Published Recommendation

Improving B-cell epitope prediction and its application to global antibody-antigen docking.

Krawczyk K, Liu X, Baker T, Shi J, Deane CM.

Andrew Martin
F1000Prime Faculty Member
F1000 Bioinformatics, Biomedical Informatics & Computational Biology
University College London, London, UK.
31 Jan 2017 | Interesting Hypothesis, New Finding
Recommended

It is estimated that at least half of all drugs in development are now biopharmaceuticals, with half of those (between a quarter and a third of all drugs in development) being antibodies. In developing an antibody as a drug, it is necessary for the antibody to bind to specific regions of the cognate antigen - for example, to a binding site - to have the desired effect on the antigen. However, a problem with all biopharmaceuticals, especially those that are not native human proteins, is that they can elicit an immune response when injected into a human host. In the best case, this leads to reduced efficacy while, in the worst case, it can lead to anaphylactic shock and death. Thus, the prediction of B-cell epitopes (BCEs) - sites where an antigen will be bound by an antibody - is important in two ways. First, if sites on a biopharmaceutical that will be recognized by host antibodies can be identified, it may be possible to modify these sites to make them less immunogenic. Second, and the subject of this paper, it may be possible to find the most likely BCE for a particular antibody given only the structure of the antigen and the sequence of the antibody. This could be useful as part of a virtual screening strategy for antibodies raised against an antigen of interest.

However, BCE prediction is a very difficult problem. Reported performance is often over-estimated, either because results are only provided from cross-validation rather than using an independent test set, or because evaluation is performed across a whole sequence (including buried residues that are not accessible as BCEs), meaning that much of the signal comes merely from the ability to predict which residues are buried. In our hands, even the very best predictors achieve a Matthews Correlation Coefficient (MCC) of less than 0.1 (where 0.0 is random and 1.0 is a perfect prediction), rising to 0.2 on certain test sets.

While some BCE predictors use only sequence data, others also employ structural information. Krawczyk et al. introduce a novel twist to the problem of BCE prediction by using the structure of the antigen, but also include information about the sequence of an antibody known to bind to the antigen in order to try to predict the precise epitope. Clearly, this cannot be used as a general strategy for identifying BCEs since the sequence of the antibody or antibodies is unknown. However, the strategy here is applied to finding the precise epitope for a given antibody and has potential to be used for in silico screening of antibody libraries. Table 1 in their paper reports the performance of the method across 30 known antibody-antigen crystal structures. Averaging the reported MCC values in the table gives a mean figure of 0.156 compared with a mean for Discotope (one of the better-performing BCE predictors, but which does not use antibody information) of 0.096. Where the structure of the antibody is also known, they have used the method to guide global antibody-antigen docking, producing a significant increase in performance of general docking methods.

In summary, the results are interesting and do suggest that antibody sequence information can improve the identification of a relevant epitope. However, while the improvements to global docking are significant, the performance for identifying the cognate epitopes is not yet good enough to address their original goal of providing a method for in silico screening of potential antibody drugs.

Disclosures