Random Energy Models for Interactions and Dynamics in the Immune Response to Viruses, Vaccines, and Cancers

Michael W. Deem
Rice University
Outline

- Background on the immune system
- The flu shot paradox
- Generalized NK model
- Localization in the immune response
- T cells and dengue fever
- T cells and cancer
Background on the Immune System

- Protection against infection
- B cells originate in bone marrow
- B cells mature and evolve in lymph nodes
- B cells produce antibodies
Function of Antibodies

- Antibodies bind pathogens and toxic products
  \[
  \text{Antibody} + \text{Antigen} \rightleftharpoons \text{Antibody : Antigen}
  \]
  \[
  K = \frac{[\text{Antibody : Antigen}]}{[\text{Antibody}][\text{Antigen}]}, \quad K = (10^4 \text{ to } 10^{11} \text{ liters/mole})
  \]

- Entry of pathogen into cells
- Phagocytic cells to ingest bacteria
- Activate other pathogen-degrading proteins
Hierarchical Creation of Antibody Diversity

- Antibody genes are created by recombination of gene segments in VDJ recombination.
- Antibodies that recognize self die.
- Antibodies that recognize disease multiply.
- The amino-acid space of disease-recognizing antibodies is searched by point mutation in somatic hypermutation.
The Flu Shot Paradox

- "A flu shot this year and not next year, may lead to a greater risk of contracting the flu next year" (costco, 1998)
- Yet flu shot does not affect susceptibility to most other diseases
- And vaccination normally provides protection against disease for multiple years
Original Antigenic Sin

- First exposure to antigen defines antibody response
- Second exposure to new antigen generates response only to coat proteins common with first antigen
- Originally shown for influenza virus in humans and animals
- Original antigenic sin also occurs in dengue fever, HIV

Model of Antibody Structure and Function


A generalized NK model

- $M = 10$ subdomains
- $N = 10$ amino acids per subdomain
- $L = 5$ different subdomain types (helices, strands, loops, turns, others)
- $Q = 5$ classes of amino acids (negative, positive, polar, hydrophobic, other)

$$U = \sum_{i=1}^{M} U^{sd}_{\alpha_i} + \sum_{i > j = 1}^{M} U^{sd-sd}_{ij} + \sum_{i=1}^{P} U^{C}_{i}$$

$$U^{sd}_{\alpha_i} = \frac{1}{[M(N-K+1)]^{1/2}} \sum_{j=1}^{N-K+1} \sigma_{\alpha_i}(a_j, a_{j+1}, \ldots, a_{j+K-1})$$

$$U^{sd-sd}_{ij} = \left[ \frac{2}{DM(M-1)} \right]^{1/2} \sum_{k=1}^{D} \sigma_{ij}^{(k)}(a_{j1}^{(i)}, \ldots, a_{jK/2}^{(i)}, a_{jK/2+1}^{(i)}, \ldots, a_{jK}^{(i)})$$

$$U^{C}_{i} = \frac{1}{\sqrt{P}} \sigma_{i}(a_{i})$$

Gaussian $\sigma$

$K = 4$

$D = 6$

$P = 5$

Energy $\leftrightarrow$ Binding constant

Parameters $\leftrightarrow$ Antigen
Parameters of the Immune Response

- VDJ recombination: Sub-domain sequence pools
- $10^3$ random recombinations bind any given antigen
- Somatic hypermutation: Point mutation and selection
- B cell division time 1/3 day: One round
- 1 mutation per variable region per B cell replication: 0.5 mutation per round
- Response period is 10 days: 30 rounds
- $K = a \exp(-b \, U^{\text{tot}})$
  - Random recombinations
    - $K = 10^4$ \quad $U^{\text{tot}} = -16.64$
  - Primary response
    - $K = 1.3 \times 10^6$ \quad $U^{\text{tot}} = -19.51$
  - Secondary response
    - $K = 10^7$ \quad $U^{\text{tot}} = -20.76$
- Mutation of influenza: Mutation of $U_{\text{sd-sd}}$ and $U_c$ with probability $p$
Possible Suboptimal Dynamics

• Localization in sequence space

• Reduction of diversity of antibody repertoire
Details of VDJ Recombination

- Theoretical heavy-chain diversity of $3 \times 10^{11}$
- In model, diversity $= (L \cdot N_{pool})^{10}$
- Thus, $N_{pool} = 3$
- Dynamics implies choose 3 out of 300 sequences for each subdomain type, $\alpha_i$
- Subdomain $\sigma_{\alpha_i}$ are fixed by structural biology
- Thus, respond with same subdomain pools (genes) to all antigens
Definition of Binding Constant

- The binding or affinity constant
  \[ K_{eq} = \frac{[\text{Antigen : Antibody}]}{[\text{Antigen}][\text{Antibody}]} \]

- Probability of antibody binding specific antigen is proportional to binding constant
- Probability of antibody binding specific antigen is proportional to antibody concentration
- Binding constant is measured as average over 1000 sequences
- Concentration of memory antibodies is $10^2$ that of random antibodies
Dynamics of Immune Response

- Evolve from 1000 random recombinations
- At each round, mutate, and then keep top 20%
- Secondary response starts with $1000 \times 10^2 \times \frac{K_m}{(10^2K_m + K_n)}$ of these memory B cells and adds $1000 \times \frac{K_n}{(10^2K_m + K_n)}$ random recombinations
- Each round of secondary response is also with a total population of 1000 B cells
Localization Does Occur

- Compare first response and second response
- Use only memory B cells on second response (no random recombination)
- Memory B cells initially better, eventually worse
Localization Occurs in Immune Response

- Compare first response and second response
- Use memory and random B cells on second response
- Memory B cells initially better, eventually worse
Localization and Diversity I

- Diversity of random recombinations allows for better response
- Both random recombinations and memory B cells participate
- Secondary response worse than primary response for $p > 0.23$
- This is a result of localization
- Original antigenic sin persists for $0.23 < p < 0.60$
Original Antigenic Sin and the Binding Constant

- Compare primary and secondary immune response
- The localization is visible in the binding constant
When does Cross-Reactivity Cease?

- Examine affinity of memory antibodies for mutated antigen
- Cross-Reactivity ceases when $K_m^{eq} < 10^2 \text{ l/mol}$, the non-specific value
- No cross-reactivity for $p > 0.36$
- Experimentally, cross-reactivity ceases for $p = 0.33 - 0.42$

Exponential Decrease in Affinity

- Property of random energy model (and Nature)
  Evolved response
  Cross response
  \( p = \text{change in sequence} \)

\[
U_{\text{end}} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sigma_i
\]

\[
U_{\text{start}} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sigma_i'
\]

\[
U_{\text{start}} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (p \sigma_i'' + (1-p) \sigma_i)
\]

Since \( \langle \sigma_i \rangle = 0 \)

Thus, linear decrease in energy

\[
\langle U \rangle_{\text{end}} - \langle U \rangle_{\text{start}} = p \langle U \rangle_{\text{end}}
\]
How Many Mutations Occur in the Dynamics?

- Mutations in primary and secondary responses
- Measure smallest distance between best evolved sequence and starting sequences
- Secondary response has fewer mutations than primary for $p < 0.31$
- More mutations in secondary than primary for $0.31 < p < 0.70$
Localization and Diversity II

- How does the number of random recombinations affect the dynamics?
- Use only 500 sequences
- Localization and reduction of diversity still apparent
- Repertoire size does not greatly affect dynamics
Localization and Diversity III

- How does length of evolution affect the localization?
- Evolve for 60 rounds. Repertoire size of 1000.
- Both random recombinations and memory B cells participate
- Secondary response worse than primary for $p > 0.23$
- Original antigenic sin persists for $0.23 < p < 0.65$
Diversity of Memory B cells versus Random Recombinations

- Memory B cells have much lower diversity than do random recombinations
- This is due to the selection step
- This reduction in diversity can limit the amino acid space that memory B cells are able to search
Effect of Reducing Diversity on Random Recombinations

- Keep top $y\%$ of random sequences
- Overwrite bottom $100-y\%$ with $y\%$ sequences
- This mimics effect of memory sequences in second response
Original Antigenic Sin

- "A flu shot this year and not next year, may lead to a greater risk of contracting the flu next year" (Costco, 1998)
- Discovered in humans (Fazekas de St. Growth and Webster, 1966)

A/Panama/2007/99 H3N2 (vac); A/Fujian/411/2002 (circulat)

Munoz and Deem, Vaccine (2004)
Background on the T Cell Immune System

- Fights cellular infection
- T cell precursors are produced in bone marrow
- T cells mature in thymus
- T cells proliferate in lymph nodes
Function of T Cells

- T cells bind pieces of antigens and kill cells containing those antigens.

\[ TCR + \text{peptide/MHCI} \rightleftharpoons TCR: \text{pMHCI} \]

\[ K = \frac{[TCR: \text{pMHCI}]}{[TCR][\text{pMHCI}]}, \quad K = 10^4 \text{ to } 10^7 \text{ liters/mole} \]
T Cell Diversity

- T cell genes are created by recombination of gene segments in VDJ recombination of $\alpha$ and $\beta$ chains
- T cells that recognize nothing die
- T cells that recognize self are killed
- T cells that recognize disease multiply
- $10^{11}$ possible T cell receptors (TCRs)
- $10^8$ at any point in time  

*Science 288* (2000) 1135
T Cell Cross Reactivity

- T cell cross reactivity leads to a number of undesirable effects—immunodominance, original antigenic sin, and anergy
- T cell cross reactivity is studied in altered peptide ligand (APL) experiments
- 1 out of 9 peptide amino acids is altered
- Response of T cells to new peptide antigen is measured
The APL experiment

- T cells are evolved against an original peptide ligand (OPL)
- These T cells are then tested against an altered peptide ligand (APL)
- The APL is changed at one amino acid only
- The change can be either conservative or non-conservative
- Typically the same MHCI express the OPL and APL
The T Cell Muturation Process

- Roughly $2.4 \times 10^7$ sequences in naïve repertoire, copy number $2.4 \times 10^4$
- T cell maturation is driven by cycles of concentration expansion
- Concentration increases $10^3$ over 10 days
- Diversity of evolved sequences is 0.5% of initial, copy number $2 \times 10^6$


- Theoretical total diversity $10^{11}$
- 1 in $10^5$ sequences bind any particular antigen

T Cell interactions with Antigen

- T cells interact via several signaling molecules
- T cells interact with p-MHCI on APC

T Cells Explore Many p-MHCl Interactions

- T cells make initial, hours-long contact with one p-MHCl
- T cells then explore many p-MHCl interactions for minutes
- Dynamic Imaging studies of T cells in lymph nodes

What Drives T Cell Selection?

- "Following clearance of interaction, >90% of activated CD8+ T cells die, leaving behind a stable pool of memory CD8+ T cells capable of responding to subsequent infections with enhanced kinetics." Concluding: "Selection of memory T cell populations is stochastic" and "Maturation of the T cell repertoire during secondary LCMV infection alters the relative magnitudes of epitope-specific responses but does not significantly modify the repertoire of T cells responding to a given epitope." Others and J. D. Altman, J. Immunol. 165 (2000) 6081.


- "The primary antiviral CD8 T cell response was similar both structurally and functionally to that of the memory pool and the secondary CD8 T cell effectors. These results suggest a stochastic selection of memory cells from the pool of CD8 T cells activated during primary infection." J. Exp. Med. 188 (1998) 71.
The Physical Model: Generalized NK (Random Energy) Model


\[ U = \sum_{i=1}^{M} U_{a_i}^{sd} + \sum_{i>j=1}^{M} U_{i,j}^{sd-sd} + \sum_{i=1}^{M} U_{i}^{pep-sd} + \sum_{i=1}^{N_b} \sum_{j=1}^{N_{CON}} U_{i,j}^{c} \]

\[ M=6, \quad N_b=3, \quad N_{CON}=3 \]

\[ U_{a_i}^{sd} = \frac{1}{\sqrt{M(N-K+1)}} \sum_{j=1}^{N-K+1} \sigma_{a_i}(a_j, a_{j+1}, \ldots, a_{j+K-1}) \]

\[ U_{i,j}^{sd-sd} = \sqrt{\frac{2}{DM(M-1)}} \sum_{k=1}^{D} \sigma_{i,j}^{(k)}(a_{j_1}^{(i)}, \ldots, a_{j_{K/2}}^{(i)}, a_{j_{K/2+1}}^{(i)}, \ldots, a_{j_K}^{(i)}) \]

\[ U_{i}^{pep-sd} = \sqrt{\frac{1}{DM}} \sum_{k=1}^{D} \sigma_{i}^{(k)}(a_{j_1}^{pep}, \ldots, a_{j_{K/2}}^{pep}, a_{j_{K/2+1}}^{(i)}, \ldots, a_{j_K}^{(i)}) \]

\[ U_{i,j}^{c} = \frac{1}{\sqrt{N_bN_{CON}}} \sigma_{i,j}(a_{j_1}^{pep}, a_{j_2}) \]

\[ N=9, \quad K=4, \quad L=5 \]

Gaussian \( \sigma \)

\( D=2 \)
Conservative and Non-Conservative Changes

- All 20 amino acids are considered
- Amino acids are grouped into 5 classes
- Within each class are all the conservatively related amino acids
- We use $\sigma = w_j + w_i / 2$
- This gives non-conservative/conservative $= (1 + \frac{1}{4})^{1/2} / (1/2) \approx 2.23$
- From PAM values:
  non-conservative / conservative = 2.34
The Binding Constant

- $K = e^{a-bu}$
- Worst binding constant $3 \times 10^5$ l/mol
- Geometric average binding constant $3 \times 10^6$ l/mol
- Gives typical best binding constant $3 \times 10^7$ l/mol
- Typical values taken from experiment

B. A. Schodin et al., *Immunity* 5 (1996) 137

Specific Lysis

- Probability that an activated T cell will recognize an antigen presenting cell that is expressing a particular peptide-MHC I complex
- Measured as a function of effector to target ratio, $E/T$
  \[ L = \frac{zE/T}{1 + zE/T} \]
- $z =$ clearance probability
  \[ z = \frac{\langle k \rangle}{3 \times 10^7} \]
- Choose energy cutoff so that $z < 1$
T Cell Maturation Process: Primary Response

- 10 rounds of selection with $x = 58\%$
- Repertoire size $N_{\text{size}} = 1000$
- Leads to $10^3$ concentration expansion and 0.5% diversity
- These cells become memory T cells (90% die)
T Cell Maturation Process: APL Experiment

- T cells extracted from spleen of mouse after primary response
- *In vitro*: spleen stimulated, so $N_{\text{size}}$ memory cells only
- *Ex vivo*: spleen is used as is, with $N_{\text{size}}$ naïve cells and $N_{\text{size}}/2$ memory cells
- One conservative change $p_{\text{epitope}} = 0.011$
- One non-conservative change $p_{\text{epitope}} = 0.022$
T Cell Maturation Process: Secondary Response

- *In vivo*
- Binding constants of memory and naïve compared
- Memory cells at 100 × elevated concentration
- If memory used, 3 rounds of selection, leading to 10 × concentration expansion
- If naïve used, dynamics is like another primary response
Specific Lysis: Conservatively Altered Peptides

- Measured for LCMV in mice
- LCMV strongly immunogenic: all memory cells are specific for LCMV
- \textit{In vitro} and \textit{ex vivo} response
  

- \textit{In vitro} > \textit{ex vivo} because memory response is better than naïve response for peptides altered by one amino acid
Specific Lysis: Non-Conservatively Altered Peptides

- *In vitro* and *ex vivo* response
  
  *J. Virol.* 71 (1997) 5764

- Conservative slightly superior to non-conservative because conservatively altered peptide is more similar to the original peptide against which T cells evolved.
Human Disease APL

- APLs that occur in disease can inhibit the immune response
- This is thought to be a big problem
- E.g. human leukemia virus type I


- 11 L → A (left) and 15 Y → A APLs
Immune Response Probability: Conservatively Altered Peptides

- Not widely measured
- For repertoire sizes of $10^3$, $5 \times 10^3$, $10^4$, $5 \times 10^4$, $10^5$, $5 \times 10^5$
- *In vitro* critical point:
  $E^*/T^* = 4.6$
- *Ex vivo* critical point:
  $E^*/T^* = 16$
- Average energies not as favorable in *ex vivo* case → $E^*/T^*$ higher
Immune Response Probability: Non-Conservatively Altered Peptides

- Random energy fluctuations in ensemble lead to smooth curve rather than step function as $N_{\text{size}} \to \infty$
- Fluctuations like variability in individual responses
Fixed point in Immune Response Probability

- Finite size effects in irp curves come from evaluation of constant $a$
- Deviation is $O(1/N_{\text{size}})$
- *Ex vivo* more symmetric than *in vitro* $\rightarrow$ higher irp fixed point
- Conservative smaller negative tail than non-conservative $\rightarrow$ higher irp fixed point
Specific Lysis: Asymmetric Response

- Perform OPL → APL (A) and APL → OPL (B)
- Experiment LCMV-WE → LCMV-8.2 (circles) and reverse (squares)
- Measure correlation in responses
- Note
- Observed Correlation is
- Between None \( \left[ \frac{P(B|A)}{P_A} = 1 \right] \) and Full \( \left[ \frac{P(B|A)}{P_A} = \frac{1}{P_A} = 8.87 \right] \)

\[
(\Delta C)^2 = C_{11} - P_A P_B
\]

\[
\frac{P(B|A)}{P_A} = 1 + (\frac{\Delta C}{P_A})^2
\]

\[
\frac{P(B|A)}{P_A} = 1.30
\]
Dengue fever: Immunodominance

- 4 Strains of dengue fever, 1 conservative mutation between each pair of strains
- Most important vector-borne human virus
- Immunodominance inhibits tetravalent vaccine

Rothman et al., Vaccine 19 (2001) 4694

Park and Deem, q-bio.BM/0404023
Zhou and Deem, 2004
Dengue fever: Noncommutivity

- The order of administration of a series of vaccines matters
- Usually, multivalent → monovalent best

Cancer

- Why so hard to treat?
- Why mouse model essentially non-predictive for human cancer?
- Because real human cancer has many thousands of mutations (Fortune, 22 March 2004)
- Immune response to such cancers suffers from immunodominance
Cancer Escape

- Immune system protects against cancer
- Many standard cancer therapies fail
- Break the immunodominance hierarchy in cancer vaccine

Schreiber et al., Cancer Biology 12 (2002) 25

Park, Yang, and Deem, 2004
Mutations

- Average number of mutations in immune response
- Multi-site vaccination forces the tumor to mutate more in order to escape
Allelic Loss Leads to Escape

- Multi-site vaccination leads to much less escape than does single-site vaccination
- Escape is easier—fewer mutations
Immunodominance in Cancer

- Assume $V = 4$ strains
- $P_{epitope} = 0.15$
Dominance Order can Change

- Probability that dominant epitope changes between naïve and end of primary
- Probability that dominant epitope changes between end of primary and end of secondary
Conclusions

- The immune response can be modeled by a random energy model.
- Localization and reduction in diversity in immune sequence space occurs.
- This localization explains original antigenic sin.
- Theory may help vaccine design take into account immunodominance.
- Theory generally applicable: flu, dengue fever, cancer, immunosenescence.
- Vaccination strategy:
  - Maximize distance between vaccinations against different strains.
  - Suppress memory response during vaccination.
  - Multi-site vaccination.
Acknowledgements

- Jeong-Man Park
- David J. Earl
- Hao Zhou
- Ming Yang
- Enrique T. Munoz
- Hans Schreiber
- Alan L. Rothman
- NSF, NIH/NIAID