Molecular Modeling of Proteins: application to cancer immunotherapy

O. Michielin[1,2,3]

(1) Centre Pluridiscipinaire d'oncologie
CHUV, Lausanne, Switzerland

(2) Ludwig Institute for Cancer Research
Epalinges, Switzerland

(3) Swiss Institute of Bioinformatics
Dorigny, Switzerland

Introduction & historical note

Theoretical milestones:

Newton (1643-1727):	Classical equations of motion: $F(t)=m\ a(t)$
Schrödinger (1887-1961):	Quantum mechanical equations of motion: $-ih\ \partial t\ \Psi(t)=H(t)\ \Psi(t)$
Boltzmann(1844-1906):	Foundations of statistical mechanics

Molecular dynamics milestones:

Liquids

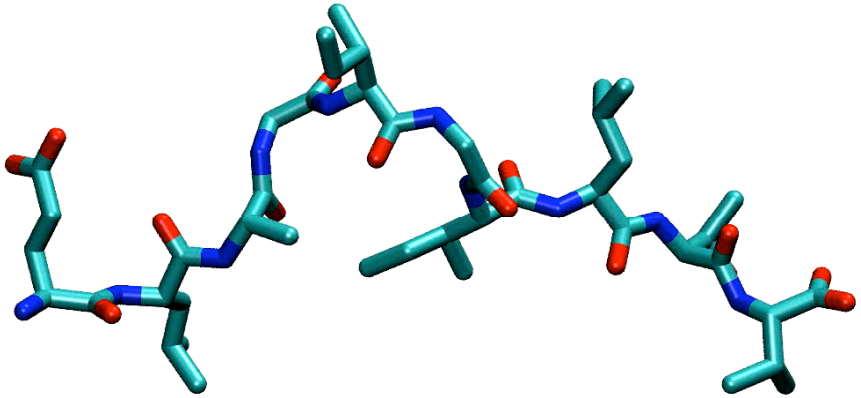
Metropolis (1953):	First Monte Carlo (MC) simulation of a liquid (hard spheres)
Wood (1957):	First MC simulation with Lennard-Jones potential
Alder (1957):	First Molecular Dynamics (MD) simulation of a liquid (hard spheres)
Rahman (1964):	First MD simulation with Lennard-Jones potential

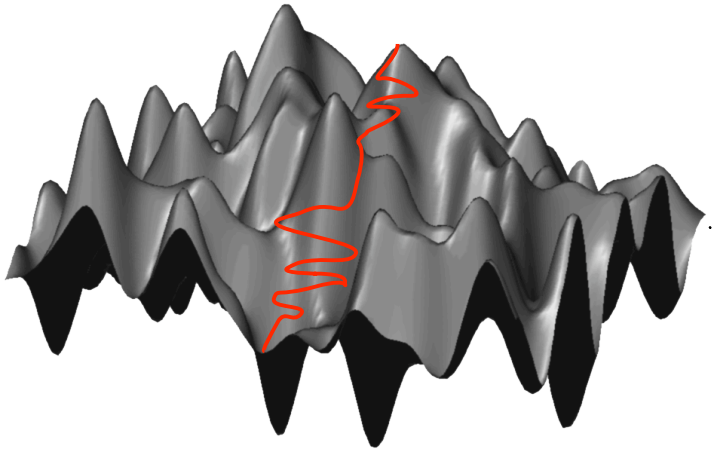
Proteins

Karplus (1977) & McCammon (1977)	First MD simulation of proteins
Karplus (1983):	CHARMM general purpose FF & MD program
Kollman(1984):	AMBER general purpose FF & MD program
Car-Parrinello(1985):	First full QM simulations
Kollmann(1986):	First QM-MM simulations

Molecular Modeling Principles

1) Modeling of molecular interactions





Free energy landscape

2) Simulation of time evolution (Newton)

3) Computation of average values

$$O = \langle O \rangle_{Ensemble} = \langle O \rangle_{Temps}$$ (Ergodicity) → Connection microscopic/ macroscopic

Macroscopic value

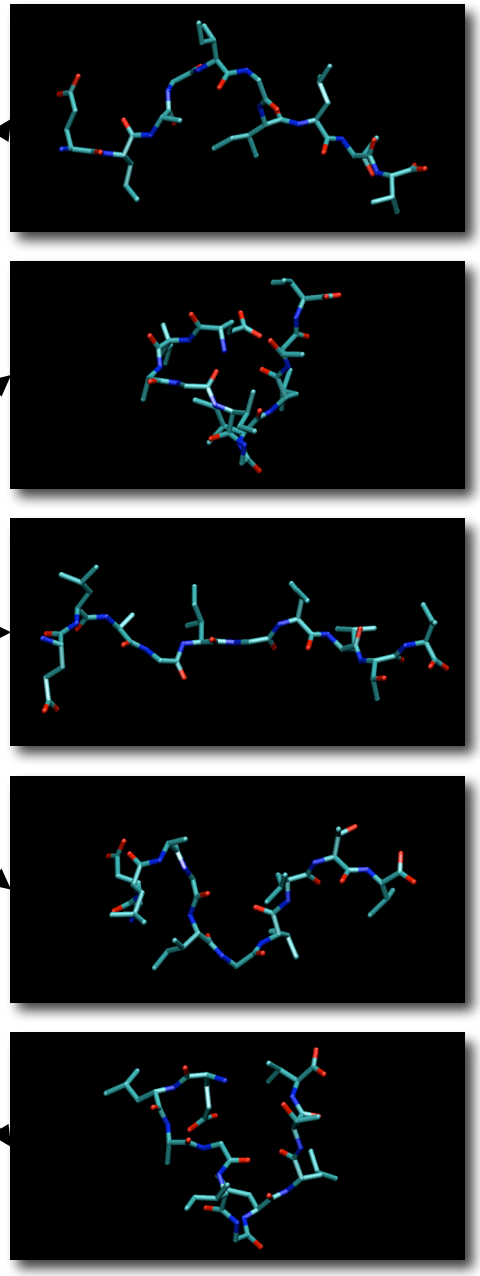
Average simulation value

Connection micro/macroscopic: intuitive view

Expectation value

$$\langle O \rangle = \frac{1}{Z} \sum_i O_i e^{-\beta E i}$$

Where $Z = \sum_i e^{-\beta E i}$
is the partition function



$E_1, P_1 \sim e^{-\beta E1}$

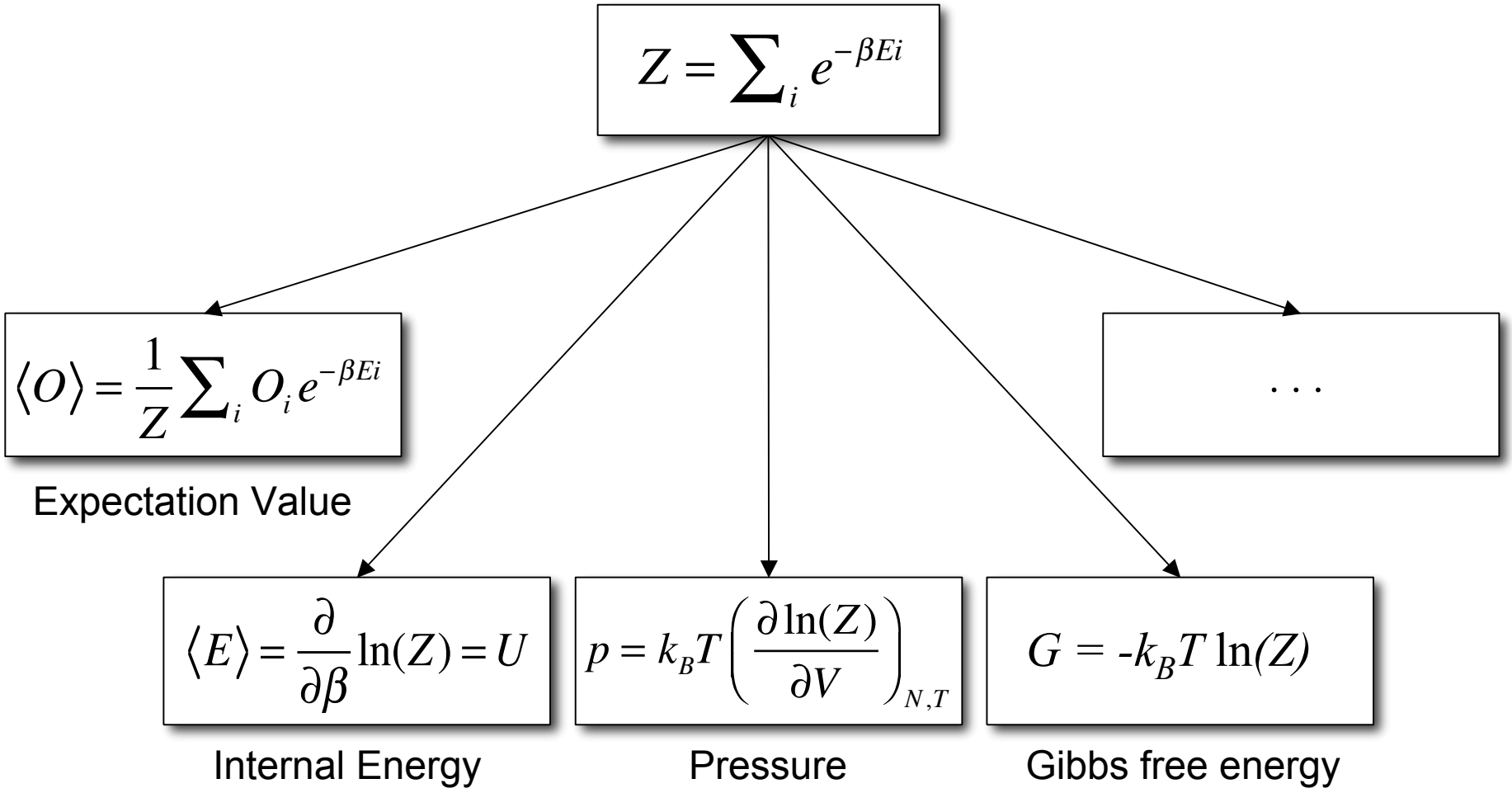
$E_2, P_2 \sim e^{-\beta E2}$

$E_3, P_3 \sim e^{-\beta E3}$

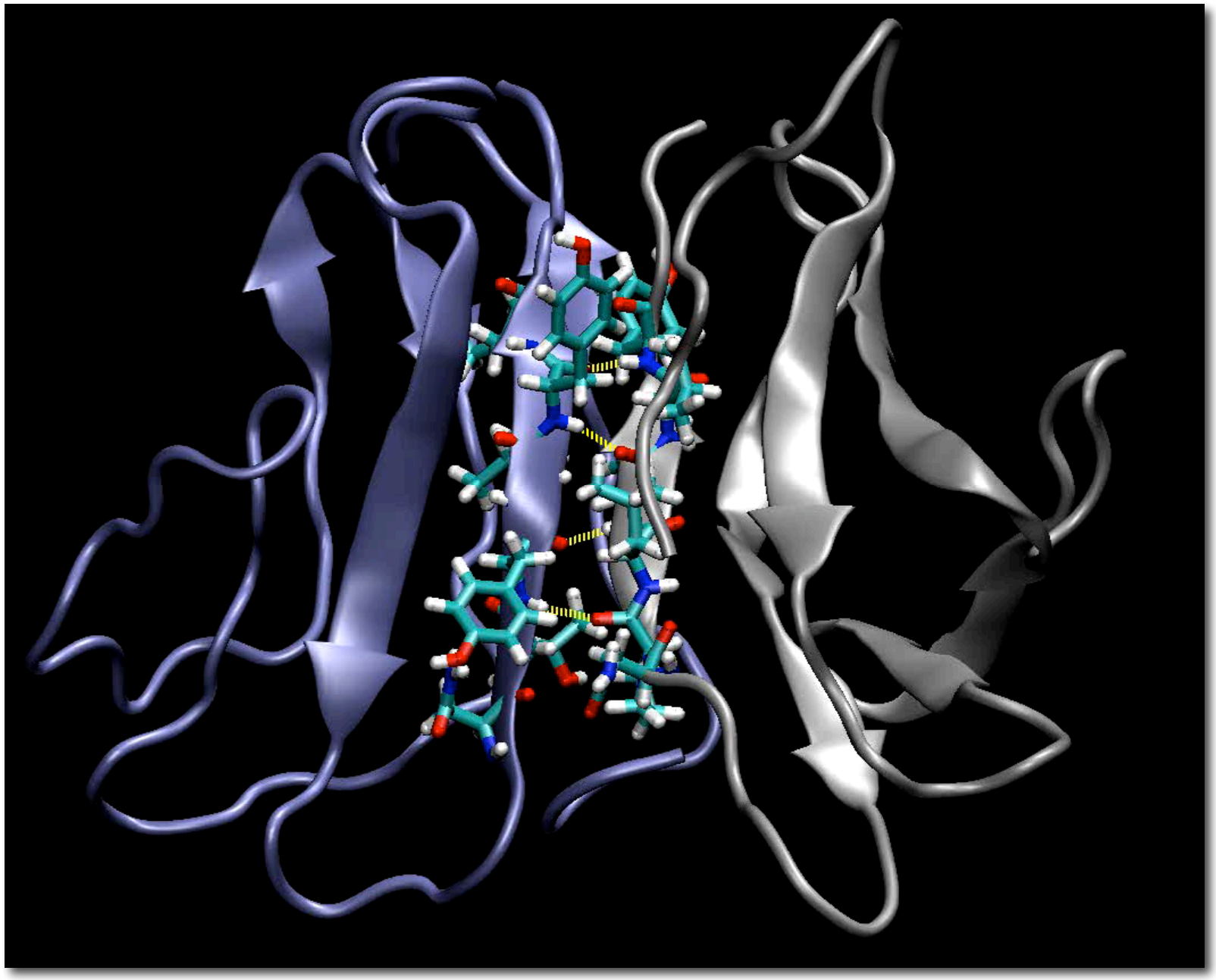
$E_4, P_4 \sim e^{-\beta E4}$

$E_5, P_5 \sim e^{-\beta E5}$

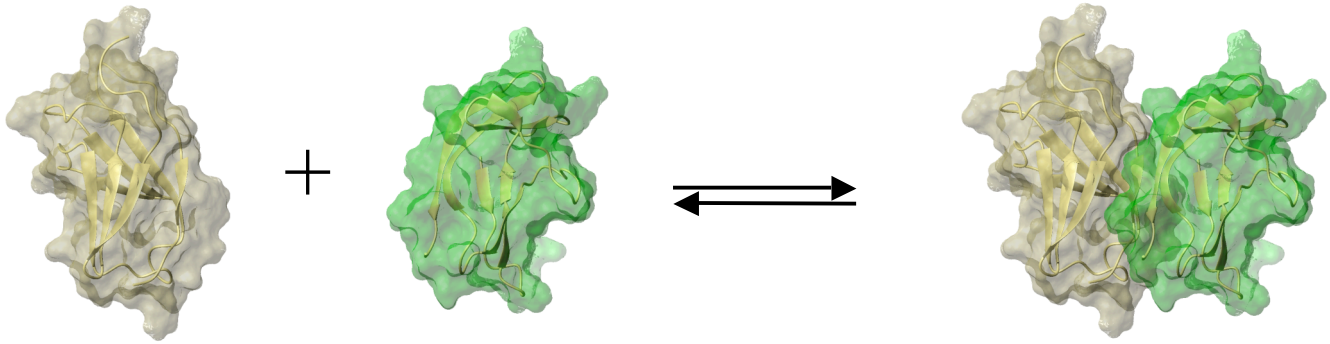
Central Role of the Partition Function



Dynamical aspects of molecular recognition



Free energy: classical definition



The free energy is the energy left for once you paid the tax to entropy:

$$\Delta G = \Delta H - T\Delta S$$

Enthalpic

- Hydrogen bonds
- Polar interactions
- Van der Waals interactions
- ...

Entropic

- Loss of degrees of freedom
- Gain of vibrational modes
- Loss of solvent/protein structure
- ...

Theoretical Predictions:

- *Approximate:* empirical formula for all contributions
- *Exact:* using statistical physics definition of *G*

Free energy: computational approaches

$$\Delta G = G_A - G_B = -k_B T \ln\left(\frac{Z_A}{Z_B}\right)$$

Free energy simulations techniques aim at computing ratios of partition functions using various techniques.

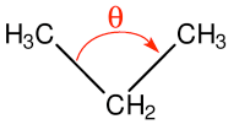
$$Z = \sum_i e^{-\beta Ei}$$

Sampling of important microstates of the system (MD, MC, GA, …)

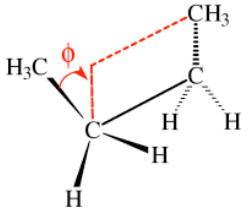
Computation of energy of each microstate (force fields, QM, CP, …)

The CHARMM Force Field

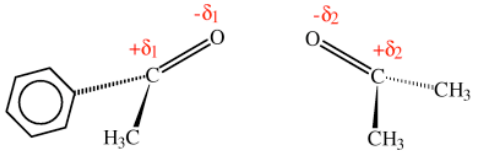
$$V = \sum_{\text{Bonds}} K_b (b - b_0)^2 + \sum_{\text{Angles}} K_\theta (\theta - \theta_0)^2$$



$$+ \sum_{\text{Impropers}} K_\omega (\omega - \omega_0)^2$$

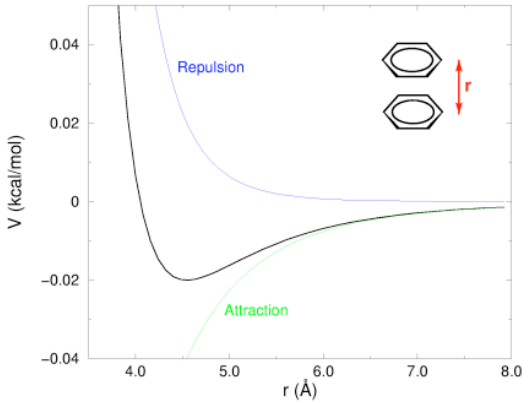


$$+ \sum_{\text{Dihedrals}} K_\phi \left[1 - \cos(n_\phi \phi - \delta_\phi)\right]$$



$$+ \sum_{i>j} \frac{q_i q_j}{4\pi\varepsilon} \frac{1}{r_{i,j}}$$

$$+ \sum_{i>j} 4\varepsilon_{ij} \left[(\sigma_{ij}/r_{ij})^{12} - (\sigma_{ij}/r_{ij})^{6}\right]$$



MD Techniques: Microcanonical sampling

For an Hamiltonian of the form $H(\boldsymbol{p},\boldsymbol{r}) = \sum_{i=1}^{3N} \frac{p_i^2}{2m_i} + \phi(r_1,\ldots,r_{3N})$

in cartesian coordinates, the Hamilton equations of motion reduce to the Newton equations

$$\dot{r}_i = \frac{p_i}{m_i} \qquad m_i a_i = \frac{-\partial}{\partial r_i}\phi(\boldsymbol{r}) \qquad i=1, \ldots, N$$

Several numerical methods have been developed to integrate these equations. One of the most stable integrator is that of ***Verlet***: for a small time increment $\mathrm{d}t$, one can use a Taylor expansion of the function ***r(t)***:

$$\boldsymbol{r}_i(t+\delta t) = \boldsymbol{r}_i(t) + \boldsymbol{v}_i(t)\,\delta t + 1/2\,\boldsymbol{a}_i(t)\,\delta t^2 + \ldots$$
$$\boldsymbol{r}_i(t-\delta t) = \boldsymbol{r}_i(t) - \boldsymbol{v}_i(t)\,\delta t + 1/2\,\boldsymbol{a}_i(t)\,\delta t^2 + \ldots$$

Adding those equations, one gets ***r(t+***$\mathrm{d}t$***)*** as a function of ***r(t)*** and ***r(t-***$\mathrm{d}t$***)***.

$$\boldsymbol{r}_i(t+\delta t) = 2\,\boldsymbol{r}_i(t) - \boldsymbol{r}_i(t-\delta t) + \boldsymbol{a}_i(t)\,\delta t^2$$

- In practice, this scheme is applied iteratively, starting from the initial conditions.
- Velocities are postcomputed as ***v(t)*** = ***[r(t+***$\mathrm{d}t$***)-r(t-***$\mathrm{d}t$***)] / 2***$\mathrm{d}t$.
- Positions are correct up to $\mathrm{d}t^4$ and velocities to $\mathrm{d}t^2$.
- This scheme conserves energy with very good accuracy.

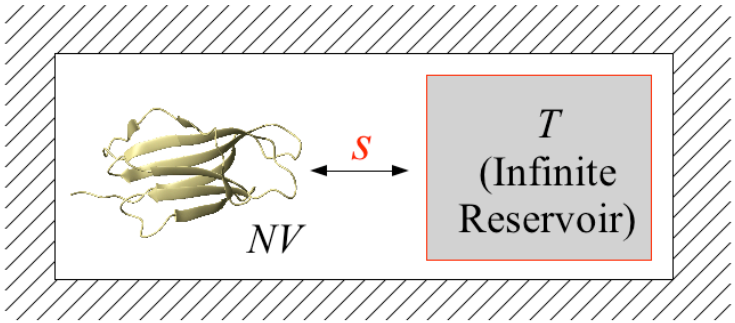
MD Techniques: Sampling of the various ensembles

1) Microcanonical ensemble (constant N, V, E)

$$H(\boldsymbol{p},\boldsymbol{q}) = \sum_{i}^{N} \frac{\boldsymbol{p}_i^2}{2 m_i s^2} + \phi(\boldsymbol{q})$$

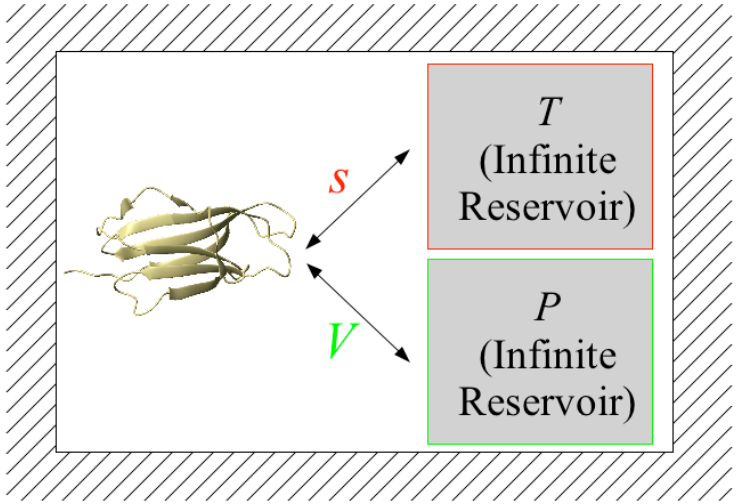
2) Canonical ensemble (constant N, V, T)

$$H(\boldsymbol{p},\boldsymbol{q},p_s,s) = \sum_{i}^{N} \frac{\boldsymbol{p}_i^2}{2 m_i s^2} + \phi(\boldsymbol{q}) + \frac{p_s^2}{2Q} + (3N+1)\,kT \ln s$$

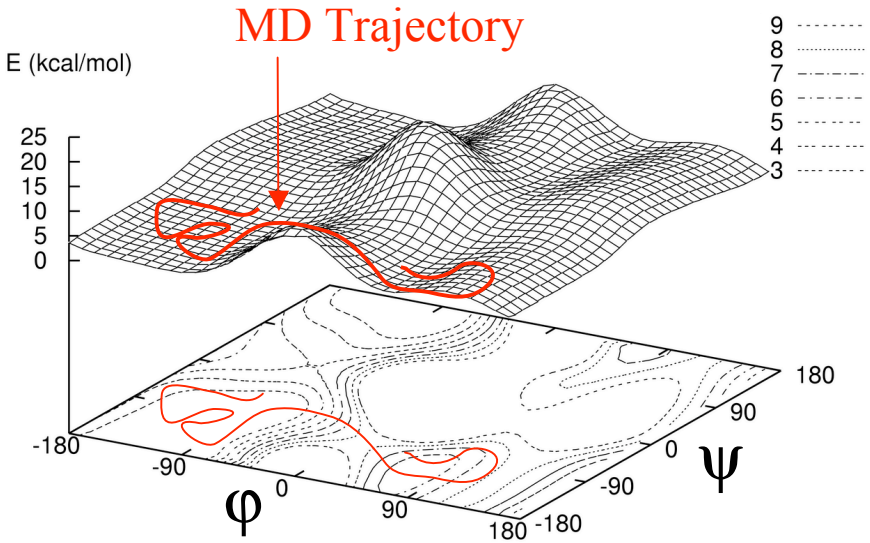


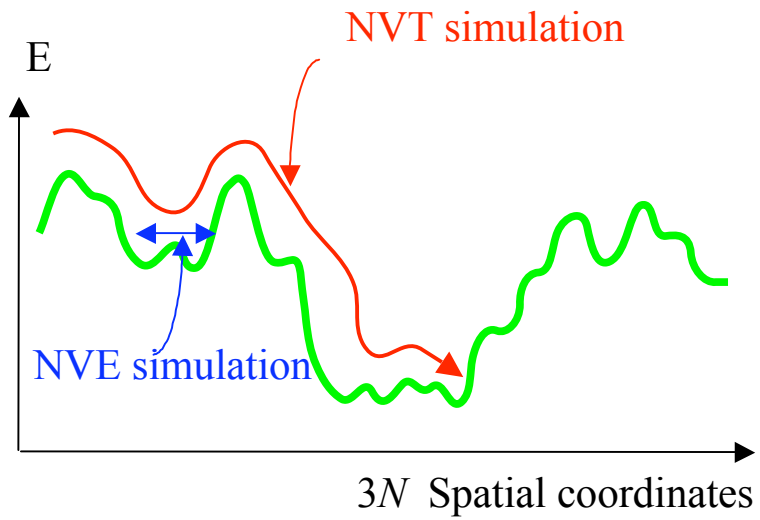
3) Isothermic-isobaric ensemble (constant N, P, T)

$$H = \sum_{i}^{N} \frac{\boldsymbol{p}_i^2}{2 m_i s^2 V^{2/3}} + \phi(V^{1/3}\boldsymbol{q}) + \frac{p_s^2}{2Q} + (3N+1)\,kT \ln s + \frac{p_V^2}{2W} + P_{ex} V$$

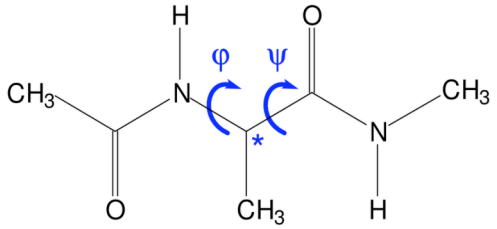


Ergodic Hypothesis

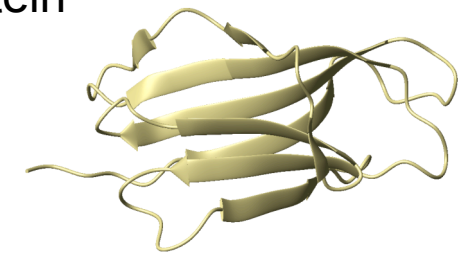




"Alanine"



Protein



$$\langle O \rangle_{Ensemble} = \frac{1}{Z} \int O(\varphi,\psi) e^{-\beta E(\varphi,\psi)} d\varphi d\psi \overset{?}{=} \frac{1}{\tau} \int_0^{\tau} O(t) dt = \langle O \rangle_{Time}$$

Free energy calculation: Main approaches

Sampling, Exact

Free Energy Perturbation (FEP)

$$\Delta G = -k_B T \ln < exp(-\beta \Delta V) >$$

Thermodynamical Integration (TI)

$$\Delta G = \int_0^1 \left\langle \frac{\partial V}{\partial \lambda} \right\rangle_\lambda d\lambda$$

Non Equilibrium Statistical Mechanics (Jarzynski)

$$\Delta G = -k_B T \ln < exp(-\beta W) >$$

Sampling, Approx.

Linear Interaction Energy (LIE)

$$\Delta G = \alpha \Delta \left\langle V_{l\ env}^{VdW} \right\rangle + \beta \Delta \left\langle V_{l\ env}^{Elec} \right\rangle$$

Molecular Mechanics/Poisson-Boltzmann/Surface area (MM-PBSA)

$$\Delta G = \langle \Delta G_{Gas} \rangle + \left\langle \Delta G_{Desolv}^{PBSA} \right\rangle - \langle T \Delta S \rangle$$

Approx.

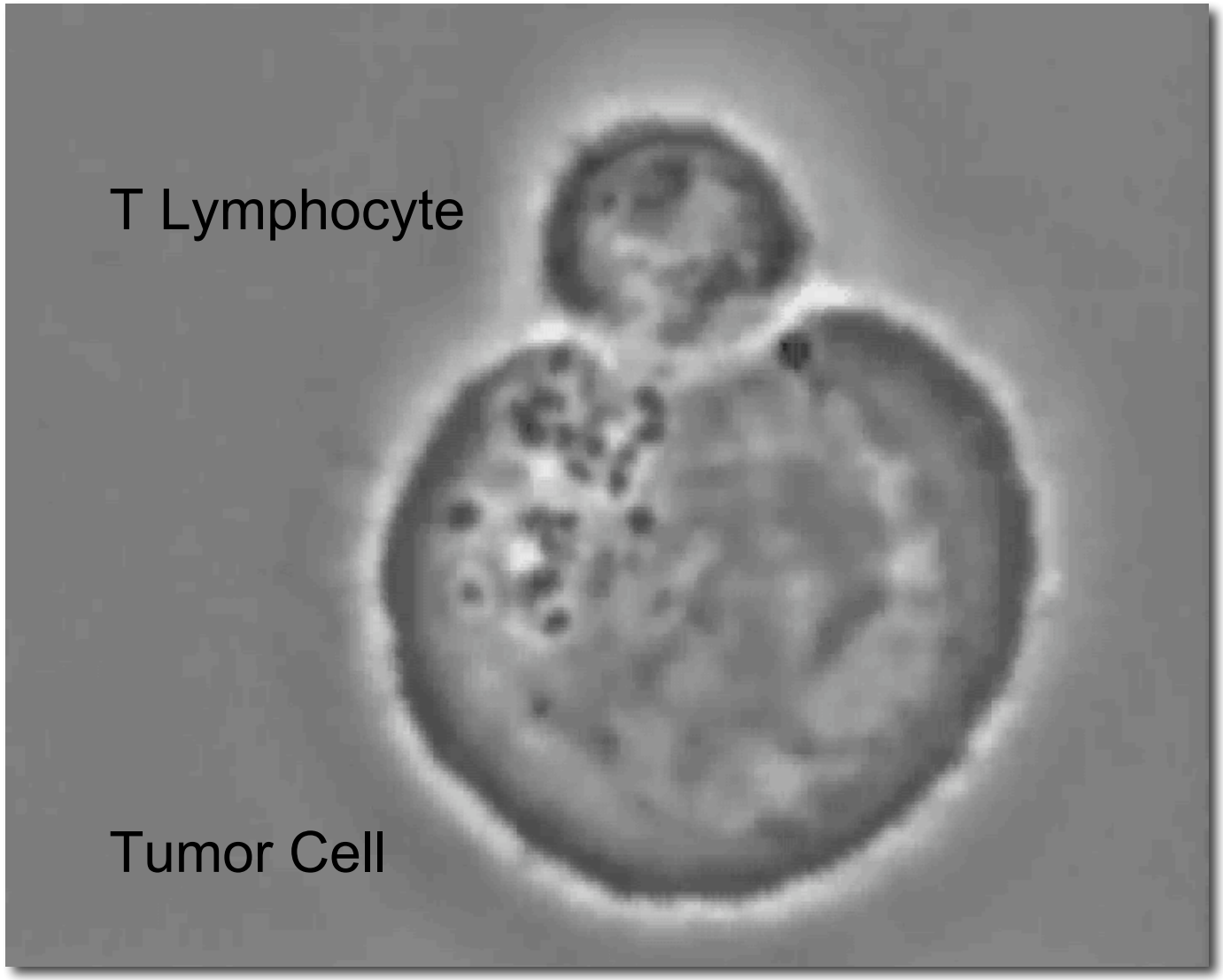
Quantitative Structure Activity Relationship (QSAR)

$$\Delta G = k_0 + \sum k_i X_i \qquad \Delta G = F(X)$$

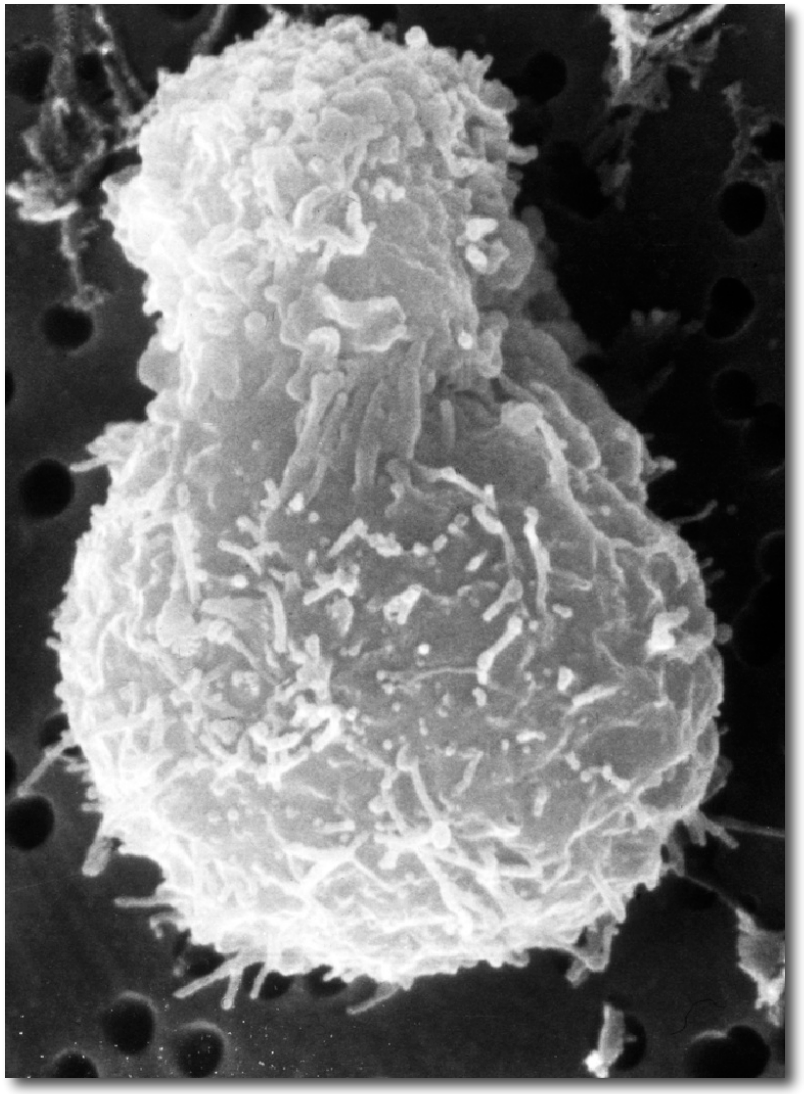
(***X*** is a descriptor)

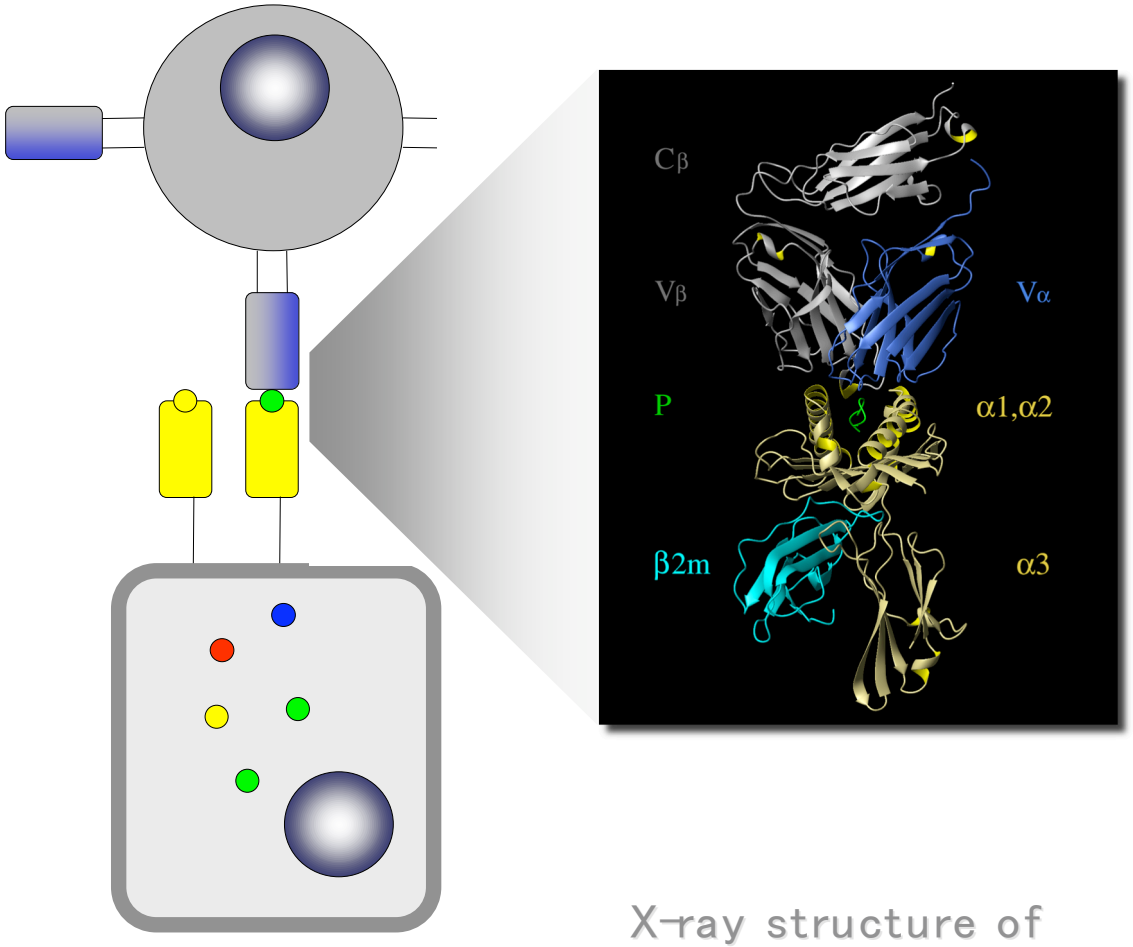
CPU Time

Medical background: Cytotoxic activity of T lymphocytes



Tumor cell recognition by CD8+ T cells: the TCR-p-MHC complex

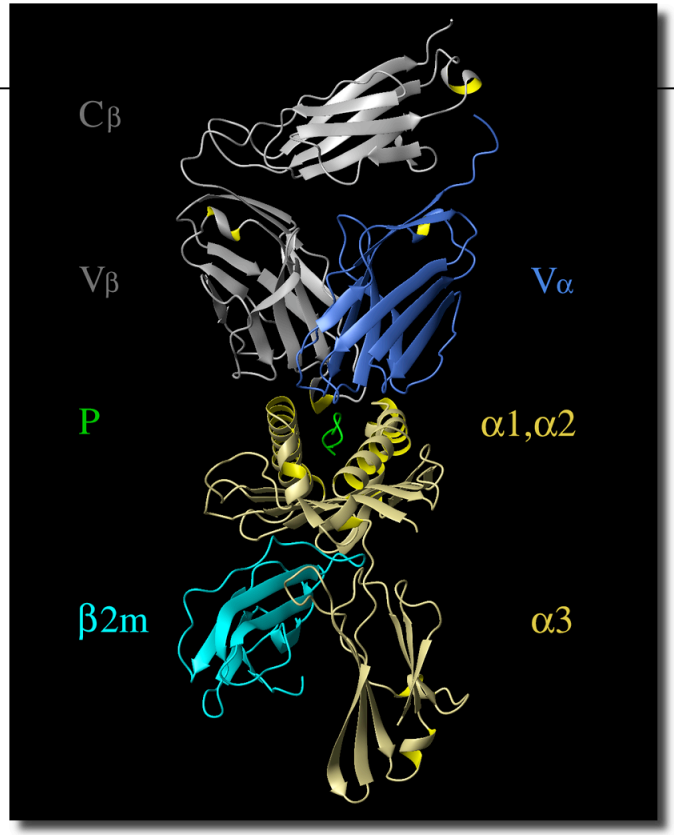




X-ray structure of bound TCR-p-MHC

Goals of the molecular modeling approach

Peptide vaccine optimization



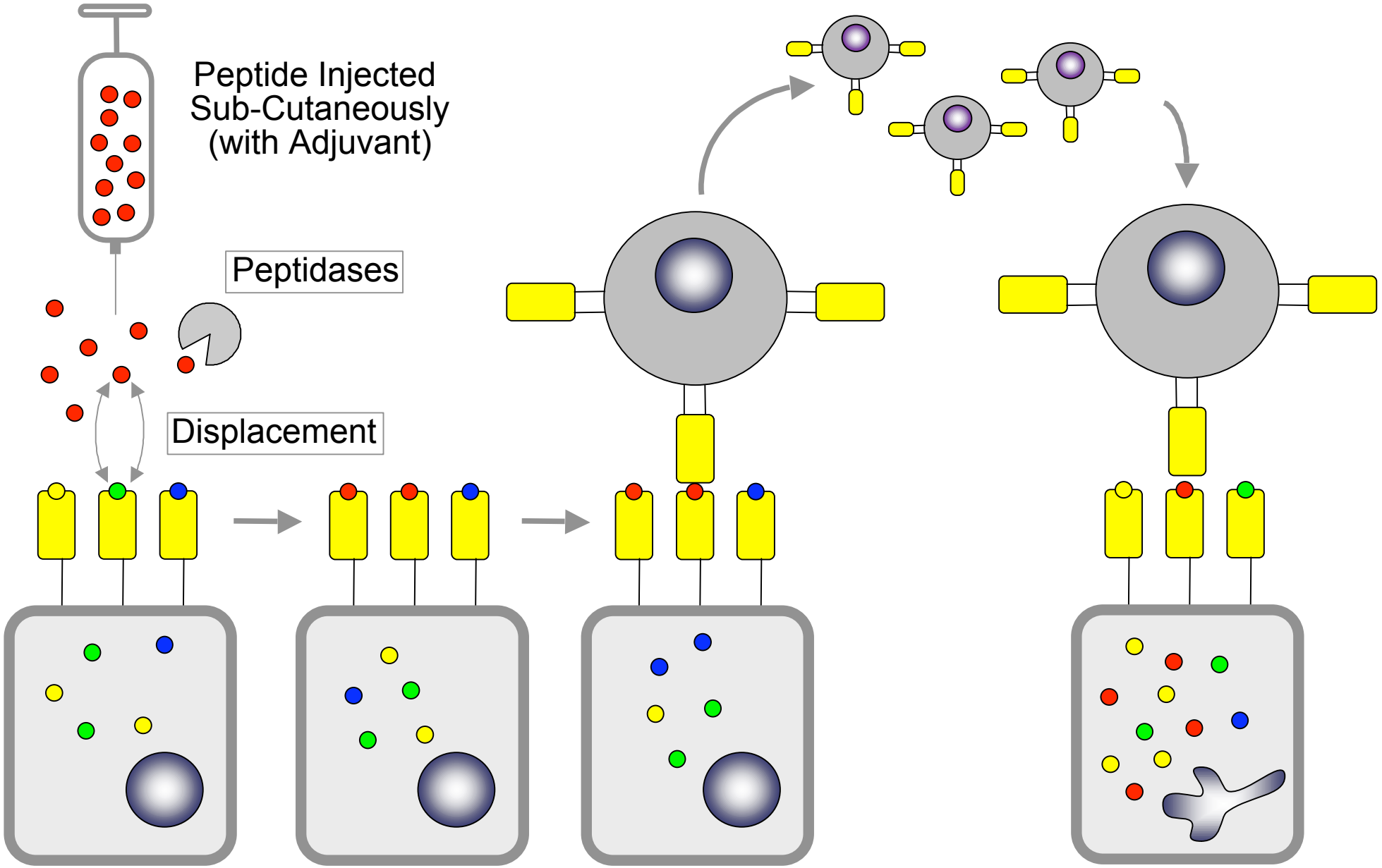
TCR sequence optimization

Optimized peptide vaccination

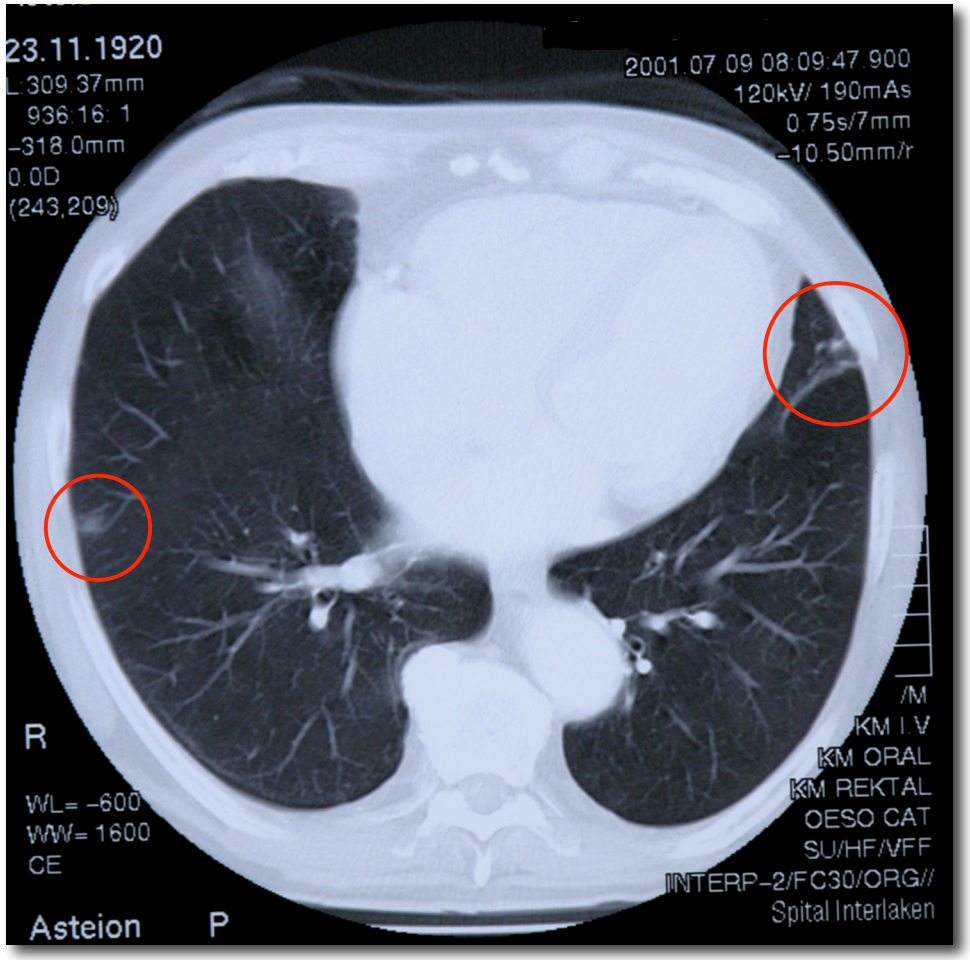
Adoptive Immunotherapy

Clinical Trials

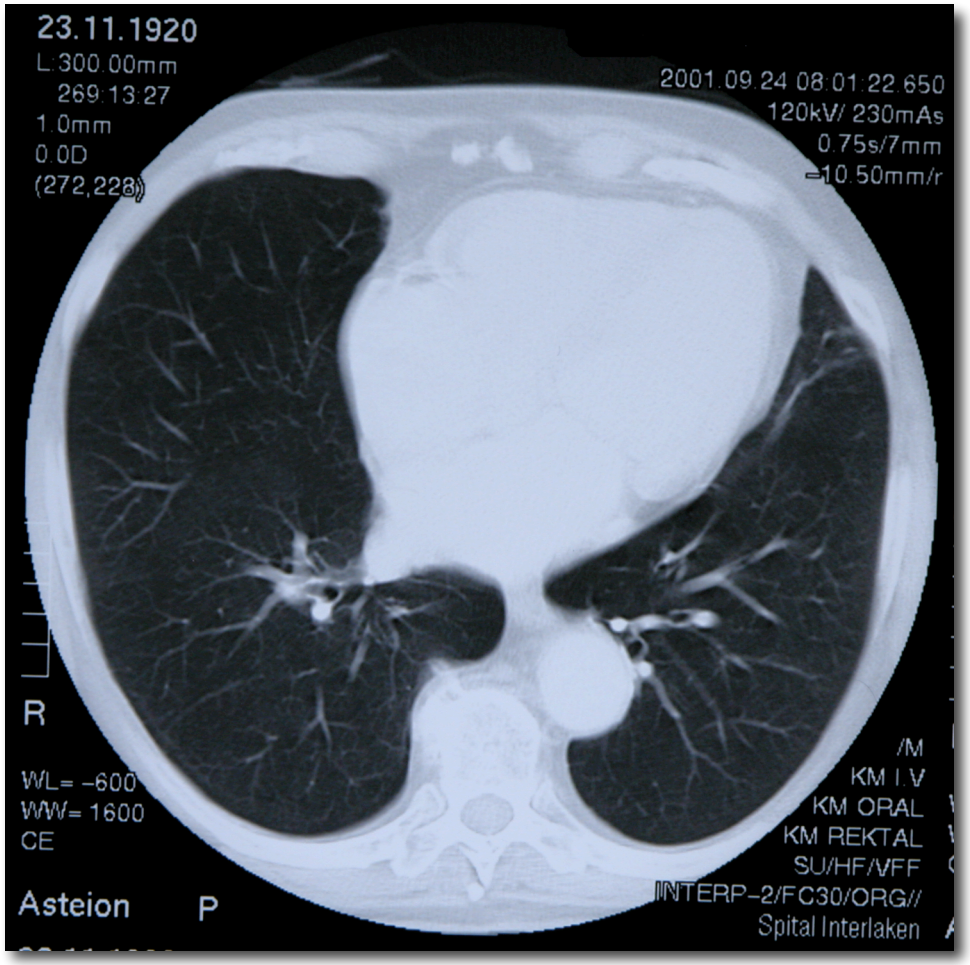
Principles of peptide based immunotherapy



Regression of pulmonary melanoma metastases after vaccination with Melan-A peptide (patient LAU 446)

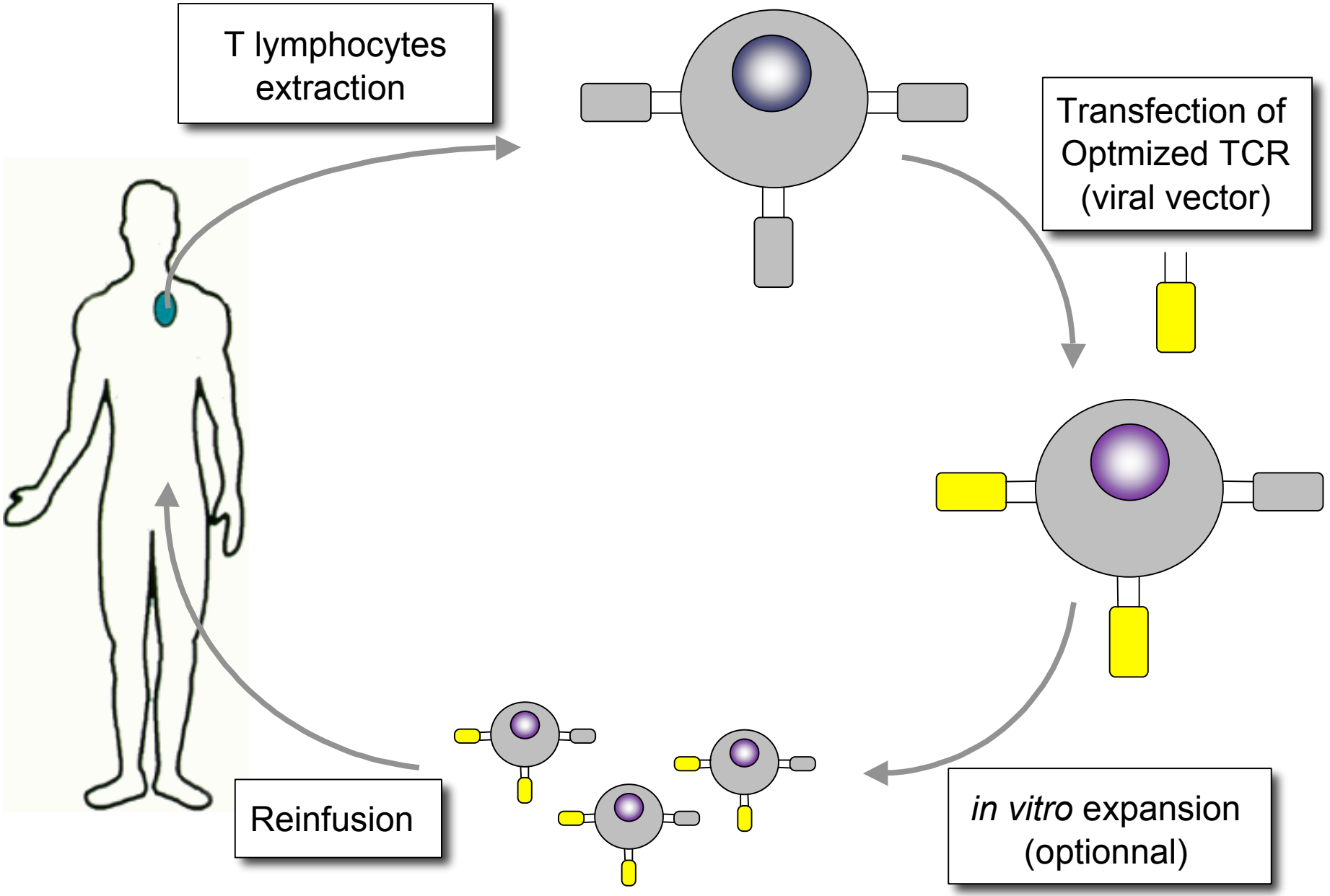


July 9, 2001
< 0.1 % of Melan-A specific
CD8+ T cells in PBL

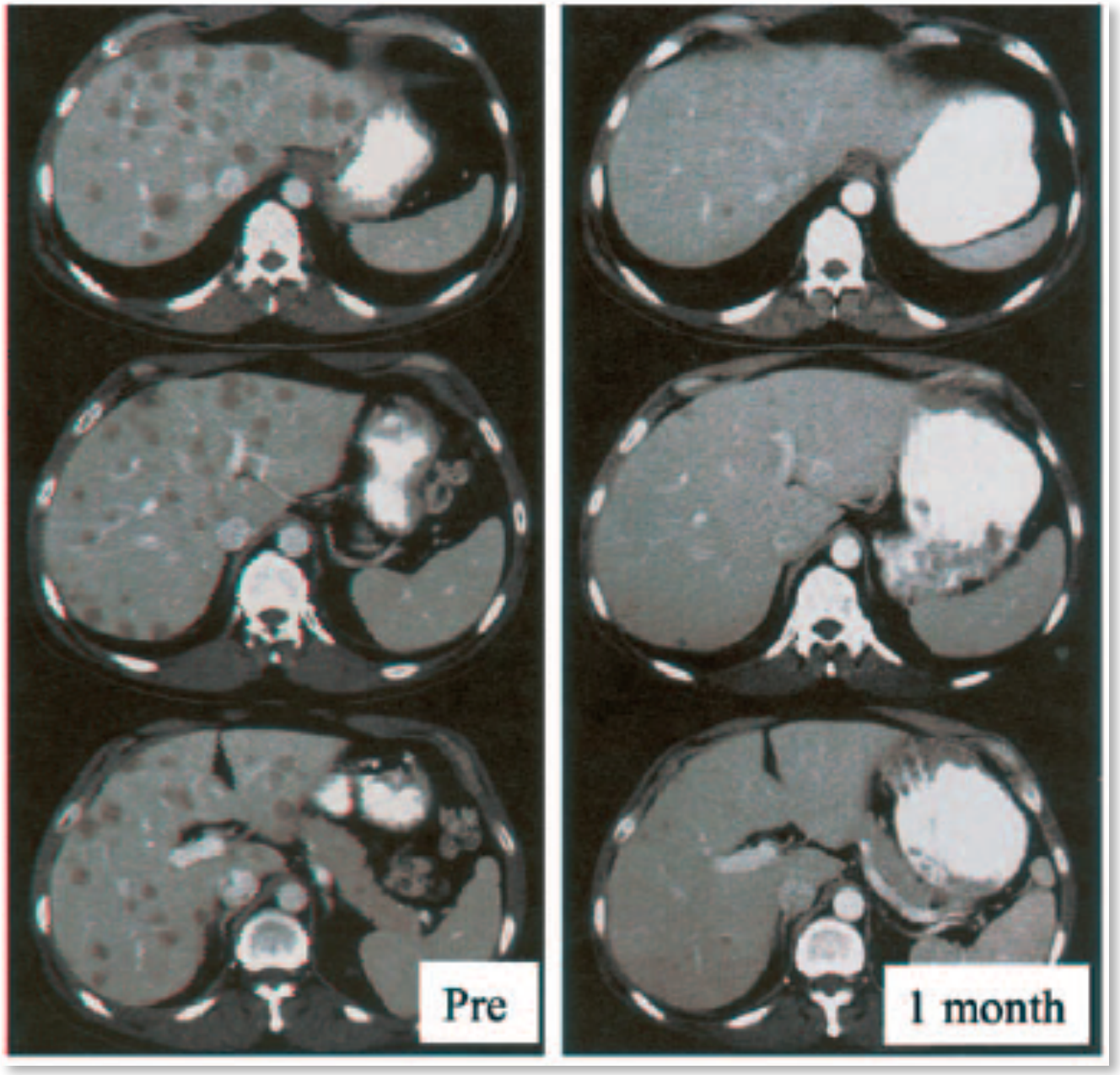


September 24, 2001
0.3 % of Melan-A specific
CD8+ T cells in PBL

Immunotherapy using adoptive transfert



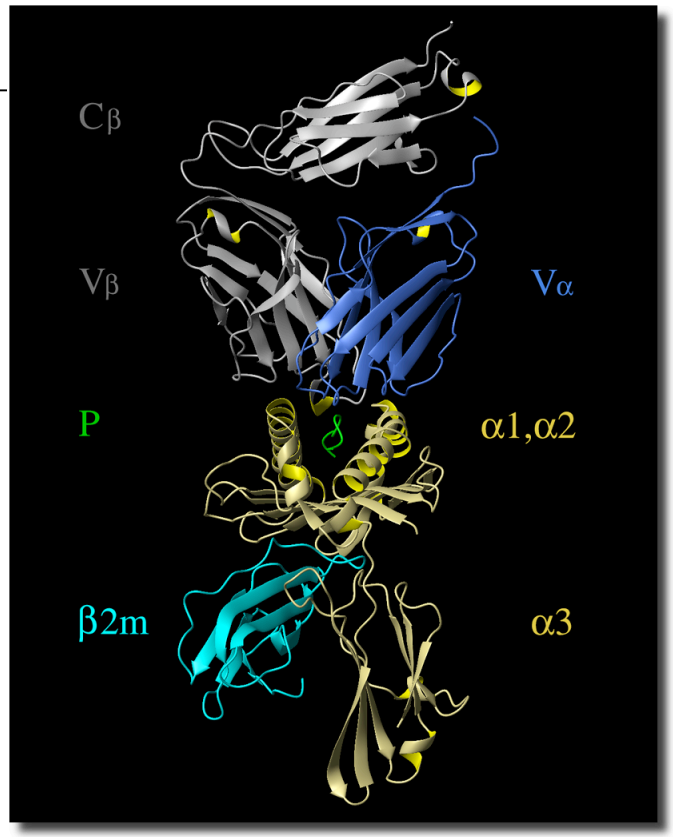
Lymphodepletion combined with adoptive transfert



Dudley & al,
JCO 2005

Goals of the molecular modeling approach

Peptide vaccine optimization



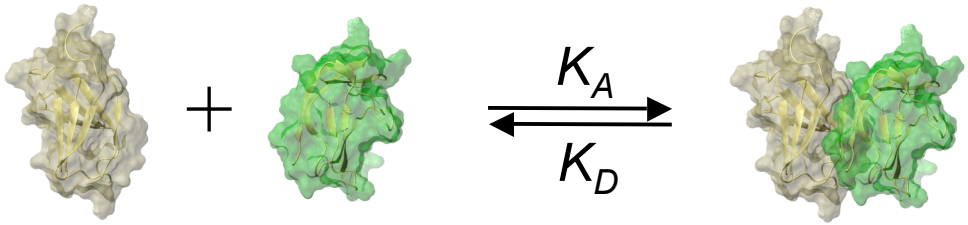
TCR sequence optimization

Optimized peptide vaccination

Adoptive Immunotherapy

Clinical Trials

Free energy calculations:



$$e^{-\Delta G/RT} = K_A$$

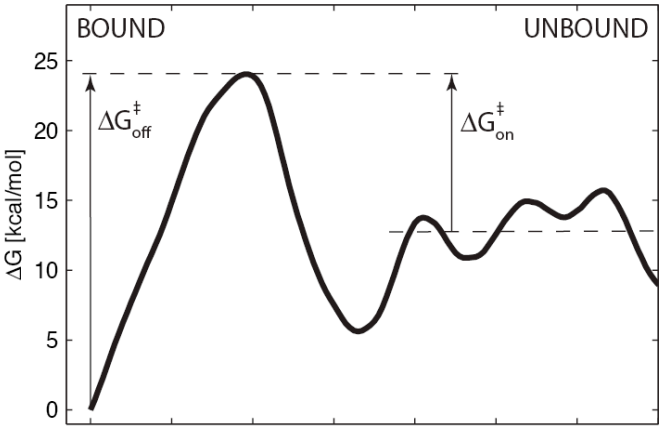
Free Energy → Association Constant

Microscopic Structure → Biological function

Absolute binding free energies: $\Delta G \rightarrow K_A$

Relative binding free energies: $\Delta\Delta G \rightarrow K_{A'}/K_A$

Binding free energy profiles: $\Delta G(\xi) \rightarrow K_A, K_{on}, K_{off}$



Free energy calculation: Main approaches

Sampling, Exact

Free Energy Perturbation (FEP)

$$\Delta G = -k_B T \ln < exp(-\beta \Delta V) >$$

Thermodynamical Integration (TI)

$$\Delta G = \int_0^1 \left\langle \frac{\partial V}{\partial \lambda} \right\rangle_\lambda d\lambda$$

Non Equilibrium Statistical Mechanics (Jarzynski)

$$\Delta G = -k_B T \ln < exp(-\beta W) >$$

Sampling, Approx.

Linear Interaction Energy (LIE)

$$\Delta G = \alpha \Delta \left\langle V_{l\ env}^{VdW} \right\rangle + \beta \Delta \left\langle V_{l\ env}^{Elec} \right\rangle$$

Molecular Mechanics/Poisson-Boltzmann/Surface area (MM-PBSA)

$$\Delta G = \langle \Delta G_{Gas} \rangle + \left\langle \Delta G_{Desolv}^{PBSA} \right\rangle - \langle T \Delta S \rangle$$

Approx.

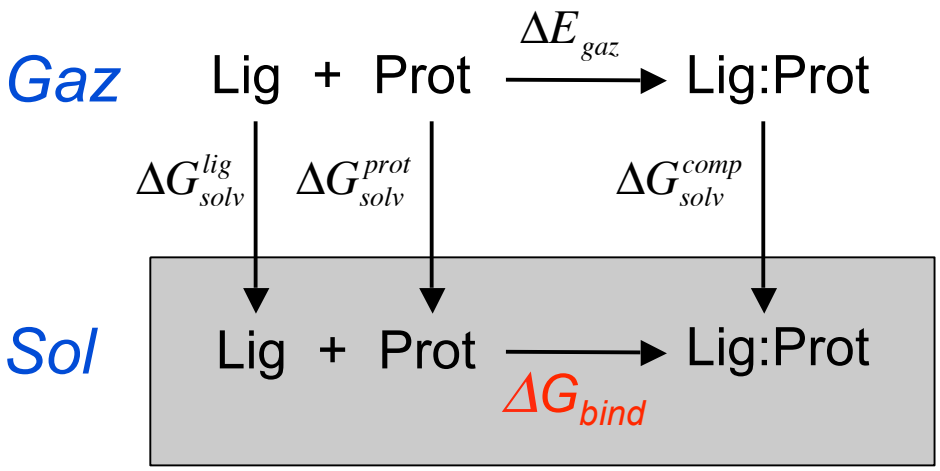
Quantitative Structure Activity Relationship (QSAR)

$$\Delta G = k_0 + \sum k_i X_i \qquad \Delta G = F(X)$$

(***X*** is a descriptor)

CPU Time

Binding free energy decomposition: MM-PBSA, MM-GBSA



Averaged over an MD simulation trajectory of the complex (and isolated parts)

$$\Delta G_{bind} = \langle \Delta E_{gaz} \rangle + \langle \Delta G_{desolv} \rangle - T\langle \Delta S \rangle$$

$$E_{gaz} = E_{elec} + E_{vdw} + \Delta E_{\text{int}\,ra}$$

$$\Delta G_{desolv} = \Delta G_{solv}^{comp} - \left(\Delta G_{solv}^{lig} + \Delta G_{solv}^{prot}\right)$$

$$-T\Delta S = -T(S^{comp} - (S^{prot} + S^{lig}))$$

$$S = S_{trans} + S_{rot} + S_{vib}$$

B. Tidor and M. Karplus, *J. Mol. Biol.*, **1994**, *238*, 405

$$\Delta G_{solv} = \Delta G_{solv,elec} + \Delta G_{solv,np}$$

$$\Delta G_{desolv} = \Delta G_{solv,elec}^{comp} - \left(\Delta G_{solv,elec}^{lig} + \Delta G_{solv,elec}^{prot}\right) + \sigma\left(SASA^{comp} - \left(SASA^{lig} + SASA^{prot}\right)\right)$$

Depending on the way $\Delta G_{solv,elec}$ is calculated:

Molecular mechanics – Poisson-Boltzmann Surface Area (MM- PBSA)

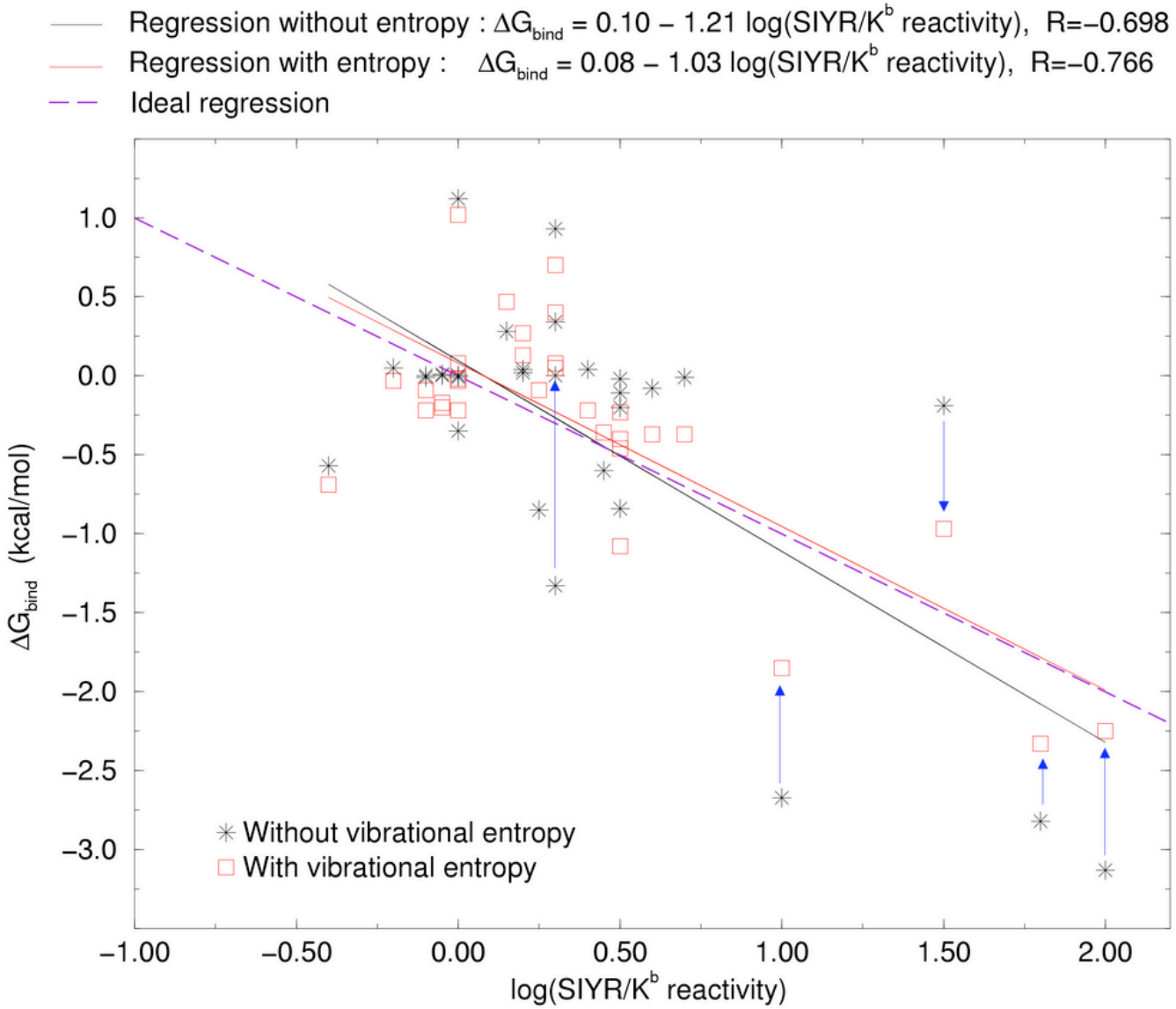
J. Srinivasan, P.A. Kollmann *et al.*, *J. Am. Chem. Soc.*, **1998**, *120*, 9401

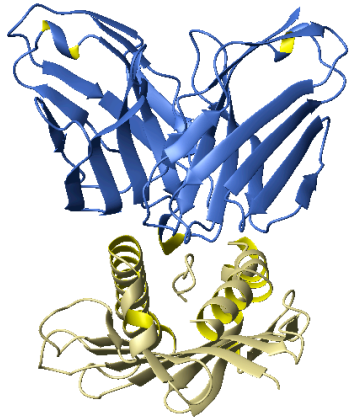
Molecular mechanics – Generalized Born Surface Area (MM- GBSA)

H. Gohlke, C. Kiel and D.A. Case, *J. Mol. Biol.*, **2003**, *330*, 891

MM-GBSA Method: application to TCR-p-MHC

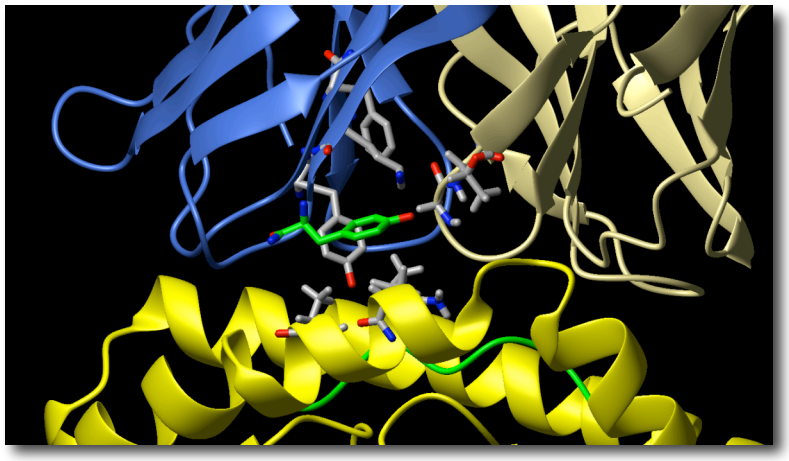
$$\Delta G_{bind} = \langle \Delta E_{gaz} \rangle + \langle \Delta G_{desolv} \rangle - T\langle \Delta S \rangle$$

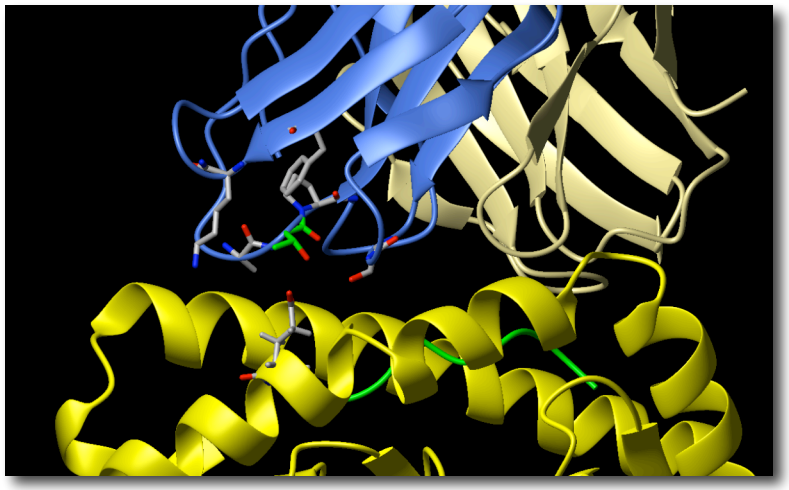




Examples of TCR optimization: 2C TCR

Residue	Domain	$\langle E^{sc}_{vdW} \rangle$	$\langle E^{sc}_{elec} \rangle$	$\langle \Delta G^{sc}_{elec,solv} \rangle$	$\langle \Delta G^{sc}_{np,solv} \rangle$	$-\langle TS^{sc}_{vib} \rangle$	$\langle \Delta G^{sc}_{bind} \rangle$
Ser93	CDR3	0.63	-9.06	3.04	-0.01	0.17	-5.23
Phe100	CDR3	-3.59	-0.74	1.36	-0.72	0.65	-3.04
Tyr31	CDR1	-3.46	-2.37	3.31	-0.61	0.75	-2.38
Tyr50	CDR2	-3.70	-4.02	5.63	-0.57	0.58	-2.08
Lys68	HV4	0.87	-56.18	53.34	-0.34	0.59	-1.72
Ser27	CDR1	-0.52	-5.14	4.32	-0.32	0.06	-1.60
Lys48	CDR2	0.63	-65.57	62.44	-0.26	1.25	-1.51
Tyr26	CDR1	-1.06	-1.46	1.79	-0.11	-0.11	-0.95
Ala28	CDR1	-0.73	0.49	-0.37	-0.33	0.13	-0.81
Ala101	CDR3	-0.40	-0.38	0.24	-0.08	0.02	-0.60
Leu104	CDR3	-0.10	-0.15	0.06	-0.00	-0.33	-0.52
Ser51	CDR2	-0.57	-1.58	1.77	-0.46	0.47	-0.37
Gln1	-	-0.16	-0.44	0.28	-0.00	0.00	-0.32
Ala103	CDR3	-0.04	-0.17	0.11	0.00	-0.19	-0.29
Pro30	CDR1	-0.10	-3.20	3.14	0.00	-0.12	-0.28
Phe66	-	-0.08	-0.10	0.18	0.00	-0.26	-0.26
Tyr49	CDR2	-0.21	0.15	-0.05	-0.00	-0.13	-0.24
Ser102	CDR3	-0.96	-4.31	5.08	-0.17	0.13	-0.23
...							
Thr29	CDR1	-0.37	-1.55	3.03	-0.18	-0.12	0.81
Asp53	CDR2	-0.10	28.56	-27.32	-0.01	0.02	1.15





Free energy calculation: Main approaches

Sampling, Exact

Free Energy Perturbation (FEP)

$$\Delta G = -k_B T \ln < exp(-\beta \Delta V) >$$

Thermodynamical Integration (TI)

$$\Delta G = \int_0^1 \left\langle \frac{\partial V}{\partial \lambda} \right\rangle_\lambda d\lambda$$

Non Equilibrium Statistical Mechanics (Jarzynski)

$$\Delta G = -k_B T \ln < exp(-\beta W) >$$

Sampling, Approx.

Linear Interaction Energy (LIE)

$$\Delta G = \alpha \Delta \left\langle V_{l\ env}^{VdW} \right\rangle + \beta \Delta \left\langle V_{l\ env}^{Elec} \right\rangle$$

Molecular Mechanics/Poisson-Boltzmann/Surface area (MM-PBSA)

$$\Delta G = \langle \Delta G_{Gas} \rangle + \left\langle \Delta G_{Desolv}^{PBSA} \right\rangle - \langle T \Delta S \rangle$$

Approx.

Quantitative Structure Activity Relationship (QSAR)

$$\Delta G = k_0 + \sum k_i X_i \qquad \Delta G = F(X)$$

(***X*** is a descriptor)

CPU Time

Computation of absolute TCR binding free energy

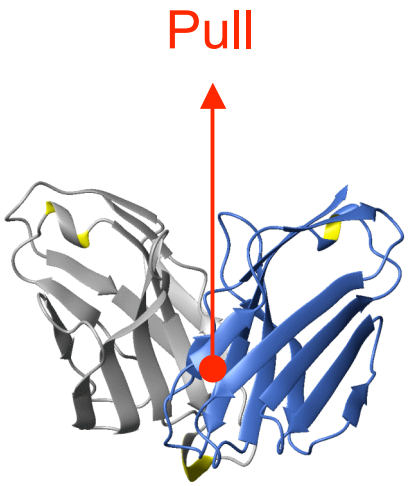
Let G be the free energy and W the work,

$W_{adia} = \Delta G \longrightarrow K_A$ (Infinitely slow)

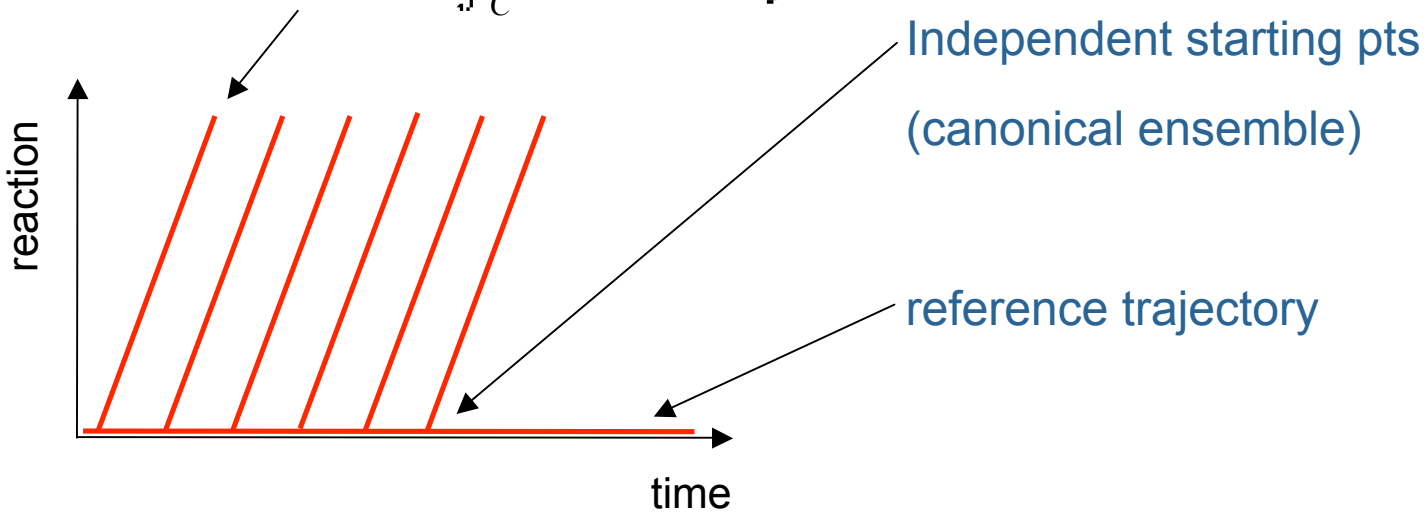
$W = W_{adia} + W_{diss}$ (Finite rate)

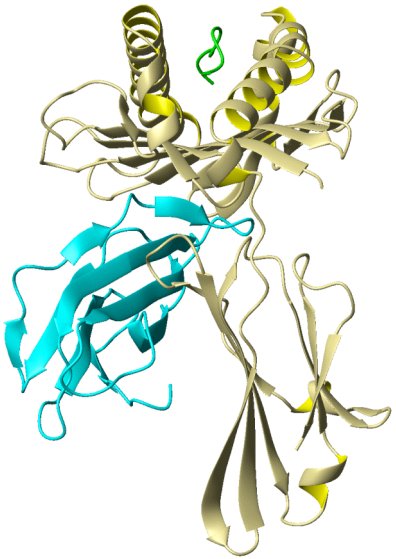
$$\left\langle e^{-\beta W} \right\rangle = e^{-\beta \Delta G}$$

(Jarzynsky)



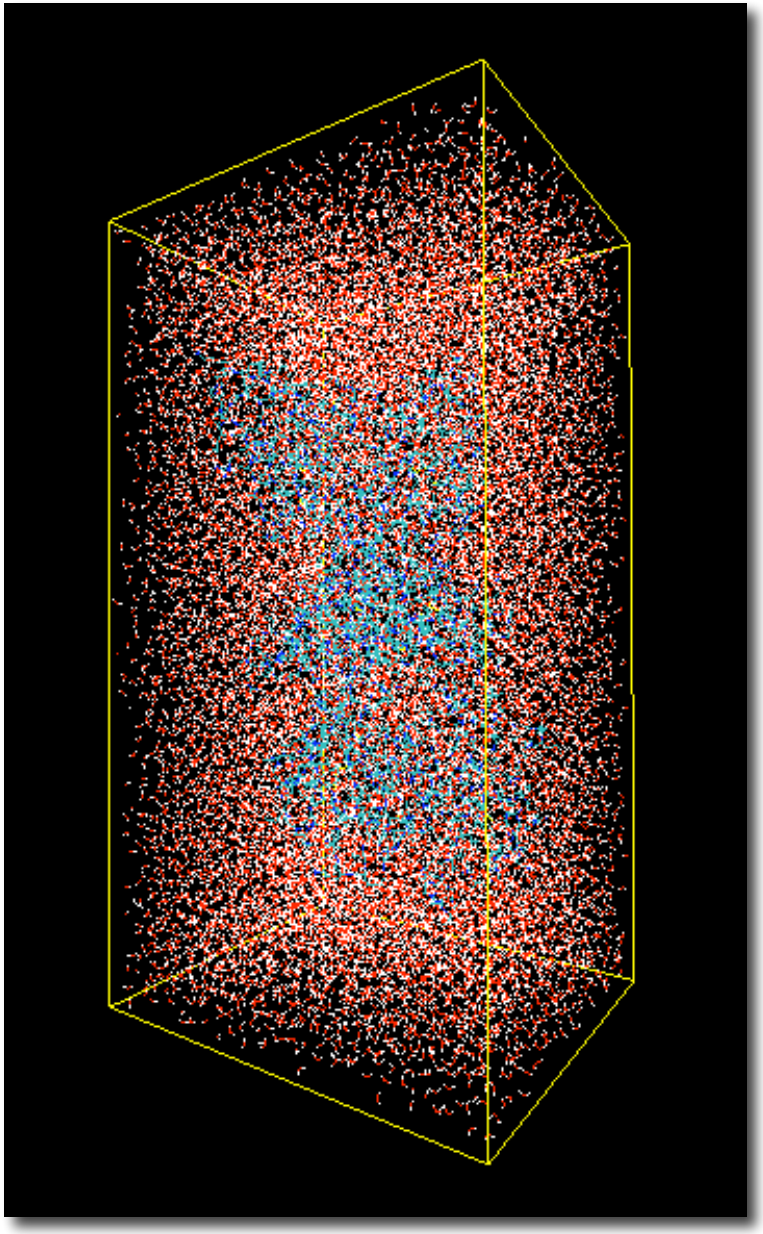
$$W = \int_C -\nabla V^{Pull} d\mathbf{r}$$



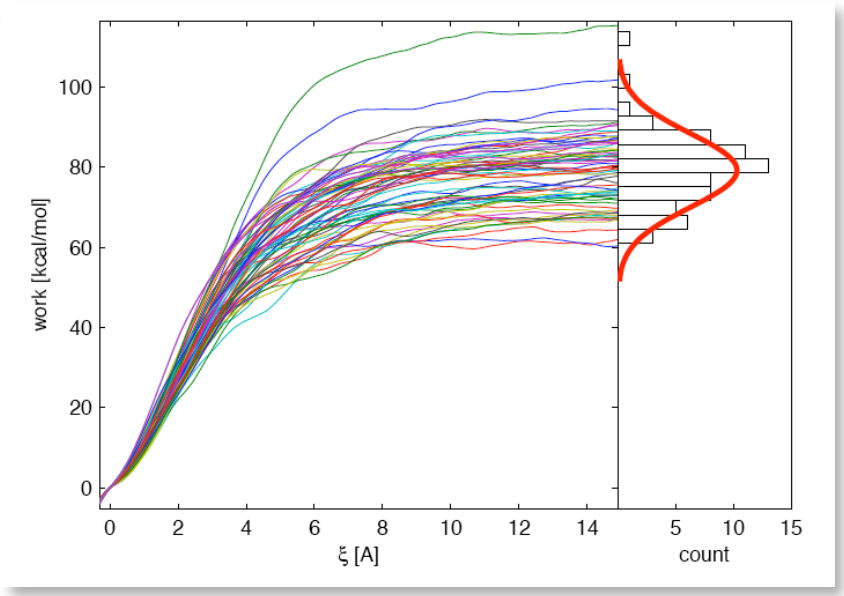


Simulation setup

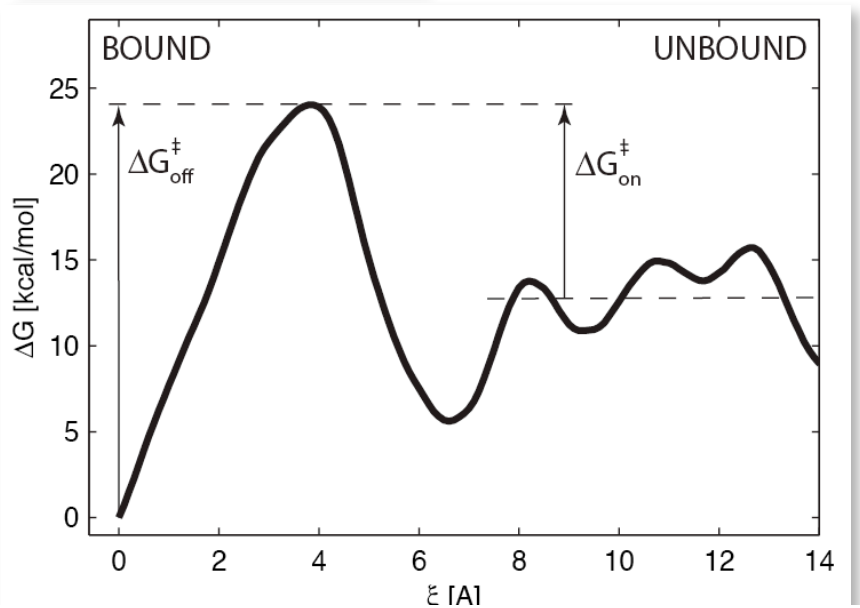
- Gromos96 Force Field
- Gromacs Engine
- Particle Mesh Ewald (PME)
- Periodic boundary conditions
- Box: 80x80x150 A
- 26000 Water molecules
- 85000 Atoms
- Hydrogen shaken
- 2 fs timestep
- 0.5 ns / 24h on 4 alpha CPU



TCR binding free energy profile:

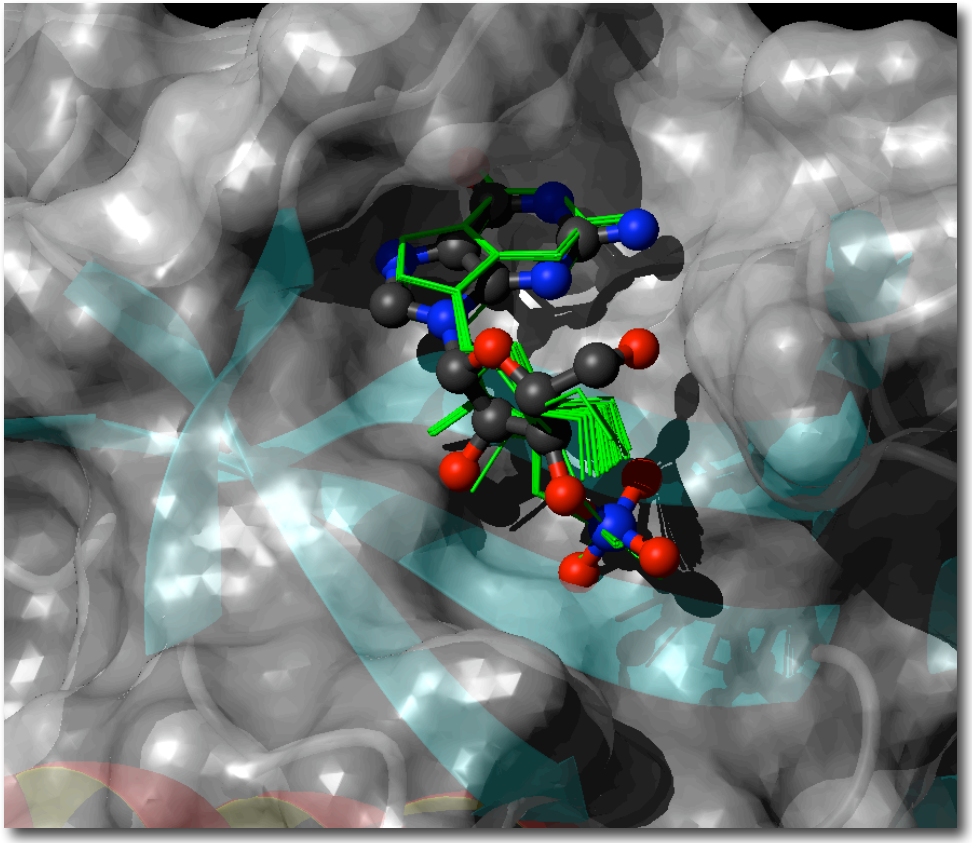


$$\left\langle e^{-\beta W} \right\rangle = e^{-\Delta G}$$

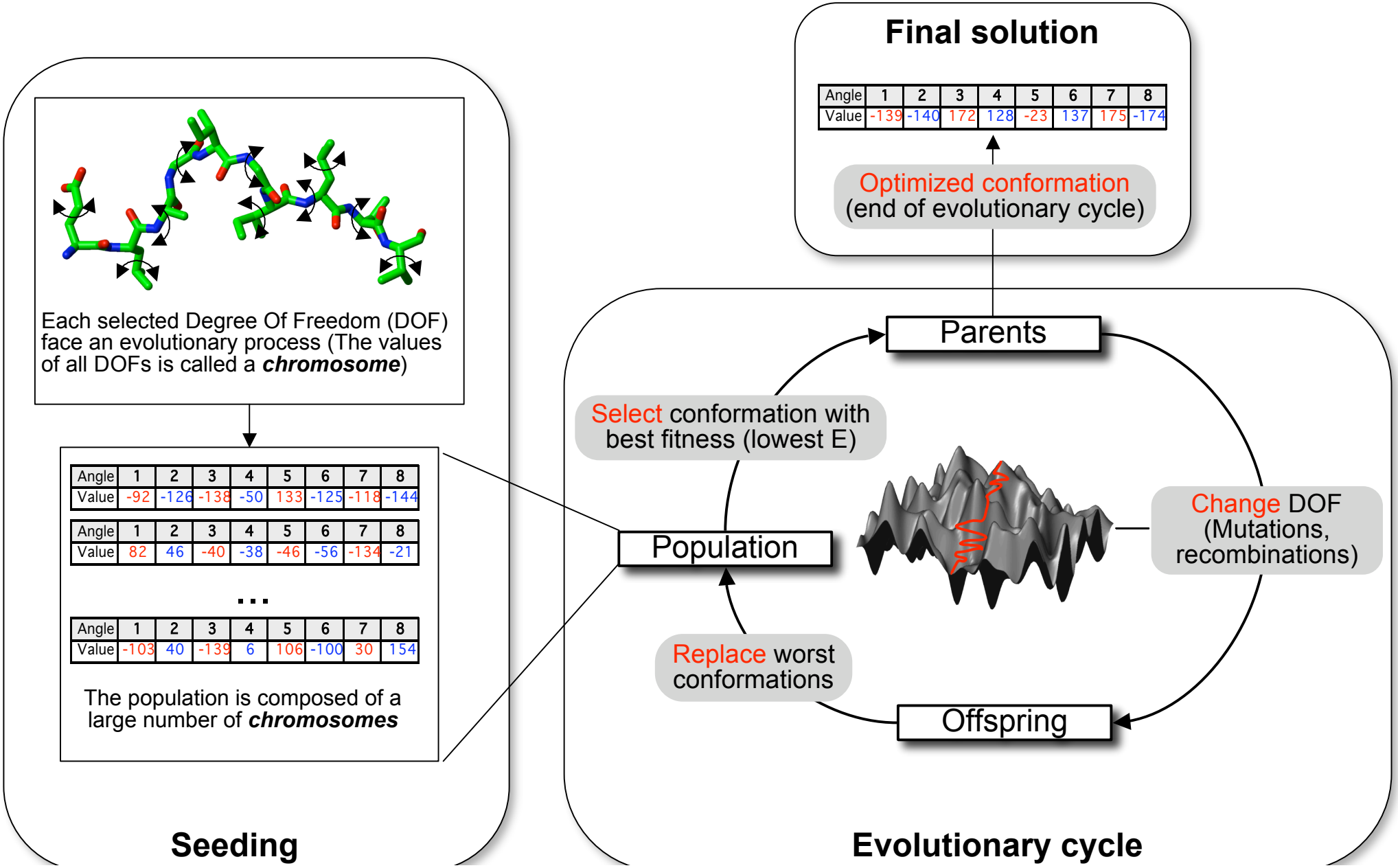


Application to the design of small molecule inhibitors

EADock



Conformational sampling using genetic algorithms



Eadock: Evolutionary Parameters

Genome

Cartesian coordinates

Fitness

Enthalpy
Enthalpy & solvation free energy (GB, PB)
Binding free energy
…

Operators

Rotations
Translations
ElectrostaticOptimizer
VanDerWaalsOptimizer
Barbatruc
LigandInterpolator
Dihedral scan
Molecular Dynamics (SA)
…

Followed by minimization

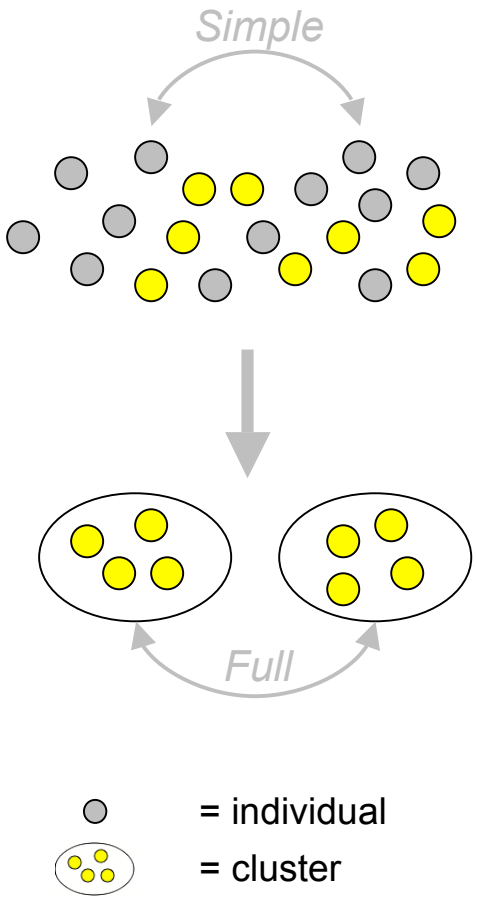
Automated Operator scheduling

Eadock: Definition of the fitness

A multi-objective fitness is used during the evolutionary process:

1) Simple fitness: CHARMM total energy with ε=4 and Rdie

2) Full fitness: CHARMM total energy with solvation free energy computed using Generalized Born implicit solvent model



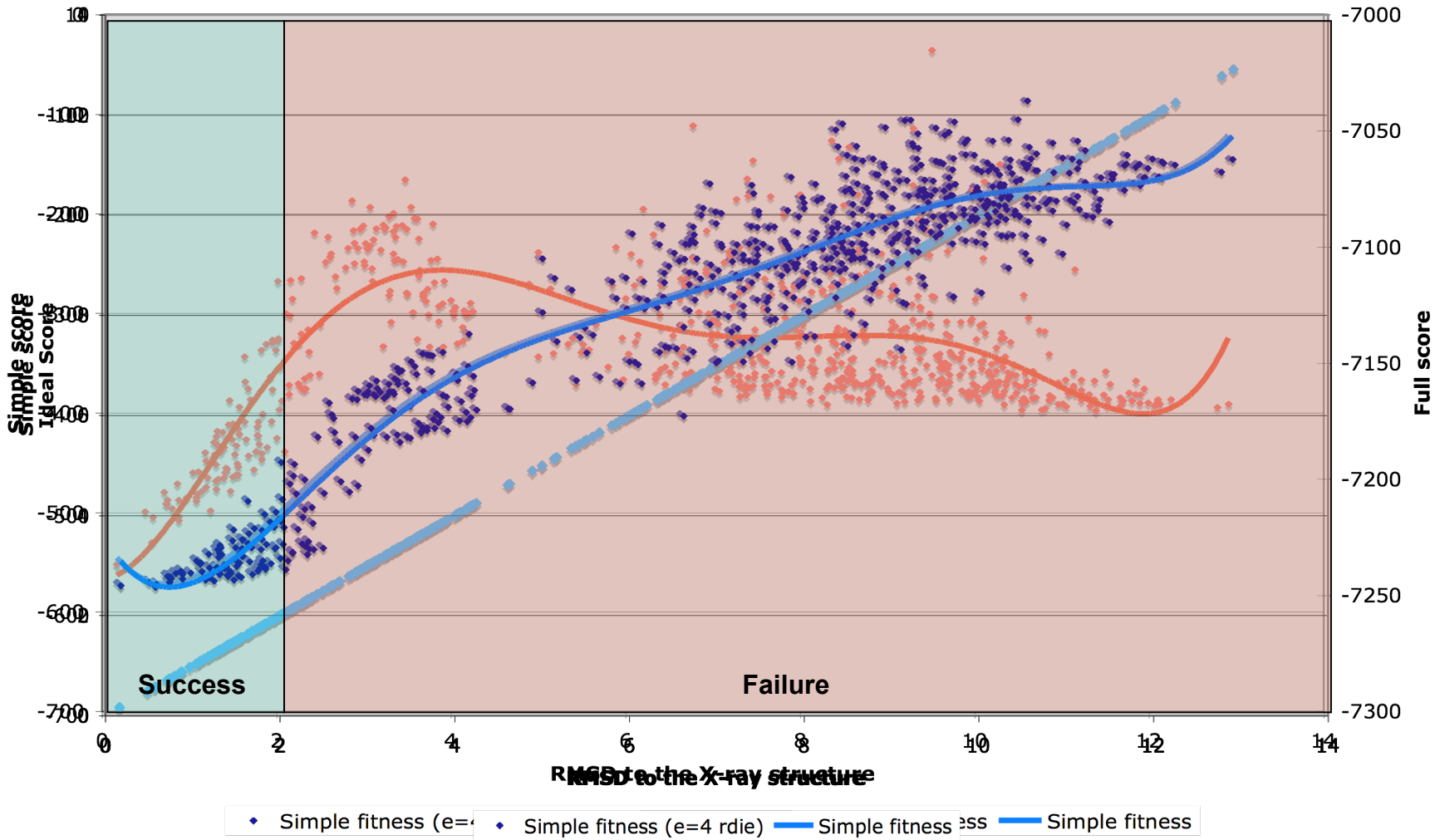
The simple fitness selects individuals

The full fitness selects between best ranked clusters

Minima of the simple fitness coincide with those of the full fitness

Choice of an optimal fitness

Analysis of 700 decoys with two solvation models



Eadock: Evolutionary Parameters

Genome

Cartesian coordinates

Fitness

Enthalpy
Enthalpy & solvation free energy (GB, PB)
Binding free energy
…

Operators

Rotations
Translations
ElectrostaticOptimizer
VanDerWaalsOptimizer
Barbatruc
LigandInterpolator
Dihedral scan
Molecular Dynamics (SA)
…

Followed by minimization

Automated Operator scheduling

Eadock: Evolutionary Parameters

Genome

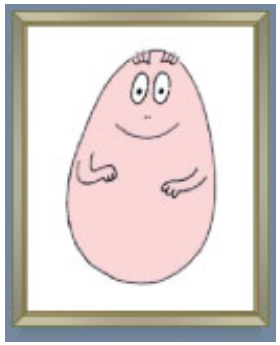
Cartesian coordinates

Fitness

Enthalpy
Enthalpy & solvation free energy (GB, PB)
Binding free energy
…

Operators

Rotations
Translations
ElectrostaticOptimizer
VanDerWaalsOptimizer
Barbatruc
LigandInterpolator
Dihedral scan
Molecular Dynamics (SA)
…

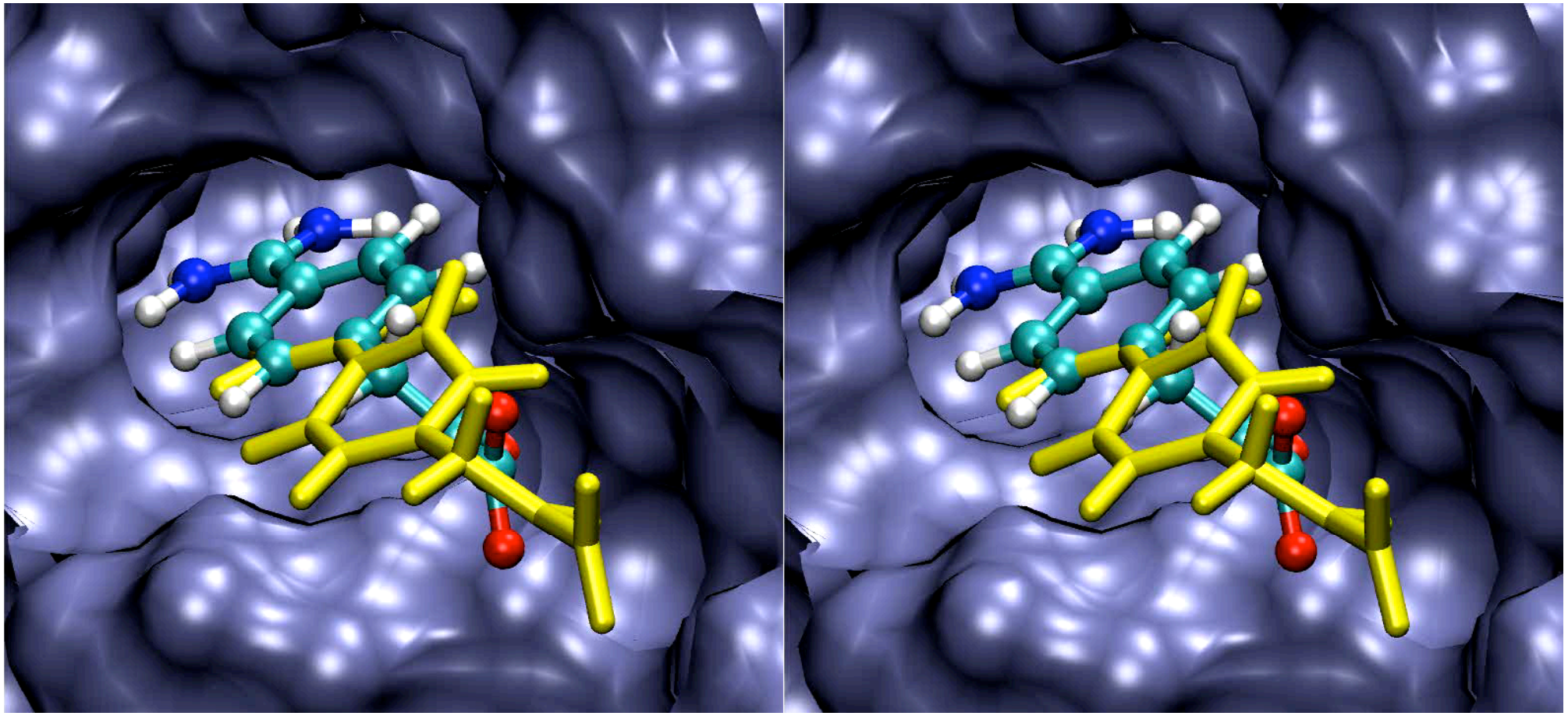


Followed by minimization

Automated Operator scheduling

Example of smart operator: Barbatruc

Starting conformation at 4.2 Å all atom RMSD



Barbatruc: final RMSD 0.7 Å

Standard minimization: final RMSD 3.2 Å

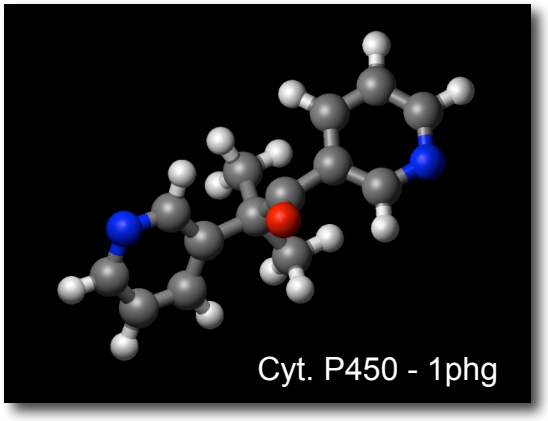
Test set for EADock benchmark

37 complexes, involving 11 different proteins

Protein	PDB	q	DoF	Hb A.	Hb D.	Mass	% B. Sur.
Anhydrase	1cil	-1	3	6	2	323.4	85.1
	1cnx	0	10	6	3	331.4	74.2
	1okl	0	2	4	1	249.3	87.7
Arabinose	1abe	0	0	5	4	150.1	100.0
	1abf	0	0	5	4	164.2	100.0
	5abp	0	1	6	5	180.2	100.0
Carbocypeptidase	1cbx	-1	3	4	1	207.2	98.2
	3cpa	0	4	4	3	238.2	97.7
	6cpa	-1	9	8	2	477.4	82.3
FABP	1icm	-1	11	2	0	227.4	95.6
	1icn	0	14	2	1	282.5	96.0
	2ifb	-1	13	2	0	255.4	96.9
Neuraminidase	1nnb	-1	4	8	5	290.3	89.7
	1nsc	-1	4	9	6	308.3	92.0
	1nsd	-1	4	8	5	290.3	92.6
Cyt. P450	1phf	0	1	1	1	144.2	100.0
	1phg	0	3	3	0	226.3	100.0
	2cpp	0	0	1	0	152.2	100.0

Protein	PDB	q	DoF	Hb A.	Hb D.	Mass	% B. Sur.
Penicillopepsin	1apt	1	17	6	5	501.7	85.9
	1apu	0	15	6	4	485.7	85.0
Ribonuclease	1gsp	0	2	9	3	360.3	80.2
	1rhl	-2	3	10	4	361.2	78.1
	1rls	-2	3	10	4	361.2	79.2
Thermolysin	3tmn	0	5	3	3	303.4	73.0
	5tln	-1	7	5	3	320.3	79.8
	6tmn	-1	11	8	3	471.5	73.2
Thrombin	1etr	0	7	6	4	504.6	87.9
	1ets	1	7	4	4	522.7	88.3
	1ett	1	7	3	3	429.6	88.2
Trypsin	1pph	1	7	3	3	429.6	69.9
	1tng	1	1	0	1	114.2	91.6
	1tni	1	4	0	1	150.2	85.6
	1tnj	1	2	0	1	122.2	92.4
	1tnk	1	3	0	1	136.2	91.0
	1tnl	1	1	0	1	134.2	92.7
	1tpp	0	2	3	2	206.2	86.9
	3ptb	1	1	0	2	121.2	94.6

$-2 \leq q \leq 1$	ligand charge
$0 \leq DoF \leq 17$	number of ligand degrees of freedom
$0 \leq Hb\ A. \leq 10$	number of ligand hydrogen bond acceptors
$0 \leq Hb\ D. \leq 6$	number of ligand hydrogen bond donnors
$114 \leq Mass \leq 523$	ligand mass (g/mol)
$69.9 \leq \%\ B.\ Sur. \leq 100$	% of ligand SASA buried upon complexation



Bursulaya et al. *JCAMD*, 2003

Docking results for the 37 test ligands

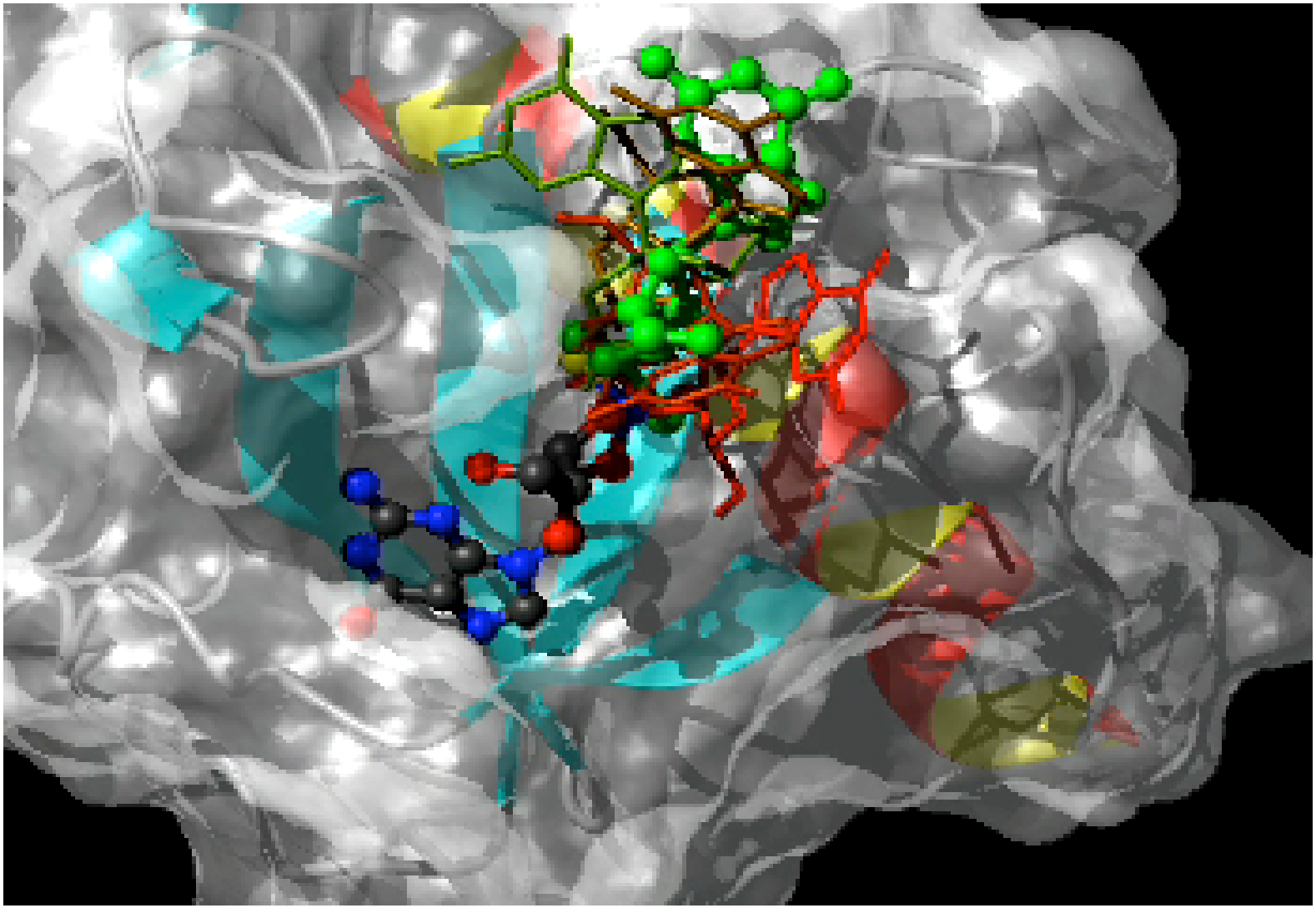
ALL RESULTS					Rank1
	Seeding w/ native binding mode				Seeding 8-11Å
Testcase Complex	**AutoDock**	**DOCK**	**FlexX**	**GOLD**	**EADock**
Trypsin					
3ptb	0.8	0.59	1.11	1.09	**0.51**
1tng	0.62	0.86	1.08	1.89	**0.27**
1tnj	1.21	1.56	1.73	1.9	**0.69**
1tnk	1.69	1.87	1.7	3.08	**1.28**
1tni	2.61	5.26	2.73	4.93	2.25
1tnl	**0.41**	2.08	3.74	1.61	0.88
1tpp	1.8	3.25	1.95	2.33	**0.38**
1pph	5.14	3.91	3.27	4.23	**0.98**
Cytochrome P-450cam					
1phf	2.09	2.39	4.68	4.42	4.58
1phg	3.52	5.57	4.87	4.2	1.68
2cpp	3.4	2.48	0.44	3.49	**0.2**
Neuraminidase					
1nsc	1.4	4.86	6	1.02	**0.48**
1nsd	1.2	4.51	1.56	0.96	**0.55**
1nnb	0.92	4.51	0.92	**0.84**	1.17
Carbocypeptidase					
1cbx	1.33	3.13	1.32	1.87	**0.42**
3cpa	2.26	6.48	1.51	1.87	0.81
6cpa	8.3	8.3	9.83	4.96	3.7
L-Arabinose					
1abe	**0.16**	1.87	0.55	0.18	0.22
1abf	0.48	3.25	0.76	0.5	**0.24**
5abp	0.48	3.89	4.68	**0.59**	0.68

ALL RESULTS					Rank1
	Seeding w/ native binding mode				Seeding 8-11Å
Testcase Complex	**AutoDock**	**DOCK**	**FlexX**	**GOLD**	**EADock**
ϵ-Thrombin					
1etr	4.61	6.66	7.26	5.99	11.07
1ets	5.06	3.93	2.11	2.39	**1.25**
1ett	8.12	1.33	6.24	1.3	5.07
Thermolysin					
3tmn	4.51	7.09	5.3	3.96	0.61
5tln	5.34	**1.39**	6.33	1.6	8.35
6tmn	8.72	7.78	4.51	8.54	8.92
Penicillopepsin					
1apt	1.89	8.06	5.95	8.82	1.65
1apu	9.1	7.58	8.43	10.7	**1.19**
Intestinal FABP					
1icm	1.8	3.99	2.94	2.3	**1.02**
1icn	3.99	3.88	2.95	2.05	1.86
2ifb	3.09	1.43	8.94	2.61	**0.6**
Ribonuclease					
1gsp	2.67	1.16	3.71	0.7	**0.39**
1rhl	0.96	**0.71**	1.15	1.08	1.02
1rls	0.98	1.75	4.33	1.16	1.01
Carbonic anhydrase					
1cil	5.81	2.78	3.52	6.04	3.48
1okl	8.54	5.65	4.22	3.55	5.79
1cnx	10.9	7.35	6.83	6.32	2.33
Overall success	46%	30%	35%	46%	**76%**

> 2.0	Unsuccessful Prediction
≤ 2.0	Successful Prediction

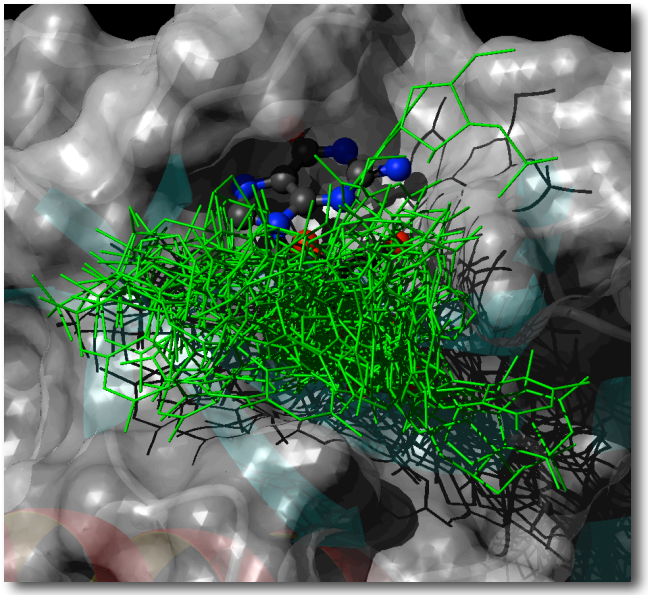
Other program data from Bursulaya et al. *JCAMD*, 2003

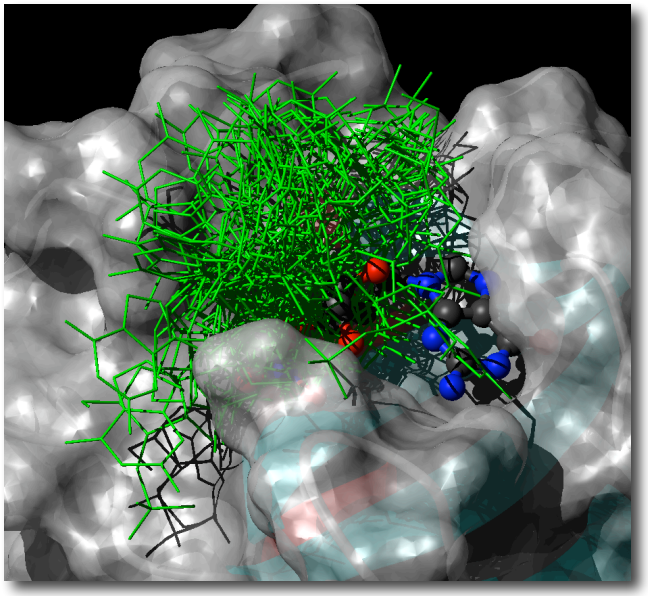
Convergence of the 5 best clusters



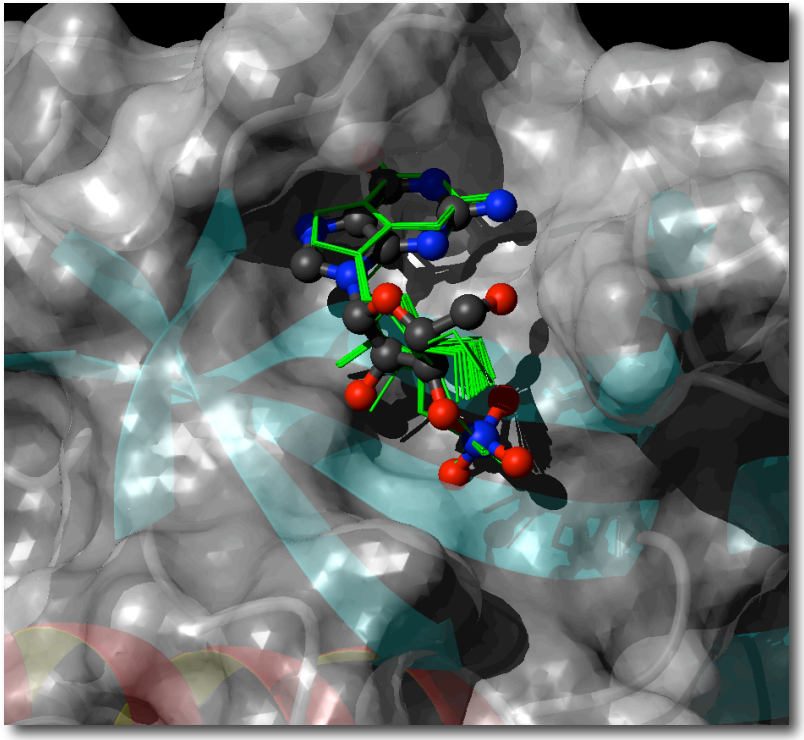
Testcase: ribonuclease (1gsp)

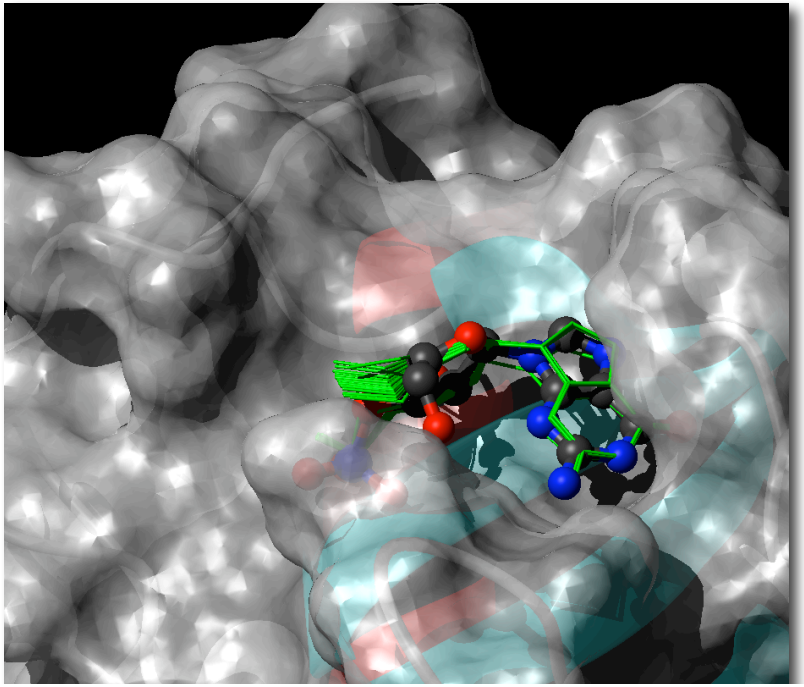
Other EAdock examples:





Evolutionnary
Process
G0 → G150





Acknowledgments

SIB

Ernest Feytmans

Vincent Zoete
Aurélien Grosdidier
Michel Cuendet
Theres Fagerberg
Antoine Leimgruber
Pierre Chodanowski
Hamid Hussain-Kahn
Muriel André

Victor Jongeneel
Roberto Fabbretti
Bruno Nyffeler

Institut Ludwig

Jean-Charles Cerottini
Pedro Romero
Daniel Speiser
Danielle Liénard

Others

Ursula Roethlisberger
John Maddocks
Horst Vogel
Paolo de Los Rios

Martin Karplus (Harvard)
Andrej Sali (UCSF)