Marcus Gastreich

3 Key Factors for Success: Fast. Visual. Easy.

Empowering Medicinal Chemistry Visually

The Past 25+ Years

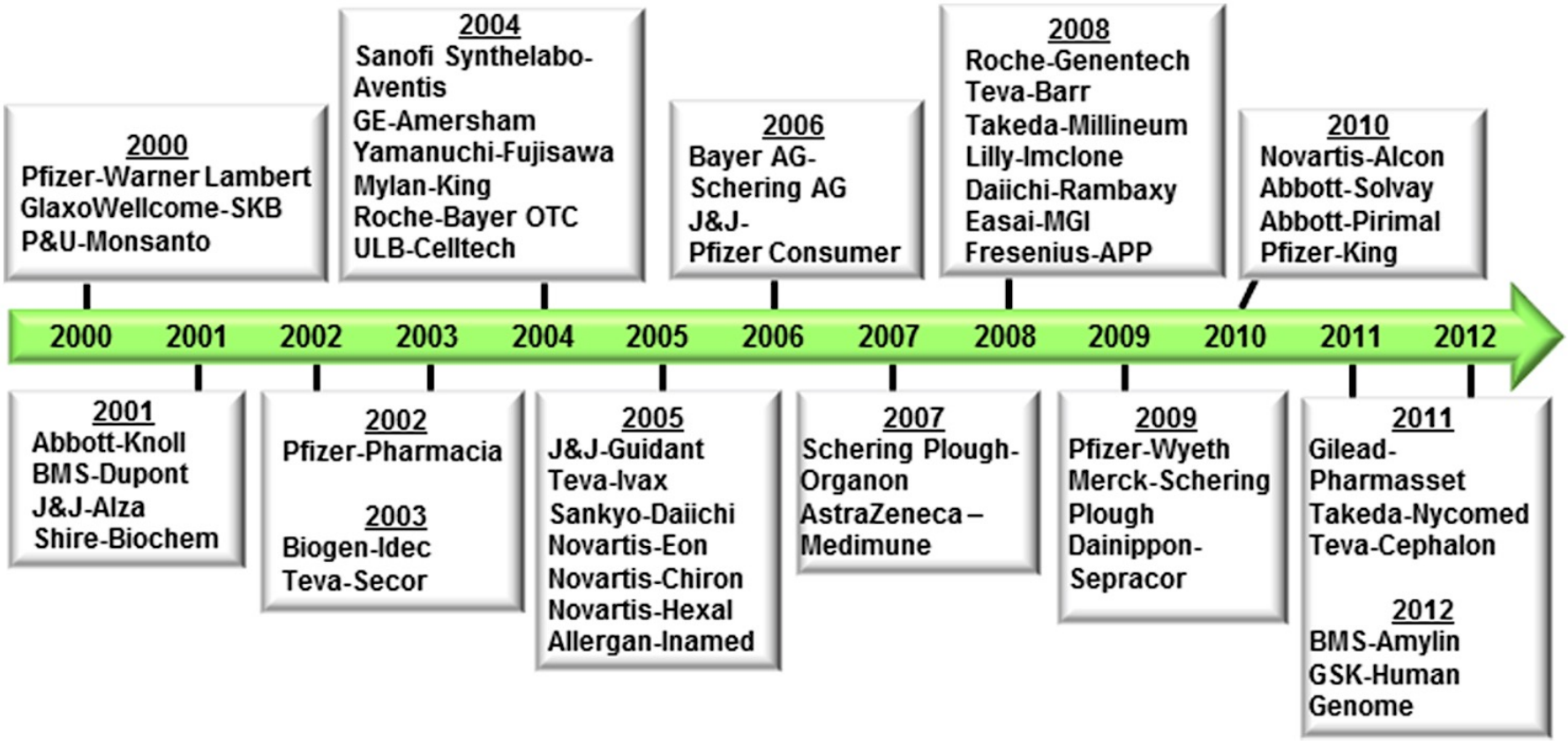
- 80ies:
  "Classical MedChem",
  almost no predictive computing

- 90ies:
  Advent of Rational DD,
  early computing, docking (FlexX1996), similarity
  (Daylight & Unity Fingerprints)

- 2000+:
  Outsourcing, merging, 300.000+ layoffs in pharma

Data from:
Abou-Gharbia et al, JMC 2013, dx.doi.org/10.1021/jm401564r

The Past 25+ Years



Timeline of mergers and acquisitions with values ≥$2 billion that occurred from 2000 to 2012.

Data from:
Abou-Gharbia et al, JMC 2013, dx.doi.org/10.1021/jm401564r

The Past 25+ Years

1988 Members of PhRMA		
Abbott Labs	G. D. Searle	Procter & Gamble
Am. Cyanamid	Glaxo	Rhone Poulenc
A. H. Robins	Hoechst	Rorer
Astra	Hoffmann-LaRoche	R. P. Scherer
BASF	ICI	Roussel
Beecham	J & J	Sandoz
Boehringer	Knoll	Schering-Plough
Boots	Eli Lilly	Smith Kline
Bristol-Myers	Marion Labs	Squibb
Carter-Wallace	Merck	Sterling
Ciba Geigy	Merrell Dow	Upjohn
Connaught	Monsanto	Warner-Lambert
DuPont	Pfizer	Wellcome
Fisons Corp	Pharmacia	Zeneca

2011 Members of PhRMA	
Abbott Labs	Eli Lilly
Astra-Zeneca	Merck
Boehringer	Novartis
Bristol-Myers Squibb	Pfizer
GlaxoSmithKline	Sanofi-Aventis
Johnson & Johnson	

Data from:
Abou-Gharbia et al, JMC 2013, dx.doi.org/10.1021/jm401564r

How This is Important for Software

Substantial changes in cheminformatics in the past years:

1. Economists/Layers manage where chemists used to
2. Medchems "model" what cheminformaticians did
3. Computers (AI) maybe taking over from... humans — ?

=> Cheminformatics and software creators must act!

- Inform managers, medchems, ... visually
- Get better decisions faster
  = improve on science, but avoid blackboxing
- Make computation transparent and very easy.

Fast.

Why? (Beyond the obvious "time winning" of course…)

- Enables non-specialists to use the tools
  (chemists are paid for lab time -> t @computer must be ~0!)
- Allows direct intervention/correction by human being
- Fast also creates fun in usage.

Algorithmic approaches

- Indexing, hashing
- Fast tree-based algorithms
- AI / neural networks / …

Fast — A Few Examples.

Examples

- Rescaffolding with on-the-fly protonation
- Traversing the chemical universe

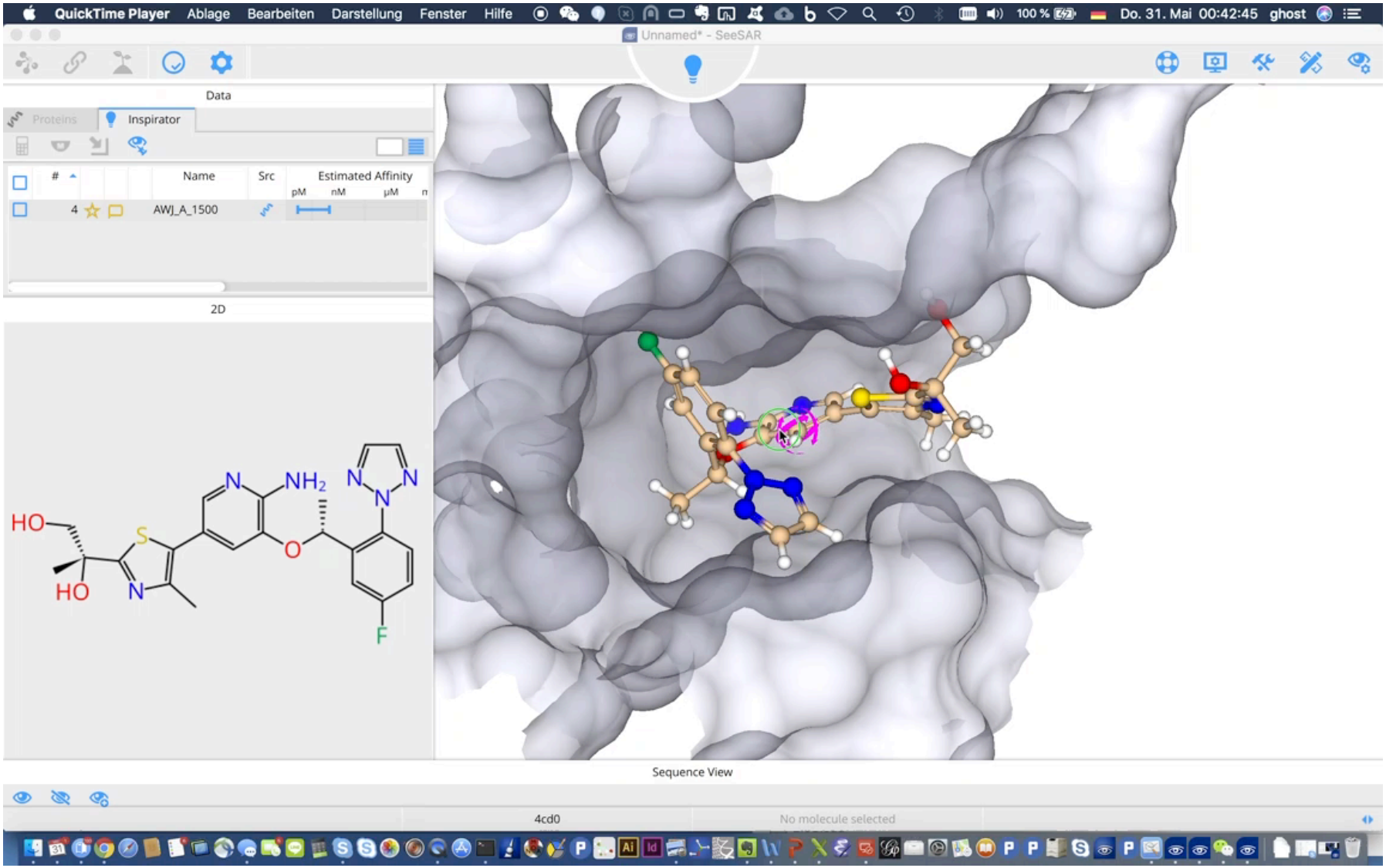
Fast — SeeSAR's Inspirator™

- **Input**: a query molecule

- **Output**:
  - from a millions of fragments…
  - the *best fitting* proposals for linking / merging / growing

- **Algorithms** used (hashing/indexing/trees):
  - ReCore and a specially adapted FlexX-like method
  - First Ideas by Maass, Rarey, and Lengauer
    JCIM **2007,** *47,* 390; J Mol Biol 1996, 261:470

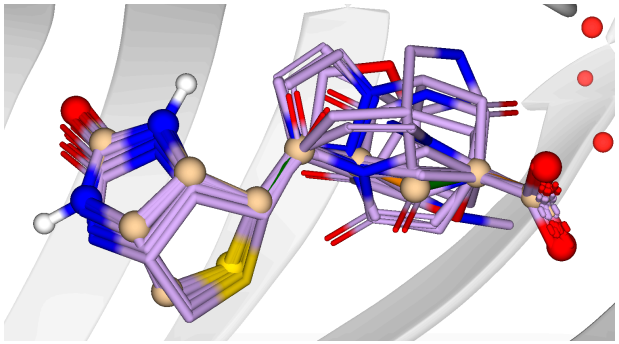
Fast — SeeSAR's Inspirator™



The Inspirator™ — Other Tasks

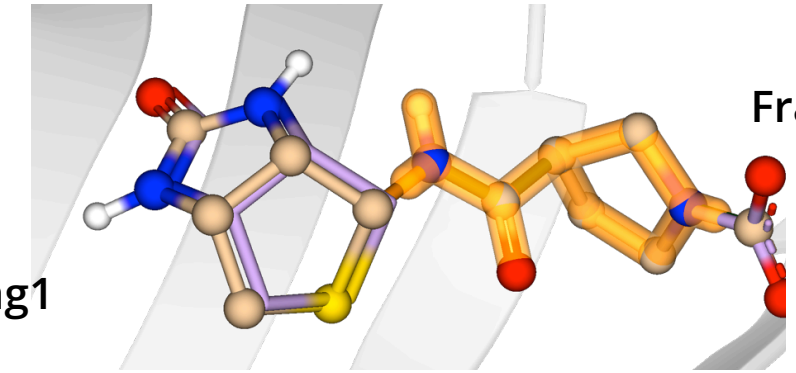
- Rescaffolding using ReCore (with Roche and Hamburg Univ.)

Replace Core!



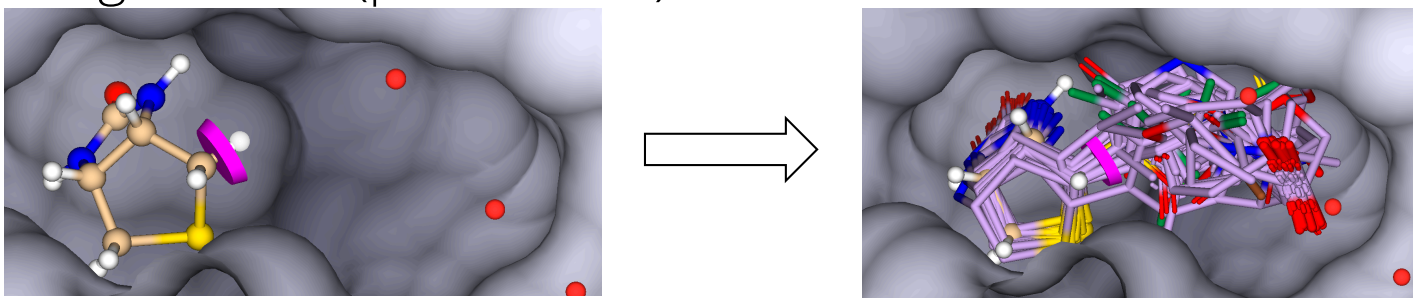
- Link/Merge (indexed)

Fit a Linker!



- Growing from a single vector (parallelized)

Grow!

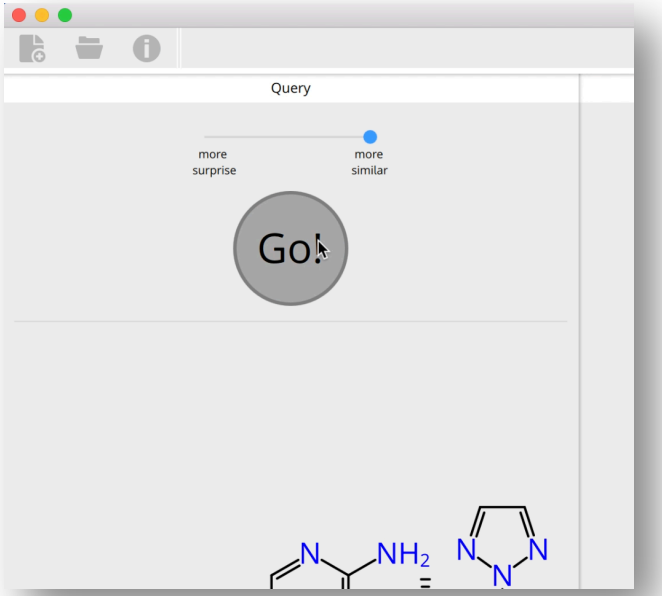


Fast — The RealSpaceNavigator 2™

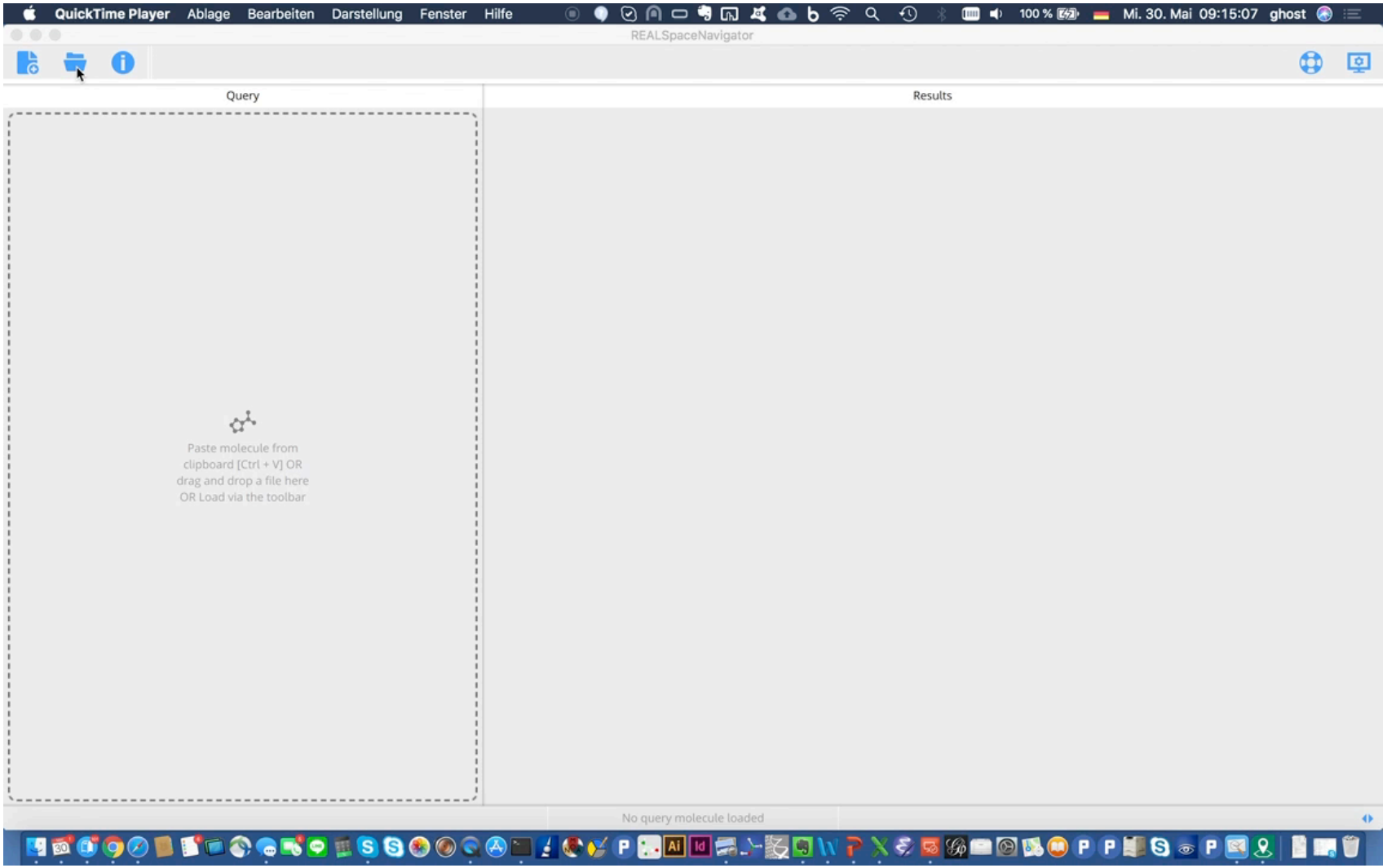
- **Input**: a query molecule

- **Output**:
  - from a 3.8billion (!) chemical space
  - 1.000 best compounds
  - purchasable on demand
  - synthesis by Enamine

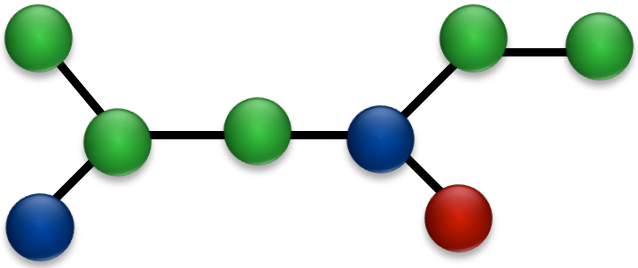
- **Fast algorithm used**:
  - FTrees-FS
  - Idea by Rarey &/Stahl, JCAMD 15: 497–520, 2001



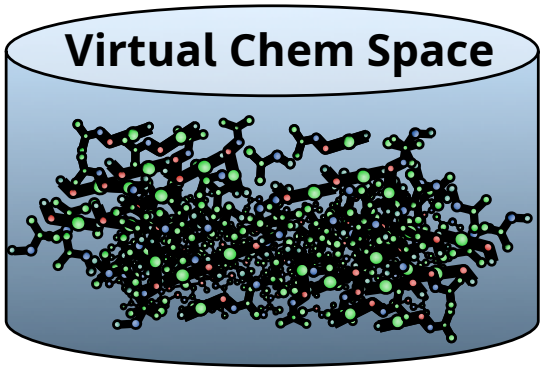
The RealSpaceNavigator2 in Action



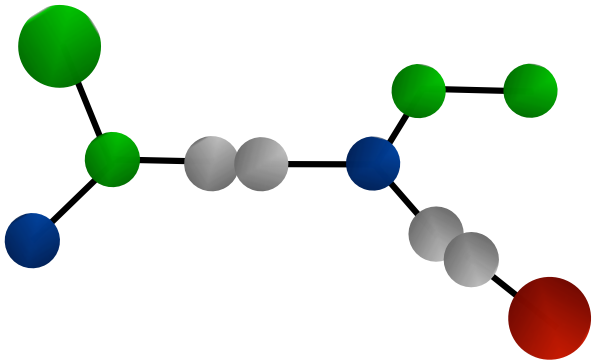
We Search Synthesis Space, Not Compounds



query



Search Algorithm



*M. Rarey, M. Stahl, *JCAMD* **2001**, 15

RSN2 — How to Run It

- Download for free from

  `biosolveit.com/RealSpaceNavigator`

- Run, select compounds that you like using the tickboxes

- “Save” an SD file

- Send email to Enamine
  - 80%+ delivery guaranteed
  - pay only what has been synthesized

Visual.

Why?

- Seeing is the ultimate key to understanding.
- Graphics convey information, esp. cross-discipline.
- Visualization leaves the power by the user.

Approaches

- Think very, very hard ☺
- Visualize so that the user can relate (less is more…)
- Avoid fitted, unphysical parameters wherever possible (distroys “understandability”)
- Use modern graphics cards & their capabilities

Why Visualization is Soooo Important

- Glues disciplines together!
  see for ex.: Martha S. Head et al. (GSK), JCAMD (2012) 26:51

- Leaves power to criticise computer in hands of scientist:

  *"I can **see** that this computer result is wrong!"*

  Examples:
  - A force field value is -17235,54 -> and now??
  - AI: Learning machine trained on arts

Robbie Barrat's Arts Learning Machines

Tousands(!) of nude portraits fed into neural network -> this is one result...:

from: pic.twitter.com/tYgzCHGfse

Just visualize, and know quickly that this result is very wrong.

http://www.neatorama.com/2018/03/29/Nude-Art-Generated-by-Artificial-Intelligence/

Visual — A Few Examples

Examples:

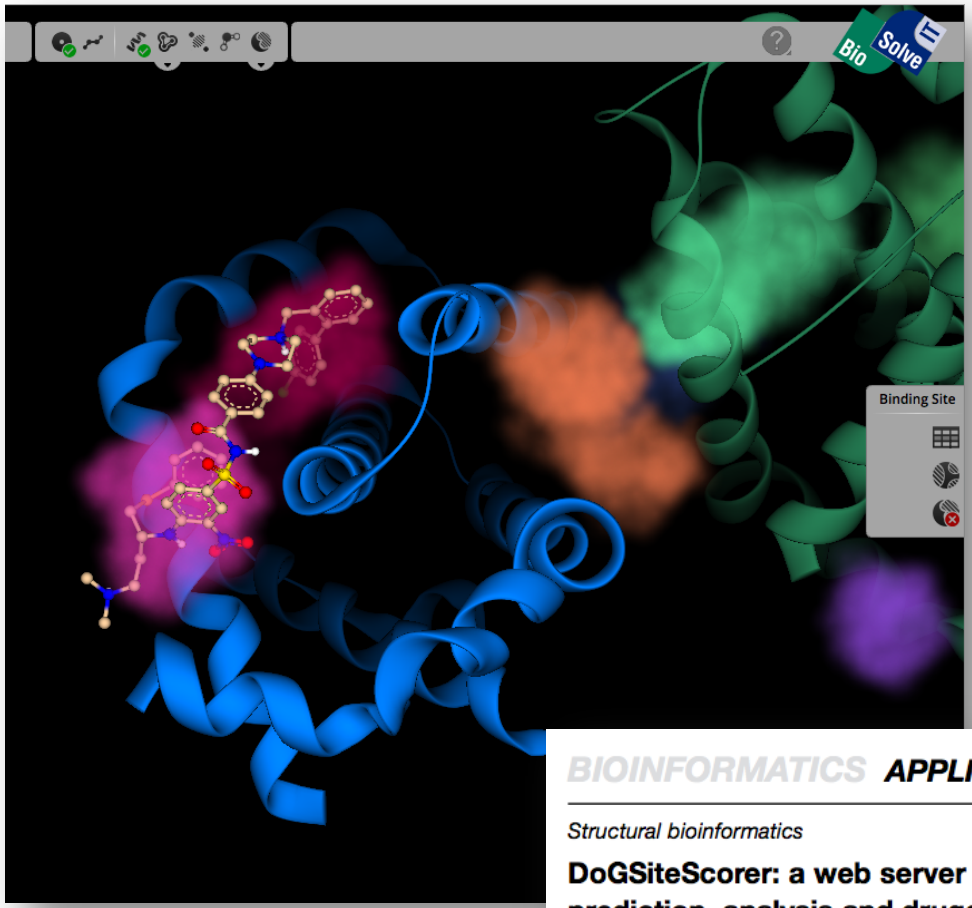
- Pocket Detection
- HYDE and torsional significance

Algorithms used:

- SVM (Pockets)
  - Volkamer et al, JCIM 2012, 52, 360−372
- Simple formulas (HYDE)
  - Reulecke et al., ChemMedChem 2008, 3, 885 – 897
  - Schneider et al., JCAMD 2013, 27, 15

Visual Pocket Proposals: DoGSite in SeeSAR

With University of Hamburg & Merck Germany



BIOINFORMATICS APPLICATIONS NOTE
Vol. 28 no. 15 2012, pages 2074–2075
doi:10.1093/bioinformatics/bts310

Structural bioinformatics

Advance Access publication May 23, 2012

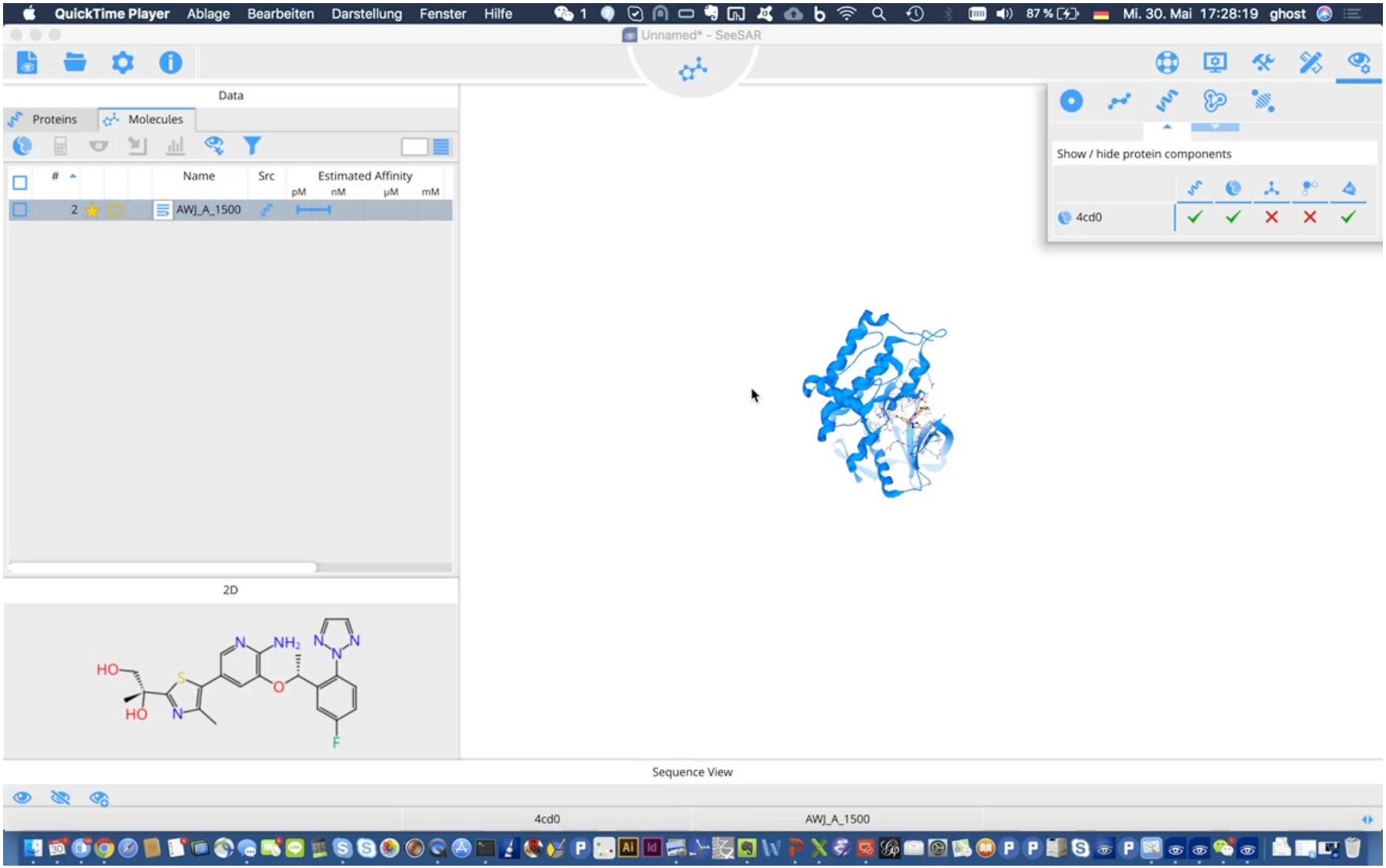
DoGSiteScorer: a web server for automatic binding site prediction, analysis and druggability assessment

Andrea Volkamer[1], Daniel Kuhn[2], Friedrich Rippmann[2] and Matthias Rarey[1,*]

[1]Center for Bioinformatics, University of Hamburg, Bundesstr and [2]Merck KGaA, Merck Serono, Global Computational Chemistry, Frankfurter Str. 250, 64293 Darmstadt, Germany

Associate Editor: Anna Tramontano

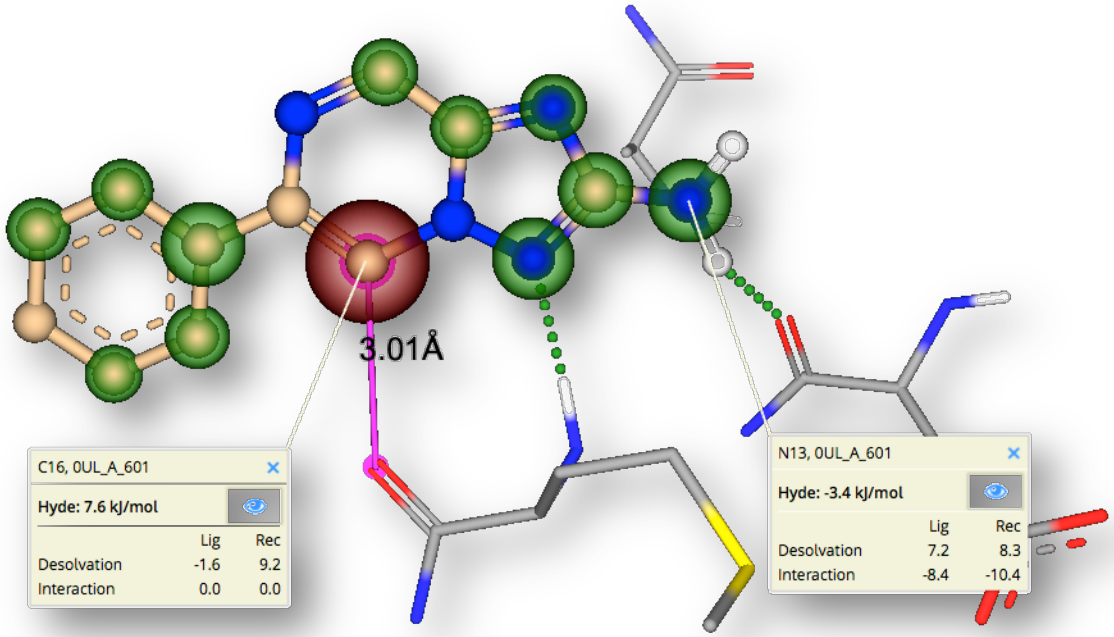
Visual — ΔG and Torsions in SeeSAR

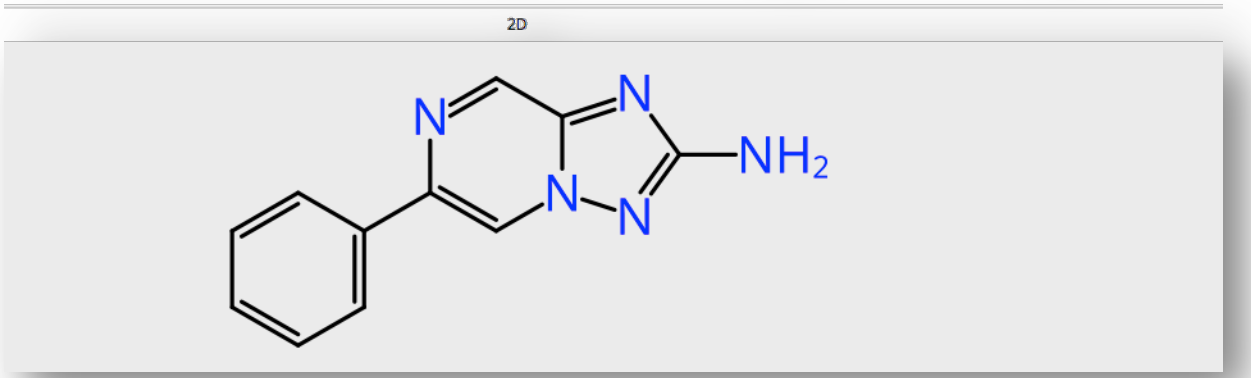


HYDE ΔG Estimates — The Coronas™

Green = good for affinity
Red = bad for affinity

Larger
= stronger contribution.





HYDE Theory in a Nutshell

$$\Delta G_{\mathrm{HYDE}} = \sum_{\text{atom } i} \Delta G^{i}_{\mathrm{dehydration}} + \Delta G^{i}_{\mathrm{H-bond}}$$

$$\Delta G^{i}_{dehydration} = -2.3RT \cdot p\log P_{atom\,i} \cdot \Delta\text{accessibility}^{i}$$

$$\Delta G^{i}_{H-bond} = 2.3RT \cdot \frac{p\log P_{atom\,i}}{f_{sat}(T)} \cdot \Delta\text{interact}^{i}$$

- logP connects dehydration and H-bonds in a natural way
  => No unphysical parameters (only 8 logPs, f_{sat})
- Computation is super fast (~ very few seconds; FEP needs days!)
  => You can explore/play with ligand modifications on-the-fly
- Atomic increments
  => You can VISUALIZE affinity !!

Collaborators:
Bayer + Univ.Hamburg

Schneider et al. 2013, JCAMD 27(1):15-29 and refs therein

Easy.

Why?

- Otherwise the software will not be used
- Young generation ‘spoiled’ by easy user interfaces (smartphones, pads, etc.)
- Quicker learning, faster results.

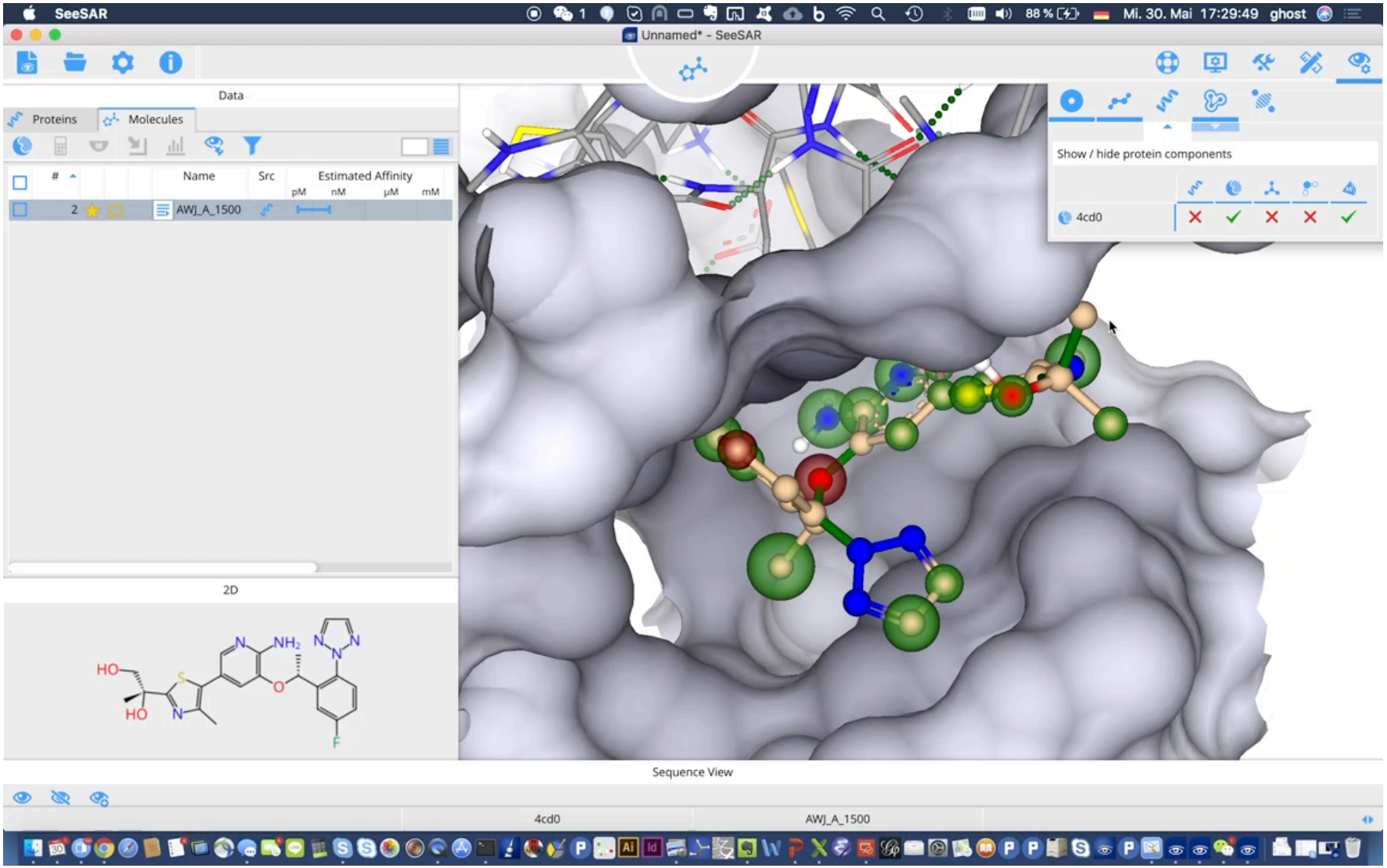
Approaches

- Less is more -> cleaner interface, smooth usage
- Align *computer screen* with *physics behind*
  e.g., drag a ligand into a pocket -> fun!
- Address user <u>needs</u> not wishes !! -> happy users

Easy — A Few Examples

- SmartAssistants in editing
- ADME Properties by Optibrium

Easy — SmartAssistants in SeeSAR™



Fast, Visual, Easy — Summary

Superfast algorithms now enable

- ΔG approximations within seconds, including desolvation
- on-the-fly protonation, tautomers, water orientation…
- 3D editing with SmartAssistants
- Searches in giant chemical spaces

Visualization…

- ..is key to understanding and knowledge transfer
- keeps decision power with the user

Easy usage…

- saves enormous amounts of time
- opens the door to all chemists/biologists/pharmacists/… to create better ideas faster

Thanks!

BioSolveIT in Japan: Contact MOLSIS ☺

BioSolveIT 社日本総代理店

株式会社 モルシス　ライフサイエンス部

〒104-0033　東京都中央区新川 1-28-38　東京ダイヤビル

Phone: 03-3553-8030　FAX: 03-3553-8031

URL: https://www.molsis.co.jp/　E-mail: support@molsis.co.jp

Our download web page:

biosolveit.com/download