



**HAL**  
open science

## Identification of a muscle-specific MuRF1-E2 network

Cécile Polge, Stéphanie Cabantous, Christiane Deval, Agnes Claustre,  
Catherine Bouchenot, Lydie Combaret, Daniel Taillandier

► **To cite this version:**

Cécile Polge, Stéphanie Cabantous, Christiane Deval, Agnes Claustre, Catherine Bouchenot, et al.. Identification of a muscle-specific MuRF1-E2 network. Understanding Protein Interactions: from Molecules to Organisms, Apr 2017, Lyon, France. 40 p., 2017, Book of Abstracts Understanding Protein Interactions: from Molecules to Organisms. hal-01606000

**HAL Id: hal-01606000**

**<https://hal.science/hal-01606000>**

Submitted on 2 Jun 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Understanding Protein Interactions: from Molecules to Organisms

24-26 Apr 2017

Lyon

France

# Table of contents

Mutational analysis of the catalytic domain of the NSD1 methyltransferases, Priya Anand [et al.] . . . . .	3
InterEvDock: structural prediction of protein-protein interactions using evolutionary information, Jessica Andreani [et al.] . . . . .	4
COCO-tools: web applications for understanding protein interactions at a molecular level, Ida Autiero [et al.] . . . . .	6
Toward a crowded protein simulation with the Martini force field, Jonathan Barnoud [et al.] . . . . .	8
Protein loop sampling using a large structural database and reinforcement learning, Amélie Barozet [et al.] . . . . .	9
ATTRACT web-interface for macromolecular docking, Isaure Chauvot De Beauchene [et al.] . . . . .	10
On the calculation of electrostatic energies in the PaLaCe coarse-grain model for proteins and protein complexes in aqueous solution: all-atom Poisson-Boltzmann energies are reproduced with a fast approach having simple analytical derivatives., Nicoletta Ceres . . . . .	12
Flexible protein docking by systematic cross-docking of rigid candidates refined with a knowledge-based potential, Sergei Grudin [et al.] . . . . .	13
Computational study of Protein-Protein Interaction (PPI) interface of bacterial Universal Stress Protein (UspC) and Chromosome Partitioning Protein (mukB), Alok Jha . . . . .	14
Using machine learning to predict protein-ligand interactions, Maria Kadukova [et al.] . . . . .	16
Use of cross-docking simulations for identification of protein-protein interactions sites: the case of proteins with multiple binding sites, Nathalie Lagarde [et al.] . . . . .	18

The KAMI Cell Signaling Modeler, Sébastien Légaré [et al.] . . . . .	20
Multiprotein complexes of eukaryotic cell membranes: integrative bioinformatic modeling of the interactome, G Khazen [et al.] . . . . .	21
Integrative Modelling goes HPC/HTC, Adrien Melquiond [et al.] . . . . .	23
BIS2Analyzer: a server for coevolution analysis of conserved protein families, Francesco Oteri [et al.] . . . . .	24
Exhaustive Analysis of Point-Group Symmetries in Protein Assemblies, Guillaume Pagès [et al.] . . . . .	25
Identification of a muscle-specific MuRF1-E2 network, Cécile Polge [et al.] . . . . .	27
JET2 Viewer: a database of predicted multiple, possibly overlapping, protein-protein interaction sites for PDB structures., Hugues Ripoché [et al.] . . . . .	28
Design of cyclic peptide inhibitors of caspase-3 using metadynamics simulations, Jay-sen Sawmynaden [et al.] . . . . .	29
Conformational landscape of cyclic RGD peptides explored by REMD simulations, Maud Jusot [et al.] . . . . .	31
PRODIGY: a web-server for the prediction of binding affinity in biomolecular complexes, Anna Vangone [et al.] . . . . .	32
A comparative study of the mutational robustness of an ordered plant factor and a disordered viral protein involved in key potyvirus biological functions, Jocelyne Walter [et al.] . . . . .	34
A full description of the conformational landscape of cyclized peptides using a robotics approach, Maud Jusot [et al.] . . . . .	35
A Docking Molecular calculations to study of properties of the D-Mannose in the natural treatment of urinary infections caused by E- Coli, Sekkal-Taleb Nezha . . . . .	37
<b>Author Index</b>	<b>37</b>

# Mutational analysis of the catalytic domain of the NSD1 methyltransferases

Priya Anand \* <sup>1</sup>, Wolfgang Wenzel

<sup>1</sup> Karlsruhe Institute of Technology – Germany

Most histone methyltransferases (HMTase) harbor a predicted Su(var)3-9, Enhancer-of-zeste, Trithorax (SET) domain, which transfers a methyl group to a lysine residue in their histone substrates. Mutations in the SET domains are reported to cause various diseases like intellectual disability syndromes such as Sotos, Weaver, or Kabuki syndromes and cancers. Despite the intense interest in NSD1 gene, very little structural and mechanistic information is available for Sotos Syndrome point mutations. Sotos syndrome is an overgrowth syndrome with intellectual disability caused by mutations and haploinsufficiency of the nuclear receptor binding SET domain protein 1 (*NSD1*) gene, an HMTase at 5q35.2-35.3. It is characterized by excessive physical growth during the first few years of life. Since the first report in 1964, more than 500 affected cases have been reported; the incidence is approximately 1 in 10,000 to 14,000 births (data from Genetic Home Reference). However, the true incidence may be closer to 1 in 5,000, because of the poor diagnosis of this disease. Here we present the effect of one novel I2007F missense mutation in the SET domain on the conformational state and stability of the protein evaluated using protein modeling. MD simulation result points to a structural mechanism underlying the pathophysiology of the *NSD1* I2007F missense mutation in the SET domain.

---

\*Speaker

# InterEvDock: structural prediction of protein-protein interactions using evolutionary information

Jessica Andreani \* <sup>1</sup>, Jinchao Yu , Marek Vavrusa , Julien Rey , Pierre Tuffery <sup>2</sup>, Raphael Guerois <sup>3</sup>

<sup>1</sup> Institute for Integrative Biology of the Cell (I2BC) – CEA, CNRS : UMR9198 – France

<sup>2</sup> Molecules Thérapeutiques in-Silico (MTi) – INSERM, Université Paris Diderot – Université Denis Diderot 35 rue Hélène Brion, case 7113, 75205 Paris Cedex 13, France

<sup>3</sup> Institute for Integrative Biology of the Cell (I2BC) – CEA, CNRS : UMR9198 – Raphael GUEROIS Laboratoire de Biologie Structurale et Radiobiologie iBiTec-S (Institut de Biologie et de Technologie de Saclay) CEA 91191 Gif sur Yvette cedex, France

Protein-protein interactions are central to most biological processes. In the past few years, our team has contributed to the improvement of methods predicting protein complex structures by combining evolutionary information with more traditional approaches. Based on our analysis of interface structure coevolution [1, 2], we developed **InterEvDock, a server for protein-protein docking designed to integrate evolutionary information in the docking process** [3, 4]. We successfully used this pipeline in the CAPRI international prediction challenge [5] and in various biological applications.

In the laptop session, we will present the InterEvDock server and illustrate how it can be used in a typical case of protein-protein interaction prediction.

Note that this laptop session is connected to the talk by Raphael Guerois.

1. Faure G, Andreani J, Guerois R. *InterEvol database: exploring the structure and evolution of protein complex interfaces*. Nucleic Acids Res, 2012. **40** (Database issue): D847-56.
2. Andreani J, Faure G, Guerois R. *Versatility and invariance in the evolution of homologous heteromeric interfaces*. PLoS Comput Biol, 2012. **8**(8): p. e1002677.
3. Andreani J, Faure G, Guerois R. *InterEvScore: a novel coarse-grained interface scoring function using a multi-body statistical potential coupled to evolution*. Bioinformatics, 2013. **29**(14): p. 1742-9.
4. Yu J, Vavrusa M, Andreani J, Rey J, Tufféry P, Guerois R. *InterEvDock: a docking server to predict the structure of protein-protein interactions using evolutionary information*. Nucleic Acids Res, 2016. **44**(W1):W542-9.
5. Yu J, Andreani J, Ochsenbein F, Guerois R. *Lessons from (co-)evolution in the docking*

---

\*Speaker

*of proteins and peptides for CAPRI Rounds 28-35.* Proteins, 2016. **85**(3):378-390.

# COCO-tools: web applications for understanding protein interactions at a molecular level

Ida Autiero \* <sup>1</sup>, Zhen Cao <sup>1</sup>, Luigi Cavallo<sup>†</sup> <sup>1</sup>, Romina Oliva<sup>‡</sup> <sup>2</sup>

<sup>1</sup> Kaust Catalysis Center, King Abdullah University of Science and Technology (KAUST) – Thuwal 23955-6900, Saudi Arabia, Saudi Arabia

<sup>2</sup> Department of Sciences and Technologies, University “Parthenope” of Naples, Centro Direzionale Isola C4 80143, Naples, Italy – Italy

Characterization of the interaction interface of protein-protein complexes is a fundamental step for understanding protein interactions at a molecular level. In the many cases where experimental structures are not available, protein-protein docking becomes the method of choice for predicting the arrangement of a complex. However, reliably scoring protein-protein docking poses remains an unsolved problem and the screening of many docking models is thus usually required in the analysis step. We introduced novel tools for the analysis of protein-protein complex structures and for the scoring of docking poses. They rely on the use of inter-residue contacts and their visualization in inter-molecular contact maps and can straightforwardly be applied also to protein-nucleic acid complexes.

**COCOMAPS (bioCOMplexes CONTACT MAPS)**<sup>1</sup> is a web server to analyse and visualize the interface in biomolecular complexes. COCOMAPS combines the traditional analyses and 3D visualization of the complexes with the effectiveness of the contact map view. It can be applied to the analysis of both experimental and predicted 3D structures.

We used contact maps also as the basis for other tools for the analysis of conformational ensembles of protein complexes: docking models, NMR conformers and MD snapshots<sup>2,3,4</sup>. In particular, CONSRANK (CONSENSUS RANKing), also available as a web server<sup>5</sup>, is devoted to score docking models. It ranks models based on their ability to match the most conserved contacts in the ensemble they belong to, thus being a pure consensus approach<sup>2</sup>. Blind testing in the latest CAPRI (Critical Assessment of PRedicted Interactions) rounds showed CONSRANK to perform competitively with the state-of-the-art energy- or knowledge-based methods<sup>6,7</sup>. A modified version of CONSRANK includes a contact-based clustering of the models as a preliminary step of the scoring process<sup>8</sup>. We are now adding weights to CONSRANK to differentiate the inter-residue contacts based on the identity of the interacting amino acids. A dedicated CONSRANK module (consrank-nmr) can be applied to: i) analyse the interface in whole ensembles of NMR complexes, ii) select a single conformer as the best representative of the overall interface<sup>3</sup>. Finally, MDcons (Molecular Dynamics CONSENSUS)<sup>4</sup> is conceived for the analysis of MD trajectories. We use the conservation of inter-residue contacts at the interface to measure

---

\*Speaker

<sup>†</sup>Corresponding author: luigi.cavallo@kaust.edu.sa

<sup>‡</sup>Corresponding author: oliva@uniparthenope.it



the similarity between different snapshots. A "consensus contact map" is provided, where the conservation of the different contacts is drawn in a grey-scale, and the complex interface area along the simulation is monitored.

Details on the above tools and results of their application to selected test-cases will be presented.

Availability:

<https://www.molnac.unisa.it/BioTools/cocomaps/>

<https://www.molnac.unisa.it/BioTools/consrank/>

<https://www.molnac.unisa.it/BioTools/mdcons/index.php/MDcons/>

## References

Vangone A, Spinelli R, Scarano V, Cavallo L, Oliva R. *Bioinformatics* 2011;**27**:2915-6.

Oliva R, Vangone A, Cavallo L. *Proteins* 2013;**81**:1571-84.

Calvanese L, et al. *J Struct Biol* 2016;**194**:317-24.

Abdel-Azeim S, Chermak E, Vangone A, Oliva R, Cavallo L. *BMC Bioinformatics* 2014;**15** Suppl 5:S1.

Chermak E, et al. *Bioinformatics* 2015;**31**:1481-3.

Lensink MF, et al. *Proteins* 2016;**84** Suppl 1:323-48.

Lensink MF, Velankar S, Wodak SJ. *Proteins* 2017;**85**:359-377.

Chermak E, et al. *PLoS One* 2016;**11**:e0166460.

# Toward a crowded protein simulation with the Martini force field

Jonathan Barnoud \* <sup>1</sup>, Siwert-Jan Marrink <sup>1</sup>

<sup>1</sup> University of Groningen – Netherlands

Proteins are found in crowded environments that affect their features and their functions. Yet, molecular dynamics simulations are, most of the time, carried out at virtually infinite dilution with a single protein in water. Simulating proteins in the context of a crowded cytosol would give a more realistic view of their interactions with their environment.

The literature reports few example of such simulations of cytosol. The work by McGuffee and Elcock (PLOS Computational Biology, 2010) showed that such large-scale simulations were possible but the use of Brownian dynamics of rigid bodies make their model difficult to apply to the study of a specific protein in its environment. The more recent work by Yu and collaborators (eLife, 2016) describes an atomistic molecular dynamics simulation of cytosol. However, the time scales that can be reached this way are short.

Coarse-grained force-fields would be a solution to increase the time scale while allowing to focus on a specific protein. In this laptop session, I will demonstrate how the Martini force field cannot currently simulate a crowded protein environment and present how we try to remedy the issues.

---

\*Speaker

# Protein loop sampling using a large structural database and reinforcement learning

Amélie Barozet <sup>\*</sup> <sup>1,2,3</sup>, Kevin Molloy <sup>1</sup>, Marc Vaisset <sup>1</sup>, Thierry Siméon <sup>1</sup>,  
Juan Cortés <sup>\*</sup> <sup>†</sup> <sup>1</sup>

<sup>1</sup> Laboratoire d'analyse et d'architecture des systèmes [Toulouse] (LAAS) – Université Paul Sabatier (UPS) - Toulouse III, CNRS : UPR8001 – 7 Avenue du colonel Roche - 31077 Toulouse Cedex 4, France

<sup>2</sup> Sanofi RD – SANOFI Recherche – 13, quai Jules Guesde - BP 14 - 94403 - Vitry-sur-Seine, France

<sup>3</sup> Institut national Des Sciences Appliquées de Toulouse (INSA Toulouse) – Institut National des Sciences Appliquées [INSA] – 135, Avenue de Rangueil - 31400 Toulouse, France

Accurate structural modeling of protein loops is key to understanding the major roles they play, particularly in many protein-protein interactions. Their high flexibility poses considerable challenges to high-resolution experimental methods such as X-ray crystallography. Computational methods are being developed in order to circumvent this problem and make predictions about the structures and motions of protein loops. In particular, loop sampling methods aim at predicting a representative ensemble of loop conformations.

Our method is inspired from robotics: the protein backbone can be seen as an articulated linkage whose only degrees of freedom are the dihedral angles. The goal then becomes to assign values to each of the dihedral angles such that the two ends of the loop keep connected to the rest of the protein, in effect closing the loop, while avoiding collisions with itself and the protein.

Our method uses reinforcement learning on a database of protein fragments in order to efficiently build a set of possible loop conformations. More precisely, we use an all-atom model of the protein and decompose the loop region into  $k$  tripeptides. We built a database of tripeptides excised from the structures of more than 27000 non-redundant experimentally solved protein domains from the SCOP database, and indexed the tripeptides by their amino-acid sequence. Our algorithm starts by selecting one tripeptide position,  $s$ . The two halves of the loop around  $s$  are built iteratively by randomly drawing tripeptides from the database. The conformation of  $s$  is then solved using analytical inverse kinematics to close the loop. The side-chains are added afterwards using an existing method such as Scwrl 4.

Instead of uniformly drawing tripeptides from the database, the tripeptides are classified based on their structural features and reinforcement learning is used to determine the conformations that work best for each position, based on knowledge from previous iterations. Sampling unexplored regions is guaranteed by the reward function applied within the reinforcement learning approach.

Regarding time efficiency, first results show the good performance of the method, which is able to generate a hundred conformations of very long loops (27 residues) in less than 10 minutes.

---

\*Speaker

†Corresponding author: jcortes@laas.fr

# ATTRACT web-interface for macromolecular docking

Isaure Chauvot De Beauchene <sup>\*</sup> <sup>1</sup>, Sjoerd De Vries <sup>2</sup>, Christina Schindler <sup>3</sup>, Martin Zacharias<sup>†</sup> <sup>4</sup>

<sup>1</sup> Laboratoire Lorrain de recherche en Informatique et ses applications (LORIA) – CNRS : UMR7503, INRIA, Université de Lorraine – 615, rue du Jardin Botanique 54600 Villers-lès-Nancy, France

<sup>2</sup> Molecules therapeutiques in silico (MTI) – Inserm : UMRS973, Université Paris Diderot - Paris 7 – 35 rue Hélène Brion, Paris, France

<sup>3</sup> Technical University of Munich (TUM) – James Franck strasse 1, 85748 Garching, Germany

<sup>4</sup> Technical University of Munich (TUM) – James Franck strasse 1, 85784 Garching, Germany

Computational protein-protein docking methods aim to predict the three-dimensional structure of complexes from structures of its protein constituents. In the past decade, web interfaces and servers have eased their usage by the biological community. I will present the last developments of a web-interface for macromolecules docking with the ATTRACT program. After explaining the principle of ATTRACT program, I'll present the different options available and when to apply them, the performances of each docking type on benchmarks, and finally I'll make an interactive demo of the usage of the web-interface.

Docking with ATTRACT starts with placing the ligand at or randomly initial positions around the receptor. The 3 positional and 3 rotational degrees of freedom are then explored by a series of minimisations of the intermolecular energy, computed as the sum of atom-atom interaction energy. Computation can be made extremely fast by two features. First, a coarse grained representation groups atoms of each residue into 3 or 4 beads, and bead-bead energies are defined by a knowledge-based force-field [1, 2]. Second, a pre-calculated grid stores the interaction energy for each possible bead with the receptor, speeding up the calculation by a factor 30-100. The final poses are re-scored, using either knowledge-based or step potentials [3].

The coarse-grained representation accounts for flexibility of small amplitudes (e.g. side-chains conformations). Options are available to account for larger conformational changes between bound and unbound molecules, for both receptor and ligand [4]: (i) normal modes can be pre-computed and added as additional degrees of freedom, (ii) ensembles of pre-computed conformations can be docked in parallel, (iii) domains with variable respective positions due to flexible linkers can be docked as separate bodies with distance restraints accounting for the possible linker conformations, (iv) final refinement of poses can be performed with partial or full flexibility. Experimental data on the complex to be modelled can also be added as distance restraints.

ATTRACT program is free and open-source, and applicable to protein – protein/peptides/nucleic acids complexes. Special features are provided for peptides (pepATTRACT program [5]) to take into account their extrem flexibility. The ATTRACT web-interface provides guidance and feedback on the various options and parameters, and performs extensive validation of the user input

---

\*Speaker

†Corresponding author: martin.zacharias@mytum.de

[6]. It is available at [www.attract.ph.tum.de](http://www.attract.ph.tum.de)

1. Zacharias. *Proteins*. 2005;60(2):252-6
2. Setny, Zacharias. *Nucleic Acids Res.* 2011;39(21):9118-29
3. Sasse, de Vries, Schindler, Chauvot de Beauchêne, Zacharias. *PLoS One*. 2017;12(1):e0170625
4. Schindler, Chauvot de Beauchêne, de Vries, Zacharias. *Proteins*. 2017;85(3):391-398
5. Schindler, de Vries, Zacharias. *Structure*. 2015;23(8):1507-15
6. de Vries, Schindler, Chauvot de Beauchêne, Zacharias. *Biophys J*. 2015;108(3):462-5

**On the calculation of electrostatic energies  
in the PaLaCe coarse-grain model for  
proteins and protein complexes in aqueous  
solution: all-atom Poisson-Boltzmann  
energies are reproduced with a fast  
approach having simple analytical  
derivatives.**

Nicoletta Ceres <sup>\*† 1</sup>

<sup>1</sup> MMSB UMR 5086 – CNRS : UMR5086 – France

TBD

---

\*Speaker

†Corresponding author: [nicoletta.ceres@ibcp.fr](mailto:nicoletta.ceres@ibcp.fr)

# Flexible protein docking by systematic cross-docking of rigid candidates refined with a knowledge-based potential

Sergei Grudinin \* <sup>1</sup>, Guillaume Pages

<sup>1</sup> NANO-D (INRIA Grenoble Rhône-Alpes / LJK Laboratoire Jean Kuntzmann) – Institut polytechnique de Grenoble (Grenoble INP), Université Joseph Fourier - Grenoble I, CNRS : UMR5224, Laboratoire Jean Kuntzmann, INRIA – Inria GIANT DRT/LETI/DACLE Batiment 51C - Minatec Campus 17 rue des Martyrs 38054 Grenoble Cedex, France

Multiple attempts have been made to tackle protein flexibility upon association into a protein complex. Currently most of the methods assume that the changes in the protein conformations upon binding are small and thus the protein partners can be treated as rigid. During the last two collaborative CASP-CAPRI computational experiments we have explored a different strategy by systematically cross-docking all the available rigid templates produced by multiple structure prediction servers. More precisely, on the first step, we systematically dock all the available (typically 150) stage-2 predictions of the CASP servers. For each of these, we exhaustively sample the rigid configurational space in 4 degrees of freedom for cyclic symmetries and in 6 degrees of freedom for other complexes. For the sampling, we use the SAM symmetry assembler, which is built on top of the Hex docking engine. On the next step, we re-score the top-60 x 150 solutions using a first-order minimization, side-chain reconstruction and the knowledge-based KSENIA potential. As a result, in CASP 12 – CAPRI experiment our approach was the only one that could find at least one interface in 7 out of 11 targets.

---

\*Speaker

# Computational study of Protein-Protein Interaction (PPI) interface of bacterial Universal Stress Protein (UspC) and Chromosome Partitioning Protein (mukB)

Alok Jha <sup>\*† 1</sup>

<sup>1</sup> Assistant Professor (Asst. Prof.) – Department of Biotechnology and Life Sciences, Institute of Bio Medical Education Research Mangalayatan University, Beswan, Aligarh, UP, India

The universal stress protein appears to belong to all *E.coli* stress and starvation stimulons under the global control of gene expression. Recent bioinformatics data suggests about the regulation and possible roles of usp paralogues in *E.coli* as well. In *E.coli* the product of *mukB* gene has features of a myosin-like motor protein and probably has a role in chromosome condensation and/or movement. We performed sequence based prediction of PPIs within bacteria using BIPS-BIANA Interolog prediction server that is based on the known interactions of the orthologous genes of other organisms (interologs). Based on the hypothesis that universal stress proteins UspC and chromosome partitioning protein, mukB (*E.coli*) have interacting domains, also, these domains could be interacting domains in the iPfam or the 3DID databases. Further, we used the STRING server version 10.0 that includes a completely redesigned prediction pipeline for inferring protein-protein associations from co-expression data, an API interface for the *R* computing environment and thus statistical analysis for Universal stress proteins, UspC and chromosomal partitioning protein, mukB from *E.coli* were performed. We generated Phylogenetic tree for UspC and chromosome partitioning protein, mukB (*E.coli*) by using Clustal Omega. The phylogenetic tree generation method includes tree format (distance matrix) and clustering method (Neighbor Joining) with distance correlation for more divergent sequences. We used KFC2 server to identify PPI interfaces. The KFC model is a machine learning approach for predicting binding hot spots within protein-protein interactions. We analyzed Universal stress proteins, UspC and chromosome partitioning protein, mukB (*E. coli*) complexes for protein-protein interaction (PPI) interface by using PDBePISA server. The PPI analyses from PDBePISA server suggested the UspC-mukB complex has a multimeric state of 2 with formula AB, composition AB and the dissociation pattern is A+B. The free energy of dissociation corresponds to the free energy difference between dissociated and associated states. Positive values of  $\Delta G_{diss}$  indicate that an external driving force should be applied in order to dissociate the assembly, therefore, the assemblies with  $\Delta G_{diss} = 0$  are thermodynamically stable. Negative  $\Delta iG$  (solvation free energy gain upon formation of interface) corresponds to hydrophobic interfaces, or positive protein affinity. The interface in complex Formation Significance Score (CSS) assigned is zero. CSS ranges from 0 to 1 as interface relevance to complex formation increases. However, achieved CSS implies that the interface does not play any role in complex formation and seems to be a

---

\*Speaker

†Corresponding author: alok.jha@mangalayatan.edu.in



result of crystal packing only. The identification of various residues in PPIs is useful not only to provide insights into mechanism of protein-protein recognition but also to indicate the areas to be targeted with small molecules. The energetic signatures in the isolated proteins are retained in the bound forms that can help determining the binding orientation of proteins on complex formation. This approach can accurately predict and characterize the binding sites for protein-protein interactions as well as provide clusters of bound small molecules on the druggable regions of the predicted binding site. These bound small molecules, peptides or peptidomimetics can be chemically combined to create candidate drugs.

# Using machine learning to predict protein-ligand interactions

Maria Kadukova \* <sup>1,2,3</sup>, Sergei Grudinin<sup>†</sup> <sup>4</sup>

<sup>1</sup> Moscow Institute of Physics and Technology (MIPT) – 141700, 9, Institutskii per., Dolgoprudny, Moscow Region, Russia, Russia

<sup>2</sup> NANO-D (INRIA Grenoble Rhône-Alpes / LJK Laboratoire Jean Kuntzmann) – Institut polytechnique de Grenoble (Grenoble INP), CNRS : UMR5224, Laboratoire Jean Kuntzmann, INRIA – Inria GIANT DRT/LETI/DACLE Batiment 51C - Minatec Campus 17 rue des Martyrs 38054 Grenoble Cedex, France

<sup>3</sup> Université Grenoble Alpes (UGA) – Université Grenoble Alpes – France

<sup>4</sup> NANO-D (INRIA Grenoble Rhône-Alpes / LJK Laboratoire Jean Kuntzmann) – Institut polytechnique de Grenoble (Grenoble INP), Université Joseph Fourier - Grenoble I, CNRS : UMR5224, Laboratoire Jean Kuntzmann, INRIA – Inria GIANT DRT/LETI/DACLE Batiment 51C - Minatec Campus 17 rue des Martyrs 38054 Grenoble Cedex, France

New methods for accurate and computationally simple description of protein-ligand interactions are still demanded by industry and academia to elucidate the activity of biological targets and find their new binders modulating functions of the living systems. **Here we present a machine learning approach to first parametrize small molecules[1] and second to train a free-shape distance-dependent protein-ligand potential called Convex-PL.** Unlike knowledge-based methods based on Boltzmann statistics, we do not impose any functional form of the potential. Instead, we use an optimization approach, accepting that the target binding energy value is decomposed into a polynomial basis with unknown expansion coefficients. These are then deduced from the structural data collected from protein-ligand complexes using a convex formulation of the optimization problem. We have already implemented this principle for successful predictions of protein-protein binding[2]. The training set consists of the complexes taken from the PDBBind database. We generate false poses with constant RMSD rigid-body deformations of the ligands inside the binding pockets. This allows the obtained potential to be unbiased towards other molecular docking methods, which are often used for training decoys generation. Convex-PL performed successfully in the CSAR 2013-2014 and D3R 2015-2016 competitions[3,4]. For a more general validation, we assessed it using data from the CASF 2013 study[5], which includes the docking, scoring, ranking, and screening tests. Our docking and ranking test results outperform the other 20 methods previously assessed in CASF 2013. Also, Convex-PL performs better than average in the scoring test and produces rather unsatisfactory results in the screening test. The parametrization is available at <https://team.inria.fr/nano-d/software/knodle/>. The protein-ligand potential is available at <https://team.inria.fr/nano-d/software/Convex-PL>.

## References

M. Kadukova & S. Grudinin. Knodle: A Support Vector Machines-Based Automatic Perception

---

\*Speaker

<sup>†</sup>Corresponding author: [sergei.grudinin@inria.fr](mailto:sergei.grudinin@inria.fr)

of Organic Molecules from 3D Coordinates. *J. Chem. Inf. Model.*, 2016, 56 (8), pp.1410-1419.

Neveu, E.; Ritchie, D.W.; Popov, P.; Grudinin, S. "PEPSI-Dock: a detailed data-driven protein–protein interaction potential accelerated by polar Fourier correlation." *Bioinformatics*, 2016

Grudinin, S.; Popov, P.; Neveu, E.; Cheremovskiy, G. "Predicting Binding Poses and Affinities in the CSAR 2013–2014 Docking Exercises Using the Knowledge-Based Convex-PL Potential." *J. Chem. Inf. Model.*, 2015

Grudinin, S.; Kadukova, M.; Eisenbarth, A.; Marillet, S.; Cazals, F. "Predicting binding poses and affinities for protein-ligand complexes in the 2015 D3R Grand Challenge using a physical model with a statistical parameter estimation." *J. Comput. Aided Mol. Des.*, 2016

Li, Y.; Han, L.; Liu, Zh.; Wang, R. "Comparative Assessment of Scoring Functions on an Updated Benchmark: 2. Evaluation Methods and General Results." *J. Chem. Inf. Model.*, 2014

# Use of cross-docking simulations for identification of protein-protein interactions sites: the case of proteins with multiple binding sites

Nathalie Lagarde <sup>\*† 1</sup>, Lydie Vamparys <sup>1</sup>, Benoist Laurent <sup>1</sup>, Alessandra Carbone <sup>2</sup>, Sophie Sacquin-Mora <sup>‡ 1</sup>

<sup>1</sup> Laboratoire de biochimie théorique (LBT) – CNRS : UPR9080 – 13 Rue Pierre et Marie Curie 75005 PARIS, France

<sup>2</sup> Laboratoire de Biologie Computationnelle et Quantitative – Université Pierre et Marie Curie (UPMC) - Paris VI – France

## Introduction

Understanding protein-protein interactions (PPI) is essential to decipher the mechanism of numerous biological functions. Some *in silico* methods can be used to investigate PPI. In particular, cross-docking simulations of large datasets of proteins can be used to predict interface residues [1-3]. Many proteins present multiple binding sites on their surface, reflecting their ability to form complexes with different partners and are particularly challenging for PPI studies. We discuss here the ability of the cross-docking method to detect multiple binding sites on protein surfaces.

## Methods

358 proteins extracted from 138 unique PDB structures were used for this study. The MAXDo algorithm [1] was used with a rigid-body docking approach and a reduced protein representation, a coarse-grain protein model developed by Zacharias [4]. Binding site predictions resulting from evolutionary sequence analysis produced with JET [5] were used to restrict the initial search space. For each surface residue, its Protein Interface Propensity (PIP) was computed and used to predict binding sites on the protein surface.

## Results

For a large number of proteins, alternative interfaces different from the reference experimental interfaces were predicted. However, about 70 % of these interfaces were not false positives but correspond to interfaces with other partners (other chain of the same PDB not included in the database, nucleic acid molecule or homo-/hetero-dimerization interface). We compared the use of two different scoring schemes accounting for multiple binding sites, for evaluating the binding sites prediction. The first score was obtained by comparing the predicted interface

---

\*Speaker

†Corresponding author: lagarde@ibpc.fr

‡Corresponding author: sacquin@ibpc.fr

with one single global reference experimental interface generated by concatenating all the existing experimental interfaces. In the second score, the predicted interface was compared to each experimental interface separately, and only the interface associated with the best performance was kept.

## Conclusion

Using cross-docking simulations on a large dataset of proteins, accurate binding sites predictions could be realized, including proteins which present multiple binding sites.

## References

Sophie Sacquin-Mora, Alessandra Carbone, and Richard Lavery. Identification of Protein Interaction Partners and Protein-Protein Interaction Sites. *Journal of Molecular Biology*, (382):1276-1289, 2008.

Anne Lopes, Sophie Sacquin-Mora, Viktoriya Dimitrova, Elodie Laine, Yann Ponty and Alessandra Carbone. Protein-Protein Interactions in Crowded Environment: An Analysis via Cross-Docking Simulations and Evolutionary Information. *PLoS Computational Biology*, (9/12):e1003369,1-18, 2013.

Lydie Vamparys, Benoist Laurent, Alessandra Carbone and Sophie Sacquin-Mora. Great interactions: How binding incorrect partners can teach us about protein recognition and function. *Proteins*, (84/10):1408-1421, 2016.

Martin Zacharias. Protein-protein docking with a reduced protein model accounting for side-chain flexibility. *Protein Science*, (12):1271-1282, 2003.

Stefan Engelen, Ladislav A Trojan, Sophie Sacquin-Mora, Richard Lavery and Alessandra Carbone. Joint Evolutionary Trees: A Large-Scale Method To Predict Protein Interfaces Based on Sequence Sampling. *PLoS Computational Biology*, (5/1):1-17, 2009.

# The KAMI Cell Signaling Modeler

Sébastien Légaré <sup>\*</sup> <sup>1</sup>, Russel Harmer <sup>†</sup> <sup>1</sup>, Yves-Stan Le Cornec <sup>1</sup>, Ievgeniia Oshurko <sup>1</sup>, Walter Fontana <sup>2</sup>

<sup>1</sup> Laboratoire de l'Informatique du Parallélisme (LIP) – École Normale Supérieure - Lyon – France

<sup>2</sup> Harvard Medical School (HMS) – United States

Cell signaling involves many crosstalking regulatory pathways that form an intricate network of protein interactions. Modeling these interaction networks would help understand numerous illnesses and find new therapies. However, while the amount experimental data on cell signaling is exploding, information on a given pathway is often spread across several non-standardized pieces of evidence. This makes assembling knowledge into comprehensive models an increasingly daunting task. **The KAMI (Knowledge Aggregator and Model Instantiator) software creates a framework for systems biologists to incrementally and intuitively accumulate information from different sources into a complete knowledge representation of signaling networks.** A pivotal aspect of KAMI is the so-called nugget, a standardized unit of knowledge about a single protein interaction. Using graph rewriting algorithms, the software then allows combining any number of nuggets into full models. Furthermore, KAMI is fully integrated with rule-based simulator KaSim. Users can hence also simulate signaling network variants to reveal key interactions in pathologies.

---

\*Speaker

†Corresponding author: russell.harmer@ens-lyon.fr

# Multiprotein complexes of eukaryotic cell membranes: integrative bioinformatic modeling of the interactome

G Khazen , A Gyulkhandanian , T Issa , R C Maroun <sup>\*† 1</sup>

<sup>1</sup> UMR-S INSERM U1204/UEVE (Structure et activité de biomolécules normales et pathologiques) – Institut National de la Santé et de la Recherche Médicale - INSERM, Université d'Evry-Val d'Essonne – France

Compared to the space of possible membrane protein complexes, the set of corresponding 3D structures is very scarce given the experimental difficulties found in obtaining them. As a consequence, protein-protein complexes have been the target of algorithms destined to their prediction from protein-protein interaction (PPI) databases. To the best of our knowledge, all those efforts, except one that deals with membrane proteins (Qi et al, 2009), deal with cytoplasmic soluble globular proteins. Yet the work of Qi et al deals only with receptors.

Our approach to the prediction of integral-to-membrane eukaryotic complexes is based on bioinformatic methods and uses the information from 3D structural data of the few experimental structures of membrane complexes. The originality of our approach is based on extracting those amino acid residues at the interface of experimental complexes and using this 3D information in the form of a linear 1D sequence, called interaction motif, in order to extract other membrane proteins susceptible of interacting between them.

The new predicted protein couples we have obtained have now been grouped in a relational database that allows for different types of queries. The database is presented in a website that will be the subject of the laptop session.

The next step will be using machine-learning methods in order to optimize the prediction of new protein complexes. For that purpose we count on taking into consideration the properties of our linear and spatial binding motifs, such as the amino acid composition, type and frequency. The algorithms thus trained, along with the integration of biochemical and biophysical data, shall allow us to propose other membrane proteins capable of interacting between them in a binary, physical mode.

---

\*Speaker

†Corresponding author: charbel.maroun@inserm.fr

Several of the proposed interactions will be tested against experimental databases and in our laboratory thanks to a new technique that uses the microtubules as platforms for probing PPIs (Boca et al, 2015).



# Integrative Modelling goes HPC/HTC

Adrien Melquiond \* <sup>1</sup>, Mikael Trellet <sup>1</sup>, Jörg Schaarschmidt <sup>1</sup>, Gydo Zundert <sup>2</sup>, Alexandre Bonvin<sup>† 1</sup>

<sup>1</sup> Computational Structural Biology, Bijvoet Center of Biomolecular Research, Faculty of Science (UU)  
– Utrecht University, Padualaan 8, 3584CH Utrecht, Netherlands  
<sup>2</sup> Schrödinger Inc., NY Stanford University, CA – United States

Within BioExcel, the European Center of Excellence (CoE) for Computational Biomolecular research, we developed new web services for the modelisation of biomolecular complexes powered by EGI HTC resources distributed worldwide.

Pioneer in the field of integrative modelling, **HADDOCK** ([haddock.science.uu.nl](http://haddock.science.uu.nl)) is an information-driven modelling service for the prediction of the three-dimensional structure of biomolecular assemblies. More than 8,500 users worldwide make use of HADDOCK's user-friendly web interface. As of March 2017, more than 150,000 user submissions have been processed by the server, which translates into more than 8 million jobs a year. The HADDOCK support center is powered by a dedicated BioExcel forum (<http://ask.bioexcel.eu>).

**DisVis** (<http://milou.science.uu.nl/services/DISVIS>) allows you to visualize and quantify the information content of distance restraints between macromolecular complexes. It performs a full and systematic 6-dimensional search of the three translational and rotational degrees of freedom to determine the number of complexes consistent with the restraints. In addition, it outputs the percentage of restraints being violated and a density that represents the center-of-mass position of the scanning chain corresponding to the highest number of consistent restraints at every position in space.

**PowerFit** (<http://milou.science.uu.nl/services/POWERFIT>) automatically fits high-resolution atomic structures in cryo-EM densities. To this end it performs a full-exhaustive 6-dimensional cross-correlation search between the atomic structure and the density. It takes as input an atomic structure in PDB-format and a cryo-EM density with its resolution; and outputs positions and rotations of the atomic structure corresponding to high correlation values. PowerFit uses the local cross-correlation function as its base score, enhanced by a Laplace pre-filter and/or core-weighting to minimize overlapping densities from neighboring subunits.

---

\*Speaker

<sup>†</sup>Corresponding author: [a.m.j.j.bonvin@uu.nl](mailto:a.m.j.j.bonvin@uu.nl)

# BIS2Analyzer: a server for coevolution analysis of conserved protein families

Francesco Oteri \* <sup>1</sup>, Francesca Nadalin<sup>†</sup>, Raphael Champeimont ,  
Alessandra Carbone<sup>‡</sup>

<sup>1</sup> UMR7238 (Laboratoire de Biologie Computationnelle et Quantitative) – Université Pierre et Marie Curie [UPMC] - Paris VI – France

Along protein sequences, coevolution analysis identifies residue pairs demonstrating either a specific coadaptation, where changes in one of the residues are compensated by changes in the other during evolution, or a less

specific external force that affects the evolutionary rates of both residues in a similar magnitude. In both cases,

independently of the underlying cause, co-evolutionary signatures within or between proteins serve as markers

of physical interactions and/or functional relationships.

Depending on the type of protein under study, the set of available homologous sequences may greatly differ

in size and amino acid variability. BIS2Analyzer, openly accessible at <http://www.lcqb.upmc.fr/BIS2Analyzer/>, is a

webserver providing the online analysis of coevolving aminoacid pairs in protein alignments, especially designed for

vertebrate and viral protein families, which typically display a few number of highly similar sequences. It is based on

BIS2, a reimplemented version of the coevolution analysis tool Blocks in Sequences (BIS) but significantly faster.

BIS2Analyzer provides a rich and interactive graphical interface to ease biological interpretation of the results.

---

\*Speaker

<sup>†</sup>Corresponding author: francesca.nadalin@upmc.fr

<sup>‡</sup>Corresponding author: Alessandra.Carbone@lip6.fr

# Exhaustive Analysis of Point-Group Symmetries in Protein Assemblies

Guillaume Pagès<sup>\*</sup> <sup>1</sup>, Sergei Grudinin<sup>†</sup> <sup>2</sup>

<sup>1</sup> NANO-D (INRIA Grenoble Rhône-Alpes / LJK Laboratoire Jean Kuntzmann) – INRIA, Laboratoire Jean Kuntzmann, Université Grenoble Alpes – Inria GIANT DRT/LETI/DACLE Batiment 51C - Minatec Campus 17 rue des Martyrs 38054 Grenoble Cedex, France

<sup>2</sup> NANO-D (INRIA Grenoble Rhône-Alpes / LJK Laboratoire Jean Kuntzmann) – Institut polytechnique de Grenoble (Grenoble INP), Université Joseph Fourier - Grenoble I, CNRS : UMR5224, Laboratoire Jean Kuntzmann, INRIA – Inria GIANT DRT/LETI/DACLE Batiment 51C - Minatec Campus 17 rue des Martyrs 38054 Grenoble Cedex, France

Many protein complexes in the Protein Data Bank (PDB) are symmetric homo-oligomers. Indeed, it appears that large symmetrical protein structures have evolved in many organisms because they carry specific morphological and functional advantages compared to small individual protein molecules (Levy et al., 2006, Levy et al., 2008). There is therefore considerable interest in studying and modelling the structures of these large bio-molecular complexes. Recently we have proposed a novel ab-initio protein docking algorithm for protein complexes with arbitrary point group symmetry (Ritchie & Grudinin, 2016). Here we report on a very fast method for the analysis of symmetric protein assemblies.

Analytical methods to perform the best structural superposition between different structures are very well known, however this alignment cannot be, generally, used to determine symmetries in protein assemblies. This is because of two underlying problems. First of all, to determine the best group of transformation for symmetrical assemblies (cyclic, dihedral, tetrahedral, octahedral and icosahedral assemblies), in addition to the correspondence between the atoms in different subunits, a correspondence between subunits is also required. Brute-force approaches will be prohibitively expensive for high-order symmetries in this case. Second, given the correspondence between the atoms and the subunits, a very fast method to estimate the quality of symmetry, or, more generally, to compute the symmetry-aware penalty function is needed.

We addressed these two challenges and developed computational tools that first find the correspondence between the subunits for all types of point-groups symmetries, and second that analytically compute the group of transformation that minimizes the symmetry-aware root-mean-square deviation (RMSD) over these transformations. In other words, our tools automatically determine the axes of symmetry for a given group by analytically minimizing the RMSD-based penalty function. The workflow of our method is split into two parts. The timing of preprocessing part, which aligns the sequences and computes the geometric matrices, grows with the number of subunits in the assembly and the size of these subunits. The subsequent symmetry analysis takes only constant time, which is negligible compared to the preprocessing timing.

---

<sup>\*</sup>Speaker

<sup>†</sup>Corresponding author: sergei.grudinin@inria.fr

Thanks to these tools, we performed an automatic exhaustive analysis of all the symmetric protein structures in PDB. Based on this analysis, we can draw some conclusions. First of all, for high-order symmetries, structural impurities do not depend on the order of the axis. For example, the average RMSD over 2-fold axes for tetrahedral structures is 0.47Å, and 0.48Å over 3 fold axes. Second, surprisingly enough, we found icosahedral complexes to be better structurally organized compared to octahedral and tetrahedral assemblies. These have an average angular error over non-crystallographic axes of 0.3° against 0.5° for octahedral and 0.9° for tetrahedral assemblies, despite the obvious entropic penalties. This allows us to propose an interaction model that makes large symmetric ensembles to be energetically favorable.

#### References:

- Levy, E.D. et al. (2006). PLoS Comput. Biol. 2, e155.
- Levy, E.D. et al. (2008). Nature, 453, 1262–1266.
- Ritchie, D. W. & Grudinin, S (2016). J. Appl. Cryst. (2016). 49, 1-10.

# Identification of a muscle-specific MuRF1-E2 network

Cécile Polge <sup>\*† 1</sup>, Stéphanie Cabantous <sup>2</sup>, Christiane Deval <sup>1</sup>, Agnès Claustre <sup>1</sup>, Catherine Bouchenot <sup>2</sup>, Lydie Combaret <sup>1</sup>, Daniel Taillandier <sup>1</sup>

<sup>1</sup> Unité de Nutrition Humaine (UNH) – Université d’Auvergne - Clermont-Ferrand I, Institut national de la recherche agronomique (INRA) : UR1019, Clermont Université – Unité de Nutrition Humaine UMR 1019 - Inra/Université Clermont 1 Centre de Recherches Inra de Clermont-Fd/Theix 63122 St Genès Champanelle FRANCE Tel. : +33(0)4 73 62 42 70 Fax : + 33(0)4 73 62 47 55  
unh@clermont.inra.fr, France

<sup>2</sup> Centre de Recherche en Cancérologie de Toulouse (CRCT) – Université Paul Sabatier (UPS) - Toulouse III, Hôpital Purpan, CHU Toulouse, Inserm : U1037 – Place du Docteur Baylac - BP 3028 - 31024 Toulouse Cedex 3, France

As a result of an imbalance between protein synthesis and proteolysis, important protein losses occur during various physiological and pathological situations. Skeletal muscle mass is largely controlled by the ubiquitin proteasome system and thus by ubiquitinating enzymes. This proteolytic system is highly regulated and controls precisely the activity of numerous cellular proteins that are first tagged by a polyubiquitin chain on an internal lysine. Ubiquitination is based on the sequential action of the E1-activating enzyme, an E2 ubiquitin-conjugating enzyme (38 members) and an E3 ligase (> 600 members). The E1 enzyme activates the ubiquitin that it is transferred to an E2. E3 ligases recruit and bind specific substrates.

MuRF1, a muscle-specific E3, is the only E3 ligase known to target contractile proteins ( $\alpha$ -actin, myosins) for degradation during catabolic situations. MuRF1 is therefore a putative target for preventing muscle wasting. However, RING E3s like MuRF1 highly depend on ubiquitin conjugating enzymes E2s for ubiquitin chain formation.

Our main objective was then to identify the E2s working in concert with MuRF1 to target myofibrillar proteins in atrophying skeletal muscles. We focused on 12 E2 enzymes that are expressed in the skeletal muscle or in atrophying skeletal muscles. Classical biochemical approach such as pull-down did not allow identifying any MuRF1 interacting partner among these E2s, suggesting that E2-MuRF1 interactions are weak and/or transient, consistent with previous studies reporting that E2-E3 interactions range from moderate to weak. We demonstrated that only sensitive and complementary interactomic approaches (Surface Plasmon Resonance, Yeast three-Hybrid and split-GFP) allowed the identification of MuRF1 E2 partners. We identified five E2 enzymes that physically interacted with MuRF1. We showed that these E2s functionally interacted with MuRF1 since, in contrast with the non-interacting E2D2, their co-expression in HEK293T cells with MuRF1 led to increased MuRF1 substrate degradation. In conclusion, we report the first MuRF1-E2s network. This may prove valuable for deciphering the precise mechanisms involved in the muscle-atrophying program and for proposing new therapeutically approaches.

---

\*Speaker

†Corresponding author: cecile.polge@inra.fr

# JET2 Viewer: a database of predicted multiple, possibly overlapping, protein-protein interaction sites for PDB structures.

Hugues Ripoche \* <sup>1</sup>, Elodie Laine \*

<sup>1</sup>, Nicoletta Ceres <sup>2</sup>, Alessandra Carbone<sup>†</sup> <sup>1,3</sup>

<sup>1</sup> Sorbonne Universités, UPMC University Paris 06, CNRS, IBPS, UMR 7238, Laboratoire de Biologie Computationnelle et Quantitative (LCQB), 75005 Paris, France – Université Pierre et Marie Curie (UPMC) - Paris VI, CNRS : UMR7238 – France

<sup>2</sup> CNRS UMR 5086/University Lyon I, Institut de Biologie et Chimie des Proteines, 69367 Lyon, France – CNRS : UMR5086, Université de Lyon – France

<sup>3</sup> Institut Universitaire de France, 75005 Paris, France – Institut universitaire de France – France

The database JET2 Viewer, openly accessible at <http://www.jet2viewer.upmc.fr/>, reports putative protein binding sites for all three-dimensional (3D) structures available in the Protein Data Bank (PDB). This knowledge base was generated by applying the computational method JET2 at large-scale on more than 20 000 chains. JET2 strategy yields very precise predictions of interacting surfaces and unravels their evolutionary process and complexity. JET2 Viewer provides an online intelligent display, including interactive 3D visualization of the binding sites mapped onto PDB structures and suitable files recording JET2 analyses. Predictions were evaluated on more than 15 000 experimentally characterized protein interfaces. This is, to our knowledge, the largest evaluation of a protein binding site prediction method. The overall performance of JET2 on all interfaces are: Sen = 52.52, PPV = 51.24, Spe = 80.05, Acc = 75.89. The data can be used to foster new strategies for protein-protein interactions modulation and interaction surface redesign.

---

\*Speaker

<sup>†</sup>Corresponding author: [alessandra.carbone@lip6.fr](mailto:alessandra.carbone@lip6.fr)

# Design of cyclic peptide inhibitors of caspase-3 using metadynamics simulations

Jaysen Sawmynaden <sup>\*</sup> <sup>1</sup>, Maud Jusot <sup>1,2</sup>, Jacques Chomilier <sup>1</sup>,  
Chahrazade El Amri <sup>3</sup>, Fabio Pietrucci <sup>1</sup>, Dirk Stratmann<sup>†</sup> <sup>1</sup>

<sup>1</sup> Institut de minéralogie, de physique des matériaux et de cosmochimie (IMPMC) – Institut de recherche pour le développement [IRD] : UR206, Université Pierre et Marie Curie (UPMC) - Paris VI, CNRS : UMR7590, Muséum National d’Histoire Naturelle (MNHN) – Tour 23 - Barre 22-23 - 4e étage - BC 115 4 place Jussieu 75252 PARIS, France

<sup>2</sup> Laboratoire d’analyse et d’architecture des systèmes [Toulouse] (LAAS) – Institut National Polytechnique de Toulouse - INPT, Université Paul Sabatier (UPS) - Toulouse III, CNRS : UPR8001, Institut National des Sciences Appliquées [INSA] - Toulouse – 7 Av du colonel Roche 31077 TOULOUSE CEDEX 4, France

<sup>3</sup> Adaptation biologique et Vieillesse, UMR 8256, IBPS, Sorbonne Universités (B2A-ERL INSERM U1164) – Sorbonnes Universités-Univ Paris 06 – 7 Quai Saint Bernard Case Courrier 256 75252 Paris Cedex 05, France

Present in all eukaryotes, caspases are a family of enzymes involved in cellular death. Exprimed during the entire life, their failure lead to cancer and Alzheimer’s disease. Most inhibitors targeting the active site are not specific to one caspase, as the active site is conserved among the family, Whereas the interface of functional caspase dimer is highly specific to one member of this family [1].

Targeting the interface, the catalytic activity of caspases can be blocked by preventing its dimerization (as shown for caspase-9 [2]). **We present here an *in silico* approach based on the use of cyclic peptides as competitive inhibitors to block the dimerization of caspase-3.**

We start with a fragment of the caspase-3’s interface, two anti-parallel beta strands that we cyclize manually N-ter to C-ter. Only one of the beta strands is in contact with the caspase-3 monomer, whereas the remaining amino acids do not play a role in the dimerization but have an effect on peptide’s structure (mainly by hydrogen bonds).

Several modifications on the non-interacting amino acids are tested. In particular, effect of an increased solubility is examined by swapping the hydrophobic amino acids to hydrophilic ones. In addition, different constructions for the turns connecting the two beta strands are tested like D-amino acid and N-methylated residues, aiming to stabilize the conformation of the fragment.

The free-form conformations of our designed cyclic peptides are first explored by temperature replica exchange MD simulations in implicit and explicit solvent. Then, the ability of the most stable peptides to bind to caspase-3 is assessed by studying the free energy of the

---

<sup>\*</sup>Speaker

<sup>†</sup>Corresponding author: dirk.stratmann@impmc.upmc.fr

protein-peptide binding using bias exchange metadynamics simulations [3]. With these two MD simulation techniques, we will be able to evaluate and select the best candidates before an *in vitro* experimentation.

(1) MacKenzie, S. H., & Clark, A. C. (2012). Death by caspase dimerization. In *Protein Dimerization and Oligomerization in Biology* (pp. 55-73). Springer New York.

(2) Huber, K. L., Ghosh, S., & Hardy, J. A. (2012). Inhibition of caspase-9 by stabilized peptides targeting the dimerization interface. *Peptide Science*, 98(5), 451-465.

(3) Pietrucci, F., Marinelli, F., Carloni, P., & Laio, A. (2009). Substrate binding mechanism of HIV-1 protease from explicit-solvent atomistic simulations. *Journal of the American Chemical Society*, 131(33), 11811-11818.



# Conformational landscape of cyclic RGD peptides explored by REMD simulations

Maud Jusot <sup>1</sup>, Matthias Lerbinger <sup>1</sup>, Eric Ngo <sup>1</sup>, Théo Torcq <sup>1</sup>, Juan Cortes <sup>2</sup>, Jacques Chomilier <sup>1</sup>, Dirk Stratmann <sup>\*† 1</sup>

<sup>1</sup> Institut de minéralogie, de physique des matériaux et de cosmochimie (IMPMC) – Institut de recherche pour le développement [IRD] : UR206, Université Pierre et Marie Curie (UPMC) - Paris VI, CNRS : UMR7590, Muséum National d'Histoire Naturelle (MNHN) – Tour 23 - Barre 22-23 - 4e étage - BC 115 4 place Jussieu 75252 PARIS, France

<sup>2</sup> LAAS-CNRS (LAAS-CNRS) – CNRS : UPR8001 – France

Cyclic peptides are an important tool to inhibit protein-protein interactions. They present important advantages over linear acyclic peptides: improved stability and bio-availability, and higher affinity for the target. One extensively studied example is the class of cyclic Arg-Gly-Asp (RGD) pentapeptides binding to integrins. Optimizing the binding affinity of the c(RGDxx) sequence experimentally with N-methylation and D-amino acids, the group of Horst Kessler developed an anti-cancer drug, c(RGDf(NMe)Val) named "cilengitide", see [1] for a recent review. The group of V.A. Voelz explored the conformational landscape with microseconds REMD simulations of 18 different c(RGDxx) designs of the Kessler group [2]. We repeated and extended this study by using an explicit solvent representation in addition to the implicit solvent used in [2]. While for most designs the results were similar for the two solvent representations, one design showed rather important differences. More recently the Kessler group improved the specificity of the c(RGDfV) motif by applying the N-methylation to two of the five residues [3]. We applied our simulation protocol to the three designs for which the Kessler group did an NMR study. Our set of microseconds REMD simulations of cyclic RGD peptides in implicit and explicit solvent will allow us to validate faster algorithms for the conformational exploration of cyclic peptides, like the robotics based algorithms that we are currently developing.

References:

- (1) T. G. Kapp *et al.*, "A Comprehensive Evaluation of the Activity and Selectivity Profile of Ligands for RGD-binding Integrins," *Sci. Rep.*, vol. 7, p. 39805, Jan. 2017.
- (2) A. E. Wakefield, W. M. Wuest, and V. A. Voelz, "Molecular Simulation of Conformational Pre-Organization in Cyclic RGD Peptides," *J. Chem. Inf. Model.*, vol. 55, no. 4, pp. 806–813, Apr. 2015.
- (3) C. Mas-Moruno *et al.*, "Increasing  $\alpha\beta3$  Selectivity of the Anti-Angiogenic Drug Cilengitide by N-Methylation," *Angew. Chem. Int. Ed.*, vol. 50, no. 40, pp. 9496–9500, Sep. 2011.

---

\*Speaker

†Corresponding author: dirk.stratmann@impmc.upmc.fr

# PRODIGY: a web-server for the prediction of binding affinity in biomolecular complexes

Anna Vangone <sup>\*† 1</sup>, Li Xue <sup>1</sup>, Joerg Shaarschmidt <sup>1</sup>, Mikael Trellet <sup>1</sup>,  
Alexandre Bonvin <sup>1</sup>

<sup>1</sup> Utrecht University - UU (NETHERLANDS) (UU) – Netherlands

Almost all critical functions in the cell rely on specific interactions between biomolecules, such as proteins, nucleic acids and small ligands. Understanding these interactions is therefore a crucial step in the investigation of biological systems and in drug design. Despite all efforts that have been devoted to unravel principles of biomolecular interactions in the past decades, we still lack a thorough understanding of the energetics of proteins association.

Recently, we introduced a simple but robust descriptor of binding affinity based only on structural properties of the protein-protein complexes (1). Using the protein-protein binding affinity benchmark of Kastritis et al. (2) we showed that the number of interfacial contacts at the interface of a protein-protein complex correlates with the experimental binding affinity. This information, combined properties of the non-interacting surface which have been shown to influence binding affinity, (3) has led to one of the best performing predictor so far reported on such a large and heterogeneous set of data (Pearson's Correlation  $r = 0.73$ ,  $p$ -value  $< 0.0001$ ; RMSE = 1.89 kcal mol<sup>-1</sup>), which furthermore is quite robust to conformational changes occurring upon binding.

Despite the importance of the topic, there are surprisingly only limited online tools for a fast and easy fast prediction of binding affinity. For this reason, based on our contact-based binding affinity predictor, we developed PRODIGY (PROtein binDing enerGY predictor), a webserver to predict the affinity of a protein-protein complex from its three-dimensional structure. PRODIGY is freely available at: <http://milou.science.uu.nl/services/PRODIGY>, (4) and will be soon available for prediction of binding affinity for protein-small ligands as well.

## References:

- (1) Vangone A and Bonvin AMJJ. Contact-based prediction of binding affinity in protein-protein complexes. *eLife*, 4, e07454 (2015).
- (2) Kastritis PL, Moal IH, Hwang H, Weng Z, Bates PA, Bonvin AMJJ, Janin J. A structure-based benchmark for protein-protein binding affinity. *Protein Science*, 20:482-491 (2011).
- (3) Kastritis PL, Rodrigues JPGLM, Folkers GE, Boelens R, Bonvin AMJJ. *Journal of Molecular Biology*, 426:2632-2652 (2014).
- (4) Xue L, Rodrigues JPGLM, Kastritis PL, Bonvin AMJJ, Vangone A. PRODIGY: a web-

---

\*Speaker

†Corresponding author: a.vangone@gmail.com

server for the prediction of binding affinity for protein-protein complexes. *Bioinformatics*, doi: 10.1093/bioinformatics/btw514 (2016).

# A comparative study of the mutational robustness of an ordered plant factor and a disordered viral protein involved in key potyvirus biological functions

Jocelyne Walter \* <sup>1</sup>, Justine Charon <sup>2</sup>, Michon Thierry <sup>1</sup>

<sup>1</sup> INRA Fruit Biology and Pathology (INRA BFP 1332) – Institut national de la recherche agronomique (INRA) : UMR1332 – 71 avenue Edouard Bourlaux 33883 Villenave d’Ornon, France

<sup>2</sup> Institut Européen de Chimie et Biologie (IECB) – CNRS : UMR5320, Inserm : U1212 – 2 Rue Robert Escarpit 33607 Pessac, France

The genome of RNA viruses codes for multifunctional and multi-interactive proteins containing an important proportion of intrinsically disordered regions (IDRs) (1). The amino acid sequence variability within viral IDRs is often observed, suggesting a relative permissiveness of these regions to mutations. Such a mutational robustness is likely to be associated to the weak steric constraints within the IDRs (2, 3). Considering that viruses quickly adapt to new environmental conditions, the exploration of a wider space of sequence possibilities in IDRs could efficiently buffer mutations effects, thus providing for the virus a means of adjustment to the host. The work presented here seeks to experimentally investigate the ability of an intrinsically disordered viral protein to buffer mutation accumulation (i.e maintain its functionality) and to compare it to that of a host protein categorized as a well-ordered one. As experimental model, we used the Potato Virus Y (PVY)/pepper pathosystem. The interaction between the viral protein (VPg) which contains IDRs (4) and the highly structured pepper translation initiation factor eIF4E is required for infection (5). Several randomized mutant libraries of VPg and eIF4E generated by error prone PCR, were screened for their interaction with the corresponding wild type partner by yeast 2-hybrid. Comparative profiles of the loss of interaction versus the accumulation of mutations were analyzed.

1-Charon et al, 2016 Mol. Biosyst. 12 :634-652

2- Tokuriki et al, 2009 Trends in biochemical sciences 34(2) :53-59

3-Brown and Johnson 2011 Current opinion in structural biology 21(3):441-46

4-Hebrard et al, 2009 Virol J. 6 :23

5-Charron et al, 2008 Plant J. 54 :56-68

---

\*Speaker

# A full description of the conformational landscape of cyclized peptides using a robotics approach

Maud Jusot <sup>\*† 1,2</sup>, Marc Vaisset <sup>2</sup>, Eric Ngo <sup>1</sup>, Matthias Lerbinger <sup>1</sup>, Théo Torcq <sup>1</sup>, Dirk Stratmann <sup>1</sup>, Jacques Chomilier <sup>1</sup>, Juan Cortés <sup>2</sup>

<sup>1</sup> Institut de minéralogie, de physique des matériaux et de cosmochimie (IMPMC) – Institut de recherche pour le développement [IRD] : UR206, Université Pierre et Marie Curie (UPMC) - Paris VI, CNRS : UMR7590, Muséum National d’Histoire Naturelle (MNHN) – Tour 23 - Barre 22-23 - 4e étage - BC 115 4 place Jussieu 75252 PARIS, France

<sup>2</sup> Laboratoire d’analyse et d’architecture des systèmes [Toulouse] (LAAS) – Institut National Polytechnique de Toulouse - INPT, Université Paul Sabatier (UPS) - Toulouse III, CNRS : UPR8001, Institut National des Sciences Appliquées [INSA] - Toulouse – 7 Av du colonel Roche 31077 TOULOUSE CEDEX 4, France

The use of peptides for pharmaceutical purposes is constantly increasing, in particular thanks to improved methods to enhance bioavailability by introducing various kinds of chemical modifications. An interesting application of peptides is the inhibition of protein-protein interactions, aiming to modulate the activity of the complex. Peptides extracted from one chain at the interface of a protein-protein complex can bind to their partner and hence compete with the dimerization process. **For this, the extracted fragments (in solution) must be able to mimic their embedded conformation in the initial complex. Their cyclization and chemical modifications are promising pathways to stabilize these highly flexible molecules in the desired conformation.** For those reasons, the study of conformational and chemical properties of peptides, alone and with their targets, is crucial for drug design.

The first step of the present study is the exploration of the peptide’s energy landscape in order to find the local minima (most stable conformations) and the transition paths between them. We are currently developing a faster exploration approach than classical simulation methods. To do so, we use a mechanistic representation of the cyclic peptide which allows us to use algorithms originating from robotics for conformational sampling (1). Our method can also treat D-form and N-methylated amino acids. In order to validate our approach, we sampled the conformational landscape of a set of 19 cyclic RGD designed pentapeptides (2)(3). An exhaustive description of these peptides has been made by complete enumeration of the backbone dihedral angles, phi and psi. The results have been compared to those obtained with Replica Exchange Molecular Dynamics simulations, showing the good performance of the robotics approach.

(1) Cortés, Juan, and Ibrahim Al-Bluwi. "A robotics approach to enhance conformational sampling of proteins." ASME 2012 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. American Society of Mechanical Engineers, 2012.

---

\*Speaker

†Corresponding author: maud.jusot@impmc.upmc.fr

- (2) Wakefield, A. E., W. M. Wuest, and V. A. Voelz. "Molecular Simulation of Conformational Pre-Organization in Cyclic RGD Peptides." *Journal of chemical information and modeling* 55.4 (2015): 806-813.
- (3) Mas-Moruno, C., Beck, J. G., Doedens, L., Frank, A. O., Marinelli, L., Cosconati, S., ... & Kessler, H. (2011). Increasing  $\alpha v\beta 3$  Selectivity of the Anti-Angiogenic Drug Cilengitide by N-Methylation. *Angewandte Chemie International Edition*, 50(40), 9496-9500.

# A Docking Molecular calculations to study of properties of the D-Mannose in the natural treatment of urinary infections caused by E- Coli

Sekkal-Taleb Nezha \* 1,2

<sup>1</sup> Laboratoire de Génie civile et Environnement, LGCE, – Université Djillali liabes de sidi bel abbes, 22000, Algeria

<sup>2</sup> Departement de Pharmacy, Faculty of Medecine – University of Djillali Liabes of Sidi bel Abbès, Algeria

D-Mannose is a natural sugar; it is present in various foods, and binds to *E. coli*, which is then discharged in urine.

The purpose of this research is to prove the efficiency of D-Mannose and provide an approach to the mean duration of treatment by performing tests on laboratory rats, by provocation of UTI contaminating rats by different ways then, administration of D-Mannose orally. A bacteriological examination of urine was carried out and the interpretation of results was based on the sterility of the culture media.

Onether aim of this research is the study of protein-protein interactions, witch have an important role to understand the process of pathogenesis of bacterial and viral infections. Bioinformatics brings also its contribution to the study of protein-protein interactions by methods and software as molecular dynamics and protein -protein docking.

We have study the interaction between D -Mannose and the Fimh protein by the use of molecular dynamics method. Initially, several structural calculations and optimizations by Hyperchem8 software were conducted on D- Mannose to understand how this natural sugar attack the Escherichia coli bacterium.

Then Docking calculations were performed by Hex6.3. Interpretation of results is based on the energy of interaction formed by ligands Alpha -D- mannose and Beta -D- mannose. The lowest energy of interaction of complex probably present a greater inhibition of Fimh protein.

---

\*Speaker

# Author Index

- Anand, Priya, 3  
Andreani, Jessica, 4  
Autiero, Ida, 6
- Barnoud, Jonathan, 8  
Barozet, Amélie, 9  
Bonvin, Alexandre, 23, 32  
Bouchenot, Catherine, 27
- Cabantous, Stéphanie, 27  
Cao, Zhen, 6  
Carbone, Alessandra, 18, 24, 28  
Cavallo, Luigi, 6  
Ceres, Nicoletta, 12, 28  
Champeimont, Raphael, 24  
CHARON, Justine, 34  
CHAUVOT DE BEAUCHENE, Isaure, 10  
Chomilier, Jacques, 29, 31, 35  
Claustre, Agnès, 27  
Combaret, Lydie, 27  
Cortés, Juan, 9, 35  
Cortes, Juan, 31
- de Vries, Sjoerd, 10  
Deval, Christiane, 27
- EL AMRI, Chahrazade, 29
- Fontana, Walter, 20
- GRUDININ, Sergei, 13, 16, 25  
GUEROIS, Raphael, 4  
GYULKHANDANIAN, A, 21
- Harmer, Russel, 20
- ISSA, T, 21
- Jha, Alok, 14  
jusot, maud, 29, 31, 35
- Kadukova, Maria, 16  
KHAZEN, G, 21
- Légaré, Sébastien, 20  
Lagarde, Nathalie, 18  
Laine, Elodie, 28  
Laurent, Benoist, 18  
Le Cornec, Yves-Stan, 20  
Lerbinger, Matthias, 31, 35
- MAROUN, R C, 21  
Marrink, Siwert-Jan, 8  
Melquiond, Adrien, 23  
Molloy, Kevin, 9
- Nadalin, Francesca, 24  
nezha, sekkal-Taleb, 37  
Ngo, Eric, 31, 35
- Oliva, Romina, 6  
Oshurko, Ievgeniia, 20  
OTERI, Francesco, 24
- Pagès, Guillaume, 25  
Pages, Guillaume, 13  
Pietrucci, Fabio, 29  
Polge, Cécile, 27
- Rey, Julien, 4  
Ripoche, Hugues, 28
- Sacquin-Mora, Sophie, 18  
SAWMYNADEN, Jaysen, 29  
Schaarschmidt, Jörg, 23  
Schindler, Christina, 10  
Shaarschmidt, Joerg, 32  
Siméon, Thierry, 9  
STRATMANN, Dirk, 29, 31, 35
- Taillandier, Daniel, 27  
Thierry, Michon, 34  
Torcq, Théo, 31, 35  
Trellet, Mikael, 23, 32  
Tuffery, pierre, 4
- Vaisset, Marc, 9, 35  
Vamparys, Lydie, 18  
Vangone, Anna, 32  
Vavrusa, Marek, 4
- WALTER, Jocelyne, 34  
Wenzel, Wolfgang, 3
- xue, li, 32
- Yu, Jinchao, 4
- Zacharias, Martin, 10  
Zundert, Gydo, 23



