

Herbert Treutlein, Andreas Windemuth und Klaus Schulten

Molecular Design

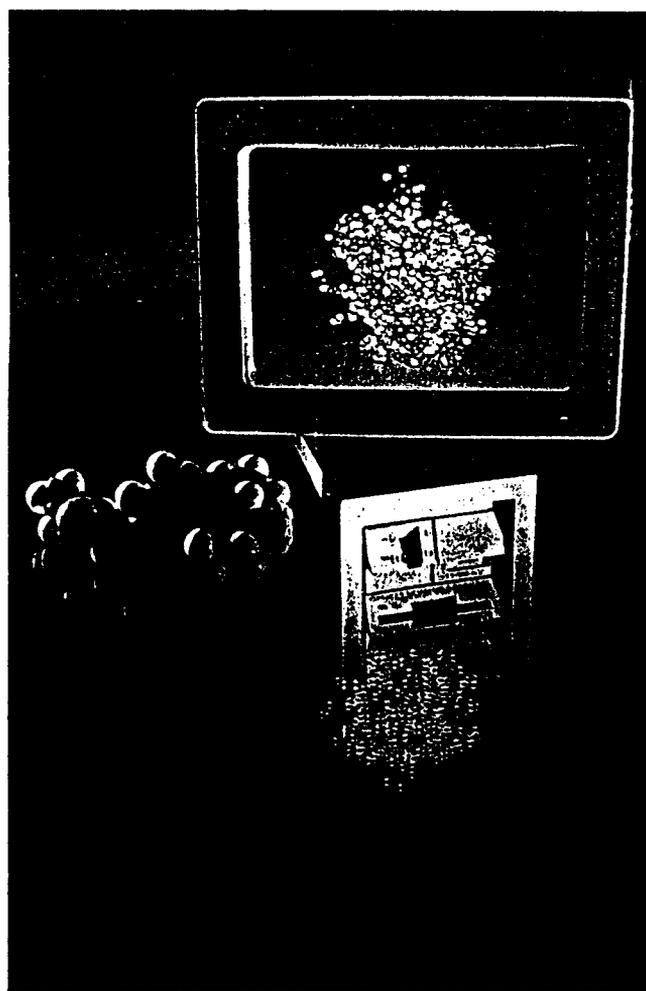
Simulation und Graphik von Biomolekülen

Bekannt ist seit einigen Jahren die Anwendung des Computers für die Entwicklung mechanischer Werkstücke, das sogenannte „Computer Aided Design“ (CAD). Während in der ersten Phase dieser Computeranwendung der Rechner nur den technischen Zeichner ersetzte und Blaupausen in der Fläche darstellte, konnte man in einer nächsten Phase Werkstücke als dreidimensionale Objekte darstellen, indem der Computer es gestattete, das Werkstück von jedem gewünschten Blickwinkel aus zu betrachten. In einer dritten Phase wurde es möglich, eine Serie von Bildern der Werkstücke in verschiedenen Funktionsphasen zu erzeugen, etwa die Kolbenbewegungen eines Verbrennungsmotors in drei Dimensionen nachzuahmen. In der jüngsten Phase des CAD konnten auch komplexe dynamische Abläufe in Großrechnern simuliert und dann graphisch dargestellt werden, eine Methode, die sich etwa in der Flugzeug- und Autoindustrie durchzusetzen beginnt.

Makromoleküle und Computer

In diesem Artikel soll über eine Anwendung ähnlicher Art berichtet werden, bei der anstelle von mechanischen Werkstücken Makromoleküle untersucht, verändert und konstruiert werden können. Ein weiterer Forschungs- und Industriezweig befaßt sich mit der gezielten Herstellung von Makromolekülen, das heißt von chemischen Stoffen, die aus Hunderten bis Zehntausenden von Atomen zusammengesetzt sind, und es erscheint offensichtlich, daß computergraphische Abbildungen hier genauso hilfreich sein können wie im CAD-Bereich. Es hat sich mittlerweile auch eine Bezeichnung für den neuen Computereinsatz

Dieser Beitrag, so scheint es uns von der Redaktion, ist ein Musterbeispiel für den Einsatz von Computern in der modernen Forschung. In diesem Beitrag geht es gleichzeitig um Molekularphysik der aufregendsten Sorte und um deren Umsetzung im Computer: Photosynthese und Computer.



durchgesetzt, nämlich „Molecular Design“. Der Einsatz des Computers in diesem Bereich wird wahrscheinlich eine noch erheblich größere Hilfe als bei der Anwendung von CAD im mechanischen Verarbeitungssektor bedeuten. Dies liegt hauptsächlich an drei Umständen: Erstens handelt es sich bei Molekülen um verschwindend kleine Objekte, die den Augen des menschl-

chen Betrachters nicht einmal mit dem besten Lichtmikroskop zugänglich sind. Ihre Ausdehnung liegt in der Größenordnung eines millionstel Millimeters. Nur falls es gelingt, Kristalle der entsprechenden Makromoleküle herzustellen, kann deren räumlicher Aufbau mit Hilfe von allerdings sehr aufwendigen Röntgenstrukturanalysen sichtbar gemacht werden. Zweitens laufen wichtige molekulare Prozesse oft in sehr kurzer Zeit (d. h. in einer millionstel bis milliardstel Sekunde) ab. Solche Prozesse können natürlich vom Auge nicht aufgelöst werden. Der Computer muß also dem Betrachter eine Welt erschließen, die seinen Sinnen nicht zugänglich ist, und über die er keine Intuition besitzt. Drittens schließlich sind Makromoleküle sehr unregelmäßige Gebilde, eine detaillierte Betrachtung ihres Innen- und Außenraums aus allen Richtungen ist daher besonders wichtig, zumal die Details der räumlichen Anordnung wie Oberflächenformen oder Hohlräume oft entscheidend für die technische, biologische oder medizinische Funktion sind.

Bis vor einigen Jahren wurden etwa zur Unterstützung des räumlichen Vorstellungsvermögens in aufwendigen Arbeitsgängen mechanische Modelle von Makromolekülen im Maßstab von etwa $1:10^9$ gebaut. Derartige Modelle haben viele Nachteile. Werden die beteiligten Atome in ihrer räumlichen

Ausdehnung dargestellt, kann man nicht in die Modelle hineinblicken. Wegen der Massivität derartiger Modelle müssen die Atome starr befestigt werden, d. h. sie können zur besseren Betrachtung nicht einfach herausgenommen oder flexibel verschoben werden. Derartige Modelle sind vielleicht eindrucksvolle Schaustücke im Museum, aber sie können nicht dem chemischen For-

scher und Entwickler bei der Arbeit dienen. Stellt man dagegen die Atome als sehr kleine Kugeln und die chemischen Bindungen als dünne Stäbe (englisch „sticks“) zwischen diesen Kugeln dar, dann bietet das Modell („stick model“) einen gewissen Überblick, ist aber notwendigerweise unrealistisch. Auch derartige Modelle müssen im wesentlichen starr gebaut werden, so daß sie keine Information über Flexibilität des Gebildes geben. Es ist auch schwierig, Teile des Modells bei der Synthesepaltung durch neue atomare Gruppen zu ersetzen oder Information über die auftretenden Verformungsenergien oder elektrostatischen Energien (die Atome tragen meist Ladung) abzuschätzen. „Molecular Design“ siedelt dagegen das Modell im Computer an, macht es damit flexibel in der Handhabung und stellt die gewünschten Aspekte am Bildschirm optisch dar.

Der Computer dient also als Mikroskop zur räumlichen Betrachtung von Molekülen, aber er kann noch die drei folgenden entscheidenden Leistungen erbringen:

Erstens kann er Ordnung in das große Durcheinander von oft zigtausend Atomen bringen, aus denen Makromoleküle bestehen, indem er durch detaillierte Färbung die Zugehörigkeit der Atome zu Untergruppen kennzeichnet oder es dem Betrachter gestattet, ganze Lagen von Atomen durch-

der Temperatur der Materialien entsprechen; je heftiger die Zitterbewegungen sind, um so heißer fühlt sich ein Material an. Wenn etwa Wasser bei 0 °C gefriert und bei 100 °C verdampft, dann werden diese Vorgänge durch die Zitterbewegungen hervorgerufen. Wenn die thermischen Zitterbewegungen im Falle des Wassers wichtig sind, dann auch im Falle von Makromolekülen. Viele biologische Substanzen müssen zum richtigen Funktionieren mit ihren Atomen eine Zitterbewegung ausführen, welche einer Temperatur zwischen 30 °C und 40 °C entspricht.

Drittens kann der Computer neben den geometrisch sichtbaren Eigenschaften wie Ort und Geschwindigkeit auch die geometrisch nicht sichtbaren Eigenschaften der Atome wie lokale elektrische Felder, Flexibilität, elektrische Leitfähigkeiten, Stressenergien, usw. sichtbar machen.

Molecular Design im Kommen

Makromolekulare Synthesen stellen einen wichtigen, aber ungeheuer forschungsintensiven und damit sehr teuren Faktor in Technik und Medizin dar, bei dem der Computer ein hervorragendes Hilfsmittel zur rationalen Synthesepaltung sein wird. „Molecular Design“ wird zwar nicht „Computer Aided Design“ mechanischer oder elektronischer Werkstücke im Umfang des

trachtung sprechen. Erstens ist man bei Einsatz des „Molecular Design“ im Falle größerer Moleküle heute und in absehbarer Zeit noch auf eine Kenntnis der geometrischen Anordnung ihrer Atome im Raum angewiesen, da sich diese Anordnung nur schwer bzw. gar nicht voraussagen läßt. Entsprechende Beobachtungen sind nicht immer möglich (wie oben bereits gesagt), was die Einsatzmöglichkeit des „Molecular Design“ begrenzt, d. h. selbst wenn man die atomare Zusammensetzung eines Makromoleküls genau kennt, ist die Methode nur einsatzfähig, wenn etwa durch Röntgenbeugung an geeigneten Kristallen auch seine Struktur ermittelt worden ist. Zweitens ist die Simulation größerer Molekülstrukturen und die beschriebene Berechnung der geometrisch nicht sichtbaren physikalischen Eigenschaften der Atome heute noch sehr langsam: Um 10^{-12} Sekunden, d. h. ein Millionstel eines Millionstels einer Sekunde Echtzeit eines Makromoleküls und seiner unmittelbaren Umgebung, bestehend aus einigen tausend Atomen, zu beschreiben, benötigt man auf dem schnellsten heute verfügbaren Rechner etwa eine Stunde Rechenzeit. Nun sind 10^{-12} Sekunden allerdings für mikroskopische Systeme keine ganz kurze Zeit (wie wir weiter unten diskutieren werden), aber man müßte die beschreibbare Zeitspanne schon auf bis zu

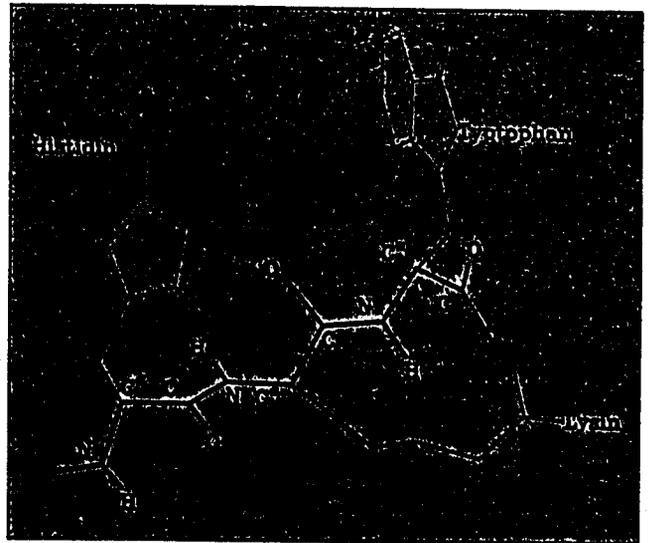
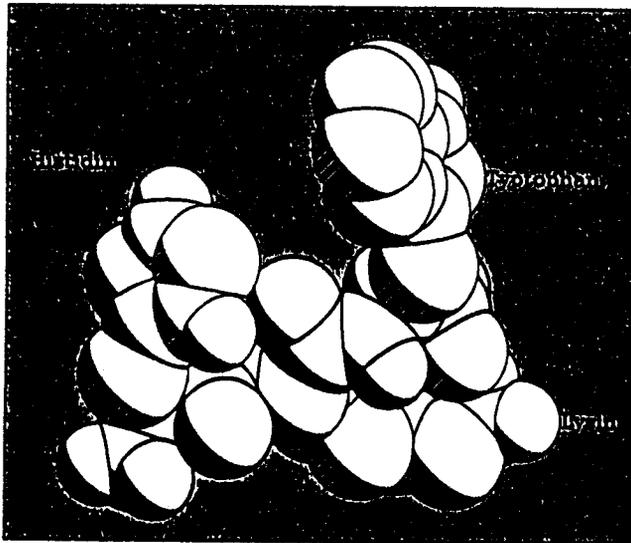


Bild 1. Ein Stück Protein als Kugel-Modell gezeichnet (a) und als „stick model“ (b)

sichtig zu machen, um ins Innere zu schauen.

Zweitens kann er die thermische Bewegung der Atome, die für die Eigenschaften der Makromoleküle oft entscheidend sind, durch Computersimulation darstellen und ihre Konsequenzen abschätzen. Es ist ja von herkömmlichen Materialien wohl bekannt, daß die Zitterbewegung ihrer Atome

technischen Einsatzes überholen, es wird aber wahrscheinlich mindestens ebenso grundlegend die entsprechenden Technologien beeinflussen, da die qualitative Hilfestellung des Computers in diesem mikroskopischen Bereich, der den Sinnen des Menschen bisher verschlossen war, noch wesentlich größer ist. Es gibt aber auch Gründe, die gegen eine so euphorische Be-

10^{-3} Sekunden ausdehnen, um den größten Teil technisch oder medizinisch relevanter Prozesse zu erfassen. Das würde eine Rechenzeit von Hunderttausend Jahren auf einer CRAY 2 bedeuten. Man sieht aber heute schon Chancen (auch Arbeiten in unserem Institut befassen sich mit diesem Problem), durch Verbesserung der Computerprogramme und durch den Einsatz von

Rechnern mit sehr vielen, parallel arbeitenden Rechenwerken Abhilfe zu schaffen. Drittens, das geht schon aus dem zweiten Punkt hervor, ist die Methode wegen des großen Verbrauchs an Rechenzeit und wegen der notwendigen aufwendigen Graphik noch sehr teuer. Nicht jede Firma kann sich eine CRAY ins Haus stellen. In diesem Punkt scheint allerdings eine Abhilfe bereits in greifbare Nähe gerückt: auf dem Markt erscheinen jetzt Rechner für Molekülgraphik und Computersimulation zu einem Preis von 200 000 bis 300 000 DM. So hat die Firma Silicon Graphics bereits einen sog. 4D-Arbeitsplatzrechner (3Ds für die räumlichen Dimensionen der Graphik, ein D (Dimension) für die Zeit) auf den Markt gebracht, der mit einem schnellen RISC-Rechner gekoppelt ist. Die Firmen Digital Equipment und Evans & Sutherland werden Anfang nächsten Jahres mit einem ähnlichen Rechner auf VAX-Basis auf den Markt kommen. Besonders leistungsfähig soll ein Rechner der Firma Stellar werden, welcher einen Minisuperrechner (in der Geschwindigkeit um etwa einen Faktor zehn unter einer CRAY 2) mit einer 4D-Graphik verbindet.

Im folgenden wollen wir die Möglichkeiten, die uns der Computer als Hilfsmittel zum Verständnis der Struktur und Funktion von Molekülen bietet, an einem konkreten Beispiel darstellen.

Ein interessantes biologisches Makromolekül: Das photosynthetische Reaktionszentrum

Das Makromolekül ist in diesem Fall ein Protein, welches in seiner atomaren Zusammensetzung und in seiner Struktur von J. Delsenhofer und H. Michel am Max-Planck-Institut für Biochemie in Martinsried bei München aufgeklärt wurde. Das Protein existiert in der Zellmembran des photosynthetischen Bakteriums *Rhodospseudomonas viridis* und übt dort eine hochinteressante und wichtige Funktion aus: es wandelt die Energie des Sonnenlichtes in ein elektrisches Membranpotential. Dazu enthält das Protein in seinem Innern eine elektronisch leitende Kette von bekannten biologischen Farbstoffmolekülen (Häm-, Chlorophyll-, Pheophytin- und Chinongruppen), die es genauso wie eine Solarbatterie funktionieren lassen: in der Mitte der Kette wird durch das Sonnenlicht ein Elektron angeregt und drifft die Leiterkette in eine Richtung entlang, während die zurückbleibende positive Ladung in die entgegengesetzte Richtung der Kette wandert. Die Trennung von negativer Elektronenladung und positiver Ladung erfolgt außerordent-

lich schnell, eine Rückvereinigung der Ladungen wird vom Protein fast vollständig verhindert, sodaß die Ladungen als Spannung oder Strom abgegriffen werden können. Fast die ganze Biosphäre auf der Erde erhält ihre Energie aus diesem Vorgang, da die Photosynthese in Pflanzen ähnlich wie im Bakterium *Rhodospseudomonas viridis* abläuft. Ziel unserer Forschung ist zu verstehen, wie das Protein die erstaunliche Effizienz beim Einfang der Energie des Sonnenlichtes bewerkstelligt.

Proteine wie das Reaktionszentrum sind spezielle Arten organischer Makromoleküle. Sie bilden die „molekularen Arbeitsmaschinen“ eines Lebewesens, welche eine Vielzahl von mikroskopischen Funktionen ausführen. Alle Proteine sind aus 24 verschiedenen Aminosäuren aufgebaut, die kettenförmig in unterschiedlicher Reihenfolge und Häufigkeit zusammengeknüpft werden. Diese „Aminosäuresequenz“ bestimmt die räumliche Struktur, zu der sich die Molekülkette faltet, und damit die Funktion eines Proteins. In *Bild 1* ist ein Stück Aminosäurekette einmal als sogenanntes *Kugel-Modell (1a)* mit kugelförmigen Atomen realistischer Ausdehnung und zum anderen als „stick model“ (*1b*) gezeichnet. In der letzteren Darstellung wird die Bindungsstruktur besser sichtbar.

Die Kette besteht aus drei Aminosäuren: Histidin, Lysin und Tryptophan. Fett gezeichnet ist in *Abbildung 1b* das „Rückgrat“ der Kette, eine sich wiederholende Folge von Kohlenstoff- (C bzw. C^α), Wasserstoff- (H), Stickstoff- (N) und Sauerstoff- (O) Atomen. Von den C^α Atomen zweigen „Seitengruppen“ ab, durch die sich die Aminosäuren unterscheiden. Die vom „Rückgrat“ des Proteins gebildete räumliche Struktur ist meist wenig flexibel, dagegen können sich die Seitengruppen leichter bewegen, z. B. können sich Ringe drehen oder lange Ketten können sich in Hohlräumen hin und her bewegen. Solche Bewegungen sind oft für die Funktion des Proteins entscheidend. *Bild 2* zeigt eine einfache Computergraphik des gesamten Reaktionszentrums mit seinen etwa 13 000 Atomen. Das Reaktionszentrum besteht aus vier Proteinen, durch unterschiedliche Farben in *Bild 2* gekennzeichnet, die sich wie die Stücke eines räumlichen Puzzles zusammenfügen. Da es zur Zeit noch unmöglich ist, die Dynamik eines Komplexes aus derartig vielen Atomen zu berechnen, haben wir etwa 9000 der Atome entfernt und die Berechnung auf denjenigen Teil des Reaktionszentrums konzentriert, in welchem die ersten entscheidenden Schritte der Photosynthese ablaufen. Der berechnete Komplex aus 3600 Atomen ist in *Bild 3* dargestellt. In diesem

Teil des Proteinkomplexes sind die oben erwähnten, für die lichtinduzierte Ladungstrennung verantwortlichen Farbstoffe (farblich gekennzeichnet) eingelagert. Es sind dies eine Hämgruppe (hellblau; dieser Farbstoff kommt auch im Blut vor und verleiht ihm die rote Farbe), vier Bakteriochlorophyllgruppen (grün) zwei Bakteriopheophytingruppen (violett), ein Chinon (gelb) und ein Eisenatom (große rote Kugel). Gehalten werden diese Farbstoffe durch die Aminosäuren der beteiligten Proteine, deren Atome schwarz oder weiß gekennzeichnet sind. Ein Teil der dem Betrachter zugewandten Proteinatome wurde unsichtbar gelassen, um einen Blick ins Innere zu erlauben.

Die Funktionsträger des Reaktionszentrums

Die Farbstoffe sind die eigentlichen Funktionsträger des Reaktionszentrums. Um sie besser betrachten und ihre Funktion verfolgen zu können, wurden in *Bild 4* alle Atome der Proteinhülle durchsichtig gemacht. Zwei der vier Chlorophyllgruppen sind jetzt rot gezeichnet. Sie bilden einen sog. „Sandwich-Komplex“ und sind für den eigentlichen Einfang des Sonnenlichtes verantwortlich. Die dabei aufgenommene Energie wird verwendet, um über eine der Pheophytingruppen (violett), über das Chinon (gelb) und dann über das Eisenatom (rot) ein Elektron wegzuleiten (in *Bild 4* nach unten) und die zurückbleibende positive Ladung über die Hämgruppe (in *Bild 4* nach oben) fortzuleiten. Nach dem Eisen wartet eine zweite, nicht gezeigte Chinongruppe auf das Elektron. Sie sammelt jeweils zwei Elektronen und verläßt dann das Reaktionszentrum. Dieses zweite Chinon wird übrigens durch Pflanzengifte (Herbizide) verdrängt, die damit dem Bakterium und entsprechend auch Pflanzen die energetische Lebensgrundlage entziehen. Von Interesse ist, daß der erste Reaktionsschritt, welcher die Lichtanregung der (roten) Chlorophylle und die Elektronenübertragung zum (violetten, in *Bild 4* rechten) Pheophytin beinhaltet, innerhalb von drei Picosekunden (eine Picosekunde sind 10⁻¹² Sekunden) erfolgt. Nach der Übertragung hat das Elektron den elektrischen Kontakt mit den (roten) Chlorophyllen praktisch verloren, es kann nicht mehr zurück und muß also seine Energie in den Metabolismus des Bakteriums einspeisen. Da schließlich der Hauptteil der Energie der Biosphäre aus der Photosynthese der Pflanzen, d. h. aus dem beschriebenen Prozeß stammt, kann man sagen, daß die Natur ihre Energieprobleme innerhalb von wenigen Picosekunden löst.

Dies ist aber gerade die Zeltspanne, die wir mit den heutigen Computern verfolgen können. Wir wollen mit ihnen herausfinden, welche Details im Protein die elektrischen Kontaktunterbrechungen zwischen den (roten) Chlorophyllen und dem (violetten) Pheophytin bewirken. Könnte man sich eine relevantere Anwendung von Computern vorstellen als zu verstehen, worauf wir Menschen letztendlich unsere biologische Existenz begründen?

Struktur im Max-Planck-Institut für Biochemie gefunden

Es sollte betont werden, daß die Verfügbarkeit der Struktur des photosynthetischen Reaktionszentrums der einzigartigen Leistung von J. Deisenhofer und H. Michel, zwei Biochemikern am Max-Planck-Institut für Biochemie in Martinsried (b. München), zu verdanken ist. Angesichts der Struktur des Proteins und seiner Funktion als biologische Solarzelle fühlen sich die Autoren dieses Artikels an die Situation erinnert, als durch Laue und Bragg zu Beginn dieses Jahrhunderts ebenfalls mit Hilfe von Röntgenstrahlen die Struktur der Kristalle aufgeklärt wurde. Damals entstand auf der Grundlage gesicherter Kenntnisse der Lagen der Atome in Kristallen die Festkörperphysik mit ihren Konzepten, etwa über die Gitterdynamik und die Elektronenleitung, und ihren spezifischen mathematischen Beschreibungen, etwa der Bändertheorie. Die Kenntnis der Lage der Atome in Proteinen fordern uns heute erneut heraus, für diese eine Physik mit eigenen Konzepten und Beschreibungen zu entwickeln. Die Heterogenität der biologischen Materie läßt dabei erwarten, daß sorgfältige Computersimulationen und graphische Analysen im Zentrum dieses neuen Zweiges der Physik stehen.

Proteindynamik im Computer

Wie bereits angedeutet ist die Dynamik eines Proteins, das heißt die Bewegungen seiner Atome, in vielen Fällen wichtig für seine Funktionsfähigkeit. In Ermangelung von Beobachtungsmethoden stellt die Computersimulation fast das einzige Instrument dar, welches es erlaubt, die Dynamik der Atome zu verfolgen. Die berechneten Bewegungen können unter verschiedenen Gesichtspunkten ausgewertet werden. Hier sei aber nochmals betont, daß die notwendigen Rechnungen sehr aufwendig sind. Nach dem heutigen Stand der Computertechnik können nur Proteine mit bis zu 8000 Atomen simuliert werden. Außerdem ist die berechenbare Zeitdauer nach oben beschränkt. Viele interessante Prozesse las-

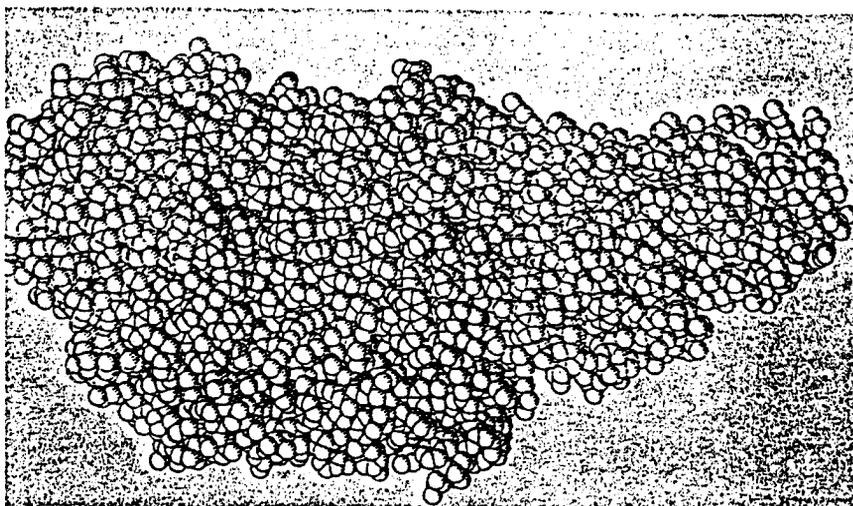


Bild 2. Eine einfache Computergrafik des gesamten photosynthetischen Reaktionszentrums. Diese Solarbatterie besteht aus etwa 13 000 Atomen

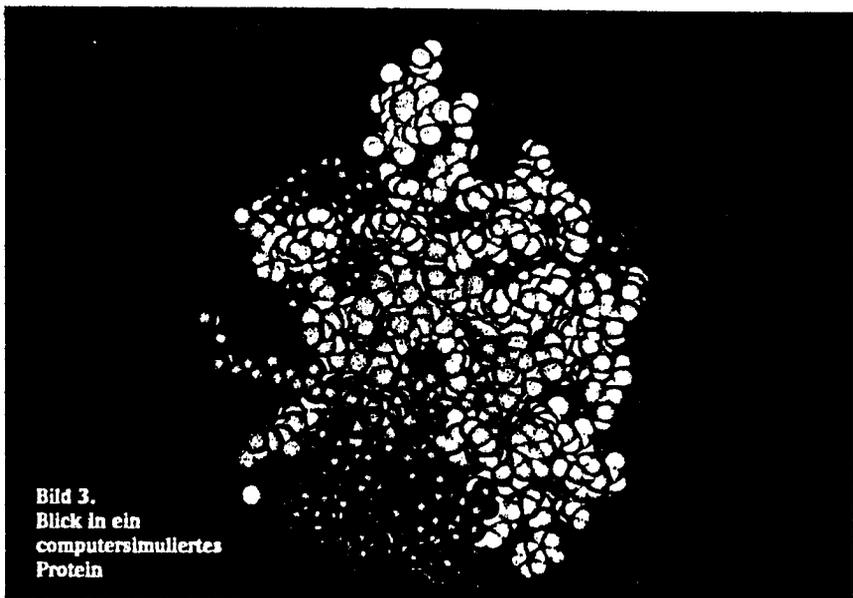


Bild 3. Blick in ein computersimuliertes Protein

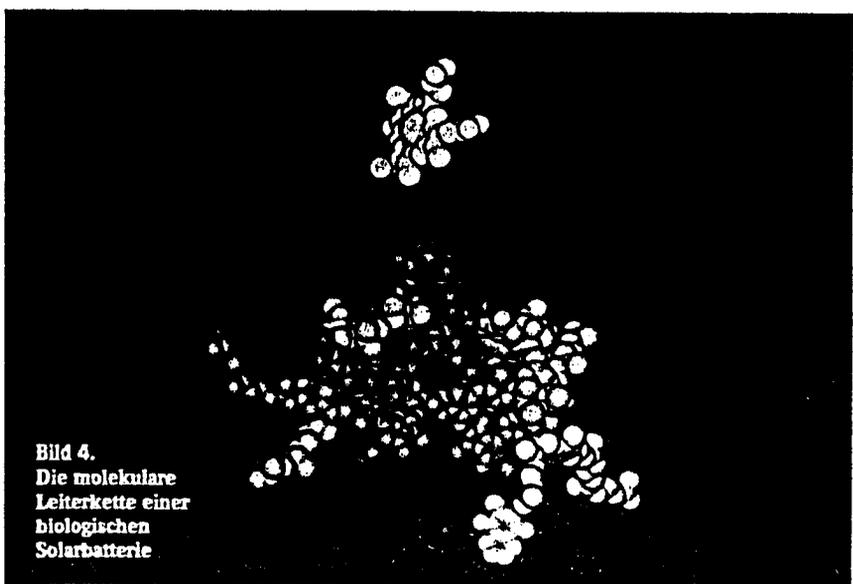


Bild 4. Die molekulare Leiterkette einer biologischen Solarbatterie

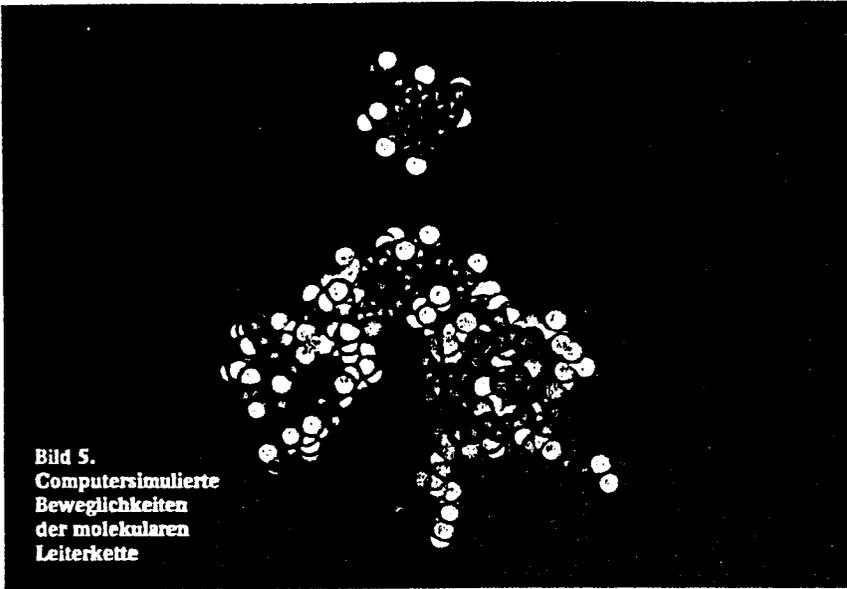


Bild 5.
Computersimulierte
Beweglichkeiten
der molekularen
Leiterkette

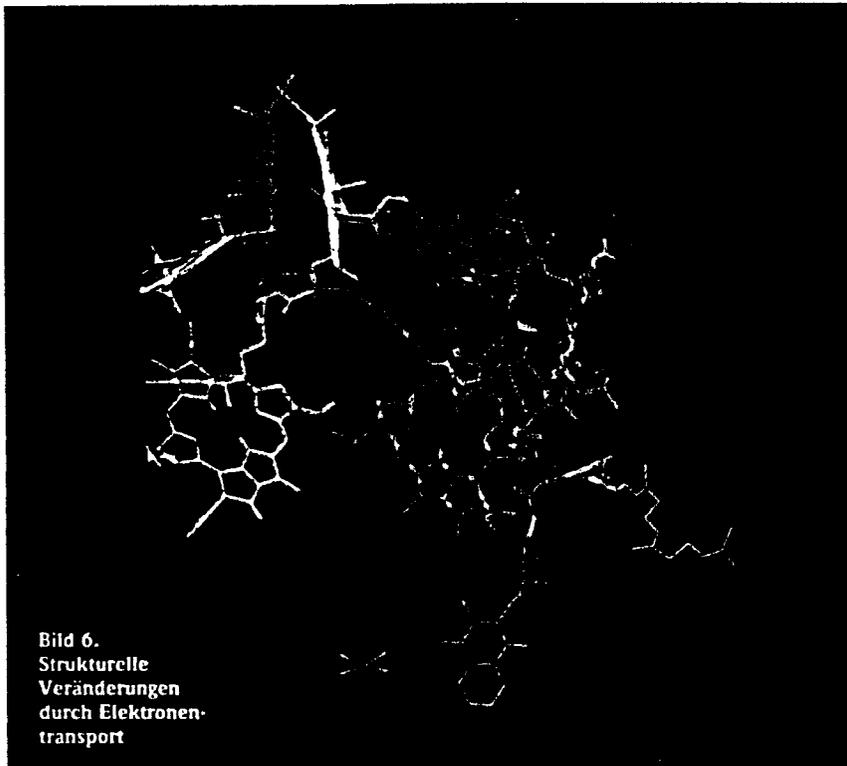


Bild 6.
Strukturelle
Veränderungen
durch Elektronen-
transport

sen sich deshalb heute noch nicht simulieren.

Bei der Simulation geht man von einem sogenannten „klassischen Modell“ des Proteins aus, was bedeutet, daß die Atome in ihrer Bewegung nicht nach den Gesetzen der modernen Atomtheorie (Quantenmechanik), sondern nach den Gesetzen der klassischen Mechanik beschrieben werden. Dabei gelten für die Positionskoordinaten (x, y, z) der Atommassenschwerpunkte die von Newton formulierten Bewegungsgleichungen ($i=1, 2, \dots, N$)

$$m_i \ddot{x}_i = - \frac{d}{dt} E(x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_N, y_N, z_N)$$

$$m_i \ddot{y}_i = - \frac{d}{dt} E(x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_N, y_N, z_N)$$

$$m_i \ddot{z}_i = - \frac{d}{dt} E(x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_N, y_N, z_N)$$

wenn x_i, y_i, z_i die Koordinaten des i -ten Atomes sind und es insgesamt N Atome gibt. Die Funktion E , die von den Koordinaten abhängt, gibt die Gesamtenergie des Protons an.

Eine derartige Beschreibung kann allerdings nur auf die in den Atomkernen konzentrierte Schwerpunktmassen m_i der Atome angewendet werden, die Elektronen bewegen sich nach den Gesetzen der Atomtheorie. Die Elektronen kontrollieren dabei die Bewegung der Atomschwerpunkte, da sie die wesentlichen Beiträge zur Energiefunktion E liefern. Es sind dies die Energien der chemischen Bindung im Protein, wobei die Bindungen im Beispiel von Abbildung 1b etwa durch die „sticks“ dargestellt werden. Die (negativen) Ableitungen der Energie nach den Atomkoordinaten liefern die Kräfte, welche auf die Atome in x, y, z -Richtung wirken. Die Elektronen führen auch zu anziehenden und abstoßenden Kräften zwischen nichtgebundenen Atomen, den sogenannten *van der Waals*-Kräften. Ferner tragen die Elektronen zu den Ladungen der Atome und damit zu *Coulomb*-Kräften zwischen den Atomen im Protein bei. Diese einzelnen Anteile zu den Kräften bzw. Energien lassen sich als Summanden der Gesamtenergiefunktion darstellen:

$$E = E_B + E_{El} + E_e + E_A + E_{v.d.W.} + E_H + E_J$$

Da die einzelnen Anteile der Energie entscheidend für das Verhalten der simulierten Proteine sind, sollen sie kurz vorgestellt werden. Der Energiebeitrag E_B beschreibt die Bindungen benachbarter Atome. Die Bindungen verhalten sich wie „Spiralfedern“, die gedehnt oder zusammengedrückt werden können. Formelmäßig lautet dieser Term:

$$E_B = \sum_{\text{Bindungen}} k_s (b - b_0)^2$$

wobei b die jeweilige Bindungslänge und k_s und b_0 Parameter sind, welche die Bindungskräfte bestimmen.

E_{El} beschreibt die elektrostatische Wechselwirkung der Atome mit den Ladungen q_i und q_j , die einen Abstand r_{ij} voneinander haben:

$$E_{El} = \sum_{\text{Atomepaare } (i,j)} \frac{q_i q_j}{r_{ij}}$$

Obwohl diese Formel sehr einfach aussieht, ist die elektrostatische Kraft die problematischste Wechselwirkung bei der Simulation. Das kommt zunächst daher, daß E_{El} wie man auch in der obigen Formel sieht, eine Summe über eigentlich alle Atompaaire beinhaltet. Die Anzahl aller Atompaaire wird aber leicht sehr groß, wenn das Molekül viele Atome besitzt. Die Anzahl der Paare wächst quadratisch mit der Anzahl der Atome, während die Anzahl der Bin-

dungen nur linear wächst. So kann zum Beispiel ein Molekül mit 66 Atomen 65 Bindungen und 2145 Atompaaire besitzen, in einem Molekül mit 3634 Atomen und 3777 Bindungen gibt es aber schon über 6,6 Millionen Paare. Entsprechend mit der Zahl der Atompaaire steigt aber auch die Rechenzeit, die zum Bestimmen der elektrostatischen Kräfte benötigt wird, d. h. für eine doppelte Zahl von Atomen benötigt man ein Vierfaches an Rechenzeit. Dies würde es aber beinahe unmöglich machen, größere Moleküle zu simulieren. Aus diesem Grund werden nicht die Energiebeiträge aller Atompaaire berechnet, sondern für ein festes Atom werden nur Beiträge zur elektrostatischen Energie von Atomen innerhalb einer Kugel mit einem bestimmten Radius berücksichtigt. Dieser Radius muß so gewählt werden, daß die Summe der Beiträge aller Atome außerhalb der Kugel fast Null ist. Diese Bedingung ist nun leider für die sehr langreichweitige elektrostatische Energie schwer zu erreichen. Dies ist der Hauptgrund, weswegen diese Wechselwirkung problematisch ist. Man versucht dieses Problem durch Korrekturformeln zu umgehen. Im Falle des Moleküls mit den 3634 Atomen wird dadurch die Zahl der Atompaaire auf etwa 300 000 beschränkt. Der Beitrag E_θ zur Gesamtenergie beschreibt Winkelschwingungen im Protein. Die Energie ist Funktion des Winkels θ zwischen jeweils benachbarten Bindungen, d. h. der Winkel zwischen zwei „sticks“ in Bild 1b. Diese Winkel können nicht beliebige Werte annehmen, sondern können nur um einen Wert θ_0 schwingen, eine Schwingung, die dabei ähnlich wie im Falle E_θ beschrieben wird. Formelmäßig ausgedrückt ist E_θ

$$E_\theta = \sum_{\text{Winkel}} k_\theta (\theta - \theta_0)^2$$

E_θ beschreibt die Energie der Dihedral-schwingungen, einer Art von Torsions-schwingungen um die einzelnen Bindungen als Drehachsen, ϕ ist dabei der Torsionswinkel, ϕ_0 die Ruhelage:

$$E_\phi = \sum_{\text{Dihedralwinkel}} k_\phi [1 - \cos(n\phi - \phi_0)]$$

E_{VdW} beschreibt die *van der Waals* Wechselwirkung, die wiederum, wie die oben beschriebene elektrostatische Wechselwirkung, zwischen nicht gebundenen Atompaairen auftritt. Diese Kraft bewirkt, daß sich zwei Atome nicht beliebig nahe kommen können:

$$E_{VdW} = \sum_{\text{Atompaaire}} \left(\frac{A}{r_{ij}^{12}} - \frac{B}{r_{ij}^6} \right)$$

Die Summierung über Atompaaire muß hier im Prinzip genauso wie im Falle der elektrostatischen Kraft durchgeführt werden. Da E_{VdW} aber sehr kurzreichweitig ist, können Kräfte (Energien) zwischen weiter entfernten Atomen ohne Schwierigkeiten vernachlässigt werden.

$$E_J = \sum_{\text{Impropere}} k_J (J - J_0)^2$$

und E_H schließlich beschreiben Kräfte, welche die sogenannte „Chiralität“ und „Planarität“, zwei quantenmechanische Effekte bestimmter Atome, sowie Wasserstoffbrücken, das sind Wechselwirkungen zwischen benachbarten, aber chemisch nicht gebundenen Wasserstoff-, Sauerstoff- und Stickstoffatomen, berücksichtigen. Die Parameter k_B, k_ϕ, \dots , welche die obigen Energiefunktionen enthalten, gewinnt man aus Experimenten und quantenmechanischen Rechnungen an kleinen Molekülen, die den Proteinsegmenten ähnlich sind. Es muß an dieser Stelle betont werden, daß die Repräsentation der Kraftverhältnisse in Makromolekülen durch die beschriebenen Energiebeiträge noch entwicklungsbedürftig sind, und daß in nächster Zukunft noch darum gerungen werden muß, durch Verbesserung der Energiefunktion eine immer realistischere Darstellung im Computer zu erreichen.

Der größte Teil der Rechenzeit bei der Simulation der Dynamik eines Proteins wird für die Berechnung der Energiefunktion bzw. der Kräfte verbraucht. Die numerische Integration der *Newtonschen* Bewegungsgleichungen wird im Rechner relativ rasch erledigt. Dabei hat sich das *Verlet*-Verfahren als numerisch am stabilsten und zeitlich am günstigsten herausgestellt. Bei diesem Verfahren wird für jeweils eine Atomkoordinate, z. B. x_i , der neue Wert $x_i(t+\Delta t)$ zum Zeitpunkt $t+\Delta t$ nach der Formel

$$x_i(t + \Delta t) = 2x_i(t) - x_i(t - \Delta t) + f_i(t)(\Delta t)^2/m_i$$

$$f_i(t) = -\frac{d}{dx_i} E(x_1(t), y_1(t), x_2(t), \dots, x_N(t), y_N(t), z_N(t))$$

bestimmt. Zur Anwendung dieser Formel werden Koordinaten $x_i(t-\Delta t)$ und $x_i(t)$ benötigt, also zu zwei vorausgehenden Zeitpunkten $t-\Delta t$ und t . Deshalb müssen auch zu Beginn der Integration jeweils zwei Koordinatenwerte $x_i(0)$ und $x_i(\Delta t)$ bekannt sein. Es ist aber auch möglich, beim Integrationsbeginn, statt von zwei aufeinanderfolgenden Koordinatenwerten, von dem Koordinatenwert $x_i(0)$ und der Geschwindigkeitskomponente $\dot{x}_i(0)$ auszugehen, da gilt: $x_i(\Delta t) = x_i(0) + \dot{x}_i(0)\Delta t$.

Als Ausgangsdaten bei der Simulation benötigt man also einen Satz von Atomkoordinaten $x_i(0), y_i(0), z_i(0), x_2(0), y_2(0), z_2(0), \dots, x_N(0), y_N(0), z_N(0)$ und einen entsprechenden Satz von Geschwindigkeiten. Den Koordinatensatz erhält man üblicherweise aus den Daten einer Röntgenstrukturanalyse. In Falle des Reaktionszentrums wurden die Koordinaten von J. Deisenhofer und H. Michel geliefert. Die Röntgendaten entsprechen einer mittleren Molekülstruktur, die je nach Auflösung und Güte der Röntgenanalyse mehr oder weniger gut geometrisch optimiert ist. Um starke *van der Waals*-Kräfte aufgrund einer schlechten Geometrie zu vermeiden, was zu großen Störungen zu Beginn der Dynamiksimulation führen würde, wird deshalb die Röntgenstruktur zunächst bezüglich der obigen Energiefunktion E minimiert. Die Anfangsgeschwindigkeiten der Struktur werden entsprechend der *Maxwellschen* Geschwindigkeitsverteilung

$$p(\dot{x}_i) = \sqrt{\frac{m_i}{2\pi k_B T}} \exp\left\{-\frac{m_i \dot{x}_i^2}{2k_B T}\right\}$$

zufällig gewählt. $p(\dot{x}_i)$ gibt die Wahrscheinlichkeit (Häufigkeit) der Geschwindigkeitskomponente \dot{x}_i an, k_B ist die *Boltzmann*-Konstante) zur gewünschten Temperatur T meistens Zimmertemperatur. Mit Hilfe der energieoptimierten Koordinaten und der gewählten Geschwindigkeiten kann nun eine kurze Dynamik-Simulation gerechnet werden, um die kinetische Energie eventuell besser als bei der ursprünglichen Geschwindigkeitswahl auf alle Atome zu verteilen. Am Ende dieser Prozedur stehen relaxierte Koordinaten des Moleküls und stochastische Anfangsgeschwindigkeiten zur Temperatur T zur Verfügung. Mit diesen Daten wird die eigentliche Simulation begonnen, die ausgewertet wird.

Der maximale Integrationsschritt Δt wird durch die schnellsten im System vorhandenen Freiheitsgrade bestimmt. Diese sind im allgemeinen die Bindungsschwingungen und Bindungswinkelschwingungen. Durch sie wird ein Integrationsschritt von einer Femtosekunde, was 10^{-15} Sekunden entspricht, festgelegt. Viele interessante Prozesse dauern aber etwa von einer Millionen Femtosekunden bis zu Sekunden. Solche Vorgänge können bislang nicht berechnet werden. So benötigt man etwa zur Simulation von 1000 Femtosekunden eines Proteins mit 3700 Atomen 800 Sekunden CPU-Zeit einer CRAY X-MP48. Zur Zeit wird in unserem Institut die Möglichkeit untersucht, die schnellen Freiheitsgrade nur in Form einer stochastischen Kraft, einem „Rauschen“ in der Rechnung zu berücksichtigen. Dadurch könnte der Integrations-

schrift um eine Größenordnung auf etwa 10 Femtosekunden verlängert werden. Es ist auch möglich, viele Schritte bei Berechnung der Kräfte und Integration auf mehreren Prozessoren gleichzeitig auszuführen. Somit können Parallelrechner, die zur Zeit entwickelt werden, und die sehr viele unabhängige Prozessoren besitzen, die Simulationen um weitere Größenordnungen beschleunigen.

Mit der Methode der stochastischen Ränder

Um nun aber doch Proteine der Größenordnung des Reaktionszentrums schon jetzt beschreiben zu können, verwendet man einen Trick, nämlich die *Methode der stochastischen Ränder*. Oft ist man ja gar nicht an dem gesamten System interessiert, sondern nur an einem kleineren Teil. So genügt es, sich beim Reaktionszentrum nur den in den Abbildungen 3 und 4 dargestellten Ausschnitt anzuschauen, wenn man nur die ersten beiden Schritte der Elektronenübertragung simulieren will, d. h. die Übertragung vom roten Chlorophyllpaar zum gelben Chinon. Diese beiden Übertragungsschritte benötigen schon etwa 200 000 Femtosekunden. Der erste Schritt vom (roten) Chlorophyllpaar zum (violettten) Pheophytin dauert allerdings nur 3000 Femtosekunden. Also simuliert man statt des gesamten Systems mit 13 000 Atomen nur ein System mit 3700 Atomen. Dabei müssen natürlich die Randatome dieses Proteinausschnitts an ihren Plätzen festgehalten werden, um eine Umstrukturierung des gesamten Systems zu vermeiden. Dies wird durch eine Zwangskraft, die auf diese Atome wirkt, erreicht. Der Einfluß der weggelassenen Atome auf die Atome im Segment wird dadurch simuliert, daß zusätzlich zur Zwangskraft auch noch eine Zufallskraft und eine geschwindigkeitsabhängige Reibungskraft zur schon vorhandenen *Newtonschen Kraft* (siehe oben) addiert wird. Gleichzeitig wird dadurch ein Energiefluß in das System hinein oder hinaus ermöglicht.

Das *Bild 5* zeigt ein Ergebnis einer Simulation über einen Zeitraum von 20 000 Femtosekunden, nämlich die über den Simulationszeitraum gemittelten Atomlagen im photosynthetischen Reaktionszentrum. Das Titelbild stellt dabei wie *Bild 3* Farbstoff- und Proteinatome des gesamten berechneten Molekülsegments dar. In *Bild 5* sind wie in *Bild 4* nur die Atome der Farbstoffe dargestellt (allerdings um 90 Grad um die vertikale Achse gedreht), während die übrigen Atome durchsichtig gehalten wurden. Farblich kodiert sind im Titelbild und in

Bild 5 aber nicht die Zugehörigkeiten der Atome zu verschiedenen Gruppen im Protein, sondern in diesem Fall die Flexibilitäten der einzelnen Atome. Die Flexibilitäten werden durch die Standardabweichungen σ der Atomlagen \vec{r} vom Mittelwert $\langle \vec{r} \rangle$ gemessen

$$\sigma = \sqrt{\langle (\vec{r} - \langle \vec{r} \rangle)^2 \rangle}$$

Blau dargestellte Atome sind unbeweglich, bei ihnen ist σ kleiner als ein halber Atomdurchmesser; gelb und rot dargestellte Atome sind umgekehrt sehr flexibel, bei den letzteren ist σ größer als ein Atomdurchmesser. Das in *Bild 4* rot dargestellte Chlorophyllpaar, welches das Licht absorbiert, ist interessanterweise sehr starr, während die langen Ketten der Farbstoffe sehr beweglich sind. Die starre Geometrie des Chlorophyllpaares ist anscheinend für die Funktion der primären Elektronenübertragung optimal, während die Flexibilität der Ketten an den Farbstoffen die Geschwindigkeit der Elektronenübertragung durch Veränderung der Abstände zu den Nachbargruppen kontrollieren können.

Das Titelbild dieser Zeitschrift ist beispielhaft für „Molecular Design“, da nicht nur die von der Biochemie bekannte Struktur dargestellt wird, sondern durch die farbliche Kodierung der Atome auch die Resultate der Computersimulation veranschaulicht werden. Es wird damit eine Information vermittelt, die ohne den Computer der Biochemie nicht zugänglich wäre.

Computergraphik und Computerexperimente

Durch Beispiele und Bilder unseres Artikels wird die zentrale Rolle der graphischen Darstellung beim „Molecular Design“ und bei der Simulation deutlich. Bei einer solchen Datenfülle – bei der Dynamikrechnung des Reaktionszentrums fallen bei der Simulation von 1000 Femtosekunden Echtzeit des Proteins etwa 1 MByte Daten an – lassen sich viele Ergebnisse nur noch graphisch analysieren. Mit den Zahlenwerten im Koordinatensatz des Moleküls alleine kann niemand unmittelbar etwas anfangen. Erst die graphische Darstellung des Moleküls macht es möglich, effektiv mit der Struktur weiterarbeiten zu können. Um eine gute räumliche Vorstellung des Moleküls zu erhalten, muß man es interaktiv am Bildschirm drehen, verschieben, vergrößern und verkleinern können. Mit den heutigen Graphikstationen ist es auch möglich Moleküle zu editieren, d. h. zusammenzubauen und zu verändern. Mit einer Maus lassen

sich am Bildschirm verschiedene Bauteile des Moleküls anwählen, verschieben oder verdrehen. Außerdem können Bestandteile, zum Beispiel Seitengruppen in Proteinen, verändert werden. Auf diese Weise kann man Gentechnologie am Computer simulieren. Die Auswirkungen der Veränderungen in der Struktur lassen sich dann mit Hilfe von Dynamiksimulationen berechnen. Solche komfortablen Arbeitsmöglichkeiten stellen natürlich hohe Ansprüche an die Hardware-Seite der Graphikstationen. Diese müssen in der Lage sein, viele (das heißt tausende) Vektoren und Polygone interaktiv, und das heißt in sehr kurzer Zeit, nicht nur zu zeichnen, sondern auch räumlich zu transformieren. Die Hardware der Graphikrechner besitzt deshalb speziell zu diesem Zweck konstruierte zusätzliche Prozessoren.

Neben der Gentechnologie am Computer sind auch noch andere Computerexperimente denkbar. Der Chemiker oder Biologe kann seine ganze Phantasie spielen lassen und viele Möglichkeiten ausprobieren, um ein Problem zu lösen. Ein Computerexperiment, das wir am Reaktionszentrum durchgeführt haben sei zum Abschluß noch vorgestellt. Wir untersuchten dabei den Einfluß der Elektronenübertragung vom Chlorophyllpaar (rot in *Bild 4*) zum Pheophytin (violett in *Bild 4*). Den eigentlichen Elektronentransport kann man mit unserem klassischen Modell leider nicht verfolgen. Wir können aber den Einfluß der Ladungverschiebung, der durch die Elektronenleitung verursacht wird, untersuchen. Dazu wird das Modell zweimal simuliert, einmal mit der neutralen Ladungsverteilung und einmal mit der verschobenen Ladung. Die strukturellen Unterschiede dieser beiden Rechnungen sind in *Bild 6* dargestellt. Das „stick mode“ zeigt durch Überlagerung der beiden Strukturen an den Orten der größten Änderung die genaue Art der Verschiebung. Man erkennt, daß das Pheophytin nach dem Elektronentransport in die Richtung des (in *Bild 4* roten) Chlorophyllpaares gezogen wird. Dabei wird die Orientierung dieser drei Farbstoffe zueinander verändert. Diese Veränderung kann z. B. zur Folge haben, daß der Rücktransport des Elektrons vom Pheophytin zum Chlorophyllpaar, der ja die Effektivität der Photosynthese vermindern würde, erschwert wird.

Molekulardynamik auf dem PC

Das Programm *Molecular_Dynamics* ist eine stark vereinfachte Version von professionellen, für *Molecular Design* eingesetzten Programmen. Es kann grob in drei Bereiche

GRUNDLAGEN

```

(.....)
* Berechnung des Potentials und der Kraefte, die auf die Atome
* wirken.
* pos      Koordinaten der Atome.
* force    wird mit den ausgerechneten Kraeften belegt.
* potential: potentielle Energie des Systems.
(.....)
Function potential( var pos , force : Vectorset ):real;
var P : real;
    fi,fx,fy,dfx,dfy,df : real;
    dx,dy,Rij : real;
    i,j : integer;
begin
    P := 0;
    for i:=1 to Atomcount do (* Zuerst die Wechselwirkung mit dem Rand *)
        begin
            fx := f( pos[i].x , Box.a-Atom[i].radius , Box.delta , dfx );
            fy := f( pos[i].y , Box.b-Atom[i].radius , Box.delta , dfy );
            P := P - Eb*fx*fy; (* Dies ist die potentielle Energie *)
            force[i].x := -Eb*dfx*fy; (* Kraft (negative Ableitung von P) *)
            force[i].y := -Eb*fx*dfy;
            for j:=1 to i-1 do (* Wechselwirkung zwischen den Atomen *)
                begin
                    dx := pos[i].x - pos[j].x;
                    dy := pos[i].y - pos[j].y;
                    Rij := sqrt( dx*dx + dy*dy ); (* Abstand der Atome *)
                    P := P + Ek*f(Rij, Atom[i].radius*Atom[j].radius,
                                Atom[i].delta +Atom[j].delta , df);
                                (* Potential *)
                    fi := Ek*df/Rij;
                    fx := fi*dx; (* Kraefte *)
                    fy := fi*dy;
                    force[i].x := force[i].x - fx;
                    force[i].y := force[i].y - fy;
                    force[j].x := force[j].x + fx; (* ...auf beide Atome *)
                    force[j].y := force[j].y + fy;
                end;
            end;
        potential := P;
    end;
(.....)
* Diese Routine vollzieht die numerische Integration der Bewegungs-
* gleichung nach den Verfahren von Verlet. 'x' ist die Lage der
* Atome und dem Parameter 'v' werden die Geschwindigkeiten zuge-
* wiesen. 'dt' gibt die Laenge des Zeitschrittes an. Die Hilfs-
* variable xold speichert die jeweils vorherigen Koordinaten ab.
(.....)
Procedure verlet( var x,v:Vectorset ; dt:real);
var force : Vectorset;
    xi : Vector;
    p : real;
    i : integer;
begin
    p := potential( x , force );
    for i:=1 to Atomcount do
        begin
            xi := xold[i];
            xold[i] := x[i];
            x[i].x := 2*x[i].x - xi.x + (dt*dt/Atom[i].mass)*force[i].x;
            x[i].y := 2*x[i].y - xi.y + (dt*dt/Atom[i].mass)*force[i].y;
            v[i].x := (0.5/dt)*(x[i].x-xi.x);
            v[i].y := (0.5/dt)*(x[i].y-xi.y);
        end;
    end;
(.....)
* Initialisierung der Konstanten und des zufaellig gewaehlten
* Anfangszustands. Hier koennen Aenderungen zum Ausprobieren
* durchgefuehrt werden, z.B. des Zeitschrittes oder der Energie-
* konstanten.
(.....)
Procedure Initialize;
var i,j:integer;

```

```

p:real;
force:Vectorset;
overlap : boolean;
begin
    Atomcount := 3; (* Anzahl der Atome *)
    for i:=1 to Atomcount do with Atom[i] do
        begin
            masse := 1.0; (* Groesse, Masse und Haerte *)
            radius := 1.0; (* der Kugeln. (Es muessen *)
            delta := 0.2; (* nicht alle gleich sein) *)
        end;
        with Box do
            begin
                a := 20.0; (* Groesse und Haerte der *)
                b := 10.0; (* Zelle. *)
                delta := 0.2;
            end;
            Eb := 20.0; (* Energiekonstanten: Hoehe des *)
            Ek := 20.0; (* Potentials ausserhalb der *)
            Dt := 0.02; (* Zelle bzw. innerhalb der Atome *)
            (* Laenge des Zeitschritts; zu *)
            (* grosse Werte ergeben Unsinn *)
            (* zu kleine lange Rechenzeiten *)
            seed := 1234; (* Anfangszustand des Zufalls- *)
            (* zahlengenerators. *)
            for i:=1 to Atomcount do
                begin
                    repeat (* Anfangskordinaten der Atome *)
                        xold[i].x := Random*(Box.a-Atom[i].radius);
                        xold[i].y := Random*(Box.b-Atom[i].radius);
                        overlap :=false;
                        for j:=1 to i-1 do
                            if abs(xold[i].x-xold[j].x)*abs(xold[i].y-xold[j].y)
                                < Atom[i].radius*Atom[j].radius then overlap:=true;
                        until not overlap; (* keine Ueberlappungen *)
                        Velocity[i].x := 5.0*Random; (* Anfangsgeschwindigkeiten *)
                        Velocity[i].y := 5.0*Random;
                    end;
                    p := potential( xold , force ); (* erster Schritt des Integrators *)
                    for i:=1 to Atomcount do (* zur Initialisierung *)
                        begin
                            Position[i].x := xold[i].x + Velocity[i].x*Dt
                                - (0.5*Dt*Dt/Atom[i].mass)*force[i].x;
                            Position[i].y := xold[i].y + Velocity[i].y*Dt
                                - (0.5*Dt*Dt/Atom[i].mass)*force[i].y;
                        end;
                    end;
                end;
            end;
(.....)
* Hauptprogramm. Hier werden einfach 2000 Schritte gerechnet.
* wobei alle 20 Schritte die Graphik aktualisiert wird.
* Steht Graphik nicht zur Verfuegung, koennen die auskommentierten
* Passagen zur numerischen Ausgabe von Positionen und Geschwindig-
* keiten verwendet werden.
(.....)
begin
    Initialize;
    for i:=1 to 100 do
        begin
            clear; (* schreibt beim Atari ST die Schrittnummer *)
            writeln(' ',i-1); (* in die linke untere Ecke. *)
            plot( Position ); (* Graphikaktualisierung *)
            for j:=1 to 20 do verlet( Position , Velocity , Dt );
                (* 20 Integrationschritte *)
            end;
            (* writeln(' Position: ');
            writec( Position );
            writeln(' Velocity: ');
            writec( Velocity ); *)
            (* Ausgabe von Positionen und Geschwindigkeiten (auskommentiert) *)
        end;
    end;
end.

```

gegliedert werden:

1. Berechnung der Kräfte (f und $potential$),
2. Integration der Bewegungsgleichung ($verlet$),
3. Darstellung der Ergebnisse ($writec$ und $plot$).

$initialize$ und $random$ sind allgemeine Hilfsroutinen.

Die Berechnung der verschiedenen Beiträge zur Energie für ein echtes, wenn auch noch so kleines Molekül würde den Rahmen dieses Artikels bei weitem sprengen. Daher wurde ein einfacheres, anschaulicheres Potential verwendet, nämlich das von freien Atomen, die sich in einem 2-dimensionalen Kasten mit elastischen Wänden befinden.

Die Datenstruktur ist wegen des Fehlens von Bindungen recht einfach und übersichtlich. Der Kasten wird durch die zwei Seitenlängen a und b beschrieben, die Atome durch einen Radius $radius$ und eine Masse $mass$. Außerdem sind beide Objekte noch mit einem Attribut $delta$ versehen, welches die Breite der Zone, in der das Potential von Null zum Maximalwert ansteigt, bestimmt. Je kleiner $delta$ ist, desto härter ist das Potential, und um so kleiner muß auch der Zeitschritt für die Integration gewählt werden. Größen wie Position, Geschwindigkeit und Kraft für einzelne Atome werden durch den Datentyp *Vektor* beschrieben, die entsprechenden Größen für alle Atome zusammen durch *Vectorset*. Die Konstanten sind durchweg als globale Variablen definiert, damit sie z. B. zur Ausführungszeit durch $readln$ eingelesen oder sonstwie verändert werden können.

Zwei Energiebeiträge lassen sich unterscheiden: Wechselwirkung der Atome mit der Wand (führt zu Reflexionen) und Wechselwirkung der Atome untereinander (führt zu Kollisionen). Die Energiefunktion ist so beschaffen, daß sich bei der Annäherung zweier Objekte aneinander ab einem gewissen Abstand die Energie kontinuierlich erhöht, dann bei einem etwas kleineren Abstand abflacht und auf einem hohen Niveau konstant bleibt. Dieses Niveau ist durch die Konstanten E_b für Atom-Atom-Wechselwirkung und E_k für Atom-Wand-Wechselwirkung gegeben. Wenn die Geschwindigkeit eines Atoms groß ist, kann es ganz aus dem Kasten entweichen, ein Effekt, auf dem z. B. das Verdunsten von Flüssigkeiten beruht. Dies läßt sich mit dem Programm durch vorsichtiges Erhöhen der Anfangsgeschwindigkeiten oder Erniedrigen von E_b sehr schön demonstrieren.

Die Rechenzeit, die für die beiden Teilpotentiale benötigt wird, hängt unterschiedlich stark von der Anzahl der Atome ab: Die

Berechnung der Atom-Wand-Wechselwirkung wird einfach für jedes Atom einzeln durchgeführt, die benötigte Rechenzeit steigt also linear mit der Anzahl der Atome; die Atom-Atom-Wechselwirkung macht jedoch eine Berücksichtigung aller Paare von Atomen erforderlich, was zu einer quadratischen Abhängigkeit von der Atomzahl führt. Dies ist genau analog zu den Wechselwirkungen im Makromolekül, wo die Rechenzeit für die elektrostatische Energie quadratisch von der Atomzahl abhängt. Dieser Effekt kann mit dem Beispielpogramm durch Erhöhung der Atomzahl leicht demonstriert werden.

Zur Integration der Bewegungsgleichungen wird das bereits beschriebene *Verlet*-Verfahren verwendet. Er ist auch in professionellen Programmen nicht anders realisiert. Allerdings beinhalten diese Programme vielfach erweiterte Aspekte der Simulation, wie zum Beispiel die Berücksichtigung von Reibungs- und Zufallskräften, die sich nicht aus der Energiefunktion ergeben. Man könnte sich auch vorstellen, gar nicht die *Newtonschen* Bewegungsgleichungen zu integrieren, sondern eine sogenannte Monte-Carlo-Simulation durchzuführen, wobei die Positionen der Atome zwar zufällig, aber mit energieabhängigen Wahrscheinlichkeiten verändert werden. Dies eröffnet die Möglichkeit, das gravierende Rechenzeitproblem durch Verlängerung des Zeitschrittes in den Griff zu bekommen.

Schon bei diesem einfachen Problem sieht man, daß auf graphische Ausgabe beim *Molecular Design* nicht verzichtet werden kann. Die numerische Ausgabe der Koordinaten und Geschwindigkeiten führt zu endlosen Zahlenkolonnen, die eigentlich nur verwirren können, während die graphische Darstellung sofort einen intuitiven Eindruck vermittelt. Im vorliegenden Fall beispielsweise denkt man gleich an einen Billardtisch:



Bild 7. Graphik des Programmes *Molecular_Dynamics*

Aus Kompatibilitätsgründen wird nur mit Liniengraphik gearbeitet, die Darstellung der Kugeln als Sechsecke ist natürlich keine befriedigende Lösung. Der graphisch interessierte Leser mag sich hier herausgefordert fühlen. Die beiden Routinen zum Lö-

sen des Bildschirms bzw. zum Zeichnen einer Linie müssen natürlich dem jeweiligen Computer angepaßt werden. Im vorliegenden Fall stand ein Atari ST mit Personal Pascal zur Verfügung; auf diesem System müßte also das Programm ohne Änderungen laufen.

In der vorliegenden Form ist das Programm für den durchschnittlichen Personal Computer viel zu langsam. Man würde gerne viele (10...20) Atome in Realzeit (ca. 10 Bilder/s) berechnen können. Die weitaus längste Zeit bringt der Prozessor in der inneren Schleife der Routine *potential* und in der Funktion f zu. Folgende Möglichkeiten könnte man für eine wesentliche Beschleunigung in Betracht ziehen:

- Einsatz eines Numerikprozessors,
- Umschreiben der Arithmetik in Integeroperationen,
- Codierung der kritischen Teile in Assembler.

Hat man bereits einen Numerikprozessor, kann man sich die zweite Möglichkeit natürlich sparen. Die Assemblerlösung dagegen dürfte auf jeden Fall eine merkliche Verbesserung zur Folge haben. Keine Probleme gibt es, wenn man eine CRAY im Bastelkeller stehen hat.

Ausblick

Vom Gesichtspunkt eines Mathematikers aus haben wir in unserem Beitrag vielleicht ein langweiliges Problem betrachtet, die numerische Lösung der *Newtonschen* Gleichungen von vielen tausend Atomen. Diese Lösungen können aber dem Menschen den Zugang zur Welt und Technologie der Makromoleküle erleichtern. Dabei stehen allerdings einer Eroberung dieser Welt durch den Computer noch beträchtliche Hindernisse im Wege: Rechenzeiten sind noch wesentlich zu lang, und es bedarf über die Entwicklung schnellerer Hardware hinaus noch neuer Ideen für Algorithmen und ihre Verteilung auf möglichst viele Rechenwerke. Ein wichtiges Problem stellt auch die enorme Information dar, welche bei der Simulation durch Höchstleistungsrechner anfällt. Durch weitere Entwicklung der graphischen Fähigkeiten des Computers muß man diese Information beherrschen lernen, da der optische Sinneskanal des Menschen im Vergleich zu seinen anderen Sinneskanälen bei weitem die größte Bandbreite, d. h. Aufnahme- und Verarbeitungskapazität pro Zeiteinheit, besitzt. Trotz Meisterleistungen der Computertechnologie und Algorithmen stünde man allerdings der Informationsfülle aus dem Computer hilflos gegenüber, hätte man nicht Konzepte der

GRUNDLAGEN

Fachdisziplinen zur Verfügung, in unserem Beispiel aus Physik, Chemie und Biologie. Der erfolgreiche Einsatz des Computers erfordert unbedingt Kenntnis sowohl der Rechner-technologie als auch der einzelnen Fachgebiete. Unsere Ausführungen sind damit typisch für das Berufsfeld des Computerfachmanns von heute und morgen, der nicht nur den Computer beherrschen, sondern auch Fähigkeiten in einem anderen Fachgebiet, sei es als Mechaniker, Ingenieur oder Mediziner, mitbringen muß, will er auf dem Markt Erfolg haben. Insofern sei es uns verziehen, daß in unserem Beitrag neben Informationen über Computer auch viel Information über Biochemie enthalten ist; die Informationsmischung scheint uns relevant für den physikalisch, chemisch, biologisch oder medizinisch interessierten Computerfachmann von morgen.

Es ist selbstverständlich auch typisch für die moderne Forschung, in die wir hier einen Einblick zu geben versucht haben, daß sie nicht an den Grenzen zwischen den Fachdisziplinen halt macht, sondern je

nach Fragestellung zwischen den Disziplinen arbeitet, also wie bei der in unserem Beitrag aufgegriffenen Fragestellung zwischen Informatik, Physik, Chemie und Biologie. Zwischen den Gebieten ist noch nicht alles „abgegrast“, die interdisziplinären Fragestellungen sind meist besonders reizvoll und anwendungsnah (Anwendungen richten sich nicht nach Disziplinen) und bieten gerade dem jungen Forscher Chancen. Moderne Forschung ist aber auch grenzüberschreitend in des Wortes engerer Bedeutung. Die in diesem Artikel dargestellten Rechnungen und Graphiken wurden in unserer Arbeitsgruppe „*Theoretische Biophysik*“ an der Technischen Universität München zum größeren Teil über Satellit am Nationalen Zentrum für Supercomputeranwendungen der USA an der University of Illinois erzielt. Dabei waren wir mit demselben *Atari*, auf dem das beigefügte Programm entwickelt wurde, direkt mit einer CRAY X-MP48 verbunden. Die Rechnungen in Illinois wurden durch die USA finanziert, die Verbindung nach den USA durch das Bundesministerium für For-

schung und Technologie im Rahmen des Deutschen Forschungsnetzes. Die Rechnungen in München wurden durch die Max-Planck-Gesellschaft und durch den Sonderforschungsbereich „*Primärprozesse der bakteriellen Photosynthese*“ ermöglicht.

Unsere Ausführungen wollen wir mit einem Blick in die Welt der Syntheseplanung und biochemischen Forschung von morgen abschließen: der Chemiker oder Biochemiker wird bei der Planung der Synthese von Makromolekülen die mikroskopische Welt im Computer nachahmen und graphisch verfolgen. Er wird Fragen an den Computer stellen, Ideen für Synthesen ausprobieren und jeweils schnelle graphische Antworten erhalten. Eine hochleistungsfähige vierdimensionale (Zeit und Raum) Farbgraphik unterstützt ihn dabei, neben seiner abstrakt geschulten chemischen Denkweise die natürliche und hervorragend ausgebildete Fähigkeit des menschlichen Gehirns zur Verarbeitung farbiger und bewegter optischer Informationen in seine Arbeit einzubringen.