UNIVERSITY of York

This is a repository copy of *Allostery in Its Many Disguises:From Theory to Applications*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/143430/

Version: Other

Article:
Wodak, Shoshana J, Paci, Emanuele, Dokholyan, Nikolay V et al. (36 more authors) (2019) Allostery in Its Many Disguises:From Theory to Applications. Structure. ISSN 1878-4186

https://doi.org/10.1016/j.str.2019.01.003

Reuse
This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Allostery in its many disguises: from theory to applications

Supplementary Material

Shoshana J. Wodak[1],Emanuele Paci[2], Nikolay V. Dokholyan[3a,3b], Igor N. Berezovsky[4], Amnon Horovitz[5], Jing Li[6], Vincent J. Hilser[6], Ivet Bahar[7], John Karanicolas[8], Gerhard Stock[9], Peter Hamm[10], Roland H. Stote[11], Jerome Eberhardt[11], Yassmine Chebaro[11], Annick Dejaegere[11], Marco Cecchini[12], Jean-Pierre Changeux[13], Peter G. Bolhuis[14], Jocelyne Vreede[14], Pietro Faccioli[15], Simone Orioli[15], Riccardo Ravasio[16] Le Yan[17], Carolina Brito[18], Matthieu Wyart[16], Paraskevi Gkeka[19], Ivan Rivalta[20], Giulia Palermo[21], J. Andrew McCammon[21], Joanna Panecka-Hofman[22], Rebecca C. Wade[23], Antonella Di Pizio[24], Masha Y. Niv[25], Ruth Nussinov[26a,26b], Chung-Jung Tsai[26a], Hyunbum Jang[26a], Dzmitry Padhorny[27], Dima Kozakov[27],Tom McLeish[28]*

[1]*VIB-VUB Center for Structural Biology, Brussels, Belgium*
[2]*Astbury Centre, University of Leeds, UK*
[3a]*Department of Biochemistry & Biophysics, University of North Carolina at Chapel Hill, NC, USA*
[3b]*Departments of Pharmacology and Biochemistry & Molecular Biology, Penn State Medical Center, Hershey, PA, USA*
[4]*Bioinformatics Institute (BII), Agency for Science, Technology and Research (A*STAR), and Department of Biological Sciences, National University of Singapore*
[5]*Department of Structural Biology, Weizmann Institute of Science, Rehovot, Israel*
[6]*Departments of Biology and T.C. Jenkins Department of Biophysics, Johns Hopkins University, Baltimore, USA*
[7]*School of Medicine, University of Pittsburgh, USA*
[8]*Fox Chase Cancer Center*, *Philadelphia, USA*
[9]*Biomolecular Dynamics, Institute of Physics, Albert Ludwigs University, Freiburg, Germany*
[10]*Department of Chemistry, University of Zurich, Switzerland*
[11]*Department of Integrative Structural Biology, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Illkirch, France*
[12]*Institut de Chimie de Strasbourg, UMR7177 CNRS & Université de Strasbourg, France*
[13]*Institut Pasteur & Collège de France, Paris, France*
[14]*van 't Hoff Institute for Molecular Sciences (HIMS), University of Amsterdam, Netherlands*
[15]*Physics Department, Università di Trento and INFN-TIFPA*
[16]*Institute of Physics, Ecole Polytechnique Fédérale de Lausanne, Switzerland*
[17]*Kavli Institute for Theoretical Physics, University of California, Santa Barbara, CA 93106*
[18]*Instituto de Física, Universidade Federal do Rio Grande do Sul, 91501-970 Porto Alegre, RS, Brazil*
[19]*Structure Design and Informatics, Sanofi R&D, Chilly-Mazarin, France*
[20]*École Normale Supérieure de Lyon, Université de Lyon, CNRS, Université Claude Bernard Lyon 1, France*
[21]*Department of Chemistry and Biochemistry, University of California, San Diego, USA*
[22]Division of Biophysics, Institute of Experimental Physics, Faculty of Physics, University of Warsaw, Poland
[23]*Molecular and Cellular Modeling Group, Heidelberg Institute for Theoretical Studies (HITS) and Center for Molecular Biology (ZMBH), DKFZ-ZMBH Alliance, Heidelberg University, Germany; Interdisciplinary Center for Scientific Computing (IWR), Heidelberg University, Germany*
[24]*Leibniz-Institute for Food Systems Biology, Technical University of Munich, Germany*

[25]*Institute of Biochemistry, Food Science and Nutrition, Robert H Smith Faculty of Agriculture Food and Environment, The Hebrew University, Israel*
[26a]*Frederick National Laboratory for Cancer Research, National Cancer Institute, Frederick, USA,*
[26b]*Sackler Institute of Molecular Medicine Department of Human Genetics and Molecular Medicine Sackler School of Medicine, Tel Aviv University, Israel*
[27]*Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY 11794*
[28]*Department of Physics, University of York, United Kingdom*

Corresponding Author:
Shoshana J. Wodak
Email: Shoshana.Wodak@gmail.com

Figure S1: Population shift between the inactive and active states of Ras-GDP/GTP.
(*Ruth Nussinov, Chung-Jung Tsa and Hyunbum Jang)*

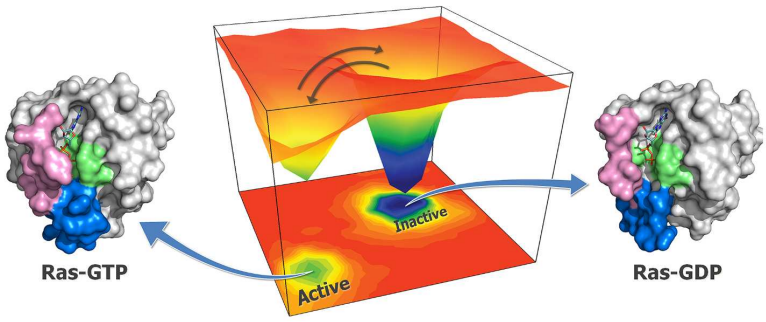


Illustration of the population shift phenomenon related to the authors' recent work on oncogenic mutants of the KRas4B isoform. Shown is the free energy landscape representing active Ras-GTP and inactive Ras-GDP states. Ras can be activated by guanine nucleotide exchange factors that exchange GDP with GTP. The active GTP-bound Ras can be inactivated through GTP hydrolysis by GTPase-activating protein (GAP).

References

-Tsai CJ, Nussinov R. A unified view of "how allostery works". PLoS Comput Biol 2014;10:e1003394
-Nussinov R, Tsai CJ. Unraveling structural mechanisms of allosteric drug action. Trends Pharmacol Sci 2014;35:256-264
-Ma B, Kumar S, Tsai CJ, et al. Transition-state ensemble in enzyme catalysis: possibility, reality, or necessity? J Theor Biol 2000;203:383-397

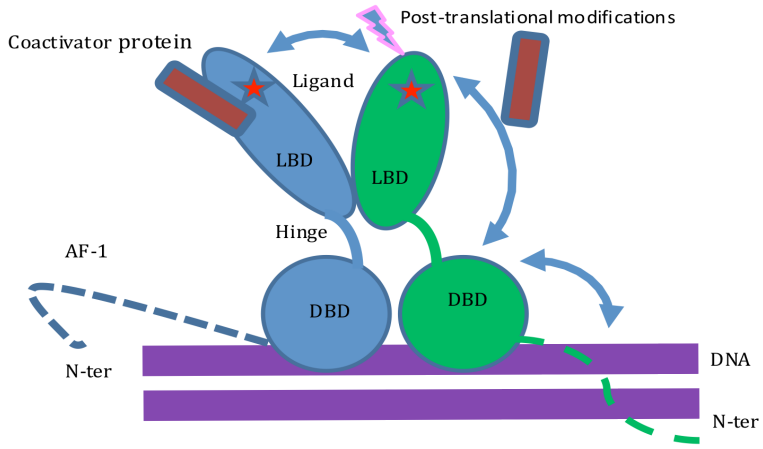


Figure S2: Allosteric regulation in nuclear receptors
(R.H. Stote, J. Eberhardt, Y. Chebaro, A. Dejaegere)

Schematic illustration of the domain structure of nuclear receptor proteins. This structure comprises the AF-1 N-terminal domain, DNA binding domain (DBD), hinge region, C-terminal ligand binding domain (LBD). Sources of allosteric signals include ligand binding and post-translational modifications, which can lead to allosteric signal transmission to the DBD affecting DNA binding. Ligand binding by one receptor can affect ligand binding by the second receptor and the DNA sequence can modulate co-activator binding.

Computational tools have been instrumental in understanding allosteric regulation of NRs [1],[2]. An early study on the PPARγ LBD identified networks of dynamically coupled amino acids that link the ligand binding pocket to the activation helix H12 and the heterodimer interface [3]. A more recently community network analysis has been applied to a larger PPAR architecture [4].

References

1. Grosdidier, S., et al., *Allosteric conversation in the androgen receptor ligand-binding domain surfaces.* Mol Endocrinol, 2012. **26**(7): p. 1078-90.
2. Burendahl, S. and L. Nilsson, *Computational studies of LXR molecular interactions reveal an allosteric communication pathway.* Proteins, 2012. **80**(1): p. 294-306.
3. Fidelak, J., et al., *Dynamic correlation networks in human peroxisome proliferator-activated receptor-gamma nuclear receptor protein.* Eur Biophys J, 2010. **39**(11): p. 1503-12.
4. Ricci, C.G., et al., *Allosteric Pathways in the PPARgamma-RXRalpha nuclear receptor complex.* Sci Rep, 2016. **6**: p. 19940.

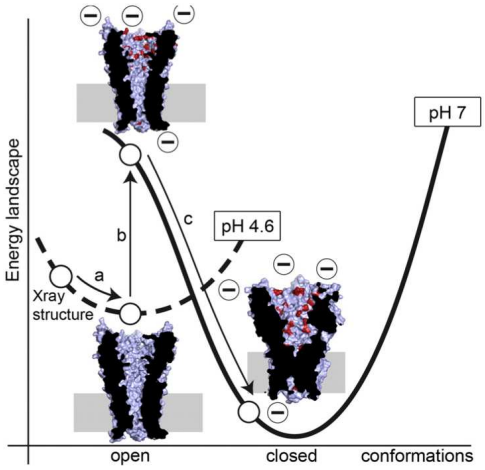


Figure S3: The Activation/relaxation MD approach to explore allosteric transitions of ligand-gated ion channels with atomic resolution: consequences for drug design
(Marco Cecchini and Jean-Pierre Changeux)

Schematic illustration of the activation/relaxation MD approach to allosteric transitions in ligand-gated ion channels. Starting from the X-ray structure of the bacterial receptor GLIC at pH 4.6 (open), a brief equilibration is carried out (a). Then, an instantaneous increase in pH is modeled (b), followed by the relaxation towards a closed conformation (c). The curves with broken and plain lines represent energy landscapes for pH 4.6 and 7.0, respectively. The protein's surface is represented in light blue with residues changing charge during the pH jump in red. Adapted with permission from Ref. [2].

The all-atom based molecular dynamics activation/relaxation approach described in the main text, relies on the hypothesis that allosteric proteins spontaneously populate a number of discrete conformational states in reversible equilibrium and that a conformational selection or *shift* of conformer population takes place upon ligand binding/unbinding [1]. In this view, modeling the addition (removal) of the agonist to (from) its binding site is expected to trigger a "vertical excitation" of the protein, whose structural relaxation by unbiased MD would capture the spontaneous transition to a distinct physiological state..

Lastly, knowledge of the transition path between pairs of physiologically relevant states at atomic resolution enables detailed analysis of the dynamics of the ligand-binding pockets. These include sites for orthosteric ligands and allosteric modulators, which are most often located at subunit or domains interfaces and typically open and close during the functional transitions between resting, active (and desensitized) states. Such analyses can be exploited for the rational design of positive or negative allosteric modulators [1], paving the way to new pharmacological strategies.

References

1. Cecchini, M. and J.P. Changeux, *The nicotinic acetylcholine receptor and its prokaryotic homologues: Structure, conformational transitions & allosteric modulation.* Neuropharmacology, 2015. **96**(Pt B): p. 137-49.

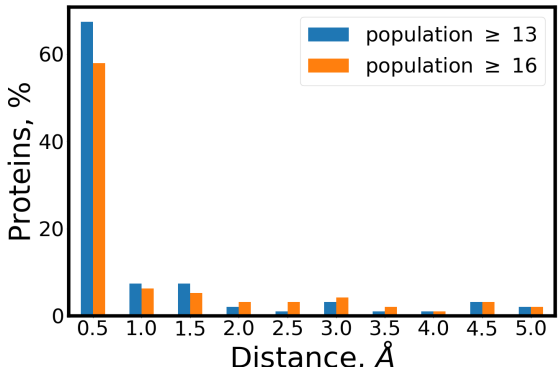


Figure S4: Structural Origins of Cryptic Binding Sites

(Dzmitry Padhorny and Dima Kozakov)

This figure shows the distribution of distances between the cryptic binding site and the closest FTsite probe cluster with above threshold population in the CryptoSite dataset [1]. Analysis is performed based on apo structures of the proteins. In the FTsite approach, clusters of size 13 and more are considered to be binding site predictors, which means that FTsite is able to correctly identify the vast majority of known cryptic binding sites. In a similar manner, clusters with more than 16 probes pinpoint "druggable" binding sites which could be targeted with sufficient affinity, thus showing that more than 80% of known cryptic binding sites are druggable.

References

1. Cimermancic, P. *et al.* CryptoSite: Expanding the Druggable Proteome by Characterization and Prediction of Cryptic Binding Sites. *J. Mol. Biol.* **428**, 709–719 (2016).

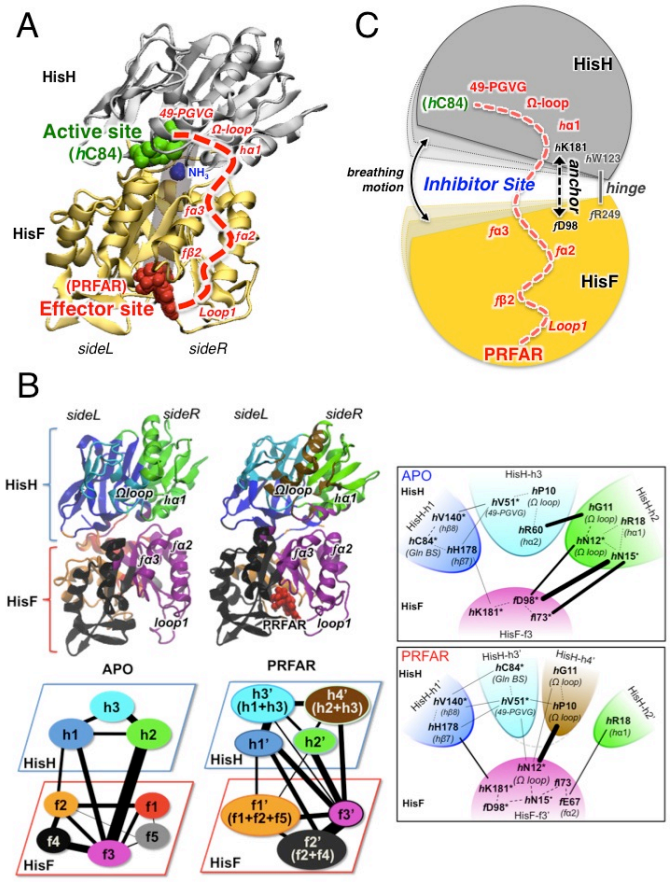


Figure S5: Community network analysis elucidating allostery regulation in the imidazole glycerol phosphate synthase (IGPS) protein
(Ivan Rivalta)

(A) Tertiary structures of *Thermotoga maritima* IGPS complex from *Thermotoga maritima*, with active site in the HisH protein (gray), allosteric site in the HisF protein (yellow) and ammonia (blue) channel (gray tube). IGPS catalyzes the hydrolysis of glutamine (Gln) at the *h* active site, upon binding of the effector PRFAR to HisF, >25 Å away. The allosteric pathways involve secondary structures (in red) on the IGPS right side (sideR). (B) Color-coded optimal community network of apo (left) and PRFAR-bound (right) complexes. The apo communities that contribute the most to the PRFAR communities are reported in brackets. Connectivity line-widths are proportional to the inter-communities total betweenness (ITB) with connections relevant to IGPS allostery depicted in the right panels. Fully and partially conserved residues are indicated with (*) and (•), respectively. Secondary structures are in brackets. (C) The allosteric communication involves structural changes induced by PRFAR binding, affecting the HisF/HisH breathing motion that has been targeted by small molecule inhibitors that bind at the interface. Community network analysis [2,3] was applied to derive optimal community networks of the apo and effector PRFAR-bound IGPS complexes (**B**), indicating that effector binding alters significantly the information flow between communities (thickness of the links between communities in **(B)**). The findings from these studies have been corroborated by measurements of NMR chemical shifts and were shown to be consistent with the expected inactive-to-active allosteric transition in IGPS. This studies have stimulated both single site mutagenesis experiments[4] and allosteric inhibitors design [5] (see **(C)**).

References

1. Manley G, Rivalta I, Loria JP (2013) Solution NMR and computational methods for understanding protein allostery. *Journal of Physical Chemistry B* 117 (11):3063-3073
2. Rivalta I, Sultan MM, Lee NS, Manley G, Loria JP, Batista VS (2012) Allosteric pathways in the imidazole glycerol phosphate synthase. *Proc Natl Acad Sci USA* 109:E1428-E1436
3. Sethi A, Eargle J, Black AA, Luthey-Schulten Z (2009) Dynamical networks in tRNA:protein complexes. *Proc Natl Acad Sci USA* 106 (16):6620-6625
4. Lisi GP, East KW, Batista VS, Loria JP (2017) Altering the allosteric pathway in igps suppresses millisecond motions and catalytic activity. *Proc Natl Acad Sci USA* 114 (17):E3414-E3423
5. Rivalta I, Lisi GP, Snoeberger N-S, Manley G, Loria JP, Batista VS (2016) Allosteric communication disrupted by a small molecule binding to the imidazole glycerol phosphate synthase protein–protein interface. *Biochemistry* 55 (47):6484-6494

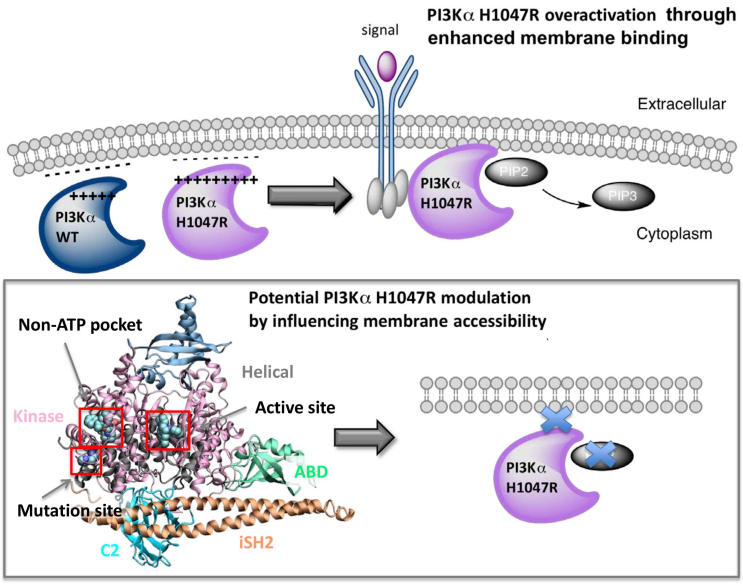


Figure S6: Allosteric modulation in drug discovery: The PI3Kα paradigm
(Paraskevi Gkeka)

Schematic illustration of the proposed allosteric mechanism phosphoinositide 3-kinase alpha (PI3Kα)[1,2].

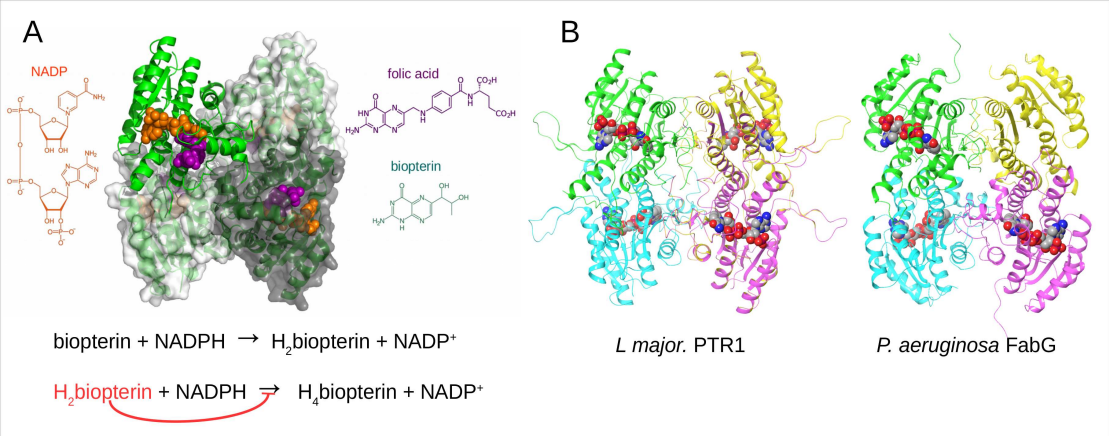
PI3Ks are lipid kinases that play key roles in many fundamental biological processes, including cell growth, proliferation, differentiation, motility, survival, and metabolism. Among the different PI3K isoforms, PI3Kα is the most important as it plays a pivotal role in cell proliferation in response to growth factor-tyrosine kinase pathway activation [3]. PI3Kα is a heterodimer that contains a p110α catalytic subunit and p85α regulatory subunit [1], and is a target of particular pharmacological interest for anti-cancer drug development [3,4]. Currently, two PI3Kα inhibitors, Alpelisib and MLN1117, are in Phase III trial for patients with HR+/HER2–advanced breast cancer and Phase II trial in patients with advanced solid tumors, respectively (https://clinicaltrials.gov, November 2017). Despite the advances in targeting the PI3Kα isoform, a major challenge remains because of the highly conserved ATP binding sites of lipid kinases. Based on molecular modeling techniques and surface plasmon resonance experiments, a multifactorial model of the over-activation mechanism of the most common *PIK3CA* mutant (H1047R), was proposed [1]. Using this model a new binding pocket distinct from the active site, and close to residue 1047,was identified. This pocket was further explored for potential allosteric modulation of the H1047R PI3Kα mutant [2]. Positional covariance, protein fluctuation analyses showed that the main areas involved in protein conformational changes upon ligand binding to the non-ATP pocket are the membrane binding domains. PI3Kα mutant activity could therefore, be modulated not in terms of the active site activity, but by altering the dynamics of protein-cell membrane interaction and subsequent substrate retrieval. Thus, the non-ATP pocket could potentially be used for the discovery of a selective inhibitor of protein−membrane interactions tailored for the H1047R mutant protein. Such an inhibitor could act by constraining PI3Kα membrane-binding motifs, influencing membrane accessibility and subsequent substrate availability, representing a promising alternative or complementary strategy for allosteric PI3Kα modulation.

References

[1] Gkeka, P; Evangelidis, T; Pavlaki, M; Lazani, V; Christoforidis, S; Agianian, B; Cournia, Z: Investigating the Structure and Dynamics of the *PIK3CA* Wild-Type and H1047R Oncogenic Mutant. *PLoS Comput Biol* **2014**, 10(10): e1003895.
[2] Gkeka P; Papafotika, A; Christoforidis, S; Cournia, Z: Exploring a Non-ATP Pocket for Potential Allosteric Modulation of PI3Kα. *J. Phys. Chem. B* **2015,** 119 (3), 1002-1016.
[3] Samuels, Y. *et al*. High frequency of mutations of the PIK3CA gene in human cancers. *Science* **2004**, 304:554.
[4] Fruman, D A.; Chiu, H; Hopkins, B D; Bagrodia, S; Cantley, L C; Abraham, R T: The PI3K Pathway in Human Disease Cell **2017**, 170(4), 605–635.

Figure S7: Allosteric modulation of pteridine reductase 1, a member of the short-chain dehydrogenases/reductases family

(Joanna Panecka-Hofman and Rebecca C. Wade)



Displayed are experimental and modeled structures of Pteridine reductase 1 (PTR1, EC 1.5.1.33), a folate pathway enzyme unique to trypanosomatid parasites. (**A**) PTR1 homotetramer structure (*T. brucei*, pdb: 3bmc), and the chemical structures of the main substrates and the NADP cofactor, and the reaction catalyzed by PTR1 with biopterin as an exemplary substrate (substrate inhibition shown in red). (**B**) Structures of PTR1 (a model based on pdb: 1e92) and FabG (pdb: 4ag3). The proteins are shown in cartoon and molecular surface representation and NADP as van der Waals spheres (only heavy atoms are shown). PTR1 catalyzes the reduction of folate and biopterin with NADPH cofactor [1]. It belongs to a large family of short-chain dehydrogenases/reductases (SDRs), which feature diverse sequences and very diverse functions [3], but have a highly conserved Rossmann fold [3]. PTR1 is a potential target for anti-parasitic drugs, but existing inhibitors require optimization [2]. Targeting allosteric sites of PTR1 is therefore an attractive alternative.

PTR1 is inhibited by semi-products of its catalytic reaction (substrate inhibition) (**A**)) at substrate concentration of ~10-mM levels, and may involve an allosteric mechanism [1]. Cooperativity between the binding sites is also postulated for other SDR enzymes, including homotetrameric bacterial 3-ketoacyl-[acyl-carrier-protein] reductase (FabG, **B**)[4], and allosteric inhibitors binding at the inter-subunit interfaces were reported for FabG [5]. The FabG tertiary and quaternary structure is similar to that of PTR1, despite low sequence identity (~29% for *L. major* PTR1 and *P. aeruginosa* FabG (**B)**). The possibility of allosteric regulation in PTR1 was investigated using normal mode analysis (NMA), indicating significant concerted movements of the substrate loops flanking the active site. In addition, Rotamerically Induced Perturbation simulations (RIP [6]), revealed flexibility hot-spots in the PTR1 homotetramer structure, and suggested dynamical allosteric couplings between distant protein residues.

References

[1] Nare B, Hardy LW, Beverley SM. J Biol Chem. 1997; 272:13883-13891. doi: 10.1074/jbc.272.21.13883

[2] Panecka-Hofman J, Pöhner I, Spyrakis F, Zeppelin T, Di Pisa F, Dello Iacono L, Bonucci A, Quotadamo A, Venturelli A, Mangani S, Costi MP, Wade RC. Biochim Biophys Acta. 2017; 1861(12):3215-3230. doi: 10.1016/j.bbagen.2017.09.012.

[3] Kavanagh KL, Jörnvall H, Persson B, Oppermann U. Cell Mol Life Sci. 2008 Dec; 65(24):3895-906. doi: 10.1007/s00018-008-8588-y.

[4] Price AC, Zhang YM, Rock CO, White SW. Biochemistry. 2001 Oct; 40(43):12772-81.

[5] Cukier CD, Hope AG, Elamin AA, Moynie L, Schnell R, Schach S, Kneuper H, Singh M, Naismith JH, Lindqvist Y, Gray DW, Schneider G. ACS Chem Biol. 2013 Nov;8(11):2518-27. doi: 10.1021/cb4005063.

[6] Ho BK, Agard DA. PLoS Comput Biol. 2009 Apr; 5(4):e1000343. doi:10.1371/journal.pcbi.1000343.

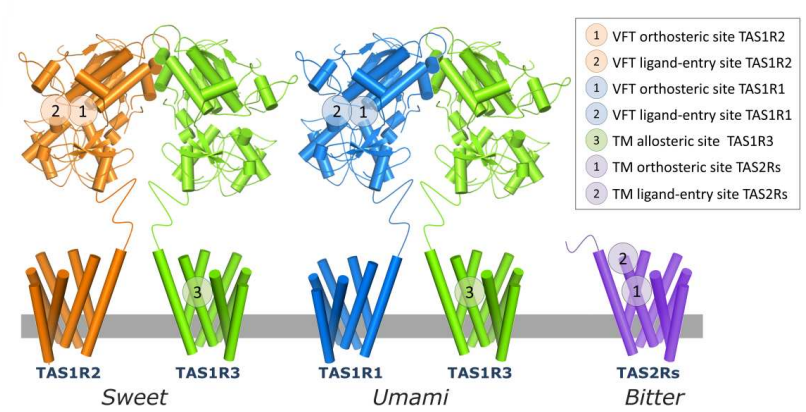


Figure S8: Allosteric modulation of taste GPCRs
(Antonella Di Pizio and Masha Y. Niv)

The Figure illustrates the ligand-binding sites of sweet (TAS1R2/TAS1R3, in orange and green), umami (TAS1R1/TAS1R3, in blue and green) and bitter (TAS2Rs, in violet) taste receptors, members of the G-protein coupled receptors (GPCRs) family.

The odorant or olfactory receptors (ORs) represent the largest family of GPCRs, with over 1000 members in mice and ~400 in humans. Sweet (elicited by sugars) and umami (elicited by glutamate) taste modalities are mediated via TAS1Rs, and bitter (elicited by caffeine, quinine, strychnine and hundreds of other compounds), is mediated by the TAS2Rs family that counts 25 subtypes in humans. Only the crystal structures of the extracellular VFT regions of the fish Tas1r2/Tas1r3 heterodimer are currently available [1], but iterative combinations of simulation and experiment have been successfully used for delineating binding modes of tastants, and for predicting additional ligands [2]. TAS1Rs (Class C GPCRs) contain an extracellular VFT as an N-terminal domain, linked to the TM domain via a short cysteine-rich (CR) domain. VFT, the orthosteric ligand-binding domain, consists of two lobes that can assume an open (inactive) or closed (active) conformation. Umami taste is mediated by a heterodimer composed of TAS1R1 and TAS1R3, and sweet is sensed by a heterodimer of TAS1R2 and TAS1R3.

The Sweet receptor, TAS1R2/TAS1R3 heterodimer is activated by diverse compounds, ranging from low molecular weight sweeteners to sweet proteins. The umami receptor, TAS1R1/TAS1R3 heterodimer, specifically responds to glutamate. Integrated computational approaches and mutagenesis studies identified the binding site for sugars in the VFT domain of TAS1R2, and of glutamate in the VFT domain of TAS1R1, as reviewed in [2]. An allosteric binding site for small molecules has also been identified in the TM domain of TAS1R3. Cyclamate and lactisole bind to an allosteric site in the TM domain, and act as allosteric enhancer and negative modulators, respectively, for both sweet and umami receptors [2]. In addition, multiple binding sites for different ligands have been identified on the sweet taste receptor. For example, small agonist molecules, such as sucralose, saccharin, aspartame and neotame, bind to the orthosteric binding site in the TAS1R2 VFT domain, but positive allosteric modulators, sweet taste enhancers, were found to bind at the entrance of the VFT domain of the TAS1R2 (ligand-entry site).

Allosteric regulation in bitter receptors is less well understood. TAS2Rs are classified as Class A GPCRs, although the sequence similarity is very low (13-29% for the TM domains). Despite low sequence identity the orthosteric binding pocket of bitter taste receptors coincides with the canonical binding site of Class A GPCRs. In addition, an extracellular or vestibular site has been described to be transiently involved in the binding process [3].

References

1. Nuemket N, Yasui N, Kusakabe Y, Nomura Y, Atsumi N, et al. (2017) Structural basis for perception of diverse chemical substances by T1r taste receptors. Nat Commun 8: 15530.
2. Di Pizio A, Niv MY (2014) Computational Studies of Smell and Taste Receptors. Israel Journal of Chemistry 54: 1205-1218.
3. Di Pizio A, Levit A, Slutzki M, Behrens M, Karaman R, et al. (2016) Comparing Class A GPCRs to bitter taste receptors: Structural motifs, ligand interactions and agonist-to-antagonist ratios. Methods Cell Biol 132: 401-427.