




Protein comparison: what can we learn from 3D superimposition

Functional annotation *in silico* by homology search

```
ADH1_SULSO  -----MRAVRLVEIGKP--LSLQEIGVPKPKGPQVLIKVEAAGVCHSDVHMRQGRFGNLRIVE
ADH_CLOBE   -----MKGFAMLGINKLG---WIEKERPVAGSYDAIVRPLAVSPCTSDIHTVFEGA-----
ADH_THEBR   -----MKGFAMLSIGKVG---WIEKEKPAPGPFDAIVRPLAVAPCTSDIHTVFEGA-----
ADH1_SOLTU  MSTTVGQVIRCKAAVAWEAGKP--LVMEEDVDVAPPQKMEVRLKILYTSLCHTDVYFWEAKG-----
ADH2_LYCES  MSTTVGQVIRCKAAVAWEAGKP--LVMEEDVDVAPPQKMEVRLKILYTSLCHTDVYFWEAKG-----
ADH1_ASPFL  ----MSIPEMQWAQVAEQKGGP--LIYKQIPVPKPGPDEILVKVRYSGVCHTDLHALKGDW-----
```

Sequence comparison is performed with alignment programs

 Sequence identity $\geq 30\%$  Similar 3D  Similar function ?

Methods for similarity searches:

BLAST, Psi-BLAST (<http://www.ncbi.nlm.nih.gov/BLAST/>)

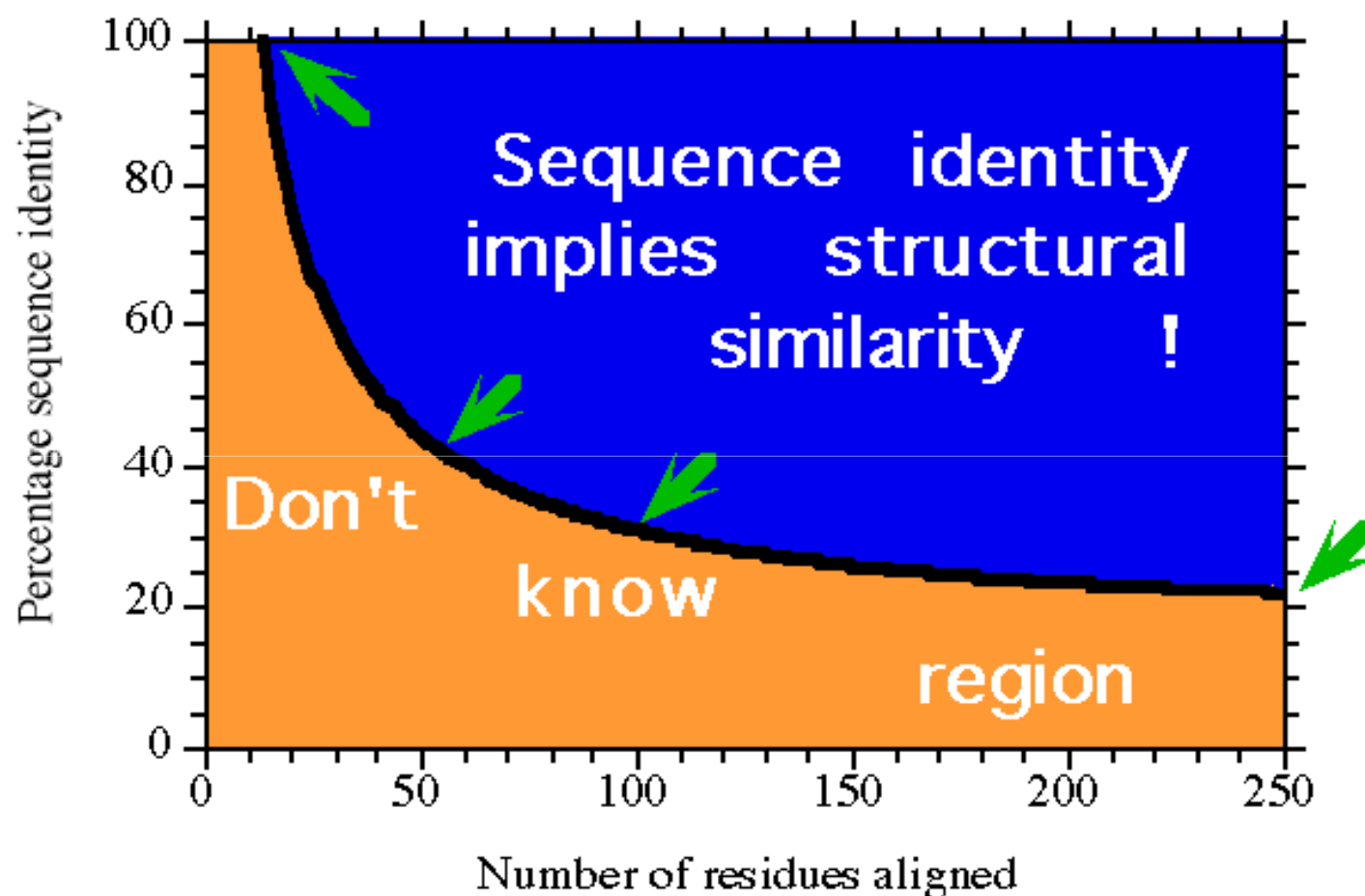
Altschul et al., (1990) J Mol Biol 215:403-410

Altschul et al., (1998) Nucleic Acids Res. 25:3389-3402

Pfam (<http://pfam.wustl.edu/hmmsearch.shtml>)

Bateman et al., (2000) Nucleic Acids Research 28:263-266

Evolution did it !



Framework for Functional Annotation through the Web

- PRIMARY STRUCTURE ANALYSIS
- SIMILARITY SEARCH (BLAST, PFAM)
- PROTEIN MOTIF SEARCH (PROSITE, INTERPRO, BLOCKS)
- MULTIPLE SEQUENCE ALIGNMENT (EVOLUTION INFORMATION)
- SECONDARY STRUCTURE PREDICTION
- 3D-STRUCTURE MODELLING
- FUNCTIONAL ANNOTATION WITH RELIABILITY MEASURE

The folding problem.....

Protein sequences are frustrated systems

KVFGRC ELAA AMKRHGLDNY RGYS LGNWVC AAKFESNFNT
QATNRNTDGS TDY GILQINS RWWCNDG RTP GSRNL CNIPC
SALLSSDITA SVNCAKKIYS DGN G MNAWVA WPMRCKGTDV
QAWIRGCRL

Too many opposite tendencies in polar solutionsThe major driving force for protein folding is the hydrophobic effect

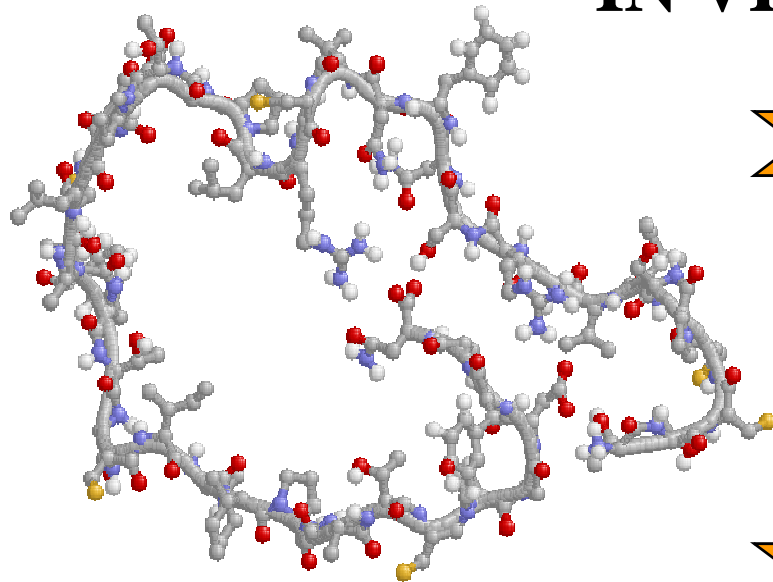
<http://webhost.bridgew.edu/fgorga/proteins/default.htm>

The folding problem

Protein sequence

Protein 3D structure

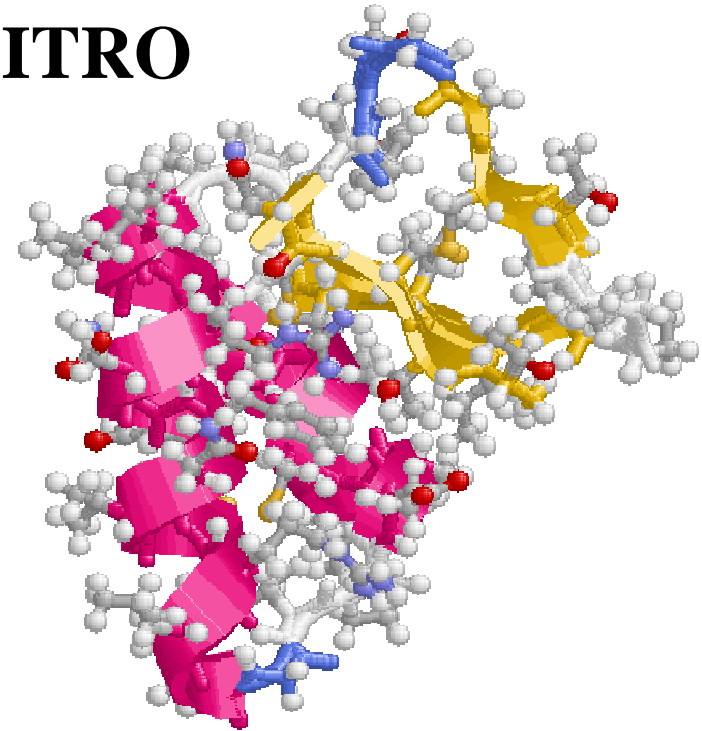
IN VIVO & IN VITRO



?



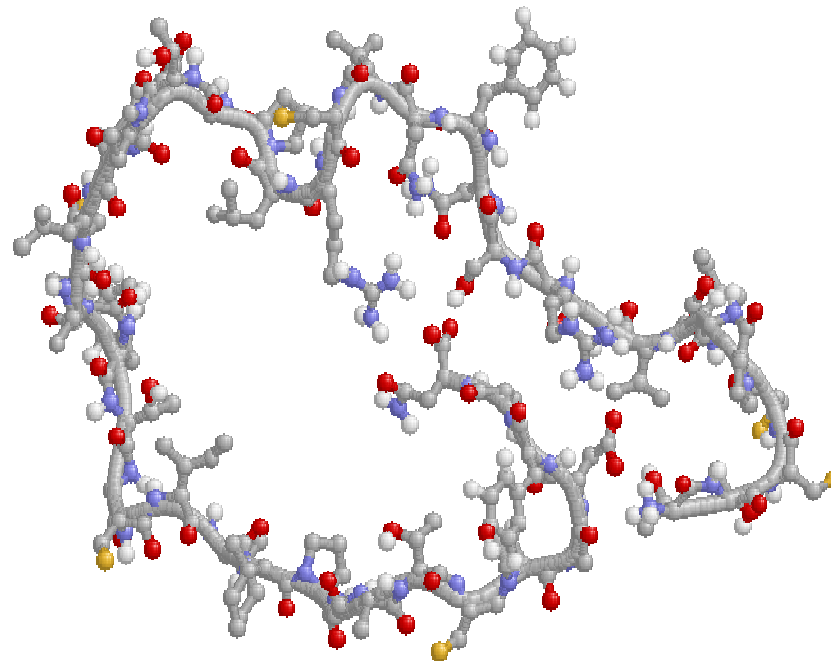
IN SILICO



The Protein Folding prediction

The folding process

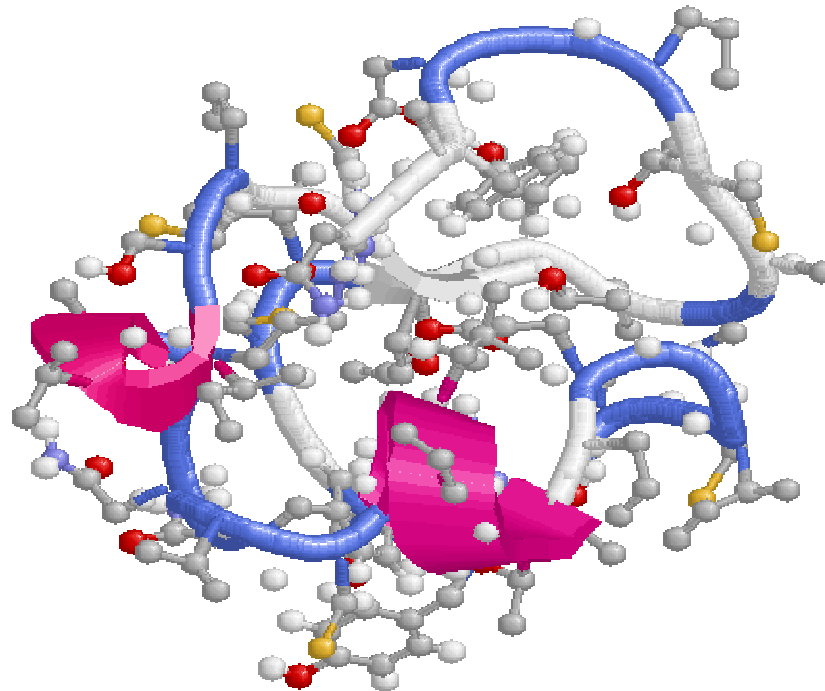
The native chain



The Protein Folding prediction

The folding process

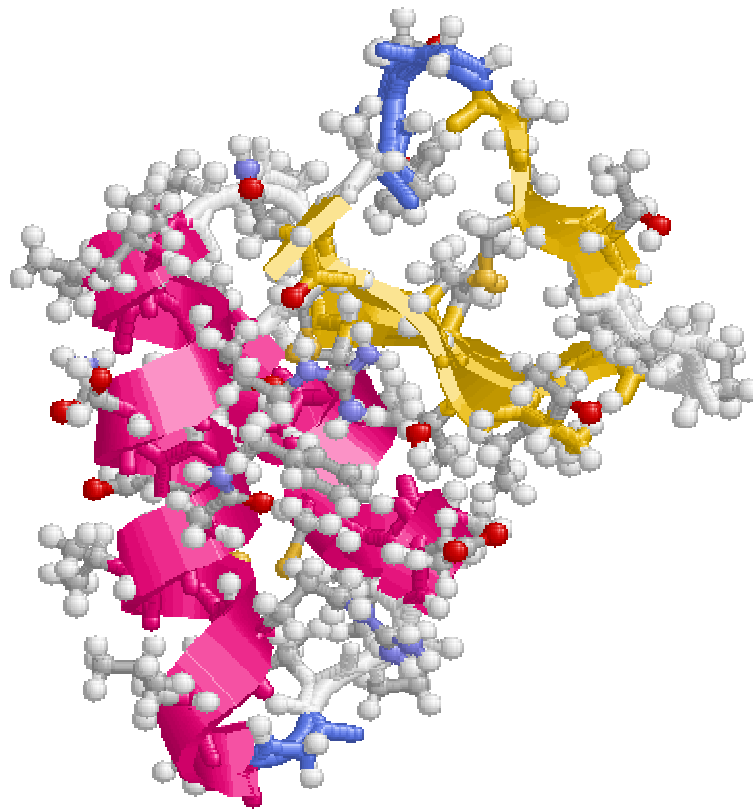
The folding initiation sites



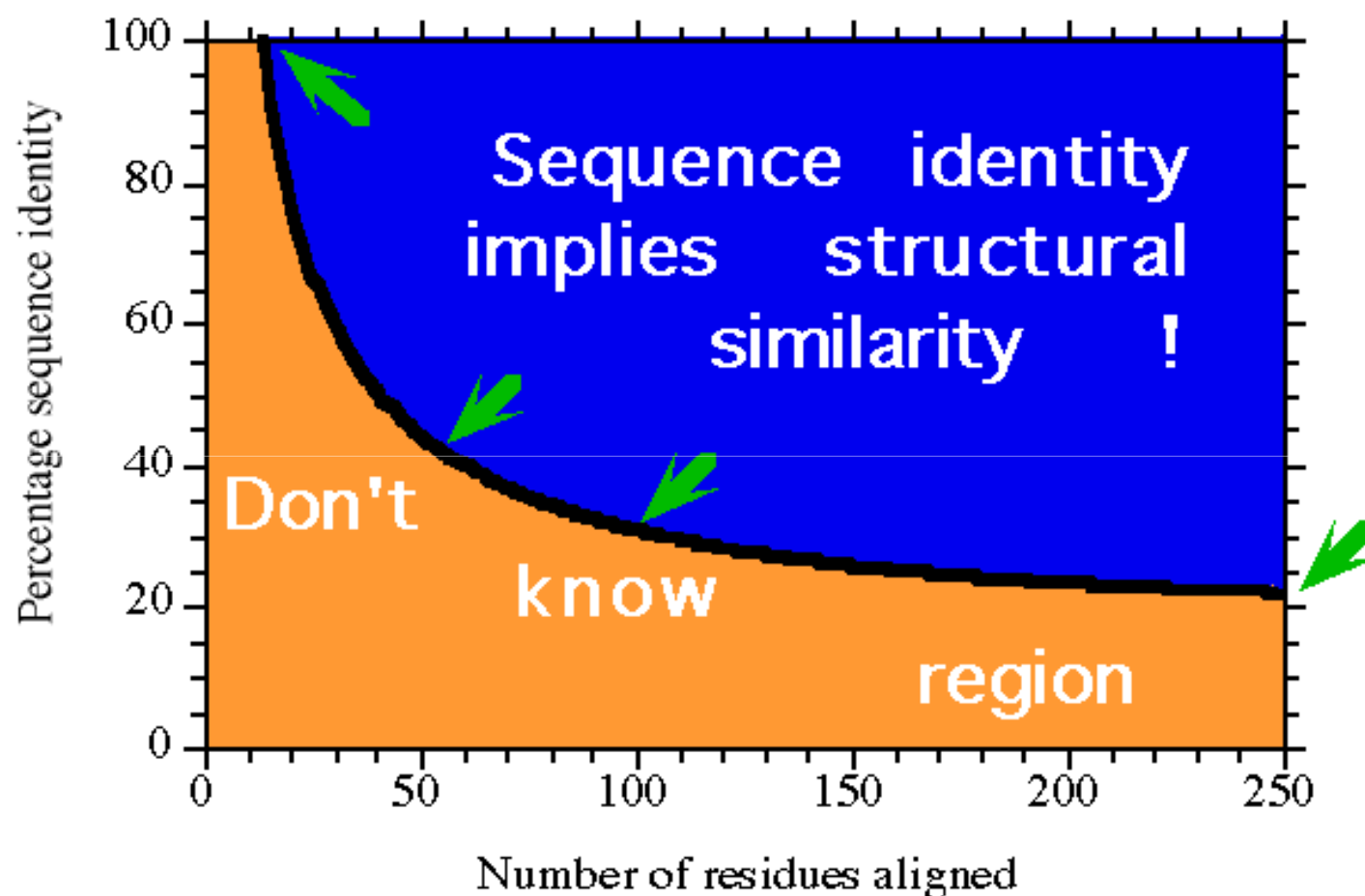
The Protein Folding prediction

The folding process

The native structure (3D)



Evolution did it !

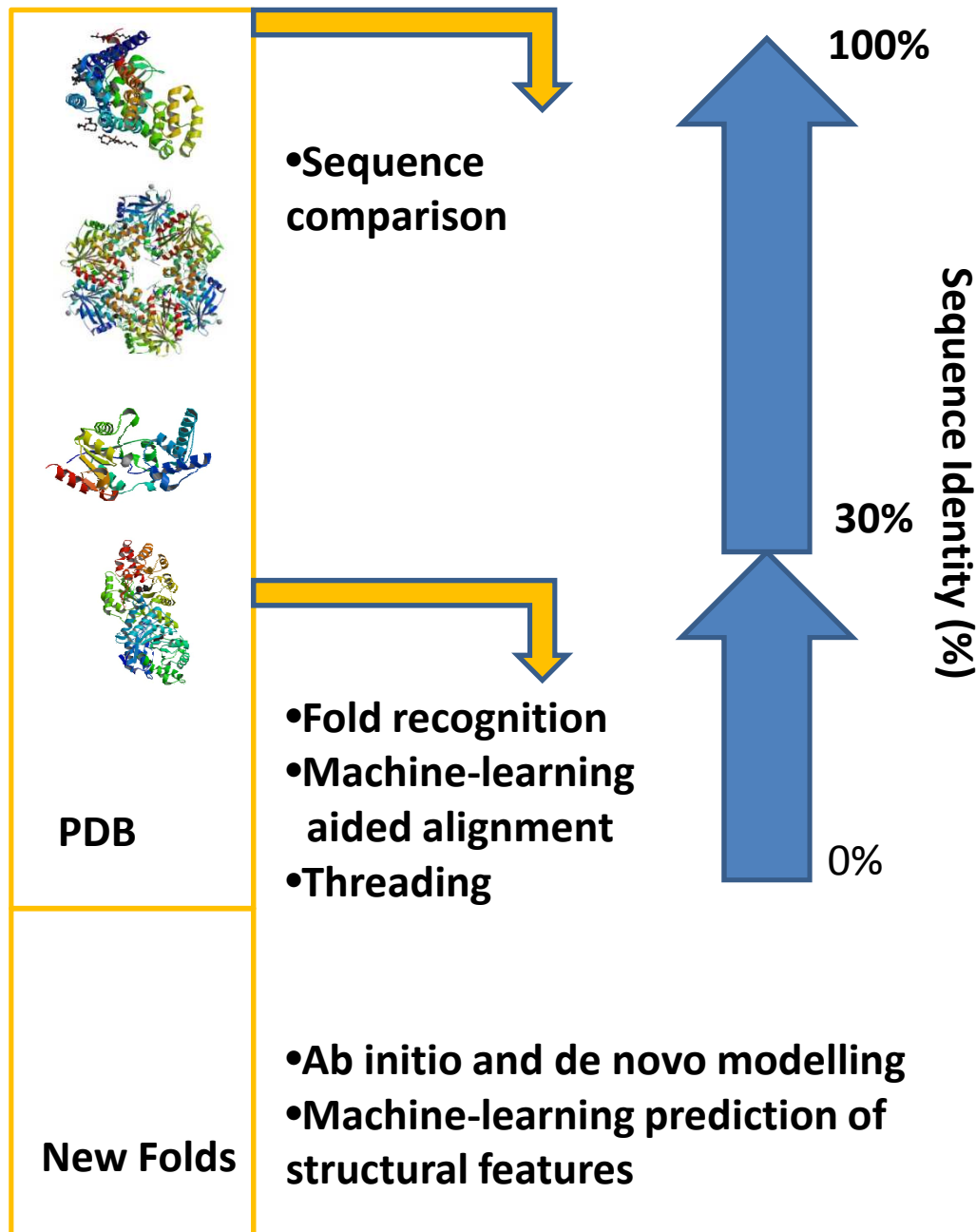


Heuristic solutions to the
computation of the protein 3D
structure from its sequence:

1) building by comparison

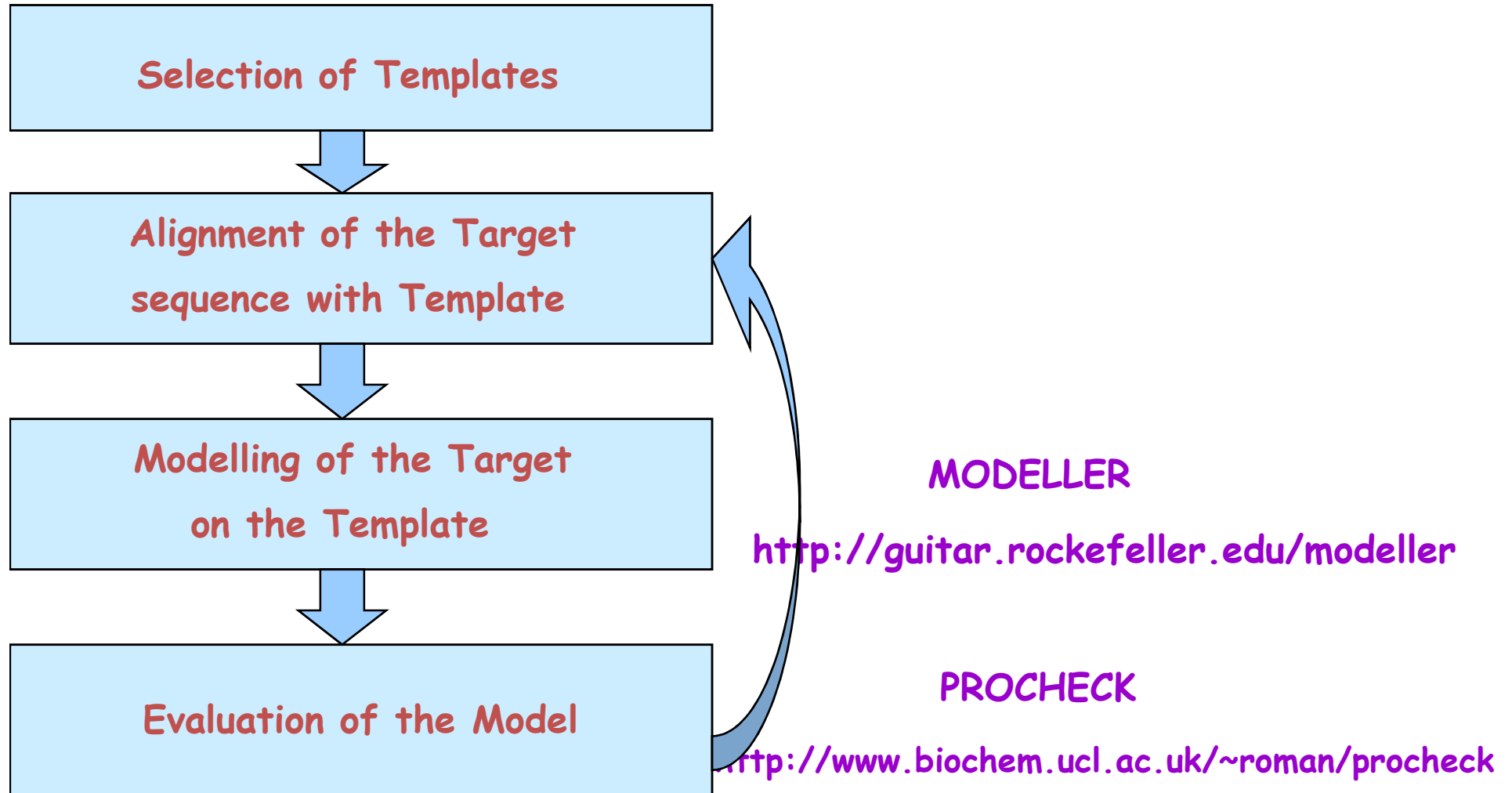
2) threading

3) *Ab initio*



Methods depend on the extent of sequence identity among the target and the template/s

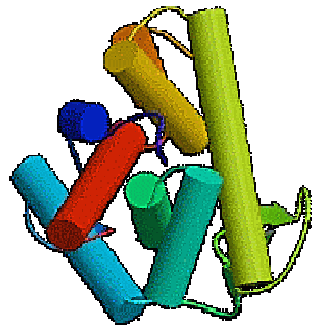
Comparative Modelling



Threading

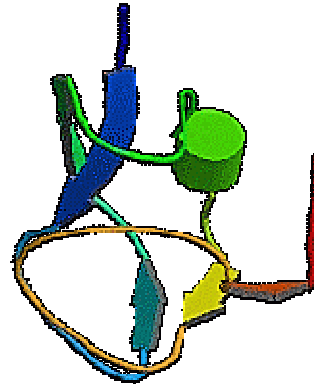
Thread the SequenceACDGGTKLMAG..... into

Model 1



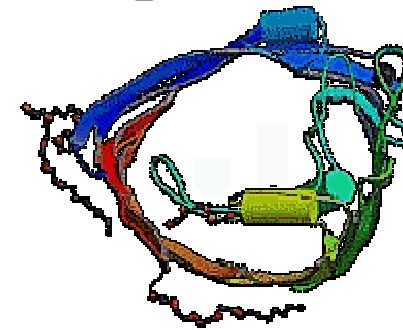
Score 1

Model 2



Score 2

Model 3



Score 3

The best scoring model is chosen as *the* candidate fold for the sequence

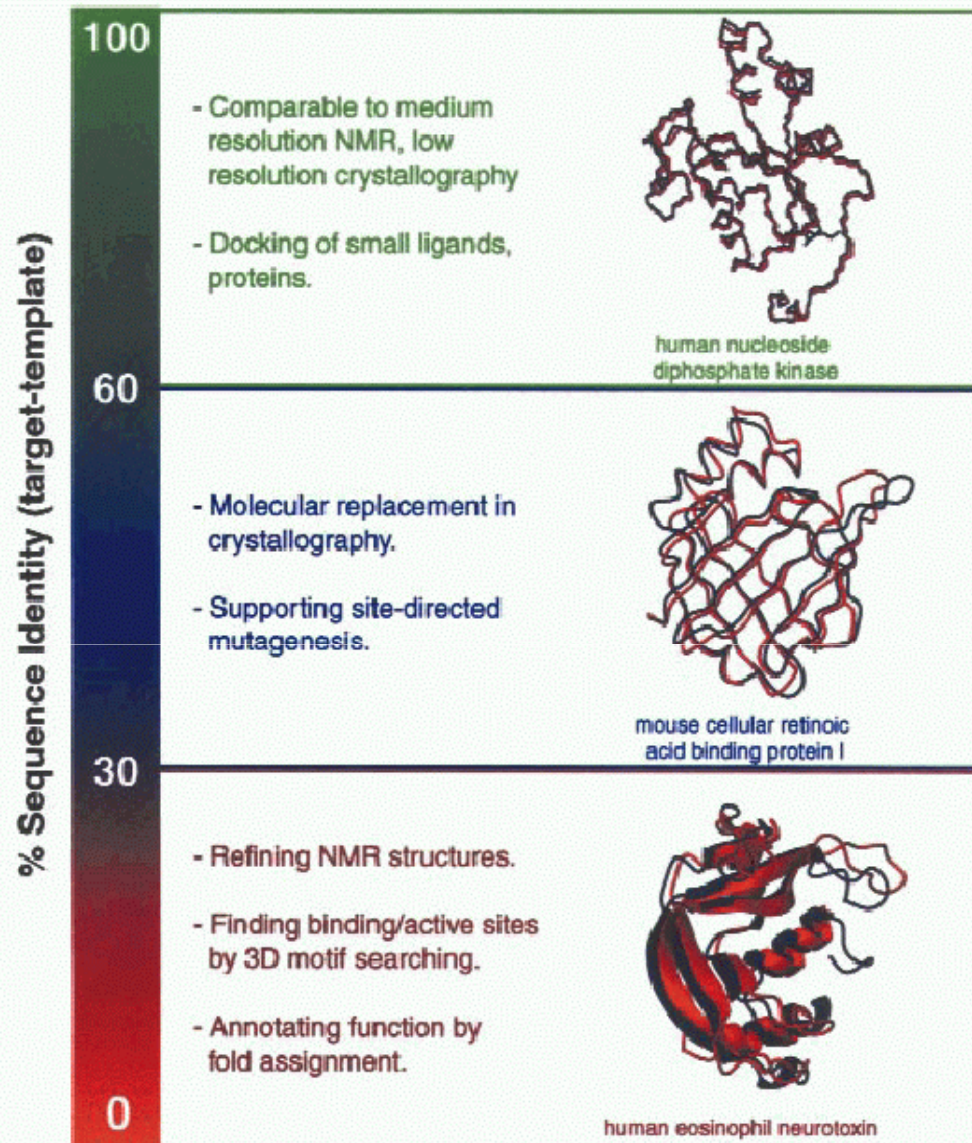


Fig. 2 Applications of comparative modeling. The potential uses of a comparative model depend on its accuracy. This in turn depends significantly on the sequence identity between the modeled sequence and the known structure on which the model was based. Sample models and corresponding experimental structures are shown on the right. Reproduced with permission from ref. 34.

Sánchez R *et al.*, Nat Struct Biol, Suppl. Nov 2000:991-994



Swiss Institute of
Bioinformatics

BIOZENTRUM

Universität Basel
The Center for Molecular Life Sciences



SWISS-MODEL

Modelling

[myWorkspace](#)

[Automated Mode](#)

[Alignment Mode](#)

[Project Mode](#)

Tools

[Template Identification](#)

[Domain Annotation](#)

[Structure Assessment](#)

[Template Library](#)

Repository

[Search by Sequence](#)

[Search by AC](#)

[Search by full text](#)

Documentation

SWISS-MODEL is a fully automated protein structure homology-modeling server, accessible via the ExPASy web server, or from the program DeepView (Swiss Pdb-Viewer). The purpose of this server is to make Protein Modelling accessible to all biochemists and molecular biologists worldwide.

What's new?

- New **Beta SWISS-MODEL** pipeline for automated model building with oligomers and ligands is now available for testing.
- Find more news on **SWISS-MODEL Blog**
- ... faster news on **Twitter**
- Follow us on **Facebook**

SWISS-MODEL Team

Torsten Schwede: Project Leader
Florian Kiefer: SWISS-MODEL Repository
Lorenza Bordoli: Method Development and support
Konstantin Arnold: SWISS-MODEL Worksp

References:

When you publish or report results using SWISS-MODEL, please cite the relevant publications:

- Arnold K., Bordoli L., Kopp J., and Schwede T. (2006). The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. *Bioinformatics*, 22,195-201.
- Kiefer F, Arnold K, Künzli M, Bordoli L, Schwede T (2009). The SWISS-MODEL Repository and associated resources *Nucleic Acids Research*. 37, D387-D392.
- Peitsch, M. C. (1995) Protein modeling *E-mail Bio/Technology* 13: 658-660.



ModBase: Database of Comparative Protein Structure Models

• [Sali Lab Home](#) • [ModWeb](#) • [ModLoop](#) • [ModBase](#) • [ModEval](#) • [PCSS](#) • [FoXS](#) • [IMP](#) • [ModPipe](#)

[ModBase Home](#) [ModBase Datasets for User:Anonymous](#) [User Login](#) [Help](#) [News](#) [Contact](#) [Current I](#)

- [General Information](#)
- [Statistics](#)
- [News](#)
- [Project Pages](#)
- [Authors and Acknowledgements](#)
- [Publications](#)
- [Related Resources](#)

Please address inquiries to:
modbase@salilab.org

ModBase Search Form

Search type  Display type 

All available datasets are selected 

Search by properties

Property 
Organism  or

[Advanced search](#)

Users of ModBase are requested to cite this article in their publications:

[MODBASE, a database of annotated comparative protein structure models and associated resources](#), Ursula Pieper, Ben M. W. Schneidman-Duhovny, Avner Schlessinger, Hannes Braberg, Zheng Yang, Elaine C. Meng, Eric F. Pettersen, Conrad C. Huang, Sampathkumar, Madhusudhan MS, Kimmer, Sielander, Thomas E. Ferrin, Stephen K. Burley, Andrei Sali *Molecular Acids Research* **30**, 46