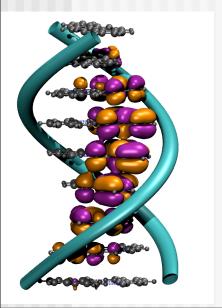
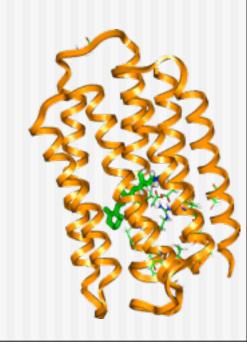
Probleme der Biophysik/Biophysikalischen Chemie:

atomistische Simulationen

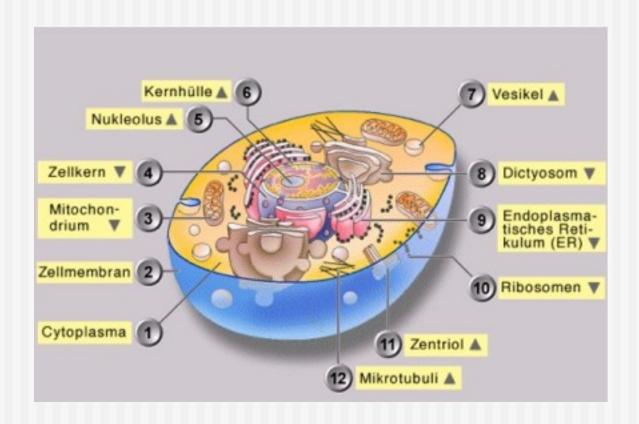
Marcus Elstner und Tomas Kubar Theoretische Chemische Biologie KIT

www.ipc.uni-karlsruhe.de/tcb

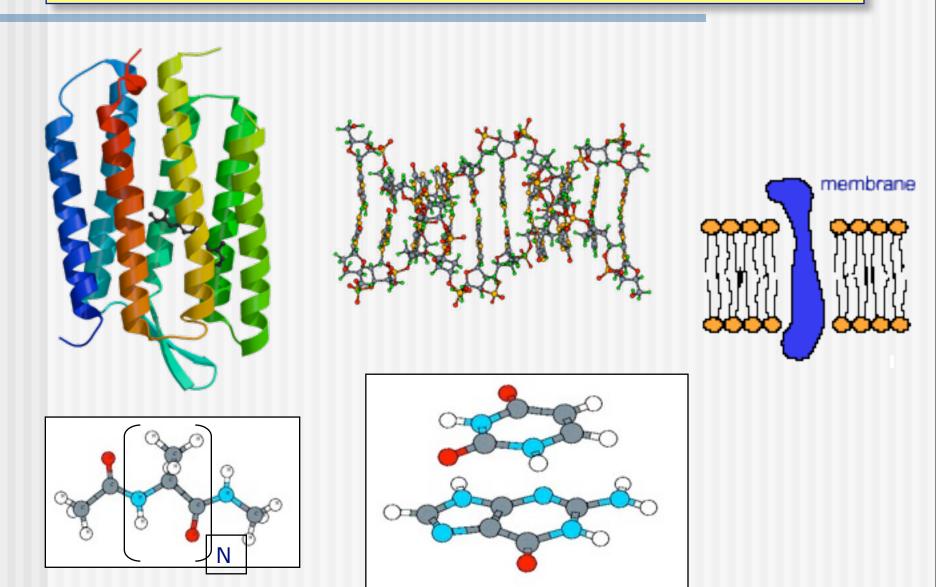




Strukturen

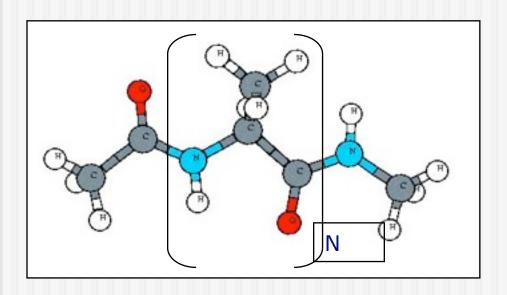


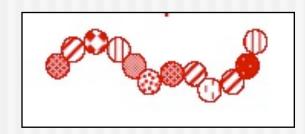
Proteine, DNA, Lipide



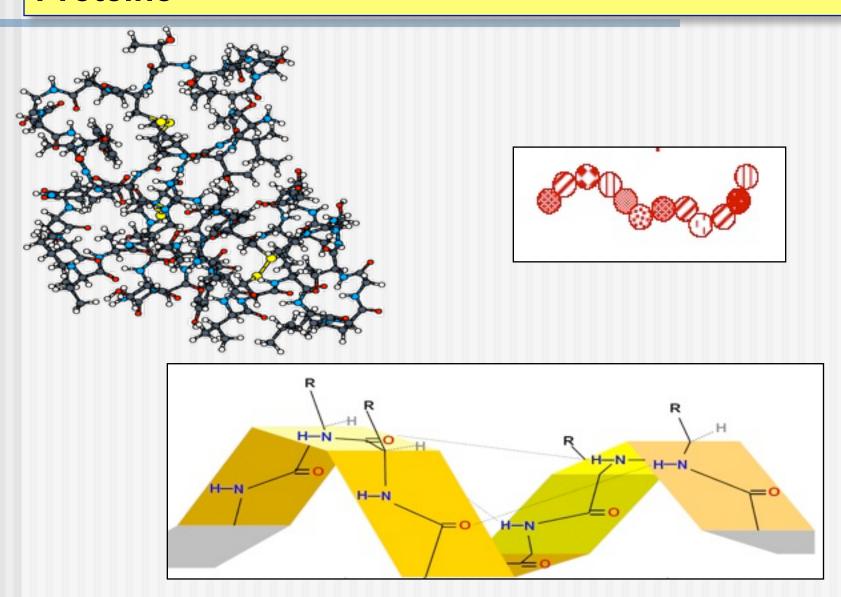
Proteine

- bestehen (hauptsächlich) aus H, C, N, O und S
- Untereinheiten: Aminosäuren

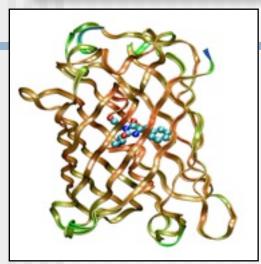




Proteine



Protonentransfer und Photobiophysik



Protonentransfer

Bakterielles Reaktionszentrum

Bakteriorhodopsin

Aquaporin

ADH, CA

Elektronische Anregungen

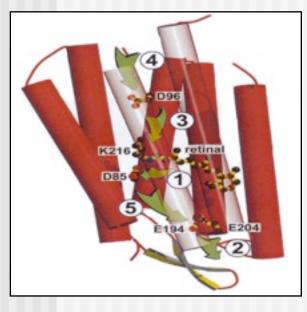
Retinalproteine

GFP

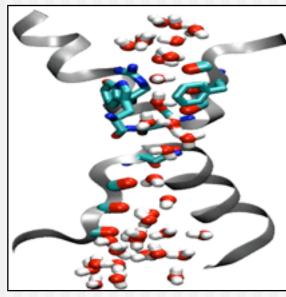
DNA-Basen

Stilben

bR



Aquaporin



Photosynthetic Reaction Center



Biophysikalische Prozesse

Bioenergetik: Energiegewinnung und

Umwandlung

Katalyse: Synthese und Abbau von chemischen

Substanzen

Transport: Austausch mit der Umgebung (Ionen,

Wasser)

Fragenstellungen der theoretischen Biophysik

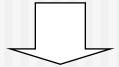
I. Dynamik komplexer Strukturen

- Proteinfaltung
- molekulare Motoren
- Protein-DNA Komplexe
- II. Transport: Wasser, Ionen, Protonen, ...
- III. Elektronentransfer
- IV. Enzyme: was macht sie so effizient?
- chemische Reaktionen: Katalyse
- Photochemie: Konversion von Licht in chemische Energie

Verständnis biophysikalischer Prozesse

Verständnis auf verschiedenen Längen- und Zeitskalen

- atomistisch: Bewegungsgleichungen für gekoppeltes
 N-Teilchenproblem (klassisch/quantenmechanisch)
- Kontinuum: elektrostatische-mechanische Eigenschaften
- Raten&Transportgleichungen: phänomenologisch



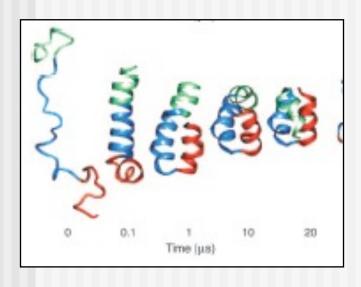
- unterschiedliche theoretische Modelle
- Kombination (in Grenzen) sinnvoll: "Multi-scale modeling"

Funktionalisierung

- Interfaces: Biokompatibilität, Lacke im Schiffsbau, Biosensoren, Hybridmaterialien etc.
- Nano-Elektronik: Moleküle als elektronische Bauelemente, DNA als nano-Leiter
- Optische Speicher, z.B. durch laseradressierbare molekulare Konformationen (bR)

= . . .

Proteinfaltung: wie 'findet' ein Protein seine Struktur?

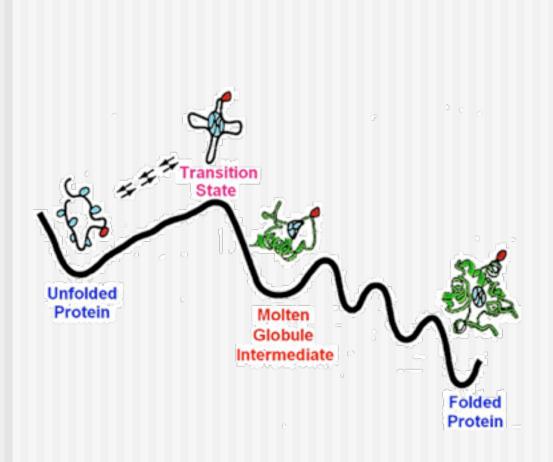


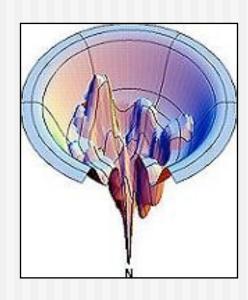
Karplus& Kuriyan, PNAS 2005

Probleme:

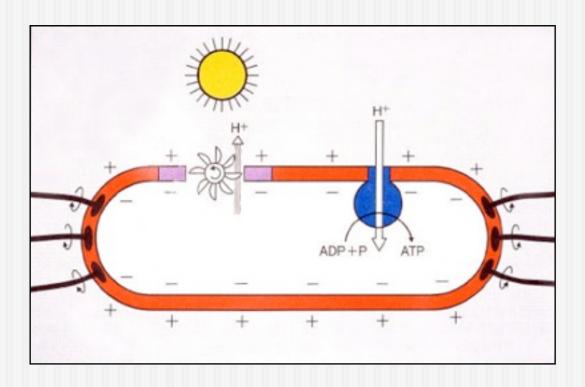
- grosse Zeitskalen > ms
- Genauigkeit von MM
- grosse Systeme(Lösungsmitteleffekt)

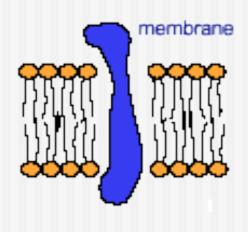
Proteinfaltung: wie 'findet' ein Protein seine Struktur?



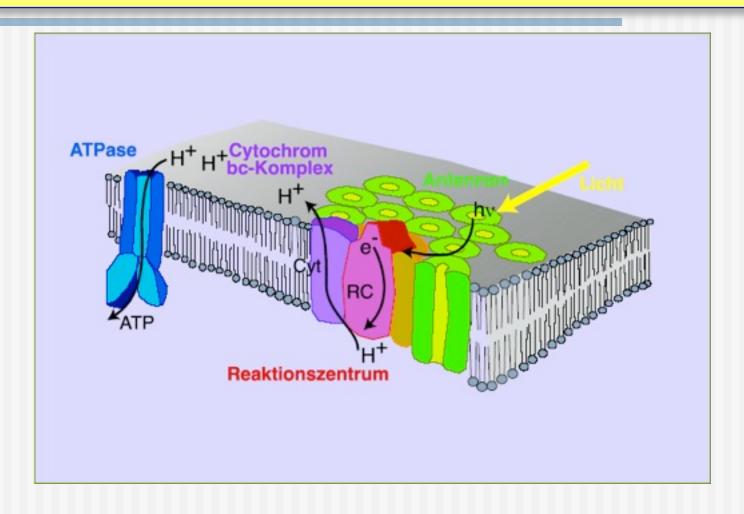


Bioenergetik

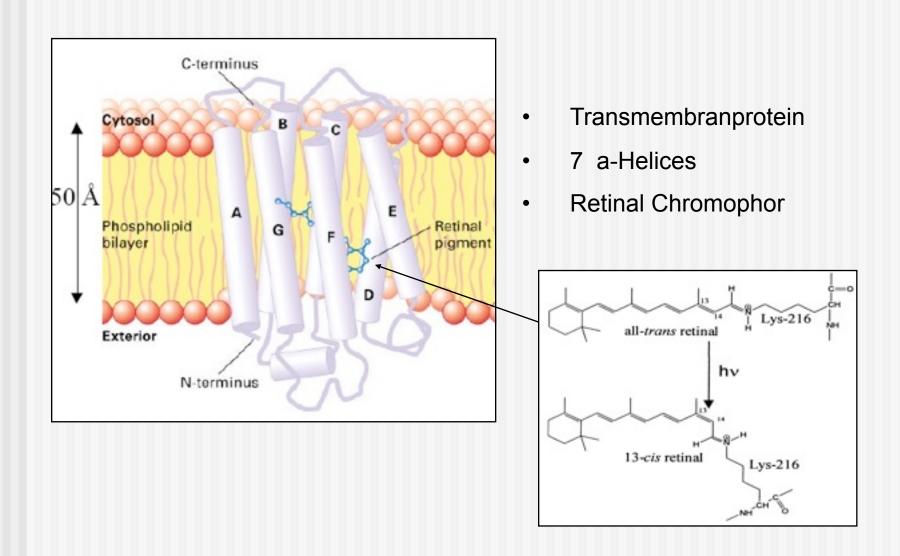




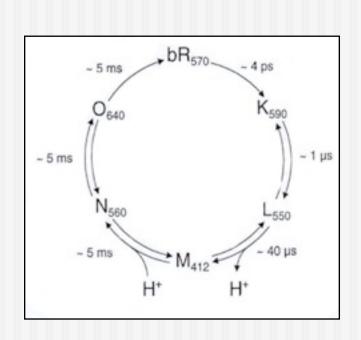
Bakterielles Reaktionszentrum

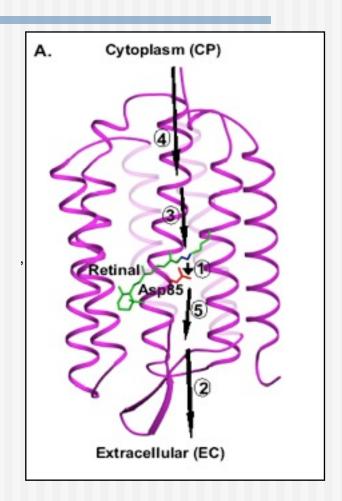


Bacteriorhodopsin



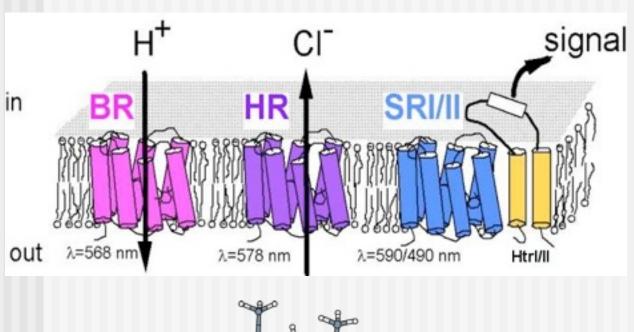
Bacteriorhodopsin

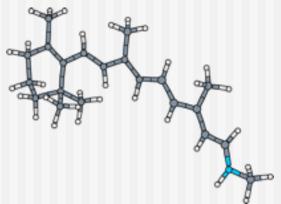


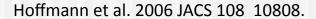


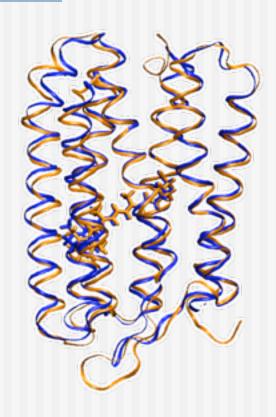
http://www.retinalproteins.org/

bR vs SRII

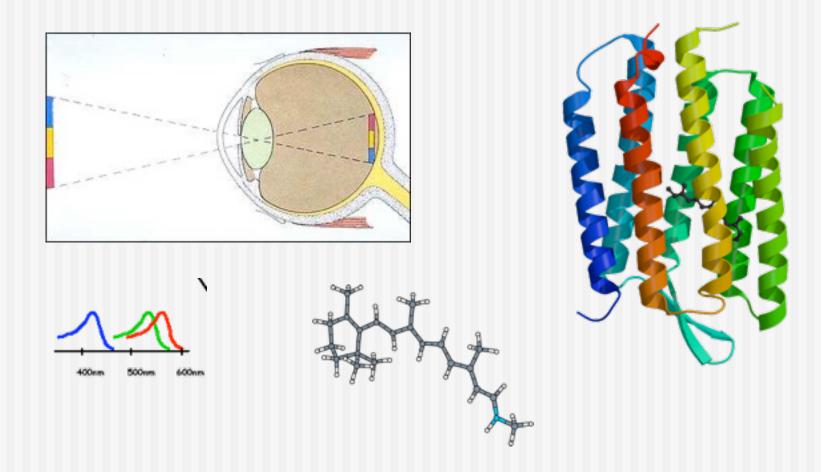






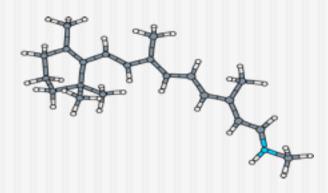


Vision: Rhodopsins

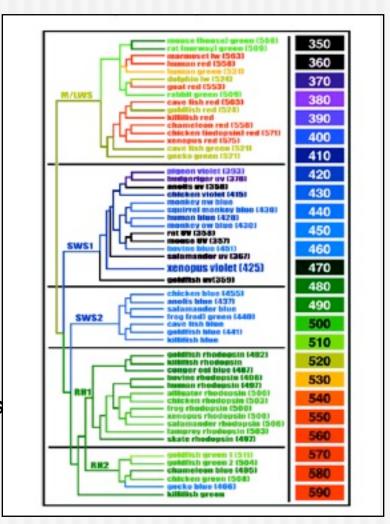


'Spectral tuning'

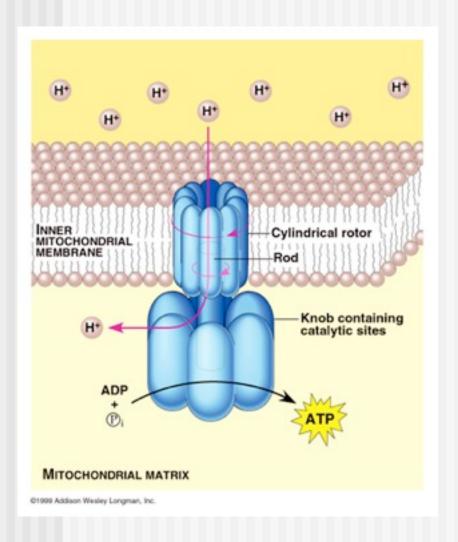
Absorption over 300 nm "Tuning" by protein environment (opsin-shift)

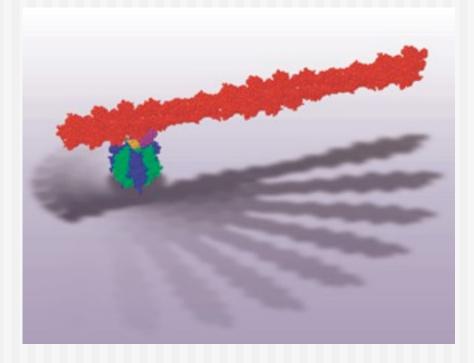


- sterical interactions: 'twist'
- interaction with polar and charged amino acids
- H-bonds with counterions



ATPase: Umsetzung von chemischer in mechanische Energie

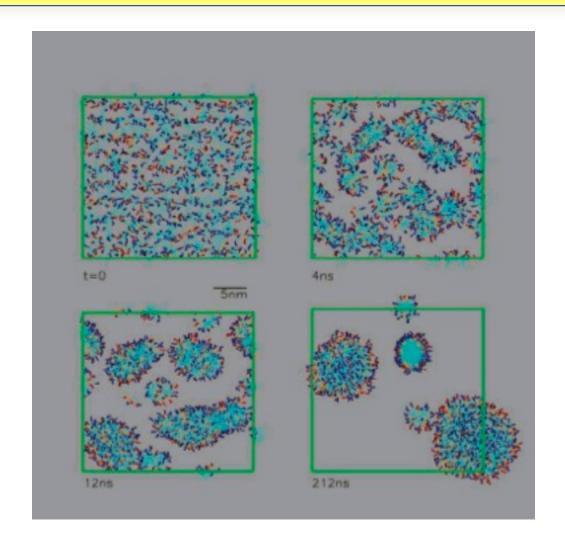




Sichtbarmachung der F1-Rotation durch Actin Filament

Rotation: ms-ms

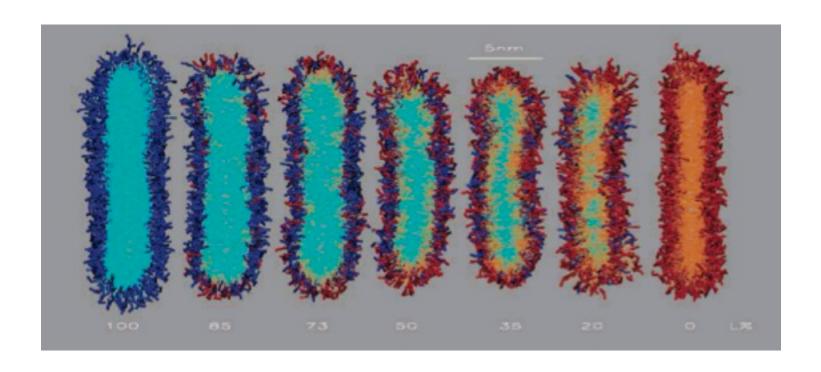
Struktur der Zellwände: Selbstorganisation



Formation of bilayers

Langmuir **2006**, **22**, **998-1005**

Simulation of lipids

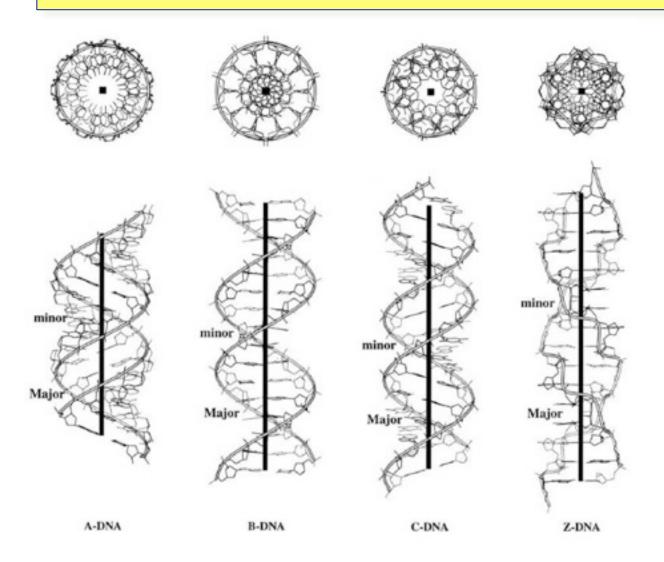


disk depending on lipid length and composition

coarse grained (CG) MD simulations, described later on

Langmuir **2006**, **22**, **998-1005**

Simulation der DNA

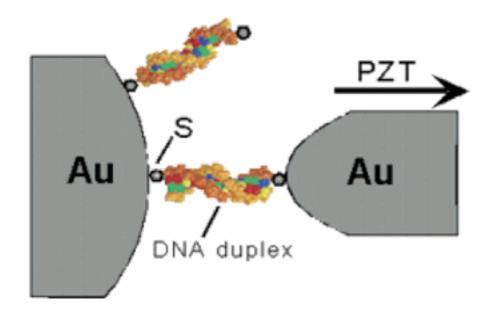


structures:

how do sequence and 3d structure relate?

DNA

Nano-electronics

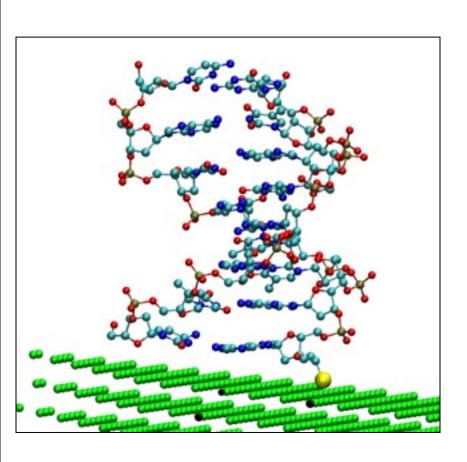


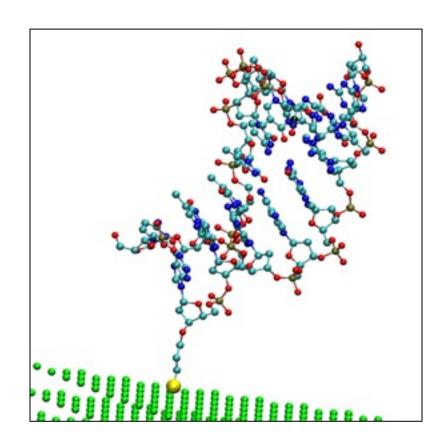
B. Xu et al., Nano Lett. 4, 1105 (2004)

Simulation of DNA

interaction of DNA with surfaces

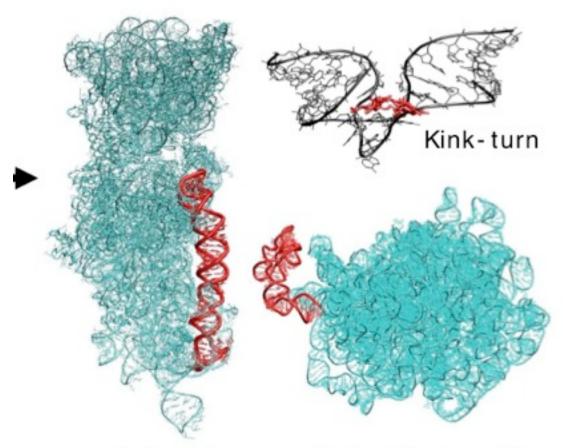
structure of DNA and contact





Even more interesting: RNA

Ribosomal RNA segments

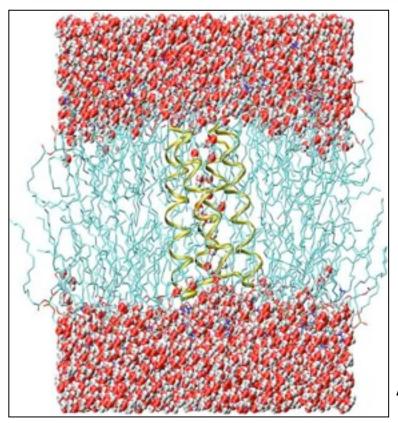


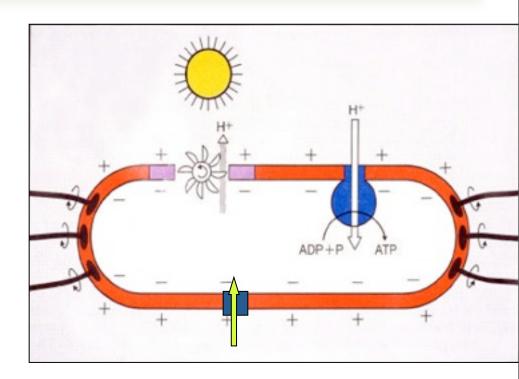
Helix 44

Factor Binding Site

see Sponer: www.ibp.cz/labs/LSDNA/

Poren



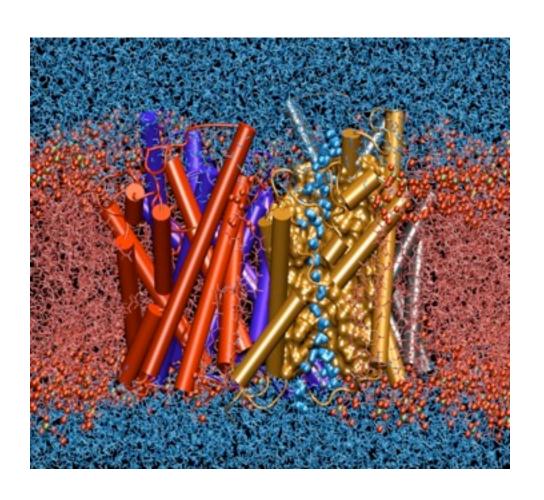


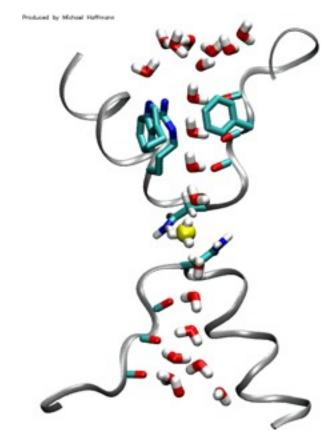
Aquaporin

Gramicidin

Poren

Druckausgleich: Wassertransport, aber Protonenblockade





Proteine

I. Struktur und Dynamik komplexer Structuren?

- auf den ersten Blick chaotisch, doch sehr strukturiert
- können wir Struktur und Dynamik verstehen?
- Faltung: wodurch wird die 3-D Struktur bestimmt?

II. Funktion

- Mechanismus?
- Struktur-Funktions Zusammenhang?

III. Mutation/Modifikation

- können wir den Effekt von Mutationen beschreiben?
- Optimierung/Design: Verbesserte Funktionalität

IV. Funktionalisierung

biologische Strukturen und Eingeschaften in Technik verwenden:
 Selbstorganisation und Selbstvermehrung

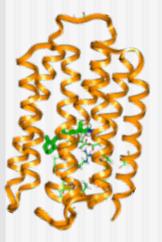
Funktionalisierung

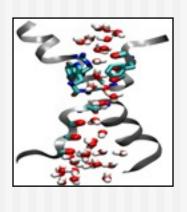
- Interfaces: Biocompatibility, new materials, biosensors, hybride materials etc.
- Nano-Elektronics: Moleculas as electronic devives, DNA as nano-rod
- Optical storage, e.g. conformational changes induced by a laser (bR)

• ...

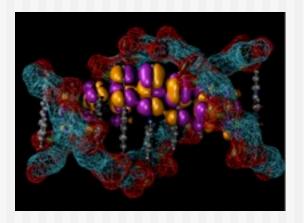
Themen

Bioenergetik: Protonentransport





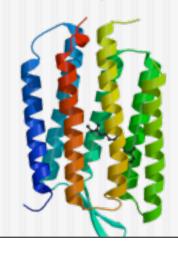
Elektronentransport in DNA

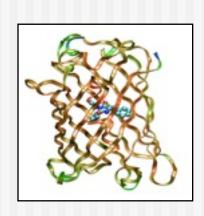


 Kombination von Quantenchemie & Biophysik

2) Methoden werden prädiktiv, sie können experimentelle Daten komplementieren

Optische Eigenschaften: 'color tuning'





Dies sind Probleme,

- in denen das molekulare Detail entscheidend ist,
- in denen die Quantenmechanik (QM) explizit von Nöten ist,
- die eigentlich zu gross für eine QM Behandlung sind

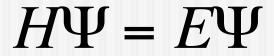
Quantenmechanik:

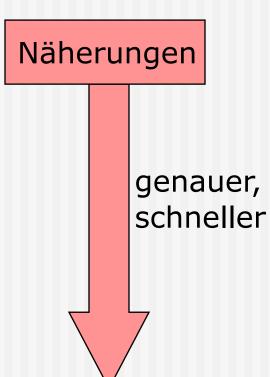
- explizit: chemische Reaktionen, angeregte Zustände ...
- **implizit**: chemische Bindung, VdW Bindung sind durch effektive Potentiale modellierbar

Anwendung der Quantenmechanik in der Chemie

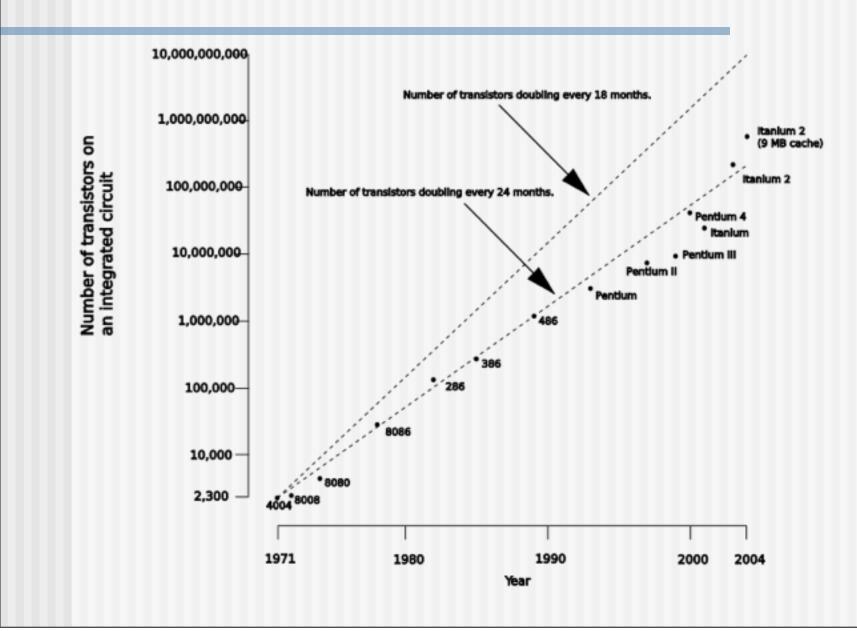
- Schrödingergleichung 1926:
- Heitler und London 1927: H2 Molekül
- Hund und Mulliken 1929: MO Theorie
- 1930 Hartree-Fock (HF)
- ab 1950: Computereinsatz: Semiempirik
- 1965 Dichtefunktionaltheorie (DFT)
- 1998 Nobelpreis für Chemie: Pople & Kohn
- viele Weiterentwicklungen:

post-HF DFT Semi-Empirik





Moore's Law



Quantenchemische Methoden

 \bigoplus

'Elektronenwolke'



2) bestimme Wellenfunktion

3) bestimme die Energie E

- 4) E →
 - Kräfte auf Atome: Struktur, Dynamik

Ψ

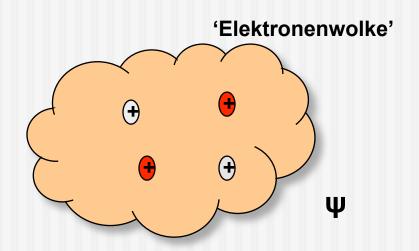
- Spektroskopie
- etc.

Rechenzeit!?

Quantenchemische Methoden (QM)

Problem ist die Rechenzeit:

- Die Schrödingergleichung (SG) ist nur für sehr kleine Moleküle exakt lösbar
- HF, DFT und post-HF sind Modelle: formalisieren Modellvorstellungen
- Weitere Näherungen: semi-empirische Modelle: MNDO, AM1, PM3 etc.



→ trotzdem:

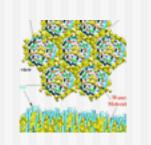
max. 100-1000 Atome

empirical force fields

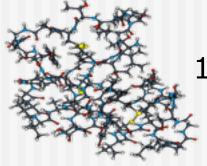
Semi-empirical methods (DFTB, MNDO etc.)

Hartree-Fock (HF), Density Functional Theory (DFT)

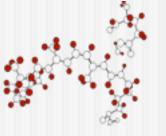
post-Hartree-Fock: MP2,CC, CI, MRCI ...



100k atoms



1000 atoms



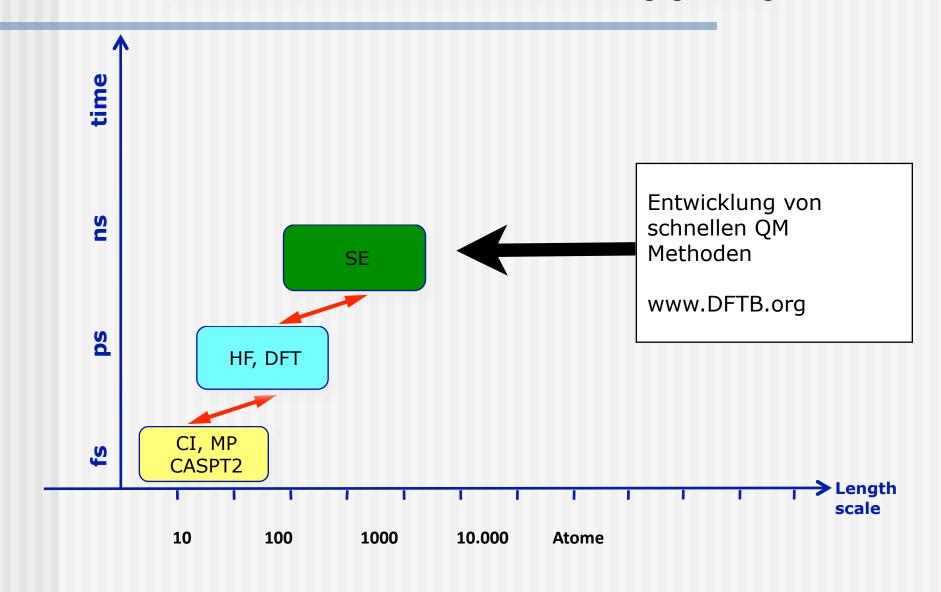
100 atoms

10-50 atoms

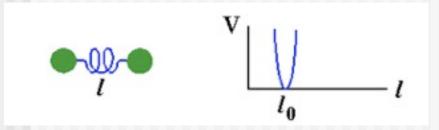
Friday, October 22, 2010

Methodenspektrum

Größe und Simulationszeit limitieren sich gegenseitig



Gehts auch einfacher? Beispiel: H2 - Molekül



$$E(I) = 0.5 k*(I-I_0)^2$$

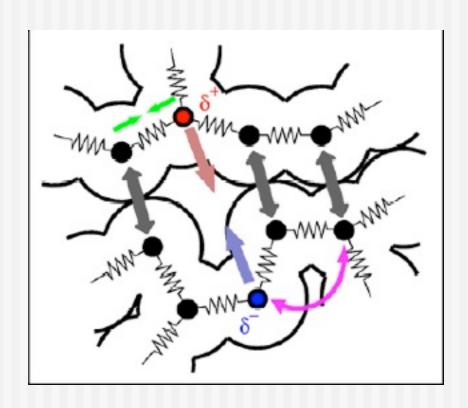
Parameter: k und lo

Geometrie: bei lo

Schwingungsfrequenz: $v = \sqrt{k/m}$

Problem: Bindungen können nicht gebrochen werden

Empirische Kraftfelder: Molecular Mechanics (MM)



© Grubmüller

$$V_{B} = \sum_{\text{Bindungen}} \frac{1}{2} K_{b} (b - b_{0})^{2}$$

$$V_{a} = \sum_{\text{Winkel}} \frac{1}{2} K_{\theta} (\theta - \theta_{0})^{2}$$

$$V_{imp} = \sum_{\text{Extraplanarwinkel}} \frac{1}{2} K_{\zeta} (\zeta - \zeta_{0})^{2}$$

$$V_{D} = \sum_{\text{Dihedralwinkel}} K_{\varphi} [1 + \cos(n\varphi - \delta)]$$

$$V_{q} = \sum_{\text{Paare}(i,j)} q_{i}q_{j}/(4\pi\epsilon_{0}\epsilon_{r}r_{ij})$$

$$V_{vdW} = \sum_{\text{Paare}(i,j)} C_{12}(i,j)/r_{ij}^{12} - C_{6}(i,j)/r_{ij}^{6}$$

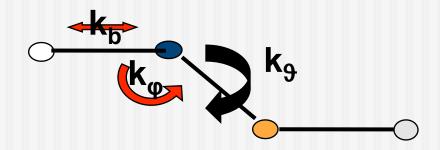
Quelle: Grubmüller MPI Göttingen

Empirische Kraftfelder: Molecular Mechanics (MM)

$$V = \sum_{bonds} k_b (b - b_0)^2 + \sum_{angles} k_\theta (\theta - \theta_0)^2 + \sum_{dihedrals} \sum_{n=1}^{N} k_\phi^{(n)} \left[1 + \cos((n\phi - \delta)) \right] + \sum_{impropers} k_\omega (\omega - \omega_0)^2$$

$$+ \sum_{i,j} 4\epsilon_{i,j} \left[\left(\frac{\sigma_{i,j}}{r_{i,j}} \right)^{\frac{1}{2}} - \left(\frac{\sigma_{i,j}}{r_{i,j}} \right)^{\frac{1}{2}} \right] + \sum_{i,j} \left(\frac{q_i q_j}{D r_{ij}} \right)^{\frac{1}{2}}$$

Für Protein- und DNA Strukturen i.A. ganz gut



- -keine chem. Reaktionen
- -feste Ladungsverteilung

Molecular Dynamics: MD

Zeitschritt: dt~1fs

→ Trajektorien

→ 1ps Dynamik: 1000 mal Energie+Kräfte

atomare Kräfte:
$$F_i = -\frac{\partial V}{\partial R_i}$$

Newtonsche Bewegungsgleichungen: F_i=m*a_i

Numerische Integration: $x(t+dt) = x(t) + vdt + 0.5 a(dt)^2$

"Molecular Dynamics": MD

Numerische Integration von Newtons Bewegungsgleichungen

Start:

berechne E und atomare Kräfte

$$F_i = -\frac{\partial E}{\partial R_i}$$









Schritt 1:

Atome werden in Richtung der Kräfte beschleunigt

a= F/m

Geschwindigkeit v







Schritt 2:

berechne E und Kräfte für neue Geometrie

$$F_i = -\frac{\partial E}{\partial R_i}$$



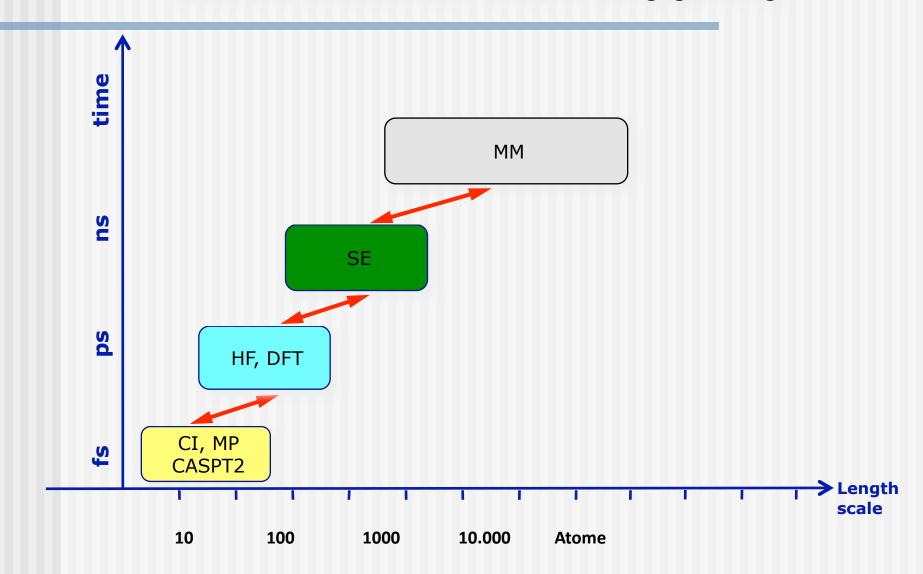






Methodenspektrum

Größe und Simulationszeit limitieren sich gegenseitig



Biophysik: Simulation biologischer Strukturen

- 1959: Vineyard, Cu-Kristall

- 1964: Rahman, flüss. Ar

- 1960er Kraftfeldmodelle

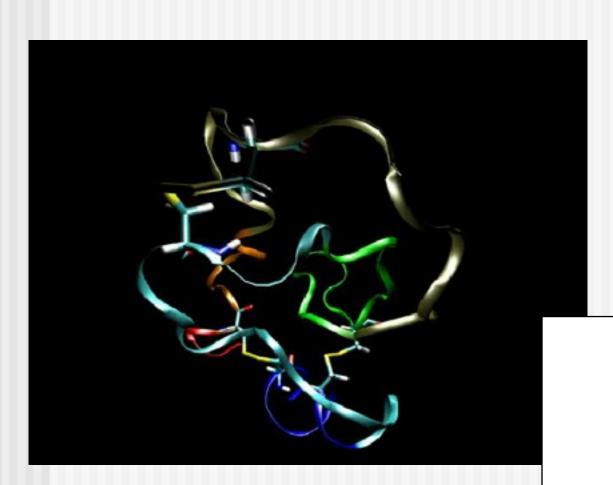
(Lifson, Sheraga, Allison)

Biophysik: Simulation biologischer Strukturen

- 1959: Vineyard, Cu-Kristall
- 1964: Rahman, flüss. Ar
- 1960er Kraftfeldmodelle
- 1977 MD von Protein

Erste Simulation der Dynamik eines Proteins: 9.2 ps

Mc Cammon, Gelin& Karplus, Nature 267, 1977



BPTI

(bovine pancreatic trypsin inhibitor)

58 Aminosäuren

Startpunkt: Kristallstruktur

BPTI: 210ps und Wasser

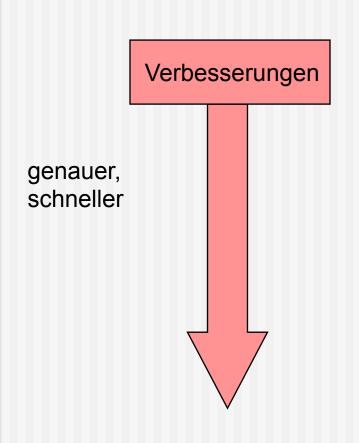
Levitt & Sharon, PNAS 85, 1988.

ABSTRACT Simulation of the molecular dynamics of a small protein, bovine pancreatic trypsin inhibitor, was found to be more realistic when water molecules were included than when in vacuo: the time-averaged structure was much more like that observed in high-resolution x-ray studies, the ampli-

heute: ~100.000 Atome

~10-100 ns

Biophysik: Simulation biologischer Strukturen



- 1959: Vineyard, Cu-Kristall
- 1964: Rahman, flüss. Ar
- 1960er Kraftfeldmodelle
- 1977 MD von Protein
- effiziente Implemetierungen
- freie Energieberechnungen
- 2005: IBM ,blue gene'

Biophysik: Simulation biologischer Strukturen

Table 9: History and extrapolated future of computer simulations of molecular dynamics. The future is deduced from extrapolation based on an observed increase of computing speed of a factor 10 every 5 years over the past decades (see Figure 31).

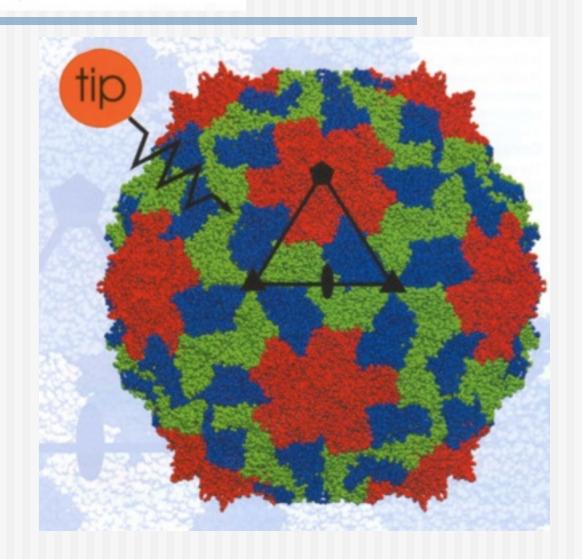
Year	Molecular system (type, size)	Length of the simula- tion [s]
1957	first molecular dynamics simulation (hard discs)	
1964	atomic liquid (argon)	10^{-11}
1971	molecular liquid (water)	5×10^{-12}
1977	protein in a vacuum	2×10^{-11}
1983	protein in water	2×10^{-11}
1989	protein-DNA complex in water	10^{-10}
1997	polypeptide folding in solvent	10^{-7}
2001	micelle formation	10^{-7}
200x	folding of a small protein	10^{-3}

van Gunsteren AC 2006

Simulation of a Complete Viral Shell

4.5 Mio Atome

 $0.5 \mu s$



Grubmüller et al. 2009

Extrapolation: Verzehnfachung der Rechenleistung alle 5 Jahre

```
And the future ...

2001 biomolecules in water (ca. 10<sup>4</sup> atoms)

2029 biomolecules in water (folding sooner?)

2034 E. coli bacteria (ca. 10<sup>11</sup> atoms)

2056 mammalian cell (ca. 10<sup>15</sup> atoms)

2080 biomolecules in water (as fast as nature)

2172 human body (ca. 10<sup>27</sup> atoms)

10<sup>-8</sup>

10<sup>-3</sup>

10<sup>-9</sup>

10<sup>-9</sup>
```

- können wir die 'Dynamik' eines Menschen genau so bestimmen, wie die eines Proteins?
- unser Programm ist **reduktionistisch** und **deterministisch!**

wir reden von Molekularen Maschinen (wenn auch metaphorisch!)

gibt es eine Grenze? Protein -. ...- ganzer Mensch

Reduktionistisches Programm

Reduktionismus: komplexe Sachverhalte können durch die elementaren Gleichungen der Physik beschrieben werden

Determinismus: alle Geschehnisse sind durch die Anfangsbedingungen und elementare Gesetze determiniert

Freiheit, Selbstbewusstsein etc. werden zu Begriffen, die 'leer' sind, sie sind nur Epiphänomene.

Wirkmächtigkeit haben nur die zu Grunde liegenden Kausalgesetze.

Aktuell: Hirnforschung (Libet Experimente)

- aber schon bei Leibnitz etc.

Einwände

Technische Probleme: mit der Systemgrösse explodieren auch die Genauigkeitsanforderungen

Quantenmechanik: Probleme mit der Unschärferelation, probabalistischer Charakter

Chaostheorie: kleine Unsicherheiten in den Anfangsbedingungen

Emergenz: treten auf grösseren Skalen neue Gesetzmässigkeiten auf, die im molekularen Bereich noch nicht vorhanden waren (z.B: R. B. Laughlin)

=> macht eine atomistische Beschreibung überhaupt Sinn?

Thermodynamik: Mikroskopische Beschreibung kennt keinen 2. HS (Zeitumkehrinvarianz)

Einwände

ein Kollege: 'und was, wenn er in der Sekunde nur gähnt?'

aber was erwarten wir?

- dass er über den Sinn des Lebens nachdenkt?
- würden wir das 'Denken' nennen?
- würden wir ihm 'Selbstbewusstsein' zuschreiben?

Problem des performativen Selbstwiderspruchs:

Die These des Determinismus ist selbstwidersprüchlich.

'Wenn alles determiniert ist, ist es auch determiniert, dass jemand dies so artikuliert' Es ist also nicht mehr wahr oder falsch, sondern einfach vorherbestimmt. Der Determinismus zerstört damit die Rede von 'wahr' und 'falsch'

Einwände

Naturbeschreibung kann unsere Intuitionen - was den Menschen ausmacht - nicht einholen.

Anwendung unserer Erkenntnisse auf uns selbst kann in die Aporie führen (Problem der Selbstbezüglichkeit?)

Theoret. Chemie

Theoret. Biophysik

- Schrödingergleichung 1926:
- Heitler und London 1927: H2 Molekül
- Hund und Mulliken 1929: MO Theorie
- 1930 Hartree-Fock (HF)
- 1950: Computereinsatz: Semiempirik
- 1965 Dichtefunktionaltheorie (DFT)
- 1998 Nobelpreis für Chemie:
 Pople und Kohn

- 1959: Vineyard, Cu-Kristall
- 1964: Rahman, flüss. Ar
- 1960er Kraftfeldmodelle
- 1977 MD von Protein
- effiziente Implemetierungen
- freie Energieberechnungen

- ..



Chemische Biologie

Theoretische Chemische Biologie

Theoretische Chemie:

Theoretische Biophysik:

genaue Quantenchemische Verfahren

MD Simulationen

DFT, post-HF ...

Grosse Systeme,

Entropie

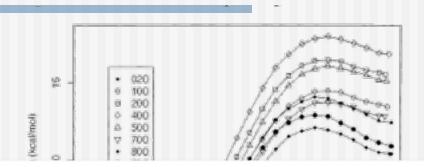




- 1. Dynamik und Entropie
- 2. Grosse Systeme: Multiskalen Methoden (effiziente Verbindung von QM und MM)

'Problem' der potentiellen Energie

Unterschiedliche Energieprofile für unterschiedliche Konformationen



A) Man muss immer über die Konformationen der Umgebung 'mitteln':
 Gesamtenergie→ innere Energie

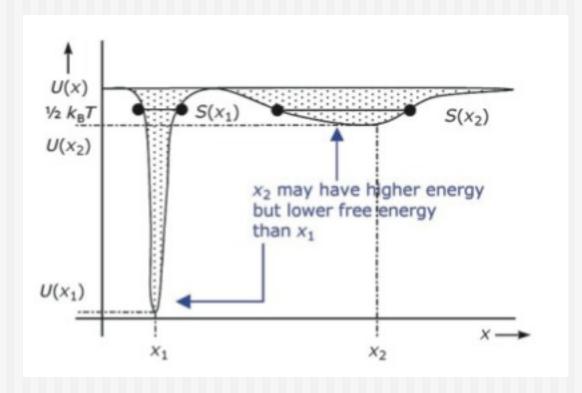
E→ U

B) Entropie ist oft genauso wichtig wie genaue Bestimmung von E:

U→ F

Lilally et al JECD 10/ (2003) 4443:

'Problem' der potentiellen Energie



van Gunsteren AC 2006

=> Methoden der Theoretischen Biophysik:

lange MD Simulationen nötig!

Theoretische Chemische Biologie

Theoretische Chemie:

Theoretische Biophysik:

genaue Quantenchemische Verfahren

MD Simulationen

DFT, post-HF ...

Grosse Systeme,

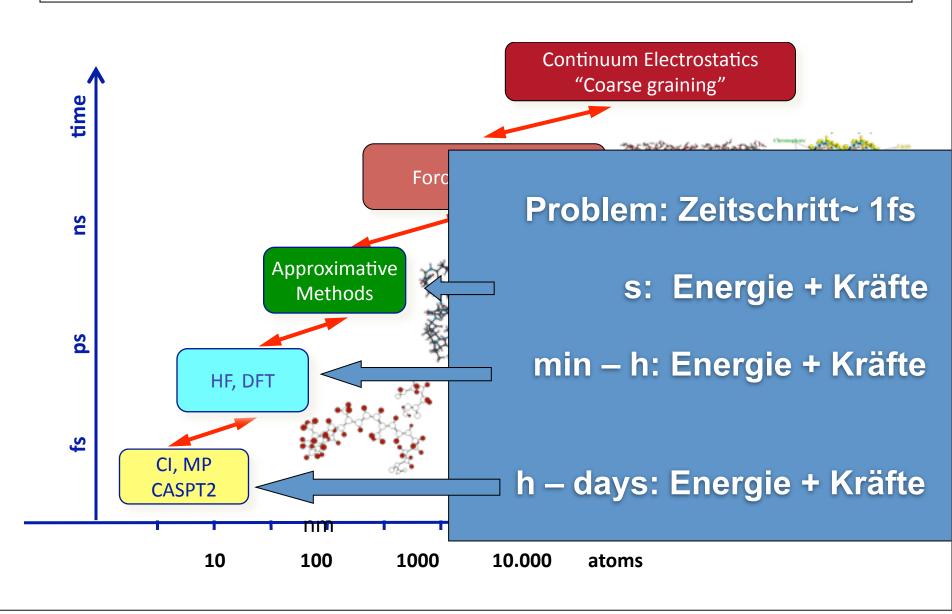
Entropie





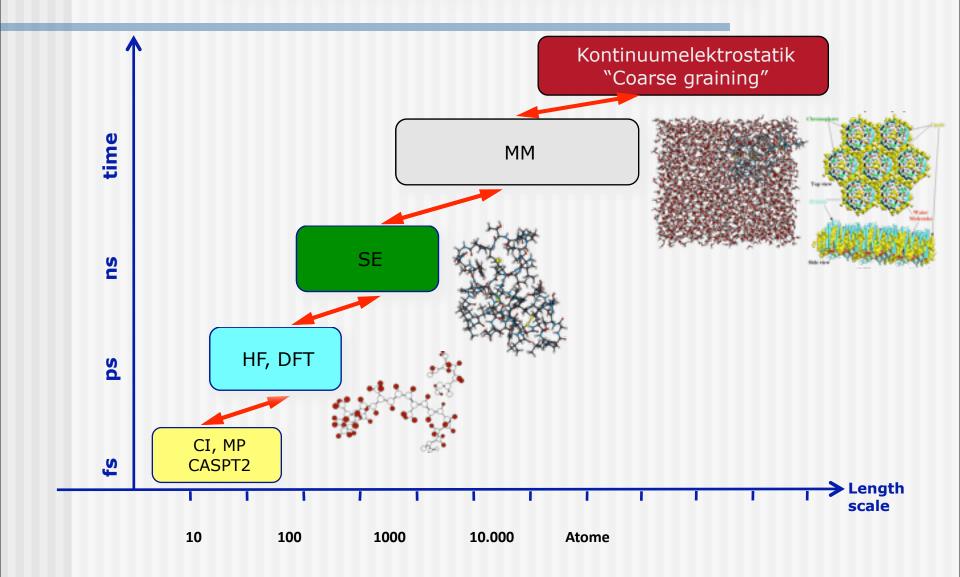
- 1. Dynamik und Entropie
- 2. Grosse Systeme: Multiskalen Methoden (effiziente Verbindung von QM und MM)

Multiskalenmodelle

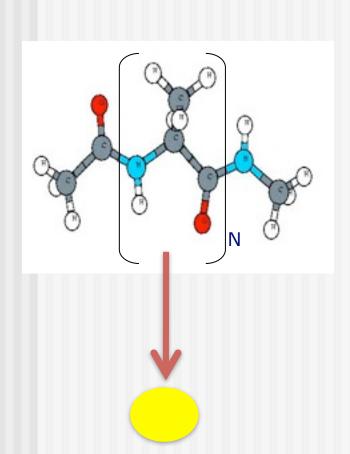


Methodenspektrum

Größe und Simulationszeit limitieren sich gegenseitig



'Coarse Graining' in klassischen Simulationen



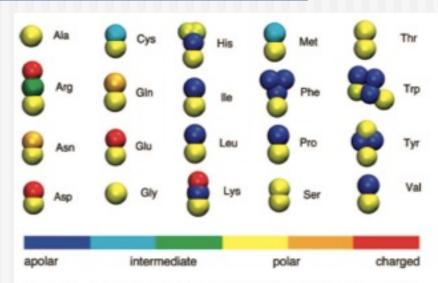


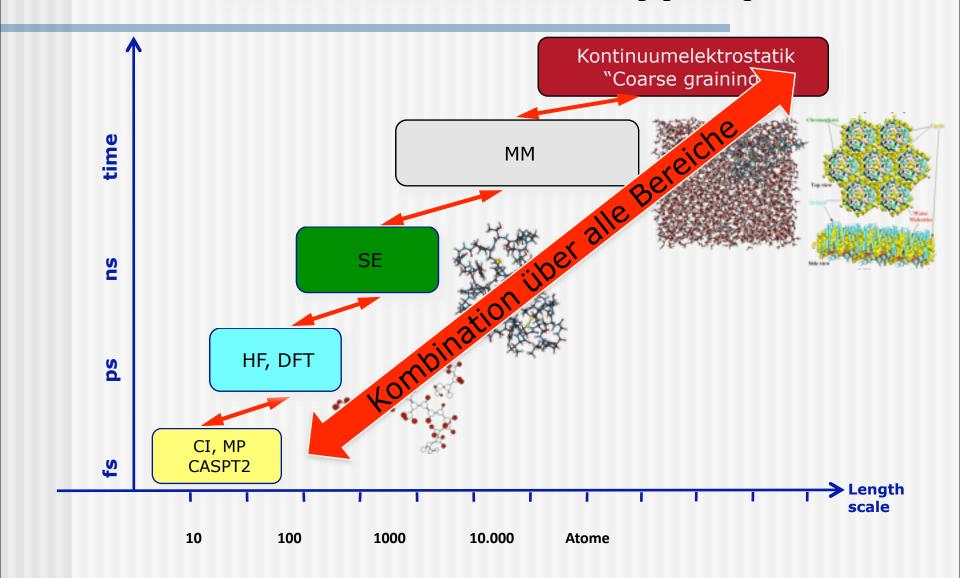
Figure 1. Coarse-grained representation of all amino acids. Different colors represent different particle types.

Martini force field, Monticelli et al. JCTC 4 (2008) 819

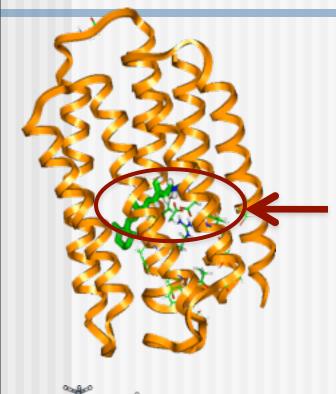
- fasse mehrere Atome zu einem effektiven 'Superatom' zusammen
- effektive Wechselwirkungen zwischen den 'Superatomen'

Methodenspektrum

Größe und Simulationszeit limitieren sich gegenseitig



QM/MM Methoden



Aktives Zentrum ist meist räumlich lokalisiert

Daher: QM Beschreibung nur für Teilbereich

Rest: schnellere Methode, etwa MM

Kombinierte QM/MM Methoden

- 1976 Warshel und Levitt
- 1986 Singh und Kollman
- 1990 Field, Bash und Karplus

QM

- Semiempirisch
- Quantenchemieprogramme: DFT, HF, MP2, LMP2
- DFT 'plane wave codes': CPMD

MM

• CHARMM, AMBER, GROMOS, SIGMA, TINKER, ...

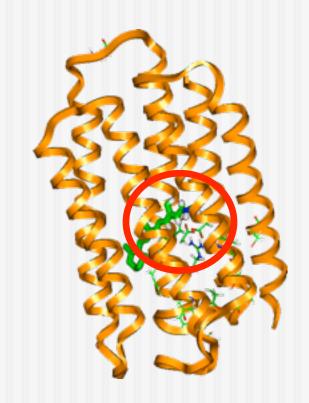
Understanding the action of enzymes

(Warshel, Annu. Rev. Biophys. Biomol. Struct. 2003. 32:425-43)

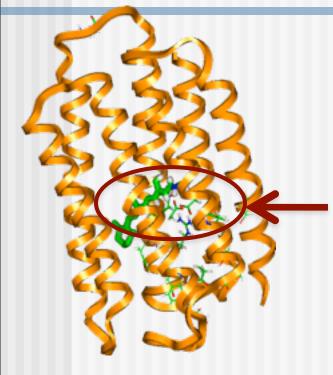
• in den meisten Proteinen: Katalytischer Effekt durch elektrostatische Wechselwirkung mit der Proteinumgebung!

Weniger wichtig:

- ,desolvation'
- sterische Effekte
- ,near attac conformation (NAC)
- ,coherent dynamics'



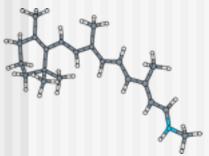
Multiskalen Methoden



Aktives Zentrum ist meist räumlich lokalisiert

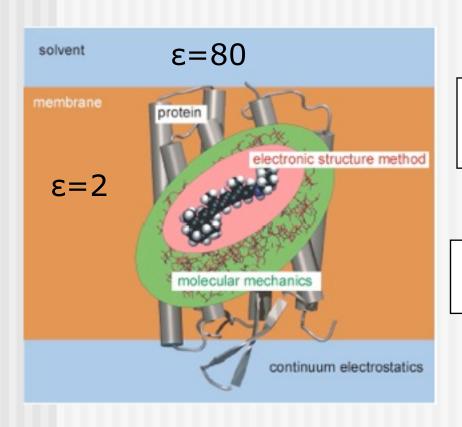
Daher: QM Beschreibung nur für Teilbereich

Rest: schnellere Methode, etwa MM



Multi-Skalen Methoden: Kombination von verschiedenen QM, MM, ,coarse graining' und Kontinuums Methoden

Multiskalen Methoden: QM/MM/Kontinuum



$$\Delta W_{elec}(\mathbf{X}) = \frac{1}{2} \sum_{\alpha} q_{\alpha} \phi_{rf}(\mathbf{r}_{\alpha}).$$

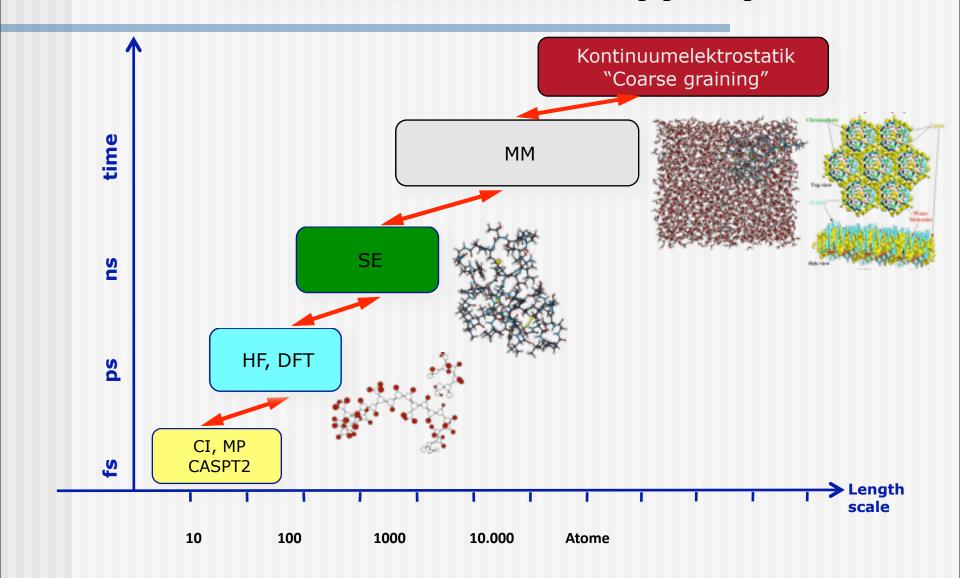
$$\nabla \cdot \left[\epsilon(\mathbf{r}) \nabla \phi(\mathbf{r}) \right] - \overline{\kappa}^2(\mathbf{r}) \phi(\mathbf{r}) = -4\pi \rho(\mathbf{r}).$$

Linearisierte PB Gleichung

Methoden für lange Zeitskalen

Methodenspektrum

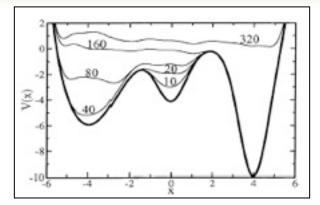
Größe und Simulationszeit limitieren sich gegenseitig



Reaktionen und langsame Prozesse

- direkte MD
- accelerated MD
 - hyperdynamics (Voter)
 - chemical flooding (Grubmüller)
 - metadynamics (Parinello)
 - replica exchange
- Reaktionspfade (reaction path methods)
 - NEB (nudged elastic band, Jonsson)
 - CPR (conjugate peak refinement, Fischer, Karplus)
 - dimer method (Jonsson)
- Freie Energie Methoden (free energy sampling techniques)
 - umbrella sampling
 - free energy perturbation
 - transition path sampling

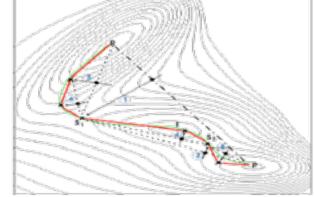
Reaktionen und langsame Prozesse

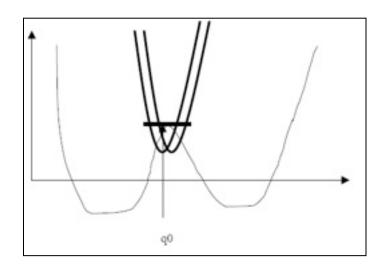


accelerated MD

reaction path methods

free energy sampling techniques

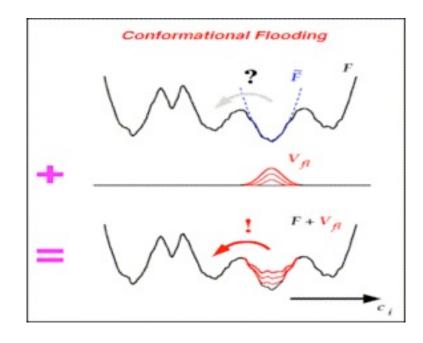




Accelerated MD

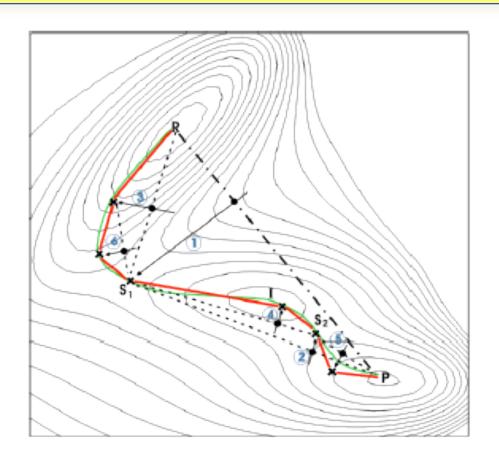
accelerated MD

- 'Chemical flooding'
- Hyperdynamics
- Metadynamics



H. Grubmüller

Reaction path methods



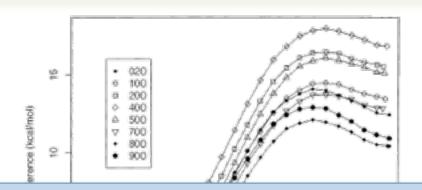
CPR: Fischer & Karplus

~10.000 Energie und Kraftberechnungen

→SE, MM Methoden

'Problem' of potential energy

Effekt der Proteindynamik



A) One always has to average over the different conformations of the environment:

Total energy → inner energy

B) Entropy is often as important as accurate total energy:



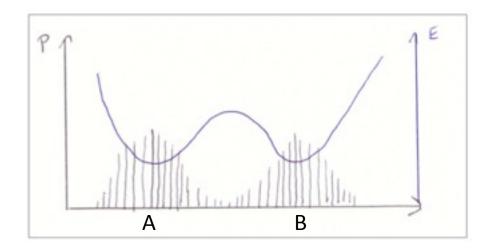
Figur

Freie Energie: 'potential of mean force'

MD für die Zustände A und B:

Freie Energie:

$$F_B - F_A = -kT ln \frac{P(q_B)}{P(q_A)}$$



bei hoher Barriere: addiere zusätzliches Potential

dieses wird nach der Simulation wieder 'abgezogen'

'Umbrella sampling'

