Two-Metal Ion Catalysis by Ribonuclease H

Edina Rosta

Department of Chemistry

King's College London

Phosphate Groups as Building Blocks

- **Biological importance**:
  - **Reproduction**:
    DNA and RNA hydrolysis, synthesis
    Regulation of gene expression
  - **Energy storage and transfer**:
    ADP/ATP equilibrium provides energy for reactions in the cells
  - **Signaling**:
    Phosphorylation activates or deactivates proteins in regulatory processes

- **Theoretical importance**:
  - **Reaction mechanism**
  - Highly charged species
  - Solvation effects
  - Charge transfer, polarization effects
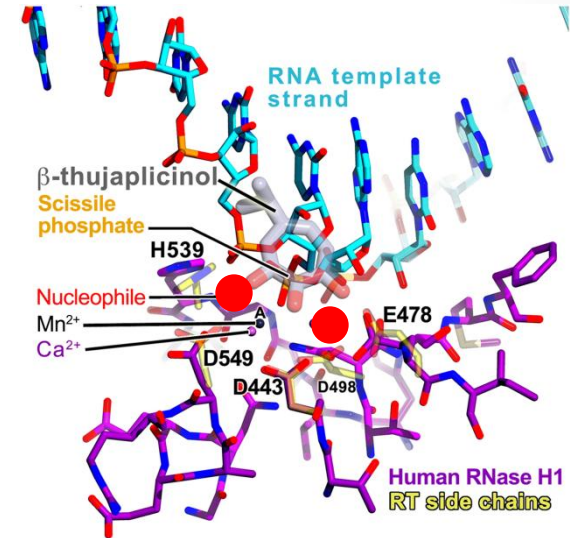 - Large entropy effects
  - **Metal ions in enzymes**
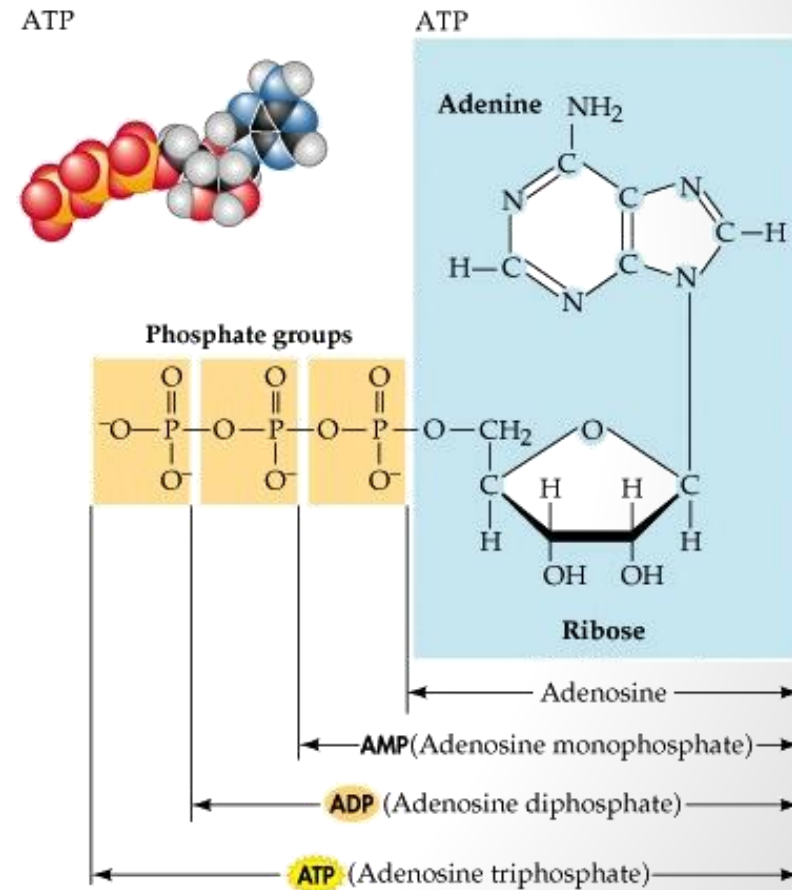
Two-Metal Ion Catalysis
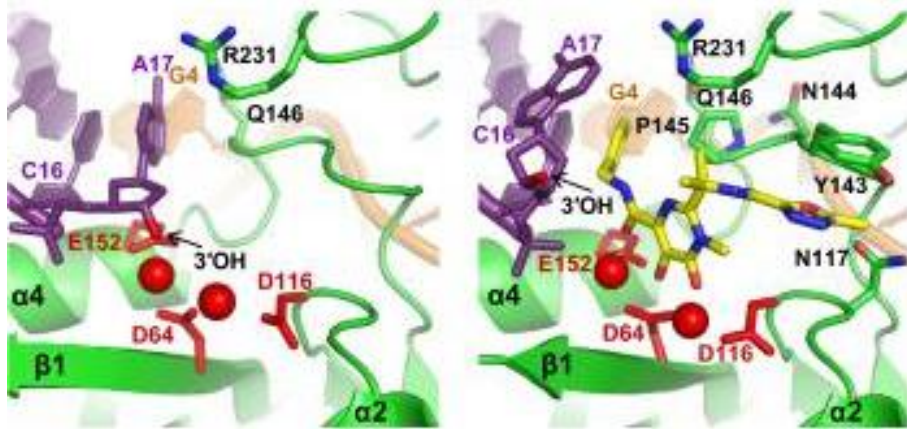
**IN active site + Raltegravir
(drug molecule for HIV-IN)**

Krishnan, et. al., *PNAS,* 2010
Hare, et. al, *Nature*, 2010

**HIV-RT RNase H active site +
inhibitor beta-thujaplicinol
superimposed with human RNase
H active site**

Nowotny, et al., *Cell,* 2007
Himmel, et al., *Cell,* 2009

- *Why are metal ions indispensable?*

- *Can we understand their catalytic roles based on quantum chemistry?*

RNase H Catalytic Reaction

- *B. halodurans* RNase H complexed with RNA/DNA duplex substrate

  Nowotny, *et. al, Cell,* 2005

- Crystal structure includes bound Mg^{2+}-ions in active site

- Enzyme catalyzes the non-specific cleavage of the RNA backbone phosphate ester bond in the RNA/DNA duplex via hydrolytic mechanism

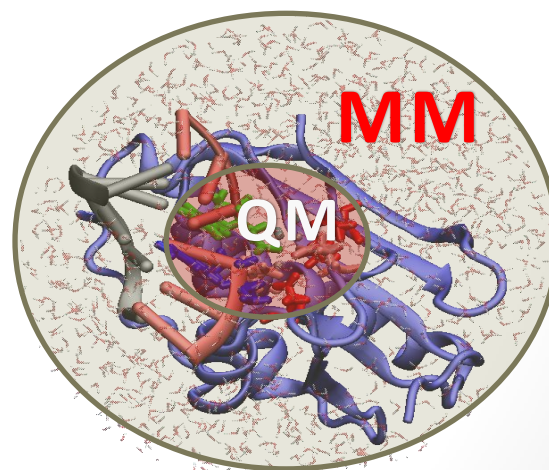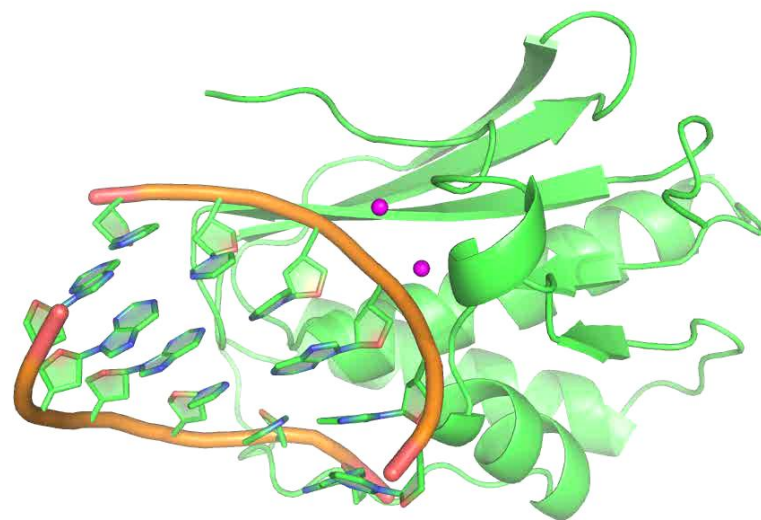- Same type of active site is major drug target in HIV-RT
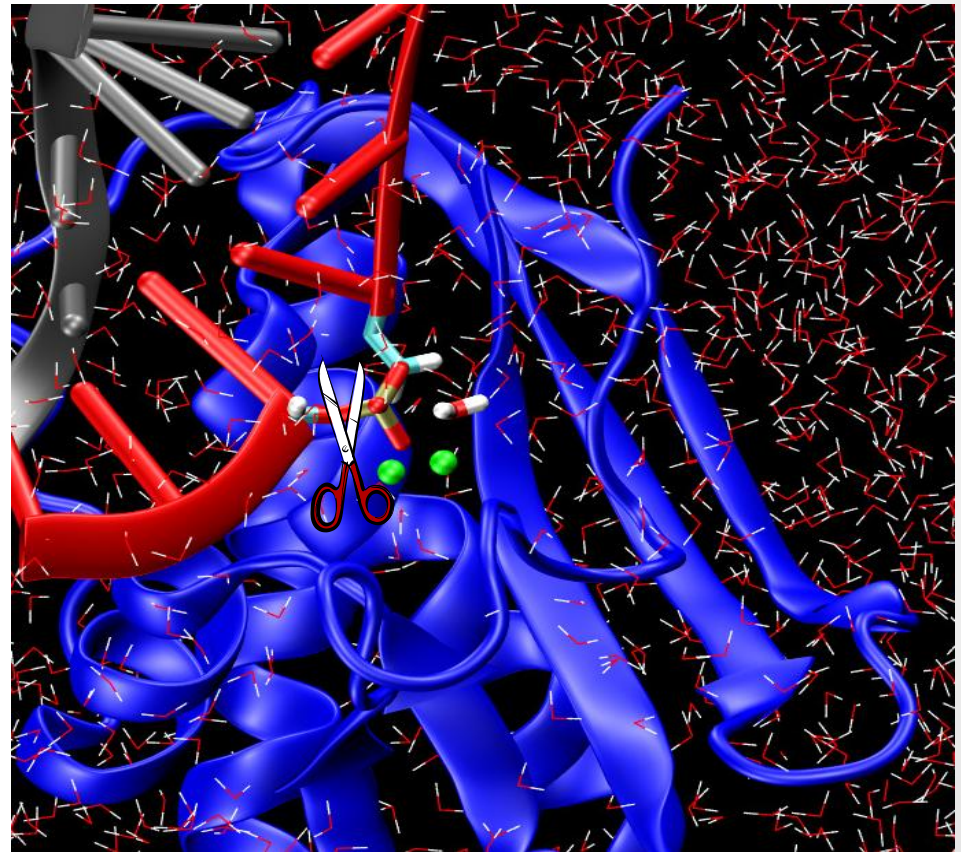
Computational Methods

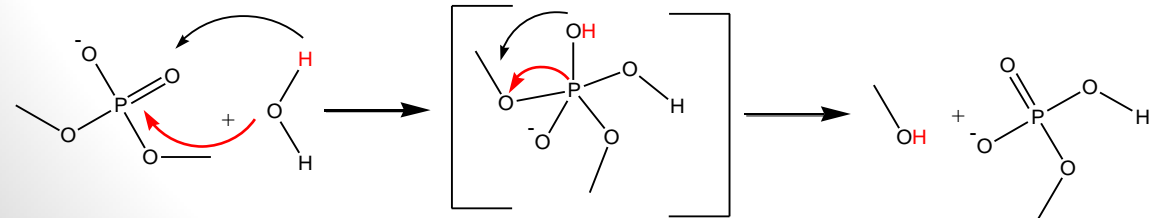- NEW QM/MM implementation with Q-Chem +CHARMM using full electrostatic embedding

  Woodcock et al., *J. Comp. Chem.*, 2007

- Phosphate-diester hydrolysis by attacking water

- DFT B3LYP method

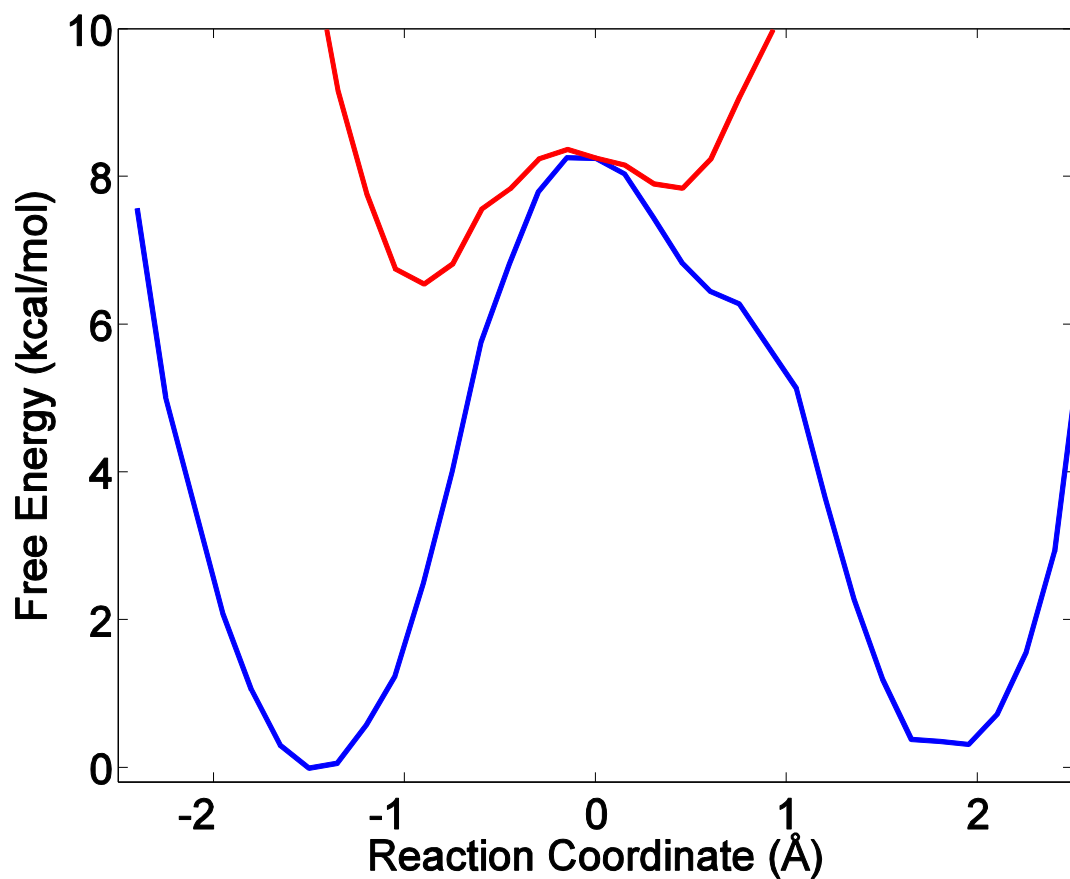- Free energy calculations using Umbrella Sampling

Shao, et al., *Phys. Chem. Chem. Phys.*, 2006

Umbrella Sampling

$$E_i\left(q_A\right) = U_{pot}\left(q_A\right) + \frac{1}{2}k_i\left(\xi_A - \xi_i\right)^2$$

- Run parallel simulations with harmonic constraints moving along the reaction coordinate

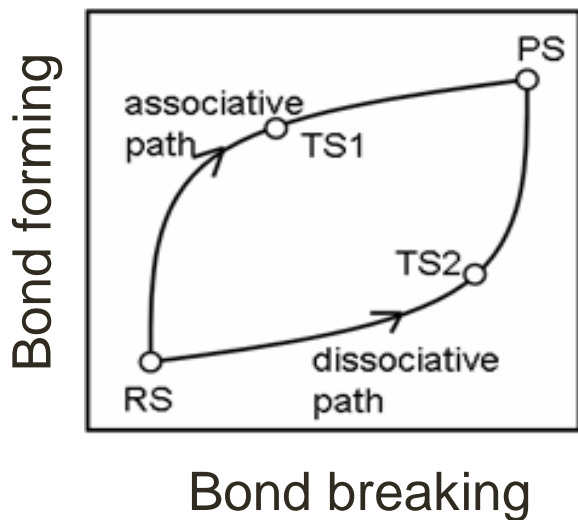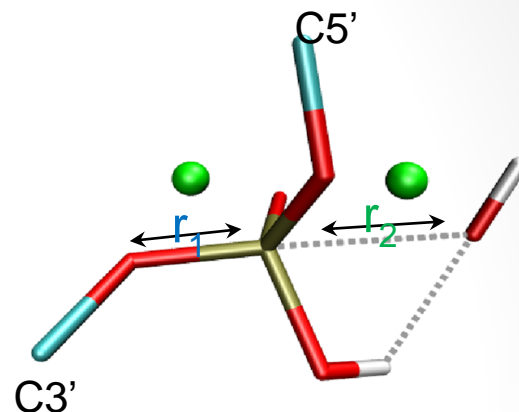- Recover the unbiased free energy surface from combined data using e.g., WHAM
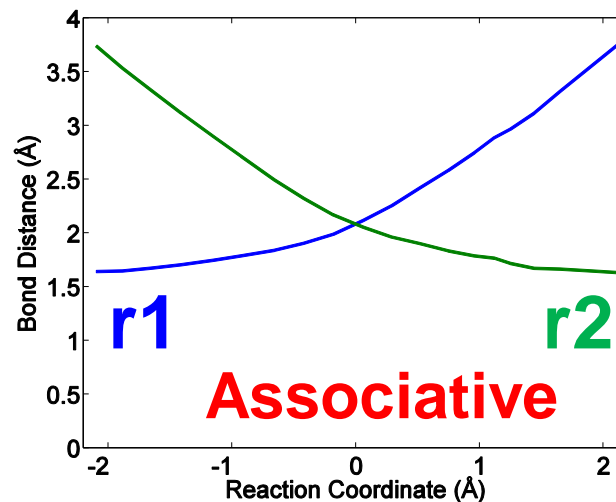
Reaction Coordinate

- 1D coordinate:

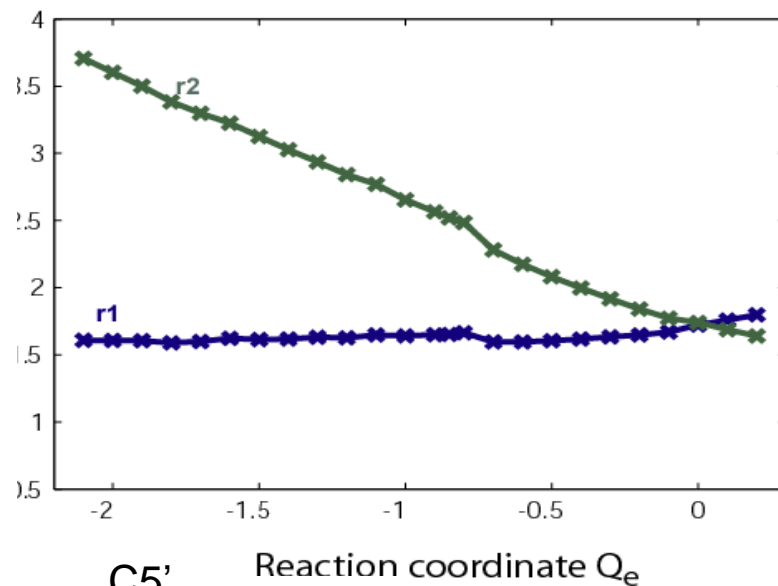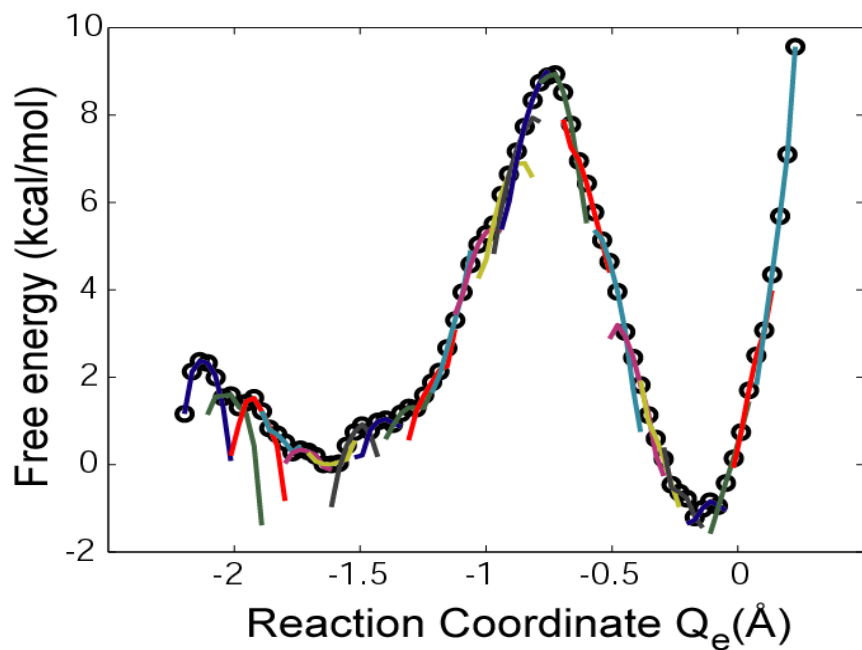 Q_e = Bond breaking (r1) – Bond forming (r2)

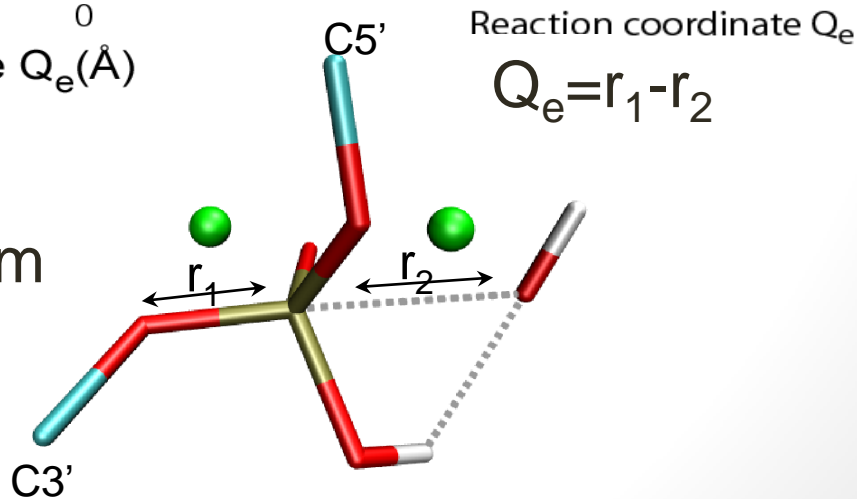- Allows to distinguish between associative/dissociative mechanisms

More O'Ferrall-Jencks diagram

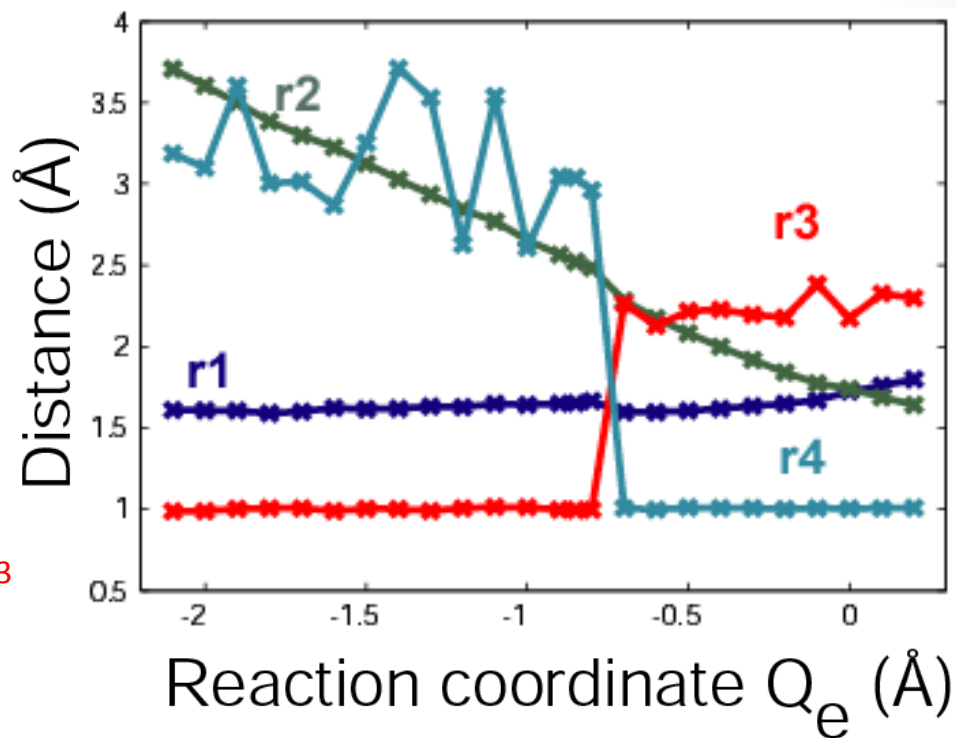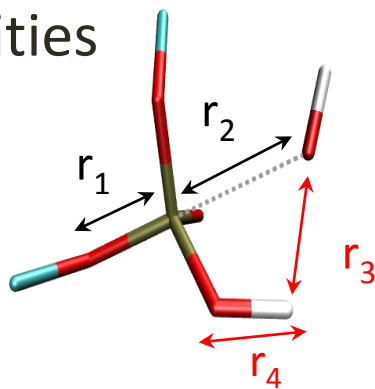# Umbrella Sampling along Q_e

$Q_e = r_1 - r_2$

- Associative mechanism

Automated search for discontinuities in atomic distances

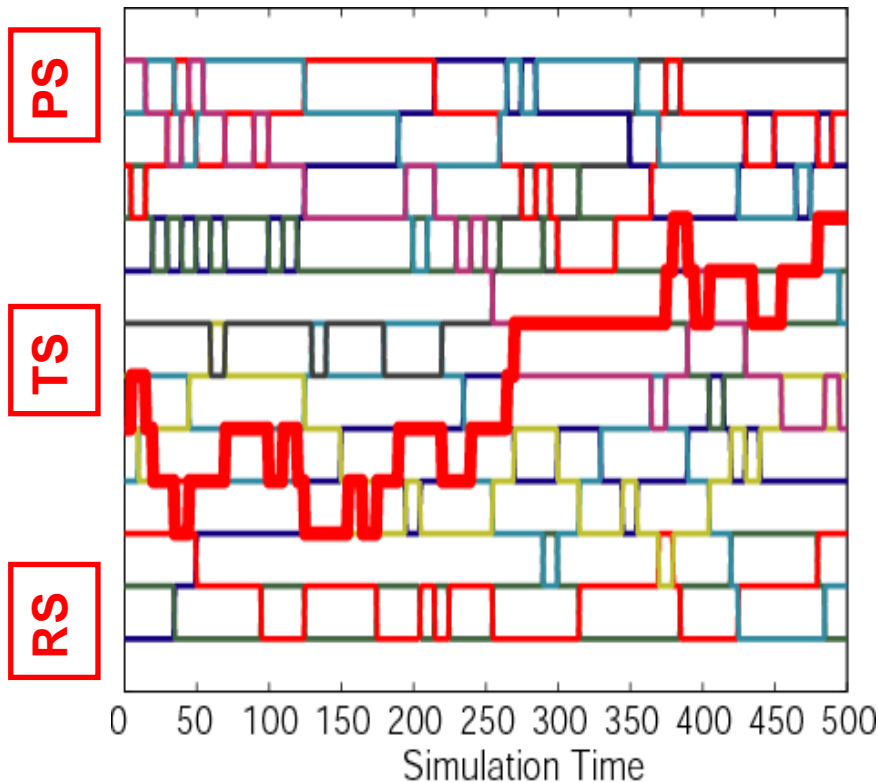Analyze atomic distances along umbrella sampling windows for discontinuities

Electron transfer: $Q_e = r_1 - r_2$
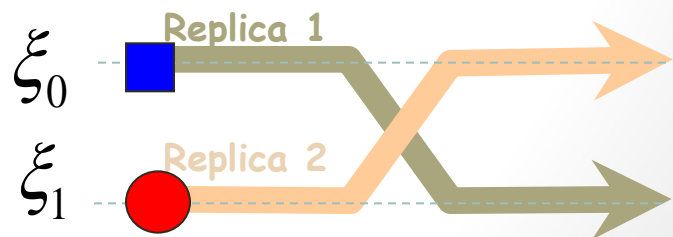
Proton transfer: $Q_p = r_3 - r_4$

Hamiltonian Replica Exchange

Why replica exchange? Can we optimize our protocol?


- Running MD at different temperatures in parallel
- Couple the runs in order to speed up lowest temperature's dynamics
- Preserve P_{eq} at each temperature
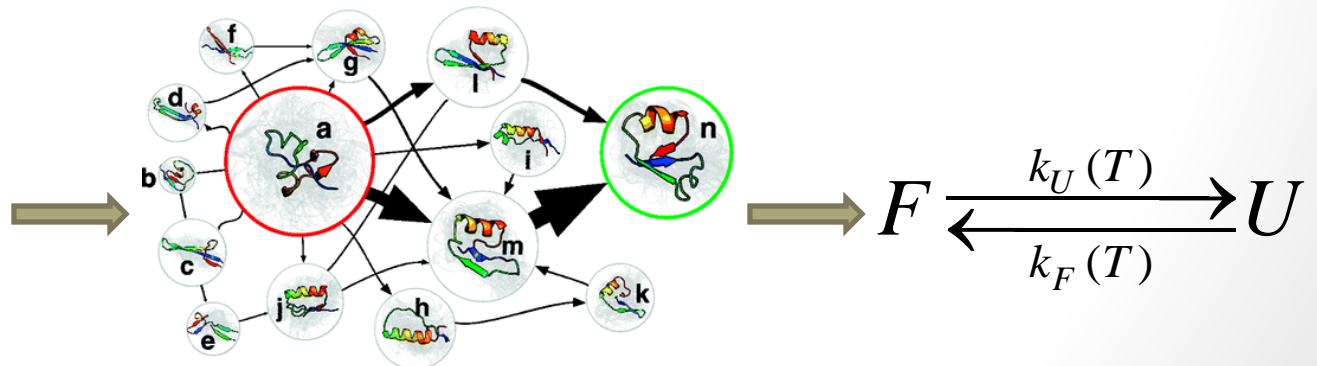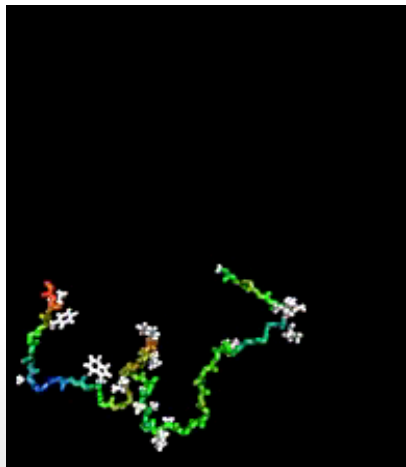- Detailed balance condition has to be satisfied

K. Hukushima and K. Nemotto, *J. Phys. Soc. Japan*, 1996
Fukunishi, Watanabe & Takada, *J. Chem. Phys.*, 2002

State Assignment in Protein Folding Simulations

$$\frac{dP_i}{dt}(t) = \sum_{\substack{j=1 \\ (j \neq i)}}^{N} k_{i \leftarrow j} P_j(t) - \sum_{\substack{j=1 \\ (j \neq i)}}^{N} k_{j \leftarrow i} P_i(t)$$
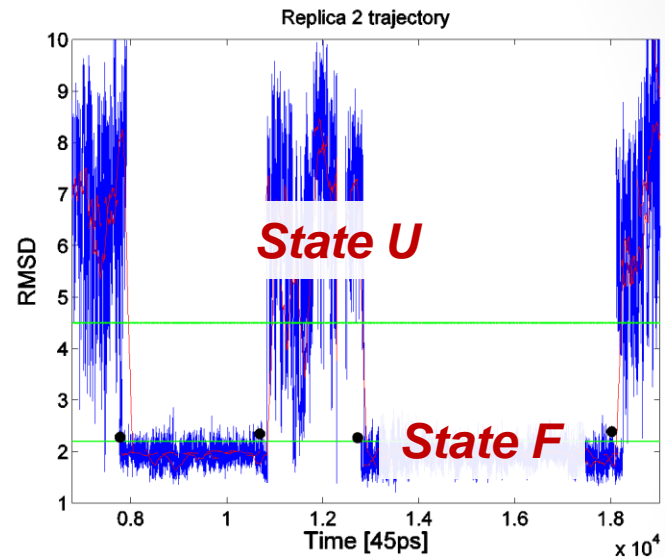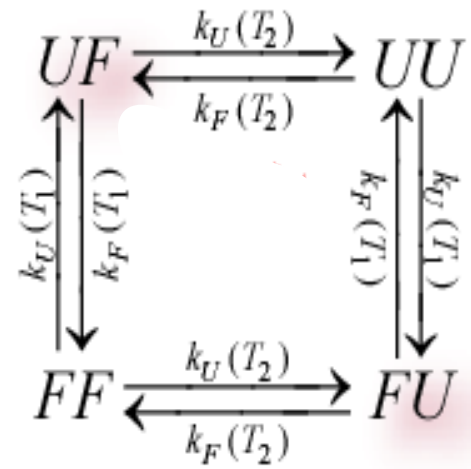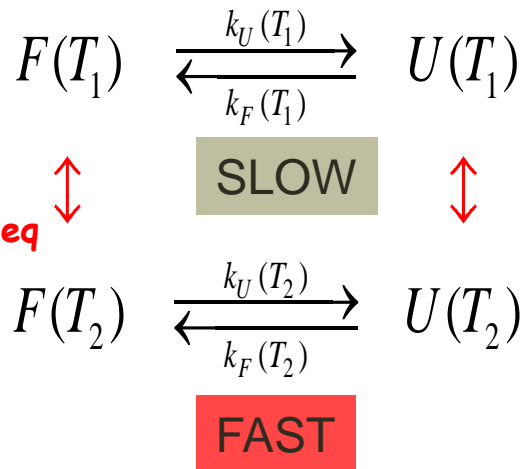
Folding@home, V. Pande, Stanford

Replica 2 trajectory

State U

State F

$$F \underset{k_F(T)}{\overset{k_U(T)}{\rightleftharpoons}} U$$
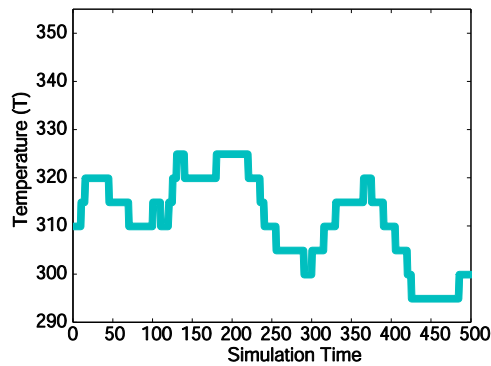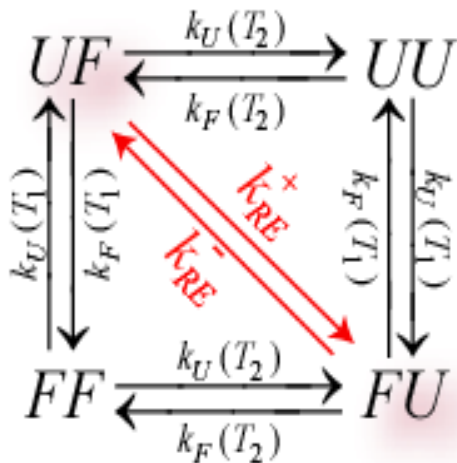
Two-State Kinetic Model of Replica Exchange and Simulated Tempering

W. Zheng, M. Andrec, E. Gallicchio and R. Levy, PNAS, 2007

$$F(T_1) \underset{k_F(T_1)}{\overset{k_U(T_1)}{\rightleftarrows}} U(T_1)$$

SLOW

preserves p_{eq}

$$F(T_2) \underset{k_F(T_2)}{\overset{k_U(T_2)}{\rightleftarrows}} U(T_2)$$
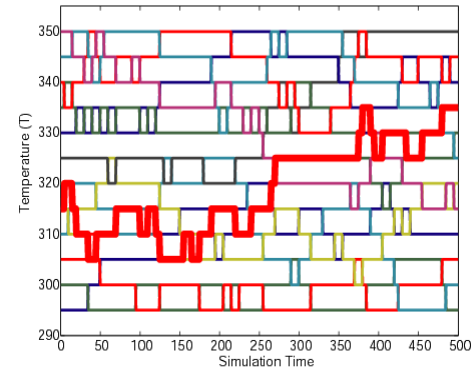
FAST

Replica Exchange Rate

$$k_{RE}\left(U_iF_j \to U_jF_i\right) = \frac{p_{acc}\left(U_iF_j \to U_jF_i\right)}{\delta t_{xc}}$$
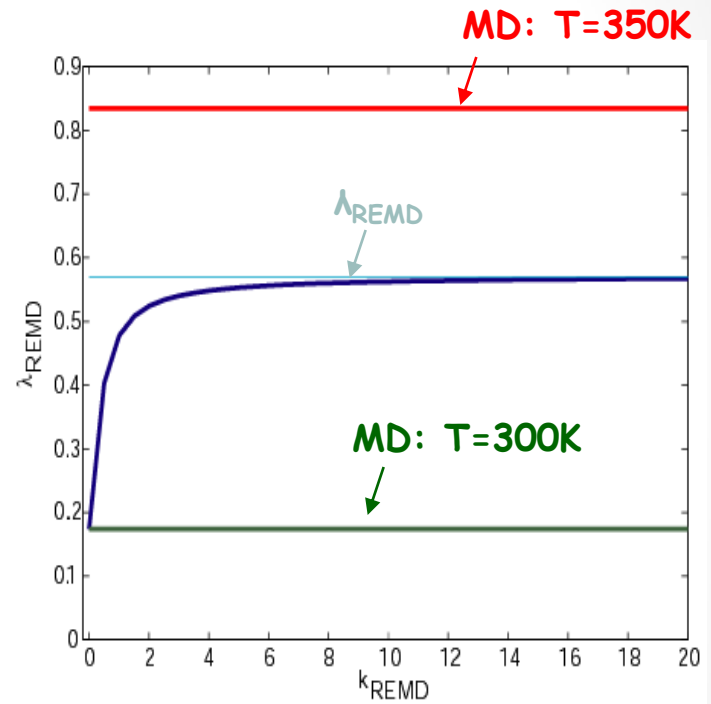
$$K = \begin{bmatrix} -\left(k_U(T_1)+k_U(T_2)\right) & k_F(T_1) & k_F(T_2) & 0 \\ k_U(T_1) & -\left(k_F(T_1)+k_U(T_2)+k_{RE}^{+}\right) & k_{RE}^{-} & k_F(T_2) \\ k_U(T_2) & k_{RE}^{+} & -\left(k_U(T_1)+k_F(T_2)+k_{RE}^{-}\right) & k_F(T_1) \\ 0 & k_U(T_2) & k_U(T_1) & -\left(k_F(T_2)+k_F(T_1)\right) \end{bmatrix}$$

Coarse Graining Replica Exchange Coupled States: Fast Exchange Limit

$$UF \xrightleftharpoons[k_F(T_2)]{k_U(T_2)} UU$$

$k_U(T_1)$ $k_F(T_1)$ k_{RE}^+ k_{RE}^- $k_F(T_1)$ $k_U(T_1)$

$$FF \xrightleftharpoons[k_F(T_2)]{k_U(T_2)} FU$$

(N=2)

(N=3)
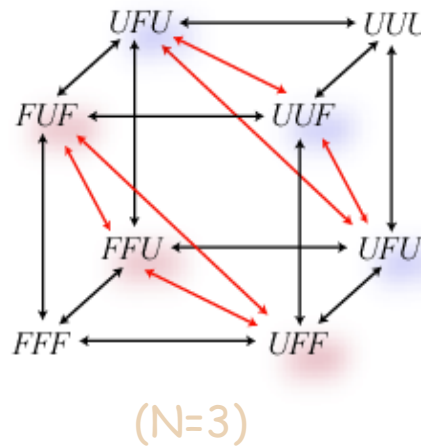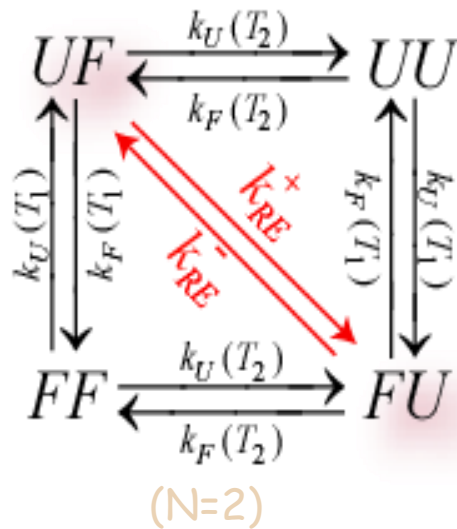
$\cdots$ *N* temperatures
(2^N microstates)

$$FF \rightleftharpoons \begin{pmatrix} UF \\ FU \end{pmatrix} \rightleftharpoons UU \qquad FFF \rightleftharpoons \begin{pmatrix} FFU \\ FUF \\ UFF \end{pmatrix} \rightleftharpoons \begin{pmatrix} FUU \\ UFU \\ UUF \end{pmatrix} \rightleftharpoons UUU \qquad 0 \rightleftharpoons 1 \rightleftharpoons \cdots \rightleftharpoons N$$
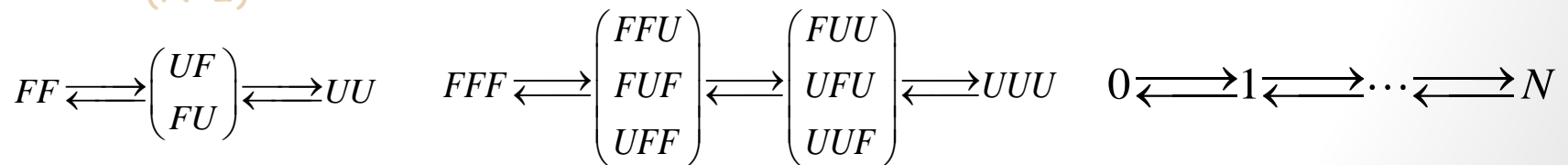
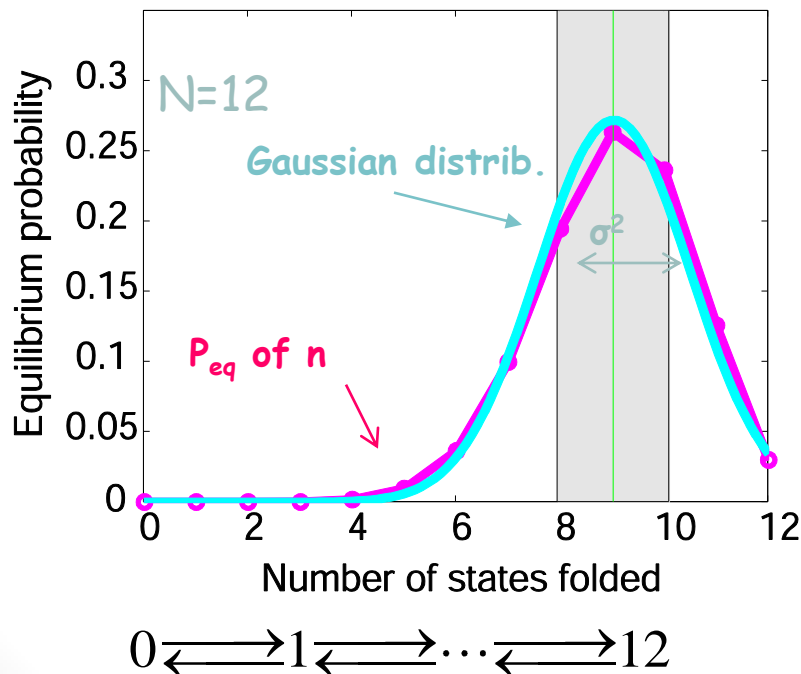# Kinetic Theory: Continuum Limit

- Smoluchowski equation for diffusion in a one-dimensional harmonic potential

- Analytic solution for the slowest relaxation rate of the system:

$$\lambda = \frac{D}{\sigma^2} = \frac{K_{n_{\max}-1,n_{\max}}}{\sigma^2}$$

$$\lambda_{REMD} = \frac{\sum_{i=1}^{N} \lambda(T_i)\, p_F(T_i)\, p_U(T_i)}{\sum_{i=1}^{N} p_F(T_i)\, p_U(T_i)}$$
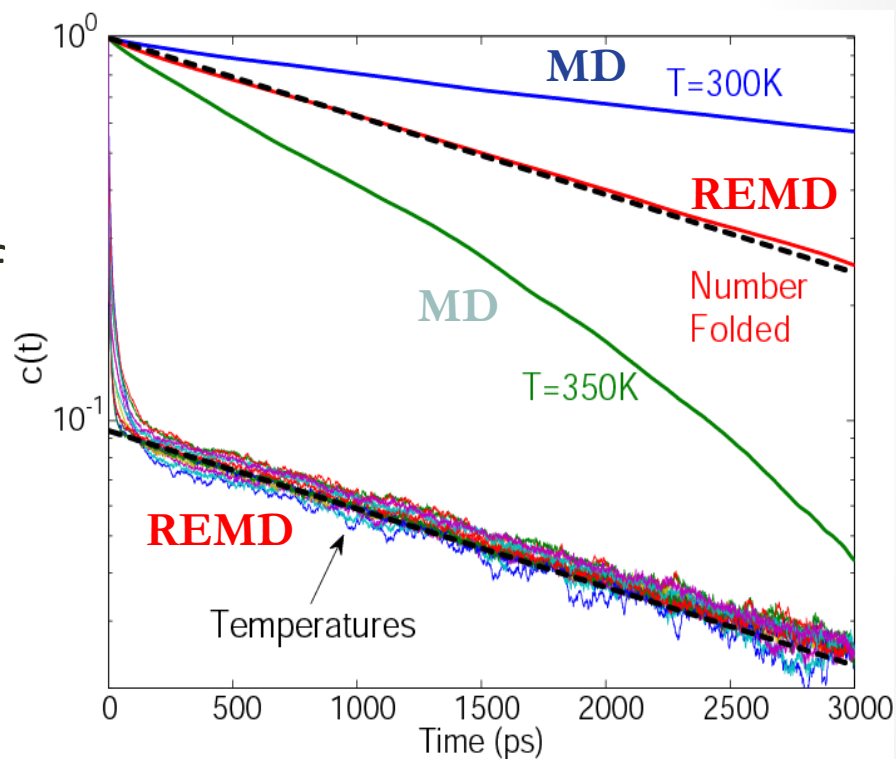
Exact for $N \rightarrow \infty$ and $k_{RE} \rightarrow \infty$.

Folding/Unfolding of Ala$_5$

- All-atom simulations of Ala$_5$ in explicit water

- State correlation function of all temperatures:

$$c(t) = \frac{\langle s(t)s(0)\rangle_T - \langle s(t)\rangle_T^2}{\langle s(t)^2\rangle_T - \langle s(t)\rangle_T^2}$$
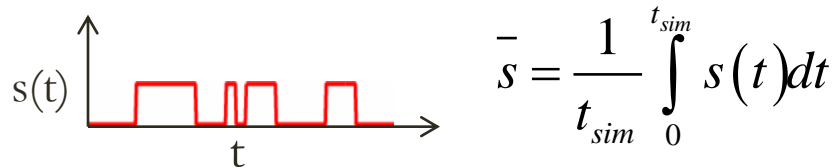

- Fit for λ matches the prediction perfectly using the corresponding folding/unfolding rates!

$$\lambda_{REMD} = \frac{\sum\limits_{i=1}^{N} \lambda(T_i)\, p_F(T_i)\, p_U(T_i)}{\sum\limits_{i=1}^{N} p_F(T_i)\, p_U(T_i)}$$

Efficiency of Replica Exchange

- Estimating the mean of the folding state function *s(t)*:

$$\bar{s} = \frac{1}{t_{sim}} \int_0^{t_{sim}} s(t)\,dt$$

- Efficiency = relative error compared to standard MD simulations using the same computational resources:

$$\eta = \frac{\sigma_{MD}^2\left(Nt_{sim}\right)}{\sigma_{REMD}^2\left(t_{sim}\right)} = \frac{\sigma_{MD}^2\left(t_{sim}\right)}{N\sigma_{REMD}^2\left(t_{sim}\right)}$$

- Estimate of the error of a general property Y:

$$Y\left(t\right) \cong \left(\langle Y\rangle_F - \langle Y\rangle_U\right)s\left(t\right) + \langle Y\rangle_U \Rightarrow \mathrm{var}\left(\bar{Y}\right) \cong \left(\langle Y\rangle_F - \langle Y\right.$$

- Error in estimating the folding probability, $\bar{s}$:

$$\sigma^2\left(t_{sim}\right) = \mathrm{var}\left(\bar{s}\right) = \frac{2}{t_{sim}^2} \int_0^{t_{sim}} \left(t_{sim} - t\right) c_s\left(t\right) dt$$
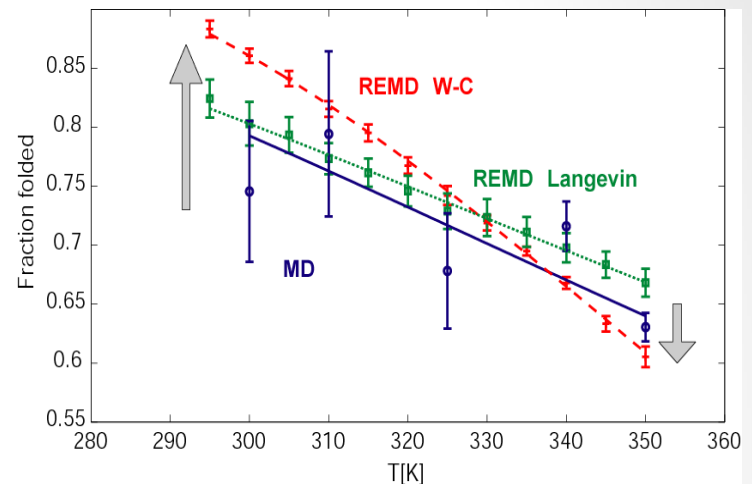
E. Rosta et al. 2009; JCTC

$$c_s\left(t\right) \cong A\exp\left(-\lambda_{REMD} t\right)$$
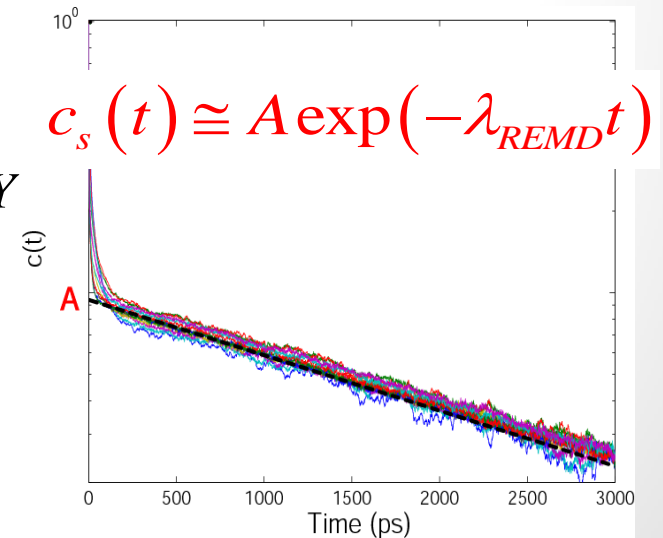
Efficiency of Replica Exchange

$$\eta = \frac{\sigma_{MD}^2\left(Nt_{sim}\right)}{\sigma_{REMD}^2\left(t_{sim}\right)} = \frac{\sigma_{MD}^2\left(t_{sim}\right)}{N\sigma_{REMD}^2\left(t_{sim}\right)}$$

$$\sigma_{MD}^2\left(t_{sim}\right) = \frac{2}{t_{sim}}\frac{p_{F_1}p_{U_1}}{\lambda_1}$$

$$\sigma_{REMD}^2\left(t_{sim}\right) = \frac{2}{t_{sim}}\frac{p_{F_1}^2 p_{U_1}^2}{\displaystyle\sum_{i=1}^{N} p_{F_i}k_{U_i}}$$

$$\left.\right\} \quad \eta_1 = \frac{\displaystyle\sum_{i=1}^{N} p_{F_i}k_{U_i}}{N p_{F_1}k_{U_1}}$$

Ala5 simulations

Legend:
- Simulation
- Fast limit
- Continuum limit
- N MD

Average number of transitions over all temperatures

Number of transitions at target temperature (T_1)

Results I: Kinetic Modeling

- Analytical efficiency expression is derived for *replica exchange & simulated tempering* simulations in the limit of fast exchange.

- Efficiency for slow exchange is obtained by numerical solution of the full kinetic rate matrix problem.

- Replica exchange & simulated tempering simulations have identical efficiencies in the limit of fast exchange.

- Model provides guidance for optimal simulation protocol to minimize the statistical error (*e.g.*, T-spacing, exchange frequency, *etc*.).

Rosta and Hummer, *J.Chem. Phys.*, 2009
Rosta and Hummer, *J.Chem. Phys.*, 2010

Computational Methods

- NEW QM/MM implementation with Q-Chem +CHARMM using full electrostatic embedding

  Woodcock et al., *J. Comp. Chem.*, 2007

- Phosphate-diester hydrolysis by attacking water

- DFT B3LYP method

- Free energy calculations using Umbrella Sampling

- COUPLED WITH HAMILTONIAN REPLICA EXCHANGE
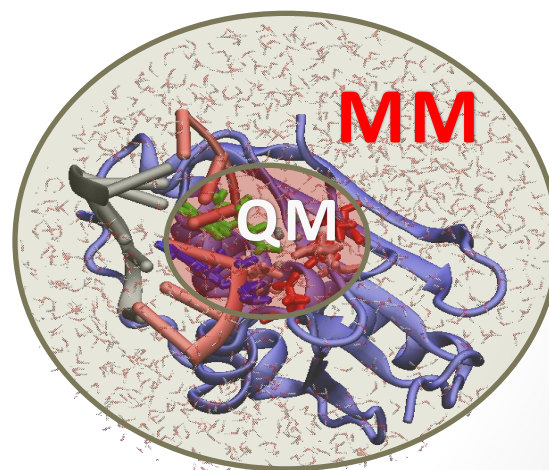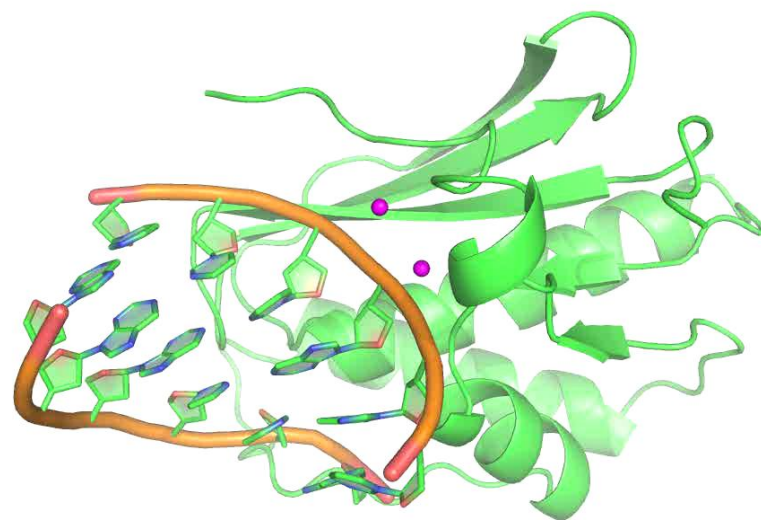
MM

QM

Q·CHEM

Shao, et al., *Phys. Chem. Chem. Phys.*, 2006

Hamiltonian replica exchange

- Smoother curves – better sampling

- Does not add extra cost to the simulations

- Discontinuity problem is not solved – could not help overcome proton transfer barrier

US only

H-REX

Rosta, Woodcock, Brooks, Hummer*, J. Comp. Chem.,* 2009
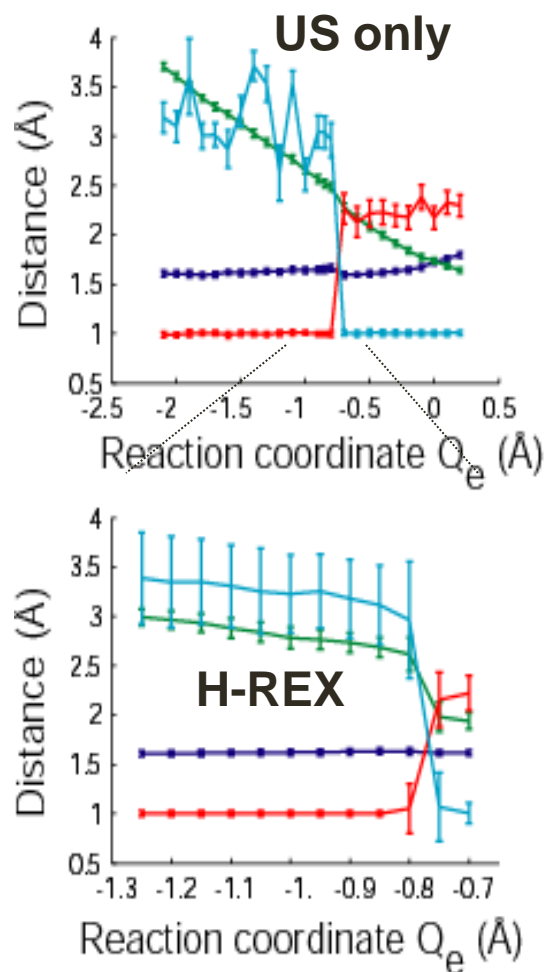
Hysteresis in low dimensional reaction coordinates
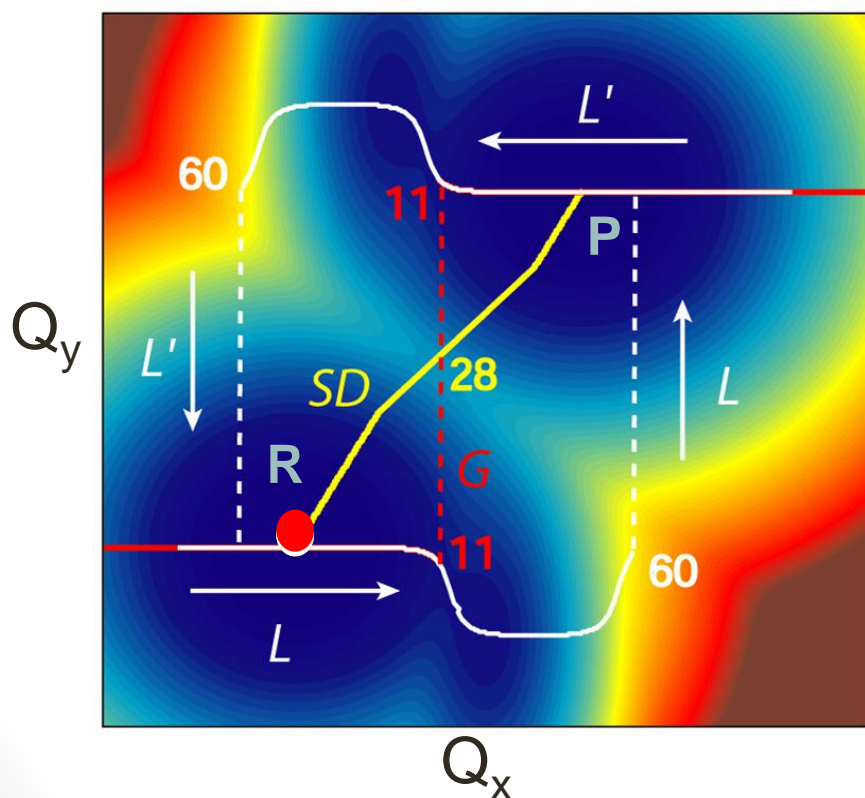
Energy minimizations along Q_x:

(Fix value along Q_x, minimize along all other coordinates.)

- **Steepest Descent (SD):**
  - Passes through transition state
  - Computationally not feasible
- **Local search (L):**
  - Overestimates barrier
  - Shows hysteresis
- **Global search (G):**
  - Underestimates barrier
  - Characterized by discontinuity

Rosta, Woodcock, Brooks, Hummer, *J. Comp. Chem.,* 2009

Hysteresis in low dimensional reaction coordinates
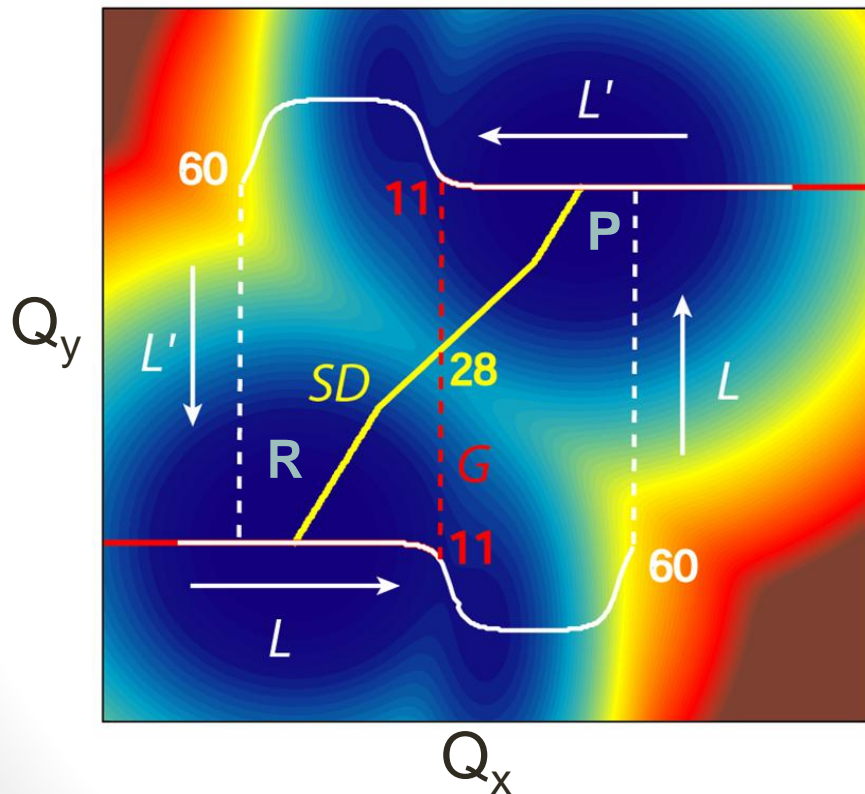
Free energy calculations:

Perfect sampling
+ wrong reaction coordinate
= underestimate the barrier of
the 1D (PMF) free energy
profile -- Characterized by
discontinuity

$$e^{-G(Q_x)/k_B T} =$$

$$\int dQ_y \, e^{-G(Q_x, Q_y)/k_B T}$$

Hamiltonian replica exchange

- Smoother curves – better sampling
- Does not add extra cost to the simulations
- Discontinuity problem is not solved – could not help overcome proton transfer barrier

Electron transfer:
$$Q_e = r_1 - r_2$$

Proton transfer:
$$Q_p = r_3 - r_4$$
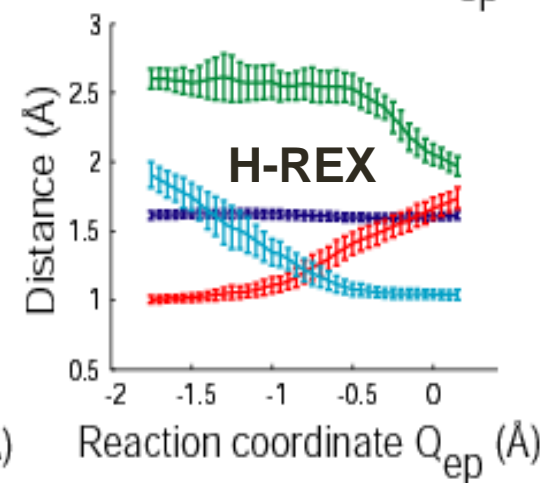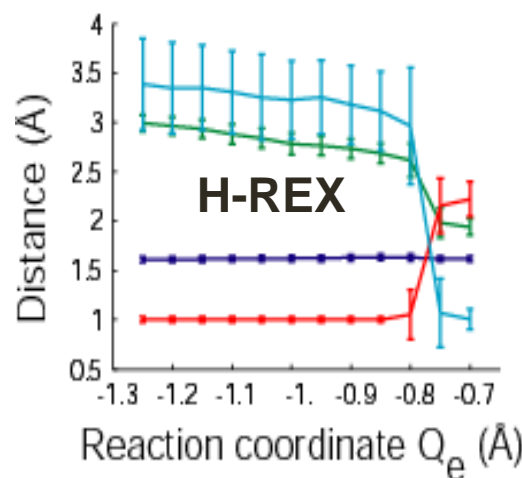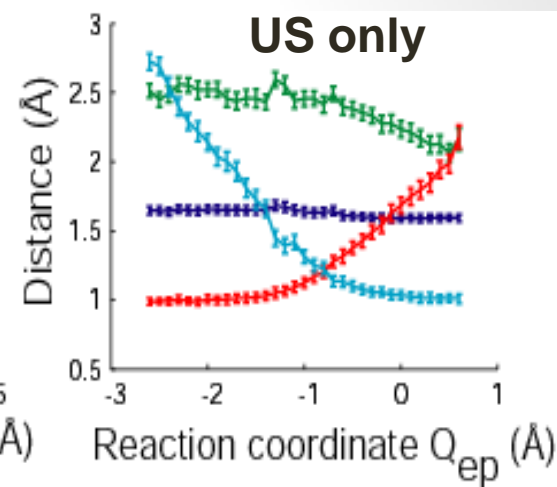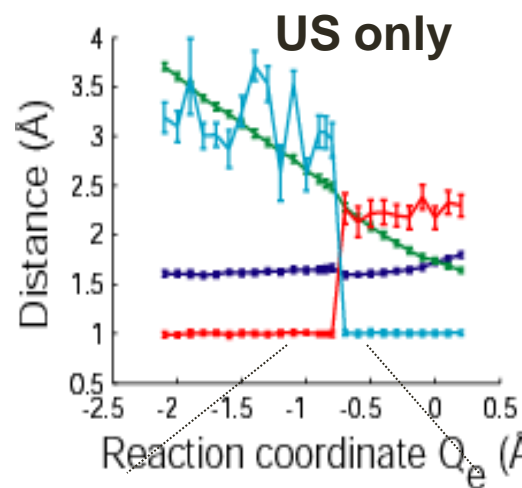
New coordinate for additional 1D umbrella sampling: $Q_{ep} = Q_e + Q_p$

2D-WHAM with proton transfer + electron transfer coordinate, Q_{ep}
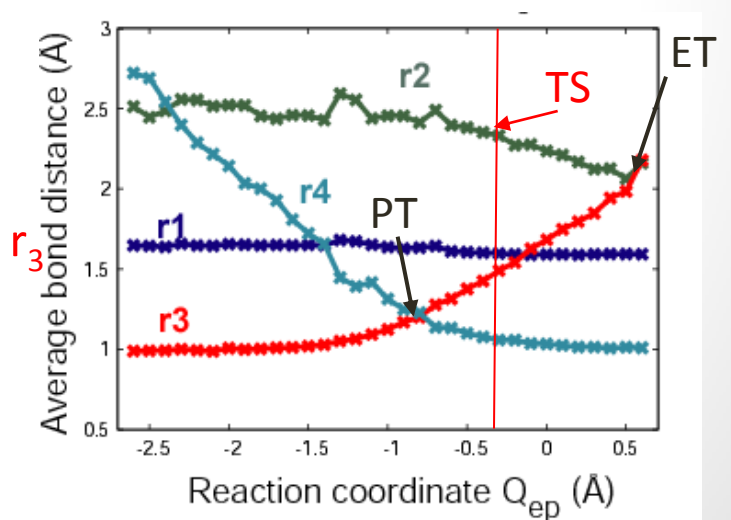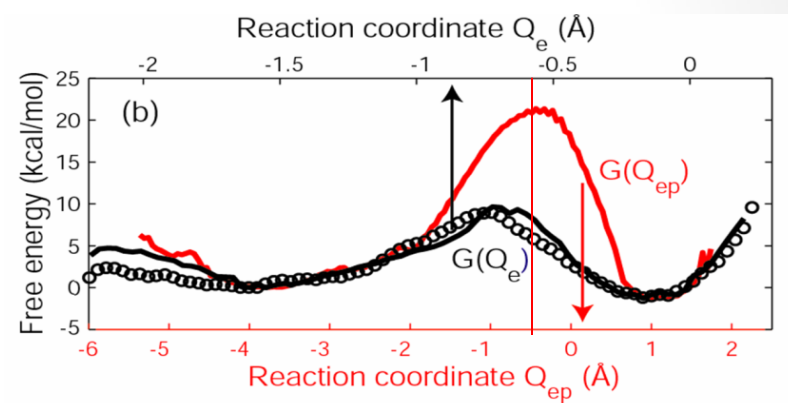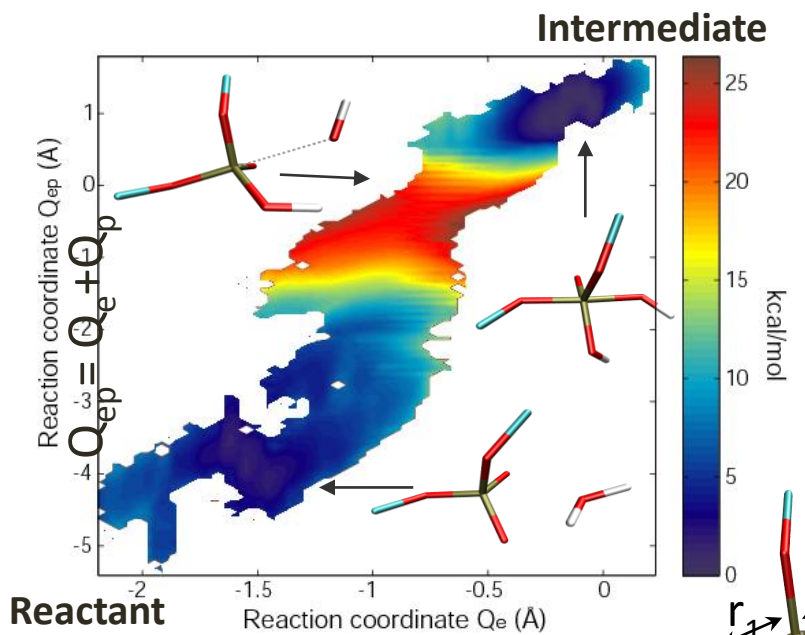
- High barrier to reach product state.
- Need to increase the QM region to account for additional pathways.

First proton transfer, then ET (bond breaking/forming at P)

Computational Methods

- NEW QM/MM implementation with Q-Chem +CHARMM using full electrostatic embedding

  Woodcock et al., *J. Comp. Chem.*, 2007

- Phosphate-diester hydrolysis by attacking water: QM region with 91 atoms

- DFT B3LYP method (6-31+G* basis)

- Free energy calculations of the reaction with enhanced sampling methods: Hamiltonian replica exchange coupled with finite temperature string method

MM

QM

Q-CHEM

Shao, Rosta, et al., *Phys. Chem. Chem. Phys.*, 2006
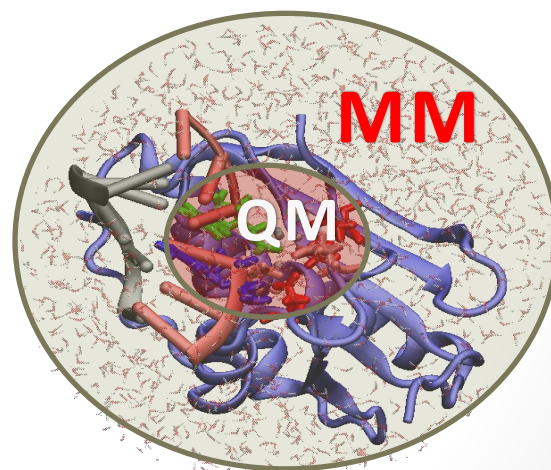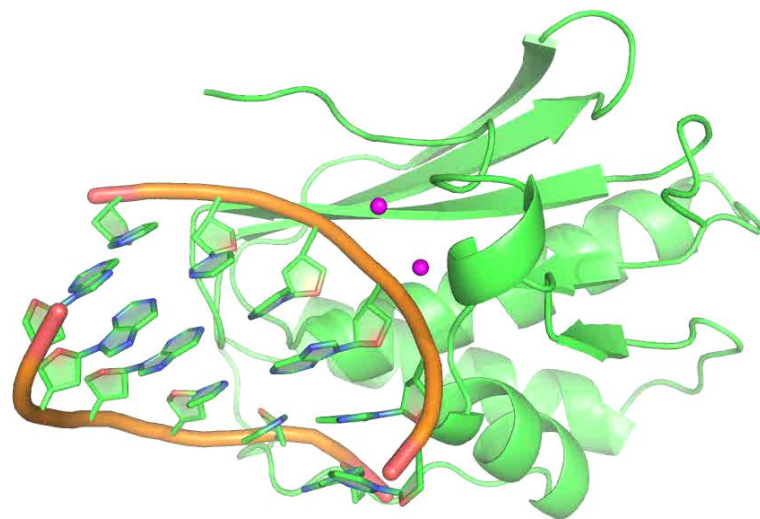
Free Energy Simulations Using the String Method

- Optimized a 1D string in the multidimensional space of the internal reaction coordinates to obtain minimum free energy path

  E, Ren, Vanden-Eijnden, *Phys. Rev. B,* 2002

- Hamiltonian replica exchange between string images
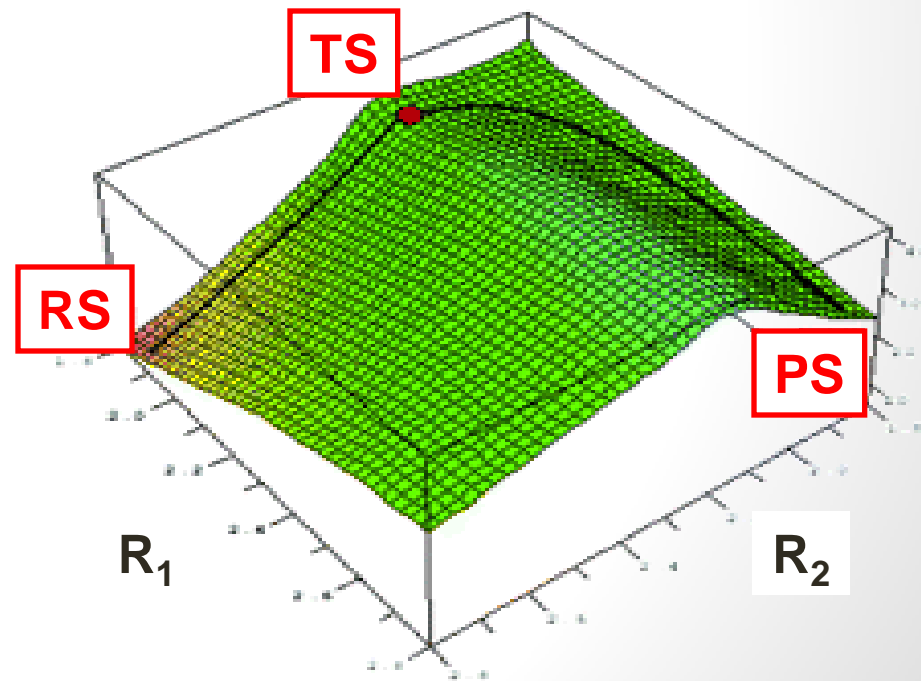
Free Energy Simulations Using the String Method

- Start with a guess for the string
- Run Umbrella Sampling simulations
- Determine forces for the images along the string
- Fit new string
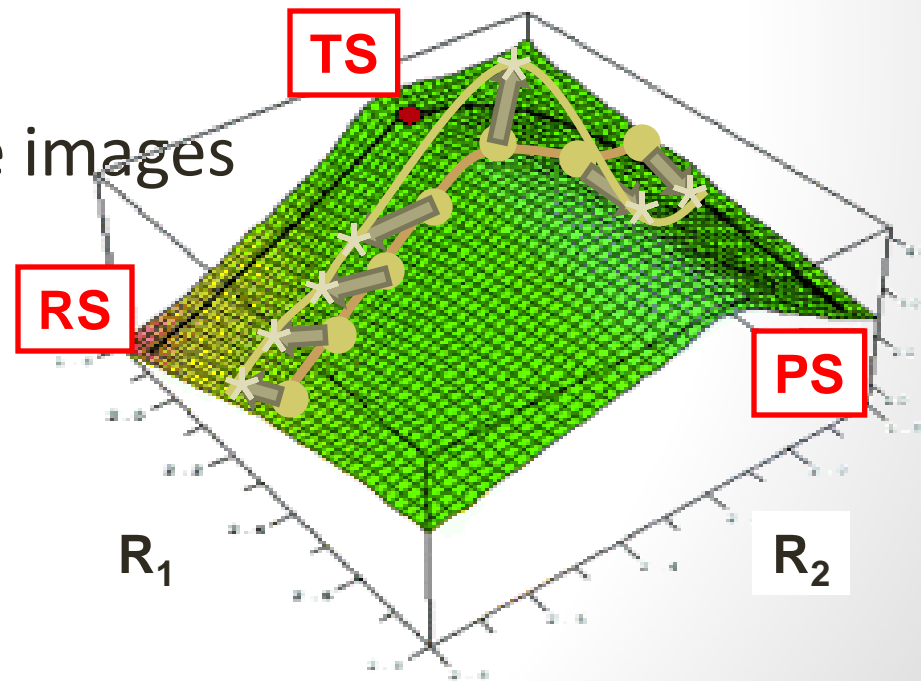
Free Energy Simulations Using the String Method

- Start with a guess for the string
- Run Umbrella Sampling simulations
- Determine forces acting on the images along the string
- Fit new string
- Redistribute images
- Run next iteration

TS

RS

PS

R_1

R_2
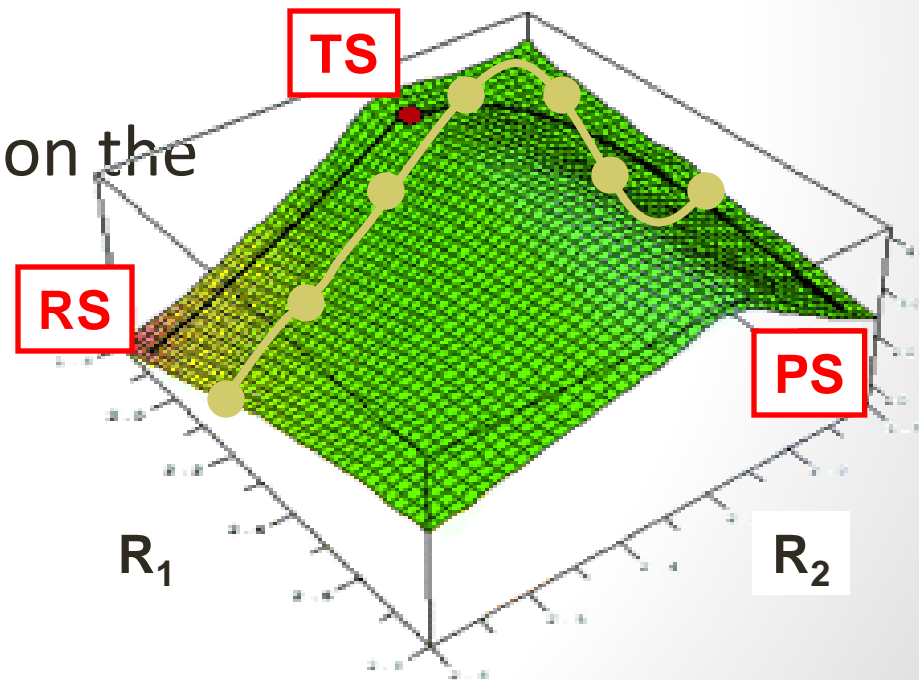
Free Energy Simulations Using the String Method

- Converged string:
  - Forces are parallel to string

- We use all data from all string simulations with **Histogram Free** implementation of WHAM (MBAR): works with **very high dimensionality**

$$f_m^{-1} = \sum_{k=1}^{NSim} \sum_{A \in NSim} \frac{c^m\left(\xi_A^k\right)}{\sum_{l=1}^{NSim} N_l f_l c^l\left(\xi_A^k\right)}$$

RS

TS

PS

R_1

R_2
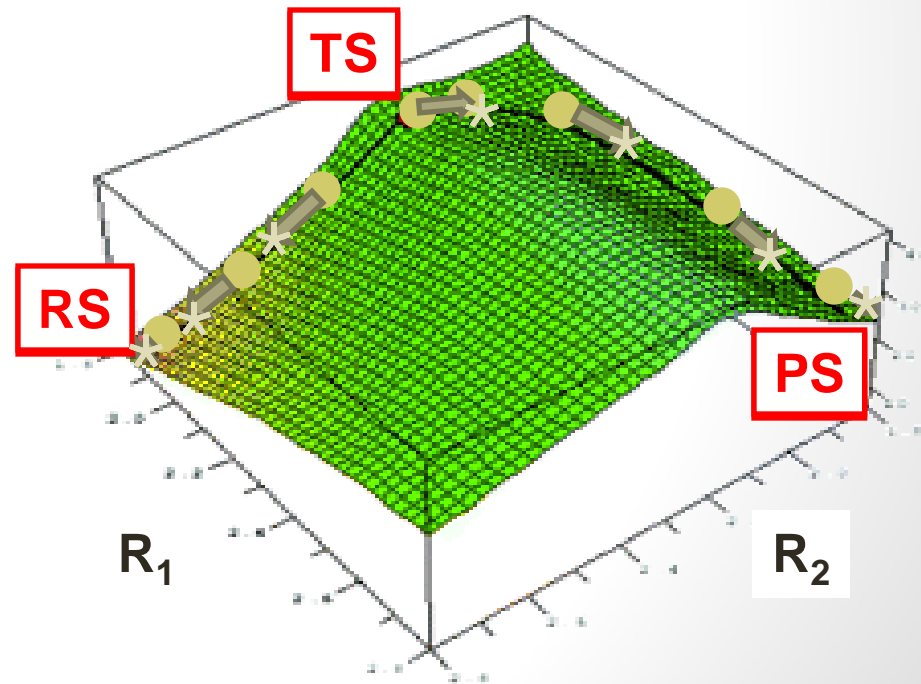
Proton Transfer Pathways

- **Step I.** **Deprotonation of water**

  a. *via* downstream phosphate group

  b. *via* Glu188

  c. *via* cleaved phosphate

     De Vivo et al. *JACS*, 2008

- **Step II.** **Protonation of leaving group**

  a. *via* 2'OH of sugar

  b. *via* cleaved phosphate

  c. *via* conserved Asp132

Thiol substitution experiments on the **accepting O** atom show a nearly ***10-fold reduction*** in reaction rate.

Haruki *et. al.*, *Biochemistry*, 2000
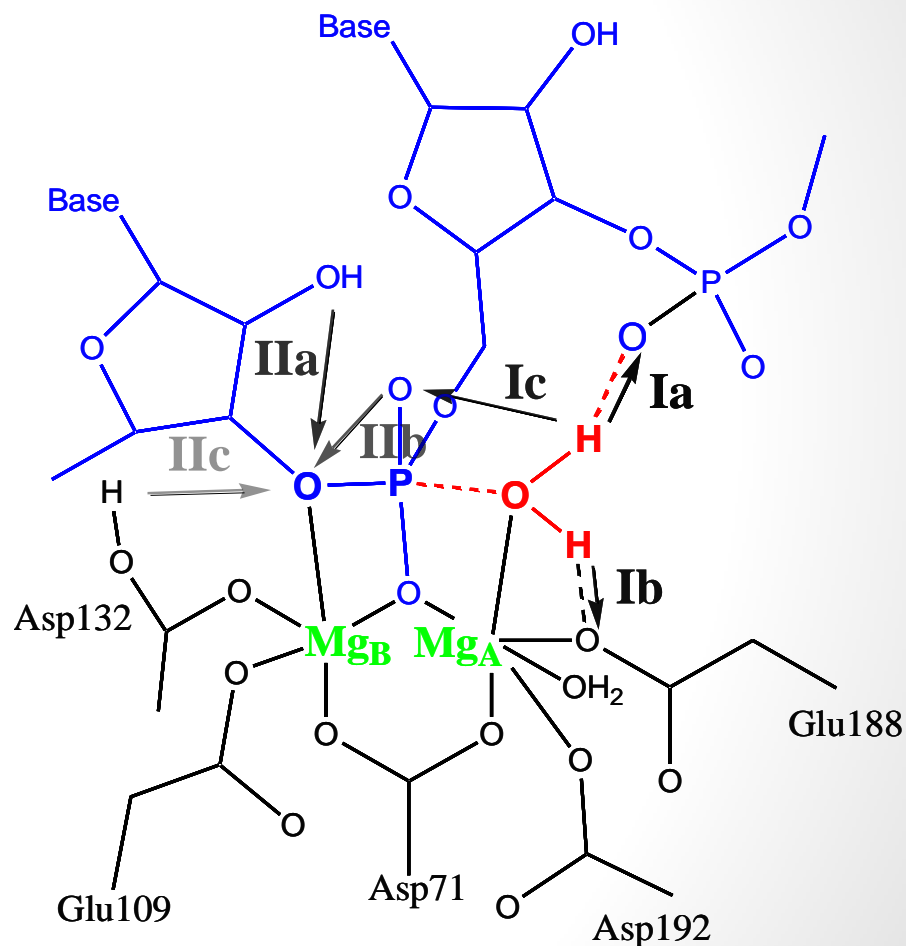
Rosta, Nowotny, Yang, Hummer, *J. Am. Chem. Soc.*, 2011

RNase H: Free Energy Surface
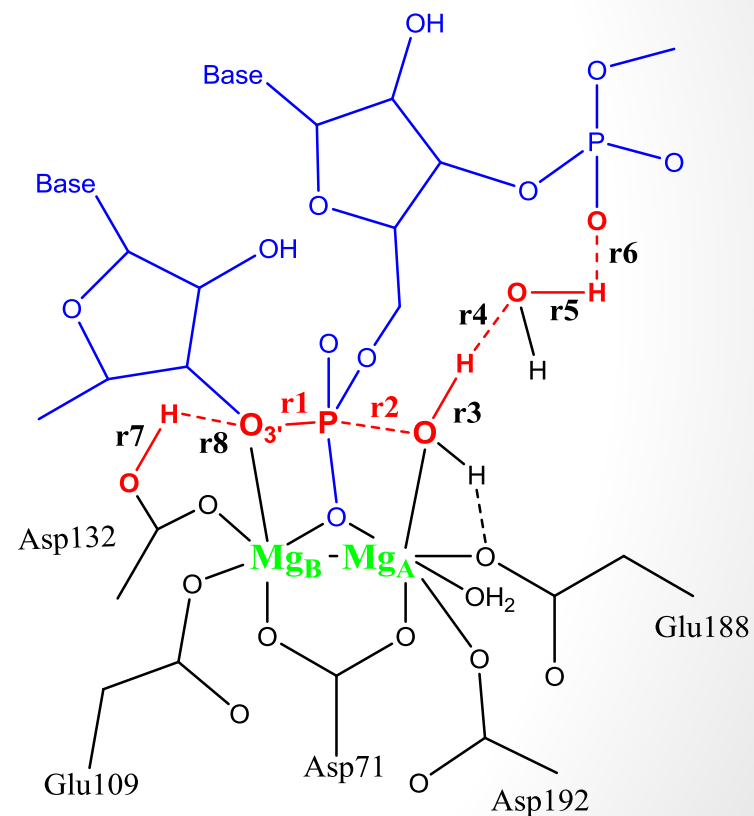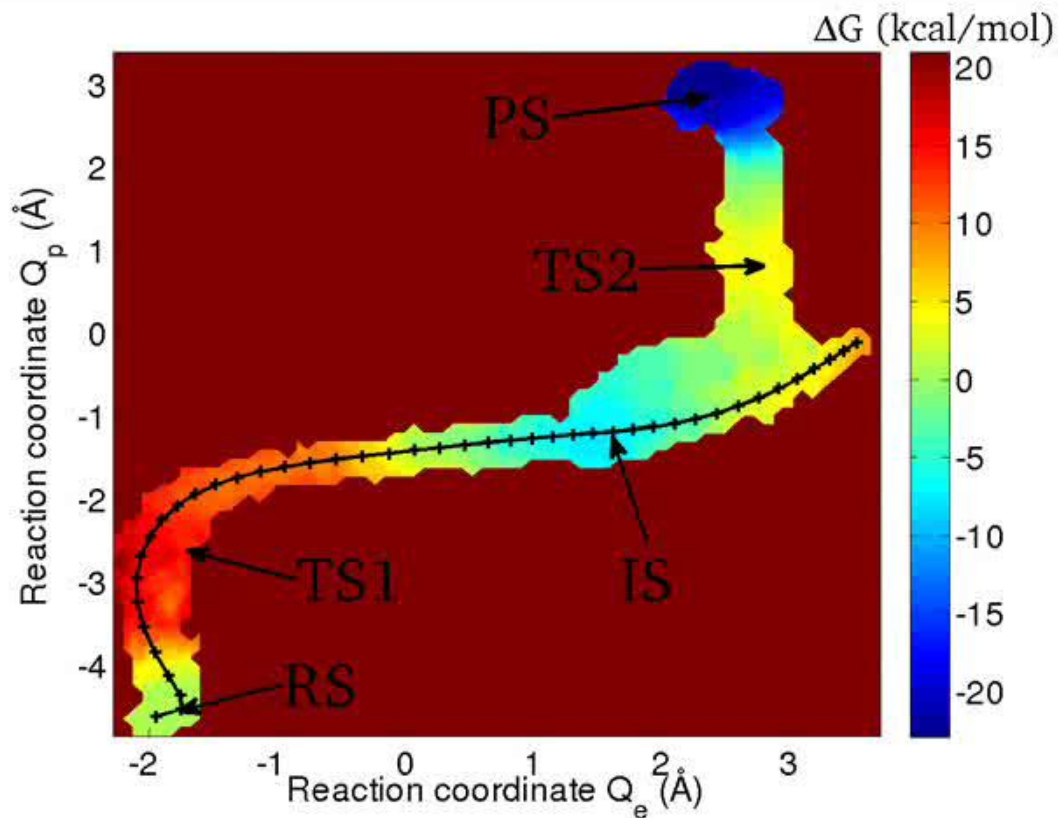
$Q_e = r_1 - r_2$

$Q_p = r_3 - r_4 + r_5 - r_6 + r_7 - r_8$

Free Energy Surface
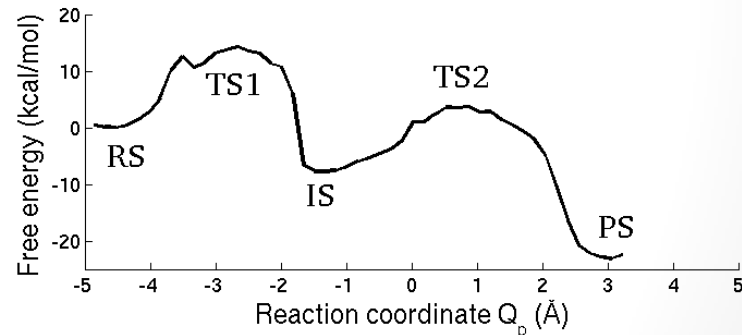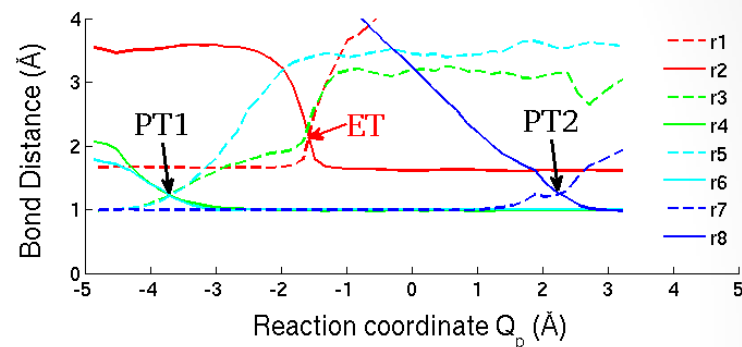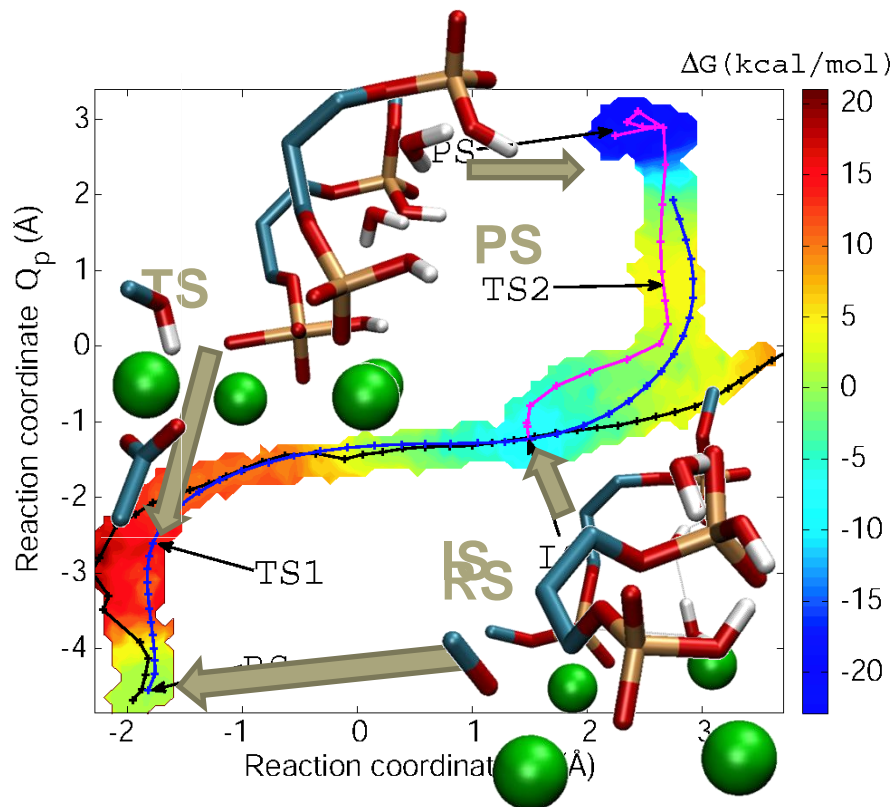
It has been thought that proton transfer is generally facile, however, we find that
barriers are dominated by proton transfer!

RNase H: Mechanism for Deprotonation (Step 1)
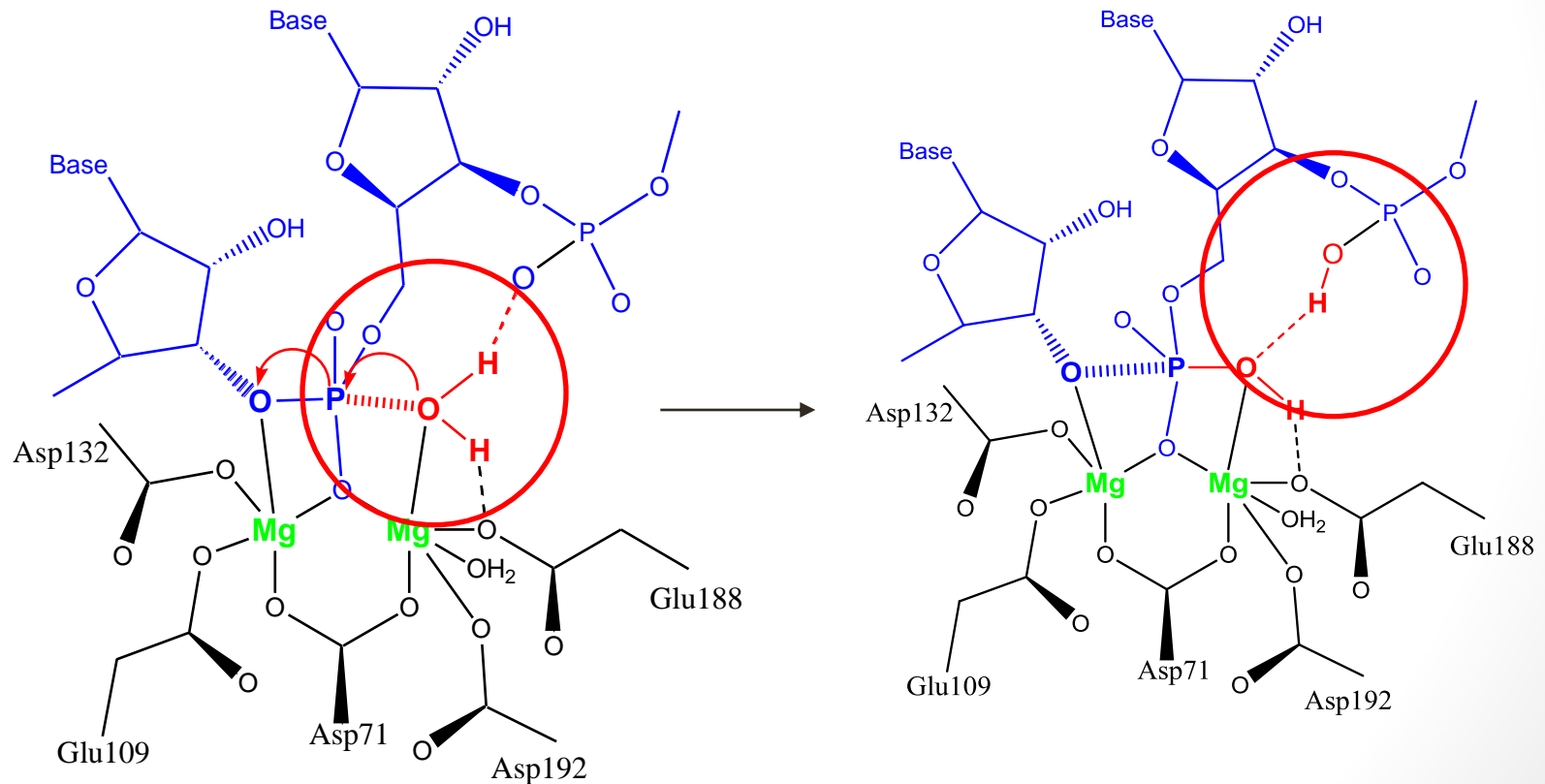
RNase H: Deprotonation *via* Downstream Phosphate
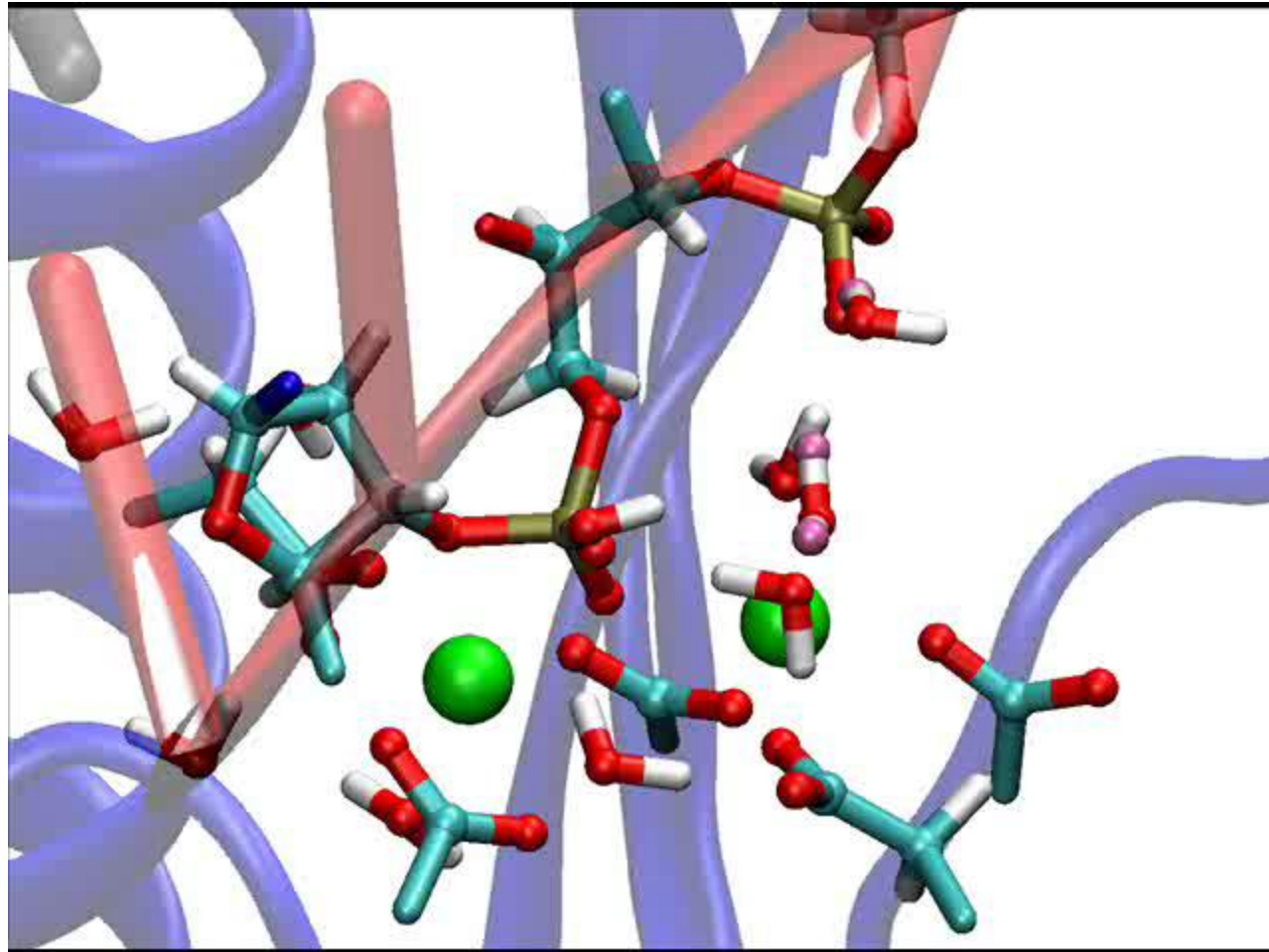
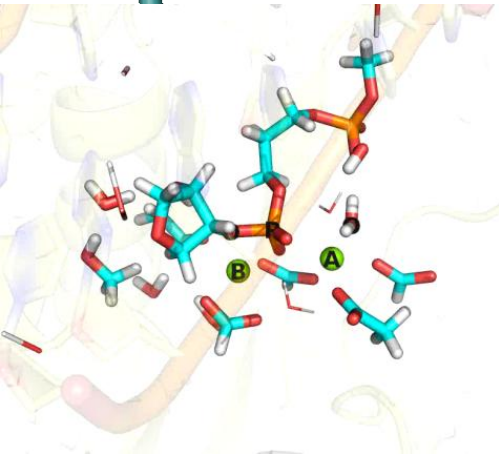# *RNase H:* Mechanism for Protonation (Step 2)
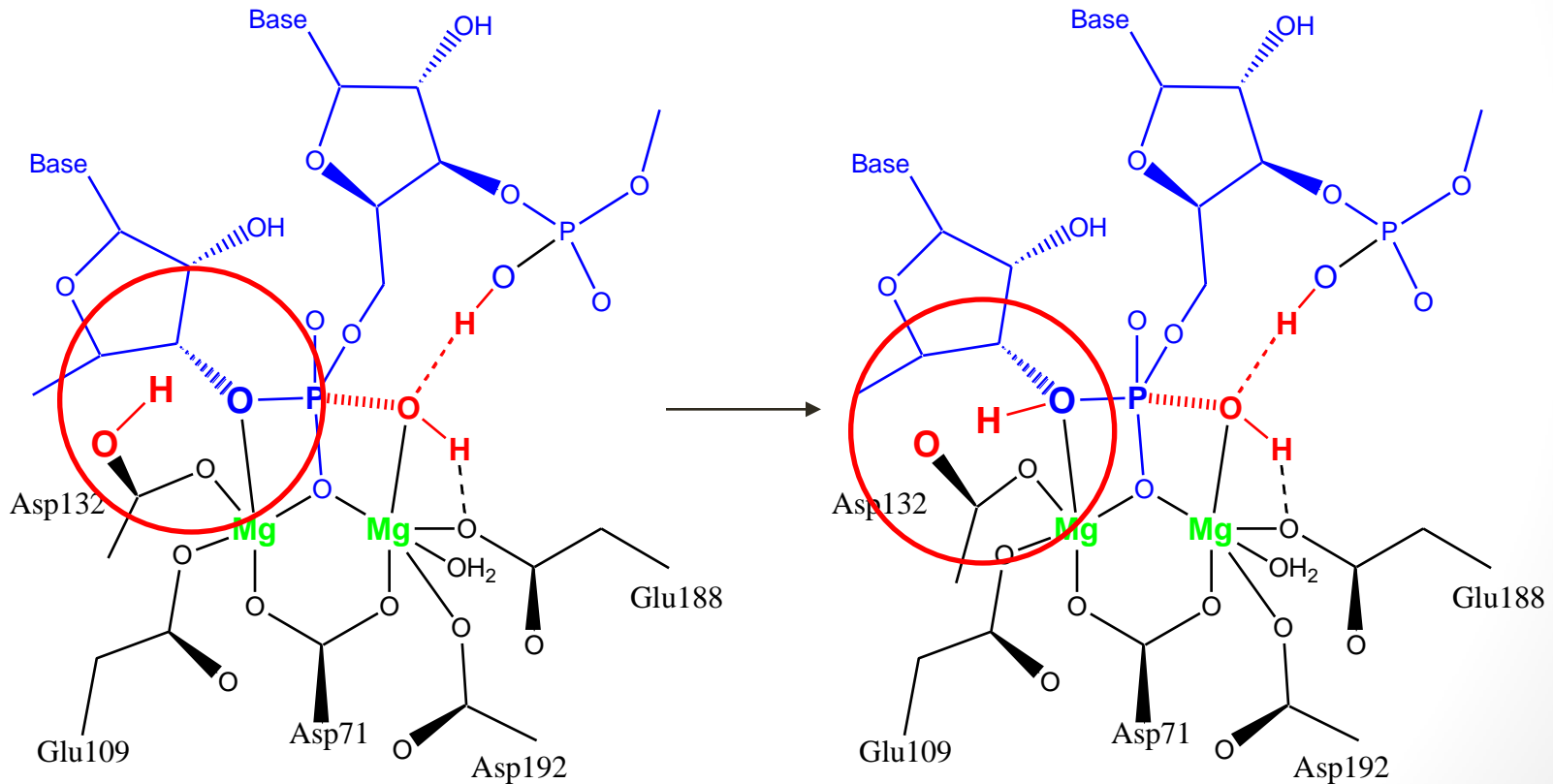
Product Formation *via* Asp132
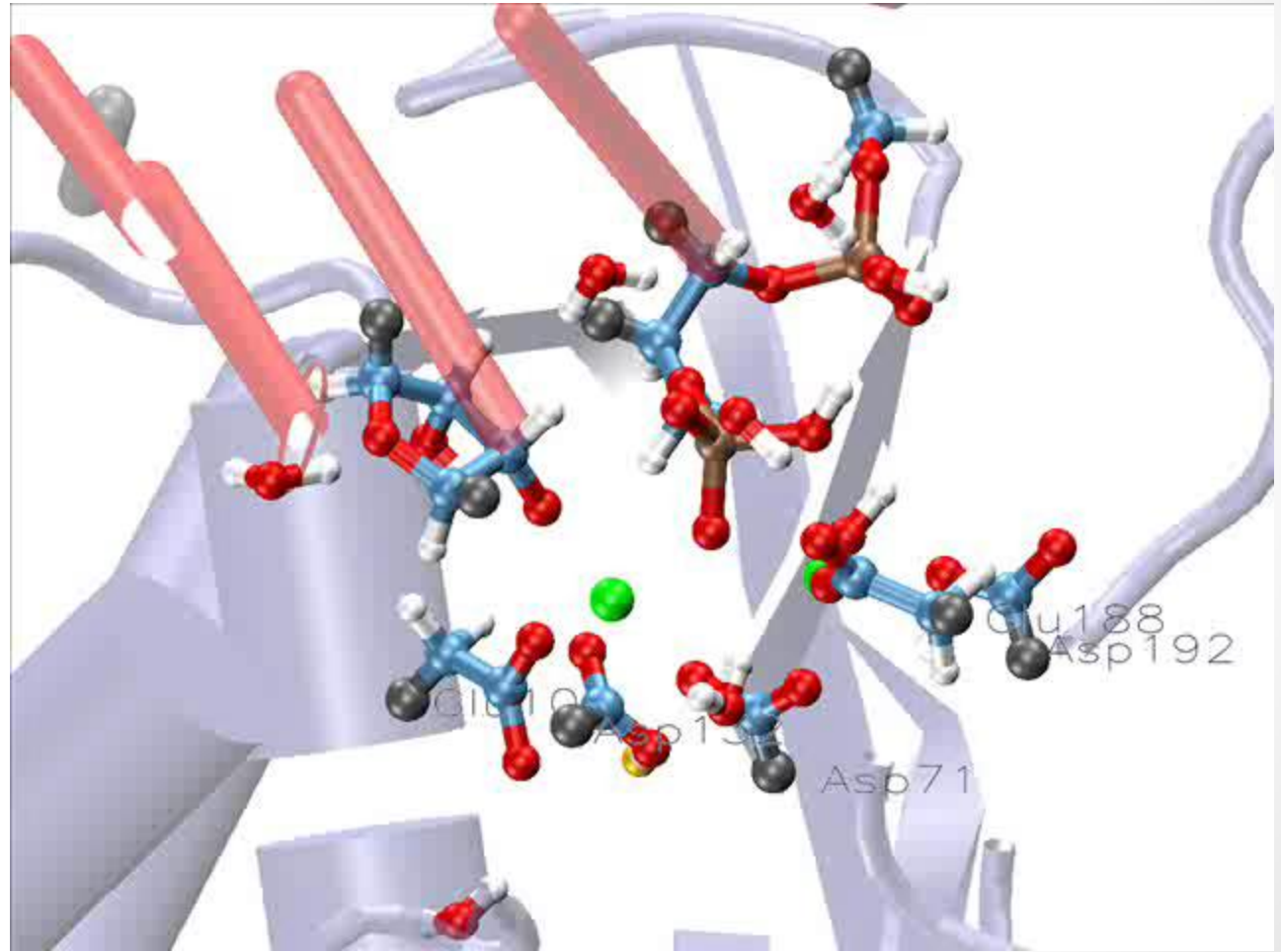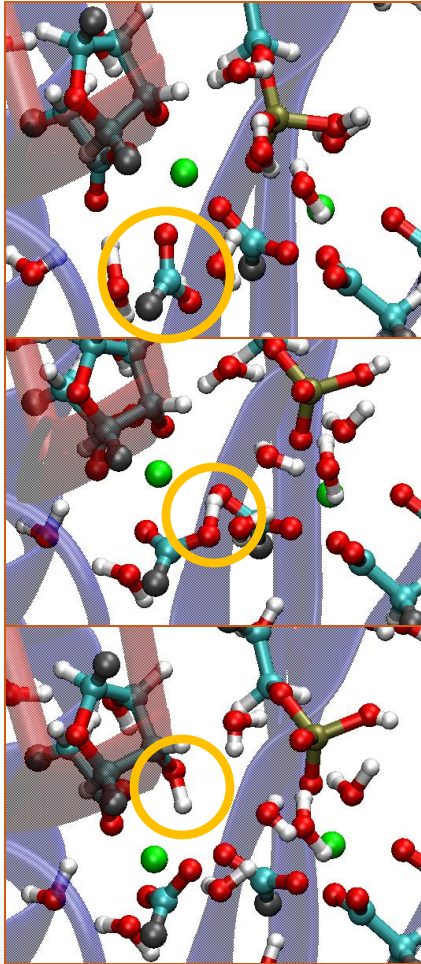
Pentacoordinated Mg^{2+}-ion B becomes more symmetrical @ TS
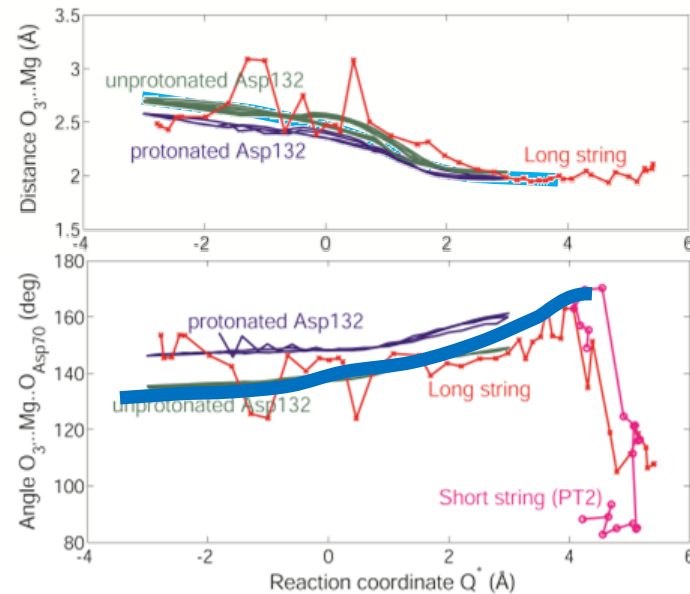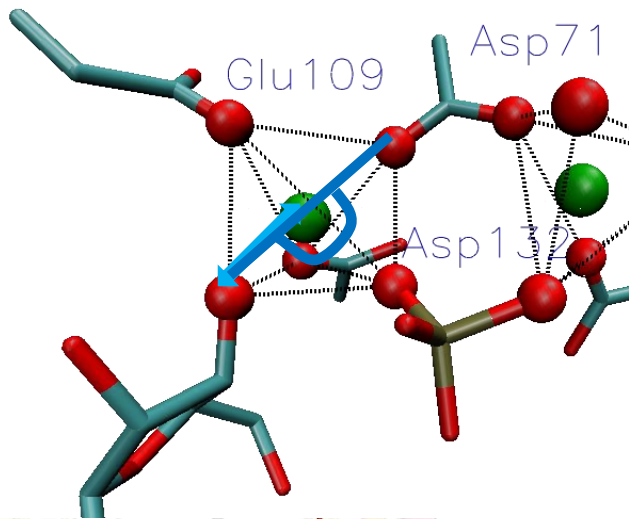
- More symmetric coordination of Mg^{2+}-ion *B* lowers the TS energy
- Shorter distance of Mg^{2+}-ion *B* to the leaving group stabilizes the negative charge by lowering the pK$_a$ of the leaving group
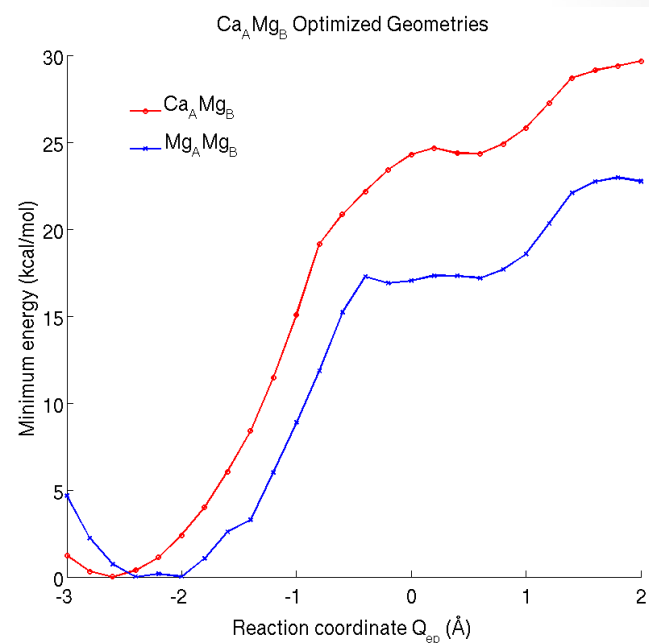
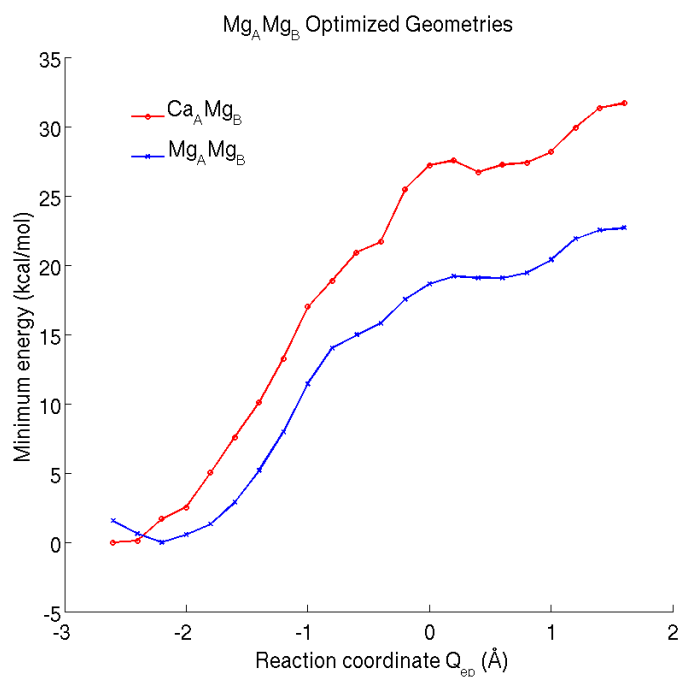# Single Metal-Ion Substitutions: QM Region and Reaction Coordinates
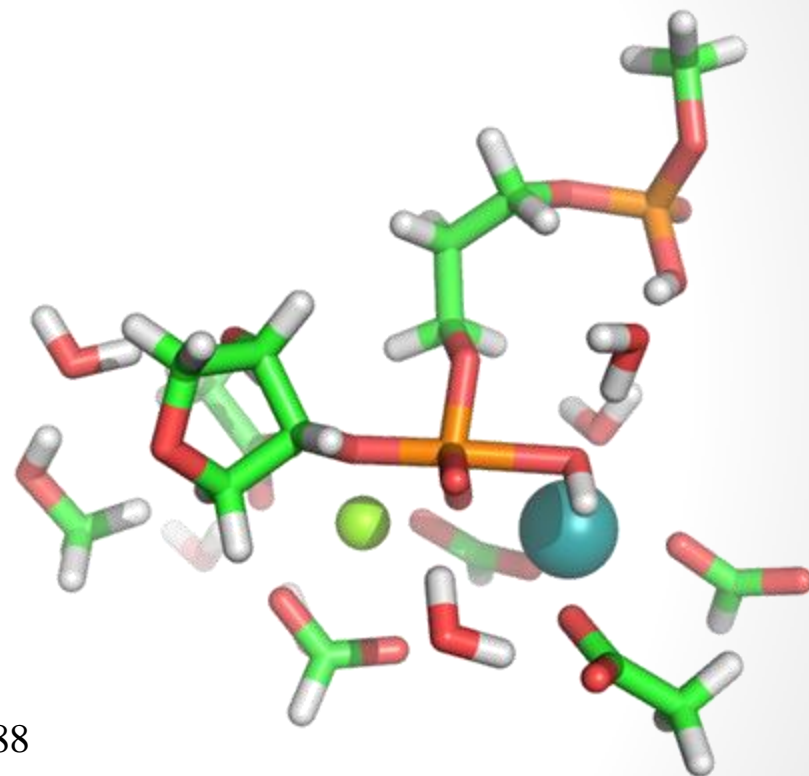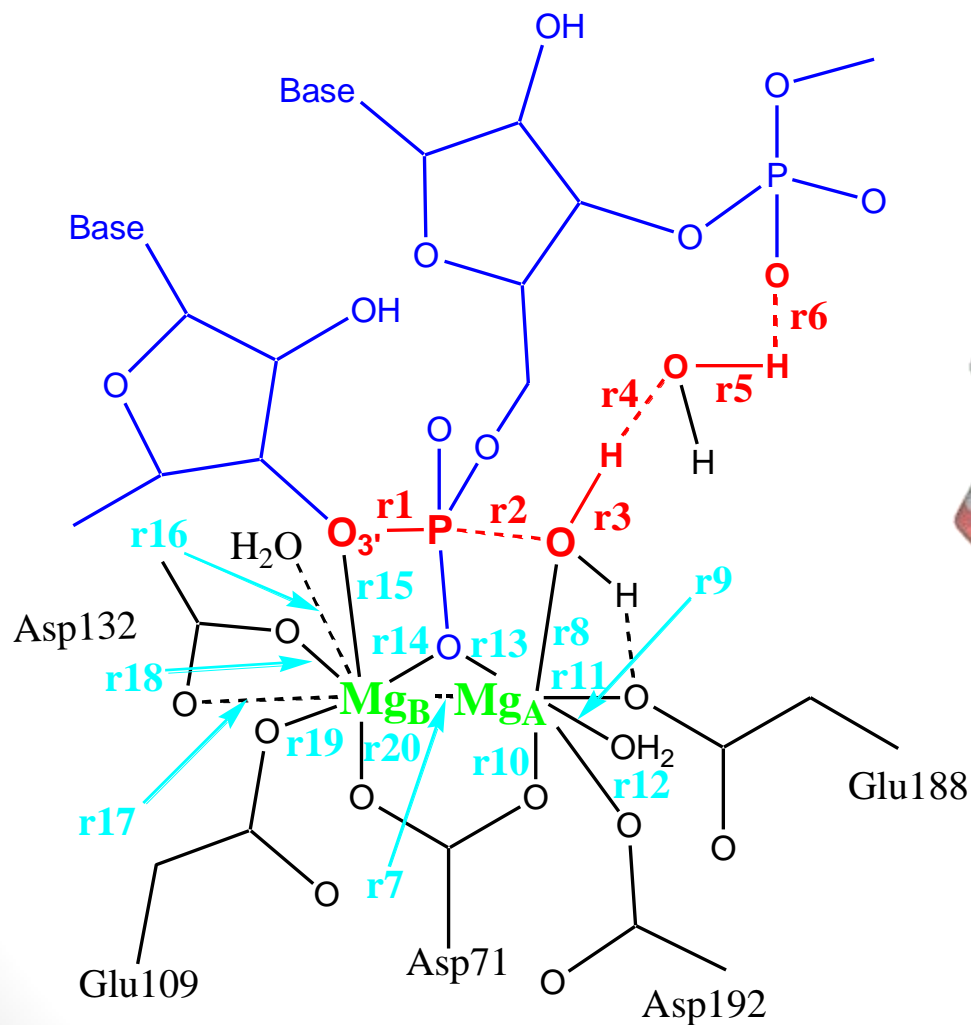
Substitutions at single metal ion sites: Mg^{2+} vs. Ca^{2+}

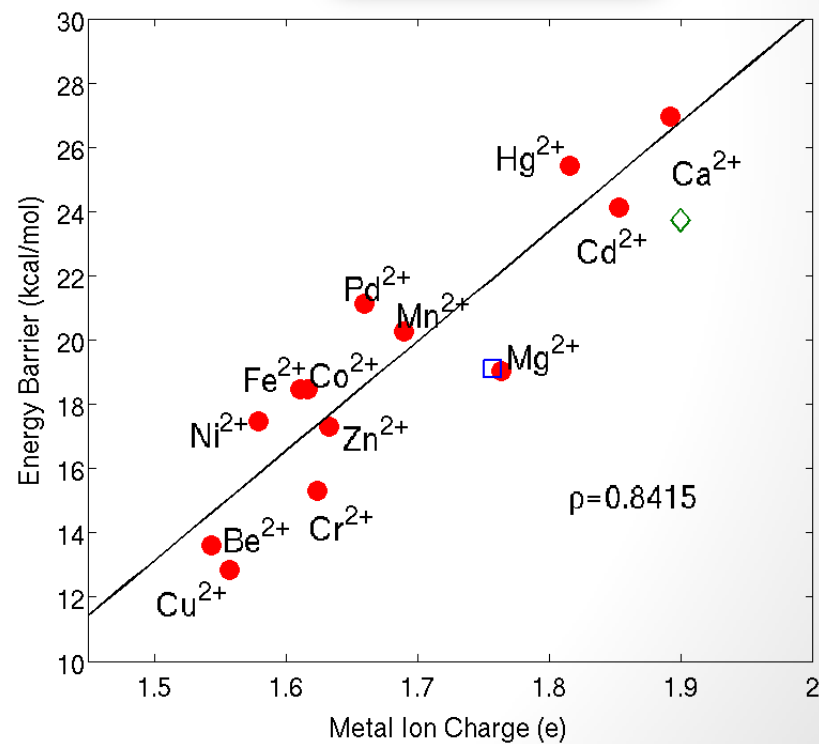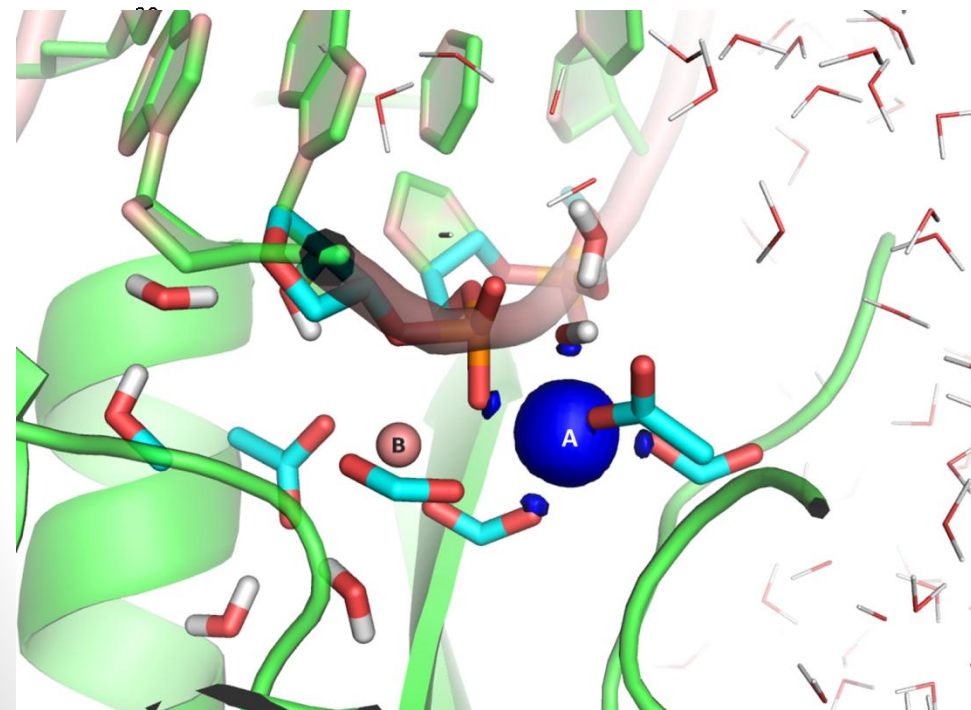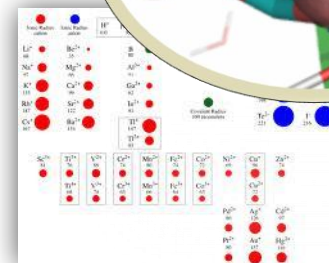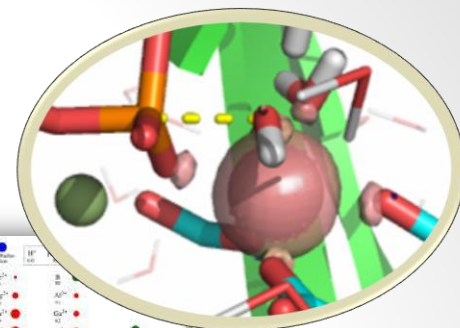Replacing either Mg^{2+} metal ions by Ca^{2+} abolishes catalysis.
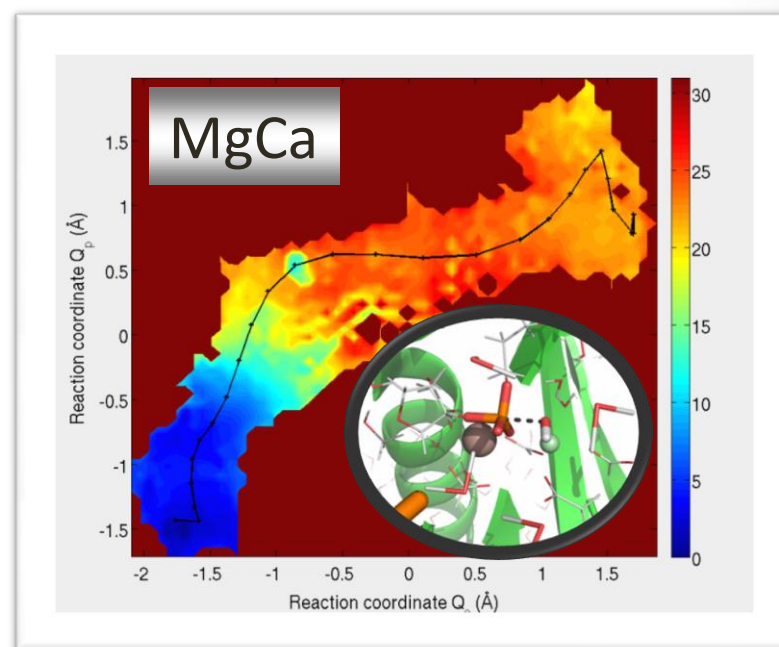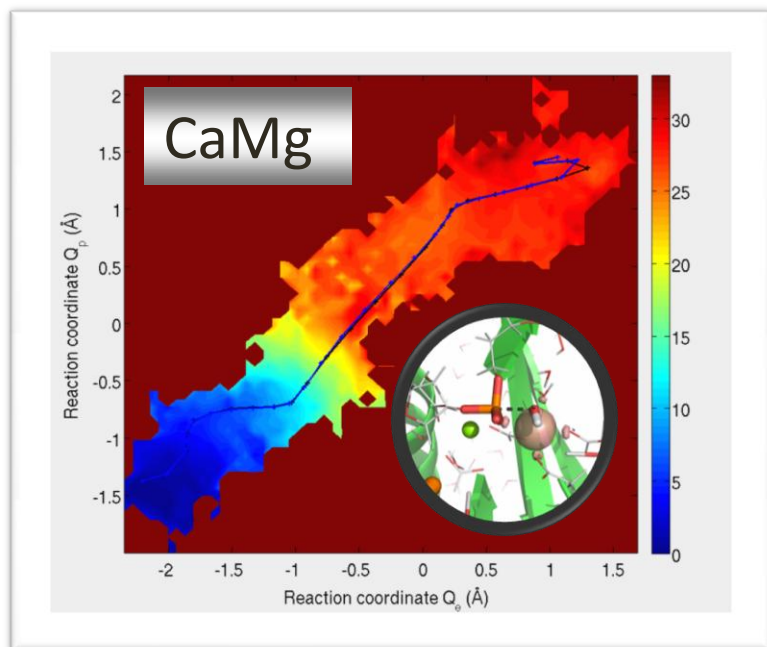
Metal Ion *A* Site: Geometric effects do not play a major role

Mg$_A$Mg$_B$ Optimized Geometries

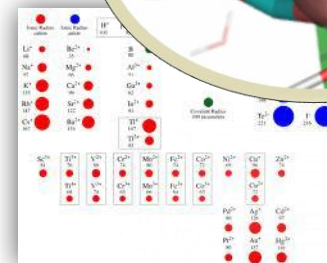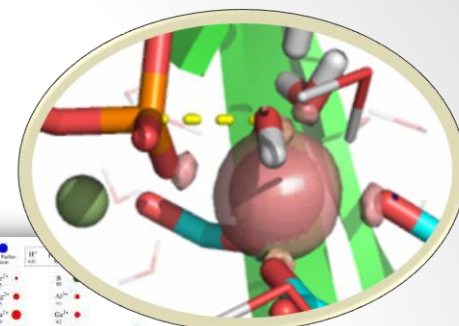Ca$_A$Mg$_B$ Optimized Geometries
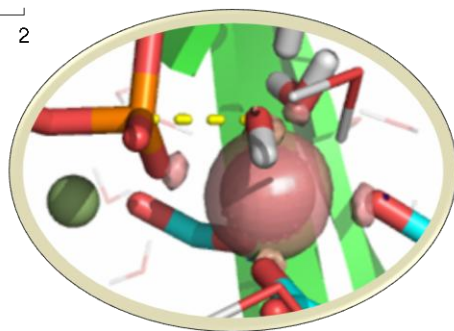
Role of partial charge transfer to metal ion *A*

Substitutions at metal ion *A* site with a series of divalent metal ions using Mg-optimized reaction pathway. (aug-cc-pVTZ(PP) basis set)
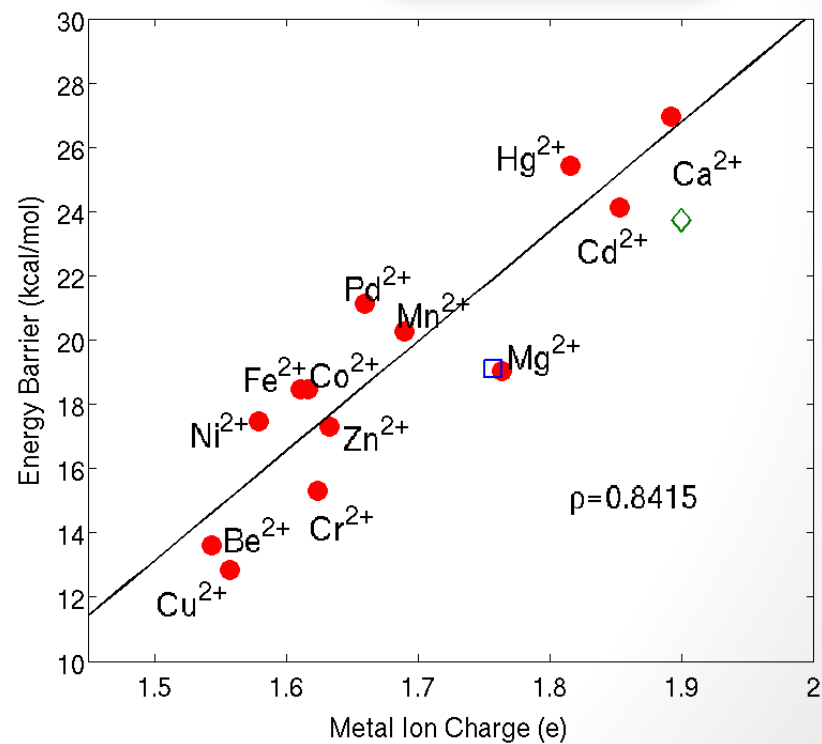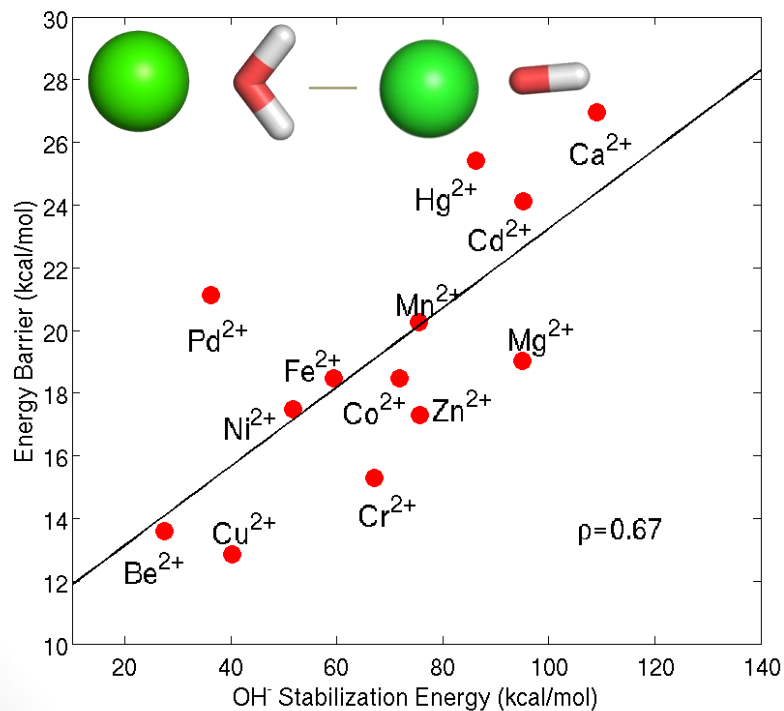
Role of metal ion *A*

Substitutions at metal ion *A* site with a series of divalent metal ions using Mg-optimized reaction pathway.
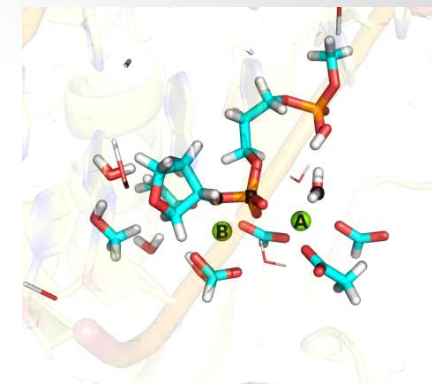
Summary

- A multidimensional finite temperature string reaction coordinate method is developed for QM/MM free energy calculations with application to RNase H.

- Hamiltonian Replica Exchange is extended to enhance sampling with *String Free Energy* simulations.

- Several PT mechanistic pathways are identified in RNase H catalytic reaction. The reaction barrier (rate) agrees well with experiment for the most probable path.

- PT dominates the reaction barrier suggesting that enzymes use metal ions to help lower the pK_a of attacking and leaving groups.

- High resolution TS structures can be used in further simulations for predicting changes in catalysis upon perturbations such *as binding of drug molecules*

Acknowledgements

Gerhard Hummer
(Laboratory Of Chemical Physics, NIDDK, NIH)

X-Ray crystallography:
Wei Yang (LMB, NIDDK, NIH)
Marcin Nowotny (IIMCB, Warsaw)
Kinetic theory:
Attila Szabo (LCP, NIDDK, NIH)
CHARMM + Q-Chem:
Bernard Brooks (NHLBI, NIH)
Lee Woodcock (USF, Tampa)
Yihan Shao (Q-Chem)