Available online at www.sciencedirect.com
ScienceDirect
Molecular Immunology 44 (2007) 1342–1351
Molecular Immunology
www.elsevier.com/locate/molimm

A study of the structural correlates of affinity maturation: Antibody affinity as a function of chemical interactions, structural plasticity and stability

Maria Pamela C. David *, Jonathan James T. Asprer, Joyce Sarah A. Ibana, Gisela P. Concepcion, Eduardo A. Padlan

Virtual Laboratory of Biomolecular Structures, Marine Science Institute, College of Science, University of the Philippines Diliman, Quezon City 1101, Philippines

Received 21 April 2006; received in revised form 15 May 2006; accepted 17 May 2006
Available online 18 July 2006

Abstract

Mutations introduced in an antibody germline sequence as a result of somatic hypermutation could cause its derivatives to have an altered affinity for its target. Affinity maturation favors the selection of the antibodies which exhibit increased affinity. The mutations in 80 high affinity anti-thyroid peroxidase sequences derived from six germlines were analysed in terms of the physicochemical properties of the replacement residues, namely hydrophilicity, size and polarizability, and charge and polarity, in the context of its position and probable solvent accessibility. The effects of these substitutions were evaluated in terms of the resultant increased chemical interactivity potential of the affinity-matured antibodies relative to the germline. The results of the analysis would be useful in the rational design of antibodies and of other proteins for improved binding properties.

Keywords: Affinity maturation; Mutational analysis; Substitution matrices; Antibody engineering

1. Introduction

The diversity of the antibody repertoire arises from the combinatorial and imprecise assembly of V, D and J gene segments, somatic hypermutation, and the selection of high affinity clones. Somatic mutations introduced in germline sequences increase affinity by increasing noncovalent interactions at the binding site, optimizing binding site conformation for the target, introducing interactions that may interfere with nonspecific binding, or a combination of these (Yin et al., 2003). These mutations are not randomly distributed on account of selection pressures, which tend to favor alterations in complementarity-determining regions (CDRs) than in framework regions (FRs).

The persistence of certain mutation patterns could be partly explained by the inherent mutability of certain codons observed even in sequences not influenced by selection (Kepler, 1997; Shapiro et al., 1999). So-called mutation 'hot spots', like the RGYW and TAA nucleotide sequence motifs, occur in antibody sequences, where these are preferentially mutated (Kepler, 1997; Shapiro et al., 1999). Other di- and tri-nucleotide motifs follow a hierarchy of mutability, which are usually reflected in most sequences (Shapiro et al., 1999). Differential codon usage is usually observed in CDRs and FRs, where highly mutable codons are preferentially used in the CDRs (Shapiro et al., 1999). However, this does not sufficiently explain either the mutations that involve minimally mutable codons, or the apparent preferences in terms of amino acid replacements after selection.

Selection, which is the key to the survival of a clone, is marked by competition. B cells that exhibit the highest affinities for the antigen would out-survive those which exhibit low or even moderate affinity. Affinity is largely determined by the character of mutations introduced in a germline. Naturally, a more chemically interactive clone would exhibit higher affinity for the antigen. However, mutations that increase chemical interactions in CDRs are not the only ones that have significant consequences on affinity. Mutations in FRs are also known to influence binding. As yet, it is difficult to pinpoint which mutations within the variable regions would contribute to affinity increases.

* Corresponding author. Tel.: +63 2 4332990; fax: +63 2 9247678.
E-mail address: pamdavid@msi.upd.edu.ph (M.P.C. David).

0161-5890/$ – see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.molimm.2006.05.006

A number of three-dimensional structural studies have been performed to evaluate the effects of mutations on antibody–ligand interactions (Cauerhff et al., 2004; Li et al., 2003; Wedemeyer et al., 1997). In one study (Wedemeyer et al., 1997), an artificial construct simulating the germline form of the antibody was used to evaluate individual contributions of mutations to structure and hapten binding. In two independent studies of anti-lysozyme antibodies (Cauerhff et al., 2004; Li et al., 2003), differences in crystal structure of lower- and high-affinity antibodies were used to demonstrate that very small structural disparities in combining site structure cause the change in affinity. The results of these structural studies apply only to the particular antibody–ligand pair studied.

It is not possible, by three-dimensional structure determination or by sequence analysis, to know exactly what changes occur during affinity maturation. This is because the sequence of the unmutated germline form of the antibody can only be guessed. Diversity at the V(D)J junctions, nucleotide additions and deletions during rearrangement, and, in the case of heavy chains, the use of more than one D gene segment and the utilization of different reading frames result in amino acids in the third CDR that cannot be distinguished from those which result from somatic hypermutation. Confining the analysis to the first and second CDRs, and to the N-terminal segment of the third CDR of the light chain, obviates this difficulty. The CDR3 of both light and heavy chains figure prominently in the binding to antigen and the inability to assess the effect of mutations in these regions of the binding site is unfortunate. Their inclusion in the analysis, however, can only result in unnecessary speculation.

Here, we report a sequence analysis of a set of mature, high-affinity immunoglobulins (Ig) in relation to germline sequences. We have chosen to study the sequences of human autoantibodies to thyroid peroxidase (TPO). Sequences for hundreds of these antibodies are available and affinity data for many of them have been determined (for example, see Chapal et al., 2001; Chazenbalk et al., 1993). Autoantibodies to TPO are known to possess high affinities, most of them having dissociation constants (K_D) in the order of 10^{-9} M or better (Chapal et al., 2001; Chazenbalk et al., 1993). The data on the anti-TPOs thus present a unique opportunity to study, at the chemical interactivity level, information on which mutations at specific positions are maintained or even preferred. Such analysis may also provide possible explanations for selection.

In most proteins, mutation preferences could be correlated with the physicochemical similarities between the mutated and replacement residues, where conservative replacements are less likely to disrupt the functional conformation of a protein (Wei et al., 1997). This, however, is not always true for antibody sequences, where a number of replacements to highly dissimilar residues contribute to functionality. This is particularly true for CDR mutations, and would sometimes hold for FR mutations. FR mutations are also of interest since FR residues play important roles in structure preservation, folding yield and overall Ig stability, which collectively affect affinity (Jung et al., 2001).

In this study, we analyzed mutations in a total of 80 anti-TPO sequences that we had assigned unequivocably to six different germlines. Of these, 47 are variable light (V_L) chain sequences and 33 are variable heavy (V_H) chain sequences. These sequences were assigned to three V_L germlines and three V_H germlines, respectively. Mutations in these anti-TPO sequences are probably representative of the point mutations retained after selection. A small percentage of the mutations, however, may have arisen from the the use of Taq polymerase, which has an average error rate of 8.0×10^{-6} mutations/base pair/duplication, in the PCR-based sequencing protocols used in the experiments (Cline et al., 1996). It is important to remark at this point that knowledge of the epitope is not required; it is sufficient to assume that all observable mutations from the germline in an affinity-matured antibody are non-destabilizing, and are also likely contributors to binding potential increases, antibody stability, or both. In support of this, the classic paper of Malby et al. has shown that the amino acid differences between cross-reactive antibodies lie outside the binding site. They also demonstrated that the principal difference between cross-reacting antibodies lie in the interactions that these form with the antigen. This implies that it is highly unlikely for antibodies recognizing a single epitope to bind to it using a limited set of interactions. It is possible that the diversity of these interactions may equal the diversity of interactions formed by antibodies binding diverse epitopes (Malby et al., 1994).

After establishing mutation frequencies, we considered corresponding changes in physicochemical factors that influence antibody binding. These include hydrophilicity, size and polarizability, and charge and polarity (De Genst et al., 2002; Quillin et al., 2000; Grantham, 1974). These directly influence antibody–antigen interface interactions, which could be broadly grouped into polar–polar, polar–hydrophobic, hydrophobic–hydrophobic, and electrostatic (Berchanski et al., 2004). Hydrophilicity influences the potential of a residue for exposure to solvent, i.e., its availability for contact (Stevens and Arkin, 1999). The hydrophilicity of CDR residues would also influence the potential interactions that may be formed with the antigen (Berchanski et al., 2004). Residue size and shape largely determine CDR topography and, consequently, complementarity to the epitope (Kumagai et al., 2003; Quillin et al., 2000; Padlan, 1994). Residue size is also a determinant of plasticity. Polarizability influences the ability to form van der Waals' interactions, while charge and polarity determine the potential for charge–charge interactions and H-bonding. These factors collectively determine binding specificity and strength.

Defining mutational trends in terms of physicochemical property changes for Ig sequences will add to our understanding of affinity maturation and provide useful information for the rational design of binding sites with improved binding properties.

2. Materials and methods

2.1. Sequence sources and notation

Anti-TPO sequences were obtained from the NCBI (http://www.ncbi.nlm.nih.gov/) and cross-checked against original papers by Chazenbalk et al. (1993) and Chapal et al. (2001). All these sequences were from antibodies with high affinity for the antigen ($K_D \leq 10^{-9}$ M). Germline sequences