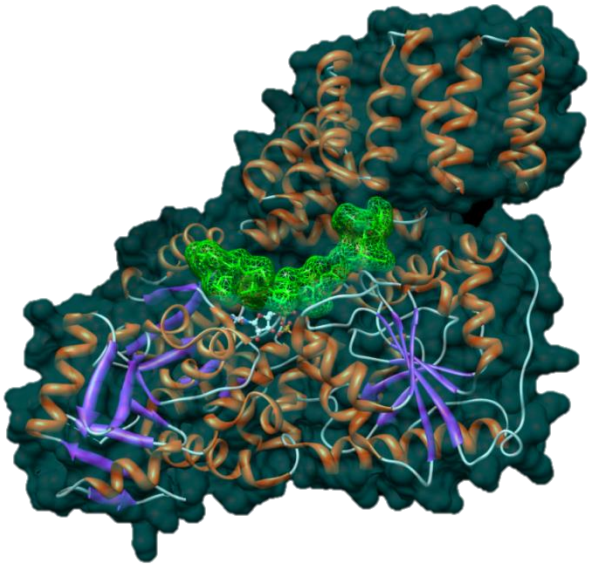
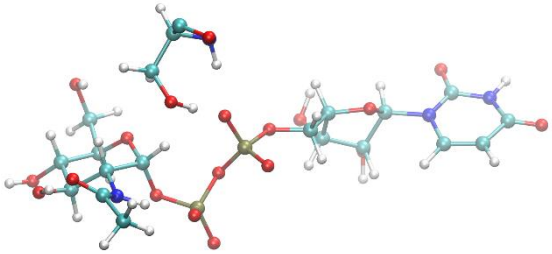
# Using the IT4I infrastructure on the enzymatic reactions studies

Stanislav Kozmon

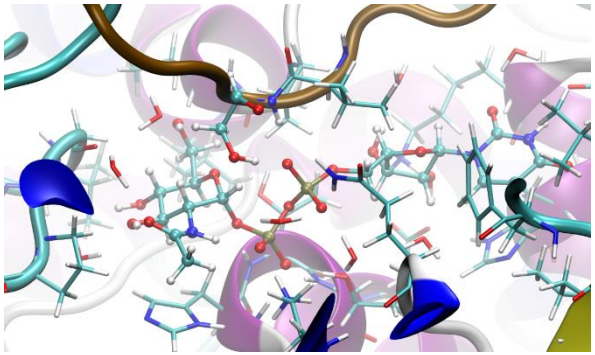


# *Molecular modelling of the biological systems*

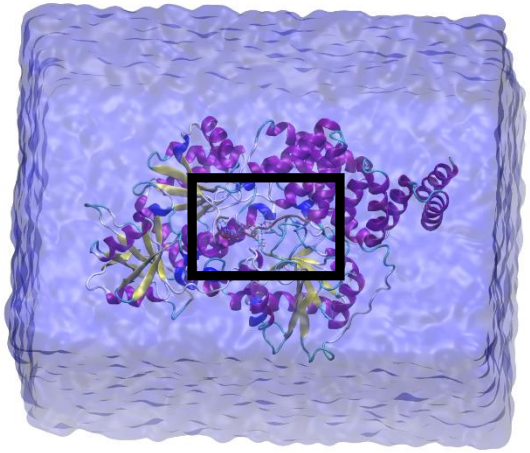
- Chemical reactions in living organisms are catalyzed by enzymes
- Enzymes are large biomolecules containing thousands of atoms
- Modelling of the enzymatic reaction:
  - Small model containing dozens of atoms treated by the QM methods
  - Model of the full enzyme treated by the hybrid QM/MM methods



Model containing only the reaction species



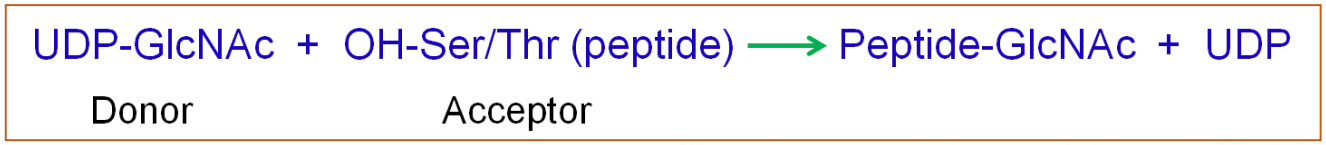
Model containing the substrates and a part of the enzyme

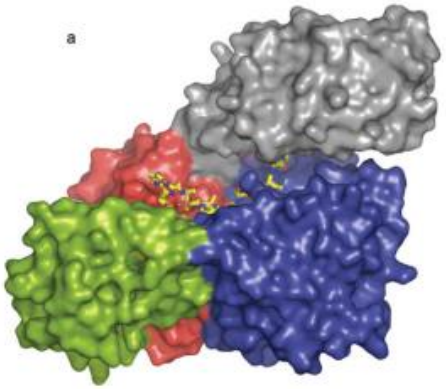


Model of the enzyme in explicit water box

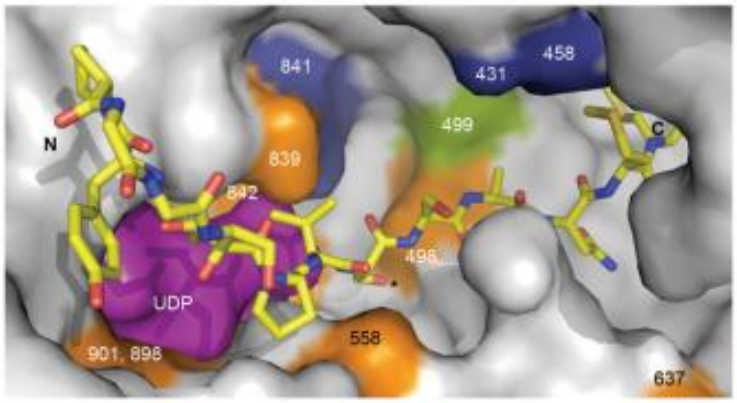
# *Introduction - OGT glycosyltransferase*

- Uridine diphospho-*N*-acetylglucosamine: polypeptide β-N-acetylaminyltransferase; PDB ID: 3PE4
- Enzymatic transfer of N-acetylglucosamine molecule on Ser/Thr residue of protein
- Inverting glycosyltransferase of the GT-B family
- Post-translational modification: first reported in 1984




**OGT Crystal Structure** (Walker et al, 2011)




**OGT Catalytic Site**

# *Introduction - OGT glycosyltransferase*

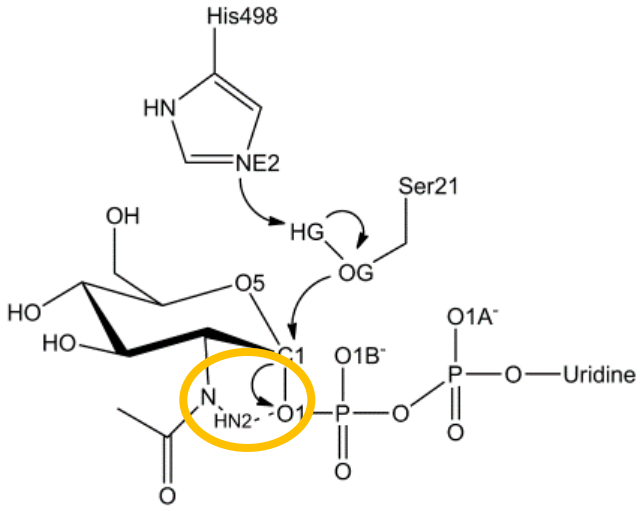
OGT biological function: Nutrient and stress sensor (cycling dynamics comparable to that of protein-phosphorylation)

OGT function abnormalities: insulin resistance, diabetic complications, neurodegenerative disorders and cancer

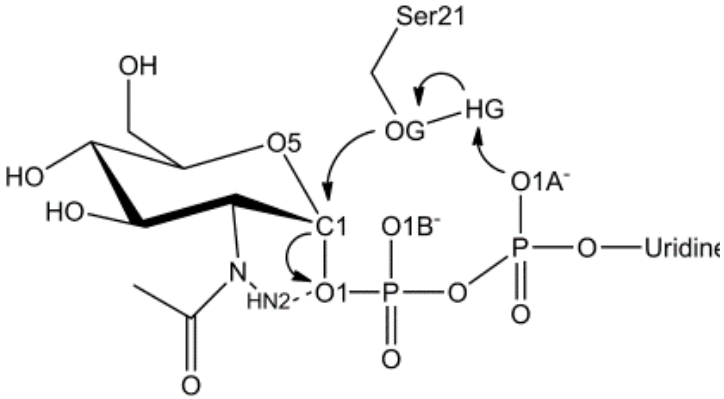
OGT as a promising drug target: The TS state analogues are the best inhibitors, however several reaction mechanisms were proposed

# *Different Mechanisms Proposed for OGT*

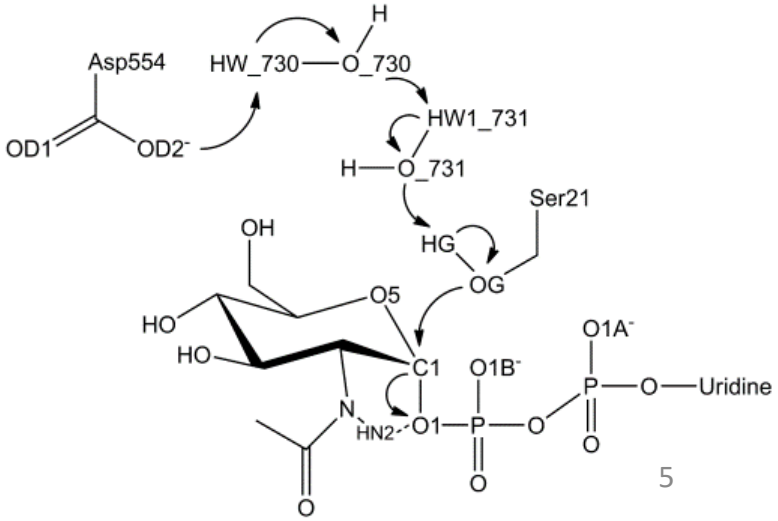
1. **His498 as catalytic base ($M_{His}$)**
   **(Lazarus et al. 2011**
   **Tvaroska et al. 2012)**




2. **α-phosphate as base ($M_{PO4}$)**
   **(Schimpl et al. 2012)**




3. **Water molecule for shunting proton to ASP554 ($M_{Asp}$)**
   **(Lazarus et al. 2012)**

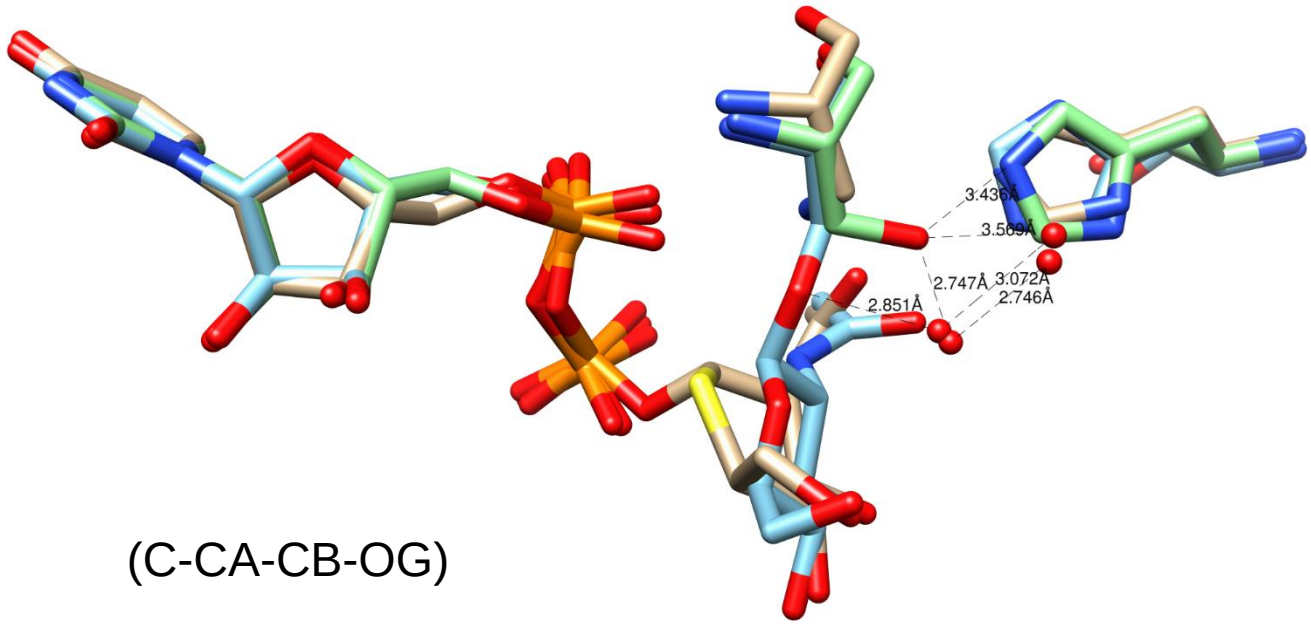


# *Structure of the Substrates in the Crystal Structures*

3TAX -> hOGT4.5, Casein Kinase II subunit alpha (Ser21), UDP; Lazarus2011; **1.88 Å**

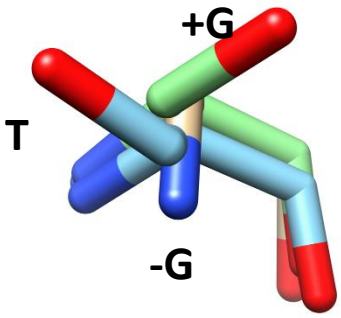
4AY6 -> hOGT(TRP fragment and CD), TGF-BETA-ACTIVATED Kinase 1 and MAP3K7-BINDING PROTEIN 1 (aminoAla1395), UDP-5S-GlcNAc; Shimpl2012; **3.3 Å**

4GYW -> hOGT4.5, Casein Kinase II subunit alpha (Ser21), UDP, Ser21-GlcNAc; Lazarus2012; **1.7 Å**



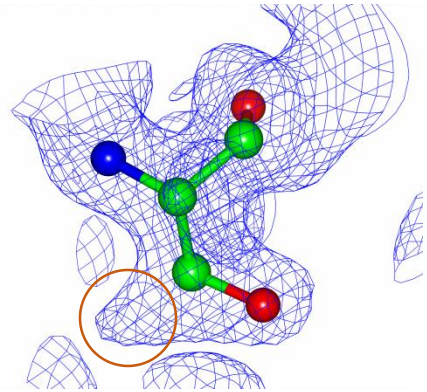

(C-CA-CB-OG)

# *Acceptor Serine Side Chain Conformations*



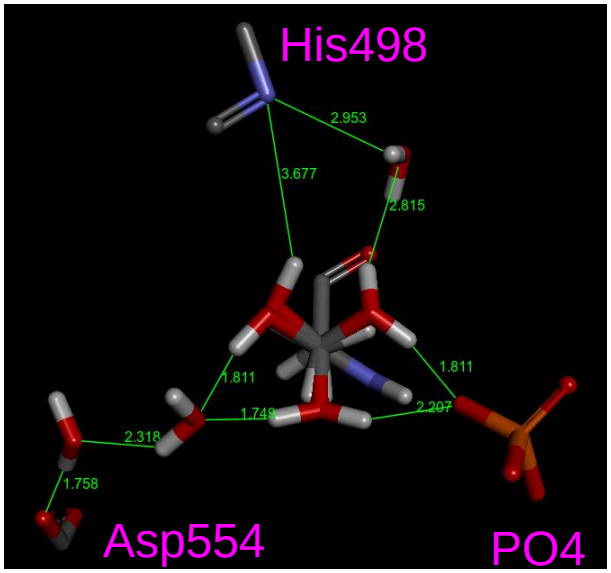

**3TAX** (**1.88 Å**) -> +g (57°)
**4AY6** (**3.30 Å**) -> -g (-65°)
**4GYW** (**1.70 Å**) -> t (169°)

3TAX x-ray electron density around the Ser21



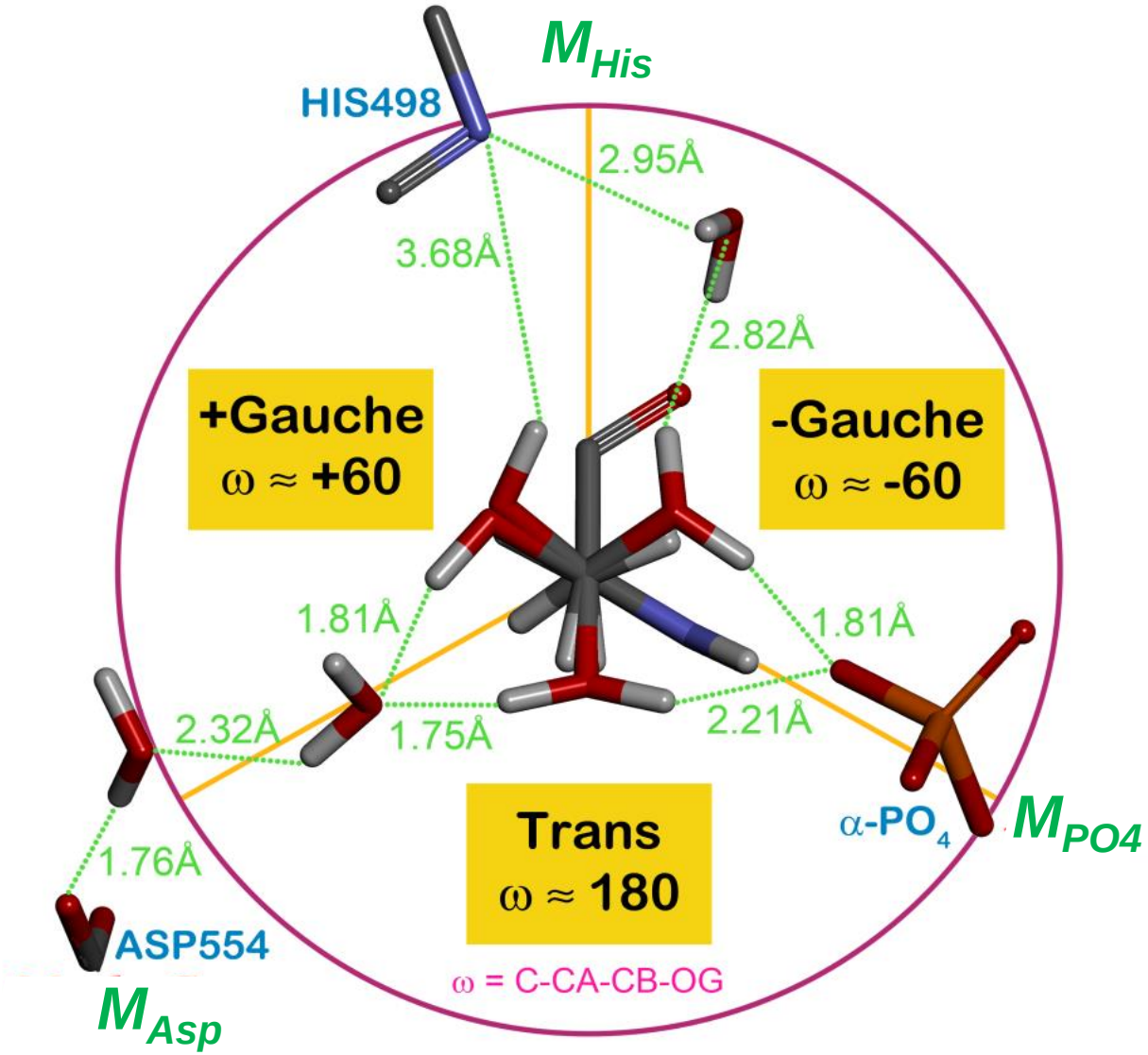
**Undescribed density for the *trans* of Ser21**

**Presumed gauche:*trans* occupancy between 2:1 and 3:1**




**Ser21 rotation: Access to Different Proton Acceptor**

# *Serine Side Chain Conformations for Diverse Mechanisms*

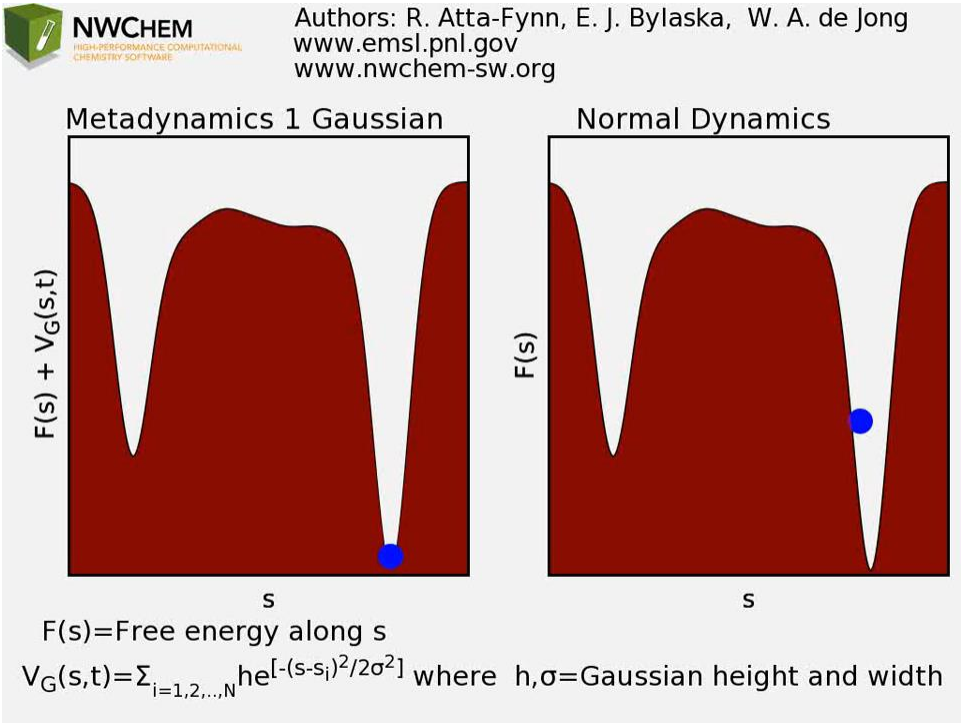


# *Computational Methodology*

- Hybrid QM/MM *ab initio* MD using CPMD/GROMOS
- Fully solvated system
- QM part treated by DFT PBE functional with Trouiller-Martins pseudopotenitials
- MM part treated by AMBER99SB force field
- Free energy reaction path optimization using the String Method on selected collective variables
- Exploring the Free Energy Surface using Metadynamics

# *Reaction Mechanism Studies - METADYNAMICS*

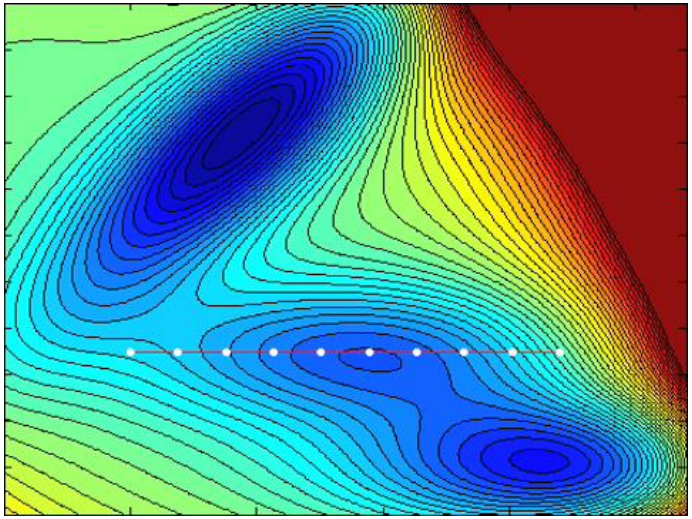
- the type of accelerated molecular dynamics, where an artificial potential is added to the site already visited to allow the molecule to explore places with higher energy
- this potential can also be applied to selected collective variables (distance, angle, etc.) for a given reaction path which helps to overcome the reaction barrier
- the difficulty of the method grows exponentially with the number of collective variables

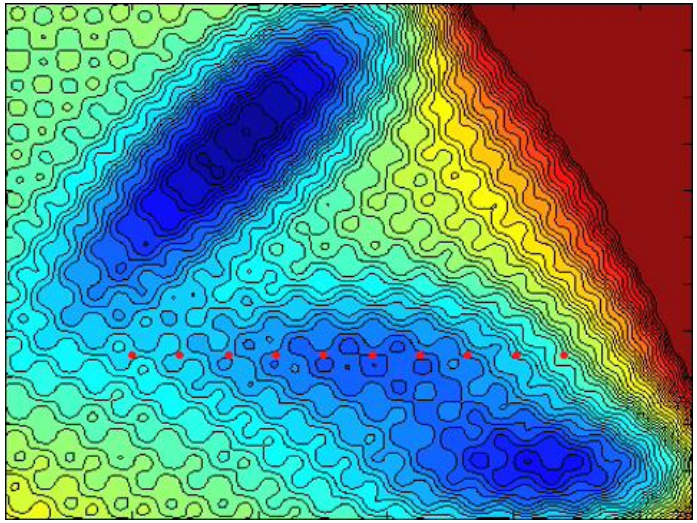



NWChem, https://www.youtube.com/watch?v=CtIrLkx6aNo

# *Reaction Mechanism Studies - String Method*

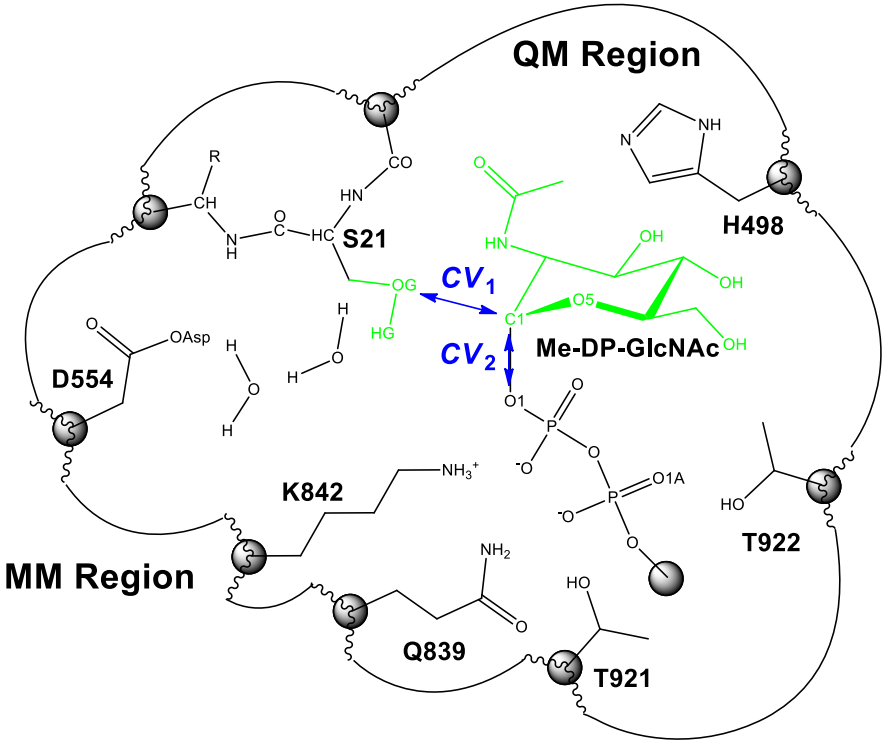
- minimum free energy reaction path optimization
- the reaction path is divided into points (beads) with defined values of the collective variable (distance, angle, dihedral angle, etc.)
- at each point there is a molecular dynamics running where the value of the collective variable is held by means of restraints
- its difficulty does not depend on the number of collective variables and is linearly dependent on the number of beads





Eric Vanden-Eijnden, http://cims.nyu.edu/~eve2/string.htm

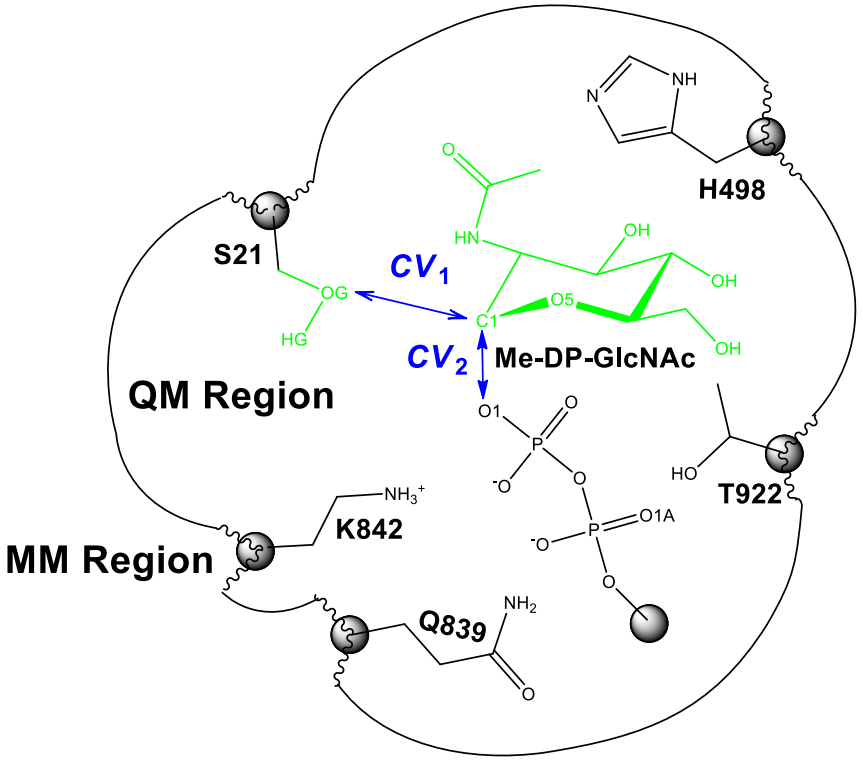
# *Partitioning of the QM and MM Zone*




## Large QM Zone

Atoms: 146 QM
Box Size: 19.7 x 27.4 x 27.1 Å
Time: 19 s/step on 80 CPUs

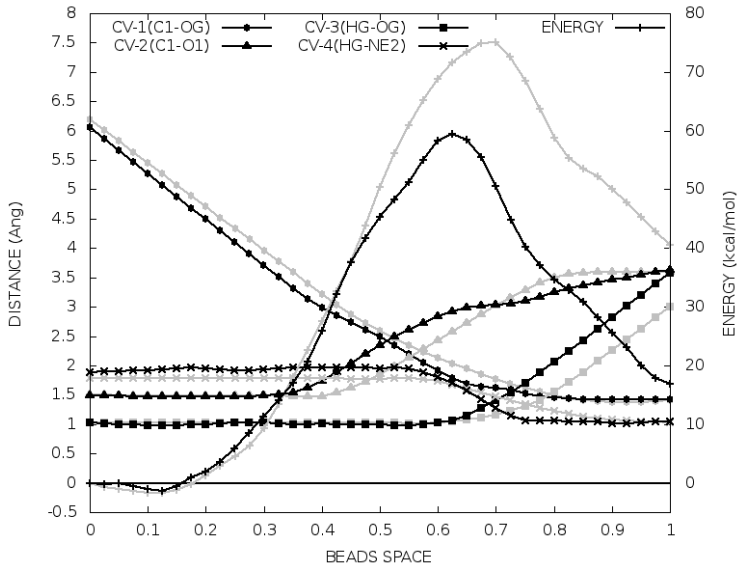
Atoms: 106 QM atoms
Box Size: 19.7 x 20.2 x 27.1 Å
Time: 11 s/step on 64 CPUs



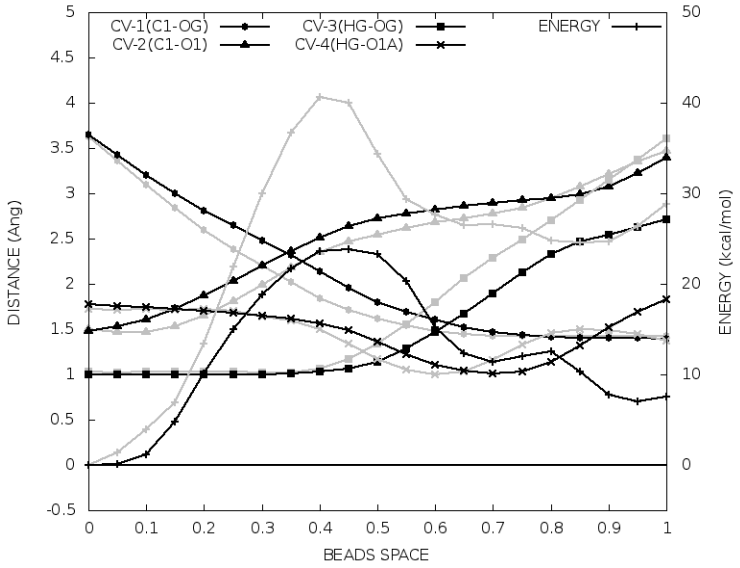

## Small QM Zone

# Collective Variable and Free Energy Evolution During Optimization

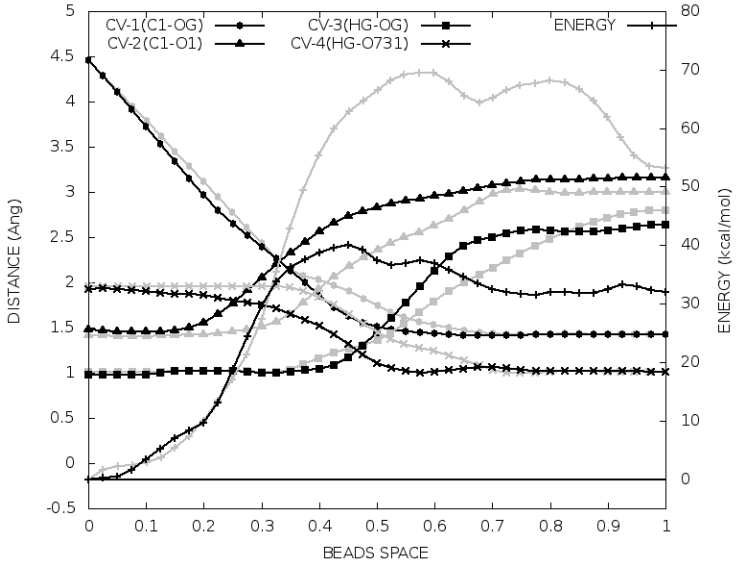
$M_{His}$




$M_{PO4}$



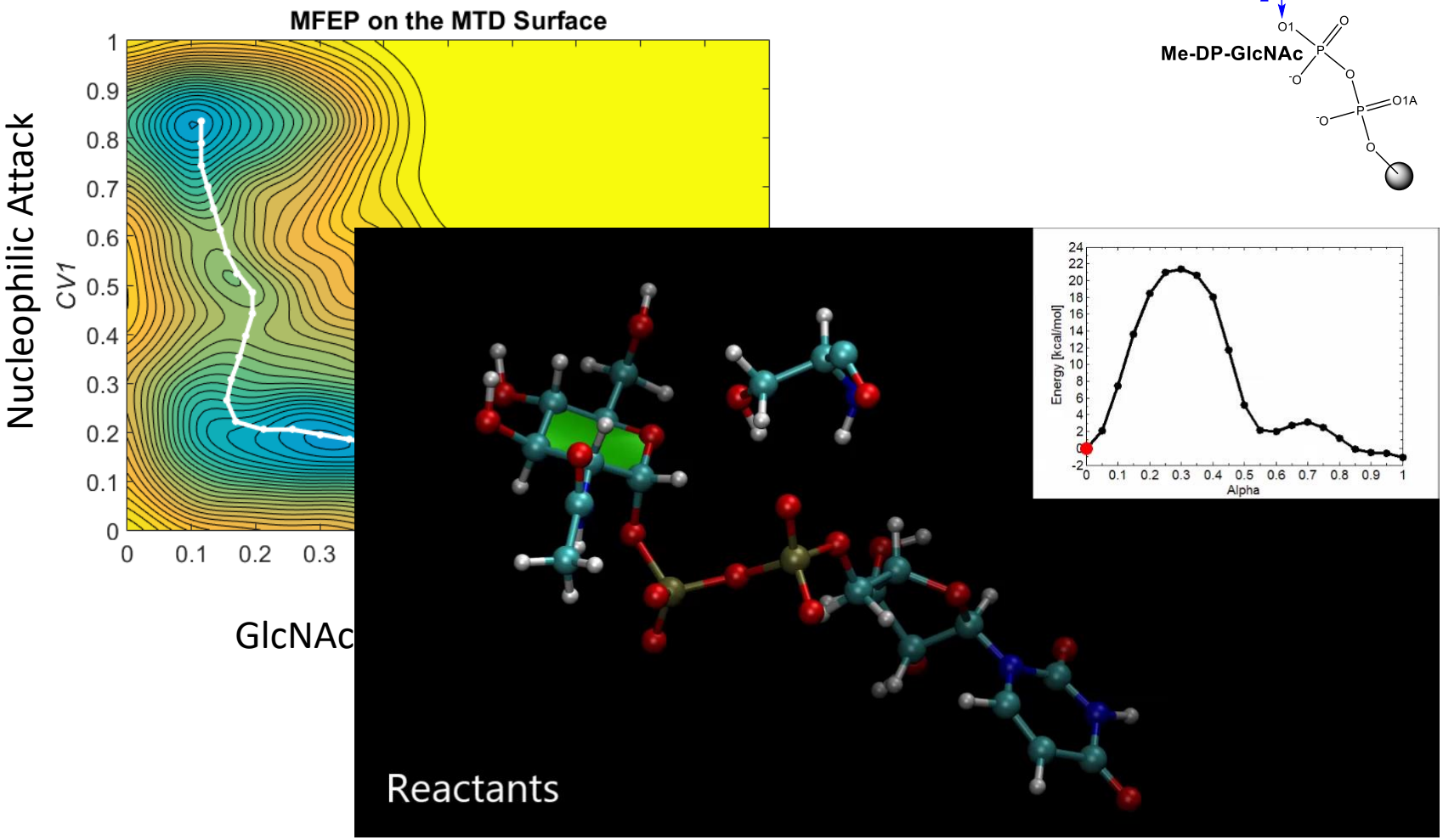

$M_{Asp}$




146 QM atoms statistics:
1 bead: 19 s / step on 80 CPUs
Steps: 2200 steps/iter
Iter: 37
Beads: 41
Overall CPU time: 1.4 M cpu/h

106 QM atoms statistics
1 bead: 11 s / step on 64 CPUs
Steps: 2200 steps/iter
Iter: 42
Beads: 21
Overall CPU time: 380 k cpu/h

# *Most Probable Reaction Path – $M_{PO4}$*



# *Summary*

- Used methodology was able to distinguish between proposed mechanisms
- Catalytic process involves nucleophilic attack, proton transfer and glycosidic bond formation, in the same order for all the mechanism and has slightly dissociative $S_N2$ character
- N-Acetyl group stabilize leaving phosphate group
- Free energy profile suggests $M_{PO4}$ as the most probable pathway having the TS barrier of ~24 kcal/mol

## *Aknowledgment*


Prof. Jaroslav Koča
Dr. Igor Tvaroška
Manju Kumari
Petr Kulhánek
Jakub Štepán
Tomáš Trnka


## Financial support


SoMoPro

JIHOMORAVSKÉ CENTRUM PRO MEZINÁRODNÍ MOBILITU

MARIE CURIE ACTIONS

Jihomoravský kraj


## Computational support


IT4Innovations national supercomputing center

Slovak Academy of Sciences Programme | S A S P R O

CEITEC
Central European Institute of Technology
BRNO | CZECH REPUBLIC


Thank You for your kind attention !