ハイブリッド分子動力学計算によるプロテイン・スプライシングの反応機構の解析

Hybrid molecular dynamics simulations of catalytic reaction of protein splicing

(Project in progress. Started: Jan. 2008)

共同研究者氏名 :

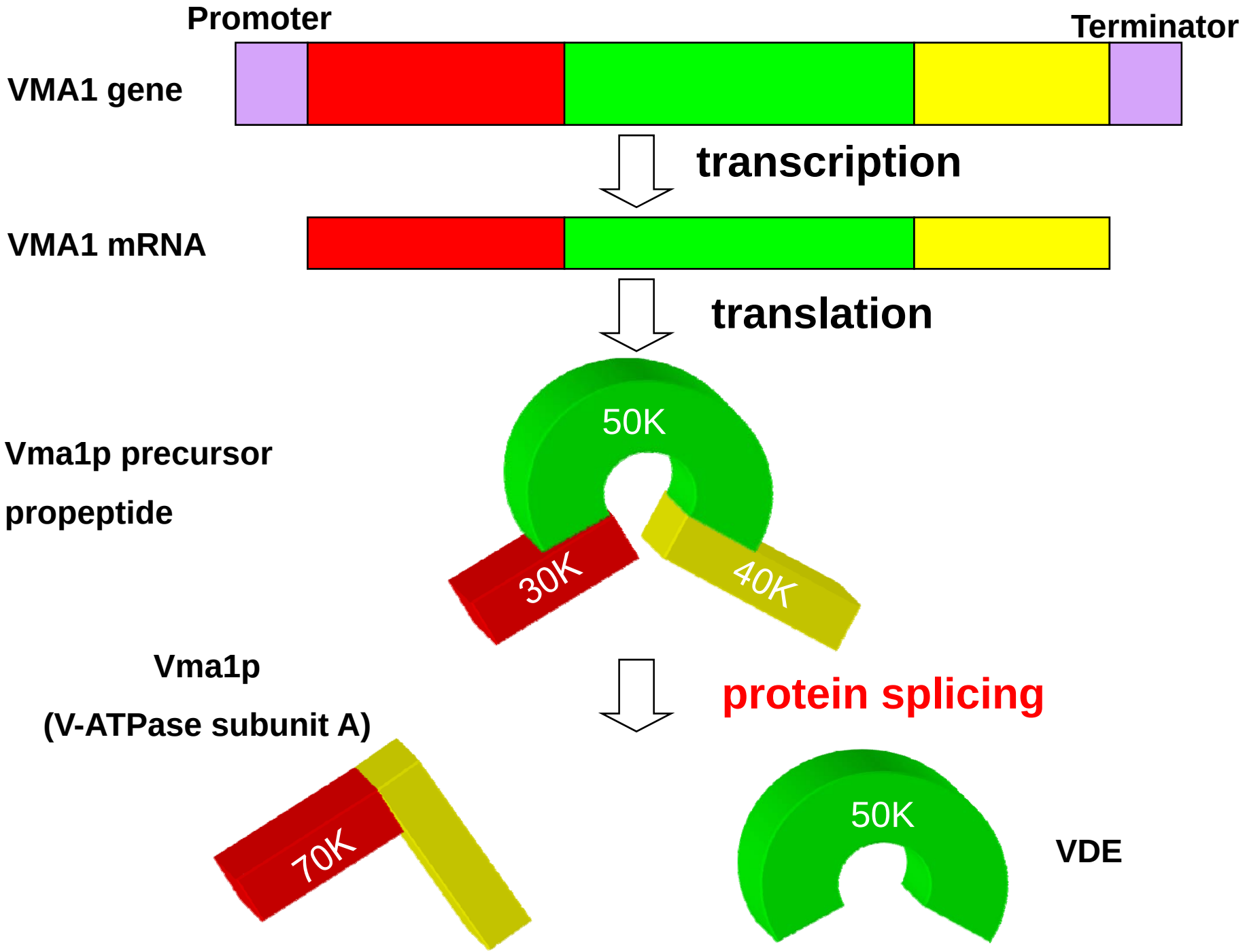
古明地　勇人　　（産業技術総合研究所　計算科学部門）

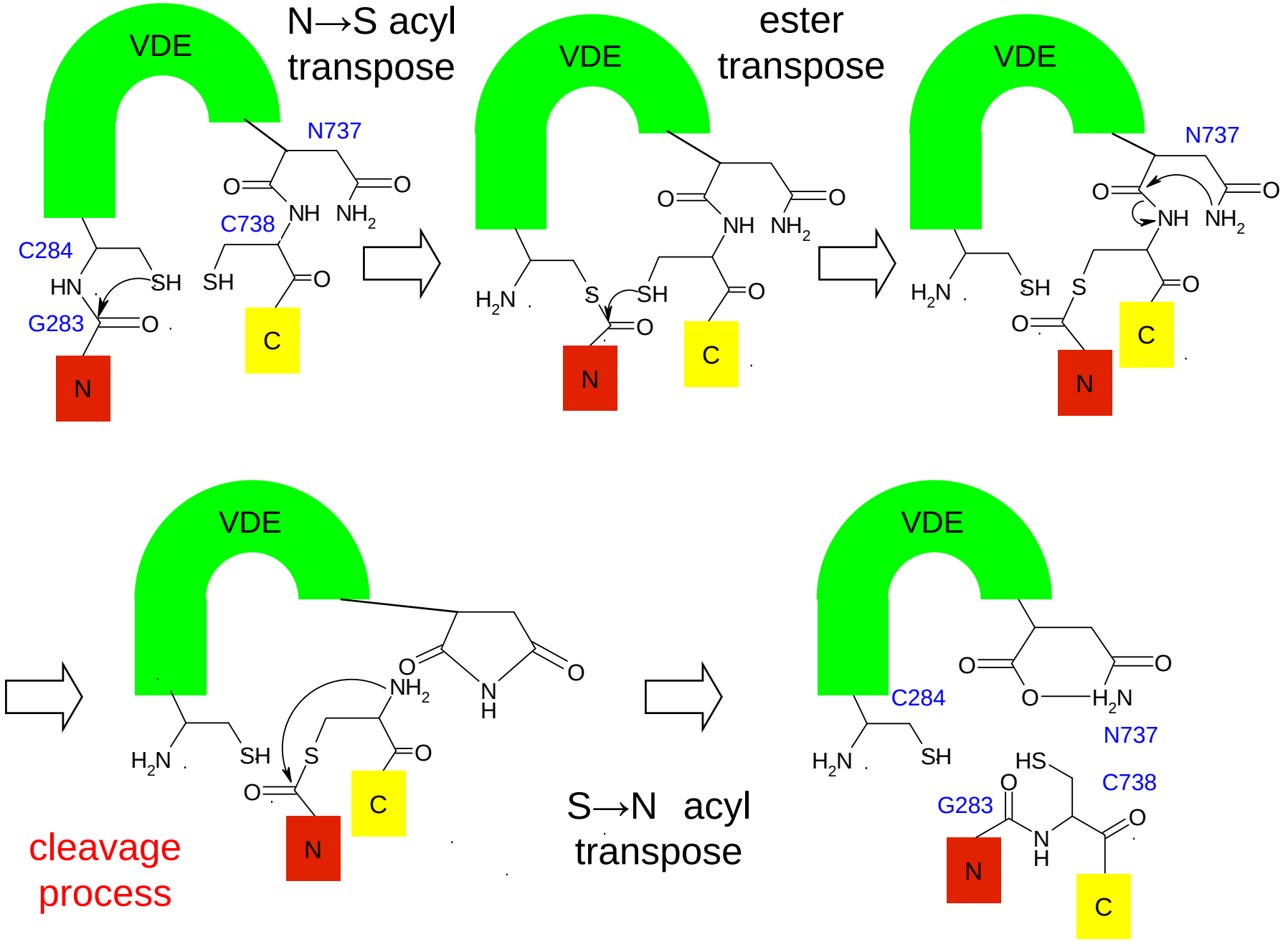
BOERO Mauro　（筑波大学計算科学研究センター）

舘野　賢　　（筑波大学計算科学研究センター）

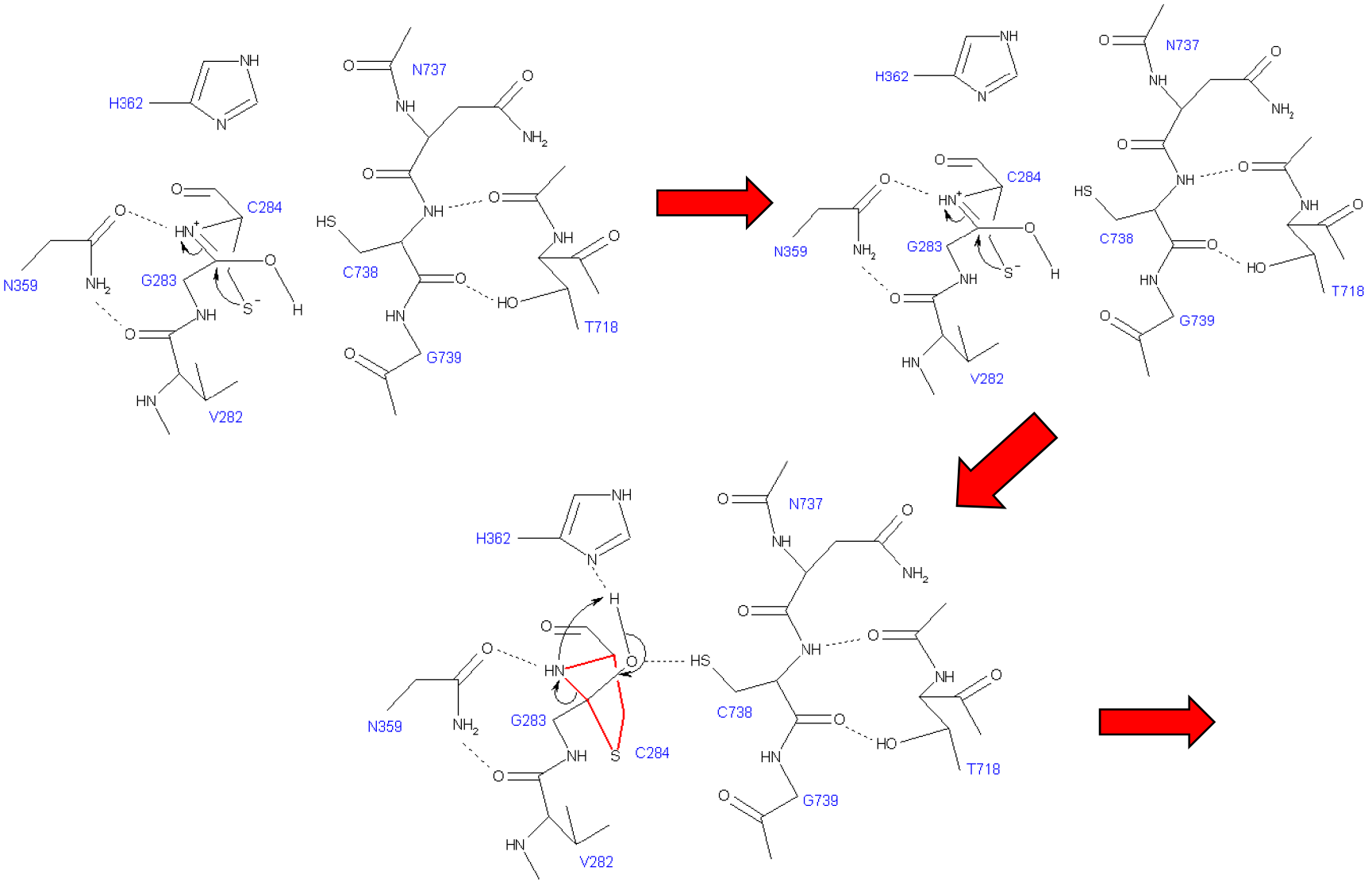
Outline of the problem

- Splicing of proteins is a common process in any bio-system, ***but*** it occurs so rapidly that the precursor protein is rarely observed in native systems (R. Mizutani et al. *J. Mol. Biol*. **316**, 919 (2002)).
- The process is a fundamental reaction in of the cell, but due to its complexity, still largely unknown.
- At least four nucleophiclic attacks by three different residues are involved (F. B. Perler, *Nucleic Acids Res*. **30**, 383 (2002)) and a microscopic picture is hard to obtain experimentally.
- The current knowledge of the splicing mechanism comes mostly from the work of Perler (M. Q. Xu and F. B. Perler, *EMBO J*. **15**, 5146 (1996); F. B. Perler, *Nature Struct. Biol*. **5**, 249 (1998); *Cell* **92**, 1 (1998)) and Paulus (H. Paulus, *Chem. Soc. Rev*. 27, 375 (1998))

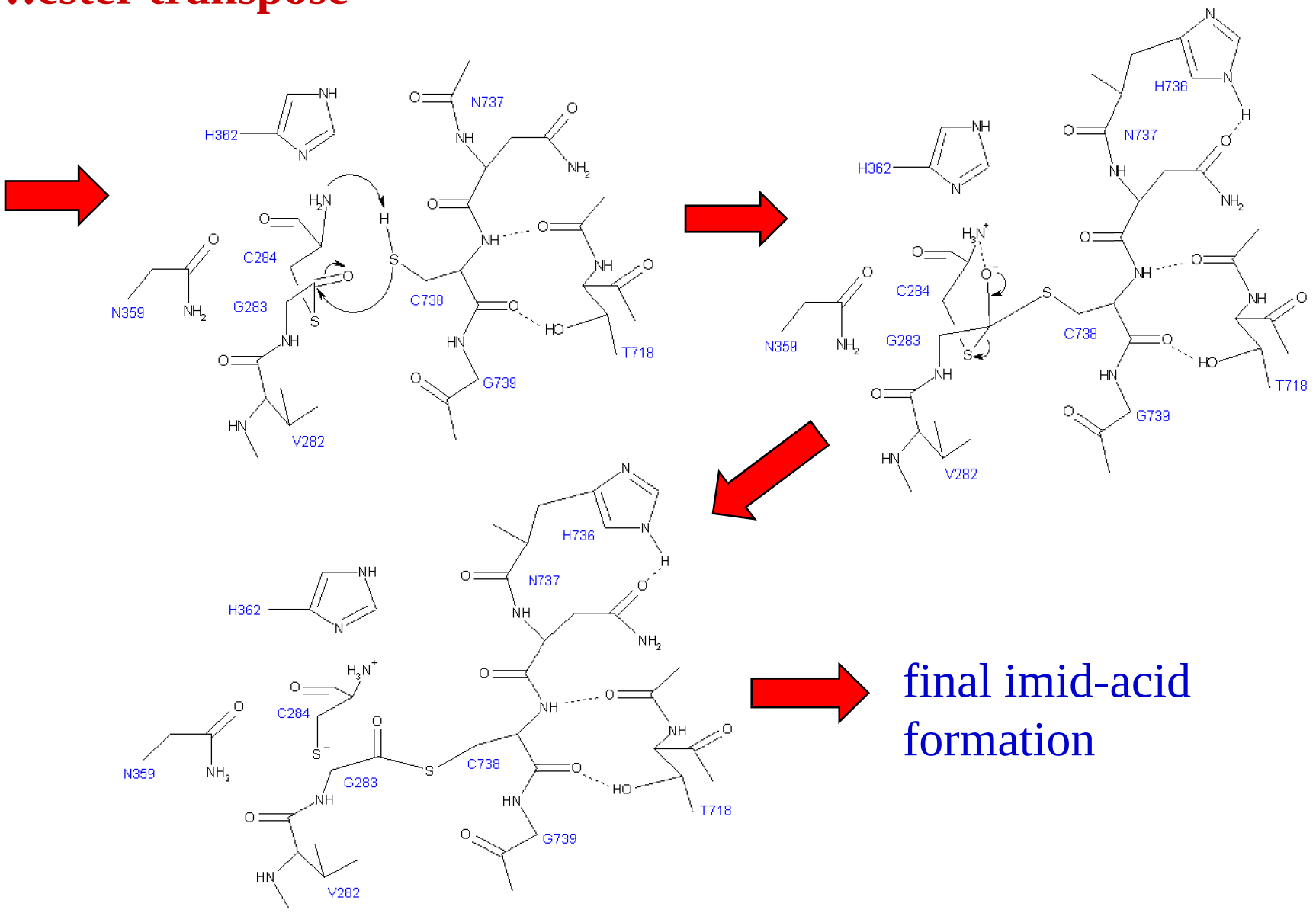

Promoter
Terminator
VMA1 gene
transcription
VMA1 mRNA
translation
Vma1p precursor
propeptide
50K
30K
40K
Vma1p
(V-ATPase subunit A)
protein splicing
70K
50K
VDE


VDE
N→S acyl transpose
N737
C738
C284
G283
HN
SH
NH
NH2
O
C
N
ester transpose
H2N
S
cleavage process
S→N acyl transpose
HS

N→S acyl transpose…



…ester transpose



Size Problem: hybrid QM/MM approach

$$H^{tot} = H^{CP/DFT} + E^{\text{int}}[\rho, \{\mathbf{r}_i\}] + H^{MM}$$

Pure QM interaction

N_{el} x N_G x N_G

with

$\psi_j(\mathbf{x}) = \sum_{\mathbf{G}} c_j(\mathbf{G})\, e^{i\mathbf{G}\mathbf{x}}$

QM/MM interface

NN x N_G

AMBER

Size Problem: electrostatic interaction

Functional form (*Ncl* = number of classical atoms):

$$E^{\text{int}}\left[\rho(\mathbf{r}),\{q_I\}\right]=\sum_{I=1}^{Ncl} q_I \int d^3r \frac{\rho(\mathbf{r})}{\left|\mathbf{r}-\mathbf{r}_I\right|}$$

Potential acting on the QM wave functions $\psi_i(\mathbf{x})$:

$$\frac{\delta E^{\text{int}}}{\delta \rho}=\sum_{I=1}^{Ncl}\frac{q_I}{\left|\mathbf{r}-\mathbf{r}_I\right|}=V^{\text{int}}(\mathbf{r})$$

Expensive if *Ncl* is large !

Forces acting on the MM charged atoms:

$$\frac{\partial E^{\text{int}}}{\partial \mathbf{r}_I}=-q_I\int d^3r \frac{\rho(\mathbf{r})}{\left|\mathbf{r}-\mathbf{r}_I\right|^3}\left(\mathbf{r}-\mathbf{r}_I\right)=\mathbf{F}_I^{\text{int}}$$

Size Problem: reduce the computational cost

Divide the world in 3 domains

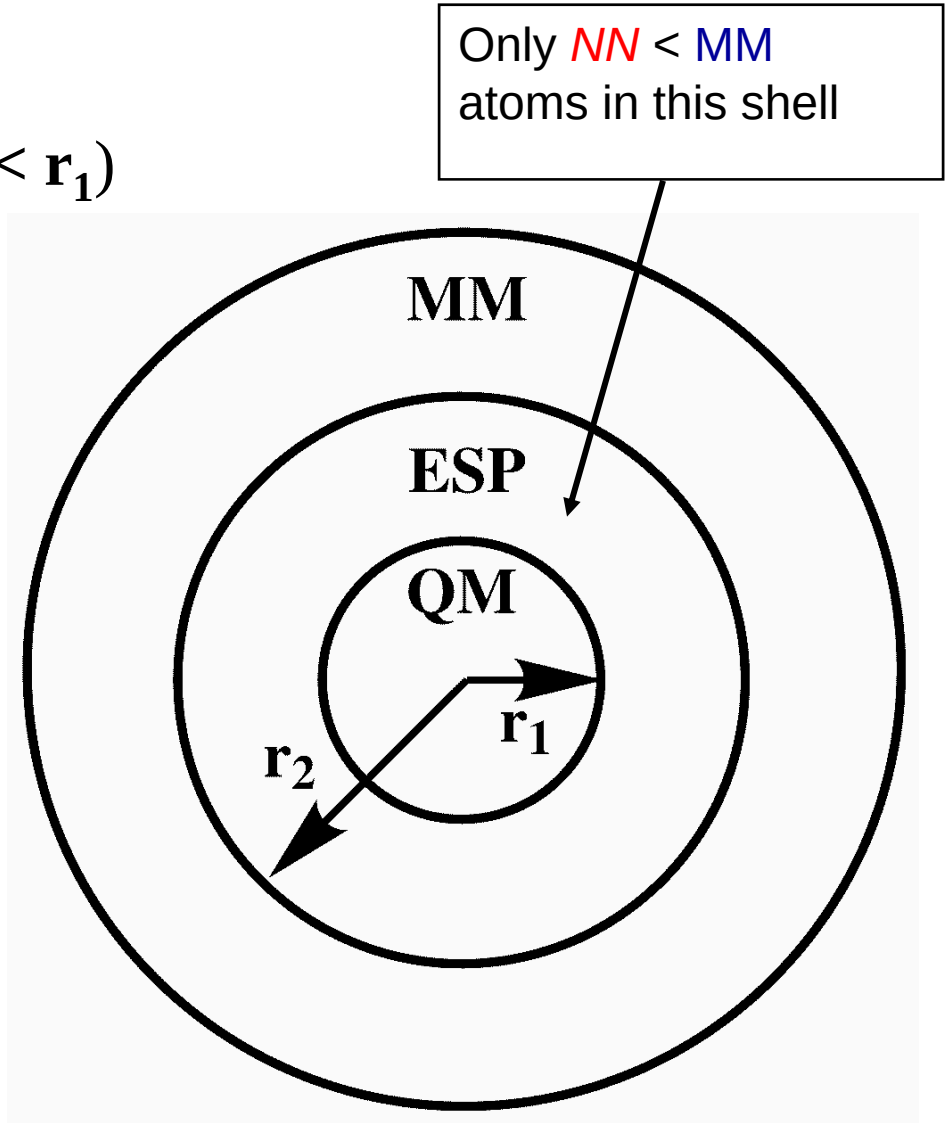
1) Close to the QM region ($\mathbf{r} < \mathbf{r}_1$)
2) Not too far, i.e. ESP region ($\mathbf{r}_1 < \mathbf{r} < \mathbf{r}_2$)
3) Far MM world ($\mathbf{r} > \mathbf{r}_2$)

Test with $r_1 = r_2 = \infty$

In all known cases (so far)

$\mathbf{r}_1$ ~ 10-12 a.u.

$\mathbf{r}_2$ ~ 20-25 a.u.



Size Problem: 3-regions scheme

Region 1: *NN* << *Ncl* → only a subset of *Ncl*

$$\sum_{I=1}^{NN} q_I \int d^3 r \frac{\rho(\mathbf{r})}{\left|\mathbf{r}-\mathbf{r}_I\right|} \qquad r < r_1$$

Region 2: Classical-RESP charges interaction:

$$\sum_{I \in NN} q_I \sum_{J \in QM} \frac{q_J^{RESP}(\rho, \mathbf{r}_I)}{\left|\mathbf{r}_I - \mathbf{r}_J\right|} \qquad r_1 < r < r_2$$

Region 3: Multipolar expansion on MM charges:

$$\sum_{I \in NN} q_I \sum_{\alpha} \frac{\wp^{\alpha}(\rho)\left(\mathbf{r}-\mathbf{r}_I\right)^{\alpha}}{\left|\mathbf{r}-\mathbf{r}_I\right|^3} + quadrupole \qquad r > r_2$$

Size Problem: Dynamical – Restrained ElectroStatic Potential (D-RESP)

• Ask the D-RESP potential to be as close as possible to the true **e**lectro**s**tatic **p**otential (ESP) V_J

• **R**estrain the charge (R-ESP) to avoid unphysical dynamical fluctuations

$$\chi = \sum_{J \in NN} \left(\sum_{I \in QM} \frac{q_I^D}{|\mathbf{r}_I - \mathbf{r}_J|} - V_J \right)^2 + w_q \sum_{I \in QM} \left(q_I^D - q_I^H \right)^2$$

ESP

$q_I^D = q_I^{RESP}$

Restrain

Size Problem: dynamic RESP charges (q_I^D)

$$\chi = \sum_{J \in NN} \left(\sum_{I \in QM} \frac{q_I^D}{|\mathbf{r}_I - \mathbf{r}_J|} - V_J \right)^2 + w_q \sum_{I \in QM} \left(q_I^D - q_I^H \right)^2$$

is minimized *on the fly* during the dynamics.

w_q = weight parameter to reduce charge fluctuations

$w_q \approx 0.10 - 0.25$

$$V_J = \int d^3 r \, \rho(\mathbf{r}) \, u\left(|\mathbf{r} - \mathbf{r}_J|\right)$$

$u(|\mathbf{r} - \mathbf{r}_J|)$ = Coulomb potential modified at short range to avoid spurious over-polarization effects

A. Laio et al. *J. Phys. Chem. B* **106**, 7300 (2002)

Size Problem: the charge restraint

$$w_q \sum_{I \in QM} \left(q_I^D - q_I^H \right)$$

q_I^H are the Hirshfeld charges* (F. L. Hirshfeld, *Theo. Chim. Acta* **44**, 129 (1977))

$$q_I^H = \int d^3r \, \rho(\mathbf{r}) \frac{\rho^{at}\left(\left|\mathbf{r} - \mathbf{r}_I\right|\right)}{\sum_K \rho^{at}\left(\left|\mathbf{r} - \mathbf{r}_K\right|\right)} - Z_I$$

r^{at} is the atomic (pseudo) valence charge density and

$$Z_I = \int d^3r \, \rho^{at}\left(\mathbf{r} - \mathbf{r}_I\right)$$

valence of the *I*-th atom

Size Problem: RESP coupling potential

$$V^{RESP}\left(\rho,\{q_J\}\right)=\sum_{J\in NN}\sum_{I\in QM}\frac{q_J q_I^D}{\left|\mathbf{r}_I-\mathbf{r}_J\right|}$$

replaces the more expensive $\sum_J q_J\int d^3r\,\rho(\mathbf{r})\,u\left(\left|\mathbf{r}-\mathbf{r}_J\right|\right)$

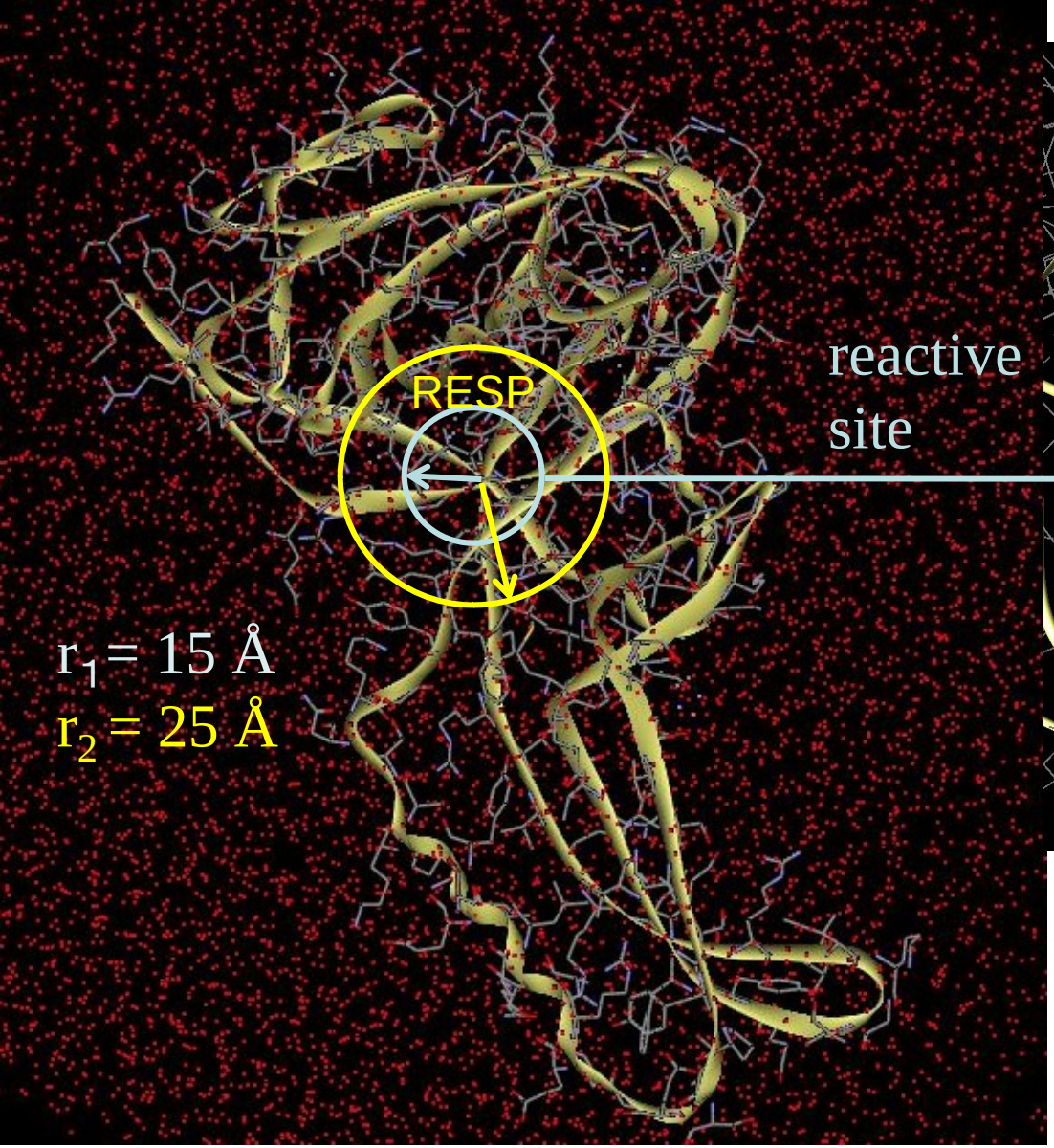
Coupling potential on electrons:

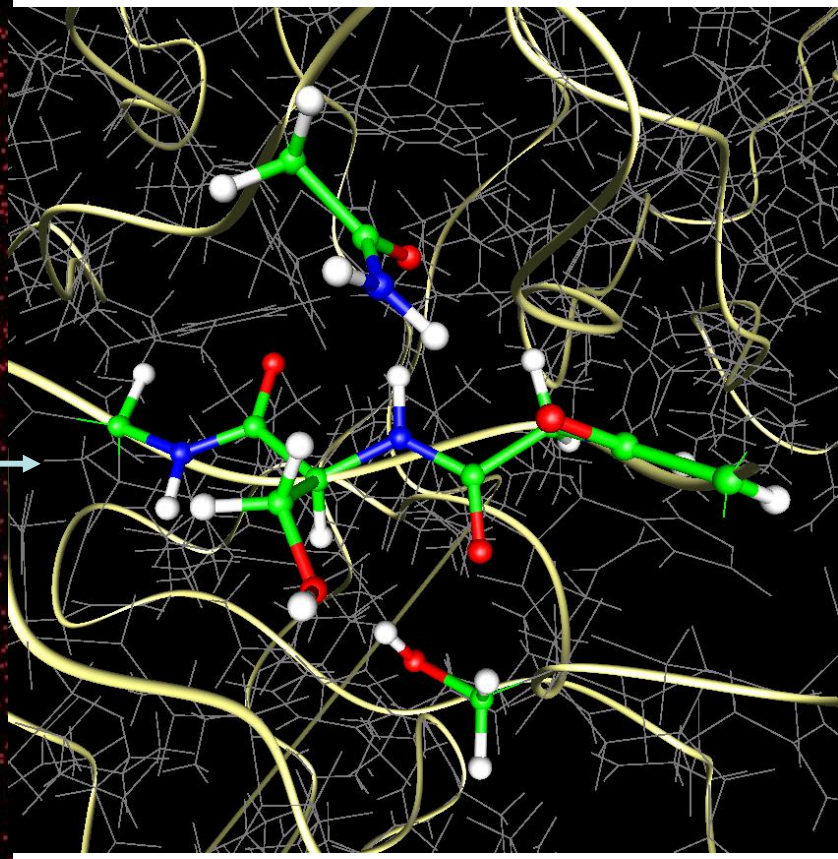
$$v(\mathbf{r})=\frac{\delta V^{RESP}}{\delta\rho(\mathbf{r})}=\sum_{I\in QM}\frac{\partial V^{RESP}}{\partial q_I^D}\frac{\delta q_I^D}{\delta\rho(\mathbf{r})}$$

Forces components on atoms:

$$\mathbf{F}_J=-\nabla_{\mathbf{r}_J}V^{RESP}=-\frac{\partial V^{RESP}}{\partial\mathbf{r}_J}-\sum_{I\in QM}\frac{\partial V^{RESP}}{\partial q_I^D}\frac{\partial q_I^D}{\partial\mathbf{r}_J}$$

Details in M.B. and M. Tateno, in *Modelling Structure and Reactivity in Biological Systems* ed. by J. J. Naidoo, J. Brady, M. Field, J. Gao and M. Hann, RSC Publishing, Cape Town, July 2006.





Protein in solution

MM: 46331 atoms

QM: 45 atoms

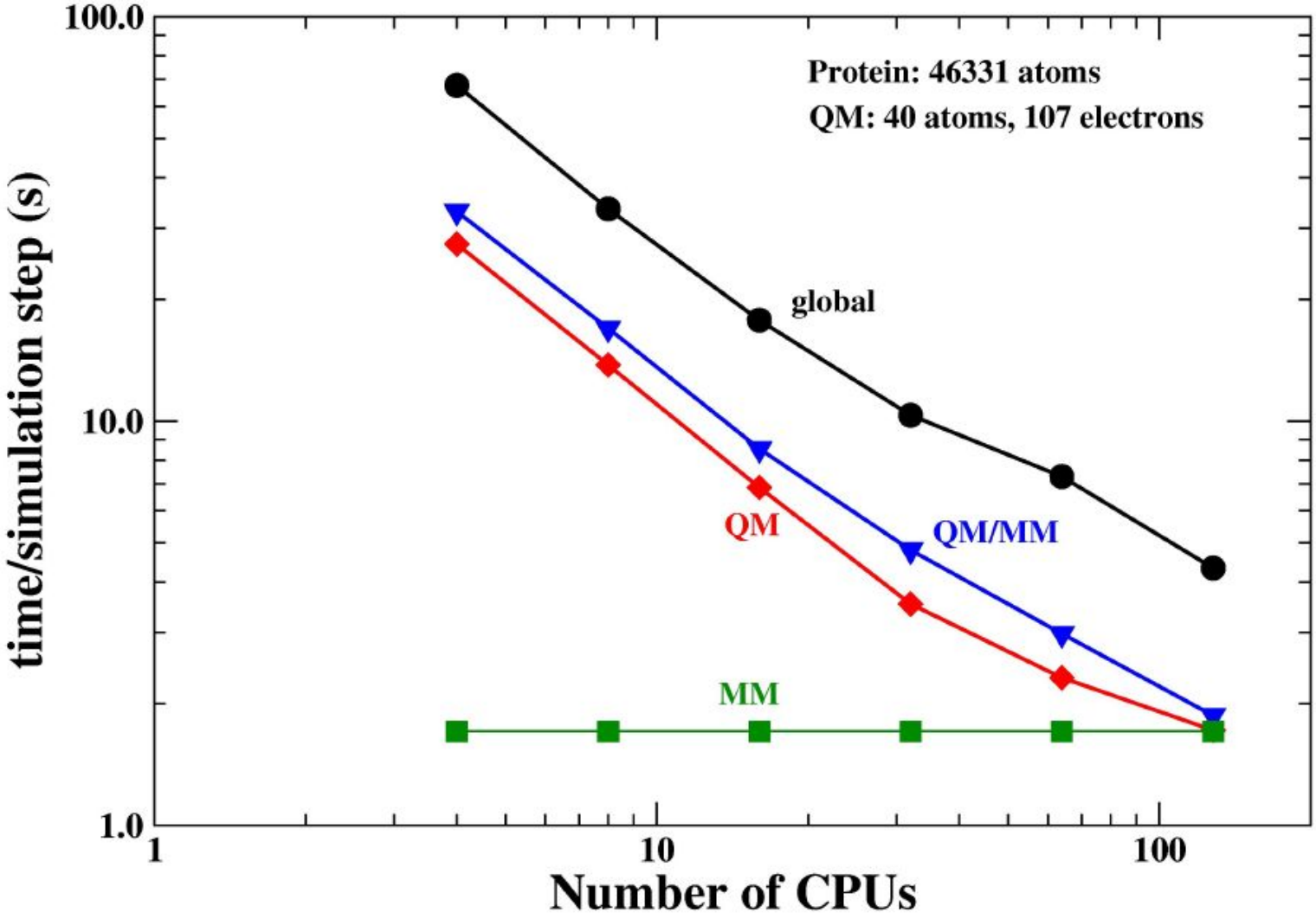
LSD-HCTH

E^{cut}=80 Ry 124 e^-

447866 PWs

Cell=22.3 x 22.3 x 22.3 Å^3

QM/MM test on PACS-CS: 1 simulation step vs. N. of CPUs



Performance on PACS-CS:

COMMUNICATION TASK	MESSAGE LENGTH	NUMBER OF CALLS
SEND/RECEIVE	17590 BYTES	253215
BROADCAST	155331 BYTES	75686
GLOBAL SUMMATION	42865 BYTES	108057
ALL TO ALL COMM	227932 BYTES	306103

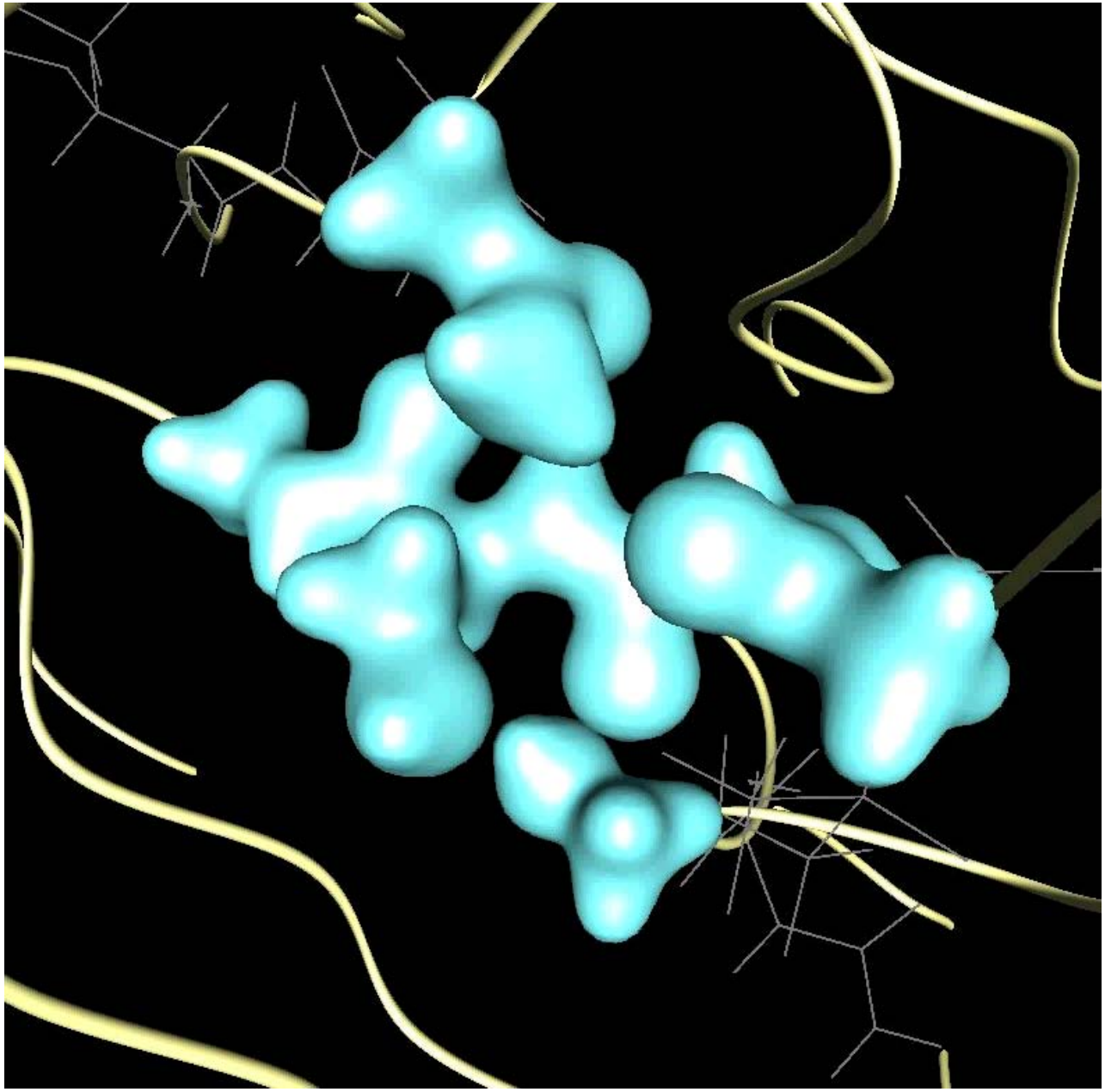
	PERFORMANCE	TOTAL TIME
SEND/RECEIVE	62.873 MB/S	70.844 s
BROADCAST	63.205 MB/S	186.003 s
GLOBAL SUMMATION	53.317 MB/S	694.990 s
ALL TO ALL COMM	23.947 MB/S	2913.515 s
SYNCHRONISATION		22.136 s

==========

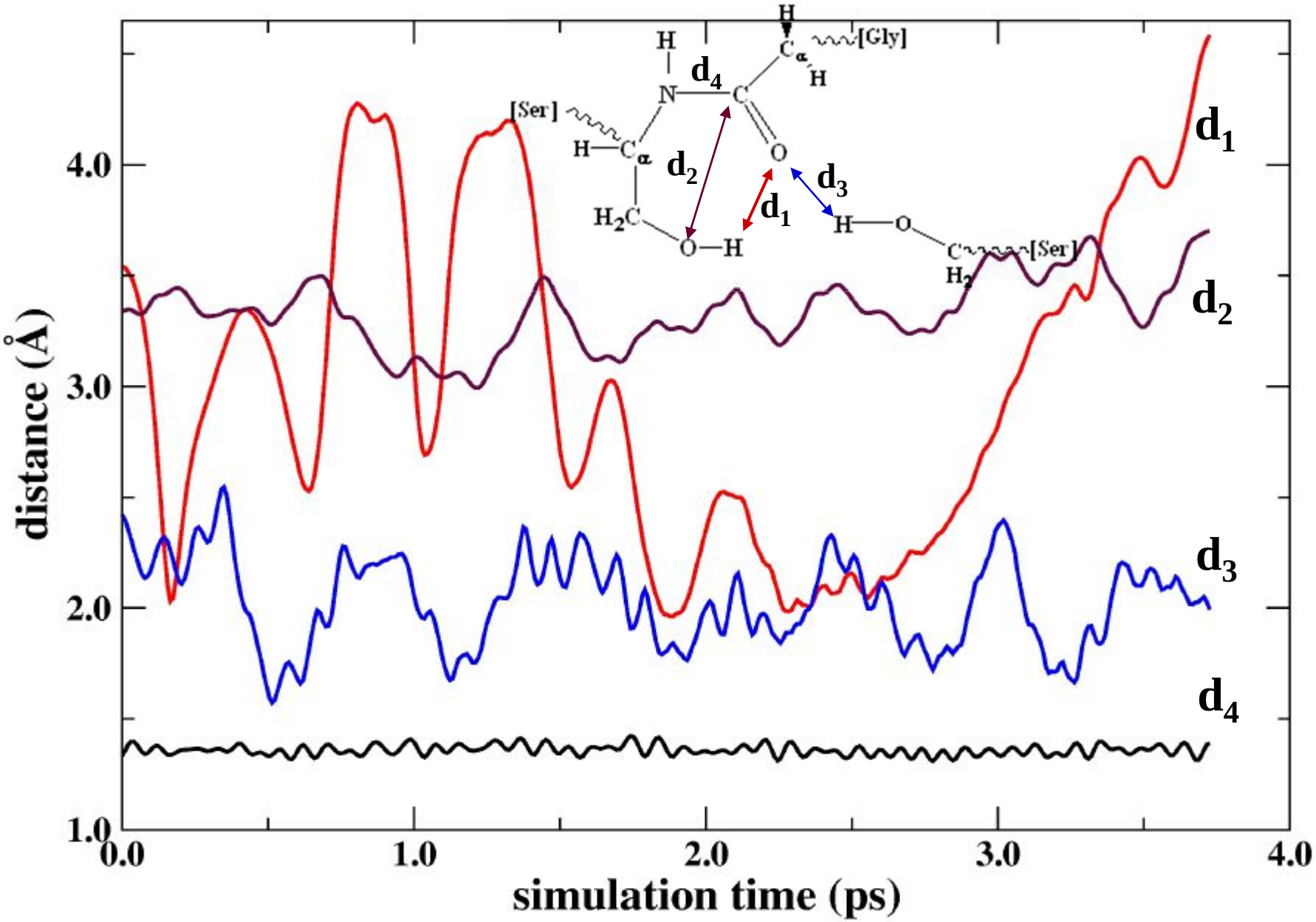
PEAK MEMORY 25221396 = 201.8 MBytes

CPMD equilibration

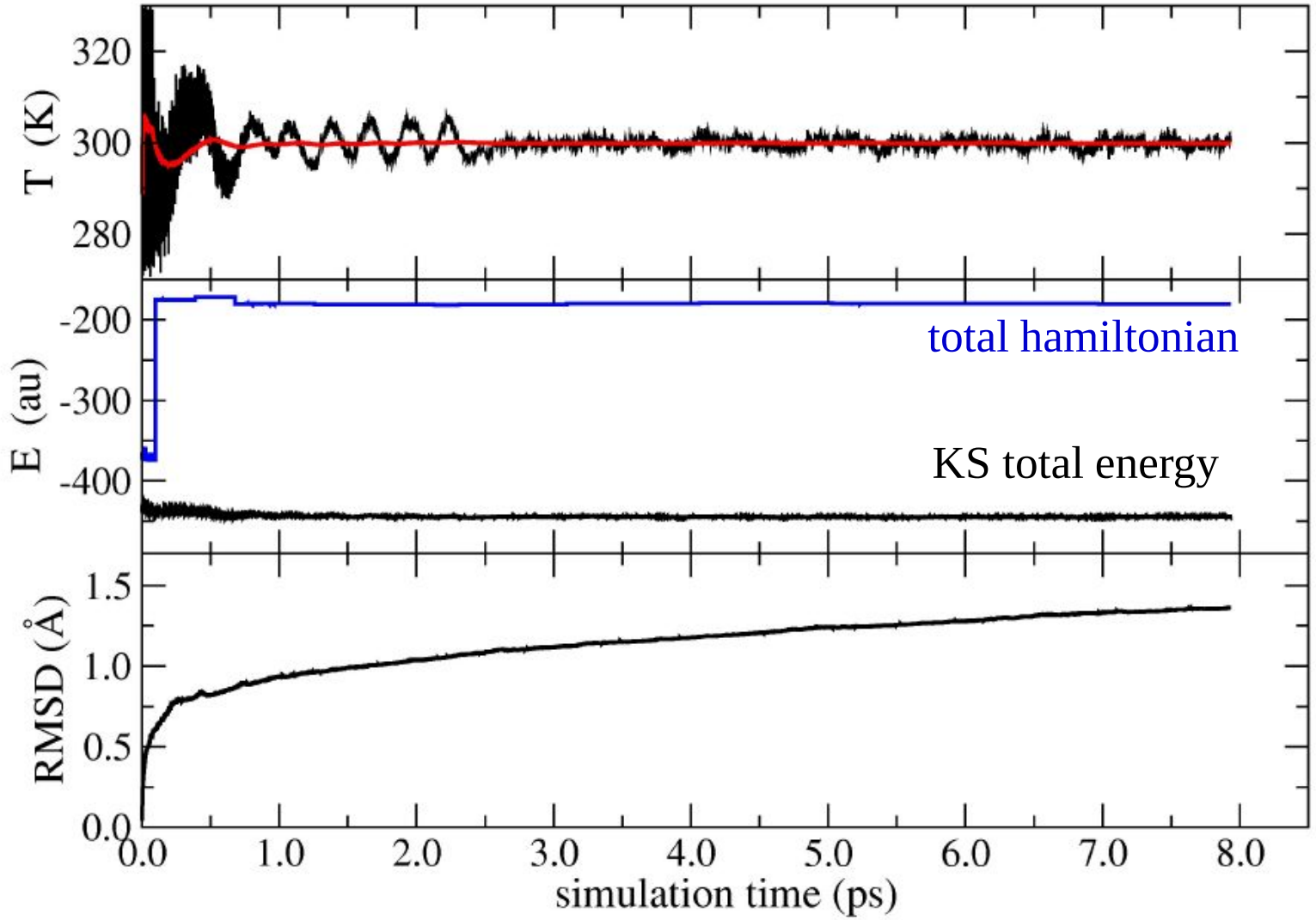
(after 40 ps of classical AMBER-MD)



CPMD equilibration (after 40 ps of classical AMBER-MD)



CPMD equilibration (after 40 ps of classical AMBER-MD)



Metadynamics: collective variables and related simulation parameters

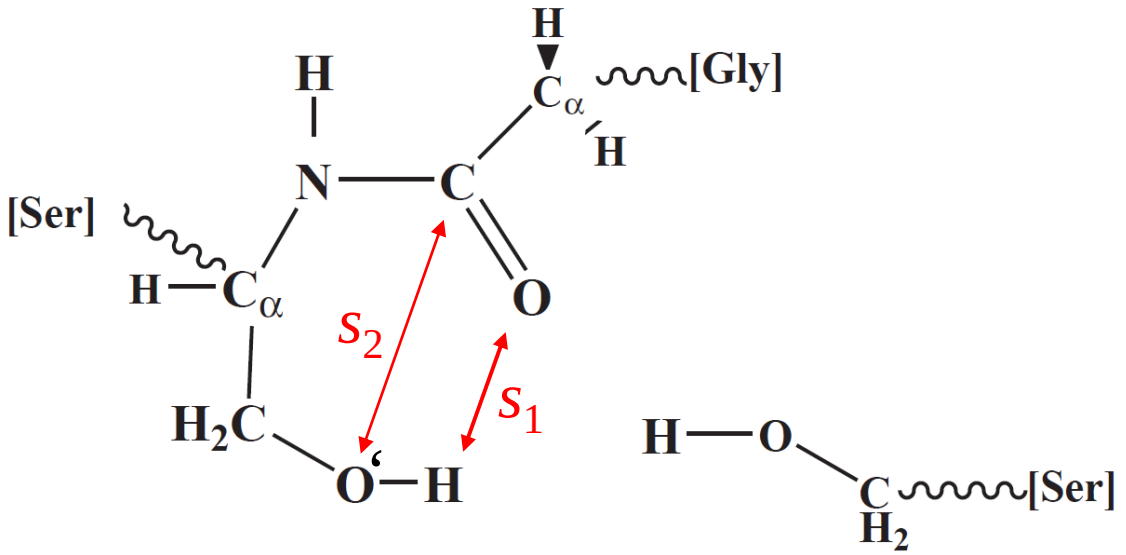
Simulation 1: s_1 =|O – H| & s_2 =|O' – C| (forming two new bonds)

M_α= 20.0 a.u.

k_α= 0.24 a.u.

max{W_i}=0.25 kcal/mol

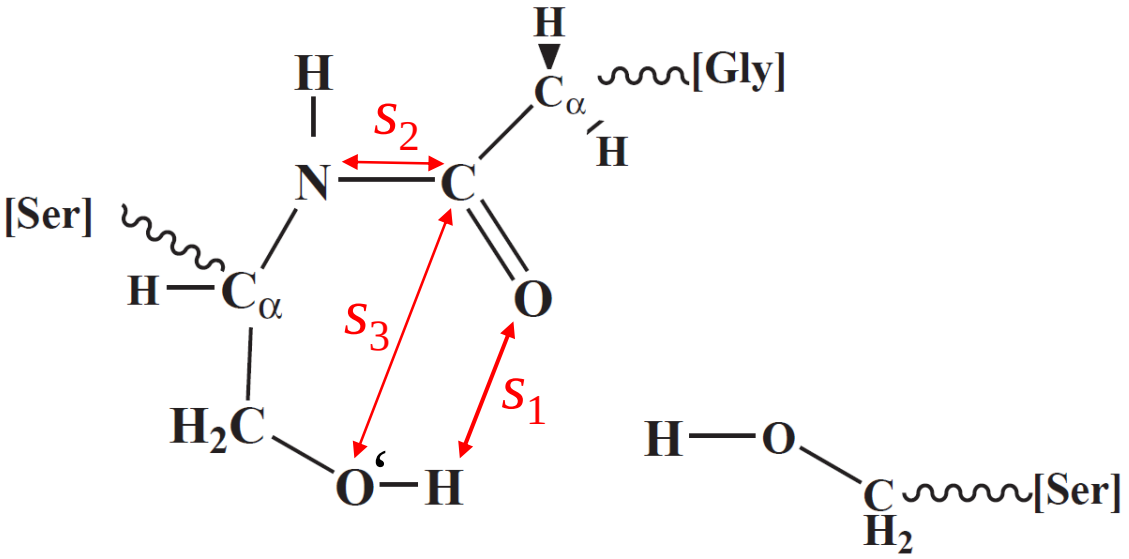
$\Delta s^{\perp} = 0.12$



Metadynamics: collective variables and related simulation parameters

Simulation 2: s_1 =|O– H| , s_2 =|N – C| & s_3 = |O' – C |
(breaking of the N-C bond included)

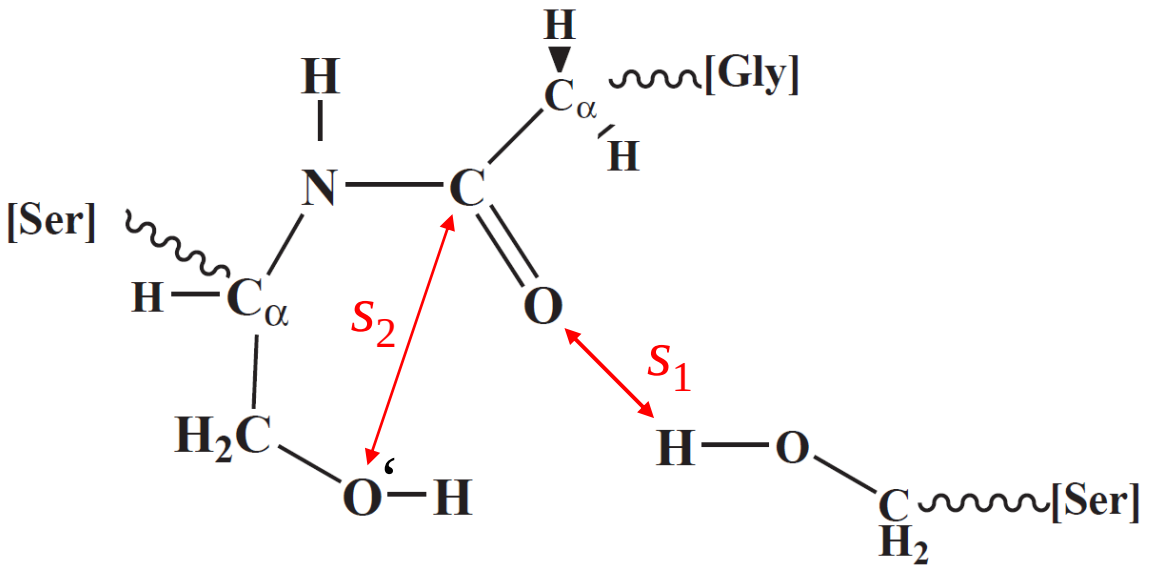
M_α= 20.0 a.u.
k_α= 0.24 a.u.
max{W_i}=0.25 kcal/mol
$\Delta s^{\perp} = 0.12$



Metadynamics: collective variables and related simulation parameters

Simulation 3: s_1 =|O – H”| & s_2 =|O’ – C| (forming two new bonds)

M_α= 20.0 a.u.
k_α= 0.24 a.u.
max{W_i}=0.25 kcal/mol
$\Delta s^\perp = 0.12$



Purpose: verify whether or not the close C-O-H, forming a H-bond with O, can participate to the splicing reaction

Conclusions (so far):

- OH groups of the catalytic site can break and reform H-bonds easily
- Other moieties are more rigid and do not show significant conformational changes on ns (MM) and ps (QM/MM) time scales
- Reactive complex obtained in the equilibration stage
- Accurate analysis of the electronic structure of the protein splicing site with explicit solvent

Related recent publications:

M. Boero, T. Ikeda, E. Ito and K. Terakura, *J. Am. Chem. Soc*. **128**, 16798 (2006)
T. Ikeda, M. Boero and K. Terakura, *J. Chem. Phys.* **127**, 074503 (2007)
K. Kamiya, M. Boero, M. Tateno, K. Shiraishi and A. Oshiyama, *J. Am. Chem. Soc*. **129**, 9663 (2007)
M. Boero, *J. Phys. Chem. A* **111**, 12248 (2007)
Y. Komeiji, T. Ishida, D.G. Fedorov, and K. Kitaura, *J. Comput. Chem*. **28**, 1750 (2007)