

Are the universes of antibodies and antigens symmetrical?

George Pieczenik

ART Institute of Washington, 101 Old Short Hills Road, West Orange, NJ 07052 and Rutgers University, Cook College, Department of Biochemistry and Microbiology, New Brunswick, NJ 08903, USA

Correspondence: e-mail: piecze@rci.rutgers.edu

Abstract

This paper discusses the assumed infinite numbers of antigens and antibodies, and the nature of their interactions, in relation to reproductive immunology. Classical models offered by Pauling, Burnet and Crick are discussed. The author's own model is described, based on selectionist principles, proposing that the numbers are defined, discrete and not infinite. The binding specificity of monoclonal antibodies of four to six amino acids limits the number on the total number of antibody-binding peptides. This limitation on size enables the total number of these peptides to define 'one-hit' or 'two-hit' protein sequences in databases of known proteins. Estimate of the combinations of variable heavy and light chains of different numerical magnitude is 2.5×10^6 . The author proposes the equivalence principle, where equivalent numbers of antigens and antibodies exist and theorizing a limited number of antigen-antibody pairs in the magnitude of 3.2×10^6 or slightly less. The nature of various protein-ligand interactions, indicated by their limited numbers of interactions, clarifies why a particular amino acid is encoded by a particular codon and the nature of encoding by tRNA. In relation to reproductive phenomena, specific proteins for contact with sperm ligands were designed using these parameters.

Keywords: antibodies, antigens, reproductive immunology, symmetry

Introduction

Many regulatory issues in reproductive immunology are still mysterious. Interactions between hormones and cell surface, spermatozoon and egg, blastocyst and uterus are highly precise, but seem infinite when considering millions of species. Are these interactions as large as they appear? The number of antigens and antibodies appears infinite and continuous.

Linus Pauling insisted that antibodies denature around an antigen to form a complementary shape. He patented this universal antibody structure, called a 'Uhr-antibody' (Pauling, 1940). His concept was to solve the question of antibody formation and diversity by structural complementarity chemistry. In Pauling's model, the structural information would flow backwards from antigen to antibody in a direct fashion.

Burnet effectively proposed a variation of Crick's central dogma, that is information flows only from nucleic acid to protein, by stating that there is no structural information flow from antigen to antibody (Burnet, 1957; Crick, 1970). The biological system to make specific antibodies for specific antigens is a selectionist model. This is Burnet's clonal selection theory. He proposed that antigens stimulate cells containing the complete a-priori repertoire of all possible antibodies. The stimulated cell is amplified and makes the antibody complementary to the antigen, by a mechanism still obscure. The selection is a posteriori. Crick's central dogma can be restructured to form more a consequence of a selectionist perspective of molecular biology rather than an unprovable dogma. This selectionist derivation of the central dogma is more consistent with known biological principles (Pieczenik, 1977).

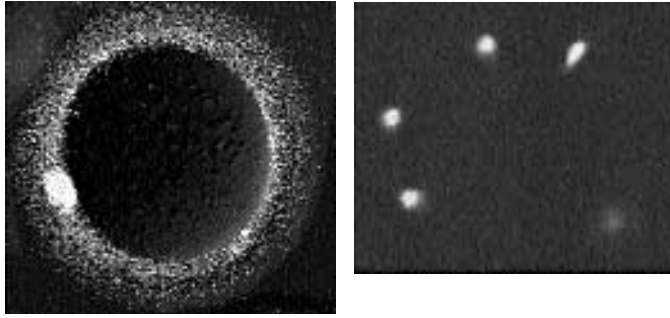
Selectionist principles in defining numbers of antibody-antigen binding pairs

In this paper, selectionist principles and simple mathematical calculations will be used. The proposition is that the number of binding antigens and the number of binding antibodies are equivalent, and that both numbers are defined and discrete and not infinite.

It is known, with the invention of combinatorial libraries of peptides, that the binding specificity of monoclonal antibodies is in the range of four to six amino acids for linear binding protein epitopes (Pieczenik, 1998, 1999). This range of binding size was also previously determined by protection experiments and genetic double hit experiments. This puts a cap on the total number of possible antibody binding peptides to 20^4 , 20^5 or 20^6 . This range is therefore anywhere from 160,000 to 64×10^6 , with a mean around 3.2×10^6 . It is also known that a pentapeptide is sufficient to define a unique 'one-hit' or at the most 'two-hit' protein sequence in the database of existing proteins. Because of this uniqueness, pentapeptide sequences can be used to address the existing known database of proteins. This is a consequence of the limited number of known coding sequences (Lander *et al.*, 2001) and a consequence of the sufficiently large number of pentapeptide sequences.

In the following paragraph, the number of coding possibilities for antibodies will be calculated.

It is also known from the V Base web site (<http://www.mrc-cpe.cam.ac.uk/vbase-ok.php?menu=901>) that the number of combinations of variable, heavy, light chain, constant and linkers are of the following numerical magnitudes: $VH \times D \times$



Figures 1 (left) and 2 (right). Oolemma peptide specific for sperm ligand contact points on the oolemma. This peptide ligand was designed without oolemma substrate but simply as a ‘complementary’ peptide to the sperm head ligand. The ability to make such functional ‘complementary’ binding peptides suggests a limited universe of ligand interactions that may mimic the limited symmetrical universe of antibody–antigen interaction. These binding ligand peptides were designed without any actual substrate but only using genomic sequence information, proprietary algorithms, and proprietary peptide complementarity determining procedures.

$JH = 51 \times 25 \times 6 = 7650$, and $VK \times JK = 40 \times 5 = 200$, $VL \times JL = 31 \times 4 = 124$. In the first calculation, for example, the number of variable heavy chain codings (VH) times the number of D region codings times the number of J junction region codings for the heavy chain (JH) equals 7650. Therefore, the total number of possible associations of heavy and light chains is $(7650 \times 200) + (7650 \times 124) = 7650 \times 324 = 2.5 \times 10^6$ (Messmer, 2000). Clearly, with the human genome database finishing completion, it is possible that this number is a lower estimate. Lehninger presents a higher number with slightly different assumptions (Lehninger, 2000). He proposes 1.5×10^7 as the number of possible IgG antibodies. Both of these estimates of the antibody repertoire are much smaller than 20^6 and much larger than 20^4 . However, they are in the magnitude of 20^5 .

Therefore, it is proposed:

(i) that the number of antibodies and antigens in the normal repertoire are equivalent. This can be termed the equivalence principle: number of antibodies = number of antigens;

(ii) that the number is in the range of $20^{4.66}$ to 20^5 . The Feigenbaum number of 4.66 may represent the actual number (Feigenbaum, 1980). The convergence to this equality may be a consequence of statistical evolutionary considerations, as explained below.

How did this equivalence come about?

Clearly from the clonal selection theory and the selectionist view of the central dogma, there could not have been any direct informational and structural constraint by the pentapeptide universe on the number of antibodies. Therefore, is this convergence another case of Nature’s serendipitous selection and dramatic success? Or is there some underlying constraint on the nature of ligand–ligand interactions and peptide codings that allow such a convergence to occur?

It is suggested that the nature of the amino acid relationship to codons (i.e. the genetic code) is such that ligand–ligand interaction is accommodated. This suggests that the universe of peptide ligand interaction is limited and that the codings for such interactions will elucidate the underlying constraint of why a particular amino acid is coded by a particular codon

(Crick *et al.*, 1976; Pieczenik, 1977, 1980a).

Dramatic and successful reproductive cytoplasmic replacement studies have shown that mitochondrial tRNA can properly code in alternate environments (Kenyon and Moraes, 1997; Pieczenik, 1980b). However, mitochondrial tRNA segments are only 51 nucleotides long, while cytoplasmic tRNA segments are over 72 nucleotides long. How could one have evolved from another? How can even one align these tRNA segments to make proper comparisons of descent? Does one add nucleotides to the mitochondria or delete them from the cytoplasmic tRNA? Clearly, there is no simple descent from a common ancestor model for these tRNA as suggested by Darwin. However, they all still code for the same codon–amino acid relationships as each other. Is this convergence, much like the shape of the lemur and human ear? That is, UUU is still Phe in both the mitochondrial and cytoplasmic codes. Yet, the physical manifestations of the genetic codes, the tRNA, in mitochondria and cytoplasm, are completely different.

Therefore, it is suggested that as a consequence of genotypic selection and ligand–ligand peptide constraints, the amino acid–codon relationship evolved to what is now known as the genetic code. The equivalence principle of the antibody–antigen universe is a consequence of this selection. This may suggest molecular points of interaction at the reproductive level, which may follow similar constraints, such as sperm ligand contact points on the oolemma. This will explain the high and accurate specificity that regulates intra-species reproduction. There may be manipulative pathways that avoid such constraints.

Figure 1 and **Figure 2** show peptides that are specific for sperm ligand contact points on the oolemma. These peptide ligands were designed without oolemma substrate but simply as a ‘complementary’ peptide to the sperm ligand. The ability to make such functional ‘complementary’ binding peptides suggest a limited universe of ligand interactions that may mimic the limited symmetrical universe of antibody–antigen interaction.

Acknowledgements

The author would like to thank Dr Jacques Cohen and Dr Dagan Wells for discussions and encouragement in presenting

these ideas. The author would like to thank Dr Henry Malter for taking the confocal picture of the oolemma peptide binder and for experiments which led to certain aspects of this work.

References

- Burnet FM 1957 A modification of Jerne's theory of antibody production using the concept of clonal selection. *Australian Journal of Science* **20**, 67–69.
- Crick FHC 1970 Central dogma of molecular biology. *Nature* **227**, 861–863.
- Crick FHC, Brenner S, Klug A, Pieczenik G 1976 A speculation on the origin of protein synthesis. *Origin of Life* **7**, 389–397.
- Feigenbaum MJ 1980 The metric universal properties of period doubling bifurcations and the spectrum for a route to turbulence. *Annals of the New York Academy of Sciences* **357**, 330–336.
- Kenyon L, Moraes C 1997 Expanding the functional human mitochondrial DNA database by the establishment of primate xenomitochondrial cybrids. *Proceedings of the National Academy of Sciences of the USA* **94**, 9131–9135.
- Lander ES and the International Human Genome Sequencing Consortium 2001 Initial sequencing and analysis of the human genome. *Nature* **409**, 860.
- Lehninger A 2000 *Principles of Biochemistry*. Worth Publishers, New York, USA; ISBN 1572599316.
- Messmer B 2000 *Toward Immunophenomics*. PhD Thesis, Rockefeller University, New York, NY, USA.
- Pauling L 1940 A theory of the structure and process of formation of antibodies. *Journal of the American Chemical Society* **62**, 2643.
- Pieczenik G 1977 *Theory of Genotypic Selection*. Congressional Record, Committee on Science and Technology, US House of Representatives, 95th Congress, No. 24, 323–340.
- Pieczenik G 1980a predicting coding function from nucleotide sequence, or survival of 'fitness' of tRNA. *Proceedings of the National Academy of Sciences of the USA* **77**, 3539–3543.
- Pieczenik G 1980b Multimers of a suppressor transfer RNA: supporting evidence for alternate conformations of the anticodon loop region. *Journal of Molecular Biology* **138**, 879–884.
- Pieczenik G 1998 *Werkwijze en middelen voor het sorteren en identificeren van biologische informatie*, April 22 1998. Netherlands Patent Office Number 0 241 487 (NL), Netherlands Industrial Property Office, NL-2280 HV Rijswijk, Netherlands.
- Pieczenik G 1999 *Method and Means for Sorting and Identifying Biological Information*. US Patent 5,866,363, February 2, 1999. General Information Services Division, United States Patent and Trademark Office, Washington, D.C. 20231.

Received 25 October 2002; refereed 21 November 2002; accepted 3 December 2002.