From Computational Design to a Protein That Binds

Bryan S. Der, *et al.*

Science **332**, 801 (2011);
DOI: 10.1126/science.1207082

BIOCHEMISTRY

From Computational Design to a Protein That Binds

Researchers use computational approaches to design two proteins from scratch that bind to the flu virus.

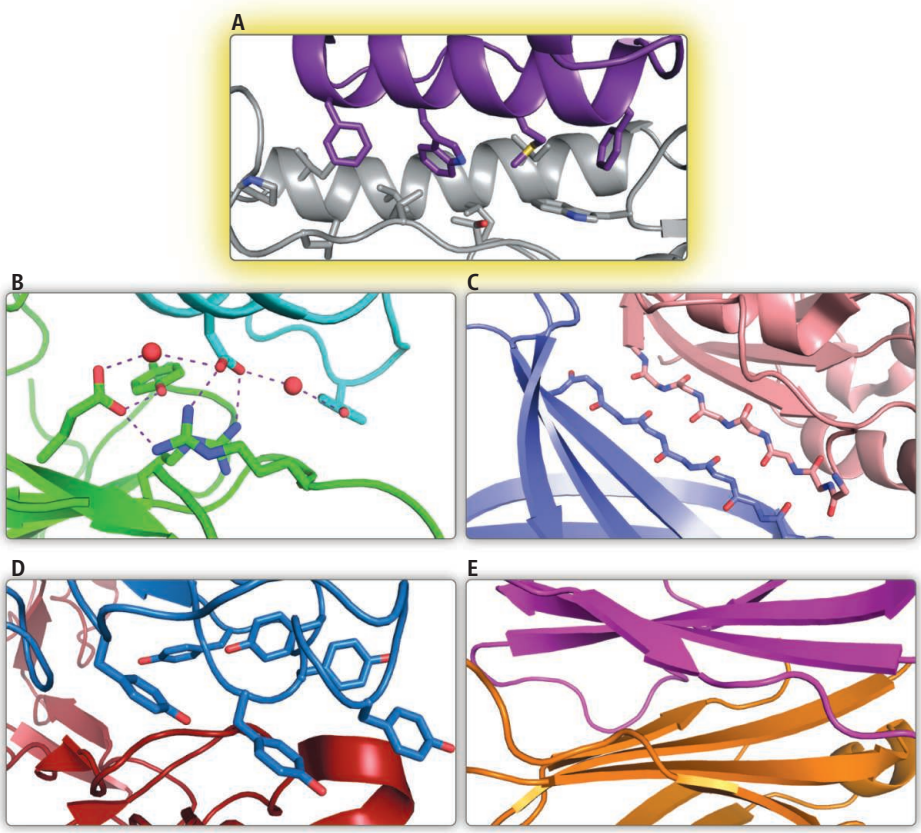
Bryan S. Der and Brian Kuhlman

Protein-protein interactions are critical for many biological processes, and over the past several decades, this importance has prompted researchers to investigate the physical and chemical bases of protein binding. How much do we now understand about how proteins interact? Perhaps the most rigorous way to answer this question is to endeavor to design, at an atomic level of detail, such an interaction from scratch. On page 816 of this issue, Fleishman *et al*. (*1*) take on this challenge. They used a computational method that enabled them to design two proteins that bind to a preselected surface on an influenza virus. This work demonstrates how far we have come in understanding and being able to predict protein-protein interactions, and the technique could prove useful for developing biosensors, reagents, and more effective drugs.

The basic task in computational protein interface design is finding low-energy amino acid sequences with configurations that enable them to "dock" with targeted protein surfaces (*2*, *3*). Candidate sequences must have side chain and backbone atoms that fit together like a jigsaw puzzle, featuring no overlaps, little empty space, and precise chemical complementarity. By designing a protein that binds the stem region of influenza hemagglutinin—a protein that enables the virus to infect a target cell—Fleishman *et al*. solved an enormously complicated jigsaw puzzle. Instead of a large table and hours of free time, however, solving their puzzle required cutting-edge software developed by ~20 groups worldwide (*4*) and 100,000 hours of highly parallel computing time. It also involved using a technique known as yeast display to screen candidate proteins and select those with high binding affinities, as well as x-ray crystallography to validate designs.

The primary novelty of this work is the use of computational methods to generate the high-affinity binders. Common existing approaches involve using animal immune systems to generate antibodies, or screening large "libraries" of candidate proteins (*5*). These two approaches are very powerful for many purposes but do not offer precise control over the binding location or orientation. This specificity is of utmost importance in efforts to develop drugs that target flu's hemagglutinin protein; previous efforts featured antibodies that ended up binding to the protein's rapidly varying head region (*6*), making it harder to prevent acquired resistance. By targeting hemagglutinin's more conserved stem region (*7*, *8*), researchers may be able to develop therapies that are effective against many virus subtypes and prevent acquired resistance. Fleishman *et al*. selected a surface patch in the conserved stem region for this reason and were successful in designing a protein that bound to the patch with the desired orientation.

Previous efforts in computational de novo protein interface design yielded binders with lower affinity (*9*, *10*) or that did not bind in the designed orientation (*11*). What did Fleishman *et al*. do differently to achieve high affinity? One answer is a rapid experimental approach for testing designs. By using yeast display (*12*), the authors were able to test 73 designs while limiting time and resources at the lab bench. Yeast can perform recombina-



CREDIT: BRYAN S. DER AND BRIAN KUHLMAN

Computational challenges ahead. The diversity of protein interface characteristics observed in nature suggests future challenges for computational design. (**A**) Fleishman *et al*. designed a hydrophobic helix (purple) to bind a hydrophobic groove (gray) with unprecedented accuracy in binding location and orientation. (**B**) The high-affinity interaction between the bacterial proteins barnase and barstar features a sophisticated hydrogen bond network that also includes water molecules. (**C**) Strand-strand pairings at an interface feature regular repeats of polar atoms. (**D**) Imitating an antibody interface that features long loops will require precise backbone conformational sampling and scoring methods. Loops provide a rich diversity of backbone conformations, such that binding can occur using only tyrosine and serine side chains (*5*). (**E**) The quaternary structure of an antibody is stabilized by a sheet-sheet interface.

Department of Biochemistry and Biophysics, University of North Carolina, Chapel Hill, NC 27599, USA. E-mail: bder@email.unc.edu; bkuhlman@email.unc.edu

Published by AAAS

tion, avoiding gene-cloning steps, and investigators can obtain initial measures of binding affinity while the protein is displayed on the yeast surface. This obviates the need to individually purify 73 different proteins. And after biophysical characterization of the two binders, the yeast are then used for a selection process called affinity maturation. The effectiveness of this experimental workflow may pave the way for future efforts to test computationally designed protein interactions.

Another possible reason for their achievement involves the computational design protocol. Previous modest-affinity binders were developed by focusing first on the placement of the two proteins, and then on high-resolution all-atom side chain design (*9*, *10*). Fleishman *et al*. took a different approach. They first focused on all-atom placement of disembodied side chains to establish critical "hot spot" interactions, and then on docking the designed protein to its target. This strategy is reminiscent of a previous approach that involved grafting key residues from a known interaction onto a new protein scaffold to generate a new binding pair (*13*). If there are no known hot spot binding motifs, however, these hot spot interactions must be designed de novo (*1*).

To create tight binders (dissociation constant < 50 nM) from their initial hits (apparent affinities of >2000 nM and >5000 nM), Fleishman *et al*. performed affinity maturation with yeast display. An examination of the stabilizing mutations suggested ways of improving their computational methods; for instance, the authors concluded that future modeling efforts should try to take into account subtle movements of sequence backbones, attractive forces known as long-range electrostatics, and the energy costs associated with protein interactions (such as the desolvation cost for "burying" polar atoms). Optimizing a protein energy function, however, presents a challenge, and the insights gained from this single study will need to be combined with results from other design and modeling studies in order to identify robust improvements.

Although Fleishman *et al*. have produced a landmark result, it is evident that computational protein interface design is not a solved problem. Researchers should not be satisfied with one or a few successes in solving these astronomically complex molecular puzzles. Each new puzzle is different from the last; for example, the region of hemagglutinin targeted in this work was hydrophobic and α-helical. Will the computational protocol developed by Fleishman *et al*. also be effective for designing binders for polar surface patches, or targeting alternative secondary structures (see the figure) such as β sheets, strands, or loops? Creating many of these interfaces will require accurate modeling of protein conformations and accurate evaluation of desolvation, electrostatics, and hydrogen bonding.

The endeavor to understand protein interactions will undoubtedly continue for decades to come. And the pursuit should remain persistent, as the impacts of rational design and manipulation of protein-protein interactions can hardly be overstated.

References

1. S. J. Fleishman *et al*., *Science* **332**, 816 (2011).
2. D. J. Mandell, T. Kortemme, *Nat. Chem. Biol.* **5**, 797 (2009).
3. J. Karanicolas, B. Kuhlman, *Curr. Opin. Struct. Biol.* **19**, 458 (2009).
4. A. Leaver-Fay *et al*., *Methods Enzymol.* **487**, 545 (2011).
5. S. S. Sidhu, S. Koide, *Curr. Opin. Struct. Biol.* **17**, 481 (2007).
6. I. A. Wilson, N. J. Cox, *Annu. Rev. Immunol.* **8**, 737 (1990).
7. D. C. Ekiert *et al*., *Science* **324**, 246 (2009).
8. J. Sui *et al*., *Nat. Struct. Mol. Biol.* **16**, 265 (2009).
9. P.-S. Huang *et al*., *Protein Sci.* **16**, 2770 (2007).
10. R. K. Jha *et al*., *J. Mol. Biol.* **400**, 257 (2010).
11. J. Karanicolas *et al*., *Mol. Cell* **42**, 250 (2011).
12. S. A. Gai, K. D. Wittrup, *Curr. Opin. Struct. Biol.* **17**, 467 (2007).
13. S. Liu *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **104**, 5330 (2007).

10.1126/science.1207082

PHYSICS

Brownian Motion Goes Ballistic

Measurements of the Brownian motion of particles in air and in water reveal a smooth "ballistic" motion at very short times.

Peter N. Pusey

When Einstein explained the origin of Brownian motion in 1905, he described the erratic movement of a microscopic particle driven by the thermal motion of liquid molecules as a random walk with sharp changes of direction between each step (*1*). He realized that this picture—the one we seem to see if we watch a particle under the microscope—must break down if we were to look more closely. A moving object would require an infinite force to change its speed or direction discontinuously. The particle actually moves "ballistically" along a smooth trajectory (*2*–*4*), as if it were a microscopic ocean liner on an erratic course (see the first figure, panels A and B). It has taken more than a century to observe this ballistic motion. The studies of Li *et al*. (*5*) were conducted on particles in air; its low viscosity allowed ballistic motion to be followed accurately for extended periods and showed that a particle's instantaneous velocities along its path obey a statistical distribution consistent with thermal motion. Huang *et al*. (*6*) used liquid water; in this higher-density medium, the transition from ballistic motion at short times to diffusive motion at longer ones could be studied in detail.

Einstein estimated (*2*–*4*) that a particle of diameter 1 μm in water would move for only about 0.1 μs over a distance of only about 2 Å before completely changing its speed and direction. The minute magnitudes of these estimates led Einstein to conclude that only the larger-scale diffusive random walk would be observed in practice. The present experiments used technologies undreamed of in Einstein's time. Single particles were held in optical traps—radiation pressure prevents them from settling under gravity. A fast position-sensitive detector (*7*) split the interference pattern formed between light scattered by the particle and the trapping laser beam itself and fed the two parts to two photodiode detectors. Lateral motion of a particle in the trap increases the intensity at one detector but decreases it at the other. The difference between the two signals measures one component of the particle's position. Coupled to fast electronics, this system can measure displacements as small as 0.3 Å in time intervals as short as 0.01 μs.

In a vacuum, a particle in an optical trap would oscillate indefinitely. Gas molecules both dampen the oscillations and introduce random impulses that induce Brownian motion. Li *et al*. determined the time evolution of the position and velocity of a 3-μm-diameter silica sphere trapped in air at about 1/36 of atmospheric pressure (see the second figure, panel A). The underlying oscillatory motion, with a period of about 300 μs, is evident, as is a degree of randomness induced by the gas molecules. Einstein (*2*–*4*) estimated the duration of ballistic Brownian motion to be the time $\tau_p = m/\gamma$

SUPA, School of Physics, University of Edinburgh, Edinburgh EH9 3JZ, UK. E-mail: p.n.pusey@ed.ac.uk