

Molecular structures

Minimization, vibrational analysis, transition states

Marcus Elstner, Tomáš Kubař & Christoph Jacob

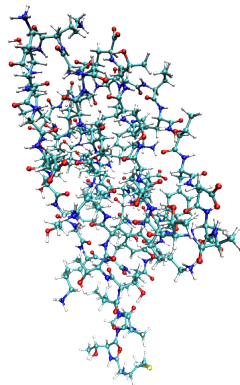
2017, June 23 and 27

Modeling of biomolecules

Potential energy surface: $E_{\text{el}} = E_{\text{el}}(\mathbf{R}_1, \dots, \mathbf{R}_N)$
(coordinates of atoms/nuclei $\mathbf{R}_1, \dots, \mathbf{R}_N$)

Approximations:

- Born–Oppenheimer approximation
 - separation of nuclei and electrons
 - E_{el} obtained for fixed positions of nuclei
- Classical description of nuclei
 - rather than quantum mechanics
for the motion of the nuclei
- Application of a force field
 - harmonic springs
 - point-charge electrostatics
 - ...



Energy with a force field

$$\begin{aligned} E_{\text{el}}(\mathbf{R}_1, \dots, \mathbf{R}_N) = & \frac{1}{2} \sum_i^{\text{bonds}} k_i (r_i - r_i^0)^2 + \frac{1}{2} \sum_i^{\text{angles}} k_i^\theta (\theta_i - \theta_i^0)^2 \\ & + \frac{1}{2} \sum_n^{\text{torsions}} \cos(n\omega_n - \gamma_n) \\ & + \sum_{I < J} \frac{1}{4\pi\epsilon_0} \frac{q_I q_J}{R_{IJ}} + \sum_{I < J} 4\epsilon_{IJ} \left(\left(\frac{\sigma_{IJ}}{R_{IJ}} \right)^{12} - \left(\frac{\sigma_{IJ}}{R_{IJ}} \right)^6 \right) \end{aligned}$$

Limitations of the force field approximation

The parameters have to be determined / fitted

- difficult for certain / unusual elements (e.g., transition metals)

Conceptual limitations

- chemical bonds cannot be broken or created
- atom types are pre-determined
- atomic charges are pre-determined and constant
 - no change of electron density can be described

Electrons are not described explicitly, so no wave function

- only ground-state energy and forces are available
- no spectroscopic properties (interaction with light)
- no photochemistry (excited states)

Quantum chemistry vs. force fields

Quantum mechanics:

electronic wave functions: $\Psi_{\text{el}} = \Psi_{\text{el}}(\mathbf{r}_1, \dots, \mathbf{r}_N)$
(coordinates of electrons $\mathbf{r}_1, \dots, \mathbf{r}_N$)

solve the electronic Schrödinger equation

$$\hat{H}^{\{R\}} \Psi^{\{R\}} = E_{\text{el}}(\mathbf{R}_1, \dots, \mathbf{R}_N) \Psi^{\{R\}}$$

Quantum chemistry

- wave function theory (WFT)
- density functional theory (DFT)

→ calculation of energy is computationally intensive

Force field methods

- evaluate E_{el} 'directly', do not look for the electronic structure
- calculation of energy is very quick and efficient

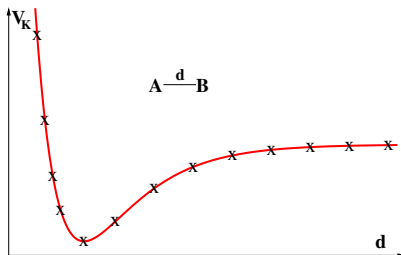
Potential energy surface

electronic energy is a function of coordinates of nuclei $\{\mathbf{R}_I\}$

$$E = E(\mathbf{R}_1, \dots, \mathbf{R}_N)$$

→ electronic energy defines the **potential energy surface** (PES)

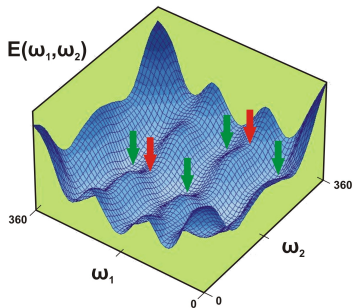
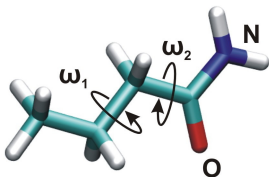
example: a diatomic molecule



→ all of the calculations only provide the PES point-wise

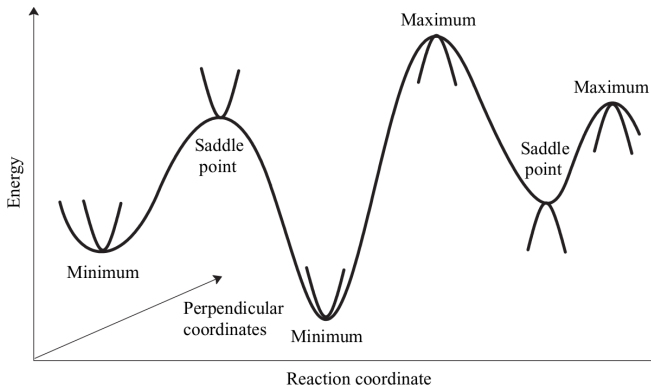
Sampling of the potential energy surface

chemically interesting:
stationary points
on the PES



- **minimum:**
‘equilibrium’ structure,
stable conformation of a molecule
- **saddle point of 1st order:**
transition state (TS),
point of maximum energy along the direction of a ‘reaction’

Potential energy surface

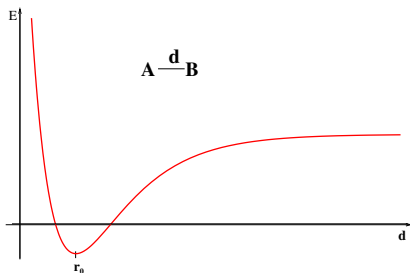


- **minimum:** any change of structure $\rightarrow$ increase of energy
- **saddle point of 1st order:** maximum along one coordinate, (reaction coordinate) $\times$ minimum along any other coordinate

Characterization of stationary points

How shall we find the interesting stationary points?

Example: Minimum of energy of a diatomic molecules



- condition for a stationary points: $\frac{dE}{dd} = 0$
- additional condition for a minimum: $\frac{d^2E}{dd^2} \geq 0$

Characterization of stationary points

How shall we find the interesting stationary points?

generally: $3N$ atom coordinates; $E = E(\mathbf{R}_1, \dots, \mathbf{R}_N)$

Condition for stationary points:

$$\text{Gradient: } \mathbf{g} = \nabla E = \left(\frac{dE}{dx_1}, \frac{dE}{dy_1}, \frac{dE}{dz_1}, \frac{dE}{dx_2}, \dots, \frac{dE}{dz_N} \right)^T = 0$$

Second derivatives:

$$\text{Hessian: } \mathbf{H} = \begin{pmatrix} \frac{d^2E}{dx_1^2} & \frac{d^2E}{dx_1dy_1} & \frac{d^2E}{dx_1dz_1} & \frac{d^2E}{dx_1dx_2} & \cdots \\ \frac{d^2E}{dy_1dx_1} & \frac{d^2E}{dy_1^2} & \frac{d^2E}{dy_1dz_1} & \frac{d^2E}{dy_1dx_2} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

- **Condition for minimum:** **all** eigenvalues of $\mathbf{H}$ are positive
- **Saddle point of 1st order:** **one** eigenvalue of $\mathbf{H}$ is negative, **all of the others** are non-negative

Characterization of stationary points

How shall we find the interesting stationary points?

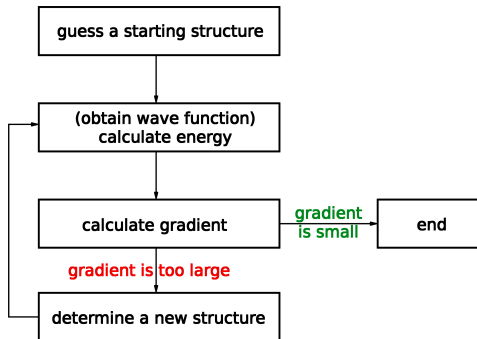
- Is it possible to sample the PES systematically?
 - example: 10 atoms, $3N - 6 = 24$) coordinates
 - 10 points per coordinate $\rightarrow 10^{24}$ calculations of energy!
 - 100 points per coordinate $\rightarrow 10^{240}$ calculations!
- $\rightarrow$ finding the global minimum – very difficult problem
- chemical intuition – concentrate on meaningful structures
- $\rightarrow$ local minima, transition states between low-energy regions

What do we need now?

- calculation of gradients (and Hessian)
- algorithm
 - to optimize the geometry (search for a local minimum)
 - to search for transition states

Geometry optimization

→ search for a local minimum, starting from a suitable structure



- starting structure?
 - depends on the chemical problem
- convergence criterion? e.g.

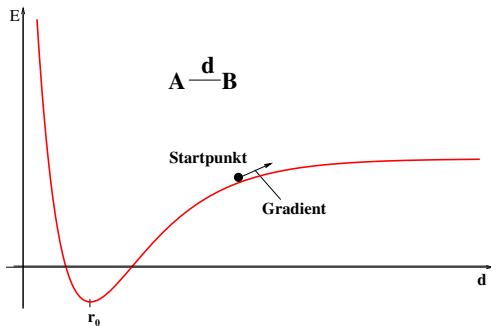
$$|\mathbf{g}| = \sqrt{\sum_{I,\alpha} g_{I,\alpha}^2} < t$$

- how shall we determine the new structure?
- try to make as few steps as possible
(calculation of energy/gradients is expensive)
- avoid any calculation of Hessian
(that is even more expensive)

Geometry optimization

How to make a step towards a minimum?

Example: geometry optimization of a diatomic molecule



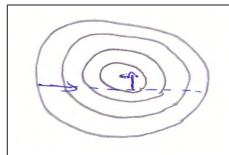
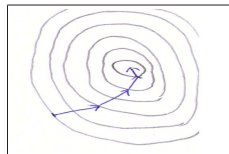
- follow the negative of gradient: $R_{i+1} = R_i + \Delta R_i = R_i - \alpha g_i$
- how shall the step length α be determined?

Steepest descents optimization

- Step along the negative of gradient

$$\Delta \mathbf{R}_i = \alpha \mathbf{d}_i \quad \mathbf{d}_i = -\mathbf{g}_i$$

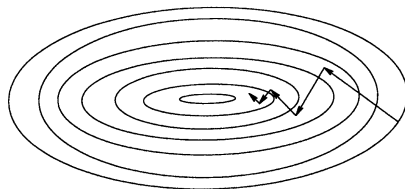
- Choice of the step length α ?
 - too short $\rightarrow$ too many steps needed
 - too long $\rightarrow$ overshoot the minimum
- 'line search':
 - choose α such that the energy in direction of gradient $\mathbf{d}_i = -\mathbf{g}_i$ is minimized
 - calculate the energy at several points along a line
- Then, the convergence is guaranteed.



Conjugate gradient optimization

Problems of steepest descents

- too many steps in similar directions (in narrow valleys)
- convergence is getting slower when close to the minimum



Solution – conjugate gradient optimization

- make the step $\mathbf{d}_i$ orthogonal to *all* of the previous ones

$$\mathbf{d}_i = -\mathbf{g}_i + \beta_i \mathbf{d}_{i-1}$$

- various schemes: Fletcher–Reeves or Polak–Ribiere:

$$\beta_i^{\text{FR}} = \frac{\mathbf{g}_i^t \mathbf{g}_i}{\mathbf{g}_{i-1}^t \mathbf{g}_{i-1}} \quad \beta_i^{\text{PR}} = \frac{(\mathbf{g}_i - \mathbf{g}_{i-1})^t \mathbf{g}_i}{\mathbf{g}_{i-1}^t \mathbf{g}_{i-1}}$$

Newton–Raphson optimization

Taylor expansion of the PES around R_i :

$$E(\mathbf{R}) = E(\mathbf{R}_i) + \mathbf{g}_i \cdot (\mathbf{R} - \mathbf{R}_i) + \frac{1}{2}(\mathbf{R} - \mathbf{R}_i)^t \cdot \mathbf{H}_i \cdot (\mathbf{R} - \mathbf{R}_i) + \dots$$

minimum of this Taylor expansion up to the 2nd order:

$$\frac{dE}{d\mathbf{R}} = \mathbf{g}_i + \mathbf{H}_i \cdot (\mathbf{R} - \mathbf{R}_i) = 0 \rightarrow \mathbf{H}_i \cdot (\mathbf{R} - \mathbf{R}_i) = -\mathbf{g}_i$$

Newton–Raphson optimization:

calculate the step toward the minimum from that expression:

$$\Delta \mathbf{R}_i = -\mathbf{H}_i^{-1} \cdot \mathbf{g}_i$$

(on a harmonic PES, this would lead to the minimum directly)

Problem: the calculation of Hessian in every step is expensive

Quasi-Newton optimization

Apply an approximated Hessian rather than the exact Hessian

- start e.g. with a unit matrix ($\rightarrow$ 1 steepest-descents step)
- other starting Hessians may be better
 - full calculation (just once), or from certain simple rules
- in every step, use the gradients $\mathbf{g}_i$
to improve the approximated Hessian $\tilde{\mathbf{H}}_i$,
then invert it for $\Delta \mathbf{R}_i = -\mathbf{H}_i^{-1} \cdot \mathbf{g}_i$
- usually converges quickly and reliably, at the cost of
storage for Hessian of $\mathcal{O}(N^2)$ and its inversion of $\mathcal{O}(N^3)$
- compare: CG is $\mathcal{O}(N)$ but converges more slowly
- a standard method in most quantum chemical / MD packages

Quasi-Newton optimization

various update algorithms are available

- Broyden–Fletcher–Goldfarb–Shanno (**BFGS**):

$$\tilde{\mathbf{H}}_{i+1} = \tilde{\mathbf{H}}_i + \frac{\Delta \mathbf{g}_i \otimes \Delta \mathbf{g}_i^t}{\Delta \mathbf{g}_i^t \cdot \Delta \mathbf{R}_i} - \frac{\tilde{\mathbf{H}}_i \cdot \Delta \mathbf{R}_i \otimes \Delta \mathbf{R}_i^t \cdot \tilde{\mathbf{H}}_i}{\Delta \mathbf{g}_i^t \cdot \tilde{\mathbf{H}}_i \cdot \Delta \mathbf{R}_i}$$

- symmetric and positive-definite
- minimizes the change in Hessian

- limited-memory BFGS (**L-BFGS**)

- propagate the inverse Hessian (not the Hessian itself)
 - the $\mathcal{O}(N^3)$ matrix inversion is eliminated
- do not store the Hessian or inverse Hessian in memory
 - rather, reconstruct the matrix on the fly
 - from $\Delta \mathbf{g}_i$ & $\Delta \mathbf{R}_i$ over a few (< 10) last steps
 - the $\mathcal{O}(N^2)$ storage requirement is eliminated

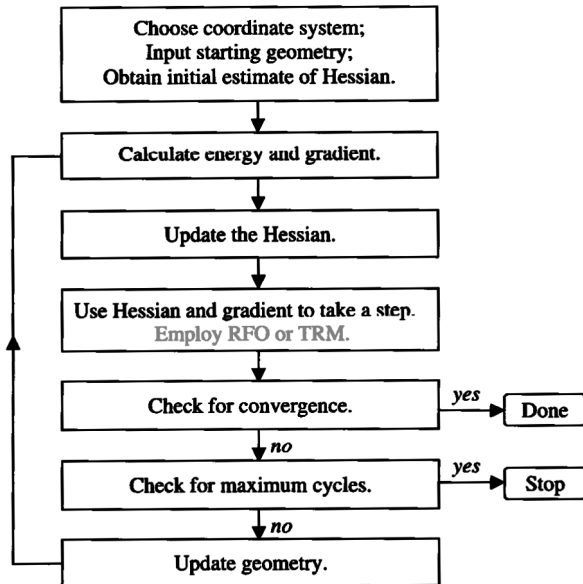
Quasi-Newton optimization

- q-N converges the better, the closer PES is to a quadratic form
- depends on **the choice of coordinate system** strongly

Possible choices of coordinate system

- cartesian coordinates
 - simple, but not adjusted to a ‘chemical’ problem
 - often slow convergence
- internal coordinates
 - use bond lengths, angles and torsional angles
 - often good convergence
 - but the definition of $3N - 6$ coordinates difficult
- redundant internal coordinates
 - use ‘too many’ internal coordinates
 - mostly good convergence
 - simple automatized definition is possible
- more complex coordinate systems possible

Quasi-Newton optimization



Summary: geometry optimization

- starting point: a chemically meaningful structure
- minimization procedures:
 - *steepest descents*: converges always, but slowly
 - better: *conjugate gradients*, *quasi-Newton* (e.g. *BFGS*)
 - all of these avoid the calculation of Hessian
- biomolecules – often very difficult to find true minima
- in the quantum chemistry
 - calculations mostly limited to a single minimum
 - starting point for the calculation of properties (spectra...)
- with the force field methods
 - starting point for MD around a minimum (and beyond)
 - pre-optimization for quantum chemical calculations

Potential energy surface

Energy is a function of atomic coordinates

$E_{\text{el},i}(\mathbf{R}_1, \dots, \mathbf{R}_N)$ – PES from quantum chemistry or force field

- Is it possible to generate the entire potential energy surface?
 - example: 10 atoms, $3N - 6 = 24$) coordinates
 - 10 points per coordinate $\rightarrow 10^{24}$ calculations of energy!
 - 100 points per coordinate $\rightarrow 10^{240}$ calculations!
- No! only concentrate on the region close to a minimum

Search for a local minimum: geometry optimization

- requires a chemically meaningful starting structure
- calculation of electronic energy and its gradients
 - in several (few) points of the coordinate space

In the region close to a local minimum

- apply Taylor expansion to the PES

Harmonic approximation

- move the origin of coordinates into the local minimum of PES:

$$\mathbf{R}_{I,0} = \mathbf{o} \text{ (null vector)}$$

→ the variables $\{\mathbf{R}_I\}$ give the deviation from minimum

- Taylor expansion of energy

$$\begin{aligned}
 E_{\text{el}}(\{\mathbf{R}_I\}) &= \underbrace{E_{\text{el}}(\{\mathbf{o}\})}_{\text{const.}} + \underbrace{\sum_{I=1}^M \sum_{\alpha=1}^3 \left(\frac{\partial E_{\text{el}}}{\partial R_{I,\alpha}} \right)_{\mathbf{R}_I=\mathbf{o}} R_{I,\alpha}}_{=0} \\
 &+ \frac{1}{2} \sum_{I,J}^M \sum_{\alpha,\beta=1}^3 R_{I,\alpha} \left(\frac{\partial^2 E_{\text{el}}}{\partial R_{I,\alpha} \partial R_{J,\beta}} \right)_{\substack{\mathbf{R}_I=\mathbf{o} \\ \mathbf{R}_J=\mathbf{o}}} R_{J,\beta} + \mathcal{O}(R_{I,\alpha}^3)
 \end{aligned}$$

→ gradients in a local minimum of PES **vanish**

Harmonic approximation

- terminate the Taylor series after the second order
- write the 2nd derivatives in a matrix form – **Hessian**:

$$\mathbf{F} = \{F_{ij}\} = \left\{ \frac{\partial^2 E_{\text{el}}}{\partial R_i \partial R_j} \right\}$$

- coupling of coordinates of different nuclei
 - the same property – covariance matrix in PCA!
- condense the atomic masses into the coordinates:

$$\begin{aligned} \mathbf{R}_I^{(m)} &= \sqrt{M_I} \mathbf{R}_I \\ \mathbf{F}^{(m)} &= \mathbf{M}^{-1/2} \cdot \mathbf{F} \cdot \mathbf{M}^{-1/2} \end{aligned}$$

- look for: a linear combination of mass-weighted coordinates that diagonalizes the Hessian → **normal coordinates** $\{q_i\}$

Harmonic approximation

- normal coordinates $\{q_i\}$, normal modes $\{Q^{(i)}\}$:

$$q_i = \sum_{j=1}^{3N} Q_j^{(i)} R_j^{(m)} \quad \text{so that} \quad \frac{\partial^2 E_{\text{el}}}{\partial q_i \partial q_j} = 0 \quad \forall i \neq j$$

$$\mathbf{Q} = (\mathbf{Q}^{(1)}, \dots, \mathbf{Q}^{(3N)})^T \quad \text{and} \quad \mathbf{Q}^{(i)} = (Q_{1,x}^{(i)}, \dots, Q_{3N,z}^{(i)})$$

$$\mathbf{F}^{(q)} = \mathbf{Q}^T \cdot \mathbf{F}^{(m)} \cdot \mathbf{Q} \quad \text{is the diagonalization problem to solve}$$

- eigenvalues of energy of the harmonic oscillator
 - vibrational frequencies ω_i follow from Hessian eigenvalues

$$\omega_i = \sqrt{F_{ii}^{(q)}}$$

note: with PCA upon MD trajectory, harmonic frequencies are obtained from the eigenvalues of the covariance matrix

- for molecules in the gas phase:
 - translations+rotations are separated from the vibrations

Summary: vibrational analysis

- starting point: local minimum on the PES
- calculate the Hessian – most computationally intensive step:

$$\mathbf{F} = \left\{ \frac{\partial^2 E_{\text{el}}}{\partial R_i \partial R_j} \right\}$$

- introduce mass weighting:

$$\mathbf{F}^{(m)} = \mathbf{M}^{-1/2} \cdot \mathbf{F} \cdot \mathbf{M}^{-1/2}$$

- diagonalize the mass-weighted Hessian:

$$\mathbf{F}^{(q)} = \mathbf{Q}^T \cdot \mathbf{F}^{(m)} \cdot \mathbf{Q}$$

- calculate the vibrational frequencies from the eigenvalues:

$$\omega_i = \sqrt{F_{ii}^{(q)}}$$

Transition states

How to describe a chemical reaction / another process of interest?

- structure/energy of reactant/product → reaction energy ΔH_r
- structure/energy of the transition state → **rate** of the process

- transition state theory (TST):

$$k = \frac{k_B T}{h} \cdot \exp \left[-\frac{\Delta G^\ddagger}{k_B T} \right]$$

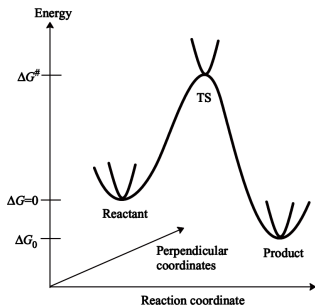
- free energy of the TS:

$$\Delta G^\ddagger = \Delta H^\ddagger - T\Delta S^\ddagger$$

- from quantum chemistry:

$$\Delta H^\ddagger = E_{\text{TS}} - E_{\text{reactant}}$$

- from vibrational analysis: $\Delta S^\ddagger$
(via stat. thermodynamics)



Optimization of transition states

‘One-structure’ / local methods

- start with a suitable initial structure
- optimize TS with a (quasi-)Newton method
- follow one eigenvector **up**
to reach one negative eigenvalue of $\mathbf{H}$

Necessary:

- starting structure already close to the TS
- good starting Hessian (for quasi-Newton)
- good choice of the coordinate system
even more important than in a usual geometry optimization

Optimization of transition states

‘One-structure’ / local methods

- start with a suitable initial structure
- optimize TS with a (quasi-)Newton method
- follow one eigenvector **up**
to reach one negative eigenvalue of H

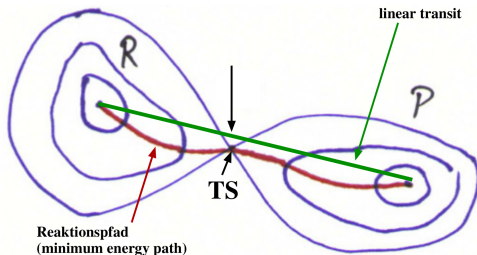
Common procedures:

- optimize the geometry of the reactant and the product
- search for a good starting structure
 - e.g. linear transit between the reactant and the product
- for quasi-Newton, usually calculate the full Hessian for the starting structure to use as a starting Hessian

Optimization of transition states

How to find a suitable starting structure for the TS search?

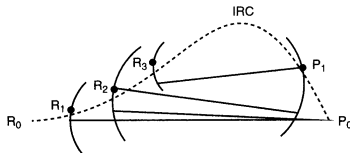
- should be good – as close as possible to the true TS
- easiest way: follow a (guessed) reaction coordinate
e.g. bond length/angle that is changing during reaction
- **linear transit**: vary this coordinate in regular steps
- in every step, optimize all of the other coordinates
(‘constraint optimization’ with the selected coordinate fixed)
- maximum of energy along this path → starting structure



Optimization of transition states

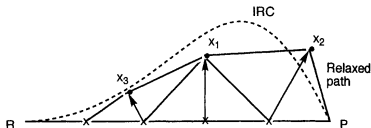
Alternative: 'two-structure' methods

- start from reactant and product; encircle the TS step-by-step



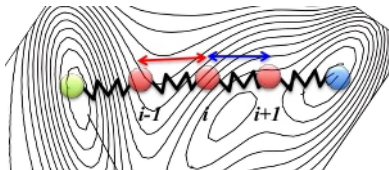
Alternative: 'multiple-structure' methods

- start with a linear interpolation between reactant and product
- optimize structuren along the entire reaction path



Nudged elastic bands

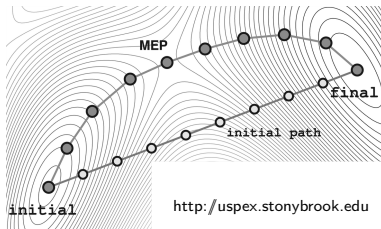
- method for finding saddle points and minimum energy paths between known reactants and products
- optimizes a number of intermediates along the reaction path



- each image finds the lowest energy possible while maintaining equal spacing to neighboring images
- restrained optimization:
 - 1 add harmonic springs between intermediates along the band
 - 2 project out the component of spring force $\perp$ to the band

Nudged elastic bands

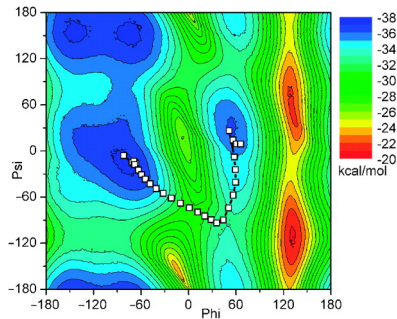
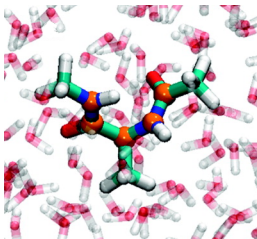
- construct a set of replicas of the system ($\approx 4 - 20$) between the initial and final state
- add a harmonic spring interaction between adjacent replicas – ensure continuity, much like an elastic ‘band’ would do
- during optimization, the ‘band’ induces forces on the replicas and brings the replicas to the minimum energy path



<http://uspex.stonybrook.edu>

Nudged elastic bands – example

- alanine dipeptide in implicit or explicit solvent
- ‘partial NEB’ – the NEB forces are applied on a part of the molecular system (solute, or its part)

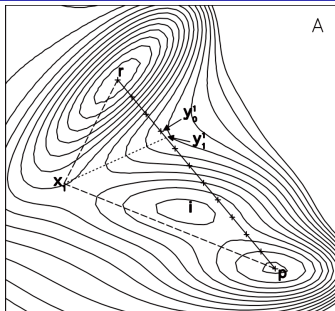


Conjugate peak refinement

CPR

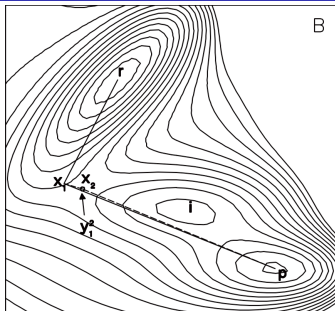
- robust method to locate TS for systems with many atoms
- iterative algorithm
- a single intermediate point is either added, refined or removed during a given iteration
- finally: the only local maxima in energy left along the path are true saddle points (TS)

Conjugate peak refinement



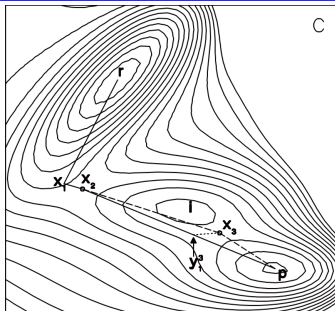
- coarse scan along the line $r - p \rightarrow$ maximum near y_0^1
(s_0 is the direction $r - p$)
- **maximization** along the line $r - p \rightarrow$ maximum y_1^1
- **minimization** along conjugate vector from $y_1^1 \rightarrow$ minimum x_1
(the conjugate vector is $s_1 = -g_1 + \frac{g_1^T h}{s_0^T h} s_0$ with $h \approx Hs_0$)
- molecular systems – minimize along several conjug. vectors

Conjugate peak refinement



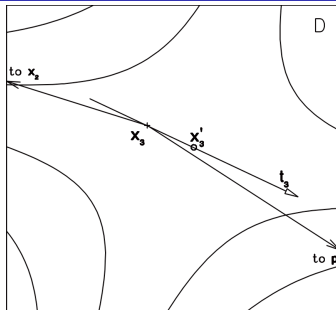
- x_1 is new intermediate on the path which now has 2 segments
- repeat the procedure along the lines $r - x_1$ and $x_1 - p$
- find the maximum y_1^2 and,
 - along the conjugate direction, the minimum x_2
- x_2 is already a saddle point – meets 2 conditions:
 - RMS gradient of the energy at x_2 is nearly zero
 - x_2 is still a maximum despite resulting from line minimizations

Conjugate peak refinement



- repeat the procedure along the line $x_2 - p$
- find the maximum y_1^3 and,
along the conjugate direction, the minimum x_3
- RMS gradient at x_3 is small, so it is close
to a stationary point (here, saddle point)

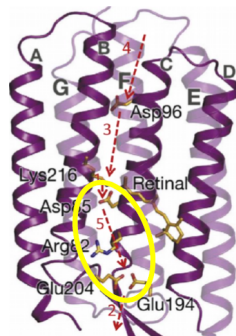
Conjugate peak refinement



- x_3 is refined to produce a saddle point
- search is performed along the line t_3
which is an average of $x_3 - x_2$ and $p - x_3$
- line maximization followed by conjugate line minimization(s)
gives the saddle point x'_3
 - RMS gradient of the energy at x'_3 is nearly zero
 - x'_3 is still a maximum despite resulting from line minimizations

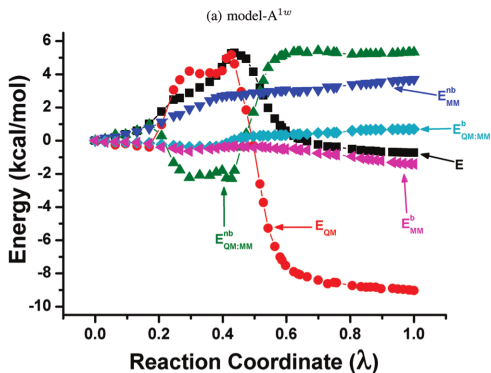
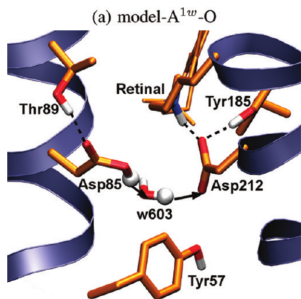
Conjugate peak refinement – example

- active site of bacteriorhodopsin
- long-range proton transfer reaction over several intermediates
- first half of the last proton transfer step
- chemical reaction → QM/MM necessary
- the various structures built ('models') differ in the number of H₂O molecules in the active site
→ very different minimum energy paths



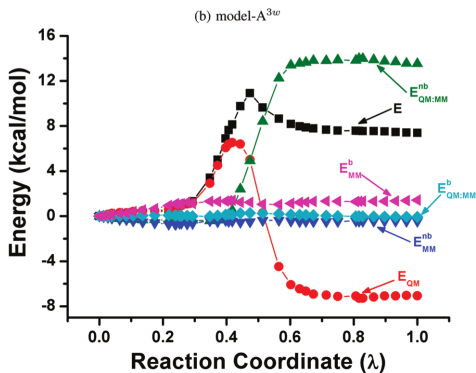
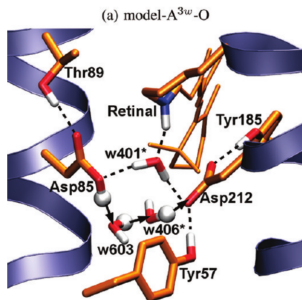
Conjugate peak refinement – example

structural model with 1 water molecule



Conjugate peak refinement – example

structural model with 3 water molecules



Summary: transition states

- Optimization of TS is much more difficult than a normal optimization of an energy minimum
- *one-structure methods*:
 - good starting structure and starting Hessian are essential
- *multiple-structure methods*:
 - usually more robust, but not always available/applicable
- after the optimization of TS:
 - always calculate the Hessian to check if it really is a TS!
- in quantum chemistry:
 - description of chemical reactions ($\rightarrow$ activation energy)
 - but, reaction rates require very accurate energies
- with force field methods:
 - chemical reactions impossible, only conformational changes
 - then, the meaning of TS not always clear