C7790 Introduction to Molecular Modelling

TSM Modelling Molecular Structures

C9087 Computational Chemistry for Structural Biology

Lesson 16

Reaction Energy I (general)

JS/2022 Distant Form of Teaching: Rev1

Petr Kulhánek

kulhanek@chemi.muni.cz

National Centre for Biomolecular Research, Faculty of Science
Masaryk University, Kamenice 5, CZ-62500 Brno

Context

macroworld **microworld**

states

(thermodynamic properties, G, T,...)

phenomenological thermodynamics

equilibrium (equilibrium constant)

kinetics (rate constant)

free energy

(Gibbs/Helmholtz)

partition function

statistical thermodynamics

microstates

(mechanical properties, E)

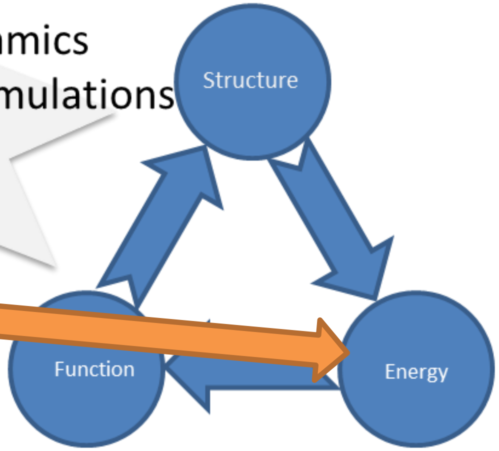
Description levels (model chemistry):

- quantum mechanics
  - semiempirical methods
  - ab initio methods
  - post-HF methods
  - DFT methods
- molecular mechanics
- coarse-grained mechanics

Simulations:

- molecular dynamics
- Monte Carlo simulations
- docking
- ...

bridge



microstate ≠ microworld

Revision: Thermodynamics & Modelling

$$aA + bB \rightleftharpoons cC + dD$$

Fundamental relation

$$\boxed{\Delta G_r^{0} = -RT\ln K}$$

$$\Delta G_r^0 = c\Delta G_{f,C}^0 + d\Delta G_{f,D}^0 - \left(a\Delta G_{f,A}^0 + b\Delta G_{f,B}^0\right)$$

$$K = \frac{[C]_r^c[D]_r^d}{[A]_r^a[B]_r^b}$$

ΔG_r^0 is the standard reaction free energy.

- its value determines the composition of the reaction mixture at the chemical equilibrium.

Revision: Thermodynamics & Modelling

Fundamental relation

$$\Delta G_r^0 = -RT \ln K$$

$$\Delta G_r^0 = c\Delta G_{f,C}^0 + d\Delta G_{f,D}^0 - \left(a\Delta G_{f,A}^0 + b\Delta G_{f,B}^0\right)$$

$$K = \frac{[C]_r^c [D]_r^d}{[A]_r^a [B]_r^b}$$

What do we need to know?

A

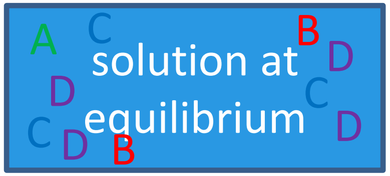
B

C

D

We only need to know the **properties of individual components** involved in the reaction at standard conditions (or at different conditions, which are well defined).

easier for modelling



We need to know the **composition of solution** at equilibrium.

It is hard or impossible to model.

Revision: Partition function and modelling

Canonical partition function:

$$Q = \sum_{j=1}^{K} e^{-\beta E_j}$$

Helmholtz energy F:

$$F = -k_B T \ln Q$$

approximation

Consider only the most important microstate

E_1, ~~E_2, E_3, ...~~

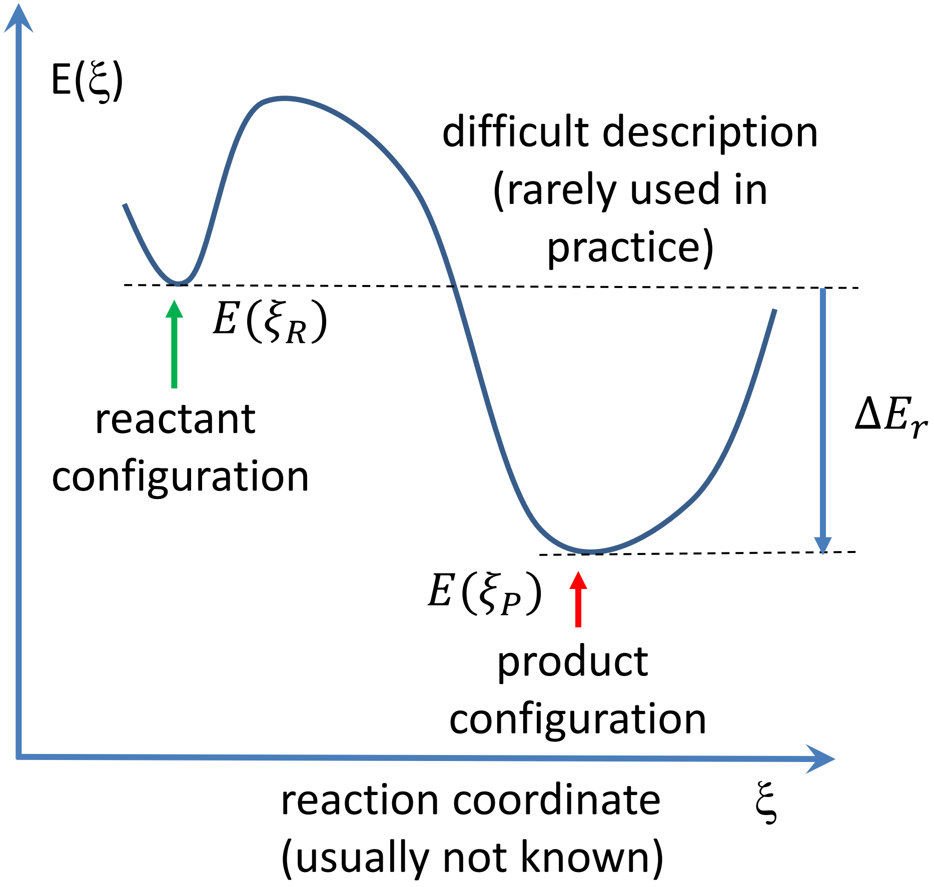
$$F = E_1$$

The most important microstate is the microstate with the lowest energy.

It is very often used for qualitative consideration or when computationally demanding methods are employed (typically quantum chemical calculations).

Reaction Energy

REMEMBER: This is 1D projection of E(R), which is a function of 3N variables (N-number of atoms).



E(ξ)

difficult description (rarely used in practice)

$E(\xi_R)$

reactant configuration

ΔE_r

$E(\xi_P)$

product configuration

reaction coordinate (usually not known) ξ

reaction energy

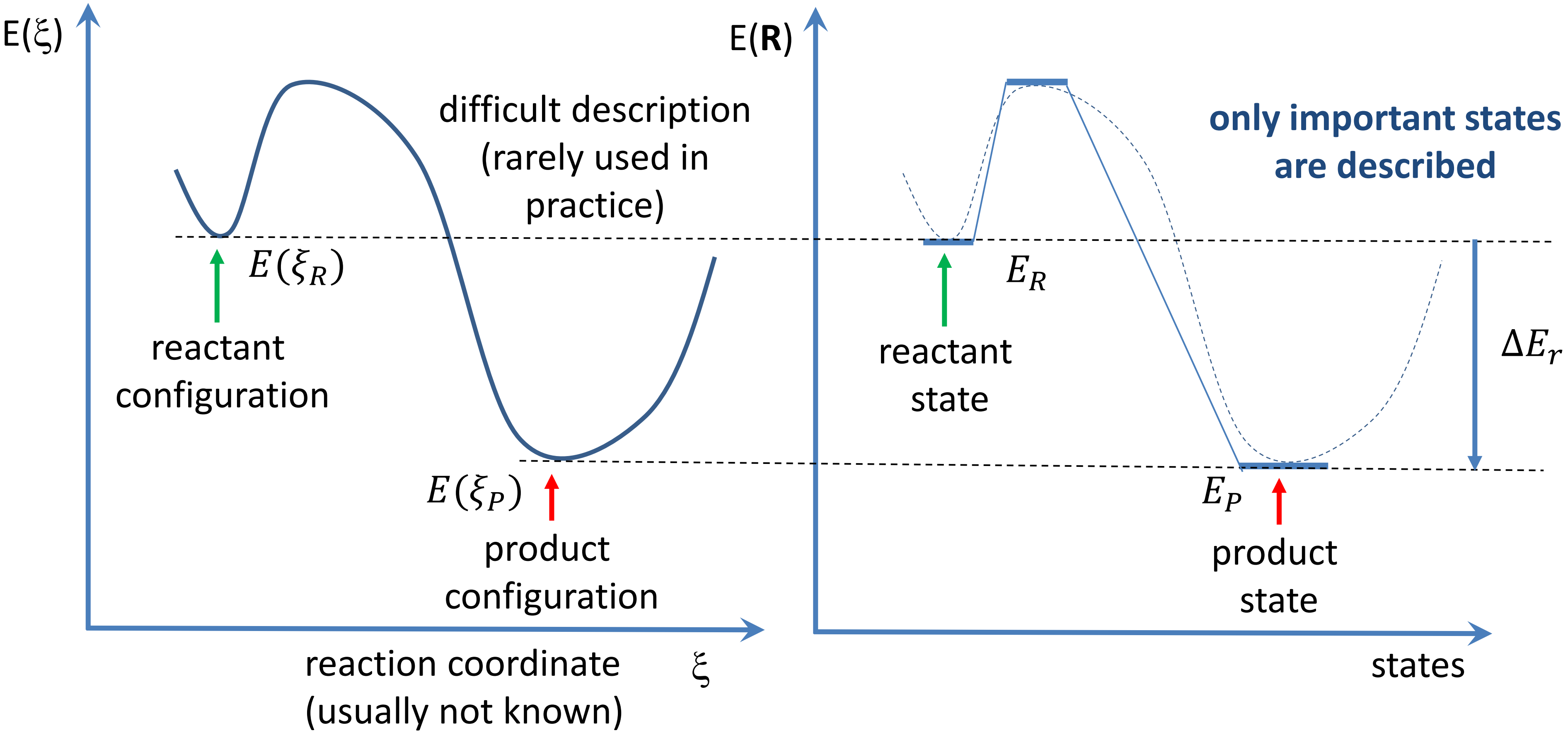
$$\Delta E_r = E(\xi_P) - E(\xi_R)$$

- **potential energy is a state function**, thus the reaction path is not necessary for reaction energy calculation
- only reactant and product states need to be characterized

the sign convention:
always use the thermodynamics convention

Reaction Energy

REMEMBER: This is 1D projection of E(R), which is a function of 3N variables (N-number of atoms).



difficult description (rarely used in practice)

only important states are described

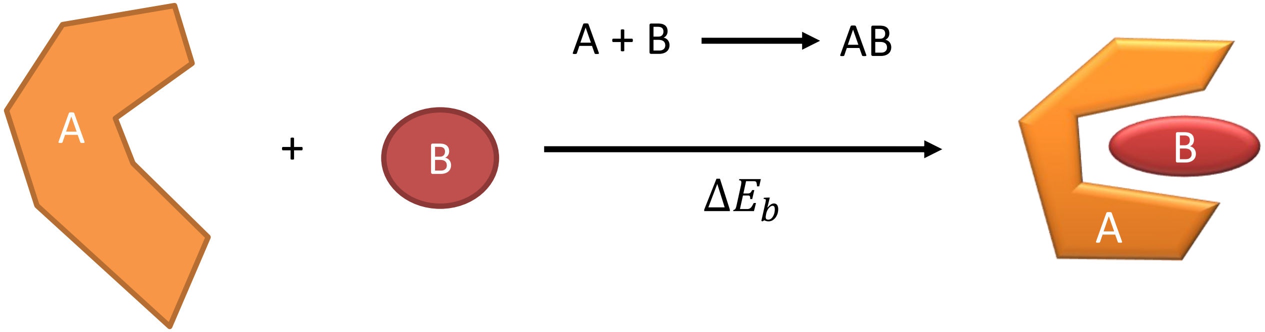
reaction energy

$$\Delta E_r = E(\xi_P) - E(\xi_R) = E_P - E_R$$

the sign convention:
always use the thermodynamics convention

Binding Energy

Binding is the special case of reaction.



$$A + B \longrightarrow AB$$

$$\xrightarrow[\Delta E_b]{}$$

Reactant state:

- model of non-interacting components

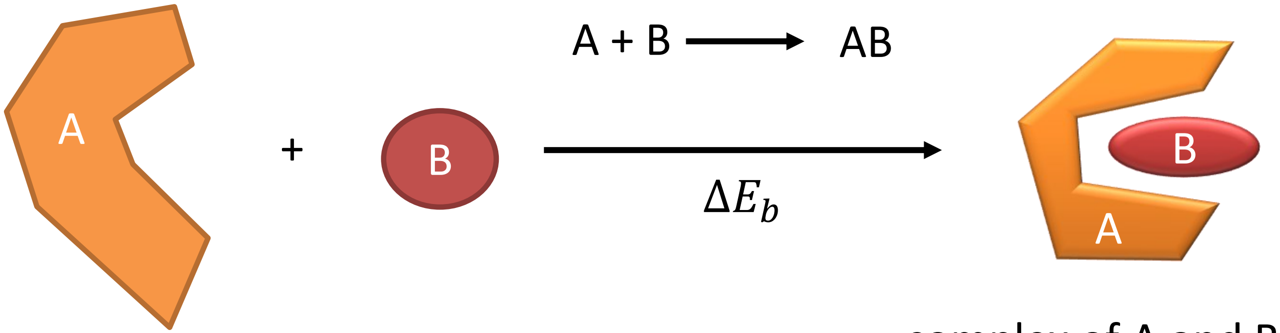
Product state:

- complex between A and B, usually non-covalently bound

Binding energy:

$$\Delta E_b = \Delta E_r = E_{AB} - \overline{(E_A + E_B)}$$

Binding Energy



$A + B \longrightarrow AB$

complex of A and B,
usually non-covalently
bound

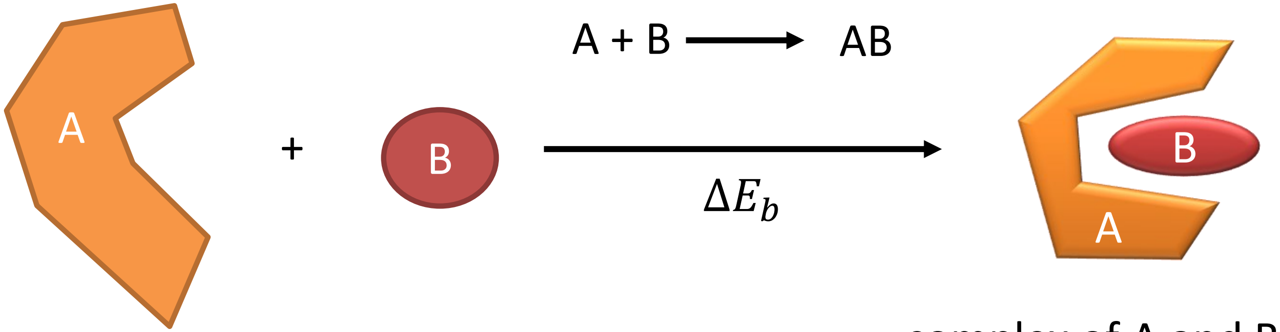
Binding energy:

$$\Delta E_b = \Delta E_r = E_{AB} - (E_A + E_B)$$

Source of energies (model chemistry):

- potential energy of geometry optimized states (QM, MM, QM/MM, etc.)
  - REMEMBER: it does not cover ZPVE (zero-point vibrational energy)
  - no thermal contributions
- ….
- various improvements to E(R): ZPVE, ideal gas model, etc. …
- ….
- free energy
- ….
- scoring functions (docking)

Binding Energy



complex of A and B,
usually non-covalently
bound

Binding energy:

$$\Delta E_b = \Delta E_r = E_{AB} - (E_A + E_B)$$

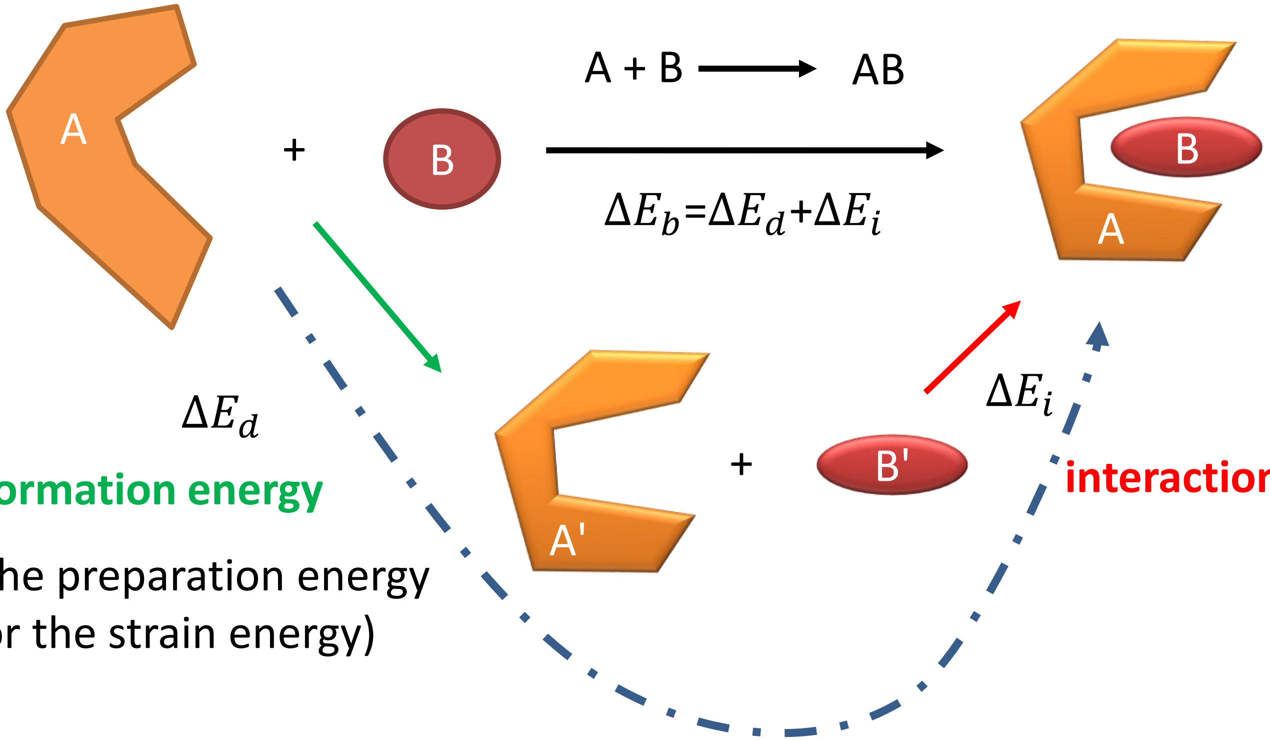
Source of energies (model chemistry):

- potential energy of geometry optimized states (QM, MM, QM/MM, etc.)
  - REMEMBER: it does not cover ZPVE (zero-point vibrational energy)
  - no thermal contributions
- ….
- various improvements to E(R)
- ….
- free energy
- ….
- scoring functions (docking)

Sometimes it is convenient to decompose the binding energy into deformation and interaction energies.

Binding vs Interaction Energy

Binding energy is the difference between the energy of the complex minus the energy of **the isolated monomers in their minima configuration**.



binding energy

A + B → AB

$\Delta E_b = \Delta E_d + \Delta E_i$

complex of A and B, usually non-covalently bound

ΔE_d

deformation energy

(also the preparation energy or the strain energy)

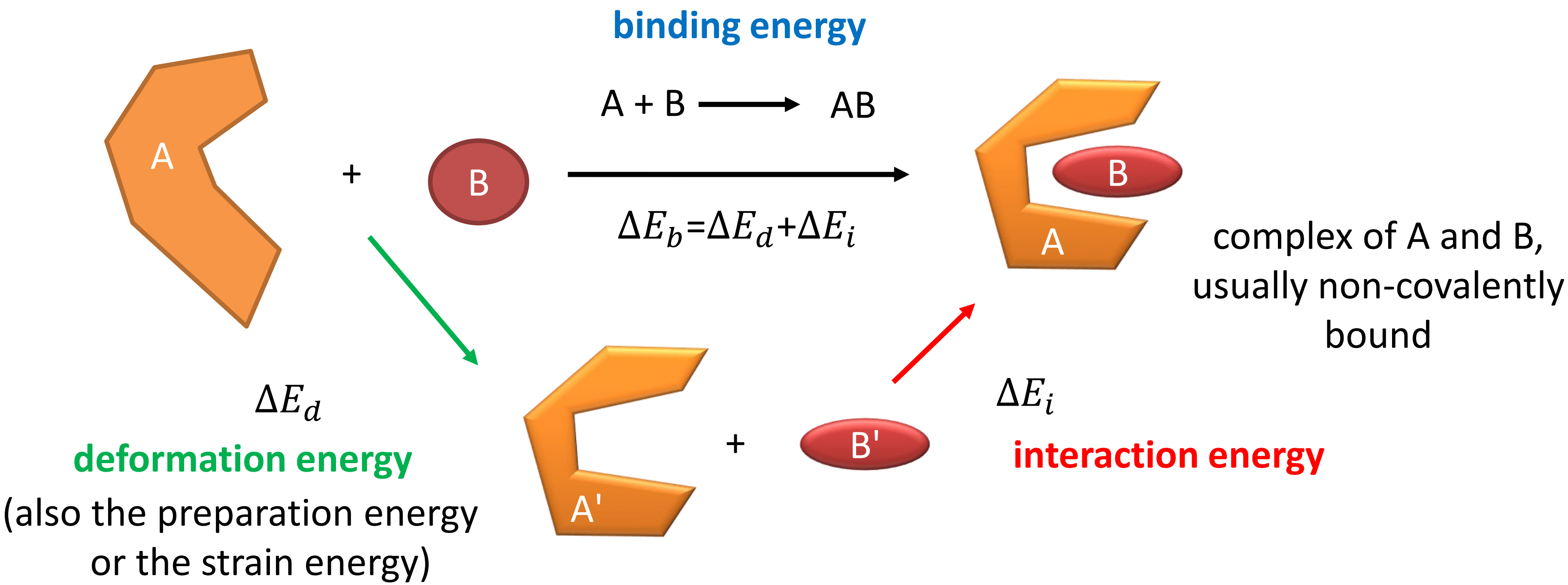
ΔE_i

interaction energy

!!! though (hypothetical) pathway, experimentally non-observable !!!
its main purpose is to better understand the binding process

Binding vs Interaction Energy

Binding energy is the difference between the energy of the complex minus the energy of **the isolated monomers in their minima configuration**.



binding energy

$A + B \longrightarrow AB$

$\Delta E_b = \Delta E_d + \Delta E_i$

complex of A and B,
usually non-covalently
bound

ΔE_d

deformation energy

(also the preparation energy
or the strain energy)

ΔE_i

interaction energy

Deformation energy is energy required to deform monomers from their relaxed geometry into the geometry observed in the complex. **This energy is ALWAYS positive**.

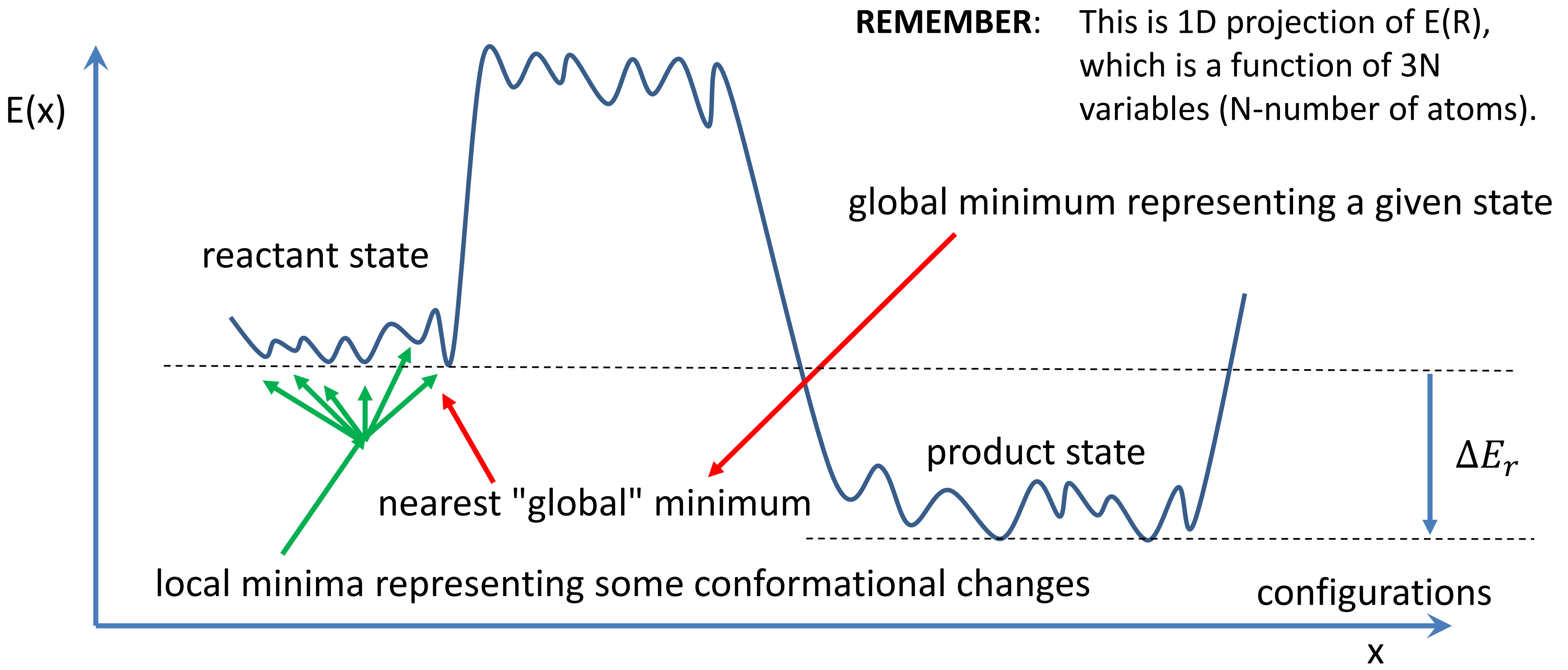
Interaction energy is the difference between the energy of the complex minus the energy of **the isolated monomers in the geometry of the complex**.

Binding vs Interaction Energy

- Terms "**binding energy**" and "**interaction energy**" are very often used interchangeably. Therefore, you have to follow the intended context of a work.

- With increasing model size, the calculation of the binding or interaction energy employing potential energy becomes difficult.

Large models



- Increasing degrees of freedom (model size) result in increased roughness of PES.
- The nearest "global" minimum representing the state could be difficult or impossible to select.
- Then, the only reasonable solution is to use statistical weighting using the free energy calculations.

Large models, cont.

- Then, the only reasonable solution is to use statistical weighting using the free energy calculations.

All microstates representing a thermodynamic state need to be considered:

$$E_1, E_2, E_3, \ldots$$

Canonical partition function:

$$\boxed{Q = \sum_{j=1}^{K} e^{-\beta E_j}}$$

Helmholtz energy F:

$$\boxed{F = -k_B T \ln Q}$$

Typical scenario:

- models employing explicit solvent models

Approches:

- molecular dynamics
- Monte Carlo simulations
- etc.

Uncharacterizable compounds

Fundamental relation

$$\Delta G_r^{0} = -RT \ln K$$

$$\begin{aligned}\Delta G_r^0 &= c\Delta G_{f,C}^0 + d\Delta G_{f,D}^0 \\ &\quad -\left(a\Delta G_{f,A}^0 + b\Delta G_{f,B}^0\right)\end{aligned}$$

A

B

C

D

We only need to know the **properties of individual components** involved in the reaction at standard conditions (or at different conditions, which are well defined).

easier for modelling

What to do if some compound is difficult to describe by modeling?

Typical example: pKa calculations

Example: pKa calculation

This reaction describes the most important chemical change, but the structure **of solvated proton is more complex than H_3O^+** in reality.

$$HA + H_2O \rightleftharpoons H_3O^+ + A^-$$

Definitions:

$$pK_a = -\log(K_a)$$

$$K_a = \frac{\{H_3O^+\}_r \, \{A^-\}_r}{\{HA\}_r \, \{H_2O\}_r}$$

$\{H_2O\}_r = 1$ (activity of bulk water is one)

pKa calculation

$$HA + H_2O \rightleftharpoons H_3O^+ + A^-$$

(this eq. is already simplified description of reality)

Definitions:

$$pK_a = -\log(K_a)$$

$$K_a = \frac{\{H_3O^+\}_r \, \{A^-\}_r}{\{HA\}_r \, \{H_2O\}_r}$$

$\{H_2O\}_r = 1$ (activity of bulk water is one)

approximation

$$HA \rightleftharpoons H^+_{solv} + A^-$$

difficult or impossible to model

$$K_a = \frac{\{H^+{}_{solv}\}_r \, \{A^-\}_r}{\{HA\}_r}$$

How to model (characterize) $H^+{}_{solv}$?

they can be modelled

pKa calculation

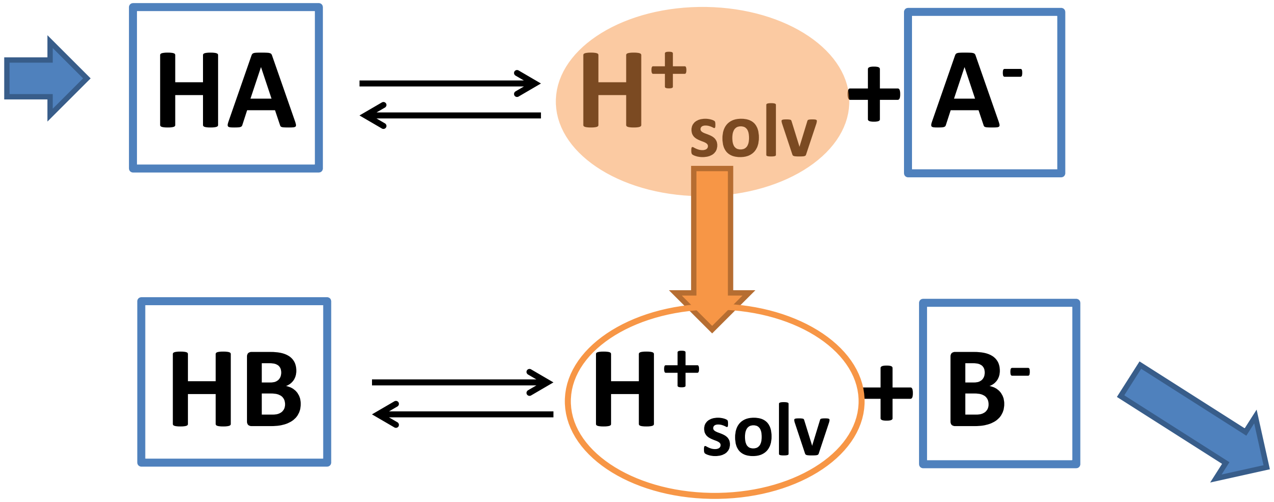
$$HA \rightleftharpoons H^{+}_{solv} + A^{-}$$

difficult or impossible to model

$$K_a = \frac{\{H^{+}_{solv}\}_r \{A^{-}\}_r}{\{HA\}_r}$$

they can be modelled

Solution:



experimentally known $pK_a(HA)$

$$HA \rightleftharpoons H^{+}_{solv} + A^{-}$$

$$HB \rightleftharpoons H^{+}_{solv} + B^{-}$$

prediction of $pK_a(HB)$

☐ modelled compounds

pKa calculation - "training"

experimentally known pK_a(HA) ⇨

$$\boxed{\mathrm{HA}} \rightleftharpoons \mathrm{H^+_{solv}} + \boxed{\mathrm{A^-}}$$

Definitions:

$$pK_a = -\log(K_a)$$

$$K_a = \frac{\{\mathrm{H^+_{solv}}\}_r\,\{\mathrm{A^-}\}_r}{\{\mathrm{HA}\}_r}$$

$$\Delta G_r^0 = -RT\ln K_a$$

$$\Delta E_r = E_{H_{solv}^{(+)}} + E_{A^{(-)}} - E_{HA}$$

Solution:

$$\Delta G_r^0 = -RT\ln K_a \longrightarrow e^{-\frac{\Delta G_r^0}{RT}} = K_a \longrightarrow \frac{\Delta G_r^0}{RT}\log(e) = pK_a$$

$$\searrow$$

$$pK_a = \frac{\Delta G_r^0}{2.303RT} \approx \frac{\Delta E_r}{2.303RT}$$

$$\uparrow$$

approximation

$$E_{H_{solv}^{(+)}} = \Delta E_r - E_{A^{(-)}} + E_{HA} \longleftarrow pK_a$$

$$\downarrow$$

$$E_{H_{solv}^{(+)}} = 2.303RT\,pK_a(HA) - E_{A^{(-)}} + E_{HA}$$

pKa calculation - "prediction"

$$\boxed{\mathrm{HB}} \rightleftarrows \mathrm{H}^{+}_{\mathrm{solv}} + \boxed{\mathrm{B}^{-}}$$

prediction of
pK_a(HB)

Definitions:

$$pK_a = -\log(K_a)$$

$$K_a = \frac{\{\mathrm{H}^{+}{}_{\mathrm{solv}}\}_r \ \{\mathrm{A}^{-}\}_r}{\{\mathrm{HA}\}_r}$$

$$\Delta G_r^0 = -RT \ln K_a \qquad \Delta E_r = E_{H_{solv}^{(+)}} + E_{A^{(-)}} - E_{HA}$$

Solution:

$$E_{H_{solv}^{(+)}} = 2.303RTpK_a(HA) - E_{A^{(-)}} + E_{HA}$$

$$E_{H_{solv}^{(+)}} + E_{B^{(-)}} - E_{HB} = 2.303RTpK_a(\mathrm{HB})$$

$$\boxed{pK_a(HA) + \underbrace{\frac{-E_{A^{(-)}} + E_{HA} + E_{B^{(-)}} - E_{HB}}{2.303RT}}_{\text{the difference between two } pK_a \text{ is obtained from the modelling}} = pK_a(\mathrm{HB})}$$

Summary

- For small size models, it is rather simple to characterize all components unambiguously on PES.
- Calculation of reaction (binding) energy is then straightforward **(BUT NOT IN QM ☹)**.
- Binding energy (as special case of reaction energy) quantifies the binding strength of two or more components, for example: protein and inhibitor.
  - The binding energy can be conveniently decomposed into the deformation and interaction energies.
  - The deformation energy is ALWAYS positive (Why?).
  - For strong binders, the interaction energy must be then large enough to counterbalance the deformation energy.
- Calculation of reaction energy can be supplemented by experimental data (typical use: pKa calculations).
- For large models, the simple approach employing PES for reaction (binding) energy calculation is not suitable and more advance sampling techniques must be employed to get reasonable data.