VACCINES BY THE NUMBERS:
Computational Approaches to Design Vaccines Faster

By Amber Dance

Like a "Wanted" poster distributed to a posse, peptide vaccines show the immune system a small sample (about eight amino acids) of a pathogen, training the body to seek and destroy viruses, cells or bacteria that tote identical peptides. But just as a criminal might dye his hair and get a nose job to avoid recognition, some vaccine targets can frequently mutate those peptides, rendering them invisible to the immune system posse. Designing vaccines to handle these shape-shifters has proven challenging using the traditional trial-and-error approach. Today, computer and physical scientists are trying to change that by developing computer programs and simulations that identify the likeliest vaccine candidates and test out their capabilities *in silico*.

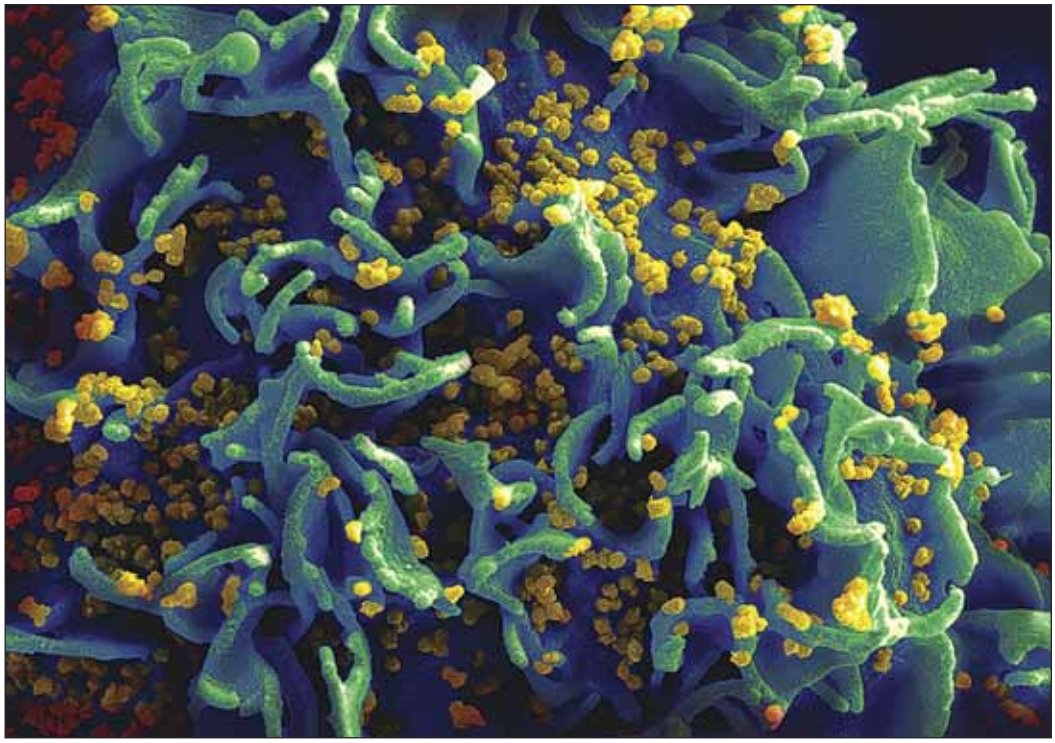
"One of my own goals…is to see how we can make vaccine design a systematic discipline," says **Arup Chakraborty, PhD**, director of MIT's Institute for Medical Engineering and Science in Cambridge. Two recent studies take steps in that direction—and offer hope of vaccines to treat HIV infection and cancer.

"One of my own goals…is to see how we can make vaccine design a systematic discipline," says Arup Chakraborty.

Hitting HIV at Its Most Vulnerable

As soon as immune cells learn to recognize HIV, the rapidly mutating virus tweaks its peptides and becomes invisible once more. Vaccine-makers have tried to foil HIV by training the immune system to recognize the virus's most crucial peptides—those that, if mutated, would weaken the virus. Thus far, this approach has failed because the virus can often make additional compensatory mutations, wiping out the disadvantage caused by the first mutation—rendering the vaccine impotent.

To design effective vaccines, Chakraborty says, scientists need a deeper understanding of how each amino acid in a key protein contributes to HIV fitness, and how multiple mutations work together. It's no simple problem: Each amino acid can mutate to 19 different alternatives. Bioengineering each possible combination would be preposterous. That's where computers can help.



HIV particles attack a human T cell in this scanning electron micrograph. Courtesy of National Institute of Allergy and Infectious Diseases (NIAID).

In a paper published in *Immunity* in March 2013, Chakraborty and his colleagues—**Bruce Walker, MD**, **Andrew Ferguson, PhD**, and **Thumbi Ndung'u, PhD**—computed the fitness landscape for the HIV polyprotein Gag, which contributes the main structural elements of the virus. Here's a simplified example of a fitness landscape: Suppose a virus contains just two amino acids. For each potential amino acid pair, the virus may be fitter or weaker. When mapped, there will be peaks for fit pairs and dips for poor ones. In designing an effective vaccine, Chakraborty says, "You want to push the virus off the hills and into the valleys."

Of course, for the Gag proteins, which together encompass 500 amino acids, this computation is more complex. The first incarnation of the model (in the *Immunity* paper) calculates the fitness of Gag sequences made up of various combinations of wild-type amino acids and alternative (mutated) residues. It considers not only Gag with single mutations, but Gag with

every possible pair of mutations, or three, four or more mutations at once. The group is currently expanding the model, calculating fitness for not just wild-type or mutation, but for any of the 20 possible residues at each place in Gag. To determine fitness, the researchers measured the prevalence of different mutations in HIV DNA sequences collected from patients. They inferred the protein sequence from the DNA and assumed that more predominant strains were the fittest. The result is a multidimensional topographical map, with peaks where the virus does well and valleys where it's weak.

To test the model *in vitro*, the researchers engineered a handful of viruses with different sequences and infected human cells with the strains. Sure enough, the viruses that the computer predicted to be least fit replicated the slowest.

Using fitness landscapes, Chakraborty can identify places where mutations would cripple HIV, and the virus could not easily compensate and evolve back to full fitness. "If you make those mutations, the virus is screwed," he says. The group has designed therapeutic vaccines that should force HIV to make just those mutations, and is working toward a trial in monkeys.

The approach combines two technologies, DNA sequencing and computation, that are becoming ever cheaper, Chakraborty notes. "I think this will be useful for any mutating virus that we have today, or that will emerge in the future," he says. Influenza, for example, is another rapidly mutating virus.

Modeling Cancer Vaccines

Therapeutic vaccines can train the immune system to attack not only infections, but also cancer. "The tumor tries to hide from the immune response," says **Robert Preissner, PhD**, of the Medical University of Berlin. The "Wanted" poster elements of cancer vaccines, designed to target individual cancers, are peptides that represent abnormal proteins on a tumor's surface.

For a vaccine peptide to work, it must interact with two immune system proteins: the major histocompatibility complex (MHC) found in all cells and the T-cell receptors found on the surface of white blood cells. MHC molecules stick out from the surface of cells, displaying sample peptides from inside the cell—including vaccines—and signaling whether the cell contains foreign or otherwise undesirable material. A therapeutic cancer vaccine trains T cells to recognize the native but anomalous proteins expressed in cancer, so they will attack a tumor.

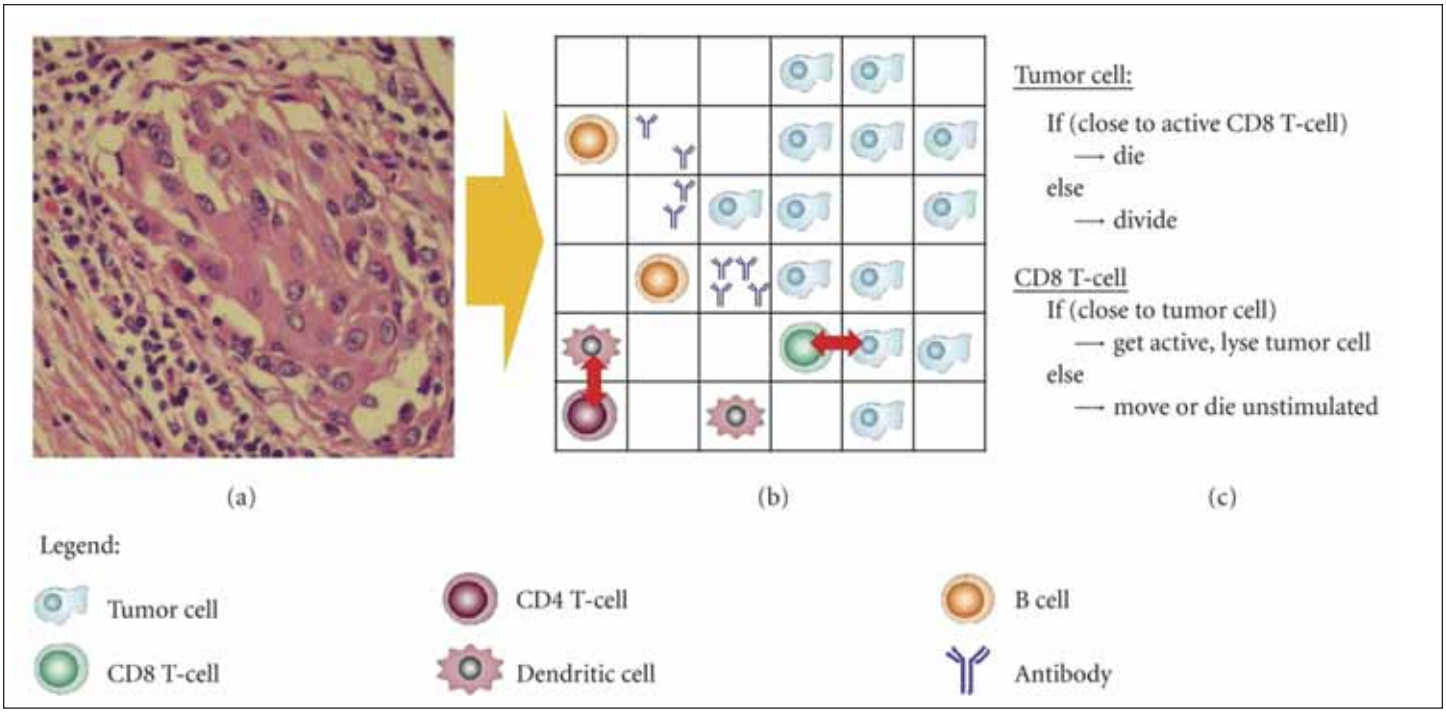
Researchers have made computer models of cancer vaccine action before, but only with simplified yes-or-no interactions between the vaccine, MHC proteins, and T cell receptors. In a study published by *BMC Bioinformatics* in April 2013, Preissner and his colleagues present an updated model, VaccImm, which calculates, in greater detail, the interactions between the specific amino acids in the vaccine peptides and those in the T-cell receptors and MHC molecules. The model should better predict whether a given peptide would help the immune system fight off cancer.

In the VaccImm model, cancer cells, immune cells, antigens and antibodies interact according to set rules. For example, if a T cell recognizes an antigen, it becomes activated and kills tumor cells. Users can input different peptide vaccines, and the program will calculate their effect on tumor growth. VaccImm is freely available online at http://bioinformatics.charite.de/vaccimm/.

In his simulations, Preissner has noticed that multiple vaccine peptides—four or more—work best. This matches the experience of immunologists using multiple peptides *in vivo*. However, only clinical trials will show if peptide cocktails that work in the simulation will work in people, Preissner noted.

VaccImm is still missing useful parameters, Preissner says. For one, there are many different types of MHC molecules. He would like to extend the list of MHC types available to users. In addition, the current incarnation does not allow the cancer cells to mutate, but it should be possible to add this feature.

Francesco Pappalardo, PhD, of the University of Catania in Italy, co-developed the original framework on which VaccImm was based. He says computational vaccine development will save time, money and the lives of experimental animals. Moreover, he says, it will help immunologists understand the biological processes that underlie vaccine success. □



Two-dimensional representation of the players in a VaccImm simulation. Image credit: Robert Preissner.