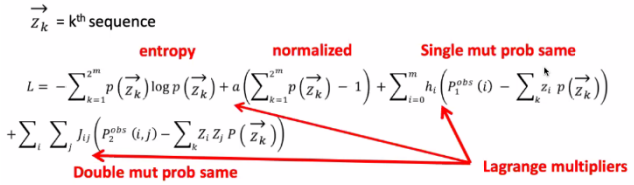
APS COVID Webinar: November 18, 2020

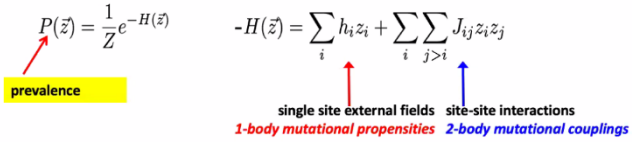
Understanding and harnessing the immune system for better vaccines: a crossroad of physics and biology

Arup Chakraborty, Massachusetts Institute of Technology, Chemical engineering, physics, chemistry, Institute for Medical Engineering and Science

- Adaptive immunity (the human immune system) mounts pathogen-specific responses
- Virus hijacks the transcriptional machinery of the cell to create new viral proteins which can go infect other cells
- We have ~10^11 B cells which have distinct receptors
    - If the specific receptors bind to the virus, a process begins with creates antibodies
    - The antibodies can then bind specifically to the spikes of the virus and dispose of it in several ways
- Humans each have 6-12 types of MHC proteins (though many exist), each binding to different peptides
- Most T cells have distinct receptors, which when the protein binds to it, activate the immune system
- After a virus has been stopped, most of those cells are wiped out, but a few stick around as memory. These are activated for a rapid response if that bug is ever seen in the body again
- All current efforts toward COVID vaccines are focused on eliciting antibodies against the spike protein of SARS-CoV-2
    - protein is made by our cells and antibodies develop against it
- Moderna and Pfizer based on delivering RNA corresponding to the spike protein
- J&J and Astra Zeneca are based on a non-replicating adenovirus that include the spike proteins DNA in its genome
    - spike protein is made by our cells and antibodies develop against it
- We don't have a good record of actually developing vaccines against viruses with high success; for example influenza and HIV
    - The sequence diversity of HIV dwarfs that of influenza
    - The high mutability presents a large challenge to vaccination as the HIV can mutate to avoid the vaccine
- How to hurt HIV with CTLs?
    - Focus on targeting conserved residues in HIV proteins
    - Strategy is that a T cell can attack the virus, it can cause a mutation, making the virus unfit
    - But this strategy can be blunted by compensatory mutations in the virus that restore the fitness and limit the effectiveness
- So we need to understand the fitness landscape of the virus to define mutational vulnerabilities
    - Vaccine induced immune response should push the virus off one of the fitness hills and block the mountain passes
- Immunogen design
    - Minimize regions with compensatory pathways
    - Maximize regions where multiple mutations are especially deleterious
    - Maxi regions that can be presented by people with divers MHCs
- Use sequences from many diverse patients to develop virus fitness landscapes

- Probability of observing a sequence represented by P(z)
- Contains probability of single mutations, double mutations, triple mutations, etc.
- Looking for the least biased model for the probability of observing a sequence, that also fits the observed one and two site mutation probabilities
- Thus least-biased is interpreted as the maximum entropy model





- ← Get a prevalence model
- Is the prevalence landscape a fitness landscape?
  - Each sequence in the ensemble is the product of a non-equilibrium host immunity-virus dynamics in an individual
  - Explore the connection between prevalence and fitness using computer simulations, Feynman variational theory, Fokker Planck equations
  - **Conclude:** The rank order is statistically the same for HIV
  - Example exploration: mapping viral evolution to a 2D Ising model
    - Resulting formulas tell us the underlying reasons that the prevalence and fitness landscape are similar
      - Great diversity of MHC genes implies that no region of the proteome is targeted by a significant fraction of people
      - Immune response evading mutations in one person reverts in infected person whose immune response does not target the same region
      - HIV has never been subjected to classes of effective natural or vaccine-induced memory immune responses
  - Comparison with in vitro experiments:
    - We can predict the fitness of viruses with different mutations
    - Show 27 strains of HIV virus, get high statistical significance for ~80% of the data
  - Comparison with clinical data:
    - Sequence viral strains as a function of time and know the CTL response
    - Asked: Which mutations emerge that escape the immune response? Can we predict this by a fitness landscape?
    - Asked: Can we predict the relative times at which the escapes occur in patients?
    - Used stochastic dynamics with fitness landscapes of a human pathogen under immune pressure
    - Conclude: Able to predict most likely or second most likely sites of escape mutations with 86% accuracy
    - Predict synergistic and antagonistic mutations coupled with the escape mutation strain
      - Explains how fast or slow a mutation will develop (escape time)

- Very successful at separating out mutations that occur quickly and those where escape is difficult (should target these!)

- One way around the delivery problem is to use the adenovirus vectors
  - These require the length of the insert to lie within a certain range (long)
  - Created a immuogen design algorithm which includes the good bits without too many unnatural junctions
    - Need to be minimize
  - Produced a 551 aa immunogen; inserted into a non-replicating viral vector (Ad26)
    - Now being used in the J&J vaccine study!
    - Tested in monkeys
    - Measure strength of the T cell response; show that this method is just as immunogenic as whole antigens; same after prime and boost
- Broadly neutralizing antibodies – HIV
  - Some people make broadly neutralizing antibodies (bnAbs) that neutralize diverse strains (at least in vitro)
  - bnAbs bind to strains that contain shielding residues
  - bnAbs usually emerge several years after infection and in low numbers
  - Tantalizing possibility here!
- Affinity maturation – Darwinian evolution in a short time
  - Mutations in B cells are introduced at high rates; B cells start to bind to the virus; obtain peptides on their surface; then compete with each other to bind to the helper T cells
    - So B cells that bind more strongly to the virus antigen are more likely to display more peptide MHC on its surface and thus more likely to interact with the helper T cells
  - This problem has been studied a lot; but mostly in the context of strain-specific antibodies
  - How can we trick this evolutionary process to create broadly neutralizing antibodies instead of very strain specific antibodies??
    - The induction of bnAbs will require immunization with more than one variant antigen
    - Many questions remain: Which variant antigens? How many variants? What should be the mutations distances? Temporal pattern of administrations? What concentrations?
    - Affinity maturation in the presence of multiple variant antigens is poorly understood!
    - Computational models can help obtain fundamental understanding which can guide this work
      - Activation -> Replication SHM -> affinity-dependent selection ->output/recycle
      - Stop when all B cells die
      - Results: immunization with a cocktail of variant antigens leads to extensive B cell death and extinction
        - The variant antigens present together act as conflicting selection forces which frustrate affinity maturation
      - Results: sequential immunization allows the system to learn how to evolve bnAbs better
      - Memory B cells form by adapting to a variant antigen

        - Results: If the mutational distance is too low, then after the second immunization you will get many memory B cells, but with few acquired mutations
            - If the distance is too high, all the B cells will die
            - There is an optimal mutation distance leads to the right kind of B cell diversity; enables many paths to success and 'clonal interference'
            - There is also an optimum for subsequent variant antigens which drives the system further off equilibrium
            - There are optimal sequential immunization strategies that maximize bnAb evolution

Question and Answer:

- Is there enough sequence data to construct a fitness landscape for COVID-19?
    - Trying to evaluate that right now.
    - There are 2 cases where you'd want to know this:
        - 1. If we do get vaccines, then we will likely see some strong mutations from the virus (like influenza)
        - 2. Can we develop a fitness landscape for all coronaviruses?
            - Maybe pie in the sky for now, but could be very helpful in understanding what is preserved about the virus even across zoonotic shifts
            - Would allow us to use all the techniques described in this talk to help develop effective vaccines
- Is it possible to include the structural info apart from the sequence info in building this computational model?
    - Yes. Have seen that some regions are very difficult to mutate which arise because of structural constraints
    - Can help us understand more mutational vulnerable regions – some work is being done on this already
- What do we know about the mutation rate of this virus, do we need to only worry about the mutations in the spike proteins…
    - SARS-CoV-2 is not mutating very much in comparison to similar viruses
        - Because it is so long, it would die off if too many mutations
    - Humans didn't develop T cells just for fun!
- How are the current vaccinations likely to perform against these mutations?
    - We don't really know how well SARS will evade the vaccine, people are starting to look into this in vitro
- Why do we use the prime and boost methodology?
    - Typically when you vaccinate, the prime activates and mutations some B cells. Then the booster encourages affinity maturation of the B cells, expands the high affinity cells
    - The J&J vaccine is being tested as a one-shot because the monkey study suggests it is potent enough in one dose and makes the logistics much more simple
- RNA vs adenovirus vaccines?
    - No vaccine in the past in either type has ever been deployed on a large scale
- Can you optimize the vaccine strategy based on specific strains for optimal efficacy?

    - The flu spike also contains relatively conserved residues which are preserved across strains… many studies are working to target those residues specifically in an effort to get a universal flu vaccine
        - Difficult because the spike proteins are so close together
- Does a universal fitness landscape really exist? Or is it personally different? Is there any way to move toward more personalized medicine approach to vaccines?
    - We can type the MHC in individuals (determine the immune dominance patterns), then we can determine who is a good candidate for a particular vaccine
- How do you decide whether to target B cells or T cells?
    - For a prophylactic vaccine, you have to have an antibody component
    - For a therapeutic vaccine, you have to have a T cell component as well
    - So most vaccines have strong B cell response, but some, like Yellow Fever, have a strong/key component of T cell targeting