Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Virology

Amino acid interaction networks provide a new lens for therapeutic antibody discovery and anti-viral drug optimization

Karthik Viswanathan, Zachary Shriver and Gregory J Babcock

Identification of epitopes on viral proteins for the design/identification of broadly-neutralizing monoclonal antibodies (bnAbs) or specific immunogens for vaccine development is hampered by target amino acid diversity. Recently, bnAbs have been isolated for variable viruses by screening B cells from infected individuals for neutralization breadth. Epitope mapping and structural analysis of bnAbs revealed, while some of these bnAbs target glycan moieties, most target protein regions that are conserved in sequence and/or structure. However, almost universally viruses develop mutations that allow escape from neutralization suggesting protein function may not be dependent on the observed conservation. An alternative method for identification of conserved amino acid sequences utilizes an amino acid network-based approach. Calculation of a significant interaction network (SIN) score allows for selection of amino acids that are conserved and constrained within the protein system. Amino acids with high SIN scores are predicted to mutate at lower frequency due to the impact mutation has on the structure/function of a protein. By ascertaining regions of high SIN score, therapeutics can be appropriately designed to target these regions of low mutability. Further, the use of atomic interaction networks to examine protein structure and protein–protein interfaces can complement existing structure-based computational approaches for therapeutic engineering.

Address
Visterra, Inc., Research, One Kendall Square, Building 300, Cambridge, MA, United States

Corresponding author: Babcock, Gregory J. (gbabcock@visterrainc.com)

Current Opinion in Virology 2015, **11**:122–129

This review comes from a themed issue on **Preventive and therapeutic vaccines**

Edited by **Mansun Law**

http://dx.doi.org/10.1016/j.coviro.2015.03.019

Introduction

The evolving nature of zoonotic viruses presents a substantial challenge for the development of therapeutics for several reasons. First, viruses have evolved into multiple subtypes, serotypes and strains within a given species necessitating the development of a therapeutic or therapeutics that can target this intrinsic diversity. Second, viruses employ a number of strategies to prevent elimination from a host population including producing latent infection (HIV), being highly transmissible (smallpox), and/or replicating to high titers in an individual (Ebola). Finally, viral genomes are not static but continuously 'updated' in response to environmental cues. For example, prior to the 2009 H1N1 pandemic, circulating influenza strains required approximately three years to develop widespread resistance to oseltamivir, a small molecule, neuraminidase inhibitor [1]. This 'resistance challenge', which can be thought of as an arms race between host and pathogen, is especially true of RNA-based viruses, such as influenza, dengue, hepatitis C, and HIV, among others. In addition to resistance development to antivirals and vaccines, viral mutation (*antigenic drift*) and genetic reassortment (*antigenic shift*) can also give rise to viral strains to which the population has no antigenic memory [2,3]. For example, the episodic advent of pandemic strains of influenza virus are a result of antigenic shift; the 2009 pandemic influenza A virus strain possessed a specific combination of gene segments that had not been seen previously [4,5].

(Re-) Emergence of passive immunization

Prior to the advent of vaccines or antiviral drugs, convalescent serum was found to provide potent protection against infection or even to treat active infection [6]. Eventually, it was discovered that most of the protective properties within the serum was due to neutralizing antibodies. Passive immunization using antibodies has been suggested to offer several benefits in comparison to other antiviral treatment options. First, passive immunity provides the opportunity to protect at-risk individuals from infection. At-risk segments of the population include first responders, those who do not mount an immune response to vaccines, the immunocompromised, those in poor health, pregnant women, and those in critical care. The potential for long-lasting protection arising from a single injection of antibodies is appealing and has been pursued for a number of infectious agents. For example, in the case of hepatitis A, prophylactic administration of immunoglobulins can protect against viral exposure. Additionally, post-exposure prophylaxis with immunoglobulin is >85% effective in preventing hepatitis A if administered within 2 weeks after viral

exposure, and efficacy is even higher when administered early in the incubation period [7]. Indeed, this type of strategy forms the basis of the sole marketed antibody for infectious diseases where palivizumab is used to prevent respiratory syncytial virus infection in high risk infants.

With the success of polyclonal antibody (and subsequently palivizumab) in combatting infectious disease a concerted effort was undertaken to identify therapeutic monoclonal antibodies for various infectious agents. Early attempts to meet this goal failed due to the fact that most single antibodies within antiserum were found to be strain-specific, with the virus evolving resistance in many cases. This dampened the enthusiasm for developing monoclonal antibodies as therapeutics for highly variable viruses. The discovery of broadly neutralizing antibodies (bNAbs) against, for example, influenza virus, dengue virus, and HIV, which bind to and neutralize multiple strains across clades, subtypes, or serotypes has renewed interest in this area. Multiple agents have entered the clinic and raise the possibility that a single antibody or limited combination of antibodies can effectively neutralize a wide variety of strains. However, a challenge with targeting conserved epitopes is that while antibodies targeting these regions may be effective in the short term, many conserved sites are still susceptible to change under immune pressure and can lead to potential escape variants that are refractory to neutralization.

As an alternative, complementary approach, several structural (instead of sequence) based approaches have been developed to measure and quantify the effect of mutations on overall protein structure and hence function of a protein system. To this end, several groups have pointed to the fact that amino acid interaction networks can be used to examine protein structure and the importance of any given amino acid in supporting that structure (Table 1). Representing the protein structure as a network map with atoms as 'nodes' and interactions between atoms as 'edges' provides a graphical view of the amino acid interactions responsible for both secondary and tertiary structure.

Of note is that network theory has substantially developed over the last 50–60 years [8,9]. Many systems from a variety of disciplines, including the social sciences [10], biology [11], and computer science [12], can be represented by networks, that is, by nodes bound by edges. In each instance, graph theory is employed to describe these systems; in general, all are attempts to explain how elements within the system interact with one another within the network and to describe the general laws which govern the observed network properties. In the specific context of describing protein structure, protein–protein interface description, and effects of mutations on the protein network, there are generally four major parameter sets that must be addressed and optimized:

1. Identification of the set of atoms to include in the network analysis. For example, a typical approach is to consider only or primarily the α-carbon atom of each amino acid comprising the protein backbone.
2. Determination of the weight of each interaction (to assign edges). Are interactions weighed the same regardless of the 'source' of this interaction, that is, π stacking, electrostatic interactions, or is differential weighting applied?
3. Application of an appropriate distance cutoff for assigning an edge. Related to weighting, at what distance are two atoms considered to be interacting and assigned an edge?
4. Identification of whether all amino acids within a protein are considered in the network or only a subset of them such as buried or solvent accessible amino acids.

In each case, a network score is determined for selected atoms within the chosen amino acid set. Using such a

Table 1

Studies describing amino acid interaction networks for examination of protein structure

Study	Description
Brinda *et al*. [36]	Network representation of protein structure as protein structure graphs (PSG) and its implication on stability
Lee *et al*. [37]	Identification of functionally important residues by residue–residue coevolution network (RRCN)
Doncheva *et al*. [38]	RINerator and RINalyzer as tools for network representation of proteins
del Sol *et al*. [39]	Representation or proteins as networks to identify residues for maintaining short paths in network communication (centrally conserved residues), whose removal increases system fragility
Cusack *et al*. [40]	Identifies functionally critical residues based on network representation of protein structure
Süel *et al*. [41]	Studies role of amino acid networks in mediating allosteric communication in proteins
Thibert *et al*. [42]	Applies network analysis to identify amino acids critical for protein function
Amitai *et al*. [43]	Describes representation of protein structures as residue interaction graphs (RIGs) for identification of functional residues
Martin *et al*. [44]	Describes RING, a web-based server, to construct physico–chemically valid residue interaction networks (RINs) from PDB files
Greene *et al*. [45]	Applies network principles to native protein structures for understanding of the underlying determinants of protein folds