Yale University

MB&B
260/266 Whitney Avenue
PO Box 208114
New Haven, CT 06520-8114

Telephone:
203 432 6105
360 838 7861 (fax)
mark@gersteinlab.org
www.gersteinlab.org

October 1 2015

Dear Editor of *Proceedings of the National Academy of Sciences*,

Please find our enclosed manuscript titled "Identifying allosteric hotspots with dynamics: application to conservation in deep sequencing", which we hope you will consider for publication in your journal. Our work is largely motivated by the growing need to transcend conventional annotations on protein structure (such as residues essential to protein-protein interactions, post-translational modifications, stability, etc.) in order to better understand the conservation patterns that are increasingly coming to light through next-generation sequencing initiatives. Deep sequencing has unearthed a class of protein elements that are strongly conserved, despite the fact that available structural annotations sometimes fail to explain such conservation. Knowledge relevant to dynamic behavior (and in particular, allostery) may often provide the missing conceptual link. In the first study of its kind, we develop an integrative and computationally tractable framework that directly employs protein structures and models of conformational change to identify potential allosteric residues on both the surface and the within the interior – followed by the identification of hundreds of alternative conformations throughout the PDB, we apply this framework to study the conservation of these residues on a large scale, with conservation being evaluated using multiple metrics and datasets to investigate inter-species and human-specific patterns of constraint. We not only find that our identified residues tend to be conserved using multiple measures and sources of data, but they may also sometimes help to explain otherwise poorly understood human disease-associated variants. Finally, we make our framework available to the scientific community with a newly introduced and accessible software tool (STRESS) that is easy to use.

We list a number of suitable reviewers for this work.

Yours sincerely,

Mark Gerstein
Albert L. Williams Professor
of Biomedical Informatics