebung (Lösung) falten sich die Proteine sehr schnell und eindeutig zu einem biologisch aktiven ‘Knäuel’ zusammen, der sog. Tertiärstruktur. Wie dieses komplexe Optimierungsproblem in der Natur so effektiv gelöst wird, ist bisher weitgehend unbekannt.

Der Faltungsprozess soll an einem einfachen Proteinmodell mit $N = 30$ Aminosäuren und ‘simulated annealing’ diskutiert werden. Das Protein wird als sich nicht schneidender ‘Faden’ auf einem Quadratgitter modelliert. Gitterpunkte stellen Aminosäuren dar, die mit festem Abstand (Gitterkonstante) und Winkel ($\pm 90^\circ$ bzw. 0° Grad) miteinander verbunden sind. Damit wird den starken kovalenten Bindungen zwischen Aminosäurepaaren Rechnung getragen. Zwei benachbarte, nicht kovalent gebundene Aminosäuren wechselwirken durch van-der-Waals-Kräfte, Wasserstoffbrücken und über das Lösungsmittel miteinander.

Im Grundzustand, d. h. in der gesuchten Tertiärstruktur, ist die Energie, gegeben als Summe

$$E = \sum_{\langle m,n \rangle} \delta_{m,n} J_{A(m),A(n)}$$

mit

$$\delta_{m,n} = \begin{cases} 1 & \text{falls } |\mathbf{r}_m - \mathbf{r}_n| = 1 \quad ; \quad |m - n| > 1 \\ 0 & \text{sonst} \end{cases}$$

über die Wechselwirkungsenergien all dieser Paare minimal. Dabei bezeichnet $A(m)$ die Aminosäure am Platz m . Die effektive Wechselwirkung zwischen den Aminosäuren i und j mit $1 \leq i, j \leq 20$ sei J_{ij} .

1. Bei hohen Temperaturen spielen die Wechselwirkungsenergien J_{ij} keine Rolle. Erzeugen Sie daher für eine gegebene Primärstruktur die Tertiärstrukturen des Proteins als ‘self-avoiding random walk’. Diese werden als Ausgangskonfigurationen für spätere Simulationen benötigt.

2. Initialisieren Sie die Wechselwirkungsmatrix durch Zufallszahlen y , d. h.

$$J_{ij} = J_{ji} = y$$

Die Wahrscheinlichkeitsverteilung der Zufallszahlen y soll einer Gauß-Kurve

$$p(y) = \frac{1}{\sqrt{\pi}} \exp \left[- (y + 3)^2 \right]$$

folgen. Welcher Verteilung folgen die Eigenwerte der symmetrischen 20×20 -Matrix J_{ij} ?

3. Die Bewegung eines Moleküls in einer Lösung wird ebenfalls durch ‘random walks’ beschrieben. Lassen Sie ein Teilchen auf einem Quadratgitter laufen. In jedem Schritt springt das Teilchen um eine Gitterkonstante, die Richtung wird zufällig ausgewählt. Alle Richtungen sind gleich wahrscheinlich. Wie hängt die mittlere quadratische Position des Teilchens (es startet am Ursprung) von der Zeit (Schrittzahl) ab? Was ändert sich, wenn sich der Weg des Teilchens nicht schneiden darf (‘selfavoiding random walk’)?

4. Untersuchen Sie das Problem der Metastabilität der Proteinfaltung.

Lassen Sie ein Protein mit gegebener Primärstruktur sich bei einer Temperatur von $T = 1$ falten und stellen Sie die Energie als Funktion der Zeit (Monte-Carlo-Schritte) dar.

Führen Sie zwei oder mehrere Simulationen durch und vergleichen Sie die Strukturen der verschiedenen gefalteten Zustände.

Schätzen Sie die Größe der Energiebarriere zwischen den verschiedenen metastabilen Zuständen ab.

5. Durch graduelles Abkühlen (‘annealing’) lassen sich schneller optimale Proteinstrukturen finden.

Simulieren Sie die Faltung des Proteins erneut, indem Sie mit $T = 10$ starten und langsam auf $T = 1$ abkühlen. Wie vergleicht sich die zeitliche Entwicklung der Energie mit der aus Aufgabe 4?

Führen Sie ebenfalls zwei oder mehrere Simulationen durch. Stellen Sie die Energie und den geometrischen Abstand der Enden des Proteins als Funktion der Temperatur dar. Dabei soll über mehrere Monte-Carlo-Schritte bei fester Temperatur (wieviele?) gemittelt werden, bevor die Temperatur wieder geändert wird.

Kühlen Sie zu tieferen Temperaturen $T < 1$ ab.

6. Studieren Sie den Einfluss der Wechselwirkungsenergien J_{ij} auf den Faltungsprozeß. Wählen Sie dazu

$$J_{ij} = \epsilon = -3$$

und berechnen Sie die Temperaturabhängigkeiten wie in Aufgabe 5.

Berechnen Sie zusätzlich die spezifische Wärme

$$C = \epsilon^2 \frac{\langle n^2 \rangle - \langle n \rangle^2}{NT^2} \quad ,$$

wobei n die Anzahl der Nächsten-Nachbar-Kontakte der Aminosäuren sind und $\langle \dots \rangle$ den Mittelwert über viele Monte-Carlo-Schritte bei fester Temperatur bedeutet. Unter einem Nächster-Nachbar-Kontakt versteht man Paare von Aminosäuren, die miteinander Wechselwirken, d. h. deren Positionen die oben angegebenen Bedingungen erfüllen.

Wie sieht die innere Energie E als Funktion von $\langle n \rangle$ aus (analytischer Ausdruck)?

Wie lange braucht man jetzt, um bei $T = 1$ einen metastabilen Zustand zu finden?

Wie verändert sich diese Zeit (Anzahl der Monte-Carlo-Schritte), wenn die Matrixelemente der Wechselwirkung gegeben sind durch

$$J_{ij} = \pm 3$$

mit zufälligem Vorzeichen? (Dies soll andeuten, dass einige Aminosäuren von Wassermolekülen angezogen andere hingegen abgestoßen werden.)

Simulierte Proteinfaltung als detailliertes Gleichgewicht in aller Kürze

Das Protein wird, wie eingangs beschrieben, als ‘Faden’ auf einem Quadratgitter modelliert. Eine neue Tertiärstruktur bei gegebener Temperatur T wird in drei Schritten gefunden:

1. Eine Aminosäure aus dem Protein, d. h., ein Platz m wird zufällig ausgewählt.
2. Eine neue mögliche Position dieser Aminosäure wird zufällig ausgewählt: Dafür kommen hier maximal zwei Plätze in Frage, da die Bindungslänge konstant ist. Anschliessend wird die Energieänderung ΔE berechnet.
3. Die lokale Faltung wird akzeptiert falls $\Delta E < 0$. Andernfalls (für $\Delta E > 0$) wird sie nur mit einer Wahrscheinlichkeit $P \sim \exp(-\Delta E/T)$ angenommen (‘detailed balance’).

Alles zusammen macht einen Monte-Carlo-Schritt aus.