ShareGrid and Medicinal Chemistry research

Paolo Tosco

Università degli Studi di Torino

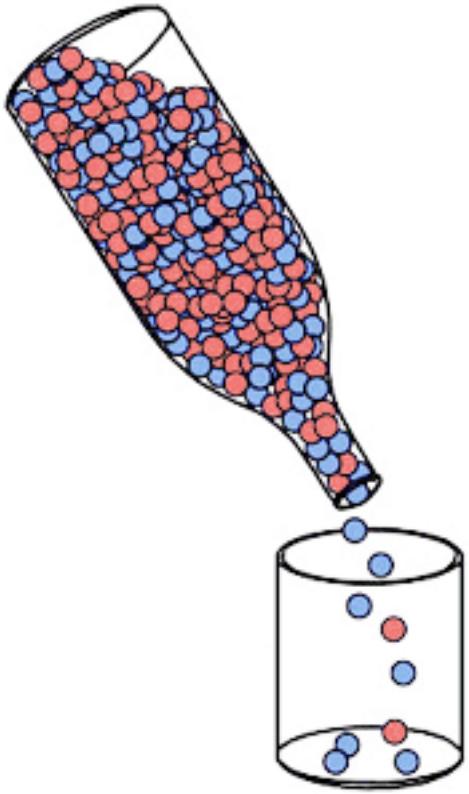
Dipartimento di Scienza e Tecnologia del Farmaco

Torino, 5th December 2007

Quest for new *lead compounds*: brute force…

Combinatorial synthesis

Many molecules characterised by high molecular diversity can be synthesised



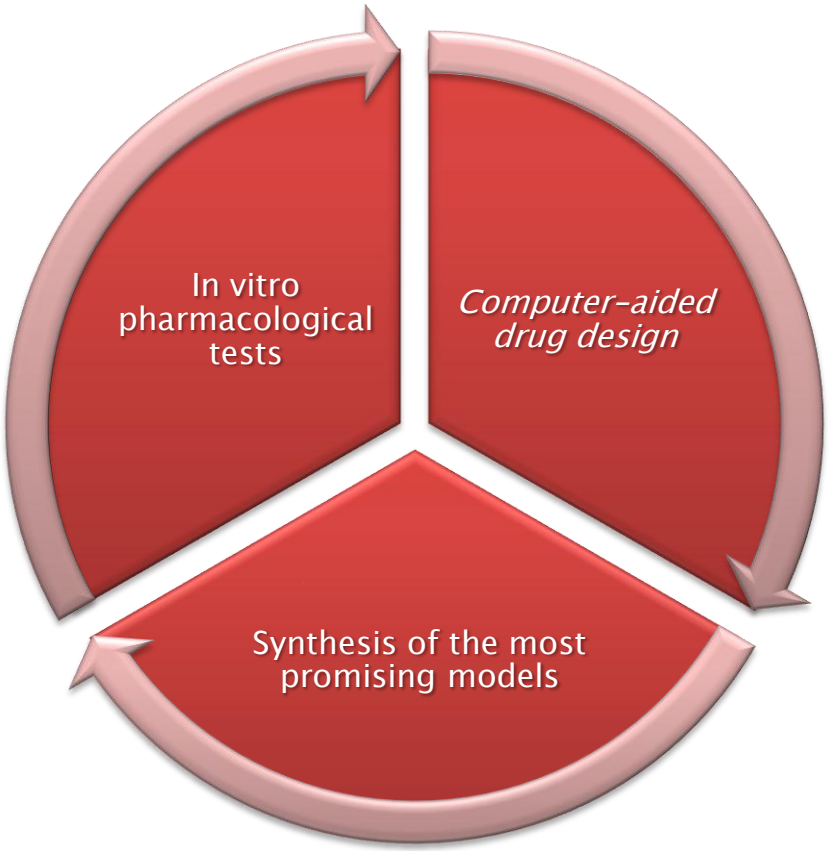
High-throughput screening

Many *in vitro* pharmacological tests can be carried out in a short time

Identifying the molecule responsible for the biologic activity without purifying reaction mixtures can be a hard task

…and rational design

Optimization of models according to the responses of *in vitro* assays



- *Virtual screening* of large compound libraries on a target whose 3D structure is known
- *Rational drug design*

Synthesis is limited to the most promising models

Docking of models into the active site of the biological target

RCSB PDB PROTEIN DATA BANK

A MEMBER OF THE PDB

An Information Portal to Biological Macromolecular Structures

As of Tuesday Nov 27, 2007 there are 47509 Structures | PDB Statistics

The **RCSB** PDB provides a variety of tools and resources for studying the structures of biological macromolecules and their relationships to sequence, function, and disease.

http://www.rcsb.org

Molecular biologists determine the 3D structure of complexes between ligands (potential drugs) and biological substrates (enzymes, receptors, nucleic acids)

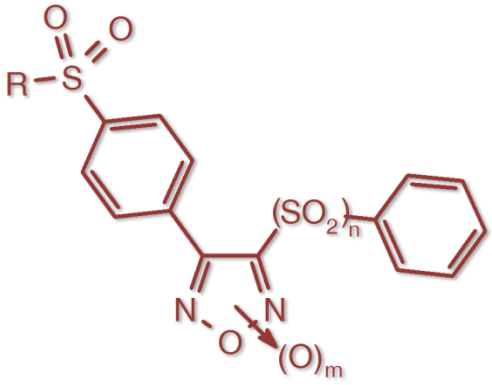
Substituting the ligand present in the experimental crystallographic complex with one's drug candidate, the possibility that it can be active on the same biological target can be estimated

Selective inhibitors of cyclooxygenase-2 enzyme (COX-2)

Del Grosso, E.; Boschi, D.; Lazzarato, L.; Cena, C.; Di Stilo, A.; Fruttero, R.; Moro, S.; Gasco, A. *Chem. Biodiv.* **2005**, *2*, 886-900

COX-1 (PDB entry 1q4g)

Click to download movie



n = 0, 1; m = 0, 1; R = NH_2, CH_3

COX-2 (PDB entry 1cx2)

Click to download movie

Flexible docking into the active site of both isozymes

"Flexible" refers only to the ligand's 3D coordinates,
while the enzyme structure is kept rigid

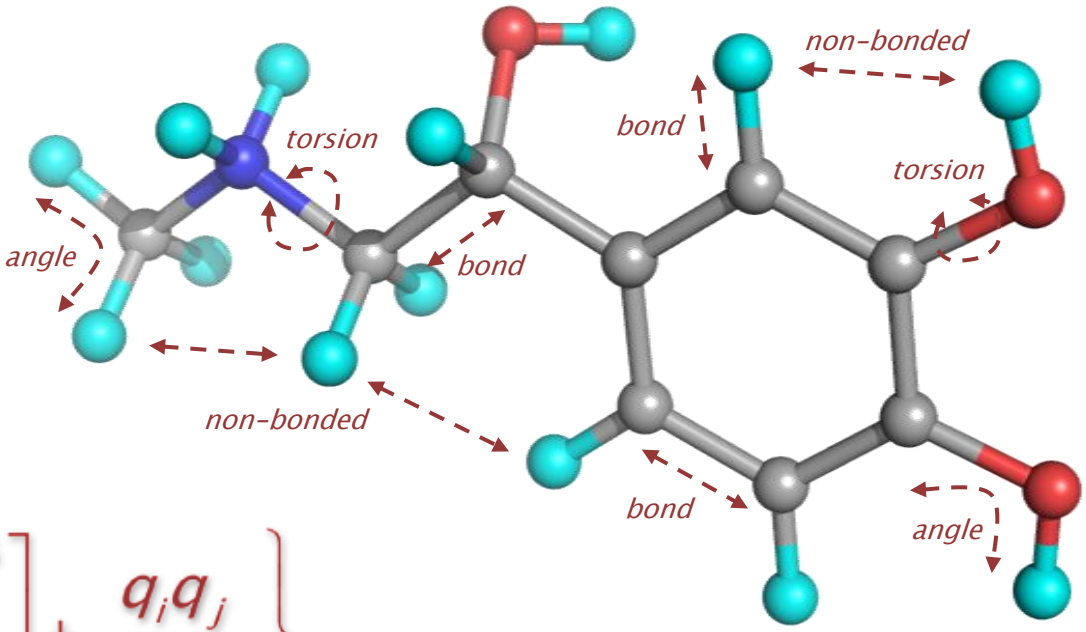
$$E_{MM} = E_{bond} + E_{angle} + E_{torsion} + E_{vdW} + E_{coulomb} =$$

$$= \sum_{bonds} \frac{1}{2} k_b (r - r_0)^2 +$$

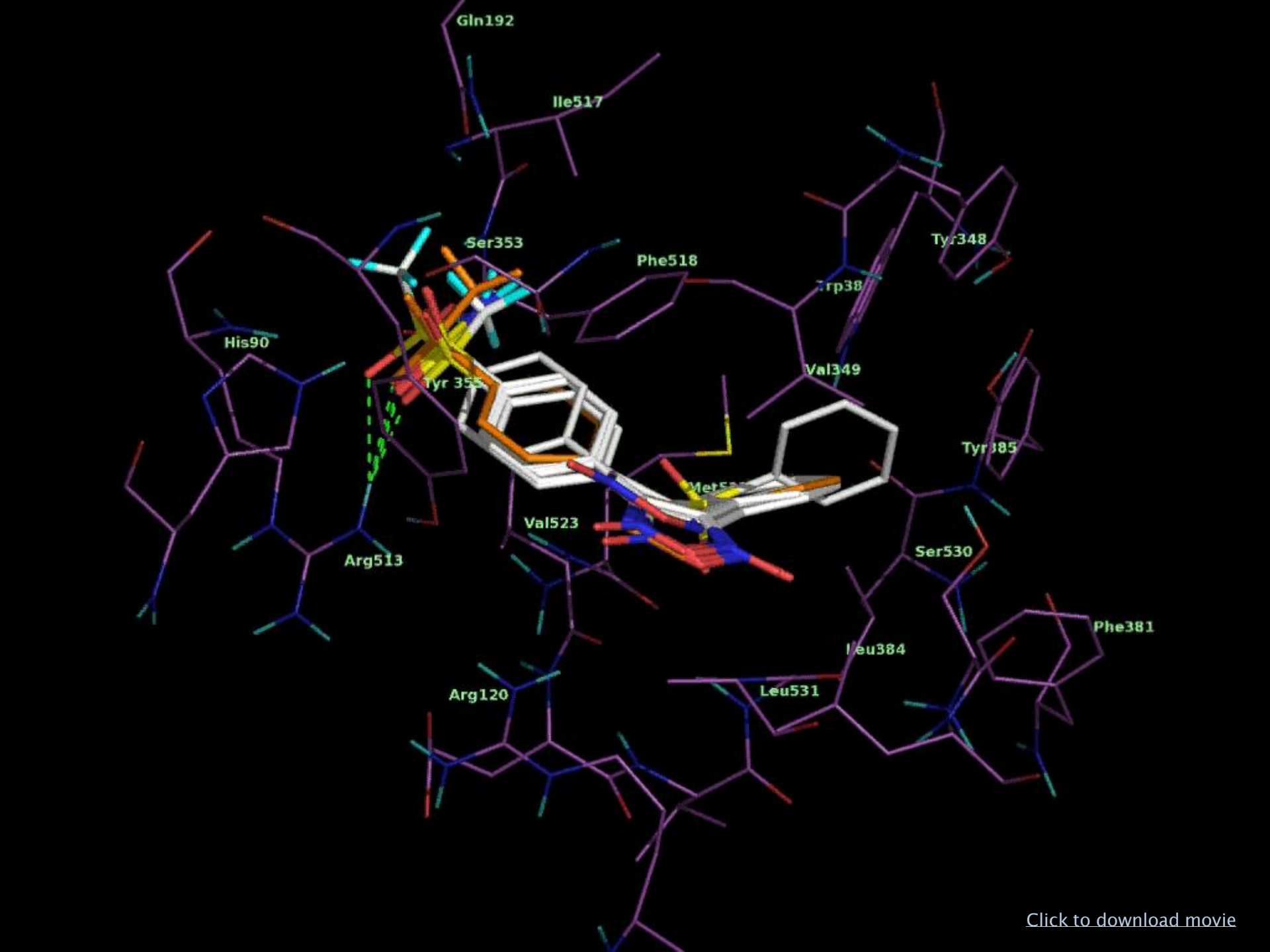
$$+ \sum_{angles} \frac{1}{2} k_a (\theta - \theta_0)^2 +$$

$$+ \sum_{torsions} \frac{1}{2} V_n \left[1 + \cos(n\omega - \gamma)\right] +$$

$$+ \sum_{j=1}^{N-1} \sum_{i=j+1}^{N} \left\{ 4\varepsilon_{i,j} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right\}$$



The function describing the potential energy of the ligand-target complex is minimized over a large number of randomly-generated ligand-enzyme complexes through a genetic algorithm


Gln192
Ile517
Ser353
Phe518
Tyr348
Trp38
His90
Tyr 355
Val349
Tyr385
Val523
Arg513
Ser530
Phe381
Leu384
Arg120
Leu531
Click to download movie

CPU-time to accomplish flexible docking of these six molecules in the active site of both isozymes

6 h on single-core CPU
(Pentium 4 EM64T 3.2 GHz)

Docking has become a widely used technique in Medicinal Chemistry

Possible issues connected with docking

Soft failure

A ligand is predicted to have some interesting pharmacological activity, but *in vitro* assays show that the latter is indeed very low or absent

Hard failure

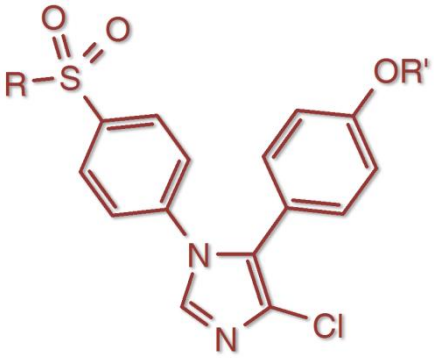
A ligand is predicted to have no pharmacological activity, but *in vitro* assays show that actually it is very active against some target

Selective inhibitors of cyclooxygenase-2 enzyme (COX-2)

Chegaev, K., Lazzarato, L., Tosco, P., Cena, C., Marini, E., Rolando, B., Carrupt, P.-A., Fruttero, R., Gasco, A. *J. Med. Chem.* **2007**, *50*, 1449-1457

COX-1 (PDB entry 1q4g)

Click to download movie

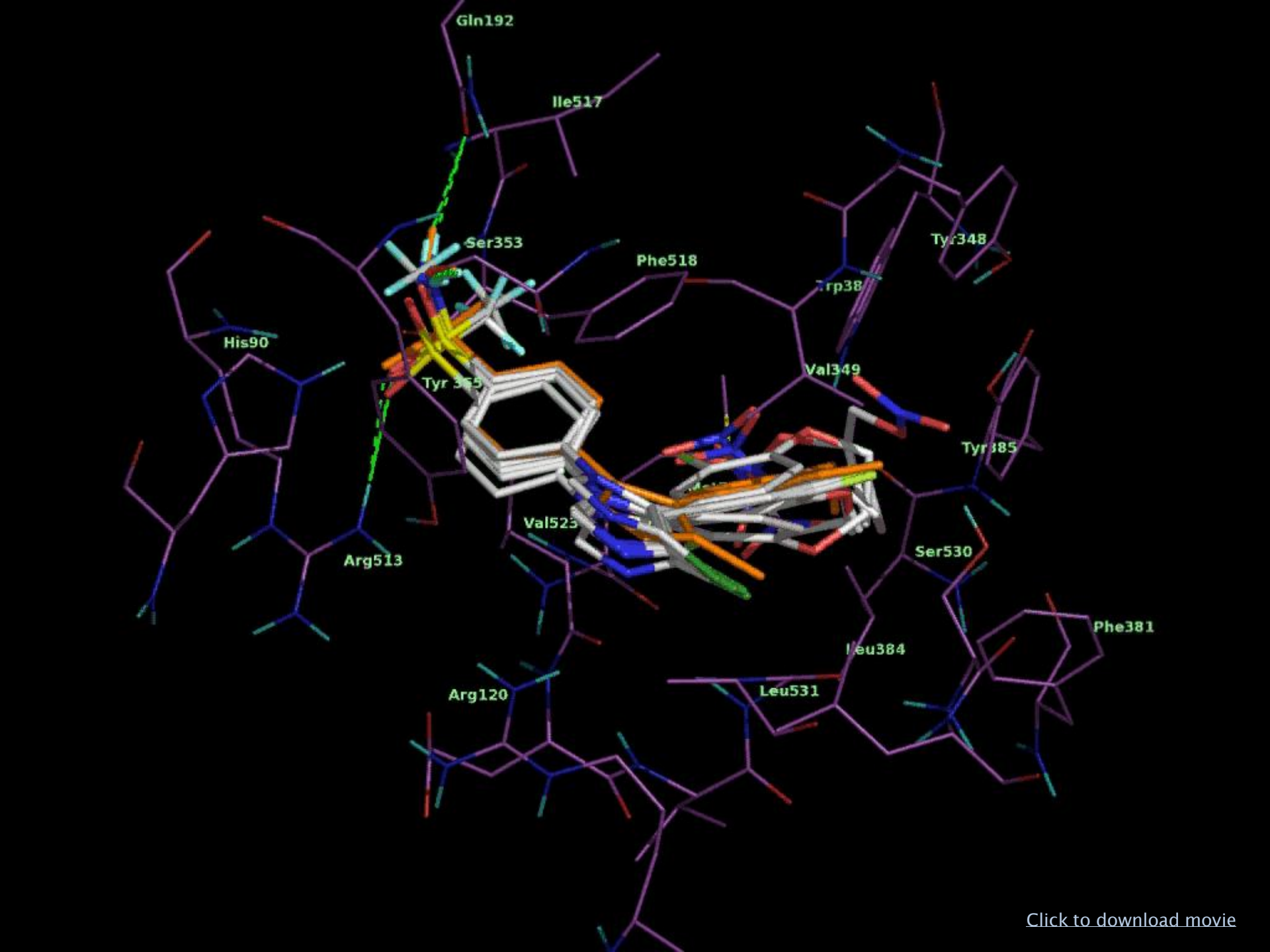


R = NH_2, CH_3
R' = $(CH_2)_2ONO_2$, $(CH_2)_3ONO_2$, $CH_2CH(ONO_2)CH_2ONO_2$

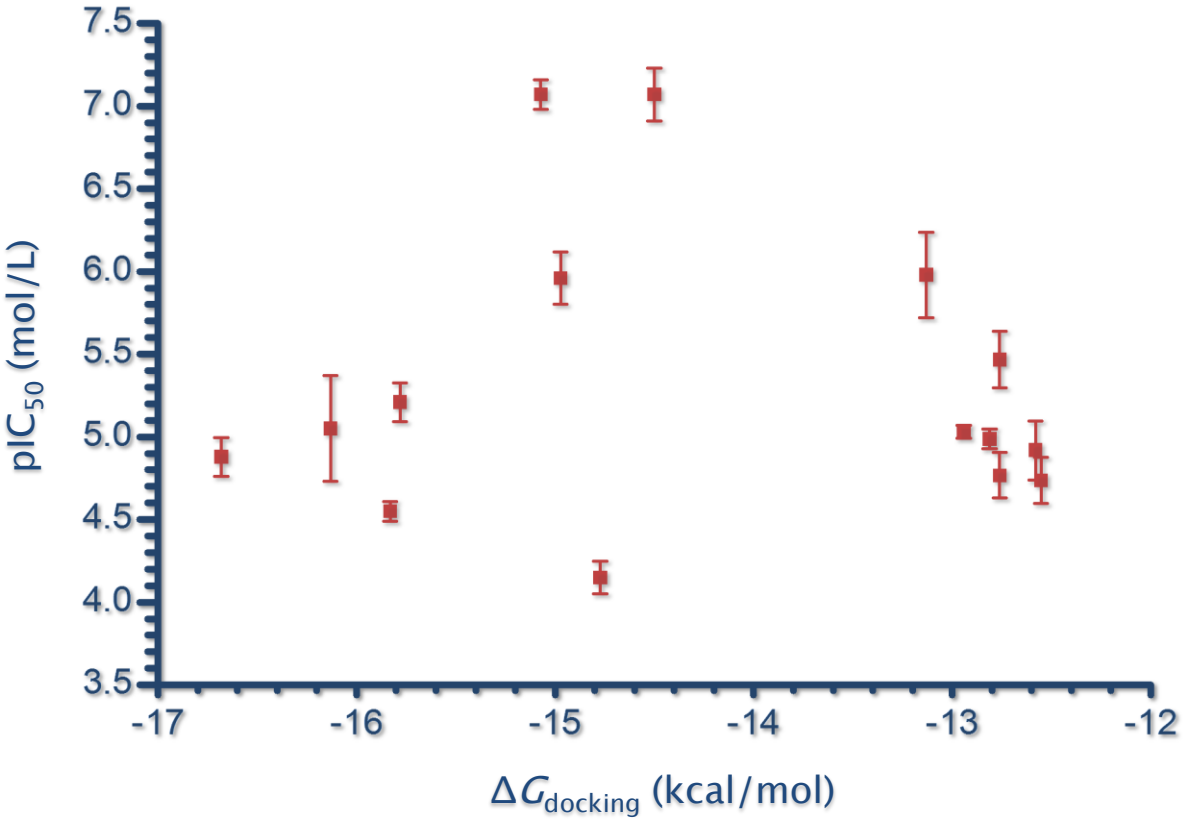
COX-2 (PDB entry 1cx2)

Click to download movie

Also this series of compounds maintains a good degree of COX-2/COX-1 selectivity, but compared to the previous series all compounds are 10 to 100-fold less active on the COX-2 isozyme


Gln192
Ile517
Ser353
Phe518
Tyr348
Trp38
His90
Val349
Tyr385
Val523
Arg513
Ser530
Phe381
Leu384
Leu531
Arg120
Click to download movie

Flexible docking: quantitative results (?)



The attempt to find a quantitative correlation between experimental pharmacological activity and the affinity for the active site predicted by a docking simulation results in a dramatic failure

Mozziconacci, J.-C.; Arnoult, E.; Bernard, P.; Do, Q.T.; Marot, C.; Morin-Allory, L. *J. Med. Chem.* **2005**, *48*, 1055-1068

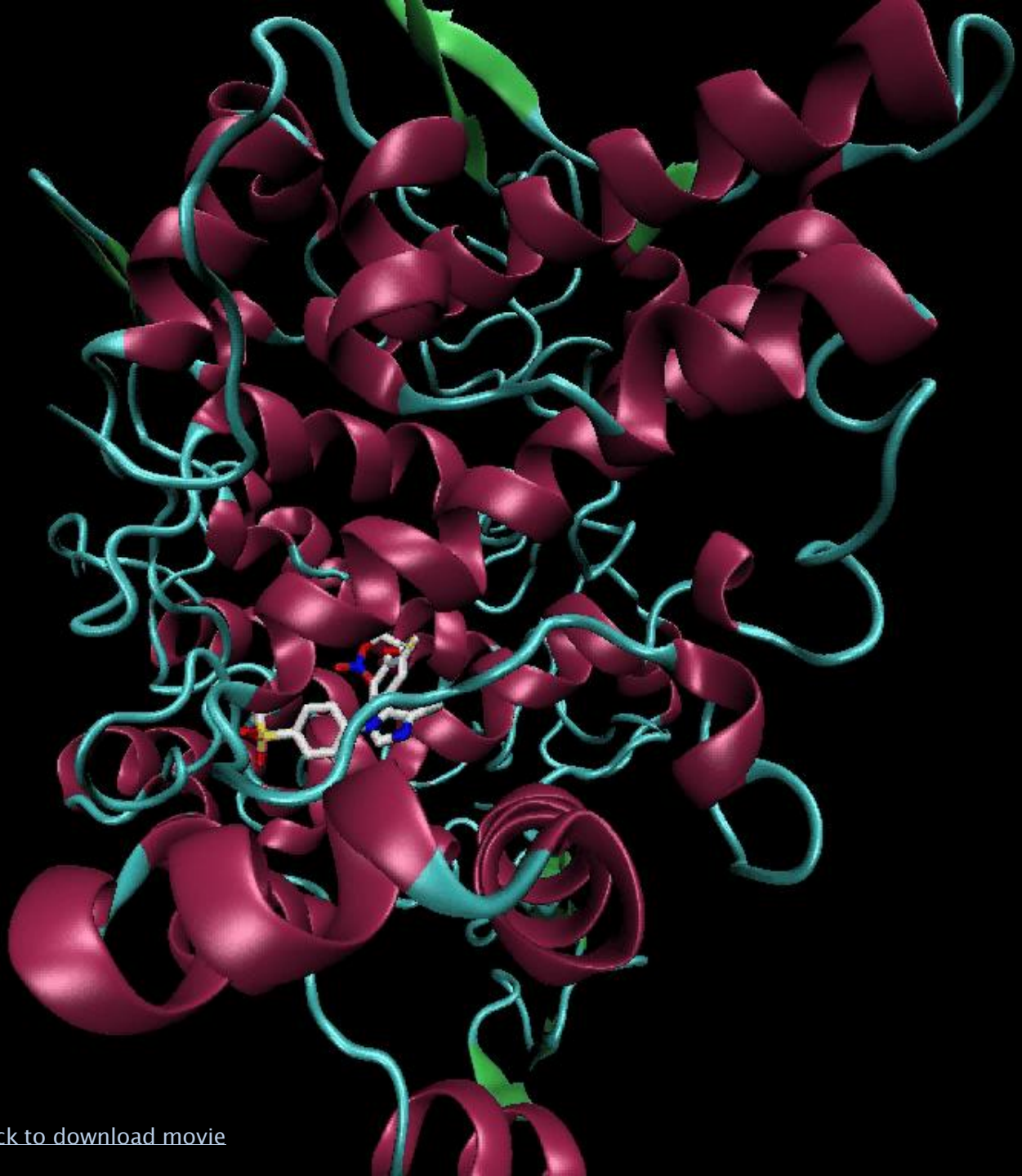
Model refinement

It is necessary to account for solvent effects

It is necessary to model the flexibility of the macromolecular target: the cyclooxygenase enzyme is not rigid, but can assume many different conformations

In addition to the 40 atoms which constitute the ligand, it is necessary to compute the coordinates accessible to the 8928 atoms which make up the protein

The computational cost scales approximately as N^2, where N is the number of atoms in the system



Click to download movie

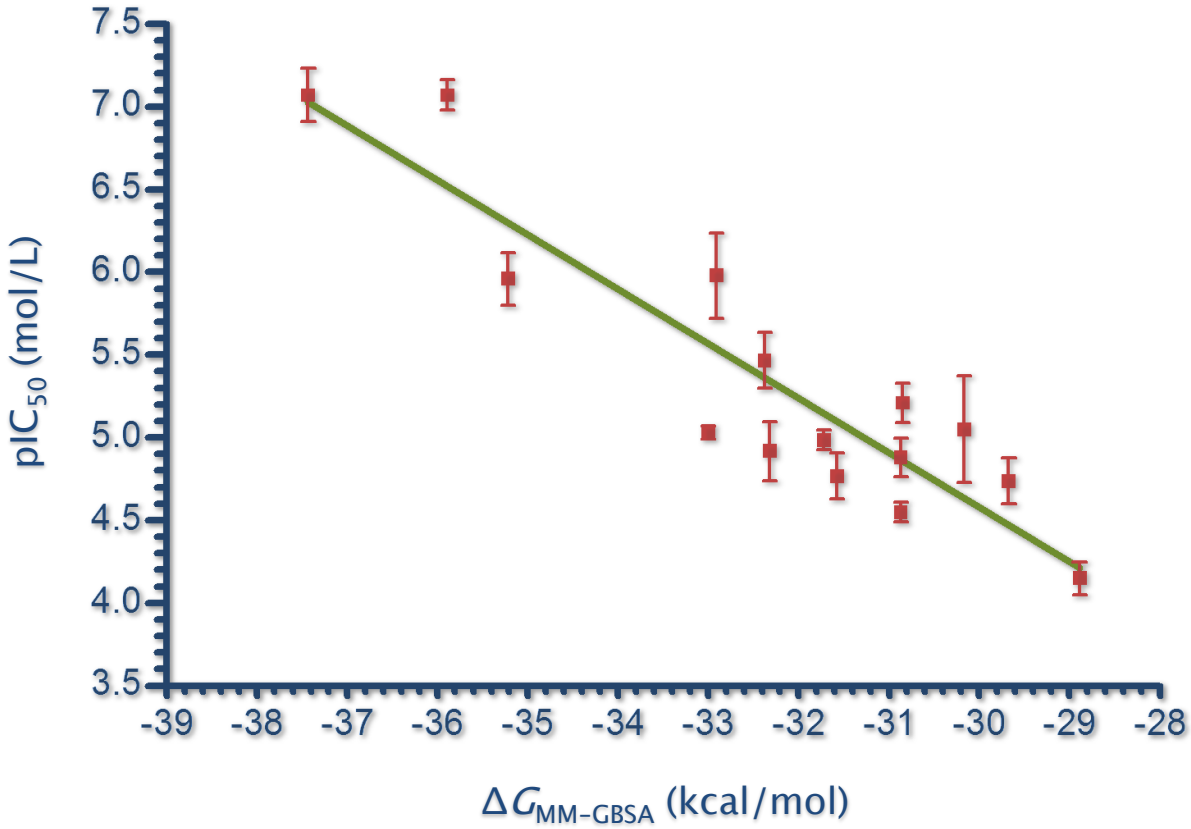
40 ps molecular dynamics using the GBSA method implemented in the SANDER module (AMBER 8)

The 400 final conformations are collected and used to compute $\Delta G_{binding}$

In order to limit CPU-time the entropic term computation is omitted, and solvent is modelled as a continuum (implicit solvent model)

$\Delta G_{binding}$ values thus computed are larger in absolute value than experimental $\Delta G_{binding}$

MM-GBSA method: quantitative results (!)



Good correlation between experimental and computed values

$pIC_{50} = -0.33 \cdot \Delta G_{MM-GBSA} - 5.27$
$F = 65.34$, $P < 0.0001$
$r^2 = 0.83$, $q^2 = 0.78$

CPU-time (setup, tuning, production) → **3 months** (8 single-core CPUs, Medicinal Chemistry research group)

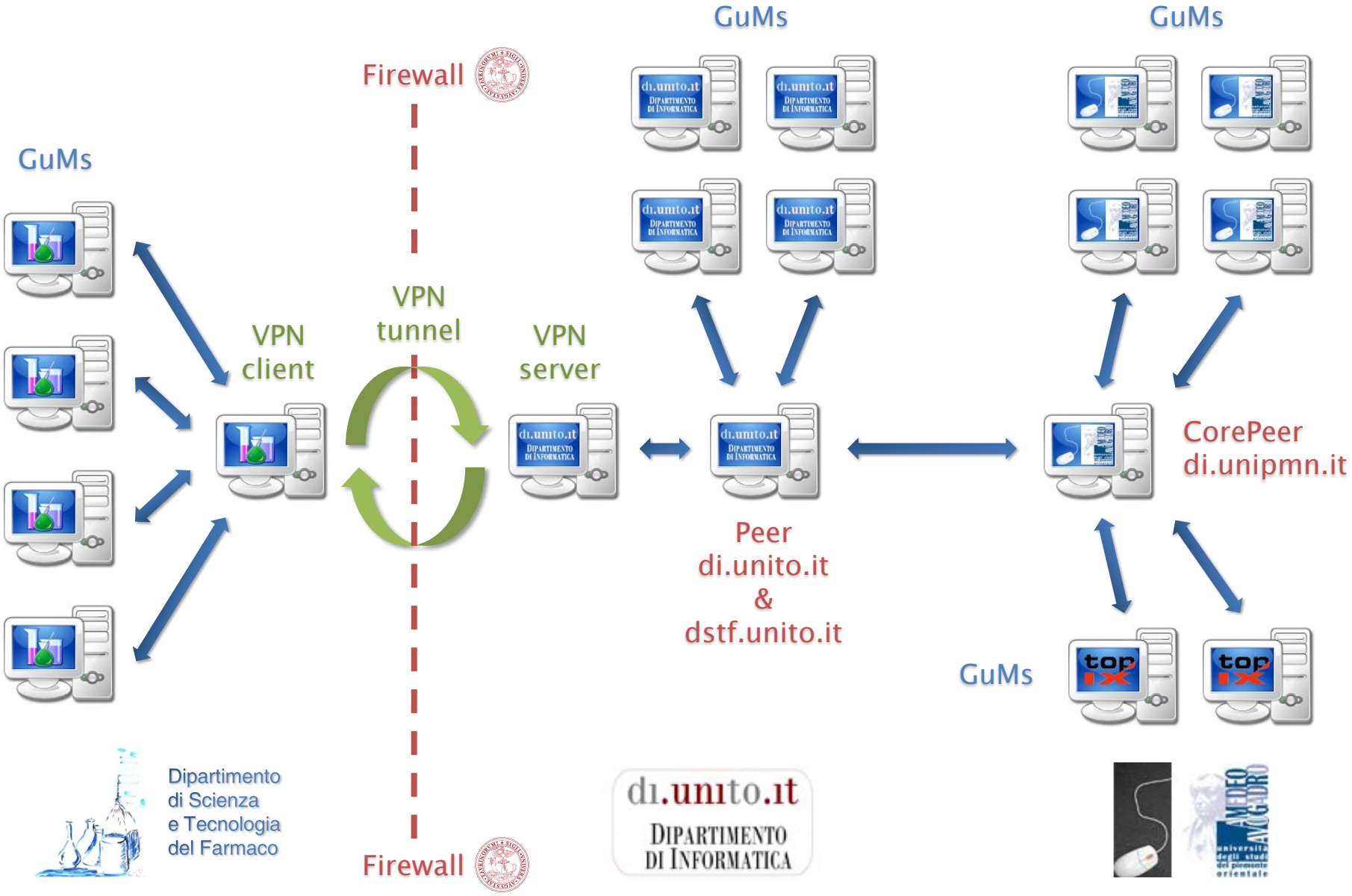
ShareGrid opens new possibilities

Installation of a small Linux partition (CentOS 5.0) on the PCs of the informatics classroom, of the library, as well as on all PCs of the Medicinal Chemistry research group

All PCs are behind a firewall, therefore it is not possible to set up a Peer reachable from the outside

A VPN is provided by the Computer Science Department (Università degli Studi di Torino)

ShareGrid and Medicinal Chemistry research



ShareGrid and MPI applications

MPI Applications

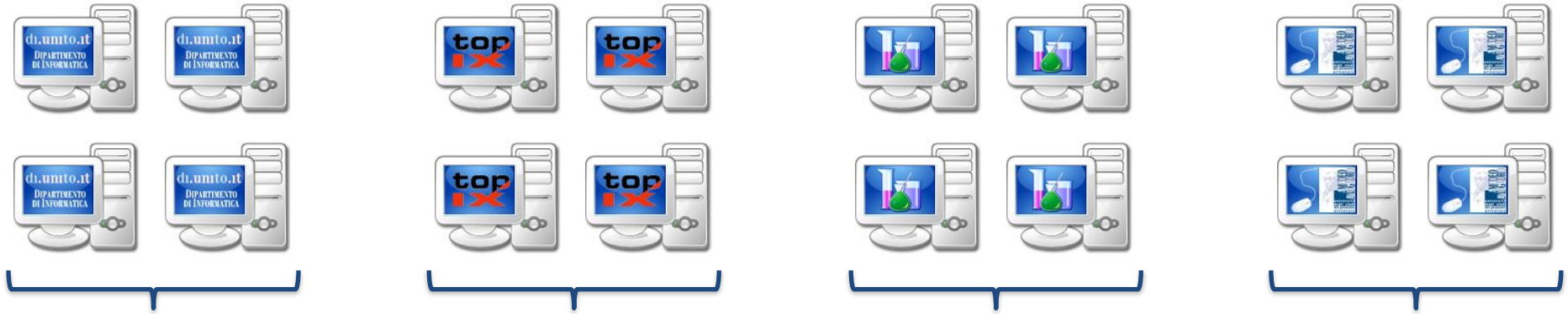
⚠ **MPI Applications is currently unavailable in OurGrid 3.3. If you fell that you shall require the use of MPI Applications, please use OurGrid 3.2.1 . Please note that this feature may not be stable on such version, thus the reason it was removed.**

Both molecular mechanics (AMBER, CHARMM) and quantum-mechanics applications (GAMESS) rely on MPI libraries for parallel execution

workaround

A client-side BASH script looks for partners available to run a MPI computation, and organizes them in sub-clusters

ShareGrid and MPI applications



Max 4 CPUs (↑network overhead on 100 Mbit LAN)

Homogenous architecture inside the sub-cluster

Free access over SSH as well as on a limited TCP/IP port range (applications statically linked with MPICH2 libraries)

Functions implemented in the script

Verification of the availability, in the ShareGrid pool, of *n* free GuMs to form an MPI sub-cluster

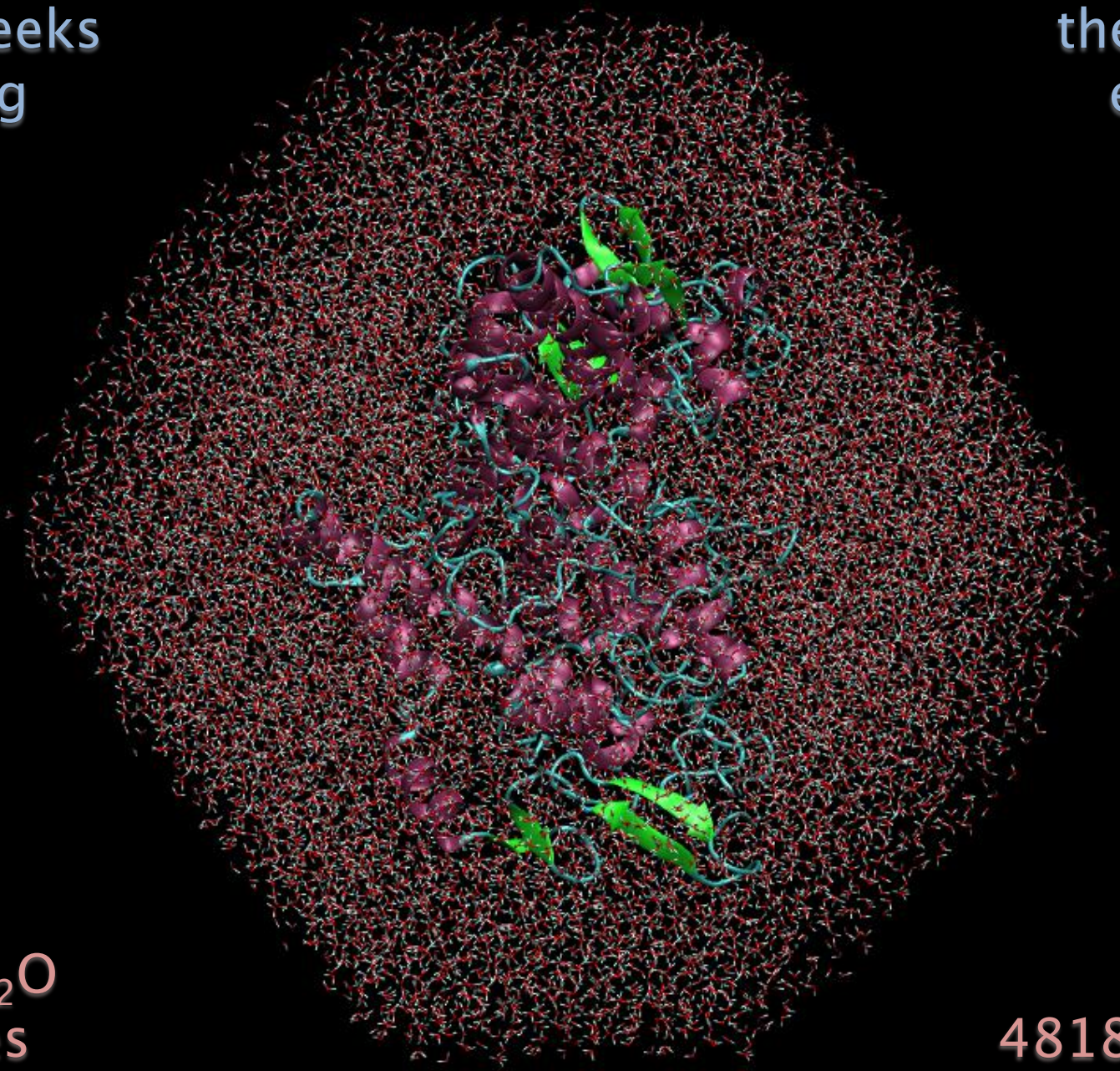
Automatic duplication of put directives included in the .jdf file on all PCs belonging to the sub-cluster if requested by the MPI application

Adherence to ShareGrid policies: even if a job is formally launched on one GuM only, the other GuMs belonging to the sub-cluster run a placeholder job, thus resulting busy and unavailable to accept more jobs

From a three-month simulation
with an implicit solvent model…

…to 2 weeks modelling

the solvent explicitly

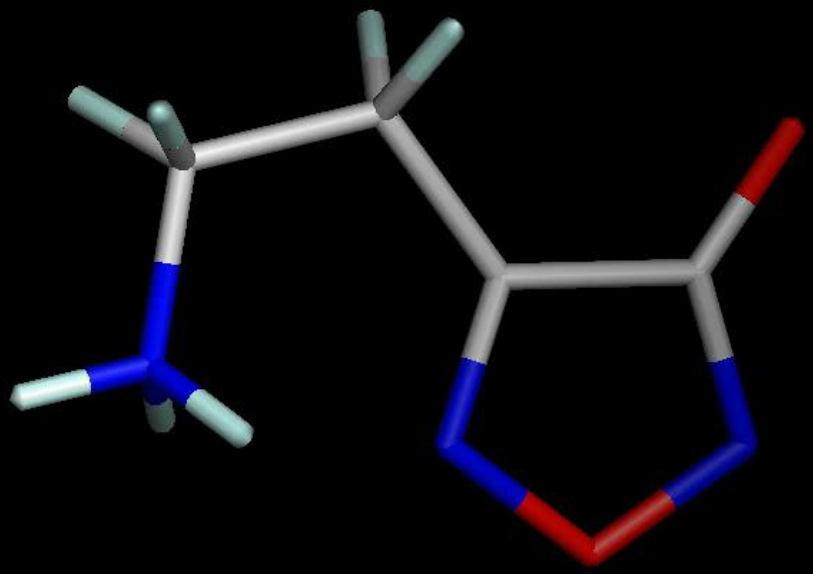


16060 H_2O molecules

that is, 48180 atoms

Click to download movie

Click to download movie



Systematic exploration of conformational space available to a small molecule through *ab initio* quantum-mechanical methods

↓

Continuous variation of the 2 degrees of freedom (torsional angles) by 15° increments, followed by minimization and potential energy computation

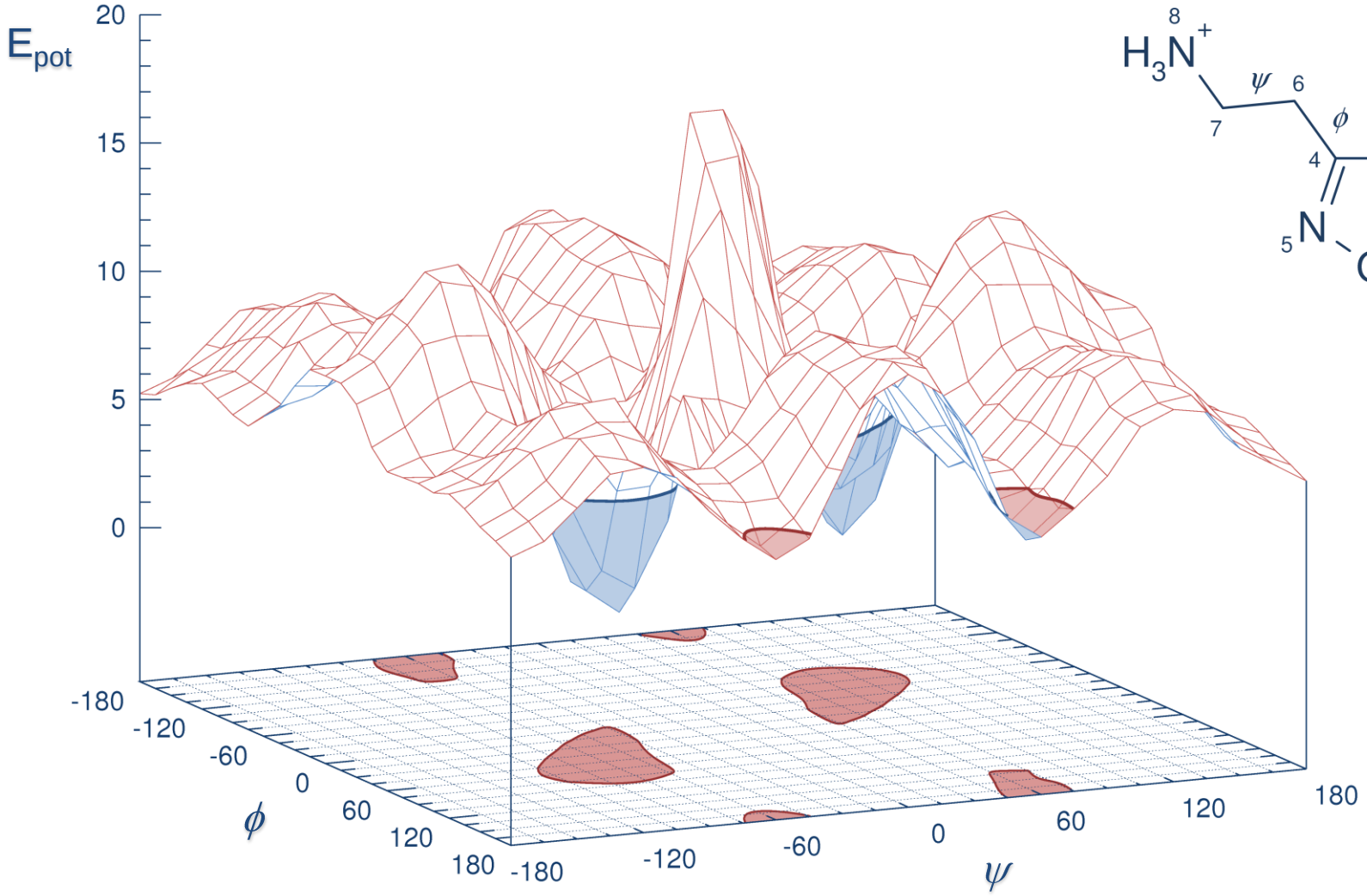
↓

576 conformers

Single CPU: 2 years

Full-power ShareGrid: 3 days

Systematic search through *ab initio* QM methods



Wishes

The continuation of the ShareGrid experience in the future

A more flexible middleware, allowing task monitoring and file transfers while jobs are running

A fault-tolerant CorePeer, so that Peers are not banned anymore following minor networking issues

Thanks

To whom founded and promoted ShareGrid as a working tool for the academic community in Piedmont

To the Computer Science Departments of "Università del Piemonte Orientale" and "Università di Torino" for the technical realisation, the end-user support, and for making available to ShareGrid almost 200 PCs