pKa Calculations of Key Ionizable Protein Residues in Acetylcholinesterase

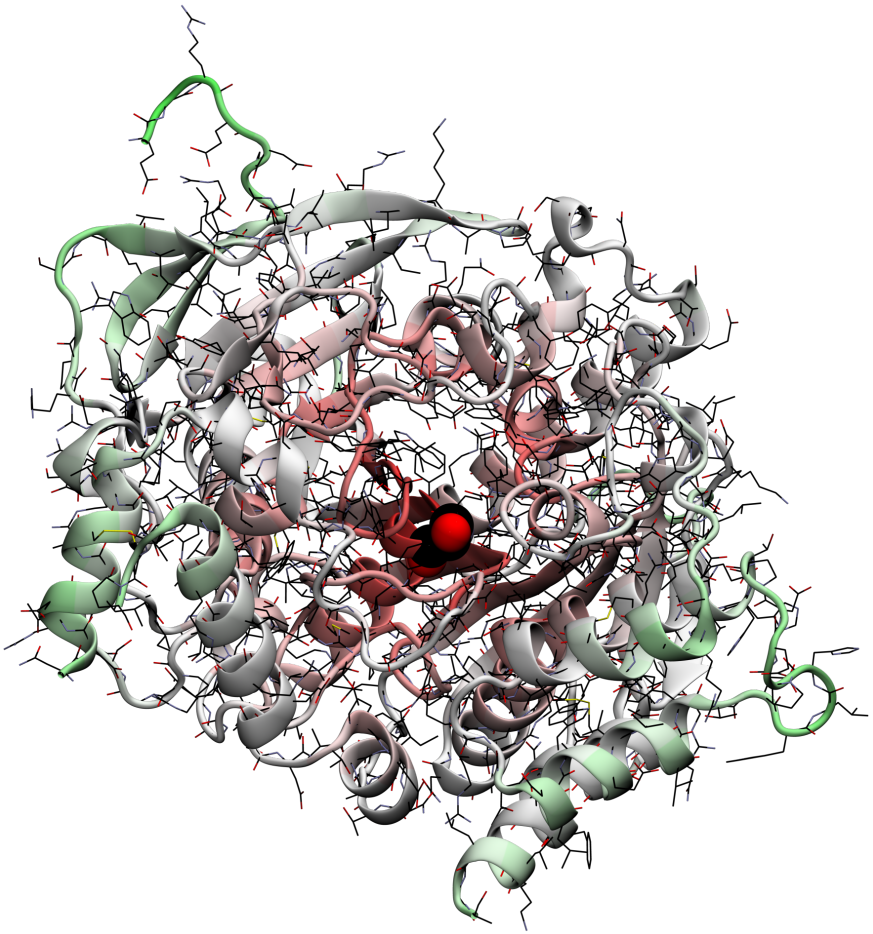
Jiří Wiesner
VOCE

Masaryk University
Czech Republic

Outline

- **Introduction**
  - *motivation, acetylcholinesterase, thioredoxin*
- **Theory**
  - *free energy, thermodynamic integration*
- **Methods**
  - *simulation set-up*
- **Suggestions**
  - *GLUE schema*
- **Results**
  - *pK_a shifts of thioredoxin*
- **Conclusions**

Acetylcholinesterase & thioredoxin



Acetylcholinesterase (2HA2) with Ser203

- **enzyme participating in the nerve signal transmission**
- **3 buried, possibly charged aminoacids in the active site**
  - *Glu334, Glu202, Glu450*
- **pK_a → protonation states**
- **Thioredoxin**
  - *benchmark protein*
  - *buried Asp26*
  - *experiment:*
    - 7.5 (4.8 kcal/mol)
  - *calculation* in Amber ff94:*
    - 12.5 (9.1 kcal/mol)

* *Simonson et al., J.Am.Chem.Soc. 2004, 126, 4167.*

Basic terms in pK_a computation

- **Molecular dynamics (MD)**
  - method to simulate motions of molecules in atomistic detail by solving Newton's equations motion numerically
  - analytical potential – force field (Amber ff99SB, CHARMM 22/CMAP)
- **Gibbs free energy**
  - amount of reversible work needed to transfer a system between an initial and final state under constant pressure and temperature conditions
- **Thermodynamic integration (TI)**
  - method how to compute free energy using molecular dynamics, when the system is transformed between two states – “potentials”
    - *transformation is non-physical – atoms may be added or removed*
- **pK_a**
  - quantitative measure of the strength of an acid in solution

pKa and Gibbs free energy

- **Brønsted-Lowry theory – acid HA and its conjugated base A^-**

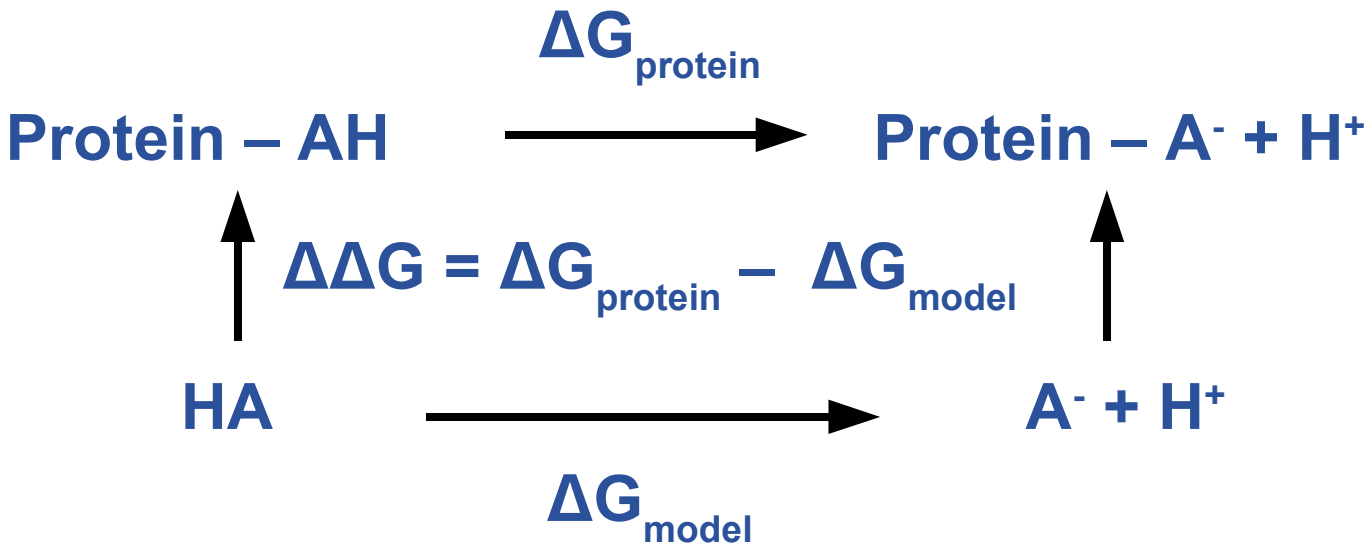
$$HA \rightarrow H^+ + A^-$$

$$K_a = \frac{[H^{+1}][A^{-1}]}{[HA]} \qquad pK_a = -\log K_a$$

$$\Delta G = -RT \ln K_a \qquad pK_a = \frac{\Delta G}{2.303 RT}$$

K_a .... equilibrium constant
G .... Gibbs free energy
R .... gas constant
T .... temperature

Thermodynamic cycle & pKa shift



$$pK_a = \frac{\Delta G}{2.303\,RT}$$

$$pK_{a,prot} = pK_{a,model} + \frac{\Delta\Delta G}{2.303\,RT}$$

pK_a shift

K_a .... equilibrium constant

G .... Gibbs free energy

R .... gas constant

T .... temperature

Thermodynamic integration (TI)

$$U(\lambda) = (1-\lambda)U_A + \lambda U_B$$

$$\Delta G = \sum_i w_i \langle \delta U / \delta \lambda \rangle$$

U_A .... potential energy of a state A

U_B .... potential energy of a state B

λ .... mixing (coupling) parameter

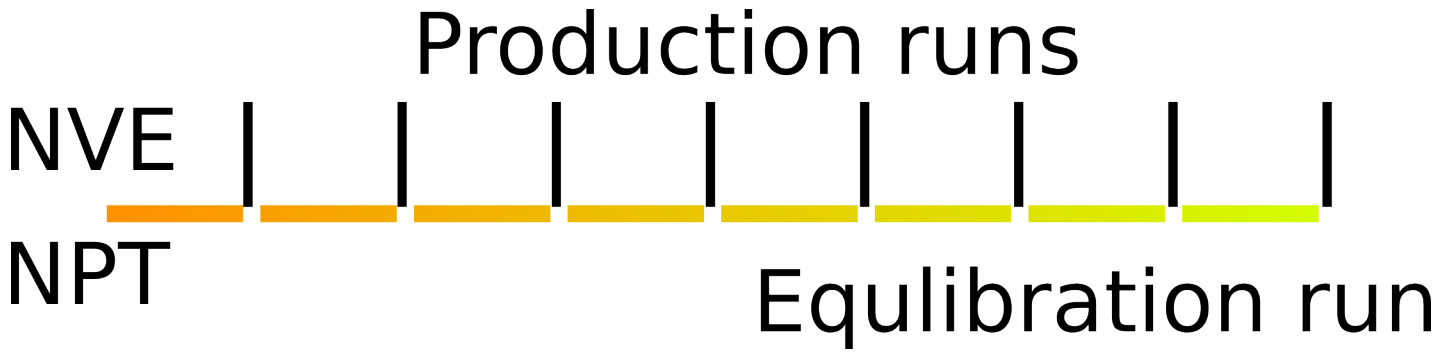
w_i .... weights in gaussian quadrature

- **MD is run at discrete points - "integration points" and δU/δλ is computed**
- **Accuracy TI of results depends on:**
  - *Convergence of δU/δλ*
  - *Distribution of δU/δλ*
  - *MD sampling, accuracy of force field*

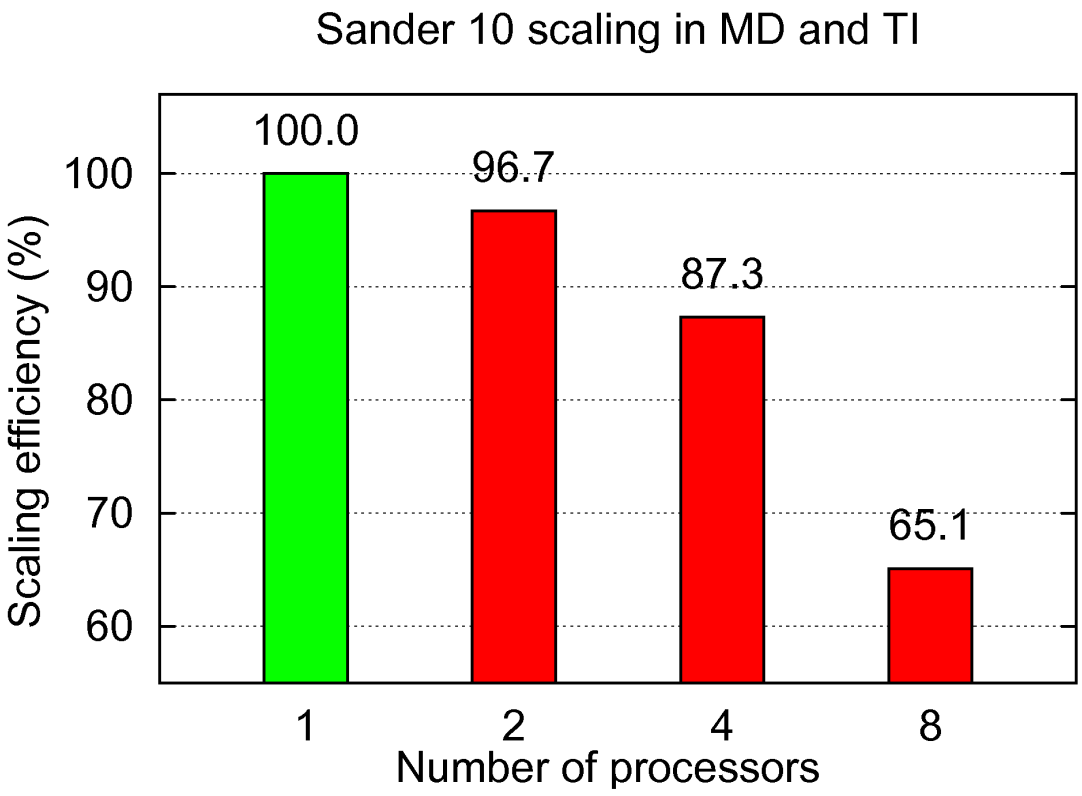
Methods: simulation set-up

- **used programs**
  - *Sander (Amber 10, patch 30) – performs TI*
    - MD - Amber ff99SB, PME, explicit water molecules
  - *MPI library (Open MPI 1.4.1) – message passing interface*
  - *MPI – start scripts – facilitate parallel execution*
    - *developed by Interactive European Grid project (I2G)*
  - *scripts for automated submission*

- **computational strategy**



Methods: gLite functions used



MD ← → **TI**

Each processor is located on a different node.

- **TI is a parallel calculation by design**
- **TI is 4x slower then the conventional MD**
- **Parallel jobs**
  - *JobType = "Normal";*
  - *CpuNumber = 2;*
- **Directed Acyclic Graph (DAG) jobs**
  - *linear – jobs are executed sequentially*

TI: automated submission details

- **equilibration run is submitted first – 8 CPUs**
- **productions runs are submitted together – 2 CPUs**
- **steps for each run:**
  - Sander input files ↑ storage element
    - *data file with LFN and GUID*
  - DAG file constructed
  - job submitted
  - job running – uploading Sander output files to SE
    - *updating data file with LFN and GUID*
  - job status check
    - *must be run by manually user because of proxy certificate on UI?*
  - job output files ↓ (data file)
  - check if the job really computed all part of the simulation
  - Sander output files ↓ → results
- **functions like SE ↑ ↓ , job submission and ↓ job output files retry in the case of an error**

Pitfalls & suggestions

MPICH job type	**X**	**Normal job type**
parallel		*single, parallel*
CEs with MPICH flag		*all CE*
should be deprecated		*parallel abilities should be documented*

- **normal job type → job can be submitted to sites without explicit support for parallel calculations**
  - *password-less SSH connection between WNs needed*
    - 6 CE out of 23 in VOCE
- **suggested GLUE schema attributes for:**
  - *password-less SSH connection between WNs*
  - *type and speed of network connection between WNs*
  - *type of filesystems at CE – shared filesystems*
- **JDL attribute to enable running parallel jobs only on processors with shared memory**

Results: pK_a shift in thioredoxin

Benchmark calculation, 18ns long TI

	Simonson, J. Am. Chem. Soc., 2004		Amber ff99SB, Asp26 "head conversion"	
	Ash-Asp model	Thioredoxin	Ash-Asp model	Thioredoxin
δU/δλ (λ = 0.11270)	-1.3	-3.1	-17.9	-5.6
δU/δλ (λ = 0.5)	-75.3	-64.5	-77.0	-73.1
δU/δλ (λ = 0.88279)	-148.6	-131.4	-140.7	-124.3
ΔG	-75.1	-66.0	-78.3	-68.6
ΔΔG		9.1		9.7
ΔΔG(exper.)		4.8		4.8

- **project is still in the benchmark phase**
- **planned pK_a calculations on acetylcholinesterase:**
  - 12 simulations for the Glu residues in active site – 10 ns each
  - simulations for other selected ionizable aminoacids
  - total about 300 ns of TI

Conclusions

WLCG/EGEE infrastructure was proven to be useful for computationally expensive TI calculations.

The existing grid environment is not fully ready for parallel jobs.

Some changes to the GLUE schema and JDL were suggested.

Acknowledgement

- Jaroslav Koča
- Petr Kulhánek

- WLCG/EGEE

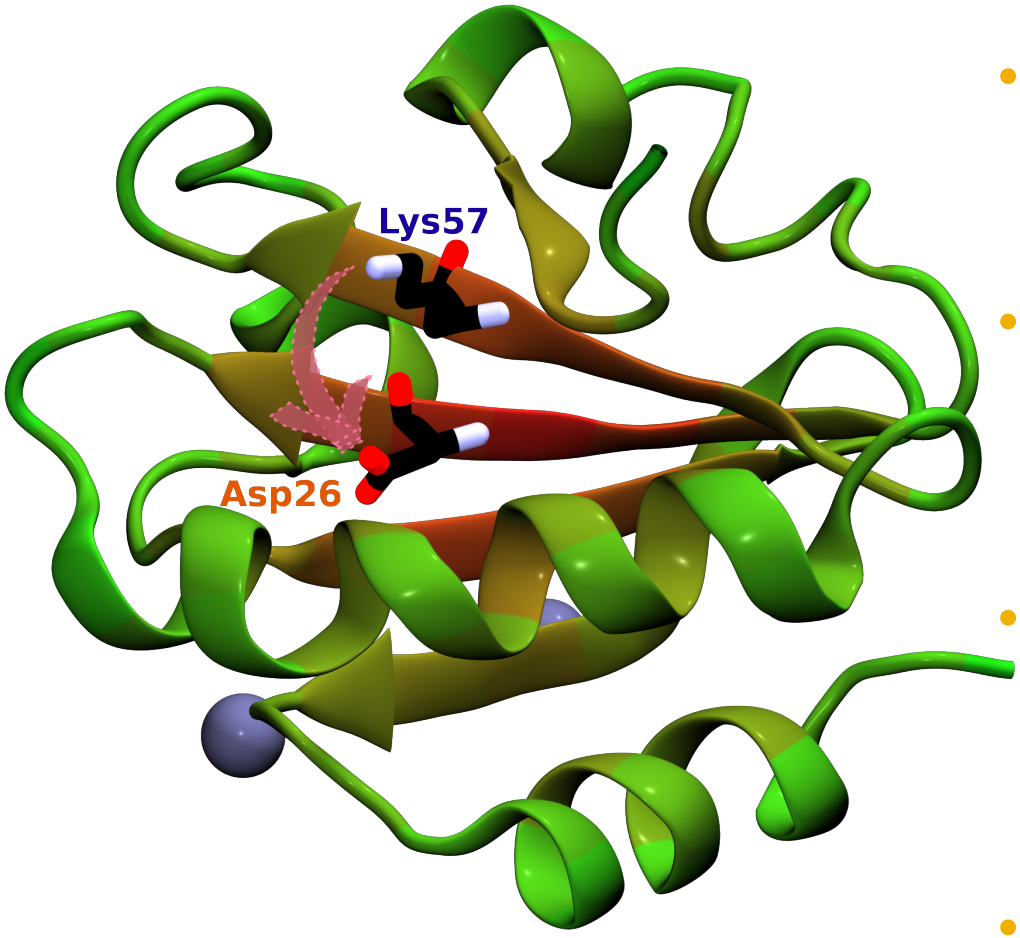
Acknowledgement

- Jaroslav Koča
- Petr Kulhánek

- WLCG/EGEE

Thank you for your attention!

Thioredoxin



- 108 residue protein
- Asp 26 – one of the largest upward pK_a shift among protein carboxylates
- Asp26 buried under Lys57
  - *Lys57 attracted to charged Asp26 according to Simonson et al.*
- Experimental pK_a is 7.5

  → *pK_a shift is 3.5 and $\Delta\Delta G$ (double free energy difference) is 4.8 kcal/mol*
- Computed $\Delta\Delta G$ is 9.1±4.1 kcal/mol