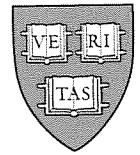


HARVARD UNIVERSITY  
DEPARTMENT OF CHEMISTRY AND CHEMICAL BIOLOGY  
12 Oxford Street, Cambridge, MA 02138, USA



November 8, 2005

Yves Brun  
Systems Biology/Microbiology Faculty Search  
Department of Biology  
Indiana University  
Jordan Hall 142, 1001 E 3rd St.  
Bloomington IN 47405-7005

Dear Dr. Brun:

This letter is written to recommend Boris Shakhnovich for a faculty position in your department.

Outright disclosure: Boris is my son. While that partially disqualifies me as his recommender it gives me some ammunition as I know him better than perhaps anybody else. The reason I am writing this letter is because Boris worked with me on a significant part of his PhD and I was fortunate enough to see him grow into a brilliant scientist.

Boris and I started to work together in 2001 when Boris joined the Bioinformatics program at BU (and of course we talked about science many times before that, perhaps from his age of 5 or so but that is less relevant to this letter). At that time we (with my postdoc Nikolay Dokholyan who is now an Assistant Professor at UNC) were trying to understand the global structure of protein Universe and possible mechanisms of its evolution. When Nikolay and I got an idea that graph-theoretical approach may be key here (where nodes are protein domains and edges are drawn according to structural similarity between domains) the concept of PDUG (Protein Domain Universe Graph) was born and its highly non-random properties were found to be suggestive of evolutionary dynamics of proteins ("Big bang" scenario). Boris joined this effort with utmost enthusiasm as he was fascinated by the concepts and their possible implications for biology. He contributed to the bioinformatics analysis of PDUGs hence he is a coauthor of the first PNAS (2002) paper on the subject but in this particular work his contribution was not defining.

After that he took off like a rocket! He quickly and by himself understood deep biological consequences of the proposed divergent evolution "Big bang" scenario that led him to prediction of special patterns related to structure-function relationships in proteins, and he discovered these patterns (which he called functional fingerprints) by looking at the distribution of GO functional annotations in PDUG clusters. This discovery (published in 2003 JMB where Boris is naturally first author) has immediate and major implication for

structural genomics and for understanding of evolution of protein function. It is noteworthy that highly visible subsequent papers in the field (e.g. KS Kim, 2004 PNAS) essentially reproduced Boris's analysis (unfortunately without reference to Boris's original contribution in 2003 MB). Boris's 2003 JMB is a landmark work that for the first time put structure-function relationship on a solid quantitative footing.

After that Boris continued his steeply rising trajectory of research. He realized, entirely on his own, that structural relationship is only one "dimension" for graph-theoretical analysis of protein evolution and went on to introduce another crucial dimensions – functional and phylogenetic. These are very beautiful ideas as he came up with a method to determine a rigorously functional distance between different protein domains by looking at GO functional distributions within sets sequences that fold into a given domain. Similarly beautiful is his definition of phylogenetic distances based on the concept of mutual information in phylogenetic analysis of appearance of domains in different organisms. Boris developed a synthetic view on these three dimensions in protein universe and came up with a truly remarkable discovery (paper published in inaugural issue of PLoS Computational Biology, Boris is sole author) that structure-function relationship is crucially determined by phylogenetic context of proteins. This is a real breakthrough that will have immediate and direct implications for structural genomics since it will significantly help in efforts to determine function from sequence and/or structural homology and will dramatically reduce the rate of misannotations (false positives).

Another instance of Boris's groundbreaking work is his study distribution among folds in sequence space (2005 Genome Research paper, featured on the cover of the issue as cover art). Here Boris used the concept of designability developed by others, and its structural determinant found recently by Jeremy England, a very talented undergraduate in my lab, to determine what defines population of gene families – an old unsolved problem in molecular biophysics. Boris introduced a new concept – functional flexibility score – and carried out a virtuoso bioinformatics analysis where he was able to discern historical and structural contributions determining gene family sizes. Moreover, Boris showed that proteins belonging to Last Universal Common Ancestor were more designable than more recent proteins – a key discovery that appears to be instrumental in our quest for understanding the molecular origin of life.

Along related lines Boris (with Jeremy England, PNAS 2003) developed a new theory of thermophilic adaptation where he demonstrated that organisms adapt to extreme conditions not only by selecting sequences but by biasing their structural repertoire in a predictable and understandable way. In this paper Boris is formally a second author but his contribution is equal to Jeremy's who is formally the first author.

Boris's more recent (yet unpublished) work on the role of functional constraints in evolution of gene families is a gem. I observed evolution of his thinking about this key problem of biology. The approach is based on the analysis of lethal vs non-lethal genes (data provided by knockout RNAi experiments) and his first intent was to correlate lethality of a gene with designability of the structure that this gene encodes. He found

such correlation indeed and standard way would have been to fire a paper to Science (and we have seen too many such superficial Science papers recently). However Boris thought deeply and understood that while correlation does exist it is not causal but rather due to common historical reasons for higher lethality and higher designability. From that a beautiful story unfolded that reveals how evolutionary pressure on the organism translates into striking signatures of molecular evolution of genes. This work provides key general insights on the relation between organismic (Darwinian) and molecular evolution. His finding stimulated me and my coworkers to develop a microscopic evolution model that could further explain these findings and we are very actively pursuing this direction of research.

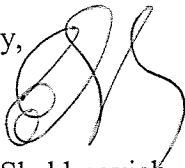
Boris is an extremely broad Computational Biologist. Besides the work that I described in this letter he is very successfully working on a number of other projects such as identification of Transcription Factor Binding Sites, evolution of transcriptional networks, service to Bioinformatics Community (developed several databases), etc.

Boris's work has been a stunning success: he made several major discoveries over a short time span of 4 years of his active work in Computational Biology Perhaps one of reasons for that, besides his outstanding ability, is his breadth of education and thinking. He is equally fluent in evolutionary biology, molecular biology, computational biology and bioinformatics, statistical mechanics, and computer science. This and his genuine interest to uncover mysteries of life is a firm guarantee that his work will be stellar for years to come. Any department that brings him on board will be fortunate to have such a prolific and brilliant colleague. He will bring multiple facets of Computational Biology expertise and will enrich intellectual life at any department that he joins. His already significant teaching experience in Computational Biology will be an asset to develop a modern Bioinformatics curriculum.

In my reference letters I usually provide calibration comparisons with my best students and postdocs that I have had in my 15+ years tenure at Harvard. However in this case I will refrain from that for obvious ethical reasons.

I hope that these comments are of help in your deliberations.

Sincerely,



Eugene Shakhnovich  
Professor of Chemistry & Chemical Biology, Biophysics

ES/jm