

A system for automated validation of GPCRs homology models against mutational data

Stefan Mordalski^{1*}, Tomasz Kosciolk¹, Kurt Kristiansen², Ingebrigt Sylte², Zdzisław Chilmonczyk³, Andrzej J Bojarski¹

¹Department of Medicinal Chemistry, Institute of Pharmacology Polish Academy of Sciences, 12 Smetna St, 31-343 Krakow, Poland

²Medical Pharmacology and Toxicology, Department of Medical Biology, Faculty of Health Science, University of Tromsø, N-9037 Tromsø, Norway

³National Medicines Institute, 30/34 Chelmska Street, 00-725 Warszawa
e-mail: stefanm@if-pan.krakow.pl

Introduction

Sequence alignment between target and template sequence is the most troublesome stage of homology modeling protocol. Misplacing amino acids responsible for interactions with ligands may lead to improper binding mode of so created model and render it useless. This is the reason of wide usage of mutational data in either aligning sequences or models verification.

In this study we present a tool allowing automated comparison of mutagenesis data retrieved from tinyGRAP [1] database with corresponding residues of the model. tinyGRAP dataset is queried for the investigated sequence and its close homologs (i.e. group members), and substitution mutations are retrieved. Query results are then checked whether appropriate residues face inside of the receptor (with some margin), and if not, the tool produces report in PyMol .pse file pointing amino acids violating mutational „constrains”.

Database preparation

In the first step the tinyGRAP database was prepared for automated usage. Since the dataset is a collection of files containing mutational data, the appropriate parser was prepared to convert it into form suitable for RDBM engine. An SQLAlchemy python module was used to provide convenient mapping the database onto the actual python class. Such approach allows picking from wide choice of database engines, from SQLite to Oracle, without any major changes in the script. In this study the SQLite engine was selected due to its simplicity.

In addition, we plan to adapt the original GRAP database in the same fashion as the above. The GRAP database provides additional information about the changes in affinity caused by the mutation. It would give additional information to assume the mutation importance, and so to evaluate model sensibility.

Model verification workflow

Verification of the input model goes in several steps (Fig. 1). First, the chain sequences retrieved from the model are located within full SwissProt sequence of the protein, and residue numbers from SwissProt record are assigned to corresponding ones in the model (A). In the next step, tinyGRAP database is queried against desired protein and resulting amino acids are localised within the model structure (A, residue in red).

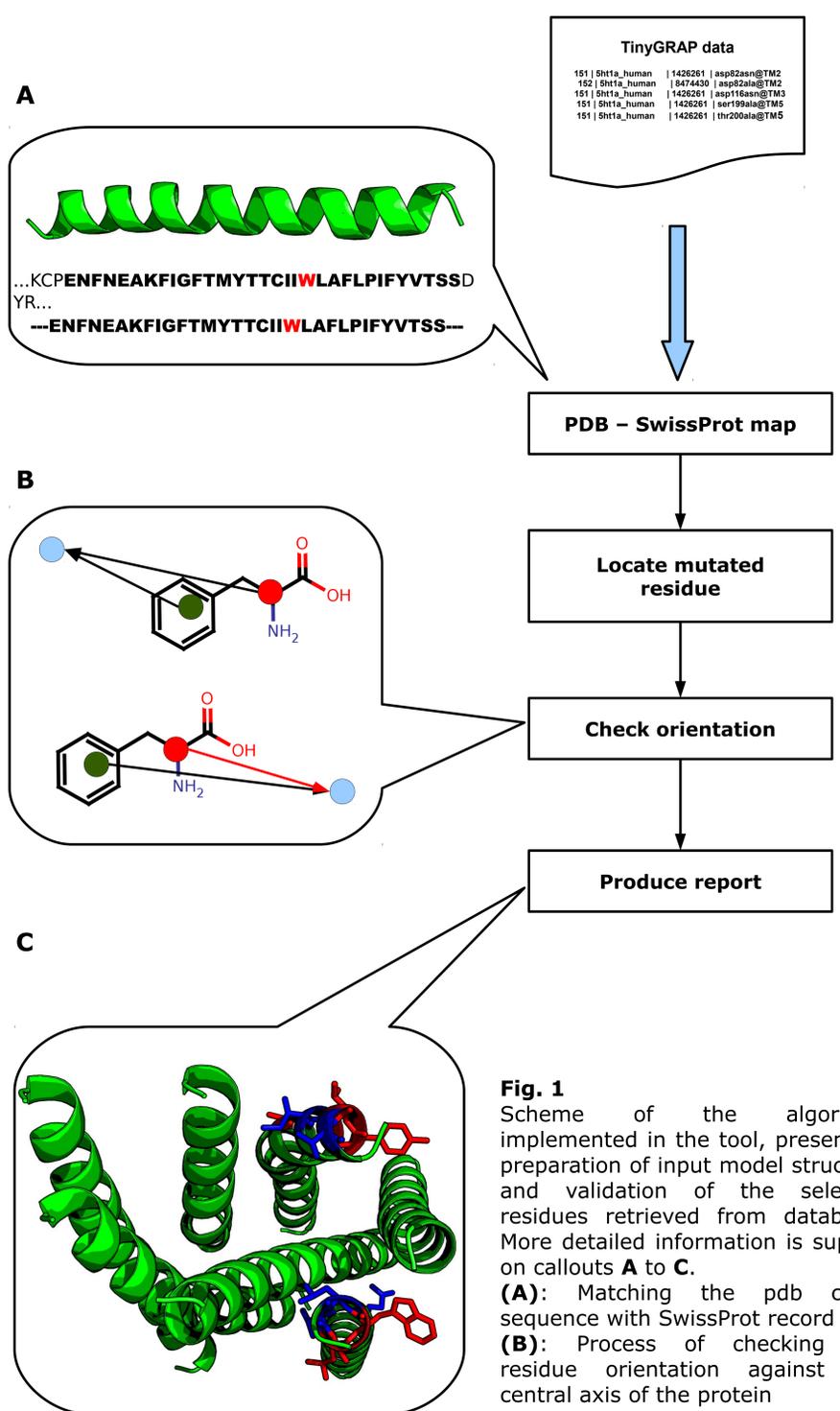
For the residues found in the structure a geometric centre of the sidechain (B, green dot) is allocated and its distance from central axis (B, blue dot) of the protein is calculated. Analogous range is computed for the alpha carbon (B, red dot) of selected amino acid. Residue is then labeled as 'good' if sidechain centre is closer to the main axis than the alpha carbon, and 'bad' flag is turned on in the opposite.

Finally, the report is produced (C). It is a PyMol session file containing input model along with the textual description of the evaluation (residue number, its verification status along with authors and title of the publication containing the point mutation used).

Conclusions and future plans

The program appears to be a very convenient tool for automating model verification against the mutational data. In rare peculiar cases it can make mistakes in flagging the amino acids, and thus visual inspection of the report is recommended. Still, the tool reduces the amount of work required to locate and check the mutations for the given homology model.

In the nearest future we plan to use the original GRAP database, providing extended mutational data. A web service versions of the tool is also in preparation.



Acknowledgments

The study was partly supported by a grant PNRF-103-AI-1/07 from Norway through the Norwegian Financial Mechanism



Polish-Norwegian
Research Fund

norway
grants

Literature

1. Beukers MW, Kristiansen K, IJzerman AP, Edvardsen I, **TinyGRAP database: a bioinformatics tool to mine G protein-coupled receptor mutant data.**, Trends Pharmacol Sci. 1999 Dec; **20(12)**:475-7.

Additional python modules used: mmlibs, Bio, SQLAlchemy.
Visualizations were prepared with PyMol

Poster can be downloaded from www.cns-platform.eu