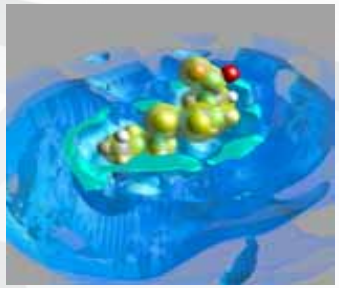
Nanoapplications molecular modeling

- Grid Enabling of Nano-Science Applications in NAREGI -

Mutsumi AOYAGI
Kyushu Univ.
aoyagi@cc.kyushu-u.ac.jp
Head of Grid-Enabling Team&
Sub-Leader of NAREGI(JAPAN)

Outline of my talk

1. Grid-Enabling through Mediator/GridMPI

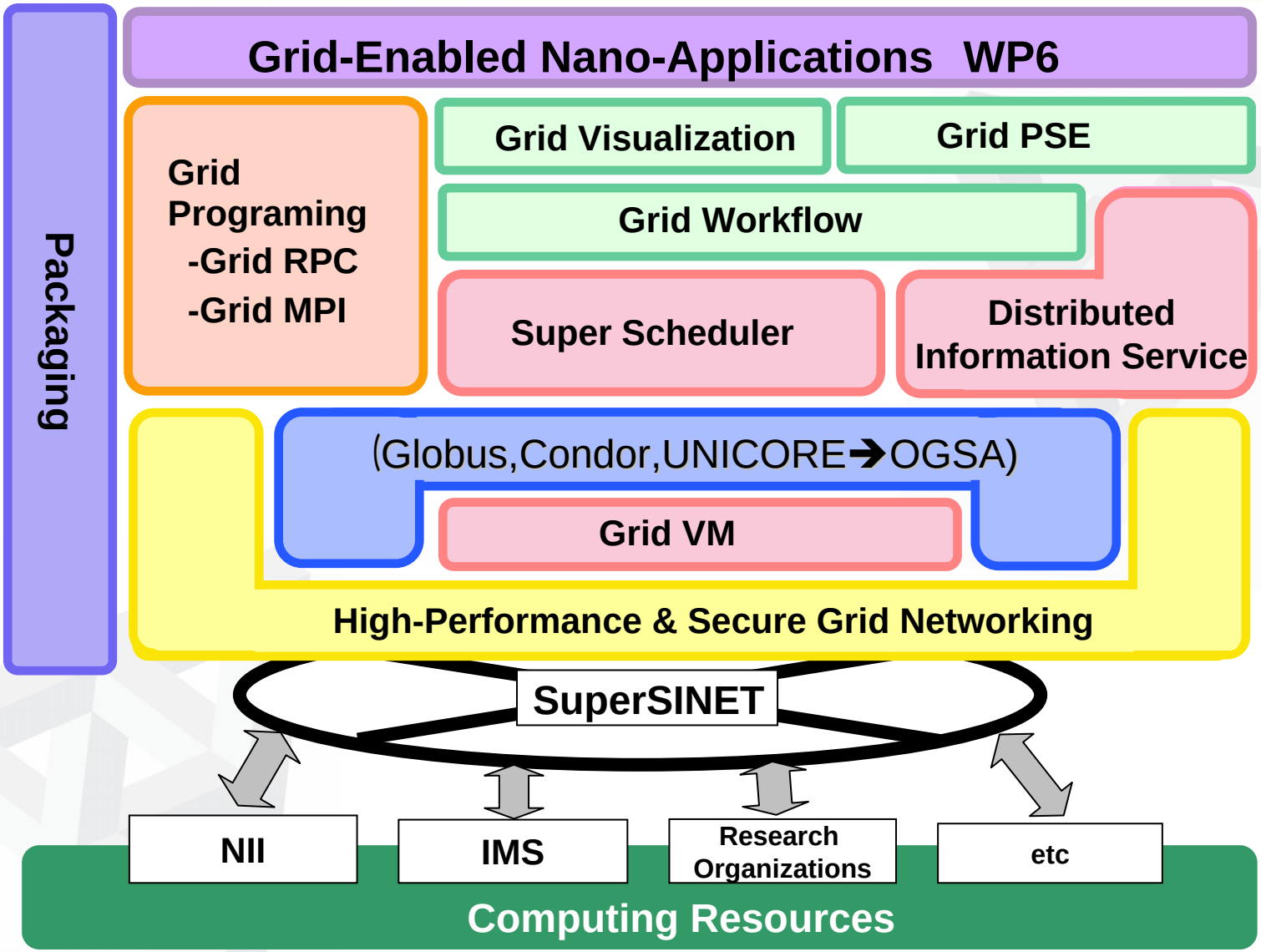


Application to Solvent interactions in hydrolysis of Lysozyme
MD-PB and RISM-FMO coupled simlations

2. NAREGI workflow tools with Loosely Coupled Application Components

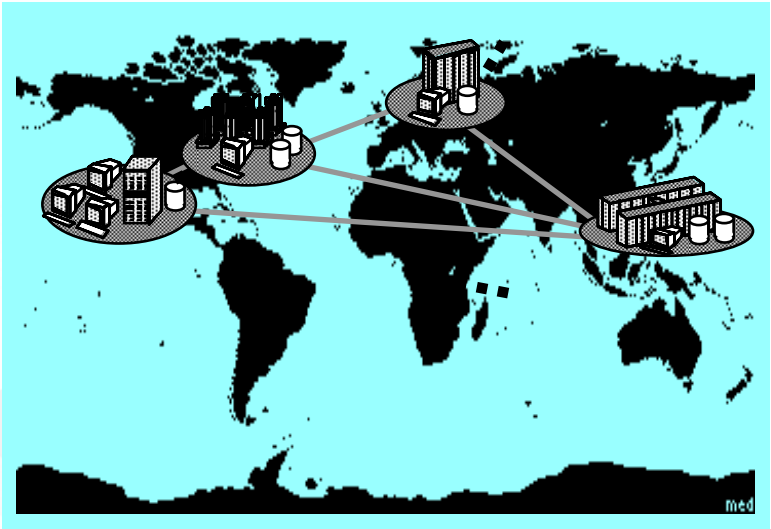
A Case study with GAMESS FMO and TINKER

NAREGI Software Stack

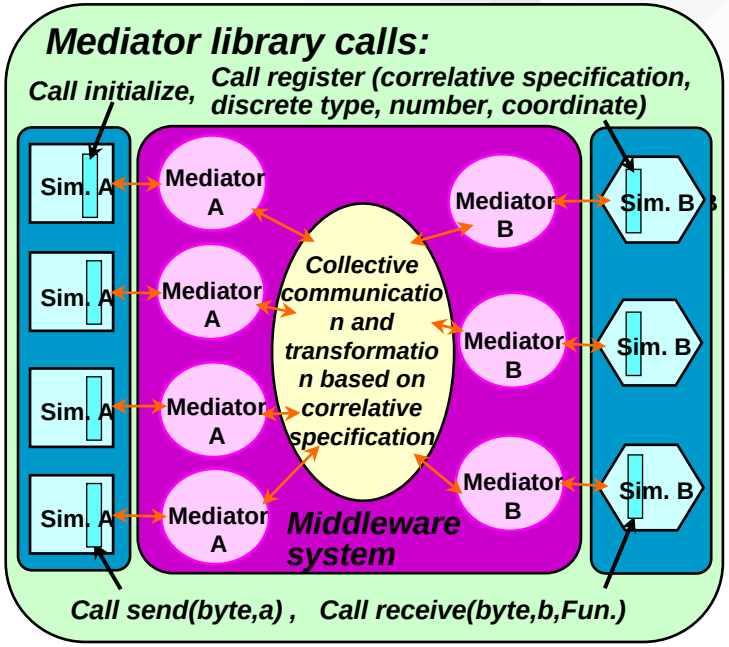


Mediator: Grid middleware for Coupled Simulations

A new grid middleware is developed which allows various kinds of Nano-application softwares to be coupled efficiently for solving multi-scale and multi-physics problems.

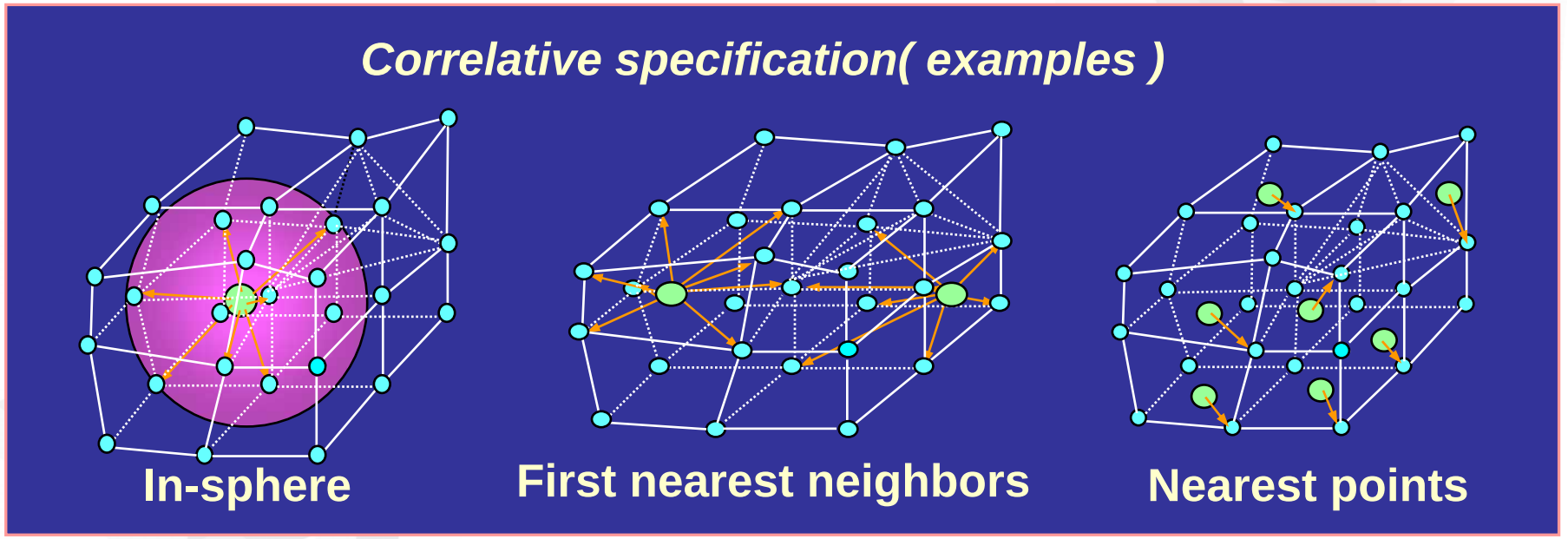


Respect Independency of each application prog.



Mediator: Grid middleware for Coupled Simulations

The mediator provides high-level transparency in data communication between different discretization methods associated with a model specific spatial and temporal scale based on our physical requirements.

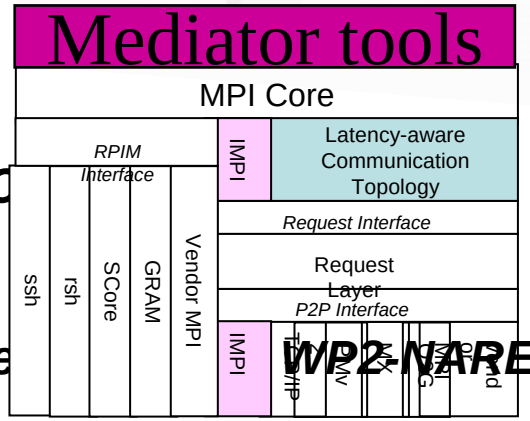


Other correlation specification can be defined as user-plug-in.

Mediator

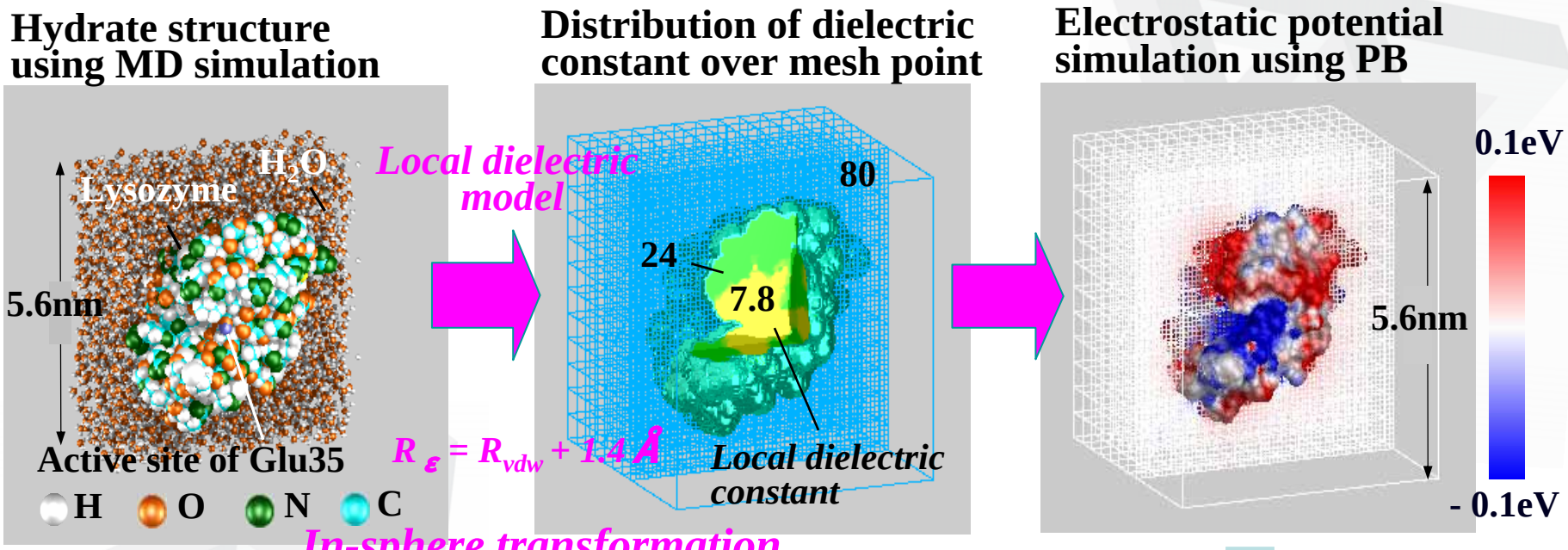
α-version released in march 2005

- **Discretization methods**
  - ***Particle method, Finite difference method, Finite element method***
- **Parallel programming style**
  - ***SPMD or MPMD in static or dynamic invoke***
- **High-level semantic transformation**
  - ***In-sphere, In-rectangle, 1st nearest neighbors, Nearest points***
- **Communication paradigm**
  - ***One-way, Variable, two-way co***
- **Interconnection**
  - ***MPICH-G2, MPICH, GridMPI(de WP2-NAREGI)***
- **Hardware architecture**
  - ***Linux clusters, AIX, Solaris, etc.***



Mediator based Molecular dynamics(MD) and Poisson-Boltzmann(PB) coupled simulation(1)

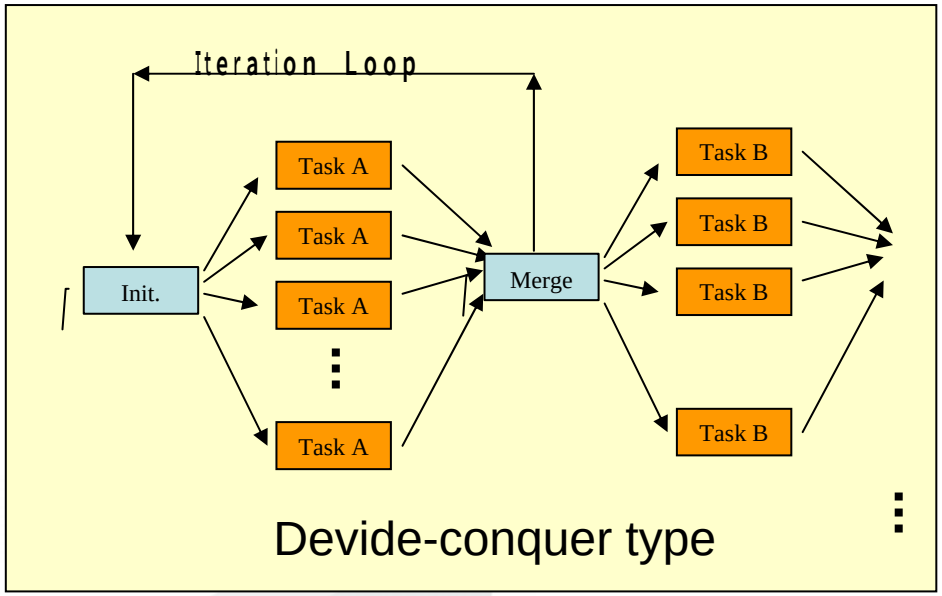
Method of Simulation

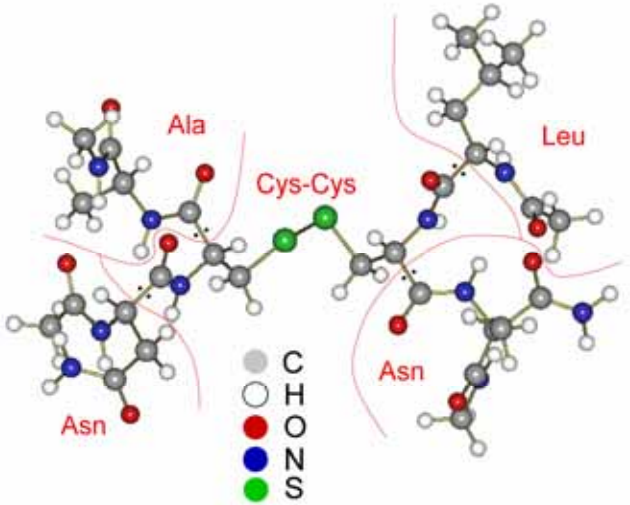


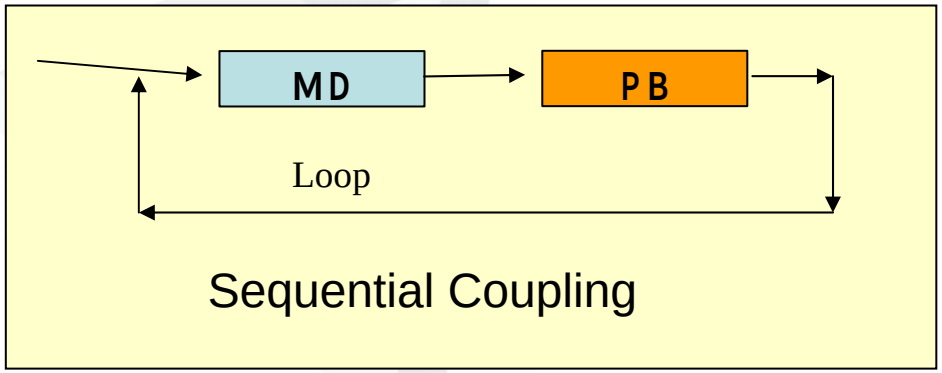
Roughly estimate "Active Site" of lysozyme and H-Dissociation energy at Glu35.

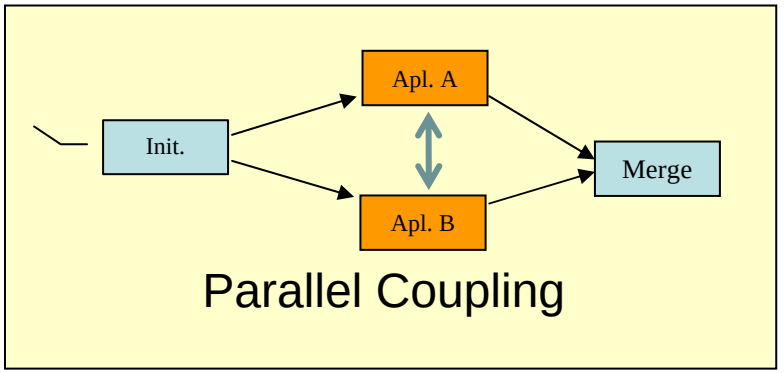
Example of Workflow

Modularize in consideration of granularity, communication overhead, latency awareness, portability, etc



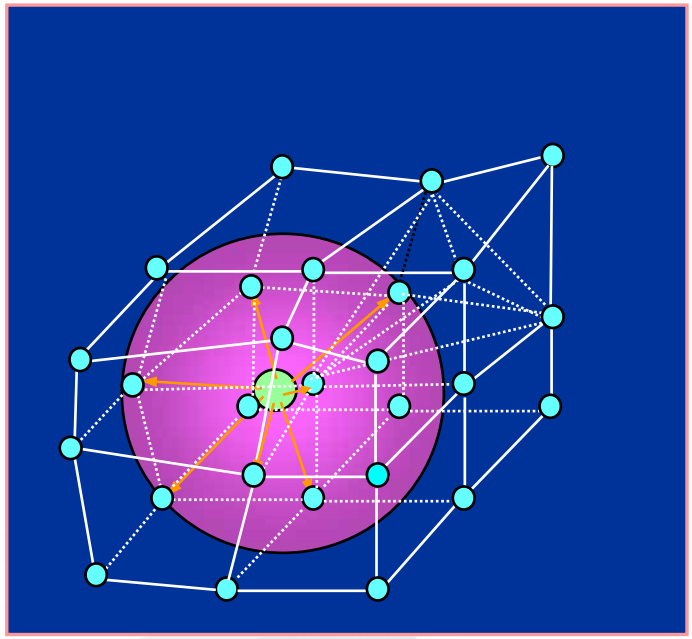






Efficient Data Transfer on Different Descretization models

- **Particle (MD) / Mesh (FDM)**



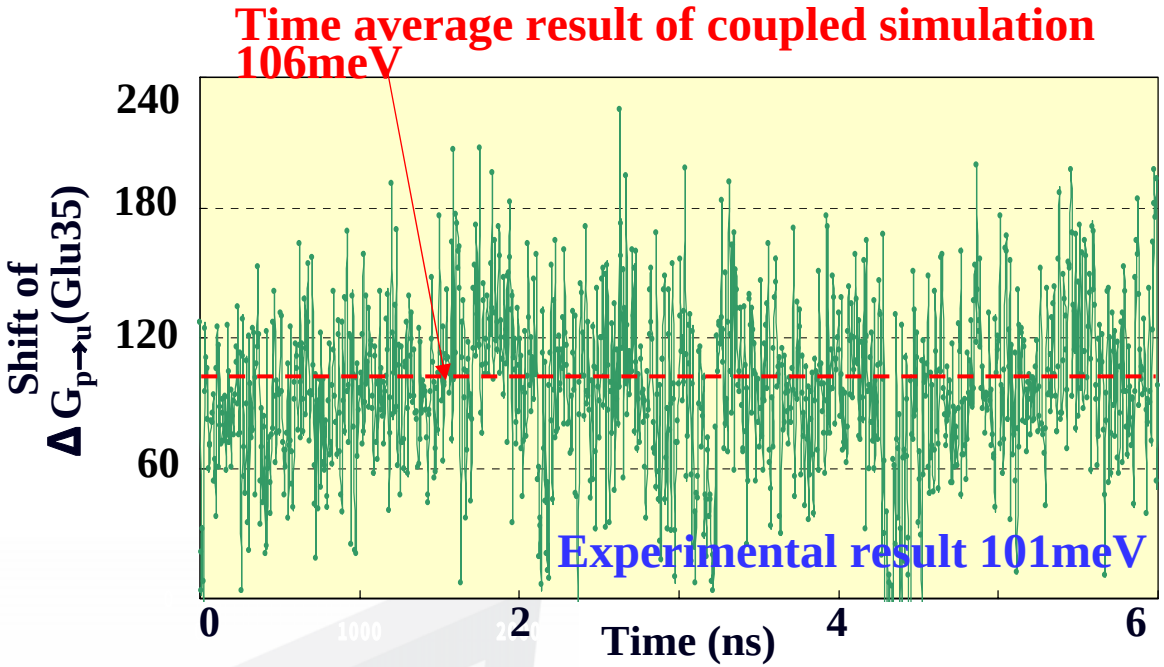
Mediator automatically generates the correlation relation over different descretization models:

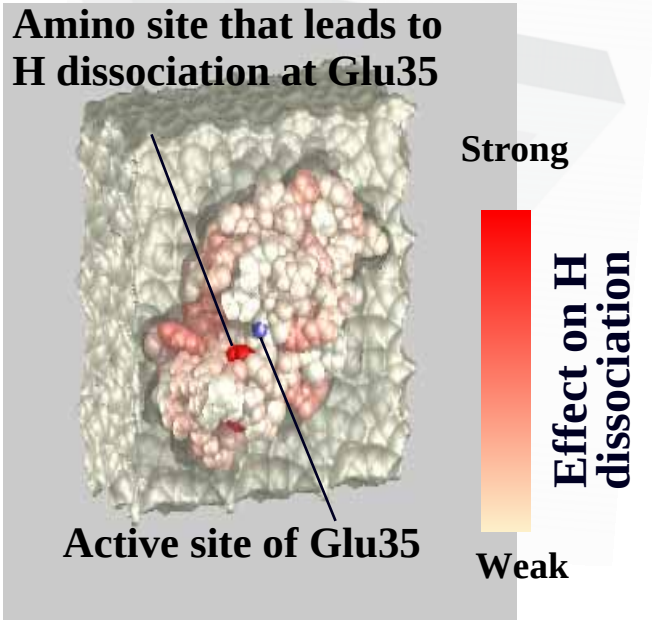
Dipole moment of particles ⇔Distribution of dielectric constant over mesh points

To keep transportabilities and independencies of APLs, Semantic transformation required in coupled simulations should be achieved in the third component as Mediator, not in the application componets themselves.

Mediator based Molecular dynamics(MD) and Poisson-Boltzmann(PB) coupled simulation(2)

Selected results



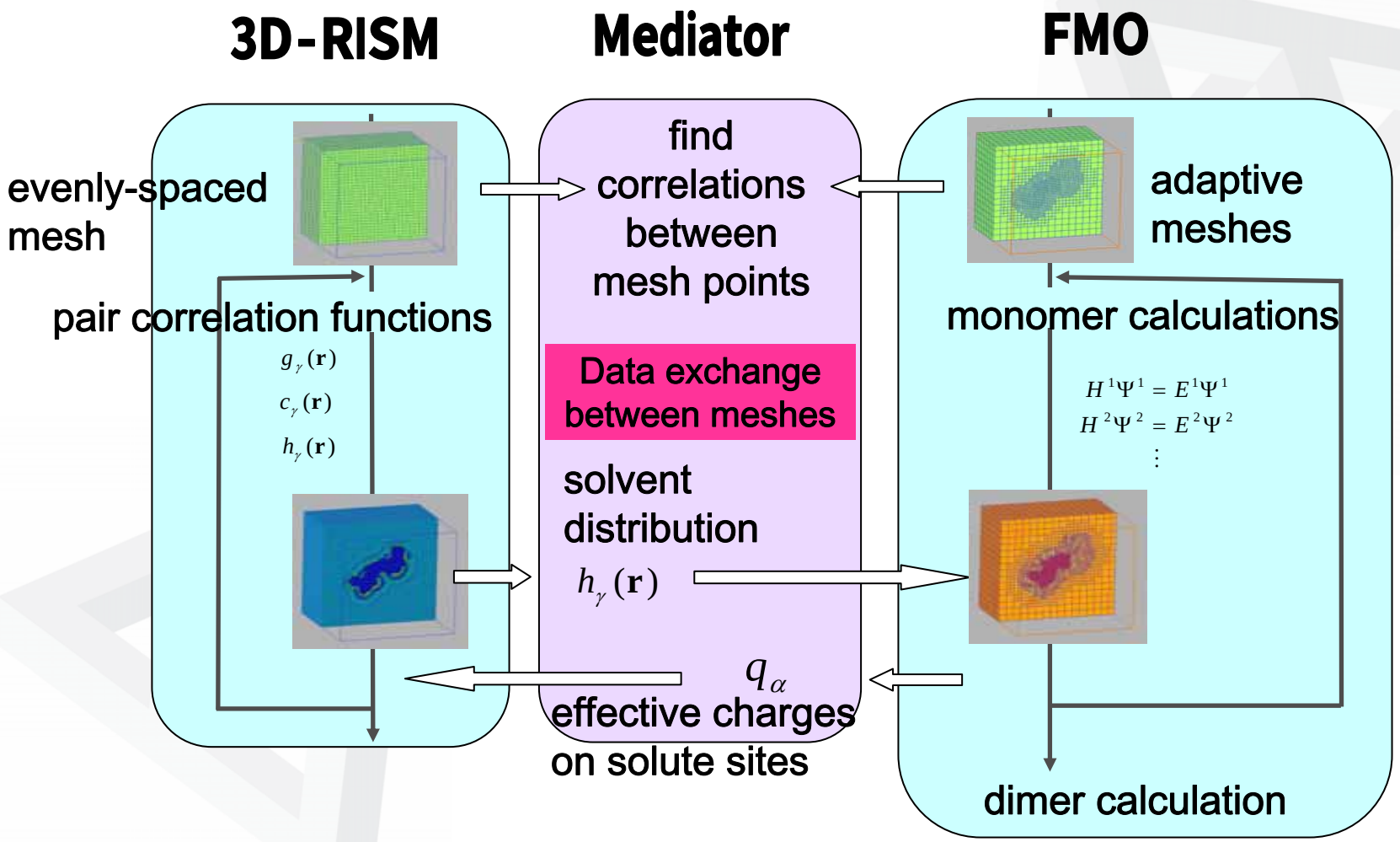


$$\Delta pKa = \Delta G_{p\to u}(protein) / 2.303\ k_B T - pKa(Glu)\ (=4.4)$$

pKa value(〜6) at COOH in Glu35
→ roughly 50 times larger probability
compared with amino sites around

2nd Example: RISM-FMO coupled Simulation

Outline of workflow:



Data Exchanges in RISM-FMO

Mesh structure of solvent

Molecular sites of solute

Mediator

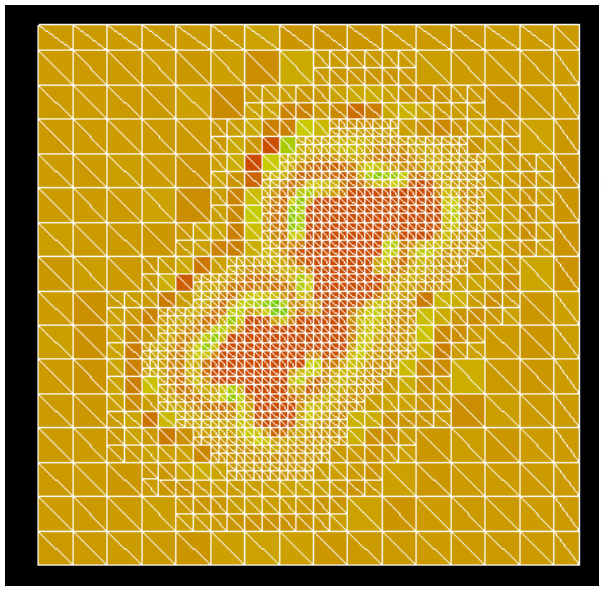
Mediator

Equally discretized mesh points

In-sphere correlation

Van der Waals and probe radius associated with molecular sites

Adaptive meshes for hydrate solvent



We respect Independency of developing each Application.

Mediator facilitate customization processes of original application program for coupled simulations.

Mediator treats with Data exchange on different descretization sheme of apl. A and apl. B.

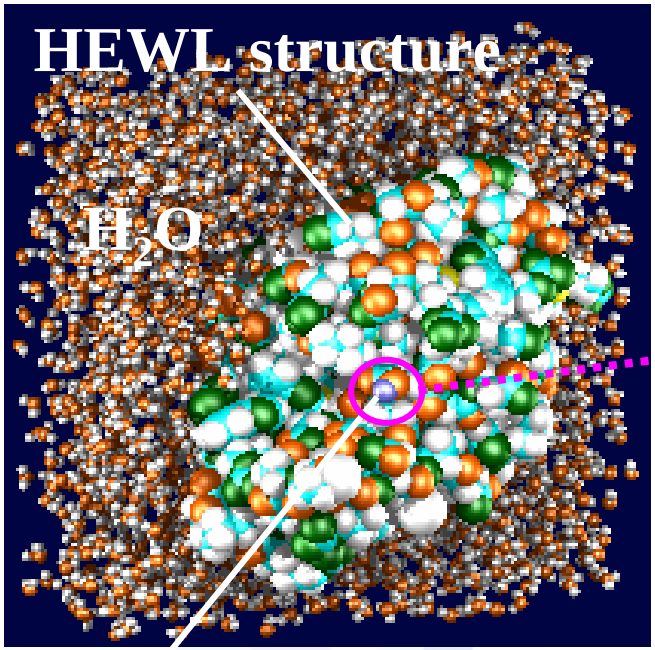
Summary of Coupled simulation model

Coupled simulation	RISM-FMO	MD-PB
Discretization method	FDM / Irregular point	Particle / FDM
Physical quantities to be transformed	Solvent charge density to charge on solvent atoms	Dielectric distribution, Charge on atoms to charge density
Correlation specification	In-rectangular	In-sphere
Transformation function	Weighted function conserving charge	Weighted function equalizing electric field
Programming style	Sequential / Master-Worker	Master-Worker / Sequential
Communication paradigm	Two-way iterative communication	One-way, Variable communication
Interconnection	GridMPI, MPICH-G2, GridFTP, MPICH, Score	MPICH, MPI2, Stampi
Server machines	Hitachi SR8000, AIX, Linux, Alpha clusters	Hitachi SR8000, SR2201, DEC, Sun clusters

Solvent interactions of hydrolysis in Lysozyme

Selected results of 2nd example

A catalytic mechanism of hydrolysis in Lysozyme is analyzed by RISM-FMO, in which proton transfer leads to bacteriolysis in peptidoglycan of bacterial cell wall.

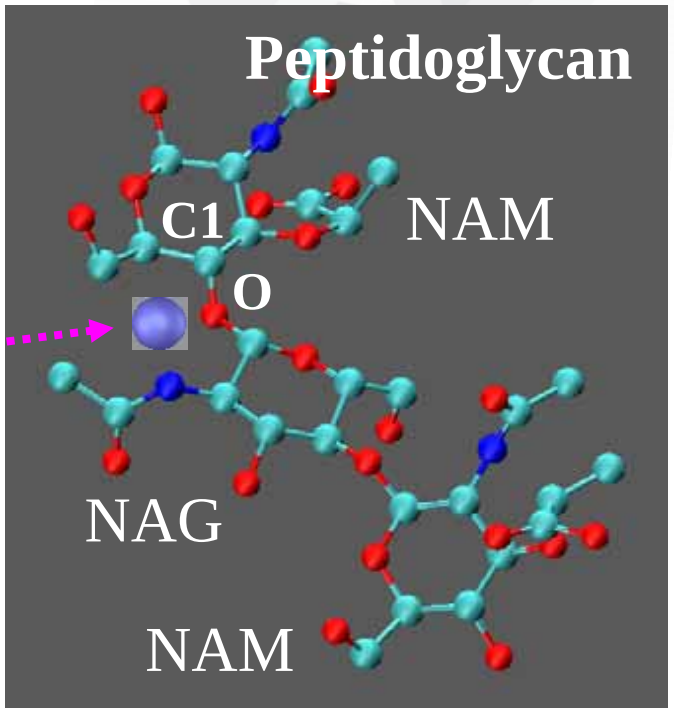


Active site of Glu35

Bacteriolysis by hydrolysis due to H dissociation

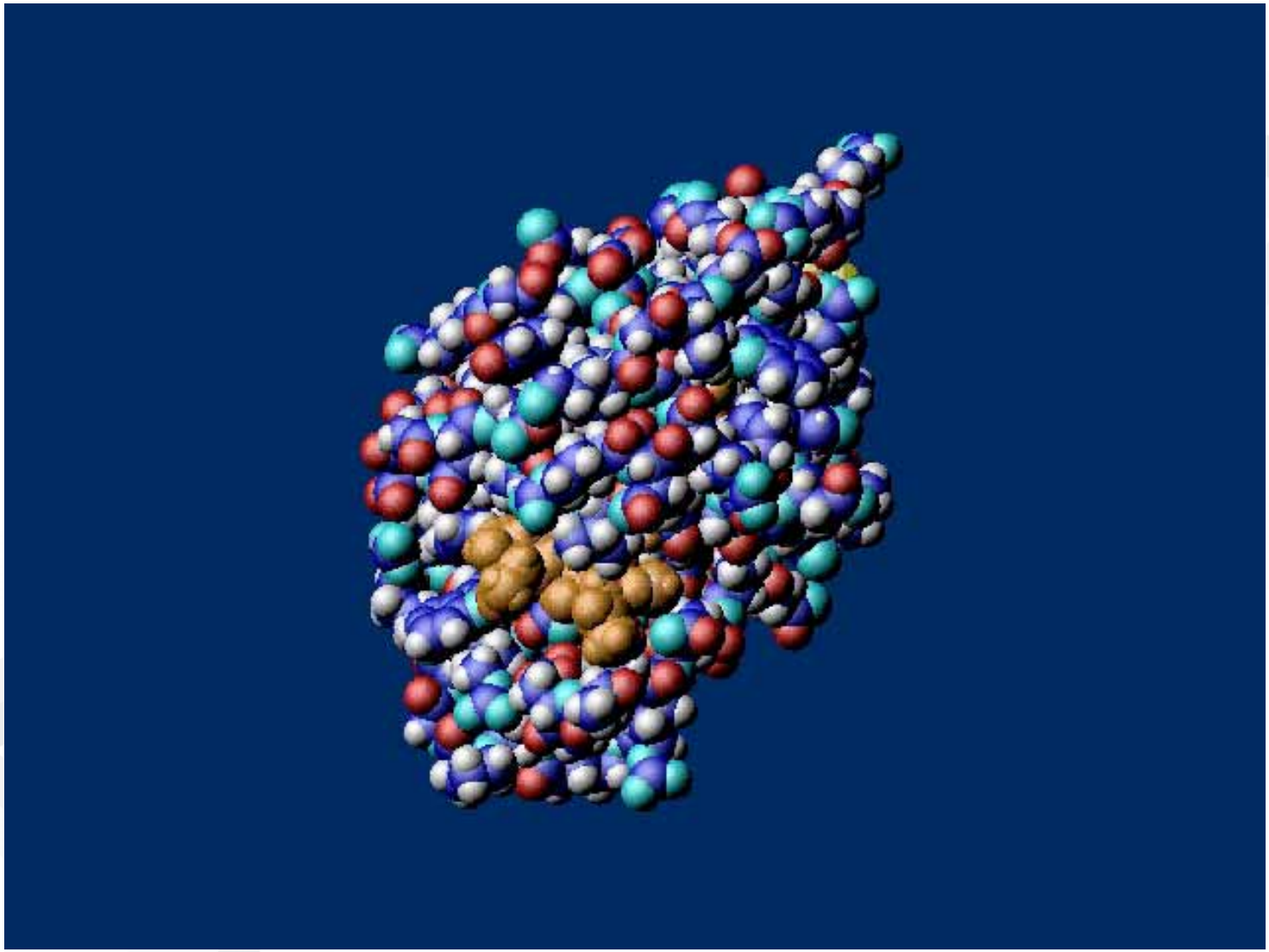
H^+

H:
C:
N:
O:



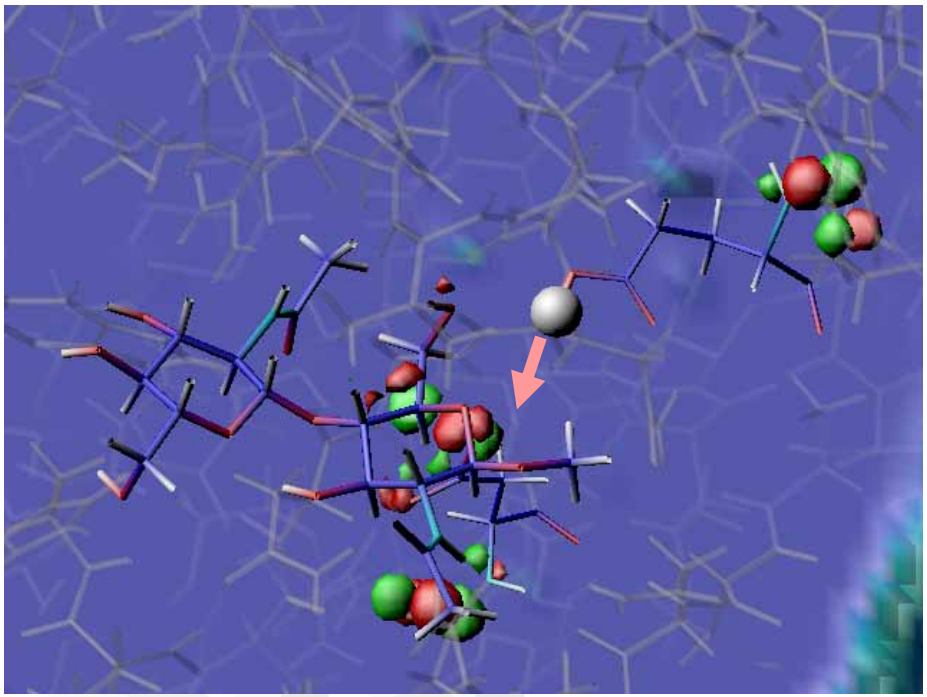
NAM: N-Acetylmuramin
NAG: N-Acetylglucosamin

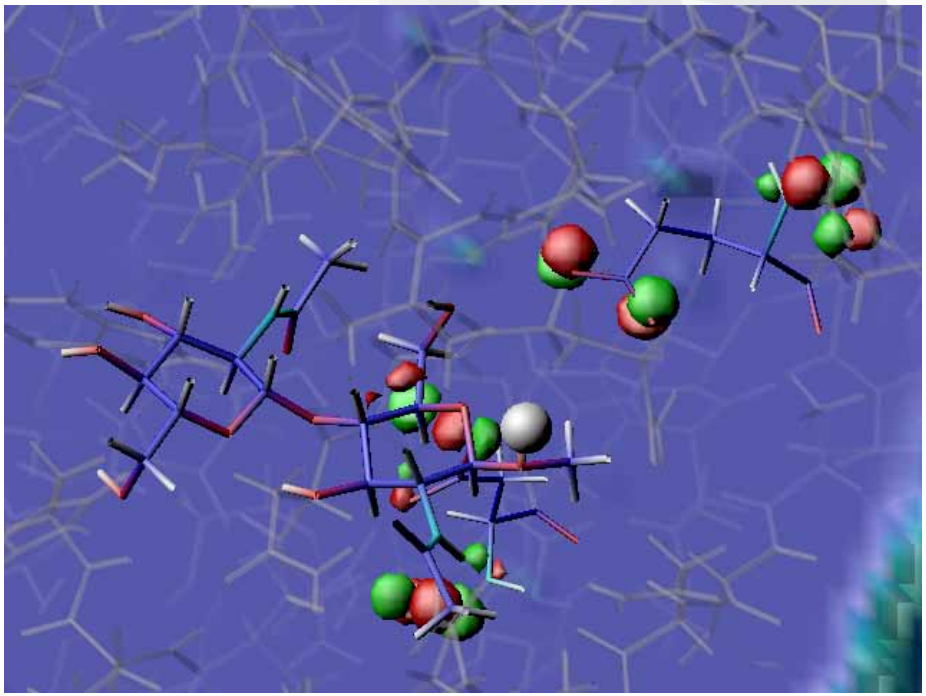
Some demonstration movies of Mediator/GridMPI Based Nano-Science Coupled Simulations



Electronic structure analysis on Proton transfer in Lysozyme (in solution)

Proton transfer from active site of Glu35 to peptidoglycan

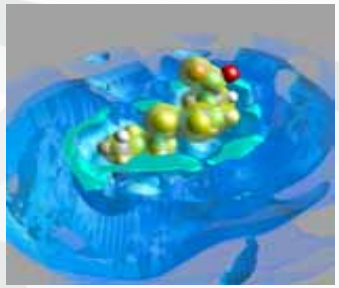




RISM-FMO coupled simulations reveal realistic microscopic mechanism of proton transfer in Lysozyme

Outline of Today's Talk

1. Grid-Enabling by Mediator/GridMPI

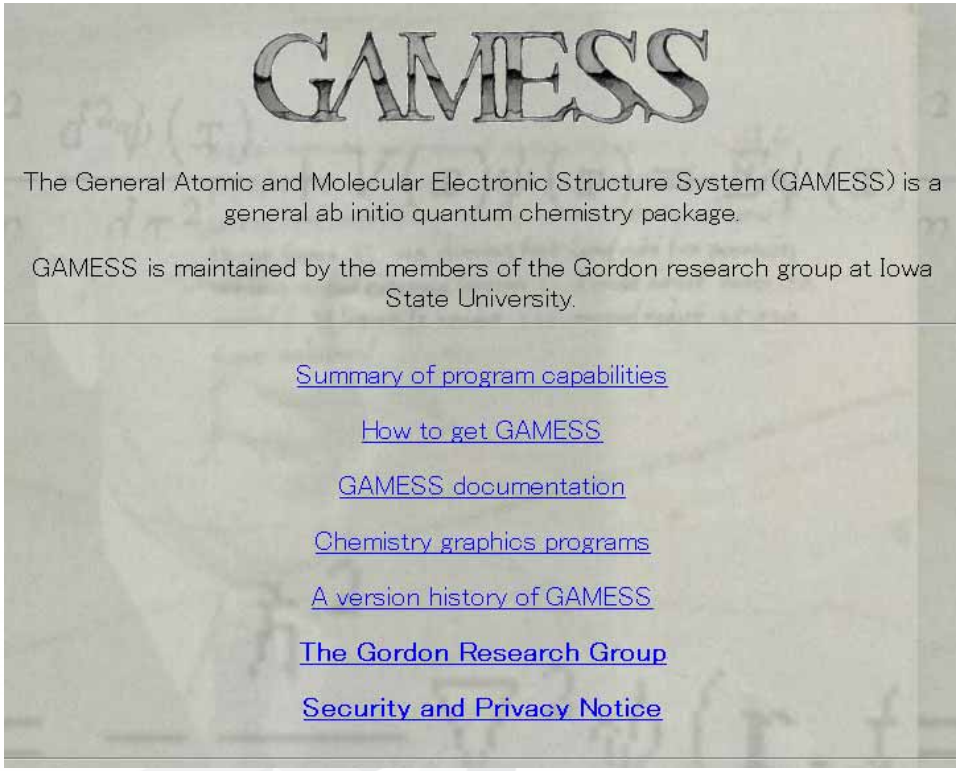


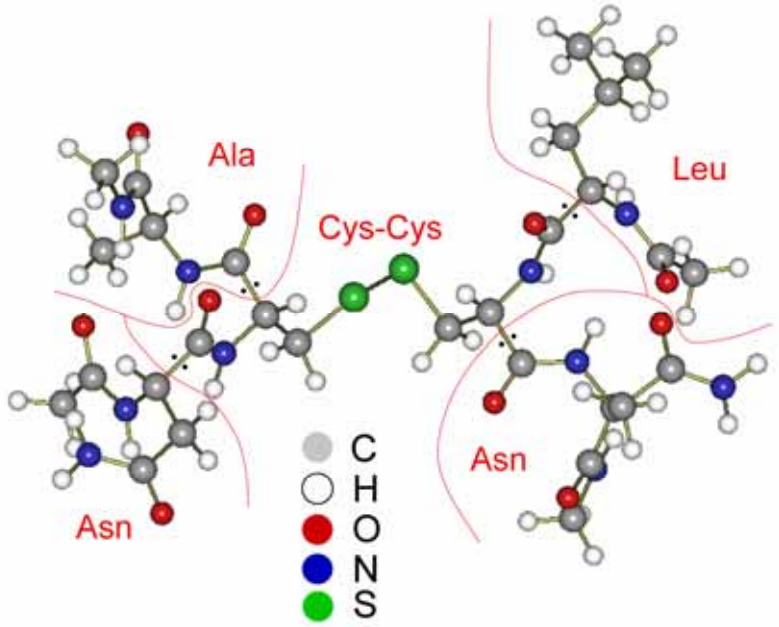
Application to Solvent interactions in hydrolysis of Lysozyme
MD-PB and RISM-FMO coupled simlations

2. NAREGI workflow tools with Loosely Coupled Application Components

A Case study with GAMESS FMO and TINKER

Grid enabling of the GAMESS FMO prog.

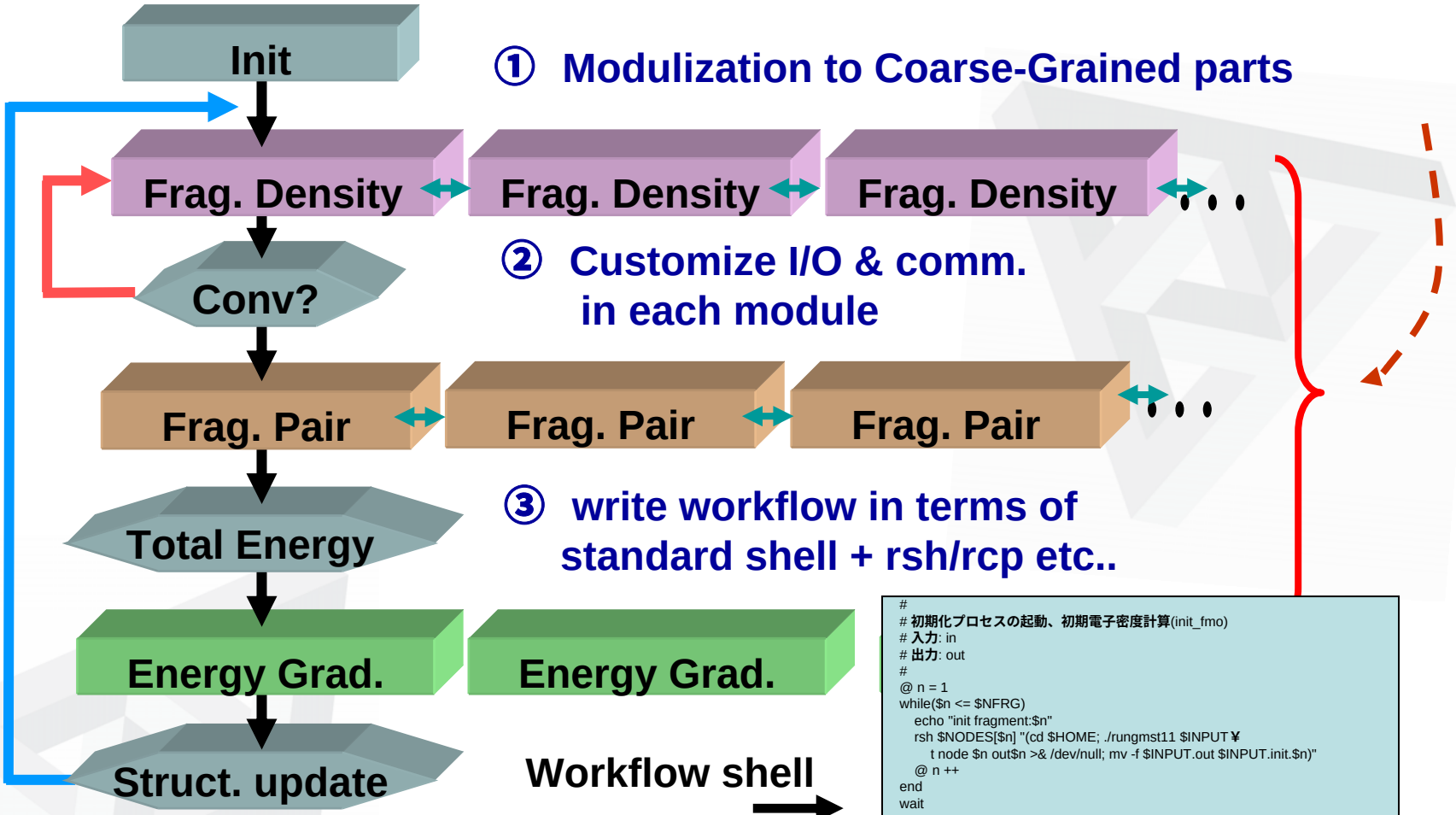




http://www.msg.ameslab.gov/GAMESS/GAMESS.html

One of the most popular open-source program used in the world-wide nano science community.

Grid enabling of the GAMESS FMO prog.



① Modulization to Coarse-Grained parts

② Customize I/O & comm. in each module

③ write workflow in terms of standard shell + rsh/rcp etc..

Workflow shell ⇒

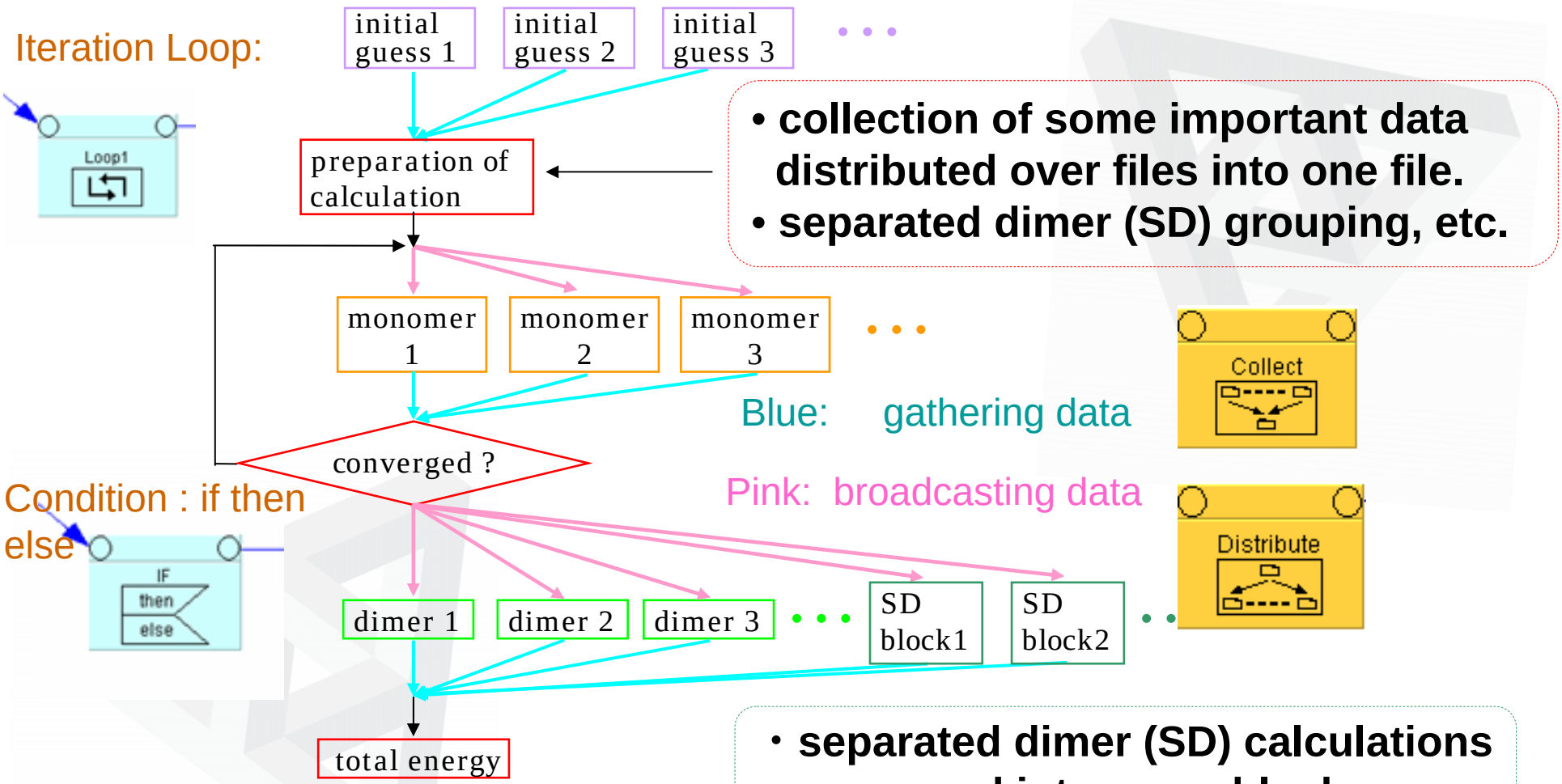
```
#
# 初期化プロセスの起動、初期電子密度計算(init_fmo)
# 入力: in
# 出力: out
#
@ n = 1
while($n <= $NFRG)
   echo "init fragment:$n"
   rsh $NODES[$n] "(cd $HOME; ./rungmst11 $INPUT¥
      t node $n out$n >& /dev/null; mv -f $INPUT.out $INPUT.init.$n)"
   @ n ++
end
wait

# データの収集
@ n = 1
while($n <= $NFRG)
    if($MY != $NODES[$n]) then
       rcp $NODES[$n]¥:$HOME/out$n out$n >& /dev/null
    endif
    @ n++
end
wait
```

Shell scripts

④ deploy each component and translate shell scripts into NAREGI workflow and submit to NAREGI SuperSchedular.

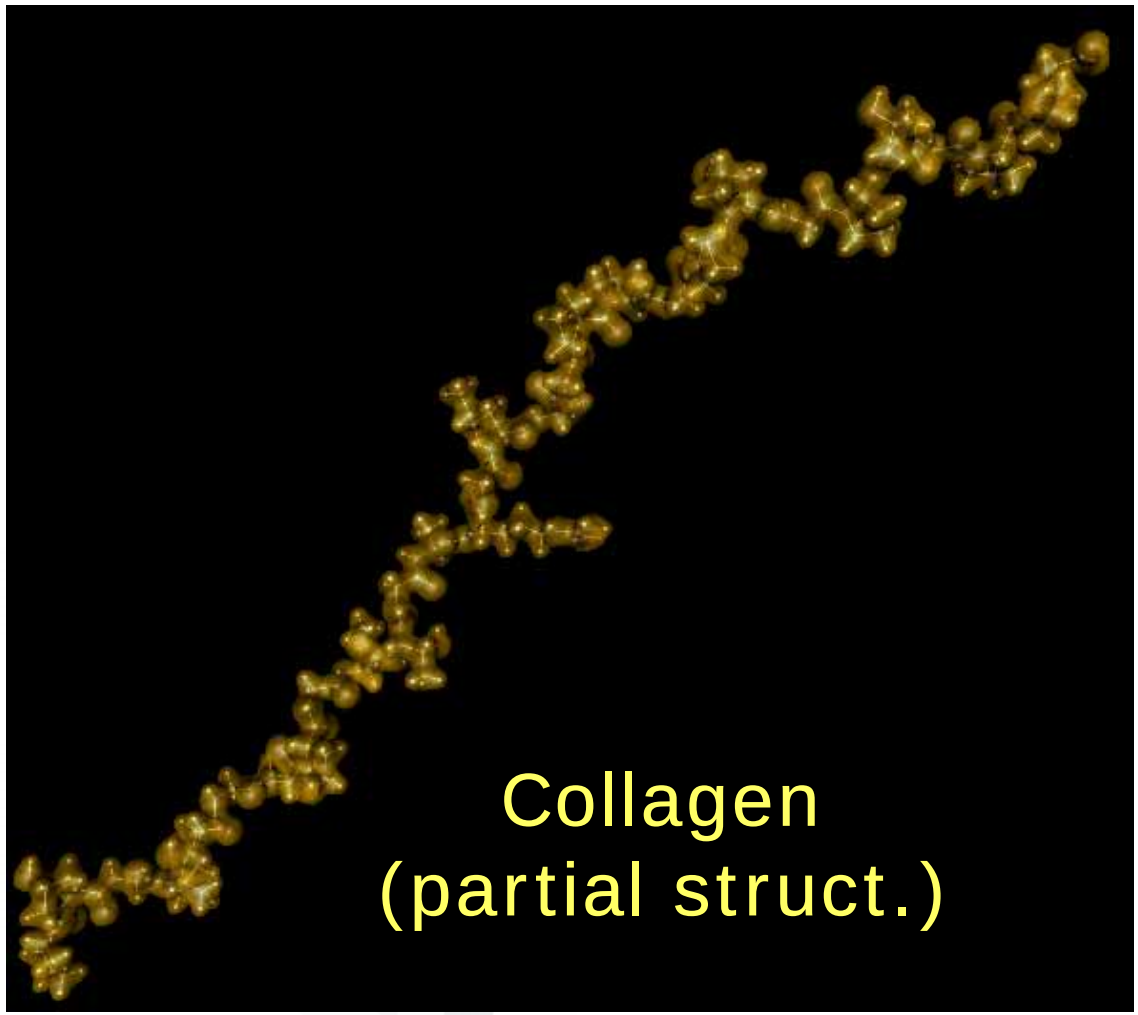
Use of NAREGI WorkFlow Tool

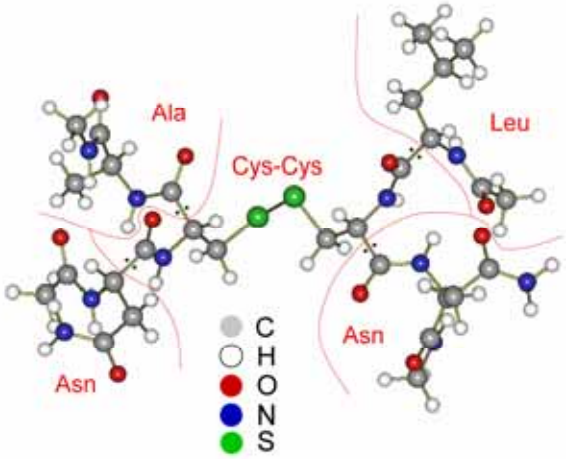


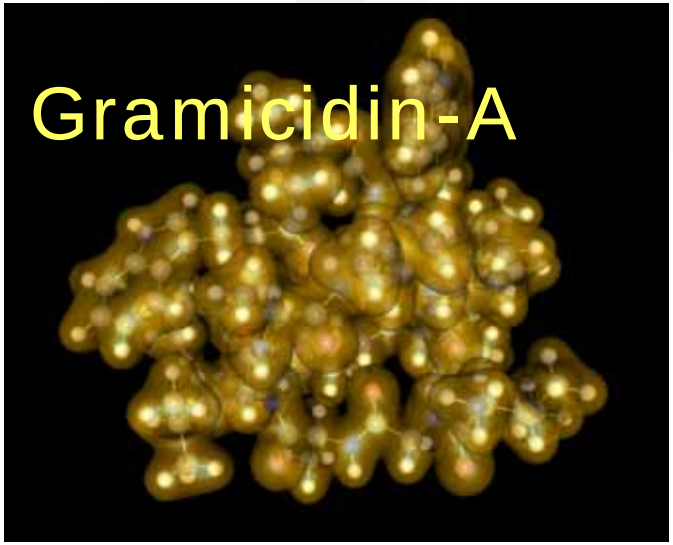
- collection of some important data distributed over files into one file.
- separated dimer (SD) grouping, etc.

- separated dimer (SD) calculations are grouped into some blocks

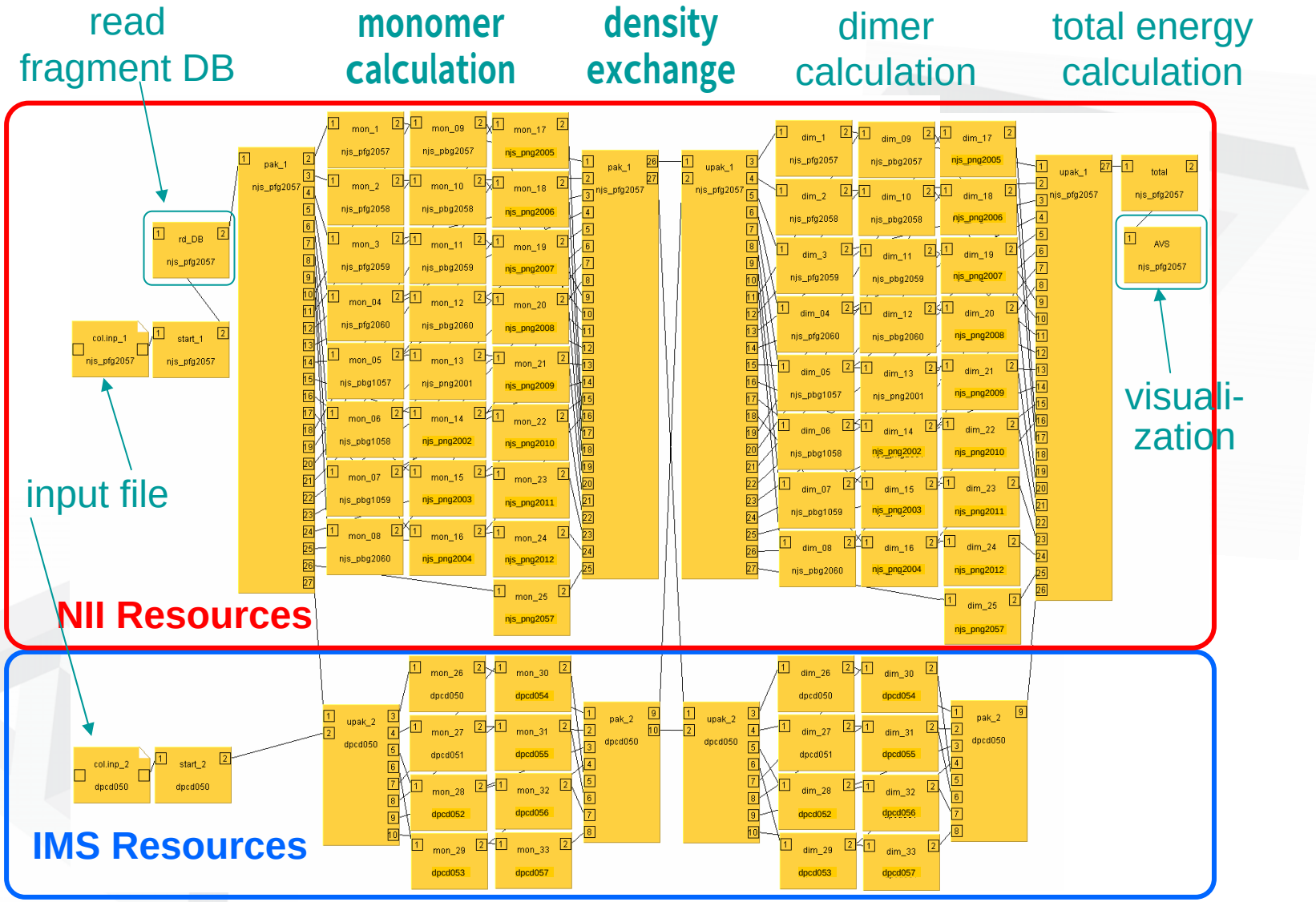
Examples:Protein Molecules examined by Loosely coupled FMO on Computational Grid



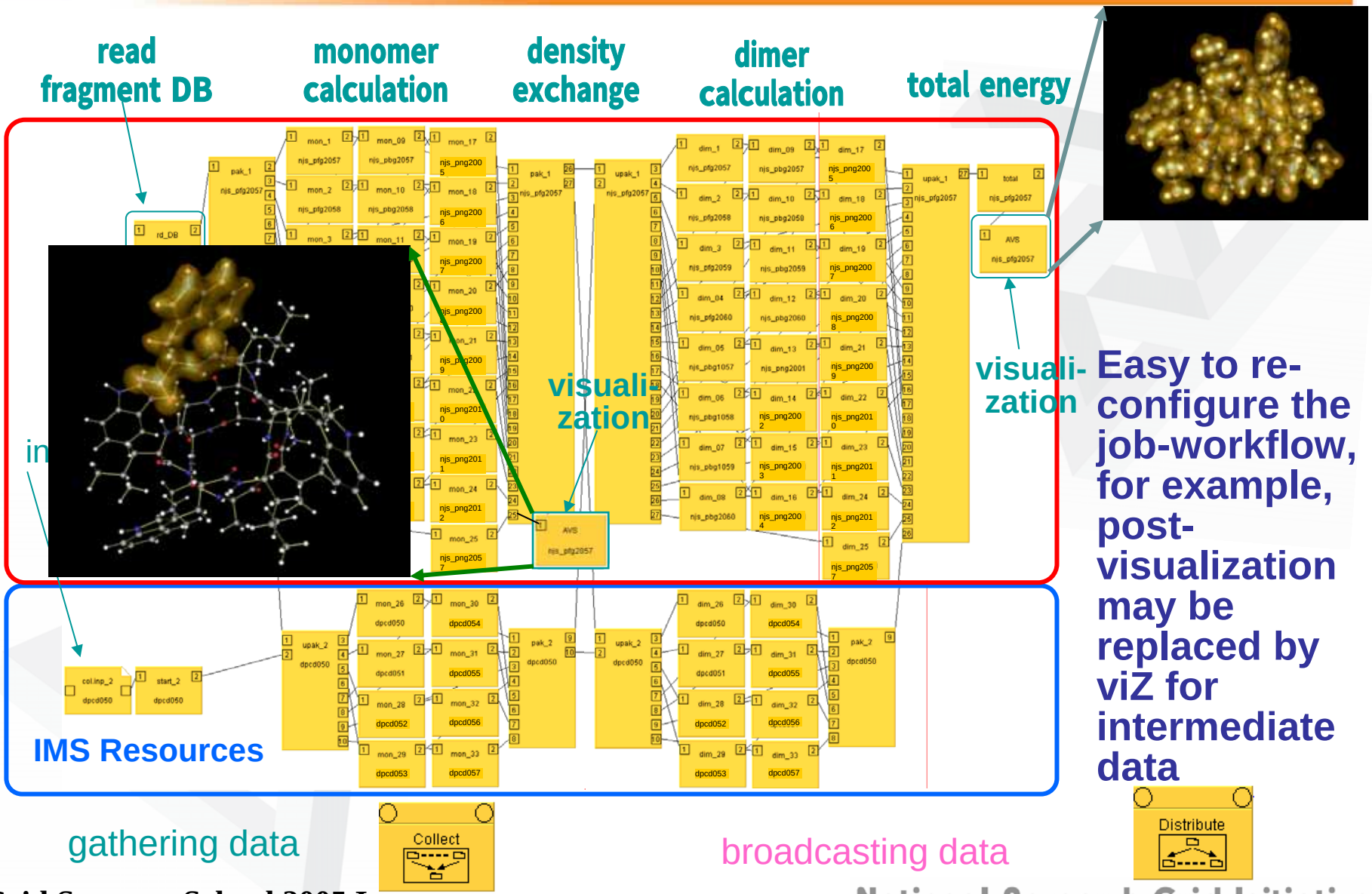




Workflow based Grid FMO Simulations of protein



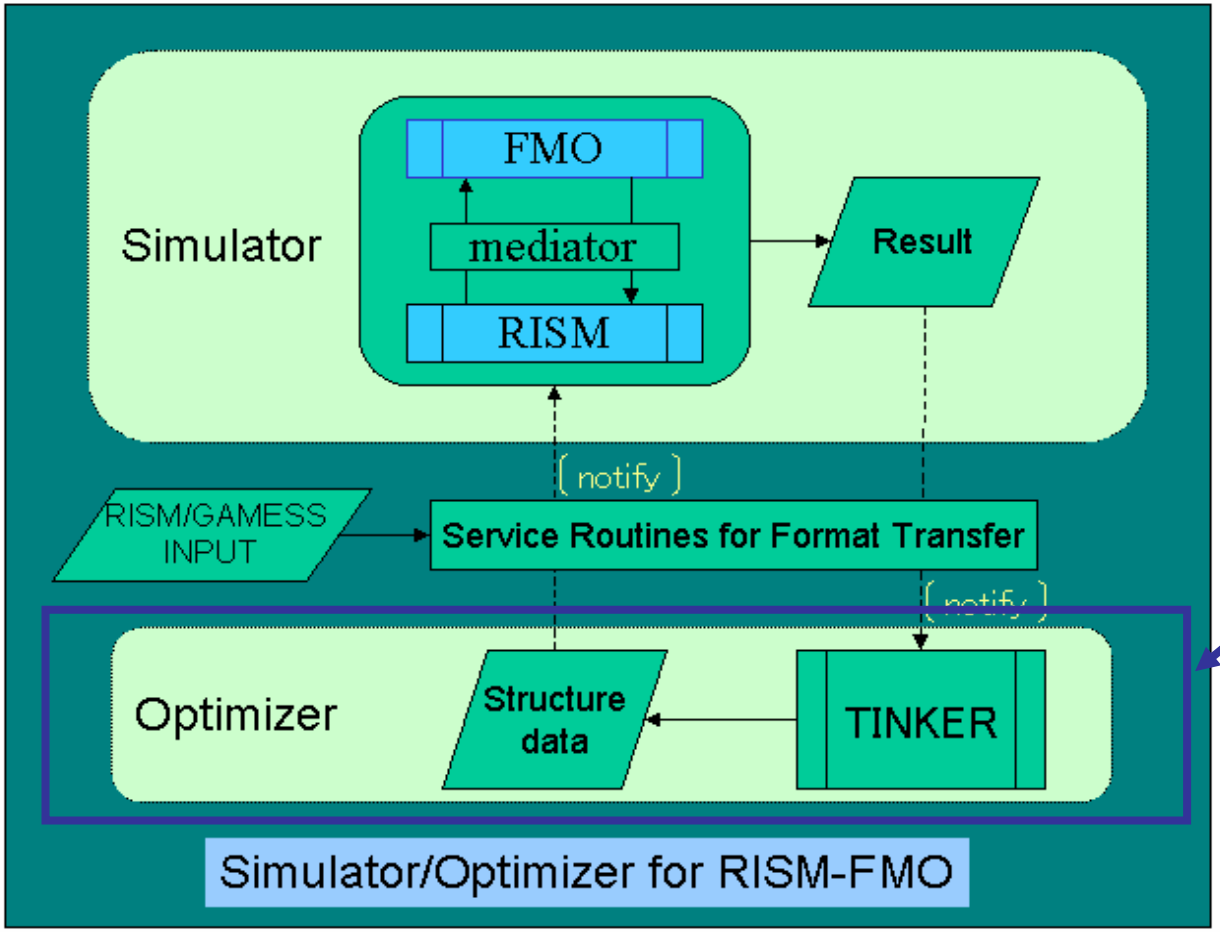
Fragment Electron Density in Grid FMO



Easy to re-configure the job-workflow, for example, post-visualization may be replaced by viZ for intermediate data

Full Geometry Optimization of Protein in water

3D-RISM/FMO with loosely coupled TINKER

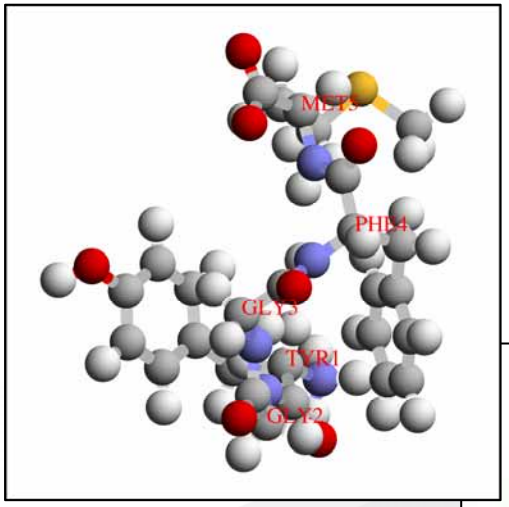


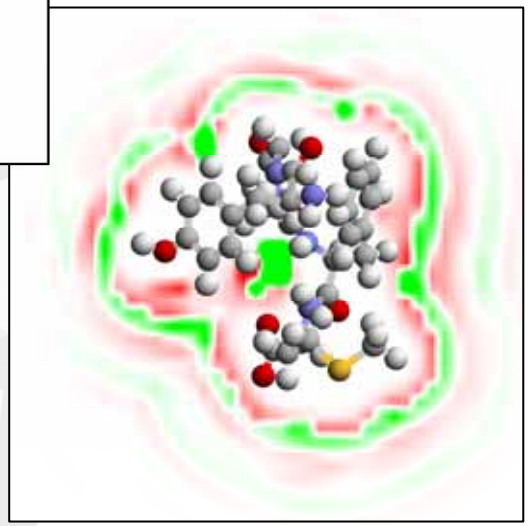
Once workflow scenario is constructed, we may easily modify and extend it with NAREGI WFT editor.

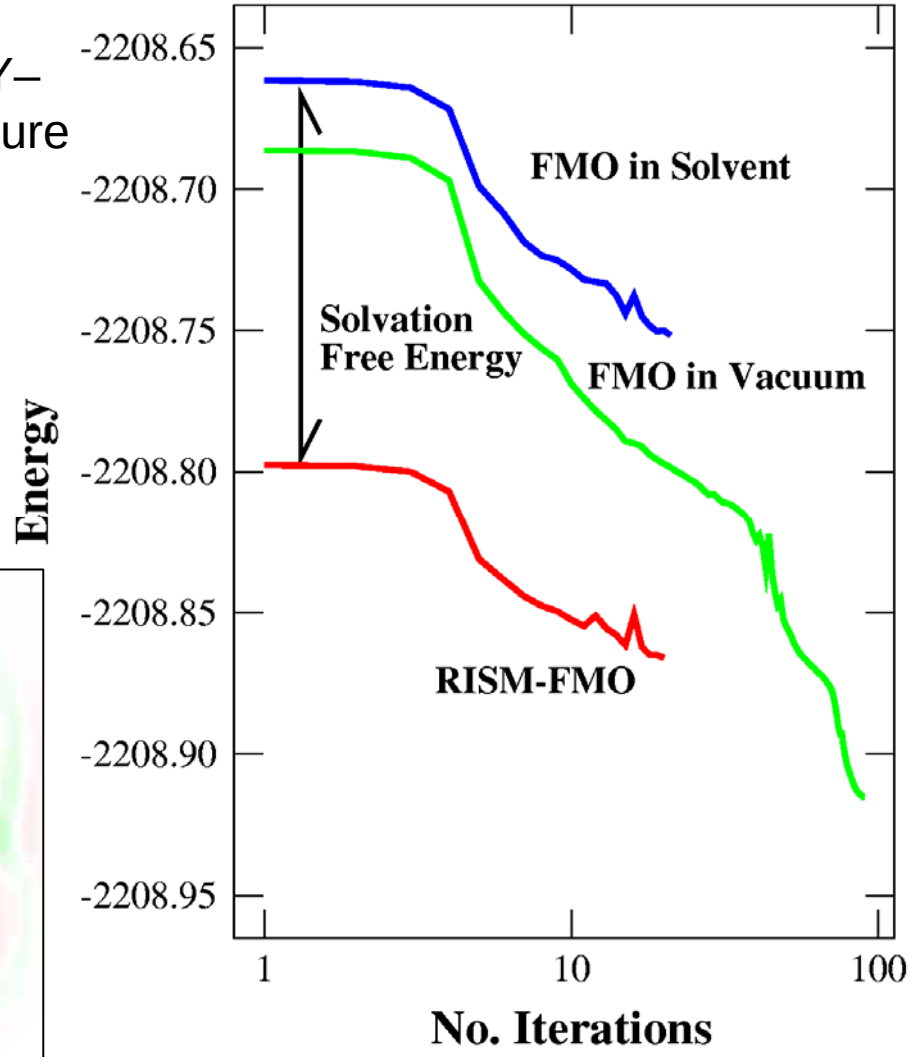
In this example, Optimizer task is attached to the main RISM-FMO coupled simulation, and construct the iteration loop for geometry OPT.

Full Geometry OPT. of Met-enkephalin

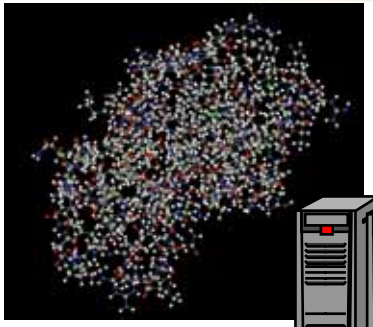
The amino-acid sequence is 'TYR–GLY–GLY–PHE–MET'. We start the optimization procedure from the Model 1 of PDB data "1plw".





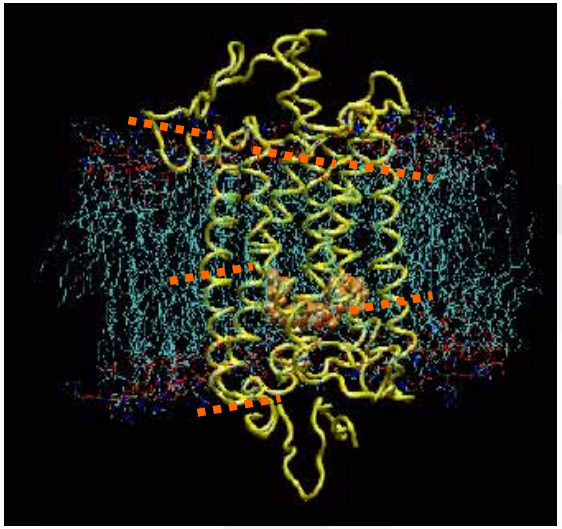


Grid Database tool for Optimized Protein Structures and Fragment density



To be used in the other FMO simulations

Retrieve

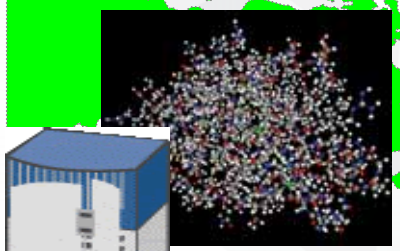


Register

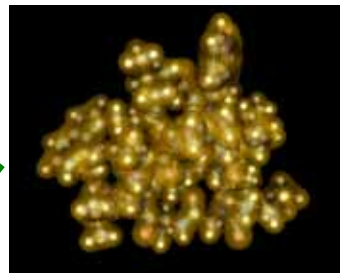
Fragment Database

The Fragment density is stored in the Fragment Database system.

Register

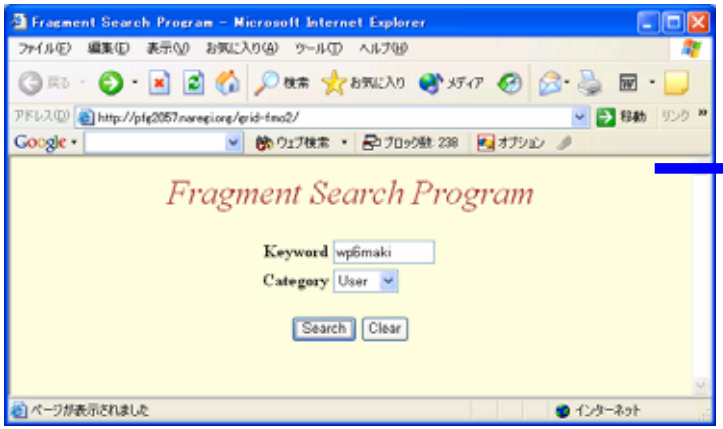


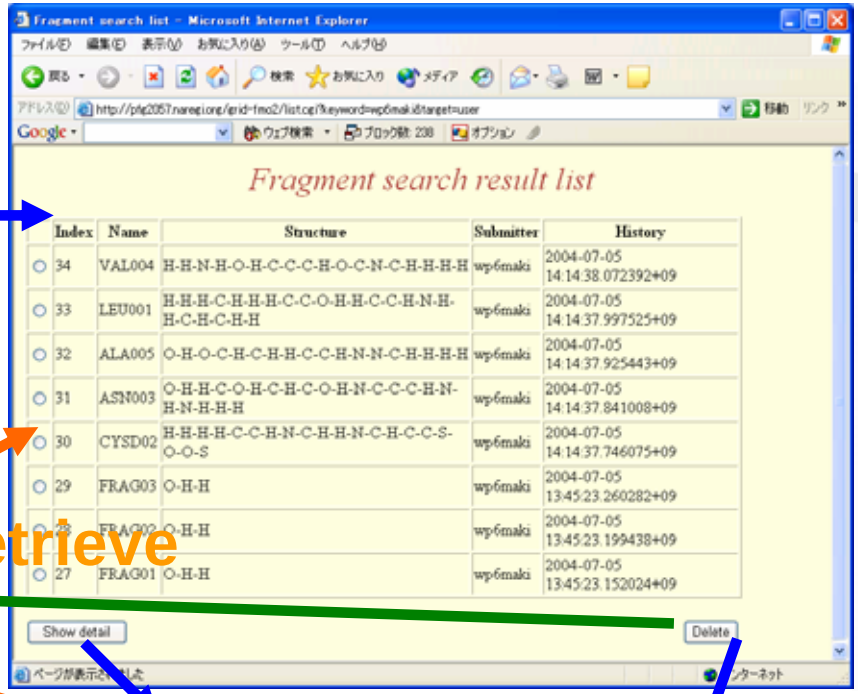
Register



Fragment Database Tools

Search





Fragment Database

Retrieve

Register

more information...

Concluding remarks

A solution based on Research Grid will have significant advantages for customizing interdisciplinary simulations,

and

Grid will make a big chance for the Collaboration between Computational Scientists and Computer/Information Scientists in HPC field.

Acknowledgement

Dr. S. Sekiguchi and Prof. S. Matsuoka
Co-workers in Hitachi and Fujitsu Co. Ltd.
NAREGI

- Members of Workflow Team
- Members of SuperScheduler Team

Special Thanks to

- Dr. Ho, Dr. Kubo (Hitachi)
- Mr. Miura, Mr. Takano (Fujitsu Kyushu)