



# Integrating Relibase+ information with other data sources

## Aim

To illustrate how a major pharmaceutical company integrated structural information from Relibase+<sup>1</sup> into their Phylosopher<sup>2</sup> database. The aim was to make structural data available to researchers in target discovery via the Phylosopher interface.

## Introduction

Relibase+ is a tool for search, retrieval and analysis of protein-ligand structures and binding sites. The data in Relibase+ is obtained from the Protein Data Bank (PDB). All the data in Relibase+ is fully curated and updates are provided on a regular basis. Furthermore, Relibase+ can be used to store in-house protein structures. Thus, both public and in-house data can be accessed through a single source.

Phylosopher is extensively used by researchers in target discovery, who retrieve gene and sequence-related data for any putative target from Phylosopher.

Along with the gene and sequence data, researchers working in target discovery at Bayer Schering Pharma were interested in obtaining information on whether or not 3D structural data for a given protein was available. The availability of a protein structure for a putative target enables an assessment of the target's druggability. It also provides opportunities to support lead finding by protein-structure-based design methods.

Historically, Bayer Schering Pharma has stored all its in-house protein structures in Relibase+. Over the years a vast range of binding sites have been investigated for many different indications. This has led to a wealth of information being stored in Relibase+. It may therefore happen that an "old" target is under consideration for a new indication. The researchers in target discovery wanted this information to be immediately accessible through Phylosopher, i.e. for each protein entry page in Phylosopher to have a link to any corresponding protein structure entries in Relibase+.



## Method

The amino acid sequences of all the proteins in Relibase+ were exported and corresponding protein entries in Phylosopher were identified using BLAST searches. For each hit, a link to the URL of the corresponding Relibase+ entry was created in Phylosopher using its Xtend API<sup>3</sup>. The process was automated by a set of BASH<sup>4</sup> scripts.

## Results

Whenever structural data for a protein is available in Relibase+, the corresponding Phylosopher pages now contain a link to the respective protein information page in Relibase+.

*“This has led to structural data being used more widely across the organisation and has resulted in increased interactions between the researchers working in target discovery and those working in drug discovery.” (Judith Günther, Bayer Schering Pharma)*

Phylosopher 8.1.2 - Proteins - Hsap36:NP\_000058.1 - Microsoft Internet Explorer bereitgestellt von Bayer Schering Pharma BLN

Adresse [http://by-bfx.bayer-ag.com/phyloProd/plsql/phyloProd/phy\\_protein.display?p\\_ges\\_ids=1003325658&p\\_summary=true](http://by-bfx.bayer-ag.com/phyloProd/plsql/phyloProd/phy_protein.display?p_ges_ids=1003325658&p_summary=true)

user: sgqon | mode: advanced | access: manager Search  Go

Genedata Phylosopher® 8.1.2 | Copyright © 1998-2010

**Hsap36:NP\_000058.1 (1 protein)**

Show 25 items on page sorted by Custom Order as summary Go ?

Target View Sequence Transcript Literature Expression Structure Related Data

NP\_000058.1 [Homo sapiens (Hsap36)]

**PDB Sequence Hits**

Query Accession	PDB Class	PDB Accession	Percent Identity	Alignment Length	Bit Score	E-Value
<input type="checkbox"/> NP_000058	rel	<input type="checkbox"/> <a href="#">pdb1ca3-A</a>	100.00	260	1803	1.2e-110
<input type="checkbox"/> NP_000058	rel	<input type="checkbox"/> <a href="#">pdb1cnw-A</a>	100.00	260	1803	1.2e-110
<input type="checkbox"/> NP_000058	rel	<input type="checkbox"/> <a href="#">pdb1cnx-A</a>	100.00	260	1803	1.2e-110
<input type="checkbox"/> NP_000058	rel	<input type="checkbox"/> <a href="#">pdb1cny-A</a>	100.00	260	1803	1.2e-110
<input type="checkbox"/> NP_000058	rel	<input type="checkbox"/> <a href="#">pdb1eou-A</a>	100.00	260	1803	1.2e-110
<input type="checkbox"/> NP_000058	rel	<input type="checkbox"/> <a href="#">pdb1g6v-A</a>	100.00	260	1803	1.2e-110
<input type="checkbox"/> NP_000058	rel	<input type="checkbox"/> <a href="#">pdb1hca-A</a>	100.00	260	1803	1.2e-110
<input type="checkbox"/> NP_000058	rel	<input type="checkbox"/> <a href="#">pdb1kwq-A</a>	100.00	260	1803	1.2e-110
<input type="checkbox"/> NP_000058	rel	<input type="checkbox"/> <a href="#">pdb1kwr-A</a>	100.00	260	1803	1.2e-110
<input type="checkbox"/> NP_000058	rel	<input type="checkbox"/> <a href="#">pdb1t9n-A</a>	100.00	260	1803	1.2e-110

The report output was limited to a subset of rows. [SHOW ALL](#)

Alignment Please contact [Judith Günther](#) to get access to Relibase

New Search Functional Analysis [Select] Go

Figure 1 – Phylosopher page containing links to protein structures in Relibase+.



Protein Information for PDB Entry 1ca3 - Microsoft Internet Explorer bereitgestellt von Bayer Schering Pharma BLN

Adresse [http://by-cc-relibase3.bayer-ag.com/display\\_protein\\_and\\_ligand.php?search\\_string=pdb1ca3&search\\_type=protein&search\\_domain=reli](http://by-cc-relibase3.bayer-ag.com/display_protein_and_ligand.php?search_string=pdb1ca3&search_type=protein&search_domain=reli)

**Relibase+**

Home Text Search Sequence Search SMILES Search Sketcher Hitlists Stored Results Help

Protein Information for PDB Entry 1ca3 Relibase+ 3.0.1

View PDB Header Save PDB File PDB Website Bookmark

### Protein Entry 1ca3

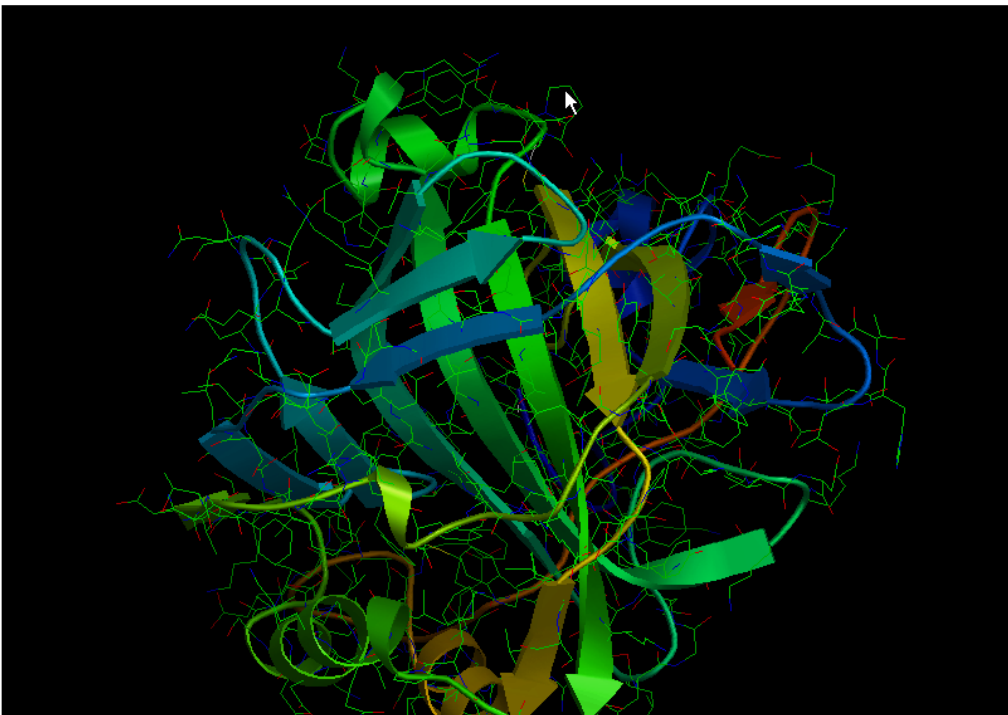


Figure 2 – Relibase+ protein page accessed from a link in Phylosopher.

## Conclusions

Relibase+ is a versatile tool for communicating structural information and it can easily be integrated with other software and databases to broaden access to protein structural data.

## References

1. M. Hendlich, A. Bergner, J. Günther, G. Klebe, *J. Mol. Biol.*, **326**, 607-620, 2003
2. <http://www.genedata.com/products/phylosopher/>
3. <http://www.genedata.com/products/phylosopher/architecture/programming-interfaces.html>
4. <http://www.gnu.org/software/bash/>



---

## Products

Relibase+ - an essential tool for searching, exploring and comparing all protein-ligand data from public and in-house data sources

For further information please contact Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, UK. Tel: +44 1223 336408, Fax: +44 1223 336033, Email: [admin@ccdc.cam.ac.uk](mailto:admin@ccdc.cam.ac.uk)