

Theoretische Chemie 1999

Molekulardynamik-Simulationen von Biomolekülen widmen sich mittlerweile Konformationsänderungen bis hin zur Proteinfaltung. Solvationen und Enzymreaktionen werden durch Kombination mit quantenmechanischen Verfahren behandelbar, und auch detaillierte Beschreibungen der Reaktionsdynamik kleiner Systeme sind routinemäßig möglich. Quanten-Monte-Carlo-Methoden sind für Elektronenstruktur- und Kerndynamikrechnungen mit vielen gekoppelten Schwingungen im Einsatz.

Molekulardynamik-Simulation von Biomolekülen

Zwischen der dreidimensionalen Struktur eines Biomoleküls, wie es im Lehrbuch oder auf dem Bildschirm erscheint, und seiner Funktion klafft eine Lücke: Unser heutiges Bild der molekularen Biologie ist geprägt von informativen, aber statischen Bildern der Biomoleküle einerseits und Darstellungen der dynamischen, vielstufigen, meist zyklischen Reaktionsschemata, die die Abläufe im Funktionszyklus der Enzyme zeigen, andererseits. Enzyme werden zum Teil in atomarer Auflösung suggestiv so dargestellt, dass sie sich uns lediglich als eine Art „Gehäuse“ katalytischer Zentren präsentieren. Tatsächlich sind jedoch weder die Struktur noch die Funktion von Enzymen oder anderen Biomolekülen zu verstehen ohne Betrachtung ihrer Dynamik. Biologische Makromoleküle sind durch nichtkovalente Kräfte strukturierte Kettenmoleküle, deren Struktur anharmonisch fluktuiert und erhebliche Strukturzyklen durchlaufen kann, die eng mit der Funktion verknüpft sind. Die Dynamik der Biomoleküle reicht von lokalen Fluktuationen bis zu erheblichen Konformationsänderungen im Funktionszyklus. Einen tiefen Blick in die „Dynamiklücke“ erlaubt die klassische Molekulardynamik-Simulation (MD) – sie wendet sich, unterstützt von neuen Computerstrukturen, den großen Bewegungen der Biomoleküle zu – sogar unter Einschluss der nativen Umgebung wie der Lipidmembran. In Verbindung mit quantenmechanischen Methoden trägt sie so zur Aufklärung von Enzymreaktionen bei.

Die Molekulardynamik-Simulation beschreibt Prozesse auf atomarer Ebene mit der klassischen Mechanik, auch Molekülmechanik (MM) genannt. Kräfte werden von empirischen Potentialen abgeleitet und ergänzt durch Terme, die Druck und Temperatur der Umgebung vermitteln. MD liefert unter anderem Trajektorien von Atompositionen und Energien über die Simulationsdauer. Für schnelle Prozesse lassen sich daraus Gleichgewichts-Mittelwerte berechnen. Für langsame Vorgänge sind die Trajektorien als Stichprobe zu werten, die eine erste Vorstellung vom Ablauf eines Prozesses liefert.

Fragen an die Proteinstruktur

Die Untersuchung von Biomolekülen mit MD konzentriert sich im wesentlichen auf Proteine und Peptide. Die Protein Data Bank, eine Sammlung von mittlerweile 11 000 dreidimensionalen Strukturen, birgt eine Menge von Problemen, denen sich die Simulation nacheinander zugewandt hat. Gleich die erste Proteinstruktur des Transportproteins Myoglobin ließ zum Beispiel einen offenen Kanal hin zum Bindeort des Sauerstoffs, dem Hämeisen, vermissen. Heute weiß man, was den Weg scheinbar verstellt hat: die schnelle Rotationsfluktuation (ψ -Bereich) von Seitenketten, die sich im Temperaturfaktor der Atompositionen widerspiegeln. Diese Fluktuationen waren seit 1977 Gegenstand der ersten Moleküldynamiksimulationen. Sie sind offenbar ein weiter verbreitetes Phänomen: In Loop-Bereichen, z.B. beim Transmembranprotein Bakteriorhodopsin, und an den Termini sind die Fluktuationen oft so groß, dass den Atomen kein definierter Bereich zugeordnet werden kann. Hier liefert die MD fehlende Strukturen und einen Eindruck in die Proteindynamik, der bei der Interpretation von Experimenten hilfreich ist.

Strukturänderungen durch Substratmoleküle verweisen auf funktionelle Konformationsänderungen, die vielfältige Ursachen haben können und von höchstem Interesse sind; bei Substratanlagerung ist zum Beispiel immer ein Induced Fit anzunehmen. Bei G-Proteinen etwa bedingt die katalytische Dephosphorylierung des Substrats Guanosintriphosphat (GTP) eine Konformationsänderung, die für die Signalweiterleitung verantwortlich ist. Effektoren behindern oder beeinflussen durch Andocken die katalytische Funktion. Auch die strukturbildende Wirkung von Metallionen gehört in diesen Bereich.

Ein weiteres großes Thema sind natürlich die Grundlagen, die hinter der Entstehung der nativen Struktur selbst stehen: Die spontane, mediumabhängige Faltung/Entfaltung von Proteinen ist die größte beobachtbare Konformationsänderung. Faltungsintermediate entziehen sich oft einer Strukturaufklärung, auch wenn sie ansonsten experimentell gut zu charakterisieren sind. Das gilt zurzeit auch noch für die meisten Transmembranproteine, die als Rezeptoren von großem Interesse sind.

Faltung und Entfaltung von Proteinen

Die Vorhersage ganzer nativer Proteinstrukturen ist zurzeit noch ein Anwendungsbereich vereinfachter Modelle und empirischer Ad hoc-Verfahren. Dennoch konnten mit MD zuletzt unter großem Aufwand spontane Faltungsereignisse bei einer kleinen Proteindomäne sichtbar gemacht werden: In einer Simulation über eine Mikrosekunde – der bisher längsten Simulation eines Peptids in Lösung – konnten Y. Duan und P. A. Kollman zeigen, wie in ihrem simulierten Protein nach schneller „Burst-Phase“ mit hydrophobem Kollaps Helixbildung eintrat und ein langsames Rearrangement der Tertiärstruktur einsetzte, das in die Nähe der (bekannten) nativen Konformation führte.¹⁾ Diese scheint von einer Reihe benachbarter Strukturen durch eine Entropiebarriere (Funnel) getrennt zu sein, wie sie auch aufgrund von Experimenten und Modellbetrachtungen gefordert wird. Mit der reversiblen Bildung von Sekundärstrukturen beschäftigten sich auch X. Daura et al. in einer Serie von 50-Nanosekunden-Simulationen eines Peptids in Methanol. Dabei beobachteten sie eine schnelle Einengung des riesigen Konformationsraums auf wenige Cluster und das Auftreten verschiedener Faltungspfade, deren Gleichgewichte sie bei Temperaturen unter und über dem Schmelzpunkt bestimmen konnten.²⁾ Die Struktur der stabilen Konformation stimmte mit experimentellen Daten gut überein – sicher ein Beweis für die Qualität heutiger empirischer Kraftfelder (hier GROMOS96).

Diskutiert wird noch, ob auch die Simulation des der Proteinfaltung entgegengesetzten Prozesses, der Entfaltung, Information über die Pfade der Proteinfaltung liefern kann. Die Entfaltung ganzer Proteine wurde in der Vergangenheit in eher experimenteller Weise untersucht, wie zum Beispiel durch die Simulation extremer Temperaturen. Bei sorgfältiger Analyse kann man dabei diejenigen Fluktuationen selektieren, die nicht der Funktion dienen, sondern die Entfaltung einleiten.³⁾ Die volle Entfaltung ist dabei allerdings, wie viele andere Prozesse von biologischer Bedeutung, selbst in langen Simulationen noch nicht zu beobachten. Dies hat zur Entwicklung von Varianten der MD geführt, die im Folgenden kurz dargestellt sind.

Dynamik mit äußerer Kraft

Wie sich ein einzelner Molekülkomplex unter dem Einfluss einer Streckkraft verhält – bis hin zum Abriss – lässt sich experimentell mit dem Atomkraftmikroskop untersuchen. Aber erst die MD kann erklären, welche Wechselwirkungen den verschiedenen Phasen der Trennung zugrunde liegen. Bei der ersten derartigen Anwendung auf Streptavidin-Biotin zeigten sich fünf Phasen, bei denen zunächst die dominierenden Wasserstoffbrücken verschwinden; anschließend werden Wasserbrücken ausgebildet, die schließlich ebenfalls reißen.⁴⁾ Ein komplizierteres Muster ergab die Trennung des Dinitrophenyl-Haptens eines monoklonalen Antikörperfragments.⁵⁾ Die Variation der Geschwindigkeit, mit der die simulierte Trennung vollzogen wird, erlaubt die Separation von Reibung und aktivierten Prozessen. Auch die Extrapolation vom simulierten ns-Bereich in den experimentellen ms-Bereich wird dadurch offenbar möglich. Simulationen, bei denen der Abstand zwischen den Termini bis hin zur Entfaltung vergrößert wird, geben einen Einblick in die Stabilität von Strukturelementen und Hinweise auf die entscheidenden Wechselwirkungen, die von Fall zu Fall variieren.⁶⁻⁸⁾ Ob aber solcherart erzwungene Prozesse mit der spontanen Faltung/Entfaltung korrelieren, bleibt noch offen. Zumindest macht die Simulation Energiebarrieren aktivierter Prozesse sichtbar, die im Experiment nicht quantitativ erfasst werden: Messungen liefern Barrieren maximaler Kraft im Gegensatz zu den Maxima der freien Energie, die für spontane Prozesse entscheidend sind.

Dynamik entlang Reaktionskoordinaten

Ein Konzept zur Behandlung von Konformationsübergängen ist die Benutzung einer geeigneten „a priori“ Reaktionskoordinate r . Sie ist so zu wählen, dass bei $r = r_{\text{start}}$ die Startstruktur, bei $r = r_{\text{ziel}}$ die Zielstruktur vorliegt und man annehmen kann, dass sich r (im Mittel) beim Konformationsübergang monoton ändert. Im Laufe einer MD-Simulation wird r durch einen Zwang kontinuierlich in r_{ziel} überführt, während alle anderen Koordinaten des Moleküls und der Umgebung frei relaxieren und fluktuieren können. Die Bezeichnung dieses für die Analyse von Reaktionen entwickelten Konzepts lautet TMD (Targeted MD); als Reaktionskoordinate dient der mittlere Abstand, der die Atome von ihren Positionen in der Zielstruktur trennt. Dieser Abstand wird im Laufe der Simulation langsam auf null reduziert. TMD-Simulationen gehören zum Bereich der „Constraint“ MD, die mit Zwangsbedingungen arbeitet.

Das Konzept funktioniert offenbar: Seine Anwendung auf ein G-Protein, Hras-p21, ergab einen Vorschlag, wie dieses

Protein zwischen seiner aktiven und inaktiven Konformation umschaltet und so zur Signaltransduktion fähig ist (Abbildung 1). Die Analyse der Simulationsdaten gibt im Detail Einblick in die Sequenz der Mikroübergänge und die Energieprofile.⁹⁾ Auffällige lokale Diederwinkelrotationen wiesen einige Residuen als „Scharniere“ aus, und eine Analyse der beteiligten Energieprofile zeigte, welche dieser Bewegungen frei oder gehemmt sind. Die Autoren konnten mittlerweile auch experimentell nachweisen, dass ihre TMD-Simulationen ein qualitativ zutreffendes Bild des Reaktionspfades liefern: Der Einbau von Glycin an die Stelle gehemmter Scharniere beschleunigt den Konformationsübergang um zwei Größenordnungen, der Einbau voluminöser Residuen an die Position freier Scharniere führt zu einer beträchtlichen Verlangsamung.^{10, 11)} Ein interessanter Aspekt der Simulationen mit festgelegter Variation der Reaktionskoordinate ist übrigens die Trennung der Simulationszeit, die zum Durchfahren der Reaktionspfades und der Bestimmung des Energieprofils aufgewendet wird, von der Prozesszeit: Man kann eine Reaktionsbarriere in einer Zeit abtasten, die kürzer ist als die Zeitdauer der spontanen Reaktion.

Für die Simulation der Proteinfaltung hat sich in einigen Experimenten der Trägheitsradius als geeignete, weil monoton wachsende Reaktionskoordinate erwiesen. Seine Benutzung in einer Constraint MD erlaubt die Untersuchung der Entfaltung auch an „schwierigen“ Proteinen – unter Randbedingungen, unter denen eine spontane Entfaltung nicht beobachtbar ist. Ein Beispiel ist das experimentell gut charakterisierte Apomyoglobin, das sich bei pH 4 zu einem stabilen Molten Globule entfaltet, der sich nicht kristallisieren lässt. In üblichen MD-Simulationen wollte die native Struktur selbst bei höherer Temperatur keine Entfaltung zeigen. Erst unter dem Zwang eines wachsenden Trägheitsradius findet man den erwarteten Übergang zu einer Struktur mit großem hydrophobem Kern und einer kleineren gestreckten Domäne (Abbildung 2).¹²⁾ Bemerkenswert ist auch in diesem Fall die gute Übereinstimmung mit experimentellen Befunden: Die Abstandsverteilung der Röntgen-Kleinwinkelstreuung, die Helizität und das Profil der Deuteriumaustausch werden gut reproduziert.

Membranen und Membranproteine

Membranproteine sind biologisch von größtem Interesse als Rezeptoren, Pumpen und Poren. Ihre Struktur ist aber nur in wenigen Fällen bekannt. Reine Transmembranproteine sind kaum zu kristallisieren – nur für Bakteriorhodopsin gibt es hochaufgelöste Strukturen –, so dass der Simulation eine wichtige Rolle zukommt. In der Vergangenheit waren nur Vakuumsimulationen möglich, inzwischen kann aber auch das komplizierte System aus dem eigentlichen Protein (stets mit internen Wassermolekülen), Wasser und der Lipiddoppelschicht in aufwendigen Rechnungen behandelt werden.¹³⁾ Die Groninger Arbeitsgruppe um H. J. C. Berendsen zum Beispiel hat nach einer Reihe von Vorstudien zuletzt den Alamethicin-Kanal, ein Bündel von sechs Helices, auf Stabilität, pK_A interner Residuen und die Dynamik des internen Wassers untersucht.¹⁴⁾ Einer ihrer interessanten Befunde ist, dass die dipolaren Helices des Kanals durch antiparalleles ausgerichtetes internes Wasser stabilisiert werden. Mit einem Kanalprotein hat sich auch I. T. Arkin beschäftigt: ihm gelang eine erfolgreiche Strukturermittlung, indem er in seinen Simulationen gewisse Orientierungsdaten als Zwang (hier vom Typ Restraint) berücksichtigte.¹⁵⁾ Diese Daten hatte er aus dem IR-Dichroismus ortsspezifisch mutierter Residuen ermittelt. Ähnlich geschieht auch die Verwertung von Abstandsdaten aus der 2D-NMR. Mit derartigen Methoden könnte man möglicherweise auch die große Menge von Strukturvorschlägen für Transmembranproteine verwerten und verbessern, die in der Vergangenheit auf der Grundlage von Hydrophobizität und ähnlichen Daten erstellt wurden.

Hybridmethoden

Die Kräfte, die in der MD wirken, werden üblicherweise aus einem empirischen Potential für die Atome abgeleitet. Bei der „Ab-initio“-MD, wie sie z.B. zur Untersuchung der Solvatisierung von Glucose in Wasser mittels Dichtefunktionaltheorie benutzt wurde, werden die Kräfte dagegen quantenmechanisch unter Einschluss elektronischer Freiheitsgrade berechnet.¹⁶⁾ Der hierzu nötige enorme Rechenaufwand verhindert leider immer noch die Anwendung auf interessante Fragestellungen aus dem Inneren der großen Biomoleküle. Um diese Probleme dennoch angehen zu können, wurden schon vor zwanzig Jahren QM/MM-Hybridmethoden eingesetzt: dabei wird ein kleines Teilsystem (mit allen Freiheitsgraden) quantenmechanisch, die Umgebung dagegen mit klassischer Molekülmechanik (MM) behandelt. Dies ist sinnvoll zum Beispiel bei Enzymreaktionen: hier gibt die Einbettung des aktiven Zentrums in das Enzym ein realistischeres Energieprofil. So lässt sich die Proteinumgebung nicht nur für einen vorgegebenen Reaktionspfad optimieren: Eine Optimierung des Pfades selbst wird möglich, mittlere Kräfte im fluktuierenden Protein und somit Profile der freien Energie werden berechenbar. Auf diesem Niveau wurde zuletzt die Amidhydrolyse in Trypsin behandelt.¹⁷⁾ Die erfolgreiche Berechnung von freien Aktivierungsenergien steht und fällt allerdings mit einer aufwendigen Behandlung des QM-Teils: Hier sind Ab-initio-Methoden wie HF und MP2/6-31+G* erforderlich. Wer sich auf semiempirische oder empirische Valence-Bond-Verfahren, die auch heute noch in der Literatur zu finden sind, verlässt, erhält seine Ergebnisse zwar schneller, aber auf Kosten der Genauigkeit. Einen Fortschritt bedeutet auch die Kombination mit MD anstelle einer Energieminimierung: Verfahren ohne explizite Dynamik erlauben zwar die Bestimmung des Reaktionspfades und des Übergangszustands und liefern ein Profil der Energie, nicht aber der freien Energie.¹⁸⁾ In der Entwicklung sind auch

Verfahren, welche die „QM-“ und „MM-Atome“ verknüpfen; nicht zuletzt von ihnen hängt die Qualität der Rechnung ab.

Fortschritte bei den QM/MM-Methoden helfen auch bei der Reproduktion (und somit Interpretation) experimenteller Vibrationsspektren. Mit einer Kombination von DFT und MD haben zum Beispiel Eichinger et al. einen methodischen Ansatz zur Berechnung von IR-Spektren prosthetischer Gruppen in Proteinen präsentiert.¹⁹⁾ Eingebettete Chromophore oder auch andere Atomgruppierungen zeigen unter dem Einfluss der Proteinumgebung Spektren, die stark von denen im Vakuum oder in Lösung abweichen. Sie sind daher geeignete Reportergruppen für die funktionellen Konformationsänderungen des Proteins. Hier könnte die MD – über die Strukturoptimierung hinaus – die inhomogene Linienverbreiterung als Folge der Dynamik liefern.

Schnellere Computer

Das große Handikap der MD ist noch immer die Beschränkung der Rechenzeiten auf den vorhandenen Plattformen. Lokale Bewegungen in Biomolekülen, die über Seitenkettenfluktuationen hinausgehen, setzen „Beobachtungszeiten“ oberhalb einer Nanosekunde, Faltungereignisse oberhalb einer Mikrosekunde voraus. Die Faltung/Entfaltung als spontanes, singuläres Ereignis kann daher bisher lediglich bei langen Simulationen an Modellsystemen beobachtet werden. Statistisch signifikante Resultate sind selbst dort kaum zu erhalten. Gleichgewichtssimulationen von biologischen Makromolekülen überstreichen derzeit lediglich einige zehn Nanosekunden.

Wie bereits dargestellt, wächst jedoch die Vielfalt der Methoden, die in geeigneter Näherung auch bei kurzen Rechenzeiten Prozesse längerer Dauer simulieren können. Eine erhebliche Verbesserung bringt zudem neuerdings der Einsatz von Parallelrechnern. Die Leistung von bis zu 256 parallel arbeitenden CPUs ermöglichte bereits die Simulation von Faltungsprozessen.¹⁾ Interessant sind aber auch „Linux-Cluster“ aus vernetzten, preiswerten PC-Prozessoren, deren Rechenleistung rasch ansteigt – sie skaliert in etwa linear bis zu rund 15 CPUs. So entstehen große Rechenkapazitäten an vielen Orten, die der Simulation vermutlich weiteren Aufschwung geben werden.

Sowohl Programme mit quantenmechanischem „Innenleben“ wie auch MD-Programme sind aufgrund ihrer Struktur gut parallelisierbar. Für einige Programmpakete sind schon angepasste Versionen erhältlich, für andere sind sie in der Entwicklung. Als Beispiel für ein klassisches Simulationspaket, dem nun durch Parallelisierung neue „Muskeln“ wachsen, sei hier nur GROMACS von der Abteilung für Biophysikalische Chemie der Universität Groningen erwähnt.²⁰⁾ In absehbarer Zeit werden sicher auch Enzymreaktionen und Rezeptorbindung mit QM/MM-Kombinationen auf schnellen Computern behandelbar sein.

Jürgen Schlitter, Lehrstuhl für Biophysik, Ruhr-Universität Bochum; E-Mail: juergen@bph.ruhr-uni-bochum.de

- 1) Y. Duan, P. A. Kollman, *Science* 1998, 282, 740.
- 2) X. Daura, B. Jaun, D. Seebach, W. F. van Gunsteren, A. Mark, *J. Biol. Mol.* 1998, 280, 925.
- 3) L. D. Crevelde, A. Amadei, R. C. Vanschaik, H. A. M. Pepermans, J. Devlieg, H. J. C. Berendsen, *Proteins* 1998, 33, 253.
- 4) H. Grubmüller, B. Heymann, P. Tavan, *Science*, 1996, 271, 997.
- 5) B. Heymann, H. Grubmüller, *Chem. Phys. Lett.* 1999, 303, 1.
- 6) H. Lu, B. Israelewitz, A. Krammer, V. Vogel, K. Schulten, *Biophys. J.* 1998, 75, 662.
- 7) H. Lu, K. Schulten, *Prot. Struct. Funct. Gen.* 1999, 35, 453.
- 8) T. Lazaridis, M. Karplus, *Prot. Struct. Funct. Gen.* 1999, 35, 453.
- 9) J. F. Diaz, B. Wroblowski, J. Schlitter, Y. Engelborghs, *Prot. Struct. Funct. Gen.* 1997, 28, 434.
- 10) S. Kuppens, J. F. Diaz, Y. Engelborghs, *Protein Sci.* 1999, 8, 1860.
- 11) J. F. Diaz, M. M. Escalona, S. Kuppens, Y. Engelborghs, *Protein Sci.* 1999, *in press*
- 12) B. Portmann, W. Swegat, J. Schlitter, *Biochemistry*, eingereicht.
- 13) P. C. Biggin, M. S. P. Sansom, *Biophys. Chem.* 1999, 76, 161.
- 14) D. P. Tieleman, H. J. C. Berendsen, M. S. P. Sansom, *Biophys. J.* 1999, 76, 1757.
- 15) A. Kukol, P. D. Adams, L. M. Rice, A. T. Brunger, I. T. Arkin, *J. Mol. Biol.* 1999, 286, 951.
- 16) C. Molteni, M. Parrinello, *J. Am. Chem. Soc.* 1998, 120, 2168.
- 17) R. V. Stanton, M. Peräkylä, D. Barkowies, P. A. Kollman, *J. Am. Chem. Soc.* 1999, 121, 5712.
- 18) V. Moliner, J. Andres, M. Oliva, S. Safont, O. Tapia, *Theor. Chem. Acc.* 1999, 101, 228.
- 19) M. Eichinger, P. Tavan, J. Hutter, M. Parrinello, *J. Chem. Phys.* 1999, 110, 10452
- 20) GROMACS <http://rugmd0.chem.rug.nl/~gmx/>

Abb. 1. **Der Übergang des Signaltransduktionsproteins Hras-p21 zwischen der aktiven und der nichtaktiven Konformation (hier übereinander gelegt). Bei der Simulation mit TMD erweisen sich die hervorgehobenen Aminosäuren als Scharniere für die Bewegung der dazwischenliegenden „Switchregionen“. Die Simulation und spätere Experimente zeigen, dass die grüne Region frei, die rote Region unter Überwindung von Energiebarrieren an den Scharnieren umschaltet.**

Abb. 2. **Entfaltung von Apomyoglobin bei pH 4. Simulation mit einer Zwangsbedingung (Zunahme des Trägheitsradius) ergibt eine Auf-faltung in einen hydrophoben Kern und eine gestreckte kleine Domäne bei relativ hohem Helixgehalt. Diese und weitere Eigenschaften waren aus Experimenten bekannt. Ein Strukturmodell konnte erst die Simulation liefern, weil das Entfaltungsintermediat nicht kristallisiert, sondern aggregiert. Die Abbildung zeigt Zwischenschritte bei der Ausbildung der flexiblen Endstruktur.**