

Research

Open Access

In silico saturation mutagenesis and docking screening for the analysis of protein-ligand interaction: the Endothelial Protein C Receptor case study

Federica Chiappori^{1,2}, Pasqualina D'Urso^{1,2}, Ivan Merelli¹, Luciano Milanese¹ and Ermanna Rovida*¹

Address: ¹Institute for Biomedical Technologies – National Research Council (ITB-CNR), via Fratelli Cervi 93, 20090 Segrate (MI), Italy and ²Consorzio Interuniversitario Lombardo per l'Elaborazione Automatica (CILEA), via Raffaello Sanzio 4, 20090 Segrate (MI), Italy

E-mail: Federica Chiappori - federica.chiappori@itb.cnr.it; Pasqualina D'Urso - pasqualina.durso@itb.cnr.it;

Ivan Merelli - ivan.merelli@itb.cnr.it; Luciano Milanese - luciano.milanese@itb.cnr.it; Ermanna Rovida* - ermanna.rovida@itb.cnr.it

*Corresponding author

from Bioinformatics Methods for Biomedical Complex Systems Applications (NETTAB2008)
Varenna, Italy 19–21 May 2008

Published: 15 October 2009

BMC Bioinformatics 2009, 10(Suppl 12):S3 doi: 10.1186/1471-2105-10-S12-S3

This article is available from: <http://www.biomedcentral.com/1471-2105/10/S12/S3>

© 2009 Chiappori et al; licensee BioMed Central Ltd.

This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: The design of mutants in protein functional regions, such as the ligand binding sites, is a powerful approach to recognize the determinants of specific protein activities in cellular pathways. For an exhaustive analysis of selected positions of protein structure large scale mutagenesis techniques are often employed, with laborious and time consuming experimental set-up. 'In silico' mutagenesis and screening simulation represents a valid alternative to laboratory methods to drive the 'in vivo' testing toward more focused objectives.

Results: We present here a high performance computational procedure for large-scale mutant modelling and subsequent evaluation of the effect on ligand binding affinity. The mutagenesis was performed with a 'saturation' approach, where all 20 natural amino acids were tested in positions involved in ligand binding sites. Each modelled mutant was subjected to molecular docking simulation and stability evaluation. The simulated protein-ligand complexes were screened for their impairment of binding ability based on change of calculated K_i compared to the wild-type.

An example of application to the Endothelial Protein C Receptor residues involved in lipid binding is reported.

Conclusion: The computational pipeline presented in this work is a useful tool for the design of structurally stable mutants with altered affinity for ligand binding, considerably reducing the number of mutants to be experimentally tested. The saturation mutagenesis procedure does not require previous knowledge of functional role of the residues involved and allows extensive exploration of all possible substitutions and their pairwise combinations. Mutants are screened by

docking simulation and stability evaluation followed by a rationally driven selection of those presenting the required characteristics. The method can be employed in molecular recognition studies and as a preliminary approach to select models for experimental testing.

Background

Structure-based site-directed mutagenesis is a widely used approach to elucidate and modify specific aspects of protein function and to investigate the properties of protein-ligand interactions. Prediction of mutants with desirable properties is often obtained by rational design of a few specific position in the region of interest. This approach requires an in-depth knowledge of physicochemical, structural and functional properties of the protein, which may not be exhaustively provided by X-ray crystallography data. A different strategy is to apply the so-called scanning mutagenesis based on the systematic replacement of all residues involved in a specific function. Several examples of alanine scanning mutagenesis experiments [1] are reported in literature that illustrate its relevance for the study of molecular determinants of ligand binding [2] and protein function [3]. Even more informative is the use of the saturation mutagenesis, where all 20 natural amino acids are tested in place of the wild type residue, in or near the functional site. 'In vitro' scanning saturation mutagenesis has been applied to several case studies and has proved to be useful to modulate protein properties such as substrate enzyme specificity [4] or to identify key residues for catalytic mechanisms [5].

A particular aspect that is amenable to investigation with random mutagenesis is the study of protein-ligand interaction. Several examples are reported in literature that show how this approach has contributed to the clarification of ligand interaction in G protein coupled receptors [6,7], in nuclear receptors [8,9] and in EphA3/ephrin-A5 [10], to cite a few examples.

However, the large number of mutants that are obtained from the systematic replacement of all the residues of a functional site makes the experimental procedure laborious and time consuming. This is especially true when dealing with large functional sites, because it requires the implementation of specific high-throughput methods for mutant screening or selection [11]. 'In silico' simulation of mutagenesis and screening may represent a possible way around this problem. By exploiting the capabilities of high performance computers, it is possible to design and test a large numbers of variants with altered properties in order to restrict the number of models that are worth testing experimentally. Computational protein design methods have been employed in protein engineering to increase enzymes efficiency [12] or antibiotic resistance: for example Hayes et al. [13] applied a protein optimization

strategy to increase the resistance of bacteria toward the antibiotic cefotaxime by optimizing TEM-1 β -lactamase. Moreover, computational design tools have been developed and have proved to be successful in the design of proteins with enhanced stability and specificity, and in engineering novel protein functions [14].

We present here a high performance semi-automated computational procedure suited to the study of key residues of protein-ligand interactions, which is able to dissect the contribution of different residues to ligand binding. The procedure performs a saturation mutagenesis of all residues involved in ligand binding, and the subsequent evaluation of the effect of amino acid substitutions on ligand affinity by docking simulation and on protein structure stability. An example of application to Endothelial Protein C Receptor is reported.

Methods

Pipeline implementation

A semi automated pipeline was designed to manage mutant modelling, docking simulations and data analysis procedures by integrating public software through a number of in-house developed Perl scripts.

Mutant modelling

'In silico' side chain replacement and modelling were carried out with the routine 'mutate_model' of MOD-ELLER 9v3 [15]. 'Mutate_model' introduces a single point mutation in a user-specified residue and optimizes the mutant side chain conformation by conjugated gradient and by molecular dynamics simulation.

Docking

AutoDock4.0 [16,17] was used for docking simulation handling both ligand and mutated side chain as flexible. AutoDockTools (ADT) (<http://autodock.scripps.edu/resources/adt>) facility supported the protein and ligand set up for docking. The automatic procedure, based on a Perl script, employs the following ADT routines:

- 'prepare_receptor4' to add polar hydrogens, assign Kollman charges and convert the protein PDB file in pdbqt (the input file for AutoDock4);
- 'prepare_flexdocking' and 'prepare_flexreceptor' are applied to protein and ligand to obtain the input files for flexible docking simulation;

– ‘prepare_gpf’ and ‘prepare_dpf’ generate the parameters file for AutoGrid (gpf) and for AutoDock (dpf), respectively.

Docking simulations were performed with Lamarkian Genetic Algorithm which is reported as the most efficient and reliable method of AutoDock [16].

Stability evaluation

The ‘stability’ routine of FoldX algorithm [18] was applied to evaluate the total energy of wild type structure and mutated models by the estimation of the interactions that contribute to the protein stability.

Computational resource

Docking simulations were performed on a shared Linux cluster of 280 Opteron AMD cores 275 at 2,2 GHz, provided by the Eurotech Group (Amaro, UD, Italy), dedicated to bioinformatics applications. The system is composed of 10 chassis of 6 diskless blades, each equipped with an Infiniband 4× network card and 8 GB of RAM. The cluster nodes are accessed through the OpenPBS queuing system.

Case study

Protein structure and preparation

The crystal structure of the Endothelial protein C receptor (PDB: 1LQV), in a complex with phosphatidylethanolamine (PTY), and the Gla domain of protein C solved by X-ray crystallography at 1.6 Å [19] was retrieved from the Protein Data Bank [20]. Chain A with the corresponding ligand (PTY) molecule was used for this study. The Gla domain, crystallization water, Ca ions, N-acetyl-D-glucosamine (NAG) and 2-(acetylamino)-2-deoxy-A-D-glucopyranose (NDG) were removed from PDB file. Polar hydrogens were added at 7.4 pH, using InsightII (Accelrys, San Diego).

Docking simulation

The ADT graphics interface was employed for manual preparation of the ligand to assign the flexibility, add polar hydrogens and load Gasteiger charges. 22 rotatable bonds were assigned to the PTY ligand.

A grid box of dimensions 76 Å × 60 Å × 92 Å was constructed around the binding site, based on the co-crystallized ligand. Ligand flexible docking simulations were performed with 100 runs and 2500000 energy evaluations per run, while the other parameters were set to the default.

Results

Pipeline description

The complete scheme of the computational procedure used in this work is depicted in Figure 1. A preliminary step is performed using LIGPLOT [21] to identify the residues

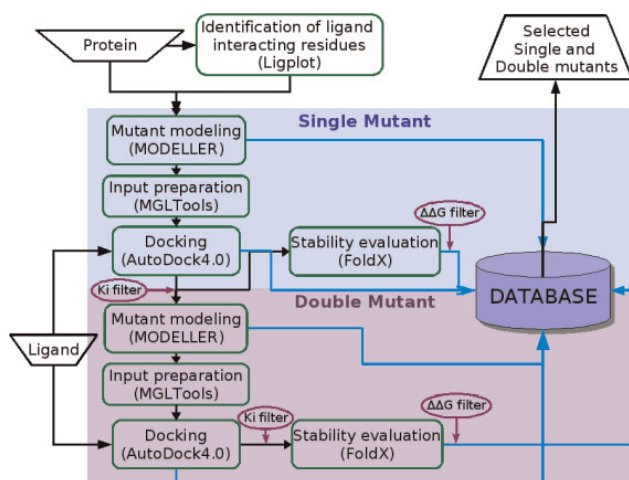


Figure 1

Schematic diagram of the pipeline. Blue panel represents the first stage of the pipeline devoted to single mutants design. Violet panel is the second stage devoted to double mutants. Steps within the panels are automatically executed. “Protein” and “Ligand” are the pipeline input files and contain the corresponding atomic coordinates. “Selected single and double mutants” are the pipeline outputs. Steps involving public software are in green boxes. Ki (inhibition constant) and $\Delta\Delta G$ (free energy difference) filters are in purple circles. Black arrows connect different steps of the pipeline, while blue arrows identify the intermediate outputs collected in the database.

involved in ligand interactions. The pipeline has a modular structure to improve its adaptability to different datasets. The backbone of the pipeline is a MySQL database where the results of each computational step are stored (blue arrows in Figure 1). The pipeline is structured in two main stages devoted to the modelling and analysis of single and double mutants respectively (blue and violet boxes).

The first stage begins by single mutant modelling starting from the receptor atomic coordinates and the list of residues identified in the preliminary step. This first step makes use of the MODELLER routine ‘mutate_model’ that performs the side chain substitution and refinement. A Perl script is used to iterate the substitution of each residue with all the amino acids to achieve the saturation mutagenesis. For each mutant, a PDB file is obtained and stored in the database.

In the following step, the pipeline prepares the input files for AutoGrid and AutoDock4.0 according to user specified parameters: PDB files are converted to PDBQT format, adding polar hydrogens and assigning Kollman charges. Substituted residues are set for flexible docking. The ligand is manually prepared as illustrated in the Methods section and given as input to the pipeline. In the next step,

AutoDock simulations are submitted to be run in parallel on the Linux cluster. Docking results are parsed to extract, for each simulation, the most representative conformation. This corresponds to the best energy conformation included in an AutoDock cluster of at least 5 elements and showing a binding energy within one Kcal/mol compared to the first ranked solution. This criterion is arbitrarily set in order to obtain a solution from a well represented group of conformations with the best scored results. However, when these conditions are not satisfied (i.e. no populated clusters are present in the top score), the first ranked solution is accepted. For each selected conformation, the corresponding values of binding energy, cluster population, inhibition constant (K_i), root mean square deviation (RMSD) and atomic coordinates are stored in the database. In order to select mutants with a modified affinity for the ligand compared to the wild type, the output data are filtered taking the wild type K_i value as threshold.

Structural stability of selected mutants is evaluated using the FoldX program that computes, for each conformation, an estimated free energy values. The free energy difference between the mutant and the wild type ($\Delta\Delta G = \Delta G_{\text{mutants}} - \Delta G_{\text{wild type}}$) is automatically calculated and stored in the database.

Pair combinations of single substitutions are analyzed in the second stage of the pipeline (violet box in Figure 1) to investigate the occurrence of synergistic effects in double mutants. Residues chosen for double mutants modelling are obtained from the results of K_i filtering in first stage screening. A Perl script manages the residues combination, removing the redundant pairs. The following pipeline steps, simulations, data analysis, filtering and storage, are performed as in the first stage.

The database is designed to collect and organize all the output data which remain available for further study. Specific queries can be applied in order to compare binding capabilities and stability of all the mutants, as illustrated in the example below.

Case study

We used the described strategy to study the ligand binding of the Endothelial Protein C Receptor (EPCR). EPCR is a transmembrane glycoprotein homologous to the major histocompatibility complex (MHC) class 1/CD1 family [22]. EPCR plays a well-characterized role in the coagulation cascade being the receptor of the anticoagulant factor Protein C (PC) and participating to its activation mechanism. Being involved in protein C pathway, EPCR also takes part in the regulation of inflammation and apoptosis in endothelial cells [23]. Similarly to CD1 molecules, it binds a phospholipid ligand in a deep hydrophobic groove

composed of an eight-stranded β -sheet floor and three almost parallel α -helices that lie above the floor and form a crevice where the lipid is located (Figure 2) [19]. The specific role of the lipid for the protein structure and function has not yet been clarified. Amino acid variants located in the lipid binding region of EPCR may provide useful information about the role of this interaction in protein function. However, mutations that impair or suppress the lipid binding have not been reported so far. In this work we predicted mutants that have reduced binding affinity for the ligand, while maintaining structural stability.

The structure of human EPCR, co-crystallized with PTY, exhibits a wide pattern of interactions that extend the length of binding groove. PTY makes contacts with 27 residues: four are located on α -helix H1, two on H2 and nine on H3, one on the loop between H1 and β -strand 5, and eleven residues on the strands of β -sheet floor (Figure 2). The binding is mostly dominated by hydrophobic interactions with the hydrocarbon chain of the ligand; a single H-bond is established between the polar group of phospholipid and ARG 156.

Employing the pipeline described above, we have carried out an 'in silico' saturation mutagenesis of the EPCR

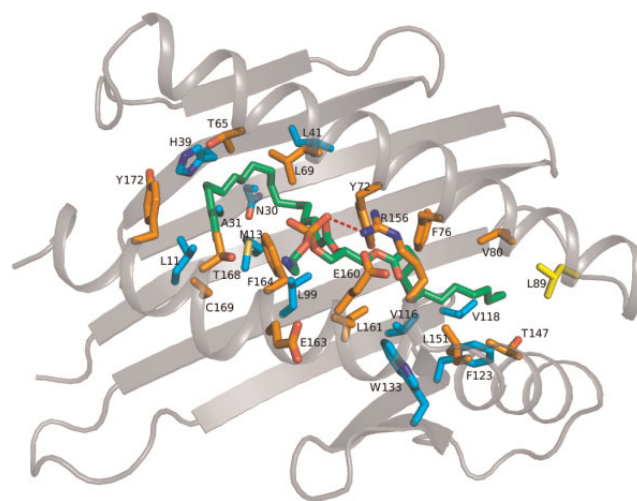


Figure 2
Cartoon representation of the 3D structure of EPCR-PTY complex. PTY ligand and EPCR residues involved in interaction, target of saturation mutagenesis, are shown in sticks. Carbon atoms are in green for the ligand, in orange for residues located on α -helices H1, H2 and H3, in cyan for residues of the eight-stranded β -sheet floor and in yellow for the loop residue. In red, the H-bond between Arg156 and the polar group of PTY. Figure is drawn with Pymol (DeLano Scientific, San Carlos, CA, USA).

residues involved in lipid binding. Each modelled mutant was subjected to molecular docking simulation and stability evaluation. The protein-ligand complexes obtained by docking simulations were screened by their inhibition constant (K_i) values in order to estimate the impairment of ligand binding ability induced by the mutation. The overall procedure and a summary of the results are depicted in Figure 3. The 27 interacting residues, identified using LIGPLOT, were saturated with the 19 remaining amino acids, producing 513 mutant models. Then, wild type and mutants were analysed by docking simulations. The wild type re-docking correctly reproduced the X-ray ligand conformation with a RMSD of 1.2 Å, a binding energy of -12.46 Kcal/mol and a K_i value of 1 nM. Positions and substitutions critical for binding affinity were identified with a K_i -based filter. Referring to the wild type K_i value, we used a threshold of 1 μ M (i.e. 3 order of magnitude higher than wild type)

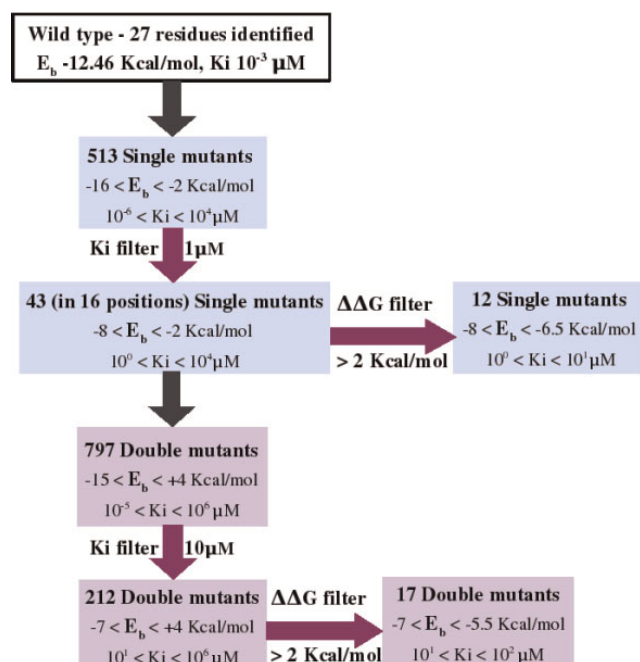


Figure 3

Summary of results for the case study. E_b (binding energy) and K_i (inhibition constant) values obtained by docking simulation of the wild type EPCR with PTY are reported. For each step of the analysis (blue and violet boxes), the number of mutants and the corresponding E_b and K_i ranges are represented. K_i and $\Delta\Delta G$ (free energy difference) filters used for mutant screening are also indicated. The analysis starts with the mutagenesis of 27 residues that yields a total of 1310 (513 single and 797 double) mutants. At the end of the screening, the number is reduced to 12 single and 17 double mutants.

to discriminate between mutants with probable impaired binding ability. The pipeline selected 43 mutants with K_i value higher than the threshold, substituted in 16 out of the 27 starting positions. Pair combinations of the 43 substitutions were generated and 797 double mutants were modelled. They were analyzed by the same docking protocol of the single mutants and filtered with a K_i threshold of 10 μ M. This value was set an order of magnitude higher than for single mutants, with the purpose of selecting those combinations with a binding ability worse than the single ones. With this filtering criterion, 212 double mutants were rescued. The large reduction in the number of double mutants (from 797 to 212), suggested that most residue combinations did not cooperate to increase the K_i value compared to the single substitutions.

The structural stability of single and double variants that passed the corresponding K_i thresholds was evaluated using the FoldX tool, directly integrated in the pipeline. Mutants ensuing a $\Delta\Delta G$ value higher than 2 kcal/mol were considered structurally unstable and removed from the selection [24]. After the screening, the pipeline retrieved 12 single and 17 double mutants with a $\Delta\Delta G < 2$ Kcal/mol. Binding energy and K_i values intervals are reported in Figure 3. The stability filter was passed only by mutants with a K_i value within 10 and 100 μ M for single and double, respectively. This value was an order of magnitude higher than the respective K_i thresholds. Mutants with high K_i values, although potential candidates for binding affinity studies, were rejected by the $\Delta\Delta G$ filter and thus considered unstable.

At the end of this analysis, 6 key positions for the ligand binding in EPCR were identified (Figure 4): position 69 and 72 are localized in the H1 α -helix, positions 156, 164 and 168 are localized in H3 α -helix and one, the position 31 on the eight-stranded β -sheet floor. The most represented substitutions of these key positions involved hydrophobic residues, such as W, in positions 31, 69 and 156, and L, in positions 31 and 164.

Pair combinations in double mutants were evaluated in terms of synergistic effect where the K_i value of double mutant was larger than the sum of the corresponding single ones. To estimate the synergistic effect, we applied the following equation:

$$K_{i_d} = K_{i_{1+2}} - (K_{i_1} + K_{i_2}) \quad (1)$$

where K_{i_1} and K_{i_2} represent the K_i value of the first and the second single mutant, and $K_{i_{1+2}}$ the K_i value of double mutant. K_{i_d} is the difference between the effect of

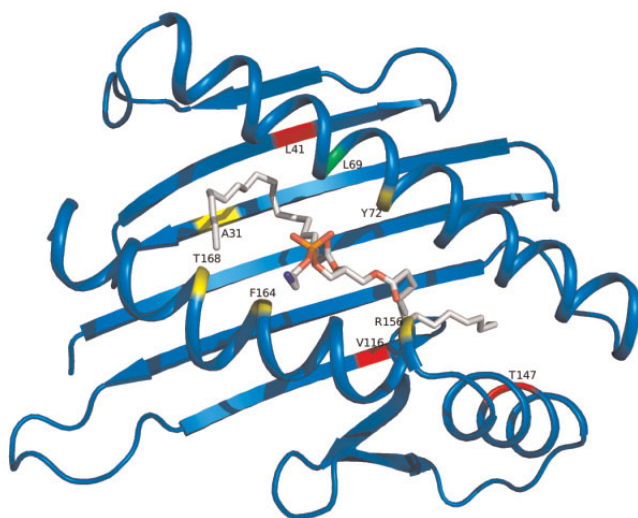


Figure 4
Key positions replaced in mutants with impaired ligand binding. Wild type EPCR residues are represented as colored spots corresponding to the C α position. L69 (in green) was found substituted only in single mutants. Residues substituted in both single and double mutants are in yellow, while those observed in double mutants only are in red. PTY ligand is represented in sticks. Figure is drawn with Pymol (DeLano Scientific, San Carlos, CA, USA).

the double mutation ($K_{i_{1+2}}$) and the sum of the two corresponding single mutations ($K_{i_1} + K_{i_2}$). Results are shown in Figure 5. We found that 8 of the 17 double mutants displayed a partial additive effect as the K_i value was included between the more damaging mutation and the sum of the two corresponding single mutants ($K_{i_1} < K_{i_{1+2}} < K_{i_1} + K_{i_2}$) (on the left of the black vertical line, in Figure 5) [25]. The remaining 9 double mutants had K_i values higher than the sum of the two corresponding single mutants ($K_{i_{1+2}} > K_{i_1} + K_{i_2}$), to which we assigned a synergistic effect compared to the single mutation (on the right of the black vertical line, in Figure 5). We observed that the highest K_i values were associated with double mutants resulting from replacement of A31 with either I or L and F164 with a small polar/apolar side chain. Between them, the highest synergistic effect was shown by A31I/F164S mutant ($K_{i_d} = 49$), while the largest effect in terms of binding energy (-5.58 Kcal/mol) and K_i (82 μ M) value was associated to A31I/F164N mutant.

In conclusion, using the described approach, we have selected a limited number of single and double mutants of EPCR for which we predict a reduced binding ability for the PTY ligand and a structural stability comparable

to the wild type. The set of 12 selected single mutants and the 9 double mutants presenting synergistic properties can be proposed as rationally constructed candidates for experimental testing in cellular models.

Pipeline performance

Due to the computational load needed to accomplish all the simulations, a parallel infrastructure was employed (see Methods). In this test case, the simulation of the single and double mutant took on average 12 and 15 hours, respectively. The whole challenge consisted of 513 docking experiments for the single mutants plus the wild type simulation and 797 for the double mutants. Depending on the resource availability, it was possible to employ 80 CPUs on average. Therefore, in a period of about ten days we collected all the results covering about 2 years of computer activity.

Conclusion

We present a high performance semi-automated computational pipeline that integrates modelling and docking simulation algorithms for large scale mutant design. Additional analysis tools such as structure stability evaluation are also included. The pipeline is organized into two main blocks for dealing with single and double mutants that are interconnected so that selected results of first analysis are used for the second part of the study. The procedure is presently devoted to protein-ligand interaction studies, but can be adapted to investigations of protein-protein interaction. The mutagenesis method is based on a saturation approach and does not require previous knowledge of functional or structural role of involved residues: it can therefore be applied to explore new binding features. The implementation of the pipeline on a parallel computer infrastructure permits a greatly reduced computational time and makes it possible to test a very large number of mutants. For example, in the reported text case we have modelled and tested 1310 variants of EPCR with the purpose of identifying single or coupled amino acid mutations that significantly impair the binding of the lipid ligand. The whole procedure reduced the number of selected models to 21, which corresponds to about 2% of the initial dataset. Selected variants provide information about the probable functional and structural effect of the mutated residues and can be proposed as subjects for experimental tests. In addition, the large amount of data stored in the database after each step of computational analysis can be further retrieved for more detailed examination and study. This pipeline can be useful for protein-ligand binding studies to identify key residues for the ligand interaction, to evaluate the effect of the substitutions on different ligands and to design selective mutants.

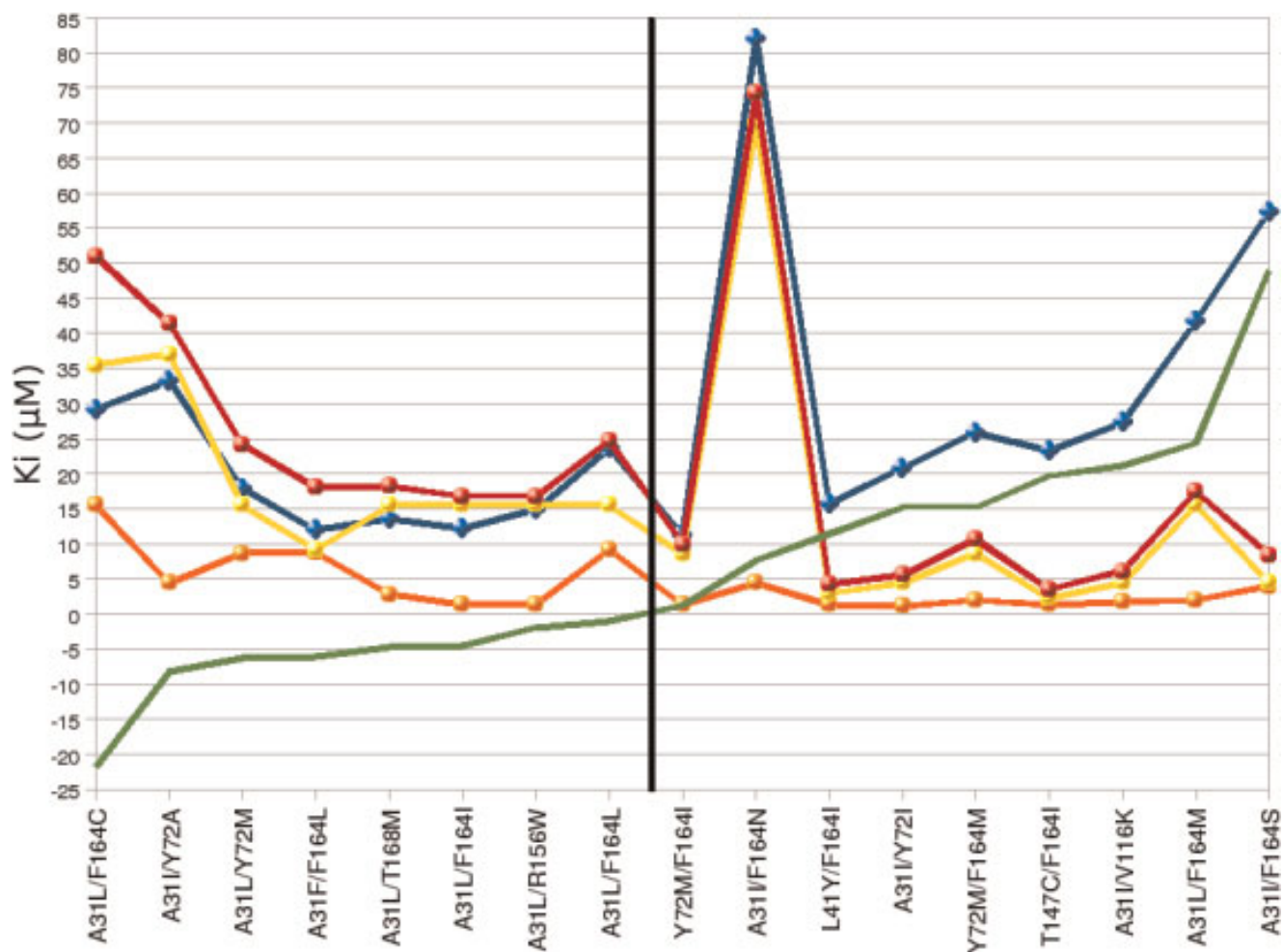


Figure 5

Synergistic effect in selected double mutants. The effect was evaluated as the difference (K_{i_d} , green line) between the K_i value of the double mutant and the K_i value of the sum of the two corresponding single mutants. $K_{i_{+2}}$: K_i values of double mutants (blue diamonds); K_{i_1} , K_{i_2} : K_i values of the two single mutants (yellow and orange dots, respectively); $K_{i_1} + K_{i_2}$: sum of the K_i values of the two corresponding single mutants (red dots).

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

FC developed the core of the pipeline, tested all the software for modelling and docking simulations, prepared the biological dataset, analyzed the output results and drafted the manuscript. PD conceived and designed the study and critically revised the results. IM contributed to developing the pipeline, designing the database and enabling the parallel computation of the docking experiments on the cluster infrastructure. LM supervised the project and made available the computational resources. ER coordinated the study and critically

revised the results and manuscript. All authors read and approved the final manuscript.

Acknowledgements

This work has been supported by the FIRB-MUR projects LITBIO (RBLA0332RH), Italian Bioinformatics Network (RBPR05ZK2Z) and Bioinformatics analysis applied to Populations Genetics (RBI-N064YAT_003), the Interdepartmental CNR-BIOINFORMATICS network and EGEE-III Enabling Grids for E-science (INFOS-RI-222667).

The authors gratefully thank Dr. E M Faioni for fruitful discussion and suggestions and Chiara Bishop for reviewing the manuscript.

This article has been published as part of *BMC Bioinformatics* Volume 10 Supplement 12, 2009: Bioinformatics Methods for Biomedical Complex System Applications. The full contents of the supplement are available online at <http://www.biomedcentral.com/1471-2105/10?issue=S12>.

References

1. Morrison KL and Weiss GA: **Combinatorial alanine-scanning.** *Curr Opin Chem Biol* 2001, **5**:302–307.
2. Williams PF, Mynarcik DC, Yu GQ and Whittaker J: **Mapping of NH₂-terminal ligand binding site of the insulin receptor by alanine scanning mutagenesis.** *J Biol Chem* 1995, **270**(7):3012–3016.
3. Hulme EC, Bee MS and Goodwin JA: **Phenotypic classification of mutants: a tool for understanding ligand binding and activation of muscarinic acetylcholine receptors.** *Biochem Soc Trans* 2007, **35**(Pt 4):742–745.
4. Geddie ML and Matsumura I: **Rapid Evolution of beta-Glucuronidase Specificity by Saturation Mutagenesis of an Active Site Loop.** *J Biol Chem* 2004, **279**(25):26462–26468.
5. Yep A, Kenyon GL and McLeish MJ: **Saturation mutagenesis of putative catalytic residues of benzoylformate decarboxylase provides a challenge to the accepted mechanism.** *Proc Natl Acad Sci USA* 2008, **105**(15):5733–5738.
6. Celic A, Connelly SM, Martin NP and Dumont ME: **Intensive mutational analysis of G protein-coupled receptors in yeast.** *Methods Mol Biol* 2004, **237**:105–120.
7. Hagemann IS, Narzinski KD, Floyd DH and Baranski TJ: **Random mutagenesis of the complement factor 5a (C5a) receptor N terminus provides a structural constraint for C5a docking.** *J Biol Chem* 2006, **281**:36783–92.
8. Lim YP and Huang JD: **Pregnane X receptor polymorphism affects CYP3A4 induction via a ligand-dependent interaction with steroid receptor coactivator-1.** *Pharmacogenet Genomics* 2007, **17**(5):369–382.
9. Dubbink HJ, Hersmus R, Verma CS, Korput van der HA, Berrevoets CA, van Tol J, Ziel-van der Made AC, Brinkmann AO, Pike AC and Trapman J: **Distinct recognition modes of FXXLF and LXXLL motifs by the androgen receptor.** *Mol Endocrinol* 2004, **18**(9):2132–2150.
10. Smith FM, Vearing C, Lackmann M, Treutlein H, Himanen J, Chen K, Saul A, Nikolov D and Boyd AW: **Dissecting the EphA3/Ephrin-A5 interactions using a novel functional mutagenesis screen.** *J Biol Chem* 2004, **279**(10):9522–31.
11. Hancock SM, Vaughan MD and Withers SG: **Engineering of glycosidases and glycosyltransferases.** *Curr Opin Chem Biol* 2006, **10**(5):509–519.
12. Tokunaga H, Arakawa T and Tokunaga M: **Engineering of halophilic enzymes: two acidic amino acid residues at the carboxy-terminal region confer halophilic characteristics to Halomonas and Pseudomonas nucleoside diphosphate kinases.** *Protein Sci* 2008, **17**(9):1603–1610.
13. Hayes RJ, Bentzien J, Ary ML, Hwang MY, Jacinto JM, Vielmetter J, Kundu A and Dahiyat BI: **Combining computational and experimental screening for rapid optimization of protein properties.** *Proc Natl Acad Sci USA* 2002, **99**(25):15926.
14. Dahiyat BI and Mayo SL: **Protein design automation.** *Protein Sci* 1996, **5**(5):895–903.
15. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, Pieper U and Sali A: **Comparative Protein Structure Modeling With MODELLER.** *Curr Protoc Bioinformatics* 2006 Unit 5.6.
16. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK and Olson AJ: **Automated Docking Using a Lamarckian Genetic Algorithm and Empirical Binding Free Energy Function.** *J Comput Chem* 1998, **19**:1639–1662.
17. Huey R, Morris GM, Olson AJ and Goodsell DS: **A Semiempirical Free Energy Force Field with Charge-Based Desolvation.** *J Comput Chem* 2007, **28**(6):1145–1152.
18. Guerois R, Nielsen JE and Serrano L: **Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations.** *J Mol Biol* 2002, **320**:369–387.
19. Oganessian V, Oganessian N, Terzyan S, Qu D, Dauter Z, Esmon NL and Esmon CT: **The crystal structure of the endothelial protein C receptor and a bound phospholipid.** *J Biol Chem* 2002, **277**:24851–4.
20. Bernstein FC, Koetzle TF, Williams GJB, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T and Tasumi M: **The Protein Data Bank: A Computer-Based Archival File for Macromolecular Structures.** *Eur J Biochem* 1977, **80**:319–324.
21. Wallace AC, Laskowski RA and Thornton JM: **LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions.** *Protein Eng* 1995, **8**:127–134.
22. Esmon CT: **The endothelial cell protein C receptor.** *Thromb Haemost* 2000, **83**:639–643.
23. Van de Wouwer M, Collen D and Conway EM: **Thrombomodulin-protein C-EPCR system: integrated to regulate coagulation and inflammation.** *Arterioscler Thromb Vasc Biol* 2004, **24**:1374–1383.
24. Tokuriki N, Stricher F, Schymkowitz J, Serrano L and Tawfik DS: **The Stability Effects of Protein Mutations Appear to be Universally Distributed.** *J Mol Biol* 2007, **369**:1318–1332.
25. Jang DS, Cha HJ, Cha SS, Hong BH, Ha NC, Lee JY, Oh BH, Lee HS and Choi KY: **Structural double-mutant cycle analysis of a hydrogen bond network in ketosteroid isomerase from Pseudomonas putida biotype B.** *Biochem J* 2004, **382**:967–973.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

