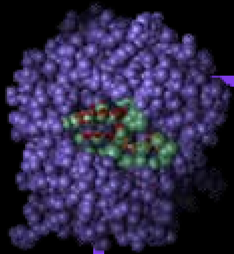
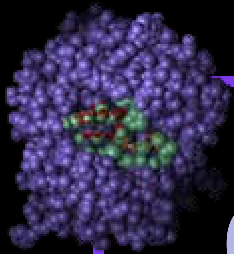


3D-bioinformatics

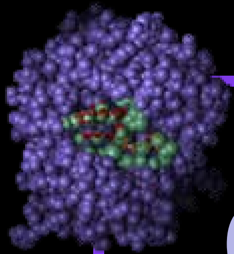
Service tools for modelling and design

Andreas Hoppe, AG Proteinstrukturtheorie,
Institut für Biochemie, Charité, Humboldt-Universität Berlin



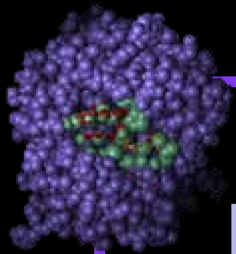
Contents

- Bioinformatics – where we are
- The course of motif searching
- Applications



Contents

- Bioinformatics – where we are
- The course of motif searching
- Applications



Bioinformatics

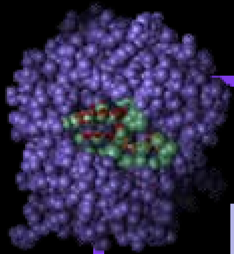
Tasks currently undertaken by biologists:

- Sequence similarity searching 35%
- Functional motif searching 11%

…

- Structure prediction 4%
- Protein analysis 3%

Stevens et al., Bioinformatics 17(2), pp. 180-188, 2001



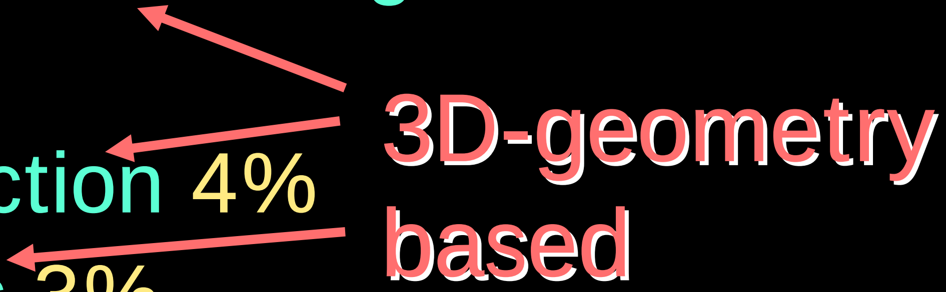
Bioinformatics

Tasks currently undertaken by biologists:

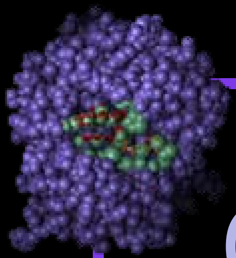
- Sequence similarity searching 35%
- Functional motif searching 11%

…

- Structure prediction 4%
- Protein analysis 3%


3D-geometry based

Stevens et al., Bioinformatics 17(2), pp. 180-188, 2001

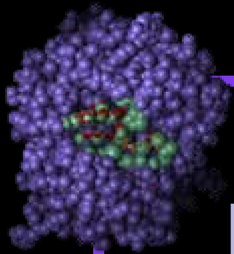


Contents

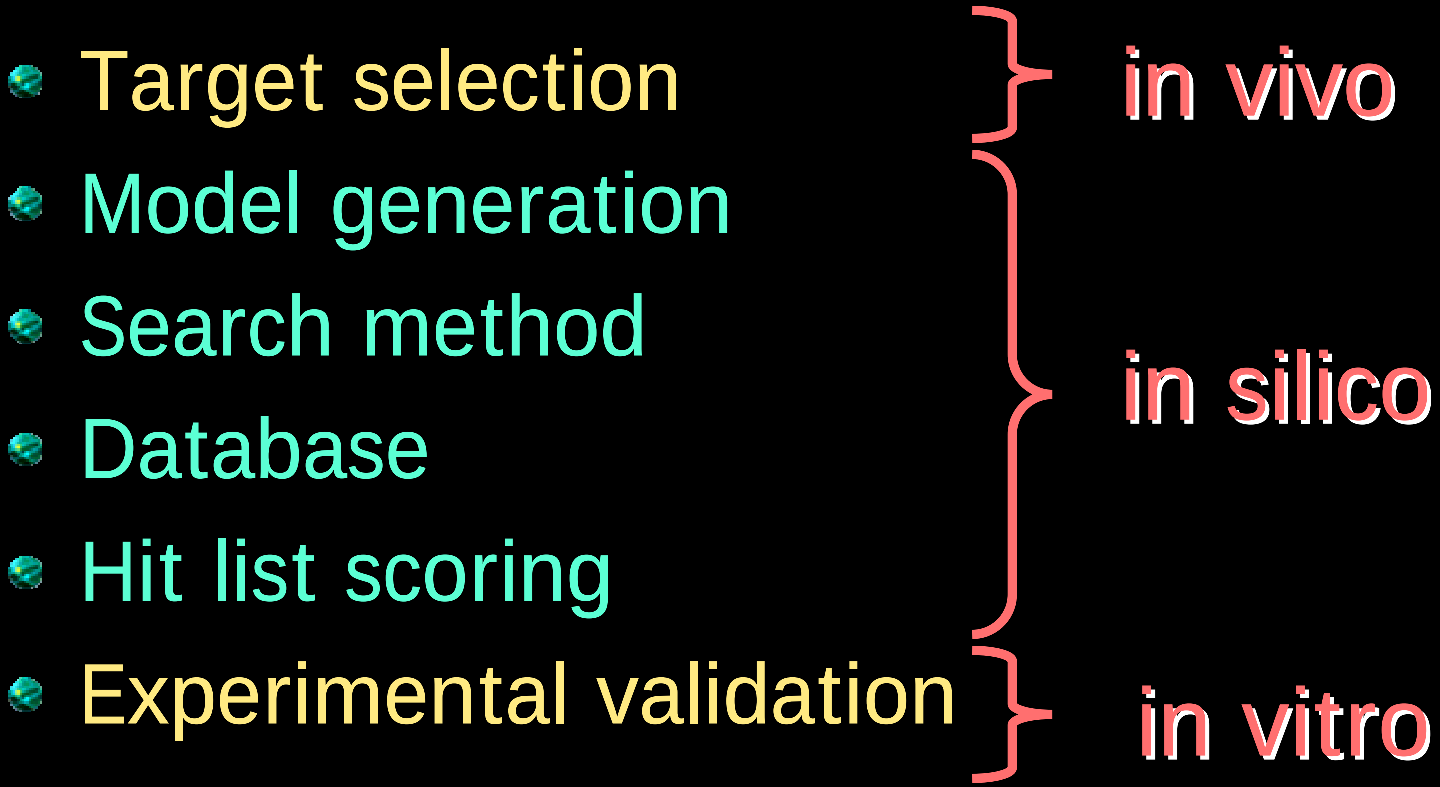
- Bioinformatics – where we are
- The course of motif searching
- Applications

Motif searching

Target selection	in vivo
Model generation	in silico
Search method	
Database	
Hit list scoring	
Experimental validation	in vitro



Motif searching

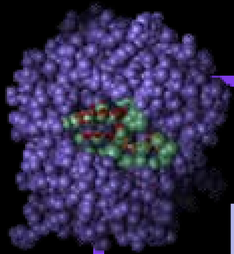


- Target selection
- Model generation
- Search method
- Database
- Hit list scoring
- Experimental validation

in vivo

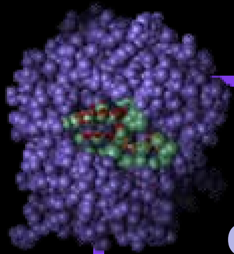
in silico

in vitro



Model generation

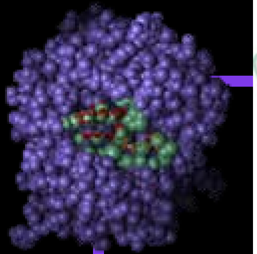
- Analysis of the target interaction: similarity graphs
- Mostly hand-work!

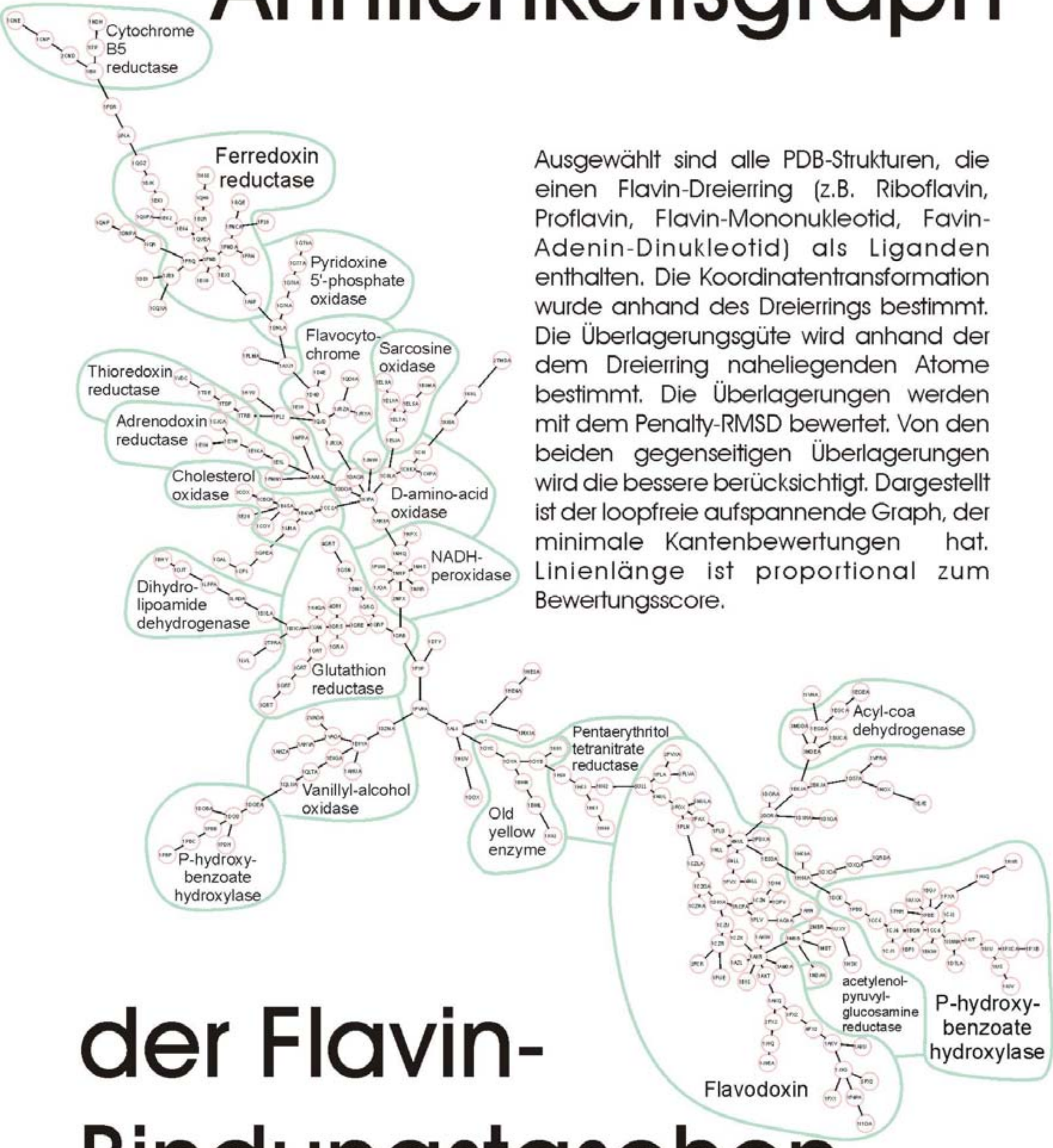


Similarity graphs

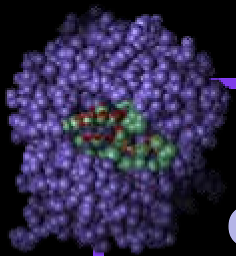
- Tool to understand binding site geometry
- Relationships of proteins based on the binding site geometry
- Aid for the model generation

Ähnlichkeitsgraph der Flavin-Bindungstaschen





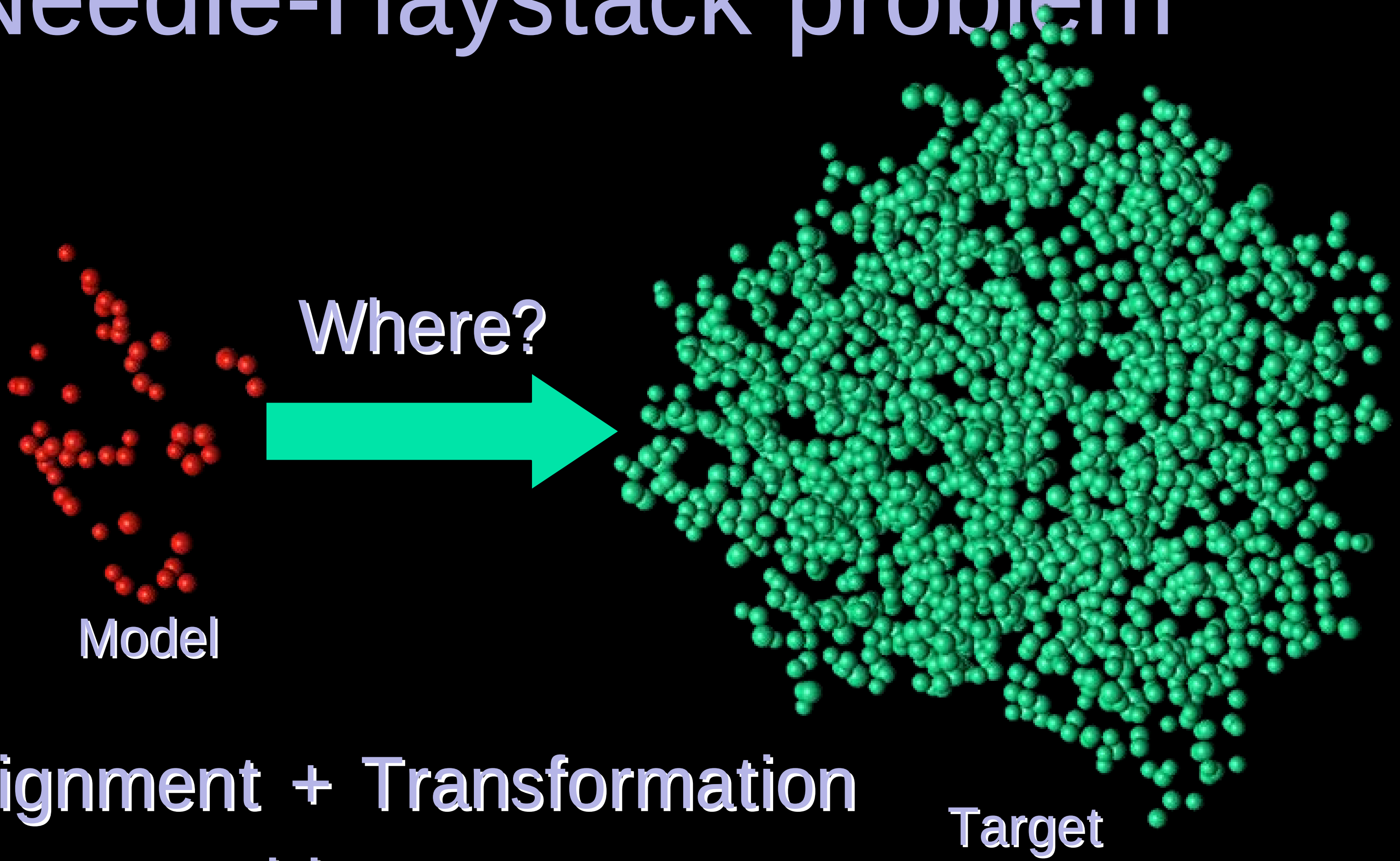
Ausgewählt sind alle PDB-Strukturen, die einen Flavin-Dreierring (z.B. Riboflavin, Proflavin, Flavin-Mononukleotid, Favin-Adenin-Dinukleotid) als Liganden enthalten. Die Koordinatentransformation wurde anhand des Dreierrings bestimmt. Die Überlagerungsgüte wird anhand der dem Dreierring naheliegenden Atome bestimmt. Die Überlagerungen werden mit dem Penalty-RMSD bewertet. Von den beiden gegenseitigen Überlagerungen wird die bessere berücksichtigt. Dargestellt ist der loopfreie aufspannende Graph, der minimale Kantenbewertungen hat. Linienlänge ist proportional zum Bewertungsscore.



Search methods

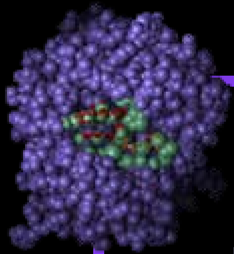
- Needle-haystack superposition
- Iteration through the database
- Focus on speed

Needle-Haystack problem



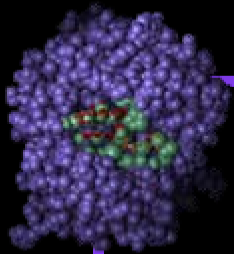
Assignment + Transformation

= Superposition



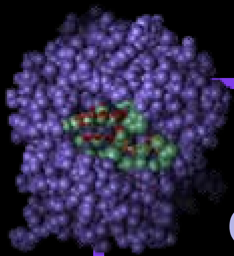
NeedleHaystack program

Available online & offline at
http://bioinf.charite.de/haystack



Databases

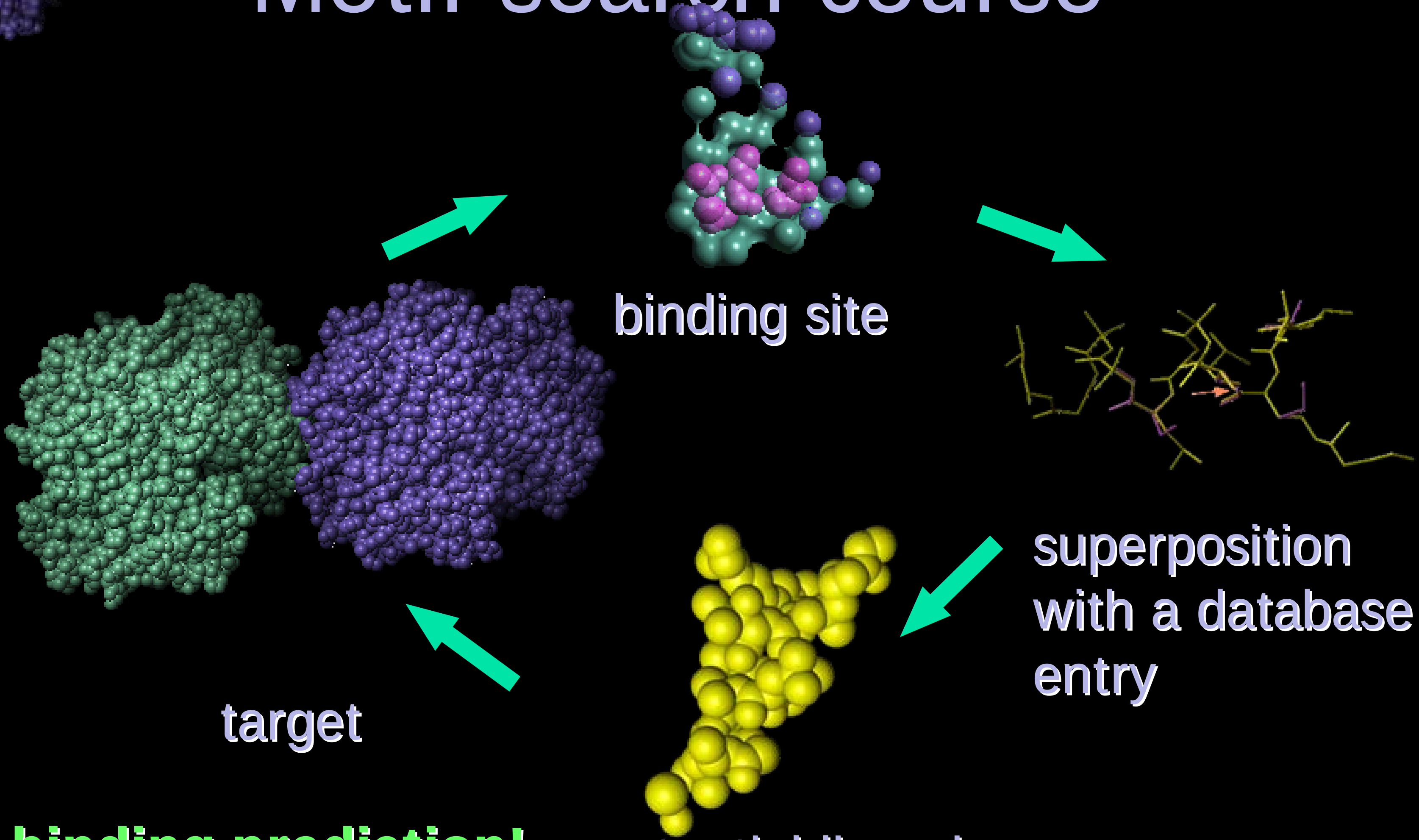
- Raw data: PDB, CSD, NCI
- Automatic error-checking, automatic generation of:
  - Complexes, monomers, chains
  - Full sets, surfaces, interfaces, backbones
  - Representative selections of homologous proteins: PISCES
  - Stereo isomers



Scoring

- Regarding binding partner
- Estimating physicochemical potential
- Visual inspection: interesting?

Motif search course



binding site

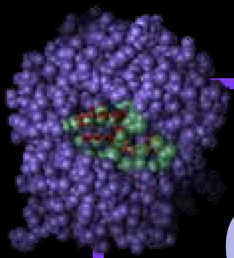
superposition
with a database
entry

potential ligand

target

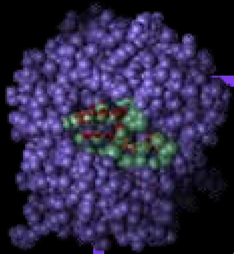
binding prediction!

NeedleHaystack ... Hoppe et al., poster at Euro QSAR 2000,
see www.charite.de/bioinf/hoppe



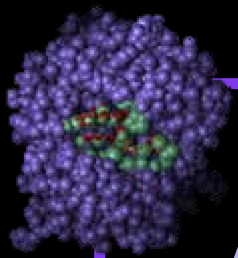
Contents

- Bioinformatics – where we are
- The course of motif searching
- → Applications



Applications

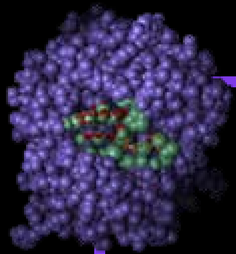
- Search putative ligands
- Find more examples for a given interaction
- Structure-based modeling



Acknowledgements

- Cornelius Frömmel, Christoph Gille, Robert Preissner, Kristian Rother
- Supported by DFG and

Charité



Thanks for your attention!

Welcome at

http://www.charite.de/bioinf/hoppe

http://bioinf.charite.de/haystack