

F. Edward Boas

<http://www.stanford.edu/~boas>

boas@stanford.edu · 650-787-1688 · 3375 Alma St #166, Palo Alto, CA 94306

Education

California medical license # A108883. National provider identity # 1487815692.

Residency, Radiology, Stanford Hospital, July 2009 – present.

Internship, General surgery, Stanford Hospital, June 2008 – June 2009.

M.D. / Ph.D., Biochemistry, Stanford University, June 2008.

A.B., Biochemistry, magna cum laude with highest honors, Harvard University, June 1999.

Selected research and publications

Metal artifact reduction in CT scans. Developed a method for reducing streak artifacts in computed tomography scans.

- Boas FE. (2008) “Method and apparatus for reducing artifacts in computed tomography images.” U.S. Patent Application 2008/0273651.

Custom binding proteins. Developed and validated a general technique for predicting binding constants, predicting structures of protein binding sites, and designing binding proteins. This was the first successful design of a binding protein using a standard physical model.

- Boas FE, Harbury PB. (2008) “Design of protein-ligand binding based on the molecular-mechanics energy model.” *Journal of Molecular Biology*. 380: 415-24.
- Boas FE and Harbury PB. (2007) “Potential energy functions for protein design.” *Current Opinion in Structural Biology*. 17: 199-204.

Red blood cell aging. Identified a chemical signal used to recognize and remove old red blood cells from the circulation.

- Boas FE, Forman L, Beutler E. (1998) “Phosphatidylserine exposure and red cell viability in red cell aging and in hemolytic anemia.” *Proc Natl Acad Sci, USA*. 95: 3077-3081.

Leadership

President, Harvard Science Review (1998 – 1999)

Worked closely with a staff of more than twenty editors, writers, layout designers, and business managers to produce Harvard’s oldest and largest undergraduate science publication. Also served terms as Editor-in-Chief and Online Manager.

President, Hippocratic Society (1998 – 1999)

Harvard Conference Director, MIT-Harvard Conference on Genetic Technology and Society. This two-day conference featured an international all-star cast of Nobel laureates, corporate CEOs, and congressmen, and drew an audience of 1,200 people.

Selected awards

Medical Scientist Training Program (1999 – 2008)

Barry M. Goldwater Scholar (1998)

United States Presidential Scholar (named by Bill Clinton, 1995)

Top ten finalist, Westinghouse (now Intel) Science Talent Search (1995)

USA Today All-USA Academic First Team (1995)

15. Boas FE, Fenn RS, Lipfert J, Harbury PB. (2008) "Physics-based design of new binding proteins." Manuscript in preparation.
14. Boas FE, Fleischmann D. (2008) "Selective algebraic reconstruction technique (SART) for reducing streak artifacts in computed tomography scans." Submitted.
13. Boas FE. (2008) *Physics-Based Design of Protein-Ligand Binding*. PhD dissertation, Department of Biochemistry, Stanford University.

Different potential energy functions have been used in protein dynamics simulations, protein design calculations, and protein structure prediction. Clearly, the same physics applies in all three cases, so the variation in potential energy functions reflects differences in how the calculations are performed. With improvements in computer power and algorithms, the same potential energy function should be applicable to all three problems.

Here we show that a standard molecular-mechanics potential energy function without any modifications can be used to engineer protein-ligand binding. A molecular-mechanics potential is used to reconstruct the coordinates of various binding sites with an average root mean square error of 0.61 Å, and to reproduce known ligand-induced side-chain conformational shifts. Within a series of 34 mutants, the calculation can always distinguish weak ($K_d > 1$ mM) and tight ($K_d < 10$ μM) binding sequences. Starting from partial coordinates of the ribose binding protein lacking the ligand and the ten primary contact residues, the molecular-mechanics potential is used to redesign a ribose binding site. Out of a search space of 2×10^{12} sequences, the calculation selects a point mutant of the native protein as the top solution (experimental $K_d = 17$ μM), and the native protein as the second best solution (experimental $K_d = 210$ nM). The quality of the predictions depends on the accuracy of the generalized Born electrostatics model, treatment of protonation equilibria, high resolution rotamer sampling, a final local energy minimization step, and explicit modeling of the bound, unbound, and unfolded states.

After this initial proof of principle experiment, we next used a standard molecular mechanics potential energy function to redesign ribose binding protein to bind a series of ligands: L-arabinose, D-xylose, indole-3-acetic acid, and estradiol. The resulting proteins have 5 – 10 mutations from the native, are stable, the predicted structures have good hydrogen bonds and shape complementarity, and they use motifs similar to natural binding proteins. All of the designed proteins bind to their target ligands with measurable but weak affinity. The affinity was improved by random mutagenesis and screening.

The application of unmodified molecular-mechanics potentials to protein design links two fields in a mutually beneficial way. Design provides a new avenue to test molecular-mechanics energy functions, and future improvements in these energy functions will presumably lead to more accurate design results.

This is the first time a single model has been used to predict structures, binding constants, and to design new small-molecule binding sites. Using a standard model should improve the generality of protein design, which could enable the creation of custom proteins for a wide variety of applications, including sensors, enzymes, and protein therapeutics.

12. Boas FE, Harbury PB. (2008) "Design of protein-ligand binding based on the molecular-mechanics energy model." *Journal of Molecular Biology*. 380: 415-24.

While the molecular-mechanics field has standardized on a few potential energy functions, computational protein design efforts are based on potentials that are unique to individual labs. Here we show that a standard molecular-mechanics potential energy function without any modifications can be used to engineer protein-ligand binding. A molecular-mechanics potential is used to reconstruct the coordinates of various binding sites with an average root mean square error of 0.61 Å, and to reproduce known ligand-induced side-chain conformational shifts. Within a series of 34 mutants, the calculation can always distinguish weak ($K_d > 1$ mM) and tight ($K_d < 10$ μM) binding sequences. Starting from partial coordinates of the ribose binding protein lacking the ligand and the ten primary contact residues, the molecular-mechanics potential is used to redesign a ribose binding site. Out of a search space of 2×10^{12} sequences, the calculation selects a point mutant of the native protein as the top solution (experimental $K_d = 17$ μM), and the native protein as the second best solution (experimental $K_d = 210$ nM). The quality of the predictions depends on the accuracy of the generalized Born electrostatics model, treatment of protonation

equilibria, high resolution rotamer sampling, a final local energy minimization step, and explicit modeling of the bound, unbound, and unfolded states. The application of unmodified molecular-mechanics potentials to protein design links two fields in a mutually beneficial way. Design provides a new avenue to test molecular-mechanics energy functions, and future improvements in these energy functions will presumably lead to more accurate design results.

11. Boas FE. (2008) "Method and apparatus for reducing artifacts in computed tomography images." U.S. Patent Application 2008/0273651.

We present an iterative method for reducing artifacts in computed tomography (CT) images. In each iteration, constraints such as non-negativity are applied, then the image is blurred to guide convergence to a smoother image. Next, the image is modified using an algebraic reconstruction algorithm to try to match the projection data to within the experimental error. A mask is calculated which specifies which parts of the image to update during each iteration. The mask allows us to first solve regions of the image that are determined by rays with low photon counts (and thus high error). Then, regions of the image determined by rays with higher photon counts (and thus lower error), are solved using those ray sums. Reducing CT scan artifacts results in clearer and higher resolution images, faster scan times, and less radiation use.

10. Boas FE and Harbury PB. (2007) "Potential energy functions for protein design." *Current Opinion in Structural Biology*. 17(2): 199-204.

Different potential energy functions have predominated in protein dynamics simulations, protein design calculations, and protein structure prediction. Clearly, the same physics applies in all three cases. The differences in potential energy functions reflect differences in how the calculations are performed. With improvements in computer power and algorithms, the same potential energy function should be applicable to all three problems. In this review, we examine energy functions currently used for protein design, and look to the molecular mechanics field for advances that could be used in the next generation of design algorithms. In particular, we focus on improved models of the hydrophobic effect, polarization and hydrogen bonding.

9. Bass JL, Bhatia A, Boas FE, Sansary J, Rauch D. (2006) "Validation of a body mass index nomogram for children as an obesity screening tool in young children." *Clinical Pediatrics*. 45(8): 718-24.

8. Arava Y, Boas FE, Brown PO, Herschlag D. (2005) "Dissecting eukaryotic translation and its control by ribosome density mapping." *Nucleic Acids Research*. 33(8):2421-32.

Translation of an mRNA is generally divided into three stages: initiation, elongation and termination. The relative rates of these steps determine both the number and position of ribosomes along the mRNA, but traditional velocity sedimentation assays for the translational status of mRNA determine only the number of bound ribosomes. We developed a procedure, termed Ribosome Density Mapping (RDM), that uses site-specific cleavage of polysomal mRNA followed by separation on a sucrose gradient and northern analysis, to determine the number of ribosomes associated with specified portions of a particular mRNA. This procedure allows us to test models for translation and its control, and to examine properties of individual steps of translation *in vivo*. We tested specific predictions from the current model for translational control of GCN4 expression in yeast and found that ribosomes were differentially associated with the uORFs elements and coding region under different growth conditions, consistent with this model. We also mapped ribosome density along the ORF of several mRNAs, to probe basic kinetic properties of translational steps in yeast. We found no detectable decline in ribosome density between the 5' and 3' ends of the ORFs, suggesting that the average processivity of elongation is very high. Conversely, there was no queue of ribosomes at the termination site, suggesting that termination is not very slow relative to elongation and initiation. Finally, the RDM results suggest that less frequent initiation of translation on mRNAs with longer ORFs is responsible for the inverse correlation between ORF length and ribosomal density that we observed in a global analysis of translation. These results provide new insights into eukaryotic translation *in vivo*.

7. Boas FE. (2000) "Linkage to Gaucher mutations in the Ashkenazi population: Effect of drift on decay of linkage disequilibrium and evidence for heterozygote selection." *Blood Cells, Molecules, and Diseases*. 26(4): 348-59.

The two most common Gaucher disease mutations in the Ashkenazi population, 1226A→G and 84G→GG in the glucocerebrosidase gene, are tightly linked to a marker in the nearby pyruvate kinase gene. This paper develops a simulation of the Ashkenazi population that considers the effects of selection and drift on the mutant allele frequency and the recombinant haplotype frequency over time. Although the fraction of mutants that are linked to the original marker decays exponentially on average, this expected value is not very likely to occur. Instead, due to random loss of the recombinant haplotype, a mutation has a significant probability of retaining complete linkage disequilibrium long after its origin, so there may be large errors in estimating the age of a mutation based on linkage data. The simulations show that the 1226G mutation probably originated between 40 and 1000 generations ago (1000 to 25,000 years ago), and the 84GG mutation probably originated between 50 and 4800 generations ago (1300 to 120,000 years ago). The recent origin of the 1226G mutation and its high current allele frequency provide strong evidence for heterozygote selection. New techniques and results developed in this paper have general applicability towards analyzing linkage disequilibrium near other mutations. For example, they potentially explain the unexpected pattern of linkage disequilibrium seen around the ΔF508 mutation of the cystic fibrosis transmembrane conductance regulator gene.

Key words: Gaucher disease, linkage disequilibrium, heterozygote advantage, genetic drift, Ashkenazi, computer simulation

6. Boas FE. (1999) "Polyvalent inhibitors of *Pseudomonas aeruginosa* adhesion." Senior honors thesis, Harvard University.

We synthesized polymers displaying multiple copies of a peptide from the cystic fibrosis transmembrane conductance regulator (residues 108-117). This peptide is the minimal sequence needed for binding of *Pseudomonas aeruginosa*. Using an *in vitro* bacterial internalization assay, our preliminary results indicate that such a polyvalent ligand can prevent *P. aeruginosa* from attaching to (and being internalized by) epithelial cells. The most effective polymer we synthesized was a polyacrylamide with CFTR peptide covalently attached to 0.5% of its acrylamide subunits. This polymer has an IC₅₀ value of 700 pM, 40,000 times lower than monovalent CFTR peptide. Polyacrylamide at the same concentration does not inhibit internalization.

We also developed a simulation to qualitatively describe how a polyvalent ligand binds to a surface. This simulation demonstrates that the polymer binds to the surface cooperatively. Binding can be enhanced by a larger degree of polymerization, a larger fraction of polymer subunits that can attach to the surface, and greater diffusibility of binding sites on the surface.

5. Demina A, Boas E, and Beutler E. (1998) "Structure and linkage relationships of the region containing the human L-type pyruvate kinase (*PKLR*) and glucocerebrosidase (*GBA*) genes." *Hematopathol Mol Hematol*. 11(2): 63-71.

Both the L-type pyruvate kinase gene (*PKLR*) and glucocerebrosidase (*GBA*) gene are on band q21 of chromosome 1 in humans. Two overlapping P1 bacteriophage clones containing *PKLR* and *GBA* were identified and mapped, defining the locations of these two genes as well as those of the *GBA* pseudogene (*ΨGBA*), metaxin (*MTX*), the *MTX* pseudogene (*ΨMTX*), and thrombospondin 3 (*THBS3*). The distance between 5' ends of *GBA* and *PKLR* was determined to be 71 kb. The direction of transcription of the *PKLR* gene was convergent to that of the *GBA* gene. All 195 Gaucher disease patients homozygous for the 1226G mutation, representing 390 chromosomes with the 1226G mutation, had a PvuII -/- *GBA* haplotype and a C/C at nt 1705 of the *PKLR* gene (-/- haplotype). All 56 Gaucher disease patients who were 1226G/84GG compound heterozygotes manifested a -/+ *GBA* haplotype and 55 of 56 patients were -/+ at *PKLR* nt 1705. Only 1 patient with 1226G/84GG genotype showed a crossover with the *PKLR* polymorphism, with a -/- haplotype at nt 1705. Similarly, 9 patients deficient in pyruvate kinase with the *PKLR* 1529A/1529A genotype were all found to have the same -/- *GBA* haplotype.

Key words: Gaucher disease, anemia, hemolytic, mapping, polymorphism, linkage.

4. Boas FE, Forman L, Beutler E. (1998) "Phosphatidylserine exposure and red cell viability in red cell aging and in hemolytic anemia." *Proc Natl Acad Sci USA*. 95(6): 3077-3081.

Phosphatidylserine (PS) normally localizes to the inner leaflet of cell membranes but becomes exposed in abnormal or apoptotic cells, signaling macrophages to ingest them. Along similar lines, it seemed possible that the removal of red cells from circulation because of normal aging or in hemolytic anemias might be triggered by PS exposure. To investigate the role of PS exposure in normal red cell aging, we used *N*-hydroxysuccinimide-biotin to tag rabbit red cells *in vivo*, then used phycoerythrin-streptavidin to label the biotinylated cells, and annexin V-fluorescein isothiocyanate (FITC) to detect the exposed PS. Flow cytometric analysis of these cells drawn at 10-day intervals up to 70 days after biotinylation indicated that older, biotinylated cells expose more PS. Furthermore, our data match a simple model of red cell senescence that assumes both an age-dependent destruction of senescent red cells preceded by several hours of PS exposure and a random destruction of red cells without PS exposure. By using this model, we demonstrated that the exposure of PS parallels the rate at which biotinylated red cells are removed from circulation. On the other hand, using an annexin V-FITC label and flow cytometry demonstrates that exposed PS does not cause the reduced red cell life span of patients with hemolytic anemia, with the possible exception of those with unstable hemoglobins or sickle cell anemia. Thus, in some cases PS exposure on the cell surface may signal the removal of red cells from circulation, but in other cases some other signal must trigger the sequestration of cells.

3. Boas FE, Forman L, and Beutler E. (1997) "Phosphatidylserine exposure and red cell viability in red cell ageing, storage, and in hemolytic anemia." *Blood*. 90(10) (Supplement 1, Part 1): 272a.
2. Boas FE, Forman L, and Beutler E. (1997) "Phosphatidylserine exposure in red cell aging." *The Scripps Research Institute Scientific Report 1996-97*. pp. 258-60.
1. Ahn CS, Boas E, and Rahn B. (1997) "The geometry and the game theory of chases." *The UMAP Journal*. 18(3): 225-242.

(We investigate the hunting and fleeing strategies of a fast predator and its agile prey in a chase of finite time.)